CAAM 453/553 · NUMERICAL ANALYSIS I

Problem Set 2

Posted Saturday 12 September 2009. Due Monday 21 September 2009.
CAAM 453 students should complete problems (1 or 3), 2, 4, 5, 6 [100 points].
CAAM 553 students should complete problems 1, 2, 3, 4, 5, 6, 7. [150 points].
Students are welcome to attempt more problems if they wish.

The problems in this set mainly addresses ways of measuring the sensitivity of a linear system $\mathbf{Ax} = \mathbf{b}$ *to changes in the matrix* $\mathbf{A}$*, and a few vagaries of floating point arithmetic. Problems 2 and 3 address the distance of a matrix from the set of singular matrices, very important quantity. Problem S2 investigates how this can be extended to measure the 'distance to instability'.*

1. [20 points]

   (a) Consider the vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^3$, with both sets $\{\mathbf{x}_1, \mathbf{x}_2\}$ and $\{\mathbf{y}_1, \mathbf{y}_2\}$ linearly independent. Hence

   $$\mathcal{X} = \text{span}\{\mathbf{x}_1, \mathbf{x}_2\}, \qquad \mathcal{Y} = \text{span}\{\mathbf{y}_1, \mathbf{y}_2\}$$

   are both two dimensional subspaces of $\mathbb{R}^3$, and must have a nontrivial intersection.

   Show how you can use three (full-size) QR factorizations of 3×2 matrices to determine some nonzero vector $\mathbf{v} \in \mathbb{R}^3$ in the intersection $\mathcal{X} \cap \mathcal{Y}$.

   [Trefethen and Bau, Problem 7.4]

   (b) Implement your algorithm in MATLAB and test it with

   $$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{y}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}.$$

2. [20 points]

   Recall the Taylor's series

   $$\frac{1}{1-x} = 1 + x + x^2 + \cdots,$$

   which converges provided $|x| < 1$ for $x \in \mathbb{C}$.

   For this problem, assume $\|\cdot\|$ is any matrix norm that is submultiplicative, $\|\mathbf{XY}\| \le \|\mathbf{X}\|\|\mathbf{Y}\|$.

   (a) Prove the matrix analogue of this formula, which is known as the *Neumann series*:
   If $\|\mathbf{X}\| < 1$, then

   $$(\mathbf{I} - \mathbf{X})^{-1} = \mathbf{I} + \mathbf{X} + \mathbf{X}^2 + \cdots.$$

   (There are two things to show here: first, that the quantity on the right is a finite matrix, i.e.,

   $$\|\mathbf{I} + \mathbf{X} + \mathbf{X}^2 + \cdots\| < \infty;$$

   second, that $(\mathbf{I} - \mathbf{X})(\mathbf{I} + \mathbf{X} + \mathbf{X}^2 + \cdots) = \mathbf{I}$.)

   (b) Suppose that $\mathbf{A}$ is invertible. Show that if $\|\mathbf{A}^{-1}\mathbf{E}\| < 1$, then $\mathbf{A} + \mathbf{E}$ is invertible.
   (To do this, you can prove that

   $$(\mathbf{A} + \mathbf{E})^{-1} = (\mathbf{I} + \mathbf{A}^{-1}\mathbf{E})^{-1}\mathbf{A}^{-1},$$

   and then explain why the matrix on the right is finite.)

   (c) Explain why part (b) implies the following lower bound on the distance of $\mathbf{A}$ from singularity:

   $$\min_{\mathbf{A} + \mathbf{E} \text{ singular}} \|\mathbf{E}\| \;\ge\; \frac{1}{\|\mathbf{A}^{-1}\|}.$$

3. [20 points]

   (a) Let $\mathbf{B} = \mathbf{x}\mathbf{y}^*$. Prove that $\|\mathbf{B}\|_2 = \|\mathbf{x}\|_2 \|\mathbf{y}\|_2$.

   For the following parts of this question, let $\mathbf{x}$ be a vector such that $\|\mathbf{A}^{-1}\mathbf{x}\|_2 = \|\mathbf{A}^{-1}\|_2$ with $\|\mathbf{x}\|_2 = 1$, define $\mathbf{y}$ to be the vector

$$\mathbf{y} = \frac{\mathbf{A}^{-1}\mathbf{x}}{\|\mathbf{A}^{-1}\|_2},$$

   and let

$$\widehat{\mathbf{E}} = -\frac{\mathbf{x}\mathbf{y}^*}{\|\mathbf{A}^{-1}\|_2}.$$

   (b) What is $\|\widehat{\mathbf{E}}\|_2$?

