COL726 A4

Raval Vedant Sanjay 2017CS10366

Ans 1 $\Rightarrow$ We know that the given non symmetric matrix A^* can be broken as : $A^* = \underbrace{\left(\frac{(A^*) + (A^*)^T}{2}\right)}_{\text{Symmetric}} + \underbrace{\left(\frac{(A^*) - (A^*)^T}{2}\right)}_{\text{Antisymmetric}}$

$\left(A = \frac{(A^*) + (A^*)^T}{2}\right)$ We call this matrix as A ahead.

$\Rightarrow$ It can also be observed that the rayleigh coefficient for a matrix is only affected by it's ~~symmetric~~ Symmetric component! (Matrix A here)

$\Rightarrow$ We know that here, the Rayleigh Quotient for a Symmetric A is given as: $\boxed{r(v) = \dfrac{v^T A v}{v^T v}}$

$\Rightarrow$ For a given non sym matrix A^*

$\boxed{A = \frac{1}{2}\left(A^* + (A^*)^T\right)}$

(Symmetric Component of given input matrix)

$\hookrightarrow$ Input matrix is A^* here.

$\Rightarrow$ Thus, We get:

$$\delta r = r(v+\delta v) - r(v) = \frac{(v+\delta v)^T A (v+\delta v)}{(v+\delta v)^T (v+\delta v)} - \frac{v^T A v}{v^T v}$$

$$v^T v + v^T \delta v + \delta v^T v + \delta v^T \delta v$$

$$\hookrightarrow v^T v + 2v^T \delta v + \|\delta v\|^2$$

$\Rightarrow$ Thus, we have: $$\delta r = \frac{(v+\delta v)^T A (v+\delta v)}{v^T v + 2v^T \delta v} - \frac{v^T A v}{v^T v}$$

Since $$\frac{1}{x+\delta x} = \frac{1}{x} - \frac{\delta x}{x^2} + O(|\delta x|^2) \Rightarrow \boxed{\frac{1}{v^T v + 2v^T \delta v} = \frac{1}{v^T v} - \frac{2v^T \delta v}{(v^T v)^2}}$$

(Ignoring 2nd order terms)

$\Rightarrow$ Thus:

$$\delta r = (v+\delta v)^T A (v+\delta v)\left(\frac{1}{v^T v} - \frac{2v^T \delta v}{(v^T v)^2}\right) - \frac{v^T A v}{v^T v}$$

$$\left(v^T A v + v^T A \delta v + \delta v^T A v\right)$$

$$\hookrightarrow \boxed{\delta r = \left(v^T A v + v^T A \delta v + \delta v^T A v\right)\left(\frac{1}{v^T v} - \frac{2v^T \delta v}{(v^T v)^2}\right) - r(v)}$$

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

+ + + + =

Q___ ___

$$\Rightarrow \delta r = \frac{v^T A v}{v^T v} + \frac{v^T A \delta v}{v^T v} + \frac{\delta v^T A v}{v^T v} - \frac{2 (v^T A v)(v^T \delta v)}{(v^T v)^2} - r(v)$$

$r(v)$

(Ignore the term involving δv twice as they would relate to $O(\|\delta v\|^2)$)

$$\Rightarrow \delta(r) = \frac{v^T A \delta v}{v^T v} + \frac{\delta v^T A v}{v^T v} - \frac{(v^T A v)(v^T \delta v)}{(v^T v)\cdot(v^T v)} - \frac{(v^T A v)(\delta v^T v)}{(v^T v)\cdot(v^T v)}$$

$r(v)$ $\quad$ $r(v)$

$$\therefore \boxed{\delta(r) = \frac{1}{v^T v}\left(\left(v^T (A\delta v - r(v)\,\delta v)\right) + \left(\delta v^T (Av - r(v)\,v)\right)\right)}$$

