NULLSPACE OF A

Def^n The nullspace $N(A)$ consists of all solutions to $A\vec{x} = \vec{0}$.

If $A \in \mathbb{R}^{m \times n}$, the $N(A)$ is a subspace of $\mathbb{R}^n$.

Why? Suppose $\vec{x}, \vec{y}$ are in the nullspace.

Then $A\vec{x} = \vec{0}$ and $A\vec{y} = \vec{0}$

Using rules for matrix multiplication, $A(\underbrace{\vec{x}+\vec{y}}) = \vec{0}$

Then $\vec{x}+\vec{y}$ is also in the nullspace.

Similarly, $c(A\vec{x}) = c\,\vec{0}$

$\Rightarrow A(c\vec{x}) = \vec{0}$

$\Rightarrow c\vec{x}$ is in the nullspace.

Linear combinations are in the nullspace. Therefore $N(A)$ is a subspace.

SOLVING $A\vec{x} = \vec{b}$ IN COMPLETE GENERALITY

Def^n A matrix is in <u>reduced row echelon form</u> (rref) if it is <u>both</u>

(i) upper triangular and

(ii) contains only a single non-zero entry - a '1' - in each pivot column.

<u>Example</u>

$$\begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 0 \end{bmatrix} \xrightarrow[\substack{-2\,\text{row}①\\ +\,\text{row}②}]{\text{(elimination top to bottom)}} \begin{bmatrix} 1 & 1 & 3 \\ 0 & -1 & -6 \end{bmatrix} \xrightarrow[]{\substack{\text{(elimination bottom to top)}\\ \text{row}② + \text{row}①}} \begin{bmatrix} 1 & 0 & -3 \\ 0 & -1 & -6 \end{bmatrix}$$

Upper triangular

$$\xrightarrow{\text{row}② \to -1\,\text{row}②} \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 6 \end{bmatrix}$$

Reduced Row Echelon Form

pivot columns   free column

Solving $A\vec{x} = \vec{0}$ when A is in reduced row echelon form

From our previous example, $A = \begin{bmatrix} 1 & 1 & 3 \\ 2 & 1 & 0 \end{bmatrix}$ and $R = \text{rref}(A) = \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 6 \end{bmatrix}$

$A\vec{x} = \vec{0}$ and $R\vec{x} = \vec{0}$ have same solution

Solving $R\vec{x} = \vec{0}$, we have $\begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

Pivot cols. free col.

z is a "free" variable

$$\begin{cases} x - 3z = 0 & \text{or} \quad x = 3z \\ y + 6z = 0 & \text{or} \quad y = -6z \end{cases}$$

<u>How to choose z?</u>

$z = 0$ (trivial solution) $x = y = 0$

$z = 1$ <u>special solution</u> $x = 3$; $y = -6$

<u>Another Example</u> Let $R = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$

free cols

pivot col pivot col

(cols. with the first 1 in each row)

reduced row echelon form
- upper triangular
- each pivot col. has only 1 as non-zero entry

Solving $R\vec{x} = \vec{0}$ $\begin{bmatrix} 1 & \boxed{2} & 0 & \boxed{1} \\ 0 & \boxed{0} & 1 & \boxed{1} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$

$\boxed{\text{2 Special solns}}$ choose $\boxed{y = 1,\ w = 0;}$ we get $x = -2$, $z = 0$

$\boxed{y = 0,\ w = 1;}$ $x = -1$, $z = -1$

Since we have 2 "free" variables or cols.

So $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} = \vec{0}$

$$\boxed{N(R) = N(A) = \text{Span}\left(\begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \end{bmatrix} \right)}$$

Column 2 is a linear combination of column 1

col. 4 is a linear combination of cols. 1 and 3

$$A = \begin{bmatrix} 2 & 1 & 4 & 4 \\ 3 & 2 & 6 & 7 \\ 1 & 3 & 2 & 7 \end{bmatrix} \qquad R = \begin{bmatrix} \boxed{\begin{matrix}1\\0\\0\end{matrix}} & \boxed{\begin{matrix}0\\1\\0\end{matrix}} & \begin{matrix}2\\0\\0\end{matrix} & \begin{matrix}1\\2\\0\end{matrix} \end{bmatrix}$$

$A \in \mathbb{R}^{3\times 4}$

pivot cols "free" cols.

