Chapter 5 Othogonality.

Recall: Vectors $\vec{v}, \vec{w}$ are said to be orthogonal with respect to an inner product if

$$\langle \vec{v}, \vec{w} \rangle = 0.$$

A basis $\{\vec{u}_1, \dots, \vec{u}_n\}$ of a vector space V is called orthogonal if, $\langle u_i, u_j \rangle = 0$ for all $i \neq j$. The basis is called orthonormal if in addition, each vector has unit length, $\|u_i\| = 1$.

Ex For $\mathbb{R}^n$

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

But this is not the only othonormal basis of $\mathbb{R}^n$!

Recall: For any non-zero vector V,

$$\vec{u} = \frac{\vec{V}}{\|\vec{V}\|} \text{ is a unit vector.}$$

Claim: If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthogonal basis of V then $\vec{u}_j = \dfrac{\vec{V}_j}{\|\vec{V}_j\|}$ $j = 1, \dots, n$ forms an orthonormal basis.

Proposition: If $\vec{v}_1, \dots, \vec{v}_k$ are mutually orthogonal $\to \langle \vec{v}_i, \vec{v}_j \rangle = 0$ for all i, j non-zero vectors then they are linearly independent.

Proof: Consider

$$\vec{v} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_k \vec{v}_k \quad \text{for } c_j \in \mathbb{R}.$$

Then

$$\begin{aligned}
\langle \vec{v}, \vec{v}_1 \rangle &= \langle c_1 \vec{v}_1 + c_2 \vec{v}_2 + \dots + c_k \vec{v}_k, \vec{v}_1 \rangle \\
&= c_1 \langle \vec{v}_1, \vec{v}_1 \rangle + c_2 \cancelto{0}{\langle \vec{v}_2, \vec{v}_1 \rangle} + \dots + c_k \cancelto{0}{\langle \vec{v}_k, \vec{v}_1 \rangle} \\
&= c_1 \underbrace{\|v_1\|^2}_{\text{non-zero}}
\end{aligned}$$

In general

$$\langle \vec{v}, \vec{v}_j \rangle = c_j \|v_j\|^2$$

Assume $\vec{v}_1, \dots, \vec{v}_k$ are linearly <u>dependent</u>. Then we can choose the constants some non-zero such that $\vec{v} = 0$.

Then $\langle \vec{v}, \vec{v}_j \rangle = \langle 0, \vec{v}_j \rangle = 0 = c_j \|v_j\|^2$

$\Rightarrow c_j = 0$, for all j.

Or. The only solution to

$$\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_k \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{pmatrix} = 0 \quad \text{is the}$$

zero solution! The vectors must be linearly independent.

Q: If $\{\vec{u}_1, \dots, \vec{u}_n\}$ is an orthonormal set in $\mathbb{R}^n$, what can be said about the matrix

$$A = (\vec{u}_1\ \vec{u}_2 \cdots \vec{u}_n)?$$

Why do we care about orthonormal bases? They are the generalization of the standard unit vectors $\{\vec{e}_j\}_{j=1}^n$. In some matrix computations using $\{\vec{e}_j\}_{j=1}^n$ is to difficult or too costly.

Orthonormal bases are customized to the inner product and for this reason the obey nice properties with respect to the corresponding norm.

<u>Theorem</u>: Let $\vec{u}_1, \dots, \vec{u}_n$ be an orthonormal basis for an inner product space V. Then one can write any element $\vec{v} \in V$ as a linear combination

$$\vec{v} = c_1\vec{u}_1 + \dots + c_n\vec{u}_n,$$

such that

$$c_i = \langle \vec{v}, \vec{u}_i \rangle, \quad i = 1, \dots, n.$$

Furthermore,

$$\|\vec{v}\| = \sqrt{c_1^2 + \dots + c_n^2} = \sqrt{\sum_{i=1}^{n} \langle \vec{v}, \vec{u}_i \rangle^2}.$$

<u>Proof</u>: Since $\{\vec{u}_1, \dots, \vec{u}_n\}$ is a basis for V then

$$\vec{v} = c_1\vec{u}_1 + \dots + c_n\vec{u}_n \quad \text{for some unique choice of } c_j,\ j=1,\dots,n.$$

