El Mètode de Newton en varies variables

Lluís Alsedà

Departament de Matemàtiques
Universitat Autònoma de Barcelona

http://www.mat.uab.cat/∼alseda

setembre 2024 (versió 1.0.0)

UAB Universitat Autònoma de Barcelona DEPARTAMENT DE **MATEMÀTIQUES**

Continguts

Deducció analítica del mètode de Newton com a mètode de Taylor directe

Suposem que $f\colon [a,b] \longrightarrow \mathbb{R}$ és de classe $\mathcal{C}^2$ i que existeix $\alpha \in (a,b)$ amb $f(\alpha) = 0$. Suposem que x és una aproximació d'α, i expandim f per Taylor fins a ordre 1 prenent x com a punt base:

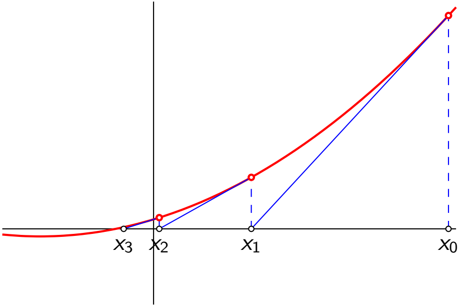
$$\begin{aligned} 0 = f(\alpha) &= f(x) + f'(x)(\alpha - x) + \mathcal{O}((\alpha - x)^2) \\ &\approx f(x) + f'(x)(\alpha - x), \end{aligned}$$

d'on
$$\alpha - x \approx -\frac{f(x)}{f'(x)}.$$

Així, $x - \frac{f(x)}{f'(x)} \approx \alpha$ i esperem que

$$\boxed{x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}}$$

sigui millor aproximació de α que x_n.



Donat x_i, aproximació d'α, prenem la recta tangent a la grafica d'f en el punt $\big(x_i, f(x_i)\big)$:

$$y = f(x_i) + f'(x_i)(x - x_i),$$

la tallem amb l'eix x (fem $y = 0$ a dalt): $\qquad x = x_i - \dfrac{f(x_i)}{f'(x_i)}$

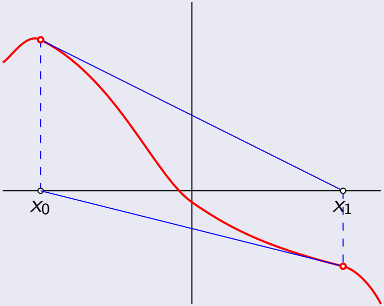
i prenem $x_{i+1} := x$ com a aproximació millorada.

Nota:

Òbviament les dues deduccions donen la mateixa fórmula, donat que l'aproximació de Taylor d'ordre 1 de f al voltant de x no és més que la recta tangent a la gràfica de f al punt $(x, f(x))$.

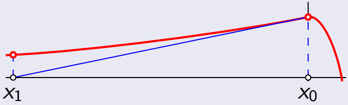
Quan som lluny de l'arrel, el mètode de Newton pot convergir molt lentament (o no convergir). És per això que *sempre sempre sempre* cal limitar el nombre màxim d'iterats.

Òrbita periòdica (de període 2)



Es perd

Aquest cas s'il·lustra analíticament a l'exemple següent



Ordre de convergència del mètode de Newton

Teorema (Ordre de convergència del mètode de Newton)

Sigui $f : [a, b] \longrightarrow \mathbb{R}$ de classe $\mathcal{C}^2$, $f(\alpha) = 0$ i $f'(\alpha) \neq 0$ (i.e. α és un zero simple de f). Si, $f''(\alpha) \neq 0$, aleshores el mètode de Newton per trobar α com a zero de $f(x) = 0$ és quadràtic amb constant asimptòtica de l'error $C = \frac{1}{2}\left|\frac{f''(\alpha)}{f'(\alpha)}\right|$. Si, $f''(\alpha) = 0$ el mètode és al menys cúbic.

