Orthonormality, Subspaces and Projections

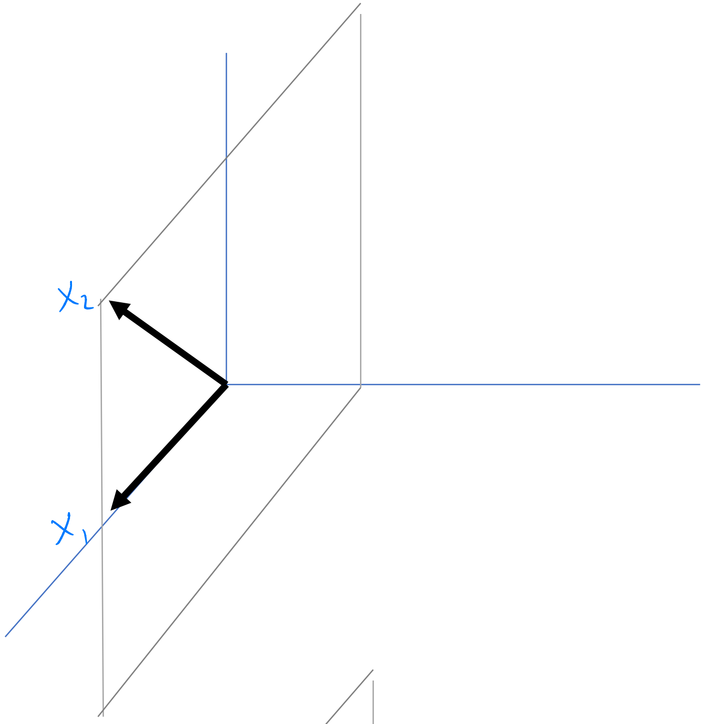
Objectives

- Define and understand orthonormal basis
- Define and understand projections
- List methods for finding orthonormal basis

Definition of orthonormal basis

An orthonormal basis for a set of vectors $\boldsymbol{x}_1, \boldsymbol{x}_2 \ldots$
is another set of vectors $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots$ such that:

1. $\boldsymbol{u}_i^T \boldsymbol{u}_j = 0$ for all $i \neq j$
2. $\boldsymbol{u}_i^T \boldsymbol{u}_i = 1$ for all i
3. $\text{span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} = \text{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m\}$

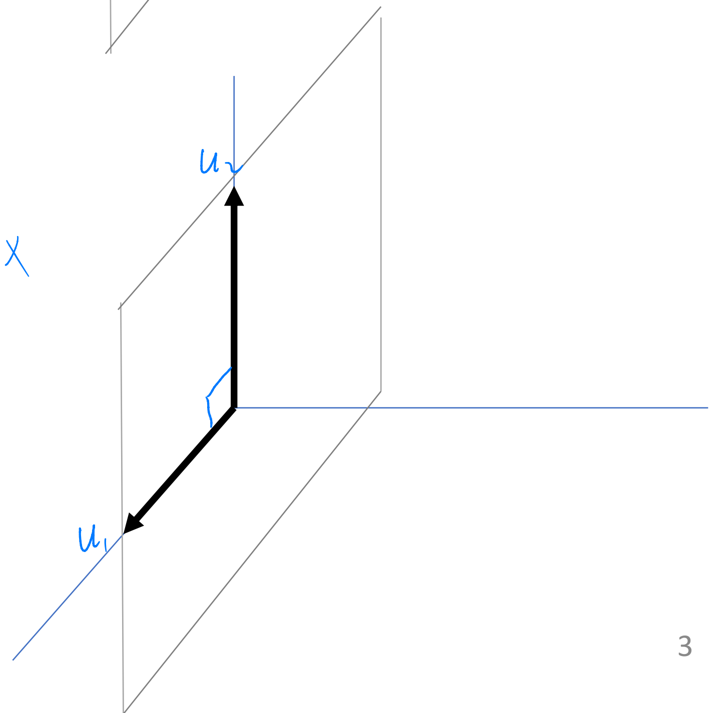


subspace

$$\text{span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\} = \left\{ \boldsymbol{x} : \boldsymbol{x} = \sum_{i=1}^{n} w_i \boldsymbol{x}_i,\ w_i \in \mathbb{R}, i = 1, \ldots, n \right\}$$

