1. System of Linear Equations

L/O: ① Determine if a given equation is linear.

↳ Has the form $a_1x_1 + a_2x_2 + \dots + a_nx_n = b$

a_i and b are constants ; a_i are coefficients & cannot be all $\emptyset$.

② Determine whether a linear system is consistent or inconsistent.

↳ Consistent: at least 1 solution

↳ Inconsistent: no solution

③ Recognize Homogeneous linear systems

↳ Linear systems where all b_i are $\emptyset$

④ Understand the types of solutions for homogeneous linear systems.

↳ Every homogeneous linear system is consistent (LHS=0, RHS=0)

↳ Trivial solution: if the system only has the above solution

↳ Infinitely many solutions: system has other non-trivial solutions in addition to the trivial one.

⑤ Represent linear systems in matrix form

↳ Coefficient matrix: matrix made up of only a_i

↳ Augmented matrix: matrix made up of both a_i and b.

⑥ Perform elementary row operations on an augmented matrix

↳ Two matrices are row equivalent if there is a sequence of ERO that transforms one matrix into another.

① Multiply a row through by a non-zero constant.
② Interchange two rows
③ Add constant times one row to another

} legal moves.

⑦ Use Gaussian elimination (to get Row Echelon Form) & Gauss-Jordan elimination (to get Reduced Row Echelon Form) to find the general solution of a linear system.

⑧ Write a given vector as a linear combination of other vectors

↳ The vector y defined by

$$y = c_1v_1 + c_2v_2 + \dots + c_nv_n$$

is called a linear combination of $\underbrace{v_1, v_2, \dots, v_n}_{\text{vectors}}$ with weights $\underbrace{c_1, c_2 \dots c_n}_{\text{scalar multiples}}$.

⑨ Determine the span of a set of vectors

↳ Span $\{v_1, v_2, \dots v_p\}$ is the set of all vectors that can be written in the form

$$c_1v_1 + c_2v_2 + \dots + c_pv_p$$

with scalars $c_1, c_2, \dots, c_p$.

⑩ Find the parametric vector equation of line and plane.

↳ Parametric vector equation of line and plane are given by $\vec{a} + s(\vec{b})$ and $\vec{a} + s(\vec{b}) + t(\vec{c})$ respectively.

⑪ Find the solution as a sum of a particular solution and of a homogeneous system. (slide 34)

↳ Every solution of a homogeneous system is a scalar multiple of another vector.

↳ Solutions of a non-homogeneous system will contain a particular solution and a scalar multiple.

(12) Determine if a set of vectors is linearly independent.

↳ An indexed set of vectors $\{v_1, \ldots, v_p\}$ in $\mathbb{R}^n$ is linearly independent if the vector equation

$$x_1 v_1 + x_2 v_2 + \cdots + x_p v_p = 0$$

has only the trivial solution (ie. you cannot represent any vector in the set as a linear combination of the others.)

↳ Linearly dependent if there exists weights $c_1 \cdots c_p$ where not all are 0 such that

$$c_1 v_1 + c_2 v_2 + \cdots + c_p v_p = 0$$

* proof on KhanAcademy

(13) Find the image of a vector under a linear transformation

↳ Similar to passing a variable (vector) through a function (transformation)

* only linear if ① $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v}$ in the domain of T.

② $T(c\vec{u}) = cT(\vec{u})$ where c is a scalar multiple.

(14) Determine the matrix corresponding to a linear transformation.

↳ Find the corresponding identity to the Transform function domain

Express the transformed columns of I_n after transforming (ie $T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}5\\-7\\2\end{bmatrix}$, $T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-3\\8\\0\end{bmatrix}$)

Pass in arbitrary values into T

$$T\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}5 & -3\\-7 & 8\\2 & 0\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$$

Theorems

1.1 — If A is an $m \times n$ matrix, with columns $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ and if $\vec{b}$ is in $\mathbb{R}^m$, then the matrix equation

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

has the same solution set as the vector equation

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b$$

which in turn, has the same solution set as the system of linear equations whose augmented matrix is

$$[a_1, a_2, \ldots, a_n \,\vdots\, b]$$

1.2 — Let A be an $m \times n$ matrix, then the following are all false or all true:

a) For each $\vec{b}$ in $\mathbb{R}^m$, the equation $Ax = \vec{b}$ has a solution

b) Each $\vec{b}$ 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

Proof of **Theorem 1.2**

(a), (b) and (c) are logically equivalent.
For an arbitrary matrix A, show that (a) and (d) are either both true or both false.

If U is an echelon form of A and $\mathbf{b}$ is in $\mathbb{R}^m$, then we can row reduce $[A \quad \mathbf{b}]$ to $[U \quad \mathbf{d}]$ for some $\mathbf{d}$ in $\mathbb{R}^m$.

Assume (d) is true. Then each row of U contains a pivot position $\Rightarrow$ No pivot in augmented column $\Rightarrow$ $A\mathbf{x} = \mathbf{b}$ has a solution for any $\mathbf{b}$ $\Rightarrow$ (a) is true.

Assume (d) is false. Then last row of U is all zeros. Let $\mathbf{d}$ be a vector with a 1 in the last entry. Then $[U \quad \mathbf{d}]$ represents an *inconsistent* system. Since row operations are reversible, $[U \quad \mathbf{d}]$ can be transformed into $[A \quad \mathbf{b}]$ $\Rightarrow$ $A\mathbf{x} = \mathbf{b}$ is also inconsistent $\Rightarrow$ (a) is false.

1.3 — If A is an $m \times n$ matrix, $\vec{u}$ and $\vec{v}$ are vectors in $\mathbb{R}^m$, and c is a scalar, then:

- $A(\vec{u}+\vec{v}) = A\vec{u} + A\vec{v}$
- $A(c\vec{u}) = c(A(\vec{u}))$

1.4 – Suppose the equation $Ax = b$ is consistent for some given b, and let p be a solution. Then the solution set of $Ax = b$ is the set of all vectors of the form $\vec{w} = \vec{p} + \vec{v_h}$ where $\vec{v_h}$ is any solution of the homogeneous equation $Ax = 0$.

1.5 – If a set contains more vectors than there are entries in each vector, then the set is linearly dependent.

Proof

Let $A = \begin{bmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_p \end{bmatrix}$. A is $n \times p$.
$A\mathbf{x} = \mathbf{0}$ corresponds to n equations in p unknowns.
$p > n \Longrightarrow$ more variables than equations $\Longrightarrow$ there must be a free variable $\Longrightarrow$
$A\mathbf{x} = \mathbf{0}$ has non trivial solution $\Longrightarrow$ columns of A are linearly dependent. ■

$$\begin{array}{cc} & p \\ n & \begin{bmatrix} * & * & * & * & * \\ * & * & * & * & * \\ * & * & * & * & * \end{bmatrix} \end{array}$$

If $p > n$, the columns are linearly dependent.

