دانشگاه صنعتی خواجه نصیرالدین طوسی

جبر خطی
(مباحث اولیه)

استاد : دکتر محمدرضا ملک
دستیار : امیرحسام طاهرزادگانی
دانشکده مهندسی ژئوماتیک و ژئودزی

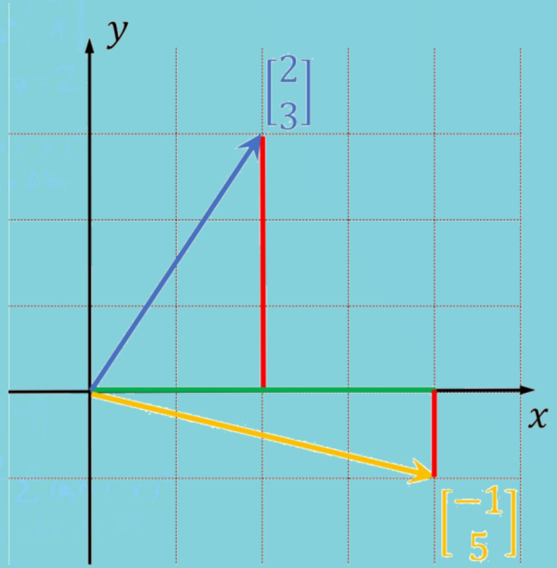
From scalar to tensor

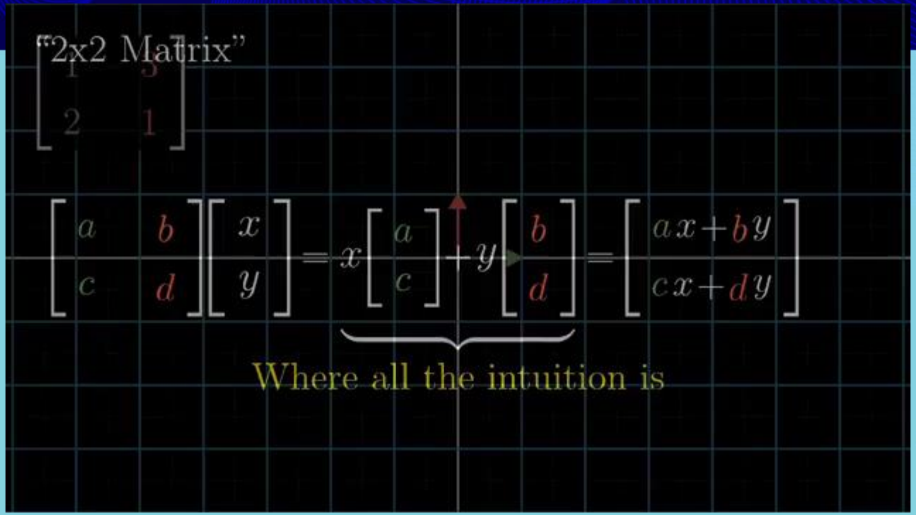
Scalars are single numbers and are an example of a 0th-order tensor. In mathematics it is necessary to describe the set of values to which a scalar belongs. The notation $x \in \mathbb{R}$ states that the (lowercase) scalar value x is an element of (or member of) the set of real-valued numbers, R.

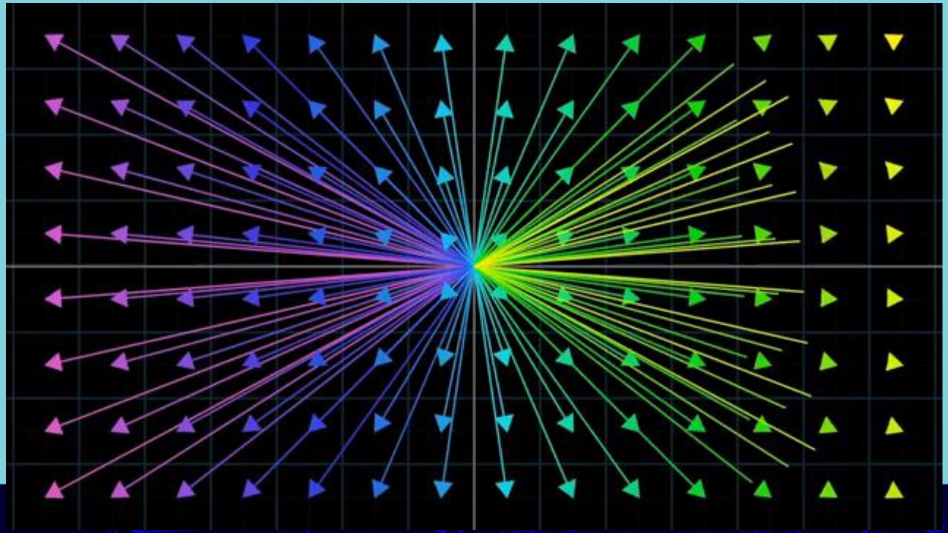
Vectors are ordered arrays of single numbers and are an example of 1st-order tensor. Vectors are members of objects known as **vector spaces**. The three-dimensional real-valued vector space, denoted by $\mathbb{R}^3$ is often used to represent our real-world notion of three-dimensional space mathematically.

Matrices are rectangular arrays consisting of numbers and are an example of 2nd-order tensors. If m and n are positive integers, that is $m, n \in N$ then the $m \times n$ matrix contains mn numbers, with m rows and n columns. If all of the scalars in a matrix are real-valued then a matrix is denoted with uppercase boldface letters, such as $A \in R^{m \times n}$. That is the matrix lives in a m×n-dimensional **real-valued** vector space.

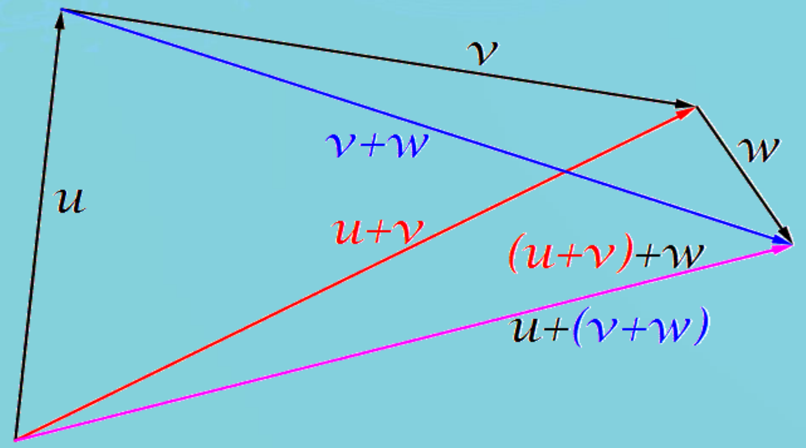
Geometric intuition

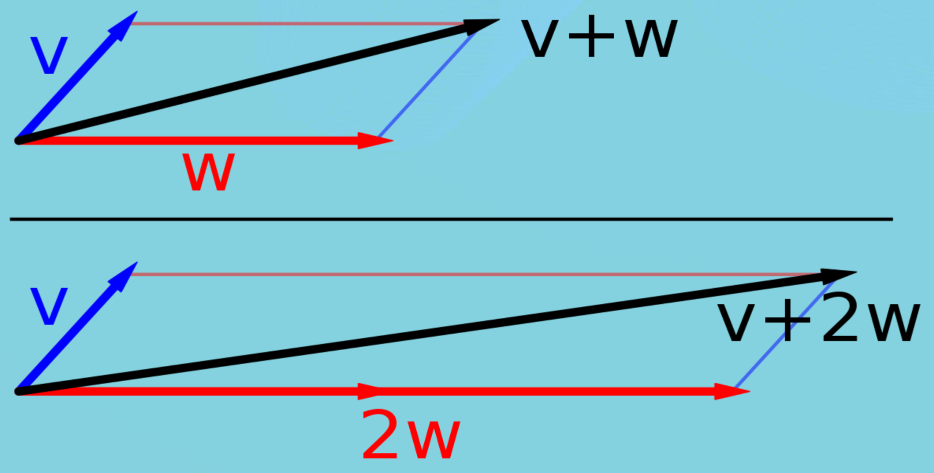






Algebraic operations





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

$$2 \cdot \begin{bmatrix} 10 & 6 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} 2 \cdot 10 & 2 \cdot 6 \\ 2 \cdot 4 & 2 \cdot 3 \end{bmatrix}$$

Adding the vectors (+ operator)

Multiplying by an scalar (. operator)

Identity matrix

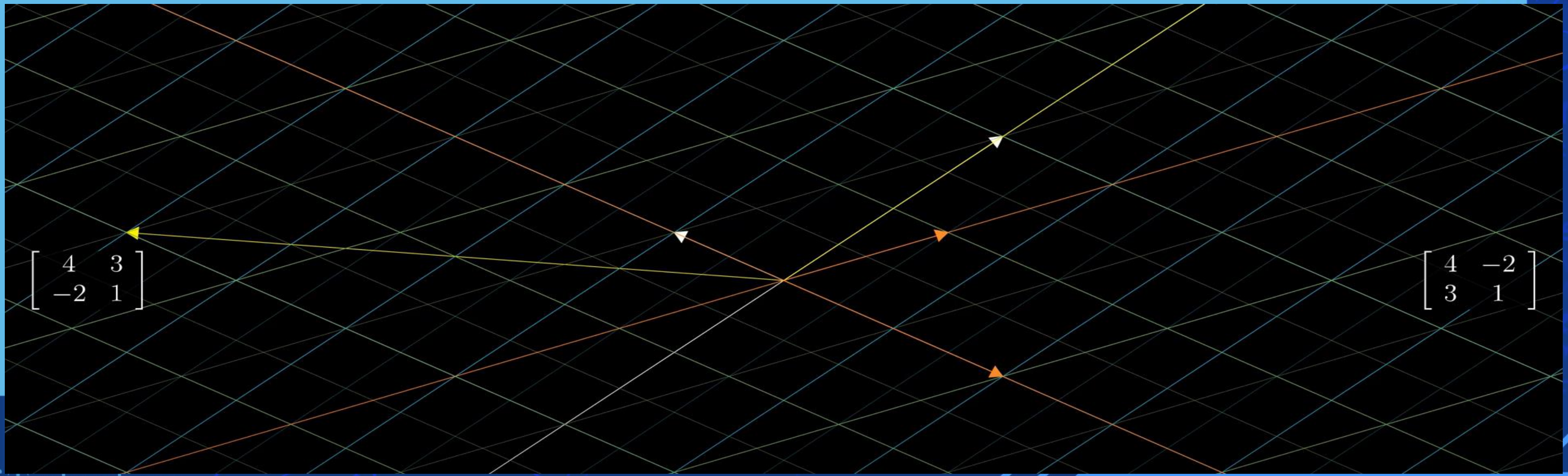
- The ith column of an identity matrix is the unit vector e_i (the vector whose ith entry is 1 and 0 elsewhere) It follows that the determinant of the identity matrix is 1, and the trace is n.

$$I_1 = \begin{bmatrix} 1 \end{bmatrix},\ I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \dots,\ I_n = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Transpose

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \\ 4 & 8 & 12 \end{bmatrix}^{\mathrm{T}}$$

The transpose of a matrix is simply a flipped version of the original matrix. We can transpose a matrix by switching its rows with its columns. We denote the transpose of matrix A by A^T .



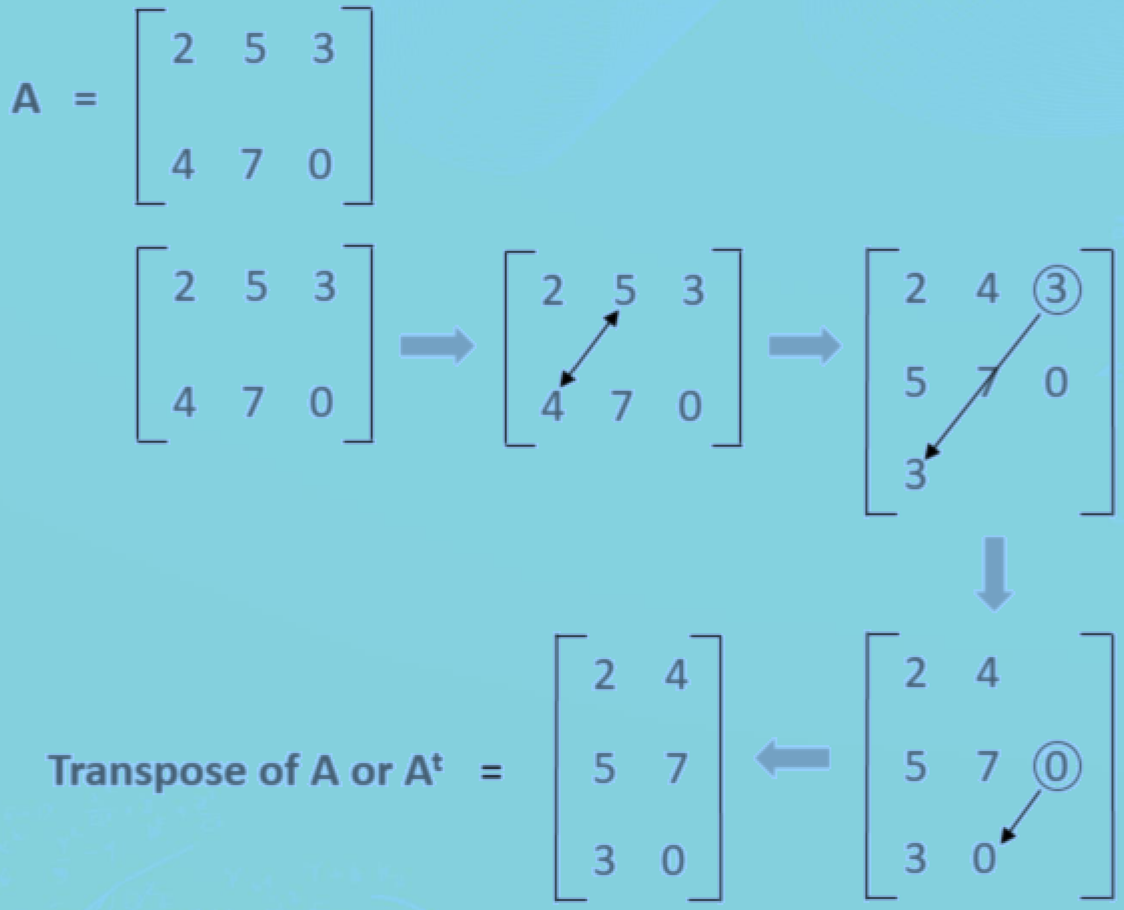
Transpose

$$(A+B)^T = B^T + A^T$$

$$(AB)^T = B^T A^T$$

$$(kA)^T = kA^T$$

$$(A^T)^T = A$$



K. N. Toosi University of Technology
Geodesy & Geomatics Faculty

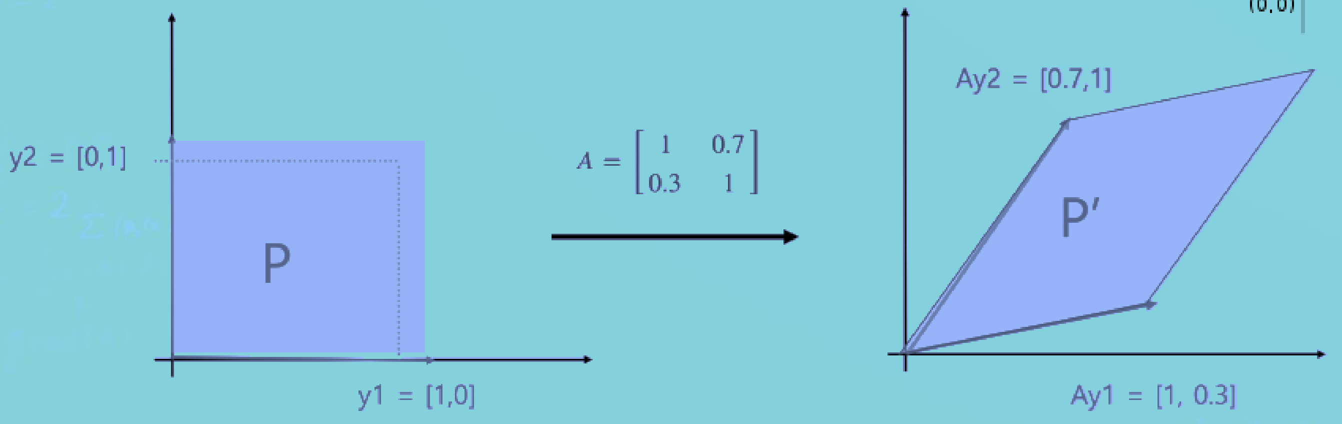
Determinant

- The determinant of a matrix is a special scalar value that can be computed from a square matrix. Given a square matrix $\boldsymbol{A}$, its determinant is often denoted as $\mathbf{det}(\boldsymbol{A})$, $\mathbf{det}\,\boldsymbol{A}$ or $|\boldsymbol{A}|$.
- From a geometrical perspective, it can be viewed as the v**olume scaling factor** of the linear transformation described by the matrix $\boldsymbol{A}$ for unit volume in a given space. The $\mathbf{det}(\boldsymbol{A})$ is the signed volume of the n -dimensional **parallelepiped** spanned by the column or row vectors of the matrix.
- Given n linearly independent vectors $y_1, y_2, \ldots, y_n$, a set **P** in below figure is called **parallelepiped**.

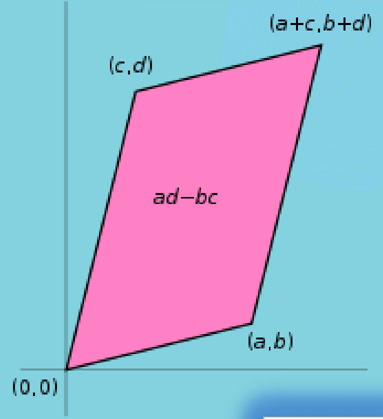
$$P = \left\{ \sum_{j=1}^{n} \lambda_j y_j \mid 0 \leq \lambda_j \leq 1 \right\}$$

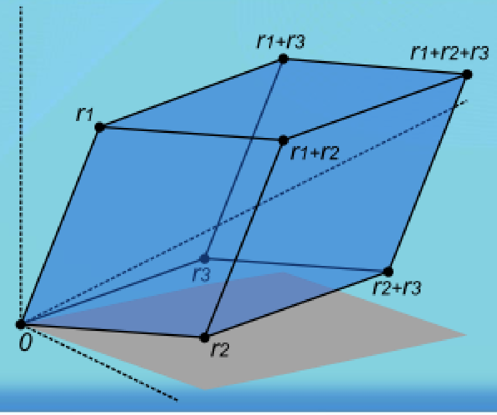
Geometric intuition of determinant

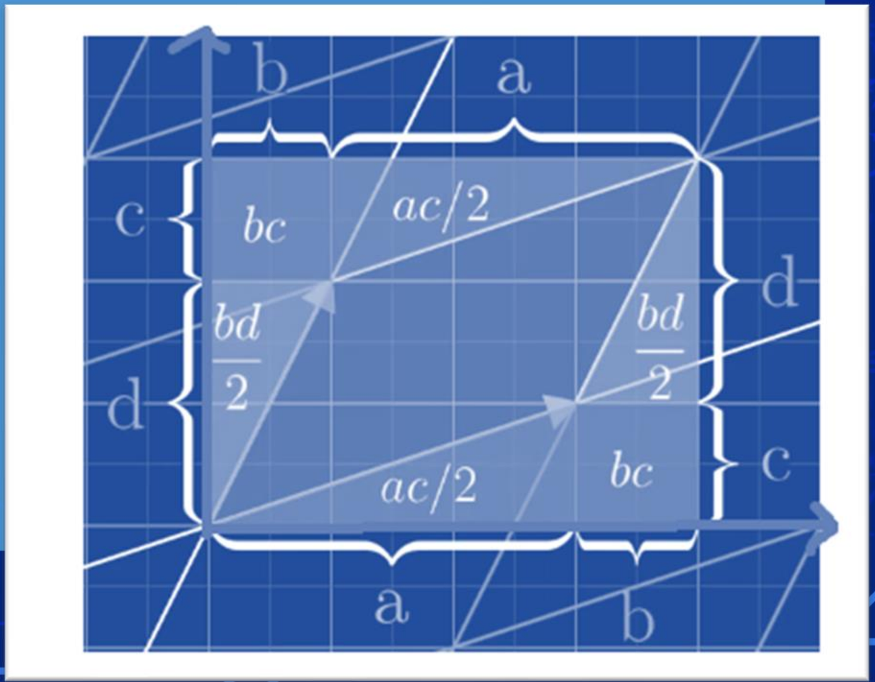
2D Vector Space V



$det(A) = 1 \times 1 - 0.3 \times 0.7 = 0.79$







K. N. Toosi University of Technology
Geodesy & Geomatics Faculty

Determinant

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

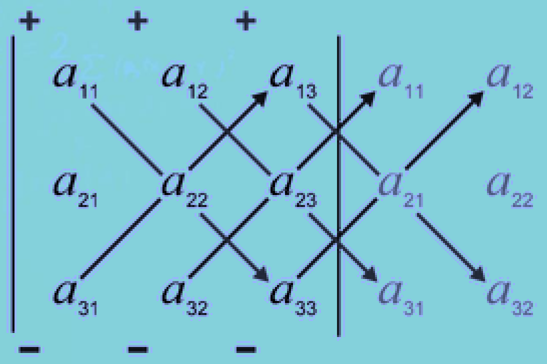
$$|A| = \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = a\begin{vmatrix} e & f \\ h & i \end{vmatrix} - b\begin{vmatrix} d & f \\ g & i \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

$$= aei + bfg + cdh - ceg - bdi - afh.$$

Leibniz formula

$$\det(A) = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma_i} \right)$$

Further Studies:
Finding determinant using permutations(Leibniz Formula)



Saruss rule

Click to download pdf

Determinant properties

1. Switching two rows or columns changes the sign.
2. Scalars can be factored out from rows and columns.
3. Multiples of rows and columns can be added together without changing the determinant's value.
4. Scalar multiplication of a row by a constant c multiplies the determinant by c.
5. A determinant with a row or column of zeros has value 0.
6. Any determinant with two rows or columns equal has value 0

Determinant properties

$$(1)\ det(I_n) = 1,\ where I_n\ is\ the\ n \times n\ identity\ matrix$$

$$(2)\ det(A^T) = det(A)$$

$$(3)\ det(A^{-1}) = \frac{1}{det(A)}$$

$$(4)\ det(AB) = det(A) \times det(B)$$

$$(5)\ det(cA) = c^n \times det(A)$$

$$(6)\ the\ determinant\ of\ a\ trianglular\ matrix\ A\ is\ the\ product\ of\ diagonal\ entries$$

$$det(A) = a_{1,1} \times a_{2,2} \times \ldots \times a_{n,n} = \prod_{i=1}^{n} a_{i,i}$$

$$(7)\ det(ca_1, a_2, \ldots, a_n) = c\,det(a_1, a_2, \ldots, a_n)$$

$$(8)\ det(a_1, a_2, \ldots, a_n) + det(a'_1, a_2, \ldots, a_n) = det(a_1 + a'_1, a_2, \ldots, a_n)$$

$$(9)\ det(a_1, a_2, .\, a_i + c \times a_j .\,., a_n) = det(a_1, a_2, \ldots, a_i, \ldots, a_n),\ where\ 1 \le i, j \le n, i \ne j$$

Matrix Multiplication

$$\begin{array}{c|c} & \begin{array}{cc} \vec{b_1} & \vec{b_2} \\ \downarrow & \downarrow \end{array} \\ \begin{array}{c} \vec{a_1} \rightarrow \\ \vec{a_2} \rightarrow \end{array} \begin{bmatrix} 1 & 7 \\ 2 & 4 \end{bmatrix} & \cdot \begin{bmatrix} 3 & 3 \\ 5 & 2 \end{bmatrix} = \begin{bmatrix} \vec{a_1} \cdot \vec{b_1} & \vec{a_1} \cdot \vec{b_2} \\ \vec{a_2} \cdot \vec{b_1} & \vec{a_2} \cdot \vec{b_2} \end{bmatrix} \end{array}$$

$$A \qquad\qquad B \qquad\qquad\qquad C$$

$$(m \times n) \cdot (n \times k) = (m \times k)$$

product is defined

Property	Example
The commutative property of multiplication does not hold!	$AB \neq BA$
Associative property of multiplication	$(AB)C = A(BC)$
Distributive properties	$A(B + C) = AB + AC$
	$(B + C)A = BA + CA$
Multiplicative identity property	$IA = A$ and $AI = A$
Multiplicative property of zero	$OA = O$ and $AO = O$
Dimension property	The product of an $m \times n$ matrix and an $n \times k$ matrix is an $m \times k$ matrix.

