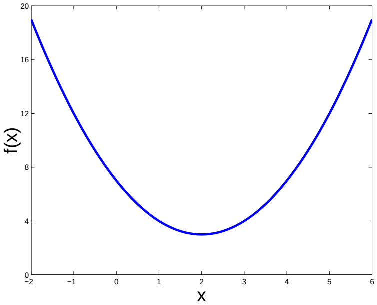
Numerical Optimization

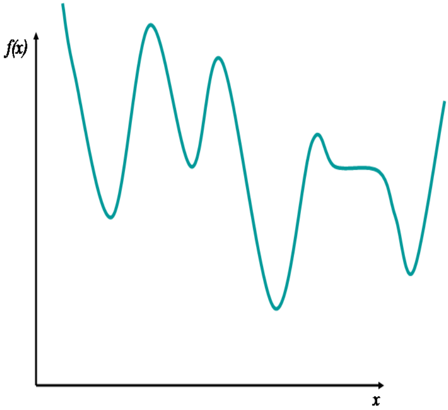
Convex Functions

Shirish Shevade

Computer Science and Automation
Indian Institute of Science
Bangalore 560 012, India.

NPTEL Course on Numerical Optimization


20
16
12
8
4
0
−2
−1
0
1
2
3
4
5
6
f(x)
x

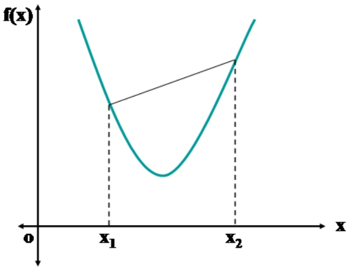

f(x)
x

Convex functions

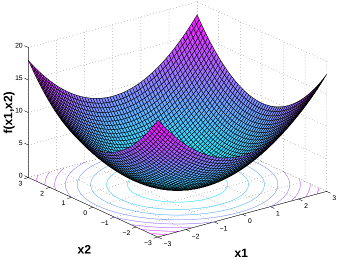
Definition

Let $C \subseteq \mathbb{R}^n$ be a convex set. A function $f : C \to \mathbb{R}$ is said to be convex if for any $\mathbf{x}_1, \mathbf{x}_2 \in C$ and $\lambda \in [0, 1]$,

$$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2).$$



- f is strictly convex if the above inequality is strict for any $\mathbf{x}_1, \mathbf{x}_2 \in C, \mathbf{x}_1 \neq \mathbf{x}_2$ and $\lambda \in (0, 1)$.

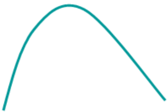


Convex function

Concave functions

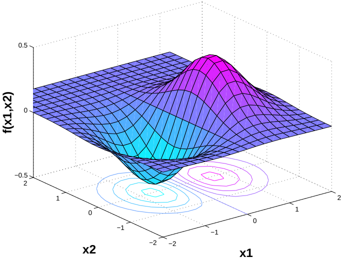
Let $C \subseteq \mathbb{R}^n$ be a convex set. A function $f : C \to \mathbb{R}$ is said to be

- concave iff $-f$ is convex
- strictly concave iff $-f$ is strictly convex.



Concave

Neither convex nor concave



Neither convex nor concave

Examples

- $f(\mathbf{x}) = \mathbf{a}^T\mathbf{x} + b$ is both convex and concave on $\mathbb{R}^n$.
- $f(x) = e^{ax}$ is convex on $\mathbb{R}$, for any $a \in \mathbb{R}$.
- $f(x) = \log x$ is concave on $\{x \in \mathbb{R} : x > 0\}$.
- $f(x) = x^3$ is neither convex nor concave on $\mathbb{R}$.
- $f(x) = |x|$ is convex on $\mathbb{R}$.

Why convex functions?

Let $X \subseteq \mathbb{R}^n$ and $f : X \to \mathbb{R}$. Consider the problem,

$$\begin{array}{rl} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in X \ldots\ldots (1) \end{array}$$

Recall the definition of a global and a local minimum.

