

41
50

AM 213A HW3

Joseph Moore

Winter 2022

Part 1

1. *used to explain how/why the decomposition works. -2*

- We can show that a positive definite matrix can be written as lower triangular matrix multiplied by its conjugate transpose. We end up with the equations 2.82 and 2.83. We've seen before how $Ax = b$ can easily be solved with substitution if A is triangular. Using this fact we have

20
25

$$A^T Ax = A^T b = \tilde{A}x = \tilde{b} = LL^*x = \tilde{b}.$$

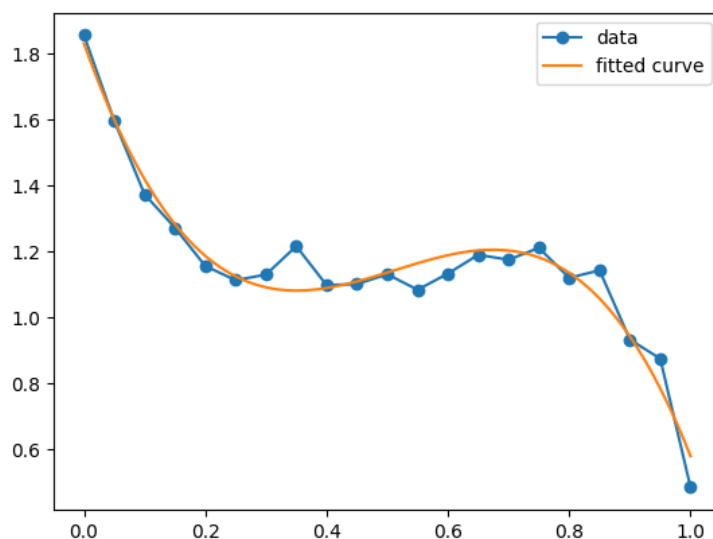
Then we solve $Ly = \tilde{b}$ for y and then $L^*x = y$ for x which is the solution to our original equation.

- The resulting equation for the third degree polynomial fit is

$$f(x) = 1.83 - 5.17x + 11.2x^2 - 7.28x^3.$$

The 2-norm of the resulting error $E = \|b - Ax\| = 6.98 \times 10^{-6}$. Below is the resulting fitted curve.

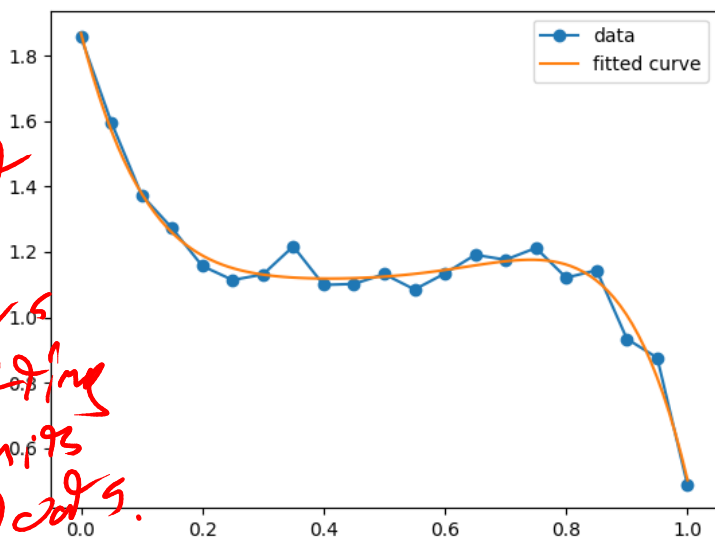
should be much larger, -1



- The resulting equation for the fifth degree polynomial fit is

$$f(x) = 1.87 - 7.25x + 28.8x^2 - 58.7x^3 + 61.0x^4 - 25.2x^5.$$

The 2-norm of the resulting error $E = \|b - Ax\| = 2.84 \times 10^{-5}$. Below is the resulting fitted curve.



On the right track, but failure occurs before hitting absolute limits of S.P. & float.

should be much larger.

Need norm of residual $\|b - Ax\|$

NOT error of

$\|A^T b - A^T A x\|$

which is likely to have values.

should discuss relative errors.

- single precision floats only store numbers from 10^{-38} up to 10^{38} . When values carry essential information up to the sixth significant digit then being raised to anything over the fifth power will result in a loss of information.

From my code I found that fitting a curve up to the sixth degree polynomial resulted in significant error.

max theoretical degree? -1

2.