Given matrix

Reduced Echelon form

Notes

- there are 2 pivots / pivot columns
- (column 3) = 2 (colum 1)

  (column 4) = 1 (colum 1) + 2 (colum 2)
- "free" columns are a linear combination of earlier pivot columns.

Def^n (RANK) The rank of A (denoted by r) is the number of pivots

Easy to see that $r \leq m$ and $r \leq n$. In our example, rank $(A) = r = 2$.

Remarks

- All columns of $A \in \mathbb{R}^{m\times n}$ are either free or pivot columns

$$\underbrace{\#\text{ columns}}_{n} = \underbrace{\#\text{ pivot columns}}_{\substack{r \\ (\text{Rank of } A)}} + \underbrace{\#\text{ free columns}}_{n-r}$$

- $N(A)$ is always "spanned" by $n-r$ vectors (corresponding to the number of free columns).

All solutions to $A\vec{x} = \vec{b}$

$$\underbrace{\begin{pmatrix} 3 & 3 \\ 2 & 2 \end{pmatrix}}_{A} \vec{x} = \vec{b}$$

First, note that $A\vec{x} = \vec{b}$ has a solution only if $\vec{b} \in C(A)$

Consider $\vec{b} = \begin{pmatrix} -6 \\ -4 \end{pmatrix}$.

Performing elimination on $[A \vdots \vec{b}]$, we have

$$\left[\begin{array}{cc:c} 3 & 3 & -6 \\ 2 & 2 & -4 \end{array}\right] \xrightarrow[-\frac{2}{3} R_1 + R_2]{} \left[\begin{array}{cc:c} 3 & 3 & -6 \\ 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1/3} \left[\begin{array}{cc:c} \boxed{1} & 1 & \boxed{-2} \\ 0 & 0 & 0 \end{array}\right] \quad \Big| \quad x_1 + x_2 = -2$$

pivot

free

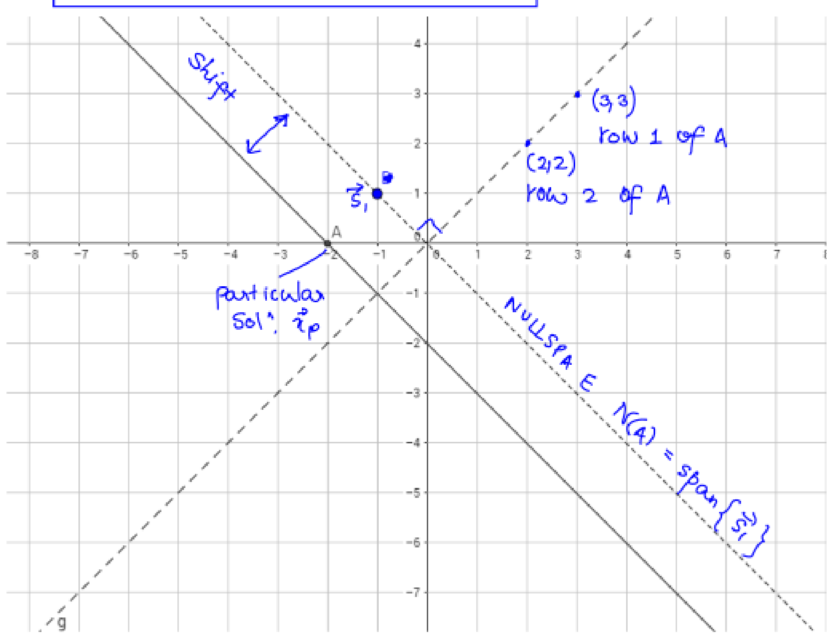
(1 free variable) Special solution $\vec{s}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ (solution to $A\vec{x} = \vec{0}$) $\left| \left[\begin{array}{cc:c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \right.$ $x_1 + x_2 = 0$

$\boxed{\text{Particular}}$ Solution choose $\underline{x_2 = 0}$, then $x_1 = -2$; i.e., $\begin{pmatrix} -2 \\ 0 \end{pmatrix}$

Complete solution is $\boxed{\vec{x} = \vec{x}_p + \vec{x}_n = \begin{pmatrix} -2 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}}$

$\boxed{\text{Note}}$

- All linear combinations of rows of A are perpendicular to $N(A)$
- the particular solution tells you where to shift/slide $N(A)$
- all solutions to $A\vec{x} = \vec{b}$ form a "hyperplane" (the nullspace $N(A)$, $n-r$ dimensional) shifted so that it passes through a point given by the particular sol^n.



