Math is Beautiful and Fun!

Sunghee Yun

July 30, 2024

Contents

List of Figures

List of Axiom, Principle

List of Definition

List of Inequality

List of Theorem, Corollary, Lemma, Proposition

List of Algorithm

1 Preamble

1.1 Kinds of fun we can enjoy with math

1.2 Notations

- sets of numbers
  - $\mathbf{N}$: set of natural numbers, $\mathbf{Z}$: set of integers, $\mathbf{Q}$: set of rational numbers
  - $\mathbf{R}$: set of real numbers, $\mathbf{R}_+$: set of nonnegative real numbers, $\mathbf{R}_{++}$: set of positive real numbers
- sequences $\langle x_i \rangle$ and like
  - finite $\langle x_i \rangle_{i=1}^n$, infinite $\langle x_i \rangle_{i=1}^\infty$ - use $\langle x_i \rangle$ when unambiguously understood
  - similarly for other operations - $\sum x_i$, $\prod x_i$, $\cup A_i$, $\cap A_i$, $\bigtimes A_i$
  - similarly for integrals - $\int f$ for $\int_{-\infty}^{\infty} f$
- sets
  - $\tilde{A}$: complement of A, $A \sim B$: $A \cap \tilde{B}$, $A \Delta B$: $A \cap \tilde{B} \cup \tilde{A} \cap B$
  - $\mathcal{P}(A)$: set of all subsets of A
- sets in metric vector spaces
  - $\overline{A}$: closure of set A

  - A°: interior of set A
  - **relint**: relative interior of set A
  - $\mathbf{bd}\, A$: boundary of set A

- set algebra
  - $\sigma(\mathcal{A})$: σ-algebra generated by $\mathcal{A}$, *i.e.*, smallest σ-algebra containing $\mathcal{A}$

- norms in $\mathbf{R}^n$
  - $\|x\|_p$ $(p \geq 1)$: p-norm of $x \in \mathbf{R}^n$, *i.e.*, $(|x_1|^p + \cdots + |x_n|^p)^{1/p}$
  - $\|x\|_2$: Euclidean norm

- matrices and vectors
  - a_i: i-th entry of vector a
  - A_{ij}: entry of matrix A at position (i, j), *i.e.*, entry in i-th row and j-th column
  - $\mathbf{Tr}(A)$: trace of $A \in \mathbf{R}^{n\times n}$, *i.e.*, $A_{1,1} + \cdots + A_{n,n}$

- symmetric, positive definite, and positive semi-definite matrices
  - $\mathbf{S}^n \subset \mathbf{R}^{n\times n}$: set of symmetric matrices
  - $\mathbf{S}^n_+ \subset \mathbf{S}^n$: set of positive semi-definite matrices - $A \succeq 0 \Leftrightarrow A \in \mathbf{S}^n_+$
  - $\mathbf{S}^n_{++} \subset \mathbf{S}^n$: set of positive definite matrices - $A \succ 0 \Leftrightarrow A \in \mathbf{S}^n_{++}$

- Python script-like notations (with serious abuse of notations!)
  - use $f : \mathbf{R} \to \mathbf{R}$ as if it were $f : \mathbf{R}^n \to \mathbf{R}^n$, *e.g.*,

    $$\exp(x) = (\exp(x_1), \ldots, \exp(x_n)) \quad \text{for } x \in \mathbf{R}^n$$

    or

    $$\log(x) = (\log(x_1), \ldots, \log(x_n)) \quad \text{for } x \in \mathbf{R}^n_{++}$$

    corresponding to Python code - `numpy.exp(x)` or `numpy.log(x)` - where `x` is instance of `numpy.ndarray`, *i.e.*, `numpy` array
  - use $\sum x$ for $\mathbf{1}^T x$ for $x \in \mathbf{R}^n$, *i.e.*

    $$\sum x = x_1 + \cdots + x_n$$

    corresponding to Python code - `x.sum()` - where `x` is `numpy` array
  - use x/y for $x, y \in \mathbf{R}^n$ for

    $$\begin{bmatrix} x_1/y_1 & \cdots & x_n/y_n \end{bmatrix}^T$$

    corresponding to Python code - `x / y` - where `x` and `y` are 1-d `numpy` arrays
    - applies to any two matrices of same dimensions

1.3 Some definitions

Definition 1.1 (infinitely often - i.o.) *statement, P_n, said to happen* infinitely often *or* i.o. *if*

$$(\forall N \in \mathbf{N})\,(\exists n > N)\,(P_n)$$

Definition 1.2 (almost everywhere - a.e.) *statement, $P(x)$, said to happen* almost everywhere *or* a.e. *or* almost surely *or* a.s. *(depending on context) associated with measure space, $(X, \mathscr{B}, \mu)$ if*

$$\mu\{x | P(x)\} = 1$$

or equivalently

$$\mu\{x | \sim P(x)\} = 0$$

1.4 Some conventions

- for some subjects, use following conventions
  - $0 \cdot \infty = \infty \cdot 0 = 0$
  - $(\forall x \in \mathbf{R}_{++})(x \cdot \infty = \infty \cdot x = \infty)$
  - $\infty \cdot \infty = \infty$

2 Algebra

2.1 Inequalities

Jensen's inequality

- strictly convex function: for any $x \neq y$ and $0 < \alpha < 1$

$$\alpha f(x) + (1-\alpha) f(y) > f(\alpha x + (1-\alpha) y)$$

- convex function: for any x, y and $0 < \alpha < 1$

$$\alpha f(x) + (1-\alpha) f(y) \geq f(\alpha x + (1-\alpha) y)$$

- *Jensen's inequality* - for strictly convex function f and *distinct* x_i and $0 < \alpha_i < 1$ with $\alpha_1 + \cdots = \alpha_n = 1$

$$\alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \geq f(\alpha_1 x_1 + \cdots + \alpha_n x_n)$$

  – equality holds *if and only if* $x_1 = \cdots = x_n$

Jensen's inequality - using probability distribution

- strictly convex function, f, and random variable, X
- discrete random variable interpretation - assume $\mathbf{Prob}(X = x_i) = \alpha_i$, then

$$\mathbf{E}\, f(X) = \alpha_1 f(x_1) + \cdots + \alpha_n f(x_n) \geq f(\alpha_1 x_1 + \cdots + \alpha_n x_n) = f\left(\mathbf{E}\, X\right)$$

- true for any random variables, X, with (general) function g

$$\mathbf{E}\, f(g(X)) \geq f(\mathbf{E}\, g(X))$$

- if probability density function (PDF), p, given

$$\int f(g(x)) p(x) dx \geq f\left(\int g(x) p(x) dx\right)$$

Proof for $n = 3$

- for any distinct x, y, z and $\alpha, \beta, \gamma > 0$ with $\alpha + \beta + \gamma = 1$

$$\begin{aligned}
\alpha f(x) + \beta f(y) + \gamma f(z) &= (\alpha+\beta)\left(\frac{\alpha}{\alpha+\beta} f(x) + \frac{\beta}{\alpha+\beta} f(y)\right) + \gamma f(z) \\
&> (\alpha+\beta) f\left(\frac{\alpha}{\alpha+\beta} x + \frac{\beta}{\alpha+\beta} y\right) + \gamma f(z) \\
&\geq f\left((\alpha+\beta)\left(\frac{\alpha}{\alpha+\beta} x + \frac{\beta}{\alpha+\beta} y\right) + \gamma z\right) \\
&= f(\alpha x + \beta y + \gamma z)
\end{aligned}$$

Proof for all integers

- use mathematical induction
- assume that Jensen's inequality holds for $1 \leq n \leq m$
- for any distinct x_i and α_i $(1 \leq i \leq m+1)$ with $\alpha_1 + \cdots + \alpha_{m+1} = 1$

$$\begin{aligned}
\sum^{m+1} \alpha_i f(x_i) &= \left(\sum^m \alpha_j\right) \sum^m \left(\frac{\alpha_i}{\sum^m \alpha_j} f(x_i)\right) + \alpha_{m+1} f(x_{m+1}) \\
&> \left(\sum^m \alpha_j\right) f\left(\sum^m \left(\frac{\alpha_i}{\sum^m \alpha_j} x_i\right)\right) + \alpha_{m+1} f(x_{m+1}) \\
&= \left(\sum^m \alpha_j\right) f\left(\frac{1}{\sum^m \alpha_j} \sum^m \alpha_i x_i\right) + \alpha_{m+1} f(x_{m+1}) \\
&\geq f\left(\sum^m \alpha_i x_i + \alpha_{m+1} x_{m+1}\right) = f\left(\sum^{m+1} \alpha_i x_i\right)
\end{aligned}$$

1st and 2nd order conditions for convexity

- 1st order condition (assuming differentiable $f : \mathbf{R} \to \mathbf{R}$) - f is strictly convex *if and only if* for any $x \neq y$
$$f(y) > f(x) + f'(x)(y - x)$$
- 2nd order condition (assuming twice-differentiable $f : \mathbf{R} \to \mathbf{R}$)
  - if $f''(x) > 0$, f is strictly convex
  - f is convex *if and only if* for any x
$$f''(x) \geq 0$$

Jensen's inequality examples

- $f(x) = x^2$ is strictly convex
$$\frac{a^2 + b^2}{2} \geq \left(\frac{a+b}{2}\right)^2$$
- $f(x) = x^4$ is strictly convex
$$\frac{a^4 + b^4}{2} \geq \left(\frac{a+b}{2}\right)^4$$
- $f(x) = \exp(x)$ is strictly convex
$$\frac{\exp(a) + \exp(b)}{2} \geq \exp\left(\frac{a+b}{2}\right)$$
- equality holds *if and only if* $a = b$ for all inequalities

1st and 2nd order conditions for convexity - vector version

- 1st order condition (assuming differentiable $f : \mathbf{R}^n \to \mathbf{R}$) - f is strict convex *if and only if* for any x, y

$$f(y) > f(x) + \nabla f(x)^T (y - x)$$

  where $\nabla f(x) \in \mathbf{R}^n$ with $\nabla f(x)_i = \partial f(x)/\partial x_i$

- 2nd order condition (assuming twice-differentiable $f : \mathbf{R}^n \to \mathbf{R}$)
  - if $\nabla^2 f(x) \succ 0$, f is strictly convex
  - f is convex *if and only if* for any x

$$\nabla^2 f(x) \succeq 0$$

  where $\nabla^2 f(x) \in \mathbf{S}^n_{++}$ with $\nabla^2 f(x)_{i,j} = \partial^2 f(x)/\partial x_i \partial x_j$

Jensen's inequality examples - vector version

- assume $f : \mathbf{R}^n \to \mathbf{R}$
- $f(x) = \|x\|_2 = \sqrt{\sum x_i^2}$ is strictly convex

$$(\|a\|_2 + 2\|b\|_2)/3 \geq \|(a + 2b)/3\|_2$$

  - equality holds *if and only if* $a = b \in \mathbf{R}^n$
- $f(x) = \|x\|_p = (\sum |x_i|^p)^{1/p}$ $(p > 1)$ is strictly convex

$$\frac{1}{k}\left(\sum_{i=1}^{k} \|x^{(i)}\|_p\right) \geq \left\|\frac{1}{k}\sum_{i=1}^{k} x^{(i)}\right\|_p$$

  - equality holds *if and only if* $x^{(1)} = \cdots = x^{(k)} \in \mathbf{R}^n$

AM $\geq$ GM

- for all $a, b > 0$

$$\frac{a+b}{2} \geq \sqrt{ab}$$

  - equality holds if and only if $a = b$
- (general form) for all $n \geq 1$, $a_i > 0$, $p_i > 0$ with $p_1 + \cdots + p_n = 1$

$$\alpha_1 a_1 + \cdots + \alpha_n a_n \geq a_1^{\alpha_1} \cdots a_n^{\alpha_n}$$

  - equality holds if and only if $a_1 = \cdots = a_n$
- let's prove these incrementally

AM $\geq$ GM - simplest case

- use fact that $x^2 \geq 0$ for any $x \in \mathbf{R}$
- for any $a, b > 0$

$$\begin{aligned} & (\sqrt{a}-\sqrt{b})^2 \geq 0 \\ \Leftrightarrow \quad & a^2 - 2\sqrt{ab} + b^2 \geq 0 \\ \Leftrightarrow \quad & a + b \geq 2\sqrt{ab} \\ \Leftrightarrow \quad & \frac{a+b}{2} \geq \sqrt{ab} \end{aligned}$$

  - equality holds if and only if $a = b$

AM $\geq$ GM - when $n = 4$ and $n = 8$

- for any $a, b, c, d > 0$

$$\frac{a+b+c+d}{4} \geq \frac{2\sqrt{ab}+2\sqrt{cd}}{4} = \frac{\sqrt{ab}+\sqrt{cd}}{2} \geq \sqrt{\sqrt{ab}\sqrt{cd}} = \sqrt[4]{abcd}$$

  - equality holds if and only if $a = b$ and $c = d$ and $ab = cd$ if and only if $a = b = c = d$

- likewise, for $a_1, \ldots, a_8 > 0$

$$\begin{aligned} \frac{a_1 + \cdots + a_8}{8} \quad & \geq \quad \frac{\sqrt{a_1a_2}+\sqrt{a_3a_4}+\sqrt{a_5a_6}+\sqrt{a_7a_8}}{4} \\ & \geq \quad \sqrt[4]{\sqrt{a_1a_2}\sqrt{a_3a_4}\sqrt{a_5a_6}\sqrt{a_7a_8}} \\ & = \quad \sqrt[8]{a_1 \cdots a_8} \end{aligned}$$

  - equality holds if and only if $a_1 = \cdots = a_8$

AM $\geq$ GM - when $n = 2^m$

- generalized to cases $n = 2^m$

$$\left(\sum_{a=1}^{2^m} a_i\right) / 2^m \geq \left(\prod_{a=1}^{2^m} a_i\right)^{1/2^m}$$

  - equality holds if and only if $a_1 = \cdots = a_{2^m}$

- can be proved by *mathematical induction*

AM $\geq$ GM - when $n = 3$

- proof for $n = 3$

$$\begin{aligned} & \frac{a+b+c}{3} = \frac{a+b+c+(a+b+c)/3}{4} \geq \sqrt[4]{abc(a+b+c)/3} \\ \Rightarrow \quad & \left(\frac{a+b+c}{3}\right)^4 \geq abc(a+b+c)/3 \\ \Leftrightarrow \quad & \left(\frac{a+b+c}{3}\right)^3 \geq abc \\ \Leftrightarrow \quad & \frac{a+b+c}{3} \geq \sqrt[3]{abc} \end{aligned}$$

  - equality holds if and only if $a = b = c = (a+b+c)/3$ if and only if $a = b = c$

AM $\geq$ GM - for all integers

- for any integer $n \neq 2^m$
- for m such that $2^m > n$

$$\begin{aligned}
& \frac{a_1 + \cdots + a_n}{n} = \frac{a_1 + \cdots + a_n + (2^m - n)(a_1 + \cdots + a_n)/n}{2^m} \\
& \geq \sqrt[2^m]{a_1 \cdots a_n \cdot ((a_1 \cdots a_n)/n)^{2^m - n}} \\
\Leftrightarrow \quad & \left(\frac{a_1 + \cdots + a_n}{n}\right)^{2^m} \geq a_1 \cdots a_n \cdot \left(\frac{a_1 \cdots a_n}{n}\right)^{2^m - n} \\
\Leftrightarrow \quad & \left(\frac{a_1 + \cdots + a_n}{n}\right)^{n} \geq a_1 \cdots a_n \\
\Leftrightarrow \quad & \frac{a_1 + \cdots + a_n}{n} \geq \sqrt[n]{a_1 \cdots a_n}
\end{aligned}$$

  - equality holds if and only if $a_1 = \cdots = a_n$

AM $\geq$ GM - rational α_i

- let

$$\alpha_i = \frac{p_i}{N}$$

  where $p_1 + \cdots + p_n = N$

- for all $a_i > 0$ and $\alpha_i > 0$ with $\alpha_1 + \cdots + \alpha_n = 1$

$$\alpha_1 a_1 + \cdots + \alpha_n a_n = \frac{p_1 a_1 + \cdots + p_n a_n}{N} \geq \sqrt[N]{a_1^{p_1} \cdots a_n^{p_n}} = a_1^{\alpha_1} \cdots a_n^{\alpha_n}$$

  - equality holds if and only if $a_1 = \cdots = a_n$

AM $\geq$ GM - real α_i

- exist n rational sequences $\{\beta_{i,1}, \beta_{i,2}, \ldots\}$ $(1 \leq i \leq n)$ such that

$$\begin{aligned}
& \beta_{1,j} + \cdots + \beta_{n,j} = 1 \ \forall \ j \geq 1 \\
& \lim_{j \to \infty} \beta_{i,j} = \alpha_i \ \forall \ 1 \leq i \leq n
\end{aligned}$$

- for all j

$$\begin{aligned}
& \beta_{1,j} a_1 + \cdots + \beta_{n,j} a_n \geq a_1^{\beta_{1,j}} \cdots a_n^{\beta_{n,j}} \\
\Rightarrow \quad & \lim_{j \to \infty} (\beta_{1,j} a_1 + \cdots + \beta_{n,j} a_n) \geq \lim_{j \to \infty} a_1^{\beta_{1,j}} \cdots a_n^{\beta_{n,j}} \\
\Leftrightarrow \quad & \alpha_1 a_1 + \cdots + \alpha_n a_n \geq a_1^{\alpha_1} \cdots a_n^{\alpha_n}
\end{aligned}$$

  - equality holds if and only if $a_1 = \cdots = a_n$

- cannot prove equality condition from above proof method

AM $\geq$ GM - proof using Jensen's inequality

- $-\log$ is strictly convex function because
$$\frac{d^2}{dx^2}\left(-\log(x)\right) = \frac{d}{dx}\left(-\frac{1}{x}\right) = \frac{1}{x^2} > 0$$
- Jensen's inequality $\Rightarrow$ for any distinct $a_i > 0$, $p_i > 0$ with $\sum p_i = 1$
$$-\log\left(\prod a_i^{\alpha_i}\right) = -\sum \log\left(a_i^{\alpha_i}\right) = \sum \alpha_i(-\log(a_i)) \geq -\log\left(\sum \alpha_i a_i\right)$$
- $-\log$ strictly decreases, hence
$$\prod a_i^{\alpha_i} \leq \sum \alpha_i a_i$$
- just proves
$$\sum \alpha_i a_i \geq \prod a_i^{\alpha_i}$$
  - equality if and only if a_i are equal

Cauchy-Schwarz inequality

- *Cauchy-Schwarz inequality* - for $a_i \in \mathbf{R}$ and $b_i \in \mathbf{R}$
$$(a_1^2 + \cdots + a_n^2)(b_1^2 + \cdots + b_n^2) \geq (a_1 b_1 + \cdots + a_n b_n)^2$$
- middle school proof
$$\begin{aligned} & \sum (t a_i + b_i)^2 \geq 0 \ \forall\, t \in \mathbf{R} \\ \Leftrightarrow \quad & t^2 \sum a_i^2 + 2t \sum a_i b_i + \sum b_i^2 \geq 0 \ \forall\, t \in \mathbf{R} \\ \Leftrightarrow \quad & \Delta = \left(\sum a_i b_i\right)^2 - \sum a_i^2 \sum b_i^2 \leq 0 \end{aligned}$$
  - equality holds if and only if $\exists t \in \mathbf{R}$, $t a_i + b_i = 0$ for all $1 \leq i \leq n$

Cauchy-Schwarz inequality - another proof

- $x \geq 0$ for any $x \in \mathbf{R}$, hence
$$\begin{aligned} & \sum_i \sum_j (a_i b_j - a_j b_i)^2 \geq 0 \\ \Leftrightarrow \quad & \sum_i \sum_j (a_i^2 b_j^2 - 2 a_i a_j b_i b_j + a_j^2 b_i^2) \geq 0 \\ \Leftrightarrow \quad & \sum_i \sum_j a_i^2 b_j^2 + \sum_i \sum_j a_j^2 b_i^2 - 2 \sum_i \sum_j a_i a_j b_i b_j \geq 0 \\ \Leftrightarrow \quad & 2 \sum_i a_i^2 \sum_j b_j^2 - 2 \sum_i a_i b_i \sum_j a_j b_j \geq 0 \\ \Leftrightarrow \quad & \sum_i a_i^2 \sum_j b_j^2 - \left(\sum_i a_i b_i\right)^2 \geq 0 \end{aligned}$$
  - equality holds if and only if $a_i b_j = a_j b_i$ for all $1 \leq i, j \leq n$

Cauchy-Schwarz inequality - still another proof

- for any $x, y \in \mathbf{R}$ and $\alpha, \beta > 0$ with $\alpha + \beta = 1$

$$
\begin{aligned}
& (\alpha x - \beta y)^2 = \alpha^2 x^2 + \beta^2 y^2 - 2\alpha\beta xy \\
& = \alpha(1-\beta)x^2 + (1-\alpha)\beta y^2 - 2\alpha\beta xy \geq 0 \\
\Leftrightarrow \quad & \alpha x^2 + \beta y^2 \geq \alpha\beta x^2 + \alpha\beta y^2 + 2\alpha\beta xy = \alpha\beta(x+y)^2 \\
\Leftrightarrow \quad & x^2/\alpha + y^2/\beta \geq (x+y)^2
\end{aligned}
$$

- plug in $x = a_i$, $y = b_i$, $\alpha = A/(A+B)$, $\beta = B/(A+B)$ where $A = \sqrt{\sum a_i^2}$, $B = \sqrt{\sum b_i^2}$

$$
\begin{aligned}
& \sum (a_i^2/\alpha + b_i^2/\beta) \geq \sum (a_i + b_i)^2 \Leftrightarrow (A+B)^2 \geq A^2 + B^2 + 2\sum a_i b_i \\
\Leftrightarrow \quad & AB \geq \sum a_i b_i \Leftrightarrow A^2 B^2 \geq \left(\sum a_i b_i\right)^2 \Leftrightarrow \sum a_i^2 \sum b_i^2 \geq \left(\sum a_i b_i\right)^2
\end{aligned}
$$

Cauchy-Schwarz inequality - proof using determinant

- almost the same proof as first one - but using 2-by-2 matrix determinant

$$
\begin{aligned}
& \sum (x a_i + y b_i)^2 \geq 0 \ \forall\, x, y \in \mathbf{R} \\
\Leftrightarrow \quad & x^2 \sum a_i^2 + 2xy \sum a_i b_i + y^2 \sum b_i^2 \geq 0 \ \forall\, x, y \in \mathbf{R} \\
\Leftrightarrow \quad & \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} \sum a_i^2 & \sum a_i b_i \\ \sum a_i b_i & \sum b_i^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \geq 0 \ \forall\, x, y \in \mathbf{R} \\
\Leftrightarrow \quad & \begin{vmatrix} \sum a_i^2 & \sum a_i b_i \\ \sum a_i b_i & \sum b_i^2 \end{vmatrix} \geq 0 \Leftrightarrow \sum a_i^2 \sum b_i^2 - \left(\sum a_i b_i\right)^2 \geq 0
\end{aligned}
$$

  - equality holds *if and only if*

$$
(\exists x, y \in \mathbf{R} \text{ with } xy \neq 0)\, (x a_i + y b_i = 0 \ \ \forall 1 \leq i \leq n)
$$

- allows *beautiful generalization* of Cauchy-Schwarz inequality

Cauchy-Schwarz inequality - generalization

- want to say something like $\sum_{i=1}^n (x a_i + y b_i + z c_i + w d_i + \cdots)^2$
- run out of alphabets ... - use double subscripts

$$
\begin{aligned}
& \sum_{i=1}^n (x_1 A_{1,i} + x_2 A_{2,i} + \cdots + x_m A_{m,i})^2 \geq 0 \ \forall\, x_i \in \mathbf{R} \\
\Leftrightarrow \quad & \sum_{i=1}^n (x^T a_i)^2 = \sum_{i=1}^n x^T a_i a_i^T x = x^T \left(\sum_{i=1}^n a_i a_i^T \right) x \geq 0 \ \forall\, x \in \mathbf{R}^m \\
\Leftrightarrow \quad & \begin{vmatrix} \sum_{i=1}^n A_{1,i}^2 & \sum_{i=1}^n A_{1,i} A_{2,i} & \cdots & \sum_{i=1}^n A_{1,i} A_{m,i} \\ \sum_{i=1}^n A_{1,i} A_{2,i} & \sum_{i=1}^n A_{2,i}^2 & \cdots & \sum_{i=1}^n A_{2,i} A_{m,i} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n A_{1,i} A_{m,i} & \sum_{i=1}^n A_{2,i} A_{m,i} & \cdots & \sum_{i=1}^n A_{m,i}^2 \end{vmatrix} \geq 0
\end{aligned}
$$

  where $a_i = \begin{bmatrix} A_{1,i} & \cdots & A_{m,i} \end{bmatrix}^T \in \mathbf{R}^m$
  - equality holds *if and only if* $\exists x \neq 0 \in \mathbf{R}^m$, $x^T a_i = 0$ for all $1 \leq i \leq n$

Cauchy-Schwarz inequality - three series of variables

- let $m = 3$

$$\left[\begin{array}{ccc} \sum a_i^2 & \sum a_i b_i & \sum a_i c_i \\ \sum a_i b_i & \sum b_i^2 & \sum b_i c_i \\ \sum a_i c_i & \sum b_i c_i & \sum c_i^2 \end{array}\right] \succeq 0$$

$$\begin{aligned} \Rightarrow \quad & \sum a_i^2 \sum b_i^2 \sum c_i^2 + 2 \sum a_i b_i \sum b_i c_i \sum c_i a_i \\ & \geq \sum a_i^2 \left(\sum b_i c_i\right)^2 + \sum b_i^2 \left(\sum a_i c_i\right)^2 + \sum c_i^2 \left(\sum a_i b_i\right)^2 \end{aligned}$$

  - equality holds if and only if $\exists x, y, z \in \mathbf{R}$, $x a_i + y b_i + z c_i = 0$ for all $1 \leq i \leq n$

- Questions for you
  - what does this imply?
  - any real-world applications?

Cauchy-Schwarz inequality - extensions

- complex numbers - for $a_i \in \mathbf{C}$ and $b_i \in \mathbf{C}$

$$\sum |a_i|^2 \sum |b_i|^2 \geq \left|\sum a_i b_i\right|^2$$

- infinite sequences - for $a_1, a_2, \ldots \in \mathbf{C}$ and $b_1, b_2, \ldots \in \mathbf{C}$

$$\sum^{\infty} |a_i|^2 \sum^{\infty} |b_i|^2 \geq \left|\sum^{\infty} a_i b_i\right|^2$$

- Hilbert space - for $f, g : [0,1] \to \mathbf{C}$

$$\int |f|^2 \int |g|^2 \geq \left|\int f g\right|^2$$

  or

$$\|f\| \|g\| \geq \langle f, g \rangle$$

  (could be derived from definition of inner products only)

2.2 Number Theory - Queen of Mathematics

Integers

- integers ($\mathbf{Z}$)
  - $\ldots - 2, -1, 0, 1, 2, \ldots$
- first defined by Bertrand Russell
- algebraic structure: commutative ring
  - addition, multiplication (not division) defined
  - addition, multiplication are associative
  - multiplication distributive over addition
  - addition, multiplication are commutative
- natural numbers ($\mathbf{N}$)
  - $1, 2, \ldots$

Division and prime numbers

- divisors for $n \in \mathbf{N}$
$$\{d \in \mathbf{N} | d \text{ divides } n\}$$
- prime numbers
  - p is primes if 1 and p are only divisors

Fundamental theorem of arithmetic

Theorem 2.1 (fundamental theorem of arithmetic) *integer $n \geq 2$ can be factored uniquely into products of primes, i.e., exist distinct primes, $p_1, \ldots, p_k$, and $e_1, \ldots, e_k \in \mathbf{N}$ such that*

$$n = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$$

Elementary quantities

- greatest common divisor (gcd) (of a and b)
$$\mathrm{gcd}(a, b) = \max\{d | d \text{ divides both } a \text{ and } b\}$$
- least common multiple (lcm) (of a and b)
$$\mathrm{lcm}(a, b) = \min\{m | \text{both } a \text{ and } b \text{ divides } m\}$$
- a and b coprime, relatively prime, mutually prime $\Leftrightarrow \mathrm{gcd}(a, b) = 1$

Are there finite number of prime numbers?

- no!
- proof
  - assume there exist finite number of prime numbers, *e.g.*, $p_1 < p_2 < \cdots < p_n$
  - but $p_1 \cdot p_2 \cdots p_n + 1$ is prime, which is greater than p_n, hence contradiction

Integers modulo n

Definition 2.1 (modulo) a is said to be equivalent to b modulo n *if n divides $a - b$, denoted by*

$$a \equiv b \pmod{n}$$

read "a congruent to b mod n"

- $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$ imply
  - $a + c \equiv b + d \pmod{n}$
  - $ac \equiv bd \pmod{n}$

Definition 2.2 (congruence class) *classes determined by modulo relation, called* congruence or residue class under modulo

Definition 2.3 (integers modulo n) *set of equivalence classes under modulo, denoted by* $\mathbf{Z}/n\mathbf{Z}$, *called* integers modulo n *or* integers mod n

Euler's theorem

Definition 2.4 (Euler's totient function) *for* $n \in \mathbf{N}$,

$$\varphi(n) = (p_1 - 1)p_1^{e_1-1} \cdots (p_k - 1)p_k^{e_k-1} = n \prod_{\text{prime } p \text{ dividing } n} (1 - 1/p)$$

called Euler's totient function, *also called* Euler φ-function

- *e.g.*, $\varphi(12) = \varphi(2^2 \cdot 3^1) = 1 \cdot 2^1 \cdot 2 \cdot 3^0 = 4$, $\varphi(10) = \varphi(2^1 \cdot 5^1) = 1 \cdot 2^0 \cdot 4 \cdot 5^0 = 4$

Theorem 2.2 (Euler's theorem - number theory) *for coprime* n *and* a

$$a^{\varphi(n)} \equiv 1 \pmod{n}$$

- *e.g.*, $5^4 \equiv 1 \pmod{12}$ whereas $4^4 \equiv 4 \not\equiv 1 \pmod{12}$
- proof not (extremely) hard, but beyond scope of presentation
- *Euler's theorem underlies RSA cryptosystem - widely used in internet communication*

3 Abstract Algebra

3.1 Why Abstract Algebra? & Some Historical View

Why abstract algebra?

- first, it's fun!
- and, lets us understand instrict structures of algebraic objects
- and, doing it allows us to solve *extremely practical problems*; well, depends on your definition of practicality
  - *e.g.*, can prove why not exist general root formulas for polynomials of order greater than 4 - is it practical enough for you?
- also prepare us for pursuing further math topics such as differential geometry, algebraic geometry, analysis, representation theory, algebraic number theory, *etc.*
- and ... it's really fun; sheer examination and pondoring upon ideas make us happy and thrilled!

Some history

- by the way, historically, often the case that application of an idea presented before extracting and presenting the idea on its own right
- *e.g.*, Galois used "quotient group" only implicitly in his 1830's investigation, and it had to wait until 1889 to be explicitly presented as "abstract quotient group" by Hölder

3.2 Groups

Monoids

Definition 3.1 (law of composition) *mapping $S \times S \to S$ for set S, called* law of composition (of S to itself)

- *when $(\forall x, y, z \in S)((xy)z = x(yz))$, composition is said to be* associative
- *$e \in S$ such that $(\forall x \in S)(ex = xe = x)$, called* unit element *- always unique*

  Proof: for unit elements, e, f $e = ef = f$, hence, $e = f$

Definition 3.2 (monoids) *set, M, with composition which is associative and having unit element, called* monoid *(so in particular, M is not empty)*

- *monoid, M, with $(\forall x, y \in M)\,(xy = yx)$, called* commutative or abelian *monoid*
- *subset, $H \subset M$, which has element and is itself monoid, called* submonoid

Groups

Definition 3.3 (groups) *monoid, G, with*

$$(\forall x \in G)\,(\exists y \in G)\,(xy = yx = e)$$

called group

- *for $x \in G$, $y \in G$ with $xy = yx = e$, called* inverse of x

- *group derived from commutative monoid, called* abelian group *or* commutative group

- *group, G, with $|G| < \infty$, called* finite group

- *(similarly as submonoid) $H \subset G$ that has unit element and is itself group, called* subgroup

- *subgroup consisting only of unit element, called* trivial

Cyclic groups, generators, and direct products

Definition 3.4 (cyclic groups) *group, G, with*

$$(\exists a \in G) \, (\forall x \in G) \, (\exists n \in \mathbf{N}) \, (x = a^n)$$

called cyclic group, *such $a \in G$, called* cyclic generator

Definition 3.5 (generators) *for group, G, $S \subset G$ with*

$$(\forall x \in G) \, (x \textit{ is arbitrary product of elements or inverse elements of } S)$$

called set of generators for G, *said to* generate G, *denoted by $G = \langle S \rangle$*

Definition 3.6 (direct products) *for two groups, G_1 and G_2, group, $G_1 \times G_2$, with*

$$(\forall (x_1, x_2), (y_1, y_2) \in G_1 \times G_2) \, ((x_1, x_2)(y_1, y_2) = (x_1 y_1, x_2, y_2) \in G_1 \times G_2)$$

whose unit element defined by (e_1, e_2) where e_1 and e_2 are unit elements of G_1 and G_2 respectively, called direct product of G_1 and G_2

Homeomorphism and isomorphism

Definition 3.7 (homeomorphism) *for monoids, M and M', mapping $f : M \to M'$ with $f(e) = e'$*

$$(x, y \in M) \, (f(xy) = f(x) f(y))$$

where e and e' are unit elements of M and M' respectively, called monoid-homeomorphism *or simple* homeomorphism

- group homeomorphism, *$f : G \to G'$, is similarly monoid-homeomorphism*

- *homeomorphism, $f : G \to G'$ where exists $g : G \to G'$ such that $f \circ g : G' \to G'$ and $g \circ f : G \to G$ are identity mappings,called* isomorphism, *sometimes denoted by* $G \approx G'$

- *homeomorphism of G into itself, called* endomorphism

- *isomorphism of G onto itself, called* automorphism

  - set of all automorphisms of G is itself group, denoted by $\mathrm{Aut}(G)$

Kernel, image, and embedding of homeomorphism

Definition 3.8 (kernels of homeomorphism) *for group-homeomorphism, $f : G \to G'$, where e' is unit element of G', $f^{-1}(\{e'\})$, which is subgroup of G, called* kernel of f*, denoted by* $\operatorname{Ker} f$

Definition 3.9 (embedding of homeomorphism) *homeomorphism, $f : G \to G'$, establishing isomorphism between G and $f(G) \subset G'$, called* embedding

Proposition 3.1 (group homeomorphism and isomorphism)

- *for group-homeomorphism, $f : G \to G'$, $f(G) \subset G'$ is subgroup of G'*
- *homeomorphism whose kernel is trivial is injective, often denoted by special arrow*

$$f : G \hookrightarrow G'$$

- *surjective homeomorphism whose kernel is trivial is* isomorphism
- *for group, G, its generators, S, and another group, G', map, $f : S \to G'$ has at most one extension to homeomorphism of G into G'*

Orthogonal subgroups

Proposition 3.2 (orthogonal subgroups) *for group, G, and two subgroups, $H, K \subset G$, with $HK = G$, $H \cap K = \{e\}$, and $(x \in H, y \in K)\,(xy = yx)$,*

$$f : H \times K \to G$$

with $(x, y) \mapsto xy$ is isomorphism

can generalize to finite number of subgroups, H_1, ..., H_n such that

$$H_1 \cdots H_n = G$$

and

$$H_{k+1} \cap (H_1 \cdots H_k) = \{e\}$$

in which case, G is isomorphic to $H_1 \cdots H_n$

Cosets of groups

Definition 3.10 (cosets of groups) *for group, G, and subgroup, $H \subset G$, aH for some $a \in G$, called* left coset of H in G*, and element in aH, called* coset representation of aH *- can define* right cosets *similarly*

Proposition 3.3 (cosets of groups) *for group, G, and subgroup, $H \subset G$,*

- *for $a \in G$, $x \mapsto ax$ induces bijection of H onto aH, hence all left cosets have same cardinality*
- *$aH \cap bH \neq \emptyset$ for $a, b \in G$ implies $aH = bH$*
- *hence, G is disjoint union of left cosets of H*
- *same statements can be made for right cosets*

Definition 3.11 (index and order of groups) *number of left cosets of H in G, called* index of H in G*, denoted by $(G : H)$ - index of trivial subgroups, called* order of G*, denoted by $(G : 1)$*

Indices and orders of groups

Proposition 3.4 (indices and orders) *(proof can be found in Proof 1) for group, G, and two subgroups, $H, K \subset G$ with $K \subset H$,*

$$(G : H)(H : K) = (G : K)$$

when K is trivial, we have

$$(G : H)(H : 1) = (G : 1)$$

hence, if $(G : 1) < \infty$, both $(G : H)$ and $(H : 1)$ divide $(G : 1)$

Normal subgroup

Definition 3.12 (normal subgroups) *subgroup, $H \subset G$, of group, G, with*

$$(\forall x \in G)\,(xH = Hx) \Leftrightarrow (\forall x \in G)\left(xHx^{-1} = H\right)$$

called normal subgroup of G*, in which case*

- *set of cosets, $\{xH|x \in G\}$, is with law of composition defined by $(xH)(yH) = (xy)H$, forms group with unit element, H, denoted by* G/H*, called* factor group of G by H*, read* G modulo H *or* G mod H
- *$x \mapsto xH$ induces homeomorphism of X onto $\{xH|x \in G\}$, called* canonical map*, kernel of which is H*

Proposition 3.5 (normal subgroups and factor groups)

- *kernel of (every) homeomorphism of G is normal subgroups of G*
- *for family of normal subgroups of G, $\langle N_\lambda \rangle$, $\bigcap N_\lambda$ is also normal subgroup*
- *every subgroup of abelian group is normal*
- *factor group of abelian group is abelian*
- *factor group of cyclic group is cyclic*

Normalizers and centralizers

Definition 3.13 (normalizers and centralizers) *for subset, $S \subset G$, of group, G,*

$$\{x \in G|xSx^{-1} = S\}$$

is subgroup, called normalizer of S*, and also called* centralizer of a *when $S = \{a\}$ is singletone;*

$$\{x \in G|(\forall y \in S)(xyx^{-1} = y)\}$$

called centralizer of S*, and centralizer of G itself, called* center of G

- *e.g.*, $A \mapsto \det A$ of multiplicative group of square matrices in $\mathbf{R}^{n\times n}$ into $\mathbf{R} \sim \{0\}$ is homeomorphism, kernel of which called *special linear group,* and (of course) is normal

$$\begin{tikzcd}
0 \arrow[r] & G' \arrow[r, "f"] \arrow[d] & G \arrow[r, "g"] \arrow[d] & G'' \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & H \arrow[r] & G \arrow[r] & G/H \arrow[r] & 0
\end{tikzcd}$$

Figure 3.1: commutative diagram for canonical map

Normalizers and congruence

Proposition 3.6 (normalizers of groups) *subgroup, $H \subset G$, of group, G, is normal subgroup of its normalizer, N_H*

- *subgroup, $H \subset G$, of group, G, is normal subgroup of its normalizer, N_H*
- *subgroup, $K \subset G$ with $H \subset K$ where H is normal in K, is contained in N_H*
- *for subgroup, $K \subset N_H$, KH is group and H is normal in KH*
- *normalizer of H is largest subgroup of G in which H is normal*

Definition 3.14 (congruence with respect to normal subgroup) *for normal subgroup, $H \subset G$, of group, G, we write*

$$x \equiv y \pmod{H}$$

if $xH = yH$, read x and y are congruent modulo H *- this notation used mostly for additive groups*

Exact sequences of homeomorphisms

Definition 3.15 (exact sequences of homeomorphisms) *below sequence of homeomorphisms with* $\mathrm{Im}\, f = \mathrm{Ker}\, g$

$$G' \xrightarrow{f} G \xrightarrow{g} G''$$

said to be exact

below sequence of homeomorphisms with $\mathrm{Im}\, f_i = \mathrm{Ker}\, f_{i+1}$

$$G_1 \xrightarrow{f_1} G_2 \xrightarrow{f_2} G_3 \longrightarrow \cdots \xrightarrow{f_{n-1}} G_n$$

said to be exact

- for normal subgroup, $H \subset G$, of group G, sequence, $H \xrightarrow{j} G \xrightarrow{\varphi} G/H$, is exact where j is inclusion and φ
- $0 \to G' \xrightarrow{f} G \xrightarrow{g} G'' \to 0$ is exact *if and only if* f injective, g surjective, and $\mathrm{Im}\, f = \mathrm{Ker}\, g$
- if $H = \mathrm{Ker}\, g$ above, $0 \to H \to G \to G/H \to 0$
- more precisely, exists commutative diagram as in Figure 3.1, in which vertical mappings are isomorphisms and rows are *exact*

$$
\begin{tikzcd}
0 \arrow[r] & H \arrow[r] \arrow[d, "\mathrm{can}"] & G \arrow[r] \arrow[d, "\mathrm{can}"] & G/H \arrow[r] \arrow[d, "\mathrm{id}"] & 0 \\
0 \arrow[r] & H/K \arrow[r] & G/K \arrow[r] & G/H \arrow[r] & 0
\end{tikzcd}
$$

Figure 3.2: commutative diagram for canonical isomorphism

Canonical homeomorphism examples all homeomorphisms described below called *canonical*

- for two groups, G and G', and homeomorphism, $f : G \to G'$, whose kernel is H, exists unique homeomorphism $f_* : G/H \to G'$ with

  $$f = f_* \circ \varphi$$

  where $\varphi : G \to G/H$ is canonical map, and f_* is injective

  - f_* can be defined by $xH \mapsto f(x)$
  - say *f_* is induced by f*
  - f_* induces isomorphism, $\lambda : G/H \to \operatorname{Im} f$
  - below sequence summarizes above statements

    $$G \xrightarrow{\varphi} G/H \xrightarrow{\lambda} \operatorname{Im} f \xrightarrow{j} G$$

    where j is inclusion

- for group, G, subgroup, $H \subset G$, intersection of all normal subgroups containing H, N (*i.e.*, smallest normal subgroup containing H), homeomorphism, $f : G \to G'$ whose kernel contains H, $N \subset \operatorname{Ker} f$ and exists unique homeomorphism, $f_* : G/N \to G'$ such that

  $$f = f_* \circ \varphi$$

  where $\varphi : G \to G/H$ is canonical map

  - f_* can be defined by $xN \mapsto f(x)$
  - say *f_* is induced by f*

- for subgroups of G, H and K, with $K \subset H$, $xK \mapsto xH$ induces homeomorphism of G/K into G/H, whose kernel is $\{xK | x \in H\}$, thus *canonical isomorphism*

  $$(G/K)/(H/K) \approx (G/K)$$

  this can be shown in Figure 3.2 where rows are exact

- for subgroup, $H, K \subset G$, with H contained in normalizer of K, $H \cap K$ is normal subgroup of H, $HK = KH$ is subgroup of G, exists surjective homeomorphism

  $$H \to HK/K$$

  with $x \mapsto xK$, whose kernel is $H \cap K$, hence *canonical isomorphism*

  $$H/(H \cap K) \approx HK/K$$

$$\begin{tikzcd} G \arrow[r] & G' \\ f^{-1}(H') \arrow[u] \arrow[r] & H' \arrow[u] \end{tikzcd}$$

Figure 3.3: commutative diagram

$$\begin{tikzcd} 0 \arrow[r] & H \arrow[r] \arrow[d] & G \arrow[r] \arrow[d, "f"] & G/H \arrow[r] \arrow[d, "\bar{f}"] & 0 \\ 0 \arrow[r] & H' \arrow[r] & G' \arrow[r] & G'/H' \arrow[r] & 0 \end{tikzcd}$$

Figure 3.4: commutative diagram for canonical homeomorphism

- for group homeomorphism, $f : G \to G'$, normal subgroup of G', H',
$$H = f^{-1}(H') \subset G$$
as shown in Figure 3.3, H is normal in G and kernel of homeomorphism
$$G \xrightarrow{f} G' \xrightarrow{\varphi} G'/H'$$
is H where φ is canonical map, hence we have injective homeomorphism
$$\bar{f} : G/H \to G'/H'$$
again called *canonical homeomorphism*, giving commutative diagram in Figure 3.4; if f is surjective, $\bar{f}$ is isomorphism

Towers

Definition 3.16 (towers of groups) *for group, G, sequence of subgroups*
$$G = G_0 \supset G_1 \supset G_2 \supset \cdots \supset G_m$$
called tower of subgroups, called normal if every G_{i+1} is normal in G_i, called abelian if normal and every factor group G_i/G_{i+1} is abelian, called cyclic if normal and every factor group G_i/G_{i+1} is cyclic

Proposition 3.7 (towers inded by homeomorphism) *for group homeomorphism, $f : G \to G'$ and normal tower*
$$G' = G'_0 \supset G'_1 \supset G'_2 \supset \cdots \supset G'_m$$
tower
$$f^{-1}(G') = f^{-1}(G'_0) \supset f^{-1}(G'_1) \supset f^{-1}(G'_2) \supset \cdots \supset f^{-1}(G'_m)$$
is

- *normal if G'_i form normal tower*
- *abelian if G'_i form abelian tower*
- *cyclic if G'_i form cyclic tower*

because every homeomorphism
$$G_i/G_{i+1} \to G'_i/G'_{i+1}$$
is injective

Refinement of towers and solvability of groups

Definition 3.17 (refinement of towers) *for tower of subgroups, tower obtained by inserting finite number of subgroups, called* refinement of tower

Definition 3.18 (solvable groups) *group having an abelian tower whose last element is trivial subgroup, said to be* solvable

Proposition 3.8 (finite solvable groups)

- *abelian tower of finite group admits cyclic refinement*
- *finite solvable group admits cyclic tower, whose last element is trivial subgroup*

Theorem 3.1 (Feit-Thompson theorem) *group whose order is prime power is solvable*

Theorem 3.2 (solvability condition in terms of normal subgroups) *for group, G, and its normal subgroup, H, G is solvable* if and only if *both H and G/H are solvable*

Commutators and commutator subgroups

Definition 3.19 (commutator) *for group, G, $xyx^{-1}y^{-1}$ for $x, y \in G$, called* commutator

Definition 3.20 (commutator subgroups) *subgroup generated by commutators of group, G, called* commutator subgroup*, denoted by G^C, i.e.*

$$G^C = \langle \{xyx^{-1}y^{-1} | x, y \in G\} \rangle$$

- G^C is normal in G
- G/G^C is commutative
- G^C is contained in kernel of every homeomorphism of G into commutative group
  - (proof can be found in Proof 2) for proof of above statements
- *commutator group is at the heart of solvability and non-solvability problems!*

Simple groups

Definition 3.21 (simple groups) *non-trivial group having no normal subgroup other than itself and trivial subgroup, said to be* simple

Proposition 3.9 (simple groups) *abelian group is simple* if and only if *cycle of prime order*

Butterfly lemma

Lemma 3.1 (butterfly lemma - Zassenhaus) *for subgroups, U and V, of group, and normal subgroups, u and v, of U and V respectively,*

$$u(U \cap v) \text{ is normal in } u(U \cap V)$$

$$(u \cap V)v \text{ is normal in } (U \cap V)v$$

and factor groups are isomorphic, i.e.

$$u(U \cap V)/u(U \cap v) \approx (U \cap V)v/(u \cap V)v$$

these shown in Figure 3.5

- indeed,
$$(U \cap V)/((u \cap V)(U \cap v)) \approx u(U \cap V)/u(U \cap v) \approx (U \cap V)v/(u \cap V)v$$

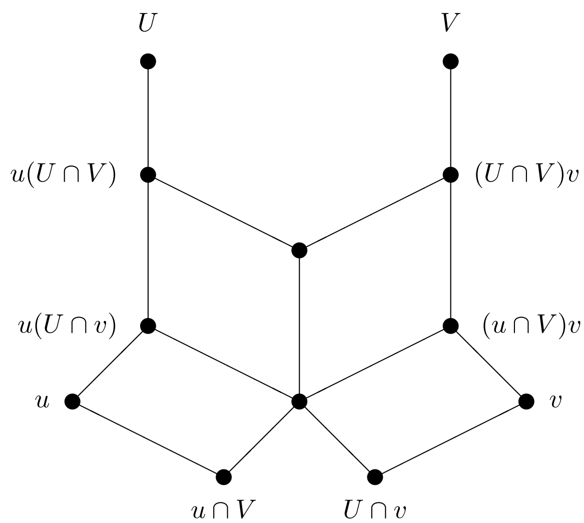

Figure 3.5: butterfly lemma

Equivalent towers

Definition 3.22 (equivalent towers) *for two normal towers of same height starting from same group ending with trivial subgroup*

$$G = G_1 \supset G_2 \supset G_3 \supset \cdots \supset G_{n+1} = \{e\}$$

$$G = H_1 \supset H_2 \supset H_3 \supset \cdots \supset H_{n+1} = \{e\}$$

with

$$G_i/G_{i+1} \approx H_{\pi(i)+1}/H_{\pi(i)}$$

for some permutation $\pi \in \mathrm{Perm}(\{1, \ldots, n\})$*, i.e., sequences of factor groups are same up to isomorphisms and permutation of indices, said to be* equivalent

Schreier and Jordan-Hölder theorems

Theorem 3.3 (Schreier) *two normal towers starting from same group and ending with trivial subgroup have equivalent refinement*

Theorem 3.4 (Jordan-Hölder theorem) *all normal towers starting from same group and ending with trivial subgroup where each factor group is non-trivial and simple are equivalent*

Cyclic groups

Definition 3.23 (exponent of groups and group elements) *for group,* G*,* $n \in \mathbf{N}$ *with* $a^n = e$ *for* $a \in G$*, called* exponent of a*;* $n \in \mathbf{N}$ *with* $x^n = e$ *for every* $x \in G$*, called* exponent of G

Definition 3.24 (period of group elements) *for group,* G*, and* $a \in G$*, smallest* $n \in \mathbf{N}$ *with* $a^n = e$*, called* period of a

Proposition 3.10 (period of elements of finite groups) *for finite group,* G*, of order,* $n > 1$*, period of every element,* $a \neq e$*, devided* n*; if* n *is prime number,* G *is cyclic, and period of every generator is* p

Proposition 3.11 (subgroups of cyclic groups) *every subgroup of cyclic group is cyclic and image of every homeomorphism of cyclic group is cyclic*

Properties of cyclic groups

Proposition 3.12 (properties of cyclic groups)

- *infinity cyclic group has exactly two generators; if a is one, a^{-1} is the other*
- *for cyclic group, G, of order, n, and generator, x, set of generators of G is*

$$\{x^m | m \text{ is relatively prime to } n\}$$

- *for cyclic group, G, and two generators, a, b, exists automorphism of G mapping a onto b; conversely, every automorphism maps a to some generator*
- *for cyclic group, G, of order, n, and $d \in \mathbf{N}$ dividing n, exists unique subgroup of order, d*
- *for cyclic groups, G_1 and G_2, of orders, n and m respectively with n and m relatively prime, $G_1 \times G_2$ is cyclic group*
- *for non-cyclic finite abelian group, G, exists subgroup isomorphic to $C \times C$ with C cyclic with prime order*

Symmetric groups and permutations

Definition 3.25 (symmetric groups and permutations) *for nonempty set, S, group G, of bijective functions of S onto itself with law of composition being function composition, called* symmetric group of S, *denoted by* $\mathrm{Perm}(S)$*; elements in* $\mathrm{Perm}(S)$ *called* permutations of S*; element swapping two disjoint elements in S leaving every others left, called* transposition

Proposition 3.13 (sign homeomorphism of finite symmetric groups) *for finite symmetric group, S_n, exits unique homeomorphism $\epsilon : S_n \to \{-1, 1\}$ mapping every transposition, τ, to -1, i.e., $\epsilon(\tau) = -1$*

Definition 3.26 (alternating groups) *element of finite symmetric group, σ, with $\epsilon(\sigma) = 1$, called* even, *element, σ, with $\epsilon(\sigma) = -1$, called* odd*; kernel of ϵ, called* alternating group, *denoted by* A_n

Theorem 3.5 (solvability of finite symmetric groups) *symmetric group, S_n with $n \geq 5$ is* not *solvable*

Theorem 3.6 (simplicity of alternating groups) *alternating group, A_n, with $n \geq 5$ is simple*

Operations of group on set

Definition 3.27 (operations of group on set) *for group, G, and set, S, homeomorphism,*

$$\pi : G \to \mathrm{Perm}(S)$$

called operation of G on S *or* action of G on S

- *S, called* G-set
- *denote $\pi(x)$ for $x \in G$ by* π_x*, hence homeomorphism denoted by* $x \mapsto \pi_x$
  - obtain mapping from such operation, $G \times S \to S$, with $(x, s) \mapsto \pi_x(s)$
  - often abbreviate $\pi_x(s)$ by xs, with which two properties satisfied
    - $(\forall x, y \in G, s \in S)\,(x(ys) = (xy)s)$
    - $(\forall s \in S)\,(es = s)$
  - conversely, for mapping, $G \times S \to S$ with $(x, s) \mapsto xs$ satisfying above two properties, $s \mapsto xs$ is permutation for $x \in G$, hence π_x is homeomorphism of G into $\mathrm{Perm}(S)$
  - thus, operation of G on S can be defined as mapping, $S \times G \to S$, satisfying above two properties

Conjugation

Definition 3.28 (conjugation of groups) *for group, G, and map, $\gamma_x : G \to G$ with $\gamma_x(y) = xyx^{-1}$, homeomorphism*

$$G \to \operatorname{Aut}(G) \text{ defined by } x \mapsto \gamma_x$$

called conjugation, which is operation of G on itself

- γ_x, called *inner*
- kernel of conjugation is *center of G*
- to avoid confusion, instead of writing xy for $\gamma_x(y)$, write
$$\gamma_x(y) = xyx^{-1} = {}^x y \text{ and } \gamma_{x^{-1}}(y) = x^{-1}yx = y^x$$
- for subset, $A \subset G$, map, $(x, A) \mapsto xAx^{-1}$, is operation of G on set of subsets of G
- similarly for subgroups of G
- two subsets of G, A and B with $B = xAx^{-1}$ for some $x \in G$, said to be *conjugate*

Translation

Definition 3.29 (translation) *operation of G on itself defined by map*

$$(x, y) \mapsto xy$$

called translation, denoted by $T_x : G \to G$ with $T_x(y) = xy$

- for subgroup, $H \subset G$, $T_x(H) = xH$ is left coset
  - denote set of left cosets also by G/H even if H is not normal
  - denote set of right cosets also by $H\backslash G$
- examples of translation
  - $G = GL(V)$, group of linear automorphism of vector space with field, F, for which, map, $(A, v) \mapsto Av$ for $A \in G$ and $v \in V$, defines operation of G on V
    - G is subgroup of group of permutations, $\operatorname{Perm}(V)$
  - for $V = F^n$, G is group of nonsingular n-by-n matrices

Isotropy

Definition 3.30 (isotropy) *for operation of group, G, on set, S,*

$$\{x \in G | xs = s\}$$

called isotropy of G, denoted by G_s, which is subgroup of G

- for conjugation operation of group, G, G_s is normalizer of $s \in G$
- isotropy groups are conjugate, *e.g.*, for $s, s' \in S$ and $y \in G$ with $ys = s'$,
$$G_{s'} = yG_s y^{-1}$$
- by definition, kernel of operation of G on S is
$$K = \bigcap_{s \in S} G_s \subset G$$
- operation with trivial kernel, said to be *faithful*
- $s \in G$ with $G_s = G$, called *fixed point*

Orbits of operation

Definition 3.31 (orbits of operation) *for operation of group, G, on set, S, $\{xs|x \in G\}$, called* orbit of s under G*, denoted by* Gs

- for $x, y \in G$ in same coset of G_s, $xs = ys$, *i.e.* $(\exists z \in G)\,(x, y \in zG_s) \Leftrightarrow xs = ys$
- hence, mapping, $G/G_s \to S$, with $x \mapsto xG_s$, is morphism of G-sets, thus

Proposition 3.14 *for group, G, operating on set, S, and $s \in S$, order of orbit, Gs, is equal to index, $(G : G_s)$*

Proposition 3.15 *for subgroup, H, of group, G, number of conjugate subgroups to H is index of normalizer of H in G*

Definition 3.32 (transitive operation) *operation with one orbit, said to be* transitive

Orbit decomposition and class formula

- orbits are disjoint
$$S = \coprod_{\lambda \in \Lambda} Gs_\lambda$$
where s_λ are elements of distinct orbits

Formula 3.1 (orbit decomposition formula) *for group, G, operating on set, S, index set, Λ, whose elements represent distinct orbits*
$$|S| = \sum_{\lambda \in \Lambda} (G : G_\lambda)$$

Formula 3.2 (class formula) *for group, G, and set, $C \subset G$, whose elements represent distinct conjugacy classes*
$$(G : 1) = \sum_{x \in C} (G : G_x)$$

Sylow subgroups

Definition 3.33 (sylow subgroups) *for prime number, p, finite group with order, p^n for some $n \geq 0$, called* p-group*; subgroup $H \subset G$, of finite group, G, with order, p^n for some $n \geq 0$, called* p-subgroup*; subgroup of order, p^n, where p^n is highest power of p dividing order of G, called* p-Sylow subgroup

Lemma 3.2 *finite abelian group of order divided by prime number, p, has subgroup of order p*

Theorem 3.7 (p-Sylow subgroups of finite groups) *finite group of order divided by prime number, p, has p-Sylow subgroup*

Lemma 3.3 (number of fixed points of group operations) *for p-group, H, operating on finite set, S*

- *number of fixed points of H is congruent to size of S modulo p, i.e.*
$$\#\text{ fixed points of } H \equiv |S| \pmod{p}$$
- *if H has exaxctly one fixed point, $|S| = 1 \pmod{p}$*
- *if p divides $|S|$, $|S| = 0 \pmod{p}$*

Sylow subgroups and solvability

Theorem 3.8 (solvability of finite p-groups) *finite p-group is solvable; if it is non-trivial, it has non-trivial center*

Corollary 3.1 *for non-trivial p-group, exists sequence of subgroups*

$$\{e\} = G_0 \subset G_1 \subset G_2 \subset \cdots \subset G_n = G$$

where G_i is normal in G and G_{i+1}/G_i is cyclic group of order, p

Lemma 3.4 (normality of subgroups of order p) *for finite group, G, smallest prime number dividing order of G, p, every subgroup of index p is normal*

Proposition 3.16 *group of order, pq, with p and q being distinct prime numbers, is solvable*

- now can prove following
  - group of order, 35, is solvable - implied by Proposition 3.8 and Proposition 3.12
  - group of order less than 60 is solvable

3.3 Rings

Rings

Definition 3.34 (rings) *set, A, together with two laws of composition called multiplication and addition which are written as product and sum respectively, satisfying following conditions, called* ring

- *A is commutative group with respect to addition - unit element denoted by* 0
- *A is monoid with respect to multiplication - unit element denoted by* 1
- *multiplication is distributive over addition, i.e.*

$$(\forall x, y, z \in A)\,((x+y)z = xz + yz \;\&\; z(x+y) = zx + zy)$$

do not assume $1 \neq 0$

- can prove, *e.g.*,
  - $(\forall x \in A)\,(0x = 0) \because 0x + x = 0x + 1x = (0+1)x = 1x = x$
  - if $1 = 0$, $A = \{0\} \because x = 1x = 0x = 0$
  - $(\forall x, y \in A)\,((-x)y = -(xy)) \because xy + (-x)y = (x + -x)y = 0y = 0$
  - $(\forall x, y \in A)\,((-x)(-y) = xy) \because (-x)(-y) + (-(xy)) = (-x)(-y) + (-x)y = (-x)(-y+y) = (-x)0 = 0$

Kinds of rings

Definition 3.35 (groups of units of rings) *subset, U, of ring, A, such that every element of U has both left and right inverses, called* group of units of A *or* group of invertible elements of A, *sometimes denoted by* A^*

Definition 3.36 (division rings) *ring with $1 \neq 0$ and every nonzero element being invertible, called* division ring

Definition 3.37 (multiplicative groups of invertible elements of rings) *we denote multiplicative groups of invertible elements of rings, A, by* A^*

Definition 3.38 (commutative rings) *ring, A, with* $(\forall x, y \in A)\,(xy = yx)$*, called* commutative ring

Definition 3.39 (fields) *commutative division ring, called* field

Definition 3.40 (subrings) *subset of ring which itself is ring with same additive and multiplicative laws of composition, called* subring

Definition 3.41 (center of rings) *subset, $C \subset A$, of ring, A such that*

$$C = \{a \in A | \forall x \in A, xa = ax\}$$

is subring*, called* center of ring, A

General distributivity and ring examples

- general distributivity - for ring, A, $\langle x_i\rangle_{i=1}^n \subset A$, $\langle y_i\rangle_{i=1}^n \subset A$

$$\left(\sum x_i\right)\left(\sum y_j\right) = \sum_i \sum_j x_i y_j$$

- ring examples
  - for set, S, and ring, A, set of all mappings of S into A, $\mathrm{Map}(S, A)$ whose addition and multiplication are defined as below, is *ring*

$$(\forall f, g \in \mathrm{Map}(S, A), x \in S)\,((f + g)(x) = f(x) + g(x))$$

$$(\forall f, g \in \mathrm{Map}(S, A), x \in S)\,((fg)(x) = f(x)g(x))$$

    - additive and multiplicative unit elements of $\mathrm{Map}(S, A)$ are constant maps whose values are additive and multiplicative of A respectively
  - for abelian group, M, set, $\mathrm{End}(M)$ of group homeomorphisms of M into itself is ring with normal addition and mapping composition as multiplication, is *ring* - not commutative in general

Group rings and convolution products

Definition 3.42 (group rings) *for group, G, and field, K, set of all formal linear combinations, $\sum_{x\in G} a_x x$, with $a_x \in K$ where a_x are zero except finite number of them, where addition is defined normally and multiplication is defined as*

$$\left(\sum_{x\in G} a_x x\right)\left(\sum_{y\in G} b_y y\right) = \sum_{z\in G}\left(\sum_{xy=z} a_x b_y xy\right)$$

called group ring*, denoted by* $K[G]$*, and $\sum_{xy=z} a_x b_y$, called* convolution product

Ideals of rings

Definition 3.43 (ideals) *subset, $\mathfrak{a}$, of ring, A, which is subgroup of additive group of A, with $A\mathfrak{a} \subset \mathfrak{a}$, called* left ideal *- indeed, $A\mathfrak{a} = \mathfrak{a}$ because A has 1;* right ideal *can be similarly defined, i.e., $\mathfrak{a}A = \mathfrak{a}$; subset which is both left and right ideal is* two-sided ideal

Definition 3.44 (principals) *for ring, A, and $a \in A$, left ideal, Aa, called* principal left ideal*; AaA, called* principal two-sided ideal

Definition 3.45 (generators of ideals) *for ring, A, and $a_1, \ldots, a_n$, set of elements of A of form*

$$\sum x_i a_i$$

with $x_i \in A$, is left ideal, and denoted by $(a_1, \ldots, a_n)$, called generators *of the left ideal; similarly for right ideals*

- above equal to smallest ideals containing a_i, *i.e.*, intersection of all ideals containing a_i $\bigcap_{a_1,\ldots,a_n \in \mathfrak{a}} \mathfrak{a}$ - just like set (σ-)algebras in set theory in §4.1 (proof can be found in Proof 3)

Definition 3.46 (principal rings) *commutative ring of which every ideal is principal and $1 \neq 0$, called* principal ring

Ring homeomorphisms

Definition 3.47 (ring homeomorphisms) *mapping of ring into ring, $f : A \to B$, such that f is monoid-homeomorphism for both additive and multiplicative structure on A and B, i.e.*

$$(a, b \in A)\,(f(a+b) = f(a) + f(b)\ \&\ f(ab) = f(a)f(b))$$

and

$$f(1) = 1\ \&\ f(0) = 0$$

called ring-homeomorphism*;* kernel*, defined to be kernel of f viewed as additive homeomorphism*

- *kernel of ring-homeomorphism*, $f : A \to B$, is ideal of A
- conversely, for ideal, $\mathfrak{a}$, can construct factor ring, $A/\mathfrak{a}$
- simply say "homeomorphism" if reference to rings is clear

Proposition 3.17 (injectivity of field homeomorphisms) *(ring) homeomorphism from field into field is injective*

Factor rings and canonical maps

Definition 3.48 (factor rings and residue classes) *for ring, A, $\mathfrak{a}$, set of cosets, $x + \mathfrak{a}$ for $x \in A$, combined with* addition *defined by viewing A and $\mathfrak{a}$ as additive groups,* multiplication *defined by*

$$(x + \mathfrak{a})(y + \mathfrak{a}) = xy + \mathfrak{a},$$

which satisfy all requirements for ring, called factor ring *or,* residue class ring*, denoted by* $A/\mathfrak{a}$*; cosets in $A/\mathfrak{a}$, called* residue classes modulo $\mathfrak{a}$*, and each coset, $x + \mathfrak{a}$, called* residue class of x modulo $\mathfrak{a}$

Definition 3.49 (canonical maps of rings) *ring-homeomorphism of ring, A, into factor ring, $A\mathfrak{a}$*

$$A \to A/\mathfrak{a}$$

called canonical map of A into $A/\mathfrak{a}$

Factor ring induced ring-homeomorphism

Proposition 3.18 (factor ring induced ring-homeomorphism) *for ring-homeomorphism, $f : A \to A'$ whose kernel contains ideal, $\mathfrak{a}$, exists unique ring-homeomorphism, $g_* : A/\mathfrak{a} \to A'$, making diagram in Figure 3.6 commutative*

- ring canonical map, $f : A \to A/\mathfrak{a}$, is universal in category of homeomorphisms whose kernel contains $\mathfrak{a}$

$$\begin{tikzcd}
A \arrow[rr, "g"] \arrow[dr, "f"'] & & A' \\
& A/\mathfrak{a} \arrow[ur, "g^*"'] &
\end{tikzcd}$$

Figure 3.6: factor-ring-induced-ring-homeomorphism

Embedding of rings

- bijective ring-homeomorphism, $f : A \to B$ is isomorphism
  - indeed, exists set-theoretic inverse, $g : B \to A$, of f, which is ring-homeomorphism
- image, $f(A)$, of ring-homeomorphism, $f : A \to B$, is subring of B

Definition 3.50 (embedding of rings) *ring-isomorphism between A and its image, established by injective ring-homeomorphism, $f : A \to B$, called* embedding of rings

Definition 3.51 (induced injective ring-homeomorphisms) *for ring-homeomorphism, $f : A \to A'$ and ideal $\mathfrak{a}'$ of A', injective ring-homeomorphism*

$$A/f^{-1}(\mathfrak{a}') \to A'/\mathfrak{a}'$$

called induced injective ring-homeomorphism

Zero divisors and entire rings

Definition 3.52 (zero divisors) *for ring, A, $x, y \in A$ with $x \neq 0$, $y \neq 0$, $xy = 0$, said to be* zero divisors

Definition 3.53 (entire rings) *commutative ring with no zero divisors for which $1 \neq 0$, said to be* entire*; entire rings sometimes called* integral domains

Ideals of commutative rings and fields

Definition 3.54 (prime ideals) *for commutative ring, A, ideal $\mathfrak{p} \neq A$ with $A/\mathfrak{p}$ entire, called* prime ideal *or just* ideal*; equivalently, ideal, $\mathfrak{p} \neq A$ is* prime if and only if

$$(\forall x, y \in A)\,(xy \in \mathfrak{p} \Rightarrow x \in \mathfrak{p} \text{ or } y \in \mathfrak{p})$$

Definition 3.55 (maximal ideals) *for commutative ring, A, ideal $\mathfrak{m} \neq A$ such that*

$$(\forall \text{ ideal } \mathfrak{a} \subset A)\,(\mathfrak{m} \subset \mathfrak{a} \Rightarrow \mathfrak{a} = A)$$

called maximal ideal

Lemma 3.5 (properties of prime and maximal ideals) *for commutative ring, A*

- *every maximal ideal is prime*
- *every ideal is contained in some maximal ideal*
- *ideal, $\{0\}$ is prime* if and only if *A is entire*
- *only ideals of field is either zero ideal or the field itself*
- *ideal, $\mathfrak{m}$, is maximal* if and only if *$A/\mathfrak{m}$ is field*
- *inverse image of prime ideal of commutative ring homeomorphism is prime*

Characteristic of rings

- consider ring-homeomorphism
$$\lambda : \mathbf{Z} \to A$$
for ring, A, such that
$$\lambda(n) = ne$$
where e is multiplicative unit element of A
  - kernel of λ is ideal (n), for $n \in \mathbf{Z}$, *i.e.*, ideal generated by integer, n
  - hence, canonical injective ring-homeomorphism, $\mathbf{Z}/n\mathbf{Z} \to A$, which is ring-isomorphism between $\mathbf{Z}/n\mathbf{Z} \to A$ and subring of A
  - when $n\mathbf{Z}$ is prime ideal, exist two cases; either $n = 0$ or $n = p$ for prime number, p

Definition 3.56 (characteristic of rings) *ring, A, with $\{0\}$ as prime ideal kernel above, said to have characteristic 0; if prime ideal kernel is $p\mathbf{Z}$ for prime number, p, A, said to have characteristic p, in which case, A contains (isomorphic image of) $\mathbf{Z}/p\mathbf{Z}$ as subring, abbreviated by $\mathbf{F}_p$*

Prime fields and prime rings

- field, K, has characteristic 0 or p for prime number, p
- K contains (isomorphic image of) $\mathbf{Q}$ or $\mathbf{F}_p$ for first and second cases respectively, as subfield

Definition 3.57 (prime fields and prime rings) *in above cases, both $\mathbf{Q}$ and $\mathbf{F}_p$, called prime field (contained in K); since prime field is smallest subfield of K containing 1 having no automorphism other than identity, identify it with $\mathbf{Q}$ or $\mathbf{F}_p$ for each case; prime ring (contained in K) means either integers, $\mathbf{Z}$, if K has characteristic 0, or $\mathbf{F}_p$ if K has characteristic p*

$\mathbf{Z}/n\mathbf{Z}$

- $\mathbf{Z}$ is ring
- every ideal of $\mathbf{Z}$ is principal (proof can be found in Proof 4), *i.e.*, either $\{0\}$ or $n\mathbf{Z}$ for some $n \in \mathbf{N}$
- ideal of $\mathbf{Z}$ is prime *if and only if* is $p\mathbf{Z}$ for some prime number, $p \in \mathbf{N}$
  - $p\mathbf{Z}$ is maximal ideal

Definition 3.58 (rings of integers modulo n) *$\mathbf{Z}/n\mathbf{Z}$, called ring of integers modulo n; abbreviated as mod n*

- $\mathbf{Z}/p\mathbf{Z}$ for prime p is *field* and denoted by $\mathbf{F}_p$

Euler phi-function

Definition 3.59 (Euler phi-function) *for $n > 1$, order of divison ring of $\mathbf{Z}/n\mathbf{Z}$, calleds Euler phi-function, denoted by $\varphi(n)$; if prime factorization of n is*
$$n = p_1^{e_1} \cdots p_r^{e_r}$$
with distinct prime numbers, p_i, and $e_i \geq 1$
$$\varphi(n) = p_1^{e_1-1}(p_1 - 1) \cdots p_r^{e_r-1}(p_r - 1)$$

Theorem 3.9 (Euler's theorem) *for x prime to n,*
$$x^{\varphi(n)} \equiv 1 \pmod{n}$$

Chinese remainder theorem

Theorem 3.10 (Chinese remainder theorem) *for ring, A and $n \geq 2$ ideals, $\mathfrak{a}_1, \ldots \mathfrak{a}_n$ with $\mathfrak{a}_i + \mathfrak{a}_j = A$ for all $1 \leq i \neq j \leq n$*

$$(\forall x_1, \ldots, x_n \in A)\,(\exists x \in A)\,(\forall 1 \leq i \leq n)\,(x \equiv x_i \pmod{\mathfrak{a}_i})$$

Corollary 3.2 (isomorphism induced by Chinese remainder theorem) *for ring, A, $n \geq 2$ ideals, $\mathfrak{a}_1, \ldots \mathfrak{a}_n$ with $\mathfrak{a}_i + \mathfrak{a}_j = A$ for all $1 \leq i \neq j \leq n$, and map of A into product induced by canonical map of A onto $A/\mathfrak{a}_i$ for each factor, i.e.,*

$$f : A \to \prod A/\mathfrak{a}_i$$

- *f is surjective,*
- $\operatorname{Ker} f = \bigcap \mathfrak{a}_i$,
- *hence, there is isomorphism*

$$A/\cap \mathfrak{a}_i \approx \prod A/\mathfrak{a}_i$$

Isomorphisms of endomorphisms of cyclic groups

Theorem 3.11 (isomorphisms of endomorphisms of cyclic groups) *for cyclic group, A, of order, n, endomorphisms of A into A, with, $x \mapsto kx$, for $k \in \mathbf{Z}$ induce*

- *ring isomorphism*

$$\mathbf{Z}/n\mathbf{Z} \approx \operatorname{End}(A)$$

  and

- *group isomorphism*

$$(\mathbf{Z}/n\mathbf{Z})^* \approx \operatorname{Aut}(A)$$

where B^* denotes multiplicative groups of invertible elements of ring, B, as defined in Definition 3.37

- *e.g.*, for group of n-th roots of unity in $\mathbf{C}$, μ_n, all automorphisms are given by

$$\xi \mapsto \xi^k$$

  for $k \in (\mathbf{Z}/n\mathbf{Z})^*$

Irreducibility and factorial rings

Definition 3.60 (irreducible ring elements) *for entire ring, A, non-unit non-zero element, $a \in A$, with*

$$(\forall b, c \in A)\,(a = bc \Rightarrow b \textit{ or } c \textit{ is unit})$$

said to be irreducible

Definition 3.61 (unique factorization into irreducible elements) *for entire ring, A, element, $a \in A$, for which exists unit, u, and irreducible elements, p_1, $\ldots$, p_r in A such that*

$$a = u \prod p_i$$

and this expression is unique up to permutation and multiplications by units, said to have unique factorization into irreducible elements

Definition 3.62 (factorial rings) *entire ring with every non-zero element has unique factorial into irreducible elements, called* factorial ring *or* unique factorization ring

Devision and greatest common divisor (g.c.d.)

