

CONSTRUCCIÓ DE SUPERFÍCIES MITJANÇANT FUNCIONS SPLINE

P. BRUNET I CROSA

La construcció d'una superfície que interpoli un conjunt de punts donats amb uns certs requeriments de suavitat i continuïtat, és un problema de gran importància tant en aplicacions de disseny gràfic com en el procés de dades experimentals. En una primera part del present article s'exposen els principals mètodes utilitzats actualment, comparant llurs prestacions i definint llurs camps d'aplicació. A la segona part, es presenta un algorisme que permet d'interpol·lar punts irregularment distribuïts i dona lloc a una superfície amb continuïtat fins a les derivades parcials segones. Es discuteixen amb detall llurs propietats i es compara amb d'altres mètodes ja existents.

1. INTRODUCCIÓ

La necessitat de construir una superfície a partir del coneixement d'alguns punts de pas pot sorgir tant en aplicacions de disseny gràfic assistit per ordinador com a l'hora de conèixer el comportament d'un cert fenomen experimental en punts no tabulats. En el primer cas, probablement caldrà disposar d'un algorisme capaç d'anar modificant la posició dels punts coneguts fins arribar a una superfície amb la forma desitjada; en el segon cas obtindrem la solució en una sola iteració. Però ambdós problemes es redueixen a l'obtenció de la superfície d'interpolació la qual cosa no és trivial. Ja en el cas d'una dimensió -interpolació de corbes- són ben coneguts els problemes inherents a la interpolació polinòmica de Lagrange. En dues dimensions, ens trobem amb problemes de baixa continuïtat al llarg de l'interpolant -cas dels mètodes locals- o de temps de càlcul elevat i possibilitat d'aplicació només en configuracions molt determinades de les dades -mètodes globals-.

En el present article, un cop presentats els mètodes més utilitzats actualment per a la construcció d'interpolants superficials, es proposa un mètode que reuneix d'una banda els avantatges dels esquemes globals (continuïtat elevada, amb el que això comporta de suavitat de la superfície resultant) i d'altra banda les dels esquemes locals (possibilitat d'interpol·lar punts irregularment dis-

tribuïts). La comparació dels resultats i temps d'execució amb els d'altres algorismes molt utilitzats actualment confirma la seva adequació a un gran nombre d'aplicacions.

2. INTERPOLACIÓ UNIDIMENSIONAL

Abans d'estudiar els algorismes existents per a la interpolació superficial començarem per a la interpolació superficial començarem per a la presentació dels esquemes de construcció de corbes. Parlarem només d'interpolacions polinòmiques a trossos, donat que la interpolació clàssica de Lagrange no té assegurada la convergència en augmentar el nombre de punts.

2.1 Corbes de Bèzier i B-Splines

Donat k i un conjunt de nodes $\tau_1, \dots, \tau_{n+k}$ tal que $\tau_i \leq \tau_j$ si $i < j$, es pot definir la base de B-Splines d'ordre k , $N_{1,k}(x), \dots, N_{n,k}(x)$ a partir de la següent relació de recurrència /1/

$$\begin{aligned} N_{i,1}(x) &= 1 \quad \text{si } \tau_i \leq x \leq \tau_{i+1} \\ N_{i,1}(x) &= 0 \quad \text{si } x \notin [\tau_i, \tau_{i+1}] \\ N_{i,k}(x) &= \frac{(x - \tau_i) N_{i,k-1}(x)}{\tau_{i+k-1} - \tau_i} + \frac{(\tau_{i+k} - x) N_{i+1,k-1}(x)}{\tau_{i+k} - \tau_{i+1}} \end{aligned} \quad (1)$$

El conjunt de funcions $N_{1,K}(x), \dots, N_{n,K}(x)$ d'ordre K (formades per trams polinòmics de

- Pere Brunet i Crosa. Catedra Metodes Informatics, Escola Tecnica Superior d Enginyers Industrials de la Universitat Politecnica de Barcelona. Diagonal, 647. Barcelona-28.

- Article rebut al març de 1981.

grau $K-1$ i tals que $N_{i,K}(x) \geq 0 \quad \forall x, i, K, \dots$
 $N_{i,K}(x) = 0 \quad \forall x \notin [\tau_i, \tau_{i+K}]$ formen base de
l'espai de dimensió n de funcions polinòmiques
a trossos definides sobre τ_1, \dots, τ_n
Donada la malla definida per les abscises crei-
xents $\xi_1 < \xi_2 < \xi_3 \dots < \xi_{1+1}$, podem trobar la base
de B-Splines, /2/, construint el conjunt
de nodes de la següent manera:

- fem $\tau_1 = \tau_2 = \dots = \tau_K = \xi_1$
- fem $\tau_{n+1} = \dots = \tau_{n+K} = \xi_{1+1}$
- per tot $i, 1 \leq i \leq 1$ cal que existeixi un r tal
que $\tau_r = \dots = \tau_{r+s_i} = \xi_i$ essent $s_i = K-1$ -
(nombre de condicions de continuïtat desit-
jades al nus ξ_i).

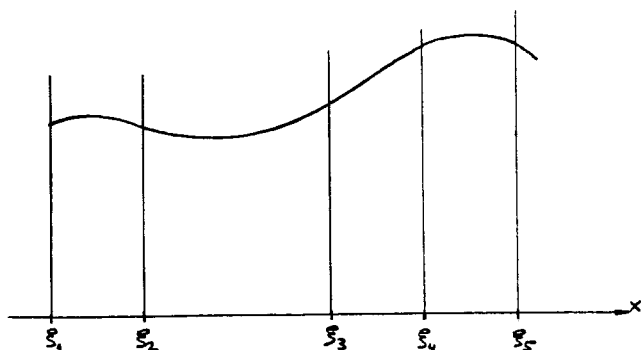
És a dir, els nodes definidors dels B-Splines
coincideixen amb els punts de la malla,
i la seva multiplicitat ens indica el grau
de continuïtat a cada un dels seus punts; a
partir del conjunt de nodes τ_j ja és im-
mediat, amb (1), calcular la base. Per exemple
amb $k = 3$, la base calculada a partir dels no-
des $0,0,0,1,3,4,5,5,5$ permet generar totes
les funcions construïdes per trams d'ordre 3
-parabòlics- als intervals de la malla d'ab-
scises $0,1,3,4,5$ connectats entre ells amb con-
tinuïtat fins la primera derivada, figura 1-a;
en canvi, a partir dels nodes $0,0,0,1,1,3,3,$
 $3,4,4,5,5,5$ trobem una base que permet gene-
rar funcions discontinües al punt d'abscisa
 $x = 3$ i amb continuïtat només de la funció -
als altres punts, com veiem a la figura 1-b.
Totes aquestes funcions es podran obtenir --
evidentment com combinació lineal dels B-Splines:

