WORKSHEET 3/5/18

(1) Find the best fit line for the points $(1,0)$, $(2,1)$, and $(3,3)$.

Solution We have $A = \begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. We calculate that $(A^T A)^{-1} = \frac{1}{6}\begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}$, so

$$\vec{x}^* = \frac{1}{6}\begin{bmatrix} 3 & -6 \\ -6 & 14 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

The best fit line is $y = x$.

(2) The following table shows the percentage of classes attended by Harvard students:

Year (y)	Percentage of Classes Attended (p)
1 (Freshman)	100
2 (Sophomore)	90
3 (Junior)	60
4 (Senior)	10

We suspect that $p(y)$ looks like ky^n for some constants k and n, and we would like to find k and n.

(a) Do you expect k and n to be positive or negative? What number should k be close to?

(b) How might we express this problem as a linear system?

Solution (a) Since $p(y)$ is positive, we expect k to be positive. Since $p(y)$ is decreasing, we expect n to be negative. Since $p(1) = 100$, k should be close to 100. (b) The equations

$$\begin{aligned} k \cdot 1^n &= 100 \\ k \cdot 2^n &= 90 \\ k \cdot 3^n &= 60 \\ k \cdot 4^n &= 10 \end{aligned}$$

aren't linear, but we can fix this by taking the natural logarithm to get

$$\begin{array}{ccccc} c & + & n \ln 1 & = \ln 100 \\ c & + & n \ln 2 & = \ln 90 \\ c & + & n \ln 3 & = \ln 60 \\ c & + & n \ln 4 & = \ln 10, \end{array}$$

where $c = \ln k$. This now a linear system in c and n.

(3) Find the best fit degree 2 polynomial for the points $(-1,1)$, $(0,0)$, $(1,2)$, and $(2,5)$.

Solution We have $A = \begin{bmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \\ 4 & 2 & 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 5 \end{bmatrix}$. We calculate that $(A^TA)^{-1} = \frac{1}{20}\begin{bmatrix} 5 & -5 & -5 \\ -5 & 9 & 3 \\ -5 & 3 & 11 \end{bmatrix}$, so

$$\vec{x}^* = \frac{1}{20}\begin{bmatrix} 5 & -5 & -5 \\ -5 & 9 & 3 \\ -5 & 3 & 11 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 & 4 \\ -1 & 0 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 2 \\ 5 \end{bmatrix} = \frac{1}{20}\begin{bmatrix} 25 \\ 8 \\ 6 \end{bmatrix}.$$

The best fit polynomial is $y = \frac{1}{20}(25x^2 + 8x + 6)$.

(4) Without using Gram-Schmidt, find the matrix of orthogonal projection onto the span of $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$. (Hint: if you forgot the formula, think of this as a least squares problem).

Solution We have $A = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}$. We calculate that $(A^TA)^{-1} = \begin{bmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{bmatrix}$, so the projection is given by

$$A(A^TA)^{-1}A^T = \begin{bmatrix} 1 & 0 \\ 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} \frac{1}{5} & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

(5) Let A be an $m \times n$ matrix.
 (a) Show that $\ker(A) = \ker(A^TA)$ (Hint: If a vector $\vec{x}$ is in the kernel of A^TA, then $A\vec{x}$ is in the image of A and in the kernel of A^T).
 (b) Use part (a) to conclude that $\ker(A) = \{\vec{0}\}$ if and only if A^TA is invertible.

Solution (a) If $A\vec{x} = \vec{0}$, then $A^TA\vec{x} = A^T\vec{0} = \vec{0}$, so the kernel of A is contained in the kernel of A^TA. In order to show that they're the same, we need to show that the kernel of A^TA is contained in the kernel of A. So suppose that $\vec{x}$ is a vector such that $A^TA\vec{x} = \vec{0}$; then $A\vec{x}$ is both in the image of A and in the kernel of A^T. Since $\ker(A^T) = \text{im}(A)^\perp$, it follows that $A\vec{x}$ is orthogonal to every vector in $\text{im}(A)$—so $A\vec{x}$ is orthogonal to itself! The only vector orthgonal to itself is the zero vector, so $A\vec{x} = \vec{0}$, which is to say that $\vec{x}$ is in the kernel of A. (b) By part (a), $\ker(A) = \vec{0}$ if and only if $\ker(A^TA) = \{\vec{0}\}$, which is equivalent to the statement that A^TA has a pivot in every column (by rank-nullity). Since A^TA is a square matrix, this condition is equivalent to invertibility.

(6) Decide whether each of the following statements is true or false. If the statement is true, explain why briefly; if the statement is false, give a counterexample (Hint: exactly one of the statements is false).
 (a) The least squares solutions of $A\vec{x} = \vec{b}$ are exactly the solutions of $A\vec{x} = \text{proj}_{\text{im}(A)}(\vec{b})$.
 (b) If $\vec{x}^*$ is a least squares solution of $A\vec{x} = \vec{b}$, then $||\vec{b}||^2 = ||A\vec{x}^*||^2 + ||\vec{b} - A\vec{x}^*||^2$.
 (c) Every linear system has a unique least squares solution.
 (d) Even if the system $A\vec{x} = \vec{b}$ is inconsistent, the system $A^TA\vec{x} = A^T\vec{b}$ is consistent.
 (e) For any matrix A, $(\ker A)^\perp = \text{im}(A^T)$.

Solution (a) True; this is the definition of a least squares solution. (b) True; if $\vec{x}^*$ is a least squares solution, then $A\vec{x}^* = \mathrm{proj}_{\mathrm{im}(A)}(\vec{b})$, so $\mathrm{proj}_{\mathrm{im}(A)}(\vec{b})$ and $\vec{b} - \mathrm{proj}_{\mathrm{im}(A)}(\vec{b})$ are orthogonal, and the equation in question follows from the Pythagorean theorem. (c) False; for example, every vector is a least squares to the system $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$. (d) True; this is equivalent to asking whether $A\vec{x} = \mathrm{proj}_{\mathrm{im}(A)}(\vec{b})$ has a solution, which is to ask whether $\mathrm{proj}_{\mathrm{im}(A)}(\vec{b})$ is in the image of A—and it is! (e) True; we know already that $\mathrm{im}(A^T)^\perp = \ker((A^T)^T) = \ker(A)$, so apply "perp" to both sides of this equation.

The **least squares solutions** to the linear system $A\vec{x} = \vec{b}$ are the exact solutions to the **normal equation**

$$A^T A\vec{x}^* = A^T\vec{b}.$$

If $\ker(A) = \vec{0}$, then there is a unique least squares solution, which is given by

$$\vec{x}^* = (A^T A)^{-1} A^T\vec{b}.$$