Capítol 3

Resolució de sistemes d'equacions no lineals

3.1 Introducció

En aquest capítol tractarem la resolució numèrica de sistemes d'equacions no lineals.

Trobar zeros d'una funció donada F, és a dir arguments $\bar{x}$ pels que $F(\bar{x}) = 0$ és un problema clàssic. Concretament, si $F : \mathbb{R}^n \to \mathbb{R}^n$ i $F = (f_1, \ldots, f_n)^T$ aleshores el problema de resoldre $f(x) = 0$, $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ es converteix en el de resoldre un sistema d'equacions no lineals

$$f_i(x_1, \ldots, x_n) = 0, \qquad i = 1, \ldots, n.$$

En general, un sistema no lineal no es pot resoldre en un nombre finit de passos. El que farem és usar mètodes iteratius per a generar una successió que tendeixi a un zero de F.

Un problema diferent, però molt relacionat és el següent:

$$\text{minimitzar } h(x), \qquad x \in \mathbb{R}^n,$$

per a una funció real $h : \mathbb{R}^n \to \mathbb{R}$ de n variables $h(x) = h(x_1, \ldots, x_n)$. La funció h s'anomena **funció objectiu** i el problema corresponent **problema de minimització o optimització sense restriccions**. Existeix un altra tipus de problema de minimització, que s'anomena **minimització amb restriccions** i pot èsser formulat de la següent manera: donada la funció objectiu h,

$$\text{minimitzar } h(x) \qquad \text{en } \Omega \subset \mathbb{R}^n.$$

Exemples remarcables d'aquest problema són aquells en els que Ω està caracteritzada per condicions com $g(x) = 0$ (restriccions amb igualtat) o $g(x) \leq 0$ (restriccions amb desigualtat), on $g : \mathbb{R}^n \to \mathbb{R}^m$, amb $m \leq n$, és una funció donada, anomenada **funció de cost**, i la condició $h(x) \leq 0$ significa $h_i(x) \leq 0$, per $i = 1, \ldots, m$. Si g és contínua i Ω és connex, el problema anterior s'anomena un **problema de programació no lineal**. Exemples notables en aquesta area són:

a) **Programació convexa** si h és convexa i g té components convexes.

b) **Programació lineal** si h i g són lineals.

c) **Programació quadràtica** si h és quadràtica i h és lineal.

Els problemes de resolució de sistemes no lineals i el d'optimització estàn íntimament relacionats. En efecte, si h és diferenciable i

$$F(x) = \begin{pmatrix} \frac{\partial h}{\partial x^1} \\ \vdots \\ \frac{\partial h}{\partial x^n} \end{pmatrix}$$

és el gradient de h, aleshores cada punt $\bar{x}$ on s'assoleix el mínim de h és un zero del gradient, $F(\bar{x}) = 0$. Recíprocament, cada zero d'una funció $F = (f_1, \ldots, f_n)$ és també el mínim d'alguna funció, com per exemple $h(x) = \|F(x)\|_2^2$.

En aquest capítol també introduirem el problema de continuació, que consisteix en el següent: Donada una funció $H : U \subset \mathbb{R}^{n+1} \to \mathbb{R}^n$, on U és un conjunt obert, un punt $x_0 \in U$ tal que $H(x_0) = 0$, trobar una corba $x(t)$, per a t definit en un cert interval I, tal que $x(t_0) = 0$, $t_0 \in I$, i $H(x(t)) = 0$, per a tot $t \in I$. Com a aplicació, veurem com ens pot ajudar la resolució d'aquest problema a la resolució de sistemes no lineals.

3.2 Construcció de mètodes iteratius

Sigui $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ una funció contínua definida en un obert $U \subset \mathbb{R}^n$. Com varem fer en el cas de sistemes lineals, volem definir una altra funció contínua $\Phi : V \subset \mathbb{R}^n \to \mathbb{R}^n$, on $V \subset \mathbb{R}^n$ és un obert tal que $U \subset V \neq \emptyset$, tal que si escollim $x^{(0)} \in \mathbb{R}^n$ adequadament i definim la successió $(x^{(k)})_{k \geq 0}$ per

$$x^{(k+1)} = \Phi(x^{(k)}), \qquad k \geq 0,$$

aleshores $(x^{(k)})_{k \geq 0} \to \bar{x}$ i $F(\bar{x}) = 0$. A la funció Φ l'anomenem **funció d'iteració**, i els mètodes obtinguts per al càlcul de zeros de f s'anomenen **mètodes iteratius d'un pas.** L'explicació del nom prové de que per a calcular un terme de la successió, només usem l'anterior. Observem que el mètode de la secant per funcions de $\mathbb{R}$ en $\mathbb{R}$ és un métode de dos passos, ja que per a calcular el següent terme de la successió necessitem els dos anteriors. En general, un mètode iteratiu s'anomena **multipas** d'ordre i, si partim de i valors inicials $x^{(0)}, x^{(1)}, \ldots, x^{(i-1)}$ i definim $x^{(k)}$ com a funció de $x^{(k-i)}$, $x^{(k-i+1)}, \ldots, x^{(k-1)}$, per a $k \geq i$.

Tornant al mètodes iteratius d'un pas, com que suposem que la funció d'iteració és contínua, si $\bar{x} = \lim_{k\to\infty} x^{(k)}$; llavors $\bar{x}$ és un **punt fix** de Φ, és a dir $\Phi(\bar{x}) = \bar{x}$. Per tant, per a que el mètode funcioni correctament, caldrà que els punts fixos de Φ siguin zeros de F.

Les qüestions que ens podem plantejar en aquest moment són:

a) Com podem trobar una funció d'iteració adequada?

b) Sota quines condicions la successió $(x^{(k)})_{k \geq 0}$ convergirà?

c) Amb quina velocitat convergirà la successió $(x^{(k)})_{k \geq 0}$?

En les properes seccions descriurem alguns mètodes basats en una generalització del mètode de Newton per a funcions de diverses variables.

