OPTIMIZATION HOMEWORK

SPRING 2025

1. **(Regression) To fit a linear regression model to a data set, we need to solve**

$$\min_{\mathbf{x}} \frac{1}{2n}\|\mathbf{y} - A\mathbf{x}\|_2^2$$

where $\mathbf{y} \in \mathbb{R}^n$ is the vector of response variables, $A \in \mathbb{R}^{n\times p}$ is the design matrix, and $\mathbf{x} \in \mathbb{R}^p$ is the vector of regression coefficients. Simulate elements of $\mathbf{y}$ and A independently from the standard normal distribution for a problem with $n = 50$, and $p = 6$ (including the intercept).

Solution. **(a).** We wish to minimize the objective

$$f(\mathbf{x}) = \frac{1}{2n}\|\mathbf{y} - A\mathbf{x}\|_2^2 = \frac{1}{2n}(\mathbf{y} - A\mathbf{x})^\top(\mathbf{y} - A\mathbf{x}).$$

Taking the gradient with respect to $\mathbf{x}$, we have

$$\nabla f(\mathbf{x}) = \frac{1}{2n}\cdot 2A^\top(A\mathbf{x} - \mathbf{y}) = \frac{1}{n}A^\top(A\mathbf{x} - \mathbf{y}).$$

Setting the gradient equal to zero gives the first-order optimality condition:

$$A^\top(A\mathbf{x} - \mathbf{y}) = \mathbf{0}.$$

Rearranging, we obtain

$$A^\top A\,\mathbf{x} = A^\top\mathbf{y}.$$

Assuming that $A^\top A$ is invertible, the unique minimizer is

$$\mathbf{x}^* = \left(A^\top A\right)^{-1}A^\top\mathbf{y}.$$

For (b)-(e) see the jupyter notebook file.

2. **Show that the solution set of a Linear Matrix Inequality (LMI) is a convex set, i.e. show that the set $\{\mathbf{x} : A(\mathbf{x}) \preceq B\}$ is a convex set, where $A(\mathbf{x}) := x_1A_1 + \cdots + x_nA_n$ and $B, A_i \in \mathbb{S}^m$ for $i = 1, \ldots, n$. Note that $A \preceq B$ means $B - A \succeq 0$.**

Solution. Let $\mathbf{x}, \mathbf{y}$ satisfy

$$A(\mathbf{x}) \preceq B \quad \text{and} \quad A(\mathbf{y}) \preceq B,$$

so that

$$B - A(\mathbf{x}) \succeq 0 \quad \text{and} \quad B - A(\mathbf{y}) \succeq 0.$$

For any $\theta \in [0, 1]$, define $\mathbf{z} = \theta\mathbf{x} + (1 - \theta)\mathbf{y}$. Since $A(\mathbf{x})$ is affine, we have

$$A(\mathbf{z}) = \theta A(\mathbf{x}) + (1 - \theta)A(\mathbf{y}).$$

Then,

$$B - A(\mathbf{z}) = \theta(B - A(\mathbf{x})) + (1 - \theta)(B - A(\mathbf{y})).$$

Since the set of positive semidefinite matrices is convex, it follows that $B - A(\mathbf{z}) \succeq 0$, as a convex combination of positive semidefinite matrices, which implies $A(\mathbf{z}) \preceq B$. Thus, the set $\{\mathbf{x} : A(\mathbf{x}) \preceq B\}$ is convex.

3. **Show that the Cartesian product is a convexity preserving operation, i.e., given $\mathcal{X}_1 \subseteq \mathbb{R}^n$, $\mathcal{X}_2 \subseteq \mathbb{R}^m$ convex, then $\mathcal{X}_1 \times \mathcal{X}_2 := \{(\mathbf{x}_1, \mathbf{x}_2) : \mathbf{x}_1 \in \mathcal{X}_1, \mathbf{x}_2 \in \mathcal{X}_2\}$ is a convex set.**

   Solution. Let $(\mathbf{x}_1, \mathbf{x}_2)$ and $(\mathbf{y}_1, \mathbf{y}_2)$ be arbitrary points in $\mathcal{X}_1 \times \mathcal{X}_2$, where $\mathbf{x}_1, \mathbf{y}_1 \in \mathcal{X}_1$ and $\mathbf{x}_2, \mathbf{y}_2 \in \mathcal{X}_2$. For any $\theta \in [0,1]$, consider the convex combination:

$$\theta(\mathbf{x}_1, \mathbf{x}_2) + (1-\theta)(\mathbf{y}_1, \mathbf{y}_2) = (\theta\mathbf{x}_1 + (1-\theta)\mathbf{y}_1,\, \theta\mathbf{x}_2 + (1-\theta)\mathbf{y}_2).$$

   Since $\mathcal{X}_1$ and $\mathcal{X}_2$ are convex, we have:

$$\theta\mathbf{x}_1 + (1-\theta)\mathbf{y}_1 \in \mathcal{X}_1 \quad \text{and} \quad \theta\mathbf{x}_2 + (1-\theta)\mathbf{y}_2 \in \mathcal{X}_2.$$

   Thus,

$$(\theta\mathbf{x}_1 + (1-\theta)\mathbf{y}_1,\, \theta\mathbf{x}_2 + (1-\theta)\mathbf{y}_2) \in \mathcal{X}_1 \times \mathcal{X}_2.$$

   This proves that $\mathcal{X}_1 \times \mathcal{X}_2$ is convex.

