

第四讲 线性方程组迭代解法



目录

4.1 定常迭代法

4.2 收敛性分析

4.3 经典迭代法的收敛性

4.4 共轭梯度法

<https://math.ecnu.edu.cn/~jypan/Teaching/SC>

4-4 | 共轭梯度法

4.4 共轭梯度法

4.4.1 最速下降法

4.4.2 共轭梯度法

<https://math.ecnu.edu.cn/~jypan/Teaching/SC>

(预处理) 共轭梯度法是当前求解 **大规模 (稀疏) 对称正定** 方程组的首选方法

线性方程组与二次规划问题

定理 4.13 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 则 x_* 是 $Ax = b$ 的解当且仅当 x_* 是**最小化问题**

$$\min_{x \in \mathbb{R}^n} \quad \Phi(x) \triangleq \frac{1}{2}x^\top Ax - b^\top x$$

的解, 即 $\Phi(x)$ 的最小值点.

(板书)

线性方程组与二次规划问题

定理 4.13 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 则 x_* 是 $Ax = b$ 的解当且仅当 x_* 是**最小化问题**

$$\min_{x \in \mathbb{R}^n} \quad \Phi(x) \triangleq \frac{1}{2}x^\top Ax - b^\top x$$

的解, 即 $\Phi(x)$ 的最小值点.

(板书)

证明概要. **必要性**. 设 x_* 是 $Ax = b$ 的解, 即 $Ax_* = b$, 并注意到 A 对称正定, 则

$$\begin{aligned}\Phi(x) - \Phi(x_*) &= \frac{1}{2}x^\top Ax - b^\top x - \frac{1}{2}x_*^\top Ax_* + b^\top x_* \\ &= \frac{1}{2}x^\top Ax - x_*^\top Ax - \frac{1}{2}x_*^\top Ax_* + x_*^\top Ax_* \\ &= \frac{1}{2}(x - x_*)^\top A(x - x_*) \geq 0,\end{aligned}$$

等号当且仅当 $x - x_* = 0$ 时成立. 因此 x_* 是 $\Phi(x)$ 的唯一最小值点.

证明概要. **充分性**. 设 x_* 是 $\Phi(x)$ 的最小值点, 则

$$\begin{aligned}\Phi(x) - \Phi(x_*) &= \frac{1}{2}(x - x_*)^\top A(x - x_*) + x^\top Ax_* - x_*^\top Ax_* - b^\top x + b^\top x_* \\ &= \frac{1}{2}(x - x_*)^\top A(x - x_*) - (x - x_*)^\top (b - Ax_*).\end{aligned}$$

记 $r \triangleq b - Ax_*$, 如果 $r \neq 0$, 则取 $x = x_* + \alpha r$, 其中 $\alpha > 0$, 于是

$$\Phi(x) - \Phi(x_*) = \frac{\alpha^2}{2} r^\top Ar - \alpha \|r\|_2.$$

当 α 足够小时, 上式右端小于 0, 这与 x_* 是 $\Phi(x)$ 的最小值点矛盾. 所以 $r = 0$, 即 $Ax_* = b$.

证明概要. **充分性**. 设 x_* 是 $\Phi(x)$ 的最小值点, 则

$$\begin{aligned}\Phi(x) - \Phi(x_*) &= \frac{1}{2}(x - x_*)^\top A(x - x_*) + x^\top Ax_* - x_*^\top Ax_* - b^\top x + b^\top x_* \\ &= \frac{1}{2}(x - x_*)^\top A(x - x_*) - (x - x_*)^\top (b - Ax_*).\end{aligned}$$

记 $r \triangleq b - Ax_*$, 如果 $r \neq 0$, 则取 $x = x_* + \alpha r$, 其中 $\alpha > 0$, 于是

$$\Phi(x) - \Phi(x_*) = \frac{\alpha^2}{2} r^\top A r - \alpha \|r\|_2.$$

当 α 足够小时, 上式右端小于 0, 这与 x_* 是 $\Phi(x)$ 的最小值点矛盾. 所以 $r = 0$, 即 $Ax_* = b$.

