CS 383C
CAM 383C/M 383E **Numerical Analysis: Linear Algebra** **Fall 2008**

Solutions to Homework 5

Lecturer: Inderjit Dhillon Date Due: Oct 22, 2008
Keywords: *Conditioning, Stability, Error Analysis, Gaussian Elimination*

1. $U = A, L = I$
   for $k = 1 : m-1$ **do**
   $\boldsymbol{l}_{k+1:m,k} = \boldsymbol{u}_{k+1:m,k}/u_{kk}$
   $U_{k+1:m,k:m} = U_{k+1:m,k:m} - \boldsymbol{l}_{k+1:m,k}U_{k,k:m}$
   end for

2. (a) Elimination by columns from left to right is equivalent to post-multiplication by an upper-triangular matrix. $m-1$ such operations will make A a lower-triangular matrix, thus:

$$AU_1U_2\dots U_{m-1} = L \Rightarrow A = LU_{m-1}^{-1}U_{m-2}^{-1}\dots U_1^{-1}.$$

   Now the inverse of a upper triangular matrix is upper-triangular and multiplication of two upper-triangular matrices is also an upper-triangular matrix. Hence, $A = LU$.

   (b) Gaussian Elimination is equivalent to pre-multiplication with lower-triangular matrices. Thus after rescaling of columns, Gaussian Elimination is equivalent to: $L_{m-1}L_{m-2}\dots L_1AD = U \Rightarrow A = LUD^{-1}$. Thus, unknowns are rescaled by D^{-1}.

   (c) Elimination by columns from left to right is equivalent to post-multiplication by an upper-triangular matrix. Thus, using part (a), $U = D\tilde{U}$. Thus, $A = LD\tilde{U}$, where $\tilde{U}$ incorporates the additional column operations.

3. (a) We prove the claim using induction. Base case follows directly from the fact that $x_1 \otimes a_1 = x_1a_1(1+\epsilon), |\epsilon| \le \epsilon_{machine}$. Now, $fl(\boldsymbol{x}^T\boldsymbol{a}) = x_1 \otimes a_1 \oplus x_2 \otimes a_2 \oplus \dots x_n \otimes a_n$. Using induction hypothesis, $fl(\boldsymbol{x}^T\boldsymbol{a}) = (\sum_{i=1}^{n-1} x_ia_i + e_{n-1}) \oplus x_n \otimes a_n, |e_{n-1}| \le (n-1)\epsilon_{machine}\sum_{i=1}^{n-1}|x_i||a_i| + O(\epsilon_{machine}^2)$. Thus, $fl(\boldsymbol{x}^T\boldsymbol{a}) = (\sum_{i=1}^{n-1} x_ia_i + e_{n-1} + x_na_n(1+\epsilon_1))(1+\epsilon_2) = \boldsymbol{x}^T\boldsymbol{a} + e_{n-1} + x_na_n\epsilon_1 + \boldsymbol{x}^T\boldsymbol{a}\epsilon_2 + e_{n-1}\epsilon_2 + x_na_n\epsilon_1\epsilon_2$, where $|\epsilon_1| \le \epsilon_{machine}, |\epsilon_2| \le \epsilon_{machine}$. Thus, $fl(\boldsymbol{x}^T\boldsymbol{a}) = \boldsymbol{x}^T\boldsymbol{a} + e_n$, where $e_n = e_{n-1} + x_na_n\epsilon_1 + \boldsymbol{x}^T\boldsymbol{a}\epsilon_2 + e_{n-1}\epsilon_2 + x_na_n\epsilon_1\epsilon_2$. Hence, $|e_n| \le |e_{n-1}| + |x_n||a_n|\epsilon_{machine} + |\boldsymbol{x}|^T|\boldsymbol{a}|\epsilon_{machine} + O(\epsilon_{machine}^2) = (n-1)\sum_i^{n-1}|x_i||a_i|\epsilon_{machine} + |x_n||a_n|\epsilon_{machine} + |\boldsymbol{x}|^T|\boldsymbol{a}|\epsilon_{machine} \le n|\boldsymbol{x}|^T|\boldsymbol{a}|\epsilon_{machine}$. Hence proved.

