APM 3F007
Convex Optimization and Optimal Control

Summary of Definitions, Propositions, and Theorems

Real Analysis Prerequisites

0.1 Real Numbers

Definition

Definition 0.1 (Supremum and infimum)**.** *Let $A \subset \mathbb{R}$. A number $S \in \mathbb{R}$ is the* supremum *(least upper bound) of A if*

$$S = \min\{\, M \in \mathbb{R} \text{ s.t. } x \leq M \ \forall x \in A \,\}.$$

A number $I \in \mathbb{R}$ is the infimum *(greatest lower bound) of A if*

$$I = \max\{\, m \in \mathbb{R} \text{ s.t. } x \geq m \ \forall x \in A \,\}.$$

We write $\sup A = S$ and $\inf A = I$. By convention, $\sup A = +\infty$ if A is not bounded above and $\inf A = -\infty$ if A is not bounded below.

If $A \subset B$, then $\sup A \leq \sup B$ and $\inf A \geq \inf B$. For a map $f : A \to \mathbb{R}$ we set

$$\sup_A f := \sup\{f(x) : x \in A\} \in \mathbb{R} \cup \{+\infty\}, \qquad \inf_A f := \inf\{f(x) : x \in A\} \in \mathbb{R} \cup \{-\infty\},$$

and

$$\arg\max_{x \in A} f := \{x \in A : f(x) = \sup_A f\}, \qquad \arg\min_{x \in A} f := \{x \in A : f(x) = \inf_A f\}.$$

Theorem

Theorem 0.1 (Bolzano–Weierstrass)**.** *Every bounded sequence of real numbers admits a convergent subsequence.*

Definition

Definition 0.2 (Minimizing / maximizing sequence)**.** *Let $A \subset \mathbb{R}$ be nonempty. A sequence $(x_n) \subset A$ is* minimizing *if $\lim_{n\to\infty} x_n = \inf A$. Similarly, for $B \subset \mathbb{R}$, $(y_n) \subset B$ is* maximizing *if $\lim_{n\to\infty} y_n = \sup B$. If $f : U \to \mathbb{R}$, a* minimizing sequence for f *is $(x_n) \subset U$ with $\lim_{n\to\infty} f(x_n) = \inf_U f$.*

0.2 Normed and Inner Product Spaces

Definition

Definition 0.3 (Norm). *Let X be a real vector space. A map $\|\cdot\| : X \to \mathbb{R}_+$ is a* norm *if for all $x, y \in X$ and $\alpha \in \mathbb{R}$:*

(i) $\|x\| = 0 \iff x = 0$*;*

(ii) $\|\alpha x\| = |\alpha|\,\|x\|$*;*

(iii) ***Triangle inequality*** $\|x + y\| \le \|x\| + \|y\|$.

It follows that $|\|x\| - \|y\|| \le \|x - y\|$.

Definition

Definition 0.4 (Balls and neighborhoods). *For $z \in X$ and $\rho > 0$ define the open ball $B_\rho(z) := \{x \in X : \|x - z\| < \rho\}$ and the closed ball $\overline{B}_\rho(z) := \{x \in X : \|x - z\| \le \rho\}$. A set $U \subset X$ is a* neighborhood *of z if $B_\rho(z) \subset U$ for some $\rho > 0$.*

Definition

Definition 0.5 (Open/closed sets; closure and interior). *A set $A \subset X$ is* open *if it is a neighborhood of each of its points, and* closed *if $X \setminus A$ is open. The* closure *$\overline{A}$ is $\{x \in X : \forall \rho > 0,\ B_\rho(x) \cap A \neq \emptyset\}$, the smallest closed set containing A. The* interior *$\operatorname{int} A$ is $\{x \in X : \exists \rho > 0,\ B_\rho(x) \subset A\}$, the largest open set contained in A.*

Definition

Definition 0.6 (Equivalent norms). *Two norms $\|\cdot\|_1, \|\cdot\|_2$ on X are* equivalent *if $\exists\, \alpha, \beta > 0$ such that $\alpha\|x\|_1 \le \|x\|_2 \le \beta\|x\|_1$ for all $x \in X$.*

Theorem

Theorem 0.2 (Equivalence in finite dimension). *If $\dim X < \infty$, then all norms on X are equivalent.*

Definition

Definition 0.7 (Inner product). *An* inner product *on X is a map $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{R}$ that is symmetric, linear in each argument, and positive definite: $\langle x, x\rangle > 0$ for all $x \neq 0$. It induces the* canonical norm *$\|x\| := \sqrt{\langle x, x\rangle}$ and satisfies Cauchy–Schwarz $|\langle x, y\rangle| \le \|x\|\,\|y\|$.*

Useful identities (all x, y in an inner product space):

$$\|x+y\|^2 = \|x\|^2+\|y\|^2+2\langle x,y\rangle, \quad \|x+y\|^2+\|x-y\|^2 = 2(\|x\|^2+\|y\|^2), \quad \langle x,y\rangle = \tfrac{1}{4}(\|x+y\|^2-\|x-y\|^2).$$

For $A \subset X$ define $A^{\perp} := \{x \in X : \langle x, y \rangle = 0 \ \forall y \in A\}$, a closed linear subspace.

Definition

Definition 0.8 (Product spaces)**.** *If X_1, X_2 are normed (resp. inner product) spaces, then $X_1 \times X_2$ is normed by $\|(x_1, x_2)\| := \sqrt{\|x_1\|^2 + \|x_2\|^2}$ (resp. $\langle (x_1, x_2), (y_1, y_2) \rangle := \langle x_1, y_1 \rangle + \langle x_2, y_2 \rangle$).*

0.3 Sequences and Compactness

Definition

Definition 0.9 (Convergence)**.** *A sequence $(x_n) \subset X$* converges *to $x \in X$ if $\forall \varepsilon > 0 \ \exists N$ such that $n \geq N \Rightarrow \|x_n - x\| \leq \varepsilon$.*

Proposition

Proposition 0.1 (Closed-set characterization)**.** *Let $A \subset E \subset X$. Then A is closed in E iff: for every sequence $(x_n) \subset A$ and every $x \in E$, $x_n \to x$ implies $x \in A$.*

Definition

Definition 0.10 (Compactness)**.** *$A \subset X$ is* compact *if every sequence in A admits a subsequence converging to a point of A.*

Theorem

Theorem 0.3 (Heine–Borel)**.** *If $\dim X < \infty$, then $C \subset X$ is compact iff it is closed and bounded.*

Proposition

Proposition 0.2 (Closedness of Minkowski sums)**.** *If A is closed and B is compact in a normed space X, then $A + B := \{a + b : a \in A, b \in B\}$ is closed.*

Definition

Definition 0.11 (Cauchy, Banach, Hilbert)**.** *A sequence (x_n) is* Cauchy *if $\forall \varepsilon > 0 \ \exists N$ s.t. $n, m \geq N \Rightarrow \|x_n - x_m\| \leq \varepsilon$. A normed space in which every Cauchy sequence converges is* complete *(a* Banach *space). A complete inner product space is a* Hilbert *space.*

