Lecture 3: Semidefinite Programming

Lecture Outline

- Part I: Semidefinite programming, examples, canonical form, and duality
- Part II: Strong Duality Failure Examples
- Part III: Conditions for strong duality
- Part IV: Solving convex optimization problems

Part I: Semidefinite Programming, Examples, Canonical Form, and Duality

Semidefinite Programming

- Semidefinite Programming: Want to optimize a linear function, can now have matrix positive semidefiniteness (PSD) constraints as well as linear equalities and inequalities
- Example: Maximize x subject to

$$\begin{bmatrix} 1 & x \\ x & 2+x \end{bmatrix} \succcurlyeq 0$$

- Answer: $x = 2$

Example: Goemans-Williamson

- First approximation algorithm using a semiefinite program (SDP)
- MAX-CUT reformulation: Have a variable x_i for each vertex i, will set $x_i = \pm 1$ depending on which side of the cut i is on.
- Want to maximize $\sum_{i,j:i<j,(i,j)\in E(G)} \frac{1-x_i x_j}{2}$ where $x_i \in \{-1,+1\}$ for all i.

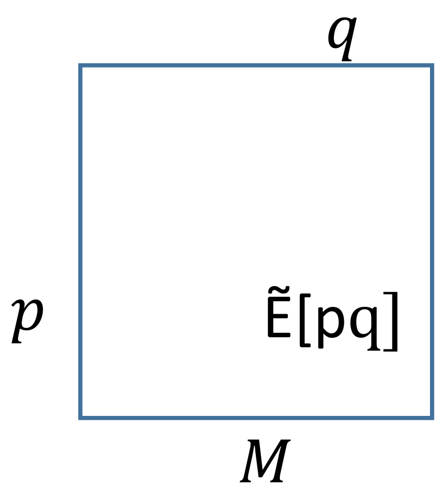
Example: Goemans-Williamson

- Idea: Take M so that $M_{ij} = x_i x_j$
- Want to maximize $\sum_{i,j:i<j,(i,j)\in E(G)} \frac{1-M_{ij}}{2}$ where $M_{ii} = 1$ for all i and $M = \mathrm{xx}^{\mathrm{T}}$.
- Relaxation: Maximize $\sum_{i,j:i<j,(i,j)\in E(G)} \frac{1-M_{ij}}{2}$ subject to
  1. $\forall i, M_{ii} = 1$
  2. $M \succcurlyeq 0$

Example: SOS Hierarchy

- Goal: Minimize a polynomial $h(x_1, \ldots, x_n)$ subject to constraints $s_1(x_1, \ldots, x_n) = 0$, $s_2(x_1, \ldots, x_n) = 0$, etc.
- Relaxation: Minimize $\tilde{\mathrm{E}}[h]$ where $\tilde{\mathrm{E}}$ is a linear map from polynomials of degree $\leq d$ to $\mathbb{R}$ satisfying:
  1. $\tilde{\mathrm{E}}[1] = 1$
  2. $\tilde{\mathrm{E}}[f s_i] = 0$ whenever $\deg(f) + \deg(s_i) \leq d$
  3. $\tilde{\mathrm{E}}\left[g^2\right] \geq 0$ whenever $\deg(g) \leq \frac{d}{2}$

The Moment Matrix



p,q are monomials of degree at most $\frac{d}{2}$.

- Indexed by monomials of degree $\leq \frac{d}{2}$
- $M_{pq} = \tilde{E}[pq]$
- Each g of degree $\leq \frac{d}{2}$ corresponds to a vector
- $\tilde{E}\left[g^2\right] = g^T M g$
- $\forall g, \tilde{E}\left[g^2\right] \geq 0 \Leftrightarrow M$ is PSD