3.2.1 Mètode de Newton-Raphson i variants

Considerem l'equació $F(x) = 0$, on $F = (f_1, \dots, f_n) : U \subset \mathbb{R}^n \to \mathbb{R}^n$ és una funció $C^1(U)$, (és a dir diferenciable amb continuïtat), que suposem que té un zero $\bar{x}$ ($F(\bar{x}) = 0$). Recordem que si $x \in U$ llavors

$$DF(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(x) & \cdots & \frac{\partial f_n}{\partial x_n}(x) \end{pmatrix}.$$

Deducció del mètode de Newton-Raphson

Si fem el desenvolupament de Taylor de F al voltant d'un $x^{(k)}$ proper a $\bar{x}$ fins a primer ordre obtenim

$$0 = F(\bar{x}) \approx F(x^{(k)}) + DF(x^{(k)})(\bar{x} - x^{(k)}).$$

Resolent el sistema $0 = F(x^{(k)}) + DF(x^{(k)})(x - x^{(k)})$ en x, obtenim una nova aproximació de $\bar{x}$, que anomenem $x^{(k+1)}$:

$$x^{(k+1)} = x^{(k)} - [DF(x^{(k)})]^{-1}F(x^{(k)}), \qquad \forall\, k \geq 0. \tag{3.1}$$

Geomètricament, donat $x^{(k)}$, obtenim $x^{(k+1)}$ com a intersecció dels hiperplans tangents a les hipersuperfícies de $\mathbb{R}^{n+1}$, que en coordenades x, y venen donades per les equacions $y = f_i(x)$, en el punt $(x^{(k)}, f_i(x^{(k)}))$, $i = 1, \dots, n$, amb l'hiperplà $y = 0$. En efecte, els hiperplans tangents tenen equació $y = Df_i(x^{(k)})(x - x^{(k)}) + f_i(x^{(k)})$, i per tant, $x^{(k+1)}$ satisfà

$$0 = Df_i(x^{(k)})(x^{(k+1)} - x^{(k)}) + f_i(x^{(k)}), \qquad i = 1, \dots, n$$

que és equivalent a la definció que hem donat.

Tornant a la fórmula (3.1), deduim el següent mètode iteratiu, anomenat **Mètode de Newton-Raphson**:

Donat $x^{(0)} \in U$, definim $x^{(k+1)}$, per a $k = 0, 1, 2, \dots$ com

$$x^{(k+1)} = x^{(k)} + z, \text{ on } z \text{ és la la solució del sistema lineal } DF(x^{(k)})z = -F(x^{(k)}). \tag{3.2}$$

Naturalment, per a que el mètode estigui ben definit cal que $\det DF(x^{(k)}) \neq 0$.

Notem que en aquest cas la funció d'iteració és $\Phi(x) = x - [DF(x)]^{-1}F(x)$ i que si $\det DF(\bar{x}) \neq 0$ i $\Phi(\bar{x}) = \bar{x}$ llavors $F(\bar{x}) = 0$.

Nota 3.2.1 *a) Quan $n = 1$, obtenim el mètode de Newton-Raphson en dimensió 1:*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}.$$

Per tant, (3.1) és una extensió a dimensió més gran que 1 del mètode anterior. Ara bé, hi ha una infinitat de mètodes en dimensió n que poden considerar-se extensions del mètode de Newton en dimensió 1. Per exemple

$$x^{(k+1)} = x^{(k)} - [DF(x^{(k)})]^{-1}F(x^{(k)}) + (n-1)G(x^{(k)}),$$

on $G : \mathbb{R}^n \to \mathbb{R}^n$ és una funció adequada.

b) *Tot i que és equivalent definir el mètode de Newton-Raphson amb la fórmula (3.1) en comptes de (3.2), des d'un punt de vista numèric sempre usarem (3.2), per a evitar el càlcul de la inversa d'una matriu, que és molt costòs.*

c) *Hem deduït el mètode de Newton-Raphson del desenvolupament de Taylor a primer ordre de F. També es pot deduir del desenvolupament F^{-1}: Si suposem que existeix $G = F^{-1}$ localment, fem el desenvolupament de Taylor de G al voltant de $x^{(k)}$ a primer ordre, al voltant de $F(x^{(k)})$:*

$$\bar{x} = G(0) \approx G(F(x^{(k)})) + DG(F(x^{(k)})(0 - F(x^{(k)})) = x^{(k+1)},$$

i com que $G(F(x)) = x$, per tot x i $DG(F(x^{(k)})) = [DF(x^{(k)})]^{-1}$, tornem a obtenir el mètode de Newton-Raphson.

d) *Una generalització natural del mètode de Newton-Raphson és considerar desenvolupaments de Taylor de F^{-1} al voltant de $F(x^{(k)})$ d'ordre més gran que 1. Per exemple, si fem el desenvolupament de Taylor fins a segon ordre, obtenim el* ***mètode de Txebichev***.

El mètode de Newton discretitzat

El mètode de Newton-Raphson pot èsser bastant costós. Cal avaluar les n components de $F(x^{(k)})$ i les n^2 derivades parcials de $DF(x^{(k)})$. A més l'avaluació de les derivades pot èsser molt més costós que el de la funció, fins i tot per $n = 1$. Per exemple, si $f(x) = x^2 3^x \cos 2x$ llavors $f'(x) = 2x3^x \cos 2x + x^2 3^x \ln 3 \cos 2x - 2x^2 3^x \sin 2x$.

Una manera d'evitar el càlcul de les derivades parcials $\partial f_i(x)/\partial x_j$ és aproximar-les per quocients incremetals $\Delta_{ij}(x, h)$ tals que

$$\lim_{h \to 0} \Delta_{ij}(x, h) = \frac{\partial f_i(x)}{\partial x_j}.$$

Per exemple podem prendre

$$\Delta_{ij}(x, h) = \frac{f_i(x + he_j) - f_i(x)}{h},$$

on e_j és el j-èssim vector de la base canònica. D'aquesta manera obtenim els **mètodes de Newton-Raphson discretitzats**:

$$x^{(k+1)} = x^{(k)} - J(x^{(k)}, h)^{-1} F(x^{(k)}), \qquad \text{on } J(x, h) = (\Delta_{ij}(x, h))_{1 \le i,j \le n}.$$

Si es vol, es pot prendre el pas h variable: per a cada component (i, j) i/o per a cada iteració k. Això permet millorar l'ordre de convergència que quan h és constant només és lineal. Per aquests mètodes hem de calcular igualment n^2 valors Δ_{ij}. L'avantatge és que són valors de la funció, no de les derivades. Cal tenir en compte que en molts casos el càlcul de les derivades és molt costòs o fins i tot impossible.

Actualització cíclica de les derivades

En aquesta variant del mètode de Newton, no calculem la diferencial de F en tots els passos, sino que la calculem només de tant en tant i utilitzem el mateix valor de la diferencial en els següents passos. Concretament, sigui $m \ge 1$. Partint d'un valor inicial $x^{(0)}$, definim, per $k \ge 0$,

$$\left.\begin{array}{rcl} x^{(k,0)} & = & x^{(k)} \\ x^{(k,i)} & = & x^{(k,i-1)} - [DF(x^{(k)})]^{-1} F(x^{(k,i-1)}),\ i = 1, \dots, m \\ x^{(k+1)} & = & x^{(k,m)} \end{array}\right\}.$$

Si en la fórmula anterior agafem $k = 0$ i $m \geq 1$ el mètodes l'anomenem **mètode de Newton simplificat**, és a dir

$$x^{(k+1)} = x^{(k)} - DF(x^{(0)})^{-1}F(x^{(k)}), \qquad k \geq 0.$$

Si $m > 1$ està fixat el mètode és equivalent, per a cada $k \geq 1$, a fer un pas de Newton més $(m-1)$ passos de Newton simplificat.