$$\boldsymbol{X} = \begin{bmatrix} | & & | \\ \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_n \\ | & & | \end{bmatrix} \qquad \boldsymbol{U} = \begin{bmatrix} | & & | \\ \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_m \\ | & & | \end{bmatrix}$$

orthonormal basis of X



Properties of orthonormal basis

$$X = \begin{bmatrix} | & & | \\ x_1 & \dots & x_n \\ | & & | \end{bmatrix} \qquad U = \begin{bmatrix} | & & | \\ u_1 & \dots & u_m \\ | & & | \end{bmatrix}$$

m is the dimension of the subspace

||

Numbers of basis vectors is the dimension of subspace

Properties of orthonormal basis:

1. $m \leq n$ $m = n$ if $x_1 \cdots$ are linearly independent
2. $m \leq \dim(x_i)$
3. $U^T U = I_{m \times m}$
4. if U is square, $U^{-1} = U^T$
5. if U is square, $UU^T = I_{m \times m}$

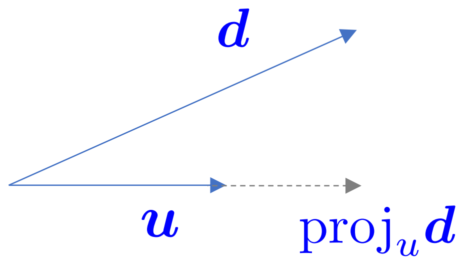
Examples of bases for $\mathbb{R}^n$:

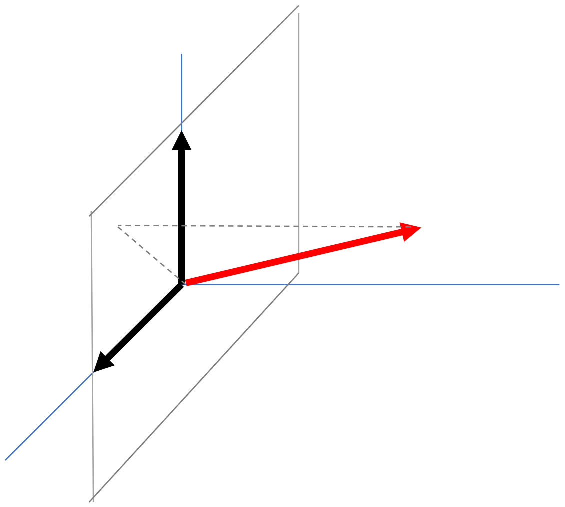
1. Euclidean basis $e_1 = [1 \; 0 \cdots 0]^T$, $e_2 = [0 \; 1 \cdots 0]^T$
2. Haar Wavelets
3. Rotation matrices
4. DFT coefficients (or Fourier basis)

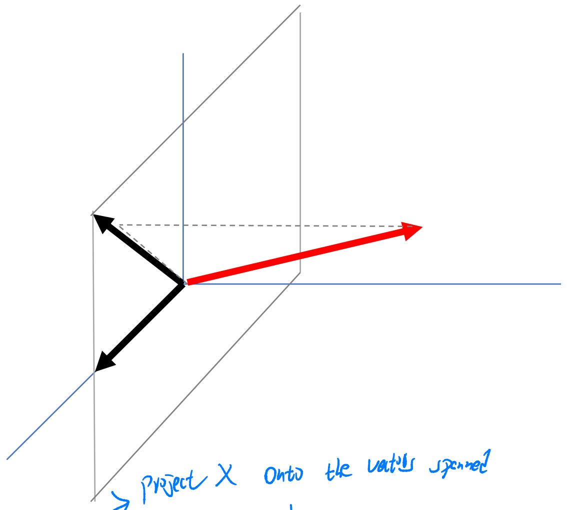
Projections

$$\boldsymbol{X} = \begin{bmatrix} | & & | \\ \boldsymbol{x}_1 & \dots & \boldsymbol{x}_n \\ | & & | \end{bmatrix} \quad \boldsymbol{U} = \begin{bmatrix} | & & | \\ \boldsymbol{u}_1 & \dots & \boldsymbol{u}_m \\ | & & | \end{bmatrix}$$

