MATH1030 Gram-Schmidt orthogonalization process.

1. Recall the definition for the notions of *orthonormal set and orthonormal basis* from the handout *Orthonormal basis and orthogonal projections.*

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k \in \mathbb{R}^n$.

   (a) *We say that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute an orthonormal set in $\mathbb{R}^n$ if and only if $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ are pairwise orthogonal and $\|\mathbf{u}_j\| = 1$ for each $j = 1, 2 \cdots, k$.*

   (b) *Suppose V is a subspace of $\mathbb{R}^n$ and . Then we say that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute an orthonormal basis for V if and only if $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute a basis for V and also constitute an orthonormal set.*

   Also recall the result ($\star$), which is a part of Theorem (C), as stated below:

   ($\star$) *Let W be a subspace of $\mathbb{R}^n$.*
   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute an orthonormal basis for W.
   Suppose $\mathbf{z} \in \mathbb{R}^n$.
   Define $\mathbf{v} \in W$ by $\mathbf{v} = \langle \mathbf{z}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{z}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{z}, \mathbf{u}_k \rangle \mathbf{u}_k$.
   Define $\mathbf{y} \in \mathbb{R}^n$ by $\mathbf{y} = \mathbf{z} - \mathbf{v}$.
   Then $\mathbf{z} = \mathbf{v} + \mathbf{y}$, and $\mathbf{y} \perp \mathbf{s}$ for any $\mathbf{s} \in W$.

2. **Lemma (G).**

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{z}$ be vectors in $\mathbb{R}^n$.

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute an orthonormal set in $\mathbb{R}^n$.

   Further suppose $\mathbf{z}$ is not a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$.

   Define $\mathbf{y} = \mathbf{z} - \langle \mathbf{z}, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{z}, \mathbf{u}_2 \rangle \mathbf{u}_2 - \cdots - \langle \mathbf{z}, \mathbf{u}_k \rangle \mathbf{u}_k$.

   Then the statements below hold:

   (a) $\|\mathbf{y}\| \neq 0$.

   (b) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$ *constitute an orthonormal set in $\mathbb{R}^n$.*

   (c) $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{z}\}) = \mathsf{Span}\ \left(\left\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}\right\}\right)$.

3. **Proof of Lemma (G).**

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{z}$ be vectors in $\mathbb{R}^n$.

   Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute an orthonormal set in $\mathbb{R}^n$.

   Further suppose $\mathbf{z}$ is not a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$.

   Define $W = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k\})$. By definition, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitutes an orthonormal basis for W.

   Define $\mathbf{y} = \mathbf{z} - \langle \mathbf{z}, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{z}, \mathbf{u}_2 \rangle \mathbf{u}_2 - \cdots - \langle \mathbf{z}, \mathbf{u}_k \rangle \mathbf{u}_k$.

   Define $\mathbf{v} = \langle \mathbf{z}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{z}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{z}, \mathbf{u}_k \rangle \mathbf{u}_k$.

   Then $\mathbf{y} = \mathbf{z} - \mathbf{v}$ by definition.

   (a) Since $\mathbf{z}$ is not a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$, we have $\mathbf{z} \neq \mathbf{v}$. Then $\mathbf{y} = \mathbf{z} - \mathbf{v} \neq \mathbf{0}$. Therefore $\|\mathbf{y}\| \neq 0$.

   (b) By the result ($\star$), $\mathbf{y} \perp \mathbf{s}$ for any $s \in W$.
   Note that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k \in W$.
   Then for each $j = 1, 2, \cdots, n$, we have $\left\langle \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}, \mathbf{u}_j \right\rangle = \dfrac{1}{\|\mathbf{y}\|} \langle \mathbf{y}, \mathbf{u}_j \rangle = 0$. Hence $\dfrac{1}{\|\mathbf{y}\|}\mathbf{y} \perp \mathbf{u}_j$.
   Also note that $\left\| \dfrac{1}{\|\mathbf{y}\|}\mathbf{y} \right\| = \dfrac{1}{\|\mathbf{y}\|} \cdot \|\mathbf{y}\| = 1$.
   It follows that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \dfrac{1}{\|\mathbf{y}\|}\mathbf{y}$ constitute an orthonormal set in $\mathbb{R}^n$.

