Outline

1. Student Learning Objectives
   - SLOs: Gram-Schmidt Process and QR Factorization
2. Gram-Schmidt Orthogonalization and QR Factorization
   - The Gram-Schmidt Orthogonalization Process
   - The QR Factorization
   - Observations
3. Suggested Problems
   - Suggested Problems 5.2
   - Lecture–Book Roadmap
4. Supplemental Material
   - Metacognitive Reflection
   - Problem Statements 5.2
   - Why Orthogonal Projections Matter
   - Example: \( V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4 \)
   - Example: 5.2.35 and Beyond — “Live Math” Discussion

Student Learning Objectives

SLOs: Gram-Schmidt Process and QR Factorization

After this lecture you should know how:

- to perform The Gram-Schmidt Orthogonalization Process on a set of vectors, and
- it can be used to compute The QR-factorization of a matrix \( A: A = QR \)
  
  \( \Rightarrow \) This builds an orthonormal basis (the columns of \( Q \)) for the subspace \( V = \text{im}(A) \), which gives us the means to compute the orthogonal projection \( \text{proj}_V(\vec{x}) \) onto \( V \).
- to orthogonally project onto any subspace.

Orthogonal Projection onto a Subspace \( V \)

From [Notes#5.1] we have:

**Theorem (Formula for the Orthogonal Projection)**

If \( V \) is a subspace of \( \mathbb{R}^n \) with an orthonormal basis \( \vec{u}_1, \ldots, \vec{u}_m \), then

\[
\text{proj}_V(\vec{x}) = \vec{x}^\parallel = (\vec{u}_1 \cdot \vec{x})\vec{u}_1 + \cdots + (\vec{u}_m \cdot \vec{x})\vec{u}_m
\]

\( \forall \vec{x} \in \mathbb{R}^n \).

*How do you project onto a subspace if/when the given basis is not orthonormal?!?*

It turns out that before we compute the projection, we have to find a new — orthonormal — basis...
Example: Doing it Right...

In this case, given a basis of \( \mathbb{R}^2 \), the answer is "obvious." Next, we develop (still so we easily visualize and use our intuition) a method for building an orthonormal basis given any starting basis. Once we have the orthonormal basis, we can use the projection formula.

We find an orthonormal basis \( \vec{v}_1, \ldots, \vec{v}_m \) so that \( V = \text{span}(\vec{v}_1, \ldots, \vec{v}_m) \) and then use the projection formula.

\[
\text{proj}_\vec{v}(x) = \frac{(\vec{v}, x)}{||\vec{v}||^2} \vec{v}
\]

Even if we remember to correct for the non-unit length of \( \vec{v} \),...
Example: Doing it Right...

(i) Rescale the first vector

\[ \vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \text{and} \quad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \]

* Divide by the original norm, \( \sqrt{2} \).

Figure: Rescale \( \vec{v}_1 \) to be norm 1, and call it \( \vec{q}_1 \).

(ii) Split the next vector into \( \parallel \) and \( \perp \) parts..

\[ \vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 \parallel \vec{q}_1 = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}; \quad \vec{v}_2 \perp \vec{q}_1 = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}; \quad \text{and} \quad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \]

\[ \vec{q}_2 = \vec{v}_2 - \vec{v}_2 \parallel \vec{q}_1 \]

Figure: Next, project \( \vec{v}_2 \) onto \( \vec{q}_1 \) and get \( \vec{v}_2 \parallel \vec{q}_1 \) and \( \vec{v}_2 \perp \vec{q}_1 \).

(iii) Rescale the \( \perp \) part, discard \( \parallel \) part

\[ \vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 \parallel \vec{q}_1 = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}; \quad \vec{v}_2 \perp \vec{q}_1 = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}; \quad \text{and} \quad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \]

\[ \vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 \parallel \vec{q}_1 = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}; \quad \vec{v}_2 \perp \vec{q}_1 = \begin{bmatrix} -0.5 \\ 0.5 \end{bmatrix}; \quad \text{and} \quad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \]

Now we have an orthonormal basis \( \Omega = \langle \vec{q}_1, \vec{q}_2 \rangle \!.

