Linear Algebra TSC

Aryaman Maithani
https://aryamanmaithani.github.io/tuts/ma-106

IIT Bombay

Spring 2022

Table of Contents

Matrices

We know what a matrix is. What a column (or row) matrix is. When (and how) we can add and multiply two matrices. What the transpose of a matrix is.

We know what a matrix is. What a column (or row) matrix is. When (and how) we can add and multiply two matrices. What the transpose of a matrix is.

A matrix is symmetric if $A^\mathsf{T} = A$ and skew-symmetric if $A^\mathsf{T} = -A$.

Matrices

We know what a matrix is. What a column (or row) matrix is. When (and how) we can add and multiply two matrices. What the transpose of a matrix is.

A matrix is symmetric if $A^{\mathsf{T}} = A$ and skew-symmetric if $A^{\mathsf{T}} = -A$.

If $\mathbf{v}$ and $\mathbf{w}$ are column vectors, then their dot product is given by $\mathbf{v}^{\mathsf{T}}\mathbf{w}$.

Matrices

We know what a matrix is. What a column (or row) matrix is. When (and how) we can add and multiply two matrices. What the transpose of a matrix is.

A matrix is symmetric if $A^{\mathsf{T}} = A$ and skew-symmetric if $A^{\mathsf{T}} = -A$.

If $\mathbf{v}$ and $\mathbf{w}$ are column vectors, then their dot product is given by $\mathbf{v}^{\mathsf{T}}\mathbf{w}$.

A square matrix A is called invertible if there exists a matrix B such that $AB = BA = \mathbf{I}$.

Matrices

We know what a matrix is. What a column (or row) matrix is. When (and how) we can add and multiply two matrices. What the transpose of a matrix is.

A matrix is symmetric if $A^{\mathsf{T}} = A$ and skew-symmetric if $A^{\mathsf{T}} = -A$.

If $\mathbf{v}$ and $\mathbf{w}$ are column vectors, then their dot product is given by $\mathbf{v}^{\mathsf{T}}\mathbf{w}$.

A square matrix A is called invertible if there exists a matrix B such that $AB = BA = \mathbf{I}$.

End of section.

Table of Contents

Linear system

Consider m linear equations in n variables $x_1, \ldots, x_n$:

Linear system

Consider m linear equations in n variables $x_1, \ldots, x_n$:

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}
$$

Linear system

Consider m linear equations in n variables $x_1, \ldots, x_n$:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m.
\end{aligned}$$

Making the obvious matrices out of the 'a_{ij}'s, 'x_i's, and 'b_i's, we can put the above in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

Linear system

Consider m linear equations in n variables $x_1, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Making the obvious matrices out of the 'a_{ij}'s, 'x_i's, and 'b_i's, we can put the above in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

where A is of size $m \times n$, $\mathbf{x}$ of size $n \times 1$, and $\mathbf{b}$ of size $m \times 1$.

Linear system

Consider m linear equations in n variables $x_1, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Making the obvious matrices out of the 'a_{ij}'s, 'x_i's, and 'b_i's, we can put the above in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

where A is of size $m \times n$, $\mathbf{x}$ of size $n \times 1$, and $\mathbf{b}$ of size $m \times 1$.

We have the augmented matrix defined by

Linear system

Consider m linear equations in n variables $x_1, \ldots, x_n$:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

Making the obvious matrices out of the 'a_{ij}'s, 'x_i's, and 'b_i's, we can put the above in matrix form as

$$A\mathbf{x} = \mathbf{b},$$

where A is of size $m \times n$, $\mathbf{x}$ of size $n \times 1$, and $\mathbf{b}$ of size $m \times 1$.

We have the augmented matrix defined by $A^+ := [A \mid \mathbf{b}]$, which completely captures the whole system.

There are three obvious things one can do without changing the set of solutions:

1. Interchanging the order of two equations.

Elementary row operations

There are three obvious things one can do without changing the set of solutions:

1. Interchanging the order of two equations.
2. Multiplying an equation by a scalar and adding it to some <u>other</u> equation of the system.

Elementary row operations

There are three obvious things one can do without changing the set of solutions:

1. Interchanging the order of two equations.
2. Multiplying an equation by a scalar and adding it to some other equation of the system.
3. Multiplying an equation by a nonzero number.

Elementary row operations

There are three obvious things one can do without changing the set of solutions:

1. Interchanging the order of two equations.
2. Multiplying an equation by a scalar and adding it to some <u>other</u> equation of the system.
3. Multiplying an equation by a <u>nonzero</u> number.

Elementary row operations

There are three obvious things one can do without changing the set of solutions:

1. Interchanging the order of two equations.
2. Multiplying an equation by a scalar and adding it to some other equation of the system.
3. Multiplying an equation by a nonzero number.

Corresponding to each of the operations above, there are obvious row operations that can be performed on A^+,

Elementary row operations

There are three obvious things one can do without changing the set of solutions:

1. Interchanging the order of two equations.
2. Multiplying an equation by a scalar and adding it to some other equation of the system.
3. Multiplying an equation by a nonzero number.

Corresponding to each of the operations above, there are obvious row operations that can be performed on A^+, called the elementary row operations (EROs).

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m \times m}$.

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m \times m}$.

Pick any ERO of your liking.

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m \times m}$.
Pick any ERO of your liking. Let E be the matrix obtained after applying that ERO on $\mathbf{I}$.
Now, let A be an arbitrary $m \times n$.

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m \times m}$.
Pick any ERO of your liking. Let E be the matrix obtained after applying that ERO on $\mathbf{I}$.
Now, let A be an arbitrary $m \times n$. Then,

EA is the same as applying *that* ERO on A.

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m \times m}$.
Pick any ERO of your liking. Let E be the matrix obtained after applying that ERO on $\mathbf{I}$.
Now, let A be an arbitrary $m \times n$. Then,

EA is the same as applying *that* ERO on A.

What this means is that one can perform row operations by pre-multiplying certain (square) matrices.

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m\times m}$.
Pick any ERO of your liking. Let E be the matrix obtained after applying that ERO on $\mathbf{I}$.
Now, let A be an arbitrary $m \times n$. Then,

EA is the same as applying *that* ERO on A.

What this means is that one can perform row operations by pre-multiplying certain (square) matrices.

A matrix that is obtained by performing an ERO on $\mathbf{I}$ is called an elementary row matrix (ERM).

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m\times m}$.
Pick any ERO of your liking. Let E be the matrix obtained after applying that ERO on $\mathbf{I}$.
Now, let A be an arbitrary $m \times n$. Then,

EA is the same as applying *that* ERO on A.

What this means is that one can perform row operations by pre-multiplying certain (square) matrices.

A matrix that is obtained by performing an ERO on $\mathbf{I}$ is called an elementary row matrix (ERM).

Note: All elementary row matrices are invertible.

EROs and ERMs

The three EROs on the previous slide can be applied on the identity matrix $\mathbf{I} = I_{m \times m}$.
Pick any ERO of your liking. Let E be the matrix obtained after applying that ERO on $\mathbf{I}$.
Now, let A be an arbitrary $m \times n$. Then,

EA is the same as applying *that* ERO on A.

What this means is that one can perform row operations by pre-multiplying certain (square) matrices.

A matrix that is obtained by performing an ERO on $\mathbf{I}$ is called an elementary row matrix (ERM).

Note: All elementary row matrices are invertible. (Why?)

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row starts with strictly more zeroes than the previous row.

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row starts with strictly more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row starts with strictly more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

$$\begin{bmatrix} \boxed{3} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 1 \end{bmatrix}$$

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row starts with strictly more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

$\begin{bmatrix} \boxed{3} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 1 \end{bmatrix}$ **is** in REF even though the first row has more zeroes in total.

Definition 2

A matrix in REF is said to be in reduced REF (RREF) if further:

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row starts with strictly more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

$\begin{bmatrix} \boxed{3} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 1 \end{bmatrix}$ **is** in REF even though the first row has more zeroes in total.