$$f(x) = \sum_{i=1}^n \alpha_i N_{i,k}(x) \quad (2)$$

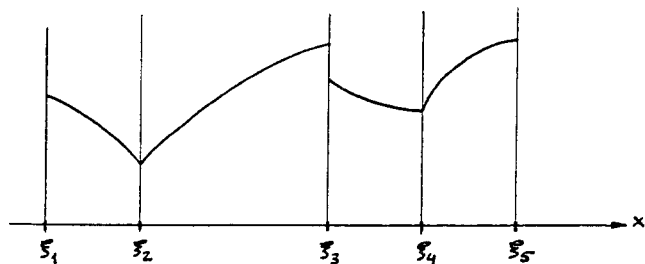
Es pot demostrar, /2/, que si considerem els
punts del pla,

$$(\alpha_1, t_1), \dots, (\alpha_n, t_n)$$

$$t_i = \frac{1}{k-1} [\tau_{i+1} + \dots + \tau_{i+k-1}] \quad (3)$$



a- Base amb nodes $0,0,0,1,3,4,5,5,5$



b- Base amb nodes $0,0,0,1,1,3,3,3,4,4,5,5,5$

Figura 1. Exemples de funcions polinòmiques a trossos generades a partir de dues bases de B-Splines

la forma de la funció $f(x)$ polinòmica a tros-
sos definida per (2) ve modelada per la poli-
gonal que uneix els punts (α_i, t_i) , com es --
pot veure a la figura 2 on s'ha representat
un cas amb $n=5$. Per tant, un cop s'ha calcu-
lat la base de B-Splines és molt senzill el
disseny de corbes, a les quals podem anar --
canviant la forma variant la poligonal. A --
més, la propietat ja anunciada que $N_{i,k}(x) =$
 $= 0 \quad \forall x \notin [\tau_i, \tau_{i+k}]$ implica que el procés --
de construcció que estem esmentant és local
en el sentit que la modificació d'un punt de
la poligonal només afecta la corba resultant
en la zona propera.

El cas particular de les funcions $f(x)$ gene-
rades segons (2) quan els B-Splines es basen
en una malla de dos punts ξ_1, ξ_2 , s'anomenen

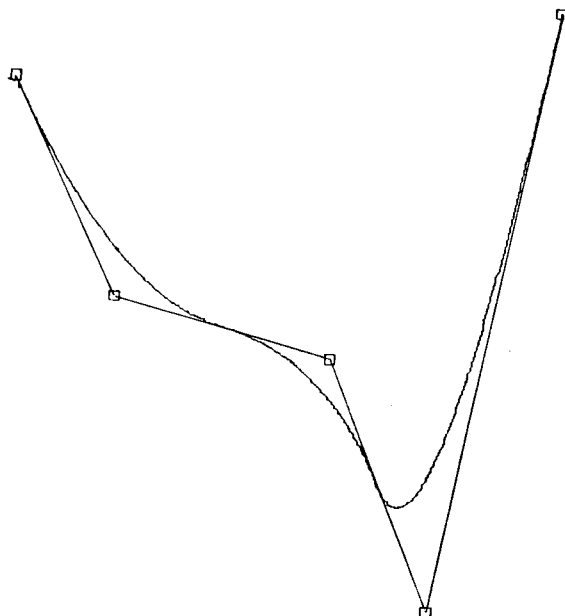


Figura 2: Poligonal construïda amb els components α_i d'una funció en la base de B-splines.

corbes de Bèzier, /3/. Veiem els nodes que serveixen per a la construcció d'aquests B-Splines:

K	nodes
3	0,0,0,1,1,1
4	0,0,0,0,1,1,1,1
5	0,0,0,0,0,1,1,1,1,1

(K és l'ordre del polinomi, o grau +1)

Les corbes de Bèzier i els B-Splines han estat dues tècniques molt utilitzades per al disseny interactiu de corbes mitjançant ordinador, degut a llur caràcter local i al fet que s'adapten a la poligonal. Ara bé, són mètodes no vàlids per a la interpolació -almenys en la forma que habitualment es presenten- ja que la corba obtinguda no passa pels punts donats.

Finalment direm que tot el dit fins ara, així com també els mètodes que anirem veient a continuació, és clar que també s'apliquen a la construcció de corbes a l'espai definides en paramètriques: $x(t)$, $y(t)$, $z(t)$. En aquest cas hauriem de substituir l'equació (2) per la seva equivalent vectorial,

$$\begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = \sum_{i=1}^n \begin{pmatrix} \alpha_x \\ \alpha_y \\ \alpha_z \end{pmatrix}_i N_{i,k}(t) \quad (4)$$

amb una poligonal definida pels punts

$$\begin{pmatrix} \alpha_x \\ \alpha_y \\ \alpha_z \end{pmatrix}_1, \dots, \begin{pmatrix} \alpha_x \\ \alpha_y \\ \alpha_z \end{pmatrix}_n$$

