Linear Algebra

Projection

Source: Introduction to Linear Algebra by Gilbert Strang

Why Projection

- We can't always solve $Ax=b$.
- When it is solvable? If b belongs to $C(A)$.
- If it is not solvable (b is not there in $C(A)$), can we get the best solution $\hat{x}$?

- Can we get the closest vector to b which is there in $C(A)$?

1 What are the projections of $\boldsymbol{b} = (2, 3, 4)$ onto the z axis and the xy plane?

2 What matrices produce those projections onto a line and a plane?

When $\boldsymbol{b}$ is projected onto a line, ***its projection p is the part of b along that line***. If $\boldsymbol{b}$ is projected onto a plane, $\boldsymbol{p}$ is the part in that plane. ***The projection p is Pb.***

There is a projection matrix P that multiplies $\boldsymbol{b}$ to give $\boldsymbol{p}$. This section finds $\boldsymbol{p}$ and P.

One projection gives $\boldsymbol{p}_1 = (0, 0, 4)$ and the other gives $\boldsymbol{p}_2 = (2, 3, 0)$. Those are the parts of $\boldsymbol{b}$ along the z axis and in the xy plane.

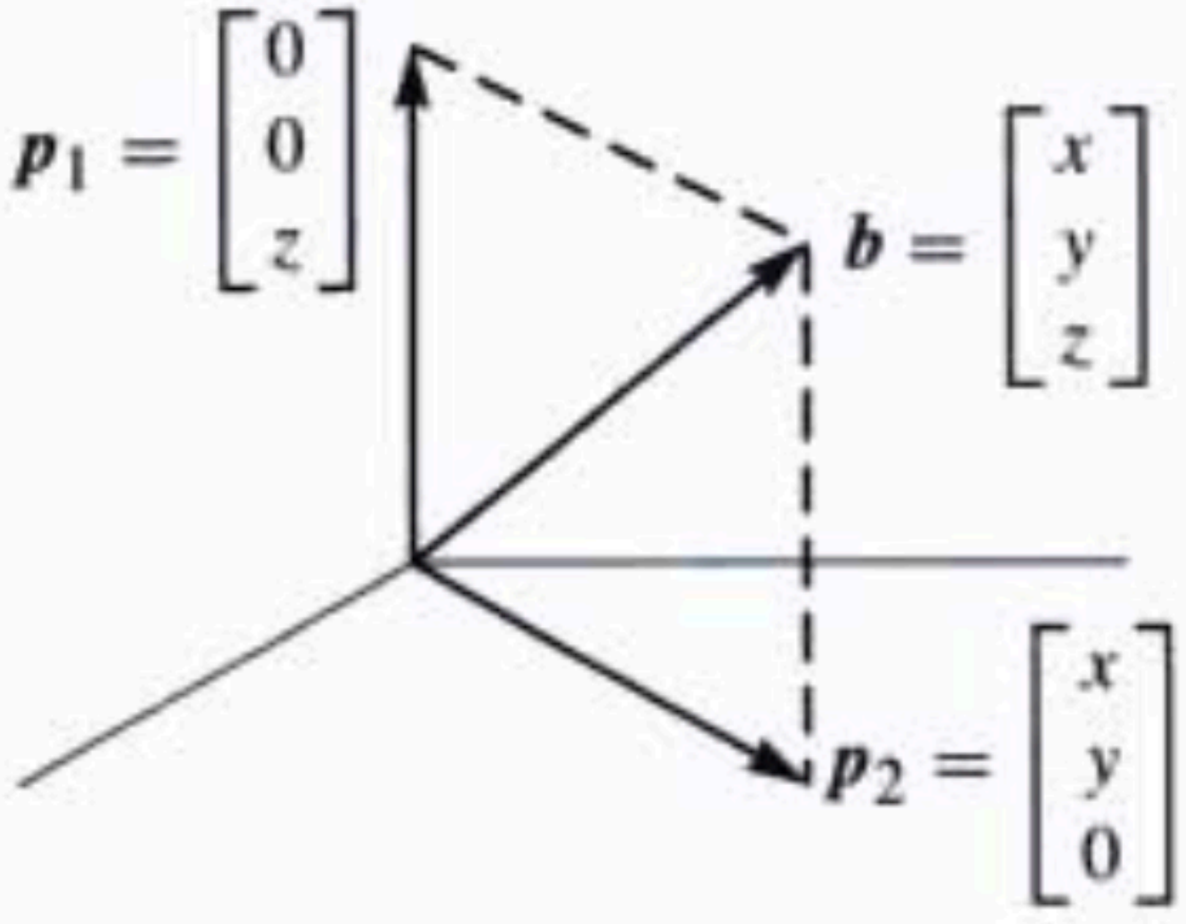

Figure 4.3 The projections of $\boldsymbol{b}$ onto the z axis and the xy plane.