Theorem

Theorem 0.4 (Banach fixed point). *Let X be Banach, $U \subset X$ closed, and $f : U \to U$ a contraction: $\exists\, 0 \le \lambda < 1$ s.t. $\|f(x) - f(y)\| \le \lambda\|x - y\|$. Then f admits a unique fixed point $\bar{x} \in U$.*

0.4 Continuity

Definition

Definition 0.12 (Limit and continuity). *Let $f : U \subset X \to Y$ between normed spaces, $z \in \overline{U}$, $\ell \in Y$. We write $\lim_{x\to z} f(x) = \ell$ if for all $\varepsilon > 0$ there exists $\eta > 0$ such that $x \in U \cap B_\eta(z) \Rightarrow f(x) \in B_\varepsilon(\ell)$. The sequential characterization holds: $\lim_{x\to z} f(x) = \ell$ iff for every $x_n \to z$ with $x_n \in U$, we have $f(x_n) \to \ell$. A map f is* continuous at *$x_0 \in U$ if $\lim_{x\to x_0} f(x) = f(x_0)$; it is* continuous *if so at every $x_0 \in U$.*

Proposition

Proposition 0.3 (Open/closed preimages). *For $f : U \subset X \to Y$, the following are equivalent:*

(i) *f is continuous;*

(ii) *For every open $A \subset Y$, $f^{-1}(A)$ is open in U;*

(iii) *For every closed $A \subset Y$, $f^{-1}(A)$ is closed in U.*

Theorem

Theorem 0.5 (Continuity preserves compactness). *If $A \subset U$ is compact and $f : U \to Y$ is continuous, then $f(A)$ is compact.*

Definition

Definition 0.13 (Lipschitz continuity). *$f : U \to Y$ is* L-Lipschitz *if $\|f(x) - f(y)\| \le L\|x - y\|$ for all $x, y \in U$.*

Proposition

Proposition 0.4 (Continuous linear maps and operator norm). *For a linear $T : X \to Y$ between normed spaces, the following are equivalent:*

T is continuous; T is continuous at 0; T is Lipschitz; $\sup_{x\neq 0} \dfrac{\|Tx\|}{\|x\|} < \infty$.

Write $\mathcal{L}(X,Y)$ for the space of continuous linear maps with norm

$$\|T\|_{\mathcal{L}(X,Y)} := \sup_{x\neq 0} \frac{\|Tx\|}{\|x\|}.$$

If Y is Banach, then $\mathcal{L}(X,Y)$ is Banach.

0.5 Differentiability

Definition

Definition 0.14 (Fréchet derivative)**.** *Let $U \subset X$ be open, $f : U \to Y$. We say that f is* Fréchet differentiable *at $x_0 \in U$ if there exists a continuous linear map $L \in \mathcal{L}(X,Y)$ such that*

$$\lim_{h\to 0} \frac{\|f(x_0+h) - f(x_0) - L(h)\|}{\|h\|} = 0.$$

The map L is unique and is denoted $df(x_0)$. If $x \mapsto df(x)$ is continuous, then f is continuously Fréchet differentiable.

Proposition

Proposition 0.5 (Jacobian in finite dimension)**.** *If $f : U \subset \mathbb{R}^n \to \mathbb{R}^m$ is Fréchet differentiable at x, then $df(x)$ is represented in canonical bases by the Jacobian*

$$Df(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix}.$$

Theorem

Theorem 0.6 (Mean value inequality in normed spaces)**.** *Let $U \subset X$ be open and convex, $f : U \to Y$ Fréchet differentiable, and $x, y \in U$ with the segment $[x,y] \subset U$. Then*

$$\|f(x) - f(y)\| \;\le\; \|x - y\| \sup_{0\le\theta\le 1} \|df(\theta x + (1-\theta)y)\|_{\mathcal{L}(X,Y)}.$$

Theorem

Theorem 0.7 (Chain rule)**.** *Let $f : U \subset X \to Y$ and $g : V \subset Y \to Z$ with $f(x_0) \in V$. If f is Fréchet differentiable at x_0 and g at $f(x_0)$, then $g \circ f$ is Fréchet differentiable at x_0 with*

$$d(g\circ f)(x_0)\, h \;=\; dg(f(x_0))\, (df(x_0)\, h).$$

Theorem

Theorem 0.8 (Implicit function theorem (Fréchet))**.** *Let X, Y, Z be Banach spaces and $F : X \times Y \to Z$ be continuously Fréchet differentiable. Let $(x_0, y_0) \in X \times Y$ with $F(x_0, y_0) = 0$ and suppose $h \mapsto dF(x_0, y_0)(0, h)$ is a linear isomorphism $Y \to Z$. Then there exist neighborhoods U of x_0 and V of y_0, and a Fréchet differentiable map $\varphi : U \to V$ such that*

$$F(x, y) = 0 \quad \Longleftrightarrow \quad y = \varphi(x) \qquad \textit{for all } (x, y) \in U \times V.$$

Chapter 1

Convex Sets and Convex Hulls

1.1 Convex Sets

Definition

Definition 1.1 (Convex set)**.** *Let X be a real vector space. A set $A \subset X$ is said to be* convex *if for all $x, y \in A$ and for all $t \in [0,1]$, we have*

$$tx + (1-t)y \in A.$$

We always assume that the underlying field is $\mathbb{R}$.

From this definition, several properties follow immediately.

Proposition

Proposition 1.1 (Intersection of convex sets)**.** *The intersection of any family of convex sets is convex:*

$$\bigcap_{i \in I} A_i \text{ is convex if each } A_i \text{ is convex.}$$

Definition

Definition 1.2 (Minkowski sum and scalar multiplication)**.** *Let $A, B \subset X$ and $\alpha \in \mathbb{R}$. Their* Minkowski sum *and* scalar multiple *are defined by*

$$A + B = \{a + b : \ a \in A, \ b \in B\}, \qquad \alpha A = \{\alpha a : \ a \in A\}.$$

Proposition

Proposition 1.3 (Operations preserving convexity)**.** *If A, B are convex subsets of X and $\alpha, \beta \in \mathbb{R}$, then:*

$$A + B \text{ is convex}, \quad \alpha A \text{ is convex.}$$

1.2 Topological Properties

Let $(X, \|\cdot\|)$ be a normed vector space.

Proposition

Proposition 1.4 (Balls are convex). *For any norm $\|\cdot\|$, both the open ball $B_r(a) = \{x \in X : \|x - a\| < r\}$ and the closed ball $\overline{B}_r(a) = \{x \in X : \|x - a\| \leq r\}$ are convex.*

Proposition

Proposition 1.5 (Closure and interior). *If $A \subset X$ is convex, then its closure $\overline{A}$ and its interior $\operatorname{int} A$ are also convex.*

1.3 Convex Combinations

Definition

Definition 1.3 (Convex combination). *Let X be a vector space. A* convex combination *of m points $x_1, \ldots, x_m \in X$ is a point*

$$x = \sum_{i=1}^{m} \theta_i x_i, \quad \text{where } \theta_i \geq 0, \quad \sum_{i=1}^{m} \theta_i = 1.$$

Proposition

Proposition 1.6. *If A is convex and $x_1, \ldots, x_m \in A$, then every convex combination of the x_i belongs to A.*

Proof (by induction). For $m = 1$ or $m = 2$, the result follows directly from the definition of convexity. Assume it holds for $m - 1$ elements. Write

$$x = \theta_1 x_1 + (1 - \theta_1) \sum_{i=2}^{m} \frac{\theta_i}{1 - \theta_1} x_i,$$

and note that the second term is a convex combination of $x_2, \ldots, x_m$. By the induction hypothesis, it belongs to A, and hence so does x.

