Optimization

1 Preliminaries

1.1 Formulation

$$\max_{\mathbf{x} \in G(\boldsymbol{\theta})} f(\mathbf{x}, \boldsymbol{\theta})$$

where typically the feasible set is defined by

$$G(\boldsymbol{\theta}) = \{\mathbf{x} \in X : g_j(\mathbf{x}, \boldsymbol{\theta}) \leq 0,\ j = 1, 2, \ldots m\}$$

▷ Note: $\min f = -\max(-f)$

▷ Functional v. set constraints

For example, nonnegativity $\mathbf{x} \geq \mathbf{0}$ can be expressed as

- $-\mathbf{x} \leq \mathbf{0}$ or
- $X = \Re^n_+$

1.2 Basic questions

▷ existence (Weierstrass Theorem 2.2)

▷ computation

▷ characterization

▷ sensitivity (comparative statics)

1.3 Terminology

▷ local v. global

If f is concave, every local optimum is a global optimum (Exercise 5.2)

▷ strict (unique) v. nonstrict

If f is strictly quasiconcave, every optimum is a strict global optimum (Exercise 5.3)

▷ interior v. boundary

▷ necessary v. sufficient conditions

2 Unconstrained optimization

$$\max_{\mathbf{x} \in X} f(\mathbf{x})$$

2.1 Basic first-order condition (Proposition 5.1)

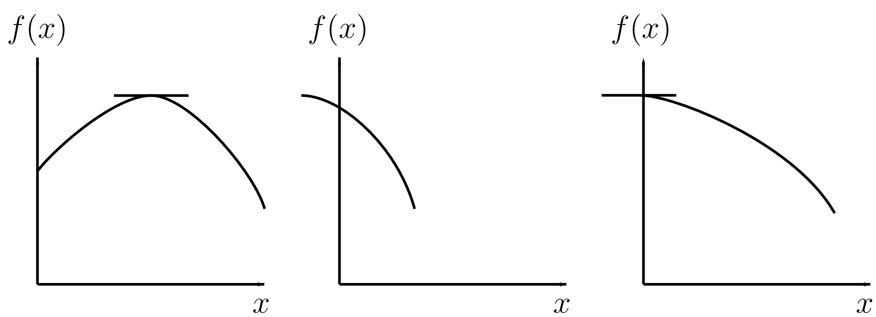
▷ If $\mathbf{x}^*$ is a local maximum of f, there exists a neighborhood S of $\mathbf{x}^*$ such that

$$f(\mathbf{x}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq f(\mathbf{x}^*) \text{ for every } \mathbf{x} \in S$$

which implies

- $\nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq 0$ for every $\mathbf{x} \in S$
- $\nabla f(\mathbf{x}^*) = \mathbf{0}$ if $\mathbf{x}^*$ is an interior point
  - That is, $D_{x_i} f[\mathbf{x}^*] = \frac{\partial f(\mathbf{x}^*)}{\partial x_i} = 0$ for every i

▷ A maximum must be **either** a stationary point of f **or** a boundary point of X **or both**.



Stationary point | Boundary point | Both

2.2 Nonnegative variables(Corollary 5.0.1)

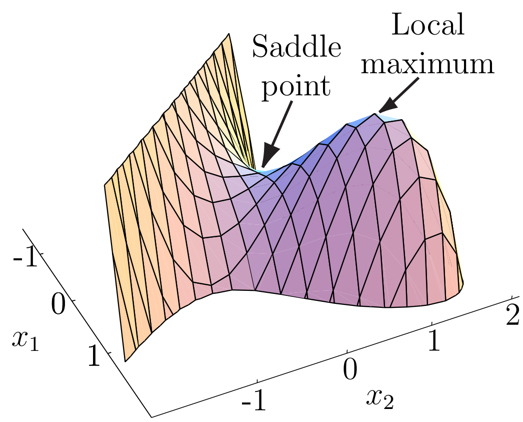
▷ For the nonnegativity constraint $\mathbf{x} \geq \mathbf{0}$, we can collapse the necessary conditions into

$$\nabla f(\mathbf{x}^*) \leq \mathbf{0} \qquad \mathbf{x}^* \geq 0 \qquad \nabla f(\mathbf{x}^*)^T \mathbf{x}^* = 0$$