Recordem que

un mètode iteratiu $x_{n+1} = g(x_n)$ per a calcular un punt fix $\alpha = g(\alpha)$, on x_0 és una aproximació inicial d'α, té *ordre p* si existeix $\varepsilon > 0$ tal que $\forall x_0 \in (\alpha - \varepsilon, \alpha + \varepsilon)$ es compleix que $x_{n+1} = g(x_n) \xrightarrow{n} \alpha$ amb ordre de convergència p.

Ordre de convergència de successions — Definició

Sigui $\{x_n\}_n \subset \mathbb{R}$ successió amb límit α. Direm que la successió té *ordre de convergència (exactament) p* (pot ser $p \notin \mathbb{N}$, $p > 0$) si $\exists C > 0$ tal que

$$\lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = C \quad (\neq 0, \infty).$$

La quantitat C s'anomena *constant asimptòtica de l'error*. En el cas $p = 1$, demanem, a més, que $C < 1$. Per $p = 1$ parlem de *convergència lineal*, per $p = 2$ de *convergència quadràtica* i, per $p = 3$, *cúbica*.

Si es compleix

$$\lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = 0$$

direm que la successió té *ordre de convergència almenys p*.

Deducció analítica (fàcil) del mètode de Newton en varies variables com a mètode de Taylor directe

Sigui $\Omega \subset \mathbb{R}^n$ obert i convex, i sigui $F\colon \Omega \longrightarrow \mathbb{R}^n$ de classe $\mathcal{C}^1$ amb $\alpha \in \Omega$ tal que $F(\alpha) = 0$. Suposem que x és una aproximació d'α, i expandim F per Taylor fins a ordre 1 al voltant de x:

$$0 = F(\alpha) \approx F(x) + J_F(x)(\alpha - x)$$

d'on

$$J_F(x)(\alpha - x) \approx -F(x).$$

Aquesta expressió dóna el mètode iteratiu

$$\boxed{J_F(x_k)\,(x_{k+1} - x_k) = -F(x_k),}$$

amb l'esperança que x_{k+1} sigui millor aproximació d'α que x_k.

Observació

Per millorar la similitud amb el mètode de Newton en una variable, l'expressió anterior també es pot escriure com

$$x_{k+1} = x_k - J_F^{-1}(x_k)\,F(x_k),$$

que no serveix de res degut a la regla bàsica: **mai mai mai** s'ha d'invertir una matriu (excepte en el cas de matrius molt concretes amb propietats especials).

El mètode de Newton en varies variables és:

donat $x_0 \in \Omega$, per $k = 0, 1, \dots,$ *fins arribar a convergència o esgotament del nombre màxim d'iterats* (*condicions de parada*)

$$\text{resoldre el sistema } J_F(x_k)\, u_k = -F(x_k)$$
$$\text{i definir } x_{k+1} := x_k + u_k.$$

Observem que, *a cada pas, cal resoldre un sistema lineal $n \times n$*.

A partir d'aquí seguim *a la nostra manera* la secció *7.1.1 Newton's Method and Its Variants* de la monografia

Numerical mathematics, Alfio Quarteroni, Riccardo Sacco, Fausto Saleri, Texts in Applied Mathematics **37**, Springer Science, 2006.

Exemple

Considerem el sistema no-lineal

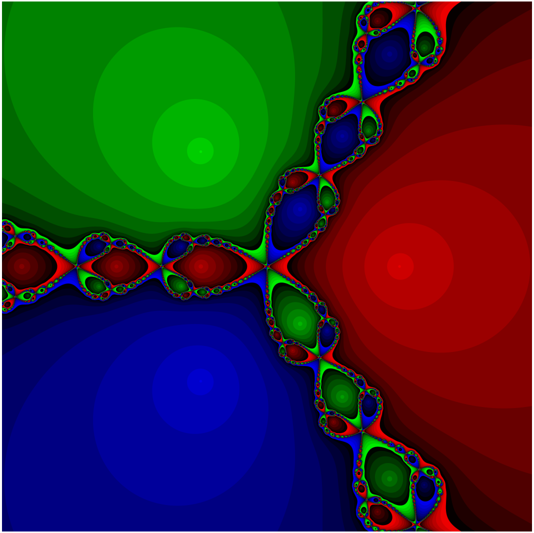
$$\begin{cases} e^{x^2+y^2} = 1, \\ e^{x^2-y^2} = 1, \end{cases}$$

que té $(0,0)$ com a única solució.

