Review. The **projection matrix** for projecting onto $\mathrm{col}(A)$ is $P = A(A^TA)^{-1}A^T$.

Projecting onto 1-dimensional spaces

When we project onto a 1-dimensional space $\mathrm{span}\{\boldsymbol{w}\}$, we usually just say that we are projecting onto $\boldsymbol{w}$.

The (orthogonal) projection of $\boldsymbol{v}$ onto $\boldsymbol{w}$ is $\dfrac{\boldsymbol{w}\cdot\boldsymbol{v}}{\|\boldsymbol{w}\|^2}\boldsymbol{w}$.

Why? Replace $\boldsymbol{b}$ with $\boldsymbol{v}$ and A with $\boldsymbol{w}$ in our general projection matrix formula to get $\boldsymbol{w}(\boldsymbol{w}^T\boldsymbol{w})^{-1}\boldsymbol{w}^T\boldsymbol{v}$, which equals $\dfrac{\boldsymbol{w}\cdot\boldsymbol{v}}{\|\boldsymbol{w}\|^2}\boldsymbol{w}$ (note that $\boldsymbol{w}^T\boldsymbol{v} = \boldsymbol{w}\cdot\boldsymbol{v}$ and $\boldsymbol{w}^T\boldsymbol{w} = \|\boldsymbol{w}\|^2$ are scalars).

Comment. If you have taken Calculus 3, you have seen that formula before. Most likely, you were deriving it using angles at that time. Namely, the dot product has the following connection to angles:

$\boldsymbol{v}\cdot\boldsymbol{w} = \|\boldsymbol{v}\|\,\|\boldsymbol{w}\|\cos\theta$ where $\theta \in [0,\pi]$ is the angle between $\boldsymbol{v}$ and $\boldsymbol{w}$

Why? You can derive this by repeating what we did, right after Definition 29 to show that $\boldsymbol{v}$ and $\boldsymbol{w}$ are orthogonal if and only if $\boldsymbol{v}\cdot\boldsymbol{w} = 0$. Just replace Pythagoras with the law of cosines ($c^2 = a^2 + b^2 - 2ab\cos\theta$ holds in any triangle!).

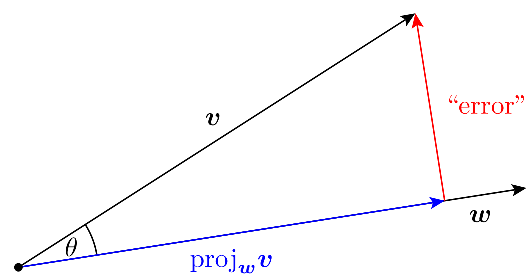
Two obvious cases. Observe that the cases $\theta = 0$ and $\theta = 90°$ are clearly true.

We will not discuss angles much further in this class. Just in case it is helpful, here is the typical argument given in Calculus 3 to determine the projection $\mathrm{proj}_{\boldsymbol{w}}\boldsymbol{v}$ of $\boldsymbol{v}$ onto $\boldsymbol{w}$:

From the sketch, we see that "error" $= \boldsymbol{v} - \mathrm{proj}_{\boldsymbol{w}}\boldsymbol{v}$ and that this error is orthogonal to $\boldsymbol{w}$.

Basic trigonometry tells us that the length of $\mathrm{proj}_{\boldsymbol{w}}\boldsymbol{v}$ is $\|\boldsymbol{v}\|\cos\theta$. Hence:

$$\mathrm{proj}_{\boldsymbol{w}}\boldsymbol{v} = \underbrace{\|\boldsymbol{v}\|\cos\theta}_{\text{length}}\ \underbrace{\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|}}_{\text{direction}} = \frac{\|\boldsymbol{v}\|\,\|\boldsymbol{w}\|\cos\theta}{\|\boldsymbol{w}\|}\frac{\boldsymbol{w}}{\|\boldsymbol{w}\|} = \left(\frac{\boldsymbol{v}\cdot\boldsymbol{w}}{\|\boldsymbol{w}\|^2}\right)\boldsymbol{w}$$



Orthogonal bases

Review. Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are a **basis** for V.

$\iff$ $V = \mathrm{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ and $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent.

