SSE5107 Optimization Theory and Algorithms
Mid-term quiz

Problem 1

1. True
2. False (Consider one dimension: $f(x) = ax^2$ is convex iff $a > 0$)
3. True
4. False
5. True

Problem 2

1. AABAA

(c) counterexample: consider two dimension: $\{(x_1, x_2) \mid x_1 x_2 \geq 0\}$ satisfies the definition, but is not convex. See Figure 1.
(d) $x^T W x \leq (c^T x)^2, c^T x \leq 0 \Rightarrow ||W^{1/2}x||_2 \leq -c^T x, c^T x \leq 0$. Here $||W^{1/2}x||_2 \leq -c^T x$ can be considered as affine mapping of a norm cone $||y|| \leq t, y = W^{1/2}x, t = -c^T x$, and thus, is convex.
(e) intersection of halfspaces is convex.

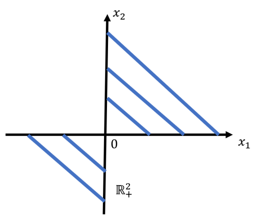

Figure 1:

2. ACAAB

(b) Let $\alpha \in [0, 1]$. Notice that

$$f(\alpha X_1 + (1-\alpha_1)X_2) = \mathrm{Tr}(A(\alpha X_1 + (1-\alpha)X_2)B) = \alpha_1 \mathrm{Tr}(AX_1B) + (1-\alpha)\mathrm{Tr}(AX_2B) = \alpha f(X_1) + (1-\alpha)f(X_2).$$

This helps verify the convexity and concavity as follows:
Convexity: $f(\alpha X_1 + (1-\alpha_1)X_2) \leq \alpha f(X_1) + (1-\alpha_1)f(X_2)$.
Concavity: Consider $-f$: $-f(\alpha X_1 + (1-\alpha_1)X_2) \leq -\alpha f(X_1) - (1-\alpha_1)f(X_2)$. So $-f$ is convex $\Rightarrow$ f is concave.
(c) Consider one dimension: $n = 1$. $f(X) = \mathrm{Tr}(AX^{-1}), \textbf{dom}\ f = \mathcal{S}^n_{++}, A \in f = \mathcal{S}^n_+ \Rightarrow f(x) = \frac{a}{x}, a > 0, x > 0$.
So far this should be enough to verify its convexity given the question is a just single-choice question.
Attach proof for general arbitrary dimension as a reference:

Let $X_0 \in \mathcal{S}_{++}^n$ and $H \in \mathcal{S}^n$ be fixed. For any $t \in \mathbb{R}$ such that $X_0 + tH \in \mathcal{S}_{++}^n$, we have

$$g(t) \equiv \operatorname{tr}\left(A\left(X_0 + tH\right)^{-1}\right) = \operatorname{tr}\left(AX_0^{-1/2}\left(I + tX_0^{-1/2}HX_0^{-1/2}\right)^{-1}X_0^{-1/2}\right)$$

Since $X_0^{-1/2}HX_0^{-1/2} \in \mathcal{S}^n$, there exist an orthogonal matrix $U \in \mathbb{R}^{n\times n}$ and a diagonal matrix $\Lambda \in \mathcal{S}^n$ such that $X_0^{-1/2}HX_0^{-1/2} = U\Lambda U^T$ (eigenvalue decomposition). Hence, we have

$$g(t) = \operatorname{tr}\left(AX_0^{-1/2}U(I + t\Lambda)^{-1}U^TX_0^{-1/2}\right) = \operatorname{tr}\left(U^TX_0^{-1/2}AX_0^{-1/2}U(I + t\Lambda)^{-1}\right)$$

Now, let $\Lambda = \operatorname{diag}\left(\lambda_1, \ldots, \lambda_n\right)$. Since $X_0 + tH \in \mathcal{S}_{++}^n$, we have $1 + t\lambda_i > 0$ for $i = 1, \ldots, n$. Thus, we may write $(I + t\Lambda)^{-1} = \operatorname{diag}\left(\left(1 + t\lambda_1\right)^{-1}, \ldots, \left(1 + t\lambda_n\right)^{-1}\right)$. It then follows that

$$g(t) = \sum_{i=1}^{n}\left(U^TX_0^{-1/2}AX_0^{-1/2}U\right)_{ii} \cdot \frac{1}{1 + t\lambda_i}$$

