Krylov子空间算法习题参考答案

2018年1月9日

习题1（课件P29） 试用T_m的三对角结构证明：

$$\mathrm{v}_{j+1} \in span\{\mathrm{v}_{j-1}, \mathrm{v}_j, \mathrm{A}\mathrm{v}_j\}. \tag{1}$$

特别的请证明如下的三项递推关系：

$$\mathrm{A}\mathrm{v}_j = \beta_{j+1}\mathrm{v}_{j+1} + \alpha_j \mathrm{v}_j + \beta_j \mathrm{v}_{j-1}. \tag{2}$$

证明： 由$\mathrm{A}\mathrm{V}_m = \mathrm{V}_m\mathrm{T}_m$和$\mathrm{T}_m$的三对角结构可以得到：

$$\mathrm{A}\mathrm{v}_j = \beta_{j+1}\mathrm{v}_{j+1} + \alpha_j \mathrm{v}_j + \beta_j \mathrm{v}_{j-1}.$$

从而可以知道：

$$\mathrm{v}_{j+1} \in span\{\mathrm{v}_{j-1}, \mathrm{v}_j, \mathrm{A}\mathrm{v}_j\}.$$

习题2（课件P33） 假设$m = 1, 2, 3, \cdots, M$时，三对角的Lanczos 矩阵

$$\mathrm{T}_m = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{m-1} & \alpha_{m-1} & \beta_m \\ & & & \beta_m & \alpha_m \end{pmatrix} \in \mathbb{R}^{m\times m}$$

的LU分解$\mathrm{T}_m = \mathrm{L}_m\mathrm{U}_m$存在。则对于$m = 2, \cdots, m$有：

1) T_m的LU分解具有双对角形式：

$$T_m = L_m U_m = \begin{pmatrix} 1 & & & & \\ \lambda_2 & 1 & & & \\ & \ddots & \ddots & & \\ & & \lambda_{m-1} & 1 & \\ & & & \lambda_m & 1 \end{pmatrix} \begin{pmatrix} \omega_2 & & & & \\ \eta_2 & \omega_3 & & & \\ & \ddots & \ddots & & \\ & & \eta_{m-1} & \omega_m & \\ & & & & \eta_m \end{pmatrix}.$$

2) 验证$\lambda_m, \omega_m, \eta_m$的递推关系式，令$\lambda_1 = 0$：

$$\omega_m = \beta_m, \qquad \lambda_m = \frac{\beta_m}{\eta_{m-1}}, \qquad \eta_m = \alpha_m - \lambda_m \omega_m.$$

并证明L_m和U_m可递归得从L_{m-1}和U_{m-1}得到：

$$L_m = \left(\begin{array}{cc|c} L_{m-1} & & 0 \\ \hline 0^T & \lambda_m & 1 \end{array}\right), \quad U_m = \left(\begin{array}{c|c} U_{m-1} & \begin{matrix} 0 \\ \omega_m \end{matrix} \\ \hline 0^T & \eta_m \end{array}\right).$$

3) 证明，若有：

$$L = \left(\begin{array}{c|c} L' & 0 \\ \hline l^T & 1 \end{array}\right), \quad U = \left(\begin{array}{c|c} U' & y \\ \hline 0^T & \eta \end{array}\right).$$

则可以得到：

$$L^{-1} = \left(\begin{array}{c|c} L'^{-1} & 0 \\ \hline -l^T L'^{-1} & 1 \end{array}\right), \quad U = \left(\begin{array}{c|c} U'^{-1} & -\frac{1}{\eta} U'^{-1} y \\ \hline 0^T & \frac{1}{\eta} \end{array}\right).$$

