Orthonormal Bases & Gram-Schmidt

If A has columns that are orthogonal:

$$A^TA = \begin{bmatrix} — (\text{col } 1)^T — \\ — (\text{col } 2)^T — \\ \vdots \\ — (\text{col } n)^T — \end{bmatrix} \begin{bmatrix} | & | & & | \\ (\text{col } 1) & (\text{col } 2) & \cdots & (\text{col } n) \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \|\text{col } 1\|_2^2 & 0 & \cdots & 0 \\ 0 & \|\text{col } 2\|_2^2 & & 0 \\ \vdots & & \ddots & 0 \\ 0 & 0 & & \|\text{col } n\|_2^2 \end{bmatrix}$$

Diagonal matrix with diagonals equal to length of columns squared.

If, in addition to being orthogonal, columns have length 1, then:
$A^TA = I$. Meaning it would be very easy to calculate the least squares solution: $\hat{x} = (A^TA)^{-1}A^Tb$ to $Ax=b$.

Definition: Vectors $q_i, q_2, \ldots, q_n$ are "orthonormal" if

$$q_i^T q_j = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$$

$\leftarrow$ unit vectors ($i=j$)
$\leftarrow$ orthogonal ($i \neq j$)

Example: Permutation matrices. (columns are orthonormal)

$$P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad P^TP = I.$$

Example:

$$A = \begin{bmatrix} \sin x & \cos x \\ \cos x & -\sin x \end{bmatrix} \qquad A^TA = I$$

* A matrix Q with orthonormal columns satisfies
$Q^TQ = I$.
If Q is square, then $Q^T = Q^{-1}$.

Example:

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

(Rotation by θ degrees).

Example: (Reflection:)

Matrix $Q = I - 2uu^T$ where u is a unit vector.

$$Q^TQ = (I-2uu^T)^T(I-2uu^T) = I - 4uu^T + 4u\underbrace{u^Tu}_{1}u^T = I$$

Important: If Q has orthonormal columns:

$$\|Qx\|^2 = (Qx)^T(Qx) = x^TQ^TQx = x^Tx = \|x\|^2$$

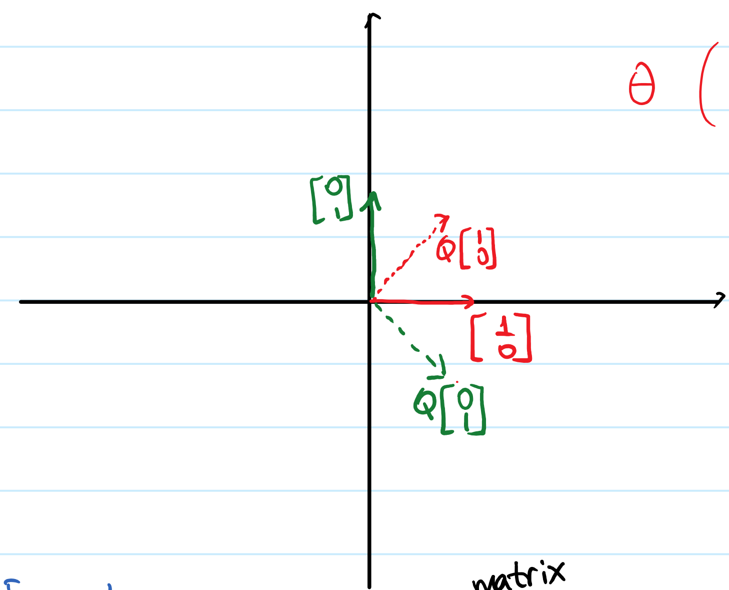
Transformation applied by Q doesn't change the length.

It preserves angles too! (Dot products).

$$(Qx)^T(Qy) = x^TQ^TQy = x^Ty.$$

Example:

$$Q = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad Q^TQ = I$$



$$\theta \left(\begin{array}{l} Q\begin{bmatrix}1\\0\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix} \\ Q\begin{bmatrix}0\\1\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix} \end{array} \right) \theta_2 \qquad \theta = \theta_2$$

$$\text{length}\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \text{length}\left(Q\begin{bmatrix}1\\0\end{bmatrix}\right)$$

Example:

If Q is a square matrix with orthonormal columns, then

$Qx = b$ has unique solution: $x = Q^{-1}b = Q^Tb$

Gram-Schmidt Process:

Create orthonormal vectors.