$n \leq \dim(\boldsymbol{x}_i)$







Direction

$$\mathrm{proj}_u \boldsymbol{d} = \boldsymbol{u}(\boldsymbol{u}^T \boldsymbol{d})$$

projection of $\boldsymbol{d}$ onto $\boldsymbol{u}$

amount of $\boldsymbol{d}$ in the direction of $\boldsymbol{u}$

$$\mathrm{proj}_{\boldsymbol{X}} \boldsymbol{d} = \sum_{i=1}^{m} \boldsymbol{u}_i (\boldsymbol{u}_i^T \boldsymbol{d})$$

subspace

(project individually on basis vectors)

$$= \begin{bmatrix} | & & | \\ \boldsymbol{u}_1 & \dots & \boldsymbol{u}_m \\ | & & | \end{bmatrix} \begin{bmatrix} \boldsymbol{u}_1^T \boldsymbol{d} \\ \vdots \\ \boldsymbol{u}_m^T \boldsymbol{d} \end{bmatrix}$$

$$= \boldsymbol{U}\boldsymbol{U}^T \boldsymbol{d}$$

project X onto the vectors spanned by X

x itself

$$\mathrm{proj}_{\boldsymbol{X}} \boldsymbol{X} = \boldsymbol{U}\boldsymbol{U}^T \boldsymbol{X} = \boldsymbol{X}$$

$$\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T$$

$$\mathrm{proj}_{\boldsymbol{X}} \boldsymbol{d} = \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{d}$$

- Finding Orthonormal Basis

$$\boldsymbol{X} = \begin{bmatrix} | & & | \\ \boldsymbol{x}_1 & \dots & \boldsymbol{x}_n \\ | & & | \end{bmatrix} \xrightarrow{\quad ? \quad} \boldsymbol{U} = \begin{bmatrix} | & & | \\ \boldsymbol{u}_1 & \dots & \boldsymbol{u}_m \\ | & & | \end{bmatrix}$$

- Gram-Schmidt orthogonalization
- scipy.linalg.orth(X), orth(X)
- The SVD (Singular Value Decomposition)

Gram-Schmidt Orthogonalization

The Gram-Schmidt procedure finds an orthonormal basis that spans the same space as vectors $\boldsymbol{a}_i, i = 1, 2, \ldots, P$. It proceeds through the vectors sequentially, starting with $\boldsymbol{a}_1$, and adds the orthonormal component associated with each $\boldsymbol{a}_i$ that is not represented by $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{i-1}$. Let $\boldsymbol{u}_i, i = 1, 2, \ldots, R$ be the orthonormal basis where $R \leq P$. The vector $\boldsymbol{u}_1$ spans the same space as $\boldsymbol{a}_1$. Then $\boldsymbol{u}_1, \boldsymbol{u}_2$ span the same space as $\boldsymbol{a}_1, \boldsymbol{a}_2$ (assuming $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are linearly independent), $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ span the same space as $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ (assuming $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ are linearly independent), and so on.

At each stage, we first find the component of $\boldsymbol{a}_i$ that lies orthogonal to $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{i-1}$, and then define $\boldsymbol{u}_i$ to be this component normalized to unit length. Let $\boldsymbol{U}_{i-1} = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_{i-1} \end{bmatrix}$. Then the component of $\boldsymbol{a}_i$ that lies in the space spanned by the columns of $\boldsymbol{U}_{i-1}$ is $\boldsymbol{b}_i = \boldsymbol{U}_{i-1}\boldsymbol{U}_{i-1}^T\boldsymbol{a}_i$, since $\boldsymbol{U}_{i-1}\boldsymbol{U}_{i-1}^T$ is a projection matrix for the space spanned by the columns of $\boldsymbol{U}_{i-1}$. Note that we may also write $\boldsymbol{b}_i = \sum_{k=1}^{i-1} \boldsymbol{u}_k(\boldsymbol{u}_k^T\boldsymbol{a}_i)$. Hence, the component of $\boldsymbol{a}_i$ that is not in the space spanned by the columns of $\boldsymbol{U}_{i-1}$ is $\boldsymbol{c}_i = \boldsymbol{a}_i - \boldsymbol{b}_i$ and we define $\boldsymbol{u}_i = \boldsymbol{c}_i/||\boldsymbol{c}_i||_2$.