We have already seen that

$$\langle \vec{v}, \vec{u}_j \rangle = c_j \|\vec{u}_j\|^2, \quad j=1,\dots,n$$

$$= c_j, \quad j=1,\dots,n \quad \text{since } \|\vec{u}_j\| = 1.$$

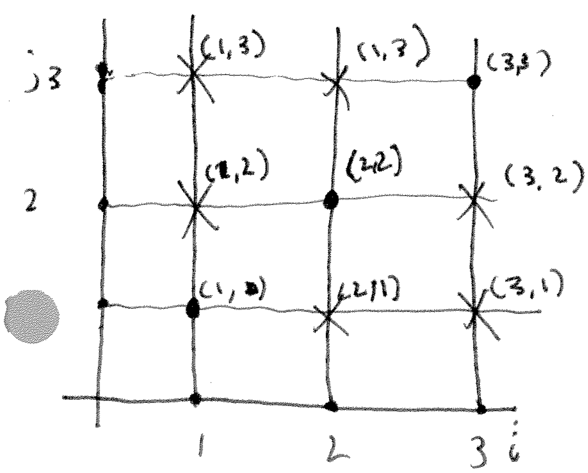
For the second statement consider

$$\|v\|^2 = \langle \vec{v}, \vec{v} \rangle = \left\langle \sum_{j=1}^{n} c_j\vec{u}_j, \sum_{i=1}^{n} c_i\vec{u}_i \right\rangle$$

$$= \sum_{j=1}^{n} \left\langle c_j\vec{u}_j, \sum_{i=1}^{n} c_i\vec{u}_i \right\rangle$$

$$= \sum_{j=1}^{n}\sum_{i=1}^{n} \langle c_j\vec{u}_j, c_i\vec{u}_i \rangle = \sum_{j=1}^{n}\sum_{i=1}^{n} c_j c_i \langle \vec{u}_j, \vec{u}_i \rangle$$

Non zero only when $i=j$



$$= \sum_{j=1}^{n} c_j^2 = \sum_{j=1}^{n} \langle \vec{v}, \vec{u}_j \rangle^2.$$

Ex Let $\{\vec{u}_1, \dots, \vec{u}_n\}$ be an orthonormal basis for $\mathbb{R}^n$.

and define

$$A = (\vec{u}_1 \ \vec{u}_2 \ \cdots \ \vec{u}_n)$$

when ever I don't specify, we use the dot product.

Write down the solution of

$$A\vec{x} = \vec{b}$$

for a general $\vec{b} \in \mathbb{R}^n$.

This is asking you to find the constants $x_1, \dots, x_n$ so that

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_n\vec{u}_n = \vec{b}$$

$$\Rightarrow \quad x_1 = \langle \vec{u}_1, \vec{b} \rangle, \ x_2 = \langle \vec{u}_2, \vec{b} \rangle, \ \dots$$

$$\Rightarrow \quad \vec{x} = \begin{pmatrix} \langle \vec{u}_1, \vec{b} \rangle \\ \langle \vec{u}_2, \vec{b} \rangle \\ \vdots \\ \langle \vec{u}_n, \vec{b} \rangle \end{pmatrix}.$$

__Note__ $\langle \vec{u}_i, \vec{b} \rangle$ requires n multiplications and $n-1$ additions.

we do this n times.

$$\Rightarrow \begin{array}{l} n^2 \text{ multiplications} \\ n^2 - n \text{ additions.} \end{array}$$

This is the same as matrix vector multiplication!

The Gram-Schmidt Process

- The fundament question here is how do we obtain an orthonormal basis from any other basis in an algorithmic way?

We start with a basis $\{\vec{w}_1, \ldots, \vec{w}_n\}$ and vector-by-vector fix it up so that we obtain a basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ that is orthonormal.
The process for doing this is called the Gram-Schmidt Process.

- In our first run through we won't worry about normality. since we know that it is easy to add normality to any basis by dividing by norms.