求解 $Ax = b \iff$ 计算 $\Phi(x)$ 的最小值点

求解最小化问题的线搜索方法

线搜索方法

求解最小化问题 $\min_{x \in \mathbb{R}^n} \Phi(x)$ 的一类常用方法是 **线搜索方法**:

给定初始值 $x^{(0)}$, 确定一个使得 $\Phi(x)$ 变小的方向 $p_1 \in \mathbb{R}^n$, 即 **下降方向**, 满足 $\|p_1\| = 1$, 然后计算步长 $\alpha_1 \in \mathbb{R}_+$, 使得 $\Phi(x)$ 沿该下降方向达到最小, 即

$$x^{(1)} = x^{(0)} + \alpha_1 p_1, \quad \text{其中} \quad \alpha_1 = \operatorname{argmin}_{\alpha > 0} \Phi(x^{(0)} + \alpha p_1).$$

怎么计算步长

由于 A 对称正定, 直接计算可得

$$\begin{aligned}\Phi(x^{(0)} + \alpha p_1) &= \frac{1}{2}(x^{(0)} + \alpha p_1)^\top A(x^{(0)} + \alpha p_1) - b^\top(x^{(0)} + \alpha p_1) \\ &= \frac{1}{2}\alpha^2 p_1^\top A p_1 + \alpha p_1^\top A x^{(0)} + \frac{1}{2}(x^{(0)})^\top A x^{(0)} - b^\top x^{(0)} - \alpha b^\top p_1 \\ &= \frac{1}{2}\alpha^2 p_1^\top A p_1 - \alpha p_1^\top r_0 + \Phi(x^{(0)}),\end{aligned}$$

其中 $r_0 = b - Ax^{(0)}$.

怎么计算步长

由于 A 对称正定, 直接计算可得

$$\begin{aligned}\Phi(x^{(0)} + \alpha p_1) &= \frac{1}{2}(x^{(0)} + \alpha p_1)^\top A(x^{(0)} + \alpha p_1) - b^\top(x^{(0)} + \alpha p_1) \\ &= \frac{1}{2}\alpha^2 p_1^\top A p_1 + \alpha p_1^\top A x^{(0)} + \frac{1}{2}(x^{(0)})^\top A x^{(0)} - b^\top x^{(0)} - \alpha b^\top p_1 \\ &= \frac{1}{2}\alpha^2 p_1^\top A p_1 - \alpha p_1^\top r_0 + \Phi(x^{(0)}),\end{aligned}$$

其中 $r_0 = b - Ax^{(0)}$.

 这是关于 α 的一元二次函数, 且二次项系数为正, 所以

$$\alpha_1 = \operatorname{argmin}_{\alpha > 0} \Phi(x^{(0)} + \alpha p_1) = \frac{p_1^\top r_0}{p_1^\top A p_1}.$$

4-4-1 | 最速下降法

根据多元函数的 Taylor 展开公式, 我们有

$$\Phi(x) = \Phi(x^{(0)}) + (x - x^{(0)})^T \nabla \Phi(x^{(0)}) + o(\|x - x^{(0)}\|).$$

4-4-1

最速下降法

根据多元函数的 Taylor 展开公式, 我们有

$$\Phi(x) = \Phi(x^{(0)}) + (x - x^{(0)})^T \nabla \Phi(x^{(0)}) + o(\|x - x^{(0)}\|).$$

记 p 为 $x - x^{(0)}$ 所在的方向, 即 $p = \frac{x - x^{(0)}}{\|x - x^{(0)}\|}$, 则

$$(x - x^{(0)})^T \nabla \Phi(x^{(0)}) = \|x - x^{(0)}\| \cdot p^T \nabla \Phi(x^{(0)}).$$

4-4-1 | 最速下降法

根据多元函数的 Taylor 展开公式, 我们有

$$\Phi(x) = \Phi(x^{(0)}) + (x - x^{(0)})^T \nabla \Phi(x^{(0)}) + o(\|x - x^{(0)}\|).$$

记 p 为 $x - x^{(0)}$ 所在的方向, 即 $p = \frac{x - x^{(0)}}{\|x - x^{(0)}\|}$, 则

$$(x - x^{(0)})^T \nabla \Phi(x^{(0)}) = \|x - x^{(0)}\| \cdot p^T \nabla \Phi(x^{(0)}).$$

 因此, 只要满足 $p^T \nabla \Phi(x^{(0)}) < 0$, 则 p 就是 **下降方向**.

