# 6.215/6.255J/15.093J/IDS.200J: Optimization Methods

Problem Set 2

Due: October 12, 2021 1:00 PM

Several of the problems are from the course textbook, *Introduction to linear optimization*, by D. Bertsimas and J. Tsitsiklis.

**Problem 1: (10 points)** Bertsimas & Tsitsiklis, Exercise 2.4

**Solution** Consider a linear program of the form

$$\begin{aligned} \text{minimize } & \mathbf{c}'\mathbf{x} \\ \text{subject to } & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned} \tag{1}$$

where $\mathbf{A}$ is a $m \times n$ matrix, $\mathbf{c}, \mathbf{x}$ are $n$-dimensional vectors, and $\mathbf{b}$ is an $m$-dimensional vector. We add slack variables to get a standard form LP:

$$\begin{aligned} \text{minimize } & \mathbf{c}'\mathbf{x} \\ \text{subject to } & \begin{pmatrix} \mathbf{A} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} = \mathbf{b} \\ & \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \geq 0. \end{aligned} \tag{2}$$

Here, $\begin{pmatrix} \mathbf{x} & \mathbf{y} \end{pmatrix}'$ is an $(n+m)$-dimensional vector. Note that while problems (1) and (2) are equivalent, the polyhedron over which each problem optimizes are different! In (1), we have a polyhedron in $\mathbf{R}^n$, and in (2), we have a polyhedron in $\mathbf{R}^{n+m}$. This is the caveat: while any linear programming problem can be converted to an equivalent problem in standard form, the two problems may be optimizing over different polyhedra. Since extreme points are a purely geometric property, we cannot use the arguments described in the problem to necessarily conclude that every nonempty polyhedron has at least one extreme point.

**Common mistake:** Showing just that the result (every polyhedron has a extreme point) is wrong. A complete answer must address the argument itself and mention for example that the dimension/geometry changes when we move to standard form and that extreme points are not the same in standard form.

**Problem 2: (10 points)** Bertsimas & Tsitsiklis, Exercise 4.1

**Solution** We are given the following linear program:

$$
\begin{aligned}
\text{minimize} \quad & x_1 - x_2 \\
\text{subject to} \quad & 2x_1 + 3x_2 - x_3 + x_4 \le 0 \\
& 3x_1 + x_2 + 4x_3 - 2x_4 \ge 3 \\
& -x_1 - x_2 + 2x_3 + x_4 = 6 \\
& x_1 \le 0 \\
& x_2 \ge 0 \\
& x_3 \ge 0
\end{aligned}
$$

The dual is:

$$
\begin{aligned}
\text{maximize} \quad & 3p_2 + 6p_3 \\
\text{subject to} \quad & 2p_1 + 3p_2 - p_3 \ge 1 \\
& 3p_1 + p_2 - p_3 \le -1 \\
& -p_1 + 4p_2 + 2p_3 \le 0 \\
& p_1 - 2p_2 + p_3 = 0 \\
& p_1 \le 0 \\
& p_2 \ge 0 \\
& p_3 \quad \text{free}
\end{aligned}
$$

**Problem 3: (15 points)** Solving primal and dual problems in Julia.

**(a)** Find the dual problem of the following LP:

$$
\begin{aligned}
\text{minimize} \quad & c'x + d'y \\
\text{subject to} \quad & Ax + By = b \\
& Dx \le f \\
& Gy \le g
\end{aligned}
$$

where $x \in \mathbf{R}^n, y \in \mathbf{R}^m$ are the decision variables, and everything else is the problem data. We have $A \in \mathbf{R}^{p_1 \times n}, B \in \mathbf{R}^{p_1 \times m}, D \in \mathbf{R}^{p_2 \times n}$, and $G \in \mathbf{R}^{p_3 \times m}$.

**Solution**

$$
\begin{aligned}
\text{maximize} \quad & b'v_1 + f'v_2 + g'v_3 \\
\text{subject to} \quad & A'v_1 + D'v_2 = c \\
& B'v_1 + G'v_3 = d \\
& v_1 : \text{free}, v_2 \le 0, v_3 \le 0,
\end{aligned}
$$

where $v_1 \in \mathbf{R}^{p_1}, v_2 \in \mathbf{R}^{p_2}, v_3 \in \mathbf{R}^{p_3}$ are the decision variables.

**(b)** For the given data in `data.jld2` (i.e., instances of $c$, $d$, $A$, $B$, $b$, $D$, $f$, $G$, $g$), solve both the primal problem and the dual problem and show that both have the same optimal solution. In this data file, $p_1 = 25, p_2 = 100, p_3 = 200, n = 50, m = 100$. Load the data into your Julia code by running the following commands:

```
using Pkg

Pkg.add("JLD2")

using JLD2

@load "data.jld2" c d A B b D f G g
```

**Solution** See the IJulia notebook `problem_3_sol.ipynb`

**Problem 4: (20 points)** In order to implement complicated nonlinear functions in a computer, sometimes polynomial approximations are used. In this exercise, we explore one way of computing these using linear programming.

