第七讲　Krylov 子空间迭代方法

子空间迭代方法的基本思想

在一个 维数较低的子空间 中寻找解析解的一个 最佳近似

子空间迭代方法的主要过程可以分解为下面三步:

(1) 寻找合适的子空间;

(2) 在该子空间中求 “最佳近似”;

(3) 若这个近似解满足精度要求, 则停止计算;

否则, 重新构造一个新的子空间, 并返回第 (2) 步.

两个关键问题

(1) 如果选择和更新子空间;

(2) 如何在给定的子空间中寻找 “最佳近似”.

- 关于第一个问题, 目前较成功的解决方案是 Krylov 子空间

1 Krylov 子空间

设 $A \in \mathbb{R}^{n\times n}$, $r \in \mathbb{R}^n$, 则由 A 和 r 生成的 m 维 **Krylov 子空间** 定义为

$$\mathcal{K}_m = \mathcal{K}_m(A, r) \triangleq \mathrm{span}\{r, Ar, A^2 r, \ldots, A^{m-1} r\}, \quad m \le n.$$

基本性质

- Krylov 子空间是嵌套的, 即: $\mathcal{K}_1 \subseteq \mathcal{K}_2 \subseteq \cdots \subseteq \mathcal{K}_m \subseteq \cdots$
- $\mathcal{K}_m(A, r) = \left\{ x = p(A) r : p \text{ 为次数不超过 } m-1 \text{ 的多项式} \right\}$

寻找 “最佳近似” $x^{(m)}$ 就转化为以下两个问题

(1) 寻找一组合适的基 $v_1, v_2, \ldots, v_m$; $\quad$ ($\rightarrow$ 单位正交)

(2) 求出 $x^{(m)}$ 在这组基下面的表出系数 $y^{(m)}$.

寻找基: Arnoldi 过程

最简单的一组基: $\{r, Ar, A^2r, \dots, A^{m-1}r\} \longmapsto$ 但非正交, 稳定性得不到保证.

Arnoldi 过程 $\to$ 将 $\{r, Ar, A^2r, \dots, A^{m-1}r\}$ 单位正交化

$v_1 = r/\|r\|_2$
for $j = 1$ to m **do**
  $z = Av_j$
  for $i = 1$ to j **do** % MGS 正交化过程
    $h_{i,j} = (v_i, z), \quad z = z - h_{i,j}v_i$
  end for
  $h_{j+1,j} = \|z\|_2$ % if $h_{j+1,j} = 0$, break, endif
  $v_{j+1} = z/h_{j+1,j}$
end for

Arnoldi 过程的矩阵表示

$$h_{j+1,j}v_{j+1} = Av_j - h_{1,j}v_1 - h_{2,j}v_2 - \cdots - h_{j,j}v_j,$$

即

$$Av_j = \sum_{i=1}^{j+1} h_{i,j}v_i = [v_1, \ldots, v_{j+1}] \begin{bmatrix} h_{1,j} \\ \vdots \\ h_{j+1,j} \end{bmatrix} = V_{m+1} \begin{bmatrix} h_{1,j} \\ \vdots \\ h_{j+1,j} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = V_{m+1}H_{m+1,m}(:, j),$$

其中 $V_m = [v_1, v_2, \ldots, v_m, v_{m+1}]$,

$$
H_{m+1,m} = \begin{bmatrix} h_{1,1} & h_{1,2} & h_{1,3} & \cdots & h_{1,m} \\ h_{2,1} & h_{2,2} & h_{2,3} & \cdots & h_{2,m} \\ & h_{3,2} & h_{3,3} & \cdots & h_{3,m} \\ & & h_{4,3} & \ddots & \vdots \\ & & & \ddots & h_{m,m} \\ & & & & h_{m+1,m} \end{bmatrix} \in \mathbb{R}^{(m+1)\times m}.
$$

所以有

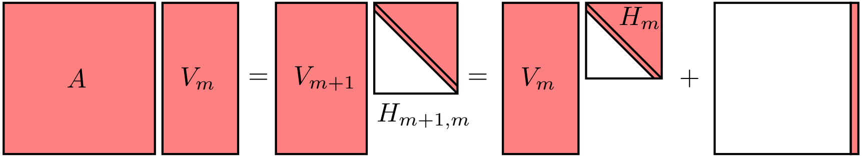
$$
\boxed{AV_m = V_{m+1}H_{m+1,m} = V_mH_m + h_{m+1,m}v_{m+1}e_m^{\mathsf{T}},} \tag{7.1}
$$

其中 $H_m = H_{m+1,m}(1:m, 1:m)$, $e_m = [0, \dots, 0, 1]^{\mathsf{T}} \in \mathbb{R}^m$.

