Càlcul en diverses variables

Joan Verdera

Gener 2015

El que el lector té a les mans és el conjunt de notes que l'autor portava a classe de "Càlcul en diverses variables i optimització" de segon de matemàtiques durant els cursos 2013–14, 2014–15 i 2015–16. Ho vaig ordenar i polir una mica amb la intenció que l'alumne tingués una ajuda per seguir el curs. Per la integral i el càlcul vectorial seguíem de molt a prop el llibre de Marsden i Tromba, un clàssic difícil de superar.

Bellaterra, 12 de setembre de 2016.

Capítol 1

L'espai euclidià

1.1 Norma i distància euclidianes

Recordem que

$$\mathbb{R}^n = \{x = (x_1, \ldots, x_n) : x_1, \ldots, x_n \in \mathbb{R}\}$$

és un espai vectorial sobre $\mathbb{R}$. També ens el podem mirar com un conjunt de punts entre els que hi ha, com veurem, una distància. Recordem que en una variable la derivada d'una funció es defineix com el límit

$$f'(a) = \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}. \tag{1.1}$$

Quan x s'acosta al punt a, però $x \neq a$, el quocient a (1.1) té sentit i aquest quocient s'acosta, si la funció és prou bona, a un cert límit que, per definició, és $f'(a)$. El nostre primer objectiu és precisar què significa que el punt $x = (x_1, \ldots, x_n)$ de $\mathbb{R}^n$ s'acosti al punt $a = (a_1, \ldots, a_n)$ de $\mathbb{R}^n$.

La norma (euclidiana) del punt (o vector) $x = (x_1, \ldots, x_n)$ de $\mathbb{R}^n$ és

$$\|x\| = \left(\sum_{j=1}^{n} x_j^2 \right)^{1/2} \tag{1.2}$$

i s'interpreta com la longitud del vector x o com la distància del punt x a l'origen. Per $n = 2$

Figura

Per $n = 3$

Figura

Per exemple,

$$\begin{aligned} &\|(1,-2)\| = \sqrt{1+4} = \sqrt{5} && \text{a } \mathbb{R}^2, \\ &\|(1,0,-2)\| = \sqrt{5} && \text{a } \mathbb{R}^3, \\ &\|(-1,1,-1,1)\| = \sqrt{4} = 2 && \text{a } \mathbb{R}^4. \end{aligned}$$

La norma (1.2) té les tres propietats següents:

1. $\|x\| \geq 0$ i $\|x\| = 0 \Leftrightarrow x = 0$.
2. $\|\lambda x\| = |\lambda| \|x\|$, $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.
3. $\|x + y\| \leq \|x\| + \|y\|$.

La propietat 2 dóna la homogeneïtat de la norma (diu com els escalars surten fora) i és conseqüència òbvia de la definició.

La propietat 3 s'anomena *la desigualtat triangular* i simplement diu

Figura

que un costat d'un triangle és més petit que la suma dels altres dos. La desigualtat triangular per la norma euclidiana no és immediata i requereix una certa feina. S'anomena la desigualtat de Minkowski.

Demostració de la desigualtat de Minkowski. Prenent el quadrat dels dos membres obtenim la desigualtat equivalent

$$\|x + y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\|.$$

Recordem que el producte escalar a $\mathbb{R}^n$ és

$$\langle x, y \rangle = \sum_{j=1}^{n} x_j y_j.$$

Doncs $\|x\|^2 = \langle x, x \rangle$. Per tant

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \langle x, x \rangle + 2\langle x, y \rangle + \langle y, y \rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y \rangle.$$

La demostració queda doncs reduïda a veure que

$$|\langle x, y\rangle| \leq \|x\|\|y\| \tag{1.3}$$

que s'anomena *desigualtat de Schwarz* (o Cauchy-Schwarz). Per demostrar-la notem que, per $t \in \mathbb{R}$

$$\|x+ty\|^2 = \langle x+ty, x+ty\rangle = \|x\|^2 + 2\langle x, y\rangle t + \|y\|^2 t^2.$$

Com que el polinomi de segon grau en t de la dreta és sempre no-negatiu el discriminant és negatiu o zero:

$$(2\langle x, y\rangle)^2 - 4\|x\|^2\|y\|^2 \leq 0,$$

que és (1.3). □

Hi ha un argument més directe. Donats els vectors x, y posem-nos en el pla que generen (si són linealment depenents la desigualtat és molt fàcil). Podem aplicar el teorema del cosinus al triangle que té per costats x, y i $y-x$. Obtenim

$$\|y-x\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos\theta,$$

on θ és l'angle determinat per x i y. Per altra banda

$$\|y-x\|^2 = \|x\|^2 + \|y\|^2 - 2\langle x, y\rangle,$$

i, per tant,

$$\langle x, y\rangle = \|x\|\|y\|\cos\theta,$$

que, òbviament, dóna la desigualtat de Schwarz. Hi ha un detall, que hem obviat, i que deixem al lector: cal veure que la norma restringida al pla generat per x i y és la norma pròpia d'aquest pla (per la qual val el teorema del cosinus).

Definició. Donats dos punts $x, y \in \mathbb{R}^n$ la distància euclidiana entre x i y és

$$d(x, y) = \|x-y\|.$$

Aquesta funció $d\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ té les propietats:

1. $d(x, y) \geq 0$ i $d(x, y) = 0 \Leftrightarrow x = y$.

2. $d(x, y) \leq d(x, z) + d(z, y)$, $x, y, z \in \mathbb{R}$, que és la desigualtat triangular.

Per exemple,

$$d\left(\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right),(0,0)\right) = \left\|\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)\right\| = 1,$$
$$d((1,-1,1),(0,1,0)) = \|(1,-2,1)\| = \sqrt{6}.$$

Hi ha una noció abstracta d'espai mètric, que no estudiarem en aquest curs però que cal saber. Un espai mètric és un conjunt E dotat d'una distància d. Una distància d a E és una aplicació $d\colon E \times E \to \mathbb{R}$ tal que

1. $d(x,y) \geq 0$, $x,y \in E$ i $d(x,y) = 0$ si i només si $x = y$.
2. $d(x,y) \leq d(x,z) + d(z,y)$, $x,y,z \in E$.

Per exemple, tot subconjunt de $\mathbb{R}^n$ amb la distància euclidiana és un espai mètric.

FALTA COLOR; FALTEN PROBLEMES

Problemes de la secció 1.1

1. Hi ha una altra demostració de la desigualtat de Schwarz, que és interessant perquè mostra realment la seva naturalesa. Fem-ho per $n = 2$ perquè l'argument en més dimensions és el mateix. Feu primer el cas particular en que $x = (1,0)$. Després reduiu el cas general a aquest fent una rotació. Haureu de notar que una aplicació lineal que conserva la norma també conserva el producte escalar.

1.2 Límits de successions i completesa de $\mathbb{R}^n$

La definició de límit d'una successió $(x_m)_{m=1}^\infty$ de punts de $\mathbb{R}^n$ és molt natural tenint present la noció de distància. Diem que el límit de la successió $(x_m)_{m=1}^\infty$ és $x \in \mathbb{R}^n$, cosa que s'escriu $x = \lim\limits_{m\to\infty} x_m$, sempre que la distància entre x_m i x tendeixi a 0 quan $m \to \infty$, és a dir, sempre que $\|x_m - x\| \xrightarrow{m\to\infty} 0$. Si volem, això es pot escriure amb ε i n_0: donat $\varepsilon > 0$ hi ha un índex m_0 tal que $\|x_m - x\| < \varepsilon$ si $m \geq m_0$.

Per exemple, a $\mathbb{R}^3$,

$$\lim_{m\to\infty}(\frac{m-1}{m+1}, e^{-m}m^3, \log\frac{m^4-2}{m^4+m+3}) = (1,0,0).$$

Això és obvi si se sap que el límit d'una successió de vectors a $\mathbb{R}^n$ és el límit per coordenades, és a dir, posant $x_m = (x_{m1}, x_{m2}, \cdots, x_{mn}) \in \mathbb{R}^n,\ m = 1, 2, \cdots$, si se sap que el límit de la successió $(x_m)_{m=1}^{\infty}$ existeix si i només si existeixen els límits de les successions de les coordenades $(x_{mk})_{m=1}^{\infty}$ per $1 \leq k \leq n$ i, en aquest cas,

$$\lim_{m\to\infty} x_m = (\lim_{m\to\infty} x_{m1}, \lim_{m\to\infty} x_{m2}, \cdots, \lim_{m\to\infty} x_{mn}).$$

Es pot demostrar aquest fet com un exercici (conseqüència senzilla del problema 2 del final d'aqiesta secció).

Una successió es diu convergent si té límit. Una successió convergent compleix la condició de Cauchy, exactament igual que en una variable. Si $\lim_{m\to\infty} x_m = x$ llavors donat $\varepsilon > 0$ hi ha un índex m_0 tal que per indexos $m \geq m_0$ tenim $\|x_m - x\| < \varepsilon$. Ara prenem dos indexos p i q amb $p, q \geq m_0$. Obtenim $\|x_p - x_q\| \leq \|x_p - x\| + \|x - x_q\| < 2\varepsilon$. Canviant ε per $\frac{\varepsilon}{2}$ obtenim la condició de Cauchy: per a tot $\varepsilon > 0$ hi ha un índex m_0 tal que per a indexos més avançats $p, q \geq m_0$ es té $\|x_p - x_q\| < \varepsilon$.

Teorema. $\mathbb{R}^n$ *és complet*

Demostració. La demostració és una senzilla reducció al cas de la recta. Si $(x_m)_{m=1}^{\infty}$ compleix la condició de Cauchy, llavors les successions de coordenades $(x_{mj})_{m=1}^{\infty},\ 1 \leq j \leq n$ també i, doncs, tenen límits $x_j = \lim_{m\to\infty} x_m j,\ 1 \leq j \leq n$. Per tant, el límit de $(x_m)_{m=1}^{\infty}$ és $x = (x_1, \ldots, x_n)$. □

Problemes de la secció 1.2

1. Demostreu que per $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$

$$|x_j| \leq \|x\| \leq |x_1| + \cdots + |x_n| \leq n \max_{1\leq j\leq n} |x_j|.$$

2. Espais metrics complets i no complets.

(1) Sigui $B = \{x \in \mathbb{R}^n : \|x\| < 1\}$. El conjunt B dotat de la distància euclidiana és un espai mètric. Volem veure que no és complet, és a dir que no tota successió de Cauchy és convergent. Considerem la successió

$$x_m = (1 - \frac{1}{m}, \frac{1}{m}, \ldots, \frac{1}{m}), \quad m = 1, 2, \ldots$$

Demostreu que aquesta successió és de Cauchy a B però no té límit a B.

(2) Sigui $\bar{B} = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$. Demostreu que $\bar{B}$ dotat de la distància euclidiana és un espai mètric complet.

1.3 Límits de funcions

Igual que en una variable, el tenir una distància permet definir la noció de límit d'una funció $f(x)$, definida a un subconjunt $D \subset \mathbb{R}^n$ i que prèn valors reals (o, fins i tot, a $\mathbb{R}^m$) en un punt a. Si $f : D \to \mathbb{R}^m$ direm que

$$\lim_{x \to a} f(x) = b \tag{1.4}$$

si $f(x)$ s'acosta a b quan x s'acosta al punt a sense ser a. Més formalment (1.4) és el mateix que dir que donat $\varepsilon > 0$ hi ha $\delta > 0$ tal que si $x \in D$ i $0 < \|x - a\| < \delta$ (x s'acosta al punt a i $x \neq a$), llavors $\|f(x) - b\| < \varepsilon$ ($f(x)$ s'acosta a b).

Aquesta definició només té una pega: si el punt a és "lluny" de D (mai s'ha dit que a hagi de ser de D), llavors $x \in D$ no pot acostar-se al punt a i la definició perd sentit. Per exemple, a $\mathbb{R}$, $D = [0, 1]$ i $a = 2$. A $\mathbb{R}^2$, $D = \{x : \|x\| < 1\}$ i $a = (2, 0)$. A $\mathbb{R}^3$, $D = \{x : \|x\| < 1\}$ i $a = (0, 2, 0)$.

Definició. Diem que un punt $a \in \mathbb{R}^n$ és un punt d'acumulació de $D \subset \mathbb{R}^n$ si tant a prop com vulgueu de a hi ha punts de D diferents de a. Més formalment, donat $\varepsilon > 0$ hi ha $x \in D$ tal que $0 < \|x - a\| < \varepsilon$.

Equivalentment, hi ha una successió $(x_m)_{m=1}^{\infty}$ de punts de $D \setminus \{a\}$ que té límit a.

Exemples

(1) Un punt de D que no és d'acumulació (se'n diu un punt aïllat de D).
A $\mathbb{R}$, $D = [0, 1] \cup \{2\}$ i $a = 2$. A $\mathbb{R}^n$, $D = \{x \in \mathbb{R}^n : x_n > 0\} \cup \{(0, \ldots, 0, -1)\}$ i $a = (0, \ldots, 0, -1)$.

(2) Un punt d'acumulació de D que no és de D.
A $\mathbb{R}$, $D = (0, 1)$ i $a = 0$ o $a = 1$. A $\mathbb{R}^n$, $D = \{x : \|x\| < 1\}$ i $a = (0, \ldots, 0, 1)$ o $\|a\| = 1$.

(3) Un punt d'acumulació de D que és de D.
A $\mathbb{R}^n$, $D = \{x : \|x\| \leq 1\}$ i a qualsevol punt de D.

La conclusió és que la definició que hem fet de (1.4) és interessant només quan a és un punt d'acumulació del domini D de la funció f.

Exemples

(1) Calculem a $\mathbb{R}^2$

$$\lim_{(x,y)\to(0,0)} \frac{\sin(x^2+y^2)}{x^2+y^2}.$$

Tenim que $\lim\limits_{t\to 0} \frac{\sin(t)}{t} = 1$. Donat $\varepsilon > 0$ hi ha δ tal que $0 < |t| < \delta$,

$$\left| \frac{\sin(t)}{t} - 1 \right| < \varepsilon.$$

Si $\|(x,y)\| = \sqrt{x^2+y^2} < \sqrt{\delta}$, llavors

$$\left| \frac{\sin(x^2+y^2)}{x^2+y^2} - 1 \right| < \varepsilon.$$

(2) Calculem

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}}.$$

El numerador és "quadràtic" i el denominador és "lineal". Sembla, doncs, que el límit hauria de ser 0. Fixem-nos que

$$|x| \leq \sqrt{x^2+y^2}$$

o

$$x^2 \leq x^2+y^2.$$

Per tant,

$$\frac{x^2}{\sqrt{x^2+y^2}} \leq \sqrt{x^2+y^2} = \|(x,y)\| \to 0$$

sempre que $(x,y) \to 0$.

(3) Calculem

$$\lim_{(x,y)\to(0,0)} \frac{|x|}{\sqrt{x^2+y^2}}.$$

Aquí tant numerador com denoninador són "lineals". Agafem $y = 0$

Figura

Llavors

$$\frac{|x|}{\sqrt{x^2+y^2}} = \frac{|x|}{|x|} = 1$$

i concloem que si el límit existeix ha de ser 1.

Agafem $x = 0$

Figura

La funció val 0 i, doncs, si el límit existeix ha de ser 0. Per tant el límit no existeix.

(4) Calculeu

$$\lim_{(x,y)\to(0,0)} \frac{2xy}{x^2+y^2}.$$

Provem $y = 0$. El límit és 0.

Figura

Provem $x = 0$. El límit es 0.

Figura

Provem $x = y$. El límit és $\frac{2x^2}{2x^2} = 1$. Doncs no hi ha límit.

Figura

Provem $y = \lambda x$,

$$\frac{2\lambda x^2}{(1+\lambda^2)x^2} = \frac{2\lambda}{1+\lambda^2}.$$

Figura

Així que el límit depèn de la direcció en que ens apropem a $(0, 0)$.

Compte, perquè pot succeir que els límits direccionals existeixin i siguin tots iguals però que el límit no existeixi. Veieu els problemes 4 i 5 d'aquesta secció.

També el límit es pot calcular per "successions": vegeu el problema 6.

Problemes de la secció 1.3

1. Trobeu els punts d'acumulació i els punts aïllats dels següents conjunts:

(1) $\{x \in \mathbb{R}^n : \|x\| < 1\}$.

(2) $\{(x,y) \in \mathbb{R}^2 : y > 0\} \cup \{(0,0),(0,-1)\}$.

(3) $\left\{\left(\frac{1}{n},(-1)^n\right) : n = 1,2,\ldots\right\} \subset \mathbb{R}^2$.

2. Demostreu que per $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$

$$|x_j| \le \|x\| \le |x_1| + \cdots + |x_n| \le n \max_{1\le j\le n} |x_j|.$$

3. Sigui $f\colon D \to \mathbb{R}^m$ la funció $f(x) = (f_1(x), f_2(x), \ldots, f_m(x))$. Sigui a un punt d'acumulació de D. Llavors $\lim_{x\to a} f(x) = b$ si i només si $\lim_{x\to a} f_j(x) = b_j$, $1 \le j \le m$.

4. Considereu la funció

$$f(x,y) = \frac{y\,x^2}{x^4+y^2}, \quad (x,y) \ne (0,0) \qquad \text{i} \qquad f(0,0) = 0.$$

Demostreu que tot els lḿits direccionals a l'origen són 0, però que el límit no existeix.

5. Considerem la funció definida a $\mathbb{R}^2$ que val 0 si $x^2 + (y-1)^2 \le 1$, si $x^2 + (y+1)^2 \le 1$ i si $y = 0$, i val 1 en els altres casos. Calculeu els límits direccionals a l'origen i el límit sobre la paràbola $y = x^2$.

6. Sigui a un punt d'acumulació de $A \subset \mathbb{R}^n$ i $f : A \to \mathbb{R}^m$ una funció. Llavors $\lim_{x\to a} f(x) = b$ si i només si sempre que $x_m \in A \setminus \{a\}$ i $x_m \xrightarrow{m\to\infty} a$, resulta que $f(x_m) \xrightarrow{m\to\infty} b$.

1.4 Gràfiques de funcions de dues variables

Es tracta d'estudiar alguns exemples de gràfiques de funcions senzilles en dues variables.

Sigui $z = f(x,y)$ una funció de 2 variables definida a un subconjunt $D \subset \mathbb{R}^2$. La gràfica de f és $\{(x,y,z) \in \mathbb{R}^3 : z = f(x,y) \text{ i } (x,y) \in D\}$. Si f és de n variables la gràfica és $\{(x, f(x) : x \in D\} \subset \mathbb{R}^{n+1}$. En el cas de 2 variables podem pensar que $z = f(x,y)$ és l'altura sobre el nivell del mar d'un perfil muntanyós sobre la regió D del mapa

Figura

D'aquesta manera si fixem una altura c el conjunt $\{(x, y) \in D : f(x, y) = c\}$ és una "corba de nivell", és a dir, el conjunt de punts (x, y) del mapa tals que el punt de la muntanya $(x, y, f(x, y))$ està a altura $c = f(x, y)$. Un mètode per dibuixar la gràfica de la funció $z = f(x, y)$ consisteix a trobar les corbes de nivell en el pla $x - y$,

Figura

i llavors aixecar-les a l'altura que els hi correspon.

Exemples

(1) La funció és constant, per exemple, $f(x, y) = 1$, $(x, y) \in \mathbb{R}^2$.
La "corba de nivell" d'altura 1 és tot el pla!

Figura

(2) $f(x, y) = x - y + 1$.
La corba de nivell $x - y + 1 = c$ és la recta $y = x + 1 - c$.