(c) By definition, we have $\frac{1}{\|\mathbf{y}\|}\mathbf{y} = \frac{1}{\|\mathbf{y}\|}\mathbf{z} - \frac{\langle \mathbf{z}, \mathbf{u}_1 \rangle}{\|\mathbf{y}\|}\mathbf{u}_1 - \frac{\langle \mathbf{z}, \mathbf{u}_2 \rangle}{\|\mathbf{y}\|}\mathbf{u}_2 - \cdots - \frac{\langle \mathbf{z}, \mathbf{u}_k \rangle}{\|\mathbf{y}\|}\mathbf{u}_k$.

Then each of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \frac{1}{\|\mathbf{y}\|}\mathbf{y}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{z}$.

We also have $\mathbf{z} = \|\mathbf{y}\| \cdot (\frac{1}{\|y\|}\mathbf{y}) + \langle \mathbf{z}, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{z}, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{z}, \mathbf{u}_k \rangle \mathbf{u}_k$.

Then each of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{z}$ is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \frac{1}{\|\mathbf{y}\|}\mathbf{y}$.

It follows that $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \mathbf{z}\}) = \mathsf{Span}\ \left(\left\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k, \frac{1}{\|\mathbf{y}\|}\mathbf{y}\right\}\right)$.

4. **Theorem (H). (Existence of orthonormal basis.)**

Suppose W is a non-zero subspace of $\mathbb{R}^n$. Then W has an orthonormal basis.

Remark. The constructive argument in the proof below, generating an orthonormal basis for W from an (arbitrary) basis for W, is referred to as the Gram-Schmidt orthogonalization process.

5. **Proof of Theorem (H).**

Suppose W is a non-zero subspace of $\mathbb{R}^n$. Write $\dim(W) = k$. By assumption, k is between 1 and n.

Pick some basis for W, which is a collection of k vectors, denoted by $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_k$.

For each $j = 1, 2, \cdots, k$, define $W_j = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_j\})$. Note that $\dim(W_j) = j$, and by definition, $\mathbf{z}_{j+1}$ does not belong to W_j.

[So now we have a sequence of subspaces of W, namely,

$$W_1, W_2, W_3, \cdots, W_{k-2}, W_{k-1}, W_k,$$

in which $W_{\ell-1}$ is a 'proper' subspace of W_ℓ, in view of $\mathbf{z}_\ell \notin W_{\ell-1}$ and $z_\ell \in W_\ell$.]

(a) Note that $\mathbf{z}_1 \neq \mathbf{0}$. Then $\|\mathbf{z}_1\| \neq 0$. (Take $\mathbf{y}_1 = \mathbf{z}_1$.)
Define $\mathbf{u}_1 = \frac{1}{\|\mathbf{z}_1\|}\mathbf{z}_1$.
We have $\|\mathbf{u}_1\| = 1$.
$\mathbf{u}_1$ and $\mathbf{z}_1$ are non-zero scalar multiples of each other.
Then $W_1 = \mathsf{Span}\ (\{\mathbf{z}_1\}) = \mathsf{Span}\ (\{\mathbf{u}_1\})$.
Therefore $\mathbf{u}_1$ constitutes an orthonormal basis for W_1.

(b) $\mathbf{z}_2$ does not belong to W_1. Then $\mathbf{z}_2$ is not a linear combination of $\mathbf{u}_1$.
Define $\mathbf{y}_2 = \mathbf{z}_2 - \langle \mathbf{z}_2, \mathbf{u}_1 \rangle \mathbf{u}_1$.
By Lemma (G), $\|\mathbf{y}_2\| \neq 0$.
Define $\mathbf{u}_2 = \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2$.
By Lemma (G), $\mathbf{u}_1, \mathbf{u}_2$ constitute an orthonormal set in $\mathbb{R}^n$.
Since $\mathsf{Span}\ (\{\mathbf{z}_1\}) = \mathsf{Span}\ (\{\mathbf{u}_1\})$, we have $W_2 = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2\}) = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{z}_2\})$.
Then, again by Lemma (G), $W_2 = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{z}_2\}) = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\})$.
Therefore $\mathbf{u}_1, \mathbf{u}_2$ constitutes an orthonormal basis for W_2.

