Assignment Two

November 9, 2019

MCSC 6020G
Fall 2019
Submitted by Derick Smith

Question One:

Prove matrix A is nonsingular

A matrix $A \in \mathbb{C}^{\text{nxn}}$, where,

$$|a_{j,j}| > \sum_{i=1;i\neq j}^{n} |a_{i,j}| \qquad (j \in [1,n]) \in \mathbb{Z}$$

a strictly column diagonally dominant matrix that is nonsingular.

Proof preparation:

If $n = 1$, $det(A) = a_{1,1}$. Otherwise:

Properties of determinants to be used in proof. First, $det(A) = det(A^T)$. If $det(A^T) \neq 0 \Rightarrow det(A) \neq 0$. Second, if $A^T$is decomposed into so that $A^T = LU$, where, U diagonal elements are all nonzero then $det(A^T) = det(L) \cdot det(U) \neq 0$.

In the process of Gaussian elimination used in LU decomposition of $A^T = \{a_{i,j}\}_{\forall i,j}$, to zero all first column elements except row one, the resulting row, $a'_{i,j}$, is defined by:

$$a'_{i,j} = a_{i,j} - \frac{a_{1,j}}{a_{1,1}} \cdot a_{i,1}$$

Observe for some arbitrary row, $a_{i,j} \qquad i \neq 1$,

$$a'_{i,1} = a_{i,1} - \frac{a_{1,1}}{a_{1,1}} \cdot a_{i,1} = a_{i,1} - a_{i,1} = 0$$

For the transpose of a strictly column diagonally dominant matrix, A^T, is strictly row diagonally dominant matrix:

$$|a_{i,i}| > \sum_{j=1;j\neq j}^{n} |a_{i,j}| \qquad (i \in [1,n]) \in \mathbb{Z}$$

Proof:
Decompose $A^T = DM$, where,

$$M = \{m_{i,j}\}_{\forall i,j} = \left\{ \frac{a_{i,j}}{a_{i,i}} \right\}$$

Our strictly row diagonally dominant inequality becomes,

$$(|m_{i,i}| = 1) > \sum_{i=1;i\neq j}^{n} |m_{i,j}| \qquad (i \in [1,n]) \in \mathbb{Z}$$

For matrix $D = \{m_{i,j}\}_{\forall i,j}$, $d_{i,i} = \dfrac{1}{a_{i,i}}$ and $d_{i,j} = 0$ for all $i \neq j$.

For any row,

$$1 > \sum_{j=1;j\neq i}^{n} |m_{1,j}| \tag{1}$$

$$|m_{i,1}| > |m_{i,1}| \cdot \sum_{j=1;i\neq j}^{n} |m_{i,j}| \tag{2}$$

$$|m_{i,1}| > \left[\sum_{j=1;i\neq j}^{n} |m_{i,j}| \cdot |m_{i,1}| \right] \tag{3}$$

From that inequality and for any row greater than one,

$$|m_{i,i}| > \left(\sum_{j=1;i\neq j}^{n} |m_{i,j}| \right) \tag{1}$$

$$|m_{i,i}| > \left(\sum_{j=2;i\neq j}^{n} |m_{i,j}| \right) + |m_{i,1}| > \left[\left(\sum_{j=2;i\neq j}^{n} |m_{i,j}| \right) + \left(\sum_{j=1;i\neq j}^{n} |m_{i,j}| \cdot |m_{i,1}| \right) \right] \tag{2}$$

$$|m_{i,i}| > \left[\left(\sum_{j=2;i\neq j}^{n} |m_{i,j}| \right) + \left(\sum_{j=2;i\neq j}^{n} |m_{i,j}| \cdot |m_{i,1}| \right) + |m_{1,j} \cdot m_{i,1}| \right] \tag{3}$$

$$|m_{i,i}| - |m_{1,j} \cdot m_{i,1}| > \left[\sum_{j=2;i\neq j}^{n} |m_{i,j}| + |m_{i,j}| \cdot |m_{i,1}| \right] + |m_{1,j} \cdot m_{i,1}| \tag{4}$$

