 LO, Li-yu 20997405 elloac@connect.ust.hk P1

21/21

5.5 Find the dual function of the LP

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Gx \preceq h \quad \leadsto \lambda \\ & Ax = b \quad \leadsto \nu \end{aligned}$$

a give dual problem b make implicit equality constraints explicit

sol^n

△ given an optimization problem in standard form

$$\begin{aligned} \text{minimize} \quad & f_0(x) \quad x \in R^n \\ \text{subject to} \quad & f_i(x) \le 0, \quad i = 1 \dots m \quad \leadsto \lambda \in R^m \\ & h_i(x) = 0, \quad i = 1 \dots p \quad \leadsto \nu \in R^p \end{aligned}$$

Lagrangian $L(x, \lambda, \nu)$ is defined as

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

Lagrange dual function

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu) = \inf_{x \in D} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

a) give dual problem

△ original LP reformulate as:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Gx - h \preceq 0 \\ & Ax - b = 0 \end{aligned}$$

△ $L(x, \lambda, \nu) = c^T x + \lambda (Gx - h) + \nu (Ax - b)$

△
$$\begin{aligned} g(\lambda, \nu) &= \inf_x L(x, \lambda, \nu) \\ &= \inf_x \left(c^T x + \lambda^T (Gx - h) + \nu^T (Ax - b) \right) \\ &= \inf_x \left(\underline{c^T x} + \underline{\lambda^T G x} - \lambda^T h + \underline{\nu^T A x} - \nu^T b \right) \\ &= \inf_x \left((c^T + \lambda^T G + \nu^T A) x - \lambda^T h - \nu^T b \right) \end{aligned}$$

$$\therefore g(\lambda, \nu) = \begin{cases} -\lambda^T h - \nu^T b & c^T + \lambda^T G + \nu^T A = 0 \\ -\infty & \text{o.w.} \end{cases}$$

(uninformative)

△ maximize $g(\lambda, \nu)$ where $g(\lambda, \nu) = \begin{cases} -\lambda^T h - \nu^T b & c^T + \lambda^T G + \nu^T A = \\ -\infty & \text{o.w.} \end{cases}$

$\lambda \succeq 0$

※

5.5 (cont'd)

soln (cont'd)

b) make implicit equality constraints explicit

$$\begin{array}{ll} \text{maximize} & -\lambda^T h - \nu^T b \\ \text{subject to} & c + G^T\lambda + A^T\nu = 0 \\ & \lambda \succeq 0 \end{array}$$

5.6.

(a) prove $\|Ax_{ls} - b\|_\infty \le \sqrt{m}\,\|Ax_{ch} - b\|_\infty$

soln

- from hint:

$$\frac{1}{\sqrt{m}}\|z\|_2 \le \|z\|_\infty \le \|z\|_2$$

$\therefore \frac{1}{\sqrt{m}}\|Ax_{ls}-b\|_\infty \le \frac{1}{\sqrt{m}}\|Ax_{ls}-b\|_2$

$\Downarrow$ as x_{ls} is solved closed-form

$\frac{1}{\sqrt{m}}\|Ax_{ls}-b\|_2 \le \frac{1}{\sqrt{m}}\|Ax_{ch}-b\|_2$

$\Downarrow$ from hint

$\frac{1}{\sqrt{m}}\|Ax_{ch}-b\|_2 \le \|Ax_{ch}-b\|_\infty$

- $\therefore \frac{1}{\sqrt{m}}\|Ax_{ls}-b\|_\infty \le \frac{1}{\sqrt{m}}\|Ax_{ls}-b\|_2 \le \frac{1}{\sqrt{m}}\|Ax_{ch}-b\|_2 \le \|Ax_{ch}-b\|_\infty$

$\therefore \|Ax_{ls}-b\|_\infty \le \sqrt{m}\,\|Ax_{ch}-b\|_\infty$

(b)

- dual problem

$$\begin{array}{ll} \text{maximize} & b^T\nu \\ \text{subject to} & \|\nu\|_1 \le 1 \\ & A^T\nu = 0 \end{array}$$

- show that $\hat{\nu} = -r_{ls}/\|r_{ls}\|_1$ & $\tilde{\nu} = r_{ls}/\|r_{ls}\|_1$ are both feasible

5-6 (cont'd)

(b) (cont'd)

soln (cont'd)

△ $\|\hat{\upsilon}\| = 1$, $\|\tilde{\upsilon}\| = 1$ — ①

△ $r_{es} = b - A x_{ls}$

$x_{ls} = (A^T A)^{-1} A^T b$

$\therefore A^T \hat{\upsilon} = A^T \left(\frac{-r_{es}}{\|r_{es}\|_1} \right)$, $A^T \tilde{\upsilon} = A^T \left(\frac{r_{es}}{\|r_{es}\|_1} \right)$

$$\begin{aligned} A^T r_{es} &= A^T (b - A x_{ls}) \\ &= A^T (b - A (A^T A)^{-1} A^T b) \\ &= A^T b - A^T A (A^T A)^{-1} A^T b \\ &= A^T b - I A^T b \\ &= 0 \end{aligned}$$

$\therefore A^T \hat{\upsilon} = A^T \tilde{\upsilon} = 0$

both feasible ✡

△ which is a better bound ?

$$b^T \hat{\upsilon} = - b^T \frac{r_{es}}{\|r_{es}\|_1}$$

$$= - \frac{(b - A x_{ls})^T r_{es}}{\|r_{es}\|_1}$$

$$= - \frac{\|r_{es}\|_2^2}{\|r_{es}\|_1}$$

$$b^T \tilde{\upsilon} = b^T \frac{r_{es}}{\|r_{es}\|_1}$$

$$= \frac{(b - A x_{ls})^T r_{es}}{\|r_{es}\|_1}$$

$$= \frac{\|r_{es}\|_2^2}{\|r_{es}\|_1}$$

$b^T \tilde{\upsilon} > b^T \hat{\upsilon}$

$\therefore \tilde{\upsilon}$ is a better bound

✡

△ compare to (a)

$$\tilde{\upsilon} \geq \frac{1}{\sqrt{m}} \| A x_{ls} - b \|_\infty$$

$$\Rightarrow \frac{\|r_{es}\|_2^2}{\|r_{es}\|_1} \geq \frac{1}{\sqrt{m}} \| r_{es} \|_\infty$$

$$\Rightarrow \|r_{es}\|_2^2 \geq \frac{1}{\sqrt{m}} \|r_{es}\|_\infty \|r_{es}\|_1$$

which is true, as

$\|r_{es}\|_2 \geq \frac{1}{\sqrt{m}} \|r_{es}\|_1$

$\|r_{es}\|_2 \geq \|r_{es}\|_\infty$

$\therefore \tilde{\upsilon}$ is a better bound than $\frac{1}{\sqrt{m}} \| A x_{ls} - b \|_\infty$ from (a)

✡

5.7

consider the convex piecewise-linear minimization problem

$$\text{minimize} \quad \max_{i=1\ldots m} (a_i^T x + b_i) \qquad x \in R^n$$

sol^n

(a) derive a dual problem based on Lagrange dual of

$$\begin{aligned} &\text{minimize} && \max_{i=1,\ldots m} y_i \\ &\text{subject to} && a_i^T x + b_i = y_i, \quad i=1\ldots m \end{aligned}$$

△ $L(x, y, \nu) = \max_{i=1\ldots m} y_i + \sum_{i=1}^{m} \nu_i (a_i^T x + b_i - y_i)$

△ $g(\nu) = \inf_{x,y} \left(\max_{i=1\ldots m} y_i + \sum_{i=1}^{m} \nu_i (a_i^T x + b_i - y_i) \right)$

△ to have infimum over $L(x, y, \nu)$, from above, it could be seen we either let terms related to x, y $\begin{cases} 0 \\ -\infty \end{cases}$ to have infimum.

yet $-\infty$, unformative, so we try to set up constraints to let x-, y- related term $= 0$

