Curs pràctic de Maple

Pràctica 13

13 Interpolació i ajust de dades

13.1 Aproximació de funcions, polinomi interpolador

Recordem que Maple pot representar gràficament un núvol de punts $(x_1, y_1), \ldots, (x_n, y_n)$ amb la funció `plot`.

Exercici 13.1
Definiu una llista `l` com els següents punts de $\mathbb{R}^2$: $(0,2)$, $(1,1)$, $(2,2)$, $(3,1)$, $(4,2)$, $(5,1)$, $(6,2)$, $(7,1)$, $(8,2)$, $(9,1)$ i $(10,2)$. Dibuixeu en una gràfica el núvol de punts corresponent.

Donats n punts $(x_1, y_1), \ldots, (x_n, y_n)$ amb $x_i \neq x_j$ si $i \neq j$ existeix un polinomi p de grau menor o igual a $n-1$ complint que $p(x_i) = y_i$ per a tot $i = 1..n$, que anomenem *polinomi interpolador*. La funció que calcula polinomis interpoladors està al paquet **CurveFitting** i s'anomena `PolynomialInterpolation`.

Exemple 13.1
Si volem calcular el polinomi que interpola els punts $(0,0), (1,1)$ i $(2,4)$, que ha de ser de grau menor o igual a 2, ho podem fer amb Maple:

```
> with(CurveFitting); ll:=[[0,0],[1,1],[2,4]]:
> PolynomialInterpolation(ll,x);
```

Exercici 13.2
Calculeu el polinomi interpolador p als punts de l'exercici anterior. Quan val el polinomi avaluat al punt 5? I al punt 9.5? Feu la representació gràfica dels punts i el polinomi interpolador.

13.2 Spline

Una altra opció per a trobar funcions que passin per punts fixats és "*agrupar*" els punts en subconjunts més petits i calcular el polinomi de grau més petit per a cada un dels subconjunts. Per exemple, podem considerar que la unió dels punts mitjançant segments es correspon a considerar els punts de dos en dos i considerar el polinomi de grau 1 (recta) que els uneix.
Dins el paquet **CurveFitting** tenim la funció `Spline` (la `S` en majúscules) per a definir aquestes noves aproximacions.

Exemple 13.2
Si volem unir els punts $(0,0), (1,1)$ i $(2,4)$ amb rectes escrivim:

```
> ll:=[[0,0],[1,1],[2,4]]:
> Spline(ll,x,degree=1);
```

On l'opció `degree=1` és per a definir el grau dels polinomis que estem calculant.

Exercici 13.3
Dibuixeu tres gràfiques amb els "*splines*" de graus 1, 2 i 3 de la llista de punts de l'exercici 13.1.

13.3 Interpolació racional

Hi ha cops en els que pot interessar obtenir no un polinomi si no, més generalment, funció racional (quocient de polinomis) per a interpolar una llista de valors coneguts. Dins el paquet **CurveFitting** hi ha la funció `RationalInterpolation` que pot fer aquesta feina.

Exemple 13.3
Si teniu la llista de punts

```
> a := [[-3, -55/14],[-2, -19/9],[-1, -1/6],[0, 1],[1, 5/6],[2, 5/9],[3, 11/14]];
> plot(a,style=point);
```

podeu buscar una expressió racional que interpoli aquests valors amb

```
> with(CurveFitting):
> RationalInterpolation(a,x);
```

(Observeu que, apart de la llista de valors per interpolar, s'ha d'especificar un nom de variable per a poder escriure l'expressió resultant).

Aquesta funció també es pot cridar posant com arguments dues llistes separades per als valors de les x's i de les y's com en el següent

```
> valx := [-3, -2, -1, 0, 1, 2, 3];
> valy := [-55/14, -19/9, -1/6, 1, 5/6, 5/9, 11/14];
> RationalInterpolation(valx,valy,x);
```

Es poden consultar altres opcions amb `?RationalInterpolation`.

13.4 Exercicis d'interpolació

Exercici 13.4
Feu una funció (o procediment) `dibuix` que fixada una llista de punts dibuixi de colors diferents el núvol de punts, els segments que uneixen els punts, el "*spline*" de grau 2, el "*spline*" de grau 3 i el polinomi interpolador en una sola gràfica.