Figura

Aixecant cada recta a la seva altura

Figura

(3) $f(x, y) = x^2 + y^2$.
Les corbes de nivell són $x^2 + y^2 = c$, que és la circumferència de centre l'origen i radi $\sqrt{c}$ (si $c < 0$ és buit). El mapa és

Figura

i la muntanya és

Figura

La gràfica és la superfície que s'anomena "paraboloide de revolució" perquè la intersecció de la superfície amb $x = 0$ és la paràbola $z = y^2$

Figura

i amb $y = 0$ és la paràbola $z = x^2$

Figura

La superfície s'obté girant $z = y^2$ al voltant de l'eix vertical.

Problemes de la secció 1.4

1. Dibuixeu la gràfica de $f(x,y) = x^2 - y^2$ (paraboloide hiperbòlic o sella de muntar).

2. Dibuixeu la gràfica de $f(x,y) = \sqrt{x^2+y^2}$ (con).

1.5 Continuïtat

Una funció

$$f\colon D \to \mathbb{R}^m, \quad D \subset \mathbb{R}^n$$

és contínua al punt $a \in D$ si a és un punt d'acumulació de D i

$$\exists \lim_{x\to a} f(x) = f(a)$$

o si a és un punt aïllat.

f és contínua a D si ho és a cada punt de D.

Com en una variable hi ha 2 raons bàsiques per les quals una funció no és contínua a un punt:

1. Existeix $\lim_{x\to a} f(x)$ però és diferent de $f(a)$. Redefinint f en el punt a com el valor de $\lim_{x\to a} f(x)$ la funció esdevé contínua. La discontinuïtat és diu evitable. Per exemple,

$$f(x) = 0, \quad x \in \mathbb{R}^n \setminus \{0\} \quad \text{i} \quad f(0) = 1$$

   té una discontinuïtat evitable a l'origen. Definint $f(0) = 0$, la funció esdevé contínua.

2. No existeix $\lim_{x\to a} f(x)$. En aquest cas, la singularitat es diu essencial. Per exemple, considerem $f\colon \mathbb{R}^n \to \mathbb{R}$ definida per $f(x) = 0$ si $\|x\| < 1$ i $f(x) = 1$ si $\|x\| \geq 1$.

En el punts on $\|x\| = 1$ hi ha un salt.

Un altre exemple és $f\colon \mathbb{R}^2 \to \mathbb{R}$,

$$f(x,y) = \sin\left(\frac{1}{\sqrt{x^2+y^2}}\right)$$

si $(x,y) \neq (0,0)$ i $f(0,0) = 0$. A l'apropar-se a l'origen la funció oscil·la entre -1 i 1, com en el cas d'una variable $y = \sin(\frac{1}{x})$. Doncs el límit a l'origen no existeix degut a l'oscil·lació de la funció.

FALTA COLOR; FALTEN PROBLEMES

1.6 Topologia de $\mathbb{R}^n$ i compacitat

Definició. Una bola oberta de radi a i centre r és el conjunt

$$B(a,r) = \{x \in \mathbb{R}^n : \|x-a\| < r\}.$$

La bola tancada és

$$\overline{B}(a,r) = \{x \in \mathbb{R}^n : \|x-a\| \leq r\}.$$

Definició. Un conjunt $U \subset \mathbb{R}^n$ es diu obert si per tot $x \in U$ hi ha una bola centrada al punt x continguda a U:

Figura

Exemples

- $B(a,r)$ és obert i $\overline{B}(a,r)$ no ho és. Donat $b \in B(a,r)$ tenim que $B(b, r - \|b-a\|) \subset B(a,r)$ com es veu fàcilment per la desigualtat triangular. Si b és tal que $\|b-a\| = r$, llavors cap bola centrada a b està continguda a $B(a,r)$.

Figura

- El semi-espai obert $\{x \in \mathbb{R}^n : x_n > 0\}$ és un obert.
- $\{a\}$ no és obert.
- $\{(x,y) : x \geq 0,\ y > 0\}$ no és obert.
- És obert $\{(x,y) : x > 0,\ y > 0\}$?

Propietats dels oberts

1. ϕ i $\mathbb{R}^n$ són oberts.
2. Si $(U_i)_{i\in I}$ és una família arbitrària d'oberts, llavors $\bigcup\limits_{i\in I} U_i$ és un obert.
3. Si $U_1, \ldots, U_m$ són oberts $\bigcap\limits_{i=1}^{m} U_i$ és obert.

Aquestes tres propietats són les que defineixen una "topologia" a un conjunt, com es veurà en altres cursos.

Tancats. Un conjunt $F \subset \mathbb{R}^n$ és tancat si el complementari $F^c = \mathbb{R}^n \setminus F$ és obert.

Exemples

- $\overline{B}(a,r)$ és tancat.
- $\{a\}$ és tancat.
- $\{(x,y) : x \geq 0,\ y > 0\}$ no és obert ni tancat.

Adherència: Donat $F \subset \mathbb{R}^n$ qualsevol posem

$$\begin{aligned}\overline{F} &= \{x \in \mathbb{R}^n : B(x,r) \cap F \neq \emptyset,\ r > 0\} \\ &= \{x \in \mathbb{R}^n : x = \lim_{m\to\infty} x_m,\ x_m \in F\}.\end{aligned}$$

Nota. $x = \lim_{m\to\infty} x_m$ si donat $\varepsilon > 0$ hi ha m_0 tal que $\|x - x_m\| < \varepsilon$, $m \geq m_0$.

Teorema. *Tenim que són equivalents per $F \subset \mathbb{R}^n$,*

(1) F és tancat.

(2) $F = \overline{F}$.

(1) $\Rightarrow$ (2) Sigui $x \in \overline{F}$. Si $x \notin F$, com que $\mathbb{R}^n \setminus F$ és obert, $B(x,r) \subset \mathbb{R}^n \setminus F$, per cert $r > 0$. Això contradiu la definició de $x \in \overline{F}$.

(2) $\Rightarrow$ (1) Sigui $x \notin F$. Llavors $x \notin \overline{F}$ i, doncs, hi ha $r > 0$, $B(x,r) \cap F = \emptyset$ o sigui $B(x,r) \subset \mathbb{R}^n \setminus F$. Per tant, $\mathbb{R}^n \setminus F$ és obert.

Definició. Un subconjunt K de $\mathbb{R}^n$ és compacte si és tancat i fitat. 'Fitat' vol dir que és un subconjunt d'una bola. Equivalentment, hi ha una constant positiva C tal que $\|x\| \leq C$, $x \in K$.

Per exemple, una bola tancada és un compacte i una bola oberta no ho és. El quadrat

$$Q = \left\{x = (x_1, \dots, x_n) \in \mathbb{R}^n : -\frac{1}{2} \leq x_j \leq \frac{1}{2},\ 1 \leq j \leq n\right\}$$

de centre l'origen i de costat 1 és compacte. $\mathbb{R}^n$ no és compacte, ni cap semi-espai tampoc.

Veiem ara que els conjunts compactes de $\mathbb{R}^n$ són els que tenen la propietat de Bolzano-Weierstrass, és a dir, aquells pels quals tota successió d'elements del conjunt té una subsuccessió (o successió parcial) convergent a un punt del conjunt.

Teorema. *Són equivalents, per $K \cap \mathbb{R}^n$,*

(1) K és tancat i fitat.

(2) Si $(x_m)_{m=1}^{\infty}$ és una successió de punts de K, hi ha una parcial $(x_{m_k})_{k=1}^{\infty}$ i un $x \in K$ tals que $x_{m_k} \xrightarrow{k\to\infty} x$.

(1) $\Rightarrow$ (2) Posem $x_m = (x_{m1}, x_{m2}, \ldots, x_{mn})$. Com que $|x_{m1}| \leq \|x_m\| \leq c$, $m = 1, 2, \ldots$ la successió de números reals $(x_{m1})_{m=1}^{\infty}$ té una parcial convergent $(x_{m_k 1})_{k=1}^{\infty}$, diguem, envers x_1. Ara $|x_{m_k 2}| \leq \|x_{m_k}\| \leq c$, $k = 1, 2, \ldots$ Doncs hi ha una parcial de $(x_{m_k 2})_{k=1}^{\infty}$ convergent, diguem, envers x_2. Inductivament trobem una parcial $x_{m_k} \xrightarrow{k\to\infty} x = (x_1, \ldots, x_n)$. Com que $x \in \overline{K}$ i K és tancat, $x \in K$ i (2) queda vist.

(2) $\Rightarrow$ (1) Veiem primer que F és fitat. Si no ho fos, per cada m hi hauria un $x_m \in F$, $\|x_m\| > m$. Llavors $\|x_m\| \xrightarrow{m\to\infty} \infty$ i la successió $(x_m)_{m=1}^{\infty}$ no té cap parcial convergent.

Veiem ara que F és tancat. Sigui $x \in \overline{F}$ i $x_m \in F$, $x_m \xrightarrow{m\to\infty} x$. Hi ha una parcial $(x_{m_k})_{k=1}^{\infty}$ amb $x_{m_k} \xrightarrow{k\to\infty} y \in F$, per hipòtesi. Doncs $x = y \in F$.

Quan estudiem problemes d'optimització farem servir sovint el següent

Teorema. *Sigui K un compacte $\subset \mathbb{R}^n$ i $f\colon K \to \mathbb{R}$ una funció contínua. Llavors f és fitada i s'assoleixen els extrems, és a dir, hi ha punts $a, b \in K$ tals que*

$$f(a) = \sup_{x\in K} f(x) \quad i \quad f(b) = \inf_{x\in K} f(x).$$

Demostració. Si f no és fitada hi ha, per a cada m, un $x_m \in K$ tal que $|f(x_m)| > m$. Agafem una parcial $x_{m_k} \xrightarrow{k\to\infty} x$ convergent cap a $x \in K$. Com que f és contínua al punt x, $f(x_{m_k}) \xrightarrow{k\to\infty} f(x)$ i obtenim una contradicció.

Veiem ara que s'assoleix $\sup\limits_{x\in K} f(x) \equiv s$. Donat $m = 1, 2, \ldots$ agafem $x_m \in K$ amb $s - \frac{1}{m} < f(x_m) \leq s$. Passant a una parcial $x_{m_k} \xrightarrow{k\to\infty} x \in K$. Doncs $f(x) = \lim\limits_{k\to\infty} f(x_{m_k}) = s$. $\square$

Problemes de la secció 1.6

1. Un subconjunt F de $\mathbb{R}^n$ és complet si tota successió de Cauchy d'elements de F és convergent envers un punt de F. Demostreu que $\mathbb{R}^n$ és complet però que $B(0,1)$ no ho és. Demostreu que F és complet si i només si és tancat.

2. Un punt a d'un subconjunt F de $\mathbb{R}^n$ és un punt frontera de F si tota bola centrada al punt a conté punts de F i de F^c. Sigui $\mathrm{Fr}(F)$ el conjunt de punts frontera. Clarament

$\mathrm{Fr}(F) \subset \overline{F}$. Demostreu que $\mathrm{Fr}(F)$ és un tancat de $\mathbb{R}^n$ i que F és tancat si i només si $\mathrm{Fr}(F) \subset F$.

3. Un punt a d'un subconjunt F de $\mathbb{R}^n$ és un punt interior de F si hi ha una bola centrada al punt continguda a F. Sigui $\overset{\circ}{F}$ el conjunt de punts interiors de F. Llavors F és obert si i només si $\overset{\circ}{F} = F$. Per a quasevol F, $\mathrm{Fr}(F) = \overline{F} \setminus \overset{\circ}{F}$.

4. Construïu un subconjunt tancat no buit F de $\mathbb{R}^n$ amb interior buit.

Sugg.: el conjunt de punts de $\mathbb{R}^n$ amb totes les coordenades racionals és numerable i és dens (la seva adherència és $\mathbb{R}^n$). Agafeu una successió de boles centrades en punts d'aquest conjunt.

Capítol 2

Diferenciabilitat

2.1 La diferencial d'una funció

Quina és la definició correcta de "derivada" d'una funció

$$f\colon \mathbb{R}^2 \to \mathbb{R} \text{ en el punt } a = (a_1, a_2) \in \mathbb{R}\ ? \tag{2.1}$$

En una variable la definició de derivada

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a}$$

es pot interpretar dient que $f'(a)$ és el pendent de la recta tangent a la gràfica de $y = f(x)$ en el punt a. Aquesta recta tangent es defineix com $y = f(a) + f'(a)(x - a)$ i es té que

$$\lim_{x \to a} \frac{|f(x) - (f(a) + f'(a)(x - a))|}{|x - a|} = 0. \tag{2.2}$$

Tornem al cas de dues variables. Un pla no vertical pel punt $(a, f(a)) \in \mathbb{R}^3$ és de la forma

$$z = f(a) + A(x - a_1) + B(y - a_2), \tag{2.3}$$

per certs nombres reals A i B. Diem que el pla (2.3) és tangent a la gràfica de $z = f(x, y)$ en el punt $(a, f(a))$ si

$$\lim_{(x,y) \to a} \frac{|f(x, y) - [f(a) + A(x - a_1) + B(y - a_2)]|}{\|x - a\|} = 0. \tag{2.4}$$

Resulta que (2.4) determina A i B. Doncs el pla tangent si existeix és únic.

La intersecció del pla $y = a_2$ amb el pla (2.3) és la recta

$$z = f(a) + A(x - a_1)$$

i (2.4) es tranforma, per punts d'aquesta recta, en

$$\lim_{x \to a_1} \frac{|f(x, a_2) - [f(a) + A(x - a_1]|}{|x - a_1|} = 0.$$

Així que la recta $z = f(a) + A(x - a_1)$ és tangent a la corba $z = f(x, a_2)$ en el punt $(a_1, f(a))$, o sigui,

$$A = \frac{d}{dx} f(x, a_2)\Big|_{x=a_1}.$$

Anàlogament trobem

$$B = \frac{d}{dx} f(a_1, y)\Big|_{y=a_2}.$$

L'argument precedent porta a la definició següent.

Definició. La derivada parcial de f respecte de la primera variable en el punt $a = (a_1, a_2)$ és

$$\frac{\partial f}{\partial x}(a) = \frac{df}{dx}(x, a_2)\Big|_{x=a_1} = \lim_{h \to 0} \frac{f(a_1 + h, a_2) - f(a)}{h}$$

i la derivada parcial respecte de la segona variable en el punt $a = (a_1, a_2)$ és

$$\frac{\partial f}{\partial y}(a) = \frac{df}{dy}(a_1, y)\Big|_{y=a_2} = \lim_{h \to 0} \frac{f(a_1, a_2 + h) - f(a)}{h}.$$

La figura és

Figura

En general, si $f \colon U \to \mathbb{R}$, U obert $\subset \mathbb{R}^n$, és una funció de n variables i $a \in U$ es defineix la derivada parcial de f respecte de la variable x_j en el punt $a = (a_1, \ldots, a_n)$ com

$$\begin{aligned} \frac{\partial f}{\partial x_j}(a) &= \frac{df}{dt} f(a_1, \ldots, a_{j-1}, t, a_{j+1}, \ldots, a_n)\Big|_{t=a_j} \\ &= \lim_{h \to 0} \frac{f(a_1, \ldots, a_{j-1}, a_j + h, a_{j+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}. \end{aligned}$$

Exemples

(1) Posem $f(x,y) = (x+1)\sin y + \cos y \sin x$, $(x,y) \in \mathbb{R}^2$. Prop de l'origen $\sin y \simeq y$, $\sin x \simeq x$ i $\cos y \simeq 1$ així que

$$f(x,y) = (x+1)y + x + E,$$

on E és un error quadràtic. Clarament

$$f(x,y) = x + y + E$$

incorporant el terme xy a l'error. Per tant, sembla que la part lineal a l'origen és $x+y$ i que $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 1$. Això és clar perquè

$$\begin{aligned}
&\frac{\partial f}{\partial x}(x,y) = \sin y + \cos y \cos x,\\
&\frac{\partial f}{\partial x}(0,0) = 1,\\
&\frac{\partial f}{\partial y}(x,y) = (x+1)\cos y - \sin y \sin x,\\
&\frac{\partial f}{\partial y}(0,0) = 1.
\end{aligned}$$

(2) Existeixen les derivades parcials a un punt però la funció no és contínua al punt. Posem

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & (x,y) \neq (0,0),\\ 0, & (x,y) = (0,0). \end{cases}$$

Les derivades parcials a $(0,0)$ són 0 perquè la funció s'anul·la als eixos. Les derivades parcials clarament existeixen per a qualsevol (x,y) perquè el denominador no s'anul·la fora de l'origen. Sobre la recta $x=y$ la funció és

$$\frac{x^2}{2x^2} = \frac{1}{2}.$$

Per tant f no és contínua a $(0,0)$.

Sabem que un candidat a "pla tangent" a la gràfica de $z = f(x,y)$ en el punt $a = (a_1, a_2)$ és

$$z = \frac{\partial f}{\partial x}(a)(x-a_1) + \frac{\partial f}{\partial y}(a)(y-a_2) + f(a)$$

(en una variable la recta tangent és $y = f'(a)(x-a) + f(a)$).

Definició. Sigui $f\colon U \to \mathbb{R}$, U obert de $\mathbb{R}^2$ i $a \in U$. La funció f és diferenciable al punt a si existeixen les derivades parcials de f al punt a i

$$\lim_{x\to a} \frac{|f(x) - [f(a) + \frac{\partial f}{\partial x}(a)(x-a_1) + \frac{\partial f}{\partial y}(a)(y-a_2)]|}{\|x-a\|} = 0.$$

Notem que això significa, com en una variable, que la diferència entre la funció i el pla tangent és

$$f(x) - \left[f(a) + \frac{\partial f}{\partial x}(a)(x-a_1) + \frac{\partial f}{\partial y}(y-a_2) \right] = \eta(x)\|x-a\|,$$

on $\eta(x) \to 0$ si $x \to a$. Per aquesta raó es diu que

$$f(a) + \frac{\partial f}{\partial x}(a)(x-a_1) + \frac{\partial f}{\partial y}(a)(y-a_2)$$

és l'aproximació lineal de la funció $f(x)$ al voltant del punt a.

Exemple. $f(x,y) = x + y^2$ és diferenciable a $\mathbb{R}^2$. A l'origen tenim $\frac{\partial f}{\partial x}(0,0) = 1$, $\frac{\partial f}{\partial y}(0,0) = 0$ i

$$\frac{|f(x,y) - x|}{\sqrt{x^2+y^2}} = \frac{|x + y^2 - x|}{\sqrt{x^2+y^2}} = \frac{y^2}{\sqrt{x^2+y^2}} \le \frac{y^2}{|y|} = |y| \longrightarrow 0 \text{ si } (x,y) \longrightarrow (0,0).$$

Si $a = (a_1, a_2)$ és qualsevol, $\frac{\partial f}{\partial x}(a) = 1$, $\frac{\partial f}{\partial y}(a) = 2a_2$,

$$\begin{aligned}
\frac{|f(x,y) - [f(a) + (x-a_1) + 2a_2(y-a_2)]|}{\sqrt{(x-a_1)^2 + (y-a_2)^2}} &= \frac{x + y^2 - a_1 - a_2^2 - (x-a_1) - 2a_2(y-a_2)}{\sqrt{(x-a_1) + (y-a_2)^2}} \\
&= \frac{(y-a_2)^2}{\sqrt{(x-a_1)^2 + (y-a_2)^2}} \\
&\le |y-a_2| \longrightarrow 0 \text{ si } (x,y) \longrightarrow (a_1,a_2).
\end{aligned}$$

La definició de funció diferenciable d'un nombre arbitrari de variables és l'extensió òbvia de la que hem fet per dues variables. Abans introduïm una terminologia molt popular.