Definition 3.63 (devision of entire ring elements) *for entire ring, A, and nonzero elements, $a, b \in A$, if exists $c \in A$ such that $ac = b$, we say* a divides b, *denoted by* $a|b$

Definition 3.64 (greatest common divisor (g.c.d.)) *for entire ring, A, and $a, b \in A$, $d \in A$ which divides a and b and satisfies*

$$(\forall c \in A \text{ with } c|a \text{ \& } c|b)\,(c|d)$$

called greatest common divisor (g.c.d.) of a and b

Proposition 3.19 (existence of greatest common divisor of principal entire rings) *for principal entire ring, A, and nonzero $a, b \in A$, $c \in A$ with $(a, b) = (c)$ is g.c.d. of a and b*

Theorem 3.12 (principal entire ring is factorial) *principal entire ring is factorial*

3.4 Polynomials

Why (rings of) polynomials?

- lays ground work for polynomials in general
- needs polynomials over arbitrary rings for diverse purposes
  - polynomials over finite field which cannot be identified with polynomial functions in that field
  - polynomials with integer coefficients; reduce them mod p for prime, p
  - polynomials over arbitrary commutative rings
  - rings of polynomial differential operators for algebraic geometry & analysis

Defintion of polynomials

- exist many ways to define polynomials over commutative ring
- here's one

Definition 3.65 (polynomials) *for ring, A, infinite cyclic group generated by X, and monoid $S = \{X^r | r \in \mathbf{Z}, r \geq 0\}$, set of functions from S into A which are equal to 0 except finite number of elements of S, called* polynomials over A, *denoted by* $A[X]$

- for every $a \in A$, define function which has value, a, on X^n, and value, 0, for every other element of S, by aX^r
- then, a polynomial can be uniquely written as

  $$f(X) = a_0X^0 + \cdots + a_nX^n$$

  for some nonnegative $n \in \mathbf{Z}$, $a_i \in A$
- a_i, called *coefficients of f*

Polynomial functions

Definition 3.66 (polynomial functions) *for two rings, A and B, with $A \subset B$, and $f \in A[X]$ with $f(X) = a_0 + a_1 X + \cdots + a_n X^n$, map, $f_B : B \to B$ defined by*

$$f_B(x) = a_0 + a_1 x + \cdots + a_n x^n$$

called polynomial function associated with $f(X)$

Definition 3.67 (evaluation homeomorphism) *for two rings, A and B, with $A \subset B$, and $b \in B$, ring homeomorphism from $A[X]$ into B with association, $\mathrm{ev}_b : f \mapsto f(b)$, called* evaluation homeomorphism*, said to be obtained by* substituting b for X in f

- hence, for $x \in B$, subring, $A[x]$ of B generated by x over A is ring of all polynomial values, $f(x)$, for $f \in A[X]$

Definition 3.68 (variables and transcendentality) *for two rings, A and B, with $A \subset B$, if x makes evaluation homeomorphism, $\mathrm{ev}_x : f \mapsto f(x)$, isomorphic, x, said to be* transcendental over A *or* variable over A

- in particular, X *is variable over* A

Polynomial examples

- consider $\alpha = \sqrt{2}$ and $\{a + b\alpha \,|a, b \in \mathbf{Z}\}$, subring of $\mathbf{Z}[\alpha] \subset \mathbf{R}$ generated by α.
  - α is *not* transcendental because $f(\alpha) = 0$ for $f(X) = X^2 - 1$, hence kernel of evaluation map of $\mathbf{Z}[X]$ into $\mathbf{Z}[\alpha]$ is not injective, hence not isomorphism
- consider $\mathbf{F}_p$ for prime number, p
  - $f(X) = X^p - X \in \mathbf{F}_p[X]$ is not zero polynomial, but because $x^{p-1} \equiv 1$ for every x prime to p by Theorem 3.9 (Euler's theorem), hence $x^p \equiv x$ for every $x \in \mathbf{F}_p$, thus for polynomial function, $f_{\mathbf{F}_p}$, $f_{\mathbf{F}_p}(x) = 0$ for every x in $\mathbf{F}_p$
  - *i.e.*, non-zero polynomial induces zero polynomial function

Reduction map

- for homeomorphism, $\varphi : A \to B$, of commutative rings, exists associated homeomorphisms of polynomials rings, $A[X] \to B[X]$, such that

$$f(X) = \sum a_i X^i \mapsto \sum \varphi(a_i) X^i = (\varphi f)(X)$$

Definition 3.69 (reduction maps) *above ring homeomorphism, $f \mapsto \varphi f$, called* reduction map

- *e.g.*, for complex conjugate, $\varphi : \mathbf{C} \to \mathbf{C}$, homeomorphism of $\mathbf{C}[X]$ into itself can be obtained by reduction map, $f \mapsto \varphi f$, which is complex conjugate of polynomials with complex coefficients

Definition 3.70 (reduction of f modulo p) *for prime ideal, $\mathfrak{p}$, of ring, A, surjective canonical map, $\varphi : A \to A/\mathfrak{p}$, reduction map, φf for $f \in A[X]$, sometimes called* reduction of f modulo p

Basic properties of polynomials in one variable

Theorem 3.13 (Euclidean algorithm) *for set of all polynomials in one variable of nonnegative degrees, $A[X]$, with commutative ring, A*

$$(\forall f, g \in A[X] \text{ with leading coefficients of } g \text{ unit in } A)$$
$$(\exists q, r \in A[X] \text{ with } \deg r < \deg g)\,(f = qg + r)$$

Theorem 3.14 (principality of polynomial rings) *polynomial ring in one variable, $k[X]$, with field, k, is principal*

Constant, monic, and irreducible polynomials

Definition 3.71 (constant and monic polynomials) *$k \in k[X]$, with field, k, called constant polynomial; $f(x) \in k[X]$ with leading coefficient, 1, called monic polynomial*

Definition 3.72 (irreducible polynomials) *polynomial $f(x) \in k[X]$ such that*

$$(\forall g(X), h(X) \in k[X])\,(f(X) = g(X)h(X) \Rightarrow g(X) \in k \text{ or } h(X) \in k)$$

said to be irreducible

Root or zero of polynomials

Definition 3.73 (roots) *for commutator ring, B, its subring, $A \subset B$, and $f(x) \in A[X]$ in one variable, $b \in B$ satisfying*

$$f(b) = 0$$

called root or zero of f;

Theorem 3.15 (roots of polynomials) *for field, k, polynomial, $f \in k[X]$ in one variable of degree, $n \geq 0$, has at most n roots in k; if a is root of f in k, $X - a$ divides $f(X)$*

Induction of zero functions

Corollary 3.3 (induction of zero functions in one variable) *for field, k, infinite subset, $T \subset k$, if polynomial, $f \in k[X]$, in one variable over k, satisfies*

$$(\forall a \in k)\,(f(a) = 0)$$

then $f(0) = 0$, i.e., f induces zero function

Corollary 3.4 (induction of zero functions in multiple variables) *for field, k, and n infinite subsets of k, $\langle S_i \rangle_{i=1}^{n}$, if polynomial in n variables over field, k, satisfies*

$$(\forall a_i \in S_i \text{ for } 1 \leq i \leq n)\,(f(a_1, \ldots, a_n) = 0)$$

then $f = 0$, i.e., f induces zero function

Corollary 3.5 (induction of zero functions in multiple variables - infinite fields) *if polynomial in n variables over infinite field, k, induces zero function in $k^{(n)}$, $f = 0$*

Corollary 3.6 (induction of zero functions in multiple variables - finite fields) *if polynomial in n variables over finite field, k, of order, q, degree of which in each variable is less than q, induces zero function in $k^{(n)}$, $f = 0$*

Reduced polynomials and uniqueness

- for field, k, with q elements, polynomial in n variables over k can be expressed as

$$f(X_1, \ldots, X_n) = \sum a_i X_1^{\nu_{i,1}} \cdots X_n^{\nu_{i,n}}$$

  for finite sequence, $\langle a_i \rangle_{i=1}^m$, and $\langle \nu_{1,j} \rangle_{j=1}^n$, ..., $\langle \nu_{m,j} \rangle_{j=1}^n$ where $a_i \in k$ and $\nu_{i,j} \geq 0 \in \mathbf{Z}$

- for each $\nu_{i,j}$, exist $g_{i,j} \geq 0$ and $0 \leq r_{i,j} < q$ such that $\nu_{i,j} = g_{i,j} q + r_{i,j}$ (by Euclidean algorithm), thus f can be rewritten as

$$f(X_1, \ldots, X_n) = \sum a_i X_1^{r_{i,1}} \cdots X_n^{r_{i,n}}$$

  because $X_i^q = 1$

- called *reduced polynomial*, denoted by f^*

Corollary 3.7 (uniqueness of reduced polynomials) *for field, k, with q elements, reduced polynomial is unique (by Corollary 3.6)*

Multiplicative subgroups and n-th roots of unity

Definition 3.74 (multiplicative subgroups of fields) *for field, k, subgroup of group, $k^* = k \sim \{0\}$, called* multiplicative subgroup of k

Theorem 3.16 (finite multiplicative subgroups of field is cyclic) *finite multiplicative subgroup of field, k, is cyclic*

Corollary 3.8 (multiplicative subgroup of finite field is cyclic) *multiplicative subgroup of finite field is cyclic*

Definition 3.75 (primitive n-th roots of unity) *generator for group of n-th roots of unity, called* primitive n-th root of unity*; group of roots of unity, denoted by* μ*; group of roots of unity in field, k, denoted by* $\mu(k)$

Algebraic closedness

Definition 3.76 (algebraically closed) *field, k, for which every polynomial in $k[X]$ of positive degree has root in k, said to be* algebraically closed

- *e.g.*, complex numbers are algebraically closed
- every field is contained in some algebraically closed field
- for algebraically closed field, k
  - every irreducible polynomial in $k[X]$ is of degree 1
  - unique factorization of polynomial of nonnegative degree can be written in form

$$f(X) = c \prod_{i=1}^{r} (X - \alpha_i)^{m_i}$$

    with $c \neq 0 \in k$ and distinct roots, $\alpha_1, \ldots, \alpha_r$

Derivatives of polynomials

Definition 3.77 (derivatives of polynomials over commutative rings) *for polynomial $f(X) = a_n X^n + \cdots + a_0 \in A[X]$ with commutative ring, A, where $a_i \in A$, map, $D : A[X] \to A[X]$ defined by*

$$Df(X) = na_n X^{n-1} + \cdots + a_1$$

called derivative of polynomial, *denoted by $f'(X)$;*

- for $f, g \in A[X]$ with commutative ring, A, and $a \in A$

$$(f + g)' = f' + g' \;\&\; (fg)' = f'g + fg' \;\&\; (af) = af'$$

Multiple roots and multiplicity

Definition 3.78 (multiplicity and multiple roots) *nonzero polynomial, $f(X) \in k[X]$, in one variable over field, k, having $a \in k$ as root can be written of form*

$$f(X) = (X - a)^m g(X)$$

with some polynomial, $g(X) \in A[X]$ relatively prime to $(X - a)$ (hence, $g(a) \neq 0$); m here, called multiplicity of a in f, *when $m > 1$, a, said to be* multiple root of f

Proposition 3.20 (necessary and sufficient condition for multiple roots) *for polynomial, f, of one variable over field, k, $a \in k$ is multiple root of f* if and only if *is root of f and $f'(a) = 0$*

Proposition 3.21 (derivatives of polynomials) *for polynomial, $f \in K[X]$ over field, K, of positive degree, $f' \neq 0$ if K has characteristic 0; if K has characteristic $p > 0$, $f' = 0$* if and only if

$$f(X) = \sum_{\nu=1}^{n} a_\nu X^\nu$$

where p divides each integer, ν, with $a_\nu \neq 0$

Fields having characteristic p

- for prime, p, $p | \binom{p}{\nu}$ for all $1 \leq \nu < p$, hence for field, K, having characteristic p, and $a, b \in K$,

$$(a + b)^p = a^p + b^p$$

- applying this resurvely r times yields

$$(a + b)^{p^r} = (a^p + b^p)^{p^{r-1}} = (a^p + b^p)^{p^{r-2}} = \cdots = a^{p^r} + b^{p^r}$$

- thus, $(X - a)^{p^r} = X^{p^r} - a^{p^r}$
- if $c \in K$ and polynomial

$$X^{p^r} - c$$

  has root $a \in K$, then $a^{p^r} = c$ and

$$X^{p^r} - c = (X - a)^{p^r}$$

- thus, has precisely one root, a, of multiplicity, p^r!

Frobenius endomorphisms

- homeomorphism of K into itself, $x \mapsto x^p$, has trivial kernel, hence injective
- hence, iterating $r \geq 1$ times yields endomorphism, $x \mapsto x^{p^r}$

Definition 3.79 (Frobenius endomorphisms) *for field, K, prime number, p, and $r \geq 1$, endomorphism of K into itself, $x \mapsto x^{p^r}$, called* Frobenius endomorphism

3.5 Algebraic Extensions

Algebraic extensions

- will show that for polynomial over field, always exists some extension of the field where the polynomial has root
- will show existence of algebraic closure for every field

Extensions of fields

Definition 3.80 (extensions of fields) *for field, E, and subfield, $F \subset E$, E said to be* extension field of F,

- *can view E vector space over F*
- *if dimension of the vector space is finite, the extension called* finite extension of F
- *if infinite, called* infinite extension of F

- extension, $F \subset E$, sometimes denoted by E/F, which should *not* confused with *factor group*, only when unambiguously understood

Algebraic over fields

Definition 3.81 (algebraic over fields) *for field, E, and subfield, $F \subset E$, $\alpha \in E$ with*

$$(\exists a_0, \ldots, a_n \ (n \geq 1) \text{ with not all } a_i \text{ zero}) \, (a_0 + a_1\alpha + \cdots + a_n\alpha^n = 0)$$

said to be algebraic over F

- *for algebraic $\alpha \neq 0$, can always find such equation like above that $a_0 \neq 0$*

- equivalent statement to Definition 3.81
  - for field, E, and subfield, $F \subset E$, $\alpha \in E$ is algebraic over F *if and only if* exists homeomorphism, $\varphi : F[X] \to E$ such that

$$(\forall x \in F) \, (\varphi(x) = x) \ \& \ \varphi(X) = \alpha \ \& \ \operatorname{Ker} \varphi \neq \{0\}$$

- in which case, $\operatorname{Ker} \varphi$ is principal ideal (by Theorem 3.14), hence generated by single element, thus exists nonzero $p(X) \in F[X]$ (with normalized leading coefficient being 1) so that

$$F[X]/(p(X)) \approx F(\alpha)$$

- $F(\alpha)$ entire, hence $p(X)$ irreducible (refer to Definition 3.54)

Definition 3.82 (THE irreducible polynomial) *normalized $p(X)$ (i.e., with leading coefficient being 1) uniquely determined by α, called* THE irreducible polynomial of α over F, *denoted by* $\operatorname{Irr}(\alpha, F, X)$

Algebraic extensions

Definition 3.83 (algebraic extensions) *for field, F, its extension field, every element of which is algebraic over F, said to be* algebraic extension of F

Proposition 3.22 (algebraic-ness of finite field extensions) *every finite extension field of field, F, is algebraic over F*

- converse is *not* true, *e.g.*, subfield of complex numbers consisting of algebraic numbers over $\mathbf{Q}$ is infinite extension of $\mathbf{Q}$

Dimension of extensions

Definition 3.84 (dimension of extensions) *for field, F, dimension of its extension, E, as vector space over F, called* dimension of E over F, *denoted by* $[E : F]$

Proposition 3.23 (dimension of finite extensions) *for field, k, and its extension fields, F and E, with $k \subset F \subset E$,*

$$[E : k] = [E : F][F : k]$$

- *if $\langle x_i \rangle_{i\in I}$ is basis for F over k, and $\langle y_j \rangle_{j\in J}$ is basis for E over F, $\langle x_i y_j \rangle_{(i,j)\in I\times J}$ is basis for E over k*

Corollary 3.9 (finite dimensions of extensions) *for field, k, and its extension fields, F and E, with $k \subset F \subset E$, extension of E over k is finite* if and only if *extension of F over k is finite and extension of E over F is finite*

Generation of field extensions

Definition 3.85 (generation of field extensions) *for field, k, its extension field, E, and $\alpha_1, \ldots, \alpha_n \in E$, smallest subfield containing k and $\alpha_1, \ldots, \alpha_n$, said to be* finitely generated over k by $\alpha_1, \ldots, \alpha_n$, *denoted by* $k(\alpha_1, \ldots, \alpha_n)$

- $k(\alpha_1, \ldots, \alpha_n)$ consists of all quotients, $f(\alpha_1, \ldots, \alpha_n)/g(\alpha_1, \ldots, \alpha_n)$ where $f, g \in k[X]$ and $g(\alpha_1, \ldots, \alpha_n) \neq 0$, *i.e.*

$$\begin{aligned} &k(\alpha_1, \ldots, \alpha_n) \\ &\quad = \quad \{f(\alpha_1, \ldots, \alpha_n)/g(\alpha_1, \ldots, \alpha_n) \,|\, f, g \in f[X], g(\alpha_1, \ldots, \alpha_n) \neq 0\} \end{aligned}$$

- *any* field extension, $k \subset E$, is union of smallest subfields containing $\alpha_1, \ldots, \alpha_n$ where $\alpha_1, \ldots, \alpha_n$ range over finite set of elements of E, *i.e.*

$$E = \bigcup_{n\in\mathbf{N}} \bigcup_{\alpha_1,\ldots,\alpha_n\in E} k(\alpha_1, \ldots, \alpha_n)$$

Proposition 3.24 (finite extension is finitely generated) *every finite extension of field is finitely generated*

Tower of fields

Definition 3.86 (tower of fields) *sequence of extension fields*

$$F_1 \subset F_2 \subset \cdots \subset F_n$$

called tower of fields

Definition 3.87 (finite tower of fields) *tower of fields, said to be* finite if and only if *each step of extensions is finite*

Algebraic-ness of finitely generated subfields

Proposition 3.25 (algebraic-ness of finitely generated subfields by single element) *for field, k and $\alpha \in k$ being algebraic over k,*

$$k(\alpha) = k[\alpha]$$

and

$$[k(\alpha) : k] = \deg \operatorname{Irr}(\alpha, k, X)$$

hence $k(\alpha)$ is finite extension of k, thus algebraic extension over k *(by Proposition 3.22)*

Lemma 3.6 (a fortiori algebraic-ness) *for field, k, its extension, F, and α being algebraic over k where $k(\alpha)$ and F are subfields of common field, α is algebraic over F*

- *indeed,* $\operatorname{Irr}(\alpha, k, X)$ *has* a fortiori *coefficients in F*
  - assume tower of fields
    $$k \subset k(\alpha_1) \subset k(\alpha_1, \alpha_2) \subset \cdots \subset k(\alpha_1, \ldots, \alpha_n)$$
    where α_i is algebraic over k
  - then, α_{i+1} is algebraic over $k(\alpha_1, \ldots, \alpha_i)$ (by Lemma 3.6)

Proposition 3.26 (algebraic-ness of finitely generated subfields by multiple elements) *for field, k, and α_1, $\ldots$, α_n being algebraic over k, $E = k(\alpha_1, \ldots, \alpha_n)$ is finitely algebraic over k (because of Proposition 3.25, Corollary 3.9, and Proposition 3.22). Indeed, $E = k[\alpha_1, \ldots, \alpha_n]$ and*

$$\begin{aligned} &[k(\alpha_1, \ldots, \alpha_n) : k] \\ &= \operatorname{Irr}(\alpha_1, k, X) \operatorname{Irr}(\alpha_2, k(\alpha_1), X) \cdots \operatorname{Irr}(\alpha_n, k(\alpha_1, \ldots, \alpha_{n-1}), X) \end{aligned}$$

(proof can be found in Proof 5)

Compositum of subfields and lifting

Definition 3.88 (compositum of subfields) *for field, k, its extensions, E and F, which are subfields of common field, L, smallest subfield of L containing both E and F, called* compositum of E and F in L, *denoted by EF*

- *cannot define compositum if E and F are not embedded in common field, L*
  - could define *compositum of set of subfields of L* as smallest subfield containing subfields in the set
  - extension, E, of k is compositum of all its finitely generated subfields over k, *i.e.*, $E = \bigcup_{n \in \mathbf{N}} \bigcup_{\alpha_1, \ldots, \alpha_n \in E} k(\alpha_1, \ldots, \alpha_n)$

Lemma 3.7 (finite generation of compositum) *for field, k, $E = k(\alpha_1, \ldots, \alpha_n)$, and F, any extension of k where both E and F are contained in common field, L,*

$$EF = F(\alpha_1, \ldots, \alpha_n),$$

i.e., compositum, EF, is finitely generated over F (proof can be found in Proof 6)

- *refer to diagra in Figure 3.7*

Definition 3.89 (lifting) *extension, EF, of F, called* translation of E to F *or* lifting of E to F

- *often draw diagram as in Figure 3.8*

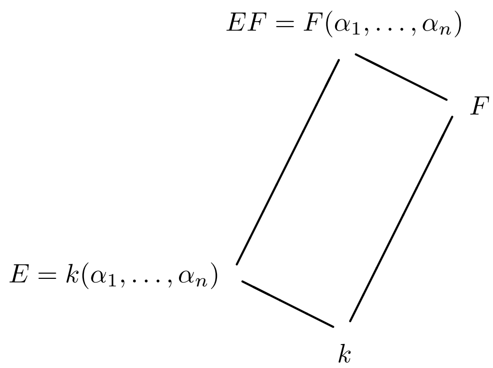

Figure 3.7: lifting or smallest fields

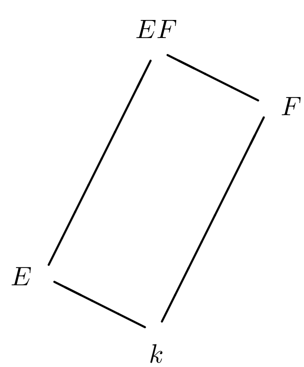

Figure 3.8: translation or lifting of fields

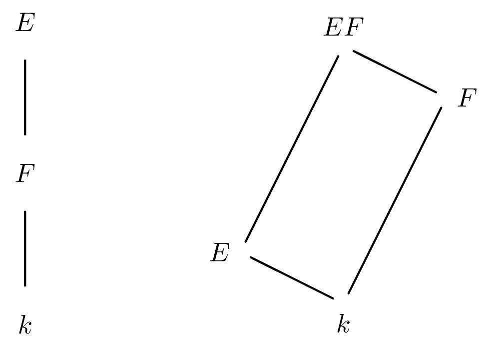

Figure 3.9: lattice diagram of fields

Distinguished classes

Definition 3.90 (distinguished class of field extensions) *for field, k, and its extension, E, class, $\mathcal{C}$, of extension fields satisfying*

- *for tower of fields, $k \subset F \subset E$, extension $k \subset E$ is in $\mathcal{C}$* if and only if *both $k \subset F$ and $F \subset E$ are in $\mathcal{C}$*
- *if $k \subset E$ is in $\mathcal{C}$, F is any extension of k, and both E and F are subfields of common field, then $F \subset EF$ is in $\mathcal{C}$*
- *if $k \subset F$ and $k \subset E$ are in $\mathcal{C}$ and both E and F are subfields of common field, $k \subset EF$ is in $\mathcal{C}$*

said to be distinguished*; Figure 3.8 illustrates first two properties - last property follows formally from first two properties*

Both algebraic and finite extensions are distinguished

Proposition 3.27 (algebraic and finite extensions are distinguished) *class of algebraic extensions is distinguished, so is class of finite extensions*

- true that finitely generated extensions form distinguished class (not necessarily algebraic extensions or finite extensions)

Field embedding and embedding extensions

Definition 3.91 (field embedding) *for two fields, F and L, injective homeomorphism $\sigma : F \to L$, called* embedding of F into L*; then σ induces isomorphism of F with its image σF, which is sometimes written as* F^σ

Definition 3.92 (field embedding extensions) *for field embedding, $\sigma : F \to L$, field extension, $F \subset E$, and embedding, $\tau : E \to L$, whose restriction to F being equal to σ, said to* be over σ *or* extend σ*; if σ is identity, embedding, τ, called* embedding of E over F*; diagrams in Figure 3.10 show these embedding extensions*

- assuming same as in Definition 3.92, if $\alpha \in E$ is root of $f \in F[X]$, then α^τ is root of f^σ for if $f(X) = \sum_{i=0}^n a_i X^i$, then $f(\alpha) = \sum_{i=0}^n a_i \alpha^i = 0$, and $0 = f(\alpha)^\tau = \sum_{i=0}^n (a_i^\tau)(\alpha^\tau)^i = \sum_{i=0}^n a_i^\sigma (\alpha^\tau)^i = f^\sigma(\alpha^\tau)$

$$
\begin{tikzcd}
E \arrow[r, "\tau"] & L \\
F \arrow[u, "\text{inc}"] \arrow[r, "\sigma"'] & L \arrow[u, "\text{id}"']
\end{tikzcd}
\qquad\qquad
\begin{tikzcd}
A \arrow[rr, "\tau"] & & A' \\
& A/\mathfrak{a} \arrow[ul, "\text{inc}"] \arrow[ur, "\text{id}"'] &
\end{tikzcd}
$$

Figure 3.10: embedding extension

Embedding of field extensions

Lemma 3.8 (field embeddings of algebraic extensions) *for field, k, and its algebraic extension, E, embedding of E into itself over k is isomorphism*

Lemma 3.9 (compositums of fields) *for field, k, and its extensions, E and F, contained in common field,*

$$E[F] = F[E] = \bigcup_{n=1}^{\infty} \{e_1 f_1 + \cdots + e_n f_n \,| e_i \in E, f_i \in F\}$$

and EF is field of quotients of these elements

Lemma 3.10 (embeddings of compositum of fields) *for field, k, its extensions, E_1 and E_2, contained in commen field, E, and embedding $\sigma : E \to L$ for some field, L,*

$$\sigma(E_1 E_2) = \sigma(E_1)\sigma(E_2)$$

Existence of roots of irreducible polynomials

- assume $p(X) \in k[X]$ irreducible polynomial and consider canonical map, which is ring homeomorphism

$$\sigma : k[X] \to k[X]/((p(X))$$

- consider $\operatorname{Ker} \sigma|k$
  - every kernel of ring homeomorphism is ideal, hence if $a \neq 0 \in \operatorname{Ker} \sigma|k$, $1 \in \operatorname{Ker} \sigma|k$ because $a^{-1} \in \operatorname{Ker} \sigma|k$, but $1 \notin (p(X))$
  - thus, $\operatorname{Ker} \sigma|k = \{0\}$, hence $p^\sigma \neq 0$
- now for $\alpha = X^\sigma$

$$p^\sigma(\alpha) = p^\sigma(X^\sigma) = (p(X))^\sigma = 0$$

- thus, α is algebraic in k^σ, *i.e.*, $\alpha \in k[X]^\sigma$ is root of p^σ in $k^\sigma(\alpha)$

Lemma 3.11 (existence of root of irreducible polynomials) *for field, k, and irreducible $p(X) \in k[X]$ with $\deg p \geq 1$, exists homeomorphism of k, $\sigma : k \to L$, such that p^σ with $\deg p^\sigma \geq 1$ has root in field extension of k^σ*

Existence of algebraically closed algebraic field extensions

Proposition 3.28 (existence of extension fields containing roots) *for field, k, and $f \in k[X]$ with* $\deg f \geq 1$, *exists extension of k in which f has root*

Corollary 3.10 (existence of extension fields containing roots) *for field, k, and $f_1, \ldots, f_n \in k[X]$ with* $\deg f_i \geq 1$, *exists extension of k in which every f_i has root*

Definition 3.93 (algebraic closedness) *field, L, for which every polynomial, $f \in L[X]$ with $\deg f \geq 1$ has root in L, said to be* algebraically closed

Theorem 3.17 (existence of algebraically closed field extensions) *for field, k, exists algebraically closed extension*

Corollary 3.11 (existence of algebraically closed algebraic field extensions) *for field, k, exists algebraically closed algebraic extension of k (proof can be found in Proof 7)*

Isomorphism between algebraically closed algebraic extensions

Proposition 3.29 (number of algebraic embedding extensions) *for field, k, α being algebraic over k, algebraically closed field, L, and embedding, $\sigma : k \to L$, # possible embedding extensions of σ to $k(\alpha)$ in L is equal to # distinct roots of $\mathrm{Irr}(\alpha, k, X)$, hence no greater than # roots of $\mathrm{Irr}(\alpha, k, X)$*

Theorem 3.18 (algebraic embedding extensions) *for field, k, its algebraic extensions, E, algebraically closed field, L, and embedding, $\sigma : k \to L$, exists embedding extension of σ to E in L; if E is algebraically closed and L is algebraic over k^σ, every such embedding extension is isomorphism of E onto L*

Corollary 3.12 (isomorphism between algebraically closed algebraic extensions) *for field, k, and its algebraically closed algebraic extensions, E and E', exists isomorphism bewteen E and E' which induces identity on k, i.e.*

$$\tau : E \to E'$$

where $\tau|k$ is identity

- thus, *algebraically closed algebraic extension is determined up to isomorphism*

Algebraic closure

Definition 3.94 (algebraic closure) *for field, k, algebraically closed algebraic extension of k, which is determined up to isomorphism, called* algebraic closure of k*, frequently denoted by* k^{a}

- examples
  - complex conjugation is automorphism of $\mathbf{C}$ (which is the only continuous automorphism of $\mathbf{C}$)
  - subfield of $\mathbf{C}$ consisting of all numbers which are algebraic over $\mathbf{Q}$ is algebraic closure of $\mathbf{Q}$, *i.e.*, $\mathbf{Q}^{\mathrm{a}}$
  - $\mathbf{Q}^{\mathrm{a}} \neq \mathbf{C}$
  - $\mathbf{R}^{\mathrm{a}} = \mathbf{C}$
  - $\mathbf{Q}^{\mathrm{a}}$ *is countable*

Theorem 3.19 (countability of algebraic closure of finite fields) *algebraic closure of finite field is countable*

Theorem 3.20 (cardinality of algebraic extensions of infinite fields) *for infinite field, k, every algebraic extension of k has same cardinality as k*

Splitting fields

Definition 3.95 (splitting fields) *for field, k, and $f \in k[X]$ with $\deg f \geq 1$, field extension, K, of k, f splits into linear factors in which, i.e.,*

$$f(X) = c(X - \alpha_1) \cdots (X - \alpha_n)$$

and which is finitely generated over k by $\alpha_1, \ldots, \alpha_n$ (hence $K = k(\alpha_1, \ldots, \alpha_n)$), called splitting field of f

- for field, k, every $f \in k[X]$ has splitting field in k^{a}

Theorem 3.21 (isomorphism between splitting fields) *for field, k, $f \in k[X]$ with $\deg f \geq 1$, and two splitting fields of f, K and E, exists isomorphism between K and E; if $k \subset K \subset k^{\mathrm{a}}$, every embedding of E into k^{a} over k is isomorphism of E onto K*

Splitting fields for family of polynomials

Definition 3.96 (splitting fields for family of polynomials) *for field, k, index set, Λ, and indexed family of polynomials, $\{f_\lambda \in k[X] | \lambda \in \Lambda, \deg f_\lambda \geq 1\}$, extension field of k, every f_λ splits into linear factors in which and which is generated by all roots of all polynomials, f_λ, called* splitting field for family of polynomials

- in most applications, deal with finite Λ
- becoming increasingly important to consider infinite algebraic extensions
- various proofs would not be simpler if restricted ourselves to finite cases

Corollary 3.13 (isomorphism between splitting fields for family of polynomials) *for field, k, index set, Λ, and two splitting fields, K and E, for family of polynomials, $\{f_\lambda \in k[X] | \lambda \in \Lambda, \deg f_\lambda \geq 1\}$, every embedding of E into K^{a} over k is isomorphism of E onto K*

Normal extensions

Theorem 3.22 (normal extensions) *for field, k, and its algebraic extension, K, with $k \subset K \subset k^{\mathrm{a}}$, following statements are equivalent*

- *every embedding of K into k^{a} over k induces automorphism*
- *K is splitting field of family of polynomials in $k[X]$*
- *every irreducible polynomial of $k[X]$ which has root in K splits into linear factors in K*

Definition 3.97 (normal extensions) *for field, k, and its algebraic extension, K, with $k \subset K \subset k^{\mathrm{a}}$, satisfying properties in Theorem 3.22, said to be* normal

- *not* true that class of normal extensions is *distinguished*
  - *e.g.*, below tower of fields is tower of normal extensions

$$\mathbf{Q} \subset \mathbf{Q}(\sqrt{2}) \subset \mathbf{Q}(\sqrt[4]{2})$$

  - but, extension $\mathbf{Q} \subset \mathbf{Q}(\sqrt[4]{2})$ is not normal because complex roots of $X^4 - 2$ are not in $\mathbf{Q}(\sqrt[4]{2})$

Retention of normality of extensions

Theorem 3.23 (retention of normality of extensions) *normal extensions remain normal under lifting; if $k \subset E \subset K$ and K is normal over k, K is normal over E; if K_1 and K_2 are normal over k and are contained in common field, $K_1 K_2$ is normal over k, and so is $K_1 \cap K_2$*

Separable degree of field extensions

- for field, F, and its algebraic extension, E
  - let L be algebraically closed field and assume embedding, $\sigma : F \to L$
    - exists embedding extension of σ to E in L by Theorem 3.18
    - such σ maps E on subfield of L which is algebraic over F^σ
    - hence, E^σ is contained in algebraic closure of F^σ which is contained in L
    - will *assume* that L is the algebraic closure of F^σ
  - let L' be another algebraically closed field and assume another embedding, $\tau : F \to L'$ - assume as before that L' is algebraic closure of F^τ
  - then Theorem 3.18 implies, exists isomorphism, $\lambda : L \to L'$ extending $\tau \circ \sigma^{-1}$ applied to F^σ
  - let S_σ & S_τ be sets of embedding extensions of σ to E in L and L' respectively
  - then λ induces map from S_σ into S_τ with $\tilde{\sigma} \mapsto \lambda \circ \tilde{\sigma}$ and λ^{-1} induces inverse map from S_τ into S_σ, hence exists bijection between S_σ and S_τ, hence have same cardinality

Definition 3.98 (separable degree of field extensions) *above cardinality only depends on extension E/F, called separable degree of E over F, denoted by $[E : F]_s$*

Multiplicativity of and upper bound on separable degree of field extensions

Theorem 3.24 (multiplicativity of separable degree of field extensions) *for tower of algebraic field extensions, $k \subset F \subset E$,*

$$[E : k]_s = [E : F]_s [F : k]_s$$

Theorem 3.25 (upper limit on separable degree of field extensions) *for finite algebraic field extension, $k \subset E$*

$$[E : k]_s \leq [E : k]$$

- *i.e.*, separable degree is at most equal to degree (*i.e.*, dimension) of field extension

Corollary 3.14 *for tower of algebraic field extensions, $k \subset F \subset E$, with $[E : k] < \infty$*

$$[E : k]_s = [E : k]$$

holds if and only if *corresponding equality holds in every step of tower, i.e., for E/F and F/k*

Finite separable field extensions

Definition 3.99 (finite separable field extensions) *for finite algebraic field extension, E/k, with $[E : k]_s = [E : k]$, E, said to be separable over k*

Definition 3.100 (separable algebraic elements) *for field, k, α, which is algebraic over k with $k(\alpha)$ being separable over k, said to be separable over k*

Proposition 3.30 (separability and multiple roots) *for field, k, α, which is algebraic over k, is separable over k* if and only if $\mathrm{Irr}(\alpha, k, X)$ *has no multiple roots*

Definition 3.101 (separable polynomials) *for field, k, $f \in k[X]$ with no multiple roots, said to be separable*

Lemma 3.12 *for tower of algebraic field extensions, $k \subset F \subset K$, if $\alpha \in K$ is separable over k, then α is separable over F*

Theorem 3.26 (finite separable field extensions) *for finite field extension, E/k, E is separable over k* if and only if *every element of E is separable over k*

Arbitrary separable field extensions

Definition 3.102 (arbitrary separable field extensions) *for (not necessarily finite) field extension, E/k, E, of which every finitely generated subextension is separable over k, i.e.,*

$$(\forall n \in \mathbf{N} \ \& \ \alpha_1, \ldots, \alpha_n \in E) \, (k(\alpha_1, \ldots, \alpha_n) \textit{ is separable over } k)$$

said to be separable over k

Theorem 3.27 (separable field extensions) *for algebraic extension, E/k, E, which is generated by family of elements, $\{\alpha_\lambda\}_{\lambda \in \Lambda}$, with every α_λ is separable over k, is separable over k*

Theorem 3.28 (separable extensions are distinguished) *separable extensions form distinguished class of extensions*

Separable closure and conjugates

Definition 3.103 (separable closure) *for field, k, compositum of all separable extensions of k in given algebraic closure k^{a}, called* separable closure of k, *denoted by* k^{s} or k^{sep}

Definition 3.104 (conjugates of fields) *for algebraic field extension, E/k, and embedding of E, σ, in k^{a} over k, E^σ, called* conjugate of E in k^{a}

- smallest normal extension of k containing E is compositum of all conjugates of E in E^{a}

Definition 3.105 (conjugates of elements of fields) *for field, k, α being algebraic over k, and distinct embeddings, σ_1, ..., σ_r of $k(\alpha)$ into k^{a} over k, α^{σ_1}, ..., α^{σ_r}, called* conjugates of α in k^{a}

- α^{σ_1}, ..., α^{σ_r} are simply distinct roots of $\mathrm{Irr}(\alpha, k, X)$
- smallest normal extension of k containing one of these conjugates is simply $k(\alpha^{\sigma_1}, \ldots, \alpha^{\sigma_r})$

Prime element theorem

Theorem 3.29 (prime element theorem) *for finite algebraic field extension, E/k, exists $\alpha \in E$ such that $E = k(\alpha)$* if and only if *exists only finite # fields, F, such that $k \subset F \subset E$; if E is separable over k, exists such element, α*

Definition 3.106 (primitive element of fields) *for finite algebraic field extension, E/k, $\alpha \in E$ with $E = k(\alpha)$, called* primitive element of E over k

Finite fields

Definition 3.107 (finite fields) *for every prime number, p, and integer, $n \geq 1$, exists finite field of order p^n, denoted by* $\mathbf{F}_{p^n}$, *uniquely determined as subfield of algebraic closure, ${\mathbf{F}_p}^{\mathrm{a}}$, which is splitting field of polynomial*

$$f_{p,n}(X) = X^{p^n} - X$$

and whose elements are roots of $f_{p,n}$

Theorem 3.30 (finite fields) *for every finite field, F, exist prime number, p, and integer, $n \geq 1$, such that $F = \mathbf{F}_{p^n}$*

Corollary 3.15 (finite field extensions) *for finite field, $\mathbf{F}_{p^n}$, and integer, $m \geq 1$, exists one and only one extension of degree, m, which is $\mathbf{F}_{p^{mn}}$*

Theorem 3.31 (multiplicative group of finite field) *multiplicative group of finite field is cyclic*

Automorphisms of finite fields

Definition 3.108 (Frobenius mapping) *mapping*

$$\varphi_{p,n} : \mathbf{F}_{p^n} \to \mathbf{F}_{p^n}$$

defined by $x \mapsto x^p$*, called* Frobenius mapping

- $\varphi_{p,n}$ is (ring) homeomorphism with $\operatorname{Ker}\varphi_{p,n} = \{0\}$ since $\mathbf{F}_{p^n}$ is field, thus is injective (Proposition 3.17), and surjective because $\mathbf{F}_{p^n}$ is finite,
- thus, $\varphi_{p,n}$ is *isomorphism leaving* $\mathbf{F}_p$ *fixed*

Theorem 3.32 (group of automorphisms of finite fields) *group of automorphisms of* $\mathbf{F}_{p^n}$ *is cyclic of degree* n*, generated by* $\varphi_{p,n}$

Theorem 3.33 (group of automorphisms of finite fields over another finite field) *for prime number,* p*, and integers,* $m, n \geq 1$*, in any* $\mathbf{F}_p{}^{\mathrm{a}}$*,* $\mathbf{F}_{p^n}$ *is contained in* $\mathbf{F}_{p^m}$ if and only if n *divides* m*, i.e., exists* $d \in \mathbf{Z}$ *such that* $m = dn$*, in which case,* $\mathbf{F}_{p^m}$ *is* normal and separable *over* $\mathbf{F}_{p^n}$ *group of automorphisms of* $\mathbf{F}_{p^m}$ *over* $\mathbf{F}_{p^n}$ *is cyclic of order,* d*, generated by* $\varphi_{p,m}^n$

3.6 Galois Theory

What we will do to appreciate Galois theory

- study
  - group of automorphisms of finite (and infinite) Galois extension (at length)
  - give examples, *e.g.*, cyclotomic extensions, abelian extensions, (even) non-abelian ones
  - leading into study of matrix representation of Galois group & classifications
- have tools to prove
  - fundamental theorem of algebra
  - insolvability of quintic polynomials
- mention unsolved problems
  - given finite group, exists Galois extension of $\mathbf{Q}$ having this group as Galois group?

Fixed fields

Definition 3.109 (fixed fields) *for field,* K*, and group of automorphisms,* G*, of* K*,*

$$\{x \in K | \forall \sigma \in G, x^\sigma = x\} \subset K$$

is subfield of K*, and called* fixed field of G*, denoted by* K^G

- K^G is subfield of K because for every $x, y \in K^G$
  - $0^\sigma = 0 \Rightarrow 0 \in K^G$
  - $(x+y)^\sigma = x^\sigma + y^\sigma = x + y \Rightarrow x + y \in K^G$
  - $(-x)^\sigma = -x^\sigma = -x \Rightarrow -x \in K^G$
  - $1^\sigma = 1 \Rightarrow 1 \in K^G$
  - $(xy)^\sigma = x^\sigma y^\sigma = xy \Rightarrow xy \in K^G$
  - $(x^{-1})^\sigma = (x^\sigma)^{-1} = x^{-1} \Rightarrow x^{-1} \in K^G$

  hence, K^G closed under addition & multiplication, and is commutative division ring, thus field
- $0, 1 \in K^G$, hence K^G *contains prime field*

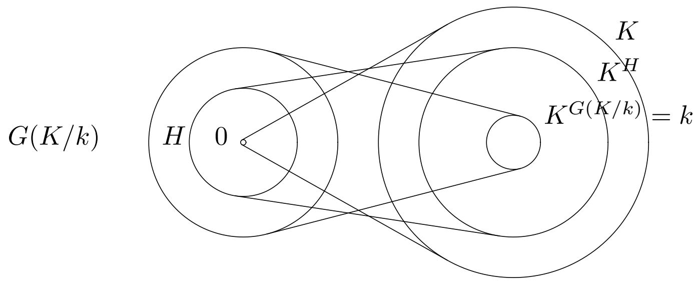

Figure 3.11: diagrams for Galois main result

Galois extensions and Galois groups

Definition 3.110 (Galois extensions) *algebraic extension, K, of field, k, which is normal and separable, said to be* Galois (extension of k) *or* Galois over k *considering K as embedded in k^{a}; for convenience, sometimes say* K/k is Galois

Definition 3.111 (Galois groups) *for field, k and its Galois extension, K, group of automorphisms of K over k, called* Galois group of K over k, *denoted by* $G(K/k)$, $G_{K/k}$, $\mathrm{Gal}(K/k)$, or (simply) G

Definition 3.112 (Galois groups of polynomials) *for field, k, separable $f \in k[X]$ with $\deg f \geq 1$, and its splitting field, K/k, Galois group of K over k (i.e., $G(K/k)$), called* Galois group of f over k

Proposition 3.31 (Galois groups of polynomials and symmetric groups) *for field, k, separable $f \in k[X]$ with $\deg f \geq 1$, and its splitting field, K/k,*

$$f(X) = (X - \alpha_1) \cdots (X - \alpha_n)$$

elements of Galois group of f over k, G, permute roots of f, hence, exists injective homeomorphism of G into S_n, i.e., symmetric group on n elements

Fundamental theorem for Galois theory

Theorem 3.34 (fundamental theorem for Galois theory) *for finite Galois extension, K/k*

- *map $H \mapsto K^H$ induces isomorphism between set of subgroups of $G(K/k)$ & set of intermediate fields*
- *subgroup, H, of $G(K/k)$, is normal* if and only if *K^H/k is Galois*
- *for normal subgroup, H, $\sigma \mapsto \sigma|K^H$ induces isomorphism between $G(K/k)/H$ and $G(K^H/k)$*

(illustrated in Figure 3.11)

- shall prove step by step

Galois subgroups association with intermediate fields

Theorem 3.35 (Galois subgroups associated with intermediate fields - 1) *for Galois extension, K/k, and intermediate field, F*

- *K/F is Galois & $K^{G(K/F)} = F$, hence, $K^G = k$*

- *map*

$$F \mapsto G(K/F)$$

induces injective *homeomorphism from set of intermediate fields to subgroups of* G

(proof can be found in Proof 8)

Definition 3.113 (Galois subgroups associated with intermediate fields) *for Galois extension,* K/k*, and intermediate field,* F*, subgroup,* $G(K/F) \subset G(K/k)$*, called* group associated with F*, said to* belong to F

Corollary 3.16 (Galois subgroups associated with intermediate fields - 1) *for Galois extension,* K/k*, and two intermediate fields,* F_1 *and* F_2*,* $G(K/F_1) \cap G(K/F_2)$ *belongs to* F_1F_2*, i.e.,*

$$G(K/F_1) \cap G(K/F_2) = G(K/F_1F_2)$$

(proof can be found in Proof 9)

Corollary 3.17 (Galois subgroups associated with intermediate fields - 2) *for Galois extension,* K/k*, and two intermediate fields,* F_1 *and* F_2*, smallest subgroup of* G *containing* $G(K/F_1)$ *and* $G(K/F_2)$ *belongs to* $F_1 \cap F_2$*, i.e.*

$$\bigcap_{G(K/F_1)\subset H, G(K/F_2)\subset H} \{H | H \subset G(K/k)\} = G(K/(F_1 \cap F_2))$$

Corollary 3.18 (Galois subgroups associated with intermediate fields - 3) *for Galois extension,* K/k*, and two intermediate fields,* F_1 *and* F_2*,*

$$F_1 \subset F_2 \text{ if and only if } G(K/F_2) \subset G(K/F_1)$$

(proof can be found in Proof 10)

Corollary 3.19 *for finite separable field extension,* E/k*, the smallest normal extension of* k *containing* E*,* K*,* K/k *is finite Galois and exist only finite number of intermediate fields*

Lemma 3.13 *for algebraic separable extension,* E/k*, if every element of* E *has degree no greater than* n *over* k *for some* $n \geq 1$*,* E *is finite over* k *and* $[E : k] \leq n$

Theorem 3.36 (Artin's theorem) *(Artin) for field,* K*, finite* $\mathrm{Aut}(K)$ *of order,* n*, and* $k = K^{\mathrm{Aut}(K)}$*,* K/k *is Galois,* $G(K/k) = \mathrm{Aut}(K)$*, and* $[K : k] = n$

Corollary 3.20 (Galois subgroups associated with intermediate fields - 4) *for finite Galois extension,* K/k*, every subgroup of* $G(K/k)$ *belongs to intermediate field*

Theorem 3.37 (Galois subgroups associated with intermediate fields - 2) *for Galois extension,* K/k*, and intermediate field,* F*,*

- F/k *is normal extension* if and only if $G(K/F)$ *is normal subgroup of* $G(K/k)$
- *if* F/k *is normal extension, map,* $\sigma \mapsto \sigma|F$*, induces homeomorphism of* $G(K/k)$ *onto* $G(F/k)$ *of which* $G(K/F)$ *is kernel, thus*

$$G(F/k) \approx G(K/k)/G(K/F)$$

Proof for fundamental theorem for Galois theory

- finally, we prove *fundamental theorem for Galois theory* (Theorem 3.34)
- assume K/k is finite Galois extension and H is subgroup of $G(K/k)$
  - Corollary 3.20 implies K^H is intermediate field, hence Theorem 3.35 implies K/K^H is Galois, Theorem 3.36 implies $G(K/K^H) = H$, thus, every H is Galois
  - map, $H \mapsto K^H$, induces homeomorphism, σ, of set of all subgroups of $G(K/k)$ into set of intermediate fields
  - σ is *injective* since for any two subgroups, H and H', of $G(K/k)$, if $K^H = K^{H'}$, then $H = G(K/K^H) = G(K/K^{H'}) = H'$
  - σ is *surjective* since for every intermediate field, F, Theorem 3.35 implies K/F is Galois, $G(K/F)$ is subgroup of $G(K/k)$, and $K^{G(K/F)} = F$, thus, $\sigma(G(K/F)) = K^{G(K/F)} = F$
  - *therefore, σ is isomorphism between set of all subgroups of $G(K/k)$ and set of intermediate fields*
  - since Theorem 3.28 implies separable extensions are distinguished, H^K/k is separable, thus Theorem 3.37 implies that K^H/k is Galois *if and only if* $G(K/K^H)$ is normal
  - lastly, Theorem 3.37 implies that if K^H/k is Galois, $G(H^K/k) \approx G(K/k)/H$

Abelian and cyclic Galois extensions and groups

Definition 3.114 (abelian Galois extensions) *Galois extension with abelian Galois group, said to be* abelian

Definition 3.115 (cyclic Galois extensions) *Galois extension with cyclic Galois group, said to be* cyclic

Corollary 3.21 *for Galois extension, K/k, and intermediate field, F,*

- *if K/k is abelian, F/k is Galois and abelian*
- *if K/k is cyclic, F/k is Galois and cyclic*

Definition 3.116 (maximum abelian extension) *for field, k, compositum of all abelian Galois extensions of k in given k^{a}, called* maximum abelian extension of k, *denoted by* k^{ab}

Theorems and corollaries about Galois extensions

Theorem 3.38 *for Galois extension, K/k, and arbitrary extension, F/k, where K and F are subfields of common field,*

- *KF/F and $K/(K \cap F)$ are Galois extensions*
- *map*

$$\sigma \mapsto \sigma|K$$

  induces isomorphism between $G(KF/F)$ and $G(K/(K \cap F))$

theorem illustrated in Figure 3.12

Corollary 3.22 *for finite Galois extension, K/k, and arbitrary extension, F/k, where K and F are subfields of common field,*

$$[KF : F] \text{ divides } [F : k]$$

Theorem 3.39 *for Galois extensions, K_1/k and K_2/k, where K_1 and K_2 are subfields of common field,*

$$
\begin{tikzcd}
& KF \arrow[dl, dash] \arrow[dr, dash, "\text{Galois}"] & \\
K \arrow[dr, dash, "\text{Galois}"'] & & F \arrow[dl, dash] \\
& K \cap F \arrow[d, dash] & \\
& k &
\end{tikzcd}
$$

Figure 3.12: diagram for Galois lifting

$$
\begin{tikzcd}
& K_1K_2 \arrow[dl, dash] \arrow[dr, dash] & \\
K_1 \arrow[dr, dash] & & K_2 \arrow[dl, dash] \\
& K_1 \cap K_2 \arrow[d, dash] & \\
& k &
\end{tikzcd}
$$

Figure 3.13: diagram for Galois two-side lifting

- *K_1K_2/k is Galois extension*
- *map*

$$\sigma \mapsto (\sigma|K_1, \sigma|K_2)$$

  of $G(K_1K_2/k)$ into $G(K_1/k) \times G(K_2/k)$ is injective; if $K_1 \cap K_2 = k$, map is isomorphism

theorem illustrated in Figure 3.13

Corollary 3.23 *for n Galois extensions, K_i/k, where K_1, …, K_n are subfields of common field and $K_{i+1} \cap (K_1 \cdots K_i) = k$ for $i = 1, \ldots, n-1$,*

- *$K_1 \cdots K_n/k$ is Galois extension*
- *map*

$$\sigma \mapsto (\sigma|K_1, \ldots, \sigma|K_n)$$

  induces isomorphism of $G(K_1 \cdots K_n/k)$ onto $G(K_1/k) \times \cdots \times G(K_n/k)$

Corollary 3.24 *for Galois extension, K/k, where $G(K/k)$ can be written as $G_1 \times \cdots \times G_n$, and K_1, …, K_n, each of which is fixed field of*

$$G_1 \times \cdots \times \underbrace{\{e\}}_{i\text{th position}} \times \cdots \times G_n$$

- *K_1/k, …, K_n/k are Galois extensions*
- *$G(K_i/k) = G_i$ for $i = 1, \ldots, n$*

- $K_{i+1} \cap (K_1 \cdots K_i) = k$ *for* $i = 1, \ldots, n-1$
- $K = K_1 \cdots K_n$

Theorem 3.40 *assume all fields are subfields of common field*

- *for two abelian Galois extensions,* K/k *and* L/k*,* KL/k *is abelian Galois extension*
- *for abelian Galois extension,* K/k*, and any extension,* E/k*,* KE/E *is abelian Galois extension*
- *for abelian Galois extension,* K/k*, and intermediate field,* E*, both* K/E *and* E/k *are abelian Galois extensions*

Solvable and radical extensions

Definition 3.117 (sovable extensions) *finite separable extension,* E/k*, such that Galois group of smallest Galois extension,* K/k*, containing* E *is solvable, said to be* solvable

Theorem 3.41 (solvable extensions are distinguished) *solvable extensions form distinguished class of extensions*

Definition 3.118 (solvable by radicals) *finite extension,* F/k*, such that it is separable and exists finite extension,* E/k*, containing* F *admitting tower decomposition*

$$k = E_0 \subset E_1 \subset \cdots \subset E_m = E$$

with E_{i+1}/E_i *is obtained by adjoining root of*

- *unity, or*
- $X^n = a$ *with* $a \in E_i$*, and* n *prime to characteristic, or*
- $X_p - X - a$ *with* $a \in E_i$ *if* p *is positive characteristic*

said to be solvable by radicals

Theorem 3.42 (extensions solvable by radicals) *separable extension,* E/k*, is solvable by radicals* if and only if *it is solvable*

Applications of Galois theory

Theorem 3.43 (insolvability of quintic polynomials) *general equation of degree,* n*, cannot be solved by radicals for* $n \geq 5$ *(implied by Definition 3.112, Proposition 3.31, Theorem 3.42, and Theorem 3.5)*

Theorem 3.44 (fundamental theorem of algebra) $f \in \mathbf{C}[X]$ *of degree,* n*, has precisely* n *roots in* $\mathbf{C}$ *(when counted with multiplicity), i.e.,* $\mathbf{C}$ *is algebraically closed*

4 Real Analysis

4.1 Set Theory

Some principles

Principle 4.1 (principle of mathematical induction)

$$P(1)\&[P(n \Rightarrow P(n+1)] \Rightarrow (\forall n \in \mathbf{N})P(n)$$

Principle 4.2 (well ordering principle) *each nonempty subset of* $\mathbf{N}$ *has a smallest element*

Principle 4.3 (principle of recursive definition) *for* $f : X \to X$ *and* $a \in X$*, exists unique infinite sequence* $\langle x_n \rangle_{n=1}^{\infty} \subset X$ *such that*

$$x_1 = a$$

and

$$(\forall n \in \mathbf{N})\,(x_{n+1} = f(x_n))$$

- note that Principle 4.1 $\Leftrightarrow$ Principle 4.2 $\Rightarrow$ Principle 4.3