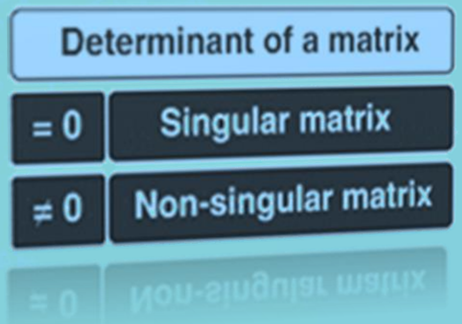
Matrix Multiplication

$$\begin{array}{cc} & \begin{array}{ccc} \vec{c}_1 & \vec{c}_2 & \vec{c}_3 \\ \downarrow & \downarrow & \downarrow \end{array} \\ A \bullet B = \begin{array}{c} \vec{r}_1 \rightarrow \\ \vec{r}_2 \rightarrow \\ \vec{r}_3 \rightarrow \end{array} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \bullet & \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 6 \\ 7 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 26 & 16 & \vec{r}_1 \bullet \vec{c}_3 \\ \vec{r}_2 \bullet \vec{c}_1 & \vec{r}_2 \bullet \vec{c}_2 & \vec{r}_2 \bullet \vec{c}_3 \\ \vec{r}_3 \bullet \vec{c}_1 & \vec{r}_3 \bullet \vec{c}_2 & \vec{r}_3 \bullet \vec{c}_3 \end{bmatrix} \end{array}$$

$$= \begin{bmatrix} 26 & 16 & 1 \bullet 1 + 2 \bullet 6 + 3 \bullet 5 \\ 4 \bullet 1 + 5 \bullet 2 + 6 \bullet 7 & 4 \bullet 2 + 5 \bullet 4 + 6 \bullet 2 & 4 \bullet 1 + 5 \bullet 6 + 6 \bullet 5 \\ 7 \bullet 1 + 8 \bullet 2 + 9 \bullet 7 & 7 \bullet 2 + 8 \bullet 4 + 9 \bullet 2 & 7 \bullet 1 + 8 \bullet 6 + 9 \bullet 5 \end{bmatrix}$$

$$= \begin{bmatrix} 26 & 16 & 28 \\ 56 & 40 & 64 \\ 86 & 64 & 100 \end{bmatrix}$$

Singularity



A singular matrix refers to a matrix whose determinant is zero. Furthermore, such a matrix has no inverse

Reason : **At least one** of the columns or rows are linear combinations of other columns or rows

If $\begin{bmatrix} 2+x & 3 & 4 \\ 1 & -1 & 2 \\ x & 1 & -5 \end{bmatrix}$ is a singular matrix

then x is (A) $\frac{13}{25}$ (B) $-\frac{25}{13}$ (C) $\frac{5}{13}$ (D) $\frac{25}{13}$

Inverse of matrix

$$A^{-1} = \frac{1}{\det(A)} \operatorname{Adj}(A)$$

Adjugate

$$\operatorname{cof}\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} \begin{vmatrix} e & f \\ h & i \end{vmatrix} & -\begin{vmatrix} d & f \\ g & i \end{vmatrix} & \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ -\begin{vmatrix} b & c \\ h & i \end{vmatrix} & \begin{vmatrix} a & c \\ g & i \end{vmatrix} & -\begin{vmatrix} a & b \\ g & h \end{vmatrix} \\ \begin{vmatrix} b & c \\ e & f \end{vmatrix} & -\begin{vmatrix} a & c \\ d & f \end{vmatrix} & \begin{vmatrix} a & b \\ d & e \end{vmatrix} \end{bmatrix}$$

$$\mathbf{C} = \left((-1)^{i+j}\mathbf{M}_{ij}\right)_{1\le i,j\le n}$$

$$\operatorname{adj}\begin{bmatrix} -3 & 2 & -5 \\ -1 & 0 & -2 \\ 3 & -4 & 1 \end{bmatrix} = \begin{bmatrix} -8 & 18 & -4 \\ -5 & 12 & -1 \\ 4 & -6 & 2 \end{bmatrix}$$

$$\mathbf{C} = \mathbf{Cof(A)} = \mathbf{adj(A)}^{\mathbf{T}}$$

Adjugate And inverse

$$\mathrm{adj}(\mathbf{I}) = \mathbf{I}$$

$$\mathrm{adj}(\mathbf{0}) = \mathbf{0}$$

$$\mathrm{adj}(\mathbf{A}) = (\det \mathbf{A})\mathbf{A}^{-1}$$

$$\mathrm{adj}(c\mathbf{A}) = c^{n-1}\,\mathrm{adj}(\mathbf{A})$$

$$\mathrm{adj}(\mathbf{A}^{\mathsf{T}}) = \mathrm{adj}(\mathbf{A})^{\mathsf{T}}$$

$$\mathrm{adj}(\mathbf{AB}) = \mathrm{adj}(\mathbf{B})\,\mathrm{adj}(\mathbf{A}).$$

$$\mathrm{adj}(\mathbf{A}^k) = \mathrm{adj}(\mathbf{A})^k.$$

$$\text{Let } A = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & -1 & 2 \end{vmatrix}$$

$$= 2\begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} - 1\begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} + 1\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}$$

$$a_{ij} = \begin{bmatrix} +\begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} & -\begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} & +\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} \\ -\begin{vmatrix} 1 & 1 \\ -1 & 2 \end{vmatrix} & +\begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} \\ +\begin{vmatrix} 1 & 1 \\ 1 & 1 \end{vmatrix} & -\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} & +\begin{vmatrix} 2 & 1 \\ 1 & 1 \end{vmatrix} \end{bmatrix}$$

$$a_{ij} = \begin{bmatrix} +(2+1) & -(2-1) & +(-1-1) \\ -(2+1) & +(4-1) & -(-2-1) \\ +(1-1) & -(2-1) & +(2-1) \end{bmatrix}$$

$$a_{ij} = \begin{bmatrix} 3 & -1 & -2 \\ -3 & 3 & 3 \\ 0 & -1 & 1 \end{bmatrix}$$

$$\boxed{\mathrm{adj}\,A = \begin{bmatrix} 3 & -3 & 0 \\ -1 & 3 & -1 \\ -2 & 3 & 1 \end{bmatrix}}$$

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

Linear combination

- If one vector is equal to the sum of scalar multiplications of other vectors, it is said to be a **linear combination** of the other vectors.

$$\alpha_1 A_1 + \alpha_2 A_2 + \alpha_3 A_3 = \alpha_1 \cdot \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \alpha_2 \cdot \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + \alpha_3 \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_1 \cdot 0 + \alpha_2 \cdot 1 + \alpha_3 \cdot 1 \\ \alpha_1 \cdot 1 + \alpha_2 \cdot 0 + \alpha_3 \cdot 0 \\ \alpha_1 \cdot 2 + \alpha_2 \cdot (-1) + \alpha_3 \cdot 0 \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_2 + \alpha_3 \\ \alpha_1 \\ 2\alpha_1 - \alpha_2 \end{bmatrix}$$

assume

$w = (18, 5\ ), \qquad u = (-9{,}2), \qquad v = (0\ , 9)$

then we have:

$w = c_1 . u + c_2 . v$

Solving for constants we will have

$w = -2u + v$

Norm

Definition 5.10. *A* **matrix norm** *must satisfy*

(a) $\|A\| \geq 0 \quad \forall A$

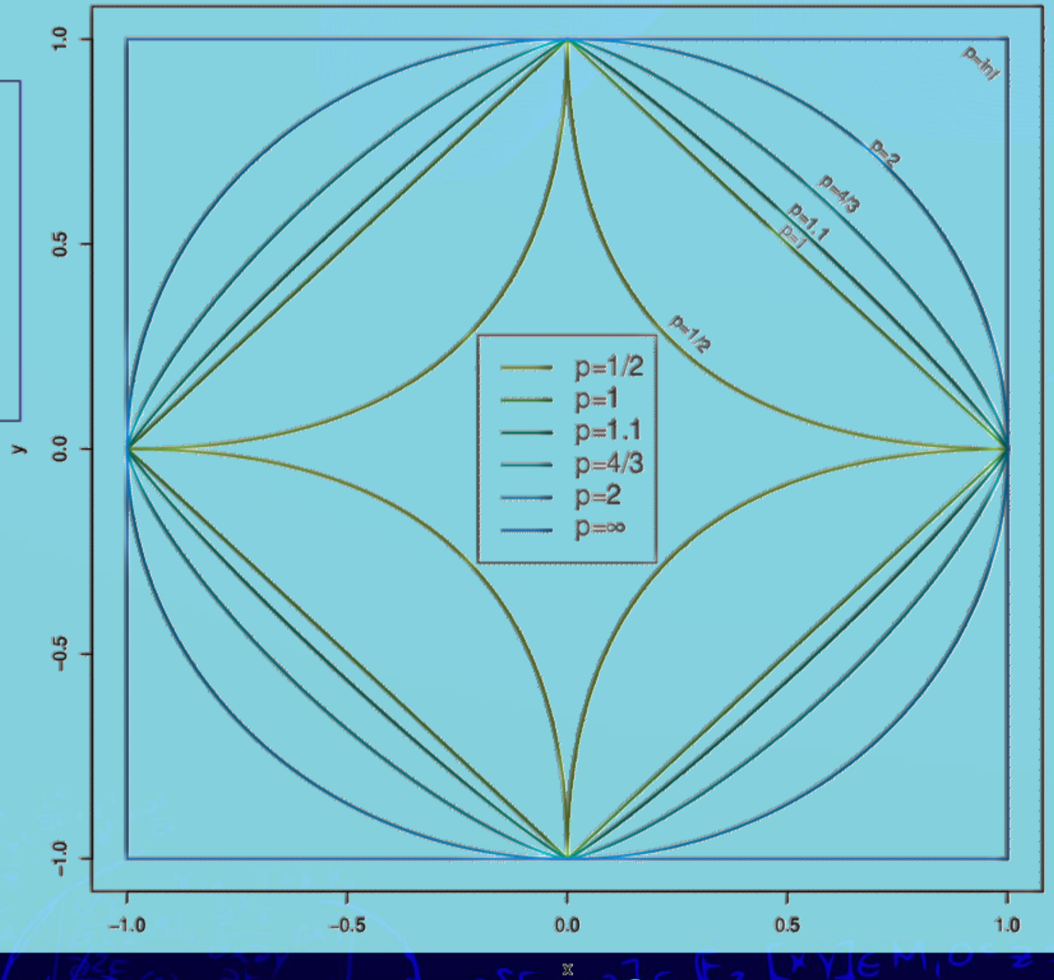
(b) $\|A\| = 0 \quad \Leftrightarrow \quad A = 0$

(c) $\|\alpha A\| = |\alpha| \|A\|$

(d) $\|A + B\| \leq \|A\| + \|B\|$

(e) $\|A \cdot B\| \leq \|A\| \cdot \|B\| \quad$ *(submultiplicativity)*

- The 1-norm (aka taxicab norm) $\|\mathbf{x}\|_1 = \sum_i |x_i|$
- The 2-norm (aka Euclidean norm) $\|\mathbf{x}\|_2 = \sqrt{\sum_i |x_i|^2}$
- The more general p-norm $\|\mathbf{x}\|_p = \left(\sum_i |x_i|^p \right)^{1/p}$
- The infinity norm $\|\mathbf{x}\|_\infty = \max\{|x_i|\}$



دانشگاه صنعتی خواجه نصیرالدین طوسی

جبر خطی
(مباحث اولیه ۲)

استاد : دکتر محمدرضا ملک
دستیار : امیرحسام طاهرزادگانی
دانشکده مهندسی ژئوماتیک و ژئودزی

Trace

The **trace** of an $n \times n$ square matrix $\mathbf{A}$ is defined as

$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}$$

$$\operatorname{tr}(\mathbf{A}) = \sum_{i} \lambda_i$$

Later we will talk about this

$$\begin{aligned} \operatorname{tr}(\mathbf{A}+\mathbf{B}) &= \operatorname{tr}(\mathbf{A}) + \operatorname{tr}(\mathbf{B}), \\ \operatorname{tr}(c\mathbf{A}) &= c \operatorname{tr}(\mathbf{A}), \\ \operatorname{tr}(\mathbf{AB}) &= \operatorname{tr}(\mathbf{BA}), \end{aligned}$$

$$\operatorname{tr}(\mathbf{AB}) = \operatorname{tr}(\mathbf{BA})$$

$$\operatorname{tr}(\mathbf{A}) = \operatorname{tr}\left(\mathbf{A}^{\mathsf{T}}\right).$$

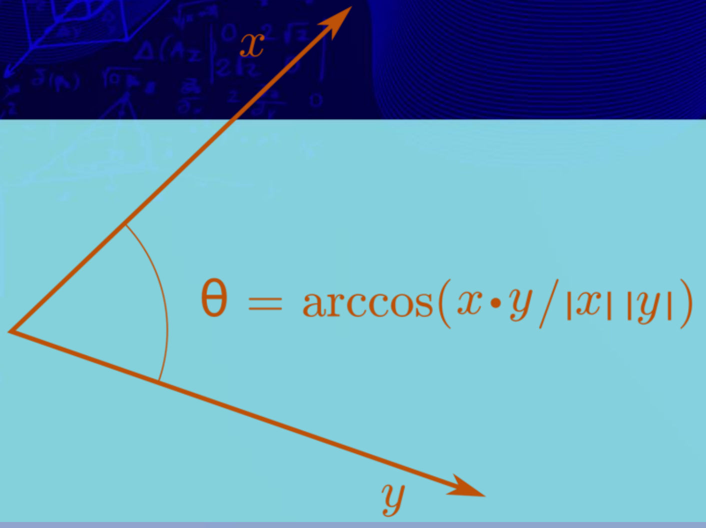
$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 3 \\ 11 & 5 & 2 \\ 6 & 12 & -5 \end{pmatrix}$$

$$\operatorname{tr}(\mathbf{A}) = \sum_{i=1}^{3} a_{ii} = a_{11} + a_{22} + a_{33} = 1 + 5 + (-5) = 1$$

Inner product

- An inner product on a real spaces V is a function that associates a number, denoted $\langle u, v \rangle$, with each pair of vectors u and v of V. This function has to satisfy the following conditions for vectors u, v, and w, and scalar c.
- 1. $\langle u, v \rangle = \langle v, u \rangle$ (symmetry axiom)
- 2. $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ (additive axiom)
- 3. $\langle cu, v \rangle = c \langle u, v \rangle$ (homogeneity axiom)
- 4. $\langle u, u \rangle \neq 0$, and $\langle u, u \rangle = 0$ if and only if $u = 0$(position definite axiom)

Dot product



Properties Of Scalar Product

1. $\bar{a}.\bar{b}$ is a scalar quantity.
2. $\bar{a}.\bar{b} = \bar{0} \Leftrightarrow \bar{a} \perp \bar{b}$
3. $\bar{a}.\bar{b} = |\bar{a}|\,|\bar{b}| \Leftrightarrow \theta = 0^\circ$
4. $\bar{a}.\bar{b} = -\,|\bar{a}|\,|\bar{b}| \Leftrightarrow \theta = 180^\circ$
5. $\bar{i}.\bar{i} = \bar{j}.\bar{j} = \bar{k}.\bar{k} = 1$
6. $\bar{i}.\bar{j} = \bar{j}.\bar{k} = \bar{k}.\bar{i} = 0$
7. $\cos\theta = \dfrac{\bar{a}.\bar{b}}{|\bar{a}|\,|\bar{b}|}$
8. $\bar{a}.\bar{b} = \bar{b}.\bar{a}$
9. $\bar{a}.(\bar{b} + \bar{c}) = \bar{a}.\bar{b} + \bar{a}.\bar{c}$
10. if $\bar{a} = a_1\bar{i} + a_2\bar{j} + a_3\bar{k}$
    and
    $\bar{b} = b_1\bar{i} + b_2\bar{j} + \bar{b}_3k$
    then
    $\bar{a}.\bar{b} = a_1b_1 + a_2b_2 + a_3b_3$
11. $(\lambda\bar{a}).\bar{b} = \lambda(\bar{a}.\bar{b})$
    $= \bar{a}.(\lambda.\bar{b})$

Lengh of vector **u**,**v**

Symbol for inner product

Angle between **u** and **v**

$$\mathbf{u} \bullet \mathbf{v} = |\mathbf{u}|\,|\mathbf{v}|\cos(\theta) \quad (1)$$

$$= x_1 \times x_2 + y_1 \times y_2 \quad (2)$$

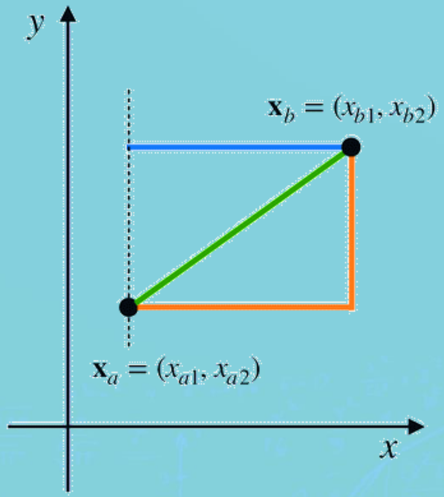
$$= \mathbf{u}\mathbf{v}^T \quad (3)$$

Transpose of vector **v**

$$\begin{bmatrix} A_x & A_y & A_z \end{bmatrix}\begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} = A_xB_x + A_yB_y + A_zB_z = \vec{A}\cdot\vec{B}$$

Norm

- In mathematics, a norm is a function from a real or complex vector space to the nonnegative real numbers that behaves in certain ways like the distance from the origin



$p = 2$ Euclidean distance

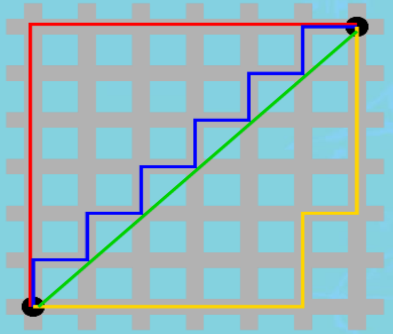
$$\|\mathbf{x}_a - \mathbf{x}_b\|_2 = (|x_{a1} - x_{b1}|^2 + |x_{a2} - x_{b2}|^2)^{\frac{1}{2}}$$

$p = 1$ Manhattan distance

$$\|\mathbf{x}_a - \mathbf{x}_b\|_M = |x_{a1} - x_{b1}| + |x_{a2} - x_{b2}|$$

$p = \infty$ Chebyshev distance

$$\|\mathbf{x}_a - \mathbf{x}_b\|_\infty = \max\{|x_{a1} - x_{b1}|, |x_{a2} - x_{b2}|\}$$



Norm

Definition 5.10. *A* **matrix norm** *must satisfy*

(a) $\|A\| \geq 0 \quad \forall A$

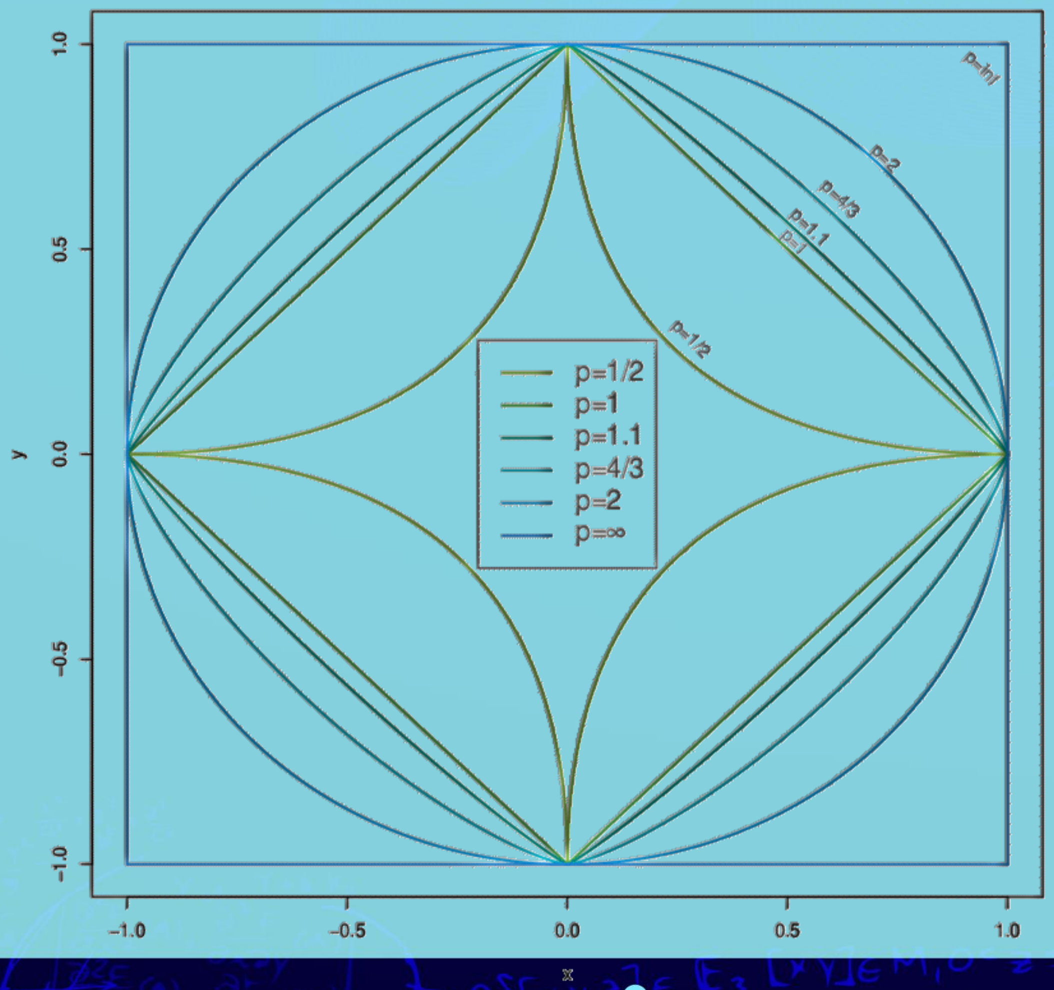
(b) $\|A\| = 0 \quad \Leftrightarrow \quad A = 0$