Figure: Finally, we can use the projection formula.

\[ \text{proj}_\Omega(\vec{x}) = (\vec{q}_1 \cdot \vec{x})\vec{q}_1 + (\vec{q}_2 \cdot \vec{x})\vec{q}_2 = \frac{5}{\sqrt{2}} \vec{q}_1 + \frac{1}{\sqrt{2}} \vec{q}_2 \]

\[ \frac{5}{\sqrt{2}} \begin{bmatrix} 1 \\ \frac{1}{\sqrt{2}} \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}. \]
Example: Doing it Right... Coordinates

In the context of Coordinates (Notes#3.4), we have

Basis: \( \Omega = \langle \vec{q}_1, \vec{q}_2 \rangle \)

Coordinates: \[ [k]_\Omega = \begin{bmatrix} \frac{5}{\sqrt{2}} \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 5 \\ 1 \end{bmatrix} \]

Milking the Example for More Details...

We have performed a Change of Basis, in this case for the purpose of making the projection onto the subspace easily (after the change of basis, that is) computable.

It is “easy” to see that

\[
A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 3 \\ 0 & \sqrt{2} \end{bmatrix},
\]

we have \( A = QR \), where \( Q \) is the new orthonormal basis, and \( R \) is an upper triangular matrix.

The entries in the \( R \) matrix are \( \sqrt{2} \): the original norm of \( \vec{v}_1 \); \( \frac{1}{\sqrt{2}} \): the dot product \( \langle \vec{q}_1, \vec{v}_2 \rangle \); \( \frac{1}{\sqrt{2}} \): the norm of \( \vec{v}_2/\sqrt{2} \). Not likely a coincidence...

Let’s Ponder Higher Dimensions

When you have more basis vectors \( \vec{v}_1, \ldots, \vec{v}_n \) needing orthogonalization (to make an orthonormal basis):

Theorem (Gram-Schmidt Process (annotated))

- Start like we did:
  - \( \vec{q}_1 = \frac{\vec{v}_1}{\|\vec{v}_1\|} \)
  - \( \vec{w}_2 = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\vec{q}_1 \), note that this is a vector in the orthogonal complement of \( \text{span}(\vec{q}_1) = \text{span}(\vec{v}_1) \).
  - \( \vec{q}_2 = \frac{\vec{w}_2}{\|\vec{w}_2\|} \)
- Each time we grab a new vector \( \langle \vec{v}_k \rangle \), find a “help vector” \( \vec{w}_k \) in the orthogonal complement of the space spanned by the previously computed \( \vec{q}_k \)-vectors:
  - \( \vec{w}_k = \vec{v}_k - (\vec{q}_1 \cdot \vec{w}_k)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 - \cdots - (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1} \)
  - Then \( \vec{q}_k = \frac{\vec{w}_k}{\|\vec{w}_k\|} \).

The QR Factorization

The Gram-Schmidt process computed a change of basis from the old basis (funky-script-A)

\[ k = (\vec{v}_1, \ldots, \vec{v}_n) \]

to a new orthonormal basis (funky-script-Q)

\[ \Omega = (\vec{q}_1, \ldots, \vec{q}_n) \]

We describe the result using the change-of-basis-Matrix \( R \) from \( k \) to \( \Omega \), writing

\[ (\vec{v}_1 \cdots \vec{v}_n) = (\vec{q}_1 \cdots \vec{q}_n) R \]
Interpretations and Relations

With $A = QR$, we have the following relations:

- $[\bar{x}]_\Omega = R[x]_\Omega$
  - Multiplication by $R$ moves us from $A$-coordinates to $Q$-coordinates.
- $\bar{x} = Q[x]_\Omega = Q(R[x]_\Omega)$
  - Multiplying the $Q$-coordinate vector by $Q$ "builds" the vector $\bar{x}$.
- $\bar{x} = A[x]_\Omega$
  - Multiplying the $A$-coordinate vector by $A$ "builds" the (same) vector $\bar{x}$.

The "burning" question is how do we construct $R$? It turn out we already have all the pieces, we just need some book-keeping.

What’s in $R$?

