10.007: Modeling the Systems World

Week 02 - Cohort 2

Convex Functions

Term 3, 2017

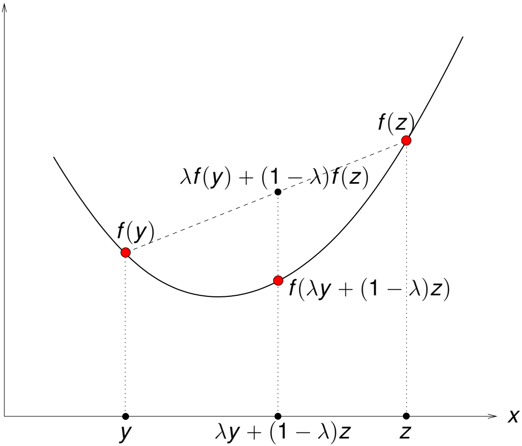
Singapore University of
Technology & Design

Learning outcomes

We will learn a lot about convexity today, which will be very useful later on!

- Definition of convex and concave functions
- Properties of convex functions
- Identfying convex functions (three ways)

Convex functions: intuition



A function is convex if the line between $(y, f(y)), (z, f(z))$ lies above the graph of the function over the segment $\overline{yz}$.

Convexity

Definition

A function $f(x) : \mathbb{R}^n \to \mathbb{R}$ is called convex if for every $y, z \in \mathbb{R}^n$ and every scalar $0 \leqslant \lambda \leqslant 1$ we have:

$$f(\lambda y + (1 - \lambda)z) \leqslant \lambda f(y) + (1 - \lambda)f(z).$$

It is called strictly convex if for every $y \neq z \in \mathbb{R}^n$ and every scalar $0 < \lambda < 1$ we have:

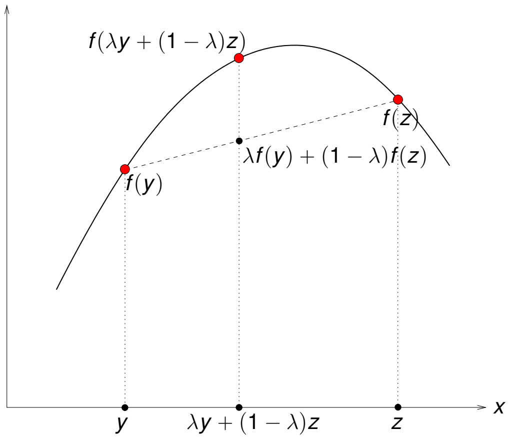
$$f(\lambda y + (1 - \lambda)z) < \lambda f(y) + (1 - \lambda)f(z).$$

The expression $\lambda y + (1 - \lambda)z$ is simply the line segment between y and z when λ ranges from 0 to 1.

Concavity

Definition

A function f is concave if $-f$ is convex.



Activity 1 (20 minutes)

(a) Using the definition of convexity/concavity, show that the linear function $f(x) = f(x_1, x_2) = a_1 x_1 + a_2 x_2$ is both convex and concave.

(b) Generalize part (a) to show that the linear function

$$f(x) = f(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_i$$

is both convex and concave.

(c) Is the product of two convex functions a convex function? Show that it is true or provide a counterexample.

Operations that preserve convexity

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^n \to \mathbb{R}, \ell : \mathbb{R}^m \to \mathbb{R}, h : \mathbb{R} \to \mathbb{R}$ be convex functions. Then the following functions of $x \in \mathbb{R}^n$ are convex:

1. $\ell(Ax + b)$, for all $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ (composition with affine function).
2. $af(x) + bg(x)$, for all $a, b \in \mathbb{R}, a, b \geqslant 0$ (sum of convex functions).
3. $\max\{f(x), g(x)\}$ (maximum of convex functions).
4. $h(f(x))$, for nondecreasing h (composition with convex nondecreasing function).

Example: $e^{x^2+y^2}$ is convex because:

- $f(x, y) = x^2$ and $g(x, y) = y^2$ are convex.
- $f(x, y) + g(x, y) = x^2 + y^2$ is convex by (2).
- $e^{x^2+y^2}$ is convex by (4).

Activity 2 (10 minutes)

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}, g : \mathbb{R}^n \to \mathbb{R}, \ell : \mathbb{R}^m \to \mathbb{R}, h : \mathbb{R} \to \mathbb{R}$ be convex functions. Then the following functions of $x \in \mathbb{R}^n$ are convex:

1. $\ell(Ax + b)$, for all $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$ (composition with affine function).
2. $af(x) + bg(x)$, for all $a, b \in \mathbb{R}, a, b \geqslant 0$ (sum of convex functions).
3. $\max\{f(x), g(x)\}$ (maximum of convex functions).
4. $h(f(x))$, for nondecreasing h (composition with convex nondecreasing function).

