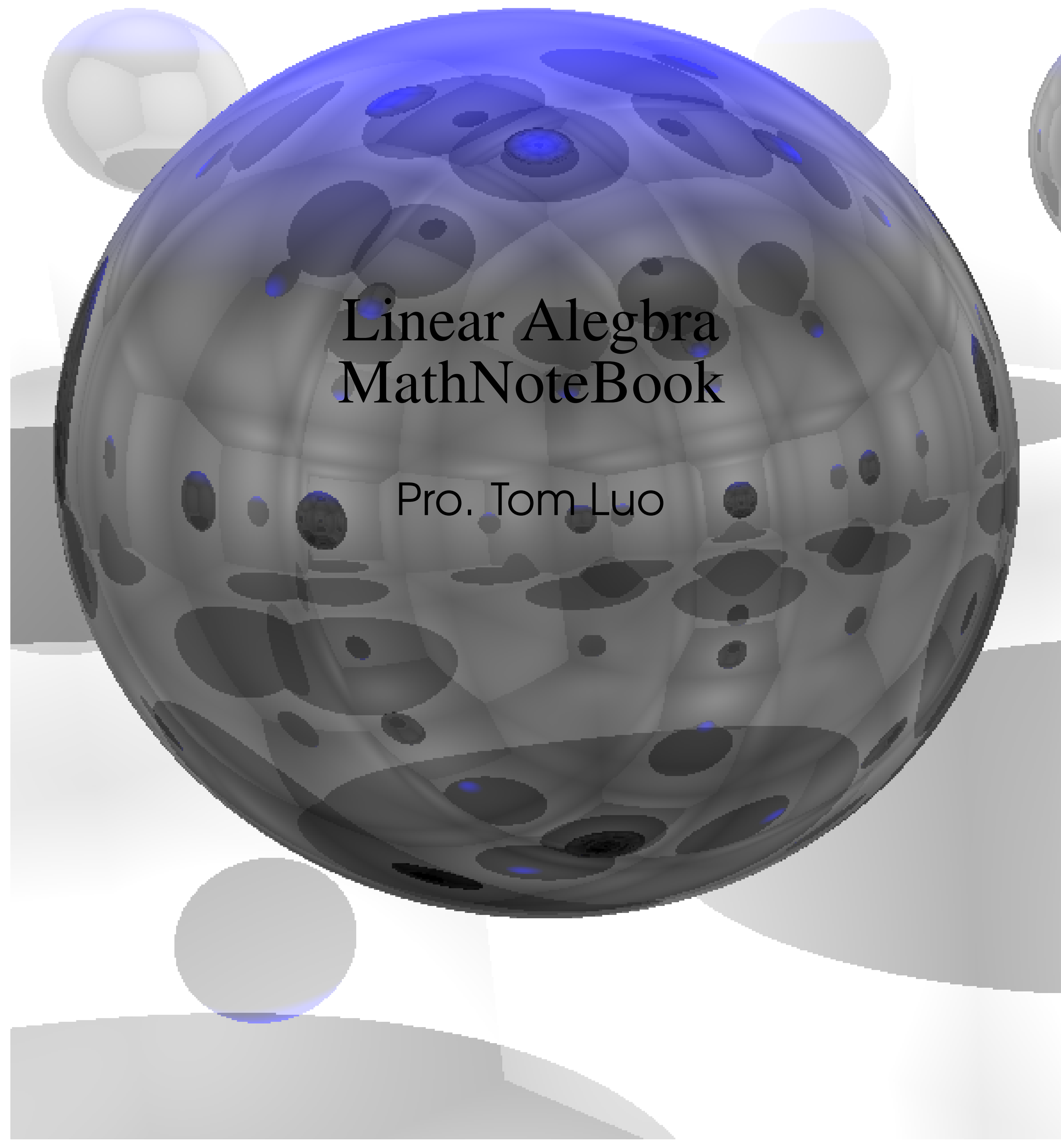

Linear Alegbra
MathNoteBook

Pro. Tom Luo

15 — Week7

15.1 Thursday

15.1.1 Review

- **eigenvalue and eigenvectors**: If for square matrix $\boldsymbol{A}$ we have

  $$\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$$

  where $\boldsymbol{x} \neq \boldsymbol{0}$, then we say λ is the *eigenvalue*, $\boldsymbol{x}$ is the *eigenvector* corresponding to λ.
- **How to compute eigenvalues and eigenvectors?** To solve the eigenvalue problem for an n by n matrix, you should follow these steps:
  - *Compute the determinant of* $\lambda\boldsymbol{I} - \boldsymbol{A}$. The determinant is a polynomial in λ of degree n.
  - *Find the roots of this polynomial*, by solving $\det(\lambda\boldsymbol{I} - \boldsymbol{A}) = 0$. The n roots are the n eigenvalues of $\boldsymbol{A}$. They make $\boldsymbol{A} - \lambda\boldsymbol{I}$ singular.
  - For each eigenvalue λ, *Solve* $(\lambda\boldsymbol{I} - \boldsymbol{A})\boldsymbol{x} = \boldsymbol{0}$ *to find an eigenvector* $\boldsymbol{x}$.

15.1.2 Similarity and eigenvalues

Which two matrices have the same eigenvalues? The similar matrices have the same eigenvalues:

Definition 15.1 — Similar. If there exists a *nonsingular* matrix $\boldsymbol{S}$ such that

$$\boldsymbol{B} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S},$$

then we say $\boldsymbol{A}$ is **similar** to $\boldsymbol{B}$.

Proposition 15.1 Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be $n \times n$ matrices. If $\boldsymbol{B}$ is *similar* to $\boldsymbol{A}$, then $\boldsymbol{A}$ and $\boldsymbol{B}$ have the same eigenvalues.

