Càlcul Diferencial en vàries variables

En aquest capítol abordem el càlcul diferencial per a funcions $f : U \longrightarrow \mathbb{R}^m$ definides en un obert $U \subset \mathbb{R}^d$. Recordem que ser obert significa intuïtivament que U té al voltant de cada punt d-graus de llibertat (les d coordenades poden variar independentment unes de les altres en uns certs intervals) i que en particular qualsevol $a \in U$ pot ser aproximable per $x \in U, x \neq a$ en qualsevol direcció.

1 Arcs a $\mathbb{R}^m$. Longitud i integració sobre arcs

Primer considerarem el cas $d = 1$, és a dir, funcions vectorials d'una variable real; com que sovint convindrà pensar que aquesta variable és el temps utilitzarem la notació t per a la variable independent. També utilitzarem la notació γ enlloc de f. Així, γ és una aplicació

$$\gamma : (a, b) \longrightarrow \mathbb{R}^m.$$

Aqui suposarem que $m > 1$. Quan $m = 1$ i tenim una funció real $y = f(x)$ d'una variable real x la visualitzem mitjançant el gràfic, el conjunt de punts del pla de la forma $(x, f(x))$. En el nostre cas no visualitzem cap gràfic a $\mathbb{R}^{m+1}$, sinó que visualitzem tan sols el recorregut $\Gamma = \gamma((a, b))$. El gràfic $y = f(x)$ d'una funció d'una variable és el recorregut de $\gamma(x) = (x, f(x))$. Si pensem que t és temps i $\gamma(t)$ és la posició en el temps t d'un mòbil, Γ és el recorregut. Diem que Γ és un arc i γ una forma de recorre'l. Fixemnos que Γ no determina γ; si composem γ amb $\phi : (c, d) \to (a, b), t = \phi(s)$ exhaustiva, $\hat{\gamma}(s) = \gamma(\phi(s))$ té el mateix recorregut Γ. Diem que *gam̂ma* és una reparametrització de γ o que s'obté fent un canvi de paràmetre. Fixemnos que tampoc demanem que γ sigui injectiva, per exemple poden haver-hi autointerseccions. En aquest sentit el terme "paràmetre" l'estem utilitzant d'una manera diferent que a la secció ??, un mateix punt pot tenir més d'un paràmetre.

Suposarem que γ és una funció contínua en cada punt; això és el mateix que dir que cadascuna de les funcions components de γ, que denotarem $x_i(t), i = 1, \cdots, m$ són contínues. En el pla i l'espai utilitzarem

$$\gamma(t) = (x(t), y(t)), \gamma(t) = (x(t), y(t), z(t)).$$

Diem que γ és un arc (continu); si γ està definida en $[a, b]$ (és a dir,γ té límit en a, b), $A = \gamma(a), B = \gamma(b)$ diem que γ va d' A a B, o que A és el punt inicial i B el punt final. Si $A = B$ diem que tenim un *arc tancat.* Si γ és injectiva en (a, b) diem que l'arc és *simple.* Intuitivament, la continuïtat significa que podem dibuixar Γ a $\mathbb{R}^m$ sense aixecar el llapis. Ens imaginem Γ com un fil no trencat a $\mathbb{R}^m$. Tanmateix, estrictament parlant això no és així, perquè hi ha arcs continus al pla que omplen un quadrat.

Per a definir arcs podem utilitzar, és clar, altres sistemes de coordenades. Si $(u_1, \cdots, u_d)$ són coordenades en A podem donar un arc dins A posant les coordenades en termes d'un paràmetre $t \in (a, b)$

$$u_1(t), u_2(t), \cdots, u_d(t), a < t < b.$$

Per exemple, en el pla, amb coordenades polars, $r = \theta$ (que vol dir que r és el paràmetre, i les equacions són $r = r, \theta = r$) descriu una espiral.

La definició de derivada la podem copiar exactament del cas $m = 1$: diem que γ és derivable en t_0 amb derivada $\gamma'(t_0)$ si

$$\lim_{h\to 0} \frac{\gamma(t_0 + h) - \gamma(t_0)}{h} = \gamma'(t_0).$$

Això és equivalent a dir que cada component $x_i(t)$ de $\gamma(t)$ és derivable en t_0 i llavors $\gamma'(t_0) = (x_1'(t_0), \cdots, x_d'(t_0))$. Com que el quocient incremental és un vector secant a Γ, $\gamma'(t_0)$ s'interpreta com a vector tangent a Γ en el punt $\gamma(t_0)$. Evidentment, si t és temps, $\gamma'(t_0)$ significa el vector velocitat instantània en la posició $\gamma(t_0)$. Si fem un canvi de paràmetre $t = \phi(s)$, amb ϕ derivable, $\hat{\gamma}(s) = \gamma(\phi(s)), t_0 = \phi(s_0)$, per la regla de la cadena

$$\hat{\gamma}'(s_0) = \phi'(s_0)\gamma'(t_0),$$

el vector tangent es multiplica per l'escalar $\phi'(s_0)$.

Quan γ és derivable en tots els punts diem que tenim un *arc regular*. Si $\gamma'(t)$ és contínua en t (és a dir, la tangent va canviant contínuament amb t) diem que γ és un arc de classe C^1.

Seguidament ens proposem definir la longitud recorreguda per un arc continu $\gamma(t), a \leq t \leq b$. Si pensem que t és temps i $\gamma(t)$ és la posició a l'instant t d'un mòbil estem definint la distància recorreguda. Considerarem particions $P : a = t_0 < t_1 < \cdots < t_n = b$ i els corresponents punts $Q_i = \gamma(t_i)$, que determinen una poligonal. La longitud d'aquesta poligonal és

$$L(P) = \sum_{i=0}^{n} \|Q_{i+1} - Q_i\| = \sum_i \|\gamma(t_{i+1}) - \gamma(t_i)\|,$$

i és evident que si afegim punts a la partició la longitud s'incrementa. Aleshores definim la longitud recorreguda per $\gamma(t), a \leq t \leq b$ com

$$L(\gamma) = \sup_P L(P).$$

El suprem és el mateix que el límit quan la partició es va fent fina. Aquest suprem pot ser infinit o finit. Per exemple, per la corba gràfic de $y = x \sin \frac{1}{x}, 0 \leq x \leq 1$, la longitud és infinita entre els punts $(0,0)$ i qualsevol altre. Els arcs pels quals la longitud és finita s'anomenen *rectificables*.

Tot arc de classe C^1 és rectificable. Això és perquè la longitud d'una poligonal és en aquest cas, essencialment, una suma de Riemann d'una funció contínua. En efecte, si $\gamma(t) = (x_1(t), \cdots, x_m(t)), a \leq t \leq b$ i la partició està donada per la partició $P : a = t_0 < t_1 < \cdots < t_n = b$, la longitud de la poligonal és

$$L(P) = \sum_{i=0}^{n} \sqrt{\sum_{j=1}^{m} (x_j(t_{i+1}) - x_j(t_i))^2}.$$

Pel teorema del valor mig, $x_j(t_{i+1}) - x_j(t_i) = x_j'(\xi_{ij})(t_{i+1} - t_i)$, amb ξ_{ij} intermedi entre t_i, t_{i+1} i per tant

$$L(P) = \sum_{i=0}^{n} (t_{i+1} - t_i) \sqrt{\sum_{j=1}^{m} |x_j'(\xi_{ij})|^2}.$$

Si els punts ξ_{ij} no depenguessin de j, $\xi_{ij} = \xi_i$, la darrera suma seria $\|\gamma'(\xi)\|$ i $L(P)$ seria una suma de Riemann de la funció contínua $\|\gamma'\|$. Es pot veure sense gaire dificultat que l'error que representa el fet que ξ_{ij} depengui de j té limit zero quan la partició es va fent fina (això depèn de la continuitat uniforme de les funcions x_j'). Amb tot això veiem que la longitud recorreguda és

$$L = \int_a^b \|\gamma'(t)\| \, dt.$$

La interpretació és ben clara: com que $\|\gamma'(t)\|$ és la velocitat instantània, $\|\gamma'(t)\| \, dt$ és la distància recorreguda entre els instants t i $t + dt$, per tant la suma (integral) és la distància total recorreguda.

Fixem-nos que en el cas del gràfic d'una funció $y = f(x)$ amb derivada contínua, $\gamma(x) = (x, f(x))$ trobem

$$L = \int_a^b \sqrt{1 + f'(x)^2} \, dx.$$

També observem que si fem un canvi de paràmetre com abans $t = \phi(s)$, amb ϕ derivable i contínua, $\hat{\gamma}(s) = \gamma(\phi(s)), t_0 = \phi(s_0)$, llavors

$$\|\hat{\gamma}'(s)\| = |\phi'(s)| \|\gamma'(\phi(s))\|,$$

i pel teorema de canvi de variable $t = \phi(s)$ en una integral

$$\int_c^d \|\hat{\gamma}'(s)\| \, ds = \int_c^d |\phi'(s)| \|\gamma'(\phi(s))\| \, ds = \int_a^b \|\gamma'(t)\| \, dt.$$

Aquesta definició s'esten naturalment als *arcs de classe C^1 a trossos*, que són aquells arcs continus que són de classe C^1 llevat d'un nombre finit de punts on la derivada presenta discontinuïtats de salt.

Fixem-nos que aquesta quantitat $L(\gamma)$ no depen només del recorregut Γ, perquè podem passar vàries vegades pel mateix punt. Per exemple, $\gamma(t) = (\cos t, \sin t), 0 \leq t \leq 4\pi$, recorre dos cops el cercle unitat del pla, i $L(\gamma) = 4\pi$. Quan γ és simple, aleshores hi ha una correspondència bijectiva entre particions de $[a, b]$ i poligonals inscrites a Γ i aleshores sí que el que hem definit depen només de Γ i el designem $L(\Gamma)$. Per tant tenim definida la longitud per a conjunts Γ de $\mathbb{R}^d$ que puguin ser recorreguts per un arc simple de classe C^1 o de classe C^1 a trossos.

Anomenem a $ds = \|\gamma'(t)\| \, dt$, l'*element de longitud*, allò infinitesimal que hem de sumar (integrar) per tenir la longitud. Hem de pensar ds com la longitud infinitesimal d'un arc de corda entre $\gamma(t)$ i $\gamma(t + dt)$, que s'aproxima per la longitud del trosset de tangent.

Per evitar complicacions, a partir d'ara suposarem que γ és simple, de classe C^1, i $\gamma'(t) \neq 0$ per a tot t. Fixem dos punts $P, Q \in \Gamma$, suposem $P = \gamma(0), Q = \gamma(t_0)$. Per a un punt intermedi $R = \gamma(t) \in \Gamma, 0 \leq t \leq t_0$, definim $s(t)$ com la longitud de l'arc entre P i R, això és

$$s(t) = \int_0^t \|\gamma'(x)\| \, dx.$$

Per tant $s'(t) = \|\gamma'(t)\|$. Com que $\gamma'(t) \neq 0$ per a tot t, l'aplicació $t \to s(t)$ és estrictament creixent entre $[0, t_0]$ i $[0, L]$ on L és la longitud entre P, Q, amb inversa $t = t(s)$ també estrictament creixent amb derivada $1/\|\gamma'(t)\|$. Llavors , la parametrització $\hat{\gamma}(s) = \gamma(t(s)), 0 \leq s \leq L$ compleix $\|\hat{\gamma}'(s)\| = 1$. Com que el paràmetre s és la longitud de l'arc, s'anomena la *parametrització pel paràmetre arc.*

Fixem-nos que hem treballat amb la parametrització γ per definir longituds, paràmetre arc, etc. però només depen de Γ entre P i Q. Per deslliurar-nos de la parametrització, faltaria ara saber dir quins conjunts Γ poden aparéixer com a recorregut d'arcs simples de classe C^1 amb vector tangent no nul. Això depen del teorema de la funció implícita, veure la secció ??

És natural preguntar-se com es calcula la longitud quan l'arc està donat en un altre sistema de coordenades. Per exemple, suposem que $r = r(t), \theta = \theta(t)$ és un arc donat en coordenades polars. En coordenades cartesianes serà

$$x(t) = r(t) \cos\theta(t), y(t) = r(t) \sin\theta(t),$$

per tant l'element de longitud és

$$ds = \sqrt{x'^2 + y'^2} \, dt = \sqrt{(r' \cos\theta - r\theta' \sin\theta)^2 + (r' \sin\theta + r\theta' \cos\theta)^2} \, dt =$$

$$= \sqrt{r'^2 + r^2(\theta')^2} \, dt.$$

La interpretació és clara: quan passem de $(r(t), \theta(t))$ a $(r(t+dt), \theta(t+dt))$ (en coordenades polars), podem pensar que passem per $(r(t+dt), \theta(t))$ (recorrent primer una distància infinitesimal $|r'(t)| \, dt$ en la direcció radial i després de $(r(t+dt), \theta(t))$ a $(r(t+dt), \theta(t+dt))$ en la direcció tangencial recorrrent un arc de $|\theta(t+dt) - \theta(t)| \approx |\theta'(t)| \, dt$ radians. La longitud d'aquest arc infinitesimal és $r(t)|\theta'(t)| \, dt$, per definició de radià. Com que aquests dos trajectes són perpendiculars (un en la direcció radial i l'altre en la tangencial) per això $ds = \sqrt{r'^2 + r^2(\theta')^2} \, dt$.

Si γ és simple, de classe C^1, i $\gamma'(t) \neq 0$ per a tot t, sabem doncs parlar de la longitud d'un tros $A \subset \Gamma$:

$$l(A) = \int_A ds.$$

A la pràctica, prendrem qualsevol parametrització $\gamma(t)$, veurem quins $t's$ corresponen a A, $I = \{t : \gamma(t) \in A\}$, i

$$l(A) = \int_I \|\gamma'(t)\| \, dt.$$

Un cop sabem parlar de longitud sobre Γ, podem definir integrals de funcions contínues sobre Γ. Si $f : \Gamma \to \mathbb{R}$ és contínua, considerem sumes de Riemann que les podem ja pensar en el domini de paràmetres

$$\sum_i f(\gamma(\xi_i)) \|\gamma(t_{i+1}) - \gamma(t_i)\|, t_i \leq \xi_i \leq t_{i+1}.$$

El límit d'aquestes expressions és el que prenem com a definició de integral de f sobre Γ. Amb el mateix argument veiem que

$$\int_\Gamma f \, ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\| \, dt.$$

Per exemple, si f és densitat de massa, $M = \int_\Gamma f \, ds$ és la massa total, i el punt de coordenades

$$\frac{1}{M} \int_\Gamma x_i \, ds = \frac{1}{M} \int_a^b x_i(t) \|\gamma'(t)\| \, dt,$$

és el centre de masses.

2 Oberts connexos, la distància euclidiana d'un obert connex

Introduirem ara la noció de arc-connexió. Direm que un obert U de $\mathbb{R}^d$ és arc-connex si donats dos punts $P, Q \in U$ hi ha un arc continu $\gamma : [a, b] \to U$ tal que $P = \gamma(a), Q = \gamma(b)$. Ara veurem que de fet podem triar l'arc de forma ben particular, consistent en segments que tenen les direccions dels eixos de coordenades (n'hi diem escaleta).

Un obert s'anomena convex si donats dos punts $P, Q \in U$ el segment que els uneix és dins U.

Proposició. *Per a un obert U són equivalents:*

1. *U és arc-connex*
2. *Dos punts qualsevols es poden unir per una escaleta.*
3. *Dos punts qualsevols de U es poden unir per un arc de classe C^1.*

Demostració. Suposem 1, i donats $P, Q \in U$ sigui $\gamma : [a, b] \to U$ tal que $P = \gamma(a), Q = \gamma(b)$. El recorrregut Γ és un conjunt compacte (imatge per l'aplicació contínua de l'interval tancat $[a, b]$). Sigui r la distància de Γ a U^c; per a cada punt x de Γ, el cub $Q(x, \frac{r}{4})$ centrat en x d'aresta $\frac{r}{2}$ està inclós dins U. Per la compacitat de Γ, un nombre finit d'aquests cubets ja cobreix Γ. Si V és la reunió d'aquests cubets, V és un obert que conté Γ. És evident per la forma de V que a dins hi podem dibuixar una escaleta que uneix P amb Q.

Per veure que 2 implica 3, tan sols hem d'arrodonir les cantonades d'una escaleta per tal que hi tinguin tangent contínua també. □

Un obert U s'anomena *connex* si no és reunió disjunta d'oberts més petits. Dit d'una altra manera, si $U = U_1 \cup U_2, U_1 \cap U_2 = \emptyset$, llavors un dels dos és buit. Això és el mateix que dir que l'únic subconjunt obert no buit $V \subset U$ que també és tancat en U és U.

Proposició. *Un obert és arc-connex si i només si és connex.*

Demostració. Per veure que si és arc-connex és connex, suposem que $U = U_1 \cup U_2, U_1 \cap U_2 = \emptyset$ i que cap d'ells és buit. Prenem $P \in U_1, Q \in U_2$ i considerem un arc continu $\gamma : [a, b] \to U$ que els ajunti, $\gamma(a) = P, \gamma(b) = Q$. Definim $A = \{t \in [a, b] : \gamma(t) \in U_1\}$. Com que U_1 és obert, conté P i γ és continua, A conté un intervalet $[a, a + \delta)$. De la mateixa forma, hi ha un intervalet $(b - \tau, b]$ tal que $\gamma(t) \in U_2, b - \tau < t \leq b$, i per tant $(b - \tau, b]$ és disjunt amb A. Llavors $\alpha = \sup A$ compleix $a < \alpha < b$. Ara, o bé $\gamma(\alpha) \in U_1$ o bé $\gamma(\alpha) \in U_2$. En el primer cas, per la continuïtat de γ hi ha un intervalet $(\alpha - \varepsilon, \alpha + \varepsilon)$ dins U_1 i dins (c, d), cosa que diu que $\alpha + \varepsilon \in A$ i contradiu que α sigui el suprem d'A. De la mateixa forma, en el segon cas hi ha un intervalet $(\alpha - \varepsilon, \alpha + \varepsilon)$ fora d'A, cosa que també contradiu que α sigui el suprem. Aquesta contradicció mostra que l'obert és connex.

Recíprocament, suposem que U és un connex no buit i volem veure que és arc-connex. Fixem un punt $P \in U$ i definim V com el conjunt de punts $Q \in U$ que es poden ajuntar amb P per un arc continu. Llavors, V és obert, perquè si $Q \in V$ i $Q' \in B(Q,r) \subset U$, empalmant un arc que va de P a Q amb el segment que va de Q a Q' tenim u arc continu que va de P a Q', per tant $B(Q,r) \subset V$ i V és obert. Però $U \setminus V$ també és obert, amb el mateix argument (si un punt no es pot ajuntar amb P cap dels propers tampoc). Com que $P \in V$, V no és buit i ha de ser $U \setminus V$ buit, per tant $U = V$.

□

Proposició. *A la recta els oberts connexos són els intervals oberts.*

Demostració. Que un interval obert és arc-connex és evident. Recíprocament, suposem que U és un obert arc-connex no buit de la recta, i definim

$$a = \inf U, b = \sup U,$$

on s'interpreta que $a = -\infty$ si U no és acotat inferiorment i $b = +\infty$ si U no és acotat superiorment. Llavors, hi ha $a' > a$ arbitràriament proper a a amb $a' \in U$, i hi ha $b' < b$arbitràriament proper a b amb $b' \in U$. Com que U és arconnex, $[a', b'] \subset U$. Com que això és cert per a tot $a' > a$ arbitràriament proper a a i per a tot b' arbitràriament proper a b, tindrem que $(a, b) = U$. □

En un obert U la relació $P \equiv Q$ si P, Q es poden unir per un arc continu dins U és una relació d'equivalència. Els mateixos arguments anteriors demostren que cada classe d'equivalència és oberta i, per definició arc-connexa. Així, tot obert U és una reunió d'oberts arc-connexs disjunts, que s'anomenen les *components arc-connexes d'U*, o també *components connexes d'U*. Aplicant la proposició ??, basada en el fet que els punts de coordenades racionals és dens i numerable, veiem que hi ha com a mot un nombre numerable de components connexes, tot obert és reunió numerable d'oberts arc-connexs disjunts. A la recta, tot obert és una reunió numerable d'intervals oberts.

Si U és un obert arc-connex, podem definir la distància euclidea, dins U, entre dos punts $P, Q \in U$,

$$d_U(P, Q) = \inf L(\gamma),$$

on l'infim és respecte totes les corbes regulars de classe C^1 (o regulars de classe C^1 a trossos) que uneixen P, Q dins U. Aquesta noció compleix les propietats de distància, és a dir,

- $d_U(P,Q) = d_U(Q,P) \geq 0$
- $d_U(P,Q) = 0$ si i només si $P = Q$.
- $d(P,Q) \leq d(P,R) + d(R,Q), P, Q, R \in U$.

Evidentment, $d_U(P,Q) \geq \|P - Q\|$, i si U és convex aleshores $d_U(P,Q) = \|P - Q\|$.

És fàcil donar exemples que mostren que en general l'ínfim anterior no és accessible, no és un mínim. De fet, si el segment que uneix P, Q no és tot ell dins U, aquest infim no és accesible, perquè donada qualsevol corba que uneix P, Q que no sigui el segment que els uneix, hi ha una poligonal més curta, sempre dins U.