△ x-related term

$\Rightarrow \inf_x \left(\sum_{i=1}^{m} \nu_i a_i^T x \right)$

$\Rightarrow \begin{cases} 0, & \sum_{i=1}^{m} \nu_i a_i = 0 \\ -\infty, & \text{o.w.} \end{cases}$

$\Rightarrow \begin{cases} 0, & A^T \nu = 0 \\ -\infty, & \text{o.w} \end{cases}$ ------- ①

△ y-related term

$\Rightarrow \inf_y \left(\max_{i=1\ldots m} y_i - \sum_{i=1}^{m} \nu_i y_i \right)$

$\Rightarrow \begin{cases} 0, & \sum_{i=1}^{m} \nu_i = 1 \text{ \& all } \nu_i \geq 0 \\ -\infty, & \text{o.w.} \end{cases}$ ⋯ ②

(as if $\sum_{i=1}^{m} \nu_i > 1$ or $\sum_{i=1}^{m} \nu_i < 1$ the infimum can easily go to $-\infty$)

(if any $\nu_i < 0$, it also leads to $-\infty$)

△ from ① & ② we have condition (constraints)

$$A^T \nu = 0, \ 1^T \nu = 1, \ \nu \succeq 0$$

$\therefore \inf_{x,y} L(x, y, \nu) = b^T \nu$

$\therefore$ dual problem :

$$\begin{aligned} &\underset{\nu}{\text{maximize}} && b^T \nu \\ &\text{subject to} && \nu \succeq 0 \\ & && 1^T \nu = 1 \\ & && A^T \nu = 0 \end{aligned}$$

✗

5.7 (cont'd)

sol^n (cont'd)

(b) formulate original problem as LP

Δ $\underset{x}{\text{minimize}} \quad \max_{i=1\ldots m} (a_i^T x + b_i)$

$\Rightarrow \underset{x,t}{\text{minimize}} \quad t$

subject to $a_1^T x + b_1 \leq t$

$a_2^T x + b_2 \leq t$

$\vdots$

$a_m^T x + b_m \leq t$

$\Rightarrow \underset{x,t}{\text{minimize}} \quad t$

subject to $Ax + b \preceq t\mathbf{1}$

Δ $L(x, t, \lambda) = t + \lambda^T (Ax + b - t\mathbf{1})$

Δ $g(\lambda) = \inf_{x,t} (t + \lambda^T (Ax + b - t\mathbf{1}))$

to have infimum over $L(x, t, \lambda)$,
from above,
we either let terms related to x, t $\begin{cases} 0 \\ -\infty \end{cases}$ to have infimum
yet $-\infty$ is unformative, so we try to set up constraints
to let x-, t-related term $= 0$

Δ x-related term

$\Rightarrow \inf_x (\lambda A x)$

$\Rightarrow \begin{cases} 0, & A^T\lambda = 0 \\ -\infty, & \text{o.w.} \end{cases}$ ...①

Δ t-related term

$\Rightarrow \inf_t (t - \lambda^T t \mathbf{1})$

$\Rightarrow \begin{cases} 0, & \mathbf{1}^T\lambda = 0 \\ -\infty, & \text{o.w.} \end{cases}$, all $\lambda_i \geq 0$ ...②

Δ $\therefore \inf_{x,t} (t + \lambda^T (Ax - b - t\mathbf{1}))$

$\Rightarrow \begin{cases} b^T\lambda, & A^T\lambda = 0, \ \mathbf{1}^T\lambda = 0, \ \lambda \succeq 0 \\ -\infty \end{cases}$ (from ① & ②)

$\therefore$ dual problem

$\underset{\lambda}{\text{maximize}} \quad b^T\lambda$

subject to $A^T\lambda = 0$

$\mathbf{1}^T\lambda = 0$

$\lambda \succeq 0$ #

Δ equivalent to dual problem in 5.7 (a) #

5.7 (cont'd)
soln (cont'd)

(c) approximate minimize $\max_{i=1\ldots m} (a_i^T x + b_i)$

as

minimize $\log \sum_{i=1}^m \exp(a_i^T x + b_i)$

△ dual problem of new formulation:

$$\begin{aligned} \text{minimize} \quad & b^T \nu - \sum_{i=1}^m \nu_i \log \nu_i \\ \text{subject to} \quad & 1^T \nu = 1 \\ & A^T \nu = 0 \\ & \nu \succeq 0 \end{aligned} \qquad \text{——} \quad ① \quad (gP)$$

△ recall (a) (c)

dual problem was:

$$\begin{aligned} \text{minimize} \quad & b^T \nu \\ \text{subject to} \quad & 1^T \nu = 1 \\ & A^T \nu = 0 \\ & \nu \succeq 0 \end{aligned} \qquad \text{——} \quad ② \quad (pwl)$$

△ from ① and ② new formulation has a lower bound in ①

let d^*_{gP} and d^*_{pwl} be the dual optimal, respectively.

$\Rightarrow \quad d^*_{gP} \leq d^*_{pwl}$

△ from ①, gP & from 5.62,

$$\begin{cases} d^* = b^T \nu^* - \sum_{i=1}^m \nu_i^* \log \nu_i^* \\ d^* = p^* \end{cases}$$

$\Rightarrow \quad p^*_{gP} = b^T \nu^* - \sum_{i=1}^m \nu_i^* \log \nu_i^*$

$\Rightarrow \quad b^T \nu^* = p^*_{gP} + \sum_{i=1}^m \nu_i^* \log \nu_i^*$

$\Rightarrow \quad d^*_{pwl} = p^*_{gP} + \sum_{i=1}^m \nu_i^* \log \nu_i^*$

5.7 (cont'd)

soln (cont'd)

(c) (cont'd)

$\log \nu + 1 = 0.$

$\triangle$ for any primal optimal value & dual optimal value

$p^* \geq d^*$

$\nu_i^* \log \nu_i$

$\therefore P^*_{pwl} \geq d^*_{pwl}$

$\Rightarrow P^*_{pwl} \geq P^*_{gp} + \sum_{i=1}^{m} \nu_i^* \log \nu_i^*$ — ③

$\delta(\nu)$

$\frac{\partial}{\partial \nu} \delta(\nu) = 0.$

$\triangle \quad P^*_{gp} + \sum_{i=1}^{m} \nu_i^* \log \nu_i^* \geq P^*_{gp} + \inf_{1^T \nu = 1} \left(\sum_{i=1}^{m} \nu_i^* \log \nu_i^* \right)$

$= P^*_{gp} + \log \frac{1}{m}$ — ④

$\log \nu + \nu \cdot \frac{1}{\nu} = 0$

$\log \nu + 1 = 0.$

$\log x + 1 = 0$

$\log x = -1$

$x = e^{-1}$

$\log \nu = -1.$

$\triangle$ from ③ & ④

$P^*_{pwl} \geq P^*_{gp} + \log \frac{1}{m}$ — ⑤

$\triangle$ from ② & ⑤

$P^*_{gp} \geq P^*_{pwl}$ as $\log \frac{1}{m} < 0$

$\therefore P^*_{gp} - P^*_{pwl} \geq 0$

& $P^*_{pwl} \geq P^*_{gp} - \log m$

$\Rightarrow \log m \geq P^*_{gp} - P^*_{pwl}$

$\therefore 0 \leq P^*_{gp} - P^*_{pwl} \leq \log m$ #

5.7 (cont'd)

soln (cont'd)

(d) consider the below :

$$\text{minimize} \quad \left(\frac{1}{r}\right) \log\left(\sum_{i=1}^{m} \exp\left(r(a_i^T x + b_i)\right)\right)$$

derive $(p^*_{gpr} - p^*_{pwl})$'s bound as in (c)

Δ original problem

$\Rightarrow$ minimize $\left(\frac{1}{r}\right) \log\left(\sum_{i=1}^{m} \exp(r y_i)\right)$

subject to $a_i^T x + b_i = y_i \quad (i=1,\dots m)$ ~~~ ν

$$\Rightarrow L(x, y, \nu) = \frac{1}{r} \log\left(\sum_{i=1}^{m} \exp(r y_i)\right) + \nu^T (Ax + b - y)$$

