



Lecture 11: Symmetric Eigen Value Problem

Prof. Krishna R. Pattipati

Dept. of Electrical and Computer Engineering

University of Connecticut

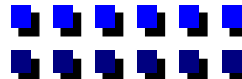
Contact: krishna@engr.uconn.edu (860) 486-2890

ECE 6435

Fall 2008

Adv Numerical Methods in Sci Comp

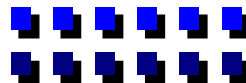
November 5, 2008





Outline of Lecture 11

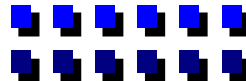
- ❑ House holder method to convert a *symmetric matrix to tridiagonal form*
- ❑ Key Ideas of Symmetric QR method
- ❑ Lanczos method for sparse symmetric matrices
 - Relationship with the conjugate gradient (CG) method
- ❑ A variety of algorithms exist for finding the Eigen values and Eigen vectors of a symmetric matrix
 - Symmetric QR
 - Jacobi method
 - Bisection method (using Sturm sequence property)
 - Power method
 - Lanczos method (for sparse symmetric matrices)
- ❑ **However, symmetric QR is the best general purpose algorithm**
 - Unless only a few Eigen values/Eigen vectors are desired
 - In the latter case, specialized methods may be useful (see Golub and Van Loan, 1996)





Symmetric QR Algorithm

- Key ideas of symmetric QR algorithm
 - If $Q_0^T A Q_0 = H$, an upper Hessenberg matrix
 - A is symmetric $\Rightarrow H = H^T$
 - **A symmetric upper Hessenberg matrix is a symmetric tridiagonal matrix**
 - Symmetry and tridiagonal band structure are preserved with single shift QR
 - Recall that $\{\lambda_i(A)\}$ are real for a symmetric matrix A
 \Rightarrow **no need for double and complex shifts**





Transformation to Tridiagonal Form - 1

Transformation to tridiagonal form via Householder

- Suppose, we have determined $W_1 W_2 \dots W_{k-2}$

- So,

$$A_{k-1} = (W_1 \dots W_{k-1})^T A W_1 \dots W_{k-1} = \begin{bmatrix} B & & \underline{0} & & \\ & & & & \\ & & & & \\ \underline{0}^T & & \underline{b} & & \underline{D} \\ & & & & \\ & & & & \end{bmatrix} \begin{matrix} k-1 \\ 1 \\ n-k \\ k-1 \\ 1 \\ n-k \end{matrix}$$

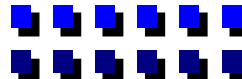
where B is tridiagonal

- \widetilde{W}_k is an order $n-k$ Householder $\ni W_k = \text{Diag}(I_k, \widetilde{W}_k)$

$$A_k = W_k A_{k-1} W_k = \begin{bmatrix} B & 0 \\ & b^T \widetilde{W}_k \\ 0 & \widetilde{W}_k b \quad \widetilde{W}_k D \widetilde{W}_k \end{bmatrix}$$

- Exploit Symmetry in generating A_k

$$\widetilde{W}_k = I_{n-k} - \frac{2}{\underline{u}^T \underline{u}} \underline{u} \underline{u}^T, \quad \underline{u} \in R^{n-k}$$





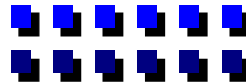
Transformation to Tridiagonal Form - 2

$$\underline{p} = \frac{2}{\underline{u}^T \underline{u}} D \underline{u} \quad \text{and} \quad \underline{q} = \underline{p} - \frac{\underline{p}^T \underline{u}}{\underline{u}^T \underline{u}} \underline{u}$$

$$\begin{aligned} \widetilde{W}_k D \widetilde{W}_k &= \widetilde{W}_k \left[D - \frac{2}{\underline{u}^T \underline{u}} D \underline{u} \underline{u}^T \right] = \widetilde{W}_k \left[D - \underline{p} \underline{u}^T \right] = \left(I - \frac{2}{\underline{u}^T \underline{u}} \underline{u} \underline{u}^T \right) (D - \underline{p} \underline{u}^T) \\ &= D - \underline{u} \underline{p}^T - \underline{p} \underline{u}^T + \frac{2}{\underline{u}^T \underline{u}} \underline{u} \underline{u}^T \underline{p} \underline{u}^T = D - \underline{u} \underline{q}^T - \underline{q} \underline{u}^T \end{aligned}$$

where $\underline{q} = \underline{p} - \frac{\underline{p} \underline{u}}{\underline{u}^T \underline{u}}$