(c) $\mathbf{z}_3$ does not belong to W_2. Then $\mathbf{z}_3$ is not a linear combination of $\mathbf{u}_1, \mathbf{u}_2$.
Define $\mathbf{y}_3 = \mathbf{z}_3 - \langle \mathbf{z}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{z}_3, \mathbf{u}_2 \rangle \mathbf{u}_2$.
By Lemma (G), $\|\mathbf{y}_3\| \neq 0$.
Define $\mathbf{u}_3 = \frac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3$.
By Lemma (G), $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitute an orthonormal set in $\mathbb{R}^n$.
Since $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2\}) = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\})$, we have $W_3 = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}) = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{z}_3\})$.
Then, again by Lemma (G),

$$W_3 = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{z}_3\}) = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}).$$

Therefore $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitutes an orthonormal basis for W_3.

(d) Let ℓ be any one of $2, 3, \cdots, k$. Suppose that the vectors $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_{\ell-1}$ and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}$ are successively defined by $\mathbf{y}_1 = \mathbf{z}_1$, $\mathbf{u}_1 = \dfrac{1}{\|\mathbf{z}_1\|}\mathbf{z}_1$ and

$$\begin{cases} \mathbf{y}_2 &= \mathbf{z}_2 - \langle \mathbf{z}_2, \mathbf{u}_1 \rangle \mathbf{u}_1, \\ \mathbf{u}_2 &= \dfrac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2, \\ \mathbf{y}_3 &= \mathbf{z}_3 - \langle \mathbf{z}_3, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{z}_3, \mathbf{u}_2 \rangle \mathbf{u}_2, \\ \mathbf{u}_3 &= \dfrac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3, \\ &\vdots \\ \mathbf{y}_{\ell-1} &= \mathbf{z}_{\ell-1} - \langle \mathbf{z}_{\ell-1}, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{z}_{\ell-1}, \mathbf{u}_2 \rangle \mathbf{u}_2 - \cdots - \langle \mathbf{z}_{\ell-1}, \mathbf{u}_{\ell-2} \rangle \mathbf{u}_{\ell-2}, \\ \mathbf{u}_{\ell-1} &= \dfrac{1}{\|\mathbf{y}_{\ell-1}\|}\mathbf{y}_{\ell-1}, \end{cases}$$

and satisfies:

- $\|\mathbf{y}_2\| \neq 0$, $\|\mathbf{y}_3\| \neq 0$, ..., $\|\mathbf{y}_{\ell-1}\| \neq 0$, and
- for each $j = 2, 3, \cdots, \ell-1$, the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_j$ constitute an orthonormal basis for for W_j.

We now note that $\mathbf{z}_\ell$ does not belong to $W_{\ell-1}$. Then $\mathbf{z}_\ell$ is not a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}$.
Define $\mathbf{y}_\ell = \mathbf{z}_\ell - \langle \mathbf{z}_\ell, \mathbf{u}_1 \rangle \mathbf{u}_1 - \langle \mathbf{z}_\ell, \mathbf{u}_2 \rangle \mathbf{u}_2 - \cdots \langle \mathbf{z}_\ell, \mathbf{u}_{\ell-1} \rangle \mathbf{u}_{\ell-1} - \langle \mathbf{z}_\ell, \mathbf{u}_2 \rangle \mathbf{u}_2$.
By Lemma (G), $\|\mathbf{y}_\ell\| \neq 0$.
Define $\mathbf{u}_\ell = \dfrac{1}{\|\mathbf{y}_\ell\|}\mathbf{y}_\ell$.
By Lemma (G), $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}, \mathbf{u}_\ell$ constitute an orthonormal set in $\mathbb{R}^n$.
Since $\mathsf{Span}\,(\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{\ell-1}\}) = \mathsf{Span}\,(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}\})$, we have $W_\ell = \mathsf{Span}\,(\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{\ell-1}, \mathbf{z}_\ell\}) = \mathsf{Span}\,(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}, \mathbf{z}_\ell\})$.
Then, again by Lemma (G), $W_\ell = \mathsf{Span}\,(\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{\ell-1}, \mathbf{z}_\ell\}) = \mathsf{Span}\,(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}, \mathbf{u}_\ell\})$.
Therefore $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{\ell-1}, \mathbf{u}_\ell$ constitutes an orthonormal basis for W_ℓ.

