Questions for the course

Numerical Linear Algebra

TMA265/MMA600

Date: 2013 October 24, Time: 14.00 - 18.00, Place: at CTH, Maskinhuset

Question 1

- 1. Find eigenvalues for a matrix A such that

$$\mathbf{A} = \begin{bmatrix} 4 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

  (1p)
- 2. Compute $||A||_\infty, ||A||_1$. Find conjugate transpose matrix A^* to the matrix A. **(1p)**
- 3. Find inverse matrix A^{-1} to the matrix A via the matrix of cofactors. **(2p)**

Question 2

- 1. Write algorithm of LU factorization of the matrix A with pivoting using conventional programming language notation. **(1p)**
- 2. Prove that following two statements are equivalent:
  1. There exists a unique unit lower triangular matrix L and nonsingular upper triangular U such that $A = LU$.
  2. All leading principal submatrices of A are nonsingular.

  (3p)

Question 3

- 1. Derive the condition number $k(A)$ of the matrix A. **(2p)**
- 2. Derive practical error bounds

$$error = \frac{||\widetilde{x} - x||_\infty}{||\widetilde{x}||_\infty}$$

  of $Ax = b$ in the terms of residual r and the approximate solution $\widetilde{x}$ of $Ax = b$.
  (2p)

Question 4

- 1. Derive normal equations $A^T Ax = A^T b$ in the method of normal equations. **(2p)**
- 2. Derive formula for x that minimizes the functional $F(x) = ||Ax - b||_2^2$ using the QR decomposition of the matrix $A = QR$. You can present any one of the three derivations of this formula. **(3p)**
- 3. Let A will be $m \times n$ matrix with $m \geq n$. Define the SVD decomposition of a matrix A. **(1p)**

Question 5

- Compute QR decomposition of the matrix A using Householder reflections:

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 0 & 3 & 1 \\ 3 & 4 & 7 \end{bmatrix}$$

  (2p)

- Compute QR decomposition of the matrix A using Given's rotations

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 3 & 3 & 1 \\ 0 & 4 & 7 \end{bmatrix}$$

  (1p)

Question 6

- 1. Give definitions of the Schur canonical form and the Real Schur canonical form.
  (1p)
- 2. Let the matrix A is diagonalizible such that $S^{-1}AS = \Lambda$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ are eigenvalues. Prove that the $S = [x_1, \ldots, x_n]$ be the nonsingular matrix of right eigenvectors, and rows of S^{-1} are conjugate transposes of the left eigenvectors y_i.
  (1p)
- 3. Let λ be a simple eigenvalue of A with right eigenvector x and left eigenvector y, normalized so that $||x||_2 = ||y||_2 = 1$. Define condition number for the eigenvalue λ.
  (1p)

Question 7

- 1. Let $A = U\Sigma V^T$ be the SVD decomposition of the m-by-n matrix A with $m \geq n$. Define the Moore-Penrose pseudoinverse matrix A^+ of the matrix A.
  (1p)
- 2. Let $A = U\Sigma V^T$ be the SVD decomposition of the m-by-n matrix A with $m \geq n$. Using definitions of A and A^+ prove that $AA^+A = A$.
  (2p)

Numerical Linear Algebra

TMA265/MMA600
Solutions to the examination at 24 October 2013

Question 1

1. We should solve characteristic equation $\det(A - \lambda I) = 0$:

$$\det \begin{bmatrix} 4-\lambda & 1 & 0 \\ 2 & -\lambda & 0 \\ 0 & 0 & 5-\lambda \end{bmatrix} = 0.$$

Solving above equation for λ we get three eigenvalues $\lambda_1 = 5, \lambda_2 = \frac{4+\sqrt{24}}{2} \approx 4.4495, \lambda_3 = \frac{4-\sqrt{24}}{2} \approx -0.4495$ which are solutions to the equation $(5-\lambda)(\lambda^2 - 4\lambda - 2) = 0$.
3. We use definition of A^*:

$$A^* = \overline{A^T}.$$

$$A^T = \begin{bmatrix} 4 & 2 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 5 \end{bmatrix}.$$

$||A||_\infty = 5, ||A||_1 = 6.$
4. By definition of an inverse matrix we have:

$$A^{-1} = \frac{1}{\det A}(C^T)_{ij}$$

Thus,

$$C = \begin{bmatrix} 0 & -10 & 0 \\ -5 & 20 & 0 \\ 0 & 0 & -2 \end{bmatrix},$$

$$C^T = \begin{bmatrix} 0 & -5 & 0 \\ -10 & 20 & 0 \\ 0 & 0 & -2 \end{bmatrix},$$

and thus $A^{-1} = \frac{1}{-10} \cdot \begin{bmatrix} 0 & -5 & 0 \\ -10 & 20 & 0 \\ 0 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 0 & 0.5 & 0 \\ 1 & -2 & 0 \\ 0 & 0 & 0.2. \end{bmatrix}$

Question 2

1. See Lecture 3 and the course book.
1. See Theorem 2.4 of the course book.
Proof.
We first show that (1) implies (2). $A = LU$ may also be written

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \times \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12} + L_{22}U_{22} \end{bmatrix}$$

where A_{11} is a j-by-j leading principal submatrix, as well as L_{11} and U_{11}. Therefore $det A_{11} = det(L_{11}U_{11}) = det L_{11} det U_{11} = 1 \cdot \prod_{k=1}^{j} (U_{11})_{kk} \neq 0$, since L is unit triangular and U is triangular.

