Problem 1

Who are your group members?

Solution. Nicholas Rees

Problem 2

The point of this exercise is to compare monomial interpolation (Section 10.2 of [A&G]) with Lagrange interpolation (Section 10.3).

(a). Let $p(x) = c_0 + c_1 x$ be the unique polynomial of degree at most 1 such that

$$p(2) = \sqrt{2}, \quad p(2.01) = \sqrt{3}$$

In exact arithmetic,

$$p(2.005) = \frac{\sqrt{2} + \sqrt{3}}{2}$$

since 2.005 is the midpoint between 2 and 2.01. Hence one can also write:

$$p(2.005) = c_0 + c_1(2.005)$$

Solve for c_0, c_1 using the Vandemonde matrix and the formula derived in class (see also page 300 of [A&G]). [Hint: you may find the following MATLAB commands useful:

```
A = fliplr( vander([2 2.01]))
y = [sqrt(2);sqrt(3)]
c = A^(-1)*y
trueVal = (y(1)+y(2))/2
monoVal = c(1) + c(2) * 2.005
```

What does MATLAB report for the absolute error in

$$(c_0 + c_1(2.005))$$

as an approximation for

$$\frac{\sqrt{2} + \sqrt{3}}{2}$$

(in absolute value)? What about the relative error?

(b). Same question, where

$$p(2) = \sqrt{2}, \quad p(2 + 10^{-6}) = \sqrt{3}$$

and you want to compute $p(2 + 10^{-6}/2)$. [Hint: Recall 5×10^{-7} in MATLAB notation is `5e-7` or `5.0e-7`.]

(c). Same question, where

$$p(2) = \sqrt{2}, \quad p(2 + 10^{-10}) = \sqrt{3}$$

and you want to compute $p(2+10^{-10}/2)$. [Hint: Recall 5×10^{-11} in MATLAB notation is `5e-11` or `5.0e-11`.]

(d). What is the L^p-condition number of A in part (c) for $p = \infty$? Do this FIRST by typing `cond(A,Inf)`, and SECOND check this by examining the values of A and A^{-1} and using the formula

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

(i.e., given in class and proven on the previous homework).

(e). Double precision for standard numbers has a relative precision error after rounding of roughly $2^{-53} = 1.1102\cdots \times 10^{-16}$ in the worst case (the reason is that a true value of $1 + 2^{-53}$ has to be stored as either 1 or $1 + 2^{-52}$ or a number farther away, resulting in a relative error of $2^{-53}/(1 + 2^{-53})$; of course, in the best case the relative error is 0). If you multiply this by the condition number of A (and this is only a very rough indication of the precision you'd expect to lose **c**...), what do you get?

(f). Now use the Lagrange formula for $p(x)$ in part (c):

$$p(x) = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0}$$

to calculate $p(2 + 10^{-10}/2)$; what are the absolute and relative errors in this calculation compared with the true value?

(g). Now use the Lagrange formula for $p(x)$ in part (c):

$$p(x) = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_1 - x_0}$$

to calculate $p(2 + 10^{-10}/3)$, and compute the true value of

$$p(2 + 10^{-10}/3) = (2/3)\sqrt{2} + (1/3)\sqrt{3}$$

via the MATLAB line `(2/3)*sqrt(2) + (1/3)*sqrt(3)`. What are the absolute and relative errors in the Lagrange formula computation as compared with the true value?

(a). *Solution.* We can compute c_0 and c_1 using the formula:

$$\begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = A^{-1}\mathbf{y} = \begin{bmatrix} x_0^0 & x_0^1 \\ x_1^0 & x_1^1 \end{bmatrix}^{-1} \begin{bmatrix} y_0 \\ y_1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 2.01 \end{bmatrix}^{-1} \begin{bmatrix} \sqrt{2} \\ \sqrt{3} \end{bmatrix}$$

We can compute the inverse to be

$$\frac{1}{2.01 - 2} \begin{bmatrix} 2.01 & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 201 & -200 \\ -100 & 100 \end{bmatrix}$$

We can then compute

$$\begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 201\sqrt{2} - 200\sqrt{3} \\ -100\sqrt{2} + 100\sqrt{3} \end{bmatrix}$$

Hence, written exactly, $c_0 = 201\sqrt{2} - 200\sqrt{3}$ and $c_1 = -100\sqrt{2} + 100\sqrt{3}$. MATLAB calculates `c0 = -62.1532` and `c1 = 31.7837`. Using these values, MATLAB can produce an approximate value of $p(2.005)$ with $p(2.005) = c_0 + c_1(2.005)$. MATLAB calculates an abosolute error of `7.9936e-15` and a relative error of `5.0813e-15`.