2.2 Funcions Spline Locals

Passem ara als algorismes de disseny de corbes que es basen en la interpolació a partir d'un conjunt donat de punts, i en els que, per tant, la corba final passa per tots ells. La gran majoria es basen en les funcions Spline cúbiques definides sobre el conjunt donat de punts $(\xi_1, y_1), \dots, (\xi_{l+1}, y_{l+1})$:

$$f(x) = P_i(x), \quad \xi_i \leq x \leq \xi_{i+1} \quad (5)$$

essent $P_1(x) \dots P_l(x)$ polinomis cúbics

$$P(\xi_i) = y_i, \quad P_i(\xi_{i+1}) = y_{i+1}$$

Les funcions spline local venen definides per les equacions (5), i a més, per:

$$P'_i(\xi_{i+1}) = P'_{i+1}(\xi_{i+1}) \quad i=1 \dots l-1 \quad (6)$$

amb la qual cosa queda assegurada la continuïtat fins la primera derivada per $x \in [\xi_1, \xi_{l+1}]$. Ara bé, les equacions (5) i (6) deixen $l+1$ graus de llibertat a l'hora de determinar els $4l$ coeficients dels polinomis P_1, \dots, P_l integrants de l'spline $f(x)$. Un dels mètodes més extesos, /4/, per tal d'eliminar els esmentats graus de llibertat, calcula un valor aproximat de les $l+1$ derivades primeres $P'_1(\xi_1), \dots, P'_l(\xi_1), P'_l(\xi_{l+1})$ fent, per al punt i , una ponderació de les derivades $d_{i,1}, d_{i,2}$ dels polinomis cúbics que passen pels punts $\xi_{i-2}, \xi_{i-1}, \xi_i, \xi_{i+1}$ i $\xi_{i-1}, \xi_i, \xi_{i+1}, \xi_{i+2}$ respectivament. El resultat és un interpolant local en el mateix sentit que els B-Splines: la modificació de l'ordenada d'un punt no produeix un canvi en la forma de tota la corba final, sinó només en la zona més propera.

2.3 Funcions Spline Cardinals

Es defineixen com splines cúbics naturals els que compleixen, a més de les condicions expressades per (5) i (6),

$$P''_i(\xi_{i+1}) = P''_{i+1}(\xi_{i+1}) \quad i=1 \dots l-1 \quad (7)$$

$$P''_1(\xi_1) = P''_l(\xi_{l+1}) = 0 \quad (8)$$

En aquest cas, hem determinat $l-1$ dels $l+1$ graus de llibertat de les funcions spline locals al imposar continuïtat fins la segona derivada, la qual cosa augmenta considerablement la suavitat de la corba. Per tal de determinar els dos graus de llibertat restants una alternativa a les equacions (8) és considerar el subespai de funcions spline tals que $P'''_1(\xi_1)$ i $P'''_l(\xi_{l+1})$ són les corresponents als polinomis d'interpolació que passen pels punts d'abscises $\xi_1, \xi_2, \xi_3, \xi_4$ i $\xi_{l-2}, \xi_{l-1}, \xi_l, \xi_{l+1}$ respectivament. D'ara en davant parlarem sempre de funcions spline que satisfan aquesta última restricció, en haver-se demostrat que dóna resultats molt més acceptables en aplicacions de disseny.

En aquests moments, podem definir els splines cardinals $C_1(x), \dots, C_{l+1}(x)$ com les funcions, /5/, que compleixen (5), (6), (7), la restricció de derivades terceres, i:

$$C_K(\xi_j) = \delta_{Kj} \quad \forall_{K,j} \quad (9)$$

Aquestes funcions formen base del subespai de funcions spline que estem considerant: donat un conjunt de punts $(\xi_1, y_1), \dots, (\xi_{l+1}, y_{l+1})$ és immediat comprovar que l'spline cúbic amb continuïtat C^2 que passa per tots ells és,

$$S_{\Delta}(x) = \sum_{i=1}^{l+1} y_i C_i(x) \quad (10)$$

on $\Delta = \{\xi_1, \dots, \xi_{l+1}\}$ és la malla

A la figura 3 es tenen les components de la base per a $l = 5$

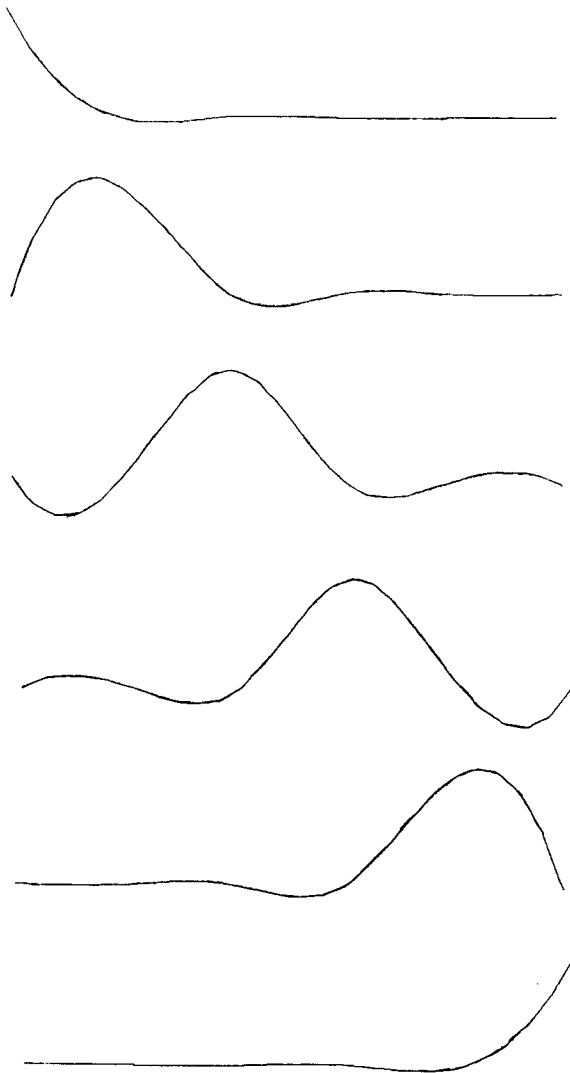


Figura 3. Els elements de la base de splines cardinals per a 6 punts igualment distanciat.

Cal observar que:

- Els splines cardinals C_1, \dots, C_{l+1} queden totalment determinats a partir de la malla

Δ . Si es modifiquen les ordenades no cal canviar la base; calculant de nou el sumatori (10) es té el nou interpolant $S_{\Delta}(x)$.