⇒ Now if $\boxed{\delta(\gamma) = 0}$

$$\longrightarrow \underbrace{v^T\left(A\,\delta v - \gamma(v)\,\delta v\right)}_{①} + \underbrace{\delta v^T\left(Av - \gamma(v)\,v\right)}_{②} = 0$$

$$② = v^T\left(A^T\delta v - \gamma(v)\,\delta v\right)$$

(Since it is a real so we can take transpose)

$$v^T\left(\frac{A+A^T}{2}\right)\delta v = v^T\gamma(v)\,\delta v \Longleftarrow$$

$\Downarrow$ b/c A is symmetric

$$\boxed{A\,\delta v = \gamma(v)\,\delta v} \leftarrow (\alpha)$$

⇒ Similarly by taking transpose of term ① rather than ② $\Rightarrow$ $\boxed{Av = \gamma(v)\,v} \leftarrow (\beta)$

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question Sub-question No.

Q___ ___

| + | + | + | + | = |

$\Rightarrow$ From α & β, we can say that these eq's are satisfied when v is an eig vector of A & dv is along v.

$\Rightarrow$ $\boxed{\text{Thus } v \text{ is an eig vector of } (A^*) + (A^*)^T \text{ \& } dv \text{ is along } v}$

(A^* was the Non Symetric matrix)

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

Q

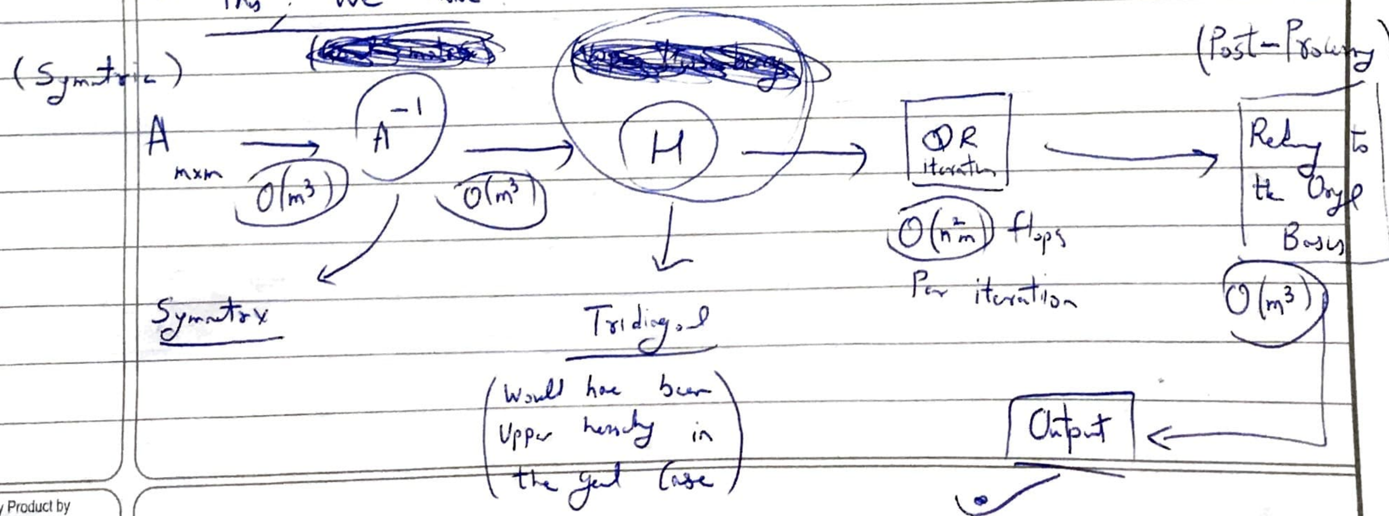
Ans 2. (a) ⇒ As described in Heath, the algorithm for QR iteration can be used to find (n) largest Eig Values (with their Eig Vectors) for a matrix A.

⇒ Now, this is equivalent to find (n) smallest Eig Values for the matrix A^{-1}. So this is what we would do.

