A Course for Undergraduate Students

Optimization Methods

Instructor: Prof. WANG Zheng
Email: wangz@seu.edu.cn

Southeast University

September, 2025

What is Optimization?

- *Optimization* is the methodology of finding the minimum (or maximum) value of an objective function.
- This course concentrates on the fundamental theory, algorithm and application of linear programming, nonlinear programming, and dynamic programming.
- Prerequisite courses: *Calculus (Mathematical Analysis), Linear Algebra*

Why do we study optimization methods?

Answer:

There exist a lot of optimization problems in engineering systems, e.g., mechanical systems, manufacturing systems, power systems, aerospace systems, chemical engineering systems, transportation systems. Especially, *optimization theory* is closely related to *control system*, which is our expertise and research area.

What to learn?

- Optimization Problems in Engineering Systems
- Linear Programming and Mixed Integer Linear Programming
- Nonlinear Programming
- Dynamic Programming and Reinforcement Learning

How to learn?

■ Theory + Project

- *Optimization Theory*: Linear programming, nonlinear programming, mixed integer programming, dynamic programming
- *Project Research*: Solving an optimization problem from real-world engineering systems by **teamwork**, including modeling, analysis, algorithm design, programming and numerical experiments, paper writing and presentation

■ Textbook + Literature

- *Textbook Reading*: Constructing systematic theory and methodology framework of optimization
- *Literature Reading*: Understanding the application of optimization theory and methodology to real-world engineering systems

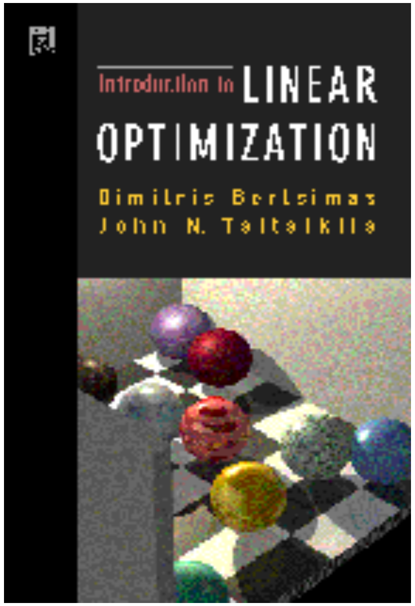
■ Learning + Presentation

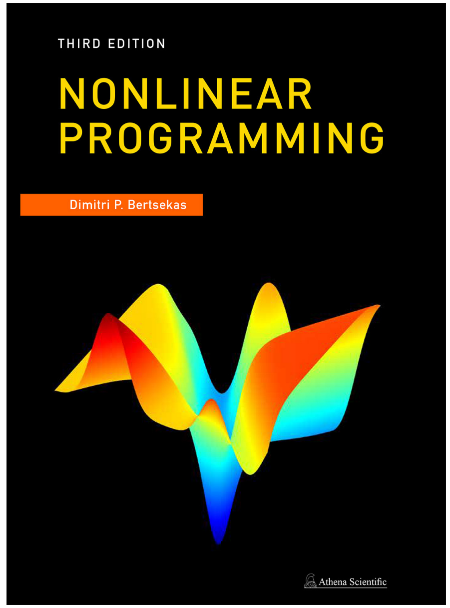
- *Learning*: Understanding what others have done or are doing
- *Presentation*: Tell others what you have done or are doing

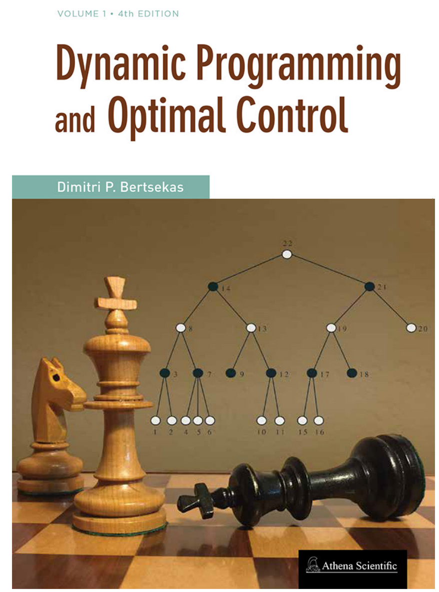
How to grade?

