Problem 1

Who are your group members?

Solution. Nicholas Rees

Problem 2

Consider a monomial interpolation $p(x) = c_0 + c_1x + c_2x^2$ to data points $(x_0, y_0), (x_1, y_1), (x_2, y_2)$, where $x_0 = 2$, $x_1 = 2 + \varepsilon$, and $x_2 = 2 - \varepsilon$ (but y_0, y_1, y_2 are arbitrary). Hence we are solving the equations:

$$\begin{bmatrix} 1 & 2 & 4 \\ 1 & 2+\varepsilon & 4+4\varepsilon+\varepsilon^2 \\ 1 & 2-\varepsilon & 4-4\varepsilon+\varepsilon^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}$$

(a). Show that c_2, in terms of y_0, y_1, y_2, is given by

$$c_2(\varepsilon) = \frac{y_1 + y_2 - 2y_0}{2\varepsilon^2} \tag{1}$$

(b). Now assume that for some twice differentiable f we have $y_i = f(x_i)$, hence

$$y_0 = f(2),\ y_1 = f(2+\varepsilon),\ y_2 = f(2-\varepsilon)$$

Use L'Hôpital's Rule or Taylor's Theorem to show that

$$\lim_{\varepsilon\to 0} c_2(\varepsilon) = f''(2)/2$$

(if you use Taylor's Theorem, for simplicity assume that f''' exists and is bounded near 2). [Hint: See Section 1.4 of the handout for a similar example.]

(c). How is your formula for c_2 related to the *centered formula for the second derivative*, page 412, Subsection 14.1.4, of [A&G]?

(a). *Solution.* One can find that the inverse of the matrix A is

$$A^{-1} = \frac{1}{2\varepsilon^2} \begin{bmatrix} -2(4-\varepsilon^2) & 2(2-\varepsilon) & 2(\varepsilon+2) \\ 8 & \varepsilon-4 & -\varepsilon-4 \\ -2 & 1 & 1 \end{bmatrix}$$

We can confirm this by multiplying it by A: $AA^{-1} =$

$$\frac{1}{2\varepsilon^2} \begin{bmatrix} -2(4-\varepsilon^2)+16-8 & 2(2-\varepsilon)+2(\varepsilon-4)+4 & 2(\varepsilon+2)-2(\varepsilon+4)+4 \\ -2(4-\varepsilon^2)+8(2+\varepsilon)-2(2+\varepsilon)^2 & 2(2-\varepsilon)+(2+\varepsilon)(\varepsilon-4)+(2+\varepsilon)^2 & 2(\varepsilon+2)-(2+\varepsilon)(\varepsilon+4)+(2+\varepsilon)^2 \\ -2(4-\varepsilon^2)+8(2-\varepsilon)-2(2-\varepsilon)^2 & 2(2-\varepsilon)+(2-\varepsilon)(\varepsilon-4)+(2-\varepsilon)^2 & 2(\varepsilon+2)-(2-\varepsilon)(\varepsilon+4)+(2-\varepsilon)^2 \end{bmatrix}$$

$$= \frac{1}{2\varepsilon^2} \begin{bmatrix} 2\varepsilon^2 & 0 & 0 \\ 0 & 2\varepsilon^2 & 0 \\ 0 & 0 & 2\varepsilon^2 \end{bmatrix}$$

$$= I$$

and inverses are unique for invertible matrices, so we need not check $A^{-1}A$. Hence, we can directly compute c_2 from

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = A^{-1} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}$$

Specifically, taking the bottom row of A^{-1} gives

$$c_2(\varepsilon) = \frac{1}{2\varepsilon^2}(-2y_0 + y_1 + y_2)$$

as desired.

(b). *Solution.* We have

$$\begin{aligned}\lim_{\varepsilon\to 0} c_2(\varepsilon) &= \lim_{\varepsilon\to 0}\frac{y_1+y_2-2y_0}{2\varepsilon^2}\\ &= \lim_{\varepsilon\to 0}\frac{f(2+\varepsilon)+f(2-\varepsilon)-2f(2)}{2\varepsilon^2}\\ &= \lim_{\varepsilon\to 0}\frac{f'(2+\varepsilon)-f'(2-\varepsilon)}{4\varepsilon}\\ &= \lim_{\varepsilon\to 0}\frac{f''(2+\varepsilon)+f''(2-\varepsilon)}{4}\\ &= \frac{1}{2}f''(2)\end{aligned}$$

where L'ôpital's can be applied twice here, since f is twice differentiable, and for the first use, $\lim_{\varepsilon\to 0} f(2+\varepsilon)+f(2-\varepsilon)-2f(2) = f(2)+f(2)-2f(2)$ and $\lim_{\varepsilon\to 0} 2\varepsilon^2 = 0$, and for the second use, $\lim_{\varepsilon\to 0} f'(2+\varepsilon)-f'(2-\varepsilon) = f'(2)-f'(2) = 0$ and $\lim_{\varepsilon\to 0} 4\varepsilon = 0$.