Step 1: Set $\vec{v}_1 = \vec{w}_1$

Step 2: Set $\vec{v}_2 = \vec{w}_2 - c\vec{v}_1$ ← fix up $\vec{w}_2$ with previously determined information

c ← to be determined.

We want $\vec{v}_1$ and $\vec{v}_2$ to be orthogonal:

$$0 = \langle \vec{v}_1, \vec{v}_2 \rangle = \langle \vec{v}_1, \vec{w}_2 \rangle - c\langle \vec{v}_1, \vec{v}_1 \rangle \Rightarrow c = \frac{\langle \vec{v}_1, \vec{w}_2 \rangle}{\|v_1\|^2}$$

Step 3:

$$\vec{v}_3 = \vec{w}_3 - c_1\vec{v}_1 - c_2\vec{v}_2$$

Orthogonality:

$$\langle \vec{v}_1, \vec{v}_3 \rangle = \langle \vec{v}_1, \vec{w}_3 \rangle - c_1 \langle \vec{v}_1, \vec{v}_1 \rangle - 0$$

$$\langle \vec{v}_2, \vec{v}_3 \rangle = \langle \vec{v}_2, \vec{w}_3 \rangle - 0 - c_2 \langle \vec{v}_2, \vec{v}_2 \rangle$$

$$\Rightarrow \quad c_1 = \frac{\langle \vec{v}_1, \vec{w}_3 \rangle}{\|v_1\|^2}, \quad c_2 = \frac{\langle \vec{v}_2, \vec{w}_3 \rangle}{\|v_2\|^2}$$

$$\vec{v}_3 = \vec{w}_3 - \left(\frac{\langle \vec{v}_1, \vec{w}_3 \rangle}{\|v_1\|^2} \right) \vec{v}_1 - \left(\frac{\langle \vec{v}_2, \vec{w}_3 \rangle}{\|v_2\|^2} \right) \vec{v}_2 .$$

Step 4:

$$\vec{v}_4 = \ \ldots$$

In general

$$\vec{v}_k = \vec{w}_k - c_1\vec{v}_1 - c_2\vec{v}_2 - \cdots - c_{k-1}\vec{v}_{k-1}$$

and inner products determine each c_j for $j = 1, \ldots, k-1$.

Ex The vectors

$$\vec{w}_1 = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}, \quad \vec{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \vec{w}_3 = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}.$$

form a basis of $\mathbb{R}^3$. Construct an orthogonal basis.

$\vec{v}_1 = \vec{w}_1$.

$$\vec{v}_2 = \vec{w}_2 - c\vec{v}_1, \qquad c = \frac{w_2 \cdot v_1}{\|v_1\|^2} = \frac{-1}{3}$$

$$= \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} - \left(-\frac{1}{3}\right)\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 4/3 \\ 1/3 \\ 5/3 \end{pmatrix}$$

$$\vec{v}_3 = \vec{w}_3 - c_1\vec{v}_1 - c_2 v_2 \qquad c_1 = \frac{w_3 \cdot v_1}{\|v_1\|^2} = -1$$

$$c_2 = \frac{w_3 \cdot v_2}{\|v_2\|^2} = 3/2$$

$$\vec{v}_3 = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix} - (-1)\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} - 3/2\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ -3/2 \\ -1/2 \end{pmatrix}.$$

To make them orthonormal we would then divide them by their lenght:

$$\vec{u}_1 = \frac{\vec{v}_1}{\|v_1\|} = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix}, \quad \vec{u}_2 = \frac{\vec{v}_2}{\|\vec{v}_2\|} = \begin{pmatrix} 4/\sqrt{42} \\ 1/\sqrt{42} \\ 5/\sqrt{42} \end{pmatrix}, \quad \vec{u}_3 = \frac{\vec{v}_3}{\|\vec{v}_3\|} = \begin{pmatrix} 2/\sqrt{14} \\ -3/\sqrt{14} \\ -1/\sqrt{14} \end{pmatrix}.$$

We have already discussed how every finite dimensional vector space admits a basis. The Gram-Schmidt Process allows one to construct an orthonormal basis if the underlying vector space is also an inner product space. This gives a beautiful constructive proof of the following theorem:

Theorem

Every non-zero finite-dimensional inner product space has an orthonormal basis.

