Episode 03
Convex Functions

Ahmad F. Taha

CE 5999-02 Special Topics — Intro to Optimization

Email: ahmad.taha@vanderbilt.edu

Webpage: http://lab.vanderbilt.edu/taha

VANDERBILT UNIVERSITY

January 9, 2023

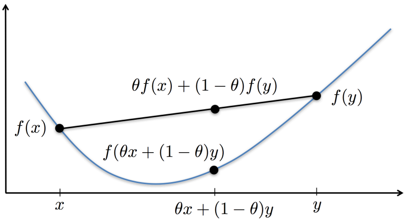
In this module

- Definition of convex and concave functions
- Examples
- Criterion for convexity based on Gradient vector and Hessian matrix
- Operations that preserve convexity
- Epigraph and sublevel sets
- Extended-value extensions
- Restriction to a line
- Quasiconvexity and quasiconcavity; examples

Definition of convexity

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if 1) its domain is a convex set; and 2) it satisfies the following property:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y), \quad \text{for all } x, y \in \mathrm{dom}(f),\ \theta \in [0,1].$$



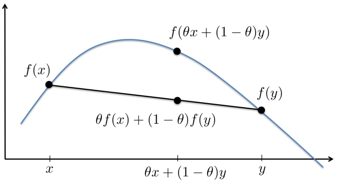
A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strictly convex* if 1) its domain is a convex set; and 2) it satisfies the following property:

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y), \quad \text{for all } x, y \in \mathrm{dom}(f),\ x \neq y,\ \theta \in (0,1).$$

Concavity

A function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if $-f$ is convex.
Equivalently, a function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if 1) its domain is a convex set; and 2) it satisfies the following property:

$$f(\theta x + (1-\theta)y) \geq \theta f(x) + (1-\theta)f(y), \quad \text{for all } x, y \in \mathrm{dom}(f),\ \theta \in [0,1].$$



A function $f : \mathbb{R}^n \to \mathbb{R}$ is *strictly concave* if 1) its domain is a convex set; and 2) it satisfies the following property:

$$f(\theta x + (1-\theta)y) > \theta f(x) + (1-\theta)f(y), \quad \text{for all } x, y \in \mathrm{dom}(f),\ x \neq y,\ \theta \in (0,1).$$

Notes

- For affine functions, we always have equality in the definition of convex/concave functions
- Hence: all affine (and also linear) functions are convex and concave
- Conversely, any function that is convex and concave is affine
- A convex function is continuous on the interior of its domain—it can have discontinuities only on its boundary

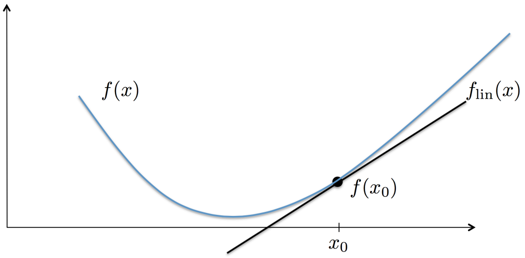
Examples on $\mathbb{R}$

- x^α is convex on $\mathbb{R}_{++}$ if $\alpha \geq 1$ or $\alpha \leq 0$; concave if $0 \leq \alpha \leq 1$
- $\log x$ is concave on $\mathbb{R}_{++}$
- e^{ax} is convex
- $|x|$ is convex
- $ax + b$ is both convex and concave

Criterion based on gradient vector (1st-order condition)

A differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if 1) its domain is a convex set; and 2) its linear approximation (1st-order Taylor approximation) at any point is a global underestimator of the function. That is,

$$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0) \quad \text{for all } x_0, x \in \mathrm{dom}(f)$$



Additional 1st-order conditions

A differentiable function[1] $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is concave if and only if 1) its domain is a convex set; and 2) its linear approximation (1st-order Taylor approximation) at any point is a global overestimator of the function. That is,

$$f(x) \leq f(x_0) + \nabla f(x_0)^T (x - x_0) \quad \text{for all } x_0, x \in \mathrm{dom}(f)$$

A differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is *strictly convex* if and only if 1) its domain is a convex set; and 2) the following condition holds:

$$f(x) > f(x_0) + \nabla f(x_0)^T (x - x_0) \quad \text{for all } x_0, x \in \mathrm{dom}(f),\ x \neq x_0$$

A differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is *strictly concave* if and only if 1) its domain is a convex set; and 2) the following condition holds:

$$f(x) < f(x_0) + \nabla f(x_0)^T (x - x_0) \quad \text{for all } x_0, x \in \mathrm{dom}(f),\ x \neq x_0$$

[1] This means that its gradient ∇f exists at each point in $\mathrm{dom}(f)$.