- *If there exists* $\mathbf{x}^* \in X$ *such that* $f(\mathbf{x}^*) \leq f(\mathbf{x})$ *for every* $\mathbf{x} \in X$*, then* $\mathbf{x}^*$ *is said to be a global minimum of* f *over* X.
- $\bar{\mathbf{x}}$ *is said to be a local minimum of* f *over* X *if there exists* $\delta > 0$ *such that* $f(\bar{\mathbf{x}}) \leq f(\mathbf{x})$ *for every* $\mathbf{x} \in X \cap B(\bar{\mathbf{x}}, \delta)$.

If f is a convex function and X is a convex set, then every local minimum of (1) is a global minimum.

Convex Programming Problem

Let $C \subseteq \mathbb{R}^n$ be a nonempty convex set and $f : C \to \mathbb{R}$ be a convex function.

Convex Programming Problem (CP):

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C \end{array}$$

Theorem

Every local minimum of a convex programming problem is a global minimum.

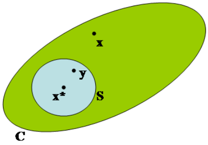
Theorem

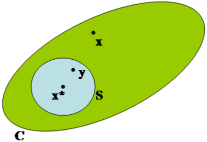
Every local minimum of a convex programming problem is a global minimum.

Proof.

(I) The theorem is trivially true if C is a singleton set.

(II) Assume that there exists $\mathbf{x}^* \in C$ which is a *local minimum* of f over C.
$\mathbf{x}^*$ is a local minimum
$\Rightarrow \; \exists\, \delta > 0 \ni f(\mathbf{x}^*) \leq f(\mathbf{x}) \; \forall\, \mathbf{x} \in C \cap B(\mathbf{x}^*, \delta)$.





Proof. (continued)

Let $S = C \cap B(\mathbf{x}^*, \delta)$.
We already have $f(\mathbf{x}^*) \le f(\mathbf{x}) \;\forall\; \mathbf{x} \in S \quad \ldots (1)$. It is enough to show that $f(\mathbf{x}^*) \le f(\mathbf{x}) \;\forall\; \mathbf{x} \in C \backslash S$. Let $\mathbf{y} \in S$, $\mathbf{y} \neq \mathbf{x}^*$.
Consider any $\mathbf{x} \in C \backslash S$ such that $\mathbf{x}$ lies on the extended line segment $LS[\mathbf{x}^*, \mathbf{y}]$.
Since C is convex, $\mathbf{y} = \lambda \mathbf{x}^* + (1-\lambda)\mathbf{x} \in C \;\forall\; \lambda \in (0, 1)$.

$$\begin{aligned} f(\mathbf{x}^*) &\le f(\mathbf{y}) \\ &= f(\lambda \mathbf{x}^* + (1-\lambda)\mathbf{x}) \\ &\le \lambda f(\mathbf{x})^* + (1-\lambda) f(\mathbf{x}) \;\; \text{(since f is convex)} \\ \therefore\; f(\mathbf{x}^*) &\le f(\mathbf{x}) \;\forall\; \mathbf{x} \in C \backslash S. \quad \ldots (2) \end{aligned}$$

From (1) and (2), $\mathbf{x}^*$ is a global minimum of f over C. □

Theorem

The set of all optimal solutions to the convex programming problem is convex.

Proof.

(I) The theorem is true if there is an unique optimal solution.
(II) Let $S = \{\mathbf{z} \in C : f(\mathbf{z}) \leq f(\mathbf{x}), \mathbf{x} \in C\}$. We need to show that S is a convex set.
Let $\mathbf{x}_1, \mathbf{x}_2 \in S, \mathbf{x}_1 \neq \mathbf{x}_2$.
$\therefore\ f(\mathbf{x}_1) = f(\mathbf{x}_2), f(\mathbf{x}_1) \leq f(\mathbf{x}), f(\mathbf{x}_2) \leq f(\mathbf{x})\ \forall\ \mathbf{x} \in C$.
Since $\mathbf{x}_1, \mathbf{x}_2 \in C$ and C is a convex set,
$\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in C\ \forall\ \lambda \in [0, 1]$.
Since f is convex, we have, for any $\lambda \in [0, 1]$,
$f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) = f(\mathbf{x}_2)$
This implies that S is a convex set.