Definició. Sigui $f\colon U \to \mathbb{R}$, U un obert de $\mathbb{R}^n$, una funció que té derivades parcials $\partial_1 f(a) = \frac{\partial f}{\partial x_1}(a), \dots, \partial_n f(a) = \frac{\partial f}{\partial x_n}(a)$ al punt $a \in U$. El gradient de f al punt a és

$$\nabla f(a) = (\partial_1 f(a), \dots, \partial_n f(a)).$$

Així doncs $\nabla f(a)$ és un vector de $\mathbb{R}^n$.

També es pot pensar com una matriu amb una fila i n columnes. L'aplicació lineal que determina és

$$\begin{aligned} \mathbb{R}^n &\longrightarrow \mathbb{R} \\ h = (h_1, \dots, h_n) &\longrightarrow \sum_{j=1}^{n} \partial_j f(a) h_j = \langle \nabla f(a), h \rangle. \end{aligned}$$

Definició. Considerem una funció $f\colon U \to \mathbb{R}$, U un obert de $\mathbb{R}^n$. Diem que f és diferenciable al punt $a \in U$ si f té derivades parcials al punt $a \in U$ i

$$\lim_{x \to a} \frac{|f(x) - (f(a) + \langle \nabla f(a), x - a \rangle)|}{\|x - a\|} = 0.$$

La funció lineal afí

$$\ell(x) = f(a) + \sum_{j=1}^{n} \partial_j f(a)(x_j - a_j)$$

és l'aproximació lineal de f al punt a i la seva gràfica, que és un hiperplà a $\mathbb{R}^{n+1}$, és l'hiperplà tangent a f al punt $(a, f(a)) \in \mathbb{R}^{n+1}$. Heu de pensar sempre en $n = 2$, perquè per $n > 2$ no podeu dibuixar els objectes que hem definit.

Exemple. Posem $f(x) = x_1^2 + \cdots + x_n^2 = \|x\|^2$, $x \in \mathbb{R}^n$. Comprovem que és diferenciable al punt $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Clarament $\frac{\partial f}{\partial x_j}(a) = 2a_j$, $1 \le j \le n$. Tenim

$$\begin{aligned} \left| f(x) - f(a) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)(x_j - a_j) \right| &= \left| \sum_{j=1}^{n} x_j^2 - a_j^2 - 2aj(x_j - a_j) \right| \\ &= \sum_{j=1}^{n} (x_j - a_j)^2 = \|x - a\|^2. \end{aligned}$$

Dividint per $\|x - a\|$ obtenim $\|x - a\|$ que tendeix a zero. Noteu que al restar el terme lineal $\sum_j 2a_j(x_j - a_j)$ hem produït un terme quadràtic.

Ara ja podem definir diferenciabilitat per funcions vectorials de n variables reals.

Definició. Sigui $f\colon U \to \mathbb{R}^m$, U obert de $\mathbb{R}^n$, $a \in U$. La funció f és diferenciable al punt a si hi ha una funció lineal $L\colon \mathbb{R}^n \to \mathbb{R}^m$ tal que

$$\frac{\|f(x) - f(a) - L(x - a)\|}{\|x - a\|} \xrightarrow{x \to a} 0. \tag{2.5}$$

L es diu la diferencial de f al punt a i es denota per $Df(a)$.

Proposició. *Si $f(x) = (f_1(x), \dots, f_m(x))$ és diferenciable al punt a, llavors cada funció component $f_j(x)$ és diferenciable i la matriu de $Df(a)$ (en les bases canòniques de $\mathbb{R}^n$ i $\mathbb{R}^m$) és*

$$Df(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \dots & \dfrac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Demostració. Posem $L(h) = (L_1(h), \dots, L_m(h))$, $h \in \mathbb{R}^n$, de manera que $L_k \colon \mathbb{R}^k \to \mathbb{R}$ és lineal per $1 \le k \le m$. Com que f és diferenciable al punt a

$$\frac{|f_k(x) - f_k(a) - L_k(x-a)|}{\|x-a\|} \xrightarrow{x \to a} 0. \tag{2.6}$$

Agafem $x = a + he_j$, $e_j = (0, \dots, 0, 1, 0, \dots, 0)$, on 1 va a la component j-èsima. Llavors (2.6) dóna

$$\frac{|f_k(a+he_j) - f_k(a) - hL_k(e_j)|}{\|x-a\|} \xrightarrow{h \to 0} 0.$$

Doncs, f_k té derivades parcials al punt a i $\frac{\partial f_k}{\partial x_j}(a) = L_k(e_j)$ que és l'entrada de la fila j i columna k de la matriu de $L = Df(a)$. Noteu que $L_k(x-a) = \sum\limits_{j=1}^{n} \frac{\partial f_k(a)}{\partial x_j}(x_j - a_j)$ i, per tant, (2.6) diu que f_k és diferenciable al punt a. □

Exercici. $f \colon U \to \mathbb{R}^m$, $f = (f_1, \dots, f_m)$ és diferenciable al punt $a \in U$ si i només si totes les components f_k ho són.

Criteris de diferenciabilitat

1. **Suma.** Si $f, g \colon U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, són diferenciables al punt $a \in U$, llavors $f+g$ és diferenciable al punt a i $D(f+g)(a) = Df(a) + Dg(a)$.

2. **Producte.** Si $f, g \colon U \to \mathbb{R}$, $U \subset \mathbb{R}^n$, són diferenciables al punt $a \in U$, llavors fg és diferenciable al punt a i

$$D(fg)(a) = f(a)Dg(a) + g(a)Df(a).$$

Per fer la demostració, que es deixa com a exercici per al lector, es necessita el

Lema. *Si $f\colon U \to \mathbb{R}^m$, U obert $\subset \mathbb{R}^n$ és diferenciable al punt $a \in U$, llavors f és contínua al punt a.*

Demostració. $\|f(x) - f(a)\| \le \|f(x) - f(a) - Df(a)(x-a)\| + \|Df(a)(x-a)\| = I + II$.
$|I| = \frac{|I|}{\|x-a\|}\|x-a\|$ i els dos factors tendeixen a 0.

Per veure que II tendeix a 0 si $x \to a$, posem $L = Df(a)$ que és una aplicació lineal $L\colon \mathbb{R}^n \to \mathbb{R}^m$ que ve donada per una matriu (c_{jk}). Doncs

$$L(x) = \left(\sum_{j=1}^{n} c_{j1} x_j, \dots, \sum_{j=1}^{n} c_{jm} x_j \right).$$

Cada component és un polinomi de primer grau en les variables $x_1, \dots, x_n$ i, doncs, és una funció contínua a $\mathbb{R}^n$. Per tant L és contínua a $\mathbb{R}^n$ i $L(x-a) = L(x) - L(a) \to 0$ si $x \to a$. $\square$

Veiem ara una condició suficient per la diferenciabilitat d'una funció real.

Criteri de diferenciabilitat. *Si una funció $f\colon U \to \mathbb{R}$, U obert de $\mathbb{R}^n$, té derivades parcials $\partial_j f(x)$ que són contínues a U per $1 \le j \le n$, llavors f és diferenciable a U.*

Demostració. Fixem $a \in U$. La definició de diferenciabilitat és local, en el sentit que només depèn dels valors de $f(x)$ per x proper al punt a. Prenem, doncs, una bola de centre a i radi r prou petit de manera que $B(a,r) \subset U$ i fixem-nos només en punts $x \in B(a,r)$.

Posem $n = 2$, per simplificar la notació. Expressem la diferència $f(x,y) - f(a_1,a_2)$ com

$$f(x,y) - f(a_1,a_2) = f(x,y) - f(x,a_2) + f(x,a_2) - f(a_1,a_2)$$

i apliquem el teorema del valor mitjà a les funcions d'una variable

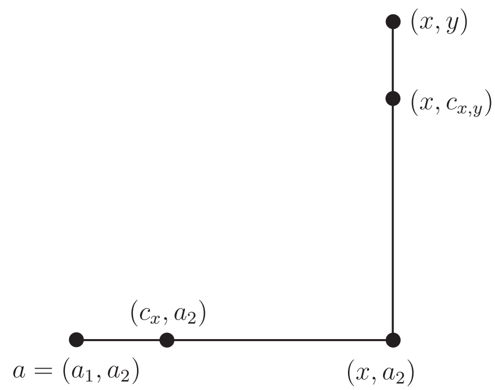
$$t \longrightarrow f(t,a_2),$$
$$t \longrightarrow f(x,t).$$

Obtenim punts c_x i $c_{x,y}$ tals que

$$f(x,a_2) - f(a_1,a_2) = \frac{\partial f}{\partial x}(c_x,a_2)(x-a_1)$$

i

$$f(x,y) - f(x,a_2) = \frac{\partial f}{\partial y}(x,c_{x,y})(y-a_2).$$



Ara

$$\begin{aligned}
&\left| f(x,y) - f(a_1,a_2) - \frac{\partial f}{\partial x}(a)(x-a_1) - \frac{\partial f}{\partial y}(a)(y-a_2) \right| \\
&\qquad \le \left| \left(\frac{\partial f}{\partial x}(c_x,a_2) - \frac{\partial f}{\partial x}(a_1,a_2) \right)(x-a_1) + \left(\frac{\partial f}{\partial y}(x,c_{x,y}) - \frac{\partial f}{\partial y}(a) \right)(y-a_1) \right| \\
&\qquad \le \left(\left| \frac{\partial f}{\partial x}(c_x,a_2) - \frac{\partial f}{\partial x}(a_1,a_2) \right| + \left| \frac{\partial f}{\partial y}(x,c_{x,y}) - \frac{\partial f}{\partial y}(a) \right| \right) \|x-a\|.
\end{aligned}$$

Si $(x,y) \to (a_1,a_2)$, llavors $x \to a_1$ i, com que c_x està entre a_1 i x, $(c_x,a_2) \to (a_1,a_2)$. Doncs

$$\left| \frac{\partial f}{\partial x}(c_x,a_2) - \frac{\partial f}{\partial x}(a_1,a_2) \right| \longrightarrow 0$$

perquè $\frac{\partial f}{\partial x}$ és contínua al punt a. També, com que $c_{x,y}$ està entre a_2 i y, $c_{x,y} \to a_2$ i, per tant, $(x,c_{x,y}) \to (a_1,a_2)$. Doncs

$$\left| \frac{\partial f}{\partial y}(x,c_{x,y}) - \frac{\partial f}{\partial y}(a_1,a_2) \right| \longrightarrow 0$$

perqué $\frac{\partial f}{\partial y}$ és contínua al punt a. □

El cas de les corbes a $\mathbb{R}^m$

Fem $n=1$ i deixem que m sigui qualsevol enter positiu. En aquest cas estem considerant funcions diferenciables

$$c\colon (a,b) \longrightarrow \mathbb{R}^m.$$

Si $m = 1$ obtenim la teoria d'una variable. Si $m = 2$ o $m > 2$, obtenim funcions d'una variable real t, que podem pensar com el temps, que prenen valors a $\mathbb{R}^m$.

La interpretació més suggestiva és que c descriu la posició d'un mòbil que a l'instant t, $a < t < b$, està a la posició $c(t) = (c_1(t), c_2(t), \dots, c_m(t))$. La definició de diferenciabilitat al punt t_0 és que hi ha una aplicació lineal $L\colon \mathbb{R} \to \mathbb{R}^m$ tal que

$$\frac{\|c(t) - c(t_0) - L(t - t_0)\|}{|t - t_0|} \xrightarrow{t \to t_0} 0. \tag{2.7}$$

Posant $v = L(1) \in \mathbb{R}^m$, tenim que $L((t - t_0)) = (t - t_0)v$ i, doncs, (2.7) és equivalent a

$$\lim_{t \to t_0} \frac{c(t) - c(t_0)}{t - t_0} = v.$$

Denotem el límit precedent per $c'(t_0)$, d'acord amb la notació unidimensional, l'anomenem derivada al punt t_0 i l'interpretem com la velocitat del nostre mòbil al temps t_0.

Exemples

(1) Recta per a i vector director $\vec{v}$.

$$c(t) = a + t\vec{v}, \quad t \in \mathbb{R}.$$

Figura

En aquest cas $c'(t) = \vec{v}$, $t \in \mathbb{R}$.

(2) Circumferència de centre a i radi r.

$$c\colon [0, 2\pi] \longrightarrow \mathbb{R}^2, \quad c(t) = a + re^{it} = (a_1 + r\cos t, a_2 + r\sin t).$$

En aquest cas la velocitat és

$$c'(t) = re^{it}i = r(\cos t + i\sin t)i = r(-\sin t + i\cos t) = r(-\sin t, \cos t).$$

Fixem-nos que $c'(t)$ és perpendicular a $c(t) - a$.

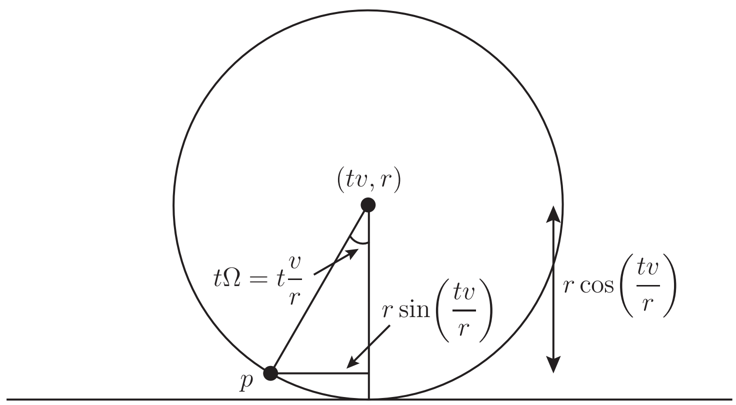
Figura

(3) L'*el·lipse* $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$.

$$\begin{aligned} c(t) &= (a\cos t, b\sin t), \quad 0 \le t \le 2\pi, \\ c'(t) &= (-a\sin t, b\cos t). \end{aligned}$$

(43) **La cicloide.** És la corba descrita per un punt d'una circumferència de radi r que es desplaça amb velocitat constant rodant sobre una superfície plana.

Deduïm l'equació de la cicloide suposant que inicialment el punt p toca el terra. El punt p es mou per la circumferència amb una velocitat angular constant Ω. Quan p torna a tocar el terra ha passat un temps $T = \frac{2\pi}{\Omega} = \frac{2\pi r}{v}$. Doncs $\Omega = \frac{v}{r}$. Al cap de t unitats de temps la situació és la següent:



El punt p sobre la cicloide és

$$(tv, r) - \left(r\sin\left(\frac{tv}{r}\right), r\cos\left(\frac{tv}{r}\right)\right).$$

Per tant

$$c(t) = \left(tv - r\sin\left(\frac{tv}{r}\right), r\left(1 - \cos\left(\frac{tv}{r}\right)\right)\right)$$

és una parametrització de la cicloide i la velocitat del punt a l'instant t és

$$\begin{aligned} c'(t) &= \left(v - v\cos\left(\frac{tv}{r}\right), v\sin\left(\frac{tv}{r}\right)\right) \\ &= v\left(1 - \cos\left(\frac{tv}{r}\right), \sin\left(\frac{tv}{r}\right)\right). \end{aligned}$$

La regla de la cadena

És la regla que diu com es deriva una composició de funcions.

Sigui $f\colon U \to \mathbb{R}^m$, U obert de $\mathbb{R}^n$ i $g\colon V \to \mathbb{R}^p$ on V és un obert de $\mathbb{R}^m$ que conté $f(U)$:

$$\begin{aligned} U \subset \mathbb{R}^n &\xrightarrow{f} \quad V \quad \xrightarrow{g} \mathbb{R}^p \\ a &\longrightarrow f(a) \longrightarrow g(f(a)). \end{aligned}$$

Suposem que f és diferenciable al punt a i g diferenciable al punt $b = f(a)$. Llavors $g \circ f$ és diferenciable al punt a i

$$D(g \circ f)(a) = D(g)(f(a)) \circ Df(a),$$

on en el membre de la dreta hi ha una composició d'aplicacions lineals. És a dir,

$$\frac{\partial (g \circ f)_k}{\partial x_j}(a) = \sum_{\ell=1}^{m} \frac{\partial g_k}{\partial y_\ell}(f(a)) \frac{\partial f_\ell}{\partial x_j}(a).$$

1. Cas $\boldsymbol{n = 1 = p}$, $\boldsymbol{m = n}$. Posem $h = g(f(t))$

$$\begin{aligned} &\mathbb{R} \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R} \\ &t \longrightarrow (f_1(t), \ldots, f_n(t)) \longrightarrow g(f_1(t), \ldots, f_n(t)). \end{aligned}$$

Aplicant la regla de la cadena

$$\frac{dh}{dt} = \left(\frac{\partial g}{\partial y_1}, \ldots, \frac{\partial g}{\partial y_n} \right) \begin{pmatrix} \dfrac{df_1}{dt} \\ \vdots \\ \dfrac{df_n}{dt} \end{pmatrix} = \frac{\partial g}{\partial y_1} \frac{df_1}{dt} + \cdots + \frac{\partial g}{\partial y_n} \frac{df_n}{dt}.$$

Noteu que això és $\langle \nabla g(f(t)), f'(t) \rangle$.

La identitat

$$\frac{dh}{dt} = \frac{\partial g}{\partial y_1} \frac{df_1}{dt} + \cdots + \frac{\partial g}{\partial y_n} \frac{df_n}{dt}$$

es pot explicar d'una manera simple: g depèn de les variables $y_1, \ldots, y_n$, així que al derivar $g(f(t))$ primer derivem g respecte de y_1 i llavors $y_1 = f_1(t)$ respecte de t; el producte $\frac{\partial g}{\partial y_1} \frac{df_1}{dt}$ representa l'efecte de la variable y_1. Operem igual amb la resta de variables y_j i després sumem els resultats.

Exemple 1. La funció g és diferenciable a $\mathbb{R}^n$ i $f(t) = a + t(b - a)$, $t \in \mathbb{R}$, que és la recta pel punt a que té vector director $b - a$ (a i b són punts de $\mathbb{R}^n$ arbitraris). Llavors

$$g(f(t)) = g(a + t(b - a))$$

i

$$\frac{d(g \circ f)(t)}{dt} = \frac{\partial g}{\partial x_1}(a + t(b - a))(b_1 - a_1) + \cdots + \frac{\partial g}{\partial x_n}(a + t(b - a))(b_n - a_n).$$

Noteu que aquesta expressió és

$$\langle \nabla g(a + t(b - a)), b - a \rangle.$$

Exemple 2. La funció $f(t)$ és una funció diferenciable qualsevol i $g(x) = \|x\|^2$ és el quadrat de la distància a l'origen. Llavors

$$g(f(t)) = f_1^2(t) + \cdots + f_n^2(t) \tag{2.8}$$

i

$$\frac{d(g \circ f)(t)}{dt} = (2x_1, \ldots, 2x_n) \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix} = 2f_1(t)f_1'(t) + \cdots + 2f_n(t)f_n'(t),$$

resultat que es recupera derivant directament respecte de t a (2.8).

