QR-Factorization

Gram-Schmidt is not only a useful algorithm; it tells something about the structure of $n \times n$ matrices.

Let $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ be independent vectors in $\mathbb{R}^n$. Then

$$A = \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

is an $n \times n$ invertible matrix. Applying Gram-Schmidt to $\vec{v}_1, \dots, \vec{v}_n$ gives an orthonormal basis $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n$ of $\mathbb{R}^n$, so

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

is an $n \times n$ orthogonal matrix. Consider the following:

$$\underbrace{\overset{Q}{\begin{bmatrix} | & & | \\ \vec{q}_1 & \cdots & \vec{q}_n \\ | & & | \end{bmatrix}}}_{n \times n} \underbrace{\overset{R}{\begin{bmatrix} \vec{q}_1 \cdot \vec{v}_1 & \vec{q}_1 \cdot \vec{v}_2 & \vec{q}_1 \cdot \vec{v}_3 & \cdots & \vec{q}_1 \cdot \vec{v}_n \\ 0 & \vec{q}_2 \cdot \vec{v}_2 & \vec{q}_2 \cdot \vec{v}_3 & & \vec{q}_2 \cdot \vec{v}_n \\ 0 & 0 & \vec{q}_3 \cdot \vec{v}_3 & & \vdots \\ \vdots & & 0 & \ddots & \\ 0 & \cdots & & & \vec{q}_n \cdot \vec{v}_n \end{bmatrix}}}_{n \times n} = \underbrace{\begin{bmatrix} | & & | \\ \vec{x}_1 & \cdots & \vec{x}_n \\ | & & | \end{bmatrix}}_{n \times n}$$

What are the vectors $\vec{x}_1, \dots, \vec{x}_n$?

$$\vec{x}_1 = \vec{q}_1(\vec{q}_1 \cdot \vec{v}_1) + \vec{q}_2(0) + \cdots + \vec{q}_n(0) = \mathrm{proj}_{\vec{q}_1}(\vec{v}_1) = \vec{v}_1.$$

$$\vec{x}_2 = \vec{q}_1(\vec{q}_1 \cdot \vec{v}_2) + \vec{q}_2(\vec{q}_2 \cdot \vec{v}_2) + \vec{q}_3(0) + \cdots + \vec{q}_n(0) = \mathrm{proj}_{\vec{q}_1}(v_2) + \mathrm{proj}_{\vec{q}_2}(v_2)$$
$$= v_2.$$

and so on: $\vec{x}_3 = \vec{v}_3, \ldots, \vec{x}_n = \vec{v}_n.$

So we obtain $\begin{bmatrix} | & & | \\ \vec{x}_1 & \cdots & \vec{x}_n \\ | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} = A.$

We have shown:

QR Factorization: Any invertible $n \times n$ matrix A can be written as

$$\boxed{A = QR}$$

where Q is orthogonal, R is upper triangular with all diagonal entries $\neq 0$.

Remarks

1) We can arrange that all diagonal entries of R are positive. In this case the factorization is unique!

2) More generally, any $n \times n$ matrix (not nec. invertible) can be written as QR where Q is orthogonal, R upper triangular (may have 0's on diag.)

Example

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \\ 0 & 1 & -1 \end{bmatrix}$$
$\vec{v}_1 \quad \vec{v}_2 \quad \vec{v}_3$

First apply Gram-Schmidt to $\vec{v}_1, \vec{v}_2, \vec{v}_3$.

$$\vec{w}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \qquad \|\vec{w}_1\| = \sqrt{2} \qquad \vec{q}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{w}_2 = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\vec{q}_1 = \vec{v}_2 - \left(\frac{1}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\|\vec{w}_2\| = \sqrt{(-1/2)^2 + (1/2)^2 + (1)^2} = \sqrt{3/2}$$

$$\vec{q}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} = \frac{\sqrt{2}}{\sqrt{3}} \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \end{bmatrix}$$

$$\vec{w}_3 = \vec{v}_3 - \underbrace{(\vec{q}_1 \cdot \vec{v}_3)}_{0}\vec{q}_1 - \underbrace{(\vec{q}_2 \cdot \vec{v}_3)}_{0}\vec{q}_2 = \vec{v}_3 = \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$\|\vec{w}_3\| = \sqrt{3}$$

$$\vec{q}_3 = \frac{\vec{w}_3}{\|\vec{w}_3\|} = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}. \quad \longrightarrow \quad Q = \begin{bmatrix} 1/\sqrt{2} & -\frac{\sqrt{2}}{2\sqrt{3}} & -1/\sqrt{3} \\ 1/\sqrt{2} & \frac{\sqrt{2}}{2\sqrt{3}} & 1/\sqrt{3} \\ 0 & \sqrt{2}/\sqrt{3} & -1/\sqrt{3} \end{bmatrix}$$

$A = QR$ here is:

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

Determinants

Recall: $\underset{2\times 2}{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ $\quad \det(A) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$

Significance: $\boxed{\det(A) \neq 0 \iff A \text{ invertible}}$

When $\det(A) \neq 0$, $\quad A^{-1} = \dfrac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

We want to generalize this story to $n\times n$ matrices.

Want a function "det": $\{n\times n \text{ matrices}\} \to \mathbb{R}$. Write

$$\det(\vec{v}_1, \dots, \vec{v}_n) = \det \begin{bmatrix} | & | & & | \\ \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \\ | & | & & | \end{bmatrix}$$

for vectors $\vec{v}_1, \dots, \vec{v}_n$ in $\mathbb{R}^n$.