- L'interpolant obtingut passa pels punts donats, a diferència dels esquemes basats en B-Splines.
- L'interpolant $S_{\Delta}(x)$ pertany a C^2 i per tant és més suau que els obtinguts amb criteris locals.
- El fet que les funcions $C_i(x)$ són no nul·les a tot l'interval $[\xi_1, \xi_{l+1}]$ -vegeu la figura 3- implica que aquest és un mètode global: el canvi d'una sola ordenada y_i modifica la forma de tota la corba.

Com a resum de tot el que hem dit respecte a la interpolació unidimensional, veiem que hi ha dos criteris que cal analitzar a l'hora d'escollir un mètode de disseny de corbes:

- Cal optar per un mètode local, on es té poca continuïtat però és possible de modificar punts sense que canviï la forma de tota la corba, o per un mètode global, que dona interpolants més suaus però tota modificació repercuteix -tot i que atenuada- al llarg de tota la corba.
- Cal optar per un esquema d'interpolació (splines local o splines cardinals) o bé per un d'aproximació a una poligonal coneguda (B-Splines, Bèzier).

En tot cas, l'opció concreta dependrà de l'aplicació específica que es vulgui realitzar.

3. DISSENY DE SUPERFÍCIES A PARTIR D'UNA MALLA RECTANGULAR

El cas més senzill de construcció de superfícies se'ns presenta quan les dades són els valors z_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m$, en els $n \times m$ punts de la malla rectangular $\Delta_x \cdot \Delta_y$ definida per les particions unidimensionals

$$\begin{aligned} \Delta_x &= \{x_1, x_2, \dots, x_n\} \\ \Delta_y &= \{y_1, y_2, \dots, y_m\} \end{aligned} \quad (11)$$

Sense considerar els esquemes provinents del treball amb funcions spline locals, parlarem a continuació de B-Splines i splines cardinals.

3.1 Superfícies construïdes amb B-Splines

En aquest cas, a partir de la informació (11) i de les condicions de continuïtat als punts de la malla, es construeixen les dues bases unidimensionals de B-Splines $N_{1,K,X}(x), \dots, N_{n,K,X}(x)$ i $N_{1,K,Y}(y), \dots, N_{m,K,Y}(y)$ (cal entendre els punts x_1, \dots, x_n i els y_1, \dots, y_m com a una generalització dels punts t_i de l'equació (3) que ens definien la poligonal inicial; ara tindrem una "poligonal" a R^2 , unió dels punts $(x_i, y_i, z_{ij}) \forall i, j$).

En aquest moment es pot definir (/6/) la superfície construïda amb B-Splines com,

$$f(x,y) = \sum_{i=1}^n \sum_{j=1}^m z_{ij} N_{i,K,X}(x) N_{j,K,Y}(y) \quad (12)$$

La superfície resultant $f(x,y)$ és un tros bicúbic a cada rectangle definit per $x_i \leq x \leq x_{i+1}$, $y_j \leq y \leq y_{j+1}$. Les condicions de continuïtat -en les derivades parcials- entre diferents trossos venen determinades per la construcció de les bases de B-Splines. Com en el cas unidimensional, la superfície no passa pels punts donats, tot i que la seva forma ve modelada per la de la poligonal (figura 4).

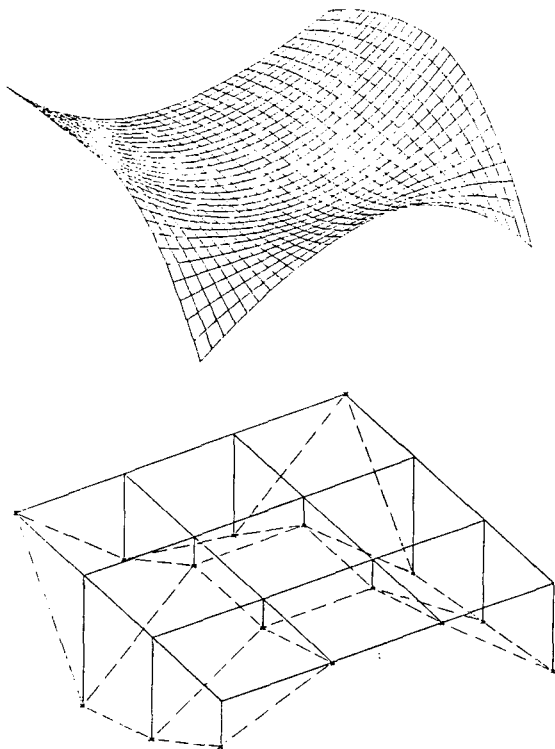


Figura 4: Superfície construïda amb B-splines i la poligonal que ha servit per a la seva generació.

3.2 Superfícies construïdes amb Splines Cardinals

Si, a partir de la informació (11) sobre la malla rectangular,

- construïm la base de splines cardinals en direcció x , $C_1(x), \dots, C_n(x)$.
- construïm la base en direcció y , $C_1(x), \dots, C_m(x)$.

podem definir la superfície d'interpolació construïda amb splines cardinals, /5/, com

$$s(x,y) = \sum_{i=1}^n \sum_{j=1}^m z_{ij} C_i(x) C_j(y) \quad (13)$$

En aquest cas,

- La superfície $s(x,y)$ està composta també de trossos bicúbics definits a cada rectangle de la malla
- Hi ha continuïtat entre els diferents trossos fins les derivades parcials segones.
- L'interpolant passa per tots els punts donats z_{ij} , i es pot demostrar que és el de màxima suavitat.

En resum, els dos mètodes exposats per a la construcció de superfícies tenen com a principal diferència el passar o no pels punts donats. En canvi, a tots dos es parteix del càlcul dels elements de les dues bases en direccions x, y , bases que només depenen de la malla i no de les altures. Per tant, un canvi en les altures implica refer només els sumatoris (12) o (13). A més, cal observar que el càlcul de les bases comporta únicament el doble de treball que en el cas d'una dimensió.

Probablement, l'esquema basat en B-Splines serà més utilitzat en processos de disseny de formes que no contenen punts de pas obligat mentre que utilitzarem els splines cardinals en aplicacions on s'imposen a priori --unes certes cotes a la superfície.

