
Aproximació de funcions i regressió

PID_00266167

Álvaro Leitao Rodríguez

Temps mínim de dedicació recomanat: 3 hores



Álvaro Leitao Rodríguez

L'encàrrec i la creació d'aquest recurs d'aprenentatge UOC han estat coordinats per la professora: Teresa Sancho (2019)

Primera edició: setembre 2019
© Álvaro Leitao Rodríguez
Tots els drets reservats
© d'aquesta edició, FUOC, 2019
Av. Tibidabo, 39-43, 08035 Barcelona
Realització editorial: FUOC

Cap part d'aquesta publicació, incloent-hi el disseny general i la coberta, no pot ser copiada, reproduïda, emmagatzemada o transmesa de cap manera ni per cap mitjà, tant si és elèctric com químic, mecànic, òptic, de gravació, de fotocòpia o per altres mètodes, sense l'autorització prèvia per escrit dels titulars del copyright.

Índex

Introducció	5
1. Aproximació de funcions pel mètode de mínims quadrats.	11
1.1. Consideracions prèvies.....	11
1.1.1. Normes i seminormes	11
1.1.2. Producte escalar	13
1.1.3. Sistemes ortogonals.....	14
1.2. Formulació del problema	14
1.3. Solució del problema d'aproximació	15
1.4. Aproximació per polinomis ortogonals	18
1.4.1. Família triangular de polinomis	18
1.4.2. Polinomis ortogonals	20
1.4.3. Polinomis de Txebixev	21
1.4.4. Polinomis de Legendre	23
1.5. Mètodes de Fourier.....	25
1.5.1. Conceptes bàsics de l'anàlisi de Fourier.....	27
1.5.2. Sèries de Fourier	30
1.5.3. Teorema integral de Fourier	31
2. Regressió	34
2.1. Regressió lineal	34
2.1.1. Quantificació de l'error de la regressió lineal	36
2.2. Linealització de relacions no lineals.....	38
2.3. Regressió lineal polinòmica	40
2.4. Regressió lineal múltiple	41
Bibliografia	43

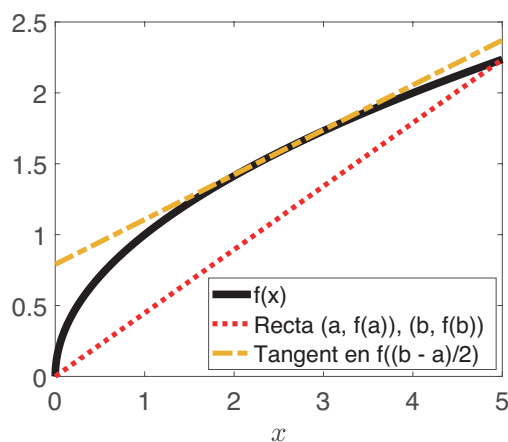
Introducció

Tal com s'ha comentat en la guia corresponent a l'activitat 2, hi ha moltes situacions en les quals és necessari aproximar una funció: per exemple, si aquesta és desconeguda, si és coneguda només en alguns punts o si la seva definició és intrínsecament complicada. L'objectiu, doncs, és aproximar una funció f per una funció alternativa \hat{f} , membre d'una classe de funcions amb les quals és fàcil treballar matemàticament (polinomis, funcions racionals, polinomis trigonomètrics). En aquesta classe cada funció està definida pels valors numèrics de certs paràmetres, però en aquest material ens centrarem en el problema d'aproximar funcions d'una variable en un interval tancat.

Funció definida en un interval $[a, b]$

En aquest exemple es presenta la corba de la funció $f(x) = \sqrt{x}$ en l'interval $[0,5]$. A més, hi incloem dues possibles aproximacions de f , dues de les més trivials: la recta que uneix els punts d'origen i final de la funció (en vermell) i la tangent en un punt ($x = \frac{b-a}{2} = 2.5$, en groc).

Figura 1

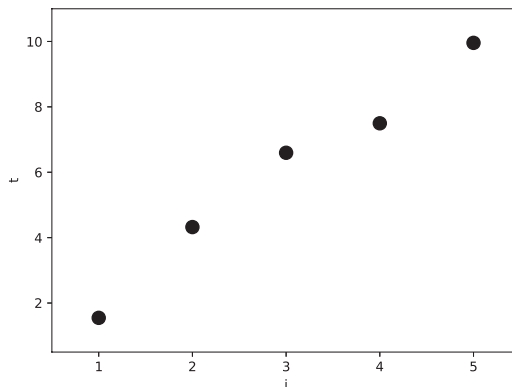


En una aproximació hi ha diferències entre la funció original i l'aproximada. Hi ha dos tipus de discrepàncies que s'han de tenir en compte: discrepàncies en les dades d'entrada i discrepàncies en el model (classe de funcions, forma, etc.) amb el qual volem ajustar les dades d'entrada. Aquestes discrepàncies s'anomenen *error de mesura* i *error en el model*, respectivament.

Mesurament experimental

Els punts en la figura 2 mostren els cinc temps de pas pel punt d'equilibri d'un pèndol. Cada temps pot ser un temps i -èsim de pas per a $i = 1, 2, 3, 4, 5$. Per les condicions de l'experiment, una relació de la forma $t = t_0 + iT$ (en què t_0 és el temps inicial i T és el temps que tarda el pèndol en cada cicle) es considera que és vàlida i d'alta precisió. En realitat, poden aparèixer alguns errors aleatoris en el procediment de mesurament, que serien els causants de les desviacions produïdes pel que fa a la linealitat (això seria el que s'espera per a un pèndol que s'estigüés balancejant). Aquestes desviacions expliquen que els valors dels paràmetres t_0 i T siguin incerts. A més, tenim cinc punts i només dos paràmetres per determinar. Així doncs, en aquest cas el problema és **sobredeterminat**.

Figura 2. Temps de pas d'un pèndol



Funció definida com a integral

Volem avaluar, de la manera més eficient possible, la funció f definida com a

$$f(x) = \int_0^x \exp(-t^2) dt,$$

per un $x \in [-1, 1]$, amb un error relatiu més petit que 10^{-4} .

Aquesta integral no pot resoldre's amb paper i llapis, mitjançant les tècniques clàssiques de càlcul de primitives, per la qual cosa la funció és complicada d'avaluar en els punts de l'interval donat. En aquest apartat veurem diferents maneres de construir una aproximació que vulgui minimitzar l'error del model i en permeti l'avaluació sense haver de resoldre la integral per a cada avalució.

Els exemples anteriors presenten dos casos diferents: en un cas tenim un conjunt de punts i volem conèixer valors intermedis; en l'altre volem avaluar una funció definida mitjançant una expressió complicada. En la pràctica, tant un conjunt de dades com la utilització de funcions senzilles (típicament polinomis) que serveixen per "modelar" la funció original són, en general, insuficients, en el sentit que un model sempre és una simplificació de la realitat i les dades poden estar distorsionades (poden contenir *soroll*). L'aproximació de funcions pot ser vista com un cas especial d'un problema més general i important: ajustar un model matemàtic a unes dades donades i a altres fets coneguts.

Ens centrarem principalment en el problema anomenat **aproximació lineal**, és a dir, en el fet que una funció desconeguda f sigui aproximada emprant una funció \hat{f} , que pot ser expressada com a combinació lineal d'unes altres.

Combinació lineal

Podem veure un polinomi de primer grau com una combinació lineal de la funció constant 1 (que no apareix explícitament) i la funció x :

$$\hat{f}(x) = c_0 + c_1x.$$

Aquesta expressió surt en moltes situacions, com en l'exemple del pèndol vist anteriorment o en l'equació de la recta que veurem en un altre exemple.

La combinació lineal de l'exemple anterior es pot generalitzar en dues direccions: el nombre de sumands o coeficients i la forma de les funcions corresponents. Així, la funció \hat{f} (que aproxima linealment f) tindrà aquesta forma genèrica:

$$\hat{f}(x) = c_0\phi_0(x) + c_1\phi_1(x) + \dots + c_n\phi_n(x),$$

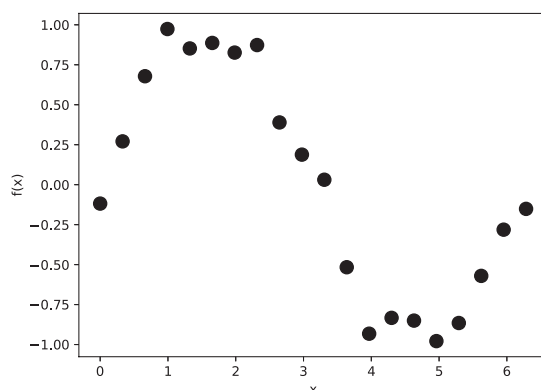
amb $n + 1$ funcions $\phi_0, \phi_1, \dots, \phi_n$, que són escollides prèviament. Els coeficients c_0, c_1, \dots, c_n són paràmetres el valor dels quals ha de ser determinat per aproximar f . Per exemple, si $\phi_0(x) = 1$, $\phi_1(x) = x$, $\phi_2(x) = x^2$, \dots , $\phi_n(x) = x^n$ o, equivalentment $\phi_i(x) = x^i$, la classe de possibles \hat{f} és la classe de polinomis de grau n . Llavors, el conjunt format per $\{1, x, x^2, \dots, x^n\}$ es denomina **base** del conjunt de tots els polinomis de grau n .