The projection matrices P_1 and P_2 are 3 by 3. They multiply $\boldsymbol{b}$ with 3 components to produce $\boldsymbol{p}$ with 3 components. Projection onto a line comes from a rank one matrix. Projection onto a plane comes from a rank two matrix:

$$\text{Onto the } z \text{ axis:} \quad P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \text{Onto the } xy \text{ plane:} \quad P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

P_1 picks out the z component of every vector. P_2 picks out the x and y components. To find $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, multiply $\boldsymbol{b}$ by P_1 and P_2 (small $\boldsymbol{p}$ for the vector, capital P for the matrix that produces it):

$$\boldsymbol{p}_1 = P_1\boldsymbol{b} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ z \end{bmatrix} \qquad \boldsymbol{p}_2 = P_2\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}.$$

The xy plane and the z axis are ***orthogonal subspaces***, like the floor of a room and the line between two walls. More than that, the line and plane are orthogonal ***complements***. Their dimensions add to $1 + 2 = 3$—every vector $\boldsymbol{b}$ in the whole space is the sum of its parts in the two subspaces. The projections $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ are exactly those parts:

$$\text{The vectors give } \boldsymbol{p}_1 + \boldsymbol{p}_2 = \boldsymbol{b}. \qquad \text{The matrices give } P_1 + P_2 = I. \tag{1}$$

The object is to find the part $\boldsymbol{p}$ in each subspace, and the projection matrix P that produces that part $\boldsymbol{p} = P\boldsymbol{b}$. Every subspace of $\mathbf{R}^m$ has its own m by m projection matrix. To compute P, we absolutely need a good description of the subspace that it projects onto.

The best description of a subspace is a basis. We put the basis vectors into the columns of A. ***Now we are projecting onto the column space of A!*** Certainly the z axis is the column space of the 3 by 1 matrix A_1. The xy plane is the column space of A_2. That plane is also the column space of A_3 (a subspace has many bases):

$$A_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{and} \quad A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad A_3 = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 0 & 0 \end{bmatrix}.$$

Our problem is ***to project onto the column space of any m by n matrix.*** Start with a line (dimension $n = 1$). The matrix A has only one column. Call it $\boldsymbol{a}$.

We are given a line through the origin, in the direction of $\boldsymbol{a} = (a_1, \ldots, a_m)$. Along that line, we want the point $\boldsymbol{p}$ closest to $\boldsymbol{b} = (b_1, \ldots, b_m)$. The key to projection is orthogonality: ***The line from b to p is perpendicular to the vector a***. This is the dotted line marked $\boldsymbol{e}$ in Figure 4.4—which we now compute by algebra.

