Convex Sets and Convex Functions

Bolei Zhang

October 22, 2025

1 Convex Sets

Problem 1: The solution set of linear equations is affine set.

Solution: Let $C = \{x|Ax = b\}$ be the solution set of linear equations, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Suppose $x_1, x_2 \in C, i.e., Ax_1 = b, Ax_2 = b$. Then for any θ, we have

$$\begin{aligned} A(\theta x_1 + (1-\theta)x_2) &= \theta A x_1 + (1-\theta)Ax_2 \\ &= \theta b + (1-\theta)b \\ &= b \end{aligned} \tag{1}$$

which shows that the affine combination $\theta x_1 + (1-\theta)x_2$ is also in C.

Problem 2: The set of symmetric semi-positive definite matrices is a convex cone.

Solution: Let $S^n_+ = \{X \in S^n|X \succeq 0\}$ be the set of symmetric semi-positive definite matrices. For any two points $X_1, X_2 \in S^n_+$, let $X = \theta_1 X_1 + \theta_2 X_2$, where $\theta_1 \geq 0, \theta_2 \geq 0$. Then, for any non-zero vector v, there is

$$\begin{aligned} v^T X v &= v^T(\theta_1 X_1 + \theta_2 X_2)v \\ &= \theta_1 v^t X_1 v + \theta_2 v^T X_2 v \\ &\geq 0 \end{aligned} \tag{2}$$

Therefore, S^n_+ is a convex cone.

Problem 3: The ball $B(x_c, r)$ is a convex set.

Proof: For any two points $x_1, x_2 \in B(x_c, r)$. Let $x = \theta x_1 + (1-\theta)x_2$, where $0 \leq \theta \leq 1$. We have

$$\begin{aligned} ||x - x_c||_2 &= ||\theta x_1 + (1-\theta)x_2 - x_c||_2 \\ &= ||\theta(x_1 - x_c) + (1-\theta)(x_2 - x_c)||_2 \\ &\leq \theta||x_1 - x_c||_2 + (1-\theta)||x_2 - x_c||_2 \quad \leq r \end{aligned} \tag{3}$$

Theorem 1: The intersection of any number of convex sets is a convex set.

Proof: Let $A = A_1 \cap A_2 \cap ... \cap A_k$, where $A_i, i = 1, ..., k$ is convex set. For any two points $x_1, x_2 \in A$. Let $x = \theta x_1 + (1-\theta)x_2$, where $0 \leq \theta \leq 1$. We have $x \in A_i, i = 1, ..., k$. Therefore, there is $x \in A$.

Problem 4: The affine operation of convex set is also convex set.

Solution: Recall that a function $f : \mathbb{R}_n \to \mathbb{R}_m$ is affine if it is a sum of a linear function and a constant, i.e., if it has the form $f(x) = Ax + b$, where $A \in \mathbb{R}_{m\times n}$ and $b \in \mathbb{R}_m$. Suppose $S \subseteq R_n$ is convex and $f : \mathbb{R}_n \to \mathbb{R}_m$ is an affine function. Then the image of S under f is

$$f(S) = \{f(x)|x \in S\}.$$

Consider any two points $f(x), f(y)$ in $f(S)$, with their origin points $x, y \in S$. With $0 \leq \theta \leq 1$, there is $f(\theta x + (1-\theta)y) = A\theta x + A(1-\theta)y + b = \theta(Ax + b) + (1-\theta)(Ay + b) = \theta f(x) + (1-\theta)f(y)$. Therefore, the line segment of $f(x)f(y)$ is in $f(S)$, which proves the statement.

Problem 5: Show that a set is convex if and only if its intersection with any line is convex.

Solution: The intersection of two convex sets is convex. Therefore, if S is a convex set, the intersection of S with a line is convex. Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and $x_2 \in S$. The intersection of S with the line through x_1 and x_2 is convex. Therefore, convex combinations of x_1 and x_2 belong to the intersection, hence also to S.