4. PUNTS IRREGULARMENT DISTRIBUÏTS; ESQUEMES LOCALS

Desafortunadament, en molts casos se'ns demana construir una superfície que passi per p punts distribuïts irregularment en el pla,

$$(x_i, y_i, z_i) \quad i = 1, \dots, p \quad (14)$$

Tot i que el problema és molt més engorrosos - que en el cas de malla rectangular, i que és aconsellable anar a parar a aquest últim sempre que sigui possible, hi ha casos que és - forçós plantejar-se el cas general (14): da- des experimentals que volem interpolar, ca- sos en els quals cal construir la superfície a partir del coneixement de les corbes de nivell, etc. En aquests casos certs esquemes parteixen d'un únic interpolant global que - és obligat a passar pels punts donats; així ho fan els splines tipus placa prima, /7/, que prenen la forma d'una placa deformada, (amb mínima energia). L'equació de l'interpolant és

$$f(x,y) = \sum_{i=1}^p \lambda_i [(x-x_i)^2 + (y-y_i)^2] \cdot \log [(x-x_i)^2 + (y-y_i)^2] + \alpha_1 x + \alpha_2 y + \alpha_3 \quad (15)$$

on els vectors $\lambda = (\lambda_1, \dots, \lambda_p)^T$ i

$\alpha = (\alpha_1, \alpha_2, \alpha_3)^T$ compleixen,

$$K\lambda + E\alpha = (z_1, \dots, z_p)^T$$

$$E^T \lambda = 0$$

$$K_{ij} = \begin{cases} [(x_i - x_j)^2 + (y_i - y_j)^2] \cdot \log [(x_i - x_j)^2 + (y_i - y_j)^2] & , i \neq j \\ 0 & \text{si } i=j \end{cases} \quad (16)$$

$$E = \begin{bmatrix} 1 & x_1 & y_1 \\ \text{-----} \\ \text{-----} \\ 1 & x_p & y_p \end{bmatrix}$$

Encara que aquest és l'esquema òptim d'interpolació, s'utilitza poc en comportar un temps de càlcul de l'ordre de p^3 . En canvi s'han utilitzat molt més ampliamt els esquemes lo- cals basats en una triangularització previa del domini. En ells,

- Es divideix el pla x-y en un cert nombre de triangles, cada un dels quals té com a vèr- texts les projeccions de tres dels punts do- nats sobre el pla x-y.
- Es construeix una superfície d'interpola- ció diferent a cada triangle, procurant unes condicions correctes de continuïtat en pas- sar d'un triangle a un altre.

En particular, Lawson, /8/, ha desenvolupat un mètode òptim de triangularització a par -

tir dels punts (x_i, y_i) , $i = 1, \dots, p$. En - ell, utilitzant un procés quasi seqüencial, s'arriba a la configuració final en què per a tot quadrilàter convex, la divisió en dos triangles maximitza el seu mínim angle inte- rior. A la figura 5 es poden veure exemples de triangularització.

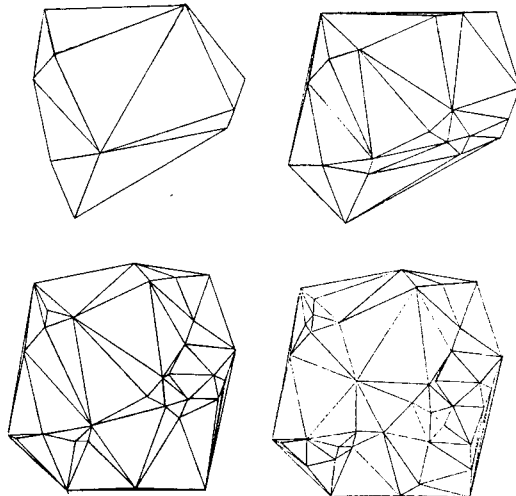


Figura 5: Triangularitzacions obtingu- des amb 10, 20, 30 i 40 punts. En cada cas s'han afegit 10 punts més a la configuració precedent.

Akima, /9/, ha publicat un algorisme d'inter- polació que es basa en una primera triangula- rització seguint l'esquema de Lawson seguida de la interpolació a cada triangle amb una - funció polinòmica de grau cinc:

$$f(x,y) = \sum_{j=0}^5 \sum_{k=0}^{5-j} q_{jk} x^j y^k \quad (17)$$

en la que els 21 coeficients q_{ij} es determi- nen a partir del coneixement de,

- els valors de la funció i de les seves de- rivades parcials primeres i segones, $z, z_x, z_y, z_{xx}, z_{yy}, z_{xy}$, a cada un dels tres vèr- texts del triangle (18 condicions)
- els valors de la derivada parcial en sentit normal als costats, calculada al punt mig d'aquests.

En relació amb l'algorisme d'Akima cal obser- var,

- Que cal realitzar un primer càlcul aproxi-

mat de les derivades parcials que són necessàries per a obtenir els q_{ij} . En aquest sentit, l'algorisme fa ús dels punts més propers al que analitza, i cal indicar-li a priori el nombre de punts addicionals a fer servir.

- Que es pot demostrar que l'interpolant obtingut pertany a C^1 (continuitat de derivades parcials primeres entre triangles), la qual cosa fa que sigui un algorisme molt utilitzat.

5. PUNTS IRREGULARMENT DISTRIBUITS: FUNCIONS SPLINE GLOBALS

Com ja hem vist, els mètodes basats en una primera triangularització del domini, donen lloc a interpolants locals (d'equació diferent a cada triangle) que com a màxim presenten continuïtat a nivell de les derivades parcials primeres. Si volem en canvi una suavitat més gran, hem d'anar a interpolants globals no basats en triangles sinó en producte tensorial d'interpolants unidimensionals. Suposem que coneixem les coordenades (14) de p punts donats, i que volem construir una superfície suau que passi per tots ells. El mètode que presentem a continuació utilitza un spline bicúbic global i segueix els passos següents:

- a- Es construeix una malla rectangular $\Delta_x \Delta_y$ amb n intervals en direcció x i m intervals en direcció y . Normalment Δ_x s'extén des de $\min(x_i)$ fins a $\max(x_i)$, i el mateix per Δ_y . D'altra banda, n i m són escollits de manera que $|n-m| \leq 1$, $(n+1) \cdot (m+1) > p$, $(n+1) \cdot (m+1) < 1.5p$ (sempre és possible si $p \geq 3$, només cal fer $n = \text{ent}(\sqrt{p})$ i $m=n$ si $p \geq n(n+1)$ o bé $m=n-1$ si $p < n(n+1)$)
- b- Es calcula la matriu F de $(n+1) \cdot (m+1)$ valors a la malla anterior, de manera que la superfície obtinguda per splines cardinals a partir de les altures f_{ij} interpoli els p punts donats.
- c- En aquest moment, hem reduït la interpolació bidimensional de p punts irregularment distribuïts, a la dels $(n+1) \cdot (m+1)$ punts de la malla $\Delta_x \cdot \Delta_y$, problema que podem resoldre fàcilment segons hem vist a l'apartat 3, equació (13).