OK, let’s rearrange the previous expression:

$$\bar{v}_k = \left( \frac{(\bar{q}_1 \cdot \bar{v}_k) \bar{q}_1 - (\bar{q}_2 \cdot \bar{v}_k) \bar{q}_2 - \cdots - (\bar{q}_{k-1} \cdot \bar{v}_k) \bar{q}_{k-1} + \bar{w}_k}{\bar{v}_k^\perp} \right) \bar{q}_k^\perp$$

The next thing we do is normalize $\bar{v}_k^\perp$ to be norm 1, and name it $\bar{q}_k$; which means we can write the relation above:

$$\bar{v}_k = \left( \frac{(\bar{q}_1 \cdot \bar{v}_k) \bar{q}_1 + (\bar{q}_2 \cdot \bar{v}_k) \bar{q}_2 + \cdots + (\bar{q}_{k-1} \cdot \bar{v}_k) \bar{q}_{k-1} + ||\bar{v}_k^\perp|| \bar{q}_k}{\bar{v}_k^\perp} \right) \bar{q}_k$$

This is the "recipe" for rebuilding the $k$th column of $A$ using the first $k$ columns of $Q$. The entries in $R$ are given by

- $r_{\ell,k} = (\bar{q}_\ell \cdot \bar{v}_k), \ell < k$; $(r_{\ell,k} = 0, \ell > k)$, and
- $r_{k,k} = ||\bar{v}_k^\perp||$. 

What’s in $R$?

If we think back to the $k$th step, we compute

$$\bar{w}_k = \bar{v}_k - \left( \frac{(\bar{q}_1 \cdot \bar{v}_k) \bar{q}_1 - (\bar{q}_2 \cdot \bar{v}_k) \bar{q}_2 - \cdots - (\bar{q}_{k-1} \cdot \bar{v}_k) \bar{q}_{k-1}}{\bar{q}_k^\perp} \right) \bar{q}_k^\perp$$

$\bar{v}_k^\perp$ is orthogonal to $V_{k-1} = \text{span}(\bar{q}_1, \ldots, \bar{q}_{k-1}) = \text{span}(\bar{v}_1, \ldots, \bar{v}_{k-1})$, and $\bar{v}_k^\perp \in \text{span}(\bar{q}_1, \ldots, \bar{q}_{k-1})$.

Note: Subspaces, Orthogonal Complements, and Bases

We are constructing a sequence of subspace-pairs

$$V_k \oplus V_k^\perp = \mathbb{R}^n; \quad \text{dim}(V_k) = k, \quad \text{dim}(V_k^\perp) = (n - k); \quad k = 1, \ldots, n$$

and orthonormal bases $\Omega_k = (\bar{q}_1, \ldots, \bar{q}_k)$ for each of the $V_k$-spaces; and we have $V_{k-1} \subset V_k$ and $V_k^\perp \subset V_{k-1}^\perp$.

We are explicitly constructing $V_k$ and $\Omega_k$; whereas we’re only concerned with a specific vector $\bar{q}_k^\perp \in V_k^\perp$.
The QR-factorization

Consider an \((n \times m)\) matrix \(A\), with linearly independent columns, \(
\vec{v}_1, \ldots, \vec{v}_m \in \mathbb{R}^n\). Then there exists an \((n \times m)\) matrix \(Q\) whose columns \(\vec{q}_1, \ldots, \vec{q}_m \in \mathbb{R}^n\) are orthonormal, and an upper triangular matrix \(R\) with positive diagonal entries such that \(A = QR\). This representation is unique.

Further:
\[
\begin{align*}
\| \vec{v}_1 \| & = r_{11}, \\
r_{kk} & = \| \vec{v}_k \| - \sum_{i=1}^{k-1} \langle \vec{v}_k, \vec{q}_i \rangle \vec{q}_i, \\
r_{\ell,k} & = \langle \vec{v}_k, \vec{q}_\ell \rangle, \quad \ell, k \in \{1, \ldots, k-1\}.
\end{align*}
\]

Note that
\[
[\text{QR-factorization}] = [\text{Gram-Schmidt}] + [\text{Bookkeeping}].
\]

Available on Learning Glass videos:
5.2 — 3, 7, 13, 31, 32, 33, 35, 39

### Observations

\[
A = [\vec{v}_1 \cdots \vec{v}_m] = QR, \quad A \in \mathbb{R}^{n \times m}
\]

- Note that \(\text{span}(\vec{q}_1, \ldots, \vec{q}_k) = \text{span}(\vec{v}_1, \ldots, \vec{v}_k), \quad k = 1, \ldots, m\) (that's the point — we are building an orthonormal set of vectors, describing the same subspaces spanned the columns of the matrix \(A\)).
- Let \(V_k = \text{span}(\vec{q}_1, \ldots, \vec{q}_k) \equiv \text{span}(\vec{v}_1, \ldots, \vec{v}_k)\); these subspaces are “nested”:

\[
V_0 \subset V_1 \subset \cdots \subset V_k,
\]

\[
\text{dim}(V_0) \leq \text{dim}(V_1) \leq \cdots \leq \text{dim}(V_k),
\]

(the maximal dimension is limited by the number of linearly independent vectors in \(\{\vec{v}_1, \ldots, \vec{v}_k\}\)).

