EECS 16A Designing Information Devices and Systems I
Spring 2021 Discussion 12B

1. Mechanical Projection

In $\mathbb{R}^n$, the vector valued projection of vector $\vec{b}$ onto vector $\vec{a}$ is defined as:

$$\text{proj}_{\vec{a}}\left(\vec{b}\right) = \frac{\left\langle \vec{a}, \vec{b} \right\rangle}{\|\vec{a}\|^2}\vec{a}.$$

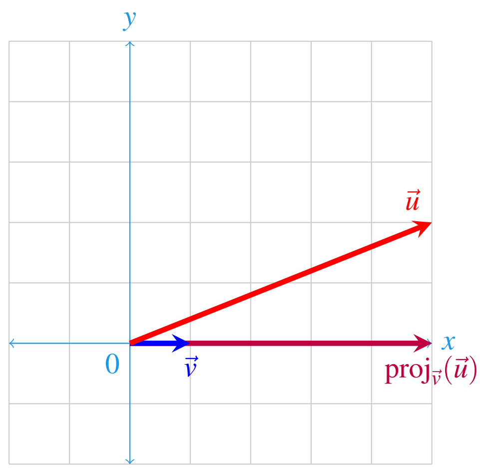
Recall $\|\vec{a}\|^2 = \langle \vec{a}, \vec{a} \rangle$.

(a) Project $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ – that is, onto the x-axis. Graph these two vectors and the projection.

Answer:

$$\vec{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \text{proj}_{\vec{v}}(\vec{u}) &= \frac{\vec{u}^\top \vec{v}}{\|\vec{v}\|^2}\vec{v} \\ &= \frac{5}{1}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix} \end{aligned}$$

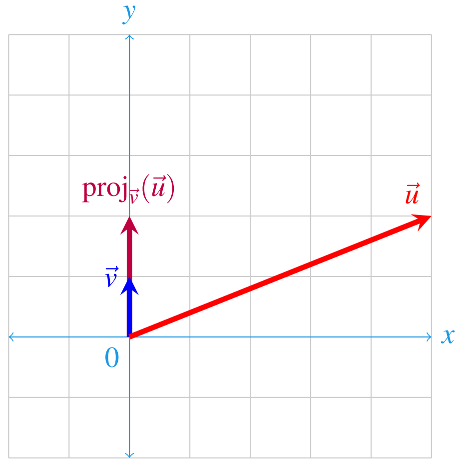


(b) Project $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ onto $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ – that is, onto the y-axis. Graph these two vectors and the projection.

Answer:

$$\vec{u} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \text{proj}_{\vec{v}}(\vec{u}) &= \frac{\vec{u}^\top \vec{v}}{\|\vec{v}\|^2}\vec{v} \\ &= \frac{2}{1}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \end{aligned}$$

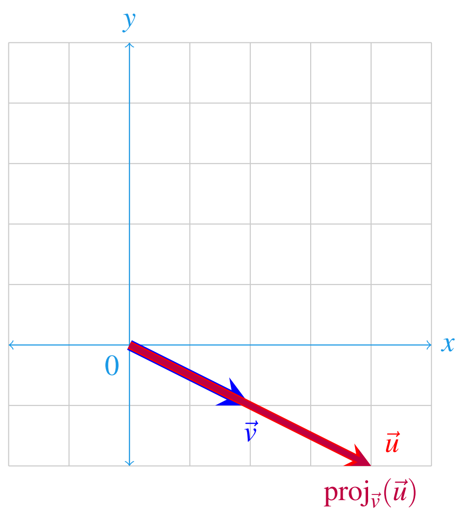


(c) Project $\begin{bmatrix} 4 \\ -2 \end{bmatrix}$ onto $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$. Graph these two vectors and the projection.

Answer:

$$\vec{u} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

$$\begin{aligned} \text{proj}_{\vec{v}}(\vec{u}) &= \frac{\vec{u}^\top \vec{v}}{\|\vec{v}\|^2} \vec{v} \\ &= \frac{10}{5} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix} \end{aligned}$$

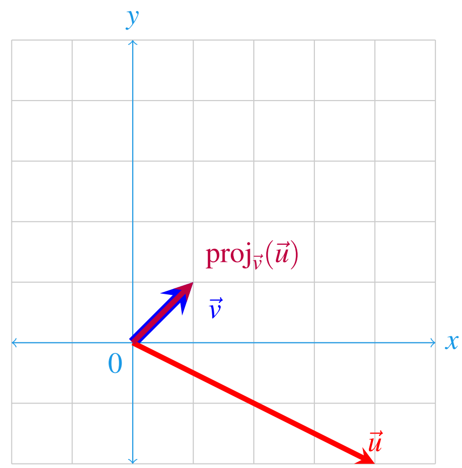


(d) Project $\begin{bmatrix} 4 \\ -2 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Graph these two vectors and the projection.

