Orthogonality — Part 2: Projections

Lecture notes on section 4.2 of Strang, *Introduction to Linear Algebra*. Read with pen and paper beside you, work each example yourself before reading its solution, and do every exercise before opening the answer.

The chapter in three parts: Part 1 · Orthogonal vectors and subspaces | Part 2 · Projections | Part 3 · Least squares and orthonormal bases

Contents

0. Before you start
1. Projection onto a line
2. Projection onto a subspace
3. Question types and model answers
4. Practice test
5. Formula sheet

0 Before you start

Why we study this

Given a line or a plane through the origin and a point $\mathbf{b}$ outside it, which point of the line or plane is nearest to $\mathbf{b}$? The answer is the **projection** of $\mathbf{b}$. A GPS receiver, a robot arm and a photo editor all solve this problem constantly, and so does every statistician: the least-squares fit of Part 3 is a projection. The whole of this part follows from one sentence: **the error from $\mathbf{b}$ to its projection is perpendicular to the subspace.**

What you need

- From Part 1: two vectors are orthogonal when their dot product is zero; $\|\mathbf{v}\|^2 = \mathbf{v}^{\mathsf{T}}\mathbf{v}$; the column space $\mathbf{C}(A)$ of a matrix.
- From Chapter 2: solving two equations in two unknowns by elimination and back substitution, and the inverse of a 2×2 matrix.
- Matrix multiplication, in particular a column times a row (which gives a matrix) and a row times a column (which gives a number).

Notation

Symbol	Meaning
$\mathbf{a}$, $A = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 \end{bmatrix}$	the direction of the line; the matrix whose columns span the plane
$\hat{x}$, $\hat{\mathbf{x}}$	the number (or vector of numbers) of coefficients chosen by the projection
$\mathbf{p}$	the projection: the closest point, $\mathbf{p} = \hat{x}\,\mathbf{a}$ or $\mathbf{p} = A\hat{\mathbf{x}}$
$\mathbf{e}$	the error $\mathbf{e} = \mathbf{b} - \mathbf{p}$, perpendicular to the line or plane
P	a projection matrix: $\mathbf{p} = P\mathbf{b}$
$\mathbf{a}^\mathsf{T}\mathbf{a}$ and $\mathbf{a}\,\mathbf{a}^\mathsf{T}$	a number (row times column) and an $n \times n$ matrix (column times row)

1 Projection onto a line (4.2)

At noon the shadow of a vertical pole falls directly below its tip: the shadow is the point of the ground closest to the tip, and the line from the shadow to the tip is perpendicular to the ground. Projection is exactly this idea. The closest point is found by making the error perpendicular, and the rest of the chapter uses this one picture again and again.

The problem. A line passes through the origin in the direction of a vector $\mathbf{a}$. Given a vector $\mathbf{b}$, which point $\mathbf{p}$ on the line is closest to $\mathbf{b}$?

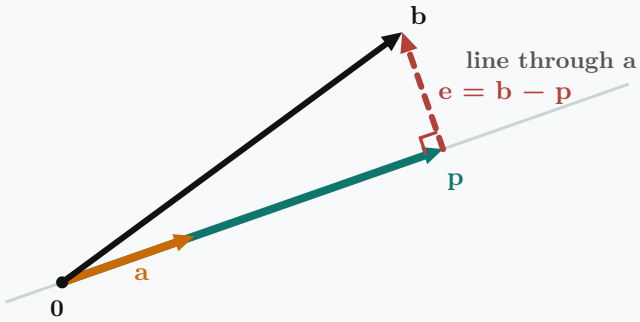


Figure 1.1 — projection onto a line. $\mathbf{p} = \hat{x}\,\mathbf{a}$ is the point of the line closest to $\mathbf{b}$; the error $\mathbf{e} = \mathbf{b} - \mathbf{p}$ meets the line at a right angle. The three vectors $\mathbf{p}$, $\mathbf{e}$, $\mathbf{b}$ form a right triangle, so $\|\mathbf{p}\|^2 + \|\mathbf{e}\|^2 = \|\mathbf{b}\|^2$.

Setting up. The point $\mathbf{p}$ lies on the line, so it is a multiple of $\mathbf{a}$: write $\mathbf{p} = \hat{x}\,\mathbf{a}$, where the number $\hat{x}$ is what we must find. The **error** is $\mathbf{e} = \mathbf{b} - \mathbf{p}$. The closest point is reached by dropping a perpendicular from $\mathbf{b}$ onto the line, so the error must be perpendicular to $\mathbf{a}$. That single condition determines $\hat{x}$.