Some definitions for functions

Definition 4.1 (functions) *for* $f : X \to Y$

- *terms,* map *and* function*, exterchangeably used*
- X *and* Y*, called* domain of f *and* codomain of f *respectively*
- $\{f(x)|x \in X\}$*, called* range of f
- *for* $Z \subset Y$*,* $f^{-1}(Z) = \{x \in X | f(x) \in Z\} \subset X$*, called* preimage *or* inverse image of Z under f
- *for* $y \in Y$*,* $f^{-1}(\{y\})$*, called* fiber of f over y
- f*, called* injective *or* injection *or* one-to-one *if* $(\forall x \neq v \in X)\,(f(x) \neq f(v))$
- f*, called* surjective *or* surjection *or* onto *if* $(\forall x \in X)\,(\exists y in Y)\,(y = f(x))$
- f*, called* bijective *or* bijection *if* f *is both injective and surjective, in which case,* X *and* Y*, said to be* one-to-one correspondece *or* bijective correspondece
- $g : Y \to X$*, called* left inverse *if* $g \circ f$ *is identity function*
- $h : Y \to X$*, called* right inverse *if* $f \circ h$ *is identity function*

Some properties of functions

Lemma 4.1 (functions) *for* $f : X \to Y$

- f *is injective* if and only if f *has left inverse*
- f *is surjective* if and only if f *has right inverse*
- *hence,* f *is bijective* if and only if f *has both left and right inverse because if* g *and* h *are left and right inverses respectively,* $g = g \circ (f \circ h) = (g \circ f) \circ h = h$
- *if* $|X| = |Y| < \infty$*,* f *is injective* if and only if f *is surjective* if and only if f *is bijective*

Countability of sets

- set A is countable if range of some function whose domain is $\mathbf{N}$
- $\mathbf{N}$, $\mathbf{Z}$, $\mathbf{Q}$: countable
- $\mathbf{R}$: *not* countable

Limit sets

- for sequence, $\langle A_n \rangle$, of subsets of X
  - *limit superior or limsup of $\langle A_n \rangle$*, defined by

$$\limsup \langle A_n \rangle = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_m$$

  - *limit inferior or liminf of $\langle A_n \rangle$*, defined by

$$\liminf \langle A_n \rangle = \bigcup_{n=1}^{\infty} \bigcap_{m=n}^{\infty} A_m$$

- always

$$\liminf \langle A_n \rangle \subset \limsup \langle A_n \rangle$$

- when $\liminf \langle A_n \rangle = \limsup \langle A_n \rangle$, sequence, $\langle A_n \rangle$, said to *converge to it*, denote

$$\lim \langle A_n \rangle = \liminf \langle A_n \rangle = \limsup \langle A_n \rangle = A$$

Algebras of sets

- collection $\mathscr{A}$ of subsets of X called *algebra* or *Boolean algebra* if

$$(\forall A, B \in \mathscr{A})(A \cup B \in \mathscr{A}) \text{ and } (\forall A \in \mathscr{A})(\tilde{A} \in \mathscr{A})$$

  - $(\forall A_1, \ldots, A_n \in \mathscr{A})(\cup_{i=1}^{n} A_i \in \mathscr{A})$
  - $(\forall A_1, \ldots, A_n \in \mathscr{A})(\cap_{i=1}^{n} A_i \in \mathscr{A})$
- algebra $\mathscr{A}$ called *σ-algebra* or *Borel field* if
  - every union of a countable collection of sets in $\mathscr{A}$ is in $\mathscr{A}$, *i.e.*,

$$(\forall \langle A_i \rangle)(\cup_{i=1}^{\infty} A_i \in \mathscr{A})$$

- given sequence of sets in algebra $\mathscr{A}$, $\langle A_i \rangle$, exists disjoint sequence, $\langle B_i \rangle$ such that

$$B_i \subset A_i \text{ and } \bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$$

Algebras generated by subsets

- *algebra generated by* collection of subsets of X, $\mathcal{C}$, can be found by

$$\mathscr{A} = \bigcap\{\mathscr{B}|\mathscr{B} \in \mathcal{F}\}$$

  where $\mathcal{F}$ is family of all algebras containing $\mathcal{C}$
  - smallest algebra $\mathscr{A}$ containing $\mathcal{C}$, *i.e.*,

$$(\forall \mathscr{B} \in \mathcal{F})(\mathscr{A} \subset \mathscr{B})$$

- *σ-algebra generated by* collection of subsets of X, $\mathcal{C}$, can be found by

$$\mathscr{A} = \bigcap\{\mathscr{B}|\mathscr{B} \in \mathcal{G}\}$$

  where $\mathcal{G}$ is family of all σ-algebras containing $\mathcal{C}$
  - smallest σ-algebra $\mathscr{A}$ containing $\mathcal{C}$, *i.e.*,

$$(\forall \mathscr{B} \in \mathcal{G})(\mathscr{A} \subset \mathscr{B})$$

Relation

- x said to *stand in relation* $\mathbf{R}$ to y, denoted by $x \mathbf{R} y$
- $\mathbf{R}$ said to *be relation on* X if $x \mathbf{R} y \Rightarrow x \in X$ and $y \in X$
- $\mathbf{R}$ is
  - transitive if $x \mathbf{R} y$ and $y \mathbf{R} z \Rightarrow x \mathbf{R} z$
  - symmetric if $x \mathbf{R} y = y \mathbf{R} x$
  - reflexive if $x \mathbf{R} x$
  - antisymmetric if $x \mathbf{R} y$ and $y \mathbf{R} x \Rightarrow x = y$
- $\mathbf{R}$ is
  - *equivalence relation* if transitive, symmetric, and reflexive, *e.g.*, modulo
  - *partial ordering* if transitive and antisymmetric, *e.g.*, "$\subset$"
  - *linear (or simple) ordering* if transitive, antisymmetric, and $x \mathbf{R} y$ or $y \mathbf{R} x$ for all $x, y \in X$
    - *e.g.*, "$\geq$" linearly orders $\mathbf{R}$ while "$\subset$" does not $\mathcal{P}(X)$

Ordering

- given partial order, $\prec$, a is
  - a first/smallest/least element if $x \neq a \Rightarrow a \prec x$
  - a last/largest/greatest element if $x \neq a \Rightarrow x \prec a$
  - a minimal element if $x \neq a \Rightarrow x \not\prec a$
  - a maximal element if $x \neq a \Rightarrow a \not\prec x$
- partial ordering $\prec$ is
  - strict partial ordering if $x \not\prec x$
  - reflexive partial ordering if $x \prec x$
- strict linear ordering $<$ is
  - *well ordering* for X if every nonempty set contains a first element

Axiom of choice and equivalent principles

Axiom 4.1 (axiom of choice) *given a collection of nonempty sets, $\mathcal{C}$, there exists $f : \mathcal{C} \to \cup_{A \in \mathcal{C}} A$ such that*

$$(\forall A \in \mathcal{C})\,(f(A) \in A)$$

- also called *multiplicative axiom* - preferred to be called to axiom of choice by Bertrand Russell for reason writte in §4.1
- no problem when $\mathcal{C}$ is finite
- need axiom of choice when $\mathcal{C}$ is not finite

Principle 4.4 (Hausdorff maximal principle) *for particial ordering $\prec$ on X, exists a maximal linearly ordered subset $S \subset X$, i.e., S is linearity ordered by $\prec$ and if $S \subset T \subset X$ and T is linearly ordered by $\prec$, $S = T$*

Principle 4.5 (well-ordering principle) *every set X can be well ordered, i.e., there is a relation $<$ that well orders X*

- note that Axiom 4.1 $\Leftrightarrow$ Principle 4.4 $\Leftrightarrow$ Principle 4.5

Infinite direct product

Definition 4.2 (direct product) *for collection of sets, $\langle X_\lambda \rangle$, with index set, Λ,*

$$\bigtimes_{\lambda \in \Lambda} X_\lambda$$

called direct product

- *for $z = \langle x_\lambda \rangle \in \bigtimes X_\lambda$, x_λ called* λ-th coordinate of z
- if one of X_λ is empty, $\bigtimes X_\lambda$ is empty
- *axiom of choice* is equivalent to converse, *i.e.*, if none of X_λ is empty, $\bigtimes X_\lambda$ is not empty
  if one of X_λ is empty, $\bigtimes X_\lambda$ is empty
- this is why Bertrand Russell prefers *multiplicative axiom* to *axiom of choice* for name of axiom (Axiom 4.1)

4.2 Real Number System

Field axioms

- field axioms - for every $x, y, z \in \mathbf{F}$
  - $(x + y) + z = x + (y + z)$ - additive associativity
  - $(\exists 0 \in \mathbf{F})(\forall x \in \mathbf{F})(x + 0 = x)$ - additive identity
  - $(\forall x \in \mathbf{F})(\exists w \in \mathbf{F})(x + w = 0)$ - additive inverse
  - $x + y = y + x$ - additive commutativity
  - $(xy)z = x(yz)$ - multiplicative associativity
  - $(\exists 1 \neq 0 \in \mathbf{F})(\forall x \in \mathbf{F})(x \cdot 1 = x)$ - multiplicative identity
  - $(\forall x \neq 0 \in \mathbf{F})(\exists w \in \mathbf{F})(xw = 1)$ - multiplicative inverse
  - $x(y + z) = xy + xz$ - distributivity
  - $xy = yx$ - multiplicative commutativity
- system (set with $+$ and $\cdot$) satisfying axiom of field called *field*
  - *e.g.*, field of module p where p is prime, $\mathbf{F}_p$

Axioms of order

- axioms of order - subset, $\mathbf{F}_{++} \subset \mathbf{F}$, of positive (real) numbers satisfies
  - $x, y \in \mathbf{F}_{++} \Rightarrow x + y \in \mathbf{F}_{++}$
  - $x, y \in \mathbf{F}_{++} \Rightarrow xy \in \mathbf{F}_{++}$
  - $x \in \mathbf{F}_{++} \Rightarrow -x \not\in \mathbf{F}_{++}$
  - $x \in \mathbf{F} \Rightarrow x = 0 \vee x \in \mathbf{F}_{++} \vee -x \in \mathbf{F}_{++}$
- system satisfying field axioms & axioms of order called *ordered field*
  - *e.g.*, set of real numbers ($\mathbf{R}$), set of rational numbers ($\mathbf{Q}$)

Axiom of completeness

- completeness axiom
  - every nonempty set S of real numbers which has an upper bound has a least upper bound, *i.e.*,

$$\{l | (\forall x \in S)(l \leq x)\}$$

    has least element.
  - use $\inf S$ and $\sup S$ for least and greatest element (when exist)
- ordered field that is complete is *complete ordered field*
  - *e.g.*, $\mathbf{R}$ (with $+$ and $\cdot$)

$\Rightarrow$ axiom of Archimedes
  - given any $x \in \mathbf{R}$, there is an integer n such that $x < n$

$\Rightarrow$ corollary
  - given any $x < y \in \mathbf{R}$, exists $r \in \mathbf{Q}$ such tat $x < r < y$

Sequences of R

- sequence of $\mathbf{R}$ denoted by $\langle x_i \rangle_{i=1}^{\infty}$ or $\langle x_i \rangle$
  - mapping from $\mathbf{N}$ to $\mathbf{R}$
- limit of $\langle x_n \rangle$ denoted by $\lim_{n\to\infty} x_n$ or $\lim x_n$ - defined by $a \in \mathbf{R}$

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(n \geq N \Rightarrow |x_n - a| < \epsilon)$$

  - $\lim x_n$ unique if exists
- $\langle x_n \rangle$ called Cauchy sequence if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(n, m \geq N \Rightarrow |x_n - x_m| < \epsilon)$$

- Cauchy criterion - characterizing complete metric space (including $\mathbf{R}$)
  - sequence converges *if and only if* Cauchy sequence

Other limits

- cluster point of $\langle x_n \rangle$ - defined by $c \in \mathbf{R}$

$$(\forall \epsilon > 0, N \in \mathbf{N})(\exists n > N)(|x_n - c| < \epsilon)$$

- limit superior or limsup of $\langle x_n \rangle$

$$\limsup x_n = \inf_n \sup_{k>n} x_k$$

- limit inferior or liminf of $\langle x_n \rangle$

$$\liminf x_n = \sup_n \inf_{k>n} x_k$$

- $\liminf x_n \leq \limsup x_n$
- $\langle x_n \rangle$ converges *if and only if* $\liminf x_n = \limsup x_n$ $(=\lim x_n)$

Open and closed sets

- O called *open* if

$$(\forall x \in O)(\exists \delta > 0)(y \in \mathbf{R})(|y - x| < \delta \Rightarrow y \in O)$$

  - intersection of finite collection of open sets is open
  - union of any collection of open sets is open
- $\overline{E}$ called *closure* of E if

$$(\forall x \in \overline{E} \;\&\; \delta > 0)(\exists y \in E)(|x - y| < \delta)$$

- F called *closed* if

$$F = \overline{F}$$

  - union of finite collection of closed sets is closed
  - intersection of any collection of closed sets is closed

Open and closed sets - facts

- *every open set is union of countable collection of disjoint open intervals*
- (Lindelöf) any collection $\mathcal{C}$ of open sets has a countable subcollection $\langle O_i \rangle$ such that

$$\bigcup_{O \in \mathcal{C}} O = \bigcup_i O_i$$

  - equivalently, any collection $\mathcal{F}$ of closed sets has a countable subcollection $\langle F_i \rangle$ such that

$$\bigcap_{O \in \mathcal{F}} F = \bigcap_i F_i$$

Covering and Heine-Borel theorem

- collection $\mathcal{C}$ of sets called *covering* of A if

$$A \subset \bigcup_{O \in \mathcal{C}} O$$

  - $\mathcal{C}$ said to *cover* A
  - $\mathcal{C}$ called *open covering* if every $O \in \mathcal{C}$ is open
  - $\mathcal{C}$ called *finite covering* if $\mathcal{C}$ is finite

- *Heine-Borel theorem* - for any closed and bounded set, every open covering has finite subcovering
- corollary
  - any collection $\mathcal{C}$ of closed sets including at least one bounded set every finite subcollection of which has nonempty intersection has nonempty intersection.

Continuous functions

- f (with domain D) called *continuous at* x if

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall y \in D)(|y - x| < \delta \Rightarrow |f(y) - f(x)| < \epsilon)$$

- f called *continuous on* $A \subset D$ if f is continuous at every point in A
- f called *uniformly continuous on* $A \subset D$ if

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall x, y \in D)(|x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon)$$

Continuous functions - facts

- f is continuous *if and only if* for every open set O (in co-domain), $f^{-1}(O)$ is open
- f continuous on closed and bounded set is uniformly continuous
- *extreme value theorem* - f continuous on closed and bounded set, F, is *bounded on F and assumes its maximum and minimum on F*

$$(\exists x_1, x_2 \in F)(\forall x \in F)(f(x_1) \leq f(x) \leq f(x_2))$$

- *intermediate value theorem* - for f continuous on $[a, b]$ with $f(a) \leq f(b)$,

$$(\forall d)(f(a) \leq d \leq f(b))(\exists c \in [a, b])(f(c) = d)$$

Borel sets and Borel σ-algebra

- *Borel set*
  - any set that can be formed from open sets (or, equivalently, from closed sets) through the operations of countable union, countable intersection, and relative complement
- *Borel algebra* or *Borel σ-algebra*
  - *smallest σ-algebra containing all open sets*
  - also
    - smallest σ-algebra containing all closed sets
    - smallest σ-algebra containing all open intervals (due to statement on page 70)

Various Borel sets

- countable union of closed sets (in $\mathbf{R}$), called *an F_σ* (F for closed & σ for sum)
  - thus, every countable set, every closed set, every open interval, every open sets, is an F_σ (note $(a,b) = \bigcup_{n=1}^\infty [a+1/n, b-1/n]$)
  - countable union of sets in F_σ again is an F_σ
- countable intersection of open sets called *a G_δ* (G for open & δ for durchschnitt - average in German)
  - complement of F_σ is a G_δ and vice versa
- F_σ and G_δ are simple types of Borel sets
- countable intersection of F_σ's is $F_{\sigma\delta}$, countable union of $F_{\sigma\delta}$'s is $F_{\sigma\delta\sigma}$, countable intersection of $F_{\sigma\delta\sigma}$'s is $F_{\sigma\delta\sigma\delta}$, *etc.*, & likewise for $G_{\delta\sigma\ldots}$
- below are all classes of Borel sets, but not every Borel set belongs to one of these classes

$$F_\sigma, F_{\sigma\delta}, F_{\sigma\delta\sigma}, F_{\sigma\delta\sigma\delta}, \ldots, G_\delta, G_{\delta\sigma}, G_{\delta\sigma\delta}, G_{\delta\sigma\delta\sigma}, \ldots,$$

4.3 Lebesgue Measure

Riemann integral

- Riemann integral
  - partition induced by sequence $\langle x_i\rangle_{i=1}^n$ with $a = x_1 < \cdots < x_n = b$
  - lower and upper sums
    * $L(f, \langle x_i\rangle) = \sum_{i=1}^{n-1} \inf_{x\in[x_i,x_{i+1}]} f(x)(x_{i+1} - x_i)$
    * $U(f, \langle x_i\rangle) = \sum_{i=1}^{n-1} \sup_{x\in[x_i,x_{i+1}]} f(x)(x_{i+1} - x_i)$
  - always holds: $L(f, \langle x_i\rangle) \leq U(f, \langle y_i\rangle)$, hence

$$\sup_{\langle x_i\rangle} L(f, \langle x_i\rangle) \leq \inf_{\langle x_i\rangle} U(f, \langle x_i\rangle)$$

  - Riemann integrable if

$$\sup_{\langle x_i\rangle} L(f, \langle x_i\rangle) = \inf_{\langle x_i\rangle} U(f, \langle x_i\rangle)$$

- every continuous function is Riemann integrable

Motivation - want measure better than Riemann integrable

- consider indicator (or characteristic) function $\chi_{\mathbf{Q}} : [0,1] \to [0,1]$

$$\chi_{\mathbf{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbf{Q} \\ 0 & \text{if } x \not\in \mathbf{Q} \end{cases}$$

- *not* Riemann integrable: $\sup_{\langle x_i\rangle} L(f, \langle x_i\rangle) = 0 \neq 1 = \inf_{\langle x_i\rangle} U(f, \langle x_i\rangle)$
- however, irrational numbers infinitely more than rational numbers, hence
  - *want to* have some integral $\int$ such that, *e.g.*,

$$\int_{[0,1]} \chi_{\mathbf{Q}}(x)dx = 0 \text{ and } \int_{[0,1]} (1 - \chi_{\mathbf{Q}}(x))dx = 1$$

Properties of desirable measure

- want some measure $\mu : \mathcal{M} \to \mathbf{R}_+ = \{x \in \mathbf{R} | x \geq 0\}$
  - defined for every subset of $\mathbf{R}$, *i.e.*, $\mathcal{M} = \mathcal{P}(\mathbf{R})$
  - equals to length for open interval
  $$\mu[b, a] = b - a$$
  - countable additivity: for disjoint $\langle E_i \rangle_{i=1}^{\infty}$
  $$\mu(\cup E_i) = \sum \mu(E_i)$$
  - translation invariant
  $$\mu(E + x) = \mu(E) \text{ for } x \in \mathbf{R}$$
- *no* such measure exists
- *not* known whether measure with first three properties exists
- want to find translation invariant *countably additive measure*
  - hence, *give up on first property*

Race won by Henri Lebesgue in 1902!

- mathematicians in 19th century struggle to solve this problem
- race won by French mathematician, *Henri Léon Lebesgue in 1902!*
- Lebesgue integral covers much wider range of functions
  - indeed, $\chi_{\mathbf{Q}}$ is Lebesgue integrable
  $$\int_{[0,1]} \chi_{\mathbf{Q}}(x) dx = 0 \text{ and } \int_{[0,1]} (1 - \chi_{\mathbf{Q}}(x)) dx = 1$$

Outer measure

- for $E \subset \mathbf{R}$, define outer measure $\mu^* : \mathcal{P}(\mathbf{R}) \to \mathbf{R}_+$
$$\mu^* E = \inf_{\langle I_i \rangle} \left\{ \sum l(I_i) \middle| E \subset \cup I_i \right\}$$
where $I_i = (a_i, b_i)$ and $l(I_i) = b_i - a_i$
- outer measure of open interval is length
$$\mu^*(a_i, b_i) = b_i - a_i$$
- countable subadditivity
$$\mu^* \left(\cup E_i\right) \leq \sum \mu^* E_i$$
- corollaries
  - $\mu^* E = 0$ if E is countable
  - $[0, 1]$ not countable

Measurable sets

- $E \subset \mathbf{R}$ called measurable if for every $A \subset \mathbf{R}$

$$\mu^* A = \mu^*(E \cup A) + \mu^*(\tilde{E} \cup A)$$

- $\mu^* E = 0$, then E measurable
- every open interval (a, b) with $a \geq -\infty$ and $b \leq \infty$ is measurable
- disjoint countable union of measurable sets is measurable, *i.e.*, $\cup E_i$ is measurable
- collection of measurable sets is σ-algebra

Borel algebra is measurable

- note
  - every open set is disjoint countable union of open intervals (page 70)
  - disjoint countable union of measurable sets is measurable (page 74)
  - open intervals are measurable (page 74)
- hence, every open set is measurable
- also
  - collection of measurable sets is σ-algebra (page 74)
  - every open set is Borel set and Borel sets are σ-algebra (page 71)
- hence, *Borel sets are measurable*
- specifically, *Borel algebra (smallest σ-algebra containing all open sets) is measurable*

Lebesgue measure

- restriction of μ^* in collection $\mathcal{M}$ of measurable sets called *Lebesgue measure*

$$\mu : \mathcal{M} \to \mathbf{R}_+$$

- countable subadditivity - for $\langle E_n \rangle$

$$\mu(\cup E_n) \leq \sum \mu E_n$$

- *countable additivity* - for disjoint $\langle E_n \rangle$

$$\mu(\cup E_n) = \sum \mu E_n$$

- for dcreasing sequence of measurable sets, $\langle E_n \rangle$, *i.e.*, $(\forall n \in \mathbf{N})(E_{n+1} \subset E_n)$

$$\mu\left(\bigcap E_n\right) = \lim \mu E_n$$

(Lebesgue) measurable sets are nice ones!

- following statements are equivalent

  - E is measurable
  - $(\forall \epsilon > 0)(\exists \text{ open } O \supset E)(\mu^*(O \sim E) < \epsilon)$
  - $(\forall \epsilon > 0)(\exists \text{ closed } F \subset E)(\mu^*(E \sim F) < \epsilon)$
  - $(\exists G_\delta)(G_\delta \supset E)(\mu^*(G_\delta \sim E) < \epsilon)$
  - $(\exists F_\sigma)(F_\sigma \subset E)(\mu^*(E \sim F_\sigma) < \epsilon)$

- if $\mu^* E$ is finite, above statements are equivalent to

$$(\forall \epsilon > 0)\left(\exists U = \bigcup_{i=1}^{n}(a_i, b_i)\right)(\mu^*(U \Delta E) < \epsilon)$$

Lebesgue measure resolves problem in movitation

- let
$$E_1 = \{x \in [0,1] | x \in \mathbf{Q}\},\ E_2 = \{x \in [0,1] | x \not\in \mathbf{Q}\}$$

- $\mu^* E_1 = 0$ because E_1 is countable, hence measurable and
$$\mu E_1 = \mu^* E_1 = 0$$

- algebra implies $E_2 = [0,1] \cap \tilde{E_1}$ is measurable

- countable additivity implies $\mu E_1 + \mu E_2 = \mu[0,1] = 1$, hence
$$\mu E_1 = 1$$

4.4 Lebesgue Measurable Functions

Lebesgue measurable functions

- for $f : X \to \mathbf{R} \cup \{-\infty, \infty\}$, *i.e.*, extended real-valued function, the followings are equivalent
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) < a\}$ is measurable
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) \leq a\}$ is measurable
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) > a\}$ is measurable
  - for every $a \in \mathbf{R}$, $\{x \in X | f(x) \geq a\}$ is measurable

- if so,
  - for every $a \in \mathbf{R} \cup \{-\infty, \infty\}$, $\{x \in X | f(x) = a\}$ is measurable

- extended real-valued function, f, called *(Lebesgue) measurable function* if
  - domain is measurable
  - any one of above four statements holds

(refer to page 120 for abstract counterpart)

Properties of Lebesgue measurable functions

- for real-valued measurable functions, f and g, and $c \in \mathbf{R}$
  - $f + c$, cf, $f + g$, fg are measurable
- for every extended real-valued measurable function sequence, $\langle f_n \rangle$
  - $\sup f_n$, $\limsup f_n$ are measurable
  - hence, $\inf f_n$, $\liminf f_n$ are measurable
  - thus, if $\lim f_n$ exists, it is measurable

(refer to page 120 for abstract counterpart)

Almost everywhere - a.e.

- statement, $P(x)$, called *almost everywhere* or *a.e.* if

$$\mu\{x| \sim P(x)\} = 0$$

  - *e.g.*, f said to be equal to g a.e. if $\mu\{x|f(x) \neq g(x)\} = 0$
  - *e.g.*, $\langle f_n \rangle$ said to converge to f a.e. if

$$(\exists E \text{ with } \mu E = 0)(\forall x \not\in E)(\lim f_n(x) = f(x))$$

- facts
  - if f is measurable and $f = g$ i.e., then g is measurable
  - if measurable extended real-valued f defined on $[a, b]$ with $f(x) \in \mathbf{R}$ a.e., then for every $\epsilon > 0$, exist step function g and continuous function h such that

$$\mu\{x||f - g| \geq \epsilon\} < \epsilon, \ \mu\{x||f - h| \geq \epsilon\} < \epsilon$$

Characteristic & simple functions

- for any $A \subset \mathbf{R}$, χ_A called *characteristic function* if

$$\chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \not\in A \end{cases}$$

  - χ_A is measurable *if and only if* A is measurable
- measurable φ called *simple* if for some distinct $\langle a_i \rangle_{i=1}^n$

$$\varphi(x) = \sum_{i=1}^n a_i \chi_{A_i}(x)$$

  where $A_i = \{x|x = a_i\}$

(refer to page 120 for abstract counterpart)

Littlewood's three principles

let $M(E)$ with measurable set, E, denote set of measurable functions defined on E

- *every (measurable) set is nearly finite union of intervals*, *e.g.*,
  - E is measurable *if and only if*

$$(\forall \epsilon > 0)(\exists \{I_i : \text{open interval}\}_{i=1}^n)(\mu^*(E\Delta(\cup I_n)) < \epsilon)$$

- *every (measurable) function is nearly continuous*, *e.g.*,
  - (Lusin's theorem)

$$(\forall f \in M[a,b])(\forall \epsilon > 0)(\exists g \in C[a,b])(\mu\{x|f(x) \neq g(x)\} < \epsilon)$$

- *every convergent (measurable) function sequence is nearly uniformly convergent*, *e.g.*,

$$\begin{aligned}&(\forall \text{ measurable } \langle f_n \rangle \text{ converging to } f \text{ a.e. on } E \text{ with } \mu E < \infty)\\ &\quad (\forall \epsilon > 0 \text{ and } \delta > 0)(\exists A \subset E \text{ with } \mu(A) < \delta \text{ and } N \in \mathbf{N})\\ &\quad (\forall n > N, x \in E \sim A)(|f_n(x) - f(x)| < \epsilon)\end{aligned}$$

Egoroff's theorem

- *Egoroff theorem* - provides stronger version of third statement on page 77

$$\begin{aligned}&(\forall \text{ measurable } \langle f_n \rangle \text{ converging to } f \text{ a.e. on } E \text{ with } \mu E < \infty)\\ &\quad (\exists A \subset E \text{ with } \mu(A) < \epsilon)(f_n \text{ uniformly converges to } f \text{ on } E \sim A)\end{aligned}$$

4.5 Lebesgue Integral

Integral of simple functions

- *canonical representation* of simple function

$$\varphi(x) = \sum_{i=1}^n a_i \chi_{A_i}(x)$$

where a_i are *distinct* $A_i = \{x|\varphi(x) = a_i\}$ - note A_i are *disjoint*

- when $\mu\{x|\varphi(x) \neq 0\} < \infty$ and $\varphi = \sum_{i=1}^n a_i \chi_{A_i}$ is canonical representation, define *integral of φ* by

$$\int \varphi = \int \varphi(x)dx = \sum_{i=1}^n a_i \mu A_i$$

- when E is measurable, define

$$\int_E \varphi = \int \varphi \chi_E$$

(refer to page 121 for abstract counterpart)

Properties of integral of simple functions

- for simple functions φ and ψ that vanish out of finite measure set, *i.e.*, $\mu\{x|\varphi(x) \neq 0\} < \infty$, $\mu\{x|\psi(x) \neq 0\} < \infty$, and for every $a, b \in \mathbf{R}$

$$\int (a\varphi + b\psi) = a \int \varphi + b \int \psi$$

(refer to page 121 for abstract counterpart)

- thus, even for simple function, $\varphi = \sum_{i=1}^n a_i \chi_{A_i}$ that vanishes out of finite measure set, not necessarily in canonical representation,

$$\int \varphi = \sum_{i=1}^n a_i \mu A_i$$

- if $\varphi \geq \psi$ a.e.

$$\int \varphi \geq \int \psi$$

Lebesgue integral of bounded functions

- for bounded function, f, and finite measurable set, E,

$$\sup_{\varphi:\text{ simple, } \varphi \leq f} \int_E \varphi \leq \inf_{\psi:\text{ simple, } f \leq \psi} \int_E \psi$$

  - if f is defined on E, f is measurable function *if and only if*

$$\sup_{\varphi:\text{ simple, } \varphi \leq f} \int_E \varphi = \inf_{\psi:\text{ simple, } f \leq \psi} \int_E \psi$$

- for bounded measurable function, f, defined on measurable set, E, with $\mu E < \infty$, define *(Lebesgue) integral of f over E*

$$\int_E f(x)dx = \sup_{\varphi:\text{ simple, } \varphi \leq f} \int_E \varphi = \inf_{\psi:\text{ simple, } f \leq \psi} \int_E \psi$$

(refer to page 121 for abstract counterpart)

Properties of Lebesgue integral of bounded functions

- for bounded measurable functions, f and g, defined on E with finite measure
  - for every $a, b \in \mathbf{R}$

$$\int_E (af + bg) = a \int_E f + b \int_E g$$

  - if $f \leq g$ a.e.

$$\int_E f \leq \int_E g$$

  - for disjoint measurable sets, $A, B \subset E$,

$$\int_{A \cup B} f = \int_A f + \int_B f$$

(refer to page 123 for abstract counterpart)

- hence,
$$\left|\int_E f\right| \leq \int_E |f| \;\&\; f = g \text{ a.e. } \Rightarrow \int_E f = \int_E g$$

Lebesgue integral of bounded functions over finite interval

- if bounded function, f, defined on $[a, b]$ is Riemann integrable, then f is measurable and
$$\int_{[a,b]} f = R\int_a^b f(x)dx$$
where $R\int$ denotes Riemann integral
- bounded function, f, defined on $[a, b]$ is Riemann integrable *if and only if* set of points where f is discontinuous has measure zero
- for sequence of measurable functions, $\langle f_n \rangle$, defined on measurable E with finite measure, and $M > 0$, if $|f_n| < M$ for every n and $f(x) = \lim f_n(x)$ for every $x \in E$
$$\int_E f = \lim \int_E f_n$$

Lebesgue integral of nonnegative functions

- for nonnegative measurable function, f, defined on measurable set, E, define
$$\int_E f = \sup_{h\text{: bounded measurable function, } \mu\{x|h(x)\neq 0\}<\infty,\ h\leq f} \int_E h$$

(refer to page 122 for abstract counterpart)

- for nonnegative measurable functions, f and g
  - for every $a, b \geq 0$
  $$\int_E (af + bg) = a\int_E f + b\int_E g$$
  - if $f \geq g$ a.e.
  $$\int_E f \leq \int_E g$$
- thus,
  - for every $c > 0$
  $$\int_E cf = a\int_E f$$

Fatou's lemma and monotone convergence theorem for Lebesgue integral

- *Fatou's lemma* - for nonnegative measurable function sequence, $\langle f_n \rangle$, with $\lim f_n = f$ a.e. on measurable set, E

$$\int_E f \leq \liminf \int_E f_n$$

  - note $\lim f_n$ is measurable (page 76), hence f is measurable (page 76)

- *monotone convergence theorem* - for nonnegative increasing measurable function sequence, $\langle f_n \rangle$, with $\lim f_n = f$ a.e. on measurable set, E

$$\int_E f = \lim \int_E f_n$$

(refer to page 122 for abstract counterpart)

- for nonnegative measure function, f, and sequence of disjoint measurable sets, $\langle E_i \rangle$,

$$\int_{\cup E_i} f = \sum \int_{E_i} f$$

Lebesgue integrability of nonnegative functions

- nonnegative measurable function, f, said to be *integrable* over measurable set, E, if

$$\int_E f < \infty$$

(refer to page 123 for abstract counterpart)

- for nonnegative measurable functions, f and g, if f is integrable on measurable set, E, and $g \leq f$ a.e. on E, then g is integrable and

$$\int_E (f - g) = \int_E f - \int_E g$$

- for nonnegative integrable function, f, defined on measurable set, E, and every ϵ, exists $\delta > 0$ such that for every measurable set $A \subset E$ with $\mu A < \epsilon$ (then f is integrable on A, of course),

$$\int_A f < \epsilon$$

Lebesgue integral

- for (any) function, f, define f^+ and f^- such that for every x

$$\begin{aligned} f^+(x) &= \max\{f(x), 0\} \\ f^-(x) &= \max\{-f(x), 0\} \end{aligned}$$

- note $f = f^+ - f^-$, $|f| = f^+ + f^-$, $f^- = (-f)^+$

- measurable function, f, said to be *(Lebesgue) integrable* over measurable set, E, if (nonnegative measurable) functions, f^+ and f^-, are integrable

$$\int_E f = \int_E f^+ - \int_E f^-$$

(refer to page 123 for Lebesgue counterpart)

Properties of Lebesgue integral

- for f and g integrable on measure set, E, and $a, b \in \mathbf{R}$
  - $af + bg$ is integral and
$$\int_E (af + bg) = a \int_E f + b \int_E g$$
  - if $f \geq g$ a.e. on E,
$$\int_E f \geq \int_E g$$
  - for disjoint measurable sets, $A, B \subset E$
$$\int_{A \cup B} f = \int_A f + \int_B g$$

(refer to page 123 for abstract counterpart)

Lebesgue convergence theorem (for Lebesgue integral)

- *Lebesgue convergence theorem* - for measurable g integrable on measurable set, E, and measurable sequence $\langle f_n \rangle$ converging to f with $|f_n| < g$ a.e. on E, (f is measurable (page 76), every f_n is integrable (page 80)) and
$$\int_E f = \lim \int_E f_n$$

(refer to page 123 for abstract counterpart)

Generalization of Lebesgue convergence theorem (for Lebesgue integral)

- *generalization of Lebesgue convergence theorem* - for sequence of functions, $\langle g_n \rangle$, integrable on measurable set, E, converging to integrable g a.e. on E, and sequence of measurable functions, $\langle f_n \rangle$, converging to f a.e. on E with $|f_n| < g_n$ a.e. on E, if
$$\int_E g = \lim \int_E g_n$$
then (f is measurable (page 76), every f_n is integrable (page 80)) and
$$\int_E f = \lim \int_E f_n$$

Comments on convergence theorems

- Fatou's lemma (page 80), monotone convergence theorem (page 80), Lebesgue convergence theorem (page 81), *all* state that under suitable conditions, we say something about
$$\int \lim f_n$$
in terms of
$$\lim \int f_n$$

- Fatou's lemma requires weaker condition than Lebesgue convergence theorem, *i.e.*, only requires "bounded below" whereas Lebesgue converges theorem also requires "bounded above"

$$\int \lim f_n \leq \liminf \int f_n$$

- monotone convergence theorem is somewhat between the two;
  - advantage - applicable even when f not integrable
  - Fatou's lemma and monotone converges theorem very clsoe in sense that can be derived from each other using only facts of positivity and linearity of integral

Convergence in measure

- $\langle f_n \rangle$ of measurable functions said to *converge f in measure* if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n > N)(\mu\{x||f_n - f| > \epsilon\} < \epsilon)$$

- thus, third statement on page 77 implies

$$(\forall \langle f_n \rangle \text{ converging to } f \text{ a.e. on } E \text{ with } \mu E < \infty)(f_n \text{ converge in measure to } f)$$

- however, the converse is *not* true, *i.e.*, exists $\langle f_n \rangle$ converging in measure to f that does not converge to f a.e.
  - *e.g.*, XXX
- Fatou's lemma (page 80), monotone convergence theorem (page 80), Lebesgue convergence theorem (page 81) *remain valid!* even when "convergence a.e." replaced by "convergence in measure"

Conditions for convergence in measure

Proposition 4.1 (necessary condition for converging in measure)

$$(\forall \langle f_n \rangle \textit{ converging in measure to } f)\,(\exists \textit{ subsequence } \langle f_{n_k} \rangle \textit{ converging a.e. to } f)$$

Corollary 4.1 (necessary and sufficient condition for converging in measure) *for sequence $\langle f_n \rangle$ measurable on E with $\mu E < \infty$*

$$\begin{aligned} & \langle f_n \rangle \textit{ converging in measure to } f \\ \Leftrightarrow \quad & (\forall \textit{ subsequence } \langle f_{n_k} \rangle) \left(\exists \textit{ its subsequence } \left\langle f_{n_{k_l}} \right\rangle \textit{ converging a.e. to } f\right) \end{aligned}$$

4.6 Spaces Overview

Diagrams for relations among various spaces

- note from Figure 4.14
  - metric should be defined to utter completeness
  - metric spaces can be induced from normed spaces

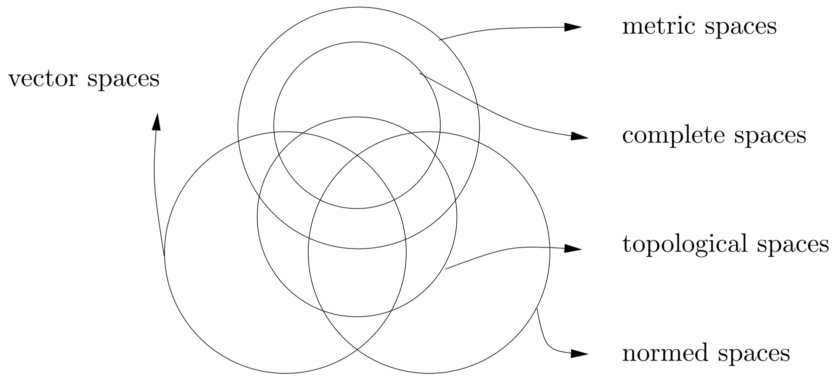

Figure 4.14: diagrams for relations among various spaces

4.7 Classical Banach Spaces

Normed linear space

- X called *linear space* if

$$(\forall x, y \in X, a, b \in \mathbf{R})(ax + by \in X)$$

- linear space, X, called *normed space* with associated norm $\|\cdot\| : X \to \mathbf{R}_+$ if
  -
$$(\forall x \in X)(\|x\| = 0 \Rightarrow x \equiv 0)$$
  -
$$(\forall x \in X, a \in \mathbf{R})(\|ax\| = |a|\|x\|)$$
  - subadditivity
$$(\forall x, y \in X)(\|x + y\| \leq \|x\| + \|y\|)$$

L^p spaces

- $L^p = L^p[0, 1]$ denotes space of (Lebesgue) measurable functions such that

$$\int_{[0,1]} |f|^p < \infty$$

- define $\|\cdot\| : L^p \to \mathbf{R}_+$

$$\|f\| = \|f\|_p = \left(\int_{[0,1]} |f|^p \right)^{1/p}$$

- L^p are *linear normed spaces* with norm $\|\cdot\|_p$ when $p \geq 1$ because
  - $|f(x)|^p + |g(x)|^p \leq 2^p(|f(x)|^p + |g(x)|^p)$ implies $(\forall f, g \in L^p)(f + g \in L^p)$
  - $|\alpha f(x)|^p = |a|^p|f(x)|^p$ implies $(\forall f \in L^p, a \in \mathbf{R})(af \in L^p)$
  - $\|f\| = 0 \Rightarrow f = 0$ a.e.
  - $\|af\| = |a|\|f\|$
  - $\|f + g\| \geq \|f\| + \|g\|$ (Minkowski inequality)

L^∞ space

- $L^\infty = L^\infty[0,1]$ denotes space of measurable functions bounded a.e.
- L^∞ is linear normed space with norm

$$\|f\| = \|f\|_\infty = \text{ess sup}|f| = \inf_{g:g=f \text{ a.e}} \sup_{x\in[0,1]} |g(x)|$$

  - thus

$$\|f\|_\infty = \inf\{M|\mu\{x|f(x) > M\} = 0\}$$

Inequalities in L^∞

- *Minkowski inequality* - for $p \in [1,\infty]$

$$(\forall f, g \in L^p)(\|f + g\|_p \leq \|f\|_p + \|g\|_p)$$

  - if $p \in (1,\infty)$, equality holds *if and only if* $(\exists a, b \geq 0 \text{ with } ab \neq 0)(af = bg \text{ a.e.})$
- Minkowski inequality for $0 < p < 1$:

$$(\forall f, g \in L^p)(f, g \geq 0 \text{ a.e.} \Rightarrow \|f + g\|_p \geq \|f\|_p + \|g\|_p)$$

- *Hölder's inequality* - for $p, q \in [1,\infty]$ with $1/p + 1/q = 1$

$$(\forall f \in L^p, g \in L^q) \left(fg \in L^1 \text{ and } \int_{[0,1]} |fg| \leq \int_{[0,1]} |f|^p \int_{[0,1]} |g|^q \right)$$

  - equality holds *if and only if* $(\exists a, b \geq 0 \text{ with } ab \neq 0)(a|f|^p = b|g|^q \text{ a.e.})$

(refer to page 125 for complete measure spaces counterpart)

Convergence and completeness in normed linear spaces

- $\langle f_n \rangle$ in normed linear space
  - said to *converge* to f, *i.e.*, $\lim f_n = f$ or $f_n \to f$, if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n > N)(\|f_n - f\| < \epsilon)$$

  - called *Cauchy sequence* if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n, m > N)(\|f_n - f_m\| < \epsilon)$$

  - called *summable* if $\sum_{i=1}^n f_i$ converges
  - called *absolutely summable* if $\sum_{i=1}^n |f_i|$ converges
- normed linear space called *complete* if every Cauchy sequence converges
- normed linear space is *complete if and only if* every absolutely summable series is summable

Banach space

- *complete normed linear space* called *Banach space*
- (Riesz-Fischer) L^p spaces are compact, hence Banach spaces
- convergence in L^p called *convergence in mean of order p*
- convergence in L^∞ implies nearly uniformly converges

Approximation in L^p

- $\Delta = \langle d_i \rangle_{i=0}^n$ with $0 = d_1 < d_2 < \cdots < d_n = 1$ called *subdivision* of $[0, 1]$ (with $\Delta_i = [d_{i-1}, d_i]$)
- $\varphi_{f,\Delta}$ for $f \in L^p$ called *step function* if

$$\varphi_{f,\Delta}(x) = \frac{1}{d_i - d_{i+1}} \int_{d_{i-1}}^{d_i} f(t)dt \text{ for } x \in [d_{i-1}, d_i)$$

- for $f \in L^p$ $(1 < p \leq \infty)$, exist $\varphi_{f,\Delta}$ and continuous function, ψ such that

$$\|\varphi_{f,\Delta_i} - f\| < \epsilon \text{ and } \|\psi - f\| < \epsilon$$

  - L^p version of Littlewood's second principle (page 77)

  (refer to page 125 for complete measure spaces counterpart)

- for $f \in L^p$, $\varphi_{f,\Delta} \to f$ as $\max \Delta_i \to 0$, *i.e.*,

$$(\forall \epsilon > 0)(\exists \delta > 0)(\max \Delta_i < \delta \Rightarrow \|\varphi_{f,\Delta} - f\|_p < \epsilon)$$

Bounded linear functionals on L^p

- $F : X \in \mathbf{R}$ for normed linear space X called *linear functional* if

$$(\forall f, g \in F, a, b \in \mathbf{R})(F(af + bg) = aF(f) + bF(g))$$

- linear functional, F, said to be *bounded* if

$$(\exists M)(\forall f \in X)(|F(f)| \leq M\|f\|)$$

- smallest such constant called *norm of F*, *i.e.*,

$$\|F\| = \sup_{f \in X, f \neq 0} |F(f)|/\|f\|$$

Riesz representation theorem

- for every $g \in L^q$ $(1 \leq p \leq \infty)$, following defines a bounded linear functional in L^p

$$F(f) = \int fg$$

  where $\|F\| = \|g\|_q$

- *Riesz representation theorem* - for every bounded linear functional in L^p, F, $(1 \leq p < \infty)$, there exists $g \in L^q$ such that

$$F(f) = \int fg$$

  where $\|F\| = \|g\|_q$

(refer to page 125 for complete measure spaces counterpart)

- for each case, L^q is dual of L^p (refer to page 109 for definition of dual)

4.8 Metric Spaces

Metric spaces

- $\langle X, \rho \rangle$ with nonempty set, X, and *metric* $\rho : X \times X \to \mathbf{R}_+$ called *metric space* if for every $x, y, z \in X$
  - $\rho(x, y) = 0 \Leftrightarrow x = y$
  - $\rho(x, y) = \rho(y, x)$
  - $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ (triangle inequality)
- examples of metric spaces
  - $\langle \mathbf{R}, |\cdot| \rangle$, $\langle \mathbf{R}^n, \|\cdot\|_p \rangle$ with $1 \leq p \leq \infty$
- for $f \subset X$, $S_{x,r} = \{y | \rho(y, x) < r\}$ called *ball*
- for $E \subset X$, $\sup\{\rho(x, y) | x, y \in E\}$ called diameter of E defined by
- ρ called *pseudometric* if 1st requirement removed
- ρ called *extended metric* if $\rho : X \times X \to \mathbf{R}_+ \cup \{\infty\}$

Cartesian product

- for two metric spaces $\langle X, \rho \rangle$ and $\langle Y, \sigma \rangle$, metric space $\langle X \times Y, \tau \rangle$ with $\tau : X \times Y \to \mathbf{R}_+$ such that

  $$\tau((x_1, y_1), (x_2, y_2)) = (\rho(x_1, x_2)^2 + \sigma(y_1, y_2)^2)^{1/2}$$

  called *Cartesian product metric space*
- τ satisfies all properties required by metric
  - *e.g.*, $\mathbf{R}^n \times \mathbf{R}^m = \mathbf{R}^{n+m}$

Open sets - metric spaces

- $O \subset X$ said to be open *open* if

  $$(\forall x \in O)(\exists \delta > 0)(\forall y \in X)(\rho(y, x) < \delta \Rightarrow y \in O)$$

  - X and $\emptyset$ are open
  - intersection of *finite* collection of open sets is open
  - union of *any* collection of open sets is open

Closed sets - metric spaces

- $x \in X$ called *point of closure of $E \subset X$* if

  $$(\forall \epsilon > 0)(\exists y \in E)(\rho(y, x) < \epsilon)$$

  - $\overline{E}$ denotes set of points of closure of E; called *closure* of E
  - $E \subset \overline{E}$
- $F \subset X$ said to be *closed* if

  $$F = \overline{F}$$

  - X and $\emptyset$ are closed
  - union of *finite* collection of closed sets is closed
  - intersection of *any* collection of closed sets is closed
- complement of closed set is open
- complement of open set is closed

Dense sets and separability - metric spaces

- $D \subset X$ said to be dense if

$$\overline{D} = X$$

- X is said to be separable if exists finite dense subset, *i.e.*,

$$(\exists D \subset X)(|D| < \infty \ \&\ \overline{D} = X)$$

- X is separable *if and only if* exists countable collection of open sets $\langle O_i \rangle$ such that for all open $O \subset X$

$$O = \bigcup_{O_i \subset O} O_i$$

Continuous functions - metric spaces

- $f : X \to Y$ for metric spaces $\langle X, \rho \rangle$ and $\langle Y, \sigma \rangle$ called *mapping* or *function* from X into Y
- f said to be *onto* if
$$f(X) = Y$$
- f said to be *continuous* at $x \in X$ if

$$(\forall \epsilon > 0)(\exists \delta > 0)(\forall y \in X)(\rho(y, x) < \delta \Rightarrow \sigma(f(y), f(x)) < \epsilon)$$

- f said to be *continuous* if f is continuous at every $x \in X$
- f is continuous *if and only if* for every open $O \subset Y$, $f^{-1}(O)$ is open
- if $f : X \to Y$ and $g : Y \to Z$ are continuous, $g \circ f : X \to Z$ is continuous

Homeomorphism

- one-to-one mapping of X onto Y (or equivalently, one-to-one correspondece between X and Y), f, said to be *homeomorphism* if
  - both f and f^{-1} are continuous
- X and Y said to be *homeomorphic* if exists homeomorphism
- *topology* is study of properties unaltered by homeomorphisms and such properties called *topological*
- one-to-one correspondece X and Y is homeomorphism *if and only if* it maps open sets in X to open sets in Y and vice versa
- every property defined by means of *open sets* (or equivalently, *closed sets*) or/and being *continuous functions* is *topological one*
  - *e.g.*, f is continuous on X is homeomorphism, then $f \circ h^{-1}$ is continuous function on Y

Isometry

- homeomorphism preserving distance called *isometry*, *i.e.*,

$$(\forall x, y \in X)(\sigma(h(x), h(y)) = \rho(x, y))$$

- X and Y said to be *isometric* if exists isometry
- (from abstract point of view) two isometric spaces are exactly *same*; it's nothing but relabeling of points
- two metrics, ρ and σ on X, said to be *equivalent* if identity mapping of $\langle X, \rho\rangle$ onto $\langle X, \sigma\rangle$ is homeomorphism
  - hence, two metrics are equivalent *if and only if* set in one metric is open whenever open in the other metric

Convergence - metric spaces

- $\langle x_n\rangle$ defined for metric space, X
  - said to *converge* to x, *i.e.*, $\lim x_n = x$ or $x_n \to x$, if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n > N)(\rho(x_n, x) < \epsilon)$$

    - equivalently, every ball about x contains all but finitely many points of $\langle x_n\rangle$
  - said to have cluster point, x, if

$$(\forall \epsilon > 0, N \in \mathbf{N})(\exists n > N)(\rho(x_n, x) < \epsilon)$$

    - equivalently, every ball about x contains infinitely many points of $\langle x_n\rangle$
    - equivalently, every ball about x contains at least one point of $\langle x_n\rangle$
- every convergent point is cluster point
  - converse not true

Completeness - metric spaces

- $\langle x_n\rangle$ of metric space, X, called *Cauchy sequence* if

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n, m > N)(\rho(x_n, x_m) < \epsilon)$$

- convergence sequence is Cauchy sequence
- X said to be *complete* if every Cauchy sequence converges
  - *e.g.*, $\langle \mathbf{R}, \rho\rangle$ with $\rho(x, y) = |x - y|$
- for incomplete $\langle X, \rho\rangle$, exists complete X^* where X is isometrically embedded in X^* as dense set
- if X contained in complete Y, X^* is isometric with $\overline{X}$ in Y

Uniform continuity - metric spaces

- $f : X \to Y$ for metric spaces $\langle X, \rho\rangle$ and $\langle Y, \sigma\rangle$ said to be *uniformly continuous* if

$$(\forall \epsilon > 0)(\exists \delta)(\forall x, y \in X)(\rho(x, y) < \delta \Rightarrow \sigma(f(x), f(y)) < \epsilon)$$

  - example of continuous, but not uniformly continuous function
    - $h : [0, 1) \to \mathbf{R}_+$ with $h(x) = x/(1 - x)$
    - h maps Cauchy sequence $\langle 1 - 1/n\rangle_{n=1}^{\infty}$ in $[0, 1)$ to $\langle n - 1\rangle_{n=1}^{\infty}$ in $\mathbf{R}_+$, which is *not* Cauchy sequence

- homeomorphism f between $\langle X, \rho\rangle$ and $\langle Y, \sigma\rangle$ with both f and f^{-1} uniformly continuous called *uniform homeomorphism*

Uniform homeomorphism

- uniform homeomorphism f between $\langle X, \rho\rangle$ and $\langle Y, \sigma\rangle$ maps every Cauchy sequence $\langle x_n\rangle$ in X mapped to $\langle f(x_n)\rangle$ in Y which is Cauchy
  - being Cauchy sequence, hence, being complete preserved by uniform homeomorphism
  - being uniformly continuous also preserved by uniform homeomorphism

- each of three properties (being Cauchy sequence, being complete, being uniformly continuous) called *uniform property*

- uniform properties are *not* topological properties, *e.g.*, h on page 89
  - is *homeomorphism* between incomplete space $[0, 1)$ and complete space $\mathbf{R}_+$
  - maps Cauchy sequence $\langle 1 - 1/n\rangle_{n=1}^{\infty}$ in $[0, 1)$ to $\langle n - 1\rangle_{n=1}^{\infty}$ in $\mathbf{R}_+$, which is not Cauchy sequence
  - its inverse maps uniformly continuous function sin back to non-uniformly continuity function on $[0, 1)$

Uniform equivalence

- two metrics, ρ and σ on X, said to be *uniformly equivalent* if identity mapping of $\langle X, \rho\rangle$ onto $\langle X, \sigma\rangle$ is uniform homeomorphism, *i.e.*,

$$(\forall \epsilon, \delta > 0, x, y \in X)(\rho(x, y) < \delta \Rightarrow \sigma(x, y) < \epsilon \ \&\ \sigma(x, y) < \delta \Rightarrow \rho(x, y) < \epsilon)$$

- example of uniform equivalence on $X \times Y$
  - any two of below metrics are uniformly equivalent on $X \times Y$

$$\begin{aligned}
\tau((x_1, y_1), (x_2, y_2)) &= (\rho(x_1, x_2)^2 + \sigma(y_1, y_2)^2)^{1/2} \\
\rho_1((x_1, y_1), (x_2, y_2)) &= \rho(x_1, x_2) + \sigma(y_1, y_2) \\
\rho_\infty((x_1, y_1), (x_2, y_2)) &= \max\{\rho(x_1, x_2), \sigma(y_1, y_2)\}
\end{aligned}$$

- for $\langle X, \rho\rangle$ and complete $\langle Y, \sigma\rangle$ and $f : X \to Y$ uniformly continuous on $E \subset X$ into Y, exists unique continuous extension g of f on $\overline{E}$, which is uniformly continuous

Subspaces

- for metric space, $\langle X, \rho \rangle$, metric space $\langle S, \rho_S \rangle$ with $S \subset X$ and ρ_S being restriction of ρ to S, called *subspace* of $\langle X, \rho \rangle$
  - *e.g.* (with standard Euclidean distance)
    - $\mathbf{Q}$ is subspace of $\mathbf{R}$
    - $\left\{ (x, y) \in \mathbf{R}^2 \middle| y = 0 \right\}$ is subspace of $\mathbf{R}^2$, which is isometric to $\mathbf{R}$
- for metric space, X, and its subspace, S,
  - $\overline{E} \subset S$ is closure of E relative to S.
  - $A \subset S$ is closure relative to S *if and only if* $(\exists \overline{F} \subset A)(A = \overline{F} \cap S)$
  - $A \subset O$ is open relative to S *if and only if* $(\exists \text{ open } O \subset A)(A = O \cap S)$
- also
  - every subspace of separable metric space is separable
  - every complete subset of metric space is closed
  - every closed subset of complete metric space is complete

Compact metric spaces

- motivation - want metric spaces where
  - conclusion of Heine-Borel theorem (page 71) are valid
  - many properties of $[0, 1]$ are true, *e.g.*, Bolzano-Weierstrass property (page 91)
- *e.g.*,
  - bounded closed set in $\mathbf{R}$ has *finite open covering property*
- metric space X called *compact metric space* if every open covering of $X, \mathcal{U}$, contains finite open covering of X, *e.g.*,

$$(\forall \text{ open covering of } X, \mathcal{U})(\exists \{O_1, \ldots, O_n\} \subset \mathcal{U})(X \in \cup O_i)$$

- $A \subset X$ called *compact* if compact as subspace of X
  - *i.e.*, every open covering of A contains finite open covering of A

Compact metric spaces - alternative definition

- collection, $\mathcal{F}$, of sets in X said to have *finite intersection property* if every finite subcollection of $\mathcal{F}$ has nonempty intersection
- if rephrase definition of compact metric spaces in terms of *closed* instead of *open*
  - X is called *compact metric space* if every collection of closed sets with empty intersection contains finite subcollection with empty intersection
- thus, X is compact *if and only if* every collection of closed sets with *finite intersection property* has nonempty intersection

Bolzano-Weierstrass property and sequential compactness

- metric space said to
  - have *Bolzano-Weierstrass property* if every sequence has cluster point
  - X said to be *sequentially compact* if every sequence has convergent subsequence
- *X has Bolzano-Weierstrass property* if and only if *sequentially compact* (proof can be found in Proof 11)

Compact metric spaces - properties

- following three statements about metric space are equivalent *(not true for general topological sets)*
  - being compact
  - having Bolzano-Weierstrass property
  - being sequentially compact
- compact metric spaces have corresponding to some of those of complete metric spaces (compare with statements on page 90)
  - every compact subset of metric space is closed *and bounded*
  - every closed subset of compact metric space is compact
- (will show above in following slides)

Necessary condition for compactness

- compact metric space is sequentially compact (proof can be found in Proof 12)
- equivalently, compact metric space has Bolzano-Weierstrass property (page 91)

Necessary conditions for sequentially compactness

- every continuity real-valued function on sequentially compact space is *bounded and assumes its maximum and minimum*
- sequentially compact space is *totally bounded*
- every open covering of sequentially compact space has *Lebesgue number*

Sufficient conditions for compactness

- metric space that is totally bounded and has Lebesgue number for every covering is compact

Borel-Lebesgue theorem

- conditions on pages 91, 91, and 91 imply the following equivalent statements
  - X is *compact*
  - X has *Bolzano-Weierstrass property*
  - X is *sequentially compact*
- above called *Borel-Lebesgue theorem*
- hence, can drop *sequentially* in every statement on page 91, *i.e.*,
  - every continuity real-valued function on ~~sequentially~~ compact space is *bounded and assumes its maximum and minimum*
  - ~~sequentially~~ compact space is *totally bounded*
  - every open covering of ~~sequentially~~ compact space has *Lebesgue number*

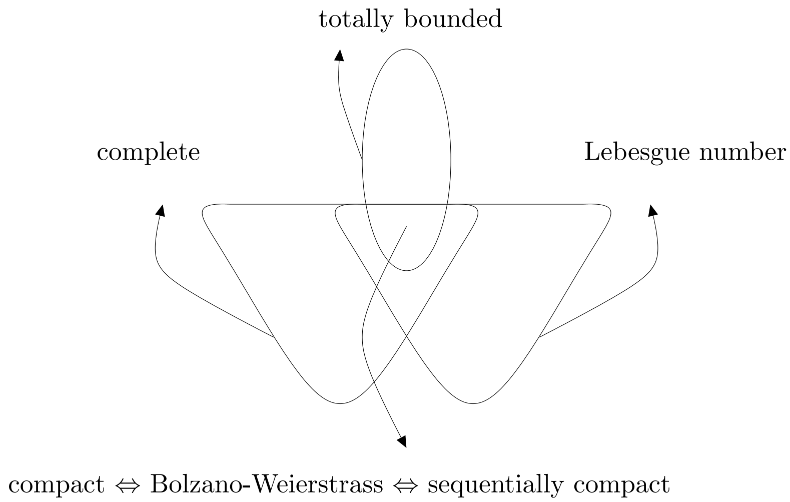

Figure 4.15: diagrams for relations among metric spaces

Compact metric spaces - other facts

- closed subset of compact space is compact
- compact subset of metric space is *closed and bounded*
  - hence, Heine-Borel theorem (page 71) implies
    set of $\mathbf{R}$ *is compact* if and only if *closed and bounded*
- metric space is compact *if and only if* it is complete and totally bounded
- thus, *compactness can be viewed as absolute type of closedness*
  - refer to page 102 for exactly same comments for general topological spaces
- continuous image of compact set is compact
- continuous mapping of compact metric space into metric space is uniformly continuous

Diagrams for relations among metric spaces

- Figure 4.15 shows relations among metric spaces stated on pages 91, 91, 91, and 92

Baire category

- do (more) deeply into certain aspects of complete metric spaces, namely, *Baire theory of category*
- subset E in metric space where $\sim(\overline{E})$ is dense, said to be *nowhere dense*
  - equivalently, $\overline{E}$ contains no nonempty open set
- union of countable collection of *nowhere open sets*, said to be *of first category or meager*
- set not of first category, said to be *of second category or nonmeager*
- complement of set of first category, called *residual or co-meager*

Baire category theorem

- *Baire theorem* - for complete metric space, X, and countable collection of dense open subsets, $\langle O_k \rangle \subset X$, the intersection of the collection
$$\bigcap O_k$$
is dense
  - refer to page 106 for locally compact space version of Baire theorem
- *Baire category theorem* - no nonempty open subset of complete metric space is of first category, *i.e.*, union of countable collection of nowhere dense subsets
- Baire category theorem is *unusual* in that *uniform property, i.e., completeness of metric spaces, implies purely topological nature*[1]

Second category everywhere

- metric or topological spaces with property that "no nonempty open subset of complete metric space is of first category", said to be *of second category everywhere* (with respect to themselves)
- Baire category theorem says *complete metric space* is of second category everywhere
- locally compact Hausdorff spaces are of second category everywhere, too (refer to page 105 for definition of locally compact Hausdorff spaces)
  - for these spaces, though, many of results of category theory follow directly from *local compactness*

Sets of first category

- collection of sets with following properties, called *a σ-ideal of sets*
  - countable union of sets in the collection is, again, in the collection
  - subset of any in the collection is, again, in the collection
- both of below collections are σ-ideal of sets
  - sets of first category in topological space
  - measure zero sets in complete measure space
- sets of first category regards as "small" sets
  - such sets in complete metric spaces no interior points
- interestingly! set of first category in $[0, 1]$ can have Lebesgue measure 1, hence complement of which is residual set of measure zero

[1] "no nonempty open subset of complete metric space is of first category" is purely topological nature because if two spaces are (topologically) homeomorphic, and no nonempty open subsets of one space is of first category, then neither is any nonempty open subset of the other space

Some facts of category theory

- for open set, O, and closed set, F, $\overline{O} \sim O$ and $F \sim F^\circ$ are nowhere dense
- closed set of first category in complete metric space is nowhere dense
- subset of complete metric space is residual *if and only if* contains dense G_δ, hence subset of complete metric space is of first category *if and only if* contained in F_σ whose complement is dense
- for countable collection of closed sets, $\langle F_n \rangle$, $\bigcup {F_n}^\circ$ is residual open set; if $\bigcup F_n$ is complete metric space, O is dense
- some applications of category theory to analysis seem almost too good to be belived; here's one:
- *uniform boundedness principle* - for family, $\mathcal{F}$, of real-valued continuous functions on complete metric space, X, with property that $(\forall x \in X)(\exists M_x \in \mathbf{R})(\forall f \in \mathcal{F})(|f(x)| \leq M_x)$

$$(\exists \text{ open } O, M \in \mathbf{R})(\forall x \in O, f \in \mathcal{F})(|f(x)| \leq M)$$

4.9 Topological Spaces

Motivation for topological spaces

- want to have something like
  - notion of open set is fundamental
  - other notions defined in terms of open sets
  - more general than metric spaces
- why not stick to metric spaces?
  - certain notions have natural meaning *not* consistent with topological concepts derived from metric spaces
    - *e.g.*. weak topologies in Banach spaces

Topological spaces

- $\langle X, \mathfrak{J} \rangle$ with nonempty set X of points and family $\mathfrak{J}$ of subsets, which we call open, having the following properties called *topological spaces*
  - $\emptyset, X \in \mathfrak{J}$
  - $O_1, O_2 \in \mathfrak{J} \Rightarrow O_1 \cap O_2 \in \mathfrak{J}$
  - $O_\alpha \Rightarrow \cup_\alpha O_\alpha \in \mathfrak{J}$
- family, $\mathfrak{J}$, is called *topology*
- for X, *always exist two topologies* defined on X
  - *trivial topology* having only $\emptyset$ and X
  - *discrete topology* for which every subset of X is an open set

Topological spaces associated with metric spaces

- can associate topological space, $\langle X, \mathfrak{J} \rangle$, to any metric space $\langle X, \rho \rangle$ where $\mathfrak{J}$ is family of open sets in $\langle X, \rho \rangle$

  $\because$ because properties in definition of topological space satisfied by open sets in metric space

- $\langle X, \mathfrak{J} \rangle$ assiaciated with metric space, $\langle X, \rho \rangle$ said to be *metrizable*
  - ρ called *metric for* $\langle X, \mathfrak{J} \rangle$

