AMS213A Project 2

Sutirtha Sengupta (Student id#:1566652)

February 16, 2017

Problem 1

The Fortran program myprog.f90 solves the least square problem using Cholesky decomposition of the corresponding Vandermond matrix A (performed by the routine Choldec in LinAl.f90 module) by casting it into the standard linear form:

$$Ax = b \tag{1}$$

where b is a column vector containing the y-values corresponding to the x vector read from the atkinson.dat file (given as an argument to the program). The vector b is replaced by the solution for x on executing the program.
Figure shows the fit to the data using with a 3^{rd} order polynomial given by

$$f(x) = 0.574662268 + 4.72584343x - 11.1281977x^2 + 7.66867352x^3 \tag{2}$$

with a rms error given by

$$E = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2} = 0.0420603156 \tag{3}$$

Figure 2 hows the fit to the data using with a 5^{th} order polynomial given by

$$f(x) = 0.511144936+7.14999056x-28.0456734x^2+51.0233231x^3-46.5010490x^4+17.7211895x^5 \tag{4}$$

with a rms error given by

$$E = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2} = 0.0306545403 \tag{5}$$

The algorithm breaks down for order of polynomial exceeding 5.

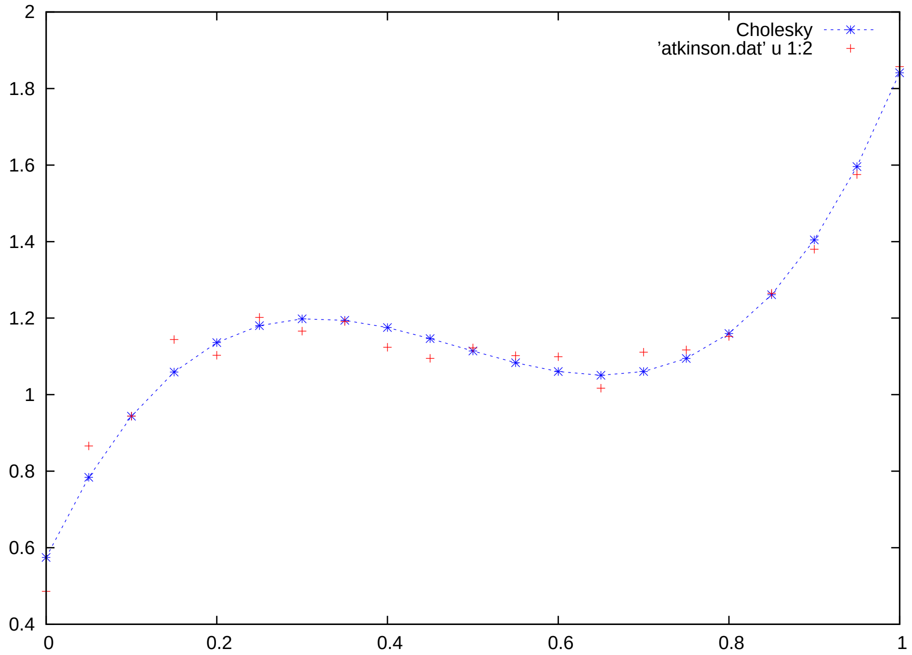

Figure 1: least square fit with 3^{rd}order polynomial (Eq. 2) to the data in atkinson.dat using Cholesky decomposition