Usarem el mètode de Newton en dimensió 2 amb la funció

$$F(x,y) = \left(e^{x^2+y^2} - 1, e^{x^2-y^2} - 1\right).$$

- Prenent $x_0 = (0.1, 0.1)$ com a llavor inicial, el mètode convergeix al punt $\left(0.61 \cdot 10^{-5}, 0.61 \cdot 10^{-5}\right)$ en 15 iterats.
- Prenent $x_0 = (10, 10)$ com a llavor inicial, es necessiten 220 iterats per a convergir a un punt similar al $\left(0.61 \cdot 10^{-5}, 0.61 \cdot 10^{-5}\right)$.
- Prenent $x_0 = (20, 20)$ com a llavor inicial no hi ha convergència.



The *Classical Newton's fractal* (at the left) is obtained by drawing the convergence pattern of Newton's method applied to the function

$$f(z) = z^3 - 1$$

in the complex plane.

If a sequence generated by applying Newton's method to this function converges, then it will be to one of the three cube roots of -1. For the Classic Newton's fractal, we color each point according to which limit point the sequence converges.

Newton Fractal, Mitch Richling, `https://www.mitchr.me/SS/newton/`, 2024.

I l'ordre de convergència?

Teorema (Ordre de convergència del mètode de Newton)

Sigui $\Omega \subset \mathbb{R}^n$ obert i convex, i sigui $F\colon \Omega \longrightarrow \mathbb{R}^n$ de classe $\mathcal{C}^1$ amb $\alpha \in \Omega$ tal que $F(\alpha) = 0$. Suposem que $J_F(\alpha)$ és no singular i que existeixen constants positives R, C i L, tals que $\left|\!\left|\!\left| J_F^{-1}(\alpha) \right|\!\right|\!\right| \leqslant C$ i

$$\left|\!\left|\!\left| J_F(x) - J_F(y) \right|\!\right|\!\right| \leqslant L \left\| x - y \right\| \qquad \forall x, y \in \mathcal{B}(\alpha; R),$$

on $|\!|\!|\cdot|\!|\!|$ és una norma de matrius consistent amb la norma de vectors $\left\|\cdot\right\|$.

Aleshores, existeix $r > 0$ tal que, per a qualsevol $x_0 \in \mathcal{B}(\alpha; r)$, la successió $\{x_k\}_{k=0}^{\infty}$ d'iterats del mètode de Newton per la funció F està ben definida i convergeix a α de manera que

$$\left\| x_{k+1} - \alpha \right\| \leqslant 2CL \left\| x_k - \alpha \right\|^2 .$$

Notació: $\mathcal{B}(\alpha; \rho)$ denota la bola oberta de centre α i radi ρ amb la norma $\left\|\cdot\right\|$.

Demostració I

Demostrarem per inducció respecte de k que $\{x_k\}_{k=0}^{\infty} \subset \mathcal{B}(\alpha; r)$, amb $r = \min\left\{R, \frac{1}{2CL}\right\}$, i que $\|x_{k+1} - \alpha\| \leqslant 2CL\,\|x_k - \alpha\|^2 \ \ \forall k$.

Recordem el Teorema

Si per alguna norma subordinada $|||M||| < 1$, *llavors* $\mathrm{Id} - M$ *és no-singular i*

$$\frac{1}{1 + |||M|||} \leqslant \left|\left|\left|(\mathrm{Id} - M)^{-1}\right|\right|\right| \leqslant \frac{1}{1 - |||M|||}.$$

