Transzcendens, illetve algebrai egyenletrendszer megoldásának egy módszere, illetve annak egy alkalmazása molekula modellek erő-állandóinak kiszámitására

Gehér István

Legyen $f_i(x_1,x_2,\ldots,x_n)$ n darab n változós függvény és legyenek $c_1,c_2,\ldots,c_n$ valós számok. Megoldandó az

$$\left.\begin{array}{l} f_1(x_1,x_2,\ldots,x_n) = c_1 \\ f_2(x_1,x_2,\ldots,x_n) = c_2 \\ \ldots\ldots\ldots \\ f_n(x_1,x_2,\ldots,x_n) = c_n \end{array}\right\} \qquad (A)$$

egyenletrendszer.

A tömörség kedvéért jelöljük az

$$\left.\begin{array}{l} y_1 = f_1(x_1,x_2,\ldots,x_n) \\ y_2 = f_2(x_1,x_2,\ldots,x_n) \\ \ldots\ldots\ldots \\ y_n = f_n(x_1,x_2,\ldots,x_n) \end{array}\right\}$$

transzformációt $Y = F(X)$-szel. Ez a transzformáció tehát az n dimenziós tér $X = (x_1,x_2,\ldots,x_n)^*$ vektorához hozzárendeli ugyanennek a térnek egy $Y = (y_1,y_2,\ldots,y_n)^*$ vektorát. Ha a $(c_1,c_2,\ldots,c_n)$ vektort C-vel jelöljük, akkor az (A) egyenletrendszer az

$$F(X) = C \ldots \qquad (A')$$

egyenletbe megy át.

Hogy az (A), vagy a vele egyenértékü (A') egyenletet meg tudjuk oldani, az F(X) transzformációra és a C vektorra bizonyos megszoritásokat kell tennünk. Megoldáson természetesen elektronikus számológépekre programozható eljárást értünk.

A továbbiakban tegyük fel, hogy az

$$Y = F(X)$$

transzformáció az n dimenziós tér egy D összefüggő és nyilt halmazát ugyanannek a térnek egy R halmazára egy egyértelmüen képezi le. Tegyük fel továbbá, hogy a

$$\frac{\partial f_i}{\partial x_j} \qquad (i=1,2,\ldots,n;\ \ j=1,2,\ldots,n)$$

parciális deriváltak a D minden pontjában léteznek és folytonosak. Végül tegyük fel, hogy a

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}, & \frac{\partial f_1}{\partial x_2}, & \ldots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1}, & \frac{\partial f_2}{\partial x_2}, & \ldots & \frac{\partial f_2}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial f_n}{\partial x_1}, & \frac{\partial f_n}{\partial x_2}, & \ldots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} = \left[\frac{\partial f_i}{\partial x_j}\right]$$

Jacobi mátrix a D minden pontjában reguláris, azaz determinánsa zérustól különböző.

Megjegyezzük, hogy a $\frac{\partial f_i}{\partial x_j}$ deriváltak folytonossága és a

$\left[\frac{\partial f_i}{\partial x_j}\right]$ mátrix regularitása miatt az R halmaz szintén összefüggő és nyilt.

Az $Y = F(X)$ transzformáció inverzét jelöljük $X = \phi(Y)$-nal.
Az $F(X) = C$ egyenletnek akkor van gyöke, ha C eleme az R halmaznak. Ebben az esetben viszont az egyenlet egyetlen gyöke $\phi(C)$.

Jelöljük a $\left[\frac{\partial f_i}{\partial x_j}\right]$ Jacobi mátrixot $\frac{dF}{dX}$ -szel, vagy

$\frac{dY}{dX}$ -szel. A láncszabály miatt

$$\frac{dF}{dX} \cdot \frac{d\phi}{dY} = \frac{dY}{dY} = I$$

ahol I az egységtranszformáció, ezért

$$\frac{d\phi}{dY} = \left[\frac{dF}{dX}\right]^{-1} \qquad (B)$$

Mivel a $\frac{dF}{dX}$ és igy az inverze is csak az $X = (Y)$-tól függ,

azért a (B) formula a $\phi(Y)$ függvényre egy általánosabb értelemben vett differenciál egyenlet.

