Section 6.4

The Gram–Schmidt Process

Motivation

The procedures in §6 start with an *orthogonal basis* $\{u_1, u_2, \ldots, u_m\}$.

- Find the $\mathcal{B}$-coordinates of a vector x using dot products:

$$x = \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i$$

- Find the orthogonal projection of a vector x onto the span W of $u_1, u_2, \ldots, u_m$:

$$\mathrm{proj}_W(x) = \sum_{i=1}^{m} \frac{x \cdot u_i}{u_i \cdot u_i}\, u_i.$$

Problem: What if your basis isn't orthogonal?

Solution: The **Gram–Schmidt process**: take any basis and *make it orthogonal*.

The Gram–Schmidt Process

Procedure

The Gram–Schmidt Process

Let $\{v_1, v_2, \ldots, v_m\}$ *be a basis* for a subspace W of $\mathbf{R}^n$. Define:

1. $u_1 = v_1$
2. $u_2 = v_2 - \text{proj}_{\text{Span}\{u_1\}}(v_2) \qquad = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1}\, u_1$
3. $u_3 = v_3 - \text{proj}_{\text{Span}\{u_1,u_2\}}(v_3) \qquad = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1}\, u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2}\, u_2$

$\vdots$

m. $u_m = v_m - \text{proj}_{\text{Span}\{u_1,u_2,\ldots,u_{m-1}\}}(v_m) = v_m - \displaystyle\sum_{i=1}^{m-1} \dfrac{v_m \cdot u_i}{u_i \cdot u_i}\, u_i$

Then $\{u_1, u_2, \ldots, u_m\}$ is an *orthogonal basis* for the same subspace W.

Remark

In fact, for every i between 1 and n, the set $\{u_1, u_2, \ldots, u_i\}$ is an *orthogonal basis* for $\text{Span}\{v_1, v_2, \ldots, v_i\}$.

The Gram–Schmidt Process

Example 1: Two vectors

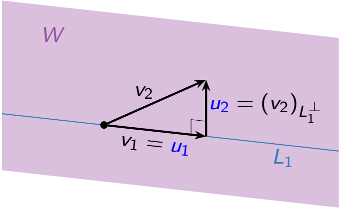
Find an orthogonal basis $\{u_1, u_2\}$ for $W = \text{Span}\{v_1, v_2\}$, where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Run Gram–Schmidt:

1. $u_1 = v_1$
2. $u_2 = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1}\, u_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \dfrac{2}{2} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$

Why does this work?



- First we take $u_1 = v_1$.
- Because $u_1 \cdot v_2 \neq 0$, we can't take $u_2 = v_2$.
- Fix: let $L_1 = \text{Span}\{u_1\}$, and let $u_2 = (v_2)_{L_1^\perp} = v_2 - \text{proj}_{L_1}(v_2)$.
- *By construction,* $u_1 \cdot u_2 = 0$, because $L_1 \perp u_2$.

Remember: This is an orthogonal basis for the *same subspace*.

$$\text{Span}\{u_1, u_2\} = \text{Span}\{v_1, v_2\}$$

The Gram–Schmidt Process

Example 2: Three vectors

Find an orthogonal basis $\{u_1, u_2, u_3\}$ for $W = \mathrm{Span}\{v_1, v_2, v_3\} = \mathbf{R}^3$, where

$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix} \qquad v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix} \qquad v_3 = \begin{pmatrix}3\\1\\1\end{pmatrix}.$$

Run Gram–Schmidt:

1. $u_1 = v_1$
2. $u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1}\, u_1 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \frac{2}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\1\end{pmatrix}$
3. $u_3 = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1}\, u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2}\, u_2$

$$= \begin{pmatrix}3\\1\\1\end{pmatrix} - \frac{4}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} - \frac{1}{1}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

Remember: This is an orthogonal basis for the *same subspace W*.

The Gram–Schmidt Process

Example 2, continued

$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix},\ v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix},\ v_3 = \begin{pmatrix}3\\1\\1\end{pmatrix} \overset{\text{G–S}}{\rightsquigarrow} u_1 = \begin{pmatrix}1\\1\\0\end{pmatrix},\ u_2 = \begin{pmatrix}0\\0\\1\end{pmatrix},\ u_3 = \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

Why does this work?

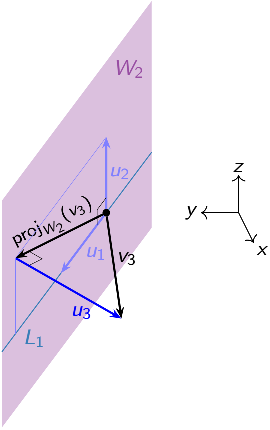
- Once we have u_1 and u_2 orthogonal,
- let $W_2 = \text{Span}\{u_1, u_2\}$, and $u_3 = (v_3)_{W_2^\perp} = v_3 - \text{proj}_{W_3}(u_3)$.
- *By construction, $W_2 \perp u_3$*, so $u_1 \cdot u_3 = 0 = u_2 \cdot u_3$.