Exercici 13.5
Considereu la llista de punts:
`lli1:=[[-9,1], [-8,1], [-7,1], [-6,1], [-5,2], [-4,3], [-3,5], [-2,8], [-1,12], [0,16], [1,12], [2,8], [3,5], [4,3], [5,2], [6,1], [7,1], [8,1], [9,1]]`.
Dibuixeu el núvol de punts en una gràfica. Apliqueu la funció `dibuix` de l'exercici anterior a aquest núvol de punts. Que observeu?

Exercici 13.6
Definiu una funció `aproxf` que depengui de 4 paràmetres `f, a, b, n` que dibuixi la funció `f` entre `a` i `b` i el polinomi de grau menor o igual a n que interpoli a `n+1` punts equiespaiats entre `a` i `b` (inclosos).

Exercici 13.7
Apliqueu la funció `aproxf` a la funció $f(x) = \dfrac{1}{1+x^2}$ entre -5 i 5 i per als valors de `n` 1, 3, 6, 10 i 15. Què observeu?

Exercici 13.8
Feu un gràfic on es puguin veure al mateix temps les interpolacions polinòmica i racional del núvol de punts `a:= [[-3, -55/14], [-2, -19/9], [-1, -1/6], [0, 1], [1, 5/6], [2, 5/9], [3, 11/14]];` (reflexioneu sobre les diferències que s'observen entre els dos resultats).

13.5 Ajust a les dades

Una qüestió que es planteja sovint amb dades experimentals és la següent: es realitza un experiment i es voldria saber si hi ha una relació entre dues de les seves propietats (una relació de proporcionalitat per exemple). És a dir, pot tenir interès descriure a través d'una funció la relació entre dues variables quantitatives X i Y. Posem $Y = f(X)$ on f és una funció que depèn d'uns paràmetres $a_1, ..., a_d$ que haurem de determinar per garantir un *bon* ajust. Naturalment, a més de disposar d'observacions, $(x_1, y_1), ..., (x_n, y_n)$, dades que suposem bivariants (x, y) $x, y \in \mathbb{R}$, hem de recorre a un criteri que valori la bondat de l'ajust. El criteri que estudiarem és l'anomenat de *mínims quadrats.*

13.5.1 Llei de Hooke

Començarem plantejant un experiment concret. La llei de Hooke a la física ens diu que hi ha una relació de proporcionalitat entre el pes que es penja d'una molla i la distància que aquesta s'estira. És a dir, si x és un pes i y representa la distància que una molla s'estira al penjar-li aquell pes, existeix una constant k tal que $y = kx$. k s'anomena la constant d'elasticitat de la molla.

L'experiment que realitzarem consisteix en determinar aquesta constant k per a una molla concreta. Per això, es realitza l'experiment que consisteix en mesurar les distàncies que la molla s'estira per a diferents pesos i obtenim la següent taula,

Pes	2	2.5	3	3.5	4	5
cm	3.2	4	5	5.07	6.5	8

Si l'experiment es realitzés en condicions perfectes (mesures de precisió perfecte, sense fregament amb l'aire,...) les dades obtingudes serien de la forma $y_i = kx_i$ i per tant, seria fàcil determinar la constant k. Però com que no podem suposar condicions perfectes, obtenim unes dades que es troben lleugerament distorsionades.

Exercici 13.9
En uns eixos de coordenades representeu la taula de punts $(x, y) = (\text{pes}, \text{cm})$ de l'experiment.

L'objectiu doncs es trobar la recta de la forma $y = kx$ que millor s'ajusta a les dades obtingudes experimentalment. Per això hem de fixar un criteri. És a dir, què vol dir que una recta s'ajusta millor que una altra al conjunt de dades?

El criteri a fer servir és el següent: buscarem k tal que la suma de les distàncies (al quadrat) dels punts obtinguts experimentalment a la recta sigui mínima. És a dir, el valor de k tal que

$$\sum_{i=1}^{6} (y_i - kx_i)^2$$

pren el valor mínim. Observeu que aquesta funció mesura l'error comés en l'experiment respecte les condicions perfectes.

Exercici 13.10
Sigui $d(k) = \sum_{i=1}^{6} (y_i - kx_i)^2$, busqueu el valor de k pel qual d pren el valor mínim. En una mateixa gràfica representeu els punts obtinguts a l'experiment i la gràfica de la recta $y = kx$ pel valor de k obtingut.

Aquest mètode que acabem de descriure per aproximar dades obtingudes experimentalment per funcions és el que es coneix com el mètode dels mínims quadrats.

