Linear and non-linear programming

Benjamin Recht

March 11, 2005

The Gameplan

- Constrained Optimization
- Convexity
- Duality
- Applications/Taxonomy

Constrained Optimization

$$
\begin{array}{lll}
\text{minimize} & f(\mathbf{x}) & \\
\text{subject to} & g_j(\mathbf{x}) \leq 0 & j = 1, \ldots, J \\
 & h_k(\mathbf{x}) = 0 & k = 1, \ldots, K \\
 & \mathbf{x} \in \Omega \subset \mathbb{R}^n &
\end{array}
$$

Exercise: formulate the halting problem as an optimization

Equivalence of Feasibility and Optimization

From a complexity point of view

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & g_j(\mathbf{x}) \leq 0 \\ & h_k(\mathbf{x}) = 0 \\ & \mathbf{x} \in \Omega \subset \mathbb{R}^n \end{array} \iff \begin{array}{ll} \text{find} & \mathbf{x} \text{ and } t \\ \text{subject to} & f(\mathbf{x}) - t \geq 0 \\ & g_j(\mathbf{x}) \leq 0 \\ & h_k(\mathbf{x}) = 0 \\ & \mathbf{x} \in \Omega \subset \mathbb{R}^n \end{array}$$

If you solve the RHS, you get a solution for the LHS. If you do bisection on t for the LHS, you solve the RHS.

Convexity: Overview

- Phrasing a problem as an optimization generally buys you nothing

- However, solving a Convex Program is generically no harder than least squares.

- The hard part is *formulating the problem*.

Convex Sets

- If $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \Omega$, a *convex combination* is a linear combination $\sum_{i=1}^{N} p_i \mathbf{x}_i$ where $p_i > 0$ and $\sum_{i=1}^{N} p_i = 1$

- The line segment between $\mathbf{x}$ and $\mathbf{y}$ is given by $(1-t)\mathbf{x} + t\mathbf{y}$. This is a convex combination of two points.

- A set $\Omega \subset \mathbb{R}^n$ is *convex* if it contains all line segments between all points. That is, $\mathbf{x}, \mathbf{y} \in \Omega$ implies $(1-t)\mathbf{x} + t\mathbf{y} \in \Omega$ for all t.

Examples of Convex Sets

- $\mathbb{R}^n$ is convex. Any vector space is convex.

- Any line segment is convex.

- Any line is convex.

- The set of psd matrices is convex. $\mathbf{Q} \succeq 0$ and $\mathbf{P} \succeq 0$ implies $t\mathbf{Q} + (1-t)\mathbf{P} \succeq 0$.

Examples of Non-convex Sets

- The integers are not convex.
- The set of bit strings of length n is not convex.
- The set of vectors with norm 1 is not convex.
- The set of singular matrices is not convex. The set of invertible matrices is not convex.

Operations that preserve convexity

- If $\Omega_1, \ldots, \Omega_m$ are convex, then $\bigcap_{i=1}^m \Omega_i$ is convex.

- If Ω_1 is convex. Then $\Omega_2 = \{\mathbf{A}\mathbf{x} + \mathbf{b} | \mathbf{x} \in \Omega_1\}$ is convex.

- If Ω_1 is convex. Then $\Omega_2 = \{\mathbf{x} | \mathbf{A}\mathbf{x} + \mathbf{b} \in \Omega_1\}$ is convex.

Convex Functions

- A function $f : \Omega \rightarrow \mathbb{R}^n$ is convex if the set

$$\text{epi}(f) = \{(\mathbf{x}, f(\mathbf{x}))|\mathbf{x} \in \Omega\}$$

  is convex

- For functions $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, f is convex iff for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$

$$f((1-t)\mathbf{x} + t\mathbf{y}) \leq (1-t)f(\mathbf{x}) + tf(\mathbf{y})$$

Checking Convexity with Derivatives

$f : \mathbb{R}^n \rightarrow \mathbb{R}$

- If f is differentiable f is convex iff $f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x})$ for all $\mathbf{y}$

- If f is twice differentiable f is convex iff $\nabla^2 f$ is positive semi-definite.

- These facts will be useful next week when we discuss optimization algorithms

Operations that preserve convexity

- If $f(\mathbf{x})$ is convex then $f(\mathbf{Ax} + \mathbf{b})$ is convex.

