Convex Functions

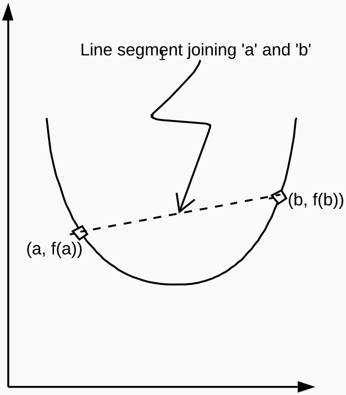
Nipun Batra

January 27, 2020

IIT Gandhinagar

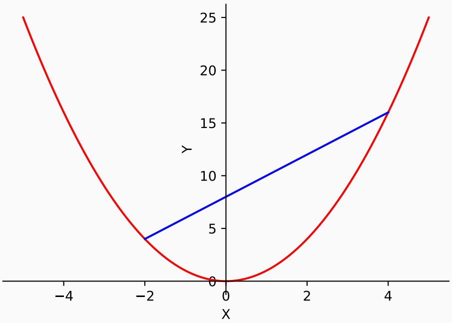
Definition

- Convexity is defined on an interval $[\alpha, \beta]$
- The line segment joining $(a, f(a)) and (b, f(b))$ should be *above or on* the function f for all points in interval $[\alpha, \beta]$.



