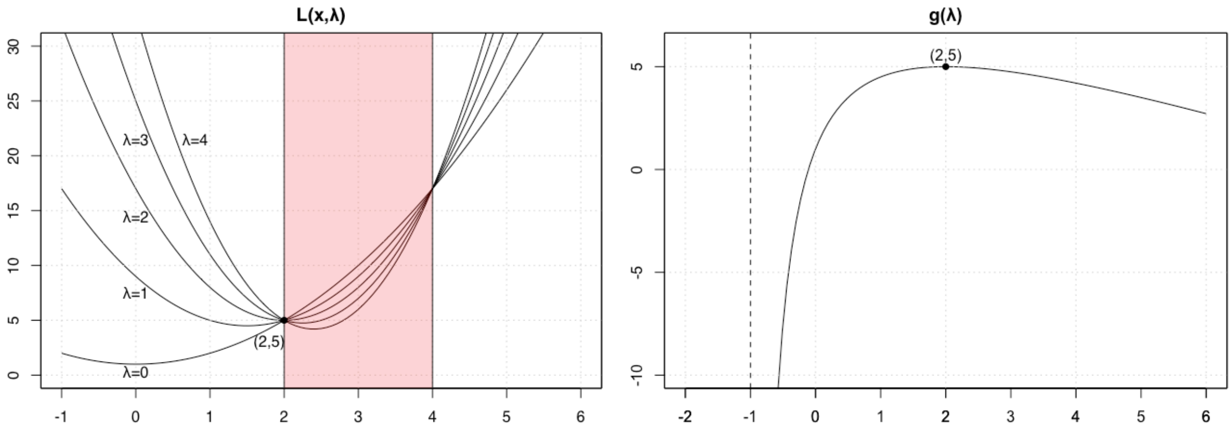
Homework 4 B00401062 羅文斌

5.1

a. Let $f_0(x) = x^2 + 1$ and $f_1(x) = (x-2)(x-4)$. First, $f_0(x)$ has no implicit constraints, and $f_1(x) = (x-2)(x-4) \le 0 \rightarrow 2 \le x \le 4$. Therefore, the feasible set is $2 \le x \le 4$.
   $f_0(x)$ is strictly increasing when $x > 0$ because $f_0'(x) = 2x > 0$ if $x > 0$. Therefore, the optimal point is at $x^* = 2$, and the optimal value $p^* = 2^2 + 1 = 5$.

b. The Lagrangian $L(x,\lambda) = (x^2+1) + \lambda(x-2)(x-4) = (\lambda+1)x^2 - 6\lambda x + (8\lambda+1)$. If $\lambda = 0$, then $L(x,\lambda) = L(x,0)$ is exactly the objective function $f_0(x)$. The optimal point (x^*, p^*) is $(2,5)$ as shown in the plot.
   To minimize $L(x,\lambda)$ over x, we let $\nabla_x L(x,\lambda) = 2(\lambda+1)x - 6\lambda = 0$, and $x = 3\lambda/(\lambda+1)$ if $\lambda \ne -1$. Plugging x into $L(x,\lambda)$ gives the dual function $g(\lambda) = \inf_x L(x,\lambda)$ $= -9\lambda^2/(\lambda+1) + (8\lambda+1), \lambda \ne -1$.
   To maximize $g(\lambda)$, we let $g'(\lambda) = -9\lambda(\lambda+2)/(\lambda+1)^2 + 8 = 0$, and $\lambda = -4, 2$. Plugging $\lambda = 2$ (because we want $\lambda \ge 0$) into $g(\lambda)$ gives $g(2) = 5$.
   Therefore, $p^* = g(2) \ge g(\lambda) = \inf_x L(x,\lambda)$. The lower bound property holds. From the plot, we can also see that lower bound property holds. The minimum of $L(x,\lambda)$ over x given λ is always no greater than p^*.