$\iff$ Any vector $\boldsymbol{w}$ in V can be written as $\boldsymbol{w} = c_1\boldsymbol{v}_1 + \ldots + c_n\boldsymbol{v}_n$ in a unique way.

The latter is the practical reason why we care so much about bases!

V could be some abstract vector space (of polynomials or Fourier series), meaning that vectors are abstract objects and not just our usual column vectors. However, as soon as we pick a basis of V, then we can represent every (abstract) vector $\boldsymbol{w}$ by the (usual) column vector $(c_1, c_2, \ldots, c_n)^T$.

This means all of our results can be used, too, when working with these abstract spaces!

Definition 59. A basis $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ of a vector space V is an **orthogonal basis** if the vectors are (pairwise) orthogonal. If, in addition, the basis vectors have length 1, then this is called an **orthonormal basis**.

Armin Straub
straub@southalabama.edu

Example 60. The standard basis $\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ is an orthonormal basis for $\mathbb{R}^3$.

Example 61. Are the vectors $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ an orthogonal basis for $\mathbb{R}^3$? Is it orthonormal?

Solution. $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 0$, $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$, $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 0$.

So, this is an orthogonal basis.

On the other hand, the vectors do not all have length 1, so that this basis is not orthonormal.

Note. Orthogonal vectors are always linearly independent (see next class). Here, this certifies that the three vectors are linearly independent (and hence a basis for $\mathbb{R}^3$).

Normalize the vectors to produce an orthonormal basis.

Solution.

$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ has length $\sqrt{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} = \sqrt{2} \implies$ normalized: $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ has length $\sqrt{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} = \sqrt{2} \implies$ normalized: $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ has length $\sqrt{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} = 1 \implies$ is already normalized: $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

The resulting orthonormal basis is $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Theorem 62. Suppose that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are nonzero and pairwise orthogonal. Then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are linearly independent.

Proof. Suppose that $c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n = \boldsymbol{0}$. In order to show that $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are independent, we need to show that $c_1 = c_2 = \ldots = c_n = 0$.

Take the dot product of $\boldsymbol{v}_1$ with both sides:

$$\begin{aligned} 0 &= \boldsymbol{v}_1 \cdot (c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n) \\ &= c_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 + c_2 \boldsymbol{v}_1 \cdot \boldsymbol{v}_2 + \ldots + c_n \boldsymbol{v}_1 \cdot \boldsymbol{v}_n \\ &= c_1 \boldsymbol{v}_1 \cdot \boldsymbol{v}_1 = c_1 \|\boldsymbol{v}_1\|^2 \end{aligned}$$

But $\|\boldsymbol{v}_1\| \neq 0$ and hence $c_1 = 0$. Likewise, we find $c_2 = 0$, …, $c_n = 0$. Hence, the vectors are independent. □

Comment. Note that this result is intuitively obvious: if the vectors were linearly dependent, then one of them could be written as a linear combination of the others. However, all these other vectors (and hence any combination of them) are orthogonal to it.

Armin Straub
straub@southalabama.edu

Orthogonal projections if we have an orthogonal basis

Lemma 63. (orthogonal projection if we have an orthogonal basis)

If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are orthogonal, then the orthogonal projection of $\boldsymbol{w}$ onto $\operatorname{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is

$$\hat{\boldsymbol{w}} = \underbrace{\frac{\boldsymbol{w} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1} \boldsymbol{v}_1}_{\substack{\text{proj of } \boldsymbol{w} \\ \text{onto } \boldsymbol{v}_1}} + \ldots + \underbrace{\frac{\boldsymbol{w} \cdot \boldsymbol{v}_n}{\boldsymbol{v}_n \cdot \boldsymbol{v}_n} \boldsymbol{v}_n}_{\substack{\text{proj of } \boldsymbol{w} \\ \text{onto } \boldsymbol{v}_n}}.$$