证明： 经过简单的带入计算便可以得到题目1）、2）、3）中结果的正确性。

习题3（课件P69） 考虑如下的矩阵：$A = I + \alpha B$，其中B为反对成矩阵 $B^T = -B$：

1）证明

$$\frac{(Ax, x)}{(x, x)} = 1, \qquad \forall x \neq 0.$$

证明： 通过计算我们有：

$$(Bx, x) = (x, B^T x) = -(x, Bx) = -(Bx, x),$$

从而我们有：

$$(\mathrm{Bx}, \mathrm{x}) = 0, \qquad \forall \mathrm{x} \neq 0,$$

那么

$$(\mathrm{Ax}, \mathrm{x}) = (\mathrm{x}, \mathrm{x}) + \alpha(\mathrm{Bx}, \mathrm{x}) = (\mathrm{x}, \mathrm{x}), \qquad \forall \mathrm{x} \neq 0.$$

2）验证用Arnoldi 算法算出的A的Hessenberg矩阵具有如下的三对角形式：

$$\mathrm{H}_m = \begin{pmatrix} 1 & -\beta_2 & & & \\ \beta_2 & 1 & -\beta_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_{m-1} & 1 & -\beta_m \\ & & & \beta_m & 1 \end{pmatrix}.$$

证明： 我们有：$\mathrm{H}_m = \mathrm{V}_m^T \mathrm{A} \mathrm{V}_m$，其中$\mathrm{V}_m^T$各列是正交的，那么根据A的性质我们有：

$$\mathrm{H}_m = \mathrm{V}_m^T \mathrm{I} \mathrm{V}_m + \alpha \mathrm{V}_m^T \mathrm{B} \mathrm{V}_m,$$

而由B是一个反对成矩阵可知$\mathrm{P}_m = \mathrm{V}_m^T \mathrm{B} \mathrm{V}_m$也是反对称的，从而可以得到：

$$P_{ij} = -P_{ji}, \quad i \neq j, \qquad P_{jj} = 0.$$

而由于H_m是上Hessenberg矩阵且 $\mathrm{V}_m^T \mathrm{I} \mathrm{V}_m = \mathrm{I}_m$，那么$\mathrm{H}_m$是三对角矩阵，且对角线元素为1，非对角元素满足反对称性质。

3）利用2）中的结论解释用CG法求解$\mathrm{Ax} = \mathrm{b}$，即使A不对称，得到的残量也是相互正交的。

证明： 由2）中的结论可知，用Arnoldi 算法算出的A的Hessenberg矩阵是一个三对角矩阵。则对于此类矩阵，类似于正定对称阵，可构造对应的D-Lanczos算法，再加上正交和共轭条件就可以得到对应的CG算法，那么其残差满足：

$$\mathrm{r}_j = \sigma_m \mathrm{v}_{m+1}, \quad \sigma_m \in \mathbb{R},$$

从而我们有：

$$(\mathrm{r}_j, \mathrm{r}_j) = 0, \quad \forall i \neq j.$$

习题4（课件P73） 若$\mathrm{r}_0 \neq 0$，而$\bar{\mathrm{H}_m} \in \mathbb{R}^{(m+1)\times m}$是由Arnoldi过程得到的上Hessenberg阵：

$$\overline{\mathrm{H}}_m = \begin{pmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1m} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2m} \\ & h_{32} & h_{33} & \cdots & h_{3m} \\ & & \ddots & \ddots & \vdots \\ & & & h_{m,m-1} & h_{m,m} \\ & & & & h_{m+1,m} \end{pmatrix}$$

设$h_{j+1,j} \neq 0, \quad j = 1, \cdots, m-1$，试证：

1） $\mathcal{K}_m = span\{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_m\}$ 且$dim\mathcal{K}_m = m$；

证明： 可利用数学归纳法证明$\mathcal{K}_m = span\{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_m\}$。当$k=1$时，$\mathrm{v}_1 = \mathrm{r}_0/|\mathrm{r}_0|_2$，从而$\mathcal{K}_1 = span\{\mathrm{v}_1\}$。假设当$k < m$时均成立。那么当$k = m$时，由于$\mathcal{K}_{m-1} = span\{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{m-1}\}$，从而存在$\alpha_j, \quad j = 1, 2, \quad, m-1$使得：

$$\mathrm{A}^{m-2}\mathrm{r}_0 = \sum_{j=1}^{m-1} \alpha_j \mathrm{v}_j, \quad \alpha_{m-1} \neq 0.$$

