suggested solutions to A3

1.

2.

3.

4. *Proof.* $(a) \Rightarrow (b)$: trivial.

   $(b) \Rightarrow (a)$: Suppose the MFCQ holds at $\bar{x}$, that is to say, $\exists d$ such that
   $$d^T \nabla g_i(\bar{x}) < 0, \ \forall i \in I(\bar{x}).$$
   Then for any $i \in I(\bar{x})$, $\exists \alpha_i > 0$ such that $g_i(\bar{x} + \alpha d) < g_i(\bar{x}) = 0$ for any $0 < \alpha \leq \alpha_i$. On the other hand, for any $i \notin I(\bar{x})$, $\exists \beta_i > 0$ such that for any $x \in \mathbb{B}(\bar{x}, \beta_i)$ $|g_i(x) - g_i(\bar{x})| < \left|\frac{g_i(\bar{x})}{2}\right|$ (and so $g_i(x) < \frac{g_i(\bar{x})}{2} < 0$) holds.

   Let $\varepsilon = \min\left\{\min_{i \in I(\bar{x})} \{\alpha_i\}, \min_{i \notin I(\bar{x})} \left\{\frac{\beta_i}{\|d\|}\right\}\right\}$. Then $g_i(\bar{x} + \varepsilon d) < 0$ for all $i = 1, 2, \cdots, p$. □

5. *Proof.* Let $\mathcal{F}$ be the feasible region. Recall that $\bar{x}$ minimizes $L(\cdot, \lambda)$ over $\mathbb{R}^n$ and $\bar{x} \in \mathcal{F}$, then $\bar{x}$ minimizes $L(\cdot, \lambda)$ over $\mathcal{F}$ (think of why). Hence for any $x \in \mathcal{F}$,
   $$f(\bar{x}) = L(\bar{x}, \lambda) \leq L(x, \lambda) \leq f(x),$$
   where the last inequality holds since $\lambda \geq 0$ and $g(x) \leq 0$. □

6. *Proof.* (a) for the auxiliary problem, the objective is continuous and the feasible region is compact, then the global optimal solutions set X_k is nonempty. Next we show that pick any $x^k \in X_k$, we have $x^k \to x^*$.

   Notice that $x^k \in \mathbb{B}(x^*, \varepsilon)$, that is to say, $\{x^k\}_{k \geq 1}$ is bounded, so there exists at least one limit point. WLOG, let $x^k \to \bar{x}$. Observe the fact that
   $$0 \leq \frac{k}{2}\left[\sum_{i=1}^{p}\left[h_i(x^k)\right]^2 + \sum_{j=1}^{q}\left[g_j^+(x^k)\right]^2\right] \leq f(x^*) - f\left(x^k\right) - \frac{1}{2}\left\|x^k - x^*\right\|^2,$$
   and
   $$0 \leq f(x^*) - f\left(x^k\right) - \frac{1}{2}\left\|x^k - x^*\right\|^2 \to f(x^*) - f(\bar{x}) - \frac{1}{2}\|\bar{x} - x^*\|^2.$$
   Then the sequence
   $$\frac{k}{2}\left[\sum_{i=1}^{p}\left[h_i(x^k)\right]^2 + \sum_{j=1}^{q}\left[g_j^+(x^k)\right]^2\right]$$
   is bounded and so $h_i(\bar{x}) = 0$ for $i = 1, 2, \cdots, p$ and $g_j(\bar{x}) \leq 0$ for $j = 1, 2, \cdots, q$.(think of why). That is to say $\bar{x}$ is feasible. Recall that
   $$0 \leq f(x^*) - f(\bar{x}) - \frac{1}{2}\|\bar{x} - x^*\|^2,$$
   $\bar{x} \in \mathcal{F} \cap \mathbb{B}(x^*, \varepsilon)$ and $f(x) \geq f(x^*)$ for any $x \in \mathcal{F} \cap \mathbb{B}(x^*, \varepsilon)$. We have $\bar{x} = x^*$.

(b) It's sufficient to show

$$\nabla \left[g_j^+(\cdot)\right]^2 = 2g_j^+(\cdot)\nabla g_j\left(\cdot\right).$$

(think of why)

(c) Notice that all $\{\lambda_{k,0}, \lambda_{k,1}, \cdots, \lambda_{k,p}, \mu_{k,1}, \cdots, \mu_{k,q}\}_{k\geq 1} \subseteq \mathbb{R}^{1+p+q}$ are located on the unit sphere.

