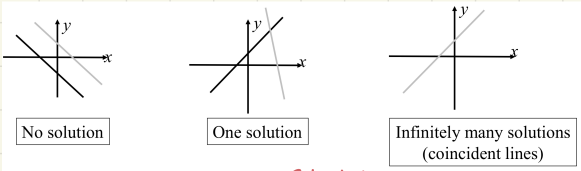
Linear System

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

if any $b \neq 0$ Non-Homogeneous

if all $b = 0$ Homogeneous → all eq^n will intersect through 0, 0, 0 … cfm 1 solution

Two Unknowns



$x - y = 4$, $3x - 3y = 6$ $\begin{bmatrix} 1 & -1 & 4 \\ 0 & 0 & -6 \end{bmatrix}$

$x - y = 1$, $2x + y = 6$ $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 3 & 4 \end{bmatrix}$

$4x - 2y = 1$, $16x - 8y = 4$ $\begin{bmatrix} 4 & -2 & 1 \\ a4 & -a2 & a4 \end{bmatrix}$ $\begin{bmatrix} 1 & -\frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 \end{bmatrix}$

Two Forms

Row Echelon Form

non-unique

$$\begin{bmatrix} 7 & 3 & 1 & 0 & 4 \\ 0 & 1 & 2 & 4 & 3 \\ 0 & 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Reduced Row Echelon Form

unique

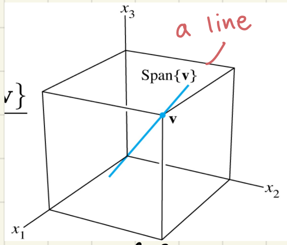
$$\begin{bmatrix} 1 & 3 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Solution? LHS/RHS

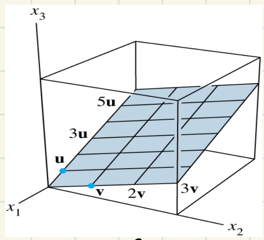
0/x NO solution

↳ Stop

SPAN $\mathbb{R}^x$



Span {v}



Span {u, v}

0/0 ↳ Reduced row echelon form, and assign free varibles to the ones w/o pivot

∞

$$\begin{bmatrix} 1 & 3 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\rightsquigarrow x_1 + 3x_2 = 4$

$x_3 = 2$

$x_4 = 3$

$\therefore x_1 = 4 - 3t$

$x_3 = 2$

$x_4 = 3$

x/0, x/y, can!

Matrix $Ax = b$

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

The equation $A\mathbf{x} = \mathbf{b}$ has a solution if and only if $\mathbf{b}$ is a linear combination of the columns of A.

Exercise:

Let $A = \begin{bmatrix} 1 & 3 & 4 \\ -4 & -2 & -6 \\ -3 & -2 & -7 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$. Is the equation $A\mathbf{x} = \mathbf{b}$ consistent for all possible values of b_1, b_2, b_3?

For $A\mathbf{x} = \mathbf{b}$ to be consistent, $b_1 - \frac{1}{2}b_2 + b_3 = 0$.

The columns of $A = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3]$ span a plane through $\mathbf{0}$ in $\mathbb{R}^3$. ■

$$\begin{bmatrix} 1 & 3 & 4 & b_1 \\ -4 & -2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{bmatrix} \to \begin{bmatrix} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & 4b_1 - b_2 \\ 0 & 0 & 0 & -2b_1 + b_2 - 2b_3 \end{bmatrix}$$

$0x_1 + 0x_2 + 0x_3 = -2b_1 + b_2 - 2b_3$

$-2b_1 + b_2 - 2b_3 = 0$ → This condition is required

Theorem 1.2. *Let A be an $m \times n$ matrix. Then the following statements are logically equivalent, i.e., for a particular A, either they are all true statements or they are all false.*

a. *For each $\mathbf{b}$ in $\mathbb{R}^m$, the equation $A\mathbf{x} = \mathbf{b}$ has a solution.*

b. *Each $\mathbf{b}$ in $\mathbb{R}^m$ is a linear combination of the columns of A.*

c. *The columns of A span $\mathbb{R}^m$.*

d. *A has a pivot position in every row.*

Conversly

$Ax \neq b$

No linear combi

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

Examples: Homo

Determine if the following homogeneous system has a nontrivial solution.

$$\begin{aligned} 3x_1 + 5x_2 - 4x_3 &= 0 \\ -3x_1 - 2x_2 + 4x_3 &= 0 \\ 6x_1 + x_2 - 8x_3 &= 0 \end{aligned}$$

$$\begin{bmatrix} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

x_3 is a free varible

$x_1 - \frac{4}{3}x_3 = 0$

$x_2 = 0$

$\therefore$ The solution $\vec{x} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix}$

$= x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$

x_3 just makes the line longer, increasing "span"

Non-Homo

Describe all solutions of $A\mathbf{x} = \mathbf{b}$ where $A = \begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix}$, and $\mathbf{b} = \begin{bmatrix} 7 \\ -1 \\ -4 \end{bmatrix}$

$$\begin{bmatrix} 3 & 5 & -4 & 7 \\ -3 & -2 & 4 & -1 \\ 6 & 1 & -8 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$x_1 - \frac{4}{3}x_3 = -1$

$x_2 = 2$

x_3 is a free varible

$\vec{x} = \begin{bmatrix} \frac{4}{3}x_3 - 1 \\ 2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \frac{4}{3}x_3 \\ 0 \\ x_3 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$

$= x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$

Linear In/dependence

Definition. An indexed set of vectors $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ in $\mathbb{R}^n$ is said to be **linearly independent** if the vector equation

$x_1 = x_2 = x_3 = x_4 = 0$

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \dots + x_p\mathbf{v}_p = \mathbf{0}$$

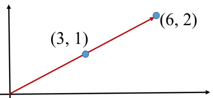
has only the trivial solution. The set $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \dots, c_p$, not all zero, such that

$3v_1 + 2v_2 = 0$

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \dots + c_p\mathbf{v}_p = \mathbf{0}$$

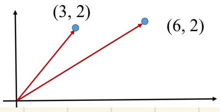
By inspection for simple cases

$\mathbf{v}_1 = \begin{bmatrix}3\\1\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}6\\2\end{bmatrix}$



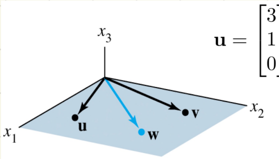
Linearly Dependent

$\mathbf{v}_1 = \begin{bmatrix}3\\2\end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix}6\\2\end{bmatrix}$



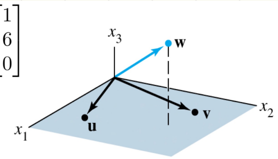
Linearly Independent

$\mathbf{u} = \begin{bmatrix}3\\1\\0\end{bmatrix}$ $\quad \mathbf{v} = \begin{bmatrix}1\\6\\0\end{bmatrix}$



Linearly dependent, $\mathbf{w}$ in Span$\{\mathbf{u}, \mathbf{v}\}$

Observe !



Linearly independent, $\mathbf{w}$ *not* in Span$\{\mathbf{u}, \mathbf{v}\}$

$A = \begin{bmatrix}1 & 2\\2 & 4\\3 & 6\end{bmatrix} \leftarrow$ linearly dependent

what non-zero x can make $Ax = 0$?

$$\begin{bmatrix}1 & 2\\2 & 4\\3 & 6\end{bmatrix}\begin{bmatrix}1\\-\frac{1}{2}\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$$

$x = \begin{bmatrix}0\\0\end{bmatrix}$ or $\begin{bmatrix}1\\-\frac{1}{2}\end{bmatrix}$

Example: $v_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, v_2 = \begin{bmatrix}4\\5\\6\end{bmatrix}, v_3 = \begin{bmatrix}2\\1\\0\end{bmatrix}$

is set $\{v_1, v_2, v_3\}$ linearly independent?

Check; does $Ax = 0$ only have trivial solution

$$\begin{bmatrix}1 & 4 & 2 & 0\\2 & 5 & 1 & 0\\3 & 6 & 0 & 0\end{bmatrix} = \begin{bmatrix}1 & 4 & 2 & 0\\0 & -3 & -3 & 0\\0 & 0 & 0 & 0\end{bmatrix} \Rightarrow$$ x_3 is free, there are trivial solutions as x_3 can take any value

$$= \begin{bmatrix}1 & 4 & 2 & 0\\0 & 1 & 1 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

$$= \begin{bmatrix}1 & 0 & -2 & 0\\0 & 1 & 1 & 0\\0 & 0 & 0 & 0\end{bmatrix} \Rightarrow \begin{matrix}x_1 - 2x_3 = 0, & x_1 = 2x_3\\ x_2 + x_3 = 0, & x_2 = -x_3\end{matrix}$$

$\therefore$ let $x_3 = 5$, then $x_1 = 10$, $x_2 = -5$

$\therefore$ $10v_1 - 5v_2 + 5v_3 = 0$, linearly dependent

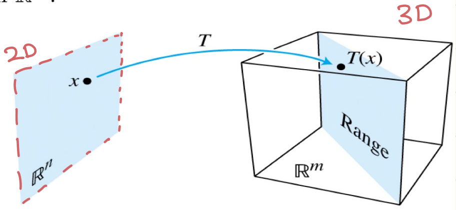
Linear Transformations

Think of $A\mathbf{x} = \mathbf{b}$ as :

matrix A as an object that "acts" on a vector $\mathbf{x}$ by multiplication to produce a new vector called $A\mathbf{x}$.

A **transformation** T from $\mathbb{R}^n$ to $\mathbb{R}^m$ (denoted by $T : \mathbb{R}^n \to \mathbb{R}^m$) is a rule that assigns to each vector $\mathbf{x}$ in $\mathbb{R}^n$ a vector $T(\mathbf{x})$ in $\mathbb{R}^m$.

- $\mathbb{R}^n$ - **domain** of T
- $\mathbb{R}^m$ - **codomain** of T
- $T(\mathbf{x})$ - **image** of $\mathbf{x}$
- Set of all images $T(\mathbf{x})$ - **range** of T



Domain, codomain, and range of $T : \mathbb{R}^n \to \mathbb{R}^m$.

43

Vector $\vec{x}$ exists in the domain $\mathbb{R}^n$

Something (A) multiplies $\vec{x}$

$A\vec{x}$ is now in the codomain $\mathbb{R}^m$

This is called transformation (T)

$T: \vec{x} \rightarrow A\vec{x}$, or $T(\vec{x})$ which means that the transformation T operates on the vector $\vec{x}$ which produces the image of $\vec{x}$.

Example:

Let $A = \begin{bmatrix}1 & -3\\3 & 5\\-1 & 7\end{bmatrix}$, $\mathbf{u} = \begin{bmatrix}2\\-1\end{bmatrix}$, $\mathbf{b} = \begin{bmatrix}3\\2\\-5\end{bmatrix}$, $\mathbf{c} = \begin{bmatrix}3\\2\\5\end{bmatrix}$, and define a transformation $T : \mathbb{R}^2 \to \mathbb{R}^3$ by $T(\mathbf{x}) = A\mathbf{x}$.

a. Find $T(\mathbf{u})$.

b. Find an $\mathbf{x}$ in $\mathbb{R}^2$ whose image under T is $\mathbf{b}$.

c. Is there more than one $\mathbf{x}$ whose image under T is $\mathbf{b}$?

d. Determine if $\mathbf{c}$ is in the range of T.

a) $T(\vec{u}) = A \cdot \vec{u}$

$$= \begin{bmatrix}1 & -3\\3 & 5\\-1 & 7\end{bmatrix}\begin{bmatrix}2\\-1\end{bmatrix}$$

$$= \begin{bmatrix}5\\1\\-9\end{bmatrix}$$

b) Find x in $\mathbb{R}^2$ that image under T is b

$A \cdot \vec{x} = b$

$$\begin{bmatrix}1 & -3\\3 & 5\\-1 & 7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}3\\2\\-5\end{bmatrix}$$

$$\begin{bmatrix}1 & -3 & 3\\3 & 5 & 2\\-1 & 7 & -5\end{bmatrix} = \begin{bmatrix}1 & 0 & -1.5\\0 & 1 & -0.5\\0 & 0 & 0\end{bmatrix} \Rightarrow \begin{matrix}x_1 = -1.5\\x_2 = -0.5\end{matrix}$$

$\therefore$ The vector $\vec{x}$ that is transformed under A to become $b = \begin{bmatrix}-1.5\\-0.5\end{bmatrix}$

c) $A\vec{x} = b$ has one unique solution

d) Range = Set of all transformed $\vec{x}$

$\therefore$ Is $c = A\vec{x}$ for some x

$$\begin{bmatrix}1 & -3\\3 & 5\\-1 & 7\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}3\\2\\5\end{bmatrix}$$

$$[Ac] \Rightarrow \begin{bmatrix}1 & -3 & 3\\3 & 5 & 2\\-1 & 7 & 5\end{bmatrix} = \begin{bmatrix}1 & -3 & 3\\0 & 1 & 2\\0 & 0 & -35\end{bmatrix}$$

The last eq^n is $0 = -35$, inconsistent, and C is not in the range of T.

Definition. A transformation T is **linear** if:

i. $T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v}$ in the domain of T;

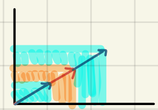
ii. $T(c\mathbf{u}) = cT(\mathbf{u})$

Generalization:

$$T(c_1\mathbf{v}_1 + \ldots + c_p\mathbf{v}_p) = c_1T(\mathbf{v}_1) + \ldots + c_pT(\mathbf{v}_p)$$

Examples:

- $T(\mathbf{x}) = r\mathbf{x}$, $\; T : \mathbb{R}^2 \to: \mathbb{R}^2$ is called a **contraction** when $0 \le r \le 1$ and a **dilation** when $r > 1$.
- Show that $\begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}$ represents a linear transformation by finding the images of $\mathbf{u} = \begin{bmatrix}4\\1\end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix}2\\3\end{bmatrix}$.



① $T(\vec{u}) = \begin{bmatrix}0 & -1\\1 & 0\end{bmatrix}\begin{bmatrix}4\\1\end{bmatrix} = \begin{bmatrix}-1\\4\end{bmatrix}$, $T(\vec{v}) = \begin{bmatrix}-3\\2\end{bmatrix}$

$T(\vec{u}) + T(\vec{v}) = \begin{bmatrix}-4\\6\end{bmatrix}$

$T(\vec{u} + \vec{v}) = T\left(\begin{bmatrix}6\\4\end{bmatrix}\right) = \begin{bmatrix}-4\\6\end{bmatrix}$

② $c\vec{u} = \begin{bmatrix}4c\\c\end{bmatrix}$; $T(c\vec{u}) = \begin{bmatrix}-c\\4c\end{bmatrix}$

$c\,T(\vec{u}) = c\begin{bmatrix}-1\\4\end{bmatrix} = \begin{bmatrix}-c\\4c\end{bmatrix}$

Finding T

When A transforms something, look @ the columns.
Create a test vector, $\begin{bmatrix}1&0\\0&1\end{bmatrix}$, $\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$, $\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$
and let $e_1=\begin{bmatrix}1\\0\\0\end{bmatrix}, e_2=\begin{bmatrix}0\\1\\0\end{bmatrix}, e_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$.

$$\left.\begin{matrix}T(e_1)\\T(e_2)\\T(e_3)\end{matrix}\right\} T=[T(e_1)\quad T(e_2)\quad T(e_3)]$$

Example:

Let $I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}$ and $\mathbf{e}_1=\begin{bmatrix}1\\0\end{bmatrix}$ and $\mathbf{e}_2=\begin{bmatrix}0\\1\end{bmatrix}$. Suppose T is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^3$ such that $T(\mathbf{e}_1)=\begin{bmatrix}5\\-7\\2\end{bmatrix}$ and $T(\mathbf{e}_2)=\begin{bmatrix}-3\\8\\0\end{bmatrix}$. Find the formula for the image of an arbitrary $\mathbf{x}$ in $\mathbb{R}^2$.

Find T for an arbitrary x

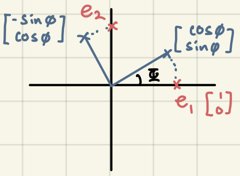
$\therefore\ \vec{x}=\begin{bmatrix}x_1\\x_2\end{bmatrix}=x_1\underbrace{\begin{bmatrix}1\\0\end{bmatrix}}_{e_1}+x_2\underbrace{\begin{bmatrix}0\\1\end{bmatrix}}_{e_2}$

$$T(\vec{x})=T(x_1e_1+x_2e_2)$$
$$=x_1T(e_1)+x_2T(e_2)$$
$$=x_1\begin{bmatrix}5\\-7\\2\end{bmatrix}+x_2\begin{bmatrix}-3\\8\\0\end{bmatrix}$$
$$=\begin{bmatrix}5x_1-3x_2\\-7x_1+8x_2\\2x_1\end{bmatrix}$$

Also, if we want to find T;
$$T=[T(e_1)\quad T(e_2)]$$
$$=\begin{bmatrix}5&-3\\-7&8\\2&0\end{bmatrix}$$

Example:

Let $T:\mathbb{R}^2\to\mathbb{R}^2$ be the transformation that rotates each point in $\mathbb{R}^2$ about the origin through an angle Φ, with counterclockwise rotation for a positive angle. Find the standard matrix A for this transformation.



As this is $\mathbb{R}^2$...
let $e_1=\begin{bmatrix}1\\0\end{bmatrix}, e_2=\begin{bmatrix}0\\1\end{bmatrix}$

$\therefore\ T(e_1)=\begin{bmatrix}\cos\phi\\\sin\phi\end{bmatrix}$

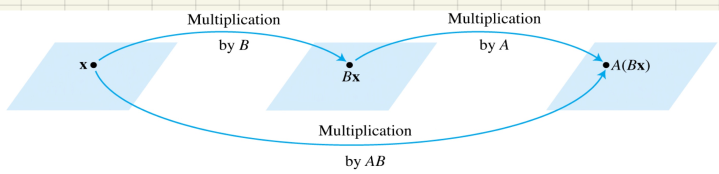
$T(e_2)=\begin{bmatrix}-\sin\phi\\\cos\phi\end{bmatrix}$

$$A=\begin{bmatrix}\cos\phi&-\sin\phi\\\sin\phi&\cos\phi\end{bmatrix}$$

If $T\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}0\\1\end{pmatrix}$, $T\begin{pmatrix}0\\1\end{pmatrix}=\begin{pmatrix}1\\0\\0\end{pmatrix}$, $T\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}3\\1\end{pmatrix}$

This is not linear trans, as $T\begin{pmatrix}0\\1\end{pmatrix}+T\begin{pmatrix}1\\0\end{pmatrix}\neq T\begin{pmatrix}1\\1\end{pmatrix}$

$\longrightarrow$ Chap 2



Multiplication by *AB*.

$$A(B\mathbf{x})=(AB)\mathbf{x}$$

Theorem 2.1. *If A is an invertible $n\times n$ matrix, then for each $\mathbf{b}$ in $\mathbb{R}^n$, the equation $A\mathbf{x}=\mathbf{b}$ has the unique solution $\mathbf{x}=A^{-1}\mathbf{b}$.*

Proof. Let $\mathbf{b}\in\mathbb{R}^n$.

Solution *exists* : Substitute $A^{-1}\mathbf{b}$ in $A\mathbf{x}=\mathbf{b}$.
LHS $=A\underbrace{\mathbf{x}}_{A^{-1}b}=A(A^{-1})\mathbf{b}=(\underbrace{AA^{-1}}_{\text{Identity}})\mathbf{b}=I\mathbf{b}=\mathbf{b}=$ RHS.

Solution is *unique* : Show that if $\mathbf{u}$ is a solution, it must be $A^{-1}\mathbf{b}$.
If $A\mathbf{u}=\mathbf{b}$, multiply both sides by A^{-1}
$A^{-1}A\mathbf{u}=A^{-1}\mathbf{b}$ or $I\mathbf{u}=A^{-1}\mathbf{b}$, i.e., $\mathbf{u}=A^{-1}\mathbf{b}$

Theorem 2.2. *Invertible matrices*

$XY=I$
$YX=I$

1. *If A is an invertible matrix, then A^{-1} is invertible and $(A^{-1})^{-1}=A$*
2. *If A and B are $n\times n$ invertible matrices, then so is AB and $(AB)^{-1}=B^{-1}A^{-1}$* also $(AB)^T=B^TA^T$
3. *If A is an invertible matrix, then so is A^T, and $(A^T)^{-1}=(A^{-1})^T$*

Proof

$A^{-1}C=I$
$CA^{-1}=I$ $\therefore C=A$

1. Find a matrix C such that $A^{-1}C=I$ and $CA^{-1}=I$. Here, C is simply A. Hence, A^{-1} is invertible and its inverse is A.
2. Find a matrix C such that $(AB)C=I$ and $C(AB)=I$. If $C=B^{-1}A^{-1}$, then $AB(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AIA^{-1}=AA^{-1}=I$. Similarly show that $(B^{-1}A^{-1})(AB)=I$.
3.

Definition. An **elementary matrix** is one that is obtained by performing a single elementary row operation on an identity matrix.