Δ $g(\nu) = \inf_{x,y} L(x, y, \nu)$

$\Rightarrow$ x-, y-related either $\begin{cases} 0 \\ -\infty \end{cases}$

Δ x-related term

$$\inf_x \nu^T (Ax) = \begin{cases} 0, & A^T \nu = 0 \\ -\infty, & \text{o.w.} \end{cases} \quad \text{— ①}$$

Δ y-related term

$$\inf_y \left(\frac{1}{r} \log\left(\sum_{i=1}^{m} \exp(r y_i)\right) - \nu^T y\right) = \begin{cases} Y, & \text{constraint } C \\ -\infty, & \text{o.w.} \end{cases}$$

Δ y-related term, constraint C with Y, is as :

$$\frac{\partial}{\partial y}\left(\frac{1}{r} \log\left(\sum_{i=1}^{m} \exp(r y_i)\right) - \nu^T y\right) = 0$$

$$\Rightarrow \cancel{r} \cdot \frac{1}{\cancel{r}} \sum_{i=1}^{m} \left(\frac{\exp(r y_i)}{\sum_{i=1}^{m} \exp(r y_i)}\right) - \nu = 0$$

$\therefore$ for y-related term to have infimum

$$\frac{\exp(r y_i)}{\sum_{i=1}^{m} \exp(r y_i)} - \nu_i = 0 \quad \text{— ②}$$

$$\therefore \sum_{i=1}^{m} \frac{\exp(r y_i)}{\sum_{i=1}^{m} \exp(r y_i)} = \sum_{i=1}^{m} \nu_i$$

$$\Rightarrow \sum_{i=1}^{m} \frac{\exp(r y_i)}{\sum_{i=1}^{m} (\exp(r y_i))} = 1 = \sum_{i=1}^{m} \nu_i$$

5.7 (cont'd)

soln (cont'd)

(d) (cont'd)

$$\therefore \inf_y \left(\frac{1}{r} \log \left(\sum_{i=1}^m \exp(r y_i) \right) - \nu^T y \right)$$

$$\Rightarrow \inf_y \left(\frac{1}{r} \log \left(\sum_{i=1}^m \exp(r y_i) \right) - \frac{1}{r} \log \exp \left(\sum_{i=1}^m \nu_i y_i r \right) \right)$$

$$\Rightarrow \inf_y \frac{1}{r} \log \left(\frac{\sum_{i=1}^m \exp(r y_i)}{\exp\left(\sum_{i=1}^m \nu_i y_i r\right)} \right)$$

$$\Rightarrow \inf_y \frac{-1}{r} \log \left(\frac{\exp\left(\sum_{i=1}^m \nu_i y_i r\right)}{\sum_{i=1}^m \exp(r y_i)} \right)$$

$$\Rightarrow \inf_y \frac{-1}{r} \log \frac{\left(e^{\nu_1 y_1 r} \cdot e^{\nu_2 y_2 r} \cdots e^{\nu_m y_m r}\right) \cdot \left(e^{r y_1} + e^{r y_2} + \cdots + e^{r y_m}\right)^{\nu_1} \cdots \left(e^{r y_1} + e^{r y_2} + \cdots + e^{r y_m}\right)^{\nu_m}}{\left(e^{r y_1} + e^{r y_2} + \cdots + e^{r y_m}\right)^{\nu_1} \cdots \left(e^{r y_1} + e^{r y_2} + \cdots + e^{r y_m}\right)^{\nu_m} \left(e^{r y_1} + e^{r y_2} + \cdots + e^{r y_m}\right)}$$

as $\frac{\exp(r y_i)}{\sum_{i=1}^m \exp(r y_i)} - \nu_i = 0$ for above to hv infinum

$$\Rightarrow \frac{-1}{r} \log \nu_1^{\nu_1} \cdot \nu_2^{\nu_2} \cdots \nu_m^{\nu_m} \cdot \frac{\left[\sum_{i=1}^m \exp(r y_i)\right]^{\nu_1 + \cdots + \nu_m}}{\sum_{i=1}^m \exp(r y_i)}$$

(from ② $\nu_1 + \cdots + \nu_m = 1$)

$$= 1$$

$$\Rightarrow \frac{-1}{r} \log \nu_1^{\nu_1} \cdot \nu_2^{\nu_2} \cdots \nu_m^{\nu_m}$$

$$\Rightarrow -\frac{1}{r} \left(\log \nu_1^{\nu_1} + \log \nu_2^{\nu_2} + \cdots + \log \nu_m^{\nu_m} \right)$$

$$= -\frac{1}{r} \sum_{i=1}^m \nu_i \log \nu_i \quad — ③$$

△ Hence, from ① & ③

$$g(\nu) = \inf_{x,y} L(x, y, \nu)$$

$$= \begin{cases} b^T \nu - \frac{1}{r} \sum_{i=1}^m \nu_i \log \nu_i, & \mathbf{1}^T \nu = 1, \ \nu \succeq 0 \\ -\infty & \end{cases}$$

$\therefore$ dual problem is

$$\begin{aligned} \underset{\nu}{\text{maximize}} \quad & b^T \nu - \frac{1}{r} \sum_{i=1}^m \nu_i \log \nu_i \\ \text{subject to} \quad & A^T \nu = 0 \\ & \mathbf{1}^T \nu = 1 \\ & \nu \succeq 0 \end{aligned}$$

5.7 (cont'd)

sol^n (cont'd)

(d) (cont'd)

△ maximize$_{\nu}$ $b^T\nu - \frac{1}{r}\sum_{i=1}^{m} \nu_i \log \nu_i$

subject to $A^T\nu = 0$

$\mathbf{1}^T\nu = 1$

$\nu \succeq 0$

(dual problem of new formulation — #

△ recall (a)

maximize $b^T\nu$

subject to $A^T\nu = 0$

$\mathbf{1}^T\nu = 1$

$\nu \succeq 0$

(original dual problem in (a)) — $(*)$

△ from (c)

$$0 \le p_{gp}^* - p_{pwl}^* \le \log m$$

$$\therefore \; 0 \le p_{gpr}^* - p_{pwl}^* \le \frac{1}{r}\log m$$

△ as for the effect of increasing $\frac{1}{r}$

$\Longrightarrow$ $p_{gpr}^* - p_{pwl}^*$ becomes smaller

$\Longrightarrow$ p_{gpr}^* gets closer to p_{pwl}^*

5.10

D-optimal design

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \log\det\left(\sum_{i=1}^{p} x_i v_i v_i^T\right)^{-1} \\ \text{subject to} & x \succeq 0 \\ & 1^T x = 1 \end{array}$$

- Domain : $\{x \mid \sum_{i=1}^{p} x_i v_i v_i^T \succ 0\}$
- $x \in R^p$, $v_1 \cdots v_p \in R^n$

→ derive dual problem new variable $X \in S^n$. $X = \sum_{i=1}^{p} x_i v_i v_i^T$

soln

△ original problem becomes :

$X - \sum_{i=1}^{p} x_i v_i v_i^T = 0$

$$\begin{array}{ll} \underset{x,X}{\text{minimize}} & \log\det(X^{-1}) \\ \text{subject to} & X = \sum_{i=1}^{p} x_i v_i v_i^T , \; x \succeq 0 , \; 1^T x = 1 \end{array}$$

△ $L(x, Z, \lambda, \nu)$

$$= \log\det(X^{-1}) + \underline{tr(ZX) - \sum_{i=1}^{p} x_i v_i^T Z v_i} \; \underline{- \lambda^T x} + \underline{\nu(1^T x - 1)}$$

$$= \underbrace{\log\det(X^{-1}) + tr(ZX)}_{a} + \underbrace{\sum_{i=1}^{p} x_i(-v_i^T Z v_i - \lambda_i + \nu)}_{b} - \nu$$

