Math 104, Summer 2010
Homework 4

1. As explained at the beginning of chapter 11 in the text, finding the best fit line \( y = ax + b \) through these points is equivalent to minimizing

\[
||Ax - b|| = \left\| \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} b \\ a \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \right\|
\]

We have seen that the solution can be found by putting \( x = (A^*A)^{-1}A^*b \). We compute

\[
A^*A = \begin{pmatrix} 3 & 3 \\ 3 & 5 \end{pmatrix}, \text{ so } (A^*A)^{-1} = \frac{1}{6} \begin{pmatrix} 5 & -3 \\ -3 & 3 \end{pmatrix}.
\]

Also \( A^*b = (-2, -5)^T \). Thus our solution is

\[
\begin{pmatrix} b \\ a \end{pmatrix} = (A^*A)^{-1}A^*b = \frac{1}{6} \begin{pmatrix} 5 \\ -9 \end{pmatrix}
\]

so the line of best fit (based on the least squares approach) is \( y = -\frac{3}{2}x + \frac{5}{6} \).

2. (a) Writing out the matrix-vector multiplications in \( a^T \hat{x} = a^T b \), we get \( n\hat{x} = b_1 + ... + b_n \), so \( \hat{x} = \frac{1}{n}(b_1 + ... + b_n) \), which is the average or mean of the components of \( b \).

(b) Writing \( \hat{x} = \frac{1}{n}(b_1 + ... + b_n) \) for brevity, the variance is \( \frac{1}{n}||r||^2 = \frac{1}{n} \sum_{i=1}^{n} (b_i - \hat{x})^2 \) while the standard deviation is \( \frac{1}{\sqrt{n}}||r|| = \frac{1}{\sqrt{n}} (\sum_{i=1}^{n} (b_i - \hat{x})^2)^{1/2} \).

3. (a) We have seen that \( ||A||_1 \) is just the “maximum column sum” of \( A \), so in this case \( ||A||_1 = |\cos \theta| + |\sin \theta| \). (If you had followed the hint, you would have basically reproven this result in this special case.)

(b) Using that \( \sin^2 \theta + \cos^2 \theta = 1 \), we can check that \( A^*A = I \), i.e., \( A \) is unitary. Hence \( ||AB||_2 = ||B||_2 \) for any \( 2 \times n \) matrix \( B \). In particular \( ||A||_2 = ||AI||_2 = ||I||_2 = 1 \) (the last equality being clear from the definition of induced matrix norms).