c. The dual problem is $g(\lambda) = \inf_x L(x,\lambda) = -9\lambda^2/(\lambda+1) + (8\lambda+1)$ subject to $\lambda \ge 0$.
   Concavity: $-g(\lambda) = 9\lambda^2/(\lambda+1) - (8\lambda+1)$, and $-g'(\lambda) = 9\lambda(\lambda+2)/(\lambda+1)^2 - 8$, and $-g''(\lambda) = 18/(\lambda+1)^3$. When $\lambda \ge 0$, $-g''(\lambda)$ is always greater than 0. Therefore, $-g(\lambda)$ is a convex function when $\lambda \ge 0$, and $g(\lambda)$ is a concave function when $\lambda \ge 0$.
   To maximize $g(\lambda)$, we let $g'(\lambda) = -9\lambda(\lambda+2)/(\lambda+1)^2 + 8 = 0$, and $\lambda = -4, 2$. Plugging $\lambda = 2$ (because we want $\lambda \ge 0$) into $g(\lambda)$ gives $g(2) = 5$.
   Therefore, $\lambda^* = 2$, and strong duality holds because $p^* = g(\lambda^*) = 5$.

5.7

To simplify the notation, we let $A = [a_1^{\mathrm{T}}, \ldots, a_m^{\mathrm{T}}] \in \mathbb{R}^{m\times n}$.

a. The dual function is $g(\nu) = \inf_{x,y} L(x,y,\nu) = \inf_{x,y}(\max\{y\} + \nu^{\mathrm{T}}(Ax+b-y))$.

To minimize over x, we let $A^{\mathrm{T}}\nu = 0$. If not so, $\inf_{x,y} L(x,y,\nu) = -\infty$. Plugging $A^{\mathrm{T}}\nu = 0$ into $L(x,y,\nu)$ gives $g(\nu) = \inf_y L(y,\nu) = \inf_y(\max\{y\} + \nu^{\mathrm{T}}(b-y)) = b^{\mathrm{T}}\nu + \inf_y(\max\{y\} - \nu^{\mathrm{T}}y)$

To minimize over y, we consider two conditions: (1) $\nu \not\succeq 0$ and (2) $1^{\mathrm{T}}\nu \neq 1$.

1. $\nu \not\succeq 0$, that is there exist some $\nu_i < 0$. If y_i goes to $-\infty$, then $\inf_y(\max\{y\} - \nu^{\mathrm{T}}y)$ $= \inf_y(\max\{y\} - \infty) = -\inf$. Therefore, to avoid $g(\nu)$ goes to $-\infty$, we need $\nu \succeq 0$.
2. $1^{\mathrm{T}}\nu \neq 1$. If $1^{\mathrm{T}}\nu > 1$, $y = t1$, and t goes to ∞, then $\inf_y(\max\{y\} - \nu^{\mathrm{T}}y) = \inf_y(t - \nu^{\mathrm{T}}t1)$ $= \inf_y t(1-\nu^{\mathrm{T}}1) = -\infty$. Otherwise, if $1^{\mathrm{T}}\nu < 1$, $y = t1$, and t goes to $-\infty$, then $\inf_y(\max\{y\} - \nu^{\mathrm{T}}y) = \inf_y(t - \nu^{\mathrm{T}}t1) = \inf_y t(1-\nu^{\mathrm{T}}1) = -\infty$. Therefore, to avoid $g(\nu)$ goes to $-\infty$, we need $1^{\mathrm{T}}\nu = 1$.

If under the restriction that $\nu \succeq 0$ and $1^{\mathrm{T}}\nu = 1$, then $\max\{y\} - \nu^{\mathrm{T}}y \geq \max\{y\} - \nu^{\mathrm{T}}\max\{y\}1$ $= \max\{y\}(1-\nu^{\mathrm{T}}1) = 0$. Therefore, $\inf_y(\max\{y\} - \nu^{\mathrm{T}}y) = 0$.

$$g(\nu) = \begin{cases} b^{\mathrm{T}}\nu \text{ if } \nu \succeq 0, 1^{\mathrm{T}}\nu = 1 \\ -\infty \text{ otherwise} \end{cases}.$$

The dual problem: maximize $g(\nu) = b^{\mathrm{T}}\nu$ subject to $A^{\mathrm{T}}\nu = 0, \nu \succeq 0, 1^{\mathrm{T}}\nu = 1$.

b. The equivalent LP problem: minimize t subject to $Ax + b \preccurlyeq t1$.

