Math 591 – Real Algebraic Geometry and Convex Optimization

Lecture 5: Convex optimization basics
Cynthia Vinzant, Spring 2019

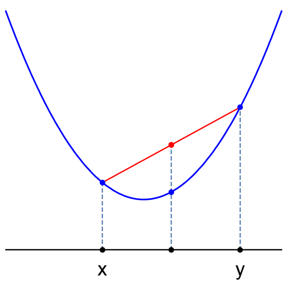
Definition. Let $C \subseteq V$ be convex. A function $f : C \to \mathbb{R}$ is **convex** if for all $x, y \in C$ and $\lambda \in [0, 1]$,

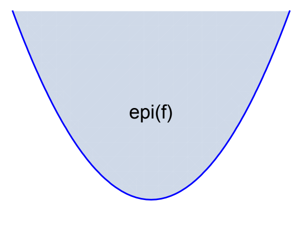
$$f(\lambda x + (1-\lambda)y) \;\leq\; \lambda f(x) + (1-\lambda)f(y).$$

An equivalent condition is that f is convex if its **epigraph**,

$$\text{epi}(f) \;=\; \{(x,t) \in C \times \mathbb{R} : f(x) \leq t\}$$

is a convex subset of $V \times \mathbb{R}$.





Some examples include

- linear functions
- $f(x) = \max\{f_1(x), \ldots, f_k(x)\}$ where $f_1, \ldots, f_k$ are convex
- for $C = \mathbb{R}^{n\times n}_{\text{sym}}$,

$$f(A) \;=\; \text{maximum eigenvalue of } A \;=\; \max_{||v||_2=1} v^T A v$$

  (a maximum of infinitely many linear functions $A \mapsto v^T A v$).

A *convex optimization problem* is one of the form

$$\min_{x\in C} f(x),$$

where C is convex and $f : C \to \mathbb{R}$ is a convex function. Note that the set of points achieving this minimum is itself convex, as its the intersection of epi(f) with $\{(x,t) : t = \text{min. value}\}$, both of which are convex.

Also, for any convex optimization problem, we can write an equivalent problem that consists of minimizing a *linear* function over a convex set. Namely:

$$\min_{x\in C} f(x) \;=\; \min_{(x,t)\in\text{epi}(f)} t.$$

It will be useful to translate even further, by viewing the convex set as an affine-linear section of a convex cone.

Conic Programming. A **conic optimization problem** has the form

$$\min_{x \in K} c(x) \quad \text{such that} \quad a_i(x) = b_i \quad \text{for} \quad i = 1, \ldots, m,$$

where $K \subset V$ is a convex cone, $c, a_1, \ldots, a_m \in V^*$ and $b_1, \ldots, b_m \in \mathbb{R}$.

Example. For $K = (\mathbb{R}_{\geq 0})^n$, the feasible sets $\{x \in (\mathbb{R}_{\geq 0})^n : a_i^T x = b_i \text{ for } i = 1, \ldots, m\}$ are exactly **polyhedra**, and methods for solving these problems are called **linear programs**.

Example. For $K = \mathrm{PSD}_n$, the feasible sets $\{X \in \mathrm{PSD}_n : \langle A_i, X \rangle = b_i \text{ for } i = 1, \ldots, m\}$ are called **spectrahedra**, and methods for solving these problems are called **semidefinite programs**.

For example, given a matrix $A \in \mathbb{R}^{n \times n}_{\text{sym}}$, we can find the maximum eigenvalue of A as a semidefinite program. Namely,

$$\text{max eigval. of } A \quad = \quad \min_{t \in \mathbb{R}} t \ \text{ s.t. } \ t \geq \text{all eigval. of } A \quad = \quad \min_{t \in \mathbb{R}} t \ \text{ s.t. } \ tI - A \in \mathrm{PSD}_n.$$

The set $\{tI - A : t \in \mathbb{R}\}$ is an affine line in the space of real symmetric matrices and we are interested in minimizing the linear function t over its intersection with PSD_n.

Example. For $K = P_{1,\leq 2d} = \{f \in \mathbb{R}[x]_{\leq 2d} : f(p) \geq 0 \text{ for all } p \in \mathbb{R}\}$, conic programming with this cone results in polynomial optimization. Suppose that we want to find the maximum value of some polynomial $f \in \mathbb{R}[x]_{\leq 2d}$ (should that maximum exist). We can rewrite this as

$$\begin{aligned} \max_{x \in \mathbb{R}} f(x) &= \min_{t \in \mathbb{R}} t \quad \text{s.t.} \quad t - f(x) \in K \\ &= f(0) + \min_{g \in K} g(0) \quad \text{s.t.} \quad \mathrm{coeff}(g, x^k) = -\mathrm{coeff}(f, x^k) \text{ for } k = 1, \ldots, 2d, \end{aligned}$$

which has the form of a conic optimization problem over K.

Reformulation into a conic optimization problems is useful because it lets us understand a large family of problems using a single convex cone. This is most useful when the cone is one that we can work with (for example, easily check membership in). Important examples of cones that we can easily work with are $\mathbb{R}^n_{\geq 0}$ and PSD_n.

It is also useful because we can easily formulate a "dual problem".

Duality in conic optimization. We call out original optimization problem the "primal" problem and formulate a "dual" optimization problem to provide a lower bound.

