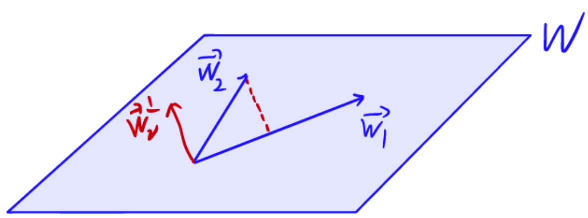
• Instructor: He Wang Email: he.wang@northeastern.edu

§5.2 Gram-Schmidt Process and QR Factorization

Example 1. Find an orthogonal basis for the subspace

$$W = \text{Span}\left\{ \vec{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{w}_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} \right\}$$



$$\text{proj}_{\vec{w}_1} \vec{w}_2 = \left(\frac{\vec{w}_2 \cdot \vec{w}_1}{\vec{w}_1 \cdot \vec{w}_1} \right) \vec{w}_1 = \left(\frac{32}{14} \right) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 16/7 \\ 32/7 \\ 48/7 \end{bmatrix}$$

$$\vec{w}_2^{\perp} = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} - \begin{bmatrix} 16/7 \\ 32/7 \\ 48/7 \end{bmatrix} = \begin{bmatrix} 12/7 \\ 3/7 \\ -6/7 \end{bmatrix}$$

$\{\vec{w}_1, \vec{w}_2^{\perp}\}$ is an orthogonal basis for W.

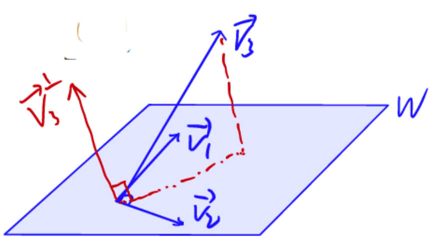
Example 2. Find an orthogonal basis for the subspace

$$V = \text{Span}\left\{ \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \\ 0 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 6 \\ 0 \end{bmatrix} \right\}$$

$\vec{v}_1 \cdot \vec{v}_2 = 0$

From Example 5 in §5.1

$$\vec{v}_3^{\perp} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 0 \end{bmatrix}$$



So $\{\vec{v}_1, \vec{v}_2, \vec{v}_3^{\perp}\}$ is an orthogonal basis for V.

The **Gram-Schmidt process** is an algorithm that produces an orthogonal (or orthonormal) basis for any subspace W of $\mathbb{R}^n$ by starting with any basis for W.

Theorem. [Gram-Schmidt (Orthogonalize)]

Let W be a subspace of $\mathbb{R}^n$ and let $\vec{b}_1, \cdots, \vec{b}_p$ be a basis for W. Define vectors $\vec{v}_1, \ldots, \vec{v}_p$ as

$$\begin{aligned}
\vec{v}_1 &= \vec{b}_1 \\
\vec{v}_2 &= \vec{b}_2 - \frac{\vec{b}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 \\
\vec{v}_3 &= \vec{b}_3 - \frac{\vec{b}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{b}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 \\
&\vdots \\
\vec{v}_p &= \vec{b}_p - \frac{\vec{b}_p \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\vec{v}_1 - \frac{\vec{b}_p \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\vec{v}_2 - \cdots - \frac{\vec{b}_p \cdot \vec{v}_{p-1}}{\vec{v}_{p-1} \cdot \vec{v}_{p-1}}\vec{v}_{p-1}
\end{aligned}$$

Then $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W and

$$\text{Span}\{\vec{b}_1, \cdots, \vec{b}_k\} = \text{Span}\{\vec{v}_1, \cdots, \vec{v}_k\}$$

for and $k = 1, \ldots, p$.

From the formulas we can see that $\vec{v}_2 = \vec{b}_2^{\perp}$ relative to $\vec{b}_1$.

$\vec{v}_3 = \vec{b}_3^{\perp}$ relative to Span$(\vec{b}_1, \vec{b}_2)$.

$\vec{v}_i = \vec{b}_i^{\perp}$ relative to Span$(\vec{b}_1, \vec{b}_2, ..., \vec{b}_{i-1})$.