Az (A) illetve a vele egyenértékü (A') egyenletet most már a következőképpen oldjuk meg:

Veszünk egy X^o vektort a D halmazból és meghatározzuk hozzá az $Y^o = F(X^o)$ vektort. Megkeressük a

$$\frac{dX}{dY} = \left[\frac{dF}{dX}\right]^{-1} \qquad (C)$$

differenciál-egyenlet azon $\phi(Y)$ megoldásának $\phi(C)$ helyettesitési értékét, amely eleget tesz a $\phi(Y^o) = X^o$ kezdeti feltételnek.

Mivel

$$\Delta X = \Delta \phi\ (Y) = \frac{dX}{dY} \Delta Y + R(\Delta Y)$$

ahol $|R(\Delta Y)| = O(|\Delta Y|)$, azért az (A') egyenlet $X = \phi(C)$ gyökének egy X^k közelitő értékét a következőképpen adjuk meg:

Felveszünk m+1 darab $Y^o = Z^o, Z^1, Z^2, \ldots, Z^{m-1}, Z^m = C$ vektort ugy, hogy a

$$Z^o, Z^1, Z^2, \ldots, Z^{m-1}, Z^m$$

törtvonal az R-ben legyen és hogy az

$$|Z^i - Z^{i-1}| < \frac{1}{k}$$

egyenlőtlenség fennálljon, ha i=1,2m...,m. Meghatározzuk az $V^o, V^1, V^2, \ldots, V^m = X^k$ vektorokat a következő rekurziós formula segitségével:

$$V^o = X^o$$

$$V^{i+1} = V^i + \left[\frac{dF}{dX}\right]^{-1}_{X=U^i} (Z^{i+1} - Z^i)$$

Belátható, hogy $X^k \longrightarrow X = \emptyset(C)$, ha $k \longrightarrow \infty$.

Mivel ez eléggé plauzibilis és itt számitástechnikai módszerről van szó, azért a bizonyitást annak hosszadalmassága miatt nem közöljük.

A gyakorlati feladatok megoldásánál általában a $\frac{\partial f_i}{\partial x_j}$ parciális deriváltat, vagyis a $\frac{dF}{dX}$ mátrixot kénytelenek vagyunk kézzel meghatározni. Mégis abban az igen fontos speciális esetben, amikor tudjuk, hogy az $f_i\ (x_1,x_2,\ldots,x_n)$ függvények olyan polinomok, hogy bármely változójukban legfeljebb m-ed fokuak, a differenciálást a számológéppel is el lehet végeztetni. Ennek az az oka, hogy egy polinom együtthatóit a differenciák segitségével ki lehet fejezni.

Legyen h egy fix szám és jelöljük Δ-val azt az operátort, amelyikre

$$\Delta g(x) = g(x+h) - g(x)$$

Ha $p(x) = C_m X^m + C_{m-1} X^{m-1} + \ldots + C_1 X + C_o$, akkor a C_i együtthatókra a következő rekurziós formula érvényes:

$$C_m = \frac{1}{m!} \frac{\Delta^m p(x)}{h^m}$$

$$C_i = \frac{1}{i!}\;\frac{\Delta^i\left[p(x) - (C_m X^m + C_{m-1} X^{m-1} + \dots + C_{i+1} X^{i+1})\right]}{h^i}$$

ahol X tetszés szerinti. A C_i együtthatók ismeretében a

$$p'(x) = m\, C_m X^{m-1} + (m-1)\, C_{m-1} X^{m-2} + \dots + 2\, C_2 X + C_1$$

polinom helyettesitési értékeinek a meghatározása nem okoz gondot.