(c) $\|\alpha A\| = |\alpha| \|A\|$

(d) $\|A + B\| \leq \|A\| + \|B\|$

(e) $\|A \cdot B\| \leq \|A\| \cdot \|B\|$ *(submultiplicativity)*

- The 1-norm (aka taxicab norm) $||\mathbf{x}||_1 = \sum_i |x_i|$
- The 2-norm (aka Euclidean norm) $||\mathbf{x}||_2 = \sqrt{\sum_i |x_i|^2}$
- The more general p-norm $||\mathbf{x}||_p = \left(\sum_i |x_i|^p\right)^{1/p}$
- The infinity norm $||\mathbf{x}||_\infty = \max\{|x_i|\}$



Entrywise matrix norms

$$\|A\|_{p,p} = \|\mathrm{vec}(A)\|_p = \left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^p\right)^{1/p}$$

General form

$$\|A\|_{\mathrm{F}} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2} = \sqrt{\mathrm{trace}(A^*A)}$$

Frobenius norm or the Hilbert–Schmidt norm

$$\|A\|_{\max} = \max_{ij}|a_{ij}|.$$

Max norm

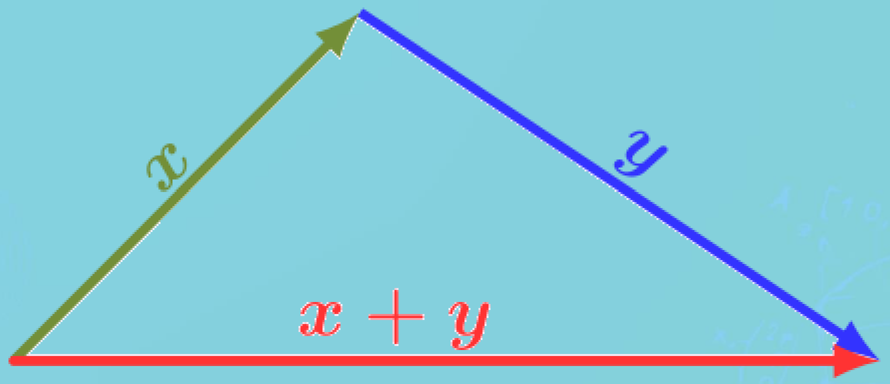
Special relations between norms

$$\|\mathbf{u}\|_\infty \leq \|\mathbf{u}\|_2 \leq \|\mathbf{u}\|_1$$

$$\|\mathbf{u}\|_\infty \leq \|\mathbf{u}\|_2 \leq \sqrt{n}\|\mathbf{u}\|_\infty$$

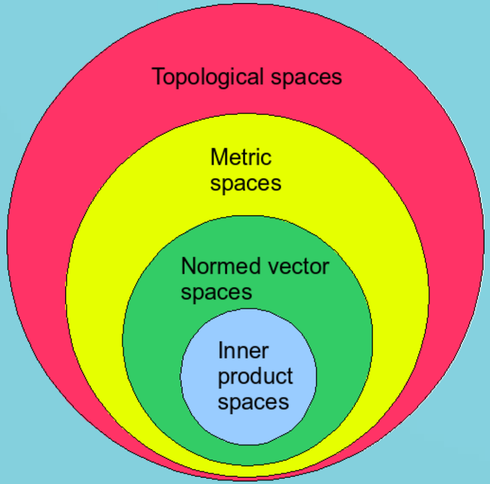
$$\|\mathbf{u}\|_2 \leq \|\mathbf{u}\|_1 \leq \sqrt{n}\|\mathbf{u}\|_2$$

In mathematics, the triangle inequality states that for any triangle, the sum of the lengths of any two sides must be greater than or equal to the length of the remaining side.



Metric Space

A metric space is an ordered pair (M,D) where M is a set and d is a metric on M, i.e. function



$$d(x,y)=0 \iff x=y$$

$$d(x,y)\geq 0$$

$$d(x,y)=d(y,x)$$

$$d(x,z)\leq d(x,y)+d(y,z)$$

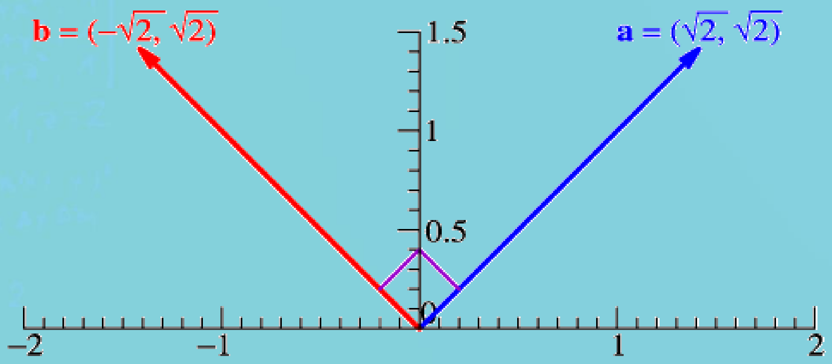
Examples :

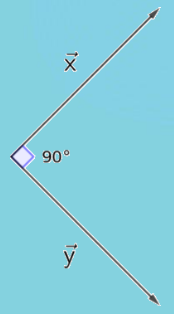
$$d(x,y)=|y-x|$$

$$d(x,y)=|\log(y/x)|$$

	Chicago	Detroit	Indianapolis	Milwaukee	St. Louis
Chicago		281	184	92	297
Detroit	283		287	382	530
Indianapolis	183	287		280	242
Milwaukee	92	382	279		373
St. Louis	298	531	244	375	

orthogonality





$$\vec{x} \cdot \vec{y} = |\vec{x}||\vec{y}|\cos(90^\circ) = 0$$

$$v^T w = 0 \; and \; \|u\|^2 + \|u\|^2 = \|u+w\|^2 = \|u-w\|^2$$

Why ?

CLICK

K. N. Toosi University of Technology
Geodesy & Geomatics Faculty

Hilbert space

- In mathematics, a Hilbert space (named for David Hilbert) generalizes the notion of Euclidean space. It extends the methods of vector algebra and calculus from the two-dimensional Euclidean plane and three-dimensional space to spaces with any finite or infinite number of dimensions. A Hilbert space is a vector space equipped with an inner product operation, which allows lengths and angles to be defined

$$\langle f,g \rangle = \int_a^b f(x)g(x)dx$$

[For Further information]

Symmetric matrix

Symmetric matrix

$$\begin{bmatrix} 1 & 2 & 6 \\ 2 & 8 & 4 \\ 6 & 4 & 5 \end{bmatrix}$$

A symmetric matrix is a square matrix that is equal to its transpose

Because equal matrices have equal dimensions, only square matrices can be symmetric

$$A \text{ is symmetric} \iff A = A^{\mathsf{T}}$$

$$A \text{ is symmetric} \iff \text{for every } i, j, \quad a_{ji} = a_{ij}$$

- The sum and difference of two symmetric matrices is again symmetric
- For integer n, A^n is symmetric if A is symmetric
- if A^{-1} exists and is symmetric then A is symmetric
- $A^T A$ is always symmetric

Skew symmetric matrix

$$A = \begin{pmatrix} 0 & -1 & -2 \\ 1 & 0 & -1 \\ 2 & 1 & 0 \end{pmatrix} \quad A^T = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix}$$

$$-A = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -2 & -1 & 0 \end{pmatrix}$$

are equal

- The sum of two skew-symmetric matrices is skew-symmetric.
- A scalar multiple of a skew-symmetric matrix is skew-symmetric.
- The elements on the diagonal of a skew-symmetric matrix are zero, and therefore its trace equals zero.
- If A is skew symmetric then $I + A$ is always symmetric

$$A \text{ skew-symmetric} \quad \Longleftrightarrow \quad A^{\mathsf{T}} = -A.$$

$$A \text{ skew-symmetric} \quad \Longleftrightarrow \quad a_{ji} = -a_{ij}.$$

Decomposition into symmetric and skew-symmetric

- Any square matrix can uniquely be written as sum of a symmetric and a skew-symmetric matrix. This decomposition is known as the Toeplitz decomposition.

$$\mathrm{Mat}_n = \mathrm{Sym}_n \oplus \mathrm{Skew}_n\,,$$

$\oplus$ Direct sum

$$\mathrm{Sym}_n \cap \mathrm{Skew}_n = \{0\}$$

$$\mathrm{Mat}_n = \mathrm{Sym}_n + \mathrm{Skew}_n$$

$$X = \frac{1}{2}\left(X + X^{\mathsf{T}}\right) + \frac{1}{2}\left(X - X^{\mathsf{T}}\right)$$

$$\tfrac{1}{2}\left(X + X^{\mathsf{T}}\right) \in \mathrm{Sym}_n \quad \tfrac{1}{2}\left(X - X^{\mathsf{T}}\right) \in \mathrm{Skew}_n$$

Example

$$A = \begin{pmatrix} 3 & 3 & -1 \\ -2 & -2 & 1 \\ -4 & -5 & 2 \end{pmatrix} \quad A^T = \begin{pmatrix} 3 & -2 & -4 \\ 3 & -2 & -5 \\ -1 & 1 & 2 \end{pmatrix}$$

$$A + A^T = \begin{pmatrix} (3+3) & (3-2) & (-1-4) \\ (-2+3) & (-2-2) & (1-5) \\ (-4-1) & (-5+1) & (2+2) \end{pmatrix}$$

$$\frac{1}{2}(A + A^T) = \frac{1}{2}\begin{pmatrix} 6 & 1 & -5 \\ 1 & -4 & -4 \\ -5 & -4 & 4 \end{pmatrix} \text{----(1)}$$

$$A - A^T = \begin{pmatrix} (3-3) & (3+2) & (-1+4) \\ (-2-3) & (-2+2) & (1+5) \\ (-4+1) & (-5-1) & (2-2) \end{pmatrix}$$

$$\frac{1}{2}(A - A^T) = \frac{1}{2}\begin{pmatrix} 0 & 5 & 3 \\ -5 & 0 & 6 \\ -3 & -6 & 0 \end{pmatrix} \text{-----(2)}$$

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

$$= \frac{1}{2}\begin{pmatrix} (6+0) & (1+5) & (-5+3) \\ (1-5) & (-4+0) & (-4+6) \\ (-5-3) & (-4-6) & (4+0) \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 6 & 6 & -2 \\ -4 & -4 & 2 \\ -8 & -10 & 4 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 3 & -1 \\ -2 & -2 & 1 \\ -4 & -5 & 2 \end{pmatrix}$$

Special matrices

Diagonal matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

$$L = \begin{bmatrix} \ell_{1,1} & & & & 0 \\ \ell_{2,1} & \ell_{2,2} & & & \\ \ell_{3,1} & \ell_{3,2} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n,1} & \ell_{n,2} & \ldots & \ell_{n,n-1} & \ell_{n,n} \end{bmatrix}$$

$$U = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \ldots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \ldots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$

PROPERTIES OF TRIANGULAR MATRICES

a) The transpose of an upper triangular matrix is a lower triangular matrix, and the transpose of a lower triangular matrix is an upper triangular matrix.

b) The product of lower triangular matrices is lower triangular; the product of upper triangular matrices is upper triangular.

c) A triangular matrix is invertible if and only if its diagonal entries are all nonzero.

d) The inverse of an invertible lower triangular matrix is lower triangular, and the inverse of an invertible upper triangular matrix is upper triangular.

دانشگاه صنعتی خواجه نصیرالدین طوسی

جبر خطی
(فضای برداری)

استاد : دکتر محمدرضا ملک
دستیار : امیرحسام طاهرزادگانی
دانشکده مهندسی ژئوماتیک و ژئودزی

Ring theory in discrete algebra

Field

Integral domain

Commutative ring

Ring

Abelian group

Group

(A1) Closure under addition:	If a and b belong to S, then $a + b$ is also in S
(A2) Associativity of addition:	$a + (b + c) = (a + b) + c$ for all a, b, c in S
(A3) Additive identity:	There is an element 0 in R such that $a + 0 = 0 + a = a$ for all a in S
(A4) Additive inverse:	For each a in S there is an element $-a$ in S such that $a + (-a) = (-a) + a = 0$
(A5) Commutativity of addition:	$a + b = b + a$ for all a, b in S
(M1) Closure under multiplication:	If a and b belong to S, then ab is also in S
(M2) Associativity of multiplication:	$a(bc) = (ab)c$ for all a, b, c in S
(M3) Distributive laws:	$a(b + c) = ab + ac$ for all a, b, c in S $(a + b)c = ac + bc$ for all a, b, c in S
(M4) Commutativity of multiplication:	$ab = ba$ for all a, b in S
(M5) Multiplicative identity:	There is an element 1 in S such that $a1 = 1a = a$ for all a in S
(M6) No zero divisors:	If a, b in S and $ab = 0$, then either $a = 0$ or $b = 0$
(M7) Multiplicative inverse:	If a belongs to S and $a \quad 0$, there is an element a^{-1} in S such that $aa^{-1} = a^{-1}a = 1$

K. N. Toosi University of Technology
Geodesy & Geomatics Faculty

fields

FIELDS

1.12 Definition A *field* is a set F with two operations, called *addition* and *multiplication*, which satisfy the following so-called "field axioms" (A), (M), and (D):

(A) Axioms for addition

(A1) If $x \in F$ and $y \in F$, then their sum $x + y$ is in F.
(A2) Addition is commutative: $x + y = y + x$ for all $x, y \in F$.
(A3) Addition is associative: $(x + y) + z = x + (y + z)$ for all $x, y, z \in F$.
(A4) F contains an element 0 such that $0 + x = x$ for every $x \in F$.
(A5) To every $x \in F$ corresponds an element $-x \in F$ such that

$$x + (-x) = 0.$$

(M) Axioms for multiplication

(M1) If $x \in F$ and $y \in F$, then their product xy is in F.
(M2) Multiplication is commutative: $xy = yx$ for all $x, y \in F$.
(M3) Multiplication is associative: $(xy)z = x(yz)$ for all $x, y, z \in F$.
(M4) F contains an element $1 \neq 0$ such that $1x = x$ for every $x \in F$.
(M5) If $x \in F$ and $x \neq 0$ then there exists an element $1/x \in F$ such that

$$x \cdot (1/x) = 1.$$

[follow the link to see the axioms]

Vector space

DEFINITION:

A vector space is a nonempty set V of objects, called vectors, on which are defined two operations, called addition and multiplication by scalars (real numbers), subject to the following 10 axioms (or rules):

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

These axioms must hold for all vectors $\bar{u}$, $\bar{v}$, and $\bar{w}$ in V and all scalars c and d.

Let V be the set of n by 1 column matrices of real numbers, let the field of scalars be $\mathbb{R}$, and define vector addition and scalar multiplication by

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix} \qquad c\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{pmatrix}$$

Let's verify that the conditions for being a vector space over $\mathbb{R}$ are satisfied in this case.

follow link to see the proof

Subspace

- Strictly speaking, A Subspace is a Vector Space included in another larger Vector Space. Therefore, all properties of a Vector Space, such as being closed under addition and scalar multiplication still hold true when applied to the Subspace.

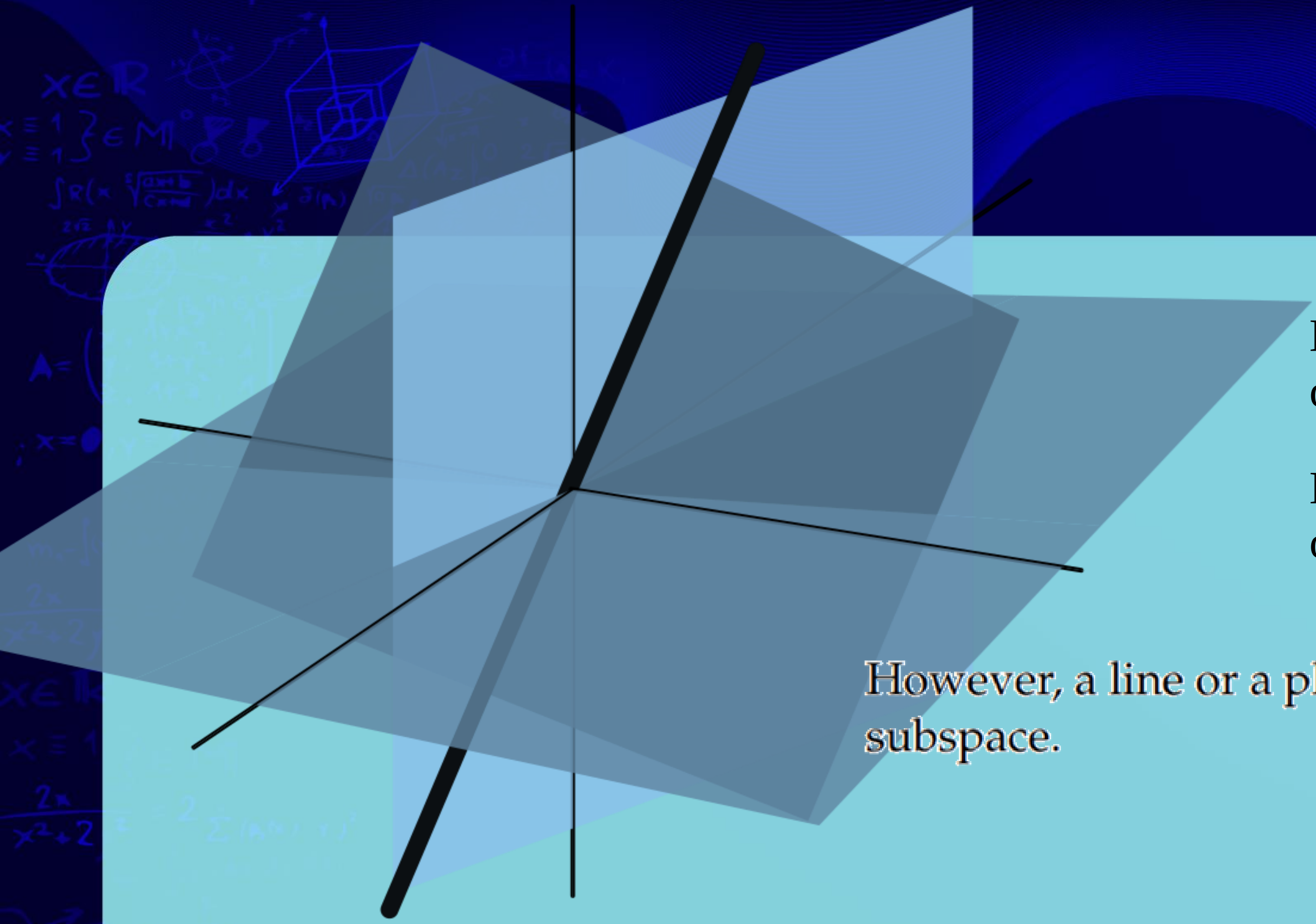
- $s = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \rightarrow a = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}, b = \begin{bmatrix} 8 \\ 9 \\ 0 \end{bmatrix} \quad \begin{cases} a + b \in s \\ ca, cb \in s \end{cases}$

The Subspace Test
To test whether or not S is a subspace of some Vector Space R you must check two things:
1. if s1 and s2 are vectors in S, their sum must also be in S
2. if s is a vector in S and k is a scalar, ks must also be in S

Other examples of Sub Spaces:

- The line defined by the equation $y = 2x$, also defined by the vector definition $\begin{pmatrix} t \\ 2t \end{pmatrix}$ is a subspace of $\mathbb{R}^2$
- The plane $z = -2x$, otherwise known as $\begin{pmatrix} t \\ 0 \\ -2t \end{pmatrix}$ is a subspace of $\mathbb{R}^3$

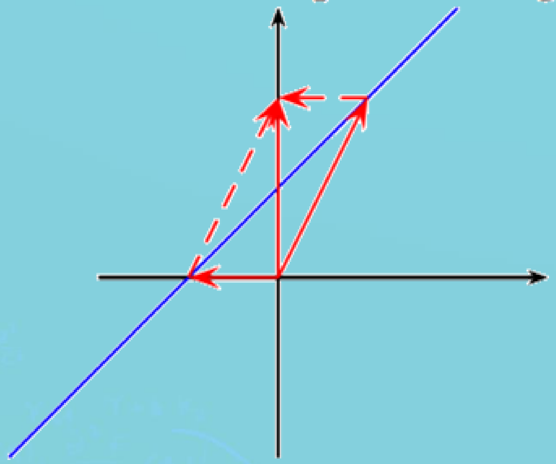
One subtlety



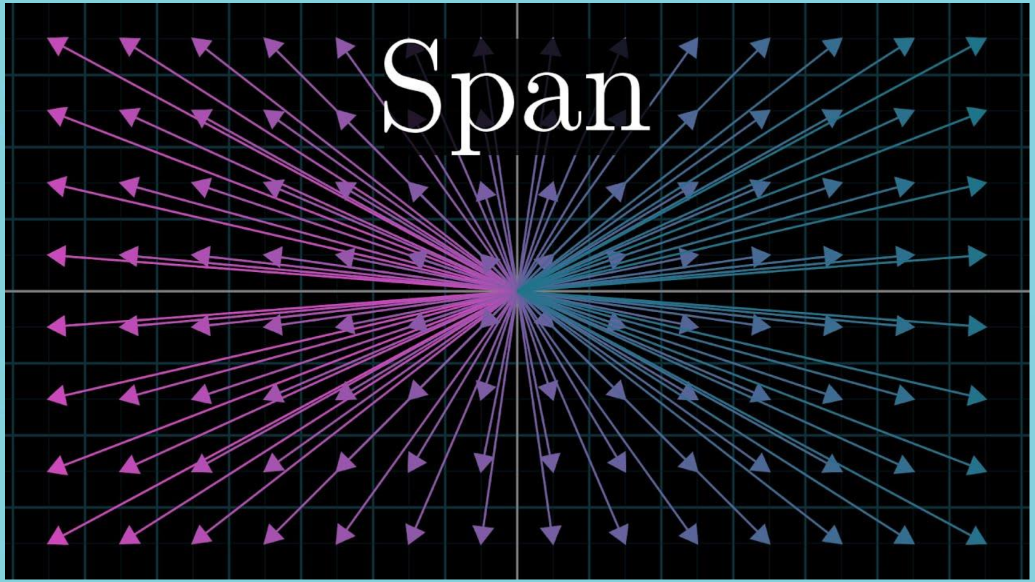
Every line in $\mathbb{R}^n$ that passes through the origin is a subspace of $\mathbb{R}^n$.

