

Why Gram-Schmidt orthogonalization works

Peter Haggstrom
mathsatbondibeach@gmail.com
<https://gotohaggstrom.com>

January 18, 2020

1 Background

The Gram-Schmidt orthogonalisation process is fundamental to linear algebra and its applications to machine learning for instance. In many courses it is presented as an algorithm that works without any real motivation.

The proof of the Gram-Schmidt (GS) orthogonalisation process relies upon a recursive procedure which replicates why it works in 2 and 3 dimensions. We can see why it works in low dimensions but in a 1000 dimensional space the fact that it works relies upon purely analytical properties and induction since you can't actually visualise the mutual orthogonality.

In what follows I will focus on vectors with an associated dot product rather than a more general inner product space since the basic elements of the GS process are conceptually the same whether we are dealing with vectors in \mathbb{R}^n or some vector space of functions with an appropriate norm eg the classic orthonormal trigonometric function sets of Fourier theory.

Let us first recapitulate some basic principles. In 2 dimensions let us construct a basis. We need two linearly independent vectors which will enable us to construct any other vector by taking appropriately scaled amounts of each basis vector and then add them to get the vector we want. That is what a basis is. The standard basis in \mathbb{R}^2 is $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}$. It is an orthonormal basis since the component vectors are normalised to unit length.

However, an equally legitimate basis is the following:

$$\left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \end{pmatrix} \right\} \tag{1}$$

We let $\vec{u}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\vec{u}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ and we can see that these two vectors are linearly independent by either simply drawing them or noting that the only way that:

$$\lambda_1 \vec{u}_1 + \lambda_2 \vec{u}_2 = \vec{0} \quad (2)$$

is if $\lambda_1 = \lambda_2 = 0$.

In this context (2) can be rewritten as:

$$\begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (3)$$

The matrix is invertible (with determinant -2) and hence by multiplying both sides by the inverse we get $\lambda_1 = \lambda_2 = 0$.

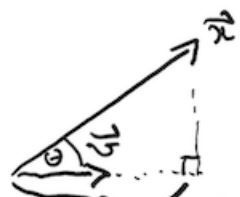
For any vector $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ to find the weights λ_1, λ_2 we simply need to solve:

$$\begin{pmatrix} 2 & 0 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \quad (4)$$

Thus:

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 \\ \frac{1}{2} & -1 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \frac{v_1}{2} \\ \frac{v_1 - 2v_2}{2} \end{pmatrix} \quad (5)$$

So we know that the vectors in (1) form a basis but if we want an orthogonal (actually orthonormal) basis we need to employ the GS process. The crux of the GS process is the concept of an orthogonal projection of one vector onto another (or a linear combination of others). The following diagram explains the essence:



projection of \vec{x} onto unit vector $\vec{v} = |\vec{x}| \cos \theta \vec{v}$

$$= \frac{(\vec{x} \cdot \vec{v}) |\vec{x}| |\vec{v}|}{|\vec{x}| |\vec{v}|}$$

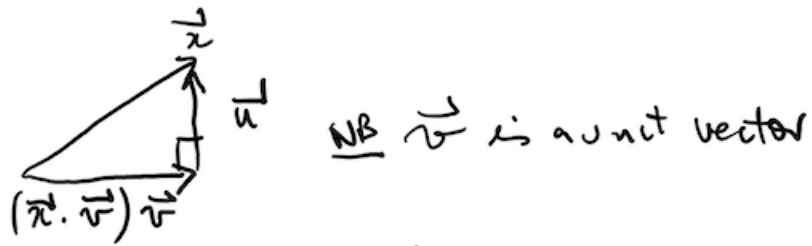
i.e. $\text{proj}_{\vec{v}} \vec{x} = (\vec{x} \cdot \vec{v}) \vec{v}$ noting that $|\vec{v}| = 1$.

