LECTURE 12

LECTURE OUTLINE

- We transition from theory to algorithms
- The next two lectures provide:
  - An overview of interesting/challenging large-scale convex problem structures
  - An overview of fundamental algorithmic ideas for large-scale convex programming
- Problem Structures
  - Separable problems
  - Integer/discrete problems – Branch-and-bound
  - Large sum problems
  - Problems with many constraints
- Conic Programming
  - Second Order Cone Programming
  - Semidefinite Programming

SEPARABLE PROBLEMS

- Consider the problem

$$\text{minimize} \quad \sum_{i=1}^{m} f_i(x_i)$$

$$\text{s. t.} \ \sum_{i=1}^{m} g_{ji}(x_i) \leq 0, \quad j = 1, \ldots, r, \quad x_i \in X_i, \quad \forall\ i$$

where $f_i : \Re^{n_i} \mapsto \Re$ and $g_{ji} : \Re^{n_i} \mapsto \Re$ are given functions, and X_i are given subsets of $\Re^{n_i}$.

- Form the dual problem

$$\text{maximize} \ \sum_{i=1}^{m} q_i(\mu) \equiv \sum_{i=1}^{m} \inf_{x_i \in X_i} \left\{ f_i(x_i) + \sum_{j=1}^{r} \mu_j g_{ji}(x_i) \right\}$$

subject to $\mu \geq 0$

- Important point: The calculation of the dual function has been decomposed into m simpler minimizations.

- Another important point: If X_i is a discrete set (e.g., $X_i = \{0, 1\}$), the dual optimal value is a lower bound to the optimal primal value. It is still useful in a branch-and-bound scheme.

LARGE SUM PROBLEMS

- Consider cost function of the form

$$f(x) = \sum_{i=1}^{m} f_i(x), \qquad m \text{ is very large}$$

- Dual cost of a separable problem.

- Data analysis/machine learning. x is parameter vector of a model; each f_i corresponds to error between data and output of the model.
  - Least squares problems (f_i quadratic).
  - ℓ_1-regularization (least squares plus ℓ_1 penalty):

    $$\min_x \sum_{j=1}^{m} (a_j' x - b_j)^2 + \gamma \sum_{i=1}^{n} |x_i|$$

    The nondifferentiable penalty tends to set a large number of components of x to 0.
  - Maximum likelihood estimation.

- Min of an expected value $E\{F(x,w)\}$, where w is a random variable taking a finite but very large number of values w_i, $i = 1, \ldots, m$, with corresponding probabilities π_i. A special case: Stochastic programming.

- Special type of algorithms, called incremental apply (they operate on a single f_i at a time).

PROBLEMS WITH MANY CONSTRAINTS

- Problems of the form

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad a_j' x \le b_j, \quad j = 1, \ldots, r, \end{aligned}$$

where r: very large.

- One possibility is a penalty function approach: Replace problem with

$$\min_{x \in \Re^n} f(x) + c \sum_{j=1}^{r} P(a_j' x - b_j)$$

where $P(\cdot)$ is a scalar penalty function satisfying $P(t) = 0$ if $t \le 0$, and $P(t) > 0$ if $t > 0$, and c is a positive penalty parameter.

- Examples:
  - The quadratic penalty $P(t) = \big(\max\{0, t\}\big)^2$.
  - The nondifferentiable penalty $P(t) = \max\{0, t\}$.

- Another possibility: Initially discard some of the constraints, solve a less constrained problem, and later reintroduce constraints that seem to be violated at the optimum (outer approximation).