$$|m_{i,i}| - |m_{1,j} \cdot m_{i,1}| > \left[\sum_{j=2;i\neq j}^{n} |m_{i,j}| + |m_{i,j}| \cdot |m_{i,1}| \right] \tag{5}$$

$$|m_{i,i} - m_{1,j} \cdot m_{i,1}| \geq \qquad |m_{i,i}| - |m_{1,j} \cdot m_{i,1}| > \left[\sum_{j=2;i\neq j}^{n} |m_{i,j}| + |m_{i,j}| \cdot |m_{i,1}| \right] \tag{6}$$

$$|m_{i,i}| - |m_{1,j} \cdot m_{i,1}| > \left[\sum_{j=2;i\neq j}^{n} |m_{i,j}| + |m_{i,j}| \cdot |m_{i,1}| \right] \tag{7}$$

$$|m_{i,i}| - |m_{1,j} \cdot m_{i,1}| > \left[\sum_{j=2;i\neq j}^{n} |m_{i,j}| + |m_{i,j}| \cdot |m_{i,1}| \right] \geq \left[\sum_{j=2;i\neq j}^{n} |m_{i,j} - m_{i,j} \cdot m_{i,1}| \right]$$

$$|m_{i,i}| - |m_{1,j} \cdot m_{i,1}| > \qquad |m_{i,i} - m_{1,j} \cdot m_{i,1}| \geq \left[\sum_{j=2;i\neq j}^{n} |m_{i,j} - m_{i,j} \cdot m_{i,1}| \right] \tag{8}$$

$$|m_{i,i} - m_{1,j} \cdot m_{i,1}| \geq \left[\sum_{j=2;i\neq j}^{n} |m_{i,j} - m_{i,j} \cdot m_{i,1}| \right] + 0 \tag{9}$$

$$|m_{i,i} - m_{1,j} \cdot m_{i,1}| \geq \left[\sum_{j=2;i\neq j}^{n} |m_{i,j} - m_{i,j} \cdot m_{i,1}| \right] + m_{i,1} \tag{10}$$

$$|m_{i,i} - m_{1,j} \cdot m_{i,1}| \geq \left[\sum_{j=1;i\neq j}^{n} |m_{i,j} - m_{i,j} \cdot m_{i,1}| \right] \tag{11}$$

Finally,

$$|m'_{i,i}| > \sum_{j=1;i\neq j}^{n} |m'_{i,j}|$$

The Gaussian elimination operations result in a matrix with first column elements equal to zero, except row one and remains strictly row diagonally dominant. This means that every diagonal element is nonzero.

The matrix A, from which M was derived, was of arbitrary dimension. Let $M' = m'_{i,j}, i \in [2,n]_{\mathbb{Z}}, j \in [2,n]_{\mathbb{Z}}$, M' is of arbitrary size and satisfies the arguments used to perform the row elimination of every row's first column elements, for all rows greater than one. The same operations could be performed on M' so that $M' \in \mathbb{C}^{\text{(n-1)x(n-1)}} \rightarrow M'' \in \mathbb{C}^{\text{(n-2)x(n-2)}}$, where M'' is strictly row diagonally dominant. Inductively, $M^k \in \mathbb{C}^{\text{(n-k)x(n-k)}} \rightarrow M^{k+1} \in \mathbb{C}^{\text{(n-k-1)x(n-k-1)}}$, $\forall k < n-1$ ($M^{n-1} \in \mathbb{C}^{\text{1x1}}$). Every iteration is strictly row diagonally dominant.

The matrix M can decomposed into a product of triangular matrices with nonzero diagonal elements and the product of all matrix determinants are nonzero, therefore, matrix A is nonsingular.

