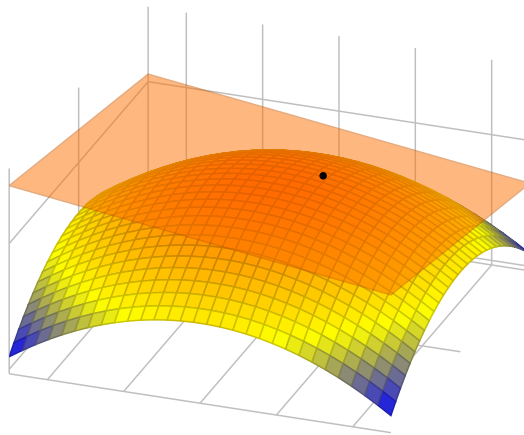


Aureli Alabert

# Optimització No Lineal

per a estudiants de  
matemàtica computacional

*teoria i exercicis*



©Aureli Alabert 2019-2024

Es permet la reproducció per a estudi privat.  
Per a qualsevol altre ús, contacteu amb l'autor.

---

# Teoria d'extremes en $\mathbb{R}^n$

---

## 1.1. El problema general d'optimització en $\mathbb{R}^n$

En aquest primer apartat veurem què es pot dir del *problema general d'optimització en  $\mathbb{R}^n$* ,

$$\min_{x \in S} f(x),$$

on  $S$  és un subconjunt qualsevol de  $\mathbb{R}^n$ , i  $f$  és una funció a valors en  $\mathbb{R}$ .

**Definició 1.1.1.** Un punt  $x^* \in S$  és un *mínim global* (o *absolut*) de  $f$  en  $S$  si

$$\forall x \in S, f(x) \geq f(x^*).$$

El mínim global s'anomena *estricte* si

$$\forall x \in S, x \neq x^* \Rightarrow f(x) > f(x^*).$$

□

Està clar que no pot haver-hi més d'un mínim global estricte.

**Definició 1.1.2.** Un punt  $x^* \in S$  és un *mínim local* (o *relatiu*) de  $f$  en  $S$  si

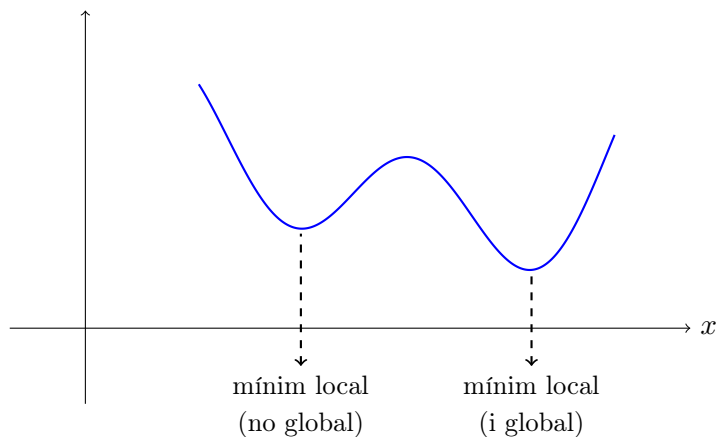
$$\exists \varepsilon > 0: \forall x \in S, \|x - x^*\| < \varepsilon \Rightarrow f(x) \geq f(x^*).$$

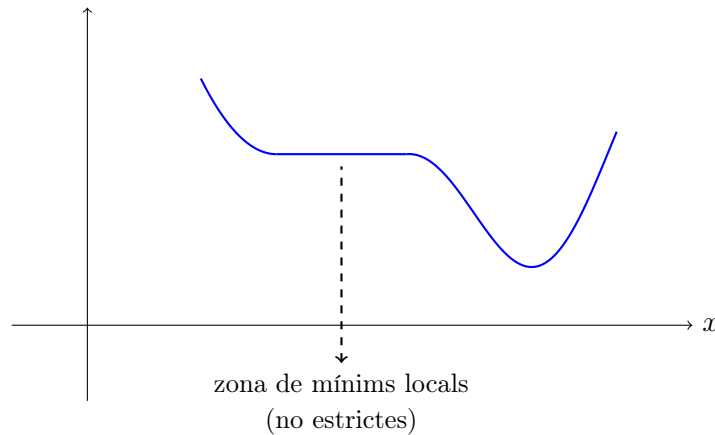
Un mínim local és *estricte* si

$$\exists \varepsilon > 0: \forall x \in S, \|x - x^*\| < \varepsilon, x \neq x^* \Rightarrow f(x) > f(x^*).$$

□

Podem visualitzar fàcilment les nocions anteriors en dimensió 1:





En principi, el que voldríem és trobar un mínim global de la funció  $f$  sobre el conjunt  $S$ . Malauradament aquest és un problema molt difícil en general. Els mètodes que veurem estan enfocats cap a la cerca de mínims locals.

Sobre l'existència de mínims, la resposta més senzilla ve donada pel Teorema de Weierstrass:

**Teorema 1.1.3** (Teorema de Weierstrass).

Si  $f$  és contínua i  $S$  és un conjunt compacte de  $\mathbb{R}^n$ , llavors existeix mínim global de  $f$  en  $S$ .  $\square$

**Definició 1.1.4.** Donat un problema d'optimització

$$\min_{x \in S} f(x),$$

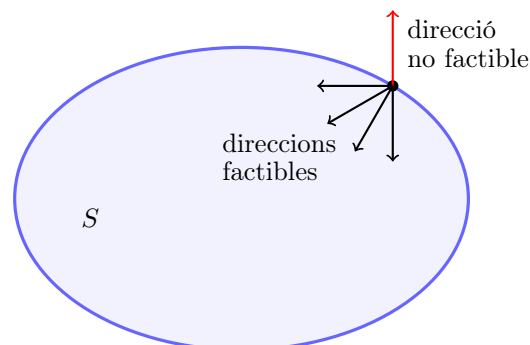
el conjunt  $S$  s'anomena *conjunt factible* i els punts  $x \in S$  són els *punts factibles* del problema. Els punts  $x \notin S$  són *punts no-factibles*.  $\square$

Veurem que els mètodes pràctics per trobar mínims consisteixen en anar-nos movent d'un punt factible a un altre punt factible millor (amb valor menor de la funció  $f$ ), tantes vegades com calgui. Això motiva la definició següent:

**Definició 1.1.5.** Donat  $x \in S$ , un vector  $d \in \mathbb{R}^n$  és una *direcció factible* en  $x$  si

$$\exists \varepsilon > 0 : \alpha \in [0, \varepsilon] \Rightarrow x + \alpha d \in S.$$

En paraules: "el segment  $x + \alpha d$  està en  $S$ , per  $\alpha > 0$  prou petits". Per exemple:



Observeu que si  $x \in \overset{\circ}{S}$  (l'interior de  $S$ ), tot vector  $d \in \mathbb{R}^n$  és una direcció factible.

**Notacions 1.1.6.** Denotarem per  $\nabla f(x)$  el *vector gradient* de  $f$  en el punt  $x$ ; és a dir:

$$\nabla f(x) := \left( \frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Recordem que, si  $d$  és un vector de  $\mathbb{R}^n$ , el producte escalar  $\nabla f(x) \cdot d$  coincideix amb la derivada direccional de  $f$  en  $x$  segons el vector  $d$ :

$$\nabla f(x) \cdot d = \lim_{h \rightarrow 0} \frac{f(x + hd) - f(x)}{h}.$$

Denotarem per  $\nabla^2 f(x)$  la *matriu hessiana* de  $f$  en el punt  $x$ ; és a dir:

$$\nabla^2 f(x) := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}.$$

Naturalment, el gradient i la matriu hessiana estan definits si existeixen totes les derivades parcials necessàries. Una hipòtesi habitual és suposar que les derivades parcials, a més d'existir, són funcions contínues de  $x$ . Quan les primeres derivades són contínues, diem que  $f$  és de classe  $C^1$ , i escrivim  $f \in C^1$ . Quan ho són les segones derivades, diem que  $f$  és de classe  $C^2$ , i escrivim  $f \in C^2$ , i així successivament. Evidentment,  $f \in C^2 \Rightarrow f \in C^1$ .  $\square$

**Teorema 1.1.7** (Condicions necessàries de mínim local).

Suposem  $f \in C^2$ .

Si  $x^*$  és mínim local de  $f$  en  $S$ , llavors per a tota direcció factible  $d$  en  $x^*$  es té:

- i)  $\nabla f(x^*) \cdot d \geq 0$ .
- ii) Si  $\nabla f(x^*) \cdot d = 0$ , llavors  $d \cdot \nabla^2 f(x^*) \cdot d \geq 0$ .<sup>1</sup>

*Demostració.*

- i) Intuïtivament, podem pensar en el que sabem en una variable: Com que  $\nabla f(x^*) \cdot d$  és la derivada direccional de  $f$  en la direcció  $d$ , si en  $x^*$  hi ha un mínim, la funció no pot decreixer en la direcció  $d$ , i aquesta derivada ha de ser més gran o igual que zero.

Formalment: Considerem la funció d'una variable  $g(\alpha) := f(x^* + \alpha d)$ , per  $\alpha \in [0, \varepsilon]$ , amb  $\varepsilon$  tan petit com calgui. Tindrem que:

- $g(\alpha) \geq g(0)$  per  $\alpha \in [0, \varepsilon]$ ,
- $g'(\alpha) = \nabla f(x^* + \alpha d) \cdot d$  per  $\alpha \in (0, \varepsilon)$  (regla de la cadena),
- $g'(0)$  existeix com a límit per la dreta en el zero de  $g'(\alpha)$  (perquè  $f \in C^1$ ) i és igual a  $\nabla f(x^*) \cdot d$ .

Podem escriure aleshores, pel Teorema de Taylor,

$$0 \leq g(\alpha) - g(0) = g'(0)\alpha + o(\alpha), \quad \alpha \in [0, \varepsilon].$$

Dividint per  $\alpha$ , que és positiu, i prenent límit quan  $\alpha \rightarrow 0$ , la desigualtat es conserva i tenim  $\nabla f(x^*) \cdot d = g'(0) \geq 0$ .

- ii) Pel Teorema de Taylor,

$$0 \leq g(\alpha) - g(0) = \frac{1}{2}g''(0)\alpha^2 + o(\alpha^2),$$

i, aplicant la regla de la cadena,  $d \cdot \nabla^2 f(x^*) \cdot d = g''(0) \geq 0$ .

$\square$

**Exercici 1.1.8.** Completeu els detalls de l'apartat ii) de la demostració anterior. En particular, comproveu l'aplicació de la regla de la cadena.  $\square$

<sup>1</sup>Notació: Si no hi ha confusió possible, quan multipliquem un vector per la dreta a una matriu evitarem posar el símbol de transposat. Formalment, hauríem d'escriure  $d^T \cdot \nabla^2 f(x^*) \cdot d \geq 0$

**Corol·lari 1.1.9** (Condicions necessàries de mínim local, cas interior).

En la situació anterior, si  $x^* \in \overset{\circ}{S}$ , llavors:

- i)  $\nabla f(x^*) = 0$ .
- ii)  $\nabla^2 f(x^*)$  és semidefinida positiva.

*Demostració.*

- i) Tota direcció  $d$  és factible i compleix  $\nabla f(x^*) \cdot d \geq 0$ . Agafant com a  $d$  els elements de la base canònica i després ells mateixos canviats de signe, obtenim  $\nabla f(x^*) = 0$ .
- ii)  $d \cdot \nabla^2 f(x^*) \cdot d \geq 0, \forall d \in \mathbb{R}^n$  és precisament la definició de semidefinida positiva. □

**Observació 1.1.10.** Els apartats i) de la Proposició i el Corol·lari anteriors són certs suposant només  $f \in C^1$ . □

**Teorema 1.1.11** (Condicions suficients de mínim local, cas interior).

Suposem  $f \in C^2$ .

Sigui  $x^* \in \overset{\circ}{S}$ . Si

- a)  $\nabla f(x^*) = 0$ ,
- b)  $\nabla^2 f(x^*)$  és definida positiva (és a dir,  $d \cdot \nabla^2 f(x^*) \cdot d > 0, \forall d \in \mathbb{R}^n, d \neq 0$ ),

aleshores  $x^*$  és mínim local de  $f$  en  $S$  (i a més, estricte). □

Aquesta última proposició no la fem en general perquè per  $x^*$  a la frontera de  $S$  és difícil d'escriure i no té prou interès fer l'esforç. A l'Apartat 1.3 ho farem per al cas en què  $S$  ve donat per restriccions funcionals, que sí és interessant.

## 1.2. Convexitat

**Definicions 1.2.1.** Un conjunt  $S \subset \mathbb{R}^n$  és *convex* si

$$\forall x_1, x_2 \in S, \forall \lambda \in (0, 1), \lambda x_1 + (1 - \lambda)x_2 \in S.$$

Sigui  $S \subset \mathbb{R}^n$  convex. Aleshores:

- $f: S \rightarrow \mathbb{R}$  és una *funció convexa* si

$$\forall x_1, x_2 \in S, \forall \lambda \in (0, 1), f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

- $f: S \rightarrow \mathbb{R}$  és *estrictament convexa* si

$$\forall x_1, x_2 \in S, x_1 \neq x_2, \forall \lambda \in (0, 1), f(\lambda x_1 + (1 - \lambda)x_2) < \lambda f(x_1) + (1 - \lambda)f(x_2).$$

- $f: S \rightarrow \mathbb{R}$  és *còncava* o *estrictament còncava* si  $-f$  és convexa o estrictament convexa, respectivament. (Equivalentment, les mateixes definicions de convexa i estrictament convexa invertint les desigualtats.) □

Per a funcions de classe  $C^2$ , la convexitat té una caracterització útil en termes de les derivades segones: La matriu hessiana ha de ser semidefinida positiva, almenys en les direccions factibles. Per demostrar aquesta caracterització, necessitem primer establir-ne una altra, que val per a funcions de classe  $C^1$ , i té una interpretació gràfica fàcil: En dimensió 1, per exemple, ens diu que les rectes tangents a la gràfica estan sempre sota la gràfica.

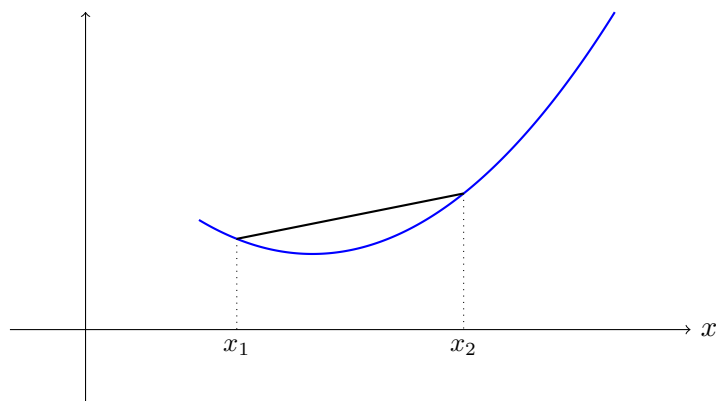


Figura 1.1: Una funció estrictament convexa en dimensió 1. Els segments que uneixen punts de la gràfica estan per sobre de la gràfica.

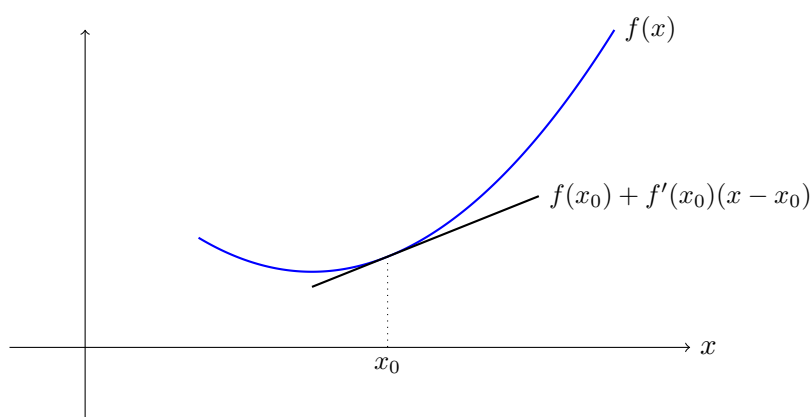


Figura 1.2: La mateixa funció convexa de la Figura 1.1. Les tangents estan per sota de la gràfica.

**Lema 1.2.2.** *Siguin  $S \subset \mathbb{R}^n$  convex i  $f: S \rightarrow \mathbb{R}$  de classe  $C^1$ . Aleshores:*

$$f \text{ és convexa en } S \Leftrightarrow \forall x, x_0 \in S, f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0) .$$

*Demostració.*

$\Rightarrow$ ) Per qualsevol  $\lambda \in (0, 1)$ ,

$$\begin{aligned} f(\lambda x + (1 - \lambda)x_0) &\leq \lambda f(x) + (1 - \lambda)f(x_0) , \Rightarrow \\ f(x_0 + \lambda(x - x_0)) &\leq f(x_0) + \lambda(f(x) - f(x_0)) , \Rightarrow \\ \frac{1}{\lambda}(f(x_0 + \lambda(x - x_0)) - f(x_0)) &\leq f(x) - f(x_0) . \end{aligned}$$

Fent  $\lambda \rightarrow 0$ , queda

$$\nabla f(x_0) \cdot (x - x_0) \leq f(x) - f(x_0) .$$

$\Leftarrow$ ) Fixem  $x_1, x_2 \in S$  i  $\lambda \in (0, 1)$ , i posem  $x := \lambda x_1 + (1 - \lambda)x_2$ . Tindrem:

$$\begin{aligned} f(x_1) &\geq f(x) + \nabla f(x) \cdot (x_1 - x) \\ f(x_2) &\geq f(x) + \nabla f(x) \cdot (x_2 - x) \end{aligned}$$

Multiplicant per  $\lambda$  la primera expressió, per  $(1 - \lambda)$  la segona, i sumant, obtenim

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2) + 0 .$$

□

**Proposició 1.2.3.** *Siguin  $S \subset \mathbb{R}^n$  convex, i  $f: S \rightarrow \mathbb{R}$  de classe  $C^2$ . Aleshores:*

$$f \text{ és convexa en } S \Leftrightarrow \forall x \in S, \forall d \text{ direcció factible en } x, d \cdot \nabla^2 f(x) \cdot d \geq 0 .$$

*Demostració.*

$\Leftarrow$ ) Donats  $x, x_0 \in S$ , existeix  $\lambda \in [0, 1]$  tal que (Teorema de Taylor)

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0) \cdot \nabla^2 f(\lambda x + (1 - \lambda)x_0) \cdot (x - x_0) .$$

L'últim terme és positiu, per la hipòtesi. Per tant,

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0) ,$$

que és equivalent a la convexitat, pel Lema 1.2.2.

$\Rightarrow$ ) Suposem que per certs punts  $x, x_0 \in S$  es tinguéssim

$$(x - x_0) \cdot \nabla^2 f(x_0) \cdot (x - x_0) < 0 .$$

Per la continuïtat de les derivades segones, podem seleccionar  $x$  prou proper a  $x_0$  de manera que

$$(x - x_0) \cdot \nabla f(\lambda x + (1 - \lambda)x_0) \cdot (x - x_0) < 0, \quad \forall \lambda \in [0, 1] .$$

Això vol dir, usant la mateixa fórmula de Taylor d'abans, que

$$f(x) < f(x_0) + \nabla f(x_0) \cdot (x - x_0) ,$$

i, pel Lema 1.2.2, la funció no seria convexa. □

**Teorema 1.2.4** (Local-Global per a funcions convexes).

*Sigui  $S \subset \mathbb{R}^n$  convex, i  $f: S \rightarrow \mathbb{R}$  convexa.*

*Si  $x^*$  és mínim local de  $f$  en  $S$ , llavors  $x^*$  és mínim global de  $f$  en  $S$ .*

*Demostració.* Si  $x^*$  no fos mínim global de  $f$  en  $S$ , existiria  $x \in S$  tal que  $f(x) < f(x^*)$ . Aleshores, per  $\forall \lambda \in (0, 1)$ ,

$$f(\lambda x + (1 - \lambda)x^*) \leq \lambda f(x) + (1 - \lambda)f(x^*) < f(x^*) ,$$

i això contradiu que  $x^*$  sigui mínim local, ja que  $\lambda$  pot ser arbitràriament proper a zero. □

**Teorema 1.2.5** (“les condicions necessàries són suficients”).

*Siguin  $S \subset \mathbb{R}^n$  convex, i  $f: S \rightarrow \mathbb{R}$  convexa i de classe  $C^1$ .*

*Si  $x^* \in S$  compleix  $\nabla f(x^*) \cdot d \geq 0$ , per a tota  $d$  direcció factible en  $x^*$ , llavors  $x^*$  és mínim (global!) de  $f$  en  $S$ .*

*En particular,*

$$\nabla f(x^*) = 0 \Rightarrow x^* \text{ és mínim global.}$$

*Demostració.* Per a qualsevol  $x \in S$ ,  $x - x^*$  és direcció factible gràcies a la convexitat de  $S$ . Per tant, usant el Lema 1.2.2,  $f(x) \geq f(x^*) + \nabla f(x^*) \cdot (x - x^*) \geq f(x^*)$ . □

**Exercici 1.2.6.** Considerem la funció  $f(x_1, x_2) = x_1^4 + x_2^4 + x_1 + x_2$ , definida sobre el conjunt

$$S = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, |x_2| \leq x_1\} .$$

- Comproveu que  $\nabla f(0, 0) \cdot d \geq 0$ , per a tota direcció factible  $d$  en  $(0, 0)$ .
- Comproveu que la funció  $f$  és convexa i que el domini  $S$  és convex.
- A partir dels apartats anteriors, què es pot afirmar del punt  $(0, 0)$  respecte la funció  $f$ ? □



### 1.3. Restriccions funcionals

Considerarem ara el cas en què la regió factible  $S$  ve donada per igualtats i desigualtats funcionals:

$$S := \text{punts } x \text{ de } \mathbb{R}^n \text{ tals que } \left\{ \begin{array}{c} h_1(x) = 0 \\ \vdots \\ h_m(x) = 0 \end{array} \right\} \quad \text{i} \quad \left\{ \begin{array}{c} g_1(x) \leq 0 \\ \vdots \\ g_r(x) \leq 0 \end{array} \right\}$$

on  $h_i: \mathbb{R}^n \rightarrow \mathbb{R}$ , i  $g_i: \mathbb{R}^n \rightarrow \mathbb{R}$ .