Note that a collection of n pairwise orthogonal vectors $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\}$ in \mathbb{R}^n is linearly independent and hence forms a basis. This is so because $\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n = 0$ implies that for any k :

$$\begin{aligned} \vec{v}_k \cdot (\lambda_1 \vec{v}_1 + \dots + \lambda_n \vec{v}_n) &= \lambda_k \vec{v}_k \cdot \vec{v}_k \\ &= \lambda_k |\vec{v}_k|^2 \\ &= 0 \implies \lambda_k = 0 \end{aligned} \tag{6}$$

Note: I am using dot product notation but the inner product notation is actually more general and one would write $\langle u, v \rangle$ for instance. The GS process is general enough to apply to functions in inner product spaces where, for example, the inner product is defined as $\langle f, g \rangle = \int_{-\pi}^{\pi} f(t)g(t) dt$.

We can now see how the orthonormal vector is obtained (see the diagram below):



$$\vec{x} = (\vec{x} \cdot \vec{u})\vec{u} + \vec{w}$$

$$\therefore \vec{w} = \vec{x} - (\vec{x} \cdot \vec{u})\vec{u}$$

$\vec{w} = \frac{\vec{x} \cdot \vec{u}}{|\vec{u}|}$ is orthogonal to $(\vec{x} \cdot \vec{u})\vec{u}$. Because:

$$\begin{aligned} \vec{w} \cdot (\vec{x} \cdot \vec{u})\vec{u} &= \frac{\vec{x} \cdot \vec{u}}{|\vec{u}|} \left[(\vec{x} - (\vec{x} \cdot \vec{u})\vec{u}) \cdot (\vec{x} \cdot \vec{u})\vec{u} \right] \\ &= \frac{\vec{x} \cdot \vec{u}}{|\vec{u}|} \left[(\vec{x} \cdot \vec{u})^2 - (\vec{x} \cdot \vec{u})^2 |\vec{u}|^2 \right] \\ &= \frac{\vec{x} \cdot \vec{u}}{|\vec{u}|} (\vec{x} \cdot \vec{u})^2 [1 - |\vec{u}|^2] = 0 \end{aligned}$$

$\because |\vec{u}| = 1$

Starting with our basis $\vec{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\vec{v}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ and applying these principles to the basis in (1) we first normalise one of the basis vectors thus:

$$\vec{u}_1 = \frac{\vec{v}_1}{|\vec{v}_1|} = \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} \quad (7)$$

The orthogonal vector is:

$$\begin{aligned} \vec{w} &= \vec{v}_2 - (\vec{v}_2 \cdot \vec{u}_1) \vec{u}_1 \\ &= \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \frac{1}{\sqrt{5}} \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{2}{5} \\ -\frac{4}{5} \end{pmatrix} \end{aligned} \quad (8)$$

We now normalise \vec{w} and obtain our second orthonormal vector (the norm $|\vec{w}| = \frac{2}{\sqrt{5}}$):

$$\vec{u}_2 = \frac{\sqrt{5}}{2} \begin{pmatrix} \frac{2}{5} \\ -\frac{4}{5} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix} \quad (9)$$

Thus our orthonormal basis is $\{\vec{u}_1, \vec{u}_2\} = \left\{ \begin{pmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix}, \begin{pmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{pmatrix} \right\}$

It is easily seen that $\vec{u}_1 \cdot \vec{u}_2 = 0$ and by construction both vectors are unit vectors.

1.1 A basic result needed to prove the GS process

We have already seen that n pairwise orthogonal (or orthonormal) vectors \vec{u}_k for $k = 1, 2, \dots, n$ is linearly independent and using this we are able to assert that, for any \vec{v} , the following vector is orthogonal to each of the \vec{u}_k :

$$\vec{w} = \vec{v} - (\vec{v} \cdot \vec{u}_1) \vec{u}_1 - (\vec{v} \cdot \vec{u}_2) \vec{u}_2 - \dots - (\vec{v} \cdot \vec{u}_n) \vec{u}_n \quad (10)$$

To prove the orthogonality for each k take the dot product (or inner product) of \vec{u}_k with each side of (10):

$$\begin{aligned} \vec{u}_k \cdot \vec{w} &= \vec{u}_k \cdot \vec{v} - (\vec{v} \cdot \vec{u}_1) \vec{u}_1 \cdot \vec{u}_k - \dots - (\vec{v} \cdot \vec{u}_k) \vec{u}_k \cdot \vec{u}_k - \dots - (\vec{v} \cdot \vec{u}_n) \vec{u}_n \cdot \vec{u}_k \\ &= \vec{u}_k \cdot \vec{v} - \vec{v} \cdot \vec{u}_k \\ &= 0 \end{aligned} \quad (11)$$

since the \vec{u}_i are orthonormal ie $\vec{u}_i \cdot \vec{u}_j = 1$ if $i = j$ and 0 otherwise.