Derivation of $\hat{x}$.

$$
\begin{aligned}
\mathbf{a}^{\mathsf{T}}\mathbf{e} &= 0 && \text{(the error is perpendicular to } \mathbf{a}\text{)} \\
\mathbf{a}^{\mathsf{T}}(\mathbf{b} - \hat{x}\,\mathbf{a}) &= 0 && \text{(}\mathbf{e} = \mathbf{b} - \mathbf{p} \text{ and } \mathbf{p} = \hat{x}\,\mathbf{a}\text{)} \\
\mathbf{a}^{\mathsf{T}}\mathbf{b} - \hat{x}\,(\mathbf{a}^{\mathsf{T}}\mathbf{a}) &= 0 && \text{(multiply out; } \hat{x} \text{ is a number, so it can be moved in front)} \\
\hat{x}\,(\mathbf{a}^{\mathsf{T}}\mathbf{a}) &= \mathbf{a}^{\mathsf{T}}\mathbf{b} && \text{(move the second term across)} \\
\hat{x} &= \frac{\mathbf{a}^{\mathsf{T}}\mathbf{b}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} && \text{(divide by } \mathbf{a}^{\mathsf{T}}\mathbf{a} = \|\mathbf{a}\|^2, \text{ which is not 0 because } \mathbf{a} \neq \mathbf{0}\text{)}
\end{aligned}
$$

Projection onto the line through $\mathbf{a}$.

$$
\begin{aligned}
\hat{x} &= \frac{\mathbf{a}^{\mathsf{T}}\mathbf{b}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} && \text{(a number)} \\
\mathbf{p} &= \hat{x}\,\mathbf{a} && \text{(the projection: the closest point on the line)} \\
\mathbf{e} &= \mathbf{b} - \mathbf{p} && \text{(the error, with } \mathbf{e}^{\mathsf{T}}\mathbf{a} = 0\text{)}
\end{aligned}
$$

Method 1.1 (four steps).

1. Compute the two dot products $\mathbf{a}^{\mathsf{T}}\mathbf{b}$ and $\mathbf{a}^{\mathsf{T}}\mathbf{a}$.
2. Compute the number $\hat{x} = \mathbf{a}^{\mathsf{T}}\mathbf{b} \,/\, \mathbf{a}^{\mathsf{T}}\mathbf{a}$. Keep it as a fraction.
3. Compute $\mathbf{p} = \hat{x}\,\mathbf{a}$: multiply every entry of $\mathbf{a}$ by $\hat{x}$.
4. Compute $\mathbf{e} = \mathbf{b} - \mathbf{p}$ and **check** that $\mathbf{e}^{\mathsf{T}}\mathbf{a} = 0$.

Example 1.1 (book). **Project $\mathbf{b} = (1, 1, 1)$ onto the line through $\mathbf{a} = (1, 2, 2)$.**

1. $\mathbf{a}^{\mathsf{T}}\mathbf{b} = 1 \cdot 1 + 2 \cdot 1 + 2 \cdot 1 = 5$ and $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + 4 + 4 = 9$.
2. $\hat{x} = \dfrac{5}{9}$.
3. $\mathbf{p} = \dfrac{5}{9}\begin{bmatrix}1\\2\\2\end{bmatrix} = \begin{bmatrix}5/9\\10/9\\10/9\end{bmatrix}$.
4. $\mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix}1\\1\\1\end{bmatrix} - \begin{bmatrix}5/9\\10/9\\10/9\end{bmatrix} = \begin{bmatrix}4/9\\-1/9\\-1/9\end{bmatrix}$. Check: $\mathbf{e}^{\mathsf{T}}\mathbf{a} = \dfrac{4}{9} - \dfrac{2}{9} - \dfrac{2}{9} = 0$. ✓

A second check with Pythagoras: $\|\mathbf{p}\|^2 = \dfrac{25 + 100 + 100}{81} = \dfrac{25}{9}$, $\|\mathbf{e}\|^2 = \dfrac{16 + 1 + 1}{81} = \dfrac{2}{9}$, and $\dfrac{25}{9} + \dfrac{2}{9} = 3 = \|\mathbf{b}\|^2$. ✓

Example 1.2 (book, worked example 4.2A). **Project $\mathbf{b} = (3, 4, 4)$ onto the line through $\mathbf{a} = (2, 2, 1)$.**

1. $\mathbf{a}^{\mathsf{T}}\mathbf{b} = 2 \cdot 3 + 2 \cdot 4 + 1 \cdot 4 = 6 + 8 + 4 = 18$ and $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 4 + 4 + 1 = 9$.
2. $\hat{x} = \dfrac{18}{9} = 2$.
3. $\mathbf{p} = 2\begin{bmatrix}2\\2\\1\end{bmatrix} = \begin{bmatrix}4\\4\\2\end{bmatrix}$.
4. $\mathbf{e} = \begin{bmatrix}3\\4\\4\end{bmatrix} - \begin{bmatrix}4\\4\\2\end{bmatrix} = \begin{bmatrix}-1\\0\\2\end{bmatrix}$. Check: $\mathbf{e}^{\mathsf{T}}\mathbf{a} = -2 + 0 + 2 = 0$. ✓

Pythagoras: $\|\mathbf{p}\|^2 + \|\mathbf{e}\|^2 = 36 + 5 = 41 = 9 + 16 + 16 = \|\mathbf{b}\|^2$. ✓

Example 1.3 (book, problem 1b). **Project $\mathbf{b} = (1, 3, 1)$ onto the line through $\mathbf{a} = (-1, -3, -1)$.**

1. $\mathbf{a}^{\mathsf{T}}\mathbf{b} = -1 - 9 - 1 = -11$ and $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + 9 + 1 = 11$.
2. $\hat{x} = \dfrac{-11}{11} = -1$.
3. $\mathbf{p} = -1 \begin{bmatrix} -1 \\ -3 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \mathbf{b}$.
4. $\mathbf{e} = \mathbf{b} - \mathbf{p} = \mathbf{0}$. The check $\mathbf{e}^{\mathsf{T}}\mathbf{a} = 0$ is automatic.