$\square$

Epigraph

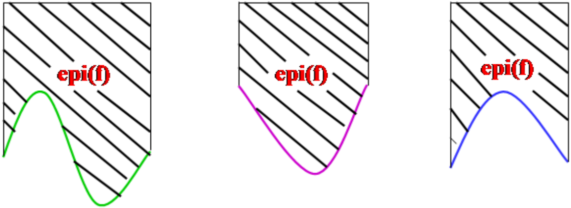
Let $X \subseteq \mathbb{R}^n$ and $f : X \to \mathbb{R}$

Describe f by its graph, $\{(\mathbf{x}, f(\mathbf{x})) : \mathbf{x} \in X\} \subseteq \mathbb{R}^{n+1}$

Definition

The epigraph of f, $\text{epi}(f)$ is a subset of $\mathbb{R}^{n+1}$ and is defined by

$$\{(\mathbf{x}, y) : \mathbf{x} \in X, y \in \mathbb{R}, y \geq f(\mathbf{x})\}$$



Characterization of a convex function

Theorem

Let $C \subseteq \mathbb{R}^n$ be a convex set and $f : C \to \mathbb{R}$. Then f is convex iff $epi(f)$ is a convex set.

Proof.

(I). Assume that f is convex. Let $(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2) \in \text{epi}(f)$.
Therefore, $y_1 \geq f(\mathbf{x}_1)$ and $y_2 \geq f(\mathbf{x}_2)$.
f is a convex function. So, for any $\lambda \in [0, 1]$, we can write,

$$\begin{aligned} f(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) &\leq \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) \\ &\leq \lambda y_1 + (1-\lambda) y_2 \end{aligned}$$

Therefore, we have $(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2, \lambda y_1 + (1-\lambda) y_2) \in \text{epi}(f)$
$\Rightarrow \text{epi}(f)$ is a convex set.

Proof (continued)

(II). Assume that epi(f) is a convex set. Let $\mathbf{x}_1, \mathbf{x}_2 \in C$.
$\therefore$ $(\mathbf{x}_1, f(\mathbf{x}_1)), (\mathbf{x}_2, f(\mathbf{x}_2)) \in \text{epi}(f)$.
$\therefore$ $(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2, \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2)) \in \text{epi}(f)$ for any $\lambda \in [0, 1]$
$\therefore$ $\lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \geq f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2)$ for any $\lambda \in [0, 1]$
$\therefore$ f is convex. $\square$

Level set

Let $C \subseteq \mathbb{R}^n$ be a convex set and $f : C \rightarrow \mathbb{R}$ be a convex function.
Define the level set of f for a given α as
$C_\alpha = \{\mathbf{x} \in C : f(\mathbf{x}) \leq \alpha, \alpha \in \mathbb{R}\}$.

Theorem

If f is a convex function, then the level set C_α is a convex set.

Proof.

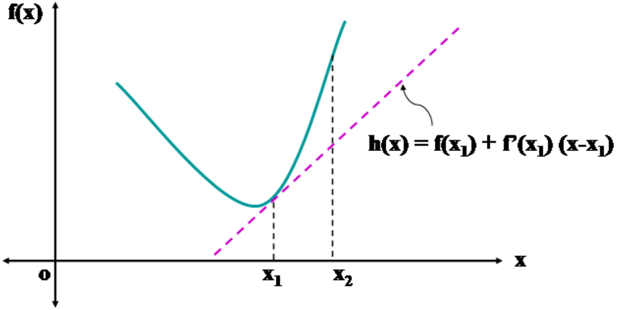
Let $\mathbf{x}, \mathbf{y} \in C_\alpha$.
$\therefore$ $\mathbf{x}, \mathbf{y} \in C$ and $f(\mathbf{x}) \leq \alpha, f(\mathbf{y}) \leq \alpha$.
Let $\mathbf{z} = \lambda\mathbf{x} + (1 - \lambda)\mathbf{y}$ where $\lambda \in (0, 1)$.
Clearly, $\mathbf{z} \in C$.
Since f is convex, $f(\mathbf{z}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y}) \leq \alpha$.
$\therefore$ $\mathbf{z} \in C_\alpha \Rightarrow C_\alpha$ is convex. $\square$

Theorem

Let $C \subseteq \mathbb{R}^n$ be a convex set and $f : C \to \mathbb{R}$ be a differentiable function. Let $g(\mathbf{x}) = \nabla f(\mathbf{x})$. Then f is convex iff

$$f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in C$. Further, f is strictly convex iff the above inequality is strict for all $\mathbf{x}_1, \mathbf{x}_2 \in C, \mathbf{x}_1 \neq \mathbf{x}_2$.