Exercise 2.2.1:

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\xrightarrow{E_1}\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix}\qquad r3\leftarrow r3-4r1$$

Let $A=\begin{bmatrix}a&b&c\\d&e&f\\g&h&i\end{bmatrix}\begin{bmatrix}1&0&0\\0&1&0\\-4&0&1\end{bmatrix}$

$E_1A=\begin{bmatrix}a&b&c\\d&e&f\\g-4a&h-4b&i-4c\end{bmatrix}$ also transform by $r3\leftarrow r3-4r1$

How did we get E_2, E_3

$E_2=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$ $r_2\leftrightarrow r_1$ swap

$E_3=\begin{bmatrix}1&0&0\\0&1&0\\0&0&5\end{bmatrix}$ $r_3\leftarrow 5r_3$

$E_2A=?\quad E_3A=?$

$E_2A=\begin{bmatrix}d&e&f\\a&b&c\\g&h&i\end{bmatrix}\qquad E_3A=\begin{bmatrix}a&b&c\\d&e&f\\5g&5h&5i\end{bmatrix}$

inverse

$E_1^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\+4&0&1\end{bmatrix}$ $E_1\rightsquigarrow r_3\leftarrow r_3-4r_1$; $E_1^{-1}\rightsquigarrow r_3\leftarrow r_3+4r_1$

$E_2^{-1}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}$ $E_2\rightsquigarrow r_1\leftrightarrow r_2$; $E_2^{-1}\rightsquigarrow r_2\leftrightarrow r_1$

$E_3^{-1}=\begin{bmatrix}1&0&0\\0&1&0\\0&0&\frac{1}{5}\end{bmatrix}$ $E_3\to r_3\leftarrow 5r_3$; $E_3^{-1}\to r_3\leftarrow\frac{1}{5}r_3$

Theorem 2.3. *An $n \times n$ matrix A is invertible if and only if A is row equivalent to I_n and in this case, any sequence of elementary row operations that reduces A to I_n also transforms I_n to A^{-1}.*

Proof. $\underset{n\times n}{A} \xrightarrow{\text{ERO}} \underset{n\times n}{I}$

$\underset{n\times n}{I} \xrightarrow{\text{same EROs}} \underset{n\times n}{A^{-1}}$ $\square$

Exercise 2.2.3:

Find the inverse of $A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 3 \\ 4 & -3 & 8 \end{bmatrix}$.

1) $[A\ I] = \left[\begin{array}{ccc|ccc} 0 & 1 & 2 & 1 & 0 & 0 \\ 1 & 0 & 3 & 0 & 1 & 0 \\ 4 & -3 & 8 & 0 & 0 & 1 \end{array}\right]$

$\underbrace{}_{\text{make this become } I}$ $\underbrace{}_{\text{This will then become } A^{-1}}$

$$\to \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & -3 & -4 & 0 & -4 & 1 \end{array}\right]$$

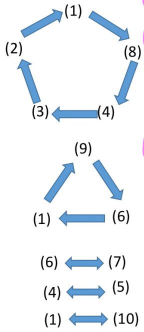
$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 3 & 0 & 1 & 0 \\ 0 & 1 & 2 & 1 & 0 & 0 \\ 0 & 0 & 2 & 3 & -4 & 1 \end{array}\right]$$

$$= \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -9/2 & 7 & -3/2 \\ 0 & 1 & 0 & -2 & 4 & -1 \\ 0 & 0 & 1 & 3/2 & -2 & 1/2 \end{array}\right] \quad \therefore A^{-1} = \begin{bmatrix} -9/2 & 7 & -3/2 \\ -2 & 4 & -1 \\ 3/2 & -2 & 1/2 \end{bmatrix}$$

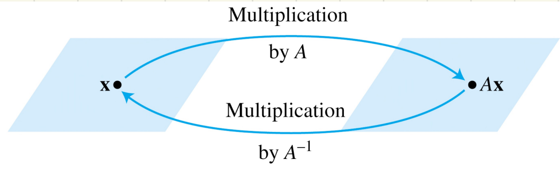
$\underbrace{}_{I}$ $\underbrace{}_{\text{This is } A^{-1}}$

Theorem 2.4. *The Invertible Matrix Theorem*
Let A be a square $n \times n$ matrix. Then the following statements are equivalent, i.e., for a given A, the statements are either all true or all false.

1. *A is an invertible matrix.*
2. *A is row equivalent to I_n.*
3. *A has n pivot positions.*
4. *$A\mathbf{x} = \mathbf{0}$ has only the trivial solution.*
5. *The columns of A form a linearly independent set.*
6. *$A\mathbf{x} = \mathbf{b}$ has at least one solution for each $\mathbf{b}$ in $\mathbb{R}^n$.*
7. *The columns of A span $\mathbb{R}^n$.*
8. *There is an $n \times n$ matrix C such that $CA=I$.*
9. *There is an $n \times n$ matrix D such that $AD=I$.*
10. *A^T is an invertible matrix.*



Invertible Linear Transformation



A^{-1} transforms $A\mathbf{x}$ back to $\mathbf{x}$.

Matrix Factorization (LU factorization)

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

$\Rightarrow LU\mathbf{x} = \mathbf{b}$

Let $\mathbf{y} = U\mathbf{x}$ — 'Forward Substitution'

$L\mathbf{y} = \mathbf{b} \to$ Solve for $\mathbf{y}$

$U\mathbf{x} = \mathbf{y} \to$ Solve for $\mathbf{x}$

} Easy to solve because L and U are triangular

'Backward Substitution'

L: Unit Lower triangular

U: Upper triangular

$$A = \underset{L}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix}} \underbrace{\underset{U}{\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & \blacksquare & * \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}}_{\text{Echelon form}}$$

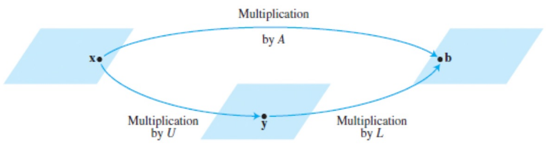


FIGURE 2 Factorization of the mapping $\mathbf{x} \mapsto A\mathbf{x}$.

Example

given L & U

Solve $A\mathbf{x} = \mathbf{b}$ if $A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \underset{L}{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}}$

and $\mathbf{b} = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$.

$A = LU$

$LUx = b$ let $y = Ux$

$Ly = b$

$$[L\ b] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -9 \\ -1 & 1 & 0 & 0 & 5 \\ 2 & -5 & 1 & 0 & 7 \\ -3 & 8 & 3 & 1 & 11 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & -9 \\ 0 & 1 & 0 & 0 & -4 \\ 0 & 0 & 1 & 0 & 5 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

$\therefore y = \begin{bmatrix} -9 \\ 5 \\ 7 \\ 11 \end{bmatrix}$

Number of multiplication - addition pairs to reduce L to I } 3 2 1

6 multiplications
6 additions

$$\begin{aligned} y_1 &= -9 \\ -y_1 + y_2 &= 5 \\ 2y_1 - 5y_2 + y_3 &= 7 \\ -3y_1 + 8y_2 + 3y_3 + y_4 &= 11 \end{aligned}$$

$x = U^{-1} \cdot y$

$Ux = y$

$$[U\ y] = \left[\begin{array}{cccc|c} 3 & -7 & -2 & 2 & -9 \\ 0 & -2 & -1 & 2 & -4 \\ 0 & 0 & -1 & 1 & 5 \\ 0 & 0 & 0 & -1 & 1 \end{array}\right] = \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 3 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 1 & 0 & -6 \\ 0 & 0 & 0 & 1 & -1 \end{array}\right]$$

To reduce U to I:
Number of divisions - 4
Number of additions - 6
Number of multiplications - 6

$$\begin{aligned} 3x_1 - 7x_2 - 2x_3 + 2x_4 &= -9 \\ -2x_2 - x_3 + 2x_4 &= -4 \\ -x_3 + x_4 &= 5 \\ -x_4 &= 1 \end{aligned}$$

Through LU factorization : 28 arithmetic operations or "flops" (floating point operations) - excluding cost of factorization
Through row reduction of $[A\ \ \mathbf{b}]$ to $[I\ \ \mathbf{x}]$: 62 flops

$x = \begin{bmatrix} 3 \\ 4 \\ -6 \\ -1 \end{bmatrix}$

How to factor ?

Find an LU factorization of $A = \begin{bmatrix} 2 & -2 & 3 \\ 6 & -7 & 14 \\ 4 & -8 & 30 \end{bmatrix}$

1) USE I matrix, find the "steps" needed to make $\begin{bmatrix} x & - & \\ 0 & x & \cdots \\ 0 & 0 & x \cdots \end{bmatrix}$

$$A = \begin{bmatrix} 2 & -2 & 3 \\ 6 & -7 & 14 \\ 4 & -8 & 30 \end{bmatrix}$$

involves row 2 & 1 $\left\{ E_{21} A = \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 6 & -7 & 14 \\ 4 & -8 & 30 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 4 & -8 & 30 \end{bmatrix} \right.$

$$E_{31}(E_{21}A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 4 & -8 & 30 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & -4 & 24 \end{bmatrix}$$

$$E_{32}(E_{31}E_{21}A) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & -4 & 24 \end{bmatrix} = \begin{bmatrix} 2 & -2 & 3 \\ 0 & -1 & 5 \\ 0 & 0 & 4 \end{bmatrix} = U$$

$A = LU$ ... how to find L? Observe.

$$\therefore\ L = \begin{bmatrix} 2 & 0 & 0 \\ 6 & -1 & 0 \\ 4 & -4 & 4 \end{bmatrix} \rightsquigarrow \text{Use row reduction}$$

$$\to \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 4 & 1 \end{bmatrix}$$

$\div 2 \quad \times -1 \quad \div 4$

Alternatively: $L = E_{21}^{-1} E_{31}^{-1} E_{32}^{-1}$

$$= \begin{bmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix}^{-1}$$

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

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

---

find LU for $\begin{bmatrix} -1 & 2 & -1 & 3 \\ -2 & 5 & 0 & 7 \\ -1 & 5 & 7 & 7 \end{bmatrix} \xrightarrow[\substack{L_{13} \\ R_3 = R_3 + R_1(-1)}]{\substack{L_{12} \\ R_2 = R_2 + R_1(-2)}} \begin{bmatrix} -1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 3 & 8 & 4 \end{bmatrix} \xrightarrow{\substack{L_{23} \\ R_3 = R_3 + R_2(-3)}} \begin{bmatrix} -1 & 2 & -1 & 3 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 1 \end{bmatrix} = U$ //

$$\therefore\ L_{12} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad L_{13} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \quad L_{23} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -3 & 1 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 3 & 1 \end{bmatrix}$$ //

using LU factorization: $\begin{bmatrix} -1 & 2 & 3 \\ -2 & 2 & 7 \\ -1 & -4 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} -1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$$

Step 1: $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix}$ $\quad x = 1,\ y = 1,\ z = 0$

Step 2: $\begin{bmatrix} -1 & 2 & 3 \\ 0 & -2 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \therefore \left. \begin{aligned} -x + 2y + 3z &= 1 \\ -2y + z &= 1 \end{aligned} \right\} \begin{aligned} x &= 4y \\ z &= 1 + 2y \end{aligned} \Leftrightarrow \left\{ \begin{pmatrix} 4t \\ t \\ 1+2t \end{pmatrix} \middle|\ t \in \mathbb{R} \right\}$

→ Chap3

Determinants

1. **The determinant of the $n \times n$ identity matrix is 1.**

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \text{ and } \begin{vmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{vmatrix} = 1$$

2. **The determinant changes sign when two rows are exchanged.**

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = - \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

If P is a permutation matrix with r row exchanges , then $|P| = 1$ for even r and $|P| = -1$ for odd r.

3. **The determinant is a linear function of each row separately.**

If 1 row of a matrix A is multiplied by t to get A', then $|A'| = t|A|$

$$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\quad |\tfrac{1}{2}A| = \begin{vmatrix} \tfrac{1}{2}a & \tfrac{1}{2}b \\ \tfrac{1}{2}c & \tfrac{1}{2}d \end{vmatrix} = \tfrac{1}{2}\begin{vmatrix} a & b \\ \tfrac{1}{2}c & \tfrac{1}{2}d \end{vmatrix} = \left(\tfrac{1}{2}\right)^2 \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

$\boxed{\begin{bmatrix} x & x & x \\ x & x & x \\ x & x & x \end{bmatrix} = A \quad |5A| = (5)^3\,|A|}$

If one row of A is added to one row of A', then the determinants add.

$$\begin{vmatrix} a + a' & b + b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$$

Important: *This rule applies only when the other rows do not change.*

4. **If two rows of A are equal, then $|A| = 0$.**

$$\begin{vmatrix} a & b \\ a & b \end{vmatrix} = 0.$$

This follows from Rule 2 (Show!).

5. **Subracting a multiple of one row from another row leaves $|A|$ unchanged.**

$$\begin{vmatrix} a & b \\ c - la & d - lb \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

r3) $\begin{vmatrix} a & b \\ c & d \end{vmatrix} - \begin{vmatrix} a & b \\ la & lb \end{vmatrix}$

r4) $= \begin{vmatrix} a & b \\ c & d \end{vmatrix} - l \underbrace{\begin{vmatrix} a & b \\ a & b \end{vmatrix}}_{0}$

This follows from Rule 3 and Rule 4.

$|A| = |U|$ without row exchanges and $|A| = \pm|U|$ with row exchanges.

6. **A matrix with a row of zeros has $|A| = 0$.**

$$\begin{vmatrix} 0 & 0 \\ c & d \end{vmatrix} = 0.$$

This follows from Rule 4 and Rule 5.

7. **If A is triangular, then $|A| = a_{11}a_{22}\ldots a_{nn}$ = product of diagonal entries.**

Consider the determinant of a diagonal matrix:

$$\begin{vmatrix} a_{11} & & & 0 \\ & a_{22} & & \\ & & \ddots & \\ 0 & & & a_{nn} \end{vmatrix} = D$$

$\begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t \begin{vmatrix} a & b \\ c & d \end{vmatrix}$

$\therefore \begin{vmatrix} a_{11} & 0 \\ 0 & a_{22} \end{vmatrix} = a_{11} \begin{vmatrix} 1 & 0 \\ 0 & a_{22} \end{vmatrix}$

Factor a_{11} from the first row. By rule 3, $D = a_{11}D'$.

Factor a_{22} from the second row. By rule 3, $D = a_{11}a_{22}D''$.

Finally, factor a_{nn} from the last row. By rule 3, $D = a_{11}a_{22}\ldots a_{nn}|I|$.

From rule 1, $|I| = 1$. So, $D = a_{11}a_{22}\ldots a_{nn}$.

Now, consider the determinants for the following triangular matrices

$$\begin{vmatrix} a & b \\ 0 & d \end{vmatrix} = D_1 \text{ and } \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} = D_2.$$

Make the off diagonal elements 0 through elimination.

$$R_1 \leftarrow R_1 - \tfrac{b}{d}R_2 : D_1' = \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} = ad$$

eliminate through $b - d\left(\frac{b}{d}\right) = 0$

$$R_2 \leftarrow R_2 - \tfrac{c}{a}R_2 : D_2' = \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} = ad$$

eliminate through $c - a\left(\frac{c}{a}\right) = 0$

If an $a_{ii} = 0$, elimination produces a zero row.

By Rule 5, determinant is unchanged and by Rule 6, determinant = 0.

Such matrices are called **singular**.

8. **If A is singular, then $|A| = 0$. If A is invertible, then $|A| \neq 0$.**

Transform A to U through elimination.
If A is singular:

- U has a zero row

Through rule 7b

- From previous rules, $|A| = |U| = 0$

If A is invertible: $|A| \neq 0$

- U has pivots along its diagonal
- From Rule 7, product of non-zero pivots ⇒ non zero determinant
- $|A| = \pm|U| = \pm$(product of pivots)

[+ for even number of row exchanges and − for odd number of row exchanges]

Pivots of a 2×2 matrix ($a \neq 0$):

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ 0 & d - (c/a)b \end{vmatrix} = ad - bc$$

Upper Δ form!

(Finally, a formula for the determinant!!)

9. **$|AB| = |A||B|$.**

Consider the ratio $D(A) = |AB|/|B|$. If $D(A)$ satisfies rules 1, 2 and 3, then it is a determinant.

- Rule 1 (*Determinant of I*)
  - If $A = I$, then $D(A) = |B|/|B| = 1$
- Rule 2(*Sign reversal*)
  - Two rows of A are exchanged ⇒ Same two rows of $|AB|$ are exchanged ⇒ $|AB|$ changes sign ⇒ $D(A)$ changes sign
- Rule 3 (*Linearity*)
  - When 1 row of A is multiplied by t ⇒ so is 1 row of AB ⇒ $|AB|$ is multiplied by t ⇒ $D(A)$ is multiplied by t
  - When 1 row of A is added to 1 row of A' ⇒ 1 row of AB is added to 1 row of $A'B$ ⇒ determinants add ⇒ dividing by B, the ratios add

The ratio $|AB|/|B|$ has the same properties that define $|A|$.

Therefore, $|AB|/|B| = |A| \Rightarrow |AB| = |A||B|$

If $|B| = 0$, B is singular ⇒ AB is singular ⇒ $|AB| = 0$

$|A||B| = 0$

Therefore $|AB| = |A||B|$

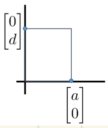
10. **$|A^T| = |A|$**

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc$$

Area / Volume

- Geometric interpretation of determinants

Theorem 3.1. *If A is a 2×2 matrix, the area of the parallelogram determined by the columns of A is $|A|$. If A is a 3×3 matrix, the volume of the parallelepiped determined by the columns of A is $|A|$.*



Proof. True for a 2×2 diagonal matrix:

$\text{abs}(\begin{vmatrix}a & 0\\0 & d\end{vmatrix}) = \text{abs}(ad)$ =area of rectangle ☐

Can we transform any 2×2 matrix $A = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2\end{bmatrix}$ into a diagonal matrix without change in area of the associated parallelogram or in $|A|$?

A can be transformed into a diagonal matrix by:

- Interchanging two columns
  - Does not change the parallelogram
  - From property 2, $|A|$ is unchanged

  Remember: properties apply to *columns* also.

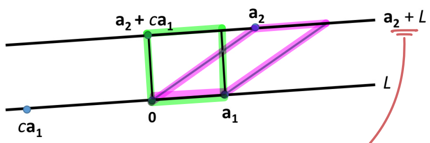
$\left|\begin{bmatrix}a & 0\\0 & d\end{bmatrix}\right| = -\left|\begin{bmatrix}0 & d\\a & 0\end{bmatrix}\right|$

$|A| = -|A'|$

- Adding a multiple of one column to another

Prove the following geometric observation:

Let $\mathbf{a}_1$ and $\mathbf{a}_2$ be nonzero vectors. Then for any scalar c, the area of a parallelogram determined by $\mathbf{a}_1$ and $\mathbf{a}_2$ equals the area of the parallelogram determined by $\mathbf{a}_1$ and $\mathbf{a}_2 + c\mathbf{a}_1$.



Assume $\mathbf{a}_2$ is not a multiple of $\mathbf{a}_1$.

- L is the line through $\mathbf{0}$ and $\mathbf{a}_1 \Rightarrow \mathbf{a}_2 + L$ is the line through $\mathbf{a}_2$ and parallel to L
- Points $\mathbf{a}_2$ and $\mathbf{a}_2 + c\mathbf{a}_1$ have the same perpendicular distance to L
- Hence, two parallelograms have the same area (base X height)

Example

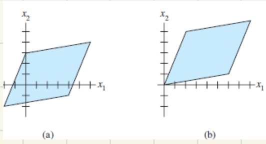
Calculate the area of the parallelogram determined by the points $(-2,-2), (0,3), (4,-1), (6,4)$.

Points are $(-2,-2), (0,3), (4-1), (6,4)$

→ Tag one point to $(0,0)$ ...

for all x & y, $+2$

∴ Points $(0,0), (2,5), (6,1), (8,6)$



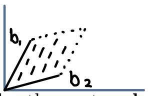
$A = \begin{bmatrix}2 & 6\\5 & 1\end{bmatrix}$

$\text{abs}(|A|) = |-28| = 28$

Transformation

Theorem 3.2. *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation determined by a 2×2 matrix A. If S is a parallelogram in $\mathbb{R}^2$, then area of $T(S) = abs(|A|)\times$ area of S.*
If T is determined by a 3×3 matrix A, and if S is a parallelepiped in $\mathbb{R}^3$, then volume of $T(S) = abs(|A|)\times$ volume of S.

Proof.
Consider the 2×2 case, $A = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2\end{bmatrix}$
A parallelogram at the origin in $\mathbb{R}^2$ determined by the vectors $\mathbf{b}_1$ and $\mathbf{b}_2$ has the form
$S = \{s_1\mathbf{b}_1 + s_2\mathbf{b}_2 : 0 \le s_1 \le 1, 0 \le s_2 \le 1\}$



The image of S under T consists of the points of the form

$$\underbrace{T(s_1\mathbf{b}_1 + s_2\mathbf{b}_2)}_{\text{Transform}} = s_1T(\mathbf{b}_1) + s_2T(\mathbf{b}_2) = s_1A\mathbf{b}_1 + s_2A\mathbf{b}_2,$$

where $0 \le s_1 \le 1, 0 \le s_2 \le 1$.