As a final note we give the general Gram-Schmidt formula:

$$\boxed{\vec{V}_k = \vec{W}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{W}_k, \vec{V}_j \rangle}{\|\vec{V}_j\|^2} V_j \ , \ k = 1, \ldots, n.}$$

Modifications and Generalization of Gram-Schmidt

Consider $\{\vec{w}_1, \ldots, \vec{w}_n\}$ (some basis)

and $\{\vec{u}_1, \ldots, \vec{u}_n\}$ the orthonormal basis from Gram-Schmidt, $\vec{u}_j = \frac{v_j}{\|v_j\|}$.

We can express

$$\vec{w}_1 = r_{11}\vec{u}_1$$
$$\vec{w}_2 = r_{12}\vec{u}_1 + r_{22}\vec{u}_2$$
$$\vec{w}_3 = r_{13}\vec{u}_1 + r_{23}\vec{u}_2 + r_{33}\vec{u}_3$$
$$\vdots$$
$$\vec{w}_n = r_{1n}\vec{u}_1 + \cdots r_{nn}\vec{u}_n$$

It is straightforward to compute the r_{ij}'s once the other vectors are known..

$$\begin{cases} \langle \vec{w}_1, \vec{u}_1 \rangle = r_{11} \\ \langle \vec{w}_2, \vec{u}_1 \rangle = r_{12} \\ \langle \vec{w}_2, \vec{u}_2 \rangle = r_{22} \end{cases}$$
$$\vdots$$

$$r_{ij} = \langle w_j, u_i \rangle$$

We know that the u_j's are orthonormal...

the question is: how does this show up?

$$\|w_j\|^2 = \|r_{1j}\vec{u}_1 + \dots + r_{jj}\vec{u}_j\|^2$$
$$= r_{1j}^2 + \dots + r_{jj}^2$$

We wish to determine the r_{ij}'s as we determine the $\vec{u}_j$, this way we can go back and forth between the two bases.

$$\vec{w}_1 = r_{11}\vec{u}_1 \quad \leftarrow \text{stage } 1$$
$$\vec{w}_{j-1} = r_{1,j-1}\vec{u}_1 + \dots + r_{j-1,j-1}\vec{u}_{j-1} \quad \leftarrow \text{stage } j-1$$
$$\vec{w}_j = r_{1j}\vec{u}_1 + \dots + r_{jj}\vec{u}_j \quad \leftarrow \text{stage } j.$$

At stage j we know

$\vec{u}_1, \dots, \vec{u}_{j-1}$ but not $\vec{u}_j$.

We find

$$r_{1j} = \langle \vec{w}_j, \vec{u}_1 \rangle$$
$$r_{2j} = \langle \vec{w}_j, \vec{u}_2 \rangle$$
$$\vdots$$
$$r_{j-1,j} = \langle \vec{w}_j, \vec{u}_{j-1} \rangle$$

What is r_{jj}?

$$\|\vec{w}_j\|^2 = r_{1j}^2 + r_{2j}^2 + \dots + r_{j-1,j}^2 + r_{jj}^2$$

$$\Rightarrow \quad r_{jj} = \sqrt{\|\vec{w}_j\|^2 - r_{1j}^2 - \dots - r_{j-1,j}^2}.$$

$$u_j = \frac{\vec{w}_j - r_{1j}\vec{u}_1 - \dots - r_{j-1,j}\vec{u}_{j-1}}{r_{jj}}$$

But ... isn't the first version easier?

Yes, we know the r_{ij}'s, but this was a lot more work.

- Usefulness: Numerical computations? Not quite, the algorithm is unstable. Round off error will accumulate.

