2 Review of convexity

2.1 Convex sets

Definition 2.1. A set $C \subset \mathbb{R}^n$ is convex if for any $x, y \in C$ and $\lambda \in [0,1]$, $\lambda x + (1-\lambda)y \in C$.

Proposition 2.1. *If C is a convex set, and $A : \mathbb{R}^n \to \mathbb{R}^m$ is a linear map, then $A(C)$ is convex. If $(C_j)_{j\in J}$ is a collection of convex sets, then $C = \cap_{j\in J} C_j$ is convex.*

Examples:
A *halfspace* is a set of the form $\{x \in \mathbb{R}^n : \langle a, x\rangle \leq b\}$ where $a \neq 0$,

$$\langle a, x\rangle = a^T x = \sum_{i=1}^n a_i x_i$$

is the Euclidean inner product.

An intersection of (any number of) halfspaces is a convex set. If C is an intersection of a *finite* number of halfspaces it is called a *convex polyhedron.*
It turns out that any *closed* convex set can be written as an intersection of halfspaces! This can be proved using the following fundamental fact about convex sets.

Theorem 2.1 (Separating hyperplane theorem)**.** *Let $C \subset \mathbb{R}^n$ be a convex set, and let $y \notin C$. Then there is $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}$ such that*

$$\langle a, x\rangle \leq b \; \forall x \in C \quad \text{and} \quad \langle a, y\rangle \geq b.$$

Proof. We give the proof when C is closed. The general case is left as an exercise. If C is closed we can define the projection map on C, namely $p_C(y) := \min\{\|y - x\|_2 : x \in C\}$ is well defined and satisfies $\langle y - p_C(y), x - p_C(y)\rangle \leq 0$ for any $x \in C$. Let $a = y - p_C(y)$ and $b = \langle a, p_C(y)\rangle + \frac{1}{2}\|a\|_2^2$. Note that $\langle a, y\rangle - b = \|a\|_2^2 - \frac{1}{2}\|a\|_2^2 > 0$. Also for any $x \in C$ we have $\langle a, x\rangle - b = \langle y - p_C(y), x - p_C(y)\rangle - \frac{1}{2}\|a\|_2^2 < 0$ which is what we wanted. □

EXERCISE: Use theorem above to prove that if C is a closed convex set, then C is equal to the intersection of halfspaces that contain it.

Supporting hyperplane: The result above can be used to prove the existence of *supporting hyperplanes.* If C is a closed convex set, and $y \in C \setminus \textbf{int}\, C$, then there is a hyperplane that supports C at y, i.e., there is $a \in \mathbb{R}^n \setminus \{0\}$ and $b \in \mathbb{R}^n$ such that $\langle a, y\rangle = b$ and $\langle a, x\rangle \leq b$ for all $x \in C$.

2.2 Convex functions

Let $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

Definition 2.2. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if for any $x, y \in \mathbb{R}^n$ and $\lambda \in [0,1]$

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

The domain of f is the set of x where $f(x)$ is finite: $\mathbf{dom}(f) = \{x \in \mathbb{R}^n : f(x) < \infty\}$. Note that $\mathbf{dom}(f)$ is a convex set by (1).

The indicator function of a convex set $C \subset \mathbb{R}^n$ is

$$I_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{else.} \end{cases}$$

If $f : C \to \mathbb{R}$ is a convex function defined on a convex set C, then we can always think of f as an (extended-valued) convex function $\mathbb{R}^n \to \overline{\mathbb{R}}$, by defining f to be $+\infty$ outside of C.

The *epigraph* of a convex function f is defined as

$$\mathbf{epi}(f) = \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq t\}.$$

It is easy to see that f is convex if, and only if, $\mathbf{epi}(f)$ is a convex set.

Observe that $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if, and only if, its restriction to any one-dimensional line is convex, i.e., for any $x \in \mathbf{dom}(f)$ and any $h \in \mathbb{R}^n$, the function $t \mapsto f(x+th)$ is convex on the interval $\{t \in \mathbb{R} : x+th \in \mathbf{dom}(f)\}$. This reduces the problem of proving convexity of a multivariate function, to that of univariate functions.

EXERCISE: Use this to show that the function $X \in \mathbf{S}^n_{++} \mapsto -\log\det X$ is convex on the set $\mathbf{S}^n_{++}$ of positive definite matrices.