⇒ So we will obtain A^{-1} from A first in $O(m^3)$ time. Then we will convert this matrix into Tridiagonal form as the the work per QR iteration would be reduced. Precisely, it will be $O(n^2 m)$ flops per iteration if we use a Tridiagonal matrix. This conversion to Tridiagonal can be done in $O(m^3)$.

⇒ Then finally, we would also need to convert the obtained Eig Vectors to the Original Basis by reverting the matrix. This can also be done with the required bound of $O(m^3)$. And so we get the proper Eigen Vectors finally.

⇒ Thus, we have:



પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

Q___ ___

Ans 3 : (a) ⇒ We know that A is an upper Hessenberg matrix.

⇒ To prove that $A^{(k)}$ is upper hessenberg & has same Eig Vals as A, we will use induction.

→ Base Case : $k=0 \Rightarrow \boxed{A^{(0)} = A}$ (True)

→ Induct hypothes : Let that statement be True for $A^{(k-1)}$.

→ To Prove : $A^{(k)}$ has the same Eig Vals as A & is upper hessenberg as well.

$\Downarrow$

→ We have : $Q^{(k)} R^{(k)} = A^{(k-1)} - \mu^{(k)} I$

$\Downarrow$

$$\left(Q^{(k)}\right)^* \left(Q^{(k)} R^{(k)}\right)\left(Q^{(k)}\right) = \left(Q^{(k)}\right)^* A^{(k-1)} Q^{(k)} - \left(Q^{(k)}\right)^* \left(\mu^{(k)} I\right) Q^{(k)}$$

$R^{(k)} Q^{(k)}$ → $A^{(k)} - \mu^{(k)} I$

$\Downarrow$

→ $A^{(k)} - \mu^{(k)} I = \left(Q^{(k)}\right)^* A^{(k-1)} Q^{(k)} - \left(\mu^{(k)} I\right)\left(\left(Q^{(k)}\right)^* Q^{(k)}\right)$

$\Downarrow$

$\boxed{A^{(k)} = \left(Q^{(k)}\right)^* A^{(k-1)} Q^{(k)}}$ ⇒ Similarity transformation

$\Downarrow$

$A^{(k)}$ & $A^{(k-1)}$ have the same Eig Values.

પ્રશ્ન / પેટા પ્રશ્ન
ક્રમાંક
Question
Sub-question No.

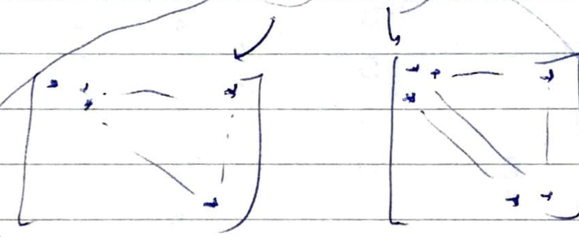
Q___ ___

$\square + \square + \square + \square + \square = \square$

→ Now, using $A^{(n-1)}$ to be Upper Hessenberg, we get:
$Q^{(n)} R^{(n)}$ is also upper Hessenberg

→ Now we know that the Q of the QR factorization for an Upper Hessenberg matrix is also an upper hessenberg matrix. So, we have $Q^{(n)}$ is an upper hessenberg matrix.

→ Now: $A^{(n)} = R^{(n)} Q^{(n)} + \mu^{(n)} I$



The Product is Upper hessenberg $\Rightarrow$ $\boxed{A^{(n)} \text{ is Upper Hessenberg}}$

* Thus we Prove the required statement by Induction.

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question Sub-question No.

