# MA110-2023-2 (1st half): Linear Algebra

H. Ananthnarayan & Rekha Santhanam
$\infty$
Department of Mathematics, I.I.T. Bombay

March 8, 2024

This set of notes is to supplement the material presented in our classes of Linear Algebra in the first half of MA110 in Spring 2024 at IIT Bombay. The primary content was developed by using the reference:
*Linear Algebra and its Applications* by G. Strang, 4th Ed., Thomson.
Most of this content in these Appendices were not covered in class, but are the result of various questions asked by curious students. These are being shared primarily to give such students more content to mull over.

The topics covered in the class were:

1. LINEAR EQUATIONS & MATRICES
2. VECTOR SPACES
3. EIGENVALUES & EIGENVECTORS
4. ORTHOGONALITY & PROJECTIONS

The topics covered in these slides are:

APPENDIX I - DETERMINANTS

APPENDIX II - ORTHOGONAL SUBSPACES

APPENDIX III - ADDITIONAL TOPICS

(a) An Application of $QR$-factorization
(b) Angles in $\mathbb{R}^n$
(c) Inner Product Spaces
(d) Symmetric Matrices & SVD

NOTE: (i) The notation in these notes is the same as that discussed in class.
(ii) Work out as many examples as you can.

# Appendix I - DETERMINANTS

**Determinants: Key Properties**

Let $A$ and $B$ $n \times n$, and $c$ a scalar.

- TRUE/FALSE: $\det(A+B) = \det(A) + \det(B)$.
- TRUE/FALSE: $\det(cA) = c \det(A)$.
- $\det(AB) = \det(A)\det(B)$.
- $\det(A) = \det(A^T)$.
- If $A$ is orthogonal, i.e., $AA^T = I$, then $\det(A) =$
- If $A = [a_{ij}]$ is triangular, then $\det(A) =$ ___
- $A$ is invertible $\Leftrightarrow \det(A) \neq 0$.

  If this happens, then $\det(A^{-1}) =$ ___
- If $B = P^{-1}AP$ for an invertible matrix $P$,

  i.e., $A$ and $B$ are similar, then $\det(B) =$ ___
- If $A$ is invertible, and $d_1, \dots, d_n$ are the pivots of $A$,

  then $\det(A) =$ ___.

**Determinants: Defining Properties**

DEFN. The determinant function $\det : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ can be defined (uniquely) by its three basic properties.

- $\det(I) = 1$.
- The sign of determinant is reversed by a row exchange.

Thus, if $B = P_{ij}A$, i.e., $B$ is obtained from $A$ by exchanging two rows, then $\det(B) = -\det(A)$. In particular, $\det(I) = 1 \Rightarrow \det(P_{ij}) = -1$.

- det is linear in each row separately, i.e. , we fix $n-1$ row vectors, say $v_2, \cdots, v_n$,

then $\det \begin{pmatrix} - & v_2 & \cdots & v_n \end{pmatrix}^T : \mathbb{R}^n \to \mathbb{R}$ is a linear function.

I.e., for $c$, $d$ in $\mathbb{R}$, and vectors $u$ and $v$, if $A_{1*} = cu + dv$, we have

$\det \begin{pmatrix} cu+dv & A_{2*} & \cdots & A_{n*} \end{pmatrix}^T = c \det \begin{pmatrix} u & A_{2*} & \cdots & A_{n*} \end{pmatrix}^T + d \det \begin{pmatrix} v & A_{2*} & \cdots & A_{n*} \end{pmatrix}^T$.

There are $n$ such equations (for $n$ choices of rows).

**Determinants: Induced Properties**

1. If two rows of $A$ are equal, then $\det(A) = 0$.

   *Proof.* Suppose $i$-th and $j$-th rows of $A$ are equal, i.e., $A_{i*} = A_{j*}$, then $A = P_{ij}A$.

   Hence $\det(A) = \det(P_{ij}A) = -\det(A) \Rightarrow \boxed{\det(A) = 0}$.

2. If $B$ is obtained from $A$ by $\boxed{R_i \mapsto R_i + aR_j,}$ then $\det(B) = \det(A)$. (Use linearity in the $i$th row)

3. If $A$ is $n \times n$, and its row echelon form $U$ is obtained without row exchanges, then $\det(U) = \det(A)$.

   **Q:** What happens if there are row exchanges? Exercise!

4. If $A$ has a zero row, then $\det(A) = 0$.

   *Proof:* Let the $i$th row of $A$ be zero, i.e., $A_{i*} = 0$.

   Let $B$ be obtained from $A$ by $\boxed{R_i = R_i + R_j}$, i.e., $B = E_{ij}(1)A$. Then $B_{i*} = B_{j*}$.

   EXERCISE: Complete the proof.

**Determinants: Special Matrices**

5. If $A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}$ is diagonal, then $\det(A) = a_1 \cdots a_n$. (Use linearity).

6. If $A = (a_{ij})$ is triangular, then $\det(A) = a_{11} \dots a_{nn}$.

   *Proof.* If all $a_{ii}$ are non-zero, then by elementary row operations, $A$ reduces to the diagonal matrix $\begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix}$ whose determinant is $a_{11} \cdots a_{nn}$.

   If at least one diagonal entry is zero, then elimination will produce a zero row $\Rightarrow$ $\det(A) = 0$.