△ $g(Z, \lambda, \nu) = \inf_x L(x, Z, \lambda, \nu)$

$\Rightarrow \inf_X L(x, Z, \lambda, \nu)$, acquire through taking the gradient & set to 0

a

$$\frac{\partial}{\partial X}\left(\log\det(X^{-1}) + tr(ZX)\right) = 0$$

$$\Rightarrow \frac{\partial}{\partial X}(-\log\det X + tr(ZX)) = 0$$

$$\Rightarrow -X^{-T} + Z = 0$$

$$\Rightarrow X^{-T} = X^{-1} = Z$$

b

$$\frac{\partial}{\partial x_i} x_i(-v_i^T Z v_i - \lambda_i + \nu) = 0$$

$$\Rightarrow -v_i^T Z v_i - \lambda_i + \nu = 0$$

$$\therefore \inf_X L(X, Z, \lambda, \nu) = \begin{cases} \log\det(Z) + tr(ZZ^{-1}) - \nu & -v_i^T Z v_i - \lambda_i + \nu = \\ -\infty & , \; o.w. \end{cases}$$

5.10 (cont'd)

soln (cont'd)

$\triangle$ dual problem :

$$\begin{array}{ll} \underset{Z, \nu, \lambda}{\text{maximize}} & \log\det Z + \mathrm{tr}(ZZ^{-1}) - \nu \\ \text{subject to} & \nu = \lambda_i + v_i^T Z v_i \\ & i = 1 \dots p \end{array}$$

$$\Rightarrow \begin{array}{ll} \underset{Z, \nu, \lambda}{\text{maximize}} & \log\det Z + n - \nu \\ & \nu = \lambda_i + v_i^T Z v_i \quad i = 1 \dots p \end{array}$$

5.11

$$\text{minimize} \quad \sum_{i=1}^{N} \|A_i x + b_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2$$

derive a dual problem

$A_i \in R^{m_i \times n}$

$b_i \in R^{m_i}$

$x_0 \in R^n$

introduce $y_i = A_i x + b_i$

soln

$\triangle$ original problem turns into :

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{N} \|\overset{R^{m_i}}{y_i}\|_2 + \frac{1}{2}\|x - x_0\|_2^2 \\ \text{subject to} & y_i = A_i x + b_i \quad i = 1, 2 \dots N \end{array}$$

ν

$\triangle$ $L(x, y, \nu) = \sum_{i=1}^{N} \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \nu_i^T (y_i - A_i x - b_i)$

where $\nu \in R^{N \times m_i}$, $\begin{bmatrix} \text{—} \nu_1 \text{—} \\ \text{—} \nu_2 \text{—} \\ \vdots \\ \text{—} \nu_N \text{—} \end{bmatrix}$

$$g(\nu) = \inf_{x, y} L(x, y, \nu)$$

$\triangle$ x-related term

$$\inf_x \left(\frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \nu_i^T A_i x \right)$$

$$\Rightarrow \frac{\partial}{\partial x}\left(\frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \nu_i^T A_i x \right) = 0$$

$$\Rightarrow \frac{1}{2}(2x - 2x_0) + \sum_{i=1}^{N} A_i^T \nu_i$$

$$\Rightarrow x = x_0 - \sum_{i=1}^{N} A_i^T \nu_i$$

$\triangle$ y-related term

$$\inf_y \sum_{i=1}^{N} \|y_i\|_2 - \sum_{i=1}^{N} \nu_i^T y_i$$

$$\Rightarrow \inf_y \sum_{i=1}^{N} (\|y_i\|_2 - \nu_i^T y_i)$$

$$= \begin{cases} 0, & \|\nu_i\|_2 \le 1 \\ -\infty \end{cases}$$

5.11 (cont'd)

sol^n (cont'd)

△ $g(\nu) = \inf_{x,y} L(x,y,\nu)$

$$\left(\begin{aligned} &= \tfrac{1}{2} \| x_0 - \sum_{i=1}^{N} A_i^T \nu_i - x_0 \|_2^2 + \sum_{i=1}^{N} \left(A_i (x_0 - \sum_{i=1}^{N} A_i^T \nu_i) + b_i \right)^T \nu_i \\ &= \tfrac{1}{2} \| \sum_{i=1}^{N} A_i^T \nu_i \|_2^2 + \sum_{i=1}^{N} \left(A_i x_0 + b_i - A_i \sum_{i=1}^{N} A_i^T \nu_i \right)^T \nu_i \end{aligned} \right)$$

neglect $-\infty$

$$= \begin{cases} \sum_{i=1}^{N} (A_i x_0 + b_i)^T \nu_i - \tfrac{1}{2} \| \sum_{i=1}^{N} A_i^T \nu_i \|_2^2 & \|\nu_i\|_2 \le 1 \\ -\infty & , \text{o.w.} \end{cases}$$

△ dual problem :

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{N} (A_i x_0 + b_i)^T \nu_i - \tfrac{1}{2} \| \sum_{i=1}^{N} A_i^T \nu_i \|_2^2 \\ \text{subject to} \quad & \|\nu_i\|_2 \le 1 , \quad i = 1, \dots, N \end{aligned}$$

☆

5.13

(a)

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \preceq b \\ & x_i (1 - x_i) = 0 \quad i = 1 \dots n \end{aligned}$$

Find the Lagrange dual

sol^n

$$\triangle \; L(x, \lambda, \nu) = c^T x + \lambda^T (Ax - b) + \sum_{i=1}^{n} (\nu x_i - \nu x_i x_i)$$

$$\Rightarrow c^T x + \lambda^T (Ax - b) + \sum_{i=1}^{n} (\nu x_i x_i - \nu x_i)$$

equivalent

$$= x^T \operatorname{diag}(\nu) x + (c + A^T \lambda - \nu)^T x - b^T \lambda$$

$$\triangle \; g(\lambda, \nu) = \inf_x \left(x^T \underbrace{\operatorname{diag}(\nu)}_{= \nu'} x + (c^T + A^T \lambda - \nu)^T x - b^T \lambda \right)$$

$$\Rightarrow \text{let} \; \frac{\partial}{\partial x} \left(x^T \nu' x + (c^T + A^T \lambda - \nu)^T x - b^T \lambda \right) = 0$$

$$\Rightarrow 2 \nu' x + (c^T + A^T \lambda - \nu) = 0$$

$$\Rightarrow x = -\tfrac{1}{2} (c^T + A^T \lambda - \nu) \nu'^{-1}$$

assume

$\operatorname{diag}(\nu) \succeq 0$

$\therefore \; \nu \succeq 0$

5.13 (cont'd)
(a) (cont'd)
soln (cont'd)

$\triangle$ when $x = -\frac{1}{2}(c + A^T\mu - \nu)V'^{-1}$

has infinum

$\therefore \; g(\lambda, \nu) = \inf L(x, \lambda, \nu)$

$$= \begin{cases} \left(-\frac{1}{2}(c + A^T\lambda - \nu)\cancel{V'^{-1}}\right)^T \cdot \cancel{V'}\left(-\frac{1}{2}(c + A^T\lambda - \nu)V'^{-1}\right) \\ + (c + A^T\lambda - \nu)^T\left(-\frac{1}{2}(c + A^T\lambda - \nu)V'^{-1}\right) - b^T\lambda & , \nu \succeq 0 \\ -\infty, \; \text{o.w} \end{cases}$$

$$= \begin{cases} -b^T\lambda + \left(\frac{1}{4} - \frac{1}{2}\right)(c + A^T\lambda - \nu)^T(c + A^T\lambda - \nu)V'^{-1} & , \nu \succeq 0 \\ -\infty, \; \text{o.w} \end{cases}$$

$$= \begin{cases} -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n}(c_i + a_i^T\lambda - \nu_i)^2\nu_i^{-1} & \nu \succeq 0 \\ -\infty, \; \text{o.w.} \end{cases}$$

$\triangle$ dual problem :

$$\underset{\lambda, \nu}{\text{maximize}} \quad -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n}(c_i + a_i^T\lambda - \nu_i)^2\nu_i^{-1}$$

subject to $\quad \nu \succeq 0$

$\quad \lambda \succeq 0 \quad$ ※

(b) show that Lagrangian relaxation & LP relaxation

are the same.

