CX 4640 Homework 2

Wenqi He

September 25, 2017

2.7

(a)

$$\begin{aligned}\det A &= \det \begin{bmatrix} 1 & 1+\epsilon \\ 1-\epsilon & 1 \end{bmatrix} \\ &= 1-(1-\epsilon)(1+\epsilon) \\ &= 1-(1-\epsilon^2) \\ &= \epsilon^2\end{aligned}$$

(b)

The smallest non-negative number representable in a normalized single-precision system is 1×2^{-126}. Therefore the computed result for determinant would be zero if

$$\epsilon^2 < 2^{-126}$$
$$|\epsilon| < 2^{-63}$$
$$-2^{-63} < \epsilon < 2^{63}$$

In a double-precision system,

$$\epsilon^2 < 2^{-1022}$$
$$-2^{-511} < \epsilon < 2^{511}$$

(c)

$$A = \begin{bmatrix} 1 & 1+\epsilon \\ 1-\epsilon & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1-\epsilon & 1 \end{bmatrix} \begin{bmatrix} 1 & 1+\epsilon \\ 0 & \epsilon^2 \end{bmatrix}$$

(d)

$$\det U = 1 \cdot \epsilon^2 = \epsilon^2$$

The matrix is singular when the computed value of ϵ^2 equals zero, that is, ϵ^2 is smaller than the smallest representable number. So the answer is the same as (b).

2.21

$$\begin{aligned}
\boldsymbol{x} &= B^{-1}(2A+I)(C^{-1}+A)\boldsymbol{b} \\
B\boldsymbol{x} &= (2A+I)(C^{-1}+A)\boldsymbol{b} \\
&= 2AC^{-1}\boldsymbol{b} + 2A^2\boldsymbol{b} + C^{-1}\boldsymbol{b} + A\boldsymbol{b}
\end{aligned}$$

Let $\boldsymbol{y} = C^{-1}\boldsymbol{b}$, then

$$C\boldsymbol{y} = \boldsymbol{b}$$

$$B\boldsymbol{x} = 2A\boldsymbol{y} + 2A^2\boldsymbol{b} + \boldsymbol{y} + A\boldsymbol{b}$$

Using Gaussian Elimination, one can solve the first equation for $\boldsymbol{y}$, and then solve the second one for $\boldsymbol{x}$ without computing the inverses of B and C.

The MATLAB code is included in `SolveForX.m`

2.26

(a)

One can compute the inverse of A using Sherman-Morrison formula:

$$\begin{aligned}
A^{-1} &= (I - uv^T)^{-1} \\
&= I^{-1} + I^{-1}u(1 - v^T I^{-1}u)^{-1}v^T I^{-1} \\
&= I + u(1 - v^T u)^{-1}v^T,
\end{aligned}$$

provided that $1 - v^T u$ is invertible, that is,

$$v^T u \neq 1$$

(b)

From (a), $\sigma = -(1 - v^T u)^{-1}$

(c)

Yes.

$$M_k = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & -m_{k+1} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & -m_n & 0 & \cdots & 1 \end{bmatrix} = I - \begin{bmatrix} 0 \\ \vdots \\ m_{k+1} \\ \vdots \\ m_n \end{bmatrix} \boldsymbol{e}_k{}^T,$$

$$u = \begin{bmatrix} 0 \\ \vdots \\ m_{k+1} \\ \vdots \\ m_n \end{bmatrix}, \quad v = \boldsymbol{e}_k, \quad \sigma = -\left(1 - \boldsymbol{e}_k{}^T \begin{bmatrix} 0 \\ \vdots \\ m_{k+1} \\ \vdots \\ m_n \end{bmatrix}\right)^{-1} = -(1-0)^{-1} = -1$$

2.4

$\|A_1^{-1}\| = 0.7097, \quad \mathrm{cond}(A_1) = 12.7742, \quad \|A_2^{-1}\| = 1.6393e{+}04, \quad \mathrm{cond}(A_2) = 4.0163e{+}06$

Using the first approach, the estimations are:

$\|A_1^{-1}\| = 0.6226, \quad \mathrm{cond}(A_1) = 11.2071, \quad \|A_2^{-1}\| = 1.3082e{+}04, \quad \mathrm{cond}(A_2) = 3.2051e{+}06$

Using the second approach (The values might vary):

$\|A_1^{-1}\| = 0.6013, \quad \mathrm{cond}(A_1) = 11.1528, \quad \|A_2^{-1}\| = 4.4131e{+}03, \quad \mathrm{cond}(A_2) = 1.9666e{+}06$

Apparently, the first approach is more accurate.

5

(Run `CreateTable` to reproduce the result.)

n	relative error	condition number
2	2.8951e-16	19.281
3	3.4571e-15	524.06
4	9.1928e-14	15514
5	8.4248e-14	4.7661e+05
6	1.0052e-10	1.4951e+07
7	3.3589e-09	4.7537e+08
8	1.1118e-08	1.5258e+10
9	3.2899e-06	4.9315e+11
10	0.00010306	1.6025e+13
11	0.0024921	5.2202e+14
12	0.068797	1.6212e+16