Theorem There's a unique function $\det: \{n\times n \text{ matrices}\} \to \mathbb{R}$ satisfying:

1. $\det(I) = 1$ where I is the $n\times n$ identity matrix
2. $\det(\vec{v}_1, \text{—}, \vec{v}_{k-1}, a\vec{v} + b\vec{w}, \vec{v}_{k+1}, \text{—}, \vec{v}_n)$

   $= a\det(\vec{v}_1, \text{—}, \vec{v}_{k-1}, \vec{v}, \vec{v}_k, \text{—}, \vec{v}_n) + b\det(\vec{v}_1, \text{—}, \vec{v}_{k-1}, \vec{w}, \vec{v}_{k+1}, \text{—}, \vec{v}_n)$

   for any scalars a, b and any vectors.
3. $\det(A) = 0$ if any two columns of A are equal.

Examples

1) $$\begin{vmatrix} 1 & 0 & 1 \\ 2 & 2 & 0 \\ 1 & 0 & 1 \end{vmatrix} \overset{2}{=} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 1 \\ 2 & 2 & 0 \\ 0 & 0 & 1 \end{vmatrix} \overset{3}{=} 0 + 0 = 0$$

(We write $\det(A) =$ "$|A|$" for short.)

2) $$\begin{vmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{vmatrix} \overset{2}{=} 3 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{vmatrix} \overset{2}{=} 3 \cdot 2 \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \overset{1}{=} 3 \cdot 2 \cdot 1 = 6$$

3) Compute $\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}$.

$$0 \overset{3}{=} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix} \overset{2}{=} \begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix}$$

$$\overset{2}{=} \cancel{\begin{vmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{vmatrix}} + \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} + \cancel{\begin{vmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{vmatrix}}$$

$$\overset{3}{=} \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} + \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}$$

$$\overset{1}{=} 1 + \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix}$$

Thus $\begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -1$.

There are many formulas for determinants.

3×3 case:

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

$$- a_{31}a_{22}a_{13} - a_{32}a_{23}a_{11} - a_{33}a_{21}a_{12}$$

Example:

$$\begin{vmatrix} 2 & 1 & 4 \\ 2 & 3 & -1 \\ 1 & 1 & 1 \end{vmatrix} = (2)(3)(1) + (1)(-1)(1) + (4)(2)(1)$$

$$- (1)(3)(4) - (1)(-1)(2) - (1)(2)(1)$$

$$= 6 - 1 + 8 - 12 + 2 - 2 = 1$$

The above formula generalizes to $n \times n$ matrices:

("Big Formula")
$$\det_{n\times n}(A) = \sum_{\substack{\text{all permutations} \\ P}} \underbrace{\det(P)}_{=\pm 1} a_{1\alpha} a_{2\beta} \cdots a_{n\omega}$$

Here a permutation P is an $n \times n$ matrix that interchanges coordinates:

$$P \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_\alpha \\ x_\beta \\ \vdots \\ x_\omega \end{bmatrix}$$

Some important properties of determinants:

(i) $\det(AB) = \det(A)\det(B)$ where A, B are both $n \times n$

(ii) $\det\left(\begin{bmatrix} a_{11} & ? & ? & ? \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & ? \\ 0 & & & a_{nn} \end{bmatrix}\right) = a_{11} a_{12} \cdots a_{nn}$. Similarly for lower triangular.

upper triangular

Examples

1)

$$\underset{A}{\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}} = \underset{L}{\overset{\text{lower triang.}}{\begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix}}} \underset{U}{\overset{\text{upper triang.}}{\begin{bmatrix} 2 & -1 & 0 \\ 0 & \frac{3}{2} & -1 \\ 0 & 0 & \frac{4}{3} \end{bmatrix}}}$$

(i) gives $\det(A) = \det(L)\det(U)$.

(ii) gives $\det(L) = (1)(1)(1) = 1$, $\det(U) = (2)(\frac{3}{2})(\frac{4}{3}) = 4$.

So we get $\det(A) = (1)(4) = 4$.

2) $\begin{vmatrix} 3 & 2 & 500 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{vmatrix} = (3)(1)(2) = 6$. $\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = (2)(2) = 4$.

Note $\det(cA) \neq c\det(A)$ in general! We have:

$$\boxed{\det(c\underset{\substack{\uparrow \\ n \times n}}{A}) = c^n \det(A)}$$

Proof of (i) from the Theorem, for A invertible:

define $f : \{\begin{smallmatrix} n\times n \\ \text{matrices}\end{smallmatrix}\} \to \mathbb{R}$ by $f(B) = \det(AB)/\det(A)$,

for some fixed invertible A, $n \times n$.

Check f satisfies ①, ②, ③. Then $f = \det$ by "uniqueness" part of the theorem. So

$$f(B) = \frac{\det(AB)}{\det(A)} = \det(B)$$

which implies $\det(AB) = \det(A)\det(B)$.

Perhaps the most important property of det is:

$$\boxed{\underset{n\times n}{A} \text{ invertible} \iff \det(A) \neq 0.}$$

There are various ways to prove this.

Here's one way. Elimination gives a way to get from $\underset{n\times n}{A}$ to its RREF (call it R) via multiplying A by

$E_{ij,\ell}$ Elimination matrices

P_{ij} Permutation matrices

$D = D(a_1, \ldots, a_n)$ Diagonal matrices.

It's not hard to compute the determinants of these.

In particular, all have det $\neq 0$.

Then using multiplicativity of determinants:

$$\det(A) \neq 0 \iff \det(R) \neq 0$$

where R is the RREF of A.

Finally recall $R = I$ (identity) $\iff$ A invertible.

In this case, $\det(I) = 1$ so $\det(A) \neq 0$.

If A is not invertible then R has a column of zeros, and it easily follows that $\det(R) = \det(A) = 0$.