Com ja hem dit, la funció f pot ser donada de diferents maneres. Una situació bastant comuna és que f només és coneguda a través d'un conjunt d'avaluacions $f(x_0), f(x_1), \dots, f(x_m)$ de valors en una malla $x_0, x_1, x_2, \dots, x_m$. Assumirem sempre que els $m + 1$ valors de la malla són diferents.

Funció observada

En aquest exemple es presenta una funció coneguda només pel seu valor en certs punts, també anomenades *observacions*.

Figura 3



Si, per contra, la funció f és originàriament expressada com una corba o a través d'una fórmula complicada, la malla de punts es pot construir.

Recta que passa per dos punts

L'equació de la recta és donada per $y = \alpha + \beta x$. Com ja sabem, el problema de trobar l'expressió d'una recta amb dos punts donats té una única solució exacta, que, a més, és fàcilment calculable. Aquest problema encaixa en la formulació general d'aproximació de funcions, simplement amb $\hat{f}(x) = y$, $n = 1$, $\phi_0 = 1$, $\phi_1 = x$, $c_0 = \alpha$ i $c_1 = \beta$. Assumint que els dos punts donats tenen coordenades $(x_0 = 1, y_0 = 1)$ i $(x_1 = 2, y_1 = 3)$, podem escriure aquest sistema d'equacions:

$$\alpha + \beta = 1$$

$$\alpha + 2\beta = 3,$$

o, equivalentment,

$$c_0 + c_1 = 1$$

$$c_0 + 2c_1 = 3.$$

D'aquí podem deduir que, com ja sabíem, el **pendent** $\beta = \frac{y_1 - y_0}{x_1 - x_0}$ i α (respectivament, $c_1 = \frac{\hat{f}(x_1) - \hat{f}(x_0)}{x_1 - x_0}$ i c_0) s'obtenen immediatament resolent el sistema: $\beta = c_1 = 2$ i $\alpha = c_0 = -1$.

Imaginem que volem determinar els $n+1$ coeficients c_0, c_1, \dots, c_n de tal manera que $\hat{f}(x_i) = f(x_i)$, o que es troba tan a prop com sigui possible, per a tots els $m+1$ punts de la malla. Aquesta situació, considerant l'aproximació lineal, ens portaria un sistema d'equacions lineal amb $n+1$ incògnites i $m+1$ equacions:

$$\phi_0(x_0)c_0 + \phi_1(x_0)c_1 + \dots + \phi_n(x_0)c_n = f(x_0)$$

$$\phi_0(x_1)c_0 + \phi_1(x_1)c_1 + \dots + \phi_n(x_1)c_n = f(x_1)$$

⋮

$$\phi_0(x_m)c_0 + \phi_1(x_m)c_1 + \dots + \phi_n(x_m)c_n = f(x_m),$$

que, escrit en forma de matriu-vector, seria

$$(\phi_0(\mathbf{x}) | \phi_1(\mathbf{x}) | \dots | \phi_n(\mathbf{x})) \cdot \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_n \end{pmatrix} = f(\mathbf{x}), \quad (1)$$

en què $\mathbf{x} = (x_0, x_1, \dots, x_m)^T$, $\phi_j(\mathbf{x}) = (\phi_j(x_0), \phi_j(x_1), \dots, \phi_j(x_m))^T$ i $f(\mathbf{x}) = (f(x_0), f(x_1), \dots, f(x_m))^T$ són vectors columna.

Si $m = n$, llavors el sistema d'equacions tindrà, en condicions normals, una única solució. En aquest cas, aquesta manera de determinar \hat{f} es denomina **interpolació**, i s'explica en la guia corresponent a l'activitat 2. La condició que cal satisfer perquè la solució sigui única és que els vectors $\phi_0(\mathbf{x}), \phi_1(\mathbf{x}), \dots, \phi_n(\mathbf{x})$ han de ser linealment independents. Aquesta condició també es pot expressar dient que les funcions ϕ_j han de ser linealment independents de la malla.

Si $m > n$, llavors només en casos excepcionals es pot aconseguir que $\hat{f}(x_i) = f(x_i)$ per a tots els valors de la malla, x_i . El sistema té més equacions que incògnites, és a dir, el sistema és **sobredeterminat**. En aquest cas, el sistema d'equacions es pot satisfer només parcialment. Tal com s'esmenta en la guia de l'activitat 1, un mètode apropiat per resoldre sistemes d'equacions sobredeterminats és, per exemple, la **descomposició o factorització QR**.

La **sobredeterminació** es fa servir per aconseguir dos tipus diferents de suavització en l'aproximació:

- Per reduir l'efecte d'errors aleatoris en els valors de la funció (filtratge).
- Per donar a la corba una forma més suau entre els valors de la malla (suavització).

Una de les tècniques més utilitzades per resoldre sistemes d'equacions lineals sobredeterminats és el **mètode de mínims quadrats**. La seva aplicació a l'aproximació de funcions implica una computació relativament simple i, en moltes situacions, pot estar motivada per arguments estadístics. El mètode es descriurà a l'apartat 1.

1. Aproximació de funcions pel mètode de mínims quadrats

Una manera d'aproximar una funció per una altra és mitjançant els anomenats *mínims quadrats*. En aquest cas es tracta de trobar la funció \hat{f} més semblant a f d'acord amb una mesura concreta, precisament la dels mínims quadrats. Per definir aquesta mesura necessitarem alguns resultats previs, que recordarem breument, per formular després el problema d'aproximació d'una manera matemàticament correcta.

A partir d'aquest moment, diferenciarem entre els casos en què la funció que volem aproximar és contínua i coneguda per una expressió complicada (cas continu) i els casos en què la funció és expressada en termes d'uns valors donats en una malla (cas discret).

1.1. Consideracions prèvies

1.1.1. Normes i seminormes

Una manera molt útil d'estudiar l'anàlisi de funcions és observant una funció com un vector. Veurem que alguns tipus de funcions, com els polinomis, es poden definir mitjançant "coordenades", la qual cosa facilita la tasca.

Si considerem el conjunt de tots els polinomis de grau n -èsim, un polinomi concret es pot escriure d'aquesta manera:

$$p(x) = c_0 + c_1x + \dots + c_nx^n,$$

i està completament determinat per tots els seus coeficients c_0, c_1, \dots, c_n . Així doncs, el polinomi $p(x)$ es pot escriure com a $(c_0, c_1, \dots, c_n)^T$. Aquest vector té $n + 1$ components i, per tant, està en un espai de dimensió $n + 1$.

Paràbola

Per exemple, la paràbola

$$y = x^2$$

es correspon amb un polinomi de grau $n = 2$ amb els coeficients $c_0 = 0, c_1 = 0$ i $c_2 = 1$. Així doncs, el vector associat és $(0,0,1)$ en un espai de dimensió 3.

Font de l'apartat

Apartat basat en l'obra següent: **Germund Dahlquist; Ake Björck** (2003). *Numerical Methods*. Dover Publications, Inc.

Mesura

Definim *mesura* com una funció que assigna un valor positiu (magnitud) a una realitat observable modelitzada com a objecte matemàtic. Per exemple, la longitud, l'àrea i el volum són tipus de mesures que fem regularment en la nostra vida diària.

En la guia corresponent a l'activitat 1 hem vist com es defineixen i es calculen les distàncies entre vectors, particularment utilitzant les **normes** vectorials. El concepte de *norma* també es pot utilitzar quan fem funcions en comptes de vectors.

Com s'ha especificat en la guia de l'activitat 1, totes les normes han de complir una sèrie de propietats que, adaptades al context de les funcions, es concreten a continuació.

Un funció $f : \mathbb{R}^n \rightarrow \mathbb{R}$ s'anomena **norma** si està definida en tot l'espai i compleix les condicions següents:

- 1) $\|f\| \geq 0$ per a totes les funcions f a l'espai.
- 2) $\|\alpha f\| = |\alpha| \cdot \|f\|$ per a tot escalar real α .
- 3) $\|f + g\| \leq \|f\| + \|g\|$ per a tot f i g a l'espai. És conegut com a **desigualtat triangular**.
- 4) $\|f\| = 0 \leftrightarrow f = 0$. En el cas particular d'espai de funcions contínues en l'interval $[a, b]$, aquesta condició significa que $f(x) = 0$ en tot l'interval $[a, b]$.

Les normes més comunament usades per a una funció f contínua en un interval tancat $[a, b]$ són:

- **Norma infinit:** $\|f\|_\infty = \max_{x \in [a, b]} |f(x)|$.
- **Norma euclidiana (L_2) o norma-2:** $\|f\|_2 = \sqrt{\int_a^b |f(x)|^2 dx}$.
- **Norma euclidiana ponderada:** $\|f\|_{2, w} = \sqrt{\int_a^b |f(x)|^2 w(x) dx}$, en què w és una funció de pes contínua i estrictament positiva en $[a, b]$.