Right now it is easier to look at columns of A. who does it? ... IDK do to column's and columns before?

$$H_j A = A - 2v_j v_j^T A = A - 2 \begin{pmatrix} 0 \\ 0 \\ a_{jj} + s_j \\ \vdots \\ a_{mj} \end{pmatrix} \begin{pmatrix} 0 & 0 & a_{jj} + s_j & \dots & a_{mj} \end{pmatrix} A$$

$$= A - 2 \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (a_{jj} + s_j)^2 & \dots & \dots & (a_{jj} + s_j)a_{mj} \\ 0 & 0 & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & (a_{jj} + s_j)a_{mj} & \dots & \dots & a_{mj}^2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

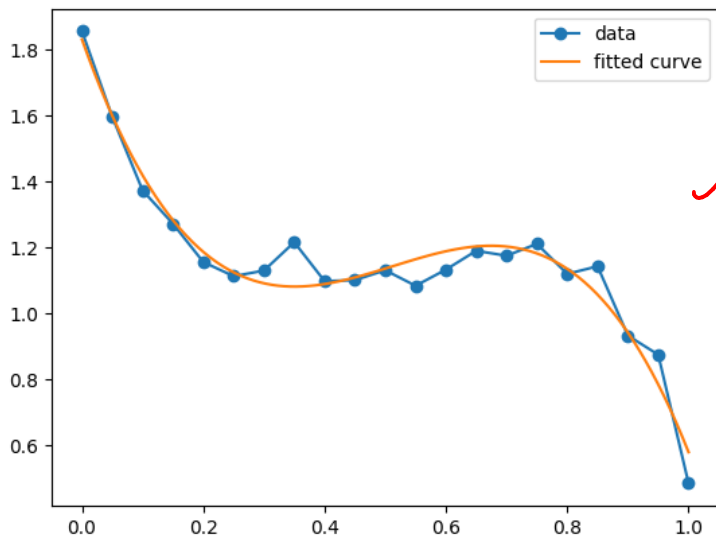
Biggest Act? How does QR work? -2

21/25

- The resulting equation for the third degree polynomial fit is

$$f(x) = 1.83 - 5.17x + 11.2x^2 - 7.28x^3.$$

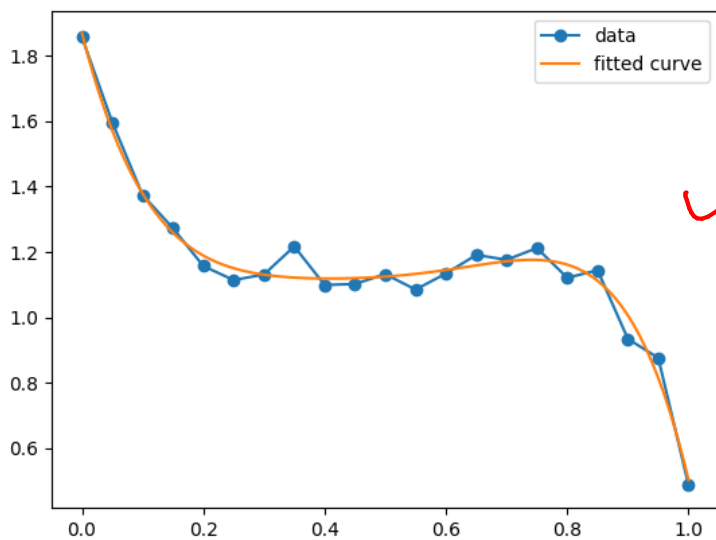
The 2-norm of the resulting error $E = \|b - Ax\| = 0.244575$. Below is the resulting fitted curve.



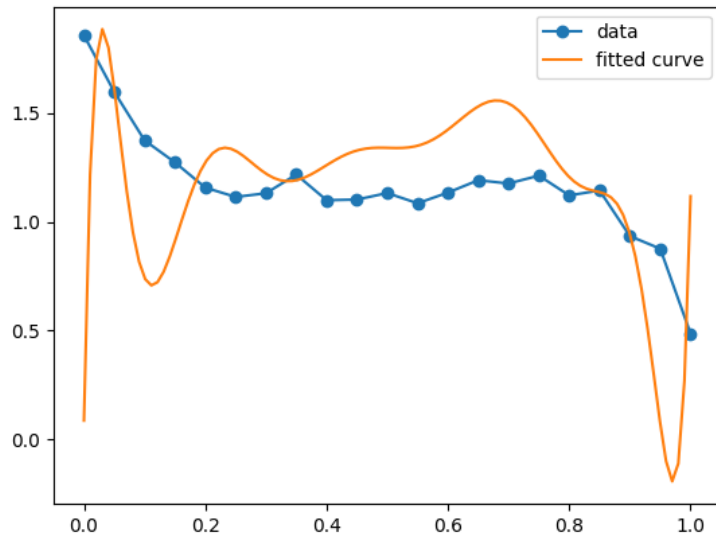
- The resulting equation for the fifth degree polynomial fit is

$$f(x) = 1.87 - 7.26x + 28.8x^2 - 58.8x^3 + 61.1x^4 - 25.2x^5.$$

The 2-norm of the resulting error $E = \|b - Ax\| = 2.84 \cdot 10^{-5}$. Below is the resulting fitted curve.



- The algorithm begins to break down around the 12th polynomial but is fine up until that point. ✓



✓
Plot!

should also be
small...