Problem 6: Show that polyhedrons are convex sets

Solution: A polyhedron is defined as the solution set of a finite number of linear equalities and inequalities:

$$P = \{x | a_j^T x \leq b_j, j = 1, ..., m, c_j^T x = d_j, j = 1, ..., p\}.$$

For any two points $x, y \in P$, $j = 1, ..., m$, $0 \leq \theta \leq 1$, there is:

$$a_j^T(\theta x + (1-\theta)y) = \theta a_j^T x + (1-\theta) a_j^T y \leq \theta b_j + (1-\theta) b_j = b_j, \tag{4}$$

and for $j = 1, ..., p$, there is:

$$c_j^T(\theta x + (1-\theta)y) = \theta c_j^T x + (1-\theta) c_j^T y = \theta d_j + (1-\theta) d_j = d_j \tag{5}$$

Therefore, $\theta x + (1-\theta)y$ is also a point in P, showing that P is a convex set.

Problem 7: Determine whether the following sets are convex sets, polyhedra, and give a proof
(1) $\{x \in R^n | \alpha \leq a^T x \leq \beta\}$
(2) $\{x \in R^n | \alpha_1^T x \leq b_1, a_2^T \leq b_2\}$.

Solution: (1) This is an intersection of two halfspaces, hence it is a convex set (and a polyhedron).

(2) This is called a rectangle, which convex set and a polyhedron because it is a finite intersection of halfspaces.

Problem 8: Show that the affine transformation of a convex set is a convex set based on its definition.

Solution: Let S be a convex set in $\mathbb{R}^n$, and $A : \mathbb{R}^n \to \mathbb{R}^m$ be an affine transformation: $A(x) = Bx + c$, for some matrix $B \in \mathbb{R}^{m \times n}$ and $c \in \mathbb{R}^m$.

For any two points $u, v \in A(S)$, the points can be expressed as $u = Bx_1 + c, v = Bx_2 + c$.the line segment connecting u, v can be expressed as $w = tu + (1-t)v = t(Bx_1 + c) + (1-t)(Bx_2 + c) = B(tx_1 + (1-t)x_2) + c$. As S is convex, $tx_1 + (1-t)x_2 \in S$. Therefore, w can be expressed as $A(tx_1 + (1-t)x_2) \in S$.

Thus, the affine transformation of a convex set is also a convex set.

Problem 9: Prove that if S_1 and S_2 are convex sets in $\mathbb{R}^m \times \mathbb{R}^n$, then their Minkowski sum is also a convex set. The Minkowski sum is defined as:

$$S = \{(x, y_1 + y_2 | x \in \mathbb{R}^m, y_1 \in S_1, y_2 \in S_2)\}$$

Solution: Consider two points $(\bar{x}, \bar{y}_1 + \bar{y}_2)$,$(\tilde{x}, \tilde{y}_1 + \tilde{y}_2)$, where $(\bar{x}, \bar{y}_1) \in S_1, (\bar{x}, \bar{y}_2) \in S_2, (\tilde{x}, \tilde{y}_1) \in S_1, (\tilde{x}, \tilde{y}_2) \in S_2$

For any $0 \leq \theta \leq 1$, there is

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

As S_1, S_2 are convex, $(\theta \bar{x} + (1-\theta)\tilde{x}, (\theta \bar{y}_1 + (1-\theta)\tilde{y}_1)) \in S_1, (\theta \bar{x} + (1-\theta)\tilde{x}, (\theta \bar{y}_2 + (1-\theta)\tilde{y}_2)) \in S_2$. Therefore, their convex combination belongs to S, and S is convex.

2 Convex Functions

Problem 10: A function $f(\cdot)$ is convex $\iff$ $\forall x \in \text{dom} f,\ g(t) = f(x+tv), t \in \{t | x+tv \in \text{dom} f\}$ is convex.

Solution: "$\Longrightarrow$": For any $t_1, t_2 \in \text{dom} g$, $x \in \text{dom} f$, there is

$$\begin{aligned} g(\theta t_1 + (1-\theta)t_2) &= f(x + \theta t_1 v + (1-\theta) t_2 v) \\ &= f(\theta(x + t_1 v) + (1-\theta)(x + t_2 v)) \\ &\leq \theta f(x + t_1 v) + (1-\theta) f(x + t_2 v)) \\ &= \theta g(t_1) + (1-\theta) g(t_2) \end{aligned} \tag{6}$$

"$\Longleftarrow$": For any $x_1, x_2 \in \text{dom} f$, we have

$$\begin{aligned} f(\theta x_1 + (1-\theta)x_2) &= f(x_1 + (1-\theta)x_2 - (1-\theta)x_1) \\ &= f(x_1 + (1-\theta)(x_2 - x_1)) \\ &= g((1-\theta)) \\ &\leq \theta g(0) + (1-\theta) g(1) \\ &= \theta f(x_1) + (1-\theta) f(x_2) \end{aligned} \tag{7}$$

Theorem 2: First order condition: For a differentiable function f, if dom f is a convex set, then f is a convex function $\iff$ $f(y) \geq f(x) + \nabla f(x)^T (y-x), \forall x, y \in \text{dom} f$.