Criterion based on Hessian matrix (2nd-order condition)

For a twice differentiable function, the **Hessian matrix** is defined as

$$\nabla^2 f(x) = \left[\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \right]$$

A twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if 1) its domain is a convex set; and 2) its Hessian matrix $\nabla^2 f(x)$ is positive semidefinite for all $x \in \mathrm{dom}(f)$.

A twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is concave if and only if 1) its domain is a convex set; and 2) its Hessian matrix $\nabla^2 f(x)$ is negative semidefinite for all $x \in \mathrm{dom}(f)$.

A twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex if 1) its domain is a convex set; and 2) its Hessian matrix $\nabla^2 f(x)$ is positive definite for all $x \in \mathrm{dom}(f)$. Converse is not true (ex: $f(x) = 1/x^2$ with domain $x \in \{\mathbb{R} - 0\}$).

Examples on $\mathbb{R}^n$

- Quadratic function $f(x) = \frac{1}{2}x^T Px + q^T x + r$
  - Convex if $P \succeq 0$; strictly convex if $P \succ 0$
  - Concave if $P \preceq 0$; strictly concave if $P \prec 0$;
  - Not convex and not concave if P is indefinite
  - Proof: The Hessian is $\nabla^2 f(x) = P$
- Norm $\|x\|$ is convex
  - Proof: Use triangle inequality
- Negative entropy $f(x) = x \log x$ for positive x is convex (domain is convex, take first then second derivative, show it's positive for all positive x)
- Max function $f(x) = \max\{x_1, \ldots, x_n\}$ is convex on $\mathbb{R}^n$
  - apply convex function for $f(x) = \max_i x_i$, i.e.,

$$f(\theta x + (1-\theta)y) = \max_i(\theta x_i + (1-\theta)y_i) \leq \ldots$$

- Log-sum-exp $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$ is convex on $\mathbb{R}^n$ (see next page)

Convexity of log-sum-exp

- $f(x) = \log(e^{x_1} + \cdots + e^{x_n})$; set $z_k = e^{x_k}$ then $\nabla f(x) = \frac{1}{\mathbf{1}^\top z} z$
- The Hessian is (see differentiation rules in Appendix A, Examples A.2 and A.4, of your textbook)

$$\nabla^2 f(x) = \frac{1}{\mathbf{1}^T z}\mathrm{diag}(z) - \frac{1}{(\mathbf{1}^T z)^2} z z^T$$

  where $\mathbf{1}$ is an $n \times 1$ vector of all ones and $\mathrm{diag}(z)$ is a diagonal matrix with the entries of z on the diagonal
- Show that $\nabla^2 f(x) \succeq 0$ by definition. For all $v \in \mathbb{R}^n$, it holds that $v^T \nabla^2 f(x) v \geq 0$:

$$v^T \nabla^2 f(x) v = \frac{1}{(\mathbf{1}^T z)^2} \left[\left(\sum_{i=1}^n z_i \right) \left(\sum_{i=1}^n v_i^2 z_i \right) - \left(\sum_{i=1}^n v_i z_i \right)^2 \right] \geq 0$$

