AM213A Homework 1

Owen Morehead

January 30, 2022

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?

   WLOG, suppose A is upper-triangular and unitary ($A^* = A^{-1}$). We know that since A is upper-triangular, so is its inverse, A^* (previously proved).

   $$a_{ij} = 0 \text{ for } i > j, \text{ and } \bar{a}_{ji} = 0 \text{ for } j > i.$$

   Therefore we have $a_{ij} \neq 0$ only if $i = j$, hence matrix A is a diagonal matrix.
   The proof is analogous if A is lower-triangular and unitary. So yes, the same does hold if $A \in \mathbb{C}^{m\times m}$ is both lower-triangular and unitary.

   $$a_{ij} = 0 \text{ for } i < j, \text{ and } \bar{a}_{ji} = 0 \text{ for } j < i.$$

   Therefore we have $a_{ij} \neq 0$ only if $i = j$, hence matrix A is a diagonal matrix. □

2. Prove the following in each problem.

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

   We can prove this with the following:

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

   Hence by definition, $1/\lambda$ 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.

   For a scalar $\lambda \neq 0$, If λ is an eigenvalue of AB, then there is $v \neq 0$ such that $ABv = \lambda v$. Let $y = Bv$ which is also a non-zero vector. So $y \neq 0$ (otherwise $y = 0$ would imply that $\lambda = 0$ or $v = 0$, and by definition both do not equal zero). Now we have:

   $$Ay = \lambda v \to BAy = BABv = B(ABv) = B(\lambda v) = \lambda Bv = \lambda y.$$

   And to finish the proof, consider the case that $\lambda = 0$ is an eigenvalue of AB. Then, utilizing the properties of determinant, $0 = det(AB) = det(A)det(B) = det(B)det(A) = det(BA)$. This implies that 0 is also an eigenvalue of BA iff it is an eigenvalue of AB. □

Or we can prove with the following:
If $det(A) \neq 0$, then:

$$det(AB - \lambda I) = det(A^{-1}A)det(AB - \lambda I) = det(A^{-1})det(AB - \lambda I)det(A)$$
$$= det(A^{-1}ABA - \lambda I) = det(BA - \lambda I).$$

$\square$

(c) Let $A \in \mathbb{R}^{m\times m}$. Show that A and A^* have the same eigenvalues. Hint, Use $det(M) = det(M^T)$ for any square matrix $M \in \mathbb{R}^{m\times m}$ in connection to the definition of characteristic polynomials. Hint, when a real-valued matrix A has a complex eigenvalue λ, then $\bar{\lambda}$ is also an eigenvalue of A.

The characteristic polynomial of a matrix A is given by $p_A(\lambda) = \sum_i c_i\lambda^i$. Since A is a real matrix, the c_i coefficients are also all real. We also know that λ is a root of $p_A(x) = 0$, i.e., $\sum_i c_i\lambda^i = 0 \to \sum_i c_i\bar{\lambda}^i = 0$. So we see that $\bar{\lambda}$ is also a root of $p_A(\lambda) = 0$ as assumed. Now for the rest of the proof.

The eigenvalues are found by solving

$$det(A - \lambda I) = 0.$$

We also know that

$$(A - \lambda I)^T = A^T - \lambda I^T = A^T - \lambda I.$$

Thus

$$det(A - \lambda I) = det(A^T - \lambda I).$$

This shows that the two matrices, A and A^T have the same characteristic polynomial, hence they share the same eigenvalues. $\square$

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

(a) Prove that all eigenvalues of A are real, i.e., $\lambda = \lambda^*$.