1.4 Convex Hull and Carathéodory Theorem

Definition

Definition 1.4 (Convex hull). *Let $A \subset X$. The* convex hull *of A, denoted* $\operatorname{conv}(A)$, *is* *the smallest convex set containing A:*

$$\operatorname{conv}(A) = \bigcap_{\substack{C \supset A \\ C \text{ convex}}} C.$$

Equivalently, it is the set of all convex combinations of finitely many points of A.

Proposition

Proposition 1.7. *If A is convex, then* $\operatorname{conv}(A) = A$*.*

Theorem

Theorem 1.1 (Carathéodory)**.** *Let X be a real vector space of dimension n. Every point of* $\operatorname{conv}(A) \subset X$ *can be expressed as a convex combination of at most $n+1$ points of A.*

Consequences.

- In $\mathbb{R}^2$, any convex combination of more than three points can be written as a convex combination of at most three points.
- Therefore, any convex set in the plane can be represented using at most three generating points (triangles are the basic convex polygons in $\mathbb{R}^2$).

1.5 Exercise Results and Applications

Proposition

Proposition 1.8 (Convexity of intersections and unions)**.** *Let $S = \{x \in \mathbb{R}^2 : \ 0 \leq x_i \leq 1, \ i = 1, 2\}$ and $D = \{x \in \mathbb{R}^2 : \|x\|_2 \leq 1\}$. Then:*

$$S \cap D \text{ is convex}, \qquad S \cup D \text{ is convex}.$$

Proposition

Proposition 1.9 (Convex hull of two segments)**.** *Let $A = [0,1] \times \{0\}$ and $B = [1,2] \times \{1\}$. Then the convex hull of $A \cup B$ is a* parallelogram*.*

Proposition

Proposition 1.10 (Closure of a convex set)**.** *If A is convex, then its closure $\overline{A}$ is convex. However, if A is only closed, $\overline{A} = A$ need not be convex.*

Theorem

Theorem 1.2 (Carathéodory in $\mathbb{R}^2$). *In dimension* 2*, every convex combination of* 5 *elements can be expressed as a convex combination of at most* 3 *elements.*

Key facts summary:

- Intersection of convex sets $\Rightarrow$ convex.
- Sum or scalar multiple of convex sets $\Rightarrow$ convex.
- Closure and interior of convex sets $\Rightarrow$ convex.
- Convex hull of any set $=$ all its convex combinations.
- In $\mathbb{R}^n$, Carathéodory $\Rightarrow n+1$ points suffice.
- In $\mathbb{R}^2$, convex combinations $\Rightarrow$ at most 3 points.

Chapter 2

Projections, Cones, and the Riesz Theorem

2.1 Projection onto a Closed Convex Set

Theorem

Theorem 2.1 (Projection theorem)**.** *Let $(X, \langle \cdot, \cdot \rangle)$ be a Hilbert space. Let $A \subset X$ be nonempty, closed, and convex. Then for every $x \in X$, there exists a unique $y \in A$ such that*

$$\|x - y\| = \min_{z \in A} \|x - z\|.$$

The point y is called the orthogonal projection of x onto A *and is denoted $P_A(x)$. It is characterized by*

$$\langle x - y, z - y \rangle \leq 0, \quad \forall z \in A.$$

Proposition

Proposition 2.1 (Uniqueness and geometric characterization)**.** *The minimizing point $y = P_A(x)$ is unique. Geometrically, the vector $x - y$ is orthogonal to the supporting half-space of A at y:*

$$\langle x - y, z - y \rangle \leq 0 \quad \forall z \in A.$$

Proposition

Proposition 2.2 (Lipschitz continuity of the projection)**.** *Under the assumptions of the theorem, the projection operator $P_A : X \to A$ is* 1*-Lipschitz:*

$$\|P_A(x) - P_A(y)\| \leq \|x - y\|, \qquad \forall x, y \in X.$$

Proposition

Proposition 2.3 (Linearity on a subspace)**.** *If $M \subset X$ is a closed linear subspace, then*

P_M is linear and called the orthogonal projection *on M. It satisfies*

$$\langle x - P_M(x), v \rangle = 0, \quad \forall v \in M.$$

Moreover $X = M \oplus M^\perp$.

Idea. Let $x, y \in X$ and use the characterization above for both x and y. By linearity of the inner product, adding the equalities shows $P_M(x+y) = P_M(x) + P_M(y)$. A similar argument gives $P_M(\alpha x) = \alpha P_M(x)$ for any $\alpha \in \mathbb{R}$. □

2.2 Cones

Definition

Definition 2.1 (Cone). *A subset C of a vector space X is called a* cone *if for all $(\alpha, x) \in \mathbb{R}_+ \times C$, we have $\alpha x \in C$. Equivalently:*

$$x \in C,\ \alpha \geq 0 \quad \Rightarrow \quad \alpha x \in C.$$

The zero vector always belongs to a cone (take $\alpha = 0$).

Proposition

Proposition 2.4 (Examples). *(i) $\{(x,t) \in \mathbb{R}^n \times \mathbb{R} : \|x\| \leq t\}$ is a convex cone (second-order or Lorentz cone).*

(ii) $\{x \in \mathbb{R}^n : Ax \geq 0\}$, for $A \in \mathbb{R}^{m \times n}$, is a closed convex cone (intersection of half-spaces).

Proposition

Proposition 2.5 (Operations preserving conicity). *If $(C_i)_{i \in I}$ are cones of X, then*

$$\bigcap_{i \in I} C_i \text{ is a cone.}$$

If C_1, C_2 are cones, then $C_1 + C_2$ is also a cone. If C is a cone, then its closure $\overline{C}$ is a cone.

Proposition

Proposition 2.6 (Convex cones). *A cone C is said to be* convex *if for any $x, y \in C$ and $t \in [0,1]$, $tx + (1-t)y \in C$. Intersections and closures of convex cones are convex cones.*

2.3 Application: Riesz Representation Theorem

Theorem

Theorem 2.2 (Riesz representation theorem). *Let X be a Hilbert space and $\varphi : X \to \mathbb{R}$ a continuous linear functional. Then there exists a unique $w \in X$ such that*

$$\langle w, x\rangle = \varphi(x), \quad \forall x \in X.$$

The vector w is called the Riesz representative *of φ.*

Sketch of proof using the projection theorem. 1) Let $H = \ker(\varphi) = \{x : \varphi(x) = 0\}$; it is a closed linear subspace (by linearity and continuity of φ).

2) If $H = X$, then $\varphi \equiv 0$ and $w = 0$.

3) Otherwise, pick $u \in X \setminus H$ and let $v = \dfrac{P_H(u) - u}{\|P_H(u) - u\|}$. Then $\varphi(v) \neq 0$.

4) For arbitrary $x \in X$, choose $\lambda = \varphi(x)/\varphi(v)$ so that $y = x - \lambda v \in H$.