The vector $\mathbf{b}$ was already on the line ($\mathbf{b} = -\mathbf{a}$), so its closest point is itself.

Two special cases.

- If $\mathbf{b}$ is on the line, $\mathbf{b} = c\,\mathbf{a}$: then $\hat{x} = c$, $\mathbf{p} = \mathbf{b}$, $\mathbf{e} = \mathbf{0}$.
- If $\mathbf{b}$ is perpendicular to $\mathbf{a}$: then $\mathbf{a}^{\mathsf{T}}\mathbf{b} = 0$, $\hat{x} = 0$, $\mathbf{p} = \mathbf{0}$, $\mathbf{e} = \mathbf{b}$.

In every case $\mathbf{p}$, $\mathbf{e}$, $\mathbf{b}$ form a right triangle, so $\|\mathbf{p}\|^2 + \|\mathbf{e}\|^2 = \|\mathbf{b}\|^2$ and $\mathbf{p}$ is never longer than $\mathbf{b}$.

The projection matrix

Projection is a linear operation, so some matrix P must produce it: $\mathbf{p} = P\mathbf{b}$. To find P, write the number $\hat{x}$ on the right of $\mathbf{a}$ and regroup:

$$\mathbf{p} = \mathbf{a}\,\hat{x} = \mathbf{a}\,\frac{\mathbf{a}^{\mathsf{T}}\mathbf{b}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} = \frac{\mathbf{a}\,\mathbf{a}^{\mathsf{T}}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}}\,\mathbf{b} = P\mathbf{b}.$$

Projection matrix onto the line through $\mathbf{a}$.

$$P = \frac{\mathbf{a}\,\mathbf{a}^{\mathsf{T}}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} \qquad (\text{an } n \times n \text{ matrix: a column times a row, divided by a number}).$$

Note the difference: $\mathbf{a}^{\mathsf{T}}\mathbf{a}$ is a row times a column, a number. $\mathbf{a}\,\mathbf{a}^{\mathsf{T}}$ is a column times a row, an $n \times n$ matrix whose (i, j) entry is $a_i a_j$.

Example 1.4. **Find the projection matrix onto the line through $\mathbf{a} = (1, 2, 2)$, and use it to recover Example 1.1.**

$$P = \frac{1}{9}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 2 & 2 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}$$

$$P\mathbf{b} = \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix} = \mathbf{p} \quad \checkmark$$

Theorem 1.2 (properties of P). The projection matrix $P = \mathbf{a}\,\mathbf{a}^{\mathsf{T}}/\mathbf{a}^{\mathsf{T}}\mathbf{a}$ satisfies

$$P^{\mathsf{T}} = P, \qquad P^2 = P, \qquad \operatorname{rank} P = 1.$$

Proof. $P^{\mathsf{T}} = (\mathbf{a}\,\mathbf{a}^{\mathsf{T}})^{\mathsf{T}}/\mathbf{a}^{\mathsf{T}}\mathbf{a} = \mathbf{a}\,\mathbf{a}^{\mathsf{T}}/\mathbf{a}^{\mathsf{T}}\mathbf{a} = P$. For P^2, regroup the middle factor, which is the number $\mathbf{a}^{\mathsf{T}}\mathbf{a}$:

$$P^2 = \frac{\mathbf{a}\,(\mathbf{a}^{\mathsf{T}}\mathbf{a})\,\mathbf{a}^{\mathsf{T}}}{(\mathbf{a}^{\mathsf{T}}\mathbf{a})^2} = \frac{\mathbf{a}\,\mathbf{a}^{\mathsf{T}}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} = P.$$

Every column of P is a multiple of $\mathbf{a}$, so the rank is 1. ∎

The meaning of $P^2 = P$: projecting a second time changes nothing, because $\mathbf{p}$ is already on the line.

Exercise 1.

a. Project $\mathbf{b} = (2, 2, 2)$ onto the line through $\mathbf{a} = (1, 0, 1)$. Give $\hat{x}$, $\mathbf{p}$, $\mathbf{e}$ and the check.

b. Write the projection matrix for that line and verify that $P\mathbf{b} = \mathbf{p}$.