Check:

$u_1 \cdot u_2 = 0$ ✓

$u_1 \cdot u_3 = 0$ ✓

$u_2 \cdot u_3 = 0$ ✓



The Gram–Schmidt Process

Example 3: Vectors in $\mathbf{R}^4$

Find an orthogonal basis $\{u_1, u_2, u_3\}$ for $W = \mathrm{Span}\{v_1, v_2, v_3\}$, where

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad v_2 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} \qquad v_3 = \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix}.$$

Run Gram–Schmidt:

1. $u_1 = v_1$
2. $u_2 = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = \begin{pmatrix} -1 \\ 4 \\ 4 \\ -1 \end{pmatrix} - \dfrac{6}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix}$
3. $u_3 = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2$

$$= \begin{pmatrix} 4 \\ -2 \\ -2 \\ 0 \end{pmatrix} - \frac{0}{24} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{-20}{25} \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix}$$

QR Factorization

Recall: A set of vectors $\{v_1, v_2, \ldots, v_m\}$ is *orthonormal* if they are orthogonal unit vectors: $v_i \cdot v_j = 0$ when $i \neq j$, and $v_i \cdot v_i = 1$.

Orthonormal

A matrix Q has orthonormal columns if and only if $Q^T Q = I$.

QR Factorization Theorem

Let A be a matrix with linearly independent columns. Then

$$A = QR$$

where **Q has orthonormal columns** and *R is upper-triangular* with positive diagonal entries.

- The columns of A are a basis for $W = \operatorname{Col} A$.
- The columns of Q are equivalent *basis coming from Gram–Schmidt* (as applied to the columns of A), *after normalizing* to unit vectors.
- The columns of *R come from the steps* in Gram–Schmidt.

Procedure: *QR* Factorization

Through an example

Find the *QR* factorization of $A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}$.

(The columns of A are the vectors v_1, v_2, v_3 from example 2.)

Step 1: Run Gram–Schmidt and *solve for v_1, v_2, v_3 in terms of u_1, u_2, u_3*.

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad v_1 = u_1$$

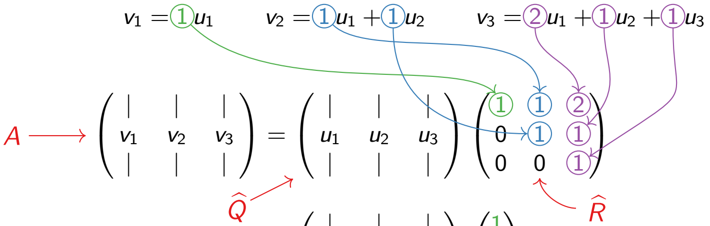
$$u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1}\, u_1 = v_2 - 1\, u_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad v_2 = u_1 + u_2$$

$$\begin{aligned} u_3 &= v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1}\, u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2}\, u_2 \\ &= v_3 - 2\, u_1 - 1\, u_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \end{aligned} \qquad v_3 = 2u_1 + u_2 + u_3$$

QR Factorization

Through an example, continued

Step 2: Write $A = \widehat{Q}\widehat{R}$, where $\widehat{Q}$ has **orthogonal columns** u_1, u_2, u_3 and $\widehat{R}$ is upper-triangular (with 1s on the diagonal) *as shown below*.



$$v_1 = \textcircled{1}u_1 \qquad v_2 = \textcircled{1}u_1 + \textcircled{1}u_2 \qquad v_3 = \textcircled{2}u_1 + \textcircled{1}u_2 + \textcircled{1}u_3$$

$$A \longrightarrow \begin{pmatrix} | & | & | \\ v_1 & v_2 & v_3 \\ | & | & | \end{pmatrix} = \underbrace{\begin{pmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix}}_{\widehat{Q}} \underbrace{\begin{pmatrix} \textcircled{1} & \textcircled{1} & \textcircled{2} \\ 0 & \textcircled{1} & \textcircled{1} \\ 0 & 0 & \textcircled{1} \end{pmatrix}}_{\widehat{R}}$$

$$\text{first column of } A = \begin{pmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 1u_1 = v_1$$

$$\text{second column of } A = \begin{pmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1u_1 + 1u_2 = v_2$$

$$\text{third column of } A = \begin{pmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} = 2u_1 + 1u_2 + 1u_3 = v_3$$

QR Factorization

Through an example, continued

$$A = \widehat{Q}\widehat{R} \qquad \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Step 3: To get Q and R, **scale the columns** of $\widehat{Q}$ **to get unit vectors**, and *scale the rows* of $\widehat{R}$ by the *opposite factor*.

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

It doesn't change the product: the entries in the ith column of Q multiply by the entries in the ith row of R.

The *final QR decomposition* is:

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

QR Factorization

Through a second example

Find the QR factorization of $A = \begin{pmatrix} 1 & -1 & 4 \\ 1 & 4 & -2 \\ 1 & 4 & -2 \\ 1 & -1 & 0 \end{pmatrix}$.

(The columns are vectors from example 3.)

