5.3

$$g(\lambda) = \begin{cases} -\infty & \lambda = 0 \\ \inf(c^T x + \lambda f(x)) & \lambda > 0 \end{cases}$$

$$= \begin{cases} -\infty & \lambda = 0 \\ -\lambda \sup_x \{-c^T/\lambda - f(x)\} & \lambda > 0 \end{cases}$$

$$= \begin{cases} -\infty & \lambda = 0 \\ -\lambda f^*(-c/\lambda) & \lambda > 0 \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maxmize} & -\lambda f^*(-c/\lambda) \\ \text{subject to} & \lambda \geq 0. \end{array}$$

5.4

(a)

$$\inf_{x \in H_w} c^T x = \begin{cases} \lambda w^T b & c = \lambda A^T w, \lambda \leq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

(b)

$$\begin{array}{ll} \text{maximize} & \lambda w^T b \\ \text{subject to} & c = \lambda A^T w \\ & \lambda \leq 0 \\ & w \succeq 0 \end{array}$$

Variables are λ and w

Obviously, this is not a convex problem because neither the objective function nor

(c)

Let y $= -\lambda w$

The dual problem of the original linear problem is

$$\begin{array}{ll} \text{maximize} & -b^T y \\ \text{subject to} & A^T y + c = 0 \\ & y \succeq 0. \end{array}$$

5.6

(a) Because the least squares problem has a closed-form solution x_{ls},

$$\|Ax_{ch} - b\|_2 \geq \|Ax_{ls} - b\|_2$$

using the fact that for all $z \in \mathbb{R}^m$

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

have

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

So

$$\|Ax_{ls} - b\|_\infty \leq \sqrt{m}\|Ax_{ch} - b\|_\infty$$

(b)

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

$$r_{ls} = b - Ax_{ls}$$

$$A^T r_{ls} = A^T(b - A(A^T A)^{-1} A^T b) = A^T b - A^T b = 0$$

$$(Ax_{ls} - b)^T r_{ls} = x_{ls} A^T r_{ls} - b^T r_{ls} = 0 - b^T r_{ls}$$

$$\text{so } b^T \hat{v} = \frac{-b^T r_{ls}}{\|r_{ls}\|_1} = \frac{(Ax_{ls} - b)^T r_{ls}}{\|r_{ls}\|_1} = -\frac{\|r_{ls}\|_2^2}{\|r_{ls}\|_1}$$

$$\text{Similarly } b^T \tilde{v} = \frac{\|r_{ls}\|_2^2}{\|r_{ls}\|_1}.$$

Therefore $\tilde{v}$ gives a better bound than $\hat{\nu}$

$$\text{There are factors that } \|x\|_1 \leq \sqrt{m}\|x\|_2, \quad \|x\|_\infty \leq \|x\|_2$$

$$\text{which hold for general } x \in \mathbb{R}^m$$

$$\text{So } \frac{\|r_{ls}\|_2^2}{\|r_{ls}\|_1} \geq \frac{1}{\sqrt{m}}\|r_{ls}\|_2 \geq \frac{1}{\sqrt{m}}\|r_{ls}\|_\infty$$

5.7

(a)

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

$$= \begin{cases} \inf_y \left(\max_{i=1,\ldots,m} y_i + \sum_{i=1}^m \lambda_i(b_i - y_i)\right) & \sum_i \lambda_i a_i = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

If $\exists \lambda_j < 0$, let $y_j \to -\infty$, then

$$\max_i y_i - \lambda^T y \to -\infty.$$

If $1^T\lambda \neq 1$, choosing $y = t1$

$$\max_i y_i - \lambda^T y = t(1 - 1^T\lambda) \to -\infty$$

$$\text{So } g(\lambda) = \begin{cases} b^T\lambda & \sum_i \lambda_i a_i = 0, \lambda \succeq 0, 1^T\lambda = 1 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & b^T\lambda \\ \text{subject to} & A^T\lambda = 0 \\ & 1^T\lambda = 1 \\ & \lambda \geq 0. \end{array}$$