2 The proof

The proof of the GS process is really nothing more than describing the algorithm plus an appeal to induction. The GS process works for any inner product space (the above treatment was confined to a vector space with a dot product) but we will stick with n independent vectors which can be viewed as the columns of an $n \times n$ matrix A of rank n . They form a basis. We want to construct orthonormal vectors $\vec{q}_1, \vec{q}_2, \dots, \vec{q}_n$. The reason for this approach is that we will end up with this relationship:

$$A = QR \quad (12)$$

where Q is the matrix of orthonormal vectors and R is an upper triangular matrix. This decomposition is critical to a number of applied contexts and I recommend Gil Strang's heavily revised textbook on linear algebra in the new era of data analysis [1] for more on the connection between GS and applications as well as his new MIT course 18.065 covering the same material [2].

We start with the independent columns $\vec{a}_1, \dots, \vec{a}_n$ and form the following normalised vector ($|\vec{q}_1| = 1$):

$$\vec{q}_1 = \frac{\vec{a}_1}{|\vec{a}_1|} \quad (13)$$

The recipe for the orthogonal basis is this: $\vec{A}_2 = \vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1$ then normalise \vec{A}_2 ie $q_2 = \frac{\vec{A}_2}{|\vec{A}_2|}$ and repeat (see (8) and the diagram that precedes it).

We should check here that we have orthogonality ie \vec{A}_2 is orthogonal to \vec{q}_1 :

$$\begin{aligned} \vec{A}_2^T \vec{q}_1 &= \left(\vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1 \right)^T \vec{q}_1 \\ &= \vec{a}_2^T \vec{q}_1 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1^T \vec{q}_1 \quad \text{note that } \vec{q}_1^T \vec{q}_1 = 1 \\ &= \vec{a}_2^T \vec{q}_1 - \vec{a}_2^T \vec{q}_1 \\ &= 0 \end{aligned} \quad (14)$$

We repeat the orthogonalisation process which involves subtracting the projections of the current vector, \vec{a}_3 in this case, onto each of the previously established orthonormal vectors and then normalising:

$$\vec{A}_3 = \vec{a}_3 - (\vec{a}_3^T \vec{q}_1) \vec{q}_1 - (\vec{a}_3^T \vec{q}_2) \vec{q}_2 \quad (15)$$

Again we need to check orthogonality:

$$\vec{A}_3^T \vec{q}_1 = \vec{a}_3^T \vec{q}_1 - \vec{a}_3^T \vec{q}_1 - (\vec{a}_3^T \vec{q}_2) \vec{q}_2^T \vec{q}_1 = 0 \quad (16)$$

since $\vec{q}_2^T \vec{q}_1 = 0$ - see (14)

Similarly, $\vec{A}_3^T \vec{q}_2 = 0$ since (recalling that \vec{q}_1 and \vec{q}_2 are orthogonal):

$$\vec{A}_3^T \vec{q}_2 = \vec{a}_3^T \vec{q}_2 - \vec{a}_3^T \vec{q}_2 = 0 \quad (17)$$

Of course $q_3 = \frac{\vec{A}_3}{|\vec{A}_3|}$.

It is important to note that at each stage $|\vec{A}_k| \neq 0$ ie $\vec{A}_k \neq \vec{0}$. The reason for this is that \vec{A}_k is a linear combination of $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k$. For instance, $\vec{A}_2 = \vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1$ where $\vec{q}_1 = \frac{\vec{a}_1}{|\vec{a}_1|}$. Because the \vec{a}_i form a basis (recall that they are assumed to be independent) the only way a linear combination can be the null vector is if all the coefficients in the sum are zero. Clearly this is not the case. Hence we can divide by $|\vec{A}_k|$ at each stage of the algorithm.