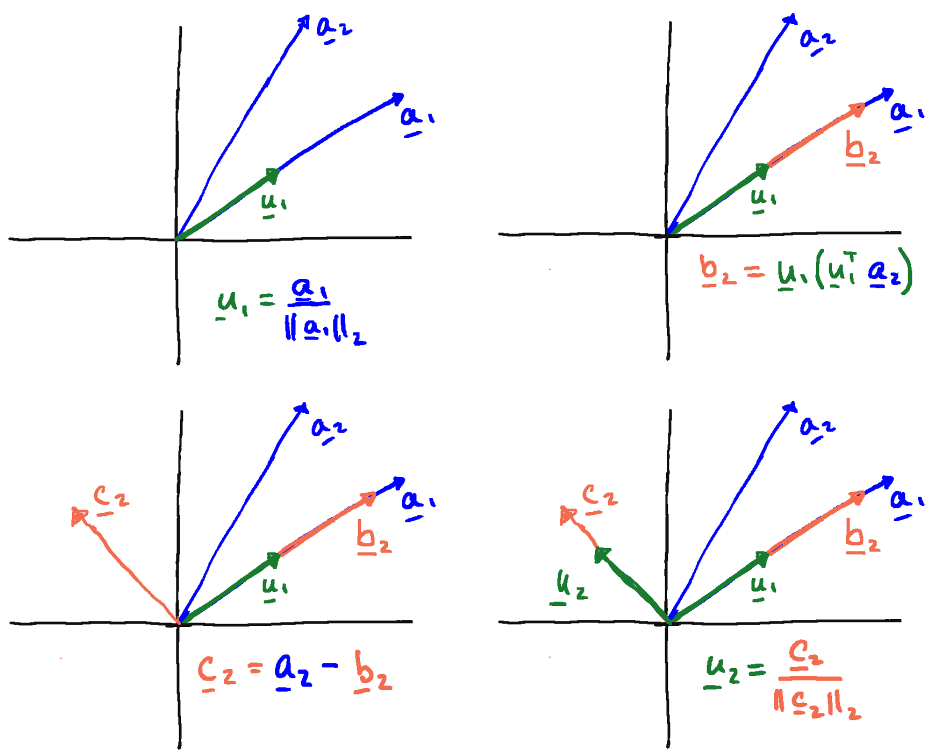


Figure 1: Gram-Schmidt orthogonalization of two vectors.

This process is illustrated in Fig. 1 for a two-dimensional problem. The upper left panel shows $\boldsymbol{u}_1$ as a normalized version of $\boldsymbol{a}_1$. Then the upper right panel shows $\boldsymbol{b}_2$ as the projection of $\boldsymbol{a}_2$ onto the space spanned by $\boldsymbol{u}_1$. The lower left panel depicts $\boldsymbol{c}_2$ as the projection of $\boldsymbol{a}_2$ onto the space orthogonal to that spanned by $\boldsymbol{u}_1$. Finally, in the lower right panel we normalize $\boldsymbol{c}_2$ to obtain $\boldsymbol{u}_2$.

Figure 2 illustrates computation of $\boldsymbol{u}_3$ for a three-dimensional subspace. The process for finding $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ is identical to that illustrated in Fig. 1. The left panel illustrates computation of $\boldsymbol{b}_3$, the projection of $\boldsymbol{a}_3$ onto the space spanned by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. The vector $\boldsymbol{b}_3$ lies in the plane corresponding to the span of $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ or $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$. Hence, $\boldsymbol{c}_3$, shown on the right panel, is the component of $\boldsymbol{a}_3$ that is orthogonal to the plane defined by span$\{\boldsymbol{u}_1, \boldsymbol{u}_2\}$. We obtain $\boldsymbol{u}_3$ by normalizing $\boldsymbol{c}_3$ to unit length.