Q___ ___

(b) $\Rightarrow$ Since $\mu(k)$ is an Eig Value of $A^{(k-1)}$, from

$$Q^{(k)} R^{(k)} = A^{(k-1)} - \mu^{(k)} I$$

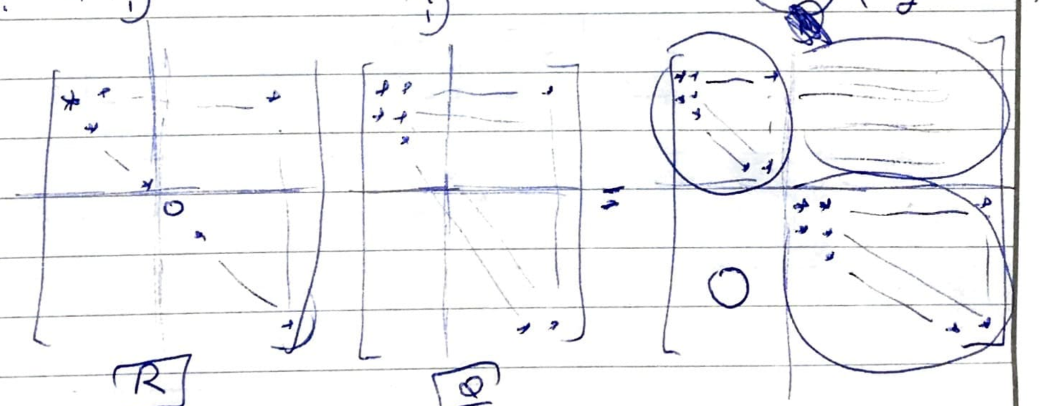
$\rightarrow$ Singular matrix

$R^{(k)}$ is a Singular matrix ($\because Q^{(k)}$ is not Singular)

$\Rightarrow$ So we can say that there would be atleast one diagonal Element in $R^{(k)}$ would be 0.

$\Rightarrow$ So we get: index (i) (i) (i+1) (Till index (i) & index 0 diagonal)

$R^{(k)} Q^{(k)} =$



$\Downarrow$

$$\begin{bmatrix} A_{11} & A_{12} \\ & A_{22} \end{bmatrix}$$

with A_{11}, A_{22} being Upper Hessenberg

Note that the multiplication with $\mu^{(k)} I$ will only alter the main diagonal of the resulting matrix & so this formulation will still hold.

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

Q__ __

No. 4:

$\Rightarrow$ We know that x^* is a fixed point for g & $G(x)$ is diagonalizable for all x.

$\Rightarrow$ Let us have: $\rho(G(x^*)) < 1$ ② & $\exists$ an induced norm s.t. $\|G(x^*)\| < 1$ ③

$\Rightarrow$ We first try to prove that ② & ③ are equivalent.

* **To Prove:** ③ $\Rightarrow$ ②

$\rightarrow \rho(G(x^*)) \Rightarrow \max\{|\lambda_1|, \ldots, |\lambda_m|\} \Rightarrow \lambda_1$ (WLOG)

max Eig Value.

$\rightarrow$ So: $\rho(G(x^*)) = |\lambda_1| = \dfrac{\|G(x^*)v_1\|}{\|v_1\|} \le \|G(x^*)\|$

v_1: Eig Vector corresponding to λ_1.

$\rightarrow$ So: $\boxed{\rho(G(x^*)) \le \|G(x^*)\|}$ $\xrightarrow{\text{Using ③}}$ $\boxed{\rho(G(x^*)) < 1}$ H.P.

* **To Prove:** ② $\Rightarrow$ ③

$\rightarrow$ Since $G(\bar{x})$ is diagonalizable $\Rightarrow$ $G(x^*) = Q^{-1}DQ$

$\rightarrow$ Let $\exists v$ s.t. $\|G(x^*)\| = \dfrac{\|G(x^*)v\|}{\|v\|}$

A Quality Product by

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question Sub-question No.

Q___ ___

$\square + \square + \square + \square + \square = \square$

→ It gn : $\|G(x^*)\| = \dfrac{\|Q^* D Q v\|}{\|v\|} = \dfrac{\|DQv\|}{\|v\|}$

($\because$ Q^* is unity)

→ Ths : $\|G(x^*)\| = \dfrac{\|DQv\|}{\|v\|} \leq \|D\| \cdot \dfrac{\|Qv\|}{\|v\|} \leq \|D\|$

$\rho(G(x^*)) \Longleftarrow \max\{|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|\}$

→ So : $\boxed{\|G(x^*)\| \leq \rho(G(x^*))}$

by (2) → $\boxed{\|G(x^*)\| < 1}$ $\boxed{\text{H.P.}}$

$\Rightarrow$ we now have ② $\Leftrightarrow$ ③, & so we need to prove ①.