Independent vectors $a, b, c, \ldots$ $\xrightarrow{\text{Gram-Schmidt}}$ Orthonormal vectors $q_1, q_2, \ldots,$

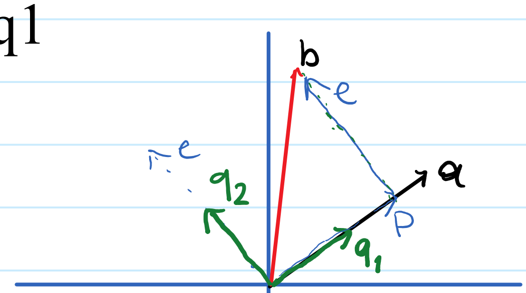
Q: How? A: Make use of projections.

1. Pick a vector a and "normalize" it: $q_1 = \dfrac{a}{\|a\|}$

2. Pick another vector b and project it on q_1, and subtract that from b (to find error vector between b & q_1)

$$p = \frac{q_1^T b}{q_1^T q_1} b \Rightarrow e = b - (q_1^T b)\, bq1$$

e is orthogonal to q_1



3. "Normalize" e: $q_2 = \dfrac{e}{\|e\|}$.

Now q_1 and q_2 are orthonormal vectors:

- $\|q_1\| = 1 \quad \|q_2\| = 1$

$$q_1^T q_2 = \frac{a^T}{\|a\|} \frac{\left(b - \frac{a^T b\, a}{a^T a}\right)}{\|e\|} = \frac{(a^T b - a^T b) \,|\, a^T a}{\|a\| \; \|e\|} = 0$$

4. Repeat for $c, d, \ldots$

(First find projections to each constructed q_i's and then subtract to find the error part orthogonal to q_i's)

For c:

$$e = c - (q_1^T c)\, q_1 - (q_2^T c)\, q_2, \text{ normalize: } q_3 = \frac{e}{\|e\|}.$$

Example:

$a = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \quad c = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ (exercise)

Closer look: (matrix form)

$$\underbrace{\begin{bmatrix} a & b & c & \cdots \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} q_1 & q_2 & q_3 & \cdots \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} q_1^T a & q_1^T b & q_1^T c \\ 0 & q_2^T b & q_2^T c \,\cdots \\ \vdots & 0 & q_3^T c \\ 0 & \vdots & 0 \end{bmatrix}}_{R}$$

$a = q_1(q_1^T a) \leftarrow a$ has non-zero projection on q_1 only.

$b = q_1(q_1^T b) + q_2(q_2^2 b)$

If A has LI columns, then we can decompose it into $A = QR$, where Q has orthonormal columns and R is rectangular.

Reason: Later q_i's are orthogonal to earlier columns of A.

Example: $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & -3 \\ 0 & -2 & 3 \end{bmatrix} = \underbrace{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ -1/\sqrt{2} & 1/\sqrt{6} & 1/\sqrt{3} \\ 0 & -2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{18} \\ & \sqrt{6} & -\sqrt{6} \\ & & \sqrt{3} \end{bmatrix}}_{R}$

Example: Least squares

$$(A^TA)\hat{x} = A^Tb \qquad A = \underset{m\times n}{Q}\,\underset{n\times n}{R}, \quad A: m\times n$$

$$(R^TQ^TQR)\hat{x} = R^TQ^Tb$$

$$(R^TR)\hat{x} = R^TQ^Tb$$

$$R\hat{x} = Q^Tb \quad \text{or} \quad \hat{x} = R^{-1}Q^Tb$$

Cost of $R\hat{x} = Q^Tb$: back-substitution.

Cost of Gram-Schmidth: mn^2 multiplications.

Determinants

Defined for square matrices, contains information about A.

"How much multiplication by A stretch or shrink the space?"

