Problem: find a line $y = mx+b$ to best fit some given data $\{(x_1, y_1), (x_2, y_2), ..., (x_k, y_k)\}$. The best solution minimizes the following sum of square differences:

$$\sum_{i=1}^{k} |mx_i + b - y_i|^2.$$

Equivalently, we are looking for a least squares solution to

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ . & . \\ . & . \\ . & . \\ x_k & 1 \end{pmatrix} \begin{pmatrix} m \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_k \end{pmatrix}.$$

1. Suppose we have the following data points in the xy-plane:

$$\{(1, 2), (3, 5), (4, 5), (6, 8), (6, 9), (7, 10)\}.$$

Find an equation for the line of best fit. Here the procedure:

(a) Find equations your data should fit if there were an exact solution. (The equations must be linear with respect to the parameters.)

(b) Write these equations as a linear system in a matrix form (inconsistent unless there is an exact solution).

(c) Find the least squares solution: left-multiply both sides by the conjugate transpose (adjoint) of the coefficient matrix. This gives the normal equation. Solve the normal equation.