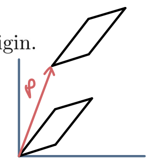
$T(S)$ is the parallelogram determined by columns of $\begin{bmatrix}A\mathbf{b}_1 & A\mathbf{b}_2\end{bmatrix} = AB$ where $B = \begin{bmatrix}\mathbf{b}_1 & \mathbf{b}_2\end{bmatrix}$.

$$\text{area of } T(S) = abs(|AB|) = (abs|A|)(abs|B|) = (abs|A|)(\text{area of } S)$$

Now for the general case:

An arbitrary parallelogram has the form $\mathbf{p} + S$
where $\mathbf{p}$ is a vector and S is a parallelogram at the origin.

$T(\mathbf{p} + S) = T(\mathbf{p}) + T(S)$



Translation does not affect the area of a set

$$\begin{aligned}\text{area of } T(\mathbf{p}+S) &= \text{area of } (T(\mathbf{p}) + T(S))\\ &= \text{area of } T(S)\\ &= abs(|A|) \times \text{area of } S\\ &= abs(|A|) \times \text{area of } \mathbf{p}+S\end{aligned}$$

Proof for 3×3 is analogous.

For triangle

Vector Spaces

Definition. A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication by scalars* (real numbers), subject to the ten axioms below. The axioms must hold for all vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ and for all scalars c and d.

sum of u+v must be in the same vector space

1. The sum of $\mathbf{u}$ and $\mathbf{v}$, denoted by $\mathbf{u}+\mathbf{v}$ is in V.
2. $\mathbf{u}+\mathbf{v}=\mathbf{v}+\mathbf{u}$.
3. $(\mathbf{u}+\mathbf{v})+\mathbf{w}=\mathbf{u}+(\mathbf{v}+\mathbf{w})$.
4. There is a **zero vector** $\mathbf{0}$ in V such that $\mathbf{u}+\mathbf{0}=\mathbf{u}$.

Some zero vector

5. For each $\mathbf{u}$ in V, there is a vector $-\mathbf{u}$ in V such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.
6. The scalar multiple of $\mathbf{u}$ by c, denoted by $c\mathbf{u}$ is in V.
7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$.
8. $(c+d)\mathbf{u}=c\mathbf{u}+d\mathbf{u}$.
9. $c(d\mathbf{u})=(cd)\mathbf{u}$.
10. $1\mathbf{u}=\mathbf{u}$.

Following simple facts can be proved from the axioms: For each $\mathbf{u}$ in V and scalar c,

- $0\mathbf{u}=\mathbf{0}$.
- $c\mathbf{0}=\mathbf{0}$.
- $-\mathbf{u}=(-1)\mathbf{u}$.

Examples of vector spaces

1. $\mathbb{R}^n$ $\quad -\infty$ to ∞

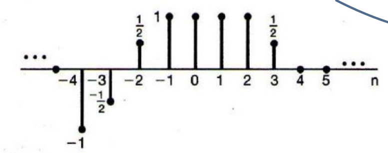
2. $\mathbb{S}$: space of all doubly infinite sequences of numbers

$$\{y_k\}=(\ldots,y_{-2},y_{-1},y_0,y_1,y_2,\ldots)$$

Define addition: If $\{z_k\}$ is another element in $\mathbb{S}$, then sum $\{y_k\}+\{z_k\}$ is the sequence $\{y_k+z_k\}$ formed by adding corresponding terms of $\{y_k\}$ and $\{z_k\}$.

$\{y_k\}+\{z_k\}=\{y_k\}+\{z_k\}$
$\vdots$
$y_{-2}+z_{-2}$
$y_{-1}+z_{-1}$
y_0+z_0
y_1+z_0

Define scalar multiplication: The scalar multiple $c\{y_k\}$ is the sequence $\{cy_k\}$.



3. $\mathbb{P}_n$: polynomials of degree n, $(n>0)$

$$\mathbf{p}(t)=a_0+a_1t+a_2t^2+\cdots+a_nt^n$$

Degree : highest power of t in $\mathbf{p}$. If $\mathbf{p}(t)=a_0\neq 0$, degree of $\mathbf{p}$ is zero.

If all the coefficients are zero, $\mathbf{p}$ is called the zero polynomial.

to make vector space valid (rule 4)

Define addition: If $\mathbf{q}(t)=b_0+b_1t+b_2t^2+\cdots+b_nt^n$, then the sum $\mathbf{p}+\mathbf{q}$ is defined by

$$\begin{aligned}(\mathbf{p}+\mathbf{q})(t) &= \mathbf{p}(t)+\mathbf{q}(t)\\ &= (a_0+b_0)+(a_1+b_1)t+\cdots+(a_n+b_n)t^n\end{aligned}$$

Define scalar multiplication: The scalar multiple $c\mathbf{p}$ is the polynomial

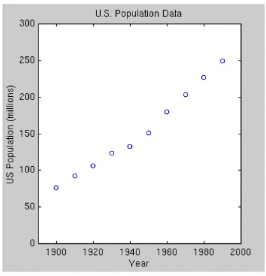
$$(c\mathbf{p})(t)=c\mathbf{p}(t)=ca_0+(ca_1)t+\cdots+(ca_n)t^n$$

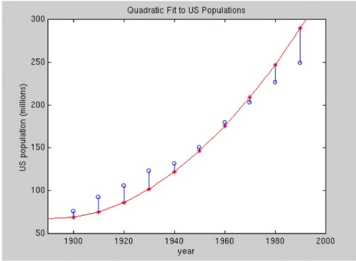
$\vec{u} + \vec{v}$ in V $\quad$ $p(t) + q(t) = v(t)$ in P_n

- Axioms 1 and 6: satisfied because $\mathbf{p} + \mathbf{q}$ and $c\mathbf{p}$ are polynomials of degree less than or equal to n
- Axioms 2, 3, and 7-10: satisfied because of properties of real numbers
- Axiom 4: zero polynomial acts as the zero vector
- Axiom 5: $(-1)\mathbf{p}$ acts as the negative of $\mathbf{p}$

4: $\vec{u} + \vec{0} = \vec{u}$
$\therefore p + 0 = p$

Non-linear curve fitting





Examples of vector spaces (contd.)

4. Let V be the set of all real-valued functions defined on a set $\mathbb{D}$

Define addition: $\mathbf{f} + \mathbf{g}$ is the function whose value at t in the domain $\mathbb{D}$ is $\mathbf{f}(t) + \mathbf{g}(t)$.

Define scalar multiplication: The scalar multiple $c\mathbf{f}$ is the function whose value at t is $c\mathbf{f}(t)$.

Check axioms are true for Vspace to be valid

- Axioms 1 and 6: obvious
- Axiom 4: zero vector is the function that is identically zero, $\mathbf{f}(t) = 0$
- Axiom 5: Negative of $\mathbf{f}$ is $(-1)\mathbf{f}$
- Rest of the axioms: satisfied because of properties of real numbers

Subspace

S is Subspace when
1) $\underset{\sim}{0} \in S$
2) $\underset{\sim}{u}, \underset{\sim}{v} \in S \Rightarrow \underset{\sim}{u} + \underset{\sim}{v} \in S$
3) $\alpha \in \mathbb{R}$, $\underset{\sim}{v} \in S \Rightarrow \alpha \underset{\sim}{v} \in S$

Definition. A **subspace** of a vector space V is a subset H of V that has three properties:

$\vec{0}$ is in vector space V, $\therefore$ $\vec{0}$ will be in subspace H

a. The zero vector of V is in H.

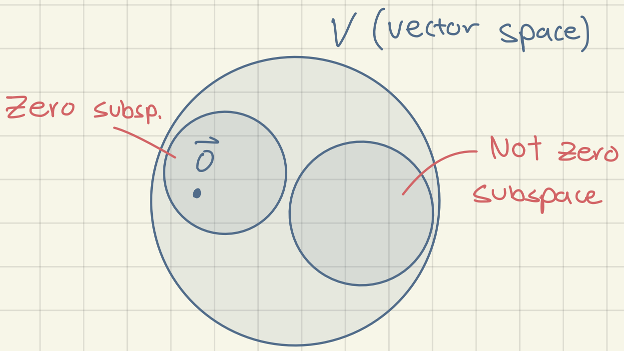
b. H is closed under vector addition, i.e., for each $\mathbf{u}$ and $\mathbf{v}$ in H, the sum $\mathbf{u}+\mathbf{v}$ is in H.

c. H is closed under scalar multiplication, i.e., for each $\mathbf{u}$ in H and each scalar c, the vector $c\mathbf{u}$ is in H.

Subspace H of V is itself a *vector space* under vector space operations already defined in V.

Axiom of V = Axiom of H

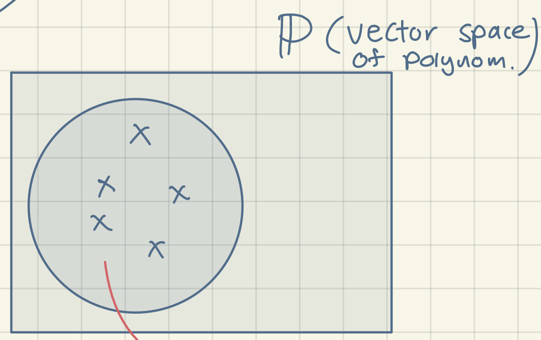
- Axioms 1, 4 and 6: same as (a), (b) and (c)
- Axioms 2, 3, and 7-10: automatically true in H because they apply to all elements of V, including those in H
- Axiom 5: if $\mathbf{u}$ is in H, then $(-1)\mathbf{u}$ is in H [by property (c) and by $(-1)\mathbf{u} = -\mathbf{u}$]



Examples of subspaces

1. The set consisting of only the zero vector in a vector space V : **zero subspace** written as $\{\mathbf{0}\}$.

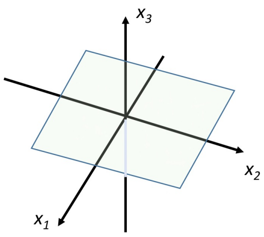
2. $\mathbb{P}$: set of all polynomials with real coefficients, with operations in $\mathbb{P}$ defined as for functions.
For each $n \geq 0$, $\mathbb{P}_n$ is a subspace of $\mathbb{P}$. [properties (a), (b) and (c)]



$\mathbb{P}_n$ subspace
1) addition of two poly is still in subsp.
$x^3 \times a = ax^3$
$x^3 + x^3 = 2x^3$

3. Is $\mathbb{R}^2$ a subspace of $\mathbb{R}^3$?

Is the set $H = \left\{ \begin{bmatrix} s \\ t \\ 0 \end{bmatrix} : s \text{ and } t \text{ are real} \right\}$ a subset of $\mathbb{R}^3$?



Recall: Span$\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is the set of all vectors that can be written as linear combinations of $\mathbf{v}_1, \cdots, \mathbf{v}_p$.

$a_1V_1 + a_2V_2 + \ldots a_nV_n$

all possible combinations of $a_1v_1 + a_2v_2$

Exercise 4.3.1

Given $\mathbf{v}_1$ and $\mathbf{v}_2$ in a vector space V, let $H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. Show that H is a subspace of V.

→ 3 conditions
a) zero exist → $0v_1 + 0v_2 = 0$ ✓
b) closed under addition? → $(a_1v_1 + a_2v_2) + (s_1v_1 + s_2v_2) = (s_1+a_1)(v_1) + (s_2+a_2)(v_2)$ ✓ ($\vec{u} + \vec{w}$ is linear combi of the subsp)
c) " " multiplication? → $(a_1v_1 + a_2v_2) * c = ca_1v_1 + ca_2v_2$ ✓

all tests passed
$\therefore$ H is subspace of V

Theorem 4.1. *If $\mathbf{v}_1, \cdots, \mathbf{v}_p$ are in a vector space V, then Span$\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is a subspace of V.*

Span$\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$ is called **the subspace spanned** by $\mathbf{v}_1, \cdots, \mathbf{v}_p$.
Given any subspace H of V, a **spanning set** for H is a set $\mathbf{v}_1, \cdots, \mathbf{v}_p$ in H such that $H = \text{Span}\{\mathbf{v}_1, \cdots, \mathbf{v}_p\}$.

Let H be the set of all vectors of the form $(a-3b, b-a, a, b)$ where a and b are in $\mathbb{R}$. Show that H is a subspace of $\mathbb{R}^4$.

Solution:

$$H = \begin{bmatrix} a-3b \\ b-a \\ a \\ b \end{bmatrix} = a \underbrace{\begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix}}_{\mathbf{v}_1} + b \underbrace{\begin{bmatrix} -3 \\ 1 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{v}_2}.$$

$H = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$.

$\therefore H$ is a subspace of $\mathbb{R}^4$.

$H = \left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}$

4 entries, or 4D

Null Space ō Matrix

System of homogeneous equations

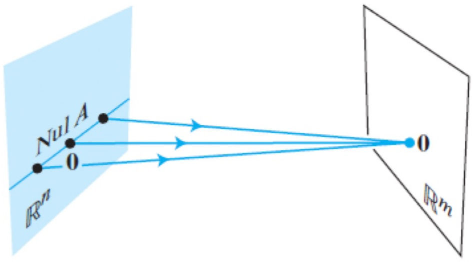
$$\begin{aligned} x_1 \quad - \quad 3x_2 \quad - \quad 2x_3 &= 0 \\ -5x_1 \quad + \quad 9x_2 \quad + \quad x_3 &= 0 \end{aligned}$$

Rewritten as $A\mathbf{x} = \mathbf{0}$, where $A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$

Null space: the set of $\mathbf{x}$ that satisfy $A\mathbf{x} = \mathbf{0}$

Definition. The **null space** of an $m \times n$ matrix A, written as $\mathbf{N}(A)$, is the set of all soutions of the homogeneous equation $A\mathbf{x} = \mathbf{0}$, i.e.,

$$\mathbf{N}(A) = \{\mathbf{x} : \mathbf{x} \text{ is in } \mathbb{R}^n \text{ and } A\mathbf{x} = \mathbf{0}\}$$



All $\mathbf{x}$ in $\mathbb{R}^n$ mapped into the zero vector in $\mathbb{R}^m$ via the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$.

Exercise 4.4.1

Let $A = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix}$ and $\mathbf{u} = \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$. Determine if $\mathbf{u}$ belongs to $\mathbf{N}(A)$.

$\rightarrow Au = \begin{bmatrix} 1 & -3 & -2 \\ -5 & 9 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 3 \\ -2 \end{bmatrix}$

$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$,

∴ u belongs to nullspace

or. $1(5) - 3(3) - 2(-2) = 0$

$-5(5) + 9(3) + 1(-2) = 0$

Exercise 4.4.2

Describe the null space of $A = \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}$.

$\begin{bmatrix} 1 & 2 & 0 \\ 3 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

Solution

Apply elimination on $A\mathbf{x} = \mathbf{0}$.

$$\begin{aligned} x_1 + 2x_2 &= 0 \\ 3x_1 + 6x_2 &= 0 \end{aligned} \rightarrow \begin{aligned} x_1 + 2x_2 &= 0 \\ 0 &= 0 \end{aligned}$$

The line $x_1 + 2x_2 = 0$ is $\mathbf{N}(A)$. It contains all solutions (x_1, x_2).

Set free variable x_2 to some value, say, 1. Then $x_1 = -2$.

$\mathbf{N}(A)$ contains all multiples of $\begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Exercise 4.4.3

Find a spanning set for the null space of the matrix

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 & 0 \\ 1 & -2 & 2 & 3 & -1 & 0 \\ 2 & -4 & 5 & 8 & -4 & 0 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$x_1 \quad -2x_2 \quad -x_4 + x_5 = 0$

$x_3 + 2x_4 - 2x_5 = 0$

$x_1 = 2x_2 + x_4 - x_5$

$x_3 = -2x_4 + 2x_5$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 2x_2 + x_4 - x_5 \\ x_2 \\ -2x_4 + x_5 \\ x_4 \\ x_5 \end{bmatrix} = x_2 \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

∴ Spanning set ō $H = \left\{ \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

Column Space ō Matrix

Definition. The **column space** of an $m \times n$ matrix A, written as $\mathbf{C}(A)$, is the set of all linear combinations of the columns of A. If $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$, then $\mathbf{C}(A) = \text{Span}\{\mathbf{a}_1 \cdots \mathbf{a}_n\}$, i.e.,

$V_1 a_1 + V_2 a_2 + \ldots V_n a_n$

$$\mathbf{C}(A) = \{\mathbf{b} : \mathbf{b} = A\mathbf{x} \text{ for some } \mathbf{x} \text{ in } \mathbb{R}^n\}$$

eg. $A = [a_1, a_2]$ → part of column space

$x_1 a_1 + x_2 a_2 = \vec{b}$ ... alternatively, $[a_1, a_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \vec{b}$ → The set of all these $\underline{b}$

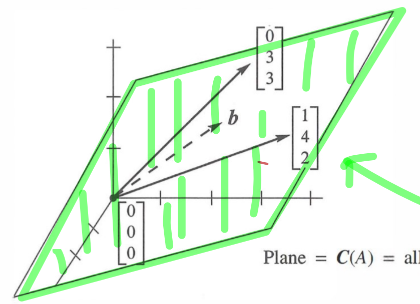
Theorem 4.3. *The column space of an $m \times n$ matrix A is a subspace of $\mathbb{R}^m$.*

Proof. $\text{Span}\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$ is a subspace [Theorem 4.1]

Columns of A are in $\mathbb{R}^m$. → $\begin{bmatrix} \vdots \\ m \end{bmatrix}$ □

Recall (from Chapter 1, slide 29)): The columns of A span $\mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$. → Similar to above!

Restating the above:
The column space of an $m \times n$ matrix A is all of $\mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for each $\mathbf{b}$ in $\mathbb{R}^m$.



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

$\boldsymbol{b} = .4 \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + .3 \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}$

$A\boldsymbol{x} = \boldsymbol{b}$ has $\boldsymbol{x} = \begin{bmatrix} .4 \\ .3 \end{bmatrix}$

Plane $= \boldsymbol{C}(A) =$ all vectors $A\boldsymbol{x}$

Linear combi of collumns of A:

$\underbrace{A\vec{x}} = x_1 \begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 3 \\ 3 \end{bmatrix}$

This forms a plane!

Exercise 4.5.1

Let $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$. 3×4

a. If the column space of A is a subspace of $\mathbb{R}^k$, what is k?

b. If the null space of A is a subspace of $\mathbb{R}^k$, what is k?

c. Find a nonzero vector in $\mathbf{C}(A)$ and a nonzero vector in $\mathbf{N}(A)$.

d. If $\mathbf{u} = \begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$

i. Determine if $\mathbf{u}$ is in $\mathbf{N}(A)$. Could $\mathbf{u}$ be in $\mathbf{C}(A)$?

ii. Determine if $\mathbf{v}$ is in $\mathbf{C}(A)$. Could $\mathbf{v}$ be in $\mathbf{N}(A)$?

a. → $a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 = \begin{bmatrix} \vec{b}_1 \end{bmatrix}_{3 \times 1}$

∴ Column Subspace in R^3

b. → $\begin{bmatrix} 2 & 4 & -2 & 1 & 0 \\ -2 & -5 & 7 & 3 & 0 \\ 3 & 7 & -8 & 6 & 0 \end{bmatrix}$

∴ There are 4 Var, nullspace x in R^4

c. C(A) → $a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 = \begin{bmatrix} \vec{b}_1 \end{bmatrix}$

possible $\begin{bmatrix} \vec{b}_1 \end{bmatrix} = 1(x_1)$

N(A) → $\begin{bmatrix} 2 & 4 & -2 & 1 & 0 \\ -2 & -5 & 7 & 3 & 0 \\ 3 & 7 & -8 & 6 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 9 & 0 & 0 \\ 0 & 1 & -5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

$x_1 = -9x_3$
$x_2 = 5x_3$
$x_4 = 0$

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = x_3 \begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}$

One vector can be $\begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}$

d. i. → Nullspace: $x \begin{bmatrix} -9 \\ 5 \\ 1 \\ 0 \end{bmatrix}$, not a multiple of $\begin{bmatrix} 3 \\ -2 \\ -1 \\ 0 \end{bmatrix}$

∴ not nullspace.

→ Column space? linear Combination ō A is only 3×1. U is 4×1. can't fit. ∴ not column subspace

ii. → V is 3×1, R^3 nullspace requives 4×1, R^4

∴ not nullspace

Column space? Check

→ $\left[\begin{array}{cccc|c} 2 & 4 & -2 & 1 & 3 \\ -2 & -5 & 7 & 3 & -1 \\ 3 & 7 & -8 & 6 & 3 \end{array}\right] \sim \begin{bmatrix} 1 & 0 & 9 & 0 & 5 \\ 0 & 1 & -5 & 0 & -\frac{36}{17} \\ 0 & 0 & 0 & 1 & \frac{1}{17} \end{bmatrix}$

→ The eq^n is consistent, ∴ $\vec{V}$ is in C(A)

Kernel & range of Linear Transform

- Generalize definition of linear transformation to include vector spaces

$A\vec{x} = \vec{b}$

Definition. A **linear transformation** T from a vector space V into a vector space W is a rule that assigns to each vector $\mathbf{x}$ in V a unique vector $T(\mathbf{x})$ in W, such that

1. $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$ for all $\mathbf{u}$, $\mathbf{v}$ in V, and
2. $T(c\mathbf{u}) = cT(\mathbf{u})$ for all $\mathbf{u}$ in V and all scalars c.

same as linear transformation

Definition. The **kernel**(or **null space**) of such a T is the set of all $\mathbf{u}$ in V such that $T(\mathbf{u}) = \mathbf{0}$ (the zero vector in W).