Proofidea. Since eigenvalues are the roots of the *characteristic polynomial*, so it suffices to prove these two polynomials are the same.

Proof. The *characteristic polynomial* for $\boldsymbol{B}$ is given by

$$\begin{aligned}
P_{\boldsymbol{B}}(\lambda) &= \det(\lambda \boldsymbol{I} - \boldsymbol{B}) \\
&= \det(\lambda \boldsymbol{I} - \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}) = \det(\boldsymbol{S}^{-1}\lambda \boldsymbol{I}\boldsymbol{S} - \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}) \\
&= \det(\boldsymbol{S}^{-1}(\lambda \boldsymbol{I} - \boldsymbol{A})\boldsymbol{S}) \\
&= \det(S^{-1})\det(\lambda \boldsymbol{I} - \boldsymbol{A})\det(\boldsymbol{S})
\end{aligned}$$

Since $\det(\boldsymbol{S}^{-1})\det(\boldsymbol{S}) = 1$, we obtain:

$$\begin{aligned}
P_{\boldsymbol{B}}(\lambda) &= \det(\lambda \boldsymbol{I} - \boldsymbol{A}) \\
&= P_{\boldsymbol{A}}(\lambda).
\end{aligned}$$

Since they have the same *characteristic polynomial*, the roots for *characteristic polynomials* of $\boldsymbol{A}$ and $\boldsymbol{B}$ must be same. Hence they have the same eigenvalues. ∎

R What is invarient? In other words, what is not changed during matrix transformation?
- **Rank** is invarient under *row transformation*.
- **Eigenvalues** is invarient undet *similar transformation*.
- Unluckily, similar matrices usually don't have the same eigenvectors. It's easy to raise a counterexample.

By using eigenvalues, we have a new proof for $\det(\boldsymbol{S}^{-1}) = \frac{1}{\det(\boldsymbol{S})}$.

