Lecture 3

Convex Functions

September 2, 2008
Outline

- Convex Functions
- Examples
- Verifying Convexity of a Function
- Operations on Functions Preserving Convexity
Convex Functions

Informally: $f$ is convex when for every segment $[x_1, x_2]$, as $x_\alpha = \alpha x_1 + (1-\alpha)x_2$ varies over the line segment $[x_1, x_2]$, the points $(x_\alpha, f(x_\alpha))$ lie below the segment connecting $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

Let $f$ be a function from $\mathbb{R}^n$ to $\mathbb{R}$, $f : \mathbb{R}^n \to \mathbb{R}$.

The domain of $f$ is a set in $\mathbb{R}^n$ defined by

$$\text{dom}(f) = \{ x \in \mathbb{R}^n | f(x) \text{ is well defined (finite)} \}$$

**Def.** A function $f$ is *convex* if

1. Its domain $\text{dom}(f)$ is a convex set in $\mathbb{R}^n$ and
2. For all $x_1, x_2 \in \text{dom}(f)$ and $\alpha \in (0, 1)$

$$f(\alpha x_1 + (1 - \alpha)x_2) \leq \alpha f(x_1) + (1 - \alpha)f(x_2)$$
More on Convex Function

Def. A function \( f \) is \textit{strictly convex} when \( \text{dom}(f) \) is convex and

\[
f(\alpha x_1 + (1 - \alpha)x_2) < \alpha f(x_1) + (1 - \alpha)f(x_2)
\]

for all \( x_1, x_2 \in \text{dom}(f) \) and \( \alpha \in (0, 1) \)

Def. A function \( f \) is \textit{concave} when \( -f \) is convex, i.e.,

(1) Its domain \( \text{dom}(f) \) is a convex set in \( \mathbb{R}^n \) and

(2) For all \( x_1, x_2 \in \text{dom}(f) \) and \( \alpha \in (0, 1) \)

\[
f(\alpha x_1 + (1 - \alpha)x_2) \geq \alpha f(x_1) + (1 - \alpha)f(x_2)
\]

Def. A function \( f \) is \textit{strictly concave} when \( -f \) is strictly convex
Examples on $\mathbb{R}$

Convex:

- Affine: $ax + b$ over $\mathbb{R}$ for any $a, b \in \mathbb{R}$
- Exponential: $e^{ax}$ over $\mathbb{R}$ for any $a \in \mathbb{R}$
- Power: $x^p$ over $(0, +\infty)$ for $p \geq 1$ or $p \leq 0$
- Powers of absolute value: $|x|^p$ over $\mathbb{R}$ for $p \geq 1$
- Negative entropy: $x \ln x$ over $(0, +\infty)$

Concave:

- Affine: $ax + b$ over $\mathbb{R}$ for any $a, b \in \mathbb{R}$
- Powers: $x^p$ over $(0, +\infty)$ for $0 \leq p \leq 1$
- Logarithm: $\ln x$ over $(0, +\infty)$
Examples: Affine Functions and Norms

- Affine functions are both convex and concave
- Norms are convex

Examples on $\mathbb{R}^n$

- Affine function $f(x) = a'x + b$ with $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$
- Euclidean, $l_1$, and $l_\infty$ norms
- General $l_p$ norms

\[
\|x\|_p = \left( \sum_{i=1}^{n} |x_i|^p \right)^{1/p} \quad \text{for } p \geq 1
\]
Examples on $\mathbb{R}^{m \times n}$

The space $\mathbb{R}^{m \times n}$ is the space of $m \times n$ matrices

- Affine function

$$f(X) = \text{tr}(A^T X) + b = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_{ij} + b$$

- Spectral (maximum singular value) norm

$$f(X) = \|X\|_2 = \sigma_{\max}(X) = \sqrt{\lambda_{\max}(X^T X)}$$

where $\lambda_{\max}(A)$ denotes the maximum eigenvalue of a matrix $A$
Verifying Convexity of a Function

We can verify that a given function $f$ is convex by

- Using the definition

- Applying some special criteria
  - Second-order conditions
  - First-order conditions
  - Reduction to a scalar function

- Showing that $f$ is obtained through operations preserving convexity
Second-Order Conditions

Let $f$ be twice differentiable and let $\text{dom}(f) = \mathbb{R}^n$ [in general, it is required that $\text{dom}(f)$ is open]

The Hessian $\nabla^2 f(x)$ is a symmetric $n \times n$ matrix whose entries are the second-order partial derivatives of $f$ at $x$:

$$
\left[ \nabla^2 f(x) \right]_{ij} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \quad \text{for } i, j = 1, \ldots, n
$$

