30

Thanks to Tim Carnes
ORIE 630
Homework #11
12/2/2005

SOLUTION SET

1. a) Suppose that $E(z_+, B_+)$ is the minimum volume ellipsoid containing

$$\{x \in E(z,B) : a^T x \leq a^T z - \alpha(a^T Ba)^{\frac{1}{2}}\},$$

where $\alpha > -1/m$ and $0 \neq a \in \mathbb{R}^m$. Show that

$$a^T z - \alpha(a^T Ba)^{\frac{1}{2}} = a^T z_+ + \frac{1}{m}(a^T B_+ a)^{\frac{1}{2}},$$

i.e., the "depth" of the constraint that was used to make the cut is exactly $-1/m$ in the new ellipsoid.

b) Suppose we apply the ellipsoid method to try to find a point in

$$\{x \in \mathbb{R}^2 : x_1 \leq \frac{1}{2}, -x_1 \leq -\frac{1}{2}, -x_2 \leq -\frac{1}{4}, x_2 \leq \frac{1}{2}\},$$

starting with $E_0 := \{x \in \mathbb{R}^2 : \|x\| \leq 1\}$. At each iteration, we choose as the cut to define the new ellipsoid the constraint $a_i^T x \leq b_i$ with maximum depth

$$\alpha_i := \frac{a_i^T z - b_i}{(a_i^T B a_i)^{\frac{1}{2}}},$$

stopping if all α_i's are nonpositive, and using the deep cut method (i.e., the ellipsoid is updated as in (a)).

(i) What are the depths of all the constraints, and what cut is chosen, at the first iteration?

(ii) What are the depths of all the constraints, and what cut is chosen, at the second iteration?

a) By Theorem 2 of the lecture of 11/17, we know

$$\begin{aligned} z_+ &= z - \tau \frac{Ba}{\sqrt{a^T Ba}} \\ B_+ &= \delta\left(B - \sigma \frac{Baa^T B}{a^T Ba}\right) \end{aligned}$$

where $\tau = \frac{1+m\alpha}{m+1}$, $\delta = \frac{(1-\alpha)^2 m^2}{m^2-1}$, and $\sigma = \frac{2(1+m\alpha)}{(m+1)(1+\alpha)}$. Now using the fact that $-1/m < \alpha < 1$ we can calculate directly

$$\begin{aligned}
a^T z_+ + \frac{1}{m}(a^T B_+ a)^{1/2} &= a^T z - \tau \frac{a^T B a}{(a^T B a)^{1/2}} + \frac{\delta^{1/2}}{m}\left(a^T B a - \sigma \frac{a^T B a a^T B a}{a^T B a}\right)^{1/2} \\
&= a^T z - \tau (a^T B a)^{1/2} + \frac{\delta^{1/2}}{m}\left((1-\sigma)(a^T B a)\right)^{1/2} \\
&= a^T z - \left[\tau - \frac{\delta^{1/2}(1-\sigma)^{1/2}}{m}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\tau - \frac{\left(\frac{(1-\alpha)^2 m^2}{m^2-1}\right)^{1/2}\left(1 - \frac{2(1+m\alpha)}{(m+1)(1+\alpha)}\right)^{1/2}}{m}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\tau - \left(\frac{(1-\alpha)^2}{m^2-1} - \frac{2(1+m\alpha)(1-\alpha)}{(m-1)(m+1)^2}\right)^{1/2}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\tau - \left(\frac{(m+1)(1-\alpha)^2 - 2(1+m\alpha)(1-\alpha)}{(m-1)(m+1)^2}\right)^{1/2}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\tau - \left(\frac{m\alpha^2 - 2m\alpha + m - \alpha^2 + 2\alpha - 1}{(m-1)(m+1)^2}\right)^{1/2}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\tau - \left(\frac{(m-1)(1-\alpha)^2}{(m-1)(m+1)^2}\right)^{1/2}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\frac{1+m\alpha}{m+1} - \frac{1-\alpha}{m+1}\right](a^T B a)^{1/2} \\
&= a^T z - \left[\frac{(m+1)\alpha}{m+1}\right](a^T B a)^{1/2} \\
&= a^T z - \alpha (a^T B a)^{1/2},
\end{aligned}$$

which is the desired result. Note that the fact that $-1/m < \alpha < 1$ was used to ensure the square roots of $(1-\sigma)$ and δ could be taken.