(b). *Solution.* Using the same method as before, MATLAB computes `c0 = -6.3567e+05` and `c1 = 3.1784e+05`. MATLAB then gives an absolute error of `6.1605e-11` and a relative error of `3.9161e-11`.

(c). *Solution.* Using the same method as before, MATLAB computes `c0 = -6.3567e+09` and `c1 = 3.1784e+09`. MATLAB then gives an absolute error of `1.2837e-06` and a relative error of `8.1599e-07`.

(d). *Solution.* First, MATLAB produces `1.2000e+11`. Secondly, recall the formula for the condition number of A, $\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$. Furthermore,

$$A^{-1} = \begin{bmatrix} 1 & 2 \\ 1 & 2 + 10^{-10} \end{bmatrix}^{-1} = \frac{1}{2 + 10^{-10} - 2} \begin{bmatrix} 2 + 10^{-10} & -2 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 2 \times 10^{10} + 1 & -2 \times 10^{10} \\ -10^{10} & 10^{10} \end{bmatrix}$$

Then we can compute

$$\begin{aligned}
\kappa_\infty(A) &= \left\| \begin{bmatrix} 1 & 2 \\ 1 & 2+10^{-10} \end{bmatrix} \right\|_\infty \left\| \begin{bmatrix} 2\times 10^{10}+1 & -2\times 10^{10} \\ -10^{10} & 10^{10} \end{bmatrix} \right\|_\infty \\
&= \max(|1|+|2|, |1|+|2+10^{-10}|)\max(|2\times 10^{10}+1|+|-2\times 10^{10}|, |-10^{10}|+|10^{10}|) \\
&= (3+10^{-10})(4\times 10^{10}+1) \\
&= 12\times 10^{10}+3+4+10^{-10} \\
&= 1.2\times 10^{11}+7+10^{-10}
\end{aligned}$$

This is approximately the same as the MATLAB value, since the 10^{11} term is so much larger than the others.

(e). *Solution.* MATLAB produces `1.3323e-05`.

(f). *Solution.* MATLAB produces an absolute and relative error of 0 in the calculation. Amazing!

(g). *Solution.* MATLAB produces an absolute error of `2.2204e-16` and a relative error of `1.4607e-16`.

Problem 3

(a). Let $p(x) = c_0 + c_1 x + c_2 x^2$ be the unique polynomial of degree at most 2 such that

$$p(2) = \sqrt{2}, \quad p(2.01) = \sqrt{3}, \quad p(2.02) = \sqrt{5}$$

Let

$$\alpha_2 = p(2.005)$$

(we will explain the subscript 2 in the notation α_2 below). Approximate α as follows: first solve for $\mathbf{c} = (c_0, c_1, c_2)$ as $\mathbf{c} = A^{-1}\mathbf{y}$ using the formula derived in class (see also page 300 of [A&G]) $A\mathbf{c} = \mathbf{y}$ where $\mathbf{y} = (y_0, y_1, y_2)$ and A is a Vandermonde matrix.

(i). What value do you get for α_2? Report this as a base 10 number $1.d_1d_2d_3d_4d_5d_6d_7\ldots$ (so drop the remaining digits, rather than round up/down, and make sure you type `format long` into MATLAB if you aren't seeing enough decimal places).

(ii). What does MATLAB report for the ∞-condition number of A? (Here a few decimal places suffice, e.g., $5.37\cdots\times 10^5$.)

You may find some of the following lines of MATLAB code helpful:

```
A = fliplr( vander([2, 2.01, 2.02]))
y = [sqrt(2);sqrt(3);sqrt(5)]
c = A^(-1)*y
% For the result below, note that MATLAB indexing
% begins with 1, not 0
monoVal = c(1) + c(2) * 2.005 + c(3) * (2.005)^2
cond(A,Inf)
```

(b). Let $q(x)$ be the unique polynomial of degree at most 2 such that

$$q(2) = \sqrt{2}, \quad q(2+10^{-6}) = \sqrt{3}, \quad q(2+10^{-6}\cdot 2) = \sqrt{5}$$

Let

$$\alpha_6 = q(2+10^{-6}/2)$$

Approximate α_6 in the same way as you did α_2 in part (a).