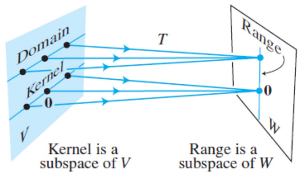
null space

Definition. The **range** of T is the set of all vectors in W of the form $T(\mathbf{x})$ for some $\mathbf{x}$ in V.

column space

If $T(\mathbf{x}) = A\mathbf{x}$, then

- Kernel $= \mathbf{N}(A)$
- Range $= \mathbf{C}(A)$



Bases

- Linear independence (again! introduced in Sec 1.9 for $\mathbb{R}^n$)

Definition. An indexed set of vectors $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ in V is said to be **linearly independent** if the vector equation

all $x = 0$, eqn $= 0$

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

has only the trivial solution. The set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is said to be **linearly dependent** if there exist weights $c_1, \ldots, c_p$, not all zero, such that

some $x \neq 0$, eqn $= 0$

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}$$

- A set containing a single vector $\mathbf{v}$ is linearly independent if and only if $\mathbf{v} \neq \mathbf{0}$. $0 * v_1 = 0$, $1 * v_1 = 0$, $\therefore v_1 == 0$
- A set of two vectors is linearly dependent if and only if one of the vectors is a multiple of other. $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$
- Any set containing the zero vector is linearly dependent.

$3(0) = 0$

$x_1 = 3$, $x_1(0) = 0$

Theorem 4.4. *An indexed set $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ of two or more vectors, with $\mathbf{v}_1 \neq \mathbf{0}$, is linearly dependent if and only if some $\mathbf{v}_j$ (with $j > 1$) is a linear combination of the preceding vectors $\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}$.*

(Skip proof)

$P_3(t) = 4P_1(t) + P_2(t)$

$\therefore$ linear dependent

Examples: Linearly independent or not?

- $\mathbf{p}_1(t) = 1$, $\mathbf{p}_2(t) = t$, and $\mathbf{p}_3(t) = 4 - t$.
- The set $\{\sin t, \cos t\}$ in $C[0,1]$, the space of all continuous functions on $0 \leq t \leq 1$ → Not multiples of each other, $\therefore$ independent
- The set $\{\sin t \cos t, \sin 2t\}$ in $C[0,1]$, the space of all continuous functions on $0 \leq t \leq 1$

$\sin 2t = 2 \sin t \cos t$

this set is linearly dependent

Definition. Let H be a subspace of a vector space V. An indexed set of vectors $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$ in V is a **basis for** H if

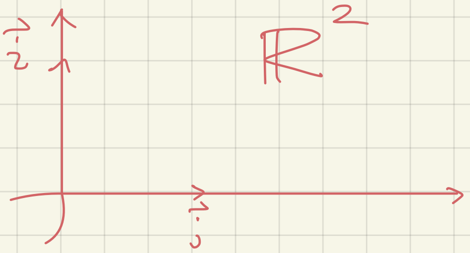
(i) $\mathcal{B}$ is a linearly independent set, and

(ii) the subspace spanned by $\mathcal{B}$ coincides with H, i.e.,

$$H = \text{Span}\{\mathbf{b}_1, \ldots, \mathbf{b}_p\}$$

Vector subspace is the span of the set of vectors

- Also true when $H = V$ since every vector space is a subspace of itself.
  $\Rightarrow$ a basis of V is a linearly independent set that spans V.



are $\vec{i}$ & $\vec{j}$ dependent?
No. they are linearly independent

However, any vector can be expressed in $\mathbb{R}^2$, e.g. $\vec{x} = a.\vec{i} + b.\vec{j}$

$\vec{j}$ & $\vec{i}$ are the basis of $\mathbb{R}^2$

☆ Basis is just the vector for each "coordinate" of vector

Examples:

- Invertible matrix $A = [\mathbf{a}_1 \ \cdots \ \mathbf{a}_n]$

  Do these columns form a basis for $\mathbb{R}^n$?

  Columns of A form a basis for $\mathbb{R}^n$

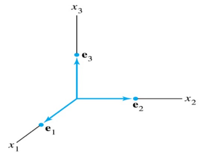
  – they are linearly independent
  – they span $\mathbb{R}^n$

  } Invertible matrix theorem (Theorem 2.4)

- Columns of I_n

$$\mathbf{e}_1 = \begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix}0\\1\\\vdots\\0\end{bmatrix}, \cdots \mathbf{e}_n = \begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$$

unit $\vec{v}$ for each coordinate



The standard basis for $\mathbb{R}^3$.

The set $\{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ is called the **standard basis** for $\mathbb{R}^n$

$$\begin{bmatrix}1 & & 0\\ & 1 & \\ 0 & & \ddots \; 1\end{bmatrix} \rightsquigarrow \text{Identity Matrix}$$

Exercise 4.7.1:

Let $\mathbf{v}_1 = \begin{bmatrix}3\\0\\-6\end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix}-4\\1\\7\end{bmatrix}$ and $\mathbf{v}_3 = \begin{bmatrix}-2\\1\\5\end{bmatrix}$. Determine if $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a basis for $\mathbb{R}^3$.

☆ Does $V_1 + V_2 + V_3$ form a matrix with valid pivot points?

$$\begin{bmatrix}3 & -4 & -2\\0 & 1 & 1\\-6 & 7 & 5\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}$$

∴ yes, it is a basis for $\mathbb{R}^3$

OR... If linearly independent, it forms a basis

$$\begin{bmatrix}3 & -4 & -2 & 0\\0 & 1 & 1 & 0\\-6 & 7 & 5 & 0\end{bmatrix} = \begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 1 & 0\end{bmatrix}$$

Only the trivial solution exists.
∴ they are the basis

Exercise 4.7.2:

Let $S = \{1, t, t^2, \ldots, t^n\}$. Verify that S is a basis for $\mathbb{P}_n$. This basis is called the **standard basis** for $\mathbb{P}_n$.

$\mathbb{P}_n$ is polynomial vector space

→ Set S spans $\mathbb{P}_n$...

∴ any polynomial can be written in $\mathbb{P}^n$: eg: $3t^2 + t' + 3$

also, it is linearly independent as

$c_1(1) + c_2 t + c_3 t^2 + \ldots + c_n t^n = 0(t)$

only the trivial solution exists.

Second half

Vectors

Norm ō vector = length = $\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$

Unit vector = $\hat{v} = \dfrac{v}{\|v\|}$ $\quad \therefore \|\hat{v}\| = 1$

THEOREM 3.2.1 *If* $\mathbf{v}$ *is a vector in* R^n, *and if* k *is any scalar, then:*

(*a*) $\|\mathbf{v}\| \geq 0$

(*b*) $\|\mathbf{v}\| = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$

(*c*) $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$

Example:

Let $\mathbf{v} = (1, -2, 2, 0)$. Find a unit vector $\mathbf{u}$ in the same direction as $\mathbf{v}$.

$V \cdot V = \|V\|\|V\| \cos(0)$
$= \|V\|^2$
$\|V\|^2 = (1)^2 + (-2)^2 + (2)^2 + (0)^2$
$= 9$
$\|V\| = \sqrt{9}$
$= 3$

$\therefore \hat{U} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{bmatrix}$

Example:

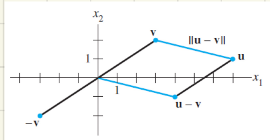
Compute the distance between the vectors $\mathbf{u} = (7, 1)$ and $\mathbf{v} = (3, 2)$.

$\vec{OU} = (7,1) \quad \vec{OV} = (3,2)$

$\vec{OU} - \vec{OV} = \vec{VU} = (4, -1)$

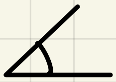
$\|\vec{VU}\|^2 = (4)^2 + (-1)^2$

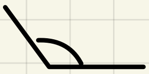
$\|\vec{VU}\| = \sqrt{17}$

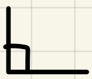


Eucleadian inner product

Dot Product: $U \cdot V = \|U\|\|V\| \cos\theta$


$U \cdot V > 0$


$U \cdot V < 0$


$U \cdot V = 0$

(*a*) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$ **[Symmetry property]**

(*b*) $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$ **[Distributive property]**

(*c*) $k(\mathbf{u} \cdot \mathbf{v}) = (k\mathbf{u}) \cdot \mathbf{v}$ **[Homogeneity property]**

(*d*) $\mathbf{v} \cdot \mathbf{v} \geq 0$ *and* $\mathbf{v} \cdot \mathbf{v} = 0$ *if and only if* $\mathbf{v} = \mathbf{0}$ **[Positivity property]**

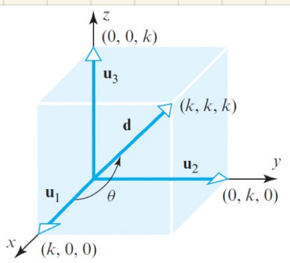
(*a*) $\mathbf{0} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{0} = 0$

(*b*) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

(*c*) $\mathbf{u} \cdot (\mathbf{v} - \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} - \mathbf{u} \cdot \mathbf{w}$

(*d*) $(\mathbf{u} - \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} - \mathbf{v} \cdot \mathbf{w}$

(*e*) $k(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (k\mathbf{v})$



▲ Figure 3.2.7

Note that the angle θ obtained in Example 7 does not involve k. Why was this to be expected?

▶ **EXAMPLE 7 A Geometry Problem Solved Using Dot Product**

Find the angle between a diagonal of a cube and one of its edges.

Solution Let k be the length of an edge and introduce a coordinate system as shown in Figure 3.2.7. If we let $\mathbf{u}_1 = (k, 0, 0)$, $\mathbf{u}_2 = (0, k, 0)$, and $\mathbf{u}_3 = (0, 0, k)$, then the vector

$$\mathbf{d} = (k, k, k) = \mathbf{u}_1 + \mathbf{u}_2 + \mathbf{u}_3$$

is a diagonal of the cube. It follows from Formula (13) that the angle θ between $\mathbf{d}$ and the edge $\mathbf{u}_1$ satisfies

$$\cos\theta = \frac{\mathbf{u}_1 \cdot \mathbf{d}}{\|\mathbf{u}_1\|\|\mathbf{d}\|} = \frac{k^2}{(k)(\sqrt{3k^2})} = \frac{1}{\sqrt{3}}$$

With the help of a calculator we obtain

$$\theta = \cos^{-1}\left(\frac{1}{\sqrt{3}}\right) \approx 54.74° \blacktriangleleft$$

Form	Dot Product	Example	
$\mathbf{u}$ a column matrix and $\mathbf{v}$ a column matrix	$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T\mathbf{v} = \mathbf{v}^T\mathbf{u}$	$\mathbf{u} = \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix}$ $\mathbf{v} = \begin{bmatrix} 5 \\ 4 \\ 0 \end{bmatrix}$	$\mathbf{u}^T\mathbf{v} = \begin{bmatrix} 1 & -3 & 5 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 0 \end{bmatrix} = -7$ $\mathbf{v}^T\mathbf{u} = \begin{bmatrix} 5 & 4 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \\ 5 \end{bmatrix} = -7$

If A is an $n \times n$ matrix and $\mathbf{u}$ and $\mathbf{v}$ are $n \times 1$ matrices, then it follows from the first row in Table 1 and properties of the transpose that

$$A\mathbf{u} \cdot \mathbf{v} = \mathbf{v}^T(A\mathbf{u}) = (\mathbf{v}^T A)\mathbf{u} = (A^T\mathbf{v})^T\mathbf{u} = \mathbf{u} \cdot A^T\mathbf{v}$$
$$\mathbf{u} \cdot A\mathbf{v} = (A\mathbf{v})^T\mathbf{u} = (\mathbf{v}^T A^T)\mathbf{u} = \mathbf{v}^T(A^T\mathbf{u}) = A^T\mathbf{u} \cdot \mathbf{v}$$

The resulting formulas

$$A\mathbf{u} \cdot \mathbf{v} = \mathbf{u} \cdot A^T\mathbf{v} \tag{26}$$

$$\mathbf{u} \cdot A\mathbf{v} = A^T\mathbf{u} \cdot \mathbf{v} \tag{27}$$

$$Au \cdot v = u \cdot A^T v$$
$$u \cdot Av = A^T u \cdot v$$

Example:

$$A = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad \mathbf{u} = \begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix}$$

$$\boxed{v^T v = v \cdot v}$$

$$AU = \begin{bmatrix} 1 & -2 & 3 \\ 2 & 4 & 1 \\ -1 & 0 & 1 \end{bmatrix}\begin{bmatrix} -1 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ 10 \\ 5 \end{bmatrix} \quad A^T v = \begin{bmatrix} 1 & 2 & -1 \\ -2 & 4 & 0 \\ 3 & 1 & 1 \end{bmatrix}\begin{bmatrix} -2 \\ 0 \\ 5 \end{bmatrix} = \begin{bmatrix} -7 \\ 4 \\ -1 \end{bmatrix}$$

$$Au \cdot v = 11 \qquad u \cdot A^T v = 11$$

ORTHOGONALITY

THEOREM 3.2.4 Cauchy–Schwarz Inequality

If $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ *and* $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ *are vectors in* R^n, *then*

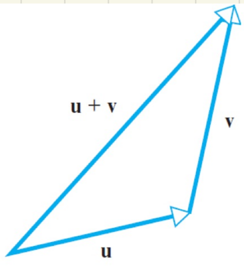
$$|\mathbf{u} \cdot \mathbf{v}| \le \|\mathbf{u}\|\|\mathbf{v}\| \tag{22}$$

or in terms of components

$$|u_1v_1 + u_2v_2 + \cdots + u_nv_n| \le (u_1^2 + u_2^2 + \cdots + u_n^2)^{1/2}(v_1^2 + v_2^2 + \cdots + v_n^2)^{1/2} \tag{23}$$

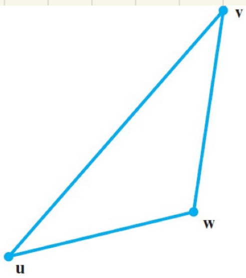
which makes sense, as

$$\underbrace{|u \cdot v| = \|u\| \; \|v\| \; \cos\theta}_{\text{always equal/smaller than } \|u\| \times \|v\|}, \qquad -1 \le \cos\theta \le 1$$



$\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$

▲ Figure 3.2.8



$d(\mathbf{u}, \mathbf{v}) \le d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$

▲ Figure 3.2.9

Proof (a)

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) = (\mathbf{u} \cdot \mathbf{u}) + 2(\mathbf{u} \cdot \mathbf{v}) + (\mathbf{v} \cdot \mathbf{v}) \\ &= \|\mathbf{u}\|^2 + 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \\ &\le \|\mathbf{u}\|^2 + 2|\mathbf{u} \cdot \mathbf{v}| + \|\mathbf{v}\|^2 && \text{Property of absolute value} \\ &\le \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 && \text{Cauchy–Schwarz inequality} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

This completes the proof since both sides of the inequality in part (*a*) are nonnegative.

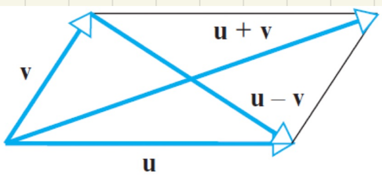
Proof (b) It follows from part (*a*) and Formula (11) that

$$\begin{aligned} d(\mathbf{u}, \mathbf{v}) &= \|\mathbf{u} - \mathbf{v}\| = \|(\mathbf{u} - \mathbf{w}) + (\mathbf{w} - \mathbf{v})\| \\ &\le \|\mathbf{u} - \mathbf{w}\| + \|\mathbf{w} - \mathbf{v}\| = d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v}) \end{aligned}$$ ◀

THEOREM 3.2.5 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* R^n, *then:*

(*a*) $\|\mathbf{u}+\mathbf{v}\| \le \|\mathbf{u}\| + \|\mathbf{v}\|$ **[Triangle inequality for vectors]**

(*b*) $d(\mathbf{u}, \mathbf{v}) \le d(\mathbf{u}, \mathbf{w}) + d(\mathbf{w}, \mathbf{v})$ **[Triangle inequality for distances]**



▲ Figure 3.2.10

THEOREM 3.2.6 Parallelogram Equation for Vectors

If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* R^n, *then*

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right) \tag{24}$$

Proof

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) + (\mathbf{u}-\mathbf{v}) \cdot (\mathbf{u}-\mathbf{v}) \\ &= 2(\mathbf{u} \cdot \mathbf{u}) + 2(\mathbf{v} \cdot \mathbf{v}) \\ &= 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right) \end{aligned}$$ ◀

THEOREM 3.2.6 Parallelogram Equation for Vectors

If $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* R^n, *then*

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\left(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\right) \tag{24}$$

THEOREM 3.2.7 *If* $\mathbf{u}$ *and* $\mathbf{v}$ *are vectors in* R^n *with the Euclidean inner product, then*

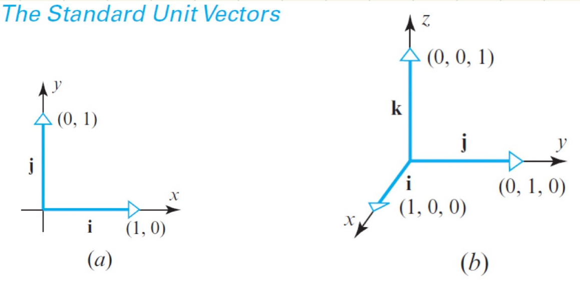
$$\mathbf{u} \cdot \mathbf{v} = \tfrac{1}{4}\|\mathbf{u}+\mathbf{v}\|^2 - \tfrac{1}{4}\|\mathbf{u}-\mathbf{v}\|^2 \tag{25}$$

Proof

$$\begin{aligned} \|\mathbf{u}+\mathbf{v}\|^2 &= (\mathbf{u}+\mathbf{v}) \cdot (\mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \\ \|\mathbf{u}-\mathbf{v}\|^2 &= (\mathbf{u}-\mathbf{v}) \cdot (\mathbf{u}-\mathbf{v}) = \|\mathbf{u}\|^2 - 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 \end{aligned}$$

from which (25) follows by simple algebra. ◀

The Standard Unit Vectors



▲ Figure 3.2.2

When a rectangular coordinate system is introduced in R^2 or R^3, the unit vectors in the positive directions of the coordinate axes are called the ***standard unit vectors***. In R^2 these vectors are denoted by

$$\mathbf{i} = (1, 0) \quad \text{and} \quad \mathbf{j} = (0, 1)$$

and in R^3 by

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \text{and} \quad \mathbf{k} = (0, 0, 1)$$

Example:

(a) Show that $\mathbf{u} = (-2, 3, 1, 4)$ and $\mathbf{v} = (1, 2, 0, -1)$ are orthogonal vectors in R^4.

(b) Let $S = \{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ be the set of standard unit vectors in R^3. Show that each ordered pair of vectors in S is orthogonal.

a) If u & v are orthogonal, $u \cdot v = 0$

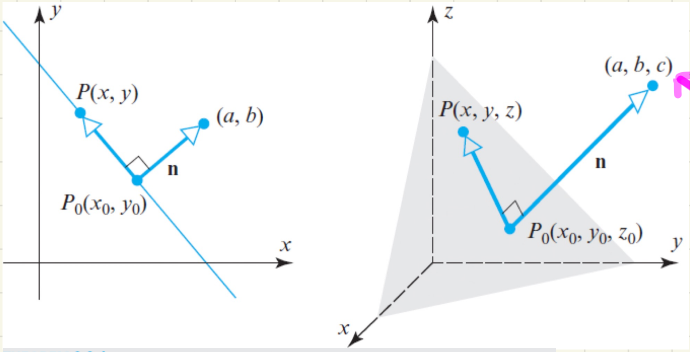
$$u \cdot v = \begin{bmatrix} -2 \\ 3 \\ 1 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 0 \\ -1 \end{bmatrix} = -2 + 6 + 0 - 4 = 0$$

b) Basically, show that $i \cdot j = i \cdot k = j \cdot k = 0$

$$i \cdot j = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 0$$

$$i \cdot k = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$

$$j \cdot k = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$$



$$\mathbf{n} \cdot \overrightarrow{P_0P} = 0$$

$$a(x - x_0) + b(y - y_0) = 0 \quad \textbf{[line]}$$

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0 \quad \textbf{[plane]}$$

▶ **EXAMPLE 2 Point-Normal Equations**

It follows from (2) that in R^2 the equation

$$6(x - 3) + (y + 7) = 0$$

line: $y = -6x + 11$

represents the line through the point $(3, -7)$ with normal $\mathbf{n} = (6, 1)$; and it follows from (3) that in R^3 the equation

Plane

$$4(x - 3) + 2y - 5(z - 7) = 0$$

represents the plane through the point $(3, 0, 7)$ with normal $\mathbf{n} = (4, 2, -5)$. ◀

THEOREM 3.3.1

(a) *If a and b are constants that are not both zero, then an equation of the form*

$$ax + by + c = 0 \tag{4}$$

represents a line in R^2 with normal $\mathbf{n} = (a, b)$.

