LECTURE 16

LECTURE OUTLINE

- Conic programming
- Semidefinite programming
- Exact penalty functions
- Descent methods for convex/nondifferentiable optimization
- Steepest descent method

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

- Second order cone programming:

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

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 $C = C_1 \times \cdots \times C_m$
- The dual problem is

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

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

SEMIDEFINITE PROGRAMMING

- Consider the symmetric $n \times n$ matrices. Inner product $< X, Y >= \text{trace}(XY) = \sum_{i,j=1}^{n} x_{ij} y_{ij}$.

- Let C be the cone of pos. semidefinite matrices.

- C is self-dual, and its interior is the set of positive definite matrices.

- Fix symmetric matrices D, $A_1, \ldots, A_m$, and vectors $b_1, \ldots, b_m$, and consider

$$\begin{aligned} &\text{minimize} \quad < D, X > \\ &\text{subject to} \quad < A_i, X >= b_i, \quad i = 1, \ldots, m, \quad X \in C \end{aligned}$$

- Viewing this as a linear-conic problem (the first special form), the dual problem (using also self-duality of C) is

$$\begin{aligned} &\text{maximize} \quad \sum_{i=1}^{m} b_i \lambda_i \\ &\text{subject to} \quad D - (\lambda_1 A_1 + \cdots + \lambda_m A_m) \in C \end{aligned}$$

- There is no duality gap if there exists primal feasible solution that is pos. definite, or there exists λ such that $D - (\lambda_1 A_1 + \cdots + \lambda_m A_m)$ is pos. definite.

EXAMPLE: MINIMIZE THE MAXIMUM EIGENVALUE

- Given $n \times n$ symmetric matrix $M(\lambda)$, depending on a parameter vector λ, choose λ to minimize the maximum eigenvalue of $M(\lambda)$.

- We pose this problem as

$$\begin{aligned} &\text{minimize} \quad z \\ &\text{subject to} \;\; \text{maximum eigenvalue of } M(\lambda) \le z, \end{aligned}$$

or equivalently

$$\begin{aligned} &\text{minimize} \quad z \\ &\text{subject to} \;\; zI - M(\lambda) \in C, \end{aligned}$$

where I is the $n \times n$ identity matrix, and C is the semidefinite cone.

- If $M(\lambda)$ is an affine function of λ,

$$M(\lambda) = D + \lambda_1 M_1 + \cdots + \lambda_m M_m,$$

the problem has the form of the dual semidefinite problem, with the optimization variables being $(z, \lambda_1, \ldots, \lambda_m)$.

EXAMPLE: LOWER BOUNDS FOR DISCRETE OPTIMIZATION

- Quadr. problem with quadr. equality constraints

$$\begin{array}{ll} \text{minimize} & x'Q_0x + a_0'x + b_0 \\ \text{subject to} & x'Q_ix + a_i'x + b_i = 0, \quad i = 1, \ldots, m, \end{array}$$

$Q_0, \ldots, Q_m$: symmetric (not necessarily ≥ 0).

- Can be used for discrete optimization. For example an integer constraint $x_i \in \{0, 1\}$ can be expressed by $x_i^2 - x_i = 0$.

- The dual function is

$$q(\lambda) = \inf_{x \in \Re^n} \big\{ x'Q(\lambda)x + a(\lambda)'x + b(\lambda) \big\},$$

where

$$Q(\lambda) = Q_0 + \sum_{i=1}^m \lambda_i Q_i,$$

$$a(\lambda) = a_0 + \sum_{i=1}^m \lambda_i a_i, \quad b(\lambda) = b_0 + \sum_{i=1}^m \lambda_i b_i$$

- It turns out that the dual problem is equivalent to a semidefinite program ...

EXACT PENALTY FUNCTIONS

- We use Fenchel duality to derive an equivalence between a constrained convex optimization problem, and a penalized problem that is less constrained or is entirely unconstrained.

- We consider the problem

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \;\; x \in X, \qquad g(x) \le 0, \end{aligned}$$

where $g(x) = \big(g_1(x), \ldots, g_r(x)\big)$, X is a convex subset of $\Re^n$, and $f : \Re^n \to \Re$ and $g_j : \Re^n \to \Re$ are real-valued convex functions.

- We introduce a convex function $P : \Re^r \mapsto \Re$, called *penalty function*, which satisfies

$$P(u) = 0, \; \forall\, u \le 0, \quad P(u) > 0, \; \text{if } u_i > 0 \text{ for some } i$$

- We consider solving, in place of the original, the "penalized" problem

$$\begin{aligned} &\text{minimize} \quad f(x) + P\big(g(x)\big) \\ &\text{subject to} \;\; x \in X, \end{aligned}$$

FENCHEL DUALITY

- We have

$$\inf_{x\in X}\{f(x)+P\big(g(x)\big)\} = \inf_{u\in\Re^r}\{p(u)+P(u)\}$$

where $p(u) = \inf_{x\in X,\, g(x)\le u} f(x)$ is the primal function.