(2) implies (1) is proved by induction on n. It is easy for 1-by-1 matrices: $a = 1 \cdot a$. To prove it for n-by-n matrices $\tilde{A}$, we need to find unique (n-1)-by-(n-1) triangular matrices L and U, unique (n-1)-by-1 vectors l and u, and unique nonzero scalar η such that

$$\tilde{A} = \begin{bmatrix} A & b \\ c^T & \delta \end{bmatrix} = \begin{bmatrix} L & 0 \\ l^T & 1 \end{bmatrix} \times \begin{bmatrix} U & u \\ 0 & \eta \end{bmatrix} = \begin{bmatrix} LU & Lu \\ l^T U & l^T u + \eta \end{bmatrix}$$

By induction unique L and U exist such that $A = LU$. Now let $u = L^{-1}b$, $l^T = c^T U^{-1}$, and $\eta = \delta - l^T u$, all of which are unique. The diagonal entries of U are nonzero by induction, and $\eta \neq 0$ since $0 \neq det\tilde{A} = det(U) \cdot \eta$.

Question 3

See Lectures 2,3 and the course book.

1.

- Consider linear system $Ax = b$,
- $\hat{x}$ such that $\hat{x} = \delta x + x$ is its computed solution.
- Suppose $(A + \delta A)\hat{x} = b + \delta b$.
- Goal: to bound the norm of $\delta x \equiv \hat{x} - x$.
- Subtract the equalities and solve them for δx
- Rearranging terms we get:

$$\delta x = A^{-1}(-\delta A \hat{x} + \delta b)$$

Taking norms and triangle inequality leads us to

$$\|\delta x\| \leq \|A^{-1}\|(\|\delta A\| \cdot \|\hat{x}\| + \|\delta b\|)$$

Rearranging inequality gives us

$$\frac{\|\delta x\|}{\|\hat{x}\|} \leq \|A^{-1}\| \cdot \|A\| \cdot (\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta b\|}{\|A\| \cdot \|\hat{x}\|})$$

where $k(A) = \|A^{-1}\| \cdot \|A\|$ is the condition number of the matrix A

2.

$$error = \frac{||\widetilde{x} - x||_\infty}{||\widetilde{x}||_\infty} \leq ||A^{-1}||_\infty \cdot \frac{||r||_\infty}{||\widetilde{x}||_\infty}, \qquad (2.13)$$

where $r = A\widetilde{x} - b$ is the residual. We estimate $||A^{-1}||_\infty$ by applying Algorithm 2.5 to $B = A^{-T}$, estimating $||B||_1 = ||A^{-T}||_1 = ||A^{-1}||_\infty$ (see parts 5 and 6 of Lemma 1.7).

Question 4

1. See Lecture 6 and the course book.

2. We will derive the formula for the x that minimizes $||Ax - b||_2$ using the decomposition $A = QR$ in three slightly different ways. First, we can always choose $m - n$ more orthonormal vectors $\tilde{Q}$ so that $[Q, \tilde{Q}]$ is a square orthogonal matrix (for example, we can choose any $m - n$ more independent vectors $\tilde{X}$ that we want and then apply Algorithm

3.1 to the n-by-n nonsingular matrix $[Q, \tilde{X}]$). Then

$$\begin{aligned}
||Ax-b||_2^2 &= ||[Q,\tilde{Q}]^T(Ax-b)||_2^2 \\
&= \left\| \begin{bmatrix} Q^T \\ \tilde{Q}^T \end{bmatrix} (QRx-b) \right\|_2^2 \\
&= \left\| \begin{bmatrix} I^{n\times n} \\ O^{(m-n)\times n} \end{bmatrix} Rx - \begin{bmatrix} Q^T b \\ \tilde{Q}^T b \end{bmatrix} \right\|_2^2 \\
&= \left\| \begin{bmatrix} Rx - Q^T b \\ -\tilde{Q}^T b \end{bmatrix} \right\|_2^2 \\
&= \left\| Rx - Q^T b \right\|_2^2 + \|\tilde{Q}^T b\|_2^2 \\
&\geq \|\tilde{Q}^T b\|_2^2.
\end{aligned}$$

We can solve $Rx - Q^T b = 0$ for x, since A and R have the same rank, n, and so R is nonsinsular. Then $x = R^{-1}Q^T b$, and the minimum value of $||Ax-b||_2$ is $||\tilde{Q}^T b||_2$.