最速下降法

最速下降方向与最速下降法

下降速度最快的方向 $\rightarrow p^T \nabla \Phi(x^{(0)})$ 最小.



由 Cauchy-Schwarz 不等式可知

$$|p^T \nabla \Phi(x^{(0)})| \leq \|p\| \cdot \|\nabla \Phi(x^{(0)})\|,$$

等号当且仅当 p 与 $\nabla \Phi(x^{(0)})$ 共线时成立.

因此当 $p = -\frac{\nabla \Phi(x^{(0)})}{\|\nabla \Phi(x^{(0)})\|}$ 时, $p^T \nabla \Phi(x^{(0)})$ 达到最小.

我们称该下降方向为**最速下降方向**, 相应的线搜索方法为**最速下降法**.

由于最速下降方向就是 $\Phi(x)$ 在当前迭代点处的负梯度方向, 因此也称为**负梯度方向法**.

最速下降法的迭代格式



最速下降法的一般格式可表示为

$$x^{(k)} = x^{(k-1)} + \alpha_k p_k, \quad k = 1, 2, \dots,$$

$$\text{其中 } p_k = -\nabla \Phi(x^{(k-1)}) = b - Ax^{(k-1)} = r_{k-1}, \quad \alpha_k = \frac{p_k^\top r_{k-1}}{p_k^\top A p_k} = \frac{r_{k-1}^\top r_{k-1}}{r_{k-1}^\top A r_{k-1}}$$

最速下降法的迭代格式



最速下降法的一般格式可表示为

$$x^{(k)} = x^{(k-1)} + \alpha_k p_k, \quad k = 1, 2, \dots,$$

$$\text{其中 } p_k = -\nabla \Phi(x^{(k-1)}) = b - Ax^{(k-1)} = r_{k-1}, \quad \alpha_k = \frac{p_k^\top r_{k-1}}{p_k^\top A p_k} = \frac{r_{k-1}^\top r_{k-1}}{r_{k-1}^\top A r_{k-1}}$$

实际计算时, 我们无需对下降方向 p_k 进行单位化

最速下降法

算法 最速下降法 (Steepest Descent Algorithm)

- 1: 给定初值 $x^{(0)}$, 令 $k = 1$
 - 2: **while** not converge **do**
 - 3: 计算负梯度方向 (残量) $r_{k-1} = b - Ax^{(k-1)}$ 和步长 $\alpha_k = (r_{k-1}^T r_{k-1}) / (r_{k-1}^T A r_{k-1})$
 - 4: 计算 $x^{(k)} = x^{(k-1)} + \alpha_k p_k$, 其中 $p_k = r_{k-1}$
 - 5: 令 $k = k + 1$
 - 6: **end while**
-

最速下降法的收敛性

定理 4.14 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 则对任意初值 $x^{(0)}$, 最速下降法都收敛, 且

$$\frac{\|x^{(k+1)} - x_*\|_A}{\|x^{(k)} - x_*\|_A} \leq \frac{\kappa - 1}{\kappa + 1},$$

其中 κ 为 A 的谱条件数, 范数 $\|x\|_A \triangleq \sqrt{x^T A x}$.