$\vec{x}_{nullspace}$	$n-r$ special solutions to $A\vec{x}_n = \vec{0}$
$\vec{x}_{particular}$	Particular solution solves $A\vec{x}_p = \vec{b}$

$$\boxed{\text{Complete solution} \begin{pmatrix} \text{one } \vec{x}_p \\ \text{many } \vec{x}_n \end{pmatrix} \quad \vec{x} = \vec{x}_p + \vec{x}_n}$$

Linear Independence

Def^n The sequence of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ is <u>linearly independent</u> if the only combination that gives the zero vector is
$0\vec{v}_1 + 0\vec{v}_2 + \cdots + 0\vec{v}_n = \vec{0}$.

Def^n (Alternate) Vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathbb{R}^m$ are <u>linearly independent</u> if $A = \begin{pmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{pmatrix} \in \mathbb{R}^{m \times n}$ has $N(A) = \{\vec{0}\}$.

Equivalently vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathbb{R}^m$ are linearly independent if <u>A has rank n</u>
(full column rank, n pivots, no free variables)

Recap

$\{\vec{a}_1, \ldots, \vec{a}_n\}$ are linearly independent if and only if

(i) $A = \begin{pmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{pmatrix}$ has $N(A) = \{\vec{0}\}$.

(ii) rank of A is n

(iii) $\nexists \ \vec{a}_j \in \text{span}\{\vec{a}_1, \ldots, \vec{a}_{j-1}, \vec{a}_{j+1}, \ldots, \vec{a}_n\}$
(there does not exist)

(iv) there is no possible way to write any column as a linear combination of the others.

(v) $C(A) = \text{span}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n\}$ can never be written as the span of fewer than n vectors.

Row Space

Def^n Let $A \in \mathbb{R}^{m \times n}$. The row space of a matrix is the subspace of $\mathbb{R}^n$ spanned by the rows.

The row space of A is $C(A^T)$.

Example: Let $A = \begin{bmatrix} 1 & 2 \\ 4 & 1 \\ 7 & 0 \end{bmatrix}$ $\quad C(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ 4 \\ 7 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} \right\}$

$(A \in \mathbb{R}^{3 \times 2})$

Also $A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 1 & 0 \end{bmatrix} \in \mathbb{R}^{2 \times 3}$

rowspace $= C(A^T) = \text{span}\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} \begin{pmatrix} 7 \\ 0 \end{pmatrix} \right\}$

Basis

Def^n A basis for a vector space is a sequence of vectors with two properties:

(i) the basis vectors are linearly independent

(ii) the basis vectors span the space.

Note

* there is <u>only</u> one way to write a given vector $\vec{v}$ in the space as a linear combination of the basis vectors.

* <u>Examples</u>

standard basis in $\mathbb{R}^2$	$\vec{i} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $\vec{j} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	independent, span $\mathbb{R}^2$	can be interpreted as columns of $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

* the columns of every invertible $n \times n$ matrix form a basis for $\mathbb{R}^n$

* Pivot columns of A are a basis for $C(A)$ - the column space
  pivot rows of A are a basis for its row space (as are the rows of echelon form R)
  special solutions of A are a basis for $N(A)$

DIMENSION OF A VECTOR SPACE

Defn The dimension of a space is the number of vectors in (any and) every basis

REMARKS:

(1) If $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m$ and $\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_n$ are both bases for the same vector space, then $m=n$.

- the number of basis vectors depends on the space; not on a particular basis

(2) For $A \in \mathbb{R}^{m \times n}$,

- $C(A)$ has dimension $= r =$ rank of A
- $N(A)$ has dimension $= n-r$

FOUR FUNDAMENTAL SUBSPACES

Let $A \in \mathbb{R}^{m \times n}$

1. The ROW SPACE is $C(A^T)$. It is a subspace of $\mathbb{R}^n$
2. The COLUMN SPACE is $C(A)$. — " — of $\mathbb{R}^m$
3. The NULLSPACE is $N(A)$ — " — of $\mathbb{R}^n$
4. The LEFT NULLSPACE is $N(A^T)$ — " — of $\mathbb{R}^m$

Note: The ROWSPACE AND COLUMNSPACE have dimension $r =$ rank of the matrix

The NULLSPACE, $N(A)$, has dimension $n-r$.

