Overview

Last time we discussed orthogonal projection. We'll review this today before discussing the question of how to find an orthonormal basis for a given subspace.

From Lay, §6.4

Orthogonal projection

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

$$\mathbf{y} = \hat{\mathbf{y}} + \mathbf{z} = \text{proj}_W \mathbf{y} + \text{proj}_{W^\perp} \mathbf{y},$$

where $\hat{\mathbf{y}} \in W$ is the closest vector in W to $\mathbf{y}$ and $\mathbf{z} \in W^\perp$. We call $\hat{\mathbf{y}}$ the *orthogonal projection of* $\mathbf{y}$ *onto* W.

Given an orthogonal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_p\}$ for W, we have a formula to compute $\hat{\mathbf{y}}$:

$$\hat{\mathbf{y}} = \frac{\mathbf{y}\cdot\mathbf{u}_1}{\mathbf{u}_1\cdot\mathbf{u}_1}\mathbf{u}_1 + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_p}{\mathbf{u}_p\cdot\mathbf{u}_p}\mathbf{u}_p.$$

If we also had an orthogonal basis $\{\mathbf{u}_{p+1}, \ldots, \mathbf{u}_n\}$ for $W^\perp$, we could find $\mathbf{z}$ by projecting $\mathbf{y}$ onto $W^\perp$:

$$\mathbf{z} = \frac{\mathbf{y}\cdot\mathbf{u}_{p+1}}{\mathbf{u}_{p+1}\cdot\mathbf{u}_{p+1}}\mathbf{u}_{p+1} + \cdots + \frac{\mathbf{y}\cdot\mathbf{u}_n}{\mathbf{u}_n\cdot\mathbf{u}_n}\mathbf{u}_n.$$

However, once we subtract off the projection of $\mathbf{y}$ to W, we're left with $\mathbf{z} \in W^\perp$. We'll make heavy use of this observation today.

Orthonormal bases

In the case where we have an orthonormal basis $\{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ for W, the computations are made even simpler:

$$\hat{\mathbf{y}} = (\mathbf{y}\cdot\mathbf{u}_1)\mathbf{u}_1 + (\mathbf{y}\cdot\mathbf{u}_2)\mathbf{u}_2 + \cdots + (\mathbf{y}\cdot\mathbf{u}_p)\mathbf{u}_p.$$

If $\mathcal{U} = \{\mathbf{u_1}, \ldots, \mathbf{u_p}\}$ is an orthonormal basis for W and U is the matrix whose columns are the $\mathbf{u_i}$, then

- $UU^T\mathbf{y} = \hat{\mathbf{y}}$
- $U^TU = I_p$

The Gram Schmidt Process

The aim of this section is to find an orthogonal basis $\{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ for a subspace W when we start with a basis $\{\mathbf{x_1}, \ldots, \mathbf{x_n}\}$ that is not orthogonal.

Start with $\mathbf{v_1} = \mathbf{x_1}$.
Now consider $\mathbf{x_2}$. If $\mathbf{v_1}$ and $\mathbf{x_2}$ are not orthogonal, we'll modify $\mathbf{x_2}$ so that we get an orthogonal pair $\mathbf{v_1}$, $\mathbf{v_2}$ satisfying

$$\mathrm{Span}\{\mathbf{x_1}, \mathbf{x_2}\} = \mathrm{Span}\{\mathbf{v_1}, \mathbf{v_2}\}.$$

Then we modify $\mathbf{x_3}$ so get $\mathbf{v_3}$ satisfying $\mathbf{v_1} \cdot \mathbf{v_3} = \mathbf{v_2} \cdot \mathbf{v_3} = 0$ and

$$\mathrm{Span}\{\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}\} = \mathrm{Span}\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}.$$

We continue this process until we've built a new orthogonal basis for W.

Example 1

Suppose that $W = \mathrm{Span}\ \{\mathbf{x}_1, \mathbf{x}_2\}$ where $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix}$. Find an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2\}$ for W.

