Notes:

1. Duration: 2.5 hours
2. Explain everything carefully. You will be graded on the clarity of your arguments.

Exercises:

1. *(2 points)* Let a convex set $C \subset \mathbb{R}^n$ and a point $x_0 \notin C$. Show that

$$\operatorname{conv}(C \cup \{x_0\}) = \{(1-\theta)x + \theta x_0 : x \in C, \theta \in [0,1]\}.$$

Solution: Observe that

$$\{(1-\theta)x + \theta x_0 : x \in C, \theta \in [0,1]\} \subseteq \operatorname{conv}(C \cup \{x_0\})$$

by the definition of the convex hull. So the interesting part of the proof is showing the inverse. To this effect, let $y \in \operatorname{conv}(C \cup \{x_0\})$. It then follows

$$y = \theta_0 x_0 + \sum_{i=1}^{n} \theta_i x_i = \theta_0 x_0 + \left(\sum_{i=1}^{n} \theta_i\right) \sum_{i=1}^{n} \left(\frac{\theta_i}{\sum_{i=1}^{n} \theta_i} x_i\right),$$

where in the above the points $x_i \in C$, for $i = 1, \ldots, n$, and $\sum_{i=0}^{n} \theta_i = 1$. The result easily follows by setting $\theta = \theta_0$, noting that then $\sum_{i=1}^{n} \theta_i = 1 - \theta$, and observing that the last summation belongs to C, as it is convex.

2. *(2 points)*

(α′) Let $f : [a, b] \to \mathbb{R}$ convex function defined on a closed interval. Show that f is bounded above by $\max\{f(a), f(b)\}$.

(β′) Generalize this property when $f : A \to \mathbb{R}$ and $A \subseteq \mathbb{R}^n$ and give its proof.

Solution:

(α′) By Jensen's inequality, for any $x \in [a, b]$ we have, for some $\theta \in [0, 1]$,

$$f(x) \leq \theta f(a) + (1-\theta) f(b) \leq \theta \max\{f(a), f(b)\} + (1-\theta)\max\{f(a), f(b)\} = \max\{f(a), f(b)\}$$

(β′) The generalization that if f is defined in closed bounded set, then there is no interior point where the function attains a global maximum greater than the supremum of all values in the boundary. Proof is easy, by assuming the property does not hold and considering a properly defined line segment connecting an interior point that

3. *(2 points)* Let the function

$$f(x, y, z) = xyz,$$

defined for $x, y, z \geq 0$. Is this function convex?

Solution: The function is not convex. Indeed, consider, e.g., the points $(1, 1, 0)$ and $(0, 1, 1)$ and any point on the line segment connecting them. Jensen's inequality will not hold for these points.

Alternative, we can study the Hessian and find negative eigenvalues for combinations of x, y, z.

$$\nabla f = \begin{bmatrix} yz \\ xz \\ xy \end{bmatrix}, \qquad \nabla^2 f = \begin{bmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{bmatrix}.$$

It is straightforward to see that the determinant of the Hessian is

$$\begin{vmatrix} -s & z & y \\ z & -s & x \\ y & x & -s \end{vmatrix} = (-s)(s^2 - x^2) - z(-zs - xy) + y(xz + sy) = -s^3 + sx^2 + sz^2 + xyz + xyz + sy^2.$$

4. *(2 points)* Consider the problem

$$\begin{array}{ll} \text{minimize:} & xyz \\ \text{subject to:} & x + y^2 + z^4 \leq 1, \qquad x + y + z = 1, \end{array}$$

defined for $x, y, z \in \mathbb{R}^3$.

(α′) Bring the problem in the standard form of an optimization problem.

(β′) Is the problem convex? Explain?

(γ′) Write the Lagrangian.

(δ′) Write the KKT conditions for this problem, but do NOT solve them.

Solution: The KKT conditions are:

$$\begin{aligned} x + y^2 + z^4 &\le 1, \\ x + y + z &= \; = 1, \\ \lambda &\ge 0, \\ \lambda(x + y^2 + z^4 - 1) &= 0, \\ \begin{bmatrix} yz \\ xz \\ xy \end{bmatrix} + \lambda \begin{bmatrix} 1 \\ 2y \\ 4z^3 \end{bmatrix} + \mu \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} &= 0. \end{aligned}$$

5. *(2 points)* Consider the problem

$$\begin{array}{ll} \text{minimize:} & \frac{1}{2}\|x - b\|^2 + t, \\ \text{subject to:} & t \ge a_i^T x + c_i, \quad i = 1, \ldots, m, \end{array}$$

where we optimize with respect to both $x \in \mathbb{R}^n$ and $t \in \mathbb{R}$, whereas $b, a_i \in \mathbb{R}^n$ and $c_i \in \mathbb{R}$ are parameters. Define the dual problem (but do not solve it).

Solution: The Lagrangian is

$$L(x, t, \lambda) \;=\; \frac{1}{2}\|x - b\|^2 + t + \sum_{i=1}^{m} \lambda_i (a_i^T x + c_i - t) \;=\; \frac{1}{2}\|x - b\|^2 + \sum_{i=1}^{m} \lambda_i a_i^T x + \left(1 - \sum_{i=1}^{m} \lambda_i\right) t + \sum_{i=1}^{m} \lambda_i c_i.$$

Observe that if $\sum_{i=1}^m \lambda_i \neq 1$, then the Lagrangian can be made arbitrarily small, by taking the value of $t \to \pm\infty$. Assuming $\sum_{i=1}^m \lambda_i = 1$, to minimize the Lagrangian we need to find x we need to minimize the quantity

$$\frac{1}{2}(x - b)^T I (x - b) + \lambda^T A^T x,$$

where we defined the matrix $A = [a_1 \; a_2 \; \ldots \; a_m]$ of size $n \times m$. The quadratic form is minimized when

$$x - b = -A\lambda \Leftrightarrow x = b - A\lambda.$$

Substituting in the Lagrangian, we find that, in this case,

$$g(\lambda) = \frac{1}{2}\lambda^T A^T A \lambda - \lambda^T A^T A \lambda + \lambda^T A^T b + \lambda c = -\frac{1}{2}\lambda^T A^T A \lambda + \lambda^T (A^T b + c),$$

and concluding, the dual function is

$$g(\lambda) = \begin{cases} -\infty, & \sum_{i=1}^m \lambda_i \neq 1, \\ -\frac{1}{2}\lambda^T A^T A \lambda + \lambda^T (A^T b + c), & \sum_{i=1}^m \lambda_i = 1, \end{cases}$$

and the dual problem becomes

$$\begin{array}{ll} \text{maximize:} & -\frac{1}{2}\lambda^T A^T A \lambda + \lambda^T (A^T b + c), \\ \text{subject to:} & \lambda_i \ge 0, \quad \sum_{i=1}^m \lambda_i = 1. \end{array}$$