Primer demostrarem que $J_F(x_0)$ és no singular per a tot $x_0 \in \mathcal{B}(\alpha; r) \subset \mathcal{B}(\alpha; R)$. Per a això usarem el teorema anterior amb

$$M := \Big(J_F(\alpha) - J_F(x_0)\Big) J_F^{-1}(\alpha).$$

Tenim

$$|||M||| = \left|\left|\left|\Big(J_F(\alpha) - J_F(x_0)\Big) J_F^{-1}(\alpha)\right|\right|\right| \leqslant$$
$$\left|\left|\left|J_F^{-1}(\alpha)\right|\right|\right| \cdot |||J_F(\alpha) - J_F(x_0)||| \leqslant CL\,\|\alpha - x_0\| \leqslant CLr \leqslant \tfrac{1}{2}.$$

Demostració II

De la definició d'M es segueix que,

$$\mathrm{Id} - M = J_F(\alpha)J_F^{-1}(\alpha) - \Big(J_F(\alpha) - J_F(x_0)\Big)J_F^{-1}(\alpha) = J_F(x_0)J_F^{-1}(\alpha).$$

Pel teorema anterior, aquesta matriu és no singular i, en conseqüència, $J_F(x_0)$ és no singular. Llavors, x_1 està ben definit. A més, com que $F(\alpha) = 0$, pel teorema del valor mitjà,

$$\begin{aligned} x_1 - \alpha = x_0 - J_F^{-1}(x_0)F(x_0) - \alpha = \\ (x_0 - \alpha) - J_F^{-1}(x_0)\big(F(x_0) - F(\alpha)\big) = \\ J_F^{-1}(x_0)\Big(F(\alpha) - F(x_0) - J_F(x_0)(\alpha - x_0)\Big) = \\ J_F^{-1}(x_0)\Big(J_F(z)(\alpha - x_0) - J_F(x_0)(\alpha - x_0)\Big) = \\ J_F^{-1}(x_0)\Big(J_F(z) - J_F(x_0)\Big)(\alpha - x_0), \end{aligned}$$

on z és un punt a la recta que uneix α amb x_0 (que és al domini per hipòtesi).

En particular, $\|z - x_0\| \leqslant \|\alpha - x_0\|$ i $z \in \mathcal{B}(\alpha; r) \subset \mathcal{B}(\alpha; R)$.

Demostració III

Passant la identitat anterior a normes,

$$\|x_1 - \alpha\| \leqslant \left\|\left| J_F^{-1}(x_0) \right|\right\| \cdot \left\|\left| J_F(z) - J_F(x_0) \right|\right\| \cdot \|\alpha - x_0\| .$$

Per afitar el segon terme de la desigualtat anterior necessitem una fita superior de $\left\|\left| J_F^{-1}(x_0) \right|\right\|$. De la cadena d'igualtats del principi de la diapositiva anterior i del teorema, es dedueix

$$(\mathrm{Id} - M)^{-1} = J_F(\alpha) J_F^{-1}(x_0),$$

i

$$\left\|\left| J_F^{-1}(x_0) \right|\right\| = \left\|\left| J_F^{-1}(\alpha)(\mathrm{Id} - M)^{-1} \right|\right\| \leqslant$$
$$\left\|\left| J_F^{-1}(\alpha) \right|\right\| \left\|\left| (\mathrm{Id} - M)^{-1} \right|\right\| \leqslant \frac{\left\|\left| J_F^{-1}(\alpha) \right|\right\|}{1 - \left\|\left| M \right|\right\|} \leqslant 2 \left\|\left| J_F^{-1}(\alpha) \right|\right\| \leqslant 2C$$

(recordem que $\left\|\left| M \right|\right\| \leqslant \frac{1}{2}$).

En conclusió (recordem que $\|z - x_0\| \leqslant \|\alpha - x_0\|$ i $z, x_0 \in \mathcal{B}(\alpha; r) \subset \mathcal{B}(\alpha; R)$),

$$\|x_1 - \alpha\| \leqslant \left|\!\left|\!\left| J_F^{-1}(x_0) \right|\!\right|\!\right| \cdot \left|\!\left|\!\left| J_F(z) - J_F(x_0) \right|\!\right|\!\right| \cdot \|\alpha - x_0\| \leqslant \\ 2C \cdot L \|z - x_0\| \cdot \|\alpha - x_0\| \leqslant 2CL \|\alpha - x_0\|^2 .$$

Això demostra que el teorema és cert per $k = 0$.

