複變數與多變數下牛頓法的討論

高中組數學科第二名

台灣省立武陵高級中學　　作　　者：王金龍
指導教師：盧澄根

一、研究動機與目的

一般初微的教科書中都有介紹方程式實數根的逼近法，然而複數根則未見提及，而且書上所提出的收斂條件也不完備，因此引發了研究的興趣，希望藉此研究擴充已知的方法，使其能夠解決一般根逼近的問題。

二、研究內容

㈠牛頓法：衆多逼近法中，牛頓法是效率很高的一種，如圖 1 。

$$\frac{f(X_n)-0}{X_n-X_{n+1}}=f'(X_n)$$

$$\Rightarrow X_{n+1}=X_n-\frac{f(X_n)}{f'(X_n)}$$

若 $X_n\to\beta$ ，則

$$\beta=\beta-\frac{f(\beta)}{f'(\beta)}$$

$$\Rightarrow f(\beta)=0$$

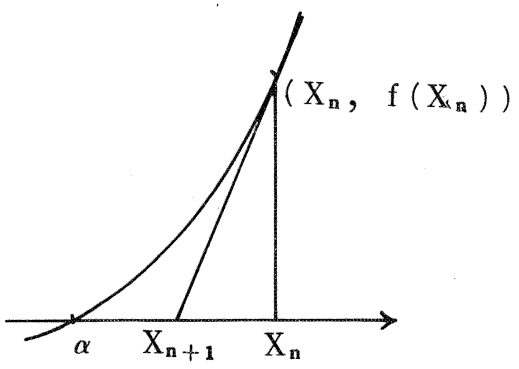

圖 1

一般書上給的收斂條件是：在根 α 近旁存在〔 a ， b 〕，$\alpha\in$〔 a ， b 〕，且 $\forall X\in$〔 a ， b 〕/ α 恒有 $f(x)\cdot f''(x)>0$ ，則 $X_n\to\alpha$ 。證明從略。但是考慮 $f(x)=\sin x$ ，則 $f'(x)=\cos x$，$f''(x)=-\sin x$ ，故 $f(x)\cdot f''(x)=-\sin^2 x\leq 0$ ，完全與前述的收斂條件不合，莫非 $x=0$ 這個根竟無法依牛頓法逼近得到？在㈡中將完整地解釋這個問題。

㈡單變數複數根之逼近

牛頓法可用來逼近複數根嗎？當然可以！由於在複變數函數的情況下，爲了可微分，因此以下都要求函數 f(Z) 在其各零點（根）可以微分（正則、解析）。

令 $Z \xrightarrow{\phi} Z - \dfrac{f(Z)}{f'(Z)}$ ， $Z_{n+1} = \phi(Z_n)$

若 $\lim\limits_{n \to \infty} Z_n$ 收斂，則必收斂至 f(Z) 之零點，此屬顯然，但問題在於：若 $f(\alpha) = 0$，能否在 α 附近找到一區間 Ω，使任何以 Ω 內之點爲起點之點列 Z_n 都收斂到 α？答案是肯定的，證明如下：

$\boxed{\text{複變函數之零點若可微分，則必可依牛頓法逼近得到。}}$

pf：由 $\phi(Z) = Z - \dfrac{f(Z)}{f'(Z)}$

$$\phi'(Z) = 1 - \frac{f'(Z)f'(Z) - f(Z)f''(Z)}{[f'(Z)]^2}$$

$$= \frac{f(Z) \cdot f''(Z)}{[f'(Z)]^2}$$

情況 1：α 是 f(Z) 之一重零點

則 $f'(\alpha) \neq 0$，故 $\phi(\alpha) = \alpha$，$\phi'(\alpha) = 0$。任何 $\varepsilon > 0$，$\exists\, \delta$，

若 $|Z - \alpha| < \delta$，則

$$\left| \frac{\phi(Z) - \phi(\alpha)}{Z - \alpha} - \phi'(\alpha) \right| < \varepsilon$$

取 $\varepsilon < 1$，且 Z_1 在 $|Z - \alpha| < \delta$ 之內，則

$$\left| \frac{Z_2 - \alpha}{Z_1 - \alpha} \right| < \varepsilon$$

$\Rightarrow |Z_2 - \alpha| < \varepsilon\, |Z_1 - \alpha| < \varepsilon\delta < \delta$