Proof: First consider the case $n = 1$: We show that a differentiable function $f : R \to R$ is convex if and only if

$$f(y) \geq f(x) + f'(x)(y-x) \tag{8}$$

for all x and y in dom$f$.

Assume first that f is convex and $x, y \in \text{dom} f$. Since dom$f$ is convex, we conclude that for all $0 < t \leq 1, x + t(y-x) \in \text{dom} f$, and by convexity of f,

$$f(x + t(y-x)) \leq (1-t) f(x) + t f(y). \tag{9}$$

If we divide both sides by t, we obtain

$$f(y) \geq f(x) + \frac{f(x + t(y-x)) - f(x)}{t}, \tag{10}$$

and taking the limit as $t \to 0$ yields (8).

To show sufficiency, assume the function satisfies (8) for all x and y in dom$f$ (which is an interval). Choose any $x \neq= y$, and $0 \leq \theta \leq 1$, and let $z = \theta x + (1-\theta) y$. Applying (8) twice yields

$$f(x) \geq f(z) + f'(z)(x-z), f(y) \geq f(z) + f'(z)(y-z). \tag{11}$$

Multiplying the first inequality by θ, the second by $1-\theta$, and adding them yields

$$\theta f(x) + (1-\theta) f(y) \geq f(z), \tag{12}$$

which proves that f is convex.

Now we can prove the general case, with $f : R^n \to R$. Let $x, y \in R^n$ and consider f restricted to the line passing through them, i.e., the function defined by $g(t) = f(ty + (1-t)x)$, so $g'(t) = \nabla f(ty + (1-t)x)^T (y-x)$.

First assume f is convex, which implies g is convex, so by the argument above we have $g(1) \geq g(0) + g'(0)$, which means

$$f(y) \geq f(x) + \nabla f(x)^T (y-x). \tag{13}$$

Now assume that this inequality holds for any x and y, so if $ty + (1-t)x \in \text{dom} f$, and $\tilde{t}y + (1-\tilde{t})x \in \text{dom} f$, we have

$$f(ty + (1-t)x) \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t-\tilde{t}), \tag{14}$$

i.e., $g(t) \geq g(t) + g'(\tilde{t})(t-\tilde{t})$. We have seen that this implies that g is convex.

Definition 1: Norm functions Given a vector space X, a norm on X is a real-valued function $p : X \to R$ with the following properties, where $|s|$ denotes the usual absolute value of a scalar s:

- Subadditivity/Triangle inequality: $p(x + y) \leq p(x) + p(y)$ for all $x, y \in X$;
- Absolute homogeneity: $p(sx) = |s|\, p(x)$ for all $x \in X$ and all scalars s.
- Positive definiteness/Point-separating: for all $x \in X$, if $p(x) = 0$, then $x = 0$.

Theorem 3: Norm functions are convex.

Proof: If $f : R^n \to R$ is an norm, and $0 \leq \theta \leq 1$,then $f(\theta x + (1-\theta)y) \leq f(\theta x) + f((1-\theta)y) = \theta f(x) + (1-\theta)f(y)$. The inequality follows from the triangle inequality, and the equality follows from homogeneity of a norm.

Problem 11: The maximum function $f(x) = \max\{x_1, x_2, ..., x_n\}$, $x \in \mathbb{R}^m$ is convex.

Solution: The function $f(x) = \max_i x_i$ satisfies, for $0 \leq \theta \leq 1$,

$$\begin{aligned} f(\theta x + (1-\theta)y) &= \max_i(\theta x_i + (1-\theta)y_i) \\ &\leq \theta \max_i x_i + (1-\theta)\max_i y_i \\ &= \theta f(x) + (1-\theta)f(y) \end{aligned} \tag{15}$$

Problem 12: Determine whether $f(x) = x^{-2}, x \neq 0$ is convex function.