Les normes euclidiana i infinit són casos especials, $p = 2$ i $p \rightarrow \infty$, respectivament, d'una família de normes anomenades **normes- p** , definides d'aquesta manera:

$$\|f\|_p = \sqrt[p]{\int_a^b |f(x)|^p dx}, \quad p \geq 1. \quad (2)$$

Quan ens trobem en el cas discret, la funció desconeguda és donada per una malla de punts, x_1, x_2, \dots, x_m , i les integrals es transformen en sumatoris. Per exemple, l'equivalent discret a la norma euclidiana de $f(\mathbf{x})$ s'expressaria així:

$$\|f(\mathbf{x})\| = \sqrt{\sum_{j=0}^m |f(x_j)|^2}.$$

Notació $x \rightarrow \infty$

Quan fem servir l'expressió $x \rightarrow \infty$ ens referim al fet que la variable x es fa molt gran, i per això es diu que x tendeix a infinit.

Tot i que es podria pensar en aquesta fórmula com una norma per a la funció f , estrictament parlant no ho és, ja que la condició 4 no es compleix. Aquesta expressió es coneix com a **seminorma** de f pel que fa a l'interval $[x_0, x_n]$. Farem servir la mateixa notació $\| \cdot \|$ per a normes i seminormes, i quan calgui farem servir subíndexs per distingir-les, per exemple, $\| \cdot \|_2$ i $\| \cdot \|_{2,w}$, respectivament.

Molts dels mètodes d'aproximació estan basats en el principi de minimitzar alguna norma o seminorma de l'error $\hat{f} - f$, en què \hat{f} aproxima f i té una forma preestablerta. És bastant important que la norma s'esculli tenint en compte la \hat{f} per la qual es vol fer servir.

1.1.2. Producte escalar

En aquest subapartat introduïm un formalisme relacionat amb les idees geomètriques esmentades anteriorment, que serà útil per estudiar l'aproximació de mínims quadrats.

El **producte escalar** de dues funcions, f i g , reals i contínues, denotat per $\langle f, g \rangle$, es defineix per la relació

$$\langle f, g \rangle = \begin{cases} \int_a^b f(x)g(x)w(x)dx, & \text{en cas continu,} \\ \sum_{i=0}^m f(x_i)g(x_i)w_i, & \text{en cas discret.} \end{cases}$$

En el cas discret, si tots els pesos $w_j = 1$, llavors $\langle f, g \rangle$ es correspon amb el producte escalar de vectors de $f(\mathbf{x})$ i $g(\mathbf{x})$ (amb la notació vectorial presentada en la introducció).

Com en el cas dels productes escalars que surten en l'àlgebra lineal, la definició genèrica de $\langle f, g \rangle$ ens permet derivar regles fonamentals associades al producte escalar de funcions. Com que f, g i ϕ són funcions, i α i β són nombres reals, tenim les propietats següents:

- **commutativitat:** $\langle f, g \rangle = \langle g, f \rangle$
- **linealitat:** $\langle \alpha f + \beta g, \phi \rangle = \alpha \langle f, \phi \rangle + \beta \langle g, \phi \rangle$
- **positivitat:** $\langle f, f \rangle \geq 0$

És fàcil deduir que $\langle f, f \rangle = \|f\|^2$, en què $\| \cdot \|$ denota la norma o seminorma euclidiana ponderada en els casos continu o discret, respectivament.

Producte escalar de vectors

Cal recordar que el producte escalar de vectors es correspon amb la suma dels productes (coordenada per coordenada) de cada vector. Per exemple, en dues dimensions $u \cdot v = (u_1, u_2) \cdot (v_1, v_2) = u_1 v_1 + u_2 v_2$.

1.1.3. Sistemes ortogonals

Dues funcions f i g s'anomenen **ortogonals** si el seu producte escalar és zero, és a dir, $\langle f, g \rangle = 0$. Una seqüència finita o infinita de funcions $\phi_0, \phi_1, \dots, \phi_n$ forma un **sistema ortogonal** si $\langle \phi_i, \phi_j \rangle = 0$ per a $i \neq j$ i $\|\phi_i\| \neq 0, \forall i$. Si, a més, $\|\phi_i\| = 1$, per a qualsevol valor de i , llavors la seqüència s'anomena **sistema ortonormal**.

Si $\langle f, g \rangle = 0$, és a dir, f i g són ortogonals, llavors es pot formular l'anomenat **teorema de Pitàgores de les funcions**:

$$\|f + g\|^2 = \|f\|^2 + \|g\|^2.$$

1.2. Formulació del problema

En aquest subapartat presentarem més concretament el problema que volem solucionar. Com hem dit en la introducció, pretenem aproximar una funció f desconeguda en l'interval $[a, b]$ per una altra \hat{f} . A més, la funció \hat{f} tindrà la forma de combinació lineal escrita com a

$$\hat{f}(x) = c_0\phi_0(x) + c_1\phi_1(x) + \dots + c_n\phi_n(x),$$

o, equivalentment,

$$\hat{f}(x) = \sum_{j=0}^n c_j\phi_j(x),$$

en què tenim $n + 1$ funcions $\phi_0(x), \phi_1(x), \dots, \phi_n(x)$ donades. Els coeficients c_0, c_1, \dots, c_n s'han d'obtenir de tal manera que una norma (o seminorma) euclidiana ponderada de la funció d'error $\hat{f} - f$ sigui el més petita possible, és a dir, s'han de calcular els coeficients de tal manera que

$$\|\hat{f} - f\|_{2,w}^2 = \int_a^b |\hat{f}(x) - f(x)|^2 w(x) dx, \quad \text{en el cas continu,}$$

$$\|\hat{f} - f\|_{2,w}^2 = \sum_{i=0}^m |\hat{f}(x_i) - f(x_i)|^2 w_i, \quad \text{en el cas discret,}$$

siguin tan petites com es pugui. Aquesta formulació es coneix com a **problema de mínims quadrats**.

1.3. Solució del problema d'aproximació

Es pot demostrar que, quan $\phi_0, \phi_1, \dots, \phi_n$ són linealment independents, el problema d'aproximació mitjançant mínims quadrats té una solució única:

$$\hat{f} = \sum_{j=0}^n \hat{c}_j \phi_j, \quad (3)$$

en què els coeficients \hat{c}_j són els que fan mínima la norma euclidiana.

A més, la solució està caracteritzada per la propietat d'ortogonalitat, que implica que $\hat{f} - f$ ha de ser ortogonal a totes les funcions ϕ_j . Així, si es compleix això, llavors \hat{f} és una solució del problema d'aproximació. Aplicant la definició d'ortogonalitat obtindríem que les $n + 1$ igualtats

$$\left\langle \sum_{j=0}^n \hat{c}_j \phi_j - f, \phi_k \right\rangle = 0, \quad k = 0, 1, 2, \dots, n,$$

s'han de complir. A més, de la propietat de linealitat del producte escalar es dedueix que

$$\left\langle \sum_{j=0}^n c_j \phi_j, \phi_k \right\rangle = \sum_{j=0}^n c_j \langle \phi_j, \phi_k \rangle.$$

Llavors, les condicions anteriors són equivalents a resoldre el sistema d'equacions

$$\begin{aligned} \langle \phi_0, \phi_0 \rangle \hat{c}_0 + \langle \phi_0, \phi_1 \rangle \hat{c}_1 + \dots + \langle \phi_0, \phi_n \rangle \hat{c}_n &= \langle \phi_0, f \rangle \\ \langle \phi_1, \phi_0 \rangle \hat{c}_0 + \langle \phi_1, \phi_1 \rangle \hat{c}_1 + \dots + \langle \phi_1, \phi_n \rangle \hat{c}_n &= \langle \phi_1, f \rangle \\ &\vdots \\ \langle \phi_n, \phi_0 \rangle \hat{c}_0 + \langle \phi_n, \phi_1 \rangle \hat{c}_1 + \dots + \langle \phi_n, \phi_n \rangle \hat{c}_n &= \langle \phi_n, f \rangle, \end{aligned}$$

que, formulat de manera compacta, és

$$\sum_{j=0}^n \langle \phi_j, \phi_k \rangle \hat{c}_j = \langle f, \phi_k \rangle, \quad k = 0, 1, 2, \dots, n. \quad (4)$$

Aquest sistema d'equacions s'anomena **equació normal**. En resoldre el sistema es determinen els coeficients \hat{c}_j . En la guia de l'activitat 1 s'han presentat diferents mètodes de resolució de sistemes.

Si ens trobem en el cas discret i la norma utilitzada és la norma euclidiana, és a dir, si els pesos són tots com la unitat $w_j = 1$, les equacions normals es poden obtenir directament sense calcular un per un tots els productes escalars necessaris. Recordem el sistema d'equacions (1) que hem vist a la introducció i denotem

$$\Phi := (\phi_0(\mathbf{x}) | \phi_1(\mathbf{x}) | \dots | \phi_n(\mathbf{x})) \quad \text{y} \quad \mathbf{c} := (c_0, c_1, \dots, c_n)^T.$$

Així doncs, ara el que volem fer és resoldre el sistema sobredeterminat $\Phi \mathbf{c} = f(\mathbf{x})$. És fàcil demostrar que podem construir les equacions normals corresponents a aquest sistema pel que fa a la norma euclidiana simplement fent la transformació següent:

$$\Phi^T \Phi \mathbf{c} = \Phi^T f(\mathbf{x}). \tag{5}$$

Un cas particular molt important és quan $\phi_0, \phi_1, \dots, \phi_n$ formen un sistema ortogonal. Llavors, les equacions normals es poden resoldre immediatament, ja que en cada equació els termes en què $j \neq k$ són zero. Tenint en compte això, cada equació normal es redueix a

$$\langle \phi_j, \phi_j \rangle \hat{c}_j = \langle f, \phi_j \rangle, \quad j = 0, 1, 2, \dots, n.$$

Si els aïllem, els coeficients es poden calcular d'una manera molt senzilla mitjançant l'expressió

$$\hat{c}_j = \frac{\langle f, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle}, \quad j = 0, 1, 2, \dots, n. \tag{6}$$

Els coeficients \hat{c}_j reben el nom de **coeficients ortogonals** (a vegades, coeficients de Fourier).

