EECS 16A Designing Information Devices and Systems I
Summer 2020 Discussion 6D

1. Least Squares: A Toy Example

Let's start off by solving a little example of least squares.

We're given the following system of equations:

$$\begin{bmatrix} 1 & 4 \\ 3 & 8 \\ 5 & 16 \end{bmatrix} \vec{x} = \begin{bmatrix} 3 \\ 1 \\ 9 \end{bmatrix},$$

where $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$.

(a) Why can we not solve for $\vec{x}$ exactly?

(b) Find $\vec{\hat{x}}$, the *least squares estimate* of $\vec{x}$, using the formula we derived in lecture.

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

2. Least Squares with Orthogonal Columns

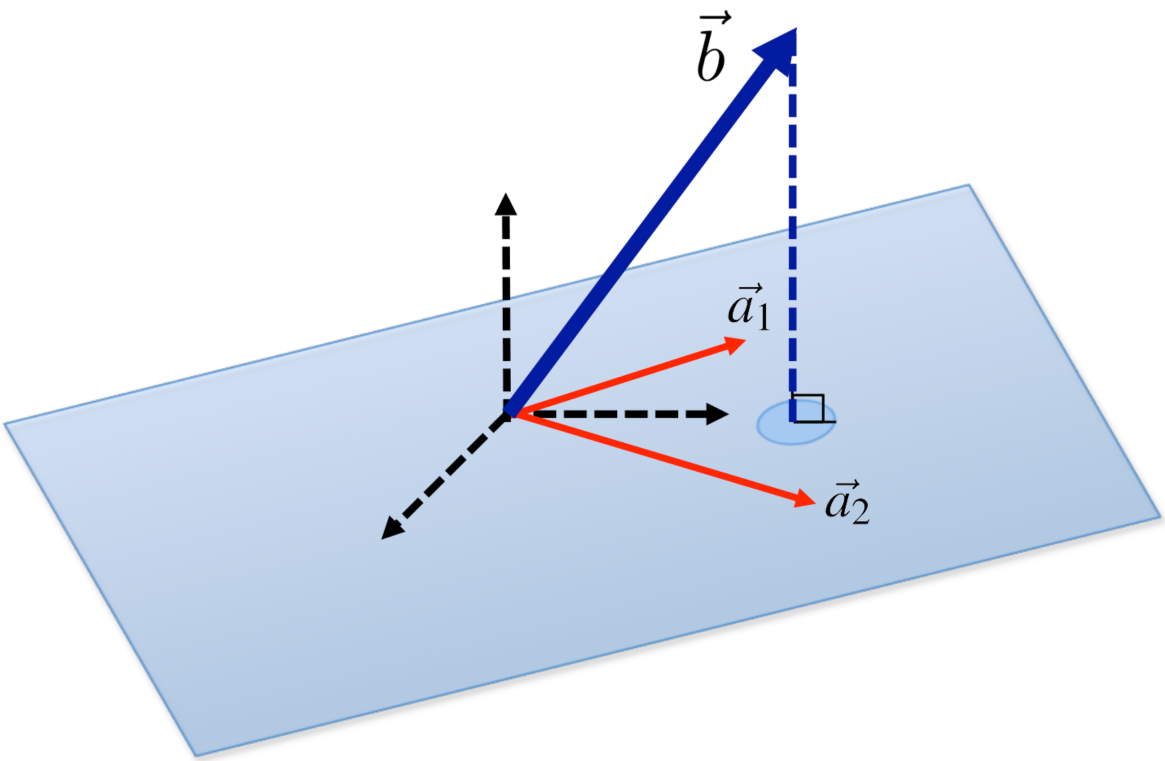
(a) Consider a least squares problem of the form

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

Let the solution be $\vec{\hat{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$.

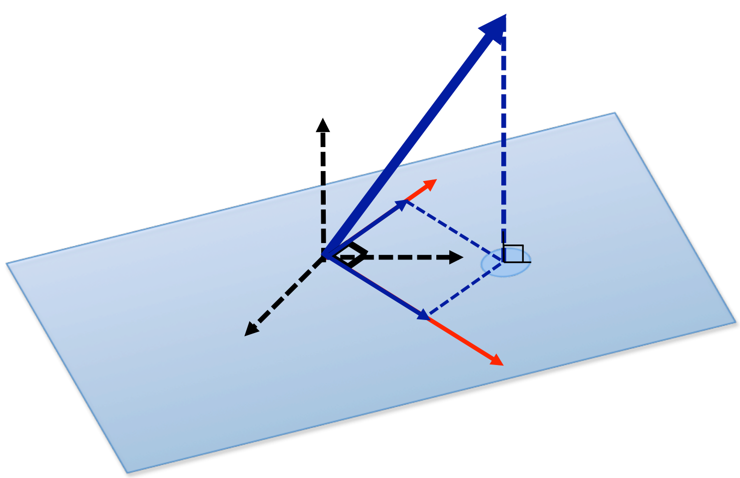
Label the following elements in the diagram below.

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



(b) We now consider the special case of least squares where the columns of $\mathbf{A}$ are orthogonal (illustrated in the figure below). Given that $\vec{\hat{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\vec{b}$ and $A\vec{\hat{x}} = \text{proj}_{\mathbf{A}}(\vec{b}) = \hat{x_1}\vec{a}_1 + \hat{x_2}\vec{a}_2$, show that

$$\text{proj}_{\vec{a}_1}(\vec{b}) = \hat{x_1}\vec{a}_1$$
$$\text{proj}_{\vec{a}_2}(\vec{b}) = \hat{x_2}\vec{a}_2$$



(c) Compute the least squares solution to

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

3. Polynomial Fitting

Let's try an example. Say we know that the output, y, is a quartic polynomial in x. This means that we know that y and x are related as follows:

$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4$$

We're also given the following observations:

x	y
0.0	24.0
0.5	6.61
1.0	0.0
1.5	−0.95
2.0	0.07
2.5	0.73
3.0	−0.12
3.5	−0.83
4.0	−0.04
4.5	6.42

(a) What are the unknowns in this question? What are we trying to solve for?

(b) Can you write an equation corresponding to the first observation (x_0, y_0), in terms of a_0, a_1, a_2, a_3, and a_4? What does this equation look like? Is it linear in the unknowns?

(c) Now, write a system of equations in terms of a_0, a_1, a_2, a_3, and a_4 using *all of the observations*.

(d) Finally, solve for a_0, a_1, a_2, a_3, and a_4 using IPython. You have now found the quartic polynomial that best fits the data!

