# Chapter 14

# Projections

Given a line $\ell$ and a point $\boldsymbol{x}$ not on the line, you may recall the geometric problem of finding the point $\boldsymbol{p}$ on $\ell$ which is closest to $\boldsymbol{x}$. (See figure 14.1.)

**Example.** ***Find the point p in the span of*** $\left\{\begin{pmatrix}2\\1\end{pmatrix}\right\}$ ***which is closest to*** $\boldsymbol{x} = \begin{pmatrix}1\\2\end{pmatrix}$.

We want to find the point $\boldsymbol{p} = a\begin{pmatrix}2\\1\end{pmatrix}$ such that $\|\boldsymbol{x} - \boldsymbol{p}\|$ is as small as possible. We will do this two different ways.

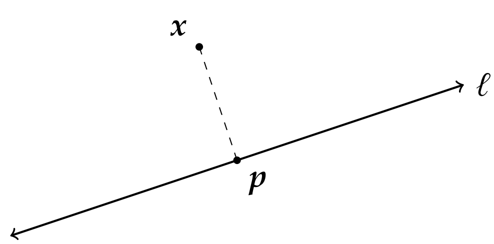


**Figure 14.1:** *The point on $\ell$ closest to $\boldsymbol{x}$.*

**Method 1.** We will find the value of $a$ which minimizes

$$\begin{aligned}\|\boldsymbol{x}-\boldsymbol{p}\|^2 &= \left\|\begin{pmatrix}2\\1\end{pmatrix} - a\begin{pmatrix}1\\2\end{pmatrix}\right\|^2\\ &= \left\|\begin{pmatrix}2-a\\1-2a\end{pmatrix}\right\|^2\\ &= (2-a)^2 + (1-2a)^2 = 5a^2 - 8a + 5.\end{aligned}$$

This second degree polynomial will have a minimum at $a = \frac{8}{2\cdot 5} = \frac{4}{5}$, and so $\boldsymbol{p} = \frac{4}{5}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}8/5\\4/5\end{pmatrix}$.

**Method 2.** In geometry, we would drop a perpendicular from $\boldsymbol{x}$ to $\ell$. We will do that here, we will find $\boldsymbol{p}$ such that $\boldsymbol{x}-\boldsymbol{p}$ is perpendicular to $\ell$; i.e., so that

$$\begin{pmatrix}2-a\\1-2a\end{pmatrix}\cdot\begin{pmatrix}2\\1\end{pmatrix} = 0.$$

This equation is simply

$$4 - 2a + 1 - 2a = 0$$

or simply

$$5 - 4a = 0.$$

The solution is $a = 4/5$, and we get the point $\boldsymbol{p} = \frac{4}{5}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}8/5\\4/5\end{pmatrix}$. ☑

**Definition.** Let $V$ be a subspace of $\mathbb{R}^n$. For any $\boldsymbol{x} \in \mathbb{R}^n$, the vector $\boldsymbol{p} \in V$ which is closest to $\boldsymbol{x}$ (i.e., which minimizes $\|\boldsymbol{x}-\boldsymbol{p}\|$) is called the ***projection*** of $\boldsymbol{x}$ onto $V$. □

As in the above example, we can also characterize the projection as the point $\boldsymbol{p}$ in $V$ such that $\boldsymbol{x}-\boldsymbol{p}$ is perpendicular to $V$; i.e., which is perpendicular to every vector in $V$. This geometric interpretation is the one we will use.

Suppose $V$ is a subspace of $\mathbb{R}^n$, and $\mathcal{B} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is an orthogonal basis for $V$. Let $\boldsymbol{x} \in \mathbb{R}^n$. What is the projection $\boldsymbol{p}$ of $\boldsymbol{x}$ onto $V$? Since

$\boldsymbol{p} \in V$, we know $\boldsymbol{p} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$ for some $c_1, \ldots, c_k$. Since $vectx - \boldsymbol{p}$ is perpendicular to any vector in $V$, we know that

$$\begin{aligned}
(\boldsymbol{x} - \boldsymbol{p}) \cdot \boldsymbol{v}_1 &= 0 \\
(\boldsymbol{x} - \boldsymbol{p}) \cdot \boldsymbol{v}_2 &= 0 \\
&\cdots \\
(\boldsymbol{x} - \boldsymbol{p}) \cdot \boldsymbol{v}_k &= 0
\end{aligned}$$

