6.4 共轭梯度法

朱升峰

华东师范大学 数学科学学院

2022 年 12 月

等价问题

共轭梯度法简称 CG (conjugate gradient) 方法，对应于求一个二次函数的极值.
设 $\boldsymbol{A} = (a_{ij}) \in \mathbb{R}^{n\times n}$ 是对称正定矩阵，$\boldsymbol{b} = (b_1, b_2, \cdots, b_n)^{\mathrm{T}}$，求解线性方程组

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}.$$

考虑二次函数 $\varphi : \mathbb{R}^n \to \mathbb{R}$:

$$\varphi(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) - (\boldsymbol{b}, \boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j - \sum_{j=1}^{n} b_j x_j.$$

函数 φ 有如下性质:

(1) 对一切 $\boldsymbol{x} \in \mathbb{R}^n, \varphi(x)$ 的梯度

$$\nabla \varphi(x) = \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}$$

(2) 对一切 $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ 及 $\alpha \in \mathbb{R}$,

$$\begin{aligned}\varphi(\boldsymbol{x}+\alpha\boldsymbol{y}) &= \frac{1}{2}(\boldsymbol{A}(\boldsymbol{x}+\alpha\boldsymbol{y}), \boldsymbol{x}+\alpha\boldsymbol{y}) - (\boldsymbol{b}, \boldsymbol{x}+\alpha\boldsymbol{y}) \\ &= \varphi(\boldsymbol{x}) + \alpha(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}, \boldsymbol{y}) + \frac{\alpha^2}{2}(\boldsymbol{A}\boldsymbol{y}, \boldsymbol{y})\end{aligned}$$

(3) 设 $\boldsymbol{x}^* = \boldsymbol{A}^{-1}\boldsymbol{b}$ 是线性方程组的解, 则有

$$\varphi\left(\boldsymbol{x}^*\right) = -\frac{1}{2}\left(\boldsymbol{b}, \boldsymbol{A}^{-1}\boldsymbol{b}\right) = -\frac{1}{2}\left(\mathbf{A}\boldsymbol{x}^*, \boldsymbol{x}^*\right)$$

且对一切 $x \in \mathbb{R}^n$, 有

$$\begin{aligned}\varphi(\boldsymbol{x}) - \varphi\left(\boldsymbol{x}^*\right) &= \frac{1}{2}(\boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}) - \left(\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}\right) + \frac{1}{2}\left(\boldsymbol{A}\boldsymbol{x}^*, \boldsymbol{x}^*\right) \\ &= \frac{1}{2}\left(\boldsymbol{A}\left(\boldsymbol{x}-\boldsymbol{x}^*\right), \boldsymbol{x}-\boldsymbol{x}^*\right).\end{aligned}$$

最速下降法

定理 1

设 $\boldsymbol{A}$ 对称正定, 则 $\boldsymbol{x}^*$ 为线性方程组 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 的解的充分必要条件是 $\boldsymbol{x}^*$ 满足

$$\varphi(\boldsymbol{x}^*)=\min_{\boldsymbol{x}\in\mathbb{R}^n}\varphi(\boldsymbol{x})$$

由定理可知, 求 $\boldsymbol{x}^*\in\mathbb{R}^n$ 使 $\varphi(x)$ 达到最小值, 这就是求解线性方程组等价的变分问题. 求解方法是构造一个向量序列 $\{\boldsymbol{x}^{(k)}\}$ 使 $\varphi(\boldsymbol{x}^{(k)})\to\varphi(\boldsymbol{x}^*)$

最速下降法

通常求 $\varphi(\boldsymbol{x})$ 的极小点 $\boldsymbol{x}^*$ 可转化为求一维问题的极小, 即从 $\boldsymbol{x}^{(0)}$ 出发, 找一个方向 $\boldsymbol{p}^{(0)}$, 令 $x^{(1)} = \boldsymbol{x}^{(0)} + \alpha \boldsymbol{p}^{(0)}$, 使 $\varphi\left(\boldsymbol{x}^{(1)}\right) = \min\limits_{\alpha \in \mathbb{R}} \varphi\left(\boldsymbol{x}^{(0)} + \alpha \boldsymbol{p}^{(0)}\right)$.