(d) Recall that

$$\begin{aligned}
0 =& \nabla f\left(x^k\right) + k\sum_{i=1}^{p} h_i\left(x^k\right)\nabla h_i\left(x^k\right) + k\sum_{j=1}^{q} g_j^+\left(x^k\right)\nabla g_j\left(x^k\right) + \left(x^k - x^*\right) \\
=& \nabla f\left(x^k\right) + \sum_{i=1}^{p} s_k\lambda_{k,i}\nabla h_i\left(x^k\right) + \sum_{j=1}^{q} s_k\mu_{k,j}\nabla g_j\left(x^k\right) + \left(x^k - x^*\right).
\end{aligned}$$

Then

$$\begin{aligned}
0 =& \lambda_{k,0}\nabla f\left(x^k\right) + \sum_{i=1}^{p}\lambda_{k,i}\nabla h_i\left(x^k\right) + \sum_{j=1}^{q}\mu_{k,j}\nabla g_j\left(x^k\right) + \frac{x^k - x^*}{s_k} \\
\to& \lambda_0\nabla f(x^*) + \sum_{i=1}^{p}\lambda_i\nabla h\left(x^*\right) + \sum_{j=1}^{q}\mu_j\nabla g\left(x^*\right).
\end{aligned}$$

On the other hand, $\mu_{k,j} \geq 0$ implies that $\mu_k \geq 0$. Finally, for any $j = 1, 2, \cdots, p$, if $g_j(x^*) = 0$, then $\mu_j g_j\left(x^*\right) = 0$; otherwise we have $g_j\left(x^k\right) < 0$ for k sufficiently large, which implies that $\mu_{k,j} = 0$ for k sufficiently large, and so $\mu_j = 0$ and $\mu_j g_j\left(x^*\right) = 0$.

(e) Recall that MFCQ $\Leftrightarrow$ NNAMCQ.

□

7. **Proof (2).** The KKT conditions for this problem are

$$\begin{cases} Px^* + q + & A^T\lambda^* = 0, \\ & Ax^* = b. \end{cases}$$

We derive

$$\begin{cases} x^* = & -P^{-1}A^T\lambda^* - P^{-1}q, \\ \lambda^* = & -(AP^{-1}A^T)^{-1}(AP^{-1}q + b). \end{cases}$$

8. **Proof:** Rewrite the problem as

$$\begin{aligned}
&\min_{x,r}\ \frac{1}{2}\left\|r\right\|_2^2 + \mu\left\|x\right\|_1, \\
&\text{s.t.}\quad Ax - b = r.
\end{aligned}$$

The Lagrange function is

$$\begin{aligned}
L(x, r, \lambda) &= \frac{1}{2}\left\|r\right\|_2^2 + \mu\left\|x\right\|_1 - \langle\lambda, Ax - b - r\rangle \\
&= \frac{1}{2}\left\|r\right\|_2^2 + \lambda^T r + \mu\left\|x\right\|_1 - (A^T\lambda)^T x + b^T\lambda.
\end{aligned}$$

$$g(\lambda) = \inf_{x,r} L(x, r, \lambda) = \begin{cases} b^T\lambda - \frac{1}{2}\left\|\lambda\right\|_2^2 & \text{if } \left\|A^T\lambda\right\|_\infty \leq \mu, \\ -\infty & \text{otherwise.} \end{cases}$$

Hence, the dual problem is

$$\begin{aligned}
&\max_{\lambda}\ b^T\lambda - \frac{1}{2}\left\|\lambda\right\|_2^2 \\
&\text{s.t.}\quad \left\|A^T\lambda\right\|_\infty \leq \mu.
\end{aligned}$$

9. **Proof (ii):** Notice that

$$\begin{aligned} x \in \bar{C} \Rightarrow & x^T A x + b^T x + c \leq 0 \text{ and } g^T x + \beta = 0 \\ \Rightarrow & x^T A x + b^T x + c \leq 0 \text{ and } \lambda x^T g g^T x + \lambda \beta g^T x = 0 \\ \Rightarrow & x^T (A + \lambda g g^T) x + (b + \lambda \beta g)^T x \leq c. \end{aligned}$$

