Problem 1 [BV 3.1]

(a) Let $\theta = \frac{b-x}{b-a}$, using convexity of f:

$$
\begin{aligned}
f(\theta a + (1-\theta)b) &\leq \theta f(a) + (1-\theta) f(b) \\
f(\frac{b-x}{b-a}a + \frac{x-a}{b-a}b) &\leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \\
f(x) &\leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)
\end{aligned}
$$

(b) LHS inequality can be derived by subtracting $f(a)$ from inequality of part (a):

$$
\begin{aligned}
f(x) - f(a) &\leq \frac{b-x}{b-a} f(a) - f(a) + \frac{x-a}{b-a} f(b) \\
f(x) - f(a) &\leq \frac{a-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) \\
f(x) - f(a) &\leq \frac{x-a}{b-a}(f(b) - f(a)) \\
\frac{f(x) - f(a)}{x-a} &\leq \frac{f(b) - f(a)}{b-a}
\end{aligned}
$$

RHS inequality can be derived by adding $f(b)$ to the negative of inequality of part (a):

$$
\begin{aligned}
\frac{x-b}{b-a} f(a) + \frac{a-x}{b-a} f(b) + f(b) &\leq f(b) - f(x) \\
\frac{x-b}{b-a} f(a) + \frac{b-x}{b-a} f(b) &\leq f(b) - f(x) \\
\frac{b-x}{b-a}(f(b) - f(a)) &\leq f(b) - f(x) \\
\frac{f(b) - f(a)}{b-a} &\leq \frac{f(b) - f(x)}{b-x}
\end{aligned}
$$

Therefore, we have shown that:

$$
\frac{f(x) - f(a)}{x-a} \leq \frac{f(b) - f(a)}{b-a} \leq \frac{f(b) - f(x)}{b-x}
$$

And it can be seen in the figure below as the slop of each line segment is shown:

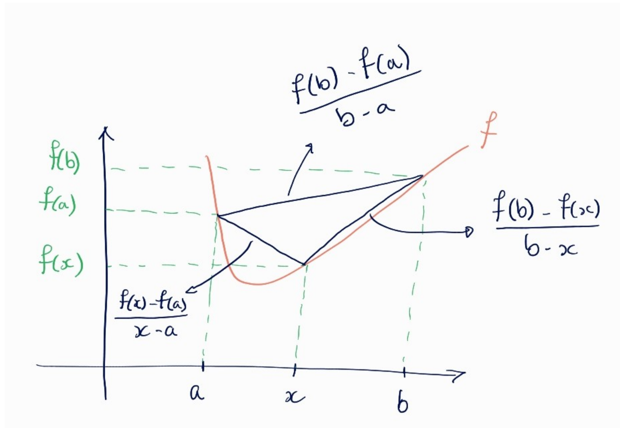

Figure 1: Slope of line segments on the convex function.

(c) Taking limit for $x \to a$ on both sides of LHS inequality of part (b):

$$\lim_{x\to a} \frac{f(x)-f(a)}{x-a} \leq \lim_{x\to a} \frac{f(b)-f(a)}{b-a}$$
$$f'(a) \leq \frac{f(b)-f(a)}{b-a}$$

Taking limit for $x \to b$ on both sides of RHS inequality of part (b):

$$\lim_{x\to b} \frac{f(b)-f(a)}{b-a} \leq \lim_{x\to b} \frac{f(b)-f(x)}{b-x}$$
$$\frac{f(b)-f(a)}{b-a} \leq f'(b)$$

Therefore, we have shown that:

$$f'(a) \leq \frac{f(b)-f(a)}{b-a} \leq f'(b)$$

(d) Using the inequality of part (c):

$$
\begin{aligned}
f'(b) - f'(a) &\geq 0 \\
\frac{f'(b) - f'(a)}{b - a} &\geq 0 \\
\lim_{b \to a} \frac{f'(b) - f'(a)}{b - a} &\geq \lim_{b \to a} 0 \\
f''(a) &\geq 0
\end{aligned}
$$

Problem 2 [BV 3.4]

$f : \mathbb{R}^n \to \mathbb{R}$ continuous

$\boxed{\Rightarrow}$ $\underline{f \text{ is convex} \Rightarrow \int_0^1 f(\lambda y + (1-\lambda)x) d\lambda \leq \frac{f(x)+f(y)}{2}}$

Since f is convex:

$$
\begin{aligned}
f(\lambda y + (1-\lambda)x) &\leq \lambda f(y) + (1-\lambda) f(x) \\
\int_0^1 f(\lambda y + (1-\lambda)x) d\lambda &\leq \int_0^1 (\lambda f(y) + (1-\lambda) f(x)) d\lambda \\
\int_0^1 f(\lambda y + (1-\lambda)x) d\lambda &\leq \frac{1}{2} f(y) + \frac{1}{2} f(x)
\end{aligned}
$$

$\boxed{\Leftarrow}$ $\underline{\int_0^1 f(\lambda y + (1-\lambda)x) d\lambda \leq \frac{f(x)+f(y)}{2} \Rightarrow f \text{ is convex}}$