- \#ProjectionFestival

\[
\text{proj}_{V_k}(\vec{x}) = (\vec{x} \cdot \vec{q}_1)\vec{q}_1 + \cdots + (\vec{x} \cdot \vec{q}_k)\vec{q}_k
\]
(5.2.3), (5.2.7)

(5.2.3) Perform the Gram-Schmidt process on the sequence of vectors given:
\[
\vec{v}_1 = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 25 \\ 0 \\ -25 \end{bmatrix}.
\]

(5.2.7) Perform the Gram-Schmidt process on the sequence of vectors given:
\[
\vec{v}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 18 \\ 0 \end{bmatrix}.
\]

(5.2.13), (5.2.31)

(5.2.13) Perform the Gram-Schmidt process on the sequence of vectors given:
\[
\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}.
\]

(5.2.31) Perform the Gram-Schmidt process on the following basis of \( \mathbb{R}^3 \):
\[
\vec{v}_1 = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} d' \\ e \\ f \end{bmatrix}.
\]
(5.2.39) Find an orthonormal basis \( \langle \vec{u}_1, \vec{u}_2, \vec{u}_3 \rangle \) of \( \mathbb{R}^3 \), such that
\[
\text{span}(\vec{u}_1) = \text{span} \left( \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right),
\]
and
\[
\text{span}(\vec{u}_1, \vec{u}_2) = \text{span} \left( \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right),
\]
Why Orthogonal Projections Matter \( \rightsquigarrow \) Solving the “Unsolvable”

Experience shows that at this point, most students tend to be a bit lost...

Known We need orthogonal bases to perform (correct) orthogonal projections to higher dimensional \( (n \geq 2) \) subspaces.

But The previous example (projecting from \( \mathbb{R}^2 \to \mathbb{R}^2 \)) was not very satisfying...

Mystery Why are orthogonal projections such a big deal? (Bad reasons include:)
- “The professor said so.” (multiple times)
- “It’ll be on the test.”

The goal of the next example is to give some idea as to why orthogonal projections can be useful... while re-visiting and connecting several “old” ideas.

Where \( \vec{w} \in \mathbb{R}^n \), \( L = \{ k \vec{w}, k \in \mathbb{R} \} \) is the (line) subspace of \( \mathbb{R}^n \).

Important Note: \( \vec{b}_{\parallel} \) is the point (in the subspace \( L \)) which is closest to \( \vec{b} \).

Now, let
\[
A = \begin{bmatrix} 1 \\ \vec{w} \end{bmatrix} \in \mathbb{R}^{n \times 1},
\]
then we are interested in solving the linear system \( A\vec{x} = \vec{b} \), where \( \vec{x} \in \mathbb{R}^1 \) (for now), and \( \vec{b} \in \mathbb{R}^n \).

The system has a solution if and only if \( \vec{b} \in \text{im}(A) = L \).

When \( \vec{b} \not\in \text{im}(A) \) we can either
- say “\( \text{you guys, I’m going home!} \)”, or
- extend the concept of a “solution” to the problem...
Why Orthogonal Projections Matter \(\Rightarrow\) Solving the “Unsolvable”

Since this is not a South Park episode, we decide to extend the concept of what it means to “solve” this problem:
We decide to look for a value \(x^*\) which makes the residual*

\[
r(x^*) = \|Ax^* - \vec{b}\|
\]
as small as possible.

In our example, that value is \(x^* = \left( \frac{\vec{b} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \right)\), which makes \(Ax^* = \vec{b}\).
and \(r(x^*) = \|\vec{b}\| - \|\vec{b}\| = \|\vec{b}\| = \|\vec{b}\|\).
It is true in general that the shortest distance between \(\vec{b}\) and a subspace \(L\), is \(\vec{b} - \text{proj}_L(\vec{b})\).

* think of is as a measure of how far we are from solving the linear system in the “traditional” sense.

Peter Blomgren ⟨blomgren@sdsu.edu⟩ 5.2. Gram–Schmidt and QR Factorization — (33/52)

Why Orthogonal Projections Matter \(\Rightarrow\) Solving the “Unsolvable”

Case \((n = 2, \text{two distinct points})\): In this case we have a unique solution. In our notation the solutions are given by

\[
\begin{bmatrix}
1 & x_1 \\
1 & x_2
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix}
= 
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\]

which gives

\[
\begin{bmatrix}
a \\
b
\end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \end{bmatrix}^{-1} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}
\]

where the inverse is guaranteed to exist when \(x_1 \neq x_2\).