则我们可以得到：

$$\mathrm{A}^{m-1}\mathrm{r}_0 = \sum_{j=1}^{m-1} \alpha_j \mathrm{A}\mathrm{v}_j,$$

而我们有：

$$\mathrm{A}\mathrm{v}_j = \sum_{i=1}^{j+1} h_{ij}\mathrm{v}_i, \qquad j = 1, \cdots, m-1,$$

从而我们有：

$$\mathrm{A}^{m-1}\mathrm{r}_0 = \sum_{j=1}^{m-1} \alpha_j \sum_{i=1}^{j+1} h_{ij}\mathrm{v}_i \in span\{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_m\},$$

另一方面，我们也可以得到：

$$\mathrm{v}_m \in \mathcal{K}_m.$$

因此，我们可以得到：

$$\mathcal{K}_m = span\{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_m\}.$$

由于v_j时相互正交的，从而$dim\mathcal{K}_m = m$。

2） $\overline{\mathrm{H}}_m$是列满秩的：$Rank\overline{\mathrm{H}}_m = m$；

证明：我们有$\mathrm{AV}_m = \mathrm{V}_m\mathrm{H}_m$，其中$A$非奇异，$V_m$的列相互正交。若$\overline{\mathrm{H}}_m$不满足列满秩，则$\mathrm{H}_m$奇异，从而有非零向量$\mathrm{y}_m$使得：

$$\mathrm{H}_m\mathrm{y}_m = \mathrm{V}_m^T\mathrm{AV}_m\mathrm{y}_m = 0.$$

由于V_m列满秩，从而由上式可以得到 $\mathrm{AV}_m\mathrm{y}_m = 0$；而A非奇异，从而$\mathrm{V}_m\mathrm{y}_m = 0$；而由于$\mathrm{V}_m$列满秩，从而可以得到$\mathrm{y}_m = 0$，从而与假设矛盾，那么$\mathrm{H}_m$非奇异，即 $\overline{\mathrm{H}}_m$是列满秩的：$Rank\overline{\mathrm{H}}_m = m$。

3）若$h_{m+1,m} = 0$，且 $\mathrm{A}^m\mathrm{r}_0 \in \mathcal{K}_m \Rightarrow \mathcal{K}_m = \mathcal{K}_{m+1} = \cdots = \mathcal{K}_n$。

证明：若$h_{m+1,m} = 0$，则利用Arnoldi方法求到的近似解即为原问题的精确解，即存在$m-1$次的多项式$q_{m-1}(x) = \sum_{k=0}^{m-1} c_k x^k$满足：

$$\mathrm{x}^* = \mathrm{x}_0 + q_{m-1}(\mathrm{A})\mathrm{r}_0.$$

而我们知道：

$$\mathrm{Ax}^* - \mathrm{Ax}_0 = \mathrm{b} - \mathrm{Ax}_0 = \mathrm{r}_0,$$

则我们有：

$$\mathrm{r}_0 - \mathrm{A}p_{m-1}(\mathrm{A})\mathrm{r}_0 = \mathrm{r}_0 - \sum_{k=1}^{m} c_{k-1}\mathrm{A}^k\mathrm{r}_0 = 0,$$

选取c_k中非零元素中下标最大的记为c_s，则上面等式两边各左乘A^{m-s-1}，则可以得到$\mathrm{A}^m\mathrm{r}_0 \in \mathcal{K}_m$，从而由上面的递归关系式可以进一步得到：

$$\mathcal{K}_m = \mathcal{K}_{m+1} = \cdots = \mathcal{K}_n.$$

习题5（课件P110） 假设在GMRES方法中，Arnoldi过程开始时取$\mathrm{v}_1 = \mathrm{Av}_0/\|\mathrm{Av}_0\|_2$，其中$\mathrm{v}_0 = \mathrm{r}_0$。此时可用正交化方法生成一组正交基$\{\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{m-1}\}$，则此时的近似解$\mathrm{x}_m$可用基底$\{\mathrm{v}_0, \mathrm{v}_1, \cdots, \mathrm{v}_{m-1}\}$表出。即$\mathrm{x}_m = \mathrm{x}_0 + \mathrm{V}_m\mathrm{y}_m$，其中$\mathrm{V}_m$的列向量为$\mathrm{v}_i$，　$0 \le i \le m-1$。

