CAAM 453/553 · NUMERICAL ANALYSIS I

Examination 1

Posted 24 October 2009 (with minor update to Problem 5 on page 6).

Due no later than 5pm on Wednesday, 28 October 2009.

Instructions:

1. Time limit: **4 uninterrupted hours**.
2. There are five questions worth 25 points each (plus a 2-point bonus).
   453 students can choose any four problems; 553 students should attempt all five.
   Please do not look at the questions until you begin the exam.
3. Read the questions carefully. Sometimes useful facts are given between steps of a problem. You only need to produce work for those parts of the problem labeled (a), (b), (c), ...
4. You *may not* use any outside resources, such as books, notes, problem sets, friends, calculators, or MATLAB.
5. Please answer the questions thoroughly and justify all your answers.
   Show all your work to maximize partial credit.
6. Print your name on the line below:

   ____

7. Time started: ______________ Time completed: ______________
8. Indicate that this is your own individual effort in compliance with the instructions above and the honor system by writing out in full and signing the traditional pledge on the lines below.

   ____

   ____

   ____

9. Staple this page to the front of your exam.

1. [25 points: (a)=5 points; (b)=5 points; (c)=5 points; (d)=5 points; (e)=5 points]

(a) Compute the full singular value decomposition of

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

(You may find it slightly easier to compute the SVD of $\mathbf{A}^*$....)

Now let $\mathbf{A} \in \mathbb{C}^{m\times n}$ be some matrix with more columns than rows, $(m < n)$ and full rank, i.e., $\text{rank}(\mathbf{A}) = m$. In this case the *pseudoinverse* is defined as

$$\mathbf{A}^+ = \mathbf{A}^*(\mathbf{A}\mathbf{A}^*)^{-1}.$$

(b) Work out a formula for $\mathbf{A}^+$ in terms of the singular values and singular vectors of $\mathbf{A}$. (You may use the matrix form or dyadic form of the SVD, as you prefer.)

(c) Describe how the singular vectors of $\mathbf{A}$ relate to the four fundamental subspaces, $\text{Ran}(\mathbf{A})$, $\text{Ker}(\mathbf{A}^*)$, $\text{Ran}(\mathbf{A}^*)$, and $\text{Ker}(\mathbf{A})$, in this situation ($m < n$ and $\text{rank}(\mathbf{A}) = m$).

Now suppose we wish to solve the *underdetermined* linear system $\mathbf{A}\mathbf{x} = \mathbf{b}$, with full-rank $\mathbf{A} \in \mathbb{C}^{m\times n}$ for $m < n$, $\mathbf{x} \in \mathbb{C}^n$, and $\mathbf{b} \in \mathbb{C}^m$. This system has infinitely many solutions: Given any one $\mathbf{x}$ such that $\mathbf{A}\mathbf{x} = \mathbf{b}$, we also have $\mathbf{A}(\mathbf{x}+\mathbf{z}) = \mathbf{b}$ for any $\mathbf{z} \in \text{Ker}(\mathbf{A})$.

(d) Use the singular value decomposition to characterize *all* solutions $\mathbf{x} \in \mathbb{C}^n$ to $\mathbf{A}\mathbf{x} = \mathbf{b}$. Which of these solutions has smallest 2-norm?

(e) Applying the result of part (d) (or otherwise – if you missed part (d)), characterize all solutions $\mathbf{x} \in \mathbb{C}^3$ to $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{A}$ is the matrix in part (a) and

$$\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Specify the solution having smallest 2-norm.

2. [25 points: (a)=5 points; (b)=6 points; (c)=5 points; (d)=5 points; (e)=4 points]

   (a) Write down a formula for the Householder reflector matrix $\mathbf{H} \in \mathbb{R}^{n\times n}$ that reflects vectors across the hyperplane orthogonal to $\mathbf{v} \in \mathbb{R}^n$, for $\mathbf{v} \neq \mathbf{0}$.

   (b) What are the eigenvalues and eigenvectors ($\mathbf{Hx} = \lambda\mathbf{x}$) of the matrix $\mathbf{H}$ in part (a)? Give a geometric interpretation of the leftmost eigenvalue and associated eigenvector (e.g., using a sketch in two dimensions).