即在 $|Z - \alpha| < \delta$ 內對 ϕ 有封閉性，故

$|Z_n - \alpha| < \varepsilon\, |Z_{n-1} - \alpha| < \cdots\cdots < \varepsilon^{n-1}\delta \to 0$，即

Z_n 以 α 爲極限。

情況 2：α 是 $f(Z)$ 的 n 重零點

$$f(Z)=P_n(Z-\alpha)^n+P_{n+1}(Z-\alpha)^{n+1}+\cdots\cdots$$

$$(P_n\neq 0)$$

$$f'(Z)=nP_n(Z-\alpha)^{n-1}+(n+1)P_{n+1}(Z-\alpha)^n+\cdots\cdots$$

$$f''(Z)=n(n-1)P_n(Z-\alpha)^{n-2}+(n+1)nP_{n+1}(Z-\alpha)^{n-1}+\cdots\cdots$$

$$\phi'(Z)=\frac{f(Z)f''(Z)}{[f'(Z)]^2}$$

$$=\frac{(Z-\alpha)^{2n-2}[P_n+P_{n+1}(Z-\alpha)+\cdots][n(n-1)P_n+(n+1)nP_{n+1}+\cdots]}{(Z-\alpha)^{2n-2}[nP_n+(n+1)P_{n+1}(Z-\alpha)+\cdots]^2}$$

得 $\phi'(\alpha)=\dfrac{n(n-1)P_n^2}{n^2P_n^2}$

$$=\frac{n-1}{n}$$

$$\phi(Z)=Z-\frac{f(Z)}{f'(Z)}$$

$$=Z-\frac{(Z-\alpha)^n[P_n+P_{n+1}(Z-\alpha)+\cdots]}{(Z-\alpha)^{n-1}[nP_n+(n+1)P_{n+1}(Z-\alpha)+\cdots]}$$

得 $\phi(\alpha)=\alpha$

任何 $\varepsilon>0$，$\exists\,\delta$，若 $|Z-\alpha|<\delta$，則：

$$\left|\frac{\phi(Z)-\phi(\alpha)}{Z-\alpha}-\phi'(\alpha)\right|<\varepsilon$$ ，即

$$\left|(\phi(Z)-\alpha)-\frac{n-1}{n}(Z-\alpha)\right|<\varepsilon|Z-\alpha|$$

兩邊加上 $\dfrac{n-1}{n}|Z-\alpha|$ 得：（引用 $|A+B|\leq|A|+|B|$）

$$|(\phi(Z)-\alpha)-\frac{n-1}{n}(Z-\alpha)|+\frac{n-1}{n}|Z-\alpha|$$

$$<(\varepsilon+\frac{n-1}{n})\cdot|Z-\alpha| \quad 即$$

$$|\phi(Z)-\alpha|<(\varepsilon+1-\frac{1}{n})\cdot|Z-\alpha|$$

取 ε 爲小於 $\frac{1}{n}$ 之正數，則 $\varepsilon+1-\frac{1}{n}=k<1$，且取 Z_1 在 $|Z-\alpha|<\delta$ 內，則 $|Z_2-\alpha|<k|Z_1-\alpha|<k\delta$，亦得封閉性，故 $|Z_n-\alpha|<k^{n-1}\delta\to 0$，即 Z_n 以 α 爲極限。

綜合情況 1、2 定理得證。

（Note：$\phi'(\alpha)\in R$，且 $\phi'(\alpha)=\frac{n-1}{n}<1$）

定理的證明是針對複變數的，對於實變數當然也成立，所以由於 $\sin X$ 是解析函數，其零點當然可用牛頓法逼近得到。$f(X)f''(X)>0$ 只是一個充分卻非必要的條件罷了！

其實利用 $Z_{n+1}=\phi(Z_n)$ 即疊代法則以逼近根者（即 ϕ 之不動點），不一定要牛頓法，譬如：

$$Z\xrightarrow{\phi}Z-f(Z)\ ,\ Z\xrightarrow{\phi}Z-\frac{f(Z)}{f''(Z)},\ \cdots\cdots\cdots\cdots$$