We know that $Ax = \lambda x$ for $x \in \mathbb{C}$ and $x \neq 0$. We also know that matrix A is Hermitian, $A = A^*$. Therefore we can expand,

$$\lambda||x||^2 = \lambda x^* x = x^*(\lambda x) = x^* Ax = (Ax)^* x = (\lambda x)^* x = x^*\lambda^* x = \lambda^*||x||^2.$$

Therefore we have $\lambda||x||^2 = \lambda^*||x||^2$, and since $x \neq 0$, it must be true that $\lambda = \lambda^*$, or in other words, all eigenvalues of A are real. $\square$

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

Since x and y are eigenvectors with respective eigenvalues λ and μ, we can write, $Ax = \lambda x$ and $Ay = \mu y$ with $\lambda \neq \mu$. In a similar fashion to part a, we can write:

$$y^*\lambda x = (y^* A)x = (Ay)^* x = (\mu y)^* x = y^*\mu x.$$

We can take the steps as shown above since A is hermitian, and as proved in part a, $\lambda = \lambda^*$, and similarly, $\mu = \mu^*$. From above we have $y^*\lambda x = y^*\mu x$, and since by definition, $\lambda \neq \mu$, we must have $y^* x = x^* y = 0$, or in other words, x and y are orthogonal. $\square$

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.

We first assume that A is Hermitian and positive definite. We use the theorem that states, if $A \in \mathbb{C}^{m\times m}$ is Hermitian ($A = A^*$), then A has real eigenvalues, λ_i where $i = 1, 2, ..., m$, not necessarily distinct, and m corresponding eigenvectors, u_i form an orthonormal basis for $\mathbb{C}^m$. For any arbitrary $x \neq 0$, we can write $x = \sum\limits_{i=1}^{m} \alpha_i u_i$.
We can now write the inner product,

$$(Ax, x) = x^* A^* x = x^* Ax = \sum_{i=1}^{m} \lambda_i \alpha_i^* u_i^* u_i \alpha_i.$$

Which simplifies to,

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

We can write this by definition of eigenvalues and eigenvectors, $Au = \lambda u$, and by the fact that eigenvectors can be scaled (x in the above equation is a set of eigenvectors).
Since we assume A is Hermitian and positive definite, then by definition:

$$x^* A^* x > 0 \to \sum_{i=1}^{m} \lambda_i |\alpha_i|^2 > 0.$$

Since by definition, $|\alpha_i|^2 > 0 \ \forall \ \alpha_i$, we can conclude that $\lambda_i > 0$, $\forall \lambda_i \in \Lambda(A)$, i.e., all the eigenvalues of a Hermitian positive definite matrix must be positive.

On the other hand, suppose A is a Hermitian matrix whose eigenvalues, $\lambda_1 \leq \lambda_2 ... \leq \lambda_n$, are all positive. Then, as done in the first part of the proof, let the orthonormal basis of eigenvectors be denoted by $u_1, u_2, ..., u_m$, such that any $x \in \mathbb{C}^m$ can be written as, $x = \sum\limits_{i=1}^{m} \alpha_i u_i$.
As already shown,

$$x^* Ax = \sum_{i=1}^{m} \lambda_i |\alpha_i|^2 \geq \lambda_1 \sum_{i=1}^{m} |\alpha_i|^2.$$

Assuming $x \neq 0$, then $0 \neq ||x||^2 = \sum\limits_{i=1}^{m} |\alpha_i|^2$. Since all the eigenvalues are assumed to be positive, we must have that A is positive definite,

$$x^* Ax > 0.$$

□

5. Suppose A is unitary ($A^{-1} = A^*$).

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

   Since A is unitary ($A^* A = I$), we can prove this as follows:

   $$x^* x = x^*(A^* A)x = (Ax)^*(Ax) = (\lambda x)^*(\lambda x) = ||\lambda||^2 x^* x.$$

   Now it must be true that $||\lambda||^2 = 1 \to |\lambda| = 1$, since $x \neq 0$. □

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

By definition $||A||_F = (\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2)^{1/2} = \sqrt{tr(A^*A)}$. Since A is unitary, $A^*A = I_n$ where n is the dimension of A. Thus we have

$$||A||_F = \sqrt{tr(I_n)} = \sqrt{n} \neq 1 \text{ for } n > 1.$$

Thus, for example, for a 2x2 unitary matrix A, $||A||_F = \sqrt{tr(I_2)} = \sqrt{2} \neq 1$. □

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.