(留作课外自习, 可以参见矩阵计算或最优化方法的相关教材)

最速下降法举例

例 4.8 用最速下降法求解 $Ax = b$, 其中 $A = \begin{bmatrix} 15 & 2 \\ 2 & 15 \end{bmatrix}$, $b = \begin{bmatrix} 17 \\ 17 \end{bmatrix}$, 初值 $x^{(0)} = \begin{bmatrix} -0.5 \\ 0 \end{bmatrix}$.

k	x	relres
1	[0.94896898, 1.06454864]	3.54e-02
2	[0.99757851, 0.99838567]	1.61e-03
3	[0.99991762, 1.00010420]	5.71e-05
4	[0.99999609, 0.99999739]	2.61e-06
5	[0.99999987, 1.00000017]	9.21e-08

“relres” 表示相对残量

$$\text{relres} = \frac{\|b - Ax^{(k)}\|_2}{\|b - Ax^{(0)}\|_2}$$

(Iter_steepest_descent_01.m)

最速下降法的局部最优性



由步长的选取方法可知, 最速下降法的迭代解具有下面的最优性

$$x^{(k)} = \operatorname{argmin}_{x \in x^{(k-1)} + \operatorname{span}\{p_k\}} \Phi(x)$$

最速下降法的局部最优性



由步长的选取方法可知, 最速下降法的迭代解具有下面的最优性

$$x^{(k)} = \operatorname{argmin}_{x \in x^{(k-1)} + \operatorname{span}\{p_k\}} \Phi(x)$$

由于舍弃了 Taylor 展开式中的高阶项, 因此最速下降方向是局部最优的

4-4-2

共轭梯度法

最速下降法的不足

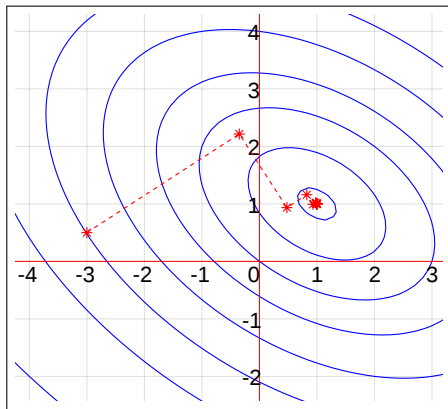
最速下降法每次选取下降方向时只考虑局部最优, 无法保证下降方向 $p_1, p_2, \dots$ 线性无关, 这意味着迭代过程可能会出现“之”字型, 即同一个下降方向可能会多次出现, 这就会导致收敛速度变得非常缓慢.

最速下降法的不足

例 4.9 用最速下降法求解 $Ax = b$, 其中 $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, 初值 $x^{(0)} = \begin{bmatrix} -3 \\ 0.5 \end{bmatrix}$.

k	x	relres
1	$[-0.3498, 2.2148]$	2.70e-01
2	$[0.4784, 0.9348]$	1.30e-01
3	$[0.8240, 1.1584]$	3.52e-02
4	$[0.9320, 0.9915]$	1.70e-02
$\vdots$	$\vdots$	$\vdots$
14	$[1.0000, 1.0000]$	6.41e-07

(Iter_steepest_descent_02.m)



A-共轭

为了改善最速下降法的收敛性质, 需要对下降方向进行改进, 由此, **共轭方向法**应运而生.

定义 4.7 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 若非零向量 $x, y \in \mathbb{R}^n$ 满足

$$y^{\top} Ax = 0,$$

则称 x 和 y 是 **A-共轭**的, 也称 **A-正交**的.

为什么 A -共轭

为什么选取 A -共轭的向量作为下降方向

设 $p_1, p_2, \dots, p_n$ 相互 A -共轭, 则 $p_1, p_2, \dots, p_n$ 线性无关, 因此构成 $\mathbb{R}^n$ 的一组基.
对任意给定的初值 $x^{(0)}$, 我们可以设

$$x_* - x^{(0)} = \alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_n p_n. \quad (4.1)$$

两边分别左乘 $p_k^\top A$, 可得

$$\alpha_k = \frac{p_k^\top A(x_* - x^{(0)})}{p_k^\top A p_k} = \frac{p_k^\top (b - A x^{(0)})}{p_k^\top A p_k}, \quad k = 1, 2, \dots, n.$$

这意味着可以通过下降方向 $p_1, p_2, \dots, p_n$ 和右端项 b , 以及 $x^{(0)}$ 将方程组的解表示出来.