愛寫幾種便有幾種，可是下面這個定理將顯示出牛頓法之優越性：

映射 ϕ 滿足 $f(\alpha)=0\Rightarrow\phi(\alpha)=\alpha$，若 $\lvert\phi'(\alpha)\rvert=k>1$，則在 ϕ 映射下之點列 Z_n 無法收斂到 α。

pf：任何 $\varepsilon>0$，$\exists\delta$，若 $|Z-\alpha|<\delta$，則

$$\left|\frac{\phi(Z)-\phi(\alpha)}{Z-\alpha}-\phi'(\alpha)\right|<\varepsilon$$

$$\Rightarrow\ |(\phi(Z)-\alpha)-\phi'(\alpha)(Z-\alpha)|<\varepsilon|Z-\alpha|$$

兩邊同加上 $|\phi(Z)-\alpha|$，則：

$$|\phi'(\alpha)(Z-\alpha)-(\phi(Z)-\alpha)|+|\phi(Z)-\alpha|$$

$< \varepsilon \mid Z-\alpha \mid + \mid \phi(Z)-\alpha \mid$

$\Rightarrow \mid \phi'(\alpha)(Z-\alpha) \mid < \varepsilon \mid Z-\alpha \mid + \mid \phi(Z)-\alpha \mid$

$\Rightarrow (k-\varepsilon) \mid Z-\alpha) < \mid \phi(Z)-\alpha \mid$

取 ε 夠小使 $k-\varepsilon = h > 1$，且 Z_1 在 $\mid Z-\alpha \mid < \delta$ 內，則

$\mid Z_2-\alpha \mid > h \mid Z_1-\alpha \mid$，$Z_2$ 距 α 比 Z_1 距 α 還遠，故知 ϕ 在 $\mid Z-\alpha \mid < \delta$ 內，不可能將點列 Z_n 收斂到 α。

拿 $\phi(Z) = Z - f(Z)$ 爲例，$\phi'(Z) = 1 - f'(Z)$，顯然可知無法保證 $\mid \phi'(\alpha) \mid < 1$，但是牛頓法卻無此顧忌。

㈢多變數聯立方程組實數根（數對）之逼近

若 $f(x,y) = 0$ 表曲線 Γ_1，$g(x,y) = 0$ 表曲線 Γ_2，則 $\Gamma_1 \cap \Gamma_2$ 即聯立方程組之解。

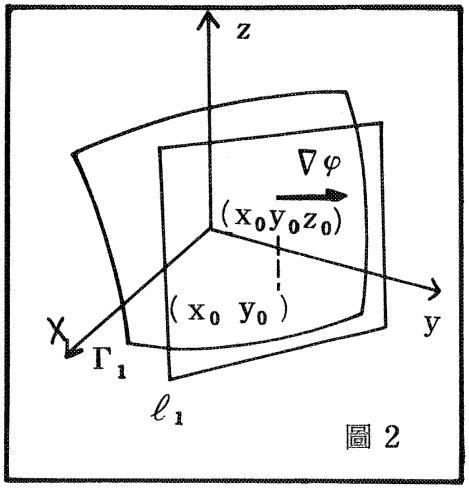

圖 2

解：設 $\varphi = f(x,y) - z$，則

$\nabla\varphi = (f_x, f_y, -1)$

$Z = f(x,y)$ 爲 φ 之零等高面，垂直於 $\nabla\varphi$，故 (x_0, y_0, z_0) 上之切平面方程式爲：

$(x-x_0, y-y_0, z-z_0)(f_x, f_y, -1) = 0$

令 $z = 0$，則得到 $x-y$ 平面上之線 ℓ_1：

$f_x(x-x_0) + f_y(y-y_0) = -f$ ……①

同樣的方法用在 $g(x,y)$ 上得線 ℓ_2：

$g_x(x-x_0) + g_y(y-y_0) = -g$ ……②

以下便用 $\ell_1 \cap \ell_2$ 來逼近 $\Gamma_1 \cap \Gamma_2$。(x_0, y_0) 是原先任意選取的，經過了上述步驟得到了 (x, y)，稱此步驟爲映射 ϕ，則有：

$$\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix} \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} = -\begin{pmatrix} f \\ g \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} - \begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}^{-1} \begin{pmatrix} f \\ g \end{pmatrix}$$

若採 Jacobi 符號，且令

$\vec{x}=\begin{pmatrix} x \\ y \end{pmatrix}$，$\vec{f}=\begin{pmatrix} f \\ g \end{pmatrix}$，$\vec{X}_{n+1}=\phi(\vec{X}_n)$，$J_0=\begin{pmatrix} f_x & f_y \\ g_x & g_y \end{pmatrix}$，