Com que el problema té ara més estructura, és d'esperar que puguem afinar la teoria de l'Apartat 1.1.

Si totes les funcions involucrades són afins, es tracta d'un *problema de Programació Lineal*, per al qual tenim una bona teoria, basada en àlgebra lineal, i un bon mètode general de resolució. Si no, diem que tenim un problema de *Programació No Lineal* (o *Optimització No Lineal*).

La funció  $f$  rep el nom de *funció objectiu*, i les igualtats i desigualtats que defineixen  $S$  són les *restriccions* del problema. Des del punt de vista de les aplicacions, interessa trobar tant el valor del mínim com un punt  $x \in S$  on s'assoleixi. A la pràctica, naturalment, una bona aproximació és suficient.

Si es tracta de trobar el màxim de  $f$  en  $S$ , el més senzill és canviar el signe de la funció i pensar sempre en minimitzar, usant que

$$\max_{x \in S} f(x) = -(\min_{x \in S} -f)$$

i que ambdós s'assoleixen en el mateix o mateixos punts.

**Definicions 1.3.1.** Donat un punt factible  $x^*$ , una restricció  $g_i(x) \leq 0$  es diu que és *activa* en  $x^*$  si  $g_i(x^*) = 0$ . És *inactiva* si  $g_i(x^*) < 0$ .

Les restriccions  $h_i(x) = 0$  són sempre *actives* en tot punt factible. □

Una funció  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$  es pot pensar com un vector de funcions reals  $h(x) := (h_1(x), \dots, h_m(x))$ , i podem escriure l'equació vectorial  $h(x) = 0$  per representar el sistema de  $m$  equacions escalars

$$\left. \begin{array}{c} h_1(x) = 0 \\ \vdots \\ h_m(x) = 0 \end{array} \right\} .$$

Anàlogament, podem escriure  $g(x) \leq 0$  per denotar el sistema d'inequacions

$$\left. \begin{array}{c} g_1(x) \leq 0 \\ \vdots \\ g_r(x) \leq 0 \end{array} \right\} .$$

**Definició 1.3.2.** Un vector  $v$  és *tangent* al conjunt  $S \subset \mathbb{R}^n$  en un punt  $x^* \in S$  si:

- a) Hi ha una successió de punts  $\{x_k\}_k \subset S$ , tal que  $\lim_{k \rightarrow \infty} x_k = x^*$ .
- b) Hi ha una successió  $\{t_k\}_k \subset \mathbb{R}$ , amb  $t_k > 0$ , decreixent amb límit zero, tal que

$$\lim_{k \rightarrow \infty} \frac{x_k - x^*}{t_k} = v .$$

□

És important visualitzar el significat d'aquesta definició. Imaginem que estem en el punt  $x^*$  i veiem acostar-se cap a nosaltres els punts  $x_k$ . La direcció des de la qual ens colpejaren és la direcció  $v$ .

La successió  $\{t_k\}_k$  només serveix per normalitzar: Si el vector  $v$  és unitari, podem prendre  $t_k = \|x_k - x^*\|$ ; i en general, si té norma  $\|v\|$ , podem prendre  $t_k = \|x_k - x^*\|/\|v\|$ .<sup>2</sup>

Queda clar que si  $v$  és un vector tangent, aleshores per a tot  $\lambda > 0$ , el vector  $\lambda \cdot v$  també és vector tangent: només cal escalar convenientment la successió  $\{t_k\}_k$ . També està clar que el vector 0 sempre és un vector tangent, agafant  $x_k \equiv x^*$ .

Una definició equivalent de vector tangent és: Un vector  $v$  és tangent a  $S$  en el punt  $x^* \in S$  si, o bé  $v = 0$ , o bé existeix una successió  $\{x_k\}_k \subset S$ , amb  $x_k \neq x^*$  per a tot  $k$ , tal que

$$\lim_{k \rightarrow \infty} x_k = x^* \quad \text{i} \quad \lim_{k \rightarrow \infty} \frac{x_k - x^*}{\|x_k - x^*\|} = \frac{v}{\|v\|}.$$

La idea intuïtiva de vector tangent és la que tenim tots de geometria elemental (recta tangent a una corba, pla tangent a una superfície, etc).

[Vegeu la Figura 11.2 de Luenberger, pàg 324: La figura (b) és el cas en què els vectors tangents a una superfície suau de  $\mathbb{R}^3$  formen un espai vectorial de dimensió 2, el *pla tangent*. La figura (a) és una corba en  $\mathbb{R}^2$  o  $\mathbb{R}^3$ , i els vectors tangents formen un espai vectorial de dimensió 1, la *recta tangent*. A la figura (c) veiem una corba en  $\mathbb{R}^3$ , determinada per la intersecció de dues superfícies, i els vectors tangents també formen un espai vectorial de dimensió 1.]

Però no és cert en general que els vectors tangents en un punt formen un espai vectorial. A la Figura 1.3 veiem dos casos. En el cas (a) tenim el conjunt  $S = \{(x_1, x_2) \in \mathbb{R}^2 : (x_1 + 1)^2 - x_2 \leq 0, (x_1 - 1)^2 - x_2 \leq 0\}$ , delimitat per dues paràboles. En el punt  $x^* = (0, 1)$ , els vectors tangents són els que apunten cap a la regió ombrejada  $T$ , inclosa la vora. En el cas (b), el conjunt  $S$  és la unió de les dues paràboles anteriors, o sigui  $S = \{(x_1, x_2) \in \mathbb{R}^2 : ((x_1 + 1)^2 - x_2)((x_1 - 1)^2 - x_2) = 0\}$ . En el punt  $(0, 1)$ , només hi ha quatre direccions de vectors tangents.

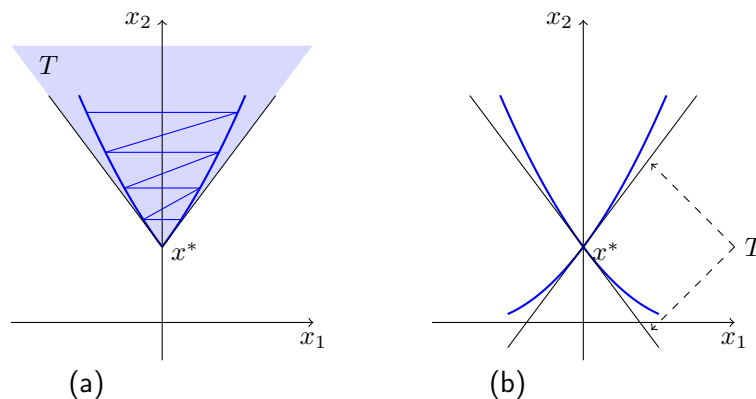


Figura 1.3: Dos casos en què els vectors tangents no formen un espai vectorial

**Definició 1.3.3.** El conjunt format pels vectors tangents en un punt s'anomena el *con tangent*.  $\square$

Es pot comprovar que és efectivament un con, segons la definició següent: Un subconjunt  $T \subset \mathbb{R}^n$  és un *con* si per a tot  $v \in T$  i per a tot  $\lambda > 0$ , el vector  $\lambda \cdot v$  també pertany al con.

**Exercici 1.3.4.** Tota direcció factible de  $S$  en un punt és un vector tangent en aquell punt. A la inversa no és cert.  $\square$

<sup>2</sup>Per ser rigorosos, caldria afegir aquí que si  $x_k = x^*$ , aleshores  $t_k$  pot ser qualsevol real positiu.

A partir d'ara suposarem que totes les funcions tenen totes les derivades contínues que calguin, sense explicitar-ho més.

**Definició 1.3.5.** Sigui  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

Un punt  $x^*$ , solució de l'equació  $h(x) = 0$ , és un *punt regular de l'equació* si  $\nabla h_1(x^*), \dots, \nabla h_m(x^*)$  són vectors linealment independents.  $\square$

El concepte de punt regular no és intrínsec del conjunt determinat per  $h(x) = 0$ , sinó de la seva representació concreta en termes de la funció  $h$ . Per exemple, en  $\mathbb{R}^2$ , les funcions  $h(x_1, x_2) = x_1$  i  $h(x_1, x_2) = x_1^2$  determinen la mateixa recta  $x_1 \equiv 0$ ; tots els punts de la recta són regulars per la primera representació, però el  $(0, 0)$  no ho és per la segona.

**Notació 1.3.6.** Si  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ , denotarem per  $\nabla h(x)$  la *matriu jacobiana* de  $h$  en el punt  $x$ , és a dir,

$$\nabla h(x) := \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x) & \dots & \frac{\partial h_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1}(x) & \dots & \frac{\partial h_m}{\partial x_n}(x) \end{pmatrix}$$

Observeu que no hi ha conflicte amb la notació per al gradient que hem introduït a 1.1.6. El gradient no és més que la matriu jacobiana d'una funció  $h: \mathbb{R}^n \rightarrow \mathbb{R}$ .  $\square$

La següent proposició diu que el con tangent a un punt regular de  $h(x) = 0$  és un espai vectorial i l'explicita de manera molt útil.

**Proposició 1.3.7.** Si  $x^*$  és un punt regular de l'equació  $h(x) = 0$ , el con tangent al conjunt  $\{x: h(x) = 0\}$  en  $x^*$  és l'espai vectorial format pels vectors ortogonals als gradients de  $h$  en  $x^*$ ; o sigui, és

$$\{v \in \mathbb{R}^n : \nabla h(x^*) \cdot v = 0\} .$$

*Demostració.* Un vector tangent  $v$  sempre és ortogonal als vectors  $\nabla h_i$ , tant si el punt és regular com si no. Veiem-ho: Considerem les successions  $\{x_k\}_k$  i  $\{t_k\}_k$  tals que

$$\lim_{k \rightarrow \infty} \frac{x_k - x^*}{t_k} = v .$$

Això es pot expressar també

$$x_k = x^* + t_k v + o(t_k) .$$

Desenvolupant per Taylor  $h_i$  fins a primer ordre  $h_i$  en el punt  $x^*$ , tenim

$$\begin{aligned} 0 &= \frac{h_i(x_k)}{t_k} = \frac{1}{t_k} \left[ h_i(x^*) + \nabla h_i(x^*)(t_k v + o(t_k)) + o(t_k) \right] \\ &= \frac{1}{t_k} \left[ t_k \nabla h_i(x^*) v + o(t_k) \right] \\ &= \nabla h_i(x^*) \cdot v + \frac{o(t_k)}{t_k} . \end{aligned}$$

Prenent límit, obtenim  $\nabla h_i(x^*) \cdot v = 0$ .

Veiem ara que si  $x^*$  és regular, un vector  $v$  ortogonal a tots els gradients  $\nabla h_i(x^*)$  és necessàriament un vector tangent. La hipòtesi implica que la matriu jacobiana  $\nabla h(x^*)$ , de dimensions  $m \times n$ , té rang màxim  $m$ . Sigui  $Z$  una matriu  $n \times (n - m)$  que tingui per columnes una base de  $\text{Ker } \nabla h(x^*)$ .

Considerem, per cada  $t \in \mathbb{R}$ , el sistema lineal d'equacions  $n \times n$  amb incògnites  $z \in \mathbb{R}^n$ ,

$$R(z, t) := \begin{pmatrix} h(z) - t\nabla h(x^*)v \\ Z^T(z - x^* - td) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Per  $t = 0$ , la única solució és  $z = x^*$ , donat que  $Z^T$  té rang màxim. La matriu jacobiana de  $R$  en aquest punt respecte les variables  $z$  és la matriu  $n \times n$

$$\nabla_z R(x^*, 0) = \begin{pmatrix} \nabla h(x^*) \\ Z^T \end{pmatrix}$$

que és no-singular.

Pel Teorema de la Funció Implícita, podem aïllar  $z$  en funció de  $t$  en un entorn de  $t = 0$ . Per tant, hi ha una solució  $z_k$  del sistema per valors de  $t_k > 0$  prou petits i formant una successió que tendeix a zero.

A més,

$$h(z_k) = t_k \nabla h(x^*) \cdot v = 0.$$

Per tant, els punts  $z_k$  estan en el conjunt determinat per  $h(x) = 0$ . Veurem que amb aquestes successions ens serveixen per demostrar que  $v$  és un vector tangent:

Aplicant el Teorema de Taylor a  $h(z_k)$ ,

$$\begin{aligned} 0 = R(z_k, t_k) &= \begin{pmatrix} h(z_k) - t_k \nabla h(x^*) \cdot v \\ Z^T(z_k - x^* - t_k v) \end{pmatrix} \\ &= \begin{pmatrix} \nabla h(x^*)(z_k - x^*) + o(\|z_k - x^*\|) - t_k \nabla h(x^*) \cdot v \\ Z^T(z_k - x^* - t_k v) \end{pmatrix} \\ &= \begin{pmatrix} \nabla h(x^*) \\ Z^T \end{pmatrix} (z_k - x^* - t_k v) + o(\|z_k - x^*\|) \end{aligned}$$

Dividint per  $t_k$ ,

$$0 = \begin{pmatrix} \nabla h(x^*) \\ Z^T \end{pmatrix} \left( \frac{z_k - x^*}{t_k} - v \right) + o\left(\frac{\|z_k - x^*\|}{t_k}\right),$$

d'on, invertint la matriu del davant,

$$\frac{z_k - x^*}{t_k} = v + o\left(\frac{\|z_k - x^*\|}{t_k}\right)$$

i es dedueix que el terme de l'esquerra tendeix a  $v$ , com volíem demostrar.  $\square$

**Teorema 1.3.8** (Condicions necessàries de mínim local). *Sigui*

$$S : \left\{ \begin{array}{ll} h_1(x) = 0 & g_1(x) \leq 0 \\ \vdots & \vdots \\ h_m(x) = 0 & g_r(x) \leq 0 \end{array} \right\}$$

*Si  $x^*$  és un mínim local de  $f$  en  $S$  i és punt regular de les restriccions actives en  $x^*$ , llavors existeixen  $\lambda = (\lambda_1, \dots, \lambda_m)$  i  $\mu = (\mu_1, \dots, \mu_r) \geq 0$  tals que*

1)

$$\left. \begin{array}{l} \nabla f(x^*) + \lambda \cdot \nabla h(x^*) + \mu \cdot \nabla g(x^*) = 0 \\ \forall i = 1, \dots, r, \mu_i g_i(x^*) = 0 \end{array} \right\}.$$

2) La matriu

$$M(x^*) := \nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i \cdot \nabla^2 h_i(x^*) + \sum_{i=1}^r \mu_i \cdot \nabla^2 g_i(x^*)$$

compleix  $v \cdot M(x^*) \cdot v \geq 0$ , per a tot  $v$  de l'espai tangent a les restriccions actives en  $x^*$ .

Les constants  $\lambda_1, \dots, \lambda_m$  s'anomenen *multiplicadors de Lagrange*, i les constants  $\mu_1, \dots, \mu_r$  s'anomenen *multiplicadors de Karush-Kuhn-Tucker*. La condició 1) és una *condició de 1r ordre*, perquè només involucra derivades primeres; la 2) és de *2n ordre*, perquè hi intervenen derivades segones. Les condicions de primer ordre es coneixen habitualment com les *condicions de Karush-Kuhn-Tucker* (KKT, per abreviar).

Per a la demostració, distingirem dos casos:

**1r cas:** No hi ha restriccions de tipus  $\leq$ . Llavors:

- La condició de primer ordre queda  $\nabla f(x^*) + \lambda \nabla h(x^*) = 0$ .
- La matriu  $M(x^*)$  és  $\nabla^2 f(x^*) + \sum_{i=1}^m \lambda_i \nabla^2 h_i(x^*)$ .

**2n cas:** Cas general.

*Demostració.*

*1r cas; condició de primer ordre:* Podem suposar que  $m < n$ ; en cas contrari, qualsevol vector es pot expressar com a combinació lineal dels vectors linealment independents  $\nabla h_1(x^*), \dots, \nabla h_m(x^*)$  i ja estaríem.

La matriu jacobiana  $\nabla h(x^*)$ , que és  $m \times n$ , tindrà rang màxim  $m$  i suposarem que les primeres  $m$  columnes formen la submatriu invertible que segur que existeix.

La variable  $x \in \mathbb{R}^n$  podem dividir-la en dues parts,  $x = (x_B, x_N)$ , on  $B$  representa els subíndexs  $1, \dots, m$  i  $N$  representa els subíndexs  $m+1, \dots, n$ . Denotarem  $\nabla_B$  la derivació respecte les variables de  $x_B$ , i  $\nabla_N$  la derivació respecte les variables de  $x_N$ . Tenim doncs que la matriu  $m \times m$

$$\nabla_B h(x^*) := \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x^*) & \dots & \frac{\partial h_1}{\partial x_m}(x^*) \\ \vdots & & \vdots \\ \frac{\partial h_m}{\partial x_1}(x^*) & \dots & \frac{\partial h_m}{\partial x_m}(x^*) \end{pmatrix}$$

és invertible.

Pel Teorema de la Funció Implícita, l'equació  $h(x_B, x_N) = 0$  permet aïllar  $x_B$  en funció de  $x_N$ , donant lloc a una funció  $\phi: V \subset \mathbb{R}^{n-m} \rightarrow \mathbb{R}^m$ , tant regular com  $h$ , en un entorn  $V$  de  $x_N^*$ , i

$$\nabla \phi(x_N) = (-\nabla_B h(\phi(x_N), x_N))^{-1} \cdot \nabla_N h(\phi(x_N), x_N).$$

Considerem la funció  $F: V \subset \mathbb{R}^{n-m} \rightarrow \mathbb{R}$ , definida així:

$$F(x_N) := f(\phi(x_N), x_N).$$

Està clar que  $x_N^*$  és un mínim local de  $F$  sense restriccions (hem "eliminat" les restriccions al usar-les per aïllar unes variables en funció de les altres). Per tant, estem en la situació del Corol·lari 1.1.9 (si no hi ha restriccions, tot punt és interior al conjunt factible), i el gradient de  $F$  ha de ser zero en  $x_N^*$ . Per la regla de la cadena,

$$0 = \nabla F(x_N^*) = \nabla_B f(x^*) \cdot (-\nabla_B h(x^*))^{-1} \cdot \nabla_N h(x^*) + \nabla_N f(x^*). \quad (1.1)$$

Definint el vector  $\lambda \in \mathbb{R}^m$  com

$$\lambda := \nabla_B f(x^*) \cdot (-\nabla_B h(x^*))^{-1} \quad (1.2)$$

podem escriure (1.1) com

$$0 = \lambda \cdot \nabla_N h(x^*) + \nabla_N f(x^*) , \quad (1.3)$$

i (1.2) com

$$\lambda \cdot \nabla_B h(x^*) + \nabla_B f(x^*) = 0 . \quad (1.4)$$

Combinant (1.3) i (1.4), obtenim

$$\lambda \cdot \nabla h(x^*) + \nabla f(x^*) = 0 . \quad (1.5)$$

*1r cas; condició de segon ordre:* Volem aplicar la condició necessària de segon ordre sense restriccions a la funció  $F$  definida abans, o sigui que la matriu Hessiana de  $F$  en  $x_N^*$  sigui semidefinida positiva:

$$y \cdot \nabla^2 F(x_N^*) \cdot y \geq 0, \quad \forall y \in \mathbb{R}^{n-m} .$$

Per trobar la matriu Hessiana de  $F$  cal tornar a fer derivades parcials en (1.1). És un càlcul molt pesat, i dóna, posant per abreviar  $B := \nabla_B h(x^*)$ ,  $N := \nabla_N h(x^*)$ ,

$$\begin{aligned} 0 \leq y^\top \cdot \nabla^2 F(x_N^*) \cdot y &= \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix}^\top \cdot \nabla^2 f(x^*) \cdot \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix} \\ &+ y^\top \cdot \left[ \sum_{j=1}^m \nabla^2 \phi_j(x_N^*) \frac{\partial f}{\partial x_j}(x^*) \right] \cdot y \end{aligned} \quad (1.6)$$

Anàlogament, si definim  $H_i: \mathbb{R}^{n-m} \rightarrow \mathbb{R}$  com  $H_i(x_N) = h_i(\phi(x_N), x_N)$ , per  $i = 1, \dots, m$ , obtenim, derivant dues vegades l'equació  $H_i(x_N) = 0$ ,

$$\begin{aligned} 0 = y^\top \cdot \nabla^2 H_i(x_N^*) \cdot y &= \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix}^\top \cdot \nabla^2 h_i(x^*) \cdot \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix} \\ &+ y^\top \cdot \left[ \sum_{j=1}^m \nabla^2 \phi_j(x_N^*) \frac{\partial h_i}{\partial x_j}(x^*) \right] \cdot y \end{aligned} \quad (1.7)$$

Multiplicant (1.7) per la  $\lambda_i$  de la condició de primer ordre, sumant per  $i = 1, \dots, m$ , i sumant també la desigualtat (1.6), obtenim

$$\begin{aligned} 0 \leq \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix}^\top \cdot \left[ \nabla^2 f(x^*) + \sum_{i=1}^m \nabla^2 h_i(x^*) \right] \cdot \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix} \\ + y^\top \cdot \left[ \sum_{j=1}^m \nabla^2 \phi_j(x_N^*) \left( \frac{\partial f}{\partial x_j}(x^*) + \sum_{i=1}^m \lambda_i \frac{\partial h_i}{\partial x_j}(x^*) \right) \right] \cdot y \end{aligned}$$

Com que, per la condició de primer ordre,

$$\frac{\partial f}{\partial x_j}(x^*) + \sum_{i=1}^m \lambda_i \frac{\partial h_i}{\partial x_j}(x^*) = 0, \quad j = 1, \dots, m ,$$

ens queda

$$0 \leq \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix}^\top \cdot [M(x^*)] \cdot \begin{pmatrix} -B^{-1}Ny \\ y \end{pmatrix} \quad (1.8)$$



**Exemple 1.3.10.** Per estudiar un teorema és tan important l'enunciat precís com els contra-exemples quan una hipòtesi falla. Què pot fallar per a què no existeixin els multiplicadors de Lagrange? Considereu el problema

$$\begin{aligned} \min \quad & f(x_1, x_2) = (x_1 - 1)^2 + x_2^2 \\ \text{subjecte a: } & \{ x_1^2 = 0 \end{aligned}$$

Fent un dibuix amb les corbes de nivell de  $f$ , es veu clar que hi ha un mínim local en  $(0, 0)$ . Però plantejant les condicions necessàries, obtenim

$$\left. \begin{aligned} 2(x_1 - 1) + 2\lambda x_1 &= 0 \\ 2x_2 &= 0 \end{aligned} \right\},$$

que conjuntament amb la restricció ens dóna un sistema incompatible. El teorema no s'aplica perquè el punt  $(0, 0)$  no és un punt regular de la restricció.