The LEFT NULLSPACE has dimension $m-r$.

FUNDAMENTAL THEOREM OF LINEAR ALGEBRA

Let A be an $m \times n$ matrix

(PART I)

> The column space and row space both have dimension r.
> The nullspaces have dimensions $n-r$ and $m-r$.

Alternatively,

$$\dim C(A) + \dim N(A) = n$$

(Sometimes called rank-nullity theorem)

$$\dim C(A^T) + \dim N(A^T) = m$$

GEOMETRY OF THE FUNDAMENTAL SUBSPACES

Similarly basis vector(s) for $N(A^T)$ and $C(A)$ are perpendicular to each other

It follows that any vector in $N(A^T)$ is perpendicular to any vector in $C(A)$.

$$\boxed{C(A) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 2 \\ \pi \end{pmatrix}, \begin{pmatrix} -1 \\ 3 \\ 0 \end{pmatrix} \right\} \quad N(A^T) = \operatorname{span} \left\{ \begin{pmatrix} -3/5 \\ -1/5 \\ 1/\pi \end{pmatrix} \right\}}$$

$\boxed{\text{Def}^{n}}$ (ORTHOGONAL SUBSPACES)

Two subspaces V and W of a vector space are <u>orthogonal</u> if every vector $\vec{v} \in V$ is perpendicular to every vector $\vec{w} \in W$.

i.e., $\vec{v} \cdot \vec{w} = \vec{v}^T \vec{w} = 0$ for all $\vec{v} \in V$, $\vec{w} \in W$.

<u>Notation:</u>

$W \perp V$

We have

> Nullspace $N(A)$ and Row space $C(A^T)$ are orthogonal subspaces of $\mathbb{R}^n$
> Left Nullspace $N(A^T)$ and Column space $C(A)$ are orthogonal subspaces of $\mathbb{R}^m$

ORTHOGONAL COMPLEMENTS

Defn The orthogonal complement of subspace V contains every vector that is perpendicular to V.

This orthogonal subspace is denoted by $\underbrace{V^{\perp}}$

pronounced "V perp"

FUNDAMENTAL THEOREM OF LINEAR ALGEBRA (PART 2)

The nullspace $N(A)$ is the orthogonal complement of the row space $C(A^T)$ (in $\mathbb{R}^n$)

The left nullspace $N(A^T)$ is the orthogonal compliment of the column space $C(A)$ (in $\mathbb{R}^m$)

Why is this useful?

Every $\vec{x}$ can be split into $\underbrace{\vec{x}_r}_{\text{row space component}} + \underbrace{\vec{x}_n}_{\text{nullspace component}}$ | $\vec{x} = \vec{x}_r + \vec{x}_n$

PROJECTION OF $\vec{b}$ ONTO THE LINE THROUGH $\vec{a}$

$$\boxed{\hat{x} = \frac{\vec{a} \cdot \vec{b}}{\vec{a} \cdot \vec{a}} = \frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}}$$

Projection of $\vec{b}$ onto the line through $\vec{a}$ is the vector $\boxed{\vec{p} = \hat{x}\,\vec{a} = \left(\frac{\vec{a}^T \vec{b}}{\vec{a}^T \vec{a}}\right) \vec{a}}$

Projection Matrix $\boxed{P = \frac{\vec{a}\,\vec{a}^T}{\vec{a}^T \vec{a}}}$

Note:
- P is rank-1 with $P^2 = P$
- we are projecting onto a one-dimensional subspace (line)
- this line is the column space of P

Special cases
1. $\vec{b} = \vec{a}$ $(\hat{x} = 1,\ \vec{p} = \vec{a})$
2. $\vec{b}$ is perpendicular to $\vec{a}$ $(\hat{x} = 0,\ \vec{p} = \vec{0})$

$\boxed{\text{PROJECTION ONTO A SUBSPACE}}$

Note: $A \in \mathbb{R}^{m\times n}$
$\vec{p} = \hat{x}_1 \vec{a}_1 + \ldots + \hat{x}_n \vec{a}_n = A\vec{\hat{x}}$