13.5.2 El mètode dels mínims quadrats

En general, aquest mètode s'utilitza quan es vol analitzar la relació funcional entre un conjunt de dades (x_i, y_i) obtingudes experimentalment. Es suposa conegut el tipus de funció f que les relaciona però no els paràmetres $a_1, \ldots, a_d$ de la funció.

El criteri de mínims quadrats proposa calcular els paràmetres $a_1, \ldots, a_d$ tals que l'expressió dels errors

$$\sum_{i=1}^{n} \big(y_i - f(x_i)\big)^2, \tag{1}$$

és mínima, és a dir, la suma dels desajustos quadràtics entre els valors y_i observats i els valors $f(x_i, ; a_1, ..., a_d)$ proposats per f.

El primer pas sempre consisteix en representar gràficament en uns eixos de coordenades els punts obtinguts experimentalment com ja heu fet en la secció anterior. D'aquest tipus de diagrama se'n diu un diagrama de dispersió. Dins el paquet **plots** hi ha la comanda `pointplot` que permet fer aquests tipus de gràfics fàcilment.

Exemple 13.4
Donades les observacions següents: $(0.70, 0.035)$, $(0.76, 0.025)$, $(0.37, -0.18)$, $(0.82, 0.045)$, $(0.29, -0.16)$, $(0.56, -0.058)$, $(0.42, -0.11)$, $(0.47, -0.085)$, el següent grup de comandes realitza la representació gràfica de les dades anteriors en un gràfic de dispersió:

```
> with(plots):
> dades:=[ [0.70,0.035],[0.76,0.025],[0.37,-0.18],[0.82,0.045],
> [0.29,-0.16],[0.56,-0.058],[0.42,-0.11],[0.47,-0.085]];
> pointplot(dades);
```

(Per a obtenir aquest gràfic també es pot utilitzar un `plot` com haureu fet abans. L'avantatge de `pointplot` és que no cal especificar ni `style=point` ni `symbol=circle`).

Un cop fet el diagrama, el següent pas consisteix en analitzar la forma del núvol de punts obtinguts per determinar quin tipus de funció s'hi ajustaria millor.

Exercici 13.11
Quin tipus de funció creieu que podríem triar per aconseguir un bon ajust a les dades de l'exemple anterior? (Proveu *a ull* una funció lineal del tipus $f(x) = a\,x + b$ i feu un gràfic de la situació).

Exercici 13.12
Donades les observacions: $(-1, 5)$, $(-0.4, 2.5)$, $(0.1, 2.1)$, $(0.8, 3.98)$, $(1.2, 6.5)$, $(1.5, 8.8)$, $(2.1, 15.2)$,

Representeu les dades en un gràfic de dispersió. En aquest cas, quina mena de funció creieu que podríem triar per ajustar a les dades? (sembla el més adequat triar $y = f(x) = a + bx + cx^2$, intenteu donar valors als paràmetres a, b i c per ajustar aquestes dades).

Un cop determinat el tipus de funció, podem calcular els paràmetres pels quals la suma dels quadrats dels errors de mesura és mínim. La comanda que calcula els paràmetres que minimitzen aquests errors és `LeastSquares` del paquet **CurveFitting**. Com a paràmetres necessita el nom de les variables, el tipus de funció i les llistes de dades experimentals.

En l'exemple següent veureu com utilitzar aquesta comanda pel cas d'ajust per una recta. Aplicarem `LeastSquares` a les dades obtingudes per l'experiment de la molla. (Anomeneu `Pes` i `Cm` les dades obtingudes en l'experiment de la molla).

```
> with(CurveFitting):
> Pes:=[2,2.5,3,3.5,4,5];
> Cm:=[3.2,4,5,5.07,6.5,8];
> PesCm:=[seq([Pes[i],Cm[i]],i=1..nops(Pes))];
> minq:=LeastSquares(Pes,Cm,pes);
> minq2:=LeastSquares(PesCm,pes);
```

(Noteu que la comanda `LeastSquares` es pot aplicar o bé al parell de llistes `Pes`, `Cm` o a la llista individual de les mesures de pes i estirament combinades. Aplicar una versió o l'altre només dependrà de la manera en la que es tinguin guardades les dades).

Exercici 13.13
Compareu aquest resultat amb l'obtingut anteriorment i feu un gràfic amb el diagrama de dispersió juntament amb la recta de mínims quadrats que acabeu de calcular.

Exercici 13.14
Calculeu els paràmetres de la recta que millor s'ajusta a les dades de l'exemple 13.4. En una mateixa gràfica representeu el diagrama de dispersió i la recta obtinguda.