由于 V_m 是列正交矩阵, 上式两边同乘 V_m^{T} 可得

$$\boxed{V_m^{\mathsf{T}} A V_m = H_m.} \tag{7.2}$$

等式 (7.1) 和 (7.2) 是 Arnoldi 过程的两个重要性质.



提前终止: 不变子空间

需要指出的是, 如果 $r, Ar, A^2r, \ldots, A^{m-1}r$ 线性相关, 则 Arnoldi 过程就会提前中断. 此时, 由 (7.1) 可知:

定理 如果 Arnoldi 过程在第 k 步时中断, 即 $h_{k+1,k} = 0$, 其中 $k < m$. 则有

$$AV_k = V_k H_k,$$

即 $\mathcal{K}_k$ 是 A 的一个不变子空间.

这对 GMRES, CG 等方法来说是一个非常受欢迎的性质.

对称情形: Lanczos 过程

若 A 对称, 则 H_m 为对称三对角矩阵. 为简单起见, 将其记为 T_m, 即

$$T_m = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{m-1} \\ & & \beta_{m-1} & \alpha_m \end{bmatrix}. \tag{7.3}$$

Lanczos 过程的性质和三项递推公式

$$AV_m = V_m T_m + \beta_m v_{m+1} e_m^{\mathsf{T}}$$

$$V_m^{\mathsf{T}} A V_m = T_m$$

$$\beta_j v_{j+1} = A v_j - \alpha_j v_j - \beta_{j-1} v_{j-1}, \quad j = 1, 2, \ldots .$$

基于三项递推公式的 Lanczos 过程

```
Set v_0 = 0 and β_0 = 0
v_1 = r/‖r‖_2
for j = 1 to m do
    z = A v_j
    α_j = (v_j, z)
    z = z − α_j v_j − β_{j−1} v_{j−1}
    β_j = ‖z‖_2
    if β_j = 0 then break,
    v_{j+1} = z/β_j
end for
```

Lanczos 向量的单位正交性

虽然在 Lanczos 过程中没有显式的正交化过程，但我们可以证明

$$\{v_1, v_2, \ldots, v_m\}$$

是单位正交的.

定理 设 $\{v_1, v_2, \ldots, v_m\}$ 是由 Lanczos 过程得到的向量组, 则

$$(v_i, v_j) = \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad i, j = 1, 2, \ldots, m.$$

(留作练习)

Krylov 子空间算法

为了充分利用迭代初值中所包含的有用信息, 我们在 仿射空间 $x^{(0)} + \mathcal{K}_m$ 中寻找 “最佳近似” 解.

Krylov 子空间算法的一般流程:

(1) 令 $m = 1$

(2) 定义 Krylov 子空间 $\mathcal{K}_m(A, r_0)$;

(3) 在 **仿射空间** $x^{(0)} + \mathcal{K}_m$ 中找出原问题的 “最佳近似” 解;

(4) 如果这个近似解满足精度要求, 则迭代结束;

否则令 $m \leftarrow m + 1$, 返回第 (2) 步.

Krylov 子空间算法的基本框架

```
选取初始向量 x^(0)
计算 r_0 = b − Ax^(0), v_1 = r_0/‖r_0‖_2
寻找“最佳近似”解: x^(1) ∈ x^(0) + K_1 = x^(0) + span{v_1}
if x^(1) 满足精度要求 then
    终止迭代
end if
for m = 2 to n do
    调用 Arnoldi 或 Lanczos 过程计算向量 v_m
    寻找“最佳近似”解: x^(m) ∈ x^(0) + K_m = x^(0) + span{v_1, ..., v_m}
    if x^(m) 满足精度要求 then
        终止迭代
    end if
end for
```

如何计算 $x^{(0)} + \mathcal{K}_m$ 中的 “最佳近似” 解

首先, 我们必须给出 “最佳” 的定义, 不同的定义会导致不同的算法.