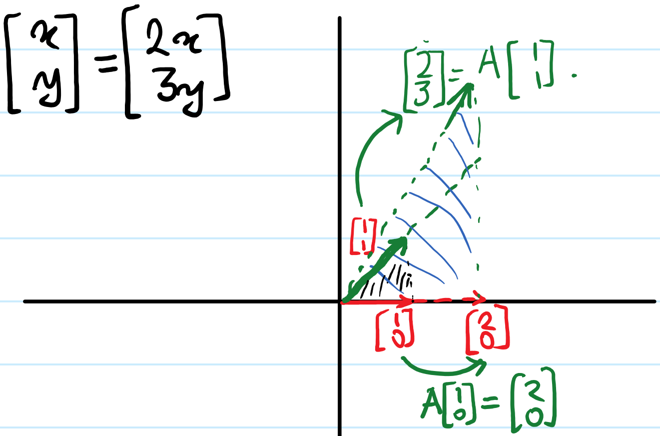
$$\underbrace{A}_{n\times n}\ \underbrace{x}_{n\times 1} = \underbrace{b}_{n\times 1}$$

We take a vector from R^n back to R^n.

What happens to its size? or two areas, volumes, etc.?

Example:

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



Area $1 = 1/2$

Area after transform: 3.

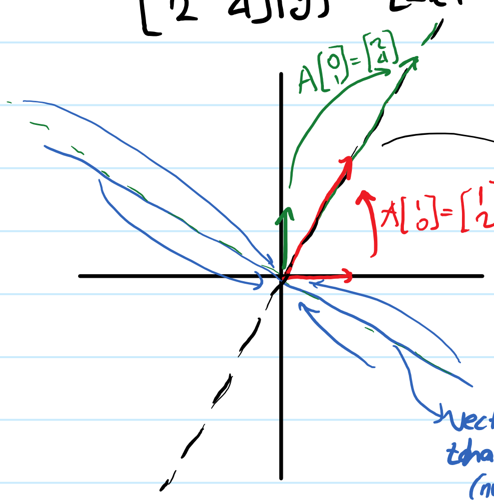
Got bigger by a factor of 6.

"Determinant" tells us how much we expanded or shrinked the space. Or whether we "flipped" the space.

Example: $\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+2y \\ 2x+4y \end{bmatrix}$

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

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



from R^2, multiplication by A reduced to a single line.

Definition:

$\det(A) = |A|$ is the determinant of A, and is a single number.

Example:

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = ad - bc$$

Notation: $\det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det A = |A|$

Three wads to calculate $\det(A)$:

① $\det(A)$ = multiplication of pivots (times 1 or -1).

$$\det(A) = \pm \det(U) = \pm(\text{product of the pivots})$$

Properties of $\det(A)$

① $\det(I) = 1$.

② determinant changes sign when rows are exchanged.

Ex: $\det(P) = \pm 1$ for permutation

③ $\det(A)$ is a linear function of each row separately. (all other rows stay fixed).

$$\det(A) = \begin{vmatrix} row1 \\ row2 \\ \vdots \\ row_i^1 + row_i^2 \\ \vdots \\ rowm \end{vmatrix} = \begin{vmatrix} row1 \\ row2 \\ \vdots \\ row_i^1 \\ \vdots \\ rowm \end{vmatrix} + \begin{vmatrix} row1 \\ row2 \\ \vdots \\ row_i^2 \\ \vdots \\ rowm \end{vmatrix}$$

(we wrote $row_i = row_i^1 + row_i^2$)

④ Equal rows: $\det(A) = 0$.

⑤ new row_i = row_i − (multiplier) row_J doesn't change the determinant. (row operations doesn't change det. Note: not row exchanges).

⑥ zero rows ⇒ $\det(A) = 0$.

⑦ Triangular ⇒ $\det(A)$ = product of diagonals.

⑧ Singular: $\det(A) = 0$. Invertible: $\det(A) \neq 0$.

9. $\det(AB) = \det(A)\det(B)$

$AA^{-1} = I \Rightarrow 1 = \det(A)\det(A^{-1})$.

10.

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

Same rules that apply to rows apply to columns as well.

Different methods to calculate determinant

Pivot formula:

$\det(A) = \pm\det(U) = \pm$ product of pivots.

