**CSE 203B W25 Midterm 10AM 2/23/2025 - 10AM 2/25/2025**

Submit your solution to Gradescope before the due time (no late submission).

**Policy of the Exam:**

1. This is an open-book take-home exam. Internet search is permitted. However, you are required to work by yourself. Consultation or discussion with any other parties is not allowed.
2. You are not required to typeset your solutions. We do expect your writing to be legible and your final answers clearly indicated. Also, please allow sufficient time to upload your solutions.
3. You are allowed to check your answers with programs in Matlab, CVX, Mathematica, Maple, NumPy, etc. Be aware that these programs may not produce the intermediate steps needed to receive credit.
4. If something is unclear, state the assumptions that seem most natural to you and proceed under those assumptions. Out of fairness, we will not be answering questions about the technical content of the exam on Piazza or by email. The solution will then be graded based on the reasonable assumptions made.

**Part I: True or False: Briefly explain your answer** (30 pts, 3 pts each)

I.1 Convex Set

The set $S = \{(x, y) \in \mathbb{R}^2 \mid y > 0, \frac{(x+2y)^2}{y} < x - y + 3\}$ is convex.

**Answer: True**

After rearrangement, we have the inequality $f(x, y) = x^2 + 5y^2 + 3xy - 3y < 0$. We also know that the Hessian matrix $\begin{bmatrix} 2 & 3 \\ 3 & 10 \end{bmatrix}$ for $f(x, y)$ is positive definite. Since $f(x, y)$ is a convex function, its sublevel set $S$ is convex.

I.2 Matrix Solver

If $\hat{\mathbf{x}}$ is an approximate solution to $\mathbf{Ax} = \mathbf{b}$, then the relative residual $\frac{\|\mathbf{b}-\mathbf{A}\hat{\mathbf{x}}\|_2}{\|\mathbf{b}\|_2}$ is always larger than the relative error $\frac{\|\hat{\mathbf{x}}-\mathbf{x}\|_2}{\|\mathbf{x}\|_2}$ for any matrix $\mathbf{A}$.

**Answer: False**

We also need to consider the condition number for matrix $A$.

I.3 Support Vector Machine

Given a set of points $\{(x_i, y_i) \mid i = 1, ..., m\}$, where $x_i \in \mathbb{R}^n$ and $y_i \in \{-1, 1\}$, we find a hyperplane with vector $a \in \mathbb{R}^n$ and bias $b \in \mathbb{R}$ by solving the optimization problem: $\min_{\mathbf{a},b} ||\mathbf{a}||_2^2$, $\mathbf{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$ subject to $y_i(\mathbf{a}^T x_i - b) \geq 1$, $\forall i = 1, ..., m$. To have a valid solution, the margin, defined as the distance from the hyperplane to the closest point across both classes, is at least 1.

**Answer: False**

The margin is defined as $\frac{1}{||\mathbf{a}||_2}$. If we consider the total width of the margin (sum of both sides), it will be $\frac{2}{||\mathbf{a}||_2}$. Thus, the margin might be smaller than 1.

I.4 Dual Cone

$K = \{\theta_1 u_1 + \theta_2 u_2 \mid u_1 = [3, -2]^T, u_2 = [1, 1]^T, \theta_1 \geq 0, \theta_2 \geq 0\}$, its dual cone is $K^* = \{x_1 u_1 + x_2 u_2 \mid u_1 = [2, 3]^T, u_2 = [-1, 1]^T, x_1 \geq 0, x_2 \geq 0\}$.
T/F:
**Answer: False**
Step 1: Compute the inner product between points $x \in K$ and $y \in K^*$

A point $x \in K$ is given by:

$$x = \theta_1 u_1 + \theta_2 u_2 = \theta_1 \begin{bmatrix} 3 \\ -2 \end{bmatrix} + \theta_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3\theta_1 + \theta_2 \\ -2\theta_1 + \theta_2 \end{bmatrix}$$

A point $y \in K^*$ is given by:

$$y = x_1 u_1 + x_2 u_2 = x_1 \begin{bmatrix} 2 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2x_1 - x_2 \\ 3x_1 + x_2 \end{bmatrix}$$

