Math 254: Introduction to Linear Algebra
Lecture Notes #5.2 — Gram-Schmidt Process and QR Factorization

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩
Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720
http://terminus.sdsu.edu/

Spring 2019

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

Student Learning Objectives SLOs: Gram-Schmidt Process and QR Factorization
SLOs 5.2 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...
Ponder what happens if we use the formula, but the given basis is **not** orthonormal...

Let’s live in $\mathbb{R}^2$, with basis $\mathcal{B} = (\vec{v}_1, \vec{v}_2)$ defined by

$$
\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \text{and} \quad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}; \quad ||\vec{x}|| = \sqrt{13}.
$$

Clearly $\vec{v}_1$ and $\vec{v}_2$ are linearly independent, and $\vec{x} = 1\vec{v}_1 + 1\vec{v}_2$, but the projection formula goes haywire:

$$
\text{proj}_\mathcal{B}(\vec{x}) = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + (\vec{v}_2 \cdot \vec{x})\vec{v}_2 = 5\vec{v}_1 + 8\vec{v}_2 = \begin{bmatrix} 13 \\ 21 \end{bmatrix}.
$$

... even if we remember to correct for the non-unit length of $\vec{v}_{1,2}$:

$$
\text{proj}_\mathcal{B}(\vec{x}) = \left( \frac{\vec{v}_1 \cdot \vec{x}}{||\vec{v}_1||^2} \right)\vec{v}_1 + \left( \frac{\vec{v}_2 \cdot \vec{x}}{||\vec{v}_2||^2} \right)\vec{v}_2 = \frac{5}{2}\vec{v}_1 + \frac{8}{5}\vec{v}_2 = \begin{bmatrix} 4.1 \\ 5.7 \end{bmatrix}.
$$

**Example: Doing it Right...**

Build an Orthonormal Basis

**Figure:** This is where we start

$$
\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}; \quad \text{and} \quad \vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.
$$

---

**Comments**

There are other ways to realize the “projection” went awry:

- **This is “life in $\mathbb{R}^2$,” and since**

  $$
  \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix},
  $$

  are linearly independent $\rightarrow$ they form a basis for $\mathbb{R}^2$. $\rightarrow$ any projection of a vector $\vec{w} \in \mathbb{R}^2$ onto the subspace $\text{span}(\vec{v}_1, \vec{v}_2) \equiv \mathbb{R}^2$ must be the original vector $\vec{w}$.

- **Even simpler, the famous Method of the Eyeball shows that** $\vec{v}_1 + \vec{v}_2 = \vec{x}$:

  $$
  \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}.
  $$

**Example: Doing it Right...**

(i) Rescale the first vector

**Figure:** Rescale $\vec{v}_1$ to be length 1, and call it $\vec{q}_1$

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

* Divide by the original length, $\sqrt{2}$.
Example: Doing it Right...

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

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

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

Example: Doing it Right...

(iii) Rescale the $\perp$ part, discard $\|\)$ part.

Figure: Next, throw away $\vec{v}_2 \parallel \vec{q}_1$ and rescale $\vec{v}_2 \perp \vec{q}_1$ to length 1, and name it $\vec{q}_2$

$\vec{q}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{q}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$; and $\vec{x} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

In the context of [3.4: Coordinates], we have

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

Coordinates: $[\vec{x}]_\Omega = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$

$= \frac{1}{\sqrt{2}} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$
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

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

\[
A = \begin{bmatrix}
\vec{v}_1 & \vec{v}_2 \\
\vec{q}_1 & \vec{q}_2
\end{bmatrix}
R
Q
\]

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 length of \( \vec{v}_i \); \( \frac{1}{\sqrt{2}} \): the dot product \( \vec{q}_1 \cdot \vec{q}_2 \); \( \frac{1}{\sqrt{3}} \): the length of \( \vec{q}_2 \). 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:**
  - \( \vec{q}_1 = \vec{v}_1 / \| \vec{v}_1 \| \)
  - \( \vec{q}_2 = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\vec{q}_1 \), note that this is a vector in the orthogonal complement of \( \text{span}(\vec{q}_1) = \text{span}(\vec{v}_1) \).
  - \( \vec{q}_2 = \vec{w}_2 / \| \vec{w}_2 \| \)

- **Each time we grab a new vector** \( \vec{v}_k \), find a “help vector” \( \vec{w}_k \) in the orthogonal complement of the space spanned by the previously computed \( \vec{q} \)-vectors:
  - \( \vec{w}_k = \vec{v}_k - (\vec{q}_1 \cdot \vec{v}_k)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_k)\vec{q}_2 - \cdots - (\vec{q}_{k-1} \cdot \vec{v}_k)\vec{q}_{k-1} \)

- **Then** \( \vec{q}_k = \vec{w}_k / \| \vec{w}_k \| \).