Ex:

$$A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \quad PA = \begin{bmatrix} 4 & 5 & 6 \\ 0 & 2 & 3 \\ 0 & 0 & 1 \end{bmatrix} \qquad \det(A) = \underset{\text{"row exchange"}}{-}\,4 \cdot 2 \cdot 1 = -8$$

Big formula:

det A = sum of determinant of $n!$ matrices. (terms).

Ex:

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

$$= ad - bc.$$

Plus or minus of each term comes from # row exchanges needed.

det A = sum over all column permutations.

Cofactors:

$M_{ij} = A$ with row i and column j deleted.

Ex: $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ $M_{23} = \begin{bmatrix} 1 & 2 \\ 7 & 8 \end{bmatrix}$ cofactor $C_{ij} = (-1)^{i+j}|M_{ij}|$

Determinant formula with cofactors:

$$\det A = a_{i1}C_{i1} + a_{i2}C_{i2} + \dots + a_{in}C_{in} \quad \text{where}$$

$$C_{ij} = (-1)^{i+j} \det |M_{ij}|, \quad \text{with} \quad M_{ij} = \begin{pmatrix} A \text{ with row } i \text{ and} \\ \text{column } j \text{ deleted} \end{pmatrix}.$$

Example:

$$A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad \det(A) = 1\,(-1)^{2+1} \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} + 2\,(-1)^{3+1} \begin{vmatrix} 1 & 0 \\ 1 & 1 \end{vmatrix}$$

Useful when matrices have many zeros.

Cramer's Rule, Inverse & Volumes

Goal: Solve $Ax=b$ algebraically:

$$Ax=b \qquad \underbrace{\begin{bmatrix} & A & \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 & & 0 \\ x_2 & 1 & \\ x_n & 0 & 1 \end{bmatrix}}_{\substack{I \text{ with } x \\ \text{for first column}}} = \underbrace{\begin{bmatrix} b & a_{12} & \dots & a_{1n} \\ & & & \\ & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{\substack{A \text{ with } b \\ \text{for first column}}}$$

$$\Rightarrow \det(A) \underbrace{\det\begin{pmatrix} I \text{ wit } x \\ \text{for first column} \end{pmatrix}}_{=x_1} = \underbrace{\det\begin{pmatrix} A \text{ with } b \\ \text{for first column} \end{pmatrix}}_{\text{call it } B_1}$$

$$\Rightarrow \quad x_1 = \frac{\det B_1}{\det A}$$

Similarly:

$$\begin{bmatrix} & A & \end{bmatrix} \begin{bmatrix} 1 & x_1 & 0 & \dots \\ 0 & x_2 & 0 & \dots \\ \vdots & x_n & 1 & \\ & & \vdots & \end{bmatrix} = \begin{bmatrix} a_{11} & b_1 & a_{13} & \dots \\ a_{21} & b_2 & a_{23} & \\ \vdots & \vdots & \vdots & \end{bmatrix} \Rightarrow x_2 = \frac{\det B_2}{\det A}.$$

Cramer's Rule:

$Ax=b$ can be solved from determinants:

$$x_1 = \frac{\det B_1}{\det A}, \quad x_2 = \frac{\det B_2}{\det A}, \quad \dots, \quad x_n = \frac{\det B_n}{\det A}$$

B_i is matrix A with i^{th} column replaced by b.