Solution: The domain of this function is not a convex set.

Problem 13: $L0$ norm: $||x||_0$ is the number of non-zero elements in x. Explain that whether $L0$ norm is a norm function, whether it is a convex function.

Solution: Let $x = (0, 1)$. There is $||x||_0 = 1$ and $||2x||_0 = 1$. However, $2||x||_0 = 2$, which violates the homogeneity. Therefore, L_0 norm is not a norm function.

Let $x = (0, 1), y = (1, 0)$. Let $\theta = 0.5$,

$$f(\theta x + (1-\theta)y) = ||(0.5, 0.5)||_0 = 2, \tag{16}$$

$$\theta f(x) + (1-\theta)f(y) = 0.5 + 0.5 = 1, \tag{17}$$

Therefore, L_0 norm is not convex.

Problem 14: Determine whether $f(x, y) = \frac{x^2}{y}$, $y > 0$ is a convex function.

Solution: The Hessian matrix of $f(x, y)$ is (for $y > 0$)

$$\nabla^2 f = \begin{bmatrix} \frac{2}{y} & -\frac{2x}{y^2} \\ -\frac{2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix} = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix}^T \begin{bmatrix} y \\ -x \end{bmatrix} \succeq 0. \tag{18}$$

Problem 15: Determine whether the following functions are convex

- $f(x) = e^x - 1, x \in \mathbb{R}$
- $f(x_1, x_2) = x_1 x_2, x_1, x_2 \in \mathbb{R}$

Solution: (1) The function is convex as $\nabla^2 f = e^x > 0$.

(2) The function is not convex as $\nabla^2 f = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is not positive semi-definite.

Problem 16: The KL-divergence is convex $D_{KL}(u, v) = \sum_{i=1}^n (u_i \log \frac{u_i}{v_i} - u_i + v_i)$

Solution: First, we show that $g(x,t) = -t\log(x/t)$ is convex on R^2_{++}: Let dom$f = C$. The domain dom$g = R^2_{++}$ is a convex set. The function $g(\cdot)$ is convex because the function is defined as a scaled and shifted version of convex function $f(x)$.

Therefore, the relative entropy of two vectors $u, v \in R^n_{++}$, defined as

$$\sum_{i=1}^{n} u_i \log(u_i/v_i)$$

is convex in (u,v), since it is a sum of relative entropies of u_i, v_i.

The KL-divergence is convex as it is the relative entropy plus a linear function of (u,v).

Problem 17: Convex-concave functions and saddle-points. We say the function $f : R^n \times R^m \to R$ is convex-concave if $f(x,z)$ is a concave function of z, for each fixed x, and a convex function of x, for each fixed z. We also require its domain to have the product form dom$f = A \times B$, where $A \subseteq R^n$ and $B \subseteq R^m$ are convex.

(1) Give a second-order condition for a twice differentiable function $f : R^n \times R^m \to R$ to be convex-concave, in terms of its Hessian $\nabla^2 f(x,z)$.

(2) Suppose that $f : R^n \times R^m \to R$ is convex-concave and differentiable, with $\nabla f(\tilde{x},\tilde{z}) = 0$. Show that the saddle-point property holds: for all x, z, we have

$$f(\tilde{x},z) \le f(\tilde{x},\tilde{z}) \le f(x,\tilde{z})$$

. Show that this implies that f satisfies the *strong max-min property*:

$$\sup_z \inf_x f(x,z) = \inf_x \sup_z f(x,z)$$

(and their common value is $f(\tilde{x},\tilde{z})$.

(3) Now suppose that $f : R^n \times R^m \to R$ is differentiable, but not necessarily convex-concave, and the saddle-point property holds at $\tilde{x},\tilde{z}$:

$$f(\tilde{x},z) \le f(\tilde{x},\tilde{z}) \le f(x,\tilde{z})$$

for all x, z. Show that $\nabla f(\tilde{x},\tilde{z}) = 0$

Solution: (1) The Hessian matrix of f is

$$\nabla^2 f = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where $A_{11} \in R^{n\times n}$, $A_{12} \in R^{n\times m}$, $A_{21} \in R^{m\times n}$, $A_{22} \in R^{m\times m}$. And there is $\nabla^2 f = A_{11}$, $\nabla^2_z f = A_{22}$. As f is convex when z is fixed, then A_{11} is positive semi-definite; As f is concave when x is fixed, then A_{22} is negative semi-definite.

