Suggested Solutions to Midterm Exam

1. (a)(b)-(f) F T F T T (g)-(k) T F F T T

2. (a) **True.**

   (b) **True.** Let

$$f(x) = \frac{1}{2}\|Ax - b\|^2 = \frac{1}{2}x^T(A^TA)x - x^TA^Tb + \frac{1}{2}b^Tb,$$

then

$$\nabla f(x) = A^TAx - A^Tb \text{ and } \nabla^2 f(x) = A^TA.$$

Note that A^TA is invertible (think of why) and $x^* = (A^TA)^{-1}A^Tb$ is the unique minimizer of f (think of why). For any initial point x^0, we have

$$x^1 = x^0 - \nabla^2 f(x^0)^{-1}\nabla f(x^0) = x^0 - (A^TA)^{-1}(A^TAx^0 - A^Tb) = x^*.$$

3. (a) Solve: Introduce slack variables and reformulate to the standard form

$$\begin{aligned} \min \quad & -4x_1 - 3x_2 - 5x_3 - 0x_4 - 0x_5 \\ s.t. \quad & 3x_1 + x_2 + 3x_3 + x_4 = 30 \\ & 2x_1 + 2x_2 + 3x_3 + x_5 = 40 \\ & x_1,\, x_2,\, x_3,\, x_4,\, x_5 \geq 0 \end{aligned}$$

the starting basis $B = [a_4,\, a_5]$, and BFS

$$x = [0,\, 0,\, 0,\, B^{-1}b] = [0,\, 0,\, 0,\, 30,\, 40]^T$$

reduced costs

$$\bar{c}^T := c^T - c_B^TB^{-1}A = [-4,\, -3,\, -5,\, 0,\, 0]^T$$

notice $c_1 = -4 < 0, c_2 = -3 < 0, c_3 = -5 < 0$, current solution is no optimal and bring $j = 3$ (or $1, 2$) into basis.
compute edge direction: $\delta^{(3)} = [0,\, 0,\, 1,\, -B^{-1}a_3]^T = [0,\, 0,\, 1,\, -3,\, -3]^T$
$\delta_B^{(3)} = [-3,\, -3]^T \ngeq 0$, so this edge direction is not unbounded.
ratio test

$$\alpha_{\min} = \min\left\{ \frac{x_i}{-\delta_i^{(j)}} \middle| i \in B,\, \delta_i^{(j)} < 0 \right\} = \min\{10,\, 40/3\} = 10$$

the "min" is achieved at $i' = 4$, so remove a_4 from basis.

updated basis: $B = [a_3,\, a_5]$ and BFS:

$$x = [0,\, 0,\, 10,\, 0,\, 10]^T$$

reduced costs

$$\bar{c}^T := c^T - c_B^TB^{-1}A = [1,\, -4/3,\, 0,\, 5/3,\, 0]^T$$

notice $\bar{c}_2 < 0$, current solution is no optimal and bring $j = 2$ into basis.
compute edge direction: $\delta^{(2)} = [0,\, 1,\, -1/3,\, 0,\, -5/3]^T$

$\delta_B^{(2)} = [-1/3,\ -5/3]^T \ngeq 0$, so this edge direction is not unbounded.
ratio test

$$\alpha_{\min} = \min\left\{ \frac{x_i}{-\delta_i^{(j)}} \middle| i \in B,\ \delta_i^{(j)} < 0 \right\} = \min\{30,\ 10\} = 10$$

the "min" is achieved at $i^{'} = 5$, so remove a_5 from basis.

updated basis: $B = [a_2,\ a_3]$ and BFS:

$$x = [0,\ 10,\ 20/3,\ 0,\ 0]^T$$

reduced costs

$$\bar{c}^T := c^T - c_B^T B^{-1} A = [-1/3,\ 0,\ 0,\ 1/3,\ 4/3]^T$$

notice $\bar{c}_1 < 0$, current solution is no optimal and bring $j = 1$ into basis.
compute edge direction: $\delta^{(1)} = [1,\ 1,\ -4/3,\ 0,\ 0]^T$
$\delta_B^{(2)} = [1,\ -4/3]^T \ngeq 0$, so this edge direction is not unbounded.
remove a_3 form basis.

updated basis: $B = [a_1,\ a_2]$ and BFS:

$$x = [5,\ 15,\ 0,\ 0,\ 0]^T$$

reduced costs

$$\bar{c}^T := c^T - c_B^T B^{-1} A = [0,\ 0,\ 1/4,\ 1/2,\ 5/4]^T \geq 0.$$

(b) Let

$$f(x,y) = x^2 + y^2 - \ln\left(x^2 y^2\right);$$
$$g_1(x,y) = x - \ln y \text{ and } g_2(x,y) = -x + 1 \text{ and } g_3(x,y) = -y + 1.$$