Equacions normals

Suposem que volem fer servir el mètode de mínims quadrats per ajustar una funció $\hat{f}(x) = c_0 + c_1 x$ a unes dades mesurades en un experiment i presentades en la taula 1. Assumim que tots els pesos són la unitat.

Taula 1

x	$f(x)$
1	-2,1
3	-0,9
4	-0,6
6	0,6
7	0,9

Dades experimentals

Fent servir la notació anterior tindriem que $\phi_0(x) = 1$ $\phi_1(x) = x$, amb $m = 5$. Així doncs, les equacions normals són donades per:

$$\langle \phi_0, \phi_0 \rangle \hat{c}_0 + \langle \phi_1, \phi_0 \rangle \hat{c}_1 = \langle f, \phi_0 \rangle$$

$$\langle \phi_0, \phi_1 \rangle \hat{c}_0 + \langle \phi_1, \phi_1 \rangle \hat{c}_1 = \langle f, \phi_1 \rangle.$$

Es necessiten els productes escalars dels vectors:

$$\langle \phi_0, \phi_0 \rangle = \sum_{i=1}^5 \phi_0(x_i) \phi_0(x_i) = \sum_{i=1}^5 1 = 5,$$

$$\langle \phi_0, \phi_1 \rangle = \sum_{i=1}^5 \phi_0(x_i) \phi_1(x_i) = \sum_{i=1}^5 x_i = 21,$$

$$\langle \phi_1, \phi_0 \rangle = \sum_{i=1}^5 \phi_1(x_i) \phi_0(x_i) = \sum_{i=1}^5 x_i = 21,$$

$$\langle \phi_1, \phi_1 \rangle = \sum_{i=1}^5 \phi_1(x_i) \phi_1(x_i) = \sum_{i=1}^5 x_i^2 = 111,$$

$$\langle f, \phi_0 \rangle = \sum_{i=1}^5 f(x_i) \phi_0(x_i) = \sum_{i=1}^5 f(x_i) = -2,1,$$

$$\langle f, \phi_1 \rangle = \sum_{i=1}^5 f(x_i) \phi_1(x_i) = \sum_{i=1}^5 f(x_i) x_i = 2,7,$$

Així doncs, s'obté aquest sistema:

$$5\hat{c}_0 + 21\hat{c}_1 = -2,1$$

$$21\hat{c}_0 + 111\hat{c}_1 = 2,7,$$

la solució del qual és $\hat{c}_0 = -2,542$ i $\hat{c}_1 = 0,5053$. El sistema resultant es podria haver obtingut directament mitjançant la transformació (5).

Mitjana com a aproximació de funcions

Considerem el cas $n = 0$, $\phi_0(x) = 1$. Recordem que

$$\langle f, g \rangle = \sum_{i=0}^m w_i f(x_i) g(x_i).$$

Les "equacions normals" en l'equació (4) es redueixen a una sola equació,

$$\langle \phi_0, \phi_0 \rangle c_0 = \langle f, \phi_0 \rangle,$$

en què

$$c_0 = \frac{\sum_{i=0}^m w_i f(x_i)}{\sum_{i=0}^m w_i}.$$

El coeficient c_0 és la **mitjana ponderada** dels valors de la funció. Si tots els pesos són iguals, llavors obtenim

$$c_0 = \frac{\sum_{i=0}^m f(x_i)}{m+1},$$

amb la qual cosa veiem que el càlcul de la mitjana és un cas especial del mètode de mínims quadrats.

És important destacar que, quan $\phi_1, \phi_2, \dots, \phi_n$ formen un sistema ortogonal, el coeficient de Fourier \hat{c}_j és independent de n (consulteu l'equació (6) per veure-ho). L'avantatge d'això és que es pot incrementar el nombre total de paràmetres sense recalculer-ne cap dels prèviament obtinguts. L'ús de sistemes ortogonals és molt avantatjós, no només perquè simplifiquen els càlculs, sinó també perquè permeten evitar problemes numèrics que poden sorgir quan es resolen sistemes d'equacions normals per a un conjunt de funcions base no ortogonals.

1.4. Aproximació per polinomis ortogonals

Una família de polinomis ortogonals és una família triangular (que definirem a continuació) de polinomis que és un sistema ortogonal segons una funció pes donada. Les **aproximacions de funcions en termes de polinomis ortogonals** són molt útils: són **fàcils de manipular**, tenen **bones propietats de convergència** i proporcionen una **representació ben condicionada** d'una funció. La teoria dels polinomis ortogonals també constitueix la base de mètodes numèrics per a molts problemes que, en principi, pot semblar que tenen poca relació amb l'aproximació de mínims quadrats (integració numèrica, fraccions contínues o el problema algebraic d'autovalors, per exemple).

Nombre de condició

El nombre de condició d'una funció mesura el grau d'afectació d'un grau de sortida davant un canvi important en els valors d'entrada. Si el valor del nombre de condició es troba a prop d'1, es diu que la funció està **ben condicionada**. Si, per contra, el nombre de condició és molt més gran que 1, es diu que està **mal condicionada**.

1.4.1. Família triangular de polinomis

Un polinomi de grau n és una funció que presenta aquesta forma:

$$p(x) = c_n x^n + c_{n-1} x^{n-1} + \dots + c_1 x + c_0.$$

El coeficient c_n s'anomena *primer coeficient*. Si $c_n \neq 0$, llavors el polinomi es denomina *polinomi genuí de grau n* . La classe de polinomi de grau n conté tots els polinomis d'un grau més petit com a casos especials. Una constant és un polinomi de grau zero.

Hi ha moltes maneres d'especificar un polinomi. Una d'aquestes és escriure els termes en forma de sumatori, també anomenat *expansió*, amb la qual cosa obtenim

$$\sum_{j=0}^n c_j x^j.$$

Tot i que la forma d'aquest sumatori pot semblar la més intuïtiva, no és l'única alternativa. Per exemple, podríem reescriure la forma anterior d'aquesta manera:

$$\sum_{j=0}^n d_j (x-h)^j,$$

per a $h \neq 0$.

Matemàticament, les expressions anteriors són equivalents, però, computacionalment, el fet de treballar amb els **valors arrodonits** dels coeficients pot suposar una gran diferència. Si es volgués emprar la segona forma per a una aproximació polinòmica en l'interval $[a, b]$, llavors s'hauria de seleccionar $h = \frac{1}{2}(a+b)$, és a dir, el punt mitjà de l'interval.

Les dues formes presentades són les més naturals, i també les primeres en què pensariem, però hi ha moltes altres representacions que són fins i tot més avantatjoses. Pensem en la seqüència de polinomis $\phi_0, \phi_1, \phi_2, \dots$ (finita o infinita)

$$\begin{aligned} \phi_0(x) &= c_{00} \\ \phi_1(x) &= c_{10} + c_{11}x \\ \phi_2(x) &= c_{20} + c_{21}x + c_{22}x^2 \\ &\vdots \\ \phi_n(x) &= c_{n0} + c_{n1}x + c_{n2}x^2 + \dots + c_{nn}x^n \\ &\vdots \end{aligned}$$

en què tots els primers coeficients són diferents de zero, és a dir, $c_{ii} \neq 0$. Aquesta representació es coneix com a **família triangular** de polinomis. Com veurem, un polinomi escrit com a família triangular en facilita el tractament i l'avaluació.

Utilitzant la forma triangular, les potències de x poden ser expressades recursivament i únicament com a combinació lineal de $\phi_0, \phi_1, \phi_2, \dots$,

$$x^0 = \frac{\phi_0(x)}{c_{00}}, \quad x^1 = \frac{\phi_1(x) - c_{10}}{c_{11}} = \frac{\phi_1(x) - c_{10} \frac{\phi_0(x)}{c_{00}}}{c_{11}}, \dots$$

1.4.2. Polinomis ortogonals

En aquest subapartat definirem formalment els polinomis ortogonals. Als subapartats següents mostrarem alguns exemples representatius de polinomis ortogonals, com ara els **polinomis de Txebixev** i els **polinomis de Legendre**.

Sigui quina sigui la distribució de pesos per al producte escalar, té un sistema ortogonal associat $\phi_0, \phi_1, \phi_2, \dots$, que és una família triangular de polinomis. La família està determinada de manera única, independentment que els primers coeficients $A_0 = c_{00}, A_1 = c_{11}, A_2 = c_{22}, \dots$ dels polinomis de la família triangular puguin ser donats per valors arbitraris diferents de zero.

Donada una distribució de pesos en una malla amb $m + 1$ punts, la família acaba amb $\phi_m(x)$. La funció $\phi_{m+1}(x)$ és zero en cada punt de la malla. En el cas continu, la família té membres infinits.

