

第五章 最优化方法 (6)

共轭梯度法

共轭梯度法 (Fletcher-Reeves 型, FR)

算法思想

最速下降法每次都沿着负梯度方向前进, 但方向之间可能互相“打架”, 导致 zig-zag (之字形) 式缓慢收敛。

共轭梯度法的思想是:

第一步沿负梯度走; 后续方向由“当前梯度 + 上一步方向”组合而成, 使得方向之间不会相互干扰, 从而提高收敛效率。

对于二次函数且 Hessian 正定的情况, 共轭梯度法最多经过 n 步即可精确到达最优点。

迭代公式

给定初始迭代点 $x^{(1)}$, 令梯度 $\nabla f(x^{(k)})$

- 初始化方向

$$d^{(1)} = -\nabla f(x^{(1)}) + \beta_1 d^{(0)}$$

其中 $\beta_1 = 0$, $d^{(0)} = 0$. 因此初始 $d^{(1)}$ 搜索方向就是梯度相反方向。

- 求步长 α_k

$$\alpha_k = \arg \min_{\alpha \geq 0} f(x^{(k)} + \alpha d^{(k)})$$

构造一元函数:

$$\varphi(\alpha) = f(x^{(k)} + \alpha d^{(k)})$$

求导并令零:

$$\varphi'(\alpha) = 0 \Rightarrow \alpha_k$$

- 更新迭代点 $x^{(k+1)}$

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)}$$

- 计算新梯度 $\nabla f(x^{(k+1)})$ 和 FR 系数 $\beta^{(k+1)}$

$$\beta_{k+1} = \frac{\|\nabla f(x^{(k+1)})\|^2}{\|\nabla f(x^{(k)})\|^2}$$

- 更新方向 $d^{(k+1)}$

$$d^{(k+1)} = -\nabla f(x^{(k)}) + \beta_{k+1}d^{(k)}$$

重复上述过程，直到： $\|\nabla f(x^{(k)})\| \leq \varepsilon$ 。

例题

已知函数和初始点

$$f(x) = x_1^2 + 2x_2^2 - 4x_1 - 2x_1x_2, \quad x^{(1)} = (1, 1)^T$$

用 FR 共轭梯度法求解，并写出迭代过程。

解答：

先求梯度

$$\nabla f(x) = (2x_1 - 2x_2 - 4, -2x_1 + 4x_2)^T$$

第一次迭代

- 梯度：

$$\nabla f(x^{(1)}) = (-4, 2)^T$$

- 初始方向：

$$d^{(1)} = -\nabla f(x^{(1)}) + \beta_1 d^{(0)} = (4, -2)^T$$

其中 $\beta_1 = 0, d^{(0)} = 0$ 。

- 求步长

$$x^{(1)} + \alpha d^{(1)} = (1 + 4\alpha, 1 - 2\alpha)^T$$

构造：

$$\varphi(\alpha) = f(1 + 4\alpha, 1 - 2\alpha) = 40\alpha^2 - 20\alpha - 3$$

求导：

$$\varphi'(\alpha) = 80\alpha - 20 = 0 \Rightarrow \alpha_1 = \frac{1}{4}$$

- 更新点

$$x^{(2)} = (1, 1)^T + \frac{1}{4}(4, -2)^T = (2, \frac{1}{2})^T$$

第二次迭代

- 新梯度：

$$\nabla f(x^{(2)}) = (-1, -2)^T$$

- 计算 FR 系数：

$$\beta_2 = \frac{|g^{(2)}|^2}{|g^{(1)}|^2} = \frac{5}{20} = \frac{1}{4}$$

- **新方向:**

$$d^{(2)} = -g^{(2)} + \beta_2 d^{(1)} = (1, 2) + \frac{1}{4}(4, -2) = (2, \frac{3}{2})$$

- **求步长**

$$x^{(2)} + \alpha d^{(2)} = (2 + 2\alpha, \frac{1}{2} + \frac{3}{2}\alpha)$$

代入:

$$\varphi_2(\alpha) = \frac{5}{2}\alpha^2 - 5\alpha - \frac{11}{2}$$

求导:

$$\varphi_2'(\alpha) = 5\alpha - 5 = 0 \Rightarrow \alpha_2 = 1$$

- **更新点**

$$x^{(3)} = (2, \frac{1}{2})^T + (2, \frac{3}{2})^T = (4, 2)^T$$

此时, 代入 $\nabla f(x)$ 发现:

$$\nabla f(4, 2) = (0, 0)^T \Rightarrow \text{满足停止条件}$$

综上, 最终答案:

$$\boxed{x^* = (4, 2)^T}$$