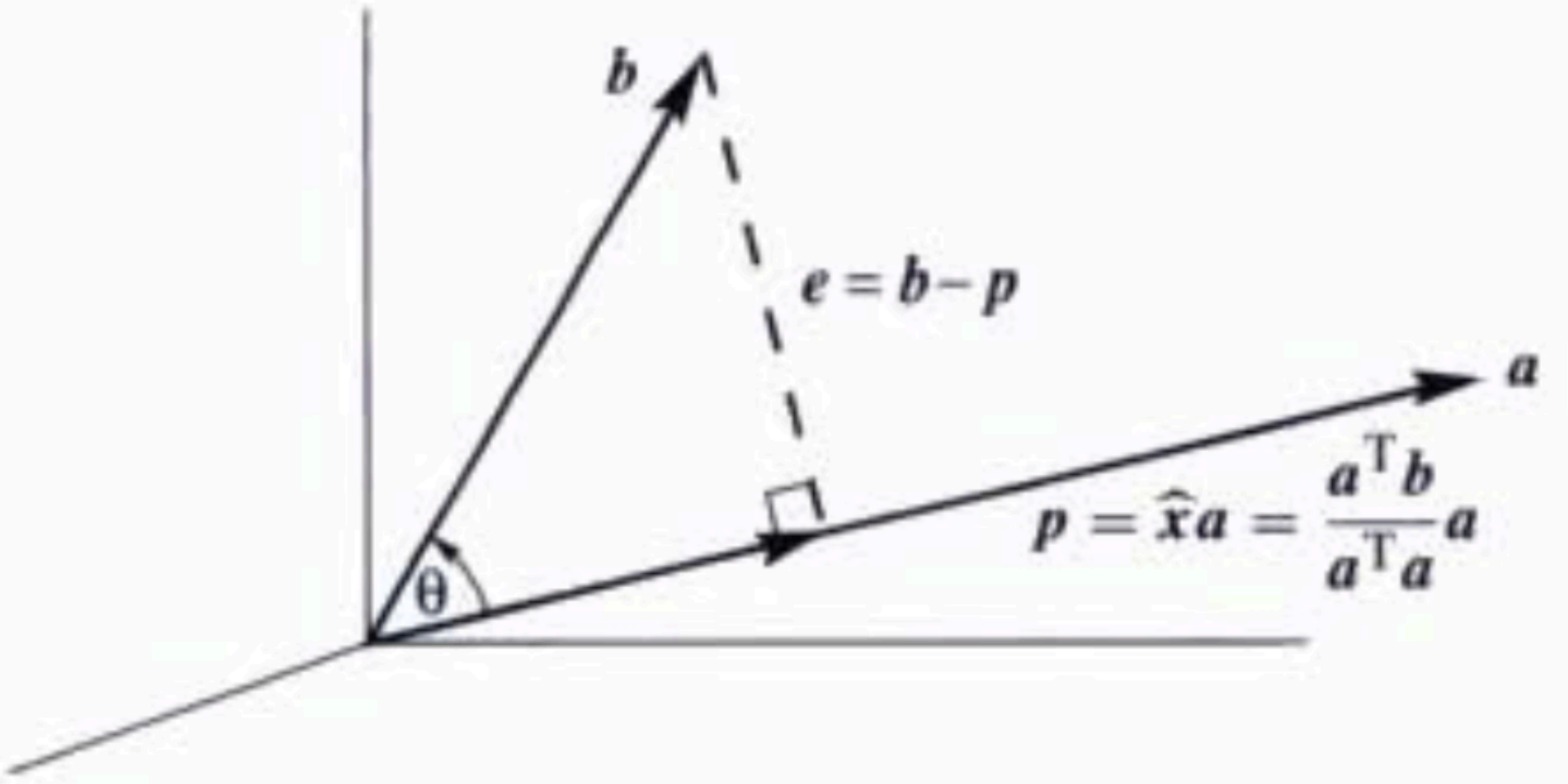

Figure 4.4 The projection $\boldsymbol{p}$, perpendicular to $\boldsymbol{e}$, has length $\|\boldsymbol{b}\| \cos\theta$.

The projection $\boldsymbol{p}$ is some multiple of $\boldsymbol{a}$. Call it $\boldsymbol{p} = \widehat{x}\boldsymbol{a}$ = "x hat" times $\boldsymbol{a}$. Our first step is to compute this unknown number $\widehat{x}$. That will give the vector $\boldsymbol{p}$. Then from the formula for $\boldsymbol{p}$, we read off the projection matrix P. These three steps will lead to all projection matrices: *find* $\widehat{x}$, *then find* ***the vector p***, *then find* ***the matrix*** P.

$$\boldsymbol{a} \cdot (\boldsymbol{b} - \widehat{x}\boldsymbol{a}) = 0 \quad \text{or} \quad \boldsymbol{a} \cdot \boldsymbol{b} - \widehat{x}\boldsymbol{a} \cdot \boldsymbol{a} = 0 \quad \text{or} \quad \boxed{\widehat{x} = \frac{\boldsymbol{a} \cdot \boldsymbol{b}}{\boldsymbol{a} \cdot \boldsymbol{a}} = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}}. \tag{2}$$

4E The projection of $\boldsymbol{b}$ onto the line through $\boldsymbol{a}$ is the vector $\boxed{\boldsymbol{p} = \widehat{x}\boldsymbol{a} = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}\boldsymbol{a}.}$

Special case 1: If $\boldsymbol{b} = \boldsymbol{a}$ then $\widehat{x} = 1$. The projection of $\boldsymbol{a}$ onto $\boldsymbol{a}$ is itself.