則 $\boxed{\vec{X}_{n+1}=\vec{X}_n-[J_0(\vec{X}_n)]^{-1}\vec{f}(\vec{X}_n)}$

這種形式和 $X_{n+1}=X_n-\dfrac{f(X_n)}{f'(X_n)}$ 完全是等價的，也就是“牛頓法向量形”。很容易了解m個變數，m個方程式聯立的情況下，也同樣有牛頓法的向量形式，證明只需涉及m＋1維空間中切面觀念罷了，無難度，故略。後面將以三變數爲例，證明此法之可行性（收斂性），在此之前，先看一個有趣的結果。

⊙若f，g是一對共軛調和函數，即符合“歌西—黎曼”方程者：

$f_x=g_y$，$g_x=-f_y$。則$F=f+g_i$形成複變解析函數。

而有：$F'(Z)=f_x+g_{xi}=-if_y+g_y$，記$z=x+y_i$，$z_0=x_0+y_{0i}$，由①＋②xi得到：

$F'(z_0)(x-x_0)+iF'(z_0)(y-y_0)=-F(z_0)$

$\Rightarrow \boxed{z=z_0-\dfrac{f(z_0)}{f'(z_0)}}$

得到了單變數複數根之牛頓法。所以複數中的牛頓法只是二維中的一個特例，不過運算的方式二者是截然不同的，前者乃直接用複數代入計算，後者則用了矩陣。

三個方程式$f(x,y,z)=0$，$g(x,y,z)=0$，$h(x,y,z)=0$，其牛頓法將是：

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \xrightarrow{\phi} \begin{pmatrix} u \\ \upsilon \\ \omega \end{pmatrix}$$

原始的形式是：

$$f_x(u-x)+f_y(\upsilon-y)+f_z(\omega-z)=-f$$
$$g_x(u-x)+g_y(\upsilon-y)+g_z(\omega-z)=-g$$
$$h_x(u-x)+h_y(\upsilon-y)+h_z(\omega-z)=-h$$

令

$$J_0 = \begin{vmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \\ h_x & h_y & h_z \end{vmatrix}$$

依克來瑪法則解 u、υ、ω 得：

$$u = x - \frac{1}{J_0}\begin{vmatrix} f & f_y & f_z \\ g & g_y & g_z \\ h & h_y & h_z \end{vmatrix}$$

$$\upsilon = y - \frac{1}{J_0}\begin{vmatrix} f_x & f & f_z \\ g_x & g & g_z \\ h_x & h & h_z \end{vmatrix}$$

$$\omega = z - \frac{1}{J_0}\begin{vmatrix} f_x & f_y & f \\ g_x & g_y & g \\ h_x & h_y & h \end{vmatrix}$$

現在要證明若 $(x_0\ y_0\ z_0) \equiv Q$ 是零點，則

$$J(Q) = \frac{D(u, \upsilon, \omega)}{D(x, y, z)} = 0$$

且 $\phi(Q) = Q$。這二個條件就如前面的 $\phi'(\alpha) = 0$ 及 $\phi(\alpha) = \alpha$ 一樣的地位。

首先，行列式形的函數，其微分等於只對其中某一行（列）微分之和。即若 $A = \det[A_1, A_2, \cdots\cdots A_n]$，$A_n$ 表行向量，則

$$\frac{\partial A}{\partial x} = \det\left[\frac{\partial A_1}{\partial x}, A_2, \cdots\cdots A_n\right] + \det\left[A_1, \frac{\partial A_2}{\partial x}, \cdots\cdots A_n\right] +$$

$$\cdots\cdots + \det\left[A_1, A_2, \cdots\cdots \frac{\partial A_n}{\partial x}\right]。$$