hint derive dual of LP relaxation

soln

minimize $\quad c^Tx$

subject to $\quad Ax \preceq b$

$\quad 0 \leq x_i \leq 1, \quad i = 1, \ldots, n$

5.13 (cont'd)

(b) (cont'd)

Sol^n (cont'd)

$\triangle$ $L(x, \lambda_1, \lambda_2, \lambda_3)$

$= c^T x + \lambda_1 (Ax - b) + \lambda_2^T(-x) + \lambda_3^T(x - 1)$

$= (c^T + (A^T\lambda_1)^T - \lambda_2^T + \lambda_3^T)\, x - b^T\lambda_1 - 1^T\lambda_3$

$\triangle$ $\inf_x f(x, \lambda_1, \lambda_2, \lambda_3)$

$$= \begin{cases} -b^T\lambda_1 - 1^T\lambda_3 & c + A^T\lambda_1 - \lambda_2 + \lambda_3 = 0 \\ -\infty & \text{o.w.} \end{cases}$$

$\triangle$ $g(\lambda_1, \lambda_2, \lambda_3)$

$$= \begin{cases} -b^T\lambda_1 - 1^T\lambda_3 & c + A^T\lambda_1 - \lambda_2 + \lambda_3 = 0 \\ -\infty & \text{o.w.} \end{cases}$$

$\therefore$ dual problem

$$\begin{aligned} \underset{\lambda_1, \lambda_2, \lambda_3}{\text{maximize}} \quad & -b^T\lambda_1 - 1^T\lambda_3 \\ \text{subject to} \quad & c + A^T\lambda_1 - \lambda_2 + \lambda_3 = 0 \\ & \lambda_1, \lambda_2, \lambda_3 \succeq 0 \end{aligned}$$

$$\Rightarrow \underset{\lambda_1, \lambda_2}{\text{maximize}} \quad -b^T\lambda_1 - 1^T(-c - A^T\lambda_1 + \lambda_2)$$

$$\lambda_1, \lambda_2 \succeq 0$$

5.13 (cont'd)

(b) (cont'd)

Sol^n (cont'd)

- recall dual problem (a)

$$\underset{\lambda, \nu}{\text{maximize}} \quad -b^T\lambda - \frac{1}{4}\sum_{i=1}^{n}(c_i + a_i^T\lambda - \nu_i)^2 \nu_i^{-1}$$

subject to $\nu \succeq 0$

$\lambda \succeq 0$

$\Rightarrow$ is equivalent to first optimize over ν

$$\sup_{\nu}\left(-\frac{(c_i + a_i^T\lambda - \nu_i)^2}{4\nu_i}, \quad \nu_i \geq 0\right) \quad \text{let } c_i + a_i^T\lambda = \alpha$$

$$\Rightarrow \text{let } \frac{\partial}{\partial \nu_i}\left(-\frac{(\alpha - \nu_i)^2}{4\nu_i}\right) = 0$$

$$\Rightarrow \quad -\frac{\nu_i^2 - \alpha^2}{4\nu_i^2} = 0$$

$$\Rightarrow \quad (\nu_i + \alpha)(\nu_i - \alpha) = 0$$

as $\nu_i \geq 0$

if $\alpha > 0$

$\nu_i = \alpha$

else if $\alpha < 0$

$\nu_i = -\alpha$

$\Rightarrow$ substitute back to

$$\sup_{\nu} -\frac{(\alpha - \nu_i)^2}{4\nu_i} = \begin{cases} c + a_i^T\lambda & \text{if } c_i + a_i^T\lambda \leq 0 \\ 0 & \text{if } c_i + a_i^T\lambda \geq 0 \end{cases}$$

$\therefore$ dual problem (a)

$$\underset{\lambda}{\text{maximize}} \quad -b^T\lambda + \sum_{i=1}^{n} \min(0, c + a_i^T\lambda)$$

subject to $\lambda \succeq 0$ — #

5.13 (cont'd)

(b) (cont'd)

sol^n (cont'd)

△ as for dual problem (b)

$$\begin{array}{ll} \underset{\lambda_1, \lambda_2}{\text{maximize}} & -b^T\lambda_1 - \mathbf{1}^T(-c - A^T\lambda_1 + \lambda_2) \\ \text{subject to} & \lambda_1 \succeq 0 \\ & \lambda_2 \succeq 0 \end{array}$$

$$\Rightarrow \begin{array}{ll} \underset{\lambda_2}{\sup} & -\mathbf{1}^T\lambda_2 \\ & \lambda_2 \succeq 0 \end{array} = 0$$

∴ is equivalent to

$$\begin{array}{ll} \underset{\lambda_1}{\text{maximize}} & -b^T\lambda_1 - \mathbf{1}^T(-c - A^T\lambda_1) \\ \text{subject to} & \lambda_1 \succeq 0 \end{array}$$

$$\Rightarrow \begin{array}{ll} \underset{\lambda_1}{\text{maximize}} & -b^T\lambda_1 + c + A^T\lambda_1 \\ \text{subject to} & \lambda_1 \succeq 0 \end{array} \quad \text{—} \ (*)$$

△ (#) & (*), hence equivalent ∀

5.14

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

→ quadratic penalty method, auxiliary function

$$\phi(x) = f_0(x) + \alpha \|Ax - b\|_2^2$$

$$\tilde{x} = \underset{x}{\operatorname{argmin}} \ \phi(x)$$

→ find dual feasible point

→ find corresponding lower bound

sol^n

5.14 (cont'd)

soln (cont'd)

$\triangle$ from given information

$\rightarrow$ problem

$$\begin{aligned} &\text{minimize} \quad f_0(x) \\ &\text{subject to} \quad Ax=b \end{aligned}$$

$\rightarrow$ quadratic penalty method

$$\phi(x) = f_0(x) + \alpha \|Ax-b\|_2^2$$

$$\tilde{x} = \underset{x}{\operatorname{argmin}} \ \phi(x)$$

$\triangle$ $\nabla\phi(\tilde{x}) = 0$, as $\tilde{x} = \underset{x}{\operatorname{argmin}} \ \phi(x)$

$$\Rightarrow \nabla\phi(\tilde{x}) = \nabla_0 f(\tilde{x}) + 2\alpha A^T(A\tilde{x}-b) = 0 \quad \text{— ①}$$

$\triangle$ $L(x, \nu) = f_0(x) + \nu^T(Ax-b)$

$\rightarrow \inf L(x, \nu) =$.

$$\Rightarrow \nabla L(x,\nu) = 0$$

$$\Rightarrow \nabla f_0(x) + A^T\nu = 0 \quad \text{— ②}$$

$\therefore$ from ① & ②

$$\nabla f_0(\tilde{x}) + 2\alpha A^T(A\tilde{x}-b) = \nabla f_0(x) + A^T\nu \quad \text{— ③}$$

$\therefore$ from ③

$$\nu = 2\alpha(A\tilde{x}-b)$$

$\therefore$ $\inf L(x,\nu)$

$= \inf L(\tilde{x}, \nu)$

$= f_0(\tilde{x}) + \nu^T A\tilde{x} - b$

$\triangle$ dual problem :

$$\underset{\nu}{\text{maximize}} \quad \inf L(x,\nu)$$

$$\Rightarrow \underset{\nu}{\text{maximize}} \quad f_0(\tilde{x}) + \nu^T A\tilde{x} - b$$

5.14 (cont'd)
soln (cont'd)

△ dual problem (cont'd)

$$\Rightarrow \underset{\nu}{\text{maximize}} \quad f_0(\tilde{x}) + 2\alpha(A\tilde{x}-b)^T(A\tilde{x}-b)$$

$$\Rightarrow \underset{\nu}{\text{maximize}} \quad f_0(\tilde{x}) + 2\alpha\|A\tilde{x}-b\|_2^2$$