- If $f_1, \ldots, f_n$ are convex, then so is $a_1 f_1 + \cdots + a_n f_n$ for any scalars a_i.

- If $f_1, \ldots, f_n$ are convex, then $\max_i f_i(\mathbf{x})$ is convex.

- If for all $\mathbf{y}$, $f(\mathbf{x}, \mathbf{y})$ is convex in x, then $\sup_{\mathbf{y}} f(\mathbf{x}, \mathbf{y})$ is convex

Examples of Convex Functions

- Any affine function $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ is convex.
- $-\log(x)$ is convex. $\exp(x)$ is convex.
- $\|\mathbf{x}\|^2$ is convex.
- A quadratic form $\mathbf{x}^\top \mathbf{Q}\mathbf{x}$ with $\mathbf{Q} = \mathbf{Q}^\top$ is convex if and only if $\mathbf{Q} \succeq 0$.

Quadratic Forms

A quadratic form $\mathbf{Q} = \mathbf{Q}^\top$ is convex if and only if $\mathbf{Q} \succeq 0$.

Proof $\quad \mathbf{Q} \succeq 0$ implies $\mathbf{Q} = \mathbf{A}^\top \mathbf{A}$ for some $\mathbf{A}$. Then

$$\mathbf{x}^\top \mathbf{Q} \mathbf{x} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{A} \mathbf{x} = \|\mathbf{A}\mathbf{x}\|^2$$

Conversely, if $\mathbf{Q}$ is not psd, let $\mathbf{v}$ be a norm 1 eigenvector corresponding to eigenvalue $\lambda < 0$. Then

$$0 = (-\mathbf{v} + \mathbf{v})^\top \mathbf{Q} (-\mathbf{v} + \mathbf{v}) > (-\mathbf{v})^\top \mathbf{Q} (-\mathbf{v}) + (\mathbf{v})^\top \mathbf{Q} (\mathbf{v}) = 2\lambda$$

■

Examples of Non-Convex Functions

- $\sin(x)$, $\cos(x)$, and $\tan(x)$ are not convex.

- x^3 is not convex

- Gaussians $p(\mathbf{x}) = \exp(-\mathbf{x}^\top \Lambda^{-1} \mathbf{x}/2)$ are not convex. However, $-\log(p)$ is convex!

Examples of Convex Constraint Sets

$\Omega_1 = \{\mathbf{x}|g(\mathbf{x}) \leq 0\}$ is convex if g is convex.

Proof Let $\mathbf{x}, \mathbf{y} \in \Omega_1$, $0 \leq t \leq 1$. If f is convex,

$$f(t\mathbf{x} + (1-t)\mathbf{y}) \leq tf(\mathbf{x}) + (1-t)\mathbf{y} \leq 0$$

proving $t\mathbf{x} + (1-t)\mathbf{y} \in \Omega$. ∎

$\Omega_2 = \{\mathbf{x}|h(\mathbf{x}) = 0\}$ is convex if $h(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$.

Proof Let $\mathbf{x}, \mathbf{y} \in \Omega_2$, $0 \leq t \leq 1$. If h is affine,

$$\begin{aligned} h(t\mathbf{x} + (1-t)\mathbf{y}) &= \mathbf{A}(t\mathbf{x} + (1-t)\mathbf{y}) + \mathbf{b} \\ &= t(\mathbf{A}\mathbf{x} + \mathbf{b}) + (1-t)(\mathbf{A}\mathbf{y} + \mathbf{b}) = 0 \end{aligned}$$

proving $t\mathbf{x} + (1-t)\mathbf{y} \in \Omega$. ∎

The Hahn-Banach Theorem

- A hyperplane is a set of the form $\{\mathbf{a}^\top \mathbf{x} = b\} \subset \mathbb{R}^n$. A half-space is a set of the form $\{\mathbf{a}^\top \mathbf{x} \leq b\} \subset \mathbb{R}^n$

- **Theorem** If Ω is convex and $x \notin \text{cl}(\Omega)$ then there exists a hyperplane separating x and Ω.

- It follows that Ω is the intersection of all half-spaces which contain it.

