SF 1684 Algebra and Geometry
Lecture 18

Patrick Meisner

KTH Royal Institute of Technology

Topics for Today

1. Orthogonal Diagonalization
2. Powers of Matrices
3. Cayley-Hamilton Theorem

Orthogonally Similar

Recall we say that two square matrices A and C are similar if and only if there is an invertible matrix, P, such that $C = P^{-1}AP$.

Orthogonally Similar

Recall we say that two square matrices A and C are similar if and only if there is an invertible matrix, P, such that $C = P^{-1}AP$.

Definition

We say that two square matrices A and C are **orthogonally similar** if there is an orthogonal matrix, P, such that $C = P^T AP$.

Orthogonally Similar

Recall we say that two square matrices A and C are similar if and only if there is an invertible matrix, P, such that $C = P^{-1}AP$.

Definition

We say that two square matrices A and C are **orthogonally similar** if there is an orthogonal matrix, P, such that $C = P^T AP$.

Recall also that if P is orthogonal that $P^T = P^{-1}$ $\quad C = P^{-1} A P$

Orthogonally Similar

Recall we say that two square matrices A and C are similar if and only if there is an invertible matrix, P, such that $C = P^{-1}AP$.

Definition

We say that two square matrices A and C are **orthogonally similar** if there is an orthogonal matrix, P, such that $C = P^T AP$.

Recall also that if P is orthogonal that $P^T = P^{-1}$ and so if two matrices are orthogonally similar then they are also similar.

Orthogonally Similar

Recall we say that two square matrices A and C are similar if and only if there is an invertible matrix, P, such that $C = P^{-1}AP$.

Definition

We say that two square matrices A and C are **orthogonally similar** if there is an orthogonal matrix, P, such that $C = P^T AP$.

Recall also that if P is orthogonal that $P^T = P^{-1}$ and so if two matrices are orthogonally similar then they are also similar.

Theorem

Two square matrices are orthogonally similar if and only if there exists orthonormal bases with respect to which the matrices represent the same linear transformation.

Orthogonal Diagonalization Problem

Since orthonormal bases are the nicest bases and diagonal matrices are the nicest matrices, this leads to an obvious next question.

Orthogonal Diagonalization Problem

Since orthonormal bases are the nicest bases and diagonal matrices are the nicest matrices, this leads to an obvious next question.

Question (Orthogonal Diagonalization Problem)

Given a square matrix A, does there exist an orthogonal matrix P for which $P^T AP$ is a diagonal matrix

Orthogonal Diagonalization Problem

Since orthonormal bases are the nicest bases and diagonal matrices are the nicest matrices, this leads to an obvious next question.

Question (Orthogonal Diagonalization Problem)

Given a square matrix A, does there exist an orthogonal matrix P for which $P^T AP$ is a diagonal matrix, and if so how does one find such a P?

Orthogonal Diagonalization Problem

Since orthonormal bases are the nicest bases and diagonal matrices are the nicest matrices, this leads to an obvious next question.

Question (Orthogonal Diagonalization Problem)

Given a square matrix A, does there exist an orthogonal matrix P for which $P^T AP$ is a diagonal matrix, and if so how does one find such a P? If such a matrix P exists, then A is said to be **orthogonally diagonalizable** and P is said to **orthogonally diagonalize** A.

Orthogonal Diagonalization Problem

Since orthonormal bases are the nicest bases and diagonal matrices are the nicest matrices, this leads to an obvious next question.

Question (Orthogonal Diagonalization Problem)

Given a square matrix A, does there exist an orthogonal matrix P for which $P^T AP$ is a diagonal matrix, and if so how does one find such a P? If such a matrix P exists, then A is said to be **orthogonally diagonalizable** and P is said to **orthogonally diagonalize** A.

Orthonormal Eigenvectors

Note that if an $n \times n$ matrix A is orthogonally diagonalizable then it would have to be diagonalizable.