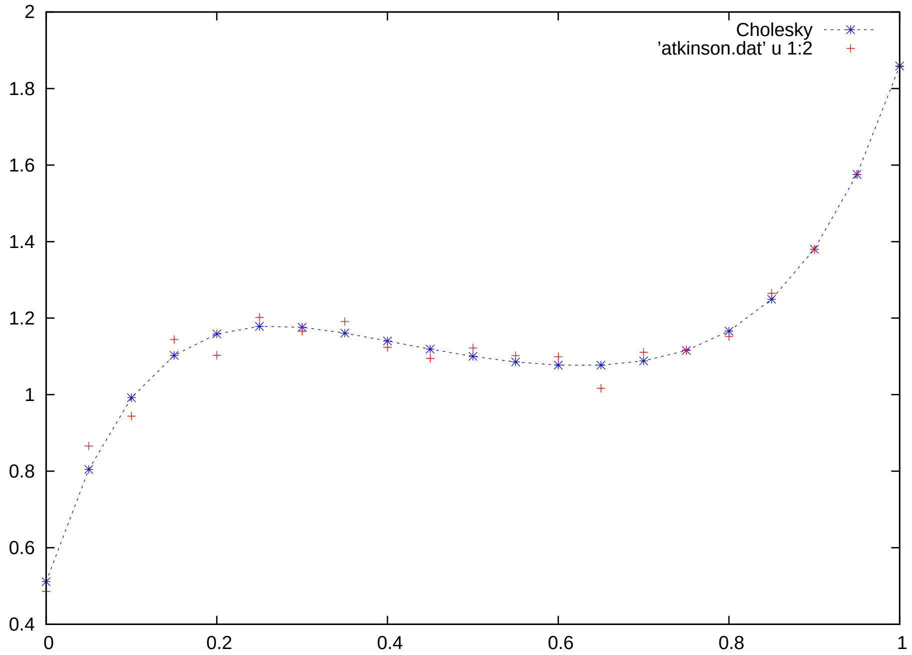

Figure 2: least square fit with 5^{th}order polynomial (Eq. 4) to the data in atkinson.dat using Cholesky decomposition

Problem 2

The Householder matrix is defined by

$$H_j = I_m - 2v_j v_j^T \tag{6}$$

with

$$\begin{aligned} v_j &= \frac{(0, ..., 0, a_{jj} + s_j, a_{j+1,j}, ..., a_{mj})^T}{||v_j||}, && (7) \\ (||v_j||^2 = (a_{jj} + s_j)^2 + &\sum_{k=j+1}^{m} a_{kj}^2 = 2s_j^2 + 2a_{jj}s_j, && (8) \\ s_j &= sign(a_{jj})(\sum_{k=j}^{m} a_{kj}^2)^{\frac{1}{2}} && (9) \end{aligned}$$

Operating on the the j-th column of A denoted by a_j, we get

$$H_j a_j = a_j - 2v_j(v_j^T a_j) \tag{10}$$

$$v_j^T a_j = \frac{1}{||v_j||}(0, ..., 0, a_{jj} + s_j, a_{j+1,j}, ..., a_{mj}) \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{jj} \\ a_{j+1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} \tag{11}$$

$$= \frac{1}{||v_j||}(a_{jj}s_j + \sum_{k=j}^{m} a_{kj}^2) \tag{12}$$

Using this in Eq.10, we have

$$H_j a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{jj} \\ a_{j+1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} - \frac{2s_j(a_{jj} + s_j)}{||v_j||^2} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ a_{jj} + s_j \\ a_{j+1,j} \\ \vdots \\ a_{m,j} \end{pmatrix} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ -s_j \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{13}$$

$$(\text{using Eq. 8}) \tag{14}$$

Thus successive application of the Householder matrices progressively zero out the sub-diagonal elements in each column of A thereby reducing A into upper triangular form.

The Fortran program myprog.f90 solves the least square problem using Householder-based QR decomposition (performed by the routine QRdec in LinAl.f90 module) of the corresponding Vandermond matrix A given by Eq. by casting it into the standard linear form:

$$Ax = b \tag{15}$$

where b is a column vector containing the y-values corresponding to the x vector read from the atkinson.dat file (given as an argument to the program). The vector b is replaced by the solution for x on executing the program.
Figure shows the fit to the data using with a 3^{rd} order polynomial given by

$$f(x) = 0.574658096 + 4.72586584x - 11.1282244x^2 + 7.66868114x^3 \tag{16}$$

with a rms error given by

$$E = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2} = 0.0420603491 \tag{17}$$

which is greater than the corresponding rms error in the fit obtained using the Cholesky decomposition method in Problem 1.