5) Using orthogonality, $\langle y, v\rangle = 0$ and $\lambda = \langle x, v\rangle$.

6) Set $w = \varphi(v)v$; then $\langle w, x\rangle = \varphi(v)\langle v, x\rangle = \varphi(x)$ for all x.

7) Uniqueness follows: if w_1, w_2 satisfy the identity, $\langle w_1 - w_2, x\rangle = 0$ for all x, hence $w_1 = w_2$. □

2.4 Adjoint Operator — Application of Riesz

Theorem

Theorem 2.3 (Adjoint operator). *Let $(X, \langle\cdot,\cdot\rangle_X)$ and $(Y, \langle\cdot,\cdot\rangle_Y)$ be Hilbert spaces, and let $L \in \mathcal{L}(X, Y)$ be a continuous linear map. Then there exists a unique $L^* \in \mathcal{L}(Y, X)$ such that*

$$\langle Lx, y\rangle_Y = \langle x, L^* y\rangle_X, \quad \forall (x, y) \in X \times Y.$$

The operator L^* is called the adjoint *of L.*

Construction. 1) For fixed $y \in Y$, define $\varphi_y : X \to \mathbb{R}$ by $\varphi_y(x) = \langle Lx, y\rangle_Y$. This map is linear and continuous because

$$|\varphi_y(x)| \leq \|L\|\|x\|_X\|y\|_Y.$$

2) By Riesz, there exists a unique $w_y \in X$ such that

$$\langle w_y, x\rangle_X = \langle Lx, y\rangle_Y, \quad \forall x \in X.$$

3) Define $L^* : Y \to X$ by $L^*(y) = w_y$. Then $\langle Lx, y\rangle_Y = \langle x, L^* y\rangle_X$ for all x, y.

4) Linearity: for $\alpha, \beta \in \mathbb{R}$ and $y_1, y_2 \in Y$, one checks

$$L^*(\alpha y_1 + \beta y_2) = \alpha L^* y_1 + \beta L^* y_2.$$

5) Continuity: for any $y \in Y$,

$$\|L^*y\|_X^2 = \langle L^*y, L^*y\rangle_X = \langle L(L^*y), y\rangle_Y \leq \|L\|\|L^*y\|_X\|y\|_Y,$$

hence $\|L^*y\|_X \leq \|L\|\|y\|_Y$. Thus L^* is continuous and $\|L^*\| \leq \|L\|$.

Uniqueness follows from Riesz representation uniqueness. □

2.5 Summary of Key Results

- Projection theorem $\Rightarrow$ existence and uniqueness of $P_A(x)$ for closed convex A in Hilbert spaces.
- P_A is 1-Lipschitz.
- For a closed subspace M, P_M is linear and orthogonal: $\langle x - P_M(x), v\rangle = 0$.
- Cones: closed under nonnegative scaling; intersections, sums, and closures of cones remain cones.
- Riesz theorem: every continuous linear functional on a Hilbert space has a unique vector representative.
- Adjoint operator: for $L \in \mathcal{L}(X, Y)$, there exists a unique $L^* \in \mathcal{L}(Y, X)$ such that $\langle Lx, y\rangle_Y = \langle x, L^*y\rangle_X$.

Chapter 3

Convex Functions and Polar/Normal/-Tangent Cones

3.1 Convex functions: epigraph, domain, strictness

Definition

Definition 3.1 (Effective domain and properness)**.** *Let X be a real vector space and $f : X \to \mathbb{R} \cup \{+\infty\}$. The* effective domain *is* $\operatorname{dom} f := \{x \in X : f(x) < +\infty\}$. *We say f is* proper *if $f \not\equiv +\infty$ and $f(x) > -\infty$ for all x.*

Definition

Definition 3.2 (Epigraph and convexity)**.** *The* epigraph *of f is* $\operatorname{epi} f := \{(x, u) \in X \times \mathbb{R} : f(x) \le u\}$. *A function f is* convex *iff* $\operatorname{epi} f$ *is a convex subset of $X \times \mathbb{R}$.*

Proposition

Proposition 3.1 (Pointwise characterization)**.** *A function $f : X \to \mathbb{R} \cup \{+\infty\}$ is convex iff for all $x, y \in \operatorname{dom} f$ and $t \in [0, 1]$,*

$$f(tx + (1-t)y) \;\le\; t\, f(x) + (1-t)\, f(y).$$

Definition

Definition 3.3 (Strict convexity)**.** *A function f is* strictly convex *if for all $x \neq y$ in* $\operatorname{dom} f$ *and $t \in (0, 1)$,*

$$f(tx + (1-t)y) \;<\; t\, f(x) + (1-t)\, f(y).$$

Proposition

Proposition 3.2 (Canonical examples). *(i) Any norm $\|\cdot\|$ is convex, but not strictly convex in general.*

(ii) In an inner-product space, $x \mapsto \|x\|^2$ is strictly convex.

Hints. (i) Use the triangle inequality $\|tx+(1-t)y\| \le t\|x\|+(1-t)\|y\|$, and show equality can hold with $x \neq y$ when y is a positive multiple of x. (ii) Expand $\|tx+(1-t)y\|^2$ with polarization to get $t\|x\|^2+(1-t)\|y\|^2-t(1-t)\|x-y\|^2$. □

Proposition

Proposition 3.3 (A discontinuous convex function). *The function $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ given by*

$$f(x) = \begin{cases} 0, & x > 0, \\ 1, & x = 0, \\ +\infty, & x < 0 \end{cases}$$

is convex, $\operatorname{dom} f = [0, +\infty)$*, and f is not continuous at* 0*.*

Proposition

Proposition 3.4 (Distance to a convex set is convex). *Let $(X, \|\cdot\|)$ be a normed space and $\Omega \subset X$ be nonempty and convex. Then $d(\cdot, \Omega) : x \mapsto \inf_{y\in\Omega} \|x-y\|$ is convex on X.*

Proposition

Proposition 3.5 (Subspace minimization (partial infimum preserves convexity)). *Let X, Y be vector spaces and $f : X \times Y \to \mathbb{R}\cup\{+\infty\}$ be convex. Define $g(x) := \inf_{y\in Y} f(x, y)$. Then g is convex.*

3.2 Cones: positive/polar cones and orthants

Definition

Definition 3.4 (Cones). *A set $C \subset X$ is a* cone *if $\alpha x \in C$ for all $x \in C$ and $\alpha \ge 0$. Intersections, sums, and closures of cones are cones.*

Definition

Definition 3.5 (Positive and polar (negative) cones). *Let X be a Hilbert space and $C \subset X$. Define*

$$C^+ := \{x \in X : \langle x, y\rangle \ge 0, \ \forall y \in C\}, \qquad C^\ominus := \{x \in X : \langle x, y\rangle \le 0, \ \forall y \in C\} = -C^+.$$

Proposition

Proposition 3.6. *C^+ and $C^\ominus$ are always closed convex cones (even if C is neither convex nor a cone).*

Proposition

Proposition 3.7 (Orthants)**.** *For the positive orthant $\mathbb{R}^n_+ = \{x : \ x_i \geq 0\}$ and the negative orthant $\mathbb{R}^n_- = -\mathbb{R}^n_+$,*

$$(\mathbb{R}^n_+)^+ = \mathbb{R}^n_+, \qquad (\mathbb{R}^n_+)^\ominus = \mathbb{R}^n_-.$$