Answer:

$$\vec{u} = \begin{bmatrix} 4 \\ -2 \end{bmatrix}, \vec{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\begin{aligned} \text{proj}_{\vec{v}}(\vec{u}) &= \frac{\vec{u}^\top \vec{v}}{\|\vec{v}\|^2}\vec{v} \\ &= \frac{2}{2}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{aligned}$$

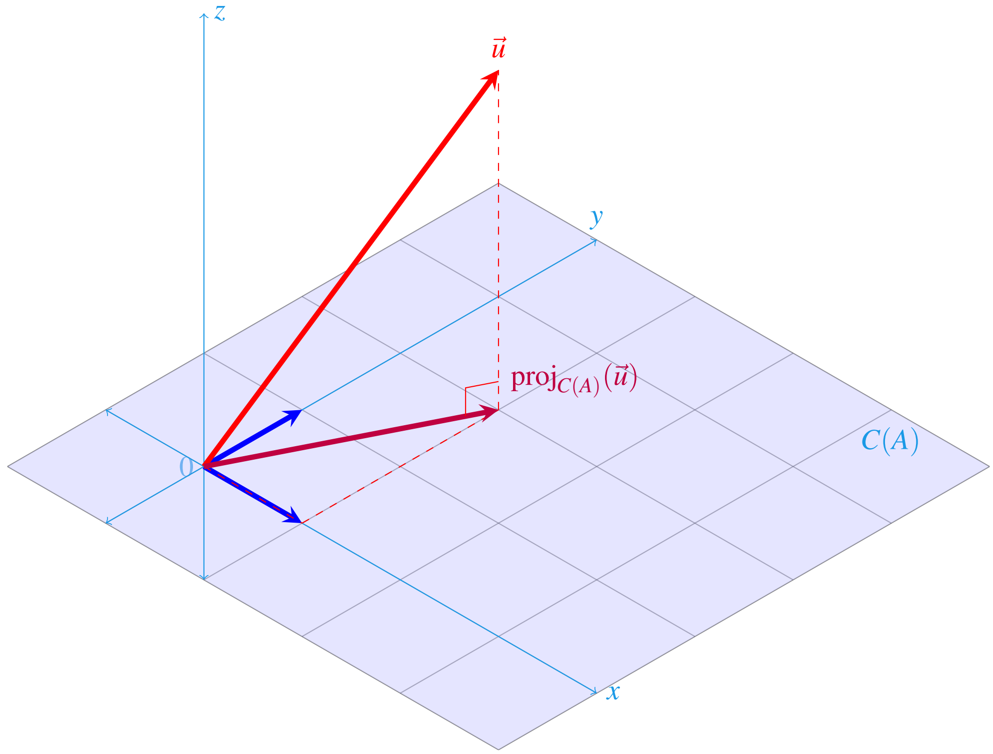


(e) (Practice) Project $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ onto the span of the vectors $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$ – that is, onto the x-y plane in $\mathbb{R}^3$. (Hint: From least squares, the matrix $A(A^\top A)^{-1}A^\top$ projects a vector into $C(A)$.)

Answer: Define $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$. We are being asked to project $\vec{u} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ onto $C(A)$.

$$\begin{aligned} \text{proj}_{C(A)}(\vec{u}) &= A(A^\top A)^{-1}A^\top \vec{u} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \right)^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 2 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} \end{aligned}$$

While you were not asked to plot, a visualization follows.



(f) (Practice) What is the geometric/physical interpretation of projection? Justify using the previous parts.

Answer:

Geometrically, the orthogonal projection yields the closest vector in the projection space, such that the difference between the original vector and the projected vector is orthogonal to the projected vector. That is, the projection is the best approximation for the original vector within the subspace contraint.