1. Prove that $f(x_1, x_2) = (x_1 - 2x_2)^4 + 2e^{(3x_1+2x_2-5)}$ is a convex function.
2. Is the sum of a convex and a nonconvex function always nonconvex?

Positive (semi)definiteness

Definition

A symmetric matrix $H \in \mathbb{R}^{n \times n}$ is positive semidefinite if $x^T Hx \geqslant 0$ for all vectors $x \in \mathbb{R}^n$. It is positive definite if $x^T Hx > 0$ for all nonzero vectors $x \in \mathbb{R}^n$. It is indefinite if $x^T Hx$ could be positive or negative depending on x.

Equivalent definition

A symmetric matrix $H \in \mathbb{R}^{n \times n}$ is positive semidefinite if all its eigenvalues are $\geqslant 0$. It is positive definite if all its eigenvalues are > 0. It is indefinite if it has both positive and negative eigenvalues.

The Hessian Test for Convexity

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with continuous second-order partial derivatives. Then f is convex over $\mathbb{R}^n$ if and only if its Hessian is positive semidefinite at all points in $\mathbb{R}^n$. If its Hessian is positive definite at all points in $\mathbb{R}^n$ then f is strictly convex over $\mathbb{R}^n$.

$$H_f(x_1, \ldots, x_n) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Remark: a function can be strictly convex while having second derivative equal to zero somewhere, e.g.: x^4.

On negative definiteness

Negative (semi)definiteness is the opposite of positive (semi)definiteness.

A symmetric matrix $H \in \mathbb{R}^{n \times n}$ is negative definite if $-H$ is positive definite. It is negative semidefinite if $-H$ is positive semidefinite.

A symmetric matrix could be positive (semi)definite, negative (semi)definite, or indefinite. It is indefinite if it has both positive and negative eigenvalues.

Proposition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function with continuous second-order partial derivatives. Then f is concave over $\mathbb{R}^n$ if and only if its Hessian is negative semidefinite at all points in $\mathbb{R}^n$. If its Hessian is negative definite at all points in $\mathbb{R}^n$ then f is strictly concave over $\mathbb{R}^n$.

Activity 3 (10 minutes)

Consider the functions:

(a) $f(x, y) = y^2$,

(b) $f(x, y) = -xy + y^2$,

(c) $f(x, y) = -xy + y^4$.

Compute their Hessians and the corresponding eigenvalues, and use this information to determine if the functions are convex over $\mathbb{R}^2$.

Activity 4 (15 minutes)

Recall the orange production problem with a supply/demand model from Cohort 1.1:

Decision variables: quantities Q_o, Q_f, Q_j respectively of oranges bought, oranges sold unprocessed, units of juice.

$$\left.\begin{array}{rrcl}
\max & P_f Q_f + P_j Q_j - P_o Q_o & & \\
\text{s.t.:} & & & \\
\text{(Availability)} & Q_f + 2Q_j - Q_o & \leqslant & 0 \\
 & Q_f, Q_j, Q_o & \geqslant & 0.
\end{array}\right\}$$

With prices given by:

$$\begin{aligned}
P_o &= 3 + 0.0005 Q_o \\
P_f &= 4 - 0.001 Q_f - 0.0002 Q_j \\
P_j &= 7.5 - 0.0002 Q_f - 0.005 Q_j.
\end{aligned}$$

Activity 4 (15 minutes)

The problem could be rewritten as:

$$\left.\begin{array}{rrcl}
\max & \multicolumn{3}{l}{4Q_f + 7.5Q_j - 3Q_o - 0.0004Q_jQ_f} \\
 & \multicolumn{3}{l}{-0.001Q_f^2 - 0.005Q_j^2 - 0.0005Q_o^2} \\
\text{s.t.:} & & & \\
\text{(Availability)} & Q_f + 2Q_J - Q_o & \leqslant & 0 \\
 & Q_f, Q_j, Q_o & \geqslant & 0.
\end{array}\right\}$$

- Is the objective function convex or concave?

Summary

- Definition of convex and concave functions.
- Three ways to show that a function is convex: Definition, operations that preserve convexity, Hessian test.

Remark: We will see next that unconstrained optimization is much easier when the objective function is convex.