   (c) Recall that in two dimensions, $\mathbb{R}^2$, the Givens rotation matrix that rotates a vector by an angle θ in the (x_1, x_2) plane takes the form

$$\mathbf{G} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

   What are the eigenvalues and eigenvectors of $\mathbf{G}$?
   (Use your geometric intuition to guess the eigenvectors.)

   (d) Now consider three dimensions, $\mathbb{R}^3$. Write down the form of the following 3×3 Givens rotation matrices

      (i) $\mathbf{G}_{1,2}$, a rotation by an angle α in the (x_1, x_2) plane;
      (ii) $\mathbf{G}_{1,3}$, a rotation by an angle β in the (x_1, x_3) plane;
      (iii) $\mathbf{G}_{2,3}$, a rotation by an angle γ in the (x_2, x_3) plane.

Many applications in mechanics require coordinate transformations, affected by the the product of the three rotation in part (d):

$$\mathbf{R} = \mathbf{G}_{1,2}\mathbf{G}_{1,3}\mathbf{G}_{2,3}.$$

(Here α, β, γ are related to the *Euler angles*, i.e., *yaw*, *pitch*, and *roll*.)

   (e) Is this matrix $\mathbf{R}$ unitary? Give a geometric interpretation of the eigenvector associated with the eigenvalue $\lambda = 1$. (*Do not* explicitly form the matrix $\mathbf{R}$ entry-by-entry to answer this question!)

3. [25 points: (a)=4 points; (b)=4 points; (c)=4 points; (d)=9 points; (e)=4 points]

Given a set of points $\ldots, x_{-1}, x_0, x_1, \ldots \in \mathbb{R}$, recall the definition of the degree-k B-spline:

$$B_{j,k}(x) = \left(\frac{x - x_j}{x_{j+k} - x_j}\right) B_{j,k-1}(x) + \left(\frac{x_{j+k+1} - x}{x_{j+k+1} - x_{j+1}}\right) B_{j+1,k-1}(x),$$

where

$$B_{j,0}(x) = \begin{cases} 1, & x \in [x_j, x_{j+1}); \\ 0, & \text{otherwise.} \end{cases}$$

(a) For what values of $x \in \mathbb{R}$ does $B_{j,k}(x) \neq 0$? (This is the *support* of $B_{j,k}$.)

(b) Let $C^p(\mathbb{R})$ denote the set of functions on $\mathbb{R}$ that are continuous and whose first p derivatives are also continuous. What is the largest integer p for which $B_{j,k} \in C^p(\mathbb{R})$, as a function of the degree k?

Suppose we wish to use cubic splines to *approximate* a function over the interval, $[0, 1]$. In particular, we wish to approximate f using some linear combination of B-splines:

$$S(x) = \sum_{j=?}^{?} c_j B_{j,3}(x). \qquad (*)$$

To be concrete, suppose we have $n = 4$ and the spline knots

$$x_j = \frac{j}{n}.$$

We want S to approximate f at the $m = 2n + 1 = 9$ approximation points

$$\xi_j = \frac{j}{2n}, \quad j = 0, \ldots, 2n.$$

That is, we want

$$S(\xi_j) \approx f(\xi_j), \qquad j = 0, \ldots, 2n.$$

(c) What should the limits on the sum $(*)$ be, so that this sum incorporates only those cubic B-splines $B_{j,3}$ that are nonzero for some $x \in [0, 1]$ with $n = 4$? (A picture might help.)

(d) Explain how to set up a discrete least squares problem to determine the coefficients c_j in the $n = 4$ case; that is, specify $\mathbf{A}$, $\mathbf{c}$, $\mathbf{f}$ in the minimization problem

$$\min_{\mathbf{c} \in \mathbb{R}^?} \|\mathbf{A}\mathbf{c} - \mathbf{f}\|_2.$$

Please clearly state the dimensions and entries of $\mathbf{A}$, $\mathbf{c}$, $\mathbf{f}$.