Note that each of the orthonormal \vec{q}_k is linear combination of $\vec{a}_1, \dots, \vec{a}_k$ or, equally validly, each \vec{a}_k is a linear combination of $\vec{q}_1, \dots, \vec{q}_k$.

For $k = 3$ we get the following:

$$\begin{aligned}
\vec{a}_1 &= |\vec{a}_1| \vec{q}_1 \\
\vec{a}_2 &= (\vec{a}_2^T \vec{q}_1) \vec{q}_1 + |\vec{A}_2| \vec{q}_2 \\
\vec{a}_3 &= (\vec{a}_3^T \vec{q}_1) \vec{q}_1 + (\vec{a}_3^T \vec{q}_2) \vec{q}_2 + |\vec{A}_3| \vec{q}_3
\end{aligned} \tag{18}$$

In general we will have that:

$$\vec{a}_n = (\vec{a}_n^T \vec{q}_1) \vec{q}_1 + (\vec{a}_n^T \vec{q}_2) \vec{q}_2 + \dots + (\vec{a}_n^T \vec{q}_n) \vec{q}_n \tag{19}$$

(18) suggests an upper triangular structure $A = QR$ where $R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix}$,

but what are the r_{ij} components?

The form of r_{ij} (see (18)) is this:

$$r_{ij} = \vec{q}_i^T \vec{a}_j \tag{20}$$

Incidentally, because r_{ij} is just a number (assumed real in this context) it is equal to its transpose (or conjugate transpose in the complex case) ie $\vec{a}_j^T \vec{q}_i$.

This means that (19) can be written as:

$$R = Q^T A \tag{21}$$

Because Q is an orthogonal (actually orthonormal) matrix $Q^T = Q^{-1}$ so that $Q^T Q = QQ^T = I$, hence $A = QR$.

We really should check that (19) makes sense so we should be able to confirm that, for instance:

$$r_{22} = |\vec{A}_2| = \vec{q}_2^T \vec{a}_2 \tag{22}$$

Since:

$$\vec{q}_2 = \frac{\vec{A}_2}{|\vec{A}_2|} \tag{23}$$

we have:

$$\begin{aligned}
r_{22} &= \vec{q}_2^T \vec{a}_2 \\
&= \frac{1}{|\vec{A}_2|} \left(\vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1 \right)^T \vec{a}_2 \\
&= \frac{1}{|\vec{A}_2|} \left(\vec{a}_2^T \vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1^T \vec{a}_2 \right)
\end{aligned} \tag{24}$$

But (noting that $\vec{q}_1^T \vec{q}_1 = 1$):

$$\begin{aligned}
|\vec{A}_2|^2 &= \vec{A}_2^T \vec{A}_2 \\
&= \left(\vec{a}_2^T - (\vec{a}_2^T \vec{q}_1) \vec{q}_1^T \right) \left(\vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1 \right) \\
&= \vec{a}_2^T \vec{a}_2 - (\vec{a}_2^T \vec{q}_1)^2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1^T \vec{a}_2 + (\vec{a}_2^T \vec{q}_1)^2 \\
&= \vec{a}_2^T \vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1^T \vec{a}_2
\end{aligned} \tag{25}$$

Putting (25) into (24) gives us:

$$r_{22} = \frac{1}{|\vec{A}_2|} |\vec{A}_2|^2 = |\vec{A}_2| \tag{26}$$

as required. Clearly, the recursive process of generating each orthonormal vector ultimately involves an appeal to induction.

3 Example

As an example consider the following basis in \mathbb{R}^3 :

$$\begin{aligned}
\vec{a}_1 &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
\vec{a}_2 &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \\
\vec{a}_3 &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}
\end{aligned} \tag{27}$$

Thus:

$$A = (\vec{a}_1 \quad \vec{a}_2 \quad \vec{a}_3) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \tag{28}$$

We start with:

$$\vec{q}_1 = \frac{\vec{a}_1}{|\vec{a}_1|} = \begin{pmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{pmatrix} \tag{29}$$

