NLA 2021-2022
Eigenproblems

Martin Sombra

24 November 2021

Orthogonal iteration

It is generalization of the power method aimed to compute higher dimensional invariant subspaces

Fix $1 \leqslant r \leqslant n$. The *orthogonal iteration* writes down as

> Choose a unitary $n \times r$ matrix Q_0
> for $k = 0, 1, 2, \ldots$
> $\quad Y_{k+1} \leftarrow A\, Q_k$
> $\quad Y_{k+1} = Q_{k+1}\, R_{k+1}$ *(QR factorization)*

For $r = 1$ it coincides with the power method:

$$y_{k+1} = x_{k+1} \, \|A\, y_{k+1}\|_2$$

is the QR factorization of y_{k+1}

$\rightsquigarrow Q_{k+1} = x_{k+1}$ for all k

Error analysis

Suppose $A = S \Lambda S^{-1}$ diagonalizable with

$$S = [s_1 \ \cdots \ s_n] \quad \text{and} \quad \Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

and that

$$|\lambda_1| \geqslant \cdots \geqslant |\lambda_r| > |\lambda_{r-1}| \geqslant \cdots \geqslant |\lambda_n|$$

Then

$$\mathrm{span}(Q_k) = \mathrm{span}(Y_k) = \cdots = \mathrm{span}(A^k Q_0)$$

$\rightsquigarrow$ Q_k gives an orthonormal basis of the linear subspace $\mathrm{span}(A^k Q_0)$

For $k \to \infty$ it converges to a basis of the invariant subspace generated by the *dominant eigenvectors* $s_1, \ldots, s_r$:

$$\mathrm{span}(Q_k) = \mathrm{span}(A^k Q_0) \longrightarrow_{k \to \infty} \mathrm{span}(s_1, \ldots, s_r)$$

Funtoriality

The QR factorization is compatible with restricting columns:

$$\left[\begin{array}{c|c} A' & A \end{array}\right] = \left[\begin{array}{c|c} Q' & Q \end{array}\right] \left[\begin{array}{c|c} R' & R \\ \hline & \end{array}\right]$$

so that if $A = Q\,R$ is the QR factorization of A, then $A' = Q'\,R'$ is the QR factorization of A'

$\rightsquigarrow$ for $1 \leqslant s \leqslant r$ the first s columns of Q_k coincide with the orthogonal iteration of dimension s starting with the first s columns of Q_0

Setting $r = n$ runs the orthogonal iteration simultaneously for all intermediate dimensions

Convergence

Suppose for simplicity that

$$|\lambda_1| > \cdots > |\lambda_n|$$

Then for a "typical" unitary $n \times n$ matrix Q_0, for $k \to \infty$ we have that $Q_k \to Q$ unitary and $Q_k^* A\, Q_k \to T$ upper triangular, giving the Schur decomposition

$$A = Q\, T\, Q^*$$

Moreover

$$\|Q - Q_k\|, \|T - Q_k^* A\, Q_k\| \leqslant O\left(\max_i \frac{|\lambda_{i+1}|}{|\lambda_i|} \right)$$

The QR iteration

Set

$$T_k = Q_k^* \, A \, Q_k \quad \text{for } k = 0, 1, 2, \ldots .$$

The QR iteration arises when considering how to compute T_k directly from T_{k-1}

From the definition of Q_k and R_k and the relation $Q_{k-1} \, A = R_k \, Q_k$:

$$T_{k-1} = Q_{k-1}^* A \, Q_{k-1} = (Q_{k-1}^* Q_k) \, R_k \tag{1}$$

$$T_k = Q_k^* \, A \, Q_k = (Q_k^* \, A \, Q_{k-1}) \, Q_{k-1}^* Q_k = R_k \, (Q_{k-1}^* Q_k) \tag{2}$$

We have that $Q_{k-1}^* Q_k$ is unitary and (1) is the QR factorization of T_{k-1}

Drawbacks

- a single iteration costs $O(n^3)$ flops
- convergence (when it exists) is only linear