(i). What value do you get for α_6? Report this as a base 10 number $1.d_1d_2d_3d_4d_5d_6d_7\ldots$ (so drop the remaining digits, rather than round up/down, and make sure you type `format long` into MATLAB if you aren't seeing enough decimal places).

(ii). What does MATLAB report for the ∞-condition number of A?

(c). Same question in part (b), with $q(x)$, 10^{-6}, α_6 respectively replaced with $r(x), 10^{-7}, \alpha_7$.

(d). Same question in part (b), with $q(x)$, 10^{-6}, α_6 respectively replaced with $s(x), 10^{-8}, \alpha_8$.

(e). Let p be the polynomial in part (a), and q that in part (b). Show that $f(y) = p(2 + y10^{-2}) - q(2 + y10^{-6})$ is a polynomial in y of degree 2 such that $f(y) = 0$ for $y = 0, 1, 2$.

(f). Use the previous part to show that (in an exact computation) $\alpha_2 = \alpha_6$.

(g). Use the ideas of the two previous parts to argue that in exact computations, $\alpha_6 = \alpha_7$.

(h). Now use the Lagrange formula for quadratic polynomials,

$$p(x) = y_0 \frac{x - x_1}{x_0 - x_1} \frac{x - x_2}{x_0 - x_2} + y_1 \frac{x - x_0}{x_1 - x_0} \frac{x - x_2}{x_1 - x_2} + y_2 \frac{x - x_0}{x_2 - x_0} \frac{x - x_1}{x_2 - x_1}$$

to calculate $\alpha_2, \alpha_7, \alpha_8$ and report all the decimal places that MATLAB's `format long` reports. You may find the following MATLAB lines helpful for the α_2 calculation:

```
n=2
x0 = 2 ; x1 = 2 + 10^(-n) ; x2 = 2 + 10^(-n) * 2;
x = 2 + 10^(-n)/2;
y0 = sqrt(2); y1 = sqrt(3); y2 = sqrt(5);
L0 = (x-x1) * (x-x2) / ( (x0-x1) * (x0-x2) );
L1 = (x-x0) * (x-x2) / ( (x1-x0) * (x1-x2) );
L2 = (x-x0) * (x-x1) / ( (x2-x0) * (x2-x1) );
p = y0 * L0 + y1 * L1 + y2 * L2
```

(a). *Solution.* We can compute c_0, c_1, c_2 using the formula:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = A^{-1}\mathbf{y} = \begin{bmatrix} x_0^0 & x_0^1 & x_0^2 \\ x_1^0 & x_1^1 & x_1^2 \\ x_2^0 & x_2^1 & x_2^2 \end{bmatrix}^{-1} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2.01 & (2.01)^2 \\ 1 & 2.02 & (2.02)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sqrt{2} \\ \sqrt{3} \\ \sqrt{5} \end{bmatrix}$$

We can have MATLAB solve for c_0, c_1, c_2 and evaluate the polynomial $p(x) = c_0 + c_1 x + c_2 x^2$ at $x = 2.005$. We find $\alpha_2 = 1.549859694375755$.

MATLAB produces $5.7371 \cdots \times 10^5$ as the ∞-condition number of A.

(b). *Solution.* We can compute c_0, c_1, c_2 using the formula:

$$\begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = A^{-1}\mathbf{y} = \begin{bmatrix} 1 & 2 & 4 \\ 1 & 2 + 10^{-6} & (2 + 10^{-6})^2 \\ 1 & 2 + 10^{-6} \cdot 2 & (2 + 10^{-6} \cdot 2)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sqrt{2} \\ \sqrt{3} \\ \sqrt{5} \end{bmatrix}$$

We can have MATLAB solve for c_0, c_1, c_2 and evaluate the polynomial $q(x) = c_0 + c_1 x + c_2 x^2$ at $x = 2 + 10^{-6}/2$. We find $\alpha_6 = 14.468750000000000$.

MATLAB produces $3.5836 \cdots \times 10^{15}$ as the ∞-condition number of A.

(c). *Solution.* Using the same formula as before, we can have MATLAB solve for c_0, c_1, c_2 and evaluate the polynomial $r(x) = c_0 + c_1 x + c_2 x^2$ at $x = 2 + 10^{-7}/2$. MATLAB computes that $\alpha_7 = 8$.

MATLAB produces $7.3367 \cdots \times 10^{17}$ as the ∞-condition number of A.