Els oberts arc-connexs s'anomenen *dominis*.

3 La diferencial d'una funció de d variables

Sigui $f : U \longrightarrow \mathbb{R}^m$ una aplicació definida en un obert U de $\mathbb{R}^d$, on ara suposem $d \geq 1$. Donat $a \in U$, ens proposem definir quelcom anàleg al concepte de derivada del cas $d = 1$. La idea fonamental d'aquest concepte és que ha de mesurar *canvi* en a. La primera cosa que cal assenyalar és que ara no podem considerar, per a x proper a a, un quocient incremental

$$\frac{f(x) - f(a)}{x - a},$$

perquè el denominador $x - a$ és ara un vector, i no té sentit dividir per un vector.

Hi ha, en principi, dues formes d'encarar aquesta definició. Com veurem, en ambdues podrem suposar que $m = 1$.

La primera és mitjançant el concepte de *derivada direccional*: donat un vector u, ens fixem en la restricció de f a la recta que passa per a i té direcció u, que en forma paràmetrica s'escriu $a+tu, t \in \mathbb{R}$ i mirem cóm canvia aquesta funció en $t = 0$, que és el paràmetre que correspon al punt a:

$$\lim_{t \to 0} \frac{f(a + tu) - f(a)}{t} = D_u f(a).$$

Si existeix aquest límit, n'hi diem la *derivada de f en a en la direcció u*. Dit d'una altra manera, si $\phi(t) = f(a + tu), D_u f(a) = \phi'(0)$, $D_u f(a)$ és el vector tangent a l'arc $\phi(t)$. Si $f = (f_1, \cdots, f_m)$ és evident que aquest límit existeix si i només existeix per a cada component f_j, i en aquest cas, $D_u f(a)$ és el vector de $\mathbb{R}^m$ que té per components $D_u f_j(a)$.

Si f pren valors escalars, com a derivada que és, el nombre $D_u f(a)$ mesura el ritme de canvi de f en la direcció u. Si és molt gran i positiu vol dir que $f(a+tu)$ creix molt ràpidament en $t=0$, etc.

També podem escriure

$$f(a+tv) = f(a) + tD_u f(a) + o(t), t \to 0$$

Recordem que la notació $A = o(B), t \to 0$, significa que $\lim_{t\to 0} \frac{\|A(t)\|}{\|B(t)\|} = 0$, i es llegeix A és d'ordre més petit que B quan $t \to 0$. La notació $o(t)$ designa genèricament qualsevol variable d'ordre més petit que t quan $t \to 0$, la notació $o(1)$ indica qualsevol cosa amb límit zero, etc. La notació $A = O(B)$ significa que el quocient $\frac{\|A\|}{\|B\|}$ es manté acotat, i en particular A tendeix a zero si B tendeix a zero.

Quan u és el vector e_i de la base canònica s'utilitza la notació $D_i f(a)$, o també $\frac{\partial f}{\partial x_i}(a)$, $\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a), \frac{\partial f}{\partial z}(a)$ en el cas $d = 2, 3$. També s'utilitza la notació f_x, f_y, f_z. Quan aquesta derivada direccional existeix en tots els punts utilitzem la notació $D_u f(x)$, o $\frac{\partial f}{\partial x}$, etc. que seran altre cop funcions de d variables. S'anomenen *les derivades parcials*. Si f està donada per una fórmula, la derivada parcial $\frac{\partial f}{\partial x_i}$ existeix en tots els punts on la fórmula té sentit, i la calculem utilitzant les regles de derivació derivant respecte la variable x_i pensant que les altres variables $x_j, j \neq i$ són constants. Per exemple, fer a una funció $f(x,y)$ de dues variables

$$\frac{\partial f}{\partial x} = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h},$$

és això: derivar respecte de x pensant que y és constant.

És immediat que si $D_u f(a)$ existeix i $\lambda \in \mathbb{R}$ aleshores també existeix $D_{\lambda u} f(a)$ i val $\lambda D_u f(a)$. Ara bé, per a direccions u, v linealment independents, pot existir $D_u f(a)$ i no existir $D_v f(a)$. Per exemple $f(x,y) = x^2 y \sin \frac{1}{y}, y \neq 0, f(x,0) = 0$, té derivada $\frac{\partial f}{\partial x} = 0$ en $(0,0)$ però no té derivada $\frac{\partial f}{\partial y}$.

Més significativament, l'existència de totes les derivades direccionals no implica ni la continuïtat. Reprenem l'exemple en $d = 2$,

$$f(x,y) = \phi(\frac{x}{y^2}), y \neq 0, f(x,0) = 0,$$

on $\phi(t)$ és qualsevol funció d'una variable $t \in \mathbb{R}$ tal que $\phi(0) = 0, \lim_{t\to\infty} \phi(t) = 0$, que no és contínua en $(0,0)$. Les derivades $\frac{\partial f}{\partial x}(0,0)$, $\frac{\partial f}{\partial y}(0,0)$ són zero; si $y = mx, m \neq 0$ és qualsevol altre recta que passi per zero, és a dir de vector director $u = (1, m)$ llavors

$$\frac{f(x, mx)}{x} = \frac{1}{x}\phi(\frac{1}{m^2 x}),$$

i si imposem que $\lim_{t\to\infty} t\phi(t) = 0$ tindrem que totes les derivades direccionals són zero.

L'existència de totes les derivades direccionals (situació que alguns texts anomemen f *derivable en* a) no és doncs la definició correcta. La definició correcta ens la suggereix mirar l'existència de derivada en el cas $d = 1$ com equivalent a l'existència d'una aproximació afí de $f(x)$. En efecte, per a funcions d'una variable $t \in \mathbb{R}$, amb valors a $\mathbb{R}^m$, l'existència del límit del quocient incremental

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = f'(a) \in \mathbb{R}^m,$$

és equivalent a

$$f(a+h) = f(a) + hf'(a) + o(h), h \to 0; \lim_{h\to 0} \frac{\|f(a+h) - f(a) - hf'(a)\|}{|h|} = 0,$$

que expressa que la funció afí d'ordre 1 en h, $f(a) + hf'(a)$ és una aproximació d'ordre 1 de f, o, equivalentment, que la funció lineal $hf'(a)$ és una aproximació lineal de l'increment $f(a+h) - f(a)$, en el sentit que la diferencia és $o(h)$. Aquesta aproximació és única, perquè si $A + Bh$ compleix

$$f(a+h) = A + Bh + o(h), h \to 0$$

llavors necessàriament $A = f(a), B = f'(a)$.

Aquest és el punt de vista que prendrem: l'existència d'una aproximació lineal a l'increment $f(a+h) - f(a)$.

Definició. Direm que $f : U \longrightarrow \mathbb{R}^m$ és diferenciable en un punt $a \in U$ si hi ha una aplicació lineal de $\mathbb{R}^d$ en $\mathbb{R}^m$, anomenada *la diferencial de* f *en* a i designada $df(a)$, tal que

$$f(a+h) = f(a) + df(a)(h) + o(h), h \to 0,$$

és a dir

$$\frac{\|f(a+h)-f(a)-df(a)(h)\|}{\|h\|}=0,$$

on $df(a)(h)$ indica l'acció de l'aplicació lineal $df(a)$ sobre el vector h.

Obviament, si f és diferenciable en a llavors f és contínua en a, perquè $df(a)(h)\to 0$ quan $h\to 0$. Amb la notació $x=a+h$,

$$f(x)=f(a)+df(a)(x-a)+o(\|x-a\|).$$

Novament, si $f=(f_1,\cdots,f_m)$, f és diferenciable en a si i només si ho és cada funció component, i en aquest cas $df(a)(h)=(df_1(a)(h),\cdots,df_m(a)(h))$. Del que hem dit abans es despren que en el cas $d=1$ f és diferenciable en a si i només si és derivable; en aquest cas, la diferencial és $df(a)(h)=hf'(a)$ (una aplicació lineal L de $\mathbb{R}$ en $\mathbb{R}^m$ queda determinada per $L(1)$).

En el cas $m=1$ la interpretació geomètrica és la mateixa que en el cas $d=1, m=1$. Que f sigui diferenciable en a significa que la varietat afi de $\mathbb{R}^{d+1}$ d'equació paramètrica $h\to f(a)+df(a)(h)$ és tangent en el punt $(a,f(a))$ a la varietat gràfic $y=f(x)$. En $d=2$, amb la notació (x,y) i (a,b) enlloc d'a, si f és diferenciable en (a,b) el pla d'equació

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

és tangent al gràfic $z=f(x,y)$ en el punt $(a,b,f(a,b))$.

Si f és diferenciable en a i prenem $h=tu$, on u és un vector direcció, com que $df(a)(tu)=tdf(a)(u)$, trobem que

$$f(a+tu)=f(a)+tdf(a)(u)+o(t),$$

que diu que $D_uf(a)$ existeix i val $df(a)(u)$. Això ens diu en particular que la diferencial és única i consisteix en assignar a cada direcció u la seva derivada direccional $D_uf(a)$, que és el vector tangent en $f(a)$ a l'arc $f(a+tu)$.

L'existència de la diferencial no només diu que f és contínua i existeixen totes les derivades direccionals, sinó que a més a més, l'assignació $u\to D_uf(a)$ és lineal. Per exemple, la funció $f(x,y)=0, y\neq 0, f(x,0)=x$ és contínua en $(0,0)$ i totes les derivades direccionals en $(0,0)$ són zero llevat de la parcial respecte de x, que val 1, amb la qual cosa $u\to D_uf(0,0)$ no és lineal. Tampoc el fet que les derivades direccionals existeixin i l'aplicació $u\to D_uf(a)$ sigui lineal assegura que f sigui diferenciable en a.

Com a conseqüència de la linealitat de $df(a)$ observem que si f és diferenciable en a i les derivades direccionals $D_{u_j}f(a)$ són zero per a d vectors linealment independents (és a dir, una base), aleshores $df(a) = 0$, i $D_u f(a) = 0$ per a tot vector u.

Per a una funció escalar $f(x)$ definida al voltant d'a, comprovem si és diferenciable fent el següent: primer ha d'existir $f(a) = \lim_{x \to a} f(x)$. Després han d'existir les d derivades parcials

$$A_i = \frac{\partial f}{\partial x_i}(a) = \lim_{t \to zero} \frac{f(a + te_i) - f(a)}{t}.$$

I finalment, l'aplicació lineal $h = (h_1, \cdots, h_d) \to L(h) = A_1 h_1 + \cdots + A_d h_d$ ha de complir

$$\|f(a+h) - f(a) - L(h)\| = o(h).$$

Una condició suficient de diferenciabilitat en termes de les derivades parcials és el següent fet:

Proposició. *Suposem que les derivades parcials $D_i f(x)$ existeixen en tots els punts $x \in B(a, \varepsilon)$ i són funcions contínues en a; aleshores f és diferenciable en a.*

Demostració. Fem la demostració en el cas $d = 2$ amb la notació $f(x, y)$, (a, b) enlloc d'a, la idea si $d > 2$ és la mateixa. Descomposem

$$f(x, y) - f(a, b) = f(x, y) - f(x, b) + f(x, b) - f(a, b).$$

Pel teorema del valor mig aplicat a les funcions d'una variable $y \to f(x, y)$, $x \to f(x, b)$ hi ha punts ξ, η, ξ entre y, b, η_y entre a, x tals que

$$f(x, y) - f(x, b) = \frac{\partial f}{\partial y}(x, \xi)(y - b), f(x, b) - f(a, b) = \frac{\partial f}{\partial x}(\eta, b)(x - a),$$

i per tant

$$f(x, y) - f(a, b) - \frac{\partial f}{\partial x}(a, b)(x - a) - \frac{\partial f}{\partial y}(a, b)(y - b) =$$

$$\left(\frac{\partial f}{\partial x}(\eta, b) - \frac{\partial f}{\partial x}(a, b)\right)(x - a) + \left(\frac{\partial f}{\partial y}(x, \xi) - \frac{\partial f}{\partial y}(a, b)\right)(y - b).$$

La continuïtat de les derivades parcials en (a, b) implica que els dos darrers termes són $o(\|h\|)$ i ja. $\square$

En particular, qualsevol funció expresada per una fórmula és diferenciable en els punts on la fórmula té sentit. D'altra banda, tots els enunciats naturals relatius a operacions entre funcions diferenciables són certs, amb proves rutinàries:

Proposició. *Suposem que $f, g : U \longrightarrow \mathbb{R}^m$ són diferenciables en a. Aleshores*

- *La suma $f + g$ també ho és amb diferencial $df(a) + dg(a)$.*
- *El producte escalar $E(x) = \langle f(x), g(x)\rangle$ és diferenciable amb diferencial*

$$dE(a)(h) = \langle f(a), dg(a)(h)\rangle + \langle df(a)(h), g(a)\rangle.$$

  En particular, per a funcions escalars, $d(fg)(a) = f(a)dg(a)+g(a)df(a)$.

- *Si $d = 3$ i designem per $V(x) = f(x) \times g(x)$ el producte vectorial, V és diferenciable en a amb diferencial*

$$dV(a)(h) = df(a)(h) \times g(a) + f(a) \times dg(a)(h).$$

Finalment un tema notacional. Clàsicament es designen per $dx_1, \cdots, dx_d$ els petits increments $h_1, \cdots, h_d$ i es pensa df com l'increment infinitesimal de f, i s'escriu

$$df = \sum_i \frac{\partial f}{\partial x_i}\, dx_i.$$

Una altra forma de veure-ho és pensar que $dx_1, \cdots, dx_d$ és la base dual de la canònica, en el sentit que l'acció de dx_i sobre un vector h és $dx_i(h) = h_i$, la seva i-sima component. L'acció de df sobre h és doncs

$$\sum_i \frac{\partial f}{\partial x_i} h_i.$$

4 La matriu jacobiana. La regla de la cadena

Si f és diferenciable en a i expressem l'increment $h = (h_1, \cdots, h_d) = \sum_i h_i e_i$, per linealitat

$$df(a)(h) = \sum_i h_i df(a)(e_i) = \sum_i h_i D_i f(a).$$

Si $f = (f_1, \cdots, f_m)$, $D_i f(a)$ és el vector de $\mathbb{R}^m$ de components

$$(D_i f_1(a), \cdots, D_i f_m(a)).$$

Tot plegat, veiem que la matriu de $df(a)$ en les bases canòniques de $\mathbb{R}^d, \mathbb{R}^m$ és

$$\begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \dots & D_d f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \dots & D_d f_2(a) \\ \vdots & \vdots & \vdots & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \dots & D_d f_m(a) \end{pmatrix}.$$

Aquesta matriu la designem també $df(a)$, i s'anomena la *matriu jacobiana* de f en a. La i-sima columna és $D_i f(a), 1 \leq i \leq d$ i la j-sima fila és $df_j(a), 1 \leq j \leq m$.

La regla de la cadena per a funcions reals de variable real estableix que $g(f(x))$ té derivada $g'(f(x))f'(x)$. Això s'explica perquè $f'(x)$ és el factor pel qual s'ha de multiplicar un increment infinitesimal Δx per tenir l'increment Δy de $y = f(x)$, i $g'(y)$ és el factor pel qual s'ha de multiplicar Δy per tenir l'increment Δz de $z = g(y)$. Ara veiem la versió general

Proposició. *Si $f : U \to \mathbb{R}^m$ és diferenciable en a i $g : V \to \mathbb{R}^p$ ho és en $b = f(a)$, la composició $M(x) = g(f(x))$ és diferenciable en a amb diferencial $dM(h) = dg(b)(df(a)(h))$, la composició de les diferencials.*

Demostració. Designem per h els petits increments de x i per k els petits increments de $y = f(x)$. Per hipòtesi,

$$f(a+h) = f(a) + df(a)(h) + R(h), g(b+k) = g(b) + dg(b)(k) + S(k),$$

amb $R(h) = o(h), S(k) = o(k)$, quan $h, k \to 0$. Podem posar $f(a+h) = b+k$ amb $k = df(a)(h) + R(h)$ (que tendeix a zero quan $h \to 0$). Per tant, substituint

$$g(f(a+h)) = g(b) + dg(b)(df(a)(h)) + dg(b)(R(h)) + S(k).$$

Hem de veure que els dos darrers termes són $o(h)$. Ara bé, com que $k = O(h)$, tot allò que sigui $o(k)$ és també $o(h)$; per tant tan sols cal veure que $dg(b)(R(h) = o(h)$. Però si C és la norma de l'aplicació lineal $dg(b)$, tenim $\|dg(b)(R(h))\| \leq C\|R(h)\| = o(h)$. $\square$

En termes de les matrius $df(a) = \left(\frac{\partial f_j}{\partial x_i}(a)\right)_{i,j}, dg(b) = \left(\frac{\partial g_k}{\partial y_j}(b)\right)_{k,j}$, la matriu $dM(a)$ és la matriu producte $dg(b)df(a)$, l'entrada $\frac{\partial M_k}{\partial x_i}$ de la fila $k = 1, \cdots, p$ i columna $i = 1, \cdots, d$ és

$$\frac{\partial M_k}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g_k}{\partial y_j}(b) \frac{\partial f_j}{\partial x_i}(a).$$

O també,

$$\frac{\partial M}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g}{\partial y_j}(b) \frac{\partial f_j}{\partial x_i}(a).$$

Enlloc de recordar-ho en termes de producte de matrius podem pensar-ho de la següent forma. Hem de derivar

$$M_k(x) = g_k(f(x)) = g_k(f_1(x), \cdots, f_m(x))$$

respecte de x_i, la i-sima component de x. Al fer un petit increment Δx_i a x_i, totes les $y_j = f_j(x)$ s'incrementen una petita quantitat $\Delta y_j = \frac{\partial f_j}{\partial x_i} \Delta x_i$. Per tant totes les variables y_j de g_k han sofert un increment; l'increment de y_j porta a l'increment $\frac{\partial g_k}{\partial y_j} \Delta y_j$ i sumant arribem a la conclusió. La regla és per tant que per derivar una composició $g(f(x))$ respecte d'una variable x_i, derivem g respecte cadascuna de les seves variables, multipliquem per la derivada d'aquesta variable respecte de x_i, i sumem.

Una conseqüència de la regla de la cadena és la següent. Havíem vist que donada una direcció u, $D_u f(a)$ és el vector tangent en $f(a)$ a l'arc $f(a + tu)$. Prenem, enlloc del segment $a + tu$, qualsevol arc regular $\gamma(t)$ que passi per $a, \gamma(0) = a$ amb tangent $u, \gamma'(0) = u$, i fem també la composició $\hat{\gamma}(t) = f(\gamma(t)) = f(\gamma_1(t)), \cdots, \gamma_d(t))$, que passa per $f(a)$. Hi té també com a vector tangent el vector $D_u f(a)$:

$$\hat{\gamma}'(0) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(a) x_i'(0) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}(a) u_i = D_u f(a).$$

Amb això tenim una clara interpretació geomètrica de com funciona la diferencial $df(a)$ com a aplicació lineal: donat u, prenem qualsevol arc γ que passi per a i tingui vector tangent u, transportem γ per f i d'aquesta corba imatge prenem el seu vector tangent en $f(a)$.

Amb aquesta interpretació geomètrica, la regla de la cadena és clara: si u és un vector direcció, $D_u M(a) = dM(a)(u)$ és el vector tangent a la corba $g(f(a+tu))$; però aquesta és la imatge per g de la corba $f(a+tu)$; pel que acabem de dir, $dM(a)(u)$ ha de ser igual a $dg(f(a)(v)$ on v és el vector tangent a $f(a+tu)$, que és $v = df(a)(u)$. O també,

$$D_u(gof)(a) = D_v g(f(a), v = D_u f(a). \tag{1}$$

5 El gradient d'una funció escalar. Punts crítics, extrems relatius

Aqui suposem que f és una funció escalar de d variables, diferenciable en tots els punts d'un obert U. Hem vist que per a un vector direcció $u = (u_1, \cdots, u_d)$,

$$D_u f(a) = df(a)(u) = \sum_{i=1}^{d} u_i D_i f(a).$$

Introduim el vector $\nabla f(a) = (D_1 f(a), \cdots, D_d f(a))$, anomenat el *gradient de f en a*, de manera que

$$\langle \nabla f(a), u \rangle = D_u f(a).$$

Notem que amb les notacions d'abans, quan $p = 1$ la regla de la cadena

$$\frac{\partial M}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g}{\partial y_j}(b) \frac{\partial f_j}{\partial x_i}(a).$$

s'escriu

$$\nabla M(a) = \sum_{j=1}^{d} \frac{\partial g}{\partial y_j}(b) \nabla f_j(a),$$

i també

$$\nabla M(a) = (df(a)^t)(\nabla g(b)),$$

on $(df(a))^t$ designa la matriu trasposta. Això és una altra forma d'escriure (1) perquè si $v = D_u f(a) = df(a)(u)$,

$$D_u M(a) = \langle nablaM(a), u \rangle = \langle df(a)^t (\nabla g(b), u \rangle = \langle \nabla g(b), df(a)(u) \rangle =$$

$$= \langle \nabla g(b), v \rangle = D_v g(b).$$

Aquest vector gradient té diverses interpretacions i utilitats. En primer lloc, pensem en una direcció donada per un vector unitari $u, \|u\| = 1$, i suposem que $\nabla f(a) \neq 0$. Aleshores per la desigualtat de Cauchy-Schwarz, $|\langle \nabla f(a), u \rangle| \leq \|\nabla f(a)\|$, és a dir,

$$-\|\nabla f(a)\| \leq D_u f(a)\| \leq \nabla f(a)\|,$$

mentre que per a $u = \frac{\nabla f(a)}{\|\nabla f(a)\|}$ la derivada direccional val $\|\nabla f(a)\|$ i per a $u = -\frac{\nabla f(a)}{\|\nabla f(a)\|}$ la derivada direccional val $-\|\nabla f(a)\|$. Això ens diu que la direcció u en la que $D_u f(a)$ és màxima és la del gradient, igual al mòdul del gradient, i la direcció en la que és mínima és l'oposada. Com $D_u f(a)$ mesura el ritme de canvi en la direcció u, tenim que en cada punt el gradient, si no és zero, ens indica la direcció de màxim creixement de f; i el seu mòdul, la velocitat màxima de creixement.