Proof.

(I). Assume that f is convex.

$\therefore\ f(\lambda\mathbf{x}_2 + (1-\lambda)\mathbf{x}_1) \leq \lambda f(\mathbf{x}_2) + (1-\lambda)f(\mathbf{x}_1)\ \forall\ \lambda \in [0,1]$

That is, $f(\mathbf{x}_1 + \lambda(\mathbf{x}_2 - \mathbf{x}_1)) \leq f(\mathbf{x}_1) + \lambda(f(\mathbf{x}_2) - f(\mathbf{x}_1))$.

$\therefore\ \frac{f(\mathbf{x}_1+\lambda(\mathbf{x}_2-\mathbf{x}_1))-f(\mathbf{x}_1)}{\lambda} \leq f(\mathbf{x}_2) - f(\mathbf{x}_1)$.

Letting $\lambda \to 0^+$, we get

$$g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) \leq f(\mathbf{x}_2) - f(\mathbf{x}_1)$$

Proof.(Continued)

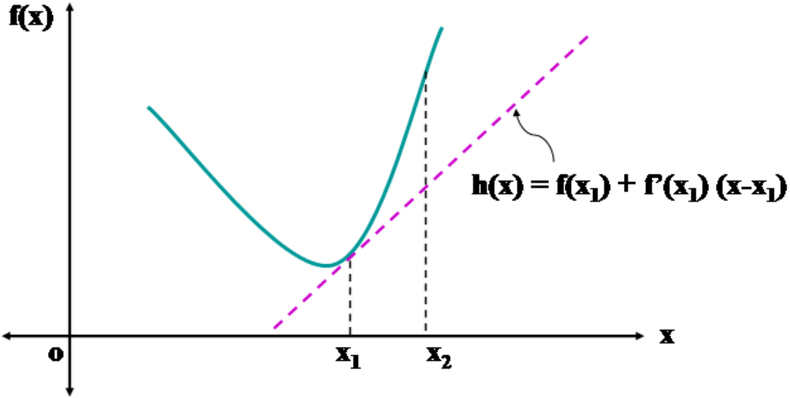
(II). Assume that $f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1)$ holds for any $\mathbf{x}_1, \mathbf{x}_2 \in C$. Let $\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$ where $\lambda \in [0, 1]$.

$$\begin{aligned} \therefore f(\mathbf{x}_1) &\geq f(\mathbf{x}) + g(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x}) \quad \ldots (a) \\ f(\mathbf{x}_2) &\geq f(\mathbf{x}) + g(\mathbf{x})^T(\mathbf{x}_2 - \mathbf{x}) \quad \ldots (b) \end{aligned}$$

Multiplying (a) by λ and (b) by $(1-\lambda)$ and adding, we get,

$$\begin{aligned} & \lambda f(\mathbf{x}_1) + (1-\lambda) f(\mathbf{x}_2) \\ \geq\ & f(\mathbf{x}) + \lambda g(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x}) + (1-\lambda) g(\mathbf{x})^T(\mathbf{x}_2 - \mathbf{x}) \\ =\ & f(\mathbf{x}) + \lambda g(\mathbf{x})^T(\mathbf{x}_1 - \mathbf{x}_2) + g(\mathbf{x})^T(\mathbf{x}_2 - \mathbf{x}) \\ =\ & f(\mathbf{x}) + g(\mathbf{x})^T(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 - \mathbf{x}) \\ =\ & f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \\ \Rightarrow\ & \text{f is convex.} \qquad \square \end{aligned}$$

- Let $C \subseteq \mathbb{R}^n$ and $f : C \rightarrow \mathbb{R}$ be a differentiable convex function on a convex set C. Then, the first order approximation of f at any $\mathbf{x}_1 \in C$ never *overestimates* $f(\mathbf{x}_2)$ for any $\mathbf{x}_2 \in C$.