一般地, 令

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)} \tag{1}$$

使

$$\varphi\left(\boldsymbol{x}^{(k+1)}\right) = \min_{a \in \mathbb{R}} \varphi\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{p}^{(k)}\right)$$

$$\varphi\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{p}^{(k)}\right) = \varphi\left(\boldsymbol{x}^{(k)}\right) + \alpha\left(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}, \boldsymbol{p}^{(k)}\right) + \frac{\alpha^2}{2}\left(\boldsymbol{A}\boldsymbol{p}^{(k)}, \boldsymbol{p}^{(k)}\right)$$

$$\frac{\mathrm{d}\varphi\left(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{p}^{(k)}\right)}{\mathrm{d}\alpha} = \left(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}, \boldsymbol{p}^{(k)}\right) + \alpha\left(\boldsymbol{A}\boldsymbol{p}^{(k)}, \boldsymbol{p}^{(k)}\right) = 0$$

于是可得

$$\alpha_k = -\frac{\left(\boldsymbol{A}\boldsymbol{x}^{(k)} - \boldsymbol{b}, \boldsymbol{p}^{(k)}\right)}{\left(\boldsymbol{A}\boldsymbol{p}^{(k)}, \boldsymbol{p}^{(k)}\right)}, \tag{2}$$

样得到的 α_k 显然满足

$$\varphi\left(\boldsymbol{x}^{(k)}+\alpha_k\boldsymbol{p}^{(k)}\right)\leqslant\varphi\left(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{p}^{(k)}\right),\quad\forall\alpha\in\mathbb{R},$$

这就是求 $\varphi(x)$ 极小点的下降算法，这里 $p^{(k)}$ 是任选的一个方向，如果我们选一个方向 $\boldsymbol{p}^{(k)}$ 使 $\varphi(\boldsymbol{x})$ 在点 $\boldsymbol{x}^{(k)}$ 沿 $\boldsymbol{p}^{(k)}$ 下降最快，从 $\boldsymbol{x}^{(k)}$ 出发，先找一个使函数值 $\varphi(\boldsymbol{x})$ 减少最快的方向，这就是正交于椭球面的函数 $\varphi(\boldsymbol{x})$ 的负梯度方向

$$-\nabla\varphi\left(\boldsymbol{x}^{(k)}\right)=-\left(\frac{\partial\varphi\left(\boldsymbol{x}^{(k)}\right)}{\partial x_1},\frac{\partial\varphi\left(\boldsymbol{x}^{(k)}\right)}{\partial x_2},\cdots,\frac{\partial\varphi\left(\boldsymbol{x}^{(k)}\right)}{\partial x_n}\right)^{\mathrm{T}}$$

由 $\nabla\varphi(x)=\boldsymbol{A}\boldsymbol{x}^{(k)}-\boldsymbol{b}$ 有

$$\boldsymbol{p}^{(k)}=-\nabla\varphi(\boldsymbol{x}^{(k)})=-(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})=\boldsymbol{r}^{(k)}.$$

α_k 的计算公式

$$\alpha_k=\frac{\left(\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)}\right)}{\left(\boldsymbol{A}\boldsymbol{r}^{(k)},\boldsymbol{r}^{(k)}\right)},\tag{3}$$

最速下降法

于是

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{r}^{(k)}, \quad k = 0, 1, \cdots, \tag{4}$$

其中 $\boldsymbol{r}^{(k)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k)}$ 为剩余向量. 由 (3) 式和(4)式计算得到的向量序列 $\{\boldsymbol{x}^{(k)}\}$ 称为解线性方程组的最速下降法. 由于

$$\begin{aligned}
\left(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k)}\right) &= \left(\boldsymbol{b} - \boldsymbol{A}\left(\boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{r}^{(k)}\right), \boldsymbol{r}^{(k)}\right) \\
&= \left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right) - \alpha_k \left(\boldsymbol{A}\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right) \\
&= 0,
\end{aligned}$$

说明两个相邻的搜索方向是正交的.