Interpretations and Relations

With \( A = QR \), we have to following relations:

- \( \vec{x} = Q\vec{x}_\Omega \)
  - Multiplication by \( R \) moves us from \( \vec{x} \)-coordinates to \( \vec{q} \)-coordinates.

- \( \vec{x} = Q[x]_\Omega = QR[x]_\Xi \)
  - Multiplying the \( Q \)-coordinate vector by \( Q \) “builds” the vector \( \vec{x} \).

- \( \vec{x} = A[x]_\Xi \)
  - Multiplying the \( A \)-coordinate vector by \( A \) “builds” the (same) vector \( \vec{x} \).

The “burning” question is how do we construct \( R \)? It turn out we already have all the pieces, we just need some book-keeping.
The Gram-Schmidt Orthogonalization Process
Suggested Problems

**Observations**

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

$$
\tilde{v}_k = v_k - \sum_{j=1}^{k-1} (\tilde{v}_j \cdot v_k) \tilde{v}_j
$$

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

Note: Subspaces, Orthogonal Complements, and Bases

We are explicitly constructing $V_k$ and $Q_k$: whereas we're only concerned with a specific vector $v_k^\perp \in V_k^\perp$.

Peter Blomgren, (blomgren.peter@gmail.com)

---

The next thing we do is normalize $v_k^\perp$ to be length 1, and name it $\tilde{q}_k$.

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 $r_{k,k} = (\tilde{q}_k \cdot v_k)$, $\ell < k$, and $r_{k,k} = \|v_k^\perp\|, (r_{k,k} = 0, \ell > k)$

Peter Blomgren, (blomgren.peter@gmail.com)

---

**The QR-factorization**

Consider an $n \times m$ matrix $A$, with linearly independent columns, $v_1, \ldots, v_m \in \mathbb{R}^n$. Then there exists an $n \times m$ matrix $Q$ whose columns $\tilde{q}_1, \ldots, \tilde{q}_m \in \mathbb{R}^n$ are orthonormal, and an upper triangular matrix $R$ with positive diagonal entries such that $A = QR$. This representation is unique.

Further

- $r_{11} = \|v_1\|$
- $r_{kk} = \|v_k^\perp \cdot v_k\|, k \in \{2, \ldots, m\}$, and
- $r_{\ell,k} = (\tilde{q}_\ell \cdot v_k), \ell \in \{1, \ldots, k-1\}$.

Note that

$$
[QR\text{-factorization}]=[\text{Gram-Schmidt}]+[\text{Bookkeeping}].
$$
Gram-Schmidt Orthogonalization and QR Factorization
Suggested Problems

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

- Note that $\text{span}(\vec{q}_1, \ldots, \vec{q}_k) = \text{span}(\vec{v}_1, \ldots, \vec{v}_k)$, $k = 1, \ldots, m$ (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”:
  
  $V_0 \subset V_1 \subset \cdots \subset V_k,$
  
  $\dim(V_0) \leq \dim(V_1) \leq \cdots \leq \dim(V_k),$

  (the maximal dimension is limited by the number of linearly independent vectors in $\{\vec{v}_1, \ldots, \vec{v}_k\}$)
- \#ProjectionFest2018

$$\text{proj}_{V_k}(\vec{x}) = (\vec{x} \cdot \vec{q}_1)\vec{q}_1 + \cdots + (\vec{x} \cdot \vec{q}_k)\vec{q}_k$$

Suggested Problems 5.2

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

Lecture – Book Roadmap

<table>
<thead>
<tr>
<th>Lecture</th>
<th>Book, [GS5–]</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>§4.1, §4.2, §4.4</td>
</tr>
<tr>
<td>5.2</td>
<td>§4.1, §4.2, §4.4</td>
</tr>
<tr>
<td>5.3</td>
<td>§4.1, §4.2, §4.4</td>
</tr>
</tbody>
</table>

Supplemental Material
Solved Problems

Metacognitive Reflection
Problem Statements 5.2
Why Orthogonal Projections Matter

Metacognitive Exercise — Thinking About Thinking & Learning

<table>
<thead>
<tr>
<th>I know / learned</th>
<th>Almost there</th>
<th>Huh?!?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Right After Lecture</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