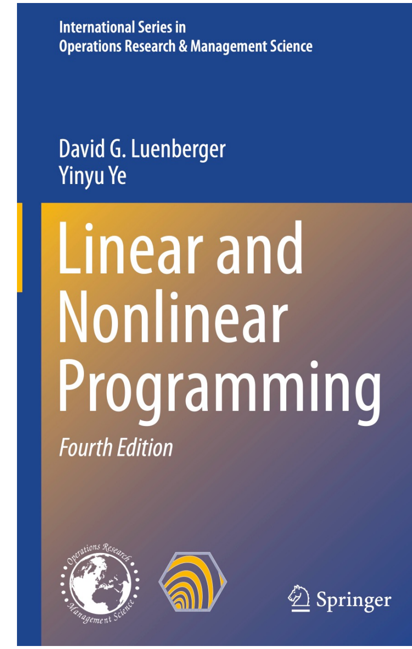
- Assignments: 15 %
- Algorithm Programming: 5%
- Project Research: 30%
- Final Exam: 50%

Textbooks and Supporting Materials









- D. Bertsimas and J. N. Tsitsiklis. *Introduction to Linear Optimization*, Athena Scientific, 1997.
- D. P. Bertsekas. *Nonlinear Programming (3rd Edition)*, Athena Scientific, 2016.
- D. P. Bertsekas. *Dynamic Programming and Optimal Control (Vol. 1, 4^{rd} Edition)*, Athena Scientific, 2017.
- D. G. Luenberger and Y. Ye. Linear and Nonlinear Programming (4^{th} Edition), Springer, 2016.

Some high quality research papers on optimization problems in engineering systems from top journals.

Chapter 1

Fundamental Theory of Optimization

Section 1.1

What is an Optimization Problem

- Definition of an Optimization Problem
- The Optimization Problem We Have Learnt Before
- Examples

Optimization Problem: Make decisions to optimize performance

When people control, operate, or design engineering systems (*e.g., mechanical systems, electricity systems, industrial and manufacturing systems, transportation systems, aeronautical and astronautical systems, etc.*), they usually need to make decisions on control, operation and design of engineering systems carefully such that their performances can be optimized (i.e., minimized or maximized).

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && g(x) \leq 0 \end{aligned}$$

Objective function

Constraint

Decision variable

$$x \in R^n, f(x) \in R, g(x) \in R^m$$

Unconstrained optimization problem

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

$$x = (x_1, \cdots, x_n)^T$$

Constrained optimization problem

$$\min_{x \in [a,b]} f(x) \qquad or \qquad \min_{x \in \Omega} f(x)$$

$$\Omega = \{x | g(x) \leq 0\}$$

minimize	$f(x)$	Objective function
subject to	$g(x) \leq 0$	Constraint
		Decision variable

$x \in R^n, f(x) \in R, g(x) \in R^m$, i.e., $g(x) = [g_1(x), \cdots, g_m(x)]^T$

Examples:

- *Electric power system*: Optimize the power flow such that the cost for electricity generation and transmission can be minimized while the loads in the electric network can be satisfied.
- *Manufacturing system*: Optimize the production plan and scheduling so as to minimize the production cost, inventory holding cost and the cost for failing to meet the demand.
- *Transportation system*: Optimize the vehicle flow in a road network such that the time for travelling can be as short as possible.

Optimization problems exist everywhere in different types of engineering systems

The Optimization Method We Have Learnt in the Course of "Calculus" (or "Mathematical Analysis")

For a differentiable function, take its first order derivative and find the solution that makes the derivative be equal to zero; for that solution, check whether the second order derivative of the function is positive …

$$f'(x) = 0 \xrightarrow{x^*} f''(x^*) > 0\ ?$$

Example 1:

$$\min \quad 2x^2$$

Example 2:

$$\begin{aligned} \min \quad & 2x^2 \\ \text{s.t.} \quad & x \in [1, 2] \end{aligned}$$

or

$$\begin{aligned} \min \quad & 2x^2 \\ \text{s.t.} \quad & x \in [-2, -1] \end{aligned}$$

Example 3:

$$\min \quad x^3 - x^2$$

Example 4:

$$\begin{aligned} \min \quad & x^3 - x^2 \\ \text{s.t.} \quad & x \in [-1, 1] \end{aligned}$$

Example 5:

Solve the following optimization problem:

$$\min \quad x^2 + y^2$$

Example 6:

Solve the following optimization problem:

$$\begin{aligned} \min \quad & x^2 + y^2 \\ \text{s.t.} \quad & x + y = 1 \end{aligned}$$

Section 1.2

Fundamental Theory of Optimization

- General Formulation of Optimization Problem
- Optimality Condition
- Algorithms for Optimization Problems

1.2.1. General Formulation of Optimization Problem

Optimization Problem

Find the minimum (or maximum) value of a given objective function over a constrained set of decision variables.

