Section 6.4

The Gram–Schmidt Process

Motivation

All of the procedures we learned in §§6.2–6.3 require an *orthogonal* basis $\{u_1, u_2, \ldots, u_m\}$.

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

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

- Finding 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

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. $u_2 = v_2 - \mathrm{proj}_{\mathrm{Span}\{u_1\}}(v_2) \quad = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1}\, u_1$
3. $u_3 = v_3 - \mathrm{proj}_{\mathrm{Span}\{u_1, u_2\}}(v_3) \quad = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1}\, u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2}\, u_2$

$\vdots$

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

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 $\mathrm{Span}\{v_1, v_2, \ldots, v_i\}$.

The Gram–Schmidt Process

Two vectors

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

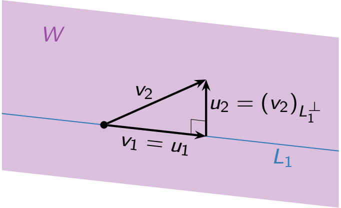
$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix}.$$

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 $u_1 = v_1$.
- Now we're sad because $u_1 \cdot v_2 \neq 0$, so we can't take $u_2 = v_2$.
- Fix: let $L_1 = \mathrm{Span}\{u_1\}$, and let $u_2 = (v_2)_{L_1^\perp} = v_2 - \mathrm{proj}_{L_1}(v_2)$.
- By construction, $u_1 \cdot u_2 = 0$, because $L_1 \perp u_2$.



Important: $\mathrm{Span}\{u_1, u_2\} = \mathrm{Span}\{v_1, v_2\} = W$: this is an *orthogonal* basis for the *same* subspace.

The Gram–Schmidt Process

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. $u_2 = v_2 - \dfrac{v_2 \cdot u_1}{u_1 \cdot u_1}\, u_1 = \begin{pmatrix}1\\1\\1\end{pmatrix} - \dfrac{2}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} = \begin{pmatrix}0\\0\\1\end{pmatrix}$
3. $u_3 = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1}\, u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2}\, u_2$
$$= \begin{pmatrix}3\\1\\1\end{pmatrix} - \frac{4}{2}\begin{pmatrix}1\\1\\0\end{pmatrix} - \frac{1}{1}\begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

Important: $\mathrm{Span}\{u_1, u_2, u_3\} = \mathrm{Span}\{v_1, v_2, v_3\} = W$: this is an *orthogonal* basis for the *same* subspace.

The Gram–Schmidt Process

Three vectors, continued

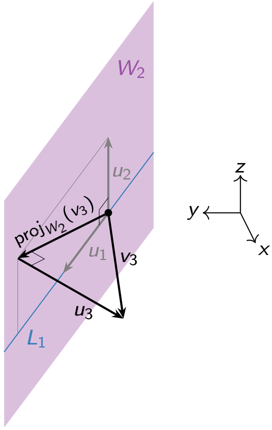
$$v_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \; v_2 = \begin{pmatrix}1\\1\\1\end{pmatrix}, \; v_3 = \begin{pmatrix}3\\1\\1\end{pmatrix} \xrightarrow{\text{G–S}} u_1 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \; u_2 = \begin{pmatrix}0\\0\\1\end{pmatrix}, \; u_3 = \begin{pmatrix}1\\-1\\0\end{pmatrix}$$

Why does this work?

- Once we have u_1 and u_2, then we're sad because v_3 is not orthogonal to u_1 and u_2.
- Fix: let $W_2 = \mathrm{Span}\{u_1, u_2\}$, and let $u_3 = (v_3)_{W_2^\perp} = v_3 - \mathrm{proj}_{W_3}(u_3)$.
- By construction, $u_1 \cdot u_3 = 0 = u_2 \cdot u_3$ because $W_2 \perp u_3$.

Check:

$$u_1 \cdot u_2 = 0 \quad ✓$$
$$u_1 \cdot u_3 = 0 \quad ✓$$
$$u_2 \cdot u_3 = 0 \quad ✓$$



The Gram–Schmidt Process

Three 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. $u_3 = v_3 - \dfrac{v_3 \cdot u_1}{u_1 \cdot u_1}\, u_1 - \dfrac{v_3 \cdot u_2}{u_2 \cdot u_2}\, u_2$

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

Poll

Poll

What happens if you try to run Gram–Schmidt on a linearly *dependent* set of vectors $\{v_1, v_2, \ldots, v_m\}$?

A. You get an inconsistent equation.

B. For some i you get $u_i = u_{i-1}$.

C. For some i you get $u_i = 0$.

D. You create a rift in the space-time continuum.

If $\{v_1, v_2, \ldots, v_m\}$ is linearly dependent, then some v_i is in $\text{Span}\{v_1, v_2, \ldots, v_{i-1}\} = \text{Span}\{u_1, u_2, \ldots, u_{i-1}\}$.

This means

$$v_i = \text{proj}_{\text{Span}\{u_1, u_2, \ldots, u_{i-1}\}}(v_i)$$
$$\implies u_i = v_i - \text{proj}_{\text{Span}\{u_1, u_2, \ldots, u_{i-1}\}}(v_i) = 0.$$

In this case, you can simply discard u_i and v_i and continue: so Gram–Schmidt produces an orthogonal basis from any spanning set!

QR Factorization

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.

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 $v_i \cdot v_i = 1$.

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

The columns of A are a basis for $W = \operatorname{Col} A$. The columns of Q come 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.

Here is the procedure for producing a QR factorization.

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 a previous example.)

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

Example, continued

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

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.

Do this by putting the above equations in matrix form:

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

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

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

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

QR Factorization

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

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

Note that the entries in the ith column of Q multiply by the entries in the ith row of R, so this doesn't change the product.

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

Another 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 a previous example.)

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

Another 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

Another 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: Normalize the columns of $\widehat{Q}$ and the rows of $\widehat{R}$ to get Q and R:

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

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

The final QR decomposition is

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

Let A be an *invertible* $n \times n$ matrix. Consider its QR factorization

$$A = QR.$$

Recall: Since Q has orthonormal columns, $Q^T Q = I_n$, so $Q^T = Q^{-1}$.
But $\det(Q^T) = \det(Q)$, so

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

It follows that $\det(Q) = \pm 1$.

(Since $\det(R) > 0$, in fact $\det(Q)$ has the same sign as $\det(A)$.)

Therefore,

$$\det(A) = \det(Q)\det(R) = \pm\det(R).$$

But R is upper-triangular, so it's easy to compute its determinant!

In fact, if $v_1, v_2, \ldots, v_n$ are the columns of A, and $u_1, u_2, \ldots, u_n$ are the vectors you obtain by applying Gram–Schmidt, then the (i, i) entry of R is $\|u_i\|$, so

$$\det(A) = \pm\|u_1\|\,\|u_2\| \cdots \|u_n\|.$$

So you can use Gram–Schmidt to compute determinants (up to sign)!

QR Factorization

Application: computing eigenvalues

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

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

Theorem

The matrices A_k converge to an upper triangular matrix, and the diagonal entries converge (quickly!) to the eigenvalues of A.

This gives a computationally efficient way (called the QR algorithm) to find the eigenvalues of a matrix.