

QR iteration II and Power Method

Friday, November 22, 2024 10:57 AM

Class 38: November 22, 2024

Recall: Last time, we covered an introduction to a series of important results relevant to the problem of finding an orthonormal basis of eigenvectors $\{u_1, u_2, \dots, u_n\}$ and the corresponding eigenvalues $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$. This is equivalent to computing the eigendecomposition $A = U D U^*$:

- We defined the concepts of eigenvalue, eigenvector and eigenpair: (λ, x) is an eigenpair if x is a non-zero solution to $(A - \lambda I)x = 0$. λ is an eigenvalue if $(A - \lambda I)$ is singular, which is equivalent to $\det(A - \lambda I) = 0$.
- We defined two important tools: the characteristic polynomial $p_A(\lambda) = \det(A - \lambda I)$ (eigenvalues are the roots of p_A) and the eigenspace $E_{\{\lambda\}}$ and its dimension $\dim(E_{\{\lambda\}})$.
- We then introduced the concepts of **similarity** and **unitary similarity**. Eigenvalues (and eigenvectors modulo a change of basis) are preserved under similarity, which makes unitary similarity transformations a crucial tool to find our eigendecomposition.

We then discussed that an eigendecomposition doesn't always exist, but a Schur decomposition always does. That is: given $A \in \mathbb{C}^{n \times n}$ with complex (or real) entries, you can always find U unitary and T triangular such that $T = U^* A U$ (and so, $A = U T U^*$). The eigenvalues (spectrum) of A is then the diagonal of T .

We used Schur to sketch a proof of the "spectral theorems":

- If A is $n \times n$ and **Hermitian (symmetric)**, then all its eigenvalues are real and T is diagonal, meaning $D = U^* A U$ for D diagonal!
- If A is $n \times n$ and **Normal**, then there are complex matrix U and complex diagonal matrix D (T is again diagonal) such that $D = U^* A U$.

This told us very important information: An algorithm to approximately find a Schur decomposition (compute U and T) will leave us in the best possible shape: If A is **eigendecomposable**, then T will be diagonal and we will get our decomposition. If A is not eigendecomposable, we can "read off" the eigenvalues from $\text{diag}(T)$ and find whatever eigenvectors we can find.

Finally, we introduced the QR iteration as an algorithm based on the QR decomposition that produces iterates A_k and Q_k such that $A_k \rightarrow T$ and $U_k \rightarrow U$. Today, we will discuss the performance of this algorithm (it is painfully slow) and how to accelerate it so it is much faster.

QR iteration :

$$\text{Start } A_0 = A$$

while (algo has not converged)

$$[Q_k, R_k] = \text{qr}(A_k)$$

$$A_{k+1} = \underbrace{R_k}_{\text{ }} \underbrace{Q_k^*}_{\text{ }} A_k \underbrace{Q_k}_{\text{ }}$$

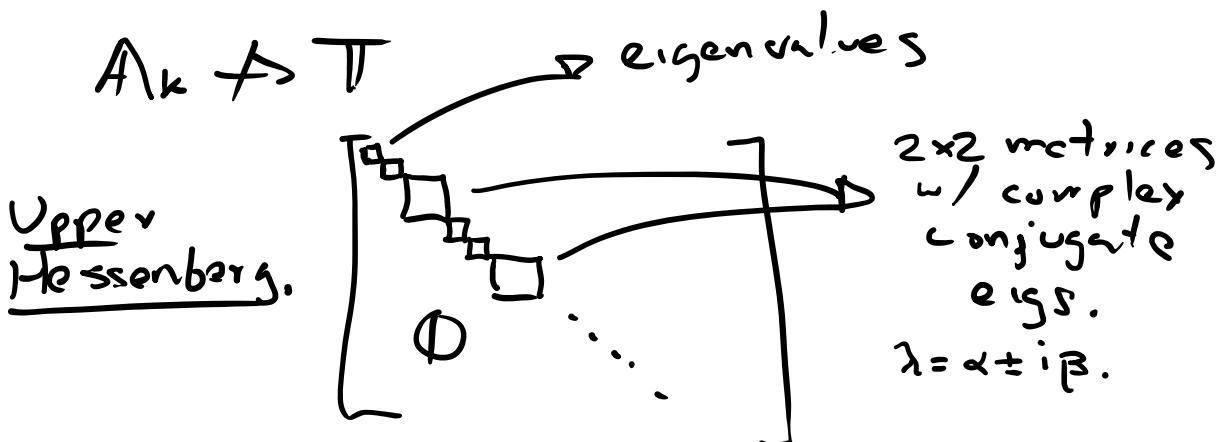
$$\dots \lim_{k \rightarrow \infty} A_k = \underbrace{\Pi}_{\text{ }} , \quad \lim_{k \rightarrow \infty} (Q_1 \dots Q_k) = \underbrace{U}_{\text{ }}$$