什么是 “最佳”

(1) 直观定义: $\|x^{(m)} - x_*\|_2$ 达到最小 $\Longrightarrow$ 由于 x_* 不知道, 因此不实用

(2) $\|r_m\|_2 = \|b - Ax^{(m)}\|_2$ 最小: **MINRES** (对称), **GMRES** (非对称)

(3) A 对称正定, 极小化 $\|x^{(m)} - x_*\|_A \rightarrow$ **CG** (共轭梯度法)

本讲主要介绍 **GMRES** 方法和 **CG** 方法.

2 GMRES 算法

GMRES 算法是目前求解**非对称线性方程组**的最常用算法之一.

“**最佳近似**” 解的判别标准:

$$\text{使得} \quad \|r_m\|_2 = \|b - Ax^{(m)}\|_2 \quad \text{最小}$$

下面我们就根据这个最优性条件来导出 GMRES 算法.

GMRES 算法的推导

对任意向量 $x \in x^{(0)} + \mathcal{K}_m$, 可设 $x = x^{(0)} + V_m y$, 其中 $y \in \mathbb{R}^m$. 于是

$$r = b - Ax = r_0 - AV_m y = V_{m+1}\left(\beta e_1 - H_{m+1,m} y\right), \quad \text{这里 } \beta = \|r_0\|_2$$

由于 V_{m+1} 列正交, 所以

$$\|r\|_2 = \|V_{m+1}(\beta e_1 - H_{m+1,m} y)\|_2 = \|\beta e_1 - H_{m+1,m} y\|_2.$$

于是最优性条件就转化为

$$\boxed{y^{(m)} = \arg\min_{y \in \mathbb{R}^m} \|\beta e_1 - H_{m+1,m} y\|_2.} \tag{7.4}$$

可以用基于 Givens 变换的 QR 分解来求解.

GMRES 方法的基本框架

算法 1 GMRES 基本框架

选取初值 $x^{(0)}$, 停机标准 $\varepsilon > 0$, 以及最大迭代步数 IterMax

$r_0 = b - Ax^{(0)}$, $\beta = \|r_0\|_2$

$v_1 = r_0/\beta$

for $j = 1$ to IterMax **do**

    $w = Av_j$

    for $i = 1$ to j **do**   % Arnoldi 过程

        $h_{i,j} = (v_i, w)$

        $w = w - h_{i,j} v_i$

    end for

```
    h_{j+1,j} = ||w||_2
    if h_{j+1,j} = 0 then
        m = j, break
    end if
    v_{j+1} = w/h_{j+1,j}
    relres = ||r_j||_2/β      % 相对残量
    if relres< ε then      % 检测是否收敛
        m = j, break
    end if
end for
解最小二乘问题 (7.4), 得到 y^(m)
计算出近似解: x^(m) = x^(0) + V_m y^(m)
```

实施细节

需要解决下面两个问题

(1) 如何计算残量 $r_m \triangleq b - Ax^{(m)}$ 的范数? $\rightarrow$ 判断算法的收敛性

(2) 如何求解最小二乘问题

$$y^{(m)} = \arg \min_{y \in \mathbb{R}^m} \|\beta e_1 - H_{m+1,m} y\|_2$$

✍ 这两个问题可以同时处理解决

$\rightarrow$ 在求解最小二乘问题的过程中可以得到残量的范数.

最小二乘问题的求解

设 $H_{m+1,m}$ 的 QR 分解为

$$H_{m+1,m} = Q_{m+1}^{\mathsf{T}} R_{m+1,m},$$

其中 Q_{m+1} 是正交矩阵, $R_{m+1,m} \in \mathbb{R}^{(m+1)\times m}$ 是上三角矩阵. 则

$$\|\beta e_1 - H_{m+1,m} y\|_2 = \|\beta Q_{m+1} e_1 - R_{m+1,m} y\|_2 = \left\| \beta q_1 - \begin{bmatrix} R_m \\ 0 \end{bmatrix} y \right\|_2,$$

其中 $R_m \in \mathbb{R}^{m\times m}$ 非奇异 (假定 $H_{m+1,m}$ 不可约). 所以

$$y^{(m)} = \beta R_m^{-1} q_1(1:m),$$

$$\|r_m\|_2 = \|b - Ax^{(m)}\|_2 = \|\beta e_1 - H_{m+1,m} y^{(m)}\|_2 = \beta \cdot |q_1(m+1)|$$

其中 $q_1(m+1)$ 表示 q_1 的第 $m+1$ 个分量.

怎么高效计算 $H_{m+1,m}$ 的 QR 分解 $\rightarrow$ 递推方法

由于 $H_{m+1,m}$ 是上 Hessenberg 矩阵, 因此我们只需采用 Givens 变换.