**Formula for Determinant:** $2 \times 2$ **case**

Write $(a, b) = (a, 0) + (0, b)$, the sum of vectors in coordinate directions.
Similarly write $(c, d) = (c, 0) + (0, d)$. By linearity,

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

For an $n \times n$ matrix, each row splits into $n$ coordinate directions,
so the expansion of $\det(A)$ has $n^n$ terms.
However, when two rows are in same coordinate direction, that term will be zero, e.g.,

$$\begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix} = 0, \quad \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = - \begin{vmatrix} c & 0 \\ 0 & b \end{vmatrix} = -bc \Rightarrow \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

The non-zero terms have to come in different columns.
So, there will be $n!$ such terms in the $n \times n$ case.

**Formula for Determinant:** $n \times n$ **case**

For $n \times n$ matrix $A = (a_{ij})$,

$$\det(A) = \sum_{\text{all permutations } P} (a_{1\alpha_1} \dots a_{n\alpha_n}) \det(P).$$

The sum is over $n!$ permutations of numbers $(1, \dots, n)$. Here a permutation $(i_1, i_2, \dots, i_n)$ of $(1, 2, \dots, n)$ corresponds to the product of permutation matrices $P = \begin{bmatrix} e_{i_1}^T \\ \vdots \\ e_{i_n}^T \end{bmatrix}$.

Then $\det(P) = +1$ if the number of row exchanges in $P$ needed to get $I$ is even, and $-1$ if it is odd.

**Cofactors:** $3\times 3$ **Case**

$$\begin{aligned}
\det(A) &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \\
&= a_{11}\,a_{22}\,a_{33}\,(1) + a_{11}\,a_{23}\,a_{32}\,(-1) + a_{12}\,a_{21}\,a_{33}\,(-1) \\
&\quad + a_{12}\,a_{23}\,a_{31}\,(1) + a_{13}\,a_{21}\,a_{32}\,(1) + a_{13}\,a_{22}\,a_{31}\,(-1) \\
&= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\
&\quad + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\
&= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + a_{12}(-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\
&= a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} \text{ where,}
\end{aligned}$$

$$C_{11} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix},\ C_{12} = (-1)\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix},\ C_{13} = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}.$$

**Cofactors:** $n\times n$ **Case**

Let $C_{1j}$ be the coefficient of $a_{1j}$ in the expansion

$$\det(A) = \sum_{\text{all permutations } P} (a_{1\alpha_1}\,\ldots\,a_{n\alpha_n})\det(P)$$

Then $\boxed{\det(A) = a_{11}C_{11} + a_{12}C_{12} + \ldots + a_{1n}C_{1n}}$ where,

$$\begin{aligned}
C_{1j} &= \Sigma a_{2\beta_2}\ldots a_{n\beta_n}\det(P) \\
&= (-1)^{1+j}\det\begin{bmatrix} a_{21} & . & a_{2(j-1)} & a_{2(j+1)} & . & a_{2n} \\ \vdots & . & \vdots & . & \vdots & \\ a_{n1} & . & a_{n(j-1)} & a_{j+1} & . & a_{nn} \end{bmatrix} \\
&= (-1)^{1+j}\det(M_{1j}),
\end{aligned}$$

where $M_{1j}$ is obtained from $A$ by deleting the $1^{\text{st}}$ row and $j^{\text{th}}$ column.

**det**$(AB) =$ **det**$(A)$ **det**$(B)$ **(Proof)**

7.

$$\boxed{\det(AB) = \det(A)\det(B)}$$

*Proof.* We may assume that $B$ are invertible. Else, $\text{rank}(AB) \le \text{rank}B \ne n$ $\Rightarrow \text{rank}(AB) \ne n \Rightarrow AB$ is not invertible.

Hint: For fixed $B$, show that the function $d$ defined by

$$d(A) = \det(AB)/\det(B)$$

satisfies the following properties

(a) $d(I) = 1$.

(b) If we interchange two rows of $A$, then $d$ changes its sign.

(c) $d$ is a linear function in each row of $A$.

Then $d$ is the unique determinant function det and $\det(AB) = \det(A)\det(B)$. $\square$

**Determinants of Transposes (Proof)**

8. $$\boxed{\det(A) = \det(A^T)}$$

   *Proof.* With $U$, $L$, and $P$, as usual write $PA = LU \Rightarrow A^T P^T = U^T L^T$ Since $U$ and $L$ are triangular, we get $\det(U) = \det(U^T)$ and $\det(L) = \det(L^T)$.

   Since $PP^T = I$ and $\det(P) = \pm 1$, we get $\det(P) = \det(P^T)$.

   Thus $\det(A) = \det(A^T)$. □

**Determinants and Invertibility (Proof)**

9. $A$ is invertible if and only if $\det(A) \neq 0$.

   By elimination, we get an upper triangular matrix $U$, a lower triangular matrix $L$ with diagonal entries 1, and a permutation matrix $P$, such that $PA = LU$.

   OBSERVATION 1: If $A$ is singular, then $\det(A) = 0$.

   This is because elimination produces a zero row in $U$ and hence $\det(A) = \pm \det(U) = 0$.

   OBSERVATION 2: If $A$ is invertible, then $\det(A) \neq 0$.

   This is because elimination produces $n$ pivots, say $d_1, \ldots, d_n$, which are non-zero. Then $U$ is upper triangular, with diagonal entries $d_1, \ldots, d_n \Rightarrow \det(A) = \pm \det(U) = \pm\, d_1 \cdots d_n \neq 0$.