(d). *Solution.* Using the same formula as before, we can have MATLAB solve for c_0, c_1, c_2 and evaluate the polynomial $s(x) = c_0 + c_1 x + c_2 x^2$ at $x = 2 + 10^{-8}/2$. MATLAB computes that $\alpha_8 =$ `Inf` and also the ∞-condition number of A is `Inf` as well.

(e). *Solution.* First, we can write $f(y)$:

$$\begin{aligned} f(y) &= c_{0,2} + c_{1,2}(2 + y10^{-2}) + c_{2,2}(2 + y10^{-2})^2 - c_{0,6} + c_{1,6}(2 + y10^{-6}) + c_{2,6}(2 + y10^{-6})^2 \\ &= c_{0,2} + 2c_{1,2} + c_{1,2}10^{-2}y + 4c_{2,2} + c_{2,2}20^{-2}y + c_{2,2}10^{-4}y^2 \\ &- c_{0,6} - 2c_{1,6} - c_{1,6}10^{-6}y - 4c_{2,6} - c_{2,6}20^{-6}y - c_{2,6}10^{-12}y^2 \\ &= d_0 + d_1 y + d_2 y^2 \end{aligned}$$

with appropriate collecting of terms. d_0, d_1, d_2 are all constants (since they are from just adding/subtracting constants), hence $f(y)$ is clearly a polynomial of degree at most 2 (we have degree less than 2 when $d_2 = 0$).

From parts (a) and (b) of this problem, we have $f(0) = p(2) - q(2) = \sqrt{2} - \sqrt{2} = 0$, $f(1) = p(2+10^{-2}) - q(2+10^{-6}) = \sqrt{3} - \sqrt{3} = 0$, and $f(2) = p(2+2\cdot 10^{-2}) - q(2+2\cdot 10^{-6}) = \sqrt{5} - \sqrt{5} = 0$. So $f(0) = f(1) = f(2) = 0$.

(f). *Solution.* From the theorem from class (Feb. 9), there is a unique degree at most 2 polynomial that passes through the three points $(0, f(0)), (1, f(1)), (2, f(2))$. f is a degree at most 2 polynomial that passes through all those points. Also 0 is a polynomial of degree at most 2 that passes through all the points. Hence, $f(y) = 0$ by uniqueness. Hence, $0 = f(1/2) = p(2 + \cdot 10^{-2}/2) - q(2 + \cdot 10^{-6}/2) = \alpha_2 - \alpha_6 \implies \alpha_2 = \alpha_6$.

(g). *Solution.* We claim that $g(y) = q(2 + y10^{-6}) - r(2 + y10^{-7})$ is a polynomial of degree at most 2 such that $g(0) = g(1) = g(2) = 0$. We expand out $g(y)$:

$$\begin{aligned} g(y) &= c_{0,6} + c_{1,6}(2 + y10^{-6}) + c_{2,6}(2 + y10^{-6})^2 - c_{0,7} + c_{1,7}(2 + y10^{-7}) + c_{2,7}(2 + y10^{-7})^2 \\ &= c_{0,6} + 2c_{1,6} + c_{1,6}10^{-6}y + 4c_{2,6} + c_{2,6}20^{-6}y + c_{2,6}10^{-12}y^2 \\ &- c_{0,7} - 2c_{1,7} - c_{1,7}10^{-7}y - 4c_{2,7} - c_{2,7}20^{-7}y - c_{2,7}10^{-14}y^2 \\ &= e_0 + e_1 y + e_2 y^2 \end{aligned}$$

with appropriate collecting of terms. e_0, e_1, e_2 are all constants (since they are from just adding/subtracting constants), hence $g(y)$ is clearly a polynomial of degree at most 2.

From parts (b) and (c) of this problem, see $g(0) = q(2) - r(2) = \sqrt{2} - \sqrt{2} = 0$, $g(1) = q(2+10^{-6}) - r(2+10^{-7}) = \sqrt{3} - \sqrt{3} = 0$, and $g(2) = q(2 + 2\cdot 10^{-6}) - r(2 + 2\cdot 10^{-7}) = \sqrt{5} - \sqrt{5} = 0$. So $g(0) = g(1) = g(2) = 0$.

