Assignment - 4

Ashish Raut
S. R. No - 14431

1. $f(t) = \sin(10t)$, $t \in [0, 1]$

Code is attached at the end.

Observation

① plot obtained by modified gram schmidt QR decomposition and Householder triangulization produced best result to the original curve.

② plot based on SVD is pretty skewed and looks like damped oscillation curve. Also python numpy default least squares method is also based on SVD. In this case also, we obtain similar curve. Hence the result was bad.

③ Further, there's no ~~way~~ default function for normal equation method in python. When I tried to solve it using Cholskey decomposition, computationally it couldn't calculate that and showed error that rank $(A^TA) = 12$ although rank $(A) = 15$ (i.e. Full rank). I'm not sure why it showed sth like this. One reason may be b/c $t \in [0, 1]$ in A^tA, there's lots of entries with very less value which results in

SVD in significantly low values in Σ. Hence python adjudges A^TA as rank deficient and could not calculate Cholskey decomposition.
However, if you directly calculate $(A^TA)^{-1}$, then it gives a result which looks like that obtained in SVD.
Also the bad performance of SVD method might be due to this (though I'm not sure).

On the other hand, modified Gram-Schmidt and Householder method performed as usual and gave expected result.

Plots and link to code are given at the end.

2. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \\ 2 & 8 & 2 \end{bmatrix}$ $\underset{\sim}{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

Least squares soln : $\underset{\sim}{x}^* = \begin{bmatrix} -1.33 & 0.67 & 0.67 \end{bmatrix}$

$B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \\ 5 & 7 & 9 \end{bmatrix}$, $\underset{\sim}{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$

Least squares soln : $\underset{\sim}{x}^* = \begin{bmatrix} -1 & 10^{-16} & 1 \end{bmatrix}$

We obtained this soln. using numpy default / SVD based method.

We also observe least squares problem always has a soln. Here the idea is just to minimize $\|A\underset{\sim}{x} - \underset{\sim}{b}\|^2$.

We know that, an orthogonal projector to range(A) can be expressed as

$$P = A(A^T A)^{-1} A^T$$

where A is full rank.

In this case $P\underset{\sim}{b} \in$ Range(A) and solution will be unique.

If A is not full rank, even then we can have a projection onto range(A) let's say $\underset{\sim}{c}$.

In that case, $A\underset{\sim}{x}^* = \underset{\sim}{c}$

So there always exists $\underset{\sim}{x}^*$ b/c there's a projection onto range(A).

And soln. is unique when A is full rank.

3. Let, A be a general square matrix.

let, $A \in \mathbb{R}^{m \times m}$

So

$$L_{m-1} P_{m-1} \cdots \cdots L_2 P_2 L_1 P_1 A = U$$

— (Gauss elimination with partial pivoting)

L_i's are lower triangular matrices with 1 at diagonal.

So $\det(L_i) = 1$.

P_i's are identity matrix with row interchange.

So $\det(P_i) = \begin{cases} 1 & \text{, if no row interchange} \\ -1 & \text{, if row interchange} \end{cases}$

So $\prod_{i=1}^{m-1} \det(P_i) \det(L_i)$ ~~det A~~ $\det A = \det U$

$\Rightarrow (-1)^k \det A = \det U$

where k = number of time row interchange has occured.

$\Rightarrow \det(A) = (-1)^{-k} \det(U)$.

4. Code is attached at the end and plots were presented. From the plot, it's clear than $\|LU - A\|_F$ is much higher in case of without pivoting than with partial pivot.

As studied, LU factorization without pivoting is stable but not backward stable. Hence $\|LU - A\|_F$ is pretty high in this case.

However, pivoting brings more numerical stability and $\|LU - A\|_F$ is about three orders of magnitude lower. Actually with partial pivoting, it's backward stable. and in practice it also works well.

---

③ Time complexity Part

$\det(A) = (-1)^{-k} \det(U)$

$\det U$ = product of diagonal element

= ~~...~~ $(m-1)$ operations

To check, whether row interchange occured or not in P_i, we need to check $(m-1)$ diagonal elements.

We have to do this for $(m-1)$ P_i's.

So this checking takes apprx. $(m-1)^2$ operations.

