UNIVERSITY OF NEW SOUTH WALES
School of Mathematics and Statistics

MATH2089 Numerical Methods and Statistics
Term 2, 2019

Numerical Methods Tutorial – Week 6 Solutions

1. Let

$$A = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

   (a) Show that A is symmetric.

   (b) Show that $A = LDL^T$.

   (c) Find the Cholesky factorization $A = R^T R$, where R is upper triangular.

   (d) Show that A is positive definite.

   (e) Using the Cholesky factorization, find the inverse of A.

   (f) Compute the condition number $\kappa_\infty(A)$.

   (g) Estimate the relative error in the computed solution to $A\mathbf{x} = \mathbf{b}$ if $\mathbf{b}$ is known to 4 significant figures,

Answer

(a) A matrix A is symmetric if and only if $A^T = A$, where $(A^T)_{ij} = A_{ji}$. Here

$$A^T = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} = A,$$

so A is symmetric. In MATLAB check `norm(A-A')`.

(b)

$$\begin{aligned} LDL^T &= \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 & 2 \\ -1 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} = A. \end{aligned}$$

Note that pre-multiplying a matrix B by a diagonal matrix D to get DB scales the ith row of B by D_{ii}, while post-multiplying a matrix C by a diagonal matrix D to get CD. scales the jth column of C by D_{jj}.

(c) If $A = LDL^T$ where D is diagonal with $D_{ii} > 0$ for $i = 1, \ldots, n$ and L is lower triangular, then let $D^{\frac{1}{2}} = \texttt{diag}(\sqrt{d_{11}}, \ldots, \sqrt{d_{nn}})$ and

$$A = LDL^T = LD^{\frac{1}{2}}D^{\frac{1}{2}}L^T = \left(LD^{\frac{1}{2}}\right)\left(D^{\frac{1}{2}}L^T\right) = \left(D^{\frac{1}{2}}L^T\right)^T\left(D^{\frac{1}{2}}L^T\right) = R^T R.$$

Thus

$$R = D^{\frac{1}{2}}L^T = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Check that

$$R^T R = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \sqrt{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 & 2 \\ -1 & 1 & -1 \\ 2 & -1 & 3 \end{bmatrix} = A.$$

(d) A is positive definite iff $\mathbf{x}^T A\mathbf{x} > 0$ for all $\mathbf{x} \neq \mathbf{0}$. Now, from the Cholesky factorization $A = R^T R$,

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T R^T R\mathbf{x} = (R\mathbf{x})^T (R\mathbf{x}) = \mathbf{y}^T\mathbf{y} = \sum_{i=1}^{n} y_i^2 \geq 0,$$

where $\mathbf{y} = R\mathbf{x}$. Also as $\det(A) = \det\left(R^T R\right) = \det(R)^2$, A is nonsingular $\iff$ R is nonsingular. Thus

$$\mathbf{x}^T A\mathbf{x} = 0 \Longrightarrow \mathbf{y} = \mathbf{0} \Longrightarrow R\mathbf{x} = \mathbf{0} \Longrightarrow \mathbf{x} = \mathbf{0}$$

so $\mathbf{x} \neq \mathbf{0} \Longrightarrow \mathbf{x}^T A\mathbf{x} > 0$.

(e) To find $X = A^{-1}$ solve $AX = R^T RX = I$ by first solving $R^T Y = I$ and then $RX = Y$.

$$R^T Y = I \Longrightarrow \begin{bmatrix} \sqrt{2} & 0 & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \sqrt{2} & 0 & 1 \end{bmatrix} \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} \\ Y_{21} & Y_{22} & Y_{23} \\ Y_{31} & Y_{32} & Y_{33} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Forward substitution gives

$$Y = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

Then

$$RX = Y \Longrightarrow \begin{bmatrix} \sqrt{2} & -\frac{1}{\sqrt{2}} & \sqrt{2} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} & X_{13} \\ X_{21} & X_{22} & X_{23} \\ X_{31} & X_{32} & X_{33} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & \sqrt{2} & 0 \\ -1 & 0 & 1 \end{bmatrix}.$$

Back-substitution yields

$$X = \begin{bmatrix} 2 & 1 & -1 \\ 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} = A^{-1}.$$

Again you can check that $AA^{-1} = A^{-1}A = I$. Also note that as A is symmetric, A^{-1} is also symmetric.

(f)

$$\|A\|_\infty = \max \begin{bmatrix} 5 \\ 3 \\ 6 \end{bmatrix} = 6, \quad \|A^{-1}\|_\infty = \max \begin{bmatrix} 4 \\ 3 \\ 2 \end{bmatrix} = 4, \quad \Longrightarrow \quad \kappa_\infty(A) = 24.$$

Note that as A is symmetric, $\kappa_1(A) = \kappa_\infty(A)$.
For reference $\|A\|_2 = 5.05$, $\|A^{-1}\|_2 = 3.25$ and $\kappa_2(A) = 16.4$.