- Note that only the upper D portion $\widetilde{W}_k D \widetilde{W}_k$ needs to be computed $\Rightarrow (n-k)^2$ operations.
- So, the process of going from A_{k-1} to A_k takes $2(n-k)^2$ operations
- Overall, it takes $O(2n^3/3)$ operations to obtain tridiagonal form
- Additional $2n^3/3$ to obtain the transformation matrix $Q_0 = W_1 W_2 \dots W_{n-2}$





Rudiments of a Symmetric QR

- Starting with $T=A_0 = Q_0^T A Q_0$, where T is tridiagonal, the symmetric QR algorithm employs the following iterations

For $k = 1, 2, \dots$

$$A_k - \mu I = QR$$

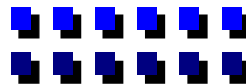
$$A_{k+1} = RQ + \mu I = Q^T (A_k - \mu I) Q + \mu I = Q^T A_k Q$$

end DO

- Key: QR decomposition of a tridiagonal matrix requires $O(n)$ operations. The resulting R has bidiagonal structure \Rightarrow diagonal and super diagonal

that is,

$$A = \begin{bmatrix} a_1 & b_2 & 0 & \dots & 0 \\ b_2 & a_2 & b_3 & \dots & 0 \\ 0 & 0 & 0 & a_{n-1} & b_n \\ 0 & 0 & 0 & b_n & a_n \end{bmatrix} \xrightarrow{Q^T} \begin{bmatrix} x & x & & & 0 \\ & x & x & & \\ & & x & x & \\ & & & x & x \\ 0 & & & & x \end{bmatrix} \rightarrow \begin{array}{l} RQ = \text{tridiagonal} \\ \text{takes } O(n) \text{ operations} \end{array}$$





Picking the Shift Factor

□ Choice of μ :

- one choice is $\mu = a_n$. Another choice is to set $\mu =$ Eigen value of 2 x 2 block

closest to a_n , where the 2 x 2 block is $\begin{pmatrix} a_{n-1} & b_n \\ b_n & a_n \end{pmatrix}$

$$\Rightarrow (\lambda - a_{n-1})(\lambda - a_n) - b_n^2 = \lambda^2 - (a_n + a_{n-1})\lambda + a_n a_{n-1} - b_n^2$$

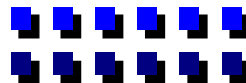
$$\lambda = \frac{a_n + a_{n-1} \pm \sqrt{(a_n + a_{n-1})^2 - 4a_n a_{n-1} + 4b_n^2}}{2}$$

$$= a_n + \frac{a_{n-1} - a_n}{2} \pm \frac{\sqrt{(a_{n-1} - a_n)^2 + 4b_n^2}}{2}$$

$$\Rightarrow \lambda = a_n + d_n \pm \sqrt{d_n^2 + b_n^2} \text{ where } d_n = \frac{a_{n-1} - a_n}{2}$$

$$\Rightarrow \mu = a_n + d_n - \operatorname{sgn}(d_n) \sqrt{d_n^2 + b_n^2}$$

- This choice of μ is called *Wilkinson shift*
- Wilkinson shift results in *cubic convergence* of A to diagonal





QR Decomposition of Tridiagonal - 1

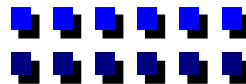
- Only question then is: how to get QR form $(A_k - \mu I)$ in $O(n)$ operations ?
 - Use Givens transformations to obtain QR decomposition of a tridiagonal matrix $(A_k - \mu I)$.
 - Consider the initial step $k = 1$
 - The problem is to find

$$c_1, s_1 \ni \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} \begin{bmatrix} a_{11} - \mu \\ b_2 \end{bmatrix} = \begin{bmatrix} * \\ 0 \end{bmatrix} = k \underline{e}_1$$

