共轭梯度法

董波
数学科学学院
大连理工大学

共轭梯度法

共轭梯度法（conjugate gradient method, CG）是以共轭方向（conjugate direction）作为搜索方向的一类算法。

共轭梯度法是由Hesteness和Stiefel于1952年为求解线性方程组而提出的。后来用于求解无约束最优化问题，它是一种重要的数学优化方法。

Fig. 1. Wolfgang Wasow (left) and Magnus Hestenes (right).

Fig. 2. Eduard Stiefel [*From Zeitschrift für angewandte Mathematik und Physik* 30, 139 (1979); used with permission].

线性方程组 $Ax = b$,其中A为n阶对称正定矩阵

$$Ax = b \Longleftrightarrow \min \phi(x) := \frac{1}{2}\left(x - x^*\right)^T A\left(x - x^*\right)$$

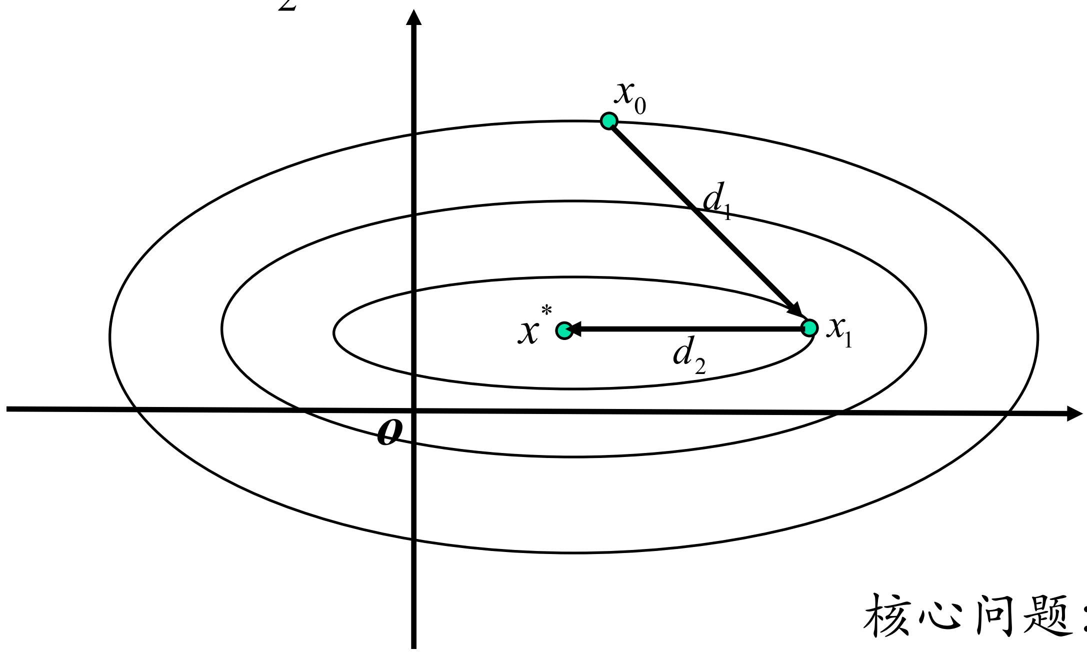
$$\Longleftrightarrow \min \frac{1}{2}\left(Ax, x\right) - \left(b, x\right) + \frac{1}{2}\left(x^*, b\right)$$

无约束最优化问题 $\min \varphi(x) := \frac{1}{2}\left(Ax, x\right) - \left(b, x\right)$

x^*是线性方程组的解 $\Longleftrightarrow$ $\varphi(x^*) = \min \varphi(x)$

$$\min \varphi(x) := \frac{1}{2}(Ax, x) - (b, x) \quad \Longleftrightarrow \quad \min \phi(x) := \frac{1}{2}\left(x - x^*\right)^T A\left(x - x^*\right)$$

函数 $\phi(x)$ 的等值面 $\frac{1}{2}(x-x^*)^T A(x-x^*) = c$ 表示以 x^* 为中心的椭球面



核心问题：确定方向、步长

最速下降法

$$\min \varphi(x) := \frac{1}{2}\left(Ax, x\right) - \left(b, x\right)$$

下降方向： 负梯度方向

$$r = -\nabla\varphi(x) = -\left(Ax - b\right) = b - Ax$$

步长： 寻求步长 α ，使得 $\varphi(x+\alpha r) = \min$

$$\varphi(x+\alpha r) = \frac{1}{2}\left(A(x+\alpha r),(x+\alpha r)\right) - \left(b,(x+\alpha r)\right)$$

$$\frac{d\varphi(x+\alpha r)}{d\alpha} = \alpha\left(Ar, r\right) - \left(r, r\right) = 0$$

