EECS 16A Designing Information Devices and Systems I
Fall 2020 Discussion 13A

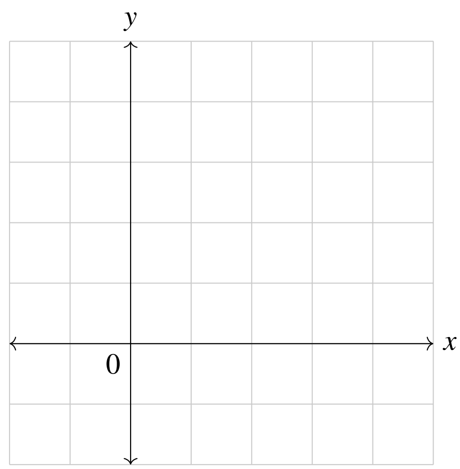
1. Mechanical Projection

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

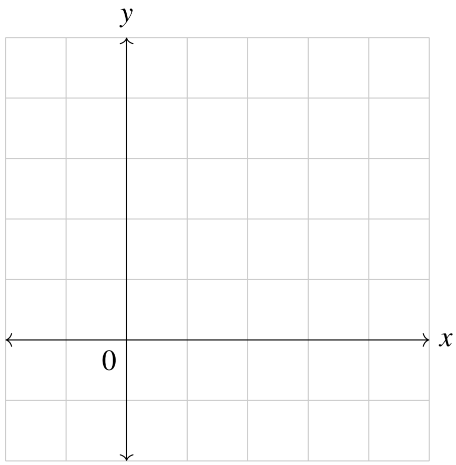
$$\text{proj}_{\vec{a}}\left(\vec{b}\right) = \frac{\left\langle \vec{a}, \vec{b} \right\rangle}{\|\vec{a}\|^2}\vec{a}.$$

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

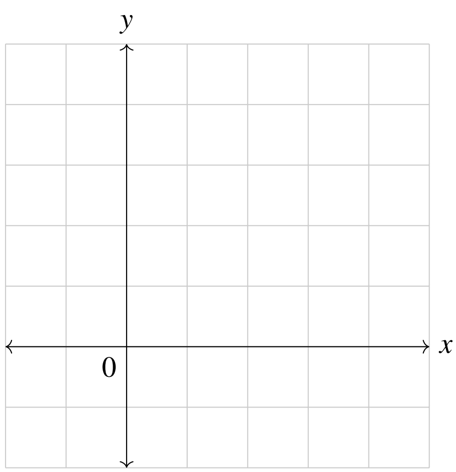
(a) Project $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ onto $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ – that is, onto the x-axis. Graph these two vectors and the projection.



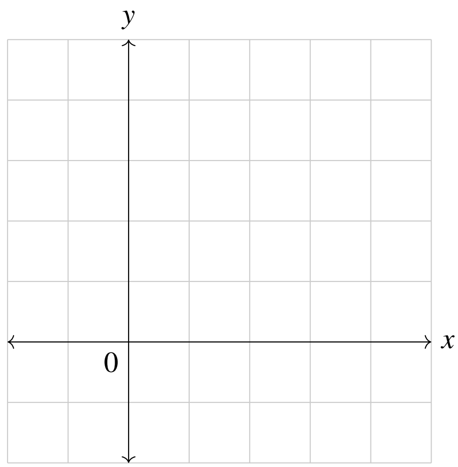
(b) Project $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$ onto $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ – that is, onto the y-axis. Graph these two vectors and the projection.



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



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



(e) 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 $C(A)$.)

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

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 $\hat{\vec{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$. Mathematically, we can express this as:

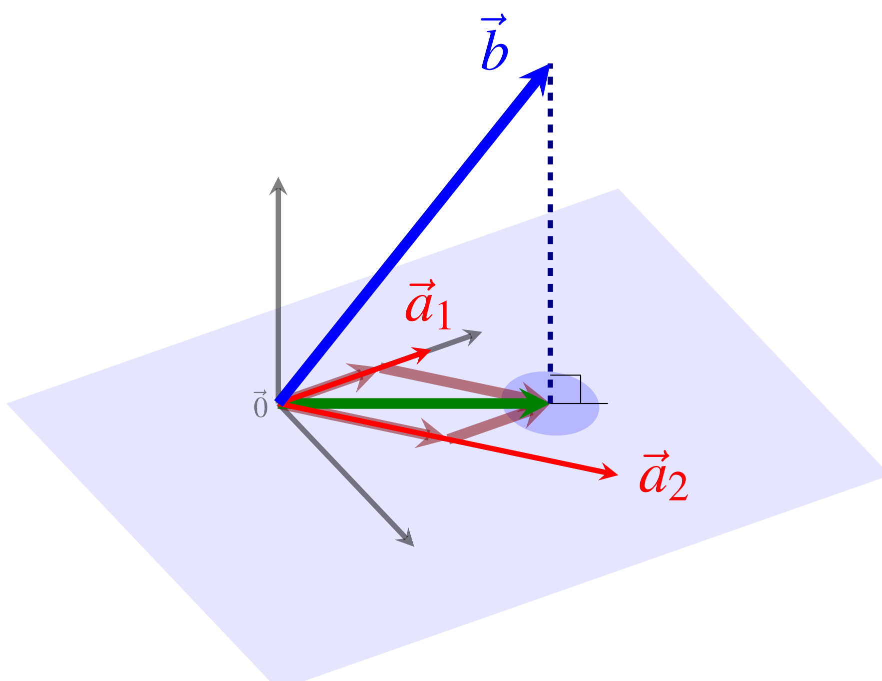
$$||\vec{b} - \mathbf{A}\hat{\vec{x}}||^2 = \min_{\vec{x}\in\mathbb{R}^2} ||\vec{b} - \mathbf{A}\vec{x}||^2 = \min_{\vec{x}\in\mathbb{R}^2} \left\| \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} - \begin{bmatrix} | & | \\ \vec{a}_1 & \vec{a}_2 \\ | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}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

To identify the solution $\hat{\vec{x}}$, we may recall the least squares formula: $\hat{\vec{x}} = (\mathbf{A}^\top \mathbf{A})^{-1}\mathbf{A}^\top \vec{b}$, 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.

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



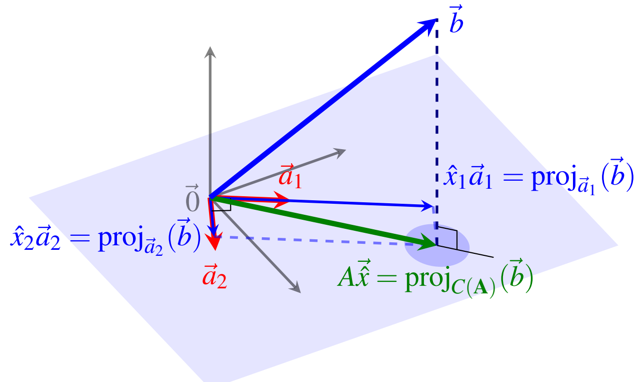
(b) Now suppose we assume a special case of the least squares problem where the columns of $\mathbf{A}$ are orthogonal (illustrated in the figure below). Given that $\hat{\vec{x}} = (\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\vec{b}$, and $\text{proj}_{C(\mathbf{A})}(\vec{b}) = \mathbf{A}(\mathbf{A}^\top\mathbf{A})^{-1}\mathbf{A}^\top\vec{b}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

$\mathbf{A}\vec{\hat{x}}$, show the following statement holds.

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

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



(c) Compute the least squares solution $\vec{\hat{x}} \in \mathbb{R}^2$ 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!!