- Make $a_{21} = b_2 \rightarrow 0$ via $J(1,2, \theta) = J_{21}$ and consider $J_{21}^T (A_k - \mu I) J_{21}$

$$\begin{bmatrix} x & x & & 0 \\ x & x & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix} \begin{bmatrix} a_1 - \mu & b_2 & & \\ b_2 & a_2 - \mu & b_3 & \\ & b_3 & a_3 - \mu & b_4 \\ 0 & & & \dots \\ & & & & a_{n-1} - \mu & b_n \\ & & & & b_n & a_n - \mu \end{bmatrix} \begin{bmatrix} x & x & & 0 \\ x & x & & \\ & & 1 & \\ 0 & & & 1 \end{bmatrix} = \begin{bmatrix} x & x & + & 0 & \dots & 0 \\ x & x & x & 0 & \dots & 0 \\ + & x & x & x & \dots & 0 \\ & 0 & \cdot & \dots & & \\ & & \cdot & & & \\ 0 & & & \dots & & 0 \end{bmatrix};$$

+ \Rightarrow unwanted elements





QR Decomposition of Tridiagonal - 2

- This initial step is called implicit Q -step, which creates unwanted elements denoted by +
- **Chase these unwanted elements away via Givens or Householder transformation:**

$$A = J_{31}^T A J_{31} = J_{31}^T J_{21}^T (A_k - \mu I) J_{21} J_{31} = \begin{bmatrix} x & x & 0 & 0 \\ x & x & x & + \\ 0 & x & x & x \\ \cdot & + & x & x & x \\ & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & 0 \end{bmatrix}$$

$$A = J_{42}^T A J_{42} = \begin{bmatrix} x & x \\ x & x & x \\ & x & x & x & + \\ & & x & x & x \\ & & + & x & x \end{bmatrix} \text{ etc.}$$

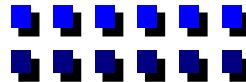
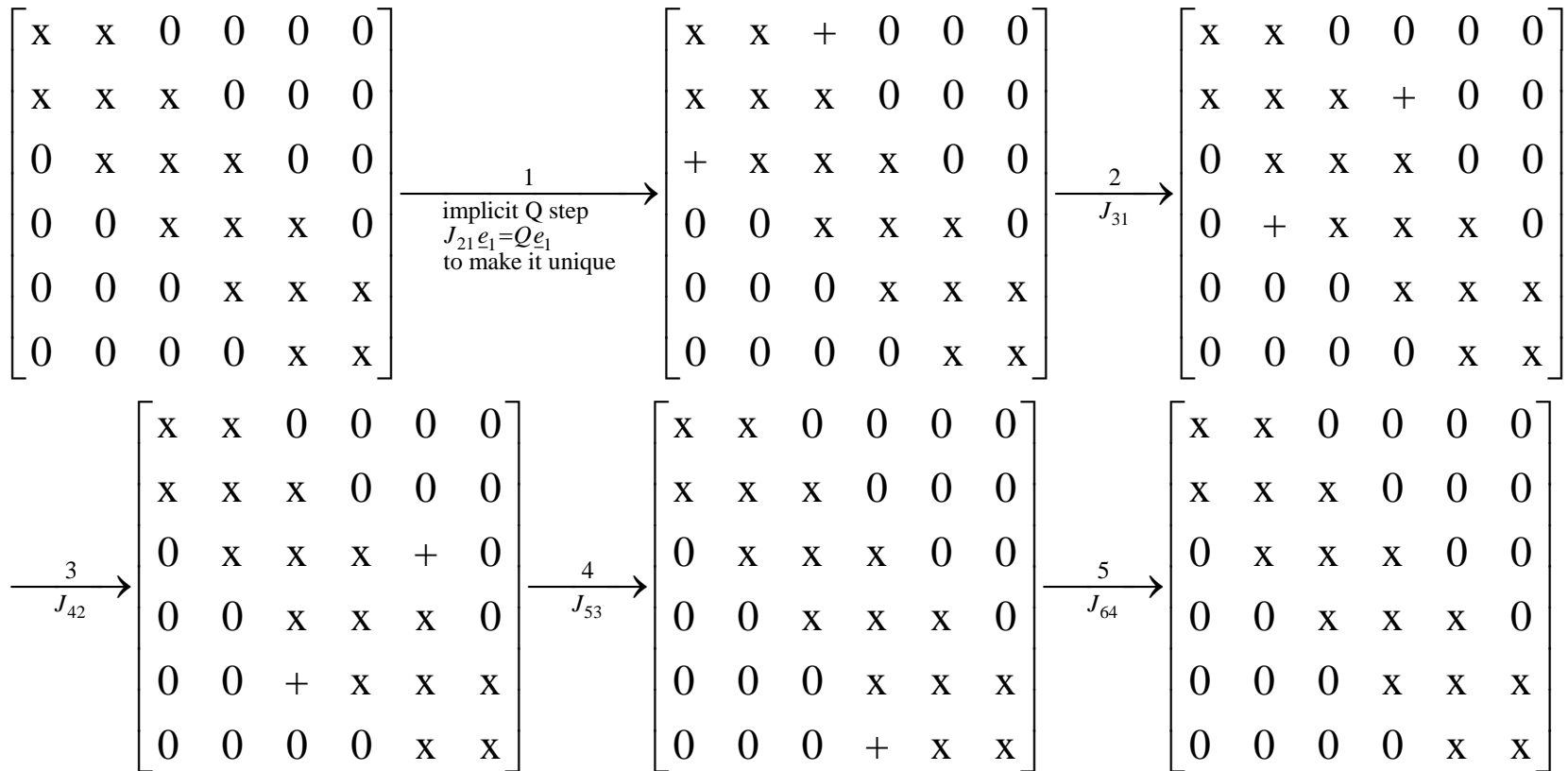