$\Rightarrow$ Here, we define $\gamma(t) = g(x + t(y-x))$

$$\Rightarrow \|g(y) - g(x)\| = \left\| \int_0^1 \gamma'(t)\, dt \right\| \leq \int_0^1 \|\gamma'(t)\|\, dt$$

$$\Downarrow$$

$$\|g(y) - g(x)\| \leq \int_0^1 \|y-x\| \cdot \|G(x + t(y-x))\|\, dt$$

$\Rightarrow$ Now as we know that $\|G(x^*)\| < 1$, so if $(x + t(y-x))$ is in a close ~~neigh~~ neighbourhood of x^* then $\|G(x + t(y-x))\|$ will also be bounded ($\leq c$)

$$\Rightarrow \text{So: } \|g(y) - g(x)\| \leq \int_0^1 \|y-x\| \cdot c\, dt \leq c\|y-x\|$$

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

	+		+		+		+		=	

Q___ ___

⇒ Here if we take $x = x^*$ ⇒ $g(x) = x^*$. Then for $y = x_k$, $g(y) = g(x_k) = x_{k+1}$.

⇒ $\boxed{\| x_{k+1} - x^* \| < c \cdot \| x_k - x^* \|}$

→ This Fixed Point itr will happen with the distance b/w x_k & x^* getting decreased with each iteration (c will be <1 here since $\| G(x^*) \| < 1$) Using ③

$\boxed{\text{The Solution Converges}}$ ⟸

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

☐ + ☐ + ☐ + ☐ + ☐ = ☐

Q__ __

Ans 5 :

$\Rightarrow$ Let us consider the Jacobian update $\boxed{B \leftarrow B + \Delta B}$.
For the initial matrix B, ΔB shall be a rank-1 matrix such that : $(B + \Delta B)s = y$

$\Rightarrow$ We know that the change ΔB by broyden's method is :

$$\boxed{\Delta B^{*} = \frac{r s^T}{s^T s}} \quad \text{where} \quad r = y - Bs$$

$\Rightarrow$ Now from $(B + \Delta B)s = y \Rightarrow \boxed{\Delta B\, s = y - Bs = r}$

$$\Downarrow$$

$$\boxed{\Delta B \left(\frac{s s^T}{s^T s}\right) = \frac{r s^T}{s^T s} = \Delta B^{*}} \Leftarrow \boxed{\Delta B\, s s^T = r s^T}$$

$\Rightarrow$ From this, we get : $\boxed{\left\| \Delta B \cdot \left(\frac{s s^T}{s^T s}\right) \right\|_F = \| \Delta B^{*} \|_F}$

$$\Downarrow$$

$$\boxed{\| \Delta B \|_F \cdot \left\| \frac{s s^T}{s^T s} \right\|_F \geq \| \Delta B^{*} \|_F}$$

$$\Downarrow$$

$$\boxed{\| \Delta B \|_F \cdot \left(\frac{\| s s^T \|_F}{|s^T s|} \right) \geq \| \Delta B^{*} \|_F}$$

A Quality Product by
સ્ક્રોલ

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question Sub-question No.