- Let $C \subseteq \mathbb{R}$ be an open convex set and $f : C \to \mathbb{R}$ be a differentiable convex function on C.
  Consider $x_1, x_2 \in C$ such that $x_1 < x_2$. We therefore have

$$f(x_1) \geq f(x_2) + f'(x_2)(x_1 - x_2)$$
$$f(x_2) \geq f(x_1) + f'(x_1)(x_2 - x_1)$$

  Hence,
  $f'(x_2)(x_2 - x_1) \geq f(x_2) - f(x_1) \geq f'(x_1)(x_2 - x_1)$.
  This implies,
  $f'(x_2) \geq f'(x_1) \; \forall \; x_2 > x_1$.

If f is a differentiable convex function of one variable defined on an open interval C, then the derivative of f is non-decreasing.

The converse of this statement is also true.

- Consider the **Convex Programming Problem (CP)**:

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C \end{array}$$

where f is differentiable.
Let $\hat{\mathbf{x}} \in C$.
The optimal objective function value of the problem,

$$\begin{array}{ll} \min & f(\hat{\mathbf{x}}) + g(\hat{\mathbf{x}})^T(\mathbf{x} - \hat{\mathbf{x}}) \\ \text{s.t.} & \mathbf{x} \in C \end{array}$$

gives a lower bound on the optimal objective function value of **CP**.

- Again, consider the **Convex Programming Problem (CP)**:

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C \end{array}$$

where f is differentiable and C is an open convex set.

  - Let $\mathbf{x}^* \in C$ such that $g(\mathbf{x}^*) = \nabla f(\mathbf{x}^*) = \mathbf{0}$.
    Then,

$$\begin{array}{rcl} f(\mathbf{x}) & \geq & f(\mathbf{x}^*) + g(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \\ \Rightarrow f(\mathbf{x}) & \geq & f(\mathbf{x}^*) \; \forall \, \mathbf{x} \in C \\ \Rightarrow \mathbf{x}^* & \text{is} & \text{a global minimum of } f \text{ over } C. \end{array}$$

Theorem

Let $f : C \rightarrow \mathbb{R}$ be a twice differentiable function on an open convex set $C \subseteq \mathbb{R}^n$. Then f is convex iff its Hessian matrix, $\mathbf{H}(\mathbf{x})$, is positive semi-definite for each $\mathbf{x} \in C$.

Proof.

(I). Let $\mathbf{x}_1, \mathbf{x}_2 \in C$ and $\mathbf{H}(\mathbf{x})$, be positive semi-definite for each $\mathbf{x} \in C$.
Let $\mathbf{x} = \lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2,\ \lambda \in (0,1)$.
Using second order truncated Taylor series, we have,
$f(\mathbf{x}_2) = f(\mathbf{x}_1) + g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) + \frac{1}{2}(\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{H}(\mathbf{x})(\mathbf{x}_2 - \mathbf{x}_1)$.
That is, $f(\mathbf{x}_2) \geq f(\mathbf{x}_1) + g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1)$ (since H is psd)
Hence, f is convex.

Proof. (continued)

(II). Let $\mathbf{H}$ be *not* positive semi-definite for some $\mathbf{x}_1 \in C$.

$\therefore\ \exists\ \mathbf{x}_2 \in C\ \ni\ (\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{H}(\mathbf{x}_1)(\mathbf{x}_2 - \mathbf{x}_1) < 0.$

Let $\mathbf{x} = \lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2,\ \lambda \in (0,1).$

Using second order truncated Taylor series, we have,

$f(\mathbf{x}_2) = f(\mathbf{x}_1) + g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1) + \frac{1}{2}(\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{H}(\mathbf{x})(\mathbf{x}_2 - \mathbf{x}_1).$

Choose $\mathbf{x}$ sufficiently close to $\mathbf{x}_1$ so that

$(\mathbf{x}_2 - \mathbf{x}_1)^T \mathbf{H}(\mathbf{x})(\mathbf{x}_2 - \mathbf{x}_1) < 0.$

$\therefore\ f(\mathbf{x}_2) < f(\mathbf{x}_1) + g(\mathbf{x}_1)^T(\mathbf{x}_2 - \mathbf{x}_1).$

This implies that f is not convex.