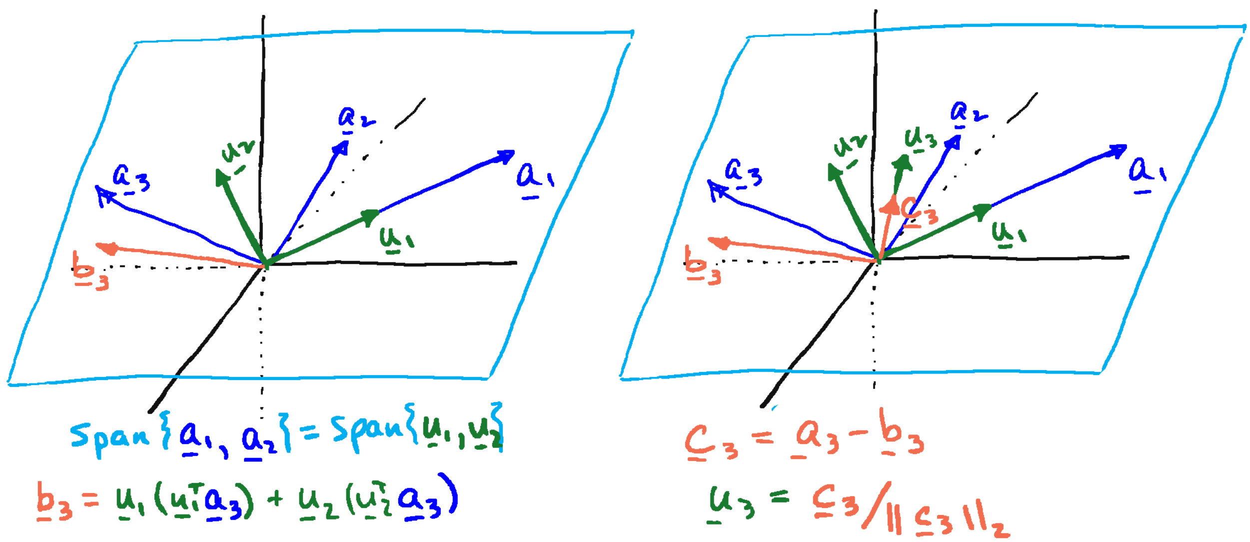


Figure 2: Gram-Schmidt orthogonalization of three vectors.

Note that if the $\boldsymbol{a}_i, i = 1, 2, \ldots, P$ are linearly dependent, then one of the $\boldsymbol{c}_i$ will be all zeros. In that case $\boldsymbol{a}_i$ lies in the space spanned by $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_{i-1}$. In such a case we move on to $\boldsymbol{a}_{i+1}$ to find the next orthonormal basis vector.

A psuedo-code for the Gram-Schmidt orthogonalization procedure is given as:

$\boldsymbol{u}_1 = \boldsymbol{a}_1/||\boldsymbol{a}_1||_2$
$\boldsymbol{U}_1 = \boldsymbol{u}_1$
$j = 1$

for i = 2 to P
  $\boldsymbol{c}_i = (\boldsymbol{I} - \boldsymbol{U}_j\boldsymbol{U}_j^T)\boldsymbol{a}_i$
  if $||\boldsymbol{c}_i||_2 >$ tol, then
    $j = j + 1$
    $\boldsymbol{u}_j = \boldsymbol{c}_i/||\boldsymbol{c}_i||_2$
    $\boldsymbol{U}_j = \left[\begin{array}{cc} \boldsymbol{U}_{j-1} & \boldsymbol{u}_j \end{array}\right]$

end if
end for

Example 1. Find an orthonormal basis for the space spanned by the columns of

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 \\ 1 & 3 \end{bmatrix}$$

Here $\boldsymbol{a}_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}^T$, so

$$\boldsymbol{u}_1 = \boldsymbol{a}_1/||\boldsymbol{a}_1||_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Next, we find $\boldsymbol{u}_2$ as follows: $\boldsymbol{c}_2 = (\boldsymbol{I} - \boldsymbol{u}_1\boldsymbol{u}_1^T)\boldsymbol{a}_2$ or

$$\boldsymbol{c}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left(\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right)$$

so

$$\boldsymbol{c}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} - \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

or

$$\boldsymbol{c}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

which implies

$$\boldsymbol{u}_2 = \boldsymbol{c}_2/||\boldsymbol{c}_2||_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Figure 3 depicts the columns of **A** and orthonormal bases.

