Problem 1
Let $A \in \mathbb{C}^{m\times m}$ be both upper triangular and unitary. Show that A is a diagonal matrix. Does the same hold if $A \in \mathbb{C}^{m\times m}$ is both lower-triangular and unitary?

Proof. Since A is unitary, by definition we have

$$AA^* = A^*A = I$$

Since A is upper triangular, it is of the form

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ 0 & a_{22} & \dots & a_{2m} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & a_{mm} \end{bmatrix}$$

Then

$$A^* = \begin{bmatrix} a_{11}^* & 0 & \dots & 0 \\ a_{12}^* & a_{22}^* & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m}^* & a_{2m}^* & \dots & a_{mm} \end{bmatrix}$$

Notice that A being upper triangular implies A^* lower triangular. If A were lower triangular, A^* would be upper triangular. Regardless, we know the product of these two matrices equals the identity. That is,

$$AA^* = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ 0 & a_{22} & \dots & a_{2m} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & a_{mm} \end{bmatrix} \begin{bmatrix} a_{11}^* & 0 & \dots & 0 \\ a_{12}^* & a_{22}^* & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m}^* & a_{2m}^* & \dots & a_{mm} \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} = I$$

So we may represent each entry of the resultant matrix as sum of the products of the rows and columns of A, A^*. Specifically, for entry b_{ij} we have the following

$$b_{ij} = \sum_{n=1}^{m} a_{in}a_{nj}^* \quad \text{for } 1 \leq i, j \leq m$$

Since we know the result is the identity matrix,

$$b_{ij} = \sum_{n=1}^{m} a_{in}a_{nj}^*, \quad 1 \leq i, j \leq m = \begin{cases} 0 & \text{when } i \neq j \\ 1 & \text{when } i = j \end{cases}$$

From this fact alone we can determine a surprising amount about our matrix A. For example, by the zero product property we know all entries $i \neq j$ must have at least one zero appear in each product. So for all $i \neq j$

$$b_{ij} = \sum_{n=1}^{m} a_{in}a_{nj}^* = 0 \implies a_{ij} = 0 \text{ or } a_{ij}^* = 0$$

Consider when $i = j$, i.e., the main diagonal. By the above fact, all products will be zero aside from when $i = j$. So for the nth diagonal entry,

$$b_{ii} = a_{in}a_{ni}^* = 1 \implies a_{in} \neq 0 \text{ and } a_{in}^* \neq 0$$

such that $1 \leq n \leq m$.

In conclusion, all entries of our matrix A that aren't on the main diagonal are zero, while all entries along the main diagonal are nonzero. This is precisely the definition of a matrix being diagonal. The same would hold if the given supposition was for a lower-triangular matrix, as then A^* would be upper triangle and the above argument applies.

$\square$

Problem 2
Prove the following in each problem.

(a) Let $A \in \mathbb{C}^{m\times m}$ be invertible and $\lambda \neq 0$ is an eigenvalue of A. Show that $1/\lambda$ is an eigenvalue of A^{-1}

(b) Let $A, B \in \mathbb{C}^{m\times m}$. Show that AB and BA have the same eigenvalues.

(c) Let $A \in \mathbb{R}^{m\times m}$. Show that A and A^* have the same eigenvalues.

2(a):

Proof. Let v denote the eigenvector associated with nonzero eigenvalue λ. Then by definition

$$Av = \lambda v.$$

Since A is invertible A^{-1} exists. Applying this to both sides gives

$$A^{-1}Av = A^{-1}\lambda v$$

$$\implies v = A^{-1}\lambda v$$

the scalar λ is nonzero, so we may divide both sides by this quantity, giving

$$\frac{1}{\lambda}v = A^{-1}v$$

which is precisely the definition of $\frac{1}{\lambda}$ being an eigenvalue of A^{-1}. □

2(b):

Proof. Consider an eigenvalue of the matrix AB, call it λ. Then with its associated eigenvector v, we have

$$ABv = \lambda v$$

For nonzero λ. Now let $x = Bv$. Note $x \in \mathbb{C}^m$. Then

$$\begin{aligned} BAx &= BABv \\ &= B(ABv) \\ &= B(\lambda v) \\ &= \lambda Bv \\ &= \lambda x \end{aligned} \tag{1}$$

Which is precisely the definition of λ being an eigenvalue for the matrix BA. That is, AB and BA share the same eigenvalues.