Special case 2: If $\boldsymbol{b}$ is perpendicular to $\boldsymbol{a}$ then $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 0$. The projection is $\boldsymbol{p} = \boldsymbol{0}$.

- What if you multiply 2 with b?
- What if you multiply 2 with a?

Example 1 Project $\boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$ onto $\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ to find $\boldsymbol{p} = \widehat{x}\boldsymbol{a}$ in Figure 4.4.

Solution The number $\widehat{x}$ is the ratio of $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = 5$ to $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 9$. So the projection is $\boldsymbol{p} = \frac{5}{9}\boldsymbol{a}$. The error vector between $\boldsymbol{b}$ and $\boldsymbol{p}$ is $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$. Those vectors $\boldsymbol{p}$ and $\boldsymbol{e}$ will add to $\boldsymbol{b}$:

$$\boldsymbol{p} = \frac{5}{9}\boldsymbol{a} = \left(\frac{5}{9}, \frac{10}{9}, \frac{10}{9}\right) \text{ and } \boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \left(\frac{4}{9}, -\frac{1}{9}, -\frac{1}{9}\right).$$

The error $\boldsymbol{e}$ should be perpendicular to $\boldsymbol{a} = (1, 2, 2)$ and it is: $\boldsymbol{e}^{\mathrm{T}}\boldsymbol{a} = \frac{4}{9} - \frac{2}{9} - \frac{2}{9} = 0$.

Look at the right triangle of $\boldsymbol{b}$, $\boldsymbol{p}$, and $\boldsymbol{e}$. The vector $\boldsymbol{b}$ is split into two parts—its component along the line is $\boldsymbol{p}$, its perpendicular part is $\boldsymbol{e}$.

Now comes the ***projection matrix***. In the formula for $\boldsymbol{p}$, what matrix is multiplying $\boldsymbol{b}$? You can see it better if the number $\widehat{x}$ is on the right side of $\boldsymbol{a}$:

$$p = a\widehat{x} = a\frac{a^{\mathrm{T}}b}{a^{\mathrm{T}}a} = Pb \quad \textit{when the matrix is} \quad P = \frac{aa^{\mathrm{T}}}{a^{\mathrm{T}}a}.$$

P is a column times a row! The column is $\boldsymbol{a}$, the row is $\boldsymbol{a}^{\mathrm{T}}$. Then divide by the number $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$. The projection matrix P is m by m, but ***its rank is one***. We are projecting onto a one-dimensional subspace, the line through $\boldsymbol{a}$.

Example 2 Find the projection matrix $P = \dfrac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}$ onto the line through $\boldsymbol{a} = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$.

Solution Multiply column times row and divide by $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a} = 9$:

$$P = \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} = \frac{1}{9}\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}\begin{bmatrix} 1 & 2 & 2 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}.$$

This matrix projects *any* vector $\boldsymbol{b}$ onto $\boldsymbol{a}$. Check $\boldsymbol{p} = P\boldsymbol{b}$ for the particular $\boldsymbol{b} = (1, 1, 1)$ in Example 1:

$$\boldsymbol{p} = P\boldsymbol{b} = \frac{1}{9}\begin{bmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \frac{1}{9}\begin{bmatrix} 5 \\ 10 \\ 10 \end{bmatrix} \quad \text{which is correct.}$$

If the vector $\boldsymbol{a}$ is doubled, the matrix P stays the same. It still projects onto the same line. If the matrix is squared, P^2 equals P. ***Projecting a second time doesn't change anything***, so $P^2 = P$. The diagonal entries of P add up to $\frac{1}{9}(1+4+4) = 1$.

The matrix $I - P$ should be a projection too. It produces the other side $\boldsymbol{e}$ of the triangle—the perpendicular part of $\boldsymbol{b}$. Note that $(I - P)\boldsymbol{b}$ equals $\boldsymbol{b} - \boldsymbol{p}$ which is $\boldsymbol{e}$. ***When P projects onto one subspace, $I - P$ projects onto the perpendicular subspace.*** Here $I - P$ projects onto the plane perpendicular to $\boldsymbol{a}$.