Theorem. [Gram-Schmidt (Normalize)]

If $\{\vec{v}_1, \ldots, \vec{v}_p\}$ is an orthogonal basis for W, then $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is an orthonormal basis for W, where, $\vec{u}_i = \dfrac{\vec{v}_i}{||\vec{v}_i||}$ for $i = 1, \ldots, p$.

$$\boxed{\text{Basis} \xrightarrow{\text{orthogonalize}} \text{Orthogonal basis} \xrightarrow{\text{normalize}} \text{Orthonormal basis.}}$$

Example 3. Using Gram-Schmidt Process to find an orthonormal basis for

$$V = \text{Span}\left\{\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}\right\}.$$

$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is already a basis for V since it is independent.

$$\vec{b}_1 = \vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

$$\vec{b}_2 = \vec{v}_2 - \left(\frac{\vec{v}_2 \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1}\right)\vec{b}_1 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} - \frac{18}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

$$\vec{b}_3 = \vec{v}_3 - \left(\frac{\vec{v}_3 \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1}\right)\vec{b}_1 - \left(\frac{\vec{v}_3 \cdot \vec{b}_2}{\vec{b}_2 \cdot \vec{b}_2}\right)\vec{b}_2$$

$$= \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix} - \frac{15}{6}\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$$

$\{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ is an orthogonal basis for V.

$$\vec{u}_1 = \frac{\vec{b}_1}{\|\vec{b}_1\|} = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix} \qquad \vec{u}_2 = \frac{\vec{b}_2}{\|\vec{b}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \qquad \vec{u}_3 = \frac{\vec{b}_3}{\|\vec{b}_3\|} = \begin{bmatrix} -1/\sqrt{3} \\ -1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis for V.

Since $\vec{v}_i$ is already used in the question, in the solution, we changed the formulas. Another (better) way is rewrite the letters in the question as $\vec{b}_i$.

Let us rewrite the solution of Example 3 here.

Rewrite Example 3 Using Gram-Schmidt Process to find an orthonormal basis for

$$V = \text{Span}\left\{\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \vec{b}_3 = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}\right\}.$$

$$\vec{v}_1 = \vec{b}_1 = \begin{bmatrix}1\\1\\2\end{bmatrix}$$

$$\vec{v}_2 = \vec{v}_2 - \left(\frac{\vec{b}_2\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\right)\vec{v}_1 = \begin{bmatrix}2\\4\\6\end{bmatrix} - \frac{18}{6}\begin{bmatrix}1\\1\\2\end{bmatrix} = \begin{bmatrix}-1\\1\\0\end{bmatrix}$$

$$\vec{v}_3 = \vec{b}_3 - \left(\frac{\vec{b}_3\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\right)\vec{v}_1 - \left(\frac{\vec{b}_3\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\right)\vec{v}_2 = \begin{bmatrix}1\\2\\6\end{bmatrix} - \frac{15}{6}\begin{bmatrix}1\\1\\2\end{bmatrix} - \frac{1}{2}\begin{bmatrix}-1\\1\\0\end{bmatrix} = \begin{bmatrix}-1\\-1\\1\end{bmatrix}$$

$\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is an orthogonal basis for V.

$$\vec{u}_1 = \frac{\vec{v}_1}{||\vec{v}_1||} = \begin{bmatrix}1/\sqrt{6}\\1/\sqrt{6}\\2/\sqrt{6}\end{bmatrix}; \ \vec{u}_2 = \frac{\vec{v}_2}{||\vec{v}_2||} = \begin{bmatrix}-1/\sqrt{2}\\1/\sqrt{2}\\0\end{bmatrix}; \ \vec{u}_3 = \frac{\vec{v}_3}{||\vec{v}_3||} = \begin{bmatrix}-1/\sqrt{3}\\-1/\sqrt{3}\\1/\sqrt{3}\end{bmatrix}$$

$\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ is an orthonormal basis for V.

