Orthonormal Basis

Let $\mathbf{S} = \{\mathbf{v}_1, \mathbf{v}_2, .., \mathbf{v}_n\}$ be a basis for an inner product space $\mathbf{V}$. Then $\mathbf{S}$ is an orthonormal basis for $\mathbf{V}$ if

a)$(\mathbf{v}_i, \mathbf{v}_j)$=0 for $i \neq j$

b)$(\mathbf{v}_i, \mathbf{v}_i)$=1 for all i

Theorem

Let $\mathbf{S} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for an inner product space $\mathbf{V}$ and let $\mathbf{v}$ be any vector in $\mathbf{V}$.

Then $\boldsymbol{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n$

where $c_i = (\mathbf{v}, \mathbf{v}_i)$ for all i

Proof

$$(\mathbf{v}, \mathbf{v}_1) = (c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_i\mathbf{v}_i + \ldots c_n\mathbf{v}_n, \mathbf{v}_i)$$

$$= (c_1\mathbf{v}_1, \mathbf{v}_i) + (c_2\mathbf{v}_2, \mathbf{v}_i) + \ldots + (c_i\mathbf{v}_i, \mathbf{v}_i) + \cdots (\mathrm{c}_n\mathbf{v}_n, \mathbf{v}_i)$$

$$= c_1(\mathbf{v}_1, \mathbf{v}_i) + c_2(\mathbf{v}_2, \mathbf{v}_i) + \ldots + c_i(\mathbf{v}_i, \mathbf{v}_i) + \cdots c_n(\mathbf{v}_n, \mathbf{v}_i)$$

$$= c_1 \bullet 0 + c_2 \bullet 0 + \ldots + c_i \bullet 1 + \ldots c_n \bullet 0$$

$$= c_i$$

Gram-Schmidt Process

If $\boldsymbol{S} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is a basis (not orthonormal) for an inner product space **V**, is there a way to convert it to an orthonormal basis?

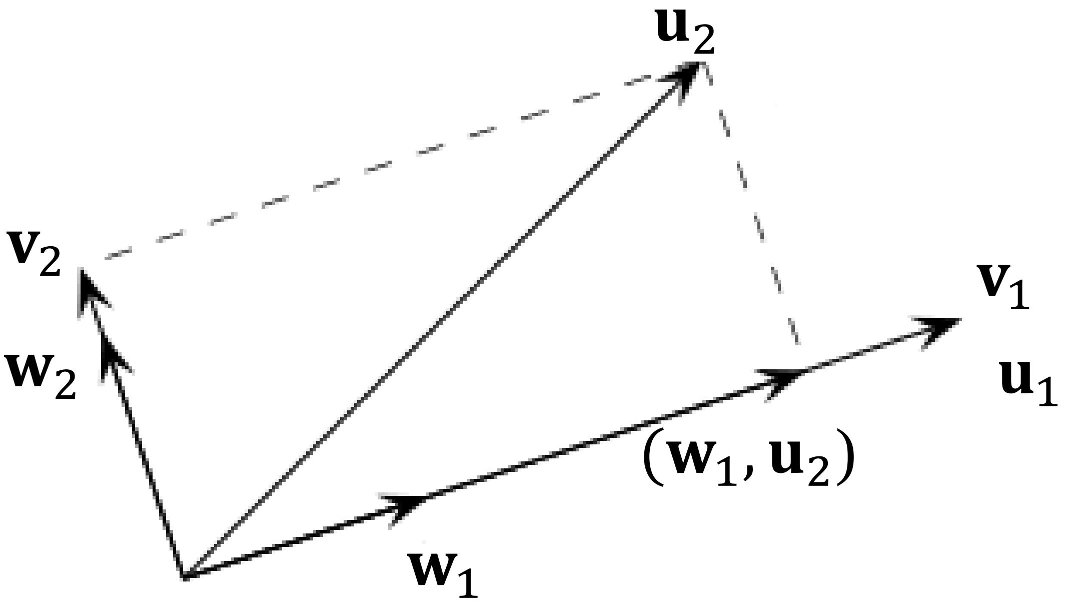
Gram-Schmidt Process

- Replace the basis $\mathbf{S} = \{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \right\}$ with an orthonormal basis $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$

$$\mathbf{v}_1 = \mathbf{u}_1 \Rightarrow \mathbf{w}_1 = \frac{\mathbf{v}_1}{||\mathbf{v}_1||} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbf{v}_2 = \mathbf{u}_2 - (\mathbf{w}_1, \mathbf{u}_2)\mathbf{w}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$$