- We have $\|A - QR\| = 2.166 \cdot 10^{-06}$ and $\|Q^T * Q - I\| = 2.658$. This means that the accuracy of our decomposition is high where as the orthogonality of Q is low. ✓

Part 2

1. We must show that $(I-2P)^*(I-2P) = (I-2P)(I-2P)^* = (I-2P)(I-2P)^{-1} = I$. We know that $P^* = P^2 = P$ and can then show that

$$(I-2P)^*(I-2P) = (I^*-2P^*)(I-2P) = (I-2P)(I-2P) = I-2P-2P+4P^2 = I-4P+4P = I.$$

2.

a) Using the Cauchy-Schwarz inequality and the fact that $P = P^2$ we can say

$$\|P\|_2 = \|P^2\|_2 \leq \|P\|_2^2.$$

Divide everything by $\|P\|_2$ and you get $1 \leq \|P\|_2$. P being orthogonal means that

$$\|Py\| = \|y\| \leq \|P\| \cdot \|y\|$$

Divide everything by $\|P\|$ and we have

$$\frac{\|Py\|}{\|P\|} = \frac{\|y\|}{\|P\|} = 1 \leq \|y\|$$

3.

a) We can write our matrix as $A = QR$ where A and Q can be written as columns $A = (a_1, a_2, a_3, \dots)$ and $Q = (q_1, q_2, q_3, \dots)$ and R as standard matrix elements r_{ij} . Then

$$A = (a_1, \dots, a_n) = \begin{pmatrix} q_1 & \dots & q_n \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & r_{nn} \end{pmatrix}$$

so that we can write

$$a_j = \sum_{i=1}^j r_{ij}q_i = \sum_{i=1}^{j-1} r_{ij}q_i + r_{jj}q_j$$

and we see that a_j is in the span of the columns of Q up to q_j . Likewise, the columns of A up to a_{j-1} are all in the span of Q up to q_{j-1} . Remembering that the columns of Q are all linearly independent we can say a_j can only be linearly independent from the previous columns if it has a nonzero contribution from q_j . In other words, a_j can only be linearly independent from the previous columns if r_{jj} is non-zero and if r_{jj} is non-zero then a_j must be linearly independent from the previous columns.

b) Because r_{jj} being non-zero means that a_j is linearly independent from all previous columns then the number of linearly independent columns must be equal to the number of non-zero diagonal elements.

4. The projection of a vector x onto a vector v is $(\frac{v^T x}{v^T v})v$ and thus the projection onto the plane orthogonal to v is simply x minus this component or $x - (\frac{v^T x}{v^T v})v$. If we want to reflect across this plane we simply subtract the parallel component twice and get $x - 2(\frac{v^T x}{v^T v})v$ which is the Householder transformation applied to x . This geometric description leads us to see that the transformation hold all dimensions constant except the one parallel to v and that v must be an eigenvector with a corresponding eigenvalue of -1. All other eigenvalues must be 1.

If $x \perp v$ then $Hx = x - 2\frac{v^T x}{v^T v}v = x - 0 = x$. Thus x is a eigenvector with corresponding eigenvalue 1. Where as if $x \parallel v$ then $Hx = x - 2\frac{\|v\|^2 x}{\|v\|^2} = x - 2x = -x$ and thus x is a eigenvector with corresponding eigenvalue of -1. Because H is symmetric we know that the singular values are $|\lambda_i|$ then they must all be 1. Because only one eigenvalue will be -1, the eigenvalue corresponding to the eigenvector orthogonal to the plane of reflection, and all the other eigenvalues will be 1, then the determinant, being the product of all the eigenvalues, will be -1.

5.

$$\kappa(A^T A) = \frac{\max_{\|x\|=1} \|A^T A x\|}{\min_{\|x\|=1} \|A^T A x\|}$$

We can make a geometric argument that the maximum Mx is when x is an eigenvector. We take the fact that $Mx = \lambda x$ and $M(Mx) = \lambda Mx = \lambda^2 x$ and take $\|A^T A x\|$ to show

$$\|A^T A x\| = (A^T A x)^T A^T A x = x^T A^T A A^T A x = x^T (A^T A)^2 x = x^T M^2 x = \lambda^2 x^T x$$

and also

$$\|Ax\| = (Ax)^T Ax = x^T A^T A x = x^T Mx = \lambda x^T x.$$

Thus these matrices will have the same maximum x corresponding to the largest eigenvalue. Similar argument can be used to show that the minimum x will correspond to the smallest eigenvalue.

It then follows that

$$\kappa(A^T A) = \frac{\lambda_{max}^2 x^T x}{\lambda_{min}^2 x^T x} = \left(\frac{\lambda_{max}}{\lambda_{min}} \right)^2$$

and

$$\kappa(A) = \frac{\lambda_{max} x^T x}{\lambda_{min} x^T x} = \frac{\lambda_{max}}{\lambda_{min}}$$

and thus

$$\kappa(A^T A) = (\kappa(A))^2.$$