1.6 – If a set $S = \{v_1, \ldots, v_p\}$ in $\mathbb{R}^n$ contains the zero vector, then the set is linearly dependent.

Proof
By renumbering the vectors, we may suppose $\mathbf{v}_1 = \mathbf{0}$. Then the equation $1\mathbf{v}_1 + 0\mathbf{v}_2 + \ldots + 0\mathbf{v}_p = \mathbf{0}$ shows that S is linearly dependent. ■

1.7 – Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, then there exists a unique matrix A such that
$T(x) = Ax$ for all x in $\mathbb{R}^n$

A is the $m \times n$ matrix whose j^{th} column is the vector $T(\vec{e_j})$, where $\vec{e_j}$ is the j^{th} column of the identity matrix in $\mathbb{R}^n$:

ie. $A = [T(\vec{e}_1) \;\; T(\vec{e}_2) \ldots T(\vec{e}_n)]$

A is called the standard matrix for the linear transformation T.

2. Matrix Algebra

L/O: (1) Apply properties of matrix inverse

→ If A is an invertible matrix, then A^{-1} is invertible and $(A^{-1})^{-1} = A$

→ If A and B are $n \times n$ invertible matrices, then so is AB and $(AB)^{-1} = B^{-1}A^{-1}$
(i.e. the product of two invertible matrices is invertible)

→ If A is an invertible matrix, then so is A^T and $(A^T)^{-1} = (A^{-1})^T$

(2) Write the elementary matrix corresponding to an ERO

→ An elemantary matrix is one that is obtained by performing a single ERO on an identity matrix. (each Elementary matrix is invertible)

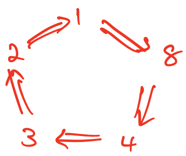
* The inverse of an Elementary matrix (E) is the elementary matrix of the same type that transform E back into I.

(3) Find the inverse of a 3x3 matrix using EROs

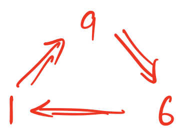
⇒ Any sequence of elementary row operations that reduces A to I also transforms I to A^{-1}

* A has to be invertible. (repeat steps on I to get A^{-1})

(4) Prove the invertible matrix theorem (theorem 2.4)



: If A is an (1) invertible matrix, there is an (2) $n \times n$ matrix C such that $CA = I$.
This implies that $Ax = 0$ has (4) only the trival solution ⇒ A has (3) n pivot positions
∴ There is a series of (2) ERO that can reduce A to I_n ⇒ A is invertible



: Since A is (1) invertible, there is an (9) $n \times n$ matrix D such that $AD = I$. This means that there (6) is exactly one solution for each $\vec{b}$ in $\mathbb{R}^n$, leads back to (3) which leads back to (1)

6<=>7 : Theorem 1.2; If $A\vec{x} = \vec{b}$ has a solution for each $\vec{b}$ in $\mathbb{R}$, the columns of A span $\mathbb{R}^n$.

4<=>5 : If $A\vec{x} = \vec{0}$ has (4) only the trivial solution, then the columns of A form a (5) linearly independent set

1 <=> 10 : If A is invertible, A^T is also invertible.

(5) Perform LU factorization with and without permutation

with permutation: Eg. $[PA = LU]$

$$A = \begin{bmatrix} 2 & 1 & 5 \\ 4 & 4 & -4 \\ 1 & 3 & 1 \end{bmatrix} \xrightarrow[R_1 <=> R_2]{P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \begin{bmatrix} 4 & 4 & -4 \\ 2 & 1 & 5 \\ 1 & 3 & 1 \end{bmatrix}$$

$R_2 \leftarrow R_2 - (\frac{1}{2})R_1$

$$\begin{bmatrix} 4 & 4 & -4 \\ 1/2 & -1 & 7 \\ 1 & 3 & 1 \end{bmatrix} \xrightarrow{R_3 \leftarrow R_3 - (\frac{1}{4})R_1} \begin{bmatrix} 4 & 4 & -4 \\ 1/2 & -1 & 7 \\ 1/4 & 2 & 2 \end{bmatrix} \xrightarrow[R_2 <=> R_3]{P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}} \begin{bmatrix} 4 & 4 & -4 \\ 1/4 & 2 & 2 \\ 1/2 & -1 & 7 \end{bmatrix}$$

$R_3 \leftarrow R_3 - (-\frac{1}{2})R_1$

$$\begin{bmatrix} 4 & 4 & -4 \\ 1/4 & 2 & 2 \\ 1/2 & -1/2 & 8 \end{bmatrix} \Rightarrow P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad L = \begin{bmatrix} 1 & 0 & 0 \\ 1/4 & 1 & 0 \\ 1/2 & -1/2 & 1 \end{bmatrix} \quad U = \begin{bmatrix} 4 & 4 & -4 \\ 0 & 2 & 2 \\ 0 & 0 & 8 \end{bmatrix}$$

Without permutation: Eg. $A = \begin{bmatrix} 2 & 1 & 5 \\ 4 & 4 & -4 \\ 1 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ \frac{1}{2} & & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 5 \\ 0 & 2 & -14 \\ 0 & \frac{5}{2} & -\frac{3}{2} \end{bmatrix}$ $\begin{array}{l} R_2 \leftarrow R_2 - (2)R_1 \\ R_3 \leftarrow R_3 - (\frac{1}{2})R_1 \end{array}$

$$= \underset{L}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ \frac{1}{2} & \frac{5}{4} & 1 \end{bmatrix}} \underset{U}{\begin{bmatrix} 2 & 1 & 5 \\ 0 & 2 & -14 \\ 0 & 0 & 16 \end{bmatrix}} \quad R_3 \leftarrow R_3 - (\tfrac{5}{4})R_2$$

Theorems

2.1 **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 $= A\mathbf{x} = A(A^{-1})\mathbf{b} = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b} =$ 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., $\mathbf{u} = A^{-1}\mathbf{b}$

□

2.2 **Theorem 2.2.** *Invertible matrices*

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 $(AB)^{-1} = B^{-1}A^{-1}$*
3. *If A is an invertible matrix, then so is A^T, and $(A^T)^{-1} = (A^{-1})^T$*

Proof

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 $(B^{-1}A^{-1})(AB) = I$.

2.3 **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}.*

ie. Given an $n \times n$ matrix A, $A \xrightarrow{\text{ERO}} I_n$ (implies invertibility)

$I_n \xrightarrow{\text{ERO}} A^{-1}$ (where sequence of ERO is the same)

3. Determinants

L/O : ① Interpret properties of determinants

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

→ The determinant changes sign when two rows are swapped.