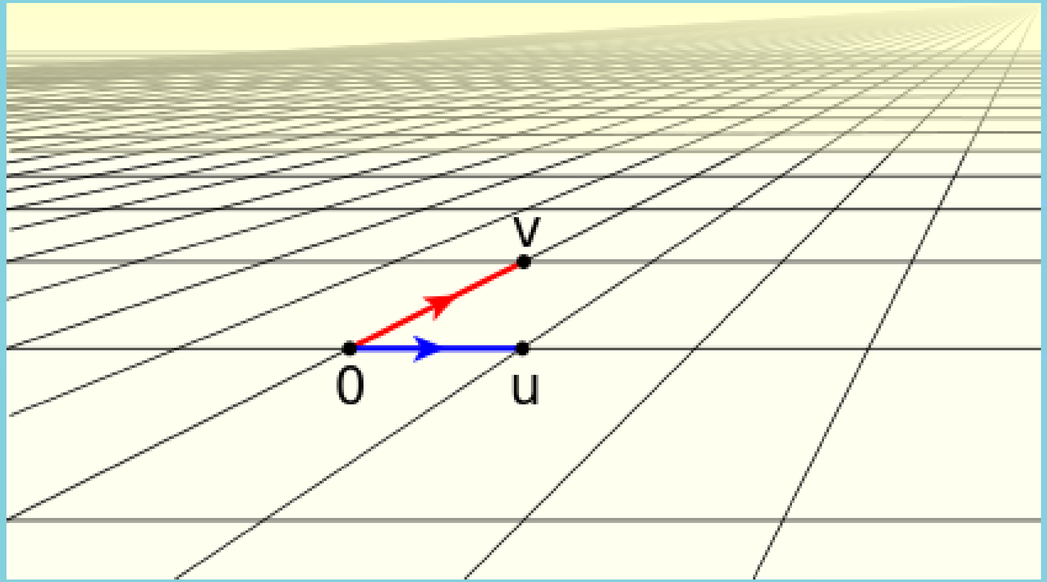
Every plane in $\mathbb{R}^n$ that passes through the origin is a subspace of $\mathbb{R}^n$.

However, a line or a plane that does not pass through the origin is not a subspace.



Span





Linear span

- the linear span (also called the linear hull or just span) of a set S of vectors (from a vector space), denoted span(S)is the smallest linear subspace that contains the set. It can be characterized either as the intersection of all linear subspaces that contain S, or as the set of linear combinations of elements of S

$$\operatorname{span}(S) = \left\{ \sum_{i=1}^{k} \lambda_i v_i \;\middle|\; k \in \mathbf{N}, v_i \in S, \lambda_i \in K \right\}.$$

$$W = \mathrm{sp}(S) \quad , \quad W = \mathrm{sp}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

$$W = \{c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n : c_1, c_2, \ldots c_n \in \Re\}$$

Span

- The span of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ consists of all matrices of the form

$$a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

This is the subspace of diagonal matrices.

- The span of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, and $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ consists of all matrices of the form

$$a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + c\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} a & c \\ c & b \end{pmatrix}.$$

This is the subspace of symmetric matrices.

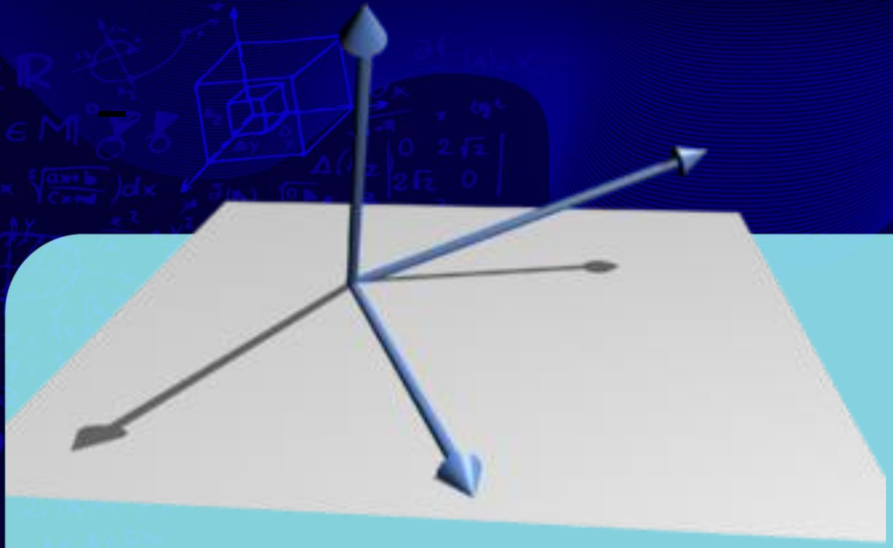
- The span of $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is the subspace of anti-symmetric matrices.

- The span of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, and $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ is the subspace of upper triangular matrices.

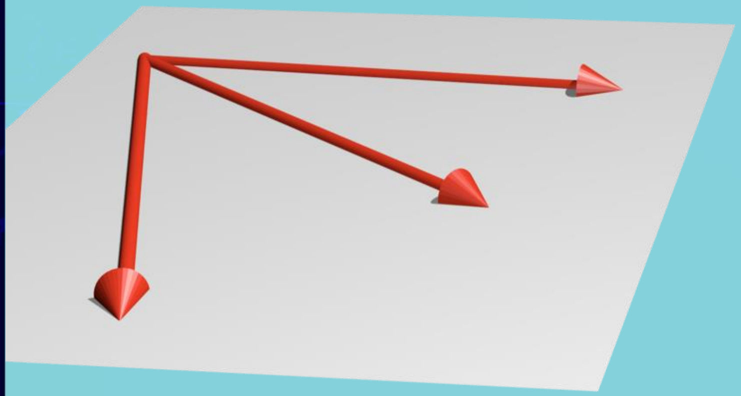
- The span of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ is the entire space $\mathcal{M}_{2,2}(\mathbb{R})$.

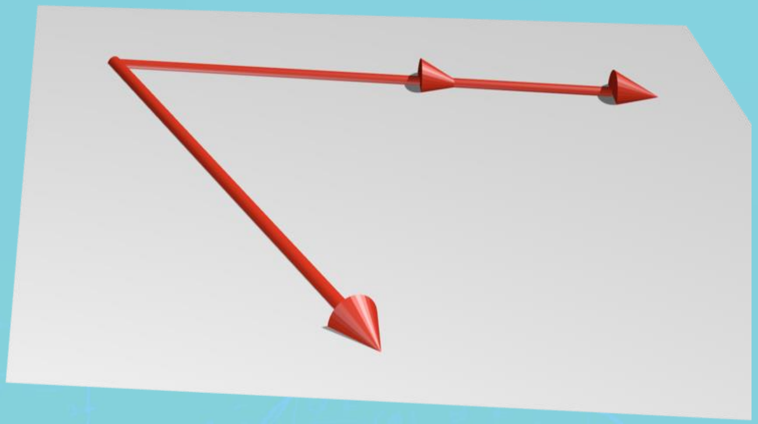
You also want to check the example **3-15** of Dr. Sedghizadeh pamphlet

Linear independence



Linear independence pdf





Change of basis

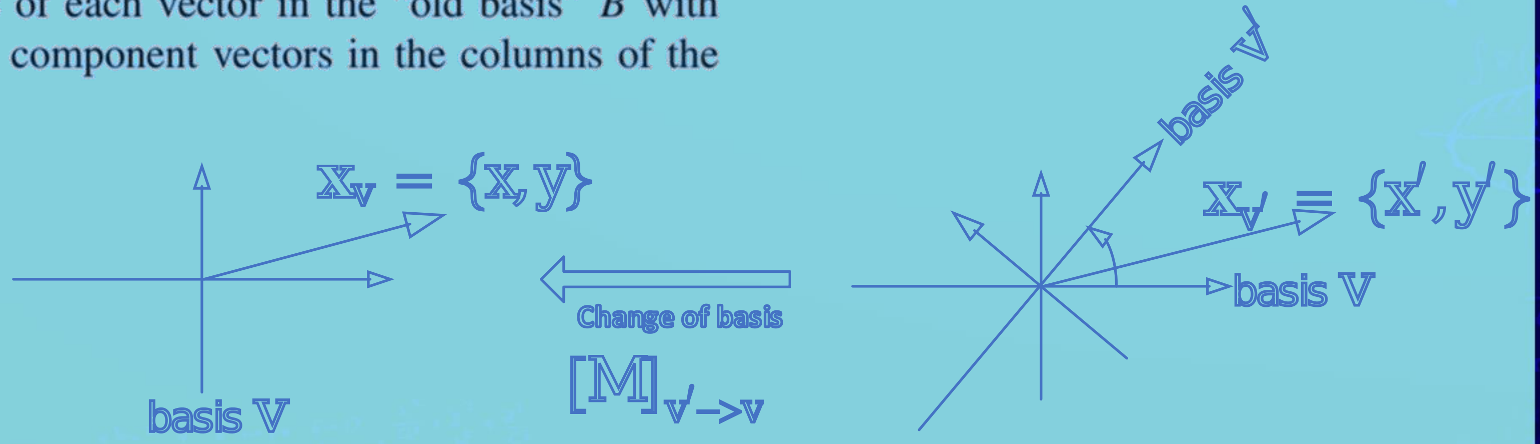
DEFINITION 4.7.6

Let V be an n-dimensional vector space with ordered bases B and C given in (4.7.4). We define the **change-of-basis matrix from** B **to** C by

$$P_{C\leftarrow B} = \left[[\mathbf{v}_1]_C, [\mathbf{v}_2]_C, \ldots, [\mathbf{v}_n]_C \right]. \quad (4.7.5)$$

In words, we determine the components of each vector in the "old basis" B with respect the "new basis" C and write the component vectors in the columns of the change-of-basis matrix.

Link to learn



Same vector X has different representation depending on which basis are used

K. N. Toosi University of Technology
Geodesy & Geomatics Faculty

Find the rank of matrix

Rank = number of non zero rows in row echelon form.

$$\begin{bmatrix} -2 & 3 & 3 \\ 3 & -4 & 1 \\ -5 & 7 & 2 \end{bmatrix} \quad R_1 \to -\frac{1}{2} R_1$$

$$\begin{bmatrix} 1 & -\frac{3}{2} & -\frac{3}{2} \\ 3 & -4 & 1 \\ -5 & 7 & 2 \end{bmatrix} \quad R_2 \to R_2 - 3R_1$$

$$\begin{bmatrix} 1 & -\frac{3}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & \frac{11}{2} \\ -5 & 7 & 2 \end{bmatrix} \quad R_3 \to R_3 + 5R_1$$

$$\begin{bmatrix} 1 & -\frac{3}{2} & -\frac{3}{2} \\ 0 & \frac{1}{2} & \frac{11}{2} \\ 0 & -\frac{1}{2} & -\frac{11}{2} \end{bmatrix} \quad R_2 \to 2R_2$$

$$\begin{bmatrix} 1 & -\frac{3}{2} & -\frac{3}{2} \\ 0 & 1 & 11 \\ 0 & -\frac{1}{2} & -\frac{11}{2} \end{bmatrix} \quad R_3 \to R_3 + \frac{1}{2} R_2$$

$$\begin{bmatrix} 1 & -\frac{3}{2} & -\frac{3}{2} \\ 0 & 1 & 11 \\ 0 & 0 & 0 \end{bmatrix}$$

Rank is 2

Rank

- In linear algebra, the rank of a matrix A is the dimension of the vector space generated (or spanned) by its columns. This corresponds to the maximal number of linearly independent columns of A. This, in turn, is identical to the dimension of the vector space spanned by its rows.

$$\begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix} \xrightarrow{2R_1+R_2 \to R_2} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 3 & 5 & 0 \end{bmatrix} \xrightarrow{-3R_1+R_3 \to R_3} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & -1 & -3 \end{bmatrix}$$

$$\xrightarrow{R_2+R_3 \to R_3} \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{-2R_2+R_1 \to R_1} \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}.$$

$$\operatorname{rank}(A) \leq \min(m, n).$$

Rank properties

- Only a zero matrix has rank zero.
- f is injective (or "one-to-one") if and only if A has rank n (in this case, we say that A has full column rank).
- f is surjective (or "onto") if and only if A has rank m (in this case, we say that A has full row rank).
- If A is a square matrix (i.e., m = n), then A is invertible if and only if A has rank n (that is, A has full rank).
- If B is any n × k matrix, then

$$\operatorname{rank}(AB) \leq \min(\operatorname{rank}(A), \operatorname{rank}(B)).$$

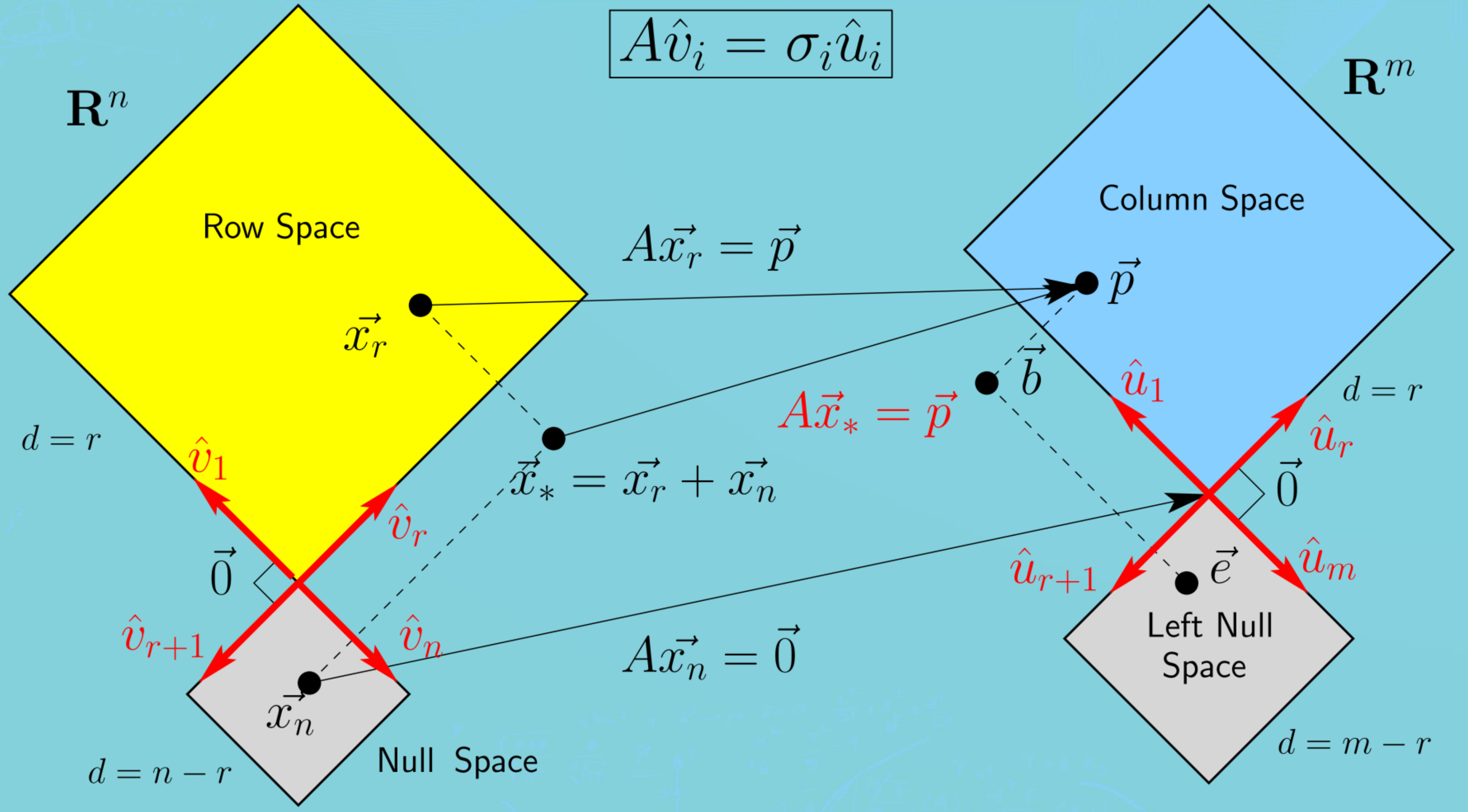
- Sylvester’s rank inequality: if A is an m × n matrix and B is n × k, then

$$\operatorname{rank}(A) + \operatorname{rank}(B) - n \leq \operatorname{rank}(AB).$$

- some other properties :

$$\operatorname{rank}(A + B) \leq \operatorname{rank}(A) + \operatorname{rank}(B)$$

Four fundamental subspace



link to follow

Row space

$$\begin{bmatrix} 1 & 8 & 13 & 12 \\ 14 & 11 & 2 & 7 \\ 4 & 5 & 16 & 9 \\ 15 & 10 & 3 & 6 \end{bmatrix}$$

Let K be a field of scalars. Let A be an m × n matrix, with row vectors r1, r2, …, rm. A linear combination of these vectors is any vector of the form

$$c_1\mathbf{r}_1 + c_2\mathbf{r}_2 + \cdots + c_m\mathbf{r}_m,$$

where c1, c2, …, cm are scalars. The set of all possible linear combinations of r1, …, rm is called the row space of A.

column space

$$\begin{bmatrix} 1 & 8 & 13 & 12 \\ 14 & 11 & 2 & 7 \\ 4 & 5 & 16 & 9 \\ 15 & 10 & 3 & 6 \end{bmatrix}$$

In linear algebra, the column space (also called the range or image) of a matrix A is the span (set of all possible linear combinations) of its column vectors. The column space of a matrix is the image or range of the corresponding matrix transformation.

Range Space or column space

- Let K be a field of scalars. Let A be an m × n matrix, with column vectors $v_1, v_2, \ldots, v_n$. A linear combination of these vectors is any vector of the form

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n,$$

he set of all possible linear combinations of $v_1, v_2, \ldots, v_n$ is called the column space of A

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 0 \end{bmatrix} \longrightarrow c_1\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + c_2\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ 2c_1 \end{bmatrix}$$

Type in z = 2x

Basis for range space

- The columns of A span the column space, but they may not form a basis if the column vectors are not linearly independent. Fortunately, elementary row operations do not affect the dependence relations between the column vectors. This makes it possible to use row reduction to find a basis for the column space.

$$A = \begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 7 & 3 & 9 \\ 1 & 5 & 3 & 1 \\ 1 & 2 & 0 & 8 \end{bmatrix}. \quad \begin{bmatrix} 1 & 3 & 1 & 4 \\ 2 & 7 & 3 & 9 \\ 1 & 5 & 3 & 1 \\ 1 & 2 & 0 & 8 \end{bmatrix} \sim \begin{bmatrix} 1 & 3 & 1 & 4 \\ 0 & 1 & 1 & 1 \\ 0 & 2 & 2 & -3 \\ 0 & -1 & -1 & 4 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 5 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

At this point, it is clear that the first, second, and fourth columns are linearly independent, while the third column is a linear combination of the first two. therefore, the first, second, and fourth columns of the original matrix are a basis for the column space:

$$\begin{bmatrix} 1 \\ 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 7 \\ 5 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \\ 1 \\ 8 \end{bmatrix}.$$

Note that the independent columns of the reduced row echelon form are precisely the columns with pivots

Row space

- The row space is not affected by elementary row operations. This makes it possible to use row reduction to find a basis for the row space.