Definition 2

A matrix in REF is said to be in reduced REF (RREF) if further: (i) all pivots are 1,

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if
each row starts with strictly more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

$\begin{bmatrix} \boxed{3} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 1 \end{bmatrix}$ **is** in REF even though the first row has more zeroes in total.

Definition 2

A matrix in REF is said to be in reduced REF (RREF) if further: (i) all pivots are 1, and (ii) the entries above each pivot is 0.

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row starts with strictly more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

$\begin{bmatrix} \boxed{3} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 1 \end{bmatrix}$ **is** in REF even though the first row has more zeroes in total.

Definition 2

A matrix in REF is said to be in reduced REF (RREF) if further: (i) all pivots are 1, and (ii) the entries above each pivot is 0.

The matrix above is not in RREF.

(R)REF

Definition 1

An $m \times n$ matrix is said to be in row echelon form (REF) if each row <u>starts</u> with <u>strictly</u> more zeroes than the previous row.
The first nonzero element in a nonzero row is called the pivot of that row.

$\begin{bmatrix} \boxed{3} & 1 & 0 & 0 \\ 0 & \boxed{1} & 1 & 1 \end{bmatrix}$ **is** in REF even though the first row has more zeroes <u>in total</u>.

Definition 2

A matrix in REF is said to be in reduced REF (RREF) if further: (i) all pivots are 1, and (ii) the entries above each pivot is 0.

The matrix above is not in RREF. It violates *both* the conditions.

Gauss

Theorem 3

Given any $m \times n$ matrix A, one can perform elementary row operations on A to convert it into RREF.

Gauss

Theorem 3

Given any $m \times n$ matrix A, one can perform elementary row operations on A to convert it into RREF.
Equivalently, there exists elementary roe matrices $E_1, \ldots, E_N$ such that

$$E_N \cdots E_1 A$$

is in RREF.

Gauss

Theorem 3

Given any $m \times n$ matrix A, one can perform elementary row operations on A to convert it into RREF.
Equivalently, there exists elementary roe matrices $E_1, \ldots, E_N$ such that

$$E_N \cdots E_1 A$$

is in RREF.
Furthermore, the RREF is unique.

Theorem 3

Given any $m \times n$ matrix A, one can perform elementary row operations on A to convert it into RREF.
Equivalently, there exists elementary roe matrices $E_1, \ldots, E_N$ such that

$$E_N \cdots E_1 A$$

is in RREF.
Furthermore, the RREF is unique.

Note: The same matrix can be converted to many distinct REFs.

Theorem 3

Given any $m \times n$ matrix A, one can perform elementary row operations on A to convert it into RREF.
Equivalently, there exists elementary roe matrices $E_1, \ldots, E_N$ such that

$$E_N \cdots E_1 A$$

is in RREF.
Furthermore, the RREF is unique.

Note: The same matrix can be converted to many distinct REFs.

It is fairly straightforward to perform EROs to turn A into an REF (and further an RREF).

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs.

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs. Note that this does not change the solution set, so we might as well assume that $A^+ = [A \mid \mathbf{b}']$ is in REF.

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs. Note that this does not change the solution set, so we might as well assume that $A^+ = [A \mid \mathbf{b}']$ is in REF. Note that in doing so, we also have that A is in REF.

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs. Note that this does not change the solution set, so we might as well assume that $A^+ = [A \mid \mathbf{b}']$ is in REF. Note that in doing so, we also have that A is in REF.

If the numbers of zero rows of A and A^+ are the same, then the system is consistent, i.e., has a solution.

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs. Note that this does not change the solution set, so we might as well assume that $A^+ = [A \mid \mathbf{b}']$ is in REF. Note that in doing so, we also have that A is in REF.

If the numbers of zero rows of A and A^+ are the same, then the system is consistent, i.e., has a solution.

The set of solutions can now be found directly by back-substitution.

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs. Note that this does not change the solution set, so we might as well assume that $A^+ = [A \mid \mathbf{b}']$ is in REF. Note that in doing so, we also have that A is in REF.

If the numbers of zero rows of A and A^+ are the same, then the system is consistent, i.e., has a solution.

The set of solutions can now be found directly by back-substitution. You will see that the variables corresponding to columns not having a pivot are "free".

Application of Gauss to solving linear equations

For this application, it suffices to turn matrices into REF.

We take the augmented matrix and convert it into REF using EROs. Note that this does not change the solution set, so we might as well assume that $A^+ = [A \mid \mathbf{b}']$ is in REF. Note that in doing so, we also have that A is in REF.

If the numbers of zero rows of A and A^+ are the same, then the system is consistent, i.e., has a solution.

The set of solutions can now be found directly by back-substitution. You will see that the variables corresponding to columns not having a pivot are "free". (Do an example to see what is happening.)

Theorem 4

Let A be a square matrix.

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$.

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$.
Equivalently, A can be written as a product of elementary row matrices.

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$.
Equivalently, A can be written as a product of elementary row matrices.

Algorithm for finding inverse:

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$.
Equivalently, A can be written as a product of elementary row matrices.

Algorithm for finding inverse:

1. Write the matrices A and $\mathbf{I}$ side-by-side.

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$.
Equivalently, A can be written as a product of elementary row matrices.

Algorithm for finding inverse:

1. Write the matrices A and $\mathbf{I}$ side-by-side.
2. Performs EROs on A to convert A into its RREF. Simultaneously perform those EROs on $\mathbf{I}$ (in the same order).

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$. Equivalently, A can be written as a product of elementary row matrices.

Algorithm for finding inverse:

1. Write the matrices A and $\mathbf{I}$ side-by-side.
2. Performs EROs on A to convert A into its RREF. Simultaneously perform those EROs on $\mathbf{I}$ (in the same order).
3. At the end – if A were invertible – the left matrix has become $\mathbf{I}$ and the right matrix is the desired inverse of A.

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$.
Equivalently, A can be written as a product of elementary row matrices.

Algorithm for finding inverse:

1. Write the matrices A and $\mathbf{I}$ side-by-side.
2. Performs EROs on A to convert A into its RREF. Simultaneously perform those EROs on $\mathbf{I}$ (in the same order).
3. At the end – if A were invertible – the left matrix has become $\mathbf{I}$ and the right matrix is the desired inverse of A.

Application of Gauss to finding inverses

Theorem 4

Let A be a square matrix. A is invertible iff the RREF of A is $\mathbf{I}$. Equivalently, A can be written as a product of elementary row matrices.

Algorithm for finding inverse:

1. Write the matrices A and $\mathbf{I}$ side-by-side.
2. Performs EROs on A to convert A into its RREF. Simultaneously perform those EROs on $\mathbf{I}$ (in the same order).
3. At the end – if A were invertible – the left matrix has become $\mathbf{I}$ and the right matrix is the desired inverse of A.

Note: The above algorithm does not require prior knowledge that A is invertible. If you perform it on a non-invertible matrix, you'll end up finding out that it is not invertible.

Table of Contents

Vector subspaces

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,

Vector subspaces

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,

Vector subspaces

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,
3. $\mathbf{v}, \mathbf{w} \in V$ implies $\mathbf{v} + \mathbf{w} \in V$.

Vector subspaces

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,
3. $\mathbf{v}, \mathbf{w} \in V$ implies $\mathbf{v} + \mathbf{w} \in V$.

Vector subspaces

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,
3. $\mathbf{v}, \mathbf{w} \in V$ implies $\mathbf{v} + \mathbf{w} \in V$.

One can replace $\mathbb{R}$ above with $\mathbb{C}$ everywhere to get the notion of a complex vector subspace.

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,
3. $\mathbf{v}, \mathbf{w} \in V$ implies $\mathbf{v} + \mathbf{w} \in V$.

One can replace $\mathbb{R}$ above with $\mathbb{C}$ everywhere to get the notion of a complex vector subspace. To consider both at once, we shall use the symbol $\mathbb{K}$ – which could stand for either $\mathbb{R}$ or $\mathbb{C}$.