   THUS we have: $A$ invertible $\Rightarrow \det(A) = \pm$(product of pivots).

   EXERCISE: If $AB$ is invertible, then so are $A$ and $B$.

   EXERCISE: $A$ is invertible if and only if $A^T$ is invertible.

**Determinant: Geometric Interpretation ($2 \times 2$)**

INVERTIBILITY: Very often we are interested in knowing when a matrix is invertible.

Consider $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Then $A$ is invertible if and only if $A$ has full rank.

If $a$, $c$ both are zero then clearly $\text{rank}(A) < 2 \Rightarrow A$ is not invertible.

Assume $a \neq 0$, else, interchange rows. The row operations $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xrightarrow{R_2 - c/aR_1} \begin{bmatrix} a & b \\ 0 & d - cb/a \end{bmatrix}$ show that $A$ is invertible if and only if $d - cb/a \neq 0$, i.e., $ad - bc \neq 0$.

AREA: The area of the parallelogram with sides as vectors $v = (a, b)$ and $w = (c, d)$ is equal to $ad - bc$. Thus,

$\boxed{\text{A } 2 \times 2 \text{ matrix } A \text{ is singular}} \Leftrightarrow \boxed{\text{its columns are on the same line}} \Leftrightarrow \boxed{\text{the area is zero.}}$

**Determinant: Geometric Interpretation**

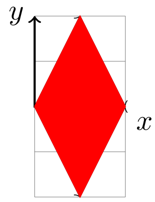
- Test for invertibility: An $n \times n$ matrix $A$ is invertible $\Leftrightarrow \det(A) \neq 0$.

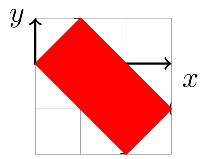
- $n$-dimensional volume: If $A$ is $n \times n$, then $|\det(A)| =$ the volume of the box (in $n$-dimensional space $\mathbb{R}^n$) with edges as rows of $A$.

Examples: (1) The volume (area) of a line in $\mathbb{R}^2 = 0$.

(2) The determinant of the matrix $A = \begin{pmatrix} 1 & 2 \\ 1 & -2 \end{pmatrix}$ is $\boxed{-4}$.

(3) Let's compute the volume of the box (parallelogram) with edges as rows of $A$ or columns of $A$.




$\boxed{= 4}$

**Expansion along the $i$-th row (Proof)**

If $C_{ij}$ is the coefficient of $a_{ij}$ in the formula of $\det(A)$, then

$\boxed{\det(A) \;=\; a_{i1}\, C_{i1} + \ldots + a_{in}\, C_{in}}$, where $C_{ij}$ is determined as follows:

By $i-1$ row exchanges on $A$, get the matrix $B = \begin{pmatrix} A_{i*} & A_{1*} & .. & A_{(i-1)*} & A_{(i+1)*} & .. & A_{n*} \end{pmatrix}^T$

Since $\det(A) = (-1)^{i-1} \det(B)$, we get

$$C_{ij}(A) = (-1)^{i-1} C_{1j}(B) = (-1)^{i-1}(-1)^{j-1} \det(M)$$

where $M$ is obtained from $B$ by deleting $1^{\text{st}}$ row and $j^{\text{th}}$ column. Here $M$ is obtained from $B$ by deleting its first row, and $j$-th column, and hence from $A$ by deleting $i$-th row and $j$-th column. Write $M$ as $M_{ij}$. Then $\boxed{C_{ij} = (-1)^{i+j} \det(M_{ij})}$

**Expansion along the $j$-th column (Proof)**

Note that $\qquad C_{ij}(A^T) = C_{ji}(A).$

Hence, if we write $A^T = (b_{ij})$, then

$$\begin{aligned} \det(A) &= \det(A^T) \\ &= b_{j1}\, C_{j1}(A^T) + \ldots + b_{jn}\, C_{jn}(A^T) \\ &= a_{1j}\, C_{1j}(A) + \ldots + a_{nj}\, C_{nj}(A) \end{aligned}$$

This is the expansion of $\det(A)$ along $j$-th column of $A$.

**Applications: 1. Computing $A^{-1}$**

If $C = (C_{ij})$ : cofactor matrix of $A$, then $\boxed{A\, C^T = \det(A)\, I}$

i.e., $\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & \ldots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \ldots & C_{nn} \end{bmatrix} = \begin{bmatrix} \det(A) & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \det(A) \end{bmatrix}$

*Proof.* We have seen that $a_{i1}C_{i1} + \ldots + a_{in}C_{in} = \det(A)$. Now $a_{11}C_{21} + a_{12}C_{22} + \ldots + a_{1n}C_{2n}$

$= \det \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ a_{11} & \ldots & a_{1n} \\ a_{31} & \ldots & a_{3n} \\ \cdot & & \cdot \\ a_{n1} & \ldots & a_{nn} \end{bmatrix} = 0.$ Similarly, if $i \neq j$, then $a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots + a_{in}C_{jn} = 0.$ $\square$

REMARK. If $A$ is invertible, then $A^{-1} = \dfrac{1}{\det(A)} C^T$ .
For $n \geq 4$, this is *not* a good formula to find $A^{-1}$. Use elimination to find $A^{-1}$ for $n \geq 4$. *This formula is of theoretical importance.*

**Applications: 2. Solving $Ax = b$**

CRAMER'S RULE: If $A$ is invertible, the $Ax = b$ has a unique solution.

$$x = A^{-1}b = \frac{1}{\det(A)} C^T b = \frac{1}{\det(A)} \begin{bmatrix} C_{11} & \cdots & C_{n1} \\ \vdots & & \vdots \\ C_{1n} & \cdots & C_{nn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

Hence $x_j = \dfrac{1}{\det(A)}(b_1 C_{1j} + b_2 C_{2j} + \cdots + b_n C_{nj}) = \dfrac{1}{\det(A)} \det(B_j)$, where $B_j$ is obtained by replacing the $j^{\text{th}}$ column of $A$ by $b$, and $\det(B_j)$ is computed along the $j^{\text{th}}$ column.