2nd-order conditions: For a twice differentiable $f$ with convex domain

- $f$ is convex if and only if
  $$
  \nabla^2 f(x) \succeq 0 \quad \text{for all } x \in \text{dom}(f)
  $$

- $f$ is strictly convex if
  $$
  \nabla^2 f(x) \succ 0 \quad \text{for all } x \in \text{dom}(f)
  $$
Examples

Quadratic function: \( f(x) = (1/2)x'Px + q'x + r \) with a symmetric \( n \times n \) matrix \( P \)

\[
\nabla f(x) = Px + q, \quad \nabla^2 f(x) = P
\]

Convex for \( P \succeq 0 \)

Least-squares objective: \( f(x) = \|Ax - b\|^2 \) with an \( m \times n \) matrix \( A \)

\[
\nabla f(x) = 2A^T(Ax - b), \quad \nabla^2 f(x) = 2A^TA
\]

Convex for any \( A \)

Quadratic-over-linear: \( f(x, y) = x^2/y \)

\[
\nabla^2 f(x, y) = \frac{2}{y^3} \begin{bmatrix} y \\ -x \end{bmatrix} \begin{bmatrix} y \\ -x \end{bmatrix}^T \succeq 0
\]

Convex for \( y > 0 \)
Verifying Convexity of a Function

We can verify that a given function $f$ is convex by

- Using the definition

- Applying some special criteria
  - Second-order conditions
  - First-order conditions
  - Reduction to a scalar function

- Showing that $f$ is obtained through operations preserving convexity
First-Order Condition

$f$ is differentiable if $\text{dom}(f)$ is open and the gradient

$$\nabla f(x) = \left( \frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

exists at each $x \in \text{dom} f$

1st-order condition: differentiable $f$ is convex if and only if its domain is convex and

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

A first order approximation is a global underestimate of $f$

Very important property used in algorithm designs and performance analysis
Restriction of a convex function to a line

\( f \) is convex if and only if \( \text{dom} f \) is convex and the function \( g : \mathbb{R} \to \mathbb{R} \),
\[ g(t) = f(x + tv), \quad \text{dom} g = \{ t \mid x + tv \in \text{dom}(f) \} \]
is convex (in \( t \)) for any \( x \in \text{dom} f, v \in \mathbb{R}^n \)

Checking convexity of multivariable functions can be done by checking convexity of functions of one variable

Example \( f : S^n \to \mathbb{R} \) with \( f(X) = -\ln \det X \), \( \text{dom} f = S^n_{++} \)

\[ g(t) = -\ln \det(X + tv) = -\ln \det X - \ln \det(I + tX^{-1/2} V X^{-1/2}) \]
\[ = -\ln \det X - \sum_{i=1}^{n} \ln(1 + t\lambda_i) \]

where \( \lambda_i \) are the eigenvalues of \( X^{-1/2} V X^{-1/2} \)
\( g \) is convex in \( t \) (for any choice of \( V \) and any \( X \succ 0 \)); hence \( f \) is concave
Operations Preserving Convexity

- Positive Scaling
- Sum
- Composition with affine function
- Pointwise maximum and supremum
- Composition
- Minimization
Scaling, Sum, & Composition with Affine Function

Positive multiple: For a convex $f$ and $\lambda > 0$, the function $\lambda f$ is convex.

Sum: For convex $f_1$ and $f_2$, the sum $f_1 + f_2$ is convex (extends to infinite sums, integrals).

Composition with affine function: For a convex $f$ and affine $g$ [i.e., $g(x) = Ax + b$], the composition $f \circ g$ is convex, where $(f \circ g)(x) = f(Ax + b)$.

Examples
- Log-barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x), \quad \text{dom} f = \{x \mid a_i^T x < b_i, i = 1, \ldots, m\}$$

- (Any) Norm of affine function: $f(x) = \|Ax + b\|$
**Pointwise maximum**

For convex functions \( f_1, \ldots, f_m \), the pointwise-max function
\[
F(x) = \max \{ f_1(x), \ldots, f_m(x) \}
\]
is convex (What is domain of \( F \)?)