b) (i) For this problem we have

$$A^T = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & -1 \\ 0 & 1 \end{pmatrix}, b = \begin{pmatrix} 1/2 \\ -1/2 \\ -1/4 \\ 1/2 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, z = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

So calculating the depths of each constraint we obtain

$$
\begin{aligned}
\alpha_1 &= \frac{\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} - 1/2}{\left[\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right]^{\frac{1}{2}}} \\
&= -1/2 \\
\alpha_2 &= \frac{\begin{pmatrix}-1 & 0\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} + 1/2}{\left[\begin{pmatrix}-1 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}-1\\0\end{pmatrix}\right]^{\frac{1}{2}}} \\
&= 1/2 \\
\alpha_3 &= \frac{\begin{pmatrix}0 & -1\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} + 1/4}{\left[\begin{pmatrix}0 & -1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}0\\-1\end{pmatrix}\right]^{\frac{1}{2}}} \\
&= 1/4 \\
\alpha_4 &= \frac{\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} - 1/2}{\left[\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\right]^{\frac{1}{2}}} \\
&= -1/2.
\end{aligned}
$$

So the cut that is chosen is

$$
\begin{aligned}
a_2^T x &\leq a_2^T z - \alpha_2 (a_2^T B a_2)^{1/2} \\
-x_1 &\leq -\frac{1}{2}.
\end{aligned}
$$

(ii) Using Theorem 2 of the lecture of 11/17, with $m = 2$, we calculate

$$
\begin{aligned}
\tau &= \frac{1+m\alpha_2}{m+1} \\
&= \frac{2}{3} \\
\delta &= \frac{(1-\alpha_2)^2 m^2}{m^2-1} \\
&= 1 \\
\sigma &= \frac{2(1+m\alpha_2)}{(m+1)(1+\alpha_2)} \\
&= \frac{8}{9} \\
z_+ &= z - \tau \frac{Ba_2}{\sqrt{a_2^T B a_2}} \\
&= -\frac{2}{3}\begin{pmatrix} -1 \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} 2/3 \\ 0 \end{pmatrix} \\
B_+ &= \delta\left(B - \sigma \frac{Ba_2 a_2^T B}{a_2^T B a_2}\right) \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{8}{9}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1/9 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}
$$

So we again calculate the depth of each constraint

$$
\begin{aligned}
\alpha_1 &= \frac{\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}2/3\\0\end{pmatrix} - 1/2}{\left[\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1/9 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}\right]^{\frac{1}{2}}}\\
&= 1/2\\
\alpha_2 &= \frac{\begin{pmatrix}-1 & 0\end{pmatrix}\begin{pmatrix}2/3\\0\end{pmatrix} + 1/2}{\left[\begin{pmatrix}-1 & 0\end{pmatrix}\begin{pmatrix}1/9 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}-1\\0\end{pmatrix}\right]^{\frac{1}{2}}}\\
&= -1/2\\
\alpha_3 &= \frac{\begin{pmatrix}0 & -1\end{pmatrix}\begin{pmatrix}2/3\\0\end{pmatrix} + 1/4}{\left[\begin{pmatrix}0 & -1\end{pmatrix}\begin{pmatrix}1/9 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}0\\-1\end{pmatrix}\right]^{\frac{1}{2}}}\\
&= 1/4\\
\alpha_3 &= \frac{\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}2/3\\0\end{pmatrix} - 1/2}{\left[\begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1/9 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}0\\1\end{pmatrix}\right]^{\frac{1}{2}}}\\
&= -1/2.
\end{aligned}
$$

So the cut that is chosen is

$$
\begin{aligned}
a_1^T x &\leq a_1^T z_+ - \alpha_1 (a_1^T B_+ a_1)^{1/2}\\
x_1 &\leq \frac{1}{2}.
\end{aligned}
$$

