Topics: Projection and least squares.

Do: Answer the following questions, and feel free to use a computer to solve *any* systems of equations you encounter throughout this problem.

1. Consider the four data points $(t_i, b_i) = (0,0)$, $(1,8)$, $(3,8)$, and $(4,20)$. Let $\boldsymbol{t} = (0,1,3,4)$ be the vector of inputs and $\boldsymbol{b} = (0,8,8,20)$ the vector of outputs.

   (a) If there were a straight line $b = C + Dt$ through these four points, then a certain equation $\boldsymbol{Ax} = \boldsymbol{b}$ would have a solution, where $\boldsymbol{x} = (C, D)$. Write this equation in matrix form (that is, find $\boldsymbol{A}$).

   (b) Instead, we wish to find the "best fit" line, which means we need to solve $\boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{p}$, where $\boldsymbol{p}$ is the projection of $\boldsymbol{b}$ onto the column space of $\boldsymbol{A}$. Write down the *normal equations* $\boldsymbol{A}^T\boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{A}^T\boldsymbol{b}$, where $\widehat{\boldsymbol{x}} = (\widehat{C}, \widehat{D})$, and solve for $\widehat{\boldsymbol{x}}$.

   (c) Check that $\boldsymbol{e} = \boldsymbol{b} - \boldsymbol{p}$ is orthogonal to both columns of $\boldsymbol{A}$. Compute $||\boldsymbol{e}||$, which is the shortest distance from $\boldsymbol{b}$ to the column space of $\boldsymbol{A}$. Sketch a diagram of $\boldsymbol{e}$, $\boldsymbol{b}$, $\boldsymbol{p}$, and the orthogonal subspaces $C(\boldsymbol{A})$ and $N(\boldsymbol{A}^T)$ to illustrate this.

   (d) Plot the four data points in $\mathbb{R}^2$ (on the tb-plane) and sketch the best fit line through them that you just found. Clearly mark what the vectors $\boldsymbol{b} = (b_1, b_2, b_3, b_4)$, $\boldsymbol{e} = (e_1, e_2, e_3, e_4)$, and $\boldsymbol{p} = (p_1, p_2, p_3, p_4)$ represent.

   (e) Write down $E := ||\boldsymbol{Ax} - \boldsymbol{b}||^2$ as a sum of four squares–the last one is $(C + 4D - 20)^2$, and compute $\partial E/\partial C$ and $\partial E/\partial D$. Set these derivatives equal to zero, and obtain scalars of the normal equations $\boldsymbol{A}^T\boldsymbol{A}\widehat{\boldsymbol{x}} = \boldsymbol{A}^T\boldsymbol{b}$.

   (f) The method above found the best fit degree-1 polynomial (line). Now, find the best fit degree-0 polynomial (horizontal line) $b = C$. Note that this will be a 4×1 system instead of a 4×2 system. Compute the vectors $\boldsymbol{p}$ and $\boldsymbol{e}$, and the (squared) error $||\boldsymbol{e}||^2$.

   (g) Find the best fit parabola (degree-2 polynomial) $b = C + Dt + Et^2$. On a new set of axes, plot the four data points and this parabola. Compute the vectors $\boldsymbol{p}$ and $\boldsymbol{e}$, and the (squared) error $||\boldsymbol{e}||^2$.

   (h) Find the best fit cubic (degree-3 polynomial) $b = C + Dt + Et^2 + Ft^3$. On a new set of axes, plot the four data points and this cubic. Compute the vectors $\boldsymbol{p}$ and $\boldsymbol{e}$, and the (squared) error $||\boldsymbol{e}||^2$.