Main Categories of Optimization Problems

– Linear Programming

– Mixed Integer Linear Programming

– Nonlinear Programming

– Dynamic Programming

An Optimization Problem can be formulated by

(P_0) $\min_{x \in \Omega} f(x)$ or (P_0) $\min \quad f(x)$

$\text{s.t.} \quad g(x) \leq 0$

$x \in R^n$ Decision variable

$f: R^n \rightarrow R$ Objective function

Ω Feasible set

$x \in \Omega$ A feasible solution

$x^* \in \Omega$ An optimal solution

$g: R^n \rightarrow R$

Constraint function

$\Omega = \{x | g(x) \leq 0\}$

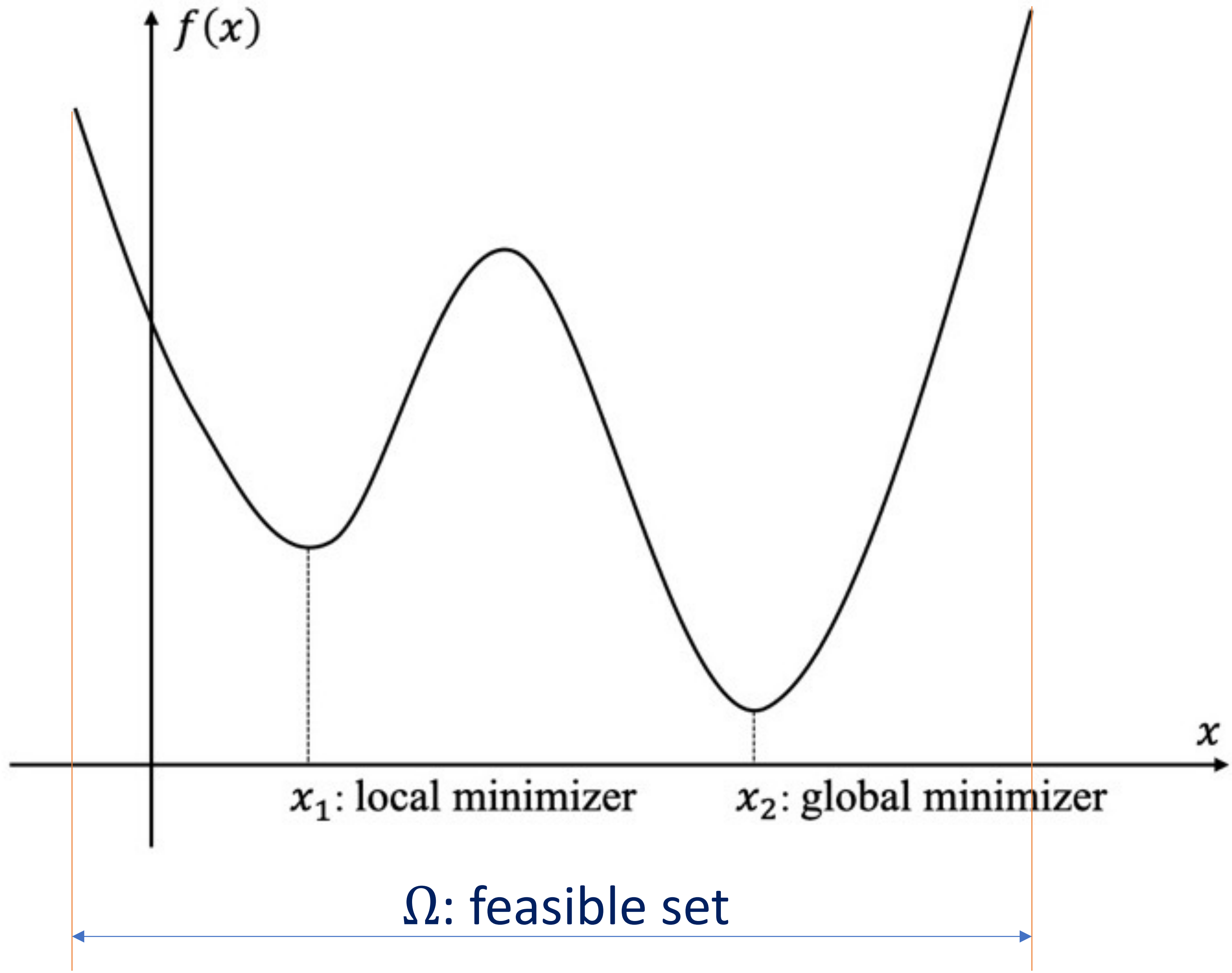
If $f(x)$ and $g(x)$ are linear in x, then P_0 is a *linear programming* (LP); otherwise, P_0 is a *nonlinear programming* (NLP)