Proof. Suppose $\det(\boldsymbol{S}) = \lambda_1\lambda_2\dots\lambda_n$, where λ_i's are eigenvalues of $\boldsymbol{S}$.
Then there exists $\boldsymbol{x}_i$ such that

$$\boldsymbol{S}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i$$

for $i = 1,\dots,n$.
Since $\boldsymbol{S}$ is invertible, all λ_i's are nonzero, and we obain:

$$\boldsymbol{x}_i = \lambda_i \boldsymbol{S}^{-1}\boldsymbol{x}_i \implies \frac{1}{\lambda_i}\boldsymbol{x}_i = \boldsymbol{S}^{-1}\boldsymbol{x}_i$$

Or equivalently, $\boldsymbol{S}^{-1}\boldsymbol{x}_i = \frac{1}{\lambda_i}\boldsymbol{x}_i$. $\frac{1}{\lambda_i}$'s are eigenvalues of $\boldsymbol{S}^{-1}$.
Since S^{-1} is $n \times n$ matrix, $\frac{1}{\lambda_i}$'s ($i = 1,\dots,n$) are the only eigenvalues of $\boldsymbol{S}^{-1}$.
Hence the determinant of $\boldsymbol{S}^{-1}$ is the product of eigenvalues:

$$\det(\boldsymbol{S}^{-1}) = \frac{1}{\lambda_1}\frac{1}{\lambda_2}\cdots\frac{1}{\lambda_n} = \frac{1}{\det(\boldsymbol{S})}.$$

∎

We can also use eigenvalue to proof the statement below:

Proposition 15.2 $\boldsymbol{A}$ is singular if and only if $\det(\boldsymbol{A}) = 0$.

Proof. Suppose $\det(\boldsymbol{A}) = \lambda_1\lambda_2\dots\lambda_n$, where λ_i's are eigenvalues of $\boldsymbol{A}$.
Thus

$$\det(\boldsymbol{A}) = 0 \iff \exists \lambda_i = 0 \iff \exists \text{ nonzero } \boldsymbol{x} \text{ s.t. } \boldsymbol{A}\boldsymbol{x} = \lambda_i \boldsymbol{x} = 0\boldsymbol{x} = \boldsymbol{0}.$$

Equivalently, $\boldsymbol{A}$ is singular. ∎

15.1.3 Diagonalization

Proposition (15.1) says if $\boldsymbol{A}$ is similar to $\boldsymbol{B}$, then they have the same eigenvalues.

- Q1: *What about the reverse direction?*
- What's the simplest form of matrix to have eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$?

  We can answer this question immediately. The matrix $\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$ has the simplest form. And we often write this matrix as $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

  Q2: What we want to ask is that *if* $\boldsymbol{A}$ *has eigenvalues* $\lambda_1, \dots, \lambda_n$, *then* $\boldsymbol{A}$ *and* $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ *have the same eigenbalues. Are they similar?*

(R) Why the matrix $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ has eigenvalues $\lambda_1, \dots, \lambda_n$?
Answer: Let's explain it with $n = 2$:

$$\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda_1 \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda_2 \end{pmatrix} = \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

General n is also easy to verify.

The answer to question 1 and 2 are both No! Let's raise a counterexample to explain it:

Example 15.1
If $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, then $P_{\boldsymbol{A}}(\lambda) = \det(\lambda \boldsymbol{I} - \boldsymbol{A}) = \begin{vmatrix} \lambda & -1 \\ 0 & \lambda \end{vmatrix}$. Hence its eigenvalues are $\lambda_1 = \lambda_2 = 0$.
And $\boldsymbol{A}$ and $\boldsymbol{D} = \operatorname{diag}(0,0)$ have the same eigenvalues. Are they similar?
We assume they are similar, which means there exists invertible matrix $\boldsymbol{S}$ such that

$$\boldsymbol{A} = \boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S} = \boldsymbol{S}^{-1} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \boldsymbol{S} = \boldsymbol{0}$$

which leads to a contradiction! So $\boldsymbol{A}$ and $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \lambda_2)$ are not similar. ■

Suppose $\boldsymbol{A}$ has eigenvalues $\lambda_1, \dots, \lambda_n$, but $\boldsymbol{A}$ and $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$ may not be similar! But which matrix is similar to its diagonal matrix $\operatorname{diag}(\lambda_1, \dots, \lambda_n)$?