Projection Onto a Subspace

Start with n vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ in $\mathbf{R}^m$. Assume that these $\boldsymbol{a}$'s are linearly independent. ***Problem: Find the combination*** $\widehat{x}_1\boldsymbol{a}_1 + \cdots + \widehat{x}_n\boldsymbol{a}_n$ ***that is closest to a given vector b***. We are projecting each $\boldsymbol{b}$ in $\mathbf{R}^m$ onto the subspace spanned by the $\boldsymbol{a}$'s.

With $n = 1$ (only one vector $\boldsymbol{a}_1$) this is projection onto a line. The line is the column space of A, which has just one column. In general the matrix A has n columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Their combinations in $\mathbf{R}^m$ are the vectors $A\boldsymbol{x}$ in the column space. We are looking for the particular combination $\boldsymbol{p} = A\widehat{\boldsymbol{x}}$ (***the projection***) that is closest to $\boldsymbol{b}$. The hat over $\widehat{\boldsymbol{x}}$ indicates the *best* choice, to give the closest vector in the column space. That choice is $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$ when $n = 1$. For $n > 1$, the best $\widehat{\boldsymbol{x}}$ is to be found.

We solve this problem for an n-dimensional subspace in three steps: *Find the vector* $\widehat{\boldsymbol{x}}$, *find the projection* $\boldsymbol{p} = A\widehat{\boldsymbol{x}}$, *find the matrix* P.