2. Cas $m = p = 1$.

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}.$$

La regla de la cadena diu

$$\left(\frac{\partial g \circ f}{\partial x_1}, \ldots, \frac{\partial g \circ f}{\partial x_n} \right) = \frac{dg}{dt}(f(x)) \left(\frac{\partial f}{\partial x_1}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right)$$

o

$$\frac{\partial (g \circ f)}{\partial x_j}(x) = \frac{dg}{dt}(f(x)) \frac{\partial f}{\partial x_j}(x), \quad 1 \le j \le n.$$

Exemple.

$$\begin{array}{ccccc} \mathbb{R}^n & \xrightarrow{f} & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & & t & \longrightarrow & e^t \end{array}$$

$$\frac{\partial}{\partial x_j} e^{f(x)} = e^{f(x)} \frac{\partial f}{\partial x_j}(x), \quad 1 \le j \le n.$$

3. Cas $n = m$, $p = 1$.

$$\mathbb{R}^n \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R}.$$

$$\left(\frac{\partial g}{\partial y_1}, \dots, \frac{\partial g}{\partial y_n}\right) \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dots & \dfrac{\partial f_n}{\partial x_n} \end{pmatrix} = \left(\dots, \frac{\partial g}{\partial y_1}\frac{\partial f_1}{\partial x_j} + \dots + \frac{\partial g}{\partial y_n}\frac{\partial f_n}{\partial x_j}, \dots\right)$$

o

$$\frac{\partial g(f(x))}{\partial x_j} = \frac{\partial g}{\partial y_1}\frac{\partial f_1}{\partial x_j} + \frac{\partial g}{\partial y_2}\frac{\partial f_2}{\partial x_j} + \dots + \frac{\partial g}{\partial y_n}\frac{\partial f_n}{\partial x_j}.$$

Exemple.

$$\begin{aligned} (0, \infty) \times (0, 2\pi) &\xrightarrow{f} \mathbb{R}^2 \xrightarrow{g} \mathbb{R} \\ (r, \theta) &\longrightarrow (r\cos\theta, r\sin\theta) \\ & \qquad\quad (x, y) \longrightarrow x^2 + y. \end{aligned}$$

Tenim

$$Df(r, \theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$$

i

$$Dg(x, y) = (2x, 1).$$

La regla de la cadena dóna:

$$\begin{aligned} \left(\frac{\partial g \circ f}{\partial r}, \frac{\partial g \circ f}{\partial \theta}\right) &= (Dg \circ Df)(r, \theta) = (2x, 1) \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} \\ &= (2x\cos\theta + \sin\theta, -r2x\sin\theta + r\cos\theta) \\ &= (2r\cos^2\theta + \sin\theta, -2r^2\cos\theta\sin\theta + r\cos\theta). \end{aligned}$$

Ara,

$$g(f(r, \theta)) = r^2\cos^2\theta + r\sin\theta$$

i derivant respecte de r i θ recuperem la fórmula precedent.

4. Cas $n = 1$, $m = n$, $p = n$.

$$\begin{aligned} \mathbb{R} &\xrightarrow{c} \mathbb{R}^n \xrightarrow{f} \mathbb{R}^n \\ t &\longrightarrow c(t) \longrightarrow f(c(t)). \end{aligned}$$

$$\begin{aligned} \frac{df(c(t))}{dt} &= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(c(t)) & \dots & \frac{\partial f_n}{\partial x_n}(c(t)) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(c(t)) & \dots & \frac{\partial f_n}{\partial x_n}(c(t)) \end{pmatrix} \begin{pmatrix} c_1'(t) \\ \vdots \\ c_n'(t) \end{pmatrix} \\ &= Df(c(t))(c'(t)). \end{aligned}$$

O sigui: la velocitat del mòbil que traça la corba $f(c(t))$ és la imatge per l'aplicació lineal $Df(c(t))$ de la velocitat del mòbil que traça la corba $c(t)$ (tot a l'instant t).

Demostració del cas 1. Com que $\mathbb{R} \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R}$, tenim que

$$\begin{aligned} \frac{g(f(t)) - g(f(t_0))}{t - t_0} &= \frac{g(f(t)) - g(f(t_0)) - \langle \nabla g(f(t_0)), f(t) - f(t_0) \rangle}{t - t_0} \\ &\quad + \left\langle \nabla g(f(t_0)), \frac{f(t) - f(t_0)}{t - t_0} \right\rangle = I + II. \end{aligned}$$

Clarament,

$$II \xrightarrow{t \to t_0} \langle \nabla g(f(t_0)), f'(t_0) \rangle.$$

Si $f(t) = f(t_0)$, I s'anul·la. Si $f(t) \neq f(t_0)$,

$$\begin{aligned} |I| = \frac{|g(f(t)) - g(f(t_0)) - \langle \nabla g(f(t_0)), f(t) - f(t_0) \rangle|}{\|f(t) - f(t_0)\|} \left\| \frac{f(t) - f(t_0)}{t - t_0} \right\| \\ \xrightarrow{t \to t_0} 0 \, \|f'(t_0)\| = 0. \quad \square \end{aligned}$$

Demostració del cas 2. Ara som al cas $\mathbb{R}^n \xrightarrow{f} \mathbb{R} \xrightarrow{g} \mathbb{R}$. Sumant i restant el terme

$$g'(f(a))(f(x) - f(a))$$

obtenim

$$\begin{aligned}|g(f(x))-g(f(a))-g'(f(a))\langle\nabla f(a),x-a\rangle| \le |g(f(x)) - g(f(a)) - g'(f(a))(f(x) - f(a))| \\ + |g'(f(a))|\,|f(x) - f(a) - \langle\nabla f(a), x - a\rangle|.\end{aligned}$$

El segon terme té límit zero al dividir-lo per $\|x - a\|$ perquè f és diferenciable al punt a. El primer terme és zero si $f(x) = f(a)$. Si $f(x) \neq f(a)$ llavors el primer terme és

$$|f(x) - f(a)| \left|\frac{g(f(x)) - g(f(a))}{f(x) - f(a)} - g'(f(a))\right|$$

i el segon factor tendeix a zero si $x \to a$ perquè f és contínua al punt a i g és diferenciable al punt $f(a)$. Al dividir el primer factor per $\|x - a\|$ obtenim

$$\frac{|f(x) - f(a)|}{\|x - a\|}\eta(x)$$

amb $\eta(x) \to 0$ si $x \to a$. L'argument s'acaba notant que hi ha $r > 0$ tal que $\frac{|f(x)-f(a)|}{\|x-a\|} \le C$, $\|x - a\| < r$. □

2.2 Gradient i derivades direccionals

Sigui $f\colon U \to \mathbb{R}$, U obert $\subset \mathbb{R}^n$, $a \in U$ i $\vec{n}$ un vector de $\mathbb{R}^n$ de norma 1. La recta que passa per a i té vector director $\vec{n}$ es pot parametritzar per

$$t \in \mathbb{R} \longrightarrow a + t\vec{n}.$$

Com que $a \in U$ i U és obert hi ha una bola centrada al punt a i de radi r prou petit de manera que $B(a, r) \subset U$. Si $t \in (-r, r)$ el punt $a + t\vec{n} \in U$ i llavors té sentit $f(a + t\vec{n})$. Per tant $f(a + t\vec{n})$ és una funció de la variable real $t \in (-r, r)$.

Definició. La derivada direccional de f al punt a en la direcció $\vec{n}$ és la derivada (sempre que existeixi) de $f(a + t\vec{n})$ al punt $t = 0$:

$$D_{\vec{n}}f(a) = \frac{d}{dt}f(a + t\vec{n})\bigg|_{t=0} = \lim_{t\to 0}\frac{f(a + t\vec{n}) - f(a)}{t}.$$

La derivada $D_{\vec{n}}f(a)$ s'ha d'entendre com la variació de f en el punt a en la direcció $\vec{n}$. Per exemple,

$$D_{e_j}f(a) = \frac{\partial f}{\partial x_j}(a), \quad 1 \le j \le n,$$

on $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ és el j-èsim vector de la base canònica de $\mathbb{R}^n$.

Exemple. Prenem $f(x,y) = x - y + y^3$. Les derivades a l'origen en les direccions e_1 i e_2 són, respectivament, 1 i -1. Calculem la derivada a l'origen en la direcció donada pel vector $(1,1)$. Tenim $\vec{n} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ i $f\left((0,0) + t\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)\right) = \left(\frac{t}{\sqrt{2}}\right)^3$. Doncs $D_{\vec{n}}f(0,0) = 0$. Calculeu com exercici la derivada a l'origen en la direcció $(1,2)$.

El càlcul d'una d'una derivada direccional es fa habitualment aplicant el següent fet.

Teorema. *Si f és diferenciable al punt a, llavors $D_{\vec{n}}f(a)$ existeix i*

$$D_{\vec{n}}f(a) = Df(a)(\vec{n}) = \langle \nabla f(a), \vec{n} \rangle.$$

Demostració. És simplement la regla de la cadena ($\vec{n} = (n_1, \ldots, n_j, \ldots, n_n)$)

$$\frac{d}{dt} f(a + t\vec{n})\bigg|_{t=0} = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) n_j = \langle \nabla f(a), \vec{n} \rangle. \qquad \square$$

Aquest resultat tant senzill permet donar una interpretació geomètrica del gradient que és molt útil.

Si $\nabla f(a) \neq 0$, el vector $\nabla f(a)$ assenyala la direcció en la qual f creix més ràpidament i $-\nabla f(a)$ la direcció en la qual f decreix més ràpidament. En efecte,

$$D_{\vec{n}}f(a) = \langle \nabla f(a), \vec{n} \rangle = \|\nabla f(a)\| \cos\theta,$$

on θ és l'angle entre $\nabla f(a)$ i $\vec{n}$. La derivada direccional és màxima si $\cos\theta = 1$ ($\theta = 0$) i mínima si $\cos\theta = -1$ ($\theta = \pi$).

Exercici. La temperatura d'una nau espacial ve donada per una funció

$$T(x,y,z) = e^{-x^2 - 2y^2 - 3z^2}.$$

Si la nau és al punt $(1,1,1)$, en quina direcció ha de navegar per tal que la temperatura disminueixi al més ràpidament possible? Resposta: (1,2,3)

Exercici. Sigui $f(x,y) = x^2 - y^2$. Quina és la direcció de màxim creixement de f en el punt $(0,1,-1)$ de la gràfica de f? I en el punt $\left(\frac{1}{2}, 1, -\frac{3}{4}\right)$?

El gradient té una altra interpretació geomètrica que ara descrivim. Prenem una funció de dues variables $z = f(x,y)$ i pensem que $f(x,y)$ és l'altura sobre el nivell del mar d'un perfil de muntanya. Hem de pensar que (x,y) és un punt del mapa tal que el punt de la muntanya $(x, y, f(x, y))$ està a altura $f(x, y)$. Fixem una altura $c \in \mathbb{R}$. Llavors el conjunt de punts del mapa tals que el corresponent punt de la muntanya està a altura c s'anomena línia o corba de nivell:

$$\{(x, y) \in \mathbb{R}^2 : f(x, y) = c\}. \tag{2.9}$$

Per exemple, si $f(x, y) = x^2 + y^2$, les corbes de nivell corresponents a altures positives són circumferències de centre l'origen i radi $\sqrt{c}$. La corba de nivell d'altura 0 és l'origen i la corba de nivell d'altura negativa és buida.

Si la funció $f(x, y)$ no és prou bona les corbes de nivell poden tenir patologies diverses. Per exemple, per $f(x, y) = xy$ i $c \neq 0$ la corba de nivell (2.9) és una hipèrbola (amb dues branques) i si $c = 0$, és la unió dels eixos coordenats.

Suposem que la corba de nivell és prou bona de manera que es pot "parametritzar" localment (és a dir, al voltant de tot punt de la corba) per una funció diferenciable definida a un interval $(\alpha, \beta) \subset \mathbb{R}$

$$\begin{aligned} c\colon (\alpha, \beta) &\longrightarrow \mathbb{R}^2 \\ t &\longrightarrow c(t), \end{aligned}$$

i

$$\{(x, y) \in \mathbb{R}^2 : f(x, y) = c\} \supset \{c(t) : t \in (\alpha, \beta)\}.$$

Llavors el vector $\nabla f(c(t))$ és "perpendicular" a la corba (2.9) en el punt $c(t)$. Això és clar, perquè

$$f(c(t)) = c, \quad t \in (\alpha, \beta),$$

i, doncs, derivant respecte de t i aplicant la regla de la cadena

$$\langle \nabla f(c(t)), c'(t) \rangle = 0, \quad t \in (\alpha, \beta).$$

Com que $c'(t)$ és el vector tangent a la corba en el punt $c(t)$, la interpretació de $\nabla f(c(t))$ com a vector perpendicular a la corba al punt $c(t)$ és correcta.

Exemple 3. $f(x, y) = x^2 + y^2$, $(x, y) \in \mathbb{R}^2$.

Llavors $\nabla f(x, y)$ és un vector perpendicular a la corba $\{(x, y) : x^2 + y^2 = c\}$, $c > 0$, en qualsevol punt d'aquesta corba. En efecte podem parametritzar tota la corba per

$$\begin{aligned} [0, 2\pi] &\longrightarrow \mathbb{R}^2 \\ t &\longrightarrow \sqrt{c}(\cos t, \sin t). \end{aligned}$$

Llavors $f(\sqrt{c}\cos t, \sqrt{c}\sin t) = c$, $0 \leq t \leq 2\pi$, i derivant respecte de t obtenim

$$\frac{\partial f}{\partial x}(\sqrt{c}\cos t, \sqrt{c}\sin t)(-\sqrt{c}\sin t) + \frac{\partial f}{\partial y}(\sqrt{c}\cos t, \sqrt{c}\sin t)\sqrt{c}\cos t = 0, \quad 0 \leq t \leq 2\pi.$$

Com que $\sqrt{c}(-\sin t, \cos t)$ és tangent a la corba, $\nabla f(x,y)$ és perpendicular a la corba en el punt $(x,y) = (\sqrt{c}\cos t, \sqrt{c}\sin t)$, que és un punt arbitrari de la corba.

Una manera més ràpida de pensar-ho és la següent. Clarament $\nabla f(x,y) = 2(x,y)$. Per altra banda, parametritzant la corba veiem que $\sqrt{c}(-\sin t, \cos t) = (-y,x)$ és tangent a la corba al punt (x,y). Com que $\langle \nabla f(x,y), (-y,x)\rangle = 2(-xy+yx) = 0$, $\nabla f(x,y)$ és perpendicular a la corba al punt (x,y).

Exemple 4. $f(x,y) = xy$, $(x,y) \in \mathbb{R}^2$.

Considerem la corba de nivell $xy = 0$. El gradient de f, $\nabla f(x,y) = (y,x)$ s'anul·la l'origen i doncs és perpendicular a qualsevol altre vector (de fet, no té gaire sentit parlar de perpendicularitat en aquest cas).

Comproveu, com a exercici, que si $(x,y) \neq (0,0)$ es pot parametritzar la corba $xy = 0$ i concloure que $\nabla f(x,y)$ és perpendicular a la corba.

Exercici. Sigui $f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$.

Demostreu que $\nabla f(x,y)$ és perpendicular a l'el·lipse $\left\{(x,y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\right\}$ en tot punt (x,y) de l'el·lipse (l'única cosa que s'ha de fer és parametritzar l'el·lipse en un entorn del punt (x,y)).

Definició. Una superfície de nivell de la funció de 3 variables $f(x,y,z)$ és un conjunt del tipus

$$S = \{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = c\},$$

on c és una constant real. Com en el cas de les corbes de nivell, si $(x,y,z) \in S$ llavors $\nabla f(x,y,z)$ es pot interpretar com un vector perpendicular a S. En efecte, considerem qualsevol corba diferenciable

$$\begin{aligned} (\alpha,\beta) &\longrightarrow S \\ t &\longrightarrow c(t) \end{aligned}$$

dins la superfície que passi pel punt $(x,y,z) \in S$. Això vol dir que $(x,y,z) = c(t_0)$ per cert $t_0 \in (\alpha,\beta)$. Llavors $f(c(t)) = c$, $t \in (\alpha,\beta)$, i derivant respecte de t

$$\langle \nabla f(c(t)), c'(t)\rangle = 0, \quad t \in (\alpha,\beta).$$

Ara $c'(t)$ és el vector tangent a la corba en el punt $c(t)$ i, per tant, és també un vector "tangent" a S en el punt $c(t)$. Per $t = t_0$, $c(t_0) = (x,y,z)$ i concloem que $\nabla f(x,y,z)$ és perpendicular a $c'(t_0)$. Si definim vector tangent a S en el punt $(x,y,z) \in S$ com un vector tangent en el punt (x,y,z) a una corba diferenciable dins S que passi pel punt, resultarà que $\nabla f(x,y,z)$ és perpendicular a tot vector tangent a S en el punt (x,y,z). Doncs $\nabla f(x,y,z)$ és "perpendicular a S en el punt (x,y,z)".

Exercicis

1. Calculeu el vector perpendicular a $x^2 + y^2 + z^2 = 1$ en qualsevol punt de la superfície.

2. Calculeu el vector perpendicular a la superfície $x^2 + y^2 - z^2 = 0$ en el punt $(0, 1, 1)$. Verifiqueu amb una figura que el càlcul correspon a la intuició geomètrica.

Exemple (El camp gravitatori)**.** Una massa M és a l'origen. Si una altra massa m és al punt $x \neq 0$, la llei de la gravitació de Newton diu que sobre m actua una força d'atracció que és proporcional a mM i a l'invers de la distància entre les masses. Llavors, per certes constants G i c_0 la força és

$$\vec{F}(x) = -\frac{GmM}{\|x\|^2}\frac{x}{\|x\|} = -c_0\frac{x}{\|x\|^3}.$$

Figura

Resulta que hi ha una funció escalar $V(x)$ (el potencial gravitatori) tal que

$$\nabla V(x) = \vec{F}(x).$$

En efecte,

$$\partial_j\left(\frac{1}{\|x\|}\right) = \partial_j(x_1^2 + \cdots + x_n^2)^{-\frac{1}{2}} = -\frac{1}{2}(\|x\|^2)^{-\frac{3}{2}}2x_j = -\frac{x_j}{\|x\|^3}$$

i doncs

$$\nabla\left(\frac{1}{\|x\|}\right) = -\frac{x}{\|x\|^3}, \quad x \neq 0.$$

Per tant $V(x) = c_0\frac{1}{|x|}$.

Les superfícies amb el mateix potencial $V(x) = c$ són esferes centrades a l'origen i $\nabla V(x)$, que apunta a l'origen, és perpendicular a l'esfera.

Exercici. Calculeu $\nabla(|x|^\alpha)$, $\alpha \in \mathbb{R}$. Feu primer els casos $\alpha = 1$, $\alpha = 2$ i després el cas general.

Derivades d'ordre superior

A vegades una funció f té derivades parcials que es poden tornar a derivar. Per exemple, si $f(x,y) = xy + y^2$, $\frac{\partial f}{\partial x} = y$. Llavors

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 0 \quad \text{i} \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = 1.$$

També $\frac{\partial f}{\partial y} = x + 2y$ i, per tant,

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = 1 \quad \text{i} \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = 2.$$

Les derivades parcials de segon ordre d'una funció $f\colon U \to \mathbb{R}$, U obert $\subset \mathbb{R}^n$, són, si existeixen les funcions (cas $n = 2$)

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right),\\
\frac{\partial^2 f}{\partial y^2} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right),\\
\frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right),\\
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right).
\end{aligned}$$

En el cas de n variables posem

$$\frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial f}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right).$$

Una funció és de classe $C^k(U)$, $0 \leq k \leq \mathbb{Z}$, si té derivades parcials d'ordre k que són contínues a U. És de classe $C^\infty(U)$ si és de classe $C^k(U)$ per tot $k = 1, 2, \ldots$

Exemple.

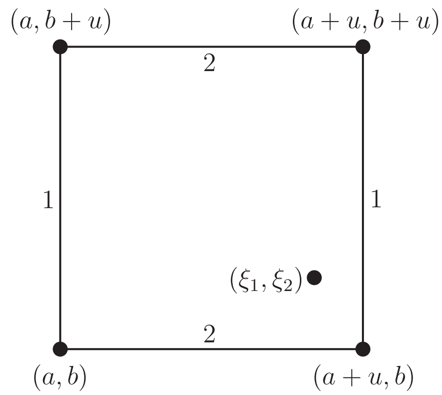
$$
\begin{aligned}
f(x,y,z) &= x\sin y \cdot \sin z - y\sin(xz),\\
\frac{\partial f}{\partial x} &= \sin y \sin z - yz\cos(xz),\\
\frac{\partial f}{\partial y} &= x\cos y \sin z - \sin xz,\\
\frac{\partial f}{\partial z} &= x\sin y \cos z - yx\cos(xz),\\
\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) &= \cos y \sin z - z\cos(xz),\\
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) &= \cos y \sin z - \cos(xz)z,\\
\frac{\partial}{\partial z}\left(\frac{\partial f}{\partial x}\right) &= \sin y \cos z - y\cos(xz) + xyz\sin(xz),\\
\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial z}\right) &= \sin y \cos z - y\cos(xz) + zyx\sin(xz).
\end{aligned}
$$

Així que les derivades segones “creuades” són iguals!