Per a $n \geq 0$, els polinomis ortogonals satisfan un fórmula recursiva de tres termes:

$$\phi_{n+1}(x) = \alpha_n(x - \beta_n)\phi_n(x) - \gamma_n\phi_{n-1}(x), \quad \phi_{-1}(x) = 0, \quad \phi_0(x) = A_0, \quad (7)$$

en què $\alpha_n = \frac{A_{n+1}}{A_n}$, β_n i γ_n són donades per les expressions

$$\beta_n = \frac{\langle x\phi_n, \phi_n \rangle}{\|\phi_n\|^2}, \quad \gamma_n = \frac{\alpha_n \|\phi_n\|^2}{\alpha_{n-1} \|\phi_{n-1}\|^2}.$$

En el càlcul dels coeficients en una expansió de la forma

$$p(x) = c_0\phi_0(x) + c_1\phi_1(x) + \dots + c_n\phi_n(x), \quad (8)$$

utilitzant la fórmula del coeficient ortogonal $c_j = \langle p, \phi_j \rangle / \|\phi_j\|^2$, es pot fer servir la fórmula recursiva de l'equació 7. En el cas discret, és aconsellable calcular recursivament els vectors $\phi_j(\mathbf{x}), j = 0, 1, 2, \dots, n$.

La manera més fàcil d'avaluar una funció definida mitjançant una expansió ortogonal és amb la **fórmula recursiva de Clenshaw**. Si fem servir la notació de l'equació 7 tenim que, per a $k = n, n - 1, \dots, 1, 0$,

$$y_k = \alpha_k(x - \beta_k)y_{k+1} - \gamma_{k+1}y_{k+2} + c_k,$$

en què $y_{n+2} = y_{n+1} = 0$. Finalment obtenim que $p(x) = A_0y_0$.

1.4.3. Polinomis de Txebixev

Els polinomis de Txebixev constitueixen una de les famílies triangulars de polinomis ortogonals més utilitzades. Les seves propietats poden ser deduïdes amb força facilitat, fins i tot sense la teoria enunciat al subapartat anterior.

Considerem la fórmula

$$\cos((n+1)\phi) + \cos(n-1)\phi = 2 \cos(\phi) \cos(n\phi), \quad n \geq 1.$$

Aquesta fórmula ens permet expressar $\cos(n\phi)$ com un polinomi en termes de $\cos(\phi)$,

$$\cos(2\phi) = 2 \cos^2(\phi) - 1$$

$$\cos(3\phi) = 2 \cos(\phi) \cos(2\phi) - \cos(\phi) = 4 \cos^3(\phi) - 3 \cos(\phi)$$

$$\cos(4\phi) = 2 \cos(\phi) \cos(3\phi) - \cos(2\phi) = 8 \cos^4(\phi) - 8 \cos^2(\phi) + 1$$

⋮

Si assumim que $x = \cos(\phi)$, on $\phi = \arccos(x)$, llavors obtenim una família triangular de polinomis, els polinomis de Txebixev, per $-1 \leq x \leq 1$, $n = 0, 1, 2, \dots$, definits per la fórmula

$$T_n(x) = \cos(n \arccos(x)).$$

Així, utilitzant les expressions anteriors per a $\cos(n\phi)$, obtenim

$$T_0(x) = 1,$$

$$T_1(x) = x,$$

$$T_2(x) = 2x^2 - 1,$$

$$T_3(x) = 4x^3 - 3x,$$

$$T_4(x) = 8x^4 - 8x^2 + 1$$

⋮

Propietats dels polinomis de Txebixev

1) Fórmula recursiva:

$$T_0(x) = 1,$$

$$T_1(x) = x,$$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n \geq 1.$$

2) Primer coeficient: és 2^{n-1} per a $n \geq 1$ i 1 per a $n = 0$.

3) Simetria:

$$T_n(-x) = (-1)^n T_n(x).$$

4) Zeros i punts crítics: $T_n(x)$ té n zeros en l'interval $[-1,1]$, que són donats per

$$x_k = \cos\left(\frac{2k+1}{n} \frac{\pi}{2}\right), \quad k = 0, 1, 2, \dots, n-1,$$

coneguts com a **abscisses de Txebixev**, i $n+1$ punts crítics,

$$x'_k = \cos\left(\frac{k\pi}{n}\right), \quad T_n(x'_k) = (-1)^k, \quad k = 0, 1, 2, \dots, n.$$

Aquests resultats es dedueixen directament pel fet que $\cos(n\phi) = 0$ per a

$$\phi = \frac{2k-1}{n} \frac{\pi}{2},$$

i que $|\cos(n\phi)|$ té màxims en $\phi = k\pi/n$.

5) Ortogonalitat, cas continu: els polinomis de Txebixev són ortogonals segons el producte escalar

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x)(1-x^2)^{-\frac{1}{2}} dx.$$

Així doncs,

$$\langle T_i, T_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ \frac{1}{2}\pi & \text{si } i = j \neq 0 \\ \pi & \text{si } i = j = 0 \end{cases}$$

6) Propietat d'ortogonalitat, cas discret: si

$$\langle f, g \rangle = \sum_{k=0}^m f(x_k)g(x_k),$$

on $\{x_k\}$ són els zeros de $T_{m+1}(x)$, llavors per a $0 \leq i \leq m$ i $0 \leq j \leq m$ tenim

$$\langle T_i, T_j \rangle = \begin{cases} 0 & \text{si } i \neq j \\ \frac{1}{2}(m+1) & \text{si } i = j \neq 0 \\ m+1 & \text{si } i = j = 0 \end{cases}$$

7) Propietat minimax: de tots els polinomis de grau n amb primer coeficient 1, el polinomi $2^{1-n}T_n$ té la norma del màxim més petita en $[-1,1]$. El valor de la seva norma del màxim és 2^{1-n} .

Les expansions en termes de polinomis de Txebixev són una ajuda important en l'estudi de funcions en l'interval $[-1,1]$. Si volem treballar en termes d'una variable t que viu en l'interval $[a, b]$, llavors s'ha de fer el canvi de variables (substitució) següent:

$$t = \frac{1}{2}(a+b) + \frac{1}{2}(b-a)x, \quad t \in [a,b] \leftrightarrow x \in [-1,1].$$

1.4.4. Polinomis de Legendre

Els polinomis de Legendre es defineixen mitjançant la fórmula

$$L_0(x) = 1,$$

$$L_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left((x^2 - 1)^n \right).$$

(9)

Nota
La derivada n -èsima d'una funció, que descrivim $f(x)d^n/dx^n(f(x))$ és la derivada de la derivada de la derivada... (n vegades).

Propietats dels polinomis de Legendre

1) Fórmula recursiva:

$$\begin{aligned}
 L_0(x) &= 1, \\
 L_1(x) &= x, \\
 L_2(x) &= \frac{1}{2}(3x^2 - 1), \\
 L_{n+1}(x) &= \frac{2n+1}{n+1}xL_n(x) - \frac{n}{n+1}L_{n-1}(x).
 \end{aligned}
 \tag{10}$$

2) Primer coeficient: com que $(x^2-1)^n$ és un polinomi de grau $2n$, $L_n(x)$ és un polinomi de grau n . El primer coeficient A_n és el mateix que el del polinomi

$$\frac{1}{2^n n!} \frac{d^n}{dx^n} (x^{2n}) = \frac{1}{2^n n!} 2n(2n-1)(2n-2) \dots (n+1)x^n.$$

Així doncs,

$$\alpha_n = \frac{A_{n+1}}{A_n} = \frac{2n+1}{n+1}.$$

3) Simetria:

$$L_n(-x) = (-1)^n L_n(x).$$

4) Ortogonalitat: els polinomis de Legendre són ortogonals pel que fa al producte escalar,

$$\langle f, g \rangle = \int_{-1}^1 f(x)g(x)dx,$$

per la qual cosa,

$$\langle L_n, L_j \rangle = \begin{cases} 0 & \text{si } n \neq j \\ \frac{2}{2n+1} & \text{si } n = j \end{cases}$$

5) Cota:

$$|L_n(x)| \leq 1, \quad x \in [-1, 1].$$

Si observem la propietat d'ortogonalitat, els polinomis de Legendre són força útils per a aproximacions amb norma euclidiana i distribució de pes $w(x) = 1$.

1.5. Mètodes de Fourier

En aquest subapartat estudiarem una aplicació particular de la teoria d'aproximació de funcions que hem vist fins ara. Es tracta d'un marc teòric altament desenvolupat, l'**anàlisi de Fourier**, en el qual s'emmarquen els anomenats *polinomis trigonomètrics*. Aquests són una eina amb molt potencial en l'aproximació de funcions, especialment pel que fa a les **funcions periòdiques** (però no només aquestes). L'anàlisi de Fourier utilitza molts dels resultats generals dels polinomis ortogonals.

Molts fenòmens naturals (acústics, òptics, etc.) presenten un comportament periòdic. És ben sabut que un so està compost per oscil·lacions regulars: el to fonamental (o pur), amb una freqüència n , i els tons harmònics, amb freqüències $2n, 3n, 4n, \dots$. La ràtio entre el to fonamental i els tons harmònics determina la nostra percepció del so. També hi pot haver sons sense tons harmònics, és a dir, basats només en el to fonamental, per exemple, en la música electrònica, en la qual es fan servir tons sinusoidals purs.