Duality

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & g_j(\mathbf{x}) \leq 0 \\ & \mathbf{x} \in \Omega \end{array}$$

The *Lagrangian* for this problem is given by

$$\mathcal{L}(\mathbf{x}, \mu) = f(\mathbf{x}) + \sum_{j=1}^{J} \mu_j g_j(\mathbf{x})$$

with $\mu \geq 0$. The μ_j and are called *Lagrange multipliers*. In calculus, we searched for values of μ by using $\nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mu) = 0$. Here, note that solving the optimization is equivalent to solving

$$\min_{\mathbf{x}} \max_{\mu \geq 0} \mathcal{L}(\mathbf{x}, \mu)$$

Duality (2)

$$\min_{\mathbf{x}} \max_{\mu \geq 0} \mathcal{L}(\mathbf{x}, \mu) \geq \max_{\mu \geq 0} \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mu)$$

The right hand side is called the *Dual Program*

Proof Let $f(\mathbf{x}, \mathbf{y})$ be *any* function with two arguments. Then $f(\mathbf{x}, \mathbf{y}) \geq \min_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})$. Taking the max w.r.t. $\mathbf{y}$ of both sides shows $\max_{\mathbf{y}} f(\mathbf{x}, \mathbf{y}) \geq \max_{\mathbf{y}} \min_{\mathbf{x}} f(\mathbf{x}, \mathbf{y})$. Now take the min of the right hand side w.r.t. $\mathbf{x}$ to prove the theorem. ∎

Duality (3)

The dual program is always concave. To see this, consider the *dual function*

$$q(\mu) \equiv \min_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \mu) = \min_{\mathbf{x}} f(\mathbf{x}) + \sum_{j=1}^{J} \mu_j g_j(\mathbf{x})$$

Now, since $\min_{\mathbf{x}}(f(\mathbf{x}) + g(\mathbf{x})) \leq (\min_{\mathbf{x}} f(\mathbf{x})) + (\min_{\mathbf{x}} g(\mathbf{x}))$, we have

$$\begin{aligned} q(t\mu_1 + (1-t)\mu_2) &= \min_{\mathbf{x}} t \left(f(\mathbf{x}) + \sum_{j=1}^{J} \mu_{1j} g_j(\mathbf{x}) \right) \\ &\quad + (1-t) \left(f(\mathbf{x}) + \sum_{j=1}^{J} \mu_{2j} g_j(\mathbf{x}) \right) \\ &\geq t q(\mu_1) + (1-t) q(\mu_2) \end{aligned}$$

Duality (4)

- The dual may be interpreted as searching over half spaces which contain the set $\{(f(x), g(x)) \in \mathbb{R}^{J+1} | x \in \Omega\}$. This is illustrated in the figures.

- When the problem is convex and strictly feasible, the dual of the dual returns the primal.

Duality Gaps

We know that the solution to the primal problem is greater than or equal to the solution of the dual problem. The *duality gap* is defined to be

$$\min_{\mathbf{x}\in\Omega} \max_{\mu\geq 0} \mathcal{L}(\mathbf{x},\mu) - \max_{\mu\geq 0} \min_{\mathbf{x}\in\Omega} \mathcal{L}(\mathbf{x},\mu)$$

- When f and g_j are convex functions, Ω is a convex set, and there is a point strictly inside Ω with $g_j(x) < 0$ for all j then the duality gap is zero.

- Otherwise, estimating the duality gap is quite hard. In many cases, this gap is infinite. Later classes will examine how to analyze when the gap is small.

Linear Programming

$$\begin{array}{ll} \text{minimize} & \mathbf{c}^\top \mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$$

Sometimes you will have equality constraints as well. Sometimes you won't have $\mathbf{x} \geq 0$.