Teorema (Schwarz)**.** *Sigui $f\colon U \to \mathbb{R}$, U obert de $\mathbb{R}^n$ i suposem que existeixen $\partial_k\partial_j f$ i $\partial_j\partial_k f$ a U i són contínues a U. Llavors*

$$\partial_k\partial_j f = \partial_j\partial_k f.$$

Demostració. Fem el cas $n = 2$. Fixem el punt $p = (a,b) \in U$. Hi ha $\delta < 0$ tal que $B(p,\delta) \subset U$. Agafo u, $0 < u < \frac{\delta}{\sqrt{2}}$ i escric la “segona diferència” (s'utilitza aquesta expressió perquè ens referim a una diferència de diferències) dividida pel quadrat de l'increment u:



$$\begin{aligned}
Q(u) &= \frac{1}{u^2}\left[f(a+u,b+u) - f(a+u,b) - (f(a,b+u) - f(a,b))\right] \\
&= \frac{1}{u^2}\left[f(a+u,b+u) - f(a,b+u) - (f(a+u,b) - f(a,b))\right].
\end{aligned}$$

Defineixo

$$g(x) = f(x, b+u) - f(x, b),$$

de manera que, pel teorema del valor mitjà aplicat dues vegades

$$\begin{aligned}
Q(u) &= \frac{1}{u^2}\left[g(a+u) - g(a)\right] \\
&= \frac{1}{u} g'(\xi_1) = \frac{1}{u}\left[\frac{\partial f}{\partial x}(\xi_1, b+u) - \frac{\partial f}{\partial x}(\xi_1, b)\right] \\
&= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(\xi_1, \xi_2)\right),
\end{aligned}$$

on $a < \xi_1 < a+u$ i $b < \xi_2 < b+u$.

Fent $u \to 0$

$$\lim_{u \to 0} Q(u) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(a,b)\right).$$

Fent servir la segona expressió per $Q(u)$ i la funció auxiliar

$$h(x) = f(a+u, x) - f(a, x)$$

trobem que

$$\lim_{u \to 0} Q(u) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(a,b)\right). \qquad \square$$

Notació per les derivades parcials

S'utilitza molt, per simplificar, $\partial_j = \dfrac{\partial}{\partial x_j}$.

Per tractar derivades successives del tipus

$$\frac{\partial^4}{\partial x_3^4}\left(\frac{\partial}{\partial x_2}\left(\frac{\partial^2 f}{\partial x_1^2}\right)\right)$$

s'utilitzen multi-indexos, és a dir elements de $\mathbb{Z}^n$ amb coordenades no negatives

$$\alpha = (\alpha_1, \ldots, \alpha_n), \quad 0 \leq \alpha_j \in \mathbb{Z}, \quad 1 \leq j \leq n.$$

S'entén que α_j és el nombre de vegades que es deriva respecte de x_j. Es posa

$$|\alpha| = \alpha_1 + \cdots + \alpha_n$$

que és el nombre total de derivades que es prenen. Finalment es posa

$$\partial^\alpha = \partial_1^{\alpha_1} \ldots \partial_n^{\alpha_n}.$$

Per exemple,

$$\begin{aligned} \partial^{(1,2)} f &= \partial_1 \partial_2^2 f = \partial_1 \partial_2 \partial_2 f, \\ \partial^{(0,2,1)} f &= \partial_2^2 \partial_3 f = \partial_2 \partial_2 \partial_3 f. \end{aligned}$$

La fórmula de Taylor

Sigui $f\colon U \to \mathbb{R}$, U obert de $\mathbb{R}^n$, $a \in U$ i suposem que f és diferenciable al punt a. Llavors

$$f(x) = f(a) + Df(a)(x-a) + R_1(x,a), \tag{2.10}$$

on R_1 és "la resta d'ordre 1" i compleix que

$$\frac{R_1(x,a)}{\|x-a\|} \xrightarrow{x \to a} 0$$

per la definició de diferenciabilitat. La identitat (2.10) és la fórmula de Taylor d'ordre 1. La fórmula de Taylor d'ordre 2 és la següent.

Teorema. *Si $f \in C^2(U)$, U obert de $\mathbb{R}^n$, llavors per a U tenim que*

$$f(x) = f(a) + Df(a)(x-a) + \frac{1}{2}\sum_{j,k=1}^{n} \partial_{jk} f(a)(x_j - a_j)(x_k - a_k) + R_2(x,a)$$

i "la resta" $R_2(x,a)$ compleix

$$\frac{R_2(x,a)}{\|x-a\|^2} \xrightarrow{x\to a} 0.$$

Demostració. Farem una reducció al cas d'una variable. Sigui $r > 0$ prou petit de manera que $B(a,r) \subset U$ i prenem $x \in B(a,r)$. Posem $h = x - a$ i definim

$$g(t) = f(a+th).$$

Noteu que $g(t)$ està ben definida per $|t|\,\|h\| < r$, és a dir a l'interval $\left(-\frac{r}{\|h\|}, \frac{r}{\|h\|}\right) \subset \mathbb{R}$. Per la fórmula de Taylor en una variable

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(0)t^2 + \rho_2(t) \tag{2.11}$$

amb $\frac{\rho_2(t)}{t^2} \xrightarrow{t\to 0} 0$.

Aplicant la regla de la cadena

$$g'(t) = \sum_{j=1}^{n} \partial_j f(a+th)h_j$$

i

$$g''(t) = \sum_{j,k=1}^{n} \partial_{jk} f(a+th)h_j h_k.$$

Així que

$$g'(0) = \langle \nabla f(a), h\rangle$$

i

$$g''(0) = \sum_{j,k=1}^{n} \partial_{jk} f(a)h_j h_k.$$

Per tant (2.11) per $t = 1 < \frac{r}{\|h\|}$ és

$$f(a+h) = f(a) + \langle \nabla f(a), h\rangle + \frac{1}{2}\sum_{j,k=1}^{n} \partial_{jk} f(a) h_j h_k + \rho_2(1). \tag{2.12}$$

Ara, pel teorema del valor mitjà aplicat dues vegades,

$$\begin{aligned}\rho_2(t) &= g(t) - g(0) - g'(0)t - \frac{g''(0)}{2}t^2 \\ &= [g'(\xi_1) - g'(0) - g''(0)\xi_1]t \\ &= [g''(\xi) - g''(0)]\xi_1 t\end{aligned}$$

amb $0 < \xi < \xi_1 < t$.

Doncs

$$\frac{|\rho_2(t)|}{t^2} \le |g''(\xi) - g''(0)|, \quad \text{amb } 0 < \xi < t,$$

i

$$|\rho_2(1)| \le |g''(\xi) - g''(0)| = \left|\sum_{j,k=1}^{n} (\partial_{jk} f(a+\xi h) - \partial_{jk} f(a)) h_j h_k\right|.$$

Com que $\partial_{jk} f(x)$ és contínua al punt a

$$\begin{aligned}\frac{|\rho_2(1)|}{\|h\|^2} &= \left|\sum_{j,k=1}^{n} (\partial_{jk} f(a+\xi h) - \partial_{jk} f(a)) \frac{h_j}{\|h\|}\frac{h_k}{\|h\|}\right| \\ &\le \sum_{j,k=1}^{n} |\partial_{j,k} f(a+\xi h) - \partial_{jk} f(a)| \xrightarrow{h\to 0} 0.\end{aligned}$$

Per tant (2.12) es converteix en

$$f(x) = f(a) + \langle \nabla f(a), x-a\rangle + \frac{1}{2}\sum_{j,k=1}^{n} \partial_{jk} f(a)(x_j - a_j)(x_k - a_k) + R_2(x,a)$$

amb $\frac{R_2(x,a)}{\|x-a\|^2} = \frac{\rho_2(1)}{\|h\|^2} \to 0$. □

Exemple 5. Considerem la funció

$$f(x,y) = 5 + 3x - 5y - x^2 + y^2 - xy + x^3 + xy^2.$$

Volem trobar els polinomis de Taylor homogenis de grau 0, 1 i 2 al voltant de l'origen. Clarament

$$\begin{aligned} 5 &= f(0,0), \\ 3x - 5y &= \partial_1 f(0)x + \partial_2 f(0)y \end{aligned}$$

és el polinomi de Taylor homogeni de grau 1 i

$$y^2 - xy = \frac{1}{2}\left[\partial_1^2 f(0)x^2 + \partial_2^2 f(0)y^2 + \partial_{12} f(0)xy\right]$$

és el polinomi de Taylor homogeni de grau 2. La resta és

$$R_2(x,y) = x^3 + xy^2$$

perquè

$$\frac{x^3 + xy^2}{x^2 + y^2} \longrightarrow 0 \text{ si } (x,y) \to (0,0).$$

Exemple 6. Prenem

$$f(x,y) = \sin x - \cos y + 3\sin^2 x \cos y + (e^x - 1)^2 \sin(xy).$$

Utilitzant

$$\begin{aligned} \sin x &= x + O(x^3) && \text{si } x \to 0 \\ \cos y &= 1 - \frac{y^2}{2} + O(y^4) && \text{si } y \to 0 \end{aligned}$$

trobem que

$$f(x,y) = -1 + x + \frac{y^2}{2} + 3x^2 + R_2(x,y)$$

amb

$$\frac{R_2(x,y)}{x^2 + y^2} \longrightarrow 0 \text{ si } (x,y) \to (0,0).$$

Exemple 7. Sigui $f(x,y) = x^2 + y^3$. Escrivim la fórmula de Taylor al voltant de $(0,1)$. Calculant trobem

$$\begin{aligned} &f(0,1) = 1, \\ &\frac{\partial f}{\partial x}(0,1) = 0, \quad \frac{\partial f}{\partial y}(0,1) = 3, \\ &\frac{\partial^2 f}{\partial x^2}(0,1) = 0, \quad \frac{\partial^2 f}{\partial y^2}(0,1) = 6, \quad \frac{\partial^2 f}{\partial y \partial x}(0,1) = 0. \end{aligned}$$

Doncs
$$f(x,y) = 1 + 3(y-1) + 3(y-1)^2 + R_2(x,y).$$
El resultat s'obté també notant que
$$y^3 = (y-1+1)^3 = (y-1)^3 + 3(y-1)^2 + 3(y-1) + 1,$$
d'on deduïm que
$$f(x,y) = x^2 + 1 + 3(y-1) + 3(y-1)^2 + (y-1)^3.$$
Llavors
$$R_2(x,y) = (y-1)^3.$$

Exemple 8. Sigui $f(x,y) = e^{x^2+y^2}$. Escrivim la fórmula de Taylor fins a l'ordre 2 a l'origen. Sabem que
$$e^t = 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} + \cdots$$
Per tant
$$e^{x^2+y^2} = 1 + x^2 + y^2 + \frac{(x^2+y^2)^2}{2} + \cdots$$
i $R_2(x,y) = \frac{(x^2+y^2)^2}{2} + \frac{(x^2+y^2)^3}{3!} + \cdots$. Calculant $\partial^\alpha f(0,0)$, $|\alpha| \leq 2$, trobaríem el polinomi de Taylor d'ordre 2 a l'origen, però no la fórmula explícita per la resta.

2.3 Extrems locals

Sigui $f\colon U \to \mathbb{R}$, U obert $\subset \mathbb{R}^n$, una funció.

Definició. Diem que f té un màxim (mínim) local al punt $a \in U$ si hi ha una bola $B(a,r)$ centrada al punt a, $B(a,r) \subset U$, tal que
$$f(x) \leq f(a), \quad x \in B(a,r) \qquad \big(f(x) \geq f(a), \quad x \in B(a,r)\big).$$
Un extrem local és un màxim o mínim local.

Recordem que en una variable si f té un extrem local en el punt a i f és diferenciable al punt a, llavors $f'(a) = 0$ (la recta tangent al punt $(a, f(a))$ és horitzontal). El mateix passa en diverses variables.

Proposició. *Si f té un extrem local al punt a i f és diferenciable al punt a, llavors $\nabla f(a) = 0$ (per $n = 2$, el pla tangent és horitzontal).*

Demostració. Sigui

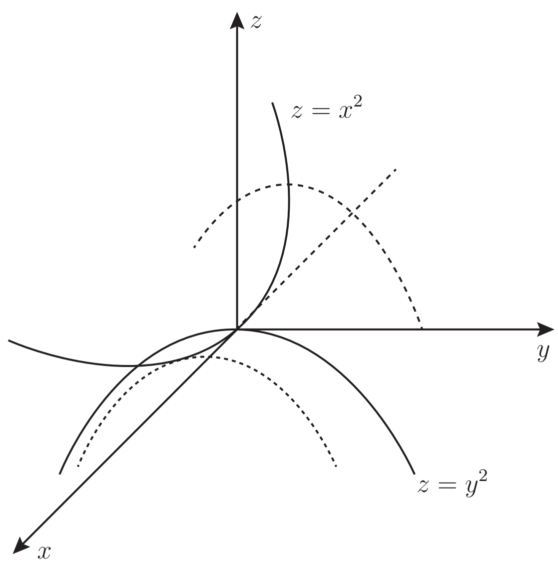
$$g(t) = f(a_1, \dots, a_{j-1}, t, a_{j+1}, \dots, a_n).$$

Llavors g està definida a un interval centrat al punt $a_j \in \mathbb{R}$ i té un extrem local a a_j. Doncs

$$0 = \left.\frac{dg}{dt}\right|_{t=a_j} = \frac{\partial f}{\partial x_j}(a). \qquad \square$$

Definició. Un punt crític de f és un punt a on $\nabla f(a) = 0$.

Com en una variable el problema és decidir, quan a és un punt crític, si a és un extrem local. Sabem que això no sempre passa com mostra la funció $y = x^3$ en una variable o la funció $z = x^2 - y^2$ en dues variables, que té un punt de sella a l'origen.



El criteri més famós d'extrem local en una variable involucra la derivada segona.

Teorema. *Si $f\colon (\alpha, \beta) \to \mathbb{R}$ és dues vegades diferenciable a (α, β) i per $a \in (\alpha, \beta)$, $f'(a) = 0$, llavors f té un màxim local al punt a si $f''(a) < 0$ i un mínim local si $f''(a) > 0$. El cas $f''(a) = 0$ és dubtós.*

Demostració. La fórmula de Taylor fins a l'ordre 2 diu que

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \eta(x)(x-a)^2$$

on $\eta(x) \to 0$ si $x \to a$.

Suposem que $f'(a) = 0$ i $f''(a) < 0$. Llavors

$$f(x) = f(a) + \left(\frac{f''(a)}{2} + \eta(x)\right)(x-a)^2.$$

Prenem $\varepsilon > 0$ prou petit de manera que $\eta(x) < \frac{-f''(a)}{4}$ per $a - \varepsilon < x < a + \varepsilon$. Llavors si $a - \varepsilon < x < a + \varepsilon$ obtenim

$$f(x) \leq f(a) + \left(\frac{f''(a)}{2} - \frac{f''(a)}{4}\right)(x-a)^2 = f(a) + \frac{f''(a)}{4}(x-a)^2 \leq f(a).$$

Si $f'(a) = 0$ i $f''(a) > 0$ prenem $\varepsilon > 0$ tal que $\eta(x) > \frac{-f''(a)}{4}$ per $a - \varepsilon < x < a + \varepsilon$. Llavors per aquests x

$$f(x) \geq f(a) + \left(\frac{f''(a)}{2} - \frac{f''(a)}{4}\right)(x-a)^2 = f(a) + \frac{f''(a)}{4}(x-a)^2 \geq f(a). \qquad \square$$

Per estendre aquest argument a $\mathbb{R}^n$ necessitem la noció de forma quadràtica. La forma quadràtica associada a una matriu $n \times n$

$$A = \begin{pmatrix} a_{11} & \dots & a_{11} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix}$$

és l'aplicació $Q \colon \mathbb{R}^n \to \mathbb{R}$

$$Q(h) = \langle A(h), h \rangle = \sum_{j,k=1}^{n} a_{jk} h_j h_k$$

on

$$A(h) = \begin{pmatrix} \sum\limits_{j=1}^{n} a_{j_1} h_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{j_n} h_j \end{pmatrix}.$$

Definició. Si f és de classe C^2 a un obert que conté el punt a la matriu hessiana de f al punt a és

$$H(f)(a) = H = \begin{pmatrix} \partial_{11} f(a) & \dots & \partial_{1n} f(a) \\ \vdots & & \vdots \\ \partial_{n1} f(a) & \dots & \partial_{nn} f(a) \end{pmatrix}.$$

Noteu que H és una matriu simètrica. La forma quadràtica associada és el hessià de f al punt a:

$$Q(h) = \sum_{j,k=1}^{n} \partial_{jk} f(a) h_j h_k = \langle H(h), h \rangle.$$

Definició. Una forma quadràtica Q és definida positiva (negativa) si $Q(h) > 0$ si $h \neq 0$ ($Q(h) < 0$ si $h \neq 0$).

Per exemple, la forma quadràtica associada a $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ és $Q(h) = \frac{1}{2}(h_1^2 + h_2^2)$ que és definida positiva. La forma quadràtica associada a $\left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ és definida negativa. La forma quadràtica associada a $\left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ no és definida, perquè $Q(h) = h_1^2 - h_2^2$ és positiva per $h = (1, 0)$ i negativa per $h = (0, 1)$.

Lema. *Si Q és una forma quadràtica definida positiva, llavors hi ha $m > 0$ tal que*

$$Q(h) \geq m \|h\|^2, \quad h \in \mathbb{R}^n.$$

Demostració. La funció $Q(h)$ és contínua a tot $\mathbb{R}^n$ i, en particular, a

$$S^{n-1} = \{h \in \mathbb{R}^n : \|h\| = 1\}.$$

Com que S^{n-1} és un compacte, la restricció de Q a S^{n-1} té un mínim:

$$Q(h) \geq m, \quad \|h\| = 1.$$

Notem que m és positiu, perquè $m = Q(h_0) > 0$ per cert $h_0 \in S^{n-1}$. Ara si $h \in \mathbb{R}^n$, $h \neq 0$,

$$Q(h) = \|h\|^2 Q\left(\frac{h}{\|h\|}\right) \geq \|h\|^2 m,$$

com volíem demostrar. $\square$

Teorema (Criteri d'extrem local)**.** *Sigui $f\colon U \to \mathbb{R}$, $U \subset \mathbb{R}^n$ obert, una funció de classe $C^2(U)$ i $a \in U$ tal que $\nabla f(a) = 0$. Si la hessiana de f al punt a és definida positiva (negativa), llavors f té un mínim (màxim) local al punt a.*

Abans de fer la demostració considerem uns exemples.

Exemple 9. Sigui $f(x, y) = (x + y)^4$. L'origen és un punt crític i la hessiana a l'origen és idènticament nul·la. Per $g(x, y) = x^4 - y^4$ passa exactament el mateix. Però f té un mínim local (de fet absolut) a l'origen, mentre que g no té un extrem local a l'origen.