For example, consider the matrix: The rows of this matrix span the row space, but they may not be linearly independent, in which case the rows will not be a basis. To find a basis, we reduce A to row echelon form:

$$A = \begin{bmatrix} 1 & 3 & 2 \\ 2 & 7 & 4 \\ 1 & 5 & 2 \end{bmatrix}.$$

$$\begin{bmatrix} 1 & 3 & 2 \\ 2 & 7 & 4 \\ 1 & 5 & 2 \end{bmatrix} \xrightarrow{\mathbf{r}_2 - 2\mathbf{r}_1 \to \mathbf{r}_2} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 0 \\ 1 & 5 & 2 \end{bmatrix} \xrightarrow{\mathbf{r}_3 - \ \mathbf{r}_1 \to \mathbf{r}_3} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 0 \\ 0 & 2 & 0 \end{bmatrix}$$

$$\xrightarrow{\mathbf{r}_3 - 2\mathbf{r}_2 \to \mathbf{r}_3} \begin{bmatrix} 1 & 3 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{\mathbf{r}_1 - 3\mathbf{r}_2 \to \mathbf{r}_1} \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Once the matrix is in echelon form, the nonzero rows are a basis for the row space. In this case, the basis is { [1, 3, 2], [2, 7, 4] }. Another possible basis { [1, 0, 2], [0, 1, 0] } comes from a further reduction

Free variables and pivots

A variable is a basic variable if it corresponds to a pivot column. Otherwise, the variable is known as a free variable. In order to determine which variables are basic and which are free, it is necessary to row reduce the augmented matrix to echelon form.

$$\begin{array}{cccc} \textit{pivot} & \textit{free} & \textit{pivot} & \textit{free} \\ \multicolumn{4}{c}{\begin{bmatrix} \boxed{1} & 3 & 1 & 2 \\ 0 & 0 & \boxed{2} & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}} \end{array}$$

Rank and Reduced Row Echelon Matrix

$$A = \begin{bmatrix} 1 & 0 & * & * & 0 & * & 0 \\ 0 & 1 & * & * & 0 & * & 0 \\ 0 & 0 & 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The Pivot columns are not combinations of earlier columns

Pivot columns

Free variables
Free Columns

Null space

- The null space of any matrix A consists of all the vectors B such that AB = 0 and B is not zero. It can also be thought as the solution obtained from AB = 0 where A is known matrix of size m × n and B is matrix to be found of size n × k. The size of the null space of the matrix provides us with the number of linear relations among attributes.

$$A = \begin{pmatrix} X_{11} & \dots & X_{1n} \\ \dots & \dots & \dots \\ X_{m1} & \dots & X_{mn} \end{pmatrix} \qquad B = \begin{pmatrix} b_1 \\ . \\ . \\ . \\ . \\ b_n \end{pmatrix}$$

$$X_{11}b_1 + X_{12}b_2 + \dots\dots\; X_{1n}b_n = 0$$
$$\dots\dots$$
$$\dots\dots$$
$$X_{m1}b_1 + X_{m2}b_2 + \dots\dots\; X_{mn}b_n = 0$$

How to find null space's basis:

1. RREF(Reduced row echelon form)
2. Matmul(RREF(A), x) = 0 and find m equations , m:= no. of pivots
3. Find the nullity := n
4. Represent n linearly indie vectors as null space basis

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

$$Ax = 0$$

$$x\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} + y\begin{bmatrix} 2 \\ 4 \\ 2 \end{bmatrix} + z\begin{bmatrix} 3 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

a basis for the nullspace of a matrix A is given by the set of "special solutions". There is a special solution for each free variable, and each free variable corresponds with a free column

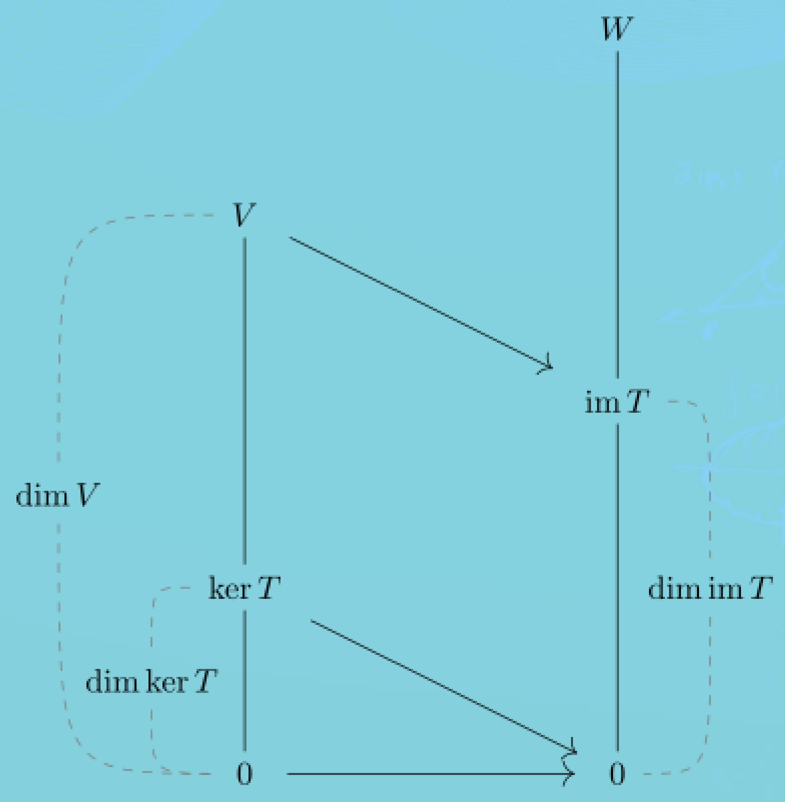
In the case of an matrix, the dimension of the domain is , the number of columns in the matrix. Thus the Rank-Nullity theorem for a given matrix immediately becomes

Rank nullity theorem

$$\mathrm{Rank}(T) + \mathrm{Nullity}(T) = \dim V,$$

$$\mathrm{Nullity}(T) := \dim(\mathrm{Ker}(T)).$$

For matrices theorem above immediately becomes
$\mathbf{dim}(\boldsymbol{Range\ space}) + \mathbf{dim}(\boldsymbol{null\ space}) = \boldsymbol{n}$
Considering the matrix $m \times n$



Left null space

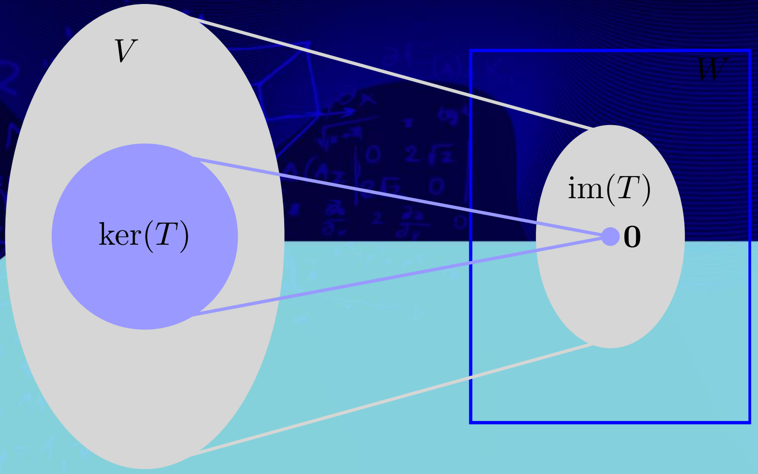
The left nullspace is the space of all vectors y such that $A^T y = 0$. It can equivalently be viewed as the space of all vectors y such that $y^T A = 0$. Thus the term “left” nullspace.

Subspace	Subspace of	Symbol	Dimension	Basis
column space	$\mathbb{R}^m$	$\mathrm{im}(A)$	r = rank	First r columns of U
nullspace (kernel)	$\mathbb{R}^n$	$\ker(A)$	$n-r$	Last $n-r$ columns of V
row space	$\mathbb{R}^n$	$\mathrm{im}(A^T)$	r	First r columns of V
left nullspace (kernel)	$\mathbb{R}^m$	$\ker(A^T)$	$m-r$	Last $m-r$ columns of U

How to find left null space’s basis:

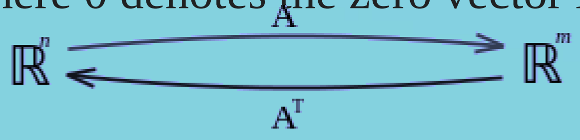
1. RREF(Reduced row echelon form of A.Transpose)
2. Matmul(RREF(A.T), x) = 0 and find m equations , m:= no. of pivots
3. Find the nullity := n
4. Represent n linearly indie vectors as null space basis

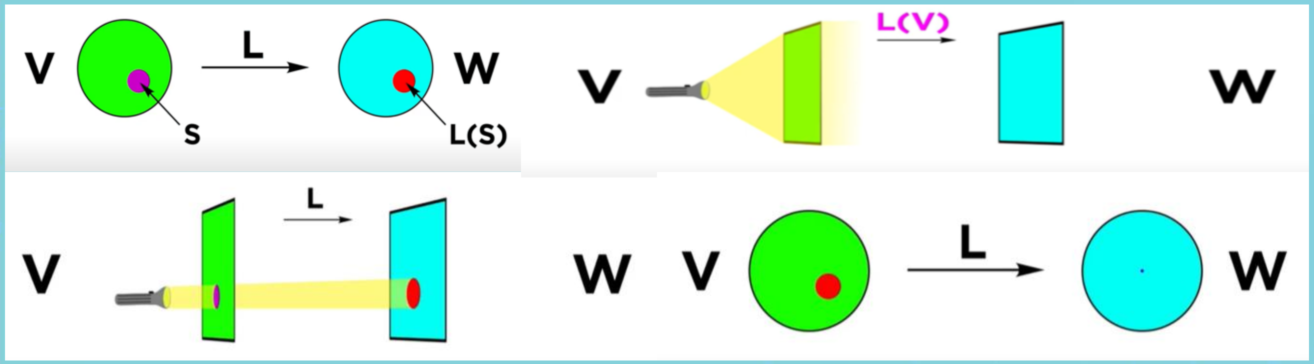
Other names and comp

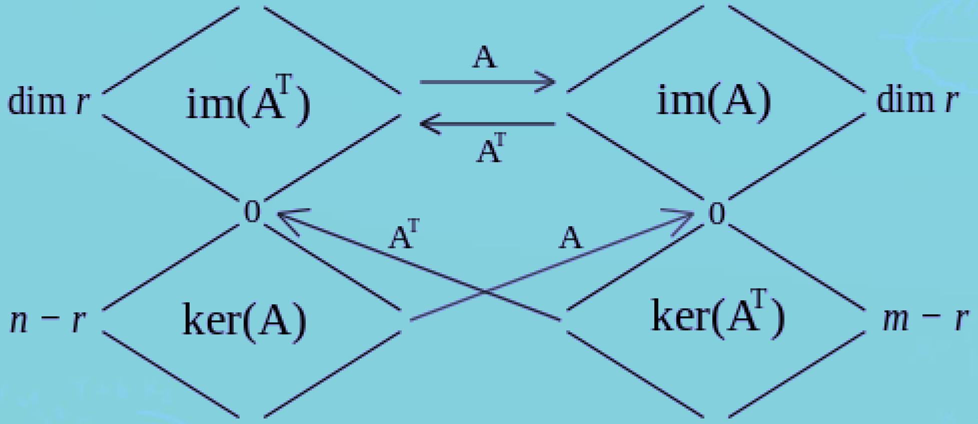


In mathematics, the kernel of a linear map, also known as the null space or nullspace, is the linear subspace of the domain of the map which is mapped to the zero vector. That is, given a linear map L : V → W between two vector spaces V and W, the kernel of L is the vector space of all elements v of V such that L(v) = 0, where 0 denotes the zero vector in W

$$\ker(L) = \{\mathbf{v} \in V \mid L(\mathbf{v}) = \mathbf{0}\}.$$







دانشگاه صنعتی خواجه نصیرالدین طوسی

جبر خطی
(تبدیلات)

استاد : دکتر محمدرضا ملک
دستیار : امیرحسام طاهرزادگانی
دانشکده مهندسی ژئوماتیک و ژئودزی

Linear Mapping

- a linear map (also called a linear mapping, linear transformation, vector space homomorphism, or in some contexts linear function) is a mapping V to W between two vector spaces that preserves the operations of vector addition and scalar multiplication.
- Let V and W be in a same field of F:

$$f(\mathbf{u}+\mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$$

Additivity / operation of addition

$$f(c\mathbf{u}) = cf(\mathbf{u})$$

Homogeneity of degree 1 / operation of scalar multiplication

$$f(c_1\mathbf{u}_1 + \cdots + c_n\mathbf{u}_n) = c_1 f(\mathbf{u}_1) + \cdots + c_n f(\mathbf{u}_n).$$

Examples

$$f : \mathbb{R} \to \mathbb{R} : x \mapsto cx$$

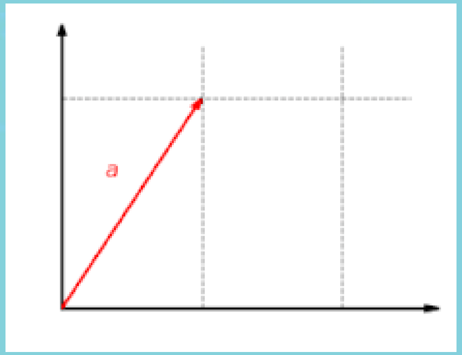
$$\frac{d}{dx}\left(c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x)\right) = c_1 \frac{df_1(x)}{dx} + c_2 \frac{df_2(x)}{dx} + \cdots + c_n \frac{df_n(x)}{dx}$$

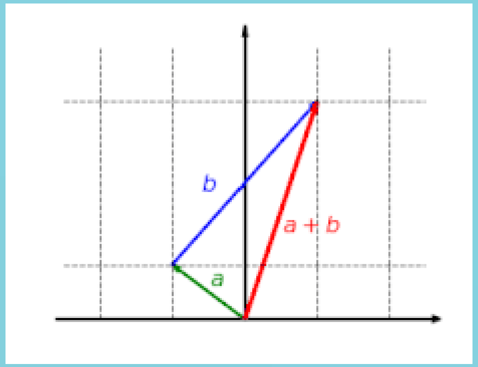
$$\int_a^b \left[c_1 f_1(x) + c_2 f_2(x) + \cdots + c_n f_n(x)\right] dx = c_1 \int_a^b f_1(x)\, dx + c_2 \int_a^b f_2(x)\, dx + \cdots + c_n \int_a^b f_n(x)\, dx.$$

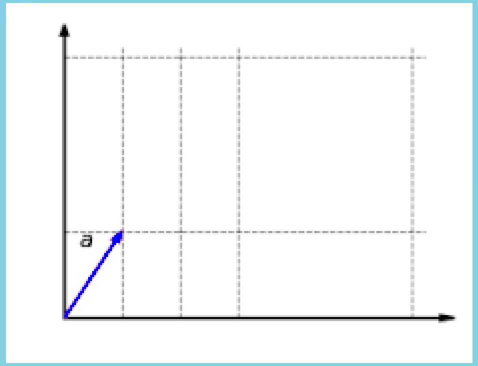
$$E[X+Y] = E[X] + E[Y] \qquad E[aX] = aE[X]$$

σ? VAR(cX)=?

Intuition

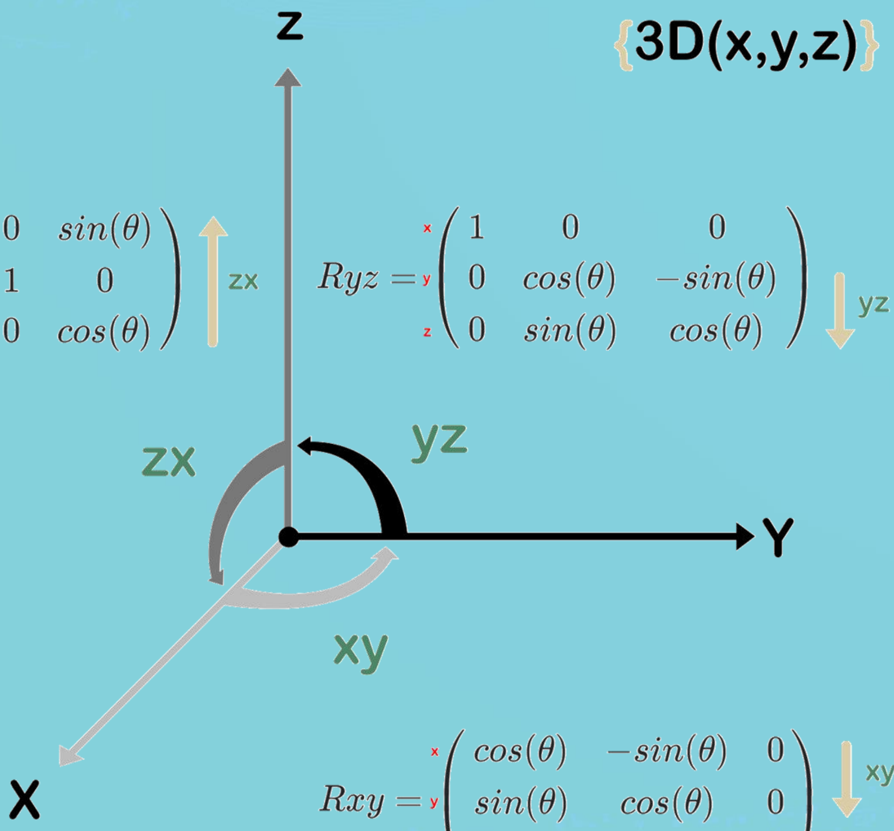






$$f(x,y) = (2x, y) \qquad f(\mathbf{a}+\mathbf{b}) = f(\mathbf{a}) + f(\mathbf{b}) \qquad f(\lambda\mathbf{a}) = \lambda f(\mathbf{a})$$

Rotation matrix

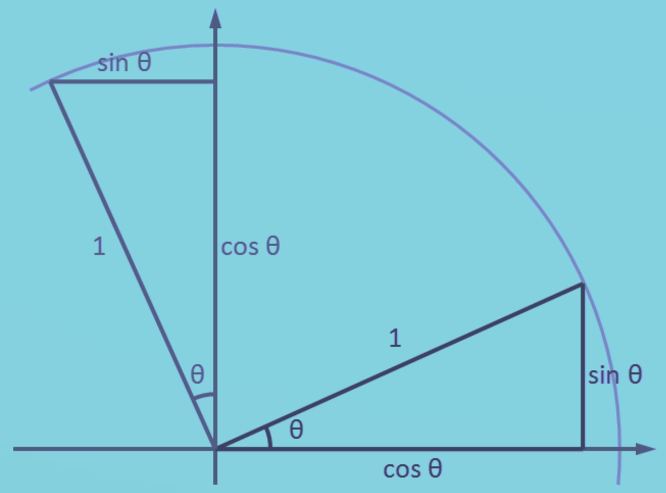


$$Rzx = \begin{array}{c} x \\ y \\ z \end{array}\begin{pmatrix} cos(\theta) & 0 & sin(\theta) \\ 0 & 1 & 0 \\ -sin(\theta) & 0 & cos(\theta) \end{pmatrix} \uparrow zx$$

$$Ryz = \begin{array}{c} x \\ y \\ z \end{array}\begin{pmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & -sin(\theta) \\ 0 & sin(\theta) & cos(\theta) \end{pmatrix} \downarrow yz$$

$$Rxy = \begin{array}{c} x \\ y \\ z \end{array}\begin{pmatrix} cos(\theta) & -sin(\theta) & 0 \\ sin(\theta) & cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \downarrow xy$$

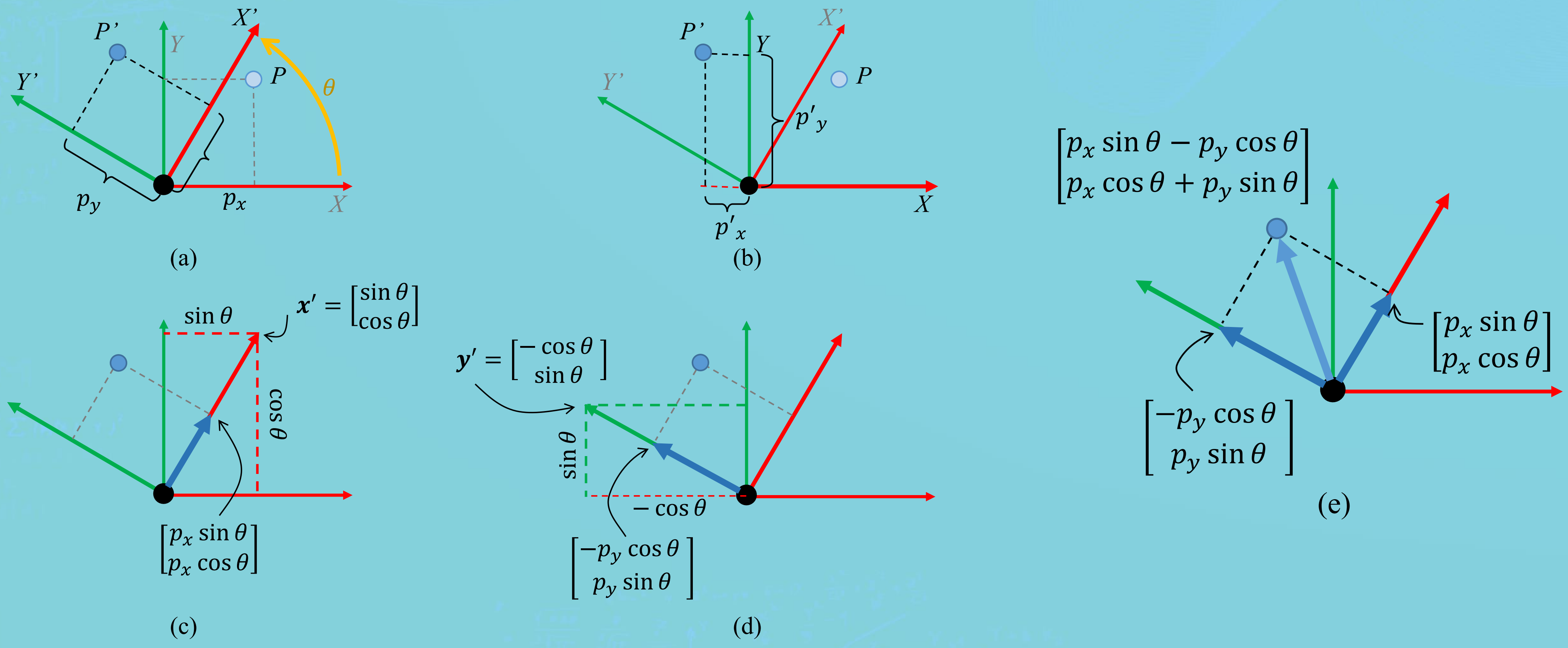
a rotation matrix is a transformation matrix that is used to perform a rotation in Euclidean space.