$\forall\ \nu$ - dual feasible ✡

△ lower bound for primal problem

$$f_0(x) = p^* \geq d^* = \underset{\nu}{\text{maximize}}\ f_0(\tilde{x}) + 2\alpha\|A\tilde{x}-b\|_2^2$$

$$= f_0(\tilde{x}) + 2\alpha\|A\tilde{x}-b\|_2^2 \quad ✡$$

5.17

Robust linear programming

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && \sup_{a\in P_i} a^T x \leq b_i, \quad i=1,\dots,m \end{aligned}$$

$x \in R^n$
$P_i = \{a \mid C_i a \preceq d_i\}$
$c \in R^n$
$C_i \in R^{m_i \times n}$
$d_i \in R^{m_i}$ $\qquad P_i \neq \emptyset$
$b \in R^m$

show the equivalence of above to below:

$$\begin{aligned} &\text{minimize} && c^T x \\ &\text{subject to} && d_i^T z_i \leq b_i \\ & && C_i^T z_i = x \qquad i=1\dots m \\ & && z_i \succeq 0 \end{aligned}$$

$x \in R^n$
$z_i \in R^m$ $\qquad i=1\dots m$

soln

△ minimize $c^T x$
subject to $\sup_{a\in P_i} a^T x \leq b_i$

$\Rightarrow$ minimize $c^T x$
subject to $f_i(x) \leq b_i$ - where $f_i(x) = \sup_{a\in P_i} a^T x$

5.17 (cont'd)
sol^n (cont'd)

△ $f_i(x) = \sup_{a \in P_i} a^T x$

$$\Rightarrow \quad \underset{a}{\text{maximize}} \quad x^T a \quad (a \in R^n)$$
$$\text{subject to} \quad C_i a \preceq d_i \quad i = 1, \dots, m$$

$$\begin{array}{c} C_1 a \preceq d_1 \\ C_2 a \preceq d_2 \\ \vdots \\ C_m a \preceq d_m \end{array} \qquad C_1 \in R^{m_1 \times n},\ C_2 \in R^{m_2 \times n},\ C_m \in R^{m_m \times n},\ a \in R^n,\ d_1 \in R^{m_1},\ d_2 \in R^{m_2},\ d_m \in R^{m_m}$$

△ as indicated by Hint
we derive dual problem:

$$\Rightarrow \quad \underset{a}{\text{minimize}} \quad -x^T a$$
$$\text{subject to} \quad C_i a \preceq d_i \quad i = 1, \dots m$$

for every constraint i, got

$$L(a, \lambda_i) = -x^T a + \lambda_i^T (C_i a - d_i)$$
$$g(\lambda_i) = \inf_a \left((-x + C_i^T \lambda)^T a - d_i^T \lambda_i \right)$$
$$= \begin{cases} -d_i^T \lambda_i, & -x + C_i^T \lambda = 0 \\ -\infty, & \text{o.w.} \end{cases}$$

dual problem:

$$\text{maximize} \; (-d_i^T \lambda_i)$$
$$\text{subject to} \quad C_i^T \lambda_i - x = 0$$
$$\lambda_i \succeq 0$$

$$\Rightarrow \quad \text{minimize} \quad d_i^T \lambda_i$$
$$\text{subject to} \quad C_i^T \lambda - x = 0$$
$$\lambda_i \succeq 0$$

△ $d_i^T \lambda_i$ also satisfies $d_i^T \lambda_i \le b_i$

∴ original becomes:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & d_i^T \lambda_i \le b_i \\ & C_i^T \lambda_i = x \\ & \lambda_i \succeq 0 \end{array} \quad \underset{\text{to}}{\overset{\text{equivalent}}{\Longleftrightarrow}} \quad \begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & d_i^T z_i \le b_i \\ & C_i^T z_i = x \\ & z_i \succeq 0 \end{array}$$

#

5-21

$$\underset{x,y}{\text{minimize}} \quad e^{-x}$$
$$\text{subject to} \quad x^2/y \leq 0$$

$D = \{ (x,y) \mid y > 0 \}$

sol^n

(a)

△ objective function e^{-x} is convex
&
x^2/y is convex, when $y > 0$

∴ it is a convex optimization problem #

△ optimal value p^*

$x^2 \geq 0$, $y > 0$
to satisfy $x^2/y \leq 0$
$x = 0$
$p^* = e^{-0} = 1$ #

(b)

△ $L(x,y,\lambda) = e^{-x} + \lambda \cdot \frac{x^2}{y}$

△ $g(\lambda) = \inf_{x,y} L(x,y,\lambda)$

$$= \begin{cases} 0, & \lambda \geq 0 \\ -\infty, & \lambda < 0 \end{cases}$$

∴ dual problem :

$$\underset{\lambda}{\text{maximize}} \quad 0$$
$$\text{subject to} \quad \lambda \geq 0$$

∴ $d^* = 0$

△ optimal duality gap :

$p^* - d^* = 1 - 0 = 1$ #

(c) Slater's condition

△ recall Slater's condition :

$x \in \text{relint } D$
s.t.
$f_i(x) < 0$, $i = 1 \dots m$
$Ax = b$

△ yet the strict feasibility does not apply in this problem

∴ it does not hold for Slater's condition #

5.21 (cont'd)
soln (cont'd)

(d) minimize e^{-x}
subject to $x^2/y \leq u$

which is the perturbed problem.

→ what is the optimal value $p^*(u)$?
→ verify that

$$p^*(u) \geq p^*(0) - \lambda^* u \quad \text{does not hold}$$

△ $p^*(u) = \inf \{ e^{-x} \mid \exists x, y \in D, \frac{x^2}{y} \leq u \}$

if $u > 0$ $\quad \inf \{ e^{-x} \mid \exists x, y \in D, \frac{x^2}{y} \leq u \}$

$= 0 \quad (\text{as } u \to \infty)$

if $u = 0$, unperturbed

$= 1$

if $u < 0$ it is unfeasible,

hence $= \infty$

$$\begin{cases} 0, & u > 0 \\ 1, & u = 0 \\ \infty, & u < 0 \end{cases}$$

△ $p^*(u) \geq p^*(0) - \lambda^* u$

does not always hold, as

⇒ dual problem of perturbed problem:

maximize$_\lambda$ $\quad 0 - \lambda u$

subject to $\quad \lambda \geq 0$

⇒ from above, if $u > 0$, & $1 > u$

$\lambda = 0$, to have maximum objective function value.

⇒ ∴ $p^*(u) \geq p^*(0) - \lambda^* u$

⇒ $0 \geq 1 - 0 \cdot u$

∴ which does not hold

5.30

→ Derive KKT conditions for

$$\underset{X}{\text{minimize}} \quad \operatorname{tr} X - \log\det X \qquad\qquad X \in S^n \qquad \operatorname{dom} X = S^n_{++}$$
$$\text{subject to} \quad Xs = y \qquad\qquad y \in R^n$$
$$s \in R^n$$
$$s^T y = 1$$

→ verify optimal solution is given by

$$X^* = I + yy^T - \frac{1}{s^T s} ss^T$$

soln

▵ recall KKT conditions :

for any optimization problem, w/ differentiable $f_0, f_1 \dots, f_m, h_1, h_2 \dots h_p$ & strong duality, and let x^*, (λ^*, ν^*) be primal & dual optimal points

we have :

$$f_i(x^*) \le 0 \qquad i = 1 \dots m$$
$$h_i(x^*) = 0 \qquad i = 1 \dots p$$
$$\lambda_i^* \ge 0 \qquad i = 1 \dots m$$
$$\lambda_i^* f_i(x^*) = 0 \qquad i = 1 \dots m$$

$$\nabla L(x^*, \lambda^*, \nu^*) = 0$$

▵ Hence, from the above described problem :

$$Xs = y \quad \text{—— ①}$$

→ for $\nabla L(X^*, \nu^*) = 0$

$$\Rightarrow \quad L(X, \nu) = \operatorname{tr} X - \log\det X + \nu^T(Xs - y)$$
$$= \operatorname{tr} X - \log\det X + \nu^T X s - \nu^T y$$

$$\Rightarrow \quad \nabla_X L(X, \nu) = I - X^{-1} + \frac{\nu s^T + s\nu^T}{2} \quad ☆$$

$$\Rightarrow \quad 0 = I - X^{-1} + \frac{\nu s^T + s\nu^T}{2}$$

$$\therefore \left.\begin{aligned} &X^* s = y \\ &0 = I - X^{*-1} + \frac{\nu^* s^T + s\nu^{*T}}{2} \end{aligned}\right\} \text{KKT conditions} \quad ☆$$

