Part 1

I Exercise 1

Consider the following optimisation problem:

$$\min f(\boldsymbol{x}) \tag{1}$$
$$g_1(\boldsymbol{x}) \leq 0 \tag{2}$$
$$g_2(\boldsymbol{x}) \leq 0 \tag{3}$$

where $f(\boldsymbol{x}) = (x_1 - 4)^2 + (x_2 - 6)^2$ and

$$g(x_1, x_2) = \begin{cases} x_1^2 - x_2, \\ x_2 - 4. \end{cases} \tag{4}$$

To solve this problem using the Karush-Kuhn-Tucker (KKT) conditions, we first form the Lagrangian:

$$\mathcal{L}(x, \lambda) = f(\boldsymbol{x}) + \lambda_1 g_1(\boldsymbol{x}) + \lambda_2 g_2(\boldsymbol{x}),$$

where λ_1 and λ_2 are Lagrange multipliers.
The KKT conditions for optimality are given by:

$$\begin{aligned} &\nabla f(\boldsymbol{x}^*) + \lambda_1 \nabla g_1(\boldsymbol{x}^*) + \lambda_2 \nabla g_2(\boldsymbol{x}^*) = 0, \qquad (5)\\ &g_1(\boldsymbol{x}^*) \leq 0, \quad g_2(\boldsymbol{x}^*) \leq 0, \\ &\lambda_1 \geq 0, \quad \lambda_2 \geq 0, \\ &\lambda_1(g_1(\boldsymbol{x}^*) - 0) = 0, \quad \lambda_2(g_2(\boldsymbol{x}^*) - 0) = 0, \\ &g_1(\boldsymbol{x}^*) \leq 0, \quad g_2(\boldsymbol{x}^*) \leq 0. \end{aligned}$$

Substituting the expressions for $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ from (1) and (4) into (5), we get the following equations:

$$\begin{cases} 2(x_1^* - 4) + 2\lambda_1 x_1^* = 0, \\ 2(x_2^* - 6) - \lambda_1 + \lambda_2 = 0. \end{cases}$$

Now, let's consider a specific solution $\boldsymbol{x}^* = (2, 4)$. Plugging in these values into the KKT conditions, we find that $\lambda_1 = 1$ and $\lambda_2 = 5$, and all other conditions are satisfied.
In summary, the solution to the optimization problem (1) with constraints (2) and (3) that satisfies the KKT conditions is $\boldsymbol{x}^* = (2, 4)$ with $\lambda_1 = 1$ and $\lambda_2 = 5$. Additionally $f(\boldsymbol{x})$ is a strictly convex function, in fact its hessian:

$$\mathbb{H}_f(\boldsymbol{x}) = 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{6}$$

so the hessian is (strictly) positive definite.
Additionally also the constraint is convex, in fact $\{\boldsymbol{x} : x_2 - 4 \leq 0\}$ is obviously convex since it is an half-plain. Similarly the set $\{\boldsymbol{x} : x_1^2 - x_2 \leq 0\}$ is convex since it is the epigraph of a convex function ($x_2 = x_1^2$). So the constraint $g(\boldsymbol{x})$ is convex as intersection of convex sets.
Both the objective function and the constraints are convex. In such cases, the Karush-Kuhn-Tucker (KKT) conditions are not only necessary but also sufficient conditions for optimality.

II Exercise 2

Let $S \subseteq \mathbb{R}^n$ be a closed and convex set, where $n \in \mathbb{N}$. Consider $\boldsymbol{x}$ and $\boldsymbol{y}$ in S, and let $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{y}}$ be their respective projections onto S. To avoid trivial cases, assume $x \neq y$, and also assume $\bar{\boldsymbol{x}} \neq \bar{\boldsymbol{y}}$. According to the Projection Theorem, for all $z \in S$, the following inequalities hold:

$$(\boldsymbol{x} - \bar{\boldsymbol{x}})^T(\boldsymbol{z} - \bar{\boldsymbol{x}}) \leq 0, \tag{7}$$
$$(\boldsymbol{y} - \bar{\boldsymbol{y}})^T(\boldsymbol{z} - \bar{\boldsymbol{y}}) \leq 0. \tag{8}$$