- distinction between metric space and associated topological space is *essential*

  $\because$ because different metric spaces associate same topological space
  - in this case, these metric spaces are equivalent

- metric and topological spaces are couples

Some definitions for topological spaces

- subset $F \subset X$ with $\tilde{F}$ is open called *closed*
- intersection of all closed sets containing $E \subset X$ called *closure* of E denoted by $\overline{E}$
  - $\overline{E}$ is smallest closed set containing E
- $x \in X$ called *point of closure* of $E \subset X$ if every open set containing x meets E, *i.e.*, has nonempty intersection with E
- union of all open sets contained in $E \subset X$ is called *interior* of E denoted by E°
- $x \in X$ called *interior point* of E if exists open set, E, with $x \in O \subset E$

Some properties of topological spaces

- $\emptyset$, X are closed
- union of closed sets is closed
- intersection of any collection of closed sets is closed
- $E \subset \overline{E}$, $\overline{\overline{E}} = \overline{E}$, $\overline{A \cup B} = \overline{A} \cup \overline{B}$
- F closed *if and only if* $\overline{F} = F$
- $\overline{E}$ is set of *points of closure* of E
- $E^\circ \subset E$, $(E^\circ)^\circ = E^\circ$, $(A \cup B)^\circ = A^\circ \cup B^\circ$
- E° is set of *interior points* of E
- $(\tilde{E})^\circ = \sim \overline{E}$

Subspace and convergence of topological spaces

- for subset of $\langle X, \mathfrak{J} \rangle$, A, define *topology $\mathfrak{S}$ for A* with $\mathfrak{S} = \{A \cap O | O \in \mathfrak{J}\}$
  - $\mathfrak{S}$ called *topology inherited from $\mathfrak{J}$*
  - $\langle A, \mathfrak{S} \rangle$ called *subspace* of $\langle X, \mathfrak{J} \rangle$
- $\langle x_n \rangle$ said to *converge* to $x \in X$ if

$$(\forall O \in \mathfrak{J} \text{ containing } x)(\exists N \in \mathbf{N})(\forall n > N)(x_n \in O)$$

  - denoted by

$$\lim x_n = x$$

- $\langle x_n \rangle$ said to have $x \in X$ as *cluster point* if

$$(\forall O \in \mathfrak{J} \text{ containing } x, N \in \mathbf{N})(\exists n > N)(x_n \in O)$$

- $\langle x_n \rangle$ has converging subsequence to $x \in X$, then x is cluster point of $\langle x_n \rangle$
  - converse is *not* true for arbitrary topological space

Continuity in topological spaces

- mapping $f : X \to Y$ with $\langle X, \mathfrak{J} \rangle$, $\langle Y, \mathfrak{S} \rangle$ said to be *continuous* if

$$(\forall O \in \mathfrak{S})(f^{-1}(O) \in \mathfrak{J})$$

- $f : X \to Y$ said to be *continuous at* $x \in X$ if

$$(\forall O \in \mathfrak{S} \text{ containing } f(x))(\exists U \in \mathfrak{J} \text{ containing } x)(f(U) \subset O)$$

- f is continuous *if and only if* f is continuous at every $x \in X$
- for continuous f on $\langle X, \mathfrak{J} \rangle$, restriction g on $A \subset X$ is continuous (proof can be found in Proof 13)
- for A with $A = A_1 \cup A_2$ where both A_1 and A_2 are either open or closed, $f : A \to Y$ with each of both restrictions, $f|A_1$ and $f|A_2$, continuous, is continuous

Homeomorphism for topological spaces

- one-to-one continuous function of X onto Y, f, with continuous inverse function, f^{-1}, called *homeomorphism* between $\langle X, \mathfrak{J} \rangle$ and $\langle Y, \mathfrak{S} \rangle$
- $\langle X, \mathfrak{J} \rangle$ and $\langle Y, \mathfrak{S} \rangle$ said to be *homeomorphic* if exists homeomorphism between them
- homeomorphic spaces are indistinguishable where homeomorphism amounting to relabeling of points (from abstract pointp of view)
- thus, below roles are same
  - role that *homeomorphism plays for topological spaces*
  - role that *isometry plays for metric spaces*
  - role that *isomorphism plays for algebraic systems*

Stronger and weaker topologies

- for two topologies, $\mathfrak{J}$ and $\mathfrak{S}$ for same X with $\mathfrak{S} \supset \mathfrak{J}$
  - $\mathfrak{S}$ said to be *stronger or finer* than $\mathfrak{J}$
  - $\mathfrak{J}$ said to be *weaker or coarser* than $\mathfrak{S}$
- $\mathfrak{S}$ is stronger than $\mathfrak{J}$ *if and only if* identity mapping of $\langle X, \mathfrak{S}\rangle$ to $\langle Y, \mathfrak{J}\rangle$ is continuous
- for two topologies, $\mathfrak{J}$ and $\mathfrak{S}$ for same X, $\mathfrak{J} \cap \mathfrak{S}$ also topology
- for any collection of topologies, $\{\mathfrak{J}_\alpha\}$ for same X, $\cap_\alpha \mathfrak{J}_\alpha$ is topology
- for nonempty set, X, and any collection of subsets of X, $\mathcal{C}$
  - *exists weakest topology containing $\mathcal{C}$, i.e.*, weakest topology where all subsets in $\mathcal{C}$ are open
  - it is intersection of all topologies containing $\mathcal{C}$

Bases for topological spaces

- collection $\mathcal{B}$ of open sets of $\langle X, \mathfrak{J}\rangle$ called *a base for topology,* $\mathfrak{J}$, of X if

$$(\forall O \in \mathfrak{J}, x \in O)(\exists B \in \mathcal{B})(x \in B \subset O)$$

- collection $\mathcal{B}_x$ of open sets of $\langle X, \mathfrak{J}\rangle$ containing x called *a base at* x if

$$(\forall O \in \mathfrak{J} \text{ containing } x)(\exists B \in \mathcal{B}_x)(x \in B \subset O)$$

  - elements of $\mathcal{B}_x$ often called *neighborhoods of* x
  - when no base given, *neighborhood of* x is an open set containing x
- thus, $\mathcal{B}$ of open sets is a base *if and only if* contains a base for every $x \in X$
- for topological space that is also metric space
  - all balls from a base
  - balls centered at x form a base at x

Characterization of topological spaces in terms of bases

- *definition of open sets in terms of base* - when $\mathcal{B}$ is base of $\langle X, \mathfrak{J}\rangle$

$$(O \in \mathfrak{J}) \Leftrightarrow (\forall x \in O)(\exists B \in \mathcal{B})(x \in B \subset O)$$

- often, convenient to specify topology for X by
  - specifying a base of open sets, $\mathcal{B}$, and
  - using above criterion to define open sets
- collection of subsets of X, $\mathcal{B}$, is base for some topology *if and only if*

$$(\forall x \in X)(\exists B \in \mathcal{B})(x \in B)$$
and
$$(\forall x \in X, B_1, B_2 \in \mathcal{B} \text{ with } x \in B_1 \cap B_2)(\exists B_3 \in \mathcal{B})(x \in B_3 \subset B_1 \cap B_2)$$

  - *condition of collection to be basis for some topology*

Subbases for topological spaces

- for $\langle X, \mathfrak{J} \rangle$, collection of open sets, $\mathcal{C}$ called *a subbase* for topology $\mathfrak{J}$ if

$$(\forall O \in \mathfrak{J}, x \in O)(\exists \langle C_i \rangle_{i=1}^n \subset \mathcal{C})(x \in \cap C_i \subset O)$$

  - sometimes convenient to define topology in terms of subbase

- for subbase for $\mathfrak{J}$, $\mathcal{C}$, collection of finite intersections of sets from $\mathcal{C}$ forms base for $\mathfrak{J}$
- any collection of subsets of X is subbase for weakest topology where sets of the collection are open

Axioms of countability

- topological space said to satisfy *first axiom of countability* if exists countable base at every point
  - every metric space satisfies first axiom of countability because for every $x \in X$, set of balls centered at x with rational radii forms base for x
- topological space said to satisfy *second axiom of countability* if exists countable base for the space
  - every metric space satisfies second axiom of countability *if and only if* separable (refer to page 87 for definition of separability)

Topological spaces - facts

- given base, $\mathcal{B}$, for $\langle X, \mathfrak{J} \rangle$
  - $x \in \overline{E}$ *if and only if* $(\exists B \in \mathcal{B})(x \in B \;\&\; B \cap E \neq \emptyset)$
- given base at x for $\langle X, \mathfrak{J} \rangle$, $\mathcal{B}_x$, and base at y for $\langle Y, \mathfrak{S} \rangle$, $\mathfrak{C}_y$
  - $f : X \to Y$ continuous at x *if and only if* $(\forall C \in \mathfrak{C}_y)(\exists B \in \mathcal{B}_x)(B \subset f^{-1}(C))$
- if $\langle X, \mathfrak{J} \rangle$ satisfies *first axiom of countability*
  - $x \in \overline{E}$ *if and only if* $(\exists \langle x_n \rangle$ from $E)(\lim x_n = x)$
  - x cluster point of $\langle x_n \rangle$ *if and only if* exists its subsequence converging to x
- $\langle X, \mathfrak{J} \rangle$ said to be *Lindelöf space* or have *Lindelöf property* if every open covering of X has countable subcover
- second axiom of countability implies *Lindelöf property*

Separation axioms

- why separation axioms
  - properties of topological spaces are (in general) quite different from those of metric spaces
  - often convenient assume additional conditions true in metric spaces
- separation axioms
  - *T_1 - Tychonoff spaces*
    - $(\forall x \neq y \in X)(\exists \text{ open } O \subset X)(y \in O, x \notin O)$

- T_2 - *Hausdorff spaces*
  - $(\forall x \neq y \in X)(\exists \text{ open } O_1, O_2 \subset X \text{ with } O_1 \cap O_2 = \emptyset)(x \in O_1, y \in O_2)$
- T_3 - *regular spaces*
  - T_1 & $(\forall \text{ closed } F \subset X, x \notin F)(\exists \text{ open } O_1, O_2 \subset X \text{ with } O_1 \cap O_2 = \emptyset)(x \in O_1, F \subset O_2)$
- T_4 - *normal spaces*
  - T_1 & $(\forall \text{ closed } F_1, F_2 \subset X)(\exists \text{ open } O_1, O_2 \subset X \text{ with } O_1 \cap O_2 = \emptyset)(F_1 \subset O_1, F_2 \subset O_2)$

Separation axioms - facts

- necessary and sufficient condition for T_1
  - topological space satisfies T_1 *if and only if* every singletone, $\{x\}$, is closed
- important consequences of normality, T_4
  - *Urysohn's lemma* - for normal topological space, X

$$(\forall \text{ disjoint closed } A, B \subset X)(\exists f \in C(X, [0,1]))(f(A) = \{0\}, f(B) = \{1\})$$

  - *Tietze's extension theorem* - for normal topological space, X

$$(\forall \text{ closed } A \subset X, f \in C(A, \mathbf{R}))(\exists g \in C(X, \mathbf{R}))(\forall x \in A)(g(x) = f(x))$$

  - *Urysohn metrization theorem* - *normal* topological space satisfying *second axiom of countability* is *metrizable*

Weak topology generated by functions

- given any set of points, X & any collection of functions of X into $\mathbf{R}$, $\mathcal{F}$, exists weakest totally on X such that all functions in $\mathcal{F}$ is continuous
  - it is weakest topology containing - refer to page 97

$$\mathcal{C} = \bigcup_{f \in \mathcal{F}} \bigcup_{O \subset \mathbf{R}} f^{-1}(O)$$

  - called *weak topology generated by* $\mathcal{F}$

Complete regularity

- for $\langle X, \mathfrak{J} \rangle$ and continuous function collection $\mathcal{F}$, *weak topology* generated by $\mathcal{F}$ is weaker than $\mathfrak{J}$
  - however, if

$$(\forall \text{ closed } F \subset X, x \notin F)(\exists f \in \mathcal{F})(f(A) = \{0\}, f(x) = 1)$$

    then, *weak topology generated by* $\mathcal{F}$ coincides with $\mathfrak{J}$
  - if condition satisfied by $\mathcal{F} = C(X, \mathbf{R})$, X said to be *completely regular* provided X satisfied T_1 (Tychonoff space)
- every normal topological (T_4) space is completely regular (Urysohn's lemma)
- every completely regular space is regular space (T_3)
- complete regularity sometimes called $T_{3\frac{1}{2}}$

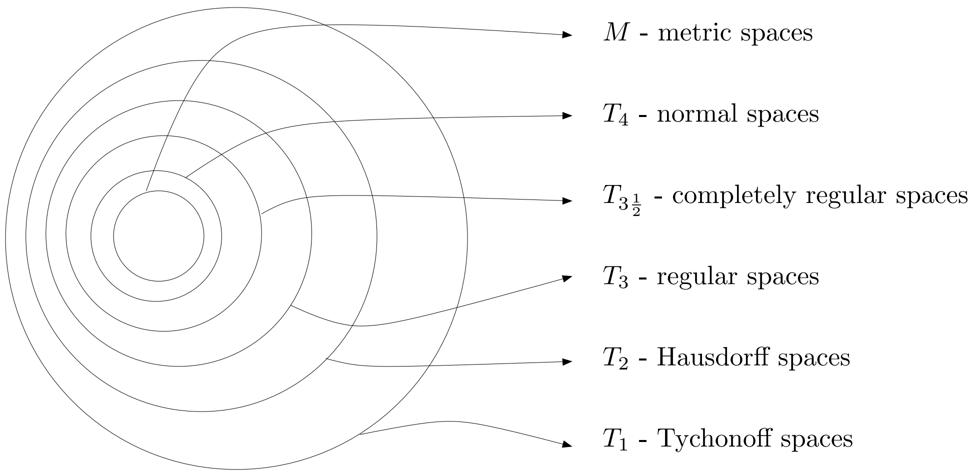

Figure 4.16: diagrams for separation axioms for topological spaces

Diagrams for separation axioms for topological spaces

- Figure 4.16 shows $T_4 \Rightarrow T_{3\frac{1}{2}} \Rightarrow T_3 \Rightarrow T_2 \Rightarrow T_1$
- every metric spaces is normal space

Topological spaces of interest

- very general topological spaces quite bizarre
  - do *not* seem to be much needed in analysis
- only topological spaces (Royden) found useful for analysis are
  - metrizable topological spaces
  - locally compact Hausdorff spaces
  - topological vector spaces
- all above are *completely regular*
- algebraic geometry, however, uses Zariski topology on affine or projective space, topology giving us compact T_1 space which is not Hausdorff

Connectedness

- topological space, X,said to be *connected* if *not* exist two nonempty disjoint open sets, O_1 and O_2, such that $O_1 \cup O_2 = X$
  - such pair, (O_1, O_2), if exist, called *separation of* X
  - pair of disjoint nonempty closed sets, (F_1, F_2), with $F_1 \cup F_2 = X$ is also *separation of* X - because they are also open
- X is connected *if and only if* only subsets that are both closed and open are $\emptyset$ and X
- subset $E \subset X$ said to be *connected* if connected in topology inherited from $\langle X, \mathfrak{J} \rangle$
  - thus, E is connected if not exist two nonempty open sets, O_1 and O_2, such that $E \subset O_1 \cup O_2$ and $E \cap O_1 \cap O_2 = \emptyset$

Properties of connected space, component, and local connectedness

- if exists continuous mapping of connected space to topological space, Y, Y is connected
- *(generalized version of) intermediate value theorem* - for $f : X \to \mathbf{R}$ where X is connected

$$(\forall x, y \in X, c \in \mathbf{R} \text{ with } f(x) < c < f(y))(\exists z \in X)(z = f(z))$$

- subset of $\mathbf{R}$ is connected *if and only if* is either interval or singletone
- for $x \in X$, union of all connected sets containing x is called *component*
  - component is *connected and closed*
  - two components containing same point coincide
  - thus, *X is disjoint union of components*
- X said to be *locally connected* if exists base for X consisting of connected sets
  - components of locally connected space are *open*
  - space *can be connected, but not locally connected*

Product topological spaces

- for $\langle X, \mathfrak{J} \rangle$ and $\langle Y, \mathfrak{S} \rangle$, topology on $X \times Y$ taking as *a base* the following

$$\{O_1 \times O_2 | O_1 \in \mathfrak{J}, O_2 \in \mathfrak{S}\}$$

  called *product topology* for $X \times Y$
  - for metric spaces, X and Y, *product topology is product metric*
- for indexed family with index set, $\mathcal{A}$, $\langle X_\alpha, \mathfrak{J}_\alpha \rangle$, product topology on $\bigtimes_{\alpha \in \mathcal{A}} X_\alpha$ defined as taking as *a base* the following

$$\left\{ \bigtimes X_\alpha \middle| O_\alpha \in \mathfrak{J}_\alpha, O_\alpha = X_\alpha \text{ except finite number of } \alpha \right\}$$

- $\pi_\alpha : \bigtimes X_\alpha \to X_\alpha$ with $\pi_\alpha(y) = x_\alpha$, *i.e.*, α-th coordinate, called *projection*
  - every π_α continuous
  - $\bigtimes X_\alpha$ *weakest topology* with continuous π_α's
- if $(\forall \alpha \in \mathcal{A})(X_\alpha = X)$, $\bigtimes X_\alpha$ denoted by $X^{\mathcal{A}}$

Product topology with countable index set

- for countable $\mathcal{A}$
  - $\bigtimes X_\alpha$ *denoted by* X^ω *or* $X^{\mathbf{N}}$ $\because$ only # elements of $\mathcal{A}$ important
    - *e.g.*, $\mathbf{2}^\omega$ is *Cantor set* if denoting discrete topology with two elements by $\mathbf{2}$
- if X is metrizable, X^ω is metrizable
- $\mathbf{N}^\omega = \mathbf{N}^{\mathbf{N}}$ *is topology space homeomorphic to* $\mathbf{R} \sim \mathbf{Q}$ when denoting discrete topology with countable set also by $\mathbf{N}$

Product topologies induced by set and continuous functions

- for $I = [0, 1]$, $I^{\mathcal{A}}$ called *cube*
- I^{ω} is metrizable, and called *Hilbert cube*
- for any set X and any collection of $f : X \to [0, 1]$, $\mathcal{F}$ with $(\forall x \neq y \in X)(\exists f \in \mathcal{F})(f(x) \neq f(y))$
  - can define *one-to-one mapping of $\mathcal{F}$ into I^X* with $f(x)$ as x-th coordinate of f
    - $\pi_x : \mathcal{F} \to I$ (mapping of $\mathcal{F}$ into I) with $\pi_x(f) = f(x)$
    - topology that $\mathcal{F}$ inherits as subspace of I^X called *topology of pointwise convergence* (because π_x is project, hence continuous)
  - can define *one-to-one mapping of X into $I^{\mathcal{F}}$* with $f(x)$ as f-th coordinate of x
    - topology of X as subspace of $I^{\mathcal{F}}$ is *weak topology generated by $\mathcal{F}$*
    - if every $f \in \mathcal{F}$ is continuous,
      - topology of X into $I^{\mathcal{F}}$ is continuous
      - if for every closed $F \subset X$ and for each $x \notin F$, exists $f \in \mathcal{F}$ such that $f(x) = 1$ and $f(F) = \{0\}$, then X *is homeomorphic to image of $I^{\mathcal{F}}$*

4.10 Compact and Locally Compact Spaces

Compact spaces

- compactness for metric spaces (page 90) can be generalized to topological spaces
  - things are very much similar to those of metrics spaces
- for subset $K \subset X$, collection of open sets, $\mathcal{U}$, the union of which K is contained in called *open covering* of K
- topological space, X, said to be *compact* if every open convering of contains finite subcovering
- $K \subset X$ said to be *compact* if compact as subspace of X
  - or equivalently, K is compact if every covering of K *by open sets of X* has finite subcovering
  - thus, Heine-Borel (page 71) says every closed and bounded subset of $\mathbf{R}$ is compact
- for $\mathcal{F} \subset \mathcal{P}(X)$ any finite subcollection of which has nonempty intersection called *finite intersection property*
- thus, topological space compact *if and only if* every collection with *finite intersection property* has nonempty intersection

Compact spaces - facts

- *compactness can be viewed as absolute type of closedness* because
  - closed subset of compact space is compact
  - compact subset of Hausdorff space is closed
- refer to page 92 for exactly the same comments for metric spaces
- thus, every compact set of $\mathbf{R}$ is closed and bounded
- continuous image of compact set is compact
- one-to-one continuous mapping of compact space into Hausdorff space is homeomorphism

Refinement of open covering

- for open covering of X, $\mathcal{U}$, open covering of X every element of which is subset of element of $\mathcal{U}$, called *refinement* of $\mathcal{U}$ or said to *refine* $\mathcal{U}$
- X is cmopact *if and only if* every open covering has finite refinement
- any two open covers, $\mathcal{U}$ and $\mathcal{V}$, have common refinement, *i.e.*,

$$\{U \cap V | U \in \mathcal{U}, V \in \mathcal{V}\}$$

Countable compactness and Lindelöf

- topological space for which every open covering has countable subcovering said to be *Lindelöf*
- topological space for which every countable open covering has finite subcovering said to be *countably compact* space
- thus, topological space is compact *if and only if* both Lindelöf and countably compact
- every second countable space is Lindelöf
- thus, countable compactness coincides with compactness if second countable (*i.e.*, satisfying second axiom of countability)
- continuous image of compact countably compact space is countably compact

Bolzano-Weierstrass property and sequential compactness

- topological space, X, said to have *Bolzano-Weierstrass property* if every sequence, $\langle x_n \rangle$, in X has at least one cluster point, *i.e.*,

$$(\forall \langle x_n \rangle)(\exists x \in X)(\forall \epsilon > 0, N \in \mathbf{N})(\exists n > N, O \subset X)(x \in O, O \text{ is open}, x_n \in O)$$

- topological space has *Bolzano-Weierstrass properties if and only if* countably compact
- topological space said to be *sequentially compact* if every sequence has converging subsequence
- sequentially compact space is countably compact
- thus, Lindelöf coincides with compactness if sequentially compact
- countably compact and first countable (*i.e.*, satisfying first axiom of countability) space is sequentially compact

Diagrams for relations among topological spaces

- Figure 4.17 shows relations among topological spaces stated on pages 103 and 103

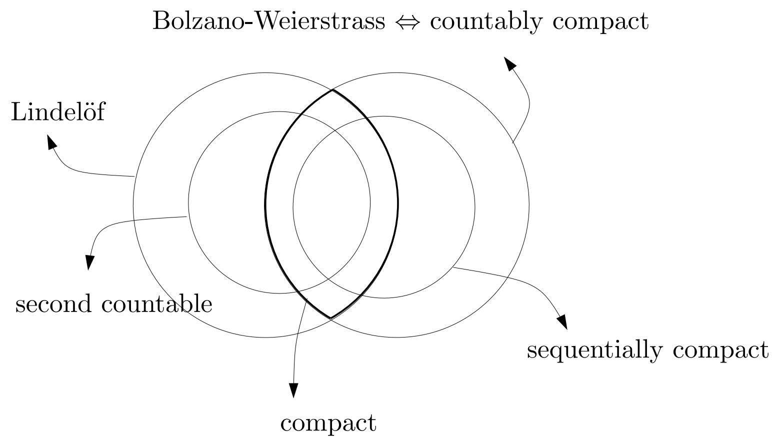

Figure 4.17: diagrams for relations among topological spaces

Real-valued functions on topological spaces

- continuous real-valued function on countably compact space is bounded and assumes maximum and minimum
- $f : X \to \mathbf{R}$ with topological space, X, called *upper semicontinuous* if $\{x \in X | f(x) < \alpha\}$ is open for every $\alpha \in \mathbf{R}$
- stronger statement - upper semicontinuous real-valued function on countably compact space is bounded (from above) and assumes maximum
- *Dini* - for sequence of upper semicontinuous real-valued functions on countably compact space, $\langle f_n \rangle$, with property that $\langle f_n(x) \rangle$ decreases monotonically to zero for every $x \in X$, $\langle f_n \rangle$ converges to zero uniformly

Products of compact spaces

- *Tychonoff theorem - (probably) most important theorem in general topology*
- most applications in analysis need only special case of product of (closed) intervals, but this special case does not seem to be easire to prove than general case, *i.e.*, Tychonoff theorem
- lemmas needed to prove Tychonoff theorem
  - for collection of subsets of X with finite intersection property, $\mathcal{A}$, exists collection $\mathcal{B} \supset \mathcal{A}$ with finite intersection property that is maximal with respect to this property, *i.e.*, no collection with finite intersection property properly contains $\mathcal{B}$
  - for collection, $\mathcal{B}$, of subsets of X that is maximal with respect to finite intersection property, each intersection of finite number of sets in $\mathcal{B}$ is again in $\mathcal{B}$ and each set that meets each set in $\mathcal{B}$ is itself in $\mathcal{B}$
- *Tychonoff theorem* - product space $\bigtimes X_\alpha$ is compact for indexed family of compact topological spaces, $\langle X_\alpha \rangle$

Locally compact spaces

- topological space, X, with

$$(\forall x \in X)(\exists \text{ open } O \subset X)(x \in O, \overline{O} \text{ is compact})$$

  called *locally compact*

- topological space is locally compact *if and only if* set of all open sets with compact closures forms base for the topological space

- every compact space is locally compact
  - but converse it *not* true
    - *e.g.*, Euclidean spaces $\mathbf{R}^n$ are locally compact, but not compact

Locally compact Hausdorff spaces

- *locally compact Hausdorff spaces* constitute one of most important classes of topological spaces

- so useful is combination of Hausdorff separation axioms in connection with compactness that French usage (following Bourbaki) reserves term 'compact space' for those compact and Hausdorff, using term 'pseudocompact' for those not Hausdorff!

- following slides devote to establishing some of their basic properties

Support and subordinateness

- for function, f, on topological spaces, closure of $\{x|f(x) \neq 0\}$, called *support* of f, *i.e.*,

$$\mathbf{support}\, f = \overline{\{x|f(x) \neq 0\}}$$

- given covering $\{O_\lambda\}$ of X, collection $\{\varphi_\alpha\}$ with $\varphi_\alpha : X \to \mathbf{R}$ satisfying

$$(\forall \varphi_\alpha)(\exists O_\lambda)(\mathbf{support}\, \varphi_\alpha \subset O_\lambda)$$

  said to be *subordinate to* $\{O_\lambda\}$

Some properties of locally compact Hausdorff spaces

- for compact subset, K, of locally compact Hausdorff space, X
  - exists open subset with compact closure, $O \subset X$, containing K
  - exists continuous nonnegative function, f, on X, with

$$(\forall x \in K)(f(x) = 1) \text{ and } (\forall x \not\in O)(f(x) = 0)$$

    if K is G_δ, may take $f < 1$ in $\tilde{K}$

- for open covering, $\{O_\lambda\}$, for compact subset, K, of locally compact Hausdorff space, exists $\langle \varphi_i \rangle_{i=1}^n \subset C(X, \mathbf{R}_+)$ subordinate to $\{O_\lambda\}$ such that

$$(\forall x \in K)(\varphi_1(x) + \cdots + \varphi_n(x) = 1)$$

Local compactness and second Baire category

- for locally compact space, X, and countable collection of dense open subsets, $\langle O_k \rangle \subset X$, the intersection of the collection
$$\bigcap O_k$$
is dense
  - analogue of Baire theorem for complete metric spaces (refer to page 93 for Baire theorem)
- thus, *every locally compact space is locally of second Baire category with respect to itself*

Local compactness, Hausdorffness, and denseness

- for countable union, $\bigcup F_n$, of closed sets containing open subset, O, in locally compact space, union of interiors, $\bigcup {F_n}^\circ$, is open set dense in O
- dense subset of Hausdorff space, X, which is locally compact in its subspace topology, is open subset of X
- subset, Y, of locally compact Hausdorff space is locally compact in its subspace topology *if and only if* Y is relatively open subset of $\overline{Y}$

Alexandroff one-point compactification

- for locally compact Hausdorff space, X, can form X^* by adding single point $\omega \notin X$ to X and take set in X^* to be open if it is either open in X or complement of compact subset in X, then
  - X^* is compact Hausdorff spaces
  - identity mapping of X into X^* is homeomorphism of X and $X^* \sim \{\omega\}$
  - X^* called *Alexandroff one-point compactification of X*
  - ω often referred to as *infinity in X^**
- continuous mapping, f, from topological space to topological space inversely mapping compact set to compact set, said to be *proper*
- proper maps from locally compact Hausdorff space into locally compact Hausdorff space are precisely those continuous maps of X into Y tha can be extended to continuous maps f^* of X^* into Y^* by taking point at infinity in X^* to point at infinity in Y^*

σ-compact spaces

- XXX - Royden p203

Manifolds

- connected Hausdorff space with each point having neighborhood homeomorphic to ball in $\mathbf{R}^n$ called n-dimensional *manifold*
- sometimes say manifold is connected Hausdorff space that is *locally Euclidean*
- thus, manifold has all local properties of Euclidean space; particularly *locally compact and locally connected*
- neighborhood homeomorphic to ball called *coordinate neighborhood* or *coordinate ball*
- pair $\langle U, \varphi \rangle$ with coordinate ball, U, with homeomorphism from U onto ball in $\mathbf{R}^n$, φ, called *coodinate chart*; φ called *coordinate map*
- coordinate (in $\mathbf{R}^n$) of point, $x \in U$, under φ said to be *coordinate of x* in the chart

Equivalent properties for manifolds

- for manifold, M, the following are equivalent
  - M is paracompact
  - M is σ-compact
  - M is Lindelöf
  - every open cover of M has star-finite open refinement
  - exist sequence of open subsets of M, $\langle O_n \rangle$, with $\overline{O_n}$ compact, $\overline{O_n} \subset O_{n+1}$, and $M = \bigcup O_n$
  - exists proper continuous map, $\varphi : M \to [0, \infty)$
  - M is second countable

4.11 Banach Spaces

Vector spaces

- set X with $+ : X \times X \to X$, $\cdot : \mathbf{R} \times X \to X$ satisfying the following properties called *vector space* or *linear space* or *linear vector space* over $\mathbf{R}$

  - for all $x, y, z \in X$ and $\lambda, \mu \in \mathbf{R}$

$$
\begin{aligned}
x + y &= y + x && \text{- additive commutativity} \\
(x + y) + z &= x + (y + z) && \text{- additive associativity} \\
(\exists 0 \in X)\ x + 0 &= x && \text{- additive identity} \\
\lambda(x + y) &= \lambda x + \lambda y && \text{- distributivity of multiplicative over addition} \\
(\lambda + \mu)x &= \lambda x + \mu x && \text{- distributivity of multiplicative over addition} \\
\lambda(\mu x) &= (\lambda\mu)x && \text{- multiplicative associativity} \\
0 \cdot x &= 0 \in X && \\
1 \cdot x &= x &&
\end{aligned}
$$

Norm and Banach spaces

- $\|\cdot\| : X \to \mathbf{R}_+$ with vector space, X, called *norm* if

  for all $x, y \in X$ and $\alpha \in \mathbf{R}$

$$
\begin{aligned}
\|x\| = 0 \Leftrightarrow x = 0 && \text{- positive definiteness / positiveness /point-separating} \\
\|x + y\| \geq \|x\| + \|y\| && \text{- triangle inequality / subadditivity} \\
\|\alpha x\| = |\alpha|\|x\| && \text{- Absolute homogeneity}
\end{aligned}
$$

- *normed vector space* that is *complete metric space* with metric induced by norm, *i.e.*, $\rho : X \times X \to \mathbf{R}_+$ with $\rho(x, y) = \|x - y\|$, called *Banach space*
  - can be said to be class of spaces endowed with both topological and algebraic structure
- examples include
  - L^p with $1 \leq p \leq \infty$ (page 84),
  - $C(X) = C(X, \mathbf{R})$, *i.e.*, space of all continuous real-valued functions on *compact* space, X

Properties of vector spaces

- normed vector space is complete *if and only if* every absolutely summable sequence is summable

Subspaces of vector spaces

- nonempty subset, S, of vector space, X, with $x, y \in S \Rightarrow \lambda x + \mu y \in S$, called *subspace* or *linear manifold*
- intersection of any family of linear manifolds is linear manifold
- hence, for $A \subset X$, exists smallest linear manifold containing A, often denoted by $\{A\}$
- if S is closed as subset of X, called *closed linear manifold*
- some definitions
  - $A + x$ defined by $\{y + x | y \in A\}$, called *translate* of A by x
  - λA defined by $\{\lambda x | x \in A\}$
  - $A + B$ defined by $\{x + y | x \in A, y \in B\}$

Linear operators on vector spaces

- mapping of vector space, X, to another (possibly same) vector space called *linear mapping*, or *linear operator*, or *linear transformation* if

$$(\forall x, y \in X, \alpha, \beta \in \mathbf{R})(A(\alpha x + \beta y y) = \alpha(Ax) + \beta(Ay))$$

- linear operator called *bounded* if

$$(\exists M)(\forall x \in X)(\|Ax\| \leq M\|x\|)$$

- least such bound called *norm* of linear operator, *i.e.*,

$$M = \sup_{x \in X, x \neq 0} \|Ax\| / \|x\|$$

  - linearity implies

$$M = \sup_{x \in X, \|x\|=1} \|Ax\| = \sup_{x \in X, \|x\| \leq 1} \|Ax\|$$

Isomorphism and isometrical isomorphism

- bounded linear operator from X to Y called *isomorphism* if exists bounded inverse linear operator, *i.e.*,

$$(\exists A : X \to Y, B : Y \to X)(AB \text{ and } BA \text{ are identity})$$

- isomorphism between two normed vector spaces that preserve norms called *isometrical isomorphism*
- from abstract point of view, isometrically isomorphic spaces are *identical*, *i.e.*, isometrical isomorphism merely amounts to *element renaming*

Properties of linear operators on vector spaces

- for linear operators, point continuity $\Rightarrow$ boundedness $\Rightarrow$ uniform continuity, *i.e.*,
  - bounded linear operator is uniformly continuous
  - linear operator continuous at one point is bounded
- *space of all bounded linear operators from normed vector space to Banach space is Banach space*

Linear functionals on vector spaces

- linear operator from vector space, X, to $\mathbf{R}$ called *linear functional*, *i.e.*, $f : X \to \mathbf{R}$ such that for all $x, y \in X$ and $\alpha, \beta \in \mathbf{R}$

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y)$$

- want to extend linear functional from subspace to whole vector space while preserving properties of functional

Hahn-Banach theorem

- *Hahn-Banach theorem* - for *vector space*, X, and linear functional, $p : X \to \mathbf{R}$ with

$$(\forall x, y \in X, \alpha \geq 0)(p(x+y) \leq p(x) + p(y) \text{ and } p(\alpha x) = \alpha p(x))$$

  and for subspace of X, S, and linear functional, $f : S \to \mathbf{R}$, with

$$(\forall s \in S)(f(s) \leq p(s))$$

  exists linear functional, $F : X \to \mathbf{R}$, such that

$$(\forall s \in S)(F(s) = f(s)) \text{ and } (\forall x \in X)(F(x) \leq p(x))$$

- corollary - for normed vector space, X, exists bounded linear functional, $f : X \to \mathbf{R}$

$$f(x) = \|f\|\|x\|$$

Dual spaces of normed spaces

- space of *bounded linear functionals* on *normed space*, X, called *dual* or *conjugate* of X, denoted by X^*
- every dual is Banach space (refer to page 108)
- dual of L^p is (isometrically isomorphic to) L^q for $1 \leq p < \infty$
  - exists natural representation of bounded linear functional on L^p by L^q (by Riesz representation theorem on page 85)
- *not* every bounded linear functionals on L^∞ has natural representation (proof can be found in Proof 14)

Natural isomorphism

- define linear mapping of normed space, X, to X^{**} (*i.e.*, dual of dual of X), $\varphi : X \to X^{**}$ such that for $x \in X$, $(\forall f \in X^*)((\varphi(x))(f) = f(x))$
  - then, $\|\varphi(x)\| = \sup_{\|g\|=1, g \in X^*} g(x) \leq \sup_{\|g\|=1, g \in X^*} \|g\|\|x\| = \|x\|$
  - by corollary on page 109, there exists $f \in X^*$ such that $f(x) = \|x\|$, then $\|f\| = 1$, and $f(x) = \|x\|$, thus $\|\varphi(x)\| = \sup_{\|g\|=1, g \in X^*} g(x) \geq f(x) = \|x\|$
  - thus, $\|\varphi(x)\| = \|x\|$, hence φ is isometrically isomorphic linear mapping of X onto $\varphi(X) \subset X^{**}$, which is subspace of X^{**}
  - φ called *natural isomorphism* of X into X^{**}
  - X said to be *reflexive* if $\varphi(X) = X^{**}$
- thus, L^p with $1 < p < \infty$ is reflexive, but L^1 and L^∞ are not
- note X may be isometric with X^{**} without reflexive

Completeness of natural isomorphism

- for natural isomorphism, φ
- X^{**} is complete, hence Banach space
  - because bounded linear functional to $\mathbf{R}$ (refer to page 108)
- thus, closure of $\varphi(X)$ in X^{**}, $\overline{\varphi(X)}$, complete (refer to page 90)
- therefore, *every normed vector space (X) is isometrically isomorphic to dense subset of Banach spaces (X^{**})*

Hahn-Banach theorem - complex version

- *Bohnenblust and Sobczyk* - for *complex* vector space, X, and linear functional, $p : X \to \mathbf{R}$ with

$$(\forall x, y \in X, \alpha \in \mathbf{C})(p(x + y) \leq p(x) + p(y) \text{ and } p(\alpha x) = |\alpha| p(x))$$

  and for subspace of X, S, and (complex) linear functional, $f : S \to \mathbf{C}$, with

$$(\forall s \in S)(|f(s)| \leq p(s))$$

  exists linear functional, $F : X \to \mathbf{R}$, such that

$$(\forall s \in S)(F(s) = f(s))$$

  and

$$(\forall x \in X)(|F(x)| \leq p(x))$$

Open mapping on topological spaces

- mapping from topological space to another topological space the image of each open set by which is open called *open mapping*
- hence, one-to-one continuous open mapping is *homeomorphism*
- (will show) continuous linear transformation of Banach space onto another Banach space is always open mapping
- (will) use above to provide criteria for continuity of linear transformation

Closed graph theorem (on Banach spaces)

- every continuous linear transformation of Banach space onto Banach space is open mapping
  - in particular, if the mapping is one-to-one, it is isomorphism
- for linear vector space, X, complete in two norms, $\|\cdot\|_A$ and $\|\cdot\|_B$, with $C \in \mathbf{R}$ such that $(\forall x \in X)(\|x\|_A \leq C\|x\|_B)$, two norms are equivalent, *i.e.*, $(\exists C' \in \mathbf{R})(\forall x \in X)(\|x\|_B \leq C'\|x\|_A)$
- *closed graph theorem* - linear transformation, A, from Banach space, A, to Banach space, B, with property that "if $\langle x_n \rangle$ converges in X to $x \in X$ and $\langle Ax_n \rangle$ converges in Y to $y \in Y$, then $y = Ax$" is continuous
  - equivalent to say, if graph $\{(x, Ax) | x \in X\} \subset X \times Y$ is closed, A is continuous

Principle of uniform boundedness (on Banach spaces)

- *principle of uniform boundedness -* for family of bounded linear operators, $\mathcal{F}$ from Banach space, X, to normed space, Y, with

$$(\forall x \in X)(\exists M_x)(\forall T \in \mathcal{F})(\|Tx\| \leq M_x)$$

  then operators in $\mathcal{F}$ is uniformly bounded, *i.e.*,

$$(\exists M)(\forall T \in \mathcal{F})(\|T\| \leq M)$$

Topological vector spaces

- just as notion of metric spaces generalized to notion of topological spaces
- *notion of normed linear space generalized to notion of topological vector spaces*
- linear vector space, X, with topology, $\mathfrak{J}$, equipped with continuous addition, $+ : X \times X \to X$ and continuous multiplication by scalars, $+ : \mathbf{R} \times X \to X$, called *topological vector space*

Translation invariance of topological vector spaces

- for topological vector space, translation by $x \in X$ is homeomorphism (due to continuity of addition)
  - hence, $x + O$ of open set O is open
  - every topology with this property said to be *translation invariant*
- for translation invariant topology, $\mathfrak{J}$, on X, and base, $\mathcal{B}$, for $\mathfrak{J}$ at 0, set

$$\{x + U | U \in \mathcal{B}\}$$

  forms *a base* for $\mathfrak{J}$ at x
- hence, sufficient to give a base at 0 to determine *translation invariance of topology*
- base at 0 often called *local base*

Sufficient and necessarily condition for topological vector spaces

- for topological vector space, X, can find base, $\mathcal{B}$, satisfying following properties

$$\begin{aligned}
&(\forall U, V \in \mathcal{B})(\exists W \in \mathcal{B})(W \subset U \cap V)\\
&(\forall U \in \mathcal{B}, x \in U)(\exists V \in \mathcal{B})(x + V \subset U)\\
&(\forall U \in \mathcal{B})(\exists V \in \mathcal{B})(V + V \subset U)\\
&(\forall U \in \mathcal{B}, x \in X)(\exists \alpha \in \mathbf{R})(x \in \alpha U)\\
&(\forall U \in \mathcal{B}, 0 < |\alpha| \leq 1 \in \mathbf{R})(\alpha U \subset U, \alpha U \subset \mathcal{B})
\end{aligned}$$

- conversely, for collection, $\mathcal{B}$, of subsets containing 0 satisfying above properties, exists topology for X making X *topological vector space* with $\mathcal{B}$ as base at 0
  - this topology is Hausdorff *if and only if*

$$\bigcap \{U \in \mathcal{B}\} = \{0\}$$

- for normed linear space, can take $\mathcal{B}$ to be set of spheres centered at 0, then $\mathcal{B}$ satisfies above properties, hence can form *topological vector space*

Topological isomorphism

- in topological vector space, can compare neighborhoods at one point with neighborhoods of another point by translation
- for mapping, f, from topological vector space, X, to topological vector space, Y, such that

$$
\begin{aligned}
&(\forall \text{ open } O \subset Y \text{ with } 0 \in O)(\exists \text{ open } U \subset X \text{ with } 0 \in U)\\
&(\forall x \in X)(f(x+U) \subset f(x) + O)
\end{aligned}
$$

  said to be *uniformly continuous*
- linear transformation, f, is uniformly continuous if continuous at one point
- continuous one-to-one mapping, φ, from X onto Y with continuous φ^{-1} called *(topological) isomorphism*
  - in abstract point of view, isomorphic spaces are *same*
- *Tychonoff* - finite-dimensional Hausdorff topological vector space is topologically isomorphic to $\mathbf{R}^n$ for some n

Weak topologies

- for vector space, X, and collection of linear functionals, $\mathcal{F}$, weakest topology generated by $\mathcal{F}$, *i.e.*, in way that each functional in $\mathcal{F}$ is continuous in that topology, called *weak topology generated by* $\mathcal{F}$
  - translation invariant
  - base at 0 given by sets

$$
\{x \in X | \forall f \in \mathcal{G}, |f(x)| < \epsilon\}
$$

    for all finite $\mathcal{G} \subset \mathcal{F}$ and $\epsilon > 0$
  - basis satisfies properties on page 111, hence, (above) weak topology makes *topological vector space*
- for *normed* vector space, X, and collection of continuous functionals, $\mathcal{F}$, *i.e.*, $\mathcal{F} \subset X^*$, weak topology generated by $\mathcal{F}$ *weaker than* (fewer open sets) norm topology of X
- metric topology generated by norm called *strong topology of X*
- weak topology generated by X^* called *weak topology of X*

Strongly and weakly open and closed sets

- open and closed sets of strong topology called *strongly open* and *strongly closed*
- open and closed sets of weak topology called *weakly open* and *weakly closed*
- wealy closed set is strongly closed, but converse not true
- however, these coincides for linear manifold, *i.e.*, linear manifold is weakly closed *if and only if* strongly closed
- every strongly convergent sequence (or net) is weakly convergent

Weak* topologies

- for normed space, *weak topology of X^** is weakest topology for which all functionals in X^{**} are continuous
- turns out that weak topology of X^* is less useful than weak topology generated by X, *i.e.*, that generated by $\varphi(X)$ where φ is the natural embedding of X into X^{**} (refer to page 109)
- (above) weak topology generated by $\varphi(X)$ called *weak* topology for X^**
  - even *weaker than* weak topology of X^*
  - thus, weak* closed subset of is weakly closed, and weak convergence implies weak* convergence
- base at 0 for weak* topology given by sets

$$\{f|\forall x \in A, |f(x)| < \epsilon\}$$

  for all finite $A \subset X$ and $\epsilon > 0$
- *when X is reflexive, weak and weak* topologies coincide*
- *Alaoglu* - unit ball $S^* = \{f \in X^* | \|f\| \geq 1\}$ is compact in weak* topology

Convex sets

- for vector space, X and $x, y \in X$

$$\{\lambda x + (1-\lambda)y | \lambda \in [0,1]\} \subset X$$

  called *segmenet joining x and y*
- set $K \subset X$ said to be *convex* or *convex set* if every segment joining any two points in K is in K, *i.e.*, $(\forall x, y \in K)$(segment joining $x, y \subset X$)
- every $\lambda x + (1-\lambda)y$ for $0 < \lambda < 1$ called *interior point of segment*
- point in $K \subset X$ where intersection with K of every line going through x contains open interval about x, said to be *internal point*, *i.e.*,

$$(\exists \epsilon > 0)(\forall y \in K, |\lambda| < \epsilon)(x + yx \in K)$$

- convex set examples - linear manifold & ball, ellipsoid in normed space

Properties of convex sets

- for convex sets, K_1 and K_2, following are also convex sets

$$K_1 \cap K_2,\ \lambda K_1,\ K_1 + K_2$$

- for linear operators from vector space, X, and vector space, Y,
  - image of convex set (or linear manifold) in X is convex set (or linear manifold) in Y,
  - inverse image of convex set (or linear manifold) in Y is convex set (or linear manifold) in X
- closure of convex set in topological vector space is convex set

Support functions of and separated convex sets

- for subset K of vector space X, $p : K \to \mathbf{R}_+$ with $p(x) = \inf \lambda | \lambda^{-1} x \in K, \lambda > 0$ called *support functions*
- for convex set $K \subset X$ containing 0 as internal point
  - $(\forall x \in X, \lambda \geq 0)(p(\lambda x) = \lambda p(x))$
  - $(\forall x, y \in X)(p(x+y) \leq p(x) + p(y))$
  - $\{x \in X | p(x) < 1\} \subset K \subset \{x \in X | p(x) \leq 1\}$
- two convex sets, K_1 and K_2 such that exists linear functional, f, and $\alpha \in \mathbf{R}$ with $(\forall x \in K_1)(f(x) \leq \alpha)$ and $(\forall x \in K_2)(f(x) \geq \alpha)$, said to be *separated*
- for two disjoint convex sets in vector space with at least one of them having internal point, exists *nonzero linear functional* that separates two sets

Local convexity

- topological vector space with base for topology consisting of convest sets, said to be *locally convex*
- for family of convex sets, $\mathcal{N}$, in vector space, following conditions are sufficient for being able to translate sets in $\mathcal{N}$ to form base for topology to make topological space into locally convex topological vector space

$$
\begin{gathered}
(\forall N \in \mathcal{N})(x \in N \Rightarrow x \text{ is internal}) \\
(\forall N_1, N_2 \in \mathcal{N})(\exists N_3 \in \mathcal{N})(N_3 \subset N_1 \cap N_2) \\
(\forall N \in \mathcal{N}, \alpha \in \mathbf{R} \text{ with } 0 < |\alpha| < 1)(\alpha N \in \mathcal{N})
\end{gathered}
$$

- conversely, for every locally convex topological vector space, exists base at 0 satisfying above conditions
- follows that
  - weak topology on vector space generated by linear functionals is locally convex
  - normed vector space is locally convex topological vector space

Facts regarding local convexity

- for locally convex topological vector space closed convex subset, F, with point, x, not in F, exists continuous linear functional, f, such that

$$
f(x) < \inf_{y \in F} f(y)
$$

- corollaries
  - convex set in locally convex topological vector space is strongly closed *if and only if* weakly closed
  - for distinct points, x and y, in locally convex Hausdorff vector space, exists continuous linear functional, f, such that $f(x) \neq f(y)$

Extreme points and supporting sets of convex sets

- point in convex set in vector space that is not interior point of any line segment lying in the set, called *extreme point*
- thus, x is extreme point of convex set, K, *if and only if* $x = \lambda y + (1-\lambda)z$ with $0 < \lambda < 1$ implies $y \notin K$ or $z \notin K$
- closed and convex subset, S, of convex set, K, with property that for every interior point of line segment in K belonging to S, entire line segment belongs to S, called *supporting set of K*
- for closed and convex set, K, set of points a continuous linear functional assumes maximum on K, is *supporting set of K*

Convex hull and convex convex hull

- for set E in vector space, intersection of all convex sets containing set, E, called *convex hull of E*, which is convex set
- for set E in vector space, intersection of all closed convex sets containing set, E, called *closed convex hull of E*, which is closed convex set
- *Krein-Milman theorem -* compact convex set in locally convex topologically vector space is *closed convex hull of its extreme points*

Hilbert spaces

- Banach space, H, with function $\langle\cdot,\cdot\rangle : H \times H \to \mathbf{R}$ satisfying following properties, called *Hilbert space*

$$
\begin{aligned}
&(\forall x, y, z \in H, \alpha, \beta \in \mathbf{R})(\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle)\\
&(\forall x, y \in H)(\langle x, y\rangle = \langle y, z\rangle)\\
&(\forall x \in H)(\langle x, x\rangle = \|x\|^2)
\end{aligned}
$$

- $\langle x, y\rangle$ called *inner product* for $x, y \in H$
  - examples - $\langle x, y\rangle = x^T y = \sum x_i y_i$ for $\mathbf{R}^n$, $\langle x, y\rangle = \int x(t)y(t)dt$ for L^2
- *Schwarz or Cauchy-Schwarz or Cauchy-Buniakowsky-Schwarz inequality -*

$$
\|x\|\|y\| \geq \langle x, y\rangle
$$

  - hence,
    - linear functional defined by $f(x) = \langle x, y\rangle$ bounded by $\|y\|$
    - $\langle x, y\rangle$ is continuous function from $H \times H$ to $\mathbf{R}$

Inner product in Hilbert spaces

- x and y in H with $\langle x, y\rangle = 0$ said to be *orthogonal* denoted by $x \perp y$
- set S of which any two elements orthogonal called *orthogonal system*
- orthogonal system called *orthonormal* if every element has unit norm
- any two elements are $\sqrt{2}$ apart, hence *if H separable, every orthonormal system in H must be countable*
- shall deal only with *separable Hilbert spaces*

Fourier coefficients

- assume orthonormal system expressed as sequence, $\langle \varphi_n \rangle$ - may be finite or infinite
- for $x \in H$

  $$a_n = \langle x, \varphi_n \rangle$$

  called *Fourier coefficients*
- for $n \in \mathbf{N}$, we have

  $$\|x\|^2 \geq \sum_{i=1}^{n} a_i^2$$

  Proof:

  $$\begin{aligned} \left\| x - \sum_{i=1}^{n} a_i \varphi_i \right\|^2 &= \left\langle x - \sum a_i \varphi_i, x - \sum a_i \varphi_i \right\rangle \\ &= \langle x, x \rangle - 2 \left\langle x, \sum a_i \varphi_i \right\rangle + \left\langle \sum a_i \varphi_i, \sum a_i \varphi_i \right\rangle \\ &= \|x\|^2 - 2 \sum a_i \langle x, \varphi_i \rangle + \sum a_i^2 \|\varphi_i\|^2 = \|x\|^2 - \sum a_i^2 \geq 0 \end{aligned}$$

Fourier coefficients of limit of x

- *Bessel's inequality* - for $x \in H$, its Fourier coefficients, $\langle a_n \rangle$

  $$\sum_{n=1}^{\infty} a_n^2 \leq \|x\|^2$$

- then, $\langle z_n \rangle$ defined by following is *Cauchy sequence* $z_n = \sum_{i=1}^{n} a_i \varphi_i$
- completeness (of Hilbert space) implies $\langle z_n \rangle$ converges - let $y = \lim z_n$

  $$y = \lim z_n = \sum_{i=1}^{\infty} a_i \varphi_i$$

- continuity of inner product implies $\langle y, \varphi_n \rangle = \lim(z_n, \varphi_n) = a_n$, *i.e.*, Fourier coefficients of $y \in H$ are a_n, *i.e.*,
- *y has same Fourier coefficients as x*

Complete orthonormal system

- orthonormal system, $\langle \varphi_n \rangle_{n=1}^{\infty}$, of Hilbert spaces, H, is said to be *complete* if

  $$(\forall x \in H, n \in \mathbf{N})(\langle x, \varphi_n \rangle = 0) \Rightarrow x = 0$$

- orthonormal system is complete *if and only if* maximal, *i.e.*,

  $$\langle \varphi_n \rangle \text{ is complete} \Leftrightarrow ((\exists \text{ orthonormal } R \subset H)(\forall n \in \mathbf{N})(\varphi_n \in R) \Rightarrow R = \langle \varphi_n \rangle)$$

  (proof can be found in Proof 15)
- Hausdorff maximal principle (Principle 4.4) implies existence of maximal orthonormal system, hence following statement

- for separable Hilbert space, H, every orthonormal system is separable and exists *a* complete orthonormal system. any such system, $\langle \varphi_n \rangle$, and $x \in H$

$$x = \sum a_n \varphi_n$$

with $a_n = \langle x, \varphi_n \rangle$, and $\|x\| = \sum a_n^2$

Dimensions of Hilbert spaces

- every complete orthonormal system of separable Hilbert space has same number of elements, *i.e.*, has same cardinality
- hence, every complete orthonormal system has either finite or countably infinite complete orthonormal system
- this number called *dimension of separable Hilbert space*
  - for Hilbert space with countably infinite complete orthonormal system, we say, $\dim H = \aleph_0$

Isomorphism and isometry between Hilbert spaces

- *isomorphism, Φ, of Hilbert space onto another Hilbert space* is linear mapping with property, $\langle \Phi x, \Phi y \rangle = \langle x, y \rangle$
- hence, every *isomorphism between Hilbert spaces is isometry*
- every n-dimensional Hilbert space is isomorphic to $\mathbf{R}^n$
- every $\aleph_0$-dimensional Hilbert space is isomorphic to l^2, which again is isomorphic to L^2
- $L^2[0,1]$ is separable and $\langle \cos(n\pi t) \rangle$ is infinite orthogonal system
- every bounded linear functional, f, on Hilbert space, H, has unique y such that

$$(\forall x \in H)(f(x) = \langle x, y \rangle)$$

and $\|f\| = \|y\|$

4.12 Measure and Integration

Purpose of integration theory

- purpose of "measure and integration" slides
  - abstract (out) most important properties of Lebesgue measure and Lebesgue integration
- provide certain *axioms that Lebesgue measure satisfies*
- base our integration theory on these axioms
- hence, our theory valid for every system satisfying the axioms

Measurable space, measure, and measure space

- family of subsets containing $\emptyset$ closed under countable union and completement, called *σ-algebra*
- mapping of sets to extended real numbers, called *set function*
- $(X, \mathscr{B})$ with set, X, and σ-algebra of X, $\mathscr{B}$, called *measurable space*
  - $A \in \mathscr{B}$, said to be *measurable (with respect to $\mathscr{B}$)*
- nonnegative set function, μ, defined on $\mathscr{B}$ satisfying $\mu(\emptyset) = 0$ and for every disjoint, $\langle E_n \rangle_{n=1}^{\infty} \subset \mathscr{B}$,

  $$\mu\left(\bigcup E_n\right) = \sum \mu E_n$$

  called *measure on* measurable space, $(X, \mathscr{B})$
- measurable space, $(X, \mathscr{B})$, equipped with measure, μ, called *measure space* and denoted by $(X, \mathscr{B}, \mu)$

Measure space examples

- $(\mathbf{R}, \mathcal{M}, \mu)$ with Lebesgue measurable sets, $\mathcal{M}$, and Lebesgue measure, μ
- $([0,1], \{A \in \mathcal{M} | A \subset [0,1]\}, \mu)$ with Lebesgue measurable sets, $\mathcal{M}$, and Lebesgue measure, μ
- $(\mathbf{R}, \mathscr{B}, \mu)$ with class of Borel sets, $\mathscr{B}$, and Lebesgue measure, μ
- $(\mathbf{R}, \mathcal{P}(\mathbf{R}), \mu_C)$ with set of all subsets of $\mathbf{R}$, $\mathcal{P}(\mathbf{R})$, and counting measure, μ_C
- interesting (and bizarre) example
  - $(X, \mathcal{A}, \mu_B)$ with any uncountable set, X, family of either countable or complement of countable set, $\mathcal{A}$, and measure, μ_B, such that $\mu_B A = 0$ for countable $A \subset X$ and $\mu_B B = 1$ for uncountable $B \subset X$

More properties of measures

- for $A, B \in \mathscr{B}$ with $A \subset B$

  $$\mu A \leq \mu B$$

- for $\langle E_n \rangle \subset \mathscr{B}$ with $\mu E_1 < \infty$ and $E_{n+1} \subset E_n$

  $$\mu\left(\bigcap E_n\right) = \lim \mu E_n$$

- for $\langle E_n \rangle \subset \mathscr{B}$

  $$\mu\left(\bigcup E_n\right) \leq \sum \mu E_n$$

Finite and σ-finite measures

- measure, μ, with $\mu(X) < \infty$, called *finite*
- measure, μ, with $X = \bigcup X_n$ for some $\langle X_n \rangle$ and $\mu(X_n) < \infty$, called *σ-finite*
  - always can take $\langle X_n \rangle$ with disjoint X_n
- Lebesgue measure on $[0,1]$ is finite
- Lebesgue measure on $\mathbf{R}$ is σ-finite
- countering measure on uncountable set is *not* σ-measure

Sets of finite and σ-finite measure

- set, $E \in \mathscr{B}$, with $\mu E < \infty$, said to be *of finite measure*
- set that is countable union of measurable sets of finite measure, said to be *of σ-finite measure*
- measurable set contained in set of σ-finite measure, is of σ-finite measure
- countable union of sets of σ-finite measure, is of σ-finite measure
- when μ is σ-finite, every measurable set is of σ-finite

Semifinite measures

- roughly speacking, nearly all familiar properties of Lebesgue measure and Lebesgue integration hold for arbitrary σ-finite measure
- many treatment of abstract measure theory limit themselves to σ-finite measures
- many parts of general theory, however, do *not* required assumption of σ-finiteness
- undesirable to have development unnecessarily restrictive
- measure, μ, for which every measurable set of infinite measure contains measurable sets of arbitrarily large finite measure, said to be *semifinite*
- every σ-finite measure is semifinite measure while measure, μ_B, on page 118 is not

Complete measure spaces

- measure space, $(X, \mathscr{B}, \mu)$, for which $\mathscr{B}$ contains all subsets of sets of measure zero, said to be *complete*, *i.e.*,
$$(\forall B \in \mathscr{B} \text{ with } \mu B = 0)(A \subset B \Rightarrow A \in \mathscr{B})$$
  - *e.g.*, Lebesgue measure is complete, but Lebesgue measure restricted to σ-algebra of Borel sets is *not*
- every measure space can be *completed* by addition of subsets of sets of measure zero
- for $(X, \mathscr{B}, \mu)$, can find *complete* measure space $(X, \mathscr{B}_0, \mu_0)$ such that
  - $\mathscr{B} \subset \mathscr{B}_0$
  - $E \in \mathscr{B} \Rightarrow \mu E = \mu_0 E$
  - $E \in \mathscr{B}_0 \Leftrightarrow E = A \cup B$ where $B, C \in \mathscr{B}, \mu C = 0, A \subset C$
  - $(X, \mathscr{B}_0, \mu_0)$ called *completion* of $(X, \mathscr{B}, \mu)$

Local measurability and saturatedness

- for $(X, \mathscr{B}, \mu)$, $E \subset X$ for which $(\forall B \in \mathscr{B} \text{ with } \mu B < \infty)(E \cap B \in \mathscr{B})$, said to be *locally measurable*
- collection, $\mathscr{C}$, of all locally measurable sets is σ-algebra containing $\mathscr{B}$
- measure for which every locally measurable set is measurable, said to be *saturated*
- every σ-finite measure is saturated
- measure can be extended to saturated measure, but (unlike completion) extension is not unique
  - can take $\mathscr{C}$ as extension for locally measurable sets, but measure can be extended on $\mathscr{C}$ in more than one ways

Measurable functions

- concept and properties of measurable functions in abstract measurable space almost identical with those of Lebesgue measurable functions (page 75)
- theorems and facts are essentially same as those of Lebesgue measurable functions
- assume measurable space, $(X, \mathscr{B})$
- for $f : X \to \mathbf{R} \cup \{-\infty, \infty\}$, following are equivalent
  - $(\forall a \in \mathbf{R})(\{x \in X | f(x) < a\} \in \mathscr{B})$
  - $(\forall a \in \mathbf{R})(\{x \in X | f(x) \leq a\} \in \mathscr{B})$
  - $(\forall a \in \mathbf{R})(\{x \in X | f(x) > a\} \in \mathscr{B})$
  - $(\forall a \in \mathbf{R})(\{x \in X | f(x) \geq a\} \in \mathscr{B})$
- $f : X \to \mathbf{R} \cup \{-\infty, \infty\}$ for which any one of above four statements holds, called *measurable* or *measurable with respect to $\mathscr{B}$*

(refer to page 75 for Lebesgue counterpart)

Properties of measurable functions

- **Theorem 4.1 (measurability preserving function operations)** *for measurable functions, f and g, and $c \in \mathbf{R}$*
  - *$f + c$, cf, $f + g$, fg, $f \vee g$ are measurable*
- **Theorem 4.2 (limits of measurable functions)** *for every measurable function sequence, $\langle f_n \rangle$*
  - *$\sup f_n$, $\limsup f_n$, $\inf f_n$, $\liminf f_n$ are measurable*
  - *thus, $\lim f_n$ is measurable if exists*

(refer to page 76 for Lebesgue counterpart)

Simple functions and other properties

- φ called *simple function* if for distinct $\langle c_i \rangle_{i=1}^n$ and measurable sets, $\langle E_i \rangle_{i=1}^n$

$$\varphi(x) = \sum_{i=1}^{n} c_i \chi_{E_i}(x)$$

(refer to page 76 for Lebesgue counterpart)

- for nonnegative measurable function, f, exists nondecreasing sequence of simple functions, $\langle \varphi_n \rangle$, *i.e.*, $\varphi_{n+1} \geq \varphi_n$ such that for every point in X

$$f = \lim \varphi_n$$

  - for f defined on σ-finite measure space, we may choose $\langle \varphi_n \rangle$ so that every φ_n vanishes outside set of finite measure
- for complete measure, μ, f measurable and $f = g$ a.e. imply measurability of g

Define measurable function by ordinate sets

- $\{x|f(x) < \alpha\}$ sometimes called *ordinate sets*, which is nondecreasing in α
- below says when given nondecreasing ordinate sets, we can find f satisfying

$$\{x|f(x) < \alpha\} \subset B_\alpha \subset \{x|f(x) \leq \alpha\}$$

- for nondecreasing function, $h : D \to \mathscr{B}$, for dense set of real numbers, D, *i.e.*, $B_\alpha \subset B_\beta$ for all $\alpha < \beta$ where $B_\alpha = h(\alpha)$, exists unique measurable function, $f : X \to \mathbf{R} \cup \{-\infty, \infty\}$ such that $f \leq \alpha$ on B_α and $f \geq \alpha$ on $X \sim B_\alpha$
- can relax some conditions and make it a.e. version as below
- for function, $h : D \to \mathscr{B}$, for dense set of real numbers, D, such that $\mu(B_\alpha \sim B_\beta) = 0$ for all $\alpha < \beta$ where $B_\alpha = h(\alpha)$, exists measurable function, $f : X \to \mathbf{R} \cup \{-\infty, \infty\}$ such that $f \leq \alpha$ a.e. on B_α and $f \geq \alpha$ a.e. on $X \sim B_\alpha$
  - if g has the same property, $f = g$ a.e.

