Marcus Rose
Math 126 HW 5
Due 3/11/23 (extension recieved)

---

1. Let

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

Find a nonzero vector $b \in \mathbb{R}^2$ and a nonzero vector $\Delta b \in \mathbb{R}^2$ so that for

$$x = A^{-1}b, \Delta x = A^{-1}\Delta b,$$

we have

$$\frac{\|\Delta x\|_1}{\|x\|_1} = \text{cond}_1(A)\frac{\|\Delta b\|_1}{\|b\|_1}.$$

Answer Note that $A^{-1} = \frac{1}{3}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ and $\text{cond}_1(A) = \|A\|_1\|A^{-1}\|_1 = 3$. We have to choose $\mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and $\|\mathbf{b}\|_1 = 2$. Then,

$$\mathbf{x} = A^{-1}\mathbf{b} \implies \mathbf{x} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ \frac{-1}{3} \end{bmatrix} \text{ and } \|\mathbf{x}\|_1 = \frac{2}{3}.$$

Note that the condition number of A is relatively large, so $\frac{\|\Delta \mathbf{x}\|_1}{\|\mathbf{x}\|_1}$ will be large too.

Let $\Delta\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ such that $\|\Delta\mathbf{b}\|_1 = 1, \Delta\mathbf{x} = A^{-1}\Delta\mathbf{b} = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{2}{3} \\ \frac{-1}{3} \end{bmatrix}$. Then,

$$\frac{\|\Delta\mathbf{x}\|_1}{\|\mathbf{x}\|_1} = \text{cond}_1(A)\frac{\|\Delta\mathbf{b}\|_1}{\|\mathbf{b}\|_1} \implies \frac{3}{2} = \text{cond}_1(A)\frac{1}{2} \implies \frac{3}{2} = \frac{3}{2}.$$

Hence, there exist nonzero $\mathbf{b} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \Delta\mathbf{b} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ such that

$$\frac{\|\Delta x\|_1}{\|x\|_1} = \text{cond}_1(A)\frac{\|\Delta b\|_1}{\|b\|_1}.$$

2. Let

$$x_1 < x_2 < ... < x_m,$$

and let $y_1, y_2, ..., y_m$ be real numbers. There exists exactly one polynomial p of degree $\leq m-1$ such that

$$p(x_i) = y_i, 1 \leq i \leq m.$$

Goal: prove this result without finding an explicit formula for the determinant of V_m.

(a) Explain: It suffices to prove that $p(x) = 0$ is the only solution when all the y_i are zero.

Answer If all y_i are zero, then

$$\begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \multicolumn{5}{c}{\dotfill} \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \mathbf{0},$$

and the set of solutions $V_m\mathbf{c} = \mathbf{0}$ spans the null space of the Vandermonde matrix. All x_i's are distinct, as the initial condition stated that x is strictly monotone. All columns of V_m are linearly independent and is hence full rank. The kernel of A must only contain the zero vector. I.e. the only $\mathbf{c}$ that satisfies $V_m\mathbf{c} = \mathbf{0}$ is $\mathbf{c} = \mathbf{0}$ and $p(x) = 0$.

(b) Assume that p were a polynomial of degree $\leq m-1$ with $p(x_i) = 0$ for all i. Show there are $m-1$ different points which the derivative p' is zero. Repeat...

Proof Using the Mean Value Theorem, for each subinterval, check to se if each $p(x)$ is satisfied for some interval $[x_i, x_{i+1}]$.

$x_i \leq x_{i+1}(\forall i)$ and for some $e_i \in (x_i, x_{i+1})$. $p(x_i) - p(x_{i+1}) = p'(e_i)(x_{i+1} - x_i)$ but $p(x_i) = 0 \forall i$. Then

$$p'(e_i)(x_{i+1} - x_i) = 0 \text{ and } x_{i+1} > x_i \forall i.$$

So $(x_{i+1} - x_i) > 0 \implies p'(e_i) = 0$. Repeating,

$$\begin{gathered}
[x_1, x_2] : p(x_1) - p(x_2) = p'(e_1)(x_2 - x_1) \implies p'(e_1) = 0. \\
[x_2, x_3] : p(x_2) - p(x_3) = p'(e_2)(x_3 - x_2) \implies p'(e_2) = 0. \\
\vdots \\
[x_i, x_{i+1}] : p(x_i) - p(x_{i+1}) = p'(e_i)(x_{i+1} - x_i) \implies p'(e_i) = 0. \\
\vdots \\
[x_{m-1}, x_m] : p(x_{m-1}) - p(x_m) = p'(e_{m-1})(x_m - x_{m-1}) \implies p'(e_{m-1}) = 0.
\end{gathered}$$