□

2(c):

Proof. Since $A \in \mathbb{R}^{m\times m}$, it follows that $A^* = A^T$. We know from lecture that for a square matrix $M \in \mathbb{R}^{m\times m}$

$$\det(M) = \det(M^T) \tag{2}$$

Now consider the characteristic equation

$$p_A(\lambda) = \det(A - \lambda I)$$

and since $A \in \mathbb{R}^{m\times m} \implies (A - \lambda I) \in \mathbb{R}^{m\times m}$ by equation (2) it follows that

$$\begin{aligned} p_A(\lambda) = \det(A - \lambda I) &= \det((A-\lambda I)^T) \\ &= \det(A^T - \lambda I^T) \\ &= \det(A^T - \lambda I) \\ &= P_{A^T}(\lambda) \end{aligned} \tag{3}$$

Thus the characteristic equations of A and A^* are the same. The zeros of these polynomials are eigenvalues, and thus A and A^* share the same eigenvalues.

$\square$

Problem 3
Let $A \in \mathbb{C}^{m\times m}$ be hermitian. Suppose that for nonzero eigenvectors of A, there exist corresponding eigenvalues λ satisfying $Ax = \lambda x$

(a) Prove that all the eigenvalues of A are real

(b) Let x and y be eigenvectors corresponding to distinct eigenvalues. Show that $(x, y) = 0$, i.e., they are orthogonal.

3(a):

Proof. We are given, for nonzero v, that

$$\begin{aligned} Av &= \lambda v \\ \implies (Av)^* &= (\lambda x)^* \\ \implies v^*A^* &= \overline{\lambda} v^* \end{aligned}$$

Now right multiply both sides by v to give

$$v^*A^*v = \overline{\lambda} v^*v$$

Notice that the product of v, v^* is a scalar, and as such we may let $v^*v = b$, some $b \in \mathbb{R}$. Since v was nonzero, b will be nonzero. Then

$$v^*A^*v = \overline{\lambda} b$$

and since $A = A^*$ by the definition of hermitian we have

$$v^*Av = \overline{\lambda} b$$

$$\implies v^*(\lambda v) = \overline{\lambda} b$$

since the reals are commutative we have

$$\begin{aligned} \lambda v^*v &= \overline{\lambda} b \\ \implies \lambda b &= \overline{\lambda} b \\ \implies \lambda &= \overline{\lambda} \end{aligned}$$

Which can only be true if $\lambda \in \mathbb{R}$, as desired. $\square$

3(b):

Proof. We are given that

$$Ax = \lambda_1 x \quad \text{and, } Ay = \lambda_2 y$$

for distinct eigenvalues λ_1, λ_2 (i.e., $\lambda_1 \neq \lambda_2$). Observe

$$\begin{aligned} Ay &= \lambda_2 y \\ \implies x^*Ay &= x^*\lambda_2 y \\ x^*Ay &= \lambda_2 x^*y \end{aligned} \tag{4}$$

Now, since A is hermitian $A^* = A$. Thus it follows that

$$\begin{aligned} x^*(Ay) = x^*A^*y &= (Ax)^*y \\ &= (\lambda_1 x)^*y \\ &= \lambda_1 x^*y \end{aligned} \tag{5}$$

where the last line follows by part (a), since we know all eigenvalues $\lambda \in \mathbb{R}$. Notice that we may equate the right hand side of equations (4) and (5), giving

$$\lambda_2 x^*y = \lambda_1 x^*y$$

Since $\lambda_1 \neq \lambda_2$, the above can only hold true when $x^*y = 0$. That is, x and y are orthogonal. $\square$

Problem 4
A matrix A is called positive definite if and only if $(Ax, x) > 0$ for all $x \neq 0$ in $\mathbb{C}^m$. Show that A is Hermitian and positive definite if and only if $\lambda_i > 0, \forall \lambda_i \in \Lambda(A)$, the spectrum of A.

Proof. $\boxed{\Rightarrow}$ Suppose A is Hermitian and positive definite. We are provided with a theorem which states that if $A \in \mathbb{C}^{m \times m}$ is Hermitian then it has real eigenvalues $\lambda_i, i = 1, 2 \ldots m$ not necessarily distinct and m corresponding eigenvectors u_i that form an orthonormal basis for $\mathbb{C}^m$. By the definition of an orthonormal basis, we can write any vector $x \neq 0$ as a linear combination of these vectors. E.g.,

$$x = \sum_{i=1}^{m} \alpha_i u_i$$

Then

$$\begin{aligned}(Ax, x) &= (x^* A^*)x = x^*(Ax) = x^*(\lambda x) \\ &= x^*(\lambda \sum_{i=1}^{m} \alpha_i u_i) \\ &= \sum_{i=1}^{m} \lambda_i |\alpha_i|^2\end{aligned}$$

