2025 年数理经济学笔记

授课: 杨佳楠老师

作者：徐靖

组织：PKU

时间：Febuary 27, 2025

Pkuhub.cn

个人笔记, 如有谬误, 欢迎指正! 联系方式: 2200012917@stu.pku.edu.cn

目录

第 1 章 Multi-Variable Unconstrained Optimization

Keywords

- First Order Condition 一阶条件
- Bisection Method 二分法
- Secant Method 割线法
- False Position Method 假位法
- Newton's Method 牛顿法

1.1 First Order Condition

An Unconstrained Optimization Problem is :

$$\min_{x \in \mathbb{R}^n} f(x)$$

定义 1.1

First Order Condition (FOC): $\nabla f(x^*) = 0$.

- x^* is a **stationary point** (驻点) of f.
- It is necessary but not sufficient.

global minimum: $f(x^*) \le f(x)$ for all $x \in \mathbb{R}^n$.

local minimum: $f(x^*) \le f(x)$ for all $x \in B(x^*, \epsilon)$ for some $\epsilon > 0$.

♣

命题 1.1

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuously differentiable function, and $\nabla f(x^*) = 0$. If $\nabla^2 f(x^*)$ is:

- positive definite, then x^* is a local minimum.
- negative definite, then x^* is a local maximum.
- indefinite, then x^* is a **saddle point**. (鞍点)

♠

1.2 Convex Optimization

定义 1.2 (Convex Function)

A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if for all $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$:

$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y)$$

♣

定理 1.1

A twice continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if and only if its Hessian matrix $\nabla^2 f(x)$ is positive semidefinite for all $x \in \mathbb{R}^n$.

♡

命题 1.2

Let f be differentiable. Then f is (strictly) convex if and only if:

$$f(y) - f(x)(>) \ge \nabla f(x) \cdot (y - x)$$

for all $x, y \in \mathbb{R}^n$.

♠

定理 1.2 (Minimum/maximum Characterization)

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex (concave) function. Then x^* is a local minimum (maximum) if and only if:

$$\nabla f(x^*) = 0$$

- If f is strictly **convex**, then x^* is a global minimum.
- If f is strictly **concave**, then x^* is a global maximum.

♡

1.3 Numerical Optimization

1.3.1 Bisection Method

定义 1.3 (Bisection Method)

A simple, robust method for finding roots of continuous functions on bounded intervals

- Start with an interval $[a, b]$ such that $f(a)f(b) < 0$.
- Compute the midpoint $c = \frac{a+b}{2}$, and evaluate $f(c)$.
- Replace a or b with c based on the sign of $f(c)$.
- Iterate until desired precision.

♣

定义 1.4 (Convergence Rate and Order 收敛速度和阶)

For iteration x_n approaching the root r, the convergence rate C and order ρ are defined as:

$$\lim_{n\to\infty} \frac{|x_{n+1} - r|}{|x_n - r|^\rho} = C$$

- Linear convergence: $\rho = 1$, $C < 1$.
- Quadratic convergence: $\rho = 2$, $C < 1$.
- Superlinear convergence: $\rho > 1$, $C < 1$.

♣

Method	Definition	Rate	Order
Bisection	Iteratively bisects an interval and selects a subinterval	Linear ($C = 0.5$)	1
Secant	Root approximation via secant line through two points	Superlinear ($C \approx 1.618$)	1.618
False Position	Bisection variant with linear interpolation updates	Linear	1
Newton-Raphson	Derivative-based iterative root-finding	Quadratic ($C \propto f''$)	2
Gradient method	Function minimization via negative gradient direction	Linear ($C \propto \kappa$)	1

表 1.1: Compact Comparison of Numerical Methods

定义 1.5 (Methods)

Secant Method:

- Compute the secant line through $(x_0, f(x_0))$ and $(x_1, f(x_1))$.
- Find the intersection with the x-axis to get the next approximation x_2.
- Iterate until convergence.