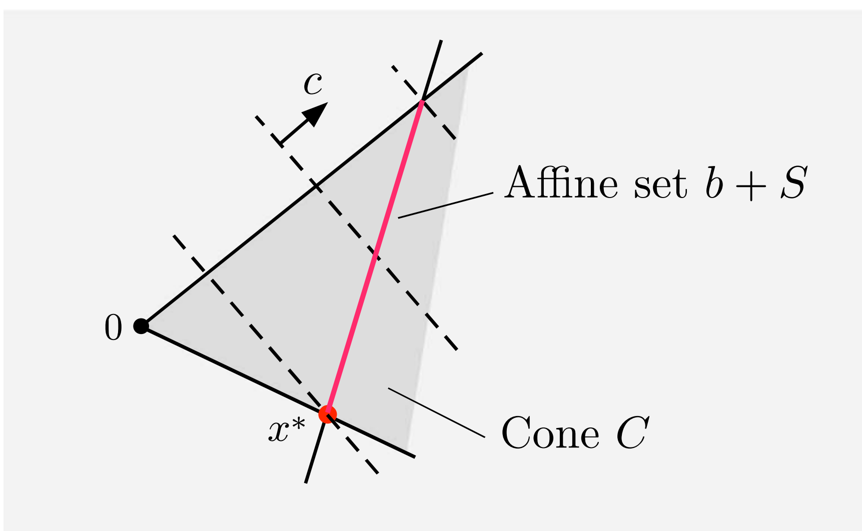
- Also inner approximation of the constraint set.

CONIC PROBLEMS

- A conic problem is to minimize a convex function $f : \Re^n \mapsto (-\infty, \infty]$ subject to a cone constraint.

- The most useful/popular special cases:
  - Linear-conic programming
  - Second order cone programming
  - Semidefinite programming

involve minimization of linear function $c'x$ over intersection of an affine set $b + S$ and a cone C.



- Can be analyzed as a special case of Fenchel duality.

- There are many interesting applications of conic problems, including in discrete optimization.

PROBLEM RANKING IN INCREASING PRACTICAL DIFFICULTY

- Linear and (convex) quadratic programming.
  - Favorable special cases (e.g., network flows).
- Second order cone programming.
- Semidefinite programming.
- Convex programming.
  - Favorable special cases (e.g., network flows, monotropic programming, geometric programming).
- Nonlinear/nonconvex/continuous programming.
  - Favorable special cases (e.g., twice differentiable, quasi-convex programming).
  - Unconstrained.
  - Constrained.
- Discrete optimization/Integer programming.
  - Favorable special cases.

CONIC DUALITY

- Consider minimizing $f(x)$ over $x \in C$, where $f : \Re^n \mapsto (-\infty, \infty]$ is a closed proper convex function and C is a closed convex cone in $\Re^n$.

- We apply Fenchel duality with the definitions

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

The conjugates are

$$f_1^\star(\lambda) = \sup_{x \in \Re^n} \bigl\{\lambda' x - f(x)\bigr\}, \; f_2^\star(\lambda) = \sup_{x \in C} \lambda' x = \begin{cases} 0 & \text{if } \lambda \in C^*, \\ \infty & \text{if } \lambda \notin C^*, \end{cases}$$

where $C^* = \{\lambda \mid \lambda' x \le 0, \, \forall \; x \in C\}$ is the polar cone of C.

- The dual problem is $\min_\lambda \bigl\{f_1^\star(\lambda) + f_2^\star(-\lambda)\bigr\}$, or

$$\begin{aligned} &\text{minimize} \quad f^\star(\lambda) \\ &\text{subject to} \;\; \lambda \in \hat{C}, \end{aligned}$$

where $f^\star$ is the conjugate of f and $\hat{C}$ is the dual cone ($= -C^*$, negative polar cone)

$$\hat{C} = \{\lambda \mid \lambda' x \ge 0, \, \forall \; x \in C\}$$

LINEAR-CONIC PROBLEMS

- Let f be affine, $f(x) = c'x$, with $\text{dom}(f)$ being an affine set, $\text{dom}(f) = b + S$, where S is a subspace.