其證明由行列式定義及微分乘法公式立刻可得。

$$u_x = 1 - \frac{1}{J_0^2}\left[\left(\begin{vmatrix} f_x & f_y & f_z \\ g_x & g_y & g_z \\ h_x & h_y & h_z \end{vmatrix} + \begin{vmatrix} f & f_{yx} & f_z \\ g & g_{yx} & g_z \\ h & h_{yx} & h_z \end{vmatrix} + \begin{vmatrix} f & f_y & f_{zx} \\ g & g_y & g_{zx} \\ h & h_y & h_{zx} \end{vmatrix}\right) J_0\right.$$

$$-\begin{vmatrix} f & f_y & f_z \\ g & g_y & g_z \\ h & h_y & h_z \end{vmatrix} \frac{\partial J_0}{\partial x} \Big]$$

$$= \frac{-1}{J_0^2} \Big[\Big(\begin{vmatrix} f \\ g \cdots\cdots \\ h \end{vmatrix} + \begin{vmatrix} f \\ g \cdots\cdots \\ h \end{vmatrix} \Big) J_0 - \begin{vmatrix} f \\ g \cdots\cdots \\ h \end{vmatrix} \frac{\partial J_0}{\partial x} \Big]$$

……型①

$$u_y = -\frac{1}{J_0^2} \Big[\Big(\begin{vmatrix} f_y & f_y & f_z \\ g_y & g_y & g_z \\ h_y & h_y & h_z \end{vmatrix} + \begin{vmatrix} f & f_{yy} & f_z \\ g & g_{yy} & g_z \\ h & h_{yy} & h_z \end{vmatrix} + \begin{vmatrix} f & f_y & f_{zy} \\ g & g_y & g_{zy} \\ h & h_y & h_{zy} \end{vmatrix} \Big) J_0$$

$$-\begin{vmatrix} f & f_y & f_z \\ g & g_y & g_z \\ h & h_y & h_z \end{vmatrix} \frac{\partial J_0}{\partial y} \Big]$$

$$= \frac{-1}{J_0^2} \Big[\Big(\begin{vmatrix} f \\ g \cdots\cdots \\ h \end{vmatrix} + \begin{vmatrix} f \\ g \cdots\cdots \\ h \end{vmatrix} \Big) J_0 - \begin{vmatrix} f \\ g \cdots\cdots \\ h \end{vmatrix} \frac{\partial J_0}{\partial y} \Big]$$

……型②

設Q爲零點 $f(Q) = g(Q) = h(Q) = 0$，故 $u_x = 0$，$u_y = 0$，而 υ_y，ω_z 屬於①，u_z，υ_x，υ_z，ω_x，ω_y 屬於②，故均爲零。即：

$$[J(Q)] = \begin{pmatrix} u_x & u_y & u_z \\ \upsilon_x & \upsilon_y & \upsilon_z \\ \omega_x & \omega_y & \omega_z \end{pmatrix} = 0$$

$$J(Q) = \frac{D(u, \upsilon, \omega)}{D(x, y, z)} = 0$$

$$\phi(Q) = Q$$

顯然可知 n 個變數時同樣有此現象。在 $\vec{X}_{n+1} = \phi(\vec{X}_n)$ 下，要考慮 X_n 之收斂性，似可用縮小映射法，即 $|J| < 1$ 之條件下，ϕ 映射會將體積縮小，實際上卻不可行，設區域 Ω 經 ϕ 映至 Ω'，則：

$$|\Omega'| = \int_{\Omega'} du d\upsilon d\omega = \int_{\Omega} J dx dy dz < \int_{\Omega} dx dy dz = |\Omega|$$

但Ω'却不見得仍落於$|J|<1$之範圍內，故縮小的特性無法連續使用，收斂的證明也就有了困難，見圖3。Q是零點，且$J(Q)=0$，當然在Q附近存在半徑δ，若$|P-Q|<\delta$，則$|J(P)|<1$。

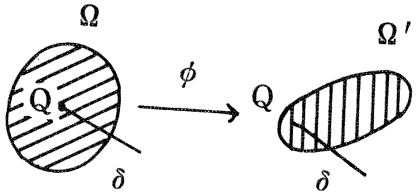

圖3

既然體積的方法行不通，改用弧長縮短來做。考慮一線段$\overline{RQ}$，記爲Γ，而$\Gamma'=\phi(\Gamma)$，顯然Γ'仍有一端點是Q，今希望$|\Gamma'|<\sqrt{\varepsilon}\,|\Gamma|$，$\varepsilon$係小於1之正數（寫成$\sqrt{\varepsilon}$是爲了平方後計算方便）。

$$du=u_x\ d_x+u_y\ d_y+u_z\ d_z$$

$$d\upsilon=\upsilon_x\ d_x+\upsilon_y\ d_y+\upsilon_z\ d_z$$

$$d\omega=\omega_x\ d_x+\omega_y\ d_y+\omega_z\ d_z$$

$$du^2+d\upsilon^2+d\omega^2=Adx^2+2Bdxdy+Cdy^2+2Ddxdz+2Edydz+Fdz^2$$

$$=(\ dxdydz\ )\begin{pmatrix}A&B&D\\B&C&E\\D&E&F\end{pmatrix}\begin{pmatrix}dx\\dy\\dz\end{pmatrix}$$