Next:

$$\begin{aligned}\vec{A}_2 &= \vec{a}_2 - (\vec{a}_2^T \vec{q}_1) \vec{q}_1 \\ &= \begin{pmatrix} -\frac{2}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}\end{aligned}\tag{30}$$

$$|\vec{A}_2| = \frac{\sqrt{6}}{3}\tag{31}$$

Hence:

$$\vec{q}_2 = \frac{\vec{A}_2}{|\vec{A}_2|} = \begin{pmatrix} -\frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{pmatrix}\tag{32}$$

Finally:

$$\begin{aligned}\vec{A}_3 &= \vec{a}_3 - (\vec{a}_3^T \vec{q}_1) \vec{q}_1 - (\vec{a}_3^T \vec{q}_2) \vec{q}_2 \\ &= \begin{pmatrix} 0 \\ -\frac{1}{2} \\ \frac{1}{2} \end{pmatrix}\end{aligned}\tag{33}$$

$$|\vec{A}_3| = \frac{1}{\sqrt{2}}\tag{34}$$

Hence:

$$\vec{q}_3 = \frac{\vec{A}_3}{|\vec{A}_3|} = \begin{pmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix}\tag{35}$$

Our coefficients for Q are as follows:

$$\begin{aligned}
r_{11} &= \vec{q}_1^T \vec{a}_1 = \left(\frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \right) \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \sqrt{3} \\
r_{12} &= \vec{q}_1^T \vec{a}_2 = \left(\frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \right) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \frac{2}{\sqrt{3}} \\
r_{13} &= \vec{q}_1^T \vec{a}_3 = \left(\frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \quad \frac{1}{\sqrt{3}} \right) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{3}} \\
r_{22} &= \vec{q}_2^T \vec{a}_2 = \left(-\frac{2}{\sqrt{6}} \quad \frac{1}{\sqrt{6}} \quad \frac{1}{\sqrt{6}} \right) \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \frac{2}{\sqrt{6}} \\
r_{23} &= \vec{q}_2^T \vec{a}_3 = \left(-\frac{2}{\sqrt{6}} \quad \frac{1}{\sqrt{6}} \quad \frac{1}{\sqrt{6}} \right) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{6}} \\
r_{33} &= \vec{q}_3^T \vec{a}_3 = \left(0 \quad -\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}} \right) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}
\end{aligned} \tag{36}$$

Thus:

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ 0 & r_{22} & r_{23} \\ 0 & 0 & r_{33} \end{pmatrix} = \begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \tag{37}$$

We now check the product QR:

$$QR = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = A \tag{38}$$

We can use Mathematica to perform the QR decomposition using the function "QRDecomposition[]". Note that Mathematica performs the decomposition as $A = Q^T R$ as can be seen from the example below.

$$\text{In[1]:= } \mathbf{m} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$$

QRDecomposition[m] // TraditionalForm

$$\text{Out[1]= } \{ \{1, 0, 0\}, \{1, 1, 0\}, \{1, 1, 1\} \}$$

Out[2]//TraditionalForm=

$$\left(\begin{array}{ccc} \left\{ \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right\} & \left\{ -\sqrt{\frac{2}{3}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}} \right\} & \left\{ 0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right\} \\ \left\{ \sqrt{3}, \frac{2}{\sqrt{3}}, \frac{1}{\sqrt{3}} \right\} & \left\{ 0, \sqrt{\frac{2}{3}}, \frac{1}{\sqrt{6}} \right\} & \left\{ 0, 0, \frac{1}{\sqrt{2}} \right\} \end{array} \right)$$

In[3]:= Map[MatrixForm, %]

$$\text{Out[3]= } \left\{ \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}, \begin{pmatrix} \sqrt{3} & \frac{2}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ 0 & \sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \right\}$$

4 Observations

GS is remarkable in its simplicity for vector and more general inner product spaces. The basic geometric unsight from 2 and 3 dimensions easily scales in an inductive way to arbitrary dimensions and the process itself guarantees the existence of the required orthonormal set.

5 References

- [1] Gilbert Strang, "Linear Algebra and Learning from Data", Wellesley-Cambridge, 2019
- [2] <https://ocw.mit.edu/courses/mathematics/18-065-matrix-methods-in-data-analysis-signal-processing-and-machine-learning-spring-2018/>

6 History

Created 18 January 2020