Integration

- many definitions and proofs of Lebesgue integral depend only on properties of Lebesgue measure which are also true for arbitrary measure in abstract measure space (page 77)
- integral of nonnegative simple function, $\varphi(x) = \sum_{i=1}^n c_i \chi_{E_i}(x)$, on measurable set, E, defined by

$$\int_E \varphi d\mu = \sum_{i=1}^n c_i \mu(E_i \cap E)$$

  - independent of representation of φ

(refer to page 77 for Lebesgue counterpart)

- for $a, b \in \mathbf{R}_{++}$ and nonnegative simple functions, φ and ψ

$$\int (a\varphi + b\psi) = a \int \varphi + b \int \psi$$

(refer to page 78 for Lebesgue counterpart)

Integral of bounded functions

- for bounded function, f, identically zero outside measurable set of finite measure

$$\sup_{\varphi:\ \text{simple},\ \varphi \leq f} \int \varphi = \inf_{\psi:\ \text{simple},\ f \leq \psi} \int \psi$$

  if and only if $f = g$ a.e. for measurable function, g

(refer to page 78 for Lebesgue counterpart)

- but, *$f = g$ a.e. for measurable function, g,* if and only if *f is measurable with respect to completion of μ, $\bar{\mu}$*
- *natural class of functions to consider for integration theory are those measurable with respect to completion of μ*
- thus, shall either assume μ is complete measure or define integral with respect to μ to be integral with respect to completion of μ depending on context unless otherwise specified

Difficulty of general integral of nonnegative functions

- for Lebesgue integral of nonnegative functions (page 79)
  - first define integral for bounded measurable functions
  - define integral of nonnegative function, f as supremum of integrals of all bounded measurable functions, $h \leq f$, vanishing outside measurable set of finite measure
- unfortunately, not work in case that measure is not semifinite
  - *e.g.*, if $\mathscr{B} = \{\emptyset, X\}$ with $\mu\emptyset = 0$ and $\mu X = \infty$, we want $\int 1 d\mu = \infty$, but only bounded measurable function vanishing outside measurable set of finite measure is $h \equiv 0$, hence, $\int g d\mu = 0$
- to avoid this difficulty, we define integral of nonnegative measurable function directly in terms of integrals of nonnegative simple functions

Integral of nonnegative functions

- for measurable function, $f : X \to \mathbf{R} \cup \{\infty\}$, on measure space, $(X, \mathscr{B}, \mu)$, define *integral of nonnegative extended real-valued measurable function*

$$\int f d\mu = \sup_{\varphi\text{: simple function, } 0 \leq \varphi \leq f} \int \varphi d\mu$$

(refer to page 79 for Lebesgue counterpart)

- however, *definition of integral of nonnegative extended real-valued measurable function* can be awkward to apply because
  - taking supremum over large collection of simple functions
  - *not clear from definition that* $\int (f+g) = \int f + \int g$
- thus, first establish some convergence theorems, and determine value of $\int f$ as limit of $\int \varphi_n$ for increasing sequence, $\langle \varphi_n \rangle$, of simple functions converging to f

Fatou's lemma and monotone convergence theorem

- *Fatou's lemma* - for nonnegative measurable function sequence, $\langle f_n \rangle$, with $\lim f_n = f$ a.e. on measurable set, E

$$\int_E f \leq \liminf \int_E f_n$$

- *monotone convergence theorem* - for nonnegative measurable function sequence, $\langle f_n \rangle$, with $f_n \leq f$ for all n and with $\lim f_n = f$ a.e.

$$\int_E f = \lim \int_E f_n$$

(refer to page 80 for Lebesgue counterpart)

Integrability of nonnegative functions

- for nonnegative measurable functions, f and g, and $a, b \in \mathbf{R}_+$
$$\int (af + bg) = a \int f + b \int g \;\&\; \int f \geq 0$$
  - equality holds *if and only if* $f = 0$ a.e.

(refer to page 78 for Lebesgue counterpart)

- monotone convergence theorem together with above yields, for nonnegative measurable function sequence, $\langle f_n \rangle$
$$\int \sum f_n = \sum \int f_n$$
- measurable nonnegative function, f, with
$$\int_E f d\mu < \infty$$
said to be *integral (over measurable set, E, with respect to μ)*

(refer to page 80 for Lebesgue counterpart)

Integral

- arbitrary function, f, for which both f^+ and f^- are integrable, said to be *integrable*
- in this case, define *integral*
$$\int_E f = \int_E f^+ - \int_E f^-$$

(refer to page 80 for Lebesgue counterpart)

Properties of integral

- for f and g integrable on measure set, E, and $a, b \in \mathbf{R}$
  - $af + bg$ is integral and
$$\int_E (af + bg) = a \int_E f + b \int_E g$$
  - if $|h| \leq |f|$ and h is measurable, then h is integrable
  - if $f \geq g$ a.e.
$$\int f \geq \int g$$

(refer to page 81 for Lebesgue counterpart)

Lebesgue convergence theorem

- *Lebesgue convergence theorem* - for integral, g, over E and sequence of measurable functions, $\langle f_n \rangle$, with $\lim f_n(x) = f(x)$ a.e. on E, if
$$|f_n(x)| \leq g(x)$$
then
$$\int_E f = \lim \int_E f_n$$

(refer to page 81 for Lebesgue counterpart)

Setwise convergence of sequence of measures

- preceding convergence theorems assume fixed measure, μ
- can generalize by allowing measure to vary
- given measurable space, $(X, \mathscr{B})$, sequence of set functions, $\langle \mu_n \rangle$, defined on $\mathscr{B}$, satisfying

$$(\forall E \in \mathscr{B})(\lim \mu_n E = \mu E)$$

  for some set function, μ, defined on $\mathscr{B}$, said to *converge setwise* to μ

General convergence theorems

- *generalization of Fatou's leamma* - for measurable space, $(X, \mathscr{B})$, sequence of measures, $\langle \mu_n \rangle$, defined on $\mathscr{B}$, converging setwise to μ, defined on $\mathscr{B}$, and sequence of nonnegative functions, $\langle f_n \rangle$, each measurable with respect to μ_n, converging pointwise to function, f, measurable with respect to μ (compare with Fatou's lemma on page 122)

$$\int f d\mu \leq \liminf \int f_n d\mu_n$$

- *generalization of Lebesgue convergence theorem* - for measurable space, $(X, \mathscr{B})$, sequence of measures, $\langle \mu_n \rangle$, defined on $\mathscr{B}$, converging setwise to μ, defined on $\mathscr{B}$, and sequences of functions, $\langle f_n \rangle$ and $\langle g_n \rangle$, each of f_n and g_n, measurable with respect to μ_n, converging pointwise to f and g, measurable with respect to μ, respectively, such that (compare with Lebesgue convergence theorem on page 123)

$$\lim \int g_n d\mu_n = \int g d\mu < \infty$$

  satisfy

$$\lim \int f_n d\mu_n = \int f \mu$$

L^p spaces

- for complete measure space, $(X, \mathscr{B}, \mu)$
  - space of measurable functions on X with with $\int |f|^p < \infty$, for which element equivalence is defined by being equal a.e., called *L^p spaces* denoted by $L^p(\mu)$
  - space of bounded measure functions, called *L^∞ space* denoted by $L^\infty(\mu)$
- norms
  - for $p \in [1, \infty)$

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{1/p}$$

  - for $p = \infty$

$$\|f\|_\infty = \operatorname{ess\,sup} |f| = \inf \{ |g(x)| | \text{ measurable } g \text{ with } g = f \text{ a.e.} \}$$

- for $p \in [1, \infty]$, spaces, $L^p(\mu)$, are Banach spaces

Hölder's inequality and Littlewood's second principle

- *Hölder's inequality* - for $p, q \in [1, \infty]$ with $1/p + 1/q = 1$, $f \in L^p(\mu)$ and $g \in L^q(\mu)$ satisfy $fg \in L^1(\mu)$ and

$$\|fg\|_1 = \int |fg| d\mu \leq \|f\|_p \|g\|_q$$

(refer to page 84 for normed spaces counterpart)

- *complete measure space version of Littlewood's second principle* - for $p \in [1, \infty)$

$$
\begin{aligned}
&(\forall f \in L^p(\mu), \epsilon > 0)\\
&\quad (\exists \text{ simple function } \varphi \text{ vanishing outside set of finite measure})\\
&\qquad (\|f - \varphi\|_p < \epsilon)
\end{aligned}
$$

(refer to page 85 for normed spaces counterpart)

Riesz representation theorem

- *Riesz representation theorem* - for $p \in [1, \infty)$ and bounded linear functional, F, on $L^p(\mu)$ and σ-finite measure, μ, exists *unique* $g \in L^q(\mu)$ where $1/p + 1/q = 1$ such that

$$F(f) = \int fg d\mu$$

  where $\|F\| = \|g\|_q$

(refer to page 85 for normed spaces counterpart)

- if $p \in (1, \infty)$, Riesz representation theorem holds without assumption of σ-finiteness of measure

4.13 Measure and Outer Measure

General measures

- consider some ways of defining measures on σ-algebra
- recall that for Lebesgue measure
  - define measure for open intervals
  - define outer measure
  - define notion of measurable sets
  - finally derive Lebesgue measure
- one can do similar things in general, *e.g.*,
  - derive measure from outer measure
  - derive outer measure from measure defined on algebra of sets

Outer measure

- set function, $\mu^* : \mathcal{P}(X) \to [0, \infty]$, for space X, having following properties, called *outer measure*
  - $\mu^* \emptyset = 0$
  - $A \subset B \Rightarrow \mu^* A \leq \mu^* B$ (monotonicity)
  - $E \subset \bigcup_{n=1}^\infty E_n \Rightarrow \mu^* E \leq \sum_{n=1}^\infty \mu^* E_n$ (countable subadditivity)
- μ^* with $\mu^* X < \infty$ called *finite*
- set $E \subset X$ satisfying following property, said to be *measurable with respect to μ^**

$$(\forall A \subset X)(\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap \tilde{E}))$$

- class, $\mathscr{B}$, of μ^*-measurable sets is σ-algebra
- restriction of μ^* to $\mathscr{B}$ is complete measure on $\mathscr{B}$

Extension to measure from measure on an algebra

- set function, $\mu : \mathscr{A} \to [0, \infty]$, defined on algebra, $\mathscr{A}$, having following properties, called *measure on an algebra*
  - $\mu(\emptyset) = 0$
  - $(\forall$ disjoint $\langle A_n \rangle \subset \mathscr{A}$ with $\bigcup A_n \in \mathscr{A})\left(\mu\left(\bigcup A_n\right) = \sum \mu A_n\right)$
- *measure on an algebra, $\mathscr{A}$*, is measure *if and only if* $\mathscr{A}$ is σ-algebra
- can extend measure on an algebra to measure defined on σ-algebra, $\mathscr{B}$, containing $\mathscr{A}$, by
  - constructing outer measure μ^* from μ
  - deriving desired extension $\bar{\mu}$ induced by μ^*
- process by which constructing μ^* from μ similar to constructing Lebesgue outer measure from lengths of intervals

Outer measure constructed from measure on an algebra

— given measure, μ, on an algebra, $\mathscr{A}$

- define set function, $\mu^* : \mathcal{P}(X) \to [0, \infty]$, by

$$\mu^* E = \inf_{\langle A_n \rangle \subset \mathscr{A},\ E \subset \bigcup A_n} \sum \mu A_n$$

- μ^* called *outer measure induced by μ*

— then

- for $A \in \mathscr{A}$ and $\langle A_n \rangle \subset \mathscr{A}$ with $A \subset \bigcup A_n$, $\mu A \leq \sum \mu A_n$
- hence, $(\forall A \in \mathscr{A})(\mu^* A = \mu A)$
- μ^* is outer measure
- every $A \in \mathscr{A}$ is measurable with respect to μ^*

Regular outer measure

- for algebra, $\mathscr{A}$
  - $\mathscr{A}_\sigma$ denote sets that are countable unions of sets of $\mathscr{A}$
  - $\mathscr{A}_{\sigma\delta}$ denote sets that are countable intersections of sets of $\mathscr{A}_\sigma$
- given measure, μ, on an algebra, $\mathscr{A}$ and outer measure, μ^* induced by μ, for every $E \subset X$ and every $\epsilon > 0$, exists $A \in \mathscr{A}_\sigma$ and $B \in \mathscr{A}_{\sigma\delta}$ with $E \subset A$ and $E \subset B$

$$\mu^* A \leq \mu^* E + \epsilon \text{ and } \mu^* E = \mu^* B$$

- outer measure, μ^*, with below property, said to be *regular*

$$(\forall E \subset X, \epsilon > 0)(\exists\ \mu^*\text{-measurable set } A \text{ with } E \subset A)(\mu^* A \subset \mu^* E + \epsilon)$$

- every outer measure induced by measure on an algebra is regular outer measure

Carathéodory theorem

— given measure, μ, on an algebra, $\mathscr{A}$ and outer measure, μ^* induced by μ

- $E \subset X$ is μ^*-measurable *if and only if* exist $A \in \mathscr{A}_{\sigma\delta}$ and $B \subset X$ with $\mu^* B = 0$ such that

$$E = A \sim B$$

  - for $B \subset X$ with $\mu^* B = 0$, exists $C \in \mathscr{A}_{\sigma\delta}$ with $\mu^* C = 0$ such that $B \subset C$
- *Carathéodory theorem -* restriction, $\bar{\mu}$, of μ^* to μ^*-measurable sets if extension of μ to σ-algebra containing $\mathscr{A}$
  - if μ is finite or σ-finite, so is $\bar{\mu}$ respectively
  - if μ is σ-finite, $\bar{\mu}$ is only measure on smallest σ-algebra containing $\mathscr{A}$ which is extension of μ

Product measures

- for countable disjoint collection of measurable rectangles, $\langle (A_n \times B_n) \rangle$, whose union is measurable rectangle, $A \times B$

$$\lambda(A \times B) = \sum \lambda(A_n \times B_n)$$

- for $x \in X$ and $E \in \mathscr{R}_{\sigma\delta}$

$$E_x = \{y | \langle x, y \rangle \in E\}$$

  is measurable subset of Y
- for $E \subset \mathscr{R}_{\sigma\delta}$ with $\mu \times \nu(E) < \infty$, function, g, defined by

$$g(x) = \nu E_x$$

  is measurable function of x and

$$\int g d\mu = \mu \times \nu(E)$$

- XXX

Carathéodory outer measures

- set, X, of points and set, Γ, of real-valued functions on X
- two sets for which exist $a > b$ such that function, φ, greater than a on one set and less than b on the other set, said to be *separated by function, φ*
- outer measure, μ^*, with $(\forall A, B \subset X$ separated by $f \in \Gamma)(\mu^*(A \cup B) = \mu^* A + \mu^* B)$, called *Carathéodory outer measure with respect to Γ*
- outer measure, μ^*, on metric space, $\langle X, \rho, , \rangle$ for which $\mu^*(A \cup B) = \mu^* A + \mu^* B$ for $A, B \subset X$ with $\rho(A, B) > 0$, called *Carathéodory outer measure for X* or *metric outer measure*
- for *Carathéodory outer measure, μ^*, with respect to* Γ, every function in Γ is μ^*-measurable
- for *Carathéodory outer measure, μ^*, for metric space,* $\langle X, \rho, \rangle$, every closed set (hence every Borel set) is measurable with respect to μ^*

5 Measure-theoretic Treatment of Probabilities

5.1 Probability Measure

Measurable functions

- denote *n-dimensional Borel sets* by $\mathscr{R}^n$

- for two measurable spaces, $(\Omega, \mathscr{F})$ and $(\Omega', \mathscr{F}')$, function, $f : \Omega \to \Omega'$ with

$$(\forall A' \in \mathscr{F}') \left(f^{-1}(A') \in \mathscr{F} \right)$$

  said to be *measurable with respect to $\mathscr{F}/\mathscr{F}'$* (thus, measurable functions defined on page 75 and page 120 can be said to be measurable with respect to $\mathcal{B}/\mathscr{R}$)

- when $\Omega = \mathbf{R}^n$ in $(\Omega, \mathscr{F})$, $\mathscr{F}$ is assumed to be $\mathscr{R}^n$, and sometimes drop $\mathscr{R}^n$
  - thus, *e.g.*, we say $f : \Omega \to \mathbf{R}^n$ is measurable with respect to $\mathscr{F}$ (instead of $\mathscr{F}/\mathscr{R}^n$)

- measurable function, $f : \mathbf{R}^n \to \mathbf{R}^m$ (*i.e.*, measurable with respect to $\mathscr{R}^n/\mathscr{R}^m$), called *Borel functions*

- $f : \Omega \to \mathbf{R}^n$ is measurable with respect to $\mathscr{F}/\mathscr{R}^n$ *if and only if* every component, $f_i : \Omega \to \mathbf{R}$, is measurable with respect to $\mathscr{F}/\mathscr{R}$

Probability (measure) spaces

- set function, $P : \mathscr{F} \to [0, 1]$, defined on algebra, $\mathscr{F}$, of set Ω, satisfying following properties, called *probability measure* (refer to page 118 for resumblance with measurable spaces)
  - $(\forall A \in \mathscr{F})(0 \leq P(A) \leq 1)$
  - $P(\emptyset) = 0, \ P(\Omega) = 1$
  - $(\forall \text{ disjoint } \langle A_n \rangle \subset \mathscr{F})(P(\bigcup A_n) = \sum P(A_n))$

- for σ-algebra, $\mathscr{F}$, $(\Omega, \mathscr{F}, P)$, called *probability measure space* or *probability space*

- set $A \in \mathscr{F}$ with $P(A) = 1$, called *a support of P*

Dynkin's π-λ theorem

- class, $\mathcal{P}$, of subsets of Ω closed under finite intersection, called *π-system*, *i.e.*,
  - $(\forall A, B \in \mathcal{P})(A \cap B \in \mathcal{P})$

- class, $\mathcal{L}$, of subsets of Ω containing Ω closed under complements and countable disjoint unions called *λ-system*
  - $\Omega \in \mathcal{L}$
  - $(\forall A \in \mathcal{L})(\tilde{A} \in \mathcal{L})$
  - $(\forall \text{ disjoint } \langle A_n \rangle)(\bigcup A_n \in \mathcal{L})$

- *class that is both π-system and λ-system is σ-algebra*

- *Dynkin's π-λ theorem* - for π-system, $\mathcal{P}$, and λ-system, $\mathcal{L}$, with $\mathcal{P} \subset \mathcal{L}$,

$$\sigma(\mathcal{P}) \subset \mathcal{L}$$

- for π-system, $\mathscr{P}$, two probability measures, P_1 and P_2, on $\sigma(\mathscr{P})$, agreeing $\mathscr{P}$, agree on $\sigma(\mathscr{P})$

Limits of Events

Theorem 5.1 (convergence-of-events) *no for sequence of subsets, $\langle A_n \rangle$,*

$$P(\liminf A_n) \leq \liminf P(A_n) \leq \limsup P(A_n) \leq P(\limsup A_n)$$

- *for $\langle A_n \rangle$ converging to A*

$$\lim P(A_n) = P(A)$$

Theorem 5.2 (independence-of-smallest-sig-alg) *no for sequence of π-systems, $\langle \mathscr{A}_n \rangle$, $\langle \sigma(\mathscr{A}_n) \rangle$ is independent*

Probabilistic independence

– given probability space, $(\Omega, \mathscr{F}, P)$

- $A, B \in \mathscr{F}$ with

$$P(A \cap B) = P(A)P(B)$$

  said to be *independent*

- indexed collection, $\langle A_\lambda \rangle$, with

$$(\forall n \in \mathbf{N},\ \text{distinct } \lambda_1, \ldots, \lambda_n \in \Lambda) \left(P\left(\bigcap_{i=1}^{n} A_{\lambda_i} \right) = \prod_{i=1}^{n} P(A_{\lambda_i}) \right)$$

  said to be *independent*

Independence of classes of events

- indexed collection, $\langle \mathcal{A}_\lambda \rangle$, of classes of events (*i.e.*, subsets) with

$$(\forall A_\lambda \in \mathcal{A}_\lambda)\,(\langle A_\lambda \rangle \text{ are independent})$$

  said to be *independent*

- *for independent indexed collection, $\langle \mathcal{A}_\lambda \rangle$, with every $\mathcal{A}_\lambda$ being π-sytem, $\langle \sigma(\mathcal{A}_\lambda) \rangle$ are independent*

- for independent (countable) collection of events, $\langle \langle A_{ni} \rangle_{i=1}^{\infty} \rangle_{n=1}^{\infty}$, $\langle \mathscr{F}_n \rangle_{n=1}^{\infty}$ with $\mathscr{F}_n = \sigma(\langle A_{ni} \rangle_{i=1}^{\infty})$ are independent

Borel-Cantelli lemmas

- **Lemma 5.1 (first Borel-Cantelli)** *for sequence of events, $\langle A_n \rangle$, with $\sum P(A_n)$ converging*

$$P(\limsup A_n) = 0$$

- **Lemma 5.2 (second Borel-Cantelli)** *for independent sequence of events, $\langle A_n \rangle$, with $\sum P(A_n)$ diverging*

$$P(\limsup A_n) = 1$$

Tail events and Kolmogorov's zero-one law

- for sequence of events, $\langle A_n \rangle$

$$\mathscr{T} = \bigcap_{n=1}^{\infty} \sigma\left(\langle A_i \rangle_{i=n}^{\infty}\right)$$

  called *tail σ-algebra associated with $\langle A_n \rangle$*; its lements are called *tail events*

- *Kolmogorov's zero-one law* - for independent sequence of events, $\langle A_n \rangle$ every event in tail σ-algebra has probability measure either 0 or 1

Product probability spaces

- for two measure spaces, $(X, \mathscr{X}, \mu)$ and $(Y, \mathscr{Y}, \nu)$, want to find product measure, π, such that

$$(\forall A \in \mathscr{X}, B \in \mathscr{Y})\, (\pi(A \times B) = \mu(A)\nu(B))$$

  - *e.g.*, if both μ and ν are Lebesgue measure on $\mathbf{R}$, π will be Lebesgue measure on $\mathbf{R}^2$

- $A \times B$ for $A \in \mathscr{X}$ and $B \in \mathscr{Y}$ is *measurable rectangle*
- *σ-algebra generated by measurable rectangles* denoted by

$$\mathscr{X} \times \mathscr{Y}$$

  - thus, *not* Cartesian product in usual sense
  - generally *much larger* than class of measurable rectangles

Sections of measurable subsets and functions

for two measure spaces, $(X, \mathscr{X}, \mu)$ and $(Y, \mathscr{Y}, \nu)$

- sections of measurable subsets
  - $\{y \in Y | (x, y) \in E\}$ is *section of E determined by x*
  - $\{x \in X | (x, y) \in E\}$ is *section of E determined by y*
- sections of measurable functions - for measurable function, f, with respect to $\mathscr{X} \times \mathscr{Y}$
  - $f(x, \cdot)$ is *section of f determined by x*
  - $f(\cdot, y)$ is *section of f determined by y*
- sections of measurable subsets are measurable
  - $(\forall x \in X, E \in \mathscr{X} \times \mathscr{Y})\, (\{y \in Y | (x, y) \in E\} \in \mathscr{Y})$
  - $(\forall y \in Y, E \in \mathscr{X} \times \mathscr{Y})\, (\{x \in X | (x, y) \in E\} \in \mathscr{X})$
- sections of measurable functions are measurable
  - $f(x, \cdot)$ is measurable with respect to $\mathscr{Y}$ for every $x \in X$
  - $f(\cdot, y)$ is measurable with respect to $\mathscr{X}$ for every $y \in Y$

Product measure

for two σ-finite measure spaces, $(X, \mathscr{X}, \mu)$ and $(Y, \mathscr{Y}, \nu)$

- two functions defined below for every $E \in \mathscr{X} \times \mathscr{Y}$ are σ-finite measures
  - $\pi'(E) = \int_X \nu\{y \in Y | (x, y) \in E\} d\mu$
  - $\pi''(E) = \int_Y \mu\{x \in X | (x, y) \in E\} d\nu$
- for every measurable rectangle, $A \times B$, with $A \in \mathscr{X}$ and $B \in \mathscr{Y}$

$$\pi'(A \times B) = \pi''(A \times B) = \mu(A)\nu(B)$$

  (use conventions in page 15 for extended real values)
- indeed, $\pi'(E) = \pi''(E)$ for every $E \in \mathscr{X} \times \mathscr{Y}$; let $\pi = \pi' = \pi''$
- π is
  - called *product measure* and denoted by $\mu \times \nu$
  - σ-finite measure
  - *only* measure such that $\pi(A \times B) = \mu(A)\nu(B)$ for every measurable rectangle

Fubini's theorem

- suppose two σ-finite measure spaces, $(X, \mathscr{X}, \mu)$ and $(Y, \mathscr{Y}, \nu)$ - define
  - $X_0 = \{x \in X | \int_Y |f(x, y)| d\nu < \infty\} \subset X$
  - $Y_0 = \{y \in Y | \int_X |f(x, y)| d\nu < \infty\} \subset Y$
- *Fubini's theorem* - for nonnegative measurable function, f, following are measurable with respect to $\mathscr{X}$ and $\mathscr{Y}$ respectively

$$g(x) = \int_Y f(x, y) d\nu, \quad h(y) = \int_X f(x, y) d\mu$$

  and following holds

$$\int_{X \times Y} f(x, y) d\pi = \int_X \left(\int_Y f(x, y) d\nu \right) d\mu = \int_Y \left(\int_X f(x, y) d\mu \right) d\nu$$

- for f, (not necessarily nonnegative) integrable function with respect to π
  - $\mu(X \sim X_0) = 0$, $\nu(Y \sim Y_0) = 0$
  - g and h are finite measurable on X_0 and Y_0 respectively
  - (above) equalities of *double integral* holds

5.2 Random Variables

Random variables

- for probability space, $(\Omega, \mathscr{F}, P)$,

- measurable function (with respect to $\mathscr{F}/\mathscr{R}$), $X : \Omega \to \mathbf{R}$, called *random variable*
- measurable function (with respect to $\mathscr{F}/\mathscr{R}^n$), $X : \Omega \to \mathbf{R}^n$, called *random vector*
  - when expressing $X(\omega) = (X_1(\omega), \ldots, X_n(\omega))$, X is measurable *if and only if* every X_i is measurable
  - thus, n-dimensional random vaector is simply n-tuple of random variables
- smallest σ-algebra with respect to which X is measurable, called *σ-algebra generated by X* and denoted by $\sigma(X)$
  - $\sigma(X)$ consists exactly of sets, $\{\omega \in \Omega | X(\omega) \in H\}$, for $H \in \mathscr{R}^n$
  - random variable, Y, is measurable with respect to $\sigma(X)$ *if and only if* exists measurable function, $f : \mathbf{R}^n \to \mathbf{R}$ such that $Y(\omega) = f(X(\omega))$ for all ω, *i.e.*, $Y = f \circ X$

Probability distributions for random variables

- probability measure on $\mathbf{R}$, $\mu = PX^{-1}$, *i.e.*,
  $$\mu(A) = P(X \in A) \text{ for } A \in \mathscr{R}$$
  called *distribution* or *law* of random variable, X
- function, $F : \mathbf{R} \to [0, 1]$, defined by
  $$F(x) = \mu(-\infty, x] = P(X \leq x)$$
  called *distribution function* or *cumulative distribution function (CDF)* of X
- Borel set, S, with $P(S) = 1$, called *support*
- random variable, its distribution, its distribution function, said to be *discrete* when has *countable* support

Probability distribution of mappings of random variables

- for measurable $g : \mathbf{R} \to \mathbf{R}$,
  $$(\forall A \in \mathscr{R}) \left(\mathbf{Prob}\,(g(X) \in A) = \mathbf{Prob}\left(X \in g^{-1}(A)\right) = \mu(g^{-1}(A))\right)$$
  hence, $g(X)$ has distribution of μg^{-1}

Probability density for random variables

- Borel function, $f : \mathbf{R} \to \mathbf{R}_+$, satisfying
  $$(\forall A \in \mathscr{R}) \left(\mu(A) = P(X \in A) = \int_A f(x) dx\right)$$
  called *density* or *probability density function (PDF)* of random variable

- above is equivalent to

$$(\forall a < b \in \mathbf{R}) \left(\int_a^b f(x)dx = P(a < X \leq b) = F(b) - F(a) \right)$$

(refer to statement on page 129)

  - note, though, F does not need to differentiate to f everywhere; only f required to integrate properly
  - if F does differentiate to f and f is continuous, *fundamental theorem of calculus* implies f indeed is density for F

Probability distribution for random vectors

- (similarly to random variables) probability measure on $\mathbf{R}^n$, $\mu = PX^{-1}$, *i.e.*,

$$\mu(A) = P(X \in A) \text{ for } A \in \mathscr{B}^k$$

called *distribution* or *law* of random vector, X

- function, $F : \mathbf{R}^k \to [0, 1]$, defined by

$$F(x) = \mu S_x = P(X \preceq x)$$

where

$$S_x = \{\omega \in \Omega | X(\omega) \preceq x\} = \{\omega \in \Omega | X_i(\omega) \leq x_i\}$$

called *distribution function* or *cumulative distribution function (CDF)* of X

- (similarly to random variables) random vector, its distribution, its distribution function, said to be *discrete* when has *countable* support

Marginal distribution for random vectors

- (similarly to random variables) for measurable $g : \mathbf{R}^n \to \mathbf{R}^m$

$$(\forall A \in \mathscr{R}^m) \left(\mathbf{Prob}\,(g(X) \in A) = \mathbf{Prob}\,\left(X \in g^{-1}(A)\right) = \mu(g^{-1}(A))\right)$$

hence, $g(X)$ has distribution of μg^{-1}

- for $g_i : \mathbf{R}^n \to \mathbf{R}$ with $g_i(x) = x_i$

$$(\forall A \in \mathscr{R})\,(\mathbf{Prob}\,(g(X) \in A) = \mathbf{Prob}\,(X_i \in A))$$

- measure, μ_i, defined by $\mu_i(A) = \mathbf{Prob}\,(X_i \in A)$, called *(i-th) marginal distribution of X*
- for μ having density function, $f : \mathbf{R}^n \to \mathbf{R}_+$, density function of marginal distribution is

$$f_i(x) = \int_{\mathscr{R}^{n-1}} f(x_{-i}) d\mu_{-i}$$

where $x_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ and similarly for $d\mu_{-i}$

Independence of random variables

- random variables, X_1, $\ldots$, X_n, with independent σ-algebras generated by them, said to be *independent*

  (refer to page 130 for independence of collections of subsets)

  - because $\sigma(X_i) = X_i^{-1}(\mathscr{R}) = \{X_i^{-1}(H)|H \in \mathscr{R}\}$, independent *if and only if*

$$(\forall H_1, \ldots, H_n \in \mathscr{R}) \left(P\left(X_1 \in H_1, \ldots, X_n \in H_n\right) = \prod P\left(X_i \in H_i\right) \right)$$

  i.e.,

$$(\forall H_1, \ldots, H_n \in \mathscr{R}) \left(P\left(\bigcap X_i^{-1}(H_i)\right) = \prod P\left(X_i^{-1}(H_i)\right) \right)$$

Equivalent statements of independence of random variables

- for random variables, X_1, $\ldots$, X_n, having μ and $F : \mathbf{R}^n \to [0,1]$ as their distribution and CDF, with each X_i having μ_i and $F_i : \mathbf{R} \to [0,1]$ as its distribution and CDF, following statements are *equivalent*

  - $X_1, \ldots, X_n$ are independent
  - $(\forall H_1, \ldots, H_n \in \mathscr{R}) \left(P\left(\bigcap X_i^{-1}(H_i)\right) = \prod P\left(X_i^{-1}(H_i)\right)\right)$
  - $(\forall H_1, \ldots, H_n \in \mathscr{R}) \left(P(X_1 \in H_1, \ldots, X_n \in H_n) = \prod P(X_i \in H_i)\right)$
  - $(\forall x \in \mathbf{R}^n) \left(P(X_1 \leq x_1, \ldots, X_n \leq x_n) = \prod P(X_i \leq x_i)\right)$
  - $(\forall x \in \mathbf{R}^n) \left(F(x) = \prod F_i(x_i)\right)$
  - $\mu = \mu_1 \times \cdots \times \mu_n$
  - $(\forall x \in \mathbf{R}^n) \left(f(x) = \prod f_i(x_i)\right)$

Independence of random variables with separate σ-algebra

- given probability space, $(\Omega, \mathscr{F}, P)$

- random variables, X_1, $\ldots$, X_n, each of which is measurable with respect to each of n independent σ-algebras, $\mathscr{G}_1 \subset \mathscr{F}$, $\ldots$, $\mathscr{G}_n \subset \mathscr{F}$ respectively, are independent

Independence of random vectors

- for random vectors, $X_1 : \Omega \to \mathbf{R}^{d_1}$, $\ldots$, $X_n : \Omega \to \mathbf{R}^{d_n}$, having μ and $F : \mathbf{R}^{d_1} \times \cdots \times \mathbf{R}^{d_n} \to [0,1]$ as their distribution and CDF, with each X_i having μ_i and $F_i : \mathbf{R}^{d_i} \to [0,1]$ as its distribution and CDF, following statements are *equivalent*

  - $X_1, \ldots, X_n$ are independent
  - $\left(\forall H_1 \in \mathscr{R}^{d_1}, \ldots, H_n \in \mathscr{R}^{d_n}\right) \left(P\left(\bigcap X_i^{-1}(H_i)\right) = \prod P\left(X_i^{-1}(H_i)\right)\right)$
  - $\left(\forall H_1 \in \mathscr{R}^{d_1}, \ldots, H_n \in \mathscr{R}^{d_n}\right) \left(P(X_1 \in H_1, \ldots, X_n \in H_n) = \prod P(X_i \in H_i)\right)$
  - $\left(\forall x_1 \in \mathbf{R}^{d_1}, \ldots, x_n \in \mathbf{R}^{d_n}\right) \left(P(X_1 \preceq x_1, \ldots, X_n \preceq x_n) = \prod P(X_i \preceq x_i)\right)$
  - $\left(\forall x_1 \in \mathbf{R}^{d_1}, \ldots, x_n \in \mathbf{R}^{d_n}\right) \left(F(x_1, \ldots, x_n) = \prod F_i(x_i)\right)$
  - $\mu = \mu_1 \times \cdots \times \mu_n$
  - $\left(\forall x_1 \in \mathbf{R}^{d_1}, \ldots, x_n \in \mathbf{R}^{d_n}\right) \left(f(x_1, \ldots, x_n) = \prod f_i(x_i)\right)$

Independence of infinite collection of random vectors

- infinite collection of random vectors for which every finite subcollection is independent, said to be *independent*
- for independent (countable) collection of random vectors, $\langle\langle X_{ni}\rangle_{i=1}^{\infty}\rangle_{n=1}^{\infty}$, $\langle \mathscr{F}_n\rangle_{n=1}^{\infty}$ with $\mathscr{F}_n = \sigma(\langle X_{ni}\rangle_{i=1}^{\infty})$ are independent

Probability evaluation for two independent random vectors

Theorem 5.3 (Probability evaluation for two independent random vectors) *for independent random vectors, X and Y, with distributions, μ and ν, in $\mathbf{R}^n$ and $\mathbf{R}^m$ respectively*

$$\left(\forall B \in \mathscr{R}^{n+m}\right)\left(\mathbf{Prob}\left((X,Y)\in B\right) = \int_{\mathbf{R}^n} \mathbf{Prob}\left((x,Y)\in B\right) d\mu_X\right)$$

and

$$\left(\forall A \in \mathscr{R}^{n}, B \in \mathscr{R}^{n+m}\right)\left(\mathbf{Prob}\left(X\in A, (X,Y)\in B\right) = \int_{A} \mathbf{Prob}\left((x,Y)\in B\right) d\mu_X\right)$$

Sequence of random variables

Theorem 5.4 (squence of random variables) *for sequence of probability measures on $\mathscr{R}$, $\langle \mu_n\rangle$, exists probability space, (X, Ω, P), and sequence of independent random variables in $\mathbf{R}$, $\langle X_n\rangle$, such that each X_n has μ_n as distribution*

Expected values

Definition 5.1 (expected values) *for random variable, X, on $(\Omega, \mathscr{F}, P)$, integral of X with respect to measure, P*

$$\mathbf{E}\,X = \int X dP = \int_{\Omega} X(\omega) dP$$

called expected value of X

- $\mathbf{E}\,X$ is
  - always defined for nonnegative X
  - for general case
    - defined, or
    - X has an expected value if either $\mathbf{E}\,X^+ < \infty$ or $\mathbf{E}\,X^- < \infty$ or both, in which case, $\mathbf{E}\,X = \mathbf{E}\,X^+ - \mathbf{E}\,X^-$
- X is integrable *if and only if* $\mathbf{E}\,|X| < \infty$
- limits
  - if $\langle X_n\rangle$ is dominated by integral random variable or they are uniformly integrable, $\mathbf{E}\,X_n$ converges to $\mathbf{E}\,X$ if X_n converges to X in probability

Markov and Chebyshev's inequalities

Inequality 5.1 (Markov inequality) *for random variable, X, on $(\Omega, \mathscr{F}, P)$,*

$$\mathbf{Prob}\,(X \geq \alpha) \leq \frac{1}{\alpha} \int_{X \geq \alpha} X dP \leq \frac{1}{\alpha}\,\mathbf{E}\,X$$

for nonnegative X, hence

$$\mathbf{Prob}\,(|X| \geq \alpha) \leq \frac{1}{\alpha^n} \int_{|X| \geq \alpha} |X|^n dP \leq \frac{1}{\alpha^n}\,\mathbf{E}\,|X|^n$$

for general X

Inequality 5.2 (Chebyshev's inequality) *as special case of Markov inequality,*

$$\mathbf{Prob}\,(|X - \mathbf{E}\,X| \geq \alpha) \leq \frac{1}{\alpha^2} \int_{|X - \mathbf{E}\,X| \geq \alpha} (X - \mathbf{E}\,X)^2 dP \leq \frac{1}{\alpha^2}\,\mathbf{Var}\,X$$

for general X

Jensen's, Hölder's, and Lyapunov's inequalities

Inequality 5.3 (Jensen's inequality) *for random variable, X, on $(\Omega, \mathscr{F}, P)$, and convex function, φ*

$$\varphi\,(\mathbf{E}\,X)\,\mathbf{Prob}\,(X \geq \alpha) \leq \frac{1}{\alpha} \int_{X \geq \alpha} X dP \leq \frac{1}{\alpha}\,\mathbf{E}\,X$$

Inequality 5.4 (Holder's inequality) *for two random variables, X and Y, on $(\Omega, \mathscr{F}, P)$, and $p, q \in (1, \infty)$ with $1/p + 1/q = 1$*

$$\mathbf{E}\,|XY| \leq (\mathbf{E}\,|X|^p)^{1/p}\,(\mathbf{E}\,|X|^q)^{1/q}$$

Inequality 5.5 (Lyapunov's inequality) *for random variable, X, on $(\Omega, \mathscr{F}, P)$, and $0 < \alpha < \beta$*

$$(\mathbf{E}\,|X|^\alpha)^{1/\alpha} \leq \left(\mathbf{E}\,|X|^\beta\right)^{1/\beta}$$

- note Hölder's inequality implies Lyapunov's inequality

Maximal inequalities

Theorem 5.5 (Kolmogorov's zero-one lawy) *if $A \in \mathscr{F} = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots)$ for independent $\langle X_n \rangle$,*

$$\mathbf{Prob}\,(A) = 0 \vee \mathbf{Prob}\,(A) = 1$$

– define $S_n = \sum X_i$

Inequality 5.6 (Kolmogorov's maximal inequality) *for independent $\langle X_i \rangle_{i=1}^n$ with $\mathbf{E}\,X_i = 0$ and $\mathbf{Var}\,X_i < \infty$ and $\alpha > 0$*

$$\mathbf{Prob}\,(\max S_i \geq \alpha) \leq \frac{1}{\alpha}\,\mathbf{Var}\,S_n$$

Inequality 5.7 (Etemadi's maximal inequality) *for independent $\langle X_i \rangle_{i=1}^n$ and $\alpha > 0$*

$$\mathbf{Prob}\,(\max |S_i| \geq 3\alpha) \leq 3 \max \mathbf{Prob}\,(|S_i| \geq \alpha)$$

Moments

Definition 5.2 (moments and absolute moments) *for random variable, X, on $(\Omega, \mathscr{F}, P)$, integral of X with respect to measure, P*

$$\mathbf{E}\, X^n = \int x^k d\mu = \int x^k dF(x)$$

called k-th moment of X or μ or F, and

$$\mathbf{E}\, |X|^n = \int |x|^k d\mu = \int |x|^k dF(x)$$

called k-th absolute moment of X or μ or F

- if $\mathbf{E}\, |X|^n < \infty$, $\mathbf{E}\, |X|^k < \infty$ for $k < n$
- $\mathbf{E}\, X^n$ defined only when $\mathbf{E}\, |X|^n < \infty$

Moment generating functions

Definition 5.3 (moment generating function) *for random variable, X, on $(\Omega, \mathscr{F}, P)$, $M : \mathbf{C} \to \mathbf{C}$ defined by*

$$M(s) = \mathbf{E}\left(e^{sX}\right) = \int e^{sx} d\mu = \int e^{sx} dF(x)$$

called moment generating function of X

- n-th derivative of M with respect to s is $M^{(n)}(s) = \frac{d^n}{ds^n} F(s) = \mathbf{E}\left(X^n e^{sX}\right) = \int x e^{sx} d\mu$
- thus, n-th derivative of M with respect to s at $s = 0$ is n-th moment of X

$$M^{(n)}(0) = \mathbf{E}\, X^n$$

- for independent random variables, $\langle X_i \rangle_{i=1}^n$, moment generating function of $\sum X_i$

$$\prod M_i(s)$$

5.3 Convergence of Random Variables

Convergences of random variables

Definition 5.4 (convergence with probability 1) *random variables, $\langle X_n \rangle$, with*

$$\mathbf{Prob}\left(\lim X_n = X\right) = P(\{\omega \in \Omega | \lim X_n(\omega) = X(\omega)\}) = 1$$

said to converge to X with probability 1 and denoted by $X_n \to X$ a.s.

Definition 5.5 (convergence in probability) *random variables, $\langle X_n \rangle$, with*

$$(\forall \epsilon > 0)\left(\lim \mathbf{Prob}\left(|X_n - X| > \epsilon\right) = 0\right)$$

said to converge to X in probability

Definition 5.6 (weak convergence) *distribution functions, $\langle F_n \rangle$, with*

$$(\forall x \textit{ in domain of } F)\left(\lim F_n(x) = F(x)\right)$$

said to converge weakly to distribution function, F, and denoted by $F_n \Rightarrow F$

Definition 5.7 (converge in distribution) *When $F_n \Rightarrow F$, associated random variables, $\langle X_n \rangle$, said to* converge in distribution *to X, associated with F, and denoted by* $X_n \Rightarrow X$

Definition 5.8 (weak convergence of measures) *for measures on $(\mathbf{R}, \mathscr{R})$, $\langle \mu_n \rangle$, associated with distribution functions, $\langle F_n \rangle$, respectively, and measure on $(\mathbf{R}, \mathscr{R})$, μ, associated with distribution function, F, we denote*

$$\mu_n \Rightarrow \mu$$

if

$$(\forall A = (-\infty, x] \textit{ with } x \in \mathbf{R}) \left(\lim \mu_n(A) = \mu(A)\right)$$

- indeed, if above equation holds for $A = (-\infty, x)$, it holds for many other subsets

Relations of different types of convergences of random variables

Proposition 5.1 (relations of convergence of random variables) *convergence with probability 1 implies convergence in probability, which implies $X_n \Rightarrow X$, i.e.*

$$\begin{aligned}
& X_n \to X \textit{ a.s., i.e., } X_n \textit{ converge to } X \textit{ with probability } 1 \\
\Rightarrow \quad & X_n \textit{ converge to } X \textit{ in probability} \\
\Rightarrow \quad & X_n \Rightarrow X, \textit{ i.e., } X_n \textit{ converge to } X \textit{ in distribution,}
\end{aligned}$$

Necessary and sufficient conditions for convergence of probability

$$X_n \text{ converge in probability}$$

if and only if

$$(\forall \epsilon > 0) \left(\mathbf{Prob}\left(|X_n - X| > \epsilon \text{ i.o}\right) = \mathbf{Prob}\left(\limsup |X_n - X| > \epsilon\right) = 0\right)$$

if and only if

$$(\forall \text{ subsequence } \langle X_{n_k} \rangle) \left(\exists \text{ its subsequence } \left\langle X_{n_{k_l}} \right\rangle \text{ converging to } f \text{ with probability } 1\right)$$

Necessary and sufficient conditions for convergence in distribution

$$X_n \Rightarrow X, \textit{i.e.}, X_n \text{ converge in distribution}$$

if and only if

$$F_n \Rightarrow F, \textit{i.e.}, F_n \text{ converge weakly}$$

if and only if

$$(\forall A = (-\infty, x] \text{ with } x \in \mathbf{R}) \left(\lim \mu_n(A) = \mu(A)\right)$$

if and only if

$$(\forall x \text{ with } \mathbf{Prob}\left(X = x\right) = 0) \left(\lim \mathbf{Prob}\left(X_n \leq x\right) = \mathbf{Prob}\left(X \leq x\right)\right)$$

Strong law of large numbers – define $S_n = \sum_{i=1}^n X_i$

Theorem 5.6 (strong law of large numbers) *for sequence of independent and identically distributed (i.i.d.) random variables with finite mean,* $\langle X_n \rangle$

$$\frac{1}{n} S_n \to \mathbf{E}\, X_1$$

with probability 1

- strong law of large numbers also called *Kolmogorov's law*

Corollary 5.1 (strong law of large numbers) *for sequence of independent and identically distributed (i.i.d.) random variables with* $\mathbf{E}\, X_1^- < \infty$ *and* $\mathbf{E}\, X_1^+ = \infty$ *(hence,* $\mathbf{E}\, X = \infty$*)*

$$\frac{1}{n} S_n \to \infty$$

with probability 1

Weak law of large numbers – define $S_n = \sum_{i=1}^n X_i$

Theorem 5.7 (weak law of large numbers) *for sequence of independent and identically distributed (i.i.d.) random variables with finite mean,* $\langle X_n \rangle$

$$\frac{1}{n} S_n \to \mathbf{E}\, X_1$$

in probability

- because convergence with probability 1 implies convergence in probability (Proposition 5.1), strong law of large numbers implies weak law of large numbers

Normal distributions – assume probability space, $(\Omega, \mathscr{F}, P)$

Definition 5.9 (normal distributions) *Random variable,* $X : \Omega \to \mathbf{R}$*, with*

$$(A \in \mathscr{R}) \left(\mathbf{Prob}\,(X \in A) = \frac{1}{\sqrt{2\pi}\sigma} \int_A e^{-(x-c)^2/2} d\mu \right)$$

where $\mu = PX^{-1}$ *for some* $\sigma > 0$ *and* $c \in \mathbf{R}$*, called* normal distribution *and denoted by* $X \sim \mathcal{N}(c, \sigma^2)$

- note $\mathbf{E}\, X = c$ and $\mathbf{Var}\, X = \sigma^2$
- called *standard normal distribution* when $c = 0$ and $\sigma = 1$

Multivariate normal distributions – assume probability space, $(\Omega, \mathscr{F}, P)$

Definition 5.10 (multivariate normal distributions) *Random variable,* $X : \Omega \to \mathbf{R}^n$*, with*

$$(A \in \mathscr{R}^n) \left(\mathbf{Prob}\,(X \in A) = \frac{1}{\sqrt{(2\pi)^n}\sqrt{\det \Sigma}} \int_A e^{-(x-c)^T \Sigma^{-1} (x-c)/2} d\mu \right)$$

where $\mu = PX^{-1}$ *for some* $\Sigma \succ 0 \in \mathbf{S}_{++}^n$ *and* $c \in \mathbf{R}^n$*, called* (n-dimensional) normal distribution*, and denoted by* $X \sim \mathcal{N}(c, \Sigma)$

- note that $\mathbf{E}\, X = c$ and covariance matrix is Σ

Lindeberg-Lévy theorem – define $S_n = \sum^n X_i$

Theorem 5.8 (Lindeberg-Levy theorem) *for independent random variables, $\langle X_n \rangle$, having same distribution with expected value, c, and same variance, $\sigma^2 < \infty$, $(S_n - nc)/\sigma\sqrt{n}$ converges to standard normal distribution in distribution, i.e.,*

$$\frac{S_n - nc}{\sigma\sqrt{n}} \Rightarrow N$$

where N is standard normal distribution

– Theorem 5.8 implies

$$S_n/n \Rightarrow c$$

Limit theorems in $\mathbf{R}^n$

Theorem 5.9 (equivalent statements to weak convergence) *each of following statements are equivalent to weak convergence of measures, $\langle \mu_n \rangle$, to μ, on measurable space, $(\mathbf{R}^k, \mathscr{R}^k)$*

- $\lim \int f d\mu_n = \int f d\mu$ *for every bounded continuous f*
- $\limsup \mu_n(C) \leq \mu(C)$ *for every closed C*
- $\liminf \mu_n(G) \geq \mu(G)$ *for every open G*
- $\lim \mu_n(A) = \mu(A)$ *for every μ-continuity A*

Theorem 5.10 (convergence in distribution of random vector) *for random vectors, $\langle X_n \rangle$, and random vector, Y, of k-dimension, $X_n \Rightarrow Y$, i.e., X_n converge to Y in distribution* if and only if

$$\left(\forall z \in \mathbf{R}^k\right) \left(z^T X_n \Rightarrow z^T Y\right)$$

Central limit theorem – assume probability space, $(\Omega, \mathscr{F}, P)$ and define $\sum^n X_i = S_n$

Theorem 5.11 (central limit theorem) *for random variables, $\langle X_n \rangle$, having same distributions with $\mathbf{E}\, X_n = c \in \mathbf{R}^k$ and positive definite covariance matrix, $\Sigma \succ 0 \in \mathcal{S}_k$, i.e., $\mathbf{E}(X_n - c)(X_n - c)^T = \Sigma$, where $\Sigma_{ii} < \infty$ (hence $\Sigma \prec M I_n$ for some $M \in \mathbf{R}_{++}$ due to Cauchy-Schwarz inequality),*

$$(S_n - nc)/\sqrt{n} \text{ converges in distribution to } Y$$

where $Y \sim \mathcal{N}(0, \Sigma)$

(proof can be found in Proof 16)

Convergence of random series

- for independent $\langle X_n \rangle$, probability of $\sum X_n$ converging is either 0 or 1
- below characterize two cases in terms of distributions of individual X_n – XXX: diagram

Theorem 5.12 (convergence with probability 1 for random series) *for independent $\langle X_n \rangle$ with $\mathbf{E}\, X_n = 0$ and $\mathbf{Var}\, X_n < \infty$*

$$\sum X_n \text{ converges with probability } 1$$

Theorem 5.13 (convergence conditions for random series) *for independent $\langle X_n \rangle$, $\sum X_n$ converges with probability* 1 if and only if *they converges in probability*

– define trucated version of X_n by $X_n^{(c)}$, *i.e.*, $X_n I_{|X_n|\leq c}$

Theorem 5.14 (convergence conditions for truncated random series) *for independent* $\langle X_n \rangle$,

$$\sum X_n \text{ converge with probability } 1$$
$$\text{if all of } \sum \mathbf{Prob}\left(|X_n| > c\right), \sum \mathbf{E}(X_n^{(c)}), \sum \mathbf{Var}(X_n^{(c)}) \text{ converge for some } c > 0$$

6 Convex Optimization

6.1 Convex Sets

Lines and line segmenets

Definition 6.1 (lines) *for some* $x, y \in \mathbf{R}^n$

$$\{\theta x + (1-\theta)y | \theta \in \mathbf{R}\}$$

called line going through x and y

Definition 6.2 (line segmenets) *for some* $x, y \in \mathbf{R}^n$

$$\{\theta x + (1-\theta)y | 0 \leq \theta \leq 1 \in \mathbf{R}\}$$

called line segment connecting x and y

Affine sets

Definition 6.3 (affine sets) *set,* $C \subset \mathbf{R}^n$*, every line going through any two points in which is contained in* C*, i.e.*

$$(\forall x, y \in C) \left(\{\theta x + (1-\theta)y | \theta \in \mathbf{R}\} \subset C\right)$$

called affine set

Definition 6.4 (affine hulls) *for set,* $C \subset \mathbf{R}^n$*, intersection of all affine sets containing* C*, called* affine hull of C*, denoted by* $\mathbf{aff}\, C$*, which is equal to set of all affine combinations of points in* C*, i.e.*

$$\bigcup_{n \in \mathbf{N}} \{\theta_1 x_1 + \cdots + \theta_n x_n | x_1, \ldots, x_n \in C, \theta_1 + \cdots + \theta_n = 1\}$$

Definition 6.5 (affine dimension) *for* $C \subset \mathbf{R}^n$*, dimension of* $\mathbf{aff}\, C$*, called* affine dimension

Relative interiors and boundaries

Definition 6.6 (relative interiors of sets) *for* $C \subset \mathbf{R}^n$*,*

$$\bigcup_{O:\text{open}, O \cap \mathbf{aff}\, C \subset C} O \cap \mathbf{aff}\, C$$

or equivalently

$$\{x | (\exists \epsilon > 0)(\forall y \in \mathbf{aff}\, C, \|y - x\| < \epsilon)(y \in C)\}$$

is called relative interior of C *or* interior relative to C*, denoted by* $\mathbf{relint}\, C$

Definition 6.7 (relative boundaries of sets) *for* $C \subset \mathbf{R}^n$*,* $\overline{C} \sim \mathbf{relint}\, C$*, called* relative boundary of C

Convex sets

Definition 6.8 (convex sets) *set,* $C \subset \mathbf{R}^n$*, every line segment connecting any two points in which is contained in* C*, i.e.*

$$(\forall x, y \in C) \, (\forall 0 \leq \theta \leq 1) \, (\theta x + (1-\theta)y \in C)$$

called convex set

Definition 6.9 (convex hulls) *for set,* $C \subset \mathbf{R}^n$*, intersection of all convex sets containing* C*, called* convex hull of C*, denoted by* $\mathbf{Conv}\, C$*, which is equal to set of all convex combinations of points in* C*, i.e.*

$$\bigcup_{n \in \mathbf{N}} \{\theta_1 x_1 + \cdots + \theta_n x_n | x_1, \ldots, x_n \in C, \theta_1 + \cdots + \theta_n = 1, \theta_1, \ldots, \theta_n > 0\}$$

- convex hull (of course) is convex set

Cones

Definition 6.10 (cones) *set, $C \subset \mathbf{R}^n$, for which*

$$(\forall x \in C, \theta \geq 0)\,(\theta x \in C)$$

called cone *or* nonnegative homogeneous

Definition 6.11 (convex cone) *set, $C \subset \mathbf{R}^n$, which is both convex and cone, called* convex cone*; C is convex cone* if and only if

$$(\forall x, y \in C, \theta, \xi \geq 0)\,(\theta x + \xi y \in C)$$

- convex cone (of course) is convex set
- examples of convex cones: $\mathbf{R}^n_+$, $\mathbf{R}^n_{++}$, $\mathbf{S}^n_+$, and $\mathbf{S}^n_{++}$

Hyperplanes and half spaces

Definition 6.12 (hyperplanes) *$n-1$ dimensional affine set in $\mathbf{R}^n$, called* hyperplane*; every hyperplane can be expressed as*

$$\{x \in \mathbf{R}^n | a^T = b\}$$

for some $a \neq 0 \in \mathbf{R}^n$ and $b \in \mathbf{R}$

Definition 6.13 (half spaces) *one of two sets divided by hyperplane, called* half space*; every half space can be expressed as*

$$\{x \in \mathbf{R}^n | a^T \leq b\}$$

for some $a \neq 0 \in \mathbf{R}^n$ and $b \in \mathbf{R}$

- hyperplanes and half spaces are convex sets

Euclidean balls and ellipsoids

Definition 6.14 (Euclidean ball) *set of all points distance of which from point, $x \in \mathbf{R}^n$, is no greater than $r > 0$, called* (Euclidean) ball centered at x with radius, r, *denoted by* $B(x,r)$, *i.e.*

$$B(x,r) = \{y \in \mathbf{R}^n | \|y - x\|_2 \leq r\}$$

Definition 6.15 (ellipsoids) *ball elongated along n orthogonal axes, called* ellipsoid*, i.e.,*

$$\{y \in \mathbf{R}^n | (y-x)^T P^{-1} (y-x) \leq 1\}$$

for some $x \in \mathbf{R}^n$ and $P \in \mathbf{S}^n_{++}$

- Euclidean balls and ellipsoids are convex sets

Norm balls and norm cones

Definition 6.16 (norm ball) *for norm, $\|\cdot\| : \mathbf{R}^n \to \mathbf{R}_+$, set of all points distance of which measured in the norm from point, $x \in \mathbf{R}^n$, is no greater than $r > 0$, called* norm ball centered at x with radius, r, associated with norm, $\|\cdot\|$, *i.e.*

$$\{y \in \mathbf{R}^n | \|y - x\| \leq r\}$$

Definition 6.17 (norm cone) *for norm, $\|\cdot\| : \mathbf{R}^n \to \mathbf{R}_+$, $x \in \mathbf{R}^n$, and $r > 0$,*

$$\{(x,y) \in \mathbf{R}^n \times \mathbf{R} | \|x\| \leq r\} \subset \mathbf{R}^{n+1}$$

called cone associated with norm, $\|\cdot\|$

Definition 6.18 (second-order cone) *norm cone associated with Euclidean norm, called* second-order cone

- norm balls and norm cones are convex sets

Polyhedra

Definition 6.19 (polyhedra) *intersection of finite number of hyperplanes and half spaces, called* polyhedron*; every polyhedron can be expressed as*

$$\{x \in \mathbf{R}^n | Ax \preceq b, Cx = d\}$$

for $A \in \mathbf{R}^{m\times n}$*,* $b \in \mathbf{R}^m$*,* $C \in \mathbf{R}^{p\times n}$*,* $d \in \mathbf{R}^p$

- polyhedron is convex set (by Proposition 6.1)

Convexity preserving set operations

Proposition 6.1 (convexity preserving set operations)

- *intersection preserves convexity*
  - *for (any) collection of convex sets,* $\mathcal{C}$*,*
    $$\bigcap_{C\in\mathcal{C}} C$$
    is convex set (proof can be found in Proof 17)
- *scalar scaling preserves convexity*
  - *for convex set* C
    $$\alpha C$$
    is convex set for any $\alpha \in \mathbf{R}$
- *sum preserves convexity*
  - *for convex sets* C *and* D
    $$C + D$$
    is convex set
- *direct product preserves convexity*
  - *for convex sets* C *and* D
    $$C \times D$$
    is convex set
- *projection preserves convexity*
  - *for convex set* $C \subset A \times B$
    $$\{x \in A | (\exists y)((x, y) \in C)\}$$
    is convex
- *image and inverse image by affine function preserve convexity*
  - *for affine function* $f : A \to B$ *and convex sets* $C \subset A$ *and* $D \subset B$
    $$f(C) \ \&\ f^{-1}(D)$$
    are convex
- *image and inverse image by linear-fractional function preserve convexity*
  - *for convex sets* $C \subset \mathbf{R}^n, D \subset \mathbf{R}^m$ *and linear-fractional function,* $g : \mathbf{R}^n \to \mathbf{R}^m$*, i.e., function defined by* $g(x) = (Ax + b)/(c^T x + d)$ *for* $A \in \mathbf{R}^{m\times n}$*,* $b \in \mathbf{R}^m$*,* $c \in \mathbf{R}^n$*, and* $d \in \mathbf{R}$
    $$g(C) \ \&\ g^{-1}(D)$$
    are convex

Proper cones and generalized inequalities

Definition 6.20 (proper cones) *closed convex cone K which is*

- *solid, i.e., $K^\circ \neq \emptyset$*
- *pointed, i.e., $x \in vK$ and $-x \in K$ imply $x = 0$*

called proper cone

- examples of proper cones: $\mathbf{R}^n_+$ and $\mathbf{S}^n_+$

Definition 6.21 (generalized inequalities) *proper cone K defines* generalized inequalities

- *(nonstrict) generalized inequality*
$$x \preceq_K y \Leftrightarrow y - x \in K$$
- *strict generalized inequality*
$$x \prec_K y \Leftrightarrow y - x \in K^\circ$$

- $\preceq_K$ and $\prec_K$ are partial orderings

Convex sets induced by generalized inequalities

- for affine function $g : \mathbf{R}^n \to \mathbf{S}^m$, *i.e.*, $f(x) = A_0 + A_1x_1 + \cdots + A_nx_n$ for some $A_0, \ldots, A_n \in \mathbf{S}^m$, $f^{-1}(\mathbf{S}^n_+)$ is convex (by Proposition 6.1), *i.e.*,
$$\{x \in \mathbf{R}^n | A_0 + A_1x_1 + \cdots + A_nx_n \succeq 0\} \subset \mathbf{R}^n$$
is convex
- can negate each matrix A_i and have same results, hence
$$\{x \in \mathbf{R}^n | A_0 + A_1x_1 + \cdots + A_nx_n \preceq 0\} \subset \mathbf{R}^n$$
is (also) convex

Separating and supporting hyperplanes

Theorem 6.1 (separating hyperplane theorem) *for nonempty disjoint convex sets C and D, exists hyperplane which separates C and D, i.e.*

$$(\exists a \neq 0 \in \mathbf{R}^n, b \in \mathbf{R})\, (\forall x \in C, y \in D)\, \left(a^T x + b \geq 0 \ \&\ a^T y + b \leq 0\right)$$

Definition 6.22 (separating hyperplanes) *for nonempty disjoint convex sets C and D, hyperplane satisfying property in Theorem 6.1, called* separating hyperplane*, said to* separate C and D

Theorem 6.2 (supporting hyperplane theorem) *for nonempty convex set C and $x \in \mathbf{bd}\, C$, exists hyperplane passing through x, i.e.,*

$$(\exists a \neq 0 \in \mathbf{R}^n)\, (\forall y \in C)\, \left(a^T (y - x) \leq 0\right)$$

Definition 6.23 (supporting hyperplanes) *for nonempty convex set C and $x \in \mathbf{bd}\, C$, hyperplane satisfied property in Theorem 6.2, called* supporting hyperplane

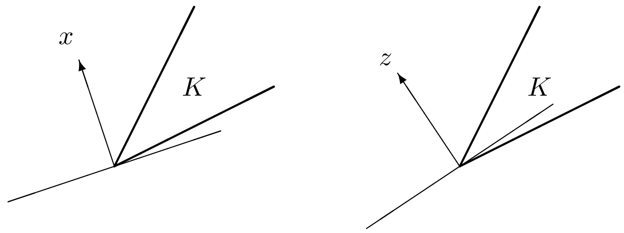

Figure 6.18: dual cone

Dual cones

Definition 6.24 (dual cones) *for cone K,*

$$\{x | \forall y \in K, y^T x \geq 0\}$$

called dual cone of K, denoted by K^*

- Figure 6.18 illustrates $x \in K^*$ while $z \notin K^*$

Dual norms

Definition 6.25 (dual norms) *for norm $\|\cdot\|$, fudnction defined by*

$$y \mapsto \sup\{y^T x | \|x\| \leq 1\}$$

called dual norm of $\|\cdot\|$, denoted by $\|\cdot\|_*$

- examples
  - dual cone of subspace $V \subset \mathbf{R}^n$ is orthogonal complement of V, $V^\perp$, where $V^\perp = \{y | \forall v \in V, v^T y = 0\}$
  - $\mathbf{R}^n_+$ and $\mathbf{S}^n_+$ are self-dual
  - *dual of norm cone* is *norm cone associated with dual norm*, *i.e.*, if $K = \{(x,t) \in \mathbf{R}^n \times \mathbf{R} | \|x\| \leq t\}$

$$K = \{(y,u) \in \mathbf{R}^n \times \mathbf{R} | \|y\|_* \leq u\}$$

Properties of dual cones

Proposition 6.2 (properties of dual cones) *for cones K, K_1, and K_2*

- *K^* is closed and convex*
- *$K_1 \subset K_2 \Rightarrow K_2^* \subset K_1^*$*
- *if $K^\circ \neq \emptyset$, K^* is pointed*
- *if $\overline{K}$ is pointed, $(K^*)^\circ \neq \emptyset$*
- *$K^{**} = (K^*)^*$ is closure of convex hull of K,*
- *K^* is closed and convex*

thus,

- *if K is closed and convex, $K^{**} = K$*
- *dual of proper cone is proper cone*
- *for proper cone K, $K^{**} = K$*

Dual generalized inequalities

- dual of proper cone is proper (Proposition 6.2), hence the dual also induces generalized inequalities

Proposition 6.3 *for proper cone K,*

- $x \preceq_K y$ if and only if $(\forall \lambda \succeq_{K^*} 0)(\lambda^T x \leq \lambda^T y)$
- $x \prec_K y$ if and only if $(\forall \lambda \succeq_{K^*} 0 \text{ with } \lambda \neq 0)(\lambda^T x < \lambda^T y)$

$K^{**} = K$, hence above are equivalent to

- $x \preceq_{K^*} y$ if and only if $(\forall \lambda \succeq_K 0)(\lambda^T x \leq \lambda^T y)$
- $x \prec_{K^*} y$ if and only if $(\forall \lambda \succeq_K 0 \text{ with } \lambda \neq 0)(\lambda^T x < \lambda^T y)$

