Take-Home Midterm Key **CSE/Math 441** **Fall 2009**

Instructions: Point values of problems are given in parentheses. This exam is worth 130 points. Note the following:

- Papers must be written neatly and legibly. Your solutions must be readable, i.e., consist of narratives with sentences and connectives between equations.
- Show your work. Solutions should be complete. Do not expect me to run any .m files to obtain your answers, although you must include such files or diaries as evidence of your work. Get Matlab help on "diary" if you're still unclear on this. When finished, you can edit your diary file with any text editor. Eliminate unnecessary output.
- All work is to be done by individuals, with no collaboration or consultation of any kind with anyone else other than me. If you use any outside sources other than the text or notes, they should be clearly cited. If you have questions about any problems, contact me by email ASAP. I will post my comments on our Blackboard discussion board if this is appropriate.
- Exam files must handed in at class time on Thursday, October 22.

(15) **1.** What point of the plane $3x + 2y + z - 6 = 0$ in $\mathbb{R}^3$ is closest to the origin when measured by the (a) 1-norm, (b) 2-norm, and (c) ∞-norm?

SOLUTION. We note that this plane has intercepts on the positive axes and that expanding balls centered at the origin of any of these three norms will first touch the plane in the first octant.

For the 1-norm, note that the face in the first octant expands with vertices on the axes, and the (Euclidean) distance from the origin to the vertex on any axis gives the radius of the 1-norm ball. Thus, a ball first contacts the plane at the nearest intercept, which is along the x-axis at $x = 2$. Hence the point nearest the origin in this norm is $(2, 0, 0)$.

For the 2-norm, just use calculus III vector arithmetic: A position vector from the origin to the point nearest the origin on a plane with outward pointing normal vector $\mathbf{n}$ and point P in the plane, is just the projection of the vector $\overrightarrow{OP}$ along $\mathbf{n}$. In this case $\mathbf{n}=\langle 3, 2, 1\rangle$ and we can take $P = (2, 0, 0)$ to obtain the vector

$$\frac{\overrightarrow{OP}\cdot\mathbf{n}}{\mathbf{n}\cdot\mathbf{n}}\mathbf{n} = \frac{6}{14}\langle 3, 2, 1\rangle = \frac{3}{7}\langle 3, 2, 1\rangle,$$

so that the closest point is $\left(\frac{9}{7}, \frac{6}{7}, \frac{3}{7}\right)$.

For the ∞-norm, balls of radius $\sqrt{2}t$ expand with outward vertex at (t, t, t) for positive t. Now the point $(t, t, t)$is on the plane when $6t = 6$, that is, $t = 1$. Hence the point nearest the origin in the ∞-norm is $(1, 1, 1)$.

(24) **2.** Determine the best approximation $p(x)$ to $f(x) = x^3$ from the subspace $\mathcal{P}_1$ of $C[-1, 1]$. Then compute the 1-norm, 2-norm and ∞-norm of the error function $e(x) = f(x) - p(x)$.

SOLUTION. We know that this polynomial must be of the form $p(x) = ax$, for some $0 < a < 1$, since it must be skew-symmetric or it does better on one side than the other. Furthermore, we can use Chebyshev polynomials to best approximate monomials. Note that the third Chebyshev polynomial is

$$\begin{aligned} T_3(x) &= 2xT_2(x) - T_1(x) \\ &= 2x(2xT_1(x) - T_0) - T_1(x) \\ &= 2x\left(2x^2 - 1\right) - x \\ &= 4x^3 - 3x. \end{aligned}$$

Hence the best approximation to x^3 from $\mathcal{P}_2$ in the max norm is $\frac{3}{4}x$. However, this polynomial is linear, so it is the best approximation from $\mathcal{P}_1$ as well. Thus

$$\left\| x^3 - \frac{3}{4}x \right\|_\infty = \frac{1}{4}.$$

For the other norms, symmetry shows we only need integrate from 0 to 1 and double the result. Now $e(x)$ changes sign at $x = \sqrt{3}/2$, so we have

$$\|e\|_1 = 2\int_0^{\sqrt{3}/2} \left(\frac{3}{4}x - x^3\right) dx + 2\int_{\sqrt{3}/2}^1 \left(x^3 - \frac{3}{4}x\right) dx = \frac{20}{64} = \frac{5}{16}$$

Also

$$\frac{1}{4}\|e\|_2^2 = \int_0^1 \left(\frac{3}{4}x - x^3\right)^2 dx = \left(\frac{3x^3}{16} - \frac{3x^5}{10} + \frac{x^7}{7}\right)\Big|_{x=0}^1 = \frac{3}{16} - \frac{3}{10} + \frac{1}{7} = \frac{17}{560}.$$

(28) **3.** Use a divided difference table to derive a formula (as we did in class for Hermite cubics) for a quadric polynomial that interpolates $f(a)$, $f'(a)$, $f''(a)$, $f(b)$ and $f'(b)$ for a given $f \in C[a,b]$. Code this up as a Matlab function with appropriate inputs and test on $f(x) = (x+1)^4 \in C[0,1]$.