Llavors, per a cada punt $a \in U$ tenim un vector gradient $\nabla f(a)$ que convé visualitzar, pel que acabem d'explicar, com un vector amb origen al punt a. Això no és res més que una aplicació

$$\nabla f : U \longrightarrow \mathbb{R}^d.$$

En general, si tenim en un obert U de $\mathbb{R}^d$ una aplicació $X : U \longrightarrow \mathbb{R}^d$, però que ens mirem $X(p)$ com un vector d'origen en p, diem que tenim un *camp vectorial*. Exemples en són els camps de forces, camps de velocitats, etc. Donat un camp vectorial X en un obert, és natural mirar als arcs $\gamma(t)$ dins U (és a dir, la trajectòria $\Gamma \subset U$) tals que $\gamma'(t) = X(\gamma(t))$. En coordenades, si les equacions de γ són $x_i(t), i = 1, \cdots, d$, i $X = (X_1, \cdots, X_d)$ tenim d equacions

$$x_i'(t) = X_i(x_1(t), x_2(t), \cdots, x_d(t)), i = 1, \cdots, d.$$

Això és un *sistema d'equacions diferencials autònom*. Si pensem en t com a temps, això diria que en cada instant t la velocitat instantània, quan el mobil està en $\gamma(t)$, és la prefixada pel camp. Aquestes corbes s'anomenen les trajectòries integrals del camp. Si X és un camp de forces i en l'instant $t = 0$ deixem anar una particula en un punt p, aquesta partícula seguirà una d'aquestes trajectòries. La nostra intuició ens diria doncs que hi ha una única trajectòria que surti d'un punt donat. Aquest fet es pot provar quan el camp X compleix determinades condicions i és el contingut del teorema d'existència i unicitat d'equacions diferencials.

Tornant al cas del camp gradient, que és un tipus especial de camp, l'exemple següent ens servirà per entendre aquest concepte. Suposem que la funció $f(x)$ indica una temperatura estacionària (constant en el temps) en el punt $x \in U$. Segons hem dit, el camp gradient en cada punt indica la direcció en la que l'augment de temperatura és més gran. Si estem en un punt $p \in U$ i som molt fredolics, ens voldrem moure sempre en aquesta direcció. És a dir, si a l'instant t estic a $\gamma(t)$ la meva velocitat ha d'apuntar en la direcció del gradient en $\gamma(t)$. Això vol dir que $\gamma(t)$ ha de complir $\gamma(0) = p$ i en cada instant

$$\gamma'(t) = \lambda(t)\nabla f(\gamma(t)).$$

Una segona interpretació del gradient, quan no és zero, és en termes dels conjunts de nivell. Tenim els $L_c = \{x \in U : f(x) = c\}$. Un punt $a \in U$ és a $L_c, c = f(a)$. Aquest conjunt de nivell ens l'imaginem com quelcom de dimensió $d-1$, però ja hem vist que això no sempre és així, no és necessàriament una varietat de dimensió $d-1$. Ara suposarem que $\nabla f(a) \neq 0$ i que dins de L_c hi tenim un arc regular de parametrització $\gamma(t)$, que passa per $a, \gamma(0) = a$,amb tangent $u = \gamma'(0)$ (de fet, després veurem que si f és de classe C^1, el fet que $\nabla f(a) \neq 0$ ja implica que n'hi ha d'aquests arcs). Llavors, derivant en $t = 0$ l'equació $f(\gamma_1(t), \cdots, \gamma_d(t)) = c$ la regla de la cadena ens diu que

$$\langle \nabla f(a), u\rangle = \sum_{i=1} D_i f(a) x_i'(0) = 0.$$

Això ens diu que $\nabla f(a)$ és perpendicular en a a tots les vectors tangents d'arcs regulars dins L_c que passen per a i per tant diem que és perpendicular al conjunt de nivell. Si tenim una intersecció de conjunts de nivell de k funcions,

$$M = \{x \in U : f_k(x) = c_k\},$$

i $a \in M$, tindrem que els k gradients $\nabla f_k(a)$ són perpendiculars a les tangents a tots els arcs din M que passen per a. Després tornarem sobre aquest tema.

Una tercera utilitat del gradient és en relació a la noció d'extrem relatiu. Un punt $a \in U$ s'anomena un *extrem relatiu* de f si hi ha una bola $B(a,r) \subset U$ tal que $f(a) \leq f(x), x \in B(a,r)$ (mínim relatiu) o $f(a) \geq f(x), x \in B(a,r)$ (màxim relatiu). En aquest cas, si f és diferenciable en a i u és qualsevol direcció, la funció $f(a+tu)$ té un extrem relatiu en $t=0$; per tant la derivada, que és $D_u f(a)$, ha de ser zero, i com que això és cert per a tot u, $\nabla f(a)=0$.

Els punts a amb $\nabla f(a)=0$ s'anomenen *punts crítics* de f, i hem vist doncs que tots els extrems relatius són punts crítics. Obviament, no tots els punts crítics són extrems relatius; per exemple, $f(x,y)=x^2-y^2$ té un punt crític al $(0,0)$ que no és ni màxim ni mínim. En la direcció de l'eix x té un mínim i en la direcció de l'eix y té un màxim; d'aquest tipus de punts n'hi direm *punts de sella*. La funció $f(x,y)=x^3+y^3$ també té un punt crític en $(0,0)$, i no és ni màxim ni mínim relatiu, ni tampoc un punt de sella. Més endavant aprendrem a classificar aquests punts.

6 Canvis de coordenades diferenciables

Reprenem ara el concepte de canvi de coordenades en un obert U de $\mathbb{R}^d$. És una aplicació $\Phi: U \to \mathbb{R}^d$ injectiva, és a dir bijectiva de U en $\Phi(U)$ i que és contínua en els dos sentits, un homeomorfisme entre U i $\Phi(U)$. Si Φ té components $(v_1(x),\cdots,v_d(x))$, pensem $v_1(x),\cdots,v_d(x)$ com a noves coordenades del punt x. Aquest nou sistema de coordenades defineix d families $\{x : v_i(x)=v_i(a), i\neq j\}$ que fan d'eixos de coordenades en cada punt (els punts que tenen les mateixes noves coordenades que a, menys la j-sima, que pot variar), que pensem com uns arcs.

Un canvi de coordenades diferenciable és aquell pel qual $\Phi(U)=V$ és obert i tant Φ com la inversa Φ^{-1} són diferenciables. També s'utilitza el terme *difeomorfisme*. Llavors existeixen les derivades parcials $\frac{\partial v_j}{\partial x_i}, \frac{\partial x_i}{\partial v_j}$. Com que la composició de Φ i Φ^{-1} és la identitat, la regla de la cadena implica que

$$d(\Phi^{-1})=(d\Phi)^{-1},$$

és a dir, la diferencial $d\Phi(x)$ és una aplicació lineal inversible de $\mathbb{R}^d$ en $\mathbb{R}^d$, la qual cosa és equivalent a dir que $\det d\Phi(x)\neq 0$ per a tot $x\in U$. Més endavant veurem com a conseqüència del teorema de la funció inversa, que si $\Phi: U\to\mathbb{R}^d$ és injectiva i $d\Phi(x)$ és inversible per a tot x, llavors $\Phi(U)=V$ és obert i Φ^{-1} és diferenciable (o sigui que es podria definir difeomorfisme com a una aplicació injectiva amb diferencial inversible a cada punt).

En particular, si $u\neq 0, D_u\Phi(x)\neq 0, D_u(\Phi)^{-1}(v)\neq 0$. Designem per $x_i(v), v=(v_1,\cdots,v_d)$, les funcions components de Φ^{-1}. Els nous eixos de coordenades s'obtenen fixant tots els valors del v_j menys un, diguem v_i, per tant són corbes regulars amb vector tangent no nul en cada punt. Provisionalment anomenen X_i aquest camp vectorial, així $X_i(x)$ és el vector tangent en x a l'i-sim eix de coordenades v_i. Té l'expressió

$$X_i(x) = (\frac{\partial x_1}{\partial v_i}(\Phi(x)), \cdots, \frac{\partial x_d}{\partial v_i}(\Phi(x))).$$

Aquesta seria la seva expressió en termes de les coordenades v_j, per tenir-lo en termes de les x_i cal saber $v_j(x)$.

En el cas de les coordenades cartesianes habituals, respecte de la base canònica, tindrem obviament $X_i(x) = \vec{e_i}$, és un camp constant. En general, el vector $X_i(x)$ pot canviar de direcció al canviar el punt.

Suposem ara que tenim una funció $f(x)$ escalar definida per a $x \in U$. Si composem amb $x = \Phi^{-1}(v)$ tindrem una funció definida en V,

$$g(v) = f(\Phi^{-1}(v)), \text{o} \quad g(\Phi(x)) = f(x).$$

La funció g no és altra que f expressada en les noves coordenades v, però de moment la designem amb una notació diferent, g. Com que Φ i la seva inversa són diferenciables, f és diferenciable si i només si ho és g.

Si apliquem la regla de la cadena, com que per construcció $X_i(x)$ es correspon amb $\frac{\partial}{\partial v_i}$ mitjançant $d\Phi(x)$, si en cada punt x fem la derivada direccional de f en la direcció d'aquest vector trobem ,

$$\frac{\partial f}{\partial X_i(x)}(x) = \sum_{j=1}^{d} \frac{\partial x_j}{\partial v_i}(\Phi(x))\frac{\partial f}{\partial x_j}(x) = \frac{\partial g}{\partial v_i}(\Phi(x)).$$

Dit d'una altra manera, $\frac{\partial f}{\partial X_i}(x)$ és $\frac{\partial g}{\partial v_i}(v)$ expressat en termes de les x.

Per aquest motiu canviem la notació i substituim la notació X_i per $\frac{\partial}{\partial v_i}$. Dit d'una altra manera, aquest derivació parcial, que en principi està definida en $\Phi(U) = V$ i actuant sobre la g, el passem a veure com a una derivació en U actuant sobre f , seguint la idea de que els punts no s'han mogut sinó que han canviat de coordenades. Per això, utilitzem la notació $f(v_1, \cdots, v_d)$ enlloc de $g(v_1, \cdots, v_d)$. Així,

$$\frac{\partial f}{\partial v_i} = \sum_{j=1}^{d} \frac{\partial f}{\partial x_j}\frac{\partial x_j}{\partial v_i}.$$

Obviament, $\frac{\partial v_i}{\partial v_j} = \delta_{i,j}$. Com que el terme esquerra és la derivada direccional de v_i en la direcció del vector $\frac{\partial}{\partial v_j}$, per tant $dv_i(\frac{\partial}{\partial v_j})$, tenim que $\frac{\partial}{\partial v_1}, \cdots, \frac{\partial}{\partial v_d}$ i $dv_1, \ldots, dv_d$ són bases duals (que van canviant amb x).

Per tant,

$$df = \sum_i df(\frac{\partial}{\partial v_i}) dv_i = \sum_i \frac{\partial f}{\partial v_i} dv_i.$$

Considerem amb tot detall el cas de les coordenades esfèriques (r, ϕ, θ). El vector $\frac{\partial}{\partial r}$ té la direcció radial en cada punt, $\frac{\partial}{\partial \phi}$ és el vector tangent als meridians i $\frac{\partial}{\partial \theta}$ és el vector tangent als paralels. Calculem-los. Com que

$$x = r \sin\phi \cos\theta, y = r \sin\phi \sin\theta, z = r \cos\phi,$$

tenim

$$\frac{\partial}{\partial r} = (\frac{\partial x}{\partial r}, \frac{\partial y}{\partial r}, \frac{\partial z}{\partial r}) = (\sin\phi\cos\theta, \sin\phi\sin\theta, r\cos\phi,$$

$$\frac{\partial}{\partial \phi} = (\frac{\partial x}{\partial \phi}, \frac{\partial y}{\partial \phi}, \frac{\partial z}{\partial \phi}) = (r\cos\phi\cos\theta, r\cos\phi\sin\theta, -r\sin\phi).$$

$$\frac{\partial}{\partial \theta} = (\frac{\partial x}{\partial \theta}, \frac{\partial y}{\partial \theta}, \frac{\partial z}{\partial \theta}) = (-r\sin\phi\sin\theta, r\sin\phi\cos\theta, 0).$$

Aquests tres vectors estan expressats en termes de les coordenades (r, ϕ, θ). Expressats en termes de x, y, z, són

$$\frac{\partial}{\partial r} = \frac{1}{\sqrt{x^2+y^2+z^2}}(x, y, z),$$

$$\frac{\partial}{\partial \phi} = \frac{1}{\sqrt{x^2+y^2}}(zx, zy, -x^2+y^2)$$

$$\frac{\partial}{\partial \theta} = (-y, x, 0).$$

7 Funcions que depenen de menys variables, antiderivades parcials

Per a una funció $f(x, y)$ definida al pla, diferenciable, quin criteri coneixem per tal que sigui només funció de x?. Obviament la resposta és que

$$\frac{\partial f}{\partial y} = 0$$

idènticament. Com que $y \to f(x,y)$ té derivada zero, és constant, per tant $f(x,y) = f(x,0)$ només depen de x. El mateix argument serveix si f està definida en un rectangle obert (producte d'intervals oberts).

De la mateixa forma, reiterant veiem que una funció diferenciable $f(x)$ definida en un rectangle obert de $\mathbb{R}^d$ (un producte cartesià d'intervals oberts, acotats o no) depen només de les variables $x_i, i \in A$ on A és un cert subconjunt d'indexs $A \subset \{1, 2, \cdots, d\}$ si i només si les derivades parcials respecte de les altres variables són zero. Per exemple, una funció $f(x,y,z)$ diferenciable de tres variables depen nomes de x si i només si les dues derivades parcials $\frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$ són idènticament zero. Si $\frac{\partial f}{\partial z} = 0$ dependrà només de x, y, etc.

Si l'obert U és un producte d'intervals oberts, podem doncs afirmar que si $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ existeixen i són zero, aleshores f és constant. Observem que això també és cert, més generalment, si la intersecció amb U de qualsevol recta paral·la a un dels eixos és un interval obert en aquesta recta (per exemple, l'obert que s'obté aplicant un moviment rígid a un rectangle). Observem també que un obert d'aquest tipus no és necessàriament convex (per exemple un obert en forma d'ela)

En el cas d'un obert arc-connex U, si f és diferenciable en U i $df(x) = 0$ per a tot $x \in U$, aleshores f és constant. En efecte, fixem $p \in U$ i ajuntem qualsevol $x \in U$ amb p mitjançant un arc $\gamma(t)$ regular, $\gamma(a) = p, \gamma(b) = x$, i mirem la funció $\phi(t) = f(\gamma(t))$; per la regla de la cadena té derivada $\phi'(t) = \langle \nabla f(\gamma(t)), \gamma'(t) \rangle = 0$, per tant és constant, i $f(x) = \phi(b) = \phi(a) = f(p)$. Més endavant veurem una versió quantitativa d'aquest fet.

La solució general de

$$\frac{\partial f}{\partial y}(x,y) = h(x,y),$$

amb h donada és

$$f(x,y) = \int h(x,y)\, dy + C(x),$$

amb C una funció arbitrària. La integral indefinida representa una primitiva de $h(x,y)$, com a funció de y, interpretant que x és constant. Per exemple, la solució general de

$$\frac{\partial f}{\partial y}(x,y) = xy,$$

és $f(x,y) = \frac{1}{2}xy^2 + C(x)$, amb C arbitrària, etc. D'això en diem antiderivades parcials.

8 Integrals depenents d'un paràmetre

Estudiarem aquí funcions definides per integrals de Riemann depenents d'un paràmetre

$$f(x) = \int_a^b F(x,t)\,dt.$$

Suposem primer que F és una funció contínua en $U \times [a,b]$, on U és un obert de $\mathbb{R}^d$, fet que assegura que f està ben definida. Afirmem que f és contínua en U. En efecte, fixat $x_0 \in U$, considerem una bola tancada $\overline{B}(x_0,r) \subset U$. Com que $\overline{B}(x_0,r) \times [a,b]$ és compacte, f és uniformement contínua: per a tot $\varepsilon > 0$ hi ha $\delta > 0$ tal que si $|t-t'| < \delta, \|x-x'\| < \delta, x, x' \in \overline{B}(x_0,r)$, llavors $|F(t,x)-F(t',x')| < \varepsilon$. En particular $|F(t,x)-F(t,x_0)| < \varepsilon$ per a tot t si $\|x-x_0\| < \delta < r$, i llavors

$$|f(x)-f(x_0)| \leq \int_a^b |F(t,x)-F(t,x_0)|\,dt \leq \varepsilon(b-a).$$

Suposem seguidament que existeix $\frac{\partial F}{\partial x_i}(t,x)$ i és una funció contínua en $U \times [a,b]$. Afirmem llavors que existeix $\frac{\partial f}{\partial x_i}(x)$ i que

$$\frac{\partial f}{\partial x_i}(x) = \int_a^b \frac{\partial F}{\partial x_i}(t,x)\,dt.$$

Per demostrar això podem suposar $d = 1$ i que U és un interval obert. Fixem $c \in U$ i prenem un interval tancat $[c-r, c+r] \subset U$. Cal veure que

$$\int_a^b \left| \frac{F(t,c+h)-F(t,c)}{h} - \frac{\partial F}{\partial x}(t,c) \right| dt \to 0, h \to 0.$$

Pel teorema del valor mig

$$\frac{F(t,c+h)-F(t,c)}{h} - \frac{\partial F}{\partial x}(t,c) = \frac{\partial F}{\partial x}(t,c+\theta h) - \frac{\partial F}{\partial x}(t,c), 0 < \theta < 1.$$

Per la continuitat uniforme de $\frac{\partial F}{\partial x}$ en $[a,b] \times [c-r,c+r]$ donat $\varepsilon > 0$ hi ha $\delta > 0$ tal que l'expressió anteior és ε per a tot t i θ si $|h| < \delta$ i ja.

Més generalment, suposem que volem derivar

$$g(x) = \int_{a(x)}^{b(x)} F(x,t)\,dt,$$

on els límits d'integració són funcions derivables de x i F és com abans. Introduim

$$f(x,y)=\int_a^y F(x,t)\,dt.$$

Llavors, pel que acabem de veure

$$\frac{\partial f}{\partial x}=\int_a^y \frac{\partial F}{\partial x}(x,t)\,dt,$$

mentre que pel teorema fonamental del càlcul

$$\frac{\partial f}{\partial y}=F(x,y).$$

Com que $g(x)=f(x,b(x))-f(x,a(x))$, per la regla de la cadena trobem

$$g'(x)=\frac{\partial f}{\partial x}(x,b(x))+\frac{\partial f}{\partial y}(x,b(x))b'(x)-\frac{\partial f}{\partial x}(x,a(x))-\frac{\partial f}{\partial y}(x,a(x))a'(x)=$$

$$=\int_{a(x)}^{b(x)}\frac{\partial F}{\partial x}F(x,t)\,dt+F(x,b(x))b'(x)-F(x,a(x))a'(x).$$

Una forma més entenedora de provar això és organitzar l'increment

$$g(x+h)-g(x)=\int_{a(x+h)}^{b(x+h)}F(x+h,t)\,dt-\int_{a(x)}^{b(x)}F(x,t)\,dt,$$

de la forma següent:

$$=\int_{a(x)}^{b(x)}(F(x+h,t)-F(x,t))\,dt+\int_{b(x)}^{b(x+h)}F(x+h,t)\,dt-\int_{a(x)}^{a(x+h)}F(x+h,t)\,dt.$$

9 Funcions amb gradient donat

En una variable $d=1$, sabem que tota funció contínua té una primitiva, una antiderivada; si h és contínua en un interval, hi ha f derivable amb $f'=h$ (la integral indefinida de h) i la solució general és $f+c$. Ens podriem plantejar si en $d>1$ tambñe podem prefixar $df(x), x\in U$, trobar funcions amb diferencial coneguda. Això és equivalent a resoldre el sistema d'equacions

$$\frac{\partial f}{\partial x_i}(x)=h_i(x), i=1,2,\cdots,d$$

amb h_i funcions prefixades. Això en general no és possible, cal que les funcions h_i compleixin una condició de compatibilitat. Expliquem primer la idea en dues variables. Suposem que volem resoldre

$$\frac{\partial f}{\partial x}(x,y) = h_1(x,y), \frac{\partial f}{\partial y}(x,y) = h_2(x,y),$$

amb h_1, h_2 de classe C^1 a tot el pla (o més generalment en un producte d'intervals oberts). De la primera veiem que

$$f(x,y) = \int_0^x h_1(t,y)\,dt + A(y),$$

amb A arbitrària. Per a aquesta f, que compleix la primera equació, i derivant sota el signe integral

$$\frac{\partial f}{\partial y}(x,y) = \int_0^x \frac{\partial h_1}{\partial y}\,dx + A'(y).$$

Si volem que això sigui $h_2(x,y)$, la diferència

$$\int_0^x \frac{\partial h_1}{\partial y}(t,y)\,dt - h_2(x,y)$$

ha de ser una funció només de y. És a dir, la derivada parcial respecte de x d'aquest diferència ha de ser zero, i trobem que

$$\frac{\partial h_1}{\partial y} = \frac{\partial h_2}{\partial x},$$

és una condició de compatibilitat necessària. L'argument que hem fet és reversible, si es compleix aquesta condició llavors

$$\frac{\partial f}{\partial y}(x,y) = \int_0^x \frac{\partial h_2}{\partial t}(t,y)\,dt + A'(y) = h_2(x,y) - h_2(0,y) + A'(y),$$

i ara tan sols cal resoldre $A'(y) = h_2(0,y)$.

