Test 1 - Thu Oct. 4th
Math 224 - Fall 2007

Dr. Mauro Maggioni
E-mail: mauro.maggioni at duke.edu

Write your name at the top right corner of each page. Provide full answers and motivations. No books, notes, or calculators are allowed. You have 1 hour.

True or False? Fully justify your answers, with proofs or counterexamples as required by your answers.

a. A stable algorithm will return a solution close to the true solution, no matter what the condition number of the problem is. As part of your answer, define what a stable algorithm is, and what the condition number of a problem is.

False. A stable algorithm will return a solution of a nearby problem. If the problem is well-conditioned, then solutions to nearby problems are close to the solutions of the original problem. But if the original problem is badly conditioned, the solutions of nearby problems are far from solutions of the original problem. This follows immediately from the definition of condition number: if the problem is formulated as $y = f(x)$, then the (relative) condition number is defined as $|y - \hat{y}|/|x - \hat{x}|$, where $\hat{y} = f(\hat{x})$ is the solution obtained by perturbing the data from x to $\hat{x}$. Is the condition number is large small perturbations $|x - \hat{x}|$ will greatly affect the solution $\hat{y}$: for a stable algorithm $|x - \hat{x}|$ (where $\hat{x}$ is obtained by backward analysis) is small, but this will not imply that $|y - \hat{y}|$ is small when the problem is ill-conditioned.

b. Pivoting in Gaussian elimination is only needed if the matrix is ill-conditioned.

False: Gaussian elimination can dramatically fail even for perfectly conditioned problems. For example it will fail immediately if a 0 is on the first diagonal entry, even if the matrix is perfectly conditioned, for example $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

c. Every positive-definite matrix symmetric A has a Cholesky factorization. [If it helps, you may assume the existence of LU factorizations]. As part of your answer, define what the Choleski factorization is.

True. This is reasonable since $A = LU$ and $A = A^T = U^T L^T$, and the LU decomposition is essentially unique up to fixing the diagonal entries of L or U (for example by forcing the diagonal entries of U to be 1). It is also clear that we need to assume A positive definite since if $A = L^T L$, which is the Cholesky decomposition, where L is lower triangular, and all diagonal entries of L are nonzero, then necessarily A needs to be positive definite (since $x^T Ax = (Lx)^T (Lx) = ||Lx||_2^2 > 0$ if $||x|| > 0$ since L is invertible). An actual proof of existence of Cholesky can done by induction on the size of the matrix, which also shows that no pivoting is necessary.

d. A least squares problem $Ax \simeq b$, with A a $m \times n$ matrix with $m > n$, always has at least one solution.

True: we need to minimize over x the quantity $||Ax - b||_2^2$. Clearly $||x||$ cannot be too big since this quantity goes to infinity as $||x||$ goes to infinity, therefore we can restrict the problem to $||x||_2 \leq M$ for M large enough (depending on b and A). But then we are minimizing a continuous function on a compact set, and certainly this has at least one solution. In fact, if A has full rank, $A^T A$ is positive definite, but it's also the Hessian of the linear function $x \mapsto Ax$, and so the minimum is unique. If A has not full rank, there are infinitely many x that give the same point Ax at which the minimum is attained.

1. Define machine precision ϵ_{mach}. Discuss whether it is similar or different from the machine underflow UFL.

These are straight from the book. A main quantitative statement here is that the relative error in rounding off is upper bounded by ϵ_{mach}.

2. What can you say about the difference between the true result of a basic arithmetic operation and the result of performing it in floating point arithmetic? What is cancellation, and why is it usually bad?

This is also straight from the book. The key quantitative statement here is that $|\text{fl}(x \,\text{op}\, y) - (x \,\text{op}\, y)|/|(x \,\text{op}\, y)| \leq \epsilon_{mach}$, where op is a basic operation (barring overflows/underflows, divisions by 0 etc...). This is very important, as it bounds the relative error between computing in FP arithmetic vs. in full arithmetic. If we didn't have this, we wouldn't know how large errors induced by floating point arithmetic would be. Rounding errors are handled in the previous question!

3. Define the relative condition number of a problem, and explain its meaning and purpose. If a problem has (relative) condition number close to 1, what does it mean? Is it a good or bad thing? Why?

For the problem $y = f(x)$, the relative condition number is defined as $\frac{|f(x)-f(\hat{x})|/|f(x)|}{|x-\hat{x}|/|x|}$, where $\hat{x}$ is a point nearby the data x (and typically one takes the supremum over all nearby $\hat{x}$). It measures how much a relative perturbation in the data can affect the solution of the problem, in relative terms. A condition number of 1 since it implies a relative error in the solution which has at most the same size as the relative perturbation in the data, which is general the best we can expect (for example it is the best one can expect if f is linear).

4. Define the condition number of a matrix, explain how it is related to the concept of singularity of a matrix, and show examples of matrices (at least 2×2) with small and high condition number. Explain and comment on the implications of the following formula:

$$\frac{||x - \hat{x}||}{||\hat{x}||} \leq \text{cond}(A)\frac{||E||}{||A||},$$

where $Ax = b$, $(A + E)\hat{x} = b$. Suppose $\epsilon_{mach} \sim 10^{-16}$ and you are solving $Ax = b$, and $\text{cond}(A) \sim 10^{6}$. What do you expect the relative error of your solution, with respect to the true solution, to be? Does the answer depend on the algorithm you are using? If yes, explain the best and worst case scenarios.

5. Describe the LU factorization of a square nonsingular matrix A, and how it can be used to solve the linear system $Ax = b$. Also, describe column pivoting and why it is expected to be less unstable than LU factorization. If you knew that A is positive definite, how would you solve the linear system?

6. What is a linear least squares problem $Ax \simeq b$? Explain, discuss existence and uniqueness of the solution, and give a geometrical interpretation.

7. What is a QR factorization of a matrix A, m by n? Why is useful to solve least square problems? List three techniques for computing a QR factorization, and discuss one of your choice in more detail.

8. Describe the SVD decomposition of a matrix A, m by n with $m > n$, and how it can be used to solve a least square problem. As part of your answer, describe the pseudo-inverse of A.