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/
Fall 2018

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
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 5.2 Gram-Schmidt Process and QR Factorization

After this lecture you should:
- Know the The Gram-Schmidt Orthogonalization Process, and
- know how it can be used to compute The QR-factorization of a matrix \( A \).

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

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

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

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...
Gram-Schmidt Process and QR Factorization
Suggested Problems

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 

\[ \vec{v}_1 + \vec{v}_2 = \vec{x}: \]

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

This is ALL WRONG!!!!!!

Ponder what happens if we use the formula, but the given basis is not orthonormal...

Let’s live in \( \mathbb{R}^2 \), with basis \( B = (\vec{v}_1, \vec{v}_2) \) 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 \( \vec{x} = \vec{v}_1 + \vec{v}_2 \), but the projection formula goes haywire:

\[ \text{proj}_B(\vec{x}) = (\vec{v}_1 \cdot \vec{x})\vec{v}_1 + (\vec{v}_2 \cdot \vec{x})\vec{v}_2 = 5\vec{v}_1 + 8\vec{v}_2 = \begin{bmatrix} 13 \\ 21 \end{bmatrix}. \]

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

\[ \text{proj}_B(\vec{x}) = \frac{(\vec{v}_1 \cdot \vec{x})}{||\vec{v}_1||^2} \vec{v}_1 + \frac{(\vec{v}_2 \cdot \vec{x})}{||\vec{v}_2||^2} \vec{v}_2 = \frac{5}{2} \vec{v}_1 + \frac{8}{5} \vec{v}_2 = \begin{bmatrix} 4.1 \\ 5.7 \end{bmatrix}. \]

Suggested Problems

The Gram-Schmidt Orthogonalization Process
The QR Factorization

Example: Doing it Right... (i) Rescale the first vector

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

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

(iii) Rescale the ⃗q part, discard ⃗q parts...

![Figure](image.png)

$v = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $v∥q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$; $v⊥q = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Now we have an orthonormal basis $Ω = \{q, q\}!$

Example: Doing it Right...

(iv) Project using the new orthonormal basis!

![Figure](image.png)

$proj_(q)(x) = (q_1 ∙ x)q_1 + (q_2 ∙ x)q_2 = \frac{5}{\sqrt{2}}q_1 + \frac{1}{\sqrt{2}}q_2$

$\frac{5}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩ Gram-Schmidt Process and QR Factorization — (9/46)

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

BASIS: $Ω = \{q_1, q_2\}$

COORDINATES: $[x]_Ω = \begin{bmatrix} 5 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 5 \\ 1 \end{bmatrix}$
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

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ \sqrt{2} & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} \\ 0 \end{bmatrix}$$

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 length of $v_1$; $\frac{1}{\sqrt{2}}$: the length of $v_1^\perp \overset{\sim}{q}_1$. 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:**
  - $\vec{q}_1 = \vec{v}_1/\|\vec{v}_1\|$
  - $\vec{w}_2 = \vec{v}_2 - (\vec{q}_1 \cdot \vec{v}_2)\vec{q}_1$, note that this is a vector in the orthogonal complement of $\text{span}(\vec{q}_1) = \text{span}(\vec{v}_1)$.
  - $\vec{q}_2 = \vec{w}_2/\|\vec{w}_2\|$
- 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:
  - $\vec{w}_k = \vec{v}_k - (\vec{q}_1 \cdot \vec{w}_k)\vec{q}_1 - (\vec{q}_2 \cdot \vec{w}_k)\vec{q}_2 - \cdots - (\vec{q}_{k-1} \cdot \vec{w}_k)\vec{q}_{k-1}$
- Then $\vec{q}_k = \vec{w}_k/\|\vec{w}_k\|$

Interpretations and Relations

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

- $[x]_\Omega = R[x]_\Xi$
  - Multiplication by $R$ moves us from $A$-coordinates to $Q$-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.
What’s in $R$? 1 of 3

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

$$
\tilde{v}_k = \tilde{v}_k - (\tilde{q}_1 \cdot \tilde{v}_k)\tilde{q}_1 - (\tilde{q}_2 \cdot \tilde{v}_k)\tilde{q}_2 - \cdots - (\tilde{q}_{k-1} \cdot \tilde{v}_k)\tilde{q}_{k-1}
$$

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

Note: Subspaces, Orthogonal Complements, and Bases

We are constructing a sequence of subspace-pairs

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

and orthonormal bases $\Omega_k = (\tilde{q}_1, \ldots, \tilde{q}_k)$ for each of the $V_k$-spaces.

