Math 472 Midterm Exam, March 14, 2019

(1) (10 pts.) Suppose the $n \times n$ matrix A is invertible. If n is large, what is the most efficient method you know of for computing its determinant? Justify your answer in terms of work estimates. Your work estimate should be of the form $w = cn^p$. Give the values of c and p. How does this compare to using cofactor expansion?

Gaussian Elimination is the most efficient means of calculating the determinant. The work estimate of $PA = LU$ is $\frac{2}{3}n^3 + O(n^2)$. $\det(P)\det(A) = \det(L)\det(U) = 1 \cdot u_{11} \cdots u_{nn}$

Total is still $\frac{2}{3}n^3 + O(n^2)$ vs. $n!$ for cofactor expansion

(2) (10 pts.) State a condition that guarantees that a square matrix A will have a factorization of the form $A = LU$ with $l_{ii} = 1$ and $u_{ii} \neq 0$.

If all the leading principal minors are nonzero then A will have a factorization of the form above.

(3) (10 pts.) Suppose A is symmetric, positive definite. Is the factorization $A = LL^T$ unique?

NO. It is unique only if we assume $L_{ii} > 0$.

Counterexample: Suppose A is s.p.d. and $A = LL^T$.
Then let $D = \{\pm 1, \pm, \ldots, \pm 1\} \Rightarrow A = L \underbrace{DD^T}_{I} L^T = (LD)(LD)^T$
which gives a whole family of other factorizations.

(4) (10 pts.) Let Q be a real orthogonal matrix. Show that $\sum_{i,j=1}^{n} q_{ij}^2 = n$.

$$\sum_{\substack{i=1\\j=1}}^{n} q_{ij}^2 = \sum_{j=1}^{n} \sum_{i=1}^{n} q_{ij}^2.$$

Since the columns of Q are orthonormal, $\sum_{i=1}^{n} q_{ij}^2 = 1$, $i = 1, \ldots, n$

$$\Rightarrow \sum_{\substack{i=1\\j=1}}^{n} q_{ij}^2 = \sum_{j=1}^{n} 1 = n.$$

(5) (25 pts.) Compute the Choleski factorization of the matrix

$$A = \begin{bmatrix} 4 & 2 & -2 \\ 2 & 10 & 5 \\ -2 & 5 & 21 \end{bmatrix} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

Use $A = LL^T$ to show that A is positive definite.

$l_{11}^2 = 4 \Rightarrow \boxed{l_{11} = 2}$. $l_{21} l_{11} = 2 \Rightarrow l_{21} = 2/2 = 1$, $l_{31} l_{11} = -2 \Rightarrow l_{31} = -1$

$l_{21}^2 + l_{22}^2 = a_{22} = 10 \Rightarrow l_{22} = \sqrt{10-1} = 3$

$l_{31} l_{21} + l_{32} l_{22} = a_{32} = 5 \Rightarrow l_{32} = \frac{5-(-1)(1)}{3} = 2$

$l_{31}^2 + l_{32}^2 + l_{33}^2 = a_{33} = 21 \Rightarrow l_{33} = \sqrt{21 - (-1)^2 - (2)^2} = 4 = a_{33}$

$$L = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ -1 & 2 & 4 \end{bmatrix}$$

$x^T A x = x^T L L^T x = \|L^T x\|_2^2 \geq 0$.

Also, if $x^T A x = 0$, then $\|L^T x\|_2 = 0 \Rightarrow L^T x = 0 \Rightarrow x = 0$ since L is invertible.

(6) (15 pts.) Given $\mathbf{v} = (1, -1, 3)^T$ compute the Householder matrix $H(\mathbf{v})$ and $H(\mathbf{v})\mathbf{b}$ with $\mathbf{b} = (2, 3, 1)^T$ in an efficient way.

$\|v\|_2^2 = 11$, $vv^T = \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} [1 \ -1 \ 3] = \begin{bmatrix} 1 & -1 & 3 \\ -1 & 1 & -3 \\ 3 & -3 & 9 \end{bmatrix}$

$\Rightarrow H(v) = I - \frac{2vv^T}{\|v\|_2^2} = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} - \frac{2}{11} \begin{bmatrix} 1 & -1 & 3 \\ -1 & 1 & -3 \\ 3 & -3 & 9 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 9 & 2 & -6 \\ 2 & 9 & 6 \\ -6 & 6 & -7 \end{bmatrix}$

$$\begin{bmatrix} .\overline{81} & .\overline{18} & -.\overline{54} \\ .\overline{18} & .\overline{81} & .\overline{54} \\ -.\overline{54} & .\overline{54} & -.\overline{63} \end{bmatrix}$$

Using matrix-vector multiplication to compute $H(v)b$ costs $2n^2$ ops. It is more efficient to use

$H(v)b = \left(I - \frac{2vv^T}{\|v\|_2^2}\right) b = b - \frac{2v^T b}{\|v\|_2^2} v$ costs $6n$ ops

$v^T b = (1,-1,3) \cdot (2,3,1) = 2$

$$\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} - \frac{2(2)}{11} \begin{bmatrix} 1 \\ -1 \\ 3 \end{bmatrix} = \frac{1}{11} \begin{bmatrix} 18 \\ 37 \\ -1 \end{bmatrix} = \begin{bmatrix} 1.\overline{63} \\ 3.\overline{36} \\ -0.\overline{09} \end{bmatrix}$$