Orthonormal Eigenvectors

Note that if an $n \times n$ matrix A is orthogonally diagonalizable then it would have to be diagonalizable.

Hence, it would have to have n linearly independent eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$ and be diagonalized by the matrix

$$P = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \end{pmatrix}$$

Note that if an $n \times n$ matrix A is orthogonally diagonalizable then it would have to be diagonalizable.

Hence, it would have to have n linearly independent eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$ and be diagonalized by the matrix

$$P = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \end{pmatrix}$$

Therefore, we see that A would be orthogonally diagonalizable if and only if P was orthogonal.

$A = P^{-1} D P$

If P ends up being orthogonal then $P^T = P^{-1}$

$A = P^T D P$

Orthonormal Eigenvectors

Note that if an $n \times n$ matrix A is orthogonally diagonalizable then it would have to be diagonalizable.

Hence, it would have to have n linearly independent eigenvectors $\vec{v}_1, \ldots, \vec{v}_n$ and be diagonalized by the matrix

$$P = \begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \ldots & \vec{v}_n \end{pmatrix}$$

if P is orthogonal $\Leftrightarrow$ $v_1 \ldots v_n$ form an orthonormal set.

Therefore, we see that A would be orthogonally diagonalizable if and only if P was orthogonal.

Theorem

An $n \times n$ matrix is orthogonally diagonalizable if and only if there exists an orthonormal set of n eigenvectors of A.

linearly independent.

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable.

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Proof.

A is orthogonaly diagonalizable if and only if $A = P^T DP$ for some orthogonal matrix P.

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Proof.

A is orthogonaly diagonalizable if and only if $A = P^T DP$ for some orthogonal matrix P. Then

$$A^T$$

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Proof.

A is orthogonaly diagonalizable if and only if $A = P^T DP$ for some orthogonal matrix P. Then

$$A^T = (P^T DP)^T$$

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Proof.

A is orthogonaly diagonalizable if and only if $A = P^T D P$ for some orthogonal matrix P. Then

$$A^T = (P^T D P)^T = P^T D^T (P^T)^T$$

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Proof.

A is orthogonaly diagonalizable if and only if $A = P^T DP$ for some orthogonal matrix P. Then

$$A^T = (P^T DP)^T = P^T D^T (P^T)^T = P^T DP$$

$D = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ $\qquad D^T = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = D$

Condition for Orthogonally Diagonalizable

We wish to find an easier property for when A is orthogonally diagonalizable. Indeed, we immediately see one condition.

Theorem

If A is orthogonally diagonalizable than A is symmetric (i.e. $A^T = A$).

Proof.

A is orthogonaly diagonalizable if and only if $A = P^T DP$ for some orthogonal matrix P. Then

$$A^T = (P^T DP)^T = P^T D^T (P^T)^T = P^T DP = A$$

$\square$

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2)$$

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2$$

$$(\lambda \vec{v}_1) \cdot (\vec{v}_1)$$

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2$$

v_1 is eigenvector of A with eigenvalue λ_1

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$

Orthogonal Eigenvectors

$A^T = A$

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$

$$= \vec{v}_1^T A \vec{v}_2$$

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$

$$= \vec{v}_1^T A \vec{v}_2 = \vec{v}_1^T (\lambda_2 \vec{v}_2)$$

V_2 is an eigenvector of A
with eigenvalue λ_2

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$

$$= \vec{v}_1^T A \vec{v}_2 = \vec{v}_1^T (\lambda_2 \vec{v}_2) = \lambda_2(\vec{v}_1 \cdot \vec{v}_2)$$

$(\vec{v}_1) \cdot (\lambda_2 \vec{v}_1)$

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$

$$= \vec{v}_1^T A \vec{v}_2 = \vec{v}_1^T (\lambda_2 \vec{v}_2) = \lambda_2(\vec{v}_1 \cdot \vec{v}_2)$$

Now, since $\lambda_1 \neq \lambda_2$, this can only happen if $\vec{v}_1 \cdot \vec{v}_2 = 0$

Orthogonal Eigenvectors

Theorem

If A is a symmetric matrix and $\vec{v}_1$ and $\vec{v}_2$ are two eigenvectors of A corresponding to two different eigenvalues λ_1 and λ_2, then $\vec{v}_1$ and $\vec{v}_2$ are orthogonal.