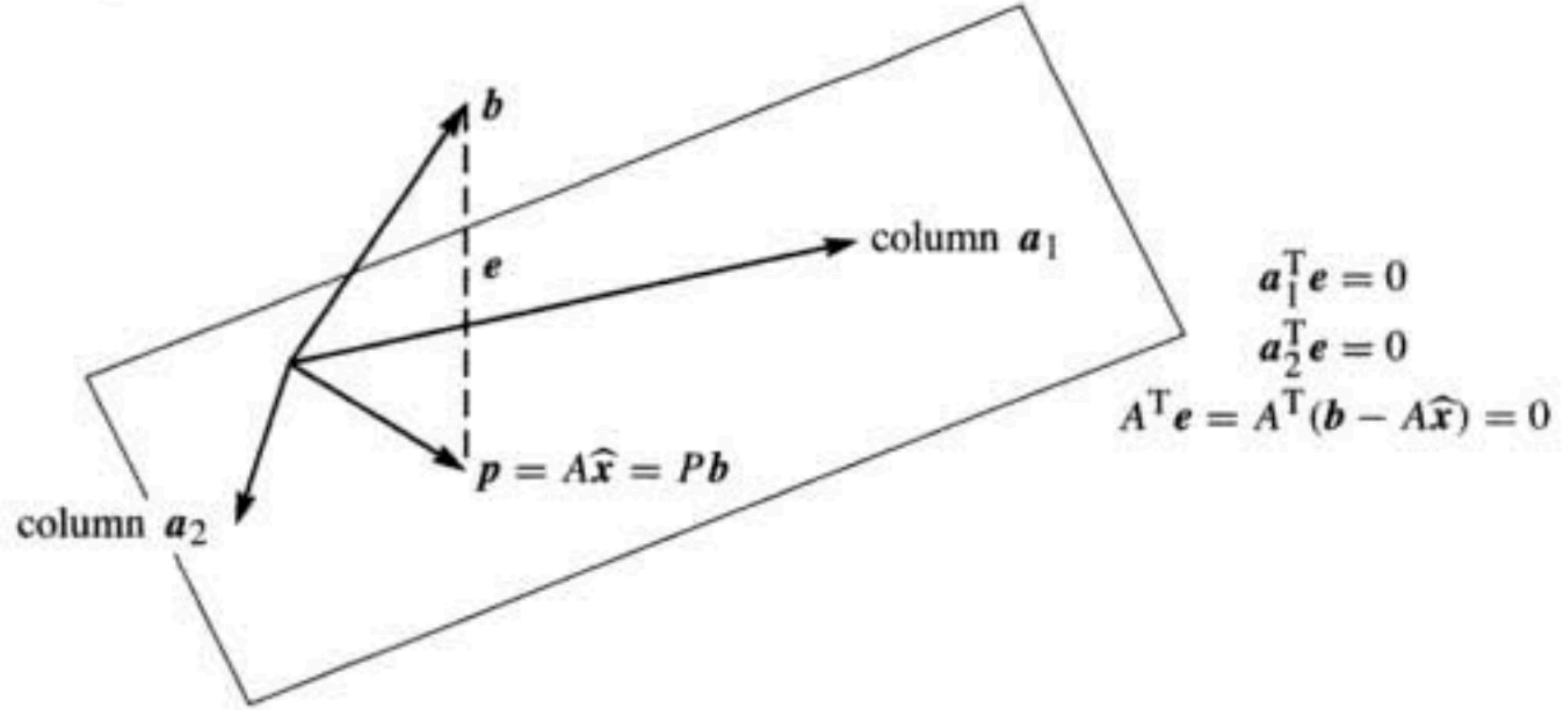

Figure 4.5 The projection $\boldsymbol{p}$ is the nearest point to $\boldsymbol{b}$ in the column space of A. The perpendicular error $\boldsymbol{e}$ must be in the nullspace of A^{T}.

The key is in the geometry! The dotted line in Figure 4.5 goes from $\boldsymbol{b}$ to the nearest point $A\widehat{\boldsymbol{x}}$ in the subspace. ***This error vector $\boldsymbol{b} - A\widehat{\boldsymbol{x}}$ is perpendicular to the subspace.*** The error $\boldsymbol{b} - A\widehat{\boldsymbol{x}}$ makes a right angle with all the vectors $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. That gives the n equations we need to find $\widehat{\boldsymbol{x}}$:

$$\begin{matrix} \boldsymbol{a}_1^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = 0 \\ \vdots \\ \boldsymbol{a}_n^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = 0 \end{matrix} \quad \text{or} \quad \begin{bmatrix} - & \boldsymbol{a}_1^{\mathrm{T}} & - \\ & \vdots & \\ - & \boldsymbol{a}_n^{\mathrm{T}} & - \end{bmatrix} \begin{bmatrix} \\ \boldsymbol{b} - A\widehat{\boldsymbol{x}} \\ \\ \end{bmatrix} = \begin{bmatrix} \\ \boldsymbol{0} \\ \\ \end{bmatrix}. \tag{4}$$

The matrix in those equations is A^{T}. The n equations are exactly $A^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \boldsymbol{0}$.

Rewrite $A^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \boldsymbol{0}$ in its famous form $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$. This is the equation for $\widehat{\boldsymbol{x}}$, and the coefficient matrix is $A^{\mathrm{T}}A$. Now we can find $\widehat{\boldsymbol{x}}$ and $\boldsymbol{p}$ and P:

4F The combination $\widehat{x}_1 \boldsymbol{a}_1 + \cdots + \widehat{x}_n \boldsymbol{a}_n = A\widehat{\boldsymbol{x}}$ that is closest to $\boldsymbol{b}$ comes from

$$A^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \mathbf{0} \quad \text{or} \quad A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}. \tag{5}$$

The symmetric matrix $A^{\mathrm{T}}A$ is n by n. It is invertible if the $\boldsymbol{a}$'s are independent. The solution is $\widehat{\boldsymbol{x}} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}$. The ***projection*** of $\boldsymbol{b}$ onto the subspace is the vector

$$\boldsymbol{p} = A\widehat{\boldsymbol{x}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}. \tag{6}$$

This formula shows the n by n ***projection matrix*** that produces $\boldsymbol{p} = P\boldsymbol{b}$:

$$P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}. \tag{7}$$

Compare with projection onto a line, when the matrix A has only one column $\boldsymbol{a}$:

$$\widehat{\boldsymbol{x}} = \frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} \quad \text{and} \quad \boldsymbol{p} = \boldsymbol{a}\frac{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}} \quad \text{and} \quad P = \frac{\boldsymbol{a}\boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}.$$

Those formulas are identical with (5) and (6) and (7)! The number $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$ becomes the matrix $A^{\mathrm{T}}A$. When it is a number, we divide by it. When it is a matrix, we invert it. The new formulas contain $(A^{\mathrm{T}}A)^{-1}$ instead of $1/\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}$. The linear independence of the columns $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ will guarantee that this inverse matrix exists.

The key step was $A^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \mathbf{0}$. We used geometry ($\boldsymbol{e}$ is perpendicular to all the $\boldsymbol{a}$'s).

