LECTURE 24: REVIEW/EPILOGUE

LECTURE OUTLINE

• Basic concepts of convex analysis
• Basic concepts of convex optimization
• Geometric duality framework - MC/MC
• Constrained optimization duality - minimax
• Subgradients - Optimality conditions
• Special problem classes
• Descent/gradient/subgradient methods
• Polyhedral approximation methods

All figures are courtesy of Athena Scientific, and are used with permission.
BASIC CONCEPTS OF CONVEX ANALYSIS

- Epigraphs, level sets, closedness, semicontinuity

- Finite representations of generated cones and convex hulls - Caratheodory’s Theorem.
- Relative interior:
  - Nonemptiness for a convex set
  - Line segment principle
  - Calculus of relative interiors
- Continuity of convex functions
- Nonemptiness of intersections of nested sequences of closed sets.
- Closure operations and their calculus.
- Recession cones and their calculus.
- Preservation of closedness by linear transformations and vector sums.
HYPERPLANE SEPARATION

- Separating/supporting hyperplane theorem.
- Strict and proper separation theorems.
- Dual representation of closed convex sets as unions of points and intersection of halfspaces.

A union of points  
An intersection of halfspaces

- Nonvertical separating hyperplanes.
CONJUGATE FUNCTIONS

- Conjugacy theorem: $f = f^{**}$
- Support functions

- Polar cone theorem: $C = C^{**}$
  - Special case: Linear Farkas’ lemma
POLYHEDRAL CONVEXITY

• Extreme points

![Extreme Points](a) ![Extreme Points](b) ![Extreme Points](c)

• A closed convex set has at least one extreme point if and only if it does not contain a line.

• Polyhedral sets.

• Finitely generated cones: $C = \text{cone}(\{a_1, \ldots, a_r\})$

• **Minkowski-Weyl Representation:** A set $P$ is polyhedral if and only if

$$P = \text{conv}(\{v_1, \ldots, v_m\}) + C,$$

for a nonempty finite set of vectors $\{v_1, \ldots, v_m\}$ and a finitely generated cone $C$.

• **Fundamental Theorem of LP:** Let $P$ be a polyhedral set that has at least one extreme point. A linear function that is bounded below over $P$, attains a minimum at some extreme point of $P$. 
BASIC CONCEPTS OF CONVEX OPTIMIZATION

- **Weierstrass Theorem** and extensions.
- Characterization of existence of solutions in terms of nonemptiness of nested set intersections.

- Role of recession cone and lineality space.

- **Partial Minimization Theorems:** Characterization of closedness of \( f(x) = \inf_{z \in \mathbb{R}^m} F(x, z) \) in terms of closedness of \( F \).

![Level Sets of \( f \)](image1)

![Optimal Solution](image2)

![F(x, z)](image3)

![epi(f)](image4)
MIN COMMON/MAX CROSSING DUALITY

- Defined by a single set $M \subset \mathbb{R}^{n+1}$.
- $w^* = \inf_{(0,w) \in M} w$
- $q^* = \sup_{\mu \in \mathbb{R}^n} q(\mu) \triangleq \inf_{(u,w) \in M} \{w + \mu'u\}$
- Weak duality: $q^* \leq w^*$
- Two key questions:
  - When does strong duality $q^* = w^*$ hold?
  - When do there exist optimal primal and dual solutions?
MC/MC THEOREMS \((\overline{M} \text{ CONVEX, } W^* < \ )\)

- **MC/MC Theorem I**: We have \(q^* = w^*\) if and only if for every sequence \(\{(u_k, w_k)\} \subset M\) with \(u_k \to 0\), there holds

\[
w^* \leq \lim \inf_{k \to \infty} w_k.
\]

- **MC/MC Theorem II**: Assume in addition that \(-\infty < w^*\) and that

\[
D = \{u \mid \text{there exists } w \in \mathbb{R} \text{ with } (u, w) \in M\}
\]

contains the origin in its relative interior. Then \(q^* = w^*\) and there exists \(\mu\) such that \(q(\mu) = q^*\).

- **MC/MC Theorem III**: Similar to II but involves special polyhedral assumptions.

1. \(M\) is a “horizontal translation” of \(\tilde{M}\) by \(-P\),

\[
M = \tilde{M} - \{(u, 0) \mid u \in P\},
\]

where \(P\): polyhedral and \(\tilde{M}\): convex.