Semidefinite Program for SOS

- Program: Minimize $\tilde{\mathrm{E}}[h]$ where $\tilde{\mathrm{E}}$ satisfies:
  1. $\tilde{\mathrm{E}}[1] = 1$
  2. $\tilde{\mathrm{E}}[f s_i] = 0$ whenever $\deg(f) + \deg(s_i) \leq d$
  3. $\tilde{\mathrm{E}}\left[g^2\right] \geq 0$ whenever $\deg(g) \leq \frac{d}{2}$
- Expressible as semidefinite program using M:
  1. $\forall h, \tilde{\mathrm{E}}[h]$ is a linear function of entries of M
  2. Constraints that $\tilde{\mathrm{E}}[1] = 1$ and $\tilde{\mathrm{E}}[f s_i] = 0$ give linear constraints on entries of M
  3. $\tilde{\mathrm{E}}\left[g^2\right] \geq 0$ whenever $\deg(g) \leq \frac{d}{2} \Leftrightarrow M \succcurlyeq 0$
  4. Also have SOS symmetry constraints

SOS symmetry

- Define $x_I = \prod_{i \in I} x_i$ where I is a multi-set
- SOS symmetry constraints: $M_{x_I x_J} = M_{x_{I'} x_{J'}}$ whenever $I \cup J = I' \cup J'$
- Example:

$$
\begin{array}{c} \\ 1 \\ x \\ y \\ x^2 \\ xy \\ y^2 \end{array}
\begin{array}{c} \begin{array}{cccccc} 1 & x & y & x^2 & xy & y^2 \end{array} \\
\begin{bmatrix} 1 & a & b & c & d & e \\ a & c & d & f & g & h \\ b & d & e & g & h & i \\ c & f & g & j & k & l \\ d & g & h & k & l & m \\ e & h & i & l & m & n \end{bmatrix} \end{array}
$$

Canonical Form

- Def: Define $X \bullet Y = \sum_{i,j} X_{ij} Y_{ij} = tr(XY^T)$ to be the entry-wise dot product of X and Y
- Canonical form: Minimize $C \bullet X$ subject to
  1. $\forall i, A_i \bullet X = b_i$ where the A_i are symmetric
  2. $X \succcurlyeq 0$

Putting Things Into Canonical Form

- Canonical form: Minimize $C \bullet X$ subject to
  1. $\forall i, A_i \bullet X = b_i$ where the A_i are symmetric
  2. $X \succcurlyeq 0$
- Ideas for obtaining canonical form:
  1. $X \succcurlyeq 0, Y \succcurlyeq 0 \Leftrightarrow \begin{bmatrix} X & 0 \\ 0 & Y \end{bmatrix} \succcurlyeq 0$
  2. Slack variables: $A_i \bullet X \leq b_i \Leftrightarrow A_i \bullet X = b_i + s_i, s_i \geq 0$
  3. Can enforce $s_i \geq 0$ by putting s_i on the diagonal of X

Semidefinite Programming Dual

- Primal: Minimize $C \bullet X$ subject to
  1. $\forall i, A_i \bullet X = b_i$ where the A_i are symmetric
  2. $X \succcurlyeq 0$
- Dual: Maximize $\sum_i y_i b_i$ subject to
  1. $\sum_i y_i A_i \preccurlyeq C$
- Value for dual lower bounds value for primal:

$$C \bullet X = \left(C - \sum_i y_i A_i\right) \bullet X + \left(\sum_i y_i A_i\right) \bullet X \geq \sum_i y_i b_i$$

Explanation for Duality

- Primal: Minimize $C \bullet X$ subject to
  1. $\forall i, A_i \bullet X = b_i$ where the A_i are symmetric
  2. $X \succcurlyeq 0$
- $= \min_{X \succcurlyeq 0} \max_{y} C \bullet X + \sum_i y_i (b_i - A_i \bullet X)$
- $= \max_{y} \min_{X \succcurlyeq 0} \sum_i y_i b_i + (C - \sum_i y_i A_i) \bullet X$
- Dual: Maximize $\sum_i y_i b_i$ subject to
  1. $\sum_i y_i A_i \preccurlyeq C$