2.3 Basic second-order condition (interior maximum)

▷

$$\max_{x_1,x_2} f(x_1, x_2) = 3x_1x_2 - x_1^3 - x_2^3$$



▷ If $\mathbf{x}^*$ is a local maximum of f, there exists a neighborhood S of $\mathbf{x}^*$ such that

$$f(\mathbf{x}) \approx f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^T \mathbf{dx} + \mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx} \leq f(\mathbf{x}^*)$$

or

$$\nabla f(\mathbf{x}^*)^T \mathbf{dx} + \mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx} \leq 0 \text{ for every } \mathbf{x} \in S$$

which requires that

- $\nabla f(\mathbf{x}^*)^T \mathbf{dx} = 0$
- $\mathbf{dx}^T H_f(\mathbf{x}^*)\mathbf{dx} \leq 0$ for every $\mathbf{x} \in S$.

2.4 Necessary conditions for an interior maximum

▷ For $\mathbf{x}^*$ to be an interior local maximum of $f(\mathbf{x})$ in X, it is necessary that

1. $\mathbf{x}^*$ be a stationary point of f, that is $\nabla f(\mathbf{x}^*) = 0$ and
2. f be locally concave at $\mathbf{x}^*$, that is $H_f(\mathbf{x}^*)$ is nonpositive definite.

▷ Examples

$$\begin{array}{lll} f(x) = x^2 & f'(0) = 0 & f''(0) = 2 \\ f(x) = x^3 & f'(0) = 0 & f''(0) = 0 \end{array}$$

2.5 Sufficient conditions for an interior maximum

▷ If

1. $\mathbf{x}^*$ is a stationary point of f, that is $\nabla f(\mathbf{x}^*) = 0$ and
2. f is locally strictly concave at $\mathbf{x}^*$, that is $H_f(\mathbf{x}^*)$ is negative definite

then $\mathbf{x}^*$ is a strict local maximum of f.

▷ Examples

$$
\begin{array}{lll}
f(x) = -x^2 & f'(0) = 0 & f''(0) = -2 \\
f(x) = -x^4 & f'(0) = 0 & f''(0) = 0
\end{array}
$$

2.6 Concave maximization (Corollary 5.1.2)

▷ Suppose f is concave and $\mathbf{x}^*$ is an interior point of X. Then $\mathbf{x}^*$ is a global maximum of f on X if and only if $\nabla f(\mathbf{x}^*) = \mathbf{0}$.

2.7 Some examples

▷ The competitive firm (Example 5.11)

▷ Monopoly (Example 5.12)

▷ Least squares regression (Example 5.13)

▷ Maximum likelihood estimation (Exercise 5.13)