Similar to the proof process of Problem 3, we can write:

$$\lambda||x||^2 = \lambda x^*x = x^*(\lambda x) = x^*Ax = -x^*A^*x = -(Ax)^*x = -(\lambda x)^*x = -\lambda^*||x||^2.$$

$$\to \lambda^* = -\lambda$$

i.e., λ is purely imaginary. □

(b) Show that $I - A$ is non-singular ($I - A$ is invertible, null$(I - A) = \{0\}$ are equivalent statements).

To prove this we want to use the property for a non-singular matrix, A, $Ax = 0$ only has the trivial solution $x = 0$. Let $(I - A)x = 0 \to x = Ax$. Then we have,

$$x^*x = (Ax)^*x = x^*A^*x = -x^*Ax = -x^*x.$$

We see that

$$x^*x = -x^*x \to x^*x = 0 \to x = 0.$$

Hence, $I - A$ is non-singular. □

7. Show that $\rho(A) \leq ||A||$, where $\rho(A)$ is the spectral radius of A (i.e., the largest absolute value $|\lambda|$ of an eigenvalue λ of A.

For any eigenvalue, λ, and corresponding eigenvector, $x \in \mathbb{C}^m\backslash 0$, we have by definition, $Ax = \lambda x$. Thus we can write,

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

$$\to |\lambda| = \frac{||Ax||}{||x||} \leq \sup_{y\in\mathbb{C}^m\backslash\{0\}} \frac{||Ay||}{||y||} = ||A||$$

where we used the definition of induced matrix norm. This leads us to $|\lambda| = \rho(A) \leq ||A||$. □

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

I will prove this statement. For any n-vector, x, we can bound $||Ax||_2$ as follows:

$$\frac{||Ax||_2}{||x||_2} = \frac{||uv^*x||_2}{||x||_2} = \frac{|v^*x|||u||_2}{||x||_2} \leq ||u||_2||v||_2.$$

The last inequality is due to the Cauchy-Schwarz inequality. Therefore, $||A||_2 \leq ||u||_2||v||_2$. If we let $x = v$ in the above equation, we can see that the equality holds, i.e.,

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

$\square$

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

I will prove this statement. Note that $A_{ij} = u_i v_j$, $||A||_F = ||uv^*||_F$. Therefore we have:

$$||A||_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}|A_{ij}|^2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}|u_i|^2|v_j|^2} = \sqrt{(\sum_{i=1}^{m}|u_i|^2)(\sum_{j=1}^{m}|v_j|^2} = ||u||_F||v||_F.$$

$\square$

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

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

First note that we have:

$$||Qx||_2 = (Qx, Qx)_2 = (x, Q^*Qx)_2 = (x, x)_2 = ||x||_2.$$

Hence we can write:

$$||AQ||_2 = \sup_{x\in\mathbb{C}^m\setminus\{0\}}\frac{||AQx||_2}{||x||_2} = \sup_{Qx\in\mathbb{C}^m\setminus\{0\}}\frac{||AQx||_2}{||Qx||_2} = \sup_{y\in\mathbb{C}^m\setminus\{0\}}\frac{||Ay||_2}{||y||_2} = ||A||_2.$$

$\square$

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

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

and similarly:

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

where we have used the fact that $Tr[XY] = Tr[YX]$ to prove that

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

$\square$

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. (Hint, note that every matrix has a SVD and their singular values are uniquely defined).

Let $A = U_1 \Sigma_1 V_1^*$ and $B = U_2 \Sigma_2 V_2^*$. If A and B are unitarily equivalent, there exists a unitary matrix Q such that $A = QBQ^*$.

$$A = U_1 \Sigma_1 V_1^* = QU_2 \Sigma_2 V_2^* Q^* = (QU_2)\Sigma_2(QV_2)^*.$$

In other words, $(QU_2)\Sigma_2(QV_2)^*$ is an SVD of A. (QU_2) and $(QV_2)^*$ are unitary matrices. Since we know that the SVD of a matrix is unique, $\Sigma_1 = \Sigma_2$, i.e., A and B have the same singular values. □

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

To show that the converse is not necessarily true, I will provide a counterexample of two matrices, $A, B \in \mathbb{C}^{m\times m}$ that have the same singular values but are not unitarily equivalent. Consider the matrices: A $= \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ and B $= \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix}$.