REMARK: For $n \geq 4$, use elimination to solve $Ax = b$. Cramer's rule is of theoretical importance.

**Applications: 3. Volume of a box**

$\boxed{\text{Assume the rows of } A \text{ are mutually orthogonal}}$. Then

$$AA^T = \begin{pmatrix} A_{1*} \\ \vdots \\ A_{n*} \end{pmatrix} \begin{pmatrix} (A_{1*})^T & \ldots & (A_{n*})^T \end{pmatrix} = \begin{pmatrix} l_1^2 & & 0 \\ & \ddots & \\ 0 & & l_n^2 \end{pmatrix}$$

where $l_i = \sqrt{(A^i)^T \cdot A^i}$ is the length of $A^i$. Since $\det(A) = \det(A^T)$, we get $\boxed{|\det(A)| = l_1 \cdots l_n}$.

Since the edges of the box spanned by rows of $A$ are at right angles, the volume of the box

$$= \text{ the product of lengths of edges} = |\det(A)|$$

.

**Applications: 4. A Formula for Pivots**

OBSERVATION: If row exchanges are not required, then the first $k$ pivots are determined by the top-left $k \times k$ submatrices $\widetilde{A}_k$ of $A$.

EXAMPLE. If $A = [a_{ij}]_{3\times 3}$, then $\widetilde{A}_1 = (a_{11})$, $\widetilde{A}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\widetilde{A}_3 = A$.

Assume the pivots are $d_1, \ldots, d_n$, obtained without row exchange. Then

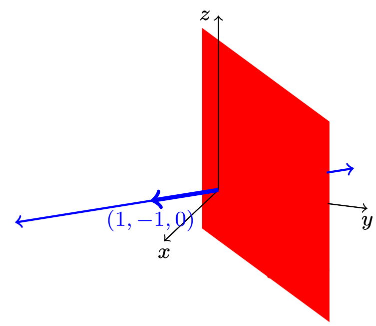
- $\det(\widetilde{A}_1) = a_{11} = d_1$
- $\det(\widetilde{A}_2) = d_1\, d_2 = \det(A_1)\, d_2$
- $\det(\widetilde{A}_3) = d_1\, d_2\, d_3 = \det(A_2)\, d_3$ *etc.*,
- If $\det(\widetilde{A}_k) = 0$, then we need a row exchange in elimination.
- Otherwise the $k$-th pivot is $\boxed{d_k = \det(\widetilde{A}_k)/\det(\widetilde{A}_{k-1})}$

# Appendix II - Orthogonal Subspaces

**Orthogonal Subspaces: Examples**

If $\{w_1, \ldots, w_r\}$ is a orthogonal set then $w_r$ is orthogonal to Span$\{w_1, \ldots, w_{r-1}\}$. Then every vector in Span$\{w_r\}$ is orthogonal to every vector in Span$\{w_1, \ldots, w_{r-1}\}$.

EXAMPLE: Let $u = (1, 1, 0)^T$, $v = (0, 0, 1)^T$, and $w = (1, -1, 0)^T$,
**P** $=$ Span$\{u, v\}$, and **L** = Span$\{w\}$.




OBSERVE: For scalars $a$, $b$, $t$, $(tw)^T(au + bv) = 0$.
The subspace **L** = $\{(t, -t, 0) \mid t \in \mathbb{R}\}$ is orthogonal to the subspace **P** = $\{(a, a, b) \mid a, b \in \mathbb{R}\}$.
OBSERVE: dim(**L**) + dim(**P**) = dim($\mathbb{R}^3$).

**Orthogonal Subspaces: Definition and Properties**

DEFN. (Orthogonal Subspaces) Let $V$ and $W$ be subspaces of $\mathbb{R}^n$. We say $V$ and $W$ are orthogonal to each other (notation: $V \perp W$) if every vector in $V$ is orthogonal to every vector in $W$, i.e., $\boxed{\text{for every } v \in V \text{ and } w \in W,\ v \cdot w = v^T w = 0.}$

NOTE: Let $V =$ Span of $\{v_1, \ldots, v_r\}$, $W =$ Span $\{w_1, \ldots, w_s\}$.
If $v_i^T w_j = 0$ for all $i, j$, then $V \perp W$.