En el cas d'ajustar altra tipus de corbes, com és el cas de l'exercici 13.12 on un polinomi de grau 2 semblava més adequat, la sintaxi és

```
> dades:=[[-1,5],[-0.4,2.5],[0.1,2.1],[0.8,3.98],[1.2,6.5],[1.5,8.8],[2.1,15.2]];
> curva:=LeastSquares(dades,x,curve=a*x^2+b*x+c);
```

Exercici 13.15
Representeu en una mateixa gràfica el diagrama de dispersió i la funció quadràtica obtinguda.

Exercici 13.16
El punt d'ebullició d'una mescla d'etanol i aigua depèn de la proporció d'etanol a la mescla. S'ha realitzat un experiment en el que s'han obtingut els resultats següents

Proporció	.0	.0190	.0721	.0966	.1238	.1661	.2337	.2608
Temperatura	100.	95.5	89.0	86.7	85.3	84.1	82.7	82.3

Proporció	.3273	.3965	.5079	.5198	.5732	.6763	.7472	.8943
Temperatura	81.5	80.7	79.8	79.7	79.3	78.74	78.41	78.15

(no cal dir que les temperatures venen donades en graus centígrads).

Si sabem que la funció que relaciona el punt d'ebullició t i la proporció d'etanol p en la mescla és de la forma $t = a\,e^{-14p} + b\,p + c$, calculeu les constants a, b, c per a l'etanol. Feu un gràfic on es mostrin els resultats de l'experiment (punts d'ebullició respecte la proporció d'etanol) i la corba que ajusta aquests resultats.

13.5.3 Relacions no polinòmiques

No tots els tipus de relacions que voldrem analitzar entre variables serà de tipus polinomial. Hi ha experiments que es modelen amb d'altres tipus de funcions matemàtiques com per exemple, la desintegració de substàncies radioactives.

Un punt clau en la determinació de l'edat d'una pintura o un fòssil és el fenomen de la radioactivitat. El físic Rutherford va determinar que els àtoms de certs elements radioactius són inestables i que, en un interval de temps donat, uns fracció fixa dels àtoms presents en aquell instant de temps es desintegra espontàniament per formar altres substàncies. Aquest principi té com a conseqüència que el nombre d'àtoms presents a cada instant de temps segueix una funció exponencial

$$N(t) = N_0 e^{-\lambda(t-t_0)},$$

on t_0 és l'instant on es van començar a mesurar el nombre d'àtoms. Es diu que λ és el coeficient de desintegració i depèn de cada substància.

L'estratègia és la mateixa que hem seguit fins ara: donades unes dades experimentals preses sobre un material concret (nombre d'àtoms a certs intervals de temps) volem obtenir el coeficient de desintegració

λ. Aquesta dada és molt útil si volem fer prediccions sobre el nombre d'àtoms en temps futurs o l'edat del material.

El primer que s'ha de fer és transformar aquest problema d'ajust de dades per una exponencial en un problema d'ajust lineal. Aquesta tècnica que ara descriurem és molt utilitzada i s'anomena la transformació logarítmica.

Sigui $y = ke^{at}$, si apliquem el logaritme als dos costats de l'equació obtenim $\ln(y) = \ln(k) + at$ que és lineal en $\ln(k)$ i a (anomenarem $K = \ln(k)$).

Exercici 13.17

S'han pres unes dades experimentals sobre dues mostres de material radioactiu. Ens asseguren que totes dues són de plom 210. Per verificar-ho prenem diferents mesures dels àtoms de plom 210 en ambdues mostres al llarg de dos anys i obtenim les dades següents, on t es mesura en mesos i n_i són grams de plom a la mostra i.

t	n_1	n_2
1	22.23	32
2	1.2	1.3
4	0.04	0.002
6	0.0001	$0.36 \cdot 10^{-5}$
8	$0.36 \cdot 10^{-7}$	$0.6 \cdot 10^{-8}$
14	$0.1 \cdot 10^{-14}$	$0.3 \cdot 10^{-16}$
16	$0.3 \cdot 10^{-17}$	$0.5 \cdot 10^{-19}$
24	$0.3 \cdot 10^{-27}$	$0.35 \cdot 10^{-30}$

Representeu les dades gràficament en uns eixos de coordenades i determineu la funció que determina la seva desintegració (una per a cada una de les mostres). Ens han enganyat?