Homework 4

Math 5424

Numerical Linear Algebra

Alexander Novotny
Zuriah Quinton

October 30, 2023

1. Let $\boldsymbol{L}$ be a lower triangular matrix and solve $\boldsymbol{L}\vec{x} = \vec{b}$ by forward substitution. Show that barring overflow or underflow, the computed solution $\vec{x}$ satisfies $(\boldsymbol{L} + \delta\boldsymbol{L})\vec{x} = \vec{b}$, where $|\delta l_{ij}| \leq n\varepsilon|l_{ij}|$, where ε is the machine precision. This means that forward substitution is backward stable. Argue that backward substitution for solving upper triangular systems satisfies the same bound.

   Proof. ☺

2. Matrix $\boldsymbol{A}$ is called *strictly column diagonally dominant*, or diagonally dominant for short, if

$$|a_{ii}| > \sum_{j=1, j\neq i}^{n} |a_{ji}|$$

   Show that Gaussian elimination with partial pivoting does not actually permute any rows, i.e., that it is identical to Gaussian elimination without pivoting. Hint: Show that after one step of Gaussian elimination, the trailing $(n-1)$-by-$(n-1)$ submatrix, the *Schur complement* of a_{11} in $\boldsymbol{A}$, is still diagonally dominant.

   Proof. Our first step of Gaussian elimination can be seen as:

$$\boldsymbol{A} = \begin{bmatrix} \alpha_{11} & \vec{c}^{\top} \\ \vec{a} & \hat{\boldsymbol{A}} \end{bmatrix} = \begin{bmatrix} 1 & \vec{0} \\ \vec{l} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \alpha_{11} & \vec{c}^{\top} \\ \vec{0} & \hat{\boldsymbol{A}} - \vec{l}\vec{c}^{\top} \end{bmatrix},$$

   where $\vec{l} = \frac{\vec{a}}{\alpha_{11}}$. Note that $c_i = \alpha_{1,i+1}$, $l_i = \frac{\alpha_{i+1,1}}{\alpha_{11}}$ and for $\boldsymbol{B} = \hat{\boldsymbol{A}} - \vec{l}\vec{c}^{\top}$, we have $b_{ij} = \alpha_{i+1,j+1} - l_i c_j = \alpha_{i+1,j+1} - \frac{\alpha_{i+1,1}\alpha_{1,j+1}}{\alpha_{11}}$. Then we have

$$|b_{ii}| = \left| \alpha_{i+1,i+1} - \frac{\alpha_{i+1,1}\alpha_{1,j+1}}{\alpha_{11}} \right|$$

$$\geq |\alpha_{i+1,i+1}| - \frac{|\alpha_{i+1,1}||\alpha_{1,i+1}|}{\alpha_{11}}$$

   ☺

3. Let $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ be matrices with dimensions such that the product $\boldsymbol{A}^{\top}\boldsymbol{C}\boldsymbol{B}^{\top}$ is well defined. Let $\mathcal{X}$ be the set of matrices $\boldsymbol{X}$ minimizing $\|\boldsymbol{A}\boldsymbol{X}\boldsymbol{B} - \boldsymbol{C}\|_F$, and let $\boldsymbol{X}_0$ be the unique member of $\mathcal{X}$ minimizing $\|\boldsymbol{X}\|_F$. Show that $\boldsymbol{X}_0 = \boldsymbol{A}^+\boldsymbol{C}\boldsymbol{B}^+$. Hint: Use the SVDs of $\boldsymbol{A}$ and $\boldsymbol{B}$.

   Proof. ☺

4. Show that the Moore—Penrose pseudoinverse of A satisfies the following identities:

$$\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} = \boldsymbol{A},$$
$$\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ = \boldsymbol{A}^+,$$
$$\boldsymbol{A}^+\boldsymbol{A} = (\boldsymbol{A}^+\boldsymbol{A})^\top,$$
$$\boldsymbol{A}\boldsymbol{A}^+ = (\boldsymbol{A}\boldsymbol{A}^+)^\top.$$

Proof. We have

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{A}^+\boldsymbol{A} &= \boldsymbol{A}(\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top\boldsymbol{A} \\
&= \boldsymbol{A}\cancel{(\boldsymbol{A}^\top\boldsymbol{A})^{-1}}\cancel{(\boldsymbol{A}^\top\boldsymbol{A})} \\
&= \boldsymbol{A}.
\end{aligned}$$

Then,

$$\begin{aligned}
\boldsymbol{A}^+\boldsymbol{A}\boldsymbol{A}^+ &= (\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top\boldsymbol{A}\boldsymbol{A}^+ \\
&= \cancel{(\boldsymbol{A}^\top\boldsymbol{A})^{-1}}\cancel{(\boldsymbol{A}^\top\boldsymbol{A})}\boldsymbol{A}^+ \\
&= \boldsymbol{A}^+.
\end{aligned}$$