*Proof.* If $v = a_1 v_1 + \cdots + a_r v_r \in V$, $w = b_1 w_1 + \cdots + b_s w_s \in W$, then

$$\begin{aligned} v^T w &= (a_1 v_1^T + \cdots + a_r v_r^T)(b_1 w_1 + \cdots + b_s w_s) \\ &= a_1 b_1 v_1^T w_1 + a_1 b_2 v_1^T w_2 + \cdots + a_1 b_s v_1^T w_s \\ &\quad + a_2 b_1 v_2^T w_1 + \cdots + a_2 b_s v_2^T w_s + \cdots + a_r b_s v_r^T w_s = 0 \text{ using bilinearity.} \end{aligned}$$

**Orthogonal Subspaces: Examples**

**Q.** Let $V$ be the $yz$-plane and $W$ be the $xz$-plane.
Are $V$ and $W$ orthogonal to each other (as subspaces of $\mathbb{R}^3$)?
**A.** No. The vector $e_3 = (0, 0, 1)$ lies in $V$ and $W$ both and $e_3^T e_3 \neq 0$.

REMARK: If $V$ and $W$ are orthogonal subspaces of $\mathbb{R}^n$, then $V \cap W = 0$.
This is a necessary condition.

**Q.** Is it a sufficient condition? **A.** No.

EXAMPLE 1: If $V = \text{Span}\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}$, and $W = \text{Span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}$. Then $V \cap W = 0$,

but $V$ and $W$ are not orthogonal, since $\begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} = 1 \neq 0$.

## Orthogonal Subspaces: Examples

EXAMPLE 2: Let $v_1 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}$, $v_2 = \begin{pmatrix}0\\1\\1\\0\end{pmatrix}$ and $w = \begin{pmatrix}1\\-1\\1\\0\end{pmatrix}$.

If $V =$ Span $\{v_1, v_2\}$, and $W =$ Span$\{w\}$, then $V \perp W$.

*Sketch of Proof.* It is enough to see that $v_1, v_2 \in V$ are orthogonal to $w \in W$.

COMPARE: In Example 1, dim $(L) +$ dim $(P) =$ dim $(\mathbb{R}^3)$.
In Example 2, dim $(V) +$ dim $(W) = 3 <$ dim $(\mathbb{R}^4)$.

**Q:** Can we enlarge $W$ to $W' \subseteq \mathbb{R}^4$ such that $V \perp W'$ and dim $(V) +$ dim $(W') =$ dim $(\mathbb{R}^4)$?

**A:** Yes!

OBSERVE: $(0, 0, 0, 1)$ is orthogonal to both $V$ and $W$.

Let $A = \begin{pmatrix}v_1^T\\v_2^T\end{pmatrix} = \begin{pmatrix}1 & 1 & 0 & 0\\0 & 1 & 1 & 0\end{pmatrix}$. Then $C(A^T) = V$.

Note that $x \in N(A) \Leftrightarrow v_1^T x = 0 = v_2^T x. \Rightarrow (1, -1, 1, 0)^T \in N(A) \Rightarrow W \subset N(A)$.

By the Rank-Nullity Theorem, rank$(A)+$ dim$(N(A)) = 4$. Since rank$(A) = 2$, we get dim$(N(A)) = 2$. Therefore if $W' = N(A)$, then $V \perp W'$ and dim $(V) +$dim $(W') = 4$.

By inspection, $W' =$ Span$\{w_1 = (1, -1, 1, 0)^T, w_2 = (0, 0, 0, 1)^T\}$ is orthogonal to $V$.
The dimension of the two spaces sum up to dim $(\mathbb{R}^4)$.

## The Four Fundamental Spaces and Orthogonality

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

1. The row space of $A$, $C(A^T)$ is orthogonal to $N(A)$.
2. $C(A)$ is orthogonal to the left nullspace of $A$, $N(A^T)$.

In fact, $N(A) =$ set of all vectors orthogonal to $C(A^T)$,
and $N(A^T) =$ set of all vectors orthogonal to $C(A)$

EXAMPLE. Let $A = \begin{pmatrix}1 & 2\\2 & 4\\3 & 6\end{pmatrix}$. Then $A$ has 1 pivot $\Rightarrow$ dim $(C(A)) = 1 =$ dim $(C(A^T)) =$.

OBSERVE: $C(A^T)$ = Span$\{(1, 2)^T\}$, $N(A)$ = Span$\{(-2, 1)^T\}$,
$C(A) =$ Span$\{(1, 2, 3)^T\}$, and
$N(A^T)$ is the plane $y_1 + 2y_2 + 3y_3 = 0$.
Moreover, dim$(C(A^T))$+ dim$(N(A)) = 2 =$ dim $(\mathbb{R}^2)$,
dim$(C(A))$+ dim$(N(A^T)) = 3 =$ dim $(\mathbb{R}^3)$.

## Fundamental Theorem of Orthogonality

DEFN. Let $W$ be a subspace of $\mathbb{R}^n$.
Define its orthogonal complement as $W^\perp = \{v \in \mathbb{R}^n \mid v^T w = 0 \text{ for all } w \in W\}$.

*Claim.* $W^\perp$ is a subspace of $\mathbb{R}^n$.
If $v_1, v_2 \in W^\perp$ and $w \in W$, then $v_1^T w = 0 = v_2^T w$.
Hence for $c_1, c_2 \in \mathbb{R}$, $(c_1 v_1 + c_2 v_2)^T w = c_1 v_1^T w + c_2 v_2^T w = 0. \Rightarrow c_1 v_1 + c_2 v_2 \in W^\perp$.

THEOREM (Fundamental Theorem of Orthogonality )
Let $A$ be an $m \times n$ matrix.

1. The row space of $A$ = orthogonal complement of $N(A)$.
2. The column space of $A$ = orthogonal complement of left nullspace $N(A^T)$.

i.e., $C(A^T) = N(A)^\perp$, and $C(A) = N(A^T)^\perp$.