Here is a second, slightly different derivation that does not use the matrix $\tilde{Q}$. Rewrite $Ax-b$ as

$$\begin{aligned}
Ax-b &= QRx - b = QRx - (QQ^T + I - QQ^T)b \\
&= Q(Rx - Q^T b) - (I - QQ^T)b.
\end{aligned}$$

Note that the vectors $Q(Rx - Q^T b)$ and $(I-QQ^T)b$ are orthogonal, because $(Q(Rx - Q^T b))^T((I-QQ^T)b) = (Rx-Q^T b)^T[Q^T(I-QQ^T)]b = (Rx-Q^T b)^T[0]b = 0$. Therefore, by the Pythagorean theorem,

$$\begin{aligned}
\|Ax-b\|_2^2 &= \|Q(Rx-Q^T b)\|_2^2 + \|(I-QQ^T)b\|_2^2 \\
&= \|Rx - Q^T b\|_2^2 + \|(I-QQ^T)b\|_2^2.
\end{aligned}$$

where we have used part 4 of Lemma 1.7 in the form This sum of squares is minimized when the first term is zero, i.e., $x = R^{-1}Q^T b$.

Finally, here is a third derivation that starts from the normal equations solution:

$$\begin{aligned}
x &= (A^T A)^{-1} A^T b \\
&= (R^T Q^T QR)^{-1} R^T Q^T b = (R^T R)^{-1} R^T Q^T b \\
&= R^{-1} R^{-T} R^T Q^T b = R^{-1} Q^T b.
\end{aligned}$$

3. See Lecture 7 and the course book.

Question 5

- 1. First, we need to find a reflection that transforms the first column of matrix A

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 0 & 3 & 1 \\ 3 & 4 & 7 \end{bmatrix}$$

We have:

$$\mathbf{u} = \mathbf{x} + \alpha \mathbf{e}_1,$$

where $\mathbf{x} = (4,0,3)^T$, $\alpha = -sign(4) \cdot ||x||$

$$\mathbf{v} = \frac{\mathbf{u}}{\|\mathbf{u}\|}.$$

Here,

$$\alpha = -5.$$

Therefore

$$\mathbf{u} = (-1, 0, 3)^T, \quad ||u|| = \sqrt{10}.$$

and $\mathbf{v} = \frac{1}{\sqrt{10}}(-1,0,3)^T$, and then

$$P_1 = I - \frac{2}{\sqrt{10}\sqrt{10}}\begin{pmatrix}-1\\0\\3\end{pmatrix}\begin{pmatrix}-1 & 0 & 3\end{pmatrix}$$
$$= I - \frac{1}{5}\begin{pmatrix}1 & 0 & -3\\0 & 0 & 0\\-3 & 0 & 9\end{pmatrix}$$
$$= \begin{pmatrix}4/5 & 0 & 3/5\\0 & 1 & 0\\3/5 & 0 & -4/5\end{pmatrix}.$$

Now observe:

$$P_1A = \begin{pmatrix}5 & 5.6 & 6.6\\0 & 3 & 1\\0 & -0.8 & -3.8\end{pmatrix},$$

so we already have almost a triangular matrix. We only need to zero the (3, 2) entry.

Take the $(1,1)$ minor, and then apply the process again to

$$A' = M_{11} = \begin{pmatrix}3 & 1\\-0.8 & -3.8\end{pmatrix}.$$

We have:

$$\mathbf{u} = \mathbf{x} + \alpha\mathbf{e}_1,$$

where $\mathbf{x} = (3,-0.8)^T$, $\alpha = -sign(3)\cdot ||x||$

$$\mathbf{v} = \frac{\mathbf{u}}{||\mathbf{u}||}.$$

Here,

$$\alpha = -3.1048.$$

Therefore

$$\mathbf{u} = (-0.1048,-0.8)^T, \quad ||u|| = 0.8068.$$

and $\mathbf{v} = \frac{1}{0.8068}(-0.1048,-0.8)^T$, and then

$$P_2' = I - \frac{2}{0.651}\begin{pmatrix}-0.1048\\-0.8\end{pmatrix}\begin{pmatrix}-0.1048 & -0.8\end{pmatrix}$$
$$= I - \frac{2}{0.651}\begin{pmatrix}0.011 & 0.0838\\0.0838 & 0.64\end{pmatrix}$$
$$= \begin{pmatrix}0.9662 & -0.2575\\0.2575 & -0.9662\end{pmatrix}.$$

Then the second matrix of the Householder transformation is

$$P_2 = \begin{pmatrix}1 & 0 & 0\\0 & 0.9662 & -0.2575\\0 & -0.2575 & -0.9662\end{pmatrix}$$

Now, we find

$$R = P_2P_1A = \begin{pmatrix} 5 & 5.6 & 6.6 \\ 0 & 3.1046 & 1.9447 \\ 0 & 0.0005 & 3.4141 \end{pmatrix}.$$

The matrix P is orthogonal and R is upper triangular, so $A = QR$ is the required QR-decomposition with $P = P_1^T P_2^T$.