- The latter follows from applying Cauchy-Schwarz inequality

$$(a^T b)^2 \leq \|a\|_2^2 \|b\|_2^2$$

  to vectors with entries $a_i = v_i \sqrt{z_i}$, $b_i = \sqrt{z_i}$

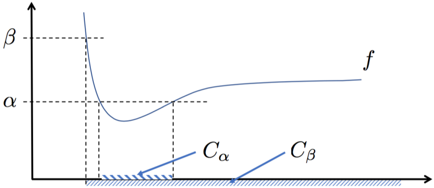
Operations that preserve convexity

1. *Nonnegative weighted sums:* If $f_1, \ldots, f_m$ are convex, and $w_1 \geq 0, \ldots, w_m \geq 0$, then $f = w_1 f_1 + \cdots + w_m f_m$ is convex
2. *Extension to integrals:* If $f(x, a)$ is convex in x for all a, and $w(a) \geq 0$ for all a, then $g(x) = \int f(x, a) w(a) da$ is convex
3. *Pointwise maximum:* If $f_1, \ldots, f_m$ are convex, then $f(x) = \max\{f_1(x), \ldots, f_m(x)\}$ is convex
    - Note that $\mathrm{dom}(f) = \bigcap_{i=1,\ldots,m} \mathrm{dom}(f_i)$ is a convex set
4. *Extension to pointwise supremum:* If $f(x, a)$ is convex in x for all $a \in \mathcal{A}$, then $g(x) = \sup_{a \in \mathcal{A}} f(x, a)$ is convex.
5. Affine transformation of domain: If $f(z)$ is convex, then $f(Ax + b)$ is convex

Sublevel sets

Sublevel set is the set of x such that $f(x)$ is below a specified level α:

$$C_\alpha = \{x \in \mathrm{dom}(f) \mid f(x) \leq \alpha\}$$



Superlevel set is the set ot x such that $f(x)$ is above a specified level α:

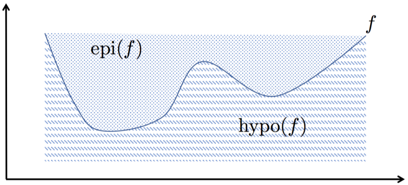
$$S_\alpha = \{x \in \mathrm{dom}(f) \mid f(x) \geq \alpha\}$$

Sublevels sets of convex functions are convex sets (the converse is not true in general).
Superlevel sets of concave functions are convex sets (the converse is not true in general).

Epigraph

Epigraph is the set "above the graph"

$$\text{epi}(f) = \{(x,t) \in \mathbb{R}^{n+1} \mid f(x) \leq t\}$$



Hypograph is the set "below the graph"

$$\text{hypo}(f) = \{(x,t) \in \mathbb{R}^{n+1} \mid f(x) \geq t\}$$

A function is convex if and only if its epigraph is a convex set.
A function is concave if and only if its hypograph is a convex set.

Extended-value extension

For a *convex* function $f : \mathbb{R}^n \to \mathbb{R}$, its extended-value extension $\tilde{f}(x)$ is a function taking values in $\mathbb{R} \cup \{\infty\}$, defined as

$$\tilde{f}(x) = \begin{cases} f(x), \text{ if } x \in \text{dom}(f) \\ \infty, \text{ if } x \notin \text{dom}(f) \end{cases}$$

For a *concave* function $f : \mathbb{R}^n \to \mathbb{R}$, its extended-value extension $\tilde{f}(x)$ is a function taking values in $\mathbb{R} \cup \{-\infty\}$, defined as

$$\tilde{f}(x) = \begin{cases} f(x), \text{ if } x \in \text{dom}(f) \\ -\infty, \text{ if } x \notin \text{dom}(f) \end{cases}$$

- Note that $\text{dom}(f) = \text{dom}(\tilde{f})$, but $\tilde{f}$ is defined over the entire $\mathbb{R}^n$
- The extended-value extension is useful because it allows us to define convex and concave functions over the entire $\mathbb{R}^n$

Extended-value extension examples

- $f(x) = \log x$, then

$$\tilde{f}(x) = \begin{cases} \log x, & \text{if } x > 0 \\ -\infty, & \text{if } x \leq 0 \end{cases}$$

- $f(x) = \frac{1}{x}$, then

$$\tilde{f}(x) = \begin{cases} \frac{1}{x}, & \text{if } x > 0 \\ \infty, & \text{if } x \leq 0 \end{cases}$$

- Indicator function of a convex set C: $I_C(x) = 0$ if $x \in C$ with $\mathrm{dom}(I_C) = C$.

$$\tilde{I}_C(x) = \begin{cases} 0, & \text{if } x \in C \\ \infty, & \text{if } x \notin C \end{cases}$$