- Assume $-\infty < q^*$ and $f^* < \infty$ so that p is proper (in addition to being convex).

- By Fenchel duality

$$\inf_{u\in\Re^r}\{p(u)+P(u)\} = \sup_{\mu\ge 0}\{q(\mu)-Q(\mu)\},$$

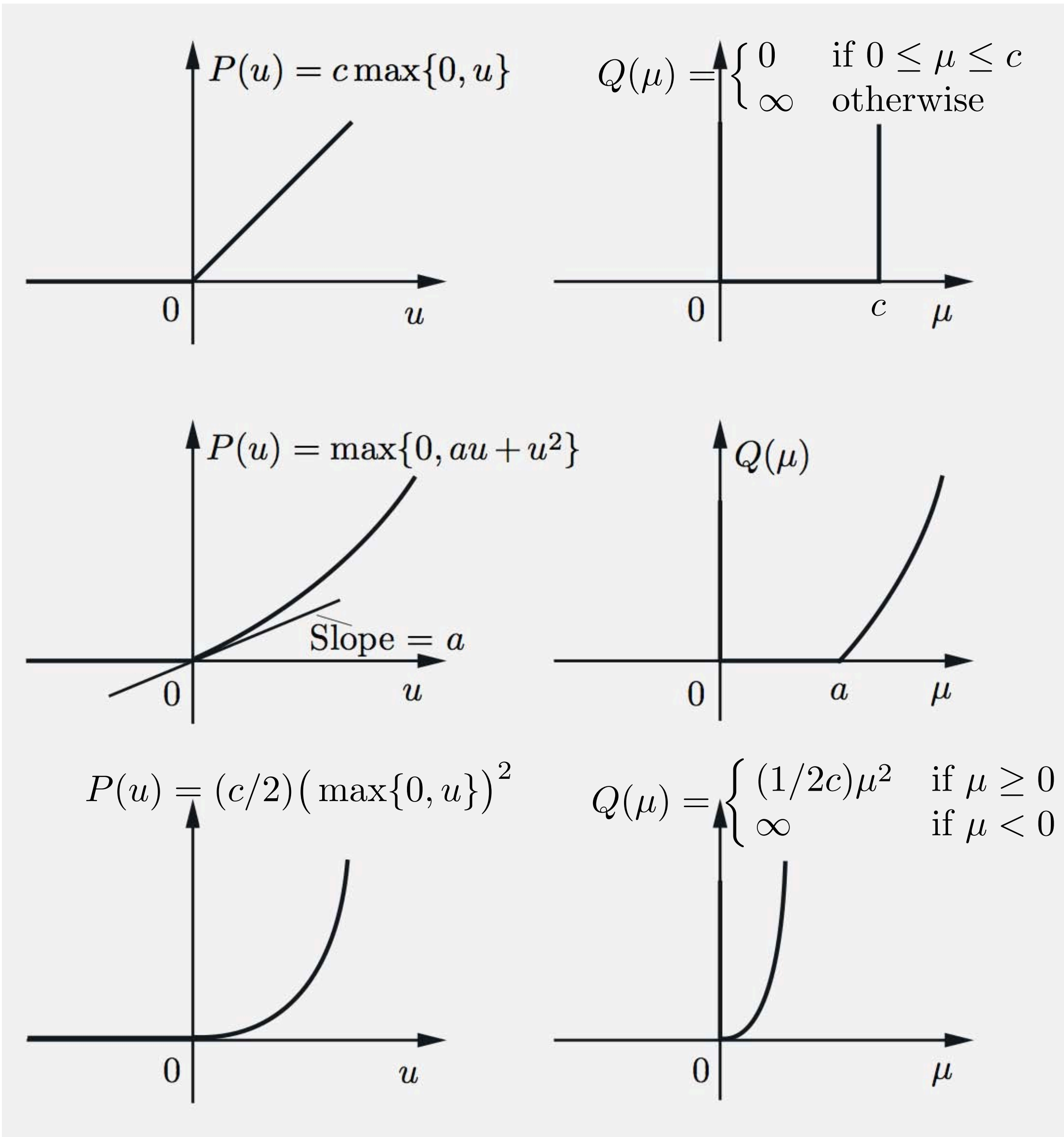
where for $\mu \ge 0$,

$$q(\mu) = \inf_{x\in X}\{f(x)+\mu' g(x)\}$$

is the dual function, and Q is the conjugate convex function of P:

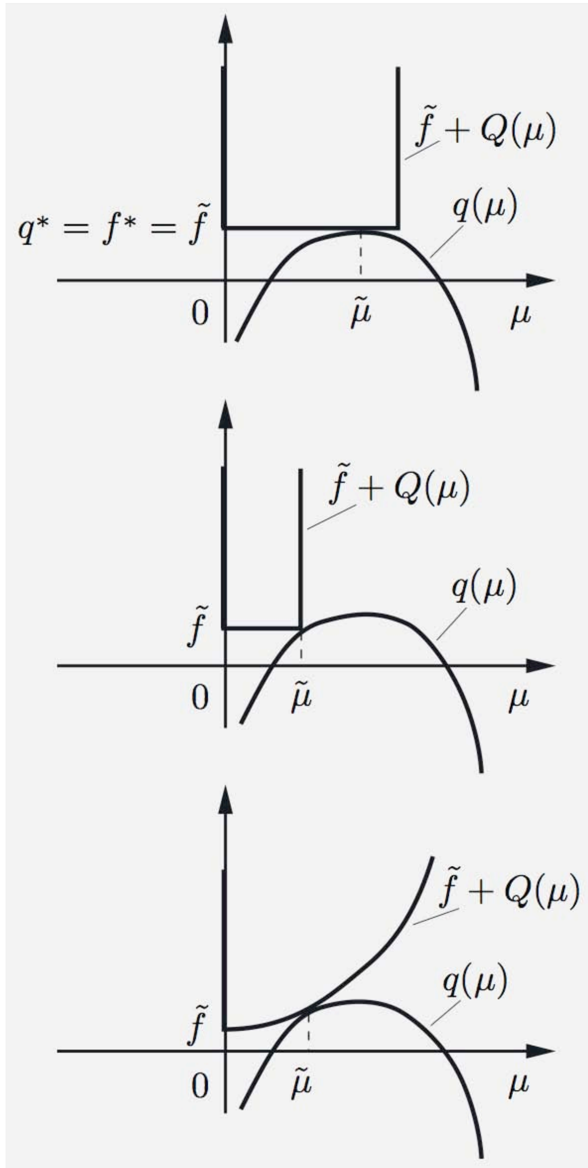
$$Q(\mu) = \sup_{u\in\Re^r}\{u'\mu - P(u)\}$$

PENALTY CONJUGATES



- **Important observation:** For Q to be flat for some $\mu > 0$, P must be nondifferentiable at 0.

FENCHEL DUALITY VIEW



- For the penalized and the original problem to have equal optimal values, Q must be "flat enough" so that some optimal dual solution μ^* minimizes Q, i.e., $0 \in \partial Q(\mu^*)$ or equivalently

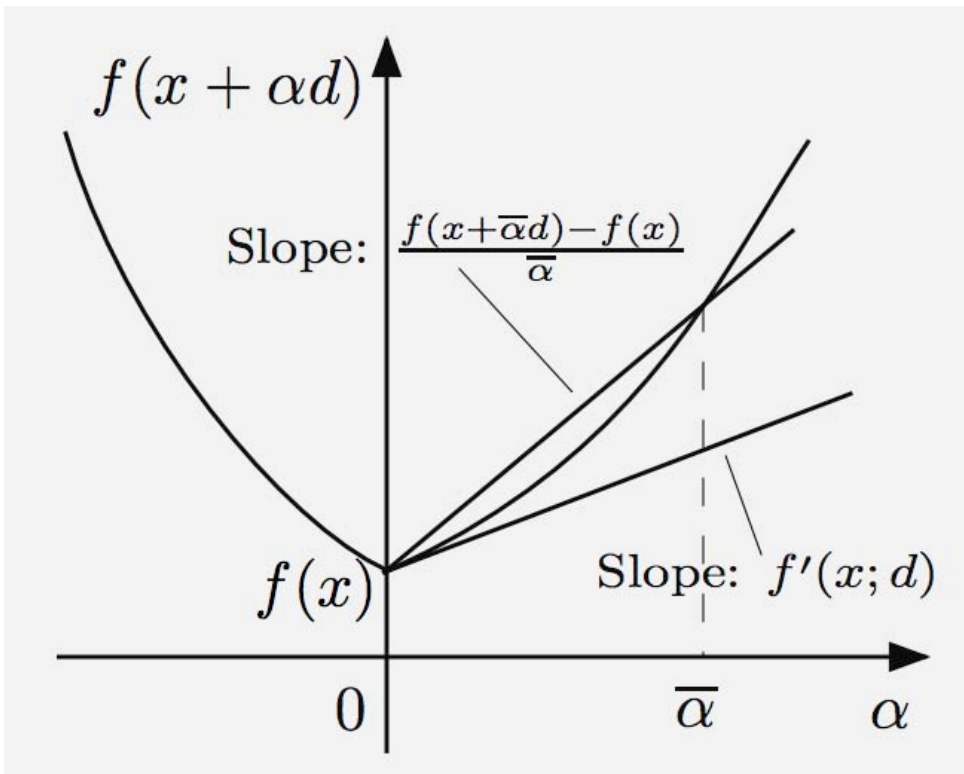
$$\mu^* \in \partial P(0)$$

- True if $P(u) = c \sum_{j=1}^{r} \max\{0, u_j\}$ with $c \geq \|\mu^*\|$ for some optimal dual solution μ^*.

