

MAT188H1S Lec0101 Burbulla

Week 12 Lecture Notes

Winter 2020

Chapter 8: Orthogonality

8.2: Orthogonal Diagonalization

Eigenvectors of Symmetric Matrices

Theorem 8.2.4: Let \vec{x}_1 and \vec{x}_2 be eigenvectors of the symmetric $n \times n$ matrix A , corresponding to the distinct eigenvalues λ_1 and λ_2 , respectively. Then \vec{x}_1 and \vec{x}_2 are orthogonal.

Proof:

$$\begin{aligned} \lambda_1(\vec{x}_1 \cdot \vec{x}_2) &= (\lambda_1 \vec{x}_1) \cdot \vec{x}_2 = (A \vec{x}_1) \cdot \vec{x}_2 = (A \vec{x}_1)^T \vec{x}_2 = (\vec{x}_1^T A^T) \vec{x}_2 \\ &= \vec{x}_1^T (A^T \vec{x}_2) = \vec{x}_1^T (A \vec{x}_2) = \vec{x}_1^T (\lambda_2 \vec{x}_2) = \vec{x}_1 \cdot (\lambda_2 \vec{x}_2) = \lambda_2 (\vec{x}_1 \cdot \vec{x}_2) \\ &\Rightarrow (\lambda_1 - \lambda_2) \vec{x}_1 \cdot \vec{x}_2 = 0 \Rightarrow \vec{x}_1 \cdot \vec{x}_2 = 0, \text{ since } \lambda_1 \neq \lambda_2. \quad \blacksquare \end{aligned}$$

This is surely one of the nicest proofs you'll see in all of linear algebra!

Example 1

Let

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

Check that $\det(xI - A) = (x - 1)(x - 2)(x + 1)$, so that the eigenvalues of A are $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = -1$. Check that corresponding eigenvectors are

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}.$$

Observe that

$$\vec{x}_1 \cdot \vec{x}_2 = 0, \vec{x}_1 \cdot \vec{x}_3 = 0, \vec{x}_2 \cdot \vec{x}_3 = 0.$$

Example 2

Let

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}.$$

Check that $\det(xI - A) = x^2(x - 6)$, so that the eigenvalues of A are $\lambda_1 = 0$, repeated, $\lambda_2 = 6$. The corresponding eigenvectors are

$$\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \text{ and } \vec{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Observe that $\vec{x}_1 \cdot \vec{x}_3 = 0$, $\vec{x}_2 \cdot \vec{x}_3 = 0$, but $\vec{x}_1 \cdot \vec{x}_2 \neq 0$. Also: $E_0(A)$ is the plane with equation $x_1 + x_2 + x_3 = 0$, and $E_6(A)$ is the line normal to the plane: $E_6(A) = (E_0(A))^\perp$.

Orthogonal Matrices

Definition: The $n \times n$ matrix P is called **orthogonal** if $P^{-1} = P^T$.

Both

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \text{ and } \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}$$

are examples of 2×2 orthogonal matrices. In fact, every 2×2 orthogonal matrix is one of the above two, which you may recognize as the matrices for a rotation, and a reflection, respectively. Note that

$$\det \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} = 1 \text{ and } \det \begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix} = -1.$$

Aside: as an exercise, prove if P is orthogonal then $\det(P) = \pm 1$.

Example 3

$$P = \begin{bmatrix} 3/7 & 2/7 & 6/7 \\ -6/7 & 3/7 & 2/7 \\ 2/7 & 6/7 & -3/7 \end{bmatrix}.$$

P is orthogonal:

$$\begin{aligned} PP^T &= \begin{bmatrix} 3/7 & 2/7 & 6/7 \\ -6/7 & 3/7 & 2/7 \\ 2/7 & 6/7 & -3/7 \end{bmatrix} \begin{bmatrix} 3/7 & -6/7 & 2/7 \\ 2/7 & 3/7 & 6/7 \\ 6/7 & 2/7 & -3/7 \end{bmatrix} \\ &= \dots = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

Characterization of Orthogonal Matrices

Theorem 8.2.1: the following statements are equivalent:

1. P is orthogonal.
2. The columns of P form an orthonormal basis of \mathbb{R}^n .
3. The rows of P form an orthonormal basis of \mathbb{R}^n .

Proof: $1 \Leftrightarrow 2$: Let $P = [\vec{v}_1 \mid \vec{v}_2 \mid \dots \mid \vec{v}_n]$; then

$$\begin{aligned} P^T P = I &\Leftrightarrow \begin{bmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \dots \\ \vec{v}_n^T \end{bmatrix} [\vec{v}_1 \mid \vec{v}_2 \mid \dots \mid \vec{v}_n] = I \\ &\Leftrightarrow \vec{v}_i^T \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \Leftrightarrow \vec{v}_i \cdot \vec{v}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases} \\ &\Leftrightarrow \{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} \text{ is an orthonormal basis of } \mathbb{R}^n. \end{aligned}$$

And 3 follows since P is orthogonal if and only if P^T is orthogonal.