En dimensió $d > 2$, un raonament anàleg mostra que l'equació $\nabla f(x) = h(x)$ (on f és escalar i h amb valors a $\mathbb{R}^d$), és a dir,

$$\frac{\partial f}{\partial x_i} = h_i(x),$$

amb h_i diferenciables, té solució si i només si per a cada parell i, j

$$\frac{\partial h_i}{\partial x_j} = \frac{\partial h_j}{\partial x_i}.$$

La solució s'obté pel procediment anterior, integrant succesivament cadascuna de les equacions respecte de la variable corresponent.

Més endavant, quan parlem de camps conservatius tornarem a aquest tema.

10 Els canvis de coordenades per a resoldre edp's

L'exemple següent resumeix alguns dels conceptes explicats fins ara. Al primer quadrant $U = \{x > 0, y > 0\}$ volem trobar la solució general de l'equació

$$f_x - f_y = (x + y)^2,$$

i també una solució que compleixi a més que $f(x, x) = 4x^2$. Introduim

$$v = x + y, u = x^2 - y^2.$$

Comprovem que u, v és un sistema de coordenades en U, és a dir que $\Phi : (x, y) \to (u, v)$ és injectiva, és a dir bijectiva sobre la imatge V, que V és obert i que l'aplicació és diferenciable en les dos sentits. Que Φ és injectiva es pot veure gràficament perquè una hiperbola $u = c_1$ i una recta de pendent -1 $v = c_2$ es tallen en un únic punt de U: Analíticament, també podem aïllar x, y en termes de u, v, perquè $u = v(x - y)$, d'on

$$x = \frac{1}{2}(v + \frac{u}{v}), y = \frac{1}{2}(v - \frac{u}{v}).$$

Amb això veiem que $V = \{(u, v) : v > 0, |u| < v^2\}$ és obert; evidentment tant Φ com la inversa són diferenciables.

Ara escriurem l'equació utilitzant les noves coordenades. Tenim

$$f_x = f_u u_x + f_v v_x = 2x f_u + f_v, f_y = -2y f_u + f_v,$$

d'on $f_x - f_y = 2(x + y) f_u = 2v f_u$ i l'equació queda $2v f_u = (x + y)^2 = v^2$, és a dir, $f_u = \frac{v}{2}$. Com que V talla els eixos u, v en intervals, la solució general d'aquesta equació és $f(u,v) = \frac{1}{2}uv + A(v)$, amb A una funció arbitrària d'una variable. Desfen el canvi, veiem que

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

és la solució general. Si a més volem que $f(x,x) = 4x^2$, ha de ser $A(2x) = 4x^2$, per tant $A(t) = t$ i la solució és

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

11 Dependència i indepèndencia funcional

Fixem un domini U de $\mathbb{R}^d$. Si f, g són funcions reals, definides en U, direm que f depen funcionalment de g en U si hi ha una funció $h : \mathbb{R} \longrightarrow \mathbb{R}$ tal que $f(x) = h(g(x))$. Això és el mateix que dir que el valor $f(x)$ tan sols depen del valor $g(x)$, per a tot x.

Considerem ara $k < d$ funcions $v_1, v_2, \cdots, v_k$ en U. Diem que una funció f definida en U depen funcionalment de $v_1, v_2, \cdots, v_k$ en U si hi ha una funció $h(y_1, \cdots, y_k)$ de k variables tal que

$$f(x) = h(v_1(x), v_2(x), \cdots, v_k(x)).$$

Això vol dir que el valor $f(x)$ només depen de quins són els valors

$$v_1(x), v_2(x), \cdots, v_k(x),$$

o que f és constant en cada conjunt del tipus $\{x \in U : v_1(x) = c_1, \cdots, v_k(x) = c_k\}$.

Per definició, si tenim un sistema de coordenades $v_1, \cdots, v_d$ en U, tota altra funció f depen funcionalment de $v_1, \cdots, v_d$. En dimensió $d = 1$, si g és injectiva en U, qualsevol altra funció f depen funcionalment de g.

La pregunta que ens formulem és: com podem saber si una funció f depen funcionalment de $v_1, \cdots, v_k$?. Ara veurem que a la categoria diferenciable, és a dir, si estem tractant amb funcions diferenciables a U, aleshores podem dir alguna cosa.

Primer observem que el concepte de dependència funcional és un concepte paralel al de dependència lineal de vectors. De fet hi té una relació. En efecte, suposem que estem a la categoria diferenciable, és a dir, que $f, v_1, v_2, \cdots, v_k$

són funcions diferenciables en U. Si f depen funcionalment de les v_k, per la regla de la cadena, tenim que

$$\frac{\partial f}{\partial x_i} = \frac{\partial h}{\partial y_1}\frac{\partial v_1}{\partial x_i} + \cdots + \frac{\partial f}{\partial y_k}\frac{\partial v_k}{\partial x_i}, i = 1, \cdots, d.$$

En forma més compacte,

$$\nabla f(x) = \frac{\partial h}{\partial y_1}(v(x))\nabla v_1(x) + \cdots + \frac{\partial h}{\partial y_k}(v(x))\nabla v_k(x).$$

Veiem per tant que una condició necessària és que en cada punt $x \in U$ el gradient de f és combinació lineal dels gradients $\nabla v_1(x), \cdots, \nabla v_k(x)$. Equivalentment, df és combinació lineal de $dv_1, \cdots, dv_k$. Dit d'una altra manera, si el rang de la matriu formada pels vectors fila $\nabla f(x), \nabla v_1(x), \cdots, \nabla v_k(x)$ és $k+1$ en un punt $x \in U$, llavors f no pot dependre funcionalment de les $v's$.

En àlgebra lineal, tenim el concepte d'independència lineal de vectors, cap d'ells es combinació lineal dels altres. Aqui tenim un concepte anàleg: diem que k funcions $v_1, v_2, \cdots, v_k$ són *funcionalment independents en un domini* U si en cap $U' \subset U$ cap d'elles és funcionalment dependent de les altres, no hi ha cap boleta dins U on una d'elles s'expressi en funció de les altres. Fixem-nos que ser funcionalment independents en U no és la negació lògica de ser funcionalment dependents en U, ho demanem per a tot obert $U' \subset U$. Hem vist que una condició que assegura això és que els gradients $\nabla v_1, \nabla v_2, \cdots, \nabla v_k$ siguin linealments independents en cada punt $x \in U$.

El problema natural és doncs, a partir d'una familia de funcions $v_1, \cdots, v_k$ amb gradients linealment independents en cada punt, donar un criteri sobre una funció f per tal que en depengui funcionalment en U. Hem vist que una condició necessaria és que en cada punt el gradient $\nabla f(x)$ ha de ser combinació lineal dels gradients $\nabla v_1(x), \cdots, \nabla v_k(x)$.

Veurem ara que aquesta condició és suficient en un cas particular, quan les k funcions $v_1, v_2, \cdots, v_k$ formen part d'un sistema de coordenades $\Phi = (v_1, v_2, \cdots, v_d)$ en l'obert U. És clar que en aquest cas són funcionalment independents, cosa que també podem veure observant que, donat que $d\Phi(x)$ és inversible, els seus vectors fila $\nabla v_1, \cdots, \nabla v_d$ són linealment independents en cada punt, per tant les $v_1, v_2, \cdots, v_k$ són funcionalment independents. En aquest cas, hem vist que

$$df = \frac{\partial f}{\partial v_1}dv_1 + \cdots + \frac{\partial f}{\partial v_d}dv_d,$$

per tant el fet que $\nabla f(x)$ sigui combinació lineal de $\nabla v_1(x), \cdots, \nabla v_k(x)$ en cada punt equival a dir que

$$\frac{\partial f}{\partial v_i} = 0, i = k+1, \cdots, d.$$

Com a conseqüència, si en les coordenades v l'obert és un rectangle, o més generalment talla qualsevol recta amb la direcció dels eixos en un interval, arribem a la conclusió que són equivalents en aquest cas els tres enunciats següents:

- La funció f depen funcionalment de $v_1, \cdots, v_k$ en U.
- $\nabla f(x)$ és combinació lineal de $\nabla v_1(x), \cdots, \nabla v_k(x)$ per a tot $x \in U$. És a dir, la matriu que té aquests gradients com files té rang k en cada punt.
- $\frac{\partial f}{\partial v_i} = 0, i = k+1, \cdots, d.$

Fixem-nos que la segona formulació no necessita coneixer quines són les altres funcions que amb $v_1, \cdots, v_k$ formen un sistema de coordenades, això tans sols és necessari per a la prova. Quan $k = d-1$, la condició és simplement que el determinant de la matriu formada pels vectors $\nabla v_j(x), j = 1, \cdots, d-1$ i $\nabla f(x)$ sigui zero en tot punt. Això ens donarà una equació en derivades parcials del tipus

$$\sum_{i=1} A_i(x) \frac{\partial f}{\partial x_i} = 0,$$

on $A_i(x)$ sónles components del producte vectorial $\nabla v_1(x) \times \cdots \times \nabla v_{d-1}(x)$. En general tindrem $d-k$ equacions com aquesta.

Una altra manera de pensar això mateix és la següent: busquem la solució general del sistema d'equacions en derivades parcials

$$\sum_{j=1}^{d} A_{ij}(x) \frac{\partial f}{\partial x_j} = 0, i = k+1, \cdots, d.$$

El que fem és escriure aquesta equació en un altre sistema de coordenades $v_1, \cdots, v_d$; si $A_{ij} = \frac{\partial x_j}{\partial v_i}$, les equacions, en el nou sistema de coordenades, es trivialitzen:

$$\frac{\partial f}{\partial v_i} = 0, i = k+1, \ldots, d,$$

i per tant la solució general és una funció de $v_1, \cdots, v_k$.

Considerem per exemple les coordenades polars (r, θ). Una funció $f(x, y)$ és una funció radial, és a dir, és de la forma $f(x, y) = h(\sqrt{x^2 + y^2})$ si i només si

$$\frac{\partial f}{\partial \theta} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial \theta} =$$

$$= \frac{\partial f}{\partial x}(-r \sin\theta) + \frac{\partial f}{\partial y} r \cos\theta = x\frac{\partial f}{\partial y} - y\frac{\partial f}{\partial x} = 0.$$

Serà una funció constant sobre semirectes si i només si

$$\frac{\partial f}{\partial r} = \cos\theta \frac{\partial f}{\partial x} + \sin\theta \frac{\partial f}{\partial y} = 0,$$

és a dir, si i només si $x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 0$.

En les coordenades esfèriques, una funció $f(x, y, z)$ depen funcionalment de $r^2 = x^2 + y^2 + z^2$ si i només si

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} + z\frac{\partial f}{\partial z} = 0,$$

i depen funcionalment de $r, \phi = \arccos\frac{z}{r}$ (és constant sobre paralels) si i només si

$$-y\frac{\partial f}{\partial x} + x\frac{\partial f}{\partial y} = 0.$$

Cóm podem reconeixer si k funcions formen part d'un sistema de coordenades en U?. Més endavant veurem el següent resultat local:

Teorema. *Un sistema de k funcions $v_1, \cdots, v_k$ de classe C^1 forma part d'un sistema de coordenades en un entorn d'a si i només si els seus gradients en a són linealment independents.*

Però a la pràctica ho podrem reconeixer directament. Considerem el cas més senzill possible, quan les funcions $v_1, \cdots, v_k$ són funcions lineals

$$v_j(x) = \sum_{i=1}^{d} a_{ij} x_i, j = 1, \cdots, k,$$

on la matriu $A = (a_{ij})$ té rang k, és a dir, els vectors columna són linealment independents, suposem que el menor (a_{ij}) té determinant diferentde zero. És evident que aquests vectors es poden ampliar a una base, per tant les funcions formen part d'un sistema de coordenades, lineal en aquest cas. Tindrem per tant que $f(x)$ és de a forma

$$f(x) = H(v_1(x), \cdots, v_k(x)),$$

si i només si els $d-k$ menors que s'obtenen orlant l'anterior són zero.

En el cas $d = 2, k = 1$, $f(x)$ és de la forma $f(x,y) = H(Ax+By)$ si i només si $Af_y - Bf_x = 0$. Ho podem pensar així: aquesta condició diu que la derivada direccional segons el vector $(-B, A)$ és zero; per tant f és constant sobre tota recta que tingui vector director $(-B, A)$, és a dir, de la forma

$$x(t) = a - Bt, y(t) = b + At, \mathrm{o} Ax + By = C.$$

Per tant f és constant allà on $Ax + By$ és constant. En el cas $d = 3, k = 2$, $f(x, y, z)$ és de la forma

$$f(x,y,z) = H(A_1x + B_1y + C_1z, A_2x + B_2y + C_2z)$$

si i només si

$$(B_1C_2 - C_1B_2)f_x + (C_1A_2 - A_1C_2)f_y + (A_1B_2 - A_2B_1)f_z = 0.$$

Novament, això ho podem interpretar així: l'anterior diu que la derivada direccional en la direcció de la recta intersecció de tot parell de plans $A_1x + B_1y + C_1z = D_1, A_2x + B_2y + C_2z = D_2$ és zero, per tant f és constant allà on $A_1x + B_1y + C_1z, A_2x + B_2y + C_2z$ són constants.

Tal com comentat abans, tot això ilustra en definitiva el perquè és útil el concepte de canvi de coordenades. Suposem que en $\mathbb{R}^3$ volem trobar la solució general de l'equació en derivades parcials

$$A(x,y,z)f_x + B(x,y,z)f_y + C(x,y,z)f_z = 0.$$

El que fem és esbrinar si en algun sistema de coordenades alternatiu (u, v, w) aquesta equació té una expressió més senzilla, i això hem vist que es pot fer en determinats casos.

12 Varietats regulars. La varietat tangent

Recordem el concepte de varietat de dimensió m de $\mathbb{R}^d$. És un conjunt M que té m graus de llibertat al voltant de cadascun dels seus punts. Expressat precisament: per a tot $p \in M$ hi ha una aplicació $H : U \longrightarrow \mathbb{R}^d$, on U és un obert de $\mathbb{R}^m$, i una bola $B(p,r)$ tal que H és un homeomorfisme entre U i $S \cap B(p,r)$ (bijecció contínua en els dos sentits). Això vol dir que $x = H(t), t = (t_1, \cdots, t_m) \in U$, parametritza $M \cap B(p,r)$.

Definició. Diem que M és una *varietat regular o varietat diferenciable de dimensió m de classe* C^1 expressada localment en forma paramètrica si per a tot $p \in M$ hi ha una aplicació $H : U \longrightarrow \mathbb{R}^d$, de classe C^1, on U és un obert de $\mathbb{R}^m$, amb $0 \in U, H(0) = p$ i una bola $B(p,r)$ tal que H és un homeomorfisme entre U i $S \cap B(p,r)$ (bijecció contínua en els dos sentits), i $dH(t)$ té rang m en tots els punts $t \in U$.

Per no complicar la notació, les d components de H les designem

$$x_i(t_1, \cdots, t_m), i = 1, \cdots, d.$$

El que H sigui un homeomorfisme significa que $t = (t_1, \cdots, t_m)$ són coordenades del punt $H(t) \in M$ al voltant de p. Si fixem tots els valors de les t_j menys un, t_i, que el deixem variar, $H(t)$ recorre un arc dins M, l'eix t_i.

Quan $m = 2$ parlem de superficies regulars i quan $m = 1$ de corbes regulars. Un cas particular d'aquesta definició és el dels gràfics, que vol dir que m de les variables $x_1, \cdots, x_d$ fan de paràmetres (suposem que són les primeres) i les altres en depenen, $x_i = h_i(x_1, \cdots, x_m), i = m+1, \cdots, d$. És a dir, H és de la forma

$$H(x_1, \cdots, x_m) = (x_1, \cdots, x_m, h_{m+1}(x_1, \cdots, x_m), \cdots, h_d(x_1, \cdots, x_m)).$$

Tornant a la definició general, és clar que és suficient imposar que el rang de $dH(0)$ sigui m; llavors vol dir que hi ha un menor d'ordre m de $dH(0)$ no nul, i per continuïtat també serà diferent de zero en una bola centrada en 0. De fet, veurem més endavant que si $dH(0)$ té rang màxim, automàticament H és injectiva en un entorn de 0. Per al cas $m = 1$ això és fàcil de demostrar directament.

Què hi afegeix, en termes geomètrics, el fet que la carta digui diferenciable ? Pel fet que H és un homeomorfisme, tot arc continu dins M és, en un entorn de p, de la forma $H(\gamma(s))$, amb $\gamma(s)$ dins U. El fet que H sigui diferenciable ens assegura senzillament que dins M *hi ha molts arcs regulars amb tangent*, prenent γ derivable. Suposant que $H(0) = p$, per a cada direcció $\vec{v}$, segons hem vist, si prenem qualsevol arc $\gamma(s)$ dins U que passi per 0 amb tangent $\vec{v}$, l'arc imatge $\hat{\gamma}(s) = H(\gamma(s))$ té un vector tangent que és $dH(0)(\vec{v}) = D_{\vec{v}}H(0)$. Com que aquesta corba és dins M, podem dir que $dH(0)(\vec{v})$ és tangent a M.

Per això definim *l'espai tangent a M en p*, designat $T_p(M)$, com el conjunt d'aquests vectors; és a dir, $T_p(M) = dH(0)(\mathbb{R}^m)$, el subespai imatge de l'aplicació lineal $dH(0)$, recull tots els vectors que són tangents a corbes dins M que passen per p. Com que per hipòtesi $dH(0)$ té rang m, l'espai tangent $T_p(M)$ té dimensió m. Una base de $T_p(M)$ consta dels vectors columna de $dH(0)$, que són $D_iH(0), i = 1, \cdots, d$; aquests vectors són els vectors tangents als eixos t_i dins M, i s'escriuen

$$D_iH(0) = (\frac{\partial x_1}{\partial t_i}, \cdots, \frac{\partial x_d}{\partial t_i}),$$

i els designem $\frac{\partial}{\partial t_i}$, d'una forma anàloga a la notació per a un canvi de coordenades.

És immediat veure que aquesta definició no depèn de la carta escollida. El que nosaltres visualitzem com a tangent és la varietat afí tangent, que és $p + T_p(M)$.

Pel cas $m = 1$, cal parar atenció a la distinció entre arcs regulars, que són senzillament aplicacions $\gamma(t)$ derivables, i corbes regulars, que són conjunts Γ que localment es parametritzen per aplicacions $\gamma(t)$ amb derivada $\gamma'(t) \neq 0$. Si $\gamma(t)$ és un arc regular, no sempre el seu recorregut Γ és una corba regular. Per exemple, l'arc regular $\gamma(t) = (t^2, t^3), -1 < t < 1$ té imatge el conjunt $\Gamma = \{(x, y) : x = y^{\frac{2}{3}}\}$ que no és una corba regular en $(0, 0)$, on hi té una punxa. Es pot veure que si $\gamma(t)$ és un arc regular amb γ' contínua i $\gamma'(t_0) \neq 0$, llavors per un $\delta > 0$ prou petit el recorregut $\gamma(t_0 - \delta, t_0 + \delta)$ és una corba regular. Això no vol dir que si $\gamma'(t_0) = 0$ el recorregut ja no pugui ser una corba regular. Per exemple, $\gamma(t) = (t^3, t^6)$ té derivada zero en zero, però el seu recorregut, que és $y = x^2$ és una corba regular.

La majoria d'exemples que coneixem, la majoria de còniques i quàdriques, compleixen això, tenen una varietat tangent en cada punt. Una excepció seria el con, el conjunt $C = \{(x, y, z); z = \sqrt{x^2 + y^2}\}$ que té pla tangent en tots els punts llevat del vertex $(0, 0, 0)$. La parametrització global $H(x, y) = (x, y, \sqrt{x^2 + y^2})$ és un homeomorfisme de $\mathbb{R}^2$ en C, però no és diferenciable en $(0, 0)$. El con desprovist del vertex sí és una superficie diferenciable.

