EE 236B HW7
Zhiyuan Cao
304397496
03/02/2017

T5.26

(a)

With inspection of Figure 1, with black circles the level curve, shaded region constraints and red dot feasible point, we can conclude that $p^\star = 1$ and $x^\star = (1, 0)$.

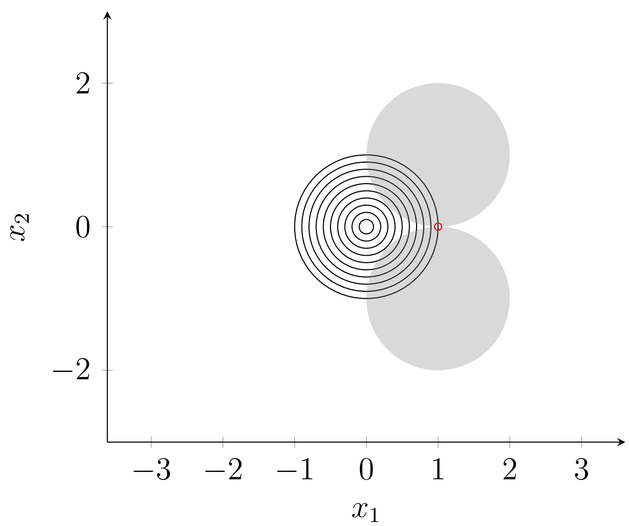

Figure 1: Level curve, constraints and feasible set

(b)

The KKT conditions are:

1. $(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1$, $(x_1 - 1)^2 + (x_2 + 1)^2 \leq 1$,
2. $\lambda_1, \lambda_2 \geq 0$,
3. $\lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 1) = 0$, $\lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 1) = 0$,
4. $(1 + \lambda_1 + \lambda_2)x_1^2 - 2(\lambda_1 + \lambda_2)x_1 = 0$, $(1 + \lambda_1 + \lambda_2)x_2^2 - 2(\lambda_1 - \lambda_2)x_2 = 0$,

where (4) is followed by convex problem. Plug in $(1, 0)$ and we have

$$\lambda_1 = 0; \qquad \lambda_2 = 0; \qquad \lambda_1 + \lambda_2 = 1.$$

Therefore, there is no solution.

(c)

The Lagrangian is

$$\begin{aligned} L(x_1, x_2, \lambda_1, \lambda_2) = x_1^2 + x_2^2 + \lambda_1(x_1^2 + x_2^2 - 2x_1 - 2x_2 + 1) \\ + \lambda_2(x_1^2 + x_2^2 - 2x_1 + 2x_2 + 1) \end{aligned}$$

Taking partial over x_1 and x_2 and we have

$$\begin{aligned} x_1 &= \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2} \\ x_2 &= \frac{\lambda_1 - \lambda_2}{1 + \lambda_1 + \lambda_2} \end{aligned}$$

Since $\lambda_1, \lambda_2 \geq 0$, we always have$1 + \lambda_1 + \lambda_2 \geq 1$, The dual is therefore given by:

$$\begin{aligned} &\textit{maximize} \quad && \frac{-\lambda_1^2 - \lambda_2^2 + 2\lambda_1\lambda_2 + \lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2} \\ &\textit{subject to} \quad && \lambda_1, \lambda_2 \geq 0 \end{aligned}$$

Since $g(\lambda_1, \lambda_2)$ is symmetric, the optimum occurs at $\lambda_1 = \lambda_2$, and therefore we have

$$g(\lambda_1) = \frac{2\lambda_1}{1 + 2\lambda_1}$$

$g(\lambda_1) = 1$ as $\lambda_1 \to \infty$, i.e., $p^\star = d^\star$. However, the dual is not attained.