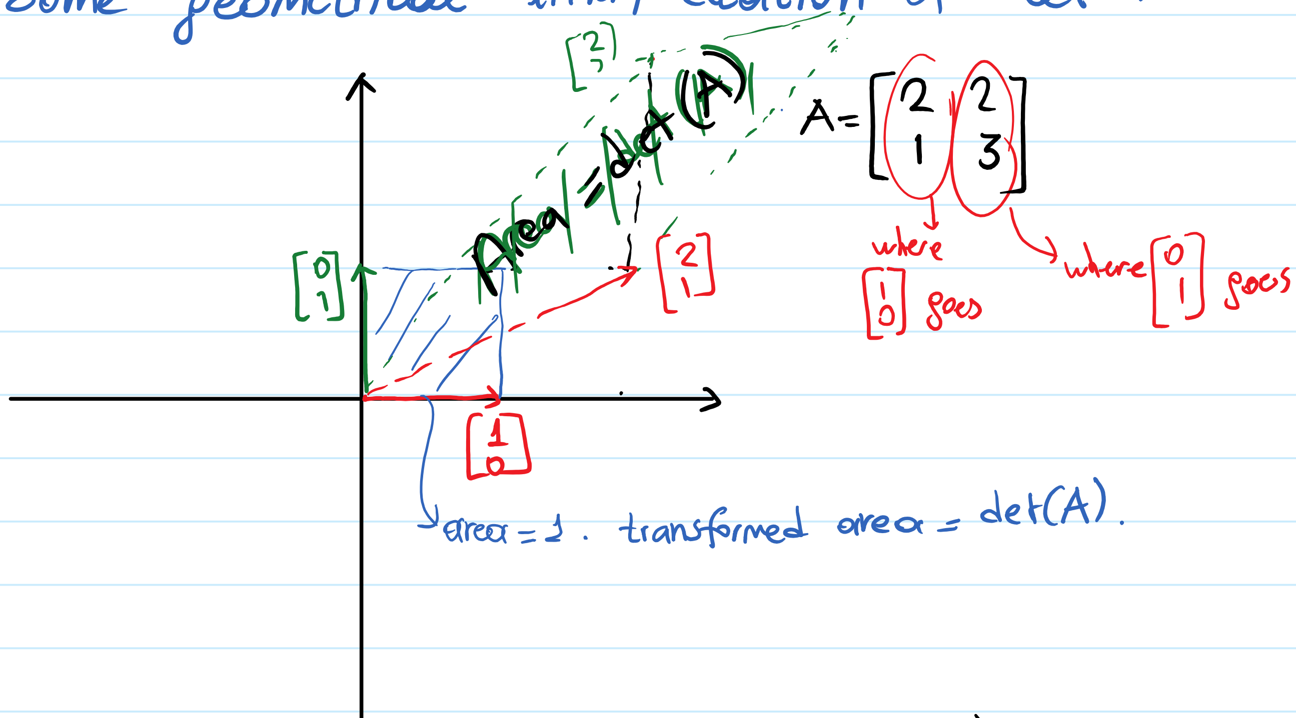
Inverse:

$$(A^{-1})_{iJ} = \frac{C_{Ji}}{\det A} \qquad A^{-1} = \frac{C^T}{\det A}$$

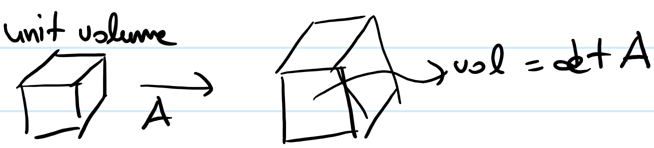
where

$$C_{iJ} = (-1)^{i+J} \det(M_{iJ})$$

Some geometrical interpretation of det A:



For 3D: similar (this times volumes).



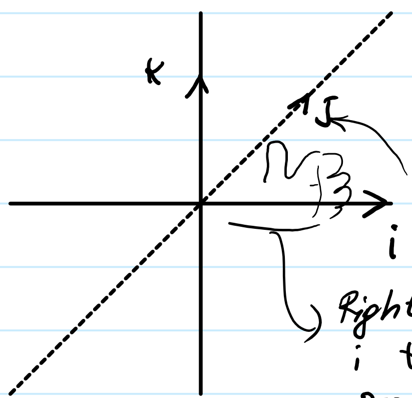
Cross product:

Assume $u = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$, $v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$

$$\underbrace{u \times v}_{\text{Cross product}} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} \times \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} u_2 v_3 - v_2 u_3 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{bmatrix} = \begin{vmatrix} i & j & k \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}$$

$u \times v$ is perpendicular to u and v.

$$\| u \times v \| = \|u\| \|v\| |\sin \theta|$$
$$| u \cdot v | = \|u\| \|v\| |\cos \theta|$$



$i \times j = k$

Right hand rule: Roll four fingers from i to j and thumb directs the cross product direction.

Volume of box with sides u, v, and w:

$$\text{volume} = (u \times v) \cdot w = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$