En un oscil·lador electrònic, un corrent elèctric es genera amb una força que depèn del temps t , d'acord amb la fórmula

$$r \sin(\omega t + \nu),$$

en què r és l'anomenada *amplitud d'oscil·lació*, ω és la freqüència angular (2π vegades la freqüència) i ν és una constant que defineix l'estat del so a temps $t = 0$.

En un altaveu, les variacions de corrent elèctric es converteixen en variacions de pressió de l'aire, que, en condicions ideals, es descriuen amb la mateixa funció. No obstant això, en la pràctica sempre es produeixen certes distorsions que provoquen que apareguin els tons harmònics. Així, els harmònics d'ordre $k-1$ contribueixen al so en la forma $r_k \sin(k\omega t + \nu_k)$. Les variacions en la pressió de l'aire que ens arriben a l'oïda es poden descriure, des d'aquest punt de vista, com aquesta suma:

$$\sum_{k=0}^{\infty} r_k \sin(k\omega t + \nu_k). \quad (11)$$

La separació d'un fenomen periòdic en un to fonamental i diversos tons harmònics no només apareix en l'àmbit acústic, sinó també en moltes altres àrees, i està relacionat amb un teorema purament matemàtic molt important desenvolupat per primera vegada per Fourier. D'acord amb aquest teorema, tota funció $f(t)$ amb període $2\pi/\omega$, en certes condicions generals, es pot expandir en una sèrie de la forma del sumatori anterior (equació (11)). Més endavant donarem una formulació més precisa d'aquest resultat.

Una funció periòdica és la que compleix que

$$f(t+p) = f(t).$$

Una expansió de la forma de l'equació (11) es pot expressar de moltes maneres equivalents. Si prenem $a_k = r_k \sin(v_k)$ i $b_k = r_k \cos(v_k)$, per les propietats del si d'una suma, podem escriure

$$f(t) = \sum_{k=0}^{\infty} (a_k \cos(k\omega t) + b_k \sin(k\omega t)), \quad (12)$$

en què a_k i b_k són constants reals. En la pràctica, el sumatori s'ha de *truncar*, és a dir, utilitzar un nombre finit de sumands, $n+1$,

$$f(t) = \sum_{k=0}^n (a_k \cos(k\omega t) + b_k \sin(k\omega t)).$$

Una altra manera de fer-ho que també pot resultar convenient es pot obtenir mitjançant la fórmula d'Euler:

$$\cos(x) = \frac{\exp(ix) + \exp(-ix)}{2}, \quad \sin(x) = \frac{\exp(ix) - \exp(-ix)}{2i},$$

en què $i = \sqrt{-1}$ és la unitat imaginària. Si per a $k = 1, 2, 3, \dots$ tenim

$$\begin{aligned} c_k &= \frac{1}{2} (a_k - ib_k), \\ c_{-k} &= \frac{1}{2} (a_k + ib_k), \end{aligned} \quad (13)$$

llavors s'obté

$$f(t) = \sum_{k=-\infty}^{\infty} c_k \exp(ik\omega t), \quad (14)$$

o, utilitzant el sumatori truncat,

$$f(t) = \sum_{k=-n}^n c_k \exp(ik\omega t).$$

Pel que fa a la resta utilitzarem el terme **sèrie de Fourier** a una expansió de la forma de l'equació (12) o l'equació (14). Les formes truncades d'aquestes sèries es coneixen com a **polinomis trigonomètrics**. Les sèries de Fourier són molt importants en l'estudi de fenòmens periòdics en el temps (vibracions, so, llum, corrent altern, etc.) o en l'espai (ones, estructura cristal·lina, etc.).

1.5.1. Conceptes bàsics de l'anàlisi de Fourier

En aquest subapartat ens centrarem en les funcions d'una variable i període 2π . Si una funció de t té període p , llavors el canvi de variables $x = 2\pi t/p$ transforma la funció de t inicial en una funció de x amb període 2π . Assumim que la funció pot tenir valors complexos, ja que l'exponencial complexa resulta molt convenient per a certes manipulacions algebraiques. El producte escalar de dues funcions que prenen valors complexos, f i g de període 2π , es defineix, en el cas continu, d'aquesta manera:

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)\bar{g}(x)dx, \quad (15)$$

i en el cas discret, d'aquesta altra:

$$\langle f, g \rangle = \sum_{\alpha=0}^M f(x_{\alpha})\bar{g}(x_{\alpha}), \quad x_{\alpha} = \frac{2\pi\alpha}{M+1}, \quad (16)$$

en què la barra horitzontal damunt de g indica conjugació complexa.

La norma de la funció f es defineix així:

$$\|f\| = \sqrt{\langle f, f \rangle}.$$

Als resultats per al producte escalar que s'han vist en apartats anteriors també se'ls aplica aquesta nova definició, amb algunes modificacions òbvies. Observeu especialment que $\langle g, f \rangle = \overline{\langle f, g \rangle}$. En el cas continu, és igual en quin interval ens trobem; sempre que tingui longitud 2π , el valor del producte escalar no canvia.

Si les funcions (que surten en l'equació (14)) són

$$\phi_j(x) = \exp(ijx), \quad j = \dots, -2, -1, 0, 1, 2, \dots,$$

es tenen les relacions d'ortogonalitat següents:

$$\langle \phi_j, \phi_k \rangle = \begin{cases} 2\pi & \text{si } j = k, \\ 0 & \text{si } j \neq k, \end{cases}$$

per al cas continu

$$\langle \phi_j, \phi_k \rangle = \begin{cases} M+1 & \text{si } \frac{j-k}{M+1} \text{ és un enter} \\ 0 & \text{en un altre cas,} \end{cases}$$

i per al cas discret.

Si se sap que f es pot expandir d'aquesta manera:

$$f = \sum_{j=a}^b c_j \phi_j,$$

en què $a = -\infty$ i $b = \infty$ en el cas continu i $b - a = M$ en el discret, llavors, formalment, per a $a \leq k \leq b$ tenim que

$$\langle f, \phi_k \rangle = \sum_{j=a}^b c_j \langle \phi_j, \phi_k \rangle,$$

ja que $\langle \phi_j, \phi_k \rangle = 0$ per a $j \neq k$. Així, canviant k per j , tenim

$$c_j = \frac{\langle f, \phi_j \rangle}{\langle \phi_j, \phi_j \rangle} = \begin{cases} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \exp(-ijx) dx & \text{en el cas continu} \\ \frac{1}{M+1} \sum_{\alpha=0}^M f(x_\alpha) \exp(-ijx_\alpha) & \text{en el cas discret.} \end{cases} \quad (17)$$

Aquests coeficients s'anomenen **coeficients de Fourier** i són un cas particular dels obtinguts al subapartat 1.3.

D'acord amb l'equació (13),

$$a_j = c_j + c_{-j}, \quad b_j = i(c_j - c_{-j}).$$

Quan $j = 0$, el coeficient a_0 es calcularia sumant dues vegades el mateix valor c_0 , la qual cosa seria incorrecta. Per tant, el coeficient a_0 apareix sempre dividit per 2 per corregir aquesta discrepància. Com que es pot deduir fàcilment, $b_0 = 0$.

Per exemple, per al cas continu tindriem que, per a $j > 0$

$$\begin{aligned}
 a_j &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(jx) dx, \\
 b_j &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(jx) dx.
 \end{aligned}
 \tag{18}$$

Dades periòdiques

A la taula 2 presentem les temperatures registrades a Washington D. C. del dia 1 de gener del 2001. Triem un model que permeti capturar les oscil·lacions que observem en les dades.

Taula 2

Hora del dia	Fracció de dia, (t)	Temperatura
00:00	0	-2.2
03:00	$\frac{1}{8}$	-2,8
06:00	$\frac{1}{4}$	-6,1
09:00	$\frac{3}{8}$	-3,9
12:00	$\frac{1}{2}$	0,0
15:00	$\frac{5}{8}$	1,1
18:00	$\frac{3}{4}$	-0,6
21:00	$\frac{7}{8}$	-1,1

D'acord amb el que hem vist en aquest subapartat, el més senzill és $f(t) = \frac{a_0}{2} + a_1 \cos(2\pi t) + b_1 \sin(2\pi t)$, que es correspon amb l'expressió (12) truncada en $n = 1$. Fem servir aquest model perquè sabem que la temperatura és més o menys periòdica amb un període de 24 hores. Utilitzant la notació de l'apartat tindriem $M = 7$ (8 dades) i $\alpha = 0, 1, \dots, 7$ (segona columna). Per tant, $x_\alpha = 2\pi t$ i $f(x_\alpha)$ són les temperatures corresponents (tercera columna). Si apliquem ara les expressions dels coeficients, c_j , per al cas discret obtenim

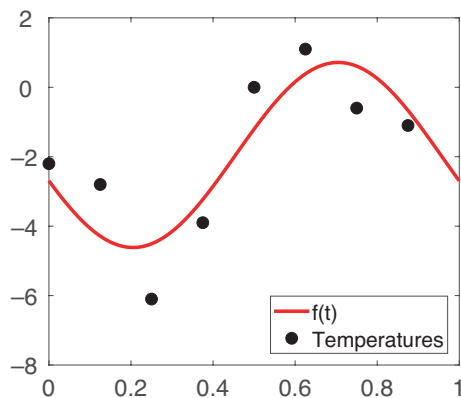
$$c_0 = -1.95, \quad c_{-1} = -0,37 - 1,28i, \quad c_1 = -0,37 + 1,28i.$$

Ara podem calcular els coeficients a_0 , a_1 i b_1 com a

$$\begin{aligned}
 a_0 &= c_0 + c_0 = -3,9, \quad \frac{a_0}{2} = -1,95, \\
 a_1 &= c_1 + c_{-1} = -0,74, \\
 b_1 &= i(c_1 - c_{-1}) = -2,56.
 \end{aligned}$$

En la figura 4 presentem l'aproximació obtinguda.