From the theorem from class, there is a unique degree at most 2 polynomial that passes through the three points $(0, g(0)), (1, g(1)), (2, g(2))$. g is a degree at most 2 polynomial that passes through all those points. Also 0 is a polynomial of degree at most 2 that passes through all the points. Hence, $g(y) = 0$ by uniqueness. Hence, $0 = g(1/2) = q(2 + \cdot 10^{-6}/2) - r(2 + \cdot 10^{-7}/2) = \alpha_6 - \alpha_7 \implies \alpha_6 = \alpha_7$.

(h). *Solution.* We compute

$$\begin{aligned} \alpha_2 &= 1.549859694379098 \\ \alpha_7 &= 1.549859695416296 \\ \alpha_8 &= 1.549859694379095 \end{aligned}$$

For the next problem(s), recall that if A is a square, invertible matrix, and if $A\mathbf{x}_{\text{true}} = \mathbf{b}_{\text{true}}$ (representing the "true values" of vector $\mathbf{x}, \mathbf{b}$) and $A\mathbf{x}_{\text{approx}} = \mathbf{b}_{\text{approx}}$ (representing the "approximate values" or "observed values by some experiment"), in class we defined the p-norm relative error (for $1 \le p \le \infty$)

$$\text{RelError}_p(\mathbf{x}_{\text{approx}}, \mathbf{x}_{\text{true}}) := \frac{\|\mathbf{x}_{\text{approx}} - \mathbf{x}_{\text{true}}\|_p}{\|\mathbf{x}_{\text{true}}\|_p} \tag{1}$$

(assuming $\mathbf{x}_{\text{true}} \neq \mathbf{0}$) and similarly with $\mathbf{x}$ replaced with $\mathbf{b}$. (See also [A&G], pages 3 and Section 5.8.) In class we proved that

$$\text{RelError}_p(\mathbf{x}_{\text{approx}}, \mathbf{x}_{\text{true}}) \le \kappa_p(A)\text{RelError}_p(\mathbf{b}_{\text{approx}}, \mathbf{b}_{\text{true}}) \tag{2}$$

where

$$\kappa_p(A) = \|A\|_p \|A^{-1}\|_p$$

and, moreover, that for any A there are $\mathbf{x}_{\text{true}}, \mathbf{b}_{\text{true}}, \mathbf{x}_{\text{approx}}, \mathbf{b}_{\text{approx}}$ for which equality holds in (2). Equivalently, if $\mathbf{x}_{\text{error}} = \mathbf{x}_{\text{approx}} - \mathbf{x}_{\text{true}}$ and similarly for $\mathbf{b}_{\text{error}}$, then (2) is equivalent to

$$\frac{\|\mathbf{x}_{\text{error}}\|_p}{\|\mathbf{b}_{\text{error}}\|_p} \frac{\|\mathbf{b}_{\text{true}}\|_p}{\|\mathbf{x}_{\text{true}}\|_p} \leq \kappa_p(A) \tag{3}$$

([A&G] refer to $\mathbf{b}_{\text{error}}$ as the *residual*, and denote it $\hat{\mathbf{r}}$.)

Conversely, for any A, p, here is a recipe for producing cases where (2) holds with equality: let $\mathbf{b}_{\text{error}}$ and $\mathbf{x}_{\text{true}}$ be arbitrary (nonzero) vectors such that

$$\frac{\|A^{-1}\mathbf{b}_{\text{error}}\|_p}{\|\mathbf{b}_{\text{error}}\|_p} = \|A^{-1}\|_p, \quad \frac{\|A\mathbf{x}_{\text{true}}\|_p}{\|\mathbf{x}_{\text{true}}\|_p} = \|A\|_p \tag{4}$$

(such vectors do exist); then (3) holds, and so working backwards we set

$$\mathbf{x}_{\text{error}} = A^{-1}\mathbf{b}_{\text{error}}, \quad \mathbf{b}_{\text{true}} = A\mathbf{x}_{\text{true}} \tag{5}$$

and

$$\mathbf{x}_{\text{approx}} = \mathbf{x}_{\text{true}} + \mathbf{x}_{\text{error}}, \quad \mathbf{b}_{\text{approx}} = \mathbf{b}_{\text{true}} + \mathbf{b}_{\text{error}} \tag{6}$$

yielding an example for which (2) holds with equality.

Problem 4

Let $\varepsilon > 0$ be a real number (which we think of as small), and let

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 2+\varepsilon \end{bmatrix} \tag{7}$$

and hence

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

(a). What are $\|A\|_\infty$ and $\|A^{-1}\|_\infty$?