Igual que per a les varietats lineals, una altra manera intuïtiva de donar un objecte de $m = d - k$ dimensions és imposar k equacions. El concepte que estem discutint és local, per tant no cal que les k funcions serveixin globalment (sí serà així a la majoria d'exemples, però).

Definició. Diem que M és una varietat regular de dimensió m de classe C^1 expressada localment en forma contínua si per a cada $p \in M$ hi ha una bola $B(p,r) \subset U$ i $k = d - m$ funcions f_j escalars definides i de classe C^1 en $B(p,r)$, amb gradients $\nabla f_j(x), j = 1, \cdots, k = d-m$ linealment independents tals que

$$M \cap B(p,r) = \{x \in B(p,r) : f_j(x) = c_j, j = 1, \cdots, k\}.$$

M està donada localment per k equacions.

Novament, és suficient suposar que els gradients són linealment independents en p, per continuïtat ho seran també en una boleta centrada en p.

D'on surt aquesta condició sobre els gradients? Perquè s'hi posa? Es pot explicar de la següent manera: si volem que cada lligam que posem redueixi la dimensió una unitat, cal que cap lligam sigui conseqüència dels altres, i també cal evitar que siguin incompatibles. Si són incompatibles, seria $M = \emptyset$, i si de les k condicions, n'hi ha alguna que és conseqüència de les altres, llavors es pot treure i estariem imposant $k - 1$ equacions, o menys, enlloc de k. La idea és doncs que aquestes k funcions siguin funcionalment independents; i com hem vist a la secció anterior el fet que els gradients siguin linealment independents en cada punt ho assegura.

De moment acceptarem el teorema següent, que és una formulació geomètrica del teorema de la funció inversa.

Teorema. *Les dues definicions són equivalents.*

Cóm es descriu $T_p(M)$ si M està donada en forma contínua? Amb les notacions anteriors, si $\hat{\gamma}(t) = (\hat{\gamma}_1(t), \cdots, \hat{\gamma}_d(t))$ és una corba dins M que passi per $p, \hat{\gamma}(0) = p$ amb tangent $\hat{\gamma}'(0) = u \in T_p(M)$, derivant en $t = 0$ l'equació $f_j(\hat{\gamma}_1(t), \cdots, \hat{\gamma}_d(t)) = c_j$ la regla de la cadena ens diu que

$$\langle \nabla f_j(p), u \rangle = \sum_{i=1} D_i f(p) x_i'(0) = 0.$$

Això ens diu que els vectors $\nabla f_j(p)$ són perpendiculars a $T_p(M)$, i d'això en diem ser perpendicular a M. Si V és el subespai generat pels $\nabla f_j(p)$, que té dimensió k, hem vist que V és ortogonal a $T_p(M), V^\perp \subset T_p(M)$, però com que $V^\perp$ i $T_p(M)$ tenen la mateixa dimensió m són iguals:

$$T_p(M) = \{u : \langle \nabla f_j(p), u \rangle = 0, j = 1, \cdots, k\},$$

V i $T_p(M)$ són subespais cadascun ortogonal de l'altre.

Per exemple, si una corba regular Γ a l'espai és la intersecció de dues superficies

$$f(x, y, z) = 0, g(x, y, z) = 0,$$

amb gradients linealment independents, la recta tangent a la corba té vector director perpendicular a $\nabla f(x)$ i $\nabla g(x)$ en cada punt $x \in \Gamma$, per tant en la direcció del producte vectorial $\nabla f(x) \times \nabla g(x)$.

13 El mètode dels multiplicadors de Lagrange. Extrems absoluts en conjunts compactes amb frontera regular

Reprenem el que s'ha dit en la secció anterior per a conjunts de nivell definits per equacions. Clàssicament, cada equació s'anomena un *lligam* que posem a les d variables. Si tenim $k = d - m$ funcions $f_1, \cdots, f_k$ definides en U, considerem $F = (f_1, \cdots, f_k)$ i

$$M = \{x \in U : f_1(x) = c_1, \cdots, f_k(x) = c_k\} = \{x \in U : F(x) = c\}.$$

Suposem que $p \in M$ i que el rang de $dF(p)$ és k (és a dir, els k gradients $\nabla f_j(p)$ són linealment independents), llavors M és una varietat regular de dimensió $m = d - k$ en un entorn d'p, i si V és el subespai vectorial generat per $\nabla f_1(p), \cdots, \nabla f_k(p)$, V i $T_p(M)$ cadascun és el ortogonal de l'altre.

Suposem ara que tenim una funció escalar h en U que, restringida a M té un extrem relatiu en p, és a dir, $h(x) - h(p)$ té signe constant per a $x \in M \cap B(p, r)$. D'aquesta situació se'n diu que h té un *extrem relatiu en p, condicionat* per les k equacions $f_k(x) = c_k$. Per contraposició, quan no hi ha lligams es parla d'extrems lliures.

Llavors, si $\gamma(s), \gamma(0) = p$ és un arc regular dins M (el que abans n'hi deiem $\hat{\gamma}$), la funció $h(\gamma(s))$ tindrà un extrem relatiu en $s = 0$, i per tant té derivada zero. Ara bé, per la regla de la cadena, aquesta derivada és $D_u h(p)$, on $u = \gamma'(0)$; com que al variar γ obtenim tots els $u \in T_p(M)$, veiem que $D_u h(p) = 0$ per a tot $u \in T_p(M)$, que és el mateix que dir que $\nabla h(p)$ és ortogonal a $T_p(M)$. Això mateix ho podem veure utilitzant una carta local $H(t), H(0) = p$, al voltant d'p: si h té un extrem relatiu condicionat en p, aleshores $h(H(t))$ té un extrem relatiu en $t = 0$ sense lligams, ha de tenir gradient zero, i això ens dóna altre cop que $\nabla h(p)$ és ortogonal a $T_p(M) = dH(0)(\mathbb{R}^m)$.

Com l'ortogonal de $T_p(M)$ és V, tenim que $\nabla h(a) \in V$, ha de ser combinació lineal dels $\nabla h_j(a), j = 1, \cdots, k$. Per tant hi ha escalars $\lambda_j, j = 1, \cdots, k$ tals que

$$\nabla h(a) = \lambda_1 \nabla f_1(a) + \cdots + \lambda_k \nabla f_k(a).$$

Els λ_j s'anomenen *multiplicadors de Lagrange*.

Aquesta anàlisi permet donar un mètode per a trobar els extrems relatius condicionats per k lligams $f_j = c_j$ si es compleix la hipòtesi per a cada punt que compleix els lligams, és a dir, si

$$M = \{x \in U : f_1(x) = c_1, \cdots, f_k(x) = c_k\} = \{x \in U : F(x) = c\},$$

és una subvarietat regular de dimensió $m = d - k$. Aquests extrems relatius han d'estar entre les solucions x del sistema de $d + k$ equacions escalars

$$\nabla h(x) = \lambda_1 \nabla f_1(x) + \cdots + \lambda_k \nabla f_k(x), f_1(x) = c_1, \cdots, f_k(x) = c_k,$$

(la primera és una igualtat vectorial, per tant d equacions escalars), on les incògnites són les d coordenades dels punts candidats x, i les $\lambda_1, \cdots, \lambda_k$. Observem que la primera equació també pot formular-se dient que x és un punt crític de $h - \sum_j \lambda_j f_j$. Entre les solucions d'aquest sistema hi ha els possibles extrems relatius condicionats, i caldrà fer un estudi particular en cadascun per decidir si ho són o no. És a dir, l'anterior és una condició necessària, però no suficient.

Alternativament, si disposem, com és el cas a la majoria de casos pràctics, d'una parametrització global de M, diguem per $H(t), t \in R^m$, els extrems relatius condicionats de h són els extrems lliures de la composició $h(H(t))$ i podem trobar-los entre els seus punts crítics directament (això és factible si $m = 1$, per exemple).

Suposem ara que U és un domini acotat i que la frontera $Fr(U) = M$ és una varietat regular de dimensió $d-1$ en les condicions anteriors, per exemple una bola o un elipsoide. En aquesta situació, hom diu que U és un *domini amb frontera regular*. Suposem que h és una funció escalar, diferenciable arreu; en el compacte $\overline{U}$, la funció h pren un valor màxim absolut M i un valor mínim absolut m, diguem en punts $a, b \in \overline{U}$. Si aquests punts són a U, llavors són extrems relatius i per tant s'hi anul·larà el gradient, són punts crítics de h. Si estan a la frontera $Fr(U)$, llavors són extrems relatius condicionats i han d'estar entre les solucions del sistema de Lagrange. Com que sabem que a, b existeixen, per trobar m, M farem el següent: d'una banda trobarem tots els punts crítics de h dins U i d'altra banda trobarem tots els punts de la frontera que són solució del sistema de Lagrange. Típicament, un i altre tindran tan sols un nombre finit de solucions. Avaluant h en cadascun d'ells veurem quin és el valor màxim M i el valor mínim m.

A la secció 8 hem vist que si $v_1, \cdots, v_k$ són funcions diferenciables que formen part d'un sistema de coordenades en el que el domini U talla totes les rectes amb direccions paraleles als eixos en intervals, aleshores una funció f depen funcionalment de les $v_1, \cdots, v_k$ si i només si el gradient $\nabla f(x)$ és combinació lineal dels gradients $\nabla v_1(x), \cdots, \nabla v_k(x)$ en cada punt. Ara, gràcies a la visió geomètrica podem generalitzar aquest resultat i entendre'l d'una altra forma.

Proposició. *Suposem que $\nabla v_1(x), \cdots, \nabla v_k(x)$ són linealment independents en cada punt, de forma que els conjunts de nivell*

$$M = \{x \in U : f_1(x) = c_1, \cdots, f_k(x) = c_k\} = \{x \in U : F(x) = c\},$$

són varietats regular de dimensió $m = d - k$. Suposem que aquests conjunts de nivell són arcconnexos, és a dir, donats dos punts $p, q \in M$, hi ha un arc regular dins M que els ajunta. Aleshores una funció f diferenciable depen funcionalment de $v_1, \cdots, v_k$ en U si i només si $\nabla f(x)$ és combinació lineal de $\nabla v_1(x), \cdots, \nabla v_k(x)$ per a tot $x \in U$.

Demostració. Que la condició és necessària és un fet general, només cal veure que és suficient. Com que el subespai generat per $\nabla v_1(x), \cdots, \nabla v_k(x)$ és l'ortogonal de $T_x(M)$, tindrem que $D_u f(x) = \langle \nabla f(x), u \rangle = 0$ per a tot $u \in T_x(M)$. Per tant f és constant al llarg de qualsevol arc dins M, i si M és arconnex, f serà constant en M, i això és el mateix que dir que f depen funcionalment de les $v's$, $f(x) = H(v_1(x), \cdots, v_k(x))$. La funció H és diferenciable, perquè localment els $v_1, \cdots, v_k$ formen part d'un sistema de coordenades.

$\square$

14 El teorema de la funció inversa. Dependència i independència funcional.

Suposem que f és de classe C^1 en un interval obert que conté a. Si $f'(a) > 0$, per continuïtat serà també $f'(x) > 0$ en un intervalet obert V que conté a, per tant f és estrictament creixent V, per tant $f(V)$ és un interval obert W que conté $f(a)$ i la funció inversa $g = f^{-1}$ és també estrictament creixent de W sobre V. Anàlogament si $f'(a) < 0$. Si $y \in W, f(x) = y$ i $k \to 0$, llavors $h = g(y+k) - g(y) \to 0$ (perquè g és contínua) i la igualtat

$$\frac{g(y+k) - g(y)}{k} = \frac{h}{f(x+h) - f(x)},$$

demostra que g és derivable en y amb derivada $g'(y) = 1/f'(x)$.

Un resultat semblant a aquest és vàlid en qualsevol dimensió d, però molt més complicat de provar. Recordem que un C^1- difeomorfisme entre dos oberts U, V de $\mathbb{R}^d$ és una bijecció f entre ells que és diferenciable de classe C^1 en els dos sentits. Si $g = f^{-1}$, la regla de la cadena implica que si $y = f(x)$, aleshores $df(x)$ i $dg(y)$ són aplicacions lineals inversa una de l'altra, i en particular amb determinant diferent de zero.

Comencem observant que hi ha homeomorfismes de classe C^1 que no són difeomorfismes. Ja en dimensió $d = 1$ tenim l'exemple trivial de $f(x) = x^3$. No és un difeomorfisme perquè té derivada zero en zero; també podem verure directament que la inversa $y^{\frac{1}{3}}$ no és derivable en zero.

Proposició.

- *Suposem que f és un homeomorfisme entre dos oberts U, V amb inversa g, i suposem que f és diferenciable en $a \in U$. Aleshores, per tal que g sigui diferenciable en $b = f(a)$ és necessari i suficient que $df(a)$ sigui inversible com a aplicació lineal, és a dir, que* $\det df(a) \neq 0$ *i, en aquest cas, $dg(b) = (df(a))^{-1}$.*

- *Sigui f un homeomorfisme de classe C^1 entre dos oberts U, V de $\mathbb{R}^d$. Per tal que f sigui un C^1 difeomorfisme, és necessari i suficient que $df(x)$ sigui inversible com a aplicació lineal per a tot $x \in U$, és a dir, que* $\det df(x) \neq 0$.

Demostració. Ja hem vist que la condició $\det df(a) \neq 0$ és necessària. Per veure que és suficient, com abans posem $x = a + h$ per a x proper a a, $y = f(a+h) = b + k$; com que f és un homeomorfisme, h, k es determinen un a l'altre i dir que $h \to 0$ és el mateix que dir $k \to 0$. De fet, per hipòtesi

$$k = f(a+h) - f(a) = df(a)(h) + \epsilon(h)\|h\|,$$

amb $\epsilon(h) \to 0$ quan $h \to 0$, en particular $k = O(h)$. Aplicant la inversa $L = df(a)^{-1}$ a aquesta igualtat trobem

$$h = L(k) - \|h\| L(\epsilon(h)),$$

és a dir,

$$g(b+k) - g(b) = L(k) - \|h\| L(\epsilon(h)).$$

Per tant acabarem si demostrem que el darrer terme és $o(k)$. Com que evidentment és $o(h)$, és suficient que provem que $h = O(k)$. Ara bé,

$$\|h\| \leq \|L(k)\| + \|h\| \|L(\epsilon(h))\| \leq \|L\| \|k\| + \|h\| \|L\| \|\epsilon(h)\|,$$

on $\|L\|$ indica la norma de l'aplicació lineal L. Com que $\epsilon(h) \to 0$ quan $h \to 0$, hi ha $\delta > 0$ tal que $\|\epsilon(h)\| \|L\| < \frac{1}{2}$ si $\|h\| < \delta$. Aleshores, si $\|h\| < \delta$,

$$\|h\| \leq \|L\| \|k\| + \frac{\|h\|}{2},$$

és a dir, $\|h\| \leq 2\|L\| \|k\|$, i $h = O(k)$.

La segona part és immediata, perquè les derivades parcials de g, les entrades de la matriu inversa de df, seran expressions en termes de les derivades parcials de f i del determinant de df, i per tant seran tambe contínues.

$\square$

El següent és el teorema de la funció inversa:

Teorema. . *Suposem que $f : U \to \mathbb{R}^d$ és de classe C^1 en un domini U de $\mathbb{R}^d$, $a \in U$ i que $df(a)$ és inversible, és a dir, $\det df(a) \neq 0$. Aleshores hi ha un obert $W \subset U$ que conté a, tal que f és un C^1 difeomorfisme entre W i un obert $V = f(W)$ de $\mathbb{R}^d$.*

Breument, diem que f és un C^1- difeomorfisme local en a. Amb el llenguatge dels sistemes de coordenades, la conclusió és que si $f = (u_1, \cdots, u_d)$, les funcions $u_1, \cdots, u_d$ són un sistema de coordenades diferenciable en W.

Abans de fer la demostració d'aquest teorema expliquem-ne la idea essencial. Per la mateixa definició de diferenciabilitat,

$$f(x) = f(a) + df(a)(x - a) + o(x - a).$$

Això vol dir que si ens mirem f com una transformació que mou punts, f es una transformació afí, $f(a) + df(a)(x - a)$, més un error $o(\|x - a\|)$. Aquesta transformació afí és bijectiva exactament quan $df(a)$ té determinant no nul. El que ens diu el teorema és que la funció f té localment en a la propietat que $df(a)$ té globalment. La demostració consisteix tota ella en *absorbir* el terme d'error $o(x - a)$.

Quelcom que farem servir és la següent observació aplicada a la diferencial $df(a)$: si T és una aplicació lineal de $\mathbb{R}^d$ en si mateix inversible, amb determinant diferent de zero, aleshores hi ha una constant $m > 0$ tal que $\|L(u)\| \geq m\|u\|$ per a tota direcció u. Que T sigui inversible és el mateix que dir que és injectiva, és a dir que $Tu = 0$ implica $u = 0$; la desigualtat anterior quantifica això. Per veure l'existència de m tan sols hem de considerar la norma M de T^{-1}; per definició es compleix

$$\|T^{-1}(u)\| \leq M\|u\|,$$

que aplicat a Tu enlloc d'u diu que $\|u\| \leq M\|Tu\|$, i $m = \frac{1}{M}$.

Demostració. Sigui, com acabem de dir, $m = \frac{1}{M}$ on M és la norma de l'aplicació lineal $df(a)^{-1}$, per tant

$$\|df(a)(\vec{u})\| \geq m\|\vec{u}\|.$$

En una primera etapa, veurem que hi ha una bola tancada $\overline{B} = \overline{B}(a, r) \subset U$ amb les tres propietats següents:

1. $\det df(x) \neq 0, x \in \overline{B}$.
2. $\|df(x) - df(a)\| \leq \frac{m}{2}$, on el terme esquerra és la norma de l'aplicació lineal $df(x) - df(a)$.
3. $\|f(x) - f(y)\| \geq \frac{m}{2}\|x - y\|, x, y \in \overline{B}$

Trobarem una bola $\overline{B}$ que compleixi els dos primers enunciats i després veurem que el segon implica el tercer. És trivial que el primer enunciat es compleix en una boleta prou petita, perquè al ser f de classe C^1 la funció

$\det df(x)$ és contínua. Per al segon també, si recordem que en ser totes les normes equivalents en un espai de dimensió finita, tindrem que

$$\|df(x) - df(a)\| \leq C \sum_{i,j} |\frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(a)|,$$

també és arbitràriament petit si les derivades parcials són contínues. Veiem ara el tercer enunciat. Considerem l'aplicació $\hat{f}(x) = f(x) - df(a)(x)$, que té diferencial $d\hat{f}(x) = df(x) - df(a)$; per tant $\|d\hat{f}(x)\| \leq \frac{m}{2}$ si $x \in \overline{B}$ i pel teorema del valor mig tindrem

$$\|f(x) - f(y) - df(a)(x-y)\| = \|\hat{f}(x) - \hat{f}(y)\| \leq \frac{m}{2}\|x-y\|, x, y \in \overline{B}.$$

Per la observació anterior, hi ha $m > 0$ tal que $\|df(a)(x-y)\| \geq m\|x-y\|$. Aleshores tenim

$$\big|\|f(x) - f(y)\| - \|df(a)(x-y)\|\big| \leq \|f(x) - f(y) - df(a)(x-y)\| \leq$$
$$\leq \frac{m}{2}\|x-y\| \leq \frac{1}{2}\|df(a)(x-y)\|,$$

la qual cosa diu que

$$\frac{1}{2}\|df(a)(x-y)\| \leq \|f(x) - f(y)\| \leq \frac{3}{2}\|df(a)(x-y)\|,$$

i per tant, si $C = \|df(a)\|$,

$$\frac{m}{2}\|x-y\| \leq \|f(x) - f(y)\| \leq \frac{3C}{2}\|x-y\|, x, y \in \overline{B}.$$

Ara considerem la frontera S de $\overline{B}$, que és compacte, per tant $f(S)$ també ho és, i no conté $f(a)$, pel punt 3. Si $d = d(f(a), f(S)) > 0$, definim $V = B(f(a), \frac{d}{2})$, de manera que $\|y - f(a)\| < \|y - f(x)\|$ per a tot $x \in S$. Definim W com l'obert

$$W = \{x \in B : f(x) \in V\}.$$

Pel punt 3, f és injectiva en B, de forma que f és injectiva de W en V. Ara veurem que per a tot punt $p \in V$ hi ha $q \in B$ tal que $f(q) = p$: si

$$h(x) = \|p - f(x)\|^2 = \sum_{i=1}^{d}(p_i - f_i(x))^2,$$

h té un mínim absolut en el compacte $\overline{B}$, i com que $h(a) < h(x), x \in S$, s'assoleix en un punt $q \in B$. Per tant $D_j h(q) = 0$, és a dir

$$\sum_i D_j f_i(q)(p_i - f_i(q)) = 0, j = 1, \cdots, d.$$

Ara bé, $df(x)$ és inversible en tot $x \in B$, i per tant $p_i = f_i(q), p = f(q)$. Tot plegat, f és una bijecció de W en V, i inversa és contínua, novament pel punt 3. La proposició anterior acaba la prova del teorema.