Definition 15.2 — Diagonalizable. An $n \times n$ matrix $\boldsymbol{A}$ is **diagonalizable** if $\boldsymbol{A}$ is similar to a *diagonal matrix*, that is to say, $\exists$ nonsingular matrix $\boldsymbol{S}$ and diagonal matrix $\boldsymbol{D}$ such that

$$\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}$$

We say $\boldsymbol{S}$ *diagonalize* $\boldsymbol{A}$. ■

(R) If $\boldsymbol{A}$ and diagonal matrix $\boldsymbol{D}$ are similar, then they have the same eigenvalues. If $n \times n$ matrix $\boldsymbol{A}$ has eigenvalues $\lambda_1, \dots, \lambda_n$, then $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

Why is diagonalizable good?

Theorem 15.1 — Diagonalization.
An $n \times n$ matrix $\boldsymbol{A}$ is *diagonalizable* iff $\boldsymbol{A}$ has n ind. eigenvectors.

Proof.
Necessity. Suppose $\boldsymbol{A}$ has n ind. eigenvectors $\boldsymbol{x}_i$ for $i = 1, \dots, n$. And we assume $\exists \lambda_i$ such that

$$\boldsymbol{A}\boldsymbol{x}_i = \lambda_i \boldsymbol{x}_i \text{ for } i = 1, \dots, n.$$

We multiply $\boldsymbol{A}$ with $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix}$. The first column of $\boldsymbol{AS}$ is $\boldsymbol{A}\boldsymbol{x}_1$, that is $\lambda_1 \boldsymbol{x}_1$. Then we obtain:

$$\textbf{\textit{A} times \textit{S}} \qquad \boldsymbol{AS} = \boldsymbol{A}\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1\boldsymbol{x}_1 & \lambda_2\boldsymbol{x}_2 & \dots & \lambda_n\boldsymbol{x}_n \end{bmatrix}.$$

The trick is to split this matrix $\boldsymbol{AS}$ into $\boldsymbol{S}$ times $\boldsymbol{D}$:

$$\textbf{\textit{S} times \textit{D}} \qquad \begin{bmatrix} \lambda_1\boldsymbol{x}_1 & \lambda_2\boldsymbol{x}_2 & \dots & \lambda_n\boldsymbol{x}_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \boldsymbol{SD}.$$

Hence we obtain $\boldsymbol{AS} = \boldsymbol{SD}$. Since $\boldsymbol{x}_i$'s are ind, there exists the inverse $\boldsymbol{S}^{-1}$.
So $\boldsymbol{D} = \boldsymbol{S}^{-1}\boldsymbol{AS}$.
Sufficiency. If $\boldsymbol{A}$ is diagonalizable, then there exists $\boldsymbol{S}$ and $\boldsymbol{D}$ such that

$$\boldsymbol{D} = \boldsymbol{S}^{-1}\boldsymbol{AS}$$

where $\boldsymbol{S}$ is nonsingular. And we assume $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.
Suppose $\boldsymbol{S} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \dots & \boldsymbol{x}_n \end{bmatrix}$, where $\boldsymbol{x}_i$'s are ind.
Then from equation $\boldsymbol{D} = \boldsymbol{S}^{-1}\boldsymbol{AS}$ we obtain $\boldsymbol{AS} = \boldsymbol{SD} \implies \boldsymbol{A}\boldsymbol{x}_i = \lambda_i\boldsymbol{x}_i$ for $i = 1, 2, \dots, n$.
Hence λ_i's are eigenvalues and $\boldsymbol{x}_i$'s are ind. eigenvectors of $\boldsymbol{A}$. ∎

For $n \times n$ matrix $\boldsymbol{A}$ which is *diagonalizable*, if its eigenvectors $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_n\}$ form a basis, then for any $\boldsymbol{y} \in \mathbb{R}^n$, there exists $(c_1, c_2, \dots, c_n)$ such that

$$\boldsymbol{y} = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \cdots + c_n\boldsymbol{x}_n$$

If we consider matrix $\boldsymbol{A}$ as representation of linear transformation, we obtain

$$\begin{aligned} \boldsymbol{Ay} &= c_1\boldsymbol{A}\boldsymbol{x}_1 + \cdots + c_n\boldsymbol{A}\boldsymbol{x}_n \\ &= c_1\lambda_1\boldsymbol{x}_1 + \cdots + c_n\lambda_n\boldsymbol{x}_n \end{aligned}$$