2. We have \(\text{ri}(\tilde{D}) \cap P \neq \emptyset\), where

\[
\tilde{D} = \{u \mid \text{there exists } w \in \mathbb{R} \text{ with } (u, w) \in \tilde{M}\}.
\]
**IMPORTANT SPECIAL CASE**

- **Constrained optimization:** $\inf_{x \in X, g(x) \leq 0} f(x)$
- **Perturbation function** (or *primal function*)
  
  \[
p(u) = \inf_{x \in X, g(x) \leq u} f(x),
  \]

- Introduce $L(x, \mu) = f(x) + \mu g(x)$. Then
  
  \[
  q(\mu) = \inf_{u \in \mathbb{R}^r} \left\{ p(u) + \mu' u \right\} \\
  = \inf_{u \in \mathbb{R}^r, x \in X, g(x) \leq u} \left\{ f(x) + \mu' u \right\} \\
  = \begin{cases} 
  \inf_{x \in X} L(x, \mu) & \text{if } \mu \geq 0, \\
  -\infty & \text{otherwise.}
  \end{cases}
  \]
**NONLINEAR FARKAS’ LEMMA**

- Let $X \subseteq \mathbb{R}^n$, $f : X \mapsto \mathbb{R}$, and $g_j : X \mapsto \mathbb{R}$, $j = 1, \ldots, r$, be convex. Assume that
  
  $$f(x) \geq 0, \quad \forall x \in X \text{ with } g(x) \leq 0$$

  Let
  
  $$Q^* = \{ \mu \mid \mu \geq 0, f(x) + \mu'g(x) \geq 0, \forall x \in X \}.$$

- **Nonlinear version:** Then $Q^*$ is nonempty and compact if and only if there exists a vector $x \in X$ such that $g_j(x) < 0$ for all $j = 1, \ldots, r$.

- **Polyhedral version:** $Q^*$ is nonempty if $g$ is linear $[g(x) = Ax - b]$ and there exists a vector $x \in \text{ri}(X)$ such that $Ax - b \leq 0$. 

![Graph](image.png)
CONSTRAINED OPTIMIZATION DUALITY

minimize \( f(x) \)
subject to \( x \in X, \ g_j(x) \leq 0, \ j = 1, \ldots, r, \)

where \( X \subset \mathbb{R}^n, \ f : X \mapsto \mathbb{R} \) and \( g_j : X \mapsto \mathbb{R} \) are convex. Assume \( f^* : \) finite.

- **Connection with MC/MC:** \( M = \text{epi}(p) \) with \( p(u) = \inf_{x \in X, g(x) \leq u} f(x) \)

- **Dual function:**

\[
q(\mu) = \begin{cases} 
\inf_{x \in X} L(x, \mu) & \text{if } \mu \geq 0, \\
-\infty & \text{otherwise}
\end{cases}
\]

where \( L(x, \mu) = f(x) + \mu'g(x) \) is the Lagrangian function.

- **Dual problem** of maximizing \( q(\mu) \) over \( \mu \geq 0 \).

- **Strong Duality Theorem:** \( q^* = f^* \) and there exists dual optimal solution if one of the following two conditions holds:

1. There exists \( x \in X \) such that \( g(x) < 0 \).
2. The functions \( g_j, \ j = 1, \ldots, r, \) are affine, and there exists \( x \in \text{ri}(X) \) such that \( g(x) \leq 0 \).
OPTIMALITY CONDITIONS

- We have $q^* = f^*$, and the vectors $x^*$ and $\mu^*$ are optimal solutions of the primal and dual problems, respectively, iff $x^*$ is feasible, $\mu^* \geq 0$, and
  
  $$x^* \in \arg \min_{x \in X} L(x, \mu^*), \quad \mu^*_j g_j(x^*) = 0, \quad \forall \ j.$$  

- For the linear/quadratic program
  
  minimize $\frac{1}{2} x' Q x + c' x$

  subject to $Ax \leq b,$

  where $Q$ is positive semidefinite, $(x^*, \mu^*)$ is a primal and dual optimal solution pair if and only if:

  (a) Primal and dual feasibility holds:

  $$Ax^* \leq b, \quad \mu^* \geq 0$$

  (b) Lagrangian optimality holds [$x^*$ minimizes $L(x, \mu^*)$ over $x \in \mathbb{R}^n$]. (Unnecessary for LP.)

  (c) Complementary slackness holds:

  $$(Ax^* - b)' \mu^* = 0,$$

  i.e., $\mu^*_j > 0$ implies that the $j$th constraint is tight. (Applies to inequality constraints only.)
Fenchel Duality

- **Primal problem:**

\[
\begin{align*}
\text{minimize} & \quad f_1(x) + f_2(x) \\
\text{subject to} & \quad x \in \mathbb{R}^n,
\end{align*}
\]

where \( f_1 : \mathbb{R}^n \rightarrow (-\infty, \infty] \) and \( f_2 : \mathbb{R}^n \rightarrow (-\infty, \infty] \) are closed proper convex functions.

- **Dual problem:**

\[
\begin{align*}
\text{minimize} & \quad f_1^*(\lambda) + f_2^*(-\lambda) \\
\text{subject to} & \quad \lambda \in \mathbb{R}^n,
\end{align*}
\]

where \( f_1^* \) and \( f_2^* \) are the conjugates.
CONIC DUALITY

• Consider minimizing \( f(x) \) over \( x \in C \), where \( f : \mathbb{R}^n \rightarrow (-\infty, \infty] \) is a closed proper convex function and \( C \) is a closed convex cone in \( \mathbb{R}^n \).