(c). *Solution.* The centered formula for the second derivative via the textbook is

$$f''(x_0) = \frac{1}{h^2}(f(x_0-h)-2f(x_0)+f(x_0+h)) + O(h^2)$$

We have just shown

$$f''(2) = 2\lim_{\varepsilon\to 0} c_2(\varepsilon) = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon^2}(f(2-\varepsilon)-2f(2)+f(2+\varepsilon))$$

Hence, we have confirmed the centered formula at $x_0 = 2$ as h gets large and so the $O(h^2)$ term becomes negligible; see this by making the substitutions $x_0 = 2$ and $h = \varepsilon$.

Problem 3

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The point of this exercise is to carefully prove that

$$\|A\|_\infty = \max(|a|+|b|, |c|+|d|)$$

Notice that for any $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$ we have

$$A\mathbf{x} = A\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ax_1+bx_2 \\ cx_1+dx_2 \end{bmatrix}$$

(a). Show that if $m = \|\mathbf{x}\|_\infty = \max(|x_1|, |x_2|)$, then

$$|ax_1+bx_2| \le m(|a|+|b|)$$

(b). Using part (a), and the same with a, b replaced with c, d show that

$$\|A\mathbf{x}\|_\infty \le \max(|a|+|b|, |c|+|d|)\|\mathbf{x}\|_\infty$$

(c). Show that there is an $\mathbf{x}$ with $\|\mathbf{x}\|_\infty = 1$ such that

$$\|A\mathbf{x}\|_\infty \ge |a|+|b|$$

[Hint: take $\mathbf{x} = (\pm 1, \pm 1)$ with appropriately chosen signs.]

(d). Conclude from all the above (and perhaps replacing a, b with c, d somewhere) that

$$\|A\|_\infty = \max(|a|+|b|, |c|+|d|)$$

(a). *Solution.* Using triangle inequality, we have

$$|ax_1 + bx_2| \leq |ax_1| + |bx_2| \leq |a||x_1| + |b||x_2| \leq |a|m + |b|m \leq m(|a| + |b|)$$

as desired.

(b). *Solution.* We have

$$\|A\mathbf{x}\|_\infty = \max(|ax_1 + bx_2|, |cx_1 + dx_2|) \leq \max(\|\mathbf{x}\|_\infty(|a| + |b|), \|\mathbf{x}\|_\infty(|c| + |d|)) = \max(|a| + |b|, |c| + |d|)\|\mathbf{x}\|_\infty$$

where the first inequality follows from using the fact that $y \leq w \implies \max(y, z) \leq \max(w, z)$ twice, since $|ax_1 + bx_2| \leq \|\mathbf{x}\|_\infty(|a| + |b|)$ and $|cx_1 + dx_2| \leq \|\mathbf{x}\|_\infty(|c| + |d|)$ by part (a); and the last equality from using the fact $\max(ky, kz) = k \max(y, z)$ where k is some constant nonnegative constant, since if $y \geq z$ then $ky \geq kz$ so $\max(ky, kz) = ky = k \max(y, z)$, and vice versa when $z \geq y$, and $\|\mathbf{x}\|_\infty \geq 0$ by definition.

(c). *Solution.* Let $\mathbf{x} = \begin{bmatrix} |a|/a \\ |b|/b \end{bmatrix}$. Clearly, $\|\mathbf{x}\|_\infty = \max(||a|/a|, ||b|/b|) = \max(1, 1) = 1$. Furthermore,

$$\|A\mathbf{x}\|_\infty = \max(a\frac{|a|}{a} + b\frac{|b|}{b}, c\frac{|a|}{a} + d\frac{|b|}{b}) \geq a\frac{|a|}{a} + b\frac{|b|}{b} = |a| + |b|$$

as desired.