Question Two:

a) For matrix $A \in \mathbb{C}^{\mathbf{n}\times\mathbf{n}}$, $A = yx^T$, show $\|A\|_2 = \|x\|_2 \|y\|_2$

Although a proof of this was found, not every step was understood so it is not included. A substantial number of computations were performed to show, as was requested[11], that $\|A\|_2 = \|x\|_2 \|y\|_2$. A 40000 computarions were run. Arrays were sized from $n = 1$ to $n = 200$, each value randomly generated per run with a uniform distribution between -1000 to 1000. Each array size was computed 200 times.

With the hypotheses:

$$H_0 : \qquad \hat{\mu} = \mu = 0$$

$$H_a : \qquad \hat{\mu} \neq \mu = 0$$

The calculated test statistic probability, $p = 2.6 \cdot 10^{-26}$, provides enough room to set the acceptance tolerance, α, extremely small. The highest scientific standard being $\alpha = 3 \cdot 10^{-7} >> p$, would fail to reject the null hypothesis, H_0. Therefore, it can be concluded that $\|A\|_2 = \|x\|_2 \|y\|_2$.

[1];)

b) Show $K_p(AB) \leq K_p(A)K_p(B)$

$$K_p(AB) = \|AB\|_p \left\|(AB)^{-1}\right\|_p \tag{1}$$

$$\|AB\|_p \left\|(AB)^{-1}\right\|_p = \|AB\|_p \left\|B^{-1}A^{-1}\right\|_p \tag{2}$$

$$\|AB\|_p \left\|B^{-1}A^{-1}\right\|_p \leq \|A\|_p \|B\|_p \left\|B^{-1}\right\|_p \left\|A^{-1}\right\|_p \tag{3}$$

$$\left[\|A\|_p \left\|A^{-1}\right\|_p\right] \left[\|B\|_p \left\|B^{-1}\right\|_p\right] = K_p(A)K_p(B) \tag{4}$$

$$\therefore \quad K_p(AB) \leq K_p(A)K_p(B) \tag{5}$$

Question Three:

$$f(x) = \sum_{i=1}^{N} \frac{a_i^2}{(b_i^2 + x)^2}$$

a) Show f **is monotonically decreasing, convex on** $[0, \infty)$

$$f(x) = \sum_{i=1}^{N} a_i^2 (b_i^2 + x)^{-2} > 0, \forall x \in [0, \infty) \tag{1}$$

$$f'(x) = -2 \sum_{i=1}^{N} a_i^2 (b_i^2 + x)^{-3} < 0, \forall x \in [0, \infty) \tag{2}$$

$$f''(x) = 6 \sum_{i=1}^{N} a_i^2 (b_i^2 + x)^{-4} > 0, \forall x \in [0, \infty) \tag{3}$$

$\therefore$ the function f is monotonically decreasing and convex on $[0, \infty)$.

b) Newton method observations

As shown in the figures in part d, the Newton method appears linear for a substatial number of iterations before curving sharply towards the tolerance threshold. This suggests that as k increases, the Newton method might in fact converge quadratically as might be intuitive for a polynomial function but not necessarily for a rational function. The number of iterations depended on δ and ranged from approximately 10 to approximately 40. The behavior of the Newton method is discussed further in part d.

c) Develop Newton-like method

Pseudo-code:

1. Find initial x_0 using a helper function; a bisection-like method can find when $f(x) < 1$ exceptionally quickly for a rational function.

2. For each k^{th} iteration, make model $g_k(x) = \dfrac{A_k}{B_k + x}$, solving for A_k and B_k using a point of intersection, $g_k(x) = f(x)$, and having the same slope, $g_k'(x) = f'(x)$:

$$\begin{aligned} g_k(x) &= A_k \cdot (B_k + x)^{-1} &&= f(x) && (1)\\ g_k'(x) &= -A_k \cdot (B_k + x)^{-2} &&= f'(x) && (2)\\ A_k &= f(x) \cdot (B_k + x) && && (3)\\ B_k &= \frac{-f(x)}{f'(x)} - x && && (4) \end{aligned}$$