   (c) Show that $\mathbf{A} + \widehat{\mathbf{E}}$ is singular, i.e., there exists some vector $\mathbf{w} \neq \mathbf{0}$ such that $(\mathbf{A} + \widehat{\mathbf{E}})\mathbf{w} = \mathbf{0}$.

   (d) Parts (b) and (c) suggest an upper bound on the 2-norm distance of $\mathbf{A}$ from singularity, i.e.,

$$\min_{\mathbf{A}+\mathbf{E}\text{ singular}} \|\mathbf{E}\|_2 \leq \|\widehat{\mathbf{E}}\|_2.$$

   What does this mean when taken together with the result of Problem 2(c)?

   [Stewart and Sun]

4. [20 points]

   This question explores the relationship between the determinant and condition number. For $\mathbf{A} \in \mathbb{C}^{n\times n}$, $\det(\mathbf{A})$ is the product of the eigenvalues of $\mathbf{A}$ (just one among several equivalent definitions) while the condition number is $\kappa(\mathbf{A}) = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$. For this problem you may use any induced norm you like (e.g., 1-norm, 2-norm, ∞-norm).

   (a) Construct a family of matrices $\{\mathbf{A}_n\}_{n=1}^{\infty}$, $\mathbf{A}_n \in \mathbb{C}^{n\times n}$, such that $\det(\mathbf{A}_n) \to 0$ as $n \to \infty$, while $\kappa(\mathbf{A}_n) = 1$ for all n.

   (b) Construct a family of matrices $\{\mathbf{A}_n\}_{n=1}^{\infty}$, $\mathbf{A}_n \in \mathbb{C}^{n\times n}$, such that $\kappa(\mathbf{A}_n) \to \infty$ as $n \to \infty$, while $\det(\mathbf{A}_n) = 1$ for all n. (The best examples will exhibit exponential growth of the condition number, i.e., $\kappa(\mathbf{A}_n) \sim \gamma^n$ for some $\gamma > 1$, but this is not necessary for full credit.)

   (c) Is $\det(\mathbf{A})$ a satisfactory measure of the distance to singularity?

5. [20 points]

   For points $x_0, \ldots, x_n \in \mathbb{R}$, the associated *Vandermonde* matrix takes the form

$$\mathbf{A} = \begin{pmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ 1 & x_2 & x_2^2 & \cdots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}.$$

(a) Write MATLAB code to generate the Vandermonde matrix for data points $x_0, \dots, x_n$ uniformly distributed over the interval $[-1, 1]$. (See `linspace`.) Please do not use MATLAB's `vander` command.

(b) To test the numerical properties of the Vandermonde matrix, take $\mathbf{y} = (1,\ 1, \dots, 1)^T$ and define $\mathbf{b} = \mathbf{Ay}$. (The jth entry in $\mathbf{b}$ is simply the sum of the jth row of $\mathbf{A}$.) Create a carefully designed plot in MATLAB (use `semilogy`) that compares the following quantities for $n = 5, 6, \dots 70$.

- $\kappa(\mathbf{A}) = \|\mathbf{A}\|_2 \|\mathbf{A}^{-1}\|_2$.
- $\|\mathbf{y} - \widehat{\mathbf{y}}\|_2$, where $\widehat{\mathbf{y}}$ is the solution to the system $\mathbf{Ay} = \mathbf{b}$ that you compute using the QR factorization ($\mathbf{c} = \mathbf{Q}^*\mathbf{b}$, $\mathbf{R}\widehat{\mathbf{y}} = \mathbf{c}$). (In exact arithmetic, we should have $\widehat{\mathbf{y}} = \mathbf{y}$.) You may use MATLAB's built-in QR factorization, or your Givens QR from Homework 1.
- $\|\widehat{\mathbf{y}} - \mathbf{y}\|_2 / \kappa(\mathbf{A})$.

Explain your results in light of conditioning and floating point arithmetic.

6. [20 points]

Consider the integral

$$I_r = \int_0^1 x^{2r} \exp(-x^2)\, \mathrm{d}x.$$

In the special case $r = 0$, this integral can be calculated using the *error function*:

$$I_0 = \tfrac{1}{2}\sqrt{\pi}\, \mathrm{erf}(1).$$

(a) Verify analytically that for integers $r \geq 1$,

$$I_r = \frac{2r-1}{2} I_{r-1} - \frac{1}{2\mathrm{e}}.$$

(b) Compute $I_1, \dots, I_{25}$ using the recurrence from part (a) and

$$I_0 = .5 * \mathtt{sqrt(pi)} * \mathtt{erf(1)}.$$

(The resulting value for I_0 is accurate to roughly 15 digits.)