Exemple 10. Sigui $f(x, y) = x^2 + y^4$. L'origen és un punt crític i la hessiana a l'origen és

$$H = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

que és semi-definida positiva:

$$\langle h, H(h) \rangle \geq 0, \quad h \in \mathbb{R}^2.$$

El mateix passa per $g(x, y) = x^2 - y^4$. Com abans f té un mínim local a l'origen, però g no té extrem local a l'origen.

Exemple 11. La funció $f(x, y) = x^2 + y^2 + x^{10} + \sin(y^{11})x^4$ té un punt crític a l'origen i aplicant el teorema veiem que té un mínim local a l'origen.

Demostració del teorema. Sigui $H = (\partial_{jk} f(a))$ la matriu hessiana de f al punt a i

$$Q(h) = \langle h, H(h) \rangle, \quad h \in \mathbb{R}^n,$$

la forma quadràtica hessiana. La fórmula de Taylor fins a l'ordre 2 al voltant del punt a és

$$\begin{aligned} f(x) &= f(a) + \langle \nabla f(a), x - a \rangle + \frac{1}{2} Q(x - a) + R_2(x, a) \\ &= f(a) + \frac{1}{2} Q(x - a) + R_2(x, a) \end{aligned}$$

amb

$$\frac{R_2(x, a)}{\|x - a\|^2} \xrightarrow{x \to a} 0.$$

Pel lema hi ha $m > 0$ tal que

$$Q(x - a) \geq m\|x - a\|^2, \quad x \in \mathbb{R}^n.$$

Sigui $r > 0$ prou petit de manera que $B(a, r) \subset U$ i

$$\frac{R_2(x, a)}{\|x - a\|^2} > -\frac{m}{4}, \quad \|x - a\| < r.$$

Llavors, si $\|x - a\| < r$, tenin que

$$\begin{aligned} f(x) &\geq f(a) + \frac{m}{2}\|x - a\|^2 - \frac{m}{4}\|x - a\|^2 \\ &= f(a) + \frac{m}{4}\|x - a\|^2 \geq f(a). \end{aligned}$$

□

Per aplicar el criteri d'extrem relatiu precedent cal saber decidir quan una forma quadràtica associada a una matriu simètrica és definida positiva. En dimensió 2 hi ha un criteri molt senzill.

Proposició. *La forma quadràtica associada a la matriu*

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix}$$

és definida positiva si i només si $a > 0$ i el determinant $ad - b^2 > 0$; és definida negativa si i només si $a < 0$ i el determinant $ad - b^2 > 0$.

Demostració. Suposem primer que $a > 0$ i $ad - b^2 > 0$. Llavors, per $h \neq 0$,

$$\left\langle h, \begin{pmatrix} a & b \\ b & d \end{pmatrix} h \right\rangle = ah_1^2 + 2bh_1h_2 + dh_2^2. \tag{2.13}$$

Si $h_2 = 0$, això és $ah_1^2 > 0$. Si $h_2 \neq 0$, l'expressió precedent és

$$h_2^2(aq^2 + 2bq + d),$$

on $q = \frac{h_1}{h_2}$. Com que el discriminant del polinomi en q és $4b^2 - 4ad < 0$, resulta que el polinomi és sempre positiu. Si la matriu és definida positiva, agafant $h = (1, 0)$ obtenim

$$0 < \left\langle (1,0), \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\rangle = a.$$

De (2.13) amb $h_1 = h$ i $h_2 = 1$ obtenim que

$$0 < ah^2 + 2bh + d, \quad h \in \mathbb{R}.$$

Doncs el discriminant és negatiu:

$$4b^2 - 4ad < 0.$$

L'altra meitat de l'enunciat queda com a exercici. □

Nota. Per recordar el resultat penseu en matrius diagonals del tipus

$$\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}.$$

La forma quadràtica és $ah_1^2 + dh_2^2$, que és definida positiva si i només si $a > 0$ i $d > 0$ o, equivalentment, si $a > 0$ i el determinant $ad > 0$. És definida negativa si i només si $a < 0$ i $d < 0$ o, equivalentment, si $a < 0$ i el determinant $ad > 0$.

Exemple. Volem determinar els extrems locals de $f(x,y) = x^2 - y^2 + xy$ al pla. Calculant s'obté

$$\nabla f(x,y) = (2x + 2y, -2y + x).$$

Només l'origen és un punt crític. El hessià a l'origen és

$$\begin{pmatrix} 2 & 1 \\ 1 & -2 \end{pmatrix}$$

que té determinant negatiu. Així que la hessiana no és definida i no es pot aplicar el criteri de les derivades segones. Hem de fer un estudi particular de la nostra funció.

$$\begin{array}{lll} \text{Per } y = 0, & f(x,y) = x^2 > 0 & \text{si } x \neq 0. \\ \text{Per } x = 0, & f(x,y) = -y^2 < 0 & \text{si } y \neq 0. \end{array}$$

Per tant f no té un extrem local a l'origen.

Ara volem calcular els extrems absoluts de la funció al disc $K = \{(x,y) : x^2 + y^2 \leq 1\}$. Sabem que en té perquè f és contínua i K és compacte. També sabem que són a la frontera de K, perquè un extrem absolut a K de l'interior seria un extrem local i ja hem vist que no n'hi ha. La frontera de K és la circumferència de centre l'origen i radi 1 que podem parametritzar per

$$\begin{aligned} [0, 2\pi] &\longrightarrow \partial K \\ \theta &\longrightarrow (\cos\theta, \sin\theta). \end{aligned}$$

Posem

$$g(\theta) = f(\cos\theta, \sin\theta) = \cos(2\theta) + \frac{1}{2}\sin(2\theta).$$

Busquem els extrems absoluts de g. Tenim que

$$g'(0) = -2\sin(2\theta) + \cos(2\theta).$$

Els punts crítics són els θ tals que $\operatorname{tg}(2\theta) = \frac{1}{2}$ i la funció en aquests punts val

$$g(\theta) = \frac{5}{4}\cos(2\theta).$$

Ara recordem que $\cos t = \pm\frac{1}{\sqrt{1+\operatorname{tg}^2 t}}$. Obtenim $\cos(2\theta) = \pm\frac{2}{\sqrt{5}}$.

El màxim de f a K és $\frac{\sqrt{5}}{2}$ i el mínim $-\frac{\sqrt{5}}{2}$.

2.4 Extrems condicionats

Sovint ens trobem amb problemes d'extrems per una funció sobre el subconjunt de punts del domini de la funció que compleixen una certa condició. Per exemple, suposem que volem trobar el rectangle d'àrea màxima que té perímetre 8. Si els costats del rectangle són x i y, l'àrea del rectangle és $f(x,y) = xy$. La condició que el perímetre sigui 8 es tradueix per $x + y = 4$. Així que hem de fer màxima la funció $f(x,y) = xy$ sobre el conjunt de punts del pla que compleixen $x + y = 4$, que resulta ser una recta. Com que una recta no és un compacte no queda clar ni que aquest màxim existeixi. Com procedim?

Aïllem $y = 4 - x$ i substituim a $f(x,y) = xy = x(4-x)$. Ara hem de fer màxima la funció d'una variable

$$g(x) = x(4-x), \quad 0 < x < \infty.$$

Derivant $g'(x) = 4 - 2x$ que s'anul·la només per $x = 2$. Notem que $y = g(x)$ és una paràbola que pren el valor màxim al punt 2. Doncs el rectangle d'àrea màxima que té perímetre 8 és el quadrat de costat 2.

El problema general a $\mathbb{R}^n$ es planteja així: donades dues funcions $f, g \in C^1(U)$, U obert de $\mathbb{R}^n$, es volen trobar els extrems de f amb la condició $g(x) = 0$. Més formalment, posem

$$S = \{x \in U : g(x) = 0\}.$$

Llavors volem trobar els extrems de la restricció $f_{|S}$ de f a S.

Teorema dels multiplicadors de Lagrange. *Amb f i g com abans, suposem que $f_{|S}$ té un extrem local al punt $a \in S$ i que $\nabla g(a) \neq 0$. Llavors hi ha $\lambda \in \mathbb{R}$ tal que*

$$\nabla f(a) = \lambda \, \nabla g(a).$$

Nota. Que $f_{|S}$ tingui un màxim (mínim) local al punt $a \in S$ vol dir que hi ha $r > 0$ tal que si $x \in S$ i $\|x - a\| < r$, llavors

$$f(x) \leq f(a) \quad (f(x) \geq f(a)).$$

Exemple. Trobeu els extrems de $f(x,y) = xy$ amb la condició $x^2 + y^2 = 1$. Plantegem el sistema de 3 equacions en les incògnites x, y λ

$$\begin{cases} \nabla f(x,y) = \lambda \nabla g(x,y) \\ x^2 + y^2 = 1 \end{cases}$$

que és

$$\begin{cases} y = 2\lambda x \\ x = 2\lambda y \\ x^2 + y^2 = 1. \end{cases}$$

Fixem-nos que ni x ni y poden ser zero. Llavors podem eliminar λ

$$2\lambda = \frac{y}{x} = \frac{x}{y}.$$

Obtenim $y^2 = x^2$ i per la condició

$$\begin{aligned} 2x^2 &= 1 \\ x &= \pm\frac{1}{\sqrt{2}}. \end{aligned}$$

Obtenim els 4 punts

$$\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \left(\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right), \quad \left(\frac{-1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \left(\frac{-1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}\right).$$

En el primer i l'últim la funció val $\frac{1}{2}$ i en els altres dos $-\frac{1}{2}$. Com que els extrems absoluts existeixen perquè f és contínua i la circumferència unitat és compacta, el màxim de f és $\frac{1}{2}$ i el mínim $-\frac{1}{2}$ i es prenen en els punts indicats.

Demostració informal del teorema dels multiplicadors de Lagrange. Ja hem vist que $\nabla g(a)$ és un vector ortogonal a S en el punt a. Sigui $c\colon (t_0-\varepsilon, t_0+\varepsilon) \to S$ una corba diferenciable tal que $c(t_0) = a$ i $c'(t_0) \neq 0$. Llavors la funció $f(c(t))$ té un extrem local al punt t_0 i, doncs,

$$0 = \frac{d}{dt} f(c(t))\bigg|_{t=t_0} = \langle \nabla f(c(t_0)), c'(t_0) \rangle,$$

que diu que $\nabla f(a)$ és ortogonal a $c'(t_0)$. Doncs $\nabla f(a)$ és ortogonal a S. Concloem que $\nabla f(a)$ i $\nabla g(a)$ estan sobre la mateixa recta i, per tant, un és un múltiple de l'altre. $\square$

Nota. No queda justificat per quina raó existeix la corba $c(t)$ que passa pel punt $a = c(t_0)$ i té vector tangent $c'(t_0) \neq 0$.

Exemple. Trobeu la caixa de sabates de volum màxim que té superfície 10. Si x, y, z són els costats, volem maximitzar $f(x, y, z) = xyz$ amb la condició $xy + xz + yz = 5$. Considerem el sistema

$$\begin{cases} \nabla f = \lambda \nabla g \\ xy + xz + yz = 5 \end{cases}$$

que és

$$\begin{cases} yz = \lambda(y+z) \\ xz = \lambda(x+z) \\ xy = \lambda(x+y) \\ xy + xz + yz = 5. \end{cases}$$

Si $x = 0$ la segona equació dóna $z = 0$ o $\lambda = 0$. Si $\lambda = 0$ la primera equació dóna $yz = 0$ i l'última $0 = 5$. Per tant $\lambda \neq 0$ i ha de ser $z = 0$, però l'última donaria $0 = 5$ una altra vegada. La conclusió és que $x \neq 0$. Com que el sistema és simètric concloem que també $y \neq 0$ i $z \neq 0$. Llavors elimino λ fent

$$\frac{1}{\lambda} = \frac{y+z}{yz} = \frac{x+z}{xz} = \frac{x+y}{xy}.$$

De la segona igualtat obtinc $x = y$. Per simetria $x = y = z$. Ara la condició dóna $x = y = z = \sqrt{\frac{5}{3}}$ i el valor màxim hauria de ser $\left(\sqrt{\frac{5}{3}}\right)^3$. Però en cap moment hem demostrat que el màxim existeix i, per tant, l'única cosa que podem concloure és que el màxim si existeix és $\left(\sqrt{\frac{5}{3}}\right)^3$. Per veure que el màxim existeix raonem de la forma següent.

Noteu que $xz \leq xy + xz + yz = 5$ i doncs $x \leq \frac{5}{z}$. Igualment $y \leq \frac{5}{z}$. Doncs $xyz \leq \frac{5}{z}\frac{5}{z}z = \frac{25}{z}$, de manera que si $z \geq 25$, tenim $xyz \leq 1$. Concloem que el màxim de xyz a S es pren al quadrat

$$Q = \{(x,y,z) \in \mathbb{R}^3 : x \leq 25,\ y \leq 25 \text{ i } z \leq 25\}.$$

Ara $Q \cap S$ és un compacte i, per tant, xyz té un màxim a $Q \cap S$. Aquest màxim és el màxim de xyz a S.

També hi ha un teorema dels multiplicadors de Lagrange quan hi ha més d'una condició. Suposem que f i $g_1, \dots, g_m$, $1 \leq m < n$, són funcions de classe $C^1(U)$, U obert $\subset \mathbb{R}^n$. Posem

$$S = \{x \in U : g_1(x) = \cdots = g_m(x) = 0\}.$$

Sigui a un punt de S on $f_{|S}$ té un extrem local i suposem que els vectors $\nabla g_1(a), \dots, \nabla g_m(a)$ són linealment independents. Llavors hi ha $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tals que

$$\nabla f(a) = \lambda_1 \nabla g_1(a) + \cdots + \lambda_m \nabla g_m(a).$$

Exemple. Trobeu la distància de la recta $x + y = 1$, $z - x - y = 0$ a l'origen. Prenem la funció (quadrat de la distància a l'origen) $f(x,y,z) = x^2 + y^2 + z^2$ i fem-la mínima amb les condicions $x + y = 1$, $z - x - y = 0$. Calculant els gradients trobem

$$(2x, 2y, 2z) = \lambda_1(1,1,0) + \lambda_2(-1,-1,1),$$

o sigui

$$\begin{aligned} 2x &= \lambda_1 - \lambda_2 \\ 2y &= \lambda_1 - \lambda_2 \\ 2z &= \lambda_2. \end{aligned}$$

Doncs $x = y$ i, de la primera condició, $x = \frac{1}{2}$. La segona dóna $z = 1$. Així que tenim l'únic candidat a extrem $\left(\frac{1}{2}, \frac{1}{2}, 1\right)$. Aquest candidat és un mínim absolut per consideracions geomètriques. Així que la distància es fa mínima en aquest punt i és $\sqrt{\frac{3}{2}}$.

2.5 El teorema del valor mitjà

Sigui $f\colon B \to \mathbb{R}$ una funció de classe C^1 a una bola oberta B (de fet, tot el que farem val si B és un obert convex). Prenem dos punts $x, y \in B$ i apliquem el teorema del valor mitjà a la funció $g(t) = f(x + t(y - x))$, $0 \le t \le 1$. Tenim que, per un cert punt z del segment $[x, y]$,

$$f(y) - f(x) = \langle \nabla f(z), y - x \rangle = Df(z)(y - x).$$

Això no val si la funció f pren valors vectorials.

Exemple. Sigui $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, $f(x, y) = (e^x \cos y, e^x \sin y)$. Llavors $f(0, 0) = (1, 0) = f(0, 2\pi)$.

La diferencial de f al punt (x, y) és

$$Df(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$$

que té determinant $e^{2x} \neq 0$. Doncs, la identitat

$$0 = f(0, 2\pi) - f(0, 0) = Df(x, y)\big((0, 2\pi)\big)$$

no pot valdre perquè $Df(x, y)$ té nucli zero en qualsevol punt (x, y).

Per funcions vectorials el teorema del valor mitjà s'ha d'enunciar en forma de desigualtat i per fer-ho amb precisió necessitem la noció de norma d'una aplicació lineal. Sabem que una aplicació lineal $L\colon \mathbb{R}^n \to \mathbb{R}^m$ és contínua. Definim

$$\|L\| = \sup_{\|x\|=1} \|L(x)\|. \tag{2.14}$$

Aquest suprem existeix perquè la funció $x \to \|L(x)\|$ és contínua i doncs és fitada a l'esfera unitat de $\mathbb{R}^n$ que és un compacte. De fet el suprem és un màxim.

Fixem-nos que si $x \neq 0$

$$\|L(x)\| = \left\| L\left(\|x\| \frac{x}{\|x\|}\right) \right\| = \|x\| \left\| L\left(\frac{x}{\|x\|}\right) \right\| \leq \|L\| \, \|x\|.$$

Si C és una constant que compleix

$$\|L(x)\| \leq C\|x\|, \quad x \in \mathbb{R}^n, \tag{2.15}$$

llavors $\|L\| = \sup\limits_{\|x\|=1} \|L(x)\| \leq C$, i, per tant, $\|L\|$ és la mínima constant que compleix (2.15):

$$\|L\| = \min\{C : \text{val } (2.15)\}. \tag{2.16}$$

La caracterització (2.16) de $\|L\|$ és tant útil com la definició (2.14).

Lema. *L'aplicació $L \to \|L\|$ és una norma a l'espai vectorial $L(\mathbb{R}^n, \mathbb{R}^m)$ de les aplicacions lineals de $\mathbb{R}^n$ en $\mathbb{R}^m$.*

Demostració. És un exercici.

En particular val la següent conseqüència de la desigualtat triangular

$$|\|L\| - \|M\|| \leq \|L - M\|, \quad L, M \in L(\mathbb{R}^n, \mathbb{R}^m),$$

que diu que $L \to \|L\|$ és una funció contínua a $L(\mathbb{R}^n, \mathbb{R}^m)$, dotat de la distància $\|L - M\|$ determinada per la norma. Observem que si $L = (a_{ij})$ és la matriu en les bases canòniques de $\mathbb{R}^n$ i $\mathbb{R}^m$ tenim

$$\|L\| \leq \left(\sum_{j,k} a_{kj}^2 \right)^{\frac{1}{2}}.$$

En efecte,

$$L(x) = \left(\sum_{j=1}^{n} a_{1j} x_j, \ldots, \sum_{j=1}^{n} a_{mj} x_j \right),$$

$$\begin{aligned}
\|L(x)\| &= \left(\left(\sum_{j=1}^{n} a_{1j} x_j \right)^2 + \cdots + \left(\sum_{j=1}^{n} a_{mj} x_j \right)^2 \right)^{\frac{1}{2}} \\
&\leq \left(\sum_{j=1}^{n} a_{1j}^2 \sum_{j=1}^{n} x_j^2 + \cdots + \sum_{j=1}^{n} a_{mj}^2 \sum_{j=1}^{n} x_j^2 \right)^{\frac{1}{2}} \\
&= \left(\sum_{j=1}^{n} \sum_{k=1}^{m} a_{kj}^2 \right)^{\frac{1}{2}} \|x\|,
\end{aligned}$$

per la desigualtat de Schwarz. Llavors, si dues matrius $L = (a_{ij})$ i $M = (b_{ij})$ tenen tots els coeficients propers, diguem $|a_{ij} - b_{ij}| < \varepsilon$, per qualssevol i, j, les aplicacions lineals L i M són properes en la mètrica de $L(\mathbb{R}^n, \mathbb{R}^m)$ perquè

$$\begin{aligned}
\|L - M\| &\leq \left(\sum_{j,k} (a_{kj} - b_{kj})^2 \right)^{\frac{1}{2}} \\
&\leq \varepsilon \sqrt{nm}.
\end{aligned} \tag{2.17}$$

Recíprocament, si L i M són aplicacions lineals properes en la mètrica de $L(\mathbb{R}^n, \mathbb{R}^m)$, llavors a_{j_k} i b_{j_k} són propers per a tots els índexs j i k. Això es veu notant que

$$L(e_k) = (a_{1k}, \ldots, a_{mk})$$

i doncs

$$|a_{j_k}| \leq \left(\sum_{j=1}^{m} a_{j_k}^2 \right)^{\frac{1}{2}} = \|L(e_k)\| \leq \|L\|.$$

Aplicant-ho a la diferència $L - M$ obtenim

$$|a_{j_k} - b_{j_k}| \leq \|L - M\|, \text{ per a tot } j \text{ i } k. \qquad \square$$

Teorema del valor mitjà. *Sigui $B \subset \mathbb{R}^n$ una bola oberta (de fet, B pot ser un obert convex) i $f\colon B \to \mathbb{R}^m$ una funció de classe $C^1(B)$. Llavors, si $x, y \in B$, tenim*

$$\|f(y) - f(x)\| \le \|Df(z)\| \, \|y - x\|$$

per cert $z \in [x, y]$.