In class exercise: SOS duality

- Exercise: What is the dual of the semidefinite program for SOS?
- Primal: Minimize $\tilde{\mathrm{E}}[h]$ where $\tilde{\mathrm{E}}$ is a linear map from polynomials of degree $\leq d$ to $\mathbb{R}$ such that:
  1. $\tilde{\mathrm{E}}[1] = 1$
  2. $\tilde{\mathrm{E}}[f s_i] = 0$ whenever $\deg(f) + \deg(s_i) \leq d$
  3. $\tilde{\mathrm{E}}\left[g^2\right] \geq 0$ whenever $\deg(g) \leq \frac{d}{2}$

In class exercise solution

- Definition: Given a symmetric matrix Q indexed by monomials x_I, we say that Q represents the polynomial
$$p_Q = \sum_J \sum_{I,I':I\cup I'=J} Q_{x_I x_{I'}} x_J$$
- Proposition 1: If $Q \succcurlyeq 0$ then p_Q is a sum of squares. Conversely, if p is a sum of squares then $\exists Q \succcurlyeq 0: p = p_Q$
- Proposition 2: If M is a moment matrix then $M \bullet Q = \tilde{\mathrm{E}}[p_Q]$

In class exercise solution continued

- $C = H$ where $p_H = h$
- Constraint that $\tilde{E}[1] = 1$ gives matrix

$$A_1 = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \text{ and } b_1 = 1$$

- Constraints that $\tilde{E}[f s_i] = 0$ give matrices A_j where $p_{A_j} = f s_i$ and $b_j = 0$
- SOS symmetry constraints give matrices A_k such that $p_{A_k} = 0$ and $b_k = 0$

In class exercise solution continued

- Recall dual: Maximize $\sum_i y_i b_i$ subject to
  1. $\sum_i y_i A_i \preccurlyeq C$
- Here: Maximize c such that

$$cA_1 + \sum_j y_j A_j + \sum_k y_k A_k \preccurlyeq H$$

- This is the answer, but let's simplify it into a more intuitive form.
- Let $Q = H - cA_1 + \sum_j y_j A_j + \sum_k y_k A_k$
- $Q \succcurlyeq 0$

In class exercise solution continued

- $H = cA_1 + \sum_j y_j A_j + \sum_k y_k A_k + Q$,

$$\mathrm{A}_1 = \begin{bmatrix} 1 & 0 & \cdots \\ 0 & 0 & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix}, Q \succcurlyeq 0$$

- View everything in terms of polynomials.
- $p_H = h, p_{A_1} = 1, p_{(\sum_j y_j A_j)} = \sum_i f_i s_i$ for some $f_i, p_{\sum_k y_k A_k} = 0, p_Q = \sum_j g_j^2$ for some g_j
- $h = c + \sum_i f_i s_i + \sum_j g_j^2$

In class exercise solution continued

- Simplified Dual: Maximize c such that

$$h = c + \sum_i f_i s_i + \sum_j g_j^2$$

- This is a Positivstellensatz proof that $h \geq c$ (see Lectures 1 and 5)

Part II: Strong Duality Failure Examples

Strong Duality Failure

- Unlike linear programming, it is not always the case that the values of the primal and dual are the same.
- However, almost never an issue in practice, have to be trying in order to break strong duality.
- We’ll give this issue its due here then ignore it for the rest of the seminar.

Non-attainability Example

- Primal: Minimize x_2 subject to $\begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix} \succcurlyeq 0$
- Dual: Maximize $2y$ subject to
  1. $\begin{bmatrix} 0 & y \\ y & 0 \end{bmatrix} \preccurlyeq \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$
- Duality demonstration:

$$\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & y \\ y & 0 \end{bmatrix}\right) \bullet \begin{bmatrix} x_1 & 1 \\ 1 & x_2 \end{bmatrix} = x_2 - 2y \geq 0$$

- Dual has optimal value 0, this is not attainable in the primal (we can only get arbitrarily close)

Duality Gap Example

- Primal: Minimize $x_2 + 1$ subject to

$$\begin{bmatrix} 1 + x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succcurlyeq 0$$