2. Least Squares with Orthogonal Columns

Suppose we would like to solve the least squares problem for $\mathbf{A} \in \mathbb{R}^{3\times 2}$ and $\vec{b} \in \mathbb{R}^3$; that is, find an optimal vector $\vec{x} \in \mathbb{R}^2$ which gets $\mathbf{A}\vec{x}$ closest to $\vec{b}$ such that the distance $||\vec{e}|| = ||\vec{b} - \mathbf{A}\vec{x}||$ is minimized. Call this optimal vector $\vec{\hat{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$. Mathematically, we can express this as:

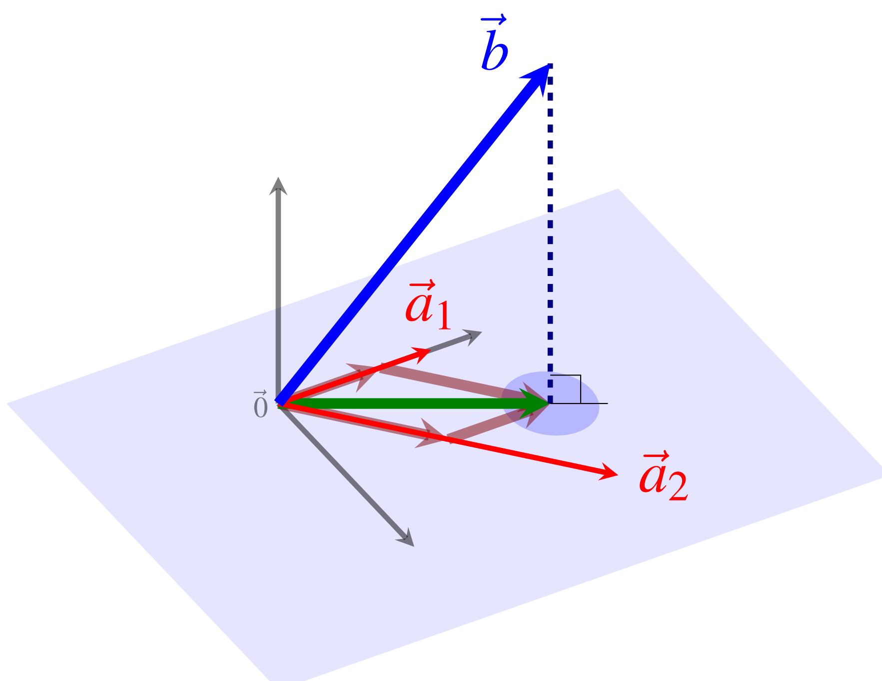
$$||\vec{b} - \mathbf{A}\vec{\hat{x}}||^2 = \min_{\vec{x}\in\mathbb{R}^2} ||\vec{b} - \mathbf{A}\vec{x}||^2 = \min_{\vec{x}\in\mathbb{R}^2} \left\| \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} - \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2 .$$

To identify the solution $\vec{\hat{x}}$, we may recall the least squares formula: $\vec{\hat{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \vec{b}$, which is applicable when $\mathbf{A}$ has linearly independent columns. We would now like to walk through the intuition behind this formula for the case when $\mathbf{A}$ has orthogonal columns: $\langle \vec{a}_1, \vec{a}_2 \rangle = 0$.

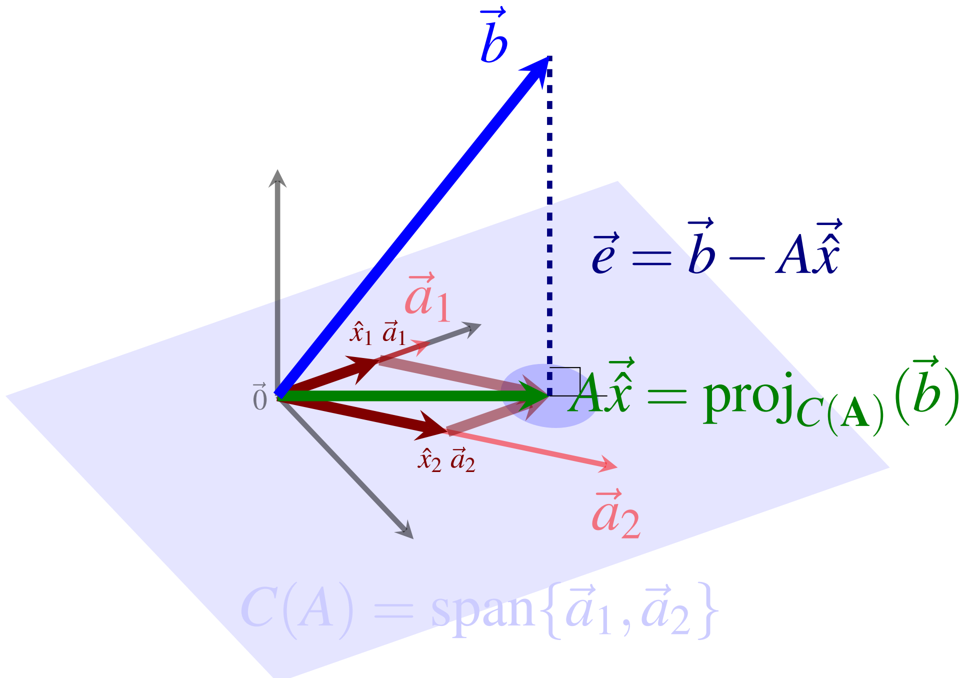
(a) On the diagram below, please label the following elements:

NOTE: For this sub-part only, the matrix $\mathbf{A}$ does not have orthogonal columns.

$\text{span}\{\vec{a}_1, \vec{a}_2\}$ $\quad$ $\mathbf{A}\,\vec{\hat{x}}$ $\quad$ $\hat{x}_1\,\vec{a}_1$ $\quad$ $\hat{x}_2\,\vec{a}_2$ $\quad$ $C(\mathbf{A})$ $\quad$ $\vec{e} = \vec{b} - \mathbf{A}\,\vec{\hat{x}}$ $\quad$ $\text{proj}_{C(\mathbf{A})}(\vec{b})$.



Answer:

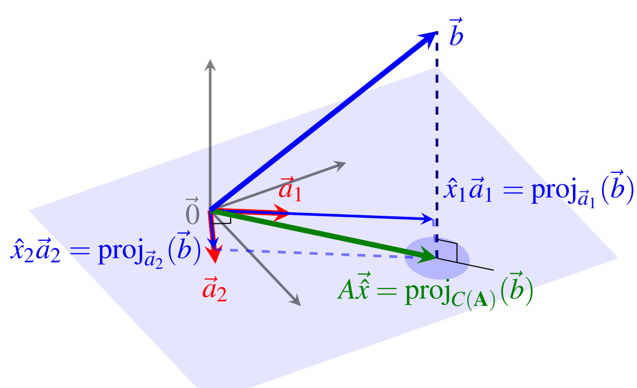


(b) Now suppose we assume a special case of the least squares problem where the columns of $\mathbf{A}$ are orthogonal (illustrated in the figure below). Given that $\vec{\hat{x}} = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\vec{b}$, and $\text{proj}_{C(\mathbf{A})}(\vec{b}) = \mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\vec{b} = \mathbf{A}\vec{\hat{x}}$, show the following statement holds.

$$\langle\vec{a}_1,\vec{a}_2\rangle = 0 \qquad \Longrightarrow \qquad \vec{\hat{x}} = \begin{bmatrix} \frac{\langle\vec{a}_1,\vec{b}\rangle}{||\vec{a}_1||^2} \\ \frac{\langle\vec{a}_2,\vec{b}\rangle}{||\vec{a}_2||^2} \end{bmatrix} \qquad \text{and} \qquad \text{proj}_{C(\mathbf{A})}(\vec{b}) = \text{proj}_{\vec{a}_1}(\vec{b}) + \text{proj}_{\vec{a}_2}(\vec{b})$$

In words, the statement says that when the columns of $\mathbf{A}$ are orthogonal, the entries of the least squares solution vector $\vec{\hat{x}}$ can be computed by using $\vec{b}$ and only the single other vector $\vec{a}_i$, and that the projection of $\vec{b}$ onto $C(\mathbf{A})$ can be computed by summing the projections of $\vec{b}$ onto the $\vec{a}_i$.

RECALL... $\quad \text{proj}_{\vec{a}_1}(\vec{b}) = \frac{\langle\vec{a}_1,\vec{b}\rangle}{||\vec{a}_1||^2}\vec{a}_1, \quad ||\vec{a}||^2 = \langle\vec{a},\vec{a}\rangle \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad \mathbf{A} = \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix}$