The Lagrangian is $L(x,t,\lambda) = t + \lambda^{\mathrm{T}}(Ax+b-t1)$.

The dual function is $g(\lambda) = \inf_{x,t} L(x,t,\lambda) = \inf_{x,t}(t + \lambda^{\mathrm{T}}(Ax+b-t1))$.

To minimize over x, we let $A^{\mathrm{T}}\lambda = 0$. If not so, $\inf_{x,y} L(x,t,\nu) = -\infty$. Plugging $A^{\mathrm{T}}\lambda = 0$ into $L(x,t,\lambda)$ gives $g(\lambda) = \inf_t L(t,\lambda) = \inf_t(t + \lambda^{\mathrm{T}}(b-t1))$. To minimize over t, we let $1^{\mathrm{T}}\lambda = 1$. If not so, $\inf_t L(t,\lambda) = -\infty$. Plugging $1^{\mathrm{T}}\lambda = 1$ into $L(t,\lambda)$ gives $g(\lambda) = b^{\mathrm{T}}\lambda$.

The dual problem: maximize $g(\lambda) = b^{\mathrm{T}}\lambda$ subject to $A^{\mathrm{T}}\lambda = 0, 1^{\mathrm{T}}\lambda = 1, \lambda \succcurlyeq 0$.

5.27

First, the Lagrangian is $L(x,\nu) = \| Ax - b \|_2^2 + \nu^{\mathrm{T}}(Gx-h) = (Ax-b)^{\mathrm{T}}(Ax-b) + \nu^{\mathrm{T}}(Gx-h)$ $= x^{\mathrm{T}}A^{\mathrm{T}}Ax + (G^{\mathrm{T}}\nu - 2A^{\mathrm{T}}b)^{\mathrm{T}}x + (b^{\mathrm{T}}b - \nu^{\mathrm{T}}h)$.

$\nabla_x L(x,\nu) = 2A^{\mathrm{T}}Ax + (G^{\mathrm{T}}\nu - 2A^{\mathrm{T}}b) = 0 \rightarrow x = -(1/2)(A^{\mathrm{T}}A)^{-1}(G^{\mathrm{T}}\nu - 2A^{\mathrm{T}}b)$.

Plugging x into $L(x,\nu)$ gives $g(\nu) = -(1/4)(G^{\mathrm{T}}\nu - 2A^{\mathrm{T}}b)^{\mathrm{T}}(A^{\mathrm{T}}A)^{-1}(G^{\mathrm{T}}\nu - 2A^{\mathrm{T}}b) - \nu^{\mathrm{T}}h$.

KKT conditions: $Gx = h$, $\nabla_x L(x,\nu) = 0$, complementary slackness always holds, and no dual constraints (because no primal inequality constraints). Since the primal problem is convex, x^* and ν^* satisfying KKT conditions are primal solutions, and strong duality holds.

With $Gx^* = h$ and $x^* = -(1/2)(A^{\mathrm{T}}A)^{-1}(G^{\mathrm{T}}\nu^* - 2A^{\mathrm{T}}b)$, we can derive:

$\nu^* = -2(G(A^{\mathrm{T}}A)^{-1}G^{\mathrm{T}})^{-1}(h - G(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b)$

$x^* = (A^{\mathrm{T}}A)^{-1}(G^{\mathrm{T}}(G(A^{\mathrm{T}}A)^{-1}G^{\mathrm{T}})^{-1}(h - G(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}b) + A^{\mathrm{T}}b)$