T5.29

The KKT condition goes as follow

$$x_1^2 + x_2^2 + x_3^2 = 1, \qquad (\nu - 3)x_1 + 1 = 0, \qquad (1 + \nu)x_2 + 1 = 0, \qquad (2 + \nu)x_3 + 1 = 0.$$

This leads to

$$(\frac{1}{3 - \nu})^2 + (-\frac{1}{1 + \nu})^2 + (-\frac{1}{2 + \nu})^2 = 1$$

Solving this equation and we have:

$$\nu = 4.04, \quad \nu = 1.89, \quad \nu = 0.22, \quad \nu = -3.15$$

Which corresponds to

$$\begin{aligned} x = [-0.97 \;\; -0.2 \;\; -0.17]^T \quad x = [-0.9 \;\; -0.35 \;\; -0.26]^T \\ x = [0.36 \;\; -0.82 \;\; -0.45]^T \quad x = [0.16 \; 0.47 \; 0.87] \end{aligned}$$

This corresponds to an objective value of

$$f_0 = -5.37, \quad f_0 = -1.6, \quad f_0 = -1.13, \quad f_0 = -4.65$$

Hence, the pair that corresponds to the optimum is $\nu^\star = 4.04, x^\star = [-0.97 \; -0.2 \; -0.17]^T$.

T5.30

With introduction of Lagrange multiplier $z \in R^n$, the KKT conditions are

1. $X \succ 0$,
2. $Xs = y$,
3. $X^{-1} = I + 1/2(zs^T + sz^T)$,

Combine 2 and 3 we have

$$s = y + \frac{1}{2}(z + sz^T y), \tag{1}$$

Multiplying y^T on both side,

$$1 - y^T y = z^T y$$

Plugging it back to (1),

$$s = y + \frac{1}{2}(z + s - y^T y s) \Rightarrow z = (1 + y^T y)s - 2y$$

Plugging it back to 3 and we have

$$X^{-1} = I + (1 + y^T y)ss^T - ys^T - sy^T.$$

To verify $X^\star = I + yy^T - ss^T/s^T s$,

$$\begin{aligned} X^{-1}X^\star &= (I + (1 + y^T y)ss^T - ys^T - sy^T)(I + yy^T - ss^T/s^T s) \\ &= I - yy^T - ss^T/s^T s + (1 + y^T y)s^T s + (1 + y^T y)sy^T \\ &\quad - (1 + y^T y)s^T s - ys^T - yy^T + ys^T - sy^T - y^T ysy^T + ss^T/s^T s \\ &= I \end{aligned}$$

Still we need to show $X^\star \succ 0$. This can be shown by

$$X^\star = I + yy^T - \frac{ss^T}{s^T s} = (I + \frac{ys^T}{||s^T s||_2} - \frac{ss^T}{s^T s})(I + \frac{ys^T}{||s^T s||_2} - \frac{ss^T}{s^T s})^T \succ 0$$

This is followed by the property of *Cholesky Decomposition*.

A4.10

(a)

Introducing Lagrange multiplier $\nu \in R^n$, and we have

$$\begin{aligned} L(x,\nu) &= x^T A^T A x - 2b^I T A x + b^T b + \sum_{i=1}^{n} \nu_i (x_i - 1) \\ &= x^T (A^T A + \mathbf{diag}(\nu)) x - 2b^T A x + b^T b - \nu^T \mathbf{1} \end{aligned}$$

$L(x,\nu)$ is bounded below if $A^T A + \mathbf{diag}(\nu) \succeq 0$ and $2b^T A \in \text{Range}(A^T A + \mathbf{diag}(\nu))$. Take partial over x and we have $x = (A^T A + \mathbf{diag}(\nu))^\dagger A^T b$ and

$$g(\nu) = -b^T A (A^T A + \mathbf{diag}(\nu))^\dagger A^T b + b^T b - \nu^T \mathbf{1}$$

With introducing $t \in R$, the dual problem is therefore equivalent to

$$\begin{array}{ll} \textit{maximize} & b^T b - t - \nu^T \mathbf{1} \\ \textit{subject to} & \begin{bmatrix} A^T A + \mathbf{diag}(\nu) & -A^T b \\ -b^T A & t \end{bmatrix} \succeq 0 \end{array}$$

With $t,\ \nu$ being the variables.

