Homework 1

Kailong Wang

October 4, 2023

Q1: Affine images and preimages of convex sets. Grade:

Let $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, C \subset \mathbb{R}^n, D \subset \mathbb{R}^m$ be convex sets. Show that following sets are convex.

(a) The image of C under the affine map $x \mapsto Ax + b$. That is

$$\{Ax + b \mid x \in C\} \subset \mathbb{R}^m .$$

(b) The preimage of D under the affine map $x \mapsto Ax + b$. That is

$$\{x \mid Ax + b \in D\} \subset \mathbb{R}^n .$$

Solution

(a) *Proof.* Let $x_1, x_2 \in C$ and $\lambda \in [0,1]$, then we have

$$\begin{aligned}\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b) &= A(\lambda x_1 + (1-\lambda)x_2) + b \\ &\in A(C) + b.\end{aligned}$$

Thus, the image of C, $A(C) + b$ is convex. □

(b) *Proof.* Let $y_1, y_2 \in A^{-1}(D-b)$ so that $Ay_1 + b \in D, Ay_2 + b \in D$ and $\lambda \in [0,1]$, then we have

$$\begin{aligned}A(\lambda y_1 + (1-\lambda)y_2) + b &= \lambda(Ay_1 + b) + (1-\lambda)(Ay_2 + b) \\ &\in \lambda D + (1-\lambda)D \\ &= D.\end{aligned}$$

Thus, The preimages of D, $A^{-1}(D-b)$ is convex. □

Q2: Affine functions. Grade:

Suppose that $f : \mathbb{R}^n \to \mathbb{R} \setminus \{-\infty, \infty\}$ always obey the convex function relation at equality, that is,

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y), \quad \forall x, y \in \mathbb{R}^n, \lambda \in [0,1]. \tag{1}$$

Show that

(a) If eq. (1) holds as stated for all $\lambda \in [0,1]$, it in fact holds for all $\lambda \in \mathbb{R}$.

(b) Any f for which eq. (1) holds must be of the form $f(x) = \langle a, x\rangle + b$ for $\lambda \in \mathbb{R}^n, b \in \mathbb{R}$ (that is, f is an inner product with some fixed vector plus a constant).

(c) Any function of this form has the property eq. (1).

Hint: given f satisfying the condition above, show that $g : x \mapsto f(x) \to f(0)$ is linear. You may then use (without proof, although the proof is very easy) that a linear function $g : \mathbb{R}^n \to \mathbb{R}$ must be of the form $x \mapsto \langle a, x\rangle$ for some $a \in \mathbb{R}^n$.

Solution

(a) *Proof.* To extend eq. (1) to $\lambda \in \mathbb{R}^n$, we need to show that eq. (1) holds for $\lambda \in (-\infty, 0) \cup (1, \infty)$.

First, let $x, y \in \mathbb{R}^n$, and for $\lambda \in (-1, 0)$ let $\alpha = -\lambda \in (0, 1)$. Then we have

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y),$$

which shows convexity for $\lambda \in (-1, 0)$. Similarly, for $\lambda \in (1, \infty)$ let $\alpha = \frac{1}{\lambda} \in (0, 1)$, and for $\lambda \in (-\infty, -1)$ let $\alpha = -\frac{1}{\lambda} \in (0, 1)$, we can prove item (a) holds for $\lambda \in (1, \infty)$ and $\lambda \in (-\infty, -1)$ respectively. □

(b) *Proof.* Let's define $g : \mathbb{R}^n \to \mathbb{R}$ as $g(x) = f(x) - f(0)$, then we have $g(0) = f(0) - f(0) = 0$. For any $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ (as proved above), we have

$$\begin{aligned} g(\lambda x + (1-\lambda)y) &= f(\lambda x + (1-\lambda)y) - f(0) \\ &\leq \lambda f(x) + (1-\lambda)f(y) - f(0) \\ &= \lambda(f(x) - f(0)) + (1-\lambda)(f(y) - f(0)) \\ &= \lambda g(x) + (1-\lambda)g(y). \end{aligned}$$

This shows g is a linear function. From the hint, we can represent g as $g(x) = \langle a, x\rangle$ for some $a \in \mathbb{R}^n$. Thus, $f(x) = \langle a, x\rangle + b$ where $b = f(0)$. □