$$x_{n+1} = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} \tag{1.1}$$

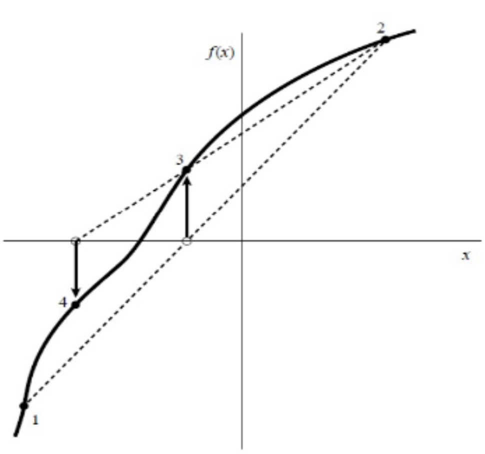

图 1.1: Secant Method

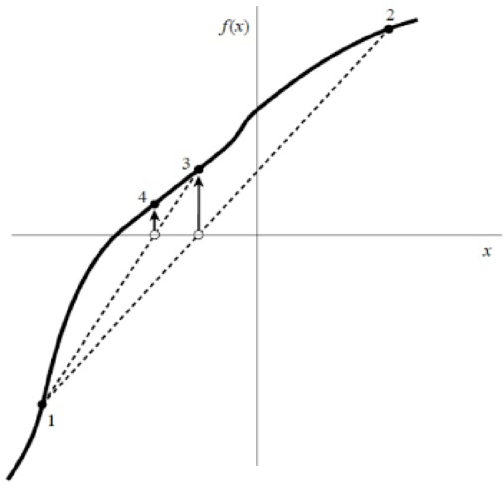

图 1.2: False Position Method

False Position Method:

- Similar to the secant method, but always keeps the interval $[a, b]$ such that $f(a)f(b) < 0$.
- Update a or b based on the sign of $f(c)$.
- Iterate until convergence.

$$c = \frac{af(b) - bf(a)}{f(b) - f(a)},$$

$$[a, b] \leftarrow [a, c] \quad \text{if} \quad f(a)f(c) < 0,$$

$$[a, b] \leftarrow [c, b] \quad \text{if} \quad f(b)f(c) < 0.$$

♣

笔记 若初始值足够接近根且函数光滑, 则 Secant Method 收敛速度优于 False Position Method, 但可能因迭代点跳出根的邻域而发散. False Position Method 保证收敛, 但多一个异号的初始条件且速度较慢.

定义 1.6 (Newton-Raphson Method)

- Start with an initial guess x_0.
- Compute the next approximation using the formula:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

- Iterate until convergence.

♣

问题 1.1 为什么牛顿法是二阶收敛的?

解 对 $f(x)$ 在 x_n 处做泰勒展开, 对于 $f(r) = 0$:

$$f(r) = f(x_n) + f'(x_n)(r - x_n) + \frac{f''(x_n)}{2}(r - x_n)^2 + O((x - x_n)^3)$$

带入牛顿法迭代公式 $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$:

$$x_{n+1} - r = (x_n - r)^2 \cdot \frac{f''(x_n)}{2f'(x_n)}$$

笔记 牛顿法初期可能出问题, 如果不满足足够接近根的假设.

牛顿法可以很好地应用到多变量上, 但过程中 Hessian 矩阵的逆矩阵计算量较大, 并且他是一个 local method.

例题 1.1 将牛顿法应用到求解二次可微函数的极值问题, 可以求解 first order condition:

$$x_{n+1} = x_n - [\nabla^2 f(x_n)]^{-1} \nabla f(x_n)$$

定义 1.7 (Gradient Method)

- Start with an initial guess x_0 and error tolerance ϵ.
- Iterate until $\|x_{n+1} - x_n\| < \epsilon$:
    - Compute the gradient $\nabla f(x_n)$.
    - Define $\phi(t) = f(x_n - t\nabla f(x_n))$.
    - Find the minimum of $\phi(t)$ using a one-variable optimization method (e.g., bisection, secant, or Newton's method).
    - Defint $x_{n+1} = x_n - t^*\nabla f(x_n)$.

♣