(b) *If a, b, and c are constants that are not all zero, then an equation of the form*

$$ax + by + cz + d = 0 \tag{5}$$

represents a plane in R^3 with normal $\mathbf{n} = (a, b, c)$.

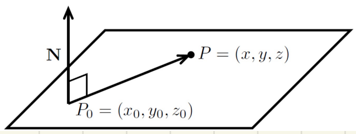
Extra Digram

$$\mathbf{N} \cdot \overrightarrow{\mathbf{P_0P}} = 0$$

$$\Leftrightarrow \quad \langle a, b, c \rangle \cdot \langle x - x_0, y - y_0, z - z_0 \rangle = 0$$

$$\Leftrightarrow \quad a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

We call this last equation the point-normal form for the plane.



Example

Example 1: Find the plane through the point (1,4,9) with normal $\langle 2, 3, 4 \rangle$.

Eqn of plane: $2(x-1) + 3(y-4) + 4(z-9) = 0$

alternatively: $2x + 3y + 4z = 50$

$W^{\perp}$

[redacted] If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W. The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^{\perp}$ (and read as "W perpendicular" or simply "W perp").

1. A vector $\mathbf{x}$ is in $W^{\perp}$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
2. $W^{\perp}$ is a subspace of $\mathbb{R}^n$.

Example of R^3: $W^{\perp}$

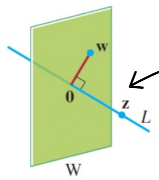


FIGURE 7

A plane and line through **0** as orthogonal complements.

For here, $L = W^{\perp}$

Can also consider $W = L^{\perp}$

Both are ⊥ to each other

Think of this like a vector that constructs a subspace where all $W \cdot W^{\perp} = 0$, no matter R^n.

✲ Orthogonal Matrix A

↳ $A^T = A^{-1}$

↳ $AA^{-1} = I \Rightarrow AA^T = I$

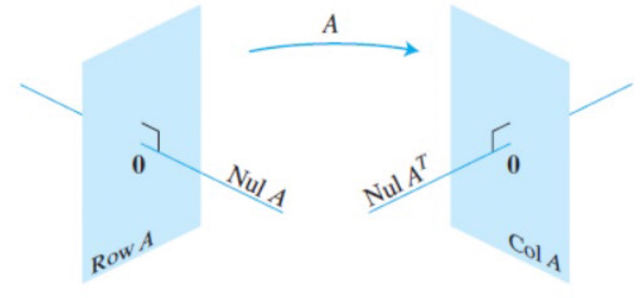


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A.

$\text{Nul } A = (\text{Row } A)^{\perp}$

$\text{Nul } A^{T} = (\text{Col } A)^{\perp}$

Orthogonal Projection

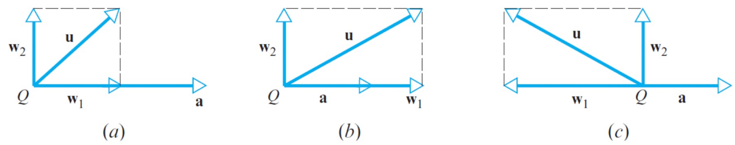


Figure 3.3.2 Three possible cases.

- Drop a perpendicular from the tip of $\mathbf{u}$ to the line through $\mathbf{a}$.
- Construct the vector $\mathbf{w}_1$ from Q to the foot of the perpendicular.
- Construct the vector $\mathbf{w}_2 = \mathbf{u} - \mathbf{w}_1$.

Since

$$\mathbf{w}_1 + \mathbf{w}_2 = \mathbf{w}_1 + (\mathbf{u} - \mathbf{w}_1) = \mathbf{u}$$

we have decomposed $\mathbf{u}$ into a sum of two orthogonal vectors, the first term being a scalar multiple of $\mathbf{a}$ and the second being orthogonal to $\mathbf{a}$.

▶ **EXAMPLE 1 The Standard Basis for R^n**

Recall from Example 11 of Section 4.2 that the standard unit vectors

$$\mathbf{e}_1 = (1, 0, 0, \ldots, 0), \quad \mathbf{e}_2 = (0, 1, 0, \ldots, 0), \ldots, \quad \mathbf{e}_n = (0, 0, 0, \ldots, 1)$$

span R^n and from Example 1 of Section 4.3 that they are linearly independent. Thus, they form a basis for R^n that we call the ***standard basis for R^n***. In particular,

$$\mathbf{i} = (1, 0, 0), \quad \mathbf{j} = (0, 1, 0), \quad \mathbf{k} = (0, 0, 1)$$

is the standard basis for R^3.

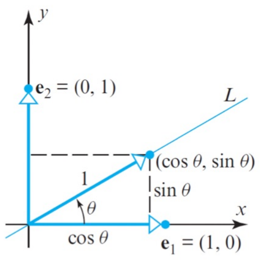
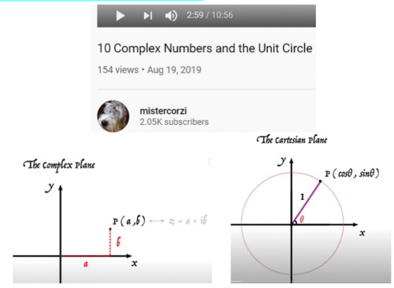


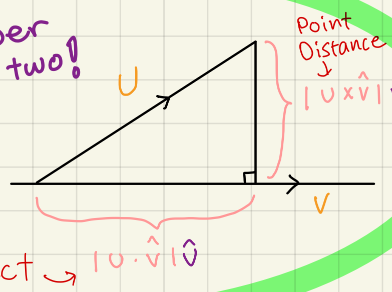
Figure 3.3.3

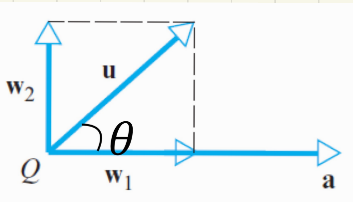


Mister Corzi: https://www.youtube.com/watch?v=UF172DZOiWo

Example above:
Commonly found in complex number of unit circle.
Where e_1 is the real line, and e_2 is the imaginary line

Remember these two!





$$w_1 = \text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \quad (\textit{vector component of } \mathbf{u} \textit{ along } \mathbf{a}) \quad = ||u||\cos\theta$$

$$w_2 = \mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u} = \mathbf{u} - \frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \quad (\textit{vector component of } \mathbf{u} \textit{ orthogonal to } \mathbf{a})$$

Example:

Let $\mathbf{u} = (2, -1, 3)$ and $\mathbf{a} = (4, -1, 2)$. Find the vector component of $\mathbf{u}$ along $\mathbf{a}$ and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$.

Solution

$$\mathbf{u}\cdot\mathbf{a} = (2)(4) + (-1)(-1) + (3)(2) = 15$$
$$\|\mathbf{a}\|^2 = 4^2 + (-1)^2 + 2^2 = 21$$

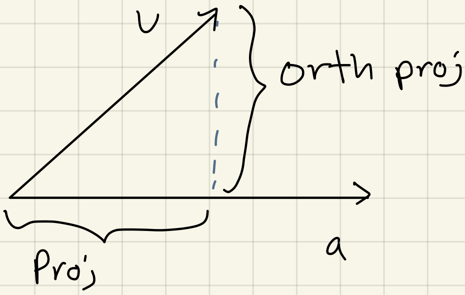
Thus the vector component of $\mathbf{u}$ along $\mathbf{a}$ is

$$\text{proj}_{\mathbf{a}}\mathbf{u} = \frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} = \tfrac{15}{21}(4, -1, 2) = \left(\tfrac{20}{7}, -\tfrac{5}{7}, \tfrac{10}{7}\right)$$

and the vector component of $\mathbf{u}$ orthogonal to $\mathbf{a}$ is

$$\mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u} = (2, -1, 3) - \left(\tfrac{20}{7}, -\tfrac{5}{7}, \tfrac{10}{7}\right) = \left(-\tfrac{6}{7}, -\tfrac{2}{7}, \tfrac{11}{7}\right)$$

As a check, you may wish to verify that the vectors $\mathbf{u} - \text{proj}_{\mathbf{a}}\mathbf{u}$ and $\mathbf{a}$ are perpendicular by showing that their dot product is zero. ◀



Norm of Projection

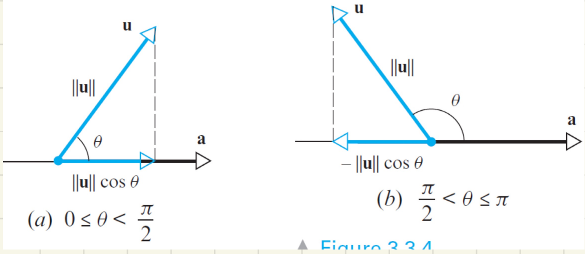


Figure 3.3.4

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \left\| \frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2}\mathbf{a} \right\| = \left| \frac{\mathbf{u}\cdot\mathbf{a}}{\|\mathbf{a}\|^2} \right| \|\mathbf{a}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|^2}\|\mathbf{a}\|$$

where the second equality follows from part (*c*) of Theorem 3.2.1 and the third from the fact that $\|\mathbf{a}\|^2 > 0$. Thus,

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|} \tag{12}$$

If θ denotes the angle between $\mathbf{u}$ and $\mathbf{a}$, then $\mathbf{u}\cdot\mathbf{a} = \|\mathbf{u}\|\|\mathbf{a}\|\cos\theta$, so (12) can also be written as

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \|\mathbf{u}\||\cos\theta| \tag{13}$$

$$\|\text{proj}_{\mathbf{a}}\mathbf{u}\| = \frac{|\mathbf{u}\cdot\mathbf{a}|}{\|\mathbf{a}\|} = |u \cdot \hat{a}|$$

Pythagoras

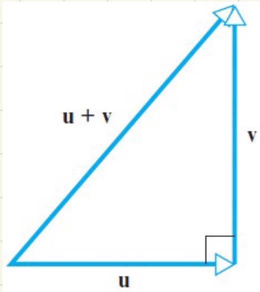
THEOREM 3.3.3 Theorem of Pythagoras in R^n

If $\mathbf{u}$ *and* $\mathbf{v}$ *are orthogonal vectors in* R^n *with the Euclidean inner product, then*

$$\|\mathbf{u}+\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 \tag{14}$$

Proof Since $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, we have $\mathbf{u}\cdot\mathbf{v} = 0$, from which it follows that

$$\|\mathbf{u}+\mathbf{v}\|^2 = (\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$$ ◀



Othorgonal Sets

A set of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ in $\mathbb{R}^n$ is said to be an **orthogonal set** if each pair of distinct vectors from the set is orthogonal, that is, if $\mathbf{u}_i \cdot \mathbf{u}_j = 0$ whenever $i \neq j$.

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

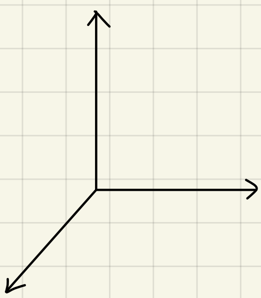
SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1 \cdot \mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$
$$\mathbf{u}_1 \cdot \mathbf{u}_3 = 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0$$
$$\mathbf{u}_2 \cdot \mathbf{u}_3 = -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Fig. 1; the three line segments there are mutually perpendicular. ■

☆ Standard basis is ortho set

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



Ortho Set is linearly independent

If $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.

PROOF If $\mathbf{0} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$ for some scalars $c_1, \ldots, c_p$, then

$$\begin{aligned} 0 = \mathbf{0}\cdot\mathbf{u}_1 &= (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= (c_1\mathbf{u}_1)\cdot\mathbf{u}_1 + (c_2\mathbf{u}_2)\cdot\mathbf{u}_1 + \cdots + (c_p\mathbf{u}_p)\cdot\mathbf{u}_1 \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) + c_2(\mathbf{u}_2\cdot\mathbf{u}_1) + \cdots + c_p(\mathbf{u}_p\cdot\mathbf{u}_1) \\ &= c_1(\mathbf{u}_1\cdot\mathbf{u}_1) \end{aligned}$$

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_2, \ldots, \mathbf{u}_p$. Since $\mathbf{u}_1$ is nonzero, $\mathbf{u}_1 \cdot \mathbf{u}_1$ is not zero and so $c_1 = 0$. Similarly, $c_2, \ldots, c_p$ must be zero. Thus S is linearly independent. ■

☆ 3 steps!

1) Zero Vector
2) Addition
3) Scalar multiplication

Vector On Subspace

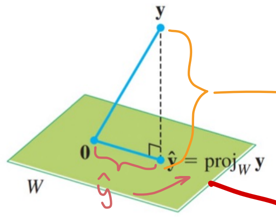


FIGURE 1

vector u

→ Here, y is a point in R^n

→ Projecting on a subspace W Produces vector $\hat{y}$

→ Vertical Ortho projection will be $\boxed{y - \hat{y}}$

Given a vector $\mathbf{y}$ and a subspace W in $\mathbb{R}^n$, there is a vector $\hat{\mathbf{y}}$ in W such that (1) $\hat{\mathbf{y}}$ is the unique vector in W for which $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to W, and (2) $\hat{\mathbf{y}}$ is the unique vector in W closest to $\mathbf{y}$. See Figure 1. These two properties of $\hat{\mathbf{y}}$ provide the key to finding least-squares solutions of linear systems, mentioned in the introductory example for this chapter.

To prepare for the first theorem, observe that whenever a vector $\mathbf{y}$ is written as a linear combination of vectors $\mathbf{u}_1, \ldots, \mathbf{u}_n$ in $\mathbb{R}^n$, the terms in the sum for $\mathbf{y}$ can be grouped into two parts so that $\mathbf{y}$ can be written as

$$\mathbf{y} = \mathbf{z}_1 + \mathbf{z}_2$$

where $\mathbf{z}_1$ is a linear combination of some of the $\mathbf{u}_i$ and $\mathbf{z}_2$ is a linear combination of the rest of the $\mathbf{u}_i$. This idea is particularly useful when $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ is an orthogonal basis.

Hack!

$$\hat{y} = |y \cdot \hat{u}|\, \hat{u}$$

Example

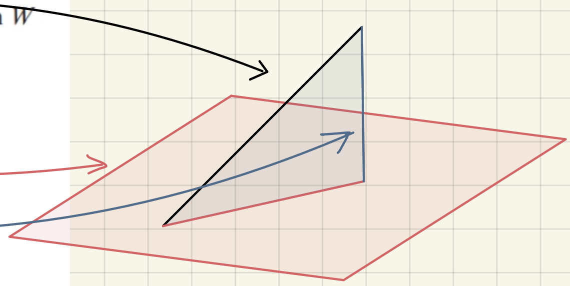
EXAMPLE 1 Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_5\}$ be an orthogonal basis for $\mathbb{R}^5$ and let

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_5\mathbf{u}_5$$

Consider the subspace $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, and write $\mathbf{y}$ as the sum of a vector $\mathbf{z}_1$ in W and a vector $\mathbf{z}_2$ in $W^\perp$.

SOLUTION Write

$$\mathbf{y} = \underbrace{c_1\mathbf{u}_1 + c_2\mathbf{u}_2}_{\mathbf{z}_1} + \underbrace{c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5}_{\mathbf{z}_2}$$



where $\quad \mathbf{z}_1 = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \quad$ is in Span $\{\mathbf{u}_1, \mathbf{u}_2\}$

and $\quad \mathbf{z}_2 = c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5 \quad$ is in Span $\{\mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5\}$.

To show that $\mathbf{z}_2$ is in $W^\perp$, it suffices to show that $\mathbf{z}_2$ is orthogonal to the vectors in the basis $\{\mathbf{u}_1, \mathbf{u}_2\}$ for W. (See Section 6.1.) Using properties of the inner product, compute

$$\begin{aligned}
\mathbf{z}_2 \cdot \mathbf{u}_1 &= (c_3\mathbf{u}_3 + c_4\mathbf{u}_4 + c_5\mathbf{u}_5) \cdot \mathbf{u}_1 \\
&= c_3\mathbf{u}_3 \cdot \mathbf{u}_1 + c_4\mathbf{u}_4 \cdot \mathbf{u}_1 + c_5\mathbf{u}_5 \cdot \mathbf{u}_1 \\
&= 0
\end{aligned}$$

z_2 is always $\perp$ to z_1 as they are originally perpendicular

because $\mathbf{u}_1$ is orthogonal to $\mathbf{u}_3$, $\mathbf{u}_4$, and $\mathbf{u}_5$. A similar calculation shows that $\mathbf{z}_2 \cdot \mathbf{u}_2 = 0$. Thus $\mathbf{z}_2$ is in $W^\perp$. ■

Othorgonial Basis

The next theorem suggests why an orthogonal basis is much nicer than other bases. The weights in a linear combination can be computed easily.

– How to find the coefficient of any U...

THEOREM 5

Let $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ be an orthogonal basis for a subspace W of $\mathbb{R}^n$. For each $\mathbf{y}$ in W, the weights in the linear combination

$$\mathbf{y} = c_1\mathbf{u}_1 + \cdots + c_p\mathbf{u}_p$$

are given by

$$\boxed{c_j = \frac{\mathbf{y}\cdot\mathbf{u}_j}{\mathbf{u}_j\cdot\mathbf{u}_j} \qquad (j = 1, \dots, p)}$$

– Because $y = c_1 u_1 + \dots + c_p u_p$

$$y \cdot u_1 = (c_1 u_1 + \dots + c_p u_p) \cdot u_1$$
$$y \cdot u_1 = c_1 u_1 \cdot u_1 + \dots + c_p u_p \cdot u_1$$
$$y \cdot u_1 = c_1 (u_1 \cdot u_1) + \dots + c_p (0)$$
$$y \cdot u_1 = c_1 (u_1 \cdot u_1)$$
$$c_1 = \frac{y \cdot u_1}{u_1 \cdot u_1}$$

PROOF As in the preceding proof, the orthogonality of $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ shows that

$$\mathbf{y}\cdot\mathbf{u}_1 = (c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_p\mathbf{u}_p)\cdot\mathbf{u}_1 = c_1(\mathbf{u}_1\cdot\mathbf{u}_1)$$

Since $\mathbf{u}_1\cdot\mathbf{u}_1$ is not zero, the equation above can be solved for c_1. To find c_j for $j = 2, \dots, p$, compute $\mathbf{y}\cdot\mathbf{u}_j$ and solve for c_j. ■

AlSO, easily solve

The Orthogonal Decomposition Theorem

Let W be a subspace of $\mathbb{R}^n$. Then each $\mathbf{y}$ in $\mathbb{R}^n$ can be written uniquely in the form

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \tag{1}$$

where $\hat{\mathbf{y}}$ is in W and $\mathbf{z}$ is in $W^\perp$. In fact, if $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is any orthogonal basis of W, then

$$\boxed{\hat{\mathbf{y}} = \underbrace{\frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}}_{\text{first}}\mathbf{u}_1 + \cdots + \underbrace{\frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}}_{\text{last}}\mathbf{u}_p} \tag{2}$$

and $\mathbf{z} = \mathbf{y} - \hat{\mathbf{y}}$.

– Similairly, Can apply above two theory to solve the projection Q! easily

Example : EASY

$$S = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Decompose $y = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ using the standard basis.

$$\begin{bmatrix} 6 \\ 1 \\ 8 \end{bmatrix} = 6\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 1\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + (-8)\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

The vector $\hat{\mathbf{y}}$ in (1) is called the **orthogonal projection of y onto** W and often is written as $\text{proj}_W\,\mathbf{y}$. See Figure 2. When W is a one-dimensional subspace, the formula for $\hat{\mathbf{y}}$ matches the formula given in Section 6.2.

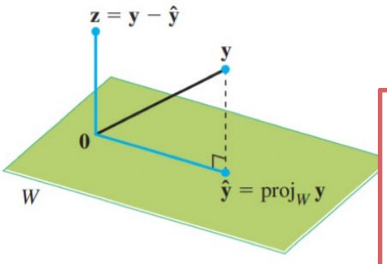


FIGURE 2 The orthogonal projection of $\mathbf{y}$ onto W.

Example : Hard

EXAMPLE 1 Show that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set, where

$$\mathbf{u}_1 = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -1/2 \\ -2 \\ 7/2 \end{bmatrix}$$

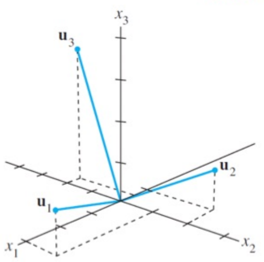


FIGURE 1

SOLUTION Consider the three possible pairs of distinct vectors, namely, $\{\mathbf{u}_1, \mathbf{u}_2\}$, $\{\mathbf{u}_1, \mathbf{u}_3\}$, and $\{\mathbf{u}_2, \mathbf{u}_3\}$.

$$\mathbf{u}_1\cdot\mathbf{u}_2 = 3(-1) + 1(2) + 1(1) = 0$$
$$\mathbf{u}_1\cdot\mathbf{u}_3 = 3\left(-\tfrac{1}{2}\right) + 1(-2) + 1\left(\tfrac{7}{2}\right) = 0$$
$$\mathbf{u}_2\cdot\mathbf{u}_3 = -1\left(-\tfrac{1}{2}\right) + 2(-2) + 1\left(\tfrac{7}{2}\right) = 0$$

Each pair of distinct vectors is orthogonal, and so $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthogonal set. See Figure 1; the three line segments there are mutually perpendicular. ■

EXAMPLE 2 The set $S = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ in Example 1 is an orthogonal basis for $\mathbb{R}^3$. Express the vector $\mathbf{y} = \begin{bmatrix} 6 \\ 1 \\ -8 \end{bmatrix}$ as a linear combination of the vectors in S.