爲正定的二次形式。

其中$A+C+F=ux^2+uy^2+uz^2+\upsilon x^2+\upsilon y^2+\upsilon z^2+\omega x^2+\omega y^2+\omega z^2$

顯然在Q點取得零值（九項元素皆零），故必可取得δ，若$|P-Q|<\delta$，則$A+C+F|p<\varepsilon$｜以上乃是基於連續性的基本假設，如此則：

$$|\Gamma'|=\int_{\Gamma}\sqrt{Adx^2+\cdots\cdots+Fdz^2}$$

$$|\Gamma|=\int_{\Gamma}\sqrt{dx^2+dy^2+dz^2}$$

只要$Adx^2+2Bdxdy+\cdots\cdots+Fdz^2<\varepsilon(dx^2+dy^2+dz^2)$便可得$|\Gamma'|<\sqrt{\varepsilon}\,|\Gamma|$，下面將證明在$A+C+F<\varepsilon$下是絕對成

立的！（用線性代數）

pf：首先因$A>0$，$C>0$，$F>0$，且$du^2+dv^2+d\omega^2$形成恒正二次形式，故將此二形式標準化得到$\lambda_1 d\bar{x}^2+\lambda_2 d\bar{y}^2+\lambda_3 d\bar{z}^2$，其中$\lambda_1$ λ_2 λ_3 均正，且$\lambda_1+\lambda_2+\lambda_3=A+C+F<\varepsilon$。

而$\varepsilon(dx^2+dy^2+dz^2)\equiv\varepsilon(d\bar{x}^2+d\bar{y}^2+d\bar{z}^2)$

故因$\lambda_1<\varepsilon$，$\lambda_2<\varepsilon$，$\lambda_3<\varepsilon$，得知

$\lambda_1 d\bar{x}^2+\lambda_2 d\bar{y}^2+\lambda_3 d\bar{z}^2<\varepsilon(d\bar{x}^2+d\bar{y}^2+d\bar{z}^2)$得證。

⊙這種討論方式對n變數時同樣適用。

綜合以上，$|\Gamma'|<\sqrt{\varepsilon}\cdot|\Gamma|=\sqrt{\varepsilon}\cdot\overline{RQ}$，故$\Gamma'$就算拉直了，也會落在$|P-Q|<\delta$內，故若$\Gamma_{n+1}=\phi(\Gamma_n)$，$\Gamma_1=\Gamma$，則$|\Gamma_{n+1}|<\sqrt{\varepsilon}^n\cdot|\Gamma|\to 0$，即多變數聯立方程組之牛頓法，其收斂性得證。

(四)多變數複數根之逼近——尚未解決：

關於這個問題可從兩個觀點來看，一是把每一個複數均拆成實部和虛部，那麼問題在(3)中便已解決了。

但是這樣做並不是很好，因爲原先是3個變數的，就要成了六個變數，試想6×6階矩陣、行列式的計算將是多麼的複雜！處理根逼近的問題，雖都是用電腦在做，但仍是太煩了。第二個想法是：牛頓法照用，亦即在多變數複變數下，依然使用牛頓法的向量形式。這就猶如單變數複數根的逼近仍用

$$Z_{n+1}=Z_n-\frac{f(Z_n)}{f'(Z_n)}$$

一般。

筆者曾在需腦上實際比較“用複數直接做”與“拆成二變數來做”的速度，前者快得多，效率較高。

我相信第二個想法是必然可行的，只可惜其嚴格的討論——收斂性，仍未能做出。今後若有機會，相信必能完成這最後一項。

三、參考資料

(一)線性代數——謝志雄。

㈡複變函數導引——華羅庚。

㈢解析概論——高木貞治。

評　語

1.對問題的提法具有相當的成熟度。

2.數學分析的能力強，邏輯推理及形式計算的能力都很強。

3.學習態度積極，不受書本知識的束縛。