EECS 16A Designing Information Devices and Systems I
Spring 2019 Discussion 13A

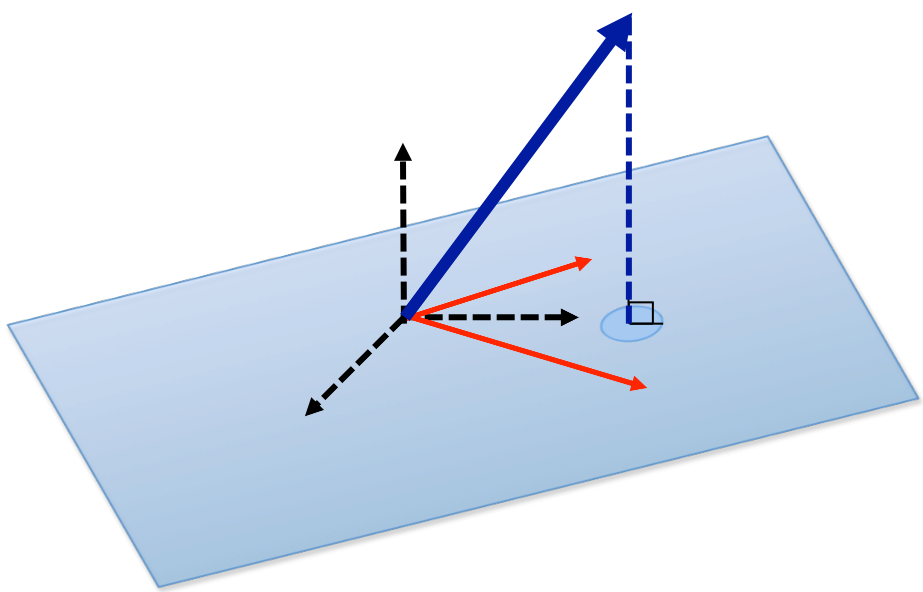
1. **Linear Least Squares with Orthogonal Columns**

(a) **Geometric Interpretation of Linear Least Squares**
Consider a linear least squares problem of the form

$$\min_{\vec{x}} \left\| \vec{b} - \mathbf{A}\vec{x} \right\|^2 = \min_{\vec{x}} \left\| \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} - \begin{bmatrix} | & | \\ A_1 & A_2 \\ | & | \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|^2$$

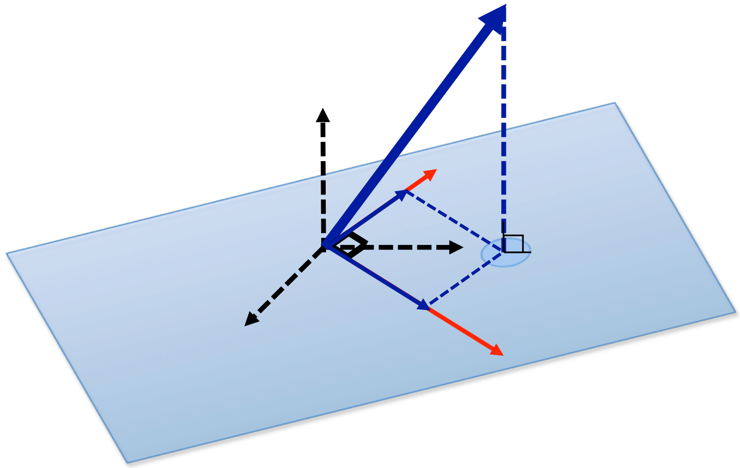
Let the solution be $\vec{\hat{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$.
Label the following elements in the diagram below.

$\vec{b}$, A_1, A_2, $\text{span}\{A_1, A_2\}$, $\vec{\hat{e}} = \vec{b} - \mathbf{A}\vec{\hat{x}}$, $\mathbf{A}\vec{\hat{x}}$, $A_1\hat{x}_1$, $A_2\hat{x}_2$



(b) We now consider the special case of linear least squares where the columns of $\mathbf{A}$ are orthogonal (illustrated in the figure below). Use the linear least squares formula $\vec{\hat{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{b}$ to show that

$$\hat{x}_1 = \text{factor by which } A_1 \text{ is scaled to produce the projection of } \vec{b} \text{ onto } A_1,$$
$$\hat{x}_2 = \text{factor by which } A_2 \text{ is scaled to produce the projection of } \vec{b} \text{ onto } A_2.$$



(c) Compute the linear least squares solution to

$$\min_{\vec{x}} \left\| \begin{bmatrix}1\\2\\3\\4\\5\end{bmatrix} - \begin{bmatrix}1&0&0\\0&0&0\\0&1&0\\0&0&0\\0&0&1\end{bmatrix} \begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix} \right\|^2.$$

2. Orthonormal Matrices and Projections

An orthonormal matrix, $\mathbf{A}$, is a matrix whose columns, $\vec{a}_i$, are:

- Orthogonal (ie. $\langle \vec{a}_i, \vec{a}_j \rangle = 0$ when $i \neq j$)
- Normalized (ie. vectors with length equal to 1, $\|\vec{a}_i\| = 1$). This implies that $\|\vec{a}_i\|_2 = \langle \vec{a}_i, \vec{a}_i \rangle = 1$.

(a) Suppose that the matrix $\mathbf{A} \in \mathbb{R}^{N \times M}$ has linearly indpendent columns. The vector $\vec{y}$ in $\mathbb{R}^N$ is not in the subspace spanned by the columns of $\mathbf{A}$. What is the projection of $\vec{y}$ onto the subspace spanned by the columns of $\mathbf{A}$?

(b) Show if $\mathbf{A} \in \mathbb{R}^{N \times N}$ is an orthonormal matrix then the columns, $\vec{a}_i$, form a basis for $\mathbb{R}^N$.

(c) When $\mathbf{A} \in \mathbb{R}^{N \times M}$ and $N \geq M$ (i.e. tall matrices), show that if the matrix is orthonormal, then $\mathbf{A}^T\mathbf{A} = \mathbf{I}_{M \times M}$.

(d) Again, suppose $\mathbf{A} \in \mathbb{R}^{N \times M}$ where $N \geq M$ is an orthonormal matrix. Show that the projection of $\vec{y}$ onto the subspace spanned by the columns of $\mathbf{A}$ is now $\mathbf{A}\mathbf{A}^T\vec{y}$.