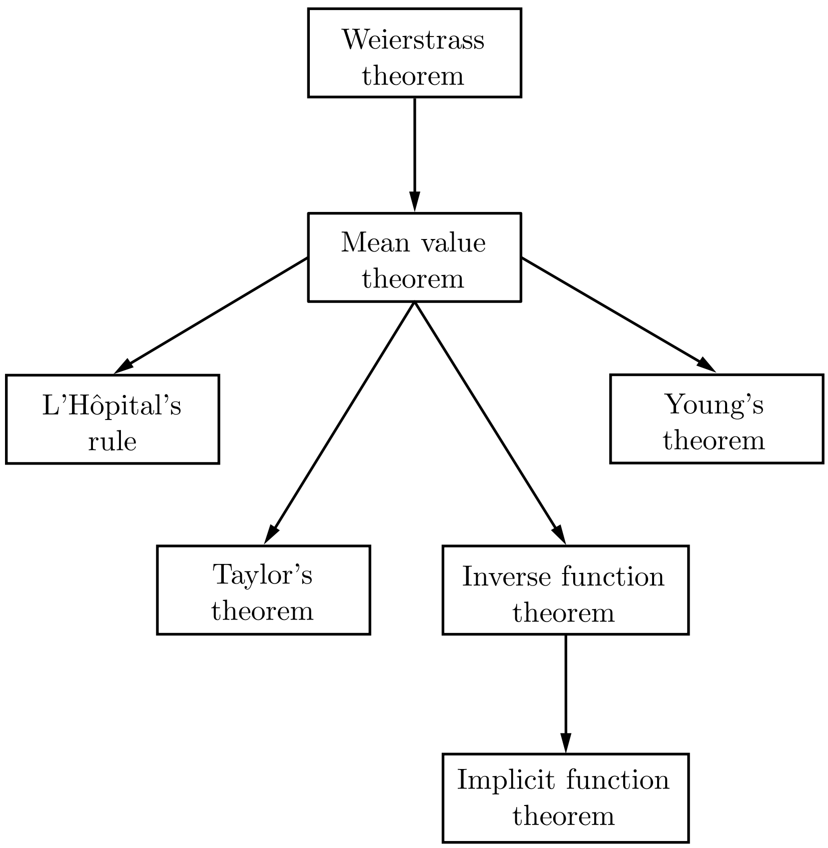
3 Homework

1. If f is concave and $G(\boldsymbol{\theta})$ is convex, every local optimum is a global optimum (Exercise 5.2).

2. Conjecture: If $f\colon \Re \to \Re$ has a local maximum at x^* which is not a strict local maximum, then f is constant in some neighbourhood of x. Prove or provide a counterexample.

3. **Rolle's theorem**. Suppose $f \in C[a, b]$ is differentiable on (a, b). If $f(a) = f(b)$, then there exists $x \in (a, b)$ where $f'(x) = 0$ (Exercise 5.8).

   Note: Rolle's theorem $\implies$ the mean value theorem, which provides all the useful theorems of calculus.



4. Solve the problem $\max_{x_1,x_2} f(x_1, x_2) = 3x_1x_2 - x_1^3 - x_2^3$ (Example 5.8).

5. Solve the problem $\max_{x_1,x_2} f(x_1, x_2) = x_1x_2 + 3x_2 - x_1^2 - x_2^2$ (Exercise 5.9)

6. Extend Corollary 5.1.2 to f pseudoconcave.

Solutions 2

1 Suppose that $\mathbf{x}^*$ is a local optimum which is not a global optimum. That is, there exists a neighborhood S of $\mathbf{x}^*$ such that

$$f(\mathbf{x}^*, \boldsymbol{\theta}) \geq f(\mathbf{x}, \boldsymbol{\theta}) \text{ for every } \mathbf{x} \in S \cap G(\boldsymbol{\theta})$$

and also another point $\mathbf{x}^{**} \in G(\boldsymbol{\theta})$ such that

$$f(\mathbf{x}^{**}, \boldsymbol{\theta}) > f(\mathbf{x}^*, \boldsymbol{\theta})$$

Since $G(\boldsymbol{\theta})$ is convex, there exists $\alpha \in (0,1)$ such that

$$\alpha \mathbf{x}^* + (1-\alpha)\mathbf{x}^{**} \in S \cap G(\boldsymbol{\theta})$$

By concavity of f

$$f(\alpha \mathbf{x}^* + (1-\alpha)\mathbf{x}^{**}, \boldsymbol{\theta}) \geq \alpha f(\mathbf{x}^*, \boldsymbol{\theta}) + (1-\alpha) f(\mathbf{x}^{**}, \boldsymbol{\theta}) > f(\mathbf{x}^*, \boldsymbol{\theta})$$

contradicting the assumption that $\mathbf{x}^*$ is a local optimum.

2 False. A counterexample is

$$f(x) = \begin{cases} x & x \leq 1 \\ 1 & x > 1 \end{cases}$$

f is constant on $[1, \infty)$, but this is not a neighbourhood of 1.

3 By the Weierstrass theorem (Theorem 2.2), f has a maximum x^* and a minimum x_* on $[a,b]$. Either

- $x^* \in (a,b)$ and $f'(x^*) = 0$ (Theorem 5.1) or
- $x_* \in (a,b)$ and $f'(x^*) = 0$ (Exercise 5.7) or
- Both maxima and minima are boundary points, that is $x^*, x_* \in \{a,b\}$ which implies that f is constant on $[a,b]$ and therefore $f'(x) = 0$ for every $x \in (a,b)$ (Exercise 4.7).