Just solve the KKT system w.r.t $(x, y, \lambda_1, \lambda_2, \lambda_3)$:

$$\begin{cases} 0 \leq g_1(x,y), 0 \leq g_2(x,y), 0 \leq g_3(x,y); \\ \lambda_1 \geq 0, \lambda_2 \geq 0, \lambda_3 \geq 0; \\ 0 = \lambda_1 g_1(x,y), 0 = \lambda_2 g_2(x,y), 0 = \lambda_3 g_3(x,y); \\ 0 = \nabla f\,(x,y) + \lambda_1 \nabla g_1(x,y) + \lambda_2 \nabla g_2(x,y) + \lambda_3 \nabla g_3(x,y) \end{cases}$$

i.e.,

$$\begin{cases} \lambda_1 \geq 0, x - \ln y \leq 0, \lambda_1(x - \ln y) = 0; \\ \lambda_2 \geq 0, -x + 1 \leq 0, \lambda_2(-x + 1) = 0; \\ \lambda_3 \geq 0, -y + 1 \leq 0, \lambda_3(-y + 1) = 0; \\ 0 = \begin{pmatrix} 2x - \frac{2}{x} \\ 2y - \frac{2}{y} \end{pmatrix} + \lambda_1 \begin{pmatrix} 1 \\ -\frac{1}{y} \end{pmatrix} + \lambda_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 2x - \frac{2}{x} + \lambda_1 - \lambda_2 \\ 2y - \frac{2+\lambda_1}{y} - \lambda_3 \end{pmatrix} \end{cases}$$

Solve:

- Case 1: $\lambda_1 = \lambda_2 = \lambda_3 = 0$. Then $x = y = 1$ while $g_1(1,1) = 1 > 0$.
- Case 2: $\lambda_1 = \lambda_2 = -y + 1 = 0$. Then $\lambda_3 = 0$ and $x = y = 1$ while $g_1(1,1) = 1 > 0$.
- Case 3: $\lambda_1 = -x + 1 = \lambda_3 = 0$. Then $\lambda_2 = 0$ and $x = y = 1$ while $g_1(1,1) = 1 > 0$.
- Case 4: $\lambda_1 = -x + 1 = -y + 1 = 0$. Then $\lambda_2 = \lambda_3 = 0$ and $x = y = 1$ while $g_1(1,1) = 1 > 0$.
- Case 5: $x - \ln y = \lambda_2 = \lambda_3 = 0$. Then $x \approx 0.47921 < 1$ and so $g_2(0.47921, \cdot) > 0$.
- Case 6: $x - \ln y = \lambda_2 = -y + 1 = 0$. No solution.
- Case 7: $x - \ln y = -x + 1 = \lambda_3 = 0$. Then $x = 1, y = e$ and $\lambda_1 = \lambda_2 = 2e^2 - 2$.
- Case 8: $x - \ln y = -x + 1 = -y + 1 = 0$. No solution.

Hence the unique solution is $(1, e)$.

4. *Proof.* (i). Pick any $\bar{x} \neq x \in \Omega$ and set $\bar{\alpha} = 1$, then for any $0 < \alpha \leq \bar{\alpha} = 1$ we have

$$\bar{x} + \alpha(x - \bar{x}) = \alpha x + (1-\alpha)\bar{x} \in \Omega$$

by the convexity of Ω. Hence $x - \bar{x} \in F(\bar{x})$.

(ii). i. $F(\bar{x}) \cup \{0\}$ is convex.
For any $d_1, d_2 \in F(\bar{x} \cup \{0\})$ and $0 \leq \lambda \leq 1$, WTS $\lambda d_1 + (1-\lambda)d_2 \in F(\bar{x}) \cup \{0\}$.