1）试证明求y_m的最小二乘问题的系数矩阵不再是Hessenberg矩阵，而是上三角阵。

证明： 根据题目中的要求，我们设：

$$\mathrm{v}_0 = \mathrm{r}_0, \qquad \mathrm{v}_1 = \frac{\mathrm{Av}_0}{\|\mathrm{Av}_0\|_2},$$

并依照下面的正交化过程构造v_j， $j = 2, 3, \cdots, m-1$：

$$\begin{aligned} \mathrm{w}_j = \mathrm{A}\mathrm{v}_j - \sum_{i=1}^{j} h_{i,j}\mathrm{v}_i, &\qquad h_{i,j} = (\mathrm{A}\mathrm{v}_j, \mathrm{v}_i), \\ h_{j+1,j} = \|\mathrm{w}_j\|_2, & \\ \mathrm{v}_{j+1} = \mathrm{w}_j / h_{j+1,j}. & \end{aligned}$$

则我们有：

$$\begin{aligned} \mathrm{A}\mathrm{v}_0 &= h_{0,1}\mathrm{v}_1, \\ \mathrm{A}\mathrm{v}_j &= \sum_{i=1}^{j+1} h_{j,i}\mathrm{v}_i, \quad j = 1, 2, \cdots, m-1. \end{aligned}$$

那么我们有：

$$\mathrm{A}\mathrm{V}_m = \overline{\mathrm{V}}_{m+1}\mathrm{H}_m, \tag{3}$$

其中，H_m是一个上三角矩阵，且满足$h_{ij} = h_{i,j-1}$， $j \geq i$；而$\overline{V}_{m+1}$则是V_{m+1}去掉第一列。

而我们得到的近似解为：

$$\mathrm{x}_m = \mathrm{x}_0 + \mathrm{V}_m\mathrm{y}_m,$$

使得它满足下面的最小二乘问题：

$$\|\mathrm{r}_m\|_2 = \|\mathrm{b} - \mathrm{A}\mathrm{x}_m\|_2 = \min_{\mathrm{x} \in \mathrm{x}_0 + \mathcal{K}_m} \|\mathrm{b} - \mathrm{A}\mathrm{x}\|_2,$$

而我们有：

$$\begin{aligned} \mathrm{b} - \mathrm{A}\mathrm{x}_m &= \mathrm{b} - \mathrm{A}\mathrm{x}_0 - \mathrm{A}\mathrm{V}_m\mathrm{y}_m \\ &= \mathrm{r}_0 - \overline{\mathrm{V}}_{m+1}\mathrm{H}_m\mathrm{y}_m, \end{aligned}$$

由于$\overline{\mathrm{V}}_{m+1}$的各列正交，从而上面的最小二乘问题等价于求解下面的问题：

$$\|\mathrm{r}_m\|_2 = \min_{\mathrm{x} \in \mathrm{x}_0 + \mathcal{K}_m} \|\mathrm{b} - \mathrm{A}\mathrm{x}\|_2 = \min_{\mathrm{y} \in \mathbb{R}^m} \|\overline{\mathrm{V}}_{m+1}^T\mathrm{r}_0 - \mathrm{H}_m\mathrm{y}\|_2. \tag{4}$$