Per altra banda,

$$\|x_1 - \alpha\| \leqslant 2CL \|x_0 - \alpha\|^2 \leqslant 2CL\, r^2 \leqslant r.$$

En conclusió, $x_1 \in \mathcal{B}(\alpha; r)$ (això és el que, en realitat, permet fer el pas d'inducció).

La demostració de que si tenim $\|x_{k+1} - \alpha\| \leqslant 2CL \|x_k - \alpha\|^2$ per algun $k \geqslant 0$, llavors $\|x_{k+2} - \alpha\| \leqslant 2CL \|x_{k+1} - \alpha\|^2$, és anàloga a la del cas $k = 0$.

Comentaris a l'ordre de convergència del mètode de Newton

El teorema de l'ordre de convergència del mètode de Newton n'assegura la convergència quadràtica però solament quan la llavor x_0 és prou propera a la solució α, i la matriu jacobiana a la solució és no-singular.

Val la pena assenyalar que l'esforç computacional necessari per resoldre el sistema lineal pot ser excessivament alt (especialment per dimensions grans).

Per altra banda, la matriu del sistema $J_F(x_k)$ pot estar mal condicionada. En aquest cas, és molt difícil obtenir solucions precises del sistema.

Hi ha diverses modificacions proposades al mètode de Newton, que pretenen reduir els problemes descrits anteriorment.

Mètodes de Newton modificats: Actualització cíclica de la Matriu Jacobiana

Una manera de millorar l'eficiència del mètode de Newton consisteix a mantenir la Matriu Jacobiana (més precisament, la seva factorització LU, QR, ...) sense canvis durant un nombre determinat $p \geqslant 2$ de passos[1].

Observació (El negoci d'en Robert amb les cabres)

Amb aquesta estratègia, en general, s'aconsegueix un deteriorament de la velocitat de convergència, amb un guany en l'eficiència computacional. Si p és massa gran pot passar que la pèrdua de velocitat provocada pel deteriorament de la convergència sigui superior al guany computacional aconseguit per la reducció computacional.

[1] En general no hi ha algorisme per a decidir la p a priori. Moltes vegades es decideix per prova i error.

Mètodes de Newton modificats: Solucions inexactes dels sistemes lineals

Una altra manera de millorar l'eficiència del mètode de Newton és resoldre el sistema lineal del mètode de Newton usant un mètode iteratiu (com Jacobi, SOR, Gauss-Seidel, Krylov, ...) on, en comptes d'iterar fins convergir prop de la solució del sistema lineal, es fixa *a priori* el nombre màxim d'iteracions admissibles del mètode iteratiu[2].

L'observació "*el negoci d'en Robert amb les cabres*" sobre l'eficiència computacional final del mètode, és d'aplicació també en aquest cas.

[2] Novament, en general no hi ha algorisme per a decidir el nombre màxim d'iteracions admissibles del mètode iteratiu a priori. Moltes vegades es decideix per prova i error.

Mètodes de Newton modificats: Aproximacions de la Matriu Jacobiana per diferències finites o centrades

La computació explícita de $J_F(x_k)$ sovint és "molt cara". En aquests cassos una bona estratègia és substituir $J_F(x_k)$ per una aproximació $J_F^{k,h}$ obtinguda per diferències finites o centrades.

Exemple (Diferències finites de primer ordre)

La columna $j \in \{1, 2, \ldots, n\}$ de l'aproximació $J_F^{k,h}$ obtinguda amb diferències finites de primer ordre és

$$\frac{F(x_k + h_{k,j}\, \boldsymbol{e}_j) - F(x_k)}{h_{k,j}}$$

on $\boldsymbol{e}_j$ és el j−èssim vector de la base canònica d'$\mathbb{R}^n$ i $h_{k,j} > 0$ són increments que s'han de triar adequadament a cada pas k de Newton.