To start the process we put $\mathbf{v}_1 = \mathbf{x}_1$.
We then find

$$\hat{\mathbf{y}} = \mathrm{proj}_{\mathbf{v}_1} \mathbf{x}_2 = \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 = \frac{4}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix}.$$

Now we define $\mathbf{v}_2 = \mathbf{x}_2 - \hat{\mathbf{y}}$; this is orthogonal to $\mathbf{x}_1 = \mathbf{v}_1$:

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \mathbf{v}_1 = \mathbf{x}_2 - \hat{\mathbf{y}} = \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix}.$$

So $\mathbf{v}_2$ is the component of $\mathbf{x}_2$ orthogonal to $\mathbf{x}_1$. Note that $\mathbf{v}_2$ is in $W = \mathrm{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$ because it is a linear combination of $\mathbf{v}_1 = \mathbf{x}_1$ and $\mathbf{x}_2$.
So we have that

$$\left\{ \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 3 \end{bmatrix} \right\}$$

is an orthogonal basis for W.

Example 2

Suppose that $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is a basis for a subspace W of $\mathbb{R}^4$. Describe an orthogonal basis for W.

- As in the previous example, we put

$$\mathbf{v}_1 = \mathbf{x}_1 \quad \text{and} \quad \mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1.$$

Then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for $W_2 = \text{Span}\,\{\mathbf{x}_1, \mathbf{x}_2\} = \text{Span}\,\{\mathbf{v}_1, \mathbf{v}_2\}$.

- Now $\text{proj}_{W_2}\mathbf{x}_3 = \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 + \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$ and

$$\mathbf{v}_3 = \mathbf{x}_3 - \text{proj}_{W_2}\mathbf{x}_3 = \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2$$

is the component of $\mathbf{x}_3$ orthogonal to W_2. Furthermore, $\mathbf{v}_3$ is in W because it is a linear combination of vectors in W.

- Thus we obtain that $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ is an orthogonal basis for W.

Theorem (The Gram Schmidt Process)

Given a basis $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p\}$ for a subspace W of $\mathbb{R}^n$, define

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{x}_1 \\
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
\mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\
&\vdots \\
\mathbf{v}_p &= \mathbf{x}_p - \frac{\mathbf{x}_p \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \ldots - \frac{\mathbf{x}_p \cdot \mathbf{v}_{p-1}}{\mathbf{v}_{p-1} \cdot \mathbf{v}_{p-1}}\mathbf{v}_{p-1}
\end{aligned}$$

Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is an orthogonal basis for W. Also

$$\text{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k\} = \text{Span}\,\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \quad \text{for } 1 \leq k \leq p.$$

Example 3

The vectors

$$\mathbf{x}_1 = \begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix}$$

form a basis for a subspace W. Use the Gram-Schmidt process to produce an orthogonal basis for W.

Step 1 Put $\mathbf{v}_1 = \mathbf{x}_1$.
Step 2

$$\begin{aligned}
\mathbf{v}_2 &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \\
&= \begin{bmatrix} -3 \\ 14 \\ -7 \end{bmatrix} - \frac{(-100)}{50}\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix}.
\end{aligned}$$

Then $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for W.

To construct an orthonormal basis for W we normalise the basis $\{\mathbf{v}_1, \mathbf{v}_2\}$:

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{50}}\begin{bmatrix} 3 \\ -4 \\ 5 \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2\|}\mathbf{v}_2 = \frac{1}{\sqrt{54}}\begin{bmatrix} 3 \\ 6 \\ 3 \end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Then $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal basis for W.

Example 4

Let $A = \begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & 3 \end{bmatrix}$. Use the Gram-Schmidt process to find an orthogonal basis for the column space of A.

Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ be the three columns of A.

Step 1 Put $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}$.

Step 2

$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} 6 \\ -8 \\ -2 \\ -4 \end{bmatrix} - \frac{(-36)}{12}\begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}.$$

Step 3

$$\begin{aligned} \mathbf{v}_3 &= \mathbf{x}_3 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{x}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 \\ &= \begin{bmatrix} 6 \\ 3 \\ 6 \\ 3 \end{bmatrix} - \frac{12}{12}\begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix} - \frac{24}{12}\begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix}. \end{aligned}$$

Thus an orthogonal basis for the column space of A is given by

$$\left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix} \right\}.$$

Example 5

The matrix A is given by

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

Use the Gram-Schmidt process to show that

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

is an orthogonal basis for Col A.

Let $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ be the three columns of A.

Step 1 Put $\mathbf{v}_1 = \mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$.

Step 2

$$\mathbf{v}_2 = \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 1/2 \\ 1 \\ 0 \end{bmatrix}.$$

For convenience we take $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix}$. (This is optional, but it makes $\mathbf{v}_2$ easier to work with in the following calculation.)

