Gram-Schmidt Orthogonalization and the QR Factorization

Linear Algebra
MATH 2076

The Gram-Schmidt Orthogonalization Procedure

This is an algorithm to produce an orthonormal basis from a basis.

We start with a basis $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ for some vector space $\mathbb{W}$.

Then we construct an orthogonal basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ for $\mathbb{W}$ with certain nice properties.

Finally, we get an orthonormal basis $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ for $\mathbb{W}$.

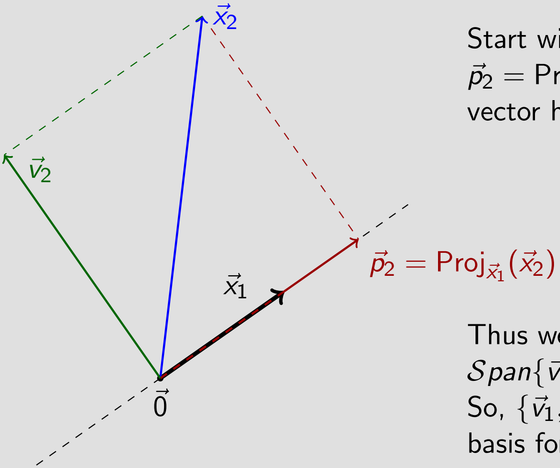
Normalization Step: For $1 \leq i \leq k$, $\vec{u}_i = \dfrac{\vec{v}_i}{\|\vec{v}_i\|}$.

First Step: $\vec{v}_1 = \vec{x}_1$.

How do we get $\vec{v}_2, \ldots, \vec{v}_k$?

Example with Basis $\{\vec{x}_1, \vec{x}_2\}$

Let $\{\vec{x}_1, \vec{x}_2\}$ be a basis for some 2-plane $\mathbb{W}$ (in some $\mathbb{R}^n$).



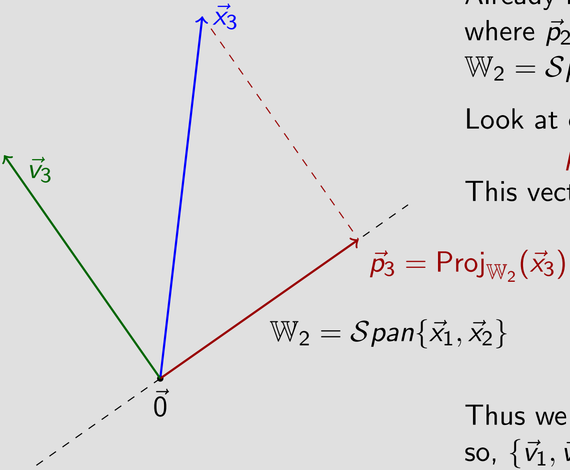
Start with orthogonal projection $\vec{p}_2 = \text{Proj}_{\vec{x}_1}(\vec{x}_2)$ of $\vec{x}_2$ onto $\vec{x}_1$. This vector has the property that

$$\vec{v}_2 = \vec{x}_2 - \vec{p}_2 \perp \vec{x}_1.$$

Thus we find that $\vec{v}_1 = \vec{x}_1 \perp \vec{v}_2$ and $\mathcal{S}pan\{\vec{v}_1, \vec{v}_2\} = \mathcal{S}pan\{\vec{x}_1, \vec{x}_2\} = \mathbb{W}$. So, $\{\vec{v}_1, \vec{v}_2\}$ is the desired orthogonal basis for $\mathbb{W}$.

Example with Basis $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$

Let $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ be a basis for some 3-plane $\mathbb{W}$ (in some $\mathbb{R}^n$).



Already have $\vec{v}_1 = \vec{x}_1$ and $\vec{v}_2 = \vec{x}_2 - \vec{p}_2$ where $\vec{p}_2 = \text{Proj}_{\vec{x}_1}(\vec{x}_2)$. Also know that $\mathbb{W}_2 = \mathcal{S}pan\{\vec{x}_1, \vec{x}_2\} = \mathcal{S}pan\{\vec{v}_1, \vec{v}_2\}$.

Look at orthogonal projection

$$\vec{p}_3 = \text{Proj}_{\mathbb{W}_2}(\vec{x}_3) \text{ of } \vec{x}_3 \text{ onto } \mathbb{W}_2.$$

This vector has the property that

$$\vec{v}_3 = \vec{x}_3 - \vec{p}_3 \perp \mathbb{W}_2.$$

Thus we find that $\vec{v}_1 \perp \vec{v}_2 \perp \vec{v}_3 \perp \vec{v}_1$, so, $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ is the desired orthogonal basis for $\mathbb{W}$.

The Gram-Schmidt Orthogonalization Procedure

An algorithm to produce an orthonormal basis from a basis.

Start with a basis $\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_k\}$ for some vector space $\mathbb{W}$.
For $1 \leq i < k$, set $\mathbb{W}_i = \mathcal{S}pan\{\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_i\}$.

Put $\vec{v}_1 = \vec{x}_1$. For $2 \leq i \leq k$:

- Set
$$\vec{p}_i = \text{Proj}_{\mathbb{W}_{i-1}}(\vec{x}_i) = \sum_{j=1}^{i-1} \text{Proj}_{\vec{v}_j}(\vec{x}_i) = \sum_{j=1}^{i-1} \frac{\vec{x}_i \cdot \vec{v}_j}{\vec{v}_j \cdot \vec{v}_j}\, \vec{v}_j = \sum_{j=1}^{i-1} (\vec{x}_i \cdot \vec{u}_j)\vec{u}_j.$$
- Put $\vec{v}_i = \vec{x}_i - \vec{p}_i$.