由此可知，求y_m的最小二乘问题的系数矩阵不再是Hessenberg矩阵，而是上三角阵。

2）证明此时残量r_m和$\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{m-1}$正交。

证明： 根据1）中的结果可知，我们需要求解(4)中的最小二乘问题。实际上，H_m是上三角矩阵，对角元素均不为0，从而可逆，则(4)中的最小二乘问题的解为：

$$\mathrm{y}_m = \mathrm{H}_m^{-1} \overline{\mathrm{V}}_{m+1}^T \mathrm{r}_0.$$

那么此时残量为：

$$\begin{aligned}
\mathrm{r}_m = \mathrm{b} - \mathrm{A}\mathrm{x}_m &= \mathrm{r}_0 - \overline{\mathrm{V}}_{m+1} \mathrm{H}_m \mathrm{y}_m \\
&= \mathrm{r}_0 - \mathrm{A}\mathrm{V}_m \mathrm{H}_m^{-1} \overline{\mathrm{V}}_{m+1}^T \mathrm{r}_0 \\
&= \mathrm{r}_0 - \overline{\mathrm{V}}_{m+1} \mathrm{H}_m \mathrm{H}_m^{-1} \overline{\mathrm{V}}_{m+1}^T \mathrm{r}_0 \\
&= \mathrm{r}_0 - \overline{\mathrm{V}}_{m+1} \overline{\mathrm{V}}_{m+1}^T \mathrm{r}_0. \qquad (5)
\end{aligned}$$

则对于任意的$\mathrm{v}_k, k = 1, 2, \cdots, m-1$，我们有：

$$\mathrm{v}_k^T \overline{\mathrm{V}}_{m+1} \overline{\mathrm{V}}_{m+1}^T = \mathrm{e}_k^T \overline{\mathrm{V}}_{m+1}^T = (\overline{\mathrm{V}}_{m+1} \mathrm{e}_k)^T = \mathrm{v}_k^T,$$

从而可以得到：

$$\mathrm{v}_k^T \mathrm{r}_m = \mathrm{v}_k^T \mathrm{r}_0 - \mathrm{v}_k^T \mathrm{r}_0 = 0,$$

即此时残量r_m和$\mathrm{v}_1, \mathrm{v}_2, \cdots, \mathrm{v}_{m-1}$正交。

3）导出不需要计算近似解x_m就能计算残量r_m的范数的公式，并写出完整的GMRES算法。

证明： 根据2）中的结果(5)，我们可以得到：

$$\mathrm{r}_m = \mathrm{r}_0 - \sum_{j=1}^{m} (\mathrm{r}_0, \mathrm{v}_j) \mathrm{v}_j,$$

则根据$\mathrm{v}_j, \quad j = 1, 2, \cdots, m$间的正交性，我们可以得到：

$$\|\mathrm{r}_m\|_2 = \sqrt{\|\mathrm{r}_0\|_2 - \sum_{j=1}^{m} (\mathrm{r}_0, \mathrm{v}_j)^2}. \qquad (6)$$

GRMES算法：

1. Compute $\mathrm{r}_0 = \mathrm{b} - \mathrm{A}\mathrm{x}_0$, set $\mathrm{v}_0 = \mathrm{r}_0$;

2. Compute $h_{1,0} = \|\mathrm{Ar}_0\|_2$, denote $\mathrm{v}_1 = \mathrm{Ar}_0/h_{1,0}$. Then compute $rr_1 = (\mathrm{v}_0, \mathrm{v}_1)$;
3. for $j = 1, \cdots, m-1$ do
4. Compute $\mathrm{w}_j = \mathrm{Av}_j$
5. for $i = 1, \cdots, j$ do
6. $h_{i,j} = (\mathrm{w}_j, \mathrm{v}_i)$
7. $\mathrm{w}_j = \mathrm{w}_j - h_{i,j}\mathrm{v}_i$
8. end for
9. $h_{j+1,j} = \|\mathrm{w}_j\|_2$
10. if $h_{j+1,j} = 0$, set $m = j$, then go to 14;
11. $\mathrm{v}_{j+1} = \mathrm{w}_j/h_{j+1,j}$
12. Compute $rr_{j+1} = (\mathrm{v}_0, \mathrm{v}_{j+1})$
13. end for
14. Define the $m \times m$ upper triangular matrix $U_m = (u_{ij})$ with $u_{ij} = h_{i,j-1}$;
15. Compute y_m as the solution of $\mathrm{U}_m\mathrm{y}_m = \mathrm{Rr} = (rr_1, rr_2, \cdots, rr_m)^T$ and set $\mathrm{x}_m = \mathrm{x}_0 + \mathrm{V}^m\mathrm{y}_m$.