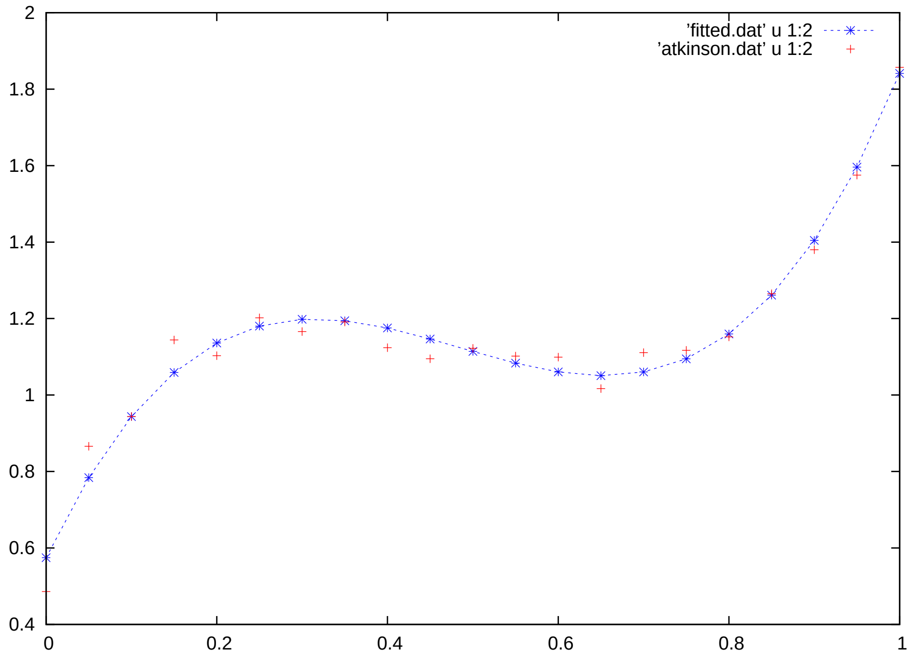

Figure 3: least square fit with 3^{rd}order polynomial (Eq. 16) to the data in atkinson.dat using Householder-based QR decomposition

Figure 4 hows the fit to the data using with a 5^{th} order polynomial given by

$$f(x) = 0.509619713 + 7.20328665x - 28.408123x^2 + 51.9441299x^3 - 47.4873962x^4 + 18.0979404x^5 \tag{18}$$

with a rms error given by

$$E = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(y_i - f(x_i))^2} = 0.0306485258 \tag{19}$$

which is smaller than the corresponding rms error in the fit obtained using the Cholesky decomposition method in Problem 1.

The algorithm gives poor fits (in term of rms error) for order of polynomial exceeding 13.

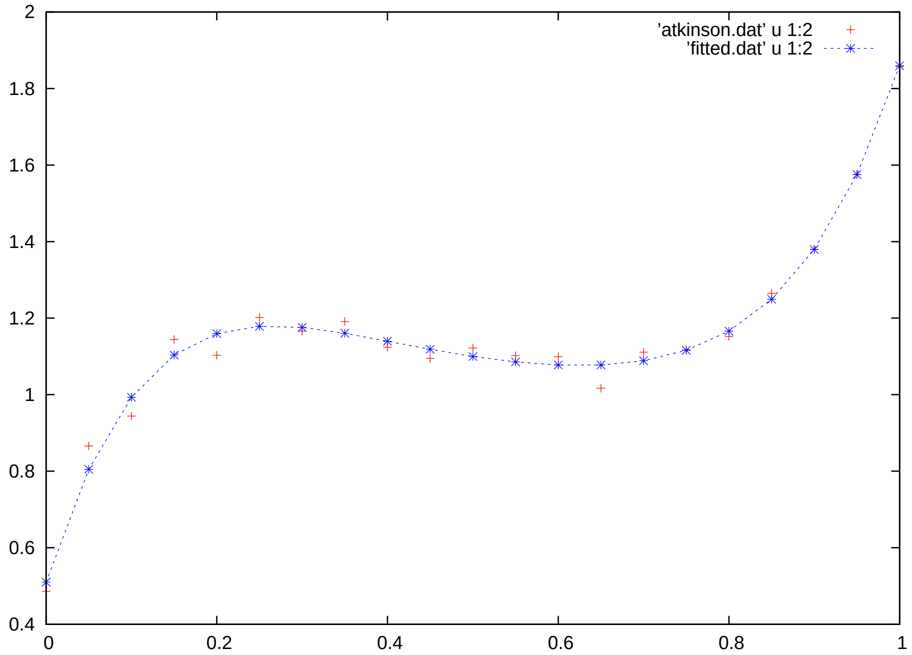

Figure 4: least square fit with 5^{th}order polynomial (Eq. 18) to the data in atkinson.dat using Householder-based QR decomposition

The Frobenius norms of $A - QR$ and $Q^T Q - I$ are given by

$$
\begin{aligned}
||A - QR||_F &= 2.05531614e - 06 \qquad (20)\\
||Q^T Q - I||_F &= 2.03146305e - 06 \qquad (21)
\end{aligned}
$$