3. Using $g_k(x_{k+1}) = 0$, solve for x_{k+1}:

$$\begin{aligned} g_k(x_{k+1}) &= A_k \cdot (B_k + x_{k+1})^{-1} = 0 && (1)\\ x_{k+1} &= \frac{A_k}{f(x_k)} - B_k && (2) \end{aligned}$$

1. If $|x_k - x_{k+1}| < TOL_x$ or $g_k(x_{k+1}) > f(x_{k+1})$, switch to Newton method.

2. Iterate until $|f(x_k) - \delta| < TOL_f$.

d) Compare two methods (include bisection)

The Newton method was never shown to surpass the Newton-like method (with helper function). With δ "large," the Newton method outperformed the Bisection method but neither had ever outperformed the Newton-like method.

As δ decreased in size the Bisection method began to cross the tolerance threshold before the Newton method. The reason this happend was the helper function which was used with the Bisection method. The helper function took advantage of an exponential term to find when $f(2^n) < 0$ so it was able to cover more ground and avoid the sharp slope of $f(x)$ near 0. The other methods, without helper functions, are bound to spend time at the start climbing away from zero.

That being said the Newton and Newton-like methods could benefit from the use of helper functions.

In the figures below: the red markers represent the Newton method; the blue markers represent the Newton-like method with squares representing the rational model $g(x) = \dfrac{A}{B+x}$ and triangles when the helper function (Newton) is called); and, the green markers represent the Bisection method. For each N, a new set of random values were generated for a_i and b_i. The tolerance of $f(x_k)$ was set to $TOL_f = 10^{-16}$.

The size of N did not seem to have much influence on the rate of convergence for any particular method. However, the figures only show a single instance at each N per δ. A statistical fit would need to accompany a large sample for each combination to make a confident statement on the impact of N.

On the other hand, δ had a noticable affect on the rates of convergence. The convergence rate of the Newton-like method might seem linear, however, this is due to the limitations of due to computer precision, $|TOL_f - \delta| \to 0$, and the selection of δ. If precision were not a contraint and δ could be chosen arbitrarily small, the rate of convergence of the Newton-like method might in fact be quadratic. A proof could be possible and an interesting pursuit but it is beyond this assignment.

Notice the point in which the blue markers switch from squares to triangles on the residual axis. It appears that just before the rational model is approaching the same magnitude of the delta, it summons the regular Newton method. It stalls out. A proof is beyond the time limitations, however, a good starting point might be the analysis of the inequality $g_k(x) \leq f(x)$ and that $f(x) \in \Theta\left(\dfrac{1}{x^2}\right) \to 0$ slower than $g_k(x) \in \Theta\left(\dfrac{1}{x}\right) \to 0$ as $x \to \infty$.