Proof.

$$\lambda_1(\vec{v}_1 \cdot \vec{v}_2) = (\lambda_1 \vec{v}_1)^T \vec{v}_2 = (A\vec{v}_1)^T \vec{v}_2 = \vec{v}_1^T A^T \vec{v}_2$$

$$= \vec{v}_1^T A\vec{v}_2 = \vec{v}_1^T(\lambda_2 \vec{v}_2) = \lambda_2(\vec{v}_1 \cdot \vec{v}_2)$$

Now, since $\lambda_1 \neq \lambda_2$, this can only happen if $\vec{v}_1 \cdot \vec{v}_2 = 0$, hence they are orthogonal. □

Symmetric is Orthogonally Diagonalizable

Theorem

A square matrix A is orthogonally diagonalizable if and only if it is symmetric.

$(\Rightarrow)$ done.

$(\Leftarrow)$ almost done...

What we haven't shown yet is that there are n linearly independent eigenvectors.

eigenvectors corresponding to the <u>same</u> eigenvalue can be be orthogonal to each other.

Symmetric is Orthogonally Diagonalizable

Theorem

A square matrix A is orthogonally diagonalizable if and only if it is symmetric.

Hard part of this theorem is to prove that if A is symmetric then it is diagonalizable.

Symmetric is Orthogonally Diagonalizable

Theorem

A square matrix A is orthogonally diagonalizable if and only if it is symmetric.

Hard part of this theorem is to prove that if A is symmetric then it is diagonalizable. We won't prove this, but Exercise P4 on pg.480 of the textbook goes through the proof step by step.

Symmetric is Orthogonally Diagonalizable

Theorem

A square matrix A is orthogonally diagonalizable if and only if it is symmetric.

Hard part of this theorem is to prove that if A is symmetric then it is diagonalizable. We won't prove this, but Exercise P4 on pg.480 of the textbook goes through the proof step by step.

However, once we know that if A is symmetric then it is diagonalizable, we can perform the Gram-Schmidt process to each eigenspace and find orthonormal bases for each eigenspace

λ is eigenvalue $\quad E_\lambda = \text{nul}(A - \lambda I_n) = \{v : Av = \lambda v\}$

$= \text{span}\{v_1, \dots, v_g\}$ (g = geometric multiplicity of λ)

performing G-S on $v_1 \dots v_g$ creates orthonormal eigenvectors

Symmetric is Orthogonally Diagonalizable

Theorem

A square matrix A is orthogonally diagonalizable if and only if it is symmetric.

Hard part of this theorem is to prove that if A is symmetric then it is diagonalizable. We won't prove this, but Exercise P4 on pg.480 of the textbook goes through the proof step by step.

However, once we know that if A is symmetric then it is diagonalizable, we can perform the Gram-Schmidt process to each eigenspace and find orthonormal bases for each eigenspace and use the previous theorem to guarantee that combining these will form a set of n orthonormal eigenvectors.

Orthogonally Diagonalizing a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix.

Orthogonally Diagonalizing a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix. Then to orthogonally diagonalize it, we do the following process

Orthogonally Diagonalizing a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix. Then to orthogonally diagonalize it, we do the following process

1. Find $\lambda_1, \ldots, \lambda_k$, the eigenvalues of A

Orthogonally Diagonalizing a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix. Then to orthogonally diagonalize it, we do the following process

1. Find $\lambda_1, \ldots, \lambda_k$, the eigenvalues of A
2. For each eigenvalue λ_i, find a basis for the eigenspace $E_{\lambda_i} = \text{span}\{\vec{v}_{i,1}, \ldots, \vec{v}_{i,g_i}\}$

