Unconstrained optimization

- Non linear objective function

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in \Omega, \end{array}$$

where f is a real-valued function and Ω, the feasible set, is a subset of E^n

  - Minimization of single valued function : some discussion

Local and global minimum

- Local minimum or relative minimum

Definition. A point $\mathbf{x}^* \in \Omega$ is said to be a *relative minimum point* or a *local minimum point* of f over Ω if there is an $\varepsilon > 0$ such that $f(\mathbf{x}) \geqslant f(\mathbf{x}^*)$ for all $\mathbf{x} \in \Omega$ within a distance ε of $\mathbf{x}^*$ (that is, $\mathbf{x} \in \Omega$ and $|\mathbf{x} - \mathbf{x}^*| < \varepsilon$). If $f(\mathbf{x}) > f(\mathbf{x}^*)$ for all $\mathbf{x} \in \Omega$, $\mathbf{x} \neq \mathbf{x}^*$, within a distance ε of $\mathbf{x}^*$, then $\mathbf{x}^*$ is said to be a *strict relative minimum point* of f over Ω.

- Global minimum

Definition. A point $\mathbf{x}^* \in \Omega$ is said to be a *global minimum point* of f over Ω if $f(\mathbf{x}) \geqslant f(\mathbf{x}^*)$ for all $\mathbf{x} \in \Omega$. If $f(\mathbf{x}) > f(\mathbf{x}^*)$ for all $\mathbf{x} \in \Omega$, $\mathbf{x} \neq \mathbf{x}^*$, then $\mathbf{x}^*$ is said to be a *strict global minimum point* of f over Ω.

- We generally focus on local minimum points
- Under certain conditions local minimum is also the global minimum

First order necessary conditions

- Feasible directions

Given $\mathbf{x} \in \Omega$, a vector $\mathbf{d}$ is a *feasible direction at* $\mathbf{x}$ if there is an $\bar{\alpha} > 0$ such that $\mathbf{x} + \alpha\mathbf{d} \in \Omega$ for all α, $0 \leqslant \alpha \leqslant \bar{\alpha}$.

Proposition 1. (First-order necessary conditions). *Let Ω be a subset of E^n and let $f \in C^1$ be a function on Ω. If $\mathbf{x}^*$ is a relative minimum point of f over Ω, then for any $\mathbf{d} \in E^n$ that is a feasible direction at $\mathbf{x}^*$, we have $\nabla f(\mathbf{x}^*)\mathbf{d} \geqslant 0$.*

Proof: Taylor's expansion

Corollary. (Unconstrained case). *Let Ω be a subset of E^n, and let $f \in C^1$ be function'on Ω. If $\mathbf{x}^*$ is a relative minimum point of f over Ω and if $\mathbf{x}^*$ is an interior point of Ω, then $\nabla f(\mathbf{x}^*) = 0$.*

n equations, n unknowns: can be solved to find the minimum point

Example

1

Example 1. Consider the problem

$$\text{minimize } f(x_1,\ x_2) = x_1^2 - x_1x_2 + x_2^2 - 3x_2.$$

2

Example 2. Consider the problem

$$\begin{array}{ll} \text{minimize} & f(x_1,\ x_2) = x_1^2 - x_1 + x_2 + x_1x_2 \\ \text{subject to} & x_1 \geqslant 0, \qquad x_2 \geqslant 0. \end{array}$$

This problem has a global minimum at $x_1 = \frac{1}{2},\ x_2 = 0.$

Second order conditions

Proposition 1. (Second-order necessary conditions). *Let Ω be a subset of E^n and let $f \in C^2$ be a function on Ω. If $\mathbf{x}^*$ is a relative minimum point of f over Ω, then for any $\mathbf{d} \in E^n$ that is a feasible direction at $\mathbf{x}^*$ we have*

$$\text{i) } \nabla f(\mathbf{x}^*)\mathbf{d} \geqslant 0 \tag{3}$$

$$\text{ii) if } \nabla f(\mathbf{x}^*)\mathbf{d} = 0, \text{ then } \mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d} \geqslant 0. \tag{4}$$

Proof:

Example 1. For the same problem as Example 2 of Section 7.1, we have for $\mathbf{d} = (d_1, d_2)$

$$\nabla f(\mathbf{x}^*)\mathbf{d} = \frac{3}{2} d_2.$$

Unconstrained case

Proposition 2. (Second-order necessary conditions—unconstrained case).

Let $\mathbf{x}^*$ be an interior point of the set Ω, and suppose $\mathbf{x}^*$ is a relative minimum point over Ω of the function $f \in C^2$. Then

$$\text{i) } \nabla f(\mathbf{x}^*) = 0 \tag{5}$$

$$\text{ii) for all } \mathbf{d},\ \mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d} \geqslant 0. \tag{6}$$

Hessian matrix $\nabla^2 f(\mathbf{x})$, **F**(x) positive semidefinite matrix

Example :

$$\begin{array}{ll} \text{minimize} & f(x_1,\ x_2) = x_1^3 - x_1^2 x_2 + 2x_2^2 \\ \text{subject to} & x_1 \geqslant 0, \quad x_2 \geqslant 0. \end{array}$$

Sufficient conditions

- A

Proposition 3. (Second-order sufficient conditions—unconstrained case).

Let $f \in C^2$ be function defined on a region in which the point $\mathbf{x}^*$ is an interior point. Suppose in addition that

$$\text{i) } \nabla f(\mathrm{x}^*) = \mathbf{0} \tag{7}$$

$$\text{ii) } \mathbf{F}(\mathbf{x}^*) \text{ is positive definite} \tag{8}$$

Then $\mathbf{x}^*$ is a strict relative minimum point of f.

