Mostafa Sabri

Lecture 24

7.1 Diagonalization of Symmetric Matrices

جامعة نيويورك أبوظبي
NYU ABU DHABI

Symmetric Matrices

A symmetric matrix A is a square matrix satisfying $A^T = A$. This implies the entries below the diagonal are a mirror of those above the diagonal.

e.g. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ is symmetric but $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $\begin{bmatrix} 1 & 2 \end{bmatrix}$ are not.

Lem: Equip $\mathbb{R}^n$ with $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u} \cdot \mathbf{v}$. Then

$$\langle \mathbf{u}, A\mathbf{v} \rangle = \langle A^T \mathbf{u}, \mathbf{v} \rangle .$$

Proof. Recall that $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v}$. So
$\langle A^T \mathbf{u}, \mathbf{v} \rangle = (A^T \mathbf{u})^T \mathbf{v} = (\mathbf{u}^T A)\mathbf{v} = \mathbf{u}^T (A\mathbf{v}) = \langle \mathbf{u}, A\mathbf{v} \rangle$. $\square$

Symmetry and Orthogonality

Symmetric matrices have a lot of advantages as we will see in this section.

Thm: If A is symmetric, eigenvectors corresponding to distinct eigenvalues are orthogonal.

Proof. If $A\mathbf{v} = \lambda_1\mathbf{v}$ and $A\mathbf{u} = \lambda_2\mathbf{u}$ then since $A^T = A$,

$$\lambda_2\langle \mathbf{v}, \mathbf{u}\rangle = \langle \mathbf{v}, A\mathbf{u}\rangle = \langle A\mathbf{v}, \mathbf{u}\rangle = \lambda_1\langle \mathbf{v}, \mathbf{u}\rangle$$

so $(\lambda_2 - \lambda_1)\langle \mathbf{v}, \mathbf{u}\rangle = 0$. If $\lambda_2 \neq \lambda_1$, we must have $\mathbf{v} \perp \mathbf{u}$. $\square$

Thm: If A is a real $n \times n$ symmetric matrix, it has n real eigenvalues, counting multiplicities.

Symmetry and spectrum

"Counting multiplicities" means that if an eigenvalue has algebraic multiplicity k, we count it k times.

<u>Proof.</u> Suppose $A\mathbf{v} = \lambda\mathbf{v}$. Then $A\overline{\mathbf{v}} = \overline{A\mathbf{v}} = \overline{\lambda}\overline{\mathbf{v}}$, so $\overline{\mathbf{v}}$ is an eigenvector of A for $\overline{\lambda}$. Hence,

$$\lambda\langle\overline{\mathbf{v}}, \mathbf{v}\rangle = \langle\overline{\mathbf{v}}, A\mathbf{v}\rangle = \langle A\overline{\mathbf{v}}, \mathbf{v}\rangle = \overline{\lambda}\langle\overline{\mathbf{v}}, \mathbf{v}\rangle,$$

so $(\overline{\lambda} - \lambda)\langle\overline{\mathbf{v}}, \mathbf{v}\rangle = 0$. But $\langle\overline{\mathbf{v}}, \mathbf{v}\rangle = \sum_{i=1}^{n} \overline{v}_i v_i$ $= \sum_{i=1}^{n} |v_i|^2 > 0$ since $\mathbf{v} \neq 0$. Hence, $\overline{\lambda} = \lambda$, so $\lambda \in \mathbb{R}$. Finally, we know that A has n complex eigenvalues. Since we showed any eigenvalue is real, then it has n real eigenvalues. $\square$

The *spectrum* of a matrix A is its set of eigenvalues. If A is real and symmetric, its spectrum is thus real.

Symmetry and Orthogonality

We say that A is *orthogonally diagonalizable* if there is an orthogonal matrix P such that $D = P^{-1}AP$ is diagonal. Recall that $P^{-1} = P^T$ if P is orthogonal.

> Thm: A square matrix A is orthogonally diagonalizable iff it is symmetric.

Proof. ($\implies$) If $D = P^{-1}AP$ with P orthogonal and D diagonal, then $A = PDP^{-1} = PDP^T$ and so $A^T = P^{TT}D^TP^T = PDP^T = A$. So A is symmetric.

($\impliedby$) Let A be symmetric.[1]
1. We approximate A by a symmetric matrix $A(t)$ that has simple spectrum. In fact, if $\{\lambda_1, \ldots, \lambda_n\}$ are the

[1] I found this argument in the lectures of Oliver Knill, Harvard.

Symmetry and Orthogonality

eigenvalues of A counting multiplicity and $\mathbf{v}_1$ is a normalized eigenvector of A for λ_1, then the eigenvalues[2] of $A_1(t) = A + t\mathbf{v}_1\mathbf{v}_1^T$ are $\{\lambda_1 + t, \lambda_2, \ldots, \lambda_n\}$. In particular, if λ_1 had multiplicity k for A, it has multiplicity $k-1$ for $A_1(t)$. Adding similar rank-one perturbations gives the $A(t)$ we seek.

