Sec 6.4 The Gram-Schmidt Process — How to turn any basis into an orthogonal basis

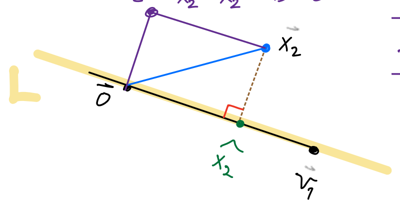
Recall Def (Sec 6.2): given vectors $\vec{x_2}$ and $\vec{v_1}$ in $\mathbb{R}^n$,

the <u>orthogonal projection of $\vec{x_2}$ onto $\vec{v_1}$</u> or <u>orthogonal projection of $\vec{x_2}$ onto W_1</u>,

where $W_1 = \text{Span}\{\vec{v_1}\}$ is the "line" containing $\vec{0}$ and $\vec{v_1}$, is equal to

$$\hat{x}_2 = \text{proj}_{W_1} x_2 = \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$$

$\vec{z} = \vec{x_2} - \hat{x}_2$ is called <u>the component of $\vec{x_2}$ orthogonal to W_1</u> or <u>the component of $\vec{x_2}$ orthogonal to $\vec{v_1}$</u>



Recall Def (Sec 6.3) "The orthogonal decomposition theorem"

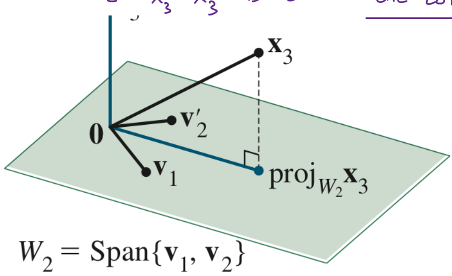
Given linearly independent vectors $\vec{x_3}$, $\vec{v_1}, \vec{v_2}$ in $\mathbb{R}^n$,

the <u>orthogonal projection of $\vec{x_3}$ onto W_2</u>, where $W_2 = \text{Span}\{\vec{v_1}, \vec{v_2}\}$

is the "plane" containing $\vec{0}, \vec{v_1}, \vec{v_2}$, is equal to

$$\hat{x}_3 = \text{proj}_{W_2} x_3 = \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$$

$\vec{z} = \vec{x_3} - \hat{x}_3$ is called <u>the component of $\vec{x_3}$ orthogonal to W_2</u>



(We use the orthogonal decomposition thm to give us an algorithm for producing an orthogonal basis from any basis, called the Gram-Schmidt process)

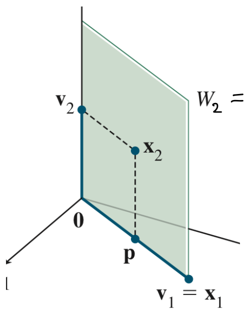
I. The Gram-Schmidt process (Thm 11)

Let $\mathcal{X} = \{x_1, x_2, x_3, \dots, x_p\}$, $p \geq 1$, be a basis for a subspace W of $\mathbb{R}^n$.

Step 1: Let $v_1 := x_1$

Step 2: Let $v_2 = x_2 - \dfrac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1$ the component of $\vec{x}_2$ orthogonal to $W_1 = \text{Span}\{\vec{v}_1\}$

projection of x_2 onto v_1



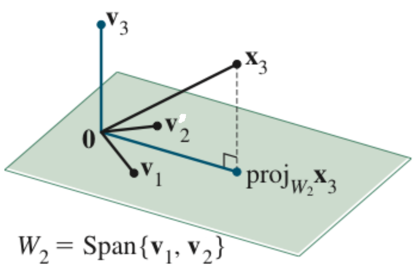
$W_2 = \text{Span}\{v_1, v_2\}$

projection of x_3 onto v_1

projection of x_3 onto v_2

Step 3: Let $v_3 = x_3 - \dfrac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2$

the component of $\vec{x}_3$ orthogonal to $W_2 = \text{Span}\{\vec{v}_1, \vec{v}_2\}$



$\vdots$

Step p: Let $v_p = x_p - \dfrac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \dfrac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \dots - \dfrac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}$.

Then $\{v_1, \dots, v_p\}$ is an orthogonal basis for W.

Note: If we want to turn an orthogonal basis $\{v_1, \dots, v_p\}$ into an orthonormal basis, scale each v_k by $\dfrac{1}{\|v_k\|}$.

EXAMPLE 2 Let $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$. Then $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ is clearly linearly independent and thus is a basis for a subspace W of $\mathbb{R}^4$. Construct an orthogonal basis for W.

SOLUTION

Step 1. Let $\mathbf{v}_1 = \mathbf{x}_1$ and $W_1 = \text{Span}\{\mathbf{x}_1\} = \text{Span}\{\mathbf{v}_1\}$.

Step 2. Let $\mathbf{v}_2$ be the vector produced by subtracting from $\mathbf{x}_2$ its projection onto the subspace W_1. That is, let

$$
\begin{aligned}
\mathbf{v}_2 &= \mathbf{x}_2 - \text{proj}_{W_1} \mathbf{x}_2 \\
&= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Since } \mathbf{v}_1 = \mathbf{x}_1 \\
&= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}
\end{aligned}
$$

As in Example 1, $\mathbf{v}_2$ is the component of $\mathbf{x}_2$ orthogonal to $\mathbf{x}_1$, and $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for the subspace W_2 spanned by $\mathbf{x}_1$ and $\mathbf{x}_2$.

