1 梯度下降法

梯度下降法（gradient descent）是求解无约束最优化问题的一种最常用的方法，有实现简单的特点。梯度下降法是迭代算法，每一步需要求解目标函数的梯度向量。

输入：目标函数 $f(x)$，梯度函数 $g(x) = \nabla f(x)$，计算精度 ε。

输出：$f(x)$ 的极小点 x^*。

（1）取初始值 $x^{(0)} \in \mathbf{R}^n$，置 $k = 0$

（2）计算 $f(x^{(k)})$ （3）计算梯度 $g_k = g(x^{(k)})$，当 $\|g_k\| < \varepsilon$ 时，停止迭代，令 $x^* = x^{(k)}$；否则令 $p_k = g(x^{(k)})$，求 λ_k，使

$$f(x^{(k)} + \lambda_k p_k) = \min_{\lambda \geq 0} f(x^{(k)} + \lambda p_k)$$

（4）置 $x^{(k+1)} = x^{(k)} + \lambda_k p_k$，计算 $f(x^{(k+1)})$

当 $\|f(x^{(k+1)}) - f(x^{(k)})\| < \varepsilon$ 或 $\|x^{(k+1)} - x^{(k)}\| < \varepsilon$ 时停止迭代，令 $x^* = x^{(k+1)}$

（5）否则，置 $k = k + 1$，转（3）。

当目标函数是凸函数时，梯度下降法的解是全局最优解。一般情况下，其解不保证是全局最优解。梯度下降法的收敛速度也未必是最快的。