## Orthogonal Complements

THEOREM (Orthogonal Complement)
Given a subspace $W \subseteq \mathbb{R}^n$, $\dim(W) + \dim(W^\perp) = n$.

*Proof.* Let $v_1, \ldots, v_r$ be a basis of $W$. Let $A$ be a matrix with rows $v_1, \ldots, v_r$.
Then $\text{rank}(A) = r$ and $W = C(A^T)$. $W^\perp = N(A)$ is of dimension $n - r$.
This proves the theorem.

OBSERVE: Let $V = \text{Span}\left\{v_1 = (1,1,0,0)^T, v_2 = (0,1,1,0)^T\right\}$
and $W = \text{Span}\{w_1 = (1,-1,1,0)^T, w_2 = (0,0,0,1)^T\}$.
- $\{v_1, v_2\}$, $\{w_1, w_2\}$ are bases for $V$ and $W$ respectively.
- $V \perp W \Rightarrow V \subseteq W^\perp$ and $\dim(V) + \dim(W) = 4 = \dim(\mathbb{R}^4)$
$\Rightarrow V = W^\perp$.
- $V \cap W = 0 \Rightarrow \mathcal{B} = \{v_1, v_2, w_1, w_2\}$ is linearly independent $\Rightarrow$
$\mathcal{B}$ is a basis of $\mathbb{R}^4 \Rightarrow V + W = \{v + w \mid v \in V, w \in W\} = \mathbb{R}^4$,
and for every $x \in \mathbb{R}^4$, $\boxed{\exists \text{ unique } v \in V = W^\perp \text{ and } w \in W \text{ such that } x = v + w.}$

## Orthogonal Complements

TO SUMMARISE: If $W$ is a subspace of $\mathbb{R}^n$ with basis $\mathcal{B}$, and $\mathcal{B}'$ is a basis of $W^\perp$,
then $\mathcal{B} \cup \mathcal{B}'$ is a basis of $\mathbb{R}^n$, $W \cap W^\perp = 0$, $\dim(W) + \dim(W^\perp) = n$, $W + W^\perp = \mathbb{R}^n$,
and for every $x \in \mathbb{R}^n$, $\boxed{\exists \text{ unique } w_1 \in W \text{ and } w_2 \in W^\perp \text{ such that } x = w_1 + w_2.}$

SOME CONSEQUENCES: Let $A$ be $m \times n$ of rank $r$. Then

1. $C(A^T) \cap N(A) = 0$ and $\mathbb{R}^n = C(A^T) + N(A)$.
Similarly $C(A) \cap N(A^T) = 0$ and $\mathbb{R}^m = C(A) + N(A^T)$.

2. If $x \in \mathbb{R}^n$, there is a unique expression $x = x_r + x_n$, where $x_r \in C(A^T)$, $x_n \in N(A)$.
Hence $Ax = Ax_r \in C(A)$. Thus

$$\boxed{\text{The matrix } A \text{ transforms its row space } C(A^T) \text{ into its column space } C(A).}$$

# Appendix III - Additional Topics

**Introduction**

Topics included (in no particular order) are:

1. An Application of $QR$-factorization
2. Angles in $\mathbb{R}^n$
3. Inner Product Spaces
4. Symmetric Matrices & SVD

## III.1 An Application of $QR$-Factorization

**Application: Least squares + QR**

Let $A$ be $n \times r$ matrix of rank $r$.

Let $A = QR$, where $Q$ is a $n \times r$ matrix with orthonormal columns and $R$ is a $k \times k$ invertible upper triangular matrix.

Recall that $Q^T Q = \begin{pmatrix} q_1^T q_1 & \dots & q_1^T q_r \\ \\ q_r^T q_1 & \dots & q_r^T q_r \end{pmatrix} = I_r$.

Let $Ax = b$ be an inconsistent system.

We apply least squares method to get best approximate solution.

$A^T A = (QR)^T QR = R^T Q^T QR = R^T R$. Therefore we solve

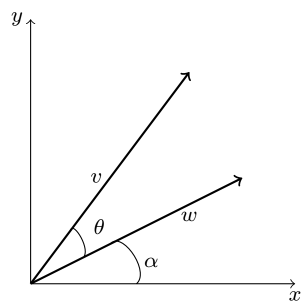
$A^T A\hat{x} = R^T R\hat{x} = A^T b = R^T Q^T b$. Since $R^T$ is invertible,

Thus least square solution is given by $\boxed{R\hat{x} = Q^T b}$.

## III.2 Angles in $\mathbb{R}^n$

**Angles in $\mathbb{R}^2$**

Note: $w \in \mathbb{R}^2$ can be written as $(||w|| \cos\alpha, ||w|| \sin\alpha)$, where




$$v = \|v\| \begin{bmatrix} \cos(\theta+\alpha) \\ \sin(\theta+\alpha) \end{bmatrix}$$
$$w = \|w\| \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}$$

$$\begin{aligned} v^T w &= \|v\|\|w\|(\cos(\theta+\alpha)\sin\alpha + \sin(\theta+\alpha)\cos\alpha) \\ &= \|v\|\|w\|\cos(\theta+\alpha-\alpha) = \|v\|\|w\|\cos\theta \end{aligned}$$

This gives a formula for the angle between two vectors in terms of the inner product and norm (length) of vectors.