Solució inexacta dels sistemes lineals

Els mètodes iteratius de resolució de sistemes lineals (Jacobi, Gauss-Seidel, S.O.R, etc.) es poden usar també per obtenir aproximacions de la solució del sistema lineal que s'ha de resoldre a cada pas del mètode de Newton. Per exemple, si usem el mètode de Jacobi, podem definir la següent variant del mètode de Newton: Descomposem $DF(x^{(k)}) = L_k + D_k + U_k$, on L_k és triangular inferior estricte, U_k és triagular superior estricte i D_k és diagonal. Partint d'un valor inicial $x^{(0)}$, definim per a $k \geq 0$, $y^{(k,0)} = 0$ i calculem

$$y^{(k,i)} \;\; = \;\; -D_k^{-1}(L_k + U_k)y^{(k,i-1)} - D_k^{-1}F(x^{(k)}),\; i = 1, \ldots m.$$

i definim $x^{(k+1)} = x^{(k)} + y^{(k,m)}$. Aquest és el mètode de **Newton-Jacobi** amb m passos. De la mateixa manera es poden deduir mètodes a partir de Gauss-Seidel i S.O.R.

Mètodes quasi-Newton

Aquests mètodes consisteixen en afegir algun tipus de factor al mètode de Newton. Per exemple:

a) $x^{(k+1)} = x^{(k)} - \omega_k DF(x^{(k)})^{-1}F(x^{(k)})$, amb ω_k adequat per a que $\|F(x^{(k+1)})\| < \|F(x^{(k)})\|$.

b) $x^{(k+1)} = x^{(k)} - (DF(x^{(k)}) + \lambda_k I)^{-1}F(x^{(k)})$, amb λ_k adequat de manera que existeixi inversa i $\|F(x^{(k+1)})\| < \|F(x^{(k)})\|$.

3.2.2 Mètodes derivats dels mètodes iteratius per a sistemes lineals

Els mètodes iteratius de resolució de sistemes lineals (Jacobi, Gauss-Seidel, S.O.R, etc.) es poden estendre al cas no lineal. Per exemple, la generalizació natural de Gauss-Seidel a sistemes no lineals és la següent: Partim d'un valor inicial $x^{(0)} = (x_1^{(0)}, \ldots, x_n^{(0)})$. En el pas k-èssim calculem, per $i = 1, \ldots, n$ $x_i^{(k+1)}$ com la solució de l'equació no lineal en z

$$f_i(x_1^{(k+1)}, \ldots, x_{i-1}^{(k+1)}, z, x_{i+1}^{(k)}, \ldots, x_n^{(k)}) = 0.$$

Per a resoldre cadascuna de les equacions en una variable podem usar qualsevol mètode, per exemple el mètode de Newton-Raphson o el de la secant. Si usem el mètode de Newton, tindrem el mètode de **Gauss-Seidel-Newton**.

Els mètodes d'aquest tipus poden ser competitius quan tenim sistemes de dimensió molt gran però en els que cada equació involucra poques variables. Per exemple per problemes de valors a la frontera no lineals usant diferencies finites o elements finits.

3.3 Teoremes generals de convergència

Donada una funció $F : \mathbb{R}^n \to \mathbb{R}^n$, i una funció d'iteració $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ associada a F, ens preguntem sobre la possible convergència de la successió $(x^{(k)})_{k\geq 0}$ a un zero $\bar{x}$ de la funció f, on $x^{(k+1)} = \Phi(x^{(k)})$, per tot $k \geq 0$. Recordem que $(x^{(k)})_{k\geq 0}$ convergeix a $\bar{x}$ sii per tot $\epsilon > 0$ existeix un enter $N(\epsilon) \geq 0$ tal que si $l \geq N(\epsilon)$ llavors $\|x^{(l)} - \bar{x}\| < \epsilon$. A més, com que $\mathbb{R}^n$ és un espai normat complet, per a veure que una successió és convergent és suficient veure que és de Cauchy: per a tot $\epsilon > 0$ existeix un enter $N(\epsilon)$ tal que si $j, l \geq N(\epsilon)$ llavors $\|x^{(j)} - x^{(l)}\| \leq \epsilon$.

Finalment, recordem que la convergència no depén de la norma escollida, ja que totes són equivalents.

En aquesta secció estudiarem la convergència per contraccions i per al mètode de Newton.

3.3.1 Convergència per a contraccions

Els següents teoremes de convergència demostren que una successió $(x^{(i)})_{i\geq 0}$ generada per $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ convergirà a un punt fix $\bar{x}$ de Φ, si Φ és una contracció. Com abans $\|\cdot\|$ representa una norma en $\mathbb{R}^n$. A més, per a $x \in \mathbb{R}^n$ i $r > 0$, $B(x; r)$ és la bola oberta de centre x i radi r, és a dir $B(x; r) = \{y \in \mathbb{R}^n \mid \|y - x\| < r\}$. La corresponent bola tancada serà $\bar{B}(x; r) = \{y \in \mathbb{R}^n \mid \|y - x\| \leq r\}$.

Abans de enunciar el primer teorema, donarem la definició de contracció:

Definició 3.3.1 *Sigui $\Phi : U \subset \mathbb{R}^n \to \mathbb{R}^n$ una aplicació. Diem que Φ és **contractiva** (o que és una contracció) si existeix $0 \leq K < 1$ tal que $\|\Phi(x) - \Phi(y)\| \leq K\|x - y\|$, per tot $x, y \in U$.*

Nota 3.3.1 *És immediat demostrar que si Φ és contractiva, llavors Φ és uniformement contínua.*

Teorema 3.3.1 *Suposem que $\Phi : B(\bar{x}; r) \subset \mathbb{R}^n \to \mathbb{R}^n$ té un punt fix $\bar{x} = \Phi(\bar{x})$ i que Φ és **contractiva** en $B(\bar{x}; r)$. Sigui $x^{(0)} \in B(\bar{x}; r)$ i $(x^{(i)})_{i\geq 0}$ una successió tal que $x^{(i+1)} = \Phi(x^{(i)})$, per a tot $i \geq 0$. Aleshores, per a tot $i \geq 0$:*

a) $x^{(i)} \in B(\bar{x}; r)$.

b) $\|x^{(i+1)} - \bar{x}\| \leq K\|x^{(i)} - \bar{x}\| \leq K^{i+1}\|x^{(0)} - \bar{x}\|$, *és a dir,* $(x^{(i)})_{i\geq 0}$ *convergeix almenys linealment a* $\bar{x}$.

Demostració:
Immediata per inducció. Si $i = 0$ a) és certa per hipòtesi. A més

$$\|x^{(1)} - \bar{x}\| = \|\Phi(x^{(0)}) - \Phi(\bar{x})\| \leq K\|x^{(0)} - \bar{x}\|,$$

el que prova b).

