Unitarily Similar matrices

Fact 15

Two matrices $A, B \in \mathbb{C}^{n\times n}$ are unitarily similar if there is a unitary matrix $U \in \mathbb{C}^{n\times n}$ such that $B = U^{-1}AU$.

Fact 16

If A, B, and U are all real, then U is orthogonal, and A and B are said to be orthogonally similar.

Fact 17

If $A = A^*$ and A is unitarily similar to B, then $B = B^*$. That is, Hermitian property is preserved under unitary similarity transformations.

Schur's Theorem

Let $A \in \mathbb{C}^{n \times n}$. Then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ and an upper triangular matrix $T \in \mathbb{C}^{n \times n}$ such that $T = U^* A U$. Equivalently, $A = UTU^*$, and it is called Schur decomposition of A. Proof on chalkboard.

Remarks on Schur's Theorem

- The main diagonal entries of T are the eigenvalues of A
  - Can we find unitary U such that A is unitarily similar to T?
- Proof of Schur's theorem (in Watkin's) book is non-constructive: does not give a way to find U
- Nevertheless, it gives a reason to hope that there may exist an algorithm to create a sequence: $A = A_0, A_1, \cdots$, that converges to upper triangular form
  - Hint: QR algorithm

Remarks on Schur's Theorem

- From the equation $T = U^* AU$: the first column of U is necessarily an eigenvector of A
  - In general, other columns of U are not necessarily eigenvectors of A

Fact 18

(Spectral Theorem for Hermitian Matrices) Suppose $A \in \mathbb{C}^{n \times n}$ is Hermitian. Then there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ and a diagonal matrix $D \in \mathbb{R}^{n \times n}$ such that $D = U^* AU$. The columns of U are eigenvectors and the main diagonal entries of D are eigenvalues of A.

Reduction to Hessenberg and Tridiagonal Form

Goal: Find an algorithm that reduces the given matrix to triangular form via similarity transform

Note: Due to Abel's theorem, can't expect the algorithm that accomplishes the goal in finite steps Idea:

- Reduce the matrix to upper Hessenberg form
- Find eigenvalues and eigenvectors of Hessenberg matrix
- For Hermitian, it reduces to tridiagonal form (Why?)

Reduction of General Matrices to Hessenberg Form

Steps to reduce a matrix A to Hessenberg form:

1. Partition A as follows:

$$A = \begin{bmatrix} a_{11} & c^T \\ b & \hat{A} \end{bmatrix}$$

2. Let $\hat{Q}_1$ be a reflector such that

$$\hat{Q}_1 b = \begin{bmatrix} -\tau_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \qquad \text{where } \tau = \|b\|_2$$

Reduction of General Matrices to Hessenberg Form

3. Set

$$Q_1 = \begin{bmatrix} 1 & 0^T \\ 0 & \hat{Q}_1 \end{bmatrix}, \quad A_{1/2} = Q_1 A = \left[\begin{array}{c|c} a_{11} & c^T \\ \hline -\tau_1 & \\ 0 & \\ \vdots & \hat{Q}_1 \hat{A} \\ 0 & \end{array}\right],$$

which has the desired zeros in the first column.

- Like the first step of QR algorithm, but less ambitious! (Why?)
  - In QR we left one non-zero in first column, but now we left two non-zeros in first column.

Reduction of General Matrices to Hessenberg Form

4. Compute $A_1 = Q_1 A Q_1^{-1}$, where (recall that) $Q_1^{-1} = Q_1$

$$A_1 = A_{1/2} Q_1 = \left[\begin{array}{c|c} a_{11} & c^T \hat{Q}_1 \\ \hline -\tau_1 & \\ 0 & \\ \vdots & \hat{Q}_1 \hat{A} \hat{Q}_1 \\ 0 & \end{array}\right] = \left[\begin{array}{c|c} a_{11} & * \cdots * \\ \hline -\tau_1 & \\ 0 & \\ \vdots & \hat{A}_1 \\ 0 & \end{array}\right]$$

- Because of the form of Q_1, above does not destroy zeros
  - Had we been more ambitious, the zeros may have got destroyed
- Next Steps: Apply the idea recursively on $\hat{A}_1$

Reduction of General Matrices to Hessenberg Form

Remaining steps: Create zeros in the second column of A_1 or in the 1st col of $\hat{A}_1$