To find $\hat{x}$ (vector, $n\times 1$)

Solve $\boxed{A^T(\vec{b} - A\vec{\hat{x}}) = 0 \quad \text{or} \quad \underline{A^TA\,\vec{\hat{x}} = A^T\vec{b}}}$ normal equations

$A^TA \in \mathbb{R}^{n\times n}$, invertible if $\vec{a}_i$ are independent

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

Projection of $\vec{b}$ onto the subspace is $\vec{p}$ (vector, $m\times 1$)

$$\boxed{\vec{p} = A\vec{\hat{x}} = A(A^TA)^{-1}A^T\vec{b}}$$

Projection matrix, $P \in \mathbb{R}^{m\times m}$

$$\boxed{P = A(A^TA)^{-1}A^T}$$

$\boxed{\text{PROPERTIES OF } P}$

$\boxed{\text{Theorem}}$ A^TA is invertible if $A \in \mathbb{R}^{m\times n}$ has rank n

or

A^TA is invertible if A has linearly independent columns.

$$* \quad P^2 = PP = \left(A(A^TA)^{-1}A^T\right)\left(A(A^TA)^{-1}A^T\right)$$
$$= A(A^TA)^{-1}\underbrace{(A^TA)(A^TA)^{-1}}_{=I}A^T)$$
$$= A(A^TA)^{-1}A^T$$
$$= P$$

$$* \quad P^T = \left(A(A^TA)^{-1}A^T\right)^T = A\left((A^TA)^{-1}\right)^TA^T = A\left((A^TA)^T\right)^{-1}A^T$$
$$= A(A^TA)^{-1}A^T$$
$$= P$$

$\Rightarrow P$ is symmetric

ORTHONORMAL BASIS

Defn An orthonormal basis $\{\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_n\}$ is a set of vectors with

$$\vec{q}_i \cdot \vec{q}_j = \vec{q}_i^{\,T} \vec{q}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

- Each vector, $\vec{q}_j$ in the orthonormal basis (ONB) must have length 1 since $\vec{q}_j \cdot \vec{q}_j = \|\vec{q}_j\| = 1$.

Example $\left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \right\}$ is an orthonormal basis for $\mathbb{R}^2$.

ORTHOGONAL MATRIX

Defn An orthogonal matrix, $Q \in \mathbb{R}^{n \times n}$, is a matrix whose column vectors are an orthonormal basis.

If $Q = \begin{pmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{pmatrix}$,

then $$Q^T Q = \begin{pmatrix} — \vec{q}_1 — \\ — \vec{q}_2 — \\ \vdots \\ — \vec{q}_n — \end{pmatrix} \begin{pmatrix} | & | & & | \\ \vec{q}_1 & \vec{q}_2 & \cdots & \vec{q}_n \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} \underbrace{\vec{q}_1 \cdot \vec{q}_1}_{1} & \cancelto{0}{\vec{q}_1 \cdot \vec{q}_2} & \cdots & \\ \cancelto{0}{\vec{q}_2 \cdot \vec{q}_1} & \underbrace{\vec{q}_2 \cdot \vec{q}_2}_{1} & \vec{q}_2 \cdot \vec{q}_3 & \cdots \\ & & \ddots & \\ \vec{q}_n \cdot \vec{q}_1 & \vec{q}_n \cdot \vec{q}_1 & \cdots & \vec{q}_n \cdot \vec{q}_n \end{pmatrix} = I \in \mathbb{R}^{n \times n}$$

Consequences of Orthogonality

① If $Q \in \mathbb{R}^{n \times n}$ is orthogonal, then $Q^T = Q^{-1}$!

(Why? $Q^T Q = I$
$\Rightarrow (Q^T Q) Q^{-1} = I Q^{-1}$
$\Rightarrow Q^T (Q Q^{-1}) = Q^{-1}$
$\Rightarrow Q^T = Q^{-1}$)

② If $Q \in \mathbb{R}^{n \times n}$ is orthogonal, then

$$\begin{aligned} \|Q\vec{x}\|^2 = (Q\vec{x}) \cdot (Q\vec{x}) &= (Q\vec{x})^T (Q\vec{x}) \\ &= \vec{x}^T \underbrace{(Q^T Q)}_{I} \vec{x} \\ &= \vec{x}^T \vec{x} \\ &= \vec{x} \cdot \vec{x} \\ &= \|\vec{x}\|^2 \end{aligned}$$