Q___ ___

$\Rightarrow$ Now : We know that : $\boxed{\|A\|_F = \sqrt{Tr(AA^T)}}$

$\Downarrow$

$$\boxed{\|SS^T\|_F = \sqrt{Tr(S(S^TS)S^T)} = \sqrt{S^TS} \cdot \sqrt{Tr(SS^T)}}$$

$\Rightarrow$ Now, if we have : $S = \begin{bmatrix} s_1 \\ \vdots \\ s_m \end{bmatrix} \Rightarrow \boxed{Tr(SS^T) = s_1^2 + \ldots + s_m^2 = S^TS}$

$\Rightarrow$ Thus, we get : $\boxed{\|SS^T\|_F = \sqrt{S^TS} \cdot \sqrt{Tr(SS^T)} = (S^TS)}$ ← ②

$\Rightarrow$ Using ② in ①, we get : $\boxed{\|\Delta B\|_F \geq \|\Delta B^*\|_F}$

↳ Thus ΔB^* used by Broyden has the smallest Frobenius norm.

A 6 : (a) $\Rightarrow$ If we have $z = x_1 + ix_2 \Rightarrow$ $\boxed{\begin{aligned} P(z) &= a_n z^n + \ldots + a_1 z + a_0 \\ &= y_1 + iy_2 \end{aligned}}$

$\Rightarrow$ The above Equation is Equivalent to having a function $f : R^2 \to R^2$ such that : $\boxed{f(x_1, x_2) = (y_1, y_2)}$

A Quality Product by

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question Sub-question No.

Q___ ___

$\Rightarrow$ Using the definition of Jacobian : $J(f(x_1,x_2)) = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} \end{bmatrix}$

$\Rightarrow$ We know the Taylor Series as : $\boxed{p(z+dz) = p(z) + p'(z)\,dz}$ (0)

$\Rightarrow$ Thus we have : $\boxed{p\big((x_1+dx_1) + i(x_2+dx_2)\big) = (y_1+dy_1) + i(y_2+dy_2)}$ (1)

$\Rightarrow$ Using eq's (0) & (1), we get :

$$\boxed{(y_1+dy_1) + i(y_2+dy_2) = (y_1 + iy_2) + (p'(z))(dx_1 + i\,dx_2)}$$

$$\Downarrow$$

$$\boxed{dy_1 + i\,dy_2 = (p'(z)).(dx_1 + i\,dx_2)}$$

$\Downarrow$ Taking $p'(z) = \alpha + i\beta$

$$\boxed{dy_1 + i\,dy_2 = (\alpha\,dx_1 - \beta\,dx_2) + i(\alpha\,dx_2 + \beta\,dx_1)}$$

$$\Downarrow$$

$$\boxed{\begin{aligned} y_1 &= \alpha\,dx_1 - \beta\,dx_2 \\ y_2 &= \beta\,dx_1 + \alpha\,dy_2 \end{aligned}} \Rightarrow \boxed{J(f(x_1,x_2)) = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}}$$

with : $\boxed{\begin{aligned} \alpha &= Re\big(p'(x_1+ix_2)\big) \\ \beta &= Im\big(p'(x_1+ix_2)\big) \end{aligned}}$

(c) $\Rightarrow$ Here the function used was $p(z) = z^3 - z - 1$.

$\Rightarrow$ From the graph it can be seen that a complex distribution of points was obtained with respect to the solution towards which the Newton's method would converge. by a complicated fractal structure.

$\Rightarrow$ Though it's all related to the closeness of the starting point to the different solutions too.

The three solutions here were R_1, c_1, $\bar{c}_1$ $\rightarrow$ Conj. of each other

(Real : x_1
Imag : x_2)

R_1 $\downarrow$ (Positive) A real

$c_1, \bar{c}_1$ $\rightarrow$ Two complex. (Negative Real Part)

$\Rightarrow$ From the plot, we can see three big partitions obtained in the different region of the $x_1 - x_2$ plane corresponding to the closeness of the origin point to the actual solution.

પ્રશ્ન / પેટા પ્રશ્ન ક્રમાંક
Question
Sub-question No.