4. **Recall the perspective function $p : \mathbb{R}^{n+1} \to \mathbb{R}^n$ with $\mathbf{dom}(p) = \mathbb{R}^n \times \mathbb{R}_{++}$ defined as $p_i(\mathbf{x}, t) := x_i/t$. Let $C \subseteq \mathbf{dom}(p)$ be a convex set. Show that $p(C)$ is a convex set. Furthermore, let $S \subseteq \mathbb{R}^n$ be a convex set. Show that the preimage of S under p, i.e., $p^{-1}(S)$ is a convex set.**

   Solution. **(i) Convexity of $p(C)$:** Let $C \subseteq \mathbb{R}^n \times \mathbb{R}_{++}$ be convex and take any $\mathbf{u}, \mathbf{v} \in p(C)$. Then, there exist $(\mathbf{x}, t), (\mathbf{y}, s) \in C$ such that

$$\mathbf{u} = \frac{\mathbf{x}}{t} \quad \text{and} \quad \mathbf{v} = \frac{\mathbf{y}}{s}.$$

   For any $\lambda \in [0,1]$, consider the point

$$(\lambda\mathbf{x} + (1-\lambda)\mathbf{y},\, \lambda t + (1-\lambda)s).$$

   Since C is convex, this point lies in C. Its image under p is

$$p\big(\lambda\mathbf{x} + (1-\lambda)\mathbf{y},\, \lambda t + (1-\lambda)s\big) = \frac{\lambda\mathbf{x} + (1-\lambda)\mathbf{y}}{\lambda t + (1-\lambda)s}.$$

   Observe that

$$\frac{\lambda\mathbf{x} + (1-\lambda)\mathbf{y}}{\lambda t + (1-\lambda)s} = \frac{\lambda t}{\lambda t + (1-\lambda)s}\frac{\mathbf{x}}{t} + \frac{(1-\lambda)s}{\lambda t + (1-\lambda)s}\frac{\mathbf{y}}{s}.$$

   Since the coefficients

$$\alpha = \frac{\lambda t}{\lambda t + (1-\lambda)s} \quad \text{and} \quad 1-\alpha = \frac{(1-\lambda)s}{\lambda t + (1-\lambda)s}$$

   are nonnegative and sum to 1, it follows that

$$\frac{\lambda\mathbf{x} + (1-\lambda)\mathbf{y}}{\lambda t + (1-\lambda)s} = \alpha\,\mathbf{u} + (1-\alpha)\,\mathbf{v}.$$

   Thus, any convex combination of points in $p(C)$ lies in $p(C)$, proving that $p(C)$ is convex.

   (ii) Convexity of $p^{-1}(S)$: Let $S \subseteq \mathbb{R}^n$ be convex and define

$$p^{-1}(S) = \{(\mathbf{x}, t) \in \mathbb{R}^n \times \mathbb{R}_{++} : \mathbf{x}/t \in S\}.$$

   Take any two points $(\mathbf{x}, t), (\mathbf{y}, s) \in p^{-1}(S)$. Then,

$$\frac{\mathbf{x}}{t} \in S \quad \text{and} \quad \frac{\mathbf{y}}{s} \in S.$$

   For any $\lambda \in [0,1]$, consider

$$(\mathbf{z}, r) = (\lambda\mathbf{x} + (1-\lambda)\mathbf{y},\, \lambda t + (1-\lambda)s).$$

Since $t, s > 0$, we have $r > 0$. By the convexity of S,

$$\frac{\mathbf{z}}{r} = \frac{\lambda\mathbf{x} + (1-\lambda)\mathbf{y}}{\lambda t + (1-\lambda)s}$$

is a convex combination of $\mathbf{x}/t$ and $\mathbf{y}/s$, and hence lies in S. Thus,

$$(\mathbf{z}, r) \in p^{-1}(S),$$

which shows that $p^{-1}(S)$ is convex.

5. **Show that the hyperbolic set $\{\mathbf{x} \in \mathbb{R}^2_+ : x_1 x_2 \geq 1\}$ is convex.**

Solution. Consider the set

$$H = \{(x_1, x_2) \in \mathbb{R}^2_{++} : x_1 x_2 \geq 1\}.$$

Taking logarithms (which is valid since $x_1, x_2 > 0$), we have

$$\log(x_1 x_2) = \log x_1 + \log x_2 \geq 0.$$

Since the function $(x_1, x_2) \mapsto \log x_1 + \log x_2$ is concave, its superlevel sets

$$\{(x_1, x_2) \in \mathbb{R}^2_{++} : \log x_1 + \log x_2 \geq 0\}$$

are convex. Hence, H is convex.