Hence, Q preserves the length of all vectors $\vec{x} \in \mathbb{R}^n$

Consequences of Orthogonality

③ Let $Q \in \mathbb{R}^{n \times n}$ be orthogonal. Then $Q\vec{x} = \vec{b}$ has solution $\vec{x} = Q^T \vec{b}$

$$\vec{x} = \begin{pmatrix} \vec{q}_1 \cdot \vec{b} \\ \vdots \\ \vec{q}_n \cdot \vec{b} \end{pmatrix}$$

(Why? $Q\vec{x} = \vec{b}$
$\Rightarrow Q^{-1}(Q\vec{x}) = Q^{-1}\vec{b}$
$\Rightarrow \underbrace{(Q^{-1}Q)}_{I} \vec{x} = Q^{-1}\vec{b}$
$\Rightarrow \vec{x} = Q^{-1}\vec{b} = Q^T \vec{b}$)

Hence $\vec{x} = \sum_{j=1}^{n} (\vec{q}_j \cdot \vec{x}) \vec{q}_j \quad \forall \vec{x} \in \mathbb{R}^n$

$\rightsquigarrow$ Rewriting $\vec{x}$ as a sum of an orthonormal basis

④ Let $Q \in \mathbb{R}^{m \times n}$ be orthogonal; then projecting onto $C(Q)$ is simple.

(Projection matrix) $P = Q \underbrace{(Q^T Q)}_{I}^{-1} Q^T = Q Q^T$

Finding Orthonormal Bases - Gram-Schmidt

Suppose we want an orthonormal basis for span$\{\vec{a}_1, \vec{a}_2, \ldots\ldots, \vec{a}_n\}$. We can build one using Gram-Schmidt procedure.

Procedure

Step 0 $\quad \vec{q}_1 = \dfrac{\vec{a}_1}{\|\vec{a}_1\|}$ (normalizing to unit length)

Step 1 $\quad \vec{q}_2 = \dfrac{\vec{a}_2 - P_{\text{span}\{\vec{q}_1\}}\vec{a}_2}{\left\| \vec{a}_2 - P_{\text{span}\{\vec{q}_1\}}\vec{a}_2 \right\|}$ we have $P_{\text{span}\{\vec{q}_1\}}\vec{a}_2 = (\vec{q}_1 \cdot \vec{a}_2)\,\vec{q}_1$

Step 2 $\quad \vec{q}_3 = \dfrac{\vec{a}_3 - P_{\text{span}\{\vec{q}_1, \vec{q}_2\}}\vec{a}_3}{\left\| \vec{a}_3 - P_{\text{span}\{\vec{q}_1, \vec{q}_2\}}\vec{a}_3 \right\|}$ where $P_{\text{span}\{\vec{q}_1, \vec{q}_2\}}\vec{a}_3 = (\vec{q}_1 \cdot \vec{a}_3)\,\vec{q}_1 + (\vec{q}_2 \cdot \vec{a}_3)\,\vec{q}_2$

...

continue

QR Factorization

Any $m \times n$ matrix with independent columns can be factored into QR.

- the matrix Q has orthonormal columns
- the square matrix R which upper triangular with positive entries on the diagonal.

In the previous problem, we have

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & 2 \\ 1 & -1 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 2 & 3 & 2 \\ 0 & 5 & -2 \\ 0 & 0 & 4 \end{bmatrix}}_{R}$$

Why do we care about QR?

Consider least squares solutions $(A^T A \hat{x} = A^T \vec{b})$ $\xrightarrow[A = QR]{\text{use}}$

$$(QR)^T (QR)\,\hat{x} = (QR)^T \vec{b} \Rightarrow R^T \underbrace{(Q^T Q)}_{I} R\,\hat{x} = R^T Q^T \vec{b} \Rightarrow \underbrace{R}_{\text{upper triangular}}\hat{x} = Q^T \vec{b}$$

or

$$\hat{x} = R^{-1} Q^T \vec{b}$$

solve by back substitution

PROPERTIES OF DETERMINANTS

Here, we list the 4 most important properties

① If $U = \begin{pmatrix} u_{11} & u_{12} & \cdots & \cdots \\ & u_{22} & \cdots & \cdots \\ & & \ddots & \\ 0 & & & u_{nn} \end{pmatrix}$ is upper triangular, then

$$|U| = u_{11} u_{22} \cdots u_{nn}.$$

The determinant of an upper triangular matrix is the product of its diagonal elements