→ 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 $|A'| = t|A|$

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

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.*

→ If two rows of A are equal, then det(A) = 0 (relates to second property)

→ Subtracting a multiple of one row from another row leaves det(A) unchanged.

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

This follows from Rule 3 and Rule 4.

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

→ A matrix with a row of '0' has det(A) = 0

→ If A is triangular, then det(A) is the product of diagonals.

→ If A is singular, then det(A) = 0, invertible otherwise (has pivots along diagonal)

→ $\det(AB) = \det(A) \cdot \det(B)$

→ $\det(A^T) = \det(A)$

② Interpret geometric properties of determinants

→ If A is a $\underline{2 \times 2}$ matrix, Area of parallelogram = det(A)

→ If A is a $\underline{3 \times 3}$ matrix, Area of parallelopiped = det(A)

③ Interpret geometry of linear transformation by determinants

→ 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 the area of $T(S)$ = det(A) · area of S. (similar for a 3×3 matrix A)

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

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

④ Compute change of area/volume using determinants.

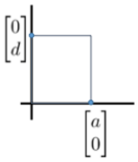
Theorems

3.1

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 parallelopiped determined by the columns of A is $|A|$.*

Proof. True for a 2×2 diagonal matrix:

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



Can we transform any 2×2 matrix $A = [\mathbf{a}_1 \quad \mathbf{a}_2]$ 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

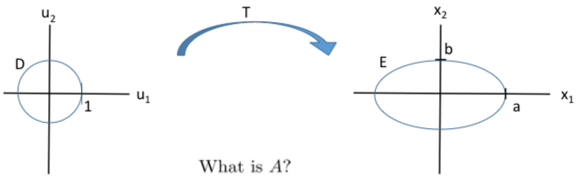
  Remember: properties apply to *columns* also.
- Adding a multiple of one column to another

3.2

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.

Thereom 3.2 is applicable for arbitrary shapes also.

Example



What is A?

$$\begin{aligned} \text{area of ellipse} &= \text{area of } T(D) \\ &= abs(|A|) \times \text{area of } D \\ &= ab \times \pi 1^2 = \pi ab \end{aligned}$$

If $\mathbf{x} = A\mathbf{u}$ with $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, equation of ellipse given by $\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$, what is $\mathbf{u}$?

Vector Spaces

L/O: ① Determine whether a given set with two operations is a vector space

1. The sum of $\vec{u}$ and $\vec{v}$, ie $\vec{u}+\vec{v}$ is in V
2. $\vec{u}+\vec{v}=\vec{v}+\vec{u}$
3. $(\vec{u}+\vec{v})+\vec{w}=\vec{u}+(\vec{v}+\vec{w})$
4. There is a zero vector in V such that $\vec{u}+\vec{0}=\vec{u}$
5. For each $\vec{u}$ in V, there is a vector $-\vec{u}$ such that $\vec{u}+(-\vec{u})=\vec{0}$
6. The scalar multiple of $\vec{u}$ by c, denoted by $c\vec{u}$ is in V
7. $c(\vec{u}+\vec{v})=c\vec{u}+c\vec{v}$.
8. $(c+d)\vec{u}=c\vec{u}+d\vec{u}$
9. $c(d\vec{u})=(cd)\vec{u}$
10. $1\vec{u}=\vec{u}$

② Determine whether a subset of a vector space is a subspace

- 3 properties: a. The zero vector of V is in subspace H.
  b. H is closed under vector addition
  c. H is closed under scalar multiplication

*subspace H is also a vector space satisfying ①

* Subspaces of $\mathbb{R}^n$ arise:

- as the set of all solutions to a system of homogeneous linear equations
- as the set of all linear combinations of specified vectors

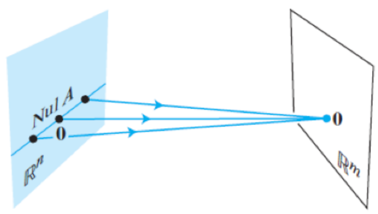
③ Determine whether a set of vectors in $\mathbb{R}^n$ span $\mathbb{R}^n$

- Theorem 4.1
  ↳ Given a vector space V, the span $\{v_1, \dots, v_p\}$ is a subspace of V (or the spanning set)
- **Find the rank of the matrix formed by the set**

④ Understand the Null Space

- Null space is the set of $\vec{x}$ that satisfies $A\vec{x}=\vec{0}$

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



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

* check yutsumura how to find a basis of nullspace etc

⑤ Understand the column space

- $C(A)$ is the set of all linear combinations of the columns of A. (ie if $A = [a_1 \dots a_n]$, then $C(A) = \text{span}\{a_1 \dots a_n\}$)

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

Exercise 4.5.1

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

$x_1\vec{a}_1 + x_2 \vec{a}_2 + x_3\vec{a}_3 + x_4\vec{a}_4 = \underset{3\times 1}{[b_1]}$

find $\vec{b}$

a. If the column space of A is a subspace of $\mathbb{R}^k$, what is k? $k = 3$

find $\vec{x}$

b. If the null space of A is a subspace of $\mathbb{R}^k$, what is k? $\underset{3\times 4}{A} \cdot \underset{4\times 1}{\vec{x}} = \vec{0}$ ↳ $\vec{x} \in \mathbb{R}^4$ $k=4$

c. Find a nonzero vector in $\mathbf{C}(A)$ and a nonzero vector in $\mathbf{N}(A)$. $\begin{bmatrix} 2 \\ -2 \\ 3 \end{bmatrix} \in C(A)$ $\begin{bmatrix} 9 \\ 5 \\ -1 \\ 0 \end{bmatrix} \in N(A)$ c part(i): take any column vector in A.

↳ solve for $A\vec{x} = 0$

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

no solution for $A\vec{u} = \vec{0}$ no; $C(A)$ is a subspace of $\mathbb{R}^3$

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 $\mathbf{N}(A)$?

