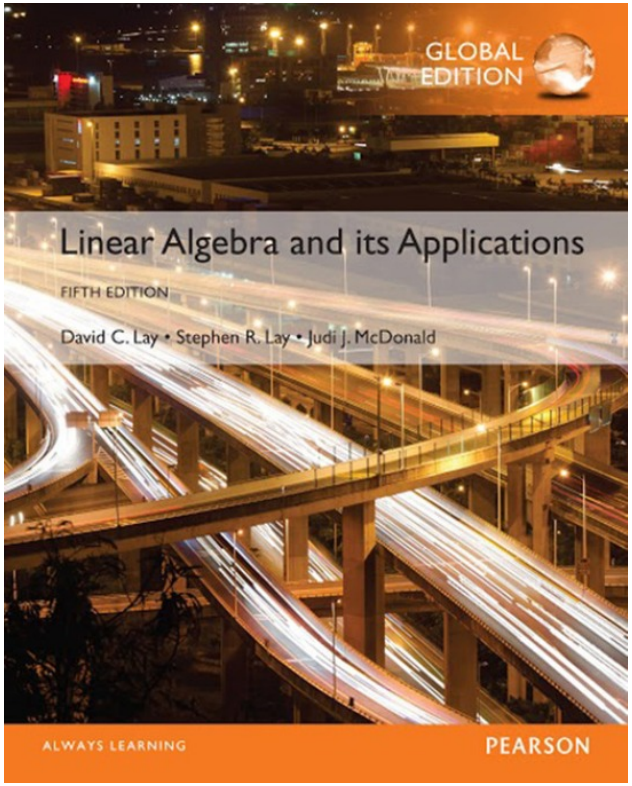
6 Orthogonality and Least Squares

6.4

THE GRAM-SCHMIDT PROCESS



PEARSON

THE GRAM-SCHMIDT PROCESS

- **Theorem 11: The Gram-Schmidt Process**
- Given a basis $\{x_1, \ldots, x_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$v_1 = x_1$$

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

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

$$\vdots$$

$$v_p = xp - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \ldots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}$$

- Then $\{v_1, \ldots, v_p\}$ is an orthogonal basis for W. In addition

$$\text{Span}\{v_1, \ldots, v_k\} = \text{Span}\{x_1, \ldots, x_k\} \quad \text{for } 1 \le k \le p \qquad (1)$$

THE GRAM-SCHMIDT PROCESS

- **Proof** For , let W_k = Span{x_1, . . . , x_k}. Set $v_1 = x_1$, so that Span{v_1} = Span {x_1}. Suppose, for some $k<p$, we have constructed v_1, . . . , v_k so that {v_1, . . . , v_k} is an orthogonal basis for W_k. Define

$$v_{k+1} = x_{k+1} - \text{proj}_{W_k} x_{k+1} \qquad (2)$$

- By Orthogonal Decomposition Theorem, v_{k+1} is orthogonal to W_k. Also, $v_{k+1} \neq 0$ because x_{k+1} is not in W_k = Span{x_1, . . . , x_k}.
- Hence {v_1 , . . . ,v_{k+1}} is an orthogonal set of nonzero vectors in the (k + 1)-dimensional space W_{k+1}. By the Basis Theorem in Section 4.5, this set is an orthogonal basis for W_{k+1}. Hence W_{k+1} = Span{v_1, . . . , v_{k+1}}. When k + 1 = p, the process stops.

THE GRAM-SCHMIDT PROCESS

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.

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

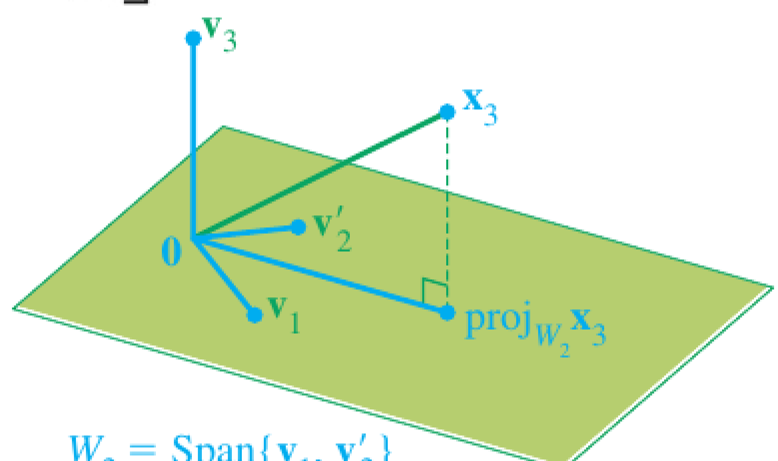
$$\begin{aligned} \mathbf{v}_2 &= \mathbf{x}_2 - \text{proj}_{W_1} \mathbf{x}_2 \\ &= \mathbf{x}_2 - \frac{\mathbf{x}_2 \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\mathbf{v}_1 \qquad \text{Since } \mathbf{v}_1 = \mathbf{x}_1 \\ &= \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{3}{4}\begin{bmatrix}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