Where the last line follows since the u_i forming an orthonormal basis implies $||u_i||_2 = 1$. Since we've supposed that A is positive definite, it follows that $(Ax, x) > 0 \ \forall x \neq 0$. Then it follows that

$$\sum_{i=1}^{m} \lambda_i |\alpha_i|^2 > 0$$

Clearly $|\alpha_i|^2 > 0$, so the above can only hold true if $\lambda_i > 0$ for each i. That is, the spectrum of A is comprised of only positive values.

$\boxed{\Leftarrow}$ For the reverse direction, we may only suppose that $\lambda_i > 0, \forall \lambda_i \in \Lambda(A)$, the spectrum of A. Then immediate we see that

$$(Ax, x) = \sum_{i=1}^{m} \lambda_i |\alpha_i|^2 > 0$$

if $A = A^*$, and thus $(Ax, x) > 0$. That is, we only need show A is Hermitian and the rest follows from the above line.

$\square$

Problem 5
Suppose A is unitary.

(a) Let (λ, x) be an eigenvalue-vector pair of A. Show λ satisfies $|\lambda| = 1$.

(b) Prove or disprove $||A||_F = 1$.

(a):

Proof. Recall a unitary matrix A is such that for $A \in \mathbb{C}^{m \times m}$,

$$AA^* = A^*A = I$$

Since we are given an eigenvalue-vector pair, it follows that

$$Ax = \lambda x$$

It is known in general that

$$x^*A^* = (\lambda x)^* = \lambda^* x^*$$

If we multiply these two expressions, we see

$$Ax(x^*A^*) = \lambda x(\lambda^* x^*)$$

Notice that xx^* is a dot product of possibly complex valued vectors, and as such the result is some constant $c \in \mathbb{R}$. Further, since the complex numbers are commutative we may rearrange the right hand side. Thus we have

$$AcA^* = \lambda\lambda^* xx^*$$
$$\implies cAA^* = \lambda\lambda^* c$$
$$\implies c = \lambda\lambda^* c$$
$$\implies 1 = \lambda\lambda^*$$

If we take the square root of both sides, we have the definition of the norm of a complex number. E.g.,

$$\sqrt{1} = \sqrt{\lambda\lambda^*} = |\lambda|$$

and thus $1 = |\lambda|$ as desired.

$\square$

(b):

Proof. Recall by definition

$$||A||_F = \sqrt{Tr(AA^*)}$$

and since A is unitary, for our particular problem

$$||A||_F = \sqrt{Tr(I)}$$

Since the trace is the sum of the main diagonal, if $A \in \mathbb{C}^{m \times m}$ then

$$\sqrt{Tr(I)} = \underbrace{\sqrt{1 + 1 + \ldots + 1}}_{\text{m times}}$$

$$\implies ||A||_F = \sqrt{m}$$

and thus the statement given is false in all generality. However, it does hold if $m = 1$, that is A is a scalar value. $\square$

Problem 6
Let $A \in \mathbb{C}^{m\times m}$ be skew-hermitian, i.e., $A^* = -A$

(a) Show that the eigenvalues of A are pure imaginary.

(b) Show that $I - A$ nonsingular.

(a):

Proof. Instead of going for this statement directly, consider the following statement. T is skew hermitian if and only if iT is hermitian. I will prove this then use it to get the desired result.

Suppose iT is hermitian. Then

$$(iT)^* = iT$$

$$\iff T^*(-i) = iT$$

$$\iff T^* = -T$$

and thus T is skew hermitian. Thus if we suppose A is skew hermitian, it follows that iA is hermitian. Note by problem 4, we invoked a theorem that states Hermitian matrices have real eigenvalues. Recall additionally that the determinant is a homomorphism and as such preserves linearity. So if iA is hermitian, and the eigenvalues are purely real, then A must have purely imaginary eigenvalues as desired. □

(b):

Proof. To show $I - A$ is nonsingular, notice that for an eigenvalue vector pair $\lambda, \;\; x \neq 0$ we know

$$Ax = \lambda x$$

$$\implies (A - \lambda I)x = 0$$

Since x is not the zero vector it must be that $x \in ker(A - \lambda I)$. In part (a) we showed that the eigenvalues of this skew-hermitian matrix are pure imaginary. Thus for $(A - \lambda I)x = 0$ to hold, it is required that the entries of λI be pure imaginary otherwise we contradict the definition of A being skew-hermitian.