3.3 Normal and tangent cones to a convex set

Definition

Definition 3.6 (Normal cone)**.** *Let $A \subset X$ be convex and $a \in A$. The* normal cone *of A at a is*

$$N_A(a) := \{x \in X : \ \langle x, y - a\rangle \leq 0, \ \forall y \in A\} = (A - a)^\ominus.$$

Definition

Definition 3.7 (Tangent cone)**.** *Let $A \subset X$ be convex and $a \in A$. The* tangent cone *at a is the smallest closed convex cone containing $A - a$:*

$$T_A(a) \ := \ \bigcap \{ C : \ C \text{ closed convex cone and } A - a \subset C \}.$$

Equivalently, $T_A(a)$ is the closure of the radial cone

$$T^\circ_A(a) := \{ \ \alpha(x - a) : \ \alpha \geq 0, \ x \in A \ \}.$$

Proposition

Proposition 3.8 (Radial cone is convex and $\overline{T^\circ_A(a)} = T_A(a)$)**.** *$T^\circ_A(a)$ is a convex cone and $\overline{T^\circ_A(a)} = T_A(a)$.*

Idea. If $x' = \alpha(x - a)$ and $y' = \beta(y - a)$ with $\alpha, \beta \geq 0$, then for any $\theta \in [0, 1]$,

$$\theta x' + (1 - \theta)y' = (\alpha\theta + \beta(1 - \theta))\Big(\frac{\alpha\theta}{\alpha\theta + \beta(1 - \theta)} x + \frac{\beta(1 - \theta)}{\alpha\theta + \beta(1 - \theta)} y - a\Big),$$

and the bracket is in A by convexity. Minimality of $T_A(a)$ among closed convex cones containing $A - a$ implies $\overline{T^\circ_A(a)} = T_A(a)$. □

Theorem

Theorem 3.1 (Normal–tangent polarity)**.** *For a convex set $A \subset X$ and $a \in A$,*

$$N_A(a) = (T_A(a))^{\ominus}.$$

Sketch. By definition $A - a \subset T_A(a)$, hence $(T_A(a))^{\ominus} \subset (A - a)^{\ominus} = N_A(a)$. Conversely, if $x \in N_A(a)$ then $\langle x, z \rangle \leq 0$ for all $z \in A - a$, and by closedness/convexity this extends to $z \in T_A(a)$, so $x \in (T_A(a))^{\ominus}$. □

3.4 Quick checklist for quizzes / problem sets

- **Epigraph test:** f convex $\Leftrightarrow$ $\operatorname{epi} f$ convex; strict convexity $\Rightarrow$ uniqueness of minimizers on convex sets.
- **Norm lore:** $\|\cdot\|$ convex, not strictly (in general); $\|\cdot\|^2$ strictly convex in inner-product spaces.
- **Distance and partial infimum:** $d(\cdot, \Omega)$ convex for convex Ω; if f convex then $x \mapsto \inf_y f(x, y)$ is convex.
- **Cones:** intersections/sums/closures preserve conicity; $C^{\ominus} = -C^{+}$; orthant polarity $(\mathbb{R}^n_+)^+ = \mathbb{R}^n_+$ and $(\mathbb{R}^n_+)^{\ominus} = \mathbb{R}^n_-$.
- **Normal & tangent:** $T^{\circ}_A(a)$ convex cone, $\overline{T^{\circ}_A(a)} = T_A(a)$, and $N_A(a) = (T_A(a))^{\ominus}$.

Chapter 4

Operations Preserving Convexity and Strong Convexity

4.1 Concavity and Convexity

Definition

Definition 4.1 (Concavity)**.** *A function $f : X \to \mathbb{R}$ is said to be* concave *if $-f$ is convex. It is* strictly concave *if $-f$ is strictly convex.*

Linear functions are both convex and concave. A maximization problem $\max f(x)$ is called *convex* when f is concave.

Proposition

Proposition 4.1 (Examples)**.**

- *$f(x) = x$ and $f(x) = \ln x$ are concave functions on $(0, \infty)$.*
- *$f(x) = x^2$ is convex and not concave.*

4.2 Level and Sublevel Sets

Definition

Definition 4.2 (Level and sublevel sets)**.** *Let $f : X \to \mathbb{R}$. For $\alpha \in \mathbb{R}$, the* sublevel set *and* level set *of f are defined by*

$$\mathrm{Lev}_{\leq\alpha}(f) := \{x \in X : \ f(x) \leq \alpha\}, \qquad \mathrm{Lev}_{=\alpha}(f) := \{x \in X : \ f(x) = \alpha\}.$$

Proposition

Proposition 4.2 (Convexity of sublevel sets)**.** *If $f : X \to \mathbb{R}$ is convex, then for every $\alpha \in \mathbb{R}$, the set $\mathrm{Lev}_{\leq\alpha}(f)$ is convex.*

Idea. Let $x_1, x_2 \in \mathrm{Lev}_{\leq\alpha}(f)$ and $\theta \in [0,1]$. Convexity gives

$$f(\theta x_1 + (1-\theta)x_2) \leq \theta f(x_1) + (1-\theta)f(x_2) \leq \alpha.$$

Hence the point lies in $\mathrm{Lev}_{\leq\alpha}(f)$. □

Definition

Definition 4.3 (Quasi-convex function). *A function $f : X \to \mathbb{R}$ is* quasi-convex *if all its sublevel sets are convex.*

Every convex function is quasi-convex, but the converse is false.

4.3 Operations Preserving Convexity

Proposition

Proposition 4.3 (Affine combinations and sums). *If $f_1, \ldots, f_m$ are convex functions and $\alpha_1, \ldots, \alpha_m \geq 0$, then*

$$f = \sum_{i=1}^{m} \alpha_i f_i$$

is convex. In particular, the sum of convex functions is convex.

Proposition

Proposition 4.4 (Affine composition). *If $f : X \to \mathbb{R}$ is convex and $A : Y \to X$ is linear (affine map $Ay + b$), then*

$$g(y) := f(Ay + b)$$

is convex on Y.