- Dual: Maximize $2y$ subject to

$$\begin{bmatrix} 2y & y_1 & y_2 \\ y_1 & 0 & -y \\ y_2 & -y & y_3 \end{bmatrix} \preccurlyeq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Duality demonstration

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 2y & y_1 & y_2 \\ y_1 & 0 & -y \\ y_2 & -y & y_3 \end{bmatrix} \right) \bullet \begin{bmatrix} 1 + x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} = (x_2 + 1) - 2y \geq 0$$

Duality Gap Example

- Primal: Minimize $x_2 + 1$ subject to

$$\begin{bmatrix} 1 + x_2 & 0 & 0 \\ 0 & x_1 & x_2 \\ 0 & x_2 & 0 \end{bmatrix} \succcurlyeq 0$$

- Has optimal value 1 as we must have $x_2 = 0$
- Dual: Maximize $2y$ subject to

$$\begin{bmatrix} 2y & y_1 & y_2 \\ y_1 & 0 & -y \\ y_2 & -y & y_3 \end{bmatrix} \preccurlyeq \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

- Has optimal value 0 as we must have $y = 0$.
- Note: This example was taken from Lecture 13, EE227A at Berkeley given on October 14, 2008.

Part III: Conditions for strong duality

Sufficient Strong Duality Conditions

- How can we rule out such a gap?
- Slater's Condition (informal):
  - If the feasible region for the primal has an interior point (in the subspace defined by the linear equalities) then the duality gap is 0. Moreover, if the optimal value is finite then it is attainable in the dual.
- Also sufficient if either the primal or the dual is feasible and bounded (i.e. any very large point violates the constraints)

Recall Minimax Theorem

- Von Neumann [1928]: If X and Y are convex compact subsets of R^m and R^n and $f: X \times Y \to R$ is a continuous function which is convex in X and concave in Y then

$$\max_{y \in Y} \min_{x \in X} f(x, y) = \min_{x \in X} \max_{y \in Y} f(x, y)$$

- Issue: X and Y are unbounded in our setting.

Minimax in the limit

- Idea: Minimax applies for arbitrarily large X, Y so long as they are bounded
- Can take the limit as X, Y get larger and larger
- Question: Do we bound X or Y first?
- If X is bounded first, get primal: $\min_{x \in X} \max_{y \in Y} f(x, y)$
- If Y is bounded first, get dual: $\max_{y \in Y} \min_{x \in X} f(x, y)$
- If we can show it doesn't matter which is bounded first, have duality gap 0.

Formal Statements

- Def: Define $B(R) = \{x : \|x\| \leq R\}$
- Minimax Theorem: For all finite R_1, R_2

$$\max_{y \in Y \cap B(R_2)} \min_{x \in X \cap B(R_1)} f(x, y) = \min_{x \in X \cap B(R_1)} \max_{y \in Y \cap B(R_2)} f(x, y)$$

- Let's call this value $opt(R_1, R_2)$
- Primal value: $\lim_{R_1 \to \infty} \lim_{R_2 \to \infty} opt(R_1, R_2)$
- Dual value: $\lim_{R_2 \to \infty} \lim_{R_1 \to \infty} opt(R_1, R_2)$
- Sufficient for 0 gap: Show that

$$\exists R : \forall R_1, R_2 \geq R, opt(R_1, R_2) = opt(R, R_2)$$

Boundedness Condition Intuition

- Assume $\exists$ feasible $x \in B(R)$ but when $\|x\| > R$, constraints on primal are violated.
- Want to show that $\exists R'$
  $$\forall R_1, R_2 \geq R', opt(R_1, R_2) = opt(R', R_2)$$
- We are in the finite setting, so we can assume x player plays first.
- Intuition: y player can heavily punish x player for violated constraints, so x player should always choose an $x \in B(R')$.
- Similar logic applies to the dual.