The real QR iteration

Matrices arising from applications have real entries! From now

$$\boxed{A \text{ real } n \times n \text{ matrix}}$$

The *real QR iteration* is defined by choosing an orthogonal $n \times n$ matrix Q_0 and setting

$$
\begin{aligned}
& H_0 \leftarrow Q_0^T A\, Q_0 \\
& \text{for } k = 1, 2, \ldots \\
& \qquad H_{k-1} = Q_k\, R_k \ \ (\textit{QR factorization}) \\
& \qquad H_k \leftarrow R_k\, Q_k
\end{aligned}
$$

The real Schur form

If the eigenvalues of A are not real, then T_k cannot converge to an upper triangular matrix

$\rightsquigarrow$ we content ourselves with convergence to the *real Schur form*:

$$A = Q\,T\,Q^T$$

with Q orthogonal and T block upper triangular with 1×1 and 2×2 diagonal blocks, that is

$$T = \begin{bmatrix} T_{1,1} & T_{1,2} & \cdots & T_{1,n} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & T_{n-1,n} \\ 0 & \cdots & 0 & T_{n,n} \end{bmatrix} \in \mathbb{R}^{n \times n}$$

where each $T_{i,i}$ is a 1×1 block or a 2×2 block with complex conjugate eigenvalues

Hessenberg reduction

To implement the QR iteration, we have to choose Q_0 carefully such that

$$T_0 = Q_0^T A Q_0$$

is an upper Hessenberg matrix, which would lower the complexity to $O(n^2)$ flops

Hessenberg reduction (cont.)

It is a variation of the QR factorization, and can be done with a sequence of Householder reflections

Let $n = 5$ and choose $P_1 = \begin{bmatrix} 1 & \mathbb{0} \\ \mathbb{0} & \widetilde{P}_1 \end{bmatrix}$ with $\widetilde{P}_1$ a Householder 4×4 matrix such that

$$P_1\,A = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}, \quad A_1 = P_1\,A\,P_1^T = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}$$

- P_1 leaves the first row of $P_1\,A$ unchanged
- P_1^T leaves the first column of $(P_1\,A)\,P_1^T$ unchanged, including the zeros

Hessenberg reduction (cont.)

Choose $P_2 = \begin{bmatrix} \mathbb{1}_2 & \mathbb{0} \\ \mathbb{0} & \widetilde{P}_2 \end{bmatrix}$ with $\widetilde{P}_2$ a 3×3 Householder such that

$$P_2\, A_1 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{bmatrix}, \quad A_2 = P_2\, A_1\, P_2^T = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{bmatrix}$$

- P_1 leaves the first and second rows of $P_2\, A_1$ unchanged
- P_2^T leaves the first and second columns of $(P_2\, A_2)\, P_2^T$ unchanged

Hessenberg reduction (cont.)

Finally choose $P_3 = \begin{bmatrix} \mathbb{1}_3 & \mathbb{0} \\ \mathbb{0} & \widetilde{P}_3 \end{bmatrix}$ with $\widetilde{P}_2$ a 2×2 Householder s.t.

$$P_3\, A_2 = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}, \quad A_3 = P_3\, A_2\, P_3^T = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \end{bmatrix}$$

This constructs an orthogonal $n \times n$ matrix

$$Q_0 = P_{n-2} \cdots P_1$$

such that $H_0 = Q_0^T A\, Q_0$ is Hessenberg

The complexity of this procedure is $5\, n^3 + O(n^2)$ flops

The Hessenberg form through the QR iteration

The Hessenberg form is preserved during the QR iteration: let

$$H = Q\,R \quad \text{and} \quad H_+ = R\,Q$$

with H Hessenberg

Since R is upper triangular, the j-th column of Q is a linear combination of the first columns of H: write

$$H = [h_1 \cdots h_n] \quad \text{and} \quad Q = [q_1 \cdots q_n]$$

Setting $S = R^{-1} = [s_{i,j}]_{i,j}$ we have that

$$q_j = \sum_{k=1}^{j} s_{k,j}\, h_k\,, \quad j = 1, \ldots, n$$

and so Q is Hessenberg

Similarly the j-th row of H_+ is a linear combination of the last j rows of Q, and so H_+ is Hessenberg