$\therefore$ $(A - I)x \neq 0$ since $\lambda = 1$ is not pure imaginary. In general $(A - I)x \neq 0 \iff -(I - A)x \neq 0$. To summarize, for arbitrary nonzero vector x, $x \notin ker(I - A)$. Then $I - A$ is invertible as well, and thus is nonsingular. □

Problem 7
Show that $\rho(A) \leq ||A||$, where $\rho(A)$ is the spectral radius of A.

Proof. Recall by definition

$$\rho(A) = max\{|\lambda_1|, ||\lambda_2|, \ldots, |\lambda_n|\}$$

Let λ denote this max without loss of generality. Then we know if the eigenvector x is associated with λ that

$$Ax - \lambda x$$

$$\implies ||Ax|| = |\lambda|||x||$$

$$\implies \frac{||Ax||}{||x||} = |\lambda|$$

Then by the definition of supremum,

$$|\lambda| \ = \frac{||Ax||}{||x||} \leq \sup_{y \in \mathbb{C}^m} \frac{||Ay||}{||y||} = ||A||$$

$$\implies \rho(A) \leq ||A||$$

□

Problem 8
Let A be defined as an outer product $A = uv^*$, where $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$

(a) Prove or disprove $||A||_2 = ||u||_2||v||_2$

(b) Prove or disprove $||A||_F = ||u||_F||v||_F$ (Note: For a vector u, $||u||_F = ||u||_2$ by definition).

Note that since A is the outer product of $u \in \mathbb{C}^m$ and $v \in \mathbb{C}^n$, it follows that $A \in \mathbb{C}^{m\times n}$.

(a):

Proof. Recall the L2 norm of a matrix A is defined to be the square root of the maximum eigenvalue of AA^*. I.e.,

$$||A||_2 = \sqrt{\rho(AA^*)}$$

Plugging in the definition of A and A^*, noting $A^* = (uv^*)^* = vu^*$ we have

$$||A||_2 = \sqrt{\rho(uv^*vu^*)}$$

The ρ function is a glorified maximum function. Luckily for us, the product of maximums is the same as the maximum of products. Meaning

$$\sqrt{\rho(uv^*vu^*)} = \sqrt{\rho(uu^*)}\sqrt{\rho(v^*v)} = ||u||_2||v||_2$$

where the commutativity invoked above is a direct result of dark magics, unholy witchcraft, and blood sacrifice. Thus assuming one has access to these boons, it follows that

$$||A||_2 = ||u||_2||v||_2$$

□

(b):

Proof. Recall by definition

$$||A||_F = \left(\sum_j^n \sum_i^m |a_{ij}|^2\right)^{1/2}$$

We've also remarked in class that this definition implies

$$||A||_F = \sqrt{Tr(AA^*)} = \sum_{j=1}^n ||c_j||_2^2$$

where c_j is a column vector of A, and thus $1 \leq j \leq n$. Since A is the inner product uv^* the columns of A are $c_j = [u_1v_j, u_2v_j, \dots, u_mv_j]^T$. Then

$$\sum_{j=1}^n ||c_j||_2^2 = \sum_{j=1}^n u_1^2v_j^2 + u_2^2v_2^2 + \ldots + u_m^2v_j^2$$

On the other hand

$$||u||_2||v||_2 = \sqrt{u_1^2 + u_2^2 \ldots + u_m^2}\sqrt{v_1^2 + v_2^2 \ldots + v_m^2}$$

which is very close, but everything is square rooted. That is, this seems to be $||u||_2||v||_2 = \sqrt{||A||_F}$.

□

Problem 9
Let $A, Q \in \mathbb{C}^{m\times m}$, where A is arbitrary and Q unitary.

(a) Show that $||AQ||_2 = ||A||_2$

(b) Show that $||AQ||_F = ||QA||_F = ||A||_F$

(a):

Proof. By the definition of the L2 norm, we know

$$||AQ||_2 = \sup_{x\neq 0} \frac{||AQx||_2}{||x||_2} = \sup_{x\neq 0} \frac{||A||_2||Qx||_2}{||x||_2}$$

Since Q is unitary, it follows that $||Qx||_2 = ||x||_2$, and thus

$$\sup_{x\neq 0} \frac{||A||_2||Qx||_2}{||x||_2} = \sup_{x\neq 0} \frac{||A||_2||x||_2}{||x||_2} = ||A||_2$$

$$\implies ||AQ||_2 = ||A||_2$$

as desired.