(c) A simple rearrangement of the recurrence in part (a) gives the *backward* recurrence

$$I_{r-1} = \frac{2}{2r-1}\left(I_r + \frac{1}{2\mathrm{e}}\right).$$

Make the (incorrect!) assumption that $I_{25} \equiv 0$ and use it to compute $I_{24}, I_{23}, \dots, I_0$.

(d) Which approach, (b) or (c), gives a more accurate value for I_{20}?
Using the value of $\varepsilon_{\mathrm{mach}}$ for IEEE double precision arithmetic, explain this result from the perspective of floating point error analysis.

N.B. A highly accurate (but more computationally expensive) value of I_r can be obtained via:

```
f_r = inline('x.^(2*r).*exp(-x.^2)', 'x','r');
Ir = quad(f_r,0,1,1e-12,[],r);
```

[Fox & Mayers; Donnelly]

7. [30 points]

One might naturally want to test the performance of a given linear algebra algorithm on a set of 'generic' or 'typical' matrices, and one might imagine that matrices with random, independent entries would be ideal. In fact, random matrices have many beautiful properties that make them anything but 'generic' or 'typical'. In this question (adapted from Trefethen and Bau's Problem 12.3) you will explore a few of these properties.

Define a random matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ to have entries with mean zero and standard deviation $1/\sqrt{n}$. Such matrices can be constructed in MATLAB via the command `A = randn(n)/sqrt(n)`. (Be sure to use `randn`, and not `rand`.) We wish to study the behavior of these matrices as the dimension n gets large.

(a) How do the eigenvalues of $\mathbf{A}$ behave as $n \to \infty$? Create four plots of the eigenvalues of $\mathbf{A}$ (computed with MATLAB's `eig` command) in the complex plane, with $n = 16, 64, 256, 1024$. In the first three plots, superimpose eigenvalues from several matrices so that a total of 1024 dots appear in each plot. (Type `hold on` to superimpose plots.) What do you conjecture about the limiting behavior?

(b) Now perform a study of $\|\mathbf{A}\|_2$. How does this quantity grow as a function of n? In Homework 1 you proved that $\|\mathbf{A}\|_2 \geq \rho(\mathbf{A})$, where $\rho(\mathbf{A})$ denotes the spectral radius. Do you expect that equality is attained in the $n \to \infty$ limit for these matrices?

(c) Perform a similar study, but for $\|\mathbf{A}^{-1}\|_2$. Can you make any predictions about growth? (If you wish to average over a number of matrices, it is better to average the logarithms of $\|\mathbf{A}^{-1}\|$ for several $\mathbf{A}$ with fixed n and exponentiate the result.)

(d) How do these results change is $\mathbf{A}$ is replaced with an upper triangular version, `A = triu(randn(n))/sqrt(n)` ?

(e) Again let `A = randn(n)/sqrt(n)`. Now plot a histogram (try `hist(data, 20)`) of the *singular values* (square roots of the eigenvalues of $\mathbf{A}^*\mathbf{A}$, computed via `svd(A)`). You should observe the 'quarter-circle law' discovered by Marčenko and Pastur.

Supplemental Problems

These extra problems go beyond the scope of the lectures and are thus optional.

S1. Note that Problem 2 holds for all submultiplicative matrix norms (and hence for all induced norms), while Problem 3 was specific to the matrix 2-norm.

Show that the conclusion from Problem 3(d) can be generalized to *any induced matrix norm.*

Hint. Let $\|\cdot\|$ be any vector norm on $\mathbb{C}^n$. Then for any $\mathbf{z} \in \mathbb{C}^n$ with $\|\mathbf{z}\| = 1$, there exists some $\mathbf{w} \in \mathbb{C}^n$ such that $\mathbf{w}^*\mathbf{z} = 1$, where the star denotes the conjugate-transpose, as usual. (This is a consequence of the Hahn–Banach Theorem.)

S2. In the study of the continuous time dynamical system $\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t)$, the matrix $\mathbf{A}$ is called *stable* if all its eigenvalues fall in the *open* left half of the complex plane. For stable $\mathbf{A}$, the *distance to instability* is of significant interest: What is the smallest-norm perturbation $\mathbf{E}$ for which $\mathbf{A} + \mathbf{E}$ has at least one eigenvalue somewhere in the *closed* right half of the complex plane? Find a formula for this smallest $\|\mathbf{E}\|$, measured in the 2-norm (or any induced matrix norm).

Hint. Use the distance to singularity, along with the fact that the eigenvalues of a matrix are a continuous function of the matrix entries. The simplest answer will be stated in terms of the minimization of some quantity over a one-dimensional subspace of $\mathbb{C}$.