5. Pick a reflector $\hat{Q}_2 \in \mathbb{C}^{(n-2)\times(n-2)}$ same way as in 1st step, except that A replaced by A_1. That is, set

$$Q_2 = \left[\begin{array}{cc|ccc} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \hline 0 & 0 & & & \\ \vdots & \vdots & & \hat{Q}_2 & \\ 0 & 0 & & & \end{array}\right]$$

Reduction of General Matrices to Hessenberg Form

6. Apply Q_2 on A_1 to get $A_{3/2}$

$$A_{3/2} = Q_2 A_1 = \left[\begin{array}{c|c|c} a_{11} & * & * \cdots * \\ \hline -\tau_1 & * & * \cdots * \\ \hline 0 & -\tau_2 & \\ \vdots & \vdots & \hat{Q}_2 \hat{A}_2 \\ 0 & 0 & \end{array}\right]$$

7. Complete the similarity transformation: $A_2 = A_{3/2} Q_2$

$$A_2 = A_{3/2} Q_2 = \left[\begin{array}{c|c|c} a_{11} & * & * \cdots * \\ \hline -\tau_1 & * & * \cdots * \\ \hline 0 & -\tau_2 & \\ \vdots & \vdots & \hat{Q}_2 \hat{A}_2 \hat{Q}_2 \\ 0 & 0 & \end{array}\right]$$

Reduction of General Matrices to Hessenberg Form

6. Next steps creates zeros in the 3rd column, and so on. After $n-2$ steps, the reduction is complete. Result is: $B = Q^*AQ$,

$$Q = Q_1 Q_2 \cdots Q_{n-2} \quad \text{and} \quad Q^* = Q_{n-2} Q_{n-3} \cdots Q_1$$

Algorithm-Reduction to Hessenberg i

for $k = 1, \ldots, n-2$ **do**
  $\beta \leftarrow \max\{|a_{ik}| \mid i = k+1, \ldots, n\}$
  $\gamma_k \leftarrow 0$
  if $\beta \neq 0$ **then**
    % Set up the reflector
    $a_{k+1:n,k} \leftarrow \beta^{-1} a_{k+1:n,k}$
    $\tau_k \leftarrow \sqrt{a_{k+1,k}^2 + \cdots + a_{n,k}^2}$
    if $a_{k+1,k} < 0$ **then**
      $\tau_k = -\tau_k$
    end if
    $\eta \leftarrow a_{k+1,k} + \tau_k$

Algorithm-Reduction to Hessenberg ii

$a_{k+1,k} \leftarrow 1$

$a_{k+2:n,k} \leftarrow a_{k+2:n,k}/\eta$

$\gamma_k \leftarrow \eta/\tau_k$

$\tau_k \leftarrow \tau_k\beta$

% Multiply on the left by $\hat{Q}_k$

$b^T_{k+1:n,1} \leftarrow a^T_{k+1:n,k}a_{k+1:n,k+1:n}$

$b^T_{k+1:n,1} \leftarrow -\gamma b^T_{k+1:n,1}$

$a_{k+1:n,k+1:n} \leftarrow a_{k+1:n,k+1:n} + a_{k+1:n,b^T_{k+1:n,1}}$

% Multiply on the right by $\hat{Q}_k$

$b_{1:n} \leftarrow a_{1:n,k+1:n}a_{k+1:n,k}$

$b_{1:n} \leftarrow -\gamma_k b_{1:n,1}$

$a_{1:n,k+1:n} \leftarrow a_{1:n,k+1:n} + b_{1:n,1}a^T_{k+1:n,k}$

Algorithm-Reduction to Hessenberg iii

```
        a_{k+1,k} ← −τ_k
    end if
end for
τ_{n−1} ← −a_{n,n−1}
```

- The input is A, and output is $B = Q^T AQ$ stored in A
- The zeros below the subdiagonal are used to store u_k
- Scalar γ_k stored in separate array γ
- Note for symmetric A, B will be triadiagonal! (See page 353 in Watkins)

The QR Algorithm

Goal: Compute eigenvalue of $A \in \mathbb{R}^{n\times n}$

Idea: Use QR algorithm as follows:

1: Set $A_0 = A$
2: **for** $i = 1, \cdots,$ **do**
3: $\quad A_{m-1} = Q_m R_m$
4: $\quad R_m Q_m = A_m$
5: **end for**

- In step 1, A_{m-1} is decomposed into Q_m and R_m (using QR)
- In step 2, the factors R_m and Q_m are multiplied in reverse order to get A_m
- $A_m = Q_m^* A_{m-1} Q_m$, the sequence $A_j \to T$, upper triangular form (Why? Later!)