So total number of operations $= (m-1)^2 + (m-1) + O(1)$

where $O(1) \rightarrow$ other small number of products or addition.

So number floating point operation : $O(m^2)$ and dominated by checking whether row interchange has occured or not.

6. $A \in C^{m \times m}$

(a) $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$

$(A - \mu I)\underset{\sim}{x} = A\underset{\sim}{x} - \mu \underset{\sim}{x}$

$= (\lambda - \mu)\underset{\sim}{x}$

So $(\lambda - \mu)$ is eigen value of $A - \mu I$.

(b) False.

e.g. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

$\det(A - \lambda I) = 0 \quad \Rightarrow \lambda = 1$

-1 is not there in eigen value.

(c) $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$

$\Rightarrow \overline{A\underset{\sim}{x}} = \overline{\lambda \underset{\sim}{x}}$

$\Rightarrow \bar{A}\,\bar{\underset{\sim}{x}} = \bar{\lambda}\,\bar{\underset{\sim}{x}}$

B/c A is real, $\bar{A} = A$

$\Rightarrow A\bar{\underset{\sim}{x}} = \lambda^* \bar{\underset{\sim}{x}}$

Hence λ^* is also eigen value of A.

(d) $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$

$\Rightarrow A^{-1}A\underset{\sim}{x} = \lambda A^{-1}\underset{\sim}{x}$

$\Rightarrow A^{-1}\underset{\sim}{x} = \frac{1}{\lambda}\underset{\sim}{x}$.

So $\frac{1}{\lambda}$ is the eigenvalue of A^{-1}!

(e) False

e.g. $A = \begin{bmatrix} 0 & 2 & 3 \\ 5 & 0 & 4 \\ 6 & 7 & 0 \end{bmatrix}$

Here $\det(A - \lambda I) = 0 \Rightarrow \lambda = 0$

But $A \neq 0$

(f) A is diagonalisable.

So $A = X D X^{-1}$ where D = diagonal matrix
X = Invertible matrix

B/c all eigen values are equal, $D = \lambda I$.

So $A = \lambda X X^{-1} = \lambda I$ = diagonal matrix.

7. (a) $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$, $\underset{\sim}{x} \in \mathbb{C}^n$.

Let, x_i be the largest entry in $\underset{\sim}{x}$.
in terms of modulus $|\cdot|$.

So $A\underset{\sim}{x} = \lambda \underset{\sim}{x}$

$$\Rightarrow \sum_{j=1}^{n} a_{ij} x_j = \lambda x_i$$

$$\Rightarrow \sum_{\substack{j=1 \\ j \neq i}}^{n} a_{ij} x_j = (\lambda - a_{ii}) x_i$$

$$\Rightarrow (\lambda - a_{ii}) = \frac{\sum_{j\neq i} a_{ij} x_j}{x_i}$$

$$\Rightarrow |\lambda - a_{ii}| \leq \sum_{j\neq i} \frac{|a_{ij}|\,|x_j|}{|x_i|} \leq \sum_{j \neq i} |a_{ij}|$$

($\because |x_i| > |x_j|$ as per assumption)

So λ is prevent within $\sum_{j\neq i} |a_{ij}|$ distance from one of the diagonal point a_{ii}. Hence it's present inside disc.

(b) Suppose, $\exists\, \underset{\sim}{x} \neq 0$ such that $A\underset{\sim}{x} = 0$.

$$\Rightarrow \sum_{j=1}^{n} a_{ij} x_j = 0 \quad \text{for all } i.$$

Let, x_k be the largest entry in $\underset{\sim}{x}$ in terms of modulus.

So $$\sum_{j=1}^{n} a_{kj} x_j = 0$$

$$\Rightarrow a_{kk} x_k = - \sum_{j\neq k} a_{kj} x_j$$

$$\Rightarrow a_{kk} = - \sum_{j \neq k} \frac{a_{kj} x_j}{x_k}$$

$$|a_{kk}| \le \sum_{j \ne k} \frac{|a_{kj}|\,|x_j|}{|x_k|} \le \sum_{j \ne k} |a_{kj}|$$

$(\because |x_k| > |x_j|$ by assumption$)$

$$|a_{kk}| \le \sum_{j \ne k} |a_{kj}|$$

This violates diagonally dominant condition.