So if we transform $\boldsymbol{y}$ into $\boldsymbol{Ay}$, it's equivalent to transform the coefficient $(c_1, \dots, c_n)$ into $(c_1\lambda_1, \dots, c_n\lambda_n)$.

$$\boldsymbol{y} \xRightarrow{\boldsymbol{A}} \boldsymbol{Ay}$$

$$(c_1, \dots, c_n) \xRightarrow{\boldsymbol{D}=\operatorname{diag}(\lambda_1,\dots,\lambda_n)} (c_1\lambda_1, \dots, c_n\lambda_n) = (c_1, \dots, c_n)\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

But is there an useful way to determine whether the eigenvectors of $\boldsymbol{A}$ is independent?

Theorem 15.2 If $\lambda_1,\dots,\lambda_k$ are *distinct* eigenvalues of a matrix $\boldsymbol{A}$ with corresponding eigenvectors $\boldsymbol{x}_1,\dots,\boldsymbol{x}_k$, then $\boldsymbol{x}_1,\dots,\boldsymbol{x}_k$ are linearly independent.

Proof. • Let's start with $k=2$. We assume $\lambda_1\neq\lambda_2$ but $\boldsymbol{x}_1,\boldsymbol{x}_2$ are dep. That is to say, $\exists(c_1,c_2)\neq\boldsymbol{0}$ s.t.

$$c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2=\boldsymbol{0}. \tag{15.1}$$

If we multiply $\boldsymbol{A}$ both sides, we obtain

$$\boldsymbol{A}(c_1\boldsymbol{x}_1+c_2\boldsymbol{x}_2)=\boldsymbol{0}\implies c_1\lambda_1\boldsymbol{x}_1+c_2\lambda_2\boldsymbol{x}_2=\boldsymbol{0}. \tag{15.2}$$

Eq(15.1)$\times\lambda_2-$Eq(15.2):

$$(c_1\lambda_2-c_1\lambda_1)\boldsymbol{x}=\boldsymbol{0}.\implies c_1(\lambda_2-\lambda_1)\boldsymbol{x}=\boldsymbol{0}.$$

Since $\lambda_1\neq\lambda_2$,$\boldsymbol{x}\neq\boldsymbol{0}$, we derive $c_2=0$.
Similarly, if we let Eq(15.1)$\times\lambda_1-$Eq(15.2) to cancel c_2, then we get $c_1=0$.
Hence $(c_1,c_2)=\boldsymbol{0}$ leads to contradiction!

• How to proof this statement for general k?
Assume there exists $(c_1,\dots,c_k)\neq\boldsymbol{0}$ s.t.

$$c_1\boldsymbol{x}_1+\cdots+c_k\boldsymbol{x}_k=\boldsymbol{0} \tag{15.3}$$

Then

$$\boldsymbol{A}(c_1\boldsymbol{x}_1+\cdots+c_k\boldsymbol{x}_k)=c_1\lambda_1+c_2\lambda_2+\cdots+c_k\lambda_k\boldsymbol{x}_k=\boldsymbol{0}. \tag{15.4}$$

We can let Eq$(11.4)-\lambda_k\times$Eq(11.3) to cancel $\boldsymbol{x}_k$:

$$c_1(\lambda_1-\lambda_k)\boldsymbol{x}_1+\cdots+c_k(\lambda_{k-1}-\lambda_k)\boldsymbol{x}_{k-1}=\boldsymbol{0}. \tag{15.5}$$

We can continue this process to cancel $\boldsymbol{x}_{k-1},\boldsymbol{x}_{k-2},\dots,\boldsymbol{x}_2$ to get:

$$c_1(\lambda_1-\lambda_k)\dots(\lambda_1-\lambda_2)\boldsymbol{x}_1=\boldsymbol{0}\qquad\text{which forces } c_1=0.$$

Similarly every $c_i=0$ for $i=1,\dots,n$. Here is the contradiction!