SOLUTION Compute

$$\mathbf{y}\cdot\mathbf{u}_1 = 11, \quad \mathbf{y}\cdot\mathbf{u}_2 = -12, \quad \mathbf{y}\cdot\mathbf{u}_3 = -33$$
$$\mathbf{u}_1\cdot\mathbf{u}_1 = 11, \quad \mathbf{u}_2\cdot\mathbf{u}_2 = 6, \quad \mathbf{u}_3\cdot\mathbf{u}_3 = 33/2$$

By Theorem 5,

$$\begin{aligned}
\mathbf{y} &= \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2 + \frac{\mathbf{y}\cdot\mathbf{u}_3}{\mathbf{u}_3\cdot\mathbf{u}_3}\mathbf{u}_3 \\
&= \frac{11}{11}\mathbf{u}_1 + \frac{-12}{6}\mathbf{u}_2 + \frac{-33}{33/2}\mathbf{u}_3 \\
&= \mathbf{u}_1 - 2\mathbf{u}_2 - 2\mathbf{u}_3
\end{aligned}$$

■

Notice how easy it is to compute the weights needed to build $\mathbf{y}$ from an orthogonal basis. If the basis were not orthogonal, it would be necessary to solve a system of linear equations in order to find the weights, as in Chapter 1.

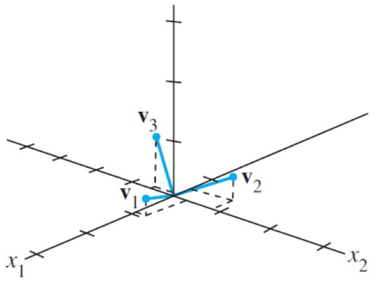
Othonormal Set

Orthonormal Sets

A set $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal set** if it is an orthogonal set of unit vectors. If W is the subspace spanned by such a set, then $\{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an **orthonormal basis** for W, since the set is automatically linearly independent, by Theorem 4.

The simplest example of an orthonormal set is the standard basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for $\mathbb{R}^n$. Any nonempty subset of $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is orthonormal, too. Here is a more complicated example.

THEOREM 4 If $S = \{\mathbf{u}_1, \dots, \mathbf{u}_p\}$ is an orthogonal set of nonzero vectors in $\mathbb{R}^n$, then S is linearly independent and hence is a basis for the subspace spanned by S.



EXAMPLE 5 Show that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal basis of $\mathbb{R}^3$, where

$$\mathbf{v}_1 = \begin{bmatrix} 3/\sqrt{11} \\ 1/\sqrt{11} \\ 1/\sqrt{11} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} -1/\sqrt{66} \\ -4/\sqrt{66} \\ 7/\sqrt{66} \end{bmatrix}$$

SOLUTION Compute

$$\mathbf{v}_1 \cdot \mathbf{v}_2 = -3/\sqrt{66} + 2/\sqrt{66} + 1/\sqrt{66} = 0$$
$$\mathbf{v}_1 \cdot \mathbf{v}_3 = -3/\sqrt{726} - 4/\sqrt{726} + 7/\sqrt{726} = 0$$
$$\mathbf{v}_2 \cdot \mathbf{v}_3 = 1/\sqrt{396} - 8/\sqrt{396} + 7/\sqrt{396} = 0$$

Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal set. Also,

$\|v\|^2 = 1$

$\therefore \|v\| = 1$

$$\mathbf{v}_1 \cdot \mathbf{v}_1 = 9/11 + 1/11 + 1/11 = 1$$
$$\mathbf{v}_2 \cdot \mathbf{v}_2 = 1/6 + 4/6 + 1/6 = 1$$
$$\mathbf{v}_3 \cdot \mathbf{v}_3 = 1/66 + 16/66 + 49/66 = 1$$

which shows that $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$ are unit vectors. Thus $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthonormal set. Since the set is linearly independent, its three vectors form a basis for $\mathbb{R}^3$. See Fig. 6. ■

An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.

$\Rightarrow$ makes sense as $U_i \cdot U_j = 0$, $U_i \cdot U_i = 1$

$$\therefore \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

PROOF To simplify notation, we suppose that U has only three columns, each a vector in $\mathbb{R}^m$. The proof of the general case is essentially the same. Let $U = [\,\mathbf{u}_1 \ \ \mathbf{u}_2 \ \ \mathbf{u}_3\,]$ and compute

$$U^TU = \begin{bmatrix} \mathbf{u}_1^T \\ \mathbf{u}_2^T \\ \mathbf{u}_3^T \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \mathbf{u}_1^T\mathbf{u}_1 & \mathbf{u}_1^T\mathbf{u}_2 & \mathbf{u}_1^T\mathbf{u}_3 \\ \mathbf{u}_2^T\mathbf{u}_1 & \mathbf{u}_2^T\mathbf{u}_2 & \mathbf{u}_2^T\mathbf{u}_3 \\ \mathbf{u}_3^T\mathbf{u}_1 & \mathbf{u}_3^T\mathbf{u}_2 & \mathbf{u}_3^T\mathbf{u}_3 \end{bmatrix} \tag{4}$$

The entries in the matrix at the right are inner products, using transpose notation. The columns of U are orthogonal if and only if

$$\mathbf{u}_1^T\mathbf{u}_2 = \mathbf{u}_2^T\mathbf{u}_1 = 0, \quad \mathbf{u}_1^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_1 = 0, \quad \mathbf{u}_2^T\mathbf{u}_3 = \mathbf{u}_3^T\mathbf{u}_2 = 0 \tag{5}$$

The columns of U all have unit length if and only if

$$\mathbf{u}_1^T\mathbf{u}_1 = 1, \quad \mathbf{u}_2^T\mathbf{u}_2 = 1, \quad \mathbf{u}_3^T\mathbf{u}_3 = 1 \tag{6}$$

The theorem follows immediately from (4)–(6). ■

Let U be an $m \times n$ matrix with orthonormal columns, and let $\mathbf{x}$ and $\mathbf{y}$ be in $\mathbb{R}^n$. Then

Preserve length

Preserve dot product

a. $\|U\mathbf{x}\| = \|\mathbf{x}\|$
b. $(U\mathbf{x})\cdot(U\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$
c. $(U\mathbf{x})\cdot(U\mathbf{y}) = 0$ if and only if $\mathbf{x}\cdot\mathbf{y} = 0$

Properties (a) and (c) say that the linear mapping $\mathbf{x} \mapsto U\mathbf{x}$ preserves lengths and orthogonality. These properties are crucial for many computer algorithms.

Important Theorem

EXAMPLE 6 Let $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ and $\mathbf{x} = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix}$. Notice that U has orthonormal columns and

$$U^TU = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 2/3 & -2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Verify that $\|U\mathbf{x}\| = \|\mathbf{x}\|$.

SOLUTION

$$U\mathbf{x} = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix}$$

$$\|U\mathbf{x}\| = \sqrt{9+1+1} = \sqrt{11}$$

$$\|\mathbf{x}\| = \sqrt{2+9} = \sqrt{11}$$ ■

Let $y = \begin{bmatrix} -\sqrt{2} \\ 6 \end{bmatrix}$, $Uy = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix} \begin{bmatrix} -\sqrt{2} \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$

$x \cdot y = \begin{bmatrix} \sqrt{2} \\ 3 \end{bmatrix} \cdot \begin{bmatrix} -\sqrt{2} \\ 6 \end{bmatrix}$
$= -2 + 18$
$= 16$

$Ux \cdot Uy = \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ -5 \\ 2 \end{bmatrix}$
$= 9 + 5 + 2$
$= 16$

Orthogonal Matrix definitions

Ref: https://en.wikipedia.org/wiki/Orthogonal_matrix **Important!**

In linear algebra, an **orthogonal matrix**, or **orthonormal matrix**, is a real square matrix whose columns and rows are orthonormal vectors.

One way to express this is

$$Q^TQ = QQ^T = I,$$

where Q^T is the transpose of Q and I is the identity matrix.

This leads to the equivalent characterization: a matrix Q is orthogonal if its transpose is equal to its inverse:

$$Q^T = Q^{-1},$$

where Q^{-1} is the inverse of Q.

Orthogonal Matrix MUST be SQUARE!

Note: confusion
If you have a mxn matrix called U *with* its column ortho-normal, and m>n (tall matrix)

1) IT IS NOT an orthogonal matrix since it satisfy ONLY $U^TU = I$ (nxn)
2) its UU^T is mxm matrix BUT it is not equals to I) .Instead UU^T is a projection matrix and has rank n.

Note: There is no standard name for "rectangular matrix with orthonormal columns"

Example:

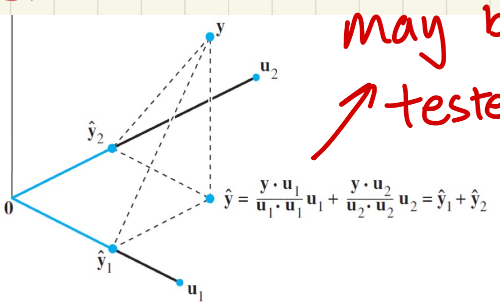


FIGURE 3 The orthogonal projection of $\mathbf{y}$ is the sum of its projections onto one-dimensional subspaces that are mutually orthogonal.

may be tested ☆

Figure 3 W is a subspace of $\mathbb{R}^3$ spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$. Here $\hat{\mathbf{y}}_1$ and $\hat{\mathbf{y}}_2$ denote the projections of $\mathbf{y}$ onto the lines spanned by $\mathbf{u}_1$ and $\mathbf{u}_2$, respectively. The orthogonal projection $\hat{\mathbf{y}}$ of $\mathbf{y}$ onto W is the sum of the projections of $\mathbf{y}$ onto one-dimensional subspaces that are orthogonal to each other.

EXAMPLE 2 Let $\mathbf{u}_1 = \begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Observe that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Write $\mathbf{y}$ as the sum of a vector in W and a vector orthogonal to W.

SOLUTION The orthogonal projection of $\mathbf{y}$ onto W is

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \frac{\mathbf{y}\cdot\mathbf{u}_2}{\mathbf{u}_2\cdot\mathbf{u}_2}\mathbf{u}_2$$

$$= \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{3}{6}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \frac{9}{30}\begin{bmatrix} 2 \\ 5 \\ -1 \end{bmatrix} + \frac{15}{30}\begin{bmatrix} -2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix}$$

Also

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} = \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

Theorem 8 ensures that $\mathbf{y} - \hat{\mathbf{y}}$ is in $W^\perp$. To check the calculations, however, it is a good idea to verify that $\mathbf{y} - \hat{\mathbf{y}}$ is orthogonal to both $\mathbf{u}_1$ and $\mathbf{u}_2$ and hence to all of W. The desired decomposition of $\mathbf{y}$ is

$$\mathbf{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2/5 \\ 2 \\ 1/5 \end{bmatrix} + \begin{bmatrix} 7/5 \\ 0 \\ 14/5 \end{bmatrix}$$

■

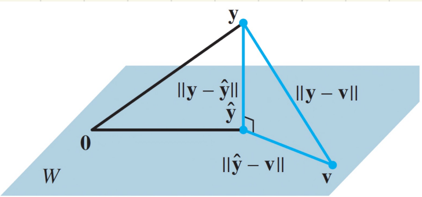


FIGURE 4 The orthogonal projection of $\mathbf{y}$ onto W is the closest point in W to $\mathbf{y}$.

The Best Approximation Theorem

Let W be a subspace of $\mathbb{R}^n$, let $\mathbf{y}$ be any vector in $\mathbb{R}^n$, and let $\hat{\mathbf{y}}$ be the orthogonal projection of $\mathbf{y}$ onto W. Then $\hat{\mathbf{y}}$ is the closest point in W to $\mathbf{y}$, in the sense that

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\| \quad (3)$$

for all $\mathbf{v}$ in W distinct from $\hat{\mathbf{y}}$.

EXAMPLE 4 The distance from a point $\mathbf{y}$ in $\mathbb{R}^n$ to a subspace W is defined as the distance from $\mathbf{y}$ to the nearest point in W. Find the distance from $\mathbf{y}$ to $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$, where

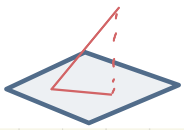
$$\mathbf{y} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix}, \quad \mathbf{u}_1 = \begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$$

SOLUTION By the Best Approximation Theorem, the distance from $\mathbf{y}$ to W is $\|\mathbf{y} - \hat{\mathbf{y}}\|$, where $\hat{\mathbf{y}} = \text{proj}_W\,\mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthogonal basis for W,

$$\hat{\mathbf{y}} = \frac{15}{30}\mathbf{u}_1 + \frac{-21}{6}\mathbf{u}_2 = \frac{1}{2}\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} - \frac{7}{2}\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix}$$

$$\mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} -1 \\ -5 \\ 10 \end{bmatrix} - \begin{bmatrix} -1 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 6 \end{bmatrix}$$

$$\|\mathbf{y} - \hat{\mathbf{y}}\|^2 = 3^2 + 6^2 = 45$$



The distance from $\mathbf{y}$ to W is $\sqrt{45} = 3\sqrt{5}$. ■

Here, Orthonomal basis $\|U\| = 1$

THEOREM 10

If $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ is an orthonormal basis for a subspace W of $\mathbb{R}^n$, then

$$\text{proj}_W\,\mathbf{y} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p \quad (4)$$

If $U = [\,\mathbf{u}_1 \;\; \mathbf{u}_2 \;\; \cdots \;\; \mathbf{u}_p\,]$, then

$$\text{proj}_W\,\mathbf{y} = UU^T\mathbf{y} \quad \text{for all } \mathbf{y} \text{ in } \mathbb{R}^n \quad (5)$$

Note : $p < n$

Let $\mathbf{u}_1 = \begin{bmatrix} -7 \\ 1 \\ 4 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} -1 \\ 1 \\ -2 \end{bmatrix}$, $\mathbf{y} = \begin{bmatrix} -9 \\ 1 \\ 6 \end{bmatrix}$, and $W = \text{Span}\{\mathbf{u}_1, \mathbf{u}_2\}$. Use the fact that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal to compute $\text{proj}_W\,\mathbf{y}$.

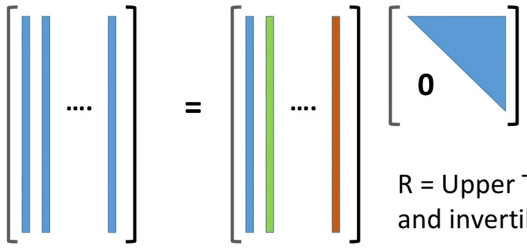
Note: above u_1, and u_2 are orthogonal, but their norm is not 1. You need to convert them to norm 1 and use Theorem 10 (orthonormality is needed). Else Use theorem 8 (only orthogonality is needed).

QR Factorization

The QR Factorization

If A is an $m \times n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where Q is an $m \times n$ matrix whose columns form an orthonormal basis for Col A and R is an $n \times n$ upper triangular invertible matrix with positive entries on its diagonal.

$$A = [\,\mathbf{x}_1 \quad \cdots \quad \mathbf{x}_n\,] = [\,Q\mathbf{r}_1 \quad \cdots \quad Q\mathbf{r}_n\,] = QR$$



R = Upper Triangle and invertible

A $\in R^{m*n}$ Q = columns are orthonormal

R is a square upper triangle matrix and Depending if

- A has independent col, then R is invertible,
- A has dependent col, then R is NOT-invertible.

It has many applications, e.g, solving least squares:

$$\boxed{Ax} = y$$
$$QRx = y$$
$$Q^TQRx = Q^Ty$$
$$Rx = Q^Ty$$

as long as linearly indep.

Is can be easily solved because R is upper triangle
(If R is not invertible, it will be more involved, see least squares chapter)

QR decomposition, given A (mxn) matrix, we can decompose it into the product of 2 matrixes,

$$A = QR$$

- Properties of Q:
  - $C(Q) == C(A)$
  - $Q^TQ = I$, i.e Q has orthonormal column (BUT not necessary square)
  - QQ^T = projection matrix into col (A)
- R is a square upper triangle matrix and Depending if
  - A has independent col, then R is invertible,
  - A has dependent col, then R is NOT-invertible.

- A is a mxn matrix, **assume columns are** Not orthogonal
- $\boldsymbol{v_i} \in \boldsymbol{R^m}, i = 1..\boldsymbol{n}$

- The columns of Q are **Orthonormal $\boldsymbol{u_i}$** $i = 1..n$
- **And spans the same space as A**

$$A = \{v_1, v_2, v_3, \cdots, v_n\} \rightarrow \boxed{\text{Gram-Schmidt Process}} \rightarrow Q = \{u_1, u_2, u_3, \cdots, u_n\}$$

$$Span(A) = Span(Q)$$

The Gram–Schmidt Process

Given a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$\mathbf{v}_1 = \mathbf{x}_1$$
$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$
$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$
$$\vdots$$
$$\mathbf{v}_p = \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_p \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 - \cdots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W. In addition

$$\text{Span}\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \text{Span}\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \qquad \text{for } 1 \le k \le p$$

$$A = QR$$

where:

$$Q = [u_1, u_2, u_3, \ldots, u_n]$$

and

$$R = \begin{pmatrix} u_1.x_1 & u_1.x_2 & u_1.x_3 & \cdots \\ 0 & u_2.x_2 & u_2.x_3 & \cdots \\ 0 & 0 & u_3.x_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

EXAMPLE 2 Let $\mathbf{x}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of $\mathbb{R}^4$. Construct an orthogonal basis for W.

SOLUTION

Step 1. Let $\mathbf{v}_1 = \mathbf{x}_1$ and $W_1 = \text{Span}\{\mathbf{x}_1\} = \text{Span}\{\mathbf{v}_1\}$.

Step 2. Let $\mathbf{v}_2$ be the vector produced by subtracting from $\mathbf{x}_2$ its projection onto the subspace W_1. That is, let

$$\mathbf{v}_2 = \mathbf{x}_2 - \text{proj}_{W_1}\mathbf{x}_2$$
$$= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Since } \mathbf{v}_1 = \mathbf{x}_1$$
$$= \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \left.\begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix}\right\} \times 4$$

As in Example 1, $\mathbf{v}_2$ is the component of $\mathbf{x}_2$ orthogonal to $\mathbf{x}_1$, and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for the subspace W_2 spanned by $\mathbf{x}_1$ and $\mathbf{x}_2$.

Note: $v_2' = v_2 * 4$ to get rid of denominator in v_2

Step 3. Let $\mathbf{v}_3$ be the vector produced by subtracting from $\mathbf{x}_3$ its projection onto the subspace W_2. Use the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2'\}$ to compute this projection onto W_2:

Projection of $\mathbf{x}_3$ onto $\mathbf{v}_1$ — Projection of $\mathbf{x}_3$ onto $\mathbf{v}_2'$

$$\text{proj}_{W_2}\mathbf{x}_3 = \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{x}_3 \cdot \mathbf{v}_2'}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\mathbf{v}_2' = \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} + \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix}$$

Then $\mathbf{v}_3$ is the component of $\mathbf{x}_3$ orthogonal to W_2, namely,

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proj}_{W_2}\mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \begin{bmatrix}0\\2/3\\2/3\\2/3\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

See Slide 3 of Chapter 6.2.5 for explanation.

Example: Gram–Schmidt on a 3x3 matrix

Example [edit]

Consider the decomposition of

$$A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix}.$$

Recall that an orthonormal matrix Q has the property

$$Q^\mathsf{T} Q = I.$$

Then, we can calculate Q by means of Gram–Schmidt as follows:

$$U = (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3) = \begin{pmatrix} 12 & -69 & -58/5 \\ 6 & 158 & 6/5 \\ -4 & 30 & -33 \end{pmatrix};$$

$$Q = \begin{pmatrix} \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} & \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} & \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \end{pmatrix} = \begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix}$$

Thus, we have

$$Q^\mathsf{T} A = Q^\mathsf{T} Q\, R = R;$$

$$R = Q^\mathsf{T} A = \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$$

Hence,

$$A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix} =$$

$$\begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix} \times \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$$

Q: Orthogonal Matrix $\quad$ R: Upper Triangular Matrix

Least Square Solution

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

a. The equation $A\mathbf{x} = \mathbf{b}$ has a unique least-squares solution for each $\mathbf{b}$ in $\mathbb{R}^m$.

b. The columns of A are linearly indpendent.

c. The matrix A^TA is invertible.