Example 2. Find an orthonormal basis for the space spanned by the columns of

$$\boldsymbol{A} = \begin{bmatrix} -1 & 1 & 3 \\ 1 & -1 & 3 \end{bmatrix}$$

Here $\boldsymbol{a}_1 = \begin{bmatrix} -1 & 1 \end{bmatrix}^T$, so

$$\boldsymbol{u}_1 = \boldsymbol{a}_1/||\boldsymbol{a}_1||_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Next, we find $\boldsymbol{u}_2$ as follows: $\boldsymbol{c}_2 = (\boldsymbol{I} - \boldsymbol{u}_1\boldsymbol{u}_1^T)\boldsymbol{a}_2$ or

$$\boldsymbol{c}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \left(\begin{bmatrix} -1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right)$$

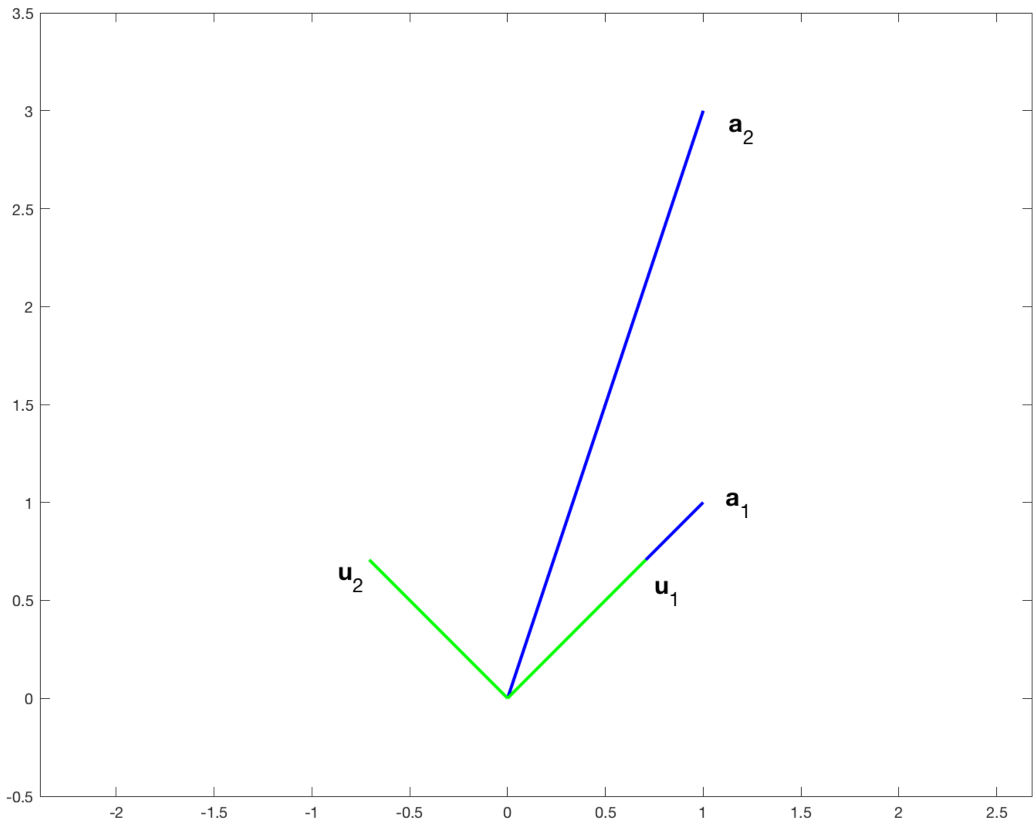


Figure 3: Gram-Schmidt orthogonalization Example 1.

so

$$c_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

or

$$c_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Thus $\boldsymbol{a}_2$ is linearly dependent on $\boldsymbol{a}_1$ and we must use $\boldsymbol{a}_3$ to find $\boldsymbol{u}_2$: $\boldsymbol{c}_3 = (\boldsymbol{I} - \boldsymbol{u}_1\boldsymbol{u}_1^T)\boldsymbol{a}_3$ or

$$\boldsymbol{c}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \end{bmatrix}\left(\begin{bmatrix} -1 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 3 \end{bmatrix}\right)$$

so

$$\boldsymbol{c}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$\boldsymbol{c}_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

Hence

$$\boldsymbol{u}_2 = \boldsymbol{c}_3/||\boldsymbol{c}_3||_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Figure 4 depicts the columns of $\mathbf{A}$ and orthonormal bases.