5-30 (cont'd)
Sol^n (cont'd)

$\triangle$ verify $X^* = I + yy^T - \frac{1}{s^Ts} ss^T$

$\rightarrow$ from above :

$$\begin{cases} 0 = I - X^{*-1} + \frac{\nu^* s^T + s\nu^{*T}}{2} \\ X^* s = y \end{cases}$$

$$\Rightarrow \begin{cases} X^{*-1} = I + \frac{\nu^* s^T + s\nu^{*T}}{2} \\ s = X^{*-1} y \end{cases} \quad — ①$$

$$\Rightarrow \begin{cases} X^{*-1} y = y + \frac{\nu^* \overset{1}{s^T y} + s\nu^{*T} y}{2} \\ X^{*-1} y = s \end{cases}$$

$$\Rightarrow \quad s = y + \frac{1}{2}(\nu^* + s\nu^{*T} y) \quad — ②$$

$$\Rightarrow \quad s^T y = y^T y + \frac{1}{2}(\nu^* + s\nu^{*T} y) y$$

$$\Rightarrow \quad 1 - y^T y = \frac{1}{2}(\nu^{*T} y + \nu^{*T} y)$$

$$\Rightarrow \quad 1 - y^T y = \nu^{*T} y$$

we try to get expression of ν^*

$\therefore$ $\nu^{*T} y = 1 - y^T y$ in ② :

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

$$\Rightarrow \quad 2s = 2y + \nu^* + (1 - y^T y) s$$

$$\Rightarrow \quad 2s - 2y - (1 - y^T y) s = \nu^*$$

$$\Rightarrow \quad \nu^* = -2y + (1 + y^T y) s \quad — ③$$

$\therefore$ use ③ in ① can then express X^{*-1}

$$\Rightarrow X^{*-1} = I + \frac{1}{2}\left([-2y + (1 + y^T y)s]s^T + s[-2y + (1 + y^T y)s]^T\right)$$

5-30 (cont'd)

soln (cont'd)

$$\Rightarrow X^{*-1} = I + \frac{1}{2}\left(-2ys^T - 2sy^T + (1+y^Ty)ss^T + (1+y^Ty)ss^T\right)$$

$$= I - ys^T - sy^T + (1+y^Ty)ss^T$$

Δ To show

$$X^* = I + yy^T - \frac{1}{s^Ts}ss^T$$

we verify

$$X^* \cdot X^{*-1} = I$$

$$\therefore \left(I + yy^T - \frac{1}{s^Ts}ss^T\right)\left(I - ys^T - sy^T + (1+y^Ty)ss^T\right)$$

$$= I - \cancel{ys^T} - sy^T + \cancel{(1+y^Ty)ss^T} + \cancel{yy^T} - \cancel{yy^Tys^T} - \cancel{yy^Tsy^T} + \cancel{(1+y^Ty)yy^Tss^T}$$
$$- \cancel{\frac{1}{s^Ts}ss^T} + \cancel{\frac{1}{s^Ts}ss^Tys^T} + \frac{1}{s^Ts}ss^Tsy^T - \frac{1}{s^Ts}ss^T\cancel{(1+y^Ty)}ss^T$$

$$= I$$

$\therefore$

$$X^* = I + yy^T - \frac{1}{s^Ts}ss^T$$

5.31

convex problem

$$\text{minimize} \quad f_0(x)$$
$$\text{subject to} \quad f_i(x) \le 0 \quad , \quad i=1,\dots m$$

and assume that it satisfies KKT condition

$$f_i(x^*) \le 0 \qquad i=1\dots m$$
$$\lambda_i^* \ge 0 \qquad i=1\dots m$$
$$\lambda_i f_i(x^*) = 0 \qquad i=1\dots m$$
$$\nabla f_0(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*) = 0$$

show that $\nabla f_0(x^*)^T (x-x^*) \ge 0$

soln

∆ our objective is to get $\nabla f_0(x^*)^T(x-x^*) \ge 0$ where we try to cook it up from above.

∆ $f_i(x^*) \le 0$

&

from first-order conditions of convex functions

$$f_i(x) \ge f_i(x^*) + \nabla f(x^*)^T (x-x^*)$$

$$\Rightarrow \quad 0 \ge f_i(x) \ge f_i(x^*) + \nabla f_i(x^*)^T (x-x^*) \qquad \text{— ①}$$

∆ as $\lambda_i f_i(x^*) = 0$

and

as $f_i(x) \ge f_i(x^*) + \nabla f_i(x^*)^T(x-x^*)$ from ①

$$\Rightarrow \sum_{i=1}^m \lambda_i^* f_i(x) \ge \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*)^T (x-x^*)$$

$$\Rightarrow \quad 0 \ge 0 + \sum_{i=1}^m \lambda_i^* \nabla f_i(x^*)^T (x-x^*) \qquad \text{— ②}$$

5.31 (cont'd)

soln (cont'd)

△ from KKT condition

$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = 0$$

$$\Rightarrow \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) = -\nabla f_0(x^*) \quad \text{--- (4)}$$

△ from (3) & (4)

$$\Rightarrow 0 \geq 0 - \nabla f_0(x^*)^T (x - x^*)$$

$$\Rightarrow 0 \geq -\nabla f_0(x^*)^T (x - x^*)$$

$$\Rightarrow \nabla f_0(x^*)^T (x - x^*) \geq 0$$

5.39

minimize $x^T W x$

subject to $x_i^2 = 1, \quad i = 1, \dots n$ primal

maximize $-1^T \nu$

subject to $W + \mathrm{diag}(\nu) \succeq 0$ dual

soln

(a) △ show original problem could be expressed as

minimize $\mathrm{tr}(WX)$

subject to $X \succeq 0, \quad \mathrm{rank}\, X = 1$

$X_{ii} = 1, \quad i = 1, \dots n$

△ let $X = xx^T$, xx^T from

$\therefore X_{ii} = (\pm 1)^2 \quad \therefore X_{ii} = 1$

$\therefore X \succeq 0$

△ $x^T W x = \mathrm{tr}(x^T W x) = \mathrm{tr}(W x x^T) = \mathrm{tr}(WX)$

△ ∴ original problem could be cast as LHS reformulation

(b)

→ why does its optimal value gives a lower bound?

△ original problem's set

$$\{ x \mid x \in \mathbb{R}^n, x_i^2 = 1, \ i = 1 \dots n \}$$

△ new SDP problem's set

$$\{ X \mid X \in S_+^n, X_{ii} = 1 \}$$

△ from above, the new set is large, which is more relaxed.

hence, yields a lower bound ✡

→ if an optimal point X^* is rank one,

then $X = xx^T$,

and gives the exact solution ✡

(c) △ maximize $-1^T \nu$

subject to $W + \operatorname{diag}(\nu) \succeq 0$ ① — original dual problem

△ minimize $\operatorname{tr}(WX)$

subject to $X \succeq 0$, rank $X = 1$

$X_{ii} = 1, \ i = 1 \dots n$ ② — new formulation of SDP relaxation

△ given by hint : relate the two via duality

rewrite ①

⇒ minimize $1^T \nu$

subject to $W + \operatorname{diag}(\nu) \succeq 0$

$$\therefore L(\nu, Z) = 1^T \nu - \operatorname{tr}(Z(W + \operatorname{diag}(\nu)))$$

$$= 1^T \nu - \operatorname{tr}(ZW) - \operatorname{tr}(Z \operatorname{diag}(\nu))$$

$$\therefore g(Z) = \inf_\nu L(\nu, Z) = \inf_\nu \left(1^T \nu - \operatorname{tr}(ZW) - \operatorname{tr}(Z \operatorname{diag}(\nu)) \right)$$

$$= \inf_\nu \left(-\operatorname{tr}(ZW) + \sum_{i=1}^n \nu_i - \nu_i Z_{ii} \right)$$

$$= \inf_\nu \left(-\operatorname{tr}(ZW) + \sum_{i=1}^n \nu_i (1 - Z_{ii}) \right)$$