- The primal problem is

$$\begin{aligned} &\text{minimize} \quad c'x \\ &\text{subject to} \;\; x - b \in S, \quad x \in C. \end{aligned}$$

- The conjugate is

$$f^\star(\lambda) = \sup_{x-b\in S} (\lambda - c)'x = \sup_{y\in S} (\lambda - c)'(y + b)$$
$$= \begin{cases} (\lambda - c)'b & \text{if } \lambda - c \in S^\perp, \\ \infty & \text{if } \lambda - c \notin S^\perp, \end{cases}$$

so the dual problem can be written as

$$\begin{aligned} &\text{minimize} \quad b'\lambda \\ &\text{subject to} \;\; \lambda - c \in S^\perp, \quad \lambda \in \hat{C}. \end{aligned}$$

- The primal and dual have the same form.

- If C is closed, the dual of the dual yields the primal.

SPECIAL LINEAR-CONIC FORMS

$$\min_{Ax=b,\ x\in C} c'x \quad \Longleftrightarrow \quad \max_{c-A'\lambda\in\hat{C}} b'\lambda,$$

$$\min_{Ax-b\in C} c'x \quad \Longleftrightarrow \quad \max_{A'\lambda=c,\ \lambda\in\hat{C}} b'\lambda,$$

where $x \in \Re^n$, $\lambda \in \Re^m$, $c \in \Re^n$, $b \in \Re^m$, $A : m\times n$.

- Proof of first relation: Let $\overline{x}$ be such that $A\overline{x} = b$, and write the problem on the left as

$$\begin{aligned} &\text{minimize} \quad c'x \\ &\text{subject to} \;\; x - \overline{x} \in \mathrm{N}(A), \quad x \in C \end{aligned}$$

- The dual conic problem is

$$\begin{aligned} &\text{minimize} \quad \overline{x}'\mu \\ &\text{subject to} \;\; \mu - c \in \mathrm{N}(A)^{\perp}, \quad \mu \in \hat{C} \end{aligned}$$

- Using $\mathrm{N}(A)^{\perp} = \mathrm{Ra}(A')$, write the constraints as $c - \mu \in -\mathrm{Ra}(A') = \mathrm{Ra}(A')$, $\mu \in \hat{C}$, or

$$c - \mu = A'\lambda, \qquad \mu \in \hat{C}, \qquad \text{for some } \lambda \in \Re^m$$

- Change variables $\mu = c - A'\lambda$, write the dual as

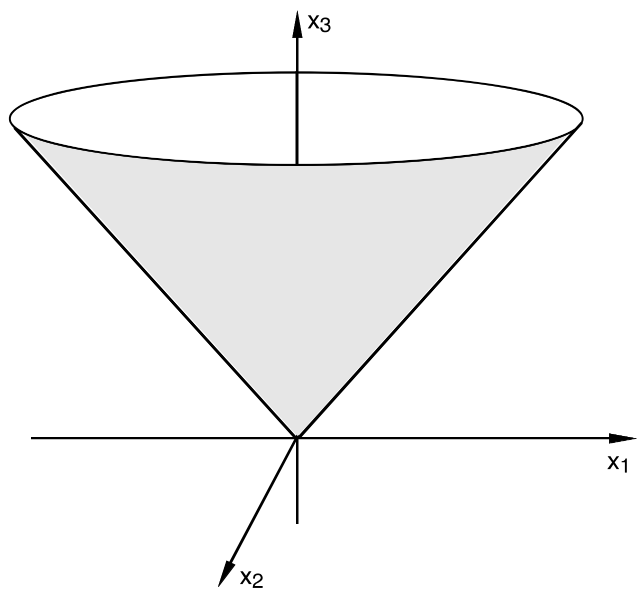
$$\begin{aligned} &\text{minimize} \quad \overline{x}'(c - A'\lambda) \\ &\text{subject to} \;\; c - A'\lambda \in \hat{C} \end{aligned}$$

discard the constant $\overline{x}'c$, use the fact $A\overline{x} = b$, and change from min to max.