(1) 当 $m=1$ 时, $H_{21}=\begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix}$, 构造 Givens 变换 G_1 使得

$$G_1 H_{21} = \begin{bmatrix} * \\ 0 \end{bmatrix} \triangleq R_{21}, \quad \text{即} \quad H_{21} = G_1^{\mathsf{T}} R_{21}.$$

(2) 假定存在 $G_1, G_2, \ldots, G_{m-1}$, 使得

$$(G_{m-1}\cdots G_2 G_1) H_{m,m-1} = R_{m,m-1},$$

即

$$H_{m,m-1} = (G_{m-1}\cdots G_2 G_1)^{\mathsf{T}} R_{m,m-1} \triangleq Q_m^{\mathsf{T}} R_{m,m-1}.$$

(✍ 为了书写方便, 这里假定 G_i 的维数自动扩张, 以满足矩阵乘积的需要)

(3) 考虑 $H_{m+1,m}$ 的 QR 分解. 易知

$$H_{m+1,m}=\begin{bmatrix} H_{m,m-1} & \boldsymbol{h}_m \\ 0 & h_{m+1,m} \end{bmatrix}, \text{其中 } \boldsymbol{h}_m=[h_{1m},h_{2m},\dots,h_{mm}]^{\mathsf{T}}.$$

所以有

$$\begin{bmatrix} Q_m & 0 \\ 0 & 1 \end{bmatrix} H_{m+1,m}=\begin{bmatrix} R_{m,m-1} & Q_m\boldsymbol{h}_m \\ 0 & h_{m+1,m} \end{bmatrix}=\begin{bmatrix} R_{m-1} & \tilde{\boldsymbol{h}}_{m-1} \\ 0 & \hat{h}_{mm} \\ 0 & h_{m+1,m} \end{bmatrix},$$

其中 $\tilde{\boldsymbol{h}}_{m-1}$ 是 $Q_m\boldsymbol{h}_m$ 的前 $m-1$ 个元素组成的向量, 而 $\hat{h}_{mm}$ 是其最后一个元素.

构造 Givens 变换 G_m:

$$G_m = \begin{bmatrix} I_{m-1} & 0 & 0 \\ 0 & c_m & s_m \\ 0 & -s_m & c_m \end{bmatrix} \in \mathbb{R}^{(m+1)\times(m+1)},$$

其中 $c_m = \dfrac{\hat{h}_{m,m}}{\tilde{h}_{m,m}}, s_m = \dfrac{h_{m+1,m}}{\tilde{h}_{m,m}}, \tilde{h}_{m,m} = \sqrt{\hat{h}_{m,m}^2 + h_{m+1,m}^2}$. 令

$$Q_{m+1} = G_m \begin{bmatrix} Q_m & 0 \\ 0 & 1 \end{bmatrix},$$

则

$$Q_{m+1}H_{m+1,m} = G_m \begin{bmatrix} R_{m-1} & \tilde{\boldsymbol{h}}_{m-1} \\ 0 & \hat{h}_{j,j} \\ 0 & h_{m+1,m} \end{bmatrix} = \begin{bmatrix} R_{m-1} & \tilde{\boldsymbol{h}}_{m-1} \\ 0 & \tilde{h}_{j,j} \\ 0 & 0 \end{bmatrix} \triangleq R_{m+1,m}$$

所以可得 $H_{m+1,m}$ 的 QR 分解 $H_{m+1,m} = Q_{m+1}^{\mathsf{T}} R_{m+1,m}$.

为了从 $H_{m,m-1}$ 的 QR 分解得到 $H_{m+1,m}$ 的 QR 分解, 我们需要

(1) 计算 $Q_m h_m$, 即将之前的 $m-1$ 个 Givens 变换作用到 $H_{m+1,m}$ 的最后一列的前 m 个元素上, 所以我们需要保留所有的 **Givens** 变换;

(2) 残量的范数: $\|r_m\|_2 = |\beta q_1(m+1)| = |\beta Q_{m+1}(m+1,1)|$, 即为

$$G_m G_{m-1} \cdots G_2 G_1 (\beta e_1)$$

的最后一个分量的绝对值. 由于在计算 r_{m-1} 时就已经计算出 $G_{m-1} \cdots G_2 G_1(\beta e_1)$, 因此这里只需做一次 **Givens** 变换即可;

(3) $y^{(m)}$ 的计算: 当相对残量满足精度要求时, 需要计算

$$y^{(m)} = R_m^{-1} q_1(1:m),$$

其中 $q_1 = G_m G_{m-1} \cdots G_2 G_1 (\beta e_1)$.