Solve for $A\vec{x} = \vec{v}$; consistent ⇒ $\vec{v} \in C(A)$

A is a 3x4 matrix
⇒ v needs to be a 4x1 matrix
∴ $v \notin N(A)$

Nul A $\equiv$ $\mathbf{N}(A)$	Col A $\equiv$ $\mathbf{C}(A)$
1. Nul A is a subspace of $\mathbb{R}^n$.	**1.** Col A is a subspace of $\mathbb{R}^m$.
2. Nul A is implicitly defined; that is, you are given only a condition ($A\mathbf{x} = \mathbf{0}$) that vectors in Nul A must satisfy.	**2.** Col A is explicitly defined; that is, you are told how to build vectors in Col A.
3. It takes time to find vectors in Nul A. Row operations on $[\,A \quad \mathbf{0}\,]$ are required.	**3.** It is easy to find vectors in Col A. The columns of A are displayed; others are formed from them.
4. There is no obvious relation between Nul A and the entries in A.	**4.** There is an obvious relation between Col A and the entries in A, since each column of A is in Col A.
5. A typical vector $\mathbf{v}$ in Nul A has the property that $A\mathbf{v} = \mathbf{0}$.	**5.** A typical vector $\mathbf{v}$ in Col A has the property that the equation $A\mathbf{x} = \mathbf{v}$ is consistent.
6. Given a specific vector $\mathbf{v}$, it is easy to tell if $\mathbf{v}$ is in Nul A. Just compute $A\mathbf{v}$.	**6.** Given a specific vector $\mathbf{v}$, it may take time to tell if $\mathbf{v}$ is in Col A. Row operations on $[\,A \quad \mathbf{v}\,]$ are required.
7. Nul $A = \{\mathbf{0}\}$ if and only if the equation $A\mathbf{x} = \mathbf{0}$ has only the trivial solution.	**7.** Col $A = \mathbb{R}^m$ if and only if the equation $A\mathbf{x} = \mathbf{b}$ has a solution for every $\mathbf{b}$ in $\mathbb{R}^m$.
8. Nul $A = \{\mathbf{0}\}$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ is one-to-one.	**8.** Col $A = \mathbb{R}^m$ if and only if the linear transformation $\mathbf{x} \mapsto A\mathbf{x}$ maps $\mathbb{R}^n$ *onto* $\mathbb{R}^m$.

⑦ Generalize definition of linear transformation to include vector spaces

- A linear transformation T from a vector space V into a vector space W is a rule that assigns to each vector $\vec{x}$ in V a unique vector $T(\vec{x})$ in W such that
  1. $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v}$ in V
  2. $T(c\vec{u}) = cT(\vec{u})$ for all $\vec{u}$ in V and all scalars c.

* Given rank(A) and rank(B), $\text{rank}(AB) \le \min(\text{rank}(A), \text{rank}(B))$

Slyvester's rank inequality: given A $(m \times n)$ B $(n \times k)$

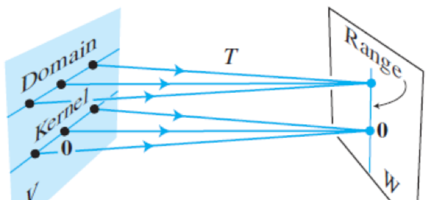
$$\text{rank}(AB) \ge \text{rank}(A) + \text{rank}(B) - n$$

⑧ Kernel and Range

- The kernel is a set of all $\vec{u}$ in V such that $T(\vec{u}) = \vec{0}$ (kernel is the domain)
- The range of T is the set of all vectors in W of the form $T(\vec{x})$ for some $\vec{x}$ in V.

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

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



Kernel is a subspace of V Range is a subspace of W

⑨ Show that a set of vectors is a basis for a vector space.

- An indexed set of vectors B in V is a basis for H if
  1. B is a linearly independent set
  2. The subspace spanned by B coincides with H. ie. $H = \text{Span}\{b_1 \dots b_p\}$

* A basis of V is a linearly dependent set that spans V.
* Spanning Set Theorem

⑩ Find the coordinates of a vector relative to a basis

- Let $S = \{v_1 \dots v_p\}$ be a set in V, and let $H = \text{Span}\{v_1 \cdots v_p\}$.
  1. If one of the vectors in S (eg v_k) is a linear combination of the remaining vectors in S, removing v_k, set S still spans H.
  2. If $H \neq \{\vec{0}\}$, some subset of S is a basis of H.

⑪ Bases for N(A) and C(A)

N(A) will contain the 0 vector
Basis of N(A) do not have to contain the 0 vector

- N(A)

When N(A) contains no zero vectors and the set of vectors are linearly independent, we can say the set is a basis for N(A)

- C(A)

Each nonpivot column of A is a linear combination of the pivot columns
$\Rightarrow$ basis of $C(A)$ is the set of pivot columns of A.

⑫ Coordinate System

- Suppose $B = \{b_1 \dots b_n\}$ is a basis for V and $\vec{x}$ is in V. The coordinates of $\vec{x}$ relative to the basis B are the weights $c_1, \dots, c_n$ such that

$$\mathbf{x} = c_1\mathbf{b}_1 + \cdots + c_n\mathbf{b}_n. \quad \text{and} \quad [\mathbf{x}]_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \Rightarrow [b_1 \ \dots \ b_n]\begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \vec{x}$$

$\vec{x}$ relative to the basis of V

⑬ Dimensions of a vector space.
- The dimensions of V (dim V) is the number of vectors in a basis of V.
  * the dimension of the zero vector space {0} is defined to be zero.

Orthogonality

① Norm of a vector
↳ length / magnitude of $\underset{\sim}{v}$ denoted by $\|\underset{\sim}{v}\|$

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

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) $\|k\mathbf{v}\| = |k|\|\mathbf{v}\|$

② Unit Length vector / unit vector
↳ Vector of length 1

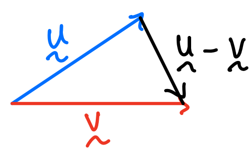
*normalizing a vector $\underset{\sim}{v}$

$$\mathbf{u} = \frac{1}{\|\mathbf{v}\|}\mathbf{v}$$ → unit vector of $\underset{\sim}{v}$ (also denoted as $\hat{\underset{\sim}{v}}$)

③ Distance between two vectors

For $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, the **distance between u and v**, written as dist$(\mathbf{u}, \mathbf{v})$, is the length of the vector $\mathbf{u} - \mathbf{v}$. That is,

$$\text{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$



④ Dot Product / inner product
↳ If $\underset{\sim}{u}$ and $\underset{\sim}{v}$ are nonzero vectors and the angle between them is θ, then the dot product $\underset{\sim}{u} \cdot \underset{\sim}{v}$ is defined as: $\underset{\sim}{u} \cdot \underset{\sim}{v} = \|\underset{\sim}{u}\|\|\underset{\sim}{v}\| \cos\theta$
* If either vector is $\underset{\sim}{0}$, then $\underset{\sim}{u} \cdot \underset{\sim}{v} = 0$ or $\underset{\sim}{u} \cdot \underset{\sim}{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n$

The sign of the dot product reveals information about the angle θ that we can obtain by rewriting Formula (12) as

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} \tag{13}$$

Since $0 \leq \theta \leq \pi$, it follows from Formula (13) and properties of the cosine function studied in trigonometry that

- θ is acute if $\mathbf{u} \cdot \mathbf{v} > 0$.
- θ is obtuse if $\mathbf{u} \cdot \mathbf{v} < 0$.
- $\theta = \pi/2$ if $\mathbf{u} \cdot \mathbf{v} = 0$.