Local minimizer

For a constrained minimization problem $\min_{x \in \Omega} f(x)$, if there exists an x^* and a $\delta > 0$ such that $f(x^*) \leq f(x)$ holds for all the x satisfying $x \in \Omega$ and $\|x - x^*\| < \delta$ (i.e., $\forall x \in N(x^*, \delta) \cap \Omega$, $N(x^*, \delta) = \{x \,|\, \|x - x^*\| < \delta\}$ is a neighborhood of x^*, then x^* is a local minimizer of the constrained optimization problem.

Global minimizer

For a constrained minimization problem $\min_{x \in \Omega} f(x)$, if there exists an x^* such that $f(x^*) \leq f(x)$ holds for all the $x \in \Omega$, then x^* is a global minimizer of the constrained optimization problem.



$f(x)$
x
x_1: local minimizer
x_2: global minimizer
Ω: feasible set

1.2.2.Optimality Condition

- *Necessary condition of a local minimizer of an unconstrained optimization problem with univariate objective function.*

Consider an unconstrained optimization problem $\min_{x \in R} f(x)$.
Suppose that the univariate objective function f is twice-differentiable. Then the necessary condition for $x^* \in R$ to be a local minimizer of the function f is

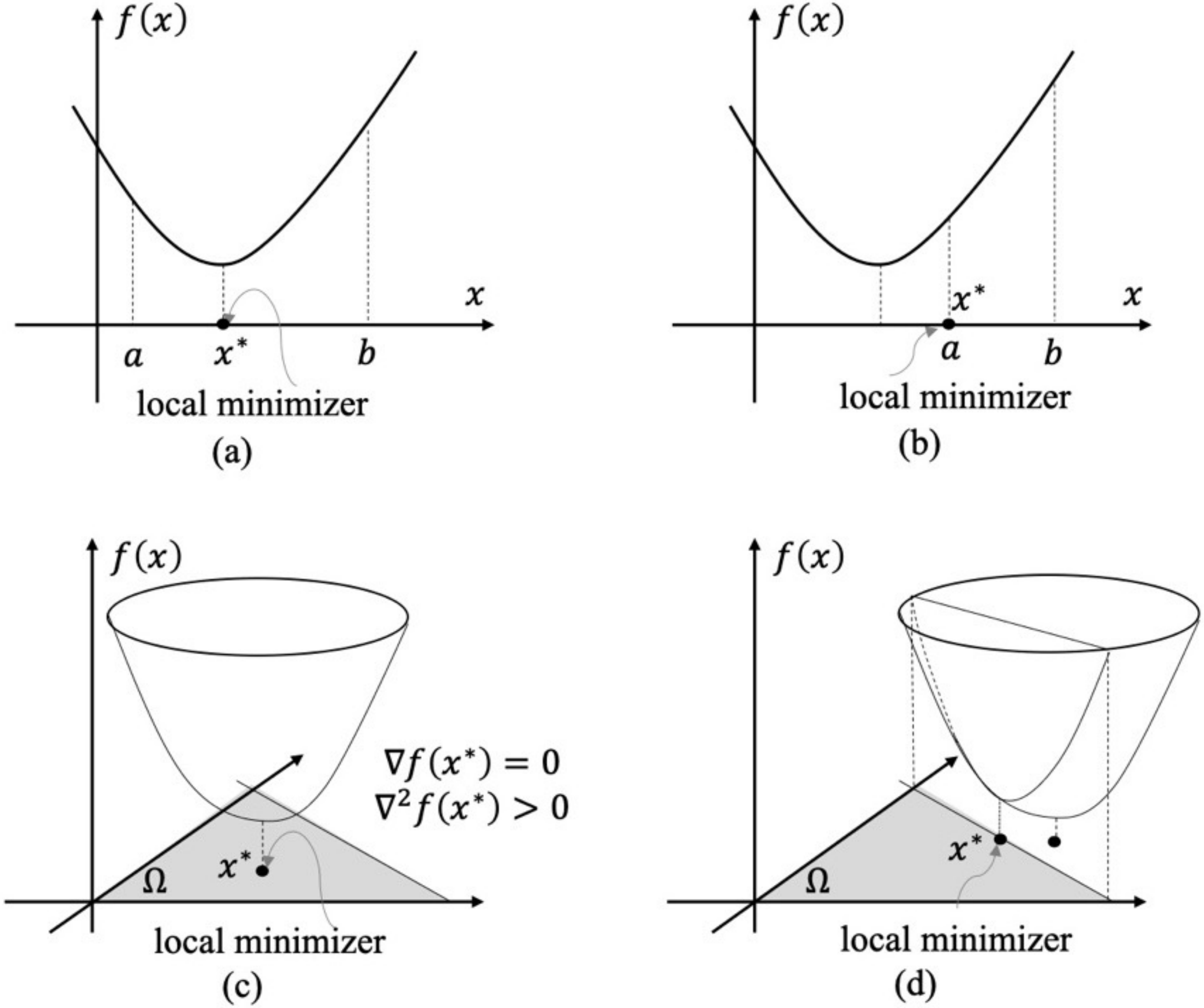
$$\left.\frac{df(x)}{dx}\right|_{x=x^*} = 0 \quad and \quad \left.\frac{d^2 f(x)}{dx^2}\right|_{x=x^*} \geq 0$$