Now, substitute $z = \bar{\boldsymbol{y}}$ into (7) and $z = \bar{\boldsymbol{x}}$ into (8), since $\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}} \in S$. By summing (7) and (8), we obtain:

$$\begin{aligned} 0 \geq\ & (\boldsymbol{x} - \bar{\boldsymbol{x}})^T(\bar{\boldsymbol{y}} - \bar{\boldsymbol{x}}) + (\boldsymbol{y} - \bar{\boldsymbol{y}})^T(\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}) = \\ & (\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}})^T(\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}} - \boldsymbol{x} + \boldsymbol{y}) \\ & \|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|^2 - (\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}})^T(\boldsymbol{y} - \boldsymbol{x}). \end{aligned} \tag{9}$$

Using (9), we can derive the desired proof:

$$\begin{aligned} \|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|^2 &\leq (\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}})^T(\boldsymbol{y} - \boldsymbol{x}) \quad \text{[by Equation (9)]} \\ &= \|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|\|\boldsymbol{x} - \boldsymbol{y}\| \cos\left(\frac{(\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}})^T(\boldsymbol{y} - \boldsymbol{x})}{\|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|\|\boldsymbol{x} - \boldsymbol{y}\|}\right) \\ &\leq \|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|\|\boldsymbol{x} - \boldsymbol{y}\|, \end{aligned} \tag{10}$$

and by dividing both sides by $\|\bar{\boldsymbol{x}} - \bar{\boldsymbol{y}}\|$, we obtain the desired inequality.

III Exercise 3

The following points are going to prove that all given sets are convex.

1 Point a

Let $S = \{\boldsymbol{x} \in \mathbb{R}^n : \alpha \leq \boldsymbol{a}^T\boldsymbol{x} \leq \beta\}$, where $\alpha, \beta \in \mathbb{R}$ and $\boldsymbol{a} \in \mathbb{R}^n$. If $\alpha > \beta$, then $S = \varnothing$, which is a convex set. Now, let's assume $\alpha \leq \beta$. If $\boldsymbol{a} = \boldsymbol{0}$, we have $S = \mathbb{R}$ if $\alpha < 0$ and $S = \varnothing$ otherwise. Thus, we can assume that $\boldsymbol{a} \neq \boldsymbol{0}$.
Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be in S, and let λ and $1-\lambda$ be positive

real numbers. Also, let $\boldsymbol{z} = \lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$. We need to prove that $z \in S$:

$$\begin{aligned}\alpha &= \alpha\lambda + \alpha(1-\lambda)\\ &= \underbrace{\leq}_{x,\boldsymbol{y}\in S} \underbrace{\lambda \boldsymbol{a}^T\boldsymbol{x} + (1-\lambda)\boldsymbol{a}^T\boldsymbol{y}}_{\boldsymbol{a}^T z}\\ &= \underbrace{\leq}_{x,\boldsymbol{y}\in S} \lambda\beta + (1-\lambda)\beta = \beta,\end{aligned}$$

and so $\boldsymbol{z} \in S$ by definition.

2 Point b

Let $S = \{x \in \mathbb{R}^n \mid \alpha_i \leq x_i \leq \beta_i, \forall i = 1, \dots, n\}$. We assume that $a_i \leq b_i$ to avoid trivial cases. However, by definition, S is a closed hypercube in $\mathbb{R}^n$ and is thus convex. The proof is straightforward. Suppose $z = \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$, where $x, y \in S$, with both λ and $(1-\lambda)$ being non-negative. For all $i = 1, \dots, n$, we have:

$$\alpha_i \leq \alpha_i\lambda + \alpha_i(1-\lambda) \leq \lambda x_i + (1-\lambda)y_i \leq \lambda\beta_i + (1-\lambda)\beta_i = \beta_i$$

and so $z \in S$.

3 Point c

In this case, $S = \{\boldsymbol{x} \in \mathbb{R} : a_1^T x \leq b_1, a_2^T x \leq b_2\}$ is the intersection of two half-planes, and as the intersection of two convex sets, it is still convex.

Let $\boldsymbol{z}$, $\boldsymbol{x}$, $\boldsymbol{y}$ and λ be as usual. We have:

$$\begin{aligned}a_1^T z &= \lambda a_1^T \boldsymbol{x} + (1-\lambda)a_1^T\boldsymbol{y} \leq \lambda b_1 + (1-\lambda)b_1 \leq b_1,\\ a_2^T z &= \lambda a_2^T \boldsymbol{x} + (1-\lambda)a_2^T\boldsymbol{y} \leq \lambda b_2 + (1-\lambda)b_2 \leq b_2.\end{aligned}$$

4 Point d

Now, consider the set S defined as follows: $S = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x}A^T\boldsymbol{y}, \boldsymbol{y} \in \mathbb{R}^m, \boldsymbol{y} \geq 0\}$, where $A \in \mathbb{R}^{m\times n}$. S is classified as a polyhedron. In mathematical terms, a polyhedron is characterised as a convex region formed by the intersection of a finite number of half-spaces. The definition of S provided here perfectly aligns with this mathematical concept.