From this point on, V will always denote a vector subspace of $\mathbb{K}^n$ (for some n).

Vector subspaces

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,
3. $\mathbf{v}, \mathbf{w} \in V$ implies $\mathbf{v} + \mathbf{w} \in V$.

One can replace $\mathbb{R}$ above with $\mathbb{C}$ everywhere to get the notion of a complex vector subspace. To consider both at once, we shall use the symbol $\mathbb{K}$ – which could stand for either $\mathbb{R}$ or $\mathbb{C}$.

From this point on, V will always denote a vector subspace of $\mathbb{K}^n$ (for some n).

Note that if $a_1, \ldots, a_k \in \mathbb{K}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, then $a_1\mathbf{v}_1 + \cdots a_k\mathbf{v}_k$ is also an element of V.

Definition 5

A subset $V \subseteq \mathbb{R}^n$ is called a (real) vector subspace if

1. $\mathbf{0} \in V$,
2. $a \in \mathbb{R}$ and $\mathbf{v} \in V$ implies $a\mathbf{v} \in V$,
3. $\mathbf{v}$, $\mathbf{w} \in V$ implies $\mathbf{v} + \mathbf{w} \in V$.

One can replace $\mathbb{R}$ above with $\mathbb{C}$ everywhere to get the notion of a complex vector subspace. To consider both at once, we shall use the symbol $\mathbb{K}$ – which could stand for either $\mathbb{R}$ or $\mathbb{C}$.

From this point on, V will always denote a vector subspace of $\mathbb{K}^n$ (for some n).

Note that if $a_1, \ldots, a_k \in \mathbb{K}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_k \in V$, then $a_1\mathbf{v}_1 + \cdots a_k\mathbf{v}_k$ is also an element of V. This element is called a linear combination of the $\mathbf{v}_i$.

Linear span

We can create vector subspaces out of linear combinations.

Linear span

We can create vector subspaces out of linear combinations.

Suppose that $\mathbf{w}_1, \ldots, \mathbf{w}_k \in \mathbb{K}^n$ are arbitrary elements.

Linear span

We can create vector subspaces out of linear combinations.

Suppose that $\mathbf{w}_1, \ldots, \mathbf{w}_k \in \mathbb{K}^n$ are arbitrary elements. Then, the set of all linear combinations

$$V := \{a_1\mathbf{w}_1 + \cdots a_k\mathbf{w}_k : a_1, \ldots, a_k \in \mathbb{K}\}$$

is a vector subspace.

Linear span

We can create vector subspaces out of linear combinations.

Suppose that $\mathbf{w}_1, \ldots, \mathbf{w}_k \in \mathbb{K}^n$ are arbitrary elements. Then, the set of all linear combinations

$$V := \{a_1\mathbf{w}_1 + \cdots a_k\mathbf{w}_k : a_1, \ldots, a_k \in \mathbb{K}\}$$

is a vector subspace.

We write $V = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$ and say that V is spanned (or generated) by $\mathbf{w}_1, \ldots, \mathbf{w}_k$.

Linear independence

Definition 6

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{K}^n$ are said to be linearly dependent

Linear independence

Definition 6

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{K}^n$ are said to be linearly dependent if there exist scalars $a_1, \ldots, a_k \in \mathbb{K}$ not all zero such that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}.$$

Linear independence

Definition 6

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{K}^n$ are said to be linearly dependent if there exist scalars $a_1, \ldots, a_k \in \mathbb{K}$ not all zero such that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}.$$

Else, they are said to be linearly independent.

Linear independence

Definition 6

Vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{K}^n$ are said to be linearly dependent if there exist scalars $a_1, \ldots, a_k \in \mathbb{K}$ not all zero such that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0}.$$

Else, they are said to be linearly independent.

Rephrasing slightly, linear independence means that

$$a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k = \mathbf{0} \Rightarrow a_1 = \cdots = a_k = 0.$$

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace.

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \text{span}(B)$.

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \text{span}(B)$.

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \operatorname{span}(B)$.

Note: One can define linear independence and span for an infinite subset also. Then the "(finite)" above can be dropped.

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \text{span}(B)$.

Note: One can define linear independence and span for an infinite subset also. Then the "(finite)" above can be dropped.

Theorem 8

Let V be any subspace of $\mathbb{K}^n$.

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \operatorname{span}(B)$.

Note: One can define linear independence and span for an infinite subset also. Then the "(finite)" above can be dropped.

Theorem 8

Let V be any subspace of $\mathbb{K}^n$.
Then, V has a basis B.

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \text{span}(B)$.

Note: One can define linear independence and span for an infinite subset also. Then the "(finite)" above can be dropped.

Theorem 8

Let V be any subspace of $\mathbb{K}^n$.
Then, V has a basis B. If B' is any other basis of V, then B and B' have the same size.

Basis

Definition 7

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A (finite) subset $B \subseteq V$ is said to be a basis for V if:

1. B is linearly independent,
2. $V = \operatorname{span}(B)$.

Note: One can define linear independence and span for an infinite subset also. Then the "(finite)" above can be dropped.

Theorem 8

Let V be any subspace of $\mathbb{K}^n$.
Then, V has a basis B. If B' is any other basis of V, then B and B' have the same size.
The size of B is called the dimension of V, denoted $\dim(V)$.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.
   If $|S| = d$, then S is a basis.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.

2. If $V = \text{span}(S)$, then $|S| \geqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can throw out vectors of S to make a basis of V.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can throw out vectors of S to make a basis of V.

The above shows that if two of the following three properties are satisfied by S, then so is the third property (and hence, S is a basis):

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can throw out vectors of S to make a basis of V.

The above shows that if two of the following three properties are satisfied by S, then so is the third property (and hence, S is a basis):

1. S is linearly independent.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can throw out vectors of S to make a basis of V.

The above shows that if two of the following three properties are satisfied by S, then so is the third property (and hence, S is a basis):

1. S is linearly independent.
2. S is of size d.

Linear independence, spanning, basis

Linear algebra is nice, it works like you would intuitively want it to. Let V be a vector subspace of dimension d, and let $S \subseteq V$. We have the following.

1. If S is linearly independent, then $|S| \leqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can extend S to a basis of V.
2. If $V = \text{span}(S)$, then $|S| \geqslant d$.
   If $|S| = d$, then S is a basis.
   Else, you can throw out vectors of S to make a basis of V.

The above shows that if two of the following three properties are satisfied by S, then so is the third property (and hence, S is a basis):

1. S is linearly independent.
2. S is of size d.
3. S spans V.

Linear independence, spanning, basis

Here is another observation from the previous slide:

Here is another observation from the previous slide: If V is spanned by k vectors, then $\dim(V) \leqslant k$.

Linear independence, spanning, basis

Here is another observation from the previous slide: If V is spanned by k vectors, then $\dim(V) \leqslant k$. This means that any $k+1$ vectors in V are linearly dependent.

Here is another observation from the previous slide: If V is spanned by k vectors, then $\dim(V) \leqslant k$. This means that any $k+1$ vectors in V are linearly dependent.

The above was the *First Fundamental Lemma in linear algebra*.

Table of Contents

Column and row ranks

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

Column and row ranks

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

The columns of A can be interpreted as elements of $\mathbb{K}^m$,

Column and row ranks

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

The columns of A can be interpreted as elements of $\mathbb{K}^m$, their linear span is called the column space of A.

Column and row ranks

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

The columns of A can be interpreted as elements of $\mathbb{K}^m$, their linear span is called the column space of A.
Similarly, we have the obvious row space of A – this is a subspace of $\mathbb{K}^n$.

Column and row ranks

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

The columns of A can be interpreted as elements of $\mathbb{K}^m$, their linear span is called the column space of A.
Similarly, we have the obvious row space of A – this is a subspace of $\mathbb{K}^n$.

The dimension of the column space of A is the column rank of A, and the row rank of A is the...

