Recall Given a subspace W of $\mathbb{R}^n$, we can write any $y \in \mathbb{R}^n$ as

$$y = \hat{y} + z$$

where $\hat{y} \in W$ and $z \in W^\perp$.

Here $\hat{y} = \mathrm{proj}_W y$ and $z = \mathrm{proj}_{W^\perp} y = y - \hat{y}$.

Given an orthogonal basis $\{u_1, \dots, u_k\}$ for W, we have

$$\hat{y} = \frac{y \cdot u_1}{u_1 \cdot u_1} u_1 + \cdots + \frac{y \cdot u_k}{u_k \cdot u_k} u_k.$$

If we have an orthonormal basis $\{u_1, \dots, u_k\}$ for W, it's even easier:

$$\hat{y} = \text{proj}_W y = (y \cdot u_1)u_1 + \cdots + (y \cdot u_k)u_k.$$

If $U = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_k \\ | & & | \end{bmatrix}$ is a matrix with orthonormal columns,

and $\text{Col}\, U = W$, then

$$UU^T y = \hat{y} = \text{proj}_W y$$

and $U^T U = I_{p \times p}$.

Today: Given a basis $\{x_1, \dots, x_k\}$ for a subspace W of $\mathbb{R}^n$, how can we produce an orthogonal basis?

The Gram-Schmidt process.

We'll construct a new basis $\{v_1, \dots, v_k\}$.

Idea: Take $v_1 = x_1$.

Next, choose v_2 so v_2 is orthogonal to v_1,

but $\operatorname{span}\{x_1, x_2\} = \operatorname{span}\{v_1, v_2\}$.

Then choose v_3 so v_3 is orthogonal to v_1 and v_2,

but $\operatorname{span}\{x_1, x_2, x_3\} = \operatorname{span}\{v_1, v_2, v_3\}$

... and continue until you've reached v_k.

Example

Let $W = \operatorname{span}\{x_1, x_2\}$, where $x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $x_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$.

Find an orthogonal basis for W.

We start with $v_1 = x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$.

We want $v_2 \in \operatorname{span}\{x_1, x_2\}$, and v_1 orthogonal to v_1.

We can find such a vector as

$$v_2 = x_2 - \operatorname{proj}_{x_1} x_2.$$

Let's calculate:

$$\operatorname{proj}_{x_1} x_2 = \frac{x_2 \cdot x_1}{x_1 \cdot x_1} x_1 = \frac{4}{2} x_1 = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}$$

$$v_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}.$$

Sure enough $v_1 \perp v_2$, and

$$\operatorname{span}\{v_1, v_2\} = \operatorname{span}\{x_1, x_2\},$$

so $\{v_1, v_2\}$ is an orthogonal basis for W.

Example

Suppose $\{x_1, x_2, x_3\}$ is a basis for a subspace W of $\mathbb{R}^4$.

Describe an orthogonal basis for W.

As before, we put

$$v_1 = x_1, \qquad v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 .$$

Now $\{v_1, v_2\}$ is an orthogonal basis for

$$W_2 = \text{span}\{x_1, x_2\} = \text{Span}\{v_1, v_2\}.$$

Then we'll take

$$v_3 = x_3 - \text{proj}_{W_2} x_3$$

$$= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 .$$

Clearly $v_3 \in W_2^{\perp}$, so v_3 is orthogonal to both v_1 and v_2.

From our formula, $\text{span}\{v_1, v_2, v_3\} = \text{span}\{x_1, x_2, x_3\} = W$.

So we've found an orthogonal basis for W.

Theorem The Gram-Schmidt process

Given a basis $\{x_1, \ldots, x_k\}$ for a subspace W of $\mathbb{R}^n$, define

$$v_1 = x_1$$

$$v_2 = x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$$v_3 = x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$$\vdots$$

$$v_k = x_k - \frac{x_k \cdot v_1}{v_1 \cdot v_1} v_1 - \cdots - \frac{x_k \cdot v_{k-1}}{v_{k-1} \cdot v_{k-1}} v_{k-1}.$$

Then $\{v_1, \ldots, v_k\}$ is an orthogonal basis for W.