For simplification, we can scale v_2 by a factor of 4 $\quad \mathbf{v}'_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}$

THE GRAM-SCHMIDT PROCESS

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

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



$W_2 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2'\}$

ORTHONORMAL BASES

- **Example 3** Example 1 constructed the orthogonal basis

$$v_1 = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix},\ v_2 = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$$

- An orthonormal basis is

$$u_1 = \frac{1}{\|v_1\|}\mathrm{v}_1 = \frac{1}{\sqrt{45}}\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix}$$

$$u_2 = \frac{1}{\|v_2\|}\mathrm{v}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

QR FACTORIZATION OF MATRICES

- **Theorem 12: The QR Factorization**
- If A is an $m*n$ matrix with linearly independent columns, then A can be factored as $A = QR$, where:
  - Q is an $m*n$ matrix whose columns form an orthonormal basis for Col A and
  - R is an $n*n$ upper triangular invertible matrix with positive entries on its diagonal.

Proof The columns of A form a basis $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ for Col A. Construct an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ for W = Col A with property (1) in Theorem 11. This basis may be constructed by e.g., the Gram-Schmidt process.

QR FACTORIZATION OF MATRICES

- Let

$$Q = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n\,]$$

- For $k = 1,\ldots, n$, $\mathbf{x}_k$ is in Span$\{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ = Span$\{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$. So there are constants, $r_{1k}, \ldots, r_{kk}$, such that

$$\mathbf{x}_k = r_{1k}\mathbf{u}_1 + \cdots + r_{kk}\mathbf{u}_k + 0 \cdot \mathbf{u}_{k+1} + \cdots + 0 \cdot \mathbf{u}_n$$

- We may assume that $r_{kk} \geq 0$.
  – If $r_{kk} < 0$, multiply both r_{kk} and $\mathbf{u}_k$ by -1
- This shows that $\mathbf{x}_k$ is a linear combination of the columns of Q using as weights the entries in the vector:

QR FACTORIZATION OF MATRICES

$$\mathbf{r}_k = \begin{bmatrix} r_{1k} \\ \vdots \\ r_{kk} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

- That is, $\mathbf{x}_k = Q\,\mathbf{r}_k$ for $k = 1, \ldots, n$. Let $R = [\mathbf{r}_1 \ldots \mathbf{r}_n]$.
- Then

$$A = [\mathbf{x}_1 \ldots \mathbf{x}_n] = [Q\mathbf{r}_1 \ldots Q\mathbf{r}_n] = QR.$$

QR FACTORIZATION OF MATRICES

- The fact that R is invertible follows easily from the fact that the columns of A are linearly independent:
  - We form: $R\mathbf{v} = 0$, which gives: $QR\mathbf{v} = 0$, and $A\mathbf{v} = 0$.
  - Since columns of A are linearly independent, $A\mathbf{v} = 0$ yields $\mathbf{v} = 0$.
  - Therefore, from $R\mathbf{v} = 0$, we concluded that $\mathbf{v} = 0$, which means R is invertible.

- Since R is clearly upper triangular, its nonnegative diagonal entries must be positive.

QR FACTORIZATION OF MATRICES

- **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 x_1, x_2, and x_3 in Example 2. An orthogonal basis for Col $A = \mathrm{Span}\{\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3\}$ was found in that example:

$$v_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -3 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{bmatrix}$$

QR FACTORIZATION OF MATRICES

- To simplify the arithmetic that follows, scale v_3 by letting $v_3 = 3v_3$. Then normalize the three vectors to obtain u_1, u_2, and 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$\{x_1, \ldots, x_k\}$.

QR FACTORIZATION OF MATRICES

- From the proof of Theorem 12, $A = QR$ for some R. To find R, observe that $Q^TQ = I$, because the columns of Q are orthonormal. Hence

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

- and

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

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