A fent emlitett differenciálási eljárásnak az a nagy előnye is megvan, hogy az $F(X)$ transzformációt explicite nem is kell ismerni, elég ha azt egy rekurziv eljárás segitségével adjuk meg. A következőkben meg fogunk emliteni egy konkrét problémát, ahol éppen ez a helyzet.

Molekula modellek erő-állandóinak kiszámitásával kapcsolatban a következő probléma merült fel:

Adva van egy A $n \times n$-es mátrix, továbbá n szám: $\gamma_1, \gamma_2, \dots, \gamma_n$.

Meghatározandó a

$$Z = \begin{bmatrix} X_1 & & & & & \\ & X_2 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & X_n \end{bmatrix}$$

diagonális mátrix ugy, hogy AZ sajátértékei $\gamma_1, \gamma_2, \dots, \gamma_n$ legyenek. A továbbiakban ezt a problémát (E)-vel fogjuk jelölni.

Egy M n-ed rendü négyzetes mátrix $|tE - M|$ karakterisztikus polinomja legyen $h_o+h_1t + h_2t^2 + \dots + h_{n-1}t^{n-1} +(-1)^n t^n$

A $(h_o,h_1,\dots,h_{n-1})^* = H$ vektort az M mátrix karakterisztikus vektorának fogjuk nevezni. A H vektort az M sajátértékei egyértelmüen meghatározzák.

Ugyanis:

$$h_{n-i} = (-1)^{n-i} \sum S_{k1} S_{k2} \dots S_{ki}$$

ahol $S_1, S_2, \dots, S_n$ az M sajátértékei.

Az E probléma tehát a következőképpen fogalmazható:

Adva van egy A n-ed rendü négyzetes mátrix, továbbá egy $C = (c_1,c_2,\dots,c_n)^*$ vektor.

Keressük a

$$Z = \begin{bmatrix} X_1 & & & & & \\ & X_2 & & & & \\ & & \cdot & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & & X_n \end{bmatrix}$$

digonális mátrixot ugy, hogy az AZ mátrix karakterisztikus vektora C legyen.

Jelölje $Y = F(X)$ azt a transzformációt, amely az $X = (x_1,x_2,\dots,x_n)^*$ vektorhoz hozzárendeli az

$$AZ = A \begin{bmatrix} x_1 & & & & \\ & x_2 & & & \\ & & \cdot & & \\ & & & \cdot & \\ & & & & \cdot \\ & & & & & x_n \end{bmatrix}$$

mátrix karakterisztikus vektorát. Az (E) probléma megoldása egyenértékü az

$$F(X) = C$$

egyenlet megoldásával.

Tudjuk, hogy egy M mátrix $H = (h_0, h_1, \ldots, h_{n-1})^*$ karakterisztikus vektorának h_{n-i} koordinátája egyenlő az összes i-edrendü főminorok összegének $(-1)^{n-i}$-szeresével.

Jelöljük az A mátrix oszlopvektorait $A^1, A^2, \ldots, A^n$-nel.

Az AZ mátrix oszlopvektorai $x_1 A^1$, $x_2 A^2, \ldots, x_n A^n$.

Az AZ mátrix főminorai nyilván az x_1, $x_2, \ldots, x_n$ olyan polinomjai, amelyek minden változóban lineárisak. Ezért az AZ mátrix karaktérisztiku vektorának a koordinátái, azaz az $f_i\ (x_1, x_2, \ldots, x_n)$ függvények minden változójukban lineárisak. A parciális deriváltakat tehát a következő egyszerü formula adja meg:

$$\frac{\partial f_i(x_1, x_2, \ldots, x_n)}{\partial x_j} = \frac{f_i(x_1, x_2, \ldots, x_j + h, \ldots, x_n)}{h} - \frac{f_i(x_1, x_2, \ldots, x_n)}{h}$$

Az F(X) transzformáció értékeinek a kiszámitása a főminorok segitségével gyakorlatilag használhatatlan, ezért F(X) meghatározására gyakorlatilag használható explicit formulánk nincs. Vannak viszont jó rekurziv eljárások, például a LEVERRIER féle eljárás, illetve annak FADDETEV által módositott formája. Mivel F(X) egy egyértelmüséget csak egy hely környezetében tudjuk biztositani, azért az eljárás csak abban az esetben konvergál, ha olyan X^o-ból indulunk ki, amely valamelyik gyökhöz közel van.