There are $m-1$ points such that $p'(x) = 0$. Note

$$p(x) = \sum_{i=0}^{m} c_{m+1}x^m \implies m \text{ coefficients}, p'(x) = \sum_{i=2}^{m} c_m x^{m-2} \implies m-1 \text{ coefficients}.$$

By the MVT, there must exist $m-1$ points such that $p'(x) = 0$. Using previous e_i,

$$\begin{gathered}
[e_1, e_2] : p'(e_1) - p'(e_2) = p''(z_1)(x_2 - x_1) \implies p'(z_1) = 0. \\
[e_2, e_3] : p(e_2) - p(e_3) = p'(z_2)(e_3 - e_2) \implies p'(z_2) = 0. \\
\vdots \\
[e_i, e_{i+1}] : p(e_i) - p(e_{i+1}) = p'(z_i)(e_{i+1} - e_i) \implies p'(z_{m-1}) = 0. \\
\vdots \\
[e_{m-1}, e_m] : p(e_{m-1}) - p(e_m) = p'(z_{m-1})(e_m - e_{m-1}) \implies p'(z_{m-2}) = 0,
\end{gathered}$$

where $z \in [e_i, e_{i+1}]$. Hence for the second derivative, there are $m-2$ unique points such that $p''(x) = 0$. Inferring for higher order derivatives, $p^{(3)}(x)$ has $m-3$ zeros and so on. I.e., $p^{(m-1)}(x)$ has $m-(m-2)$ zeros. That is, it must be that there are $m-1$ points such that p' is zero.

□

3. If you want to approximate a function, $f(x)$, on $[-1, 1]$, by a polynomial, and you have the luxury of picking the interpolation points x_i, it is possible to avoid some of the big 'wiggles' at the endpoints by picking the points in a smart way. The Chebyshev interpolation points are given in equation (12.10). So, assuming you know $f(x)$, once you've chosen the x_i, you compute $y_i = f(x_i), i = 1, ..., m$, and you have the data from which to build your $m-1$ degree polynomial that goes through those points (i.e. enforce $p(x_i) = y_i$).

   (a) Plot the Chebyshev points on the interval $[1, 1]$ for $m = 10$, then compute (numerically) the associated Vandermonde matrix V_{10}, such that $(V_{10})_{ij} = x_i^{j-1}$. What is the condition number (with respect to the 2 norm) of determining the coefficients of the polynomial $p(x) = c_0 + c_1 x + ... + c_9 x^9$ when one knows the values of $p(x_i) = y_i$ at the Chebyshev points?

Answer Here is the plot of Chebyshev points over the interval $[-1, 1]$:

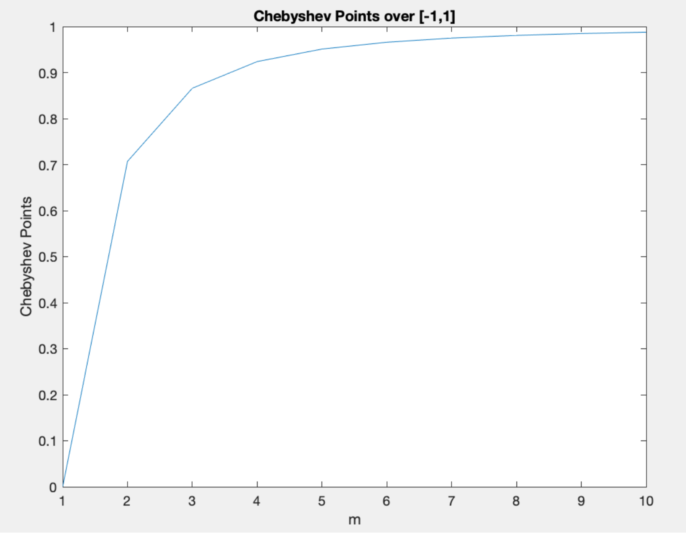
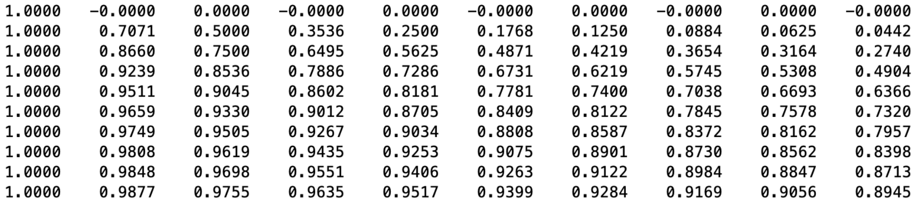