Equivalence of Representations

To turn unsigned variables into nonnegative variables:

$$\mathbf{x} = \mathbf{x}_+ - \mathbf{x}_- \quad \mathbf{x}_\pm \geq 0$$

To turn equality constraints into inequalities:

$$\mathbf{Ax} = \mathbf{b} \iff \mathbf{Ax} \leq \mathbf{b} \quad \text{and} \quad \mathbf{Ax} \geq \mathbf{b}$$

To turn inequalities into equalities

$$\mathbf{Ax} \leq \mathbf{b} \iff \mathbf{Ax} + \mathbf{s} = \mathbf{b} \quad \text{and} \quad \mathbf{s} \geq 0$$

Such $\mathbf{s}$ are called *slack variables*

Linear Programming Duality

Set up the Lagrangian

$$\begin{aligned}\mathcal{L}(\mathbf{x}, \mu) &= \mathbf{c}^\top \mathbf{x} + \mu^\top (\mathbf{b} - \mathbf{A}\mathbf{x}) \\ &= (\mathbf{c} - \mathbf{A}^\top \mu)^\top \mathbf{x} + \mathbf{b}^\top \mu\end{aligned}$$

Minimize with respect to $\mathbf{x}$

$$\inf_{\mathbf{x} \geq 0} \mathcal{L}(\mathbf{x}, \mu) = \begin{cases} \mu^\top \mathbf{b} & \mathbf{c}^\top - \mu^\top \mathbf{A} \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Linear Programming Duality

The dual program

$$\begin{array}{ll} \text{maximize} & b^\top \mu \\ \text{subject to} & A^\top \mu \le c \\ & \mu \ge 0 \end{array}$$

The dual of a linear program is a linear program. It has the same number of variables as the primal has constraints. It has the same number of constraints as the primal has variables.

Basic Feasible Solutions

Consider the LP

$$\begin{array}{ll} \text{min} & c^\top \mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$$

where $\mathbf{A}$ is $m \times n$ and has m linearly independent columns.

Let $\mathbf{B}$ be an $m \times m$ matrix formed by picking m linearly independent columns from $\mathbf{A}$ basic solution of the LP is given by

$$x_j = \begin{cases} [\mathbf{B}^{-1}\mathbf{b}]_k & j \text{ is the kth column of } \mathbf{B} \\ 0 & \mathbf{a}_j \notin \mathbf{B} \end{cases}$$

If $\mathbf{x}$ is feasible, it is called a basic feasible solution (BFS).

The Simplex Algorithm

FACT: If an optimal solution to an LP exists, then an optimal BFS exists.

Simplex Algorithm (sketch):

- Find a BFS
- Find a column which improves the cost or break
- Swap this column in and find a new BFS
- Goto step 2

Chebyshev approximation

$$\min_{\mathbf{x}} \max_{i=1,\dots,N} |\mathbf{a}_i^\top \mathbf{x} - b_i|$$

Is equivalent to the LP

$$\begin{aligned} \text{min} \quad & t \\ \text{s.t.} \quad & \mathbf{a}_i^\top \mathbf{x} - b_i \leq t \qquad i = 1, \dots, N \\ & -\mathbf{a}_i^\top \mathbf{x} + b_i \leq t \quad i = 1, \dots, N \end{aligned}$$

L_1 approximation

$$\min_{\mathbf{x}} \sum_{i=1}^{N} |\mathbf{a}_i^\top \mathbf{x} - b_i|$$

Is equivalent to the LP

$$\begin{array}{lll} \text{min} & \sum_{i=1}^{N} t_i & \\ \text{s.t.} & \mathbf{a}_i^\top \mathbf{x} - b_i \le t_i & i = 1, \dots, N \\ & -\mathbf{a}_i^\top \mathbf{x} + b_i \le t_i & i = 1, \dots, N \end{array}$$

Probability

The set of probability distributions forms a convex set. For example, the set of probabilities for N events is

$$\sum_{i=1}^{N} p_i = 1 \quad p_i \geq 0$$

The entropy is a concave function of a probability distribution

$$H[p] \equiv -\sum_{i=1}^{N} p_i \log p_i$$

Maximum Entropy Distributions

Let f be some random variable. Then the problem

$$\begin{array}{ll} \max_p & H[p] \\ \text{s.t.} & \mathbb{E}_p[f] = \bar{f} \end{array}$$

is a convex program. This is the maximum entropy distribution with the desired expected values.

Using the Lagrangian one can show

$$p_i \propto \exp(\lambda f_i)$$

and the dual is

$$\min_{\lambda} \quad \log \sum_{i=1}^{N} \exp(\lambda f_i) - \lambda \bar{f}$$

Semidefinite Programming

If A and B are symmetric $n \times n$ matrices then

$$\mathrm{Tr}(\mathbf{AB}) = \sum_{i,j=1}^{n} A_{ij} B_{ij}$$

providing an inner product on matrices.