□

Si la hipòtesi $\det df(x) \neq 0$ es compleix en tot punt $x \in U$ tindrem que f és localment injectiva i $f(U)$ és un conjunt obert, per tant

Corol·lari. *Una aplicació $f : U \to \mathbb{R}$ de classe C^1 és un difeomorfisme si i només si és injectiva i $\det df(x) \neq 0$ per a tot $x \in U$.*

La injectivitat global no és en general conseqüència d'hipòtesis de tipus local, com $\det df(x) \neq 0$. La funció en el pla $f(x, y) = (x^2 - y^2, 2xy)$ (que és $f(z) = z^2$ amb la identificació $z = x + iy$) compleix $\det df(x, y) \neq 0$ fora de l'origen, és localment injectiva i no és globalment injectiva. En canvi, en dimensió $d = 1$ sí que $f'(x) \neq 0$ implica que f és injectiva, perquè aleshores o bé $f'(x) > 0$ per a tot x i f és estrictament creixent en l'interval U, o bé $f'(x) < 0$ i f és estrictament decreixent.

Des del punt de vista dels sistemes de coordenades, el teorema de la funció inversa assegura que si $v_1, \cdots, v_d$ són funcions de classe C^1 amb gradients $\nabla v_j(a)$ linealment independents, aleshores formen un sistema de coordenades en una boleta centrada en a, en diem un *sistema de coordenades local* en a.

Proposició. *Un sistema de k funcions $v_1, \cdots, v_k$ de classe C^1 forma part d'un sistema de coordenades local en a si i només si tenen gradients (o diferencials) linealment independents en a.*

Demostració. La matriu que té per files els gradients $\nabla v_j(a)$ tindrà un menor d'ordre k no nul. Suposant sense pèrdua de generalitat que és el de les primeres k columnes, $\det(D_i v_j(a)) \neq 0$, el sistema $(v_1, \cdots, v_k, x_{k+1}, \cdots, x_d)$ té gradients linealment independents en a i per tant és un sistema de coordenades local. □

El següent corol·lari és conseqüència de la proposició anterior i de l'explicat a la secció 8.

Corol·lari. *Si un sistema de k funcions $v_1, \cdots, v_k$ de classe C^1 té gradients linealment independents en a, una funció f de classe C^1 definida en un entorn d'a depen funcionalment de $v_1, \cdots, v_k$ en un entorn d'a si i només si $\nabla f(x)$ és combinació lineal de $\nabla v_1(x), \cdots, \nabla v_k(x)$ en un entorn d'a.*

15 Teorema de la funció implícita

El teorema de la funció implícita té dues lectures, una analítica i una geomètrica, equivalents. La primera tracta el concepte de funció implícita. Per a una funció $f(x, y)$ de dues variables en un domini U, diem que

$$f(x, y) = 0,$$

defineix implícitament $y = h(x)$ si el conjunt $M = \{(x, y) \in U : f(x, y) = 0\}$ és el gràfic $y = h(x)$ d'una certa funció h, la qual cosa vol dir que M talla tota recta vertical com a molt en un punt. En moltes circumstàncies concretes podem veure gràficament que hi ha funció implícita.

El teorema de la funció implícita dóna un criteri *local* que assegura l'existència de la funció implícita i, el que és més important, que la funció h té el mateix grau de diferenciabilitat que f. En molts exemples no ens cal el teorema per veure que l'equació defineix una funció implícita, ho veiem gràficament; el teorema ens diu que té el mateix grau de diferenciabilitat que l'equació. En aquesta secció, sovint posarem $\mathbb{R}^d = \mathbb{R}^k \times \mathbb{R}^m, m = d - k$ i designarem les variables per $x = (x', x'')$.

Teorema. *Siguin $v_1, v_2, \cdots, v_k$ funcions de classe C^1 definides al voltant del punt a amb gradients $\nabla v_j(a)$ linealment independents. Aleshores (suposant $v_j(a) = 0$) el sistema d'equacions*

$$\{x \in U : v_1(x) = 0, \cdots, v_k(x) = 0\},$$

defineix unívocament m de les variables com a funcions implícites de les restants en un entorn del punt a. Més precisament, i suposant com abans que $D_i v_j(a))_{i,j=1}^k$ té determinant diferent de zero, i si $a = (a', a'')$, hi ha un obert W que conté a, un obert U'' de $\mathbb{R}^m$ que conté a'' i una funció $h : U'' \to \mathbb{R}^k$ de classe C^1 tal que

$$M = \{x \in W : v_j(x) = 0, j = 1, \cdots, k\} =$$
$$= \{x = (x', x'') \in W : x' = h(x''), x'' \in U''\}.$$

Demostració. Podem suposar que $a = 0$. Considerem com abans un obert W on $(v_1, \cdots, v_k, x_{k+1}, \cdots, x_d)$ formen un sistema de coordenades. Això vol dir que

$$\Phi(x) = (v_1(x), \cdots, v_k(x), x_{k+1}, \cdots, x_d)$$

és un C^1-difeomorfisme de W en un obert V que conté zero. La idea és que en aquestes coordenades, els punts de M són els que tenen $v_1 = \cdots = v_k = 0$, i per tant estan parametritzats per les altres coordenades $x_{k+1}, \cdots, x_d$. Més precisament, considerem l'inversa Ψ de Φ, que també serà de la forma

$$\Psi(y) = (u_1(y), \cdots, u_k(y), y_{k+1}, \cdots, y_d),$$

amb les u_j de classe C^1. Aleshores si $x \in M, \Phi(x) = (0, x'')$ i $\Phi(M)$ consta per tant dels punts de la forma $(0, y'')$ de V. Posem $U'' = \{y'' \in \mathbb{R}^m : (0, y'') \in V\}$, que també podem escriure

$$U'' = \{x'' \in \mathbb{R}^m : \text{hi ha } x' \in \mathbb{R}^k, (x', x'') \in M \cap W\},$$

la projecció de $M \cap W$ en les darreres m coordenades. Aleshores U'' és un obert de $\mathbb{R}^m$ i

$$M = \{x = (x', x'') \in W : x_1 = u_1(0, x''), \cdots, x_k = u_k(0, x''), y''), y'' \in U''\},$$

és a dir, tan sols cal definir $h(x'') = (u_1(0, x''), \cdots, u_k(0, x''))$. $\square$

Una altra forma d'enunciar el teorema és la següent: tenim un sistema d'equacions en un obert

$$v_1(x) = 0, \cdots, v_k(x) = 0,$$

i un punt $a \in U$ que les compleix. El teorema diu que si en la matriu que té per files els $\nabla v_j(a)$ el menor d'ordre k que té per columnes les derivades parcials respecte k de les variables $x_i, i = 1, \cdots, d$ no és zero, aleshores aquestes k variables que intervenen en el menor queden automàticament definides com a funció implícita de les restants. En termes col·loquials diríem que aquestes variables es poden aïllar com a funció de les restants: en un entorn d'a' hi ha una única funció $h(x'')$ que compleix $h(a'') = a'$ i

$$v_j(h(x''), x'') = 0, j = 1, \cdots, k,$$

idènticament, o bé que $v_j(x', x'') = 0, j = 1, \cdots, k$ equival a $x' = h(x'')$.

Cal parar atenció al fet que la funció implícita és única, sota les hipòtesis anteriors, si compleix $h(a'') = a'$. Ara bé, per a un mateix a'' poden haver-hi diversos a' diferents complint $h(a', a'') = 0$, i si en cadascun d'aquests punts es compleix la hipòtesi, hi haurà més d'una funció implícita definida al voltant de a''. Per exemple, per a l'equació $x^2 + y^2 = 1$ hi ha dues funcions definides implícitament al voltant de 0: la funció $y = \sqrt{1 - x^2}$ que correspon al punt $(0, 1)$ i la funció $y = -\sqrt{1 - x^2}$ que correspon al punt $(0, -1)$.

De fet s'ha demostrat quelcom més general: posem

$$y_i = v_i(x', x''), i = 1, \cdots, k, y' = (y_1, \cdots, y_k).$$

Llavors

$$\Phi(x) = (y', x''), x = \Psi(y) = (u_1(y', x''), \cdots, u_k(y', x''), x'').$$

és a dir, $x'_i = u_i(y', x''), i = 1, \cdots, k$. Dit d'una altra manera, $y_i = v_i(x', x'')$ defineix unívocament x' en funció de $y_1, \cdots, y_k, x''$, per a tots els valors de $y_1, \cdots, y_k$, no solament 0.

16 Interpretació geomètrica del teorema de la funció implícita. Teorema del rang constant

En aquesta secció veurem l'equivalència entre les diverses maneres de definir subvarietats regulars de dimensió m. Per a M dins $\mathbb{R}^d$ tenim tres nocions locals:

1. La primera és que M estigui definit *localment per* $k = d - m$ *equacions funcionalment independents*: per a tot $p \in M$ hi ha una bola $B(p, r) \subset U$ i $k = d - m$ funcions v_j escalars definides i de classe C^1 en $B(p, r)$, amb gradients $\nabla v_j(x), j = 1, \cdots, k = d - m$ linealment independents tals que

$$M \cap B(p, r) = \{x \in B(p, r) : v_j(x) = 0, j = 1, \cdots, k\}.$$

2. La segona és que sigui localment un gràfic d'una funció de m variables: per a tot $p \in M$ hi ha una bola $B(p, r)$ i m de les variables variables (suposem que són les m darreres $x_{k+1}, \cdots, x_d$), tals que $M \cap B(p, r)$ és el gràfic d'una

funció $h : U'' \to \mathbb{R}^k$ de classe C^1 definida en un obert U'' de $\mathbb{R}^m$ que conté p''.

3. La tercera és que estigui parametritzat localment per una carta: per a tot $p \in M$ hi ha una aplicació $H : U \longrightarrow \mathbb{R}^d$, de classe C^1, on U és un obert de $\mathbb{R}^m$, amb $0 \in U, H(0) = p$ i una bola $B(p, r)$ tal que H és un homeomorfisme entre U i $S \cap B(p, r)$ (bijecció contínua en els dos sentits), i $dH(t)$ té rang m en tots els punts $t \in U$.

Aquestes tres definicions són equivalents. Que 1 implica 2 és exactament l'enunciat del teorema de la funció implícita. Que el menor B' de la matriu B que té per files $\nabla v_j(p)$, corresponent a les derivades parcials respecte k de les les coordenades sigui no nul significa geomètricament que l'espai tangent $T_p(M)$, que és l'ortogonal del subespai generat per aquests gradients, és parametritzable per les altres: suposant que són les k primeres, els vectors $X \in T_p(M)$ compleixen

$$BX' + B''X'' = 0,$$

i per tant es parametritza globalment per X'': $X' = -B^{-1}B''X''$. Així, geomètricament, el teorema diu que les coordenades que globalment parametritzen l'espai tangent $T_p(M)$ parametritzen localment M.

Recíprocament, en la situació de 2, si $h = (h_1, \cdots, h_k)$, de forma que M consta dels punts $(h(x''), x'')$, M està definit localment per

$$v_j(x_1, \cdots, x_d) = x_j - h_j(x_{k+1}, \cdots, x_d) = 0, j = 1, \cdots, k.$$

També és evident que 2 implica 3, perquè un gràfic $x' = h(x'')$ dóna lloc a la parametrització $H(x'') = (h(x''), x'')$, i $dH(x'')$ té obviament rang m en tots els punts. Per tant tan sols falta veure que 3 implica 2; de fet la prova mostrarà que és suficient suposar que H és de classe C^1 i que $dH(0)$ té rang m, automàticament això implica, fent U més petit, que H és un homeomorfisme. Si $H = (h_1, \cdots, h_d)$, com que $dH(0)$ té rang m hi haurà un menor d'ordre m amb determinant no nul, suposem que és el format per les darreres m components $H'' = (h_{k+1}, \cdots, h_d)$. Llavors $\det H''(0) \neq 0$, i pel teorema de la funció inversa, H'' és un C^1-difeomorfisme entre un obert que continuem anomenant U i un obert V de $\mathbb{R}^m$ que conté p''. Això ja implica que H és un homeomorfisme sobre la imatge i $dH(x)$ té rang m en tot punt. Sigui $G : V \to U$ l'invers de H''; els punts de M al voltant de p són els de la forma

$$(H'(t), H''(t)), t \in U.$$

Si posem $x'' = H''(t)$, dir que t varia en U és el mateix que dir que x'' varia en V, i si ara prenem x'' com a paràmetre veiem que M es descriu per

$$(H'(G(x'')), x''), x'' \in V,$$

i és per tant el gràfic de $h(x'') = H'(G(x''))$.

Naturalment, les subvarietats de dimensió d de $\mathbb{R}^d$ són els oberts i les de dimensió zero, els punts.

Hem vist en particular que a la definició 3 és suficient que $dH(0)$ tingui rang màxim. En particular tenim que si U és un obert de $\mathbb{R}^m$ i

$$H : U \to \mathbb{R}^d$$

és de classe C^1 i $dH(t)$ té rang m per a tot $t \in U$, aleshores la imatge $H(U)$ és una subvarietat regular de dimensió m. Notem en particular el cas $m = 1$: si $\gamma(t), a < t < b$ és de classe C^1 i $\gamma'(t) \neq 0$, aleshores el recorregut Γ és una suvarietat regular de dimensió 1 (corba regular) de classe C^1. Recíprocament, pot veure's que si Γ és una corba regular de classe C^1 (que vol dir que localment és d'aquesta forma) i és arconex, aleshores té una carta global, és a dir, hi ha $\gamma(t)$ definida en algun (a, b), injectiva amb $\gamma'(t) \neq 0$ amb recorregut Γ.

Més generalment hom té el següent resultat, conegut com a *teorema del rang constant*.

Teorema. *Suposem que U és un obert de $\mathbb{R}^m, H : U \to \mathbb{R}^d$ és de classe C^1 i que $dH(t)$ té rang constant n per a tot $t \in U$. Aleshores $H(U)$ és una subvarietat regular de dimensió n.*

Demostració. Fixem un punt $a \in U$; si $H = (h_1, \cdots, h_d)$ hi haurà n, suposem que són les n primeres, entre les h_j tals que els ∇h_j són linealment independents en a i per tant en un entorn V que conté a formen part d'un sistema de coordenades $v_1, \cdots, v_m$; suposen que són les n primeres, és a dir, $h_j = v_j, j = 1, \cdots, n$. Com que el rang és constant igual a n, els gradients $\nabla h_{n+1}, \cdots, \nabla h_d$ depenen linealment dependents dels $\nabla v_j, j = 1, \cdots, n$, i pel darrer corol·lari de la secció 18 en depenen funcionalment, és a dir, $h_j = \phi_j(v_1, \cdots, v_n), j = n + 1, \cdots, d$. En conseqüència, $H(U)$, al voltant de $H(a)$ està parametritzat per les $v_1, \cdots, v_n$ per l'aplicació

$$G(v_1, \cdots, v_n) = (v_1, \cdots, v_n, \phi_{n+1}(v_1, \cdots, v_n), \cdots, \phi_d(v_1, \cdots, v_n)).$$

$\square$

Recapitulant, donada una familia $H = (h_1, \cdots, h_k)$ de funcions de classe C^1 en un domini U de $\mathbb{R}^d$, hem vist

- Si el rang $dH(x)$ és k (llavors $k \leq d$) en tot punt $x \in U$, aleshores les funcions són funcionalment independents en U i $H(U)$ és una varietat de dimensió k de $\mathbb{R}^d$ (un obert si $k = d$).

- Si el rang $dH(x)$ és constant a $n, n < k$, aleshores les funcions són localment funcionalment dependents, i la imatge $H(U)$ és una varietat de dimensió n.

Acabem aquesta secció amb la següent observació. Suposem que Φ és un C^1- difeomorfisme entre dos dominis U, V de $\mathbb{R}^d$ i que $M \subset U$ és una subvarietat regular de dimensió k de U. Aleshores la imatge $\Phi(M)$ també és una subvarietat regular de dimensió m. En efecte, si anomenem x els punts de U i y la variable en V i M esta definida localment per $k = d - m$ equacions $v_1(x) = c_1, \cdots, v_k(x) = c_k$ amb gradients l.i., aleshores $\Phi(M)$ està definida localment per $u_j(y) = v_j(\Phi^{-1}(y) = c_j$, i els gradients $\nabla u_1(y), \cdots, \nabla u_k(y)$ són linealment independents perquè per la regla de la cadena $\nabla v_j(x) = (d\Phi(x))^t(\nabla u_j(y)$ (en ser $d\Phi(x)$ un isomorfisme lineal, vectors l:i. vam a vectors l.i.). Pel mateix motiu, la interpretació geomètrica de la regla de la cadena, tindrem que

$$d\Phi(a)(T_a(M)) = T_{\Phi(a)}(\Phi(M)),$$

és a dir, els espais tangents es corresponen mitjançant la diferencial.

Els difeomorfismes preserven tots els conceptes del càlcul diferencial (punts singulars, punts de sella, subvarietats, extremes condicionats, etc.). Des del punt de vista dels sistemes de coordenades, qualsevol sistema de coordenades $u_1, \cdots, u_d$ té el mateix valor i ús que el cartesià respecte de la base canònica. Per exemple, si en un cert sistema de coordenades M està definit per $u_1 = c_1, \cdots, u_k = c_k$ aleshores M és una subvarietat regular de dimensió $m = d - k$; i també ho és si està donat posant k de les coordenades en termes de les m restants, etc.

17 Teoremes del valor mig

Per a funcions $f : [a, b] \to \mathbb{R}$ contínues i derivables en (a, b) hom té el ben conegut teorema del valor mig

$$f(b) - f(a) = f'(c)(b - a), c \in (a, b).$$

L'anàleg del teorema del valor mig en una situació general seria que per a $f : U \to \mathbb{R}^m$ diferenciable i $a, b \in U$ hom tingués

$$f(b) - f(a) = df(c)(b - a)$$

per un cert punt c.

Això és cert si f és escalar, $m = 1$, i el segment que uneix a, b és dins U:aplicant el teorema del valor mig anterior a $g(t) = f(a+t(b-a)), 0 \leq t \leq 1$, per a la qual $g'(t) = df(a + t(b - a)))(b - a)$, i

$$f(b) - f(a) = g(1) - g(0) = g'(c) = df(z)(b - a),$$

amb $z = a + c(b - a)$ un punt del segment entre a, b.

Per a funcions amb valors vectorials ($m > 1$) una igualtat d'aquest tipus no pot pas ser certa, com ho mostren els dos exemples següents. La funció $g(t) = (\cos t, \sin t), 0 \leq t \leq 2\pi$ val el mateix en 0 i en 2π però no hi ha pas cap punt on $g'(t) = 0$ (qualsevol arc regular tancat podria servir). La funció $f(x, y) = (e^x \cos y, e^x \sin y)$ val el mateix en $(0, 0)$ i en $(0, 2\pi)$ però la diferencial $df(x, y)$ té determinant $e^{2x} \neq 0$ per tant té nucli trivial.

En el cas $m > 1$ hi ha però desigualtats.

Teorema. *Si $f : U \to \mathbb{R}^m$ és diferenciable, $\|df(x)\| \leq M$, i si el segment que uneix a, b és tot ell dins U llavors*

$$\|f(b) - f(a)\| \leq M\|b - a\|.$$

En particular, si U és convex, això val per a tot parell $a, b \in U$, és a dir, f és Lipschitziana en U.

Demostració. Si $h(t) = f((1 - t)a + tb), h'(t) = df((1 - t)a + tb)(q - p)$ té longitud

$$\|h'(t)\| \leq \|df((1 - t)a + tb)\|\|b - a\| \leq M\|b - a\|.$$

Si f és escalar, h també, apliquem el teorema del valor mig a h i ja. Si $m > 1$,

$$\|f(b) - f(a)\| = \|h(1) - h(0)\| = \max_{\|v\|=1} \langle h(1) - h(0), v\rangle.$$

La funció $\langle h(t), v\rangle$ és escalar, amb derivada $\langle h'(t), v\rangle$, per tant per a tot v

$$\langle h(1) - h(0), v\rangle = \langle h'(c_v), v\rangle \leq \|h'(c_v)\| \leq M(b - a)$$

$\square$

Per obtenir una versió vàlida per a dos punts qualsevols $a, b \in U$, considerem com a la secció 2 una poligonal P determinada pels punts $a = p_0, p_1, \cdots, p_N = b$ que uneixi a, b dins U. Aplicant el resultat anterior a cada segment que va de p_i a p_{i+1} tindrem

$$\|f(b) - f(a)\| \leq \sum_{i=0}^{N-1} \|f(p_i) - f(p_{i+1})\| \leq M \sum_{i=0}^{N-1} \|p_i - p_{i+1}\| = ML(P).$$

Si ara prenem l'infim respecte la poligonal obtenim que en les condicions del teorema, amb U arc-connex,

$$\|f(b) - f(a)\| \leq M d_U(a, b).$$

Corol·lari. *Si f és diferenciable en un obert arc-connex i $df(x) = 0, x \in U$, és a dir, $D_i f(x) = 0, i = 1, \cdots, d$, f és constant.*

18 Derivades d'ordre superior

Si f és diferenciable en tots els punts d'una bola B centrada en a, podem considerar l'aplicació diferencial

$$df : B \to L(\mathbb{R}^d, \mathbb{R}^m),$$

on $L(\mathbb{R}^d, \mathbb{R}^m)$ designa l'espai de totes les aplicacions lineals de $\mathbb{R}^d$ en $\mathbb{R}^m$, o matrius $m \times d$. Si l'identifiquem amb $\mathbb{R}^{dm}$, té sentit considerar si aquesta aplicació és diferenciable en a.