Example 4

Find the standard matrix of $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$, a rotation of $\pi/3$ in \mathbb{R}^3 around the line parallel to the vector $\vec{d} = [1 \ 1 \ 1]^T$.

Solution: let A be the matrix of T . We have $A\vec{d} = \vec{d}$. (why?) Let $U = \{\vec{x} \in \mathbb{R}^3 \mid \vec{d} \cdot \vec{x} = 0\}$. In the plane U the effect of T is to rotate every vector by $\pi/3$. Pick $\vec{v}_1, \vec{v}_2, \vec{v}_3 \in U$

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}.$$

Then

$$\cos \theta_1 = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|} = \frac{1}{2} \quad \text{and} \quad \cos \theta_2 = \frac{\vec{v}_2 \cdot \vec{v}_3}{\|\vec{v}_2\| \|\vec{v}_3\|} = \frac{1}{2}.$$

So the angle in each case between these pairs of vectors is $\theta = \pi/3$.

Thus

$$A\vec{d} = \vec{d}, \quad A\vec{v}_1 = \vec{v}_2, \quad A\vec{v}_2 = \vec{v}_3.$$

Put this information into one matrix equation and solve for A :

$$\begin{aligned} A \begin{bmatrix} \vec{d} & \vec{v}_1 & \vec{v}_2 \end{bmatrix} &= \begin{bmatrix} \vec{d} & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \Leftrightarrow A = \begin{bmatrix} \vec{d} & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \begin{bmatrix} \vec{d} & \vec{v}_1 & \vec{v}_2 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}^{-1} = \frac{1}{3} \begin{bmatrix} 2 & -1 & 2 \\ 2 & 2 & -1 \\ -1 & 2 & 2 \end{bmatrix}. \end{aligned}$$

Note that A is orthogonal. This is not a coincidence. Rotations preserve length and

Theorem: if P is orthogonal then $\|P\vec{x}\| = \|\vec{x}\|$ for all $\vec{x} \in \mathbb{R}^n$.

Proof: we shall use $P^T P = I$.

$$\|P\vec{x}\|^2 = (P\vec{x}) \cdot (P\vec{x}) = (P\vec{x})^T (P\vec{x}) = \vec{x}^T P^T P \vec{x} = \vec{x}^T \vec{x} = \vec{x} \cdot \vec{x} = \|\vec{x}\|^2.$$

Orthogonal Diagonalization

Definition: An $n \times n$ matrix A is orthogonally diagonalizable if there is a diagonal matrix D and an orthogonal matrix P such that

$$D = P^T A P.$$

Comments:

1. Orthogonal diagonalization is a special case of regular diagonalization, since for an orthogonal matrix, $P^{-1} = P^T$.
2. If A is orthogonally diagonalizable then A must be symmetric:

$$\begin{aligned} D = P^T A P &\Rightarrow A = P D P^T, \text{ since } P^{-1} = P^T \\ &\Rightarrow A^T = (P D P^T)^T \Rightarrow A^T = P D^T P^T \\ &\Rightarrow A^T = P D P^T, \text{ since } D \text{ is diagonal} \\ &\Rightarrow A^T = A \end{aligned}$$

Principal Axes Theorem or Spectral Theorem

Theorem 8.2.2: Let A be an $n \times n$ matrix. The following statements are equivalent:

1. A is orthogonally diagonalizable.
2. \mathbb{R}^n has an orthonormal basis consisting of eigenvectors of A .
3. A is symmetric.

Proof: Some of the details have already been covered. For the rest of the details, see the book. ¶ Implicit in the theorem is the fact that all the eigenvalues of a symmetric matrix must be real numbers, not complex numbers. This is the hardest part of the details we have not covered. The name of the theorem comes from the basis of orthonormal eigenvectors, which are called principal axes.

Example 5: Example 1 Revisited

$A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$ has eigenvalues $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -1$, with

eigenvectors $\vec{x}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$. Divide

each eigenvector by its length to obtain orthonormal vectors. Then

$$P = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \end{bmatrix} \text{ and } D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

So A is orthogonally diagonalizable and

$$D = P^T A P.$$

Example 6: Example 2 Revisited

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

has eigenvalues $\lambda_1 = 0$, repeated, and $\lambda_2 = 6$. We found that

$$E_6(A) = \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}.$$

$E_0(A)$ is a little trickier.

$$E_0(A) = \text{null} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}.$$

Note that I have picked an orthogonal basis of $E_0(A)$. Then take

$$P = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix} \text{ and } D = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Bonus Topic: Eigenvalues of $A^T A$.