Example. If $f(x_1, x_2) = x_1^2 + x_2^2$ and $A(y_1, y_2) = (y_1 - 8y_2 - 3)$, then $g(y) = f(Ay + b) = 2e^{5y_1 - 8y_2 - 3}$ is convex as a composition of convex and affine maps. □

Proposition

Proposition 4.5 (Composition with monotone convex maps). *Let $f : X \to \mathbb{R}$ be convex and $g : \mathrm{range}(f) \to \mathbb{R}$ be convex and nondecreasing. Then $g \circ f$ is convex.*

Counterexample when monotonicity is missing. Take $f(x) = x^2$ (convex) and $g(t) = t^{1/4}$, which is not convex for $t > 0$. Then $g \circ f(x) = |x|^{1/2}$ is not convex near 0. □

Proposition

Proposition 4.6 (Pointwise supremum). *If $\{f_i : X \to \mathbb{R}\}_{i \in I}$ are convex, then the pointwise supremum*

$$f(x) := \sup_{i \in I} f_i(x)$$

is convex, with

$$\mathrm{epi}(f) = \bigcap_{i \in I} \mathrm{epi}(f_i).$$

Remark. The *hinge loss* $g(x) = \max(0, 1 - x)$, used in machine learning, is convex as the maximum of two convex functions. □

Proposition

Proposition 4.7 (Restriction to subspaces). *Let $f : X \times Y \to \mathbb{R}$ be convex, and fix $y \in Y$. Define $g_y(x) := f(x, y)$. Then g_y is convex on X.*

Counterexample for the converse. $f(x, y) = xy$ has all affine (hence convex) restrictions $g_y(x) = xy$, but f itself is not convex on $\mathbb{R}^2$. □

4.4 Indicator Functions and Convex Sets

Definition

Definition 4.4 (Indicator function). *Given $C \subset X$, define its* indicator function

$$I_C(x) = \begin{cases} 0, & x \in C, \\ +\infty, & x \notin C. \end{cases}$$

Theorem

Theorem 4.1 (Convexity criterion). *A set $C \subset X$ is convex $\iff$ its indicator function I_C is convex.*

Sketch. ($\Rightarrow$) If C is convex, take $x, y \in C$ and $\theta \in [0, 1]$; then $\theta x + (1 - \theta)y \in C$, hence $I_C(\theta x + (1 - \theta)y) = 0 \leq \theta I_C(x) + (1 - \theta)I_C(y)$. If either x or y is outside C, the RHS is $+\infty$, inequality holds trivially. ($\Leftarrow$) If I_C is convex and $x, y \in C$, then the convexity inequality forces $\theta x + (1 - \theta)y \in C$. □

4.5 First-Order Characterization of Convexity

Theorem

Theorem 4.2 (First-order condition). *Let $f : X \to \mathbb{R}$ be Fréchet differentiable on an open convex domain in a Hilbert space X. Then f is convex iff*

$$f(y) \;\geq\; f(x) + \langle \nabla f(x), y - x \rangle, \quad \forall x, y \in \mathrm{dom}\, f.$$

Geometrically, the graph of f lies above all its tangent hyperplanes. This is the foundation of first-order optimality conditions.

4.6 Strong Convexity

Definition

Definition 4.5 (Strong convexity)**.** *A differentiable function* $f : X \to \mathbb{R}$ *is* μ-strongly convex *if there exists* $\mu > 0$ *such that*

$$f(y) \;\geq\; f(x) + \langle \nabla f(x), y - x\rangle + \frac{\mu}{2}\|y - x\|^2, \quad \forall x, y \in \operatorname{dom} f.$$

Theorem

Theorem 4.3 (Equivalent characterizations)**.** *Let* f *be differentiable and* $\mu > 0$. *The following are equivalent:*

(i) f *satisfies the strong convexity inequality above.*

(ii) $g(x) := f(x) - \frac{\mu}{2}\|x\|^2$ *is convex.*

(iii) $\langle \nabla f(x) - \nabla f(y),\, x - y\rangle \geq \mu\|x - y\|^2$ *for all* x, y.

(iv) $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y) - \alpha(1-\alpha)\frac{\mu}{2}\|x - y\|^2$.

The constant μ measures the *curvature* or *strongness* of convexity. For $\mu = 0$, we recover standard convexity.

Proposition

Proposition 4.8 (Gradient monotonicity)**.** *For a* μ*-strongly convex differentiable function,*

$$\langle \nabla f(x) - \nabla f(y),\, x - y\rangle \geq \mu\|x - y\|^2,$$

which means the gradient mapping is strongly monotone.

4.7 Midpoint Convexity

Definition

Definition 4.6 (Midpoint convexity)**.** *A set* C *is* midpoint convex *if* $x, y \in C \implies \frac{x+y}{2} \in C$.

Proposition

Proposition 4.9. *Every convex set is midpoint convex, but the converse need not hold.*

Proposition

Proposition 4.10 (Counterexample)**.** *The set $\mathbb{Q} \subset \mathbb{R}$ of rational numbers is midpoint convex but not convex, since for $x = 0$, $y = 1$, we have $(x+y)/2 = 1/2 \in \mathbb{Q}$, but irrational points between 0 and $\sqrt{2}$ are missing.*

Theorem

Theorem 4.4 (Closed midpoint convex sets are convex)**.** *If C is closed and midpoint convex, then C is convex.*

Sketch. Given $x, y \in C$ and $\alpha \in [0, 1]$, repeatedly take midpoints between x and y to approximate $\alpha x + (1 - \alpha)y$. All these points stay in C by midpoint convexity. Closedness ensures that the limit point also lies in C. □

4.8 Summary for Quizzes and Problem Sets

- f convex $\Rightarrow$ sublevel sets convex $\Rightarrow$ quasi-convex.
- Operations preserving convexity: affine maps, nonnegative sums, positive scalar multiples, pointwise supremum, and monotone convex compositions.
- Indicator I_C convex $\iff$ C convex.
- Differentiable convex $\iff$ $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$.
- Strong convexity $\iff$ quadratic lower bound; equivalent to convexity of $f - \frac{\mu}{2}\|x\|^2$.
- Closed midpoint convex $\Rightarrow$ convex (important geometric result).

Chapter 5

Continuity, Lower Semicontinuity, and the Role of Convexity in Optimization

5.1 Continuity of Convex Functions

Proposition

Proposition 5.1 (Continuity criterion)**.** *Let $f : X \to \mathbb{R}$ be convex, and $a \in X$ such that f is upper-bounded in a neighborhood of a. Then f is continuous at a.*

Theorem

Theorem 5.1 (Continuity on the interior of the domain)**.** *Let $f : X \to \mathbb{R}$ be convex on a finite-dimensional normed space. Then f is continuous at every $x_0 \in (\mathrm{dom} f)^\circ$ where $f(x_0) \in \mathbb{R}$.*

Idea. Choose the max norm $\|x\|_\infty = \max_i |x_i|$ in $\mathbb{R}^n$. Let $x_0 \in (\mathrm{dom} f)^\circ$ and pick $\omega > 0$ such that $B_\omega(x_0) \subset \mathrm{dom} f$. Define C as the convex hull of the 2^n vertices of the cube centered at x_0 with edge 2ω. By convexity, f is upper-bounded on C, and Proposition 1 implies continuity at x_0. □

Hence convex functions are automatically continuous in the interior of their effective domain, though they may diverge to $+\infty$ at the boundary.

5.2 Lower Semicontinuity and Closed Epigraphs

Definition

Definition 5.1 (Lower semicontinuity). *A function $f : X \to \mathbb{R}$ is* lower semicontinuous (l.s.c.) *at x if for every sequence $x_n \to x$,*

$$f(x) \le \liminf_{n\to\infty} f(x_n).$$

Equivalently, f does not jump downward at x.

Theorem

Theorem 5.2 (Equivalent characterizations). *For any function $f : X \to \mathbb{R}$, the following are equivalent:*

(i) f is lower semicontinuous,

(ii) $\operatorname{epi} f$ is closed in $X \times \mathbb{R}$,

(iii) for every $\gamma \in \mathbb{R}$, the sublevel set $\{x : f(x) \le \gamma\}$ is closed.