Figure 1: $\delta = 10^{-5}$ and $N = 10, 10^2, 10^3, 10^4$

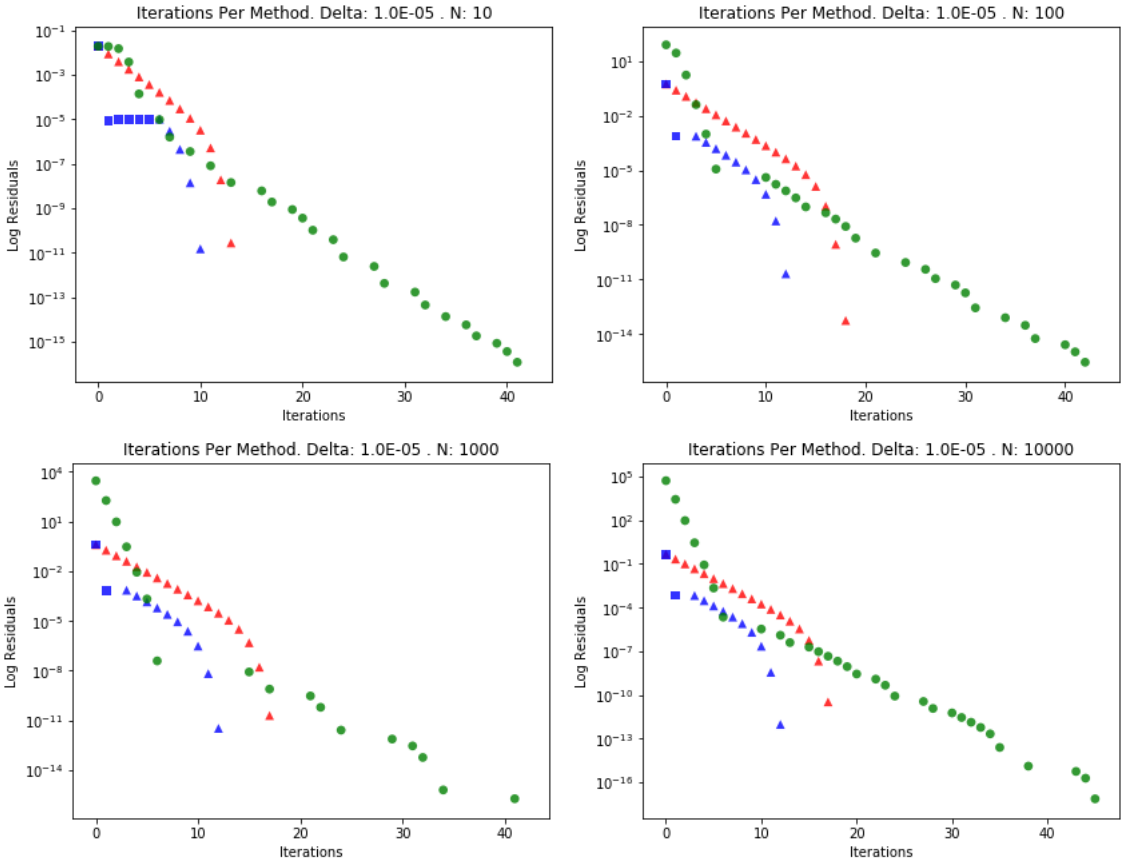


Figure 2: $\delta = 10^{-8}$ and $N = 10, 10^2, 10^3, 10^4$