Hence W has an orthonormal basis, namely $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$.

6. **Gram-Schmidt orthogonalization process.**

Suppose W is a subspace of $\mathbb{R}^n$, and $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \cdots, \mathbf{z}_k$ constitute a basis for W.

The argument in the proof of Theorem (H) provides an algorithm for obtaining an orthonormal basis $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ for W, for which the equality $\mathsf{Span}\,(\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_j\}) = \mathsf{Span}\,(\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_j\})$ holds for each $j = 1, 2, \cdots, k$:

- **Step (1).**
  We define $\mathbf{y}_1 = \mathbf{z}_1$.
- **Step (2).**
  We define $\mathbf{y}_2, \mathbf{y}_3, \cdots, \mathbf{y}_k$ inductively by

$$\mathbf{y}_j = \mathbf{z}_j - \frac{\langle \mathbf{z}_j, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{\langle \mathbf{z}_j, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 - \cdots - \frac{\langle \mathbf{z}_j, \mathbf{y}_{j-1} \rangle}{\|\mathbf{y}_{j-1}\|^2}\mathbf{y}_{j-1} \quad \text{for each } j = 2, 3, \cdots, k.$$

  When written out explicitly, $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_k$ are given recursively by:

$$\begin{cases} \mathbf{y}_1 &= \mathbf{z}_1 \\ \mathbf{y}_2 &= \mathbf{z}_2 - \dfrac{\langle \mathbf{z}_2, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 \\ \mathbf{y}_3 &= \mathbf{z}_3 - \dfrac{\langle \mathbf{z}_3, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\langle \mathbf{z}_3, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 \\ \mathbf{y}_4 &= \mathbf{z}_4 - \dfrac{\langle \mathbf{z}_4, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\langle \mathbf{z}_4, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 - \dfrac{\langle \mathbf{z}_4, \mathbf{y}_3 \rangle}{\|\mathbf{y}_3\|^2}\mathbf{y}_3 \\ &\vdots \\ \mathbf{y}_k &= \mathbf{z}_k - \dfrac{\langle \mathbf{z}_k, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\langle \mathbf{z}_k, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 - \dfrac{\langle \mathbf{z}_k, \mathbf{y}_3 \rangle}{\|\mathbf{y}_3\|^2}\mathbf{y}_3 - \cdots - \dfrac{\langle \mathbf{z}_k, \mathbf{y}_{k-1} \rangle}{\|\mathbf{y}_{k-1}\|^2}\mathbf{y}_{k-1} \end{cases}$$

- **Step (3).**

  For each $j = 1, 2, \cdots, k$, define $\mathbf{u}_j = \dfrac{1}{\|\mathbf{y}_j\|}\mathbf{y}_j$.

For each $\ell = 1, 2, \cdots, k$, the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_\ell$ constitute an orthonormal basis for $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_\ell\})$.

In particular, the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ constitute an orthonormal basis for W.

7. **Illustrations on the Gram-Schmidt orthogonalization process.**

(a) Let $\mathbf{z}_1 = \begin{bmatrix}1\\2\\2\end{bmatrix}$, $\mathbf{z}_2 = \begin{bmatrix}-1\\1\\4\end{bmatrix}$, $\mathbf{z}_3 = \begin{bmatrix}-1\\-2\\1\end{bmatrix}$.

Take for granted that $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ are linearly independent.

We proceed to find an orthonormal basis for $W = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\}) = \mathbb{R}^3$.

- Take $\mathbf{y}_1 = \mathbf{z}_1$. Then $\mathbf{y}_1 = \begin{bmatrix}1\\2\\2\end{bmatrix}$, and $\|\mathbf{y}_1\|^2 = 9$.

  Take $\mathbf{u}_1 = \dfrac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1$. Then $\mathbf{u}_1 = \begin{bmatrix}1/3\\2/3\\2/3\end{bmatrix}$.

- Take $\mathbf{y}_2 = \mathbf{z}_2 - \dfrac{\langle \mathbf{z}_2, \mathbf{y}_1\rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1$.

