Convexity Sets and Functions

Andrew Nobel

April, 2023

Convex Sets

Convex Sets

Definition: A set $C \subseteq \mathbb{R}^d$ is *convex* if for every $x, y \in C$ and every $\alpha \in [0, 1]$ the point $\alpha x + (1 - \alpha)y \in C$.

Interpretation

- Vector $\alpha x + (1 - \alpha)y$ called *convex combination* of x, y with weight α
- Set $\{\alpha x + (1 - \alpha)y : \alpha \in [0, 1]\}$ is just the line between x and y
- So C is convex if the line between any two points in C is contained in C

Examples of Convex Sets

- Simple examples: $C = \mathbb{R}^d$, $\emptyset$, $\{0\}$
- Open norm ball $B(x_0, r) := \{x : ||x - x_0|| < r\}$
- Halfspace $H(w, b) = \{x : w^t x \geq b\}$ with direction w and offset b
- Hyperplane $\partial H(w, b) = \{x : w^t x = b\}$ (n-1)-dimensional
- Polyhedron $\{x : Ax \leq c\}$ where $\leq$ understood componentwise
- Probability simplex $\{u : u_i \geq 0$ and $\sum_{i=1}^d u_i = 1\}$

Basic Properties of Convex Sets

Fact

1. If C_λ with $\lambda \in \Lambda$ are convex sets then so is their intersection $\cap_{\lambda \in \Lambda} C_\lambda$
2. If $A, B \subseteq \mathbb{R}^d$ are convex then so is $A + B := \{u + v : u \in A \text{ and } v \in B\}$
3. If $A \subseteq \mathbb{R}^d$ is convex and $c \in \mathbb{R}$ then $cA := \{cu : u \in A\}$ is convex

Convex Functions

Convex Functions

Definition: Let $C \subseteq \mathbb{R}^d$ be convex. A function $f : C \to \mathbb{R}$ is *convex* if for every $x, y \in C$ and every $\alpha \in (0, 1)$,

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

Convexity of C ensures that $f(\cdot)$ is defined at $\alpha x + (1-\alpha)y$

Interpretation: For each $x, y \in C$ the line connecting $(x, f(x))$ and $(y, f(y))$ lies *above* the graph $\{(u, f(u)) : u \in C\} \subseteq \mathbb{R}^{d+1}$ of f

Related Definitions

- $f : C \to \mathbb{R}$ is *strictly convex* if $(*)$ holds with $\leq$ replaced by $<$
- $f : C \to \mathbb{R}$ is *concave* if $(*)$ holds with $\leq$ replaced by $\geq$
- $f : C \to \mathbb{R}$ is *strictly concave* if $(*)$ holds with $\leq$ replaced by $>$

Verifying Convexity and Concavity

1. Check the definition: In many cases it is possible to directly check the definition

2. Second derivative condition: Let $C \subseteq \mathbb{R}^d$ be convex

- A function $f : C \to \mathbb{R}$ is convex if the matrix $\nabla^2 f(x)$ of second partial derivatives is well-defined and non-negative definite for each $x \in C$
- A function $f : C \to \mathbb{R}$ is concave if the matrix $\nabla^2 f(x)$ of second partial derivatives is well-defined and non-positive definite for each $x \in C$

Special case: If $d = 1$ then f is convex if $f'' \geq 0$ and concave if $f'' \leq 0$

Examples of Convex/Concave Functions

Case $d = 1$

- $f(x) = |x|$ is convex, but *not* strictly convex
- $f(x) = x^2$, e^x, e^{-x} are strictly convex on $\mathbb{R}$
- x^{-1}, and $x \log x$ are strictly convex on $[0, \infty)$
- $f(x) = \log x$, $\sqrt{x}$ are strictly concave on $(0, \infty)$

Case $d \geq 2$

- $f(x) = ||x||$ is convex
- $f(x) = \langle x, u \rangle + b$, affine function, is convex and concave
- $f(x) = \sup_{u \in A} \langle x, u \rangle$, where $A \subseteq \mathbb{R}^d$ is bounded, is convex
- $f(x) = x^t A x$ is convex if $A \geq 0$, concave if $A \leq 0$

Basic Properties of Convex Functions

Fact: Let $C \subseteq \mathbb{R}^d$ be convex

(a) $f : C \to \mathbb{R}$ is convex if $-f$ is concave, and vice-versa

(b) If $f : C \to \mathbb{R}$ is convex or concave, it is continuous on the interior of C

(c) If $\{f_\lambda : \lambda \in \Lambda\}$ are convex functions on C then so is $f(x) = \sup_{\lambda \in \Lambda} f_\lambda(x)$