Figura 4



1.5.2. Sèries de Fourier

Tota funció f contínua i fragmentada amb període 2π es pot associar amb una sèrie de Fourier de les dues formes que hem vist anteriorment:

$$\frac{1}{2}a_0 + \sum_{j=1}^{\infty} (a_j \cos(jx) + b_j \sin(jx)) \quad \text{o} \quad \sum_{j=-\infty}^{\infty} c_j \exp(ijx).$$

Els coeficients c_j , a_j i b_j es poden calcular mitjançant les equacions (17) i (18), respectivament. Si f i la seva primera derivada, f' , són contínues en tot el domini, llavors la sèrie convergeix en tot el domini en $f(x)$. Es permet fins i tot que f i f' tinguin un nombre finit de discontinuïtats en cada període. En aquestes discontinuïtats, la sèrie dona la mitjana dels valors límit per la dreta i per l'esquerra en el punt estudiat. La suma parcial o truncada de la sèrie és la millor aproximació possible a $f(x)$, en el sentit dels mínims quadrats, mitjançant l'ús de polinomis trigonomètrics.

Tenint en compte l'equació (18), és fàcil deduir les afirmacions següents:

- Si f és una funció parell, és a dir, $f(x) = f(-x)$, per a qualsevol x , llavors $b_j = 0$, per a tot j . Aleshores, la sèrie de Fourier és una sèrie de cosinus.
- Si f és una funció imparell, és a dir, $f(x) = -f(-x)$, per a qualsevol x , llavors $a_j = 0$, per a tot j . Aleshores, la sèrie de Fourier és una sèrie de sinus.

Expansió de Fourier de l'ona quadrada

Considerem aquesta funció d'ona quadrada (figura 5)

$$f(x) = \begin{cases} -1 & \text{si } -\pi < x < 0, \\ 1 & \text{si } 0 < x < \pi. \end{cases}$$

Imaginem que podem fer una continuació periòdica de $f(x)$ fora de l'interval $(-\pi, \pi)$. Així doncs, f és una funció imparell i ja sabem que tots els coeficients de cosinus són zero, $a_j = 0$. Els coeficients de sinus, b_j , es calculen així:

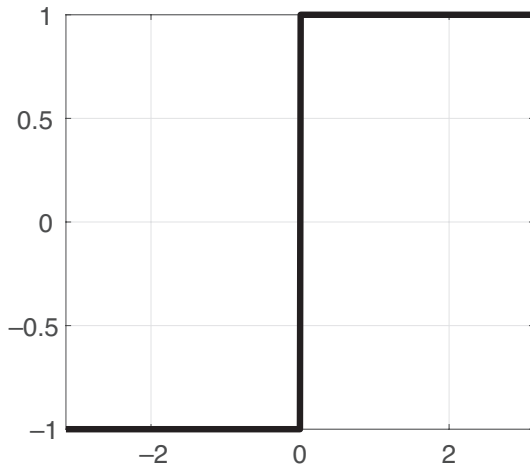
$$\begin{aligned} b_j &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(jx) dx = -\frac{1}{\pi} \int_{-\pi}^0 \sin(jx) dx + \frac{1}{\pi} \int_0^{\pi} \sin(jx) dx \\ &= \frac{2}{\pi} \int_0^{\pi} \sin(jx) dx = \frac{2}{\pi} \frac{1 - \cos(j\pi)}{j} = \begin{cases} 0 & \text{si } j \text{ és parell,} \\ \frac{4}{j\pi} & \text{si } j \text{ és imparell.} \end{cases} \end{aligned}$$

Així doncs,

$$f(x) = \frac{4}{\pi} \left(\sin(x) + \frac{\sin(3x)}{3} + \frac{\sin(5x)}{5} + \dots \right).$$

La suma de la sèrie és zero en els punts en què f presenta una discontinuïtat. Això concorda amb el fet que la suma ha de ser igual que la mitjana dels valors límit per la dreta i l'esquerra en la discontinuïtat. En la figura 6 presentem l'aproximació obtinguda en truncar la sèrie infinita en n termes, considerant diferents valors de n . Observem que, com més gran és n , l'aproximació clarament millora, és a dir, convergeix en f .

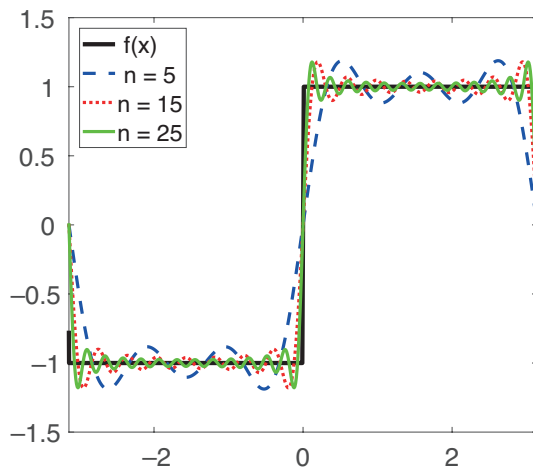
Figura 5



1.5.3. Teorema integral de Fourier

Imaginem que la funció $\psi(\xi)$ definida en tot el domini real i que satisfà les propietats de regularitat requerides perquè les sèries de Fourier siguin convergents és

Figura 6



Siguin

$$x = \frac{2\pi\xi}{L}, \quad \psi(\xi) = f(x).$$

Si tenim

$$c_j = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \exp(-ijx) dx = \frac{1}{L} \int_{-\frac{1}{2}L}^{\frac{1}{2}L} \psi(\xi) \exp(-2\pi i \xi j/L) d\xi,$$

llavors, considerant les sèries de Fourier per a $x \in (-\pi, \pi)$ o, equivalentment, $\xi \in (-\frac{1}{2}L, \frac{1}{2}L)$, resulta

$$f(x) = \sum_{j=-\infty}^{\infty} c_j \exp(ijx), \quad \psi(\xi) = \sum_{j=-\infty}^{\infty} c_j \exp(2\pi i \xi j/L).$$

Si ara tenim que

$$g_L(t) = \frac{1}{L} \int_{-\frac{1}{2}L}^{\frac{1}{2}L} \psi(\xi) \exp(-2\pi i \xi t) d\xi, \quad (19)$$

llavors obtenim

$$\psi(\xi) = \frac{1}{L} \sum_{j=-\infty}^{\infty} g_L\left(\frac{j}{L}\right) \exp(2\pi i \xi j/L), \quad \xi \in \left(-\frac{1}{2}L, \frac{1}{2}L\right). \quad (20)$$

Passant pel límit, és a dir, $L \rightarrow \infty$, s'evita fer una continuació periòdica artificial de la funció fora de l'interval finit. El sumatori en l'equació (20) és una suma de rectangles, semblant a la que apareix en la definició d'integral definida. No obstant això, aquí l'argument varia entre $-\infty$ i ∞ , i la funció $g_L(t)$ depèn de L . Mitjançant un pas al límit no gaire ortodox, el parell de fórmules en les equacions (19) i (20) es converteixen en

$$g(t) = \int_{-\infty}^{\infty} \psi(\xi) \exp(-2\pi i \xi t) d\xi \iff \psi(t) = \int_{-\infty}^{\infty} g(\xi) \exp(2\pi i \xi t) dt. \quad (21)$$

De fet, seguint una anàlisi matemàtica força complicada, es pot demostrar que el resultat anterior és correcte. La demostració requereix, a més de les condicions de regularitat ja esmentades (de caràcter local), unes condicions de regularitat "global",

$$\int_{-\infty}^{\infty} |\psi(\xi)| d\xi, \quad \text{és convergent.}$$

La relació (gairebé simètrica) presentada en l'equació (21) s'anomena **teorema integral de Fourier**. Aquest teorema (i algunes de les seves versions) és molt important tant en les matemàtiques pures com en les aplicades. La funció g es coneix com la **transformada de Fourier** de ψ .

Transformada de Fourier

La funció $\psi(\xi) = \exp(-|\xi|)$ té la transformada de Fourier,

$$\begin{aligned}g(t) &= \int_{-\infty}^{\infty} \exp(-|\xi|) \exp(-2\pi i \xi t) d\xi = \int_{-\infty}^{\infty} (\exp(-(1 + 2\pi i t)\xi) + \exp(-(1 - 2\pi i t)\xi)) d\xi \\ &= \frac{1}{1 + 2\pi i t} + \frac{1}{1 - 2\pi i t} = \frac{1}{1 + 4\pi^2 t^2}.\end{aligned}$$

Si fem servir l'equació (21) obtenim

$$\int_{-\infty}^{\infty} \frac{1}{1 + 4\pi^2 t^2} \exp(2\pi i \xi t) dt = \exp(-|\xi|),$$

o, si consideramos $x = 2\pi t$,

$$\int_0^{\infty} \frac{1}{1 + x^2} \cos(\xi x) dx = \frac{\pi}{2} \exp(-|\xi|), \quad \xi \in \mathbb{R}.$$

2. Regressió

La *regressió* és una aplicació (o un cas particular) d'aproximació de funcions. De fet, aquesta tècnica també es coneix com a *regressió per mínims quadrats*, en al·lusió al fet que proporciona una solució òptima sobre la base del marc teòric presentat a l'apartat anterior.

