Problem 8.1.3

1.
$$\min f(x_1,x_2) = x_1^2 + x_2^2$$
$$\text{s.t. } g(x_1,x_2) = -2x_1 + x_2 = -5$$

$$\Rightarrow \mathcal{L} = \max_{(\lambda_1,\lambda_2)\in\mathbb{R}_+} \min_{x_1,x_2} x_1^2 + x_2^2 + \lambda_1(-2x_1+x_2+5) + \lambda_2(-5+2x_1-x_2)$$

Setting $\lambda = \lambda_1 - \lambda_2 \Rightarrow \max_{\lambda\in\mathbb{R}} x_1^2 + x_2^2 + \lambda(-2x_1 + x_2 + 5)$

$$\text{s.t. } 2x_1 - 2\lambda = 0$$
$$2x_2 + \lambda = 0$$

$\nabla_{x_1}\mathcal{L} = 2x_1 - 2\lambda = 0 \Rightarrow x_1 = \lambda$
$\nabla_{x_2}\mathcal{L} = 2x_2 + \lambda = 0 \Rightarrow x_2 = -\lambda/2$ } Sub back into $\mathcal{L}$

$$\Rightarrow \max_\lambda \mathcal{L}(\lambda) = \lambda^2 + \frac{\lambda^2}{4} + \lambda\left(-2\lambda - \frac{\lambda}{2} + 5\right) = -\frac{5}{4}\lambda^2 + 5\lambda$$

2. $\mathcal{L}'(\lambda) = -\frac{5}{2}\lambda + 5 = 0 \Rightarrow \mathcal{L}'(2) = -\frac{5}{2}\cdot 2 + 5 = 0 \Rightarrow 0 = 0$
$\Rightarrow -\frac{5}{2}\lambda = -5 \Rightarrow \lambda = 2$

$\mathcal{L}''(2) = -\frac{5}{2} \leq 0 \Rightarrow x_1 = \lambda = 2,\ x_2 = -\frac{2}{2} = -1$ are the optimal solutions

3. $\mathcal{L}_\sigma(x,\lambda) = x_1^2 + x_2^2 + \lambda(-2x_1 + x_2 + 5) + \frac{\sigma}{2}\|-2x_1 + x_2 + 5\|^2$

$\nabla_{x_1}\mathcal{L}_\sigma = 2x_1 - 2\lambda - 2\sigma(-2x_1 + x_2 + 5) = 0$
$\Rightarrow x_1 - \lambda + 2\sigma x_1 - \sigma x_2 - 5\sigma = 0$
$\Rightarrow x_1(1+2\sigma) = \lambda + \sigma x_2 + 5\sigma \Rightarrow x_1 = \dfrac{\lambda + \sigma x_2 + 5\sigma}{1+2\sigma}$

$\nabla_{x_2}\mathcal{L}_\sigma = 2x_2 + \lambda + \sigma(-2x_1 + x_2 + 5) = 0 \Rightarrow x_2(2+\sigma) + \lambda - 2\sigma x_1 + 5\sigma = 0$
$\Rightarrow x_2(2+\sigma) = 2\sigma x_1 - 5\sigma - \lambda \Rightarrow x_2 = \dfrac{2\sigma x_1 - 5\sigma - \lambda}{2+\sigma}$

$x_2 \Rightarrow x_1$
$$\Rightarrow x_1 = \frac{\lambda}{1+2\sigma} + \frac{2\sigma^2 x_1 - 5\sigma^2 - \lambda\sigma}{(1+2\sigma)(2+\sigma)} + \frac{5\sigma}{1+2\sigma}$$

$\Rightarrow$
$$x_1 - \frac{2\sigma^2 x_1}{(1+2\sigma)(2+\sigma)} = \frac{\lambda(2+\sigma) - 5\sigma^2 - \lambda\sigma + 5\sigma(2+\sigma)}{(1+2\sigma)(2+\sigma)}$$