  We have $\langle \mathbf{z}_2, \mathbf{y}_1\rangle = 9$.

  Then $\mathbf{y}_2 = \begin{bmatrix}-1\\1\\4\end{bmatrix} - \dfrac{9}{9}\begin{bmatrix}1\\2\\2\end{bmatrix} = \begin{bmatrix}-2\\-1\\2\end{bmatrix}$, and $\|\mathbf{y}_2\|^2 = 9$.

  Take $\mathbf{u}_2 = \dfrac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2$. Then $\mathbf{u}_2 = \begin{bmatrix}-2/3\\-1/3\\2/3\end{bmatrix}$.

- Take $\mathbf{y}_3 = \mathbf{z}_3 - \dfrac{\langle \mathbf{z}_3, \mathbf{y}_1\rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\langle \mathbf{z}_3, \mathbf{y}_2\rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$.

  We have $\langle \mathbf{z}_3, \mathbf{y}_1\rangle = -3$, $\langle \mathbf{z}_3, \mathbf{y}_2\rangle = 6$.

  Then $\mathbf{y}_3 = \begin{bmatrix}-1\\-2\\1\end{bmatrix} - \dfrac{-3}{9}\begin{bmatrix}1\\2\\2\end{bmatrix} - \dfrac{6}{9}\begin{bmatrix}-2\\-1\\2\end{bmatrix} = \begin{bmatrix}2/3\\-2/3\\1/3\end{bmatrix}$, and $\|\mathbf{y}_3\|^2 = 1$.

  Take $\mathbf{u}_3 = \dfrac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3$. Then $\mathbf{u}_3 = \begin{bmatrix}2/3\\-2/3\\1/3\end{bmatrix}$.

$\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitute an orthonormal basis for W.

Also note that, by construction, $\mathsf{Span}\ (\{\mathbf{u}_1\}) = \mathsf{Span}\ (\{\mathbf{z}_1\})$ and $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\}) = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2\})$.

(b) Let $\mathbf{z}_1 = \begin{bmatrix}1\\0\\1\\0\end{bmatrix}$, $\mathbf{z}_2 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$, $\mathbf{z}_3 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}$.

Take for granted that $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ are linearly independent.

We proceed to find an orthonormal basis for $W = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\})$.

- Take $\mathbf{y}_1 = \mathbf{z}_1$. Then $\mathbf{y}_1 = \begin{bmatrix}1\\0\\1\\0\end{bmatrix}$, and $\|\mathbf{y}_1\|^2 = 2$.

  Take $\mathbf{u}_1 = \dfrac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1$. Then $\mathbf{u}_1 = \begin{bmatrix}1/\sqrt{2}\\0\\1/\sqrt{2}\\0\end{bmatrix}$.

- Take $\mathbf{y}_2 = \mathbf{z}_2 - \dfrac{\langle \mathbf{z}_2, \mathbf{y}_1\rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1$.

We have $\langle \mathbf{z}_2, \mathbf{y}_1 \rangle = 2$.

Then $\mathbf{y}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, and $\|\mathbf{y}_2\|^2 = 2$.

Take $\mathbf{u}_2 = \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2$. Then $\mathbf{u}_2 = \begin{bmatrix} 0 \\ 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$.

- Take $\mathbf{y}_3 = \mathbf{z}_3 - \frac{\langle \mathbf{z}_3, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{\langle \mathbf{z}_3, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$.

  We have $\langle \mathbf{z}_3, \mathbf{y}_1 \rangle = 1$, $\langle \mathbf{z}_3, \mathbf{y}_2 \rangle = 2$.

  Then $\mathbf{y}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/2 \\ 0 \\ 1/2 \\ 0 \end{bmatrix}$, and $\|\mathbf{y}_3\|^2 = \frac{1}{2}$.

  Take $\mathbf{u}_3 = \frac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3$. Then $\mathbf{u}_3 = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$.

$\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitute an orthonormal basis for W.

Also note that, by construction, $\mathsf{Span}\ (\{\mathbf{u}_1\}) = \mathsf{Span}\ (\{\mathbf{z}_1\})$ and $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\}) = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2\})$.