5.39 (cont'd)
(b) (cont'd)
Sol^n (cont'd)

$$g(Z) = \inf_{\nu} \left(-\mathrm{Tr}(ZW) + \sum_{i=1}^{n} \nu_i (1 - Z_{ii}) \right)$$

$$= \begin{cases} -\mathrm{Tr}(ZW) & Z_{ii} = 1 \\ -\infty & \text{o.w.} \end{cases}$$

$\therefore$ dual problem :

maximize $\quad -\mathrm{tr}(ZW)$

subject to $\quad Z \succeq 0$

$\quad Z_{ii} = 1, \; i = 1 \dots, n$

which is equivalent as the problem defined in (a)

$\therefore$ hence gives the relationship between them ✕

$\triangle$ from above SDP problem is equivalent to "dual problem" of "dual problem" of two-way partitioning problem, hence, SDP problem has a lower bound below ✕

5.43

dual prob. of minimize $f^T x$

subject to $\|A_i x + b_i\|_2 \le c_i^T x + d_i, \quad i = 1, \dots, m$

can be expressed as

maximize $\displaystyle\sum_{i=1}^{m} (b_i^T u_i - d_i v_i)$

subject to $\displaystyle\sum_{i=1}^{m} (A_i^T u_i - c_i v_i) + f = 0$ —— ⊕

$\|u_i\|_2 \le v_i, \quad i = 1, \dots m$

$u_i \in \mathbb{R}^{n_i}$, $v_i \in \mathbb{R}$, $A_i \in \mathbb{R}^{n_i \times n}$, $b_i \in \mathbb{R}^{n_i}$, $d_i, c_i \in \mathbb{R}$, $f \in \mathbb{R}^n$

sol^n :

(a)

$\triangle$ as instructed,

introduce $y_i = A_i x + b_i$

$t_i = c_i^T x + d_i$

5.43 (cont'd)

(a) (cont'd)

sol'n (cont'd)

△ original problem becomes :

$$\begin{aligned} \underset{x,y,t}{\text{minimize}} \quad & f^T x \\ \text{subject to} \quad & \|y_i\|_2 \le t_i \\ & y_i = A_i x + b_i \\ & t_i = c_i^T x + d_i \end{aligned} \Bigg\} \quad i=1 \dots m$$

△ $L(x, y, t, \lambda, \nu, \mu)$

$$= f^T x + \sum_{i=1}^m \lambda_i (\|y_i\|_2 - t_i) + \sum_{i=1}^m \nu_i^T (y_i - A_i x - b_i) + \sum_{i=1}^m \mu_i (t_i - c_i^T x - d_i)$$

$$= \left(f^T - \sum_{i=1}^m (A_i^T \nu_i)^T - \sum_{i=1}^m \mu_i c_i^T \right) x \qquad \text{(x-related term)}$$

$$+ \sum_{i=1}^m (\lambda_i \|y_i\|_2 + \nu_i^T y_i) \qquad \text{(y-related term)}$$

$$+ \sum_{i=1}^m (\mu_i - \lambda_i) t_i \qquad \text{(t-related term)}$$

$$+ \sum_{i=1}^m (-\nu_i^T b_i - \mu_i d_i)$$

△ $g(\lambda, \nu, \mu) = \inf_{x,y,t} L(x, y, t, \lambda, \nu, \mu)$

$$\inf_x L_x(\cdots) = \begin{cases} 0, & f - \sum_{i=1}^m (A_i^T \nu_i - \mu_i c_i) = 0 \\ -\infty, & \text{o.w} \end{cases}$$

$$\inf_y L_y(\cdots) = \begin{cases} 0, & \|\nu_i\|_2 \le \lambda_i \\ -\infty, & \text{o.w} \end{cases}$$

$$\inf_t L_t(\cdots) = \begin{cases} 0, & \mu_i - \lambda_i = 0 \\ -\infty, & \text{o.w} \end{cases}$$

$$\therefore \; g(\lambda, \nu, \mu) = \begin{cases} -\sum_{i=1}^m (b_i^T \nu_i + d_i \mu_i) & \begin{cases} f - \sum_{i=1}^m (A_i^T \nu_i - \mu_i c_i) = 0 \\ \|\nu_i\|_2 \le \lambda_i \\ \mu_i - \lambda_i = 0 \end{cases} \\ -\infty & \text{o.w} \end{cases}$$

5.43 (cont'd)
(a) (cont'd)
sol^n (cont'd)

△ dual problem :

$$\underset{\lambda, \nu, \mu}{\text{maximize}} \quad -\sum_{i=1}^{m} (b_i^T \nu_i + d_i \mu_i)$$

$$\text{subject to} \quad f - \sum_{i=1}^{m} (A_i^T \nu_i - \mu_i c_i) = 0$$
$$\|\nu_i\|_2 \le \lambda_i$$
$$\mu_i - \lambda_i = 0$$

f here and f in ㊉ (denoted as $f_\#$) : $f = -f_\#$

∴ it is equivalent to problem (dual) in ㊉ ✕

(b) express SOCP as conic form

$$\text{minimize} \quad f^T x$$
$$\text{subject to} \quad -(\underbrace{A_i x + b_i}_{\mathbb{R}^n}, \underbrace{c_i^T x + d_i}_{\mathbb{R}}) \preceq_{K_i} 0 \quad i=1, \dots, m$$

△ $L(x, \lambda, z) = f^T x - \sum_{i=1}^{m} \lambda_i^T (A_i x + b_i) - \sum_{i=1}^{m} z (c_i^T x + d_i)$

△ $$g(\lambda, z) = \inf_x \left(f^T x - \sum_{i=1}^{m} \lambda_i^T (A_i x + b_i) - \sum_{i=1}^{m} z_i (c_i^T x + d_i) \right)$$

$$= \inf_x \left(\left(f^T - \sum_{i=1}^{m} (A_i^T \lambda_i)^T - \sum_{i=1}^{m} z_i c_i^T \right) x - \sum_{i=1}^{m} \lambda_i^T b_i - \sum_{i=1}^{m} z_i d_i \right)$$

$$= \begin{cases} -\sum_{i=1}^{m} (\lambda_i^T b_i + z_i d_i), & f - \sum_{i=1}^{m} (A_i^T \lambda_i + z_i c_i) = 0 \\ -\infty, & \text{o.w.} \end{cases}$$

△ dual problem :

$$\underset{\lambda, z}{\text{maximize}} \quad -\sum_{i=1}^{m} (b_i^T \lambda_i + z_i d_i)$$

$$\text{subject to} \quad f - \sum_{i=1}^{m} (A_i^T \lambda_i + z_i c_i) = 0$$
$$(\lambda_i, z_i) \succeq_{K_i^*} 0$$

→ as suggested by the hint, SOC is self-dual

5.43 (cont'd)
(b) (cont'd)
sol^n (cont'd)

$$\Rightarrow \quad \underset{\lambda, z}{\text{maximize}} \quad -\sum_{i=1}^{m} (b_i^T \lambda_i + z_i d_i)$$
$$\text{subject to} \quad f - \sum_{i=1}^{m} (A_i^T \lambda_i + z_i c_i) = 0$$
$$(\lambda_i, z_i) \succeq_{K_i} 0$$

$$\Rightarrow \quad \underset{\lambda, z}{\text{maximize}} \quad -\sum_{i=1}^{m} (b_i^T \lambda_i + z_i d_i)$$
$$\text{subject to} \quad f - \sum_{i=1}^{m} (A_i^T \lambda_i + z_i c_i) = 0$$
$$\|\lambda_i\|_2 \le z_i$$

$\Rightarrow$ again f here and f in # (denoted as $f_\#$):

$$f = -f_\#$$

$\therefore$ it is equivalent to problem (dual) defined in #

#.