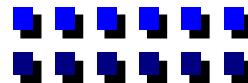


Illustration of QR → Tridiagonal

□ Illustration of the process on a 6 x 6 matrix



(n-1) steps including the implicit Q-step



Wilkinson Shift

- Implicit Symmetric QR – step with Wilkinson shift

$$\mu = a_{nn} + d - \operatorname{sgn}(d)\sqrt{d^2 + a_{n,n-1}^2} = a_{nn} - a_{n,n-1}^2 / (d + \operatorname{sgn}(d)\sqrt{d^2 + a_{n,n-1}^2})$$

$$d = (a_{n-1,n-1} - a_{nn}) / 2; \quad x = a_{11} - \mu; \quad z = a_{21}$$

For $k = 1, 2, \dots, n-1$

Determine c_k and $s_k \ni$

$$\begin{pmatrix} c_k & s_k \\ -s_k & c_k \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} = \begin{pmatrix} * \\ 0 \end{pmatrix}$$

If $k = 1$

$$\text{form } A \leftarrow J_{k+1,k}^T A J_{k+1,k}$$

else

$$\text{form } J_{k+1,k-1}^T A J_{k+1,k-1}$$

end if

If $k < n-1$ then

$$x = a_{k+1,k}; \quad z = a_{k+2,k}$$

end if

end DO

- The procedure requires $O(14n)$ operations and n square roots



Overall Symmetric QR

- ❑ To form Q, requires an additional $4n^2$ operations
- ❑ The overall Symmetric QR Algorithm
 - Find tridiagonal matrix $A_0 = T = Q_0^T A Q_0$ via Householder transformation
 - Do for ever

$$|a_{i+1,i}| \leq \varepsilon (|a_{ii}| + |a_{i+1,i+1}|)$$

For $i = 1, 2, \dots, n-1$

find largest q and smallest $p \ni$ if

$$A = \begin{bmatrix} A_{11} & & 0 \\ & A_{22} & \\ 0 & & A_{33} \end{bmatrix} \begin{matrix} p \\ n-p-q \\ q \end{matrix}$$

where A_{33} is diagonal and A_{22} has no zero subdiagonal elements

if $q = n$ then

quit

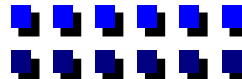
else

$$\text{Apply Wilkinson shift; } A = \text{diag}(I_p Q I_q)^T A \text{diag}(I_p Q I_q)$$

end if

end DO

- ❑ The algorithm requires $\approx 2n^3/3$ operations for $\{\lambda_i(A)\}$ and approximately $5n^3$ operations for Eigen vectors also



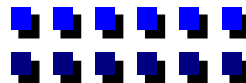


Lanczos Methods

- ❑ Lanczos Methods for sparse symmetric matrices
 - Useful for finding the Eigen values of sparse symmetric matrices
 - Also, related to the conjugate gradient method

- ❑ Key fact: if a matrix A is cyclic then
 - $\{\underline{u}, A\underline{u}, \dots, A^{n-1}\underline{u}\}$ form a basis.
 - The subspaces formed by $\{\underline{u}, A\underline{u}, \dots, A^{k-1}\underline{u}\}$ for $k = 1, 2, \dots, n$ are called Krylov subspaces $K(A, \underline{u}, k)$
 - We know by the power method that $A^k \underline{u}$ goes towards the Eigen vector corresponding to the largest Eigen value.