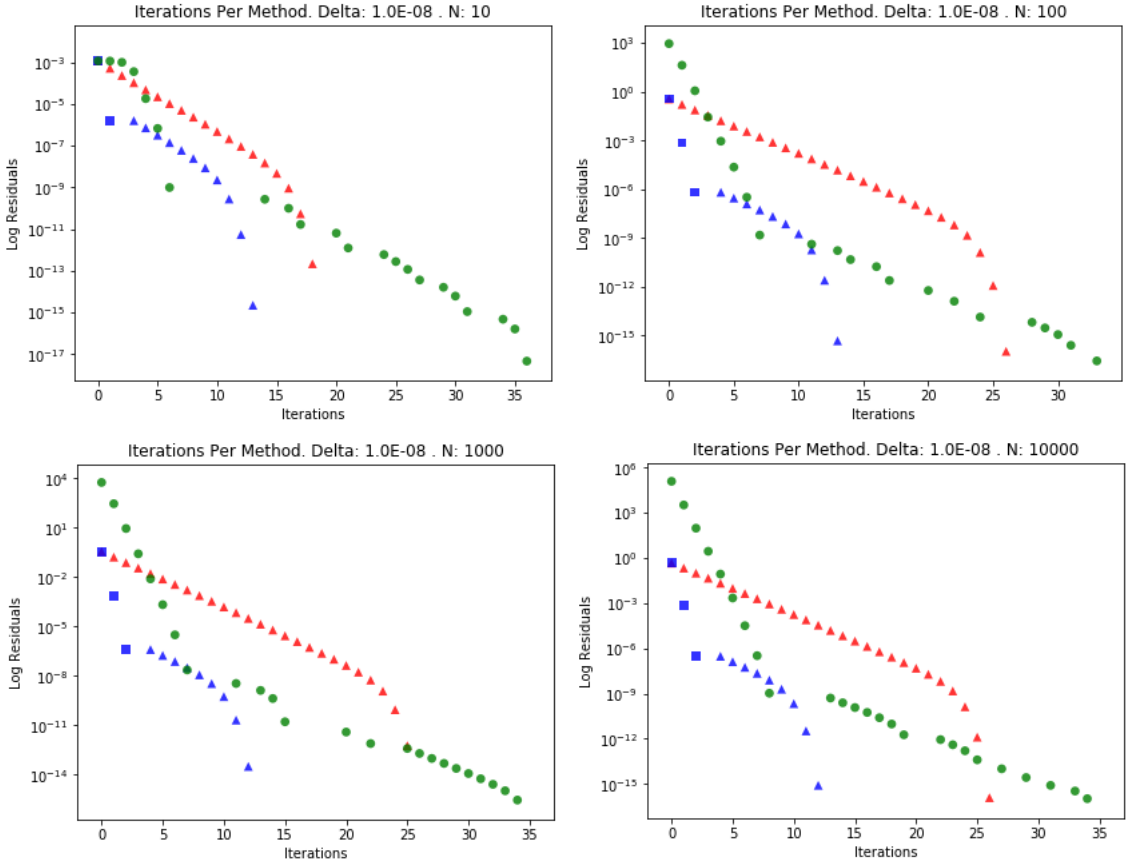


Figure 3: $\delta = 10^{-11}$ and $N = 10, 10^2, 10^3, 10^4$

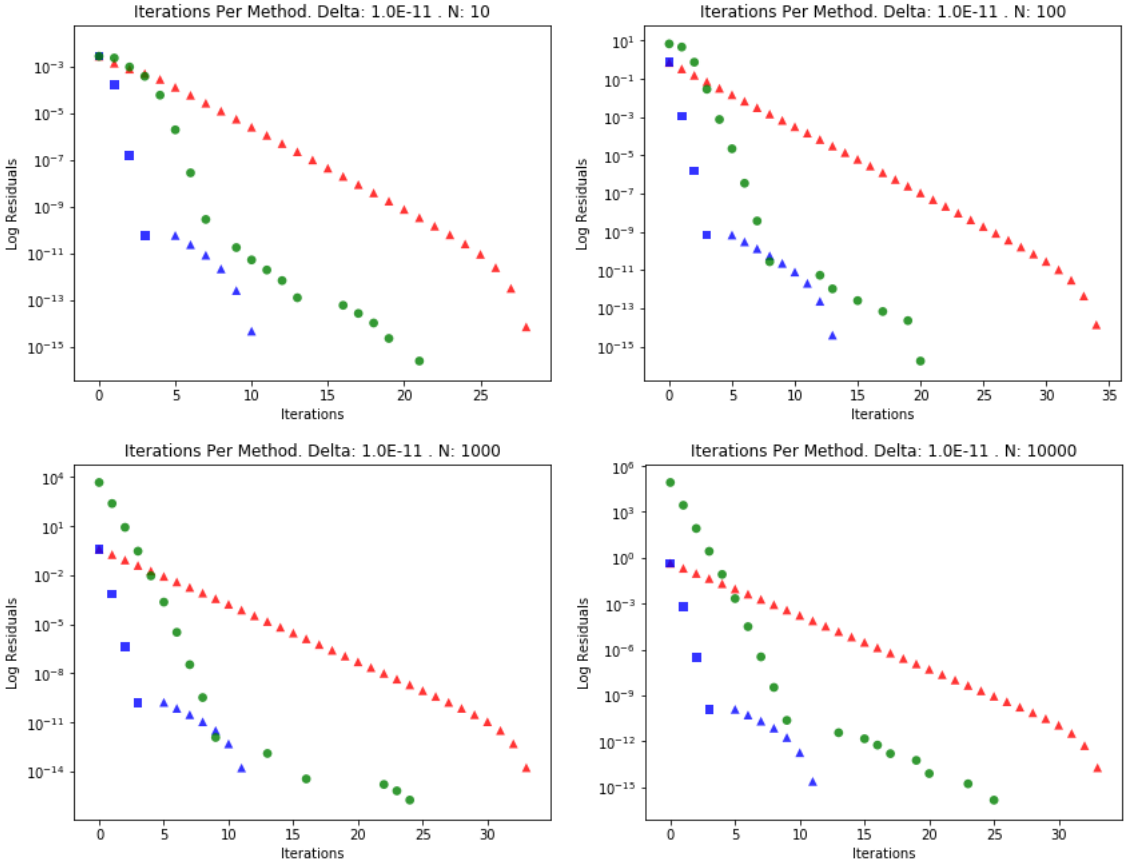