- case 1: $d_1 = 0$ or $d_2 = 0$. Trivial.
- case 2: $\lambda d_1 + (1-\lambda)d_2 = 0$. Trivial.
- case 3. $d_1 \neq 0$ and $d_2 \neq 0$ and $\lambda d_1 + (1-\lambda)d_2 \neq 0$. Then $\exists \bar{\alpha}_1, \bar{\alpha} > 0$ such that

$$\begin{aligned} &\bar{x} + \alpha d_1 \in \Omega \text{ for any } 0 < \alpha \leq \bar{\alpha}_1; \\ &\bar{x} + \alpha d_2 \in \Omega \text{ for any } 0 < \alpha \leq \bar{\alpha}_2; \end{aligned}$$

Then for any $0 < \alpha \leq \min\{\bar{\alpha}_1, \bar{\alpha}_2\}$, we have

$$\bar{x} + \alpha\left[\lambda d_1 + (1-\lambda_2)d_2\right] = \lambda\left(\bar{x} + \alpha d_1\right) + (1-\lambda)(\bar{x} + \alpha d_2) \in \Omega,$$

where the inclusion holds since $d_1, d_2 \in F(\bar{x})$ and Ω is convex. Hence $\lambda d_1 + (1-\lambda_2)d_2 \in F(\bar{x})$.

ii. $F(\bar{x}) \cup \{0\}$ is a cone. For any $d \in F(\bar{x}) \cup \{0\}$, just show that $\lambda d \in F(\bar{x}) \cup \{0\}$ for $\lambda \geq 0$. Note that the case where $d = 0$ or $\lambda = 0$ is trivial. So let $\lambda > 0$ and $0 \neq d \in F(\bar{x}) \cup \{0\}$, or equivalently, $0 \neq d \in F(\bar{x})$, then we have $\bar{x} + \alpha d \in \Omega$ for $0<\alpha<\bar{\alpha}$. Hence we have $\bar{x} + \alpha(\lambda d)$ for $0 < \alpha < \frac{\bar{\alpha}}{\lambda}$, i.e., $\lambda d \in F(\bar{x})$.

(iii). For any $\bar{x} \neq x \in \Omega$, (i) implies that $x - \bar{x} \in F(\bar{x})$ and hence $x - \bar{x} \notin D(\bar{x})$. Hence $\nabla f(\bar{x})^T(x - \bar{x}) \geq 0$, which is exactly the FOC.

(iv). Note that

$$\text{cone} F(\bar{x}) \cap D(\bar{x}) = \text{cone}\left(F(\bar{x}) \cup \{0\}\right) \cap D(\bar{x}) = \left(F(\bar{x}) \cup \{0\}\right) \cap D(\bar{x}) = F(\bar{x}) \cap D(\bar{x}),$$

where the first equality holds by the definition of convex cone, the second equality holds by the hint of (iv) and the last equality holds by the distributive property of set operations.
On the other hand, we assume that $F(\bar{x}) \cap D(\bar{x}) \neq \varnothing$. Choose $d \in F(\bar{x}) \cap D(\bar{x})$. The fact that $d \in F(\bar{x})$ implies that there exist $\bar{\alpha} > 0$ such that $\bar{x} + \bar{\alpha} d \in \Omega$. Then $\nabla f(\bar{x})^T(\bar{x} + \bar{\alpha} d - \bar{x}) \geq 0$ by FOC, so $\nabla f(\bar{x})^T d \geq 0$, which implies that $d \notin D(\bar{x})$. Contradiction.

□

5. *Proof.* (a) $x^{k+1} = x^k - \frac{1}{L_f}\nabla f\left(x^k\right)$.

(b) Use the descent lemma:

$$\begin{aligned} f\left(x^{k+1}\right) \leq & f\left(x^k\right) + \nabla f\left(x^k\right)^T\left(x^{k+1} - x^k\right) + \frac{L_f}{2}\left\|x^{k+1} - x^k\right\|^2 \\ = & f\left(x^k\right) - \frac{1}{L_f}\left\|\nabla f\left(x^k\right)\right\|^2 + \frac{1}{2L_f}\left\|\nabla f\left(x^k\right)\right\|^2 \\ = & f\left(x^k\right) - \frac{1}{2L_f}\left\|\nabla f\left(x^k\right)\right\|^2. \end{aligned}$$

Then

$$f\left(x^k\right) - f\left(x^{k+1}\right) \geq \frac{1}{2L_f}\left\|\nabla f\left(x^k\right)\right\|^2.$$

(c)

$$
\begin{aligned}
\min_{n=0,1,\cdots,k} \|\nabla f(x^n)\| =& \sqrt{\min_{n=0,1,\cdots,k} \|\nabla f(x^n)\|^2} \leq \sqrt{\frac{\sum_{n=0}^{k} \|\nabla f(x^n)\|^2}{k+1}} \\
\leq& \sqrt{\frac{2L_f \sum_{n=0}^{k} \left(f(x^n) - f(x^{n+1})\right)}{k+1}} \\
=& \sqrt{\frac{2L_f \left(f(x^0) - f(x^{k+1})\right)}{k+1}} \\
\leq& \sqrt{\frac{2L_f \left(f(x^0) - f^*\right)}{k+1}}.
\end{aligned}
$$

$\square$