(proof by contradiction) Suppose f is not convex, then $\exists\ x, y \in \mathbb{R}^n$ and $\theta_0 \in (0,1)$ such that

$$
f(\theta_0 x + (1-\theta_0)y) > \theta_0 f(x) + (1-\theta_0) f(y)
$$

Let continuous function $F(\theta)$ be defined as

$$
F(\theta) = f(\theta x + (1-\theta)y) - \theta f(x) - (1-\theta) f(y)
$$

where $F(0) = F(1) = 0$ and $F(\theta_0) > 0$. Let a and b denote the nearest zero-crossings of F around θ_0, then $F(a) = 0$ and $F(b) = 0$ and $F(\theta') > 0$ for all $\theta' \in (a, b)$. As a result, F is strictly positive over the line segment between $ax + (1-a)y := x'$ and $bx + (1-b)y := y'$. Therefore $F(\gamma x' + (1-\gamma)y') > 0$ for all $\gamma \in (0, 1)$, which means that $f(\theta_0 x + (1-\theta_0)y)$.
As $\gamma x' + (1-\gamma)y' = \gamma(ax + (1-a)y) + (1-\gamma)(bx + (1-b)y) = (\gamma a + b - \gamma b)x + (1 - \gamma a + \gamma b - b)y$ indicates a line segment between x and y, with change of notation we have:

$$\int_0^1 f(\zeta x + (1-\zeta)y)d\zeta > \int_0^1 (\zeta f(x) + (1-\zeta)f(y))d\zeta = \frac{f(x) + f(y)}{2}$$

which is a contradiction with the hypothesis.

Problem 3 [BV 3.7]

(proof by contradiction) Suppose f is not constant, then $\exists\, x, y \in \mathbb{R}^n$ such that $f(x) < f(y)$, and let the function g be defined as $g(t) = f(x + \theta(y - x))$, which is convex. We also have $g(0) < g(1)$. For all $t > 1$:

$$g(1) \leq \frac{t-1}{t} g(0) + \frac{1}{t} g(t)$$

Then

$$g(t) \geq tg(1) - (t-1)g(0) = g(0) + t(g(1) - g(0)),$$

as $t \to \infty$, $g \to \infty$, meaning that f is unbounded. This is a contradiction with our hypothesis.

Problem 4 [BV 9.1]

(a) f is twice differentiable, its hessian is equal to $\nabla^2 f(x) = P$. Since $P \not\succeq 0$ then f is not convex by definition. On the other hand, since symmetric matrix is not positive semi-definite, then $\exists\, v \in \mathbb{R}^n$ such that $v^T P v < 0$, then:

$$\lim_{a \to \infty} f(av) = \lim_{a \to \infty} \frac{1}{2} a^2 v^T P v + q^T v + r = -\infty$$

therefore, f is unbounded below in this case.

(b) Since P is positive semi-definite and P is rank deficient, then $\dim \operatorname{null} P > 0$. Therefore $\exists\, u \in \mathbb{R}^n$ such that $Pu = 0$, and also obviously $q^T u \neq 0$, then:

$$f(au) = \frac{1}{2}a^2 u^T P u + q^T u + r = a q^T u + r$$

therefore, depending on the sign of $q^T u$, f is unbounded below either $a \to \infty$ or $a \to -\infty$.

Problem 5 [BV 9.2]

(a) Since $\operatorname{dom} f$ is an open set, if $f(x)$ on the $\operatorname{bd} \operatorname{dom} f$ goes to infinity, then f is closed. The numerator $\|Ax - b\|_2^2$ is lower bounded by some value greater than zero, since $b \notin \operatorname{range} A$.
On the other hand, boundary of $\operatorname{dom} f$ is designated by equation $c^T x + d = 0$, which makes the denominator of f to go to zero.
Therefore value of $f(x)$, as x approaches the boundary of the $\operatorname{dom} f$, goes to infinity.
Hence f is closed.

(b) To have $\nabla f(x)|_{x^\star} = 0$:

$$\begin{aligned}\nabla f(x) &= \frac{2A^T(Ax-b)(c^T x - d) - \|Ax-b\|_2^2 c}{(c^T x - d)^2}\\ &= \frac{2}{c^T x - d}(A^T A x - A^T b) - \frac{\|Ax-b\|_2^2}{(c^T x - d)^2} c = 0\end{aligned}$$

having $A_{m\times n}$ full rank $(n < m)$, multiplying both sides of equality from left, by

$(A^TA)^{-1}$:

$$\begin{aligned}\frac{2}{c^Tx-d}(x-(A^TA)^{-1}A^Tb)-\frac{\|Ax-b\|_2^2}{(c^Tx-d)^2}(A^TA)^{-1}c=0\\ \frac{2}{c^Tx-d}(x-x_1)-\frac{\|Ax-b\|_2^2}{(c^Tx-d)^2}x_2=0\\ \frac{2}{c^Tx-d}(x-x_1)=\frac{\|Ax-b\|_2^2}{(c^Tx-d)^2}x_2\\ x-x_1=\frac{1}{2}\frac{\|Ax-b\|_2^2}{c^Tx-d}x_2\\ x=x_1+\frac{1}{2}\frac{\|Ax-b\|_2^2}{c^Tx-d}x_2\end{aligned}$$

therefore by definition of $x=x_1+tx_2$, we have $t=\frac{1}{2}\frac{\|Ax-b\|_2^2}{c^Tx-d}$:

$$\begin{aligned}t&=\frac{1}{2}\frac{\|Ax-b\|_2^2}{c^Tx-d}\\ &=\frac{1}{2}\frac{\|A(x_1+tx_2)-b\|_2^2}{c^T(x_1+tx_2)-d}\\ &=\frac{1}{2}\frac{\|Ax_1-b+tAx_2\|_2^2}{c^Tx_1+tc^Tx_2-d}\end{aligned}$$

The last equation can be written as:

$$(c^Tx_2)t^2+(2c^Tx_1-d)t-\|Ax_1-b\|_2^2=0$$

therefore, t can be calculated by solving the quadratic equation above while satisfying $c^T(x_1+tx_2)>0$ to have $x^\star\in \mathrm{dom} f$.