Restriction to a line

- A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if the function $g : \mathbb{R} \to \mathbb{R}$

$$g(t) = f(x + tv), \quad \mathrm{dom}(g) = \{t \in \mathbb{R} \,|\, x + tv \in \mathrm{dom}(f)\}$$

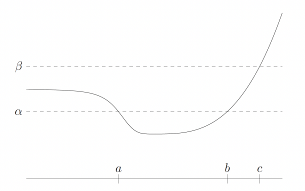
  is convex in t for all $x \in \mathrm{dom}(f)$ and $v \in \mathbb{R}^n$
- $x + tv$: Line passing through x in direction v
- Advantage: Check convexity of f by checking convexity of a 1-D function
- Disadvantage: May be tricky to find $\mathrm{dom}(g)$
- Useful formula: $g''(t) = v^T \nabla^2 f(x + tv) v$

Examples of functions of matrices

- Affine function: $\mathrm{Tr}(A^TX)+b$ is convex and concave on $\mathbb{R}^{m\times n}$
  Proof: $\mathrm{Tr}(A^TX)+b=\sum_{i=1}^{m}\sum_{j=1}^{n}A_{ij}X_{ij}+b$
- $\log\det X$ is concave on $\mathbb{S}^n_{++}$
  Proof: Restriction to a line by defining $g(t)=\log\det(X+tV)$ such that $X+tV$ is a pd matrix then defining $X=X^{1/2}X^{1/2}$ then finding the derivative w.r.t. t...
- $\log\det X^{-1}=-\log\det X$ is convex on $\mathbb{S}^n_{++}$
- $\lambda_{\max}(X)$ is convex on $\mathbb{S}^n$
  Proof: $\lambda_{\max}(X)=\sup_{\|u\|_2=1}u^TXu$
- $\|X\|_F=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}X_{ij}^2}$ (Frobenius norm) is convex on $\mathbb{R}^{m\times n}$
- $\|X\|_*=\sum_{i=1}^{\mathrm{rank}(A)}\sigma_i(A)=\sum_{i=1}^{\min\{m,n\}}\sigma_i(A)$ (nuclear norm) is convex on $\mathbb{R}^{m\times n}$

Quasiconvexity and quasiconavity

A function $f : \mathbb{R}^n \to \mathbb{R}$ is called ***quasiconvex*** if 1) its domain is a convex set; and 2) *for all* $\alpha \in \mathbb{R}$, the sublevel set $S_\alpha = \{x \in \mathrm{dom}(f) \mid f(x) \leq \alpha\}$ is convex. For a function on $\mathbb{R}$, quasiconvexity requires that each sublevel set be an interval (including, possibly, an infinite interval). An example of a quasiconvex function:



A function $f : \mathbb{R}^n \to \mathbb{R}$ is called ***quasiconcave*** if $-f$ is quasiconvex. Equivalently, 1) its domain is a convex set; and 2) *for all* $\alpha \in \mathbb{R}$, the superlevel set $S_\alpha = \{x \in \mathrm{dom}(f) \mid f(x) \geq \alpha\}$ is convex.

A function is called ***quasilinear*** if it is both quasiconvex and quasiconcave

- Convex functions have convex sublevel sets
- So convex functions are quasiconvex (the converse is not true in general)
- Likewise, concave functions are quasiconcave

Examples of quasiconvex and quasiconcave functions

- Logarithm $f(x) = \log x$ is quasilinear on $\mathbb{R}_{++}$
- Ceiling function is also quasilinear
- Linear-fractional function: $f(x) = \frac{a^T x + b}{c^T x + d}$ is quasilinear on $\mathrm{dom}(f) = \{x \,|\, c^T x + d > 0\}$
  - How? First construct the sublevel set S_α then showcase that S_α is indeed quaslilinear
- $f(x) = x_1 x_2$ with $\mathrm{dom}(f) = \mathbb{R}^2_{++}$ is quasiconcave, not convex, not concave, not quasiconvex
  - How? Prove it. :)

Questions And Suggestions?

Thank You!

Please visit
https://lab.vanderbilt.edu/taha/
IFF you want to know more ☺