Suposem que a) i b) són certes per $i - 1 \geq 0$. Aleshores

$$\|x^{(i)} - \bar{x}\| = \|\Phi(x^{(i-1)}) - \Phi(\bar{x})\| \leq K\|x^{(i-1)} - \bar{x}\| < r,$$

ja que $K < 1$ i $\|x^{(i-1)} - \bar{x}\| < r$. Això prova a) per i. Per altra banda,

$$\|x^{(i+1)} - \bar{x}\| = \|\Phi(x^{(i)}) - \Phi(\bar{x})\| \leq K\|x^{(i)} - \bar{x}\| \leq K^{i+1}\|x^{(0)} - \bar{x}\|,$$

amb el que hem demostrat b) per $i + 1$. □

Nota 3.3.2 *Si no sabem si Φ és contractiva però és diferenciable amb continuïtat i $\rho(D\Phi(\bar{x})) < 1$ es pot demostrar que existeix $s \leq r$ tal que si $x^{(0)} \in B(\bar{x}; s)$ llavors $x^{(i)} \in B(\bar{x}; s)$, per a tot $i \geq 0$, i $\lim_{i\to\infty} x^{(i)} = \bar{x}$.*

En càlculs reals no és possible evaluar la funció ϕ exactament. Això vol dir que en comptes de generar una successió $(x^{(i)})_{i\geq 0}$ tal que $x^{(i+1)} = \Phi(x^{(i)})$ generararem $(X^{(i)})_{i\geq 0}$ tal que $X^{(i+1)} = \Phi(X_{(i)}) + \delta_i$, on suposarem que $|\delta_i| \leq \delta$ per tot $i \geq 0$. Llavors tenim el següent teorema

Teorema 3.3.2 *Suposem que Φ satisfà les hipòtesis del teorema anterior. Sigui $X^{(0)} \in \mathbb{R}^n$ un punt tal que*

$$\|X^{(0)} - \bar{x}\| \leq r_0,$$

on

$$0 < r_0 < r - \frac{\delta}{1-K}.$$

Aleshores els iterats $X^{(i)}$ satisfàn

$$\|X^{(i)} - \bar{x}\| < r, \qquad \|X^{(i)} - \bar{x}\| \leq \frac{\delta}{1-K} + K^i\left(r_0 - \frac{\delta}{1-K}\right),$$

on K^i convergeix a 0 quan $i \to \infty$.

Demostració:
Està clar que

$$\|\bar{x} - X^{(0)}\| \leq r_0 < r - \frac{\delta}{1-K} \leq r.$$

Per a una demostració inductiva, suposem que $\|\bar{x} - X^{(j)}\| < r$, per $0 \leq j \leq i-1$. Llavors,

$$\|\bar{x} - X^{(i)}\| = \|\Phi(\bar{x}) - \Phi(X^{(i-1)}) - \delta_{i-1}\| \leq \|\phi(\bar{x}) - \phi(X^{(i-1)})\| + \delta.$$

Per tant, com que Φ és K-contractiva:

$$\begin{aligned}
\|\bar{x} - X^{(i)}\| &\leq K\|\bar{x} - X^{(i-1)}\| + \delta \\
&\leq K^2\|\bar{x} - X^{(i-2)}\| + K\delta + \delta \\
&\leq K^3\|\bar{x} - X^{(i-3)}\| + K^2\delta + K\delta + \delta \\
&\;\vdots \\
&\leq K^i\|\bar{x} - X^{(0)}\| + K^{i-1}\delta + \cdots + K\delta + \delta \\
&\leq K^i r_0 + \frac{1-K^i}{1-K}\delta \\
&\leq K^i r_0 + \frac{\delta}{1-K} - K^i\frac{\delta}{1-K} \\
&\leq r_0 + \frac{\delta}{1-K} \\
&< r
\end{aligned}$$

D'aquesta manera queda demostrat el teorema. □

Nota 3.3.3 *El teorema anterior demostra que el mètode és tant convergent com és possible, és a dir que els errors computacionals que provenen de les evaluacions de $\Phi(x)$ poden produir un error acumulatiu de magnitud com a màxim $\delta/(1-K)$. Està també clar que tals errors limiten el tamany de la fita de l'error independentment del nombre d'iteracions. Per tant, en calculs reals, és innecessari iterar fins que $K^i r_0 << \delta/(1-K)$. De fet, si coneixem estimacions raonables de K, δ i r_0, és un procediment eficient tenir els dos tipus d'error de la mateixa magnitud, és a dir,*

$$K^i r_0 \approx \frac{\delta}{1-K}.$$

El número requerit d'iteracions és aleshores aproximadament

$$i \approx \log\left[\frac{\delta}{(1-K)r_0}\right](\log K)^{-1}.$$

El següent teorema és més precís. Notis que no es suposa l'existència d'un punt fix a priori.

Teorema 3.3.3 *Sigui $\Phi : U \subset \mathbb{R}^n \to \mathbb{R}^n$ una funció d'iteració, i $x^{(0)} \in U$ un valor inicial. Suposem que existeixen $r > 0$ i $0 < K < 1$ tals que*

a) $\bar{B}(x^{(0)}; r) \subset U$.

b) $\|\Phi(x) - \phi(y)\| \leq K\|x - y\|$ *per a* $x, y \in \bar{B}(x^{(0)}; r)$.

c) $\|x^{(1)} - x^{(0)}\| = \|\Phi(x^{(0)}) - x^{(0)}\| \leq (1-K)r < r$.

Aleshores, si definim $x^{(i+1)} = \Phi(x^{(i)})$, per $i \geq 0$, tenim que

1) $x^{(i)}$ *està ben definit per a tot* $i \geq 0$ *i* $x^{(i)} \in B(x^{(0)}; r)$.

2) Φ *té exactament un punt fix* $\bar{x}$ $(\Phi(\bar{x}) = \bar{x})$ *en* $\bar{B}(x^{(0)}; r)$ *i*

$$\lim_{i\to\infty} x^{(i)} = \bar{x}, \qquad \|x^{(i+1)} - \bar{x}\| \leq K\|x^{(i)} - \bar{x}\|,$$

així com

$$\|x^{(i)} - \bar{x}\| \leq \frac{K^i}{1-K}\|x^{(1)} - x^{(0)}\|.$$

Demostració:

1) Ho demostrarem per inducció. En efecte, per $i = 0$, és obviament cert i per $i = 1$ és cert per hipòtesi. Suposem que és cert per a tot $j \leq i$, on $i \geq 1$. Llavors $x^{(j)} \in B(x^{(0)}; r)$, $j \leq i$ i

$$\|x^{(i+1)} - x^{(i)}\| = \|\Phi(x^{(i)}) - \Phi(x^{(i-1)})\| \leq K^i\|x^{(1)} - x^{(0)}\|, \tag{3.3}$$

i per la desigualtat triangular

$$\|x^{(i+1)} - x^{(0)}\| \leq \|x^{(i+1)} - x^{(i)}\| + \|x^{(i)} - x^{(i-1)}\| + \cdots + \|x^{(1)} - x^{(0)}\| \leq (K^i + K^{i-1} + \cdots + 1)\|x^{(1)} - x^{(0)}\| \leq$$

$$\leq (1 + K + \cdots + K^i)(1-K)r = (1 - K^{i+1})r < r,$$

el que prova que $x^{(i+1)} \in B(x^{(0)}; r)$.