ANSWER

a. $\mathbf{a}^{\mathsf{T}}\mathbf{b} = 2 + 0 + 2 = 4$, $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + 0 + 1 = 2$, so $\hat{x} = 4/2 = 2$.
$\mathbf{p} = 2(1, 0, 1) = (2, 0, 2)$. $\mathbf{e} = (2, 2, 2) - (2, 0, 2) = (0, 2, 0)$. Check:
$\mathbf{e}^{\mathsf{T}}\mathbf{a} = 0 + 0 + 0 = 0$. ✓

b. $\mathbf{a}\,\mathbf{a}^{\mathsf{T}} = \begin{bmatrix}1\\0\\1\end{bmatrix}\begin{bmatrix}1 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 1\\0 & 0 & 0\\1 & 0 & 1\end{bmatrix}$ and $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 2$, so $P = \frac{1}{2}\begin{bmatrix}1 & 0 & 1\\0 & 0 & 0\\1 & 0 & 1\end{bmatrix}$. Then
$P\mathbf{b} = \frac{1}{2}(2 + 0 + 2,\ 0,\ 2 + 0 + 2) = (2, 0, 2) = \mathbf{p}$. ✓

2 Projection onto a subspace (4.2)

Now the “ground” is a plane through the origin, spanned by two vectors $\mathbf{a}_1$ and $\mathbf{a}_2$. The closest point to $\mathbf{b}$ is a combination $\mathbf{p} = \hat{x}_1\mathbf{a}_1 + \hat{x}_2\mathbf{a}_2$, so there are two numbers to find instead of one. The idea is unchanged: the error must be perpendicular to the plane, which means perpendicular to *both* $\mathbf{a}_1$ and $\mathbf{a}_2$. Two conditions, two unknowns.

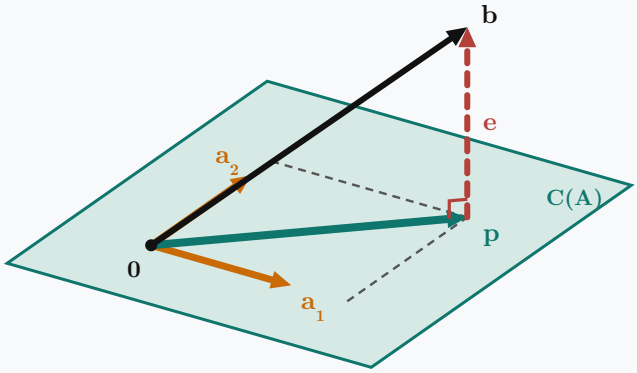


Figure 2.1 — projection onto a plane. The plane is the column space of $A = \begin{bmatrix}\mathbf{a}_1 & \mathbf{a}_2\end{bmatrix}$. The projection $\mathbf{p} = \hat{x}_1\mathbf{a}_1 + \hat{x}_2\mathbf{a}_2$ lies in the plane (dashed parallelogram); the error $\mathbf{e} = \mathbf{b} - \mathbf{p}$ stands on the plane at a right angle, so it is orthogonal to $\mathbf{a}_1$ and to $\mathbf{a}_2$.

Setting up. Put $\mathbf{a}_1, \mathbf{a}_2$ as the columns of a matrix A ($m \times 2$). The plane is the column space $\mathbf{C}(A)$, and a point of it is $A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2$. We look for $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2)$ such that $\mathbf{p} = A\hat{\mathbf{x}}$ is closest to $\mathbf{b}$.

Derivation of the normal equations.

$$
\begin{aligned}
\mathbf{a}_1^{\mathsf{T}}\mathbf{e} = 0 \quad \text{and} \quad \mathbf{a}_2^{\mathsf{T}}\mathbf{e} = 0 \qquad & (\mathbf{e} \text{ is perpendicular to both columns}) \\
A^{\mathsf{T}}\mathbf{e} = \mathbf{0} \qquad & (\text{both equations at once; see below}) \\
A^{\mathsf{T}}(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0} \qquad & (\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}) \\
A^{\mathsf{T}}\mathbf{b} - A^{\mathsf{T}}A\,\hat{\mathbf{x}} = \mathbf{0} \qquad & (\text{multiply out}) \\
A^{\mathsf{T}}A\,\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b} \qquad & (\text{the normal equations})
\end{aligned}
$$

The second line is the first line written as one matrix equation: the rows of A^{T} are $\mathbf{a}_1^{\mathsf{T}}$ and $\mathbf{a}_2^{\mathsf{T}}$, so $A^{\mathsf{T}}\mathbf{e} = \mathbf{0}$ says exactly that $\mathbf{e}$ is perpendicular to both columns.

Projection onto the column space of A.

$$
\begin{aligned}
A^{\mathsf{T}}A\,\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b} \qquad & (\text{solve this small system for } \hat{\mathbf{x}}) \\
\mathbf{p} = A\hat{\mathbf{x}} = \hat{x}_1\mathbf{a}_1 + \hat{x}_2\mathbf{a}_2 \qquad & (\text{the projection}) \\
\mathbf{e} = \mathbf{b} - \mathbf{p} \qquad & (\text{the error, with } A^{\mathsf{T}}\mathbf{e} = \mathbf{0})
\end{aligned}
$$

Compare with a line. There $A^{\mathsf{T}}A$ was the single number $\mathbf{a}^{\mathsf{T}}\mathbf{a}$ and $A^{\mathsf{T}}\mathbf{b}$ the single number $\mathbf{a}^{\mathsf{T}}\mathbf{b}$. Now $A^{\mathsf{T}}A$ is a 2×2 matrix holding the dot products of the columns with each other, and $A^{\mathsf{T}}\mathbf{b}$ is a 2×1

vector holding the dot products of the columns with $\mathbf{b}$.

Method 2.1.