(g) An estimate of the relative error in the computed solution to $A\mathbf{x} = \mathbf{b}$ is the condition number of A times the relative error in the input data (A and $\mathbf{b}$). Here the entries of A are integers, so will be represented exactly. As $\mathbf{b}$ is known to 4 significant figures, the relative error in the elements of $\mathbf{b}$ is less than 0.5×10^{-4}. Thus, using the infinity norm,

$$\mathrm{re}\,(\mathbf{x}) \le \kappa_\infty(A)\,(\mathrm{re}\,(A) + \mathrm{re}\,(\mathbf{b})) = 24\left(0 + 0.5 \times 10^{-4}\right) = 1.2 \times 10^{-3} = 0.12 \times 10^{-2}.$$

Thus we can only guarantee two significant figures in the computed solution $\mathbf{x}$, even though the condition number of A is small (on a scale of $1 = 10^0$ to $1/\epsilon \approx 10^{16}$).

2. The MATLAB script `tut06q2.m`

```
% MATH2089: File = tut06q2.m
format compact

A = [1  1  3  0; 1  3  0  0; -2  0  0  1; 3  0  0  2];
b = [0; 5; -6; 2];

[L1, U1, p1] = lu(A, 'vector')
chk = norm(A-L1*U1, 1)
x1 = U1 \ (L1 \ b(p1)); x1T = x1'

pc = colamd(A)
B = A(:,pc);
[L2, U2, p2] = lu(B, 'vector')
x2 = U2 \ (L2 \ b(p2)); x2T = x2'
```

produces the output

```
L1 =
    1.0000         0         0         0
    0.3333    1.0000         0         0
```

```
    0.3333    0.3333    1.0000         0
   -0.6667         0         0    1.0000
U1 =
    3.0000         0         0    2.0000
         0    3.0000         0   -0.6667
         0         0    3.0000   -0.4444
         0         0         0    2.3333
p1 =
     4     2     1     3
chk =
    10
x1T =
    2.0000    1.0000   -1.0000   -2.0000
pc =
     3     2     1     4
L2 =
    1.0000         0         0         0
         0    1.0000         0         0
         0         0    1.0000         0
         0         0   -0.6667    1.0000
U2 =
    3.0000    1.0000    1.0000         0
         0    3.0000    1.0000         0
         0         0    3.0000    2.0000
         0         0         0    2.3333
p2 =
     1     2     4     3
x2T =
   -1.0000    1.0000    2.0000   -2.0000
```

(a) What row operations does MATLAB do to produce zeros in the first column of A?

(b) Why is `chk` not equal to zero?

(c) Calculate the sparsity of A

(d) What is the value of `B`?

(e) Why are `x1T` and `x2T` not the same?

Answer

(a) The coefficient matrix A is

$$A = \begin{bmatrix} 1 & 1 & 3 & 0 \\ 1 & 3 & 0 & 0 \\ -2 & 0 & 0 & 1 \\ 3 & 0 & 0 & 2 \end{bmatrix}.$$

MATLAB's `lu` factorization uses partial pivoting, so the element with largest magnitude in column 1 is used as the pivot element. Thus, if R_i denotes the ith row of A, the row operations are

$$R_1 \leftrightarrow R_4$$

then

$$\begin{aligned} R_2 &\leftarrow R_2 - \frac{1}{3}R_1 \\ R_3 &\leftarrow R_3 - \frac{-2}{3}R_1 \\ R_4 &\leftarrow R_4 - \frac{1}{3}R_1 \end{aligned}$$

which produces

$$\begin{bmatrix} 3 & 0 & 0 & 2 \\ 1 & 3 & 0 & 0 \\ -2 & 0 & 0 & 1 \\ 1 & 1 & 3 & 0 \end{bmatrix}, \quad \text{then} \quad \begin{bmatrix} 3 & 0 & 0 & 2 \\ 0 & 3 & 0 & -\frac{2}{3} \\ 0 & 0 & 0 & \frac{7}{3} \\ 0 & 1 & 3 & -\frac{2}{3} \end{bmatrix},$$

and the elements in the first column of L are $L_{21} = \frac{1}{3}$, $L_{3,1} = -\frac{2}{3}$ and $L_{41} = \frac{1}{3}$. Note that partial pivoting ensures that the elements of the unit lower triangular matrix L satisfy $|L_{ij}| \leq 1$.

(b) In the script `chk` is the value of $\|A - LU\|_1$, which will only be zero if there has been no row re-ordering. To check that the calculated `L1` and `U1` are correct use

```
chk = norm(A(p1,:) - L1*U1, 1)
```

Note that `p1` is a vector specifying the re-ordering of the rows/equations rather than a permutation matrix.

(c) The sparsity of A is

$$\frac{\text{Number of non-zero elements in } A}{\text{Total number of elements in } A} \times 100 = \frac{9}{4 \times 4} \times 100 = 56.25\%.$$

Sparsity is important for large matrices, where the number of non-zeros can be obtained using MATLAB's `nnz` function or from a spy plot.