(b). Show that

$$\left\| A \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty = \|A\|_\infty \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty$$

and for any $\delta \in \mathbb{R}$

$$\left\| A^{-1} \begin{bmatrix} \delta \\ -\delta \end{bmatrix} \right\|_\infty = \|A^{-1}\|_\infty \left\| \begin{bmatrix} \delta \\ -\delta \end{bmatrix} \right\|_\infty$$

(c). Use the previous part to show that

$$\mathbf{b}_{\text{error}} = \begin{bmatrix} \delta \\ -\delta \end{bmatrix}, \quad \mathbf{x}_{\text{true}} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

satisfy (4); then let $\mathbf{x}_{\text{error}}$ satisfying (5), and show that the resulting $\mathbf{x}_{\text{approx}}$ is

$$\mathbf{x}_{\text{approx}}(\delta) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 4+\varepsilon \\ -2 \end{bmatrix} \frac{\delta}{\varepsilon} \tag{8}$$

(d). Show that $\mathbf{x}_{\text{approx}}(0)$ equals $\mathbf{x}_{\text{true}}$ above.

(e). Now check your work: let $\mathbf{x}_{\text{approx}}(\delta)$ be as in (8), and let $\delta \neq 0$.

(i). Evaluate

$$\text{RelError}_\infty(\mathbf{x}_{\text{approx}}, \mathbf{x}_{\text{true}}) = \frac{\|\mathbf{x}_{\text{approx}}(\delta) - \mathbf{x}_{\text{approx}}(0)\|_\infty}{\|\mathbf{x}_{\text{approx}}(0)\|_\infty}$$

(ii). Evaluate

$$\text{RelError}_\infty(A\mathbf{x}_{\text{approx}}, A\mathbf{x}_{\text{true}}) = \frac{\|A\mathbf{x}_{\text{approx}}(\delta) - A\mathbf{x}_{\text{approx}}(0)\|_\infty}{\|A\mathbf{x}_{\text{approx}}(0)\|_\infty}$$

(iii). Divide the result in (i) and (ii) and show that the result is equal to

$$\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$$

(which you should find to be $(3+\varepsilon)(4+\varepsilon)/\varepsilon$, using part (a)).

(a). *Solution.* Recall the formula we proved on the previous homework (and mentioned in problem 2(d) of this homework):

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

We can then compute (using the fact that $\varepsilon > 0$ when appropriate):

$$\begin{aligned} \|A\|_\infty &= \max(|1|+|2|, |1|+|2+\varepsilon|) = 3+\varepsilon \\ \|A^{-1}\|_\infty &= \max(|(2+\varepsilon)/\varepsilon| + |-2/\varepsilon|, |-1/\varepsilon| + |1/\varepsilon|) = \max((4+\varepsilon)/\varepsilon, 2/\varepsilon) = \frac{4+\varepsilon}{\varepsilon} \end{aligned}$$

(b). *Solution.* See that

$$\left\| A \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} 1+2 \\ 1+2+\varepsilon \end{bmatrix} \right\|_\infty = 3+\varepsilon = \|A\|_\infty \cdot 1 = \|A\|_\infty \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty$$

If $\delta \in \mathbb{R}$, then

$$\left\| A^{-1} \begin{bmatrix} \delta \\ -\delta \end{bmatrix} \right\|_\infty = \left\| \frac{1}{\varepsilon} \begin{bmatrix} 2\delta + \delta\varepsilon + 2\delta \\ -2\delta \end{bmatrix} \right\|_\infty = \frac{4+\varepsilon}{\varepsilon}\delta = \|A^{-1}\|_\infty \cdot \delta = \|A^{-1}\|_\infty \left\| \begin{bmatrix} \delta \\ -\delta \end{bmatrix} \right\|_\infty$$

(c). *Solution.* To verify that these choices of $\mathbf{b}_{\text{error}}$ and $\mathbf{x}_{\text{true}}$ satisfy (4), see

$$\frac{\|A^{-1}\mathbf{b}_{\text{error}}\|_\infty}{\|\mathbf{b}_{\text{error}}\|_\infty} = \frac{\|A^{-1}\|_\infty \left\| \begin{bmatrix} \delta \\ -\delta \end{bmatrix} \right\|_\infty}{\left\| \begin{bmatrix} \delta \\ -\delta \end{bmatrix} \right\|_\infty} = \|A^{-1}\|_\infty$$

$$\frac{\|A\mathbf{x}_{\text{true}}\|_\infty}{\|\mathbf{x}_{\text{true}}\|_\infty} = \frac{\|A\|_\infty \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty}{\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty} = \|A\|_\infty$$