(b)

The problem is equivalent to the Linear Problem:

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & Ax + b \preceq t1. \end{array}$$

$g(\lambda) = \inf_{t,x}\{t + \lambda(Ax + b - t1)\}$

$$\text{So } g(\lambda) = \begin{cases} b^T\lambda & A^T\lambda = 0, \lambda \succeq 0, 1^T\lambda = 1 \\ -\infty & \text{otherwise.} \end{cases}$$

which is identical to the dual derived in (a).

(c)

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

Equivalent to

$$\begin{aligned} &\text{minimize} && \log(\sum_{i=1}^{m} \exp(y_i)) \\ &\text{subject to} && a_i^T x + b_i = y_i \end{aligned}$$

$$L(x, y, \lambda) = \log(\sum_{i=1}^{m} \exp(y_i)) + \lambda(Ax + b - y)$$

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

$$\lambda^T A = 0 \text{ must be true, otherwise the value of } \lambda^T Ax \text{ is infinity}$$

Setting the gradient with respect to y equal to zero, $y_i = \log(\sum_{i=1}^{m} \exp(y_i)) + \log(\lambda_i)$

After substituting into the formula, we get:

$$\log(\sum_{i=1}^{m} \exp(y_i))(\sum_{i=1}^{m} \lambda_i - 1) - \sum_{i=1}^{m} \lambda_i \log(\lambda_i) + \lambda^T b$$

$$\sum_{i=1}^{m} \lambda_i = 1 \text{ must hold, otherwise the } g(\lambda) \text{ is infinity}$$

The dual problem is

$$\text{maxmize } b^T\lambda - \sum_{i=1}^{m} \lambda_i \log(\lambda_i)$$

$$\text{subject to } \lambda^T A = 0, \sum_{i=1}^{m} \lambda_i = 1, \lambda \succeq 0$$

Suppose $\lambda^*$is the optimal for the formulation above,so we get

$$b^T\lambda^* - \sum_{i=1}^{m} \lambda_i^* \log(\lambda_i^*) = p_{gp}^*$$

From (a), we know that, $p_{pwl}^* \geq b^T\lambda^*$

So $p_{pwl}^* \geq p_{gp}^* + \sum_{i=1}^{m} \lambda^* \log(\lambda^*)$

Through the Lagrange multiplier method, we can obtain the minimum value of

$$\inf \sum_{i=1}^{m} \lambda^* \log(\lambda^*) = -\log m$$

$$\text{So } p_{\text{pwl}}^* \geq p_{\text{gp}}^* + \sum_{i=1}^{m} \lambda_i^* \log \lambda_i^* \geq p_{\text{gp}}^* - \log m.$$

$$\text{Obviously } \max_i (a_i^T x + b_i) \leq \log \sum_i \exp(a_i^T x + b_i)$$

In conclusion,

$$p_{\text{gp}}^* - \log m \leq p_{\text{pwl}}^* \leq p_{\text{gp}}^*.$$

(d)

$$\begin{array}{ll} \text{minimize} & \left(\frac{1}{\gamma}\right) \log\left(\sum_{i=1}^{m} \exp(\gamma y_i)\right) \\ \text{subject to} & Ax + b = y. \end{array}$$

The Lagrangian is

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

The dual fuction is

$$g(\mu) = \inf_{x,y} L(x, y, \mu) = b^T\mu - \frac{1}{\gamma} \sum_{i=1}^{m} \mu_i \log \mu_i,$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & b^T\mu - \left(\frac{1}{\gamma}\right)\sum_{i=1}^{m}\mu_i \log \mu_i \\ \text{subject to} & A^T\mu = 0 \\ & 1^T\mu = 1. \end{array}$$

Let $p^*_{gp}(\gamma)$be the optimal value of the GP.