Irodalom:

D.E. Mann, T. SHIMANOUCHI, H.MEAL, L.FANO: NORMAL COORDINATE ANALYSIS OF HALOGENATED ETHYLENES.

THE JOURNAL OF CHEMICAL PHYSICS VOLUME 27. NUMBER 1. /1956/

ERICH W SCHMID: MOLEKÖLMODELLRECHNUNGEN MIT ELEKTRONISCHEN RECHENANLAGEN.

ZEITSCHRIFT FÜR ELEKTRONENCHEMIE. BD. 64. NR4. /1960/

Summary

A method for solving the system of transcendent equations

Let $Y = F(X)$ denote a transformation that maps the domain D in the n dimensional space onto a set R in the same n dimensional space. The equation $F(X) = C$ is to be solved.

Let us suppose that the domain D is open and connected, moreover the map $Y = F(X)$ is one to one, continuous, having continuous partial derivatives

$\frac{\partial f_i}{\partial x_j}$ at all the points of the domain D. Besides we will apply for the restriction that the Jacobian determinant $\left| \frac{\partial f_i}{\partial x_j} \right|$ is not equal to zero at any point of the domain D.

The continuity of the partial derivatives $\frac{\partial f_i}{\partial x_j}$ and the not vanishing of the Jacobian determinant involves for the set R to be open and connected.

Let $\emptyset(Y)$ denote the inverse transormation of $Y = F(X)$. The equation $F(X) = C$ has got a root if and only if C belongs to R. In this case the equation mentioned above has got a unique rott $\emptyset(C)$.

Let us denotet the Jacobian matrix $\left[\frac{\partial f_i}{\partial x_j} \right]$ with $\frac{dF}{dX}$. Owing to the chain-rule

$$\frac{d\emptyset}{dY} = \left[\frac{dF}{dX} \right]^{-1}$$

This is a differential equation in a general meaning for the function $\emptyset(Y)$. The solution of that differential equation can be found as follows:

We take a vector X^o of the domain D and then we define the vector $Y^o = F(X^o)$. Let $\emptyset(Y)$ be the solution of the differential equation

$$\frac{dX}{dY} = \left[\frac{dF}{dX}\right]^{-1}$$

that satisfies the initial value condition $\emptyset(Y^o) = X^o$.

$\emptyset(C)$ is the unique root of the equation $F(X) = C$.

Due to the formula

$$\Delta X = \Delta\, \emptyset(X) = \frac{dX}{dY}\, \Delta\, Y + R(\Delta\, Y),$$

where $|R(\Delta\, Y)| = \text{ordo}\,(|\Delta\, Y|)$, the approximate value of the root $\emptyset(C)$ can be calculated in the following way:

We take m+1 vectors $Y^o = Z^o,\ Z^1,\ Z^2,\ \ldots,\ Z^{m-1},\ Z^m$ so that the path $Z^o,Z^1,\ldots,Z^{m-1},Z^m$ with a straight line between two consecutive points should belong to the domain R and the inequality

$$|Z^i - Z^{i-1}| < \frac{1}{k}$$

should be satisfied, if $i=1,2,\ldots,m$. We define the vectors $V^o,V^1,\ldots,V^m = X^k$ by the following recursive formula:

$$V^o = X^o$$

$$V^{i+1} = V^i + \left[\frac{dF}{dX}\right]^{-1}_{X=Vi}$$

It is not difficult to justify that X^k converges to $X = \emptyset(C)$ if $k \longrightarrow \infty$ This method is particularly simple when the functions $f_i(x_1,x_2,\ldots,x_n)$ are polynoms.
In this case the coefficients of the polynoms can be calculated by means of the differences of higher order.