Iterative 3

王胤雅

25114020018

yinyawang25@m.fudan.edu.cn

2025 年 10 月 21 日

PROBLEM I In the Householder implementation of the Arnoldi algorithm, show the following points of detail:

(a) Q_{j+1} is unitary and its inverse is Q_{j+1}^T.

(b) $Q_{j+1}^T = P_1 P_2 \cdots P_{j+1}$.

(c) $Q_{j+1}^T e_i = v_i$ for $i < j$.

(d) $Q_{j+1} A V_m = V_{m+1}[e_1, e_2, \ldots, e_{j+1}]\bar{H}_m$, where e_i is the i-th column of the $n \times n$ identity matrix.

(e) The vectors $v_1, v_2, \ldots, v_j$ are orthonormal.

(f) The vectors $v_1, \ldots, v_j$ are equal to the Arnoldi vectors produced by the Gram-Schmidt version, except possibly for a scaling factor.

SOLUTION.

1. 因为每个 P_k 都是 Householder 反射矩阵，满足
$$P_k^T P_k = I,$$
即 P_k 是正交矩阵。正交矩阵的乘积仍为正交矩阵，因此
$$Q_{j+1} = P_{j+1} P_j \cdots P_1$$
也是正交矩阵。对于实矩阵而言，$\overline{H} = H$，故
$$Q_{j+1}^{-1} = Q_{j+1}^T.$$

2. 每个 P_k 都满足 $P_k^T = P_k$，因此
$$Q_{j+1}^T = (P_{j+1} \cdots P_1)^T = P_1 P_2 \cdots P_{j+1}.$$

3. 对 $i < j$，证明 $Q_{j+1}^T e_i = v_i$。

这里的 e_i 是标准基向量。首先说明 v_i 的意义：在 Householder 版本的 Arnoldi 过程中，逐步构造出向量序列 $\{v_i\}$，第 i 步时通过前 i 个反射矩阵将相应的向量对齐到第 i 个标准方向。

第一步中，构造 P_1 使得 $P_1 x = \alpha e_1$，于是 $v_1 = P_1 e_1 = Q_{j+1}^T e_1$ （当 $j+1 \geq 1$ 时一致）。

假设对所有 $i \leq t$，都有 $Q_{t+1}^T e_i = v_i$，则在第 $t+1$ 步，P_{t+1} 只作用于第 $t+1$ 及之后的分量，不改变前 t 个标准基向量，因此有

$$Q_{j+1}^T e_i = P_1 P_2 \cdots P_{j+1} e_i = P_1 \cdots P_i e_i = v_i.$$

因此命题对所有 $i < j$ 成立。

4. 证明关系式

$$Q_{j+1} A V_m = [e_1, e_2, \ldots, e_{m+1}] \bar{H}_m.$$

标准 Arnoldi 关系为

$$A V_m = V_{m+1} \bar{H}_m,$$

其中 $V_m \in \mathbb{R}^{n \times m}$ 为 Arnoldi 正交基，$V_{m+1} \in \mathbb{R}^{n \times (m+1)}$，而 $\bar{H}_m$ 为 $(m+1) \times m$ 上 Hessenberg 矩阵。由于在 (c) 中已知 $v_i = Q_{j+1}^T e_i$，可得

$$V_{m+1} = Q_{j+1}^T [e_1, e_2, \ldots, e_{m+1}].$$

两边左乘 Q_{j+1}，得到

$$Q_{j+1} A V_m = [e_1, e_2, \ldots, e_{m+1}] \bar{H}_m.$$

这就是所需的结果。

5. 证明向量 $v_1, v_2, \ldots, v_j$ 正交归一。由 (c) 有 $v_i = Q_{j+1}^T e_i$。由于 Q_{j+1}^T 是正交矩阵，它保持内积不变，因此

$$\langle v_i, v_\ell \rangle = \langle Q_{j+1}^T e_i, Q_{j+1}^T e_\ell \rangle = \langle e_i, e_\ell \rangle = \delta_{i\ell}.$$

于是 $\{v_1, \ldots, v_j\}$ 构成一组正交归一向量。

6. 证明 $\{v_1, \ldots, v_j\}$ 与 Gram–Schmidt 版本的 Arnoldi 向量一致，至多相差一个常数因子。Gram–Schmidt 版和 Householder 版 Arnoldi 都在同一 Krylov 子空间

$$\mathcal{K}_m(A, b) = \operatorname{span}\{b, Ab, \ldots, A^{m-1} b\}$$

中构造正交基，且都满足 Arnoldi 关系 $A V_m = V_{m+1} \bar{H}_m$。因此两种方法得到的每一步新向量方向相同，仅可能因反射方向不同而相差一个符号（或归一化常数）。由于在 (e) 中已证明 Householder 构造的 $\{v_i\}$ 也是单位正交的，所以两者在数值上最多相差 ± 1 的符号因子。故结论成立。

$\square$

PROBLEM II To derive the basic version of GMRES, we use the standard formula

$$\tilde{x} = x_0 + V \left(W^T A V \right)^{-1} W^T r_0, \tag{1}$$

where $V = V_m$ and $W = A V_m$.

SOLUTION. GMRES 要求选择 y 使得残差的二范数最小，即求解

$$y = \arg_{z\in\mathbb{R}^m} \min |0r_0 - AVz|0_2,$$

其中 $W \equiv AV$。

$$\min_{z\in\mathbb{R}^m} |0r_0 - Wz|0_2.$$

对平方范数对 z 求导并令梯度为零，

$$W^T W\, y = W^T r_0,$$

即

$$(AV)^T (AV)\, y = (AV)^T r_0.$$

由于 A 为非奇异矩阵，则 $W^T W$ 可逆，则正规方程的解为

$$y = (W^T W)^{-1} W^T r_0 = \big((AV)^T (AV)\big)^{-1} (AV)^T r_0.$$

将 y 代回近似解的表达式 $\tilde{x} = x_0 + Vy$，得到

$$\boxed{\tilde{x} = x_0 + V(W^T W)^{-1} W^T r_0 = x_0 + V\big((AV)^T (AV)\big)^{-1} (AV)^T r_0}$$

□

PROBLEM III Let a matrix A have the form

$$A = \begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}.$$

Assume that (full) GMRES is used to solve a linear system with the coefficient matrix A. What is the maximum number of steps that GMRES would require to converge?

SOLUTION. □

PROBLEM IV Consider a matrix of the form

$$A = I + \alpha B \tag{2}$$

where B is skew-symmetric (real), i.e., such that $B^T = -B$.

1. Show that $\dfrac{(Ax, x)}{(x, x)} = 1$ for all nonzero x.
2. Consider the Arnoldi process for A. Show that the resulting Hessenberg matrix will have the following tridiagonal form

$$H_m = \begin{pmatrix} 1 & -\eta_2 & & & \\ \eta_2 & 1 & -\eta_3 & & \\ & \eta_3 & 1 & \ddots & \\ & & \ddots & \ddots & -\eta_m \\ & & & \eta_m & 1 \end{pmatrix}.$$

3. Using the result of the previous question, explain why the CG algorithm applied as is to a linear system with the matrix A, which is nonsymmetric, will still yield residual vectors that are orthogonal to each other.

SOLUTION . $\square$