Consider a scalar function $f(x)$, which we are trying to approximate over the interval $[a, b]$ with a polynomial $p(x)$ of degree $d$. As a measure of how well the polynomial approximates the function, we can use the norm

$$||f - p||_\infty := \sup_{x \in [a,b]} |f(x) - p(x)|.$$

The *minimax* or *Chebyshev* polynomial approximation of degree $d$ of $f(x)$ is then defined as

$$\min_{p \in P_d} ||f(x) - p(x)||_\infty,$$

where $P_d$ is the set of polynomials of degree less than or equal to $d$. Because the true $|| \cdot ||_\infty$ norm can sometimes be troublesome to compute, throughout this exercise we will use instead a discrete approximation given by:

$$||g||_{\infty,N} := \max_{x_i \in [a,b]} |g(x_i)|,$$

where the $x_i$ is a set of $N$ points equispaced on the interval.

**(a)** (10 points) Give a linear programming formulation of the Chebyshev approximation problem in the $|| \cdot ||_{\infty,N}$ norm.

**Solution** A polynomial of degree $d$ can be written as $p(x) = \sum_{i=0}^d a_i x^i$, where $a_i$, $i = 0, \ldots, d$ are $d + 1$ coefficients. $N$ sample points equispaced on the interval $[a, b]$ are $x_i = a + (b - a)i/(N - 1)$, $i = 0, \ldots, N - 1$. The *Chebyshev approximation* problem in the norm $\| \cdot \|_{\infty,N}$ is then

$$\min_{a_0,\ldots,a_d} \max_{0 \le i \le N-1} |f(x_i) - p(x_i)|$$

Let $z = \max_{0 \le i \le N-1} |f(x_i) - p(x_i)|$, we have:

$$\begin{cases} z \ge f(x_i) - p(x_i) & \text{for } i = 0, 1, \ldots, N - 1 \\ z \ge -f(x_i) + p(x_i) & \text{for } i = 0, 1, \ldots, N - 1 \end{cases}$$

The linear programming formulation can then be written as follows

$$\begin{aligned} \min_{z,a_0,\ldots,a_d} \quad & z \\ \text{s.t.} \quad & z + \sum_{j=0}^d x_i^j a_j \ge f(x_i) \quad i = 0, 1, \ldots, N - 1 \\ & z - \sum_{j=0}^d x_i^j a_j \ge -f(x_i) \quad i = 0, 1, \ldots, N - 1 \end{aligned}$$

**Remark:** Minimizing over the set of all polynomial functions is not a linear programming. However, the form of polynomials $\sum_{i=0}^d a_i x^i$ allows us to transform the problem into a linear programming, by minimizing over the choices of the coefficients $a_i$ instead.

**Common mistake:** Giving an optimization problem where the variable is a polynomial. This is not linear.

**(b)** (5 points) Write the corresponding dual problem.

**Solution** Let $p_i^+$ and $p_i^-$ be the dual variables for two primal constraints above. The corresponding dual problem can be formulated as follows

$$\begin{aligned}
\max \quad & \sum_{i=0}^{N-1} f(x_i)(p_i^+ - p_i^-) \\
\text{s.t.} \quad & \sum_{i=0}^{N-1} (p_i^+ + p_i^-) = 1 \\
& \sum_{i=0}^{N-1} x_i^j (p_i^+ - p_i^-) = 0 \quad j = 0, 1, \ldots, d \\
& p_i^+, p_i^- \geq 0 \quad i = 0, 1, \ldots, N-1
\end{aligned}$$

**(c)** (5 points) Give an interpretation of the complementary slackness conditions.

**Solution** The complementary slackness conditions are:

$$\begin{cases} (z + \sum_{j=0}^{d} x_i^j a_j - f(x_i))p_i^+ = 0 & \text{for } i = 0, 1, \ldots, N-1 \\ (z - \sum_{j=0}^{d} x_i^j a_j + f(x_i))p_i^- = 0 & \text{for } i = 0, 1, \ldots, N-1 \end{cases}$$

If $p_i^+ > 0$ or $p_i^- > 0$, then we have $z = |p(x_i) - f(x_i)|$. That is, at $x_i$, the maximum approximation error is achieved. The dual problem is in standard form with $d+2$ constraints. Therefore, there are at most $d+2$ positive values in the optimal basic solution, which means there are at most $d+2$ points $x_i$, where the maximum approximation errors is achieved.

**Common mistake:** A complete answer has to mention these two aspects: where max error is attained and how many points attain it.