(c) *Proof.* If $f(x) = \langle a, x\rangle + b$, then for any $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$

$$\begin{aligned} f(\lambda x + (1-\lambda)y) &= \langle a, \lambda x + (1-\lambda)y\rangle + b \\ &= \lambda\langle a, x\rangle + (1-\lambda)\langle a, y\rangle + b \\ &= \lambda(\langle a, x\rangle + b) + (1-\lambda)(\langle a, y\rangle + b) \\ &\leq \lambda f(x) + (1-\lambda)f(y). \end{aligned}$$

□

Q3: Convex hulls. Grade:

Show that for any set $X \subseteq \mathbb{R}^n$, the convex hull $conv(X)$ of X (the intersection of all convex sets containing X) is equal to the set of all convex combinations of points in X.
Hint: Define Y to be the set of all convex combinations of points from X, that is,

$$Y = \left\{\sum_{i=1}^{m} \lambda_i x_i \mid m \geq 1, \lambda_i > 0, \sum_{i=1}^{m} \lambda_i = 1\right\},$$

and then prove that both $Y \subseteq conv(X)$ (which may be accomplished by showing that it is convex and contains X), and $conv(X) \subseteq Y$ (which may be accomplished by showing that every convex set containing X also contains Y).

Solution

1. $Y \subseteq conv(X)$

Proof. Let $y_1, y_2 \in Y$. By definition of Y,

$$y_1 = \sum_{i=1}^{m_1} \alpha_i x_i, \qquad \sum_{i=1}^{m_1} \alpha_i = 1,$$
$$y_2 = \sum_{j=1}^{m_2} \beta_j x_j, \qquad \sum_{j=1}^{m_2} \beta_j = 1.$$

For any $\lambda \in [0, 1]$, consider the point $y = \lambda y_1 + (1 - \lambda) y_2$. Then

$$\begin{aligned} y &= \lambda \sum_{i=1}^{m_1} \alpha_i x_i + (1-\lambda) \sum_{j=1}^{m_2} \beta_j x_j \\ &= \sum_{i=1}^{m_1} (\lambda \alpha_i) x_i + \sum_{j=1}^{m_2} ((1-\lambda)\beta_j) x_j \end{aligned}$$

where

$$\begin{aligned} &\sum_{i=1}^{m_1} \lambda \alpha_i + \sum_{j=1}^{m_2} (1-\lambda)\beta_j \\ &= \lambda \sum_{i=1}^{m_1} \alpha_i + (1-\lambda) \sum_{j=1}^{m_2} \beta_j \\ &= \lambda + (1-\lambda) \\ &= 1. \end{aligned}$$

Clearly, $y \in Y$, which shows Y is convex. Also, every point $x_i \in X$ is in Y with $\lambda_i = 1$, which shows $X \subseteq Y$. Since Y is convex and contains X, then it must contain $conv(X)$ as $conv(X)$ is the intersection of all convex sets containing X. □

2. $conv(X) \subseteq Y$

Proof. Let Z be any convex set containing X. We want to show that Z also contains Y. Take any $y \in Y$, since Z is convex and contains X, Z must contain y, the convex combination of points in X. Thus, Z contains Y. Since arbitrary Z contains Y, $conv(X)$ must contain Y as $conv(X)$ is the intersection of all convex sets containing X. Therefore, $conv(X) \subseteq Y$. □

Q4: Affine sets and hulls. Grade:

The scalars λ_i in this problem may take negative values.

(a) The textbook defines a set $X \subseteq \mathbb{R}^n$ as being affine if it is of the form $S + x = \{s + x \mid s \in S\}$ for some $x \in \mathbb{R}^n$ and linear subspace S of $\mathbb{R}^n$. Show that X is affine according to this definition if and only if X is

$$\left.\begin{array}{l} x_1, x_2, \ldots, x_m \in X \\ \lambda_1, \lambda_2, \ldots, \lambda_m \in \mathbb{R} \\ \sum_{i=1}^m \lambda_i = 1 \end{array}\right\} \Rightarrow \sum_{i=1}^m \lambda_i x_i \in X$$

Hint: For the "if", take any $x \in X$ and show that the set $S = X - x = \{x' - x \mid x' \in X\}$ is a linear subspace of $\mathbb{R}^n$.