□ + □ + □ + □ + □ = □

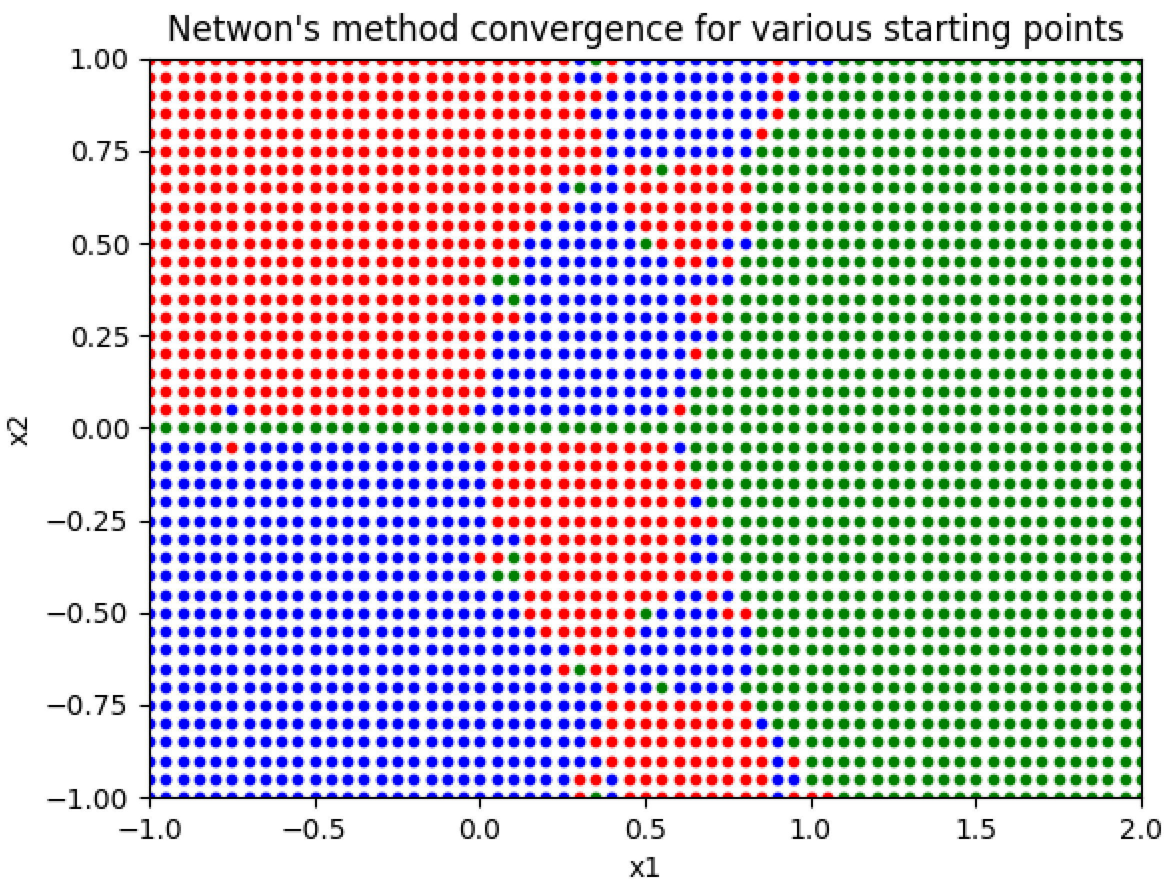
Q___ ___

$\Rightarrow$ The Plot even Seems as if the line of $X_1 = 0$ acts as a mirror to this distribution which also Supports the claim.

$\Rightarrow$ Even the Points on $X_1 = 0$ Seem to Converge to the Solution of R_1.

$\Rightarrow$ Also, it is not the only reason for such a Complicated fractal Structure. The other reasons lie in the way the different updates are Performed wrt to the Jacobian which isn't so Straightforward to imply that all the Points Converge to the nearest actual Solution.

$\Rightarrow$ Thus, even though a majority of Orig Points Converge to the nearest actual Solution, it is not the Case in general because of the irregularities within the Newton's method updates & So the Sensitive Points act as exceptions.


Netwon's method convergence for various starting points
x2
1.00
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
x1