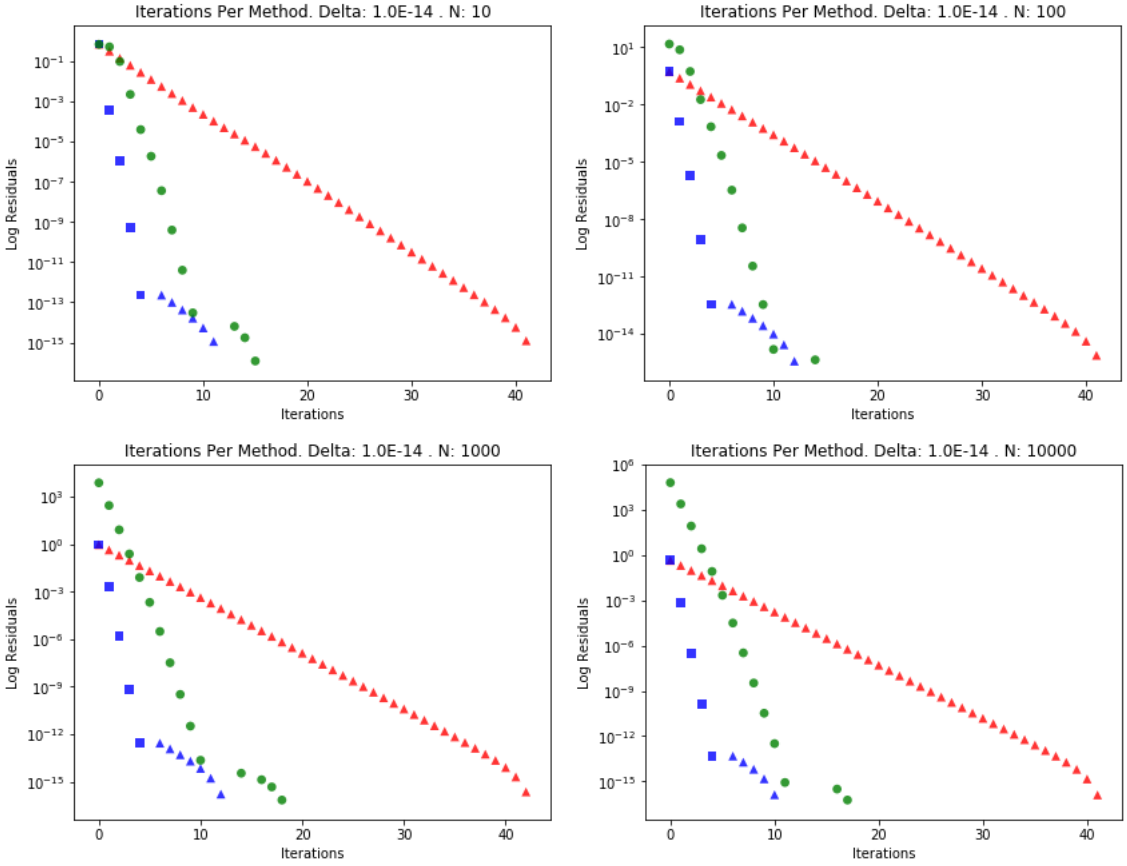


Figure 4: $\delta = 10^{-14}$ and $N = 10, 10^2, 10^3, 10^4$



Sources:

1. Professor
2. Classmates