**Exercici 1.3.11.** Considereu el problema en  $\mathbb{R}^2$

$$\begin{aligned} \min \quad & f(x) = x_1 + x_2 \\ \text{subjecte a: } & \left\{ \begin{aligned} (x_1 - 1)^2 + x_2^2 - 1 &= 0 \\ (x_1 - 2)^2 + x_2^2 - 4 &= 0 \end{aligned} \right. \end{aligned}$$

Feu un dibuix que representi el problema. Veureu que en realitat només hi ha un punt factible, que per tant és el mínim de la funció automàticament. Calculeu i dibuixeu els vectors gradients de les restriccions en el mínim i observeu que no són linealment independents; comproveu que efectivament no existeixen els multiplicadors de Lagrange, intentant resoldre el sistema KKT, i també visualment dibuixant el gradient de  $f$ .

Comproveu que la condició de punt regular, si bé és suficient, no és necessària per a l'existència dels multiplicadors: Si  $f(x) = x_1$ , sí que existeixen (el gradient de  $f$  apunta "casualment" en la direcció correcta).  $\square$

**Exemple 1.3.12.**

[Estudieu l'Exemple 12.2, pàgina 310 de Nocedal-Wright. Aquest és el segon exemple fonamental a entendre a fons, després de l'exemple 12.1. Aneu amb compte amb la notació, lleugerament diferent de la nostra: Els multiplicadors KKT aquí es denoten per  $\lambda$ ; es posen a l'altra banda del signe igual en la condició necessària de primer ordre, i les desigualtats s'escriuen  $c(x) \geq 0$  en lloc de  $g(x) \leq 0$ . Recomano que us aneu traduint l'exemple, inclosos els dibuixos, a les nostres notacions.

El que és més interessant aquí és observar que el signe del multiplicador és molt important. Intenteu entendre la resta de comentaris fins a la pàgina 313.]  $\square$

**Teorema 1.3.13** (Condicions suficients de mínim local).

*Siguin  $S$  i  $f$  com en el Teorema 1.3.8.*

*Donat un punt  $x \in S$ , si existeixen  $\lambda = (\lambda_1, \dots, \lambda_m)$  i  $\mu = (\mu_1, \dots, \mu_r) \geq 0$  tals que*

*a)*

$$\left. \begin{aligned} \nabla f(x) + \lambda \cdot \nabla h(x) + \mu \cdot \nabla g(x) &= 0 \\ \forall i = 1, \dots, r, \mu_i g_i(x) &= 0 \end{aligned} \right\},$$

*b) La matriu  $M(x)$  del Teorema 1.3.8 compleix  $v \cdot M(x) \cdot v > 0$ , per a tot  $v \neq 0$  del subespai*

$$\{v : \nabla h(x) \cdot v = 0; \nabla g_i(x) \cdot v = 0, \forall i \text{ tal que } g_i(x) = 0 \text{ i } \mu_i > 0\}, \quad (1.9)$$

*aleshores  $x$  és mínim local (estricte) de  $f$  en  $S$ .*  $\square$



Observeu que la condició  $b)$  no és només que  $M(x)$  sigui definida positiva sobre l'espai tangent de les restriccions actives en  $x$ . Si traiem el  $\mu_i > 0$ , el subespai (1.9) seria més petit, i l'exigència sobre  $M(x)$  seria més fluixa. La demostració de la Proposició 1.4.1 conté un exemple de la importància d'aquest detall.

**Exemple 1.3.14.** Vegem un exemple complet en què la teoria que hem vist és aplicable de manera directa:

$$\begin{aligned} \min \quad & f(x) = 2x_1^2 + 2x_1x_2 + x_2^2 - 10x_1 - 10x_2 \\ \text{subjecte a:} \quad & \begin{cases} x_1^2 + x_2^2 \leq 5 \\ 3x_1 + x_2 \leq 6 \end{cases} \end{aligned}$$

Imposem les condicions necessàries de 1r ordre, per trobar candidats a mínim local.

$$\begin{aligned} \nabla f(x) &= (4x_1 + 2x_2 - 10, 2x_1 + 2x_2 - 10) \\ \nabla g(x) &= \begin{pmatrix} 2x_1 & 2x_2 \\ 3 & 1 \end{pmatrix}. \end{aligned}$$

$$\left. \begin{aligned} 4x_1 + 2x_2 - 10 + \mu_1 \cdot 2x_1 + \mu_2 \cdot 3 &= 0 \\ 2x_1 + 2x_2 - 10 + \mu_1 \cdot 2x_2 + \mu_2 &= 0 \\ \mu_1 \cdot (x_1^2 + x_2^2 - 5) &= 0 \\ \mu_2 \cdot (3x_1 + x_2 - 6) &= 0 \\ \mu_1 \geq 0, \mu_2 \geq 0 \end{aligned} \right\}$$

- Si  $g_1(x) \leq 0$  i  $g_2(x) \leq 0$  són inactives, llavors  $\mu_1 = 0$  i  $\mu_2 = 0$  i queda

$$\left. \begin{aligned} 4x_1 + 2x_2 - 10 &= 0 \\ 2x_1 + 2x_2 - 10 &= 0 \end{aligned} \right\} \rightarrow x_1 = 0, x_2 = 5$$

Satisfà  $(0, 5)$  les restriccions? No. Per tant, aquest punt no val.

(Observeu que aquest cas correspon a resoldre el problema inicial sense restriccions.)

- Si  $g_1(x) \leq 0$  és activa i  $g_2(x) \leq 0$  és inactiva, llavors  $\mu_2 = 0$ , i queda

$$\left. \begin{aligned} 4x_1 + 2x_2 - 10 + 2\mu_1x_1 &= 0 \\ 2x_1 + 2x_2 - 10 + 2\mu_1x_2 &= 0 \\ x_1^2 + x_2^2 - 5 &= 0 \end{aligned} \right\}.$$

La resolució d'aquest sistema no és fàcil, però encara es pot fer raonablement. Aïllem  $\mu_1$  de la primera equació:

$$\mu_1 = \frac{-4x_1 - 2x_2 + 10}{2x_1}.$$

(Podem suposar que  $x_1 \neq 0$ , perquè si no es veu ràpidament que queda un sistema incompatible.) Substituint a la segona equació,

$$\begin{aligned} 2x_1 + 2x_2 - 10 + \frac{-4x_1 - 2x_2 + 10}{x_1}x_2 &= 0, \Rightarrow \\ 2x_1^2 - 2x_1x_2 - 2x_2^2 - 10x_1 + 10x_2 &= 0. \end{aligned}$$

De

$$x_1^2 + x_2^2 - 5 = 0$$

obtenim

$$x_2 = \pm \sqrt{5 - x_1^2},$$

que substituït a l'altre equació i simplificant (això involucra elevar al quadrat l'equació) dona

$$5x_1^4 - 30x_1^3 + 25x_1^2 + 100x_1 - 100 = 0.$$

Aquest polinomi resulta tenir dues arrels reals:

$$x_1 = \begin{cases} 1 \\ \frac{1}{3} [5 - (145 + 30\sqrt{6})^{1/3} - (145 - 30\sqrt{6})^{1/3}] \end{cases}.$$

Per la primera,

$$x_2 = \begin{cases} 2 \\ -2 \end{cases}$$

Per la segona,

$$x_2 = \begin{cases} \sqrt{5 - x_1^2} \\ -\sqrt{5 - x_1^2} \end{cases}$$

Per al punt  $(1, 2)$ , s'obté  $\mu_1 = \frac{-4 - 4 + 10}{2} = 1 > 0$ .

Per al punt  $(1, -2)$  s'obté  $\mu_1 = \frac{-4 + 4 + 10}{2} = 5 > 0$ .

Per a les altres dues solucions,  $\mu_1 \approx -4.07395$ ,  $\mu_1 \approx -8.62383$ , que no són admissibles.

Però  $x_1 = 1$ ,  $x_2 = -2$ ,  $\mu_1 = 5$  no és solució del sistema. (és una solució espúria afegida al elevar al quadrat per obtenir l'equació de 4t grau.) Només ens queda un únic candidat:  $x_1 = 1$ ,  $x_2 = 2$ ,  $\mu_1 = 1$ , que sí és solució del sistema.

Investiguem les condicions suficients:

$$\begin{aligned} \nabla^2 f(x) &= \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}, & \nabla^2 g_1(x) &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, & \nabla^2 g_2(x) &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \\ M(x) &= \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix} + 1 \cdot \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}. \end{aligned}$$

$M$  és definida positiva en tot l'espai. Per tant,  $(1, 2)$  és mínim local estricte del problema. De fet, per arribar a aquesta conclusió seria suficient que  $M$  fos definida positiva sobre

$$\{v : \nabla g_1(1, 2) \cdot v = 0\} = \{v : (2, 4) \cdot v = 0\} = \{v : 2v_1 + 4v_2 = 0\}.$$

En efecte,

$$(v_1, -\frac{1}{2}v_1) \cdot \begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ -\frac{1}{2}v_1 \end{pmatrix} = 5v_1^2 > 0,$$

si  $(v_1, v_2) \neq (0, 0)$ .

Observeu, a més, que:

- $f$  és convexa (estricta), perquè  $\nabla^2 f(x) \equiv \begin{pmatrix} 4 & 2 \\ 2 & 2 \end{pmatrix}$  és definida positiva.
- El conjunt factible és convex:

$$\left. \begin{aligned} x_1^2 + x_2^2 \leq 5 &\leftarrow \text{cercle (convex)} \\ 3x_1 + x_2 \leq 6 &\leftarrow \text{semiespai (convex)} \end{aligned} \right\} \text{Intersecció de convexos és convex.}$$

Pel Teorema local-global 1.2.4,  $(1, 2)$  és mínim global estricte del problema. Ja no cal continuar. Si haguéssim de continuar, faríem ara els casos:

- Si  $g_1(x) \leq 0$  és inactiva i  $g_2(x) \leq 0$  és activa, llavors  $\mu_1 = 0$  i queda

$$\left. \begin{aligned} 4x_1 + 2x_2 - 10 + 3\mu_2 &= 0 \\ 2x_1 + 2x_2 - 10 + \mu_2 &= 0 \\ 3x_1 + x_2 - 6 &= 0 \end{aligned} \right\}.$$

Dóna una solució no admissible:  $x_1 = \frac{6}{15}$ ,  $x_2 = \frac{72}{15}$ ,  $\mu_2 = \frac{-6}{15}$ .

- Si totes dues són actives, hauríem de resoldre

$$\left. \begin{aligned} 4x_1 + 2x_2 - 10 + 2\mu_1 x_1 + 3\mu_2 &= 0 \\ 2x_1 + 2x_2 - 10 + 2\mu_1 x_2 + \mu_2 &= 0 \\ x_1^2 + x_2^2 - 5 &= 0 \\ 3x_1 + x_2 - 6 &= 0 \end{aligned} \right\}.$$

□

**Exercici 1.3.15.** Estudieu el problema amb dues restriccions de desigualtat

$$\begin{aligned} \min \quad & f(x) = x_1 + x_2 \\ \text{subjecte a:} \quad & \begin{cases} x_1^2 + x_2^2 - 2 \leq 0 \\ -x_2 \leq 0 \end{cases} \end{aligned}$$

resolent les condicions KKT.

[Comproveu després el resultat amb el Nocedal-Wright, Exemple 12.3, pàgina 313. Si podeu, llegiu tots els comentaris de l'exemple.]

□

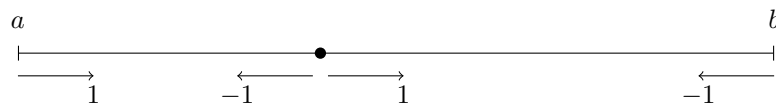
## 1.4. Extrems en $\mathbb{R}$

Veiem què ens diu la teoria vista en el cas de funcions de una variable:

$$\min_{x \in S} f(x)$$

on  $S = [a, b]$  és un interval tancat i acotat de  $\mathbb{R}$ .

Direccions factibles:



En qualsevol punt de l'interval obert  $(a, b)$ , hi ha dues direccions factibles, que corresponen als vectors unidimensionals  $(+1)$  (anar cap a la dreta) i  $(-1)$  (anar cap a l'esquerra). En el punt  $a$  només és factible la direcció  $(+1)$  (cap a la dreta), i en el  $b$  la direcció  $(-1)$  (cap a l'esquerra).

Condicions necessàries:

Es dedueixen de la teoria general de l'Apartat 1.1 (concretament, dels enunciats 1.1.7, 1.1.9) els fets següents:

- $x \in (a, b)$  mínim local  $\Rightarrow f'(x) = 0$  i  $f''(x) \geq 0$ .
- $a$  mínim local  $\Rightarrow f'(a) \geq 0$  i  $[f'(a) = 0 \Rightarrow f''(a) \geq 0]$ .
- $b$  mínim local  $\Rightarrow f'(b) \leq 0$  i  $[f'(b) = 0 \Rightarrow f''(b) \geq 0]$ .

Condicions suficients:

Es dedueix de la Proposició 1.1.11 que:

-  $x \in (a, b)$ ,  $f'(x) = 0$  i  $f''(x) > 0 \Rightarrow x$  mínim local.

Es dedueixen de les condicions suficients de l'Apartat 1.3 (concretament, del Teorema 1.3.13):

-  $f'(a) > 0$  o  $[f'(a) = 0$  i  $f''(a) > 0] \Rightarrow a$  mínim local.

-  $f'(b) < 0$  o  $[f'(b) = 0$  i  $f''(b) > 0] \Rightarrow b$  mínim local.

Veiem a la Figura 1.4 diverses d'aquestes situacions.

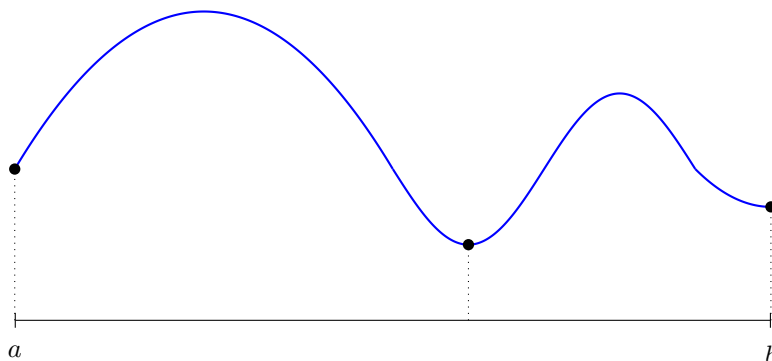


Figura 1.4: En aquesta situació, tant els extrems de l'interval com el punt marcat a l'interior són mínims locals. Observem que en  $a$  la derivada no és zero. En el punt del interior i en  $b$  s'aprecia que la derivada és zero. Del dibuix no es pot deduir si la derivada segona és més gran estrictament que zero, però sí que és més gran o igual que zero.

Anem a detallar la deducció d'aquests dos últims fets. Ho farem pel primer; el segon és anàleg.

**Proposició 1.4.1.** *Sigui  $f: [a, b] \rightarrow \mathbb{R}$  de classe  $C^2$ .*

*Si  $f'(a) > 0$  o bé  $[f'(a) = 0$  i  $f''(a) > 0]$ , aleshores  $a$  és mínim local.*

*Demostració.* El problema

$$\min_{x \in [a, b]} f(x)$$

es pot formular com

$$\begin{aligned} & \min f(x) \\ \text{subjecte a: } & \begin{cases} g_1(x) := x - b \leq 0 \\ g_2(x) := -x + a \leq 0 \end{cases} \end{aligned}$$

Segons el Teorema de condicions suficients de mínim local (Teorema 1.3.13), si existeixen  $\mu_1 \geq 0$ ,  $\mu_2 \geq 0$  tals que

a)

$$\left. \begin{aligned} f'(a) + \mu_1 g_1'(a) + \mu_2 g_2'(a) &= 0 \\ \mu_1 g_1(a) &= 0 \\ \mu_2 g_2(a) &= 0 \end{aligned} \right\}$$

b)  $M(a) := f''(a) + \mu_1 g_1''(a) + \mu_2 g_2''(a)$  compleix  $v \cdot M(a) \cdot v > 0$ , per a tot  $v \neq 0$  del subespai

$$\{v : g_i'(a) \cdot v = 0, \forall i \text{ tal que } g_i(a) = 0 \text{ i } \mu_i > 0\},$$

aleshores  $a$  és mínim local.

De la segona equació de a) es dedueix  $\mu_1 = 0$ , i la primera queda  $f'(a) + \mu_2 \cdot (-1) = 0$ , d'on obtenim  $\mu_2 = f'(a)$ .

En b), tenim que  $M(a) = f''(a)$ . Hi ha dos casos:

i) Si  $\mu_2 > 0$ , queda la condició

$$v \cdot f''(a) \cdot v > 0, \forall v \in \{v \neq 0 : (-1) \cdot v = 0\} = \emptyset,$$

és a dir, cap condició. És el cas corresponent a la possibilitat  $f'(a) > 0$ .

ii) Si  $\mu_2 = 0$ , queda

$$v \cdot f''(a) \cdot v > 0, \forall v \in \{v \neq 0 : -\} = \mathbb{R} - \{0\},$$

és a dir,  $v \cdot f''(a) \cdot v > 0, \forall v \neq 0$ , és a dir,  $f''(a) > 0$ . És el cas corresponent a  $f'(a) = 0$  i  $f''(a) > 0$ .

□

## 1.5. Exercicis addicionals

**Exercici 1.5.1.** Considereu el problema del display (12.36), pàg 322 de Nocedal-Wright il·lustrat en la Figura 12.11. Mirant la figura, heu de veure clar que el mínim està en  $(1, 0)$ . Per tant, només són actives dues de les restriccions. Els multiplicadors KKT de les altres dues restriccions seran zero. Trobeu els multiplicadors de les restriccions actives, comprovant que són positius. (Allà mateix teniu la solució.)

A més, comproveu que el punt  $(1, 0)$  és regular. □

**Exercici 1.5.2.** Es tracta de trobar el punt de la paràbola  $5y = (x - 1)^2$  que està més a prop del punt  $(1, 2)$ . Plantegeu-ho com un problema d'optimització amb restriccions. (En l'Exercici 12.18, pàg 354 de Nocedal-Wright, teniu el planteig.) Intenteu fer un dibuix.

Plantegeu les condicions necessàries de mínim local. Resoleu el sistema. De les dues solucions que surten una no és vàlida, i l'altra sí.

Comproveu que la solució vàlida és un punt regular.

Comproveu que es compleixen les condicions suficients de mínim local en el punt trobat.