为什么 A -共轭

为什么选取 A -共轭的向量作为下降方向

设 $p_1, p_2, \dots, p_n$ 相互 A -共轭, 则 $p_1, p_2, \dots, p_n$ 线性无关, 因此构成 $\mathbb{R}^n$ 的一组基.
对任意给定的初值 $x^{(0)}$, 我们可以设

$$x_* - x^{(0)} = \alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_n p_n. \quad (4.1)$$

两边分别左乘 $p_k^\top A$, 可得

$$\alpha_k = \frac{p_k^\top A(x_* - x^{(0)})}{p_k^\top A p_k} = \frac{p_k^\top (b - A x^{(0)})}{p_k^\top A p_k}, \quad k = 1, 2, \dots, n.$$

这意味着可以通过下降方向 $p_1, p_2, \dots, p_n$ 和右端项 b , 以及 $x^{(0)}$ 将方程组的解表示出来.



如果采用一组正交基作为下降方向, 也可以得到类似结论, 但 α_i 的表达式中包含 x_* .

共轭方向法

考虑到算法的实用性, 当 n 很大时, 通常不需计算所有的 α_k , 而只要计算前面部分, 得到一个满足精度要求的近似解即可. 因此, 我们把 $x_* = x^{(0)} + \alpha_1 p_1 + \cdots$ 看作是一个迭代过程:

$$x^{(k)} = x^{(k-1)} + \alpha_k p_k, \quad k = 1, 2, \dots$$

共轭方向法

考虑到算法的实用性, 当 n 很大时, 通常不需计算所有的 α_k , 而只要计算前面部分, 得到一个满足精度要求的近似解即可. 因此, 我们把 $x_* = x^{(0)} + \alpha_1 p_1 + \cdots$ 看作是一个迭代过程:

$$x^{(k)} = x^{(k-1)} + \alpha_k p_k, \quad k = 1, 2, \dots$$

共轭方向法

以相互 A -共轭的向量作为下降方向的线搜索方法就称为 **共轭方向法**.

共轭方向法

考虑到算法的实用性, 当 n 很大时, 通常不需计算所有的 α_k , 而只要计算前面部分, 得到一个满足精度要求的近似解即可. 因此, 我们把 $x_* = x^{(0)} + \alpha_1 p_1 + \cdots$ 看作是一个迭代过程:

$$x^{(k)} = x^{(k-1)} + \alpha_k p_k, \quad k = 1, 2, \dots$$

共轭方向法

以相互 A -共轭的向量作为下降方向的线搜索方法就称为 **共轭方向法**.

➤ 改写 α_k 的计算公式: 由于 $x^{(k-1)} - x^{(0)} = \alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_{k-1} p_{k-1}$, 所以 $p_k^\top A(x^{(k-1)} - x^{(0)}) = 0$, 即 $p_k^\top A x^{(0)} = p_k^\top A x^{(k-1)}$, 因此

$$\alpha_k = \frac{p_k^\top (b - A x^{(0)})}{p_k^\top A p_k} = \frac{p_k^\top (b - A x^{(k-1)})}{p_k^\top A p_k} = \frac{p_k^\top r_{k-1}}{p_k^\top A p_k}.$$

共轭方向法的全局最优性



相比于最速下降法的局部最优性:

$$x^{(k)} = \operatorname{argmin}_{x \in x^{(k-1)} + \operatorname{span}\{p_k\}} \Phi(x),$$

共轭方向法具有全局最优性.

定理 4.15 设 $x^{(k)}$ 是由共轭方向法得到的迭代解, 则

$$x^{(k)} = \operatorname{argmin}_{x \in x^{(0)} + \operatorname{span}\{p_1, p_2, \dots, p_k\}} \Phi(x).$$

(留作课外自习, 可以参见最优化方法的相关教材)

有限步终止

显然, 由 (4.1) 可知, 在不考虑舍入误差的情况下, 共轭方向法具有有限步终止性质.