La fase b- del procés no té solució única,

doncs existeixen $(n+1) \cdot (m+1) - p$ graus d'indeterminació. L'elecció dels valors de n i m que hem exposat a l'apartat a- minimitza aquesta indeterminació així com l'ocupació de memòria per part de la matriu intermitja F . (Ara bé, és evident que tot el procés continua essent vàlid amb qualsevol elecció de n , m tals que $N = (n+1) \cdot (m+1) > p$). Si definim f com el vector de dimensió N que conté totes les components de F , la naturalesa lineal de la interpolació bidimensional per splines cardinals ens porta a què

$$z = Bf \quad \text{amb} \quad z = (z_1, \dots, z_p)^T \quad (18)$$

B matriu de $p \times N$

on B és una matriu funció de la malla i de (x_i, y_i) , $i = 1, \dots, p$. Per tal de determinar unívocament f , imposarem que a més es minimitzi un cert funcional quadràtic

$$v = f^T M f \quad (19)$$

Discutirem a continuació criteris de minimització de la norma o de la variància de f (ja que s'ha demostrat, /5/, que la minimització de la corbatura entesa com integral de quadrats de derivades parcials segones, produeix interpolants abruptes i és per tan un criteri no desitjable).

Criteri 1 (mínima norma): De tots els interpolants definits per vectors f que compleixen (18), escollirem el que té mínima norma $f^T f$. En altres paraules, volem minimitzar (19) amb $M = I$.

Criteri 2 (criteri quadràtic respecte la mitjana, o de mínima variància): De tots els interpolants definits per (18) escollim el que té mínima la norma del vector de diferències entre f i la seva mitjana \hat{f} :

$$\hat{f} = \alpha U^T f, \quad \alpha = \frac{1}{N}, \quad u^T = (1, \dots, 1) \quad (20)$$

$$f - u\hat{f} = (I - \alpha U) f = u u^T$$

i ens cal minimitzar (19) amb $M = (I - \alpha U)^2 = I - \alpha U$, ja que $I - \alpha U$ és idempotent.

Criteri 3 (criteri quadràtic respecte l'interpolant lineal): D'entre els interpolants definits per (18) escollim el que té mínima la norma del

vector de diferències entre f i els valors als punts de la malla $\Delta_x \Delta_y$ de l'interpolant lineal a trossos que anomenarem π . Aquest últim es defineix mitjançant el següent algorisme: donat un punt (x, y) , l'altura z de l'interpolant és la del pla que passa pels 3 punts (x_i, y_i, z_i) , (x_j, y_j, z_j) , (x_k, y_k, z_k) més propers al donat, en la seva projecció sobre el pla $x-y$. Dient g al vector de valors d'aquests interpolant en els N punts de la malla, volem minimitzar, en lloc de (19),

$$v = (f - g)^T (f - g) \quad \text{o bé} \quad (21)$$

$$v_1 = f^T f - 2f^T g$$

Utilitzant els multiplicadors de Lagrange, i dient λ al vector (de dimensió p) de multiplicador, el càlcul de f es concreta en:

1. Criteri de mínima norma

$$f = B^T \lambda$$

$$Bf = z \quad (22)$$

2. Criteri de mínima variància

$$Mf = B^T \lambda, \quad M = I - \alpha U$$

$$Bf = z \quad (23)$$

3. Criteri de mínima variància respecte

$$f = g + B^T \lambda$$

$$Bf = z \quad (24)$$

En definitiva, per tant, en qualsevol dels tres casos és important poder generar el subespai de R^N definit per $B^T \lambda$, on λ és un vector qualsevol de R^p . Així com en els criteris 1 i 3 la solució és immediata

Solució-Criteri 1

$$\lambda = (BB^T)^{-1} z$$

$$f = B^T \lambda \quad (25)$$

Solució-Criteri 3

$$\lambda = (BB^T)^{-1} (z - Bg)$$

$$= B^T \lambda \quad (26)$$

hem d'estudiar més a fons les propietats de la transformació $M = I - \alpha U$ per tal d'obtenir la solució de (23) amb un temps de càlcul menor que el que tindriem on resoldre directa-

ment el sistema,

$$\left[\begin{array}{c|c} M & B^T \\ \hline B & 0 \end{array} \right] \begin{pmatrix} f \\ -\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ z \end{pmatrix} \quad (27)$$

En aquest sentit, podem dir:

- Que la matriu $M = I - \alpha U$ és singular. En efecte, $\det(M - I) = -\alpha \det(U) = 0$ i 1 és un valor propi de M amb multiplicitat $N-1$; com que la traça de M és $N-1$, l'altre valor propi és zero. I el vector propi corresponent és u : $Mu = 0$.
- En conseqüència, donada la transformació directa h ,

$$h = Mf$$
 no és possible anar enrera per tal de retrobar f .
- Donat h , existeix una simple infinitat f^* de vectors tal que $h = Mf^*$, on

$$f^* = h + \mu u \quad (\mu \text{ qualsevol escalar})$$

En efecte:

$$Mf^* = (I - \alpha U)(h + \mu u) = h + \mu u - \alpha U h - \alpha U \mu u = h - \alpha U h = Mh = M^2 f = Mf = h \quad (28)$$

ja que $M = M^2$, i $u = \alpha U u$.