(d). *Solution.* Recall the definition of $\|A\|_\infty$: it is the smallest real $C > 0$ such that $\|A\mathbf{x}\|_\infty \leq C\|\mathbf{x}\|_\infty$ for all $\mathbf{x} \in \mathbb{R}^2$.

If $C = \max(|a| + |b|, |c| + |d|)$, we have already shown that the inequality is true for arbitrary $\mathbf{x}$ in part (b) of this problem. It remains to show that this is the smallest value.

Now, for the sake of contradiction, assume that there is some C where $0 < C < \max(|a| + |b|, |c| + |d|)$ and $\|A\mathbf{x}\|_\infty \leq C\|\mathbf{x}\|_\infty$ for all $\mathbf{x} \in \mathbb{R}^2$. If $\max(|a| + |b|, |c| + |d|) = |a| + |b|$, then we have that there exists some $\mathbf{x}$ where $\|\mathbf{x}\|_\infty = 1$ and $\|A\mathbf{x}\|_\infty \geq |a| + |b|$ by part (c) of this problem, so

$$C\|\mathbf{x}\|_\infty = C \geq \|A\mathbf{x}\|_\infty \geq |a| + |b| = \max(|a| + |b|, |c| + |d|) > C$$

which is a contradiction, since $C < C$ is impossible. If $\max(|a| + |b|, |c| + |d|) = |c| + |d|$, using an identical proof that what was done in part (c) (except with $\mathbf{x} = \begin{bmatrix} |c|/c \\ |d|/d \end{bmatrix}$) we have that there is some $\mathbf{x}$ where $\|\mathbf{x}\|_\infty = 1$ and $\|A\mathbf{x}\|_\infty \geq |c| + |d|$, so

$$C\|\mathbf{x}\|_\infty = C \geq \|A\mathbf{x}\|_\infty \geq |c| + |d| = \max(|a| + |b|, |c| + |d|) > C$$

which is again a contradiction, since we cannot have $C > C$. Hence, regardless of the matrix A, we get a contradiction.

Hence, we have proven that $C = \max(|a| + |b|, |c| + |d|)$ is minimal, therefore $\|A\|_\infty = \max(|a| + |b|, |c| + |d|)$.

Problem 4

Let $1 \leq p, q \leq \infty$ satisfy $(1/p) + (1/q) = 1$ (so $(p, q) = (2, 2)$ is one possibility, but we also allow (p, q) equal to $(1, \infty)$ and $(\infty, 1)$). Then it is known that for any $m \times n$ matrix A (with real entries) we have

$$\|A^T\|_p = \|A\|_q \tag{2}$$

where A^T is the transpose of A. Using this fact, and the previous exercise, for

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

find a formula for $\|A\|_1$. [Hint: it should match the formula in Section 6 of the handout.] [This formula for $\|A\|_1$ is not too hard to prove from scratch, but it is probably easier to use the previous exercise and (2).]

Solution. Recall that $A^T = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Since we are allowing $p = \infty, q = 1$ for equation (2), we have

$$\|A\|_1 = \|A^T\|_\infty = \left\| \begin{bmatrix} a & c \\ b & d \end{bmatrix} \right\|_\infty = \max(|a| + |c|, |b| + |d|)$$

where the last equality was due to question 2. Hence, we have proven the formula, $\|A\|_1 = \max(|a| + |c|, |b| + |d|)$, which matches the formula from Section 6.

Problem 5

There is a standard formula for the determinant of a Vandermonde matrix: namely, if

$$X = \begin{bmatrix} 1 & x_0 & x_0^2 & \cdots & x_0^n \\ 1 & x_1 & x_1^2 & \cdots & x_1^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^n \end{bmatrix}$$

then

$$\det(X) = \prod_{0 \le i < j \le n} (x_j - x_i)$$

(This formula is not hard to prove by induction on n, if you note that replacing x_n by a variable x, then the above determinant is a degree n polynomial in x with roots $x_0, \ldots, x_{n-1}$.) This implies that if $x_0, \ldots, x_n$ are distinct, then X is invertible, and a standard formula for X^{-1} (the formula is $X^{-1} = \det(X)\text{adjugate}(X)$, where the adjugate matrix is formed by X's *cofactors*, i.e., determinants of submatrices of X where a single row and a single column are discarded) then implies that the bottom right entry of X^{-1} is

$$(X^{-1})_{n+1,n+1} = \prod_{0 \le i \le n-1} \frac{1}{x_n - x_i} \tag{3}$$

(and similarly, up to $\pm$, for all entries of the bottom row of X^{-1}). Consider the special case

$$A(\varepsilon) = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2+\varepsilon & 4+4\varepsilon+\varepsilon^2 \\ 1 & 2-\varepsilon & 4-4\varepsilon+\varepsilon^2 \end{bmatrix}$$