Theorem of alternative for linear strict generalized inequalities

Theorem 6.3 (theorem of alternative for linear strict generalized inequalities) *for proper cone $K \subset \mathbf{R}^m$, $A \in \mathbf{R}^{m\times n}$, and $b \in \mathbf{R}^m$,*

$$Ax \prec_K b$$

is infeasible if and only if *exist nonzero $\lambda \in \mathbf{R}^m$ such that*

$$\lambda \neq 0,\ \lambda \succeq_{K^*} 0,\ A^T\lambda = 0,\ \lambda^T b \leq 0$$

Above two inequality systems are alternative*, i.e., for any data, A and b, exactly one of them is feasible. (proof can be found in Proof 18)*

6.2 Convex Functions

Convex functions

Definition 6.26 (convex functions)

- *function $f : \mathbf{R}^n \to \mathbf{R}$ the domain of which is convex and which satisfies*

  $$(\forall x, y \in \mathbf{dom}\, f, 0 \leq \theta \leq 1)\, (f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y))$$

  said to be convex

- *function $f : \mathbf{R}^n \to \mathbf{R}$ the domain of which is convex and which satisfies*

  $$(\forall \text{ distinct } x, y \in \mathbf{dom}\, f, 0 < \theta < 1)\, (f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y))$$

  said to be strictly convex

Definition 6.27 (concave functions)

- *function $f : \mathbf{R}^n \to \mathbf{R}$ the domain of which is convex where $-f$ is convex, said to be* concave
- *function $f : \mathbf{R}^n \to \mathbf{R}$ the domain of which is convex where $-f$ is strictly convex, said to be* strictly concave

Extended real-value extensions of convex functions

Definition 6.28 (extended real-value extension of convex functions) *for convex function f, function $\tilde{f}: \mathbf{R}^n \to \mathbf{R} \cup \{\infty\}$ defined by*

$$\tilde{f}(x) = \left\{ \begin{array}{ll} f(x) & \text{if } x \in \mathbf{dom}\, f \\ \infty & \text{if } x \not\in \mathbf{dom}\, f \end{array} \right.$$

called extended real-value extension of f

- using extended real-value extensions of convex functions, can drop "$\mathbf{dom}\, f$" in equations, *e.g.*,
  - f is convex *if and only if* its extended-value extension $\tilde{f}$ satisfies

$$(\forall x, y \in \mathbf{dom}\, f, 0 \leq \theta \leq 1) \left(f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y)\right)$$

  - f is strictly convex *if and only if* its extended-value extension $\tilde{f}$ satisfies

$$(\forall \text{ distinct } x, y \in \mathbf{dom}\, f, 0 < \theta < 1) \left(f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y)\right)$$

First-order condition for convexity

Theorem 6.4 (first-order condition for convexity) *differentiable f, i.e., $\mathbf{dom}\, f$ is open and gradient ∇f exists at every point in $\mathbf{dom}\, f$, is*

- *convex* if and only if $\mathbf{dom}\, f$ *is convex and*

$$(\forall x, y \in \mathbf{dom}\, f) \left(f(y) \geq f(x) + \nabla f(x)^T (y - x)\right)$$

- *strictly convex* if and only if $\mathbf{dom}\, f$ *is convex and*

$$(\forall \text{ distinct } x, y \in \mathbf{dom}\, f) \left(f(y) > f(x) + \nabla f(x)^T (y - x)\right)$$

- Theorem 6.4 implies that for convex function f
  - first-order Taylor approximation is *global underestimator*
  - can derive global information from local information
    - *e.g.*, if $\nabla f(x) = 0$, x is global minimizer
    - *explains remarkable properties of convex functions and convex optimization problems*

Second-order condition for convexity

Theorem 6.5 (second-order condition for convexity) *twice-differentiable f, i.e., $\mathbf{dom}\, f$ is open and Hessian $\nabla^2 f$ exists at every point in $\mathbf{dom}\, f$, is convex* if and only if $\mathbf{dom}\, f$ *is convex and*

$$(\forall x \in \mathbf{dom}\, f) \left(\nabla^2 f(x) \succeq 0\right)$$

- *if* $\mathbf{dom}\, f$ *is convex and*

$$(\forall x \in \mathbf{dom}\, f) \left(\nabla^2 f(x) \succ 0\right)$$

  it is strictly convex

Convex function examples

- assume function $f : \mathbf{R}^n \to \mathbf{R}$ and $\mathbf{dom}\, f = \mathbf{R}^n$ unlesss specified otherwise

- affine function, *i.e.*, $f(x) = a^T x + b$ for some $a \in \mathbf{R}^n$ and $b \in \mathbf{R}$, is convex
- quadratic functions - if $f(x) = x^T P x + q^T x$ for some $P \in \mathbf{S}^n$ and $q \in \mathbf{R}^n$
  - f is convex *if and only if* $P \succeq 0$
  - f is strictly convex *if and only if* $P \succ 0$
- exponential function, *i.e.*, $f(x) = \exp(a^T x + b)$ for some $a \in \mathbf{R}^n$ and $b \in \mathbf{R}$, is convex
- power, *i.e.*, $f(x) = x^a$ for some $a \geq 1$, is convex on $\mathbf{R}_{++}$
- power of absolute value, *i.e.*, $f(x) = |x|^a$ for some $a \geq 1$, is convex on $\mathbf{R}$
- logarithm function, *i.e.*, $f(x) = \log x$, is concave on $\mathbf{R}_{++}$
- negative entropy, *i.e.*,

$$f(x) = \begin{cases} x \log x & \text{if } x > 0 \\ 0 & \text{if } x = 0 \end{cases}$$

  is convex on $\mathbf{R}_+$
- norm as function is convex (by definition of norms, *i.e.*, triangle inequality & absolute homogeneity)
- max function, *i.e.*, $f(x) = \max(x_1, \ldots, x_n\}$, is convex
- quadratic-over-linear function, $f(x, y) = x^2/y$, is convex on $\mathbf{R} \times \mathbf{R}_{++}$
- log-sum-exp, $f(x) = \log(\exp(x_1) + \cdots + \exp(x_n))$, is convex
- geometric mean, $f(x) = (\prod_{i=1}^n x_i)^{1/n}$, is concave on $\mathbf{R}^n_{++}$
- log-determinant, $f(X) = \log\det X$, is concave on $\mathbf{S}^n_{++}$

Sublevel sets and superlevel sets

Definition 6.29 (sublevel sets) *for function f and $\alpha \in \mathbf{R}$,*

$$\{x \in \mathbf{dom}\, f | f(x) \leq \alpha\}$$

called α-sublevel set of f

Definition 6.30 (superlevel sets) *for function f and $\alpha \in \mathbf{R}$,*

$$\{x \in \mathbf{dom}\, f | f(x) \geq \alpha\}$$

called α-superlevel set of f

Proposition 6.4 (convexity of level sets)

- *every sublevel set of convex function is convex*
- *and every superlevel set of concave function is convex*

  - note, however, converse is not true
    - *e.g.*, every sublevel set of log is convex, but log is concave

Epigraphs and hypographs

Definition 6.31 (epigraphs) *for function f,*

$$\{(x,t)|x \in \mathbf{dom}\, f, f(x) \leq t\}$$

called epigraph of f*, denoted by* $\mathbf{epi}\, f$

Definition 6.32 (hypographs) *for function f,*

$$\{(x,t)|x \in \mathbf{dom}\, f, f(x) \geq t\}$$

called hypograph of f*, denoted by* $\mathbf{hypo}\, f$

Proposition 6.5 (graphs and convexity)

- *function is convex* if and only if *its epigraph is convex*
- *function is concave* if and only if *its hypograph is convex*

Convexity preserving function operations

Proposition 6.6 (convexity preserving function operations)

- *nonnegative weighted sum preserves convexity*
  - *for convex functions f_1, ..., f_n and nonnegative weights $w_1, \ldots, w_n$*

    $$w_1 f_1 + \cdots w_n f_n$$

    is convex
- *nonnegative weighted integration preserves convexity*
  - *for measurable set Y, $w : Y \to \mathbf{R}_+$, and $f : X \times Y$ where $f(x, y)$ is convex in x for every $y \in Y$ and measurable in y for every $x \in X$*

    $$\int_Y w(y) f(x, y) dy$$

    is convex
- *pointwise maximum preserves convexity*
  - *for convex functions f_1, ..., f_n*

    $$\max\{f_1, \ldots, f_n\}$$

    is convex
- *pointwise supremum preserves convexity*
  - *for indexed family of convex functions $\{f_\lambda\}_{\lambda \in \Lambda}$*

    $$\sup_{\lambda \in \Lambda} f_\lambda$$

    is convex (one way to see this is $\mathbf{epi} \sup_\lambda f_\lambda = \bigcap_\lambda \mathbf{epi}\, f_\lambda$)
- *composition*
  - *suppose $g : \mathbf{R}^n \to \mathbf{R}^k$, $h : \mathbf{R}^k \to \mathbf{R}$, and $f = h \circ g$*

- *f convex if h convex & nondecreasing in each argument, and g_i convex*
- *f convex if h convex & nonincreasing in each argument, and g_i concave*
- *f concave if h concave & nondecreasing in each argument, and g_i concave*
- *f concave if h concave & nonincreasing in each argument, and g_i convex*

- *minimization*
  - *for function $f(x,y)$ convex in (x,y) and convex set C*

$$\inf_{y \in C} f(x,y)$$

is convex provided it is bounded below where domain is $\{x|(\exists y \in C)((x,y) \in \mathbf{dom}\, f)\}$ (proof can be found in Proof 19)

- *perspective of convex function preserves convexity*
  - *for convex function $f : X \to \mathbf{R}$, function $g : X \times \mathbf{R} \to \mathbf{R}$ defined by*

$$g(x,t) = tf(x/t)$$

with $\mathbf{dom}\, g = \{(x,t)|x/t \in \mathbf{dom}\, f, t > 0\}$ is convex

Convex functions examples Proposition 6.6 implies

- piecewise-linear function is convex, *i.e.*
  - $\max\{a_1^T x + b_1, \ldots, a_m^T x + b_m\}$ for some $a_i \in \mathbf{R}^n$ and $b_i \in \mathbf{R}$ is convex
- sum of k largest components is convex, *i.e.*
  - $x_{[1]} + \cdots + x_{[k]}$ where $x_{[i]}$ denotes i-th largest component, is convex (since $f(x) = \max\{x_{i_1} + \cdots + x_{i_r} | 1 \leq i_1 < i_2 < \cdots < i_r \leq n\}$)
- support function of set, *i.e.*,
  - $\sup\{x^T y | y \in A\}$ for $A \subset \mathbf{R}^n$ is convex
- distance (when measured by arbitrary norm) to farthest point of set
  - $\sup\{\|x - y\| | y \in A\}$ for $A \subset \mathbf{R}^n$ is convex
- least-squares cost as function of weights
  - $\inf_{x \in \mathbf{R}^n} \sum_{i=1}^n w_i (a_i^T x - b_i)^2$ for some $a_i \in \mathbf{R}^n$ and $b_i \in \mathbf{R}$ is concave
    - note that above function equals to $\sum_{i=1}^n w_i b_i^2 - \sum_{i=1}^n w_i^2 b_i^2 a_i^T \left(\sum_{j=1}^n w_j a_j a_j^T \right)^{-1} a_i$ but not clear whether it is concave

- maximum eigenvalue of symmetric matrix
  - $\lambda_{\max}(F(x)) = \sup\{y^T F(x) y | \|y\|_2 \leq 1\}$ where $F : \mathbf{R}^n \to \mathbf{S}^m$ is linear function in x
- norm of matrix
  - $\sup\{u^T G(x) v | \|u\|_2 \leq 1, \|v\|_2 \leq 1\}$ where $G : \mathbf{R}^n \to \mathbf{R}^{m \times n}$ is linear function in x

- distance (when measured by arbitrary norm) to convex set
    - for convex set C, $\inf\{\|x-y\| | y \in C\}$
- infimum of convex function subject to linear constraint
    - for convex function h, $\inf\{h(y)|Ay=x\}$ is convex (since it is $\inf_y(h(y)+I_{Ay=x}(x,y))$)
- perspective of Euclidean norm squared
    - map $(x,t) \mapsto x^Tx/t$ induces convex function in (x,t) for $t>0$
- perspective of negative log
    - map $(x,t) \mapsto -t\log(x/t)$ induces convex function in $(x,t) \in \mathbf{R}^2_{++}$

- perspective of convex function
    - for convex function $f : \mathbf{R}^n \to \mathbf{R}$, function $g : \mathbf{R}^n \to \mathbf{R}$ defined by

$$g(x) = (c^Tx+d)f((Ax+b)/(c^Tx+d))$$

    from some $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$, $c \in \mathbf{R}^n$, and $d \in \mathbf{R}$ with $\mathbf{dom}\, g = \{x|(Ax+b)/(c^Tx+d) \in \mathbf{dom}\, f, c^Tx+d>0\}$ is convex

Conjugate functions

Definition 6.33 (conjugate functions) *for function f*

$$\sup_{y\in\mathbf{dom}\, f}(x^Ty - f(y))$$

called conjugate function of f, denoted by f^*

- conjugate function is convex for any function f because it is supremum of linear (hence convex) functions (in x) (Proposition 6.6)

Inequality 6.1 (Fenchel's inequality) *definition of conjugate function implies*

$$f(x) + f^*(y) \geq x^Ty$$

sometimes called Young's inequality

Proposition 6.7 (conjugate of conjugate) *for convex and closed function f*

$$f^{**} = f$$

where closed function f is defined by function with closed $\mathbf{epi}\, f$

Conjugate function examples

- strictly convex quadratic function
  - for $f : \mathbf{R}^n \to \mathbf{R}_+$ defined $f(x) = x^T Qx/2$ where $Q \in \mathbf{S}^n_{++}$,

$$f^*(x) = \sup_y (y^T x - y^T Qy/2) = (y^T x - y^T Qy/2)|_{y=Q^{-1}x} = x^T Q^{-1} x/2$$

    which is also strictly convex quadratic function

- log-determinant
  - for function $f : \mathbf{S}^n_{++} \to \mathbf{R}$ defined by $f(X) = \log\det X^{-1}$

$$f^*(X) = \sup_{Y \in \mathbf{S}^n_{++}} (\mathbf{Tr}\, XY + \log\det Y) = \log\det(-X)^{-1} - n$$

    where $\mathbf{dom}\, f^* = -\mathbf{S}^n_{++}$

- indicator function
  - for indicator function $I_A : \mathbf{R}^n \to \{0, \infty\}$ with $A \subset \mathbf{R}^n$

$$I_A^*(x) = \sup_y (y^T x - I_A(y)) = \sup\{y^T x | y \in A\}$$

    which is support function of A

- log-sum-exp function
  - for function $f : \mathbf{R}^n \to \mathbf{R}$ defined by $f(x) = \log(\sum_{i=1}^n \exp(x_i))$

$$f^*(x) = \sum_{i=1}^n x_i \log x_i + I_{x \succeq 0, \mathbf{1}^T x = 1}(x)$$

- norm
  - for norm function $f : \mathbf{R}^n \to \mathbf{R}_+$ defined by $f(x) = \|x\|$

$$f^*(x) = \sup_y (y^T x - \|y\|) = I_{\|x\|_* \leq 1}(x)$$

- norm squared
  - for function $f : \mathbf{R} \to \mathbf{R}_+$ defined by $f(x) = \|x\|^2/2$

$$f^*(x) = \|x\|_*^2/2$$

- differentiable convex function
  - for differentiable convex function $f : \mathbf{R}^n \to \mathbf{R}$

$$f^*(x) = (y^*)^T \nabla f(y^*) - f(y^*)$$

    where $y^* = \operatorname{argsup}_y (x^T y - f(y))$

- sum of independent functions
  - for function $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ defined by $f(x, y) = f_1(x) + f_2(y)$ where $f_1 : \mathbf{R}^n \to \mathbf{R}$ and $f_2 : \mathbf{R}^m \to \mathbf{R}$

$$f^*(x, y) = f_1^*(x) + f_2^*(y)$$

Convex functions with respect to generalized inequalities

Definition 6.34 (K-convex functions) *for proper cone K,*

- *function f satisfying*

$$(\forall x, y \in \mathbf{dom}\, f, 0 \leq \theta \leq 1)\, (f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta) f(y))$$

  called K-convex

- *function f satisfying*

$$(\forall x \neq y \in \mathbf{dom}\, f, 0 < \theta < 1)\, (f(\theta x + (1-\theta)y) \prec_K \theta f(x) + (1-\theta) f(y))$$

  called strictly K-convex

Proposition 6.8 (dual characterization of K-convexity) *for proper cone K*

- *function f is K-convex* if and only if *for every $w \succeq_{K^*} 0$, $w^T f$ is convex*
- *function f is strictly K-convex* if and only if *for every nonzero $w \succeq_{K^*} 0$, $w^T f$ is strictly convex*

Matrix convexity

Definition 6.35 (matrix convexity) *function of $\mathbf{R}^n$ into $\mathbf{S}^m$ which is K-convex where $K = \mathbf{S}^m_+$, called* matrix convex

- examples of matrix convexity
  - function of $\mathbf{R}^{n\times m}$ into $\mathbf{S}^n_+$ defined by $X \mapsto XX^T$ is matrix convex
  - function of $\mathbf{S}^n_{++}$ into itself defined by $X \mapsto X^p$ is matrix convex for $1 \leq p \leq 2$ or $-1 \leq p \leq 0$, and matrix concave for $0 \leq p \leq 1$
  - function of $\mathbf{S}^n$ into $\mathbf{S}^n_{++}$ defined by $X \mapsto \exp(X)$ is *not* matrix convex
  - quadratic matrix function of $\mathbf{R}^{m\times n}$ into $\mathbf{S}^n$ defined by $X \mapsto X^T AX + B^T X + X^T B + C$ for $A \in \mathbf{S}^m$, $B \in \mathbf{R}^{m\times n}$, and $C \in \mathbf{S}^n$ is matrix convex when $A \succeq 0$

6.3 Convex Optimization Problems

Optimization problems

Definition 6.36 (optimization problems) *for $f : F \to \mathbf{R}$, $q : Q \to \mathbf{R}^m$, $h : H \to \mathbf{R}^p$ where F, Q, and H are subsets of common set X*

$$\begin{array}{ll} \textit{minimize} & f(x) \\ \textit{subject to} & q(x) \preceq 0 \\ & h(x) = 0 \end{array}$$

called optimization problem *where x is* optimization variable

- *f, q, and h are* objective function, inequality & equality contraint function
- *$q(x) \preceq 0$ and $h(x) = 0$ are* inequality contraints *and* equality contraints
- *$\mathcal{D} = F \cap Q \cap H$ is* domain *of optimization problem*
- *$\mathcal{F} = \{x \in \mathcal{D} | q(x) \preceq 0, h(x) = 0\}$, called* feasible set, *$x \in \mathcal{D}$, said to be* feasible *if $x \in \mathcal{F}$, optimization problem, said to be* feasible *if $\mathcal{F} \neq \emptyset$*
- *$p^* = \inf\{f(x) | x \in \mathcal{F}\}$, called* optimal value *of optimization problem*
- *if optimization problem is* infeasible, $p^* = \infty$ *(following convention that infimum of empty set is ∞)*
- *if $p^* = -\infty$, optimization problem said to be* unbounded

Global and local optimalities

Definition 6.37 (global optimality) *for optimization problem in Definition 6.36*

- *$x \in \mathcal{F}$ with $f(x) = p^*$, called* (global) optimal point
- *$X_{\text{opt}} = \{x \in \mathcal{F} | f(x) = p^*\}$, called* optimal set
- *when $X_{\text{opt}} \neq \emptyset$, we say optimal value is* attained *or* achieved *and optimization problem is* solvable

- optimization problem is *not* solvable if $p^* = \infty$ or $p^* = -\infty$ (converse is not true)

Definition 6.38 (local optimality) *for optimization problem in Definition 6.36 where X is metric space, $x \in \mathcal{F}$ satisfying $\inf\{f(z) | z \in \mathcal{F}, \rho(z, x) \leq r\}$ where $\rho : X \times X \to \mathbf{R}_+$ is metric, for some $r > 0$, said to be* locally optimal, *i.e., x solves*

$$\begin{array}{ll} \textit{minimize} & f(z) \\ \textit{subject to} & q(z) \preceq 0 \\ & h(z) = 0 \\ & \rho(z, x) \leq r \end{array}$$

Equivalent optimization problems

Definition 6.39 (equivalent optimization problems) *two optimization problems where solving one readily solve the other, said to be* equivalent

- below two optimization problems are equivalent
  - $$\begin{array}{ll} \text{minimize} & -x - y \\ \text{subject to} & 2x + y \leq 1 \\ & x + 2y \leq 1 \end{array}$$
  - $$\begin{array}{ll} \text{minimize} & -2u - v/3 \\ \text{subject to} & 4u + v/3 \leq 1 \\ & 2u + 2v/3 \leq 1 \end{array}$$

  since if (x^*, y^*) solves first, $(u, v) = (x^*/2, 3y^*)$ solves second, and if (u^*, v^*) solves second, $(x, y) = (2u^*, v^*/3)$ solves first

Change of variables

- given function $\phi : \mathcal{Z} \to X$, optimization problem in Definition 6.36 can be rewritten as

$$\begin{array}{ll} \text{minimize} & f(\phi(z)) \\ \text{subject to} & q(\phi(z)) \preceq 0 \\ & h(\phi(z)) = 0 \end{array}$$

  where $z \in \mathcal{Z}$ is optimization variable

- if ϕ is injective and $\mathcal{D} \subset \phi(\mathcal{Z})$, above optimization problem and optimization problem in Definition 6.36 are equivalent, *i.e.*
  - X_{opt} is optimal set of problem in Definition 6.36 $\Rightarrow$ $\phi^{-1}(X_{\text{opt}})$ is optimal set of above problem
  - Z_{opt} is optimal set of above problem $\Rightarrow$ $\phi(Z_{\text{opt}})$ is optimal set of problem in Definition 6.36

- two optimization problems said to be related by *change of variable or substitution of variable $x = \phi(z)$*

Convex optimization

Definition 6.40 (convex optimization) *optimization problem in Definition 6.36 where X is Banach space, i.e., complete linear normed vector space, f & q are convex functions, and h is affine function, called* convex optimization problem

- *when $X = \mathbf{R}^n$, optimization problem can be formulated as*

$$\begin{array}{ll} \textit{minimize} & f(x) \\ \textit{subject to} & q(x) \preceq 0 \\ & Ax = b \end{array}$$

  for some $A \in \mathbf{R}^{p \times n}$ and $b \in \mathbf{R}^p$

  - domain of convex optimization problem is *convex*
    - since domains of f, q, and h are convex (by definition of convex functions) and intersection of convex sets is convex
  - feasible set of convex optimization problem is *convex*
    - since sublevel sets of convex functions are convex, feasible sets for affine function is either empty set, singleton, or affine sets, all of which are convex sets

Optimality conditions for convex optimization problems

Theorem 6.6 (local optimality implies global optimality) *for convex optimization problem (in Definition 6.40), every local optimal point is global optimal point*

Theorem 6.7 (optimality conditions for convex optimality problems) *for convex optimization problem (in Definition 6.40), when f is differentiable (i.e., $\mathbf{dom}\, f$ is open and ∇f exists everywhere in $\mathbf{dom}\, f$)*

- *$x \in \mathcal{D}$ is optimal* if and only if *$x \in \mathcal{F}$ and*

$$(\forall y \in \mathcal{F}) \left(\nabla f(x)^T (y - x) \geq 0\right)$$

- *for unconstrained problems, $x \in \mathcal{D}$ is optimal* if and only if

$$\nabla f(x) = 0$$

Optimality conditions for some convex optimization problems

- unconstrained convex quadratic optimization

$$\text{minimize} \quad f(x) = (1/2)x^T P x + q^T x$$

  where $F = \mathbf{R}^n$ and $P \in \mathbf{S}^n_+$

  - x is optimal *if and only if*

$$\nabla f(x) = Px + q = 0$$

    exist three cases
    - if $P \in \mathbf{S}^n_{++}$, exists unique optimum $x^* = -P^{-1} q$
    - if $q \in \mathcal{R}(P)$, $X_{\text{opt}} = -P^\dagger q + \mathcal{N}(P)$
    - if $q \not\in \mathcal{R}(P)$, $p^* = -\infty$

- analytic centering

$$\text{minimize} \quad f(x) = -\sum_{i=1}^m \log(b_i - a_i^T x)$$

  where $F = \{x \in \mathbf{R}^n | Ax \prec b\}$

  - x is optimal *if and only if*

$$\nabla f(x) = \sum_{i=1}^m \frac{1}{b_i - a_i^T x} a_i = 0$$

    exist three cases

    - exists unique optimum, which happens *if and only if* $\{x | b_i - a_i^T x\}$ is nonempty and bounded
    - exist infinitely many optima, in which case, X_{opt} is affine set
    - exists no optimum, which happens *if and only if* f is unbounded below

- convex optimization problem with equality constraints only

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

  where $X = \mathbf{R}^n$

  - x is optimal *if and only if*

$$\nabla f(x) \perp \mathcal{N}(A)$$

    or equivalently, exists $\nu \in \mathbf{R}^p$ such that

$$\nabla f(x) = A^T \nu$$

Linear programming

Definition 6.41 (linear programming) *convex optimization problem in Definition 6.40 with $X = \mathbf{R}^n$ and linear f & q, called linear program (LP), which can be formulated as*

$$\begin{array}{ll} \textit{minimize} & c^T x \\ \textit{subject to} & Cx \preceq d \\ & Ax = b \end{array}$$

where $c \in \mathbf{R}^n$, $C \in \mathbf{R}^{m \times n}$, $d \in \mathbf{R}^m$, $A \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^p$

- *can transform above LP into standard form LP*

$$\begin{array}{ll} \textit{minimize} & \tilde{c}^T \tilde{x} \\ \textit{subject to} & \tilde{A}\tilde{x} = \tilde{b} \\ & \tilde{x} \succeq 0 \end{array}$$

LP examples

- diet problem - find amount of n different food to minimize purchase cost while satisfying nutrition requirements
  - assume exist n food and m nutritions, c_i is cost of food i, A_{ji} is amount of nutrition j contained in unit quantity of food i, b_j is amount requirement for nutrition j
  - diet problem can be formulated as LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \succeq b \\ & x \succeq 0 \end{array}$$

- Chebyshev center of polyhedron - find largest Euclidean ball contained in polyhedron
  - assume polyhedron is $\{x \in \mathbf{R}^n | a_i^T x \leq b_i, i = 1, \ldots, m\}$
  - problem of finding Chebyshev center of polyhedron can be formulated as LP

$$\begin{array}{ll} \text{maximize} & r \\ \text{subject to} & a_i^T x + r\|a_i\|_2 \leq b_i \end{array}$$

    where optimization variables are $x \in \mathbf{R}^n$ and $r \in \mathbf{R}$

- piecewise-linear minimization - minimize maximum of affine functions
  - assume m affine functions $a_i^T x + b_i$
  - piecewise-linear minimization problem can be formulated as LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & a_i^T x + b_i \leq t, \quad i = 1, \ldots, m \end{array}$$

- linear-fractional program

$$\begin{array}{ll} \text{minimize} & (c^T x + d)/(e^T x + f) \\ \text{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

  - if feasible set is nonempty, can be formulated as LP

$$\begin{array}{ll} \text{minimize} & c^T y + dz \\ \text{subject to} & Gy - hz \preceq 0 \\ & Ay - bz = 0 \\ & e^T y + fz = 1 \\ & z \geq 0 \end{array}$$

Quadratic programming

Definition 6.42 (quadratic programming) *convex optimization problem in Definition 6.40 with $X = \mathbf{R}^n$ and convex quadratic f and linear q, called quadratic program (QP), which can be formulated as*

$$\begin{array}{ll} \textit{minimize} & (1/2)x^T P x + q^T x \\ \textit{subject to} & Gx \preceq h \\ & Ax = b \end{array}$$

where $P \in \mathbf{S}_+^n$, $q \in \mathbf{R}^n$, $G \in \mathbf{R}^{m \times n}$, $h \in \mathbf{R}^m$, $A \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^p$

- when $P = 0$, QP reduces to LP, hence *LP is specialization of QP*

QP examples

- least-squares (LS) problems
  - LS can be formulated as QP

$$\text{minimize} \quad \|Ax - b\|_2^2$$

- distance between two polyhedra
  - assume two polyhedra $\{x \in \mathbf{R}^n | Ax \preceq b, Cx = d\}$ and $\{x \in \mathbf{R}^n | \tilde{A}x \preceq \tilde{b}, \tilde{C}x = \tilde{d}\}$
  - problem of finding distance between two polyhedra can be formulated as QP

$$\begin{array}{ll} \text{minimize} & \|x - y\|_2^2 \\ \text{subject to} & Ax \preceq b, \quad Cx = d \\ & \tilde{A}y \preceq \tilde{b}, \quad \tilde{C}y = \tilde{d} \end{array}$$

Quadratically constrained quadratic programming

Definition 6.43 (quadratically constrained quadratic programming) *convex optimization problem in Definition 6.40 with $X = \mathbf{R}^n$ and convex quadratic f & q, called* quadratically constrained quadratic program (QCQP)*, which can be formulated as*

$$\begin{array}{ll} \textit{minimize} & (1/2)x^T P_0 x + q_0^T x \\ \textit{subject to} & (1/2)x^T P_i x + q_i^T x + r_i \leq 0, \quad i = 1, \ldots, m \\ & Ax = b \end{array}$$

where $P_i \in \mathbf{S}_+^n$, $q_i \in \mathbf{R}^n$, $r_i \in \mathbf{R}$, $A \in \mathbf{R}^{p \times n}$, $b \in \mathbf{R}^p$

- when $P_i = 0$ for $i = 1, \ldots, m$, QCQP reduces to QP, hence *QP is specialization of QCQP*

Second-order cone programming

Definition 6.44 (second-order cone programming) *convex optimization problem in Definition 6.40 with $X = \mathbf{R}^n$ and linear f and convex q of form*

$$\begin{array}{ll} \textit{minimize} & f^T x \\ \textit{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, m \\ & Fx = g \end{array}$$

where $f \in \mathbf{R}^n$, $A_i \in \mathbf{R}^{n_i \times n}$, $b_i \in \mathbf{R}^{n_i}$, $c_i \in \mathbf{R}^n$, $d_i \in \mathbf{R}$, $F \in \mathbf{R}^{p \times n}$, $g \in \mathbf{R}^p$ called second-order cone program (SOCP)

- when $b_i = 0$, SOCP reduces to QCQP, hence *QCQP is specialization of SOCP*

SOCP examples

- robust linear program - minimize $c^T x$ while satisfying $\tilde{a}_i^T x \leq b_i$ for every $\tilde{a}_i \in \{a_i + P_i u \mid \|u\|_2 \leq 1\}$ where $P_i \in \mathbf{S}^n$
  - can be formulated as SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x + \|P_i^T x\|_2 \leq b_i \end{array}$$

- linear program with random constraints - minimize $c^T x$ while satisfying $\tilde{a}_i^T x \leq b_i$ with probability no less than η where $\tilde{a} \sim \mathcal{N}(a_i, \Sigma_i)$
  - can be formulated as SOCP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & a_i^T x + \Phi^{-1}(\eta)\|\Sigma_i^{1/2} x\|_2 \leq b_i \end{array}$$

Geometric programming

Definition 6.45 (monomial functions) *function $f : \mathbf{R}_{++}^n \to \mathbf{R}$ defined by*

$$f(x) = c x_1^{a_1} \cdots x_n^{a_n}$$

where $c > 0$ and $a_i \in \mathbf{R}$, called monomial function *or simply* monomial

Definition 6.46 (posynomial functions) *function $f : \mathbf{R}_{++}^n \to \mathbf{R}$ which is finite sum of monomial functions, called* posynomial function *or simply* posynomial

Definition 6.47 (geometric programming) *optimization problem*

$$\begin{array}{ll} \textit{minimize} & f(x) \\ \textit{subject to} & q(x) \preceq 1 \\ & h(x) = 1 \end{array}$$

for posynomials $f : \mathbf{R}^n_{++} \to \mathbf{R}$ & $q : \mathbf{R}^n_{++} \to \mathbf{R}^m$ and monomials $h : \mathbf{R}^n_{++} \to \mathbf{R}^p$, called geometric program (GP)

Geometric programming in convex form

- geometric program in Definition 6.47 is not convex optimization problem (as it is)
- however, can be transformed to equivalent convex optimization problem by change of variables and transformation of functions

Proposition 6.9 (geometric programming in convex form) *geometric program (in Definition 6.47) can be transformed to equivalent convex optimization problem*

$$\begin{array}{ll} \textit{minimize} & \log\left(\sum_{k=1}^{K_0} \exp((a_k^{(0)})^T y + b_k^{(0)})\right) \\ \textit{subject to} & \log\left(\sum_{k=1}^{K_i} \exp((a_k^{(i)})^T y + b_k^{(i)})\right) \leq 0 \quad i = 1, \ldots, m \\ & Gy = h \end{array}$$

for some $a_k^{(i)} \in \mathbf{R}^n$, $b_k^{(i)} \in \mathbf{R}$, $G \in \mathbf{R}^{p \times n}$, $h \in \mathbf{R}^p$ where optimization variable is $y = \log(x) \in \mathbf{R}^n$

Convex optimization with generalized inequalities

Definition 6.48 (convex optimization with generalized inequality constraints) *convex optimization problem in Definition 6.40 with inequality constraints replaced by generalized inequality constraints, i.e.*

$$\begin{array}{ll} \textit{minimize} & f(x) \\ \textit{subject to} & q_i(x) \preceq_{K_i} 0 \quad i = 1, \ldots, q \\ & h(x) = 0 \end{array}$$

where $K_i \subset R^{k_i}$ are proper cones and $q_i : Q_i \to \mathbf{R}^{k_i}$ are K_i-convex, called convex optimization problem with generalized inequality constraints

- problem in Definition 6.48 reduces to convex optimization problem in Definition 6.40 when $q = 1$ and $K_1 = \mathbf{R}^m_+$, hence *convex optimization is specialization of convex optimization with generalized inequalities*
- like convex optimization
  - feasible set is $\mathcal{F} = \{x \in \mathcal{D} | q_i(x) \preceq_{K_i} 0, Ax = b\}$ is convex
  - local optimality implies global optimality
  - optimality conditions in Theorem 6.7 applies without modification

Conic programming

Definition 6.49 (conic programming) *convex optimization problem with generalized inequality constraints in Definition 6.48 with linear f and one affine q*

$$\begin{array}{ll} \textit{minimize} & f(x) \\ \textit{subject to} & q(x) \preceq_K 0 \\ & h(x) = 0 \end{array}$$

called conic program (CP)

- *can transform above CP to* standard form CP

$$\begin{array}{ll} \textit{minimize} & \tilde{f}(X) \\ \textit{subject to} & \tilde{h}(X) = 0 \\ & X \succeq_K 0 \end{array}$$

- cone program is one of simplest convex optimization problems with generalized inequalities

Semidefinite programming

Definition 6.50 (semidefinite programming) *conic program in Definition 6.49 with $X = \mathbf{R}^n$ and $K = \mathbf{S}^n_+$*

$$\begin{array}{ll} \textit{minimize} & c^T x \\ \textit{subject to} & x_1 F_1 + \cdots + x_n F_n + G \preceq 0 \\ & Ax = b \end{array}$$

where $F_1, \ldots, F_n, G \in \mathbf{S}^k$ and $A \in \mathbf{R}^{p \times n}$, called semidefinite program (SDP)

- *above inequality, called* linear matrix inequality (LMI)
- *can transform SDP to standard form SDP*

$$\begin{array}{lll} \textit{minimize} & \mathbf{Tr}(CX) & \\ \textit{subject to} & \mathbf{Tr}(A_i X) = b_i & i = 1, \ldots, p \\ & X \succeq 0 & \end{array}$$

  where $X = \mathbf{S}^n_+$ and $C, A_1, \ldots, A_p \in \mathbf{S}^n$ and $b_i \in \mathbf{R}$

SDP examples

- LP
  - if $k = m$, $F_i = \mathbf{diag}(C_{1,i}, \ldots, C_{m,i})$, $G = -\mathbf{diag}(d_1, \ldots, d_m)$ in Definition 6.50, SDP reduces to LP in Definition 6.41
  - hence, LP is specialization of SDP
- SOCP
  - SOCP in Definition 6.44 is equivalent to

$$\begin{array}{ll} \text{minimize} & f^T x \\ \text{subject to} & Fx = g \\ & \begin{bmatrix} c_i^T x + d_i & x^T A_i^T + b_i^T \\ A_i x + b_i & (c_i^T x + d_i) I_{n_i} \end{bmatrix} \succeq 0 \quad i = 1, \ldots, m \end{array}$$

    which can be transformed to SDP in Definition 6.50, thus, SDP reduces to SOCP
  - hence, SOCP is specialization of SDP

Determinant maximization problems

Definition 6.51 (determinant maximization problems) *convex optimization problem with generalized inequality constraints in Definition 6.48 with $X = \mathbf{R}^n$ of form*

$$\begin{array}{ll} \textit{minimize} & -\log\det(x_1 C_1 + \cdots + x_n C_n + D) + c^T x \\ \textit{subject to} & x_1 F_1 + \cdots + x_n F_n + G \preceq 0 \\ & -x_1 C_1 - \cdots - x_n C_n - D \prec 0 \\ & Ax = b \end{array}$$

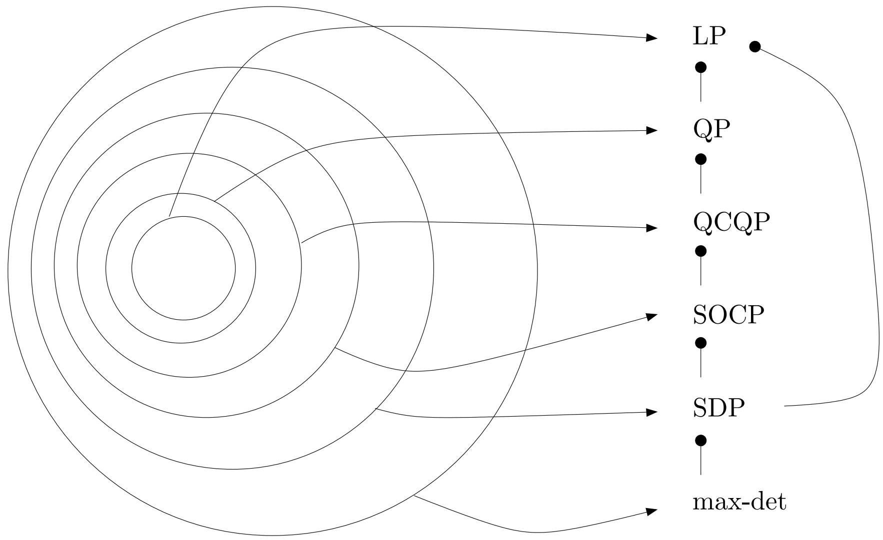

Figure 6.19: diagrams for containment of convex optimization problems

where $c \in \mathbf{R}^n$, $C_1, \ldots, C_n, D \in \mathbf{S}^l$, $F_1, \ldots, F_n, G \in \mathbf{S}^k$, *and* $A \in \mathbf{R}^{p \times n}$, *called* determinant maximization problem *or simply* max-det problem *(since it maximizes determinant of (positive definite) matrix with constraints)*

- if $l = 1$, $C_1 = \cdots = C_n = 0$, $D = 1$, max-det problem reduces to SDP, hence *SDP is specialization of max-det problem*

Diagrams for containment of convex optimization problems

- Figure 6.19 shows containment relations among convex optimization problems
- vertical lines ending with filled circles indicate existence of direct reductions, *i.e.*, optimization problem transformations to special cases

6.4 Duality

Lagrangian

Definition 6.52 (Lagrangian) *for optimization problem in Definition 6.36 with nonempty domain* $\mathcal{D}$, *function* $L : \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$ *defined by*

$$L(x, \lambda, \nu) = f(x) + \lambda^T q(x) + \nu^T h(x)$$

called Lagrangian *associated with the optimization problem where*

- λ, *called* Lagrange multiplier associated inequality constraints $q(x) \preceq 0$
- λ_i, *called* Lagrange multiplier associated i-th inequality constraint $q_i(x) \leq 0$
- ν, *called* Lagrange multiplier associated equality constraints $h(x) = 0$
- ν_i, *called* Lagrange multiplier associated i-th equality constraint $h_i(x) = 0$
- λ *and* ν, *called* dual variables *or* Lagrange multiplier vectors *associated with the optimization problem*

Lagrange dual functions

Definition 6.53 (Lagrange dual functions) *for optimization problem in Definition 6.36 for which Lagrangian is defined, function* $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R} \cup \{-\infty\}$ *defined by*

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f(x) + \lambda^T q(x) + \nu^T h(x) \right)$$

called Lagrange dual function *or just* dual function *associated with the optimization problem*

- g is *(always) concave function* (even when optimization problem is not convex)
  - since is pointwise infimum of linear (hence concave) functions is concave
- $g(\lambda, \nu)$ provides lower bound for optimal value of associated optimization problem, *i.e.*,

$$g(\lambda, \nu) \leq p^*$$

  for every $\lambda \succeq 0$ (proof can be found in Proof 20)
- $(\lambda, \nu) \in \{(\lambda, \nu) | \lambda \succeq 0, g(\lambda, \nu) > -\infty\}$, said to be *dual feasible*

Dual function examples

- LS solution of linear equations

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$

  - Lagrangian - $L(x, \nu) = x^T x + \nu^T (Ax - b)$
  - Lagrange dual function

$$g(\nu) = -\frac{1}{4} \nu^T A A^T \nu - b^T \nu$$

- standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array}$$

  - Lagrangian - $L(x, \lambda, \nu) = c^T x - \lambda^T x + \nu^T (Ax - b)$
  - Lagrange dual function

$$g(\lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

    - hence, set of dual feasible points is $\{(A^T \nu + c, \nu) | A^T \nu + c \succeq 0\}$
- *maximum cut*, sometimes called *max-cut*, problem, which is NP-hard

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1 \end{array}$$

  where $W \in \mathbf{S}^n$
  - Lagrangian - $L(x, \nu) = x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T x$
  - Lagrange dual function

$$g(\nu) = \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

    - hence, set of dual feasible points is $\{\nu | W + \mathbf{diag}(\nu) \succeq 0\}$

- some trivial problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x = 0 \end{array}$$

  - Lagrangian - $L(x,\nu) = f(x) + \nu^T x$
  - Lagrange dual function

$$g(\nu) = \inf_{x \in \mathbf{R}^n} (f(x) + \nu^T x) = -\sup_{x \in \mathbf{R}^n} ((-\nu)^T x - f(x)) = -f^*(-\nu)$$

    - hence, set of dual feasible points is $-\mathbf{dom}\, f^*$, and for every $f : \mathbf{R}^n \to \mathbf{R}$ and $\nu \in \mathbf{R}^n$

$$-f^*(-\nu) \le f(0)$$

- minimization with linear inequality and equality constraints

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d \end{array}$$

  - Lagrangian - $L(x,\lambda,\nu) = f(x) + \lambda^T(Ax - b) + \nu^T(Cx - d)$
  - Lagrange dual function

$$g(\nu) = -b^T\lambda - d^T\nu - f^*(-A^T\lambda - C^T\nu)$$

    - hence, set of dual feasible points is $\{(\lambda,\nu) | -A^T\lambda - C^T\nu \in \mathbf{dom}\, f^*, \lambda \succeq 0\}$

- equality constrained norm minimization

$$\begin{array}{ll} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{array}$$

  - Lagrangian - $L(x,\nu) = \|x\| + \nu^T(Ax - b)$
  - Lagrange dual function

$$g(\nu) = -b^T\nu - \sup_{x \in \mathbf{R}^n} ((-A^T\nu)^T x - \|x\|) = \begin{cases} -b^T\nu & \|A^T\nu\|_* \le 1 \\ -\infty & \text{otherwise} \end{cases}$$

    - hence, set of dual feasible points is $\{\nu | \|A^T\nu\|_* \le 1\}$

- entropy maximization

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

  where domain of objective function is $\mathbf{R}^n_{++}$

  - Lagrangian - $L(x,\lambda,\nu) = \sum_{i=1}^n x_i \log x_i + \lambda^T(Ax - b) + \nu(\mathbf{1}^T x - 1)$
  - Lagrange dual function

$$g(\lambda,\nu) = -b^T\lambda - \nu - \exp(-\nu - 1) \sum_{i=1}^n \exp(a_i^T \lambda)$$

    obtained using $f^*(y) = \sum_{i=1}^n \exp(y_i - 1)$ where a_i is i-th column vector of A

- minimum volume covering ellipsoid

$$\begin{array}{ll} \text{minimize} & -\log\det X \\ \text{subject to} & a_i^T X a_i \leq 1 \quad i = 1, \ldots, m \end{array}$$

  where domain of objective function is $\mathbf{S}_{++}^n$

  - Lagrangian - $L(X, \lambda) = -\log\det X + \sum_{i=1}^m \lambda_i (a_i^T X a_i - 1)$
  - Lagrange dual function

$$g(\lambda) = \left\{ \begin{array}{ll} \log\det(\sum_{i=1}^m \lambda_i a_i a_i^T) - \mathbf{1}^T \lambda + n & \sum_{i=1}^m \lambda_i a_i a_i^T \succ 0 \\ -\infty & \text{otherwise} \end{array} \right.$$

    obtained using $f^*(Y) = -\log\det(-Y) - n$

Best lower bound

- for every (λ, ν) with $\lambda \succeq 0$, Lagrange dual function $g(\lambda, \nu)$ (in Definition 6.53) provides lower bound for optimal value p^* for optimization problem in Definition 6.36
- natural question to ask is
  - how good is the lower bound?
  - what is best lower bound we can achieve?
- these questions lead to definition of *Lagrange dual problem*

Lagrange dual problems

Definition 6.54 (Lagrange dual problems) *for optimization problem in Definition 6.36, optimization problem*

$$\begin{array}{ll} \textit{maximize} & g(\lambda, \nu) \\ \textit{subject to} & \lambda \succeq 0 \end{array}$$

called Lagrange dual problem *associated with problem in Definition 6.36*

- *original problem in Definition 6.36, (somestime) called* primal problem
- *domain is* $\mathbf{R}^m \times \mathbf{R}^p$
- dual feasibility *defined in page 164, i.e.,* (λ, ν) *satisfying* $\lambda \succeq 0 \quad g(\lambda, \nu) > -\infty$ *indeed means feasibility for Lagrange dual problem*
- $d^* = \sup\{g(\lambda, \nu) | \lambda \in \mathbf{R}^m,\ \nu \in \mathbf{R}^p,\ \lambda \succeq 0\}$*, called* dual optimal value
- $(\lambda^*, \nu^*) = \operatorname{argsup}\{g(\lambda, \nu) | \lambda \in \mathbf{R}^m,\ \nu \in \mathbf{R}^p,\ \lambda \succeq 0\}$*, said to be* dual optimal *or called* optimal Lagrange multipliers *(if exists)*

- Lagrange dual problem in Definition 6.54 is convex optimization (even though original problem is not) since $g(\lambda, \nu)$ is always convex

Making dual constraints explicit dual problems

- (out specific) way we define Lagrange dual function in Definition 6.53 as function g of $\mathbf{R}^m \times \mathbf{R}^p$ into $\mathbf{R} \cup \{-\infty\}$, *i.e.*, $\mathbf{dom}\, g = \mathbf{R}^n \times \mathbf{R}^p$
- however, in many cases, feasible set $\{(\lambda, \nu) | \lambda \succeq 0 \quad g(\lambda, \nu) > -\infty\}$ is proper subset of $\mathbf{R}^n \times \mathbf{R}^p$
- can make this implicit feasibility condition explicit by adding it as constraint (as shown in following examples)

Lagrange dual problems associated with LPs

- standard form LP
  - primal problem
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & x \succeq 0 \end{array}$$
  - Lagrange dual problem
$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) = \left\{ \begin{array}{ll} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{array} \right. \\ \text{subject to} & \lambda \succeq 0 \end{array}$$
    (refer to page 164 for Lagrange dual function)
    - can make dual feasibility explicit by adding it to constraints as mentioned on page 166
$$\begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & \lambda \succeq 0 \\ & A^T \nu - \lambda + c = 0 \end{array}$$
    - can further simplify problem
$$\begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \nu + c \succeq 0 \end{array}$$
  - last problem is *inequality form LP*
  - all three problems are equivalent, but *not* same problems
  - will, however, with abuse of terminology, refer to all three problems as Lagrange dual problem
- inequality form LP
  - primal problem
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \end{array}$$
  - Lagrangian
$$L(x, \lambda) = c^T x + \lambda^T (Ax - b)$$
  - Lagrange dual function
$$g(\lambda) = -b^T \lambda + \inf_{x \in \mathbf{R}^n} (c + A^T \lambda)^T x = \left\{ \begin{array}{ll} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{array} \right.$$
  - Lagrange dual problem
$$\begin{array}{ll} \text{maximize} & g(\lambda) = \left\{ \begin{array}{ll} -b^T \lambda & A^T \lambda + c = 0 \\ -\infty & \text{otherwise} \end{array} \right. \\ \text{subject to} & \lambda \succeq 0 \end{array}$$
    - can make dual feasibility explicit by adding it to constraints as mentioned on page 166
$$\begin{array}{ll} \text{maximize} & -b^T \nu \\ \text{subject to} & A^T \lambda + c = 0 \\ & \lambda \succeq 0 \end{array}$$
  - dual problem is *standard form LP*
- thus, dual of standard form LP is inequality form LP and vice versa
- also, for both cases, dual of dual is same as primal problem

Lagrange dual problem of equality constrained optimization problem

- equality constrained optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

- dual function

$$\begin{aligned} g(\nu) &= \inf_{x \in \mathbf{dom}\, f} (f(x) + \nu^T(Ax - b)) = -b^T\nu - \sup_{x \in \mathbf{dom}\, f} (-\nu^T Ax - f(x)) \\ &= -b^T\nu - f^*(-A^T\nu) \end{aligned}$$

- dual problem

$$\text{maximize} \quad -b^T\nu - f^*(-A^T\nu)$$

Lagrange dual problem associated with equality constrained quadratic program

- strictly convex quadratic problem

$$\begin{array}{ll} \text{minimize} & f(x) = x^T Px + q^T x + r \\ \text{subject to} & Ax = b \end{array}$$

  - conjugate function of objective function

$$f^*(x) = (x - q)^T P^{-1}(x - q)/4 - r = x^T P^{-1}x/4 - q^T P^{-1}x/2 + q^T P^{-1}q/4 - r$$

  - dual problem

$$\text{maximize} \quad -\nu^T(AP^{-1}A^T)\nu/4 - (b + AP^{-1}q/2)^T\nu - q^T P^{-1}q/4 + r$$

Lagrange dual problems associated with nonconvex quadratic problems

- primal problem

$$\begin{array}{ll} \text{minimize} & x^T Ax + 2b^T x \\ \text{subject to} & x^T x \leq 1 \end{array}$$

  where $A \in \mathbf{S}^n$, $A \notin \mathbf{S}^n_+$, and $b \in \mathbf{R}^n$

  - since $A \not\succeq 0$, not convex optimization problem
  - sometimes called *trust region problem* arising minimizing second-order approximation of function over bounded region

- Lagrange dual function

$$g(\lambda) = \begin{cases} -b^T(A + \lambda I)^\dagger b - \lambda & A + \lambda I \succeq 0,\ b \in \mathcal{R}(A + \lambda I) \\ -\infty & \text{otherwise} \end{cases}$$

  where $(A + \lambda I)^\dagger$ is pseudo-inverse of $A + \lambda I$

- Lagrange dual problem

$$\begin{array}{ll} \text{maximize} & -b^T(A + \lambda I)^\dagger b - \lambda \\ \text{subject to} & A + \lambda I \succeq 0,\ b \in \mathcal{R}(A + \lambda I) \end{array}$$

  where optimization variable is $\lambda \in \mathbf{R}$

  - note we do not need constraint $\lambda \geq 0$ since it is implied by $A + \lambda I \succeq 0$

- though not obvious from what it appears to be, it is (of course) convex optimization problem (by definition of Lagrange dual function, *i.e.*, Definition 6.53)
- can be expressed ar

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^n (q_i^T b)^2/(\lambda_i + \lambda) - \lambda \\ \text{subject to} & \lambda \geq -\lambda_{\min}(A) \end{array}$$

where λ_i and q_i are eigenvalues and corresponding orthogormal eigenvectors of A, when $\lambda_i + \lambda = 0$ for some i, we interpret $(q_i^T b)^2/0$ as 0 if $q_i^T 0$ and ∞ otherwise

Weak duality

- since $g(\lambda, \nu) \leq p^*$ for every $\lambda \succeq 0$, we have

$$d^* = \sup\{g(\lambda, \nu) | \lambda \in \mathbf{R}^m,\ \nu \in \mathbf{R}^p,\ \lambda \succeq 0\} \leq p^*$$

Definition 6.55 (weak duality) *property that that optimal value of optimization problem (in Definition 6.36) is always no less than optimal value of Lagrange daul problem (in Definition 6.54)*

$$d^* \leq p^*$$

called weak duality

- *d^* is best lower bound for primal problem that can be obtained from Lagrange dual function (by definition)*
- *weak duality holds even when d^* or/and p^* are not finite, e.g.*
  - if primal problem is unbounded below *so that $p^* = -\infty$, must have $d^* = -\infty$, i.e.,* dual problem is infeasible
  - *conversely,* if dual problem is unbounded above *so that $d^* = \infty$, must have $p^* = \infty$, i.e.,* primal problem is infeasible

Optimal duality gap

Definition 6.56 (optimal duality gap) *difference between optimal value of optimization problem (in Definition 6.36) and optimal value of Lagrange daul problem (in Definition 6.54), i.e.*

$$p^* - d^*$$

called optimal duality gap

- sometimes used for lower bound of optimal value of problem which is difficult to solve
  - for example, dual problem of max-cut problem (on page 164), which is NP-hard, is

$$\begin{array}{ll} \text{minimize} & -\mathbf{1}^T \nu \\ \text{subject to} & W + \mathbf{diag}(\nu) \succeq 0 \end{array}$$

    where optimization variable is $\nu \in \mathbf{R}^n$
    - the dual problem can be solved very efficiently using polynomial time algorithms while primal problme *cannot* be solved unless n is very small

Strong duality

Definition 6.57 (strong duality) *if optimal value of optimization problem (in Definition 6.36) equals to optimal value of Lagrange daul problem (in Definition 6.54), i.e.*

$$p^* = d^*$$

strong duality *said to hold*

- strong duality does *not* hold in general
  - if it held always, max-cut problem, which is NP-hard, can be solved in polynomial time, which would be one of biggest breakthrough in field of theoretical computer science
  - may mean some of strongest cryptography methods, *e.g.*, homeomorphic cryptography, can be broken

Slater's theorem

- exist many conditions which guarantee strong duality, which are called *constraint qualifications* - one of them is Slater's condition

Theorem 6.8 (Slater's theorem) *if optimization problem is convex (Definition 6.40), and exists feasible $x \in \mathcal{D}$ contained in* **relint** $\mathcal{D}$ *such that*

$$q(x) \prec 0 \quad h(x) = 0$$

strong duality *holds (and* dual optimum is attained when $d^* > -\infty$*)*

- *such condition, called* Slater's condition
- *such point, (sometimes) said to be* strictly feasible

when there are affine inequality constraints, can refine Slater's condition - if first k inequality constraint functions q_1, ..., q_k are affine, Slater's condition can be relaxed to

$$q_i(x) \leq 0 \;\; i = 1, \ldots, k \quad q_i(x) < 0 \;\; i = k+1, \ldots, m \quad h(x) = 0$$

Strong duality for LS solution of linear equations

- primal problem

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$

- dual problem

$$\text{maximize} \quad g(\nu) = -\frac{1}{4}\nu^T AA^T \nu - b^T \nu$$

  (refer to page 164 for Lagrange dual function)

- "dual is always feasible" and "primal is feasible $\Rightarrow$ Slater's condition holds", thus Slater's theorem (Theorem 6.8) implies, exist only three cases
  - $(d^* = p^* \in \mathbf{R})$ or $(d^* \in \mathbf{R}$ & $p^* = \infty)$ or $(d^* = p^* = \infty)$
- if primal is infeasible, though, $b \not\in \mathcal{R}(A)$, thus exists z, such that $A^T z = 0$ and $b^T z \neq 0$, then line $\{tz | t \in \mathbf{R}\}$ makes dual problem unbounded above, hence $d^* = \infty$
- hence, *strong duality always holds*, *i.e.*, $(d^* = p^* \in \mathbf{R})$ or $(d^* = p^* = \infty)$

Strong duality for LP

- every LP either is infeasible or satisfies Slater's condition
- dual of LP is LP, hence, Slater's theorem (Theorem 6.8) implies
  - if primal is feasible, either ($d^* = p^* = -\infty$) or ($d^* = p^* \in \mathbf{R}$)
  - if dual is feasible, either ($d^* = p^* = \infty$) or ($d^* = p^* \in \mathbf{R}$)
  - only other case left is ($d^* = -\infty$ & $p^* = \infty$)
    - indeed, this pathological case can happen

Strong duality for entropy maximization

- primal problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

- dual problem (refer to page 165 for Lagrange dual function)

$$\begin{array}{ll} \text{maximize} & -b^T\lambda - \nu - \exp(-\nu - 1)\sum_{i=1}^n \exp(a_i^T\lambda) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

- dual problem is feasible, hence, Slater's theorem (Theorem 6.8) implies, if exists $x \succ 0$ with $Ax \preceq b$ and $\mathbf{1}^T x = 1$, strong duality holds, and indeed $d^* = p^* \in \mathbf{R}$
- by the way, can simplify dual problem by maximizing dual objective function over ν

$$\begin{array}{ll} \text{maximize} & -b^T\lambda - \log\left(\sum_{i=1}^n \exp(a_i^T\lambda)\right) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

  which is geometry program in convex form (Proposition 6.9) with nonnegativity contraint

Strong duality for minimum volume covering ellipsoid

- primal problem

$$\begin{array}{lll} \text{minimize} & -\log\det X & \\ \text{subject to} & a_i^T X a_i \leq 1 & i = 1, \ldots, m \end{array}$$

  where $\mathcal{D} = \mathbf{S}_{++}^n$
- dual problem

$$\begin{array}{ll} \text{maximize} & \begin{cases} \log\det(\sum_{i=1}^m \lambda_i a_i a_i^T) - \mathbf{1}^T\lambda + n & \sum_{i=1}^m \lambda_i a_i a_i^T \succ 0 \\ -\infty & \text{otherwise} \end{cases} \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

  (refer to page 166 for Lagrange dual function)
- $X = \alpha I$ with large enough $\alpha > 0$ satisfies primal's constraints, hence Slater's condition *always* holds, thus, *strong duality always holds*, *i.e.*, ($d^* = p^* \in \mathbf{R}$) or ($d^* = p^* = -\infty$)
- in fact, $\mathcal{R}(a_1, \ldots, a_m) = \mathbf{R}^n$ *if and only if* $d^* = p^* \in \mathbf{R}^n$

Strong duality for trust region nonconvex quadratic problems

- one of rare occasions in which *strong duality obtains for nonconvex problems*
- primal problem

$$\begin{array}{ll} \text{minimize} & x^T A x + 2b^T x \\ \text{subject to} & x^T x \leq 1 \end{array}$$

  where $A \in \mathbf{S}^n$, $A \not\in \mathbf{S}^n_+$, and $b \in \mathbf{R}^n$

- Lagrange dual problem (page 168)

$$\begin{array}{ll} \text{maximize} & -b^T (A + \lambda I)^\dagger b - \lambda \\ \text{subject to} & A + \lambda I \succeq 0, \; b \in \mathcal{R}(A + \lambda I) \end{array}$$

- *strong duality always holds* and $d^* = p^* \in \mathbf{R}$ (since dual problem is feasible - large enough λ satisfies dual constraints)
- in fact, exists stronger result - *strong dual holds* for optimization problem with quadratic objective and *one* quadratic inequality constraint, provided Slater's condition holds

Matrix games using mixed strategies

- matrix game - consider game with two players A and B
  - player A makes choice $1 \leq a \leq n$, player B makes choice $1 \leq b \leq m$, then player A makes payment of P_{ab} to player B
  - matrix $P \in \mathbf{R}^{n \times m}$, called *payoff matrix*
  - player A tries to pay as little as possible & player B tries to received as much as possible
  - players use *randomized or mixed strategies*, *i.e.*, each player makes choice randomly and independently of other player's choice according to probability distributions

$$\mathbf{Prob}(a = i) = u_i \; i = 1 \leq i \leq n \quad \mathbf{Prob}(b = j) = v_j \; i = 1 \leq j \leq m$$

- expected payoff (from player A to player B)

$$\sum_i \sum_j u_i v_j P_{ij} = u^T P v$$

- assume player A's strategy is known to play B
  - player B will choose v to maximize $u^T P v$

$$\sup\{u^T P v | v \succeq 0, \; \mathbf{1}^T v = 1\} = \max_{1 \leq j \leq m} (P^T u)_j$$

  - player A (assuming that player B will employ above strategy to maximize payment) will choose u to minimize payment

$$\begin{array}{ll} \text{minimize} & \max_{1 \leq j \leq m} (P^T u)_j \\ \text{subject to} & u \succeq 0 \quad \mathbf{1}^T u = 1 \end{array}$$

- assume player B's strategy is known to play A
  - then player B will do same to maximize payment (assuming that player A will employ such strategy to minimize payment)

$$\begin{array}{ll} \text{maximize} & \min_{1 \leq i \leq n} (P v)_i \\ \text{subject to} & v \succeq 0 \quad \mathbf{1}^T v = 1 \end{array}$$

Strong duality for matrix games using mixed strategies

- in matrix game, can guess in frist came, player B has advantage over player A because A's strategy's exposed to B, and vice versa, hence optimal value of first problem is greater than that of second problem
- surprising, no one has advantage over the other, *i.e.*, optimal values of two problems are *same* - will show this
- first observe both problems are (convex) piecewise-linear optimization problems
- formulate first problem as LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & u \succeq 0 \quad \mathbf{1}^T u = 1 \quad P^T u \preceq t\mathbf{1} \end{array}$$

  - Lagrangian

$$L(u, t, \lambda_1, \lambda_2, \nu) = \nu + (1 - \mathbf{1}^T \lambda_1)t + (P\lambda_1 - \nu\mathbf{1} - \lambda_2)^T u$$

  - Lagrange dual function

$$g(\lambda_1, \lambda_2, \nu) = \left\{ \begin{array}{ll} \nu & \mathbf{1}^T \lambda_1 = 1 \;\&\; P\lambda_1 - \nu\mathbf{1} = \lambda_2 \\ -\infty & \text{otherwise} \end{array} \right.$$

- Lagrange dual problem

$$\begin{array}{ll} \text{maximize} & \nu \\ \text{subject to} & \mathbf{1}^T \lambda_1 = 1 \quad P\lambda_1 - \nu\mathbf{1} = \lambda_2 \\ & \lambda_1 \succeq 0 \quad \lambda_2 \succeq 0 \end{array}$$

- eliminating λ_2 gives below Lagrange dual problem

$$\begin{array}{ll} \text{maximize} & \nu \\ \text{subject to} & \lambda_1 \succeq 0 \quad \mathbf{1}^T \lambda_1 = 1 \quad P\lambda_1 \succeq \nu\mathbf{1} \end{array}$$

  which is equivalent to second problem in matrix game
- weak duality confirms "player who knows other player's strategy has advantage or on par"
- moreoever, primal problem satisfies Slater's condition, hence *strong duality always holds*, and dual is feasible, hence $d^* = p^* \in \mathbf{R}$, *i.e.*, regardless of who knows other player's strategy, no player has advantage

Geometric interpretation of duality

- assume (not necessarily convex) optimization problem in Definition 6.36
- define graph

$$G = \{(q(x), h(x), f(x)) | x \in \mathcal{D}\} \subset \mathbf{R}^m \times \mathbf{R}^p \times \mathbf{R}$$

- for every $\lambda \succeq 0$ and ν

$$\begin{array}{rl} p^* = & \inf\{t | (u, v, t) \in G, u \preceq 0, v = 0\} \\ \geq & \inf\{t + \lambda^T u + \nu^T v | (u, v, t) \in G, u \preceq 0, v = 0\} \\ \geq & \inf\{t + \lambda^T u + \nu^T v | (u, v, t) \in G\} = g(\lambda, \nu) \end{array}$$

  where second inequality comes from $\{(u, v, t) | (u, v, t) \in G, u \preceq 0, v = 0\} \subset G$