Això vol dir que existeixi una aplicació lineal

$$d^2 f(a) : \mathbb{R}^d \to L(\mathbb{R}^d, \mathbb{R}^m)$$

tal que

$$\|df(a + h) - df(a) - d^2 f(a)(h)\| = o(\|h\|).$$

En aquest cas diem que f és *dues vegades diferenciable* en a. Com sempre, això és equivalent a que ho siguin les components, per tant suposarem a partir d'ara que $m = 1$, i estem dient que

$$\nabla f : U \to \mathbb{R}^d,$$

sigui diferenciable. Això és el mateix que dir que totes les components $D_j f$ de ∇f són diferenciables en a. La diferencial de ∇f en a tindrà matriu

$$(D_i D_j f(a)), i, j = 1, \cdots, d.$$

Ens mirem aquesta matriu com definint una forma bilineal

$$d^2 f(a)(u, v) = \sum_{i,j} D_i D_j f(a) u_i v_j,$$

que és el mateix que

$$d^2 f(a)(u, v) = D_u(D_v f)(a)$$

és a dir, es tracta de la derivada direccional de segon ordre, que també designem $D_{uv} f(a)$. Quan són els vectors de la base canònica utilitzem la notació D_{ij} o $\frac{\partial^2 f}{\partial x_j \partial x_i}$

Si prenem qualsevol fórmula, i derivem respecte dues de les variables, constatarem que en fer una derivada segona el resultat no depèn de l'ordre, és a dir, $D_{uv} f = D_{vu} f$. Això és un fet general:

Teorema. *Si f és dues vegades diferenciable en a, llavors $d^2 f(a)$ és una aplicació bilineal simètrica: $D_{uv} f(a) = D_{vu} f(a)$.*

Demostració. És suficient fer-ho per a dos vectors de la base canònica, per tant és suficient considerar el cas $m = 1, d = 2$ on utilitzem la notació $f(x, y)$ i $a = (0, 0)$. Considerem la segona diferència, per a $h > 0$ petit

$$Q(h) = f(h, h) - f(h, 0) - (f(0, h) - f(0, 0)).$$

Si posem $A(t) = f(t, h) - f(t, 0)$, tenim que $Q(h) = A(h) - A(0)$ que pel teorema del valor mig és $A'(\theta h)h = h(D_1 f(\theta h, h) - D_1 f(\theta h, 0))$. Ara bé, com que $D_1 f$ és diferenciable podem escriure

$$D_1 f(x, y) = D_1 f(0, 0) + x D_{11} f(0, 0) + y D_{21} f(0, 0) + o(\sqrt{x^2 + y^2}),$$

per tant

$$D_1 f(\theta h, h) - D_1 f(\theta h, 0)) = h D_{21} f(0, 0) + o(h),$$

i trobem que $Q(h) = h^2 D_{21} f(0, 0) + o(h^2)$ o el que és el mateix

$$\lim_{h \to 0} \frac{Q(h)}{h^2} = D_{21} f(0, 0).$$

Si ara escrivim Q en la forma $Q(h) = f(h, h) - f(0, h) - (f(h, 0) - f(0, 0))$ i raonem de la mateixa forma trobarem que el mateix límit val $D_{12} f(0, 0)$. $\square$

Evidentment la cosa es pot anar iterant. Si f és dues vegades diferenciable en una bola B centrada en a direm que f és tres vegades diferenciable en a si d^2f, que mirem com una funció a valors en l'espai de les aplicacions bilineals de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$ és diferenciable en a. En aquest cas $d^3f(a)$ és una aplicació trilineal simètrica, la que asigna a tres vectors (u, v, w) la derivada de tercer ordre

$$d^3f(a)(u, v, w) = D_u(D^2_{v,w}f)(a),$$

que designem $D_{uvw}f(a)$ o $D_{ikl}f(a)$ per a direccions de la base canònica.

En general, $d^rf(a)$ és una aplicació r-lineal simètrica (invariant per permutacions) que asigna a una r-pla de vectors $u^1, \dots, u^r$ la derivada d'ordre r prenent successives derivades direccionals segons les direccions $u^1, \dots, u^r$.

En coordenades, l'acció de $d^rf(a)$ sobre la r-pla $(u^1, \dots, u^r)$ és

$$\sum_{i_1=1}^{d} \sum_{i_2=1}^{d} \cdots \sum_{i_r=1}^{d} D_{i_1 i_2 \dots i_r} f(a) u^1_{i_1} u^2_{i_2} \cdots u^r_{i_r}.$$

Tota aplicació r-lineal simètrica està determinada per la seva restricció a la diagonal

$$d^rf(a)(u, u, \dots, u) = \sum_{i_1=1}^{d} \sum_{i_2=1}^{d} \cdots \sum_{i_r=1}^{d} D_{i_1 i_2 \dots i_r} f(a) u_{i_1} u_{i_2} \cdots u_{i_r}.$$

Tenint en compte la simetria, d'aquests d^r sumands n'hi ha molts d'iguals i és convenient introduir una notació que no tingui en compte l'ordre.

Un *multiindex* α és $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_d)$, amb $\alpha_i \in \mathbb{N}$. Designem $|\alpha| = \alpha_1 + \alpha_2 + \cdots + \alpha_d$ i designem $\alpha! = \alpha_1!\alpha_2! \dots \alpha_d!$. Designem

$$D^\alpha f(a) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_d^{\alpha_d}} f(a),$$

la derivada d'ordre $|\alpha|$ on derivem α_1 cops respecte x_1, α_2 cops respecte x_2, ..., α_d cops respecte x_d. Per a un vector $x \in \mathbb{R}^d$ i un multiindex α posem $(x-a)^\alpha = (x_1 - a_1)^{\alpha_1}(x_2 - a_2)^{\alpha_2} \cdots (x_d - a_d)^{\alpha_d}$.

En l'expressió anterior de $d^rf(a)$ el nombre de r-ples $(i_1, i_2, \dots, i_r)$ on hi apareixen (sense tenir en compte l'ordre) α_1 uns, α_2 dosos, etc, α_d vegades d és exactament

$$\frac{r!}{\alpha_1!\alpha_2! \dots \alpha_d!}.$$

En efecte, hi ha $\frac{r!}{\alpha_1!(r-\alpha_1)!} = \binom{r}{\alpha_1}$ maneres diferents de seleccionar els α_1 indexos entre els $i_1, i_2, \dots, i_r$ que seran 1. Després hi ha $\binom{r-\alpha_1}{\alpha_2}$ maneres diferents de triar els α_2 indexos entre els $r-\alpha_1$ restants que seran 2, etc. En definitiva podem escriure amb aquestes notacions

$$d^r f(a)(u, u, \dots, u) = \sum_{|\alpha|=r} \frac{r!}{\alpha!} D^\alpha f(a) u^\alpha.$$

Per exemple, en el cas $r = 2$ i $d = 2$, utilitzant (x, y) enlloc de (u_1, u_2) i $a = (0, 0)$ tindriem

$$D_{20} f(0) x^2 + 2 D_{11} f(0) xy + D_{02} f(0) y^2.$$

En el cas $r = 2, d = 3$ amb coordenades (x, y, z) seria

$$D_{200} f(0) x^2 + D_{020} f(0) y^2 + D_{002} f(0) z^2 + 2 D_{110} f(0) xy + 2 D_{101} f(0) xz + 2 D_{011} f(0) yz.$$

El cas $r = 3, d = 2$ seria

$$D_{30} f(0) x^3 + D_{03} f(0) y^3 + 3 D_{21} f(0) x^2 y + 3 D_{12} f(0) x y^2,$$

i el cas $r = 3, d = 3$ és

$$D_{300} f(0) x^3 + D_{030} f(0) y^3 + D_{003} f(0) z^3 + 3 D_{210} f(0) x^2 y + 3 D_{201} f(0) x^2 z +$$

$$+ 3 D_{120} f(0) x y^2 + 3 D_{021} f(0) y^2 z + 3 D_{102} f(0) x z^2 + 3 D_{012} f(0) y z^2 + 6 D_{111} f(0) xyz.$$

Quan f té derivades parcials contínues fins a l'ordre p diem que f és de classe C^p i si les té de tots els ordres diem que f és de classe C^∞ (per exemple les funcions donades per fórmules).

Si $(v_1, \cdots, v_d)$ és un sistema de coordenades i les funcions v_j són de classe C^∞, teniem per a les primeres derivades parcials

$$\frac{\partial f}{\partial v_i} = \sum_{j=1}^{d} \frac{\partial f}{\partial x_j} \frac{\partial x_j}{\partial v_i}.$$

Per tant

$$\frac{\partial^2 f}{\partial v_i \partial v_k} = \sum_{j=1}^{d} \frac{\partial f}{\partial x_j} \frac{\partial^2 x_j}{\partial v_i \partial v_k} + \sum_{j=1}^{d} \frac{\partial x_j}{\partial v_i} \sum_{l=1}^{d} \frac{\partial^2 f}{\partial x_j \partial x_l} \frac{\partial x_l}{\partial v_k}.$$

La darrera suma és
$$d^2f(\frac{\partial}{\partial v_i},\frac{\partial}{\partial v_k}),$$
que són les entrades de la matriu de $d^2f(x)$ en la base $\frac{\partial f}{\partial v_i}$. Per tant, en general, la matriu de $d^2f(x)$ en aquesta base no és $(\frac{\partial^2 f}{\partial v_i \partial v_k})_{ik}$; si que ho és quan el canvi de coordenades és lineal, perquè aleshores $\frac{\partial^2 x_j}{\partial v_i \partial v_k}=0$.

Igual que per a derivades de primer ordre, si d'una funció f coneixem una derivada d'ordre superior, la podem trobar fent antiderivades parcials reiterades. Per exemple, en el pla, si
$$\frac{\partial^2 f}{\partial x \partial y}=0,$$
la funció f_x serà nomes una funció de x, per tant $f(x,y)$ és de la forma
$$f(x,y)=A(x)+B(y).$$
Recíprocament, si f és d'aquesta forma, compleix l'equació $f_{xy}=0$. Si
$$\frac{\partial^2 f}{\partial x^2}=xy,$$
tindrem $f_x=\frac{1}{2}x^2y+A(y)$ i per tant $f(x,y)$ és de la forma $\frac{1}{6}x^3y+xA(y)+B(y)$.

Igual que per a equacions de primer ordre, un canvi de coordenades pot simplificar una equació en derivades parcials. Per exemple, a l'*equació d'ona*
$$\frac{\partial^2 f}{\partial x^2}-\frac{\partial^2 f}{\partial y^2}=0,$$
utilitzarem el canvi $u=x+y, v=x-y$. Llavors $f_x=f_uu_x+f_vv_x=f_u+f_v, f_y=f_u-f_v$ i
$$f_{xx}=f_{uu}u_x+f_{uv}v_x+f_{vu}u_x+f_{vv}v_x=f_{uu}+2f_{uv}+f_{vv},$$
$$f_{yy}=f_{uu}-2f_{uv}+f_{vv}.$$
Aleshores l'equació s'escriu $f_{uv}=0$, d'on resulta que f ha de tenir la forma $A(u)+B(v)$, és a dir,
$$f(x,y)=A(x+y)+B(x-y),$$
amb A,B generals és la solució general de l'equació d'ona.

19 La fórmula de Taylor

La finalitat de la fórmula de Taylor és l'aproximació local per polinomis. Comencem amb l'observació que la notació amb multindexs és també util per a escriure polinomis. Així, un polinomi de grau $\leq N$ s'escriu

$$P(x) = \sum_{|\alpha| \leq N} c_\alpha (x-a)^\alpha.$$

Observem també que

$$c_\alpha = \frac{1}{\alpha!} D^\alpha P(a).$$

Suposem que $f : U \to \mathbb{R}$ és N- vegades diferenciable en $a \in U$ i que $d^r f(a) = 0, r = 0, 1, \cdots, N$, és a dir, la funció i totes les derivades d'ordre $\leq N$ de f són zero en a. Diem que f *s'anul·la a l'ordre* N *en* a. Afirmem llavors que per a tot N

$$|f(x)| = o(\|x-a\|^N).$$

En efecte: quan $N = 1$ això és la mateixa definició de diferenciabilitat, atès que $f(a) = 0, df(a) = 0$. En general ho provem per inducció: suposem que aquest enunciat és cert per a $N-1$ i suposem que f s'anul·la a l'ordre N en a. Llavors df s'anul·la a l'ordre $N-1$. Volem veure que donat $\varepsilon > 0$ hi ha $\delta > 0$ tal que $\|f(x)-f(a)\| \leq \varepsilon \|x-a\|^N$ si $\|x-a\| \leq \delta$. Donat aquest ε, com que df s'anul·la a l'ordre $N-1$ hi ha $\delta > 0$ tal que $\|df(x)\| \leq \varepsilon \|x-a\|^{N-1}$ si $\|x-a\| \leq \delta$. Si $\|x-a\| \leq \delta$, llavors per a tot y en la bola de centre a i radi $\|x-a\|$ es compleix per la hipòtesi de inducció que

$$\|df(y)\| \leq \varepsilon \|y-a\|^{N-1} \leq \varepsilon \|x-a\|^{N-1}.$$

El teorema del valor mig aplicat a la bola de centre a i radi $\|x-a\|$ amb $M = \varepsilon \|x-a\|^{N-1}$ ens diu llavors que

$$\|f(x)\| = \|f(x)-f(a)\| \leq M\|x-a\| = \varepsilon \|x-a\|^N,$$

que és el que voliem veure.

Si dues funcions f, g compleixen que $d^r f(a) = d^r g(a), r = 0, 1 \cdots, N$ diem que *són iguals a l'ordre* N *en* a o que tenen un ordre de contacte N en a. Llavors

$$\|f(x) - g(x)\| = o(\|x-a\|^N).$$

En seguida veurem que el recíproc també és cert.

És obvi que si f és N vegades diferenciable en a hi ha un únic polinomi P de grau més petit o igual que N que és igual a f a l'ordre N en a, el donat pels coeficients $c_\alpha = \frac{1}{\alpha!} D^\alpha f(a)$. El polinomi

$$P(x) = \sum_{|\alpha| \leq N} \frac{1}{\alpha!} D^\alpha f(a)(x-a)^\alpha,$$

s'anomena *el polinomi de Taylor de f en a d'ordre N*. Segons acabem de veure compleix

$$f(x) = P(x) + o(\|x-a\|^N),$$

És l'únic polinomi de grau $\leq N$ complint això: si P, Q ho compleixen, llavors el polinomi $P(x) - Q(x)$ és $o(\|x-a\|^N)$ i obviament ha de ser idènticament zero. Veiem doncs que si f és N vegades diferenciable en a, f s'anul·la a l'ordre N si i només si $f(x) = o(\|x-a\|^N)$.

En correspondència amb les notacions de la secció anterior, notem que una forma alternativa d'escriure P és

$$P(x) = \sum_{r=0}^{N} \frac{1}{r!} d^r f(a)(x-a, \cdots, x-a) =$$

$$\sum_{r=0}^{N} \frac{1}{r!} \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_r=1}^{n} D_{i_1 i_2 \ldots i_r} f(a)(x_{i_1} - a_{i_1})(x_{i_2} - a_{i_2}) \cdots (x_{i_r} - a_{i_r}).$$

Observem que si fix a, x posem $h(t) = f(a + t(x-a)), 0 \leq t \leq 1$, llavors $h'(t) = \sum_{i_1=1}^{n} D_{i_1} f(a + t(x-a))(x_{i_1} - a_{i_1}) = df(a + t(x-a))(x-a)$,

$$h''(t) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} D_{i_1 i_2} f(a + t(x-a))(x_{i_1} - a_{i_1})(x_{i_2} - a_{i_2}) =$$

$$= d^2 f(a + t(x-a))(x-a, x-a),$$

i en general $h^{(r)}(t) = d^r f(a+t(x-a))(x-a, \cdots, x-a)$. Per tant P coincideix amb el polinomi de Taylor de h en zero d'ordre N avaluat en $t = 1$.

La diferència $f(x) - P(x)$, que és $o(\|x-a\|^N)$ s'anomena el *rest o error d'ordre N*. L'observació anterior ens permet trobar una fórmula per a aquest rest, la forma de Lagrange, sota la hipòtesi addicional que existeix $d^{N+1} f(x)$

en un entorn d'a. En efecte, en aquest cas h és $N+1$ vegades derivable i sabem que

$$h(t) = \sum_{r=0}^{N} \frac{1}{r!} h^{(r)}(0) t^r + \frac{1}{(N+1)!} h^{(N+1)}(\theta t) t^{N+1}.$$

Ara bé, el mateix càlcul diu que

$$h^{(N+1)}(\theta t) = d^{N+1} f(a + \theta t(x-a))(x-a, x-a, \cdots, x-a).$$

Concluim així que si f és $N+1$ vegades diferenciable en un entorn d'a tenim

$$f(x) = P(x) + \frac{1}{(N+1)!} d^{N+1} f(\xi)(x-a, \ldots, x-a),$$

on ξ és un punt en el segment que uneix a, x. Una altra forma d'aquest rest, en termes de multiindexs serà

$$\sum_{|\alpha|=N+1} \frac{1}{\alpha_1! \alpha_2! \cdots \alpha_n!} D^\alpha f(\xi)(x-a)^\alpha.$$

20 Extrems lliures

En aquesta secció utilitzarem la fórmula de Taylor per a estudiar el caràcter d'un punt singular, $\nabla f(a) = 0$. En aquest cas tenim

$$f(x) - f(a) = \frac{1}{2} d^2 f(a)(x-a, x-a) + o(\|x-a\|^2).$$

La idea essencial és que el signe de $f(x) - f(a)$ hauria de ser el del primer terme, però això només serà així quan aquest primer terme té signe constant i absorbeixi el rest $o(\|x-a\|^2)$. Per això posem primer unes definicions. Per simplificar les notacions suposarem que $a = 0$.

Considerem doncs un polinomi homogeni de grau dos, també anomenada forma quadràtica

$$Q(x) = \sum_{i,j=1}^{d} a_{i,j} x_i x_j, x \in \mathbb{R}^d,$$

amb $a_{i,j} = a_{j,i}$ (en el nostre cas seria $a_{i,j} = D_i D_j f(0)$). Podem representar-ho en forma matricial

$$Q(X) = X^t A X,$$

on A és la matriu simètrica de les $(a_{i,j})$. Diem que Q (o A) és definida positiva si $Q(x) \geq 0$ i que és estrictament definida positiva si $Q(x) > 0$ per $x \neq 0$; anàlogament tenim els conceptes de definida negativa i definida estrictament negativa.

La primera cosa que observem és que si Q és estrictament definida positiva, la desigualtat $Q(x) > 0, x \neq 0$ s'automillora. En efecte, en l'esfera unitat $S = \{x : \|x\| = 1\}$, que és compacte, $Q(x)$ és una funció contínua, i tindrà per tant un mínim absolut estrictament positiu: $Q(x) \geq m > 0$ si $\|x\| = 1$. Si apliquem això a $x/\|x\|, x \neq 0$, trobem

$$Q(x) \geq m\|x\|^2.$$

Proposició. *Si f és dues vegades diferenciable en a i hi té un mínim (resp. màxim) relatiu, aleshores $d^2f(a)$ és definida positiva (resp. negativa). En el sentit contrari, si $d^2f(a)$ es definida estrictament positiva (estrictament negativa), aleshores f té un mínim (màxim) relatiu en a. De fet, en el primer cas hi ha una bola $B(a,\delta)$ i $c > 0$ tal que*

$$f(x) \geq f(a) + c\|x-a\|^2, \|x-a\| < \delta$$

i

$$f(x) \leq f(a) - c|x-a\|^2, \|x-a\| < \delta$$

en el segon cas.

Demostració. Suposem que $a = 0$ és un mínim relatiu, aleshores si $Q = d^2f(0)$

$$Q(x) + R(x) \geq 0, \|x\| \leq \delta,$$

on $R(x) = o(\|x\|^2)$. Fixat $x \neq 0$ qualsevol aplicant això a tx per t petit, tindem que $t^2Q(x) + o(t^2) \geq 0$ i per tant $Q(x) \geq 0$. Si suposem que $Q(x)$ és estrictament definida, per l'observació anterior tindrem

$$f(x) - f(0) \geq m\|x\|^2 + R(x), \frac{f(x)-f(0)}{\|x\|^2} \geq m + \frac{R(x)}{\|x\|^2}.$$

Com el darrer terme té límit zero en 0, serà més petit que $\frac{m}{2}$ (en valor absolut) per a x prou proper a 0 i llavors l'expressió anterior és $\geq \frac{m}{2}$.