(d) If $f_1, \ldots, f_m$ are convex and $\alpha_1, \ldots, \alpha_m \geq 0$ then $f = \sum_{i=1}^m \alpha_i f_i$ is convex

(e) If $f : C \to \mathbb{R}$ is convex and $g : \mathbb{R} \to \mathbb{R}$ is convex and non-dec, then $g \circ f$ is convex

Subgradients and Jensen’s Inequality

Subgradients of Convex Functions

Fact: Let $C \subseteq \mathbb{R}^d$ be convex. If $f : C \to \mathbb{R}$ is convex, then for every $u \in C^o$ there is a vector $v \in \mathbb{R}^d$ such that

$$f(x) \;\geq\; f(u) + \langle v, x - u \rangle \;\text{ for each }\; x \in C$$

The vector v is called a *subgradient* of f at u. The set of all subgradients of f at u is denoted by $\partial f(u)$

Note

- Lower bound $h(x) := f(u) + \langle v, x - u \rangle$ is affine with $h(u) = f(u)$
- Graph of function h is a hyperplane supporting the graph of f at u
- Function f can be expressed as the maximum of its affine lower bounds

Properties of the Subgradient

Fact: Let $f, g : C \to \mathbb{R}$ be convex. For each $u \in C$

- $\partial f(u)$ is a closed convex set, and is non-empty if $u \in C^o$
- $\partial f(u) = \{v\}$ is a singleton iff f is differentiable at u
- If $\partial f(u) = \{v\}$ then $v = \nabla f(u)$ is the gradient of f at u
- $\partial(\alpha f)(u) = \alpha \, \partial f(u)$
- $\partial(f + g)(u) = \partial f(u) + \partial g(u)$
- $\partial f(Ax + b)(u) = A^t \, \partial f(Au + b)$

Jensen's Inequality

Recall: The expected value of a random vector $X = (X_1, \ldots, X_d)^t$ is defined by

$$\mathbb{E}X \;=\; (\mathbb{E}X_1, \ldots, \mathbb{E}X_d)^t \in \mathbb{R}^d$$

Jensen's Inequality: Let $C \subseteq \mathbb{R}^d$ be convex and suppose that $X \in C$. Provided that all expectations are well-defined, the following hold.

(1) The expectation $\mathbb{E}X \in C$

(2) If $f : C \to \mathbb{R}$ is convex then $f(\mathbb{E}X) \leq \mathbb{E}f(X)$. If f is strictly convex and X is not constant then the inequality is strict.

(3) If $f : C \to \mathbb{R}$ is concave then $f(\mathbb{E}X) \geq \mathbb{E}f(X)$. If f is strictly concave and X is not constant then the inequality is strict.

Note: Definition of convexity is a special case of (2) for a random vector $X \in C$ with $\mathbb{P}(X = x) = \alpha$ and $\mathbb{P}(X = y) = 1 - \alpha$

Applications of Jensen's Inequality

Case $d = 1$

- $\mathbb{E}X^2 \geq (\mathbb{E}X)^2 \quad \mathbb{E}e^X \geq e^{\mathbb{E}X} \quad \mathbb{E}(X \log X) \geq (\mathbb{E}X) \log(\mathbb{E}X)$
- $\mathbb{E} \log X \leq \log \mathbb{E}X \quad \mathbb{E}\sqrt{X} \leq \sqrt{\mathbb{E}X}$

Case $d \geq 2$

- $\mathbb{E}||X|| \geq ||\mathbb{E}X||$
- $\mathbb{E}(X^t AX) \leq (\mathbb{E}X)^t A(\mathbb{E}X)$ if $A \leq 0$

AM-GM inequality: If $a_1, \ldots, a_n$ are positive then $(\prod_{i=1}^n a_i)^{1/n} \leq n^{-1} \sum_{i=1}^n a_i$

Holder's Inequality

Fact: Let $a, b \geq 0$ and $1 < p, q < \infty$ be such that $1/p + 1/q = 1$. Then

$$\frac{1}{p}\, a^p \;+\; \frac{1}{q}\, b^q \;\geq\; ab$$

Holder's Inequality: Let $1 < p, q < \infty$ be such that $1/p + 1/q = 1$. If X, Y are random variables such that $\mathbb{E}|X|^p, \mathbb{E}|Y|^q$ are finite then

$$|\mathbb{E}XY| \;\leq\; \mathbb{E}|XY| \;\leq\; (\mathbb{E}|X|^p)^{1/p}\, (\mathbb{E}|Y|^q)^{1/q}$$

