**Convex Functions**

Let $C \subseteq \mathbb{R}^n$ be a convex set.

**Reminder:** $f : C \to \mathbb{R}$ is **convex** if
\[
f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y) \quad \text{for any } x, y \in C.
\]

Convex combination, i.e. $\lambda + \mu = 1, \lambda, \mu \geq 0$.

**Concave** if $\lambda \leq 1$.

Strictly convex or concave if $\text{sign} < 0$ whenever $x \neq y$.

NB works also for functions with values in $\mathbb{R}^+ \cup \{0\}$ (for convex functions).

In univariate situation, $f$ is convex
- in multivariate case
  - a function is convex if its restrictions to any straight line are.

Epigraph of $f : C \to \mathbb{R}$:
\[
\{(x, a) : f(x) \leq a\} \subseteq C 	imes \mathbb{R}
\]

Immediate from definition: $f$ is convex $\iff$ epigraph is convex

Subsets of $\mathbb{R}^n \times \mathbb{R}$

Another useful property: $f$ convex $\Rightarrow$ $f \leq c^2$ is convex in $\mathbb{R}^n$.

Reverse equivalence not true:

**Lemma (Jensen's inequality)** $f$ convex $\Rightarrow f(\sum \lambda_k x_k) \leq \sum \lambda_k f(x_k)$

**Proof:** automatic

Thus of usage, e.g. $\frac{1}{n} \sum \lambda_k x_k^2 \leq \left(\frac{1}{n} \sum \lambda_k x_k\right)^2$. Or $\exp \left(\frac{\sum \lambda_k x_k}{n}\right) \leq \frac{1}{n} \sum \lambda_k \exp x_k$
Properties of convex functions:
- Linear functions are convex (and concave)
- \( f, g \) convex \( \Rightarrow \) \( f + g \) is; \( \max(f, g) \) is. [Check!]
- \( \| \cdot \| \) is convex for any norm \( \| \cdot \| \).
- \( \lambda f \) is convex if \( f \) is, and \( \lambda > 0 \).

Characterization - convex functions:

Then Gateaux diff. \( f \) is convex on \( C \subset \mathbb{R}^n \) iff
\[
\forall y, z \in C : \quad f(y + \lambda z) \geq f(y) + \lambda \langle \nabla f(y), z \rangle.
\]
Proof. Let \( z = \lambda x + y \) convex comb.
\[
\forall y, z \in C : \quad f(y + \lambda z) \geq f(y) + \lambda \langle \nabla f(y), z \rangle.
\]

Second order characterization of convex functions

Thus, if \( f \) is in \( C^2(\Omega, \mathbb{R}) \), \( \Omega \) - convex, open, then:

a) \( f \) is convex \( \Leftrightarrow \) Hessian of \( f \) is positive semidefinite
b) \( \text{Hess} f \) is positive definite \( \Rightarrow \) \( f \) is strictly convex.

Proof: a) \( \Rightarrow \) If \( \text{Hess} f(x)[\xi] < 0 \), then, by Taylor formula,
\[
\frac{1}{2} \left[ f(x + \xi) + f(x) - 2f(x) \right] < f(x) \quad \text{contradictory convexity}
\]
\[
\leq \quad f(y) = f(x) + \langle \nabla f(x), y-x \rangle + \frac{1}{2} \text{Hess} f(x)[y-x]^2.
\]

b) Proved same way as a \( \Leftarrow \)
Optimization on convex sets

Key result:

Then if \( f \) is convex on a convex set \( C \), then any local minimizer is global. If \( f \) is strictly convex, global minimizer is unique.

Example: For any convex closed \( C \subset \mathbb{R}^n \), and a point \( x \in \mathbb{R}^n \), there is unique \( y \in C \) closest to \( x \):

\[
\min_{y \in C} \|x - y\|
\]

Proof: If \( x_+ \in C \) is local minimizer, \( x_+ \in C \), then \([x_-, x_+] \in C \), and \( f|_{[x_-, x_+]} \) is convex. If \( f(x) < f(x_+) \)

then \( f(x_+ + \varepsilon(x - x_+)) \leq (1 - \varepsilon)f(x_+) + \varepsilon f(x) < f(x_+) \)

\( \Rightarrow x_+ \) cannot be a local minimizer.

Also, if \( f(x) = f(x_+) \), and \( x_+ \) is local minimizer, then \( f|_{[x_-, x_+]} \equiv f(x_+) \), and \( f \) cannot be strictly convex.

The following result is a key tool for finding minimizers (minimizers) on convex sets.

Then \( C \subset \mathbb{R}^n \) is convex, \( f \) - Gateaux differentiable. Then

a) if \( x_+ \) is a local minimizer, \( (\nabla f(x_+), y - x_+) \geq 0 \) for all \( y \in C \)

b) if \( (\nabla f(x_+), y - x_+) \geq 0 \) for all \( y \in C \), \& \( f \) is convex then \( x_+ \) is local minimizer.

Proof: a) \( f(x) \leq f(x_+ + \varepsilon(y - x_+)) = f(x) + \varepsilon(\nabla f(x_+), y - x_+) + o(\varepsilon) \)

b) Property of convex function.
Example: If \( y \in P \) — polyhedron — is the closest point to \( x \in \mathbb{R}^n \), then \( y - x \perp \) to any vector in the facet of \( P \) containing \( y \).