(e) How many nonzero entries will there be in the rows of $\mathbf{A}$? Would it be preferable to use Householder reflectors or Givens rotations to compute the $\mathbf{QR}$ factorization of $\mathbf{A}$ (particularly for larger n)? Briefly explain (no need for operation counting).

[from C. Beattie]

4. [25 points: (a)=5 points; (b)=5 points; (c)=6 points; (d)=9 points]

(a) Suppose $f \in C^{n+1}[a,b]$ is interpolated at distinct $x_0, \ldots, x_n \in [a,b]$ by the degree-n polynomial p. Write down the formula for the error $f(x) - p(x)$ at a given $x \in [a,b]$.

Let $p \in \mathcal{P}_1$ denote the linear polynomial interpolant to f at x_0 and x_1, i.e., $p(x_0) = f(x_0)$ and $p(x_1) = f(x_1)$. Now, suppose we only know *approximations* f_0 and f_1 to $f(x_0)$ and $f(x_1)$ that are accurate within a factor of δ:

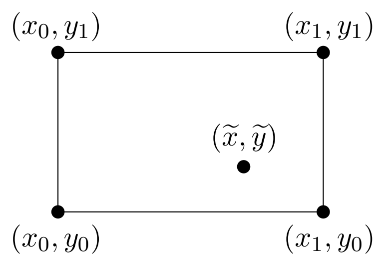
$$|f(x_0) - f_0| \le \delta, \qquad |f(x_1) - f_1| \le \delta.$$

Let $\widehat{p} \in \mathcal{P}_1$ denote the linear polynomial that interpolates f_0 and f_1 at x_0 and x_1.

(b) Give a simple bound for $\displaystyle\max_{x\in[x_0,x_1]} |p(x) - \widehat{p}(x)|$. (Hint: draw a picture.)

(c) Use your answers from parts (a) and (b) to develop a bound for $\displaystyle\max_{x\in[x_0,x_1]} |f(x) - \widehat{p}(x)|$.

We shall now use these results to derive a bound on the error present in two-dimensional interpolation. In particular, suppose we have a function $f(x,y)$ whose value is only known at the four distinct points (x_0, y_0), (x_0, y_1), (x_1, y_0), (x_1, y_1).



We shall estimate the value of $f(\widetilde{x}, \widetilde{y})$ for some $\widetilde{x} \in (x_0, x_1)$ and $\widetilde{y} \in (y_0, y_1)$ by the following procedure. (The steps that follow are merely descriptive: you do not have any work to do until part (d) below.)

- Construct the polynomial $p \in \mathcal{P}_1$ such that $p(x_0) = f(x_0, y_0)$ and $p(x_1) = f(x_1, y_0)$. (This will give $p(x) \approx f(x, y_0)$.)
- Construct the polynomial $q \in \mathcal{P}_1$ such that $q(x_0) = f(x_0, y_1)$ and $q(x_1) = f(x_1, y_1)$. (This will give $q(x) \approx f(x, y_1)$.)
- Construct the polynomial $r \in \mathcal{P}_1$ such that $r(y_0) = p(\widetilde{x})$ and $r(y_1) = q(\widetilde{x})$. (This will give $r(y) \approx f(\widetilde{x}, y)$.)

This final step implies that $r(\widetilde{y}) \approx f(\widetilde{x}, \widetilde{y})$, the value we aim to approximate.

(d) Use your results from parts (a) and (c) to derive an error bound for

$$|r(\widetilde{y}) - f(\widetilde{x}, \widetilde{y})|$$

involving various partial derivatives of f and the interpolation points. (Assume that f is sufficiently smooth to ensure existence and continuity of any derivatives you need.)

5. [25 points: (a)=6 points; (b)=6 points; (c)=6 points; (d)=7 points]

(a) Given an inner product on $C[a,b]$ and some nonzero function $\phi \in C[a,b]$, we define the operator P that maps $f \in C[a,b]$ according to the rule

$$Pf = \frac{\langle f, \phi\rangle}{\langle \phi, \phi\rangle}\phi.$$

Show that P is a projector onto $\operatorname{span}\{\phi\}$; that is, show $P^2 f = Pf$ and $Pf \in \operatorname{span}\{\phi\}$ for all $f \in C[a,b]$. (In fact, P is an *orthogonal* projector, but you need not show that.)