Now,

$$\begin{aligned}
0 &= (vectx - \boldsymbol{p}) \cdot \boldsymbol{v}_1 \\
&= (\boldsymbol{x} - c_1\boldsymbol{v}_1 - \cdots - c_k\boldsymbol{v}_k) \cdot \boldsymbol{v}_1 \\
&= \boldsymbol{x} \cdot \boldsymbol{v}_1 - c_1\boldsymbol{v}_1 \cdot \boldsymbol{v}_1 - \cdots - c_k\boldsymbol{v}_k \cdot \boldsymbol{v}_1 \\
&= \boldsymbol{x} \cdot \boldsymbol{v}_1 - c_1\|vectv_1\|^2 - 0 - \cdots - 0
\end{aligned}$$

since the $\boldsymbol{v}_i$ are orthogonal. This means that

$$c_1\|\boldsymbol{v}_1\|^2 = \boldsymbol{x} \cdot \boldsymbol{v}_1,$$

and so

$$c_1 = \frac{\boldsymbol{x} \cdot \boldsymbol{v}_1}{\|\boldsymbol{v}_1\|^2}.$$

We can similarly find the other constants.

**Proposition.** *Let $V$ be a subspace of $\mathbb{R}^n$ with orthogonal basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$. For any $\boldsymbol{x} \in \mathbb{R}^n$, the projection of $\boldsymbol{x}$ onto $V$ is*

$$\boldsymbol{p} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k,$$

*where*

$$c_i = \frac{\boldsymbol{x} \cdot \boldsymbol{v}_i}{\|\boldsymbol{v}_i\|^2}$$

*for $i = 1, \ldots, k$.* □

**Example.** ***Let $V$ be the subspace of $\mathbb{R}^3$ with orthogonal basis $\{v_1, v_2\}$ for $v_1 = \begin{pmatrix}1\\2\\3\end{pmatrix}$ and $v_2 = \begin{pmatrix}-1\\-1\\1\end{pmatrix}$. Find the projection of $x = \begin{pmatrix}2\\3\\4\end{pmatrix}$ onto $V$.***

The projection will be

$$\boldsymbol{p} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2.$$

To find $c_1$ and $c_2$, we need

$$\begin{aligned}
\boldsymbol{x} \cdot \boldsymbol{v}_1 &= 22 \\
\|\boldsymbol{v}_1\|^2 &= 14 \\
\boldsymbol{x} \cdot \boldsymbol{v}_2 &= -1 \\
\|\boldsymbol{v}_1\|^2 &= 3.
\end{aligned}$$

So $c_1 = 22/14 = 11/7$ and $c_2 = -1/3$. So the projection is

$$\boldsymbol{p} = \frac{11}{7}\begin{pmatrix}1\\2\\3\end{pmatrix} - \frac{1}{3}\begin{pmatrix}-1\\-1\\1\end{pmatrix} = \begin{pmatrix}40/21\\73/21\\92/21\end{pmatrix}.$$

☑

Suppose, next, that we don't have an orthogonal basis for $V$? We could turn any basis into an orthogonal basis, using a technique called the Gram-Schmidt procedure. But for now, we'll look at a different approach.

Suppose $V$ is a subspace of $\mathbb{R}^n$ with basis $\mathcal{B} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$, which is not necessarily orthogonal. For $\boldsymbol{x} \in \mathbb{R}^n$, what is the projection $\boldsymbol{p}$ of $\boldsymbol{x}$ onto $V$? We will again have $\boldsymbol{p} = c_1\boldsymbol{v}_1 + \cdots + c_k\boldsymbol{v}_k$ for some $c_1, \dots, c_k$, and again $vectx - \boldsymbol{p}$ will be perpendicular to any vector in $V$. So we have

$$\begin{aligned}
\boldsymbol{v}_1 \cdot (\boldsymbol{x} - \boldsymbol{p}) &= 0 \\
\boldsymbol{v}_2 \cdot (\boldsymbol{x} - \boldsymbol{p}) &= 0 \\
&\dots \\
\boldsymbol{v}_k \cdot (\boldsymbol{x} - \boldsymbol{p}) &= 0.
\end{aligned}$$

Let's right the dot product $\boldsymbol{a} \cdot \boldsymbol{b}$ as a matrix product $\boldsymbol{a}^T\boldsymbol{b}$. The above equations become

$$\begin{aligned}
\boldsymbol{v}_1^T(\boldsymbol{x} - \boldsymbol{p}) &= 0 \\
\boldsymbol{v}_2^T(\boldsymbol{x} - \boldsymbol{p}) &= 0 \\
&\dots \\
\boldsymbol{v}_k^T(\boldsymbol{x} - \boldsymbol{p}) &= 0.
\end{aligned}$$

