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

Lemma 50. If the columns of a matrix A are independent, then $A^T A$ is invertible.

Proof. Assume $A^T A$ is not invertible, so that $A^T A\boldsymbol{x} = \mathbf{0}$ for some $\boldsymbol{x} \neq \mathbf{0}$. Multiply both sides with $\boldsymbol{x}^T$ to get

$$\boldsymbol{x}^T A^T A\boldsymbol{x} = (A\boldsymbol{x})^T A\boldsymbol{x} = \|A\boldsymbol{x}\|^2 = 0,$$

which implies that $A\boldsymbol{x} = 0$. Since the columns of A are independent, this shows that $\boldsymbol{x} = \mathbf{0}$. A contradiction! □

Example 51. If P is a projection matrix, then what is P^2?

For instance. For P as in Example 49, $P^2 = \frac{1}{4}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix}^2 = \frac{1}{2}\begin{bmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} = P$.

Solution. Can you see why it is always true that $P^2 = P$?
[Recall that P projects a vector onto a space W (actually, $W = \mathrm{col}(P)$). Hence P^2 takes a vector $\boldsymbol{b}$, projects it onto W to get $\hat{\boldsymbol{b}}$, and then projects $\hat{\boldsymbol{b}}$ onto W again. But the projection of $\hat{\boldsymbol{b}}$ onto W is just $\hat{\boldsymbol{b}}$ (why?!), so that P^2 always has the exact same effect as P. Therefore, $P^2 = P$.]

Example 52. True or false? If P is the matrix for projecting onto W, then $W = \mathrm{col}(P)$.

Solution. True!
Why? The columns of P are the projections of the standard basis vectors and hence in W. On the other hand, for any vector $\boldsymbol{w}$ in W, we have $P\boldsymbol{w} = \boldsymbol{w}$ so that $\boldsymbol{w}$ is a combination of the columns of P.
[This may take several readings to digest but do read (or ask) until it makes sense!]

In particular. $\mathrm{rank}(P) = \dim W$ (because, for any matrix, $\mathrm{rank}(A) = \dim \mathrm{col}(A)$)

Review.

- Vectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ are **linearly independent**.

  $\iff$ $c_1\boldsymbol{v}_1 + \ldots + c_n\boldsymbol{v}_n = \mathbf{0}$ only has the (trivial) solution $c_1 = c_2 = \ldots = c_n = 0$.

- 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!

Armin Straub
straub@southalabama.edu

Orthogonal bases

Theorem 53. 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

$$c_1\boldsymbol{v}_1 + \ldots + c_n\boldsymbol{v}_n = \boldsymbol{0}.$$

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.

Definition 54. 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**.

Example 55. 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 56. 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.
Note that we do not need to check that the three vectors are independent. That follows from their orthogonality (see Theorem 53).
On the other hand, the vectors do not all have length 1, so that this basis is not orthonormal.

Normalize the vectors to produce an orthonormal basis.

Solution.
$\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ has length $\sqrt{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}} = \sqrt{2}$ $\Longrightarrow$ normalized: $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$
$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$ has length $\sqrt{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}} = \sqrt{2}$ $\Longrightarrow$ normalized: $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$
$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ has length $\sqrt{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}} = 1$ $\Longrightarrow$ is already normalized: $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$
The resulting orthonormal basis is $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Armin Straub
straub@southalabama.edu

Lemma 57. (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

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

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

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

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.

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

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

□

Example 58. 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:

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

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.

Armin Straub
straub@southalabama.edu