Demostració. Posem $b = \dfrac{f(y) - f(x)}{\|f(y) - f(x)\|}$ (si $f(y) = f(x)$ no hi ha res per demostrar). Llavors

$$\|f(y) - f(x)\| = \langle f(y) - f(x), b\rangle = \sum_{k=1}^{m} (f_k(y) - f_k(x)) b_k.$$

Aplicant el teorema fonamental del càlcul obtenim

$$\begin{aligned} f_k(y) - f_k(x) = \int_0^1 \frac{d}{dt} f_k(x + t(y - x))\, dt &= \int_0^1 \sum_{j=1}^{n} \partial_j f_k(x + t(y - x))(y_j - x_j)\, dt \\ &= \int_0^1 Df_k(x + t(y - x))(y - x)\, dt. \end{aligned}$$

Per tant

$$\|f(y) - f(x)\| = \int_0^1 [Df(x + t(y - x))(y - x)]_k b_k \, dt,$$

on hem indicat per $[Df(x + t(y - x))(y - x)]_k$ la component k-èsima del vector de $\mathbb{R}^m$ $Df(x + t(y - x))(y - x)$. Doncs

$$\begin{aligned} \|f(y) - f(x)\| &= \int_0^1 \langle Df(x + t(y - x))(y - x), b\rangle \, dt \\ &\le \int_0^1 |\langle Df(x + t(y - x))(y - x), b\rangle| \, dt \\ &\le \text{(Per Schwarz)} \\ &\quad \int_0^1 \|Df(x + t(y - x))(y - x)\| \, \|b\| \, dt \\ &\le \text{(Per la definició de norma d'una aplicació lineal)} \\ &\quad \int_0^1 \|Df(x + t(y - x))\| \, \|y - x\| \, dt \\ &\le \sup_{0 \le t \le 1} \|Df(x + t(y - x))\| \, \|y - x\|. \end{aligned}$$

Com que $t \to \|Df(x + t(y - x))\|$ és contínua a $[0, 1]$,

$$\sup_{0 \le t \le 1} \|Df(x + t(y - x))\| = \|Df(z)\|$$

per cert $z \in [x, y]$. □

2.5.1 El teorema de la funció inversa

Necessitem dos lemes.

Lema 1. *Sigui $f: U \to \mathbb{R}^m$, U obert $\subset \mathbb{R}^n$, $f \in C^1(U)$. Donats $a \in U$ i $\varepsilon > 0$ hi ha $B(a, r) \subset U$ tal que*

$$\|f(y) - f(x)\| \le (\|Df(a)\| + \varepsilon)\|y - x\|, \quad x, y \in B(a, r).$$

Demostració. Pel teorema del valor mitjà, si $x, y \in B(a, r)$

$$\|f(y) - f(x)\| \le \|Df(z)\| \, \|y - x\|$$

per cert $z \in [x, y]$. Si prenem $r > 0$ prou petit

$$\|Df(z) - Df(a)\| < \varepsilon, \quad \|z - a\| < r,$$

per la continuïtat de les derivades parcials i per (2.17). □

Lema 2. *Sigui $f: U \to \mathbb{R}^n$, U obert $\subset \mathbb{R}^n$, $f \in C^1(U)$. Suposem que per cert $a \in U$ el jacobià de f al punt a no s'anul·la:*

$$Jf(a) = \det Df(a) \neq 0.$$

Llavors hi ha una bola $B(a, r) \subset U$ i una constant $c > 0$ tals que

$$\|f(y) - f(x)\| \ge c\|y - x\|, \quad y, x \in B(a, r).$$

En particular, f és injectiva a $B(a, r)$.

Demostració. Sigui $L = Df(a)$. Per hipòtesi L té una inversa L^{-1} que també és lineal. Tenim

$$\|u - v\| = \|L^{-1}(L(u) - L(v))\| \le \|L^{-1}\| \, \|L(u) - L(v)\|. \tag{2.18}$$

Posem

$$f(x) = f(a) + Df(a)(x-a) + R(x), \quad x \in U.$$

Llavors

$$f(y) - f(x) = Df(a)(y-x) + R(y) - R(x)$$

i, per (2.18),

$$\begin{aligned} \|f(y) - f(x)\| &\geq \|Df(a)(y-x)\| - \|R(y) - R(x)\| \\ &\geq \frac{1}{\|Df(a)^{-1}\|}\|y-x\| - \|R(y) - R(x)\|. \end{aligned} \tag{2.19}$$

Ara

$$R(x) = f(x) - f(a) - Df(a)(x-a)$$

i, doncs,

$$DR(x) = Df(x) - Df(a) \quad \text{i} \quad DR(a) = 0.$$

Pel Lema 1 aplicat a $R(x)$, donat $\varepsilon = \frac{1}{2\|Df(a)^{-1}\|}$ hi ha una bola $B(a,r) \subset U$ tal

$$\|R(y) - R(x)\| \leq \varepsilon\|y-x\|, \quad x, y \in B(a,r).$$

Així que, per (2.19),

$$\|f(y) - f(x)\| \geq \frac{1}{2\|Df(a)^{-1}\|}\|y-x\|, \quad x, y \in B(a,r),$$

i definint $c = \frac{1}{2\|Df(a)^{-1}\|}$ la demostració queda completa. □

Teorema de la funció inversa. *Sigui $f\colon U \to \mathbb{R}^n$, U obert de $\mathbb{R}^n$, $f \in C^1(U)$ i $a \in U$ amb $Jf(a) \neq 0$. Llavors hi ha una bola $B = B(a,r) \subset U$ tal que*

(1) f és injectiva a B.

(2) $f(B) = V$ és un obert de $\mathbb{R}^n$.

(3) La inversa de f a B, $g\colon V \to B$ és diferenciable amb continuïtat a V.

Cal fer uns comentaris i considerar uns exemples. Suposem que la funció f de l'enunciat del teorema de la funció inversa compleix que $Jf(x) \neq 0$, $x \in U$. Llavors f és una aplicació oberta (és a dir, si W és un obert contingut a U, llavors $f(W)$ és obert), com es veu fàcilment pel teorema, però f no és necessàriament invertible globalment (és a dir, f no és injectiva). Això és sorprenent perquè en una variable, si una funció diferenciable té derivada no nul·la a un interval, és injectiva en aquest interval. Això és pel teorema del valor mitjà.

Exemple 12. Sigui la funció
$$\begin{aligned} \mathbb{C} &\xrightarrow{f} \mathbb{C} \\ z &\longrightarrow z^2. \end{aligned}$$
Posant $z = x + iy$ trobem que
$$\begin{aligned} f\colon \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longrightarrow (x^2 - y^2, 2xy). \end{aligned}$$
La diferencial és
$$Df(x,y) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$
i el jacobià
$$Jf(x,y) = 4x^2 + 4y^2 \neq 0 \text{ a } \mathbb{R}^2 \setminus \{(0,0)\}.$$
Ara f no és injectiva a $U = \mathbb{R}^2 \setminus \{(0,0)\}$ perquè
$$f(1,0) = f(-1,0) = (1,0).$$
Comprovem que f té una inversa a $B((1,0),1)$. Posem
$$u = x^2 - y^2, \quad v = 2xy.$$
Operant trobem
$$x = \frac{1}{\sqrt{2}}\sqrt{\sqrt{u^2+v^2} - u}$$
$$y = \frac{1}{\sqrt{2}}\frac{v}{\sqrt{\sqrt{u^2+v^2} - u}},$$
que és la inversa.

Exemple 13. Sigui la funció
$$\begin{aligned} \mathbb{C} &\xrightarrow{f} \mathbb{C} \\ z &\longrightarrow e^z \end{aligned}$$
o
$$\begin{aligned} f\colon \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longrightarrow (e^x \cos y, e^x \sin y). \end{aligned}$$

La diferencial és

$$Df(x,y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}$$

i el jacobià

$$Jf(x,y) = e^{2x} \neq 0, \quad (x,y) \in \mathbb{R}^2.$$

La funció f no és injectiva perquè

$$f(x, y + 2\pi) = f(x,y).$$

A l'obert $\{(x,y) : |y| < \pi/2\}$, f té una inversa, perquè posant

$$u = e^x \cos y, \quad v = e^x \sin y$$

trobem $u^2 + v^2 = e^{2x}$

$$x = \log \sqrt{u^2 + v^2}$$

i

$$\frac{v}{u} = \operatorname{tg} y, \quad y = \operatorname{arctg}\left(\frac{v}{u}\right).$$

Observem que $f(\{(x,y) : |y| < \frac{\pi}{2}\}) = \{(u,v) : u > 0\}$.

Exemple 14. Posem

$$f(x,y) = (\cos h(x) \cos y, \sin hx \sin y).$$

Llavors

$$Jf(x,y) = \sin h^2 x + \sin^2 y.$$

Al semiplà de la dreta $\{(x,y) : x > 0\}$ el jacobià no s'anul·la, però f no és injectiva.

Demostració del teorema de la funció inversa. (1) és el lema 2.

Per (2) prenem $y_0 \in V = f(B)$. He de veure que hi ha una bola $B(y_0, \rho) \subset f(B)$. Sigui $x_0 \in B$ tal que $f(x_0) = y_0$ i sigui $r_0 > 0$ tal que $\overline{B(x_0, r_0)} \subset B$. Llavors si $\|x - x_0\| = r$ el lema 2 dóna

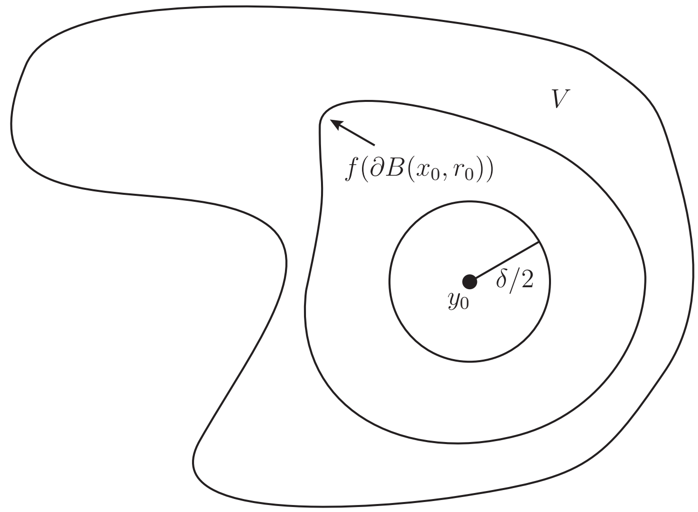
$$\|f(x) - f(x_0)\| \geq c\,\|x - x_0\| = c\,r_0 \equiv \delta > 0.$$

Veiem que

$$B\left(y_0, \frac{\delta}{2}\right) \subset f(B)$$

(de manera que $\rho = \frac{\delta}{2}$).

Agafem $y \in B(y_0, \frac{\delta}{2})$.



La funció $h(x) = \|f(x) - y\|^2$ pren el seu mínim a la bola $\overline{B(x_0, r_0)}$ en un punt de la bola oberta $B(x_0, r_0)$. En efecte si $\|x - x_0\| = r_0$ tenim que

$$\begin{aligned}\|f(x) - y\| &\geq \|f(x) - f(x_0)\| - \|f(x_0) - y\| \\ &\geq \delta - \|y_0 - y\| > \delta - \frac{\delta}{2} = \frac{\delta}{2} \\ &> \|y_0 - y\| = \|f(x_0) - y\|.\end{aligned}$$

Així que el mínim de h a l'esfera $\partial B(x_0, r_0)$ és estrictament més gran que el mínim de h a $B(x_0, r_0)$. Sigui $x \in B(x_0, r_0)$ un punt on es pren el mínim de h a la bola $\overline{B(x_0, r_0)}$. Doncs x és un mínim local de h i

$$0 = \nabla h(x) = \left(\dots, 2 \sum_{k=1}^{n} (f_k(x) - y_k) \partial_j f_k(x), \dots \right)$$

o sigui

$$\sum_{k=1}^{n} (f_k(x) - y_k) \partial_j f_k(x) = 0, \quad 1 \leq j \leq n. \tag{2.20}$$

Ara, (2.20) és un sistema lineal en les incògnites $z_k = f_k(x) - y_k$, i coeficients $\partial_j f_k(x)$, $1 \leq j$, $k \leq n$. Puc agafar el radi r de la bola $B = B(a, r)$ tan petit que $Jf(x) \neq 0$ a B. Per tant el sistema té determinant no nul i l'única solució és

$$0 = z_k = f_k(x) - y_k, \quad 1 \leq k \leq n.$$

Doncs $y = f(x)$ i
$$B\left(y_0, \frac{\delta}{2}\right) \subset f(B).$$

Veiem ara (3): la inversa $g\colon V \to B$ de f és diferenciable amb continuïtat a V. Veurem que $D(g)(y_0) = Df(x_0)^{-1}$ si $f(x_0) = y_0 \in V$. Posant $L = Df(x_0)$ i $x = g(y)$ obtenim

$$\begin{aligned} g(y) - g(y_0) - L^{-1}(y - y_0) &= x - x_0 - L^{-1}(f(x) - f(x_0)) \\ &= L^{-1}(L(x - x_0) - (f(x) - f(x_0))) \\ &= -L^{-1}(f(x) - f(x_0) - L(x - x_0)). \end{aligned}$$

Pel lema 2
$$c\|x - x_0\| \leq \|f(x) - f(x_0)\|$$
o
$$\|x - x_0\| \leq c^{-1}\|y - y_0\|,$$
i, per tant, $x \to x_0$ si $y \to y_0$. Doncs
$$\frac{\|g(y) - g(y_0) - L^{-1}(y - y_0)\|}{\|y - y_0\|} \leq \frac{\|L^{-1}\|}{c} \frac{\|f(x) - f(x_0) - L(x - x_0)\|}{\|x - x_0\|}$$
té límit 0 si $y \to y_0$. □

2.6 La funció implícita

Si tenim una equació del tipus
$$f(x, y) = 0, \quad x, y \in \mathbb{R},$$
esperem que, en condicions favorables, el conjunt de solucions
$$\{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}$$
sigui una “corba” i que sobre aquesta corba es pugui aïllar y en termes de x, $y = g(x)$, o x en termes de y, $x = h(y)$, per certes funcions g i h.

Exemple. Considerem l’equació

$$x^2 + y^2 = 1, \quad x, y \in \mathbb{R}. \tag{2.21}$$

Sigui (x_0, y_0) una solució de (2.21) amb $-1 < x_0 < 1$, i posem $\delta = 1-|x_0|$. Llavors si (x, y) és una solució de (2.21) amb $x_0 - \delta < x < x_0 + \delta$, tenim que $y = \sqrt{1-x^2}$ si $y_0 > 0$ i $y = -\sqrt{1-x^2}$ si $y_0 < 0$. Per tant, independentment del signe de y_0, podem aïllar y en termes de x a

$$B((x_0, y_0), \delta) \cap S, \text{ on } S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

Si el punt de S és $(1, 0)$, llavors $(x, y) \in B((1, 0), 1) \cap S$ si i només si $x = \sqrt{1-y^2}$.

Si el punt de S és $(-1, 0)$, llavors $(x, y) \in B((-1, 0), 1)$ si i només si $x = -\sqrt{1-y^2}$.

La conclusió és que sempre puc aïllar una variable en termes de l'altra a un entorn d'un punt $(x_0, y_0) \in S$ donat.

Exemple. No sempre es pot aïllar una coordenada en termes de l'altra. Per exemple si

$$S = \{(x, y) \in \mathbb{R}^2 : xy = 0\}$$

i el punt és l'origen.

El teorema de la funció implícita dóna una condició suficient per tal que d'una equació del tipus

$$f(x_1, \dots, x_n) = 0$$

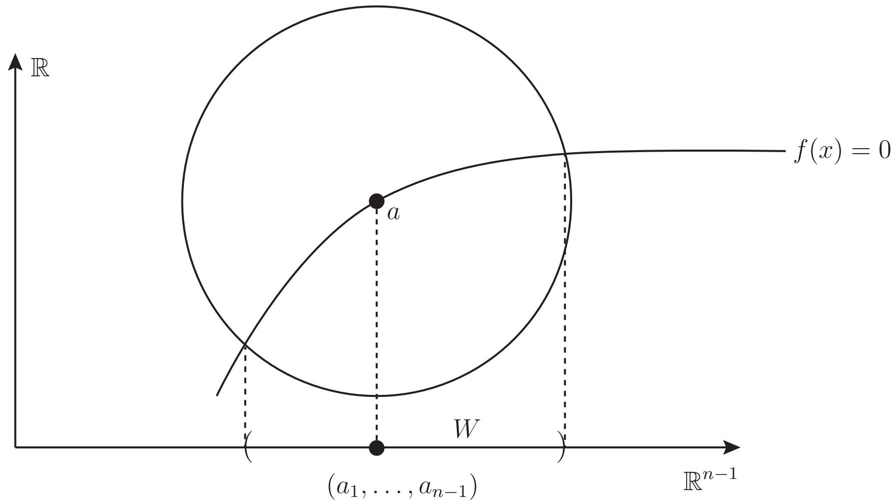
es pugui aïllar una variable en termes de les altres $n - 1$.

Teorema de la funció implícita. *Sigui U obert $\subset \mathbb{R}^n$, $f\colon U \to \mathbb{R}$ una funció de classe $C^1(U)$ i $a \in U$ tal que $f(a) = 0$ i $\frac{\partial f}{\partial x_n}(a) \neq 0$. Llavors hi ha una bola $B(a, r) \subset U$, un obert $W \subset \mathbb{R}^{n-1}$ que conté el punt $(a_1, \dots, a_{n-1})$ i una funció $g\colon W \to \mathbb{R}$, $g \in C^1(W)$, tals que*

$$x \in B(a, r) \text{ i } f(x) = 0$$

si i només si

$$(x_1, \dots, x_{n-1}) \in W \text{ i } x_n = g(x_1, \dots, x_{n-1}).$$



Nota. Naturalment que si f compleix les hipòtesis del teorema excepte l'última ($\frac{\partial f}{\partial x_n}(a) \neq 0$), però $\frac{\partial f}{\partial x_j}(a) \neq 0$, $1 \leq j \leq n$, llavors puc aïllar $x_j = g(x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_j)$ en funció de les altres $n-1$ variables.