   (b) $(XA)_{ij} = \boldsymbol{x}_i^T\boldsymbol{a}_j$, where $\boldsymbol{x}_i^T$ represents i-th row of X and $\boldsymbol{a}_j$ represents j-th column of A. Using part (a), $fl((XA)_{ij}) = (XA)_{ij} + e_{ij}$, where $|e_{ij}| \le n|\boldsymbol{x}_i|^T|\boldsymbol{a}_j|\epsilon_{machine} + O(\epsilon_{machine}^2)$. Thus $\|fl(XA) - XA\|_F \le \sqrt{\sum_{ij} n^2\epsilon_{machine}^2(|X||A|)_{ij}^2} + O(\epsilon_{machine}^2) = n\epsilon_{machine}\|\,|X|\,|A|\,\|_F + O(\epsilon_{machine}^2) \le n\epsilon_{machine}\|\,|X|\,\|_F\|\,|A|\,\|_F + O(\epsilon_{machine}^2) = n\epsilon_{machine}\|X\|_F\|A\|_F + O(\epsilon_{machine}^2)$, as $\|AB\|_F \le \|A\|_F\|B\|_F$ and $\|\,|A|\,\|_F = \|A\|_F$.

   (c) Using part (b), $fl(XA) = XA + E$, where $\|E\|_F \le n\|X\|_F\|A\|_F O(\epsilon_{machine})$. Now $fl(XA) = X(A + X^{-1}E) = X(A+\delta A)$, where $\delta A = X^{-1}E$. Thus, $\|\delta A\|_F \le \|X^{-1}\|_F\|E\|_F \le n\|X^{-1}\|_F\|X\|_F\|A\|_F O(\epsilon_{machine})$. Thus, $\frac{\|\delta A\|_F}{\|A\|_F} \le n\kappa(X)O(\epsilon_{machine})$.

4. $(A+\delta A)(\boldsymbol{x}+\delta\boldsymbol{x}) = A\boldsymbol{x} + A\delta\boldsymbol{x} + \delta A\boldsymbol{x} + \delta A\delta x = \boldsymbol{b} + \delta\boldsymbol{b}$. Now $A\boldsymbol{x} = \boldsymbol{b}$, hence $(A+\delta A)\delta\boldsymbol{x} = \delta\boldsymbol{b} - \delta A\boldsymbol{x}$. Thus, $\|\delta\boldsymbol{x}\| = \|(A+\delta A)^{-1}(\delta\boldsymbol{b} - \delta A\boldsymbol{x})\| \le \|(A+\delta A)^{-1}\|(\|\delta\boldsymbol{b}\| + \|\delta A\boldsymbol{x}\|) \le \|(I + A^{-1}\delta A)^{-1}\|\|A^{-1}\|(\|\delta\boldsymbol{b}\| + \|\delta A\boldsymbol{x}\|)$.

Using Homework 2, Problem 2(c): if $\|A^{-1}\delta A\| \leq 1$, then $\|(I + A^{-1}\delta A)^{-1}\| \leq \frac{1}{1-\|A^{-1}\delta A\|} \leq \frac{1}{1-\|A^{-1}\|\|\delta A\|}$. Thus, $\|\delta\boldsymbol{x}\| \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|\delta A\|}(\|\delta\boldsymbol{b}\| + \|\delta A\boldsymbol{x}\|)$. Now $\boldsymbol{b} = A\boldsymbol{x}$, thus $\|\boldsymbol{b}\| \leq \|A\|\|\boldsymbol{x}\|$. Hence,

$$\|\delta\boldsymbol{x}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\|\|\delta A\|}\left(\frac{\|\delta\boldsymbol{b}\|\|A\|\|\boldsymbol{x}\|}{\|\boldsymbol{b}\|} + \frac{\|\delta A\|\|\boldsymbol{x}\|\|A\|}{\|A\|}\right),$$

i.e.,

$$\frac{\|\delta\boldsymbol{x}\|}{\|\boldsymbol{x}\|} \leq \frac{\kappa(A)}{1 - \kappa(A)\frac{\|\delta A\|}{\|A\|}}\left(\frac{\|\delta A\|}{\|A\|} + \frac{\|\delta\boldsymbol{b}\|}{\|\boldsymbol{b}\|}\right).$$