Proof. Since $\mathbf{F}(\mathbf{x}^*)$ is positive definite, there is an $a > 0$ such that for all $\mathbf{d}$, $\mathbf{d}^T\mathbf{F}(\mathbf{x}^*)\mathbf{d} \geqslant a|\mathbf{d}|^2$. Thus by the Taylor's Theorem (with remainder)

Convex and concave functions

- Some restrictions on the function *f* is required to characterize global minimum

- Convex

Definition. A function f defined on a convex set Ω is said to be *convex* if, for every $\mathbf{x}_1, \mathbf{x}_2 \in \Omega$ and every α, $0 \leqslant \alpha \leqslant 1$, there holds

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \leqslant \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2).$$

If, for every α, $0 < \alpha < 1$, and $\mathbf{x}_1 \neq \mathbf{x}_2$, there holds

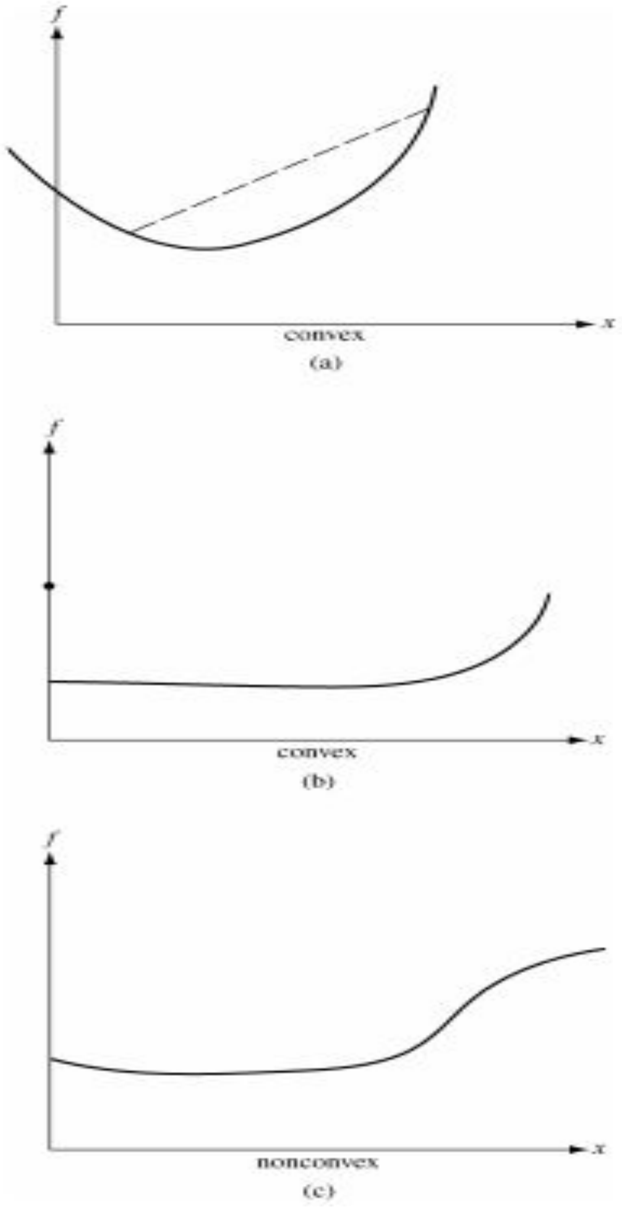
$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) < \alpha f(\mathbf{x}_1) + (1-\alpha)f(\mathbf{x}_2),$$

then f is said to be *strictly convex*.

- Concave

Definition. A function g defined on a convex set Ω is said to be *concave* if the function $f = -g$ is convex. The function g is *strictly concave* if $-g$ is strictly convex.

Examples



Combination of convex functions

- Properties

Proposition 1. *Let f_1 and f_2 be convex functions on the convex set Ω. Then the function $f_1 + f_2$ is convex on Ω.*

Proposition 2. *Let f be a convex function over the convex set Ω. Then the function af is convex for any $a \geqslant 0$.*

Proposition 3. *Let f be a convex function on a convex set Ω. The set $\Gamma_c = \{\mathbf{x} : \mathbf{x} \in \Omega,\ f(\mathbf{x}) \leqslant c\}$ is convex for every real number c.*

Differentiable convex functions

Characterization of convex function using differentials

Proposition 4. *Let $f \in C^1$. Then f is convex over a convex set Ω if and only if*

$$f(\mathbf{y}) \geqslant f(\mathbf{x}) + \nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \tag{9}$$

for all $\mathbf{x},\ \mathbf{y} \in \Omega$.

Proposition 5. *Let $f \in C^2$. Then f is convex over a convex set Ω containing an interior point if and only if the Hessian matrix* $\mathbf{F}$ *of f is positive semidefinite throughout Ω.*

Minimization of convex functions

Local minimum is also global minimum

Theorem 1. *Let f be a convex function defined on the convex set Ω. Then the set Γ where f achieves its minimum is convex, and any relative minimum of f is a global minimum.*

Proof: Assuming c_0 is the minimum of f.

$$\Gamma = \{\mathbf{x} : f(\mathbf{x}) \leqslant c_0,\ \mathbf{x} \in \Omega\}$$

Theorem 2. *Let $f \in C^1$ be convex on the convex set Ω. If there is a point $\mathbf{x}^* \in \Omega$ such that, for all $\mathbf{y} \in \Omega$, $\nabla f(\mathbf{x}^*)(\mathbf{y} - \mathbf{x}^*) \geqslant 0$, then $\mathbf{x}^*$ is a global minimum point of f over Ω.*