8.1.3

3. (cont)

$$\Rightarrow x_1\left(1-\frac{2\sigma^2}{(1+2\sigma)(2+\sigma)}\right) = \frac{2\lambda+\lambda\sigma-5\sigma^2-\lambda\sigma+10\sigma+5\sigma^2}{(1+2\sigma)(2+\sigma)}$$

$$\Rightarrow x_1\left(\frac{2+5\sigma+2\sigma^2-2\sigma^2}{(2+\sigma)(1+2\sigma)}\right) = \frac{2\lambda+10\sigma}{(1+2\sigma)(2+\sigma)} \Rightarrow x_1\left(\frac{2+5\sigma}{(1+2\sigma)(2+\sigma)}\right) = \frac{2(\lambda+5\sigma)}{(1+2\sigma)(2+\sigma)}$$

$$\Rightarrow x_1 = \frac{2(\lambda+5\sigma)}{2+5\sigma} \quad \therefore \lambda = 2 \Rightarrow x_1 = 2$$

$$x_2 = \frac{2\sigma x_1 - 5\sigma - \lambda}{2+\sigma} \quad \therefore \lambda = 2,\ x_1 = 2 \Rightarrow \frac{4\sigma-5\sigma-2}{2+\sigma} \Rightarrow x_1 = \frac{-2-\sigma}{2+\sigma} = -1$$

$\therefore x_1 = 2,\ x_2 = -1$

4. $$\nabla^2 \mathcal{L}_\sigma = \begin{bmatrix} \nabla^2_{x_1}\mathcal{L}_\sigma & \nabla_{x_1}\nabla_{x_2}\mathcal{L}_\sigma \\ \nabla_{x_2}\nabla_{x_1}\mathcal{L}_\sigma & \nabla^2_{x_2}\mathcal{L}_\sigma \end{bmatrix}$$

$$\nabla_{x_1}\mathcal{L}_\sigma = 2x_1 - 2\lambda - 2\sigma(-2x_1 + x_2 + 5) = 2x_1 - 2\lambda + 4\sigma x_1 - 2\sigma x_2 - 10\sigma$$

$$\nabla^2_{x_1}\mathcal{L}_\sigma = 2+4\sigma,\ \nabla_{x_1}\nabla_{x_2} = -2\sigma$$

$$\nabla_{x_2}\mathcal{L}_\sigma = 2x_2 + \lambda + \sigma(-2x_1 + x_2 + 5) = 2x_2 + \lambda - 2\sigma x_1 + \sigma x_2 + 5\sigma$$

$$\nabla^2_{x_2}\mathcal{L}_\sigma = 2+\sigma,\ \nabla_{x_2}\nabla_{x_1}\mathcal{L}_\sigma = -2\sigma$$

$$\nabla^2\mathcal{L}_\sigma = \begin{bmatrix} 2+4\sigma & -2\sigma \\ -2\sigma & 2+\sigma \end{bmatrix} \Rightarrow \left|\nabla^2\mathcal{L}_\sigma - \lambda\cdot I\right| = (2+4\sigma-\lambda)(2+\sigma-\lambda) - 4\sigma^2 = 0$$

$$\Rightarrow 4 + 2\sigma - 2\lambda + 8\sigma + 4\sigma^2 - 4\sigma\lambda - 2\lambda - \sigma\lambda + \lambda^2 - 4\sigma^2 = 0$$

$$\Rightarrow 4 + 10\sigma - 5\sigma\lambda - 4\lambda + \lambda^2 \Rightarrow \lambda^2 + \lambda(-5\sigma-4) + 10\sigma + 4$$

$$\Rightarrow \frac{-b \pm \sqrt{b^2-4ac}}{2a} \Rightarrow \frac{1}{2}\left[5\sigma+4 \pm \sqrt{25\sigma^2+40\sigma+16-40\sigma-16}\right]$$

$$= \frac{1}{2}\left[5\sigma+4 \pm 5\sigma\right] \Rightarrow 5\sigma+2,\ 2 > 0 \text{ for } \sigma > 0 \Rightarrow \nabla^2\mathcal{L}_\sigma \text{ is positive-definite}$$

8.1.3

5.

$\nabla_{x_1} L_\sigma(x,\lambda) = 0 \Rightarrow x_1 = \dfrac{\lambda + \sigma x_2 + 5\sigma}{1+2\sigma}$ as in part 3

$\nabla_{x_2} L_\sigma(x,\lambda = 0 \Rightarrow x_2 = \dfrac{2\sigma x_1 - 5\sigma - \lambda}{2+\sigma}$ Sub $x_2 \Rightarrow x_1 \Rightarrow x_1 = \dfrac{2(\lambda+5\sigma)}{2+5\sigma}$

(as in part 3) $\therefore \lim\limits_{\substack{\sigma\to\infty \\ \lambda \text{ fixed}}} x_1 = 2$, $\lim\limits_{\substack{\lambda\to 2 \\ \sigma \text{ fixed}}} x_1 = 2\dfrac{(2+5\sigma)}{2+5\sigma} = 2$