Figure 1: Chebyshev Points over [-1,1]

The Vandermonde matrix is



```
1.0000  -0.0000   0.0000  -0.0000   0.0000  -0.0000   0.0000  -0.0000   0.0000  -0.0000
1.0000   0.7071   0.5000   0.3536   0.2500   0.1768   0.1250   0.0884   0.0625   0.0442
1.0000   0.8660   0.7500   0.6495   0.5625   0.4871   0.4219   0.3654   0.3164   0.2740
1.0000   0.9239   0.8536   0.7886   0.7286   0.6731   0.6219   0.5745   0.5308   0.4904
1.0000   0.9511   0.9045   0.8602   0.8181   0.7781   0.7400   0.7038   0.6693   0.6366
1.0000   0.9659   0.9330   0.9012   0.8705   0.8409   0.8122   0.7845   0.7578   0.7320
1.0000   0.9749   0.9505   0.9267   0.9034   0.8808   0.8587   0.8372   0.8162   0.7957
1.0000   0.9808   0.9619   0.9435   0.9253   0.9075   0.8901   0.8730   0.8562   0.8398
1.0000   0.9848   0.9698   0.9551   0.9406   0.9263   0.9122   0.8984   0.8847   0.8713
1.0000   0.9877   0.9755   0.9635   0.9517   0.9399   0.9284   0.9169   0.9056   0.8945
```

Figure 2: Vandermonde Matrix

The condition number w.r.t. to the 2 norm is extremely large at $\text{cond}_2(V) = 7.1923 \times 10^{15}$, hinting that the problem is ill-conditioned.

(b) Plot the condition number of V_m as a function of m for $m = 1, ..., 20$. Is the polynomial interpolation problem well-conditioned when we use the Chebyshev points?

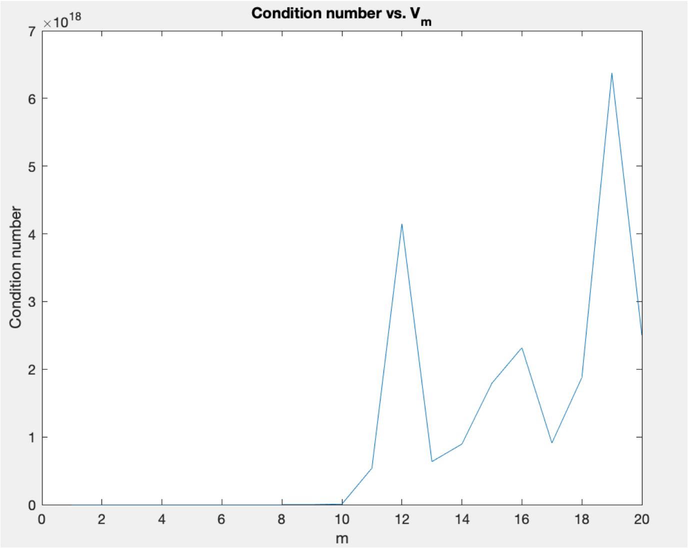


Figure 3: Condition number vs. V$_m$

Answer In the case where Chebyshev points are used, the interpolation problem becomes well conditioned when with small m, but around $m = 10$, there is a dramatic increase in the condition number before slightly decreasing and shooting up again at around 18.

(c) Repeat the last experiment using the uniform grid, $x_i = -1 + (i - 1) \times h, i = 1, ..., m$ where $h = \frac{2}{m-1}$ and report on any differences between the condition numbers computed for these choices of interpolation points.