Now, we compute the inner product $\langle x, y \rangle$:

$$\langle x, y \rangle = \begin{bmatrix} 3\theta_1 + \theta_2 \\ -2\theta_1 + \theta_2 \end{bmatrix} \cdot \begin{bmatrix} 2x_1 - x_2 \\ 3x_1 + x_2 \end{bmatrix}$$

Expanding the dot product:

$$\langle x, y \rangle = (3\theta_1 + \theta_2)(2x_1 - x_2) + (-2\theta_1 + \theta_2)(3x_1 + x_2)$$

Distribute and simplify:

$$= (3\theta_1)(2x_1) + (3\theta_1)(-x_2) + (\theta_2)(2x_1) + (\theta_2)(-x_2) + (-2\theta_1)(3x_1) + (-2\theta_1)(x_2) + (\theta_2)(3x_1) + (\theta_2)(x_2)$$

$$= 6\theta_1 x_1 - 3\theta_1 x_2 + 2\theta_2 x_1 - \theta_2 x_2 - 6\theta_1 x_1 - 2\theta_1 x_2 + 3\theta_2 x_1 + \theta_2 x_2$$

Combine like terms:

$$= (6\theta_1 x_1 - 6\theta_1 x_1) + (2\theta_2 x_1 + 3\theta_2 x_1) + (-3\theta_1 x_2 - 2\theta_1 x_2) + (-\theta_2 x_2 + \theta_2 x_2)$$

$$= 5\theta_2 x_1 - 5\theta_1 x_2$$

Thus, the inner product becomes:

$$\langle x, y \rangle = 5(\theta_2 x_1 - \theta_1 x_2)$$

Step 2: Check if the inner product satisfies the condition

For the dual cone condition to hold, the inner product must be non-negative:

$$5(\theta_2 x_1 - \theta_1 x_2) \geq 0$$

Since $\theta_1 \geq 0, \theta_2 \geq 0, x_1 \geq 0, x_2 \geq 0$, we cannot guarantee that this inequality will always hold, as the terms $\theta_2 x_1$ and $\theta_1 x_2$ might not be in the correct relationship to satisfy this inequality.

The given dual cone $K^*$ does not satisfy the condition for being the dual cone of $K$, therefore the statement is **False**.

I.5 Convex Function

If $f_1(x)$ and $f_2(x)$ are convex functions, their weighted sum $f(x) = w_1 f_1(x) + w_2 f_2(x)$ is **always convex**, where $w_1, w_2 \in R$.

T/F:

**Answer: False**

In general, the weighted sum of convex functions $f(x) = w_1 f_1(x) + w_2 f_2(x)$ is convex **only if** the weights $w_1$ and $w_2$ are **non-negative** $(w_1, w_2 \geq 0)$. If either $w_1$ or $w_2$ is negative, the weighted sum may not remain convex.

This is because convexity is preserved under non-negative linear combinations of convex functions, but not necessarily under arbitrary real-weighted sums.

I.6 Conjugate Function

Given function $f(x) = x_1^2 + 5x_1x_2 - x_2^2$, where $x \in \mathbb{R}^2$, then the conjugate of the conjugate function, $f^{**}(x)$, is equal to itself, i.e., $f^{**}(x) = f(x)$.

T/F:

**Answer: False**

The given function is:

$$f(x) = x_1^2 + 5x_1x_2 - x_2^2, \quad x \in \mathbb{R}^2$$

We are asked whether the conjugate of the conjugate function, $f^{**}(x)$, is equal to the original function $f(x)$.

To determine this, we need to check if the function $f(x)$ is convex. A function is convex if its Hessian matrix is positive semi-definite, meaning that all the eigenvalues of the Hessian are non-negative.

Gradient of $f(x)$: The gradient of $f(x)$ is:

$$\nabla f(x) = \begin{bmatrix} 2x_1 + 5x_2 \\ 5x_1 - 2x_2 \end{bmatrix}$$

Hessian of $f(x)$: The Hessian matrix of $f(x)$ is:

$$H_f(x) = \begin{bmatrix} 2 & 5 \\ 5 & -2 \end{bmatrix}$$

Determinant of the Hessian: The determinant of the Hessian is:

$$\det(H_f(x)) = (2)(-2) - (5)(5) = -4 - 25 = -29$$

Since the determinant is negative, the Hessian is not positive semi-definite. Therefore, the function $f(x)$ is not convex. Since $f(x)$ is not convex, its conjugate's conjugate is not equal to $f(x)$. Therefore, the statement is **False**.