Orthogonally Diagonalizing a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix. Then to orthogonally diagonalize it, we do the following process

1. Find $\lambda_1, \ldots, \lambda_k$, the eigenvalues of A
2. For each eigenvalue λ_i, find a basis for the eigenspace $E_{\lambda_i} = \text{span}\{\vec{v}_{i,1}, \ldots, \vec{v}_{i,g_i}\}$
3. Perform the Gram-Schmidt process on each basis of E_{λ_i} to find an orthonormal basis for $E_{\lambda_i} = \text{span}\{\vec{u}_{i,1}, \ldots, \vec{u}_{i,g_i}\}$

Orthogonally Diagonalizing a Symmetric Matrix

Let A be an $n \times n$ symmetric matrix. Then to orthogonally diagonalize it, we do the following process

1. Find $\lambda_1, \dots, \lambda_k$, the eigenvalues of A
2. For each eigenvalue λ_i, find a basis for the eigenspace $E_{\lambda_i} = \text{span}\{\vec{v}_{i,1}, \dots, \vec{v}_{i,g_i}\}$
3. Perform the Gram-Schmidt process on each basis of E_{λ_i} to find an orthonormal basis for $E_{\lambda_i} = \text{span}\{\vec{u}_{i,1}, \dots, \vec{u}_{i,g_i}\}$
4. The resulting set of n eigenvectors

$$\{\underbrace{\vec{u}_{1,1}, \dots, \vec{u}_{1,g_1}}_{E_{\lambda_1}}, \underbrace{\vec{u}_{2,1}, \dots}_{E_{\lambda_1}}, \dots, \underbrace{\vec{u}_{k,g_k}}_{E_{\lambda_k}}\}$$

are orthonormal.

Example

Exercise

Find a matrix P that orthogonally diagonalizes the matrix

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

Example

Exercise

Find a matrix P that orthogonally diagonalizes the matrix

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

First: check that it is symmetric and so can be orthogonally diagonalized.

Example

Exercise

Find a matrix P that orthogonally diagonalizes the matrix

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

First: check that it is symmetric and so can be orthogonally diagonalized.

Now, find the eigenvectors.

Example

Exercise

Find a matrix P that orthogonally diagonalizes the matrix

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

First: check that it is symmetric and so can be orthogonally diagonalized.

Now, find the eigen~~vectors~~values. A routine computation shows that

char poly of A

$$\det(A - tI_3) = (t-2)^2(t-8)$$

Example

Exercise

Find a matrix P that orthogonally diagonalizes the matrix

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

Since A is symmetric it is diagonalizable so the geometric multiplicity must be the same as the arithmetic multi. Hence, we should 2 linear ind eigenvectors for the eigenvalue 2 and only 1 for 8

First: check that it is symmetric and so can be orthogonally diagonalized.

Now, find the eigenvectors. A routine computation shows that

$$\det(A - tI_3) = (t-2)^2(t-8)^{1}$$

arithmetic multiplicity of 2 is 2
arithmetic multiplicity of 8 is 1

and so the eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = 8$.

Example 2

Row reduce

$$A - 2I_3 = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \qquad \text{and} \qquad A - 8I_3 = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}$$

$E_2 = \text{null}(A - 2I_3)$ $\qquad$ $E_8 = \text{null}(A - 8I_3)$

Example 2

Row reduce

$$A - 2I_3 = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \quad \text{and} \quad A - 8I_3 = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}$$

to find that $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$ and $E_8 = \text{span}\{\vec{v}_3\}$ where

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

Example 2

Row reduce

$$A - 2I_3 = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix} \qquad \text{and} \qquad A - 8I_3 = \begin{pmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{pmatrix}$$

to find that $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$ and $E_8 = \text{span}\{\vec{v}_3\}$ where

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

Note: a previous theorem tells us that $\vec{v}_3$ should be orthogonal to $\vec{v}_1$ and $\vec{v}_2$ a quick calculation confirms this. v_1 is not orthogonal to v_2.