- ❑ Lanczos Idea
 - What if we orthogonalize the vectors associated with Krylov subspaces? This is the basic idea of Lanczos !!
 - It turns out that we can derive a **three term recursion for the orthogonal vectors**





Lanczos Vectors are Orthogonal

□ Note:

- 1) $\underline{q}_0 = \underline{0}$
- 2) If $\underline{b}_j = 0$, then \underline{q}_{j+1} is any vector orthogonal to $\underline{q}_1 \dots \underline{q}_j$.
- 3) If this happens, T splits into two subblocks
- 4) $A \underline{q}_j \perp^r$ to $\underline{q}_1 \dots \underline{q}_{j-2}$
- 5) Vector \underline{q}_{j+1} is orthogonal to $\{\underline{q}_1 \dots \underline{q}_j\}$

This is because (recall $A=A^T$)

$$\underline{q}_i^T A \underline{q}_j = (A \underline{q}_i)^T \underline{q}_j = (b_{i-1} \underline{q}_{i-1} + a_i \underline{q}_i + b_i \underline{q}_{i+1})^T \underline{q}_j = 0$$

$$\forall i = 1, 2, \dots, j-2$$

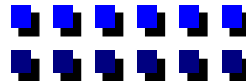
For $i \leq j-2$, we have

$$\underline{q}_i^T A \underline{q}_j = (A \underline{q}_i)^T \underline{q}_j = (b_{i-1} \underline{q}_{i-1} + A \underline{q}_i + b_i \underline{q}_{i+1})^T \underline{q}_j = 0$$

$$\text{From } A \underline{q}_j = +b_j \underline{q}_{j+1} + a_j \underline{q}_j + b_{j-1} \underline{q}_{j-1}$$

$$\text{Since } \underline{q}_i^T A \underline{q}_j = 0 \quad \forall i \leq j-2, \text{ we have } \underline{q}_i^T \underline{q}_{j+1} = 0$$

$$\Rightarrow \underline{q}_{j+1} \text{ is } \perp^r \text{ to all the previous } \{\underline{q}_i\}$$





Lanczos Iteration

- So, the Lanczos vectors $\{q_j\}$ are such that
 - 1) Each q_{j+1} is \perp^r to $\{q_1, q_2, \dots, q_j\}$
 - 2) Each q_{j+1} is a combination of $\{q_1, Aq_1, \dots, A^j q_1\}$
 - 3) Each q_{j+1} is orthogonal to all combinations of $\{q_1, Aq_1, \dots, A^{j-1} q_1\}$
- Lanczos iteration to find the vectors $\{q_j\}$:

$$j = 0; \underline{q}_0 = \underline{0}; b_0 = 1$$

$$\underline{r}_0 = \underline{q}_1; \underline{q}_{j+1} = \underline{r}_j / b_j; j = j + 1$$

$$a_j = \underline{q}_j^T A \underline{q}_j; \underline{r}_j = (A - a_j I) \underline{q}_j - b_{j-1} \underline{q}_{j-1}; b_j = \|\underline{r}_j\|$$

- Requires $O(cn + 4)$ operations where $c = \#$ of non-zero elements per row
 - Prone to round off errors ... Need to do selective re-orthogonalization of r_j against previous Lanczos vectors.
 - Find Eigen values of the tridiagonal matrix by symmetric QR
- Relationship with the conjugate gradient method
 - Residuals $\{r_j = b - A x_j\}$, where x_j is the solution at iteration j of conjugate gradient method, are multiples of the Lanczos vectors $\{q_j\}$



Summary

- ❑ House holder method to convert a *symmetric matrix to tridiagonal form*
- ❑ Lanczos method for sparse symmetric matrices
 - Relationship with the conjugate gradient (CG) method
- ❑ A variety of algorithms exist for finding the Eigen values and Eigen vectors of a symmetric matrix
 - Symmetric QR
 - Jacobi method
 - Bisection method (using Sturm sequence property)
 - Power method
 - Lanczos method (for sparse symmetric matrices)
- ❑ However, symmetric QR is the best general purpose algorithm
 - Unless only a few Eigen values/Eigen vectors are desired
 - In the latter case, specialized methods may be useful (see Golub and Van Loan, 1989)