• Convergence QR iter is Linear

(\Rightarrow ratios of $\lambda_i / \lambda_{i+1}$ (order by abs value)).

- Seems to order them.

- Seems to order them.
- = Non-symmetric real matrices



Convergence / Stopping criteria

QR .terv: Cost per iter $O(n^3)$

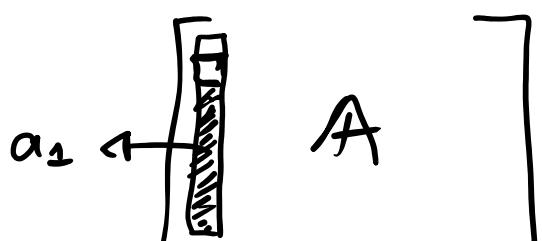
Large # (linear, slow)

Accelerations:

\rightarrow Cost per iter - $O(n^3)$.

- Maybe $A_0 = A$ is not a good idea.
 ↳ Pick $A_0 \sim A$ but $qr(A_0)$ is cheap?

Use householder reflectors, but as similarity transf.



$$\left[\begin{array}{c|ccccc} w_1 & & & & & \\ \hline & H_{11} & H_{12} & H_{13} & \dots & H_{1n} \end{array} \right]$$

Do H_{11}, \dots, H_{1n} from Householder QR, it does not work. $\rightarrow A$ is not triangular (has no zeroes).

\rightarrow Make A Upper-Hessenberg.

$$H_{11} = \begin{bmatrix} 1 & 0 \\ 0 & H_{w_1} \end{bmatrix} \quad H_{w_1} = I - 2w_1 w_1^*$$

$$A_0 = (H_{n-2} \cdots H_2 H_1) A (H_1 H_2 \cdots H_{n-2})$$

$$A_0 = \begin{bmatrix} \text{Upper triangular} & \\ \text{Zero} & \end{bmatrix} \rightarrow \text{Cost } qr \quad \underline{\mathcal{O}(n^2)}$$

What if A symmetric? $A_0 = \begin{bmatrix} \text{Upper triangular} & \\ \text{Symmetric} & \end{bmatrix}$
 \hookrightarrow Cost $qr \rightarrow \underline{\mathcal{O}(n)}$

Modified algo:

$$A_0 = \text{Upper Hess}(A) \quad \mathcal{O}(n^3)$$

$$\text{QR-step}(A_0) \rightarrow \mathcal{O}(n^2) \text{ p/stp}$$

(A sym - $\mathcal{O}(n)$ p/stp)

"

(A sym - O(n) p/st)

① "Deflation"

$$A_k = \begin{bmatrix} & & \\ & & \\ & & \ddots \\ & & 0 & \end{bmatrix}$$

① Borrow ideas from Power Methods

→ Shifted Inv. Power Method

↳ Variable shifts

$$[Q_k, P_k] = q \sqrt{(A - \mu_k I)}$$

cubic convergence?

One kind of shift for symmetric
↳ "Rayleigh shifts"

Different for non-symm - Wilkinson shifts.