According to the conclusion in (c):

$$p^*_{gp}(\gamma) - \frac{1}{\gamma}\log m \le p^*_{pwl} \le p^*_{gp}(\gamma).$$

So $p^*_{gp}(\gamma)$ approaches p^*_{pwl} as γ increases.

5.9

(a)

$$\left[\begin{array}{cc} \sum_{k=1}^{m} a_k a_k^T & a_i \\ a_i^T & 1 \end{array}\right] = \left[\begin{array}{cc} \sum_{k\neq i} a_k a_k^T & 0 \\ 0 & 0 \end{array}\right] + \left[\begin{array}{c} a_i \\ 1 \end{array}\right]\left[\begin{array}{cc} a_i^T & 1 \end{array}\right]$$

Obviously the original formula is equal to the addition of two positive semi-definite matrices, so it is still a positive semi-definite matrix.

According to the Schur complement:

$$\left[\begin{array}{cc} A & B \\ B^T & C \end{array}\right] \succeq 0 \Longleftrightarrow A - BC^{-1}B^T \succeq 0$$

So

$$\sum_{k=1}^{m} a_k a_k^T - a_i^T a_i \geq 0 = 1 - a_i^T\left(\sum_{k=1}^{m} a_k a_k^T\right)^{-1} a_i \geq 0,$$

(b)

$\lambda = t1$

So the dual fuction is

$$g(\lambda) = \log\det\left(\sum_{i=1}^{m} a_i a_i^T\right) + n\log t - mt + n.$$

Derive t and let the derivative be 0, we get t= n/m, then the dual fuction is

$$g(\lambda) = \log\det\left(\sum_{i=1}^{m} a_i a_i^T\right) + n\log(n/m).$$

so the duality gap associated with X_{sim} and λ is $n\log(m/n)$.

X_{sim} is no more than nlog(m/n) suboptimal.

5.10

(a)

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

$$L(x, Z, \lambda, \nu) = \log\det(X^{-1}) + < Z, (X - \sum_{i=1}^{p} x_i v_i x_i^T) > -\lambda^T x + \nu(1^T x - 1)$$

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

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

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

only if $-v_i Z v_i - \lambda_i + \nu = 0,$ the minimum over xi is bounded below

$$\nabla(\log\det(X^{-1}) + tr(ZX)) = -X^{-1} + Z$$

$$\text{Let the gradient=0,so } Z = X^{-1}$$

$$\text{So } g(Z, \lambda, \nu) = \log\det Z + n - \nu, \nu = \lambda_i + v_i Z v_i^T$$

The dual problem is

$$
\begin{array}{ll}
\text{maximize} & \log\det Z + n - \nu \\
\text{subject to} & v_i^T Z v_i \le \nu, \quad i = 1, \ldots, p,
\end{array}
$$

Let $W=Z/\nu$

$$
\begin{array}{ll}
\text{maximize} & \log\det W + n + n\log\nu - \nu \\
\text{subject to} & v_i^T \hat{W} v_i \le 1, \quad i = 1, \ldots, p,
\end{array}
$$

Let the gradient of $n\log\nu - \nu = 0,\ \nu = n$

The final problem is

$$
\begin{array}{ll}
\text{maximize} & \log\det W + n\log n \\
\text{subject to} & v_i^T W v_i \le 1, \quad i = 1, \ldots, p.
\end{array}
$$