1. Our subspace is the column space of A.

2. The error vector $\boldsymbol{b} - A\widehat{\boldsymbol{x}}$ is perpendicular to that column space.

3. Therefore $\boldsymbol{b} - A\widehat{\boldsymbol{x}}$ is in the left nullspace. This means $A^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = \mathbf{0}$.

The left nullspace is important in projections. This nullspace of A^{T} contains the error vector $\boldsymbol{e} = \boldsymbol{b} - A\widehat{\boldsymbol{x}}$. The vector $\boldsymbol{b}$ is being split into the projection $\boldsymbol{p}$ and the error $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$. Figure 4.5 shows the right triangle with sides $\boldsymbol{p}$, $\boldsymbol{e}$, and $\boldsymbol{b}$.

Example 3 If $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix}$ and $\boldsymbol{b} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$ find $\widehat{\boldsymbol{x}}$ and $\boldsymbol{p}$ and P.

Solution Compute the square matrix $A^{\mathrm{T}}A$ and also the vector $A^{\mathrm{T}}\boldsymbol{b}$:

$$A^{\mathrm{T}}A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \quad \text{and} \quad A^{\mathrm{T}}b = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix}.$$

Now solve the normal equation $A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}$ to find $\widehat{\boldsymbol{x}}$:

$$\begin{bmatrix} 3 & 3 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \end{bmatrix} \quad \text{gives} \quad \widehat{\boldsymbol{x}} = \begin{bmatrix} \widehat{x}_1 \\ \widehat{x}_2 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix}. \tag{8}$$

The combination $\boldsymbol{p} = A\widehat{\boldsymbol{x}}$ is the projection of $\boldsymbol{b}$ onto the column space of A:

$$\boldsymbol{p} = 5 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - 3 \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 5 \\ 2 \\ -1 \end{bmatrix}. \quad \text{The error is} \quad \boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}. \tag{9}$$

That solves the problem for one particular $\boldsymbol{b}$. To solve it for every $\boldsymbol{b}$, compute the matrix $P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$. The determinant of $A^{\mathrm{T}}A$ is $15 - 9 = 6$; $(A^{\mathrm{T}}A)^{-1}$ is easy. Then multiply A times $(A^{\mathrm{T}}A)^{-1}$ times A^{T} to reach P:

$$(A^{\mathrm{T}}A)^{-1} = \frac{1}{6} \begin{bmatrix} 5 & -3 \\ -3 & 3 \end{bmatrix} \quad \text{and} \quad P = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}. \tag{10}$$

4G $A^{\mathrm{T}}A$ is invertible if and only if A has linearly independent columns.

Proof $A^{\mathrm{T}}A$ is a square matrix (n by n). For every matrix A, we will now show that $A^{\mathrm{T}}A$ ***has the same nullspace as*** A. When the columns of A are linearly independent, its nullspace contains only the zero vector. Then $A^{\mathrm{T}}A$, with this same nullspace, is invertible.

Let A be any matrix. If $\boldsymbol{x}$ is in its nullspace, then $A\boldsymbol{x} = \mathbf{0}$. Multiplying by A^{T} gives $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$. So $\boldsymbol{x}$ is also in the nullspace of $A^{\mathrm{T}}A$.

Now start with the nullspace of $A^{\mathrm{T}}A$. From $A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0}$ we must prove that $A\boldsymbol{x} = \mathbf{0}$. We can't multiply by $(A^{\mathrm{T}})^{-1}$, which generally doesn't exist. Just multiply by $\boldsymbol{x}^{\mathrm{T}}$:

$$(\boldsymbol{x}^{\mathrm{T}})A^{\mathrm{T}}A\boldsymbol{x} = 0 \quad \text{or} \quad (A\boldsymbol{x})^{\mathrm{T}}(A\boldsymbol{x}) = 0 \quad \text{or} \quad \|A\boldsymbol{x}\|^2 = 0.$$

The vector $A\boldsymbol{x}$ has length zero. Therefore $A\boldsymbol{x} = \mathbf{0}$. Every vector $\boldsymbol{x}$ in one nullspace is in the other nullspace. If A has dependent columns, so does $A^{\mathrm{T}}A$. If A has independent columns, so does $A^{\mathrm{T}}A$. This is the good case:

When A has independent columns, $A^{\mathrm{T}}A$ is ***square, symmetric,*** and ***invertible.***