We can show that A and B share the same singular values: $A^T A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$

A$^T A - \lambda I = 0 \rightarrow \begin{pmatrix} 2-\lambda & 0 \\ 0 & 0-\lambda \end{pmatrix} = 0 \rightarrow \lambda = \{2\} \rightarrow$ singular values $= \{\sqrt{2}\}$.

$B^T B = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$

B$^T B - \lambda I = 0 \rightarrow \begin{pmatrix} 2-\lambda & 0 \\ 0 & 0-\lambda \end{pmatrix} = 0 \rightarrow \lambda = \{2\} \rightarrow$ singular values $= \{\sqrt{2}\}$.

We see that A and B share the same singular values.
We can perform SVD on each matrix to evaluate further. For matrix A we have:

$\Sigma = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{pmatrix}$, $det(\Sigma) = 0$.
For $\lambda = 2$ we have:
$(A^T A - \lambda I)\vec{x_1} = 0 \rightarrow \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = 0$
$0 = 0 \rightarrow \vec{x_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{x_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$U = (x_1 \; x_2) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
$V = \frac{1}{\sigma_1} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \mathrm{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{pmatrix}$

Since we only have one nonzero σ_i, and we need one more vector, v_2, we can find the orthogonal vector to all the vectors by finding the null space of the matrix whose rows are the vectors, $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Normalizing v_2, we are left with: $V = \begin{pmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{pmatrix}$

Thus we have the SVD of matrix A:

$$\begin{pmatrix}1 & 1\\0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix}\frac{\sqrt{2}}{2}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}$$

For matrix B we have:

$\Sigma = \begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix}$, $det(\Sigma) = 0$.
For $\lambda = 2$ we have:
$(A^TA - \lambda I)\vec{x_1} = 0 \rightarrow \begin{pmatrix}0 & 0\\0 & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} = 0$
$0 = 0 \rightarrow \vec{x_1} = \begin{pmatrix}1\\0\end{pmatrix}$, $\vec{x_2} = \begin{pmatrix}0\\1\end{pmatrix}$

$U = (x_1\, x_2) = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$
$V = \frac{1}{\sigma_1}\begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix}\mathrm{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$

Since we only have one nonzero σ_i, and we need one more vector, v_2, we can find the orthogonal vector to all the vectors by finding the null space of the matrix whose rows are the vectors, $\begin{pmatrix}0\\1\end{pmatrix}$. Normalizing v_2, we are left with: $V = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$ Thus we have the SVD of matrix B:

$$\begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}\sqrt{2} & 0\\0 & 0\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

Thus we see that although A and B share the same singular values, they cannot be unitarily equivalent. □

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

$$J_{1\times 2} = (\frac{\partial f}{\partial x_1} + \frac{\partial f}{\partial x_2}) = (1, 1)$$
$$||J||_\infty = 2$$
$$\kappa(x_1, x_2) = \frac{||J||_\infty ||x_0||_\infty}{||f(x_1, x_2)||_\infty} \text{ where } x_0 = (x_1, x_2)^T \rightarrow \kappa(x_1, x_2) = \frac{2(max\{|x_1|, |x_2|\})}{|x_1 + x_2|}$$

This quantity is large if $|x_1 + x_2| \approx 0$. So there is concern of this problem becoming ill-conditioned if this is true. This is true when $x_1 \approx -x_2$.