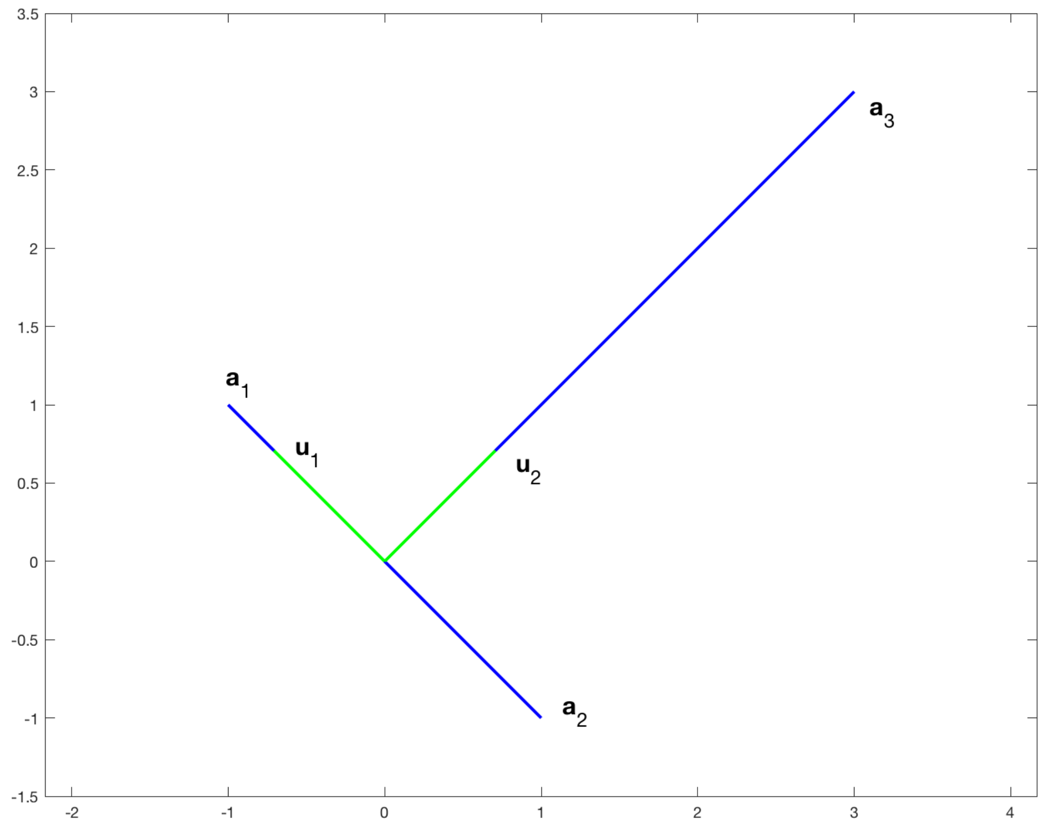


Figure 4: Gram-Schmidt orthogonalization Example 2.

Example 3. Find an orthonormal basis for the space spanned by the columns of

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

Here $\boldsymbol{a}_1 = \begin{bmatrix} -1 & 1 & 0 \end{bmatrix}^T$, so

$$\boldsymbol{u}_1 = \boldsymbol{a}_1/||\boldsymbol{a}_1||_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Next, we find $\boldsymbol{u}_2$ as follows: $\boldsymbol{c}_2 = (\boldsymbol{I} - \boldsymbol{u}_1\boldsymbol{u}_1^T)\boldsymbol{a}_2$ or

$$\boldsymbol{c}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\left(\begin{bmatrix} -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}\right)$$

so

$$\boldsymbol{c}_2 = \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

or

$$\boldsymbol{c}_2 = \begin{bmatrix} -0.5 \\ -0.5 \\ 1 \end{bmatrix}$$

Thus

$$\boldsymbol{u}_2 = \boldsymbol{c}_2/||\boldsymbol{c}_2||_2 = \sqrt{\frac{2}{3}}\begin{bmatrix} -0.5 \\ -0.5 \\ 1 \end{bmatrix}$$