I.7 Fenchel's Inequality

Fenchel's inequality states that for any convex function $f(x)$, its conjugate function $f^*(y)$ satisfies: $f(x) + f^*(y) \geq \langle x, y \rangle, \quad \forall x, y$.

Consider the following statement: If $f(x)$ is strictly convex and differentiable,

then Fenchel's inequality attains equality if and only if $y = \nabla f(x)$.
T/F:
**Answer: True**
Fenchel's inequality states that for any convex function $f(x)$, its conjugate function $f^*(y)$ satisfies:

$$f(x) + f^*(y) \geq \langle x, y \rangle, \quad \forall x, y.$$

If $f(x)$ is strictly convex and differentiable, then Fenchel's inequality attains equality if and only if $y = \nabla f(x)$.
Explanation: - Fenchel's inequality is a fundamental result in convex analysis, providing a bound between a convex function $f(x)$ and its conjugate $f^*(y)$. - For strictly convex and differentiable functions, the equality condition in Fenchel's inequality holds when $y = \nabla f(x)$. This is a direct consequence of the properties of the subdifferential of convex functions. - Strict convexity ensures the uniqueness of the gradient, making the equality in Fenchel's inequality hold only when $y$ equals the gradient of $f(x)$, i.e., $y = \nabla f(x)$. Therefore, the statement is **True**.

I.8 Geometric Programming
The geometric programming formulation can incorporate the posynomial **equality** constraints, i.e. $\sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}} = 1$, where $K \geq 1$ and $c_k > 0$.
**Answer: False.**
Geometric programming (GP) can handle posynomial inequality constraints of the form:

$$\sum_{k=1}^{K} c_k x_1^{a_{1k}} x_2^{a_{2k}} \cdots x_n^{a_{nk}} \leq 1$$

However, it cannot directly handle posynomial equality constraints. The reason is that posynomials are sums of monomials with positive coefficients, and GP formulations require constraints to be either monomial equalities or posynomial inequalities.

I.9 Duality
For any primal optimization problem, taking the dual of its dual problem and finding its optimal value always gives back the value equal to the optimal value of the original primal problem.
T/F:
**Answer: False.**
When strong duality does not hold, the dual of the dual would not equal to the primal

I.10 Min Max Problem
In the context of duality theory, if a minimax problem has a saddle point, the duality gap must be zero. That is, given Lagrangian $\mathcal{L}(x, \lambda)$, consider the minimax objective:

$$\min_x \max_{\lambda \succeq 0} \mathcal{L}(x, \lambda)$$

Assume that there exists a saddle point $(x^*, \lambda^*)$ for the above objective, then the duality gap for the corresponding original optimization problem is zero. You can assume that the saddle point satisfies the KKT conditions).

**Answer: True.**

This is a fundamental result in minimax theory and duality. The existence of a saddle point means that the min-max optimization (primal) value equals the max-min optimization (dual) value, which is the definition of a zero duality gap. (An explanation along these lines is expected; a detailed explanation is given below).

If $(x^*, \lambda^*)$ is a **saddle point** of the minimax objective it satisfies, the following inequality for all feasible $x$ and $\lambda$:

$$\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*) \quad \forall x, \lambda \succeq 0$$

Therefore,

$$\max_{\lambda \succeq 0} \mathcal{L}(x^*, \lambda) = \mathcal{L}(x^*, \lambda^*) = \min_x \mathcal{L}(x, \lambda^*)$$

But (by definition)

$$\max_{\lambda \succeq 0} \mathcal{L}(x^*, \lambda) = \max_{\lambda \succeq 0} \min_x \mathcal{L}(x, \lambda)$$

$$\min_x \mathcal{L}(x, \lambda^*) = \min_x \max_{\lambda \succeq 0} \mathcal{L}(x, \lambda)$$

So,

$$\max_{\lambda \succeq 0} \min_x \mathcal{L}(x, \lambda) = d^* = \mathcal{L}(x^*, \lambda^*) = \min_x \max_{\lambda \succeq 0} \mathcal{L}(x, \lambda) = p^*$$

Therefore, the duality gap, $p^* - d^* = 0$

**Part II: Problem Solving: Show your process**

Problem 1. Conjugate Function. (20 pts)

Find the conjugate function of the following functions.