Column and row ranks

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

The columns of A can be interpreted as elements of $\mathbb{K}^m$, their linear span is called the column space of A.
Similarly, we have the obvious row space of A – this is a subspace of $\mathbb{K}^n$.

The dimension of the column space of A is the column rank of A, and the row rank of A is the...

Theorem 9

Row rank $=$ Column rank $=$ number of pivots in any REF.

The common quantity above is called the rank of A.

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns.

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns.

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns. For example, if columns 1, 3, 4 were linearly (in)dependent, then they continue to be so.

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns. For example, if columns 1, 3, 4 were linearly (in)dependent, then they continue to be so.

Note that changing "row" and "column" above is disastrous!

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns. For example, if columns 1, 3, 4 were linearly (in)dependent, then they continue to be so.

Note that changing "row" and "column" above is disastrous! If you are asked to find a basis for column space of A,

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns. For example, if columns 1, 3, 4 were linearly (in)dependent, then they continue to be so.

Note that changing "row" and "column" above is disastrous! If you are asked to find a basis for column space of A, you should convert A to REF, find the pivotal columns there,

EROs, spaces, and ranks

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns. For example, if columns 1, 3, 4 were linearly (in)dependent, then they continue to be so.

Note that changing "row" and "column" above is disastrous! If you are asked to find a basis for column space of A, you should convert A to REF, find the pivotal columns there, and then pick the columns **from the original matrix**.

Theorem 10

Let A be an $m \times n$ matrix. Performing EROs on A does not change the following:

1. Row space.
2. Linear independence of columns. For example, if columns 1, 3, 4 were linearly (in)dependent, then they continue to be so.

Note that changing "row" and "column" above is disastrous! If you are asked to find a basis for column space of A, you should convert A to REF, find the pivotal columns there, and then pick the columns **from the original matrix**.

On the other hand, if you are asked to find a basis for the row space, then you pick the nonzero rows from any REF of A.

Rank-nullity

Given an $m \times n$ matrix, we have the following subspace of $\mathbb{K}^n$, called the null space of A:

Rank-nullity

Given an $m \times n$ matrix, we have the following subspace of $\mathbb{K}^n$, called the null space of A:

$$\mathcal{N}(A) := \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}.$$

Rank-nullity

Given an $m \times n$ matrix, we have the following subspace of $\mathbb{K}^n$, called the null space of A:

$$\mathcal{N}(A) := \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}.$$

The nullity of A is defined as $\dim(\mathcal{N}(A))$.

Rank-nullity

Given an $m \times n$ matrix, we have the following subspace of $\mathbb{K}^n$, called the null space of A:

$$\mathcal{N}(A) := \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}.$$

The nullity of A is defined as $\dim(\mathcal{N}(A))$.

Theorem 11 (Rank-nullity theorem)

$\operatorname{rank}(A) + \operatorname{nullity}(A) = n$.

Rank-nullity

Given an $m \times n$ matrix, we have the following subspace of $\mathbb{K}^n$, called the null space of A:

$$\mathcal{N}(A) := \{\mathbf{x} \in \mathbb{K}^n : A\mathbf{x} = \mathbf{0}\}.$$

The nullity of A is defined as $\dim(\mathcal{N}(A))$.

Theorem 11 (Rank-nullity theorem)

$\operatorname{rank}(A) + \operatorname{nullity}(A) = n$.

Note that n is the number of *columns*.

A remark on null space

Let $A \in \mathbb{K}^{m \times n}$ and $\mathbf{b} \in \mathbb{K}^m$. Consider the system

$$A\mathbf{x} = \mathbf{b}. \tag{$*$}$$

A remark on null space

Let $A \in \mathbb{K}^{m \times n}$ and $\mathbf{b} \in \mathbb{K}^m$. Consider the system

$$A\mathbf{x} = \mathbf{b}. \tag{$*$}$$

Suppose that $(*)$ has a particular solution $\mathbf{x}_0$.

A remark on null space

Let $A \in \mathbb{K}^{m \times n}$ and $\mathbf{b} \in \mathbb{K}^m$. Consider the system

$$A\mathbf{x} = \mathbf{b}. \tag{$*$}$$

Suppose that $(*)$ has a particular solution $\mathbf{x}_0$. *Then*, the complete set of solutions of $(*)$ is $\mathbf{x}_0 + \mathcal{N}(A)$.

A remark on null space

Let $A \in \mathbb{K}^{m \times n}$ and $\mathbf{b} \in \mathbb{K}^m$. Consider the system

$$A\mathbf{x} = \mathbf{b}. \tag{$*$}$$

Suppose that $(*)$ has a particular solution $\mathbf{x}_0$. *Then*, the complete set of solutions of $(*)$ is $\mathbf{x}_0 + \mathcal{N}(A)$.

Moreover, note that $(*)$ has a solution iff $\mathbf{b}$ is in the column space of A.

A remark on null space

Let $A \in \mathbb{K}^{m \times n}$ and $\mathbf{b} \in \mathbb{K}^m$. Consider the system

$$A\mathbf{x} = \mathbf{b}. \tag{$*$}$$

Suppose that $(*)$ has a particular solution $\mathbf{x}_0$. *Then*, the complete set of solutions of $(*)$ is $\mathbf{x}_0 + \mathcal{N}(A)$.

Moreover, note that $(*)$ has a solution iff $\mathbf{b}$ is in the column space of A.

True/False: Suppose $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions, and fix $\mathbf{b} \in \mathbb{K}^m$. Then, $A\mathbf{x} = \mathbf{b}$ also has infinitely many solutions.

A remark on null space

Let $A \in \mathbb{K}^{m \times n}$ and $\mathbf{b} \in \mathbb{K}^m$. Consider the system

$$A\mathbf{x} = \mathbf{b}. \tag{$*$}$$

Suppose that $(*)$ has a particular solution $\mathbf{x}_0$. *Then*, the complete set of solutions of $(*)$ is $\mathbf{x}_0 + \mathcal{N}(A)$.

Moreover, note that $(*)$ has a solution iff $\mathbf{b}$ is in the column space of A.

True/False: Suppose $A\mathbf{x} = \mathbf{0}$ has infinitely many solutions, and fix $\mathbf{b} \in \mathbb{K}^m$. Then, $A\mathbf{x} = \mathbf{b}$ also has infinitely many solutions.

What if "infinitely many" is replaced with "unique"?

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables. Let us say that $x_{i_1}, \ldots, x_{i_k}$ are the free variables.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables. Let us say that $x_{i_1}, \ldots, x_{i_k}$ are the free variables.

We get our first basis vector by putting $x_{i_1} = 1$ and the other free variables as 0.

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables. Let us say that $x_{i_1}, \ldots, x_{i_k}$ are the free variables.

We get our first basis vector by putting $x_{i_1} = 1$ and the other free variables as 0. We can now solve to get the explicit values of the pivotal variables as well.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables. Let us say that $x_{i_1}, \ldots, x_{i_k}$ are the free variables.

We get our first basis vector by putting $x_{i_1} = 1$ and the other free variables as 0. We can now solve to get the explicit values of the pivotal variables as well. This is our first basis vector.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables. Let us say that $x_{i_1}, \ldots, x_{i_k}$ are the free variables.

We get our first basis vector by putting $x_{i_1} = 1$ and the other free variables as 0. We can now solve to get the explicit values of the pivotal variables as well. This is our first basis vector.

After this, we put $x_{i_2} = 1$ and other free variables as 0.

Finding a basis for the null space

Let us discuss how to find a basis for the null space of an $m \times n$ matrix A. Since performing EROs does not affect the solution space, we may assume that we have converted A to REF. As discussed, we can back-substitute and get the pivotal variables in terms of the free variables. Let us say that $x_{i_1}, \ldots, x_{i_k}$ are the free variables.

We get our first basis vector by putting $x_{i_1} = 1$ and the other free variables as 0. We can now solve to get the explicit values of the pivotal variables as well. This is our first basis vector.

After this, we put $x_{i_2} = 1$ and other free variables as 0. Continue in the obvious manner.