(b)

$$
\begin{array}{ll}
\text{minimize} & \mathrm{tr}(X^{-1}) \\
\text{subject to} & X = \left(\sum_{i=1}^{p} x_i v_i v_i^T\right) \\
& x \succeq 0, \\
& 1^T x = 1.
\end{array}
$$

$$
L(X, Z, \lambda, \nu) = \mathrm{tr}(X^{-1}) + \mathrm{tr}(ZX) - \sum_{i=1}^{p} x_i v_i^T Z v_i - \lambda^T x + \nu(1^T x - 1)
$$

$$
= \mathrm{tr}(X^{-1}) + \mathrm{tr}(ZX) + \sum_{i=1}^{p} x_i(-v_i^T Z v_i - \lambda_i + \nu) - \nu.
$$

only if $-v_i Z v_i - \lambda_i + \nu = 0,$ the minimum over xi is bounded below

$$
\nabla(tr(X^{-1}) + tr(ZX)) = -X^{-2} + Z
$$

Let the gradient=0, $Z = X^{-2}$

$$
\inf_{X \succeq 0} (\mathrm{tr}(X^{-1}) + \mathrm{tr}(ZX)) = \begin{cases} 2\mathrm{tr}(Z^{1/2}) & Z \succeq 0 \\ -\infty & \text{otherwise.} \end{cases}
$$

$g(Z, \lambda, \nu) = -\nu + 2tr(Z^{1/2})\ Z \succeq 0, v_i^T Z v_i + \lambda = \nu$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -\nu + 2\mathrm{tr}(Z^{1/2}) \\ \text{subject to} & v_i^T Z v_i \leq \nu, \quad i = 1, \ldots, p \\ & Z \succeq 0. \end{array}$$

Let $\mathrm{W} = \mathrm{Z}/\nu$

$$\begin{array}{ll} \text{maximize} & -\nu + 2\sqrt{\nu}\mathrm{tr}(W^{1/2}) \\ \text{subject to} & v_i^T W v_i \leq 1, \quad i = 1, \ldots, p \\ & W \succeq 0. \end{array}$$

By optimizing over (ν > 0)

$$\begin{array}{ll} \text{maximize} & (\mathrm{tr}(W^{1/2}))^2 \\ \text{subject to} & v_i^T W v_i \leq 1, \quad i = 1, \ldots, p \\ & W \succeq 0. \end{array}$$

5.11

$$L(x, y, \nu) = \sum_{i=1}^{N} \|\mathbf{y}_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \nu_i^T(\mathbf{y}_i - A_i x - \mathbf{b}_i).$$

$$|\nu_i^T y_i| \leq ||\nu_i||_2 ||y_i||_2$$

$$||y_i||_2 + \nu_i^T y_i \geq (1 - ||\nu_i||_2)||y_i||_2$$

$$\text{So } \inf_{y_i}\{||y_i||_2 + \nu_i^T y_i\} = \begin{cases} 0 & ||\nu_i|| \leq 1 \\ -\infty & \text{otherwise.} \end{cases}$$

Setting the gradient with respect to (x) equal to zero,get

$$x = x_0 + \sum_{i=1}^{N} A_i^T \nu.$$

$$g(\nu) = \begin{cases} \sum_{i=1}^{N}(A_i x_0 + b_i)^T \nu_i - \frac{1}{2}\left\|\sum_{i=1}^{N} A_i^T \nu_i\right\|^2 & \text{if } \|\nu\|_2 \leq 1, \; i = 1, \ldots, N \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is

$$
\begin{array}{ll}
\text{maximize} & \sum_{i=1}^{N}(A_i x_0 + b_i)^T \nu_i - \frac{1}{2}\left\|\sum_{i=1}^{N} A_i^T \nu_i\right\|^2 \\
\text{subject to} & \|\nu_i\|_2 \le 1,\ i = 1, \dots, N.
\end{array}
$$

5.12

$$
\begin{array}{ll}
\text{minimize} & -\sum_{i=1}^{m} \log y_i \\
\text{subject to} & y = b - Ax,
\end{array}
$$

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

$\nu^T A = 0$ must be true, otherwise the value of $\nu^T Ax$ is infinity

Setting the gradient with respect to y equal to zero, $y_i = 1/\nu_i$ only if $\nu_i > 0$

The dual function:

$$
g(\nu) = \begin{cases} \sum_{i=1}^{m} \log \nu_i + m - b^T \nu & \text{if } A^T \nu = 0, \nu > 0 \\ -\infty & \text{otherwise} \end{cases}
$$