Theorem: If A is a real matrix, then $A^T A$ has nonnegative eigenvalues.

Proof: Let $A^T A \vec{x} = \lambda \vec{x}$. Then

$$\begin{aligned} \|A \vec{x}\|^2 &= (A \vec{x}) \cdot (A \vec{x}) \\ &= (A \vec{x})^T (A \vec{x}) \\ &= \vec{x}^T A^T A \vec{x} \\ &= \vec{x}^T \lambda \vec{x} \\ &= \lambda \vec{x} \cdot \vec{x} \\ &= \lambda \|\vec{x}\|^2, \end{aligned}$$

implying that $\lambda \geq 0$, since $\vec{x} \neq \vec{0}$. ◻

Bonus Topic: Matrices of Reflections in \mathbb{R}^3

Observe that if \vec{u}, \vec{v} are orthogonal unit vectors in \mathbb{R}^3 then $\{\vec{u}, \vec{v}, \vec{u} \times \vec{v}\}$ is an orthonormal basis of \mathbb{R}^3 , and so the matrix

$$A = [\vec{u} \quad \vec{v} \quad \vec{u} \times \vec{v}]$$

must be an orthogonal matrix. Consequently $AA^T = I$.

Theorem: Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a reflection in the plane passing through the origin normal to the unit vector \vec{v} . Then $T(\vec{x}) = Q\vec{x}$ for

$$Q = I - 2\vec{v}\vec{v}^T.$$

Proof: pick another unit vector \vec{u} which is orthogonal to \vec{v} . Then $\{\vec{u}, \vec{v}, \vec{u} \times \vec{v}\}$ is an orthonormal basis of \mathbb{R}^3 and $\vec{u}, \vec{u} \times \vec{v}$ span the plane passing through the origin with normal vector \vec{v} . Thus

$$Q\vec{u} = \vec{u}, \quad Q\vec{v} = -\vec{v} \quad \text{and} \quad Q(\vec{u} \times \vec{v}) = \vec{u} \times \vec{v}.$$

In terms of matrix multiplication, with A as above, we have

$$QA = [\vec{u} \mid -\vec{v} \mid \vec{u} \times \vec{v}].$$

Solve for Q :

$$\begin{aligned} Q &= [\vec{u} \quad -\vec{v} \quad \vec{u} \times \vec{v}] A^T \\ &= ([\vec{u} \quad \vec{v} \quad \vec{u} \times \vec{v}] - 2 [\vec{0} \quad \vec{v} \quad \vec{0}]) A^T \\ &= AA^T - 2 [\vec{0} \quad \vec{v} \quad \vec{0}] \begin{bmatrix} \vec{u}^T \\ \vec{v}^T \\ (\vec{u} \times \vec{v})^T \end{bmatrix} \\ &= I - 2\vec{v}\vec{v}^T. \quad \blacksquare \end{aligned}$$

Aside: $R = -Q$ is the matrix of a rotation of π around the axis through the origin parallel to \vec{v} . Both Q and R are symmetric.

Example 7

If $\vec{v} = [v_1 \ v_2 \ v_3]^T$ is a unit column vector, then Q looks like

$$Q = I - 2\vec{v}\vec{v}^T = \begin{bmatrix} 1 - 2v_1^2 & -2v_1v_2 & -2v_1v_3 \\ -2v_2v_1 & 1 - 2v_2^2 & -2v_2v_3 \\ -2v_3v_1 & -2v_3v_2 & 1 - 2v_3^2 \end{bmatrix}.$$

If $\vec{v}_1 = \vec{e}_1$, $\vec{v}_2 = \vec{e}_2$, and $\vec{v}_3 = (\vec{e}_1 - \vec{e}_3)/\sqrt{2}$, then

$$Q_1 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad Q_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

represent, respectively, a reflection in the plane $x = 0$, a reflection in the plane $y = 0$, and a reflection in the plane $x = z$.

Bonus Topic: Matrices of Rotations in \mathbb{R}^3

Theorem: Let $T : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ be a rotation through θ around an axis parallel to the unit vector \vec{u} , such that $T(\vec{x}) = R\vec{x}$. Then $R =$

$$\begin{bmatrix} \cos\theta + u_1^2(1 - \cos\theta) & u_1u_2(1 - \cos\theta) - u_3\sin\theta & u_1u_3(1 - \cos\theta) + u_2\sin\theta \\ u_2u_1(1 - \cos\theta) + u_3\sin\theta & \cos\theta + u_2^2(1 - \cos\theta) & u_2u_3(1 - \cos\theta) - u_1\sin\theta \\ u_3u_1(1 - \cos\theta) - u_2\sin\theta & u_3u_2(1 - \cos\theta) + u_1\sin\theta & \cos\theta + u_3^2(1 - \cos\theta) \end{bmatrix}$$