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space.

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are x_2 and x_4.

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are x_2 and x_4.

First solution: $x_2 = 1$ and $x_4 = 0$:

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are x_2 and x_4.

First solution: $x_2 = 1$ and $x_4 = 0$: We get $x_3 = 0$ and $x_1 = -3$.

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are x_2 and x_4.

First solution: $x_2 = 1$ and $x_4 = 0$: We get $x_3 = 0$ and $x_1 = -3$.

Second solution: $x_2 = 0$ and $x_4 = 1$:

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are x_2 and x_4.

First solution: $x_2 = 1$ and $x_4 = 0$: We get $x_3 = 0$ and $x_1 = -3$.

Second solution: $x_2 = 0$ and $x_4 = 1$: We get $x_3 = -2$ and $x_1 = 1$.

An example

Suppose that we have

$$A = \begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix}.$$

Note that we wish to solve $A\mathbf{x} = \mathbf{0}$ to get the null space. Thus, the equations obtained are

$$2x_3 + 4x_4 = 0 \quad \text{and} \quad x_1 + 3x_2 + x_3 + x_4 = 0.$$

The free variables are x_2 and x_4.

First solution: $x_2 = 1$ and $x_4 = 0$: We get $x_3 = 0$ and $x_1 = -3$.

Second solution: $x_2 = 0$ and $x_4 = 1$: We get $x_3 = -2$ and $x_1 = 1$.

Thus, we get a basis as $\left\{\begin{bmatrix}-3 & 1 & 0 & 0\end{bmatrix}^T, \begin{bmatrix}1 & 0 & -2 & 1\end{bmatrix}^T\right\}$.

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Then, using back-substitution, it is easy to find one particular solution

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Then, using back-substitution, it is easy to find one particular solution (you can put both free variables as 0 to make life easy).

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Then, using back-substitution, it is easy to find one particular solution (you can put both free variables as 0 to make life easy). This gives us a particular solution as

$$\mathbf{x}_0 = \begin{bmatrix} -1 & 0 & 2 & 0 \end{bmatrix}^{\mathsf{T}}.$$

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Then, using back-substitution, it is easy to find one particular solution (you can put both free variables as 0 to make life easy). This gives us a particular solution as

$$\mathbf{x}_0 = \begin{bmatrix} -1 & 0 & 2 & 0 \end{bmatrix}^{\mathsf{T}}.$$

The general solution now is $\mathbf{x}_0 + \mathcal{N}(A)$.

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Then, using back-substitution, it is easy to find one particular solution (you can put both free variables as 0 to make life easy). This gives us a particular solution as

$$\mathbf{x}_0 = \begin{bmatrix} -1 & 0 & 2 & 0 \end{bmatrix}^T.$$

The general solution now is $\mathbf{x}_0 + \mathcal{N}(A)$. We had already found a basis earlier.

An example (continued)

Suppose that we were asked to find the general solution set of

$$\begin{bmatrix} \boxed{1} & 3 & 1 & 1 \\ 0 & 0 & \boxed{2} & 4 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 1 \\ 4 \end{bmatrix}.$$

Then, using back-substitution, it is easy to find one particular solution (you can put both free variables as 0 to make life easy). This gives us a particular solution as

$$\mathbf{x}_0 = \begin{bmatrix} -1 & 0 & 2 & 0 \end{bmatrix}^T.$$

The general solution now is $\mathbf{x}_0 + \mathcal{N}(A)$. We had already found a basis earlier. Thus, the complete set of solutions is given by

$$\left\{ \begin{bmatrix} -1 \\ 0 \\ 2 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \end{bmatrix} : x_2, x_4 \in \mathbb{R} \right\}.$$

Determinants

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$.

Determinants

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^\mathsf{T})$.

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^{\mathsf{T}})$.

Theorem 12 (Determinantal rank)

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

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^{\mathsf{T}})$.

Theorem 12 (Determinantal rank)

Let A be an $m \times n$ matrix. Let k be such that A has a $k \times k$ submatrix with nonzero determinant

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^{\mathsf{T}})$.

Theorem 12 (Determinantal rank)

Let A be an $m \times n$ matrix. Let k be such that A has a $k \times k$ submatrix with nonzero determinant but every $(k+1) \times (k+1)$ submatrix has zero determinant.

Determinants

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^{\mathsf{T}})$.

Theorem 12 (Determinantal rank)

Let A be an $m \times n$ matrix. Let k be such that A has a $k \times k$ submatrix with nonzero determinant but every $(k+1) \times (k+1)$ submatrix has zero determinant.
Then, $k = \operatorname{rank}(A)$.

Determinants

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^{\mathsf{T}})$.

Theorem 12 (Determinantal rank)

Let A be an $m \times n$ matrix. Let k be such that A has a $k \times k$ submatrix with nonzero determinant but every $(k+1) \times (k+1)$ submatrix has zero determinant.
Then, $k = \operatorname{rank}(A)$.

In other words, the rank is the size of the largest square submatrix with nonzero determinant.

Determinants

We know what the determinant of a matrix is. Furthermore, we know its basic properties: it is multilinear, alternating, and $\det(\mathbf{I}) = 1$. Moreover, $\det(AB) = \det(A)\det(B)$ and $\det(A) = \det(A^\mathsf{T})$.

Theorem 12 (Determinantal rank)

Let A be an $m \times n$ matrix. Let k be such that A has a $k \times k$ submatrix with nonzero determinant but every $(k+1) \times (k+1)$ submatrix has zero determinant.
Then, $k = \operatorname{rank}(A)$.

In other words, the rank is the size of the largest square submatrix with nonzero determinant. Note that there may still be *some* $k \times k$ submatrices with zero determinant.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\text{rank}(A) = n$.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\text{rank}(A) = n$.
3. $\text{nullity}(A) = 0$.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\operatorname{rank}(A) = n$.
3. $\operatorname{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\text{rank}(A) = n$.
3. $\text{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.
5. $\det(A) \neq 0$.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\text{rank}(A) = n$.
3. $\text{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.
5. $\det(A) \neq 0$.
6. Columns (or rows) of A are linearly independent.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\text{rank}(A) = n$.
3. $\text{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.
5. $\det(A) \neq 0$.
6. Columns (or rows) of A are linearly independent.
7. RREF of A is $\mathbf{I}$.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\operatorname{rank}(A) = n$.
3. $\operatorname{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.
5. $\det(A) \neq 0$.
6. Columns (or rows) of A are linearly independent.
7. RREF of A is $\mathbf{I}$.
8. A is a product of ERMs.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\text{rank}(A) = n$.
3. $\text{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.
5. $\det(A) \neq 0$.
6. Columns (or rows) of A are linearly independent.
7. RREF of A is $\mathbf{I}$.
8. A is a product of ERMs.

Characterisations of invertibility

Let A be an $n \times n$ square matrix. TFAE:

1. A is invertible.
2. $\operatorname{rank}(A) = n$.
3. $\operatorname{nullity}(A) = 0$.
4. $A\mathbf{x} = \mathbf{0}$ has a unique solution, namely $\mathbf{x} = 0$.
5. $\det(A) \neq 0$.
6. Columns (or rows) of A are linearly independent.
7. RREF of A is $\mathbf{I}$.
8. A is a product of ERMs.

True/False: Let A be a square matrix such that $A\mathbf{x} = \mathbf{0}$ has a unique solution, and fix $\mathbf{b} \in \mathbb{K}^n$. Does $A\mathbf{x} = \mathbf{b}$ also have a unique solution?

Applications of determinants

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors.

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors. Construct the matrix G whose (i,j)-th entry is the dot product $\mathbf{v}_i^T \mathbf{v}_j$.

Applications of determinants

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors. Construct the matrix G whose (i, j)-th entry is the dot product $\mathbf{v}_i^T \mathbf{v}_j$. Then, the k vectors are linearly independent iff $\det(G) \neq 0$.