2): Primer demostrarem que $(x^{(i)})_{i\geq 0}$ és una successió de Cauchy. Per (3.3) i per b) es segueix que per $m > l$

$$\|x^{(m)} - x^{(l)}\| \leq \|x^{(m)} - x^{(m-1)}\| + \|x^{(m-1)} - x^{(m-2)}\| + \cdots + \|x^{(l+1)} - x^{(l)}\| \leq \tag{3.4}$$

$$\leq K^l(1 + K + \cdots + K^{m-l-1})\|x^{(1)} - x^{(0)}\| < \frac{K^l}{1-K}\|x^{(1)} - x^{(0)}\| < K^l r.$$

Com que $0 < K < 1$, tenim que $K^l r < \epsilon$ per l prou gran. Per tant, $(x^{(i)})_{i\geq 0}$ és una successió de Cauchy, i en conseqüència, existeix un límit

$$\lim_{i\to\infty} x^{(i)} = \bar{x}.$$

Com que $x^{(i)} \in B(x^{(0)}; r)$, per a tot i, $\bar{x} \in \bar{B}(x^{(0)}; r)$. A més, $\bar{x}$ és un punt fix de Φ, doncs per tot $i \geq 0$

$$\|\Phi(\bar{x}) - \bar{x}\| \leq \|\Phi(\bar{x}) - \Phi(x^{(i)})\| + \|\Phi(x^{(i)}) - \bar{x}\| \leq K\|\bar{x} - x^{(i)}\| + \|x^{(i+1)} - \bar{x}\|.$$

Com que $\lim_{i\to\infty} \|x^{(i)} - \bar{x}\| = 0$, tenim que $\|\Phi(\bar{x}) - \bar{x}\| = 0$, i per tant $\Phi(\bar{x}) = \bar{x}$.

Si $\bar{y} \in \bar{B}(x^{(0)}; r)$ és un altre punt fix de Φ, aleshores

$$\|\bar{x} - \bar{y}\| = \|\Phi(\bar{x}) - \Phi(\bar{y})\| \leq K\|\bar{x} - \bar{y}\|,$$

$0 < K < 1$, el que implica que $\|\bar{x} - \bar{y}\| = 0$, o $\bar{x} = \bar{y}$.

Finalment, (3.4) implica, fent tendir $m \to \infty$

$$\|\bar{x} - x^{(l)}\| = \lim_{m\to\infty} \|x^{(m)} - x^{(l)}\| \leq \frac{K^l}{1-K}\|x^{(1)} - x^{(0)}\|$$

i

$$\|x^{(i+1)} - \bar{x}\| = \|\Phi(x^{(i)}) - \Phi(\bar{x})\| \leq K\|x^{(i)} - \bar{x}\|,$$

que conclou la demostració. □

Nota 3.3.4 *a) La velocitat de convergència de $(x^{(i)})_{i\geq 0}$ cap a $\bar{x}$ depèn de K. A més la fita*

$$\|x^{(i)} - \bar{x}\| \leq K^i(1-K)^{-1}\|x^{(1)} - x^{(0)}\|$$

permet estimar a priori la quantitat d'iteracions necessàries per a obtenir una precisió determinada. Però, només val a $B(x^{(0)}; r)$ i sovint és molt grollera, i per tant, pessimista.

b) Si Φ és diferenciable amb continuïtat i $\|D\Phi(x)\| \leq K < 1$ per tot $x \in B(\bar{x}; r)$, llavors Φ és K contractiva, i podem substituir aquesta condició en l'enunciat del teorema. En efecte, pel teorema del valor mitjà, si $x, y \in B(\bar{x}; r)$ llavors

$$\|\Phi(x) - \Phi(y)\| \leq \sup_{z\in\overline{xy}} \|D\Phi(z)\|\|x - y\| \leq K\|x - y\|,$$

on $\overline{xy} \subset B(\bar{x})$ és el segment que uneix x i y.

3.3.2 Convergència del mètode de Newton

En aquesta secció veurem la convergència quadràtica local del mètode de Newton per sistemes d'equacions no lineals. Per abreujar, denotarem la matriu jacobiana associada a una funció $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$, U obert, en un punt $x \in U$ com $J(x)$, és a dir $J(x) = DF(x)$.

Teorema 3.3.4 *Sigui $F : D \subset \mathbb{R}^n \to \mathbb{R}^n$ una funció diferenciable amb continuïtat en un conjunt obert i convex $D \subset \mathbb{R}^n$. Suposem que existeix $\bar{x} \in \mathbb{R}^n$ i $r, \beta > 0$ tals que $B(\bar{x}; r) \subset D$, $F(\bar{x}) = 0$, existeix $J(\bar{x})^{-1}$ amb $\|J(\bar{x})^{-1}\| \leq \beta$, i per tot $x, y \in B(\bar{x}; r)$*

$$\|J(x) - J(y)\| \leq \gamma \|x - y\|.$$

Si $x^{(0)} \in S_\epsilon(\bar{x})$, on $\epsilon = \min\{r, (2\beta\gamma)^{-1}\}$, llavors la successió $(x^{(k)})_{k\geq 0}$ definida per

$$x^{(k+1)} = x^{(k)} - J(x^{(k)})^{-1} F(x^{(k)}), \qquad k = 0, 1, \dots$$

està ben definida, convergeix a $\bar{x}$ i

$$\|x^{(k+1)} - \bar{x}\| \leq \beta\gamma \|x^{(k)} - \bar{x}\|^2.$$

El següent teorema no usa l'existència de cap zero.

Teorema 3.3.5 *(Newton-Kantorovich) Sigui $F : U \subset \mathbb{R}^n \to \mathbb{R}^n$ una funció de classe $C^2(U)$ i $x^{(0)} \in U$, i suposem que existeixen constants $a, b, c > 0$ verificant:*

a) $\bar{B}(x^{(0)}; 2b) \subset U$.

b) $F(x^{(0)})$ és regular i $\|DF(x^{(0)})^{-1}\| \leq a$,

c) $\|x^{(1)} - x^{(0)}\| = \|DF(x^{(0)})^{-1} F(x^{(0)})\| \leq b$, i

d) per tot x verificant $\|x - x^{(0)}\| \leq 2b$ es té $\|D^2 F(x)\| \leq c$.

Aleshores, si $h = abc \leq \frac{1}{2}$, es verifica:

1) Els iterats del mètode de Newton

$$x^{(k+1)} = x^{(k)} - DF(x^{(k)})^{-1} F(x^{(k)}),$$

estan ben definits i compleixen $\|x^{(k)} - x^{(0)}\| \leq 2b$.