Considereu aquesta altra idea: Substituint la restricció a la funció objectiu podem eliminar la variable  $x$  i ens queda una funció de segon grau en  $y$ , de la qual podem calcular el mínim sense restriccions. Podem resoldre el problema així? □

**Exercici 1.5.3.** Considereu l'Exercici 12.20, pàg 354 de Nocedal-Wright. Trobeu els quatre punts que satisfan les condicions necessàries de primer ordre, i els multiplicadors de Lagrange corresponents.

Avaluant la funció objectiu en els quatre punts, decideu quins dos segur que no són mínims. Amb els altres dos, comproveu que les condicions suficients no es compleixen. No obstant, penseu un motiu pel qual han de ser mínims locals tots dos. □

**Exercici 1.5.4.** Semblant a l'anterior, però amb restricció de desigualtat, considereu l'Exercici 12.21. Cal canviar el signe de la  $f$  perquè l'enunciat parla de trobar màxims. Aprofitant els càlculs del problema anterior, trobeu els dos punts candidats (el signe del multiplicador descarta els altres dos). Raoneu igual que abans que els dos candidats són òptims. □

**Exercici 1.5.5.** Considereu el problema del display (12.72), pàg 335 de Nocedal-Wright. Observeu que la restricció assenyala *fora* del cercle. La funció no és acotada, perquè podem anar a  $-\infty$  fent créixer  $x_1$ . Però igualment poden haver-hi mínims locals.

Un dels candidats a mínim és el punt  $(1, 0)$ . Comproveu que compleix les condicions suficients. (Ho teniu resolt a la pàgina 336.)  $\square$

**Exercici 1.5.6.** Considereu l'Exercici 12.14, pàg 353 de Nocedal-Wright. Es tracta de trobar el punt d'un semiespai de  $\mathbb{R}^n$  que té norma euclidiana (o sigui, distància a l'origen) més petita. La distància a l'origen d'un punt  $(x_1, \dots, x_n)$  és l'arrel quadrada de  $\sum_{i=1}^n x_i^2$ . Però com que l'arrel quadrada és una funció creixent, podem simplement minimitzar la suma de quadrats i el punt òptim serà el mateix (penseu-ho). Per tant, seguint les notacions de l'exercici, tenim el problema de minimitzar  $x_1^2 + \dots + x_n^2$  subjecte a  $-(a_1x_1 + \dots + a_nx_n + \alpha) \leq 0$ .

Aquest problema es pot resoldre completament amb les condicions necessàries i distingint casos segons el signe de  $\alpha$ .  $\square$

---

---

## Algorismes d'optimització en $\mathbb{R}$

La teoria de condicions necessàries i suficients del Capítol 1 ens diu que per trobar els mínims locals d'una funció d'una variable  $f(x)$  en un interval  $[a, b]$ , cal buscar les solucions de  $f'(x) = 0$  en  $(a, b)$ , i estudiar apart els punts frontera  $a$  i  $b$ .

Però resoldre una equació  $f'(x) = 0$  és, en general, un problema tan difícil com l'original. El que es fa és utilitzar procediments iteratius que van obtenint, a partir d'un punt factible, un altre punt millor, a partir d'aquest un altre millor, i així successivament, amb l'esperança de convergir a un mínim local.

---

### 2.1. El mètode de Newton

Una idea molt natural és substituir la funció original  $f: \mathbb{R} \rightarrow \mathbb{R}$  per una funció que se li assembla, que tingui segur un mínim global, i que aquest mínim sigui fàcil de calcular. Si estem prop del mínim i l'aproximació de  $f$  és bona, el nou punt que s'obtindrà segurament estarà encara més a prop del mínim de  $f$ . I podem iterar el procediment per acostar-nos encara més.

La funció més fàcil complint les condicions anteriors és una funció quadràtica convexa, que en un punt determinat coincideixi amb  $f$ , i que les derivades primera i segona també coincideixin. És a dir, el polinomi de Taylor de segon ordre  $P(x)$ . Si partim d'un punt  $x_1$ , tindrem

$$P(x) = f(x_1) + f'(x_1) \cdot (x - x_1) + \frac{1}{2} f''(x_1) \cdot (x - x_1)^2,$$

suposant que  $f \in C^2$ . Calculem el mínim de  $P(x)$  i l'anomenem  $x_2$ :

$$x_2 = x_1 - \frac{f'(x_1)}{f''(x_1)}.$$

Repetint

$$x_3 = x_2 - \frac{f'(x_2)}{f''(x_2)}.$$

En general, obtenim la fórmula iterativa

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}. \quad (2.1)$$

[Mireu la Figura 8.5, pàgina 222 de Luenberger. La gràfica discontinua és la paràbola que aproxima en el punt  $x_k$ . El següent punt és el mínim d'aquesta paràbola (una mica mal assenyalat en el dibuix).]

Observeu que a la fórmula del mètode de Newton no intervé el valor de  $f$  en el punt  $x_k$ . Podríem per exemple sumar una constant a  $f$  i les iteracions serien les mateixes. Estem en realitat treballant amb  $g(x) := f'(x)$ , i buscant una solució de  $g(x) = 0$ . Sota aquesta perspectiva, el mètode de Newton és un algorisme per trobar zeros de funcions, amb la fórmula

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

[Mireu la Figura 8.6, pàgina 223 de Luenberger. La funció  $g$  representada és la derivada de la  $f$  de la figura anterior. La recta és la derivada de la paràbola. El punt de tall amb l'eix  $x$  d'aquesta recta és el nou punt  $x_{k+1}$ . El mètode de Newton també s'anomena *mètode de la tangent*, per aquesta interpretació.]

El problema amb el mètode de Newton és que no està garantit que funcioni. D'entrada, si en el punt inicial, o en qualsevol punt  $x_k$  durant el procés, es té  $f''(x_k) < 0$ , la paràbola aproximadora serà còncaua i no estarem fent res útil. Però fins i tot amb funcions  $f$  estrictament convexes, el mètode de Newton pot fallar.

Sota certes condicions, si comencem prou a prop d'un mínim local, hi convergirem. Això és el que afirma la proposició següent, de la qual es despendrà que en cas afirmatiu la convergència és molt ràpida.

**Proposició 2.1.1.** *Suposem que  $f: \mathbb{R} \rightarrow \mathbb{R}$  és de classe  $C^3$ .*

*Sigui  $x^*$  un punt tal que  $f'(x^*) = 0$ ,  $f''(x^*) \neq 0$ .*

*Aleshores, existeix  $\delta > 0$  tal que si  $|x_1 - x^*| < \delta$ , la successió definida per (2.1) i començant en  $x_1$  convergeix a  $x^*$ .*

*Demostració.* Denotem  $g(x) := f'(x)$ . Usant  $g(x^*) = 0$  podem escriure

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - \frac{g(x_k) - g(x^*)}{g'(x_k)} \\ &= \frac{g(x^*) - g(x_k) - g'(x_k)(x^* - x_k)}{g'(x_k)} . \end{aligned}$$

Per Taylor, el numerador és  $\frac{1}{2}g''(\xi)(x^* - x_k)^2$ , per un cert punt  $\xi$  entre  $x_k$  i  $x^*$ . Tindrem

$$x_{k+1} - x^* = \frac{g''(\xi)}{2g'(x_k)}(x^* - x_k)^2 . \quad (2.2)$$

En un cert entorn de  $x^*$ , per continuïtat,  $g'$  serà diferent de zero i  $g''$  estarà acotada. Per tant,

$$|x_{k+1} - x^*| \leq C \cdot |x_k - x^*|^2 , \quad (2.3)$$

per una certa constant  $C$ , independent de  $k$ . Si estem encara més a prop per tal que  $M := C|x_1 - x^*| < 1$ , obtenim

$$|x_{k+1} - x^*| < M^k |x_1 - x^*| \longrightarrow 0 \quad \text{quan } k \rightarrow \infty .$$

□

Malgrat que la Proposició 2.1.1 ens assegura la convergència si estem prou a prop del mínim local, és impossible saber si estem prou a prop, en general. Per tant, és un mètode a aplicar amb precaució i que no és apte com a algorisme general a usar de manera cega.



**Exercici 2.1.2.** L'arrel quadrada de 2 és una de les solucions de l'equació  $x^2 - 2 = 0$ . Equivalentment, és un mínim local de  $f(x) = \frac{1}{3}x^3 - 2x$ . Apliqueu a mà el mètode de Newton (amb una calculadora de butxaca), fent veure que només sabem que l'arrel està prop de  $x_1 = 1.5$ . Observeu que si aneu, per exemple, fent el càlcul amb sis decimals, després de tres iteracions ja tenim els sis decimals correctes.

Aquesta manera de calcular arrels quadrades ja es coneixia des del segle I almenys (Heron d'Alexandria).  $\square$

**Exemples 2.1.3.** És fàcil trobar exemples de funcions on el mètode falla. Considerem  $f(x) = |x|^{4/3}$ . La funció és contínua i derivable una vegada al voltant del zero, que és el mínim global. La segona derivada no existeix en  $x^* = 0$ . Aplicant la fórmula (2.1), tenim

$$x_{k+1} = x_k - \frac{\frac{4}{3}x_k^{1/3}}{\frac{4}{9}x_k^{-2/3}} = x_k - 3x_k = -2x_k .$$

Veiem que, independentment d'on comencem, cada cop ens allunyem més de la solució.

Encara que es compleixin les hipòtesis de la Proposició 2.1.1, també pot ser que la successió  $x_k$  oscil·li indefinidament entre dues regions. Per exemple, podeu comprovar molt ràpidament que  $f(x) = \frac{1}{4}x^4 - x^2 + 2x$  va oscil·lant entre 0 i 1 si comencem en algun d'aquests dos punts.  $\square$

**Exercici 2.1.4.** [Per fer amb un programa `Newton.c` que us passaré.] Considereu la funció  $f(x) = \log(1 + e^{2x}) - x$ . Calculeu les seves derivades primera i segona i comproveu que  $f$  és estrictament convexa a tota la recta real. El seu punt mínim és  $x = 0$ . Començant el mètode de Newton en  $x_1 = 1$ , el mètode convergeix. Començant en 1.1 en canvi, divergeix.  $\square$

## 2.2. Ordres de convergència

Un cop demostrada la convergència d'una successió a un cert límit  $x^*$ , com ara hem fet a la Proposició 2.1.1, un es pot preguntar per la velocitat d'aquesta convergència. La velocitat de convergència d'un algorisme és un criteri important de comparació d'algorismes. Suposem que  $\{x_k\}_k$  convergeix a  $x^*$  i que

$$|x_{k+1} - x^*| \leq C \cdot |x_k - x^*| , \quad \text{amb } C < 1 . \quad (2.4)$$

Per exemple, si  $C = 0.5$ , deduïm que la distància al límit es redueix a la meitat a cada iteració, almenys; si  $C = 0.1$ , a cada iteració reduïm la distància a una desena part, i per tant guanyem un dígit bo cada vegada almenys.

Observem que la relació (2.4) per sí sola ja demostra la convergència *global* (a partir de qualsevol  $x_1$ ) de la successió  $\{x_k\}_k$  al límit  $x^*$ . I observem que això no ho teníem en la demostració de la Proposició 2.1.1; ha estat necessari començar en un  $x_1$  complint  $C \cdot |x_1 - x^*| < 1$ , i llavors sí arribàvem a aquesta situació, amb la constant  $M$ .

Quan es compleix (2.4) per a tot  $k$  a partir d'un cert  $k_0$ , diem que tenim una *convergència lineal* (o *d'ordre 1*). La desigualtat (2.4) també es pot expressar

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|} \leq C , \quad \text{amb } C < 1 .$$

Quan, a més,

$$\lim_{k \rightarrow \infty} \frac{|x_{k+1} - x^*|}{|x_k - x^*|} = 0 ,$$

diem que tenim *convergència superlineal*.

Quan, a més,

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|^2} \leq C ,$$

(amb  $C$  no necessàriament més petita que 1!), diem que tenim *convergència quadràtica* (o *d'ordre 2*).

**Exercici 2.2.1.** Demostreu que  $x_k := 1 + k^{-k}$  convergeix superlinealment a 1. Demostreu que  $x_k := 1 + (\frac{1}{2})^{2^k}$  convergeix quadràticament a 1.  $\square$

La desigualtat (2.3), un cop sabem que la successió  $\{x_k\}_k$  convergeix a  $x^*$ , demostra que el mètode de Newton convergeix quadràticament. Prop del límit, a cada iteració es dobla aproximadament la quantitat de dígitos bons, perquè el quadrat d'un nombre petit és molt més petit. És per tant un mètode extremament eficient, quan convergeix.

**Exercici 2.2.2.** La condició  $f''(x^*) \neq 0$  no és necessària per tal que el mètode de Newton convergeixi. Però sí que, en cas contrari, la convergència quadràtica es pot convertir en lineal. Demostreu que això és el que passa amb la funció que té per derivada  $f'(x) = x^2$ .

De pas, observeu que en aquest cas,  $x^* = 0$  no és un mínim local de  $f$ . Però podem modificar l'exemple una mica per tenir un mínim i la convergència segueix essent només lineal: Agafeu  $f'(x) = \begin{cases} x^2, & \text{si } x \geq 0 \\ -x^2, & \text{si } x < 0 \end{cases}$ , que és la derivada de  $f(x) = \frac{1}{3}|x|^3$ , amb un mínim obvi en zero.  $\square$

En general, parlem de *convergència d'ordre  $p$*  quan tenim

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|^p} \leq C .$$

Tots aquests conceptes són igualment vàlids en  $\mathbb{R}^n$  per successions de vectors  $\{x_k\}_k$ , sense més que canviar els valors absoluts per normes.

### 2.3. Mètode de la posició falsa

Per ajustar la funció quadràtica donada pel polinomi de Taylor en el mètode de Newton s'usen tres dades:  $f(x_k)$ ,  $f'(x_k)$  i  $f''(x_k)$ . Es poden ajustar altres funcions quadràtiques amb tres dades diferents. Per exemple,  $f(x_k)$ ,  $f'(x_k)$  i  $f'(x_{k-1})$ . El polinomi de segon grau  $P(x) = a + b(x - x_k) + \frac{1}{2}c(x - x_k)^2$  que comparteix aquests tres valors amb la funció  $f$  és

$$P(x) = f(x_k) + f'(x_k) \cdot (x - x_k) + \frac{1}{2} \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}} \cdot (x - x_k)^2 .$$

Derivant i igualant a zero, obtenim la fórmula iterativa

$$x_{k+1} = x_k - f'(x_k) \frac{x_k - x_{k-1}}{f'(x_k) - f'(x_{k-1})} . \quad (2.5)$$

Veiem que tampoc intervé el valor de la funció en  $x_k$ , com en el mètode de Newton, i que es pot pensar que hem substituït la segona derivada per una aproximació (quocient incremental de  $f'$  entre  $x_{k-1}$  i  $x_k$ ).

Igual que en el mètode de Newton, aquest mètode intenta resoldre l'equació  $g(x) = 0$ , on  $g = f'$ . La fórmula iterativa es pot escriure llavors

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

[Mireu la Figura 8.3, pàgina 219 de Luenberger. La gràfica discontinua és la paràbola que aproxima  $f$  usant les derivades en  $x_k$  i  $x_{k-1}$ . El següent punt  $x_{k+1}$  és el mínim d'aquesta paràbola.

A la Figura 8.4, pàgina 220, veieu la mateixa situació en termes de la funció  $g$ . La recta discontinua és la derivada de la paràbola anterior, i s'anul·la en el nou punt  $x_{k+1}$ . Aquest mètode s'anomena també *mètode de la secant*.]

La proposició següent enuncia que aquest mètode convergeix si el punt inicial està prou a prop de la solució, sota les mateixes hipòtesis del mètode de Newton.

**Proposició 2.3.1.** *Suposem que  $f: \mathbb{R} \rightarrow \mathbb{R}$  és de classe  $C^3$ .*

*Signi  $x^*$  un punt tal que  $f'(x^*) = 0$ ,  $f''(x^*) \neq 0$ .*

*Aleshores, existeix  $\delta > 0$  tal que si  $|x_1 - x^*| < \delta$ , la successió definida per (2.5) i començant en  $x_1$  convergeix a  $x^*$ .*

*Demostració.* (Només idea.)

Semblantment a la demostració de la Proposició 2.1.1, però amb una manipulació una mica més complicada, s'arriba a una expressió semblant a (2.2). Concretament,

$$x_{k+1} - x^* = \frac{g''(\xi)}{2g'(\eta)} \cdot (x_k - x^*) \cdot (x_{k-1} - x^*) ,$$

on  $\xi$  i  $\eta$  són punts que estan en el segment que uneix  $x_k$ ,  $x_{k-1}$ ,  $x^*$ .

Sota les hipòtesis donades, prou a prop de  $x^*$  hi ha una constant  $C$ , independent de  $k$ , tal que

$$|x_{k+1} - x^*| \leq C \cdot |x_k - x^*| \cdot |x_{k-1} - x^*|$$

i raonant igual que en el mètode de Newton, si  $C \cdot (x_1 - x^*) < 1$  tenim la convergència de la successió  $\{x_k\}_k$  a  $x^*$ .  $\square$

L'ordre de convergència és més difícil d'analitzar en aquest cas. Es pot veure que és superlineal, però no arriba a quadràtica. Concretament,

$$\frac{|x_{k+1} - x^*|}{|x_k - x^*|^p} \leq C ,$$

amb  $p = \frac{1+\sqrt{5}}{2} \approx 1.618034$ , invers del número conegut com *la raó àuria*, que apareix en diferents contextos.

El mètode de la posició falsa es pot fer que convergeixi encara que el comencem lluny de  $x^*$  si podem forçar la situació en què, a cada iteració,  $f'(x_k)$  i  $f'(x_{k-1})$  tinguin els signes apropiats: Suposant  $x_{k-1} < x_k$ , si  $f'(x_{k-1}) < 0$  i  $f'(x_k) > 0$ , aleshores entre  $x_{k-1}$  i  $x_k$  hi haurà un mínim (vegeu altre cop les Figures 8.3 i 8.4 de Luenberger per convèncer-vos que és així).

Si els dos primers punts  $x_1$  i  $x_2$  compleixen aquesta condició, aleshores podem mantenir-la agafant, a la fórmula (2.5), en lloc de  $x_{k-1}$ , l'últim punt trobat  $x_n$  que compleixi  $f'(x_k)f'(x_n) < 0$ . A canvi d'assegurar la convergència, però, l'ordre de convergència es torna només lineal.

## 2.4. Ajust quadràtic de tres punts

Si no volem usar derivades, necessitem tres punts per ajustar una paràbola, usant el valor de  $f$  en els tres punts. Suposem que tenim tres punts  $x_1 < x_2 < x_3$  que estan en "posició convexa":

$$f(x_1) \geq f(x_2) \leq f(x_3) .$$

Això implica, sabent només que  $f$  és contínua, l'existència d'un mínim local entre  $x_1$  i  $x_3$ .

Ajustem la funció quadràtica  $P(x) = ax^2 + bx + c$  que passa per  $(x_1, f(x_1))$ ,  $(x_2, f(x_2))$  i  $(x_3, f(x_3))$ . Els coeficients  $(a, b, c)$  són la solució del sistema

$$\left. \begin{aligned} ax_1^2 + bx_1 + c &= f(x_1) \\ ax_2^2 + bx_2 + c &= f(x_2) \\ ax_3^2 + bx_3 + c &= f(x_3) \end{aligned} \right\} .$$

El mínim de  $P$  és el punt  $x = -\frac{b}{2a}$ . Amb aquest  $x$  i dos dels punts  $x_1, x_2, x_3$ , podem sempre construir una nova terna de punts  $x_1 < x_2 < x_3$  en posició convexa, amb la qual tornem a iterar: Per exemple, si  $x_1 \leq x \leq x_2 < x_3$ , aleshores

- Si  $f(x) \leq f(x_2)$ , agafem com a nova terna  $x_1, x, x_2$ .
- Si  $f(x) \geq f(x_2)$ , agafem com a nova terna  $x, x_2, x_3$ .

I si  $x_1 < x_2 \leq x \leq x_3$ , es fa anàlogament (feu un dibuix).

**Exercici 2.4.1.** Considereu la funció  $f(x) = e^x - 2x$ . Avalueu-la en els punts  $x_1 = 0$ ,  $x_2 = 0.5$ ,  $x_3 = 1$ , comprovant que estan en “posició convexa”, i que per tant segur que hi ha un mínim local en l'interval  $[0, 1]$ .

Feu una iteració del mètode d'ajust quadràtic de tres punts. Us ha de sortir com a nou punt (amb quatre decimals)  $x = 0.6673$ . Avalueu  $f$  en aquest punt i comproveu que la nova terna de punts estarà formada per  $x_1 = 0.5$ ,  $x_2 = 0.6673$ ,  $x_3 = 1$ .