(b)

First we write it into minimization form

$$\begin{array}{ll} \textit{minimize} & t + \nu^T \mathbf{1} - b^T b \\ \textit{subject to} & \begin{bmatrix} A^T A + \mathbf{diag}(\nu) & -A^T b \\ -b^T A & t \end{bmatrix} \succeq 0 \end{array}$$

With introduction of Lagrange multiplier

$$\begin{bmatrix} Z & z \\ z^T & \lambda \end{bmatrix},$$

The Lagrangian can be written as:

$$\begin{aligned} L(t,\nu,Z,z,\lambda) &= t + \mathbf{1}^T \nu - b^T b - \mathbf{tr}(Z(A^T A + \mathbf{diag}(\nu))) + 2z^T A^T b - t\lambda \\ &= (1-\lambda)t + (\mathbf{1} - \mathbf{diag}(Z))^T \nu - \mathbf{tr}(Z A^T A) + 2b^T A z - b^T b \\ &= \begin{cases} -\mathbf{tr}(Z A^T A) + 2b^T A z - b^T b, & \text{if } \mathbf{diag}(Z) = \mathbf{1}, \lambda = 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Writing the dual function in minimization form and we have

$$\begin{array}{ll} \textit{minimize} & \mathbf{tr}(ZA^TA) - 2b^TAz + b^Tb \\ \textit{subject to} & \mathbf{diag}(Z) = \mathbf{1} \\ & \begin{bmatrix} Z & z \\ z^T & 1 \end{bmatrix} \succeq 0 \end{array}$$

To see this is a relaxation of the original problem. First we have

$$\begin{aligned} ||Ax - b||_2^2 &= x^TA^TAx - 2b^TAx + b^Tb \\ &= \mathbf{tr}(x^TA^TAx) - 2b^TAx + b^Tb \\ &= \mathbf{tr}(A^TAxx^T) - 2b^TAx + b^Tb \end{aligned}$$

This problem is therefore equivalent to

$$\begin{array}{ll} \textit{minimize} & \mathbf{tr}(ZA^TA) - 2b^TAz + b^Tb \\ \textit{subject to} & \mathbf{diag}(Z) = \mathbf{1} \\ & Z = zz^T \end{array}$$

This is replaced by a weaker constraint $Z \succeq zz^T$ in the SDP and therefore it is a relaxation of the original problem. When

$$\text{rank}(\begin{bmatrix} Z & z \\ z^T & 1 \end{bmatrix}) = 1$$

We have

$$[Z \ z] = q[z \ 1]$$

and obviously a solution of $q = z$ and we have $Z = zz^T$.

(c)

$$\begin{aligned} \mathbf{E}||A\nu - b||_2^2 &= \mathbf{E}[\nu^TA^TA\nu - 2b^TA\nu + b^Tb] \\ &= \mathbf{E}[\nu^TA^TA\nu] - \mathbf{E}[2b^TA\nu] + b^Tb \\ &= \mathbf{tr}(\mathbf{E}[\nu\nu^T]A^TA) - 2b^TA\mathbf{E}\nu + b^Tb \end{aligned}$$

Therefore, the equivalence is followed by $Z = \mathbf{E}[\nu\nu^T]$ and $z = \mathbf{E}\nu$.

(d)

The optimal values are as followed:

s	$f(x_a)$	$f(x_b)$	$f(x_c)$	$f(x_d)$	$d^\star$
0.5	4.1623	4.1623	4.1623	4.1623	4.0524
1.0	12.7299	8.3245	8.3245	8.3245	7.8678
2.0	30.1419	16.6490	16.6490	16.6490	15.1814
3.0	33.9339	25.9555	25.9555	24.9735	22.1139

See matlab code as follow.

```
clc;clear;
m = 50;
n = 40;

for i=1:4
    randn('state',0)
    s = [.5 1 2 3];
    A = randn(m,n);
    xhat = sign(randn(n,1));
    b = A*xhat + s(i)*randn(m,1);
    cvx_begin sdp
        variable z(n);
        variable Z(n,n) symmetric;
        minimize (trace(A'*A*Z)-2*b'*A*z + b'*b)
        subject to
            [Z z; z' 1] >= 0
            diag(Z) == 1;
    cvx_end
    clc;
    % x_a
    f1(i) = norm(A*sign((A\b))-b);
    % x_b
    f2(i) = norm(A*sign(z)-b);
    % x_c
    [v, ~] = eig([Z z; z' 1]);
    f3(i) = norm(A*sign(v(1:n, n+1))-b);
    % x_d
    f4(i) = 1000;
    for j = 1:100
        f4_temp = norm(A*sign(mvnrnd(z,Z-z*z')')-b);
        if (f4_temp <= f4(i))
            f4(i) = f4_temp;
        end
    end
```

```
    % dual
    dual(i) = sqrt(trace(A'*A*Z)-2*b'*A*z + b'*b);
end
```