It is easy to verify that for $i = 1, \ldots, n$, the function $t \mapsto \left(1 + t\lambda_i\right)^{-1}$ is convex over the region $\left\{t \in \mathbb{R} : t > \lambda_i^{-1} \text{ for } i = 1, \ldots, n\right\}$. Moreover, since $U^TX_0^{-1/2}AX_0^{-1/2}U \in \mathcal{S}_+^n$, we have $\left(U^TX_0^{-1/2}AX_0^{-1/2}U\right)_{ii} \geq 0$ for $i = 1, \ldots, n$. It follows that g is a non-negative linear combination of convex functions, which implies that g is convex. This in turn implies that f is convex.
(d) conjugate function is always convex.

Problem 3

1. - Introduce slack variable $s \succeq 0$:

$$\begin{aligned} \min_{x\in\mathbb{R}^n, s\in\mathbb{R}^k} \quad & c^Tx \\ s.t. \quad & Ax = b \\ & Gx + s = h \\ & s \succeq 0 \end{aligned}$$

   - Decompose x as $x^+ - x^-$, where $x^+ \succeq 0$ and $x^- \succeq 0$:

$$\begin{aligned} \min_{x^+, x^-\in\mathbb{R}^n, s\in\mathbb{R}^k} \quad & c^T(x^+ - x^-) \\ s.t. \quad & Ax^+ - Ax^- = b \\ & Gx^+ - Gx^- + s = h \\ & x^+ \succeq 0,\ x^- \succeq 0,\ s \succeq 0 \end{aligned}$$

2. • Introduce slack variable $S \succeq 0$:

$$\begin{aligned}
\min_{x \in \mathbb{R}^n, S \in \mathbb{S}^k} \quad & c^T x \\
s.t. \quad & Ax = b \\
& \sum_{i=1}^n F_i x_i + S = -G \\
& S \succeq 0
\end{aligned}$$

• Decompose x as $x^+ - x^-$, where $x^+ \succeq 0$ and $x^- \succeq 0$:

$$\begin{aligned}
\min_{x^+, x^- \in \mathbb{R}^n, S \in \mathbb{S}^k} \quad & c^T (x^+ - x^-) \\
s.t. \quad & Ax^+ - Ax^- = b \\
& \sum_{i=1}^n F_i (x_i^+ - x_i^-) + S = -G \\
& x^+ \succeq 0,\ x^- \succeq 0,\ S \succeq 0
\end{aligned}$$

• Construct a large variable consisting of x^+, x^-, and S:

$$\begin{aligned}
\min_{x^+, x^- \in \mathbb{R}^n, S \in \mathbb{S}^k} \quad & c^T (x^+ - x^-) \\
s.t. \quad & Ax^+ - Ax^- = b \\
& \sum_{i=1}^n F_i (x_i^+ - x_i^-) + S = -G \\
& \begin{bmatrix} x_1^+ & & & & & & \\ & \dots & & & & & \\ & & x_n^+ & & & & \\ & & & x_1^- & & & \\ & & & & \dots & & \\ & & & & & x_n^- & \\ & & & & & & S \end{bmatrix} \succeq 0
\end{aligned}$$

Problem 4

1. $\nabla f(x^*)^T (x - x^*) \geq 0$ for all $x \in \mathcal{X}$.

2. By convexity of f, we have $f(x) \geq f(x^*) + \nabla f(x^*)^T (x - x^*)$, for all y. Since $\nabla f(x^*)^T (x - x^*) \geq 0$, we have $f(x) \geq f(x^*)$ for all y, which shows the optimality of x^*.

3. $\nabla f_i(x^*) \begin{cases} \geq 0, & x_i^* = l_i, \\ = 0, & x_i^* \in (l_i, u_i), \\ \leq 0, & x_i^* = u_i. \end{cases}$