A semidefinite program is a linear program over the positive semidefinite matrices.

$$\begin{array}{ll} \text{minimize} & \mathrm{Tr}(\mathbf{A}_0 \mathbf{Z}) \\ \text{subject to} & \mathrm{Tr}(\mathbf{A}_i \mathbf{Z}) = c_i \quad k = 1, \ldots, K \\ & \mathbf{Z} \succeq 0 \end{array}$$

Semidefinite Programming Duality

Set up the Lagrangian

$$\mathcal{L}(\mathbb{Z}, \mu) = \text{Tr}(\mathbf{A}_0\mathbf{Z}) + \sum_{K=1}^{K} \mu_k(\text{Tr}(\mathbf{A}_k\mathbf{Z}) - c_i)$$

$$= \text{Tr}\left(\left(\mathbf{A}_0 + \sum_{K=1}^{K} \mu_k\mathbf{A}_k\right)\mathbf{Z}\right) - \mathbf{c}^\top \mu$$

Minimize with respect to $\mathbf{Z}$

$$\inf_{\mathbf{Z}\succeq 0} \mathcal{L}(\mathbf{Z}, \mu) = \begin{cases} -\mathbf{c}^\top \mu & \mathbf{A}_0 + \sum_{K=1}^{K} \mu_k\mathbf{A}_k \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Semidefinite Programming Duality

The dual program

$$\begin{array}{ll} \text{min} & \mathbf{c}^\top \mu \\ \text{s.t.} & \mathbf{A}_0 + \sum_{K=1}^{K} \mu_k \mathbf{A}_k \succeq 0 \end{array}$$

We can put this back into the standard form by noting that the constraint set without the positivity condition is an affine set and hence can be written as an intersection of hyperplanes

$$C = \{\mathbf{W} |\, \text{Tr}(\mathbf{W} G_i) = b_i,\, i = 1, \ldots, T\}$$

for some symmetric matrices G_i and scalars b_i. But it is important to recognize both forms as semidefinite programs.

Linear Programming as SDPs

$$\begin{array}{ll} \min & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} \geq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}$$

Let $\mathbf{a}_i$ denote the ith column of $\mathbf{A}$.

is equivalent to the SDP

$$\begin{array}{ll} \min & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \operatorname{diag}(\mathbf{a}_i^\top \mathbf{x} - b_i) \succeq 0 \end{array}$$

Quadratic Programs as SDPs

A quadratically constrained convex quadratic program is the optimization

$$\begin{array}{ll} \text{min} & f_0(\mathbf{x}) \\ \text{s.t.} & f_i(\mathbf{x}) \le 0 \end{array}$$

$f_i(\mathbf{x}) = \mathbf{x}^\top \mathbf{A}_i \mathbf{x} - 2\mathbf{b}_i^\top \mathbf{x} + c_i$ and $\mathbf{A}_i \succeq 0$

Let $\mathbf{Q}_i^\top \mathbf{Q} = \mathbf{A}_i$. This is equivalent to the semidefinite program

$$\begin{array}{ll} \text{min} & t \\ \text{s.t.} & \begin{bmatrix} \mathbb{1} & \mathbf{Q}_0 \mathbf{x} \\ \mathbf{x}^\top \mathbf{Q}_0 & 2\mathbf{b}_0^\top \mathbf{x} - c_0 + t \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} \mathbb{1} & \mathbf{Q}_i \mathbf{x} \\ \mathbf{x}^\top \mathbf{Q}_i & 2\mathbf{b}_i^\top \mathbf{x} - c_i \end{bmatrix} \succeq 0 \end{array}$$

Logarithmic Chebyshev Approximation

$$\min_{\mathbf{x}} \max_{i=1,\dots,N} |\log(\mathbf{a}_i^\top \mathbf{x}) - \log(b_i)|$$