■ *Necessary condition of a local minimizer of an unconstrained optimization problem with multivariate objective function.*

Consider an unconstrained optimization problem $\min_{x \in R^n} f(x)$. Suppose that the objective function f is twice-differentiable. Then the necessary condition for $x^* \in R^n$ to be a local minimizer of the function $f(x)$ is

$$\nabla f(x^*) = 0 \quad and \quad \nabla^2 f(x^*) \geq 0$$

where $\nabla f(x) = \left(\frac{\partial f(x)}{\partial x_1}, \cdots, \frac{\partial f(x)}{\partial x_n}\right)^T$ is the gradient of $f(x)$ and $\nabla^2 f(x) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{n \times n}$ is the Hessian matrix of $f(x)$.



Optimality conditions of univariate and multivariate functions

(a) The local minimizer (x^*) is in the interior of the feasible set $[a, b]$; (b) the local minimizer (x^*) is on the boundary of the feasible set $[a, b]$; (c) the local minimizer (x^*) satisfies $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*) > 0$ and is in the interior of the feasible set (Ω); (d) the local minimizer (x^*) is on the boundary of the feasible set (Ω).

1.2.3. Algorithms for Optimization Problems

Algorithm: An iterative procedure to approach the optimal solution.

An optimization algorithm can be captured as

$$x_{k+1} = x_k + \alpha_k p_k$$

k : index of iteration;

x_k: solution in the *k*th iteration;

p_k: optimization direction in the *k*th iteration;

α_k: step size in the *k*th iteration.

Initial solution: x_0

Termination condition:

$$\|x_{k+1} - x_k\| < \varepsilon \text{ or } |f(x_{k+1}) - f(x_k)| < \varepsilon$$