$$v.w = ||v|| \times ||w|| \times cos\theta$$

$$v.w = \underbrace{(||w|| \times cos\theta)}_{\text{Length of projection of w onto v}} \times \underbrace{||v||}_{\text{Length of vector v}}$$

$$v.w = \underbrace{(||v|| \times cos\theta)}_{\text{Length of projection of v onto w}} \times \underbrace{||w||}_{\text{Length of vector w}}$$

THEOREM 3.2.2 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* R^n, *and if* k *is a scalar, then:*

(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]**

THEOREM 3.2.3 *If* $\mathbf{u}$, $\mathbf{v}$, *and* $\mathbf{w}$ *are vectors in* R^n, *and if* k *is a scalar, then:*

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

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

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

where A is a n×n matrix

* ⑤ Cauchy-Schwarz Inequality

↳ From the dot product formula, we see that:

$$u \cdot v = \|u\|\,\|v\| \cos\theta \Rightarrow \theta = \cos\left(\frac{u \cdot v}{\|u\|\,\|v\|}\right) \Rightarrow \underbrace{-1 \le \frac{u \cdot v}{\|u\|\,\|v\|} \le 1}$$

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*

$$|\mathbf{u} \cdot \mathbf{v}| \le \|\mathbf{u}\|\,\|\mathbf{v}\| \qquad (22)$$

or in terms of components

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

* ⑥ Triangle Inequality

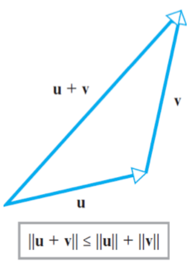


Figure 3.2.8

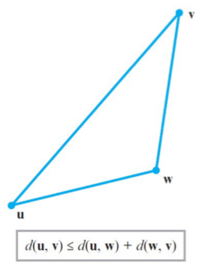


Figure 3.2.9

THEOREM 3.2.5 *If* **u**, **v**, *and* **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]**

Proof (*a*)

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

⑦ Parallelogram eqn for vectors

↳ Sum of squares of 2 diagonals of a parallelogram = sum of squares of the 4 sides.

THEOREM 3.2.6 Parallelogram Equation for Vectors

If **u** *and* **v** *are vectors in* R^n, *then*

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

⑧ Perpendicularity / Orthogonality.

↳ Two non-zero vectors are orthogonal if their dot product = 0

↳ the zero vector is orthogonal to every vector in the same dimension

THEOREM 3.3.1

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

$$ax + by + c = 0 \qquad (4)$$

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

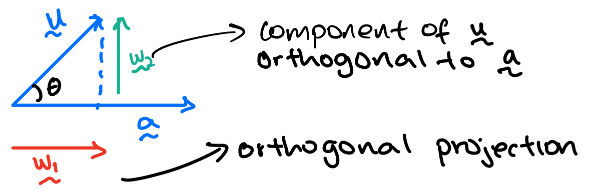
(*b*) *If a, b, and c are constants that are not all zero, then an equation of the form*

$$ax + by + cz + d = 0 \qquad (5)$$

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

plane equation ax + by + cz = d has normal (a,b,c)!

line equation ax + by = c has normal (a,b)!



* ⑨ Orthogonal Projections (projection vector)

THEOREM 3.3.2 Projection Theorem

If **u** *and* **a** *are vectors in* R^n, *and if* $\mathbf{a} \ne 0$, *then* **u** *can be expressed in exactly one way in the form* $\mathbf{u} = \mathbf{w}_1 + \mathbf{w}_2$, *where* $\mathbf{w}_1$ *is a scalar multiple of* **a** *and* $\mathbf{w}_2$ *is orthogonal to* **a**.

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

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

$$\Rightarrow \text{proj}_{a}\, u = \frac{u \cdot a}{\|a\|\,\|a\|}\, a = (u \cdot \hat{a})(\hat{a})$$

$$\Rightarrow \|\text{proj}_a u\| = \|u\| \times |\cos\theta|$$

* Vectors written without squiggly lines from now on.

(10) Orthogonal sets & basis

↳ A set of vectors $\{u_1 \cdots u_n\}$ is an orthogonal set if each pair of distinct vectors are orthogonal

THEOREM 4 $Note. p \leq n$

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. ■

Let $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ 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

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

PROOF As in the preceding proof, the orthogonality of $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ 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 $j = 2, \ldots, p$, compute $\mathbf{y}\cdot\mathbf{u}_j$ and solve for c_j. ■

* If set is an orthonormal basis, then the coefficients that will give vector y can be calculated using $c_i = (y \cdot u_i)$ since $u_i \cdot u_i = 1$

(11) Orthonormal Sets and Matrices.

↳ An orthogonal set is an orthonormal set if vectors within have a norm of 1.

* An $m \times n$ matrix U has orthonormal columns if $U^T U = I$ (Theorem 6)

↳ $u_i \cdot u_i = 1^2 \Rightarrow \text{norm}(u_i, u_i) = 1$

↳ Implies that $U^{-1} = U^T \Rightarrow U^T$ is also orthogonal

THEOREM 7

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

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 $\mathbf{x}\cdot\mathbf{y} = 0$

Theorems 6 and 7 are particularly useful when applied to *square* matrices. An **orthogonal matrix** is a square invertible matrix U such that $U^{-1} = U^T$. By Theorem 6, such a matrix has orthonormal columns.[1] It is easy to see that any *square* matrix with orthonormal columns is an orthogonal matrix. Surprisingly, such a matrix must have orthonormal *rows*, too.

(12) Orthogonal Decomposition.

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

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} \qquad (1)$$

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

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p \qquad (2)$$

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

⇒ This is saying that for any y that is lying outside of subspace we can find a $\hat{y}$ (projection of y onto the subspace W) such that the distance between y and $\hat{y}$ is the smallest possible.

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

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

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

Conclusion: if y is in the space of W, then we can show $\hat{y}$ to be equal to y using theorem 8. Else, it's the best approximation according to theorem 9.

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

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

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

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

$Note: p < n$

PROOF Formula (4) follows immediately from (2) in Theorem 8. Also, (4) shows that $\text{proj}_W\,\mathbf{y}$ is a linear combination of the columns of U using the weights $\mathbf{y}\cdot\mathbf{u}_1, \mathbf{y}\cdot\mathbf{u}_2, \ldots, \mathbf{y}\cdot\mathbf{u}_p$. The weights can be written as $\mathbf{u}_1^T\mathbf{y}, \mathbf{u}_2^T\mathbf{y}, \ldots, \mathbf{u}_p^T\mathbf{y}$, showing that they are the entries in $U^T\mathbf{y}$ and justifying (5). ■

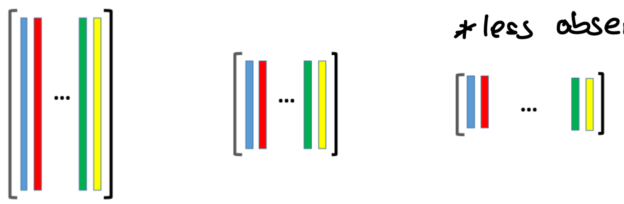
⑬ Gram - Schmidt Process (Review of prev concepts)

↳ For an M×N matrix A,

$Ax = b \Rightarrow x \in R^N \;\&\; b \in R^M$

Based on M & N, there exist three cases:

*less observed



$M \gg N$	$M = N$	$M \ll N$
More equations, less unknowns.		**Less equations, more unknowns.**
Hence, **over-determined**!		Hence, **under-determined**!