$$\alpha = \frac{\left(r, r\right)}{\left(Ar, r\right)}$$

算法及收敛性分析

算法流程: 对于给的初始向量x_0

$$r_k = b - Ax_k$$

$$\alpha_k = (r_k, r_k)/(Ar_k, r_k) \qquad k = 0,1,2,\ldots$$

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

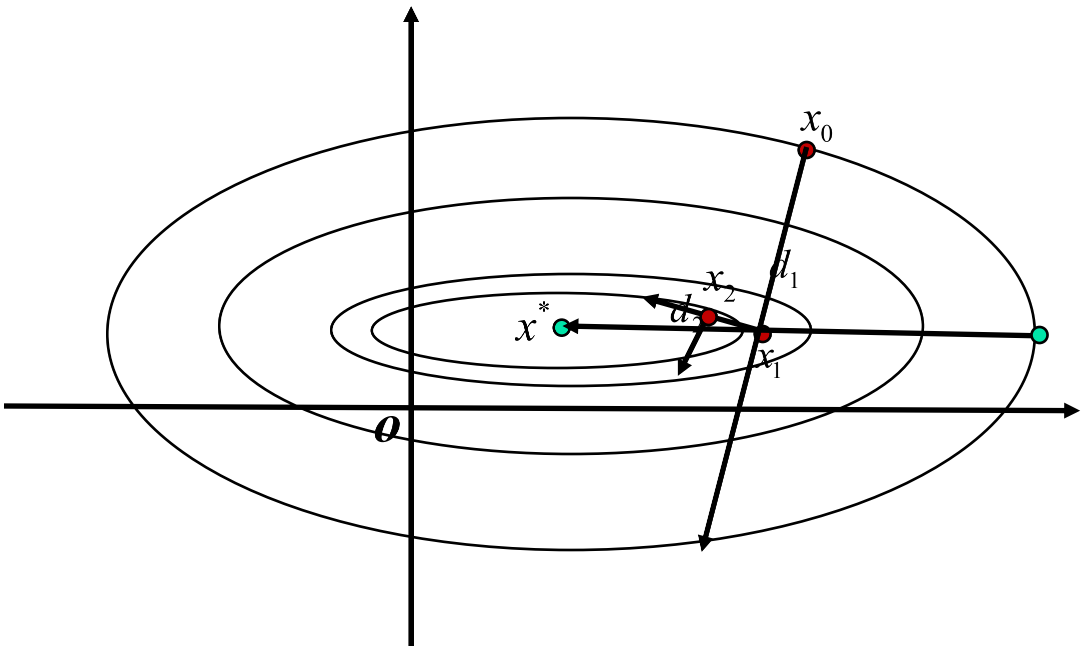
收敛性定理:

$$(r_{k+1}, r_k) = 0, \qquad k = 0,1,2,\ldots$$

$$\|x_k - x^*\|_A \le \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \|x_0 - x^*\|_A$$

其中 λ_1, λ_n 分别为矩阵A的最大最小特征值，$\|x\|_A = \sqrt{(Ax, x)}$

最速下降法收敛速度



共轭梯度法

$$\min \varphi(x) := \frac{1}{2}\left(Ax, x\right) - \left(b, x\right)$$

下降方向： 上步下降方向与当前残量的线性组合

$$p_k = r_k + \beta_{k-1} p_{k-1}$$

步长： 寻求步长α，使得 $\varphi(x + \alpha p) = \min$

$$\alpha = \frac{\left(r, p\right)}{\left(Ap, p\right)}$$

下降方向的计算

$$\min \varphi(x) := \frac{1}{2}\left(Ax, x\right) - \left(b, x\right)$$

对于 $p_k = r_k + \beta_{k-1} p_{k-1}$，选择 β_{k-1}，使得

$$\varphi\left(x_k + \alpha_k \left(r_k + \beta_{k-1} p_{k-1}\right)\right) = \min$$

$$\frac{d\varphi\left(x_k + \alpha_k \left(r_k + \beta p_{k-1}\right)\right)}{d\beta} = \left(\alpha_k^2 p_{k-1}^T A p_{k-1}\right)\beta - \alpha_k \left(p_{k-1}^T r_k - \alpha_k p_{k-1}^T A r_k\right) = 0$$