**Angles in $\mathbb{R}^n$**

To generalise the angle between two vectors to $\mathbb{R}^n$, we use the notion of inner product and norm in $\mathbb{R}^n$. But to do this, we need the Cauchy-Schwartz inequality for $v, w \in \mathbb{R}^n$:

$$|v^T w| \leq \|v\|\|w\|$$

To verify this expand the inequality $\left\| w - \dfrac{v^T w}{\|v\|\|w\|} v \right\|^2 \geq 0$.

Define the angle between any two vectors $v, w \in \mathbb{R}^n$ as

$$\theta = \cos^{-1}\left(\frac{v^T w}{\|v\|\|w\|}\right).$$

THUS: $\theta$ is $90^\circ$ if and only if $v^T w = 0$.

EXAMPLE Angle between $(0,1,1)$ and $(0,2,0)$ is $\cos^{-1}\left(\frac{2}{2\sqrt{2}}\right) = 45^\circ$.

## III.3 General Inner Product Spaces

**Inner Product**

We have defined orthogonality only in $\mathbb{R}^n$.
Can we do something similar for other vector spaces?

If $V$ is $n$-dimensional, we can use coordinate vectors and the inner product in $\mathbb{R}^n$.
POINT TO PONDER:
Can different bases and coordinate vectors be used to define new inner products on $\mathbb{R}^n$?

NOTE: On the infinite dimensional space $\mathbb{R}^\infty$, the space of all real sequences,

the dot product does not generalise.
If $x_n = \frac{1}{\sqrt{n}}$, then $(x_n) \cdot (x_n) = \Sigma_n \frac{1}{n}$ which does not converge, and hence is not finite.

However, this generalization works for sequences $(x_n)$ and $(y_n)$ if $\Sigma_n (x_n y_n)$ is convergent.

Define $\ell_2(\mathbb{R}) = \{(x_1, \ldots, x_n, \ldots) \in \mathbb{R}^\infty \mid \Sigma_i x_i^2 < \infty\}$. E.g., $(1, 1/2, 1/3, \ldots, 1/n, \ldots) \in \ell_2(\mathbb{R})$.
Which of the following are in $\ell_2(\mathbb{R})$? $(1, 0, \ldots, 0, \ldots), (1, 0, 10, 1, 0, \ldots), (1, 1/2, 1/2^2, 1/2^3, \ldots)$

**Inner Product Spaces : Definition**

If $(x_n), (y_n) \in \ell_2(\mathbb{R})$ then define $\langle (x_n), (y_n) \rangle = \Sigma_n x_n y_n$.

Why is this bounded? $(x_k - y_k)^2 \geq 0 \implies x_k y_k \leq \dfrac{1}{2}(x_k^2 + y_k^2)$.

Note that, similar to the usual dot product on $\mathbb{R}^n$, $\langle \_ , \_ \rangle$ has the following properties:
For all $x = (x_n), y = (y_n), z = (z_n) \in \ell_2(\mathbb{R})$ and $c \in \mathbb{R}$.

- (Symmetry) $\langle x\,, y\rangle = \langle y\,, x\rangle$
- (Positive Definiteness) $\langle x\,, x\rangle \geq 0$ and $\langle x\,, x\rangle = 0 \iff x = 0$.
- (Bilinearity) $\langle x + y\,, z\rangle = \langle x\,, z\rangle + \langle y\,, z\rangle$ and $\langle cx\,, y\rangle = c\langle x\,, y\rangle = \langle x\,, cy\rangle$.

An inner product on a vector space $V$ is one which gives a real number $\langle u\,, v\rangle$ for every pair $u, v \in V$ such that it has the three above-mentioned properties.

This helps us define length of a vector in such spaces

Define length (norm) of $v$ as $\|v\| = \sqrt{\langle v\,, v\rangle}$.

POINT TO PONDER:

What would one need to define angle between two non-zero vectors in $V$?

**Inner Product Spaces : Properties**

Let $V$ be a vector space with an inner product.

Then it satisfies the Cauchy-Schwartz inequality, that is, for all $u, v \in V$,

$$|\langle u\,, v\rangle| \leq \|u\|\|v\|,$$

and the Pythagoras theorem, that is, for $u, v \in V$,

$$\langle u\,, v\rangle = 0 \iff \|u - v\|^2 = \|u\|^2 + \|v\|^2.$$

We define two vectors $u, v \in V$ to be orthogonal to each other if $\langle u\,, v\rangle = 0$.

**Inner Product spaces : Another Example**

Let $V = \mathcal{C}[0, 2\pi]$ be the vector space of continuous functions on $[0, 2\pi]$.

We now define an inner product on $V$ as follows: For $f,\ g \in \mathcal{C}[0, 2\pi]$

$$\langle f\,, g\rangle = \int_0^{2\pi} f(t)g(t)\ dt.$$

This satisfies all the axioms of an inner product.

Some of the well understood functions are the cosine and sine functions on $[0, 2\pi]$, and it is often useful to write periodic functions in terms of sines and cosines.

This idea is a simple application of taking orthogonal projections in the vector space $\mathcal{C}[0, 2\pi]$ onto the subspace $W = \text{Span}\{\cos nx, \sin mx \mid m, n \in \mathbb{Z}, m, n \geq 0\}$.

In order to do take orthogonal projections we need an orthogonal basis of $W$.