Answer: The derivation follows if we directly plug into our least-squares formula:

$$
\begin{aligned}
\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix} &= \left(\begin{bmatrix} — & \vec{a}_1^\top & — \\ — & \vec{a}_2^\top & — \end{bmatrix} \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix} \right)^{-1} \left(\begin{bmatrix} — & \vec{a}_1^\top & — \\ — & \vec{a}_2^\top & — \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \right) \\
&= \begin{bmatrix} \langle \vec{a}_1, \vec{a}_1 \rangle & \langle \vec{a}_1, \vec{a}_2 \rangle \\ \langle \vec{a}_2, \vec{a}_1 \rangle & \langle \vec{a}_2, \vec{a}_2 \rangle \end{bmatrix}^{-1} \begin{bmatrix} \langle \vec{a}_1, \vec{b} \rangle \\ \langle \vec{a}_2, \vec{b} \rangle \end{bmatrix} \\
&= \begin{bmatrix} \langle \vec{a}_1, \vec{a}_1 \rangle & 0 \\ 0 & \langle \vec{a}_2, \vec{a}_2 \rangle \end{bmatrix}^{-1} \begin{bmatrix} \langle \vec{a}_1, \vec{b} \rangle \\ \langle \vec{a}_2, \vec{b} \rangle \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\langle \vec{a}_1, \vec{a}_1 \rangle} & 0 \\ 0 & \frac{1}{\langle \vec{a}_2, \vec{a}_2 \rangle} \end{bmatrix} \begin{bmatrix} \langle \vec{a}_1, \vec{b} \rangle \\ \langle \vec{a}_2, \vec{b} \rangle \end{bmatrix} \\
&= \begin{bmatrix} \frac{\langle \vec{a}_1, \vec{b} \rangle}{\langle \vec{a}_1, \vec{a}_1 \rangle} \\ \\ \frac{\langle \vec{a}_2, \vec{b} \rangle}{\langle \vec{a}_2, \vec{a}_2 \rangle} \end{bmatrix} = \begin{bmatrix} \frac{\langle \vec{a}_1, \vec{b} \rangle}{||\vec{a}_1||^2} \\ \\ \frac{\langle \vec{a}_2, \vec{b} \rangle}{||\vec{a}_2||^2} \end{bmatrix}
\end{aligned}
$$

From this point we can quickly see that the projection onto $C(\mathbf{A})$ becomes the sum of projections onto $\vec{a}_1$ and $\vec{a}_2$ individually

$$
\text{proj}_{C(\mathbf{A})}(\vec{b}) = \mathbf{A}\vec{\hat{x}} = \hat{x}_1 \vec{a}_1 + \hat{x}_2 \vec{a}_2 = \left(\frac{\langle \vec{a}_1, \vec{b} \rangle}{||\vec{a}_1||^2} \right) \vec{a}_1 + \left(\frac{\langle \vec{a}_2, \vec{b} \rangle}{||\vec{a}_2||^2} \right) \vec{a}_2 = \text{proj}_{\vec{a}_1}(\vec{b}) + \text{proj}_{\vec{a}_2}(\vec{b}). \qquad \square
$$

(c) Compute the least squares solution $\vec{x} \in \mathbb{R}^2$ to the following system:

$$
\min_{\vec{x} \in \mathbb{R}^2} \left\| \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2 .
$$

HINT: Notice that the columns of $\mathbf{A}$ are orthogonal!!

Answer: As we have shown in the previous step, there are two ways to find $\vec{\hat{x}}$:

(1) Use the least-squares formula: $\vec{\hat{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \vec{b}$.

(2) Use the projection property for each term: $\hat{x}_j \vec{a}_j = \text{proj}_{\vec{a}_j}(\vec{b}) = \left(\frac{\langle \vec{a}_j, \vec{b} \rangle}{\langle \vec{a}_j, \vec{a}_j \rangle} \right) \vec{a}_j$.

IMPORTANT: This second bullet only works when the columns of $\mathbf{A}$ are orthogonal.
Since the second bullet is far easier to compute, we will employ this method here:

$$
\begin{aligned}
\hat{x}_1 &= \frac{\langle \vec{a}_1, \vec{b} \rangle}{\langle \vec{a}_1, \vec{a}_1 \rangle} = \frac{1 \cdot 1 + 0 \cdot 2 + 0 \cdot 3}{1 \cdot 1 + 0 \cdot 0 + 0 \cdot 0} = 1 \\
\hat{x}_2 &= \frac{\langle \vec{a}_2, \vec{b} \rangle}{\langle \vec{a}_2, \vec{a}_2 \rangle} = \frac{0 \cdot 1 + 1 \cdot 2 + 1 \cdot 3}{0 \cdot 0 + 1 \cdot 1 + 1 \cdot 1} = \frac{5}{2}
\end{aligned}
$$

Thus our result is

$$
\vec{\hat{x}} = \begin{bmatrix} 1 \\ 5/2 \end{bmatrix}. \qquad \square
$$