SOLUTION. Here is the difference table for a Hermite quartic polynomial interpolating $f(x)$ at a, a, a, b, b. We use the formula of text (5.27) that says $f[x_j, x_{j+1}, \ldots, x_{j+k}] = f^{(k)}(x_j)/k!$ for $x_{j+i} = x_j$.

x_j	$f[x_j]$	$f[x_j, x_{j+1}]$	$f[x_j, x_{j+1}, x_{j+2}]$	$f[x_j, x_{j+1}, x_{j+2}, x_{j+3}]$	$f[x_j, x_{j+1}, x_{j+2}, x$
a	$f(a)$				
		$f'(a)$			
a	$f(a)$		$\frac{f''(a)}{2}$		
		$f'(a)$		$\frac{f[a,b] - f'(a) - (b-a)f''(a)/2}{(b-a)^2}$	
a	$f(a)$		$\frac{f[a,b] - f'(a)}{b-a}$		$f'(b) + 2f'(a) - 3f[a,b]$ over $(b-a)^3$
		$f[a,b]$		$\frac{f'(b) - 2f[a,b] + f'(a)}{(b-a)^2}$	
b	$f(b)$		$\frac{f'(b) - f[a,b]}{b-a}$		
		$f'(b)$			
b	$f(b)$				

Thus, the quartic looks like

$$\begin{aligned} p(x) &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f[a,b] - f'(a) - (b-a) f''(a)}{(b-a)^2}(x-a)^3 \\ &\quad + \frac{f'(b) + 2f'(a) - 3f[a,b] + (b-a) f''(a)}{(b-a)^3}(x-a)^3 (x-b) \end{aligned}$$

If we code this up in Matlab, we need to write the term $(x-b)$ as

$$x - b = (x-a) - (b-a)$$

so that the fourth degree coefficient contributes to both the third and fourth powers of $(x-a)$. Here is the function:

```
  function retval = midprob3(fa,fpa,fppa,fb,fpb,a,b)
  % usage:  poly = midprob3(fa,fpa,fppa,fb,fpb,a,b)
  % description:  returns cofficients of a quartic % polynomial in (x-a) that
interpolates f(x),
  % f’(x), f”(x) at x=a and f(x), f’(x) at x=b.
  retval = zeros(1,5);
  bma = b - a;
  fab = (fb-fa)/bma;
  retval(5) = fa;
  retval(4) = fpa;
  retval(3) = fppa/2;
  retval(2) = (fab-fpa-bma*fppa/2)/bma^2;
  retval(1) = (fpb+2*fpa-3*fab+bma*fppa/2)/bma^3;
  % correct the coefficient of (x-a)^3
  retval(2) = retval(2)-bma*retval(1);
```

Now test this with

$$\begin{aligned} f(x) &= (1+x)^4 = x^4 + 4x^3 + 6x^2 + 4x + 1 \\ f'(x) &= 4(1+x)^3 \\ f''(x) &= 12(1+x)^2 \end{aligned}$$

so that $a = 0$, $b = 1$, $f(a) = 1$, $f'(a) = 4$, $f''(a) = 12$, $f(b) = 16$, and $f'(b) = 32$.

The result is

```
  octave:1> poly = midprob3(1,4,12,16,32,0,1)

  poly =
     1 4 6 4 1
```

(28) **4.** Suppose we want to approximate $f(x) = \sin x$ for all x.

(a) Use the Matlab file PPfcn.m to compute a Hermite p.p. approximation to the function of (a) with knots at 0, $\pi/4$, $\pi/2$ and compute the infinity norm of the error.

(b) Compare this error the the error of a best minimax polynomial (use minimaxd.m) whose evaluation requires the same (or nearly so) number of flops as in (a). Who wins?