1. Compute $A^{\mathsf{T}}A$ (2×2) and $A^{\mathsf{T}}\mathbf{b}$ (2×1).
2. Solve $A^{\mathsf{T}}A\hat{\mathbf{x}} = A^{\mathsf{T}}\mathbf{b}$ by Gaussian elimination on the augmented matrix $\left[\, A^{\mathsf{T}}A \mid A^{\mathsf{T}}\mathbf{b} \,\right]$, then back substitution.
3. Compute $\mathbf{p} = A\hat{\mathbf{x}} = \hat{x}_1\mathbf{a}_1 + \hat{x}_2\mathbf{a}_2$.
4. Compute $\mathbf{e} = \mathbf{b} - \mathbf{p}$ and **check** $\mathbf{a}_1^{\mathsf{T}}\mathbf{e} = 0$ and $\mathbf{a}_2^{\mathsf{T}}\mathbf{e} = 0$.

Example 2.1 (book, Example 3). **Project $\mathbf{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ onto the column space of $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$.**

1. $A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$ and $A^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 6+0+0 \\ 0+0+0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$.
2. Elimination and back substitution:

$$\left[\begin{array}{cc|c} 3 & 3 & 6 \\ 3 & 5 & 0 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 3 & 3 & 6 \\ 0 & 2 & -6 \end{array}\right] \qquad \begin{aligned} 2\hat{x}_2 &= -6 && \Rightarrow\ \hat{x}_2 = -3 \\ 3\hat{x}_1 + 3(-3) &= 6 && \Rightarrow\ \hat{x}_1 = 5 \end{aligned}$$

3. $\mathbf{p} = 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$.
4. $\mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$. Checks: $(1,1,1)\cdot\mathbf{e} = 1 - 2 + 1 = 0$ and $(0,1,2)\cdot\mathbf{e} = 0 - 2 + 2 = 0$. ✓

Example 2.2 (book, Problem 12a). **Project $\mathbf{b} = (2, 3, 4)$ onto the column space of $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.**

1. $A^{\mathsf{T}}A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $A^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 2 \\ 2+3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}$.
2. $\left[\begin{array}{cc|c} 1 & 1 & 2 \\ 1 & 2 & 5 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 1 & 1 & 2 \\ 0 & 1 & 3 \end{array}\right]$, so $\hat{x}_2 = 3$ and $\hat{x}_1 = 2 - 3 = -1$.
3. $\mathbf{p} = -1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 0 \end{bmatrix}$.
4. $\mathbf{e} = (2, 3, 4) - (2, 3, 0) = (0, 0, 4)$. Checks: $(1, 0, 0) \cdot \mathbf{e} = 0$ and $(1, 1, 0) \cdot \mathbf{e} = 0$. ✓

The column space is the xy-plane, so the projection simply drops the third coordinate, as the shadow picture predicts.

Example 2.3 (book, Problem 12b). **Project $\mathbf{b} = (4, 4, 6)$ onto the column space of $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}$.**

1. $A^{\mathsf{T}}A = \begin{bmatrix} 2 & 2 \\ 2 & 3 \end{bmatrix}$ and $A^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 4+4 \\ 4+4+6 \end{bmatrix} = \begin{bmatrix} 8 \\ 14 \end{bmatrix}$.
2. $\left[\begin{array}{cc|c} 2 & 2 & 8 \\ 2 & 3 & 14 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 2 & 2 & 8 \\ 0 & 1 & 6 \end{array}\right]$, so $\hat{x}_2 = 6$ and $2\hat{x}_1 + 2(6) = 8$, $\hat{x}_1 = -2$.
3. $\mathbf{p} = -2 \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + 6 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 6 \end{bmatrix} = \mathbf{b}$.
4. $\mathbf{e} = \mathbf{0}$.

This $\mathbf{b}$ was already in the column space, so the projection is $\mathbf{b}$ itself and $A\mathbf{x} = \mathbf{b}$ is solvable exactly, with $\mathbf{x} = (-2, 6)$.

Remark (solving the small system). The multiplier in the elimination step is (entry below the pivot) $\div$ (pivot). In Example 2.1 it is $3 \div 3 = 1$, so we subtract row 1 once. If row 2 began with 8 and the pivot were 4, we would subtract 2 times row 1. After the step, read $\hat{x}_2$ from row 2 and substitute it into row 1 to get $\hat{x}_1$. Finally put both numbers back into the original two equations as a check.

The projection matrix

Projection matrix onto $\mathbf{C}(A)$. Solving the normal equations in symbols,

$$\hat{\mathbf{x}} = (A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\mathbf{b}$$
$$\mathbf{p} = A\hat{\mathbf{x}} = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}}\mathbf{b}$$
$$P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}} \qquad \text{with } P^{\mathsf{T}} = P \text{ and } P^2 = P.$$

For the 2×2 inverse we use the formula from Chapter 2:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Swap the diagonal entries, change the signs of the other two, and divide by the determinant $ad - bc$.

Example 2.4. **Find the projection matrix onto the column space of $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ (Example 2.1) and check it on $\mathbf{b} = (6, 0, 0)$.**

$$(A^{\mathsf{T}}A)^{-1} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}^{-1} = \frac{1}{15 - 9}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$$

$$A(A^{\mathsf{T}}A)^{-1} = \frac{1}{6}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ 2 & 0 \\ -1 & 3 \end{bmatrix}$$

$$P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ 2 & 0 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

$$P\mathbf{b} = \frac{1}{6}\begin{bmatrix} 30 \\ 12 \\ -6 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \mathbf{p} \quad \checkmark$$

P is symmetric, as it must be.