定理 4.16 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 如果不考虑舍入误差, 则共轭方向法至多经过 n 次迭代后, 近似解 $x^{(k)}$ 就是精确解 x_* .

 在共轭方向法中, 虽然我们仍然称 p_k 为下降方向或搜索方向, 但我们不要求 p_k 是单位向量, 即不需要进行单位化, 因为这对算法的构造和实施没有任何影响.

怎么构造共轭搜索方向?

如何构造相互 A -共轭的下降方向

- 首先需要指出的是, 这样的向量组并不唯一.

事实上, 与 Gram-Schmidt 正交化过程类似, 任何一组线性无关的向量组都可以构造出一组相互 A -共轭的向量组.

- 目前最成功的是共轭梯度 (Conjugate Gradient) 下降方向.

它是在最速下降方向的基础上, 通过递推方法来构造的, 可以看作是将最速下降方向进行 A -共轭化, 或者是对最速下降方向的改进.

共轭梯度下降方向的构造

❶ 给定 $x^{(0)}$, 取 $\Phi(x)$ 在 $x^{(0)}$ 处的负梯度方向:

$$p_1 = -\nabla \Phi(x^{(0)}) = r_0 = b - Ax^{(0)}$$

共轭梯度下降方向的构造

① 给定 $x^{(0)}$, 取 $\Phi(x)$ 在 $x^{(0)}$ 处的负梯度方向:

$$p_1 = -\nabla\Phi(x^{(0)}) = r_0 = b - Ax^{(0)}$$

② 以 p_1 为下降方向, 计算近似解 $x^{(1)}$:

$$x^{(1)} = x^{(0)} + \alpha_1 p_1, \text{ 其中 } \alpha_1 = \frac{p_1^\top r_0}{p_1^\top A p_1}$$

共轭梯度下降方向的构造

- ① 给定 $x^{(0)}$, 取 $\Phi(x)$ 在 $x^{(0)}$ 处的负梯度方向:

$$p_1 = -\nabla\Phi(x^{(0)}) = r_0 = b - Ax^{(0)}$$

- ② 以 p_1 为下降方向, 计算近似解 $x^{(1)}$:

$$x^{(1)} = x^{(0)} + \alpha_1 p_1, \text{ 其中 } \alpha_1 = \frac{p_1^T r_0}{p_1^T A p_1}$$

- ③ 计算 $\Phi(x)$ 在 $x^{(1)}$ 处的负梯度方向: $-\nabla\Phi(x^{(1)}) = r_1 = b - Ax^{(1)} = r_0 - \alpha_1 A p_1$,

将其与 p_1 进行 A -共轭化, 得到下降方向 p_2 :

$$p_2 = r_1 + \beta_1 p_1, \text{ 其中 } \beta_1 = -\frac{r_1^T A p_1}{p_1^T A p_1}$$

共轭梯度下降方向的构造

- ① 给定 $x^{(0)}$, 取 $\Phi(x)$ 在 $x^{(0)}$ 处的负梯度方向:

$$p_1 = -\nabla\Phi(x^{(0)}) = r_0 = b - Ax^{(0)}$$

- ② 以 p_1 为下降方向, 计算近似解 $x^{(1)}$:

$$x^{(1)} = x^{(0)} + \alpha_1 p_1, \text{ 其中 } \alpha_1 = \frac{p_1^\top r_0}{p_1^\top A p_1}$$

- ③ 计算 $\Phi(x)$ 在 $x^{(1)}$ 处的负梯度方向: $-\nabla\Phi(x^{(1)}) = r_1 = b - Ax^{(1)} = r_0 - \alpha_1 A p_1$,

将其与 p_1 进行 A -共轭化, 得到下降方向 p_2 :

$$p_2 = r_1 + \beta_1 p_1, \text{ 其中 } \beta_1 = -\frac{r_1^\top A p_1}{p_1^\top A p_1}$$

- 以此类推, 我们就可以得到相互 A -共轭的下降方向 $p_1, p_2, \dots, p_k, \dots$
该过程同时也把近似解 $x^{(k)}$ 和残量 r_k 计算出来了.