Example 4. Using Gram-Schmidt Process to find an orthonormal basis for

$$V = \text{Span}\left\{\vec{b}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}, \vec{b}_2 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix}\right\}.$$

$$\vec{v}_1 = \vec{b}_1 = \begin{bmatrix}1\\1\\1\\1\end{bmatrix}$$

$$\vec{v}_2 = \vec{b}_2 - \left(\frac{\vec{b}_2\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\right)\vec{v}_1 = \begin{bmatrix}0\\1\\1\\1\end{bmatrix} - \frac{3}{4}\begin{bmatrix}1\\1\\1\\1\end{bmatrix} = \begin{bmatrix}-3/4\\1/4\\1/4\\1/4\end{bmatrix}$$

$\{\vec{v}_1, \vec{v}_2\}$ is an orthogonal basis for V.

$$\vec{u}_1 = \frac{\vec{v}_1}{||\vec{v}_1||} = \begin{bmatrix}1/2\\1/2\\1/2\\1/2\end{bmatrix}; \ \vec{u}_2 = \frac{\vec{v}_2}{||\vec{v}_2||} = \begin{bmatrix}-3/\sqrt{12}\\1/\sqrt{12}\\1/\sqrt{12}\\1/\sqrt{12}\end{bmatrix};$$

$\{\vec{u}_1, \vec{u}_2\}$ is an orthonormal basis for V.

- Note that the formula for computing $\vec{v}_i$ for any $i = 2, 3, \ldots, p$ can be written as

$$\begin{aligned}\vec{v}_i &= \vec{b}_i - \text{proj}_{\vec{v}_1}(\vec{b}_i) - \text{proj}_{\vec{v}_2}(\vec{b}_i) - \cdots - \text{proj}_{\vec{v}_{i-1}}(\vec{b}_i)\\ &= \vec{b}_i - \text{proj}_{\text{Span}\{\vec{v}_1,\ldots,\vec{v}_{i-1}\}}(\vec{b}_i).\end{aligned}$$

  So, $\vec{v}_i = \vec{b}_i^{\perp}$ respect to $\text{Span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}\}$.

- This formula is *inductive* in that the computation of $\vec{v}_i$ relies on the vectors $\vec{v}_1, \ldots, \vec{v}_{i-1}$.

QR-Factorization.

QR-Factorization is the matrix version of the Gram-Schmidt process.

Recall the Gram-Schmidt process:

$$\text{Basis } \mathscr{B} = \{\vec{b}_1, \ldots, \vec{b}_p\} \xrightarrow{\text{orthogonalize}} \text{Orthogonal basis } \mathscr{V} = \{\vec{v}_1, \ldots, \vec{v}_p\} \xrightarrow{\text{normalize}} \text{Orthonormal basis } \mathscr{U} = \{\vec{u}_1, \ldots, \vec{u}_p\}.$$

Given a vector in W, let's compare their coordinates:

$$\begin{tikzcd}
{[\vec{x}]_{\mathscr{B}}} \arrow[dr, "M"] \arrow[dd, "R"'] & \\
& \vec{x} \\
{[\vec{x}]_{\mathscr{U}}} \arrow[ur, "Q"'] &
\end{tikzcd}$$

Each matrix defines an isomorphism. So, $M = QR$.

Here $M = [\vec{b}_1 \ \ldots \ \vec{b}_p]$ and $Q = [\vec{u}_1, \ldots, \vec{u}_p]$.

Theorem.

Given a $n \times p$ matrix $M = [\vec{b}_1 \ \ldots \ \vec{b}_p]$ with independent columns. There is a unique decomposition

$$M = QR$$

where, $Q = [\vec{u}_1, \ldots, \vec{u}_p]$ has orthonormal columns and R is an $p \times p$ upper triangular matrix with

$$r_{ii} = ||\vec{v}_i|| \text{ for } i = 1, \ldots, p \quad \text{and} \quad r_{ij} = \vec{u}_i \cdot \vec{b}_j \quad \text{for} \quad i < j.$$