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1} \vec{v}_2$$

we do this so that

w_1 & w_2 are orthogonal

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2} \vec{v}_1$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\begin{bmatrix}-1\\1\\2\end{bmatrix} = \vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2}\vec{v}_1 = \begin{bmatrix}-1\\-1\\2\end{bmatrix}$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2} \vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|} \vec{w}_1$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

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

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2}\vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$\|w_1\| = \sqrt{2}$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2} \vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|} \vec{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\|\vec{w}_2\|} \vec{w}_2$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} = \vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2} \vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|} \vec{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\|\vec{w}_2\|} \vec{w}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

$\|w_2\| = \sqrt{6}$

$E_2 = \text{span}\{u_1, u_2\}$ & $\{u_1, u_2\}$ is an orthonormal basis for E_2.

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2}\vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

Performing Gram-Schmidt of $E_8 = \text{span}\{\vec{v}_3\}$, we get

$$\vec{w}_3 = \vec{v}_3$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1}\vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2}\vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|}\vec{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\|\vec{w}_2\|}\vec{w}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

Performing Gram-Schmidt of $E_8 = \text{span}\{\vec{v}_3\}$, we get

$$\vec{w}_3 = \vec{v}_3 \qquad \vec{u}_3 = \frac{1}{\|\vec{w}_3\|}\vec{w}_3$$

Example 3

Performing Gram-Schmidt on $E_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$, we get

$$\vec{w}_1 = \vec{v}_1 \qquad \vec{w}_2 = \vec{v}_2 - \text{proj}_{\vec{v}_1} \vec{v}_2 = \vec{v}_2 - \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\|^2} \vec{v}_1 = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

$$\vec{u}_1 = \frac{1}{\|\vec{w}_1\|} \vec{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \qquad \vec{u}_2 = \frac{1}{\|\vec{w}_2\|} \vec{w}_2 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

Performing Gram-Schmidt of $E_8 = \text{span}\{\vec{v}_3\}$, we get

$\{u_1, u_2, u_3\}$

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \vec{w}_3 = \vec{v}_3 \qquad \vec{u}_3 = \frac{1}{\|\vec{w}_3\|} \vec{w}_3 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

forms an orthonormal basis for $\mathbb{R}^3$.

Hence, we see that

$$P$$

Example 4

Hence, we see that

$$P = \begin{array}{c} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix} \\ \begin{array}{ccc} u_1 \quad & \quad u_2 \quad & \quad u_3 \end{array} \end{array}$$

Example 4

Hence, we see that

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

diagonalizes A

Example 4

Hence, we see that

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

diagonalizes A and, in fact

$$P^T AP$$

Hence, we see that

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

diagonalizes A and, in fact

$$P^T A P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$u_1 \quad u_2 \quad u_3$

P^T

Hence, we see that

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

Exercise: do this triple matrix product to see that you indeed get the diagonal!

diagonalizes A and, in fact

$$P^T A P = \begin{bmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix} \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Calculating Powers of Diagonalizable Matrices

It is common that we wish to multiply matrices together.

Calculating Powers of Diagonalizable Matrices

It is common that we wish to multiply matrices together. However, this is computationally taxing as you have seen by now.

Calculating Powers of Diagonalizable Matrices

It is common that we wish to multiply matrices together. However, this is computationally taxing as you have seen by now. However, in the special case of raising diagonalizable matrices to a power, it becomes somewhat easy.

Calculating Powers of Diagonalizable Matrices

It is common that we wish to multiply matrices together. However, this is computationally taxing as you have seen by now. However, in the special case of raising diagonalizable matrices to a power, it becomes somewhat easy.