1.1 $f(x) = 2x^2 - 5x + 9$, where $x \in R$.

$$f^*(y) = \sup_x (yx - f(x)) = \sup_x \left(yx - 2x^2 + 5x - 9\right)$$

Consider $g(x, y) = yx - f(x) = yx - 2x^2 + 5x - 9 = -2x^2 + (y + 5)x - 9$.

Note that this is bounded over all $\mathbb{R}$, since it is a downward-facing parabola – i.e., for any value of $y$, this can never be $+\infty$. Also note that it will always have a maximum. Thus, our domain for the conjugate function is $\mathbb{R}$.

$$\nabla_x g(x, y) = -4x + (y + 5) = 0 \quad \Rightarrow \quad x = \frac{y + 5}{4}$$

Since $\nabla_x^2 g(x, y) = -4 < 0$, the function $f^*(y)$ achieves its global maximum at $x = \frac{y+5}{4}$ with value:

$$-2\left(\frac{y + 5}{4}\right)^2 + \frac{(y + 5)^2}{4} - 9 = \frac{(y + 5)^2}{8} - 9$$

Therefore, the conjugate function for $f(x)$ is:

$$f^*(y) = \frac{(y + 5)^2}{8} - 9 = \frac{y^2 + 10y - 47}{8}, \quad y \in \mathbb{R}$$

1.2 Consider the function

$$f(x) = \begin{cases} 2||x||_2^2, & \text{if } ||x||_2 \leq b, \\ b(4||x||_2 - b), & \text{if } ||x||_2 > b, \end{cases}$$

where variable $x \in \mathbb{R}^n$ and constant $b \in \mathbb{R}_{++}$.

**Answer:** The conjugate of a function $f(x)$ is given by:

$$f^*(y) = \sup_x \left(y^T x - f(x)\right).$$

We will consider the two cases for $f(x)$.
**Case 1:** When $||x||_2 \leq b$,

$$f(x) = 2||x||_2^2.$$

Thus, the conjugate becomes:

$$f^*(y) = \sup_{x:||x||_2 \leq b} \left(y^T x - 2||x||_2^2\right).$$

The expression $y^T x - 2||x||_2^2$ is maximized when $x$ is aligned with $y$. Let $x = \lambda y$, where $\lambda$ is a scalar. Then, we have:

$$y^T x = \lambda ||y||_2^2, \quad ||x||_2 = \lambda ||y||_2.$$

The constraint $||x||_2 \leq b$ implies that $\lambda \leq \frac{b}{||y||_2}$. Substituting this into the objective:

$$f^*(y) = \sup_{\lambda \leq \frac{b}{||y||_2}} \left(\lambda ||y||_2^2 - 2\lambda^2 ||y||_2^2\right).$$

This is a quadratic function of $\lambda$, and the supremum occurs at $\lambda = \frac{1}{4}$, resulting in:

$$f^*(y) = \frac{||y||_2^2}{8} \quad \text{for} \quad ||y||_2 \leq 4b.$$

**Case 2:** When $||x||_2 > b$,

$$f(x) = b(4||x||_2 - b).$$

The conjugate becomes:

$$f^*(y) = \sup_{x:||x||_2 > b} \left( y^T x - b(4||x||_2 - b) \right).$$

For large $||x||_2$, this term grows unbounded, so the supremum will be infinite when $||y||_2 > 4b$.
Thus, the conjugate function is:

$$f^*(y) = \begin{cases} \frac{\|y\|_2^2}{8}, & \text{if } \|y\|_2 \le 4b, \\ \infty, & \text{otherwise.} \end{cases}$$

Problem 2. Linear Programming. (20 pts)
Given

$$A = \begin{bmatrix} 9 & 3 & -1 & 2 & 0 & 1 & 0 \\ -3 & 1 & 1 & 1 & 1 & 1 & 1 \\ 5 & 0 & 1 & 1 & 0 & 0 & -1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & -1 & -1 & 0 \end{bmatrix},$$

$$b^T = \begin{bmatrix} 1 & 5 & 2 & 3 & 2 \end{bmatrix},$$

$$c^T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \end{bmatrix},$$

and $n = 7$, perform steps A, B, and C for problems 2.1, 2.2, 2.3, and 2.4.