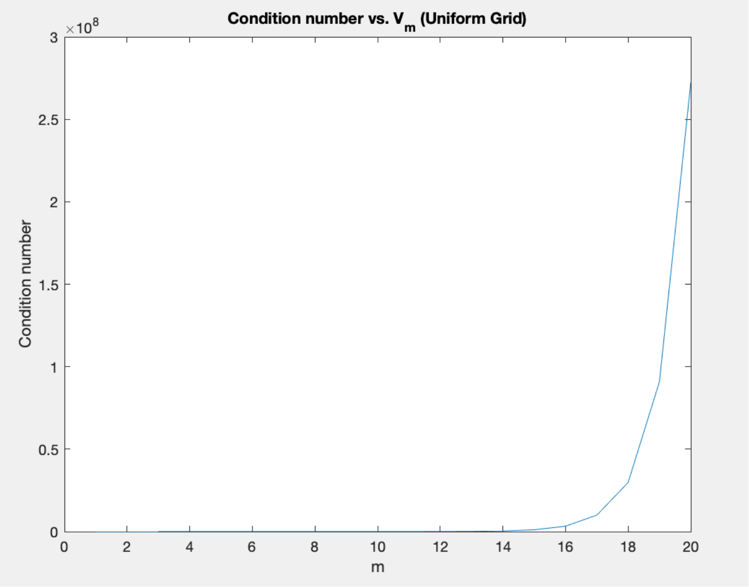


Figure 4: Part (b) plot, but on a uniform grid

Answer Note that using a uniform grid for the interpolation, the condition number stays relatively small for larger values of m when compared to the Chebyshev points in part (b) above. The condition number shoots up at around $m = 16$ in an exponential manner. This provides evidence that the problem is better conditioned for a slightly larger range of m values when conditioning via the uniform grid. However, for even larger $m > 16$, the Chebyshev points provide better conditioning.

4. Suppose that you were given a function $f = f(x)$, $0 \leq x \leq 1$, and wanted to find a polynomial

$$p(x) = \sum_{j=1}^{m} c_j x^{j-1}$$

that is as close to f as possible. Lagrange interpolation would be one approach, and we discussed that in Chapter 12. Another approach is to choose the coefficients c_j such that

$$\int_0^1 (p(x) - f(x))^2 dx = \int_0^1 \big(\sum_{j=1}^{m} c_j^{j-1} - f(x)\big)^2 dx$$

is as small as possible.

Assume that $c_j, 1 \leq j \leq m$, minimize the above integral equation. Show that

$$c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

solves

$$H_m c = r,$$

where H_m is the $m \times m$ Hilbert matrix, and $r \in \mathbb{R}^m$ is the vector defined by

$$r_j = \int_0^1 x^{j-1} f(x) dx, j = 1, 2, ..., m.$$

Proof

$$\int_0^1 (p(x) - f(x))^2 dx = \int_0^1 ([c_1 + c_2 x + c_3 x^2 + \ldots + c_m x^{m-1}] - f(x))^2 dx.$$

By expanding the square and reorganizing the values such that each of the coefficients of the same index are grouped together, we get

$$\int_0^1 [c_1^2 + (c_2 x)^2 + (c_3 x^2)^2 + \ldots (c_m x^{m-1})^2] + f(x)^2 + c_1 c_2 x + c_1 c_3 x^2 + \ldots + cf(x) + \ldots) dx.$$

Note that there is a finite continuation of the above sum; however the condensed simplification is

$$\sum_{i=1}\sum_{j=1}\int_0^1 c_i c_j x^{i+j-2}dx - 2c_j \sum_{i=1}^{m}\int_0^1 x^{j-1}f(x)dx + \int_0^1 (f(x))^2 dx.$$

And this simplifies to

$$\sum_{i=1}\sum_{j=1}\frac{c_i c_j}{i+j-1} - 2c_j \sum_{i=1}^{m}\int_0^1 x^{j-1}f(x)dx = c^T(H_m)c - 2c^T r + \int_0^1 (f(x))^2 dx.$$

Because we are told that c_j is the minimizer of the integrand specified in the problem, this turns this into a Lagrange multiplier problem, i.e.

$$\frac{\partial}{\partial c_i}(c^T(H_m)c - 2c^T r) = 0 \implies 2\sum_{j=1}^{m}\frac{c_j}{-1+i+j} = 2H_m c_i = 2r_i.$$

And because this is for all i components in the respective c and r vectors,

$$H_m c = r$$

is solved with solution c.