We can write this system of matrix equations as a single equation:

$$\begin{pmatrix} \boldsymbol{v}_1^T \\ \boldsymbol{v}_2^T \\ \vdots \\ \boldsymbol{v}_k^T \end{pmatrix} (\boldsymbol{x} - \boldsymbol{p}) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

or simply

$$A^T(\boldsymbol{x} - \boldsymbol{p}) = 0$$

for

$$A = (\boldsymbol{v}_1 \boldsymbol{v}_2 \ldots \boldsymbol{v}_k).$$

This gives us

$$A^T\boldsymbol{x} = A^T\boldsymbol{p}.$$

Since

$$\boldsymbol{p} = A\boldsymbol{c}$$

we want to find the vector $\boldsymbol{c}$ such that

$$A^T\boldsymbol{x} = A^T A\boldsymbol{c}.$$

While $A$ may not be a square matrix, $A^T A$ will be square.

**Proposition.** *If the columns of the matrix $A$ are linearly independent, then $A^T A$ is invertible.* □

Since $A^T A$ is invertible, we can solve $A^T\boldsymbol{x} = A^T A\boldsymbol{c}$ for $\boldsymbol{c}$ and get $c = (A^T A)^{-1} A^T \boldsymbol{x}$, and so the projection is

$$\boldsymbol{p} = Ac = A(A^T A)^{-1} A^T \boldsymbol{x}.$$

**Proposition.** *Let $V$ be a subspace of $\mathbb{R}^n$ with basis $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$, and let $A = (\boldsymbol{v}_1 \boldsymbol{v}_2 \ldots \boldsymbol{v}_k)$. For any $\boldsymbol{x} \in \mathbb{R}^n$, the projection of $\boldsymbol{x}$ onto $V$ is*

$$\boldsymbol{p} = P\boldsymbol{x},$$

*where*

$$P = A(A^T A)^{-1} A^T.$$ □

**Example.** *Let $V$ be the subspace of $\mathbb{R}^3$ with basis $\left\{\begin{pmatrix}1\\0\\1\end{pmatrix},\begin{pmatrix}1\\1\\0\end{pmatrix}\right\}$. Find a matrix $P$ such that for any $x \in \mathbb{R}^3$, the projection of $x$ onto $V$ is given by $Px$.*
Here,

$$A = \begin{pmatrix}1 & 1\\0 & 1\\1 & 0\end{pmatrix},$$

so

$$A^T = \begin{pmatrix}1 & 0 & 1\\1 & 1 & 0\end{pmatrix}$$

and

$$A^TA = \begin{pmatrix}2 & 1\\1 & 2\end{pmatrix}$$

So

$$(A^TA)^{-1} = \frac{1}{3}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}$$

and we get

$$\begin{aligned}P &= A(A^TA)^{-1}A^T\\ &= \frac{1}{3}\begin{pmatrix}1 & 1\\0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}2 & -1\\-1 & 2\end{pmatrix}\begin{pmatrix}1 & 0 & 1\\1 & 1 & 0\end{pmatrix}\\ &= \frac{1}{3}\begin{pmatrix}2 & 1 & 1\\1 & 2 & -1\\1 & -1 & 2\end{pmatrix}\end{aligned}$$

Let's find the projections for some specific $\boldsymbol{x}$s. For $\boldsymbol{x} = \begin{pmatrix}1\\2\\3\end{pmatrix}$, we get the projection

$$P\boldsymbol{x} = \frac{1}{3}\begin{pmatrix}7\\2\\3\end{pmatrix}.$$

For $\boldsymbol{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}$, we get the projection

$$P\boldsymbol{x} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Note that the projection is the same as the vector $\boldsymbol{x}$, indicating that $\boldsymbol{x}$ is already in the space $V$. ☑

Now let's apply projections to solving equations. Suppose the system

$$A\boldsymbol{x} = \boldsymbol{b}$$

has no solutions. That means that $\boldsymbol{b}$ is not in the column space of $A$. We might not want to leave it at that, though; perhaps there *should* be a solution, but there isn't because of some rounding errors or mismeasurements. We will then try to find $\boldsymbol{x}$ which makes $A\boldsymbol{x}$ as close as possible to $\boldsymbol{b}$; in other words, find $\boldsymbol{x}$ for which $A\boldsymbol{x}$ is the projection of $\boldsymbol{b}$ on the column space of $A$. For that, we want $A\boldsymbol{x} - \boldsymbol{b}$ to be orthogonal to the column space of $A$, which means that

$$A^T(A\boldsymbol{x} - \boldsymbol{b}) = 0.$$

This will always have a solution.