**(d) Bonus:** (8 points) Using the finite approximation described, compute with Julia the minimax polynomial approximant of the function $f(x) = e^x$ in the interval $[-1, 1]$, for $d = 0, 1, 2, 3, 4$ and $N = 100$. Plot the resulting approximating polynomials, as well as the approximation errors.
*You do not need to provide your code for this question. You are only required to write the resulting errors and show the plots.*

**Solution**

- $d = 1$ The Chebyshev polynomial is $p(x) = 1.2643 + 1.1752x$ with the maximum error of $z = 0.2788$
- $d = 2$ The Chebyshev polynomial is $p(x) = 0.9890 + 1.1302x + 0.5540x^2$ with the maximum error of $z = 0.0450$
- $d = 3$ The Chebyshev polynomial is $p(x) = 0.9946 + 0.9957x + 0.5430x^2 + 0.1795x^3$ with the maximum error of $z = 0.0055$
- $d = 4$ The Chebyshev polynomial is $p(x) = 1.0001 + 0.9973x + 0.4988x^2 + 0.1773x^3 + 0.0442x^4$ with the maximum error of $z = 0.0005$

**(e) Bonus:** Graphically compare the results with the $d$th order Taylor expansion around the origin $\sum_{i=0}^{d} x^i/(i!)$. What do you observe?

**Solution** The Taylor expansion of $f(x) = e^x$ at $x = 0$ up to degree $d$ is $p_T(x) = \sum_{i=0}^{d} x^i/(i!)$. As $d$ increases, the Taylor expansion approximates $f(x)$ accurately with larger interval around $x = 0$. However, the approximation errors are still very high outside that interval whereas Chebyshev approximation reduces the maximum approximation error. For example, when $d = 4$, the maximum approximation errors are $z_T = 10^{-2}$ and $z = 5 \times 10^{-4}$ for Taylor expansion and Chebyshev polynomial respectively.

**Problem 5: (15 points)** Bertsimas & Tsitsiklis, Exercise 4.28

**Solution** We want to show that the following two statements are equivalent:

(a) For all $\mathbf{x} \geq 0$, we have $\mathbf{a}'\mathbf{x} \leq \max_i \mathbf{a}_i'\mathbf{x}$.
(b) There exist nonnegative coeffcients $\lambda_i$ that sum to 1 and such that $\mathbf{a} \leq \sum_{i=1}^m \lambda_i \mathbf{a}_i$.

First, we show that (b) $\Rightarrow$ (a). Suppose (b) is true. Then for every $\mathbf{x} \geq 0$, we have

$$\mathbf{a}'\mathbf{x} \leq \sum_{i=1}^m \lambda_i \mathbf{a}_i'\mathbf{x} \leq \max_i \mathbf{a}_i'\mathbf{x},$$

since for any set of values, the convex combination is always less than or equal to the maximum. Therefore, (b) $\Rightarrow$ (a).

Now we show (a) $\Rightarrow$ (b). Suppose (a) is true. Consider the following linear program:

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^m 0 \cdot \lambda_i \\ \text{subject to} \quad & \sum_{i=1}^m \lambda_i \mathbf{a}_i \geq \mathbf{a} \\ & \sum_{i=1}^m \lambda_i = 1 \\ & \lambda_i \geq 0 \quad i = 1, 2, \ldots, m \end{aligned}$$

and its dual

$$\begin{aligned} \text{maximize} \quad & \mathbf{a}'\mathbf{x} + y \\ \text{subject to} \quad & \mathbf{a}_i'\mathbf{x} + y \leq 0 \quad i = 1, 2, \ldots, m \\ & \mathbf{x} \geq 0 \end{aligned}$$

Note that for any dual feasible solution $(\mathbf{x}, y)$

$$\mathbf{a}'\mathbf{x} + y \leq \mathbf{a}'\mathbf{x} - \mathbf{a}_i'\mathbf{x} \quad \text{for } i = 1, 2, \ldots, m. \tag{3}$$

Since (a) is true, $\mathbf{a}'\mathbf{x} \leq \mathbf{a}_i'\mathbf{x}$ for *some* $i$. Using this fact and (3), we can bound the dual objective value from above:

$$\mathbf{a}'\mathbf{x} + y \leq 0 \quad \text{for all dual feasible solutions } (\mathbf{x}, y).$$

Note that $(\mathbf{0}, 0)$ is a dual feasible solution with objective value 0. Since the dual objective value is bounded above by zero, $(\mathbf{0}, 0)$ must be a dual optimal solution. By strong duality, we can conclude that the primal problem also has an optimal (and feasible!) solution. Note that any primal feasible solution satisfies (b). Therefore, (a) implies (b).

**Common mistake:** Just writing the primal and dual and then QED is not a complete answer!! A complete answer must then argue that we then use weak duality as the primal is feasible AND bounded. Another common mistake is confusing $\max_i(a_i'x)$ and $(\max_i a_i)'x$. In general, $\max_i(a_i'x) \neq (\max_i a_i)'x$.