Then $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ is an *orthogonal* basis for $\mathbb{W}$ with
$$\mathbb{W}_i = \mathcal{S}pan\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_i\}.$$

Finally, we get an *orthonormal* basis $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ for $\mathbb{W}$ by setting
$$\vec{u}_i = \frac{\vec{v}_i}{\|\vec{v}_i\|} \text{ for } 1 \leq i \leq k.$$

Start with $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix}$. First, $\vec{v}_1 = \vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

Next,

$$\vec{v}_2 = \vec{x}_2 - \vec{p}_2 \quad \text{where } \vec{p}_2 = \text{Proj}_{\vec{v}_1}(\vec{x}_2) \text{ and}$$

$$\vec{v}_3 = \vec{x}_3 - \vec{p}_3 \quad \text{where } \vec{p}_3 = \text{Proj}_{\mathbb{W}_2}(\vec{x}_3).$$

$$\vec{p}_2 = \text{Proj}_{\vec{v}_1}(\vec{x}_2) = \frac{\vec{x}_2 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\,\vec{v}_1 = \frac{8}{4}\vec{v}_1 = 2\vec{x}_1, \qquad \text{so, } \vec{v}_2 = \vec{x}_2 - 2\vec{x}_1 = \begin{bmatrix} -3 \\ 2 \\ 2 \\ -1 \end{bmatrix}$$

$$\vec{p}_3 = \text{Proj}_{\mathbb{W}_2}(\vec{x}_3) = \text{Proj}_{\vec{v}_1}(\vec{x}_3) + \text{Proj}_{\vec{v}_2}(\vec{x}_3) = \vec{x}_1 - \tfrac{2}{3}\vec{v}_2$$

$$\text{Proj}_{\vec{v}_1}(\vec{x}_3) = \frac{\vec{x}_3 \cdot \vec{v}_1}{\vec{v}_1 \cdot \vec{v}_1}\,\vec{v}_1 = \frac{4}{4}\vec{v}_1 = \vec{x}_1$$

$$\text{Proj}_{\vec{v}_2}(\vec{x}_3) = \frac{\vec{x}_3 \cdot \vec{v}_2}{\vec{v}_2 \cdot \vec{v}_2}\,\vec{v}_2 = \frac{-12}{18}\vec{v}_2 = \frac{-2}{3}\vec{v}_2$$

So $\vec{v}_3 = \vec{x}_3 - \vec{x}_1 + \frac{2}{3}\vec{v}_2$

and $\vec{v}_3 = \frac{1}{3}\begin{bmatrix} 3 \\ -5 \\ 7 \\ -5 \end{bmatrix}$.

Finally, check orthogonality of

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -3 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 3 \\ -5 \\ 7 \\ -5 \end{bmatrix}.$$

Started with basis $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ where $\vec{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} -1 \\ 4 \\ 4 \\ 1 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 4 \\ -2 \\ 2 \\ 0 \end{bmatrix}$.

Got orthogonal basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ where

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} -3 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \vec{v}_3 = \begin{bmatrix} 3 \\ -5 \\ 7 \\ -5 \end{bmatrix}.$$

To get orthonormal basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, just *normalize*!

Find that $\vec{u}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} -3 \\ 2 \\ 2 \\ -1 \end{bmatrix}, \vec{u}_3 = \frac{1}{6\sqrt{3}} \begin{bmatrix} 3 \\ -5 \\ 7 \\ -5 \end{bmatrix}$.

Starting with basis $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$ where $\vec{x}_1 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$.

Get orthogonal basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ where

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

To get orthonormal basis $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, just *normalize*!

Find that $\vec{u}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \vec{u}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}$.

The QR Factorization

Let A be a matrix with linearly independent columns, say

$$A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_k \end{bmatrix} \text{ where } \vec{a}_j = \text{Col}_j(A) \text{ is in } \mathbb{R}^n.$$

Then $\mathcal{A} = \{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_k\}$ is a basis for the column space $\mathcal{CS}(A)$.

Gram-Schmidt $\mathcal{A}$ to get an orthon basis $\mathcal{U} = \{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k\}$ for $\mathcal{CS}(A)$.
Then each $\vec{a}_j$ can be written as a LC of the $\mathcal{U}$ vectors, so we get

$$\vec{a}_j = \sum_{i=1}^{k} r_{ij}\vec{u}_i = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_k \end{bmatrix} \begin{bmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{kj} \end{bmatrix}.$$

Thus $\boxed{A = QR}$ where
$Q = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_k \end{bmatrix}$
and $R = [r_{ij}]$.

Since $\mathcal{U}$ is *orthonormal*, $Q^T Q = I$, and therefore $\boxed{R = Q^T A}$.

A QR Factorization Example

Let $A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. Get orthon basis $\mathcal{U} = \{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ for $\mathcal{CS}(A)$

where

$$\vec{u}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 2 \\ -1 \\ 0 \end{bmatrix}, \vec{u}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 2 \end{bmatrix}.$$

Then

$$Q = \frac{1}{\sqrt{6}} \begin{bmatrix} -\sqrt{2} & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ \sqrt{2} & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

so

$$R = Q^T A = \frac{1}{\sqrt{6}} \begin{bmatrix} -\sqrt{2} & \sqrt{2} & \sqrt{2} & 0 \\ 1 & 2 & -1 & 0 \\ 1 & 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} -1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \frac{1}{\sqrt{6}} \begin{bmatrix} 3\sqrt{2} & 2\sqrt{2} & -\sqrt{2} \\ 0 & -2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$