Is equivalent to the SDP

$$\begin{array}{ll} \min & t \\ \text{s.t.} & \begin{bmatrix} t - \mathbf{a}_i^\top \mathbf{x}/b_i & 0 & 0 \\ 0 & \mathbf{a}_i^\top \mathbf{x}/b_i & 1 \\ 0 & 1 & t \end{bmatrix} \succeq 0 \quad i = 1, \dots, N \end{array}$$

Finding the maximum singular value

Let $\mathbf{A}(\mathbf{x}) = \mathbf{A}_0 + \mathbf{A}_1 x_1 + \ldots \mathbf{A}_k x_k$ be an $n \times m$ matrix valued function. Which value of $\mathbf{x}$ attains the matrix with the maximum singular value? Solve with an SDP

$$\begin{array}{ll} \text{min} & t \\ \text{s.t.} & \left[\begin{array}{cc} t\mathbb{1} & \mathbf{A}(\mathbf{x}) \\ \mathbf{A}(\mathbf{x})^\top & t\mathbb{1} \end{array} \right] \succeq 0 \end{array}$$

Problem 1: Examples of Convex Functions (Bertsekas Ex 1.5)

Show that the following are convex on $\mathbb{R}^n$

- $f_1(\mathbf{x}) = -(x_1 x_2 \cdots x_n)^{1/n}$ on $\{\mathbf{x} \in \mathbb{R}^n | x_i > 0\}$
- $f_2(\mathbf{x}) = \log \sum_{i=1}^{N} \exp(x_i)$
- $f_3(\mathbf{x}) = \|\mathbf{x}\|^p$ with $p \geq 1$.
- $f_4(\mathbf{x}) = \frac{1}{f(\mathbf{x})}$ where f is concave and positive for all $\mathbf{x}$.

Problem 2: Zero-Sum Games (Bertsekas Ex 6.6)

Let $\mathbf{A}$ be an $n \times m$ matrix. Consider the *zero sum game* where player 1 picks a row of $\mathbf{A}$ and player 2 picks a column of $\mathbf{A}$. Player 1 has the goal of picking as small an element as possible and Player 2 has the goal of picking as large an element as possible.

This problem will use duality to prove that the optimal strategy is independent of who goes first. That is

$$\max_{\mathbf{z} \in Z} \min_{\mathbf{x} \in X} \mathbf{x}^\top \mathbf{A} \mathbf{z} = \min_{\mathbf{x} \in X} \max_{\mathbf{z} \in Z} \mathbf{x}^\top \mathbf{A} \mathbf{z}$$

where $X = \{\mathbf{x} | \sum x_i = 1 x_i \geq 0\} \subset \mathbb{R}^n$ and $Z = \{\mathbf{z} | \sum z_i = 1 z_i \geq 0\} \subset \mathbb{R}^m$.

- For a fixed $\mathbf{z}$, show

$$\max_{\mathbf{z}\in Z}\min_{\mathbf{x}\in X}\mathbf{x}^\top\mathbf{A}\mathbf{z} = \max_{\mathbf{z}\in Z}\min\{[\mathbf{A}\mathbf{z}]_1,\ldots,[\mathbf{A}\mathbf{z}]_n\} = \max_{\mathbf{z}\in Z,[\mathbf{A}\mathbf{z}]_i\geq t} t$$

- In a similar fashion, show

$$\min_{\mathbf{x}\in X}\max_{\mathbf{z}\in Z}\mathbf{x}^\top\mathbf{A}\mathbf{z} = \min_{\mathbf{x}\in X,[\mathbf{A}^\top\mathbf{x}]_i\leq u} u$$

- Finally, show that the linear programs

$$\max_{\mathbf{z}\in Z,[\mathbf{A}\mathbf{z}]_i\geq t} t \quad \text{and} \quad \min_{\mathbf{x}\in X,[\mathbf{A}^\top\mathbf{x}]_i\leq u} u$$

are dual to each other.

Problem 3: Duality Gaps

Consider the non-convex quadratic program

$$\begin{array}{ll} \text{min} & x^2 + y^2 + z^2 + 2xy + 2yz + 2zx \\ \text{s.t.} & x^2 = y^2 = z^2 = 1 \end{array}$$

- Show that the dual problem is a semidefinite program (Hint: write the program in matrix form in terms of quadratic forms.)

- show that the dual optimum is zero

- By trying cases, show that the minimum of the primal is equal to one.