Si teniu eines per resoldre automàticament sistemes d'equacions, podeu fer un parell d'iteracions més i trobareu les ternes  $(0.6673, 0.6827, 1)$ , i després  $(0.6827, 0.6929, 1)$ . El mínim exacte (amb quatre decimals bons) és  $x^* = 0.6931$ .  $\square$

Com amb els altres mètodes vistos, no es pot assegurar en general que aquest mètode convergeixi. Si la successió produïda convergeix, i la funció és de classe  $C^1$ , el que es pot assegurar és que el límit és un punt de derivada nul·la. A més, pot eventualment passar que el nou punt  $x$  generat coincideixi amb algun dels punts anteriors. Per tant, alguna modificació s'hi ha de fer si es vol assegurar la convergència (en veurem una a l'Apartat 2.6).

Suposant que el mètode convergeixi, en general només es pot assegurar un ordre de convergència lineal. Sota certes hipòtesis, no molt exigents, es pot assegurar convergència si estem prou a prop del mínim, i amb un ordre de convergència una mica millor. Per exemple, es té el resultat següent:

**Proposició 2.4.2.** *Si  $f$  té derivada tercera Lipschitz,  $f'(x^*) = 0$ ,  $f''(x^*) \neq 0$ , i partim de tres punts  $x_1, x_2, x_3$  prou a prop de  $x^*$ , aleshores la successió  $\{x_k\}_k$  definida pel mètode d'ajust quadràtic convergeix a  $x^*$ , amb un ordre de convergència  $p \approx 1.324718$  (l'arrel real de l'equació  $p^3 - p - 1 = 0$ ).*  $\square$

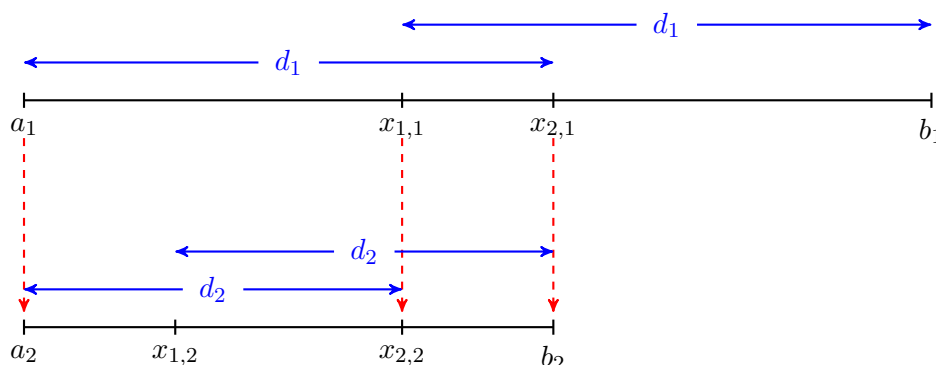
## 2.5. Mètodes de Fibonacci i de la raó àuria

Suposem que sabem que hi ha un mínim local en l'interval  $[a_1, b_1]$ . Hem vist en el mètode de l'ajust quadràtic que si tenim la funció  $f$  avaluada en  $a_1, b_1$  i dos punts  $a_2 < b_2$  de l'interior de l'interval, podem assegurar que hi ha un mínim en  $[a_1, b_2]$ , o bé en  $[a_2, b_1]$ .

Per tant, a cada pas podem reduir la longitud de l'interval que conté el mínim. Podem reduir la longitud de manera que convergim segur a un mínim? Ho podem fer de manera òptima (amb la quantitat més petita possible d'avaluacions de la funció)?

Suposem que tenim dret a avaluar la funció només en dos punts dins de l'interval  $[a_1, b_1]$ . Si volem assegurar-nos de reduir al màxim el nou interval que sortirà, és evident que els hem d'agafar ben junts al mig, per exemple  $\frac{a_1+b_1}{2} - \varepsilon$ ,  $\frac{a_1+b_1}{2} + \varepsilon$ . El nou interval tindrà una longitud màxima de  $\frac{1}{2}(b_1 - a_1) + \varepsilon$ .

Suposem que tenim dret a avaluar  $f$  en tres punts. Ja no és tant evident quins punts utilitzar. Però per simetria podem suposar que els dos primers cal agafar-los simètrics dins l'interval  $[a_1, b_1]$ . Anomenem aquests dos punts  $x_{1,1}$  i  $x_{2,1}$ , i sigui  $d_1$  les distàncies (a determinar, iguals i més grans que  $\frac{1}{2}(b_1 - a_1)$ ) marcades en el dibuix següent:



Després d'avaluar  $f$  en  $x_{1,1}$  i  $x_{2,1}$ , suposem que ens quedem amb l'interval  $[a_1, x_{2,1}]$ , que ara anomenem  $[a_2, b_2]$ , i el punt que tenim dins l'anomenem  $x_{2,2}$ . La longitud de l'interval inicial s'ha reduït en un factor  $r_1$ , amb  $d_1 = r_1(b_1 - a_1)$ .

L'últim punt l'agafarem en  $[a_2, b_2]$ , simètric respecte el  $x_{2,2}$ , perquè no tenim raons per privilegiar una o altra meitat de l'interval. Aquest punt serà el  $x_{1,2}$  del dibuix (no importa si està a una banda o a l'altre del  $x_{2,2}$ ).

L'interval  $[a_2, b_2]$  quedarà reduït en un factor  $r_2$ , amb  $d_2 = r_2(b_2 - a_2) = r_2 d_1$ . Respecte l'interval original, haurem reduït la longitud en un factor  $r_1 r_2$ . Volem que aquest número sigui el més petit possible. Però  $r_1, r_2$  estan lligats:

Per la part d'abaix del dibuix, observem que

$$b_2 - x_{2,2} = (1 - r_2)(b_2 - a_2) = (1 - r_2)r_1(b_1 - a_1) .$$

Per la part de dalt, aquesta mateixa longitud és la de

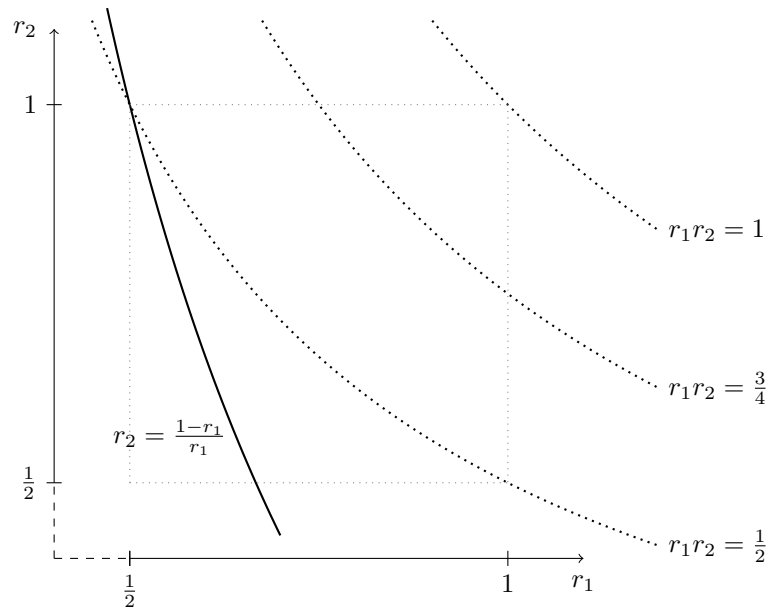
$$\begin{aligned} x_{2,1} - x_{1,1} &= (b_1 - a_1) - (x_{1,1} - a_1) - (b_1 - x_{2,1}) \\ &= (b_1 - a_1) - (1 - r_1)(b_1 - a_1) - (1 - r_1)(b_1 - a_1) \\ &= (2r_1 - 1)(b_1 - a_1) . \end{aligned}$$

Per tant,  $(1 - r_2)r_1 = 2r_1 - 1$ , que es pot escriure  $r_2 = \frac{1-r_1}{r_1}$ .

Busquem doncs resoldre aquest problema d'optimització de dues variables:

$$\begin{aligned} &\min r_1 r_2 \\ &\text{subjecte a: } \begin{cases} r_2 = \frac{1-r_1}{r_1} \\ \frac{1}{2} \leq r_1 \leq 1 \\ \frac{1}{2} \leq r_2 \leq 1 \end{cases} \end{aligned}$$

Aquest problema es pot resoldre geomètricament i la solució és  $r_1 = \frac{2}{3}$ ,  $r_2 = \frac{1}{2}$ . Vegeu la figura següent:



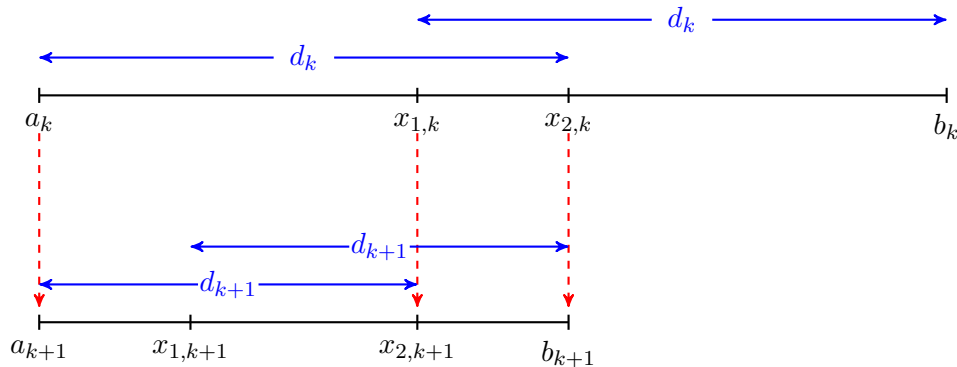
D'aquí obtenim que els dos primers punts han de dividir l'interval inicial en tres trossos iguals, i que el tercer punt hauria de coincidir amb un dels anteriors. Això era esperable, perquè quan només podem agafar un punt més, estem en la situació que descrivíem en el tercer paràgraf d'aquest apartat. Per tant, a la pràctica, l'últim interval el dividirem per la meitat aproximadament.

Si poguéssim avaluar la funció en quatre punts, arribaríem, amb els mateixos raonaments, a resoldre

$$\begin{aligned} & \min r_1 r_2 r_3 \\ & \text{subjecte a: } \begin{cases} r_2 = \frac{1-r_1}{r_1} \\ r_3 = \frac{1-r_2}{r_2} \\ \frac{1}{2} \leq r_k \leq 1 \end{cases} \end{aligned}$$

La solució és  $r_1 = \frac{3}{5}$ ,  $r_2 = \frac{2}{3}$ ,  $r_3 = \frac{1}{2}$ .

En general, entre una iteració  $k$  i la iteració  $k + 1$ , tenim l'esquema



A cada iteració, la longitud es redueix en un factor  $\frac{1}{2} \leq r_k \leq 1$ , les distàncies es van obtenim mitjançant  $d_{k+1} = r_k d_k$ , i el lligam entre  $r_k$  i  $r_{k+1}$  ve donat per

$$r_{k+1} = \frac{1-r_k}{r_k} .$$

Per tant, si podem avaluar la funció en  $N$  punts (apart dels extrems inicials), arribem al problema

d'optimització

$$\begin{aligned} & \min r_1 r_2 \cdots r_{N-1} \\ \text{subjecte a: } & \begin{cases} r_{k+1} = \frac{1-r_k}{r_k}, & k = 1, \dots, N-2 \\ \frac{1}{2} \leq r_k \leq 1, & k = 1, \dots, N-1 \end{cases} \end{aligned}$$

I la solució resulta ser

$$r_k = \frac{F_{N-k}}{F_{N-k+1}}$$

on  $F_n$  és el  $n$ -èssim terme de la successió de Fibonacci començant en  $F_0 = 1$ ,  $F_1 = 1$ ,  $F_2 = 2$ , etc.

Al final del procés, l'interval inicial queda reduït en un factor  $r_1 \cdots r_{N-1} = \frac{F_{N-1}}{F_N} \cdots \frac{F_1}{F_2} = \frac{1}{F_N}$ . Com hem vist en els exemples de pocs punts, l'última avaluació s'ha de fer al costat de la penúltima, i per tant l'interval final tindrà una longitud (en el pitjor cas) lleugerament més gran que  $\frac{1}{F_N}(b_1 - a_1)$ . Excepte aquest detall, podem determinar d'entrada quants punts hem d'avaluar per aconseguir una longitud final fixada d'entrada: Si volem una tolerància d'error de  $\pm \varepsilon$  haurem d'agafar  $N - 1$  punts, on  $N$  és el natural més petit tal que  $F_N > \frac{b_1 - a_1}{\varepsilon}$ .

Els números de Fibonacci compleixen que

$$\lim_{N \rightarrow \infty} \frac{F_{N-1}}{F_N} = \varphi,$$

on  $\varphi$  és la solució positiva de l'equació  $\varphi^2 + \varphi - 1 = 0$ , que és  $\varphi = \frac{-1 + \sqrt{5}}{2} \approx 0.618034$ , conegut com la *raó àuria*. Si en el mètode de Fibonacci hem d'avaluar molts punts, al final estarem agafant a cada pas un  $r_k$  semblant a  $\varphi$ . El *mètode de la raó àuria* consisteix en agafar ja d'entrada tots els factors  $r_k$  iguals a  $\varphi$ . No serà el mètode òptim per reduir al màxim els intervals, però és més senzill i no cal saber d'entrada en quants punts podem avaluar la funció. Amb aquest mètode garantim que a cada pas, la longitud del nou interval redueix en un factor  $\varphi$  la longitud de l'interval anterior.

A diferència dels mètodes de Newton, posició falsa i ajust quadràtic de tres punts, els de Fibonacci i la raó àuria no consisteixen en aproximar  $f$  per una altra funció senzilla, sinó en anar reduint progressivament l'interval que conté el mínim en un cert factor predeterminat. Per això s'anomenen "bracketing methods". Quant a l'ordre de convergència, està clar que la raó àuria convergeix amb un ordre lineal.

**Exercici 2.5.1.** Feu un parell d'iteracions del mètode de la raó àuria a la funció de l'Exercici 2.4.1. □

## 2.6. Mètode de Brent i la GSL

La biblioteca GSL conté els algorismes de la raó àuria i el d'ajust quadràtic modificat per assegurar la convergència. Aquest últim s'anomena *mètode de Brent*. La següent és una descripció amb més o menys (i evitant detalls) la mateixa notació que apareix en el fitxer `brent.c` de la GSL. Bàsicament, és el mètode d'ajust quadràtic, sense exigir els punts en posició convexa, i intercalant alguns passos de la raó àuria quan el nou punt no és admissible per algun motiu.

- Al principi de cada iteració, se suposa que hi ha un interval  $(a, b)$  a l'interior del qual se sap que hi ha un mínim local. No se suposa que s'hagi avaluat  $f$  en els extrems. Els punts següents  $x, w, u, v$  no necessàriament són tots diferents.
- $x$  és el valor més petit conegut de  $f$  (el més recent avaluat en cas d'empat).

- Si  $\max\{x - a, b - x\} < 2 \cdot \text{tol}$ , on “tol” és una tolerància predefinida, s’acaba el procediment i es retorna  $x$ .
- $w$  és el segon punt amb valor més petit.
- $v$  és el valor de  $w$  en la iteració anterior.
- $u$  és l’últim punt on s’ha avaluat  $f$  (indefinit en la inicialització).
- S’interpola una paràbola pels punts  $(v, f(v))$ ,  $(w, f(w))$ ,  $(x, f(x))$ . L’extrem de la paràbola està representat en `brent.c` pel quocient  $p/q$  de dos números, amb  $q > 0$ .
- Si el vèrtex de la paràbola està fora de  $(a, b)$ , o la paràbola degenerava a una recta (perquè dos o més punts coincidien), o la diferència  $|e|$  entre  $x$  i el vèrtex de la paràbola és més petit que la tolerància, o  $|e|$  és més gran que la meitat del  $|e|$  del pas anterior, aleshores es fa un pas de raó àuria.
- S’actualitza  $u$  al nou punt determinat en el pas anterior (sigui com a vèrtex de la paràbola o amb la raó àuria). Es calcula  $f(u)$ .
- S’actualitzen  $a, b, x, w, u, v$  perquè tinguin el significat que han de tenir.
- En el pas inicial,  $v = w = x = a + (1 - \varphi)(b - a)$ , on  $\varphi$  és la raó àuria. És a dir, el primer que es fa és un pas del mètode de la raó àuria.

En el manual de la GSL, l’ús dels algorismes de minimització en dimensió 1 està explicat al Capítol 37. Com veureu allà, la biblioteca proporciona iteracions individuals de l’algorisme en qüestió, i l’usuari ha de controlar l’execució amb un codi propi, accedint al detall de les iteracions a través de determinades funcions de la biblioteca. Això és el que hem d’aprendre a fer.

En particular, cal controlar en quin moment acabem l’execució de l’algorisme. La funció `gsl_min_test_interval(double x_lower, double x_upper, double epsabs, double epsrel)` ens diu si l’interval actual  $[a, b] = [x\_lower, x\_upper]$  on se sap que hi ha el mínim, compleix la condició d’aturada. Si abreviem  $\varepsilon_a = \text{epsabs}$ , i  $\varepsilon_r = \text{epsrel}$ , la condició és

$$|a - b| < \varepsilon_a + \varepsilon_r \cdot \begin{cases} \min\{|a|, |b|\} , & \text{si } 0 \notin [a, b] \\ 0 , & \text{si } 0 \in [a, b] \end{cases}$$

Els números  $\varepsilon_a$ ,  $\varepsilon_r$  que nosaltres passem a la funció controlen l’error absolut i l’error relatiu que volem. És possible posar  $\varepsilon_r = 0$  i controlar simplement l’error absolut: Com que, en el punt actual  $x$ , la distància al mínim  $x^*$  serà  $|x - x^*| < |a - b| < \varepsilon_a$ , estarem dins aquesta cota de l’error si ens aturem aquí.

També és possible posar  $\varepsilon_a = 0$ , i controlar només l’error relatiu: Si  $0 \notin [a, b]$ ,

$$\begin{aligned} |x - x^*| < |a - b| < \varepsilon_r \cdot \min\{|a|, |b|\} < \varepsilon_r \cdot |x^*| \\ \implies \frac{|x - x^*|}{|x^*|} < \varepsilon_r . \end{aligned}$$

Mentre 0 estigui dins l’interval  $[a, b]$ , l’error relatiu no es pot controlar d’aquesta manera i per això es canvia per zero. Si se sospita que el mínim està exactament en  $x^* = 0$ , aleshores cal posar un valor positiu a  $\varepsilon_a$ , o potser no es complirà mai el criteri d’aturada.



---

## Resum de mètodes

Mètode	Requereix calcular	Convergeix	Ordre de convergència	Tipus de mètode
Newton	$f', f''$	Prop del mínim	2	Ajust de funcions
Posició falsa	$f'$	Prop del mínim	$\approx 1.618$	Ajust de funcions
Ajust quadràtic de tres punts	$f$	Prop del mínim	$\approx 1.325$	Ajust de funcions
Fibonacci	$f$	NA	NA	Bracketing
Raó àuria	$f$	Sempre	1	Bracketing
Brent	$f$	Sempre	1 (típicament, $\approx 1.325$ )	Ajust de funcions + Bracketing



---

---

## Optimització sense restriccions en $\mathbb{R}^n$

Resoldre explícitament les condicions KKT de primer ordre (Teorema 1.3.8, 1), per trobar candidats a mínim local és impossible llevat de casos molt senzills. Per aquest motiu, s'han anat desenvolupant algorismes pràctics, com els que hem vist al Capítol 2 per  $\mathbb{R}$ , per obtenir bones aproximacions de la solució.

En aquest capítol considerarem el problema d'optimització en  $\mathbb{R}^n$  sense restriccions. Per tant, les condicions KKT es redueixen al sistema d'equacions  $\nabla f(x) = 0$ , o sigui

$$\begin{cases} \frac{\partial f}{\partial x_1}(x_1, \dots, x_n) = 0 \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_1, \dots, x_n) = 0 \end{cases}$$

Encara que no es puguin resoldre exactament, les condicions KKT proporcionen la base d'inspiració per als algorismes, i un mecanisme per comprovar si podem estar prop d'un punt candidat. Per altra banda, les condicions suficients de segon ordre, que ara es redueixen a demanar que  $\nabla^2 f(x^*)$  sigui definida positiva, juguen un paper en la velocitat de convergència dels algorismes. Ho hem vist en el cas del mètode de Newton en  $\mathbb{R}$  (recordeu l'Exercici 2.2.2).

Típicament, els algorismes parteixen d'un punt inicial  $x_1 \in \mathbb{R}^n$ , subministrat per l'usuari, i generen una successió de punts  $\{x_k\}_k \subset \mathbb{R}^n$ . L'usuari també ha de decidir quan aturar l'algorisme, si veu que no hi ha progrés o creu que ja s'està prou a prop del mínim buscat. Per passar del punt  $x_k$  al punt  $x_{k+1}$  s'utilitza informació de la funció objectiu en el punt  $x_k$ , i potser també en els punts anteriors.

Tots els mètodes d'aquest capítol s'enfoquen a buscar un mínim local. No està garantit que sigui el mínim global, excepte si la funció és convexa.