∎

2. Let $A \in \mathbb{R}^{m\times n}$ have rank m, and let $P_A := I - A^T(AA^T)^{-1}A$.

a) Show that $P_A = P_A^T = P_A^2$ and hence that $u^T P_A u = \|P_A u\|^2$ for every $u \in \mathbb{R}^n$. (So P_A is positive semidefinite: $u^T P_A u \geq 0$ for all u.)

b) Show that $P_A v = 0$ for every v in the range space of A^T, and $P_A v = v$ for every v in the null space of A.

a) We have

$$\begin{aligned} P_A^T &= (I - A^T(AA^T)^{-1}A)^T \\ &= I^T - A^T(AA^T)^{-T}(A^T)^T \\ &= I - A^T((A^T)^TA^T)^{-1}A \quad \checkmark \\ &= I - A^T(AA^T)^{-1}A \\ &= P_A, \end{aligned}$$

and

$$\begin{aligned} P_A^2 &= (I - A^T(AA^T)^{-1}A)(I - A^T(AA^T)^{-1}A) \\ &= I - A^T(AA^T)^{-1}A - A^T(AA^T)^{-1}A + A^T(AA^T)^{-1}AA^T(AA^T)^{-1}A \\ &= P_A - A^T(AA^T)^{-1}A + A^T(AA^T)^{-1}A \\ &= P_A. \end{aligned}$$

Using these results we obtain

$$\begin{aligned} u^TP_Au &= u_TP_A^2u \\ &= u_TP_AP_Au \\ &= u_TP_A^TP_Au \\ &= (P_Au)^T(P_Au) \\ &= \|P_Au\|^2. \end{aligned}$$

b) If v is in the range space of A^T, then there is some $x \in \mathbb{R}^m$ such that $A^Tx = v$. Thus

$$\begin{aligned} P_Av &= P_AA^Tx \\ &= (I - A^T(AA^T)^{-1}A)A^Tx \\ &= A^Tx - A^T(AA^T)^{-1}AA^Tx \\ &= A^Tx - A^Tx \\ &= 0. \end{aligned}$$

✓

Now if v is in the null space of A then $Av = 0$ so

$$\begin{aligned} P_Av &= (I - A^T(AA^T)^{-1}A)v \\ &= v - A^T(AA^T)^{-1}Av \\ &= v - A^T(AA^T)^{-1}0 \\ &= v. \end{aligned}$$

■

3. Consider the standard-form LP problem and its dual, where $A \in \mathbb{R}^{m \times n}$ has rank m, and suppose $x \in \mathcal{F}^0(P)$ and $(y,s) \in \mathcal{F}^0(D)$. Let $\mu = x^T s/n$, and suppose that $x_j s_j \geq \gamma\mu$ for all j, for some positive γ. Suppose $(\Delta x, \Delta y, \Delta s)$ is the solution to

$$\begin{array}{rcccl} & A^T\Delta y & + & \Delta s & = 0, \\ A\Delta x & & & & = 0, \\ S\Delta x & & + & X\Delta s & = \sigma\mu e - XSe, \end{array}$$

for some $0 \leq \sigma \leq 1$. Let $(x(\alpha), y(\alpha), s(\alpha)) := (x,y,s) + \alpha(\Delta x, \Delta y, \Delta s)$ for $0 \leq \alpha \leq 1$.

a) Show that $\Delta x^T \Delta s = 0$ and that $\mu(\alpha) := x(\alpha)^T s(\alpha)/n = (1 - \alpha + \alpha\sigma)\mu$.

b) Let $\bar{\alpha} := \max\{\hat{\alpha} \in [0,1] : X(\alpha)S(\alpha)e \geq \gamma\mu(\alpha)e$ for all $\alpha \in [0,\hat{\alpha}]\}$, and let $(x_+, y_+, s_+) := (x(\bar{\alpha}), y(\bar{\alpha}), s(\bar{\alpha}))$. Show that either x_+ is optimal in (P) and (y_+, s_+) in (D), or $x_+ \in \mathcal{F}^0(P)$ and $(y_+, s_+) \in \mathcal{F}^0(D)$, with only the second possibility if $\sigma > 0$.