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

The relative condition number is:

$$\kappa(x_1, x_2) = \frac{||J||_\infty ||x_0||_\infty}{||f(x_1, x_2)||_\infty} = \frac{(|x_1| + |x_2|)(max\{|x_1|, |x_2|\})}{|x_1 x_2|}$$

Recall that for $x_1, x_2 \in \mathbb{C}$, we have that $|x_1 x_2| = |x_1||x_2|$. Now we must distinguish the two cases. If $|x_1| > |x_2|$, we have:

$$\kappa(x) = \frac{|x_1|^2 + |x_2||x_1|}{|x_2||x_1|} = \frac{|x_1|}{|x_2|} + 1.$$

If $|x_1| \leq |x_2|$, we have:

$$\kappa(x) = \frac{|x_2|^2 + |x_2||x_1|}{|x_2||x_1|} = \frac{|x_2|}{|x_1|} + 1.$$

We see that this problem becomes ill-conditioned when $|x_1| \gg |x_2|$ or $|x_2| \gg |x_1|$.

(c) $f(x) = (x - 2)^9$

$$\kappa(x) = \frac{||J_0|| \cdot ||x||}{||f(x_0)||}$$

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

$$\kappa(x) = \frac{|9(x-2)^8| \cdot |x|}{|(x-2)^9|} = \frac{|9x|}{|x-2|}.$$

We see this problem becomes ill conditioned when $x = 2$.

12. Note that the function $f(x) = (x - 2)^9$ in Pt. c of Problem 9 can also be expressed as $g(x) = x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512$. Mathematically, these two functions are identical.

From part c in Problem 11 we found that $f(x) = (x - 2)^9$ becomes ill conditioned when $x = 2$, i.e., the condition number, $\kappa(x = 2)$, is large. This condition number tells us how much precision of accuracy is lost by arithmetic methods when we calculate values from the function. This problem is ill conditioned at $x = 2$, meaning the loss of accuracy represented by the condition number is high enough to mess up calculations as we can see by comparing $f(x)$ and $g(x)$ for values close to $x = 2$. Computing $g(x)$ involves many more floating point operations, which is why the function is much less accurate compared to $f(x)$ near this ill-conditioned point. In general terms, for small changes in the inputs near $x = 2$, there is a large change in the resulting function which involves more floating point operations because that function is ill-conditioned. This is why we get such different y values for $g(x)$ when compared to $f(x)$ near $x = 2$.

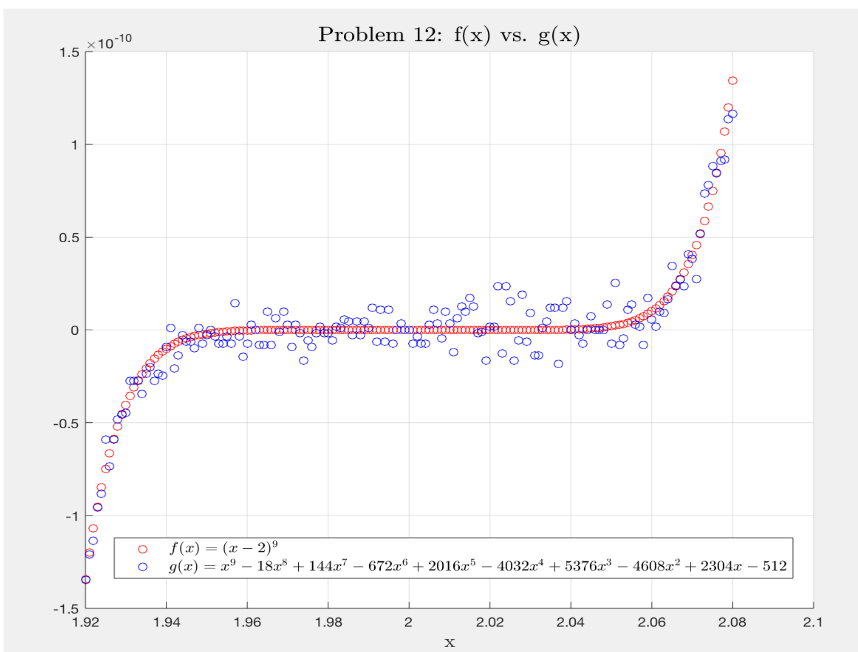

Figure 1: Plot showing comparison between f(x) and g(x).