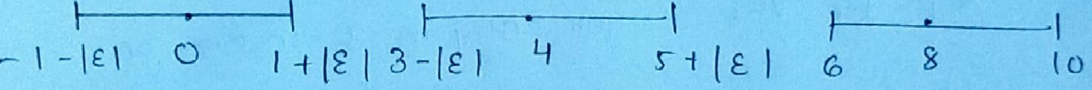
Hence our assumption of $\underset{\sim}{x} \ne 0$ and $A\underset{\sim}{x} = 0$ is wrong.

Only $\underset{\sim}{x} = 0$ can lead to $A\underset{\sim}{x} = 0$

Hence A is invertible.

(c) $A = \begin{bmatrix} 8 & 2 & 0 \\ 1 & 4 & \varepsilon \\ 0 & \varepsilon & 1 \end{bmatrix}$ $\quad |\varepsilon| < 1$

Greshgorin's Disks



$|\varepsilon| < 1$.

So $\lambda \in (-2, 2) \cup (2, 6) \cup (6, 10)$

But λ can't exactly take values like $2, 6, 10$.

5. (a) A is a non-singular square matrix $\in R^{m \times m}$

$A = QR \quad , \quad A^T = R^T Q^T$

$A^T A = R^T Q^T Q R = R^T R \quad (\because Q^T Q = I)$
$(\because \text{orthogonal})$

$A^T A = U^T U \quad - \text{(given)}$

Also
$\underset{\sim}{x}^T A^T A \underset{\sim}{x} \geq 0$
$\Uparrow$
$\Rightarrow \quad \|A \underset{\sim}{x}\|^2 \geq 0$

So $\quad A^T A$ is pos. semidefinite.

Now $\quad A^T A \underset{\sim}{x} = 0 \quad \Rightarrow \underset{\sim}{x} = 0 \quad$ b/c A is nonsingular.

Hence $A^T A$ has inverse.

So $\|A \underset{\sim}{x}\|^2 > 0$ for $\underset{\sim}{x} \neq 0$

So $\quad A^T A$ is positive definite.

Also $\quad (A^T A)^T = A^T A$ = symmetric matrix.

We know for real symmetric positive definite matrix, there's unique Cholskey factorization.

Hence $\quad R^T R = U^T U$

$\Rightarrow \quad R = U$

(c) Case - 1

$A = B^T B$

$\underset{\sim}{x}^T A \underset{\sim}{x} = \underset{\sim}{x}^T B^T B \underset{\sim}{x} = \|B \underset{\sim}{x}\|^2 \geq 0$

$\|B\underset{\sim}{x}\| = 0$ only if $\underset{\sim}{x} = 0$ ($\because$ B is full rank)

Hence for all $\underset{\sim}{x} \neq 0$, $\underset{\sim}{x}^T A \underset{\sim}{x} > 0$

So A is symmetric positive definite.

Case - 2

$A \in R^{n\times n}$ is symmetric positive definite.

So $A = Q^T D Q$ where Q = orthogonal matrix
D = Diagonal matrix

B/c A is pos. definite, all eigen values are positive.
So $D^{1/2}$ is diagonal with all +ve elements.

So $A = Q^T D^{1/2^T} D^{1/2} Q = (D^{1/2}Q)^T (D^{1/2}Q)$

$= B^T B$

Under some conditions, some part of these square
(rank limited)
B will be zero and they can be $m \times n$ matrix.

~~From this for~~ This choice of B is not unique b/c
we know for sure there's another decomposition
(Cholskey) : $A = R^T R$. $R \in$ upper triangular.

(b) $M_{ij} = \int_{-1}^{1} \phi_i(x)\,\phi_j(x)\,dx$ (inner product)