In particular when $p = 3$, the upper triangular matrix R is

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{bmatrix} = \begin{bmatrix} ||\vec{v}_1|| & \vec{u}_1 \cdot \vec{b}_2 & \vec{u}_1 \cdot \vec{b}_3 \\ 0 & ||\vec{v}_2|| & \vec{u}_2 \cdot \vec{b}_3 \\ 0 & 0 & ||\vec{v}_3|| \end{bmatrix}$$

Proof(for $p = 3$): From Gram-Schmidt process, write $\vec{b}_i$ as linear combinations of $\vec{u}_i$.

$$\begin{aligned} \vec{b}_1 &= \vec{v}_1 = ||\vec{v}_1||\vec{u}_1 \\ \vec{b}_2 &= \vec{v}_2 + \frac{\vec{b}_2 \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 = ||\vec{v}_2||\vec{u}_2 + (\vec{b}_2 \cdot \vec{u}_1)\vec{u}_1 \\ \vec{b}_3 &= \vec{b}_3 - \frac{\vec{b}_3 \cdot \vec{v}_1}{||\vec{v}_1||^2}\vec{v}_1 - \frac{\vec{b}_3 \cdot \vec{v}_2}{||\vec{v}_2||^2}\vec{v}_2 = ||\vec{v}_3||\vec{u}_3 + (\vec{b}_3 \cdot \vec{u}_1)\vec{u}_1 + (\vec{b}_3 \cdot \vec{u}_2)\vec{u}_2 \end{aligned}$$

So,

$$[\vec{b}_1\ \vec{b}_2\ \vec{b}_3] = [\vec{u}_1\ \vec{u}_2\ \vec{u}_3] \begin{bmatrix} ||\vec{v}_1|| & \vec{u}_1 \cdot \vec{b}_2 & \vec{u}_1 \cdot \vec{b}_3 \\ 0 & ||\vec{v}_2|| & \vec{u}_2 \cdot \vec{b}_3 \\ 0 & 0 & ||\vec{v}_3|| \end{bmatrix}$$

Example 5. Find the QR-factorization of the matrix $M = \begin{bmatrix} 1 & 2 & 1 \\ 1 & 4 & 2 \\ 2 & 6 & 6 \end{bmatrix}$. (Use modified Example 3,)

$\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix}, \vec{b}_3 = \begin{bmatrix} 1 \\ 2 \\ 6 \end{bmatrix}$. and $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

From Example 3. we have $Q = \begin{bmatrix} \frac{1}{\sqrt{6}} & -1/\sqrt{2} & -1/\sqrt{3} \\ \frac{1}{\sqrt{6}} & 1/\sqrt{2} & -1/\sqrt{3} \\ \frac{2}{\sqrt{6}} & 0 & 1/\sqrt{3} \end{bmatrix}$

So $r_{11} = ||\vec{v}_1|| = \sqrt{6}$ $\quad r_{22} = ||\vec{v}_2|| = \sqrt{2}$ $\quad r_{33} = ||\vec{v}_3|| = \sqrt{3}$

$r_{12} = \vec{u}_1 \cdot \vec{b}_2 = 18/\sqrt{6}$ $\quad r_{13} = \vec{u}_1 \cdot \vec{b}_3 = 15/\sqrt{6}$ $\quad r_{23} = \vec{u}_2 \vec{b}_3 = 1/\sqrt{2}$

So $R = \begin{bmatrix} \sqrt{6} & 18/\sqrt{6} & 15/\sqrt{6} \\ 0 & \sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & \sqrt{3} \end{bmatrix}$

Example 6. Using Gram-Schmidt Process to find the QR-factorization of the matrix $M = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$.

From Example 4, $Q = \begin{bmatrix} \frac{1}{2} & \frac{-3}{2\sqrt{3}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} \\ \frac{1}{2} & \frac{1}{2\sqrt{3}} \end{bmatrix}$

$r_{11} = \|\vec{v}_1\| = \sqrt{4} = 2 \qquad r_{22} = \|\vec{v}_2^{\perp}\| = \sqrt{3}/2$

$r_{12} = \vec{u}_1 \cdot \vec{v}_2 = \frac{3}{2}$

So $R = \begin{bmatrix} 2 & 3/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}$