Step 1: Run Gram–Schmidt and *solve for v_1, v_2, v_3 in terms of u_1, u_2, u_3*:

$$u_1 = v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \qquad v_1 = u_1$$

$$u_2 = v_2 - \frac{v_2 \cdot u_1}{u_1 \cdot u_1} u_1 = v_2 - \frac{3}{2} u_1 = \begin{pmatrix} -5/2 \\ 5/2 \\ 5/2 \\ -5/2 \end{pmatrix} \qquad v_2 = \frac{3}{2} u_1 + u_2$$

$$u_3 = v_3 - \frac{v_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{v_3 \cdot u_2}{u_2 \cdot u_2} u_2 = v_3 + \frac{4}{5} u_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ -2 \end{pmatrix} \qquad v_3 = -\frac{4}{5} u_2 + u_3$$

QR Factorization

Through a second example, continued

$$v_1 = 1\, u_1 \qquad v_2 = \frac{3}{2}\, u_1 + 1\, u_2 \qquad v_3 = 0\, u_1 - \frac{4}{5}\, u_2 + 1\, u_3$$

Step 2: Write $A = \widehat{Q}\widehat{R}$, where $\widehat{Q}$ has *orthogonal columns u_1, u_2, u_3* and $\widehat{R}$ is upper-triangular with 1s on the diagonal.

$$\widehat{Q} = \begin{pmatrix} | & | & | \\ u_1 & u_2 & u_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & -5/2 & 2 \\ 1 & 5/2 & 0 \\ 1 & 5/2 & 0 \\ 1 & -5/2 & -2 \end{pmatrix}$$

$$\widehat{R} = \begin{pmatrix} 1 & 3/2 & 0 \\ 0 & 1 & -4/5 \\ 0 & 0 & 1 \end{pmatrix}$$

QR Factorization

Through a second example, continued

$$A = \widehat{Q}\widehat{R} \qquad \widehat{Q} = \begin{pmatrix} 1 & -5/2 & 2 \\ 1 & 5/2 & 0 \\ 1 & 5/2 & 0 \\ 1 & -5/2 & -2 \end{pmatrix} \qquad \widehat{R} = \begin{pmatrix} 1 & 3/2 & 0 \\ 0 & 1 & -4/5 \\ 0 & 0 & 1 \end{pmatrix}$$

Step 3: To get Q and R, *normalize the columns* of $\widehat{Q}$ *and scale the rows* of $\widehat{R}$:

$$Q = \begin{pmatrix} | & | & | \\ u_1/\|u_1\| & u_2/\|u_2\| & u_3/\|u_3\| \\ | & | & | \end{pmatrix}$$

$$R = \begin{pmatrix} 1 \cdot \|u_1\| & 3/2 \cdot \|u_1\| & 0 \cdot \|u_1\| \\ 0 & 1 \cdot \|u_2\| & -4/5 \cdot \|u_2\| \\ 0 & 0 & 1 \cdot \|u_3\| \end{pmatrix}$$

The **final *QR* decomposition** is

$$A = QR \qquad Q = \begin{pmatrix} 1/2 & -1/2 & 1/\sqrt{2} \\ 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & -1/\sqrt{2} \end{pmatrix} \qquad R = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 5 & -4 \\ 0 & 0 & 2\sqrt{2} \end{pmatrix}.$$

Extra: computing determinants

Consider the QR factorization of an *invertible* $n \times n$ matrix: $A = QR$.

- $\det(R)$ is easy to compute because it *is upper-triangular*
- $\det(Q) = \pm 1$ (see below)

Why:

Q is orthonormal, $Q^T Q = I_n$, so $Q^T = Q^{-1}$. Also $\det(Q^T) = \det(Q)$,

$$1 = \det(I_n) = \det(Q^T Q) = \det(Q^T)\det(Q) = \det(Q)^2;$$

so $\det(Q)$ can take only two values: ± 1.

Determinant (up to sign)

If $v_1, v_2, \ldots, v_n$ are the columns of A, and $u_1, u_2, \ldots, u_n$ are the vectors *obtained by applying Gram–Schmidt*, then

$$\det(A) = \det(Q)\det(R) = \pm\|u_1\|\,\|u_2\|\cdots\|u_n\|;$$

Because the (i,i) entry of R is $\|u_i\|$.

- Moreover, $\det(R) > 0$ so $\det(Q)$ has the same sign as $\det(A)$.

Extra: computing eigenvalues

The QR algorithm

Let A be an $n \times n$ matrix with real eigenvalues. Here is the algorithm:

$$\begin{aligned} A &= Q_1 R_1 && \text{QR factorization} \\ A_1 &= R_1 Q_1 && \text{swap the Q and R} \\ &= Q_2 R_2 && \text{find its QR factorization} \\ A_2 &= R_2 Q_2 && \text{swap the Q and R} \\ &= Q_3 R_3 && \text{find its QR factorization} \\ & && \text{et cetera} \end{aligned}$$

Theorem

The matrices A_k *converge* to an upper triangular matrix whose *diagonal entries are the eigenvalues* of A. Moreover, the convergence is fast!

The QR algorithm

The algorithm above gives a computationally efficient way to find the eigenvalues of a matrix.