② If $L \in \mathbb{R}^{n \times n}$ is lower triangular, then

$$|L| = l_{11} l_{12} \cdots l_{nn}$$

$$L = \begin{pmatrix} l_{11} & & & 0 \\ l_{21} & l_{22} & & \\ \cdots & \cdots & \ddots & \\ \cdots & \cdots & \cdots & l_{nn} \end{pmatrix}$$

The determinant of a lower triangular matrix is the product of its diagonal elements.

PROPERTIES OF DETERMINANTS

③ Let $A, B \in \mathbb{R}^{n \times n}$. Then $|AB| = |A|\,|B|$.

④ Let $S_{i \leftrightarrow j}$ be the swap or row exchange matrix that exchanges rows i and j. Then

$$|S_{i \leftrightarrow j}| = \begin{cases} -1 & \text{if } i \neq j \\ 1 & \text{else} \end{cases}$$

Example: $S_{1 \leftrightarrow 2} \in \mathbb{R}^{3 \times 3}$ is $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ (identity matrix with rows 1 and 2 swapped)

$$|S_{1 \leftrightarrow 2}| = -1$$

⇝ these rules separately imply that $|\underbrace{I}_{\text{identity matrix}}| = 1$

Computing the Determinant

We know that every $A \in \mathbb{R}^{n\times n}$ has an LU decomposition after reordering its rows.

$$\underbrace{P}_{\substack{\text{permutation}\\ \text{matrix}}} A = \underbrace{L}_{\substack{\text{lower triangular}\\ \text{with 1's on its}\\ \text{diagonal}}} \underbrace{U}_{\text{upper triangular}}$$

with $P = S_{i_1 \leftrightarrow j_1} S_{i_2 \leftrightarrow j_2} \cdots S_{i_q \leftrightarrow j_q}$

Using the 4 rules above,

$$|PA| = |LU| \overset{\text{rule 3}}{\rightsquigarrow} |P||A| = |L||U|$$

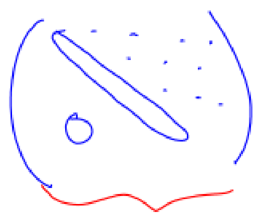
$$\Rightarrow (-1)^q |A| = 1\,|U| \quad \left(\begin{array}{l}\text{rule 4 - determinant of } S_{i\leftrightarrow j}\\ \text{rule 2 - det of } L\end{array}\right)$$

$$\Rightarrow |A| = (-1)^q \underbrace{|U|}$$

U is the upper triangular form obtained on elimination

Determinants and Invertibility

Suppose $A \in \mathbb{R}^{n\times n}$ has rank $r < n$. What is $|A|$?

$A \xrightarrow{\text{elimination}}$  $\Rightarrow \det A = |A| = 0\,!$

upper triangular matrix with at least one zero on the diagonal

$$\boxed{|A| \neq 0 \text{ if and only if } A \text{ is full rank if only if } A \text{ is invertible}}$$

Suppose $A \in \mathbb{R}^{n\times n}$ is invertible. Then $A^{-1}A = I \iff |A^{-1}A| = |I| = 1$

$$\iff |A^{-1}||A| = 1 \quad (\text{by rule 3})$$

Hence $\boxed{|A^{-1}| = \dfrac{1}{|A|}}$

DETERMINANT OF A^T

Let $A \in \mathbb{R}^{n\times n}$.