---

### 3.1. Mètode de Newton

El mètode de Newton en  $\mathbb{R}^n$  és una generalització directa del mateix mètode que hem vist en dimensió 1. Donat un punt  $x_k$  i els valors de  $f(x_k)$ ,  $\nabla f(x_k)$  i  $\nabla^2 f(x_k)$ , podem construir la funció quadràtica que millor aproxima la funció  $f$  en el punt  $x_k$ , que no és més que el polinomi de Taylor de grau 2:

$$P(x) = f(x_k) + \nabla f(x_k) \cdot (x - x_k) + \frac{1}{2}(x - x_k) \cdot \nabla^2 f(x_k) \cdot (x - x_k)$$

El nou punt  $x_{k+1}$  serà el mínim de  $P$ : Calculem el seu gradient,<sup>1</sup>

$$\nabla P(x) = \nabla f(x_k) + \nabla^2 f(x)(x - x_k) ,$$

l'igualem a zero i aïllem  $x$ :

$$x_{k+1} := x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k) . \quad (3.1)$$

Anàlogament al cas d'una variable, si el mètode de Newton convergeix, ho farà a un punt  $x^*$  tal que  $\nabla f(x) = 0$  (*punt estacionari*). També com en dimensió 1, l'ordre de convergència és quadràtic, en principi, i per tant l'algorisme convergeix molt ràpidament prou a prop del mínim.

Compareu la proposició següent amb la que vam veure en una variable.

**Proposició 3.1.1.** *Suposem que  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  és de classe  $C^3$ .*

*Sigui  $x^*$  un punt tal que  $\nabla f(x^*) = 0$ , i  $\nabla^2 f(x^*)$  és definida positiva.*

*Aleshores, existeix  $\delta > 0$  tal que si  $\|x_1 - x^*\| < \delta$ , la successió definida per (3.1) i començant en  $x_1$  convergeix a  $x^*$ .*

En el cas de dimensió 1, vèiem a la demostració d'aquesta proposició que la convergència era quadràtica, amb una constant  $C$  que depenia de la derivada segona de  $f$  (la  $g'$  que apareixia dividint). Aquest paper ara el juga el valor propi més petit de la matriu  $\nabla^2 f$  prop del mínim, que és el mateix que el valor propi més gran de la matriu inversa. Igual que en el cas d'una variable, es dedueix la convergència quadràtica

$$\frac{\|x_{k+1} - x^*\|}{\|x_k - x^*\|^2} \leq C ,$$

on la constant  $C$  és proporcional al valor propi més gran de  $[\nabla^2 f(x^*)]^{-1}$ . Intuïm, per analogia amb el cas d'una variable, que si la matriu  $\nabla^2 f(x^*)$  no és invertible, la convergència, si n'hi ha, no serà quadràtica. I això és cert també.

L'inconvenient del mètode de Newton en  $\mathbb{R}^n$ , apart dels que ja sabem en  $\mathbb{R}$ , és que cal calcular totes les derivades segones, i després invertir la matriu, que és un procediment costós i numèricament problemàtic.

Òbviament també, cal que a cada pas la matriu Hessiana  $\nabla^2 f(x_k)$  sigui definida positiva, per estar segurs que la funció aproximadora és convexa (un paraboloides el·líptic convex) i que el que estem calculant a cada iteració és el seu mínim. Aquesta condició és tant important que en les variants del mètode de Newton que ara veurem, es demana de manera essencial que la matriu que substitueixi a la inversa de la Hessiana sigui definida positiva, encara que no tingui gaire a veure amb les derivades segones.

Considerem el mètode iteratiu genèric

$$x_{k+1} = x_k - B_k^{-1} \nabla f(x_k) . \quad (3.2)$$

**Proposició 3.1.2.** *Si  $B_k$  és una matriu definida positiva, aleshores  $d_k := -B_k^{-1} \nabla f(x_k)$  és una direcció de descens.*

<sup>1</sup>Si no veieu clar perquè el gradient té aquesta expressió, escriviu-vos el polinomi de Taylor desplegant els productes:

$$P(x) = f(x_k) + \sum_i \nabla f(x_k)_i \cdot (x - x_k)_i + \frac{1}{2} (x - x_k)_i \cdot \nabla^2 f(x_k)_{ij} \cdot (x - x_k)_j ,$$

i calculeu les derivades parcials  $\frac{\partial P}{\partial x_r}(x)$ .

*Demostració.* Només cal observar que

$$\nabla f(x_k) \cdot (-B_k^{-1} \nabla f(x_k)) < 0$$

perquè si  $B_k$  és definida positiva,  $B_k^{-1}$  també, i recordar que les direccions de descens  $d_k$  són precisament les que compleixen  $\nabla f(x_k) \cdot d_k < 0$ .  $\square$

## 3.2. Mètode del gradient

El mètode *mètode del gradient*, o del *descens de més gran pendent*, o *steepest descent*, consisteix en agafar  $B_k$  igual a la identitat. Tindrem, en principi, la fórmula iterativa

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

Aquesta elecció sembla molt natural, perquè precisament la direcció  $-\nabla f(x_k)$  és la de màxim descens des del punt  $x_k$ . Ara bé, anirem a parar més o menys lluny de  $x_k$  segons la magnitud del gradient. Encara que anem en una direcció de molt descens, no està garantit que el punt d'arribada  $x_{k+1}$  compleixi  $f(x_{k+1}) < f(x_k)$ . I aquesta desigualtat seria molt interessant que es complís a cada iteració, per assegurar almenys que anem millorant la funció objectiu.

Per tant, és lògic escollir  $-\nabla f(x_k)$  com a direcció de moviment, però també introduir un paràmetre  $\alpha_k$  que ens determini la longitud del pas que farem en aquesta direcció:

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k) .$$

Idealment, volem descendir el més possible. Si abreviem  $d_k := -\nabla f(x_k)$ , el que volem és trobar  $\alpha > 0$  que minimitzi la funció

$$\phi(\alpha) = f(x_k + \alpha d_k) .$$

[Vegeu la Figura 8.9 de Luenberger, pàg 234: La figura representa les corbes de nivell d'una funció que té el mínim en  $x^*$ . El gradient és perpendicular a les corbes de nivell. Si agafem el  $x_0$  com a punt inicial i ens movem en la direcció del gradient negatiu, la funció va descendant fins que arriba al punt  $x_1$ , a partir del qual començaria a créixer altre cop. Per tant ens aturem allà i en la següent iteració escollim com a direcció el gradient negatiu en  $x_1$ . La funció va descendant i ens aturem allà on torna a pujar, que serà el punt  $x_2$ . I així successivament. En el dibuix sembla que cada direcció de moviment és perpendicular a l'anterior. I això és cert, com ara veurem.]

**Proposició 3.2.1.** *La direccions  $d_k$  que s'obtenen amb el procediment descrit, compleixen  $d_{k+1} \cdot d_k = 0$  (cada direcció és ortogonal a l'anterior).*

*Demostració.* Per la regla de la cadena,

$$\phi'(\alpha) = \nabla f(x_k + \alpha d_k) \cdot d_k .$$

El punt  $\alpha_k$  que minimitza  $\phi$  compleix que aquesta derivada és zero:

$$0 = \phi'(\alpha_k) = \nabla f(x_k + \alpha_k d_k) \cdot d_k = \nabla f(x_{k+1}) \cdot d_k = -d_{k+1} \cdot d_k .$$

$\square$

El mètode del gradient convergeix segur a un punt de gradient nul. Però la convergència pot ser molt lenta. En el dibuix de Luenberger que acabem de veure observem que al acostar-nos al mínim es perden moltes iteracions amb moviments petits. En canvi, si les corbes de nivell fossin circulars, arribaríem al mínim en una sola iteració. La ziga-zaga característica d'aquest mètode és degut al diferent escalat de les variables.

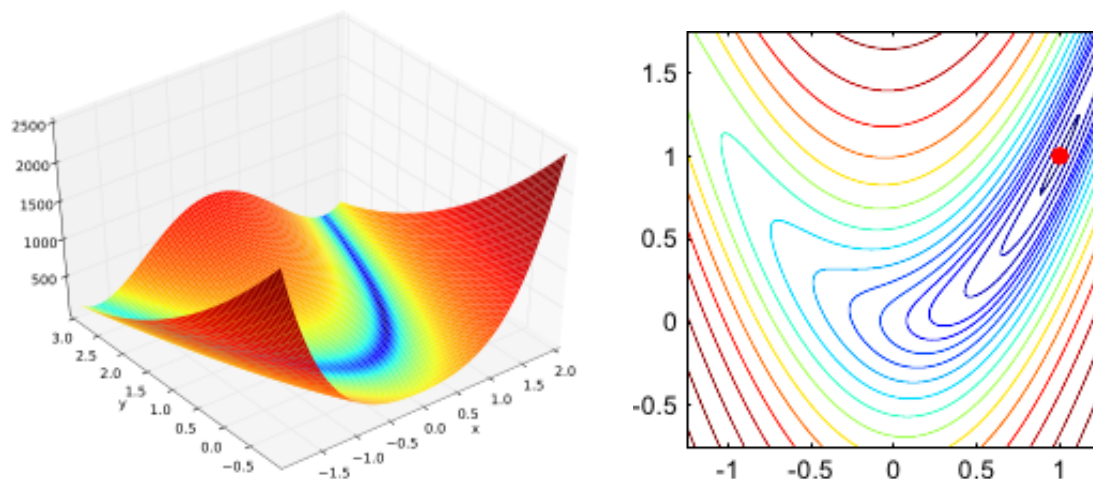


Figura 3.1: La banana de Rosenbrock. Representació 3d (Wikipedia) i corbes de nivell (www.brnt.eu)

Per exemple, la funció  $f(x_1, x_2) = 100x_1^2 + x_2^2$  té les variables mal escalades. Si fem  $z_1 = 10x_1$ ,  $z_2 = x_2$ , obtenim  $f(z_1, z_2) = z_1^2 + z_2^2$  i el mètode del gradient donaria el mínim exacte en una sola iteració. El mètode de Newton en canvi és insensible a l'escalat, perquè la matriu Hessiana ja codifica la curvatura.

**Exemple 3.2.2.** La funció  $f(x_1, x_2) = (x_1 - 1)^2 + 100(x_1^2 - x_2)^2$  s'anomena la *banana de Rosenbrock* (Figura 3.1). És una funció test que s'utilitza per provar i comparar algorismes (es poden posar altres valors en lloc del 1 i el 100). És evident que té un únic mínim en el punt  $(1, 1)$ , on la funció val zero, i es fàcil veure que la matriu Hessiana és definida positiva en el mínim. Però la funció no és convexa i està molt mal escalada prop del mínim.

**Exercici 3.2.3.** Considereu la funció de tres variables  $f(x_1, x_2, x_3) = e^{x_1^2} + x_2^2 + x_3^2 - x_2x_3$ .

- Feu una iteració del mètode de Newton a partir del punt  $(0, 2, 1)$ .
- Feu una iteració del mètode del gradient a partir del mateix punt.

### 3.3. Cerca lineal inexacta

En el mètode del gradient, o en qualsevol altre, no és bona idea gastar molts recursos per trobar el pas òptim  $\alpha_k$  en la direcció de moviment. Això pot requerir moltes avaluacions de la funció  $f$ , amb algun dels mètodes que hem estudiat al Capítol 2, només per fer una sola iteració del algorisme. Ha d'haver-hi un compromís entre el nombre d'iteracions d'un algorisme i la complexitat de cada iteració. Normalment es preferible invertir en fer més iteracions i no entretenir-se molt en cadascuna d'elles.

Això ens porta a les tècniques de *cerca lineal inexacta*.<sup>2</sup> Es tracta de trobar a cada iteració un valor  $\alpha_k$  que sigui bo, però sense buscar gaire.

La condició bàsica exigible és:

$$\mathbf{0)} \quad \phi(\alpha_k) < \phi(0).$$

Això és equivalent a dir que  $f(x_{k+1}) = f(x_k + \alpha_k d_k) < f(x_k)$ , o sigui que el valor de la funció objectiu millorarà. Si  $d_k$  és una direcció de descens, segur que existeix un  $\alpha_k$  complint això.

Aquesta condició sola no assegura, però, que el mètode convergeixi a un punt amb  $\nabla f(x) = 0$ . Hem de demanar més coses:

<sup>2</sup>En anglès, *inexact line search*. Aquí, *line* vol dir "en línia recta, unidimensional"; *inexact* es refereix a "imprecisa, poc acurada". La traducció és dolenta, però és l'habitual.

1) Fixant un  $\varepsilon$  amb  $0 < \varepsilon < 1$ , volem que

$$\phi(\alpha_k) \leq \phi(0) + \varepsilon \phi'(0) \alpha_k .$$

Equivalentment:  $f(x_k + \alpha_k d_k) \leq f(x_k) + \varepsilon \nabla f(x_k) \cdot d_k \cdot \alpha_k$  .

Vegeu la Figura 3.2. Hi ha una línia discontinua quasi horitzontal que representa la funció afí  $\alpha \mapsto \phi(0) + \varepsilon \phi'(0) \alpha$ . Com que  $\varepsilon < 1$ , aquesta funció té un pendent més moderat que  $\phi$  en el zero, i la seva gràfica anirà, almenys al començament, per sobre de la gràfica de  $\alpha \mapsto \phi(\alpha)$ , representada per la línia sòlida blava.

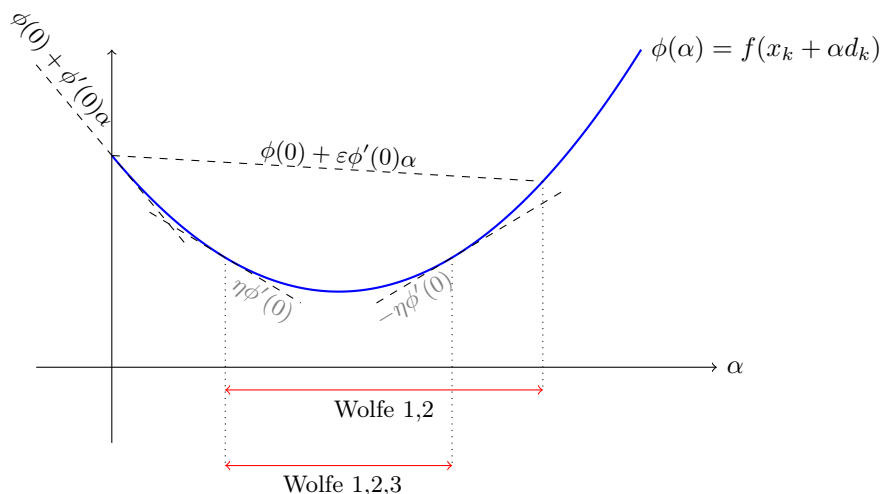


Figura 3.2: Il·lustració de les condicions de Wolfe. En gris els pendents de les rectes dibuixades.

Està clar que la condició la compleixen tots els  $\alpha$  prou petits. Típicament, s'agafa un valor petit per  $\varepsilon$ , com ara  $\varepsilon = 10^{-3}$ .

Com que tampoc volem fer passos molt petits, imposarem una altra condició:

2) Fixant un  $\eta$  amb  $\varepsilon < \eta < 1$ , volem que

$$\phi'(\alpha_k) \geq \eta \phi'(0) .$$

Equivalentment,  $\nabla f(x_k + \alpha_k d_k) \cdot d_k \geq \eta \nabla f(x_k) \cdot d_k$  .

La idea és no quedar-nos massa a prop del punt 0. Si la funció  $f$  està descendent ràpid en aquesta direcció, anem tirant endavant fins que ja no descendeixi tant ràpid com al principi. Se sol agafar un valor de  $\eta$  més aviat gran, com ara  $\eta = 0.9$  . A la Figura 3.2 veieu que hem d'anar més enllà del punt en què el pendent de  $\phi$  és  $\eta \phi'(0)$

Aquesta condició no descarta que  $\phi'(\alpha_k)$  sigui positiva (la funció ha descendit però ja ha tornat a pujar). Per tal que no sigui molt positiva, podem afegir opcionalment aquesta altra condició:

3) Amb el mateix  $\eta$  d'abans,

$$\phi'(\alpha_k) \leq -\eta \phi'(0) .$$

Equivalentment:  $\nabla f(x_k + \alpha_k d_k) \cdot d_k \leq -\eta \nabla f(x_k) \cdot d_k$  .

Les condicions 2) i 3) es poden resumir en  $|\phi'(\alpha_k)| \leq \eta |\phi'(0)|$  .

**Definició 3.3.1.** Les condicions 1) i 2) s'anomenen *the Wolfe conditions*.

Les condicions 1), 2) i 3) s'anomenen *the strong Wolfe conditions*.

[Les Figures 3.3, 3.4 i 3.5 de Nocedal-Wright, pàgines 33, 34 i 35, il·lustren les condicions de Wolfe 1 i 2 amb una funció  $\phi$  de forma més estranya. En particular es pot veure que poden haver valors de  $\alpha$  acceptables en intervals disjunts.]

Veurem que sempre existeixen passos  $\alpha_k$  complint les condicions de Wolfe, normals i fortes. Només cal demanar que la funció objectiu  $f$  estigui acotada inferiorment, cosa que suposarem sempre però ja no explicitarem, perquè en cas contrari segur que no hi ha mínim.

**Proposició 3.3.2.** *Existeix un interval de valors  $\alpha$  complint les condicions de Wolfe 1), 2), 3).*

*Demostració.* Fixem  $0 < \varepsilon < \eta < 1$ .

Signin  $\phi(\alpha) := f(x_k + \alpha d_k)$ , i  $\ell(\alpha) := \phi(0) + \varepsilon \phi'(0)\alpha$ , amb  $d_k$  direcció de descens de  $f$  en el punt  $x_k$ .

Signi  $\bar{\alpha}$  la primera intersecció de  $\phi$  i  $\ell$ ; qualsevol punt en  $(0, \bar{\alpha})$  compleix Wolfe-1. I tindrem la igualtat

$$\phi(\bar{\alpha}) = \phi(0) + \varepsilon \phi'(0)\bar{\alpha} . \quad (3.3)$$

Per altra banda, pel Teorema de Taylor d'ordre 1 (Teorema del Valor Mitjà), existeix  $\xi \in (0, \bar{\alpha})$  tal que

$$\phi(\bar{\alpha}) = \phi(0) + \phi'(\xi)\bar{\alpha} . \quad (3.4)$$

Igualant (3.3) i (3.4), obtenim

$$\phi'(\xi) = \varepsilon \phi'(0) > \eta \phi'(0) \quad (3.5)$$

(usant que  $\eta > \varepsilon$  i que  $\phi'(0) < 0$ ), i per tant  $\xi$  compleix Wolfe-2.

Finalment, de (3.4) també es dedueix que  $\phi'(\xi)$  és negatiu. Per tant,

$$\phi'(\xi) < 0 < -\eta \phi'(0) \quad (3.6)$$

i  $\xi$  compleix Wolfe-3.

Com que les tres condicions es compleixen amb desigualtat estricta, per continuïtat deduïm que les condicions se satisfan en tot un interval al voltant de  $\xi$ .  $\square$

Ens preguntem ara si, per a què un mètode d'optimització convergeixi a un punt  $x^*$  amb gradient nul, és suficient que escollim a cada iteració una direcció de descens  $d_k$ , i després un pas  $\alpha_k$  que compleixi les condicions de Wolfe. La resposta és que no. Sabem que les direccions de descens  $d_k$  compleixen  $\nabla f(x_k) \cdot d_k < 0$ , però si aquest producte escalar va tendint cap a zero, pot ser que al final l'algorisme s'encalli abans d'arribar al mínim.

Per altra banda, està clar que amb el mètode de descens de més gran pendent, això no pot passar, perquè  $\nabla f(x_k) \cdot d_k = -\|\nabla f(x_k)\|^2$ , que només s'acostarà a zero quan el gradient tendeixi a zero, que ja és el que volem.