We are explicitly constructing $V_k$ and $V_k^\perp$; whereas we’re only concerned with a specific vector $\tilde{v}_k^\perp \in V_k^\perp$.

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩ Gram-Schmidt Orthogonalization and QR Factorization — (19/46)

Gram-Schmidt Orthogonalization and QR Factorization
Suggested Problems
The Gram-Schmidt Orthogonalization Process
The QR Factorization

What’s in $R$? 2 of 3

OK, let’s rearrange the previous expression:

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

The next thing we do is normalize $\tilde{v}_k^\perp$ 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 $r_{\ell,k} = (\tilde{q}_\ell \cdot \tilde{v}_k)$, $\ell < k$, and $r_{k,k} = \|\tilde{v}_k^\perp\|$. ($r_{\ell,k} = 0, \ell > k$)

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩ Gram-Schmidt Orthogonalization and QR Factorization — (20/46)

Gram-Schmidt Orthogonalization and QR Factorization
Suggested Problems
The Gram-Schmidt Orthogonalization Process
The QR Factorization

Summarizing \(\rightarrow\) The QR-factorization 3 of 3

Theorem (QR-Factorization)

Consider an $n \times m$ matrix $A$, with linearly independent columns, $\tilde{v}_1, \ldots, \tilde{v}_m \in \mathbb{R}^n$. 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

- $r_{11} = \|\tilde{v}_1\|$
- $r_{k,k} = \|\tilde{v}_k^\perp\| \in \text{span}(\tilde{q}_1, \ldots, \tilde{q}_{k-1})$, $k \in \{2, \ldots, m\}$, and
- $r_{\ell,k} = (\tilde{q}_\ell \cdot \tilde{v}_k)$, $\ell \in \{1, \ldots, k-1\}$.

Note that

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

Peter Blomgren, ⟨blomgren.peter@gmail.com⟩ Gram-Schmidt Orthogonalization and QR Factorization — (21/46)
**Available on Learning Glass videos:**
5.2 — 3, 7, 13, 31, 32, 33, 35, 39

---

**Suggested Problems 5.2**

### (5.2.3)
Perform the Gram-Schmidt process on the sequence of vectors given:

\[
\tilde{v}_1 = \begin{bmatrix} 4 \\ 0 \\ 3 \end{bmatrix}, \quad \tilde{v}_2 = \begin{bmatrix} 0 \\ 25 \\ -25 \end{bmatrix}.
\]

### (5.2.7)
Perform the Gram-Schmidt process on the sequence of vectors given:

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

\[
\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 2 \\ 1 \\ -1 \end{bmatrix}.
\]

(5.2.31) Perform the Gram-Schmidt process on the following basis of \( \mathbb{R}^3 \):

\[
\vec{v}_1 = \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} b \\ c \\ 0 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} d \\ e \\ f \end{bmatrix}.
\]

(5.2.39) 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 \( \twoheadrightarrow \text{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.
Why Orthogonal Projections Matter \implies Solving the “Unsolvable”

Recall our old cartoon of orthogonal projections:

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 \( \vec{b} \).

Peter Blomgren, \( \langle \text{blomgren.peter@gmail.com} \rangle \) Gram-Schmidt Process and QR Factorization — (29/46)

---

**Why Orthogonal Projections Matter \implies 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 \( \vec{x^*} \) which makes the residual\(^*\)

\[ r(\vec{x}) = \|A\vec{x} - \vec{b}\| \]

as small as possible.