■

Corollary 15.1 If all eigenvalues of $\boldsymbol{A}$ are *distinct*, then $\boldsymbol{A}$ is *diagonalizable*

15.1.4 Powers of A

If $\boldsymbol{A}=\boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S}$, then $\boldsymbol{A}^2=(\boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S})(\boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S})=\boldsymbol{S}^{-1}\boldsymbol{D}^2\boldsymbol{S}$.
In general, $\boldsymbol{A}^k=(\boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S})\dots(\boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S})=\boldsymbol{S}^{-1}\boldsymbol{D}^k\boldsymbol{S}$.
We may ask if eigenvalues of $\boldsymbol{A}$ are $\lambda_1,\dots,\lambda_n$, then what is the eigenvalues of $\boldsymbol{A}^k$? The answer is intuitive, the eigenvalues of $\boldsymbol{A}^k$ are $\lambda_1^k,\dots,\lambda_n^k$. But you may use the wrong way to proof this statement:

Proposition 15.3 If eigenvalues of $n\times n$ matrix $\boldsymbol{A}$ are $\lambda_1,\dots,\lambda_n$, then eigenvalues of $\boldsymbol{A}^k$ are $\lambda_1^k,\dots,\lambda_n^k$.

Wrong proof 1: Assume $\boldsymbol{A} = \boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S}$, then $\boldsymbol{A}^k = \boldsymbol{S}^{-1}\boldsymbol{D}^k\boldsymbol{S}$. Suppose $\boldsymbol{D} = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$, then $\boldsymbol{D}^k = \operatorname{diag}(\lambda_1^k, \dots, \lambda_n^k)$. Hence eigenvalues of $\boldsymbol{A}^k$ are $\lambda_1^k, \dots, \lambda_n^k$.
This proof is wrong, because $\boldsymbol{A}$ may not be *diagonalizable*, which means $\boldsymbol{A}$ may not have the form $\boldsymbol{A} = \boldsymbol{S}^{-1}\boldsymbol{D}\boldsymbol{S}$. ■

Wrong proof 2: If $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x}$, then $\boldsymbol{A}^2\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{A}\boldsymbol{x}) = \boldsymbol{A}(\lambda\boldsymbol{x}) = \lambda(\boldsymbol{A}\boldsymbol{x}) = \lambda^2\boldsymbol{x}$.
Hence for general k, $\boldsymbol{A}^k\boldsymbol{x} = \lambda^k\boldsymbol{x}$.
This proof only states that if λ is the eigenvalue of $\boldsymbol{A}$, then λ^k is the eigenvalues of $\boldsymbol{A}^k$. But it cannot derive this proposition.
Let's raise a counterexample: Let eigenvalues of $\boldsymbol{A}$ be $\lambda_1 = 1, \lambda_2 = 1, \lambda_3 = 2$, the eigenvalues of $\boldsymbol{A}^2$ be $1^2, 2^2, 2^2$. Then obviously, this $\boldsymbol{A}$ and $\boldsymbol{A}^2$ is a contradiction for this proof. Because $1, 2$ are the eigenvalues of $\boldsymbol{A}$, but this proof fails to determine its multiplicity! ■

15.1.5 Nondiagonalizable Matrices

Sometimes we face some matrices that have too few eigenvalues. (don't count with multiplicity)
For example, if $\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, it's easy to verify that its eigenvalue is $\lambda = 0$ and eigenvectors are of the form $\boldsymbol{x} = \begin{bmatrix} c \\ 0 \end{bmatrix}$.
However, this 2×2 matrix cannot be diagonalized. Why? Let's introduce a definition:

Definition 15.3 — Eigenspace. Suppose $\boldsymbol{A}$ has k distinct eigenvalues $\lambda_1, \dots, \lambda_k$. Then the eigenspace for $\boldsymbol{A}$ is the union of all eigenvectors. Or say, the eigenspace is the union of all null space $N(\lambda_i\boldsymbol{I} - \boldsymbol{A})$ for $i = 1, \dots, k$. ■

Why this 2 by 2 matrix $\boldsymbol{A}$ cannot be diagonalizable? Because it has two repeated eigenvalues $\lambda_1 = \lambda_2 = 0$. And its eigenspace is of dimension $1 < 2$. In general, if a eigenspace for a $n \times n$ matrix has dimension $k < n$, then it cannot be diagonalizable.