A numerically stable version of Gram-Schmidt:

$$\vec{u}_1 = \vec{w}_1 / \|\vec{w}_1\|$$

$$\vec{w}_k = \vec{w}_k - \langle \vec{w}_k, \vec{u}_1 \rangle \vec{u}_1 \qquad (\text{for } k = \underline{\underline{2}}, \dots, n)$$

$$\vec{u}_2 = \vec{w}_2 / \|w_2\|$$

$$\vec{w}_k = \vec{w}_k - \langle \vec{w}_k, \vec{u}_2 \rangle \vec{u}_2 \qquad (\text{for } k = \underline{\underline{3}}, \dots, n)$$

$$\vec{u}_3 = \vec{w}_3 / \|\vec{w}_3\|.$$

$$\vdots$$

In coding language:

Start with $\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_n\}$

$\vec{w}_1 = \vec{w}_1 / \|\vec{w}_1\|$ % redefine

$\vec{w}_k = \vec{w}_k - \langle \vec{w}_k, \vec{w}_1 \rangle \vec{w}_1$ for $k = 2, \dots, n$

$\vec{w}_2 = \vec{w}_2 / \|\vec{w}_2\|$

$\vec{w}_k = \vec{w}_k - \langle \vec{w}_k, \vec{w}_2 \rangle \vec{w}_2$ for $k = 3, \dots, n$.

We have seen ~~two~~ three ways of performing Gram-Schmidt

1) $\vec{v}_k = \vec{w}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_j \rangle}{\|v_j\|^2} v_j, \quad k = 1, \dots, n$

$\vec{u}_j = \vec{v}_j / \|\vec{v}_j\|$

2) $\vec{w}_1 = r_{11} \vec{u}_1$

$\vec{w}_2 = r_{12} \vec{u}_1 + r_{22} \vec{u}_2$

$\vdots$

$\vec{w}_n = r_{1n} \vec{u}_1 + \dots + r_{nn} \vec{u}_n$

$$\begin{cases} r_{ij} = \langle \vec{w}_j, \vec{u}_i \rangle, & i < j \\ r_{jj} = \sqrt{\|\vec{w}_j\|^2 - r_{1j}^2 - \dots - r_{j-1,j}^2} \end{cases}$$

3) $\vec{u}_1 = \vec{w}_1 / \|\vec{w}_1\|$

$\vec{w}_k = \vec{w}_k - \langle \vec{w}_k, \vec{u}_1 \rangle \vec{u}_1 \quad$ for $k = 2, \dots, n$

$\vec{u}_2 = \vec{w}_2 / \|\vec{w}_2\|$

$\vec{w}_k = \vec{w}_k - \langle \vec{w}_k, u_2 \rangle u_2 \quad$ for $k = 3, \dots, n$

$\vec{u}_3 = \vec{w}_3 / \|\vec{w}_3\|$

Orthogonal Matrices

A definition:

A square matrix is called orthogonal if it satisfies

$$Q^T Q = I.$$

Thus $\quad Q^T = Q^{-1}.$

Claim: A matrix Q is orthogonal if and only if its columns form an orthonormal basis with respect the Euclidean dot product on $\mathbb{R}^n$.

Proof: First, lets look at a 2×2 matrix

$$Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \vec{u}_1 = \begin{pmatrix} a \\ c \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} b \\ d \end{pmatrix}$$

$$Q^T Q = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^2 + c^2 & ab + cd \\ ab + cd & c^2 + d^2 \end{pmatrix}.$$

$$= \begin{pmatrix} \vec{u}_1 \cdot \vec{u}_1 & \vec{u}_1 \cdot \vec{u}_2 \\ \vec{u}_1 \cdot u_2 & \vec{u}_2 \cdot \vec{u}_2 \end{pmatrix}$$

If $\quad Q^T Q = I \quad$ we get $\left.\begin{aligned} \vec{u}_1 \cdot \vec{u}_1 &= 1 \\ \vec{u}_2 \cdot \vec{u}_2 &= 1 \\ \vec{u}_2 \cdot \vec{u}_1 &= 0 \end{aligned}\right\}$ othonormal basis for $\mathbb{R}^2$.

Facts:

- An othogonal matrix has $\det Q = \pm 1$

$$\det I = \det (Q^T \cdot Q) = \det Q^T \det Q$$
$$= (\det Q)^2$$

- The product of two othogonal matrices is also an orthogonal matrix.