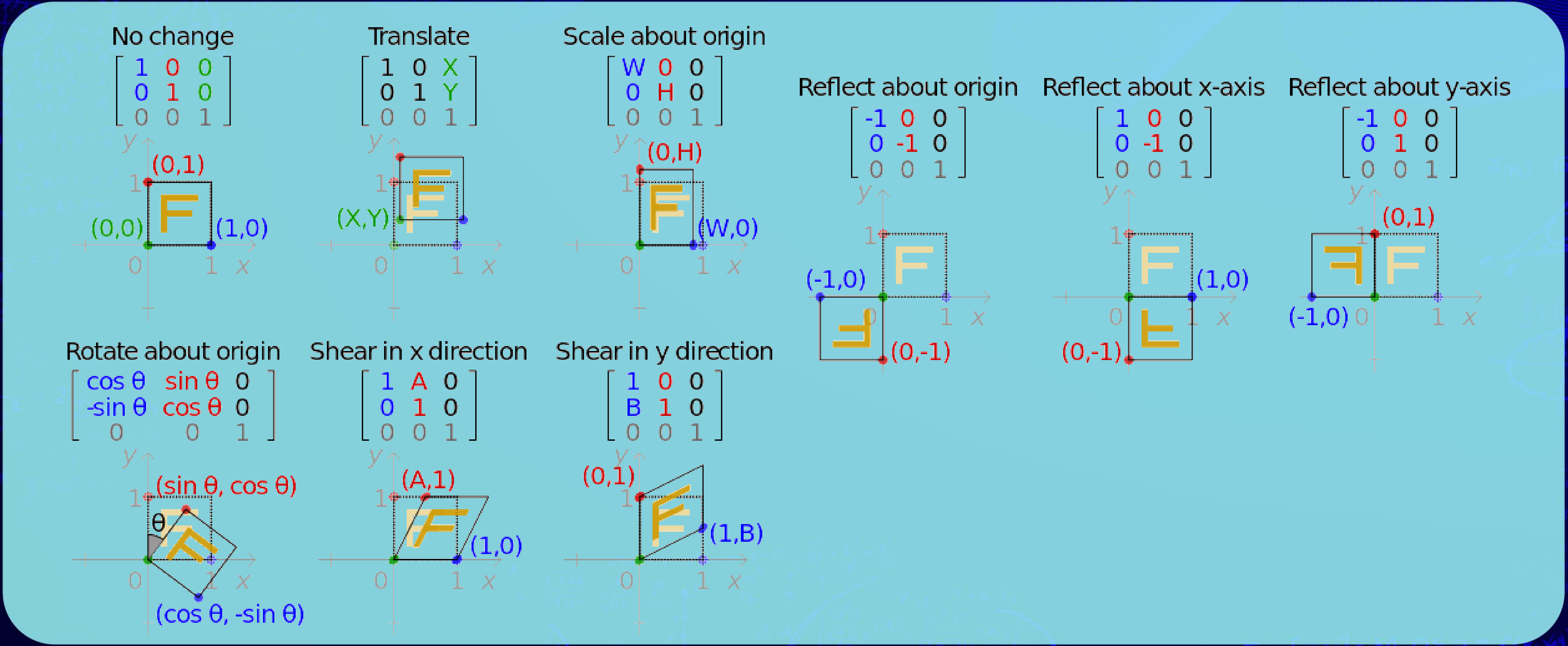
$$R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

K. N. Toosi University of Technology
Geodesy & Geomatics Faculty

2d intuition



2D Transformations recap



جبر خطی
(متعامدسازی)

دانشگاه صنعتی خواجه نصیرالدین طوسی

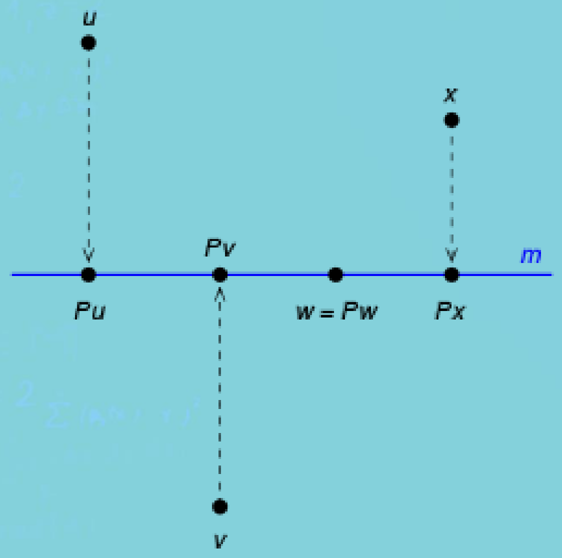
استاد : دکتر محمدرضا ملک

دستیار : امیرحسام طاهرزادگانی

دانشکده مهندسی ژئوماتیک و ژئودزی

Projections

a projection is a linear transformation P from a vector space to itself such that $P^2 = P$



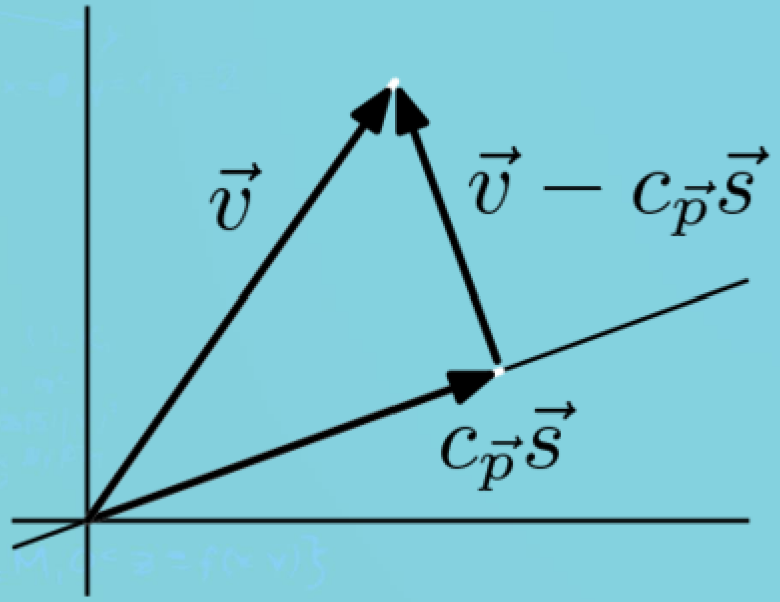
This whole part is a scalar.
This shows “How much you have to extend **y** to make it superimpose onto the vector **u**”

Inner product of v and y, which is a scalar.

$$proj_L(\mathbf{v}) = \boxed{\frac{\boxed{\mathbf{v} \bullet \mathbf{y}}}{\boxed{\mathbf{y} \bullet \mathbf{y}}}}\,\mathbf{y}$$

This still remain as a vector

Inner product of y and y, which is a scalar.



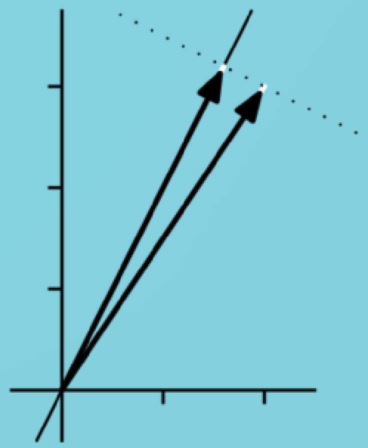
$s.(v - c_p s) = 0 \rightarrow s.v - c_p.s.s = 0$

$c_{\vec{p}} = \vec{v} \cdot \vec{s} / \vec{s} \cdot \vec{s}$

$$\text{proj}_{[\vec{s}\,]}(\vec{v}) = \frac{\vec{v} \cdot \vec{s}}{\vec{s} \cdot \vec{s}} \vec{s}$$

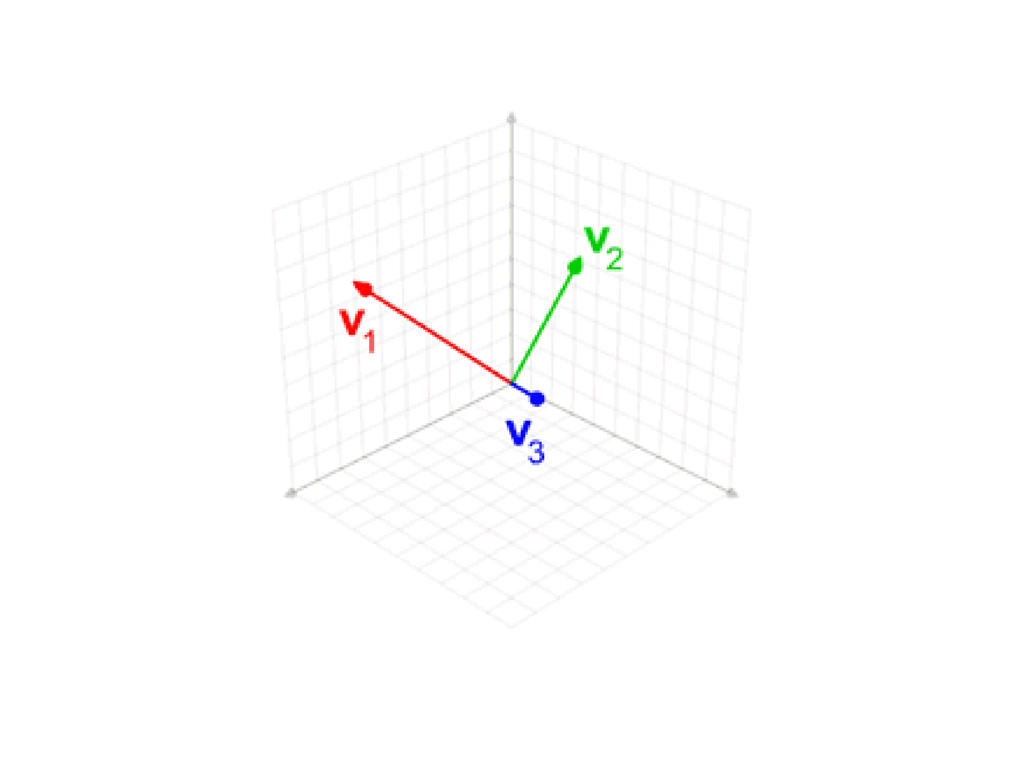
Example:

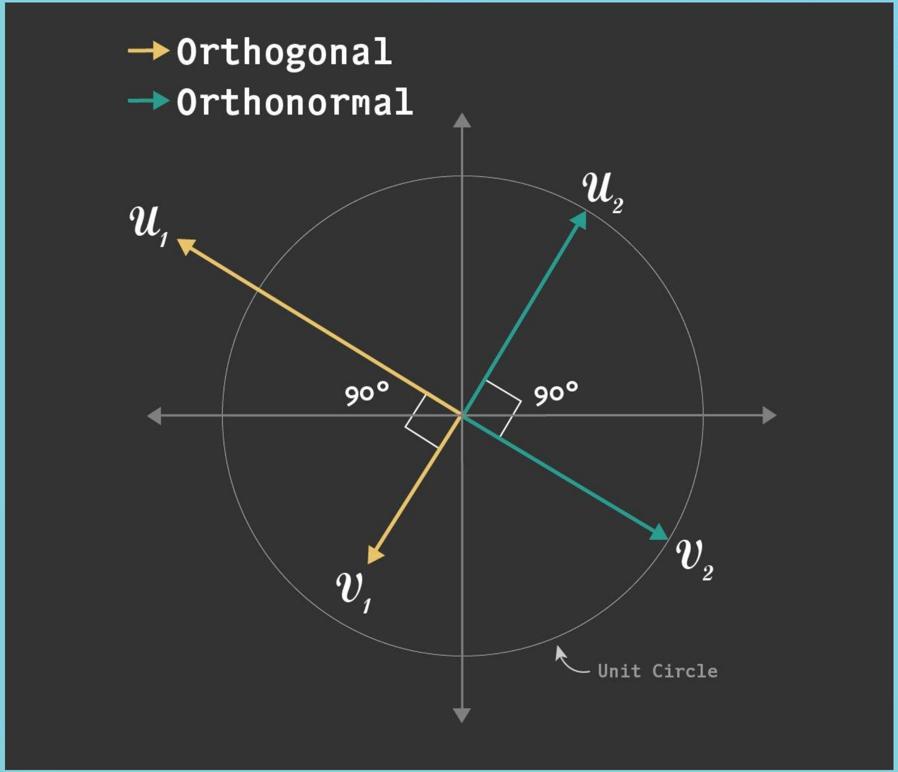
Project $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ onto $y = 2x$



$$\frac{\begin{pmatrix} 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}}{\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{8}{5} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 8/5 \\ 16/5 \end{pmatrix}$$

Gram Schmidt





Gram-Schmidt

- The Gram-Schmidt process (or procedure) is a sequence of operations that allow to transform a set of linearly independent vectors into a set of orthonormal vectors that span the same space spanned by the original set.

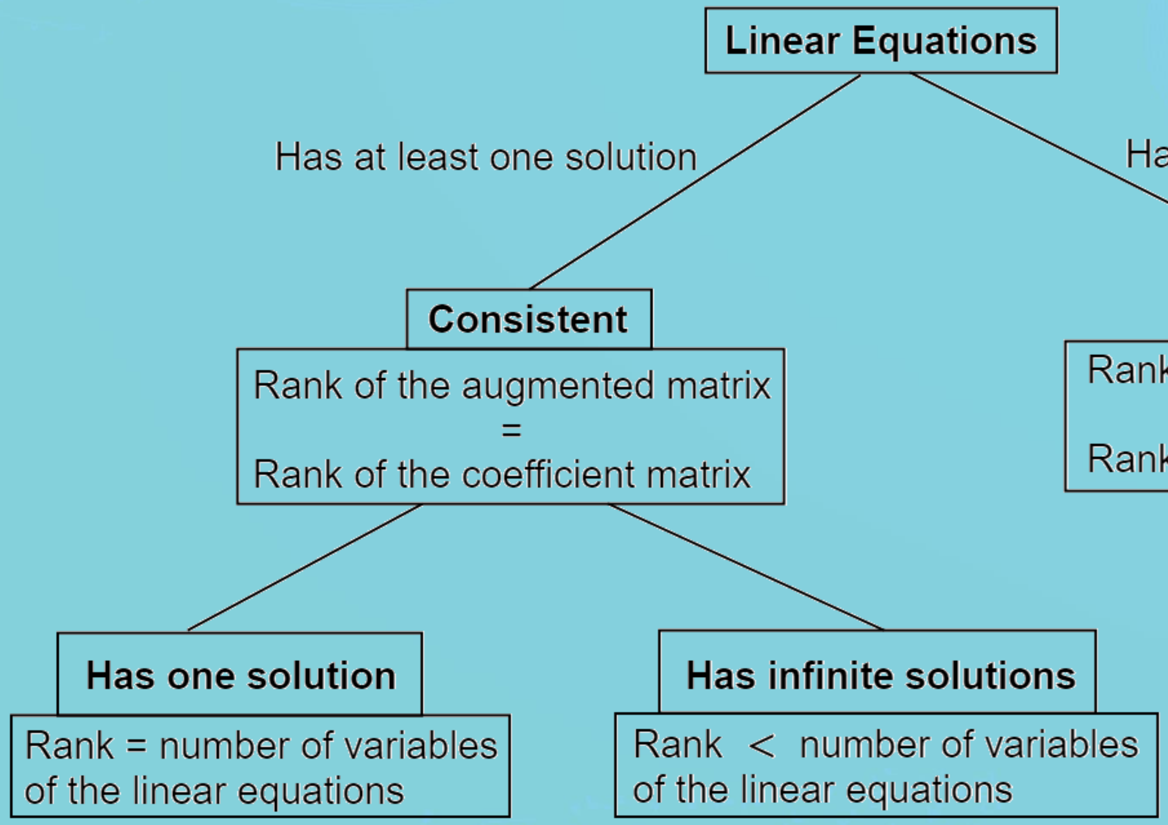
vectors of an orthonormal set are linearly independent.

for Further info and examples click

Gram-Schmidt

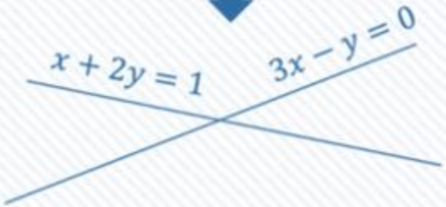
$$
\begin{aligned}
\mathbf{u}_1 &= \mathbf{v}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\
\mathbf{u}_2 &= \mathbf{v}_2 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_2), & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\
\mathbf{u}_3 &= \mathbf{v}_3 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_3), & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \\
\mathbf{u}_4 &= \mathbf{v}_4 - \mathrm{proj}_{\mathbf{u}_1}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_2}(\mathbf{v}_4) - \mathrm{proj}_{\mathbf{u}_3}(\mathbf{v}_4), & \mathbf{e}_4 &= \frac{\mathbf{u}_4}{\|\mathbf{u}_4\|} \\
&\vdots & &\vdots \\
\mathbf{u}_k &= \mathbf{v}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}(\mathbf{v}_k), & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}.
\end{aligned}
$$

Consistency of linear equations



Consistency in 2 dimension

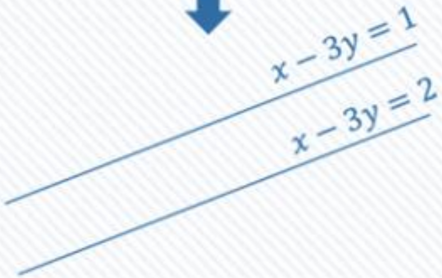
$$x + 2y = 1$$
$$3x - y = 0$$



System of equations is **consistent.** It has **one solution.**

The corresponding matrix $\begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix}$ is **non-singular.**

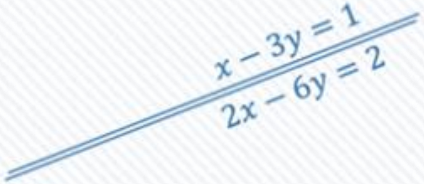
$$x - 3y = 1$$
$$x - 3y = 2$$



System of equations is **inconsistent.** It has **no solutions.**

Matrix $\begin{pmatrix} 1 & -3 \\ 1 & -3 \end{pmatrix}$ is **singular.**

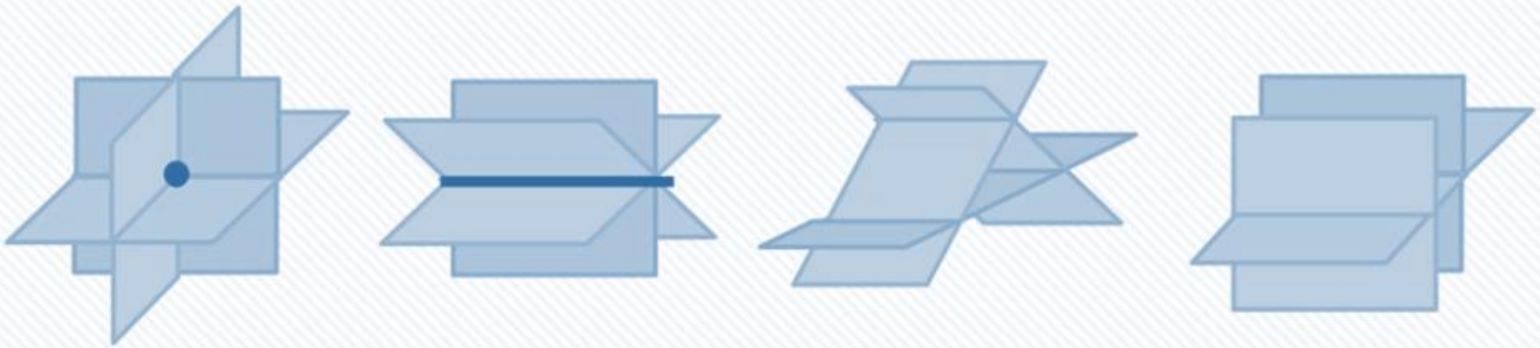
$$x - 3y = 1$$
$$2x - 6y = 2$$



System of equations is **consistent.** It has **infinitely many solutions.**

Matrix $\begin{pmatrix} 1 & -3 \\ 1 & -3 \end{pmatrix}$ is **singular.**

Consistency in 3 dimension



Scenario 1: Planes all meet at a single point.
System of equations consistent, and one solution.

Scenario 2: Planes form a sheaf.
They have a line of intersection consisting of infinitely many points. System of equations consistent and infinitely many solutions.

Scenario 3: Planes form a prism.
While planes intersect in pairs, they don't all intersect at any point. System of equations is inconsistent.

Scenario 4: Two of more planes parallel and non-identical.
Again, inconsistent, as the parallel planes never intersect, and thus all equations can't be satisfied.

Any rows in the corresponding matrix which are multiples of each other will be parallel.

Two scenarios are left. You tell me.

Categorizing systems of equations

m = n	- In this case the number of equations and variables are same - It is the easiest case to solve
m > n	- In this case the number of equations are greater than the variables - Usually solution do not exist for this case
m < n	- In this case the number of equations is less than the number of variables - Usually there is more than one solution

SUPER IMPORTANT

Consistency

System of equations:

$$\begin{aligned} 2x + 5y &= 10 \\ 3x + 4y &= 24 \end{aligned}$$

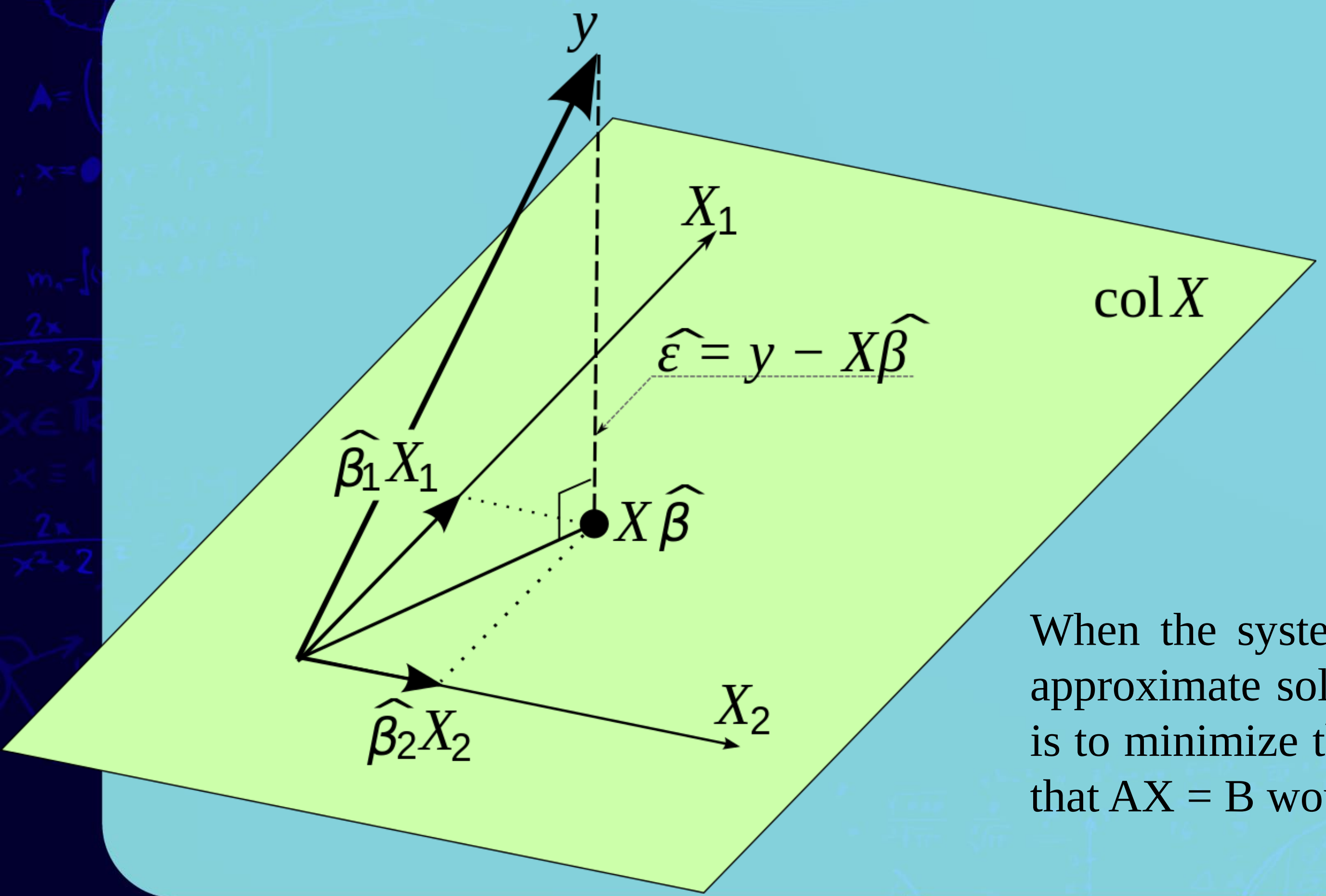
Augmented matrix:

$$\begin{array}{cccl} \left[\begin{array}{cc|c} 2 & 5 & 10 \\ 3 & 4 & 24 \end{array}\right] & \begin{array}{l} \longleftarrow \text{Eq. 1} \\ \longleftarrow \text{Eq. 2} \end{array} \\ \begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ x & y & \text{constants} \end{array} & \end{array}$$