DIRECTIONAL DERIVATIVES

- Directional derivative of a proper convex f:

$$f'(x;d) = \lim_{\alpha \downarrow 0} \frac{f(x+\alpha d) - f(x)}{\alpha}, \quad x \in \text{dom}(f),\ d \in \Re^n$$



- The ratio

$$\frac{f(x+\alpha d) - f(x)}{\alpha}$$

is monotonically nonincreasing as $\alpha \downarrow 0$ and converges to $f'(x;d)$.

- For all $x \in \text{ri}\big(\text{dom}(f)\big)$, $f'(x;\cdot)$ is the support function of $\partial f(x)$.

STEEPEST DESCENT DIRECTION

- Consider unconstrained minimization of convex $f : \Re^n \mapsto \Re$.
- A descent direction d at x is one for which $f'(x; d) < 0$, where

$$f'(x; d) = \lim_{\alpha \downarrow 0} \frac{f(x + \alpha d) - f(x)}{\alpha} = \sup_{g \in \partial f(x)} d'g$$

is the directional derivative.

- Can decrease f by moving from x along descent direction d by small stepsize α.
- Direction of steepest descent solves the problem

$$\begin{aligned} &\text{minimize} \quad f'(x; d) \\ &\text{subject to} \quad \|d\| \leq 1 \end{aligned}$$

- **Interesting fact:** The steepest descent direction is $-g^*$, where g^* is the vector of minimum norm in $\partial f(x)$:

$$\begin{aligned} \min_{\|d\| \leq 1} f'(x; d) &= \min_{\|d\| \leq 1} \max_{g \in \partial f(x)} d'g = \max_{g \in \partial f(x)} \min_{\|d\| \leq 1} d'g \\ &= \max_{g \in \partial f(x)} \big(-\|g\|\big) = -\min_{g \in \partial f(x)} \|g\| \end{aligned}$$

STEEPEST DESCENT METHOD

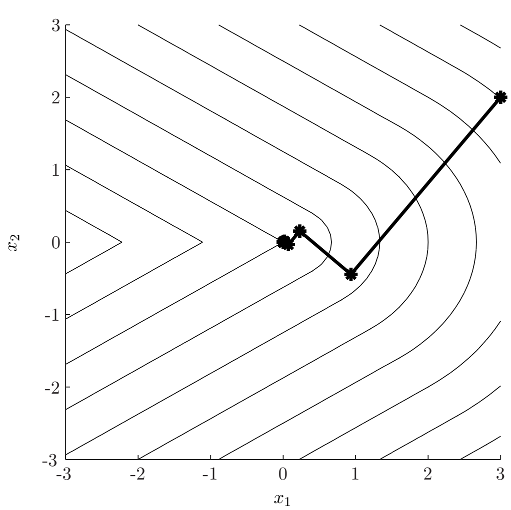
- Start with any $x_0 \in \Re^n$.

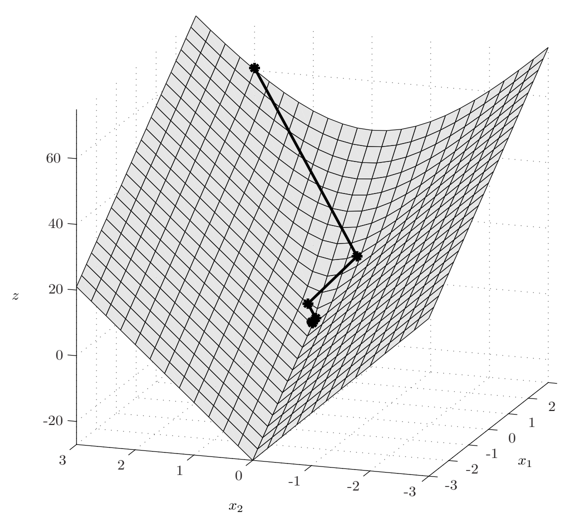
- For $k \geq 0$, calculate $-g_k$, the steepest descent direction at x_k and set

$$x_{k+1} = x_k - \alpha_k g_k$$

- **Difficulties:**
  - Need the entire $\partial f(x_k)$ to compute g_k.
  - Serious convergence issues due to discontinuity of $\partial f(x)$ (the method has no clue that $\partial f(x)$ may change drastically nearby).

- Example with α_k determined by minimization along $-g_k$: $\{x_k\}$ converges to nonoptimal point.





6.253 Convex Analysis and Optimization
Spring 2010