By the positive semidefiniteness of $A + \lambda g g^T$, we have $x_1, x_2 \in \bar{C}$ implies that $\alpha x_1 + (1-\alpha) x_2 \in \bar{C}$ for $\alpha \in [0,1]$. Then

$$(\alpha x_1 + (1-\alpha)x_2)^T \left(A + \lambda g g^T\right)(\alpha x_1 + (1-\alpha)x_2) + (b + \lambda\beta g)^T (\alpha x_1 + (1-\alpha)x_2) \leq c,$$

or equivalently,

$$\begin{aligned} & (\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) + b^T (\alpha x_1 + (1-\alpha)x_2) \\ & + g^T (\alpha x_1 + (1-\alpha)x_2) \left(g^T (\alpha x_1 + (1-\alpha)x_2) + \beta\right) \leq c, \end{aligned}$$

Moreover, we have

$$\begin{aligned} x_1, x_2 \in \bar{C} \Rightarrow & g^T x_1 + \beta = 0 \text{ and } g^T x_2 + \beta = 0 \\ \Rightarrow & g^T (\alpha x_1 + (1-\alpha)x_2) + \beta = 0. \end{aligned}$$

Hence

$$(\alpha x_1 + (1-\alpha)x_2)^T A (\alpha x_1 + (1-\alpha)x_2) + b^T (\alpha x_1 + (1-\alpha)x_2) \leq c,$$

i.e., $(\alpha x_1 + (1-\alpha)x_2) \in C$.
Hence, $(\alpha x_1 + (1-\alpha)x_2) \in \bar{C}$.

10. **Proof**

(a) Obviously, $\theta(\mu) \in \mathbb{R}$ is nondecreasing on μ, so either $\lim_{\mu} \theta(\mu)$ exists or equals to infinity. Moreover, $\lim_{\mu} \theta(\mu) = \sup_{\mu} \theta(\mu)$.

(b) Let x^* be a local minimum, then $x^* \in X$ and for any μ,

$$\begin{aligned} \theta(\mu) &= \min_{x \in X} f(x) + \mu \left(\sum_{i=1}^{p} h_i^2(x) + \sum_{j=1}^{q} \max\{0, g_j(x)\}^2 \right) \\ &\leq f(x^*) + \mu \left(\sum_{i=1}^{p} h_i^2(x^*) + \sum_{j=1}^{q} \max\{0, g_j(x)\}^2 \right) \\ &= f(x^*) = f^*. \end{aligned}$$

So $\sup_{\mu} \theta(\mu) \leq f^*$.

(c) See that for any $\mu \geq 0$, we have $x_\mu \in X$ and

$$\theta(\mu) = f(x_\mu) + \mu \left(\sum_{i=1}^{p} h_i^2(x_\mu) + \sum_{j=1}^{q} \max\{0, g_j(x_\mu)\}^2 \right).$$

Let $0 \leq \mu_1 < \mu_2$, and define

$$P(x) = \sum_{i=1}^{p} h_i^2(x_\mu) + \sum_{j=1}^{q} \max\{0, g_j(x_\mu)\}^2.$$

Then we have

$$
\begin{aligned}
f\left(x_{\mu_1}\right)+\mu_1 P\left(x_{\mu_1}\right)= & \theta\left(x_{\mu_1}\right) \leq f\left(x_{\mu_2}\right)+\mu_1 P\left(x_{\mu_2}\right) \\
f\left(x_{\mu_2}\right)+\mu_2 P\left(x_{\mu_2}\right)= & \theta\left(x_{\mu_1}\right) \leq f\left(x_{\mu_1}\right)+\mu_2 P\left(x_{\mu_1}\right)
\end{aligned}
$$

Summing the above inequality,

$$
\left(\mu_1-\mu_2\right)\left[P\left(x_{\mu_1}\right)-P\left(x_{\mu_2}\right)\right] \leq 0.
$$

Notice that $\mu_1<\mu_2$, so $P\left(x_{\mu_1}\right)-P\left(x_{\mu_2}\right) \geq 0$. Then recall the first inequality,

$$
f\left(x_{\mu_1}\right)-f\left(x_{\mu_2}\right) \leq \mu_1\left[P\left(x_{\mu_2}\right)-P\left(x_{\mu_1}\right)\right] \leq 0.
$$