还可证明由(3) 式和 (4) 式得到的 $\{\varphi(\boldsymbol{x}^{(k)})\}$ 是单调下降有下界的序列，它存在极限，满足

$$\lim_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^* = \boldsymbol{A}^{-1}\boldsymbol{b}$$

而且

$$\left\|\boldsymbol{x}^{(k)} - \boldsymbol{x}^*\right\|_{\boldsymbol{A}} \leqslant \left(\frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\right)^k \left\|\boldsymbol{x}^{(0)} - \boldsymbol{x}^*\right\|_{\boldsymbol{A}},$$

其中 λ_1, λ_n 分别为对称正定矩阵 $\boldsymbol{A}$ 的最大与最小特征值. $\|\boldsymbol{u}\|_{\boldsymbol{A}} = (\boldsymbol{A}\boldsymbol{u}, \boldsymbol{u})^{\frac{1}{2}}$，当 $\lambda_1 \gg \lambda_2$，时收敛是很慢的，而且当 $\left\|\boldsymbol{r}^{(k)}\right\|$ 很小时，由于舍入误差影响，计算将出现不稳定，这个算法实际中很少使用，需要寻找对整体而言下降更快的算法.

CG 方法是一种求解大型稀疏对称正定方程组十分有效的方法. 仍然选择一组搜索方向 $\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(1)}, \cdots$. 但它不再是具有正交性的 $\boldsymbol{r}^{(0)}, \boldsymbol{r}^{(1)}, \cdots$ 方向. 如果按方向 $\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(1)}, \cdots, \boldsymbol{p}^{(k-1)}$ 已进行 k 次一维搜索, 求得 $\boldsymbol{x}^{(k)}$, 下一步确定 $\boldsymbol{p}^{(k)}$ 方向能使 $\boldsymbol{x}^{(k+1)}$ 更快地求得 $\boldsymbol{x}^*$, 在 $\boldsymbol{p}^{(k)}$ 确定后. 仍按(1)式和(2)式的下降算法求得 α_k, 若已算出 $\boldsymbol{x}^{(k)}$ (不失一般性设 $\boldsymbol{x}^{(0)} = \mathbf{0}$), 则由(1)式有

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)},$$

$$\boldsymbol{x}^{(k)} = \alpha_0 \boldsymbol{p}^{(0)} + \alpha_1 \boldsymbol{p}^{(1)} + \cdots + \alpha_{k-1} \boldsymbol{p}^{(k-1)}$$

开始可取 $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$, 当 $k \geqslant 1$ 时确定 $\boldsymbol{p}^{(k)}$ 除了使

$$\varphi\left(\boldsymbol{x}^{(k+1)}\right) = \min_{\alpha} \varphi(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{p}^{(k)})$$

还希望 $\{\boldsymbol{p}^{(k)}\}$ 的选择使

$$\varphi\left(\boldsymbol{x}^{(k+1)}\right)=\min_{x\in\operatorname{span}\left(p^{(0)},p^{(1)},\ldots,p^{(k)},\right.}\varphi(\boldsymbol{x}), \tag{5}$$

这里 $x\in\operatorname{span}\left\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\cdots,\boldsymbol{p}^{(k)}\right\}$ 可表示为

$$\boldsymbol{x}=\boldsymbol{y}+\alpha\boldsymbol{p}^{(k)},\quad \boldsymbol{y}\in\operatorname{span}\left\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\cdots,\boldsymbol{p}^{(k-1)}\right\},\quad \alpha\in\mathbb{R}. \tag{6}$$

所以有

$$\varphi(\boldsymbol{x})=\varphi\left(\boldsymbol{y}+\alpha\boldsymbol{p}^{(k)}\right)=\varphi(\boldsymbol{y})+\alpha(\boldsymbol{A}\boldsymbol{y},\boldsymbol{p}^{(k)})-\alpha(\boldsymbol{b},\boldsymbol{p}^{(k)})+\frac{\alpha^2}{2}(\boldsymbol{A}\boldsymbol{p}^{(k)},\boldsymbol{p}^{(k)}) \tag{7}$$

(6)式表示在 $\boldsymbol{y}$ 已确定的情况下, 选 $\boldsymbol{p}^{(k)}$ 使 $\boldsymbol{x}$ 在整个空间 $\left\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\cdots,\boldsymbol{p}^{(k)}\right\}$ 中小, 为了使 (5)式极小化, 需要对 α 及 y 分别求极小.