A can be decomposed into $A = QR$

↳ Q is an M×N mat with orthonormal columns

↳ R is an upper triangular N×N mat.

The problem to find x can then be easily solved by:

$$\begin{aligned} Ax &= b \\ QRx &= b \\ Q^TQRx &= Q^Tb, \text{ hence} \\ Rx &= Q^Tb \end{aligned}$$

Since R is upper triangle, x can be quickly found by back-substitution.

Note: if b is not in C(A), then the found x will only result in the orthogonal projection of b onto C(A).

* ⑭ Gram - Schmidt Process 2.1

↳ orthogonalises a set of vectors

The Gram–Schmidt Process

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

always set $v_1 = x_1$ → $\mathbf{v}_1 = \mathbf{x}_1$

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1$$

$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

$$\vdots$$

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

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

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

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

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

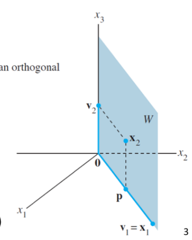


FIGURE 1 Construction of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.

1s (4th Edition)
t Squares

$$\underbrace{\text{Let } \mathbf{x}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}, \text{ and } \mathbf{x}_3 = \begin{bmatrix}0\\0\\1\\1\end{bmatrix}}_{\text{forms the subspace W of } R^4}$$

Orthogonal Basis :

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

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

$$\begin{aligned} \mathbf{v}_2 &= \mathbf{x}_2 - \text{proj}_{W_1}\mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Since } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix} \xrightarrow{\times 4} \overset{v_2'}{\begin{bmatrix}-3\\1\\1\\1\end{bmatrix}} \end{aligned}$$

$$V_3 = x_3 - \frac{x_3 \cdot V_1}{V_1 \cdot V_1}V_1 - \frac{x_3 \cdot V_2'}{V_2' \cdot V_2'}V_2'$$

$$= \begin{bmatrix}0\\0\\1\\1\end{bmatrix} - \frac{2}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} - \frac{2}{12}\begin{bmatrix}-3\\1\\1\\1\end{bmatrix} = \begin{bmatrix}0\\-2/3\\1/3\\1/3\end{bmatrix}$$

⑮ Using Gram-Schmidt for QR factorising (point ⑬)

If $A \in R^{m \times n}$ has full column rank,
i.e, A has linearly independent columns,
A can be factored as follows:

$$A = QR$$

$$A = [q_1 \;\; q_2 \;\; \cdots \;\; q_n]\begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1n} \\ 0 & R_{22} & \cdots & R_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & R_{nn} \end{bmatrix}$$

The Q Factor (economy qr):

- Q is $m \times n$ with orthonormal columns ($Q^TQ = I$)
- If A is square ($m = n$), then Q is orthogonal, i.e, $Q^TQ = QQ^T = I$

The R Factor:

- R is $n \times n$ upper triangular, with nonzero diagonal elements
- R is nonsingular (diagonal elements are nonzero)

* A=QR only if columns of A are linearly independent.

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

$$Q^T Q = I.$$

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

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

normalizing

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

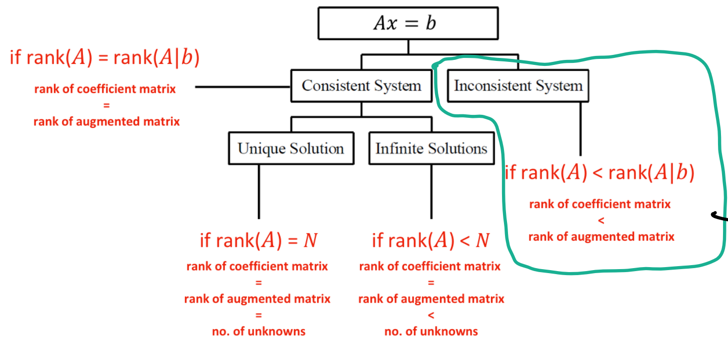
$$A = QR \Rightarrow Q^TA = Q^TQR = R = \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$$

$$\therefore \; A = \begin{pmatrix} 12 & -51 & 4 \\ 6 & 167 & -68 \\ -4 & 24 & -41 \end{pmatrix} = \begin{pmatrix} 6/7 & -69/175 & -58/175 \\ 3/7 & 158/175 & 6/175 \\ -2/7 & 6/35 & -33/35 \end{pmatrix} \times \begin{pmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{pmatrix}$$

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

Least Squares

① Consistency in a system of Linear Equation.



→ Least Squares Problem Focuses on this

② Least Squares Problem

↳ For an $m \times n$ matrix A and $Ax = b$ where $m \gg n$

↳ Equations may be inconsistent ⇒ no solution

∴ We find an $\hat{x}$ such that $A\hat{x}$ is as close to b as possible

* b does not lie in the same Columnspace of A. ⇒ We need an approximation
if b lies in the column space of A, then the error is 0 in this approximation.

③ Best Approximation Theorem & Normal Equation

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 $\text{proj}_W \mathbf{b}$.

*same as point ⑫ in Orthogonality

Proof: 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,

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

(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

$$A^TA\mathbf{x} = A^T\mathbf{b} \tag{3}$$

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 $\hat{\mathbf{x}}$.

* Normal Equation: $A^TAx = A^Tb$

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

* 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 independent.
c. The matrix A^TA is invertible.

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

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

If A^TA is not invertible, we can do Gaussian Elimination to find $\hat{x}$ but the solution is general ⇒ we need to remove redundant columns.