(c) Let $\mathbf{z}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{z}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix}$, $\mathbf{z}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

Take for granted that $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3$ are linearly independent.

We proceed to find an orthonormal basis for $W = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\})$.

- Take $\mathbf{y}_1 = \mathbf{z}_1$. Then $\mathbf{y}_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$, and $\|\mathbf{y}_1\|^2 = 2$.

  Take $\mathbf{u}_1 = \frac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1$. Then $\mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \\ 0 \end{bmatrix}$.

- Take $\mathbf{y}_2 = \mathbf{z}_2 - \frac{\langle \mathbf{z}_2, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1$.

  We have $\langle \mathbf{z}_2, \mathbf{y}_1 \rangle = 2$.

  Then $\mathbf{y}_2 = \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$, and $\|\mathbf{y}_2\|^2 = 4$.

  Take $\mathbf{u}_2 = \frac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2$. Then $\mathbf{u}_2 = \begin{bmatrix} -1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{bmatrix}$.

- Take $\mathbf{y}_3 = \mathbf{z}_3 - \frac{\langle \mathbf{z}_3, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{\langle \mathbf{z}_3, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$.

  We have $\langle \mathbf{z}_3, \mathbf{y}_1 \rangle = 2$, $\langle \mathbf{z}_3, \mathbf{y}_2 \rangle = 2$.

  Then $\mathbf{y}_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \frac{2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \frac{2}{4} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$, and $\|\mathbf{y}_3\|^2 = 1$.

Take $\mathbf{u}_3 = \dfrac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3$. Then $\mathbf{u}_3 = \begin{bmatrix} 1/2 \\ 1/2 \\ -1/2 \\ 1/2 \end{bmatrix}$.

$\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ constitute an orthonormal basis for W.

Also note that, by construction, $\mathsf{Span}\ (\{\mathbf{u}_1\}) = \mathsf{Span}\ (\{\mathbf{z}_1\})$ and $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\}) = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2\})$.

(d) Let $\mathbf{z}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 4 \end{bmatrix}$, $\mathbf{z}_2 = \begin{bmatrix} -2 \\ 6 \\ 2 \\ 9 \end{bmatrix}$, $\mathbf{z}_3 = \begin{bmatrix} 9 \\ -2 \\ -4 \\ 7 \end{bmatrix}$, $\mathbf{z}_4 = \begin{bmatrix} -3 \\ -1 \\ -3 \\ 9 \end{bmatrix}$.

Take for granted that $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4$ are linearly independent.

We proceed to find an orthonormal basis for $W = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \mathbf{z}_4\}) = \mathbb{R}^4$.

- Take $\mathbf{y}_1 = \mathbf{z}_1$. Then $\mathbf{y}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ 4 \end{bmatrix}$, and $\|\mathbf{y}_1\|^2 = 25$.

  Take $\mathbf{u}_1 = \dfrac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1$. Then $\mathbf{u}_1 = \begin{bmatrix} 1/5 \\ 2/5 \\ 2/5 \\ 4/5 \end{bmatrix}$.

- Take $\mathbf{y}_2 = \mathbf{z}_2 - \dfrac{\langle \mathbf{z}_2, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1$.

  We have $\langle \mathbf{z}_2, \mathbf{y}_1 \rangle = 50$.

  Then $\mathbf{y}_2 = \begin{bmatrix} -2 \\ 6 \\ 2 \\ 9 \end{bmatrix} - \dfrac{50}{25}\begin{bmatrix} 1 \\ 2 \\ 2 \\ 4 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ -2 \\ 1 \end{bmatrix}$, and $\|\mathbf{y}_2\|^2 = 25$.

  Take $\mathbf{u}_2 = \dfrac{1}{\|\mathbf{y}_2\|}\mathbf{y}_2$. Then $\mathbf{u}_2 = \begin{bmatrix} -4/5 \\ 2/5 \\ -2/5 \\ 1/5 \end{bmatrix}$.

- Take $\mathbf{y}_3 = \mathbf{z}_3 - \dfrac{\langle \mathbf{z}_3, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\langle \mathbf{z}_3, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2$.