Theorem 2.2. $A^{\mathsf{T}}A$ is invertible exactly when the columns of A are independent.

Proof. $A^{\mathsf{T}}A$ has the same nullspace as A. If $A\mathbf{x} = \mathbf{0}$ then certainly $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$. Conversely, if $A^{\mathsf{T}}A\mathbf{x} = \mathbf{0}$, multiply by $\mathbf{x}^{\mathsf{T}}$: $\mathbf{x}^{\mathsf{T}}A^{\mathsf{T}}A\mathbf{x} = \|A\mathbf{x}\|^2 = 0$, so $A\mathbf{x} = \mathbf{0}$. Independent columns mean the nullspace of A is $\{\mathbf{0}\}$; then the nullspace of the square matrix $A^{\mathsf{T}}A$ is $\{\mathbf{0}\}$, so $A^{\mathsf{T}}A$ is invertible. ∎

Two special cases.

- If $\mathbf{b}$ is in the column space: $\mathbf{p} = \mathbf{b}$ and $\mathbf{e} = \mathbf{0}$ (Example 2.3).
- If $\mathbf{b}$ is perpendicular to the column space: $A^{\mathsf{T}}\mathbf{b} = \mathbf{0}$, so $\hat{\mathbf{x}} = \mathbf{0}$ and $\mathbf{p} = \mathbf{0}$.

Exercise 2. Project $\mathbf{b} = (1, 1, 7)$ onto the column space of $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$. Give $\hat{\mathbf{x}}$, $\mathbf{p}$, $\mathbf{e}$ and the two checks.

ANSWER

1. $A^{\mathsf{T}}A = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$ (as in Example 2.1) and $A^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 1+1+7 \\ 0+1+14 \end{bmatrix} = \begin{bmatrix} 9 \\ 15 \end{bmatrix}$.
2. $\left[\begin{array}{cc|c} 3 & 3 & 9 \\ 3 & 5 & 15 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 3 & 3 & 9 \\ 0 & 2 & 6 \end{array}\right]$, so $\hat{x}_2 = 3$ and $3\hat{x}_1 + 3(3) = 9$, $\hat{x}_1 = 0$. $\hat{\mathbf{x}} = (0, 3)$.
3. $\mathbf{p} = 0\,(1,1,1) + 3\,(0,1,2) = (0, 3, 6)$.
4. $\mathbf{e} = (1,1,7) - (0,3,6) = (1, -2, 1)$. Checks: $(1,1,1) \cdot \mathbf{e} = 1 - 2 + 1 = 0$ and $(0,1,2) \cdot \mathbf{e} = 0 - 2 + 2 = 0$. ✓

3 Question types and model answers

Almost every question on this part is one of the three types below. Each model answer is written out the way I expect to see it on paper: the formula first, then the numbers, the arithmetic, the check, and a one-line conclusion.

Type 1 — Projection onto a line, and the matrix P

Question. **Project $\mathbf{b} = (3, 4, 4)$ onto the line through $\mathbf{a} = (2, 2, 1)$. Find the projection matrix and verify $P^2 = P$.**

$$\mathbf{a}^{\mathsf{T}}\mathbf{b} = 2 \cdot 3 + 2 \cdot 4 + 1 \cdot 4 = 18, \qquad \mathbf{a}^{\mathsf{T}}\mathbf{a} = 4 + 4 + 1 = 9$$

$$\hat{x} = \frac{\mathbf{a}^{\mathsf{T}}\mathbf{b}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} = \frac{18}{9} = 2$$

$$\mathbf{p} = \hat{x}\,\mathbf{a} = 2\begin{bmatrix}2\\2\\1\end{bmatrix} = \begin{bmatrix}4\\4\\2\end{bmatrix}$$

$$\mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix}3\\4\\4\end{bmatrix} - \begin{bmatrix}4\\4\\2\end{bmatrix} = \begin{bmatrix}-1\\0\\2\end{bmatrix}, \qquad \text{check: } \mathbf{e}^{\mathsf{T}}\mathbf{a} = -2 + 0 + 2 = 0 \;\checkmark$$

$$P = \frac{\mathbf{a}\,\mathbf{a}^{\mathsf{T}}}{\mathbf{a}^{\mathsf{T}}\mathbf{a}} = \frac{1}{9}\begin{bmatrix}4 & 4 & 2\\4 & 4 & 2\\2 & 2 & 1\end{bmatrix}$$

$$P^2 = \frac{1}{81}\begin{bmatrix}36 & 36 & 18\\36 & 36 & 18\\18 & 18 & 9\end{bmatrix} = \frac{1}{9}\begin{bmatrix}4 & 4 & 2\\4 & 4 & 2\\2 & 2 & 1\end{bmatrix} = P \;\checkmark$$

Type 2 — Projection onto a column space (normal equations)