在 (7) 式中出现的“交叉项”$(\boldsymbol{A}\boldsymbol{y}, \boldsymbol{p}^{(k)})$ 必须令它为 0，即

$$\left(\boldsymbol{A}\boldsymbol{y}, \boldsymbol{p}^{(k)}\right) = 0, \quad \forall \boldsymbol{y} \in \operatorname{span}\left\{\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(1)}, \cdots, \boldsymbol{p}^{(k-1)}\right\},$$

也就是 $\left(\boldsymbol{A}\boldsymbol{p}^{(j)}, \boldsymbol{p}^{(k)}\right) = 0, \quad j = 0, 1, \cdots, k-1.$

如果对 $k = 1, 2, \cdots$ 每步都如此选择 $\boldsymbol{p}^{(k)}$, 则它符合以下定义.

定义 2

设 A 对称正定, 若 $\mathbb{R}^n$ 中向量组 $\left\{\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(1)}, \cdots, \boldsymbol{p}^{(m)}\right\}$ 满足

$$\left(\boldsymbol{A}\boldsymbol{p}^{(i)}, \boldsymbol{p}^{(j)}\right) = 0, \quad i \neq j, i, j = 0, 1, \cdots, m,$$

则称它为 $\mathbb{R}^n$ 中一个 $\boldsymbol{A}$-共轭向量组或称 $\boldsymbol{A}$-正交向量组.

显然, 当 $m<n$ 时, 不含零向量的 $\boldsymbol{A}$-共轭向量组线性无关, 当 $\boldsymbol{A}=\boldsymbol{I}$ 时 $\boldsymbol{A}$-共轭性就是一般的正交性.

若取 $\left\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\cdots\right\}$ 是 $\boldsymbol{A}$-共轭的, 考虑(5) 式的解, $\boldsymbol{p}^{(k)}$ 使 (7) 式中 $\left(\boldsymbol{A}\boldsymbol{y},\boldsymbol{p}^{(k)}\right)=0$ 于是问题 (5)可分离为两个极小问题, 由 (7)式可得

$$
\begin{aligned}
\min_{x\in\operatorname{span}\left\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\ldots,\boldsymbol{p}^{(k)}\right\}}\varphi(x) &= \min_{\alpha,y}\varphi\left(\boldsymbol{y}+\alpha\boldsymbol{p}^{(k)}\right)\\
&= \min_{y}\varphi(\boldsymbol{y})+\min_{\alpha}\left[\frac{\alpha^2}{2}(\boldsymbol{A}\boldsymbol{p}^{(k)},\boldsymbol{p}^{(k)})+\alpha(\boldsymbol{A}\boldsymbol{y},\boldsymbol{p}^{(k)})-\alpha(\boldsymbol{b},\boldsymbol{p}^{(k)})\right]
\end{aligned}
$$

第一个极小 $\boldsymbol{y}\in\operatorname{span}\left\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\cdots,\boldsymbol{p}^{(k-1)}\right\}$ 的解 $\boldsymbol{y}=\boldsymbol{x}^{(k)}$

第二个极小就是 (1)式的极小, 由 $\boldsymbol{r}^{(k)}=\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^{(k)}$ 及 (2)式得

$$
\alpha_k=\frac{\left(\boldsymbol{r}^{(k)},\boldsymbol{p}^{(k)}\right)}{\left(\boldsymbol{A}\boldsymbol{p}^{(k)},\boldsymbol{p}^{(k)}\right)} \tag{8}
$$