Theorem

If A is diagonalizable by P with diagonal matrix D, then $A = PDP^{-1}$ and for any k,

$$A^k = PD^kP^{-1}$$

k times

$$A^k = (PDP^{-1})^k = (PDP^{-1})(PDP^{-1})(PDP^{-1})\cdots(PDP^{-1})$$

k times

$$= P D \underbrace{P^{-1} P}_{I} D \underbrace{P^{-1} P}_{I} D \underbrace{P^{-1}\cdots}_{I} \cdots \underbrace{\cdots P}_{I} D P^{-1} = P\, D D D \cdots D\, P^{-1} = P D^k P^{-1}$$

(Tempting to say $(PDP^{-1})^k = P^k D^k P^{-1^k}$ but this not correct!!!!)

Powers of Diagonals

This is useful since calculating D^k when D is diagonal is easy:

Powers of Diagonals

This is useful since calculating D^k when D is diagonal is easy:

$$\left(\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}\right)^k$$

Powers of Diagonals

This is useful since calculating D^k when D is diagonal is easy:

$$\left(\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}\right)^k = \begin{pmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{pmatrix}$$

only true for diagonal!!!

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^k \neq \begin{pmatrix} a^k & b^k \\ c^k & d^k \end{pmatrix}$$

Powers of Diagonals

This is useful since calculating D^k when D is diagonal is easy:

$$\left(\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}\right)^k = \begin{pmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{pmatrix}$$

Exercise

Use these ideaes to compute A^{13} for

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$$

Powers of Diagonals

This is useful since calculating D^k when D is diagonal is easy:

$$\left(\begin{pmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{pmatrix}\right)^k = \begin{pmatrix} d_1^k & 0 & \dots & 0 \\ 0 & d_2^k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n^k \end{pmatrix}$$

Exercise

Use these ideaes to compute A^{13} for

$$A = \begin{pmatrix} 0 & 0 & -2 \\ 1 & 2 & 1 \\ 1 & 0 & 3 \end{pmatrix}$$

Example Solution

Using all the techniques we have developed so far, one can show that

$$A = \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}$$

Example Solution

Using all the techniques we have developed so far, one can show that

$$A = \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}$$

Thus,

$$A^{13} = \left(\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1} \right)^{13}$$

Using all the techniques we have developed so far, one can show that

$$A = \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}$$

Thus,

$$A^{13} = \left(\overset{P}{\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}} \overset{D}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}} \overset{P^{-1}}{\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}} \right)^{13}$$

$$= \overset{P}{\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}} \overset{D^{13}}{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2^{13} & 0 \\ 0 & 0 & 2^{13} \end{pmatrix}} \overset{P^{-1}}{\begin{pmatrix} -2 & -1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}^{-1}}$$

Cayley-Hamilton Theorem

Recall that the characteristic polynomial of a square matrix is $\det(A - tI_n)$.

t is a variable

Cayley-Hamilton Theorem

Recall that the characteristic polynomial of a square matrix is $\det(A - tI_n)$.

Theorem (Cayley-Hamilton Theorem)

If we write

$$\det(A - tI_n) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$$

Recall that the characteristic polynomial of a square matrix is $\det(A - tI_n)$.

Theorem (Cayley-Hamilton Theorem)

If we write

$$\det(A - tI_n) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$$

no t

then we get that

include I_n here.

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n = 0$$

" $\det(A - AI_n) = \det(A - A) = \det(0) = 0$ "

0 Matrix

not the number zero

Cayley-Hamilton Theorem

Recall that the characteristic polynomial of a square matrix is $\det(A - tI_n)$.

Theorem (Cayley-Hamilton Theorem)

If we write

$$\det(A - tI_n) = t^n + c_{n-1}t^{n-1} + \cdots + c_1 t + c_0$$

then we get that

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n = 0$$

We then say the "every square matrix satisfies it's characteristic equation".

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^k P^{-1}$.

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^kP^{-1}$. We can then use this to show that

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I_n$$

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^kP^{-1}$. We can then use this to show that

requires a little work.