2. All eigenvalues of $A(t)$ are distinct, so it is diagonalizable, and its normalized eigenvectors $\mathbf{p}_i(t)$ are orthonormal as $A(t)$ is symmetric. Let $P(t)$ be the matrix with columns $\mathbf{p}_i(t)$. Then $D(t) = P(t)^{-1}A(t)P(t)$ is diagonal. Taking $t \to 0$ gives a diagonalization $D = P^{-1}AP$ for A with P orthogonal. Indeed, $\langle \mathbf{p}_i(t), \mathbf{p}_j(t)\rangle = \delta_{ij}\ \forall t > 0 \implies \langle \mathbf{p}_i(0), \mathbf{p}_j(0)\rangle = \delta_{ij}$. $\square$

---

[2] Ding-Zhou, *Eigenvalues of rank-one updated matrices with some applications*, 2007. The proof is simple.

Spectral Theorem

Here $\delta_{ij} = 1$ if $i = j$ and 0 if $i \neq j$.

Note that if A was not symmetric, the eigenvectors of $A(t)$ in step 2 would only be linearly independent. The limit of such vectors can be dependent, making $P(0)$ singular. For example, $\left\{ \frac{1}{\sqrt{2}}(1,1), \frac{1}{\sqrt{2(1+t^2)}}(1+t, 1-t) \right\}$ are independent for any $t > 0$, but become dependent when $t = 0$. The theorem is false if A is not symmetric.

Orthogonally diagonalize $A = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$ and

$B = \begin{bmatrix} 2 & 2 & -2 \\ 2 & -1 & 4 \\ -2 & 4 & -1 \end{bmatrix}$.

Spectral Theorem

Spectral Thm: Let A be a symmetric $n \times n$ matrix.

1. A has n real eigenvalues λ_i, counting multiplicities.
2. The algebraic and geometric multiplicities coincide for each λ_i.
3. The eigenspaces are mutually orthogonal.
4. There is an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ of $\mathbb{R}^n$ consisting of eigenvectors of A, so $A\mathbf{u}_i = \lambda_i \mathbf{u}_i$.
5. A is orthogonally diagonalizable.

Spectral Theorem

6. A has a *spectral decomposition*

$$A = \sum_{i=1}^{n} \lambda_i \mathcal{P}_i \,,$$

   where $\mathcal{P}_i = \mathbf{u}_i \mathbf{u}_i^T$.

7. $\mathcal{P}_i \mathbf{x} = (\mathbf{x} \cdot \mathbf{u}_i)\mathbf{u}_i = \mathrm{proj}_{W_{\mathbf{u}_i}}(\mathbf{x})$ is the orthogonal projection of $\mathbf{x}$ onto $\mathbf{u}_i$ and

$$I_n = \sum_{i=1}^{n} \mathcal{P}_i \,.$$

<u>Proof.</u> We already proved 1. and 5. Point 5. implies 2. (Lec17). Point 3. means that $E_{\lambda_1} \perp E_{\lambda_2}$ if $\lambda_1 \neq \lambda_2$, i.e. $\mathbf{u} \perp \mathbf{v}$ if $\mathbf{u} \in E_{\lambda_1}$ and $\mathbf{v} \in E_{\lambda_2}$, which we also proved since $\mathbf{u}, \mathbf{v}$ are then eigenvectors for distinct eigenvalues.

Spectral Theorem

Using Gram-Schmidt, we can find an orthonormal basis for each E_{λ_i}. These bases put together form an orthogonal set of eigenvectors by 3., and they span $\mathbb{R}^n$ by 2., which proves 4. It only remains to prove 6. and 7.

By 5., $D = P^T AP$, so $A = PDP^T$ with D diagonal and P orthogonal, with columns $\mathbf{u}_i$ that are orthonormal eigenvectors of A. Now $PD = P\begin{bmatrix}\lambda_1\mathbf{e}_1 & \cdots & \lambda_n\mathbf{e}_n\end{bmatrix} = \begin{bmatrix}\lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n\end{bmatrix}$. Thus,

$$A = PDP^T = \begin{bmatrix}\lambda_1\mathbf{u}_1 & \cdots & \lambda_n\mathbf{u}_n\end{bmatrix}\begin{bmatrix}\mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_n^T\end{bmatrix} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T .$$

Examples

This proves 6. Finally, $P_i\mathbf{x} = \mathbf{u}_i\mathbf{u}_i^T\mathbf{x} = \mathbf{u}_i(\mathbf{u}_i^T\mathbf{x}) = \mathbf{u}_i(\mathbf{u}_i \cdot \mathbf{x}) = (\mathbf{x} \cdot \mathbf{u}_i)\mathbf{u}_i$ as required. Moreover, since $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ is an orthonormal basis of $\mathbb{R}^n$, we have the expansion $\mathbf{x} = \sum_{i=1}^n (\mathbf{x} \cdot \mathbf{u}_i)\mathbf{u}_i = \sum_{i=1}^n \mathcal{P}_i\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$. This implies that $I_n = \sum_{i=1}^n \mathcal{P}_i$. □

1. Find a spectral decomposition for the previous matrices A and B.

2. Find a spectral decomposition for $C = \begin{bmatrix} 1 & 3 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$.