Sketch. *(i)* $\Rightarrow$ *(ii)*: if $(x_n, \mu_n) \in \operatorname{epi}(f)$ and $(x_n, \mu_n) \to (x, \mu)$, then $f(x_n) \le \mu_n$ and by l.s.c. $f(x) \le \liminf f(x_n) \le \lim \mu_n = \mu$. *(ii)* $\Rightarrow$ *(iii)*: closedness of $\operatorname{epi}(f)$ implies closedness of each horizontal slice at height γ. *(iii)* $\Rightarrow$ *(i)*: if $x_n \to x$ and $f(x_n) \to y$, then x lies in every closed sublevel $\{f \le y + \varepsilon\}$, hence $f(x) \le y$. ☐

5.3 Directional Derivatives and Differentiability

Definition

Definition 5.2 (Directional derivative). *Let $f : (X, \|\cdot\|) \to \mathbb{R}$ and $x, d \in X$. The* directional derivative of f at x along d *is*

$$f'(x; d) = \lim_{t\downarrow 0} \frac{f(x + td) - f(x)}{t},$$

if the limit exists.

Proposition

Proposition 5.2 (Existence for convex functions). *If f is convex, then for every $x, d \in X$ the right-hand limit defining $f'(x; d)$ exists (possibly $+\infty$). Moreover, $d \mapsto f'(x; d)$ is convex.*

Idea. For convex f, the quotient $\frac{f(x+td)-f(x)}{t}$ is nondecreasing in $t > 0$; hence its limit as $t \downarrow 0$ exists. The convexity of $f'(x; \cdot)$ follows from convexity of f in the direction variable. ☐

Definition

Definition 5.3 (Gateaux and Fréchet differentiability)**.**
- *f is* Gateaux differentiable *at x if $f'(x; d)$ exists for all d and the map $d \mapsto f'(x; d)$ is linear and continuous. We then denote $df(x)(d) = f'(x; d)$.*
- *f is* Fréchet differentiable *at x if there exists $Df(x) \in \mathcal{L}(X, \mathbb{R})$ such that*

$$\lim_{h \to 0} \frac{|f(x+h) - f(x) - Df(x)(h)|}{\|h\|} = 0.$$

Fréchet differentiability $\Rightarrow$ Gateaux differentiability.

Proposition

Proposition 5.3 (Example of Gateaux but not Fréchet)**.** *Define*

$$f(x, y) = \begin{cases} 1, & y = x^2,\ x \neq 0, \\ 0, & \textit{otherwise}. \end{cases}$$

Then f is Gateaux differentiable at $(0, 0)$ but not Fréchet differentiable.

5.4 Differentiability and Convexity

Theorem

Theorem 5.3 (First-order condition for convexity)**.** *Let $f : X \to \mathbb{R}$ be Fréchet differentiable on an open convex domain. Then f is convex $\iff$*

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle, \quad \forall x, y.$$

Theorem

Theorem 5.4 (Gradient monotonicity criterion)**.** *A Fréchet differentiable function f is convex $\iff$*

$$\langle \nabla f(y) - \nabla f(x),\ y - x \rangle \geq 0, \quad \forall x, y \in X.$$

Idea. Add the inequalities $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$ and $f(x) \geq f(y) + \langle \nabla f(y), x - y \rangle$ to deduce $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$. Conversely, integrating this inequality along the line segment between x and y recovers convexity. □

Proposition

Proposition 5.4 (Quadratic form criterion)**.** *Let $A \in \mathbb{R}^{n \times n}$ be symmetric and define $f(x) = \langle Ax, x \rangle$. Then $\nabla f(x) = 2Ax$, and f is convex $\iff$ A is positive semidefinite.*

5.5 Optimization and the Role of Convexity

Definition

Definition 5.4 (Local and global minimizers)**.** *For* $f : X \to \mathbb{R}$*:*

- $x^\star$ *is a* global minimizer *if* $f(x^\star) \leq f(x)$ *for all* $x \in X$*.*
- $x^\star$ *is a* local minimizer *if* $\exists \varepsilon > 0$ *such that* $f(x^\star) \leq f(x)$ *for all* $x \in B(x^\star, \varepsilon)$*.*

Theorem

Theorem 5.5 (Convexity and optimality)**.** *If* f *is convex, every local minimizer is a global minimizer. If* f *is strictly convex and* $\mathrm{dom} f \neq \emptyset$*, then the global minimizer—if it exists—is unique.*

Sketch. For convex f, suppose $x^\star$ is local. For any y and $t \in (0,1)$ small enough, $x_t = (1-t)x^\star + ty$ remains in the local neighborhood, and

$$f(x^\star) \leq f(x_t) \leq (1-t)f(x^\star) + tf(y) \;\Rightarrow\; f(x^\star) \leq f(y).$$

Uniqueness for strictly convex f follows since equality in the convexity inequality implies $x = y$. □

Proposition

Proposition 5.5 (Examples)**.**

- $f(x) = x^2$ *is convex and has a unique minimizer* $x = 0$*.*
- $f(x) = 0$ *for* $x \in [0,2]$ *and* $f(x) = x^2$ *otherwise is convex, with infinitely many global minimizers in* $[0,2]$*.*

5.6 Summary for Quizzes and Problem Sets

- Convex $\Rightarrow$ continuous on $(\mathrm{dom} f)^\circ$.
- Lower semicontinuous $\iff$ closed epigraph $\iff$ closed sublevel sets.
- Convex functions admit directional derivatives; $d \mapsto f'(x; d)$ is convex.
- Fréchet differentiability $\Rightarrow$ convexity via tangent-plane inequality.
- Gradient monotonicity: $\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0$.
- Quadratic form convex $\iff$ symmetric matrix $A \succeq 0$.
- Local minima of convex f are global; strict convexity $\Rightarrow$ uniqueness.

Chapter 6

Revision: Problem-Solving Tools for Convex Analysis and Optimization

6.1 Core Definitions and Quick Tests

Definition

Definition 6.1 (Convex set test)**.** *A set C is convex $\iff$*

$$x, y \in C,\ t \in [0,1] \ \Rightarrow\ tx + (1-t)y \in C.$$

Equivalent formulations:

- $C = \operatorname{conv}(C)$ *(closed under convex combinations);*
- C *midpoint convex and closed;*
- I_C *(indicator function) is convex.*

Definition

Definition 6.2 (Convex function test)**.** *For $f : X \to \mathbb{R} \cup \{+\infty\}$:*

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y), \quad \forall x, y,\ t \in [0,1].$$

Equivalent characterizations:

- $\operatorname{epi}(f)$ *convex.*
- f *satisfies the first-order inequality $f(y) \ge f(x) + \langle \nabla f(x), y - x\rangle$.*
- $\langle \nabla f(y) - \nabla f(x), y - x\rangle \ge 0.$

Definition

Definition 6.3 (Strong convexity test)**.** f *is μ-strongly convex $\iff$ any of the following* *hold:*

$$
\begin{aligned}
f(y) &\geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2}\|y - x\|^2, \\
\langle \nabla f(y) - \nabla f(x), y - x \rangle &\geq \mu \|y - x\|^2, \\
f(x) - \frac{\mu}{2}\|x\|^2 &\text{ is convex.}
\end{aligned}
$$

Definition

Definition 6.4 (Lower semicontinuity tests). *f is lower semicontinuous (l.s.c.) $\iff$*

$$f(x) \leq \liminf_{n\to\infty} f(x_n) \text{ whenever } x_n \to x.$$

Equivalent to:

- $\mathrm{epi}(f)$ *is closed;*
- $\mathrm{Lev}_{\leq\alpha}(f)$ *is closed* $\forall\alpha$.