When these statements are true, the least-squares solution $\hat{\mathbf{x}}$ is given by

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b} \tag{4}$$

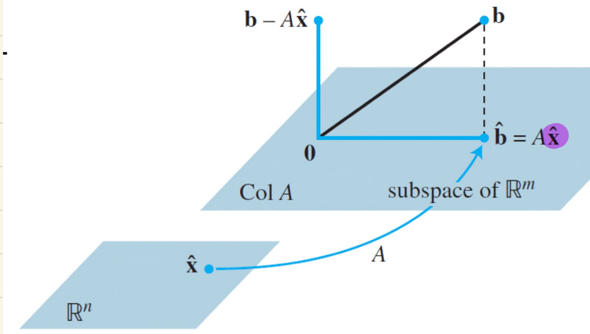


FIGURE 2 The least-squares solution $\hat{\mathbf{x}}$ is in $\mathbb{R}^n$.

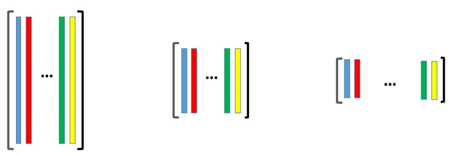
Therefore if $\quad \hat{x} = (A^TA)^{-1}A^Tb, \quad$ then

$\hat{b} = A\hat{x}$, means

$\hat{b} = A(A^TA)^{-1}A^T\ b$

And this matrix $A(A^TA)^{-1}A^T$ is a projection matrix, projecting b into the column space of A.

Based on M & N, there exist three cases:



$M \gg N$ $\quad$ $M \approx N$ $\quad$ $M \ll N$

More equations, less unknowns.

Hence, **over-determined**! **Typically** this will result in inconsistent system of equations

Less equations, more unknowns.

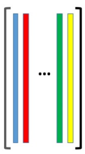
Hence, **under-determined**! **Typically,** this will result in infinite solutions

$$Ax = b$$

Consistent (solutions exist)

- b is in column space of A, i.e,
  b is formed by linear combinations of A's columns.
- Rank (A) = Rank ($A|b$), i.e,
  rank of A is same as that of the augmented matrix.

Inconsistent (solutions don't exist)

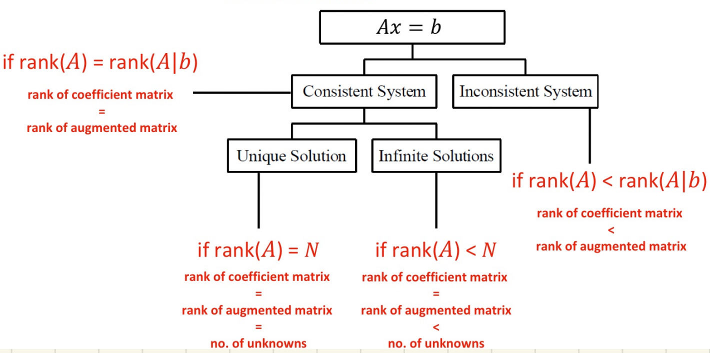


$M \gg N$

- b is NOT in column space of A, i.e,
  b is NOT formed by linear combinations of A's columns.
- **Typically** occurs when $M \gg N$ (**over-determined**), i.e,
  there exist more equations than unknowns.
- The rows of A are dependent but,
  their corresponding b values are not consistent.
- Rank (A) < Rank ($A|b$), i.e,
  rank of A is less than that of the augmented matrix.

A system of equations can be consistent or inconsistent. What does that mean?

A system of equations $Ax = b$ is consistent if there is a solution, and it is inconsistent if there is no solution. However, consistent system of equations does not mean a unique solution, that is, a consistent system of equation may have a unique solution or infinite solutions.



Examples of Rank Calculation

rank(A) = rank($A|b$) = N

Consistent and Unique Solution

a) The system of equations

$$\begin{bmatrix} 2 & 4 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

is a consistent system of equations as it has a unique solution, that is,

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

```
>> A_b = [ 2 4 6; 1 3 4]

A_b =

     2     4     6
     1     3     4

>> rank(A_b)

ans =

     2
```

Consistent and Having Infinite Solutions

b) The system of equations

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix}$$

is also a consistent system of equations but it has infinite solutions as given as follows.

Expanding the above set of equations,

$$2x + 4y = 6$$
$$x + 2y = 3$$

you can see that they are the same equation. Hence any combination of (x, y) that satisfies

$$2x + 4y = 6$$

is a solution. For example $(x, y) = (1,1)$ is a solution and other solutions include $(x, y) = (0.5, 1.25)$, $(x, y) = (0,\ 1.5)$ and so on.

```
>> A_b = [ 2 4 6; 1 2 3]

A_b =

     2     4     6
     1     2     3

>> rank(A_b)

ans =

     1
```

rank(A) = rank($A|b$) < N

Inconsistent and No solutions Exist

c) The system of equations

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$$

```
>> A = [2 4; 1 2]

A =

     2     4
     1     2

>> rank(A)

ans =

     1
```

```
>> A_b = [ 2 4 6; 1 2 4]

A_b =

     2     4     6
     1     2     4

>> rank(A_b)

ans =

     2
```

rank(A) < rank($A|b$)

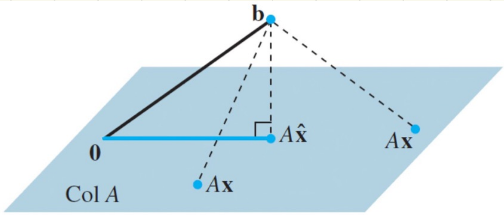


FIGURE 1 The vector $\mathbf{b}$ is closer to $A\hat{\mathbf{x}}$ than to $A\mathbf{x}$ for other $\mathbf{x}$.

The most important aspect of the least-squares problem is that no matter what $\mathbf{x}$ we select, the vector $A\mathbf{x}$ will necessarily be in the column space, Col A. So we seek an $\mathbf{x}$ that makes $A\mathbf{x}$ the closest point in Col A to $\mathbf{b}$. See Fig. 1. (Of course, if $\mathbf{b}$ happens to be in Col A, then $\mathbf{b}$ *is* $A\mathbf{x}$ for some $\mathbf{x}$, and such an $\mathbf{x}$ is a "least-squares solution.")

If a linear system is consistent, then its exact solutions are the same as its least squares solutions, in which case the least squares error is zero.

NOTE:
When the linear system $Ax = b$ is inconsistent, b does not lie in the column space of A.

To explain the terminology in this problem, suppose that the column form of $\mathbf{b} - A\mathbf{x}$ is

$$\mathbf{b} - A\mathbf{x} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}$$

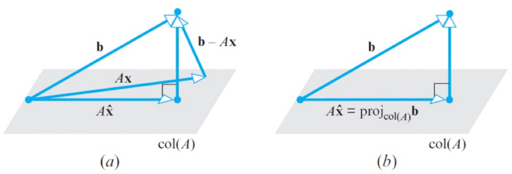
The term "least squares solution" results from the fact that minimizing $\|\mathbf{b} - A\mathbf{x}\|$ also has the effect of minimizing $\|\mathbf{b} - A\mathbf{x}\|^2 = e_1^2 + e_2^2 + \cdots + e_m^2$.

If A is $m \times n$ and $\mathbf{b}$ is in $\mathbb{R}^m$, a **least-squares solution** of $A\mathbf{x} = \mathbf{b}$ is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}\|$$

for all $\mathbf{x}$ in $\mathbb{R}^n$.

Think of $A\mathbf{x}$ as an *approximation* to $\mathbf{b}$. The smaller the distance between $\mathbf{b}$ and $A\mathbf{x}$, given by $\|\mathbf{b} - A\mathbf{x}\|$, the better the approximation. The **general least-squares problem** is to find an $\mathbf{x}$ that makes $\|\mathbf{b} - A\mathbf{x}\|$ as small as possible. The adjective "least-squares" arises from the fact that $\|\mathbf{b} - A\mathbf{x}\|$ is the square root of a sum of squares.



THEOREM 6.4.1 Best Approximation Theorem

If W is a finite-dimensional subspace of an inner product space V, and if $\mathbf{b}$ is a vector in V, then $\text{proj}_W \mathbf{b}$ *is the* ***best approximation*** *to* $\mathbf{b}$ *from W in the sense that*

$$\|\mathbf{b} - \text{proj}_W \mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\|$$

for every vector $\mathbf{w}$ *in W that is different from* $\text{proj}_W \mathbf{b}$.

Proof For every vector $\mathbf{w}$ in W, we can write

$$\mathbf{b} - \mathbf{w} = (\mathbf{b} - \text{proj}_W \mathbf{b}) + (\text{proj}_W \mathbf{b} - \mathbf{w})$$

But $\text{proj}_W \mathbf{b} - \mathbf{w}$, being a difference of vectors in W, is itself in W; and since $\mathbf{b} - \text{proj}_W \mathbf{b}$ is orthogonal to W, the two terms on the right side of (1) are orthogonal. Thus, it follows from the Theorem of Pythagoras (Theorem 6.2.3) that

$$\|\mathbf{b} - \mathbf{w}\|^2 = \|\mathbf{b} - \text{proj}_W \mathbf{b}\|^2 + \|\text{proj}_W \mathbf{b} - \mathbf{w}\|^2$$

If $\mathbf{w} \ne \text{proj}_W \mathbf{b}$, it follows that the second term in this sum is positive, and hence that

$$\|\mathbf{b} - \text{proj}_W \mathbf{b}\|^2 < \|\mathbf{b} - \mathbf{w}\|^2$$

Since norms are nonnegative, it follows (from a property of inequalities) that

$$\|\mathbf{b} - \text{proj}_W \mathbf{b}\| < \|\mathbf{b} - \mathbf{w}\|$$

$$Ax = b$$

Multiplying both sides by A^T

$$A^TA\mathbf{x} = A^T\mathbf{b}$$

Normal Equation!

The set of least-squares solutions of $A\mathbf{x} = \mathbf{b}$ coincides with the nonempty set of solutions of the normal equations $A^TA\mathbf{x} = A^T\mathbf{b}$.

Proved in the next slide!

Solution of the General Least-Squares Problem

Given A and $\mathbf{b}$ as above, apply the Best Approximation Theorem in Section 6.3 to the subspace Col A. Let

$$\hat{\mathbf{b}} = \text{proj}_{\text{Col}\,A}\,\mathbf{b}$$

Because $\hat{\mathbf{b}}$ is in the column space of A, the equation $A\mathbf{x} = \hat{\mathbf{b}}$ *is* consistent, and there is an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ such that

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}} \tag{1}$$

Since $\hat{\mathbf{b}}$ is the closest point in Col A to $\mathbf{b}$, a vector $\hat{\mathbf{x}}$ is a least-squares solution of $A\mathbf{x} = \mathbf{b}$ if and only if $\hat{\mathbf{x}}$ satisfies (1). Such an $\hat{\mathbf{x}}$ in $\mathbb{R}^n$ is a list of weights that will build $\hat{\mathbf{b}}$ out of the columns of A. See Fig. 2. [There are many solutions of (1) if the equation has free variables.]

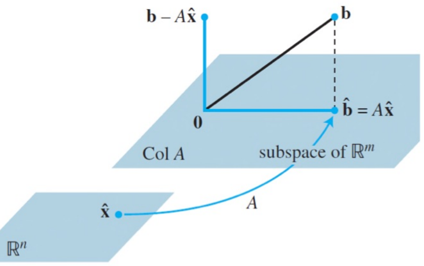


FIGURE 2 The least-squares solution $\hat{\mathbf{x}}$ is in $\mathbb{R}^n$.

Suppose $\hat{\mathbf{x}}$ satisfies $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. By the Orthogonal Decomposition Theorem in Section 6.3, the projection $\hat{\mathbf{b}}$ has the property that $\mathbf{b} - \hat{\mathbf{b}}$ is orthogonal to Col A, so $\mathbf{b} - A\hat{\mathbf{x}}$ is orthogonal to each column of A. If $\mathbf{a}_j$ is any column of A, then $\mathbf{a}_j \cdot (\mathbf{b} - A\hat{\mathbf{x}}) = 0$, and $\mathbf{a}_j^T(\mathbf{b} - A\hat{\mathbf{x}}) = 0$. Since each $\mathbf{a}_j^T$ is a row of A^T,

$$A^T(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0} \qquad (2)$$

(This equation also follows from Theorem 3 in Section 6.1.) Thus

$$A^T\mathbf{b} - A^TA\hat{\mathbf{x}} = \mathbf{0}$$
$$A^TA\hat{\mathbf{x}} = A^T\mathbf{b}$$

These calculations show that each least-squares solution of $A\mathbf{x} = \mathbf{b}$ satisfies the equation

$$\boxed{A^TA\mathbf{x} = A^T\mathbf{b}} \qquad (3)$$

The matrix equation (3) represents a system of equations called the **normal equations** for $A\mathbf{x} = \mathbf{b}$. A solution of (3) is often denoted by $\hat{\mathbf{x}}$.

EXAMPLE 1 Find a least-squares solution of the inconsistent system $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix}$$

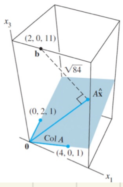
SOLUTION To use normal equations (3), compute:

$$A^TA = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}$$

$$A^T\mathbf{b} = \begin{bmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$

Then the equation $A^TA\mathbf{x} = A^T\mathbf{b}$ becomes

$$\begin{bmatrix} 17 & 1 \\ 1 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 19 \\ 11 \end{bmatrix}$$



Row operations can be used to solve this system, but since A^TA is invertible and 2×2, it is probably faster to compute

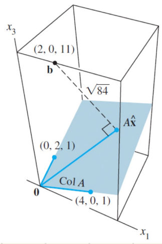
$$(A^TA)^{-1} = \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}$$

and then to solve $A^TA\mathbf{x} = A^T\mathbf{b}$ as

$$\hat{\mathbf{x}} = (A^TA)^{-1}A^T\mathbf{b}$$
$$= \frac{1}{84}\begin{bmatrix} 5 & -1 \\ -1 & 17 \end{bmatrix}\begin{bmatrix} 19 \\ 11 \end{bmatrix} = \frac{1}{84}\begin{bmatrix} 84 \\ 168 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

In many calculations, A^TA is invertible, but this is not always the case. The next

EXAMPLE 3 Given A and $\mathbf{b}$ as in Example 1, determine the least-squares error in the least-squares solution of $A\mathbf{x} = \mathbf{b}$.



SOLUTION From Example 1,

$$\mathbf{b} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} \quad \text{and} \quad A\hat{\mathbf{x}} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix}$$

Hence

$$\mathbf{b} - A\hat{\mathbf{x}} = \begin{bmatrix} 2 \\ 0 \\ 11 \end{bmatrix} - \begin{bmatrix} 4 \\ 4 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 8 \end{bmatrix}$$

and

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| = \sqrt{(-2)^2 + (-4)^2 + 8^2} = \sqrt{84}$$

The least-squares error is $\sqrt{84}$. For any $\mathbf{x}$ in $\mathbb{R}^2$, the distance between $\mathbf{b}$ and the vector $A\mathbf{x}$ is at least $\sqrt{84}$. See Fig. 3. Note that the least-squares solution $\hat{\mathbf{x}}$ itself does not appear in the figure.

Non-Unique least squared

EXAMPLE 2 Find a least-squares solution of $A\mathbf{x} = \mathbf{b}$ for

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix}$$

Note the linear dependency in the rows and columns of A:

- Column 1 = Column 2 + Column 3 + Column 4
- Rows 1 & 2 are same, but their corresponding b values are different (inconsistent)
- Rows 3 & 4 are same, but their corresponding b values are different (inconsistent)
- Rows 5 & 6 are same, but their corresponding b values are different (inconsistent)

SOLUTION Compute

$$A^TA = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} = \underbrace{\begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix}}_{A^TA}$$

$$A^T\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}\begin{bmatrix} -3 \\ -1 \\ 0 \\ 2 \\ 5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -4 \\ 2 \\ 6 \end{bmatrix}$$

Note that A^TA is always a square matrix.

The augmented matrix for $A^TA\mathbf{x} = A^T\mathbf{b}$ is

$$\left[\begin{array}{cccc|c} 6 & 2 & 2 & 2 & 4 \\ 2 & 2 & 0 & 0 & -4 \\ 2 & 0 & 2 & 0 & 2 \\ 2 & 0 & 0 & 2 & 6 \end{array}\right] \overset{\text{Reduced to}}{\sim} \begin{bmatrix} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & -1 & -5 \\ 0 & 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

A^TA $\quad A^Tb$

free varible

The general solution is $x_1 = 3 - x_4$, $x_2 = -5 + x_4$, $x_3 = -2 + x_4$, and x_4 is free. So the general least-squares solution of $A\mathbf{x} = \mathbf{b}$ has the form

$$\hat{\mathbf{x}} = \begin{bmatrix} 3 \\ -5 \\ -2 \\ 0 \end{bmatrix} + x_4\begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Note: Here, there are infinitely many solutions with the same least square error.

Note: Here, A^TA is not invertible (its determinant is 0).

A^TA may not be invertible if:

- some columns are linearly dependent (i.e. we have redundant features) (**as in this example**)
  - solution: remove the linear dependency
- too many features ($m < n$)
  - solution: delete some features, there are too many features for the amount of data we have

Ref: http://mlwiki.org/index.php/Normal_Equation

In general, the image of a vector $\mathbf{x}$ under multiplication by a square matrix A differs from $\mathbf{x}$ in both magnitude and direction. However, in the special case where $\mathbf{x}$ is an eigenvector of A, multiplication by A leaves the direction unchanged. For example,

EXAMPLE 1 Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. The images of $\mathbf{u}$ and $\mathbf{v}$ under multiplication by A are shown in Fig. 1. In fact, $A\mathbf{v}$ is just $2\mathbf{v}$. So A only "stretches," or dilates, $\mathbf{v}$. ■

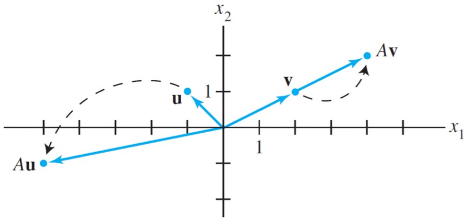


FIGURE 1 Effects of multiplication by A.

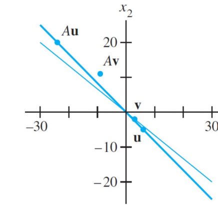
EXAMPLE 2 Let $A = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} 6 \\ -5 \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$. Are $\mathbf{u}$ and $\mathbf{v}$ eigenvectors of A?

SOLUTION

$$A\mathbf{u} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 6 \\ -5 \end{bmatrix} = \begin{bmatrix} -24 \\ 20 \end{bmatrix} = -4\begin{bmatrix} 6 \\ -5 \end{bmatrix} = -4\mathbf{u}$$

$$A\mathbf{v} = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ -2 \end{bmatrix} = \begin{bmatrix} -9 \\ 11 \end{bmatrix} \neq \lambda\begin{bmatrix} 3 \\ -2 \end{bmatrix}$$

Thus $\mathbf{u}$ is an eigenvector corresponding to an eigenvalue (-4), but $\mathbf{v}$ is not an eigenvector of A, because $A\mathbf{v}$ is not a multiple of $\mathbf{v}$. ■



$A\mathbf{u} = -4\mathbf{u}$, but $A\mathbf{v} \neq \lambda\mathbf{v}$.

EXAMPLE 3 Show that 7 is an eigenvalue of matrix A in Example 2, and find the corresponding eigenvectors.

SOLUTION The scalar 7 is an eigenvalue of A if and only if the equation

$$A\mathbf{x} = 7\mathbf{x} \tag{1}$$

has a nontrivial solution. But (1) is equivalent to $A\mathbf{x} - 7\mathbf{x} = \mathbf{0}$, or

$$(A - 7I)\mathbf{x} = \mathbf{0} \tag{2}$$

To solve this homogeneous equation, form the matrix

$$A - 7I = \begin{bmatrix} 1 & 6 \\ 5 & 2 \end{bmatrix} - \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}$$

The columns of $A - 7I$ are obviously linearly dependent, so (2) has nontrivial solutions. Thus 7 *is* an eigenvalue of A. To find the corresponding eigenvectors, use row operations:

rref $$\begin{bmatrix} -6 & 6 & 0 \\ 5 & -5 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The general solution has the form $x_2\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Each vector of this form with $x_2 \neq 0$ is an eigenvector corresponding to $\lambda = 7$. ■

Eigenspace

EXAMPLE 4 Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. An eigenvalue of A is 2. Find a basis for the corresponding eigenspace.

SOLUTION Form

$$A - 2I = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix} - \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{bmatrix}$$

and row reduce the augmented matrix for $(A - 2I)\mathbf{x} = \mathbf{0}$:

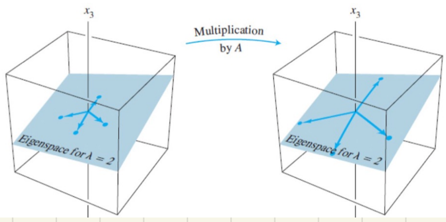
$$\begin{bmatrix} 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \\ 2 & -1 & 6 & 0 \end{bmatrix} \sim \begin{bmatrix} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

At this point, it is clear that 2 is indeed an eigenvalue of A because the equation $(A - 2I)\mathbf{x} = \mathbf{0}$ has free variables. The general solution is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_2\begin{bmatrix} 1/2 \\ 1 \\ 0 \end{bmatrix} + x_3\begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}, \quad x_2 \text{ and } x_3 \text{ free}$$

The eigenspace, shown in Fig. 3, is a two-dimensional subspace of $\mathbb{R}^3$. A basis is

$$\left\{\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}\right\}$$ ■



Find Eigenvectors (det = 0)

EXAMPLE 1 Find the eigenvalues of $A = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix}$.