2) La successió $(x^{(k)})_{k\geq 0}$ convergeix a $\bar{x}$ tal que $f(\bar{x}) = 0$ i

$$\|x^{(k)} - \bar{x}\| \leq \frac{(2h)^{2^k - 1}}{2^{k-1}} b.$$

3) $\bar{x}$ és l'únic zero de F en $\bar{B}(x^{(0)}; 2b)$.

La següent proposició és útil per a usar el teorema de Newton-Kantorovich:

Proposició 3.3.1 *Sigui $F = (f_1, \ldots, f_n) : U \subset \mathbb{R}^n \to \mathbb{R}^n$, on U és un subconjunt obert, una funció de classe C^2. Si per $x \in U$*

$$\sum_{k=1}^{n} \left| \frac{\partial^2 f_i(x)}{\partial x_j \partial x_k} \right| \le \frac{c}{n} \qquad per \qquad i, j = 1, 2, \ldots, n,$$

aleshores $\|D^2 F(x)\|_\infty \le c$, per a tot $x \in U$.

Demostració:
Per definició

$$\|D^2F(x)\|_\infty = \max_{\|u\|_\infty \le 1, \|v\|_\infty \le 1} \|D^2F(x)(u, v)\|_\infty.$$

Però

$$\|D^2F(x)(u, v)\|_\infty \le \sum_{j=1}^{n} \|D^2F(x)(e_j, v)\|_\infty \le \sum_{j=1}^{n} \sum_{k=1}^{n} \|D^2F(x)(e_j, e_k)\|_\infty \le c.$$

□

3.4 Mètodes de continuació

Els mètodes de continuació serveixen per obtenir solucions globals del problema general

$$H(x, \lambda) = 0, \tag{3.5}$$

on $H : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$. És a dir, cerquem una solució com a una corba parametritzada $c(s) = (x(s), \lambda(s))$ tal que $H(c(s)) = 0$ per a s en un cert interval de $\mathbb{R}$.

Els mètodes de homotopia es poden considerar com a precursors dels mètodes de continuació, i consisteixen en el següent: Suposem que volem obtenir una solució d'un sistema de n equacions no lineals amb n variables, per exemple

$$F(x) = 0,$$

on $F : \mathbb{R}^n \to \mathbb{R}^n$ és una aplicació suficientment diferenciable. Si no tenim un coneixement a priori respecte als zeros de F, no podrem usar mètodes iteratius com el mètode de Newton, degut a que és precís partir d'una bona condició inicial. Un possible remei a aquest problema és definir una **homotopia** o deformació $H : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ tal que

$$H(x, 0) = G(x), \qquad H(x, 1) = F(x),$$

on $G : \mathbb{R}^n \to \mathbb{R}^n$ és una aplicació que té zeros coneguts (normalment G és una simplificació de F en la que s'ha suprimit els termes "difícils" de F). Típicament hom pot escollir una **homotopia convexa** tal que:

$$H(x, \lambda) = \lambda F(x) + (1 - \lambda)G(x),$$

i tractar de traçar una corba $x = c(s) \in H^{-1}(0)$ definida implícitament des de un punt inicial $(\tilde{x}, 0)$ a un punt solució $(\bar{x}, 1)$. Si es té èxit, obtenim un zero de F. Una altra deformació standard que s'usa freqüentment és la **homotopia global**

$$H(x, \lambda) = F(x) - (1 - \lambda)F(\tilde{x}).$$

Volem fer notar que en un mètode general de continuació ens interessa tota la corba $c(s)$, mentre que en un mètode d'homotopia només ens interessa $x(1)$, encara que també es calculen valors intermitjos.

Arrel de la descripció del mètode de homotopia, ens podem plantejar les següents qüestions:

a) Quan podem assegurar que la corba $c(s) \in H^{-1}(0)$ amb $(\tilde{x}, 0)$ pertanyent al rang de $c(s)$, existeix i és diferenciable?

b) Si la corba $c(s)$ existeix, quan podrem assegurar que tindrà intersecció amb el nivell $\lambda = 1$ de l'homotopia?

c) Com podem calcular numèricament la corba $c(s)$?.

La idea més simple per a resoldre aquest problema és la següent:

Considerem com a paràmetre $s = \lambda$, amb el que hem de trobar una corba $x(\lambda)$ tal que $H(x(\lambda), \lambda) = 0$. Usant mètodes numèrics, podem esperar trobar un conjunt discret de valors $x(\lambda_i)$, $i = 1, 2, 3, \ldots$ Si H és bastant regular (per exemple C^1) i si fixem dos paràmetres diferents i pròxims $\lambda_1 \approx \lambda_2$ és d'esperar que les equacions que cal resoldre $H(x, \lambda_1) = 0$ i $H(x, \lambda_2) = 0$ siguin molt semblants i que les seves solucions respectives siguin pròximes. La idea de la continuació és aprofitar la solució d'una per a resoldre l'altra, i successivament, altres corresponents a paràmetres diferents: $\lambda_1, \lambda_2, \lambda_3, \ldots$. i així calcular la corba $x(\lambda)$ en un conjunt discret de punts. El mètode més simple per a resoldre aquest problema s'anomena **mètode primari de continuació**, que descrivim a continuació:

Sigui $H : \mathbb{R}^n \times [a, b] \to \mathbb{R}^n$, i sigui $x^{(0)}$ una solució coneguda per $\lambda = a$, és a dir, $H(x^{(0)}, a) = 0$. Llavors:

a) Considerem una partició suficientment fina de l'interval $[a, b]$:

$$a = \lambda^{(0)} < \lambda^{(1)} < \cdots < \lambda^{(N)} = b,$$

i el conjunt discret d'equacions associat:

$$H(x, \lambda^{(i)}) = 0, \qquad i = 0, \ldots, N.$$

b) Per tot $i = 1 \ldots, N$, resolem l'equació i-èssima usant algún mètode iteratiu d'un pas localment convergent, agafant com aproximació inicial, la solució ja coneguda de l'equació $(i-1)$-èssima, és a dir, trobem $x^{(i)}$ tal que $H(x^{(i)}, \lambda^{(i)}) = 0$ usant un métode iteratiu amb valor inicial $x^{(i-1)}$.

Per exemple, usant el mètode de Newton, per $i = 1, \ldots, N$ fem

$$\left\{ \begin{array}{rcl} x^{(i,0)} & = & x^{(i-1)}, \\ x^{(i,k)} & = & x^{(i,k-1)} - (D_x H(x^{(i,k-1)}, \lambda^i))^{-1} H(x^{(i,k-1)}, \lambda^i), \qquad k = 1, \ldots, m_i, \\ x^{(i)} & = & x^{(i,m_i)}. \end{array} \right.$$

El nombre d'iteracions m_i necessàries en cada pas fins a obtenir la precisió desitjada depén del pas i.