6.2 Geometric Tools and Cones

Definition

Definition 6.5 (Cones). *C is a cone $\iff$ $\lambda x \in C$ for all $\lambda \geq 0$, $x \in C$.*

Proposition

Proposition 6.1 (Operations). *Intersections, sums, closures $\Rightarrow$ cones. If C_1, C_2 are convex cones, so are $C_1 + C_2$, $C_1 \cap C_2$, $\overline{C_1}$.*

Definition

Definition 6.6 (Polar and positive cones).

$$C^+ = \{x : \langle x, y \rangle \geq 0, \ \forall y \in C\}, \qquad C^\ominus = \{x : \langle x, y \rangle \leq 0, \ \forall y \in C\}.$$

Definition

Definition 6.7 (Tangent and normal cones to convex sets). *For $A \subset X$ convex and $a \in A$:*

$$
\begin{aligned}
T_A(a) &= \overline{\{\alpha(x - a) : \alpha \geq 0, \ x \in A\}}, \\
N_A(a) &= \{z : \langle z, y - a \rangle \leq 0, \ \forall y \in A\} = (T_A(a))^\ominus.
\end{aligned}
$$

Quick polarity table:

Object	$\longleftrightarrow$	Polar relation
Convex cone C	$\leftrightarrow$	$C^{\ominus}$
$T_A(a)$	$\leftrightarrow$	$N_A(a)$
$\mathbb{R}^n_+$	$\leftrightarrow$	$\mathbb{R}^n_-$
Subspace M	$\leftrightarrow$	$M^\perp$

6.3 Projection and Riesz Tools

Theorem

Theorem 6.1 (Projection theorem). *If A is closed and convex in a Hilbert space, then for all x there exists a unique*

$$P_A(x) = \arg\min_{y\in A} \|x-y\|,$$

characterized by $\langle x - P_A(x), z - P_A(x)\rangle \le 0$ for all $z \in A$.

Proposition

Proposition 6.2 (Projection properties).

- *P_A is 1-Lipschitz: $\|P_A(x) - P_A(y)\| \le \|x-y\|$.*
- *If $A = M$ is a closed subspace: P_M is linear, $X = M \oplus M^\perp$.*
- *Orthogonality: $\langle x - P_M(x), v\rangle = 0$ for all $v \in M$.*

Theorem

Theorem 6.2 (Riesz representation). *For every continuous linear $\varphi : X \to \mathbb{R}$ on a Hilbert space, there exists a unique $w \in X$ such that*

$$\varphi(x) = \langle w, x\rangle, \quad \forall x \in X.$$

Theorem

Theorem 6.3 (Adjoint operator). *For $L \in \mathcal{L}(X,Y)$ between Hilbert spaces, there exists a unique $L^* \in \mathcal{L}(Y,X)$ such that*

$$\langle Lx, y\rangle_Y = \langle x, L^*y\rangle_X.$$

6.4 Functional Tools and Continuity

- Convex $\Rightarrow$ continuous on $\mathrm{int}(\mathrm{dom} f)$.
- L.s.c. $\iff$ closed epigraph $\iff$ closed sublevel sets.

- Directional derivative $f'(x; d)$ exists for convex f (as a finite or infinite limit).
- f Gateaux differentiable $\Rightarrow$ directional derivative linear in d.
- f Fréchet differentiable $\Rightarrow f(y) = f(x) + \langle \nabla f(x), y - x \rangle + o(\|y - x\|)$.

Theorem

Theorem 6.4 (Gradient monotonicity)**.** *For convex differentiable f:*

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq 0.$$

For μ-strongly convex f:

$$\langle \nabla f(y) - \nabla f(x), y - x \rangle \geq \mu \|y - x\|^2.$$

Proposition

Proposition 6.3 (Quadratic form test)**.** *If $f(x) = x^\top Ax$ with $A = A^\top$, then $\nabla f(x) = 2Ax$. f convex $\iff A \succeq 0$ (positive semidefinite).*

6.5 Optimization Patterns and Midterm Problem Strategies

- **Checking convexity:**
  1. Verify convexity of $\mathrm{epi}(f)$ or use the first-order inequality.
  2. For quadratic forms: check $A \succeq 0$ via eigenvalues $\lambda_i \geq 0$.
  3. For norms or distance functions: use triangle inequality.
- **Continuity + convexity combo:**
  - To prove continuity at x_0: find a local upper bound $\Rightarrow$ convex $\Rightarrow$ continuous.
  - For finite-dimensional X: convex $\Rightarrow$ continuous on $\mathrm{int}(\mathrm{dom} f)$.
- **Optimization questions:**
  - For convex f, local minima $\Rightarrow$ global.
  - For strictly convex f, minimizer (if exists) is unique.
  - First-order optimality condition:

$$\nabla f(x^\star) = 0 \quad \text{if } f \text{ differentiable and convex.}$$

- **Strong convexity:**
  - Add $\frac{\mu}{2}\|x\|^2$ to make f strictly convex.
  - Gives quadratic lower bound on growth around the minimizer.
- **Geometric intuition:**
  - Convex sets $\Rightarrow$ one supporting hyperplane per boundary point.
  - Tangent and normal cones: orthogonal geometry governs constraint qualification.

6.6 Quick Formula Library

- **Convex combination:** $\sum_i \theta_i x_i$, with $\theta_i \geq 0$, $\sum_i \theta_i = 1$.
- **Convex hull:** $\operatorname{conv}(A) = \{\text{finite convex combinations of points in } A\}$.
- **Distance:** $d(x, A) = \inf_{y\in A} \|x - y\|$ (convex if A convex).
- **Projection:** $P_A(x) = \arg\min_{y\in A} \|x - y\|$.
- **Sublevel set:** $\{f \leq \alpha\}$ convex $\iff$ f convex.
- **Epigraph:** $\operatorname{epi}(f) = \{(x, u) : f(x) \leq u\}$.
- **Strong convexity gradient test:** $\langle \nabla f(x) - \nabla f(y), x - y\rangle \geq \mu\|x - y\|^2$.
- **Monotone gradient:** ∇f monotone $\iff$ f convex.

6.7 Checklist for the Midterm

1. Identify if the object is a **set** or a **function**.
2. For sets: test convexity via combinations, intersections, closure, or support lines.
3. For functions: use epigraph, first-order test, or matrix positivity.
4. When proving continuity: look for boundedness around the point.
5. Always mention domain and properness (avoid $f \equiv +\infty$).
6. If f is convex and differentiable, minimize by solving $\nabla f(x) = 0$.
7. For geometric problems, recall:
$$N_A(a) = (T_A(a))^{\ominus}, \quad P_A(x) : \ \langle x - P_A(x), z - P_A(x)\rangle \leq 0.$$
8. Remember: convex + closed $\Rightarrow$ well-behaved projections and continuity.