  We have $\langle \mathbf{z}_3, \mathbf{y}_1 \rangle = 25$, $\langle \mathbf{z}_3, \mathbf{y}_2 \rangle = -25$.

  Then $\mathbf{y}_3 = \begin{bmatrix} 9 \\ -2 \\ -4 \\ 7 \end{bmatrix} - \dfrac{25}{25}\begin{bmatrix} 1 \\ 2 \\ 2 \\ 4 \end{bmatrix} - \dfrac{-25}{25}\begin{bmatrix} -4 \\ 2 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ -8 \\ 4 \end{bmatrix}$, and $\|\mathbf{y}_3\|^2 = 100$.

  Take $\mathbf{u}_3 = \dfrac{1}{\|\mathbf{y}_3\|}\mathbf{y}_3$. Then $\mathbf{u}_3 = \begin{bmatrix} 2/5 \\ -1/5 \\ -4/5 \\ 2/5 \end{bmatrix}$.

- Take $\mathbf{y}_4 = \mathbf{z}_4 - \dfrac{\langle \mathbf{z}_4, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \dfrac{\langle \mathbf{z}_4, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 - \dfrac{\langle \mathbf{z}_4, \mathbf{y}_3 \rangle}{\|\mathbf{y}_3\|^2}\mathbf{y}_3$.

  We have $\langle \mathbf{z}_4, \mathbf{y}_1 \rangle = 25$, $\langle \mathbf{z}_4, \mathbf{y}_2 \rangle = 25$, $\langle \mathbf{z}_4, \mathbf{y}_3 \rangle = 50$.

  Then $\mathbf{y}_4 = \begin{bmatrix} -3 \\ -1 \\ -3 \\ 9 \end{bmatrix} - \dfrac{25}{25}\begin{bmatrix} 1 \\ 2 \\ 2 \\ 4 \end{bmatrix} - \dfrac{25}{25}\begin{bmatrix} -4 \\ 2 \\ -2 \\ 1 \end{bmatrix} - \dfrac{50}{100}\begin{bmatrix} 4 \\ -2 \\ -8 \\ 4 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 1 \\ 2 \end{bmatrix}$, and $\|\mathbf{y}_4\|^2 = 25$.

  Take $\mathbf{u}_4 = \dfrac{1}{\|\mathbf{y}_4\|}\mathbf{y}_4$. Then $\mathbf{u}_4 = \begin{bmatrix} -2/5 \\ -4/5 \\ 1/5 \\ 2/5 \end{bmatrix}$.

$\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4$ constitute an orthonormal basis for W.

Also note that, by construction, $\mathsf{Span}\ (\{\mathbf{u}_1\}) = \mathsf{Span}\ (\{\mathbf{z}_1\})$, $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2\}) = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2\})$and $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}) = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3\})$.

8. **Gram-Schmidt orthogonalization, presented as QR-decomposition.**

Suppose Z is an $(n \times k)$-matrix, with $n \geq k$. For each $j = 1, 2, \cdots, k$, the j-th column of Z by $\mathbf{z}_j$ for.

Suppose $\mathbf{z}_1, \mathbf{z}_2, \mathbf{z}_3, \cdots, \mathbf{z}_k$ are linearly independent.

We define $\mathbf{y}_1 = \mathbf{z}_1$, $\mathbf{u}_1 = \dfrac{1}{\|\mathbf{y}_1\|}\mathbf{y}_1$, and define $\mathbf{y}_2, \mathbf{y}_3, \cdots, \mathbf{y}_k$ inductively by

$$\mathbf{y}_j = \mathbf{z}_j - \frac{\langle \mathbf{z}_j, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 - \frac{\langle \mathbf{z}_j, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 - \cdots - \frac{\langle \mathbf{z}_j, \mathbf{y}_{j-1} \rangle}{\|\mathbf{y}_{j-1}\|^2}\mathbf{y}_{j-1}, \quad \mathbf{u}_j = \frac{1}{\|\mathbf{y}_j\|}\mathbf{y}_j \quad \text{for each } j = 2, 3, \cdots, k.$$

According to the argument for Theorem (H), these vectors $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \cdots, \mathbf{y}_k$ are well-defined

Define the $(n \times k)$-matrix $Q = \left[\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_k\ \right]$.