The dual problem

$$
\begin{array}{ll}
\text{maximize} & \sum_{i=1}^{m} \log \nu_i - b^T \nu + m \\
\text{subject to} & A^T \nu = 0.
\end{array}
$$

5.13

(a)

The Lagrangian is

$$
\begin{aligned}
L(x, \lambda, \nu) &= c^T x + \lambda^T (Ax - b) + x^T \operatorname{diag}(\nu) x - \nu^T x \\
&= x^T \operatorname{diag}(\nu) x + (c + A^T \lambda - \nu)^T x - b^T \mu.
\end{aligned}
$$

Setting the gradient with respect to x equal to zero, $y_i = 1/\nu_i$ only if $\nu_i > 0$

$$g(\lambda, \nu) = \begin{cases} -b^T\lambda - \frac{1}{4}\sum_{i=1}^n \left(c_i + a_i^T\lambda - \nu_i\right)^2 / \nu_i, & \nu \succeq 0 \\ -\infty, & \text{otherwise} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -b^T\lambda - \dfrac{1}{4}\displaystyle\sum_{i=1}^n \frac{(c_i + a_i^T\lambda - \nu_i)^2}{\nu_i} \\ \text{subject to} & \nu \geq 0. \end{array}$$

$$\text{From } \frac{a^2}{x} + x \geq 2|a|, x > 0, \text{ we have:}$$

$$\sup_{\nu_i \geq 0}\left(-\frac{(c_i + a_i^T\lambda - \nu_i)^2}{\nu_i}\right) = \begin{cases} (c_i + a_i^T\lambda) & \text{if } c_i + a_i^T\lambda \leq 0 \\ 0 & \text{if } c_i + a_i^T\lambda > 0 \end{cases} = \min\{0, (c_i + a_i^T\lambda)\}.$$

$$\text{The dual problem is simplified to}$$

$$\begin{array}{ll} \text{maximize} & -b^T\lambda + \displaystyle\sum_{i=1}^n \min\{0, c_i + a_i^T\lambda\} \\ \text{subject to} & \lambda \geq 0. \end{array}$$

(b)

$$\text{The largrangian for (5.107) is}$$
$$\begin{aligned} L(x, u, v, w) &= c^T x + u^T(Ax - b) - v^T x + w^T(x - 1) \\ &= (c + A^T u - v + w)^T x - b^T u - 1^T w \end{aligned}$$

$$g(u, v, w) = \begin{cases} -b^T u - 1^T w & \text{if } A^T u - v + w + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -b^T u - 1^T w \\ \text{subject to} & A^T u - v + w + c = 0 \\ & u \geq 0, v \geq 0, w \geq 0, \end{array}$$

is equivalent to the Lagrange relaxation problem derived above

so the two relaxations give the same value.

5.21

(1)

$\nabla^2 e^{-x} = e^{-x} > 0$

So the objective function is convex.

The inequality constraint is obviously a convex set.

So this is a convex optimization problem.

$x^2 \geq 0$ and the domain D requires $y > 0$

So the optimal value is 1 where x=0 obviously.

(2)

The Lagrangian is

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

The dual fuction is

$$g(\lambda) = \inf_{x,y>0} \{e^{-x} + \lambda x^2/y\} = 0$$

The dual problem is

$$\text{maxmize } 0$$

$$\text{subject to } \lambda \geq 0$$

$\lambda^* \geq 0$, The optimal value is 0.The optimal duality gap is 1-0=1

(c)

Slater's condition is not satisfied

(d)

If $\mu = 0, p^*(\mu) = 1$ as the problem (a)

If $\mu > 0$, x can go to infinity ,so $p*(\mu) = 0$

Let $\lambda^* = 0$,

$p^*(\mu) = 0 \geq p^*(0) = 1$ do not hold.