实用 GMRES 算法

算法 2 实用 GMRES 算法

给定初值 $x^{(0)}$, 停机标准 $\varepsilon > 0$, 最大迭代步数 IterMax
$r_0 = b - Ax^{(0)}$, $\beta = \|r_0\|_2$
if $\beta/\|b\|_2 < \varepsilon$ **then**
    停止计算, 输出近似解 $x^{(0)}$
end if
$v_1 = r_0/\beta, \quad \xi = \beta e_1$ % ξ 即为 q_1
for $j = 1$ to IterMax **do**
    $w = Av_j$
    for $i = 1$ to j **do**   % Arnoldi 过程
        $h_{i,j} = (v_i, w), \quad w = w - h_{i,j}v_i$
    end for
    $h_{j+1,j} = \|w\|_2$

if $h_{j+1,j} = 0$ **then** % 迭代中断

  $m = j$, break

end if

$v_{j+1} = w/h_{j+1,j}$

for $i = 1$ to $j-1$ **do** % 计算 $G_{j-1}\cdots G_2 G_1 H_{j+1,j}(1:j,j)$

$$\begin{bmatrix} h_{ij} \\ h_{i+1,j} \end{bmatrix} = \begin{bmatrix} c_i & s_i \\ -s_i & c_i \end{bmatrix} \begin{bmatrix} h_{ij} \\ h_{i+1,j} \end{bmatrix}$$

end for

if $|h_{jj}| > |h_{j+1,j}|$ **then** % 构造 Givens 变换 G_j

  $\tau = h_{j+1,j}/h_{jj}$, $c_j = 1/\sqrt{1+\tau^2}$, $s_j = c_j\tau$

else

  $\tau = h_{jj}/h_{j+1,j}$, $s_j = 1/\sqrt{1+\tau^2}$, $c_j = s_j\tau$

end if

$h_{jj} = c_j h_{jj} + s_j h_{j+1,j}$ % 计算 $G_j H_{j+1,j}(1:j,j)$

$h_{j+1,j} = 0$

$$\begin{bmatrix} \xi_j \\ \xi_{j+1} \end{bmatrix} = \begin{bmatrix} c_j & s_j \\ -s_j & c_j \end{bmatrix} \begin{bmatrix} \xi_j \\ 0 \end{bmatrix}$$ % 计算 $G_j(\beta G_{j-1} \cdots G_2 G_1 e_1)$

    relres = $|\xi_{j+1}|/\beta$ % 相对残量

    if relres $< \varepsilon$ **then** % 收敛检测

        $m = j$, break

    end if

end for

$m = j$

$y^{(m)} = H(1:m, 1:m) \backslash \xi(1:m)$ % 最小二乘问题, 回代求解

$x^{(m)} = x^{(0)} + V_m y^{(m)}$

if relres $< \varepsilon$ **then**

    输出近似解 $x^{(m)}$ 及相关信息

else

    输出算法失败信息

end if

GMRES 算法的中断

在上面的 GMRES 算法中, 当执行到某一步时有 $h_{k+1,k} = 0$, 则算法会中断 (breakdown). 如果出现这种中断, 则我们就找到了精确解.

定理 设 $A \in \mathbb{R}^{n\times n}$ 非奇异且 $r_0 \neq 0$. 若 $h_{i+1,i} \neq 0$, $i = 1, 2, \ldots, k-1$, 则 $h_{k+1,k} = 0$ 当且仅当 $x^{(k)}$ 是方程组的精确解. (不考虑舍入误差)

带重启的 GMRES 算法

由于随着迭代步数的增加, GMRES 算法的每一步所需的运算量和存储量都会越来越大. 因此当迭代步数很大时, GMRES 算法就不太实用.

重启技术

事先设定一个重启迭代步数 k, 如果 GMRES 达到这个迭代步数时仍不收敛, 则计算出 $x^{(0)} + \mathcal{K}_k$ 中的最佳近似解 $x^{(k)}$, 然后令 $x^{(0)} = x^{(k)}$, 重新开始新的 GMRES 迭代.

带重启的 GMRES 算法需要注意的问题

(1) 如何选取合适的重启步数 k?

