Optimization

Lusine Poghosyan

YSU

October 10, 2020

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 $\nabla f(x_1, x_2, x_3)$ at $x = [-2, 0, 0]^T$.

Example

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = a^T x$, where $a, x \in \mathbb{R}^n$. Compute the gradient $\nabla f(x)$.

Example

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = x^T Ax$, where A is $n \times n$ symmetric matrix. Compute the gradient $\nabla f(x)$.

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

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

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

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

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

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

is called the Hessian matrix of f at x_0.

If all second order partial derivatives of f are 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 $x_0 = (1, 1, 0)^T$.

Example

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = a^T x$, where $a, x \in \mathbb{R}^n$. Compute the Hessian matrix $\nabla^2 f(x)$.

Example

Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined by $f(x) = x^T A x$, where A is $n \times n$ symmetric matrix. Compute the Hessian matrix $\nabla^2 f(x)$.

Definition

Assume $A = \left[a_{ij}\right]_{i,j=1}^{n}$ is an $n \times n$ symmetric matrix, i.e. $a_{ij} = a_{ji}$. A function $QF_A : \mathbb{R}^n \to \mathbb{R}$ is called a **quadratic form** associated to the matrix A if

$$QF_A(y) = y^T A y = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} y_i y_j.$$

Example

Construct the quadratic form associated to the matrix A if

$$A = \begin{pmatrix} 4 & 3 & 1 \\ 3 & 5 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

Definition

We will say that the symmetric $n \times n$ matrix A or the quadratic form QF_A is

- **positive definite** if $QF_A(y) > 0$, $\forall y \in \mathbb{R}^n$ and $y \neq 0$;
- **positive semidefinite** if $QF_A(y) \geq 0$, $\forall y \in \mathbb{R}^n$;
- **negative definite** if $QF_A(y) < 0$, $\forall y \in \mathbb{R}^n$ and $y \neq 0$;
- **negative semidefinite** if $QF_A(y) \leq 0$, $\forall y \in \mathbb{R}^n$;
- **indefinite** if there exist $y_1, y_2 \in \mathbb{R}^n$ such that $QF_A(y_1) > 0$ and $QF_A(y_2) < 0$.

Note. If A is positive (negative) definite, then it is also positive (negative) semidefinite.

Example

Determine whether the matrix A is positive definite (semidefinite), negative definite (semidefinite) or indefinite if

a.

$$A = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 5 & 0 \\ 0 & 0 & 2 \end{pmatrix};$$

$$QF_A(y) = y^T A y = 4y_1^2 + 5y_2^2 + 2y_3^2 + 4y_1 y_2 = (2y_1 + y_2)^2 + 4y_2^2 + 2y_3^2$$

$$QF_A(y) \geq 0, \quad \forall y \in \mathbb{R}^3, \quad \text{therefore} \quad A \succeq 0.$$

Here we can show that $A \succ 0$. To do that we need to solve $QF_A(y) = 0$ and get $y = [0, 0, 0]^T$.