Sub $x_1 \Rightarrow x_2$ $\Rightarrow x_2 = \dfrac{2\sigma}{2+\sigma}\left(\dfrac{\lambda + \sigma x_2 + 5\sigma}{1+2\sigma}\right) - \dfrac{5\sigma}{2+\sigma} - \dfrac{\lambda}{2+\sigma}$

$\Rightarrow x_2 - \dfrac{2\sigma^2 x_2}{(2+\sigma)(1+2\sigma)} = \dfrac{2\sigma\lambda - 5\sigma(1+2\sigma) - \lambda(1+2\sigma) + 10\sigma^2}{(2+\sigma)(1+2\sigma)}$

$\Rightarrow x_2\left(1 - \dfrac{2\sigma^2}{(2+\sigma)(1+2\sigma)}\right) = \dfrac{2\sigma\lambda - 5\sigma + 10\sigma^2 - \lambda - 2\sigma\lambda + 10\sigma^2}{(2+\sigma)(1+2\sigma)}$

$\Rightarrow x_2\left(\dfrac{2+5\sigma+2\sigma^2-2\sigma^2}{(2+\sigma)(1+2\sigma)}\right) = \dfrac{-5\sigma-\lambda}{(2+\sigma)(1+2\sigma)} \Rightarrow x_2\left(\dfrac{2+5\sigma}{(2+\sigma)(1+2\sigma)}\right) = \dfrac{-(\lambda+5\sigma)}{(2+\sigma)(1+2\sigma)}$

$\Rightarrow x_2 = \dfrac{-(\lambda+5\sigma)}{2+5\sigma}$ $\therefore \lim\limits_{\substack{\sigma\to\infty \\ \lambda \text{ fixed}}} x_2 = -1$, $\lim\limits_{\substack{\lambda\to 2 \\ \sigma \text{ fixed}}} x_2 = -\dfrac{(2+5\sigma)}{2+5\sigma} = -1$

$\therefore \lim\limits_{\substack{\sigma\to\infty \\ \lambda \text{ fixed}}} (x_1(\lambda,\sigma), x_2(\lambda,\sigma)) \Rightarrow x^* = (2,-1)$

$\lim\limits_{\substack{\lambda\to 2 \\ \sigma \text{ fixed}}} (x_1(\lambda,\sigma), x_2(\lambda,\sigma)) \Rightarrow x^* = (2,-1)$

Problem 8.2.1

$$\min_{(x,z)} \{ f(x) + g(z) \mid Ax - z = 0 \}$$

$f(x) = c^T x$, $g(z) = \delta_{C = \{z \mid z \le b\}}$

$x^{k+1} \in \underset{x \in \{x \mid Ax - z\}}{\arg\min} \left[c^T x + \lambda^T (Ax - z) + \frac{\rho}{2} \|Ax - z\|^2 \right]$

$L_{\rho^k}(x, \lambda^k, z^k) = c^T x + \lambda^T (Ax - z) + \frac{\rho}{2} \|Ax\|^2 - \frac{\rho}{2} A x^T z + \|z\|^2$

$\nabla_x L_{\rho^k}(x, \lambda^k, z^k) = c^T I + A^T \lambda^k + \rho^k A^T A x^{k+1} - \rho^k A^T z^k = 0$

$\Rightarrow \quad \rho^k (A^T A) x^{k+1} = \rho^k A^T z^k - A^T \lambda^k - c^T I$

$\Rightarrow \quad (A^T A) x^{k+1} = A^T z^k - A^T \nu^k - \frac{1}{\rho^k} c^T I$

$\Rightarrow \quad x^{k+1} = (A^T A)^{-1} A^T (z^k - \nu^k) - \frac{1}{\rho^k} (A^T A)^{-1} c^T I$

$\quad = (A^T A)^{-1} A^T \left(z^k - \frac{\lambda^k}{\rho^k} \right) - \frac{1}{\rho^k} (A^T A)^{-1} c^T I$

$z^{k+1} = \text{Proj}_C (A x^{k+1} + \nu^k)$ where $C = \{ z \mid z \le b \}$

$\quad = \min \{ A x^{k+1} + \nu^k, b \} = \min \{ A x^{k+1} + \frac{\lambda^k}{\rho^k}, b \}$

$\lambda^{k+1} =$

$\quad = \lambda^k + \rho^k (A x^{k+1} - z^{k+1})$