Steepest Descent Method

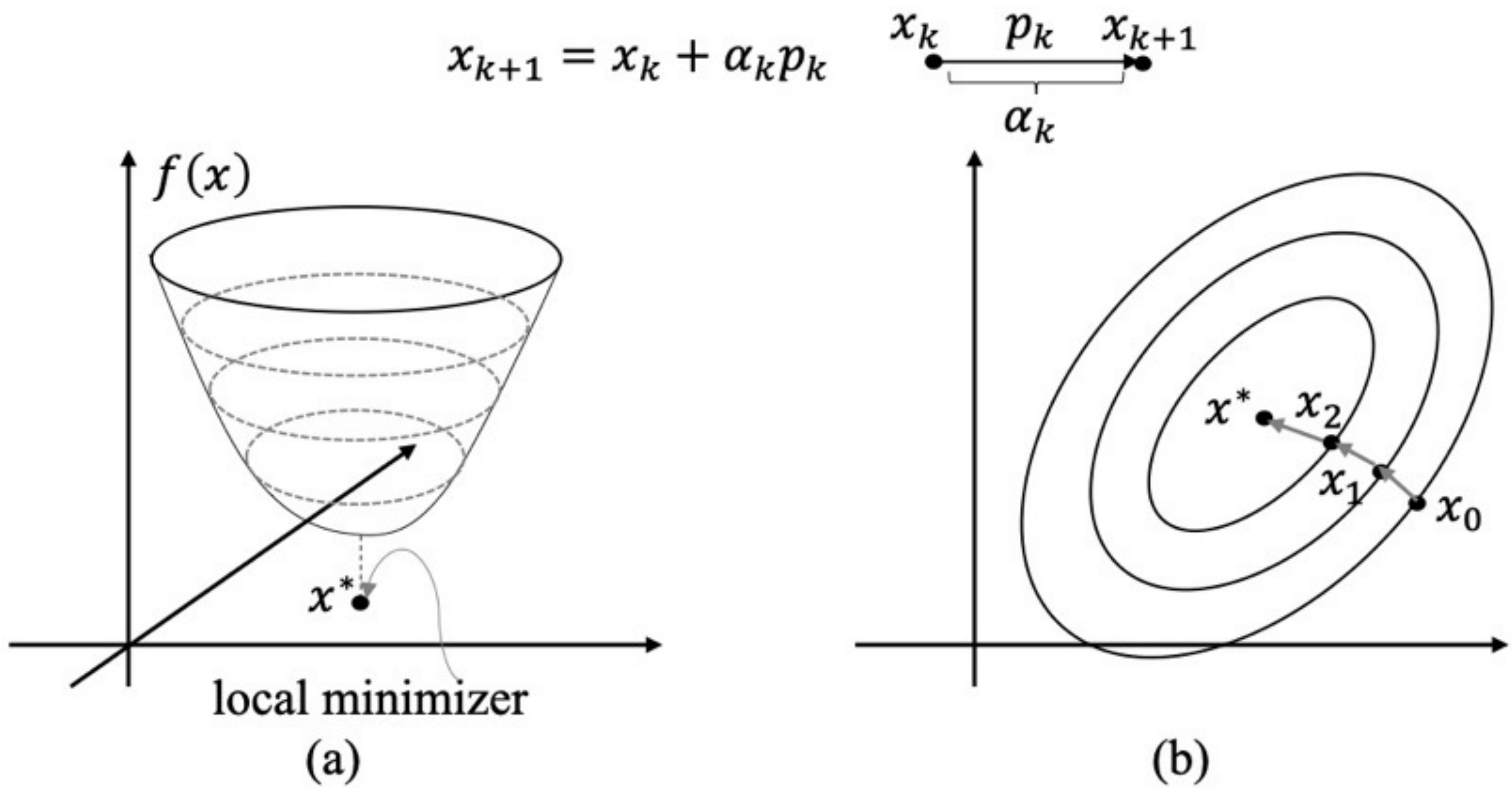
$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$
$$p_k = -\nabla f(x_k)$$

Brief Explanation: expand $f(x_{k+1})$ to its Taylor series

$$f(x_{k+1}) = f(x_k + \alpha_k p_k)$$
$$= f(x_k) + \alpha_k [\nabla f(x_k)]^T p_k + o(\alpha_k \|p_k\|)$$

$$[\nabla f(x_k)]^T p_k = \|\nabla f(x_k)\| \cdot \|p_k\| \cdot \cos\theta$$

- $\theta = \pi$ makes $\cos\theta = -1$ and minimizes $[\nabla f(x_k)]^T p_k$ and $f(x_{k+1})$;
- $\theta = \pi$ means that the optimization direction is exactly opposite to the gradient direction of the objective function at x_k ;
- $p_k = -\nabla f(x_k)$ is the direction that makes the objective function $f(x)$ decrease fastest (i.e., *steepest descent*).



General formulation of an optimization algorithm

(a) The local minimum x^* of $f(x)$; (b) the procedure of the optimization algorithm is illustrated on the contour map of $f(x)$.

The step size α_k can be obtained by solving the following one-dimensional optimization problem

$$\alpha_k = \underset{\alpha>0}{\operatorname{argmin}}\, f[x_k - \alpha \nabla f(x_k)]$$

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k)$$