Convexity

– Combinations (of vectors in a vector space)
  ↳ functions (convex/concave)
  ↳ sets (convex)

Notation For a vector $x \in \mathbb{R}^n$, we denote by x_i the i^{th} one-dimensional component of x. In particular, $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$x^i \in \mathbb{R}^n$ denotes the i^{th} vector in the vector space.

Def^n (Linear Combination)
Given $x^1, x^2 \in \mathbb{R}^n$, the **linear combination** of x^1 and x^2 is represented as $\bar{x} = \lambda_1 x^1 + \lambda_2 x^2$, $\lambda_1, \lambda_2 \in \mathbb{R}$

In general, for $x^i \in \mathbb{R}^n$, $i = 1 \dots m$, the linear combination of x^i, $i = 1 \dots m$ is represented as $\boxed{\bar{x} = \sum_{i=1}^{m} \lambda_i x^i}$, $\lambda_i \in \mathbb{R}$ $\forall i = 1 \dots m$

Set of all linear combinations of x^i = LINEAR HULL of x^i
Linear hull of x^1 and $x^2 \in \mathbb{R}^2$ as shown below = span $\{x^1, x^2\}$ = lin $\{x^1, x^2\}$ = $\mathbb{R}^2$

Def^n (Affine combination)
For $x^i \in \mathbb{R}^n$, $i = 1 \dots m$, the **affine combination** of x^i, $i = 1 \dots m$ is represented as

$$\bar{x} = \sum_{i=1}^{m} \lambda_i x^i, \quad \lambda_i \in \mathbb{R} \ \forall i = 1 \dots m \quad \text{and} \quad \sum_{i=1}^{m} \lambda_i = 1$$

For x^1 and x^2 as mentioned above, the set of all affine combinations of x^1 and x^2 is called the affine hull of x^1 and x^2 and is denoted by aff $\{x^1, x^2\}$.
For the above example, the **affine hull** is the line joining the two points.

Given $x^1, x^2, x^3 \in \mathbb{R}^2$, three mutually non-collinear points in $\mathbb{R}^2$. What is aff $\{x^1, x^2, x^3\}$ geometrically?

Ans The entire plane $\in \mathbb{R}^2$. Look at it as $\lambda_1 x^1 + \lambda_2 x^2 + (1 - \lambda_1 - \lambda_2) x^3$ with no constraints on λ_i anymore. This is sufficient to span $\mathbb{R}^2$ by itself!

Def^n (Convex Combination)
For $x^i \in \mathbb{R}^n$, $i = 1 \dots m$ the **convex combination** of x^i, $i = 1 \dots m$ is represented as

$$\bar{x} = \sum_{i=1}^{m} \lambda_i x^i, \quad \lambda_i \in \mathbb{R}_+ \ \forall i = 1 \dots m, \quad \sum_{i=1}^{m} \lambda_i = 1$$

Set of all convex combinations = **convex hull** = conv $\{x_1 \dots x_m\}$

the convex hull of x^1 and x^2 is the line segment joining x^1 and x^2.

For three mutually collinear $x^1, x^2, x^3 \in \mathbb{R}^2$, conv $\{x_1, x_2, x_3\}$?

Ans

$\bar{x} = \lambda_1 x^1 + \lambda_2 x^2 + \lambda_3 x^3$, $\lambda_1 + \lambda_2 + \lambda_3 = 1$, $\lambda_i \geq 0$

Consider $\tilde{x} = \frac{\lambda_2}{\lambda_2 + \lambda_3} x^2 + \frac{\lambda_3}{\lambda_2 + \lambda_3} x^3$

$\tilde{x}$ is any point on the segment joining x^2 and x^3.

Now rewrite $\bar{x} = \lambda_1 x^1 + (\lambda_2 + \lambda_3) \tilde{x}$. This is still a convex combination!

$\bar{x}$ is basically any point on the segment joining x^1 and $\tilde{x}$ (any point on the opposite base).

Thus $\bar{x}$ comprises all points in the interior (+ boundary) of the triangle formed by the segments joining x^1, x^2 and x^3 pairwise.

Def^n (Convex set)
$S \subseteq \mathbb{R}^n$ is a **convex set** if $\lambda x^1 + \lambda_2 x^2 \in S \ \forall x^1, x^2 \in S$ and $\lambda_1 + \lambda_2 = 1$, $\lambda_1, \lambda_2 \geq 0$

In other words, $S \subseteq \mathbb{R}^n$ is a convex set if conv $\{x^1 \dots x^m\} \in S$ $\forall x_1 \dots x^m \in S$

Def^n (Linear function/map)
A **linear function** is a map between two vector spaces that preserves vector addition and scaler multiplication.
Specifically, $\forall x, y \in \mathbb{R}^n$,
$f(x + y) = f(x) + f(y)$ addition and $f(\alpha x) = \alpha f(x)$ scaler multiplication

eg ① $f: \mathbb{R}^2 \to \mathbb{R}$, $f(x) = 3x_1 + 2x_2 + 3$ non-linear
② $f: \mathbb{R} \to (-1, 1)$, $f(x) = \cos(x)$ non-linear
③ $f: \mathbb{R}^n \to \mathbb{R}$, $c \in \mathbb{R}^n$, $f(x) = c^T x$ linear!