$\square$

f is strictly convex on C if the Hessian matrix $\mathbf{H}(\mathbf{x})$ of f is positive definite for all $\mathbf{x} \in C$.

Examples

- Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be defined as

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} + \mathbf{b}^T\mathbf{x} + c$$

  where $\mathbf{A}$ is a symmetric matrix in $\mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.
  The Hessian matrix of f is $\mathbf{A}$ at any $\mathbf{x} \in \mathbb{R}^n$.
  $\therefore$ f is convex iff $\mathbf{A}$ is positive semi-definite.
- Let $f(x) = x \log x$ be defined on $C = \{x \in \mathbb{R} : x > 0\}$.
  $f'(x) = 1 + \log x$ and $f''(x) = \frac{1}{x} > 0 \; \forall \; x \in C$
  So, $f(x)$ is convex.

- $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|^2$
  Or, $f(\mathbf{x}) = \frac{1}{2}(\mathbf{Ax} - \mathbf{b})^T(\mathbf{Ax} - \mathbf{b})$
  $\nabla^2 f(\mathbf{x}) = \mathbf{A}^T\mathbf{A}$ which is positive semi-definite.
  $\therefore f$ is convex.
- $f(x) = \log(x)$ defined on $C = \{x \in \mathbb{R} : x > 0\}$.
  $f'(x) = \frac{1}{x}$ and $f''(x) = -\frac{1}{x^2} < 0 \; \forall \, x \in C$.
  So, f is concave.

Jensen's inequality

Jensen's inequality

If $f : C \rightarrow \mathbb{R}$ is a function on a convex set $C \subseteq \mathbb{R}^n$. Then f is convex iff

$$f(\sum_{i=1}^{k} \lambda_i \mathbf{x}_i) \leq \sum_{i=1}^{k} \lambda_i f(\mathbf{x}_i) \quad \ldots \text{(JI)}$$

where $\mathbf{x}_1, \ldots, \mathbf{x}_k \in C$, $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$.

- Useful in deriving many inequalities like AM-GM inequality or Hölder inequality

Proof.

(I) Suppose f is a convex function.
Let us prove the inequality by induction on k.
If $k = 2$ the inequality (JI) holds for a convex function.

Proof. (continued)

Let $k > 2$ and the inequality (JI) holds for any collection of $k-1$ points in C.
Now, consider $f(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \ldots + \lambda_k\mathbf{x}_k)$ where $\lambda_1, \ldots, \lambda_k \geq 0$ and $\sum_{i=1}^{k} \lambda_i = 1$.
Let $\delta = \sum_{i=1}^{k-1} \lambda_i$. Note that $\delta + \lambda_k = 1$.

$$
\begin{aligned}
& f(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \ldots + \lambda_k\mathbf{x}_k) \\
= \; & f(\delta(\frac{\lambda_1}{\delta}\mathbf{x}_1 + \ldots \frac{\lambda_{k-1}}{\delta}\mathbf{x}_{k-1})) + \lambda_k\mathbf{x}_k) \\
\leq \; & \delta f(\frac{\lambda_1}{\delta}\mathbf{x}_1 + \ldots + \frac{\lambda_{k-1}}{\delta}\mathbf{x}_{k-1}) + \lambda_k f(\mathbf{x}_k) \\
\leq \; & \delta(\frac{\lambda_1}{\delta}f(\mathbf{x}_1) + \ldots + \frac{\lambda_{k-1}}{\delta}f(\mathbf{x}_{k-1})) + \lambda_k f(\mathbf{x}_k) \\
= \; & \lambda_1 f(\mathbf{x}_1) + \ldots + \lambda_k f(\mathbf{x}_k)
\end{aligned}
$$

(II) The converse is easy to prove. $\square$

- Arithmetic-geometric mean inequality can be derived using Jensen's inequality.