[CLICK](#)

[Super important](#)

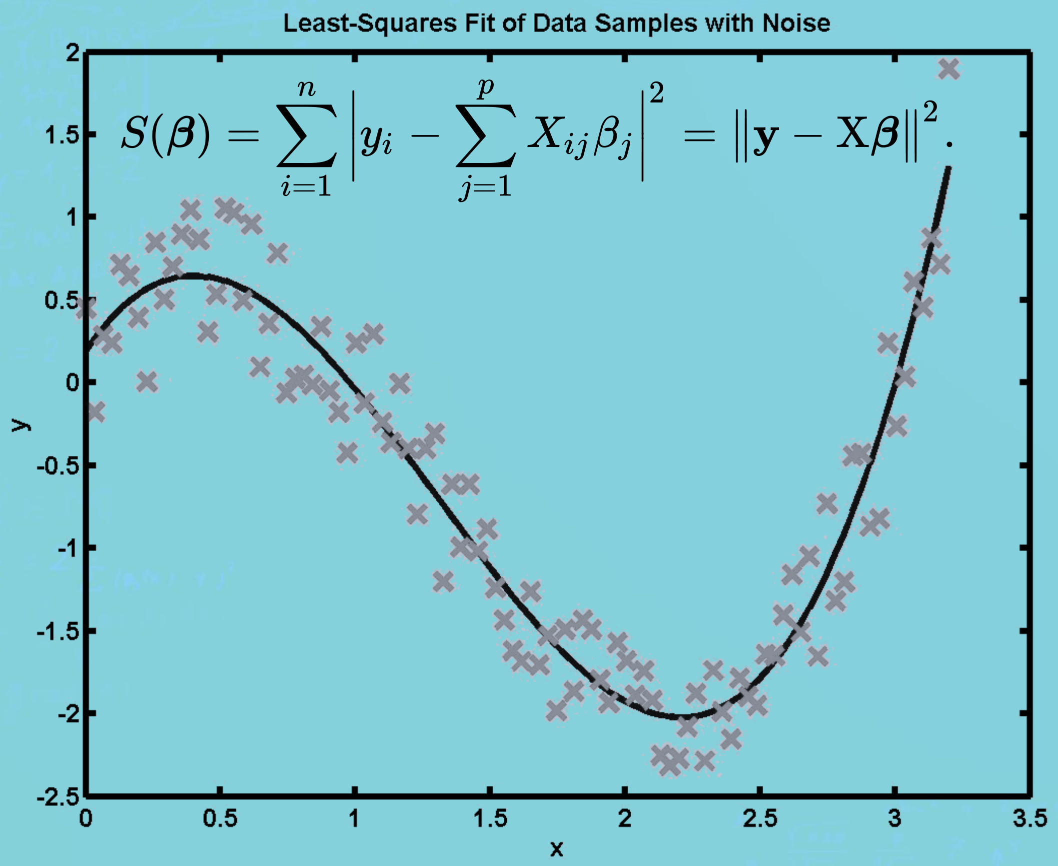
Least squares



In an Linear algebra approach, The column matrix "B" should be in the column space of matrix A, so if the system of equations is inconsistent, the statement before doesn't uphold. It is clear that the closest distance for "B" the residual or error which is the difference of orthogonal projection with the actual B. $\varepsilon = y - A\hat{X}$

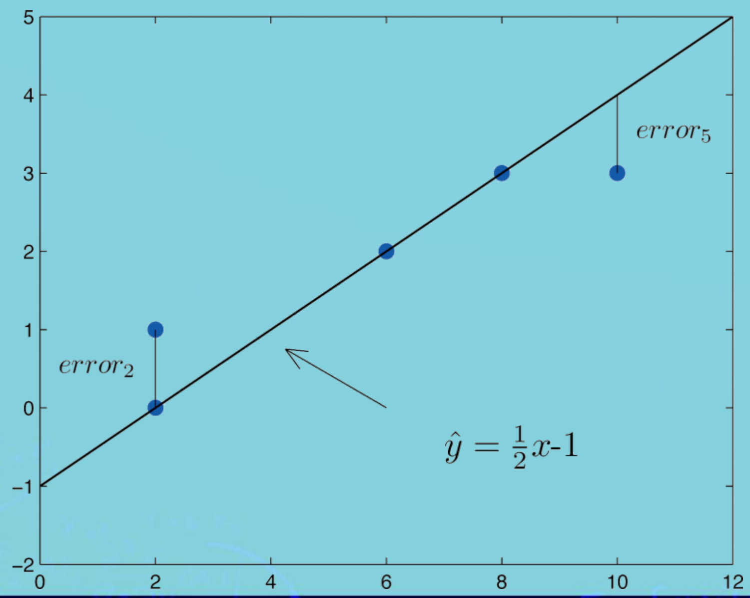
When the system of equation does not have an answer the best approximate solution is called the least-squares solution. The goal is to minimize the norm of $\|B - AX\|$, which comes from the fact that AX = B would be satisfied in a consistent situation

Least squares



$$S(\boldsymbol{\beta}) = \sum_{i=1}^{n} \left| y_i - \sum_{j=1}^{p} X_{ij}\beta_j \right|^2 = \left\| \mathbf{y} - \mathrm{X}\boldsymbol{\beta} \right\|^2.$$

The method of least squares is a standard approach in regression analysis to approximate the solution of overdetermined systems (sets of equations in which there are more equations than unknowns) by minimizing the sum of the squares of the residuals made in the results of every single equation.



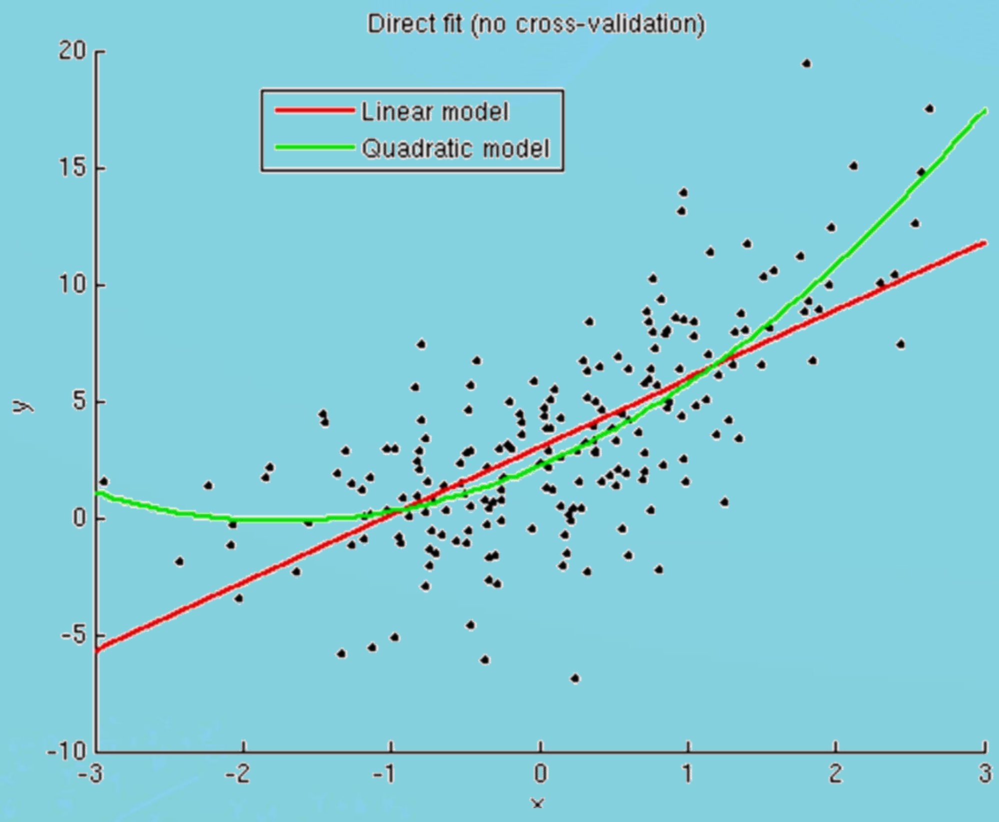
Linear regression

$$A = \frac{(\Sigma y)(\Sigma x^2) - (\Sigma x)(\Sigma xy)}{n(\Sigma x^2) - (\Sigma x)^2}$$

$$B = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{n(\Sigma x^2) - (\Sigma x)^2}$$

Assuming y = AX + B
A referred to as weight or slope,
B referred to as bias or intersection

Online Curve fitter



Normal linear regression equations

- Ordinary least squares (OLS) is the most common estimator. OLS estimates are commonly used to analyze both experimental and observational data.
- In equation below in vector y the ith element is the ith observation
- There are unique implementations for this equation below using Numpy, Sympy, Sklearn and other typical python modules

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{\top}\mathbf{X})^{-1}\mathbf{X}^{\top}\mathbf{y},$$

- In the scope of this course usually the overdetermined equations that will be solved will be solved through the equation above

Some conventions in Geomatics

- Usually each observations is denoted by l_i and the vector containing them Is L
- The variables are usually denoted by x_i and the vectors containing them Is X
- The Jacobian matrix of variables or $\frac{\partial Equations}{\partial X}$ is denoted as A
- The Jacobian matrix of observations or $\frac{\partial Equations}{\partial L}$ is denoted as B
- The hat over any variable means its estimated
- Linearity is a relative attribute

One sided inverses

- For $A : m \times n \mid m > n$ we have left inverses; for example, $\underbrace{\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}}_{A_{\text{left}}^{-1}}A = I_n$
- For $A : m \times n \mid m < n$ we have right inverses; for example, $A\underbrace{A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}}_{A_{\text{right}}^{-1}} = I_m$

Right inverse

$$A : 2 \times 3 = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

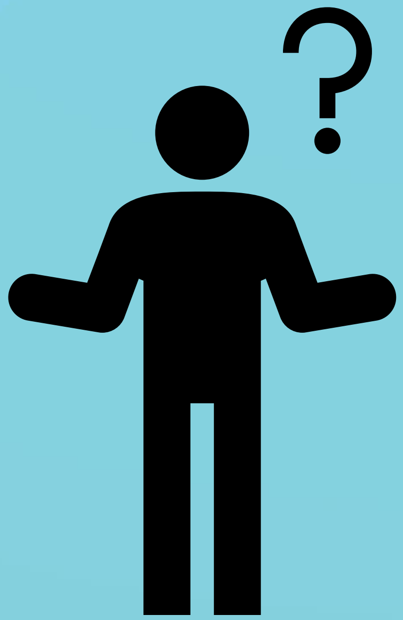
$$AA^{\mathrm{T}} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}$$

$$\left(AA^{\mathrm{T}}\right)^{-1} = \begin{bmatrix} 14 & 32 \\ 32 & 77 \end{bmatrix}^{-1} = \frac{1}{54} \begin{bmatrix} 77 & -32 \\ -32 & 14 \end{bmatrix}$$

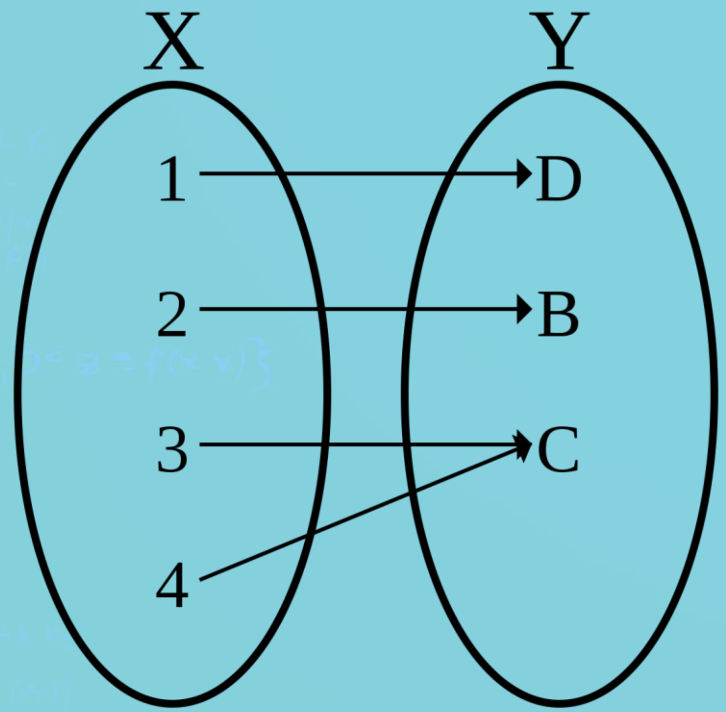
$$A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1} = \frac{1}{54} \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \cdot \begin{bmatrix} 77 & -32 \\ -32 & 14 \end{bmatrix} = \frac{1}{18} \begin{bmatrix} -17 & 8 \\ -2 & 2 \\ 13 & -4 \end{bmatrix} = A^{-1}_{\mathrm{right}}$$

Left inverse for a wide matrix

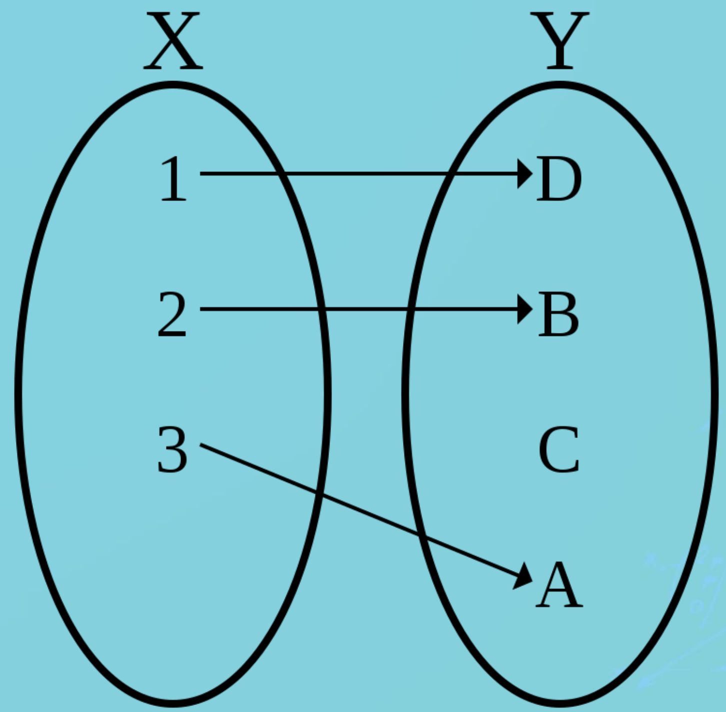
$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} = \begin{bmatrix} 17 & 22 & 27 \\ 22 & 29 & 36 \\ 27 & 36 & 45 \end{bmatrix}$$



Surjection



injection



Formal definition

1. $A^\top A$ is singular when A has full row rank, and more columns than rows. So, there is no inverse in the case. But $AA^\top$ is ok...
2. one sided inverses reflect these properties of linear maps:
   - A is surjective iff there is B such that $AB = I$
   - A is injective iff there is B such that $BA = I$

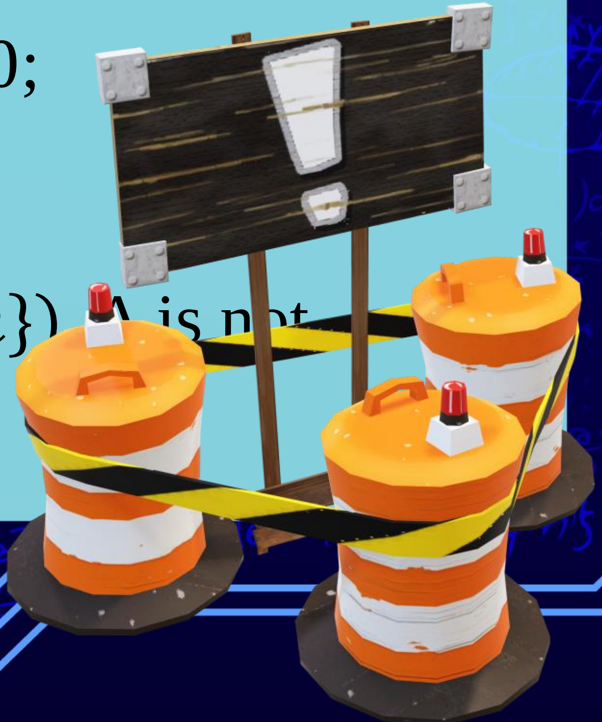
Since linear maps can be of injective/surjective independently, this makes one sided inverses natural.

Important takeaways

Wide matrix → left nullspace = {} so its subjective

tall matrix → nullspace = {} so its injective

- **In general for an $\boldsymbol{m \times n - matrix\ A}$:**
- If the matrix has full rank ($rankA = \min\{m, n\}$), A is:
- injective if $m \geq n = rank(A)$, in that case $\dim(kerA) = 0$;
- surjective if $n \geq m = rank(A)$;
- bijective $if\ m = n = rank(A)$.
- If the matrix does not have full rank ($rank(A) < \min\{m, n\}$), A is not injective/surjective.



Cholesky

- The Cholesky decomposition of a Hermitian positive-definite matrix A is the decomposing a matrix into two part of one lower triangular matrix and one part of upper triangular matrix

$$\mathrm{A} = LL^T$$

$$\begin{pmatrix} 4 & 12 & -16 \\ 12 & 37 & -43 \\ -16 & -43 & 98 \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 6 & 1 & 0 \\ -8 & 5 & 3 \end{pmatrix} \begin{pmatrix} 2 & 6 & -8 \\ 0 & 1 & 5 \\ 0 & 0 & 3 \end{pmatrix}.$$

$$L_{j,j} = \sqrt{A_{j,j} - \sum_{k=0}^{j-1} (L_{j,k})^2}$$

Finding it with block matrices

Notation:
L_{k-1}: the $k-1 \times k-1$ upper left corner of L
a_k: the first $k-1$ entries in column k of A
ℓ_k: the first $k-1$ entries in column k of L^T
a_{kk} and ℓ_{kk}: the kk entries of A and L, respectively.

0) Initialize $L_1 = \sqrt{a_{11}}$
1) For $k = 2, \ldots, n$
   1.1) Solve $L_{k-1}\ell_k = a_k$ for ℓ_k (L_{k-1} is $k-1 \times k-1$: for $k = 2$ this is a 1×1 or scalar equation)
   1.2) $\ell_{kk} = \sqrt{a_{kk} - \ell_k^T \ell_k}$.
   1.3) $L_k = \begin{pmatrix} L_{k-1} & 0 \\ \ell_k^T & \ell_{kk} \end{pmatrix}$.

Cholesky example

For $A = \begin{pmatrix} 16 & 4 & 4 & -4 \\ 4 & 10 & 4 & 2 \\ 4 & 4 & 6 & -2 \\ -4 & 2 & -2 & 4 \end{pmatrix}$ and $b = \begin{pmatrix} 32 \\ 26 \\ 20 \\ -6 \end{pmatrix}$ construct a Cholesky decompostion and solve $Ax = b$

Cholesky example

k = 1: $L_1 = \sqrt{16} = 4$

k=2: $L_1 = 4, a_2 = 4, a_{22} = 10$, Solve the 1×1 system $L_1 \ell_2 = a_2$ or $4\ell_2 = 4$ so $\ell_2 = 1$. $\ell_{22} = \sqrt{10-1} = 3$. Therefore $L_2 = \begin{pmatrix} 4 & 0 \\ 1 & 3 \end{pmatrix}$.

k=3: $L_2 = \begin{pmatrix} 4 & 0 \\ 1 & 3 \end{pmatrix}, a_3 = \begin{pmatrix} 4 \\ 4 \end{pmatrix}, a_{33} = 6$. Solve the 2×2 system $L_2 \ell_3 = a_3$ or $\begin{pmatrix} 4 & 0 \\ 1 & 3 \end{pmatrix} \ell_3 = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$ so $\ell_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. $\ell_{33} = \sqrt{6 - \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}} = 2$. Therefore $L_3 = \begin{pmatrix} 4 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix}$

k=4: $L_3 = \begin{pmatrix} 4 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix}$, $a_4 = \begin{pmatrix} -4 \\ 2 \\ -2 \end{pmatrix}, a_{44} = 4$. Solve the 3×3 system $L_3 \ell_4 = a_4$ or $\begin{pmatrix} 4 & 0 & 0 \\ 1 & 3 & 0 \\ 1 & 1 & 2 \end{pmatrix} \ell_4 = \begin{pmatrix} -4 \\ 2 \\ -2 \end{pmatrix}$ so $\ell_4 = \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}$. $\ell_{44} = \sqrt{4 - \begin{pmatrix} -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 1 \\ -1 \end{pmatrix}} = 1$.

Therefore $L = L_4 = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 1 & 2 & 0 \\ -1 & 1 & -1 & 1 \end{pmatrix}$

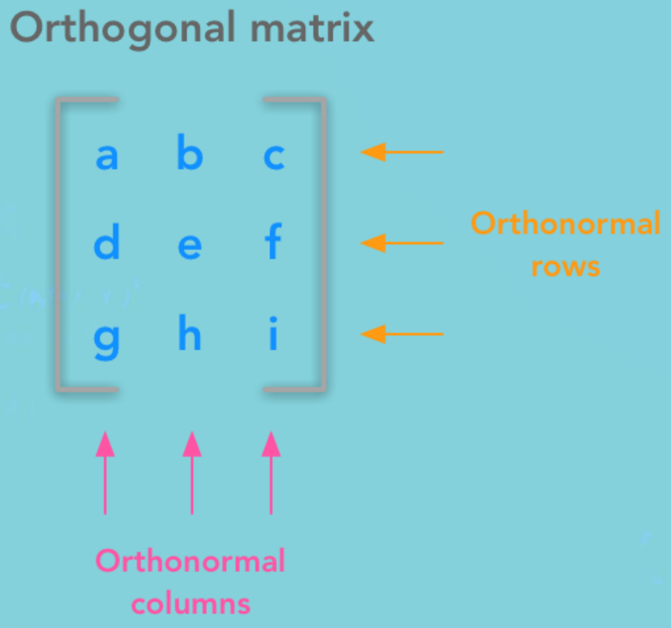
What now?