Notem que aquest mètode té dos problemes. El primer és que per assegurar la convergència pot caldre una partició molt fina, amb el conseqüent cost en nombre d'operacions. Això es pot resoldre construint una millor aproximació inicial $x^{(i,0)}$ a cada pas. Per altra banda, la corba pot no està ben definida o tenir punts de retorn. En aquest últim cas, caldrà canviar el paràmetre λ per un altre de més adient.

3.4.1 Corbes definides implícitament

Considerem l'equació $H(x,\lambda)=0$, $x\in\mathbb{R}^n$, $\lambda\in\mathbb{R}$, on H és de classe C^r, per r gran o C^∞. Si no volem distingir un paràmetre en particular, escriurem $H(x)=H(x_1,\dots,x_{n+1})$, i escriurem $H:\mathbb{R}^{n+1}\to\mathbb{R}$. El nostre primer objectiu és veure quan existeix una funció $x=x(\lambda)$ tal que $H(x(\lambda),\lambda)=0$, per a λ contingut en un cert interval.

L'eina bàsica per a assegurar que la corba solució de $H(x,\lambda)=0$,es pot continuar, és el teorema de la funció implícita:

Cas de dimensió 1 ($n=1$)

Considerem, per a fixar idees l'equació $x^3+x-1=0$. Per a obtenir solucions d'aquesta equació podem definir una funció

$$h(x,\lambda)=x^5+x-\lambda=0.$$

Quan $\lambda=0$ tenim la solució $x=0$ i ens preguntem quina serà la solució si $\lambda=1$. Denotem per $x(\lambda)$ la solució per a cada λ de $h(x,\lambda)=0$, és a dir $h(x(\lambda),\lambda)=0$. Ens podem preguntar si la corba $x(\lambda)$ existeix. La resposta des de un punt de vista local la dona el teorema de la funció implícita:

Teorema 3.4.1 *Sigui (x_0,λ_0) tal que $h(x_0,\lambda_0)=0$. Si $\frac{\partial h}{\partial x}(x_0,\lambda_0)\neq 0$, llavors existeix un interval $(\lambda_0-\delta,\lambda_0+\delta)$ i una funció*

$$\bar{x}:(\lambda_0-\delta,\lambda_0+\delta)\to\mathbb{R}$$

de classe C^r tal que

a) $\bar{x}(\lambda_0)=x_0$.

b) $h(\bar{x}(\lambda),\lambda)=0$, $\forall\,\lambda\in(\lambda_0-\delta,\lambda_0+\delta)$.

Si ho apliquem al nostre exemple, agafant $(x_0,\lambda_0)=(0,0)$, com que $\frac{\partial h}{\partial x}=5x^4+1$, tenim que $\frac{\partial h}{\partial x}(0,0)=1$, i aplicant el teorema, existeix un interval $(-\delta,\delta)$ i una funció $\bar{x}:(-\delta,\delta)\to\mathbb{R}$ tal que

$$\left.\begin{array}{rcl}\bar{x}(0)&=&0\\ h(\bar{x}(\lambda),\lambda)&=&0\quad\forall\,\lambda\in(-\delta,\delta)\end{array}\right\}$$

Observem, a més, que podem calcular $\bar{x}'(0)$. Com que $h(\bar{x}(\lambda),\lambda)=0$, derivant respecte de λ obtenim

$$\frac{\partial h}{\partial x}(\bar{x}(\lambda),\lambda)\frac{d\bar{x}}{d\lambda}+\frac{\partial h}{\partial\lambda}(\bar{x}(\lambda),\lambda)=0.$$

Si agafem $(\bar{x}(\lambda_0),\lambda_0)=(x_0,\lambda)$, llavors

$$\frac{\partial h}{\partial x}(x_0,\lambda_0)\frac{d\bar{x}}{d\lambda}(\lambda_0)+\frac{\partial h}{\partial\lambda}(x_0)=0.$$

Per tant,

$$\frac{d\bar{x}}{d\lambda}(\lambda_0)=-\left(\frac{\partial h}{\partial x}(x_0,\lambda_0)\right)^{-1}\frac{\partial h}{\partial\lambda}(x_0,\lambda_0).$$

És a dir, que si usem el desenvolupament de Taylor a primer ordre, tenim

$$\bar{x}(\lambda)=\bar{x}(\lambda_0)+\frac{d\bar{x}}{d\lambda}(\lambda_0)(\lambda-\lambda_0)+O(|\lambda-\lambda_0|^2).$$

Si ho apliquem al nostre exemple, tenim $\lambda_0 = 0$, $x_0 = 0$, i hem vist que $\frac{\partial h}{\partial x}(0,0) = 1$. Per tant,

$$\frac{d\bar{x}}{d\lambda}(0) = -\frac{\partial h}{\partial \lambda}(0,0) = 1,$$

i

$$\bar{x}(\lambda) = \lambda + O(|\lambda|^2), \qquad \lambda \to 0.$$

Per exemple, si agafem $\lambda = 0.1$, tenim que la solució està a prop de $0.1 : 0.09999000499\ldots$.

Cas dimensió qualsevol

Si la dimensió és n tot és similar: $H(x,\lambda) = 0$, i sabem que per $x_0 \in \mathbb{R}^n$ i $\lambda_0 \in \mathbb{R}$, $H(x_0,\lambda_0) = 0$.

Teorema 3.4.2 *(Funció Implícita) Si* $\det D_x H(x_0,\lambda_0) \neq 0$ *llavors existeix un interval* $(\lambda_0-\delta, \lambda_0+\delta)$ *i una funció* C^r

$$\bar{x} : (\lambda_0 - \delta, \lambda_0 + \delta) \to \mathbb{R}^n$$

tal que

a) $\bar{x}(\lambda_0) = x_0$,

b) $H(\bar{x}(\lambda),\lambda) = 0$, $\forall\, \lambda \in (\lambda_0 - \delta, \lambda_0 + \delta)$.