Consider the convex function $f(x) = -\log(x)$ defined on $C = \{x \in \mathbb{R} : x > 0\}$.
Let $x_1, x_2, \ldots, x_k \in$C.
Letting $\lambda_1 = \ldots = \lambda_k = \frac{1}{k}$ and applying Jensen's inequality , we get

$$\begin{aligned}
-\log(\sum_{i=1}^{k} \lambda_i x_i) &\leq -\frac{1}{k}(\sum_{i=1}^{k} \log(x_i)) \\
\therefore \ \log(\frac{x_1 + \ldots + x_k}{k}) &\geq \frac{1}{k}\log(x_1 x_2 \ldots x_k) \\
\therefore \ \frac{x_1 + \ldots + x_k}{k} &\geq (x_1 x_2 \ldots x_k)^{\frac{1}{k}}
\end{aligned}$$

Operations that preserve convexity

- If $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function and $\alpha > 0$, then, αf is a convex function.

f is a convex function. Therefore, for any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ and $\lambda \in [0, 1]$,

$$f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2).$$

Multiplying both sides by α gives the result.

Operations that preserve convexity

- Let $f_1, \ldots, f_k : \mathbb{R}^n \rightarrow \mathbb{R}$ be convex functions. Then, $f(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i f_i(\mathbf{x})$ where $\alpha_i > 0 \; \forall \; i = 1, \ldots, k$ is a convex function.

Consider two convex functions f_1 and f_2 and let
$f(\mathbf{x}) = f_1(\mathbf{x}) + f_2(\mathbf{x})$.
For any $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ and $\lambda \in [0, 1]$,

$$\begin{aligned} f_1(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) &\leq \lambda f_1(\mathbf{x}_1) + (1-\lambda) f_1(\mathbf{x}_2) \\ f_2(\lambda \mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) &\leq \lambda f_2(\mathbf{x}_1) + (1-\lambda) f_2(\mathbf{x}_2) \end{aligned}$$

Adding the two inequalities, we get that $f_1 + f_2$ is a convex function.
Easy to extend the idea to the general result.

Operations that preserve convexity

- Let $h : \mathbb{R} \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$. Consider the function $\psi(x) = h(f(x))$. Under what conditions is ψ convex?

Let f and h be twice differentiable.
Need to find the conditions under which $\psi''(x) \geq 0$.

$$\psi''(x) = h''(f(x))f'(x)^2 + h'(f(x))f''(x)$$

- ψ is convex if h is convex and non-decreasing, and f is convex,
- ψ is convex if h is convex and non-increasing, and f is concave.

Theorem

Let $C \subset \mathbb{R}^n$ be a compact convex set and $f : C \to \mathbb{R}$ be a convex function. Then the maximum of f occurs at a boundary point of C.

Proof.

Suppose the maximum exists at a point $\mathbf{x}^*$ which is in the interior of the set C. That is, $f(\mathbf{x}^*) \geq f(\mathbf{x}) \ \forall \ \mathbf{x} \in C$ and $\mathbf{x}^*$ is in the interior of C.

Draw a line through $\mathbf{x}^*$ cutting the boundary of C at $\mathbf{x}_1$ and $\mathbf{x}_2$.

We can write $\mathbf{x}^* = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$ for some $\lambda \in (0, 1)$.

Since f is convex, $f(\mathbf{x}^*) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2)$.

(i) $f(\mathbf{x}_1) < f(\mathbf{x}_2) \Rightarrow f(\mathbf{x}^*) < f(\mathbf{x}_2) \Rightarrow \mathbf{x}^*$ is not a global max.

(ii) $f(\mathbf{x}_1) > f(\mathbf{x}_2) \Rightarrow f(\mathbf{x}^*) < f(\mathbf{x}_1) \Rightarrow \mathbf{x}^*$ is not a global max.

(iii) $f(\mathbf{x}_1) = f(\mathbf{x}_2) \Rightarrow f(\mathbf{x}^*) \leq f(\mathbf{x}_1) = f(\mathbf{x}_2) \Rightarrow$ either $f(\mathbf{x}_1) = f(\mathbf{x}_2) = f(\mathbf{x}^*)$ or $\mathbf{x}^*$ is not a global maximum.

$\therefore$ The maximum of f occurs at a boundary point.

$\square$