(7) (25 pts.) A 3×2 matrix A has its Q, R factors given by

$$Q = \begin{bmatrix} -.26726 & .31315 & -.91132 \\ -.53452 & -.83507 & -.13019 \\ -.80178 & .45233 & .39057 \end{bmatrix}, \quad R = \begin{bmatrix} -3.74166 & -5.58576 \\ 0 & -.61586 \\ 0 & 0 \end{bmatrix}.$$

Use Q and R to find the Least-Squares solution $\mathbf{u}$ of the system $A\mathbf{x} = \mathbf{b}$ with $\mathbf{b} = (1, 1, 1)^T$. Also compute the error $\|A\mathbf{u} - \mathbf{b}\|_2$.

we have: $\|Ax-b\|_2^2 = \|QRx - b\|_2^2 = \|Q(Rx - Q^T b)\|_2^2 = \|Rx - Q^T b\|_2^2$.

$$Q^T b = \begin{bmatrix} -1.60356 \\ -0.06959 \\ \hline -0.65094 \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \hline \tilde{b}_2 \end{bmatrix}.$$

The least squares solution is given by

$$\begin{bmatrix} -3.74166 & -5.58576 \\ 0 & -.61586 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1.60356 \\ -0.06959 \end{bmatrix}$$

$$x_2 = \frac{-0.06959}{-.61586} = \boxed{.11300}\ ; \quad x_1 = \frac{-1.60356 + 5.58576\,(x_2)}{-3.74166} = \boxed{.25988}$$

Error is $\|\tilde{b}_2\|_2 = |-.65094| = .65094$

(8) (25 pts.) Consider the system $\begin{bmatrix} 10 & -7 & 0 \\ -3 & 2.099 & 6 \\ 5 & -1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 3.901 \\ 6 \end{bmatrix}$.

The solution is $\mathbf{x} = (0, -1, 1)^T$.

(a) Argue that A has the $A = LU$ factorization.

(b) Solve the system using Naive Gaussian elimination (no pivoting) using 5-decimal digit arithmetic with rounding.

(c) How do you explain the large error in $\mathbf{x}$?

(a) leading principal minors are $10 \neq 0$, $10(2.099) - 21 = -0.01 \neq 0$, $-150.05 \neq 0$

$\Rightarrow$ has an $A = LU$ factorization in exact arithmetic.

Also, the solution is $(0, -1, 1)$

(b)

$$\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ -3 & 2.099 & 6 & 3.901 \\ 5 & -1 & 5 & 6 \end{array}$$

$2.099 + .3(-7) = 2.099 - 2.100 = -0.001$ No roundoff

$3.901 + .3(7) = 3.901 + 2.100 = 6.001$ No roundoff

$\downarrow$ $.3 \times r_1 + r_2$, $-.5 \times r_1 + r_2$

$-1 + (-.5)(-7) = -1 + 3.5 = 2.5$ No roundoff

$6 + (-.5)(7) = 6 - 3.5 = 2.5$ No roundoff

$$\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & -0.001 & 6 & 6.001 \\ 0 & 2.5 & 5 & 2.5 \end{array}$$

$5 + 2{,}500 \times 6 = 5 + 15{,}000 = 15{,}005$ No roundoff

(*) $\downarrow$ $2{,}500 \times r_2 + r_3$

$\boxed{2.5 + 2{,}500 \times 6.001}$

first we multiply $2{,}500 \times 6.001 = 15002.5 \xrightarrow{\text{rounds}} 15{,}003$

add 2.5 to $15003 = 15005.5 \xrightarrow{\text{rounds}} 15{,}006$

$$\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & -.001 & 6 & 6.001 \\ 0 & 0 & 15{,}005 & 15{,}006 \end{array}$$

Back solution phase

$$x_3 = \frac{15{,}006}{15{,}005} = 1.00006644 \xrightarrow{\text{rounds}} \boxed{1.0001 = x_3}$$

$$x_2 = \frac{6.001 - 6x_3}{-0.001} = \frac{6.001 - 6(1.0001)}{-0.001}$$

first we multiply: $6(1.0001) = 6.0006$ No error

subtract from 6.001 $\Rightarrow$ $6.0010 - 6.0006 = 0.0004$ No error

Divide by $-0.001 \Rightarrow \dfrac{0.0004}{-0.001} = -0.4$ No error

Thus, computed value of $\boxed{x_2 \text{ is } -0.4}$. Compared to the exact value of -1, this is a huge error.

$$x_1 = \frac{7 + 7x_2}{10} = \frac{7 + 7(-0.4)}{10} = \frac{7 - 2.8}{10} = \frac{4.2}{10} = \boxed{0.42 = x_1} \text{ vs. } 0.$$

(c) We compute $K_\infty(A)$ to find 17.56. This is not large. So the large errors in the solution are not due to ill-conditioning, but rather the large multiplier $m_{32} = 2{,}500$ which magnified small roundoff errors.

Indeed, doing a row interchange in (*), we get

$$\left[\begin{array}{ccc|c} 10 & -7 & 0 & 7 \\ 0 & 2.5 & 5 & 2.5 \\ 0 & -0.001 & 6 & 6.001 \end{array}\right].$$

Doing elimination followed by back solving yields the computed solution $(0, -1, 1)$ which is exact.