(d) The matrix `B = A(:,pc)` is obtained from re-ordering the columns of A according to the vector `pc` obtained using the `colamd` (Column adaptive minimum degree re-ordering). As `pc = [3 2 1 4]`, the first and third columns of A have been swapped, giving

$$B = \begin{bmatrix} 3 & 1 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 3 & 2 \end{bmatrix}.$$

The idea is to choose a re-ordering of the columns of A so that its factorization (L and U in this case) has fewer non-zero elements.

(e) Re-ordering the columns of A corresponds to re-ordering the variables. To get back to the original variables, you must reverse this reordering. Thus swapping the first and third elements of `x2T` gives `x1T` as expected.

3. Consider the spy plots of the 156 by 156 matrix A from the chemical plant model
`http://math.nist.gov/MatrixMarket/data/Harwell-Boeing/chemwest/west0156.html`

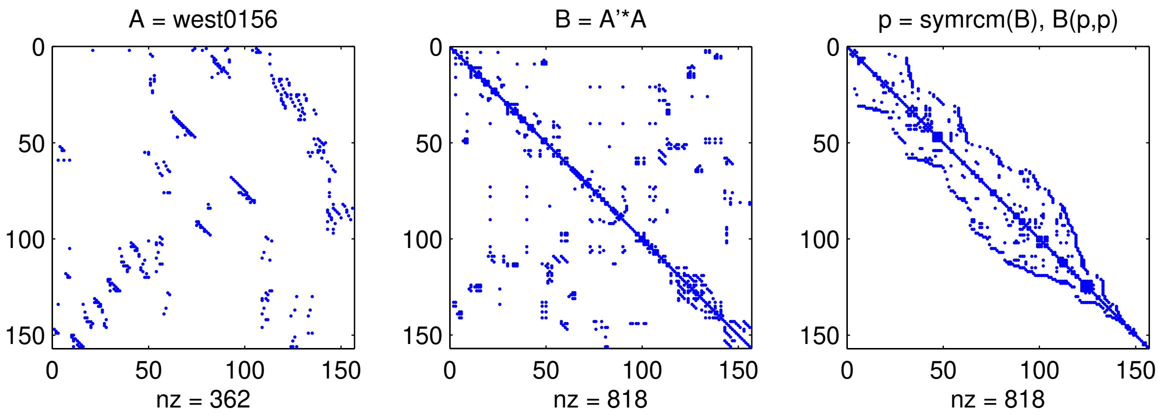

Figure 1: Spy plots of non-zero elements of matrices in a chemical plant model

(a) Are A and $B = A^T A$ symmetric?

(b) How many non-zero elements do A and B have?

(c) Calculate the sparsity of A and B.

(d) What is the sparsity of $B(p,p)$?

(e) What does $B(p,p)$ give and why is it useful?

Answer

(a) A is symmetric $\iff$ $A^T = A$ or $A_{ij} = A_{ji}$ for all $i \neq j$. A spy plot, which indicates the non-zero elements of a matrix, does not tell you the actual values, so you cannot decide if A is symmetric. However you may be able to determine that A is **not** symmetric if there exist i, j with $A_{ij} \neq 0$, but $A_{ji} = 0$. For the `west0156` matrix this is the case as the pattern of non-zero elements in the bottom left hand corner (around $i = 150, j = 10$) is clearly different from that in the top right hand corner. Thus A is not symmetric.
Any matrix of the form $B = A^T A$ is symmetric as

$$B^T = (A^T A)^T = A^T (A^T)^T = A^T A = B.$$

This does not depend on the structure of the matrix A. This is reflected in the symmetric structure of the spy plot in the middle figure.

(b) The number of non-zero elements is given below each spy plot. Thus A has 362 non-zero elements while B has 818 non-zero elements. Note that the process of forming the matrix product $B = A^T A$ has caused **fill-in**, that is non-zero elements have been created where A originally had zero elements.

(c) As A is a 156 by 156 matrix, so is B. Thus

$$\begin{aligned}
\text{Sparsity of } A &= \frac{362}{156 \times 156} \times 100 \approx 1.5\% \\
\text{Sparsity of } B &= \frac{818}{156 \times 156} \times 100 \approx 3.4\%
\end{aligned}$$

(d) The sparsity of $B(p,p)$ is exactly the same as the sparsity of B, as $B(p,p)$ is just a re-ordering of the rows and columns of B, so none of the values of the elements changes, just their positions.

(e) The aim of the reordering is to produce a matrix for which there is less fill-in when the matrix is factored. As B is symmetric, the same re-ordering must be applied to both the rows and the columns of B to keep $B(p,p)$ symmetric. Here the reverse Cuthill-McKee (MATLAB `symrcm`) reordering which tries to move the non-zero elements closer to the main diagonal has been used. This is because fill-in can then only occur within the band of non-zeros around the main diagonal.