A4.14

(a)

The KKT conditions are

- $x \succeq 0$, $\mathbf{1}^T x = 1$,
- $(\nu - a_i/a^T x - b_i/b^T x)x_i = 0$, for $i = 1, ..., n$,
- $\nu \mathbf{1} \geq a/a^T x + b/b^T x$

From complimentary slackness, for $x = (1/2, 0, ..., 0, 1/2)$, $\nu = 2$. To see optimality, $x = (1/2, 0, ..., 0, 1/2)$, $\nu = 2$ satisfies the first and second condition of KKT. To see it satisfies the last inequality,

$$\begin{aligned}
&\nu = 2 \geq a_i/a^T x + b_i/b^T x = \frac{2a_i}{a_1 + a_n} + \frac{2b_i}{b_1 + b_n} = 2\frac{a_i + a_1 a_n/a_i}{a_1 + a_n} \\
&\Rightarrow a_1(a_i - a_n) \geq a_i(a_i - a_n)
\end{aligned}$$

This is follow by $a_n \leq a_i \leq a_1$ for $i = 1, ..., n$. Therefore $x = (1/2, 0, ..., 0, 1/2)$ is indeed optimal.

(b)

$$\begin{aligned}
\log(2(u^T A u)^{1/2}(u^T A^{-1} u)^{1/2}) &= \log 2 + \frac{1}{2}\log(z^T \Lambda z) + \frac{1}{2}\log(z^T \Lambda^{-1} z) \\
&\leq \log 2 + \log(\frac{1}{2}a_1 + \frac{1}{2}a_n) + \log(\frac{1}{2}a_1^{-1} + \frac{1}{2}a_n^{-1}) \\
&= \log 2 + \frac{1}{2}\log(2 + \frac{a_1}{a_n} + \frac{a_n}{a_1}) \\
&= \log(\sqrt{\frac{a_n}{a_1}} + \sqrt{\frac{a_1}{a_n}}),
\end{aligned}$$

with first equality followed by eigendecomposition $u^T A u = u^T Q \Lambda Q^T u$ and let $z = Q^T u$ and second inequality followed by letting $a_k = \lambda_k$. Finally, since taking the exponential on both sides does not change the inequality, the *Kantorovich inequality* holds.

A4.17

(a)

Using eigendecomposition, $A = Q\Lambda Q^T$ and $\mathbf{tr}(AX)$=$\mathbf{tr}(\Lambda Q^T XQ)$. Set $Y = Q^T XQ$ and we have,

$$\begin{array}{ll} \textit{maximize} & \sum_{i=1}^{n} Y_{ii}\lambda_i \\ \textit{subject to} & \sum_{i=1}^{n} Y_{ii} = r \\ & 0 \le Y_{ii} \le 1, \qquad i = 1, ..., n \end{array}$$

This is followed by trace only involves the diagonal elements of matrix Y. Through inspection, the optimal value of this SDP is equal to $f(A)$ by taking $Y_{ii} = 1$ for $i = 1, ..., r$ and 0 otherwise.

(b)

$f(A)$ is convex since it is the pointwise supremum of a family of linear function.

(c)

$$\begin{aligned} L(X, \nu, U, V) &= -\mathbf{tr}(AX) + \nu(\mathbf{tr}X - r) - \mathbf{tr}(UX) + \mathbf{tr}(V(X - I)) \\ &= -\mathbf{tr}((A - \nu I - U + V)X) - r\nu - \mathbf{tr}V \\ &= \begin{cases} -r\nu - \mathbf{tr}V & \text{if } -A - \nu I - U + V = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Therefore, the dual function

$$\begin{array}{ll} \textit{maximize} & -r\nu - \mathbf{tr}V \\ \textit{subject to} & A + \nu I \preceq V \\ & V \succeq 0, \end{array}$$

or in minimization form,

$$\begin{array}{ll} \textit{minimize} & r\nu + \mathbf{tr}V \\ \textit{subject to} & A + \nu I \preceq V \\ & V \succeq 0, \end{array}$$

Through strong duality, this optimal value of this SDP equals to $f(A(x))$.