Cauchy-Schwartz: If $\mathbb{E}X^2$, $\mathbb{E}Y^2$ finite then $\mathbb{E}|XY| \leq \sqrt{\mathbb{E}X^2\, \mathbb{E}Y^2}$

General version: If $p, q \geq 0$ satisfy $1/p + 1/q = 1$ and $f, g, h : \mathcal{X} \to \mathbb{R}$ with $h \geq 0$ then

$$\int |f(x)g(x)|h(x)\, dx \;\leq\; \left(\int |f(x)|^p h(x)\, dx\right)^{1/p} \left(\int |g(x)|^q h(x)\, dx\right)^{1/q}$$

Convexity and Optimization

General Optimization Problem

Problem: Minimize a function $f : \mathbb{R}^d \to \mathbb{R}$ over a set $A \subseteq \mathbb{R}^d$ of interest. Often expressed in the form of a *mathematical program*:

$$\min f(x) \text{ subject to } x \in A$$

- Function f called *objective function*
- Set A represents *constraints* on the arguments x of interest
- Points $x \in A$ called *feasible*
- Usually interested in $\min_A f(x)$ and $\operatorname{argmin}_A f(x)$

General Optimization Problem, cont.

Global and local minima

- Feasible $x \in A$ is a *global minimum* of f if $f(x) \leq f(y)$ for all $y \in A$
- Feasible $x \in A$ is a *local minimum* of f if there exists an $r > 0$ such that $f(x) \leq f(y)$ for all $y \in A$ with $||x - y|| \leq r$

Notes: A global minimum is a local minimum. Other issues

- Is there a global min? Is it unique?
- Is there a closed form solution for the global min?
- Are there good iterative or approximate solutions?
- Does f have many local minima?

Convexity and Optimization

Fact: If $C \subseteq \mathbb{R}^d$ is convex and $f : C \to \mathbb{R}$ is convex then

1. Any local minimum is a global minimum
2. If f is strictly convex any global minimum is unique

In general: If $C \subseteq \mathbb{R}^d$ and $f : C \to \mathbb{R}$ are convex then there are efficient iterative methods to find the global minimum of f when it exists

Convex Hulls and Extreme Points

Convex Hulls

Definition: The convex hull of a set $A \subseteq \mathbb{R}^n$ is the intersection of all convex sets containing A, formally

$$\text{cvx}(A) \ = \ \bigcap \{C \subseteq \mathbb{R}^n : A \subseteq C \text{ and } C \text{ convex}\}$$

- $\text{cvx}(A)$ is convex, and is the smallest convex set containing A
- $A \subseteq \text{cvx}(A)$ with equality iff A is convex
- $\text{cvx}(A)$ can be open, closed, or neither

Fact: $\text{cvx}(A)$ equal to the set of all finite convex combinations of points in A

$$\text{cvx}(A) \ = \ \left\{ \sum_{j=1}^{k} \alpha_j x_j : k \geq 1, x_1^k \in A, \alpha_j \geq 0, \sum_{j=1}^{k} \alpha_j = 1 \right\}$$

Convex Hulls and Subgradients

Fact: If $f_1, \ldots, f_n : C \to \mathbb{R}$ are convex and $f = \max(f_1, \ldots, f_n)$ then

$$\partial f(u) \;=\; \text{cvx}\left(\bigcup_{i : f_i(u) = f(u)} \partial f_i(u)\right)$$

That is, the subgradient of f is the convex hull of the subgradients of the functions f_i that achieve the maximum

Extreme Points

Definition: Let $C \subseteq \mathbb{R}^n$ be convex. A element $x \in C$ is an extrema point of C if for all $u, v \in C$ and $\lambda \in (0, 1)$, the relation $x = \lambda u + (1 - \lambda)v$ implies $u = v = x$

- Definition says that an extreme point cannot be expressed as a non-trivial convex combination of other points in C
- Let $\mathcal{E}(C) =$ extreme points of C

Examples:

1. $C = [0, 1], [0, 1), (0, 1)$
2. $C = \{x : ||x|| \leq 1\}, \{x : ||x|| < 1\}$ (closed, open unit ball)
3. $C =$ convex hull of a finite set of points

More on Extreme Points

Theorem (Krein-Millman): If $\emptyset \neq C \subseteq \mathbb{R}^n$ is compact (closed and bounded) then

1. $\mathcal{E}(C) \neq \emptyset$
2. $C = \text{cvx}(\mathcal{E}(C))$ (C is the convex hull of its extreme points)

Fact: If $f : C \to \mathbb{R}$ is strictly convex, then

$$\operatorname*{argmax}_{x \in C} f(x) \subseteq \mathcal{E}(C))$$