WORKSHEET #15, 10/16/07

MATH 54, FALL 2007

1. Suppose I have a list of data $b_1, b_2, \ldots, b_n$ and I want to approximate it by a number c (you could think of this as a constant function) such that $\sum_i (b_i - c)^2$ is smallest.

(a) Let $\overrightarrow{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$ and $\overrightarrow{x} = [x]$. Reinterpret the above problem in terms of finding the least-squares solution to $A\overrightarrow{x} = \overrightarrow{b}$. What is the matrix A?

(b) Find (in terms of the b_i's) the least squares solution $x^* = [c]$. Is it what you would expect (i.e. does it make sense intuitively)?

Note: In this case the error $\sum_i (b_i - c)^2$ is called the variance of the set of data and the distance from $A[c]$ to $\overrightarrow{b}$ (i.e. $||\overrightarrow{b} - A[c]||$) is $\sigma = \sqrt{\sum_i (b_i - c)^2}$, which is known as the standard deviation.

2. (a) Write down the matrix for the transformation T from P_1 to $\mathbb{R}^4$ given by $T(p(x)) = \begin{bmatrix} p(-1) \\ p(0) \\ p(1) \\ p(2) \end{bmatrix}$ with respect to the bases $\beta := (1, x)$ for P_1 and the standard basis $\gamma := (e_1, e_2, e_3, e_4)$ for $\mathbb{R}^4$.

(b) Find the "best fit" (i.e. least-squares solution) line to the data $(-1, 1)$,$(0, 2)$, $(1, 2)$, $(2, 5)$. (Hint: Is the matrix from part (a) relevant?)

(c) Make a sketch of your data and your best fit line.

3. Is $\ker(A) = \ker(A^T A)$ necessarily true? (Hint: Can you use $\ker(A^T) = \text{im}(A)^\perp$?)

4. (a) Show that if the columns of A are orthogonal and of unit length, then the least squares solution to $A\overrightarrow{x} = \overrightarrow{b}$ is simply given by $\overrightarrow{x}^* = A^T\overrightarrow{b}$.

(b) Relate this to orthogonal projection to the subspace spanned by the columns of A.