(b) Suppose we have linearly independent functions $p_0, \ldots, p_n \in C[a,b]$. Use part (a) to write the Gram–Schmidt process that will generate an orthogonal basis for $\operatorname{span}\{p_0, \ldots, p_n\}$.

In class we developed a method for constructing polynomial approximations to a function that minimized the L^2-norm of the error. Now we investigate optimal *trigonometric polynomial* approximations to a continuous 2π-periodic function f over $[0, 2\pi]$. As with trigonometric interpolation, we approximate from the space

$$\mathcal{T}_n = \operatorname{span}\{e^{i0x}, e^{i1x}, e^{-i1x}, \ldots, e^{inx}, e^{-inx}\}.$$

As these basis functions are complex-valued, we need a complex-valued inner product:

$$\langle f, g\rangle = \int_0^{2\pi} f(x)\overline{g(x)}\,dx,$$

where $\overline{g(x)}$ denotes the complex-conjugate of $g(x)$. This inner product induces a norm:

$$\|g\|_{L^2} = \sqrt{\langle g, g\rangle}.$$

(c) Show that any two distinct basis functions, e^{ikx} and $e^{i\ell x}$ with $k \neq \ell$, are orthogonal in the complex inner product specified above. (See the 'useful facts' on the next page, and an alternative if you prefer to stay in real arithmetic.)

A polynomial $t_n \in \mathcal{T}_n$ is the optimal L^2-approximation to f from $\mathcal{T}_n$ if and only if the error is orthogonal to the approximating subspace:

$$\langle f - t_n, e^{ikx}\rangle = 0, \qquad \text{for all } k = -n, \ldots, n.$$

(Equivalently, t_n comes from a Gram–Schmidt step that orthogonalizes f against $\{e^{-inx}, \ldots, e^{inx}\}$.)

(d) Use the orthogonality of the error to find a formula for the optimal L^2-approximation

$$t_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}.$$

That is, produce a formula for the optimal values of c_k, $k = -n, \ldots, n$.

Useful facts for Problem 5 (which you can use without proof).

- $e^{i\theta} = \cos(\theta) + i\sin(\theta)$ for all $\theta \in \mathbb{R}$;
- $e^{i\theta}e^{i\phi} = e^{i(\theta+\phi)}$ for all $\theta, \phi \in \mathbb{R}$;
- $\overline{e^{i\theta}} = e^{-i\theta}$ for all $\theta \in \mathbb{R}$.

If you are entirely uncomfortable with complex analysis, you can replace the complex exponential basis with the equivalent form $\mathcal{T}_n = \text{span}\{1, \sin(x), \cos(x), \ldots, \sin(nx), \cos(nx)\}$. In this case, you can change parts (c) and (d) in the following manner.

- For part (c), show that the functions 1, $\sin(x)$, $\cos(x)$, $\ldots$, $\sin(nx)$, $\cos(nx)$ are orthogonal with respect to the usual L^2 inner product

  $$\langle f, g\rangle = \int_0^{2\pi} f(x)g(x)\,dx.$$

- For part (d), write

  $$t_n(x) = c_0 + \sum_{k=1}^{n} c_k \sin(kx) + \sum_{k=1}^{n} c_{-k}\cos(kx)$$

  and determine the c_k values from the fact that t_n is the optimal L^2 approximation to f from $\mathcal{T}_n$ if and only if

  $$\langle f - t_n, 1\rangle = 0, \qquad \langle f - t_n, \sin(kx)\rangle = 0, \qquad \langle f - t_n, \cos(kx)\rangle = 0$$

  for all $k = 1, \ldots, n$.

bonus. [2 points]

Suppose the vector $\mathbf{x} \in \mathbb{R}^5$ has, as its jth entry, your score on problem j of this exam. Consider three different ways at arriving at a total score on the exam: $\|\mathbf{x}\|_1$, $\|\mathbf{x}\|_2$, and $\|\mathbf{x}\|_\infty$.

Which measure is most fair? Explain.