□

(b):

Proof. From lecture note that $\sqrt{Tr(BB^*)} = \sqrt{Tr(B^*B)}$ for arbitrary linear transformation B. By the definition of Frobenius norm, we know

$$||AQ||_F = \sqrt{Tr(AQ(AQ)^*)} = \sqrt{Tr(AQQ^*A^*)}$$

Since Q is unitary, $QQ^* = Q^*Q = I$, and thus

$$||AQ||_F = \sqrt{Tr(AIA^*)} = \sqrt{Tr(AA^*)} = ||A||_F.$$

Similarly,

$$||QA||_F = \sqrt{Tr((QA)^*QA)} = \sqrt{Tr(A^*Q^*QA)} = \sqrt{Tr((A^*A)} = ||A||_F$$

Thus as desired,

$$||A||_F = ||AQ||_F = ||QA||_F$$

□

Problem 10
We say that $A, B \in \mathbb{C}^{m\times m}$ are unitarily equivalent if $A = QBQ^*$ for some unitary $Q \in \mathbb{C}^{m\times m}$.

(a) Show that if A and B are unitarily equivalent, then they have the same singular values.

(b) Show that the converse of part (a) is not necessarily true.

(a):

Proof. Suppose A and B are unitarily equivalent. Then $A = QBQ^*$ implies $Q^*AQ = B$. Invoking the SVD theorem, we know a general matrix can be decomposed into a product of three matrices. It is best to think of this as a sandwich. Our bread unitary, the delicious vegetables a diagonal matrix.

In terms involving less sandwiches, let

$$B = U\Sigma V^*$$

where U, V^* are unitary and Σ is a diagonal matrix comprised of the singular values in descending order. Then

$$A = QBQ^* \implies A = Q(U\Sigma V^*)Q^*$$

Since matrix multiplication is associative we have

$$A = (QU)\Sigma(V^*Q^*)$$

Notice this is almost a singular value decomposition for A, and if it were we would be done. Σ represents the singular values of A, which can be seen to be identical to that of B. However, we need to (or at least should) show that QU and V^*Q^* are unitary. This can be done in generality if we can show that the product of unitary matrices is unitary.

To that end, let C, D denote two arbitrary unitary matrices. Note that $C^*C = CC^* = I = DD^* = D^*D$. Now consider

$$(CD)(CD)^* = CDD^*C^* = CIC^* = CC^* = I$$

which is percisely the definition of matrix CD being unitary. Thus the product of unitary matrices are unitary, and so indeed

$$A = (QU)\Sigma(V^*Q^*)$$

is a valid singular value decomposition, and we can plainly see that A and B share the same singular values. □

(b):

Proof. The converse of part (a) is as follows: If two matrices A and B have the same singular values, then they are unitarily equivalent. To show that this converse need not be true in all generality, I'll try to find a counter example. Consider two matrices A and B that have the same singular values, with their following singular value decomposition's.

$$A = U_1\Sigma V_1^*, \quad B = U_2\Sigma V_2^*$$

However, it could be that A and B are not even the same dimensions in this case. As we have seen in lecture, this Σ is sometimes comprised of not only SVD on the main diagonal, but zeros if the dimensions do not correspond. Then it is impossible for A and B to be unitarily equivalent since unitary matrices are square, and no product of square matrices with nonsquare matrices results in a square matrix.

□

Problem 11
Find the relative condition number of the following functions and discuss if there is any concern of being ill-conditioned. If so, discuss when.

(a) $f(x_1, x_2) = x_1 + x_2$

(b) $f(x_1, x_2) = x_1 x_2$

(c) $f(x_1, x_2) = (x - 2)^9$

Solution: Recall that if f is differentiable at $x = x_0$, the relative condition number is given by

$$k(x_0) = \frac{||J_0||||x_0||}{||f(x_0)||}$$

under some norm $||\cdot||$. For this problem I'll use the infinity norm $||\cdot||_\infty$.

(a): Notice $f(x_1, x_2) = x_1 + x_2$ is differentiable for all values of x_1, x_2. Then

$$J_f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

$$\implies ||J_0||_\infty = 2$$

$$\implies k(x_1, x_2) = \frac{2 \cdot max\{|x_1|, |x_2|\}}{|x_1 + x_2|}$$

$\implies$ Ill conditioned. As $x_1 \to -x_2$ our relative condition number will approach infinity.