To summarize; to find the least squares solution of

$$A\boldsymbol{x} = \boldsymbol{b}$$

solve

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b}.$$

**Example.** ***The system of equations***

$$\begin{array}{rcrcr} 2x & + & y & = & -1 \\ 3x & - & y & = & 2 \\ x & - & y & = & -3 \end{array}$$

***has no solution. Find the least squares solution***

Our system is

$$A\begin{pmatrix} x \\ y \end{pmatrix} = \boldsymbol{b},$$

where

$$A = \begin{pmatrix} 2 & 1 \\ 3 & -1 \\ 1 & -1 \end{pmatrix}$$

and

$$\boldsymbol{b} = \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix}.$$

We want to solve

$$A^T A \begin{pmatrix} x \\ y \end{pmatrix} = A^T \boldsymbol{b}.$$

Since

$$A^T = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -1 & -1 \end{pmatrix},$$

we get

$$A^T A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 3 & -1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 14 & -2 \\ -2 & 3 \end{pmatrix}$$

and

$$A^T \boldsymbol{b} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Our new system

$$A^T A \begin{pmatrix} x \\ y \end{pmatrix} = A^T \boldsymbol{b}$$

has augmented matrix

$$\begin{pmatrix} 14 & -2 & 1 \\ -2 & 3 & 0 \end{pmatrix},$$

which reduces to the echelon matrix

$$\begin{pmatrix} 1 & -3/2 & 0 \\ 0 & 1 & 1/19 \end{pmatrix}.$$

The system

$$\begin{aligned} x \; - \; \tfrac{3}{2} y &= 0 \\ y &= \tfrac{1}{19} \end{aligned}$$

has solution $x = 3/38$, $y = 1/19$. ☑

A standard use of this is finding lines of best fit. Suppose you have several points, $(x_1, y_1), \ldots, (x_n, y_n)$ which are supposed to lie on a line $y = mx + b$. If the points are not collinear, what line comes closest to passing through all of them? And what does that even mean?

Putting the points into the formula $y = mx + b$, we get the simultaneous equations in the unknowns $m$ and $b$:

$$\begin{array}{ccccc} mx_1 & + & b & = & y_1 \\ \vdots & & & & \vdots \\ mx_n & + & b & = & y_n \end{array}$$

We can find the least squares solution to these equations. This system can be written

$$A \begin{pmatrix} m \\ b \end{pmatrix} = \boldsymbol{y},$$

where

$$A = \begin{pmatrix} x_1 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{pmatrix}$$

and

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

The least squares solution will be the solution of

$$A^T A \begin{pmatrix} m \\ b \end{pmatrix} = A^T \boldsymbol{y}.$$

Solving this will lead us to formulas for $m$ and $b$ which can be found in many statistic textbooks, but is may be just as simple to work with the matrices.

**Example.** ***Find the line $y = mx + b$ which best fits the points in the following table:***

| $x$ | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| $y$ | 2 | 4 | 4 | 6 |

Our matrix $A$ is

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 4 & 1 \end{pmatrix}$$

and

$$\boldsymbol{y} = \begin{pmatrix} 2 \\ 4 \\ 4 \\ 6 \end{pmatrix}.$$

So

$$A^T = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 1 & 1 & 1 \end{pmatrix}.$$

So

$$A^T A = \begin{pmatrix} 30 & 10 \\ 10 & 4 \end{pmatrix}$$

Since

$$A^T \boldsymbol{y} = \begin{pmatrix} 46 \\ 16 \end{pmatrix}$$

and

$$(A^T A)^{-1} = \frac{1}{20} \begin{pmatrix} 4 & -10 \\ -10 & 30 \end{pmatrix},$$

the solution of

$$A^T A \begin{pmatrix} m \\ b \end{pmatrix} = A^T \boldsymbol{y}$$

is

$$\begin{pmatrix} m \\ b \end{pmatrix} = (A^T A)^{-1} A^T \boldsymbol{y} = \begin{pmatrix} 6/5 \\ 1 \end{pmatrix}.$$

So the line is

$$y = \frac{6}{5}x + 1.$$