Question. **Project $\mathbf{b} = (6, 0, 0)$ onto the column space of $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$.**

$$A^{\mathsf{T}}A = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}, \qquad A^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} 3 & 3 & 6 \\ 3 & 5 & 0 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 3 & 3 & 6 \\ 0 & 2 & -6 \end{array}\right]$$

$$2\hat{x}_2 = -6 \;\Rightarrow\; \hat{x}_2 = -3, \qquad 3\hat{x}_1 + 3(-3) = 6 \;\Rightarrow\; \hat{x}_1 = 5$$

$$\mathbf{p} = A\hat{\mathbf{x}} = 5\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 3\begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}$$

$$\mathbf{e} = \mathbf{b} - \mathbf{p} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

checks: $(1, 1, 1) \cdot \mathbf{e} = 1 - 2 + 1 = 0, \qquad (0, 1, 2) \cdot \mathbf{e} = 0 - 2 + 2 = 0$ ✓

Type 3 — The projection matrix onto a column space

Question. **Find the projection matrix onto the column space of $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$, and check it on $\mathbf{b} = (6, 0, 0)$.**

$$A^{\mathsf{T}}A = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}, \qquad (A^{\mathsf{T}}A)^{-1} = \frac{1}{15 - 9}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix}$$

$$A(A^{\mathsf{T}}A)^{-1} = \frac{1}{6}\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ 2 & 0 \\ -1 & 3 \end{bmatrix}$$

$$P = A(A^{\mathsf{T}}A)^{-1}A^{\mathsf{T}} = \frac{1}{6}\begin{bmatrix} 5 & -3 \\ 2 & 0 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

$$P\mathbf{b} = \frac{1}{6}\begin{bmatrix} 30 \\ 12 \\ -6 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix} = \mathbf{p}\ \checkmark \qquad (P \text{ is symmetric, as it must be})$$

4 Practice test

Six questions, about fifty minutes of work. Write full answers on paper first and open each answer only afterwards. Question 6 is theory.

Q1. Project $\mathbf{b} = (1, 2, 3)$ onto the line through $\mathbf{a} = (1, 1, 1)$. Give $\hat{x}$, $\mathbf{p}$, $\mathbf{e}$ and the check.

ANSWER

$\mathbf{a}^{\mathsf{T}}\mathbf{b} = 1 + 2 + 3 = 6$, $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + 1 + 1 = 3$, so $\hat{x} = 6/3 = 2$.
$\mathbf{p} = \hat{x}\,\mathbf{a} = 2(1, 1, 1) = (2, 2, 2)$.
$\mathbf{e} = \mathbf{b} - \mathbf{p} = (1, 2, 3) - (2, 2, 2) = (-1, 0, 1)$.
Check: $\mathbf{e}^{\mathsf{T}}\mathbf{a} = -1 + 0 + 1 = 0$. ✓ Also $\|\mathbf{p}\|^2 + \|\mathbf{e}\|^2 = 12 + 2 = 14 = \|\mathbf{b}\|^2$.

Q2. Project $\mathbf{b} = (1, 0, 0)$ onto the line through $\mathbf{a} = (1, 1, 1)$. Give $\hat{x}$, $\mathbf{p}$, $\mathbf{e}$ and the check.

ANSWER

$\mathbf{a}^{\mathsf{T}}\mathbf{b} = 1 + 0 + 0 = 1$, $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 3$, so $\hat{x} = \dfrac{1}{3}$.

$$\mathbf{p} = \frac{1}{3}(1, 1, 1) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right).$$

$$\mathbf{e} = \mathbf{b} - \mathbf{p} = \left(\frac{2}{3}, -\frac{1}{3}, -\frac{1}{3}\right).$$

Check: $\mathbf{e}^{\mathsf{T}}\mathbf{a} = \dfrac{2}{3} - \dfrac{1}{3} - \dfrac{1}{3} = 0$. ✓ Fractions are normal here; keep them as fractions.

Q3. Write the projection matrix onto the line through $\mathbf{a} = (1, 2)$ and verify $P^2 = P$.

ANSWER

$\mathbf{a}\,\mathbf{a}^{\mathsf{T}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ and $\mathbf{a}^{\mathsf{T}}\mathbf{a} = 1 + 4 = 5$, so $P = \dfrac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$.

$$P^2 = \frac{1}{25}\begin{bmatrix} 1\cdot 1 + 2\cdot 2 & 1\cdot 2 + 2\cdot 4 \\ 2\cdot 1 + 4\cdot 2 & 2\cdot 2 + 4\cdot 4 \end{bmatrix} = \frac{1}{25}\begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} = P.$$ ✓ P is symmetric and has rank 1.

Q4. Project $\mathbf{b} = (3, 2, 7)$ onto the column space of $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$. Give $\hat{\mathbf{x}}$, $\mathbf{p}$, $\mathbf{e}$, and the checks.

ANSWER

$A^{\mathsf{T}}A = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix}$, $A^{\mathsf{T}}\mathbf{b} = \begin{bmatrix} 3+2+7 \\ 0+2+14 \end{bmatrix} = \begin{bmatrix} 12 \\ 16 \end{bmatrix}$.