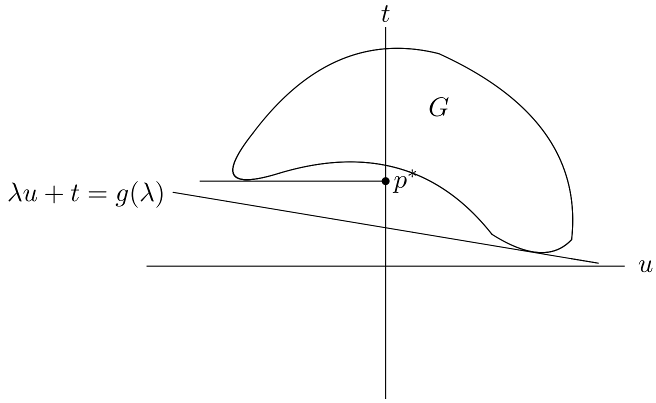

Figure 6.20: geometric interpretation of duality - 1

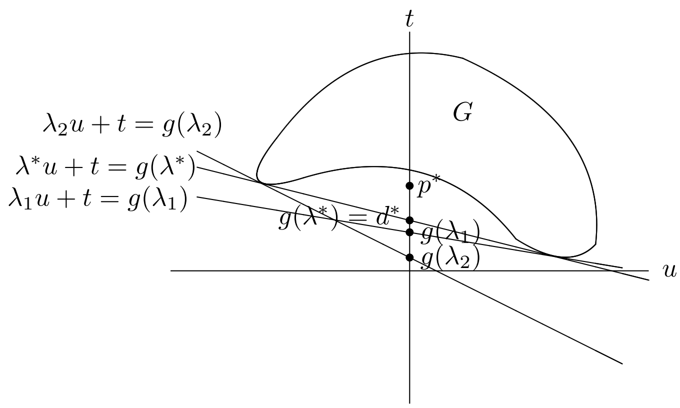

Figure 6.21: geometric interpretation of duality - 2

- above establishes *weak duality using graph*
- last equality implies that

$$(\lambda, \nu, 1)^T (u, v, t) \geq g(\lambda, \nu)$$

  hence if $g(\lambda, \nu) > -\infty$, $(\lambda, \nu, 1)$ and $g(\lambda, \nu)$ define *nonvertical supporting hyperplane* for G - nonvertical because third component is nonzero
- Figure 6.20 shows G as area inside closed curve contained in $\mathbf{R}^m \times \mathbf{R}^p \times \mathbf{R}$ where $m = 1$ and $p = 0$ as primal optimal value p^* and supporting hyperplane $\lambda u + t = g(\lambda)$
- Figure 6.21 shows three hyperplanes determined by three values for λ, one of which λ^* is optimal solution for dual problem

Epigraph interpretation of duality

- define extended graph over G - sort of epigraph of G

$$\begin{aligned} H &= G + \mathbf{R}_+^m \times \{0\} \times \mathbf{R}_+ \\ &= \{(u, v, t) | x \in \mathcal{D}, q(x) \preceq u, h(x) = v, f(x) \leq t\} \end{aligned}$$

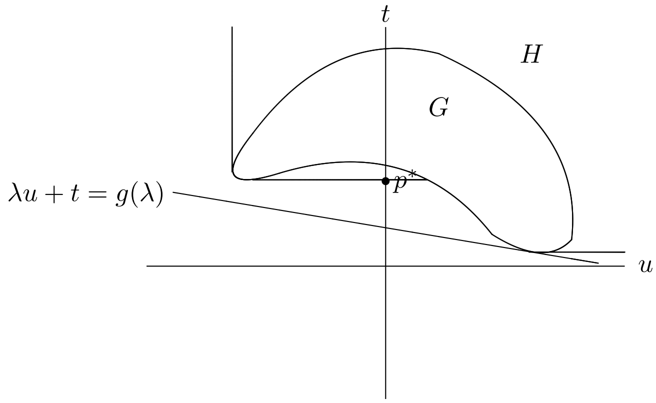

Figure 6.22: geometric interpretation of duality - 3

- if $\lambda \succeq 0$, $g(\lambda, \nu) = \inf\{(\lambda, \nu, 1)^T (u, v, t) | (u, v, t) \in H\}$, thus

$$(\lambda, \nu, 1)^T (u, v, t) \geq g(\lambda, \nu)$$

  defines nonvertical supporting hyperplane for H

- now $p^* = \inf\{t | (0, 0, t) \in H\}$, hence $(0, 0, p^*) \in \mathbf{bd}\, H$, hence

$$p^* = (\lambda, \nu, 1)^T (0, 0, p^*) \geq g(\lambda, \nu)$$

- once again establishes *weak duality*
- Figure 6.22 shows epigraph interpretation

Proof of strong duality under constraint qualification

- now we show proof of strong duality - this is one of rare cases where proof is shown in main slides instead of "selected proofs" section like Galois theory since - (I hope) it will give you some good intuition about why strong duality holds for (most) convex optimization problems
- assume Slater's condition holds, *i.e.*, f and q are convex, h is affine, and exists $x \in \mathcal{D}$ such that $q(x) \prec 0$ and $h(x) = 0$
- further assume $\mathcal{D}$ has interior (hence, $\mathbf{relint}\, \mathcal{D} = \mathcal{D}^\circ$ and $\mathbf{rank}\, A = p$
- assume $p^* \in \mathbf{R}$ - since exists feasible x, the other possibility is $p^* = -\infty$, but then, $d^* = -\infty$, hence strong duality holds
- H is convex (proof can be found in Proof 21)
- now define

$$B = \{(0, 0, s) \in \mathbf{R}^m \times \mathbf{R}^p \times \mathbf{R} | s < p^*\}$$

- then $B \cap H = \emptyset$, hence Theorem 6.1 implies exists separable hyperplane with $(\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ and α such that

$$\begin{aligned} (u, v, t) \in H \quad &\Rightarrow \quad \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha \\ (u, v, t) \in B \quad &\Rightarrow \quad \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha \end{aligned}$$

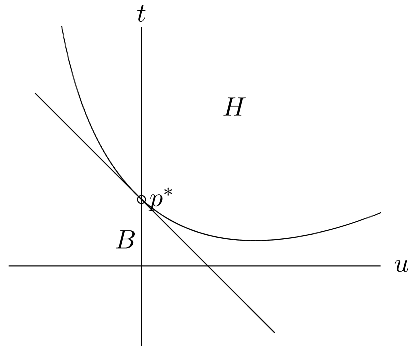

Figure 6.23: geometric interpretation of duality - 4

- then $\tilde{\lambda} \succeq 0$ & $\mu \geq 0$ - assume $\mu > 0$
  - can prove when $\mu = 0$, but kind of tedius, plus, whole purpose is provide good intuition, so will not do it here
- above second inequality implies $\mu p^* \leq \alpha$ and for some $x \in \mathcal{D}$

$$\mu L(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu) = \tilde{\lambda}^T q(x) + \tilde{\nu}^T h(x) + \mu f(x) \geq \alpha \geq \mu p^*$$

  thus,

$$g(\tilde{\lambda}/\mu, \tilde{\nu}/\mu) \geq p^*$$

- finally, weak duality implies

$$g(\lambda, \nu) = p^*$$

  where $\lambda = \tilde{\lambda}/\mu$ & $\nu = \tilde{\nu}/\mu$

Max-min characterization of weak and strong dualities

- note

$$\begin{aligned} \sup_{\lambda \geq 0, \nu} L(x, \lambda, \nu) &= \sup_{\lambda \geq 0, \nu} \left(f(x) + \lambda^T q(x) + \nu^T h(x) \right) \\ &= \begin{cases} f(x) & x \in \mathcal{F} \\ \infty & \text{otherwise} \end{cases} \end{aligned}$$

- thus $p^* = \inf_{x \in \mathcal{D}} \sup_{\lambda \succeq 0, \nu} L(x, \lambda, \nu)$ whereas $d^* = \sup_{\lambda \succeq 0, \nu} \inf_{x \in \mathcal{D}} L(x, \lambda, \nu)$
- weak duality means

$$\sup_{\lambda \succeq 0, \nu} \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \leq \inf_{x \in \mathcal{D}} \sup_{\lambda \succeq 0, \nu} L(x, \lambda, \nu)$$

- strong duality means

$$\sup_{\lambda \succeq 0, \nu} \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \sup_{\lambda \succeq 0, \nu} L(x, \lambda, \nu)$$

Max-min inequality

- indeed, inequality $\sup_{\lambda\succeq 0}\inf_{x\in\mathcal{D}} L(x,\lambda,\nu) \leq \inf_{x\in\mathcal{D}}\sup_{\lambda\succeq 0} L(x,\lambda,\nu)$ holds for general case

Inequality 6.2 (max-min inequality) *for $f : X \times Y \to \mathbf{R}$*

$$\sup_{y\in Y}\inf_{x\in X} f(x,y) \leq \inf_{x\in X}\sup_{y\in Y} f(x,y)$$

(proof can be found in Proof 22)

Definition 6.58 (strong max-min property) *if below equality holds, we say f (and X and Y) satisfies* strong max-min property *or* saddle point property

$$\sup_{y\in Y}\inf_{x\in X} f(x,y) = \inf_{x\in X}\sup_{y\in Y} f(x,y)$$

- this happens, *e.g.*, $X = \mathcal{D}$, $Y = \mathbf{R}_+^m \times \mathbf{R}^p$, f is Lagrangian of optimization problem (in Definition 6.36) for which strong duality holds

Saddle-points

Definition 6.59 (saddle-points) *for $f : X \times Y \to \mathbf{R}$, pair $x^* \in X$ and $y^* \in Y$ such that*

$$(\forall x \in X, y \in Y)\,(f(x^*,y) \leq f(x^*,y^*) \leq f(x,y^*))$$

called saddle-point for f (and X and Y)

- if assumption in Definition 6.59 holds, x^* minimizes $f(x,y^*)$ over X and y^* maximizes $f(x^*,y)$ over Y

$$\sup_{y\in Y} f(x^*,y) = f(x^*,y^*) = \inf_{x\in X} f(x,y^*)$$

  - strong max-min property (in Definition 6.58) holds with $f(x^*,y^*)$ as common value

Saddle-point interpretation of strong duality

- for primal optimum x^* and dual optimum (λ^*,ν^*)

$$g(\lambda^*,\nu^*) \leq L(x^*,\lambda^*,\nu^*) \leq f(x^*)$$

- if strong duality holds, for every $x \in \mathcal{D}$, $\lambda \succeq 0$, and ν

$$L(x^*,\lambda,\nu) \leq f(x^*) = L(x^*,\lambda^*,\nu^*) = g(\lambda^*,\nu^*) \leq L(x,\lambda^*,\nu^*)$$

  - thus x^* and (λ^*,ν^*) form saddle-point of Lagrangian

- conversely, if $\tilde{x}$ and $(\tilde{\lambda},\tilde{\nu})$ are saddle-point of Lagrangian, *i.e.*, for every $x \in \mathcal{D}$, $\lambda \succeq 0$, and ν

$$L(\tilde{x},\lambda,\nu) \leq L(\tilde{x},\tilde{\lambda},\tilde{\nu}) \leq L(x,\tilde{\lambda},\tilde{\nu})$$

  - hence $g(\tilde{\lambda},\tilde{\nu}) = \inf_{x\in\mathcal{D}} L(x,\tilde{\lambda},\tilde{\nu}) = L(\tilde{x},\tilde{\lambda},\tilde{\nu}) = \sup_{\lambda\succeq 0,\nu} L(\tilde{x},\lambda,\nu) = f(\tilde{x})$, thus $g(\lambda^*,\nu^*) \leq g(\tilde{\lambda},\tilde{\nu})$ & $f(\tilde{x}) \leq f(x^*)$
  - thus $\tilde{x}$ and $(\tilde{\lambda},\tilde{\nu})$ are primal and dual optimal

Game interpretation

- assume two players play zero-sum game with payment function $f : X \times Y \to \mathbf{R}$ where player A pays player B amount equal to $f(x, y)$ when player A chooses x and player B chooses y
- player A will try to minimize $f(x, y)$ and player B will try to maximize $f(x, y)$
- assume player A chooses first then player B chooses after learning opponent's choice
  - if player A chooses x, player B will choose $\mathrm{argsup}_{y \in Y} f(x, y)$
  - knowing that, player A will first choose $\mathrm{arginf}_{x \in X} \sup_{y \in Y} f(x, y)$
  - hence payment will be $\inf_{x \in X} \sup_{y \in Y} f(x, y)$
- if player B makes her choise first, opposite happens, *i.e.*, payment will be $\sup_{y \in Y} \inf_{x \in X} f(x, y)$
- max-min inequality of Ineq 6.2 says

$$\sup_{y \in Y} \inf_{x \in X} f(x, y) \leq \inf_{x \in X} \sup_{y \in Y} f(x, y)$$

  i.e., whowever chooses later has advantage, which is similar or rather same as matrix games using mixed strategies on page 172
- saddle-point for f (and X and Y), (x^*, y^*), called *solution of game* - x^* is optimal choice for player A and x^* is optimal choice for player B

Game interpretation for weak and strong dualities

- assume payment function in zero-sum game on page 178 is Lagrangian of optimization problem in Definition 6.36
- assume that $X = \mathcal{X}$ and $Y = \mathbf{R}_+^n \times \mathbf{R}^p$
- if player A chooses first, knowing that player B will choose $\mathrm{argsup}_{(\lambda,\nu) \in Y} L(x, \lambda, \nu)$, she will choose $x^* = \mathrm{arginf}_{x \in X} \sup_{(\lambda,\nu) \in Y} L(x, \lambda, \nu)$
- likewise, player B will choose $(\lambda^*, \nu^*) = \mathrm{argsup}_{(\lambda,\nu) \in Y} \inf_{x \in X} L(x, \lambda, \nu)$
- optimal dualtiy gap $p^* - d^*$ equals to advantage player who goes second has
- if strong dualtiy holds, (x^*, λ^*, ν^*) is solution of game, in which case no one has advantage

Certificate of suboptimality

- dual feasible point (λ, ν) degree of suboptimality of current solution
- assume x is feasible solution, then

$$f(x) - p^* \leq f(x) - g(\lambda, \nu)$$

  guarantees that $f(x)$ is no further than $\epsilon = f(x) - g(\lambda, \nu)$ from optimal point point x^* (even though we do not know optimal solution)
- for this reason, (λ, ν), called *certificate of suboptimality*
- x is ϵ-suboptimal for primal problem and (λ, ν) is ϵ-suboptimal for dual problem
- strong duality means we *could* find arbitrarily small certificate of suboptimality

Complementary slackness

- assume strong duality holds for optimization problem in Definition 6.36 and assume x^* is primal optimum and (λ^*, ν^*) is dual optimum, then

$$f(x^*) = L(x^*, \lambda^*, \nu^*) = f(x^*) + {\lambda^*}^T q(x^*) + {\nu^*}^T h(x^*)$$

- $h(x^*) = 0$ implies ${\lambda^*}^T q(x^*) = 0$
- then $\lambda^* \succeq 0$ and $q(x^*) \preceq 0$ imply
$$\lambda_i^* q_i(x^*) = 0 \quad i = 1, \ldots, m$$

Proposition 6.10 (complementary slackness) *when strong duality holds, for primal and dual optimal points x^* and (λ^*, ν)*

$$\lambda_i^* q_i(x^*) = 0 \quad i = 1, \ldots, m$$

this property, called complementary slackness

KKT optimality conditions

Definition 6.60 (KKT optimality conditions) *for optimization problem in Definition 6.36 where f, q, and h are all differentiable, below conditions for $x \in \mathcal{D}$ and $(\lambda, \nu) \in \mathbf{R}^m \times \mathbf{R}^p$*

$$\begin{array}{rcll} q(x) & \preceq & 0 & \text{- \textit{primal feasibility}} \\ h(x) & = & 0 & \text{- \textit{primal feasibility}} \\ \lambda & \succeq & 0 & \text{- \textit{dual feasibility}} \\ \lambda^T q(x) & = & 0 & \text{- \textit{complementary slackness}} \\ \nabla_x L(x, \lambda, \nu) & = & 0 & \text{- \textit{vanishing gradient of Lagrangian}} \end{array}$$

called Karush-Kuhn-Tucker (KKT) optimality conditions

KKT necessary for optimality with strong duality

Theorem 6.9 (KKT necessary for optimality with strong duality) *for optimization problem in Definition 6.36 where f, q, and h are all differentiable, if strong duality holds, primal and dual optimal solutions x^* and (λ^*, ν) satisfy KKT optimality conditions (in Definition 6.60), i.e., for every optimization problem*

- *when strong duality holds, KKT optimality conditions are necessary for primal and dual optimality*

  or equivalently

- *primal and dual optimality with strong duality imply KKT optimality conditions*

KKT and convexity sufficient for optimality with strong duality

- assume convex optimization problem where f, q, and h are all differentiable and $x \in \mathcal{D}$ and $(\lambda, \nu) \in \mathbf{R}^m \times \mathbf{R}^p$ satisfying KKT conditions, *i.e.*

$$q(x) \preceq 0,\ h(x) = 0,\ \lambda \succeq 0,\ \lambda^T q(x) = 0,\ \nabla_x L(x, \lambda, \nu) = 0$$

- since $L(x, \lambda, \nu)$ is convex for $\lambda \succeq 0$, *i.e.*, each of $f(x)$, $\lambda^T q(x)$, and $\nu^T h(x)$ is convex, vanishing gradient implies x achieves infimum for Lagrangian, hence

$$g(\lambda, \nu) = L(x, \lambda, \nu) = f(x) + \lambda^T q(x) + \nu^T h(x) = f(x)$$

- thus, strong duality holds, *i.e.*, x and (λ, ν) are primal and dual optimal solutions with zero duality gap

Theorem 6.10 (KKT and convexity sufficient for optimality with strong duality) *for convex optimization problem in Definition 6.40 where f, q, and h are all differentiable, if $x \in \mathcal{D}$ and $(\lambda, \nu) \in \mathbf{R}^m \times \mathbf{R}^p$ satisfy KKT optimality conditions (in Definition 6.60), they are primal and dual optimal solutions having zero duality gap i.e.*

- *for convex optimization problem, KKT optimality conditions are sufficient for primal and dual optimality with strong duality*

  or equivalently

- *KKT optimality conditions and convexity imply primal and dual optimality and strong duality*

- Theorem 6.9 together with Theorem 6.10 implies that for convex optimization problem
  - *KKT optimality conditions are necessary and sufficient for primal and dual optimality with strong duality*

Solving primal problems via dual problems

- when strong duality holds, can retrieve primal optimum from dual optimum since primal optimal solution is minimize of

$$L(x, \lambda^*, \nu^*)$$

  where (λ^*, ν^*) is dual optimum

- example - entropy maximization ($\mathcal{D} = \mathbf{R}^n_{++}$)
  - primal problem - min. $f(x) = \sum_{i=1}^n x_i \log x_i$ s.t. $Ax \preceq b$, $\sum x = 1$
  - dual problem - max. $-b^T\lambda - \nu - \exp(-\nu - 1) \sum \exp(A^T\lambda)$ s.t. $\lambda \succeq 0$
  - provided dual optimum (λ^*, ν^*), primal optimum is

$$x^* = \operatorname*{argmin}_{x \in \mathcal{D}} \left(\sum x_i \log x_i + {\lambda^*}^T (Ax - b) + \nu^* (\mathbf{1}^T x - 1) \right)$$

  - $\nabla_x L(x, \lambda^*, \nu^*) = \log x + A^T\lambda^* + (1 + \nu^*)\mathbf{1}$, hence

$$x^* = \exp(-(A^T\lambda^* + (1 + \nu^*)\mathbf{1}))$$

Perturbed optimization problems

- original problem in Definition 6.36 with perturbed constraints

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & q(x) \preceq u \\ & h(x) = v \end{array}$$

  where $u \in \mathbf{R}^m$ and $v \in \mathbf{R}^p$

- define $p^*(u, v)$ as optimal value of above *perturbed* problem, *i.e.*

$$p^*(u, v) = \inf\{f(x) | x \in \mathcal{D}, q(x) \preceq u, h(x) = v\}$$

  which is convex when problem is convex optimization problem (proof can be found in Proof 21) - note $p^*(0, 0) = p^*$

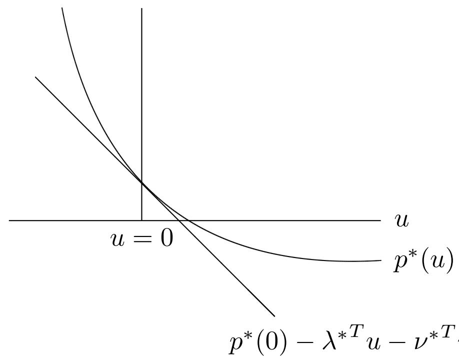

Figure 6.24: sensitivity analysis of optimal value

- assume and dual optimum (λ^*, ν^*), if strong duality holds, for every feasible x for perturbed problem

$$p^*(0,0) = g(\lambda^*, \nu^*) \leq f(x) + {\lambda^*}^T q(x) + {\nu^*}^T h(x) \leq f(x) + {\lambda^*}^T u + {\nu^*}^T v$$

  thus

$$p^*(0,0) \leq p^*(u,v) + {\lambda^*}^T u + {\nu^*}^T v$$

  hence

$$p^*(u,v) \geq p^*(0,0) - {\lambda^*}^T u - {\nu^*}^T v$$

- Figure 6.24 shows this for optimization problem with one inequality constraint and no equality constraint

Global sensitivity analysis via perturbed problems

- recall

$$p^*(u,v) \geq p^*(0,0) - {\lambda^*}^T u - {\nu^*}^T v$$

- interpretations
  - if λ_i^* is large, when i-th inequality constraint is tightened, optimal value increases a lot
  - if λ_i^* is small, when i-th inequality constraint is relaxed, optimal value decreases not a lot
  - if $|\nu_i^*|$ is large, reducing v_i when $\nu_i^* > 0$ or increasing v_i when $\nu_i^* < 0$ increases optimval value a lot
  - if $|\nu_i^*|$ is small, increasing v_i when $\nu_i^* > 0$ or decreasing v_i when $\nu_i^* < 0$ decreases optimval value not a lot
- it only gives lower bounds - will explore local behavior

Local sensitivity analysis via perturbed problems

- assume $p^*(u,v)$ is differentiable with respect to u and v, *i.e.*, $\nabla_{(u,v)} p^*(u,v)$ exist
  - then

$$\frac{\partial}{\partial u_i} p^*(0,0) = \lim_{h \to 0^+} \frac{p^*(he_i, 0) - p^*(0,0)}{h} \geq \lim_{h \to 0^+} \frac{-{\lambda^*}^T (he_i)}{h} = -\lambda_i$$

    and

$$\frac{\partial}{\partial u_i} p^*(0,0) = \lim_{h \to 0^-} \frac{p^*(he_i, 0) - p^*(0,0)}{h} \leq \lim_{h \to 0^-} \frac{-{\lambda^*}^T (he_i)}{h} = -\lambda_i$$

- obtain same result for v_i, hence

$$\nabla_u\, p^*(0,0) = -\lambda \quad \nabla_v\, p^*(0,0) = -\nu$$

- so larger λ_i or $|\nu_i|$ means larger change in optimal value of perturbed problem when u_i or v_i change a bit and vice versa quantitatively, - λ_i an ν_i provide exact ratio and direction

Different dual problems for equivalent optimization problems - 1

- introducing new variables and equality constraints for unconstrained problems
  - unconstrained optimization problem

$$\text{minimize} \quad f(Ax+b)$$

    - dual Lagrange function is $g = p^*$, hence strong duality holds, which, however, does not provide useful information
  - reformulate as equivalent optimization problem

$$\begin{array}{ll} \text{minimize} & f(y) \\ \text{subject to} & Ax+b=y \end{array}$$

    - Lagrangian - $L(x,y,\nu) = f(y) + \nu^T(Ax+b-y)$
    - Lagrange dual function - $g(\nu) = -I(A^T\nu = 0) + b^T\nu - f^*(\nu)$
    - dual optimization problem

$$\begin{array}{ll} \text{maximize} & b^T\nu - f^*(\nu) \\ \text{subject to} & A^T\nu = 0 \end{array}$$

- examples
  - unconstrained geometric problem

$$\text{minimize} \quad \log\left(\sum_{i=1}^m \exp(a_i^T x + b_i)\right)$$

    - reformulation

$$\begin{array}{ll} \text{minimize} & \log\left(\sum_{i=1}^m \exp(y_i)\right) \\ \text{subject to} & Ax+b=y \end{array}$$

    - dual optimization problem

$$\begin{array}{ll} \text{maximize} & b^T\nu - \sum_{i=1}^m \nu_i \log \nu_i \\ \text{subject to} & \mathbf{1}^T\nu = 1 \\ & A^T\nu = 0 \\ & \nu \succeq 0 \end{array}$$

    which is entropy maximization problem
  - norm minimization problem

$$\text{minimize} \quad \|Ax-b\|$$

    - reformulation

$$\begin{array}{ll} \text{minimize} & \|y\| \\ \text{subject to} & Ax-b=y \end{array}$$

    - dual optimization problem

$$\begin{array}{ll} \text{maximize} & b^T\nu \\ \text{subject to} & \|\nu\|_* \le 1 \\ & A^T\nu = 0 \end{array}$$

Different dual problems for equivalent optimization problems - 2

- introducing new variables and equality constraints for constrained problems
  - inequality constrained optimization problem

$$\begin{array}{ll} \text{minimize} & f_0(A_0x+b_0) \\ \text{subject to} & f_i(A_ix+b_i) \leq 0 \quad i=1,\ldots,m \end{array}$$

  - reformulation

$$\begin{array}{ll} \text{minimize} & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0 \quad i=1,\ldots,m \\ & A_ix+b_i = y_i \quad i=0,\ldots,m \end{array}$$

  - dual optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{i=0}^m \nu_i^T b_i - f_0^*(\nu_0) - \sum_{i=1}^m \lambda_i f_i^*(\nu_i/\lambda_i) \\ \text{subject to} & \sum_{i=0}^m A_i^T \nu_i = 0 \\ & \lambda \succeq 0 \end{array}$$

- examples
  - inequality constrained geometric program

$$\begin{array}{ll} \text{minimize} & \log\left(\sum \exp(A_0x+b_0)\right) \\ \text{subject to} & \log\left(\sum \exp(A_ix+b_i)\right) \leq 0 \quad i=1,\ldots,m \end{array}$$

    where $A_i \in \mathbf{R}^{K_i\times n}$ and $\exp(z) := (\exp(z_1),\ldots,\exp(z_k))) \in \mathbf{R}^n$ and $\sum z := \sum_{i=1}^k z_i \in \mathbf{R}$ for $z \in \mathbf{R}^k$
    - reformulation

$$\begin{array}{ll} \text{minimize} & \log\left(\sum \exp(y_0)\right) \\ \text{subject to} & \log\left(\sum \exp(y_i)\right) \leq 0 \quad i=1,\ldots,m \\ & A_ix+b_i = y_i \quad i=0,\ldots,m \end{array}$$

    - dual optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{i=0}^m b_i^T \nu_i - \nu_0^T \log(\nu_0) - \sum_{i=1}^m \nu_i^T \log(\nu_i/\lambda_i) \\ \text{subject to} & \nu_i \succeq 0 \quad i=0,\ldots,m \\ & \mathbf{1}^T\nu_0 = 1,\ \mathbf{1}^T\nu_i = \lambda_i \quad i=1,\ldots,m \\ & \lambda_i \geq 0 \quad i=1,\ldots,m \\ & \sum_{i=0}^m A_i^T \nu_i = 0 \end{array}$$

    where and $\log(z) := (\log(z_1),\ldots,\log(z_k))) \in \mathbf{R}^n$ for $z \in \mathbf{R}_{++}^k$
    - simplified dual optimization problem

$$\begin{array}{ll} \text{maximize} & \sum_{i=0}^m b_i^T \nu_i - \nu_0^T \log(\nu_0) - \sum_{i=1}^m \nu_i^T \log(\nu_i/\mathbf{1}^T\nu_i) \\ \text{subject to} & \nu_i \succeq 0 \quad i=0,\ldots,m \\ & \mathbf{1}^T\nu_0 = 1 \\ & \sum_{i=0}^m A_i^T \nu_i = 0 \end{array}$$

Different dual problems for equivalent optimization problems - 3

- transforming objectives
  - norm minimization problem
$$\text{minimize} \quad \|Ax - b\|$$
  - reformulation
$$\begin{array}{ll} \text{minimize} & (1/2)\|y\|^2 \\ \text{subject to} & Ax - b = y \end{array}$$
  - dual optimization problem
$$\begin{array}{ll} \text{maximize} & -(1/2)\|\nu\|_*^2 + b^T\nu \\ \text{subject to} & A^T\nu = 0 \end{array}$$

Different dual problems for equivalent optimization problems - 4

- making contraints implicit
  - LP with box constraints
$$\begin{array}{ll} \text{minimize} & c^Tx \\ \text{subject to} & Ax = b, \ l \preceq x \preceq u \end{array}$$
  - dual optimization problem
$$\begin{array}{ll} \text{maximize} & -b^T\nu - \lambda_1^T u + \lambda_2^T l \\ \text{subject to} & A^T\nu + \lambda_1 - \lambda_2 + c = 0, \ \lambda_1 \succeq 0, \ \lambda_2 \succeq 0 \end{array}$$
  - reformulation
$$\begin{array}{ll} \text{minimize} & c^Tx + I(l \preceq x \preceq u) \\ \text{subject to} & Ax = b \end{array}$$
  - dual optimization problem for reformulated primal problem
$$\text{maximize} \quad -b^T\nu - u^T(A^T\nu + c)^- + l^T(A^T\nu + c)^+$$

6.5 Theorems of Alternatives

Weak alternatives

Theorem 6.11 (weak alternatives of two systems) *for $q : Q \to \mathbf{R}^m$ & $h : H \to \mathbf{R}^p$ where Q and H are subsets of common set X, which is subset of Banach space, assuming $\mathcal{D} = Q \cap H \neq \emptyset$, and $\lambda \in \mathbf{R}^m$ & $\nu \in \mathbf{R}^p$, below two systems of inequalities and equalities are weak alternatives, i.e., at most one of them is feasible*
$$q(x) \preceq 0 \quad h(x) = 0$$
and
$$\lambda \succeq 0 \quad \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right) > 0$$

- can prove Theorem 6.11 using duality of optimization problems
- consider primal and dual problems
  - primal problem
$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & q(x) \preceq 0 \\ & h(x) = 0 \end{array}$$

– dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

where

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right)$$

- then $p^*,\ d^* \in \{0, \infty\}$
- now assume *first system of Theorem 6.11 is feasible,* then $p^* = 0$, hence weak duality applies $d^* = 0$, thus there exist no λ and ν such that $\lambda \succeq 0$ and $g(\lambda, \nu) > 0$ *i.e., second system is infeasible,* since otherwise there exist λ and ν making $g(\lambda, \nu)$ arbitrarily large; if $\tilde{\lambda} \succeq 0$ and $\tilde{\nu}$ satisfy $g(\lambda, \nu) > 0$, $g(\alpha\tilde{\lambda}, \alpha\tilde{\nu}) = \alpha g(\tilde{\lambda}, \tilde{\nu})$ goes to ∞ when $\alpha \to \infty$
- assume *second system is feasible,* then $g(\lambda, \nu)$ can be arbitrarily large for above reasons, thus $d^* = \infty$, hence weak duality implies $p^* = \infty$, which implies *first system is infeasible*
- therefore two systems are weak alternatives; at most one of them is feasible

  (actually, not hard to prove it without using weak duality)

Weak alternatives with strict inequalities

Theorem 6.12 (weak alternatives of two systems with strict inequalities) *for $q : Q \to \mathbf{R}^m$ & $h : H \to \mathbf{R}^p$ where Q and H are subsets of common set X, which is subset of Banach space, assuming $\mathcal{D} = Q \cap H \neq \emptyset$, and $\lambda \in \mathbf{R}^m$ & $\nu \in \mathbf{R}^p$, below two systems of inequalities and equalities are weak alternatives, i.e., at most one of them is feasible*

$$q(x) \prec 0 \quad h(x) = 0$$

and

$$\lambda \succeq 0 \quad \lambda \neq 0 \quad \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right) \geq 0$$

Strong alternatives

Theorem 6.13 (strong alternatives of two systems) *for convex $q : Q \to \mathbf{R}^m$ & affine $h : H \to \mathbf{R}^p$ where Q and H are subsets $\mathbf{R}^n$ assuming $\mathcal{D} = Q \cap H \neq \emptyset$ and $\lambda \in \mathbf{R}^m$ & $\nu \in \mathbf{R}^p$, if exists $x \in \mathbf{relint}\, \mathcal{D}$ with $h(x) = 0$, below two systems of inequalities and equalities are strong alternatives, i.e., exactly one of them is feasible*

$$q(x) \preceq 0 \quad h(x) = 0$$

and

$$\lambda \succeq 0 \quad \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right) > 0$$

Strong alternatives with strict inequalities

Theorem 6.14 (strong alternatives of two systems with strict inequalities) *for convex $q : Q \to \mathbf{R}^m$ & affine $h : H \to \mathbf{R}^p$ where Q and H are subsets $\mathbf{R}^n$ assuming $\mathcal{D} = Q \cap H \neq \emptyset$ and $\lambda \in \mathbf{R}^m$ & $\nu \in \mathbf{R}^p$, if exists $x \in \mathbf{relint}\, \mathcal{D}$ with $h(x) = 0$, below two systems of inequalities and equalities are strong alternatives, i.e., exactly one of them is feasible*

$$q(x) \prec 0 \quad h(x) = 0$$

and

$$\lambda \succeq 0 \quad \lambda \neq 0 \quad \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right) \geq 0$$

- proof - consider convex optimization problem and its dual
  - primal problem

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & q(x) - s\mathbf{1} \preceq 0 \\ & h(x) = 0 \end{array}$$

  - dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda,\nu) \\ \text{subject to} & \lambda \succeq 0 \quad \mathbf{1}^T\lambda = 1 \end{array}$$

    where $g(\lambda,\nu) = \inf_{x\in\mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right)$

- first observe Slater's condition holds for primal problem since by hypothesis of Theorem 6.14, exists $y \in \mathbf{relint}\, \mathcal{D}$ with $h(y) = 0$, hence $(y, q(y)) \in Q \times \mathbf{R}$ is primal feasible satisifying Slater's condition
- hence Slater's theorem (Theorem 6.8) implies $d^* = p^*$
- assume first system is feasible, then primal problem is strictly feasible and $d^* = p^* < 0$, hence second system infeasible since otherwise feasible point for second system is feasible point of dual problem, hence $d^* \geq 0$
- assume first system is infeasible, then $d^* = p^* \geq 0$, hence Slater's theorem (Theorem 6.8) implies exists dual optimal (λ^*, ν^*) (whether or not $d^* = \infty$), hence (λ^*, ν^*) is feasible point for second system of Theorem 6.14
- therefore two systems are strong alternatives; each is feasible *if and only if* the other is infeasible

Strong alternatives for linear inequalities

- dual function of feasibility problem for $Ax \preceq b$ is

$$g(\lambda) = \inf_{x\in\mathbf{R}^n} \lambda^T(Ax - b) = \left\{ \begin{array}{ll} -b^T\lambda & A^T\lambda = 0 \\ -\infty & \text{otherwise} \end{array} \right.$$

- hence alternative system is $\lambda \succeq 0,\ b^T\lambda < 0,\ A^T\lambda = 0$
- thus Theorem 6.13 implies below systems are strong alternatives

$$Ax \preceq b \quad \& \quad \lambda \succeq 0 \quad b^T\lambda < 0 \quad A^T\lambda = 0$$

- similarly alternative system is $\lambda \succeq 0,\ b^T\lambda < 0,\ A^T\lambda = 0$ and Theorem 6.13 implies below systems are strong alternatives

$$Ax \prec b \quad \& \quad \lambda \succeq 0 \quad \lambda \neq 0 \quad b^T\lambda \leq 0 \quad A^T\lambda = 0$$

Farkas' lemma

Theorem 6.15 (Farkas' lemma) *below systems of inequalities and equalities are strong alternatives*

$$Ax \preceq 0 \quad c^Tx < 0 \quad \& \quad A^Ty + c = 0 \quad y \succeq 0$$

- will prove Theorem 6.15 using LP and its dual
- consider LP $\left(\text{minimize } c^Tx \quad \text{subject to } Ax \preceq 0\right)$
- dual function is $g(y) = \inf_{x\in\mathbf{R}^n} \left(c^Tx + y^TAx\right) = \left\{ \begin{array}{ll} 0 & A^Ty + c = 0 \\ -\infty & \text{otherwise} \end{array} \right.$

- hence dual problem is $\left(\text{maximize } 0 \quad \text{subject to } A^T y + c = 0,\ y \succeq 0\right)$
- assume first system is feasible, then homogeneity of primal problem implies $p^* = -\infty$, thus d^*, *i.e.*, dual is infeasible, hence second system is infeasible
- assume first system is infeasible, since primal is always feasible, $p^* = 0$, hence strong duality implies $d^* = 0$, thus second system is feasible

6.6 Convex Optimization with Generalized Inequalities

Optimization problems with generalized inequalities

Definition 6.61 (optimization problems with generalized inequalities) *for $f : F \to \mathbf{R}$, $q : Q \to \bigtimes_{i=1}^{m} \mathbf{R}^{k_i}$, $h : H \to \mathbf{R}^p$ where F, Q, and H are subsets of common set X*

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & q(x) \preceq_{\mathcal{K}} 0 \\ & h(x) = 0 \end{array}$$

called optimization problem with generalized inequalities where $\mathcal{K} = \bigtimes K_i$ is proper cone with m proper cones $K_1 \subset \mathbf{R}^{k_1}, \ldots, K_n \subset \mathbf{R}^{k_m}$

- *every terminology and associated notation is same as of optimization problem in Definition 6.36 such as objective & inequality & equality contraint functions, domain of optimization problem $\mathcal{D}$, feasible set $\mathcal{F}$, optimal value p^**
- *note that when $K_i = \mathbf{R}_+$ (hence $\mathcal{K} = \mathbf{R}_+^m$), above optimization problem coincides with that in Definition 6.36, i.e., optimization problems with generalized inequalities subsume (normal) optimization problems*

Lagrangian for generalized inequalities

Definition 6.62 (Lagrangian for generalized inequalities) *for optimization problem in Definition 6.61 with nonempty domain $\mathcal{D}$, function $L : \mathcal{D} \times \bigtimes_{i=1}^{m} \mathbf{R}^{k_i} \times \mathbf{R}^p \to \mathbf{R}$ defined by*

$$L(x, \lambda, \nu) = f(x) + \lambda^T q(x) + \nu^T h(x)$$

called Lagrangian associated with the optimization problem where

- *every terminology and associated notation is same as of optimization problem in Definition 6.52 such as dual variables or Lagrange multipliers λ and ν.*
- *Lagrangian for generalized inequalities subsumes (normal) Lagrangian (Definition 6.52)*

Lagrange dual functions for generalized inequalities

Definition 6.63 (Lagrange dual functions for generalized inequalities) *for optimization problem in Definition 6.61 for which Lagrangian is defined, function $g : \bigtimes \mathbf{R}^{k_i} \times \mathbf{R}^p \to \mathbf{R} \cup \{-\infty\}$ defined by*

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f(x) + \lambda^T q(x) + \nu^T h(x)\right)$$

called Lagrange dual function or just dual function associated with optimization problem

- *Lagrange dual functions for generalized inequalities subsume (normal) Lagrange dual functions (Definition 6.53)*

- g is *concave function*
- $g(\lambda, \nu)$ is lower bound for optimal value of associated optimization problem *i.e.*,

  $$g(\lambda, \nu) \leq p^*$$

  for every $\lambda \succeq_{\mathcal{K}}^* 0$ where $\mathcal{K}^*$ denotes dual cone of $\mathcal{K}$, *i.e.*, $\mathcal{K}^* = \bigtimes K_i^*$ where $K_i^* \subset \mathbf{R}^{k_i}$ is dual cone of $K_i \subset \mathbf{R}^{k_i}$
- (λ, ν) with $\lambda \succeq_{\mathcal{K}} 0$ and $g(\lambda, \nu) > -\infty$ said to be *dual feasible*

Lagrange dual problems for generalized inequalities

Definition 6.64 (Lagrange dual problems for generalized inequalities) *for optimization problem in Definition 6.61, optimization problem*

$$\begin{array}{ll} \textit{maximize} & g(\lambda, \nu) \\ \textit{subject to} & \lambda \succeq_{\mathcal{K}^*} 0 \end{array}$$

where $\mathcal{K}^*$ denotes dual cone of $\mathcal{K}$, i.e., $\mathcal{K}^* = \bigtimes K_i^*$ where $K_i^* \subset \mathbf{R}^{k_i}$ is dual cone of $K_i \subset \mathbf{R}^{k_i}$, called Lagrange dual problem *associated with problem in Definition 6.61*

- *every terminology and related notation is same as that in Definition 6.54 such as dual feasibility, dual optimal value d^*, optimal Lagrange multipliers (λ^*, ν^*)*
- *Lagrange dual problems for generalized inequalities subsume (normal) Lagrange dual problems (Definition 6.54)*

- Lagrange dual problem in Definition 6.64 is convex optimization since $g(\lambda, \nu)$ is convex

Slater's theorem for generalized inequalities

Theorem 6.16 (Slater's theorem for generalized inequalities) *if optimization problem in Definition 6.61 is convex, i.e., f is convex, q is $\mathcal{K}$-convex (i.e., every q_i is K_i-convex) (Definition 6.34), and exists feasible $x \in \mathcal{D}$ contained in* **relint** $\mathcal{D}$ *such that*

$$q(x) \prec_{\mathcal{K}} 0 \quad h(x) = 0$$

strong duality *holds (and* dual optimal value is attained when $d^* > -\infty$*)*

- *such condition, called* Slater's condition
- *such point, (sometimes) said to be* strictly feasible
- *note resemblance with Slater's theorem in Theorem 6.8*

Duality for SDP

- (inequality form) SDP

  $$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots + x_n F_n + G \preceq 0 \end{array}$$

  where $F_1, \ldots, F_n, G \in \mathbf{S}^k$ and $\mathcal{K} = \mathbf{S}_+^k$
- Lagrangian

  $$L(x, Z) = c^T x + (x_1 F_1 + \cdots + x_n F_n + G) \bullet Z = \sum x_i (F_i \bullet Z + c_i) + G \bullet Z$$

  where $X \bullet Y = \mathbf{Tr}\, XY$ for $X, Y \in \mathbf{S}^k$

- Lagrange dual function

$$g(Z) = \inf_{x \in \mathbf{R}^n} L(x, Z) = \begin{cases} G \bullet Z & F_i \bullet Z + c_i = 0 \quad i = 1, \ldots, n \\ -\infty & \text{otherwise} \end{cases}$$

- Lagrange dual problem

$$\begin{array}{ll} \text{maximize} & G \bullet Z \\ \text{subject to} & F_i \bullet Z + c_i = 0 \quad i = 1, \ldots, n \\ & Z \succeq 0 \end{array}$$

  where fact that $\mathbf{S}_+^k$ is self-dual, *i.e.*, $\mathcal{K}^* = \mathcal{K}$

- Slater's theorem (Theorem 6.16) implies if primal problem is strictly feasible, *i.e.*, exists $x \in \mathbf{R}^n$ such that $\sum x_i F_i + G \prec 0$, strong duality holds

KKT optimality conditions for generalized inequalities

Definition 6.65 (KKT optimality conditions for generalized inequalities) *for optimization problem in Definition 6.61 where f, q, and h are all differentiable, below conditions for $x \in \mathcal{D}$ and $(\lambda, \nu) \in \bigtimes \mathbf{R}^{k_i} \times \mathbf{R}^p$*

$$\begin{array}{rcll} q(x) & \preceq_{\mathcal{K}} & 0 & \text{- primal feasibility} \\ h(x) & = & 0 & \text{- primal feasibility} \\ \lambda & \succeq_{\mathcal{K}^*} & 0 & \text{- dual feasibility} \\ \lambda^T q(x) & = & 0 & \text{- complementary slackness} \\ \nabla_x L(x, \lambda, \nu) & = & 0 & \text{- vanishing gradient of Lagrangian} \end{array}$$

called Karush-Kuhn-Tucker (KKT) optimality conditions

- *note KKT optimality conditions for generalized inequalities subsume (normal) KKT optimality conditions (Definition 6.60)*

KKT conditions and optimalities for generalized inequalities

- for every optimization problem with generalized inequalities (Definition 6.61), every statement for normal optimization problem (Definition 6.36), regarding relations among KKT conditions, optimality, primal and dual optimality, and strong duality, is *exactly the same*
  - for every optimization problem with generalized inequalities (Definition 6.61)
    - if strong duality holds, primal and dual optimal points satisfy KKT optimality conditions in Definition 6.65 (same as Theorem 6.9)
    - if optimization problem is convex and primal and dual solutions satisfy KKT optimality conditions in Definition 6.65, the solutions are optimal with strong duality (same as Theorem 6.10)
    - therefore, for convex optimization problem, *KKT optimality conditions are necessary and sufficient for primal and dual optimality with strong duality*

Perturbation and sensitivity analysis for generalized inequalities

- original problem in Definition 6.61 with perturbed constraints

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & q(x) \preceq_{\mathcal{K}} u \\ & h(x) = v \end{array}$$

  where $u \in \mathbf{R}^m$ and $v \in \mathbf{R}^p$

- define $p^*(u,v) = p^*(u,v) = \inf\{f(x)|x \in \mathcal{D}, q(x) \preceq u, h(x) = v\}$, which is convex when problem is convex optimization problem - note $p^*(0,0) = p^*$
- as for normal optimization problem case (page 180), if and dual optimum (λ^*, ν^*), if strong duality holds,
$$p^*(u,v) \geq p^*(0,0) - {\lambda^*}^T u - {\nu^*}^T v$$
and
$$\nabla_u\, p^*(0,0) = -\lambda \quad \nabla_v\, p^*(0,0) = -\nu$$

Sensitivity analysis for SDP

- assume inequality form SDP and its dual problem on page 188 and page 189
- consider perturbed SDP
$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots + x_n F_n + G \preceq U \end{array}$$
for some $U \in \mathbf{S}^k$
  - define $p^* : \mathbf{S}^k \to \mathbf{R}$ such that $p^*(U)$ is optimal value of above problem
- assume $x^* \in \mathbf{R}^n$ and $Z^* \in \mathbf{S}^k_+$ are primal and dual optimum with zero dualty gap
- then
$$p^*(U) \geq p^* - Z^* \bullet U$$
- if $\nabla_U p^*$ exists at $U = 0$
$$\nabla_U p^*(0) = -Z^*$$

Weak alternatives for generalized inequalities

Theorem 6.17 (weak alternatives for generalized inequalities) *for $q : Q \to \bigtimes \mathbf{R}^{k_i}$ & $h : H \to \mathbf{R}^p$ where Q and H are subsets of common Banach space assuming $\mathcal{D} = Q \cap H \neq \emptyset$, and $\lambda \in \bigtimes \mathbf{R}^{k_i}$ & $\nu \in \mathbf{R}^p$, below pairs of systems are strong alternatives*
$$\begin{array}{lllll} q(x) \preceq_{\mathcal{K}} 0 & h(x) = 0 & \& & \lambda \succeq_{\mathcal{K}^*} 0 & g(\lambda,\nu) > 0 \\ q(x) \prec_{\mathcal{K}} 0 & h(x) = 0 & \& & \lambda \succeq_{\mathcal{K}^*} 0 & \lambda \neq 0 \quad g(\lambda,\nu) \geq 0 \end{array}$$
where $\mathcal{K} = \bigtimes K_i$ with proper cones $K_i \subset \mathbf{R}^{k_i}$ and function $g : \bigtimes \mathbf{R}^{k_i} \times \mathbf{R}^p \to \mathbf{R}$ defined by
$$g(\lambda,\nu) = \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right)$$

note this theorem subsumes Theorem 6.11 and Theorem 6.12

Strong alternatives for generalized inequalities

Theorem 6.18 (strong alternatives for generalized inequalities) *for $\mathcal{K}$-convex $q : Q \to \bigtimes \mathbf{R}^{k_i}$ & affine $h : H \to \mathbf{R}^p$ where Q and H are subsets of $\mathbf{R}^n$ assuming $\mathcal{D} = Q \cap H \neq \emptyset$, and $\lambda \in \bigtimes \mathbf{R}^{k_i}$ & $\nu \in \mathbf{R}^p$, if exists $x \in$ **relint** $\mathcal{D}$ with $h(x) = 0$, below pairs of systems are strong alternatives*
$$\begin{array}{lllll} q(x) \preceq_{\mathcal{K}} 0 & h(x) = 0 & \& & \lambda \succeq_{\mathcal{K}^*} 0 & g(\lambda,\nu) > 0 \\ q(x) \prec_{\mathcal{K}} 0 & h(x) = 0 & \& & \lambda \succeq_{\mathcal{K}^*} 0 & \lambda \neq 0 \quad g(\lambda,\nu) \geq 0 \end{array}$$
where $\mathcal{K} = \bigtimes K_i$ with proper cones $K_i \subset \mathbf{R}^{k_i}$ and function $g : \bigtimes \mathbf{R}^{k_i} \times \mathbf{R}^p \to \mathbf{R}$ defined by
$$g(\lambda,\nu) = \inf_{x \in \mathcal{D}} \left(\lambda^T q(x) + \nu^T h(x)\right)$$

note this theorem subsumes Theorem 6.13 and Theorem 6.14

Strong alternatives for SDP

- for $F_1, \ldots, F_n, G \in \mathbf{S}^k$, $x \in \mathbf{R}^n$, and $Z \in \mathbf{S}^k$
  - below systems are strong alternatives
  $$x_1 F_1 + \cdots + x_n F_n + G \prec 0$$
  and
  $$Z \succeq 0 \quad Z \neq 0 \quad G \bullet Z \geq 0 \quad F_i \bullet Z = 0 \; i = 1, \ldots, n$$
  - if $\sum v_i F_i \succeq 0 \Rightarrow \sum v_i F_i = 0$, below systems are strong alternatives
  $$x_1 F_1 + \cdots + x_n F_n + G \preceq 0$$
  and
  $$Z \succeq 0 \quad G \bullet Z > 0 \quad F_i \bullet Z = 0 \; i = 1, \ldots, n$$

6.7 Unconstrained Minimization

Unconstrained minimization

- consider unconstrained convex optimization problem, *i.e.*, $m = p = 0$ in Definition 6.40
  $$\text{minimize} \quad f(x)$$
  where domain of optimization problem is $\mathcal{D} = F \subset \mathbf{R}^n$
- assume
  - f is twice-differentiable (hence by definition F is open)
  - optimal solution x^* exists, *i.e.*, $p^* = \inf_{x \in \mathcal{D}} f(x) = f(x^*)$
- Theorem 6.4 implies x^* is optimal solution *if and only if*
  $$\nabla f(x^*) = 0$$
- can solve above equation directly for few cases, but usually depend on iterative method, *i.e.*, find sequence of points $x^{(0)}, x^{(1)}, \ldots \in F$ such that $\lim_{k \to \infty} f(x^{(k)}) = p^*$

Requirements for iterative methods

- requirements for iterative methods
  - initial point $x^{(0)}$ should be in domain of optimization problem, *i.e.*
  $$x^{(0)} \in F$$
  - sublevel set of $f(x^{(0)})$
  $$S = \left\{ x \in F \,\middle|\, f(x) \leq f(x^{(0)}) \right\}$$
  should be closed
- *e.g.*
  - sublevel set of $f(x^{(0)})$ is closed for all $x^{(0)} \in F$ if f is closed, *i.e.*, all its sublevel sets are closed
  - f is closed if $F = \mathbf{R}^n$ and f is continuous
  - f is closed if f is continuous, F is open, and $f(x) \to \infty$ as $x \to \mathbf{bd}\, F$

Unconstrained minimization examples

- convex quadratic problem

$$\text{minimize} \quad f(x) = (1/2)x^T Px + q^T x$$

  where $P \in \mathbf{S}^n_+$ and $q \in \mathbf{R}^n$

  – solution obtained by solving

$$\nabla f(x^*) = Px^* + q = 0$$

    - if solution exists, $x^* = -P^\dagger q$ (thus $p^* > -\infty$)
    - otherwise, problem is unbounded below, *i.e.*, $p^* = -\infty$

  – *ability to analytically solve quadratic minimization problem is basis for Newton's method, power method for unconstrained minimization*

  – least-squares (LS) is special case of convex quadratic problem

$$\text{minimize} \quad (1/2)\|Ax - b\|_2^2 = (1/2)x^T(A^TA)x - b^TAx + (1/2)\|b\|_2^2$$

    - optimal always exists, can be obtained via normal equations

$$A^TAx^* = b$$

- unconstrained GP

$$\text{minimize} \quad f(x) = \log\left(\sum \exp(Ax + b)\right)$$

  for $A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$

  – solution obtained by solving

$$\nabla f(x^*) = \frac{\sum A^T \exp(Ax^* + b)}{\sum \exp(Ax^* + b)} = 0$$

  – need to resort to iterative method - since $F = \mathbf{R}^n$ and f is continuous, f is closed, hence every point in $\mathbf{R}^n$ can be initial point

- analytic center of linear inequalities

$$\text{minimize} \quad f(x) = -\sum \log(b - Ax)$$

  where $F = \{x \in \mathbf{R}^n | b - Ax \succ 0\}$

  – need to resort to iterative method - since F is open, f is continuous, and $f(x) \to \infty$ as $x \to \mathbf{bd}\, F$, f is closed, hence every point in F can be initial point

  – f, called *logarithmic barrier* for inequalities $Ax \preceq b$

- analytic center of LMI

$$\text{minimize} \quad f(x) = -\log\det F(x) = \log\det F(x)^{-1}$$

  where $F : \mathbf{R}^n \to \mathbf{S}^k$ is defined by

$$F(x) = x_1F_1 + \cdots + x_nF_n$$

  where $F_i \in \mathbf{S}^k$ and $F = \{x \in \mathbf{R}^n | F(x) \succ 0\}$

  – need to resort to iterative method - since F is open, f is continuous, and $f(x) \to \infty$ as $x \to \mathbf{bd}\, F$, f is closed, hence every point in F can be initial point

  – f, called *logarithmic barrier* for LMI

Strong convexity and implications

- function f is strongly convex on S

$$(\exists m > 0)\, (\forall x \in S)\, \big(\nabla^2 f(x) \succeq mI\big)$$

- strong convexity implies for every $x, y \in S$

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + (m/2)\|y - x\|_2^2$$

  - which implies gradient provides optimality certificate and tells us how far current point is from optimum, *i.e.*

$$f(x) - p^* \leq (1/2m)\|\nabla f(x)\|_2^2 \quad \|x - x^*\|_2 \leq (2/m)\|\nabla f(x)\|_2$$

- first equation implies sublevel sets contained in S is bounded, hence continuous function $\nabla^2 f(x)$ is also bounded, *i.e.*, $(\exists M > 0)\, \big(\nabla^2 f(x) \preceq MI\big)$, then

$$f(x) - p^* \geq \frac{1}{2M}\|\nabla f(x)\|_2^2$$

Iterative methods

Definition 6.66 (iterative meethods) *numerical method generating sequence of points* $x^{(0)}, x^{(1)}, \ldots \in S \subset \mathbf{R}^n$ *to make* $f(x^{(k)})$ *approaches to some desired value from some* $f : S \to \mathbf{R}$*, called* iterative method

Definition 6.67 (iterative meethods with search directions) *iterative method generating search direction* $\Delta x^{(k)} \in \mathbf{R}^n$ *and step length* $t^{(k)} > 0$ *at each step* k *such that*

$$x^{(k+1)} = x^{(k)} + t^{(k)} \Delta x^{(k)}$$

called iterative method with search direction *where* $\Delta x^{(k)}$*, called* search direction*,* $t^{(k)}$*, called* step length *(which actually is not length)*

Definition 6.68 (descent methods) *for function* $f : S \to \mathbf{R}$*, iterative method reducing function value, i.e.*

$$f(x^{(k+1)}) \leq f(x^{(k)})$$

for $k = 0, 1, \ldots$*, called* descent method

Line search methods

Definition 6.69 (line search method) *for iterating method with search directions, determining search direction* $\Delta x^{(k)}$ *and step length* $t^{(k)}$ *for each step, called* line search method

Algorithm 6.1 (exact line search) *for descent iterating method with search directions, determine* t *by*

$$t = \operatorname*{argmin}_{s > 0} f(x + s\Delta x)$$

Algorithm 6.2 (backtracking line search) *for descent iterating method with search directions, determine* t *by*

Require: f, $\Delta x^{(k)}$, $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$
  $t := 1$
  while $f(x^{(k)} + t\Delta x^{(k)}) > f(x^{(k)}) + \alpha t \nabla f(x^{(k)})^T \Delta x^{(k)}$ ***do***
    $t := \beta t$
  end while

Gradient descent method

Algorithm 6.3 (gradient descent method)
Require: *f, initial point $x \in \textbf{dom} f$*
  repeat
    search direction - $\Delta x := -\nabla f(x)$
    do line search to choose $t > 0$
    update - $x := x + t\Delta x$
  until *stopping criterion satisfied*

Summary of gradient descent method

- gradient method often exhibits approximately linear convergence, *i.e.*, error $f(x^{(k)}) - p^*$ converges to zero approximately as geometric series
- choice of backtracking parameters α and β has noticeable but not dramatic effect on convergence
- exact line search sometimes improves convergence of gradient method, but not by large, hence mostly not worth implementation
- converge rate depends greatly on condition number of Hessian or sublevel sets - when condition number if large, gradient method can be useless

Newton's method - motivation

- second-order Taylor expansion of f - $\hat{f}(\Delta x) = f(x + \Delta x) = f(x) + \nabla f(x)^T \Delta x + \frac{1}{2}\Delta x^T \nabla^2 f(x) \Delta x$
- minimum of Taylor expansion achieved when $\nabla \hat{f}(\Delta x) = \nabla f(x) + \nabla^2 f(x) v = 0$
- solution called *Newton step*
$$\Delta x_{\text{nt}}(x) = -\nabla^2 f(x)^{-1} \nabla f(x)$$
  assuming $\nabla^2 f(x) \succ 0$
- thus Newton step minimizes local quadratic approximation of function
- difference of current and quadratic approximation minimum
$$f(x) - \hat{f}(\Delta x_{\text{tn}}(x)) = \frac{1}{2}\Delta x_{\text{nt}}^T \nabla^2 f(x) \Delta x_{\text{nt}} = \frac{1}{2}\lambda(x)^2$$
- *Newton decrement*
$$\lambda(x) = \sqrt{\Delta x_{\text{nt}}(x)^T \nabla^2 f(x) \Delta x_{\text{nt}}(x)} = \sqrt{\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)}$$

Newton's method

Algorithm 6.4 (Newton's method) *damped descent method using Newton step*
Require: *f, initial point $x \in \textbf{dom} f$, tolerance $\epsilon > 0$*
  loop
    computer Newton step and descrement
$$\Delta x_{\text{nt}}(x) := -\nabla^2 f(x)^{-1} \nabla f(x) \quad \lambda(x)^2 := \nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x)$$
    stopping criterion - quit if $\lambda(x)^2/2 < \epsilon$
    do line search to choose $t > 0$

update - $x := x + t\Delta x_{\text{nt}}$
end loop

- Newton step is descent direction since

$$\left(\frac{d}{dx} f(x + t\Delta x_{\text{nt}})\right)\bigg|_{t=0} = \nabla f(x)^T \Delta x_{\text{nt}} = -\lambda(x)^2 < 0$$

Assumptions for convergence analysis of Newton's method

- assumptions
  - strong convexity and boundedness of Hessian on sublevel set

$$(\exists\, m, M > 0)\, (\forall x \in S)\, \big(mI \preceq \nabla^2 f(x) \preceq MI\big)$$

  - Lipschitz continuity of Hessian on sublevel set

$$(\exists L > 0)\, (\forall x, y \in S)\, \big(\|\nabla^2 f(x) - \nabla^2 f(y)\|_2 \leq L\|x - y\|_2\big)$$

- Lipschitz continuity constant L plays critical role in performance of Newton's method
  - intuition says Newton's method works well for functions whose quadratic approximations do not change fast, *i.e.*, when L is small

Convergence analysis of Newton's method

Theorem 6.19 (convergence analysis of Newton's method) *for function f satisfying strong convexity, Hessian continuity & Lipschitz continuity with $m, M, L > 0$, exist $0 < \eta < m^2/L$ and $\gamma > 0$ such that for each step k*

- *damped Newton phase - if $\|\nabla f(x^{(k)})\|_2 \geq \eta$,*

$$f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$$

- *quadratic convergence phase - if $\|\nabla f(x^{(k)})\|_2 < \eta$, backtracking line search selects step length $t^{(k)} = 1$*

$$\frac{L}{2m^2}\|\nabla f(x^{(k+1)})\|_2 \leq \left(\frac{L}{2m^2}\|\nabla f(x^{(k)})\|_2\right)^2$$

 # iterations of Newton's method required to satisfy stopping criterion $f(x^{(k)}) - p^* \leq \epsilon$ is

$$\frac{f(x^{(0)}) - p^*}{\gamma} + \log_2 \log_2(\epsilon_0/\epsilon) \quad \textit{where } \epsilon_0 = 2m^3/L^2$$

Summary of Newton's method

- Newton's method is *affine invariant*, hence *performance is independent of condition number unlike gradient method*
- once entering quadratic convergence phase, Newton's method converges extremely fast
- performance not much dependent on choice of algorithm parameters
- big disadvantage is computational cost for evaluating search direction, *i.e.*, solving linear system

Self-concordance

Definition 6.70 (self-concordance) *convex function $f : X \to \mathbf{R}$ with $X \subset \mathbf{R}^n$ such that for all $x \in X, v \in \mathbf{R}^n$, $g(t) = f(x+tv)$ with $\mathbf{dom}\, g = \{t \in \mathbf{R} | x+tv \in X\}$ satisfies*

$$(\forall t \in \mathbf{dom}\, g) \left(|g'''(t)| \leq 2g''(t)^{3/2}\right)$$

Proposition 6.11 (self-concordance for logarithms) *if convex function $g : X \to \mathbf{R}$ with $X \subset \mathbf{R}_{++}$ satisfies*

$$|g'''(x)| \leq 3g''(x)/x$$

function f with $\mathbf{dom}\, f = \{x \in \mathbf{R}_{++} | g(x) < 0\}$ defined by

$$f(x) = -\log(-g(x)) - \log x$$

and function h with $\mathbf{dom}\, h = \{x \in \mathbf{R}_{++} | g(x) + ax^2 + bx + c < 0\}$ with $a \geq 0$ defined by

$$h(x) = -\log(-g(x) - ax^2 - bx - c) - \log x$$

are self-concordant

Why self-concordance?