④ Projection matrix

↳ We have seen in prev chapter that:

$v v^T \in R^{N\times1} \times R^{1\times N} = R^{N\times N}$ (matrix)

$v^T v \in R^{1\times N} \times R^{N\times 1} = R^{1\times 1}$ (scalar)

$u \in R^{N\times 1}$

$$proj_v u = \frac{v \cdot u}{v \cdot v} v$$

$v \cdot u = v^T u$

$v \cdot v = v^T v$

$$= \frac{1}{v^T v} v(v^T u)$$

$$= \frac{1}{v^T v}(v v^T) u = \boxed{\frac{v v^T}{v^T v}} u$$

The expression $\mathbf{v}\mathbf{v}^T$ is called an **outer product** (the transpose operator is outside the product versus its inside position in the inner product). If we define $P = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}$, then the projection formula becomes

$$proj_{\mathbf{v}}\mathbf{u} = P\mathbf{u}, \text{ where } P = \frac{\mathbf{v}\mathbf{v}^T}{\mathbf{v}^T\mathbf{v}}.$$

The matrix P is called the projection matrix. You can project any vector onto the vector $\mathbf{v}$ by multiplying by the matrix P.

gives a $N\times 1$ column vector

* With the projection matrix, $proj_v y = P y$

Projection matrix for the least squares solution:

→ We know $A^T A \hat{x} = A^T b$

if $A^T A$ is invertible, $\hat{x} = (A^T A)^{-1} A^T b \Rightarrow A\hat{x} = \hat{b} = A(A^T A)^{-1} A^T b$

Which is "equivalent" to $\hat{b} = \boxed{\frac{A A^T}{A^T A}} b$ where $\hat{b} = proj_{col(A)} b = \boxed{P} b$

⑤ Projection Matrix Properties

1. P is a square matrix
2. $P^T = P$
3. $P^N = P$ (Idempotent Property)

Eigenvalues & Eigenvectors

An **eigenvector** of an $n \times n$ matrix A is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ. A scalar λ is called an **eigenvalue** of A if there is a nontrivial solution $\mathbf{x}$ of $A\mathbf{x} = \lambda\mathbf{x}$; such an $\mathbf{x}$ is called an *eigenvector corresponding to* λ.[1]

to avoid the unimportant case $A\underset{\sim}{0} = \lambda \underset{\sim}{0}$ which holds for every A and λ.

→ $A\underset{\sim}{x}$ is a transformation of $\underset{\sim}{x}$

→ If $A\underset{\sim}{x}$ maps $\underset{\sim}{x}$ to a scaled version of itself, (ie $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$), then λ is an eigenvalue of A while $\underset{\sim}{x}$ is is λ's eigenvector.

① How to find eigenvectors if eigenvalue is known

-

$Ax = 7x \quad (\lambda = 7)$

$(Ax - 7x) = \underset{\sim}{0}$

$(A - 7I)x = \underset{\sim}{0}$

Then we find $A - 7I$ and solve for x using the augmented matrix

* solution should be general with infinite solutions. (non-trivial)

② How to find eigenvectors if eigenvalue unknown

- Similar to ①:

$$(A\underset{\sim}{x} - \lambda \underset{\sim}{x}) = \underset{\sim}{0}$$
$$(A - \lambda I)\underset{\sim}{x} = \underset{\sim}{0} \quad \text{—— (a)}$$

* λ is an eigenvalue of A $\longleftrightarrow$ (a) has the non-trivial solution
  ↳ solution to eigenvectors is the Nullspace of $A - \lambda I$ (or kernel)
    ↳ This set is the eigenspace of A corresponding to λ.
      * The zero vector in this subspace is not an eigenvector
  * Eigenspace can also span a plane

① + ② Summary

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. ③

b) Second, use Gaussian Elimination (row reduction)
   to find the solution of the homogenous equation
   $(A - \lambda I)x = 0$
   for each eigenvalue λ.

③ How to find eigenvalues

$$\begin{aligned} Ax &= \lambda x \\ Ax - \lambda x &= 0 \\ (A - \lambda I)x &= 0 \\ (\lambda I - A)x &= 0 \end{aligned}$$

Since eigenvector cannot be the zero vector, then the solution to this equation cannot be trivial
⇒ $(A - \lambda I)$ is a matrix with linearly dependent columns
⇒ $(A - \lambda I)$ is a singular matrix (Det = 0)

With this info, we can simplify the equation to a polynomial with 1 unknown λ (Characteristic Equation & Polynomial)

* ④ Characteristic equation

- First compute the matrix $A - \lambda I$
- Find $\det(A - \lambda I)$ and equate the equation to 0 to solve for λ (solutions are the eigen values of A)

Theorem Given a square matrix A and a scalar λ, the following statements are equivalent:

- λ is an eigenvalue of A,
- $N(A - \lambda I) \neq \{\mathbf{0}\}$,
- the matrix $A - \lambda I$ is singular,
- $\det(A - \lambda I) = 0$.

* If $\lambda = 0$, then A is not invertible

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.

⑤ Eigenvalues of triangular matrices

The eigenvalues of a triangular matrix are the entries on its main diagonal.

- The determinant of a triangular matrix is the product of its diagonal entries

⑥ Algebraic multiplicity

- A.M is the number of times an eigenvalue appears as a root of the characteristic polynomial

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

SOLUTION Form $A - \lambda I$, and use Theorem 3(d):

$$\det(A - \lambda I) = \det \begin{bmatrix} 5-\lambda & -2 & 6 & -1 \\ 0 & 3-\lambda & -8 & 0 \\ 0 & 0 & 5-\lambda & 4 \\ 0 & 0 & 0 & 1-\lambda \end{bmatrix}$$
$$= (5-\lambda)(3-\lambda)(5-\lambda)(1-\lambda)$$

The characteristic equation is

$$(5-\lambda)^2(3-\lambda)(1-\lambda) = 0$$

A.M = 2   AM = 1

The **set of solutions** $(\lambda_1, \lambda_2, \dots, \lambda_{N_\lambda})$, that is, the **eigenvalues**, is called the **spectrum** of A. The charaacteristic polynomial can be factored as follows:

$$p(\lambda) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_{N_\lambda})^{n_{N_\lambda}} = 0.$$

The integer n_i is termed the **algebraic multiplicity** of eigenvalue λ_i. It is the number of times an eigenvalue appears as a root of the characteristic polynomial.

The algebraic multiplicities sum to N (the number of rows in A):

$$\sum_{i=1}^{N_\lambda} n_i = N.$$

For each eigenvalue λ_i, there is a corresponding EigenSpace $E(\lambda_i)$

11

⑦ Similarity and Diagonalization ($A = PDP^{-1}$)

Given $n \times n$ A and B, $P^{-1}AP = B$, $A = PBP^{-1}$ } similarity transformation

* If B is a diagonal matrix, A is diagonalizable

Table 1 Similarity Invariants

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.

Diagonal matrices have nice properties:

1) Eigenvalues of diagonal matrixes are its diagonal element
2) Determinant == product of diagonal entries
3) Rank == number of non-zero entries in the diagonal
4) Multiplication: given A and diagonal matrix D (AD and DA):
   - when we pre-multiply A by a diagonal matrix D, the rows of A are multiplied by the diagonal elements of D;
   - when we post-multiply A by D, the columns of A are multiplied by the diagonal elements of D.
5) A diagonal matrix's inverse is reciprocal of diagonal elements
6) Product of diagonal matrices are easy to compute.