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I_n$$

$$= P(\underline{D^n + c_{n-1}D^{n-1} + \cdots + c_1D + c_0I_n})P^{-1}$$

Further, since D is diagonal, we get that the inner matrix will also be diagonal.

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^kP^{-1}$. We can then use this to show that

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n$$

$$= P(D^n + c_{n-1}D^{n-1} + \cdots + c_1 D + c_0 I_n)P^{-1}$$

Further, since D is diagonal, we get that the inner matrix will also be diagonal. Moreover, the diagonal entries of D will be the eigenvalues of A: $\lambda_1, \ldots, \lambda_n$.

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^kP^{-1}$. We can then use this to show that

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I_n$$

$$= P(D^n + c_{n-1}D^{n-1} + \cdots + c_1D + c_0I_n)P^{-1}$$

Further, since D is diagonal, we get that the inner matrix will also be diagonal. Moreover, the diagonal entries of D will be the eigenvalues of A: $\lambda_1, \ldots, \lambda_n$. So, the diagonal entries of the inner matrix will be

$$\lambda_i^n + c_{n-1}\lambda_i^{n-1} + \cdots + c_1\lambda_i + c_0$$

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^kP^{-1}$. We can then use this to show that

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1A + c_0I_n$$

$$= P(D^n + c_{n-1}D^{n-1} + \cdots + c_1D + c_0I_n)P^{-1}$$

Further, since D is diagonal, we get that the inner matrix will also be diagonal. Moreover, the diagonal entries of D will be the eigenvalues of A: $\lambda_1, \ldots, \lambda_n$. So, the diagonal entries of the inner matrix will be

$$\lambda_i^n + c_{n-1}\lambda_i^{n-1} + \cdots + c_1\lambda_i + c_0 = \det(A - \lambda_i I_n)$$

$\det(A - tI_n) = t^n + c_{n-1}t^{n-1} + \cdots + c_0$

Sketch of Proof for Diagonalizable A

If $A = PDP^{-1}$ is diagonalizable then we get $A^k = PD^kP^{-1}$. We can then use this to show that

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I_n$$

$$= P(\underbrace{D^n + c_{n-1}D^{n-1} + \cdots + c_1 D + c_0 I_n}_{*})P^{-1}$$

Further, since D is diagonal, we get that the inner matrix will also be diagonal. Moreover, the diagonal entries of D will be the eigenvalues of A: $\lambda_1, \ldots, \lambda_n$. So, the diagonal entries of the inner matrix will be

$$\lambda_i^n + c_{n-1}\lambda_i^{n-1} + \cdots + c_1\lambda_i + c_0 = \det(A - \lambda_i I_n) = 0$$

by definition of an eigenvalue.

$*$ is a diagonal whose i^{th} diagonal is 0 for all i
Hence $* = 0$ & so $P * P^{-1} = 0$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$.

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$\det(A - tI_2)$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right)$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2\times 3$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2 \times 3 = t^2 - 5t - 2$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2\times 3 = t^2 - 5t - 2$$

Hence,

$$A^2 - 5A - 2I_2$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2 \times 3 = t^2 - 5t - 2$$

Hence,

$$A^2 - 5A - 2I_2 = \underbrace{\underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}}_{A}}_{A^2} - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2 \times 3 = t^2 - 5t - 2$$

Hence,

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix}$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2 \times 3 = t^2 - 5t - 2$$

Hence,

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix}$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2 \times 3 = t^2 - 5t - 2$$

Hence,

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Example

Let $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$. Then it's characteristic polynomial will be

$$\det(A - tI_2) = \det\left(\begin{pmatrix} 1-t & 2 \\ 3 & 4-t \end{pmatrix}\right) = (1-t)(4-t) - 2 \times 3 = t^2 - 5t - 2$$

Hence,

$$A^2 - 5A - 2I_2 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 5\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$= \begin{bmatrix} 7 & 10 \\ 15 & 22 \end{bmatrix} - \begin{pmatrix} 5 & 10 \\ 15 & 20 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$