CG 法中向量组 $\{\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(1)}, \cdots\}$ 的选择, 可令 $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}, \boldsymbol{p}^{(k)}$ 选为 $\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(1)}, \cdots, \boldsymbol{p}^{(k-1)}$ 的 $\boldsymbol{A}$-轭, 它并不唯一, 可选为 $\boldsymbol{r}^{(k)}$ 与 $\boldsymbol{p}^{(k-1)}$ 的线性组合, 不妨设

$$\boldsymbol{p}^{(k)} = \boldsymbol{r}^{(k)} + \beta_{k-1}\boldsymbol{p}^{(k-1)}, \tag{9}$$

利用 $(\boldsymbol{p}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{k-1}) = 0$, 可定出

$$\beta_{k-1} = -\frac{\left(\boldsymbol{r}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{(k-1)}\right)}{\left(\boldsymbol{p}^{(k-1)}, \boldsymbol{A}\boldsymbol{p}^{(k-1)}\right)}, \tag{10}$$

这样由(9)式和(10)式得到的 $\boldsymbol{p}^{(k)}$ 与 $\boldsymbol{p}^{(k-1)}$ 是 $\boldsymbol{A}$-共轭的.

根据以上分析, 取 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n, \boldsymbol{r}^{(0)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(0)}, \boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$ 可按(1) 式和(8) 式求得 $\alpha_0, \boldsymbol{x}^{(1)}$, 再由(9) 式和 (10)式求得 $\beta_0, \boldsymbol{p}^{(1)}$, 从而可得到序列 $\{\boldsymbol{x}^{(k)}\}$, 这就是 CG 算法.

下面对(8) 式作进一步简化. 由

$$\boldsymbol{r}^{(k+1)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(k+1)} = \boldsymbol{r}^{(k)} - \alpha_k \boldsymbol{A}\boldsymbol{p}^{(k)} \tag{11}$$

有

$$\left(\boldsymbol{r}^{(k+1)}, \boldsymbol{p}^{(k)}\right) = \left(\boldsymbol{r}^{(k)}, \boldsymbol{p}^{(k)}\right) - \alpha_k \left(\boldsymbol{A}\boldsymbol{p}^{(k)}, \boldsymbol{p}^{(k)}\right) = 0$$

$$\left(\boldsymbol{r}^{(k)}, \boldsymbol{p}^{(k)}\right) = \left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)} + \beta_{k-1}\boldsymbol{p}^{(k-1)}\right) = \left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right)$$

再代回(8), 有

$$\alpha_k = \frac{\left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right)}{\left(\boldsymbol{p}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right)} \tag{12}$$

由此看出, 当 $\boldsymbol{r}^{(k)} \neq \boldsymbol{0}$ 时, $\alpha_k > 0$.

定理 3

由 (1)式、(9)-(12) 式组成的 CG 算法得到的序列 $\{\boldsymbol{r}^{(k)}\}$ 及 $\{\boldsymbol{p}^{(k)}\}$ 以下性质:

(1) $\left(\boldsymbol{r}^{(i)}, \boldsymbol{r}^{(j)}\right) = 0 (i \neq j)$, 即 $\{\boldsymbol{r}^{(k)}\}$ 构成 $\mathbb{R}^n$ 中的正交向量组.

(2) $\left(\boldsymbol{A}\boldsymbol{p}^{(i)}, \boldsymbol{p}^{(j)}\right) = \left(\boldsymbol{p}^{(i)}, \boldsymbol{A}\boldsymbol{p}^{(j)}\right) = 0 (i \neq j)$, 即 $\{\boldsymbol{p}^{(k)}\}$ 为一个 $\boldsymbol{A}$-共轭向量组.

根据定理的证明还可化简 β_k 的计算, 由(10)

$$\begin{aligned} \beta_k &= -\frac{\left(\boldsymbol{r}^{(k+1)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right)}{\left(\boldsymbol{p}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right)} = \frac{-\left(\boldsymbol{r}^{(k+1)}, \alpha_k^{-1}\left(\boldsymbol{r}^{(k)} - \boldsymbol{r}^{(k+1)}\right)\right)}{\left(\boldsymbol{r}^{(k)} + \beta_{k-1}\boldsymbol{p}^{(k-1)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right)} \\ &= \frac{\left(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k+1)}\right)}{\alpha_k\left(\boldsymbol{r}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right)} = \frac{\left(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k+1)}\right)}{\left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right)}. \end{aligned} \tag{13}$$

此可见, 若 $\boldsymbol{r}^{(k+1)} \neq \mathbf{0}$, 则 $\beta_k > 0$. 根据 (4.17) 式和 (4. 18) 式可将 CG 算法归纳如下.