Peter Blomgren ⟨blomgren@sdsu.edu⟩ 5.2. Gram–Schmidt and QR Factorization — (35/52)

Why Orthogonal Projections Matter \(\Rightarrow\) Solving the “Unsolvable”

Case \((n = 3, \text{three distinct points})\): In this case we have no solution. In our notation the solutions would be given by

\[
\begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
1 & x_3
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix}
= 
\begin{bmatrix}
y_1 \\
y_2 \\
y_3
\end{bmatrix}
\]

which gives

\[
\begin{bmatrix}
a \\
b
\end{bmatrix} = \text{MAGIC MATRIX} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}
\]

There is no solution, unless the 3 points are on a common line...

Peter Blomgren ⟨blomgren@sdsu.edu⟩ 5.2. Gram–Schmidt and QR Factorization — (36/52)
Why Orthogonal Projections Matter ↦ Solving the “Unsolvable”

Case (n = large, many (distinct) points):
In this case we have no solution. In our notation the solutions would be given by

\[
\begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\]

which gives

\[
\begin{bmatrix}
a \\
b
\end{bmatrix} = \text{MAGIC MATRIX} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}
\]

There is no solution, unless the ALL points are on a common line...

Peter Blomgren ⟨blomgren@sdsu.edu⟩
5.2. Gram-Schmidt and QR Factorization — (37/52)

Why Orthogonal Projections Matter ↦ Solving the “Unsolvable”

Now, if we are looking for a best-extended-concept-of-solution candidate; we compute \( proj_P(\vec{y}) \equiv \vec{y}^\parallel \), and the system

\[
\begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} = \text{proj}_P(\vec{y})
\]

does have a unique solution, call it \( \vec{c}^* \); and the residual

\[
r(\vec{c}^*) = \|A\vec{c}^* - \vec{y}\| = \|\vec{y}\parallel - \vec{y}\parallel = \|\vec{y}^\perp\|
\]
is minimized.

Peter Blomgren ⟨blomgren@sdsu.edu⟩
5.2. Gram-Schmidt and QR Factorization — (39/52)

Why Orthogonal Projections Matter ↦ Solving the “Unsolvable”

Staying in the general \( n = \text{large} \) case, with

\[
\begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix}
\begin{bmatrix}
a \\
b
\end{bmatrix} =
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\]

In our linear algebra language, we “know” that \( P = \text{im}(A) \) is a 2-dimensional subspace of \( \mathbb{R}^n \) (the two columns are different, unless all the \( x_k \)s coincide)... and, of course, we only have a solution if/when \( \vec{y} \) can be written as a linear combination of the columns of \( A \iff \vec{y} \in \text{im}(A) \).

Peter Blomgren ⟨blomgren@sdsu.edu⟩
5.2. Gram-Schmidt and QR Factorization — (40/52)
Supplemental Material

Solved Problems

Example: \( V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4 \)

What is your problem?!?

Find an orthonormal basis for the subspace
\[
V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4,
\]
then project the vectors
\[
\vec{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \text{ and } \vec{y}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}
\]
onto \( V \).

First, we need a basis for \( V \); finding \( \ker(\begin{bmatrix} 1 & 1 & 1 \end{bmatrix}) \) will do the trick.

Since \( A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \) already is in rref, we can identify the solutions to \( \vec{A} \vec{x} = 0 \):
\[
\vec{x} = s \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix},
\]
so our basis is
\[
B_V = (\vec{v}_1, \vec{v}_2, \vec{v}_3) = \left( \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right); \quad A = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]
as an added bonus we will compute the QR-factorization of \( A \).

\[
\vec{q}_1 \cdot \vec{y}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{2}} ((-1)^2 + 1 	imes 0 + 0 	imes 1 + 0 	imes 0) = \frac{1}{\sqrt{2}}
\]

\[
\vec{v}_2^\perp = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\vec{q}_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \left( \frac{1}{\sqrt{2}} \right) \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}
\]

\[
||\vec{v}_2^\perp|| = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 + 1 + 4 + 0 \end{bmatrix} = \sqrt{6} \frac{1}{2}
\]

\[
\vec{q}_2 = \frac{1}{||\vec{v}_2^\perp||} \vec{v}_2^\perp = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}
\]

\[
Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{6} & -1/\sqrt{2} & 0 \\ 0 & 2/\sqrt{6} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{6}/2 \end{bmatrix}
\]
Example: \( V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4 \)
Example: 5.2.35 and Beyond — “Live Math” Discussion

\[ V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4 \] 5 of 8

Projections!