Demostració. Definim

$$\begin{aligned} F\colon U &\longrightarrow \mathbb{R}^n \\ x &\longrightarrow F(x) = (x_1, \ldots, x_{n-1}, f(x)), \end{aligned}$$

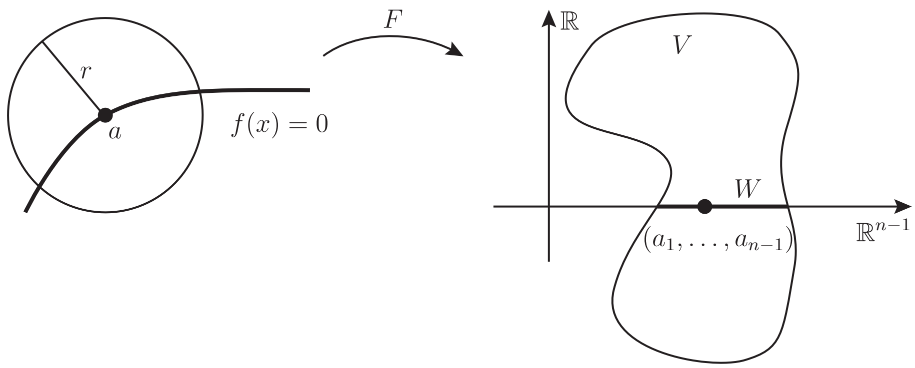
que és de classe $C^1(U)$ i té per diferencial

$$\begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & 1 & \vdots \\ \partial_1 f(x) & \partial_2 f(x) & \ldots & \ldots & \partial_n f(x) \end{pmatrix}$$

i per jacobià

$$JF(x) = \frac{\partial f}{\partial x_n}(x).$$

Com que $JF(a) \neq 0$ podem aplicar el teorema de la funció inversa i existeix una bola $B(a,r) \subset U$ tal que $F_{|_{B(a,r)}}$ té una inversa G definida a $V = F(B(a,r))$.



Posem $W = \{(x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1} : (x_1, \ldots, x_{n-1}, 0) \in V\}$. És un exercici comprovar que W és obert a $\mathbb{R}^{n-1}$ (això és degut al fet que V és obert a $\mathbb{R}^n$). Si $x \in B(a, r)$

$$x = G(F(x)) = (G_1(x_1, \ldots, x_{n-1}, f(x)), \ldots, G_n(x_1, \ldots, x_{n-1}, f(x)))$$

que dóna

$$x_n = G_n(x_1, \ldots, x_{n-1}, f(x)).$$

Així

$$x_n = G_n(x_1, \ldots, x_{n-1}, 0)$$

si $x \in B(a, r)$ i $f(x) = 0$. Definim

$$g(x_1, \ldots, x_{n-1}) = G_n(x_1, \ldots, x_{n-1}, 0)$$

de manera que si $x \in B(a, r)$ i $f(x) = 0$ llavors $(x_1, \ldots, x_{n-1}) \in W$ i $x_n = g(x_1, \ldots, x_{n-1})$. Ara cal veure que, recíprocament, si $(x_1, \ldots, x_{n-1}) \in W$ i $x_n = g(x_1, \ldots, x_{n-1})$ llavors

$$(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) \in B(a, r) \quad \text{i} \quad f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = 0.$$

En efecte,

$$\begin{aligned}
(x_1, \ldots, x_{n-1}, 0) &= F(G(x_1, \ldots, x_{n-1}, 0)) \\
&= (G_1(x_1, \ldots, x_{n-1}, 0), \ldots, G_{n-1}(x_1, \ldots, x_{n-1}, 0), \\
&\quad f(G_1(x_1, \ldots, x_{n-1}, 0), \ldots, G_{n-1}(x_1, \ldots, x_{n-1}, 0), G_n(x_1, \ldots, x_{n-1}, 0))) \\
&= (x_1, \ldots, x_{n-1}, f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})),
\end{aligned}$$

i, per tant,

$$f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = 0.$$

Noteu també que

$$(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = G(x_1, \ldots, x_{n-1}, 0) \in B(a, r). \qquad \square$$

Hi ha una fórmula molt pràctica per calcular les derivades d'una funció implícita. Com que

$$f(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})) = 0, \quad (x_1, \dots, x_{n-1}) \in W,$$

obtenim, per la regla de la cadena,

$$\frac{\partial f}{\partial x_j}(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})) \\ + \frac{\partial f}{\partial x_n}(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))\frac{\partial g(x_1, \dots, x_{n-1})}{\partial x_j} = 0, \quad 1 \le j \le n-1.$$

Doncs

$$\frac{\partial g}{\partial x_j}(x_1, \dots, x_{n-1}) = -\frac{\dfrac{\partial f}{\partial x_j}(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))}{\dfrac{\partial f}{\partial x_n}(x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1}))}.$$

Exemple. Sigui $f(x, y) = x^2 - y^2$, de manera que $\nabla f(0, 0) = (0, 0)$. Fixem-nos que podem trobar $g(x)$ tal que $f(x, g(x)) = 0$, $x \in \mathbb{R}$. Només cal posar $g(x) = x$. Però no podem trobar $y = g(x)$ i $r > 0$ tals que

$$Z = \{(x, y) \in B((0, 0), r) : f(x, y) = 0\} = \{(x, g(x)) : -r < x < r\},$$

simplement perquè Z no és la gràfica de cap funció.

Exemple. Donada l'equació $z^3 + xz^2 + yz + 1 = 0$, podem aïllar z en funció de x, y al voltant del punt $(1, -3, 1)$? Fixem-nos que l'equació és cúbica en z i, per tant, hi ha una fórmula que expressa la solució en termes dels coeficients. Fem-ho aplicant el teorema de la funció implícita.

Posem $f(x, y, z) = z^3 + xz^2 + yz + 1$. Tenim

$$\frac{\partial f}{\partial z}(x, y, z) = 3z^2 + 2xz + y$$

i, doncs,

$$\frac{\partial f}{\partial z}(1, -3, 1) = 3 + 2 - 3 = 2 \neq 0.$$

Així que podem aïllar $z = g(x, y)$. També podem calcular les derivades

$$\frac{\partial g}{\partial x}(x, y) = -\frac{z^2}{3z^2 + 2xz + y}$$

i

$$\frac{\partial g}{\partial y}(x,y) = -\frac{z}{3z^2+2xz+y}$$

que en el punt $(1,-3)$ valen

$$\frac{\partial g}{\partial x}(1,-3) = \frac{\partial g}{\partial y}(1,-3) = -\frac{1}{2}.$$

Definició. Una hipersuperfície és un subconjunt de $\mathbb{R}^n$ del tipus

$$S = \{x \in \mathbb{R}^n : f(x) = 0\} \tag{2.22}$$

on $f \in C^1(\mathbb{R}^n)$ és tal que $\nabla f(x) \neq 0$, $x \in S$.

En particular la gràfica d'una funció $g(x_1,\dots,x_{n-1}) \in C^1(\mathbb{R}^{n-1})$ és una hipersuperfície, perquè només cal posar $f(x) = x_n - g(x_1,\dots,x_{n-1})$ i notar que $\frac{\partial f}{\partial x_n} = 1$, $x \in \mathbb{R}^n$. Ja hem definit el pla tangent al punt $a = (a_1,\dots,a_{n-1},g(a_1,\dots,a_{n-1}))$ de la gràfica d'una funció $x_n = g(x_1,\dots,x_{n-1})$ com el pla

$$x_n = g(a_1,\dots,a_{n-1}) + \sum_{j=1}^{n-1}\frac{\partial g}{\partial x_j}(a_1,\dots,a_{n-1})(x_j-a_j).$$

Com es pot definir el pla tangent a la hipersuperfície (2.22) al punt $a \in S$? Suposant que $\frac{\partial f}{\partial x_n}(a) \neq 0$, pel teorema de la funció implícita hi ha una funció $x_n = g(x_1,\dots,x_{n-1})$ definida a un entorn de $(a_1,\dots,a_{n-1})$ tal que $S \cap B(a,r)$ és la gràfica de g (per un cert $r > 0$ prou petit).

Definim el pla tangent a S al punt a com el pla tangent a la gràfica de g al punt $a = (a_1,\dots,a_{n-1},g(a_1,\dots,a_{n-1}))$. Aquest pla és

$$x_n - g(a_1,\dots,a_{n-1}) = \sum_{j=1}^{n-1}\frac{\partial g}{\partial x_j}(a_1,\dots,a_{n-1})(x_j-a_j). \tag{2.23}$$

Sabem que

$$\frac{\partial g}{\partial x_j}(a_1,\dots,a_{n-1}) = -\frac{\dfrac{\partial f}{\partial x_j}(a)}{\dfrac{\partial f}{\partial x_n}(a)}$$

i, doncs, (2.23) es converteix en

$$\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}(a)(x_j-a_j) = 0. \tag{2.24}$$

Així que (2.24) és l'hiperpla tangent a S al punt a.

Tornem al problema d'extrems condicionats per fer-ne un breu tractament rigorós amb el teorema de la funció implícita. Prenem dues funcions $f, g\colon U \to \mathbb{R}$, U obert $\subset \mathbb{R}^n$, $f, g \in C^1(U)$. Suposem que f té un extrem local al punt a amb la condició $g(x) = 0$ i que $\nabla g(a) \neq 0$. Si posem $S = \{x \in U : g(x) = 0\}$ sabem que al voltant del punt $a \in S$, S és la gràfica d'una funció de $n-1$ variables. Concretament si $\frac{\partial g}{\partial x_n}(a) \neq 0$ llavors a prop de a, S és la gràfica de $x_n = h(x_1, \dots, x_{n-1})$ on h és de classe C^1 a un obert W de $\mathbb{R}^{n-1}$ que conté el punt $(a_1, \dots, a_{n-1})$. Com que la funció

$$f(x_1, \dots, x_{n-1}, h(x_1, \dots, x_{n-1}))$$

té un extrem local al punt $(a_1, \dots, a_{n-1})$ resulta que el gradient s'anul·la al punt $(a_1, \dots, a_{n-1})$. Derivant per la regla de la cadena

$$\begin{aligned} 0 = {} & \frac{\partial f}{\partial x_j}(a_1, \dots, a_{n-1}, h(a_1, \dots, a_{n-1})) \\ & + \frac{\partial f}{\partial x_n}(a_1, \dots, a_{n-1}, h(a_1, \dots, a_{n-1}))\frac{\partial h}{\partial x_j}(a_1, \dots, a_{n-1}). \end{aligned}$$

Doncs

$$\frac{\partial f}{\partial x_j}(a) = -\frac{\partial h}{\partial x_j}(a_1, \dots, a_{n-1})\frac{\partial f}{\partial x_n}(a), \quad 1 \leq j \leq n-1.$$

Per altra banda per la regla de diferenciació implícita tenim la mateixa fórmula amb f canviada per g. Per tant

$$\frac{\partial f}{\partial x_j}(a) = -\frac{\partial h}{\partial x_j}(a_1, \dots, a_{n-1})\frac{\partial f}{\partial x_n}(a) = \frac{\dfrac{\partial g}{\partial x_j}(a)}{\dfrac{\partial g}{\partial x_n}(a)}\frac{\partial f}{\partial x_n}(a), \quad 1 \leq j \leq n-1$$

i

$$\frac{\partial f}{\partial x_n}(a) = \frac{\dfrac{\partial g}{\partial x_n}(a)}{\dfrac{\partial g}{\partial x_n}(a)}\frac{\partial f}{\partial x_n}(a).$$

Posant

$$\lambda = \frac{\dfrac{\partial f}{\partial x_n}(a)}{\dfrac{\partial g}{\partial x_n}(a)}$$

obtenim que

$$\nabla f(a) = \lambda \nabla g(a).$$

Exemple. Considerem l'equació $f(x,y,z) = z^2 + yz + x = 0$. Sabem que podem aïllar $z = \frac{1}{2}(-y \pm \sqrt{y^2 - 4x})$ sempre que $y^2 - 4x \geq 0$. Clarament

$$\frac{\partial f}{\partial z}(x,y,z) = 2z + y$$

que no s'anul·la precisament quan el discriminant $y^2 - 4x \neq 0$. Si $y_0^2 - 4x_0 > 0$, llavors el teorema de la funció implícita ens diu que podem aïllar z en termes de x, y al voltant del punt (x_0, y_0) i sabem que la funció que dóna z en termes de x, y és

$$z = g(x,y) = \frac{1}{2}\left(-y \pm \sqrt{y^2 - 4x}\right).$$

La bola centrada en (x_0, y_0) on es podrà aïllar z serà tal que valgui $y^2 - 4x > 0$.

Suposem ara que $y_0^2 - 4x_0 = 0$. Llavors

$$z^2 + y_0 z + x_0 = z^2 + y_0 z + \frac{y_0^2}{4} = \left(z + \frac{y_0}{2}\right)^2,$$
$$f\left(\frac{y_0^2}{4}, y_0, -\frac{y_0}{2}\right) = 0,$$

i també

$$\frac{\partial f}{\partial z}\left(\frac{y_0^2}{4}, y_0, -\frac{y_0}{2}\right) = 0.$$

Així que no podem aplicar el teorema de la funció implícita. Per exemple, si $x_0 = y_0 = 0$ tenim que

$$0 = f(0,0,0) = \frac{\partial f}{\partial z}(0,0,0).$$

Ara, podria ser que existís una funció $z = g(x,y)$ definida a una bola centrada a $(0,0)$ tal que

$$g(x,y)^2 + y\,g(x,y) + x = 0$$

per qualsevol (x,y) d'aquesta bola. Això és impossible, com es veu considerant punts del tipus $(x, \sqrt{x})$ amb $x > 0$ petit.

Si $y_0^2 - 4x_0 < 0$, no hi ha cap punt z que compleixi l'equació $z^2 + y_0 z + x_0 = 0$.

Fem ara, per acabar, una breu incursió en el cas general del Teorema de la funció implícita. Tenim m funcions $f_j \in C^1(U)$, U obert $\subset \mathbb{R}^n$, $1 \leq j \leq m < n$ i volem considerar l'estructura del conjunt

$$S = \{x \in U : f_1(x) = \cdots = f_m(x) = 0\}.$$

Per exemple, a $\mathbb{R}^3$ podem prendre $f_1(x,y,z) = x + y + z$ i $f_2(x,y,z) = x + 2y - z$.

El sistema lineal

$$\begin{aligned} x + y + z &= 0 \\ x + 2y - z &= 0 \end{aligned}$$

té per matriu

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & -1 \end{pmatrix}$$

que té rang 2. Resolent trobem $y = -\frac{2}{3}x$ i $z = -\frac{1}{3}x$. La conclusió és que podem expressar dues variables en termes de la tercera i el conjunt de solucions és "1-dimensional": una recta en el nostre exemple.

Com a segon exemple prenem $f(x, y, z) = z - x^2 - y^2$ i $g(x, y, z) = x^2 + y^2 + z^2 - 1$. Eliminant $x^2 + y^2$ trobem $z = \frac{-1+\sqrt{5}}{2}$ i, doncs, el conjunt de solucions S en aquest cas és la circumferència $x^2 + y^2 = \frac{-1+\sqrt{5}}{2}$ aixecada a l'altura $z = \frac{-1+\sqrt{5}}{2}$. Així que S és una corba i la conclusió és que podem aïllar dues variables en termes de la tercera. Per exemple, en un entorn del punt $(0, 1, \frac{-1+\sqrt{5}}{2})$ tenim que $(x, y, z) \in S$ si i només si $y = \sqrt{1 - x^2}$ i $z = \frac{-1+\sqrt{5}}{2}$.

Reprenem la discussió del cas general del Teorema de la funció implícita. Tenim m equacions

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0 \\ f_2(x_1, \ldots, x_n) &= 0 \\ &\vdots \\ f_m(x_1, \ldots, x_n) &= 0 \end{aligned}$$

i voldríem aïllar m variables en termes de les altres $n - m$. LLavors el conjunt S de solucions del sistema de m equacions de dalt tindrà "dimensió" $n - m$. Per exemple, dues equacions a l'espai donaran una corba, tal com hem vist en els exemples. Considerem la matriu jacobiana de $f = (f_1, \ldots, f_m) \colon U \to \mathbb{R}^m$, que és la matriu

$$\begin{pmatrix} \partial_1 f_1(x) & \ldots & \partial_n f_1(x) \\ \vdots & & \vdots \\ \partial_1 f_m(x) & \ldots & \partial_n f_m(x) \end{pmatrix}$$

amb n columnes i m files. Suposem que al punt $a \in U$ la matriu precedent té rang m.

En altres paraules, hi ha m columnes linealment independents, que sense pèrdua de generalitat, podem suposar que són les últimes m. Per tant

$$\det \begin{pmatrix} \partial_{n-m+1} f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & & \vdots \\ \partial_{n-m+1} f_m(a) & \dots & \partial_n f_m(a) \end{pmatrix} \neq 0. \tag{2.25}$$

Demostrarem que hi ha una bola $B(a,r) \subset U$, un obert W a $\mathbb{R}^{n-m}$ que conté $(a_1, \dots, a_{n-m})$ i funcions $g_{n-m+1}(x), \dots, g_n(x)$ de classe $C^1(W)$ tals que $x \in B(a,r)$ i $f_j(x) = 0$, $1 \le j \le m$ és equivalent a $(x_1, \dots, x_{n-m}) \in W$ i $x_k = g_k(x_1, \dots, x_{n-m})$, $n-m+1 \le k \le n$.

Demostració. Definim $F \colon U \to \mathbb{R}^n$ per $F(x) = (x_1, \dots, x_{n-m}, f_1(x), \dots, f_m(x))$, de manera que

$$DF(x) = \begin{pmatrix} 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \vdots & 1 & \vdots & \vdots & \vdots \\ \partial_1 f_1(x) & \dots & \dots & \partial_{n-m+1} f_1(x) & \dots & \partial_n f_1(x) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \dots \\ \partial_1 f_m(x) & \dots & \partial_{n-m+1} f_m(x) & \dots & \dots & \partial_n f_m(x) \end{pmatrix}.$$

El determinant de $DF(a)$ és (2.25) i, per tant, podem agafar $r > 0$ prou petit de manera que $F \colon B(a,r) \to V$ sigui invertible. Sigui G la inversa. Definim

$$W = \{(x_1, \dots, x_{n-m}) \in \mathbb{R}^{n-m} : (x_1, \dots, x_{n-m}, 0, \dots, 0) \in V\}$$

que és obert perquè V ho és. Com que $x = G(F(x))$ deduïm que

$$x_k = G_k(x_1, \dots, x_{n-m}, f_1(x), \dots, f_m(x)), \quad n-m+1 \le k \le m$$

i que si $x \in S \cap B(a,r)$

$$x_k = G_k(x_1, \dots, x_{n-m}, 0, \dots, 0), \quad n-m+1 \le k \le n.$$

Definim

$$g_k(x_1, \dots, x_{n-m}) = G_k(x_1, \dots, x_{n-m}, 0, \dots, 0), \quad (x_1, \dots, x_{n-m}) \in W,$$

que té sentit per la definició de W. L'única cosa que falta per comprovar és que si $(x_1, \dots, x_{n-m}) \in W$

$$f_k(x_1, \dots, x_{n-m}, g_{n-m+1}(x_1, \dots, x_{n-m}), \dots, g_n(x_1, \dots, x_{n-m})) = 0, \quad 1 \le k \le m. \tag{2.26}$$

Tenim, si $(x_1,\dots,x_{n-m}) \in W$ i posem $X=(x_1,\dots,x_{n-m},0,\dots,0) \in V$,

$$\begin{aligned}
&(x_1,\dots,x_{n-m},0,\dots,0) = F(G(X)) \\
&\quad = (G_1(X),\dots,G_{n-m}(X),f_1(G(X)),\dots,f_m(G(X))) \\
&\quad = (x_1,\dots,x_{n-m},f_1(x_1,\dots,x_{n-m},g_{n-m+1}(x_1,\dots,x_{n-m}),\dots,g_n(x_1,\dots,x_{n-m})), \\
&\qquad\qquad \dots,f_m(x_1,\dots,x_{n-m},g_{n-m+1}(x_1,\dots,x_{n-m}),\dots,g_n(x_1,\dots,x_{n-m}))).
\end{aligned}$$

Igualant les m últimes coordenades s'obté (2.26). $\square$