CG 算法

(1) 任取 $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$, 计算 $\boldsymbol{r}^{(0)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(0)}$, 取 $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)}$.

(2) 对 $k = 0, 1, \cdots$, 计算

$$\alpha_k = \frac{\left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right)}{\left(\boldsymbol{p}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right)}$$

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

$$\boldsymbol{r}^{(k+1)} = \boldsymbol{r}^{(k)} - \alpha_k \boldsymbol{A}\boldsymbol{p}^{(k)}, \beta_k = \frac{\left(\boldsymbol{r}^{(k+1)}, \boldsymbol{r}^{(k+1)}\right)}{\left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right)}$$

$$\boldsymbol{p}^{(k+1)} = \boldsymbol{r}^{(k+1)} + \beta_k \boldsymbol{p}^{(k)}$$

(3) 若 $\boldsymbol{r}^{(k)} = \boldsymbol{0}$, 或 $\left(\boldsymbol{p}^{(k)}, \boldsymbol{A}\boldsymbol{p}^{(k)}\right) = 0$, 计算停止, 则 $\boldsymbol{x}^{(k)} = x^*$. 由于 $\boldsymbol{A}$ 正定, 故当 $(\boldsymbol{p}^{(k)}, \boldsymbol{A}p^{(k)}) = 0$ 时, $\boldsymbol{p}^{(k)} = \boldsymbol{0}$, 而 $\left(\boldsymbol{r}^{(k)}, \boldsymbol{r}^{(k)}\right) = \left(\boldsymbol{r}^{(k)}, \boldsymbol{p}^{(k)}\right) = 0$, 也即 $\boldsymbol{r}^{(k)} = \boldsymbol{0}$.

例 用 CG 法解线性方程组

$$\begin{cases} 3x_1 + x_2 = 5 \\ x_1 + 2x_2 = 5 \end{cases}$$

解 显然 $\boldsymbol{A} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ 是对称正定的. 取 $\boldsymbol{x}^{(0)} = (0,0)^{\mathrm{T}}$, 则 $\boldsymbol{p}^{(0)} = \boldsymbol{r}^{(0)} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^{(0)} = (5,5)^{\mathrm{T}}$

$$\alpha_0 = \frac{\left(\boldsymbol{r}^{(0)}, p^{(0)}\right)}{\left(\boldsymbol{A}\boldsymbol{p}^{(0)}, \boldsymbol{p}^{(0)}\right)} = \frac{2}{7}$$

$$\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha_0 \boldsymbol{p}^{(0)} = \left(\frac{10}{7}, \frac{10}{7}\right)^{\mathrm{T}}$$

$$\boldsymbol{r}^{(1)} = \boldsymbol{r}^{(0)} - \alpha_0 \boldsymbol{A}\boldsymbol{p}^{(0)} = \left(-\frac{5}{7}, \frac{5}{7}\right)^{\mathrm{T}}$$

$$\beta_0 = \frac{\left(\boldsymbol{r}^{(1)}, \boldsymbol{r}^{(1)}\right)}{\left(\boldsymbol{r}^{(0)}, \boldsymbol{r}^{(0)}\right)} = \frac{1}{49}$$

$$\boldsymbol{p}^{(1)} = \boldsymbol{r}^{(1)} + \beta_0 \boldsymbol{p}^{(0)} = \left(-\frac{30}{49}, \frac{40}{49}\right)^{\mathrm{T}}$$

类似可以算出 $\alpha_1 = \frac{7}{10}, \boldsymbol{x}^{(2)} = (1, 2)^{\mathrm{T}}$ 为方程的解.