- convergence analysis of Newton's method depends on assumptions about function characteristics, *e.g.*, $m, M, L > 0$ for strong convexity, continuity of Hessian, *i.e.*

$$mI \preceq \nabla^2 f(x) \preceq MI \quad \|\nabla^2 f(x) - \nabla^2 f(y)\| \leq L\|x-y\|$$

- *self-concordance* discovered by Nesterov and Nemirovski (who gave name self-concordance) plays important role for reasons such as
  - convergence analysis does not depend any function characterizing paramters
  - many barrier functions which are used for interior-point methods, which are important class of optimization algorithms are self-concordance
  - property of self-concordance is affine invariant

Self-concordance preserving operations

Proposition 6.12 (self-concordance preserving operations) *self-concordance is preserved by* positive scaling, addition, and affine transformation, *i.e., if $f, g : X \to \mathbf{R}$ are self-concordant functions with $X \subset \mathbf{R}^n$, $h : H \to \mathbf{R}^n$ with $H \subset \mathbf{R}^m$ are affine functions, and $a > 0$*

$$af, \quad f+g, \quad f \circ h$$

are self-concordant where $\mathbf{dom}\, f \circ h = \{x \in H | h(x) \in X\}$

Self-concordant function examples

- negative logarithm - $f : \mathbf{R}_{++} \to \mathbf{R}$ with

$$f(x) = -\log x$$

is self-concordant since

$$|f'''(x)|/f''(x)^{3/2} = \left(2/x^3\right) / \left((1/x^2)^{3/2}\right) = 2$$

- negative entropy plus negative logarithm - $f : \mathbf{R}_{++} \to \mathbf{R}$ with

$$f(x) = x \log x - \log x$$

is self-concordant since

$$|f'''(x)|/f''(x)^{3/2} = (x+2)/(x+1)^{3/2} \leq 2$$

- log barrier for linear inequalities - for $A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$

$$f(x) = -\sum \log(b - Ax)$$

with $\mathbf{dom}\, f = \{x \in \mathbf{R}^n | b - Ax \succ 0\}$ is self-concordant by Proposition 6.12, *i.e.*, f is affine transformation of sum of self-concordant functions

- log-determinant - $f : \mathbf{S}^n_{++} \to \mathbf{R}$ with

$$f(X) = \log\det X^{-1} = -\log\det X$$

is self-concordant since for every $X \in \mathbf{S}^n_{++}$ and $V \in \mathbf{S}^n$ function $g : \mathbf{R} \to \mathbf{R}$ defined by $g(t) = -\log\det(X + tV)$ where $\mathbf{dom}\, f = \{t \in \mathbf{R} | X + tV \succeq 0\}$ is self-concordant since

$$\begin{aligned} g(t) &= -\log\det(X^{1/2}(I + tX^{-1/2}VX^{-1/2})X^{1/2}) \\ &= -\log\det X - \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= -\log\det X - \sum \log(1 + t\lambda_i(X,V)) \end{aligned}$$

where $\lambda_i(X,V)$ is i-th eigenvalue of $X^{-1/2}VX^{1/2}$ is self-concordant by Proposition 6.12, *i.e.*, g is affine transformation of sum of self-concordant functions

- log of concave quadratic - $f : X \to \mathbf{R}$ with

$$f(x) = -\log(-x^TPx - q^Tx - r)$$

where $P \in \mathbf{S}^n_+$ and $X = \{x \in \mathbf{R}^n | x^TPx + q^Tx + r < 0\}$

- function $f : X \to \mathbf{R}$ with

$$f(x) = -\log(-g(x)) - \log x$$

where $\mathbf{dom}\, f = \{x \in \mathbf{dom}\, g \cap \mathbf{R}_{++} | g(x) < 0\}$ and function $h : H \to \mathbf{R}$

$$h(x) = -\log(-g(x) - ax^2 - bx - c) - \log x$$

where $a \geq 0$ and $\mathbf{dom}\, h = \{x \in \mathbf{dom}\, g \cap \mathbf{R}_{++} | g(x) + ax^2 + bx + c < 0\}$ are self-concordant if g is one of below

  - $g(x) = -x^p$ for $0 < p \leq 1$
  - $g(x) = -\log x$
  - $g(x) = x\log x$
  - $g(x) = x^p$ for $-1 \leq p \leq 0$
  - $g(x) = (ax+b)^2/x$ for $a, b \in \mathbf{R}$

since above g satisfy $|g'''(x)| \leq 3g''(x)/x$ for every $x \in \mathbf{dom}\, g$ (Proposition 6.11)

- function $f : X \to \mathbf{R}$ with $X = \{(x,y) | \|x\|_2 < y\} \subset \mathbf{R}^n \times \mathbf{R}_{++}$ defined by

$$f(x,y) = -\log(y^2 - x^Tx)$$

is self-concordant - can be proved using Proposition 6.11

- function $f : X \to \mathbf{R}$ with $X = \{(x,y) | |x|^p < y\} \subset \mathbf{R} \times \mathbf{R}_{++}$ defined by

  $$f(x,y) = -2\log y - \log(y^{2/p} - x^2)$$

  where $p \geq 1$ is self-concordant - can be proved using Proposition 6.11

- function $f : X \to \mathbf{R}$ with $X = \{(x,y) | \exp(x) < y\} \subset \mathbf{R} \times \mathbf{R}_{++}$ defined by

  $$f(x,y) = -\log y - \log(\log y - x)$$

  is self-concordant - can be proved using Proposition 6.11

Properties of self-concordant functions

Definition 6.71 (Newton decrement) *for convex function $f : X \to \mathbf{R}$ with $X \subset \mathbf{R}^n$, function $\lambda : \tilde{X} \to \mathbf{R}_+$ with $\tilde{X} = \{x \in X | \nabla^2 f(x) \succ 0\}$ defined by*

$$\lambda(x) = (\nabla f(x)^T \nabla^2 f(x)^{-1} \nabla f(x))^{1/2}$$

called Newton decrement

- *note*

$$\lambda(x) = \sup_{v \neq 0} \left(v^T \nabla f(x) / \left(v^T \nabla^2 f(x) v \right)^{1/2} \right)$$

Theorem 6.20 (optimality certificate for self-concordant functions) *for strictly convex self-concordant function $f : X \to \mathbf{R}^n$ with $X \subset \mathbf{R}^n$, Hessian is positive definition everywhere (hence Newton decrement is defined everywhere) and for every $x \in X$*

$$p^* \geq f(x) - \lambda(x)^2 \quad \Leftrightarrow \quad f(x) - p^* \leq \lambda(x)^2$$

if $\lambda(x) \leq 0.68$

Stopping criteria for self-concordant objective functions

- recall $\lambda(x)^2$ provides *approximate* optimality certificate, (page 194) *i.e.*, assuming f is well approximated by quadratic function around x

  $$f(x) - p^* \lessapprox \lambda(x)^2/2$$

- however, strict convexity together with self-concordance provides proven bound (by Theorem 6.20)

  $$f(x) - p^* \leq \lambda(x)^2$$

  for $\lambda(x) \leq 0.68$

- hence can use following stopping criterion for guaranteed bound

  $$\lambda(x)^2 \leq \epsilon \quad \Rightarrow \quad f(x) - p^* \leq \epsilon$$

  for $\epsilon \leq 0.68^2$

Convergence analysis of Newton's method for self-concordant functions

Theorem 6.21 (convergence analysis of Newton's method for self-concordant functions) *for strictly convex self-concordant function f, exist $0 < \eta \leq 1/4$ and $\gamma > 0$ (which depend only on line search parameters) such that*

- *damped Newton phase - if $\lambda(x^{(k)}) > \eta$*

$$f(x^{(k+1)}) - f(x^{(k)}) \leq -\gamma$$

- *quadratic convergence phase - if $\lambda(x^{(k)}) \leq \eta$ backtracking line search selects step length $t^{(k)} = 1$*

$$2\lambda(x^{(k+1)}) \leq \left(2\lambda(x^{(k)})\right)^2$$

 # iterations required to satisfy stopping criterion $f(x^{(k)}) - p^* \leq \epsilon$ is

$$\left(f(x^{(0)}) - p^*\right)/\gamma + \log_2 \log_2(1/\epsilon)$$

where $\gamma = \alpha\beta(1-2\alpha)^2/(20-8\alpha)$

6.8 Equality Constrained Minimization

Equality constrained minimization

- consider equality constrained convex optimization problem, *i.e.*, $m = 0$ in Definition 6.40

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & Ax = b \end{array}$$

  where $A \in \mathbf{R}^{p\times n}$ and domain of optimization problem is $\mathcal{D} = F \subset \mathbf{R}^n$

- assume
  - $\mathbf{rank}\, A = p < n$, *i.e.*, rows of A are linearly independent
  - f is twice-differentiable (hence by definition F is open)
  - optimal solution x^* exists, *i.e.*, $p^* = \inf_{x\in\mathcal{F}} f(x) = f(x^*)$ and $Ax^* = b$

Solving KKT for equality constrained minimization

- Theorem 6.10 implies $x^* \in F$ is optimal solution *if and only if* exists $\nu^* \in \mathbf{R}^p$ satisfy KKT optimality conditions, *i.e.*,

$$\begin{aligned} Ax^* &= b && \textit{primal feasibility equations} \\ \nabla f(x^*) + A^T\nu^* &= 0 && \textit{dual feasibility equations} \end{aligned}$$

- solving equality constrained problem is equivalent to solving KKT equations
  - handful types of problems can be solved analytically
- using unconstrained minimization methods
  - can eliminate equality constraints and apply unconstrained minimization methods
  - solving dual problem using unconstrained minimization methods and retrieve primal solution (refer to page 180)
- will discuss Newton's method directly handling equality constraints
  - preserving problem structure such as sparsity

Equality constrained convex quadratic minimization

- equality constrained convex quadratic minimization problem

$$\begin{array}{ll}\text{minimize} & f(x) = (1/2)x^T P x + q^T x \\ \text{subject to} & Ax = b\end{array}$$

  where $P \in \mathbf{S}_+^n$ and $A \in \mathbf{R}^{p \times n}$

- important since basis for extension of Newton's method to equality constrained problems

- *KKT system*

$$Ax^* = b \,\&\, Px^* + q + A^T\nu^* = 0 \;\Leftrightarrow\; \underbrace{\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}}_{\textit{KKT matrix}} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

- exist primal and dual optimum (x^*, ν^*) *if and only if* KKT system has solution; otherwise, problem is unbounded below

Eliminating equality constraints

- can solve equality constrained convex optimization by
  - eliminating equality constraints and
  - using optimization method for solving unconstrained optimization

- note

$$\mathcal{F} = \{x | Ax = b\} = \{Fz + x_0 | z \in \mathbf{R}^{n-p}\}$$

  for some $F \in \mathbf{R}^{n \times (n-p)}$ where $\mathcal{R}(F) = \mathcal{N}(A)$

- thus original problem equivalent to

$$\text{minimize} \quad f(Fz + x_0)$$

- if z^* is optimal solution, $x^* = Fz^* + x_0$

- optimal dual can be retrieved by

$$\nu^* = -(AA^T)^{-1}A\nabla f(x^*)$$

Solving dual problems

- Lagrange dual function of equality constrained problem

$$\begin{aligned} g(\nu) &= \inf_{x \in \mathcal{D}} \left(f(x) + \nu^T(Ax - b)\right) = -b^T\nu - \sup_{x \in \mathcal{D}} \left((-A^T\nu)^T x - f(x)\right) \\ &= -b^T\nu - f^*(-A^T\nu) \end{aligned}$$

- dual problem

$$\text{maximize} \quad -b^T\nu - f^*(-A^T\nu)$$

- by assumption, strong duality holds, hence if ν^* is dual optimum

$$g(\nu^*) = p^*$$

- if dual objective is twice-differentiable, can solve dual problem using unconstrained minimization methods

- primal optimum can be retrieved using method on page 180)

Newton's method with equality constraints

- finally discuss Newton's method which directly handles equality constraints
  - similar to Newton's method for unconstrained minimization
  - initial point, however, should be feasible, *i.e.*, $x^{(0)} \in F$ and $Ax^{(0)} = b$
  - Newton step tailored for equality constrained problem

Newton step via second-order approximation

- solve original problem approximately by solving

$$\begin{array}{ll} \text{minimize} & \hat{f}(x+\Delta x) = f(x) + \nabla f(x)^T \Delta x + (1/2)\Delta x^T \nabla^2 f(x) \Delta x \\ \text{subject to} & A(x+\Delta x) = b \end{array}$$

  where $x \in \mathcal{F}$

- *Newton step for equality constrained minimization problem*, defined by solution of KKT system for above convex quadratic minimization problem

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x_{\mathrm{nt}} \\ w \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ 0 \end{bmatrix}$$

  only when KKT system is nonsingular

Newton step via solving linearized KKT optimality conditions

- recall KKT optimality conditions for equality constrained convex optimization problem

$$Ax^* = b \quad \& \quad \nabla f(x^*) + A^T \nu^* = 0$$

- linearize KKT conditions

$$\begin{array}{rl} & A(x+\Delta x) = b \quad \& \quad \nabla f(x) + \nabla^2 f(x)\Delta x + A^T w = 0 \\ \Leftrightarrow & A\Delta x = 0 \quad \& \quad \nabla^2 f(x) \Delta x + A^T w = -\nabla f(x) \end{array}$$

  where $x \in \mathcal{F}$

- Newton step defined by above equations is equivalent to that obtained by second-order approximation

Newton decrement for equality constrained minimization

- *Newton descrement for equality constrained problem* is defined by

$$\lambda(x) = \left(\Delta x_{\mathrm{nt}} \nabla^2 f(x) \Delta x_{\mathrm{nt}}\right)^{1/2}$$

- same expression as that for unconstrained minimization, but is *different* since Newton step Δx_{nt} is different from that for unconstrained minimization, *i.e.*, $\Delta x_{\mathrm{nt}} \neq -\nabla^2 f(x)^{-1} \nabla f(x)$ (refer to Definition 6.71)

- however, as before,

$$f(x) - \inf_{\Delta x \in \mathbf{R}^n} \{\hat{f}(x+\Delta x) | A(x+\Delta x) = b\} = \lambda(x)^2/2$$

  and

$$\left.\left(\frac{d}{dt} f(x + t\Delta x_{\mathrm{nt}})\right)\right|_{t=0} = \nabla f(x)^T \Delta x_{\mathrm{nt}} = -\lambda(x)^2 < 0$$

Feasible Newton's method for equality constrained minimization

Algorithm 6.5 (feasible Newton's method for equality constrained minimization)

Require: *f, initial point $x \in \mathbf{dom}\, f$ with $Ax = b$, tolerance $\epsilon > 0$*

loop

computer Newton step and descrement $\Delta x_{\mathrm{nt}}(x)$ & $\lambda(x)$

stopping criterion - quit if $\lambda(x)^2/2 < \epsilon$

do line search on f to choose $t > 0$

update - $x := x + t\Delta x_{\mathrm{nt}}$

end loop

- Algorithm 6.5
  - assumes KKT matrix is nonsingular for every step
  - is *feasible descent method* since all iterates are feasible with $f(x^{(k+1)}) < f(x^{(k)})$

Assumptions for convergence analysis of feasible Newton's method for equality constrained minimization

- feasibility of initial point - $x^{(0)} \in \mathbf{dom}\, f$ & $Ax^{(0)} = b$
- sublevel set $S = \{x \in \mathbf{dom}\, f | f(x) \leq f(x^{(0)}),\ Ax = b\}$ is closed
- boundedness of Hessian on S

$$(\exists M > 0)\, (\forall x \in S)\, \left(\nabla^2 f(x) \preceq MI\right)$$

- boundedness of KKT matrix on S - corresponds to strong convexity assumption in unconstrained minimization

$$(\exists K > 0)\, (\forall x \in S) \left(\left\| \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}^{-1} \right\|_2 \leq K \right)$$

- Lipschitz continuity of Hessian on S

$$(\exists L > 0)\, (\forall x, y \in S)\, \left(\left\|\nabla^2 f(x) - \nabla^2 f(y)\right\|_2 \leq L\|x - y\|_2\right)$$

Convergence analysis of feasible Newton's method for equality constrained minimization

- convergence analysis of Newton's method for equality constrained minimization can be done by analyzing unconstrained minimization after eliminating equality constraints
- thus, yield *exactly same* results as for unconstrained minimization (Theorem 6.19) (with different parameter values), *i.e.*,
  - consists of damped Newton phase and quadratic convergence phase
  - # iterations required to achieve $f(x^{(k)}) - p^* \leq \epsilon$ is

$$\left(f(x^{(0)}) - p^*\right)/\gamma + \log_2 \log_2(\epsilon_0/\epsilon)$$

- for # iterations required to achieve $f(x^{(k)}) - p^* \leq \epsilon$ for self-concordant functions is also same as for unconstrained minimization (Theorem 6.21)

$$\left(f(x^{(0)}) - p^*\right)/\gamma + \log_2 \log_2(1/\epsilon)$$

  where $\gamma = \alpha\beta(1 - 2\alpha)^2/(20 - 8\alpha)$

Newton step at infeasible points

- only assume that $x \in \mathbf{dom}\, f$ (hence, can be infeasible)
- (as before) linearize KKT conditions

$$
\begin{array}{rll}
& A(x+\Delta x_{\mathrm{nt}}) = b & \& \quad \nabla f(x) + \nabla^2 f(x)\Delta x_{\mathrm{nt}} + A^T w = 0 \\
\Leftrightarrow & A\Delta x_{\mathrm{nt}} = b - Ax & \& \quad \nabla^2 f(x)\Delta x_{\mathrm{nt}} + A^T w = -\nabla f(x) \\
\Leftrightarrow & \left[\begin{array}{cc} \nabla^2 f(x) & A^T \\ A & 0 \end{array}\right] \left[\begin{array}{c} \Delta x_{\mathrm{nt}} \\ w \end{array}\right] = -\left[\begin{array}{c} \nabla f(x) \\ Ax - b \end{array}\right] &
\end{array}
$$

- same as feasible Newton step *except second component on RHS of KKT system*

Interpretation as primal-dual Newton step

- update both primal and dual variables x and ν
- define $r : \mathbf{R}^n \to \mathbf{R}^p \to \mathbf{R}^n \times \mathbf{R}^p$ by

$$
r(x,\nu) = (r_{\mathrm{dual}}(x,\nu), r_{\mathrm{pri}}(x,\nu))
$$

where

$$
\begin{array}{rcl}
\textit{dual residual} & - & r_{\mathrm{dual}}(x,\nu) = \nabla f(x) + A^T\nu \\
\textit{primal residual} & - & r_{\mathrm{pri}}(x,\nu) = Ax - b
\end{array}
$$

Equivalence of infeasible Newton step to primal-dual Newton step

- linearize r to obtain primal-dual Newton step, *i.e.*

$$
\begin{array}{rc}
& r(x,\nu) + D_{x,\nu} r(x,\nu) \left[\begin{array}{c} \Delta x_{\mathrm{pd}} \\ \Delta \nu_{\mathrm{pd}} \end{array}\right] = 0 \\
\Leftrightarrow & \left[\begin{array}{cc} \nabla^2 f(x) & A^T \\ A & 0 \end{array}\right] \left[\begin{array}{c} \Delta x_{\mathrm{pd}} \\ \Delta \nu_{\mathrm{pd}} \end{array}\right] = -\left[\begin{array}{c} \nabla f(x) + A^T\nu \\ Ax - b \end{array}\right]
\end{array}
$$

- letting $\nu^+ = \nu + \Delta\nu_{\mathrm{pd}}$ gives

$$
\left[\begin{array}{cc} \nabla^2 f(x) & A^T \\ A & 0 \end{array}\right] \left[\begin{array}{c} \Delta x_{\mathrm{pd}} \\ \nu^+ \end{array}\right] = -\left[\begin{array}{c} \nabla f(x) \\ Ax - b \end{array}\right]
$$

  - equivalent to infeasible Newton step
  - reveals that current value of dual variable not needed

Residual norm reduction property

- infeasible Newton step is *not* descent direction (unlike feasible Newton step) since

$$
\begin{aligned}
& \left.\left(\frac{d}{dt} f(x + t\Delta x_{\mathrm{pd}})\right)\right|_{t=0} = \nabla f(x)^T \Delta x_{\mathrm{pd}} \\
= \quad & -\Delta x_{\mathrm{pd}}^T \left(\nabla^2 f(x)\Delta x_{\mathrm{pd}} + A^T w\right) = -\Delta x_{\mathrm{pd}}^T \nabla^2 f(x)\Delta x_{\mathrm{pd}} + (Ax-b)^T w
\end{aligned}
$$

which is not necessarily negative

- however, norm of residual decreases in infeasible Newton direction

$$
\begin{aligned}
\left(\frac{d}{dx}\|r(y+t\Delta y_{\mathrm{pd}})\|_2^2\right)\Bigg|_{t=0} &= -2r(y)^T r(y) = -2\|r(y)\|_2^2 \\
\Leftrightarrow \quad \left(\frac{d}{dx}\|r(y+t\Delta y_{\mathrm{pd}})\|_2\right)\Bigg|_{t=0} &= \frac{-2\|r(y)\|_2^2}{2\|r(y)\|_2} = -\|r(y)\|_2
\end{aligned}
$$

  where $y=(x,\nu)$ and $\Delta y_{\mathrm{pd}} = (\Delta x_{\mathrm{pd}}, \Delta \nu_{\mathrm{pd}})$

- can use $r(x^{(k)}, \nu^{(k)})$ to measure optimization progress for infeasible Newton's method

Full and damped step feasibility property

- assume step length is t at some iteration, then

$$
r_{\mathrm{pri}}(x^+, \nu^+) = Ax^+ - b = A(x + t\Delta x_{\mathrm{pd}}) - b = (1-t) r_{\mathrm{pri}}(x, \nu)
$$

- hence $l > k$

$$
r^{(l)} = \left(\prod_{i=k}^{l-1} (1 - t^{(i)})\right) r^{(k)}
$$

  - primal residual reduced by $1 - t^{(k)}$ at step k
  - Newton step becomes feasible step once full step length ($t = 1$) taken

Infeasible Newton's method for equality constrained minimization

Algorithm 6.6 (infeasible Newton's method for equality constrained minimization)

Require: *f, initial point $x \in$* **dom** *f & ν, tolerance $\epsilon_{\mathrm{pri}} > 0$ & $\epsilon_{\mathrm{dual}} > 0$*
 repeat
  computer Newton step and descrement $\Delta x_{\mathrm{pd}}(x)$ & $\Delta \nu_{\mathrm{pd}}(x)$,
  do line search on $r(x,\nu)$ to choose $t > 0$
  update - $x := x + t\Delta x_{\mathrm{pd}}$ & $\nu := \nu + t\Delta \nu_{\mathrm{pd}}$
 until $\|r_{\mathrm{dual}}(x,\nu)\| \leq \epsilon_{\mathrm{dual}}$ *&* $\|Ax - b\| \leq \epsilon_{\mathrm{pri}}$

- note similarity and difference of Algorithm 6.6 & Algorithm 6.5
  - line search done not on f, but on primal-dual residuals $r(x,\nu)$
  - stopping criteria depends on $r(x,\nu)$, not on Newton decrementa $\lambda(x)^2$
  - primal and dual feasibility checked separately - here norm in $\|Ax - b\|$ can be any norm, *e.g.*, $\|\cdot\|_0$, $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$, depending on specific application

Line search methods for infeasible Newton's method

- line search methods for infeasible Newton's method, *i.e.*, Algorithm 6.1 & Algorithm 6.2 same with f replaced by $\|r(x,\nu)\|_2$,
- but they have special forms (of course) - refer to below special case descriptions

Algorithm 6.7 (exact line search for infeasible Newton's method)

$$
t = \operatorname*{argmin}_{s>0} \|r(x + s\Delta x_{\mathrm{pd}}, \nu + s\Delta \nu_{\mathrm{pd}})\|_2
$$

Algorithm 6.8 (backtracking line search for infeasible Newton's method)

Require: Δx, $\Delta \nu$, $\alpha \in (0, 0.5)$, $\beta \in (0, 1)$
  $t := 1$
  while $\|r(x + t\Delta x_{\mathrm{pd}}, \nu + t\Delta \nu_{\mathrm{pd}})\|_2 > (1 - \alpha t)\|r(x, \nu)\|_2$ ***do***
    $t := \beta t$
  end while

Pros and cons of infeasible Newton's method

- pros
  - do not need to find feasible point separately, *e.g.*
    - "minimize $-\log(Ax) + b^T x$"

    can be solved by converting to
    - "minimize $-\log(y) + b^T x \quad$ s.t. $y = Ax$"

    and solved by infeasible Newton's method
  - if step length is one at any iteration, following steps coincides with feasible Newton's method - could switch to feasible Newton's method
- cons
  - exists no clear way to detect feasibility - primal residual decreases slowly (phase I method in interior point method resolves this problem)
  - convergence of infeasible Newton's method can be very slow (until feasibility is achieved0

Assumptions for convergence analysis of infeasible Newton's method for equality constrained minimization

- sublevel set $S = \left\{(x, \nu) \in \mathbf{dom}\, f \times \mathbf{R}^m \left| \|r(x, \nu)\|_2 \leq \|r(x^{(0)}, \nu^{(0)})\|_2 \right.\right\}$ is closed, which always holds because $\|r\|_2$ is closed
- boundedness of KKT matrix on S

$$(\exists K > 0)\, (\forall x \in S) \left(\left\|Dr(x, \nu)^{-1}\right\|_2 = \left\| \begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix}^{-1} \right\|_2 \leq K \right)$$

- Lipschitz continuity of Hessian on S

$$(\exists L > 0)\, (\forall (x, \nu), (y, \mu) \in S)\, (\|Dr(x, \nu) - Dr(y, \mu)\|_2 \leq L\|(x, \nu) - (y, \mu)\|_2)$$

- above assumptions imply $\{x \in \mathbf{dom}\, f | Ax = b\} \neq \emptyset$ and exist optimal point (x^*, ν^*)

Convergence analysis of infeasible Newton's method for equality constrained minimization

- very simliar to that for Newton's method for unconstrained minimization
- consist of two phases - like unconstrained minimization or infeasible Newton's method (refer to Theorem 6.19 or page 202)
  - damped Newton phase - if $\|r(x^{(k)}, \nu^{(k)})\|_2 > 1/(K^2 L)$

$$\|r(x^{(k+1)}, \nu^{(k+1)})\|_2 \leq \|r(x^{(k)}, \nu^{(k)})\|_2 - \alpha\beta/K^2 L$$

- quadratic convergence damped Newton phase - if $\|r(x^{(k)}, \nu^{(k)})\|_2 \leq 1/(K^2L)$

$$\left(K^2L\|r(x^{(k)}, \nu^{(k)})\|_2/2\right) \leq \left(K^2L\|r(x^{(k-1)}, \nu^{(k-1)})\|_2/2\right)^2 \leq \cdots \leq (1/2)^{2^k}$$

- # iterations of infeasible Newton's method required to satisfy $\|r(x^{(k)}, \nu^{(k)})\|_2 \leq \epsilon$

$$\|r(x^{(0)}, \nu^{(0)})\|/(\alpha\beta/K^2L) + \log_2\log_2(\epsilon_0/\epsilon) \quad \text{where } \epsilon_0 = 2/(K^2L)$$

- $(x^{(k)}, \nu^{(k)})$ converges to (x^*, ν^*)

6.9 Barrier Interior-point Methods

Interior-point methods

- want to solve inequality constrained minimization problem
- interior-point methods solve convex optimization problem (Definition 6.40) or KKT optimality conditions (Definition 6.60) by
  - applying Newton's method to sequence of
    - equality constrained problems or
    - modified versions of KKT optimality conditions
- discuss interior-point *barrier method* & interior-point *primal-dual method*
- hierarchy of convex optimization algorithms
  - simplest - linear equality constrained quadratic program - can solve analytically
  - Newton's method - solve linear equality constrained convex optimization problem by solving sequence of linear equality constrained quadratic programs
  - interior-point methods - solve linear equality & convex inequality constrained problem by solving sequence of lienar equality constrained convex optimization problem

Indicator function barriers

- approxmiate general convex inequality constrained problem as linear equality constrained problem
- make inequality constraints implicit in objective function

$$\begin{array}{ll} \text{minimize} & f(x) + \sum I_-(q(x)) \\ \text{subject to} & Ax = b \end{array}$$

where $I_- : \mathbf{R} \to \mathbf{R}$ is indicator function for nonpositive real numbers, *i.e.*

$$I_-(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases}$$

Logarithmic barriers

- approximate indicator function by logarithmic function

$$\hat{I}_- = -(1/t)\log(-u) \quad \mathbf{dom}\,\hat{I}_- = -\mathbf{R}_{++}$$

  for $t > 0$ to obtain

$$\begin{array}{ll} \text{minimize} & f(x) + \sum -(1/t)\log(-q(x)) \\ \text{subject to} & Ax = b \end{array}$$

- objective function is convex due to composition rule for convexity preservation (page 151), and differentiable
- hence, can use Newton's method to solve it
- function ϕ defined by

$$\phi(x) = -\sum \log(-q(x))$$

  with $\mathbf{dom}\,\phi\{x \in X | q(x) \prec 0\}$ called *logarithmic barrier* or *log barrier*

- solve sequence of log barrier problems as we increase t

Central path

- optimization problem

$$\begin{array}{ll} \text{minimize} & tf(x) + \phi(x) \\ \text{subject to} & Ax = b \end{array}$$

  with $t > 0$ where

$$\phi(x) = -\sum \log(-q(x))$$

- solution of above problem, called *central point*, denoted by $x^*(t)$, set of central points, called *central path*
- intuition says $x^*(t)$ will converge to x^* as $t \to \infty$
- KKT conditions imply

$$Ax^*(t) = b \quad q(x^*(t)) \prec 0$$

  and exists $\nu^*(t)$ such that

$$\begin{aligned} 0 &= t\nabla f(x^*(t)) + \nabla\phi(x^*(t)) + tA^T\nu^*(t) \\ &= t\nabla f(x^*(t)) - \sum \frac{1}{q_i(x^*(t))}\nabla q_i(x^*(t)) + tA^T\nu^*(t) \end{aligned}$$

- thus if we let $\lambda^*(t) = -1/tq_i(x^*(t))$, $x^*(t)$ minimizes

$$L(x, \lambda^*(t), \nu^*(t)) = f(x) + \lambda^*(t)^T q(x) + \nu^*(t)^T (Ax - b)$$

  where L is Lagrangian of original problem in Definition 6.40

- hence, dual function $g(\lambda^*(t), \nu^*(t))$ is finite and

$$
\begin{aligned}
g(\lambda^*(t), \nu^*(t)) &= \inf_{x \in X} L(x, \lambda^*(t), \nu^*(t)) = L(x^*(t), \lambda^*(t), \nu^*(t)) \\
&= f(x^*(t)) + \lambda^*(t)^T q(x^*(t)) + \nu^*(t)^T (Ax^*(t) - b) = f(x^*(t)) - m/t
\end{aligned}
$$

and

$$
f(x^*(t)) - p^* \leq f(x^*(t)) - g(\lambda^*(t), \nu^*(t)) = m/t
$$

that is,

$x^*(t)$ is no more than m/t-suboptimal

which confirms out intuition that $x^*(t) \to x^*$ as $t \to \infty$

Central path interpretation via KKT conditions

- previous arguments imply that x is central point, *i.e.*, $x = x^*(t)$ for some $t > 0$ *if and only if* exist λ and ν such that

$$
\begin{array}{rcll}
Ax = b \quad q(x) & \preceq & 0 & \text{- primal feasibility} \\
\lambda & \succeq & 0 & \text{- dual feasibility} \\
-{\lambda_i}^T q_i(x) & = & 1/t & \text{- complementary } 1/t\text{-slackness} \\
\nabla_x L(x, \lambda, \nu) & = & 0 & \text{- vanishing gradient of Lagrangian}
\end{array}
$$

  called *centrality conditions*

- only difference between centrality conditions and KKT conditions in Definition 6.60 is *complementary $1/t$-slackness*
  - note that I've just made up term "complementary $1/t$-slackness" - you won't be able to find terminology in any literature
- for large t, $\lambda^*(t)$ & $\nu^*(t)$ *very closely* satisfy (true) complementary slackness

Central path interpretation via force field

- assume exist no equality constraints
- interpret ϕ as potential energy by some force field, *e.g.*, electrical field and tf as potential energy by some other force field, *e.g.*, gravity
- then
  - force by first force field (in n-dimensional space), which we call *barrier force*, is

$$
-\nabla \phi(x) = \sum \frac{1}{q_i(x)} \nabla q_i(x)
$$

  - force by second force field, which we call *objective force*, is

$$
-\nabla (tf(x)) = -t \nabla f(x)
$$

- $x^*(t)$ is point where two forces exactly balance each other
  - as x approach boundary, barrier force pushes x harder from barriers,
  - as t increases, objective force pushes x harder to point where objective potential energy is minimized

Equality constrained problem using log barrier

- central point $x^*(t)$ is m/t-suboptimal point guaranteed by optimality certificate $g(\lambda^*(t), \nu^*(t))$
- hence solving below problem provides solution with ϵ-suboptimality

$$
\begin{array}{ll}
\text{minimize} & (m/\epsilon) f(x) + \phi(x) \\
\text{subject to} & Ax = b
\end{array}
$$

- but works only for small problems since for large m/ϵ, objective function ill behaves

Barrier methods

Algorithm 6.9 (barrier method)

Require: *strictly feasible x, $t > 0$, $\mu > 1$, tolerance $\epsilon > 0$*
  repeat
    centering step - find $x^*(t)$ by minimizing $tf + \phi$ subject to $Ax = b$ starting at x
    (optionally) compute $\lambda^*(t)$ & $\nu^*(t)$
    stopping criterion - quit if $m/t < \epsilon$
    increase t - $t := \mu t$
    update x - $x := x^*(t)$
  until

- *barrier method*, also called *path-following method*, solves sequence of equality constrained optimization problem with log barrier
  - when first proposed by Fiacco and McCormick in 1960s, it was called *sequential unconstrained minimization technique (SUMT)*
- *centering step* also called *outer iteration*
- each iteration of algorithm used for equality constrained problem, called *inner iteration*

Accuracy in centering in barrier method

- accuracy of centering
  - only goal of centering is getting close to x^*, hence exact calculation of $x^*(t)$ not critical as long as approximates of $x^*(t)$ go to x^*
  - while cannot calculate $g(\lambda, \nu)$ for this case, below provides dual feasible point when Newton step Δx_{nt} for optimization problem on page 207 is small, *i.e.*, for nearly centered

$$
\tilde{\lambda}_i = -\frac{1}{t q_i(x)} \left(1 - \frac{\nabla q_i(x)^T \Delta x_{\mathrm{nt}}}{q_i(x)} \right)
$$

Choices of parameters of barrier method

- choice of μ
  - μ determines aggressiveness of t-update
    - larger μ, less outer iterations, but more inner iterations
    - smaller μ, less outer iterations, but more inner iterations
  - values from 10 to 20 for μ seem to work well

- candidates for choice of initial t - choose $t^{(0)}$ such that

$$m/t^{(0)} \approx f(x^{(0)}) - p^*$$

or make central path condition on page 207 maximally satisfied

$$t^{(0)} = \operatorname*{arginf}_{t} \inf_{\tilde{\nu}} \left\| t\nabla f(x^{(0)}) + \nabla \phi(x^{(0)}) + A^T \tilde{\nu} \right\|$$

Convergence analysis of barrier method

- assuming $tf + \phi$ can be minimized by Newton's method for $t^{(0)}$, $\mu t^{(0)}$, $\mu^2 t^{(0)}$, ...
- at k'th step, duality gap achieved is $m/\mu^k t^{(0)}$
- # centering steps required to achieve accuracy of ϵ is

$$\left\lceil \frac{\log\left(m/\epsilon t^{(0)}\right)}{\log \mu} \right\rceil$$

  plus one (initial centering step)
- for convergence of centering
  - for feasible centering problem, $tf + \phi$ should satisfy conditions on page 202, *i.e.*, initial sublevel set is closed, associated inverse KKT matrix is bounded & Hessian satisfies Lipschitz condition
  - for infeasible centering problem, $tf + \phi$ should satisfy conditions on page 205

6.10 Primal-dual Interior-point Methods

Primal-dual & barrier interior-point methods

- in primal-dual interior-point methods
  - both primal and dual variables are updated at each iteration
  - search directions are obtained from Newton's method, applied to modified KKT equations, *i.e.*, optimality conditions for logarithmic barrier centering problem
  - primal-dual search directions are similar to, but not quite the same as, search directions arising in barrier methods
  - primal and dual iterates are not necessarily feasible
- primal-dual interior-point methods
  - often more efficient than barrier methods especially when high accuracy is required - can exhibit better than linear convergence
  - (customized versions) outperform barrier method for several basic problems, such as, LP, QP, SOCP, GP, SDP
  - can work for feasible, but *not* strictly feasible problems
  - still active research topic, but show great promise

Modified KKT conditions and central points

- modified KKT conditions (for convex optimization in Definition 6.40) expressed as

$$r_t(x,\lambda,\nu) = \begin{bmatrix} \nabla f(x) + Dq(x)^T\lambda + A^T\nu \\ -\mathbf{diag}(\lambda)f(x) - (1/t)\mathbf{1} \\ Ax - b \end{bmatrix}$$

  where

$$\begin{aligned} \textit{dual residual} \quad &- \quad r_{\text{dual}}(x,\lambda,\nu) = \nabla f(x) + Dq(x)^T\lambda + A^T\nu \\ \textit{centrality residual} \quad &- \quad r_{\text{cent}}(x,\lambda,\nu) = -\mathbf{diag}(\lambda)f(x) - (1/t)\mathbf{1} \\ \textit{primal residual} \quad &- \quad r_{\text{pri}}(x,\lambda,\nu) = Ax - b \end{aligned}$$

- if x, λ, ν satisfy $r_t(x,\lambda,\nu) = 0$ (and $q(x) \prec 0$), then
  - $x = x^*(t)$, $\lambda = \lambda^*(t)$, $\nu = \nu^*(t)$
  - x is primal feasible and λ & ν are dual feasible with duality gap m/t

Primal-dual search direction

- assume current (primal-dual) point $y = (x,\lambda,\nu)$ and Newton step $\Delta y = (\Delta x, \Delta\nu, \Delta\lambda)$
- as before, linearize equation to obtain Newton step, *i.e.*,

$$r_t(y+\Delta y) \approx r_t(y) + Dr_t(y)\Delta y = 0 \quad \Leftrightarrow \quad \Delta y = -Dr_t(y)^{-1}r_t(y)$$

  hence

$$\begin{bmatrix} \nabla^2 f(x) + \sum \lambda_i \nabla^2 q_i(x) & Dq(x)^T & A^T \\ -\mathbf{diag}(\lambda)Df(x) & -\mathbf{diag}(f(x)) & 0 \\ A & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta\lambda \\ \Delta\nu \end{bmatrix} = -\begin{bmatrix} r_{\text{dual}} \\ r_{\text{cent}} \\ r_{\text{pri}} \end{bmatrix}$$

- above equation determines *primal-dual search direction* $\Delta y_{\text{pd}} = (\Delta x_{\text{pd}}, \Delta\lambda_{\text{pd}}, \Delta\nu_{\text{pd}})$

Surrogate duality gap

- iterates $x^{(k)}$, $\lambda^{(k)}$, and $\nu^{(k)}$ of primal-dual interior-point method are *not* necessarily feasible
- hence, cannot easily evaluate duality gap $\eta^{(k)}$ as for barrier method
- define *surrogate duality gap* for $q(x) \prec 0$ and $\lambda \succeq 0$ as

$$\hat{\eta}(x,\lambda) = -q(x)^T\lambda$$

- $\hat{\eta}$ would be duality gap if x were primal feasible and λ & ν were dual feasible
- value t corresponding to surrogate duality gap $\hat{\eta}$ is $m/\hat{\eta}$

Primal-dual interior-point method

Algorithm 6.10 (primal-dual interior-point method)

Require: *initial point* x *with* $q(x) \prec 0$, $\lambda \succ 0$, $\mu > 1$, $\epsilon_{\text{pri}} > 0$, $\epsilon_{\text{dual}} > 0$, $\epsilon > 0$

  repeat

    set $t := \mu m/\hat{\eta}$

    computer primal-dual search direction $\Delta y_{\text{pd}} = (\Delta x_{\text{pd}}, \Delta\lambda_{\text{pd}}, \Delta\nu_{\text{pd}})$

    do line search to choose $s > 0$

    update - $x := x + s\Delta x_{\text{pd}}$, $\lambda := \lambda + s\Delta\nu_{\text{pd}}$, $\nu := \nu + s\Delta\nu_{\text{pd}}$

  until $\|r_{\text{pri}}(x,\lambda,\nu)\|_2 \le \epsilon_{\text{pri}}$, $\|r_{\text{dual}}(x,\lambda,\nu)\|_2 \le \epsilon_{\text{dual}}$, $\hat{\eta} \le \epsilon$

- common to choose small ϵ_{pri}, ϵ_{dual}, & ϵ since primal-dual method often shows faster than linear convergence

Line search for primal-dual interior-point method

- liner search is standard backtracking line search on $r(x, \lambda, \nu)$ similar to that in Algorithm 6.7 except making sure that $q(x) \prec 0$ and $\lambda \succ 0$
- note initial s in Algorithm 6.11 is largest s that makes $\lambda + s\Delta\lambda_{\mathrm{pd}}$ positive

Algorithm 6.11 (backtracking line search for primal-dual interior-point method)

Require: Δx_{pd}, $\Delta\lambda_{\mathrm{pd}}$, $\Delta\nu_{\mathrm{pd}}$, $\alpha \in (0.01, 0.1)$, $\beta \in (0.3, 0.8)$

  $s := 0.99 \sup\{s \in [0,1] | \lambda + s\Delta\lambda \succeq 0\} = 0.99 \min\{1, \min\{-\lambda_i/\Delta\lambda_i | \Delta\lambda_i < 0\}\}$

  while $q(x + s\Delta x_{\mathrm{pd}}) \not\prec 0$ ***do***

    $t := \beta t$

  end while

  while $\|r(x + s\Delta x_{\mathrm{pd}}, \lambda + s\Delta\lambda_{\mathrm{pd}}, \nu + s\Delta\nu_{\mathrm{pd}})\|_2 > (1 - \alpha s)\|r(x, \lambda, \nu)\|_2$ ***do***

    $t := \beta t$

  end while

7 Selected Proofs

Selected proofs

- **Proof 1** *(Proof for "relation among coset indices" on page 29)*

  Let $\{h_1, \ldots, h_n\}$ and $\{k_1, \ldots, k_m\}$ be coset representations of H in G and K in H respectively. Then $n = (G : H)$ and $m = (H : K)$. Note that $\bigcup_{i,j} h_i k_j K = \bigcup_i h_i H = G$, and if $h_i k_j K = h_k k_l K$ for some $1 \leq i, k \leq n$ and $1 \leq j, k \leq m$, $h_i k_j KH = h_k k_l KH \Leftrightarrow h_i k_j H = h_k k_l H \Leftrightarrow h_i H = h_j H \Leftrightarrow h_i = h_j$, thus $k_j K = k_l K$, hence $k_j = k_l$. Thus $\{h_i k_j | 1 \leq i \leq n, 1 \leq j \leq m\}$ is cosets representations of K in G, therefore $(G : K) = mn = (G : H)(H : K)$. ∎

- **Proof 2** *(Proof for "normality and commutativity of commutator subgroups" on page 33)*

  - For $a, x, y \in G$,

    $$\begin{aligned} axyx^{-1}y^{-1} &= ax(a^{-1}x^{-1}xa)yx^{-1}y^{-1}(a^{-1}a) \\ &= (axa^{-1}x^{-1})(x(ay)x^{-1}(ay)^{-1})a \end{aligned}$$

    and

    $$\begin{aligned} xyx^{-1}y^{-1}a &= (aa^{-1})xyx^{-1}(ay^{-1}ya^{-1})y^{-1}a \\ &= a((a^{-1}x)y(a^{-1}x)^{-1}y^{-1})(ya^{-1}y^{-1}a), \end{aligned}$$

    hence commutator subgroup of G propagate every element of G from fron to back and vice versa. Therefore for every $a \in G$, $aG^C = G^C a$.
  - For $x, y \in G$, $xG^C yG^C = xyG^C = G^C xy = (G^C x)(G^C y)$, hence G/G^C is commutative.
  - For a homeomorphism of G, f, into a commutative group, and $x, y \in G$,

    $$f(xyx^{-1}y^{-1}) = f(x)f(y)f(x^{-1})f(y^{-1}) = f(x)f(x^{-1})f(y)f(y^{-1}) = e$$

    thus $xyx^{-1}y^{-1} \in \operatorname{Ker} f$, hence $G^C \subset \operatorname{Ker} f$.

  ∎

- **Proof 3** *(Proof for "ideal generated by elements of ring" on page 40)*

  For all $x \in (a_1, \ldots, a_n)$, and $y \in A$ $yx = y\left(\sum x_i a_i\right) = \sum (yx_i) a_i$ for some $\langle x_i \rangle_{i=1}^n \subset A$, hence $yx \in A$, and $(a_1, \ldots, a_n)$ is additive group, thus is ideal of A, hence

  $$\bigcap_{\mathfrak{a}:\text{ideal containing } a_1,\ldots,a_n} \mathfrak{a} \subset (a_1, \ldots, a_n)$$

  Conversely, if $\mathfrak{a}$ contains $a_1, \ldots, a_n$, $Aa_i \subset \mathfrak{a}$, hence for every sequence, $\langle x_i \rangle_{i=1}^n \subset A$, $\sum x_i a_i \subset \mathfrak{a}$ because $\mathfrak{a}$ is additive subgroup of A, thus $(a_1, \ldots, a_n)$ is contained in every ideal containing a_1, …, a_n, hence

  $$(a_1, \ldots, a_n) \subset \bigcap_{\mathfrak{a}:\text{ideal containing } a_1,\ldots,a_n} \mathfrak{a}$$

  ∎

- **Proof 4** *(Proof for "nonzero ideals of integers are principal" on page 42)*

  Suppose $\mathfrak{a}$ is a nonzero ideal of $\mathbf{Z}$. Because if negative integer, n, is in $\mathfrak{a}$, $-n$ is also in $\mathfrak{a}$ because $\mathfrak{a}$ is an additive group in the ring, $\mathbf{Z}$. Thus, $\mathfrak{a}$ has at least one positive integer. By Principle 4.2, there exists the smallest positive integer in $\mathfrak{a}$. Let n be that integer. Let $m \in \mathfrak{a}$. By Theorem 3.13, there exist $q, r \in \mathbf{Z}$ such that $m = qn + r$ with $0 \le r < n$. Since by the definition of ideals of rings, *i.e.*, Definition 3.43, $\mathfrak{a}$ is an additive group in $\mathbf{Z}$, hence $m - qn = r$ is also in $\mathfrak{a}$, thus r should be 0 because we assume n is the smallest positive integer in $\mathfrak{a}$. Thus $\mathfrak{a} = \{qn | q \in \mathbf{Z}\} = n\mathbf{Z}$. Therefore the ideal is either $\{0\}$ or $n\mathbf{Z}$ for some $n > 0$. Both $\{0\}$ and $n\mathbf{Z}$ are ideal. ∎

- **Proof 5** *(Proof for "algebraic-ness of smallest subfields" on page 51)*

  Proposition 3.25 implies that $k(\alpha_1) = k[\alpha_1]$ and $[k(\alpha_1) : k] = \deg \operatorname{Irr}(\alpha_1, k, X)$. Because α_2 is algebraic over k, hence algebraic over $k(\alpha_1)$ *a fortiori*, thus, the same proposition implies $k(\alpha_1, \alpha_2) = (k(\alpha_1))[\alpha_2] = (k[\alpha_1])[\alpha_2] = k[\alpha_1, \alpha_2]$ and $[k(\alpha_1, \alpha_2) : k(\alpha_1)] = \operatorname{Irr}(\alpha_2, k(\alpha_1), X)$, hence $[k(\alpha_1, \alpha_2) : k] = \operatorname{Irr}(\alpha_1, k, X) \operatorname{Irr}(\alpha_2, k(\alpha_1), X)$. The mathematical induction will surely prove that $k(\alpha_1, \ldots, \alpha_n) = k[\alpha_1, \ldots, \alpha_n]$ and

  $$\begin{aligned} &[k(\alpha_1, \ldots, \alpha_n) : k] \\ &\quad = \operatorname{Irr}(\alpha_1, k, X) \operatorname{Irr}(\alpha_2, k(\alpha_1), X) \cdots \operatorname{Irr}(\alpha_n, k(\alpha_1, \ldots, \alpha_{n-1}), X). \end{aligned}$$

  ∎

- **Proof 6** *(Proof for "finite generation of compositum" on page 51)*

  First, it is obvious that $E = k(\alpha_1, \ldots, \alpha_n) \subset F(\alpha_1, \ldots, \alpha_n)$ and $F \subset F(\alpha_1, \ldots, \alpha_n)$, hence $EF \subset F(\alpha_1, \ldots, \alpha_n)$ because EF is defined to be the smallest subfield that contains both E and F. Now every subfield containing both E and F contains all $f(\alpha_1, \ldots, \alpha_n)$ where $f \in F[X]$, hence all $f(\alpha_1, \ldots, \alpha_n)/g(\alpha_1, \ldots, \alpha_n)$ where $f, g \in F[X]$ and $g(\alpha_1, \ldots, \alpha_n) \neq 0$. Thus, $F(\alpha_1, \ldots, \alpha_n) \subset EF$ again by definition. Therefore $EF = F(\alpha_1, \ldots, \alpha_n)$. ∎

- **Proof 7** *(Proof for "existence of algebraically closed algebraic extensions" on page 55)*

  Theorem 3.17 implies there exists an algebraically closed extension of k. Let E be such one. Let K be union of all algebraic extensions of k contained in E, then K is algebraic over k. Since k is algebraic over itself, K is not empty. Let $f \in K[X]$ with $\deg f \ge 1$. If α is a root of f, $\alpha \in E$. Since $K(\alpha)$ is algebraic over K and K is algebraic over k, $K(\alpha)$ is algebraic over k by Proposition 3.27. Therefore $K(\alpha) \subset K$ and $\alpha \in K$. Thus, K is algebraically closed algebraic extension of k. ∎

- **Proof 8** *(Proof for "theorem - Galois subgroups associated with intermediate fields" on page 61)*

  Suppsoe $\alpha \in K^G$ and let $\sigma : k(\alpha) \to K^{\mathrm{a}}$ be an embedding inducing the identity on k. If we let $\tau : K \to K^{\mathrm{a}}$ extend σ, τ is automorphism by normality of K/k (Definition 3.97), hence $\tau \in G$, thus τ fixed α, which means σ is the identity, which is the only embedding extension of the identity embedding of k onto itself to $k(\alpha)$, thus, by Definition 3.98,

  $$[k(\alpha) : k]_s = 1.$$

  Since K is separable over k, α is separable over k (by Theorem 3.26), and $k(\alpha)$ is separable over k (by Definition 3.100), thus $[k(\alpha) : k] = [k(\alpha) : k]_s = 1$, hence $k(\alpha) = k$, thus $\alpha \in k$, hence

  $$K^G \subset k.$$

  Since by definition, $k \subset K^G$, we have $K^G = k$.

  Now since K/k is a normal extension, K/F is also a normal extension (by Theorem 3.23). Also, since K/k is a separable extension, K/F is also separable extension (by Theorem 3.28 and Definition 3.90). Thus, K/F is Galois (by Definition 3.110).

  Now let F and F' be two intermediate fields. Since $K^{G(K/k)} = k$, we have $K^{G(K/F)} = F$ and $K^{G(K/F')} = F'$, thus if $G(K/F) = G(K/F')$, $F = F'$, hence the map is injective. ∎

- **Proof 9** *(Proof for "Galois subgroups associated with intermediate fields - 1" on page 61)*

  First, K/F_1 and K/F_2 are Galois extensions by Theorem 3.35, hence $G(K/F_1)$ and $G(K/F_2)$ can be defined. Also, Theorem 3.23 and Theorem 3.28 imply that K/F_1F_2 is Galois extension, hence $G(K/F_1F_2)$ can be defined, too.

  Every automorphism of G leaving both F_1 and F_2 leaves F_1F_2 fixed, hence $G(K/F_1) \cap G(K/F_2) \subset G(K/F_1F_2)$. Conversely, every automorphism of G leaving F_1F_2 fxied leaves both F_1 and F_2 fixed, hence $G(K/F_1F_2) \subset G(K/F_1) \cap G(K/F_2)$.

  Now we can do the same thing using rather mathematically rigorous terms. Assume that $\sigma \in G(K/F_1) \cap G(K/F_2)$. Then

  $$(\forall x \in F_1, y \in F_2)\, (x^\sigma = x \;\&\; y^\sigma = y)\,,$$

  thus

  $$\begin{aligned}&(\forall n, m \in \mathbf{N})\\ &\quad(\forall x_1, \ldots, x_n, x_1', \ldots, x_m' \in F_1, y_1, \ldots, y_n, y_1', \ldots, y_m' \in F_2)\\ &\quad\left(\left(\frac{x_1y_1 + \cdots + x_ny_n}{x_1'y_1' + \cdots + x_m'y_m'}\right)^\sigma = \frac{x_1y_1 + \cdots + x_ny_n}{x_1'y_1' + \cdots + x_m'y_m'}\right),\end{aligned}$$

  hence $\sigma \in G(K/F_1F_2)$, thus $G(K/F_1) \cap G(K/F_2) \subset\in G(K/F_1F_2)$. Conversely if $\sigma \in G(K/F_1F_2)$,

  $$(\forall x \in F_1, y \in F_2)\, (x^\sigma = x \;\&\; y^\sigma = y)\,,$$

  hence $\sigma \in G(K/F_1) \cap G(K/F_2)$, thus $G(K/F_1) \cap G(K/F_2) \subset G(K/F_1F_2)$. ∎

- **Proof 10** *(Proof for "Galois subgroups associated with intermediate fields - 3" on page 61)*

  First, K/F_1 and K/F_2 are Galois extensions by Theorem 3.35, hence $G(K/F_1)$ and $G(K/F_2)$ can be defined.

  If $F_1 \subset F_2$, every automorphism leaving F_2 fixed leaves F_1 fixed, hence it is in $G(K/F_1)$, thus $G(K/F_2) \subset G(K/F_1)$. Conversely, if $G(K/F_2) \subset G(K/F_1)$, every intermediate field $G(K/F_1)$ leaves fixed is left fixed by $G(K/F_2)$, hence $F_1 \subset F_2$.

  Now we can do the same thing using rather mathematically rigorous terms. Assume $F_1 \subset F_2$ and that $\sigma \subset G(K/F_2)$. Since Theorem 3.35 implies that

  $$F_1 \subset F_2 = \{x \in K | (\forall \sigma \in G(K/F_2))(x^\sigma = x)\},$$

  hence $(\forall x \in F_1)\, (x^\sigma = x)$, thus $\sigma \in G(K/F_1)$, hence

  $$G(K/F_2) \subset G(K/F_1).$$

  Conversely, assume that $G(K/F_2) \subset G(K/F_1)$. Then

  $$\begin{aligned}F_1 &= \{x \in K | (\forall \sigma \in G(K/F_1))(x^\sigma = x)\}\\ &\subset \{x \in K | (\forall \sigma \in G(K/F_2))(x^\sigma = x)\} = F_2\end{aligned}$$

  ∎

- **Proof 11** *(Proof for "Bolzano-Weierstrass-implies-seq-compact" on page 91)*

  if sequence, $\langle x_n \rangle$, has cluster point, x, every ball centered at x contains at one least point in sequence, hence, can choose subsequence converging to x. conversely, if $\langle x_n \rangle$ has subsequence converging to x, x is cluster point. ∎

- **Proof 12** *(Proof for "compact-in-metric-implies-seq-compact" on page 91)*

  for $\langle x_n \rangle$, $\langle \overline{A_n} \rangle$ with $A_m = \langle b_n \rangle_{n=m}^\infty$ has *finite intersection property* because any finite subcollection $\{A_{n_1}, \ldots, A_{n_k}\}$ contains x_{n_k}, hence

  $$\bigcap \overline{A_n} \neq \emptyset,$$

  thus, there exists $x \in X$ contained in every A_n. x is cluster point because for every $\epsilon > 0$ and $N \in \mathbf{N}$, then $x \in \overline{A_{N+1}}$, hence there exists $n > N$ such that x_n contained in ball about x with radius, ϵ. hence it's sequentially compact. ∎

- **Proof 13** *(Proof for "restriction-of-continuous-topology-continuous" on page 96)*

  because for every open set O, $g^{-1}(O) \in \mathfrak{J}$, $A \cap g^{-1}(O)$ is open by definition of *inherited topology*. ∎

- **Proof 14** *(Proof for "l-infinity-not-have-natural-representation" on page 109)*

  $C[0,1]$ is *closed* subspace of $L^\infty[0,1]$. define $f(x)$ for $x \in C[0,1]$ such that $f(x) = x(0) \in \mathbf{R}$. f is linear functional because $f(\alpha x + \beta y) = \alpha x(0) + \beta y(0) = \alpha f(x) + \beta(y)$. because $|f(x)| = |x(0)| \leq \|x\|_\infty$, $\|f\| \leq 1$. for $x \in C[0,1]$ such that $x(t) = 1$ for $0 \leq t \leq 1$, $|f(x)| = 1 = \|x\|_\infty$, hence achieves supremum, thus $\|f\| = 1$.

  if we define linear functional p on $L^\infty[0,1]$ such that $p(x) = f(x)$, $p(x+y) = x(0)+y(0) = p(x)+p(y) \leq p(x) + p(y)$, $p(\alpha x) = \alpha x(0) = \alpha p(x)$, and $f(x) \leq p(x)$ for all $x, y \in L^\infty[0,1]$ and $\alpha \geq 0$, and $f(s) = p(s) \leq p(s)$ for all $s \in C[0,1]$. Hence, Hahn-Banach theorem implies, exists $F : L^\infty[0,1] \to \mathbf{R}$ such that $F(x) = f(x)$ for every $x \in C[0,1]$ and $F(x) \leq f(x)$ for every $x \in L^\infty[0,1]$.

  Now assume $y \in L^1[0,1]$ such that $F(x) = \int_{[0,1]} xy$ for $x \in C[0,1]$. If we define $\langle x_n \rangle$ in $C[0,1]$ with $x_n(0) = 1$ vanishing outside $t = 0$ as $n \to \infty$, then $\int_{[0,1]} x_n y \to 0$ as $n \to \infty$, but $F(x_n) = 1$ for all n, hence, contradiction. Therefore there is not natural representation for F. ∎

- **Proof 15** *(Proof for "orthonormal-system" on page 116)*

  Assume $\langle \varphi_n \rangle$ is complete, but not maximal. Then there exists orthonormal system, R, such that $\langle \varphi_n \rangle \subset R$, but $\langle \varphi_n \rangle \neq R$. Then there exists another $z \in R$ such that $z \notin \langle \varphi_n \rangle$. But definition $\langle z, \varphi_n \rangle = 0$, hence $z = 0$. But $\|z\| = 0$, hence, cannot be member of orthonormal system. contraction, hence proved right arrow, *i.e.*, sufficient condition (of the former for the latter).

  Now assume that it is maximal. Assume there exists $z \neq 0 \in H$ such that $\langle z, \varphi_n \rangle = 0$. Then $\langle \varphi_n \rangle_{n=0}^\infty$ with $\varphi_0 = z/\|z\|$ is anoter orthogonal system containing $\langle \varphi_n \rangle$, hence contradiction, thus proved left arrow, *i.e.*, necessarily condition. ∎

- **Proof 16** *(Proof for "central limit theorem" on page 141)*

  Let $Z_n(t) = t^T(X_n - c)$ for $t \in \mathbf{R}^k$ and $Z(t) = t^T Y$. Then $\langle Z_n(t) \rangle$ are independent random variables having same distribution with $\mathbf{E}\, Z_n(t) = t^T(\mathbf{E}\, X_n - c) = 0$ and

  $$\mathbf{Var}\, Z_n(t) = \mathbf{E}\, Z_n(t)^2 = t^T\, \mathbf{E}(X_n - c)(X_n - c)^T t = t^T \Sigma t$$

  Then by Theorem 5.8 $\sum^n Z_i(t)/\sqrt{nt^T\Sigma t}$ converges in distribution to standard normal random variable. Because $\mathbf{E}\, Z(t) = 0$ and $\mathbf{Var}\, Z(t) = t^T\, \mathbf{E}\, YY^T z = t^T \Sigma t$, for $t \neq 0$, $Z(t)/\sqrt{t^T\Sigma t}$ is standard normal random variable. Therefore $\sum^n Z_i(t)/\sqrt{nt^T\Sigma t}$ converges in distribution to $Z/\sqrt{t^T\Sigma t}$ for every $t \neq 0$, thus, $\sum^n Z_i(t)/\sqrt{n} = t^T(\sum^n X_i - nc)/\sqrt{n}$ converges in distribution to $Z(t) = t^T Y$ for every $t \in \mathbf{R}$. Then Theorem 5.10 implies $(S_n - nc)/\sqrt{n}$ converges in distribution to Y. ∎

- **Proof 17** *(Proof for "intersection of convex sets is convex set" on page 145)*

  Suppose $\mathcal{C}$ is a collection of convex sets. Suppose $x, y \in \bigcap_{C \in \mathcal{C}} C$ and $0 < \theta < 1$. Then for each $C \in \mathcal{C}$ and $\theta x + (1-\theta)y \in C$, hence, $\theta x + (1-\theta)y \in \bigcap_{C \in \mathcal{C}} C$, $\bigcap_{C \in \mathcal{C}} C$ is a convex set. ∎

- **Proof 18** *(Proof for "theorem of alternative for linear strict generalized inequalities" on page 148)*

  Suppose $Ax \prec_K b$ is infeasible. Then $\{b - Ax | x \in \mathbf{R}^n\} \cap K^\circ = \emptyset$. Theorem 6.1 implies there exist nonzero $\lambda \in \mathbf{R}^n$ and $c \in \mathbf{R}$ such that

  $$(\forall x \in \mathbf{R}^n) \left(\lambda^T(b - Ax) \leq c\right) \tag{1}$$

  and

  $$(\forall y \in K^\circ) \left(\lambda^T y \geq c\right). \tag{2}$$

  The former equation (1) implies $\lambda^T A = 0$ and $\lambda^T b \leq c$. and the latter $a \succeq_{K^*} 0$. If $c > 0$, there exists $y \in K^\circ$ such that $\lambda^T y \geq c > 0$. Then $\lambda^T((c/2\lambda^T y)y) = c/2 < c$, but $(c/2\lambda^T y)y \in K^\circ$, hence contradiction. Thus, $c \leq 0$. If $\lambda^T y < 0$ for some $y \in K^\circ$, then $\alpha y \in K^\circ$ for any $\alpha > 0$, thus there exists $z \in K^\circ$ which makes $\lambda^T z$ arbitrarily large toward $-\infty$. Therefore $\lambda^T y$ is nonnegative for every $y \in K^\circ$. Then the latter equation (2) implies $(\forall y \in K^\circ) \left(\lambda^T y \geq 0\right)$, hence $\lambda \in K^*$ (by Definition 6.24). Therefore we have

  $$\lambda \neq 0,\ \lambda \succeq_{K^*} 0,\ A^T\lambda = 0,\ \lambda^T b \leq 0.$$

  Conversely, assume that all of above are satisfied. Then for every $x \in \mathbf{R}^n$, there exists nonzero $\lambda \succeq_{K^*} 0$ such that

  $$\lambda^T(Ax) \geq \lambda^T b,$$

  thus Proposition 6.3 implies $Ax \not\prec_K b$. ∎

- **Proof 19** *(Proof for "convexity of infimum of convex function" on page 152)*

  Note

  $$\begin{aligned}
  \mathbf{epi} \inf_{y\in C} f(x,y) &= \{(x,t) | (\forall \epsilon > 0)(\exists y \in C)(f(x,y) \leq t + \epsilon)\} \\
  &= \bigcap_{n\in\mathbf{N}} \{(x,t)\, | (\exists y \in C)(f(x,y,t+1/n) \in \mathbf{epi}\, f)\,\} \\
  &= \bigcap_{n\in\mathbf{N}} (\{(x,t)\, | (\exists y \in C)(f(x,y,t) \in \mathbf{epi}\, f)\,\} - (0,1/n))
  \end{aligned}$$

  where $\{(x,t)\, | (\exists y \in C)(f(x,y,t) \in \mathbf{epi}\, f)\,\} - (0,1/n)$ for each n since $\mathbf{epi}\, f$ is convex and projection of a convex set is convex. Since the intersection of any collection of convex sets is convex, $\mathbf{epi} \inf_{y\in C} f(x,y)$ is convex, thus $\inf_{y\in C} f(x,y)$ is convex function. ∎

- **Proof 20** *(Proof for "Lagrange dual is lower bound for optimal value" on page 164)*

  For every $\lambda \succeq 0$ and $y \in \mathcal{F}$

  $$g(\lambda,\nu) \leq f(y) + \lambda^T q(y) + \nu^T h(y) \leq f(y) \leq \inf_{x\in\mathcal{F}} f(x) = p^*.$$

  ∎

- **Proof 21** *(Proof for "epigraph of convex optimization is convex" on page 180)*

  Assume $(u_1, v_1, t_1), (u_2, v_2, t_2) \in H$. Then there exist $x_1, x_2 \in \mathcal{D}$ such that $q(x_1) \preceq u_1$, $h(x_1) = v_1$, $f(x_1) \leq t_1$, $q(x_2) \preceq u_2$, $h(x_2) = v_2$, and $f(x_2) \leq t_2$. Then for every $0 \leq \theta \leq 1$

  $$\begin{aligned}
  q(\theta x_1 + (1-\theta)x_2) &\preceq \theta q(x_1) + (1-\theta)q(x_2) = \theta u_1 + (1-\theta)u_2 \\
  h(\theta x_1 + (1-\theta)x_2) &= \theta h(x_1) + (1-\theta)h(x_2) = \theta v_1 + (1-\theta)v_2 \\
  f(\theta x_1 + (1-\theta)x_2) &\preceq \theta f(x_1) + (1-\theta)f(x_2) = \theta t_1 + (1-\theta)t_2
  \end{aligned}$$

  thus $\theta(u_1, v_1, t_1) + (1-\theta)(u_2, v_2, t_2) \in H$, hence H is a convex set. ∎

- **Proof 22** *(Proof for "max-min inequality" on page 177)*

  For every $x \in X, y \in Y$

  $$f(x,y) \leq \sup_{x' \in X} f(x',y)$$

  hence for every $x \in X$

  $$\inf_{y'' \in Y} f(x,y'') \leq \inf_{y' \in Y} \sup_{x' \in X} f(x',y')$$

  i.e., $\inf_{y' \in Y} \sup_{x' \in X} f(x',y')$ is upper bound of $\inf_{y'' \in Y} f(x,y'')$, hence

  $$\sup_{x \in X} \inf_{y'' \in Y} f(x,y'') \leq \inf_{y' \in Y} \sup_{x' \in X} f(x',y')$$

  ∎

Index