一般只能依靠经验来选取, 如 $k = 20, 50$.

(2) 不带重启的 GMRES 算法能保证算法的收敛性, 但带重启的 GMRES 算法却无法保证, 有时可能出现停滞现象 (*stagnation*).

3 共轭梯度法 (CG)

“**最佳近似**” 解的判别标准:

$$\text{使得}\quad \|x^{(m)} - x_*\|_A \quad \text{或} \quad \|b - Ax^{(m)}\|_{A^{-1}} \quad \text{达到最小}$$

定理 设 A 对称正定, 则

$$x^{(m)} = \arg \min_{x \in x^{(0)} + \mathcal{K}_m} \|x - x_*\|_A$$

当且仅当

$$x^{(m)} \in x^{(0)} + \mathcal{K}_m \quad \text{且} \quad b - Ax^{(m)} \perp \mathcal{K}_m.$$

(板书)

Lanczos 过程的性质和三项递推公式

$$AV_m = V_m T_m + \beta_m v_{m+1} e_m^\mathsf{T}$$

$$V_m^\mathsf{T} A V_m = T_m$$

$$\beta_j v_{j+1} = A v_j - \alpha_j v_j - \beta_{j-1} v_{j-1}, \quad j = 1, 2, \ldots.$$

由于 V_m 的列向量构成 $\mathcal{K}_m$ 的一组基, 所以

$$\boxed{b - Ax^{(m)} \perp \mathcal{K}_m \Longleftrightarrow V_m^\mathsf{T}(b - Ax^{(m)}) = 0}$$

CG 算法的推导

设 $x^{(m)} = x^{(0)} + V_m z^{(m)}$, 其中 $z^{(m)} \in \mathbb{R}^m$, 则

$$0 = V_m^{\mathsf{T}}(b - Ax^{(m)}) = V_m^{\mathsf{T}}(r_0 - AV_m z^{(m)}) = \beta e_1 - T_m z^{(m)},$$

$$\boxed{z^{(m)} = T_m^{-1}(\beta e_1)}$$

设 T_m 的 $\mathrm{LDL}^{\mathsf{T}}$ 分解为 $T_m = L_m D_m L_m^{\mathsf{T}}$, 于是

$$x^{(m)} = x^{(0)} + V_m z^{(m)} = x^{(0)} + V_m T_m^{-1}(\beta e_1) = x^{(0)} + (V_m L_m^{-T})(\beta D_m^{-1} L_m^{-1} e_1)$$

如果 $x^{(m)}$ 满足精度要求, 则计算结束. 否则我们需要计算

$$x^{(m+1)} = x^{(0)} + V_{m+1} T_{m+1}^{-1}(\beta e_1) = x^{(0)} + (V_{m+1} L_{m+1}^{-T})(\beta D_{m+1}^{-1} L_{m+1}^{-1} e_1).$$

这里 $T_{m+1} = L_{m+1} D_{m+1} L_{m+1}^{\mathsf{T}}$.

利用 T_{m+1} 和 T_m 之间的关系, 可以得到 L_{m+1}, D_{m+1} 与 L_m 和 D_m 之间的递推关系. 并由此得到 $x^{(m+1)}$ 的递推公式.

$$\boxed{\begin{aligned} & x^{(m+1)} = x^{(m)} + \xi_{m+1} p_{m+1} \\ & r_{m+1} = r_m - \xi_{m+1} A p_{m+1} \\ & p_{m+1} = r_m + \mu_m p_m \end{aligned}}$$

其中

$$\mu_m = \frac{r_m^{\mathsf{T}} r_m}{r_{m-1}^{\mathsf{T}} r_{m-1}}, \quad \xi_{m+1} = \frac{r_m^{\mathsf{T}} r_m}{p_{m+1}^{\mathsf{T}} A p_{m+1}}.$$

共轭梯度法

CG : $Ax = b$

1: 给定初值 $x^{(0)}$

2: 计算 $r_0 = b - Ax^{(0)}$

3: $p_1 = r_0$

4: **for** $k = 1, 2, \dots$ **do**

5: $\quad x^{(k)} = x^{(k-1)} + \xi_k p_k, \quad \xi_k = \dfrac{(r_{k-1}, r_{k-1})}{(Ap_k, p_k)}$

6: $\quad r_k = r_{k-1} - \xi_k A p_k$

7: $\quad p_{k+1} = r_k + \mu_k p_k, \quad \mu_k = \dfrac{(r_k, r_k)}{(r_{k-1}, r_{k-1})}$

8: **end for**

CG 方法的两个重要性质:

(1) $r_1, r_2, \dots, r_m$ 相互正交;

(2) $p_1, p_2, \dots, p_m$ 相互 A-共轭 (A-正交), 即当 $i \neq j$ 时有 $p_i^{\mathsf{T}} A p_j = 0$.