□

5. Compute the condition numbers for $f(x) = e^x, x^2,$ and $\ln(x)$, respectively.

Answer For $f(x) = e^x$,

$$\text{cond}(f) = |(e^x)'\frac{x}{e^x}| = |e^x\frac{x}{e^x}| = |x|.$$

For $f(x) = x^2$,

$$\text{cond}(f) = |(x^2)'\frac{x}{x^2}| = 2.$$

For $f(x) = \ln(x)$,

$$\text{cond}(f) = |(\ln(x))'\frac{x}{\ln(x)}| = |\frac{1}{\ln(x)}|.$$

6. Use the chain rule for derivatives to show that

$$\text{cond}(f_1 \circ f_2, x) = \text{cond}(f_1, f_2(x)) * \text{cond}(f_2, x).$$

What do you conclude about the condition number of m compositions $\text{cond}(f_1 \circ \ldots \circ f_m, x)$?

Proof Note that

$$\mathrm{cond}(f_1 \circ f_2, x) = |(f_1 \circ f_2)'(x) \frac{x}{(f_1 \circ f_2)(x)}|,$$

where $(f_1 \circ f_2)'(x) = f_1'(f_2(x))f_2'(x)$. Then, the above is

$$|f_1'(f_2(x))f_2'(x)\frac{x}{f_1(f_2(x))}| = |f_1'(f_2(x))f_2'(x)\frac{x}{f_1(f_2(x))}| \cdot \frac{f_2(x)}{f_2(x)},$$

and with this judicious multiplication of 1,

$$= |f_1'(f_2(x))\frac{f_2(x)}{f_1(f_2(x))}f_2'(x)\frac{x}{f_2(x)}| = \mathrm{cond}(f_1, f_2(x)) \cdot \mathrm{cond}(f_2, x).$$

Using the aforementioned, proved statement, we can see that the condition number would be multiplied together over intervals of two going to m when a set of m functions is composed with other functions. Note that we can conclude that as m increases, $\mathrm{cond}(f_1 \circ \ldots \circ f_m, x)$ increases rapidly and as such, this composition of functions is ill conditioned.

□

7. Let $p = 8$ digits of precision, so $\epsilon_{\text{mach}} = 5 \times 10^{-7}$. Let $x = 3.14159285$. First, find $\tilde{x} = fl(x)$ and then find δ such that $\tilde{x} = x(1+\delta)$, verifying that $|\delta| \le \epsilon_{\text{mach}}$. Next, let $z = 3.1415914$, so $fl(z) = \tilde{z} = z$ in this problem. The desired output is $y = f(x)$, but the algorithm returns $\hat{y} = fl(fl(x) - fl(z)) = fl(\tilde{x} - z)$. Show that this is still a backward stable algorithm (this time relative to x, not $\tilde{x}$. Then use the condition number and relative error in the inputs to explain the relative error you see between y and $\tilde{y}$.

Answer

$$|\delta| = \frac{|\hat{x} - x|}{x} = \frac{|3.1415929 - 3.14159285|}{|3.14159285|} = 1.587 \times 10^8.$$

With $z = 3.1415914$ and $fl(z) = \hat{z} = z$,

$$\hat{y} = fl(fl(x) - fl(z)) = fl(\hat{x} - z) = (\hat{x} - z)(1+\epsilon) = x - z + (x\delta - z)(1+\epsilon).$$

Looking at $\hat{y} - y = (x\delta - z)(1+\epsilon) \le |x\delta - z|\epsilon_{\text{mach}}$,

$$\frac{|\hat{y} - y|}{|y|} \le |x\delta - z|\epsilon_{\text{mach}}.$$

This implies that the algorithm is still a backward stable algorithm. Next,

$$|\frac{f'(x)x}{f(x)}| = |\frac{x}{\hat{x} - z}| = \frac{3.14159285}{3.1415929 - 3.1415914} < 2{,}000{,}000.$$

This implies that the problem is ill conditioned: there is an incredibly large condition number due to a very small value in the computed denominator.

Acknowledgements: I worked with Lulu Rajguru on problems 4, 6, and 7.