Optimization

Lusine Poghosyan

AUA

January 27, 2019

Uniqueness of solution

Definition

A set $\Omega \subset \mathbb{R}^n$ is a **convex set** if $\alpha x + (1-\alpha)y \in \Omega$, $\forall x, y \in \Omega$ and $\forall \alpha \in [0, 1]$.

Example

Check if the set Ω is convex, if

a. $\Omega = \{x = [x_1, x_2]^T : x_1^2 + x_2^2 \leq 1\}$;

b. $\Omega = \{x = [x_1, x_2]^T : x_1^2 + x_2^2 \leq 1 \quad \text{and} \quad x_1 x_2 \geq 0\}$.

Definition

A function $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$ is a **convex function** if Ω is a convex set and $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y)$, $\forall x, y \in \Omega$, $x \neq y$ and $\forall \alpha \in (0,1)$.

If in the definition above we replace "$\leq$" with "$<$", then we have the definition of **strictly convex function.**

Definition

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

Definition

A function f is a **strictly concave function** if $-f$ is strictly convex.

Example

Show that the linear function $f(x) = a^T x + b$, where $a, x \in \mathbb{R}^n$ and $b \in \mathbb{R}$, is a convex function.

Example

Let $f(x) = x_1^2 + x_2^2 + ... + x_n^2$, $x \in \mathbb{R}^n$. Show that f is strictly convex function.

Theorem

Let $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$. If f is a convex function and x_0 is a local minimum point of f over Ω, then x_0 is a global minimum point of f in Ω, i.e. $x_0 = arg\ min_{x \in \Omega} f(x)$.

Theorem

Let $f : \Omega \to \mathbb{R}$, $\Omega \subset \mathbb{R}^n$. If f is a strictly convex function and x_0 is a local minimum point of f over Ω, then x_0 is the unique global minimum point of f on Ω, i.e. $x_0 = arg\ min_{x \in \Omega} f(x)$.

Convexity by using the first order derivative

Let's denote by $\nabla f(x)$ the gradient of f at x i.e.

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T.$$

Example

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f(x_1, x_2, x_3) = x_1^3 + 4x_2^2 + \cos(x_3^2 x_2)$.
Compute the gradient $\nabla f(x_1, x_2, x_3)$.

Theorem

If $\Omega \subset \mathbb{R}^n$ is an open, convex set and $f \in \mathbb{C}^1(\Omega)$, then f is convex if and only if

$$f(x) \geq f(x_0) + \nabla f(x_0)^T (x - x_0), \quad \forall x, x_0 \in \Omega.$$

Theorem

If $\Omega \subset \mathbb{R}^n$ is an open, convex set and $f \in \mathbb{C}^1(\Omega)$, then f is strictly convex if and only if

$$f(x) > f(x_0) + \nabla f(x_0)^T (x - x_0), \quad \forall x, x_0 \in \Omega, x \neq x_0.$$

Theorem

If $\Omega \subset \mathbb{R}^n$ is an open convex set and $f \in \mathbb{C}^1(\Omega)$, then f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0, \quad \forall x, y \in \Omega.$$

Theorem

If $\Omega \subset \mathbb{R}^n$ is an open convex set and $f \in \mathbb{C}^1(\Omega)$, then f is strictly convex if and only if

$$(\nabla f(x) - \nabla f(y))^T (x - y) > 0, \quad \forall x, y \in \Omega, x \neq y.$$

Convexity by using the second order derivative

Theorem

Let $f : (a, b) \to \mathbb{R}$. If f is twice differentiable function on (a, b) then f is convex if and only if $f''(x) \geq 0$, for all $x \in (a, b)$.

Theorem

Let $f : (a, b) \to \mathbb{R}$. If f is twice differentiable function on (a, b) and $f''(x) > 0$, for all $x \in (a, b)$, then f is strictly convex.

Example

Check if the following functions are convex (strictly convex), concave (strictly concave), if

a. $\frac{1}{1+x^2}, \quad x \in \mathbb{R};$

b. $\cos x, \quad x \in \left(0, \frac{\pi}{2}\right);$

c. $(x+2)^6, \quad x \in \mathbb{R}.$

Definition

Assume $f : \Omega \to \mathbb{R}$, $\Omega \to \mathbb{R}^n$ and $x_0 \in \Omega$. If all second order partial derivatives of f exist at x_0, then the following matrix

$$\nabla^2 f(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x_0) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x_0) \\ \vdots & & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{pmatrix}.$$

is called the Hessian matrix of f at x_0.

If second order partial derivatives of f are all continuous at x_0 then Hessian matrix is symmetric.

Example

Let $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by $f(x_1, x_2, x_3) = x_1^3 + 4x_2^2 x_1 + \sin(x_3)$.
Compute the Hessian matrix $\nabla^2 f(x_0)$ at $x_0 = (1, 1, 0)^T$.