怎么简化计算量?



在计算 p_{k+1} 时, 需要将 r_k 与所有 $p_1, p_2, \dots, p_k$ 进行 A -共轭化, 计算成本很大.

怎么简化计算量?



在计算 p_{k+1} 时, 需要将 r_k 与所有 $p_1, p_2, \dots, p_k$ 进行 A -共轭化, 计算成本很大.

事实上, 由于 A 对称正定, 我们只需将 r_k 与 p_k 进行 A -共轭化即可, 即

$$p_{k+1} = r_k + \beta_k p_k, \quad \text{其中} \quad \beta_k = -\frac{r_k^\top A p_k}{p_k^\top A p_k}.$$

这样不仅可以简化计算, 同时可以证明 p_{k+1} 与 $p_1, p_2, \dots, p_{k-1}$ 都 A -共轭.

共轭搜索方向的性质

定理 4.17 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 共轭梯度法迭代 m 步后 ($m < n$), 对任意 k ($k \leq m$) 有

$$(1) \quad r_k^\top r_i = 0, \quad i = 0, 1, 2, \dots, k-1;$$

$$(2) \quad r_k^\top p_i = 0, \quad i = 1, 2, \dots, k;$$

$$(3) \quad p_{k+1}^\top A p_i = 0, \quad i = 1, 2, \dots, k;$$

$$(4) \quad \text{span}\{r_0, r_1, \dots, r_k\} = \text{span}\{p_1, p_2, \dots, p_{k+1}\} = \text{span}\{r_0, A r_0, \dots, A^k r_0\}.$$

(留作课外自习, 可以参见矩阵计算或最优化方法的相关教材)

共轭梯度法迭代格式

于是共轭梯度法可描述为

$$\begin{cases} x^{(k)} = x^{(k-1)} + \alpha_k p_k, & \text{其中 } \alpha_k = \frac{p_k^\top r_{k-1}}{p_k^\top A p_k}, \\ r_k = b - A x^{(k)} = r_{k-1} - \alpha_k A p_k, \\ p_{k+1} = r_k + \beta_k p_k, & \text{其中 } \beta_k = -\frac{r_k^\top A p_k}{p_k^\top A p_k}, \end{cases} \quad k = 1, 2, \dots,$$

共轭梯度法迭代格式

于是共轭梯度法可描述为

$$\begin{cases} x^{(k)} = x^{(k-1)} + \alpha_k p_k, & \text{其中 } \alpha_k = \frac{p_k^\top r_{k-1}}{p_k^\top A p_k}, \\ r_k = b - A x^{(k)} = r_{k-1} - \alpha_k A p_k, \\ p_{k+1} = r_k + \beta_k p_k, & \text{其中 } \beta_k = -\frac{r_k^\top A p_k}{p_k^\top A p_k}, \end{cases} \quad k = 1, 2, \dots,$$



进一步简化计算: 利用 r_k 的正交性和 p_k 的 A -共轭性, 可得

推论 4.18 共轭梯度法中的 α_k 和 β_k 可以通过下面的公式计算

$$\alpha_k = \frac{r_{k-1}^\top r_{k-1}}{p_k^\top A p_k}, \quad \beta_k = \frac{r_k^\top r_k}{r_{k-1}^\top r_{k-1}}.$$

共轭梯度法的算法描述

算法 共轭梯度法 (Conjugate Gradient Algorithm)

- 1: 给定初值 $x^{(0)}$
 - 2: 计算 $r_0 = b - Ax^{(0)}$, 并令 $p_1 = r_0, k = 1$
 - 3: **while** not converge **do**
 - 4: 计算 $x^{(k)} = x^{(k-1)} + \alpha_k p_k$, 其中 $\alpha_k = \frac{r_{k-1}^\top r_{k-1}}{p_k^\top A p_k}$
 - 5: 计算 $r_k = r_{k-1} - \alpha_k A p_k$
 - 6: 计算 $p_{k+1} = r_k + \beta_k p_k$, 其中 $\beta_k = \frac{r_k^\top r_k}{r_{k-1}^\top r_{k-1}}$
 - 7: **end while**
-