Next we find $\boldsymbol{u}_3$ by removing the components in the space spanned by $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$ from $\boldsymbol{a}_3$:

$$\boldsymbol{c}_3 = (\boldsymbol{I} - (\boldsymbol{u}_1\boldsymbol{u}_1^T + \boldsymbol{u}_2\boldsymbol{u}_2^T))\boldsymbol{a}_3 = \boldsymbol{a}_3 - \boldsymbol{u}_1(\boldsymbol{u}_1^T\boldsymbol{a}_3) - \boldsymbol{u}_2(\boldsymbol{u}_2^T\boldsymbol{a}_3)$$

$$\boldsymbol{c}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}\left(\begin{bmatrix} -1 & 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) - \frac{2}{3}\begin{bmatrix} -0.5 \\ -0.5 \\ 1 \end{bmatrix}\left(\begin{bmatrix} -0.5 & -0.5 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right)$$

or

$$\boldsymbol{c}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \frac{2}{3}\begin{bmatrix} -0.25 \\ -0.25 \\ 0.5 \end{bmatrix}$$

so

$$\boldsymbol{c}_3 = \begin{bmatrix} 2/3 \\ 2/3 \\ 2/3 \end{bmatrix}$$

Hence

$$\boldsymbol{u}_3 = \boldsymbol{c}_3/||\boldsymbol{c}_3||_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

Figure 5 depicts the columns of $\mathbf{A}$ and orthonormal bases.

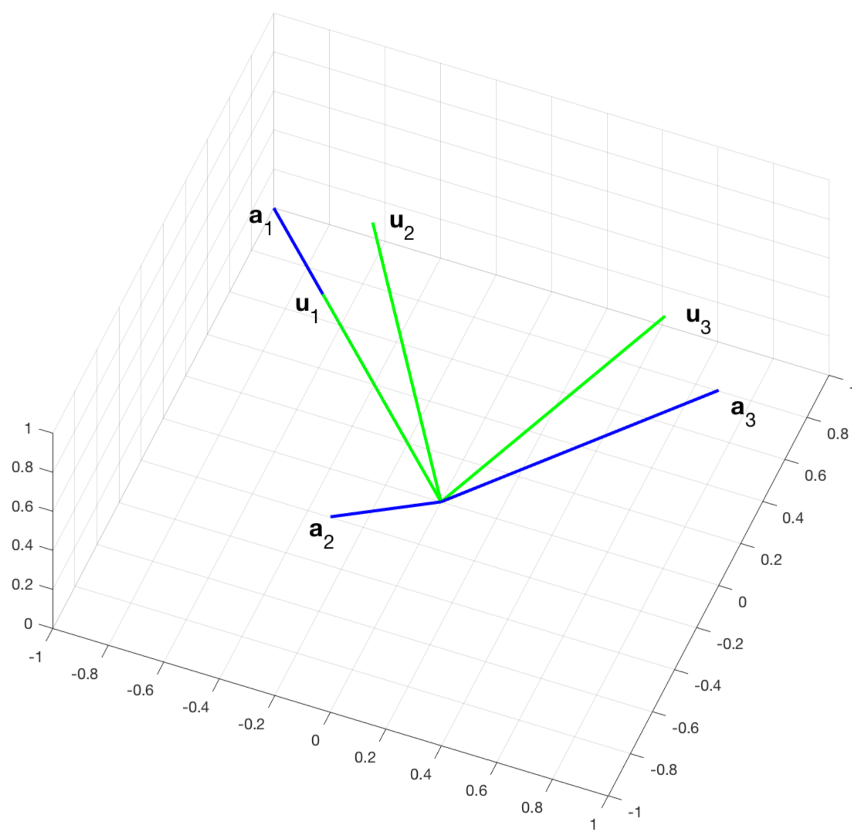


Figure 5: Gram-Schmidt orthogonalization Example 3.