SOME EXAMPLES

- Nonnegative Orthant: $C = \{x \mid x \geq 0\}$
- The Second Order Cone: Let

$$C = \left\{ (x_1, \ldots, x_n) \mid x_n \geq \sqrt{x_1^2 + \cdots + x_{n-1}^2} \right\}$$



- The Positive Semidefinite Cone: Consider the space of symmetric $n \times n$ matrices, viewed as the space $\Re^{n^2}$ with the inner product

$$< X, Y >= \text{trace}(XY) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} y_{ij}$$

Let C be the cone of matrices that are positive semidefinite.

- All these are self-dual, i.e., $C = -C^* = \hat{C}$.

SECOND ORDER CONE PROGRAMMING

- Second order cone programming is the linear-conic problem

$$\begin{aligned} &\text{minimize} \quad c'x \\ &\text{subject to} \;\; A_i x - b_i \in C_i, \;\; i = 1, \ldots, m, \end{aligned}$$

where c, b_i are vectors, A_i are matrices, b_i is a vector in $\Re^{n_i}$, and

$$C_i: \;\text{the second order cone of } \Re^{n_i}$$

- The cone here is

$$C = C_1 \times \cdots \times C_m$$

and the constraints $A_i x - b_i \in C_i, \; i = 1, \ldots, m$, can be lumped into a single constraint

$$Ax - b \in C$$

SECOND ORDER CONE DUALITY

- Using the generic duality form

$$\min_{Ax-b\in C} c'x \qquad \Longleftrightarrow \qquad \max_{A'\lambda=c,\ \lambda\in\hat{C}} b'\lambda,$$

and self duality of C, the dual problem is

$$\begin{aligned} &\text{maximize} \quad \sum_{i=1}^{m} b_i'\lambda_i \\ &\text{subject to} \quad \sum_{i=1}^{m} A_i'\lambda_i = c, \quad \lambda_i \in C_i,\ i=1,\ldots,m, \end{aligned}$$

where $\lambda = (\lambda_1,\ldots,\lambda_m)$.

- The duality theory is no more favorable than the one for linear-conic problems.

- There is no duality gap if there exists a feasible solution in the interior of the 2nd order cones C_i.

- Generally, 2nd order cone problems can be recognized from the presence of norm or convex quadratic functions in the cost or the constraint functions.

- There are many applications.

EXAMPLE: ROBUST LINEAR PROGRAMMING

$$\begin{aligned} &\text{minimize} \;\; c'x \\ &\text{subject to} \;\; a_j'x \le b_j, \quad \forall\, (a_j, b_j) \in T_j, \quad j = 1, \ldots, r, \end{aligned}$$

where $c \in \Re^n$, and T_j is a given subset of $\Re^{n+1}$.

- We convert the problem to the equivalent form

$$\begin{aligned} &\text{minimize} \;\; c'x \\ &\text{subject to} \;\; g_j(x) \le 0, \qquad j = 1, \ldots, r, \end{aligned}$$

where $g_j(x) = \sup_{(a_j, b_j) \in T_j} \{a_j'x - b_j\}$.

- For the special choice where T_j is an ellipsoid,

$$T_j = \left\{ (\overline{a}_j + P_j u_j, \overline{b}_j + q_j' u_j) \mid \|u_j\| \le 1,\ u_j \in \Re^{n_j} \right\}$$

we can express $g_j(x) \le 0$ in terms of a SOC:

$$\begin{aligned} g_j(x) &= \sup_{\|u_j\| \le 1} \left\{ (\overline{a}_j + P_j u_j)'x - (\overline{b}_j + q_j' u_j) \right\} \\ &= \sup_{\|u_j\| \le 1} (P_j'x - q_j)'u_j + \overline{a}_j'x - \overline{b}_j, \\ &= \|P_j'x - q_j\| + \overline{a}_j'x - \overline{b}_j. \end{aligned}$$

Thus, $g_j(x) \le 0$ iff $(P_j'x - q_j, \overline{b}_j - \overline{a}_j'x) \in C_j$, where C_j is the SOC of $\Re^{n_j+1}$.