(Primal) $$\min_{x \in K} c(x) \qquad \text{s.t.} \qquad a_i(x) = b_i \ \text{ for } \ i = 1, \ldots, m$$

(Dual) $$\max_{y \in \mathbb{R}^m} \sum_{i=1}^{m} b_i y_i \qquad \text{s.t.} \qquad c - \sum_{i=1}^{m} a_i y_i \ \in \ K^*$$

Note that $\{c - \sum_{i=1}^m a_i y_i : y \in \mathbb{R}^m\}$ parametrizes an affine linear space in V^*. So the dual problem still has form of optimizing a linear function over the intersection of a convex cone and an affine space.

We say that a point $x \in V$ is *primal feasible* if $x \in K$ and $a_i(x) = b_i$ for $i = 1, \ldots, m$. Similarly we say $y \in \mathbb{R}^m$ is *dual feasible* if $c - \sum_{i=1}^m a_i y_i \in K^*$.

Theorem (Weak duality)**.** *For any primal feasible* x *and dual feasible* y*,*

$$c(x) \;\geq\; \sum_{i=1}^{m} b_i y_i.$$

Moreover, if $c(\hat{x}) = \sum_{i=1}^{m} b_i \hat{y}_i$*, then both* $\hat{x}$ *and* $\hat{y}$ *are optimal.*

Proof. Suppose x, y are feasible. Then

$$\begin{aligned} c(x) - \sum_{i=1}^{m} b_i y_i \;&=\; c(x) - \sum_{i=1}^{m} a_i(x) y_i \\ &=\; (c - \sum_{i=1}^{m} y_i a_i)(x) \;\geq\; 0. \end{aligned}$$

Here the last inequality follows from the fact that $c - \sum_{i=1}^{m} y_i a_i \in K^*$ and $x \in K$.

If $c(\hat{x}) = \sum_{i=1}^{m} b_i \hat{y}_i$, then $c(\hat{x})$ is an upper bound for $\sum_i b_i y_i$ over all feasible y. Since $\hat{y}$ achieves this upper bound, it must be optimal. Similarly, $\sum_i b_i \hat{y}_i$ is a lower bound for $c(x)$, which is achieved by $\hat{x}$. □

This gives a method for *certifying* the optimal value!

Example. Take $K = \mathbb{R}^3_{\geq 0}$ and $m = 1$. We can choose $c = (7, 2, 4)$, $a = (1, 1, 1)$, $b = 1$. Then the primal optimization problem is

$$\min_{x \in \mathbb{R}^3_{\geq 0}} \; 7x_1 + 2x_2 + 4x_3 \qquad \text{s.t.} \qquad x_1 + x_2 + x_3 = 1$$

It's not hard to see that the minimum value is 2, which is achieved by $x = (0, 1, 0)$. We can certify this by writing the dual problem:

$$\max_{y \in \mathbb{R}} y \qquad \text{s.t.} \qquad (7, 2, 4) - y(1, 1, 1) \;\in\; \mathbb{R}^3_{\geq 0}.$$

Here we see the maximum of the dual problem is also 2, achieved by $y = 2$. Also, as in the proof of the theorem of weak duality, $(c - ya)(x) = 0$ for x, y achieving the optimal value:

$$(c - ya)(x) \;=\; \langle (7, 2, 4) - 2(1, 1, 1)), (0, 1, 0) \rangle \;=\; \langle (5, 0, 2), (0, 1, 0) \rangle = 0.$$

Example. Take $K = \mathrm{PSD}_3$ and $m = 3$.

$$C = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad A_i = e_i e_i^T, \quad \text{and} \quad b_i = 1 \quad \text{for} \quad i = 1, 2, 3.$$

Then the primal optimization problem is

$$\begin{aligned} &\min_{X \in \mathrm{PSD}_3} \langle C, X \rangle && \text{s.t.} \qquad \langle A_i, X \rangle = b_i \quad \text{for} \quad i = 1, 2, 3 \\ = &\min_{X \in \mathrm{PSD}_3} 2(X_{12} + X_{13} + X_{23}) && \text{s.t.} \qquad X_{ii} = 1 \quad \text{for} \quad i = 1, 2, 3, \end{aligned}$$

where X_{ij} is the (i, j)th entry of X.

We might make a lucky guess and note that the matrix

$$X \;=\; \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix}$$

is positive semidefinite and gives the values $\langle C, X\rangle = -3$.

The dual problem is

$$\begin{aligned} & \max_{y\in\mathbb{R}^3} \sum_i b_i y_i && \text{s.t.} \quad C - \sum_i y_i A_i \ \in \ \mathrm{PSD}_3 \\ = \ & \max_{y\in\mathbb{R}^3} \ y_1 + y_2 + y_3 && \text{s.t.} \quad \begin{pmatrix} -y_1 & 1 & 1 \\ 1 & -y_2 & 1 \\ 1 & 1 & -y_3 \end{pmatrix} \in \ \mathrm{PSD}_3. \end{aligned}$$

Here we observe that for $y = (-1, -1, -1)$, the matrix $C - \sum_i y_i A_i$ is positive semidefinite and $y_1 + y_2 + y_3 = -3$. This shows that both the matrix X above and $y = (-1, -1, -1)$ are optimal points.

As in the proof of weak duality, for these optimal X, y,

$$\langle C - \sum_i y_i A_i, X\rangle \ = \ \left\langle \begin{pmatrix} 1 & -1/2 & -1/2 \\ -1/2 & 1 & -1/2 \\ -1/2 & -1/2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \right\rangle \ = \ 0.$$