Now assume that $\mathbf{x}_{\text{error}} = A^{-1}\mathbf{b}_{\text{error}}$. We can then find $\mathbf{x}_{\text{approx}}$ as a function of δ:

$$\mathbf{x}_{\text{approx}}(\delta) = \mathbf{x}_{\text{true}}(\delta) + \mathbf{x}_{\text{error}}(\delta) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + A^{-1} \begin{bmatrix} \delta \\ -\delta \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{1}{\varepsilon} \begin{bmatrix} 2\delta + \delta\varepsilon + 2\delta \\ -2\delta \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 4+\varepsilon \\ -2 \end{bmatrix} \frac{\delta}{\varepsilon}$$

(d). *Solution.* We can compute $\mathbf{x}_{\text{approx}}(0) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} 4+\varepsilon \\ -2 \end{bmatrix} \frac{0}{\varepsilon} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \mathbf{x}_{\text{true}}$.

(e). *Solution.* Let $\delta \neq 0$. We can compute

$$\text{RelError}_\infty(\mathbf{x}_{\text{approx}}, \mathbf{x}_{\text{true}}) = \frac{\|\mathbf{x}_{\text{approx}}(\delta) - \mathbf{x}_{\text{approx}}(0)\|_\infty}{\|\mathbf{x}_{\text{approx}}(0)\|_\infty} = \frac{\left\| \begin{bmatrix} 4+\varepsilon \\ -2 \end{bmatrix} \frac{\delta}{\varepsilon} \right\|_\infty}{\left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\|_\infty} = (4+\varepsilon)\frac{\delta}{\varepsilon}$$

See that $A\mathbf{x}_{\text{approx}}(0) = A\begin{bmatrix}1\\1\end{bmatrix} = \begin{bmatrix}1+2\\1+2+\varepsilon\end{bmatrix} = \begin{bmatrix}3\\3+\varepsilon\end{bmatrix}$ and using the linearity of matrix multiplication, we also get $A\mathbf{x}_{\text{approx}}(\delta) = A\left(\begin{bmatrix}1\\1\end{bmatrix} + \begin{bmatrix}4+\varepsilon\\-2\end{bmatrix}\frac{\delta}{\varepsilon}\right) = A\mathbf{x}_{\text{approx}}(0) + \frac{\delta}{\varepsilon}A\begin{bmatrix}4+\varepsilon\\-2\end{bmatrix} = A\mathbf{x}_{\text{approx}}(0) + \frac{\delta}{\varepsilon}\begin{bmatrix}4+\varepsilon-4\\4+\varepsilon-4-2\varepsilon)\end{bmatrix} = A\mathbf{x}_{\text{approx}}(0) + \delta\begin{bmatrix}1\\-1\end{bmatrix}$. Then we can compute

$$\begin{aligned}\text{RelError}_\infty(A\mathbf{x}_{\text{approx}}, A\mathbf{x}_{\text{true}}) &= \frac{\|A\mathbf{x}_{\text{approx}}(\delta) - A\mathbf{x}_{\text{approx}}(0)\|_\infty}{\|A\mathbf{x}_{\text{approx}}(0)\|_\infty} \\ &= \frac{\left\|A\mathbf{x}_{\text{approx}}(0) + \begin{bmatrix}\delta\\-\delta\end{bmatrix} - A\mathbf{x}_{\text{approx}}(0)\right\|_\infty}{\left\|\begin{bmatrix}3\\3+\varepsilon\end{bmatrix}\right\|_\infty} \\ &= \frac{\delta}{3+\varepsilon}\end{aligned}$$

Finally, we have that

$$\frac{\text{RelError}_\infty(\mathbf{x}_{\text{approx}}, \mathbf{x}_{\text{true}})}{\text{RelError}_\infty(A\mathbf{x}_{\text{approx}}, A\mathbf{x}_{\text{true}})} = \frac{(4+\varepsilon)\delta/\varepsilon}{\delta/(3+\varepsilon)} = (3+\varepsilon)\frac{4+\varepsilon}{\varepsilon} = \|A\|_\infty\|A^{-1}\|_\infty = \kappa_\infty(A)$$

where we have inserted the values of $\|A\|_\infty, \|A^{-1}\|_\infty$ from part (a) of this problem.