Minimization of Squared Error

Calculus allows us to minimize $\|e\|^2$ with respect to x_1 and x_2:

$$\begin{aligned} f(x_1, x_2) = \|e\|^2 = \|b - p\|^2 &= (b - x_1 a_1 + x_2 a_2, b - x_1 a_1 + x_2 a_2) \\ &= \|b\|^2 - 2x_1(a_1, b) - 2x_2(a_2, b) + 2x_1 x_2 (a_1, a_2) \\ &+ x_1^2 \|a_1\|^2 + x_2^2 \|a_2\|^2 \end{aligned}$$

$$0 = \frac{\partial f}{\partial x_1} = -2(a_1, b) + 2x_2(a_1, a_2) + 2x_1 \|a_1\|^2$$

$$0 = \frac{\partial f}{\partial x_2} = -2(a_2, b) + 2x_1(a_1, a_2) + 2x_2 \|a_2\|^2$$

That is:

$$\begin{bmatrix} \|a_1\|^2 & (a_1, a_2) \\ (a_1, a_2) & \|a_2\|^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (a_1, b) \\ (a_2, b) \end{bmatrix}$$

Now we just need to solve for x_1 and x_2.