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors. Construct the matrix G whose (i, j)-th entry is the dot product $\mathbf{v}_i^\top \mathbf{v}_j$. Then, the k vectors are linearly independent iff $\det(G) \neq 0$.

Recall that the adjugate of a square matrix A is the transpose of the cofactor matrix, and is denoted by $\operatorname{adj}(A)$.

Applications of determinants

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors. Construct the matrix G whose (i, j)-th entry is the dot product $\mathbf{v}_i^\top \mathbf{v}_j$. Then, the k vectors are linearly independent iff $\det(G) \neq 0$.

Recall that the adjugate of a square matrix A is the transpose of the cofactor matrix, and is denoted by $\mathrm{adj}(A)$. We have $\mathrm{adj}(A)A = \det(A)\mathbf{I}$.

Applications of determinants

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors. Construct the matrix G whose (i, j)-th entry is the dot product $\mathbf{v}_i^\top \mathbf{v}_j$. Then, the k vectors are linearly independent iff $\det(G) \neq 0$.

Recall that the adjugate of a square matrix A is the transpose of the cofactor matrix, and is denoted by $\operatorname{adj}(A)$. We have $\operatorname{adj}(A)A = \det(A)\mathbf{I}$. If A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

Applications of determinants

Theorem 13

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ be vectors. Construct the matrix G whose (i, j)-th entry is the dot product $\mathbf{v}_i^\top \mathbf{v}_j$. Then, the k vectors are linearly independent iff $\det(G) \neq 0$.

Recall that the adjugate of a square matrix A is the transpose of the cofactor matrix, and is denoted by $\operatorname{adj}(A)$. We have $\operatorname{adj}(A)A = \det(A)\mathbf{I}$. If A is invertible, then $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$. In this case, the unique solution of $A\mathbf{x} = \mathbf{b}$ is given by

$$\mathbf{x} = \frac{(\operatorname{adj}(A))\mathbf{b}}{\det(A)}.$$

The above is essentially Cramer's rule.

Table of Contents

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace.

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,
5. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$,

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,
5. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$,
6. $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,
5. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$,
6. $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,
5. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$,
6. $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

The norm of $\mathbf{v}$ is defined as $\|v\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,
5. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$,
6. $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

The norm of $\mathbf{v}$ is defined as $\|v\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Theorem 15 (Parallelogram law)

$\|\mathbf{v}_1 + \mathbf{v}_2\|^2 + \|\mathbf{v}_1 - \mathbf{v}_2\|^2 = 2(\|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2)$.

Inner products

Definition 14

Let $V \subseteq \mathbb{K}^n$ be a vector subspace. A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}$ is called an inner product if

1. $\langle \mathbf{v}, \mathbf{u} + \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$,
2. $\langle \mathbf{v} + \mathbf{w}, \mathbf{u} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{w}, \mathbf{u} \rangle$,
3. $\langle \alpha \mathbf{v}, \mathbf{w} \rangle = \alpha \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \overline{\alpha} \mathbf{w} \rangle$,
4. $\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$,
5. $\langle \mathbf{v}, \mathbf{v} \rangle \geqslant 0$,
6. $\langle \mathbf{v}, \mathbf{v} \rangle = 0 \Rightarrow \mathbf{v} = \mathbf{0}$.

The norm of $\mathbf{v}$ is defined as $\|v\| := \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle}$.

Theorem 15 (Parallelogram law)

$\|\mathbf{v}_1 + \mathbf{v}_2\|^2 + \|\mathbf{v}_1 - \mathbf{v}_2\|^2 = 2(\|\mathbf{v}_1\|^2 + \|\mathbf{v}_2\|^2)$.

Similarly, one has Pythagoras theorem and Cauchy Schwarz:
$|\langle \mathbf{v}_1, \mathbf{v}_2 \rangle| \leqslant \|\mathbf{v}_1\| \cdot \|\mathbf{v}_2\|$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V,

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1/\|\mathbf{v}_1\|$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1 / \|\mathbf{v}_1\|$.
2. Then, define $\mathbf{x}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1 / \|\mathbf{v}_1\|$.
2. Then, define $\mathbf{x}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   In other words, subtract the component of $\mathbf{v}_2$ in the direction of $\mathbf{w}_1$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1/\|\mathbf{v}_1\|$.
2. Then, define $\mathbf{x}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   In other words, subtract the component of $\mathbf{v}_2$ in the direction of $\mathbf{w}_1$.
   Now, set $\mathbf{w}_2 := \mathbf{x}_2/\|\mathbf{x}_2\|$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1/\|\mathbf{v}_1\|$.
2. Then, define $\mathbf{x}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   In other words, subtract the component of $\mathbf{v}_2$ in the direction of $\mathbf{w}_1$.
   Now, set $\mathbf{w}_2 := \mathbf{x}_2/\|\mathbf{x}_2\|$.
3. Next, define $\mathbf{x_3} := \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 - \langle \mathbf{v}_3, \mathbf{w}_1 \rangle \mathbf{w}_1$.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1/\|\mathbf{v}_1\|$.
2. Then, define $\mathbf{x}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   In other words, subtract the component of $\mathbf{v}_2$ in the direction of $\mathbf{w}_1$.
   Now, set $\mathbf{w}_2 := \mathbf{x}_2/\|\mathbf{x}_2\|$.
3. Next, define $\mathbf{x_3} := \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 - \langle \mathbf{v}_3, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   Thus, we are subtracting the component in the directions of $\mathbf{w}_1$ and $\mathbf{w}_2$ both.

Gram Schmidt

Given a finite set $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ in an inner product space V, we wish to find an orthogonal subset $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ such that

$$\operatorname{span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \operatorname{span}\{\mathbf{w}_1, \ldots, \mathbf{w}_k\}$$

for all $1 \leqslant k \leqslant n$.

The idea is to do the following:

1. First define $\mathbf{w}_1 := \mathbf{v}_1/\|\mathbf{v}_1\|$.
2. Then, define $\mathbf{x}_2 := \mathbf{v}_2 - \langle \mathbf{v}_2, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   In other words, subtract the component of $\mathbf{v}_2$ in the direction of $\mathbf{w}_1$.
   Now, set $\mathbf{w}_2 := \mathbf{x}_2/\|\mathbf{x}_2\|$.
3. Next, define $\mathbf{x_3} := \mathbf{v}_3 - \langle \mathbf{v}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 - \langle \mathbf{v}_3, \mathbf{w}_1 \rangle \mathbf{w}_1$.
   Thus, we are subtracting the component in the directions of $\mathbf{w}_1$ and $\mathbf{w}_2$ both.
   Set $\mathbf{w}_3 := \mathbf{x}_3/\|\mathbf{x}_3\|$.
   Continue similarly.

Gram Schmidt

The earlier process will work fine if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent.

Gram Schmidt

The earlier process will work fine if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent. In this case, the set obtained $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ will be ortho<u>normal</u>.

Gram Schmidt

The earlier process will work fine if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent. In this case, the set obtained $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ will be ortho<u>normal</u>.

Otherwise, we will get that some of the $\mathbf{x}_i$ in the process are 0. In that case, we simply discard those.

Gram Schmidt

The earlier process will work fine if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent. In this case, the set obtained $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ will be orthonormal.

Otherwise, we will get that some of the $\mathbf{x}_i$ in the process are 0. In that case, we simply discard those.

The benefit of the above is that given any basis B of an inner product space V, we can get an orthonormal basis B'.

Gram Schmidt

The earlier process will work fine if $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent. In this case, the set obtained $\{\mathbf{w}_1, \dots, \mathbf{w}_n\}$ will be ortho<u>normal</u>.

Otherwise, we will get that some of the $\mathbf{x}_i$ in the process are 0. In that case, we simply discard those.

The benefit of the above is that given any basis B of an inner product space V, we can get an orthonormal basis B'.

Theorem 16

Any orthogonal set of nonzero vectors is linearly independent.