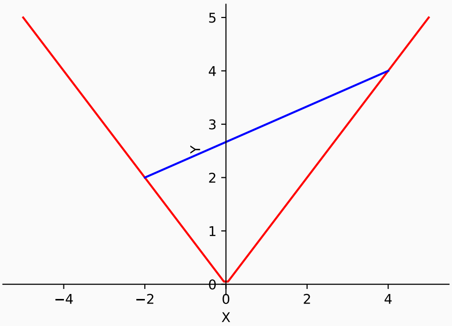
Example: $y = x^2$

Convex on the entire real line i.e. $(-\infty, \infty)$



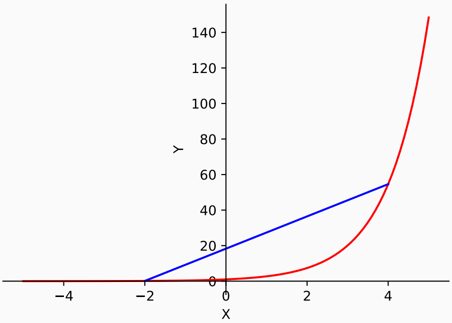
Example: $y = |x|$

Convex on the entire real line i.e. $(-\infty, \infty)$



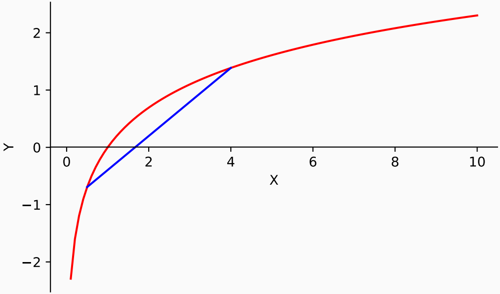
Example: $y = e^x$

Convex on the entire real line i.e. $(-\infty, \infty)$



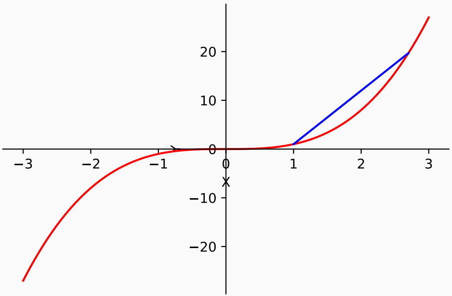
Example: $y = log_e x$

Not convex on the entire real line i.e. $(-\infty, \infty)$



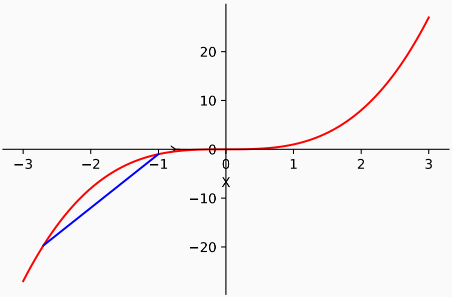
Example: $y = x^3$

It is convex for the interval $[0, \infty)$



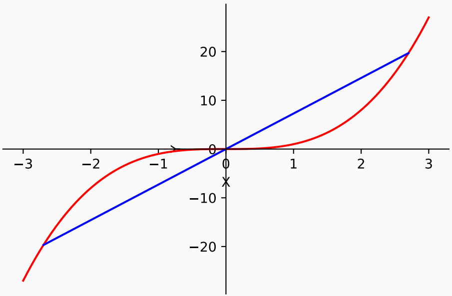
Example: $y = x^3$

It is concave for the interval $(-\infty, 0]$



Example: $y = x^3$

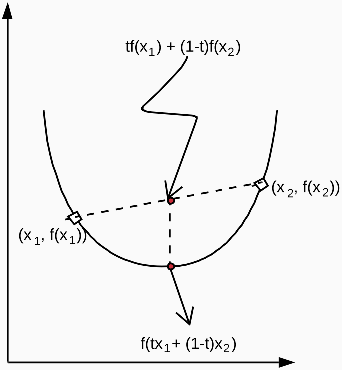
But, it is not convex for the interval $(-\infty, \infty)$



Mathematical Formulation

Function f is convex on set X, if $\forall x_1, x_2 \in X$ and $\forall t \in [0,1]$

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$



Question: Prove that $f(x) = x^2$ is convex

Question: Prove that $f(x) = x^2$ is convex

To prove:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

Question: Prove that $f(x) = x^2$ is convex

To prove:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

LHS = $f(tx_1 + (1-t)x_2)$ = $t^2x_1^2 + (1-t)^2x_2^2 + 2t(1-t)x_1x_2$
RHS = $tf(x_1) + (1-t)f(x_2)$ = $tx_1^2 + (1-t)x_2^2$

Question: Prove that $f(x) = x^2$ is convex

To prove:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

LHS = $f(tx_1 + (1-t)x_2)$ = $t^2x_1^2 + (1-t)^2x_2^2 + 2t(1-t)x_1x_2$
RHS = $tf(x_1) + (1-t)f(x_2)$ = $tx_1^2 + (1-t)x_2^2$

Here,
LHS - RHS = $(t^2 - t)x_1^2 + [(1-t)^2 - (1-t)]x_2^2 + 2t(1-t)x_1x_2$
= $(t^2 - t)x_1^2 + (t^2 - t)x_2^2 - 2(t^2 - t)x_1x_2$
= $(t^2 - t)(x_1 - x_2)^2$

Question: Prove that $f(x) = x^2$ is convex

To prove:

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

LHS = $f(tx_1 + (1-t)x_2)$ = $t^2x_1^2 + (1-t)^2x_2^2 + 2t(1-t)x_1x_2$
RHS = $tf(x_1) + (1-t)f(x_2)$ = $tx_1^2 + (1-t)x_2^2$

Here,
LHS - RHS = $(t^2 - t)x_1^2 + [(1-t)^2 - (1-t)]x_2^2 + 2t(1-t)x_1x_2$
= $(t^2 - t)x_1^2 + (t^2 - t)x_2^2 - 2(t^2 - t)x_1x_2$
= $(t^2 - t)(x_1 - x_2)^2$

Here, $(t^2 - t) \leq 0$ since $t \in [0, 1]$ and $(x_1 - x_2)^2 \geq 0$
Hence, LHS -RHS ≤ 0
Hence LHS $\leq$ RHS
Hence proved.

Alternative ways to prove convexity

The Double-Derivative Test

If f”(x) > 0, the function is convex.