4 收敛性分析

4.1 CG 方法的收敛性

最优性质

$$\|x^{(k)} - x_*\|_A = \min_{x \in x^{(0)} + \mathcal{K}_k(A, r_0)} \|x - x_*\|_A$$

定理 设 $A \in \mathbb{R}^{n\times n}$ 对称正定, 则由 CG 方法得到的近似解 $x^{(k)}$ 满足

$$\frac{\|x^{(k)} - x_*\|_A}{\|x^{(0)} - x_*\|_A} \leq \min_{q\in\mathbb{P}_k,\, q(0)=1} \max_{\lambda\in\sigma(A)} |q(\lambda)|,$$

其中 $\mathbb{P}_k$ 为所有次数不超过 k 的实系数多项式组成的集合.

利用 Chebyshev 多项式的性质, 我们可以得到下面的结论.

定理 设 $A \in \mathbb{R}^{n\times n}$ 对称正定, 且其最大与最小特征值分别为 $\lambda_{\max}$ 和 $\lambda_{\min}$, 则

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

其中 $\kappa(A) = \lambda_{\max}/\lambda_{\min}$ 是 A 的 (谱) 条件数.

4.2 GMRES 方法的收敛性

GMRES 最优性质

$$\|b - Ax^{(k)}\|_2 = \min_{x \in x^{(0)} + \mathcal{K}_k(A, r_0)} \|b - Ax\|_2$$

正规矩阵情形

定理 设 $A \in \mathbb{R}^{n\times n}$ 是正规矩阵, $x^{(m)}$ 是 GMRES 得到的近似解, 则

$$\frac{\|b - Ax^{(m)}\|_2}{\|r_0\|_2} \leq \min_{q\in\mathbb{P}_m,\ q(0)=1} \ \max_{\lambda\in\sigma(A)} |q(\lambda)|.$$

非正规情形

设 $A \in \mathbb{R}^{n\times n}$ 可对角化, 即 $\boxed{A = X\Lambda X^{-1}}$, 则

$$\|b - Ax^{(k)}\|_2 = \min_{x\in x^{(0)}+\mathcal{K}_k(A,r_0)} \|b - Ax\|_2 = \min_{q\in\mathbb{P}_k,\ q(0)=1} \|q(A)r_0\|_2.$$

定理 设 $A = X\Lambda X^{-1}$, 其中 $X \in \mathbb{C}^{n\times n}$ 非奇异, Λ 是对角矩阵, $x^{(k)}$ 是 GMRES 算法得到的近似解, 则

$$\begin{aligned}\frac{\|b - Ax^{(k)}\|_2}{\|r_0\|_2} &\leq \|X\|_2\|X^{-1}\|_2 \min_{q\in\mathbb{P}_k,\ q(0)=1} \max_{1\leq i\leq n} |q(\lambda_i)| \\ &= \kappa(X) \min_{q\in\mathbb{P}_k,\ q(0)=1} \max_{1\leq i\leq n} |q(\lambda_i)|,\end{aligned}$$

其中 $\kappa(X)$ 是 X 的谱条件数.

5 其它 Krylov 子空间方法

对称	CG (1952)	对称正定, 正交投影法 (Galerkin)
	MINRES (1975)	对称不定, 斜投影法 (Petrov-Galerkin)
	SYMMLQ (1975)	对称不定
	SQMR (1994)	对称不定
非对称	FOM (1981)	正交投影法, Arnoldi
	GMRES (1984)	斜投影法 (Petrov-Galerkin), Arnoldi
	BiCG (1976)	双正交 (biorthogonalization)
	QMR (1991)	双正交 (biorthogonalization)
	CGS (1989)	Transpose free
	BiCGStab (1992)	Transpose free, smoother convergence than CGS
	TFQMR (1993)	Transpose free, smoother convergence than CGS
	FGMRES (1993)	
正规方程	CGLS (1982)	最小二乘 (法方程)
	LSQR (1982)	最小二乘 (法方程)


谢 谢
THANK YOU