Gram Schmidt

The earlier process will work fine if $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is linearly independent. In this case, the set obtained $\{\mathbf{w}_1, \ldots, \mathbf{w}_n\}$ will be ortho<u>normal</u>.

Otherwise, we will get that some of the $\mathbf{x}_i$ in the process are 0. In that case, we simply discard those.

The benefit of the above is that given any basis B of an inner product space V, we can get an orthonormal basis B'.

Theorem 16

Any orthogonal set of nonzero vectors is linearly independent.
In particular, any orthonormal set is linearly independent.

Table of Contents

Definition 17

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

Definition 17

Let A be an $n \times n$ matrix. Let $\mathbf{v} \in \mathbb{K}^n$ be a nonzero vector such that

Definitions

Definition 17

Let A be an $n \times n$ matrix. Let $\mathbf{v} \in \mathbb{K}^n$ be a nonzero vector such that

$$A\mathbf{v} = \lambda\mathbf{v}$$

for some $\lambda \in \mathbb{K}$.

Definitions

Definition 17

Let A be an $n \times n$ matrix. Let $\mathbf{v} \in \mathbb{K}^n$ be a nonzero vector such that

$$A\mathbf{v} = \lambda\mathbf{v}$$

for some $\lambda \in \mathbb{K}$.
Then, $\mathbf{v}$ is said to be an eigenvector of A with eigenvalue λ.

Definition 17

Let A be an $n \times n$ matrix. Let $\mathbf{v} \in \mathbb{K}^n$ be a <u>nonzero</u> vector such that

$$A\mathbf{v} = \lambda\mathbf{v}$$

for some $\lambda \in \mathbb{K}$.
Then, $\mathbf{v}$ is said to be an eigenvector of A with eigenvalue λ.

Proposition 18

λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$.

Definitions

Definition 17

Let A be an $n \times n$ matrix. Let $\mathbf{v} \in \mathbb{K}^n$ be a nonzero vector such that

$$A\mathbf{v} = \lambda \mathbf{v}$$

for some $\lambda \in \mathbb{K}$.
Then, $\mathbf{v}$ is said to be an eigenvector of A with eigenvalue λ.

Proposition 18

λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$.

The polynomial $p(x) = \det(A - xI)$ is called the characteristic polynomial of A.

Definitions

Definition 17

Let A be an $n \times n$ matrix. Let $\mathbf{v} \in \mathbb{K}^n$ be a nonzero vector such that

$$A\mathbf{v} = \lambda \mathbf{v}$$

for some $\lambda \in \mathbb{K}$.
Then, $\mathbf{v}$ is said to be an eigenvector of A with eigenvalue λ.

Proposition 18

λ is an eigenvalue of A iff $\det(A - \lambda I) = 0$.

The polynomial $p(x) = \det(A - xI)$ is called the characteristic polynomial of A. The above proposition says that the eigenvalues of A are precisely the roots of the characteristic polynomial.

Multiplicities

Definition 19

Let λ be an eigenvalue of A, and let $p(x)$ be the characteristic polynomial of A.

Multiplicities

Definition 19

Let λ be an eigenvalue of A, and let $p(x)$ be the characteristic polynomial of A.
Then, we can write $p(x) = (x - \lambda)^m q(x)$ for some $m \geqslant 1$ with $q(\lambda) \neq 0$.

Multiplicities

Definition 19

Let λ be an eigenvalue of A, and let $p(x)$ be the characteristic polynomial of A.
Then, we can write $p(x) = (x - \lambda)^m q(x)$ for some $m \geqslant 1$ with $q(\lambda) \neq 0$.
m is called the algebraic multiplicity of λ.

Multiplicities

Definition 19

Let λ be an eigenvalue of A, and let $p(x)$ be the characteristic polynomial of A.
Then, we can write $p(x) = (x - \lambda)^m q(x)$ for some $m \geqslant 1$ with $q(\lambda) \neq 0$.
m is called the algebraic multiplicity of λ.
The geometric multiplicity of λ is defined as $\text{nullity}(A - \lambda I)$.

Theorem 20

Let λ be an eigenvalue of an $n \times n$ matrix A. Let g and m denote the geometric and algebraic multiplicities of λ respectively. Then,

$$1 \leqslant g \leqslant m \leqslant n.$$

Theorem 21

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues.

Diagonalisability

Theorem 21

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are linearly independent.

Theorem 22

Let A be an $n \times n$ matrix over $\mathbb{K}$. TFAE:

Diagonalisability

Theorem 21

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are linearly independent.

Theorem 22

Let A be an $n \times n$ matrix over $\mathbb{K}$. TFAE:

1. There exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ consisting of eigenvectors of A.

Diagonalisability

Theorem 21

Let $\mathbf{v}_1, \dots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are linearly independent.

Theorem 22

Let A be an $n \times n$ matrix over $\mathbb{K}$. TFAE:

1. There exists a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ consisting of eigenvectors of A.
2. There exists an invertible matrix $P \in \mathbb{K}^{n \times n}$ such that $P^{-1}AP$ is a diagonal matrix.

Diagonalisability

Theorem 21

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are linearly independent.

Theorem 22

Let A be an $n \times n$ matrix over $\mathbb{K}$. TFAE:

1. There exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ consisting of eigenvectors of A.
2. There exists an invertible matrix $P \in \mathbb{K}^{n \times n}$ such that $P^{-1}AP$ is a diagonal matrix.

Diagonalisability

Theorem 21

Let $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are linearly independent.

Theorem 22

Let A be an $n \times n$ matrix over $\mathbb{K}$. TFAE:

1. There exists a basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ of $\mathbb{K}^n$ consisting of eigenvectors of A.
2. There exists an invertible matrix $P \in \mathbb{K}^{n \times n}$ such that $P^{-1}AP$ is a diagonal matrix.

If A satisfies either (and hence, both) condition, then A is said to be diagonalisable over $\mathbb{K}$.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Note that if $\mathbb{K} = \mathbb{C}$, then $\sum m_i = n$ is always true

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Note that if $\mathbb{K} = \mathbb{C}$, then $\sum m_i = n$ is always true and the above condition just says that we must have $g_i = m_i$ for all i.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Note that if $\mathbb{K} = \mathbb{C}$, then $\sum m_i = n$ is always true and the above condition just says that we must have $g_i = m_i$ for all i.

If $\mathbb{K} = \mathbb{R}$, then $\sum m_i < n$ is possible

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Note that if $\mathbb{K} = \mathbb{C}$, then $\sum m_i = n$ is always true and the above condition just says that we must have $g_i = m_i$ for all i.

If $\mathbb{K} = \mathbb{R}$, then $\sum m_i < n$ is possible and in that case, A is automatically not diagonalisable <u>over $\mathbb{R}$</u>.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Note that if $\mathbb{K} = \mathbb{C}$, then $\sum m_i = n$ is always true and the above condition just says that we must have $g_i = m_i$ for all i.

If $\mathbb{K} = \mathbb{R}$, then $\sum m_i < n$ is possible and in that case, A is automatically not diagonalisable <u>over $\mathbb{R}$</u>. However, A may still be diagonalisable over $\mathbb{C}$.

Criteria for diagonalisability

Let $A \in \mathbb{K}^{n \times n}$ be a square matrix, and let $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ be the distinct eigenvalues of A.
Let g_i and m_i denote the geometric and algebraic multiplicities of λ_i.

A is diagonalisable <u>over $\mathbb{K}$</u> iff $\boxed{\sum g_i = n}$.

Note that if $\mathbb{K} = \mathbb{C}$, then $\sum m_i = n$ is always true and the above condition just says that we must have $g_i = m_i$ for all i.

If $\mathbb{K} = \mathbb{R}$, then $\sum m_i < n$ is possible and in that case, A is automatically not diagonalisable <u>over $\mathbb{R}$</u>. However, A may still be diagonalisable over $\mathbb{C}$.