$\left[\begin{array}{cc|c} 3 & 3 & 12 \\ 3 & 5 & 16 \end{array}\right] \xrightarrow{R_2 - R_1} \left[\begin{array}{cc|c} 3 & 3 & 12 \\ 0 & 2 & 4 \end{array}\right]$, so $\hat{x}_2 = 2$ and $3\hat{x}_1 + 3(2) = 12$, $\hat{x}_1 = 2$. $\hat{\mathbf{x}} = (2, 2)$.

$\mathbf{p} = 2(1,1,1) + 2(0,1,2) = (2, 4, 6)$.

$\mathbf{e} = (3,2,7) - (2,4,6) = (1, -2, 1)$.

Checks: $(1,1,1) \cdot \mathbf{e} = 1 - 2 + 1 = 0$ and $(0,1,2) \cdot \mathbf{e} = 0 - 2 + 2 = 0$. ✓

Q5. Find the projection matrix onto the column space of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$. Check that it is symmetric and that $P\mathbf{a}_1 = \mathbf{a}_1$ for the first column.

ANSWER

$A^{\mathsf{T}}A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, $(A^{\mathsf{T}}A)^{-1} = \dfrac{1}{4-1}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \dfrac{1}{3}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$.

$A(A^{\mathsf{T}}A)^{-1} = \dfrac{1}{3}\begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \dfrac{1}{3}\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}$, and

$P = \dfrac{1}{3}\begin{bmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \dfrac{1}{3}\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{bmatrix}$.

P is symmetric. $P\mathbf{a}_1 = \dfrac{1}{3}(2+1,\ 1+2,\ 1-1) = (1, 1, 0) = \mathbf{a}_1$ ✓ (a vector already in the column space is its own projection).

Q6 (theory). (a) Starting from "the error is perpendicular to every column of A", derive the normal equations. (b) Explain in one sentence why $P^2 = P$ for a projection matrix. (c) When is $A^\mathsf{T}A$ invertible?

ANSWER

(a) The error is $\mathbf{e} = \mathbf{b} - A\hat{\mathbf{x}}$. Perpendicular to each column $\mathbf{a}_i$ means $\mathbf{a}_i^\mathsf{T}(\mathbf{b} - A\hat{\mathbf{x}}) = 0$ for every i. Stacking these equations (the rows of A^T are the $\mathbf{a}_i^\mathsf{T}$) gives $A^\mathsf{T}(\mathbf{b} - A\hat{\mathbf{x}}) = \mathbf{0}$, that is $A^\mathsf{T}\mathbf{b} - A^\mathsf{T}A\hat{\mathbf{x}} = \mathbf{0}$, so $A^\mathsf{T}A\hat{\mathbf{x}} = A^\mathsf{T}\mathbf{b}$.

(b) $P\mathbf{b}$ already lies in the subspace, and projecting a vector that is already in the subspace leaves it unchanged, so $P(P\mathbf{b}) = P\mathbf{b}$ for every $\mathbf{b}$.

(c) Exactly when the columns of A are independent: $A^\mathsf{T}A$ has the same nullspace as A, because $A^\mathsf{T}A\mathbf{x} = \mathbf{0}$ gives $\mathbf{x}^\mathsf{T}A^\mathsf{T}A\mathbf{x} = \|A\mathbf{x}\|^2 = 0$, so $A\mathbf{x} = \mathbf{0}$.

5 Formula sheet

Topic	Formulas
Projection onto a line	$\hat{x} = \dfrac{\mathbf{a}^\mathsf{T}\mathbf{b}}{\mathbf{a}^\mathsf{T}\mathbf{a}}$, $\mathbf{p} = \hat{x}\,\mathbf{a}$, $\mathbf{e} = \mathbf{b} - \mathbf{p}$ check $\mathbf{e}^\mathsf{T}\mathbf{a} = 0$; also $\|\mathbf{p}\|^2 + \|\mathbf{e}\|^2 = \|\mathbf{b}\|^2$
Projection matrix (line)	$P = \dfrac{\mathbf{a}\,\mathbf{a}^\mathsf{T}}{\mathbf{a}^\mathsf{T}\mathbf{a}}$ (column times row, divided by a number) $P^\mathsf{T} = P$, $P^2 = P$, rank 1
Projection onto $\mathbf{C}(A)$	$A^\mathsf{T}A\hat{\mathbf{x}} = A^\mathsf{T}\mathbf{b}$ (normal equations), $\mathbf{p} = A\hat{\mathbf{x}}$, $\mathbf{e} = \mathbf{b} - \mathbf{p}$ check $A^\mathsf{T}\mathbf{e} = \mathbf{0}$: $\mathbf{e}$ is perpendicular to every column
Projection matrix (subspace)	$P = A(A^\mathsf{T}A)^{-1}A^\mathsf{T}$, $P^\mathsf{T} = P$, $P^2 = P$ $A^\mathsf{T}A$ is invertible exactly when the columns of A are independent
2×2 inverse	$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \dfrac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$
Special cases	$\mathbf{b}$ in the subspace: $\mathbf{p} = \mathbf{b}$, $\mathbf{e} = \mathbf{0}$ $\mathbf{b}$ perpendicular to the subspace: $\mathbf{p} = \mathbf{0}$, $\mathbf{e} = \mathbf{b}$