Step 3

$$\mathbf{v}_3 = \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2}\mathbf{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - 0 - \frac{2}{6}\begin{bmatrix} -1 \\ 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -1/3 \\ 1/3 \\ 1 \end{bmatrix}$$

For convenience we take $\mathbf{v}_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 3 \end{bmatrix}$.

QR factorisation of matrices

If an $m \times n$ matrix A has linearly independent columns $\mathbf{x}_1, \dots, \mathbf{x}_n$, then $A = QR$ for matrices

- Q is an $m \times n$ matrix whose columns are an orthonormal basis for $\mathrm{Col}(A)$, and
- R is an $n \times n$ upper triangular invertible matrix.

This factorisation is used in computer algorithms for various computations.

In fact, finding such a Q and R amounts to applying the Gram Schmidt process to the columns of A.

(The proof that such a decomposition exists is given in the text.)

Example 6

Let

$$A = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix}, \quad Q = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix}$$

where the columns of Q are obtained by applying the Gram-Schmidt process to the columns of A and then normalising the columns.
Find R such that $A = QR$.

As we have noted before, $Q^T Q = I$ because the columns of Q are orthonormal. If we believe such an R exists, we have

$$Q^T A = Q^T(QR) = (Q^T Q)R = IR = R.$$

Therefore $R = Q^T A$.

In this case,

$$\begin{aligned} R &= Q^T A \\ &= \begin{bmatrix} 5/6 & 1/6 & -3/6 & 1/6 \\ -1/6 & 5/6 & 1/6 & 3/6 \end{bmatrix} \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} \\ &= \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix} \end{aligned}$$

An easy check shows that

$$QR = \begin{bmatrix} 5/6 & -1/6 \\ 1/6 & 5/6 \\ -3/6 & 1/6 \\ 1/6 & 3/6 \end{bmatrix} \begin{bmatrix} 6 & 12 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 9 \\ 1 & 7 \\ -3 & -5 \\ 1 & 5 \end{bmatrix} = A.$$

Example 7

In Example 4 we found that an orthogonal basis for the column space of the matrix

$$A = \begin{bmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & 3 \end{bmatrix}$$

is given by

$$\left\{ \begin{bmatrix} -1 \\ 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \\ 3 \\ 4 \end{bmatrix} \right\}$$

Normalising the columns gives

$$Q = \begin{bmatrix} -1/\sqrt{12} & 3/\sqrt{12} & 1/\sqrt{30} \\ 3/\sqrt{12} & 1/\sqrt{12} & -2/\sqrt{30} \\ 1/\sqrt{12} & 1/\sqrt{12} & 3/\sqrt{30} \\ 1/\sqrt{12} & -1/\sqrt{12} & 4\sqrt{30} \end{bmatrix}.$$

As in the last example

$$\begin{aligned} R &= Q^T A \\ &= \begin{bmatrix} \sqrt{12} & \sqrt{12} & \sqrt{12} \\ 0 & \sqrt{12} & 2\sqrt{12} \\ 0 & 0 & \sqrt{30} \end{bmatrix}. \end{aligned}$$

It is left as an exercise to check that $QR = A$.

Matrix decompositions

We've seen a variety of matrix decompositions this semester:

- $A = PDP^{-1}$
- $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = S_t R_\theta$
- $A = QR$

In each case, we go to some amount of computation work in order to express the given matrix as a product of terms we understand well. The advantages of this can be either conceptual or computational, depending on the context.

Example 8

An orthogonal basis for the column space of the matrix

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

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 construct Q we normalise the orthogonal vectors. These become the columns of Q:

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

Since $R = Q^T A$, we solve

$$\begin{aligned} R = Q^T A &= \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ -1/\sqrt{6} & 1/\sqrt{6} & 2/\sqrt{6} & 0 \\ 1/\sqrt{12} & -1/\sqrt{12} & 1/\sqrt{12} & 3/\sqrt{12} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 3/\sqrt{6} & 2/\sqrt{6} \\ 0 & 0 & 4/\sqrt{12} \end{bmatrix} \end{aligned}$$

Check:

$$\begin{aligned} QR &= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{12} \\ 1/\sqrt{2} & 1/\sqrt{6} & -1/\sqrt{12} \\ 0 & 2/\sqrt{6} & 1/\sqrt{12} \\ 0 & 0 & 3/\sqrt{12} \end{bmatrix} \begin{bmatrix} 2/\sqrt{2} & 1/\sqrt{2} & 0 \\ 0 & 3/\sqrt{6} & 2/\sqrt{6} \\ 0 & 0 & 4/\sqrt{12} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$