Quan val $d\bar{x}/d\lambda$ en aquest cas? Com hem fet abans, escrivim $H(x(\lambda),\lambda) = 0$ i derivem respecte de λ :

$$D_x H(x(\lambda),\lambda)\frac{d\bar{x}}{d\lambda}(\lambda) + D_\lambda H(x(\lambda),\lambda) = 0.$$

Per tant,

$$D_x H(x(\lambda),\lambda)\frac{d\bar{x}}{d\lambda} = -D_\lambda H(x(\lambda),\lambda).$$

Si ho apliquem al punt (x_0,λ_0) obtenim

$$\frac{d\bar{x}}{d\lambda}(\lambda_0) = -\left[D_x H(x_0,\lambda_0)\right]^{-1} D_\lambda H(x_0,\lambda_0).$$

De fet, resoldrem el sistema lineal

$$D_x H(x_0,\lambda_0) z = -D_\lambda H(x_0,\lambda_0),$$

on $z = \frac{d\bar{x}}{d\lambda}(\lambda_0)$. Aleshores

$$\bar{x}(\lambda) = x_0 + \frac{d\bar{x}}{d\lambda}(\lambda_0)(\lambda - \lambda_0) + O(|\lambda - \lambda_0|^2).$$

El mètode de continuació pot fallar en algun pas degut a l'existència de singularitats de la corba. Prop d'aquests punts existeix més d'una solució i el teorema de la funció implícita no és vàlid. En la figura 3.1 hem dibuixat les gràfiques corresponents als tres exemples següents: a) $x^3 - x + \lambda = 0$, (punt de retorn), b) $x^3 + \lambda x = 0$, i c) $x^2 - \lambda^2 = 0$ (punts singulars o forcacions). En l'exemple a) tenim que $D_x H(x_\pm,\lambda_\pm) = 0$ i $H(x_\pm,\lambda_\pm) = 0$, on

$$x_\pm = \pm\frac{1}{\sqrt{3}}, \qquad \lambda_\pm = \pm\frac{2}{3\sqrt{3}},$$

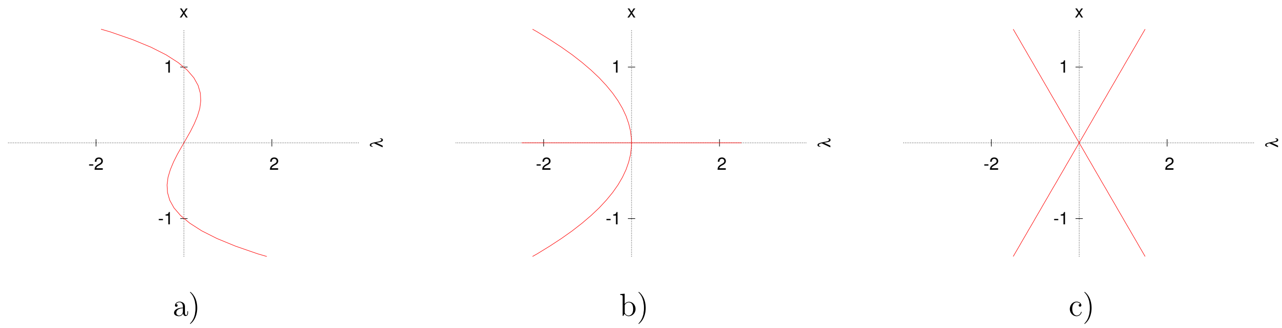

Figura 3.1: Punts de retorn i forcacions en tres exemples

però

$$DH(x_\pm, \lambda_\pm) = (0\ 1),$$

és a dir que el rang de $DH(x_\pm, \lambda_\pm)$ és màxim (igual a 1).

En l'exemple b) tenim que $D_x H(0,0) = 0$, però també $DH(0,0) = (0\ 0)$. De la mateixa manera en c) tenim que $DH(0,0) = (0\ 0)$.

Observant aquests exemples, podem dir que hi haurà un **punt de retorn** en (x_0, λ_0) si $H(x_0, \lambda_0) = 0$, $D_x H(x_0, \lambda_0)$ no és regular (com a matriu $n \times n$), mentre que $DH(x_0, \lambda_0)$ com a matriu $n \times (n+1)$ té rang màxim. Notem que en aquest cas la corba es pot continuar si canviem els paràmetres. Direm que hi ha un punt singular o **bifurcació** si $DH(x_0, \lambda_0)$ no té rang màxim. En el cas de que hi hagi una bifurcació no podrem continuar la corba de cap manera. Una vegada detectada una bifurcació, es fa un estudi local específic per a determinar quantes branques hi ha prop de (x_0, λ_0) i quina forma tenen.

La determinació d'un mètode específic de continuació correspon a les diverses eleccions dels ingredients següents:

a) **Mètode predictor** per trobar una aproximació inicial $(x^{(i,0)}, \lambda_{i,0})$ d'un nou punt de la corba.

b) **Parametrització** usada per a identificar punts de la corba (que pot èsser diferent del paràmetre **natural** λ del problema, el que pot èsser pot adequat, com hem vist abans.

c) **Mètode corrector** iteratiu per a **refinar** l'aproximació inicial fins a obtenir un punt $(x^{(i)}, \lambda_i)$ sobre la corba (amb la tolerància desitjada).

d) **Control de pas** per a optimitzar el cost de la continuació.

El mètode de continuació que descriurem, basat en les idees anteriors, és el **mètode predictor-corrector** amb **pseudo-paràmetre arc**.

Sigui $y = (x, \lambda) \in \mathbb{R}^{n+1}$ i $H : \mathbb{R}^{n+1} \to \mathbb{R}$. El sistema d'equacions $H(y) = 0$ es pot veure geomètricament com la intersecció de n hipersuperfícies a $\mathbb{R}^{n+1}$. Genèricament, aquesta intersecció serà una corba. Suposem que tenim un punt d'aquesta corba $y_0 = (x_0, \lambda_0)$ i volem seguir la corba en l'espai (x, λ). El primer pas serà calcular el vector tangent a la corba. Diem $y(s)$ a una parametrizació (per exemple el paràmetre arc) d'aquesta corba. Llavors $H(y(s)) = 0$, i derivant respecte de s :

$$D_y H(y(s)) \frac{dy}{ds}(s) = 0.$$

Notem que la matriu $D_yH(y(s))$ té n files i $n+1$ columnes, i per tant té nucli. Si el nucli és de dimensió 1 (corba) llavors calculem un vector v d'aquest nucli, que agafarem normalitzat[1]. Notem que tenim encara dues opcions: v i $-v$. Agafarem el que toqui segons cap a on volem fer anar la solució. Així obtenim una aproximació d'un punt de la corba a distància h de y_0 :

$$\tilde{y}_1 = y_0 + hv,$$

Direm que $\tilde{y}_1$ és una predicció d'un punt de la corba a distància h de y_0.

Per a la correcció de $\tilde{y}_1$ usarem el mètode de Newton. Com que el sistema a resoldre no és quadrat, una opció és fer-lo quadrat afegint una equació més:

$$\left.\begin{array}{rcl} H(y) & = & 0 \\ \|y-y_0\|_2^2 - h^2 & = & 0 \end{array}\right\}$$

D'aquesta manera, usant el mètode de Newton amb valor inicial $\tilde{y}_1$, obtindrem un nou punt de la corba y_1 a distància h del punt anterior y_0. Si iterem aquest procés, obtindrem punts $y_2, y_3, \ldots,$ de la corba tals que $\|y_{i+1} - y_i\|_2 = h$. Si estem usant un mètode d'homotopia, continuarem fins que la darrera component de y_i sigui igual a 1. Si volem calcular una corba tancada continuarem fins completar la corba.

[1]Recordem que si s és el paràmetre arc, llavors $\|dy/ds\|_2 = 1$.