共轭梯度法的收敛性

定理 4.19 设 $A \in \mathbb{R}^{n \times n}$ 对称正定, 则对任意初值 $x^{(0)}$, 共轭梯度法都收敛, 且

$$\frac{\|x^{(k)} - x_*\|_A}{\|x^{(0)} - x_*\|_A} \leq 2 \left(\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1} \right)^k,$$

其中 κ 为 A 的谱条件数, 范数 $\|x\|_A \triangleq \sqrt{x^T A x}$.

(留作课外自习, 可以参见矩阵计算的相关教材)

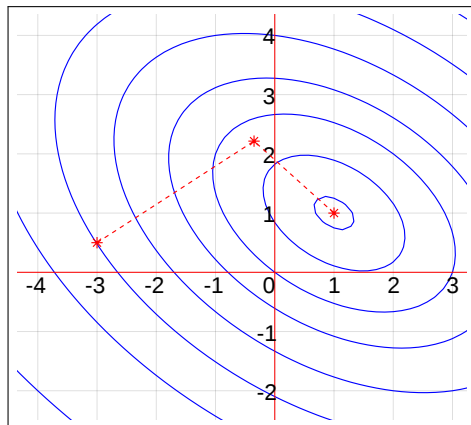
 由此可知, 共轭梯度法每次迭代的误差下降率为 $\frac{\sqrt{\kappa} - 1}{\sqrt{\kappa} + 1}$.

共轭梯度法举例（一）

例 4.10 用共轭梯度法求解 $Ax = b$, 其中 $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, $b = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$, 初值 $x^{(0)} = \begin{bmatrix} -3 \\ 0.5 \end{bmatrix}$.

k	x	relres
1	$[-0.3498, 2.2148]$	2.70e-01
2	$[1.0000, 1.0000]$	4.39e-17

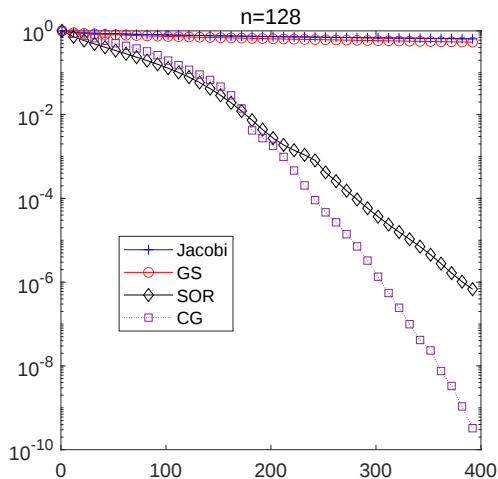
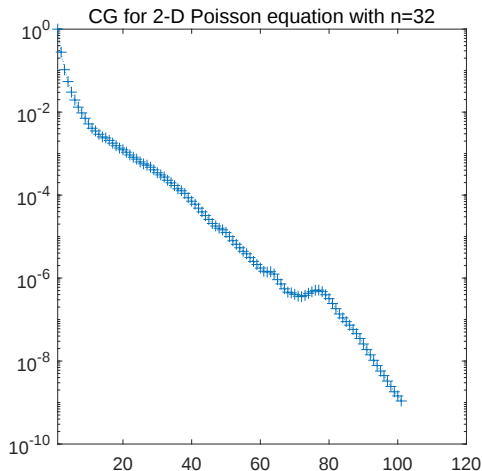
(Iter_CG_01.m)



共轭梯度法举例（二）

例 4.11 用共轭梯度法求解离散二维 Poisson 方程.

(Iter_CG_gallery.m), (Iter_CG_Jacobi_GS_SOR_Poisson2D.m)



谢谢
THANK YOU