SOLUTION We must find all scalars λ such that the matrix equation

$$(A - \lambda I)\mathbf{x} = \mathbf{0}$$

has a nontrivial solution. By the Invertible Matrix Theorem in Section 2.3, this problem is equivalent to finding all λ such that the matrix $A - \lambda I$ is *not* invertible, where

$$A - \lambda I = \begin{bmatrix} 2 & 3 \\ 3 & -6 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} = \begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix}$$

By Theorem 4 in Section 2.2, this matrix fails to be invertible precisely when its determinant is zero. So the eigenvalues of A are the solutions of the equation

$$\det(A - \lambda I) = \det\begin{bmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{bmatrix} = 0$$

Recall that

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

So

$$\begin{aligned} \boxed{\det(A - \lambda I)} &= (2-\lambda)(-6-\lambda) - (3)(3) \\ &= -12 + 6\lambda - 2\lambda + \lambda^2 - 9 \\ &\boxed{= \lambda^2 + 4\lambda - 21} \\ &= (\lambda - 3)(\lambda + 7) \end{aligned}$$

If $\det(A - \lambda I) = 0$, then $\lambda = 3$ or $\lambda = -7$. So the eigenvalues of A are 3 and -7. ■

Characteristic Eqn

Characteristic Polynomial

If A is not invertible, 0 can be eigenvector

EXAMPLE 5 Let $A = \begin{bmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{bmatrix}$ and $B = \begin{bmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{bmatrix}$. The eigenvalues of A are 3, 0, and 2. The eigenvalues of B are 4 and 1. ■

What does it mean for a matrix A to have an eigenvalue of 0, such as in Example 5? This happens if and only if the equation

$$A\mathbf{x} = 0\mathbf{x} \tag{4}$$

has a nontrivial solution. But (4) is equivalent to $A\mathbf{x} = \mathbf{0}$, which has a nontrivial solution if and only if A is not invertible. Thus 0 *is an eigenvalue of A if and only if A is not invertible*. This fact will be added to the Invertible Matrix Theorem in Section 5.2.

THEOREM **The Invertible Matrix Theorem (continued)**

Let A be an $n \times n$ matrix. Then A is invertible if and only if:

s. The number 0 is *not* an eigenvalue of A.

t. The determinant of A is *not* zero.

Diag. matrix, PDP^{-1}

2. Let $A = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}$, $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, and $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Suppose you are told that $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of A. Use this information to diagonalize A.

Sol:

2. Compute $A\mathbf{v}_1 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \cdot \mathbf{v}_1$, and

$$A\mathbf{v}_2 = \begin{bmatrix} -3 & 12 \\ -2 & 7 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 3 \cdot \mathbf{v}_2$$

So, $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors for the eigenvalues 1 and 3, respectively. Thus

$$A = PDP^{-1}, \quad \text{where} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

Property	Description
Determinant	A and $P^{-1}AP$ have the same determinant.
Invertibility	A is invertible if and only if $P^{-1}AP$ is invertible.
Rank	A and $P^{-1}AP$ have the same rank.
Nullity	A and $P^{-1}AP$ have the same nullity.
Trace	A and $P^{-1}AP$ have the same trace.
Characteristic polynomial	A and $P^{-1}AP$ have the same characteristic polynomial.
Eigenvalues	A and $P^{-1}AP$ have the same eigenvalues.
Eigenspace dimension	If λ is an eigenvalue of A (and hence of $P^{-1}AP$) then the eigenspace of A corresponding to λ and the eigenspace of $P^{-1}AP$ corresponding to λ have the same dimension.

To find the eigenvalue and vectors of a matrix, 2 steps:

a) First find the eigen value of the matrix.

- slide 8.1.2 shows how to find the eigen values.

b) Second, use Gaussian Elimination (row reduction) to find the solution of the homogenous equation $(A - \lambda I)x = 0$ for each eigenvalue λ.

Diagonalize matrix

Given a matrix A size NxN, to diagonalize it to D, perform:

1) Find the eigenvalues of A.
2) For each eigenvalue, find the eigenvectors of corresponding λ_i
3) If there are N independent eigenvectors v_i, then the matrix A can be represented as:

$$\begin{aligned} AP &= PD \\ A &= PDP^{-1} \\ P^{-1}AP &= D \end{aligned}$$

Where D = diagonal matrix with eigenvalues λ_i
And P is a matrix with columns that are corresponding eigenvectors v_i.

When is A diagonalizable? Sufficient condition: If A has n Distinct EigenValues -> Diagonalizable

But the oposite is not always T

THEOREM 6 Lay, 4thEd, pg 284, Ch 5.3

An $n \times n$ matrix with n distinct eigenvalues is diagonalizable.

PROOF Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be eigenvectors corresponding to the n distinct eigenvalues of a matrix A. Then $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is linearly independent, by Theorem 2 in Section 5.1. Hence A is diagonalizable, by Theorem 5. ■

It is not *necessary* for an $n \times n$ matrix to have n distinct eigenvalues in order to be diagonalizable. The 3×3 matrix in Example 3 is diagonalizable even though it has only two distinct eigenvalues.

Eigenspaces

Let λ be an eigenvalue of A. Recall that the eigenvectors of A for λ are the nonzero vectors in the nullspace of $A - \lambda I$. We call the nullspace $A - \lambda I$ the **eigenspace** of A for λ denoted by $\mathcal{E}_A(\lambda)$. In other words, $\mathcal{E}_A(\lambda)$ consists of all the eigenvectors of A for λ and the zero vector.

Example

Let $A = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$. Note that -1 is an eigenvalue of A. Then $A - (-1)I_2 = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$. The nullspace of this matrix is spanned by the single vector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. Hence, $\mathcal{E}_A(-1)$ is the span of $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Algebraic multiplicity vs geometric multiplicity

The **geometric multiplicity** of an eigenvalue λ of A is the dimension of $\mathcal{E}_A(\lambda)$. In the example above, the geometric multiplicity of -1 is 1 as the eigenspace is spanned by one nonzero vector.

In general, determining the geometric multiplicity of an eigenvalue requires no new technique because one is simply looking for the dimension of the nullspace of $A - \lambda I$.

The **algebraic multiplicity** of an eigenvalue λ of A is the number of times λ appears as a root of p_A. For the example above, one can check that -1 appears only once as a root. Let us now look at an example in which an eigenvalue has multiplicity higher than 1.

Diagonalizable matrix

EXAMPLE 3 Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

That is, find an invertible matrix P and a diagonal matrix D such that $A = PDP^{-1}$.

SOLUTION There are four steps to implement the description in Theorem 5.

Step 1. Find the eigenvalues of A. As mentioned in Section 5.2, the mechanics of this step are appropriate for a computer when the matrix is larger than 2×2. To avoid unnecessary distractions, the text will usually supply information needed for this step. In the present case, the characteristic equation turns out to involve a cubic polynomial that can be factored:

$$\begin{aligned} 0 = \det(A - \lambda I) &= -\lambda^3 - 3\lambda^2 + 4 \\ &= -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$.

Step 2. Find three linearly independent eigenvectors of A. Three vectors are needed because A is a 3×3 matrix. This is the critical step. If it fails, then Theorem 5 says that A cannot be diagonalized. The method in Section 5.1 produces a basis for each eigenspace:

$$\text{Basis for } \lambda = 1: \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Basis for } \lambda = -2: \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

You can check that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is a linearly independent set.

Step 3. Construct P from the vectors in step 2. The order of the vectors is unimportant. Using the order chosen in step 2, form

$$P = [\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \mathbf{v}_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Step 4. Construct D from the corresponding eigenvalues. In this step, it is essential that the order of the eigenvalues matches the order chosen for the columns of P. Use the eigenvalue $\lambda = -2$ twice, once for each of the eigenvectors corresponding to $\lambda = -2$:

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

verify $AP=PD$ instead of $A=PDP^{-1}$

It is a good idea to check that P and D really work. To avoid computing P^{-1}, simply verify that $AP = PD$. This is equivalent to $A = PDP^{-1}$ when P is invertible. (However, be sure that P is invertible!) Compute

$$AP = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

$$PD = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$ ■

$$\begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} =$$

$x + y + z = 0$

$x = -y - z$

$$y\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + z\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{matrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{matrix}$$

$x_1 - x_3 = 0$

$x_2 + x_3 = 0$

$x_3 = x_1$

$-x_2$

Non Diagonalizable

EXAMPLE 4 Diagonalize the following matrix, if possible.

$$A = \begin{bmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

SOLUTION The characteristic equation of A turns out to be exactly the same as that in Example 3:

$$0 = \det(A - \lambda I) = -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$. However, it is easy to verify that each eigenspace is only one-dimensional:

$$\text{Basis for } \lambda = 1: \quad \mathbf{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$\text{Basis for } \lambda = -2: \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

There are no other eigenvalues, and every eigenvector of A is a multiple of either $\mathbf{v}_1$ or $\mathbf{v}_2$. Hence it is impossible to construct a basis of $\mathbb{R}^3$ using eigenvectors of A. By Theorem 5, A is *not* diagonalizable. ■

A^K

In many cases, the eigenvalue–eigenvector information contained within a matrix A can be displayed in a useful factorization of the form $A = PDP^{-1}$ where D is a diagonal matrix. In this section, the factorization enables us to compute A^k quickly for large values of k, a fundamental idea in several applications of linear algebra. Later, in Sections 5.6 and 5.7, the factorization will be used to analyze (and *decouple*) dynamical systems.

The following example illustrates that powers of a diagonal matrix are easy to compute.

EXAMPLE 1 If $D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$, then $D^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}$

and

$$D^3 = DD^2 = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix} = \begin{bmatrix} 5^3 & 0 \\ 0 & 3^3 \end{bmatrix}$$

In general,

$$D^k = \begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix} \quad \text{for } k \geq 1$$ ■

Finding A^k from $A = PDP^{-1}$

EXAMPLE 2 Let $A = \begin{bmatrix} 7 & 2 \\ -4 & 1 \end{bmatrix}$. Find a formula for A^k, given that $A = PDP^{-1}$, where

$$P = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}$$

SOLUTION The standard formula for the inverse of a 2×2 matrix yields

$$P^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Then, by associativity of matrix multiplication,

$$A^2 = (PDP^{-1})(PDP^{-1}) = PD\underbrace{(P^{-1}P)}_{I}DP^{-1} = PDDP^{-1}$$

$$= PD^2P^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

Again,

$$A^3 = (PDP^{-1})A^2 = (PD\underbrace{P^{-1})P}_{I}D^2P^{-1} = PDD^2P^{-1} = PD^3P^{-1}$$

In general, for $k \geq 1$,

$$A^k = PD^kP^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}\begin{bmatrix} 5^k & 0 \\ 0 & 3^k \end{bmatrix}\begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 \cdot 5^k - 3^k & 5^k - 3^k \\ 2 \cdot 3^k - 2 \cdot 5^k & 2 \cdot 3^k - 5^k \end{bmatrix}$$ ■

A square matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, if $A = PDP^{-1}$ for some invertible matrix P and some diagonal matrix D.

PRACTICE PROBLEMS

1. Compute A^8, where $A = \begin{bmatrix} 4 & -3 \\ 2 & -1 \end{bmatrix}$.

SOLUTIONS TO PRACTICE PROBLEMS

1. $\det(A - \lambda I) = \lambda^2 - 3\lambda + 2 = (\lambda - 2)(\lambda - 1)$. The eigenvalues are 2 and 1, and the corresponding eigenvectors are $\mathbf{v}_1 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Next, form

$$P = \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad P^{-1} = \begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix}$$

Since $A = PDP^{-1}$,

$$\begin{aligned} A^8 = PD^8P^{-1} &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2^8 & 0 \\ 0 & 1^8 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 256 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -2 & 3 \end{bmatrix} \\ &= \begin{bmatrix} 766 & -765 \\ 510 & -509 \end{bmatrix} \end{aligned}$$

EXAMPLE 2 Plot several trajectories of the dynamical system $\mathbf{x}_{k+1} = A\mathbf{x}_k$, when

$$A = \begin{bmatrix} .80 & 0 \\ 0 & .64 \end{bmatrix}$$

SOLUTION The eigenvalues of A are .8 and .64, with eigenvectors $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. If $\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2$, then

$$\mathbf{x}_k = c_1(.8)^k\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.64)^k\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

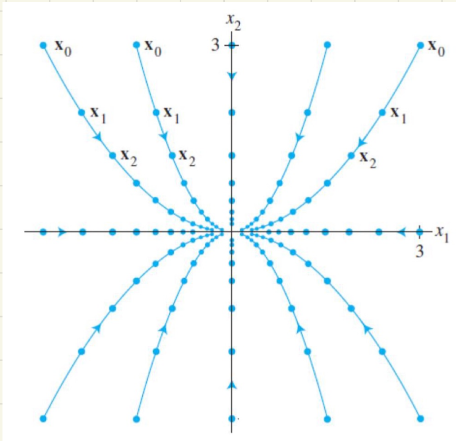


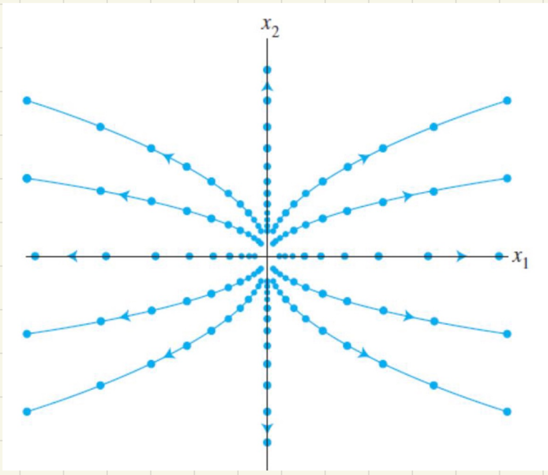
FIGURE 1 The origin as an attractor.

EXAMPLE 3 Plot several typical solutions of the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$, where

$$A = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.2 \end{bmatrix}$$

SOLUTION The eigenvalues of A are 1.44 and 1.2. If $\mathbf{x}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then

$$\mathbf{x}_k = c_1(1.44)^k\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(1.2)^k\begin{bmatrix} 0 \\ 1 \end{bmatrix}$$



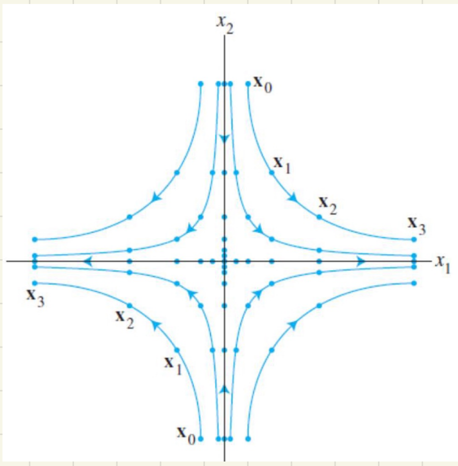
EXAMPLE 4 Plot several typical solutions of the equation $\mathbf{y}_{k+1} = D\mathbf{y}_k$, where

$$D = \begin{bmatrix} 2.0 & 0 \\ 0 & 0.5 \end{bmatrix}$$

(We write D and $\mathbf{y}$ here instead of A and $\mathbf{x}$ because this example will be used later.) Show that a solution $\{\mathbf{y}_k\}$ is unbounded if its initial point is not on the x_2-axis.

SOLUTION The eigenvalues of D are 2 and .5. If $\mathbf{y}_0 = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$, then

$$\mathbf{y}_k = c_1 2^k\begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2(.5)^k\begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{8}$$



Lay Example: Dynamical System

EXAMPLE 5 Let $A = \begin{bmatrix} .95 & .03 \\ .05 & .97 \end{bmatrix}$. Analyze the long-term behavior of the dynamical system defined by $\mathbf{x}_{k+1} = A\mathbf{x}_k$ $(k = 0, 1, 2, \ldots)$, with $\mathbf{x}_0 = \begin{bmatrix} .6 \\ .4 \end{bmatrix}$.

SOLUTION The first step is to find the eigenvalues of A and a basis for each eigenspace. The characteristic equation for A is

$$\begin{aligned} 0 &= \det \begin{bmatrix} .95-\lambda & .03 \\ .05 & .97-\lambda \end{bmatrix} = (.95-\lambda)(.97-\lambda) - (.03)(.05) \\ &= \lambda^2 - 1.92\lambda + .92 \end{aligned}$$

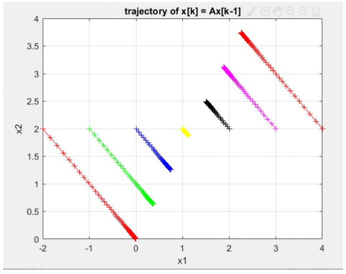
By the quadratic formula

$$\begin{aligned} \lambda &= \frac{1.92 \pm \sqrt{(1.92)^2 - 4(.92)}}{2} = \frac{1.92 \pm \sqrt{.0064}}{2} \\ &= \frac{1.92 \pm .08}{2} = 1 \quad \text{or} \quad .92 \end{aligned}$$

It is readily checked that eigenvectors corresponding to $\lambda = 1$ and $\lambda = .92$ are multiples of

$$\mathbf{v}_1 = \begin{bmatrix} 3 \\ 5 \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

respectively.




```matlab
A  = [0.95 0.03; 0.05 0.97];
N  = 10000

close all;
figure(idxFig);
figID =1;
idxFig = 0;
for  (y = 2:-1:2)
    for (x = -2:1:4)
     x0 = [x y];
     bruteForce_computeAx(A,x0,N,figID,idxFig);
```

```matlab
function bruteForce_computeAx(A,x0
colormapStr = 'rgbykm';
x  = x0(:);
hist_x = [];
% brute force way to compute x[k]
for k=0:N
   hist_x = [hist_x; x'];
   x = A*x;
end
tt = mod(idxFig,length(colormapStr
colorStr = [colormapStr(tt),'-+'];
```

The next step is to write the given $\mathbf{x}_0$ in terms of $\mathbf{v}_1$ and $\mathbf{v}_2$. This can be done because $\{\mathbf{v}_1, \mathbf{v}_2\}$ is obviously a basis for $\mathbb{R}^2$. (Why?) So there exist weights c_1 and c_2 such that

$$\mathbf{x}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 = [\,\mathbf{v}_1 \;\; \mathbf{v}_2\,]\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \tag{3}$$

In fact,

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= [\,\mathbf{v}_1 \;\; \mathbf{v}_2\,]^{-1}\mathbf{x}_0 = \begin{bmatrix} 3 & 1 \\ 5 & -1 \end{bmatrix}^{-1}\begin{bmatrix} .60 \\ .40 \end{bmatrix} \\ &= \frac{1}{-8}\begin{bmatrix} -1 & -1 \\ -5 & 3 \end{bmatrix}\begin{bmatrix} .60 \\ .40 \end{bmatrix} = \begin{bmatrix} .125 \\ .225 \end{bmatrix} \end{aligned} \tag{4}$$

Very important concept!!!
We decompose x into
the eigen basis

Because $\mathbf{v}_1$ and $\mathbf{v}_2$ in (3) are eigenvectors of A, with $A\mathbf{v}_1 = \mathbf{v}_1$ and $A\mathbf{v}_2 = .92\mathbf{v}_2$, we easily compute each $\mathbf{x}_k$:

$$\begin{aligned} \mathbf{x}_1 = A\mathbf{x}_0 &= c_1A\mathbf{v}_1 + c_2A\mathbf{v}_2 && \text{Using linearity of } \mathbf{x} \mapsto A\mathbf{x} \\ &= c_1\mathbf{v}_1 + c_2(.92)\mathbf{v}_2 && \mathbf{v}_1 \text{ and } \mathbf{v}_2 \text{ are eigenvectors.} \\ \mathbf{x}_2 = A\mathbf{x}_1 &= c_1A\mathbf{v}_1 + c_2(.92)A\mathbf{v}_2 \\ &= c_1\mathbf{v}_1 + c_2(.92)^2\mathbf{v}_2 \end{aligned}$$

and so on. In general,

$$\mathbf{x}_k = c_1\mathbf{v}_1 + c_2(.92)^k\mathbf{v}_2 \quad (k = 0, 1, 2, \ldots)$$

Using c_1 and c_2 from (4),

$$\mathbf{x}_k = .125\begin{bmatrix} 3 \\ 5 \end{bmatrix} + .225(.92)^k\begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad (k = 0, 1, 2, \ldots) \tag{5}$$

This explicit formula for $\mathbf{x}_k$ gives the solution of the difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$. As $k \to \infty$, $(.92)^k$ tends to zero and $\mathbf{x}_k$ tends to $\begin{bmatrix} .375 \\ .625 \end{bmatrix} = .125\mathbf{v}_1$. ■