Examples

- Piecewise-linear function: \( f(x) = \max_{i=1,\ldots,m} (a_i^T x + b_i) \) is convex
- Sum of \( r \) largest components of a vector \( x \in \mathbb{R}^n \):

\[
f(x) = x[1] + x[2] + \cdots + x[r]
\]
is convex (\( x[i] \) is \( i \)-th largest component of \( x \))

\[
f(x) = \max_{(i_1,\ldots,i_r) \in I_r} \{ x_{i_1} + x_{i_2} + \cdots + x_{i_r} \}
\]

\[
I_r = \{ (i_1,\ldots,i_r) \mid i_1 < \ldots < i_r, \ i_j \in \{1,\ldots,m\}, \ j = 1,\ldots,n \}
\]
Pointwise Supremum

Let \( A \subseteq \mathbb{R}^p \) and \( f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R} \). Let \( f(x, z) \) be convex in \( x \) for each \( z \in A \). Then, the supremum function over the set \( A \) is convex:

\[
g(x) = \sup_{z \in A} f(x, z)
\]

Examples

- **Set support function is convex** for a set \( C \subset \mathbb{R}^n \),
  \[
  S_C : \mathbb{R}^n \to \mathbb{R}, \quad S_C(x) = \sup_{z \in C} z^T x
  \]

- **Set farthest-distance function is convex** for a set \( C \subset \mathbb{R}^n \),
  \[
  f : \mathbb{R}^n \to \mathbb{R}, \quad f(x) = \sup_{z \in C} \|x - z\|
  \]

- **Maximum eigenvalue function of a symmetric matrix is convex**
  \[
  \lambda_{\text{max}} : S^n \to \mathbb{R}, \quad \lambda_{\text{max}}(X) = \sup_{\|z\|=1} z^T X z
  \]
Composition with Scalar Functions

Composition of \( g : \mathbb{R}^n \to \mathbb{R} \) and \( h : \mathbb{R} \to \mathbb{R} \) with \( \text{dom}(g) = \mathbb{R}^n \) and \( \text{dom}(h) = \mathbb{R} \):

\[
f(x) = h(g(x))
\]

\( f \) is convex if

(1) \( g \) is convex, \( h \) is nondecreasing and convex

(2) \( g \) is concave, \( h \) is nonincreasing and convex

Examples

- \( e^{g(x)} \) is convex if \( g \) is convex

- \( \frac{1}{g(x)} \) is convex if \( g \) is concave and positive
Composition with Vector Functions

Composition of \( g : \mathbb{R}^n \rightarrow \mathbb{R}^p \) and \( h : \mathbb{R}^p \rightarrow \mathbb{R} \) with \( \text{dom}(g) = \mathbb{R}^n \) and \( \text{dom}(h) = \mathbb{R}^p \):

\[
f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_p(x))
\]

\( f \) is convex if

(1) each \( g_i \) is convex, \( h \) is convex and nondecreasing in each argument

(2) each \( g_i \) is concave, \( h \) is convex and nonincreasing in each argument

Example

• \( \sum_{i=1}^{m} e^{g_i(x)} \) is convex if \( g_i \) are convex
Extended-Value Functions

A function \( f \) is an extended-value function if \( f : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty, +\infty\} \)

Example: consider \( f(x) = \inf_{y \geq 0} xy \) for \( x \in \mathbb{R} \)

**Def.** The *epigraph* of a function \( f \) over \( \mathbb{R}^n \) is the following set in \( \mathbb{R}^{n+1} \):

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

**General Convex Function Def.** A function \( f \) is *convex* if its epigraph \( \text{epi} f \) is a convex set in \( \mathbb{R}^{n+1} \)

This definition is equivalent to the one we have used so far (when reduced to the function class we have considered thus far). How?

For an \( f \) with domain \( \text{dom} f \), we associate an extended-value function \( \tilde{f} \) defined by

\[
\tilde{f}(x) = \begin{cases} 
  f(x) & \text{if } x \in \text{dom} f \\
  +\infty & \text{otherwise}
\end{cases}
\]

\( \text{dom} f \) is the projection of \( \text{epi} f \) on \( \mathbb{R}^n \); convexity of \( f \) by letting \( w = f(x) \)
Minimization

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

Let $f : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ be a convex function [in $(x, z) \in \mathbb{R}^n \times \mathbb{R}^p$]. Then

$$g(x) = \inf_{z \in C} f(x, z) \quad \text{is convex}$$

Example

- Distance to a set: for a nonempty convex $C \subset \mathbb{R}^n$,
  $$\text{dist}(x, C) = \inf_{z \in C} \|x - z\| \quad \text{is convex}$$

Proof: Let $x_1, x_2 \in \mathbb{R}^n$ and $\alpha \in (0, 1)$ be arbitrary. Let $\epsilon > 0$ be arbitrarily small. Then, there exist $z_1, z_2 \in C$ such that $f(x_1, z_1) \leq g(x_1) + \epsilon$ and $f(x_2, z_2) \leq g(x_2) + \epsilon$. Consider $f(\alpha x_1 + (1 - \alpha)x_2, \alpha z_1 + (1 - \alpha)z_2)$ and use convexity of $f$ and $C$. 

Convex Optimization 20