Least Squares and Condition Number

Your work should be presented in the form of a typed report using clear and properly punctuated English. Where appropriate include full program listings and output. If you choose to work in a group of two, please turn in independently prepared reports.

1a. The data in `file01.dat` consists of m values for x_i and y_i one pair of values per line. Plot the points (x_i, y_i) from this file on a graph.

1b. Create the matrix A and vector y given by

$$A = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}$$

where $n = 6$ and m is the number of lines in `file01.dat`. Find the condition number of A^tA. You may use the Matlab command `cond` to do this. Clearly state and define which matrix norm was used to compute the condition number.

1c. The normal equations for the least-squares fit of the polynomial

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

to the data in `file01.dat` are given by $A^tAv = A^ty$ where $v = [\,a_0 \quad a_1 \quad \dots \quad a_n\,]^t$. Solve these equations to find a numeric approximation of v. Print your results to 15 digits with `format long` in Matlab. Use the condition number of A^tA to estimate how many digits are correct in your approximation of the least-squares minimizer.

1d. List the Legendre polynomials $P_k(x)$ of degree k for $k = 0, 1, \dots, 6$.

1e. Let $\beta = \max\{\, x_i : i = 1, \dots, n \,\}$, $\alpha = \min\{\, x_i : i = 1, \dots, n \,\}$, $c = (\beta + \alpha)/2$ and $d = (\beta - \alpha)/2$. Let $Q_k(x) = P_k((x - c)/d)$ to obtain polynomials that are orthogonal on the interval $[\alpha, \beta]$. Create the matrix

$$B = \begin{bmatrix} Q_0(x_1) & Q_1(x_1) & Q_2(x_1) & \dots & Q_n(x_1) \\ Q_0(x_2) & Q_1(x_2) & Q_2(x_2) & \dots & Q_n(x_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Q_0(x_m) & Q_1(x_m) & Q_2(x_m) & \dots & Q_n(x_m) \end{bmatrix}.$$

where $n = 6$ and m is the number of lines in `file01.dat`. Find the condition number of the matrix B^tB.

1f. The normal equations for the least-squares fit of the polynomial

$$p(x) = b_0Q_0(x) + b_1Q_1(x) + \cdots + b_nQ_n(x)$$

are given by $B^tBw = B^ty$ where $w = [\,b_0 \quad b_1 \quad \dots \quad b_n\,]^t$. Solve these equations to find a numeric approximation of w. Use the condition number of B^tB to estimate how many digits are correct in your approximation of the least-squares minimizer.

1g. [Extra Credit and for Math/CS 667] Use the definition of $Q_k(x)$ to find values for the a_i's in part 1c directly from the b_i's computed in part 1f. Which way of computing the a_i's results in a more accurate calculation? Relate the differences in the two ways of computing the a_i's to the condition numbers of A^tA and B^tB.