Further,

$$\begin{aligned}
\boldsymbol{A}^+\boldsymbol{A} &= (\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top\boldsymbol{A} \\
&= \cancel{(\boldsymbol{A}^\top\boldsymbol{A})^{-1}}\cancel{(\boldsymbol{A}^\top\boldsymbol{A})} \\
&= \boldsymbol{I},
\end{aligned}$$

and

$$\begin{aligned}
(\boldsymbol{A}^+\boldsymbol{A})^\top &= \boldsymbol{A}^\top(\boldsymbol{A}^+)^\top \\
&= \boldsymbol{A}^\top((\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top)^\top \\
&= \boldsymbol{A}^\top\boldsymbol{A}\left((\boldsymbol{A}^\top\boldsymbol{A})^{-1}\right)^\top \\
&= \boldsymbol{A}^\top\boldsymbol{A}\left((\boldsymbol{A}^\top\boldsymbol{A})^\top\right)^{-1} \\
&= \cancel{(\boldsymbol{A}^\top\boldsymbol{A})}\cancel{(\boldsymbol{A}^\top\boldsymbol{A})^{-1}} \\
&= \boldsymbol{I},
\end{aligned}$$

so $\boldsymbol{A}^+\boldsymbol{A} = (\boldsymbol{A}^+\boldsymbol{A})^\top$. Finally,

$$\begin{aligned}
(\boldsymbol{A}\boldsymbol{A}^+)^\top &= (\boldsymbol{A}^+)^\top\boldsymbol{A}^\top \\
&= ((\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top)^\top\boldsymbol{A}^\top \\
&= \boldsymbol{A}\left((\boldsymbol{A}^\top\boldsymbol{A})^{-1}\right)^\top\boldsymbol{A}^\top \\
&= \boldsymbol{A}\left((\boldsymbol{A}^\top\boldsymbol{A})^\top\right)^{-1}\boldsymbol{A}^\top \\
&= \boldsymbol{A}(\boldsymbol{A}^\top\boldsymbol{A})^{-1}\boldsymbol{A}^\top \\
&= \boldsymbol{A}\boldsymbol{A}^+.
\end{aligned}$$

☺

5. (a) Describe a variant of Gaussian elimination that introduces zeros into the columns of $\boldsymbol{A}$ in the order $n : \text{-}1 : 2$ and which produces the factorization $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{L}$ where $\boldsymbol{U}$ is the unit upper triangular and $\boldsymbol{L}$ is lower triangular.

Answer.

(b) Based on your algorithm, prove/provide the necessary and sufficient determinant conditions for the existence of the UL decomposition.

Answer.

(c) Write a Matlab code to implement the UL decomposition and apply it to

$$\begin{bmatrix} 1 & 0 & 2 & 1 \\ -4 & 5 & 3 & -1 \\ -1 & 3 & 1 & 1 \\ 0 & 2 & 0 & 1 \end{bmatrix}$$

to verify that your code generates the required decomposition $A = UL$.

Answer.

6. Even though we rarely need to compute the inverse of a matrix, let us think about it in this problem. Let $A \in \mathbb{R}^{n \times n}$ be an invertible matrix. Describe an algorithm (based on the LU Decomposition/Gaussian Elimination) that computes A^{-1} with an operation count of $8n^3/3$ flops (ignoring the lower order terms).

Answer.

7. Inspired by the presentation in [Trefethen and Bau, SIAM Press, 1997], in this problem, we will numerically investigate the growth factor in LU with partial pivoting. In the class (see October 13th notes), we showed that the growth factor ρ_{pp} could be as large as $\rho_{pp} = 2n - 1$. Indeed we had found an example where this upper bound is attained. However, as we mentioned, the algorithm behaves much better in practice. Here, we will try LU with partial pivoting on random matrices with varying dimensions and plot the observations.

Use the command `n = ceil(logspace(1, 3, 1000))` to create (approximately) logarithmically spaced matrix dimensions between 10 and 1000. Some of the dimensions will be repeated. The variable `n` is a vector of size 1000 with entries ranging from 10 to 1000. Then, for every entry `n(i)` of `n`, i.e., for $i = 1, 2, \ldots, 1000$, create a random matrix A using `A= randn(n(i), n(i))/sqrt(n(i))`. So, we are creating a random matrix of varying dimensions with normally distributed entries having mean zero and standard deviation $\sqrt{n(i)}$. Then, compute the growth factor ρ_{pp} for every A. At the end you will have a vector of size 1000 whose entries corresponding to the growth factor for every random A. Using the `loglog` command (logarithmic scale both in the horizontal and 2 vertical axes), plot the growth factor vs the matrix dimensions `n`. On the same plot, by using the `hold on` command, plot the growth rate of $\sqrt{n}$. How is the observed/numerical growth behaving with respect to the theoretical upper bound $2n - 1$ and with respect to $\sqrt{n}$? Comment on your observations.

Answer.