For each $j = 1, 2, \cdots, k$, we have

$$\begin{aligned}
\mathbf{z}_j &= \frac{\langle \mathbf{z}_j, \mathbf{y}_1 \rangle}{\|\mathbf{y}_1\|^2}\mathbf{y}_1 + \frac{\langle \mathbf{z}_j, \mathbf{y}_2 \rangle}{\|\mathbf{y}_2\|^2}\mathbf{y}_2 + \cdots + \frac{\langle \mathbf{z}_j, \mathbf{y}_{j-1} \rangle}{\|\mathbf{y}_{j-1}\|^2}\mathbf{y}_{j-1} + \mathbf{y}_j \\
&= \langle \mathbf{z}_j, \mathbf{u}_1 \rangle \mathbf{u}_1 + \langle \mathbf{z}_j, \mathbf{u}_2 \rangle \mathbf{u}_2 + \cdots + \langle \mathbf{z}_j, \mathbf{u}_{j-1} \rangle \mathbf{u}_{j-1} + \|\mathbf{y}_j\| \cdot \mathbf{u}_j + 0 \cdot \mathbf{u}_{j+1} + \cdots + 0 \cdot \mathbf{u}_k = Q \begin{bmatrix} \langle \mathbf{u}_1, \mathbf{z}_j \rangle \\ \langle \mathbf{u}_2, \mathbf{z}_j \rangle \\ \vdots \\ \langle \mathbf{u}_{j-1}, \mathbf{z}_j \rangle \\ \|\mathbf{y}_j\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}.
\end{aligned}$$

Define the $(n \times n)$-square matrix R, whose (i, j)-th entry is denoted by r_{ij} and given by

$$r_{ij} = \begin{cases} \langle \mathbf{u}_i, \mathbf{z}_j \rangle & \text{if } \ i < j \\ \|\mathbf{y}_j\| & \text{if } \ i = j \\ 0 & \text{if } \ i > j \end{cases}$$

(So, for each $j = 1, 2, \cdots, n$, the j-th column of R is $\begin{bmatrix} \langle \mathbf{u}_1, \mathbf{z}_j \rangle \\ \langle \mathbf{u}_2, \mathbf{z}_j \rangle \\ \vdots \\ \langle \mathbf{u}_{j-1}, \mathbf{z}_j \rangle \\ \|\mathbf{y}_j\| \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.)

Then $Z = QR$.

This 'factorization' of Z into the product QR is called the 'QR-decomposition' for Z.

Note that $\mathcal{C}(Z) = \mathcal{C}(Q)$ and the columns of Q is an orthonormal basis for $\mathcal{C}(Z)$.

The matrix R encodes the Gram-Schmidt orthogonalization process from which we obtain the orthonormal set $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_k$ from the linearly independent set $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_k$.

9. **Illustrations of QR-decomposition.**

Refer to *Illustrations on the Gram-Schmidt orthogonalization process* above. The respective constructions can be displayed as the 'factorizations' below:

(a) $$\begin{bmatrix} 1 & -1 & -1 \\ 2 & 1 & -2 \\ 2 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} 3 & 3 & -1 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

(b) $$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} & 1/\sqrt{2} \\ 0 & \sqrt{2} & \sqrt{2} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}.$$

(c) $$\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/2 & -1/2 \\ 0 & 1/2 & 1/2 \\ 1/\sqrt{2} & 1/2 & -1/2 \\ 0 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2} & \sqrt{2} \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

(d) $$\begin{bmatrix} 1 & -2 & 9 & -3 \\ 2 & 6 & -2 & -1 \\ 2 & 2 & -4 & -3 \\ 4 & 9 & 7 & 9 \end{bmatrix} = \begin{bmatrix} 1/5 & -4/5 & 2/5 & -2/5 \\ 2/5 & 2/5 & -1/5 & -4/5 \\ 2/5 & -2/5 & -4/5 & 1/5 \\ 4/5 & 1/5 & 2/5 & 2/5 \end{bmatrix} \begin{bmatrix} 5 & 10 & 5 & 5 \\ 0 & 5 & -5 & 5 \\ 0 & 0 & 10 & 5 \\ 0 & 0 & 0 & 5 \end{bmatrix}.$$