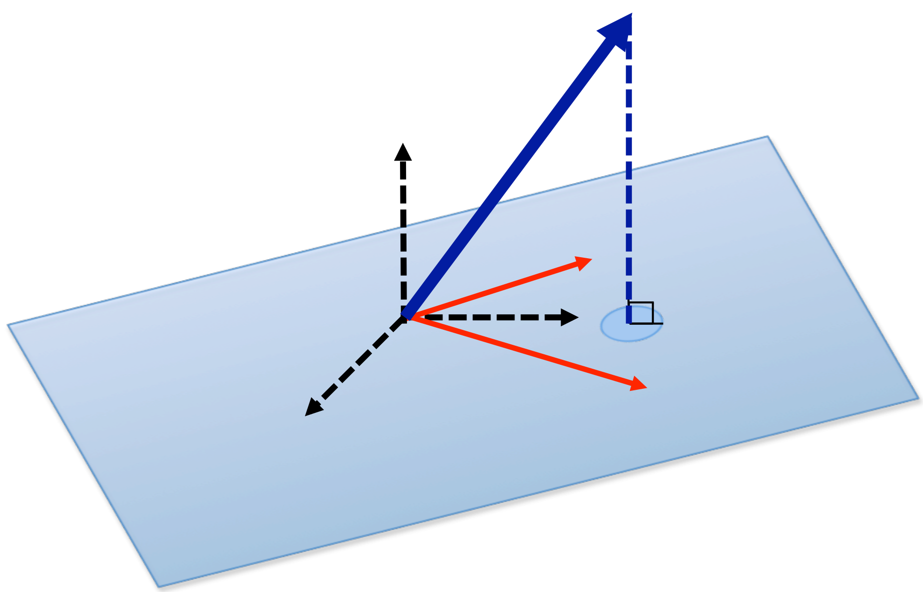
1. Linear Least Squares with Orthogonal Columns

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

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

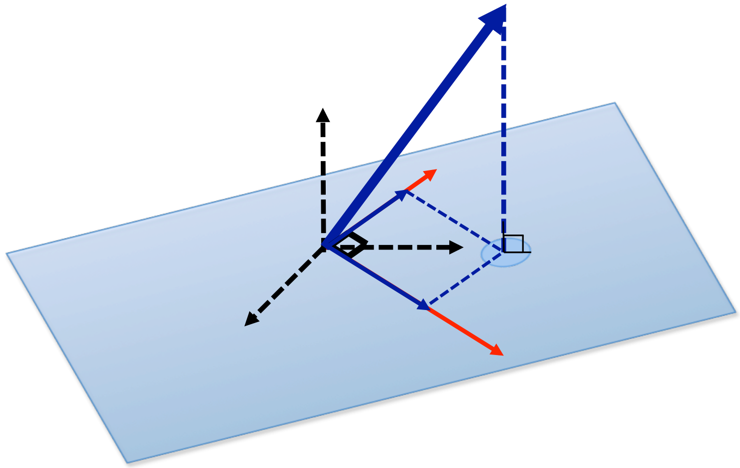
Let the solution be $\vec{\hat{x}} = \begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \end{bmatrix}$.
Label the following elements in the diagram below.

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



(b) We now consider the special case of linear least squares where the columns of **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. Polynomial Fitting

Least squares may seem rather boring at first glance – we're just using it to "solve" systems of linear equations, after all. But, at further glance, it actually comes in a variety of sizes and flavors! For instance, you can solve problems that have decidedly non-linear elements in them using least squares. Let's see how.

Last discussion, we had seen how to "fit" data in the form of (x,y) to a line. This made sense because the input-output relationship was fundamentally linear (Ohm's law).

But what if this relationship was not linear? For instance, the equation of the orbit of a planet around the sun is an ellipse. The equation for the trajectory of a projectile is a parabola. In these sorts of scenarios, how does one fit observation data to the correct curve?

In particular, 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:

$$y = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$$

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?

(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!

(e) We will now do another example in the IPython notebook and see how to do polynomial fitting quickly using IPython!