Mètodes de Newton modificats: Aproximacions de la Matriu Jacobiana per diferències finites o centrades

Exemple (Diferències centrades de primer ordre)

Ara la columna $j \in \{1, 2, \ldots, n\}$ de l'aproximació $J_F^{k,h}$ és

$$\frac{F(x_k + h_{k,j}\,\boldsymbol{e}_j) - F(x_k - h_{k,j}\,\boldsymbol{e}_j)}{2h_{k,j}}.$$

Es pot demostrar que existeixen condicions sobre els increments $h_{k,j}$ de manera que el mètode de Newton amb la matriu Jacobiana $J_F(x_k)$ substituïda per la matriu $J_F^{k,h}$ calculada amb diferències finites de primer ordre està ben definit i convergeix linealment. Oimés, si existeix una constant positiva κ tal que $\max_j |h_{k,j}| \leqslant \kappa \,\|F(x_k)\|$ per tot k, llavors la convergència és quadràtica.

Mètodes de Newton modificats: Aproximacions de la Matriu Jacobiana per diferències finites o centrades

En general no hi ha cap indicació constructiva sobre com calcular els increments $h_{k,j}$.

Tot i així, es ben sabut que l'error de truncament de les derivades aproximades amb diferències es pot reduir disminuint els $h_{k,j}$. D'altra banda, valors massa petits dels $h_{k,j}$ poden provocar grans errors d'arrodoniment[3].

Per tant, cal un compromís entre la necessitat de limitar els errors de truncament i, al mateix temps, assegurar una certa precisió en els càlculs. A la literatura es troben estratègies de càlcul (aproximadament) òptim dels $h_{k,j}$.

[3] Aquest és un problema clàssic al món dels mètodes numèrics.

Mètode de Broyden: Introducció

El Mètode Broyden pertany a la família dels *mètodes tipus secant*, que es construeixen a partir del mètode de la secant definit per funcions reals.

Per tant, en certa manera, el mètode de Broyden és un mètode de Newton modificat.

En particular, com a mètode de tipus secant, a cada pas s'usen els *dos* darrers iterats per a generar el següent element de la successió d'iterats.

El mètode de Broyden

Donada una llavor inicial $x_0 \in \Omega$ es defineix

$$x_1 := x_0 + u_1 \qquad \text{amb} \quad Q_0\, u_1 = -F(x_0),$$

on Q_0 és $J_F(x_0)$ o una bona aproximació de $J_F(x_0)$ i, per $k = 1, 2, \ldots,$ *fins arribar a convergència o esgotament del nombre màxim d'iterats*

es resol el sistema $Q_k\, u_{k+1} = -F(x_k)$

i es defineix $x_{k+1} := x_k + u_{k+1},$

on Q_k és una matriu $n \times n$ tal que

$$Q_k\, u_k = Q_k\,(x_k - x_{k-1}) = F(x_k) - F(x_{k-1}) \tag{1}$$

És a dir, Q_k generalitza el pendent de la "secant entre $\big(x_{k-1}, F(x_{k-1})\big)$ i $\big(x_k, F(x_k)\big)$".

Si oblidem les manipulacions de la matriu del sistema, el mètode de Broyden és exactament el mètode de Newton.

Observem que (1) no és suficient per determinar la matriu Q_k de manera única (tenim un sistema $n \times n$ on la incògnita és la matriu). Una possible solució per a resoldre aquest problema és requerir que, per $k \geqslant n$, la matriu Q_k compleixi

$$Q_k \left(x_k - x_{k-j}\right) = F(x_k) - F(x_{k-j})$$

per a tot $j = 1, 2, \ldots, n$.

Si els vectors $\left(x_k - x_{k-1}\right), \left(x_k - x_{k-2}\right), \ldots, \left(x_k - x_{k-n}\right)$ són linealment independents, es pot determinar unívocament Q_k resolent n sistemes lineals $n \times n$.