□

Necessitem per tant un criteri algebraic per saber quan una forma quadràtica és definida o estrictament definida. Haviem vist en el capítol 1 que A diagonalitza en una base ortonormal. Això vol dir que en unes certes coordenades cartesianes $y_1, \cdots, y_d$ (on per tant la norma és $\sqrt{\sum_i y_i^2}$) l'expressió de Q és

$$Q(y) = \lambda_1 y_1^2 + \cdots + \lambda_d y_d^2.$$

Llavors és evident que Q és definida positiva si i només si tots els valors propis són positius, i definida estrictament positiva si i només i són tots > 0. Això implica que si Q és definida positiva i $\det A \neq 0$, llavors és definida estrictament positiva.

Però és possible saber això sense necessitat de diagonalitzar, gràcies al teorema de Silvester següent, on per cert es torna a provar que diagonalitza. En l'enunciat, $\Phi(x, y) = \langle Ax, y\rangle$ és la forma bilineal associada.

Teorema. *Una forma bilineal simètrica $\Phi(x, y)$, que té matriu $A = (a_{ij})$ en la base canònica $e_1, \cdots, e_d$, és definida estrictament positiva si i només si els d determinants principals $\Delta_k = \det(a_{ij})_{i,j=1,\dots,k}, k = 1, \cdots, d$ són estrictament positius, i en aquest cas hi ha una base $u_1, \cdots, u_d$ en la que Φ té matriu identitat.*

Demostració. Anomenem H_d aquest enunciat, que provarem per inducció. Evidentement H_1 és cert; suposem cert H_{d-1} i provem H_d.

Suposem primer que Φ és definida estrictament positiva; restringim Φ al subespai E_{d-1} generat per $e_1, \cdots, e_{d-1}$ i apliquem la hipòtesi d'inducció: tindrem que els determinants principals $\Delta_k, k = 1, \cdots, d-1$ són estrictament positius, i hi ha una base $u_1, \cdots, u_{d-1}$ de E_{d-1} tal que $\Phi(u_i, u_j) = \delta_{ij}, i, j = 1, \cdots, d-1$. És clar llavors que $u_1, u_2, \cdots, u_{d-1}, e_d$ és una base en la que Φ té matriu B donada per

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & \Phi(u_1, e_d) \\ 0 & 1 & \cdots & 0 & \Phi(u_2, e_d) \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \cdots & 0 & 1 & \Phi(u_{d-1}, e_d) \\ \Phi(e_d, u_1) & \cdots & \cdots & \Phi(e_d, u_{d-1} & \Phi(e_d, e_d) \end{pmatrix}$$

També obtenim una base canviant e_d per

$$u_d' = e_d - \sum_{j=1}^{d-1} \Phi(e_d, u_j) u_j.$$

Tenim que

$$\Phi(u'_d, u_i) = \Phi(e_d, u_i) - \sum_{j=1}^{d-1} \Phi(e_d, u_j)\Phi(u_j, u_i) = \Phi(e_d, u_i) - \Phi(e_d, u_i) = 0.$$

Per tant en la base $u_1, u_2, \cdots, u'_d$, Φ té matriu diagonal amb valors

$$1, 1, \cdots, 1, \Phi(u'_d, u'_d)$$

a la diagonal. Com que $\Phi(u'_d, u'_d) > 0$, posant $u_d = \frac{u'_d}{\sqrt{\Phi(u'_d, u'_d)}}$ tenim que en la base $u_1, u_2, \cdots, u_d$, Φ té matriu identitat, amb determinant 1. Llavors el determinant d'A ha de ser estrictament positiu perquè si M és la matriu de canvi de base, hom té $A = M^t I M$, i $\det A = (\det M)^2 > 0$.

Si suposem que $\Delta_k > 0, k = 1, 2, \cdots, d$ procedim anàlogament per inducció tot observant que

$$\Phi(u'_d, u'_d) = \Phi(u'_d, e_d) = \Phi(e_d, e_d) - \sum_{j=1}^{d-1} \Phi(e_d, u_j)^2.$$

Ara bé, si a la matriu B li restem a la fila F_n la fila F_j multiplicada per $\Phi(e_n, u_j), j = 1, \cdots, n-1$, veiem que aquesta darrera expressió és el determinant de B, que té el mateix signe que el determinant d'A, per tant $\Phi(u'_n, u'_n) > 0$ i seguim com abans. En la base $u_1, u_2, \cdots, u_n$ Φ té matriu identitat, i per tant és estrictament definida positiva.

□

Evidentment, la forma bilineal serà definida estrictament negativa si i només si els determinants principals van alternant de signe començant amb $\Delta_1 < 0$. Una manera de recordar el criteri és pensar en la matriu ja diagonalitzada, amb els valors propis a la diagonal.

Un altre resultat que es pot utilitzar és el *teorema de Descartes*, segons el qual si un polinomi $P(\lambda)$ té totes les arrels reals, (en el nostre cas el polinomi característic), el nombre d'arrels estrictament positives és igual al nombre de canvis de signe dels seus coeficients. Amb aquest criteri podem saber quants valors propis estrictament positius hi ha, quants d'estrictament negatius, i quina és la multiplicitat del zero.

Quan hi ha dos valors propis diferents de zero de signe contrari, vol dir que hi ha dos vectors unitaris i perpendiculars u, v de forma que

$$Q(tu + sv) = at^2 - bs^2,$$

amb $a, b > 0$ i tenim un *un punt de sella.*

21 Condicions de segon ordre per a extrems condicionats

En aquesta secció reprenem el tema dels multiplicadors de Lagrange, això és, els extrems relatius d'una funció h condicionats per k lligams $f_j = c_j$, sempre sota la hipòtesi que

$$M = \{x \in U : f_1(x) = c_1, \cdots, f_k(x) = c_k\} = \{x \in U : F(x) = c\},$$

sigui una subvarietat regular de dimensió $m = d - k$. Hem vist que aquests extrems relatius han d'estar entre les solucions x del sistema de $d + k$ equacions escalars

$$\nabla h(x) = \lambda_1 \nabla f_1(x) + \cdots + \lambda_k \nabla f_k(x), f_1(x) = c_1, \cdots, f_k(x) = c_k,$$

on les incògnites són les d coordenades dels punts candidats x, i les $\lambda_1, \cdots, \lambda_k$ són els multiplicadors de Lagrange.

Ens preguntem si igual que per a extrems lliures és possible formular alguna condició suficient en termes de derivades d'ordre dos. Comencem primer observant que si $p \in M$ és un d'aquests punts que compleix la condició necessària, podem reformular el problema en termes d'una carta local $H(t_1, t_2, \cdots, t_m), H(0) = p$, amb $dH(0)$ de rang $m = d - k$. Evidentment, el fet que h tingui un extrem relatiu condicionat en p és el mateix que dir que $g(t) = h(H(t))$ tingui un extrem lliure en $t = 0$. Que $\nabla g(0) = 0$ és exactament el sistema de Lagrange; podriem llavors considerar el hessià de g en zero i aplicar-hi els criteris per a extrems lliures. El que veurem en aquesta secció és que aquests criteris poden formular-se novament en termes de la funció de Lagrange

$$L(x) = h(x) - \sum_j \lambda_j f_j(x),$$

sense necessitat de cap carta local.

Hom podria pensar, per analogia amb el cas dels extrems lliures, que una condició suficient seria que el hessià de h sigui definit estrictament positiu o negatiu sobre l'espai tangent $T_p(M)$. Ara bé, si això fos així, una funció estrictament convexa arreu, amb hessià definit estrictament positiu en tots els punts (per exemple x^2+2y^2), només tindria mínims relatius condicionats, cosa evidentment absurda. Això significa que el criteri que busquem no pot dependre tan sols del hessià de h (segones derivades de h) i de $T_p(M)$ (és a dir dels $\nabla f_j(p)$), sinó que les segones derivades dels lligams també han d'intervenir. Geomètricament això és veu si considerem els extrems condicionats de l'exemple anterior x^2+2y^2 sobre $x^2+y^2=1$. Els màxims, que ho són absoluts, són $(0,1),(0,-1)$ i els minims són $(1,0),(-1,0)$. En els punts $(0,1),(0,-1)$ la corba de nivell $x^2+y^2=1$ és dins de la corba de nivell $x^2+2y^2=2$, mentre que en els punts $(1,0),(-1,0)$ és al revés. La posició relativa d'aquestes corbes depen de les segones derivades i no solament de les primeres.

Per veure quina és la condició suficient correcta prenem la carta local $H(t)$, considerem $g(t)=h(H(t))$, i calculem una segona derivada $\frac{\partial^2 g}{\partial t_i \partial t_j}(0)$. Primer, si designem per $x_l=x_l(t)$ les components de H

$$\frac{\partial g}{\partial t_j}=\sum_{l=1}^{d}\frac{\partial h}{\partial x_l}(H(t))\frac{\partial x_l}{\partial t_j},$$

i tornant a derivar

$$\frac{\partial^2 g}{\partial t_i \partial t_j}=\sum_{l,n=1}^{d}\frac{\partial^2 h}{\partial x_l \partial x_n}\frac{\partial x_n}{\partial t_i}\frac{\partial x_l}{\partial t_j}+$$

$$+\sum_{l=1}^{d}\frac{\partial h}{\partial x_l}(H(t))\frac{\partial^2 x_l}{\partial t_j \partial t_i}.$$

Ara bé, tenim que $f_j(H(t))=c_j$ per tant la mateixa expressió anterior però amb f_j enlloc de h és zero; d'altra banda, avaluant en $t=0$, sabem que $\nabla h(p)=\sum_n \lambda_n \nabla f_n(p)$. Per tant la darrera suma és

$$\sum_{l=1}^{d}\sum_{n=1}^{d}\lambda_n\frac{\partial f_n}{\partial x_l}(p)\frac{\partial^2 x_l}{\partial t_j \partial t_i}(0)=$$

$$-\sum_{l,n=1}^{d}\lambda_n\frac{\partial^2 f_n}{\partial x_l \partial x_n}(p)\frac{\partial x_n}{\partial t_i}(0)\frac{\partial x_l}{\partial t_j}(0),$$

i finalment

$$\frac{\partial^2 g}{\partial t_i \partial t_j}(0) = d^2L(p)(\frac{\partial}{\partial t_i}, \frac{\partial}{\partial t_j})$$

el hessià de L en p actuant sobre la base $\frac{\partial}{\partial t_i}$ de $T_p(M)$.

Per tant és el hessià de la funció de Lagrange L i no el de h el que cal considerar: si hi ha un mínim (resp. màxim) relatiu condicionat, aquest hessià ha de ser definit positiu (resp.negatiu) en $T_p(M)$ i en sentit contrari, si aquest hessià és estrictament definit positiu en $T_p(M)$ (resp. estrictament definit negatiu) aleshores hi ha un màxim (resp. mínim) condicionat.

El que ens falta ara és un criteri algebraic per saber quan una forma quadràtica $Q(x) = X^tAX$ en $\mathbb{R}^d$ és estrictament definida sobre un subespai definit per k equacions linealment independents $BX = 0$ amb B una matriu $k \times d$ de rang k. Fem primer el cas més senzill possible $d = 2, k = 1$: volem per exemple per a un mínim relatiu que

$$Ax^2 + 2Bxy + Cy^2 > 0, \text{si}\, Ex + Fy = 0, x, y \neq 0$$

Posant $x = -\frac{F}{E}y$ tenim

$$(\frac{AF^2}{E^2} - \frac{2BF}{E} + C)y^2 > 0,$$

per tant cal que $AF^2 - 2BEF + CE^2 > 0$, això és el determinant de la matriu

$$\begin{pmatrix} 0 & E & F \\ E & A & B \\ F & B & C \end{pmatrix},$$

ha de ser negatiu. Això mateix és possible en el cas general, que convé tractar amb notació matricial. En primer lloc, suposem que el menor d'ordre k de B no nul és el B' de les primeres k columnes i posem $x' = (x_1, \cdots, x_k), x'' = (x_{k+1}, \cdots, x_d)$. El lligam $BX = B'X' + B''X'' = 0$ significa

$$X' = -(B')^{-1}B''X'' = CX''$$

(és a dir, estem parametritzant el subespai amb les X''). Considerem la mateixa partició en blocs de la matriu A

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{12}^t & A_{22} \end{pmatrix},$$

on A_{11} és $k\times k$ etc. Hi associem el *hessià orlat* que és la matriu $(d+k)\times(d+k)$ definida per

$$A^* = \begin{pmatrix} 0 & B' & B'' \\ (B')^t & A_{11} & A_{12} \\ (B'')^t & A_{12}^t & A_{22} \end{pmatrix}.$$

La matriu de la restricció de Q al subespai és

$$D = \begin{pmatrix} C^t & I \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^t & A_{22} \end{pmatrix} \begin{pmatrix} C \\ I \end{pmatrix} = C^t A_{11} C + C^t A_{12} + A_{12}^t C + A_{22}.$$

D'altra banda si apliquem un canvi de coordenades a A^*

$$\begin{pmatrix} I_k & 0 & 0 \\ 0 & I_k & 0 \\ 0 & C^t & I_{d-k} \end{pmatrix} \begin{pmatrix} 0 & B' & B'' \\ (B')^t A_{11} & A_{12} & \\ (B'')^t & A_{12}^t & A_{22} \end{pmatrix} \begin{pmatrix} I_k & 0 & 0 \\ 0 & I_k & C \\ 0 & 0 & I_{d-k} \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & B' & 0 \\ (B')^t & A_{11} & C_{12} \\ 0 & C_{12}^t & E \end{pmatrix}$$

per a una certa matriu C_{12}. Designem per $E_l, 0 \le l \le d-k-1$ els menors d'E que s'obtenen eliminant les darreres l files i columnes, el signe dels quals és el que ens interessa; anàlogament A_l^* significa el menor de A^* que s'obté eliminant les darreres l files i columnes. La relació anterior implica que

$$(-1)^k \det A_l^* = (\det B')^2 \det E_l.$$

Per tant la restricció de Q al subespai és estrictament definida positiva si i només si

$$(-1)^k \det A_l^* > 0, l = 0, \cdots, d-k-1,$$

i és definida estrictament negativa si aquestes mateixes quantitats van alternant de signe, el signe + corresponent a $l = d-k-1$, és a dir, $\det A_l^*$ té signe $(-1)^{d+l+1}$.

22 Desenvolupament de Taylor de la funció implícita

Naturalment, no sempre serà possible d'una manera explícita escriure una fórmula per a la funció implícita. Si podem però calcular explícitament la diferencial $dh''(a'')$, és a dir, les derivades parcials de la funció implícita en a''. Expliquem això en el cas més senzill possible, d'una sola equació $f(x, y) = 0$ al pla. Si $\frac{\partial f}{\partial y}(a, b) \neq 0$ el teorema diu que hi ha una única funció $y = h(x)$ que compleix $h(a) = b$ i

$$f(x, h(x)) = 0,$$

idènticament. Aleshores, si derivem respecte de x trobem

$$\frac{\partial f}{\partial x}(x, h(x)) + \frac{\partial f}{\partial y}(x, h(x))h'(x) = 0,$$

i avaluant en a,

$$\frac{\partial f}{\partial x}(a, b) + \frac{\partial f}{\partial y}(a, b)h'(a) = 0,$$

d'on

$$h'(a) = -\frac{\frac{\partial f}{\partial x}(a, b)}{\frac{\partial f}{\partial y}(a, b)}.$$

Si tenim una equació en tres variables $f(x, y, z) = 0$ i tenim $f(a, b, c) = 0$, $\frac{\partial f}{\partial z}(a, b, c) \neq 0$ hi ha una única funció $z = h(x, y)$ que compleix $h(a, b) = c$ i

$$f(x, y, h(x, y)) = 0.$$

Prenent derivades parcials respecte x, y obtenim

$$\frac{\partial f}{\partial x}(x, y, h(x, y)) + \frac{\partial f}{\partial z}(x, y, h(x, y))\frac{\partial h}{\partial x}(x, y) = 0,$$
$$\frac{\partial f}{\partial y}(x, y, h(x, y)) + \frac{\partial f}{\partial z}(x, y, h(x, y))\frac{\partial h}{\partial y}(x, y) = 0,$$

i avaluant en (a, b),

$$\frac{\partial h}{\partial x}(a, b) = -\frac{\frac{\partial f}{\partial x}(a, b, c)}{\frac{\partial f}{\partial z}(a, b, c)}, \frac{\partial h}{\partial y}(a, b) = -\frac{\frac{\partial f}{\partial y}(a, b, c)}{\frac{\partial f}{\partial z}(a, b, c)}.$$

Si tenim dues equacions $f(x,y,z)=0, g(x,y,z)=0, f(a,b,c)=g(a,b,c)=0$ i el menor

$$\begin{pmatrix} \frac{\partial f}{\partial y}(a,b,c) & \frac{\partial f}{\partial z}(a,b,c) \\ \frac{\partial g}{\partial y}(a,b,c) & \frac{\partial f}{\partial z}(a,b,c) \end{pmatrix},$$

té determinant diferent de zero, aleshores hi ha dues úniques funcions $y=h(x), z=k(x)$ tals que $h(a)=b, k(a)=c$ que compleixen

$$f(x,h(x),k(x))=0, g(x,h(x),k(x))=0.$$

Si derivem respecte de x

$$\frac{\partial f}{\partial x}(x,h(x),k(x))+\frac{\partial f}{\partial y}(x,h(x),k(x))h'(x)+\frac{\partial f}{\partial z}(x,h(x),k(x))k'(x)=0$$

$$\frac{\partial g}{\partial x}(x,h(x),k(x))+\frac{\partial g}{\partial y}(x,h(x),k(x))h'(x)+\frac{\partial g}{\partial z}(x,h(x),k(x))k'(x)=0$$

i avaluem en $x=a$ arribem al sistema lineal

$$\frac{\partial f}{\partial x}(a,b,c)+\frac{\partial f}{\partial y}(a,b,c)h'(a)+\frac{\partial f}{\partial z}(a,b,c)k'(a)=0$$

$$\frac{\partial g}{\partial x}(a,b,c)+\frac{\partial g}{\partial y}(a,b,c)h'(a)+\frac{\partial g}{\partial z}(a,b,c)k'(a)=0,$$

en les incògnites $h'(a), k'(a)$. Com que la matriu d'aquest sistema és precisament el menor anterior, podem resoldre explícitament i trobar $h'(a), k'(a)$.

En general, si designem $h=(h_1,\cdots,h_k)$, i si derivem respecte una de les variables x'', per exemple la darrera x_d, les k equacions

$$v_j(h(x''),x'')=0, j=1,\cdots,k,$$

i avaluem en $x''=a''$, trobarem un sistema lineal de k equacions en les k incògnites $\frac{\partial h_j}{\partial x_d}(a''), j=1,\cdots,k$ que té per matriu el menor no nul.

Aquestes fórmules valen naturalment no solament en el punt a sinó en tots els punts (x',x'') d'un entorn, i per tant veiem que si f és de classe C^p aleshores també ho és h.

Les derivades successives de la funció implícita h poden obtenir-se pel mateix procediment que per a les primeres. Per exemple, per a una sola equació $f(x,y)=0$ amb funció implícita $y=h(x)$ derivant altre cop

$$\frac{\partial f}{\partial x}(x,h(x))+\frac{\partial f}{\partial y}(x,h(x))h'(x)=0,$$

i avaluant en $x = a$ obtenim

$$\frac{\partial^2 f}{\partial x^2}(a,b) + 2\frac{\partial^2 f}{\partial x \partial y}(a,b)h'(a) + \frac{\partial^2 f}{\partial y^2}(h'(a))^2 + \frac{\partial f}{\partial y}(a,b)h''(a) = 0,$$

i, com que $h'(a)$ ja ha estat prèviament calculat, podem obtenir $h''(a)$.

Per a una equació $f(x,y,z) = 0$, amb les notacions anteriors, tornant a derivar

$$\frac{\partial f}{\partial x}(x,y,h(x,y)) + \frac{\partial f}{\partial z}(x,y,h(x,y))\frac{\partial h}{\partial x}(x,y) = 0,$$

respecte de x i avaluant en (a,b) obtindrem

$$E + \frac{\partial f}{\partial z}(a,b,c)\frac{\partial^2 h}{\partial x^2}(a,b) = 0,$$

on E és una expressió en termes de derivades parcials de segon ordre de f en (a,b,c) i de la primera derivada $\frac{\partial h}{\partial x}(a,b)$, per tant ja coneguda, i podrem calcular $\frac{\partial^2 h}{\partial x^2}(a,b) = 0$. Anàlogament procediriem amb les altres dues derivades de segon ordre de h en (a,b).

Finalment en la tercera situació considerada, $f(x,y,z) = 0, g(x,y,z) = 0$ amb funcions implícites $y = h(x), z = k(x)$, tornant a derivar les dues expressions anteriors i avaluant en $x = a$ trobem

$$E_1 + \frac{\partial f}{\partial y}(a,b,c)h''(a) + \frac{\partial f}{\partial z}(a,b,c)k''(a) = 0$$
$$E_2 + \frac{\partial g}{\partial y}(a,b,c)h''(a) + \frac{\partial g}{\partial z}(a,b,c)k''(a) = 0,$$

d'on trobem resolent el sistema les segones derivades $h''(a), k''(a)$.

D'aquesta forma, si les equacions són de classe C^∞ es poden obtenir desenvolupaments de Taylor arbitràriament llargs de les funcions implícites.