a) Since we know that $A\Delta x = 0$ we have

$$\begin{aligned} \Delta x^T \Delta s &= \Delta y^T A \Delta x + \Delta s^T \Delta x \\ &= (A^T \Delta y)^T \Delta x + \Delta s^T \Delta x \\ &= (A^T \Delta y + \Delta s)^T \Delta x \\ &= (0)^T \Delta x \\ &= 0. \end{aligned}$$

✓

Note that if we sum the n component-wise equations given in

$$S\Delta x + X\Delta s = \sigma\mu e - XSe$$

we obtain

$$s^T \Delta x + x^T \Delta s = n\sigma\mu - x^T s.$$

Using this and the previous result we obtain

$$\begin{aligned} \mu(\alpha) &:= x(\alpha)^T s(\alpha)/n \\ &= (x + \alpha\Delta x)^T (s + \alpha\Delta s)/n \\ &= \frac{x^T s + \alpha \Delta x^T s + \alpha x^T \Delta s + \alpha^2 \Delta x^T \Delta s}{n} \\ &= \mu + \alpha \frac{s^T \Delta x + x^T \Delta s}{n} \\ &= \mu + \alpha \frac{n\sigma\mu - x^T s}{n} \\ &= \mu + \alpha\sigma\mu - \alpha\mu \\ &= (1 - \alpha + \alpha\sigma)\mu. \end{aligned}$$

✓

b) Since $x \in \mathcal{F}^0(P)$ we have

$$Ax_+ = Ax + A\bar{\alpha}\Delta x = b + \bar{\alpha}A\Delta x = b,$$

and since $(y, s) \in \mathcal{F}^0(D)$ we have

$$A^T y_+ + s_+ = A^T y + A^T \bar{\alpha}\Delta y + s + \bar{\alpha}\Delta s = A^T y + s + \bar{\alpha}(A^T \Delta y + \Delta s) = c.$$

Now by part (a) we know $\mu(\bar{\alpha}) = (1 - \bar{\alpha} + \bar{\alpha}\sigma)\mu$, and also $\mu > 0$. Let us consider two cases:

Case 1: Suppose either $\sigma > 0$ or else $\sigma = 0$ and $\bar{\alpha} < 1$. In either case we know that $1 - \bar{\alpha} + \bar{\alpha}\sigma > 0$ and thus $\mu(\bar{\alpha}) > 0$. We are given that $\gamma > 0$, so this implies $X_+S_+ > 0$. Since this is true of $X(\alpha)S(\alpha)$ for any $\alpha \in [0, \bar{\alpha}]$, then we know that $x_+, s_+ > 0$. Therefore $x_+ \in \mathcal{F}^0(P)$ and $(y_+, s_+) \in \mathcal{F}^0(D)$.

Case 2: If the above case does not hold, then it must be true that $\sigma = 0$ and $\bar{\alpha} = 1$. This implies that $\mu(\bar{\alpha}) = \mu(1) = (1 - 1 + 1 \cdot 0)\mu = 0$. So in this case $X_+S_+ \geq 0$ and by the same reasoning as in Case 1 we have $x_+, s_+ \geq 0$, which means $x_+ \in \mathcal{F}(P)$ and $(y_+, s_+) \in \mathcal{F}(D)$. Furthermore, since $\sigma = 0$, then by the equality constraint

$$S\Delta x + X\Delta s = \sigma\mu e - XSe$$

we have $s_j \Delta x_j + x_j \Delta s_j = -x_j s_j$ for all j. So in particular ✓

$$\begin{aligned} x_{+_j} s_{+_j} &= (x_j + \Delta x_j)(s_j + \Delta s_j) \\ &= x_j s_j + s_j \Delta x_j + x_j \Delta s_j + \Delta x_j \Delta s_j \\ &= x_j s_j - x_j s_j \\ &= 0. \end{aligned}$$

This implies $x_+^T s_+ = 0$, and since we had already established feasibility then we have that x_+ is optimal in (P) and (y_+, s_+) in (D).

■