(Also, $\operatorname{span}\{x_1, \ldots, x_\ell\} = \operatorname{span}\{v_1, \ldots, v_\ell\}$ for each $1 \leq \ell \leq k$.)

Example

Let $W = \text{span}\{x_1, x_2\}$,

$$x_1 = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}.$$

Find an orthogonal basis for W.

Take $v_1 = x_1$.

Define $v_2 = x_2 - \dfrac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$

$$= \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix} - \frac{-100}{50} \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix}.$$

If we wanted an orthonormal basis, we could normalise after we've done:

Define $u_1 = \dfrac{v_1}{\|v_1\|} = \dfrac{1}{\sqrt{50}} \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}$

$$u_2 = \frac{v_2}{\|v_2\|} = \frac{1}{3\sqrt{6}} \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Then $\{u_1, u_2\}$ is an <u>orthonormal</u> basis for W.

Example

Find an orthogonal basis for Col A, where

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

Let's write

$$A = \begin{bmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{bmatrix}.$$

Set $v_1 = a_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$.

Define $v_2 = a_2 - \dfrac{a_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \dfrac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \\ 0 \end{bmatrix}$.

Actually, for convenience, let's take

$v_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}$ instead. (You're allowed to multiply any of the v_i by a scalar, as long as you always use the same v_i!)

Then $v_3 = a_3 - \dfrac{a_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{a_3 \cdot v_2}{v_2 \cdot v_2} v_2$

$$= \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - 0 - \frac{2}{6}\begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{1}{3} \\ \frac{1}{3} \\ 1 \end{bmatrix}.$$

Again, let's take $v_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 3 \end{bmatrix}$ instead.

Next QR decomposition of matrices

Theorem If A is an $m \times n$ matrix with linearly independent columns, we can write

$$A = QR$$

where Q is an $m \times n$ matrix with orthonormal columns, and $\operatorname{Col} Q = \operatorname{Col} A$, and

R is an $n \times n$ upper triangular matrix.

You can find Q by applying Gram-Schmidt to the columns of A, normalising the basis, and using the resulting vectors as the columns of Q.

Since Q has orthonormal columns,

$$Q^T Q = I_{n \times n}$$

We can use this to find R:

$$R = IR = Q^T Q R = Q^T A.$$

The textbook explains why this R is upper triangular — but you can see it in the fact that in Gram-Schmidt, $v_i \in \operatorname{span}\{x_1, \ldots, x_i\}$, but does not depend on the later x_j's.

Example

Recall an orthogonal basis for Col A, where $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

is given by

$$\left\{ \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 1 \\ 3 \end{bmatrix} \right\}.$$

Find a QR decomposition of A.

To find Q, we need to normalise these vectors, and use these as the columns of Q.

$$Q = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{12}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{12}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{12}} \\ 0 & 0 & \frac{3}{\sqrt{12}} \end{bmatrix}.$$

Then

$$R = Q^T A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 0 \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{12}} & -\frac{1}{\sqrt{12}} & \frac{1}{\sqrt{12}} & \frac{3}{\sqrt{12}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{3}{\sqrt{6}} & \frac{2}{\sqrt{6}} \\ 0 & 0 & \frac{4}{\sqrt{12}} \end{bmatrix}$$

(which, happily, is upper triangular)

Matrix decompositions

We've seen several matrix decompositions:

- $A = PDP^{-1}$ (diagonalisation: D is diagonal, with entries the eigenvalues, P has as columns the eigenvectors)

- $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} \sqrt{a^2+b^2} & 0 \\ 0 & \sqrt{a^2+b^2} \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ where $r = \sqrt{a^2+b^2}$, $\theta = \cos^{-1}\frac{a}{r}$.

- $A = QR$.