4 The first-order conditions for a maximum are

$$D_1 f = 3x_2 - 3x_1^2 = 0$$
$$D_2 f = 3x_1 - 3x_2^2 = 0$$

or

$$x_2 = x_1^2 \qquad x_1 = x_2^2$$

These equations have two solutions: $x_1 = x_2 = 0$ and $x_1 = x_2 = 1$. Therefore $(0,0)$ and $(1,1)$ are the only stationary points of f.
The Hessian of f is

$$H(\mathbf{x}) = \begin{pmatrix} -6x_1 & 3 \\ 3 & -6x_2 \end{pmatrix}$$

At $(1,1)$ this evaluates to

$$H(\mathbf{x}) = \begin{pmatrix} -6 & 3 \\ 3 & -6 \end{pmatrix}$$

which is negative definite. Therefore, we conclude that the function f attains a local maximum at $(1,1)$ (Corollary 5.1.1). At the other stationary point $(0,0)$, the Hessian evaluates to

$$H(\mathbf{x}) = \begin{pmatrix} 0 & 3 \\ 3 & 0 \end{pmatrix}$$

which is indefinite. $(0,0)$ is in fact a saddle point (Section 3.7.4). Hence $\mathbf{x}^* = (1,1)$ is the unique local maximum of f, where the function attains the value $f(1,1) = 1$. Note however that $(1,1)$ is not a global maximum of f since for example $f(-1,-1) = 5 > f(1,1)$. In fact, f has no global maximum on $\Re^2$ since for any x_2, $f(x_1, x_2) \to \infty$ as $x_1 \to -\infty$.

5 The first-order conditions for a maximum are

$$D_{x_1} f(x_1, x_2) = x_2 - 2x_1 = 0$$
$$D_{x_2} f(x_1, x_2) = x_1 + 3 - 2x_2 = 0$$

which have the unique solution $x_1^* = 1, x_2^* = 2$. $(1,2)$ is the only stationary point of f and hence the only possible candidate for a maximum. To verify that $(1,2)$ satisfies the second-order condition for a maximum, we compute the Hessian of f

$$H(\mathbf{x}) = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

which is negative definite everywhere. Therefore $(1,2)$ is a strict local maximum of f. Further, since f is strictly concave (Proposition 4.1), we conclude that $(1,2)$ is a strict global maximum of f (Exercise 5.2), where it attains its maximum value $f(1,2) = 3$ (Figure 1).

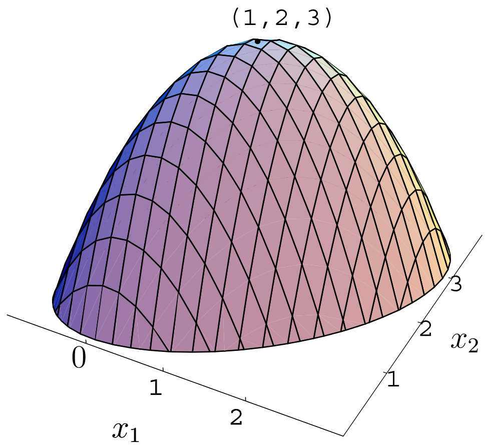


Figure 1: The strictly concave function $f(x_1, x_2) = x_1x_2 + 3x_2 - x_1^2 - x_2^2$ has a unique global maximum.

6 Assume $f\colon S \to \Re$ is pseudoconcave and $\nabla f(\mathbf{x}^*) = \mathbf{0}$. Pseudoconcavity means that

$$f(\mathbf{x}) > f(\mathbf{x}^*) \implies \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) > 0$$

for every $\mathbf{x}, \mathbf{x}^* \in S$, which is equivalent to

$$\nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq 0 \implies f(\mathbf{x}) \leq f(\mathbf{x}^*)$$

Therefore stationarity

$$\nabla f(\mathbf{x}^*) = \mathbf{0} \implies \nabla f(\mathbf{x}^*)^T(\mathbf{x} - \mathbf{x}^*) \leq 0 \implies f(\mathbf{x}) \leq f(\mathbf{x}^*) \text{ for every } \mathbf{x} \in S$$

$\mathbf{x}^*$ is a global optimum.
Conversely, if $\mathbf{x}^*$ is an interior global optimum, it must be an interior local optimum. Therefore (Theorem 5.1), $\nabla f(\mathbf{x}^*) = \mathbf{0}$.