• We apply Fenchel duality with the definitions

\[
f_1(x) = f(x), \quad f_2(x) = \begin{cases} 0 & \text{if } x \in C, \\ \infty & \text{if } x \notin C. \end{cases}
\]

• **Linear Conic Programming:**

\[
\begin{align*}
\text{minimize} & \quad c'x \\
\text{subject to} & \quad x - b \in S, \quad x \in C.
\end{align*}
\]

• The **dual linear conic** problem is equivalent to

\[
\begin{align*}
\text{minimize} & \quad b'\lambda \\
\text{subject to} & \quad \lambda - c \in S^\perp, \quad \lambda \in \hat{C}.
\end{align*}
\]

• **Special Linear-Conic Forms:**

\[
\begin{align*}
\min_{Ax=b, \ x \in C} c'x & \iff \max_{c-A'\lambda \in \hat{C}} b'\lambda, \\
\min_{Ax-b \in C} c'x & \iff \max_{A'\lambda=c, \ \lambda \in \hat{C}} b'\lambda,
\end{align*}
\]

where \( x \in \mathbb{R}^n, \ \lambda \in \mathbb{R}^m, \ c \in \mathbb{R}^n, \ b \in \mathbb{R}^m, \ A : m \times n. \)
$\partial f(x) = \emptyset$ for $x \in \text{ri}(\text{dom}(f))$.

**Conjugate Subgradient Theorem:** If $f$ is closed proper convex, the following are equivalent for a pair of vectors $(x, y)$:

(i) $x'y = f(x) + f^*(y)$.

(ii) $y \in \partial f(x)$.

(iii) $x \in \partial f^*(y)$.

**Characterization of optimal solution set**

$X^* = \arg \min_{x \in \mathbb{R}^n} f(x)$ of closed proper convex $f$:

(a) $X^* = \partial f^*(0)$.

(b) $X^*$ is nonempty if $0 \in \text{ri}(\text{dom}(f^*))$.

(c) $X^*$ is nonempty and compact if and only if $0 \in \text{int}(\text{dom}(f^*))$. 
CONSTRANDED OPTIMALITY CONDITION

- Let $f : \mathbb{R}^n \mapsto (-\infty, \infty]$ be proper convex, let $X$ be a convex subset of $\mathbb{R}^n$, and assume that one of the following four conditions holds:

  (i) $\text{ri}(\text{dom}(f)) \cap \text{ri}(X) \neq \emptyset$.

  (ii) $f$ is polyhedral and $\text{dom}(f) \cap \text{ri}(X) \neq \emptyset$.

  (iii) $X$ is polyhedral and $\text{ri}(\text{dom}(f)) \cap X = \emptyset$.

  (iv) $f$ and $X$ are polyhedral, and $\text{dom}(f) \cap X \neq \emptyset$.

Then, a vector $x^*$ minimizes $f$ over $X$ iff there exists $g \in \partial f(x^*)$ such that $-g$ belongs to the normal cone $N_X(x^*)$, i.e.,

$$g'(x - x^*) \geq 0, \quad \forall x \in X.$$
COMPUTATION: PROBLEM RANKING IN INCREASING COMPUTATIONAL DIFFICULTY

- Linear and (convex) quadratic programming.
  - Favorable special cases.
- Second order cone programming.
- Semidefinite programming.
- Convex programming.
  - Favorable cases, e.g., separable, large sum.
  - Geometric programming.
- Nonlinear/nonconvex/continuous programming.
  - Favorable special cases.
  - Unconstrained.
  - Constrained.
- Discrete optimization/Integer programming
  - Favorable special cases.
- Caveats/questions:
  - Important role of special structures.
  - What is the role of “optimal algorithms”? 
  - Is complexity the right philosophical view to convex optimization?
DESCENT METHODS

- **Steepest descent method**: Use vector of min norm on $-\partial f(x)$; has convergence problems.

- **Subgradient method**:

- **Incremental** (possibly randomized) variants for minimizing large sums.

- **$\epsilon$-descent method**: Fixes the problems of steepest descent.
**APPROXIMATION METHODS I**

- **Cutting plane:**

  ![Cutting Plane Diagram]

  

- **Instability problem:** The method can make large moves that deteriorate the value of $f$.

- **Proximal Minimization method:**

  ![Proximal Minimization Diagram]

- **Proximal-cutting plane-bundle methods:** Combinations cutting plane-proximal, with stability control of proximal center.
APPROXIMATION METHODS II

- **Dual Proximal - Augmented Lagrangian methods:** Proximal method applied to the dual problem of a constrained optimization problem.

![Primal Proximal Iteration](image1)

![Dual Proximal Iteration](image2)

- **Interior point methods:**