A. Solve the following linear programming problems twice, once using the primal formulation and once using the dual formulation.
B. Check the feasibility of the solution. If a solution is not found, explain why a solution is not available and suggest how to mitigate the issue if you are the project leader. (For this exam, there is no need to solve the mitigated problem unless you feel the explanation is not convincing enough.)
C. Compare the primal and dual solutions. If the primal and dual formulation solutions are different, explain the difference.

2.1. minimize $f_0(x) = c^T x$ subject to $Ax \le b$, $x \in R^n$.

**A.** Solve in primal formulation.
The minimum value is $-\infty$, which can be achieved by one feasible point $x^T = [0\ 0\ 0\ 0\ 0\ 0\ t]$, where $t \le 0$. $\lim_{t\to-\infty} c^T x = \lim_{t\to-\infty} t = -\infty$
Solve in the dual formulation: Refers to the formula (5.22) of the textbook Chapter 5.2.1, pp. 225.

The dual problem is equivalent to $\max -b^T\lambda$ s.t. $A^T\lambda + c = 0$, $\lambda \geq 0$. When the constraints are infeasible, we regard $-\infty$ as the value of this dual problem.

However, this is not the definition of but is equivalent to the dual problem. The dual problem is defined as

$$\max g(\lambda) \text{ s.t. } \lambda \geq 0,$$

where $g(\lambda) = \begin{cases} -b^T\lambda & \text{if } A^T\lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$ is the Lagrange dual function.

The textbook Chapter 5.2.1, pp. 224-225 shows the derivation.

Now check if the constraints $A^T\lambda + c = 0$, $\lambda \geq 0$ are feasible. Now assume one of the two constraints holds, which is $A^T\lambda + c = 0$

We can use the Python library sympy's rref() to achieve Gaussian elimination for $[A^T | -c]$:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & 0 & 0 & | & 0 \\ 0 & 0 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & 0 & | & 1 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{bmatrix}$$

From the first five equations, we get $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = \lambda_5 = 0$. But the sixth equation is not satisfied. Therefore, the dual problem is infeasible and

$$g(\lambda) = -\infty$$

Thus, the dual problem is infeasible and its value is $-\infty$, which is the same as the primal problem.

**B.** The primal problem is feasible but its value is unbounded below. The dual problem is infeasible and its value is also unbounded below.

**C.** The primal and dual optimal values are the same, which is $-\infty$. They are the same since the primal problem is strictly feasible.

2.2. minimize $f_0(x) = c^T x$ subject to $Ax = b$, $x \in R^n$.

**Primal formulation:**

$$\min c^T x$$
$$\text{s.t. } Ax = b$$

**Dual formulation:**

$$\max g(\nu)$$
$$\text{s.t. } A^T\nu + c = 0$$

where,

$$g(\nu) = \begin{cases} -\nu^T b & \text{if } A^T\nu + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

**A.** To find the solution to the primal problem, we find the RREF of $[A \mid b]$:

$$[A \mid b] = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & -\frac{1}{13} & -\frac{2}{13} & -\frac{3}{26} \\ 0 & 1 & 0 & 0 & 0 & \frac{1}{39} & \frac{28}{39} & \frac{55}{78} \\ 0 & 0 & 1 & 0 & 0 & -\frac{11}{39} & \frac{4}{39} & \frac{68}{39} \\ 0 & 0 & 0 & 1 & 0 & \frac{2}{3} & -\frac{1}{3} & \frac{5}{6} \\ 0 & 0 & 0 & 0 & 1 & \frac{14}{39} & \frac{2}{39} & \frac{107}{78} \end{bmatrix}$$

Let $x_6 = t$ and $x_7 = 0$, for any $t \leq 0$, then:

$$x = \left[\frac{t}{13} - \frac{3}{26}, -\frac{t}{39} + \frac{55}{78}, \frac{11t}{39} + \frac{68}{39}, -\frac{2t}{3} + \frac{5}{6}, -\frac{14t}{39} + \frac{107}{78}, t, 0\right]$$

satisfies $Ax = b$. This form results in $c^T x = \frac{94t}{39} - \frac{392}{39}$, and $\lim_{t \to -\infty} c^T x = \lim_{t \to -\infty} t = -\infty$. Therefore, the primal problem is feasible but unbounded.

The dual problem is infeasible and $g(\nu) = -\infty$ (can be shown as in 2.1).