- 2. To obtain QR decomposition of the matrix A

$$\mathbf{A} = \begin{bmatrix} 4 & 4 & 3 \\ 3 & 3 & 1 \\ 0 & 4 & 7 \end{bmatrix}$$

using Given's rotation we have to zero out $(2,1)$ and $(3,2)$ elements of the matrix A.

1. First, we zero out element $(2,1)$ of the matrix A.
To do that we compute c, s from the known $a = 4$ and $b = 3$ as

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

to get formulas:

$$\begin{aligned} r &= \sqrt{a^2 + b^2} = \sqrt{4^2 + 3^2} = 5, \\ c &= \frac{a}{r} = 0.8, \\ s &= \frac{-b}{r} = -0.6. \end{aligned}$$

The first Given's matrix will be

$$\mathbf{G_1} = \begin{bmatrix} c & -s & 0 \\ s & c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

or

$$\mathbf{G_1} = \begin{bmatrix} 0.8 & 0.6 & 0 \\ -0.6 & 0.8 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Then

$$\mathbf{G_1} \cdot \mathbf{A} = \begin{bmatrix} 5 & 5 & 3 \\ 0 & 0 & -1 \\ 0 & 4 & 7 \end{bmatrix}$$

2. Next step is to construct second Given's matrix G_2 in order to zero out $(3,2)$ element of the matrix $G_1 \cdot A$.
To do that we compute c, s from the known $a = 0$ and $b = 4$ as

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

to get formulas:

$$r = \sqrt{a^2 + b^2} = \sqrt{0^2 + 4^2} = 4,$$
$$c = \frac{a}{r} = 0,$$
$$s = \frac{-b}{r} = -1.$$

The second Given's matrix will be

$$\mathbf{G_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}$$

or

$$\mathbf{G_2} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$$

Then upper triangular matrix R in the QR decomposition will be

$$\mathbf{R} = \mathbf{G_2} \cdot \mathbf{G_1} \cdot \mathbf{A} = \begin{bmatrix} 5 & 5 & 3 \\ 0 & 4 & 7 \\ 0 & 0 & 1 \end{bmatrix}$$

Then $A = G_1^T \cdot G_2^T \cdot R = QR$ will be QR decomposition of the matrix A with $Q = G_1^T \cdot G_2^T$ given by

$$\mathbf{Q} = \begin{bmatrix} 0.8 & 0 & 0.6 \\ 0.6 & 0 & -0.8 \\ 0 & 1 & 0 \end{bmatrix}$$

Question 6

- 1. *Schur canonical form.* Given A, there exists a unitary matrix Q and an upper triangular matrix T such that $Q^* A Q = T$. The eigenvalues of A are the diagonal entries of T.

  Real Schur canonical form. If A is real, there exists a real orthogonal matrix V such that $V^T A V = T$ is quasi-upper triangular. This means that T is block upper triangular with 1-by-1 and 2-by-2 blocks on the diagonal. Its eigenvalues are the eigenvalues of its diagonal blocks. The 1-by-1 blocks correspond to real eigenvalues, and the 2-by-2 blocks to complex conjugate pairs of eigenvalues.
- 2. Let $S = [x_1, \dots, x_n]$ the nonsingular matrix of right eigenvectors. and we know that A is diagonalizible and thus $AS = S\Lambda$, where $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$, since the columns x_i of S are eigenvectors. This is equivalent to $AS^{-1} = S^{-1}\Lambda$, so the rows of S^{-1} are conjugate transposes of the left eigenvectors y_i.
- 3. The expression $\sec\Theta(y, x) = 1/|y^* x|$ is the condition number of the eigenvalue λ.

Question 7

- 1. Let A be the m-by-n matrix with $m \geq n$ and has a full rank such that $A = U\Sigma V^T$. Then Moore-Penrose pseudoinverse of A is $A^+ = (A^T A)^{-1} A^T$. If $m < n$ then $A^+ = A^T (AA^T)^{-1}$.
- 2.

$$\begin{aligned} AA^+A &= U\Sigma V^T (V\Sigma U^T U\Sigma V^T)^{-1} V\Sigma U^T U\Sigma V^T \\ &= U\Sigma V^T (V^{-1}\Sigma^{-2} V^{-T})\Sigma^2 = U\Sigma V^{-1} = U\Sigma V^T = A. \end{aligned}$$