- $AX = B \rightarrow LL^T X = B \rightarrow L \times Temp = B \rightarrow L^T X = Temp$
- $X = (L^T)^{-1} \times Temp \rightarrow X = (L^T)^{-1} B L^{-1}$

$$Lc = b \Rightarrow c = \begin{pmatrix} 8 \\ 6 \\ 3 \\ -1 \end{pmatrix} \text{ and } L^T x = c \Rightarrow x = \begin{pmatrix} 1 \\ 2 \\ 1 \\ -1 \end{pmatrix}.$$

QR

- In linear algebra, a QR decomposition, also known as a QR factorization or QU factorization, is a decomposition of a matrix A into a product A = QR of an **orthogonal matrix** Q and an **upper triangular** matrix R.



$$A = \begin{bmatrix} \mathbf{a}_1 & \cdots & \mathbf{a}_n \end{bmatrix}$$

$$\begin{aligned}
\mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\
\mathbf{u}_2 &= \mathbf{a}_2 - \mathrm{proj}_{\mathbf{u}_1}\,\mathbf{a}_2, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\
\mathbf{u}_3 &= \mathbf{a}_3 - \mathrm{proj}_{\mathbf{u}_1}\,\mathbf{a}_3 - \mathrm{proj}_{\mathbf{u}_2}\,\mathbf{a}_3, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \\
&\ \vdots & &\ \vdots \\
\mathbf{u}_k &= \mathbf{a}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{u}_j}\,\mathbf{a}_k, & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|}
\end{aligned}$$

QR

$$
\begin{aligned}
\mathbf{a}_1 &= \langle \mathbf{e}_1, \mathbf{a}_1 \rangle\, \mathbf{e}_1 \\
\mathbf{a}_2 &= \langle \mathbf{e}_1, \mathbf{a}_2 \rangle\, \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_2 \rangle\, \mathbf{e}_2 \\
\mathbf{a}_3 &= \langle \mathbf{e}_1, \mathbf{a}_3 \rangle\, \mathbf{e}_1 + \langle \mathbf{e}_2, \mathbf{a}_3 \rangle\, \mathbf{e}_2 + \langle \mathbf{e}_3, \mathbf{a}_3 \rangle\, \mathbf{e}_3 \\
&\vdots \\
\mathbf{a}_k &= \sum_{j=1}^{k} \langle \mathbf{e}_j, \mathbf{a}_k \rangle\, \mathbf{e}_j
\end{aligned}
$$

$$Q = \begin{bmatrix} \mathbf{e}_1 & \cdots & \mathbf{e}_n \end{bmatrix}$$

$$R = \begin{bmatrix} \langle \mathbf{e}_1, \mathbf{a}_1 \rangle & \langle \mathbf{e}_1, \mathbf{a}_2 \rangle & \langle \mathbf{e}_1, \mathbf{a}_3 \rangle & \cdots \\ 0 & \langle \mathbf{e}_2, \mathbf{a}_2 \rangle & \langle \mathbf{e}_2, \mathbf{a}_3 \rangle & \cdots \\ 0 & 0 & \langle \mathbf{e}_3, \mathbf{a}_3 \rangle & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}.$$

$$A = QR$$

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

QR example

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

$$Q^\top A = Q^\top Q\, R = R;$$

$$R = Q^\top A = \begin{bmatrix} 14 & 21 & -14 \\ 0 & 175 & -70 \\ 0 & 0 & 35 \end{bmatrix}.$$

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

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

Apply gram-schmidt
Make a matrix by stacking the normalized vectors in columns
That's your Q ,
Transpose it and matmul it to A from left and that's your R,
You're good to go

What now?

- $A^TAx = A^Tb \rightarrow (QR)^T(QR)x = (QR)^Tb \rightarrow R^TQ^TQRx = R^TQ^Tb$
- $R^TRx = R^TQ^Tb \rightarrow R^TRx = A^Tb \rightarrow Temp = Q^Tb \rightarrow x = R^{-1}Temp$

---

1. **(A few properties of orthogonal matrices)** Let A and B be two $n \times n$ orthogonal matrices so that $A^TA = I_n$ and $B^TB = I_n$. Prove the following properties:

   (a) $\det(A) = 1$ or -1.

   (b) A^{-1} is an orthogonal matrix.

   (c) AB is an orthogonal matrix.

   (d) A^T is an orthogonal matrix.

دانشگاه صنعتی خواجه نصیرالدین طوسی

جبر خطی

(فضای برداری ویژه)

استاد : دکتر محمدرضا ملک

دستیار : امیرحسام طاهرزادگانی

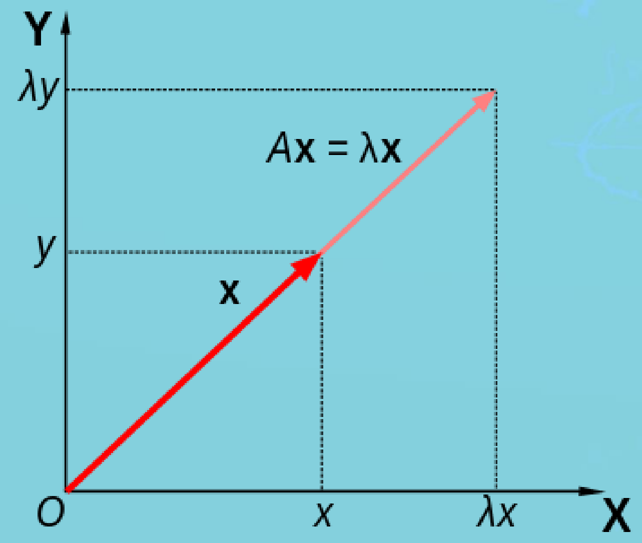
دانشکده مهندسی ژئوماتیک و ژئودزی

Definition

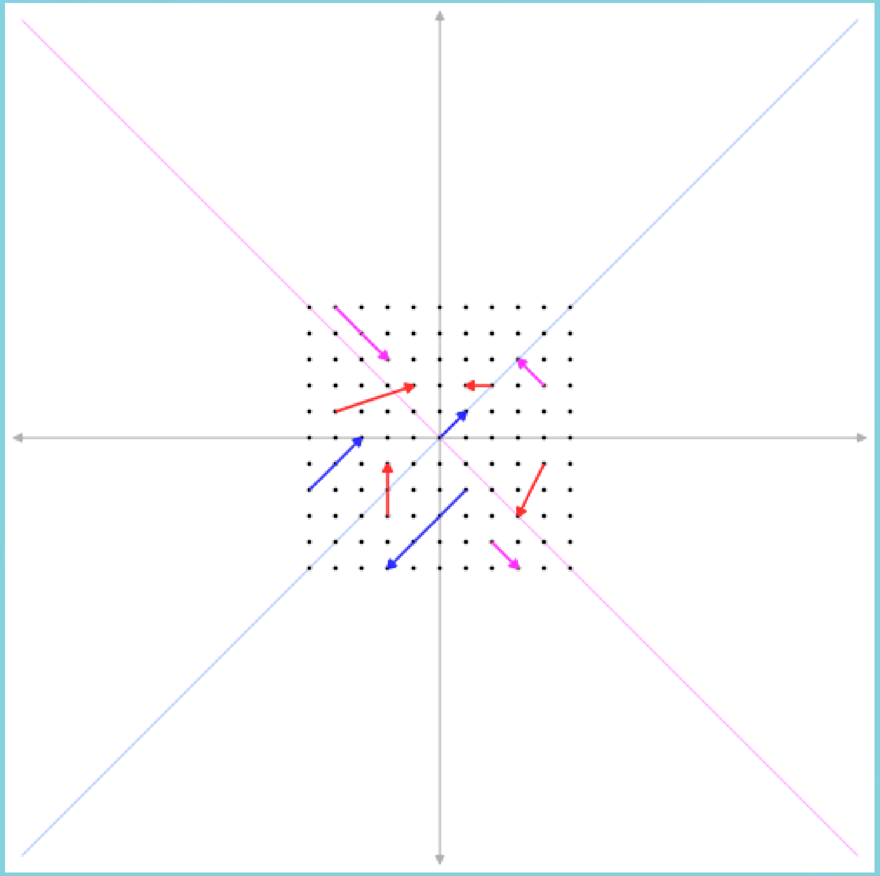
- In linear algebra, an eigenvector or characteristic vector of a linear transformation is a nonzero vector that changes at most by a scalar factor when that linear transformation is applied to it. The corresponding eigenvalue, often denoted by λ is the factor by which the eigenvector is scaled.
- Geometrically, an eigenvector, corresponding to a real nonzero eigenvalue, points in a direction in which it is stretched by the transformation and the eigenvalue is the factor by which it is stretched. If the eigenvalue is negative, the direction is reversed

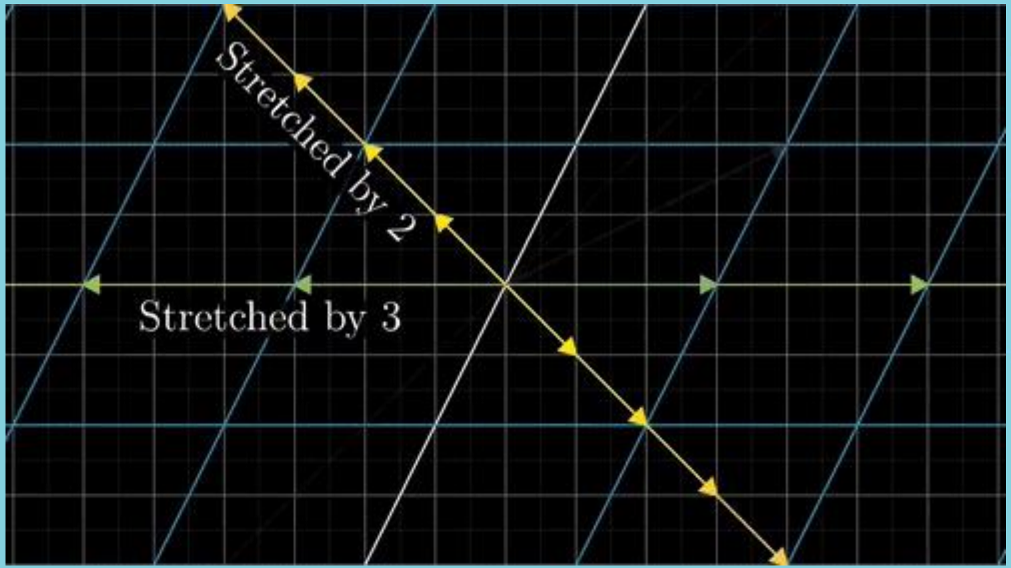
Simply, Any vector and its scaled versions that doesn't rotate in a transformation form an eigenvector and the corresponding coefficient that the vector is stretched by is called eigenvalue, so the goal is:

$$A\mathbf{u} = \lambda\mathbf{u}.$$



interpretation





Formula

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

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

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 3 - 4\lambda + \lambda^2.$$

$$\lambda_1 = 1, \quad \lambda_2 = 3$$

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

online calculator

$$(A - \lambda I)\,\mathbf{v} = \mathbf{0}$$ $\longrightarrow$ Only and if only, has a non zero solution is the determinant of $A - \lambda I$ is zero $\longrightarrow$ $|A - \lambda I| = (\lambda_1 - \lambda)(\lambda_2 - \lambda)\cdots(\lambda_n - \lambda),$

Properties of eigenvalues

Eigenvalues of $\mathbf{A}$ and $\mathbf{A}^T$ are equal

Singular matrix has at least one zero eigenvalue

Eigenvalues of $\mathbf{A}^{-1}$: $1/\lambda_1, 1/\lambda_2, \ldots, 1/\lambda_n$

Eigenvalues of diagonal and triangular matrices are equal to the diagonal elements

Trace

$$Tr(\mathbf{A}) = \sum_{j=1}^{n} \lambda_j$$

Determinant

$$|\mathbf{A}| = \prod_{j=1}^{n} \lambda_j$$

An eigen-basis is a basis in which every vector is an eigenvector.

If λ is an eigenvalue of A and $\alpha \in \mathbb{R}$, then $\lambda + \alpha$ is an eigenvalue of $A + \alpha I$, where I is the identity matrix.

If λ is an eigenvalue of A and $\alpha \in \mathbb{R}$, then $\alpha\lambda$ is an eigenvalue of αA.

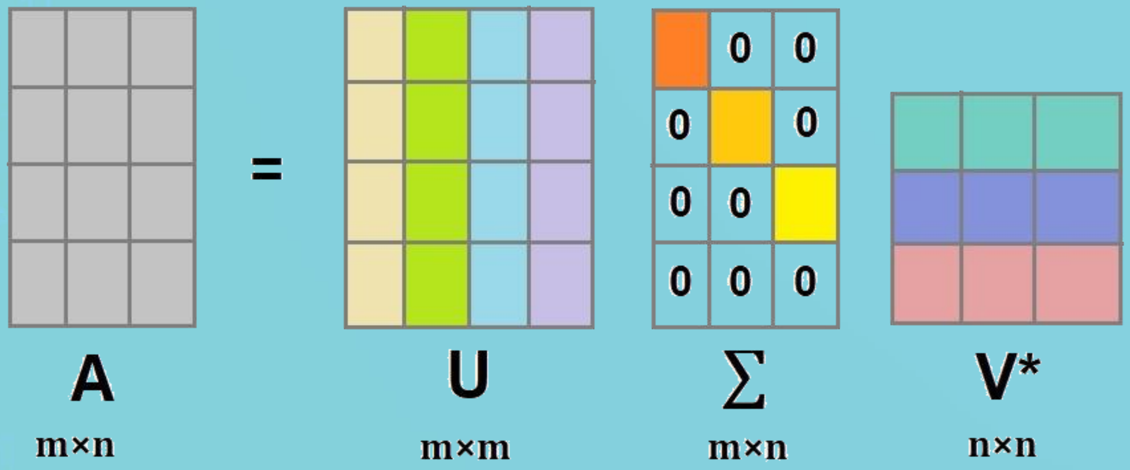
If λ is an eigenvalue of A, then for any positive integer k, λ^k is an eigenvalue of A^k.

دانشگاه صنعتی خواجه نصیرالدین طوسی

جبر خطی
(تجزیه مقادیر منفرد)

استاد : دکتر محمدرضا ملک
دستیار : امیرحسام طاهرزادگانی
دانشکده مهندسی ژئوماتیک و ژئودزی

Svd



Example

- In linear algebra, the singular value decomposition (SVD) is a factorization of a real or complex matrix. It generalizes the eigen-decomposition of a square normal matrix with an orthonormal eigen-basis to any $m \times n$ matrix. where U is an $m \times m$ complex unitary matrix, Σ is an $m \times n$ times n rectangular diagonal matrix with non-negative real numbers on the diagonal, and V is an $n \times n$ times n complex unitary matrix. If M is real , U and V are real matrices.
- $\sigma_{ii} = \Sigma_{ii}$, for $\sigma_{ii} = \sqrt{\lambda_{ii}}$

Svd applications

Rank: the *rank* of a matrix is equal to:

- number of linearly independent columns
- number of linearly independent rows

(Remarkably, these are always the same!).

For an $m \times n$ matrix, the rank must be less than or equal to $\min(m, n)$. The rank can be thought of as the *dimensionality* of the vector space spanned by its rows or its columns.

Lastly, the rank of A is equal to the number of non-zero singular values!

Svd applications

Consider the SVD of a matrix A that has rank k:

$$A = USV^\top$$

Column space: Since A is rank k, the first k left singular vectors, $\{\vec{u}_1, \ldots \vec{u}_k\}$ (the columns of U), provide an orthonormal basis for the column space of A.

Row space: Similarly, the first k right singular vectors, $\{\vec{v}_1, \ldots \vec{v}_k\}$ (the columns of V, or the rows of $V^\top$), provide an orthonormal basis for the row space of A.

Null space: The last right singular vectors, $\{\vec{v}_{k+1}, \ldots \vec{v}_n\}$ (the last columns of V, or the last rows of $V^\top$), provide an orthonormal basis for the null space of A.

Svd applications

Norm of a matrix is equal to square root of the largest value $\lambda_{\max}$ that makes the $A^T A - \lambda I$ to become singular in other word, it's determinant equals zero.

The following values are the eigenvalues of $A^T A$ and the largest one among them will be $\|A\|_2 = \sqrt{\lambda_{\max}} = \sigma_{\max}$

The condition number of a function measures how much the output value of the function can change for a small change in the input argument.

$$\text{Condition number} = \frac{\Delta y}{y} \Big/ \frac{\Delta x}{x}$$

$$\Rightarrow \text{Condition number} = \frac{\Delta x\, f'(x)}{f(x)} \times \frac{x}{\Delta x}$$

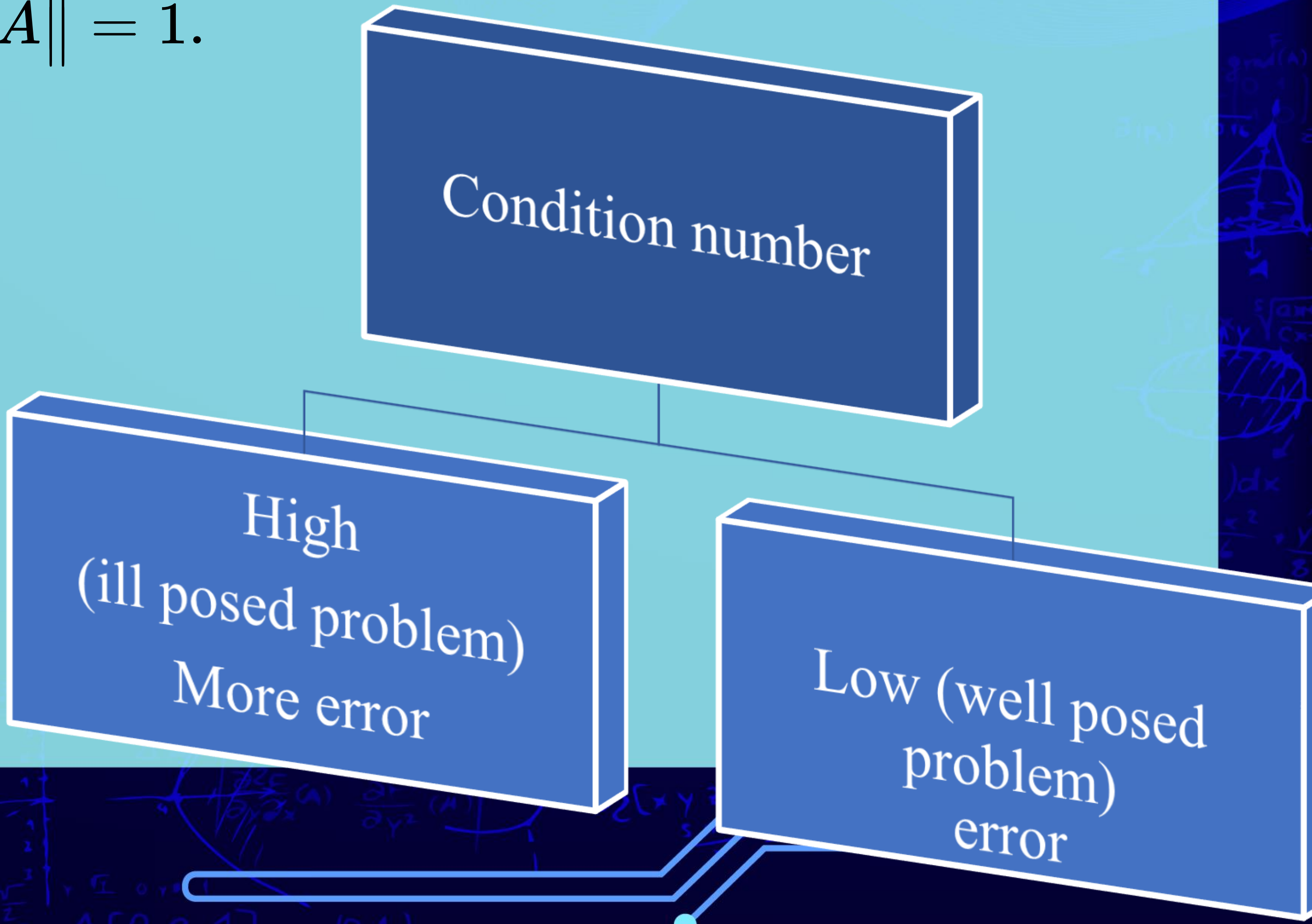
$$\boxed{\Rightarrow \text{Condition number} = \frac{x\, f'(x)}{f(x)}}$$

Condition number

$$\kappa(A) = \left\|A^{-1}\right\| \, \|A\| \geq \left\|A^{-1}A\right\| = 1.$$

$$\kappa(A) = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)},$$

$$\kappa(A) = \frac{|\lambda_{\max}(A)|}{|\lambda_{\min}(A)|},$$



Generalized inverse

- a generalized inverse (or, g-inverse) of an element x is an element y that has some properties of an inverse element but not necessarily all of them

Table 3.5. *Equivalent forms for the pseudoinverse matrix.*

	shape	row, col, rank	equivalence	form
1.	square	$m = n = \rho$	$\mathbf{A}^\dagger = \mathbf{A}^{-1}$	
2.	tall	$m > n,\ n = \rho$	$\mathbf{A}^\dagger = \mathbf{A}^{-L}$	$= (\mathbf{A}^* \mathbf{A})^{-1}\mathbf{A}^*$
3.	wide	$m = \rho,\ n > m$	$\mathbf{A}^\dagger = \mathbf{A}^{-R}$	$= \mathbf{A}^* (\mathbf{A}\mathbf{A}^*)^{-1}$
4.	arbitrary	$m \neq n,\ m \neq \rho,\ n \neq \rho$	$-$	

A generalized inverse is an extension of the concept of inverse that applies to square singular matrices and rectangular matrices. A generalized inverse always exists although it is not unique in general.

Here is the only constraint for general inverse not moore-pennrose

$$AA^{g}A = A.$$

Pseudo inverse

$$AA^{g}A = A$$

$$A^{g}AA^{g} = A^{g}$$

$$(AA^{g})^{*} = AA^{g}$$

$$(A^{g}A)^{*} = A^{g}A,$$

If the generalized inverse also gets constrained with the second to last equation, the output will definitely be a **unique** matrix which we denote as **Moore-pennrose or Pseudo inverse**

- For $A : m \times n \mid m > n$ we have left inverses; for example, $\underbrace{\left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}}_{A_{\text{left}}^{-1}}A = I_n$
- For $A : m \times n \mid m < n$ we have right inverses; for example, $A\underbrace{A^{\mathrm{T}}\left(AA^{\mathrm{T}}\right)^{-1}}_{A_{\text{right}}^{-1}} = I_m$