Example: $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is diagonalisable over $\mathbb{C}$ but not over $\mathbb{R}$.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

Now, define the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{K}^{n \times n}$.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

Now, define the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{K}^{n \times n}$. Note that P is invertible.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

Now, define the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{K}^{n \times n}$. Note that P is invertible. Moreover,

$$D := P^{-1}AP$$

is a diagonal matrix.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

Now, define the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{K}^{n \times n}$. Note that P is invertible. Moreover,

$$D := P^{-1}AP$$

is a diagonal matrix. The i-th diagonal entry will be the eigenvalue corresponding to $\mathbf{v}_i$.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

Now, define the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{K}^{n \times n}$. Note that P is invertible. Moreover,

$$D := P^{-1}AP$$

is a diagonal matrix. The i-th diagonal entry will be the eigenvalue corresponding to $\mathbf{v}_i$. Each eigenvalue will appear in D according to its multiplicity.

The P and D

Suppose that A is diagonalisable over $\mathbb{K}$ and that we have found out the distinct eigenvalues $\lambda_1, \ldots, \lambda_k$.

For each λ_i, we can find a basis for the null space of $A - \lambda_i \mathbf{I}$ using REF. Find a basis for each $\mathcal{N}(A - \lambda_i I)$ and put it all in a list: $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$. (Why do we get n vectors?)

Now, define the matrix $P = \begin{bmatrix} \mathbf{v}_1 & \cdots & \mathbf{v}_n \end{bmatrix} \in \mathbb{K}^{n \times n}$. Note that P is invertible. Moreover,

$$D := P^{-1}AP$$

is a diagonal matrix. The i-th diagonal entry will be the eigenvalue corresponding to $\mathbf{v}_i$. Each eigenvalue will appear in D according to its multiplicity. (Which multiplicity?)

Table of Contents

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

For a unitary matrix, we also have $A^*A = \mathbf{I}$.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

For a unitary matrix, we also have $A^*A = \mathbf{I}$. Note that all matrices above are normal.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

For a unitary matrix, we also have $A^*A = \mathbf{I}$. Note that all matrices above are normal. Also note that unitary matrices are invertible.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

For a unitary matrix, we also have $A^*A = \mathbf{I}$. Note that all matrices above are normal. Also note that unitary matrices are invertible.

Also recall that he standard inner product on $\mathbb{K}^n$ is given by
$\langle \mathbf{v}, \mathbf{w} \rangle := \mathbf{w}^*\mathbf{v}$.

Adjoint

Given a complex matrix A, we denote its conjugate transpose by A^*. This is called the adjoint of A.

$A \in \mathbb{C}^{n \times n}$ is said to be ____ if ____:

1. normal; $AA^* = A^*A$,
2. Hermitian; $A^* = A$,
3. skew-Hermitian; $A^* = -A$,
4. unitary; $AA^* = \mathbf{I}$,
5. orthogonal; A is unitary and real.

For a unitary matrix, we also have $A^*A = \mathbf{I}$. Note that all matrices above are normal. Also note that unitary matrices are invertible.

Also recall that he standard inner product on $\mathbb{K}^n$ is given by $\langle \mathbf{v}, \mathbf{w} \rangle := \mathbf{w}^*\mathbf{v}$. This is what we shall refer to from now on.

Some remarks

Let $A \in \mathbb{C}^{n \times n}$ be a matrix. TFAE:

Let $A \in \mathbb{C}^{n \times n}$ be a matrix. TFAE:

1. A is unitary.

Some remarks

Let $A \in \mathbb{C}^{n\times n}$ be a matrix. TFAE:

1. A is unitary.
2. The rows of A are orthonormal.

Some remarks

Let $A \in \mathbb{C}^{n \times n}$ be a matrix. TFAE:

1. A is unitary.
2. The rows of A are orthonormal.
3. The columns of A are orthonormal.

Some remarks

Let $A \in \mathbb{C}^{n \times n}$ be a matrix. TFAE:

1. A is unitary.
2. The rows of A are orthonormal.
3. The columns of A are orthonormal.

Some remarks

Let $A \in \mathbb{C}^{n \times n}$ be a matrix. TFAE:

1. A is unitary.
2. The rows of A are orthonormal.
3. The columns of A are orthonormal.

Also note that a diagonal matrix is Hermitian iff it is real.
Similarly, a diagonal matrix is skew-Hermitian iff it is purely imaginary.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable,

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable, i.e., there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}AU$ is diagonal.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable, i.e., there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}AU$ is diagonal.
Equivalently, there is an basis of $\mathbb{C}^n$ consisting of orthonormal eigenvectors.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable, i.e., there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}AU$ is diagonal. Equivalently, there is an basis of $\mathbb{C}^n$ consisting of orthonormal eigenvectors.

Note that $U^{-1} = U^*$ above, since P is unitary.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable, i.e., there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}AU$ is diagonal. Equivalently, there is an basis of $\mathbb{C}^n$ consisting of orthonormal eigenvectors.

Note that $U^{-1} = U^*$ above, since P is unitary. Note that even if A is a real normal matrix, we cannot guarantee that U can be chosen to be normal.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable, i.e., there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}AU$ is diagonal. Equivalently, there is an basis of $\mathbb{C}^n$ consisting of orthonormal eigenvectors.

Note that $U^{-1} = U^*$ above, since P is unitary. Note that even if A is a real normal matrix, we cannot guarantee that U can be chosen to be normal. Indeed, if A has a nonreal eigenvalue in $\mathbb{C}$, then P cannot be chosen real.

Spectral theorem

Theorem 23

Let A be normal, and $\mathbf{v}_1, \ldots, \mathbf{v}_k$ be eigenvectors of A corresponding to distinct eigenvalues. Then, they are orthogonal.

Theorem 24 (Spectral Theorem)

Let $A \in \mathbb{C}^{n \times n}$ be a normal matrix. Then, A is unitarily diagonalisable, i.e., there exists a unitary matrix $U \in \mathbb{C}^{n \times n}$ such that $U^{-1}AU$ is diagonal. Equivalently, there is an basis of $\mathbb{C}^n$ consisting of orthonormal eigenvectors.

Note that $U^{-1} = U^*$ above, since P is unitary. Note that even if A is a real normal matrix, we cannot guarantee that U can be chosen to be normal. Indeed, if A has a nonreal eigenvalue in $\mathbb{C}$, then P cannot be chosen real.

The converse of the above is true as well: if A is unitarily diagonalisable (or has an orthonormal eigenbasis), then A is normal.

Nature of eigenvalues

Let A be a normal matrix, and $\lambda \in \mathbb{C}$ be an eigenvalue of A.

Nature of eigenvalues

Let A be a normal matrix, and $\lambda \in \mathbb{C}$ be an eigenvalue of A. We have the following table giving us more information about the nature of A.

Nature of A	Nature of λ
Hermitian	$\lambda \in \mathbb{R}$
Skew-Hermitian	$\lambda \in \iota\mathbb{R}$
Unitary	$\lvert\lambda\rvert = 1$

Nature of eigenvalues

Let A be a normal matrix, and $\lambda \in \mathbb{C}$ be an eigenvalue of A. We have the following table giving us more information about the nature of A.

Nature of A	Nature of λ
Hermitian	$\lambda \in \mathbb{R}$
Skew-Hermitian	$\lambda \in \iota\mathbb{R}$
Unitary	$\lvert\lambda\rvert = 1$

In particular, we have a spectral theorem for real symmetric matrices.

Nature of eigenvalues

Let A be a normal matrix, and $\lambda \in \mathbb{C}$ be an eigenvalue of A. We have the following table giving us more information about the nature of A.

Nature of A	Nature of λ
Hermitian	$\lambda \in \mathbb{R}$
Skew-Hermitian	$\lambda \in \iota\mathbb{R}$
Unitary	$\lvert\lambda\rvert = 1$

In particular, we have a spectral theorem for real symmetric matrices.

Theorem 25

Let $A \in \mathbb{R}^{n\times n}$ be symmetric. Then, there exists an orthogonal matrix $O \in \mathbb{R}^{n\times n}$ such that $O^\top AO$ is diagonal.