Recall that a univariate differentiable function $f : \mathbb{R} \to \mathbb{R}$ is convex if, and only if, f' is nondecreasing. When f is twice differentiable this is equivalent to $f'' \geq 0$.

Operations that preserve convexity The following elementary facts are very useful to prove convexity.

1. Linear transformations: If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex and $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^m$, then $g(y) = f(Ax+b)$ defined on $\mathbb{R}^m$ is convex.

2. Pointwise supremum: If $(f_j)_{j\in J}$ is a collection of convex functions defined on $\mathbb{R}^n$, then $f(x) = \sup_{j\in J} f_j(x)$ is convex.

3. Perspective: If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex, then $P_f : (x,t) \in \mathbb{R}^n \times \mathbb{R}_{++} \to \overline{\mathbb{R}}$ defined by $P_f(x,t) = tf(x/t)$ is convex. (P_f is known as the perspective of f.)

4. Partial minimization: If $g : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$ is convex, then $f(x) = \inf_y g(x,y)$ is convex on $\mathbb{R}^n$.

EXERCISE: Show that $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(x) =$ sum of k largest components of x, is a convex function.
EXERCISE: Show that $f : \mathbf{S}^n \to \mathbb{R}$, $f(X) = \lambda_{\max}(X)$ (largest eigenvalue of the real symmetric matrix X) is convex.
EXERCISE: Let $f : \mathbb{R} \to \mathbb{R}$ be a function defined on the real line. For $a \neq b$, let $f[a,b] = \frac{f(b)-f(a)}{b-a}$ be the first order finite difference, and for a,b,c distinct, let $f[a,b,c] = (f[a,b] - f[b,c])/(a-c)$ be the second order finite differences. Show that f is convex if, and only if, $f[a,b,c] \geq 0$ for all $a,b,c \in \mathbf{dom}\, f$.

Differentiable functions A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is *differentiable* at $x \in \mathbf{int\,dom}(f)$ if there is a vector, denoted $\nabla f(x)$ and called the *gradient* of f at x, s.t.

$$f(x+h) = f(x) + \langle \nabla f(x), h\rangle + o(\|h\|), \qquad (h \to 0). \tag{2}$$

Because we are using the Euclidean inner product $\langle u, v\rangle = \sum_{i=1}^n u_i v_i$, we have

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{bmatrix}.$$

- If f is convex and differentiable at x, then

$$f(y) \geq f(x) + \langle \nabla f(x), y - x\rangle \qquad \forall y \in \mathbb{R}^n. \tag{3}$$

  Note that this is a *global* inequality, holding for all $y \in \mathbb{R}^n$.

- As such, if $\nabla f(x) = 0$, then $f(y) \geq f(x)$ for all $y \in \mathbb{R}^n$, and x is a global minimum of f.

- If $\mathbf{dom}(f)$ is open and f is differentiable everywhere on its domain, then inequality (3) is a sufficient condition for convexity.

Second derivatives If a function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is twice continuously differentiable around x then

$$f(x + h) = f(x) + \langle \nabla f(x), h\rangle + \frac{1}{2}\left\langle h, \nabla^2 f(x) h\right\rangle + o(\|h\|^2), \qquad (h \to 0).$$

for some $n \times n$ symmetric matrix $\nabla^2 f(x)$, called the *Hessian* of f at x. Note that

$$[\nabla^2 f(x)]_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

Recall that a $n \times n$ symmetric matrix A is *positive semidefinite* if $\langle x, Ax\rangle \geq 0$ for all $x \in \mathbb{R}^n$; equivalently, if all the eigenvalues of A are nonnegative. A matrix is *positive definite* if $\langle x, Ax\rangle > 0$ for all $x \in \mathbb{R}^n \setminus \{0\}$, or equivalently if all the eigenvalues of A are positive.

- If f is convex, and twice continuously differentiable at x, then $\nabla^2 f(x)$ is positive semidefinite, we write $\nabla^2 f(x) \succeq 0$.

- Conversely, if $\mathbf{dom}(f)$ is open, f is twice continuously differentiable everywhere in its domain with $\nabla^2 f(x) \succeq 0$, then f is convex.