We observe, $M_{ii} = \int_{-1}^{1} \phi_i(x)\,\phi_i(x)\,dx$

$= \langle \phi_i(x), \phi_i(x) \rangle > 0$

$(\because \phi_i(x) \neq 0)$

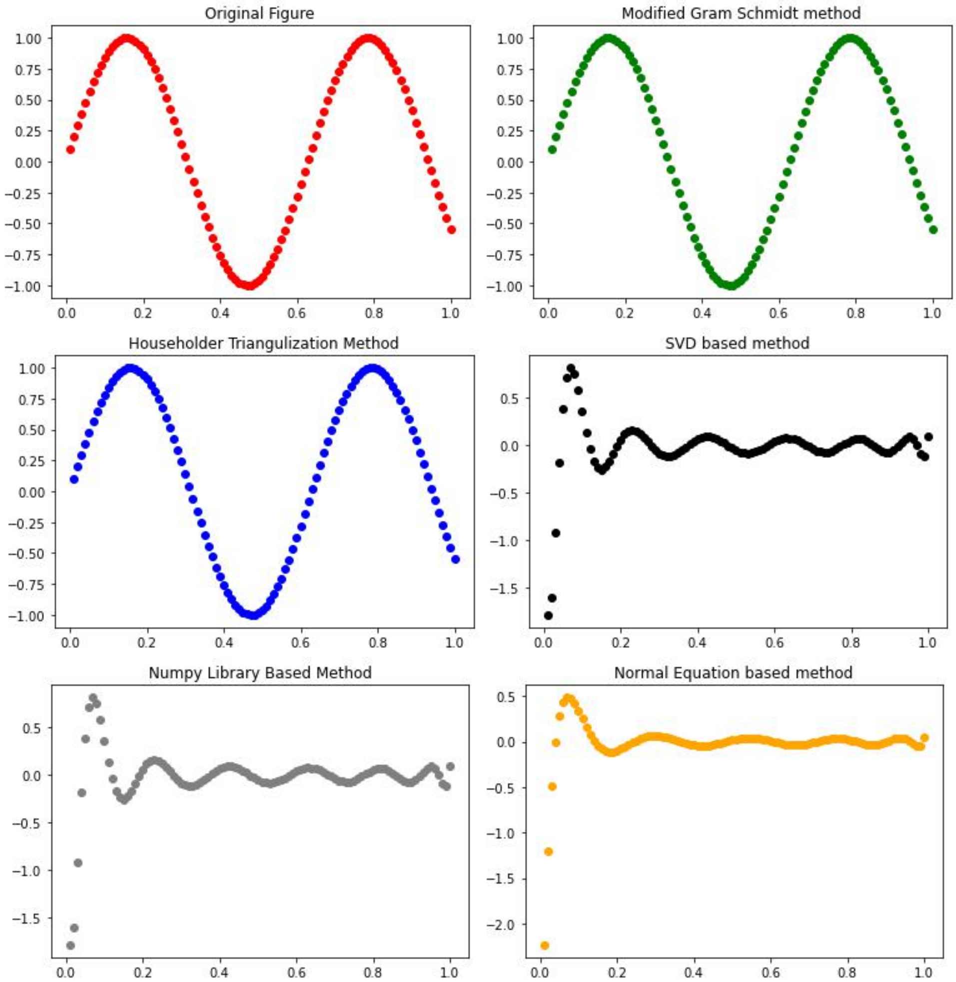
$M_{ij} = M_{ji}$ b/c ϕ_j and ϕ_i are interchangable here.

Sylvester's criterion for positive definiteness : All the leading principal minors should be positive.

First principal minor $= \langle \phi_1(x), \phi_1(x) \rangle > 0$

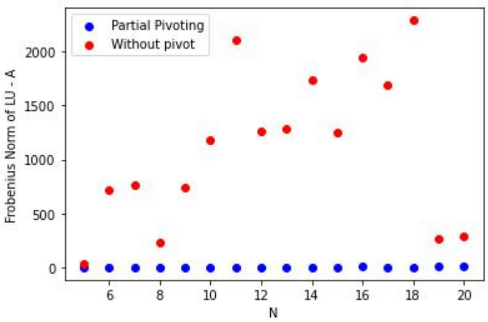
Second principal minor $= \langle \phi_1(x), \phi_1(x) \rangle \langle \phi_2(x), \phi_2(x) \rangle - \langle \phi_1(x), \phi_2(x) \rangle^2$

Q. 1.



Here I used own code in Modified Gram-Schmidt method and Householder triangularization . For others, we used library functions as asked in the question. Discrepancy in those things might be related to small values in matrix A. (More on it in theory part)

4.



Clearly we observe the benefit of backward stability in LU Factorization with pivoting.

Code for 1 and 4 in the link.