**B.** The primal problem is feasible but its value is unbounded below. The dual problem is infeasible and its value is also unbounded below.

**C.** The primal and dual optimal values are the same, which is $-\infty$. They are the same since the primal problem is strictly feasible.

2.3. minimize $f_0(x) = c^T x$ subject to $Ax \leq b$, $x \in R^n_+$.

**Primal formulation:**

$$\begin{aligned} &\min c^T x \\ &\text{s.t. } Ax \leq b \\ &x \geq 0 \end{aligned}$$

**Dual formulation:**

$$\begin{aligned} &\max g(\lambda) \\ &\text{s.t. } A^T\lambda + c \geq 0 \\ &\lambda \geq 0 \end{aligned}$$

where,

$$g(\lambda) = \begin{cases} -\lambda^T b & \text{if } A^T\lambda + c \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

**A.** The primal's optimal value is 0. Since $c \geq 0$, it must be true that $\forall x \in \mathbb{R}^n_+, c^T x \geq 0$. And since $c^T x = 0$ at the feasible point 0, the minimum possible value of $c^T x$ is 0.

The dual's optimal value is also 0. Since $b \geq 0$, it must be true that $\forall \lambda \in \mathbb{R}^5_+, \lambda^T b \geq 0$, i.e., $-\lambda^T b \leq 0$. And since $-\lambda^T b = 0$ at the feasible point 0, the maximum possible value of $-\lambda^T b$ is 0.

**B.** Both the primal and dual problems are feasible.

**C.** The primal and dual optimal values are the same, which is 0.

2.4. minimize $f_0(x) = c^T x$ subject to $Ax = b$, $x \in R^n_+$.

**Note:** You need to show valid solving for each of the parts.
**A.**
For the primal formulation, the primal solution remains $x^T = [0\ 0.6\ 2.1\ 0\ 0.9\ 1.3\ 0.1]$, giving:

$$c^T x = \begin{bmatrix}1 & 2 & 3 & 4 & 5 & 6 & 7\end{bmatrix} . \begin{bmatrix}0\ 0.6\ 2.1\ 0\ 0.9\ 1.3\ 0.1\end{bmatrix} = 20.5$$

For the dual formulation: The dual solution remains $\lambda^T = [-1.5\ 6.5\ -0.5\ -2.5\ -1]$, giving:

$$-b^T \lambda = -\begin{bmatrix}1 & 5 & 2 & 3 & 2\end{bmatrix} . \begin{bmatrix}-1.5\ 6.5\ -0.5\ -2.5\ -1\end{bmatrix} = 20.5$$

**B.** Both primal and dual solutions are feasible. The primal solution provides a minimum value of 20.5 at $x^T = [0\ 0.6\ 2.1\ 0\ 0.9\ 1.3\ 0.1]$, while the dual solution provides a maximum value of $-20.5$ at $\lambda^T = [-1.5\ 6.5\ -0.5\ -2.5\ -1]$. These solutions satisfy all constraints and complementary slackness conditions.

**C.** The primal and dual optimal values are equal while the primal minimizes $c^T x$. This demonstrates strong duality, as the optimal values match (considering the sign change in the dual formulation). The solutions are optimal as they satisfy both feasibility and complementary slackness conditions.

Problem 3. KKT Conditions. (30 pts)
Imagine that you work for a bank handling its trading portfolio. Your task is to minimize the risk associated with the portfolio while generating decent returns. A variant of this problem can be formulated as a convex optimization problem.

For the objective function, we have a covariance matrix $\Sigma \in S^n_{++}$ associated with the risk and a vector $x \in R^n$ associated with the investment portfolio. For inequality constraint, we have a minimum return threshold $b \in R$, and a vector $\alpha \in R^n$, which represents the average rate of return of the stocks. For the equality constraint. the total investment amount is fixed and normalized to one.

$$\begin{aligned} \min_x \ & \frac{1}{2} x^T \Sigma x \\ \text{s.t. } & \alpha^T x \geq b, \\ & \mathbf{1}^T x = 1. \end{aligned} \tag{1}$$

Taking the above primal problem, you need to describe the following.

a) Write the Lagrangian of the problem.
b) Write the KKT conditions for the optimal $x$ for the problem.
c) Write the dual problem.
d) Model this convex optimization problem in the convex solver of your choice. Describe the numerical value of the vector $x$ and the minimum value of the problem. You are given the following:

$$\Sigma = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}, \alpha = \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix}, b = 3$$

. **Ans:** The lagrangian for the above problem would be

$$L(x, \lambda, \mu) = \frac{1}{2} x^T \Sigma x + \lambda(b - x^T \alpha) + \mu(1 - \mathbf{1}^T x)$$

where $\lambda \geq 0, \mu \in R$.