SOLUTION. (a) Here is the Octave transcript for the Hermite p.p. approximation:

```
octave:1> PPfcns
octave:2> xknots = [0,pi/4,pi/2];
octave:3> f = sin(xknots);
octave:4> fp = cos(xknots);
octave:5> pp = HCpp(xknots,f,fp) pp =
{
  order = 4
  knum = 3
  knots =
    0.00000 0.78540 1.57080
  coefs =
    -0.15162 -0.00784 1.00000 0.00000
    -0.06280 -0.37617 0.70711 0.70711
}
octave:6> x = 0:.001:pi/2;
octave:7> norm(sin(x)-PPeval(pp,x),inf)
ans = 9.0622e-04
```

(b) The cost of evaluating a cubic p.p. with only two knots is the cost of evaluating a cubic plus the cost of searching for the right interval, which will only involve one extra flop. So a comparable polynomial would be a fourth degree at best. Here is the Matlab construction of the polynomial, and the level error will give the norm of the error of approximation:

```
octave:8> [poly,levelerr] = minimaxd(sin(x),x,4)
poly =
  2.8409e-02 -2.0264e-01 1.9058e-02 9.9643e-01 1.0750e-04
levelerr = 1.0750e-04
```

It looks like the fourth degree polynomial is the winner here.

(20) **5.** Let $\mathcal{A}$ be the set of multiples of the function $g(x) = x$, $0 \leq x \leq 1$, and the operator $X : C\,[0,1] \to \mathcal{A}$ be given by

$$(Xf)(x) = 2\int_0^1 xf\,(y)\,dy.$$

where $f \in C[0,1]$, a normed space with the infinity norm. Show that X is a linear projection operator (see pp. 22-23 of text) and estimate its norm.

Solution. For linearity, we suppose that $f, g \in C[0,1]$ and that c, d are arbitrary constants. Calculate

$$\begin{aligned}(X(cf+dg)(x) &= 2\int_0^1 x\left(cf(y)+dg(y)\right)dy \\ &= 2c\int_0^1 xf(y)\,dy + 2d\int_0^1 xg(y)\,dy \\ &= c(Xf)(x)+d(Xg)(x),\end{aligned}$$

where the second equation follows from the linearity of definite integrals. Next, suppose that $f(x) = cx$ is an arbitrary element of $\mathcal{A}$. We need to show that $Xf = f$. Calculate

$$(Xf)(x) = 2\int_0^1 x(cy)\,dy = 2x\frac{cy^2}{2}|_{y=0}^1 = x(c\cdot 1 - c\cdot 0) = cx = f(x).$$

Since this is true for all $x \in [0,1]$, we conclude that $Xf = f$, so that X is a projection operator.

To estimate the norm of X with the infinity norm on functions, note that for $\|f\|_\infty \leq 1$ we have $|f(x)| \leq 1$, for all $0 \leq x \leq 1$, so that

$$|(Xf)(x)| = \left|2\int_0^1 xf(y)\,dy\right| \leq 2\int_0^1 x|f(y)|\,dy \leq 2x\int_0^1 dy \leq 2\cdot 1 = 2.$$

so our estimate is

$$\|X\|_\infty \leq 2.$$

In fact, if we take $f(x) = 1$ and we obtain that $\|Xf\|_\infty = 2$, so the inequality above is an equality.

(15) **6.** Use any of the key facts from Chapter 7 to explain why the levelled reference error is guaranteed to increase from step to step if the current reference is not optimal (in the exchange algorithm).

Solution. The essential fact is Theorem 7.7, which asserts that if $p(x) \in \mathcal{P}_n$ and $\{x_i\}_{i=0}^n$ is a reference on which the differences $(f(x_i) - p(x_i))$ oscillate in sign, then we have the inequality

$$\min_{0\leq i\leq n+1} |f(x_i) - p(x_i)| \leq \min_{q\in\mathcal{P}_n} \max_{0\leq i\leq n+1} |f(x_i) - q(x_i)|,$$

where the inequality is strict if the reference is not levelled for $p(x)$. Now suppose that we have a non-optimal reference. Then we cannot have that all $|f(x_i) - p(x_i)| = \|f-p\|_\infty$,for otherwise $p(x)$ would be optimal by Theorem 7.2 (Characterization Theorem). Therefore, the exchange algorithm has us replace one of the reference points by an adjacent one x_i^* for which $|f(x_i^*) - p(x_i^*)| > |f(x_i) - p(x_i)|$ and sign alternation is preserved. Then the new reference set is not level for $p(x)$. Next, we level the new reference by finding an appropriate new polynomial $q(x) \in \mathcal{P}_n$ for which the new reference set is level. Therefore, the max of the above inequality on the right for this particular q is exactly the common value $|f(x_i) - p(x_i)|$, which is at least as large as the right-hand side minimum, which is in turn strictly larger than the left-hand side. This the new level reference error is larger than the old.