Doncs en efecte, només cal exigir que els productes escalars  $\nabla f(x_k) \cdot d_k$  no se'n vagin cap a zero. Això serà conseqüència del següent teorema important:

**Teorema 3.3.3** (Teorema de Zoutendijk). *Si a cada iteració d'un algorisme d'optimització la direcció  $d_k$  és de descens i el pas  $\alpha_k$  compleix les condicions Wolfe-1 i Wolfe-2, aleshores*

$$\sum_{k=1}^{\infty} \cos^2 \theta_k \cdot \|\nabla f(x_k)\|^2 < \infty , \quad (3.7)$$

on  $\theta_k$  és l'angle entre  $\nabla f(x_k)$  i  $d_k$ , o sigui

$$\cos \theta_k = \frac{\nabla f(x_k) \cdot d_k}{\|\nabla f(x_k)\| \cdot \|d_k\|} . \quad (3.8)$$



*Demostració.* Per Wolfe-2, restant  $\nabla f(x_k) \cdot d_k$  a banda i banda de l'expressió en termes de  $f$ , tenim les desigualtats:

$$\begin{aligned} (\eta - 1)\nabla f(x_k) \cdot d_k &\leq (\nabla f(x_k + \alpha_k d_k) - \nabla f(x_k)) \cdot d_k \\ &\leq \|\nabla f(x_k + \alpha_k d_k) - \nabla f(x_k)\| \cdot \|d_k\| \\ &\leq L \cdot \|\alpha_k d_k\| \cdot \|d_k\| = L\alpha_k \|d_k\|^2 \end{aligned}$$

per certa constant  $L$ , pel Teorema del Valor Mitjà. Per tant,

$$\alpha_k \geq \frac{\eta - 1}{L} \cdot \frac{\nabla f(x_k) \cdot d_k}{\|d_k\|^2}. \quad (3.9)$$

Per Wolfe-1 i la desigualtat (3.9),

$$\begin{aligned} f(x_k + \alpha_k d_k) &\leq f(x_k) + \varepsilon \nabla f(x_k) \cdot d_k \cdot \alpha_k \\ &\leq f(x_k) - \varepsilon \cdot \frac{1 - \eta}{L} \cdot \frac{(\nabla f(x_k) \cdot d_k)^2}{\|d_k\|^2}. \end{aligned}$$

Anomenant  $C := \varepsilon \frac{1-\eta}{L}$ , i usant (3.8) per introduir el  $\cos \theta_k$ , la desigualtat queda

$$f(x_k + \alpha_k d_k) \leq f(x_k) - C \cdot \cos^2 \theta_k \cdot \|\nabla f(x_k)\|^2.$$

Sumant aquestes desigualtats per  $k = 1, \dots, n$ , i usant que  $x_{k+1} = x_k + \alpha_k d_k$ , tenim que

$$f(x_{n+1}) - f(x_1) \leq -C \sum_{k=1}^n \cos^2 \theta_k \cdot \|\nabla f(x_k)\|^2,$$

i per tant

$$\sum_{k=1}^n \cos^2 \theta_k \cdot \|\nabla f(x_k)\|^2 \leq \frac{f(x_1) - f(x_{n+1})}{C} \leq M,$$

per una certa constant  $M$  puix que se suposa que  $f$  està acotada inferiorment. Per tant, al fer  $n \rightarrow \infty$ , surt una sèrie convergent.  $\square$

**Corollari 3.3.4.** *En la situació del Teorema 3.3.3, si a més  $\cos^2 \theta_k \geq \delta > 0$ , per algun  $\delta$ , aleshores l'algorisme fa que  $\lim_{k \rightarrow \infty} \nabla f(x_k) = 0$ .*

*Demostració.* Si la sèrie (3.7) és convergent, en particular el terme general convergeix a zero:  $\lim_{k \rightarrow \infty} \cos^2 \theta_k \|\nabla f(x_k)\|^2 = 0$ . Si el cosinus està acotat per sota lluny del zero, ha de ser el gradient que convergeixi a zero.  $\square$

Podem formular el mateix corollari posant la condició en una successió parcial dels angles, i obtindrem la mateixa conclusió per a la corresponent successió parcial dels gradients.

**Corollari 3.3.5.** *En la situació del Teorema 3.3.3, si a més  $\cos^2 \theta_{k_n} \geq \delta > 0$ , per algun  $\delta$  i successió parcial  $\{\theta_{k_n}\}_n$ , aleshores per a la mateixa parcial  $\lim_{n \rightarrow \infty} \nabla f(x_{k_n}) = 0$ .*  $\square$

En la situació d'aquest últim corollari, ja podem considerar que l'algorisme és convergent a un punt estacionari, encara que només sigui mitjançant una parcial. Una conseqüència d'això és que, en un algorisme en què:

- les direccions de moviment són sempre de descens;
- les cerques lineals es fan complint les condicions de Wolfe;

només que de tant en tant introduïm una iteració del mètode del gradient, podem assegurar que l'algorisme convergeix a un punt estacionari.

Veurem ara com trobar un pas  $\alpha_k$  complint les condicions de Wolfe. Observem primer que donat un  $\alpha > 0$ , pot passar exactament una d'aquestes tres coses:

a) Es compleixen les condicions de Wolfe:

$$\begin{aligned}\phi(\alpha) &\leq \phi(0) + \varepsilon\phi'(0)\alpha \\ \phi'(\alpha) &\geq \eta\phi'(0)\end{aligned}$$

b) Hem anat massa lluny i Wolfe-1 no es compleix:

$$\phi(\alpha) > \phi(0) + \varepsilon\phi'(0)\alpha$$

c) No estem prou lluny i Wolfe-2 no es compleix:

$$\begin{aligned}\phi(\alpha) &\leq \phi(0) + \varepsilon\phi'(0)\alpha \\ \phi'(\alpha) &< \eta\phi'(0)\end{aligned}$$

Comencem establint els números  $\alpha_{\min} := 0$ , i  $\alpha_{\max}$  indeterminat (teòricament  $+\infty$ ). Quan passem pel cas b), posem  $\alpha_{\max} := \alpha$ . Quan passem pel cas c), posem  $\alpha_{\min} := \alpha$ .

Quan hàgim passat alguna vegada pel cas b), ja tindrem un *bracket*  $[\alpha_{\min}, \alpha_{\max}]$ , dins el qual busquem un bon  $\alpha$ . Podem provar un nou  $\alpha$  de manera simple, com ara el punt mig  $\alpha = \frac{\alpha_{\min} + \alpha_{\max}}{2}$ , o d'alguna altra manera, i tornar a mirar en quina de les tres situacions estem amb el nou  $\alpha$ . Quan arribem al cas a), haurem acabat.

Un cop tenim el bracket, segur que acabarem trobant un bon  $\alpha$  si per exemple aquest es va reduint a la meitat a cada pas, perquè sabem que hi ha tot un interval de punts que compleixen les condicions de Wolfe.

Si no tenim encara el bracket, el primer que hem de fer és anar augmentant  $\alpha$  de manera consistent. Si la funció és acotada inferiorment, per força acabarem trobant un  $\alpha_{\max}$  (vegeu altre cop la Figura 3.2 i imagineu la situació).

**Exercici 3.3.6.** Considereu la mateixa funció de l'Exercici 3.2.3, i la direcció de descens de més gran pendent des del punt  $(0, 2, 1)$  que heu trobat a l'apartat b). Demostreu que tots els  $\alpha$  de l'interval  $[0.05, 0.99]$  compleixen les condicions de Wolfe, amb  $\varepsilon = 0.01$  i  $\eta = 0.9$ . Demostreu que tots els  $\alpha$  de l'interval  $[0.05, 0.95]$  compleixen les condicions de Wolfe fortes amb els mateixos  $\varepsilon$  i  $\eta$ .  $\square$

### 3.4. Mètodes quasi-Newton: BFGS

Tornant a la fórmula iterativa genèrica (3.2), la idea dels mètodes quasi-Newton és buscar una matriu simètrica definida positiva  $B_k$  que s'assembli a la matriu Hessiana, però que no necessiti calcular derivades segones, que sigui fàcil d'actualitzar d'iteració en iteració, i que proporcioni un ordre de convergència millor que el descens de més gran pendent, que és lineal, i encara que no sigui tan bo com Newton, que és quadràtic.

Dels diversos mètodes quasi-Newton, el que es considera millor actualment és el de Broyden-Fletcher-Goldfarb-Shanno (BFGS), que tenim programat a la GSL.

Quan tinguem  $B_k$ , la usarem per trobar la direcció de descens, amb

$$d_k = -B_k^{-1}\nabla f(x_k).$$

Després es busca un pas  $\alpha_k$  complint les condicions de Wolfe, per obtenir finalment el punt següent:

$$x_{k+1} = x_k + \alpha_k d_k .$$

Anem a veure com es calculen les matrius  $B_k$  en el mètode BFGS. En la iteració inicial, podem agafar com a  $B_1$  la matriu identitat, si no tenim altre informació rellevant, o la matriu Hessiana, si no és difícil de calcular en el punt inicial  $x_1$ .

Escrivim la fórmula de Taylor d'ordre 1 per al gradient de  $f$  en el punt  $x_k$ , avaluada en  $x_{k+1}$ :

$$\nabla f(x_{k+1}) = \nabla f(x_k) + \nabla^2 f(x_k) \cdot (x_{k+1} - x_k) + o(\|x_{k+1} - x_k\|)$$

(surt d'aplicar la típica fórmula de Taylor d'una funció real, component a component).

Té sentit per tant fer l'aproximació

$$\nabla^2 f(x_k) \cdot (x_{k+1} - x_k) \approx \nabla f(x_{k+1}) - \nabla f(x_k)$$

(observeu l'analogia amb el quocient incremental en una variable:  $f''(x_k) \approx \frac{f'(x_{k+1}) - f'(x_k)}{x_{k+1} - x_k}$ .)

Buscarem una matriu  $B_{k+1}$  que compleixi aquesta relació amb igualtat:

$$B_{k+1} \cdot (x_{k+1} - x_k) = \nabla f(x_{k+1}) - \nabla f(x_k)$$

Si denotem, per abreviar,

$$\begin{aligned} s_k &:= x_{k+1} - x_k \\ y_k &:= \nabla f(x_{k+1}) - \nabla f(x_k) \\ H_{k+1} &:= B_{k+1}^{-1} , \end{aligned}$$

estem demanant que

$$H_{k+1} y_k = s_k . \tag{3.10}$$

A més volem que  $H_{k+1}$  sigui simètrica i definida positiva. Existeix una matriu que compleixi aquestes tres propietats? En principi hi ha molts coeficients per escollir, però definida positiva implica, multiplicant (3.10) per l'esquerra per  $y_k$ , que

$$y_k s_k = y_k H_{k+1} y_k > 0$$

En general el producte escalar  $y_k s_k$  no té per què ser positiu, però veurem més avall que si el pas de  $x_k$  a  $x_{k+1}$  s'ha fet complint les condicions de Wolfe, aleshores és cert. Suposant això, encara queda una certa llibertat per escollir la matriu  $H_{k+1}$ .

El mètode BFGS proposa

$$H_{k+1} = \left( I - \frac{s_k y_k^\top}{y_k^\top s_k} \right) H_k \left( I - \frac{y_k s_k^\top}{y_k^\top s_k} \right) + \frac{s_k s_k^\top}{y_k^\top s_k} , \tag{3.11}$$

on  $H_k$  és la matriu del pas anterior,  $I$  és la matriu identitat, els vectors són columnes, i  $^\top$  denota transposició. És fàcil comprovar que si  $H_k$  és simètrica i definida positiva, i suposant  $y_k^\top s_k > 0$ , aleshores  $H_{k+1}$  també és simètrica i definida positiva.

Com veiem a la fórmula (3.11), és molt fàcil calcular la nova matriu a partir de l'anterior, usant només primeres derivades de la funció  $f$ .

**Proposició 3.4.1.** *En la situació anterior, si el  $\alpha_k$  de  $x_{k+1} = x_k + \alpha_k d_k$  compleix les condicions de Wolfe, aleshores la matriu  $H_{k+1}$  de (3.11) és definida positiva.*

*Demostració.* Hem de veure en primer lloc que el producte escalar de  $y_k = \nabla f(x_{k+1}) - \nabla f(x_k)$  per  $s_k = x_{k+1} - x_k$  és positiu.

Wolfe-2 implica

$$\begin{aligned} \nabla f(x_k + \alpha_k d_k) \cdot d_k &\geq \eta \nabla f(x_k) \cdot d_k \\ \Rightarrow \nabla f(x_{k+1}) \cdot s_k &\geq \eta \nabla f(x_k) \cdot s_k \\ \Rightarrow (\nabla f(x_{k+1}) - \nabla f(x_k)) \cdot s_k &\geq \eta \nabla f(x_k) \cdot s_k - \nabla f(x_k) \cdot s_k \\ \Rightarrow y_k s_k &\geq (\eta - 1) \nabla f(x_k) \cdot s_k = (\eta - 1) \nabla f(x_k) \cdot \alpha_k d_k > 0, \end{aligned}$$

la última desigualtat perquè tant  $(\eta - 1)$  com  $\nabla f(x_k) \alpha_k d_k$  són negatius, gràcies a que  $d_k$  és direcció de descens.

Ara, si  $z$  és un vector qualsevol diferent de zero, multiplicant (3.11) per  $z$  per l'esquerra i per la dreta,

$$\begin{aligned} z^\top H_{k+1} z &= z^\top \left( I - \frac{s_k y_k^\top}{y_k^\top s_k} \right) H_k \left( I - \frac{y_k s_k^\top}{y_k^\top s_k} \right) z + \frac{z^\top s_k s_k^\top z}{y_k^\top s_k} \\ &= \left[ z \left( I - \frac{y_k s_k^\top}{y_k^\top s_k} \right) \right]^\top H_k \left[ z \left( I - \frac{y_k s_k^\top}{y_k^\top s_k} \right) \right] + \frac{\|z^\top s_k\|^2}{y_k^\top s_k}. \end{aligned}$$

El primer sumand és estrictament positiu perquè  $H_k$  es definida positiva i estem multiplicant a esquerra i dreta pel mateix vector; el segon sumand és més gran o igual que zero altre cop perquè  $y_k^\top s_k > 0$ , gràcies a les condicions de Wolfe.  $\square$

**Exercici 3.4.2.** Demostreu que si  $f$  és estrictament convexa, aleshores per qualssevol dos punts  $x$  i  $y$  es té

$$(\nabla f(x) - \nabla f(y)) \cdot (x - y) > 0$$

Per tant, en aquest cas, no caldrien les condicions de Wolfe per assegurar l'existència de la matriu definida positiva  $H_{k+1}$  del raonament anterior.

Idea: Useu el Lemma 1.2.2 amb desigualtats estrictes, amb els punts  $x$  i  $y$  i després intercanviant els seus papers.  $\square$

**Exercici 3.4.3.** Si apliquem el mètode BFGS en dimensió 1, com queda la fórmula 3.11 (posant  $\alpha_k = 1$ )? On l'hem vista abans?  $\square$

Quant a la convergència del mètode, donant per suposat que usem les condicions de Wolfe per a la cerca lineal, només cal demostrar que el cosinus de l'angle entre  $-\nabla f(x_k)$  i  $d_k$  es manté eventualment lluny del zero per tenir la convergència  $\nabla f(x_k) \rightarrow 0$  quan  $k \rightarrow \infty$  aplicant el Teorema de Zoutendijk. Efectivament, es pot demostrar d'aquesta manera que almenys una successió parcial compleix això. La demostració és complicada perquè s'ha d'anar veient com canvia la matriu  $H_k$  a cada iteració.

Respecte l'ordre de convergència, es demostra que si en el límit  $x^*$  la Hessiana  $\nabla^2 f(x^*)$  és definida positiva, i es respecten les condicions de Wolfe amb  $\varepsilon < 1/2$ , aleshores la convergència és superlineal.

### 3.5. Mètodes de direccions conjugades

Els mètodes de direccions conjugades implementen una idea diferent als quasi-Newton. És interessant veure primer el cas particular en què  $f$  és una funció quadràtica.

Una funció quadràtica  $f: \mathbb{R}^n \rightarrow \mathbb{R}$  té la forma

$$f(x) = \frac{1}{2} x A x - b x, \quad x \in \mathbb{R}^n, \quad (3.12)$$

on  $A$  és una matriu simètrica, que suposarem definida positiva, el que ens assegura que  $f$  és estrictament convexa i té un únic mínim. Podríem afegir un terme independent  $c \in \mathbb{R}$ , però això no canvia res del problema d'optimització.

Calculant el gradient i igualant-lo a zero, obtenim el sistema d'equacions lineals

$$Ax = b .$$

**Exercici 3.5.1.** Comproveu que el mètode de Newton aplicat a aquest problema quadràtic ens ve a dir que resollem el sistema lineal  $Ax = b$  invertint la matriu  $A$ .  $\square$

Un sistema lineal es pot resoldre de moltes maneres. Per exemple, eliminació Gaussiana és segurament la millor possibilitat en dimensions petites. Veurem com resoldre'l amb el mètode de direccions conjugades.

**Definició 3.5.2.** Si  $A$  és una matriu definida positiva, un conjunt de vectors no nuls  $\{d_1, \dots, d_m\}$  s'anomena *conjugat respecte A* si

$$d_i A d_j = 0 , \quad \text{si } i \neq j .$$

$\square$

**Exercici 3.5.3.** Demostreu que si  $\{d_1, \dots, d_m\}$  són conjugats respecte  $A$ , aleshores són linealment independents.  $\square$

**Teorema 3.5.4.** Si  $\{d_1, \dots, d_n\}$  és un conjunt de direccions conjugades,  $x_1 \in \mathbb{R}^n$ , i fem

$$x_{k+1} = x_k + \alpha_k d_k$$

amb  $\alpha_k$  el mínim exacte en la cerca lineal, es convergeix al mínim  $x^*$  de la funció (3.12) en  $n$  iteracions com a màxim.

*Demostració.* Per l'Exercici 3.5.3, els vectors  $\{d_1, \dots, d_n\}$  formen base. Podem escriure

$$x^* - x_1 = \sigma_1 d_1 + \dots + \sigma_n d_n \tag{3.13}$$

per certs coeficients  $\sigma_1, \dots, \sigma_n$ . L'objectiu és veure que aquests coeficients són precisament els  $\alpha_k$  que resulten de fer la minimització exacta en la direcció  $d_k$ , i ja haurem vist que partint de  $x_1$  i sumant successivament  $\alpha_1 d_1, \dots, \alpha_n d_n$  s'arriba a  $x^*$ .

Multiplicant per l'esquerra per  $d_k A$ , usant la propietat de conjugació (3.5.2),

$$d_k A(x^* - x_1) = d_k A(\sigma_1 d_1 + \dots + \sigma_n d_n) = \sigma_k d_k A d_k ,$$

d'on

$$\sigma_k = \frac{d_k A(x^* - x_1)}{d_k A d_k} .$$

En aquesta expressió podem posar  $x_k$  en lloc de  $x_1$ , perquè

$$d_k A(x^* - x_1) = d_k A(x^* - x_k + x_k - x_1) = d_k A(x^* - x_k) + d_k A(x_k - x_1)$$

i l'últim sumand és zero, altre cop per la propietat de conjugació:

$$x_k - x_1 = \sigma_1 d_1 + \dots + \sigma_{k-1} d_{k-1} \Rightarrow d_k A(x_k - x_1) = 0 .$$

Ara calculem els  $\alpha_k$  òptims de la cerca lineal exacta. Si  $\phi(\alpha) = f(x_k + \alpha d_k)$ , derivant i igualant a zero obtenim

$$0 = (A(x_k + \alpha d_k) - b) d_k = (A(x_k + \alpha d_k) - A x^*) d_k ,$$

i d'aquí surt, aïllant  $\alpha$ ,

$$\alpha_k = \frac{d_k A(x^* - x_k)}{d_k A d_k} . \tag{3.14}$$

$\square$

[Observeu la Figura 5.1 de Nocedal-Wright, pàg 104: Representa les corbes de nivell d'una funció quadràtica en  $\mathbb{R}^2$ , que són el·lipses amb els eixos coincidint amb els eixos de coordenades. Això vol dir que la matriu  $A$  és diagonal. Les successives direccions conjugades són les dels eixos, i això implica que a cada iteració es fixa una coordenada correcta del punt mínim: Passant de  $x_0$  a  $x_1$  és determina la primera coordenada, i passat de  $x_1$  a  $x_2$  la segona. Si estiguéssim en dimensió superior seguiria la seqüència. En general, les el·lipses no estaran alineades amb els eixos de coordenades, però igualment a cada iteració és fixa un subespai del subespai anterior, fins que queda individualitzat el mínim en  $n$  iteracions.]

Veiem ara com trobar les direccions conjugades. Seguim amb el cas de  $f$  quadràtica. Una possibilitat és calcular els vectors propis de  $A$ , que són conjugats respecte  $A$ , però això és molta feina, equivalent a fer eliminació Gaussiana en el sistema  $Ax = b$ .

Denotem

$$r(x_k) := Ax_k - b ,$$

que anomenarem el *residu*, perquè és el que “sobra” per arribar a la solució del sistema. En realitat no és més que el gradient  $\nabla f(x_k)$  en el punt  $x_k$ . Observem que els  $\alpha_k$  òptims de (3.14) es poden expressar també

$$\alpha_k = \frac{-r(x_k)d_k}{d_k A d_k} .$$

Suposant que ens hem mogut per direccions conjugades fins la  $d_{k-1}$ , imposarem que  $d_k$  tingui la forma

$$d_k = -r(x_k) + \beta_k d_{k-1} ,$$

amb  $\beta_k$  uns escalars a determinar de forma que  $d_{k-1}$  i  $d_k$  siguin conjugades. És fàcil trobar els  $\beta_k$ :

$$0 = d_{k-1} A d_k = -d_{k-1} A r(x_k) + \beta_k d_{k-1} A d_{k-1} ,$$

i d'aquí

$$\beta_k = \frac{d_{k-1} A r(x_k)}{d_{k-1} A d_{k-1}} .$$

Amb aquest procediment,  $d_k$  serà automàticament conjugada de totes les direccions anteriors també (Teorema següent 3.5.5). Resumim primer l'algorisme:

$x_1$  punt inicial

$$r(x_1) = Ax_1 - b$$

$$d_1 = -r(x_1)$$

Mentre  $r(x_k) \neq 0$ , fer

$$\alpha_k = \frac{-r(x_k)d_k}{d_k A d_k}$$

$$x_{k+1} = x_k + \alpha_k d_k$$

$$r(x_{k+1}) = Ax_{k+1} - b$$

$$\beta_{k+1} = \frac{d_k A r(x_{k+1})}{d_k A d_k}$$

$$d_{k+1} = -r(x_{k+1}) + \beta_{k+1} d_k$$

**Teorema 3.5.5.** *En qualsevol iteració del algorisme anterior, es té que  $d_k A d_i = 0$ , per a tot  $i < k$ ; és a dir, el conjunt de direccions generades és conjugat respecte la matriu  $A$ . A més, els residus generats són ortogonals:  $r(x_k)r(x_i) = 0$ , per a tot  $i < k$ .*

La demostració és una mica llarga i no la farem.<sup>3</sup> Usant la ortogonalitat dels residus es pot refinar l'algorisme per estalviar algunes operacions. Podeu veure la versió refinada a la pàgina

<sup>3</sup>És el Teorema 5.3 de Nocedal-Wright.