Ex: $Q = I$ is orthogonal

$$Q = \begin{pmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{pmatrix}$$

The QR Factorization ("Gram-Schmidt w/ book keeping")

Let $\{\vec{w}_1, \dots, \vec{w}_n\}$ be a basis of $\mathbb{R}^n$ and let $\{\vec{u}_1, \dots, \vec{u}_n\}$ be the orthonormal basis that results from the Gram - Schmidt process.

Set

$$A = (\vec{w}_1, \dots, \vec{w}_n), \qquad Q = (\vec{u}_1, \dots, \vec{u}_n).$$

We want to know how A gets transformed into Q. Find R such that $A = QR$.

It turns out that R will be upper triangular. What are the elements of R?

We actually already looked at this:

We expressed:

$$\vec{w}_1 = r_{11}\vec{u}_1$$
$$\vec{w}_2 = r_{12}\vec{u}_1 + r_{22}\vec{u}_2$$
$$\vec{w}_3 = r_{13}\vec{u}_1 + r_{23}\vec{u}_2 + r_{33}\vec{u}_3$$
$$\vdots$$

And found r_{ij} in terms of u_i and w_j.

$$\boxed{r_{ij} = \langle w_j, u_i \rangle}$$

This gives us a way to find them but how would we actually do this in practice?

Ex

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

Find $A = QR$.

We will use

$$\vec{v}_k = \vec{w}_k - \sum_{j=1}^{k-1} \frac{\langle \vec{w}_k, \vec{v}_j \rangle}{\|\vec{v}_j\|^2} \vec{v}_j$$

OR, rearranging

$$\vec{w}_k = \vec{v}_k + \sum_{j=1}^{k-1} \underbrace{\frac{\langle \vec{w}_k, \vec{v}_j \rangle}{\|\vec{v}_j\|^2}}_{s_{jk}} \vec{v}_j .$$

We will find an intermediate factorization

$A = PS$ and then normalize

$$\begin{pmatrix} \vec{w}_1 & \vec{w}_2 & \cdots & \vec{w}_n \end{pmatrix} = \underbrace{\begin{pmatrix} \vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_n \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} 1 & s_{12} & s_{13} & \cdots \\ & 1 & s_{23} & \cdots \\ & & 1 & \\ & & & 1 \end{pmatrix}}_{S}$$

$\vec{w}_1 = \vec{v}_1$,

$$\begin{pmatrix} 2 & & \\ 1 & \ddots & \\ 0 & & \end{pmatrix} \begin{pmatrix} 1 & & \\ 0 & \ddots & \\ 0 & & \end{pmatrix}$$

$\vec{w}_1 = \vec{v}_1$

$$\begin{pmatrix} 2 & -3/5 & 10/35 \\ 1 & 6/5 & -20/35 \\ 0 & 1 & 30/35 \end{pmatrix}, \begin{pmatrix} 1 & \frac{4}{5} & 1/5 \\ 0 & 1 & 8/7 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\vec{v}_2 = \vec{w}_2 - \frac{\langle \vec{w}_2, \vec{v}_1 \rangle}{\|v_1\|^2} \vec{v}_1$$

$$= \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - \frac{4}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$$

$$= \begin{pmatrix} -3/5 \\ 6/5 \\ 1 \end{pmatrix}$$

$$\vec{v}_3 = \vec{w}_3 - \frac{\langle \vec{w}_3, \vec{v}_1 \rangle}{\|\vec{v}_1\|^2} \vec{v}_1 - \frac{\langle \vec{w}_3, \vec{v}_2 \rangle}{\|\vec{v}_2\|^2} \vec{v}_2$$

$$= \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{16/5}{9/25 + \frac{36}{25} + \frac{25}{25}} \begin{pmatrix} -3/5 \\ 6/5 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{16/5}{70/25} \begin{pmatrix} -3/5 \\ 6/5 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \left(\frac{1}{5}\right) \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{8}{7} \begin{pmatrix} -3/5 \\ 6/5 \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} -2/5 \\ -1/5 \\ 0 \end{pmatrix} + \begin{pmatrix} 24/35 \\ -48/35 \\ -40/35 \end{pmatrix}$$

$$= \begin{pmatrix} 10/35 \\ -20/35 \\ 30/35 \end{pmatrix}$$

Now, we have a factorization

$$A = PS.$$

But P is not orthogonal.

multiplies columns.