In our example, that value is \( \vec{x^*} = \frac{\vec{b} \cdot \vec{w}}{\vec{w} \cdot \vec{w}} \vec{w} \), which makes \( A\vec{x^*} = \vec{b} \parallel \),
and \( r(\vec{x^*}) = \| \vec{b} \parallel - \vec{b} \parallel = \| \vec{b} \parallel - \vec{b} \perp = \| \vec{b} \perp \| \).

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

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

Peter Blomgren, \( \langle \text{blomgren.peter@gmail.com} \rangle \) Gram-Schmidt Process and QR Factorization — (31/46)

---

**Why Orthogonal Projections Matter \implies Solving the “Unsolvable”**

Now, let

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

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

- say “\( \exists \) you guys, I’m going home!” 😁, or
- extend the concept of a “solution” to the problem...

Next we consider a slightly different category of problems: fitting a straight line \( y = ax + bx \) 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

\[ \begin{bmatrix} 1 \\ \vec{a} \end{bmatrix} \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \]

which gives

\[ \begin{bmatrix} \vec{a} \\ \vec{b} \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -x_1 \\ 1 \end{bmatrix} \]

Peter Blomgren, \( \langle \text{blomgren.peter@gmail.com} \rangle \) Gram-Schmidt Process and QR Factorization — (32/46)
Why Orthogonal Projections Matter \(\Rightarrow\) Solving the “Unsolvable”

Case \(n = 2\), 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 \(x_1 \neq x_2\).

Case \(n = 3\), 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

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

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

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

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

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 ↼ Solving the “Unsolvable”

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

$$
\begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
\vdots \\
c
\end{bmatrix} = \text{proj}_R(\vec{y})
$$

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.

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”

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 $V$.

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

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

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

so our basis is

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

as an added bonus we will compute the $QR$-factorization of $A$. 
Gram-Schmidt Process and QR Factorization

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

1. \( \| \vec{v}_1 \| = \sqrt{(-1)^2 + 1^2 + 0^2 + 0^2} = \sqrt{2} \)
2. \( \vec{q}_1 = \frac{1}{\| \vec{v}_1 \|} \vec{v}_1 \)

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

We move on to \( \vec{v}_2 \)...

3. \( \vec{q}_1 \cdot \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \)

4. \( \vec{q}_2 = \frac{1}{\| \vec{v}_2 \|} \vec{v}_2 \)

\[
Q = \begin{bmatrix}
-1 / \sqrt{3} & 1 / \sqrt{6} & -1 / \sqrt{2} \\
1 / \sqrt{3} & -1 / \sqrt{6} & 0 \\
0 & 2 / \sqrt{6} & 0 \\
\end{bmatrix}, \quad R = \begin{bmatrix}
\sqrt{3} / 3 & 1 / \sqrt{2} / \sqrt{3} \\
0 & 1 / \sqrt{6} / \sqrt{3} & 0 \\
0 & 0 & 3 / \sqrt{12} / \sqrt{3} \\
\end{bmatrix}
\]
\[ V = \{ x_1 + x_2 + x_3 + x_4 = 0 \} \subset \mathbb{R}^4 \]

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

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

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

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

5. \[ \text{proj}_V(\vec{y}_1) = \vec{0} \]

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 \([1 1 1 1] \vec{x} = \vec{0}\))

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

6. \[ \vec{y}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \]

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

8. \[ \vec{q}_2 \cdot \vec{y}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{6}} (-1 - 2 + 0 + 0) = \frac{3}{\sqrt{6}} \]

9. \[ \vec{q}_3 \cdot \vec{y}_2 = \frac{1}{\sqrt{12}} \begin{bmatrix} -1 \\ -1 \\ 3 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \frac{1}{\sqrt{12}} (-1 - 2 - 3 + 12) = \frac{6}{\sqrt{12}} \]

10. \[ \text{proj}_V(\vec{y}_2) = \frac{1}{2} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} + \frac{1}{6} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} + \frac{6}{7} \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -3 \\ -3 \\ 3 \\ 3 \end{bmatrix} \]