After Thinking / Office Hours / SI-session

After Reviewing for Midterm/Final
(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:
\[ \vec{v}_1 = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} -2 \\ 1 \\ 2 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 18 \\ 0 \\ 0 \end{bmatrix}. \]

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

(5.2.31) Perform the Gram-Schmidt process on the following basis of \( \mathbb{R}^3 \):
\[ \vec{v}_1 = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} d \\ e \\ f \end{bmatrix}. \]

(5.2.33) Find an orthonormal basis for the kernel of the matrix
\[ A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 \end{bmatrix}. \]

(5.2.35) Find an orthonormal basis for the image of the matrix
\[ A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 1 \\ 2 & -2 & 0 \end{bmatrix}. \]
Why Orthogonal Projections Matter ◄ 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 \(\mathbb{R}^2 \rightarrow \mathbb{R}^2\)) 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.

Now, let

\[
A = \begin{bmatrix} \mathbf{w} \\ \end{bmatrix} \in \mathbb{R}^{n \times 1},
\]

then we are interested in solving the linear system \(Ax = \tilde{b}\), where \(x \in \mathbb{R}^1\) (for now), and \(\tilde{b} \in \mathbb{R}^n\).

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

When \(\tilde{b} \notin \text{im}(A)\) we can either

- say “"you guys, I’m going home!"”, or
- extend the concept of a “solution” to the problem...

Recall our old cartoon of orthogonal projections:

\[
\text{where \(\tilde{w} \in \mathbb{R}^n\), } L = \{ k\tilde{w}, \; k \in \mathbb{R}\} \text{ is the (line) subspace of } \mathbb{R}^n.\]

Important Note: \(\tilde{b}||\) is the point (in the subspace \(L\)) which is closest to \(\tilde{b}\).

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* \(r(x) = \|Ax - \tilde{b}\|\) as small as possible.

In our example, that value is \(x^* = \left(\frac{\tilde{b} \cdot \tilde{w}}{\tilde{w} \cdot \tilde{w}}\right)\), which makes \(Ax^* = \tilde{b}||\), and \(r(x^*) = \|\tilde{b}\| - \|\tilde{b}|| = \|\tilde{b}|| - \|\tilde{b}|| = \|\tilde{b}||\).

It is true in general that the shortest distance between \(\tilde{b}\) and a subspace \(L\), is \(\tilde{b}^\perp = \tilde{b} - \text{proj}_L(\tilde{b})\).

* think of is as a measure of how far we are from solving the linear system in the “traditional” sense.
**Why Orthogonal Projections Matter → Solving the “Unsolvable”**

Next we consider a slightly different category of problems: fitting a straight line \( y = a + bx \) to some number of given points in the \( x-y \)-plane, \( \{(x_k, y_k)\}_{k=1}^n \).

**Case \((n = 1, \text{ a single point})\):** In this case we have infinitely many solutions. In our notation the solutions are given by

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

which gives

\[
\begin{bmatrix}
a \\
b
\end{bmatrix}
= [y_1] + s \begin{bmatrix}
-x_1 \\
1
\end{bmatrix}
\]

**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

\[
\begin{bmatrix}
a \\
b
\end{bmatrix}
= \text{Magic Matrix} \begin{bmatrix}
y_1 \\
y_2 \\
y_3
\end{bmatrix}
\]

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

**Case \((n = \text{large}, \text{ many (distinct) points})\):**

In this case we have no solution. In our notation the solutions would be given by

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

which gives

\[
\begin{bmatrix}
a \\
b
\end{bmatrix}
= \text{Magic Matrix} \begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\]

There is no solution, unless the ALL points are on a common line...
Why Orthogonal Projections Matter  \(\leadsto\) Solving the “Unsolvable”

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

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

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 \(A \iff \vec{y} \in \text{im}(A)\).

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

We have defined a new type of “solution” for inconsistent non-square (matrix) problems.

The way we have discussed it, the best name would be a

- “Minimum Residual Solution”

However, the most common mathematical name is the

- “Least Squares Solution”

In many applications (related to statistics), the most common name is the

- “Linear Regression Solution”

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

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

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

\[
r(\vec{c}) = \|A\vec{c} - \vec{y}\| = \|\vec{y}\| - \|\vec{y}^{\perp}\| = \|\vec{y}^{\perp}\|
\]

is minimized.