Proof: suppose $\{\vec{u}, \vec{v}, \vec{u} \times \vec{v}\}$ is an orthonormal basis of \mathbb{R}^3 . Then $R\vec{u} = \vec{u}$. Check that within the plane normal to \vec{u} and passing through the origin,

$$R\vec{v} = \cos\theta\vec{v} + \sin\theta\vec{u} \times \vec{v} \quad \text{and} \quad R(\vec{u} \times \vec{v}) = \cos\theta\vec{u} \times \vec{v} - \sin\theta\vec{v},$$

Let $A = [\vec{u} \ \vec{v} \ \vec{u} \times \vec{v}]$, as above. Then

$$\begin{aligned}
RA &= \begin{bmatrix} R\vec{u} & R\vec{v} & R(\vec{u} \times \vec{v}) \end{bmatrix} \\
&= \begin{bmatrix} \vec{u} & \cos\theta \vec{v} + \sin\theta \vec{u} \times \vec{v} & \cos\theta \vec{u} \times \vec{v} - \sin\theta \vec{v} \end{bmatrix} \\
&= \begin{bmatrix} \vec{u} & \cos\theta \vec{v} & \cos\theta \vec{u} \times \vec{v} \end{bmatrix} + \begin{bmatrix} \vec{0} & \sin\theta \vec{u} \times \vec{v} & -\sin\theta \vec{v} \end{bmatrix} \\
\Rightarrow R &= \left(\begin{bmatrix} \vec{u} & \cos\theta \vec{v} & \cos\theta \vec{u} \times \vec{v} \end{bmatrix} + \begin{bmatrix} \vec{0} & \sin\theta \vec{u} \times \vec{v} & -\sin\theta \vec{v} \end{bmatrix} \right) A^T \\
&= \begin{bmatrix} \vec{u} & \cos\theta \vec{v} & \cos\theta \vec{u} \times \vec{v} \end{bmatrix} \begin{bmatrix} \vec{u}^T \\ \vec{v}^T \\ (\vec{u} \times \vec{v})^T \end{bmatrix} + \begin{bmatrix} \vec{0} & \sin\theta \vec{u} \times \vec{v} & -\sin\theta \vec{v} \end{bmatrix} \begin{bmatrix} \vec{u}^T \\ \vec{v}^T \\ (\vec{u} \times \vec{v})^T \end{bmatrix} \\
&= \vec{u}\vec{u}^T + \cos\theta \vec{v}\vec{v}^T + \cos\theta (\vec{u} \times \vec{v})(\vec{u} \times \vec{v})^T + \sin\theta (\vec{u} \times \vec{v})\vec{v}^T - \sin\theta \vec{v}(\vec{u} \times \vec{v})^T \\
&= (1 - \cos\theta)\vec{u}\vec{u}^T + \cos\theta (\vec{u}\vec{u}^T + \vec{v}\vec{v}^T + (\vec{u} \times \vec{v})(\vec{u} \times \vec{v})^T) + \sin\theta ((\vec{u} \times \vec{v})\vec{v}^T - \vec{v}(\vec{u} \times \vec{v})^T) \\
&= (1 - \cos\theta)\vec{u}\vec{u}^T + \cos\theta I + \sin\theta ((\vec{u} \times \vec{v})\vec{v}^T - \vec{v}(\vec{u} \times \vec{v})^T), \text{ since } AA^T = I.
\end{aligned}$$

Finally we claim that

$$(\vec{u} \times \vec{v})\vec{v}^T - \vec{v}(\vec{u} \times \vec{v})^T = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix} :$$

first observe that for all unit vectors \vec{v} , $(\vec{u} \times \vec{v})\vec{v}^T - \vec{v}(\vec{u} \times \vec{v})^T$ is a matrix such that $((\vec{u} \times \vec{v})\vec{v}^T - \vec{v}(\vec{u} \times \vec{v})^T)\vec{v} = \vec{u} \times \vec{v}$. Then let $\vec{v} = \vec{e}_1, \vec{e}_2, \vec{e}_3$ in turn, to find its cols. ◻

Example 8

Find the matrices that represent the two possible rotations of 90° around the y -axis.

Solution: Let $\vec{u} = \vec{e}_2$, so that $u_1 = 0, u_2 = 1, u_3 = 0$. For θ we can take $\pm\pi/2$: one choice represents a rotation clockwise around the y -axis; the other represents a rotation counter-clockwise around the y -axis. The two possibilities for R are:

$$R_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \text{ or } R_2 = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix},$$

as you can check. Notice these two matrices are inverses of each other.