Malauradament, a la pràctica, els vectors anteriors tendeixen a ser linealment dependents i el càlcul de la matriu Q_k és molt inestable, per no parlar de la necessitat per emmagatzemar sempre els $n+1$ iterats anteriors $x_k, x_{k-1}, x_{k-2}, \ldots, x_{k-n}$.

L'*estratègia Broyden* per a resoldre la indeterminació de l'equació (1) en calcular la matriu Q_k és la de *millorar* Q_{k-1} de la manera següent:

$$Q_k := Q_{k-1} + \frac{F(x_k) - F(x_{k-1}) - Q_{k-1}\left(x_k - x_{k-1}\right)}{\left(x_k - x_{k-1}\right)^\top \left(x_k - x_{k-1}\right)} \left(x_k - x_{k-1}\right)^\top \quad (2)$$

Nota de Notació: En tot aquest text hem considerat els vectors com a matrius d'una columna (és a dir, matrius $n \times 1$). Per tant, el segon sumand de la igualtat anterior és una matriu $n \times n$.

Amb la definició (2), Q_k verifica l'equació (1):

$$Q_k\left(x_k - x_{k-1}\right) = Q_{k-1}\left(x_k - x_{k-1}\right) + \frac{F(x_k) - F(x_{k-1}) - Q_{k-1}\left(x_k - x_{k-1}\right)}{\left(x_k - x_{k-1}\right)^{\top}\left(x_k - x_{k-1}\right)}\left(x_k - x_{k-1}\right)^{\top}\left(x_k - x_{k-1}\right) = F(x_k) - F(x_{k-1}).$$

Denotem per $\mathcal{A}_k$ el conjunt de totes les matrius A, $n \times n$, per les quals l'equació (1) es verifica: $A\left(x_k - x_{k-1}\right) = F(x_k) - F(x_{k-1})$. Pel que acabem de dir, $Q_k \in \mathcal{A}_k$.

Proposició (El perquè de tot plegat)

La matriu Q_k definida mitjançant la fórmula (2) verifica

$$\left\|Q_k - Q_{k-1}\right\|_F = \min_{A \in \mathcal{A}_k} \left\|A - Q_{k-1}\right\|_F$$

Norma de Frobenius?

La *norma de Frobenius d'una matriu* $m \times n$ de nombres complexos és la norma $L_{p,q}$ de la matriu amb $p = q = 2$. Aquesta norma es defineix com

$$\|A\|_F = \|A\|_{2,2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\text{traça}(A^\star A)}.$$

on $A^\star$ o A^{H} denota la *transposta Hermitiana d'A* o *transposta conjugada d'A*.

La *transposta Hermitiana d'A* o *transposta conjugada d'A* és la matriu $n \times m$ que s'obté a partir d'A transposant-la i conjugant els seus elements. Es a dir,

$$A^\star = A^{\mathsf{H}} = \left(\overline{a_{ij}}\right)^\top.$$

Òbviament, per a les matrius reals, $A^\star = A^{\mathsf{H}} = A^\top$.

Teorema

Suposem que es compleixen les hipòtesis del Teorema de l'ordre de convergència del mètode de Newton. Llavors, existeixen dues constants positives ε *i* γ *tals que la successió d'iterats* $\{x_k\}_k$ *generada pel mètode de Broyden està ben definida i convergeix superlinealment a* α, *per tot* x_0 *i* Q_0 *tals que*

$$\|x_0 - \alpha\| \leqslant \varepsilon \qquad i \qquad \|\!|Q_0 - J_F(\alpha)|\!\| \leqslant \gamma.$$

Definició (Convergència superlineal)

Es diu que una successió $\{x_k\}_k$ *convergeix superlinealment a* α si

$$\|x_k - \alpha\| \leqslant c_k \|x_{k-1} - \alpha\|,$$

i les constants c_k compleixen $\lim_{k\to\infty} c_k = 0$.

Comentaris

Amb hipòtesis addicionals és possible demostrar que la successió $\{Q_k\}$ convergeix a $J_F(\alpha)$.

Hi ha diverses variants del mètode de Broyden que tenen com a objectiu reduir (encara més) el cost computacional, però solen ser menys estables.