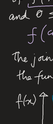
Def^n (Affine function)
An **affine function** is a translated linear function.
$f: \mathbb{R}^n \to \mathbb{R}$ is affine iff $f(x) = \alpha g(x) + \beta$, $\alpha, \beta \in \mathbb{R}$ and $g: \mathbb{R}^n \to \mathbb{R}$.

Note Every linear function is an affine function. ✳
However every affine function need not be a linear function.
eg $f(x) = 3x_1 + 2x_2 + 3$ is not linear but is affine

Def^n (Convex function)
Let $S \subseteq \mathbb{R}^n$ be convex, $f: S \to \mathbb{R}$. f is called a **convex function** on S if $f(\lambda x^1 + (1-\lambda) x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2) \ \forall x^1, x^2 \in S$ and $0 \leq \lambda \leq 1$

f (conv $\{x^1, x^2\}$) $\leq$ conv $\{f(x^1), f(x^2)\}$

The join of two points on a convex function lies above the graph of the function between those two points.



f(x) is convex on S but not on S'!

Def^n (Strictly convex function)
A convex function f is **strictly convex** if $f(\lambda x^1 + (1-\lambda) x^2) < \lambda f(x^1) + (1-\lambda) f(x^2)$ $\forall x^1, x^2 \in S$ and $\alpha < \lambda < 1$

Functions that are convex but not strictly convex – for example, affine functions!

Def^n (Concave functions)
Consider $S \subseteq \mathbb{R}^n$ that is a convex set and $f: S \to \mathbb{R}$ that is a convex function on S. Then f is a **concave function** if $-f$ is convex on S.
Similarly, f is said to be **strictly concave** on S iff $-f$ is strictly convex on S.

Why care about convexity?

Thm Let $S \subseteq \mathbb{R}^n$ be a convex set and $f: \mathbb{R}^n \to \mathbb{R}$ be convex on S. Consider the optimization problem (P).

$$\max \{f(x) : x \in S\}$$

(a) If $\bar{x}$ is a local minimizer of (P) i.e.
$\bar{x} = \text{argmin} \{f(x) : x \in S \cap N_\varepsilon(\bar{x})\}$ then,
$\bar{x}$ is a global minimizer of (P) i.e
$\bar{x} = \text{argmin} \{f(x) : x \in S\}$ ← solution to the problem (P)

(b) If f is strictly convex on S then (P) has at most one global minimizer

Proof
(a) Let $\bar{x}$ be a local minimizer of (P). In particular,
$\bar{x} = \text{argmin} \{f(x) : x \in S \cap N_\varepsilon(\bar{x})\}$
Let $\exists y \in S$ s.t $f(y) < f(\bar{x})$ such that x is NOT a global minimizer
Let z be the strict convex combination of $\bar{x}$ and y,
$z = \lambda \bar{x} + (1-\lambda) y$, $0 < \lambda < 1$
As f is convex on S, $f(z) \leq \lambda f(\bar{x}) + (1-\lambda) f(y)$
$< \lambda f(\bar{x}) + (1-\lambda) f(\bar{x})$
$= f(\bar{x})$
$\therefore f(z) < f(\bar{x}) \ \forall \ 0 < \lambda < 1$ (arbitrarily close to $\bar{x}$)
As $\lambda \to 1$, $z \in S \cap N_\varepsilon(\bar{x})$ ← bring z to the ε nbd of $\bar{x}$
Thus violates local minimality of $\bar{x}$ → contradiction!
Thus, every local minimizer of (P) is also a global minimizer of (P)!

(b) Assume that distinct $\exists \bar{x}, \bar{y} \in S$ s.t $\bar{x}, \bar{y} = \text{argmin} \{f(x) : x \in S\}$
Let $f(\bar{x}) = f(\bar{y}) = \alpha$
Consider $z \in S$ s.t $z = \lambda \bar{x} + (1-\lambda) \bar{y}$, $0 < \lambda < 1$
$f(z) < \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$ ← strict convexity
$f(z) < \lambda \alpha + (1-\lambda) \alpha = \alpha$
Thus z is the global minimizer of (P), not $\bar{x}$ or $\bar{y}$ ↳ contradiction
If f is strictly convex on S, then f has a unique minimizer on S, if one exists!