(b): Observe $f(x_1, x_2) = x_1 x_2$ is differentiable for all x_1, x_2.

$$J_f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} x_2 & x_1 \end{bmatrix}$$

$$\implies ||J_0||_\infty = |x_1| + |x_2|$$

$$\implies k(x_1, x_2) = \frac{(|x_1| + |x_2|) \cdot max\{|x_1|, |x_2|\}}{|x_1 x_2|}$$

$\implies$ Ill conditioned. As if either $x_2 \to 0$ or $x_1 \to 0$ we have blow up. That is,

$$\lim_{x_1 \to 0} k(x_1, x_2) = \lim_{x_2 \to 0} k(x_1, x_2) = \infty$$

(c): Observe $f(x_1, x_2) = (x - 2)^9$ is differentiable for all x. Then

$$J_f = \frac{df}{dx} = 9(x-2)^8$$

and

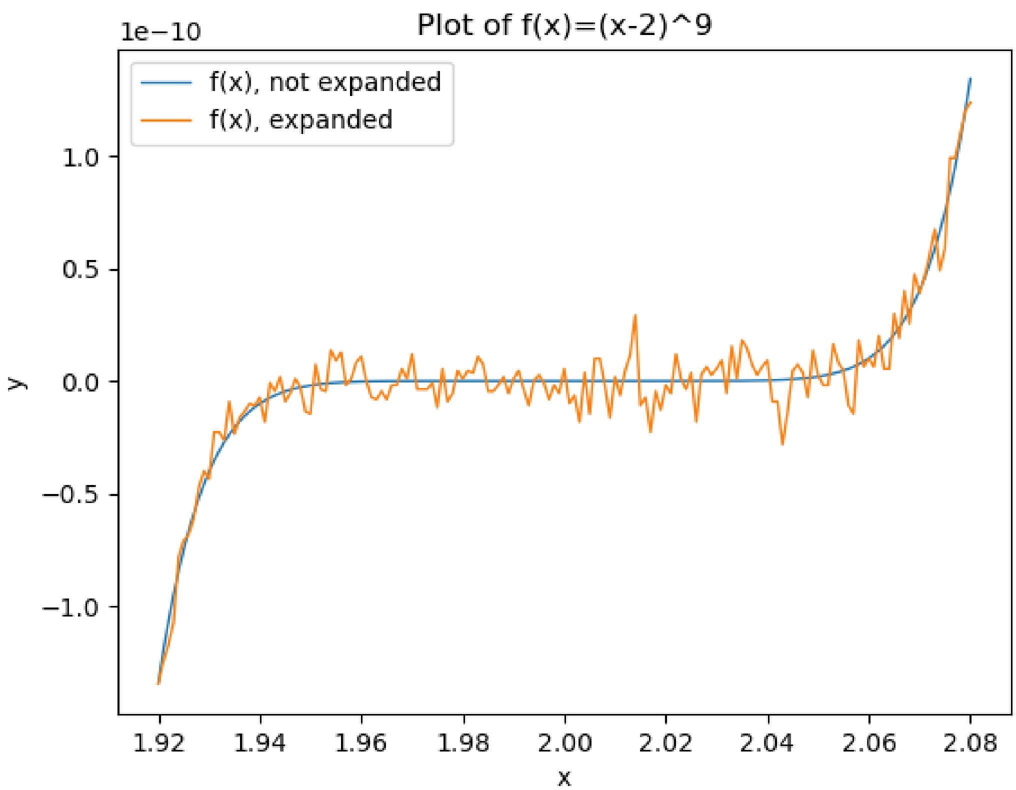
$$k(x) = \left| \frac{9(x-2)^8 x}{(x-2)^9} \right|$$

$$= \frac{|9x|}{|(x-2)|}$$

$\implies$ Ill conditioned, as when $x \to 2$, $k(x) \to \infty$

Problem 12
Plot the function $f(x) = (x-2)^9$, both in its expanded and given form. Overlay the graphs and discuss what you see.

Solution: Consider the following graph generated by python,



We can see that in its expanded form $f(x)$ varies wildly from the expected smooth plot of the unexpanded form. This confirms what we saw algebraically in problem 11, as we discovered that the function is ill-conditioned around $x = 2$. I actually expected the unexpanded form to be the inaccurate one, but upon reflecting I think this makes sense. That is because the expanded form has significantly more computations than one subtraction and nine multiplications. As such, the expanded form introduces significantly more floating point errors. This is exacerbated by $f(x)$ being very close to zero (and hence machine epsilon).