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

\[
\vec{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \vec{y}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}
\]

onto \(V\).
Supplemental Material
Solved Problems

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

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

1. First, we need a basis for \( V \); finding \( \ker([1 \ 1 \ 1\ 1]) \) will do the trick.

2. Since \( A = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \) already is in rref, we can identify the solutions to \( \bar{A} \vec{x} = 0 \):

\[
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4
\end{bmatrix} = s \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\
1 \\ -1 \\ 1 \\ 0 \\
0 \\ 0 \\ 0 \\ 1
\end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\
0 \\ 1 \\ 0 \\ 1
\end{bmatrix} + u \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1
\end{bmatrix},
\]

so our basis is

\[
B_V = (\vec{v}_1, \vec{v}_2, \vec{v}_3) = \left( \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\
1 \\ -1 \\ 1 \\ 0 \\
0 \\ 1 \\ 0 \\ 1
\end{bmatrix} \right); \quad A = \begin{bmatrix} -1 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1
\end{bmatrix}
\]

as an added bonus we will compute the QR-factorization of \( A \).

---

\[
\bar{\vec{v}_1} = \frac{1}{\|\vec{v}_1\|} \vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0
\end{bmatrix}.
\]

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

\[
\|\vec{v}_2\| = \sqrt{1^2 + 1^2 + 2^2 + 0^2} = \frac{\sqrt{6}}{2}
\]

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

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

---

**Peter Blomgren, (blomgren.peter@gmail.com) Gram-Schmidt Process and QR Factorization — (43/47)**
Additional Example: $V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4$

$V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4$

- $\vec{v}_3^\perp = \vec{v}_3 - (\vec{q}_1 \cdot \vec{v}_3)\vec{q}_1 - (\vec{q}_2 \cdot \vec{v}_3)\vec{q}_2$: 
  \[
  \begin{bmatrix}
  -1 \\
  0 \\
  1 
  \end{bmatrix} - \left( \frac{1}{\sqrt{2}} \right) \begin{bmatrix}
  1 \\
  1 \\
  0 
  \end{bmatrix} \left( \frac{1}{\sqrt{2}} \right) + \left( \frac{1}{\sqrt{6}} \right) \begin{bmatrix}
  -1 \\
  1 \\
  2 \\
  0 
  \end{bmatrix} = \begin{bmatrix}
  1/3 \\
  1 \\
  -1/3 
  \end{bmatrix}
  \]
  \[
  \| \vec{v}_3^\perp \| = \frac{1}{3} \sqrt{1^2 + 1^2 + 9} = \frac{\sqrt{12}}{3}
  \]

- $\vec{q}_3 = \frac{1}{\| \vec{v}_3^\perp \|} \vec{v}_3^\perp = \begin{bmatrix}
  -1/\sqrt{2} \\
  -1/\sqrt{6} \\
  1/\sqrt{2} \\
  2/\sqrt{6} \\
  0 \\
  3/\sqrt{12}
  \end{bmatrix}$

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

- $\vec{y}_1 = \begin{bmatrix}
  1 \\
  1 \\
  1 \\
  1
  \end{bmatrix}$

- $\vec{q}_1 \cdot \vec{y}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix}
  1 \\
  1 \\
  1 \\
  1
  \end{bmatrix} \cdot \begin{bmatrix}
  1 \\
  1 \\
  1 \\
  1
  \end{bmatrix} = \frac{1}{\sqrt{2}} (-1 + 1 + 0) = 0$

- $\vec{q}_2 \cdot \vec{y}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix}
  -1 \\
  -1 \\
  2 \\
  3
  \end{bmatrix} \cdot \begin{bmatrix}
  1 \\
  1 \\
  1 \\
  1
  \end{bmatrix} = \frac{1}{\sqrt{6}} (-1 - 1 + 2 + 0) = 0$

- $\vec{q}_3 \cdot \vec{y}_1 = \frac{1}{\sqrt{12}} \begin{bmatrix}
  -1 \\
  -1 \\
  -1 \\
  3
  \end{bmatrix} \cdot \begin{bmatrix}
  1 \\
  1 \\
  1 \\
  1
  \end{bmatrix} = \frac{1}{\sqrt{12}} (-1 - 1 + 3) = 0$

- $\text{proj}_V(\vec{y}_2) = \vec{0}$

- Of course! We constructed $B_V = (\vec{q}_1, \vec{q}_2, \vec{q}_3)$ by finding all vectors orthogonal to $\vec{y}_1$ ((Solving $[1 1 1 1]\vec{x} = \vec{0}$))