* ⑧ When is a matrix Diagonalizable

The Diagonalization Theorem

An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

* When we have a repeated eigenvalues, we need to further check if there are repeated eigen vectors

- A can be diagonalizable even with repeated e.values

→ Eigenvector basis

A is diagonalizable ⟷ There are enough eigenvectors to form a basis R^n

PROOF First, observe that if P is any $n \times n$ matrix with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$, and if D is any diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$, then

$$AP = A[\mathbf{v}_1 \ \ \mathbf{v}_2 \ \ \cdots \ \ \mathbf{v}_n] = [A\mathbf{v}_1 \ \ A\mathbf{v}_2 \ \ \cdots \ \ A\mathbf{v}_n] \tag{1}$$

while

$$PD = P\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} = [\lambda_1\mathbf{v}_1 \ \ \lambda_2\mathbf{v}_2 \ \ \cdots \ \ \lambda_n\mathbf{v}_n] \tag{2}$$

Now suppose A is diagonalizable and $A = PDP^{-1}$. Then right-multiplying this relation by P, we have $AP = PD$. In this case, equations (1) and (2) imply that

$$[A\mathbf{v}_1 \ \ A\mathbf{v}_2 \ \ \cdots \ \ A\mathbf{v}_n] = [\lambda_1\mathbf{v}_1 \ \ \lambda_2\mathbf{v}_2 \ \ \cdots \ \ \lambda_n\mathbf{v}_n] \tag{3}$$

Equating columns, we find that

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, \quad A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \quad \dots, \quad A\mathbf{v}_n = \lambda_n\mathbf{v}_n \tag{4}$$

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

$$A = PDP^{-1}, \quad \text{where} \quad P = \begin{bmatrix} 3 & 2 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$$

* Rather than computing P^{-1}, we check using $AP = PD$

⑨ Efficiently compute A^k

- In ⑦ & ⑧, we have learnt if A is diagonalizable, $A = PDP^{-1}$ where:

  P is formed by the eigenvectors of A
  D is formed by the corresponding eigenvalues of A.

$\therefore$ Since $A = PDP^{-1}$ and $PP^{-1} = I = P^{-1}P$,
we realise that:

$$A^k = \underbrace{PDP^{-1}\,PDP^{-1}\dots PDP^{-1}}_{k \text{ times}}$$

$$= PDIDID\dots IDP^{-1}$$

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

It is easy to find D^k since D^k is diagonal and its entries are the diagonals of D to the k^{th} power.

⑩ Similar Matrix and Linear Transformation.

Given A and $x[0]$, and $x[n+1] = Ax[n]$

we convert this problem to B-basis (eigenvectors of A if exists)

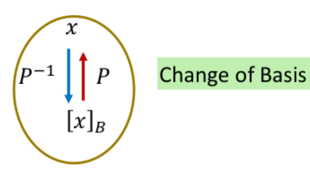
$$x[1] = Ax[0]$$
$$x[2] = Ax[1] = AAx[0] = A^2x[0]$$

$$\Rightarrow x[k] = A^kx[0]$$

using the new method we learnt in ⑦ & ⑧, we can calculate A^k efficiently such that $A^k = PD^kP^{-1}$

$$\Rightarrow x[k] = A^kx[0] = PD^kP^{-1}x[0]$$
$$\Rightarrow x[k] = P\,D^k\;x[0]_B$$

$\Longleftarrow$



$$x = P[x]_B$$
$$P^{-1}x = [x]_B$$

⑪ Dynamical System (Application of ⑩)

The equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ determines an infinite collection of equations. Beginning with an initial vector $\mathbf{x}_0$, we have

$$\begin{aligned} \mathbf{x}_1 &= A\mathbf{x}_0 \\ \mathbf{x}_2 &= A\mathbf{x}_1 \\ \mathbf{x}_3 &= A\mathbf{x}_2 \\ &\vdots \end{aligned}$$

The set $\{\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots\}$ is called a trajectory of the system. Note that $\mathbf{x}_k = A^k\mathbf{x}_0$.

Eigenvalues and eigenvectors provide the key to understanding the long-term behavior, or *evolution*, of a dynamical system described by a difference equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$.

ie the previous state x_j, when passed into A, will lead to an x_k such that

$$x_k = Ax_j = A^kx_0$$

Graphical Description of Solutions

When A is 2×2, algebraic calculations can be supplemented by a geometric description of a system's evolution. We can view the equation $\mathbf{x}_{k+1} = A\mathbf{x}_k$ as a description of what happens to an initial point $\mathbf{x}_0$ in $\mathbb{R}^2$ as it is transformed repeatedly by the mapping $\mathbf{x} \mapsto A\mathbf{x}$. The graph of $\mathbf{x}_0, \mathbf{x}_1, \dots$ is called a **trajectory** of the dynamical system.

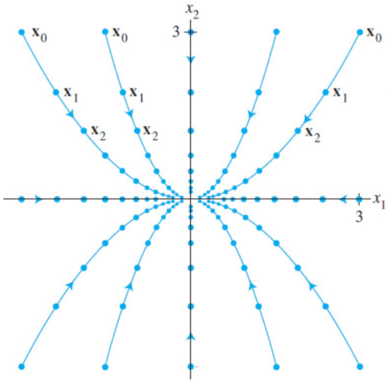


FIGURE 1 The origin as an attractor.

occurs when both eigenvalues < 1

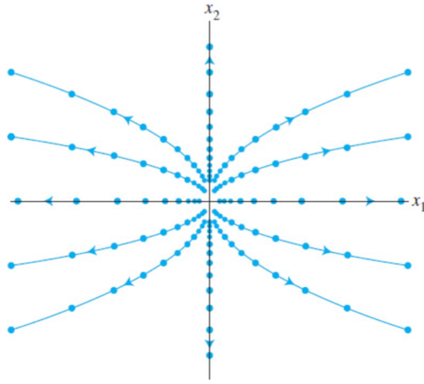


FIGURE 2 The origin as a repeller.

occurs when both eigenvalues > 1

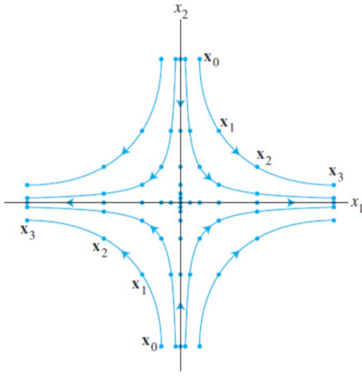


FIGURE 3 The origin as a saddle point.

occurs when e.val$_1$ > 1 & e.val$_2$ < 1

* Very important, go see last few slides of Part 6