(2) For the first inequality, as $f(x,z)$ is convex of x, there is

$$f(x,\tilde{z}) \ge f(\tilde{x},\tilde{z}) + \nabla_x f(\tilde{x},\tilde{z})(x - \tilde{x}) = f(\tilde{x},\tilde{z})$$

The same is true for the left part of the inequality.

Next we prove $\sup_z \inf_x f(x,z) = f(\tilde{x},\tilde{z})$. First

$$\inf_x f(x,z) \le f(\tilde{x},z) \le f(\tilde{x},\tilde{z}),$$

As

$$\sup_z \inf_x f(x,z) \le f(\tilde{x},\tilde{z}),$$

and

$$\sup_z \inf_x f(x,z) \ge \inf_x f(x,\tilde{z}) \ge f(\tilde{x},\tilde{z}),$$

Therefore, there is $\sup_z \inf_x f(x,z) = f(\tilde{x},\tilde{z})$. The same is true for $\inf_x \sup_z f(x,z) = f(\tilde{x},\tilde{z})$, which proves the equality.

(3) We need to prove $\nabla_x f(\tilde{x}, \tilde{z}) = 0$ and $\nabla_z f(\tilde{x}, \tilde{z}) = 0$. We prove by contradiction. Assume $\nabla_x f(\tilde{x}, \tilde{z}) \neq 0$, let $v = (\nabla_x f(\tilde{x}, \tilde{z})^T, 0)^T$. Get the first order of $f(\tilde{x} + tv, \tilde{z})$ at $(\tilde{x}, \tilde{z})$ $(t \neq 0)$,

$$f(\tilde{x} + tv, \tilde{z}) = f(\tilde{x}, \tilde{z}) + t||\nabla_x f(\tilde{x}, \tilde{z})||^2 + O(t^2).$$

Take $t < 0$ with small enough absolute value, there is $f(\tilde{x} + tv, \tilde{z}) < f(\tilde{x}, \tilde{z})$, which contradicts with the assumption. Therefore, there is $\nabla_x f(\tilde{x}, \tilde{z}) = 0$ and $\nabla_z f(\tilde{x}, \tilde{z}) = 0$.

Problem 18: Compute the conjugate of the following functions:
(1) $f(x) = ax + b$, $\mathrm{dom} f = \mathbb{R}$;
(2) $f(x) = -\log x$, $\mathrm{dom} f = \mathbb{R}_{++}$.

Solution: (1) The is $yx - ax - b$. This function is bouned if and only if $y = a$, in which case it is constant. Therefore, the domain of the conjugate function f^* is the singleton $\{a\}$, and $f^*(a) = -b$.

(2) The function is $xy + \log x$. This function is unbounded above if $y \geq 0$ and reaches its maximum at $x = -1/y$ otherwise. Therefore, $\mathrm{dom} f^* = \{y | y < 0\} = -\mathbb{R}_{++}$ and $f^*(y) = -\log(-y) - 1$ for $y < 0$.

Problem 19: Show that $f(Ax + b)$ is convex if $f(x)$ is a convex function.

Solution: The domain of $f(Ax + b)$ is the same with $f(x)$, which is a convex set. For any two points $x, y \in \mathrm{dom} f$, $0 \leq \theta \leq 1$, there is,

$$f(A(\theta x + (1-\theta)y) + b) = f(\theta(Ax + b) + (1-\theta)(Ay + b)) \leq \theta f(Ax + b) + (1-\theta) f(Ay + b). \quad (19)$$

Therefore, $f(Ax + b)$ is convex.

Problem 20: Conjugate of convex plus affine function: Define $g(x) = f(x) + c^T x + d$, where f is convex. Express g^* in terms of f^* (and c, d).

Solution:

$$\begin{aligned} g^*(y) &= \sup(y^T x - f(x) - c^T x - d) \\ &= \sup((y - c)^T x - f(x)) - d \\ &= f^*(y - c) - d \end{aligned} \quad (20)$$