**An Application: Fourier Series**

The set $S_n = \{\cos(kx), \sin(lx) \mid 0 \leq k, l \leq n\}$ is orthogonal. Why?

For example, $\langle \cos x\,, \sin x\rangle = \int_0^{2\pi} \cos t \sin t\ dt = \dfrac{1}{2}\int_0^{2\pi} \sin 2t\ dt = 0$.

$\langle \cos x\,, \cos x\rangle = \int_0^{2\pi} \cos^2 t\ dt = \dfrac{1}{2}\int_0^{2\pi}(1 - \cos 2t)\ dt = 2\pi$.

Then $S_n$ is an orthogonal basis of $W_n = \text{Span}(S_n) \subseteq W$. Let $f \in \mathcal{C}[0, 2\pi]$. Then

$$f_n = \text{proj}_{W_n}(f(x)) = a_0 + a_1\cos(x) + b_1\sin(x) + \ldots + a_n\cos(nx) + b_n\sin(nx)$$

where $a_0 = \dfrac{\langle 1\,, f(x)\rangle}{2\pi}$, $a_k = \dfrac{\langle \cos(kx)\,, f(x)\rangle}{2\pi}$, $b_k = \dfrac{\langle \sin(kx)\,, f(x)\rangle}{2\pi}$, for $1 \leq k \leq n$.

Thus $(f_n)$ is exactly the sequence of partial sums of the Fourier series expansion of $f$.

## III.4 Symmetric Matrices & SVD

### (Real) Symmetric Matrices

• Let $A$ be a symmetric $n \times n$ matrix. Then all its eigenvalues are real.

*Proof.* Let $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n$ be eigenvalue and eigenvector of $A$.
Then $Av = \lambda v \Rightarrow (\bar{A}\bar{v})^T = \bar{\lambda}\bar{v}^T$.
Now $\bar{v}^T A v = \lambda \bar{v}^T v$. Alternately, $A = A^T$ and $\bar{A}^T = A^T$ implies $\bar{v}^T A v = \bar{v}^T \bar{A}^T v = \bar{\lambda}\bar{v}^T v$.

Since $\bar{v}^T v > 0$ for a $v$ non-zero, we have, $\lambda = \bar{\lambda} \Rightarrow \lambda$ is real.

• $\lambda$ has an associated eigenvector with real entries.
*Proof.* If $x \notin \mathbb{R}^n$, write $x = u + iv$, where $u$, $v \in \mathbb{R}^n$.
Then $\lambda \in \mathbb{R} \Rightarrow u$ and $v$ are also eigenvectors associated to $\lambda$.

• If $\lambda_1$ and $\lambda_2$ are distinct eigenvalues of $A$, with associated eigenvectors $v_1$ and $v_2 \in \mathbb{R}^n$.
Then $v_1$ and $v_2$ are orthogonal.
*Proof.* Want to prove: $v_1^T v_2 = 0$. Now

$$\lambda_1(v_1^T v_2) = (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 = (v_1^T A^T)v_2 = v_1^T(Av_2) = v_1^T(\lambda_2 v_2) = \lambda_2(v_1^T v_2).$$

Since $\lambda_1 \neq \lambda_2$, we have $v_1^T v_2 = 0$.

SPECTRAL THEOREM FOR A REAL SYMMETRIC MATRIX:

$$\boxed{\text{Every real symmetric matrix } A \text{ can be diagonalised.}}$$

In particular, there is an orthogonal matrix $Q$ and a diagonal matrix $\Lambda$ such that

$$A = Q\Lambda Q^T.$$

Given that $A$ can be diagonalised, then $A = Q\Lambda Q^T$ follows from the previous slide , together with the Gram-Schmidt process applied to each eigenspace of $A$.

### Application: Singular Value Decomposition (SVD)

SVD: Given an $m \times n$ matrix $A$, there exists an $m \times n$ "diagonal" matrix $\Sigma$, orthogonal matrices $U$ ($m \times m$) and $V$ ($n \times n$), such that $A = U\Sigma V^T$.

NOTE: If $U$, $V$ and $\Sigma$ are as above, they satisfy:
$AA^T = (U\Sigma V^T)(V\Sigma^T U^T) = U\Sigma^2 U^T$, and $A^T A = V\Sigma^2 V^T$. Thus:
1. The non-zero diagonal entries in $\Sigma$, called the singular values of $A$, are the square roots of the common eigenvalues of $AA^T$ and $A^T A$.
2. Columns of $U$ are eigenvectors of $AA^T$, and those of $V$ are eigenvectors of $A^T A$.
3. If the first $r$ columns of $\Sigma$ are non-zero, then $\{U_{*1}, \ldots, U_{*r}\}$ and $\{V_{*1}, \ldots, V_{*r}\}$ are orthonormal bases for the column space of $A$, $C(A)$ and the row space of $A$, $C(A^T)$ respectively. Furthermore, $AV = U\Sigma \Rightarrow AV_{*j} = \sigma_j U_{*j}$ for each $j \leq r$.

SIMPLE-MINDED APPLICATION: Can be used in image compression,
e.g., If an image is represented by $A$, find $U$, $\Sigma$, $V$ as in SVD. Replace "small" singular values in $\Sigma$ by $0$ to $\widetilde{\Sigma}$. Then $\widetilde{A} = U\widetilde{\Sigma}V^T$ can be used to represent the compressed image.

$\sim\sim\sim$ *fin* $\sim\sim\sim$