\[ \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \]

\[ \vec{v}_1 \cdot \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} (-1 + 1 + 0 + 0) = 0 \]

\[ \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \]

\[ \vec{v}_2 = \vec{y}_1 - (\vec{v}_1 \cdot \vec{y}_1) \frac{\vec{v}_1}{\|\vec{v}_1\|^2} \]

\[ \|\vec{v}_1\|^2 = \frac{1}{3} ] + 1 + 1 + 9 = \sqrt{12} \]

\[ \vec{v}_3 = \frac{1}{\|\vec{v}_3\|} \frac{\vec{v}_1}{\sqrt{12}} = \frac{1}{\sqrt{12}} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \]

\[ Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 0 & 2/\sqrt{6} \\ 1/\sqrt{2} & -1/\sqrt{6} & 0 & 2/\sqrt{6} \\ 1/\sqrt{6} & 1/\sqrt{6} & 0 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} 2/\sqrt{6} & 1/\sqrt{6} & 0 & 0 \\ -1/\sqrt{6} & -1/\sqrt{6} & 0 & 0 \end{bmatrix} \]

\[ \text{proj}_V(\vec{v}_1) = 0 \]

Of course! We constructed \( \mathbf{B}_V = (\vec{v}_1, \vec{v}_2, \vec{v}_3) \) by finding all vectors orthogonal to \( \vec{y}_1 \)
((Solving \( [1 1 1 1]^T \vec{x} = \vec{0} \)))
Given $A$, find an orthonormal basis for the invariant subspace $\text{im}(A)$, and the QR-factorization $QR = A$:

\[
A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 2 & -2 & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}, \quad R = \begin{bmatrix} 0 & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix}.
\]

$v_1 := \sqrt{1^2 + 2^2 + 2^2} = \sqrt{9} = 3, \quad q_1 = \frac{v_1}{\|v_1\|} = \frac{1}{3} \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$

\[
Q = \begin{bmatrix} 1/3 & \cdot & \cdot \\ 2/3 & \cdot & \cdot \\ 2/3 & \cdot & \cdot \end{bmatrix}, \quad R = \begin{bmatrix} 3 & \cdot & \cdot \\ 0 & \cdot & \cdot \\ 0 & \cdot & \cdot \end{bmatrix}.
\]

$v_2 := v_2 - (q_1 \cdot v_2)q_1 = \begin{bmatrix} 2 \\ -2 \end{bmatrix} - \left( \begin{bmatrix} 2 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 \\ -2 \end{bmatrix} \right) \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

\[
\|v_2\| = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3, \quad q_2 = \frac{v_2}{\|v_2\|} = \frac{1}{3} \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}
\]

\[
Q = \begin{bmatrix} 1/3 & 2/3 & \cdot \\ 2/3 & 1/3 & \cdot \\ 2/3 & -2/3 & \cdot \end{bmatrix}, \quad R = \begin{bmatrix} 3 & 0 & \cdot \\ 0 & 3 & \cdot \\ 0 & 0 & \cdot \end{bmatrix}.
\]

\[v_3^\perp = 0 \text{ means that } v_3 \text{ is a linear combination of } v_1 \text{ and } v_2.\]

\[\text{Therefore } \text{im}(A) = \text{span}(v_1, v_2) = \text{span}(q_1, q_2).\]

\[\text{We have 2 options for the QR-factorization:} \]

\[
A = \begin{bmatrix} 1/3 & 2/3 & 0 \\ 2/3 & 1/3 & 0 \\ 2/3 & -2/3 & 0 \end{bmatrix}, \quad \text{or} \quad \begin{bmatrix} 1/3 & 2/3 & 3 \\ 2/3 & 1/3 & 0 \\ 2/3 & -2/3 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 3 \\ 1 & 3 & 0 \\ 1 & -3 & 0 \end{bmatrix}.
\]

"Economy Size" QR-factorization

"Full" QR-factorization

Note that in the second version, we have added a third orthonormal vector to the $Q$-matrix, and a row of zeros to the $R$-matrix.