Linear Algebra

Sartaj Ul Hasan

Department of Mathematics
Indian Institute of Technology Jammu
Jammu, India - 181221

Email: sartaj.hasan@iitjammu.ac.in

Lecture 40

(Nov 09, 2019)

Diagonalization of Symmetric Matrices

Diagonalization of Symmetric Matrices

- **Definition:** A matrix A is said to be **symmetric** if $A = A^T$.
A symmetric matrix is necessarily square. For the time being, we will restrict our interest to matrices and vectors with real entries only.
- **Proposition 55:** If A is symmetric, then any two eigenvectors from different eigenspaces (i.e. eigenvectors corresponding to different eigenvalues) are orthogonal.
- **Note:** Recall an earlier theorem for square matrices: Eigenvectors from different eigenspaces (i.e. corresponding to different eigenvalues) are linearly independent. For symmetric matrices, we get the stronger result above

Proof of Proposition 55

Proof of Proposition 55: Let u_1 and u_2 be eigenvectors corresponding to different eigenvalues λ_1 and λ_2. Then:

$$
\begin{aligned}
\lambda_1 < u_1, u_2 > &=< \lambda_1 u_1, u_2 > \\
&= (\lambda_1 u_1)^T u_2 \quad \text{(definition of dot product)} \\
&= (Au_1)^T u_2 \quad \text{(Since } u_1 \text{ is an eigenvector)} \\
&= {u_1}^T A^T u_2 \\
&= {u_1}^T A u_2 \quad \text{(Since } A \text{ is symmetric)} \\
&= {u_1}^T (\lambda_2 u_2) \quad \text{(Since } u_2 \text{ is an eigenvector)} \\
&=< u_1, \lambda_2 u_2 > \quad \text{(definition of dot product)} \\
&= \lambda_2 < u_1, u_2 >
\end{aligned}
$$

Since $\lambda_1 \neq \lambda_2$, we must have $< u_1, u_2 >= 0$, as desired.

Orthogonal Matrices

- **Definition:** A square matrix P is said to be **orthogonal** if its columns are orthonormal (*please note this slight inconsistency in terminology*).
- **Proposition 56:** An orthogonal matrix is necessarily invertible and $P^{-1} = P^T$.

  Proof Suppose $P = [v_1 \ v_2 \ \dots \ v_n]$ with the v_i being orthonormal column vectors. Then:

$$P^T P = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_n^T \end{bmatrix} \begin{bmatrix} v_1 & v_2 & \dots & v_n \end{bmatrix} = \begin{bmatrix} v_1^T v_1 & v_1^T v_2 & \dots & v_1^T v_n \\ v_2^T v_1 & v_2^T v_2 & \dots & v_2^T v_n \\ \vdots & \vdots & \dots & \vdots \\ v_n^T v_1 & v_n^T v_2 & \dots & v_n^T v_n \end{bmatrix}$$

  Since $v_i^T v_j = v_i . v_j =< v_i, v_j >= \delta_{ij}$ (Kronecker delta), we get that $P^T P$ is the identity matrix – hence the result.

Diagonalization of Symmetric Matrices (Cont'd)

- **Definition:** A square matrix A is said to be **orthogonally diagonalizable** if there is an orthogonal matrix P and a diagonal matrix D such that $A = PDP^{-1} = PDP^T$.
- **Note:** For an $n \times n$ matrix to be orthogonally diagonalizable, it should have n linearly independent and orthonormal eigenvectors. We see that this happens only in the following case:
- **Proposition 57:** If an $n \times n$ matrix A is orthogonally diagonalizable, then A is symmetric.

  Proof: Suppose $A = PDP^{-1} = PDP^T$ is orthogonally diagonalizable. Then $A^T = (PDP^{-1})^T = (PDP^T)^T = (P^T)^T D^T P^T = (PDP^T)$, since D is a diagonal matrix. But $PDP^T = A$, so $A^T = A$, and so A is symmetric.

Diagonalization of Symmetric Matrices (Cont'd)

- **Definition:** The set of eigenvalues of a matrix A is called the **spectrum** of A.
- **Theorem 7 (Spectral Theorem for Symmetric Matrices):** An $n \times n$ symmetric matrix A has the following properties:
  - (a) The eigenspaces are mutually orthogonal (i.e. eigenvectors corresponding to different eigenvalues are orthogonal)
  - (b) A has n real eigenvalues, counting (algebraic) multiplicities
  - (c) A is orthogonally diagonalizable
  - (d) The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ (as a root of the characteristic equation) i.e., the geometric multiplicity is equal to the algebraic multiplicity.

Remarks about the Spectral Theorem

- The proof of Statement (a) has already been given above (Proposition 55).
- It can be shown that for symmetric real matrices, every eigenvalue is real. The Statement (b) follows from exercises. (Exercise: see questions 23 and 24 on page 341 of Lay which contain hints.)
- The proof of Statement (c) is difficult and will be omitted. However, taking Statement (c) of the Spectral Theorem together with the result on an earlier slide (Proposition 57), we have the result that: A **is orthogonally diagonalizable** $\iff$ A **is symmetric**.
- Statement (d) follows from Statement (c) by using the Diagonalization Theorem.