A més, segons (23) cal que $B^T \lambda$ sigui de la forma Mf ; per tant, $MB^T \lambda = B^T \lambda$, la qual cosa és equivalent a dir $UB^T \lambda = 0$ i $u^T B^T \lambda = 0$. En definitiva, podem explicitar f en les equacions (23) de la forma:

$$f = B^T \lambda + \mu u \quad (29)$$

$$Bf = z$$

amb

$$u^T B^T \lambda = U^T B^T (BB^T)^{-1} [z - \mu Bu] = 0 \quad (30)$$

Per obtenir la solució en el criteri 2, un cop calculat el valor de μ amb (30), tenim, a partir de (29),

$$(BB^T) \lambda = z - \mu Bu$$

$$f = B^T \lambda + \mu u \quad (31)$$

En resum, el procés d'interpolació consisteix en

- Construir la malla $\Delta_x \Delta_y$ de $N = (n+1) \cdot (m+1)$ punts (pas a)
- Calcular el vector de valors f del spline als punts de la malla. Cal utilitzar les

equacions (25), (30)-(31) o (26) segons el criteri d'optimalitat escollit (pas b)

- Determinar les 2 bases de splines cardinals en direccions x e y
- Calcular valors de l'interpolant amb l'equació (13), on z_{ij} seran els valors dels components de f .

6. RESULTATS I DISCUSSIÓ

A l'apartat anterior hem exposat tres criteris diferents per a obtenir interpolants òptims amb continuïtat fins les derivades parcials segones a partir d'un conjunt qualsevol de punts donats. De fet, l'elecció d'un d'entre aquells criteris dependrà de l'aplicació concreta i de què es demana a l'interpolant. En particular,

- El criteri 1 (norma) dona lloc a superfícies que tendeixen a concentrar-se prop del pla $x-y$, a part de passar evidentment pels punts donats.
- El criteri 2 (variància) dona lloc a superfícies poc allunyades del valor mig. Evidentment, els resultats coincideixen amb els del criteri 1 si la mitjana és nul·la.
- El criteri 3 (aproximació a l'interpolant lineal) porta a superfícies més suaus que els altres, i a més, com és fàcil de comprovar, és l'únic que interpola plans donant plans. Per aquestes raons serà normalment el més utilitzat, a no ser que es tinguin raons clares a favor d'un dels altres dos.

Les figures 6 i 7 mostren els interpolants obtinguts pel mètode d'Akima (/9/) i el criteri 3, respectivament, a partir d'un mateix conjunt de dades. A les dues figures s'ha representat amb petites fletxes la posició dels 15 punts inicials. De la comparació d'ambdues figures es pot observar que les superfícies són força coincidents excepte, com és lògic, a la zona d'extrapolació -a l'esquerra de les figures-. A més, a la zona que envolta el punt de màxima altura s'observa una major discontinuïtat a la figura 6 -- que a la 7. El criteri basat en splines cardinals donen lloc per tant a interpolants més suaus.

A part, s'ha efectuat un estudi comparatiu

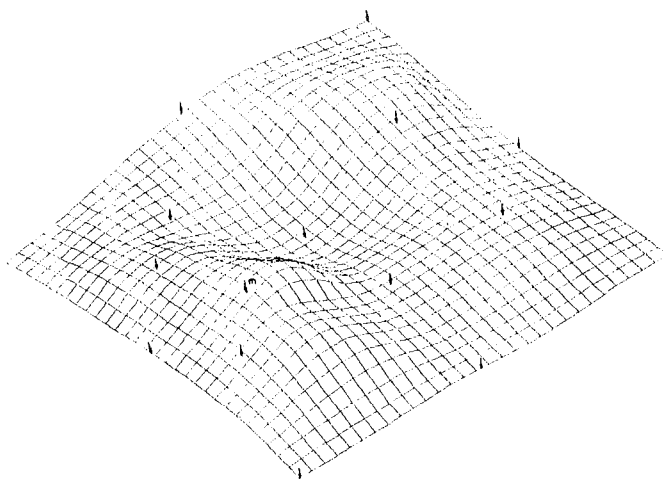


Figura 6: Interpolació c^1 per triangularització del conjunt de 15 punts indicats per fletxes. El punt d'altura màxima és el m.

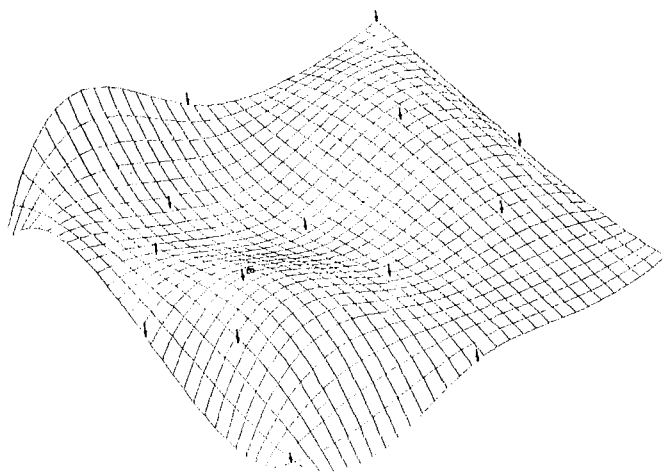


Figura 7: Interpolació c^2 segons el criteri 3, del conjunt de 15 punts indicats per fletxes. El punt d'altura màxima és el m.

basat en la interpolació de la superfície $f(x,y) = \exp[(x+2y)/3]$, a partir de 10, 20 i 30 punts a l'atzar dins del rectangle $0 < x < 1, 0 < y < 1$. La distribució dels punts és la indicada a la figura 5; per a cada conjunt de punts es partia de l'anterior incrementant-lo en 10 punts més a l'atzar. A la taula I es presenta una comparació de resultats:

TAULA I

Comparació de Variàncies i Errors Quadràtics en funció del Criteri i del Nombre de Punts.

P	σ_T	σ_2	σ_3	σ_{A3}	σ_{A5}	ϵ_2	ϵ_3	ϵ_{A3}	ϵ_{A5}
10	.113	.101	.109	.117	.115	.0064	.00062	.0006	.00026
20	.118	.107	.113	.118	.120	.0087	.0039	.00003	.00015
30	.131	.126	.127	.05	.03	.0056	.0019	2.88	2.87

on,

σ_T és la variància dels punts considerats de superfície analítica inicial

σ_2, ϵ_2 són la variància i error quadràtic mig respecte a la superfície analítica, en el criteri 2.