(a). Use (3) to show that for all $|\varepsilon| < 1$,

$$\|A^{-1}(\varepsilon)\|_\infty \ge 1/(2\varepsilon^2);$$

you may use the analog of Problem 3 for 3×3 matrices and/or the fact from class and the handout (page 12) that if M is the maximum absolute value of an entry of an $n \times n$ matrix B, then $M \le \|B\|_p \le nM$ for any $1 \le p \le \infty$.

(b). Use the formula (1) to determine the entire bottom row of $A^{-1}(\varepsilon)$, and hence double check the formula (3) in this case.

(c). Show that for some constant, $C > 0$, we have for all $|\varepsilon| < 1$,

$$\|A(\varepsilon)\|_\infty \|A^{-1}(\varepsilon)\|_\infty \ge C/\varepsilon^2$$

(a). *Solution.* The analog of Problem 3 for 3×3 gives

$$\begin{aligned} \|A^{-1}(\varepsilon)\|_\infty = \max(&|(A^{-1})_{1,1}| + |(A^{-1})_{1,2}| + |(A^{-1})_{1,3}|, \\ &|(A^{-1})_{2,1}| + |(A^{-1})_{2,2}| + |(A^{-1})_{2,3}|, \\ &|(A^{-1})_{3,1}| + |(A^{-1})_{3,2}| + |(A^{-1})_{3,3}|) \end{aligned}$$

And so $\|A^{-1}(\varepsilon)\|_\infty \geq |(A^{-1})_{3,1}| + |(A^{-1})_{3,2}| + |(A^{-1})_{3,3}| \geq |(A^{-1})_{3,3}| \geq (A^{-1})_{3,3}$ since this is just the sum of positive elements. Thus, assuming $|\varepsilon| > 0$ (I don't see a way around this fact) ensures that x_0, x_1, x_2 are distint so we can use (3) to get

$$\|A^{-1}(\varepsilon)\|_\infty \geq \prod_{0\leq i\leq 1} \frac{1}{x_2 - x_i} = \frac{1}{(x_2-x_1)(x_2-x_0)} = ((2-\varepsilon-2-\varepsilon)(2-\varepsilon-2))^{-1} = ((-2\varepsilon)(-\varepsilon))^{-1} = 1/(2\varepsilon^2)$$

as desired.

(b). *Solution.* From formula (1), since $\mathbf{c} = A^{-1}(\varepsilon)\mathbf{y}$, we have that

$$c_2(\varepsilon) = \frac{y_1 + y_2 - 2y_0}{2\varepsilon^2} = y_0(A^{-1})_{3,1} + y_1(A^{-1})_{3,2} + y_2(A^{-1})_{3,3}$$

Hence, if the y_0, y_1, y_2 are nonzero (if they are, we don't get any information for that entry of A), then

$$(A^{-1})_{3,1} = -(1/\varepsilon^2)$$
$$(A^{-1})_{3,2} = 1/(2\varepsilon^2)$$
$$(A^{-1})_{3,3} = 1/(2\varepsilon^2)$$

Since $|\varepsilon| < 1$ ensures that $(A^{-1})_{3,3} = 4 - 4\varepsilon + \varepsilon^2 = 4(1-\varepsilon) + \varepsilon^2 > 0$, our derivation works, and so we confirm what we found from formula (3), namely that $(A^{-1})_{3,3} = 1/(2\varepsilon^2)$.

(c). *Solution.* Note that $\|A(\varepsilon)\|_\infty > 0$, since $\|A(\varepsilon)\|_\infty \geq |(A)_{1,1}| + |(A)_{1,1}| + |(A)_{1,1}| = 1 + 2 + 4 = 7 > 0$ (using the Analog of Problem 3 again). Hence, if we let $C = \|A(\varepsilon)\|_\infty/2 > 0$, since $0 < 1/(2\varepsilon^2) \leq \|A^{-1}(\varepsilon)\|_\infty$ when $|\varepsilon| < 1$ from part (a), we see

$$\|A(\varepsilon)\|_\infty \|A^{-1}(\varepsilon)\|_\infty \geq \|A(\varepsilon)\|_\infty/(2\varepsilon^2) = C/\varepsilon^2$$