Gram-Schmidt Process



Gram-Schmidt Process

$$\mathbf{w}_2 = \frac{\mathbf{v}_2}{||\mathbf{v}_2||} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\mathbf{v}_3 = \mathbf{u}_3 - (\mathbf{u}_3, \mathbf{w}_1)\mathbf{w}_1 - (\mathbf{u}_3, \mathbf{w}_2)\mathbf{w}_2$$

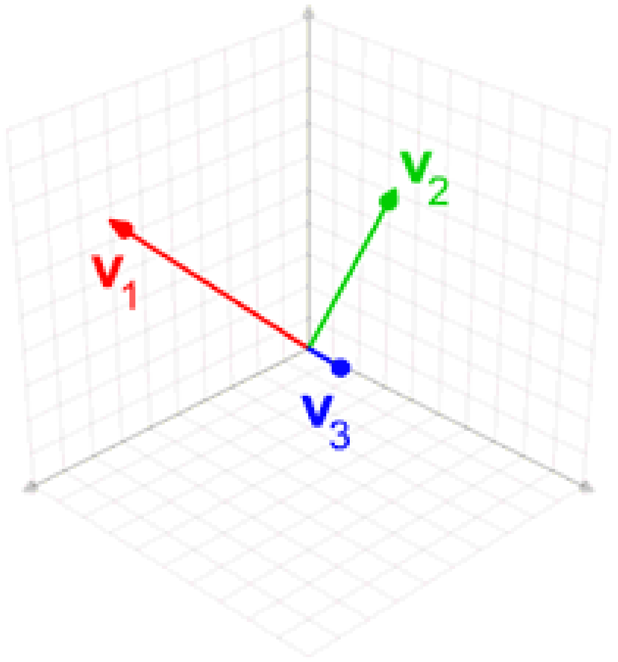
$$= \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} - \sqrt{2} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} - \sqrt{2} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Gram-Schmidt Process

$$\mathbf{w}_3 = \frac{\mathbf{v}_3}{||\mathbf{v}_3||} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Orthonormal set is $\left\{ \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$

Gram-Schmidt Process



Comments

- Key idea in Gram-Schmidt is to subtract from every new vector, $\mathbf{u}_k$, its components in the directions already determined, $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}\}$
- When doing Gram-Schmidt by hand, it simplifies the calculation to multiply the newly computed $\mathbf{v}_k$ by an appropriate scalar to clear fractions in its components. The resulting vectors are normalized at the end of the computation

QR Factorization

In the Gram-Schmidt example, the basis

$$\left\{\begin{bmatrix}1\\0\\1\end{bmatrix},\begin{bmatrix}1\\0\\0\end{bmatrix},\begin{bmatrix}2\\1\\0\end{bmatrix}\right\}$$ is transformed to

$$\left\{\begin{bmatrix}\frac{1}{\sqrt{2}}\\0\\\frac{1}{\sqrt{2}}\end{bmatrix},\begin{bmatrix}\frac{1}{\sqrt{2}}\\0\\-\frac{1}{\sqrt{2}}\end{bmatrix},\begin{bmatrix}0\\1\\0\end{bmatrix}\right\}$$

This is called the QR-Factorization of **A**

QR Factorization

Interpreting these vectors as column vectors of matrices, the following result holds

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

This is called the QR-Factorization of **A**

Comments

- Computer programs that compute the QR Factorization use an algorithm that is different from that of the proof, which is essentially Gram-Schmidt.

Comments

- MATLAB's implementation of QR-Factorization of an mxn matrix **A** returns an mxm matrix **Q** with orthonormal columns and an mxn matrix **R** of the form ⇒

The first n columns of **Q** form a basis for the column space of **A** and $\mathbf{A} = \mathbf{QR}$

$$\mathbf{R} = \begin{bmatrix} * & * & * & * \\ \mathbf{0} & * & * & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Definitions

- A *square* matrix $\mathbf{Q}$ that has orthonormal columns is called an <u>orthogonal matrix</u>
- Because of the orthonormal columns,

  $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$. Therefore $\mathbf{Q}^{-1} = \mathbf{Q}^{T}$