Moreover, it's worth noting that if we continue to use the same notation as previously introduced, and for any u and v within S, it holds true that $\boldsymbol{w} = \lambda\boldsymbol{u} + (1-\lambda)\boldsymbol{v} \in S$ for any non-negative values of λ and $(1-\lambda)$.

Additionally, if $\boldsymbol{z}$ and $\boldsymbol{y}$ are solutions to the linear equation, any subspace generated by $\boldsymbol{z}$ and $\boldsymbol{y}$ will also be a solution to the equation. Therefore, it follows that any convex combination of these sub-spaces, including $\boldsymbol{x} = \lambda\boldsymbol{z} + (1-\lambda)\boldsymbol{x}$, remains a solution to the equation.

IV Exercise 4

1 Point a

Let be $\lambda \in [0, 1]$, so it is easy to show that:

$$\begin{aligned}f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &:= g(A(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + \boldsymbol{b})\\ &= g(A\lambda\boldsymbol{x} + A(1-\lambda)\boldsymbol{y} + \boldsymbol{b})\\ &= g(\lambda(A\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)(A\boldsymbol{y} + \boldsymbol{b}))\\ &\underbrace{\leq}_{g \text{ is convex}} \lambda g(A\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)g(A\boldsymbol{y} + \boldsymbol{b})\\ &= \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y}).\end{aligned}$$

2 Point b

If $h(\boldsymbol{x})$ is a convex function, then it is evident that $\alpha h(\boldsymbol{x})$ remains convex when $\alpha \geq 0$ and becomes concave otherwise. When we introduce a constant term, such as $f(\boldsymbol{x}) = \alpha h(\boldsymbol{x}) + \beta$, we are essentially translating the function along the y-axis. Consequently, the epigraph of the function remains convex. It's important to note that a function is convex if and only if its epigraph is a convex set.

3 Point c

Let A be a positive semi-definite symmetric $n \times n$ matrix. Since A defines a non-degenerate scalar product, the function $g(\boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x}$ is convex. We want to prove this by showing that for any $\boldsymbol{x}$ and $\boldsymbol{y}$ in the domain of g and for any λ in the interval $[0, 1]$, the following inequality holds:

$$g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leq \lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y}) \tag{11}$$

Starting from the expression for $g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y})$:

$$\begin{aligned}&g(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) =\\ &= \lambda^2\boldsymbol{x}^T A\boldsymbol{x} + 2\lambda(1-\lambda)\boldsymbol{x}^T A\boldsymbol{y} + (1-\lambda)^2\boldsymbol{y}^T A^T\boldsymbol{y}\\ &= \lambda\boldsymbol{x}^T A(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + (1-\lambda)\boldsymbol{y}^T A(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y})\end{aligned}$$

Now, notice that the expression inside the parentheses is a convex combination of $\boldsymbol{x}$ and $\boldsymbol{y}$, which means it is a point within the line segment between $\boldsymbol{x}$ and $\boldsymbol{y}$. Since convex functions have the property that the value at any point is less than or equal to the value at any point within the convex hull, we can substitute $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ for a point within that line segment.

So, in the expression:

$$\lambda\boldsymbol{x}^T A(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + (1-\lambda)\boldsymbol{y}^T A(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y})$$

we have effectively replaced $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ with a point within the line segment between $\boldsymbol{x}$ and $\boldsymbol{y}$. Since $\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ lies within the convex hull of $\boldsymbol{x}$ and $\boldsymbol{y}$, it follows that

$$\lambda \boldsymbol{x}^T A(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + (1-\lambda)\boldsymbol{y}^T A(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\ \leq \lambda \boldsymbol{x}^T A \boldsymbol{x} + (1-\lambda)\boldsymbol{y}^T A \boldsymbol{y}$$

This completes the proof, demonstrating that the function $g(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$ is indeed convex. Besides also the function $\beta \boldsymbol{x}^T A \boldsymbol{x}$ is convex if $\beta \geq 0$. At this point is sufficient to note that the function e^x is convex too(is a well-known exponential function). So $f(\boldsymbol{x}) = \mathrm{e}^{\beta \boldsymbol{x}^T A \boldsymbol{x}}$ is a convex function as a composition of convex functions.