112 de Nocedal–Wright. Essencialment, el càlcul de  $\beta_{k+1}$  es pot simplificar a

$$\beta_{k+1} = \frac{r(x_{k+1})r(x_{k+1})}{r(x_k)r(x_k)}.$$

Passem ara al cas general de funcions objectiu no quadràtiques. Hi ha dues possibilitats semblantment bones, totes dues programades a la GSL: els mètodes de Fletcher–Reeves i de Polak–Ribière. Sobre funcions quadràtiques tots dos són idèntics al que ja hem vist.

En el mètode de Fletcher–Reeves, fem la substitució natural dels residus  $r(x_k)$  pels gradients  $\nabla f(x_k)$ , i fem cerca lineal inexacta. Ens queda:

$x_1$  punt inicial

$$d_1 = -\nabla f(x_1)$$

Si estem en  $x_k$ , fer

$$\alpha_k = (\text{Wolfe})$$

$$x_{k+1} = x_k + \alpha_k d_k$$

$$\beta_{k+1} = \frac{\nabla f(x_{k+1})\nabla f(x_{k+1})}{\nabla f(x_k)\nabla f(x_k)}$$

$$d_{k+1} = -\nabla f(x_{k+1}) + \beta_{k+1}d_k$$

fins que  $\nabla f(x_k) = 0$ .

Com veiem, no calen matrius, i només cal calcular un gradient nou a cada iteració. El problema és saber si les  $d_k$  obtingudes així són direccions de descens o no. Si multipliquem la última línia del bucle per  $\nabla f(x_{k+1})$ , obtenim

$$\nabla f(x_{k+1})d_{k+1} = -\|\nabla f(x_{k+1})\|^2 + \beta_{k+1}\nabla f(x_{k+1})d_k.$$

Si la cerca lineal és exacta,  $\nabla f(x_{k+1})d_k = 0$  (vegeu la demostració de la Proposició 3.2.1), i clarament queda  $\nabla f(x_{k+1})d_{k+1} < 0$ . Si la cerca no es exacta, es pot demostrar que de tota manera surten direccions de descens si es compleixen les condicions de Wolfe fortes amb  $\eta < \frac{1}{2}$ , i que s'obté la convergència a zero dels gradients sobre una successió parcial:  $\nabla f(x_{k_n}) \rightarrow 0$ .

El mètode de Polak–Ribière defineix d'una altra manera els  $\beta_k$  i la resta és igual:

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

Està considerat experimentalment millor que Fletcher–Reeves, però en aquest cas les condicions de Wolfe fortes no garanteixen tenir direccions de descens, llevat que la funció  $f$  sigui estrictament convexa. La manera d'assegurar la convergència és fer cada cert nombre d'iteracions un “restart”, posant  $\beta_k = 0$ , que equival a fer una iteració del mètode del gradient. Com ja hem vist, això garanteix que per a una successió parcial hi ha la convergència dels gradients a zero.

En l'algorisme de Fletcher–Reeves també és convenient programar “restart” periòdics, perquè quan una direcció és “poc de descens” (gairebé ortogonal al gradient), la següent iteració també tendeix a tenir el mateix defecte, i la convergència es fa molt lenta.

### 3.6. Mètodes sense derivades: Nelder–Mead

Anàlogament als problemes de dimensió 1, podem intentar buscar mínims locals d'una funció de diverses variables sense usar derivades. El mètode bàsic és el de Nelder–Mead, també anomenat

“mètode del símplex”, però no el confongueu amb el mètode del símplex d’Optimització Lineal, que no té absolutament res a veure. Aquí el *símplex* es refereix a  $n+1$  punts de  $\mathbb{R}^n$  que no estiguin continguts en un hiperplà (per exemple, en  $\mathbb{R}^2$ , tres punts que no estiguin alineats).

Els valors de la funció objectiu en els punts del símplex ens donen una idea de la forma de la funció a les rodalies. Segons quins siguin aquests valors, se substitueix un dels punts del símplex per un nou punt. I es continua el procediment amb el nou símplex. El tenim programat a la GSL. És un mètode del que no es pot assegurar en general que convergeixi a un punt estacionari (hi ha contraexemples rebuscats), però que acostuma a funcionar bé i s’utilitza molt.

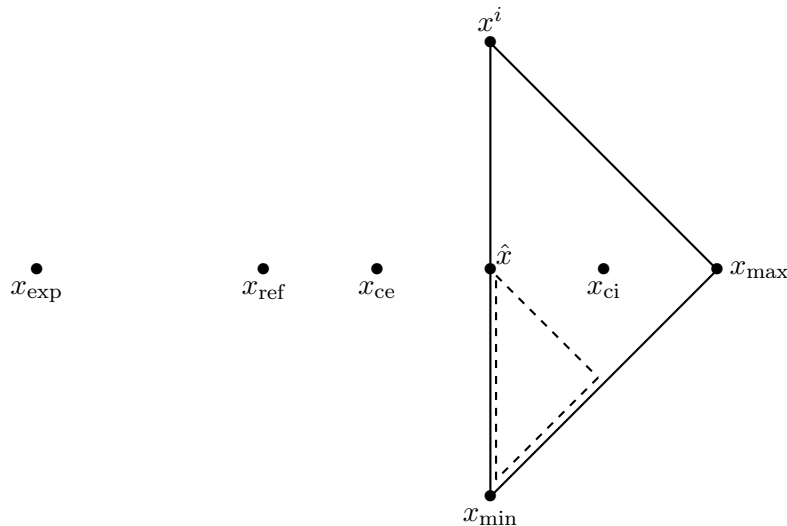


Figura 3.3: Il·lustració del mètode de Nelder–Mead en dimensió 2

Tenim un símplex de  $n+1$  punts  $\{x^0, \dots, x^n\}$  de  $\mathbb{R}^n$ .

Denotem per  $x_{\min}$  i  $x_{\max}$  el millor i el pitjor punts del símplex.

Denotem per  $\hat{x}$  el punt mig (centre de masses) de tots els punts excepte el pitjor, o sigui,

$$\hat{x} := \frac{1}{n} \left( -x_{\max} + \sum_{i=0}^n x^i \right)$$

Calculem el “punt reflectit” de  $x_{\max}$  a través de  $\hat{x}$ :

$$x_{\text{ref}} := 2\hat{x} - x_{\max}$$

Avaluem  $f$  en  $x_{\text{ref}}$  i trobem tres possibles casos: <sup>4</sup>

(a) El valor de  $f$  en  $x_{\text{ref}}$  està entre el millor valor i el segon pitjor valor del símplex:

$$f(x_{\min}) < f(x_{\text{ref}}) < \max\{f(x^i) : x^i \neq x_{\max}\}$$

Lavors canviem  $x_{\max}$  per  $x_{\text{ref}}$ .

(b) El valor de  $f$  en  $x_{\text{ref}}$  és el millor de tots:  $f(x_{\text{ref}}) < f(x_{\min})$ . Intentem una “expansió”:

$$x_{\text{exp}} := 2x_{\text{ref}} - \hat{x}$$

- Si  $f(x_{\text{exp}}) < f(x_{\text{ref}})$ , canviem  $x_{\max}$  per  $x_{\text{exp}}$ .
- Si  $f(x_{\text{exp}}) > f(x_{\text{ref}})$ , canviem  $x_{\max}$  per  $x_{\text{ref}}$ .

<sup>4</sup>Totes les desigualtats que segueixen són estrictes per claredat. En cas d’empat es pot seguir qualsevol de les possibilitats.



(c) El valor de  $f$  en  $x_{\text{ref}}$  és pitjor que el segon pitjor valor del símplex:

$$f(x_{\text{ref}}) > \max\{f(x^i) : x^i \neq x_{\text{max}}\}$$

- Si  $f(x_{\text{ref}}) < f(x_{\text{max}})$ , calculem  $x_{\text{ce}} := \frac{1}{2}(x_{\text{ref}} + \hat{x})$ . Si  $f(x_{\text{ce}}) < f(x_{\text{ref}})$ , fem “contracció exterior”: Canviem  $x_{\text{max}}$  per  $x_{\text{ce}}$ .
- Si  $f(x_{\text{ref}}) > f(x_{\text{max}})$  calculem  $x_{\text{ci}} := \frac{1}{2}(x_{\text{max}} + \hat{x})$ . Si  $f(x_{\text{ci}}) < f(x_{\text{max}})$ , fem “contracció interior”: Canviem  $x_{\text{max}}$  per  $x_{\text{ci}}$ .
- Si no s’ha pogut fer contracció exterior o interior, canviem cada  $x^i$  per  $\frac{1}{2}(x_{\text{min}} + x^i)$ .



---



---

## Optimització amb restriccions a $\mathbb{R}^n$

Com podeu comprovar fullejant el Nocedal-Wright, Capítol 15 i següents, les restriccions augmenten força la complexitat d'un problema d'optimització. Hi ha idees molt diverses per dissenyar algorismes d'optimització amb restriccions, i es fa difícil fins i tot trobar una manera de classificar-les.

Hi ha mètodes que funcionen bé quan les restriccions són només d'igualtat, altres que tenen sentit només amb restriccions de desigualtat, i també hi ha maneres de convertir una igualtat en dues desigualtats ( $h(x) = 0$  equival evidentment a  $h(x) \leq 0$  i  $-h(x) \leq 0$ ), o convertir una desigualtat difícil en una igualtat més una desigualtat fàcil ( $g(x) \leq 0$  equival a  $g(x) + x_{n+1} = 0$ ,  $x_{n+1} \geq 0$ , amb  $x_{n+1}$  una nova variable).

Potser en algunes situacions és poden eliminar variables per substitució directa, però *això és perillós*, i cal estar segurs que es pot fer. A la llista de problemes extra que teniu sobre el Capítol 1 hi ha un exercici que il·lustra que en general no es pot fer.

[L'apartat 15.3, pàg 426, del Nocedal-Wright comença amb un exemple d'una funció objectiu qualsevol i restriccions lineals, que es pot reduir d'un problema de quatre variables amb dues restriccions a un problema de dues variables sense restriccions. És clau que les dues restriccions siguin lineals, i que no hi hagi desigualtats.

En canvi l'exemple següent (el 15.2, pàgina 427) ens mostra un cas en què seria erroni fer substitucions.]

Veurem molt per sobre unes idees sobre com resoldre un problema d'optimització no lineal amb restriccions. A la GSL no hi ha algorismes per aquest tipus de problema. Hi ha una altra biblioteca de funcions en C, la NLOpt, dedicada només a optimització, que sí en té.

Una idea natural és deixar que l'algorisme pugui passar per punts no factibles, i intentar reduir a la vegada la funció objectiu i la violació de les restriccions. Això ens porta a considerar *funcions de mèrit*. Per exemple, si tenim el problema

$$\begin{aligned} & \min f(x) \\ & \text{subjecte a: } \begin{cases} h(x) = 0 \\ g(x) \leq 0 \end{cases} \end{aligned} \quad (4.1)$$

amb  $f: \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $g: \mathbb{R}^n \rightarrow \mathbb{R}^p$ , considerem la funció

$$\Phi(x; \mu) := f(x) + \mu \sum_{i=1}^m |h_i(x)| + \mu \sum_{i=1}^p [g_i(x)]^+, \quad (4.2)$$

on  $[z]^+ = \max\{0, z\} \geq 0$  és la *part positiva* de  $z$ .

La funció  $\Phi$  combina la funció  $f$  amb una penalització sobre cada restricció que no es compleixi. Es d'esperar que si anem trobant el mínim de  $\Phi$  (sense restriccions), amb un paràmetre de penalització  $\mu$  cada cop més gran, els mínims successius vagin violant menys les restriccions i al final convergeixin al mínim del problema original.

**Teorema 4.0.1.** *Si  $x^*$  és un mínim local estricte de (4.1), existeix un  $\bar{\mu}$  tal que  $x^*$  és un mínim local de (4.2) per a tot  $\mu > \bar{\mu}$ .*

Aquest teorema demostra que si augmentem suficientment la penalització, en algun moment l'optimització de  $\Phi$  coincidirà amb la de  $f$ . El que no es pot saber a priori és fins on cal augmentar la penalització. A més, la funció  $\Phi$  no és derivable, degut al valor absolut i a la part positiva; per tant, no li podem aplicar mètodes d'optimització sense restriccions que involucrin derivades.

**Exemple 4.0.2.** Considereu l'Exemple 1.3.12: Minimitzar  $f(x_1, x_2) = x_1 + x_2$  sobre la circumferència  $x_1^2 + x_2^2 - 2 = 0$ . La funció de mèrit és

$$\Phi(x; \mu) := x_1 + x_2 + \mu|x_1^2 + x_2^2 - 2| \quad (4.3)$$

Es pot comprovar que quan  $\mu > 0.5$ , el mínim de  $\Phi$  (sense restriccions) coincideix amb el mínim del problema original. Sabem que aquest mínim està en  $(-1, -1)$ . Observeu la Figura 17.4 de Nocedal-Wright, amb les corbes de nivell de  $\Phi$  i localitzeu el mínim.  $\square$

Per poder utilitzar mètodes amb derivades, necessitem penalitzar de manera que  $\Phi$  sigui més suau. Podem penalitzar amb termes quadràtics:

$$\Phi(x; \mu) := f(x) + \mu \sum_{i=1}^m h_i(x)^2 + \mu \sum_{i=1}^p ([g_i(x)]^+)^2 \quad (4.4)$$

Les segones derivades són encara discontinües, però la primera existeix i és contínua, i per tant és possible usar un mètode sense restriccions que involucri primeres derivades.

En el cas de (4.2), podem descriure un algorisme marc de la manera següent:

Decidir un valor inicial de  $\mu_0 > 0$ , una tolerància  $\tau > 0$ , i un punt inicial  $x_{0,0}$ ;

Fer, per  $k = 0, 1, 2, \dots$ ,

Trobar  $x_k$  mínim de  $\Phi(x; \mu_k)$ , començant en  $x_{k,0}$ ;

Si  $\sum_{i=1}^m |h_i(x)| + \sum_{i=1}^p [g_i(x)]^+ < \tau$ , donar resultat final  $x_k$ ; PARAR

Escollir un nou paràmetre de penalització  $\mu_{k+1} > \mu_k$

Escollir nou punt inicial  $x_{k+1,0}$

Aquesta és la idea general, però hi ha molts detalls a precisar: Com s'escull la seqüència de  $\mu$  (per exemple, es pot agafar  $\mu_{k+1}$  com l'anterior  $\mu_k$  multiplicat per una constant, com ara 5 o 10); com s'escullen els punts inicials  $x_{k,0}$  (una idea natural és agafar el mínim del problema anterior); fins a quin punt hem de trobar un mínim molt aproximat de cada problema, o passar ràpid a la iteració següent (sobre tot en els primers cicles), etc. La tolerància  $\tau$  també es pot anar reduint a cada cicle.

Per al cas de (4.4), el mateix esquema serveix, amb la possibilitat addicional que si s'utilitza un mètode amb derivades, el criteri de parada en cada cicle pot està basat en una tolerància sobre el gradient, tipus  $\|\nabla_x \Phi(x; \mu_k)\| < \tau_k$ , on  $\tau_k$  és una successió de toleràncies decreixent. No es pot assegurar un resultat idèntic al Teorema 4.0.1, però sí que hi ha convergència a un punt estacionari del problema original, quan la successió de toleràncies tendeix a zero i la successió de penalitzacions tendeix a infinit.

**Exemple 4.0.3.** Tornant a l'exemple bàsic 1.3.12, amb penalització quadràtica tindrem la funció de mèrit

$$\Phi(x; \mu) := x_1 + x_2 + \mu(x_1^2 + x_2^2 - 2)^2$$

Les corbes de nivell d'aquesta funció estan a les Figures 17.1 i 17.2 del Nocedal-Wright. Observeu que les corbes de nivell quasi circulars prop del centre de la figura assenyalen un màxim local; no un mínim.  $\square$

**Exercici 4.0.4.** Per una funció objectiu que, considerada sobre tot  $\mathbb{R}^n$ , no estigui acotada inferiorment, pot ser que la funció de mèrit  $\Phi$  segueixi essent no acotada si la penalització és massa petita.

Considerem la minimització de  $-5x_1^2 + x_2^2$ , subjecte a  $x_1 = 1$ . Resolent les condicions necessàries de Karush-Kuhn-Tucker, comproveu que el mínim està en  $(1, 0)$ . Escriviu la funció  $\Phi$  amb penalització quadràtica i comproveu que per  $\mu$  per sota d'un cert valor (quin?) el problema segueix essent no acotat. En aquests casos es veuria una ràpida divergència en l'optimització de  $\Phi$  sense restriccions, i caldria aturar-la i augmentar la  $\mu$ .  $\square$

Suposem ara que només tenim restriccions d'igualtat.

Es pot veure que la violació d'una restricció  $h_i(x) = 0$  en el punt òptim  $x_k$  obtingut amb la penalització  $\mu_k$  en el mètode de penalització quadràtica és  $h_i(x_k) \approx \lambda_i^*/\mu_k$  (aproximació de primer ordre), on  $\lambda_i^*$  és el multiplicador de Lagrange de la restricció en el punt òptim del problema original.

El mètode del Lagrangiana Augmentat es basa en aquesta observació. Es tracta de introduir aquest error sistemàtic en la funció  $\Phi$  per eliminar-lo. Els valors de  $\lambda_i^*$  no es coneixen; però suposem que en tenim una estimació  $\lambda_i$ . Podem intentar minimitzar sense restriccions

$$\Phi(x, \lambda; \mu) := f(x) + \sum_{i=1}^m \lambda_i h_i(x) + \mu \sum_{i=1}^m h_i(x)^2 \quad (4.5)$$

Aquesta funció és com la que apareix en els teoremes de condicions necessàries i suficients del Capítol 1, però amb el terme de penalització quadràtica addicional.

Suposem que a la iteració  $k$ -èsima, corresponent a minimitzar (4.5) amb penalització  $\mu_k$  i estimació dels multiplicadors  $\lambda_i^k$ , hem obtingut el punt òptim  $x_k$ . Es pot demostrar que la violació de la restriccions ve ara donada per  $h(x_k) = \frac{1}{\mu_k}(\lambda_i^* - \lambda_i^k)$ . Per tant, sembla que podem millorar l'estimació dels multiplicadors de Lagrange, fent

$$\lambda_i^{k+1} = \lambda_i^k + \mu_k h_i(x_k) \quad (4.6)$$

L'algorisme ens queda:

Decidir un valor inicial de  $\mu_0 > 0$ , una tolerància inicial  $\tau_0 > 0$ , un punt inicial  $x_{0,0}$ , i una estimació inicial  $\lambda_i^0$ ;

Fer, per  $k = 0, 1, 2, \dots$ ,

Trobar  $x_k$  mínim aproximat de  $\Phi(x, \lambda^k; \mu_k)$ , començant en  $x_{k,0}$ , i acabant quan  $\|\nabla_x \Phi(x_k, \lambda^k; \mu_k)\| < \tau_k$ ;

Si la violació de les restriccions està dins una tolerància admissible, donar resultat final  $x_k$ ; PARAR

Actualitzar l'estimació dels multiplicador de Lagrange amb (4.6)

Escollir un nou paràmetre de penalització  $\mu_{k+1} > \mu_k$

Escollir nou punt inicial  $x_{k+1,0} = x_k$

Escollir nova tolerància  $\tau_{k+1}$

Aquesta mètode millora el mètode de penalització quadràtica en el sentit que existeix un  $\bar{\mu}$  a partir del qual l'òptim de  $\Phi$  coincideix amb l'òptim del problema original, i els multiplicadors de Lagrange també convergeixen. Està considerat actualment una de les millors opcions i està programat a la biblioteca NLOpt.

**Exemple 4.0.5.** Continuant amb l'exemple 1.3.12, en el mètode del Lagrangia Augmentat usarem la funció de mèrit

$$\Phi(x, \lambda; \mu) := x_1 + x_2 + \lambda(x_1^2 + x_2^2 - 2) + \mu(x_1^2 + x_2^2 - 2)^2$$

Les corbes de nivell d'aquesta funció estan a la Figures 17.5 del Nocedal-Wright. S'assemblen a les de la Figura 17.1, però amb la mateixa penalització s'obté un punt molt més pròxim a l'òptim del problema original.  $\square$

Si tenim restriccions de desigualtat, la manera més senzilla d'incorporar-les és convertir-les en igualtats mitjançant noves variables restringides a ser positives, com hem mencionat al principi del capítol. Estarem introduint noves restriccions de desigualtat, però que són fàcils de tractar en cada subproblema. També és possible mantenir dins dels subproblemes les restriccions (d'igualtat i desigualtat) que siguin lineals. Això és el que fa el mètode del Lagrangia Augmentat en la NLOpt.