

优化方法基础

[机器学习（入门）](#) [DC学院](#)

优化方法：

指的是改变 x 以最小化或最大化某个函数 $f(x)$ 的任务，机器学习通常以最小化 $f(x)$ 指代大多数优化问题（最大化优化 $-f(x)$ ）

机器学习中把要最小化或最大化的函数称为**目标函数**，对其最小化时，又称为**代价函数**、**损失函数**、**误差函数**

导数：

假设我们有一个函数 $y = f(x)$ ，其中 x 和 y 是实数。这个函数的导数记为 $f'(x)$ 或 $\frac{dy}{dx}$ 。导数 $f'(x)$ 代表 $f(x)$ 在点 x 处的斜率。换句话说，它表明如何缩放输入的小变化才能在输出获得相应的变化： $f(x + \epsilon) \approx f(x) + \epsilon f'(x)$ 。

因此导数对于最小化一个函数很有用，因为它告诉我们如何更改 x 来略微地改善 y 。例如，我们知道对于足够小的 ϵ 来说， $f(x - \epsilon \text{sign}(f'(x)))$ 是比 $f(x)$ 小的。因此我们可以将 x 往导数的反方向移动一小步来减小 $f(x)$ 。

正定矩阵： positive definite matrix

如果对于任意的 $x \neq 0$ ，都有 $x^T A x > 0$ （或 $x^T A x \geq 0$ ），则称矩阵 A 是正定矩阵（半正定矩阵），记为 $A > 0$ （或 $A \geq 0$ ）。

负定矩阵：

如果对于任意的 $x \neq 0$ ，都有 $x^T A x < 0$ （或 $x^T A x \leq 0$ ），则称矩阵 A 是负定矩阵（半负定矩阵），记为 $A < 0$ （或 $A \leq 0$ ）。