σ_3, ϵ_3 id. en el criteri 3

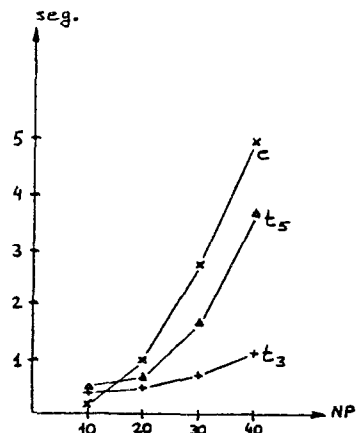
$\sigma_{A3}, \epsilon_{A3}$ id. en el mètode d'Akima que utilitza 3 punts addicionals per al càlcul de derivades.

$\sigma_{A5}, \epsilon_{A5}$ id. en el mètode d'Akima amb 5 punts addicionals.

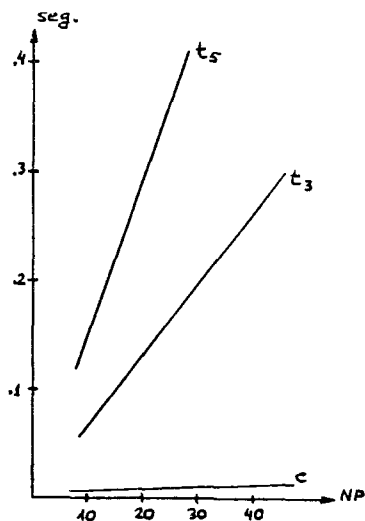
D'aquests resultats es pot deduir,

- Que, com era d'esperar, sempre $\sigma_T > \sigma_3 > \sigma_2$. (La variància és el que hom minimitza en el criteri 2). Normalment a més, σ_{A3} i σ_{A5} són més grans que σ_T .
- Que l'error quadràtic és menor a A3 i A5 que als criteris 2 i 3, encara que aquests són acceptablement petits per a la majoria d'aplicacions.
- Que els criteris 2 i 3 presenten un comportament més homogeni i estable que els altres. Això queda palès en el cas $p = 30$, en el que els interpolants d'Akima són molt dolents, amb errors quadràtics realment importants. La raó cal cercar-la en el fet que aquest últim mètode queda molt més afectat per la distribució dels punts, en ser local. En canvi els criteris 2 i 3 produeixen interpolants globals i de comportament més uniforme.
- Comparant els criteris 2 i 3, l'últim és sempre millor quan a l'error quadràtic.

Finalment, a la figura 8 es presenta una comparació de temps de procés entre els diferents mètodes, en funció del nombre p de punts. Com es pot veure, tot i que el temps inicial -necessari pel càlcul del vector f és més elevat que en els altres mètodes, posteriorment és molt fàcil el càlcul de cada nou valor interpolat. És interessant d'observar que el temps de la segona fase creix



Temps de càlcul del primer punt de la superfície.



Temps de càlcul del valor interpolat a cada nou punt.

Figura 8: Comparació de temps de procés. c: criteris 2 i 3 (els temps són molt semblants) t_5, t_3 : criteri d'Akima amb 5 i 3 punts addicionals. Temps de procés a un ordinador Data General Eclipse S-130.

linealment amb p a tots els criteris.

Per exemple, si comparem els criteris c_i i t_3 per a $p = 30$, el temps de càlcul per a n_i valors interpolats és respectivament $2.7 + 0.01 n_i$, i $0.7 + 0.2 n_i$; per tant només és més ràpid el criteri t_3 si $n_i < 11$. Donat que la fase costosa és la primera els criteris c són sobretot aconsellables quan cal interpolat a molts punts.

7. CONCLUSIONS

Després d'una presentació dels principals mètodes existents a l'actualitat per a la construcció de corbes i superfícies, en el present article s'ha exposat un algorisme per a la interpolació de punts irregularment distribuïts, que dona lloc a superfícies bicúbiques a trossos i amb continuïtat fins les derivades parcials segones. De la seva comparació amb el mètode d'Akima se'n dedueix que dona lloc a interpolants més suaus i que presenten errors quadràtics respecte la superfície teòrica acceptablement petits. El temps de càlcul en funció del nombre de punts inicials el fa un algorisme molt aconsellable - en casos en què cal obtenir molts punts interpolats a partir d'un nombre no massa elevat de punts coneguts.

8. REFERÈNCIES

- /1/ ROGERS, D.F. "Mathematical Elements for Computer Graphics"; Mc. Graw Hill, New York 1976.
- /2/ DE BOOR, C. "A Practical Guide to Splines". Springer Verlag, New York, 1978.
- /3/ GORDON, W.J. and RIESENFELD, R.F. "Bernstein-Bézier Methods for the Computer-Aided Design of Free-Form Curves and Surfaces", Journal of the A.C.M. 21 n2, apr. 1974, pp. 293-310
- /4/ ELLIS, T.M., McLAIN, D.M., "A New Method of Cubic Curve Fitting Using Local Data" A.C.M. Trans. on Math. Software 3 n2, 1977, pp. 175-178.
- /5/ BRUNET, P. "Surface Fitting by Means of Splines". Proc. of Eurographics-80, Ginebra, Set. 1980.
- /6/ WU, J.F., ABEL, D.P., GREENBERG, D.P. "An Interactive Computer Graphics Approach

to Surface Representation", Comm. of the A.C.M. 20 n210, 1977.

- /7/ PAIHUA, L. "Quelques Méthodes Numériques pour le calcul de Fonctions Spline a Une et Plusieurs Variables". Thèse, Lab. Math. Appl. Grenoble, mai 1978.
- /8/ LAWSON, C.L. "Software for C^1 Surface Interpolation", in Math. Software III, edited by J. Rice, Ac. Press 1977.
- /9/ AKIMA, H., "A Method of Bivariate Interpolation and Smooth Surface Fitting for Irregularly Distributed Data Points", A.C.M. Trans. on Math. Software, 4 n2, June 1978, pp. 148-159.