We have $A = \underbrace{\left(S_{i_1\leftrightarrow j_1} \; S_{i_2\leftrightarrow j_2} \cdots S_{i_q\leftrightarrow j_q} \right)}_{\substack{\text{swap matrices} \\ \text{(recall: these are symmetric)}}} L\, U$ (LU factorization)

Therefore $A^T = U^T L^T \left(S_{i_q\leftrightarrow j_q} \cdots S_{i_2\leftrightarrow j_2} \; S_{i_1\leftrightarrow j_1} \right)$ (using $(ABC)^T = C^T B^T A^T$ and $(S_{i\leftrightarrow j})^T = S_{i\leftrightarrow j}$)

$\Rightarrow \; |A^T| = |U^T| \cdot 1 \cdot (-1)^q$ (using rule 3 and structure of L (1's on diagonal))

$= |U| \, (-1)^q$

$$\boxed{|A^T| = |A|}$$

OTHER IMPORTANT PROPERTIES

* Let $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Consider rescaling a row by $c \in \mathbb{R}$

Then $\left| \begin{pmatrix} ca_{11} & ca_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| = \left| \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \right| = \left| \begin{pmatrix} c & 0 \\ 0 & 1 \end{pmatrix} A \right|$

$= c|A|$

In general, rescaling by a scalar 'c' rescales the determinant by c.

* adding one row to another

Let $c \in \mathbb{R}$

$$\left| \begin{pmatrix} a_{11} & a_{12} \\ a_{21}+ca_{11} & a_{22}+ca_{12} \end{pmatrix} \right| = \left| \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} A \right| = \left| \begin{pmatrix} 1 & 0 \\ c & 1 \end{pmatrix} \right| |A| = |A|$$

DETERMINANT BY COFACTORS

(Recall)

3x3 Determinant

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = + a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

$$= a_{11}\left(a_{22}a_{33} - a_{23}a_{32}\right) + a_{12}\left(a_{23}a_{31} - a_{21}a_{33}\right) + a_{13}\left(a_{21}a_{32} - a_{22}a_{31}\right)$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Cofactors

Defⁿ Let $A \in \mathbb{R}^{n\times n}$. Then Cofactor $C_{ij} = (-1)^{i+j} \det M_{ij}$

where $M_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ is a matrix obtained by deleting row i and column j from A

$$\boxed{\det A = \sum_{j=1}^{n} a_{ij} C_{ij} \quad \text{for any } A \in \mathbb{R}^{n\times n} \text{ for any row } i}$$

Moreover, since $\det A = \det A^T$

$$\boxed{\det A = \sum_{i=1}^{n} a_{ij} C_{ij} \quad \text{for any column } j}$$

CRAMER'S RULE

Def^n

Given $A \in \mathbb{R}^{n\times n}$, $A = \begin{pmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \\ | & | & & | \end{pmatrix}$, we solve $A\vec{x} = \vec{b}$ by defining

$$B_j = \begin{pmatrix} | & | & & | & & | \\ \vec{a}_1 & \vec{a}_2 & \cdots & \vec{b} & \cdots & \vec{a}_n \\ | & | & & | & & | \end{pmatrix}$$

j^{th} column replaced by $\vec{b}$

Then $$x_j = \frac{\det B_j}{\det A}$$

Note: to solve an $n\times n$ system, Cramer's rule evaluates $n+1$ determinants

INVERSES

Let $A \in \mathbb{R}^{n\times n}$ be invertible. Then $AA^{-1} = I \Rightarrow A\begin{pmatrix} | & | & & | \\ \vec{a}_1^{\,-1} & \vec{a}_2^{\,-1} & \cdots & \vec{a}_n^{\,-1} \\ | & | & & | \end{pmatrix} = I$

to solve for the inverse, we solve $A\vec{a}_1^{\,-1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $A\vec{a}_2^{\,-1} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ ......

Hence $$(A^{-1})_{11} = \frac{\begin{vmatrix} 1 & | & & | \\ 0 & \vec{a}_2 & \cdots & \vec{a}_n \\ \vdots & & & \\ 0 & | & & | \end{vmatrix}}{|A|} = \frac{C_{11}}{|A|}$$

first entry of $\vec{a}_1^{\,-1}$

Cofactor

(CRAMER'S RULE)

Similarly $$(A^{-1})_{21} = \frac{\begin{vmatrix} | & 1 & | & | & & | \\ \vec{a}_1 & 0 & \vec{a}_2 & \vec{a}_4 & \cdots & \vec{a}_n \\ & \vdots & & & & \\ | & 0 & | & | & & | \end{vmatrix}}{|A|} = \frac{C_{12}}{|A|}$$

(Recall) $n=3$

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

$|A|$

In general, $$\boxed{(A^{-1})_{ij} = \frac{C_{ji}}{\det A} \quad \text{and} \quad A^{-1} = \frac{C^T}{\det A}}$$