Step 2′ (optional). If appropriate, scale $\mathbf{v}_2$ to simplify later computations. Since $\mathbf{v}_2$ has fractional entries, it is convenient to scale it by a factor of 4 and replace $\{\mathbf{v}_1, \mathbf{v}_2\}$ by the orthogonal basis

$$
\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}
$$

Step 3. Let $\mathbf{v}_3$ be the vector produced by subtracting from $\mathbf{x}_3$ its projection onto the subspace W_2. Use the orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2'\}$ to compute this projection onto W_2:

$$
\text{proj}_{W_2} \mathbf{x}_3 = \underbrace{\frac{\mathbf{x}_3 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1}_{\substack{\text{Projection of} \\ \mathbf{x}_3 \text{ onto } \mathbf{v}_1}} + \underbrace{\frac{\mathbf{x}_3 \cdot \mathbf{v}_2'}{\mathbf{v}_2' \cdot \mathbf{v}_2'}\mathbf{v}_2'}_{\substack{\text{Projection of} \\ \mathbf{x}_3 \text{ onto } \mathbf{v}_2'}} = \frac{2}{4}\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \frac{2}{12}\begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}
$$

Then $\mathbf{v}_3$ is the component of $\mathbf{x}_3$ orthogonal to W_2, namely

$$
\mathbf{v}_3 = \mathbf{x}_3 - \text{proj}_{W_2} \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}
$$

We can also use $\begin{bmatrix} 0 \\ -2 \\ 1 \\ 1 \end{bmatrix}$

See Figure 2 for a diagram of this construction. Observe that $\mathbf{v}_3$ is in W, because $\mathbf{x}_3$ and $\text{proj}_{W_2}\mathbf{x}_3$ are both in W. Thus $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ is an orthogonal set of nonzero vectors and hence a linearly independent set in W. Note that W is three-dimensional since it was defined by a basis of three vectors. Hence, by the Basis Theorem in Section 4.5, $\{\mathbf{v}_1, \mathbf{v}_2', \mathbf{v}_3\}$ is an orthogonal basis for W. ■

Idea: Applying the Gram-Schmidt process (+ scaling) to columns of A corresponds to factoring $A = QR$

II. The QR Factorization of matrices (Thm 12)

Let A be an $m \times \overset{n \text{ columns}}{n}$ matrix w/ linearly independent columns. Then A can be factored as $A = QR$, where

- Q is an $m\times n$ matrix whose columns form an orthonormal basis for Col A
- R is an $n \times n$ upper triangular matrix w/ positive entries on its diagonal

(note: R is invertible since its rank is n)

EXAMPLE 4 Find a QR factorization of $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

SOLUTION The columns of A are the vectors $\mathbf{x}_1$, $\mathbf{x}_2$, and $\mathbf{x}_3$ in Example 2 (above). An orthogonal basis for Col $A = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$ was found in that example:

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_2' = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

To simplify the arithmetic that follows, scale $\mathbf{v}_3$ by letting $\mathbf{v}_3' = 3\mathbf{v}_3$. Then normalize the three vectors to obtain $\mathbf{u}_1$, $\mathbf{u}_2$, and $\mathbf{u}_3$, and use these vectors as the columns of Q:

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

By construction, the first k columns of Q are an orthonormal basis of Span $\{\mathbf{x}_1, \dots, \mathbf{x}_k\}$. From the proof of Theorem 12, $A = QR$ for some R. To find R, observe that $\underbrace{Q^TQ = I}_{\text{Thm 6 Sec 6.2 pg 363}}$, because the columns of Q are orthonormal. Hence

$$Q^TA = Q^T(QR) = IR = R$$

and

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

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

Thm 6

An $m \times n$ matrix U has orthonormal columns if and only if $U^TU = I$.

Practice Problems

1. Let $W = \text{Span}\{\mathbf{x}_1, \mathbf{x}_2\}$, where $\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{x}_2 = \begin{bmatrix} 1/3 \\ 1/3 \\ -2/3 \end{bmatrix}$. Construct an orthonormal basis for W.
2. Suppose $A = QR$, where Q is an $m \times n$ matrix with orthogonal columns and R is an $n \times n$ matrix. Show that if the columns of A are linearly dependent, then R cannot be invertible.

Solution to Practice Problems

1. Let $\mathbf{v}_1 = \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ and $\mathbf{v}_2 = \mathbf{x}_2 - \dfrac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 = \mathbf{x}_2 - 0\mathbf{v}_1 = \mathbf{x}_2$. So $\{\mathbf{x}_1, \mathbf{x}_2\}$ is already orthogonal. All that is needed is to normalize the vectors. Let

$$\mathbf{u}_1 = \frac{1}{\|\mathbf{v}_1\|}\mathbf{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

Instead of normalizing $\mathbf{v}_2$ directly, normalize $\mathbf{v}_2' = 3\mathbf{v}_2$ instead:

$$\mathbf{u}_2 = \frac{1}{\|\mathbf{v}_2'\|}\mathbf{v}_2' = \frac{1}{\sqrt{1^2 + 1^2 + (-2)^2}}\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

Then $\{\mathbf{u}_1, \mathbf{u}_2\}$ is an orthonormal basis for W.
2. Since the columns of A are linearly dependent, there is a nontrivial vector $\mathbf{x}$ such that $A\mathbf{x} = \mathbf{0}$. But then $QR\mathbf{x} = \mathbf{0}$. Applying Theorem 7 from Section 6.2 results in $\|R\mathbf{x}\| = \|QR\mathbf{x}\| = \|\mathbf{0}\| = 0$. But $\|R\mathbf{x}\| = 0$ implies $R\mathbf{x} = \mathbf{0}$, by Theorem 1 from Section 6.1. Thus there is a nontrivial vector $\mathbf{x}$ such that $R\mathbf{x} = \mathbf{0}$ and hence, by the Invertible Matrix Theorem, R cannot be invertible.