Proof of Theorem 7 (b)

Theorem 7 (b): Every eigenvalue of a real symmetric matrix is real, i.e., if A is an $n \times n$ real symmetric matrix, then its characteristic polynomial has only real roots (no complex roots).

Step 1 (Lay: Problem 23/Page 341): Let A be an $n \times n$ real symmetric matrix, and let $X \in \mathbb{C}^n$. Put $q = \overline{X}^T AX$, where $\overline{X}$ is complex conjugate of X. Then q is real.

Proof: We have:

$$\begin{aligned}
\overline{q} &= \overline{\overline{X}^T AX} \quad \text{(by definition)} \\
&= X^T \overline{AX} \quad (\text{since } \overline{\overline{Y}} = Y \text{ for any } Y \in \mathbb{C}^n) \\
&= X^T A\overline{X} \quad (\overline{A} = A, \text{ conjugate of a product is product of conjugates}) \\
&= (X^T A\overline{X})^T \quad \text{(since transpose of a scalar is equal to itself)} \\
&= \overline{X}^T A^T X \quad (\text{since}(AB)^T = B^T A^T) \\
&= \overline{X}^T AX \quad (\text{since } A \text{ is symmetric}) = q. \text{ Since } \overline{q} = q,\ q \text{ is real.}
\end{aligned}$$

Proof of Theorem 7 (b) (Cont'd)

Step 2 (Lay: Problem 24/Page 341): Show that if $AX = \lambda X$ for non-zero vector X in $\mathbb{C}^n$, then λ is in fact real, and the real part of X is in fact an eigenvector of A (in $\mathbb{R}^n$).

Proof: Consider $q = \overline{X}^T AX$. By Step 1, q is known to be real.

$$\begin{aligned} \text{Now, } q &= \overline{X}^T(\lambda X) \quad \text{(since } X \text{ is an eigenvector)} \\ &= \lambda(\overline{X}^T X) \end{aligned} \tag{1}$$

Claim: $\overline{X}^T X$ **is real and** > 0

Put $X = \begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$, where $z_i \in \mathbb{C}$

Proof of Theorem 7 (b) (Cont'd)

Then $\overline{X}^T X = [\overline{z_1} \ \ldots \ \overline{z_n}] \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} = z_1\overline{z_1} + \cdots + z_n\overline{z_n} > 0$, since X is non-zero. This proves the claim.

Putting $\overline{X}^T X = r \in \mathbb{R}, r > 0$, we get $q = \lambda r$ from (1), whence $\lambda = \frac{q}{r}$ is again real, as desired.
Finally, if $X = U + iV$, where $U, V \in \mathbb{R}^n$. Then

$$\begin{aligned} AX &= A(U + iV) = AU + iAV \\ &= \lambda X = \lambda(U + iV) = \lambda U + i\lambda V \end{aligned}$$

Equating real and imaginary parts, $AU = \lambda U$, and we are done.

The Spectral Theorem in Practice

- In numerical examples, we first factorize the characteristic polynomial. We will always get as many real roots (counting multiplicities) as the dimension of the matrix, i.e. complex roots will not occur.
- While row reducing the matrix $A - \lambda I$ for any eigenvalue λ to solve the associated homogeneous system, we get as many free variables as the algebraic multiplicity of λ. Thus we get the desired number of basis vectors.
- For each eigenspace of dimension greater than one, we obtain an orthogonal basis by using the Gram-Schmidt process.
- Finally we normalize all the basis vectors.

Numerical Example for Symmetric Matrices

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}, \quad \det(A - \lambda I) = -(\lambda - 1)^2(\lambda - 4)$$

Putting $\lambda = 4$**:** $A - \lambda I = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$

Solving the system, we get $u_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ as an eigenvector.

After normalizing it, we obtain: $v_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$.

Numerical Example for Symmetric Matrices (Cont'd)

Putting $\lambda = 1$: $A - \lambda I = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Solving the system, we get:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} x_2 + \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} x_3$$

We thus get

$u_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ $(x_2 = -1, x_3 = 0)$ and $u_3 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$ $(x_2 = 0, x_3 = -1)$ as eigenvectors.

We see that $< u_1, u_2 >=< u_1, u_3 >= 0$, but $< u_2, u_3 >= 1 \neq 0$.

So now we have to apply Gram-Schmidt orthogonalization process to the basis $\{u_2, u_3\}$ of eigenspace $E_{\lambda=1}$ corresponding to eigenvalue $\lambda = 1$.

Numerical Example for Symmetric Matrices (Cont'd)

$$u_2' = u_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$u_3' = u_3 - \frac{< u_3, u_2' >}{< u_2', u_2' >} u_2' = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1/2 \\ 1/2 \\ -1 \end{bmatrix}$$

Now, normalize u_2', u_3' to get: $v_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ and $v_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$

Numerical Example for Symmetric Matrices (Cont'd)

Check: We must have $A = PDP^{-1} = PDP^T$ or $AP = PD$, where P has v_1, v_2 and v_3 as columns.

$$AP = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix}$$

$$\text{and, } PD = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{4}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{4}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \end{bmatrix}$$