The KKT conditions for the above case are:

**Primal Constraints:**

$$\alpha^T x \geq b$$

$$\mathbf{1}^T x = 1$$

**Dual Constraints**

$$\lambda \geq 0$$

**Complementary Slackness**

$$\lambda(b - x^T \alpha) = 0$$

**Gradient should be 0 wrt x**

$$\frac{\partial L}{\partial x} = \Sigma x - \lambda \alpha - \mu \mathbf{1} = \mathbf{0}$$

Now the dual function would be

$$g(\lambda, \mu) = \min_x L(x, \lambda, \mu)$$

To find the optimal x, we can use the gradient condition

$$x = \Sigma^{-1}(\lambda \alpha + \mu \mathbf{1})$$

Using this, the lagrange dual function would be

$$\begin{aligned} L(\lambda, \mu) = max_{\lambda,\mu} \frac{1}{2}(\lambda\alpha + \mu\mathbf{1})^T \Sigma^{-1} \Sigma \Sigma^{-1} (\lambda\alpha + \mu\mathbf{1}) \\ + \lambda(b - (\lambda\alpha + \mu\mathbf{1})^T \Sigma^{-1} \alpha) + \mu(1 - \mathbf{1}^T \Sigma^{-1}(\lambda\alpha + \mu\mathbf{1})) \end{aligned} \quad (2)$$

.

where $\lambda \geq 0$

We can simplify the following problem by assuming $z = \lambda\alpha + \mu\mathbf{1}$

$$L(\lambda, \mu) = \max_{\lambda,\mu} \frac{1}{2} z^T \Sigma^{-1} z + \lambda b + \mu - z^T \Sigma^{-1}(\lambda\alpha + \mu\mathbf{1})$$

$$L(\lambda, \mu) = \max_{\lambda,\mu} \frac{-1}{2} z^T \Sigma^{-1} z + \lambda b + \mu$$

$$L(\lambda, \mu) = \max_{\lambda,\mu} \frac{-1}{2} (\lambda\alpha + \mu\mathbf{1})^T \Sigma^{-1}(\lambda\alpha + \mu\mathbf{1}) + \lambda b + \mu$$

Hence the dual problem is:

$$\begin{aligned} &\max_{\lambda,\mu} \frac{-1}{2} (\lambda\alpha + \mu\mathbf{1})^T \Sigma^{-1}(\lambda\alpha + \mu\mathbf{1}) + \lambda b + \mu \\ &\text{s.t. } \lambda \geq 0, \end{aligned} \tag{3}$$

.

Their are multiple variants of simplifying the dual, we will accept them if they are equivalent. (Also the cases where people took $\mathbf{1}^T x - 1$ in the lagrangian is correct and would get the correct score. For that case the dual comes out to have different sign of $\mu$ in the dual equation).

We need to have the final equation of the dual and not just $g(\lambda, \mu)$ since its possible to evaluate it in this case

Now for part d, we can use cvxpy to solve this problem

The minimum value for the problem is: 0.75

The optimal value for x is : $[-1, 2.5, -0.5]$

Below is the code to generate it:

```
import numpy as np
import cvxpy as cp
✓ 0.0s

SIGMA = np.array([[10,5,2],[5,3,2],[2,2,3]])
ALPHA = np.array([2,4,7])
ONES = np.ones(3)
BB = 3

✓ 0.0s

x = cp.Variable((len(ALPHA),1))
✓ 0.0s

prob  =cp.Problem(cp.Minimize(0.5*cp.quad_form(x, SIGMA)), [ALPHA.T @ x >= BB, ONES.T @ x == 1])
✓ 0.0s

prob.solve(cp.CLARABEL)
✓ 9.1s
0.7499999999999982

x.value
✓ 0.0s
array([[-1. ],
       [ 2.5],
       [-0.5]])
```