(b) In the text, the *affine hull* $aff(Y)$ of a set Y is defined to be the intersection of all affine sets containing Y. Show that

$$aff(Y) = \left\{ \sum_{i=1}^{m} \lambda_i y_i \mid m \geq 1, \lambda_i \in \mathbb{R}, \sum_{i=1}^{m} \lambda_i = 1, y_i \in Y \right\},$$

that is, the affine hull of Y is the set of all affine combinations of points in Y.

Solution

(a) • "if"

Proof. Assume that the condition holds for X. We want to show that X is affine by showing S is a linear subspace of $\mathbb{R}^n$. Take any $x \in X$ and let $S = X - x$, we have

(1) $0 \in S$ because $x' - x = 0$ for $x' = x$ and $x' \in X$.

(2) For $s_1, s_2, \ldots, s_m \in S$, $\sum_{i=1}^m \lambda_i s_i = \sum_{i=1}^m \lambda_i (x_i - x) = (\sum_{i=1}^m \lambda_i x_i) - x \in S$ for $\sum_{i=1}^m \lambda_i = 1$.

(3) For any $s \in S$ and any scalar λ, $\lambda s = \lambda(x' - x) = [\lambda(x') + (1-\lambda)x] - x \in S$.

Thus, S is a linear subspace of $\mathbb{R}^n$, which shows X is affine. □

• "only if"

Proof. Suppose X is affine as $X = S + x$ for some $x \in \mathbb{R}^n$ and some linear subspace S of $\mathbb{R}^n$, we want to show that when the conditions hold, $\sum_{i=1}^m \lambda_i x_i \in X$. Since $X = S + x$, for any point $x_i = s_i + x$, we have

$$\sum_{i=1}^{m} \lambda_i x_i = \sum_{i=1}^{m} \lambda_i (s_i + x) = \sum_{i=1}^{m} \lambda_i s_i + \sum_{i=1}^{m} \lambda_i x = \sum_{i=1}^{m} \lambda_i s_i + x \in X.$$

□

(b) The proof of this is similar to Q3 without condition $\sum_{i=1}^m \lambda_i = 1$.

Q5: Arithmetic-Geometric Mean Inequality. Grade:

Show that if $\lambda_1, \lambda_2, \ldots, \lambda_m$ are positive scalars with $\sum_{i=1}^m \lambda_i = 1$, then for every set of positive scalars $x_1, x_2, \ldots, x_m$, we have

$$\prod_{i=1}^{m} x_i^{\lambda_i} \leq \sum_{i=1}^{m} \lambda_i x_i,$$

with equality if and only if $x_1 = x_2 = \cdots = x_m$.
Hint: Show that $-\ln x$ is a strictly convex function on $(0, \infty)$.

Solution

Consider the function $f(x) = -\ln x$, then $f''(x) = \frac{1}{x^2} > 0$ for $x > 0$. Thus, $f(x)$ is strictly convex on $(0, \infty)$. By Jensen's inequality, we have

$$\begin{aligned} -\ln\left(\sum_{i=1}^{m} \lambda_i x_i\right) &= f\left(\sum_{i=1}^{m} \lambda_i x_i\right) \\ &\leq \sum_{i=1}^{m} \lambda_i f(x_i) \\ &= \sum_{i=1}^{m} \lambda_i (-\ln x_i) \end{aligned}$$

$$
\begin{aligned}
&= -\ln\left(\prod_{i=1}^{m} x_i^{\lambda_i}\right) \\
\Rightarrow \exp\left\{-\ln\left(\sum_{i=1}^{m} \lambda_i x_i\right)\right\} &\le \exp\left\{-\ln\left(\prod_{i=1}^{m} x_i^{\lambda_i}\right)\right\} \\
\Rightarrow \sum_{i=1}^{m} \lambda_i x_i &\ge \prod_{i=1}^{m} x_i^{\lambda_i}.
\end{aligned}
$$

Since $f(x)$ is strictly convex, the equality holds if and only if $x_1 = x_2 = \cdots = x_m$.