Slater's Condition Intuition

- Idea: Strictly feasible point x shows it is bad for y player to play a very large y.
- Primal: Minimize $h(x)$ subject to $\{g_i(x) \leq c_i\}$ where the g_i are convex.
- $f(x, y) = \sum_i y_i(g_i(x) - c_i) + h(x)$ (we'll restrict ourselves to non-negative y)
- Dual: Doesn't seem to have a nicer form than

$$\max_{y \geq 0} \min_x \left(\sum_i y_i(g_i(x) - c_i) + h(x)\right)$$

Slater's Condition Intuition

- Primal: Minimize $h(x)$ subject to $\{g_i(x) \leq c_i\}$
- Dual: $\max_{y \geq 0} \min_x \left(\sum_i y_i (g_i(x) - c_i) + h(x) \right)$
- Key observation: If $\forall i, g_i(x) < c_i$, x punishes very large y. Thus, y is effectively bounded.

Strong Duality Conditions Summary

- Strong duality may fail for semidefinite programs.
- However, strong duality holds if the program is at all robust (Slater’s condition is satisfied) or either the primal or dual is feasible and bounded (any very large point violates the constraints)
- Note: working over the hypercube satisfies boundedness.

Part III: Solving convex optimization problems

Solving Convex Optimization Problems

- In practice: simplex methods or interior point methods work best
- First polynomial time guarantee: ellipsoid method
- This seminar: We'll use algorithms as a black-box and assume that semidefinite programs of size n^d can be solved in time $n^{O(d)}$.
- Note: Can fail to be polynomial time in pathological cases (see Ryan O'Donnell's note), almost never an issue in practice

Usefulness of Convexity

- Want to minimize a convex function f over a convex set X.
- All local minima are global minima: If $f(x) <$ $f(y)$ then $f(y)$ is not a local minima as

$$\forall \epsilon > 0, f(\epsilon x + (1-\epsilon)y) \leq \epsilon f(x) + (1-\epsilon)\mathrm{f}(\mathrm{y})$$

- Can always go from the current point x towards a global minima.

Reduction to Feasibility Testing

- Want to minimize a convex function f over a convex set X
- Testing whether we can achieve $f(x) \leq c$ is equivalent to finding a point in the convex set $\mathrm{X_c} = X \cap \{x : f(x) \leq c\}$
- If we can solve this feasibility problem, we can use binary search to approximate the optimal value.

Cutting-plane Oracles

- Let X be a convex set. Given a point $x_0 \notin X$, a cutting-plane oracle returns a hyperplane H passing through x_0 such that X is entirely on one side of H.
- Intuition for obtaining a cutting-plane oracle: If $x_0 \notin X$ then there is a constraint x_0 violates. This constraint is of the form $f(x_0) < c$ where f is convex. X must be inside the half-space
$$\nabla f \cdot (x - x_0) \leq 0$$

Ellipsoid Method Sketch

- Algorithm: Let X be the feasible set
  1. Keep track of an ellipsoid S containing X
  2. At each step, query center c of S
  3. If $c \in X$, output c. Otherwise, cutting-plane oracle gives a hyperplane H passing through c and X is on one side of H. Use H to find a smaller ellipsoid S'.
- Initial Guarantees:
  1. $X \subseteq B(R)$ where $B(R) = \{x \in R^n : \|x\| \leq R\}$
  2. X contains a ball of radius r.
- Note: Not polynomial time if $\frac{R}{r}$ is $2^{(n^{\omega(1)})}$

References

- L. E. Ghaoui. Lecture 13: SDP Duality. Berkeley EE227A Convex Optimization and Applications. 2008
- [GW95] M. X. Goemans and D. P. Williamson. Improved Approximation Algorithms for Maximum Cut and Satisfiability Problems Using Semidefinite Programming. J. ACM, 42(6):1115–1145, 1995.
- [O'Don17] R. O'Donnell. SOS is not Obviously Automatizable, Even Approximately. ITCS 2017.
- [Sla50] M. Slater. Lagrange Multipliers Revisited. Cowles Commission Discussion Paper: Mathematics 403, 1950.