由于 $p_{k-1}^T r_k = 0$，

$$\beta_{k-1} = \frac{p_{k-1}^T A r_k}{p_{k-1}^T A p_{k-1}}$$

共轭梯度法性质

$$\left(p_i, r_k\right) = 0, \quad i = 0,1,\ldots,k\text{-}1$$

$$\left(r_i, r_k\right) = 0, \quad i = 0,1,\ldots,k\text{-}1$$

$$\left(Ap_i, p_k\right) = 0, \quad i = 0,1,\ldots,k\text{-}1$$

共轭梯度法最多迭代n步即可求得问题的精确解

$$\alpha_k = \frac{\left(r_k, p_k\right)}{\left(Ap_k, p_k\right)} = \frac{\left(r_k, r_k + \beta_{k-1} p_{k-1}\right)}{\left(Ap_k, p_k\right)} = \frac{\left(r_k, r_k\right)}{\left(Ap_k, p_k\right)}$$

$$\beta_{k-1} = \frac{p_{k-1}^T A r_k}{p_{k-1}^T A p_{k-1}} = \frac{r_k^T \left(r_{k-1} - r_k\right)}{p_{k-1}^T A p_{k-1}} \frac{1}{\alpha_{k-1}} = \frac{\left(r_k, r_k\right)}{\left(r_{k-1}, r_{k-1}\right)}$$

算法流程

对于给的初始向量x_0，$r_0 = p_0 = b - Ax_0$

$$\alpha_k = (r_k, r_k)/(Ap_k, p_k)$$

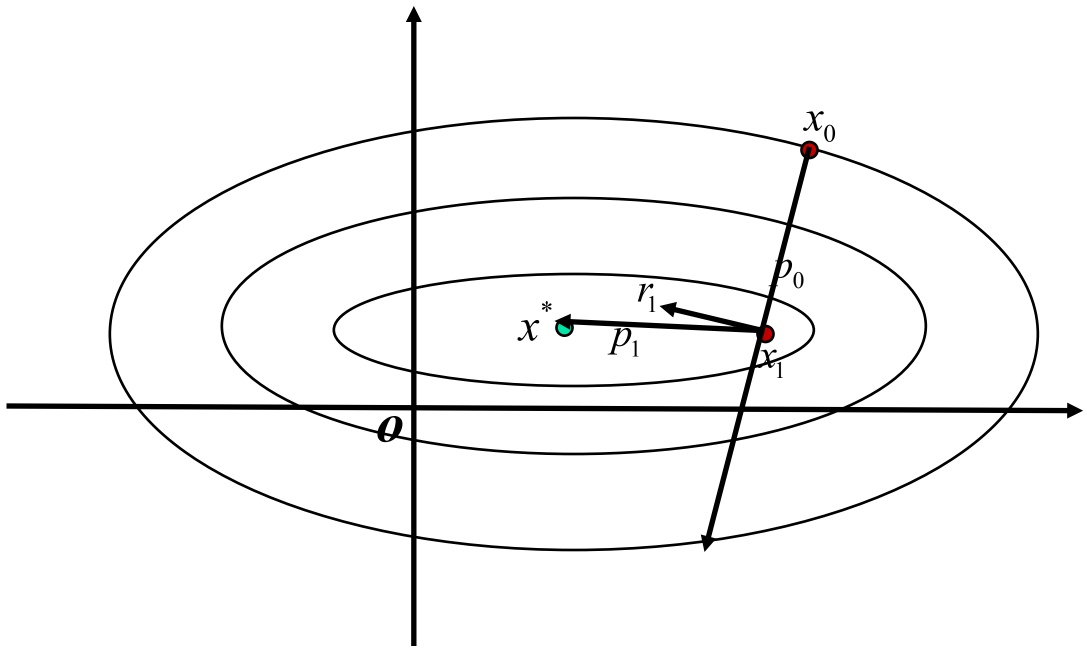
$$x_{k+1} = x_k + \alpha_k p_k$$

$$r_{k+1} = r_k - \alpha_k A p_k \qquad k = 0, 1, 2, \ldots$$

$$\beta_k = (r_{k+1}, r_{k+1})/(r_k, r_k)$$

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

共轭梯度法收敛速度



例题

共轭梯度法求解线性方程组 $\begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\quad x_0 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$

$$r_0 = p_0 = b - Ax_0 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$$

$$\alpha_0 = (r_0, r_0)/(p_0, Ap_0) = 3/10$$

$$x_1 = x_0 + \alpha_0 p_0 = \begin{pmatrix} 3/10 & 3/10 & 3/10 \end{pmatrix}^T$$

$$r_1 = b - Ax_1 = \begin{pmatrix} 1/10 & 1/10 & -2/10 \end{pmatrix}^T$$

$$\beta_0 = (r_1, r_1)/(r_0, r_0) = 1/50$$

$$p_1 = r_1 + \beta_0 p_0 = \begin{pmatrix} 3/25 & 3/25 & -9/50 \end{pmatrix}^T$$

$$\alpha_1 = (r_1, r_1)/(p_1, Ap_1) = 5/3$$

$$x_2 = x_1 + \alpha_1 p_1 = \begin{pmatrix} 1/2 & 1/2 & 0 \end{pmatrix}^T$$

$$r_2 = b - Ax_2 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$$