Proof. It suffices to show that the error $\boldsymbol{w} - \hat{\boldsymbol{w}}$ is orthogonal to each $\boldsymbol{v}_i$. Indeed:

$$(\boldsymbol{w} - \hat{\boldsymbol{w}}) \cdot \boldsymbol{v}_i = \left(\boldsymbol{w} - \frac{\boldsymbol{w} \cdot \boldsymbol{v}_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1} \boldsymbol{v}_1 - \ldots - \frac{\boldsymbol{w} \cdot \boldsymbol{v}_n}{\boldsymbol{v}_n \cdot \boldsymbol{v}_n} \boldsymbol{v}_n \right) \cdot \boldsymbol{v}_i = \boldsymbol{w} \cdot \boldsymbol{v}_i - \frac{\boldsymbol{w} \cdot \boldsymbol{v}_i}{\boldsymbol{v}_i \cdot \boldsymbol{v}_i} \boldsymbol{v}_i \cdot \boldsymbol{v}_i = 0.$$

Alternatively, can you deduce the formula (say, in the case of an orthonormal basis) from our earlier formula for the projection matrix? □

Important consequence. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ is an orthogonal basis of V, and $\boldsymbol{w}$ is in V, then

$$\boldsymbol{w} = c_1 \boldsymbol{v}_1 + \ldots + c_n \boldsymbol{v}_n \qquad \text{with} \quad c_j = \frac{\boldsymbol{w} \cdot \boldsymbol{v}_j}{\boldsymbol{v}_j \cdot \boldsymbol{v}_j}.$$

If the $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are a basis, but not orthogonal, then we have to solve a system of equations to find the c_i. That is a lot more work than simply computing a few dot products.

Note. In other words, $\boldsymbol{w}$ decomposes as the sum of its projections onto each basis vector.

Note. If $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are orthonormal, then the denominators are all 1.

Example 64. What is the projection of $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ onto $W = \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ with $\boldsymbol{v}_1 = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$, $\boldsymbol{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$?

Comment. We know how to do this using least squares. (Do it for practice!)

However, realizing that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are orthogonal makes things easier.

[Actually, here, it is obvious what the projection is going to be if we realized that W is the x-y-plane.]

Solution. (using orthogonality) Because $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are orthogonal, the projection is

$$\underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{\text{projection onto } \boldsymbol{v}_1} + \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\text{projection onto } \boldsymbol{v}_2} = \frac{-4}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{10}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 0 \end{bmatrix}.$$

Important note. Note that, at this point, we can easily extend $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ to an orthogonal basis of $\mathbb{R}^3$:

That is because the error $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} - \begin{bmatrix} 3 \\ 7 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$ is orthogonal to both of the existing basis vectors.

Therefore $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$ is an orthogonal basis of $\mathbb{R}^3$.

This observation underlies the Gram-Schmidt process, which we will discuss next class.

Armin Straub
straub@southalabama.edu

Example 65. Express $\underbrace{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}}_{\boldsymbol{x}}$ in terms of the basis $\underbrace{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{\boldsymbol{v}_1}, \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\boldsymbol{v}_2}, \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\boldsymbol{v}_3}$.

Solution. Because $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ is an orthogonal basis of $\mathbb{R}^3$, we get (much as in the previous example):

$$\begin{aligned} \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} &= c_1 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\ &= \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}}_{\text{projection of } \boldsymbol{x} \text{ onto } \boldsymbol{v}_1} + \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\text{projection of } \boldsymbol{x} \text{ onto } \boldsymbol{v}_2} + \underbrace{\frac{\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\text{projection of } \boldsymbol{x} \text{ onto } \boldsymbol{v}_3} \\ &= \frac{-4}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + \frac{10}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \frac{4}{1} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{aligned}$$

Because we spelled out all the details this looks more involved than it is. We only computed 6 dot products!

Alternative. We could have solved $\begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ to also find $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} -2 \\ 5 \\ 4 \end{bmatrix}$.

The numbers are particularly easy here but in general, to find this solution, we have to go through the entire process of Gaussian elimination. On the other hand, if we have an orthogonal basis, the former approach requires less work, because it is just computing a few dot products.

Example 66. Express $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$ in terms of the basis $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$.

Solution. This is not an orthogonal basis, so we cannot proceed as in the previous example.

To write $\begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$, we need to solve $\begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \\ 4 \end{bmatrix}$.

Solving that system (do it!), we find $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 0 \end{bmatrix}$.