For example,

$$\frac{\partial^2(x^2)}{\partial x^2} = 2 > 0 \Rightarrow x^2 \text{ is a Convex function.}$$

Alternative ways to prove convexity

The double derivate test for multi-parameter function is equal to using the Hessian Matrix

A function $f(x_1, x_2, \ldots, x_n)$ is convex iff its $n \times n$ Hessian Matrix is positive semidefinite for all posible values of $(x_1, x_2, \ldots, x_n)$

$$\mathsf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

Alternative ways to prove convexity

Show that $f(x_1, x_2) = x_1^2 + x_2^2$ is convex.

Alternative ways to prove convexity

Show that $f(x_1, x_2) = x_1^2 + x_2^2$ is convex.

$$H = \begin{bmatrix} \frac{\partial^2(x_1^2 + x_2^2)}{\partial x_1^2} & \frac{\partial^2(x_1^2 + x_2^2)}{\partial x_1 \partial x_2} \\ \frac{\partial^2(x_1^2 + x_2^2)}{\partial x_2 \partial x_1} & \frac{\partial^2(x_1^2 + x_2^2)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Alternative ways to prove convexity

Show that $f(x_1, x_2) = x_1^2 + x_2^2$ is convex.

$$H = \begin{bmatrix} \frac{\partial^2(x_1^2+x_2^2)}{\partial x_1^2} & \frac{\partial^2(x_1^2+x_2^2)}{\partial x_1 \partial x_2} \\ \frac{\partial^2(x_1^2+x_2^2)}{\partial x_2 \partial x_1} & \frac{\partial^2(x_1^2+x_2^2)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Eigen Values of H are 2 and $2 > 0 \Rightarrow$ H is positive semi-definite.
Hence, $f(x_1, x_2) = x_1^2 + x_2^2$ is convex.

Convexity of linear least squares

Prove the convexity of linear least squares i.e. $f(\theta) = ||y - X\theta||^2$

Convexity of linear least squares

Prove the convexity of linear least squares i.e. $f(\theta) = ||y - X\theta||^2$

We will use the double derivate (Hessian)

Convexity of linear least squares

Prove the convexity of linear least squares i.e. $f(\theta) = ||y - X\theta||^2$

We will use the double derivate (Hessian)

$$\frac{df}{d\theta} = \frac{d(||y^2|| - 2y^T X\theta + ||X\theta||^2)}{d\theta} = -2y^T X + 2(X\theta)^T X$$

Convexity of linear least squares

Prove the convexity of linear least squares i.e. $f(\theta) = ||y - X\theta||^2$

We will use the double derivate (Hessian)

$$\frac{df}{d\theta} = \frac{d(||y^2|| - 2y^T X\theta + ||X\theta||^2)}{d\theta} = -2y^T X + 2(X\theta)^T X$$

$$\frac{d^2 f}{d\theta^2} = H = 2X^T X$$

Convexity of linear least squares

Prove the convexity of linear least squares i.e. $f(\theta) = ||y - X\theta||^2$

We will use the double derivate (Hessian)

$$\frac{df}{d\theta} = \frac{d(||y^2|| - 2y^T X\theta + ||X\theta||^2)}{d\theta} = -2y^T X + 2(X\theta)^T X$$

$$\frac{d^2f}{d\theta^2} = H = 2X^T X$$

$X^T X$ is positive semi-definite for any $X \in \mathbb{R}^{m \times n}$.
Hence, linear least squares function is convex.

Properties of Convex Functions

- If $f(x)$ is convex, then $kf(x)$ is also convex, for some constant k

Properties of Convex Functions

- If $f(x)$ is convex, then $kf(x)$ is also convex, for some constant k
- If $f(x)$ and $g(x)$ are convex, then $f(x) + g(x)$ is also convex.

Properties of Convex Functions

- If $f(x)$ is convex, then $kf(x)$ is also convex, for some constant k
- If $f(x)$ and $g(x)$ are convex, then $f(x) + g(x)$ is also convex.

Using this we can say that,

- $(y - x\theta)^T(y - x\theta) + \theta^T\theta$ is convex
- $(y - x\theta)^T(y - x\theta) + |\theta|$ is convex