Compute

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1/c_1 & 0 \\ 0 & 1/c_2 \end{pmatrix} = \begin{pmatrix} a/c_1 & b/c_2 \\ c/c_1 & d/c_2 \end{pmatrix}$$

So
$$Q = P \underbrace{\begin{pmatrix} 1/\|\vec{v}_1\| & 0 & 0 \\ 0 & 1/\|\vec{v}_2\| & 0 \\ 0 & 0 & 1/\|\vec{v}_3\| \end{pmatrix}}_{= D}$$
is orthogonal!

Write
$$A = \underbrace{PD}_{Q}\,\underbrace{D^{-1}S}_{R}$$

mutiplies rows!

Compute

$$\begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} c_1 a & c_1 b \\ c_2 c & c_2 c \end{pmatrix}$$

Conclusion ?

Divide the colums of P by the norms $\Rightarrow Q$

Multiply the rows of S by the norms $\Rightarrow R$

$$\begin{pmatrix} 2 & -3/5 & 10/35 \\ 1 & 6/5 & -20/35 \\ 0 & 1 & 30/35 \end{pmatrix} \cdot \begin{pmatrix} 1 & 4/5 & 1/5 \\ 0 & 1 & 8/7 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\|\vec{v}_1\| = \sqrt{5} \qquad \|\vec{v}_2\| = \sqrt{\tfrac{70}{25}} = \sqrt{\tfrac{14}{5}} \qquad \|\vec{v}_3\| = 2\sqrt{\tfrac{2}{7}}$$

$$Q = \begin{pmatrix} \frac{2}{\sqrt{5}} & -3/\sqrt{70} & 1/\sqrt{14} \\ 1/\sqrt{5} & 6/\sqrt{70} & -\sqrt{2/7} \\ 0 & 5/\sqrt{70} & 3/\sqrt{14} \end{pmatrix}, \quad R = \begin{pmatrix} \sqrt{5} & 4/\sqrt{5} & 1/\sqrt{5} \\ 0 & \sqrt{\frac{14}{5}} & 8\sqrt{\frac{2}{35}} \\ 0 & 0 & 2\sqrt{\frac{2}{7}} \end{pmatrix}$$

Remark: There is no reason that we have to have 3 vectors in the previous problem:

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

then
$$A = \begin{pmatrix} 2 & -3/5 \\ 1 & 6/5 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 4/5 \\ 0 & 1 \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} 2/\sqrt{5} & -3/\sqrt{70} \\ 1/\sqrt{5} & 6/\sqrt{70} \\ 0 & 5/\sqrt{70} \end{pmatrix}}_{Q} \underbrace{\begin{pmatrix} \sqrt{5} & 4/\sqrt{5} \\ 0 & \sqrt{14}/6 \end{pmatrix}}_{R}$$

Note that $\underset{2\times 3}{Q^T}\ \underset{3\times 2}{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

An application of QR:

Consider finding a least-squares solution of

$$\underset{m\times n}{A}\ \underset{n\times 1}{\vec{x}} = \underset{m\times 1}{\vec{b}}.$$

Assuming A has rank n.

Assume we know

$$\underset{m\times n}{A} = \underset{m\times n}{Q}\ \underset{n\times n}{R}$$

We need to solve the normal equations.

$$A^T A \vec{x} = A^T \vec{b}$$

$$(QR)^T QR \vec{x} = A^T \vec{b}$$

$$\underset{n\times n}{R^T}\ \underbrace{\underset{n\times m}{Q^T}\ \underset{m\times n}{Q}}_{n\times n \text{ identity}}\ \underset{n\times n}{R}\ \vec{x} = A^T \vec{b}$$

$$\underset{\text{lower triangular}}{R^T}\ \underset{\text{upper triangular}}{R}\ \vec{x} = A^T \vec{b} \quad \leftarrow \text{ this can be solved by forward / back substitution.}$$

Moral of the story: QR factorization of A gives the LU factorization of $A^T A$.