Plan to Compute All Eigenvalues of a Matrix

Steps to compute eigenvalues and eigenvectors:

1. Reduce the matrix to upper Hessenberg form by applying unitary similarity transforms (See previous slides)
2. Apply QR algorithm to reduce this upper hessenberg matrix to upper triangular form

Speeding up Eigenvalue Computations

Gaol: Speedup the computation of eigenvalues

Idea: As the iteration progresses, the entries of the matrix A^m goes to zero, and the diagonal entries tend to get closer to eigenvalues of A, so choose appropriate shifts.

Fact 19

The subdiagonal entries $a^m_{i+1,i} \to 0$ as $m \to \infty$. More precisely, $|\lambda_i| > |\lambda_{i+1}|$, then $a^m_{i+1,i} \to 0$ linearly with convergence ratio $|\lambda_{i+1}/\lambda_i|$, as $m \to \infty$.

Steps to Speedup Computation of Eigenvalues

- If $\lambda_n \neq \lambda_{n-1}$, then we may want to make the ratio $(\lambda_n - \rho)/(\lambda_{n-1} - \rho)$ provided we have can find a ρ close to λ_n.
- Since, in the QR iteration, diagonal entries are getting closer to eigenvalues, choose

$$\rho_n = a^m_{n,n}$$

and apply QR iteration on shifted matrix: $A - \rho_n I$

Shifted QR Algorithm

Require: A

$\rho_0 = 1$
for $i = 1, \ldots$ **do**
    $A_{m-1} - \rho_{m-1} I = Q_m R_m$
    $R_m Q_m + \rho_{m-1} I =: A_m$
    $\rho_m = a_{n,n}$ (Choose shift)
    if $a_{n,n-1} < tol$ **then**
        $A_m = A_m(1 : n-1, 1 : n-1)$ (Delete last row and column)
    end if
end for

- It may happen that one of the subdiagonal entries other than last one becomes zero, in that case, the matrix becomes block upper triangular

Applications of Eigenvalue Algorithms: SVD

Recall basic LA

Recall the following spaces:

$$\mathcal{N}(A) = \{x \in \mathbb{R}^m \mid Ax = 0\}$$
$$\mathcal{R}(A) = \{Ax \mid x \in \mathbb{R}^m\}$$

- The null space is a subspace of $\mathbb{R}^m$
- The range space is a subspace of $\mathbb{R}^n$

Recall the rank-nullity theorem:

$$m = \dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A))$$

Properties of AA^T and A^TA

Fact 20

Prove that A^TA and AA^T are symmetric and positive definite.

Fact 21

$$\mathcal{N}(A^TA) = \mathcal{N}(A)$$

Fact 22

$$\text{rank}(A^TA) = \text{rank}(A) = \text{rank}(A^T) = \text{rank}(AA^T)$$

Fact 23

If v is an eigenvector of $A^T A$ associated with a non-zero eigenvalue λ, then Av is an eigenvector of AA^T associated with the same eigenvalue.

Fact 24

Let v_1 and v_2 be eigenvectors of $A^T A$. If v_1 and v_2 are orthogonal, then Av_1 and Av_2 are also orthogonal.

Fact 25

Let $B \in \mathbb{R}^{n \times n}$ be a symmetric matrix with eigenvectors u_i and u_j associated with eigenvalues λ_i and λ_j, with $\lambda_i \neq \lambda_j$. Then u_i and u_j are orthogonal.

Geometric SVD

Fact 26

Let $A \in \mathbb{R}^{n\times m}$ be a nonzero matrix with rank r. Then $\mathbb{R}^m$ has an orthonormal basis $v_1, v_2, \ldots, v_m$, $\mathbb{R}^n$ has an orthonormal basis $u_1, \ldots, u_n$, and there exists $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ such that

$$Av_i = \begin{cases} \sigma_i u_i & i = 1, \ldots, r \\ 0 & i = r+1, \ldots, m \end{cases}$$

and

$$A^T u_i = \begin{cases} \sigma_i v_i & i = 1, \ldots, r \\ 0 & i = r+1, \ldots, n \end{cases}$$