En aquest apartat ens centrarem en l'anomenada *regressió lineal*, en la qual el concepte *lineal* es refereix a la relació entre el que seran les incògnites del problema: els coeficients de l'expansió. També veurem que hi ha diferents mètodes que encaixen en aquesta definició.

La regressió és una metodologia que ajuda a tractar conjunts de dades experimentals, especialment quan aquestes no es presenten de manera convenient en quantitat o qualitat, i a extreure'n algun tipus de mesura o desenvolupar relacions predictives. Pretén fer una avaluació descriptiva de les dades que permeti extreure conclusions útils per estudiar-les. En aquest sentit, quan fem servir el terme *regressió* és perquè estem en el cas discret d'aproximació de funcions.

Hi ha dos problemes típics quan parlem de dades experimentals: anàlisi de tendència i contrast d'hipòtesi.

L'**anàlisi de tendència** es basa a fer servir un patró observat en les dades per treure conclusions generals de l'experiment analitzat o fer prediccions futures de la seva evolució.

En el **contrast d'hipòtesi** es compara un model matemàtic existent amb les dades observades. Si els paràmetres del model són desconeguts, pot ser necessari determinar-los per tenir el millor ajustament possible amb les dades mesurades. D'altra banda, si els coeficients ja estan disponibles, pot ser convenient comparar els valors predits pel model i els valors reals observats, per avaluar l'adequació del model. Sovint es comparen diversos models per aconseguir el millor ajustament a un conjunt d'observacions donat.

2.1. Regressió lineal

Quan tenim dades que són de naturalesa variable, inestable o pseudoaleatòria, o que poden presentar errors de mesura (soroll), la millor estratègia per des-

Font de l'apartat

Apartat basat en l'obra següent: **Steven C. Chapra** (2018). *Applied Numerical Methods with MATLAB for Engineers and Scientists, Fourth Edition*. Nova York: McGraw-Hill Education.

criure les observacions sol ser una línia (funció) que pugui reproduir la forma o tendència general de les dades, més que ajustar cada punt individualment, la qual cosa es correspondria amb un problema d'interpolació.

Intuïtivament, la regressió seria l'equivalent basat, objectivament i matemàticament, en una metodologia més visual que consisteix a dibuixar, a partir de les dades representades en una gràfica, la "millor" línia sobre els punts que els descriu. El criteri que s'utilitzaria aplicant el sentit comú és el de dibuixar una línia que minimitzi d'alguna manera la discrepància entre els punts i la corba. Llavors, aprofitant el marc teòric de l'aproximació de funcions, podem concloure que la manera de mesurar aquesta distància perquè la minimització ens proporcioni la millor solució és la dels **mínims quadrats**. Recordem que, entre molts altres avantatges, la solució que s'obté és única, la qual cosa ja ens ha portat a pensar que és la solució òptima.

Pensem ara en el cas més senzill: descriure unes dades donades per una línia recta. Suposem que tenim un conjunt de dades per parells, (x_1, y_1) , (x_2, y_2) , \dots , (x_m, y_m) . L'expressió matemàtica d'una recta és

$$y = c_0 + c_1 x,$$

en què els coeficients c_0 i c_1 són l'**ordenada en origen** i el **pendent**, respectivament. Assumint que la recta no aproxima totes les observacions perfectament, per a cada parell (y_i, x_i) cometem un error ϵ_i , és a dir:

$$\epsilon_i = y_i - c_0 - c_1 x_i,$$

conegut també com a *residu*. En considerar mínims quadrats, particularment el cas discret, obtenim:

$$\sum_{i=1}^m \epsilon_i^2 = \sum_{i=1}^m (y_i - c_0 - c_1 x_i)^2.$$

Ara cal recordar la solució que es presenta a l'apartat 1, que hem d'adaptar al context de la regressió mitjançant una recta, prenent $n = 1$, $\phi_0(x) = 1$, $\phi_1(x) = x$, $f(x_i) = y_i$ i $w_i = \frac{1}{m}$. D'aquesta manera, l'obtenció dels coeficients se simplifica molt:

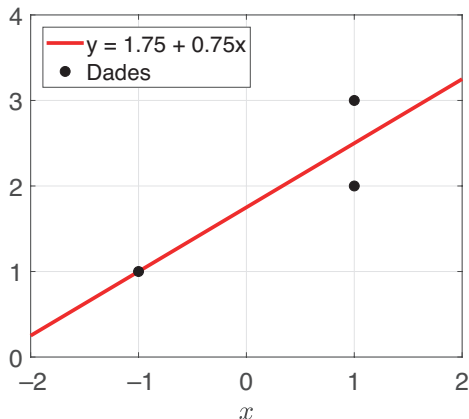
$$c_1 = \frac{m \sum_{i=1}^m x_i y_i - \sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m \sum_{i=1}^m x_i^2 - \left(\sum_{i=1}^m x_i\right)^2},$$

$$c_0 = \frac{1}{m} \sum_{i=1}^m y_i - c_1 \frac{1}{m} \sum_{i=1}^m x_i.$$

Regressió lineal

Trobeu la recta que millor s'ajusta als punts (1,2),(-1,1),(1,3). En resoldre el sistema d'equacions, obtenim que $c_0 = \frac{7}{4}$ i $c_1 = \frac{3}{4}$. En la figura 7 en podem veure el resultat.

Figura 7



Ara podem calcular els residus o errors que es cometen en l'aproximació, presentats en la taula 3.

Taula 3

x_i	y_i	Recta, y	Residu, ϵ_i
1	2	2,5	-0,5
-1	1	1,0	0,0
1	3	2,5	0,5

2.1.1. Quantificació de l'error de la regressió lineal

Com ja hem mostrat a l'apartat 1, la solució basada en mínims quadrats és única i, per tant, qualsevol altra recta produiria una suma de residus més gran. Una sèrie de propietats addicionals es poden deduir simplement examinant amb més detall la manera de calcular els residus. Si en definim la suma, S_r , com a

$$S_r = \sum_{i=1}^m (y_i - c_0 - c_1 x_i)^2,$$

veurem que s'assembla a una expressió molt coneguda, la de la desviació típica de y amb les mostres y_i ,

$$s_y = \sqrt{\frac{S_t}{m-1}}, \quad S_t = \sum_{i=1}^m (y_i - \bar{y})^2,$$

en què \bar{y} representa la mitjana. Sabem que la desviació típica expressa la discrepància de les mostres pel que fa a un valor central, la mitjana. S_r pot ser vist com la discrepància de les dades pel que fa a una altra mesura central, en aquest cas la línia recta.

L'analogia es pot portar encara més lluny en casos en què, d'una banda, les dades se separen de la línia amb una magnitud similar en tot el domini d'aquests i, de l'altra, la distribució dels punts al voltant de la línia és normal. En aquesta situació es pot demostrar que la solució de mínims quadrats és la que millor estima c_0 i c_1 (fins i tot utilitzant altres mesures). En estadística es coneix com el *principi de màxima versemblança*. A més, si les condicions prèvies es compleixen, es pot calcular una desviació típica per a la regressió lineal, definida en

$$s = \sqrt{\frac{S_r}{m-2}},$$

que es coneix com a *error estàndard de l'estimació*. Observeu que, a diferència de la variància estadística (o la desviació típica), en el denominador tenim $m-2$, ja que estimem dos valors c_0 i c_1 , per la qual cosa hem perdut dos graus de llibertat.

Com passa en el cas de la desviació típica estadística, l'error estàndard de l'estimació quantifica la dispersió de les dades, tot i que en aquest cas ho fa al voltant de la recta.

Aquests conceptes es poden fer servir per quantificar la "bondat" de l'ajustament, la qual cosa resulta molt útil a l'hora de comparar diferents tipus de regressions. Un cas particular d'especial importància és quan considerem les mesures de dispersió S_r i S_t , que corresponen a la dispersió pel que fa a la recta i a la mitjana, respectivament. La diferència $S_t - S_r$ quantifica la reducció de l'error, ja que descriu les dades en termes de la mitjana o una recta. Com que la magnitud d'aquesta mesura depèn de l'escala, ho normalitzem en S_t i obtenim l'anomenat *coeficient de determinació*:

$$r^2 = \frac{S_t - S_r}{S_t}, \quad (22)$$

l'arrel quadrada del qual, r , és l'anomenat *coeficient de correlació*. Si $r^2 = 1$ ($S_r = 0$), tenim que la línia recta explica perfectament el 100% de la variabilitat de les dades. No obstant això, per a $r^2 = 0$ ($S_r = S_t$), la línia recta no millora la mitjana a l'hora de descriure les dades. Una formulació alternativa de r que resulta ser més eficient per a la seva implementació en l'ordinador és

$$r = \frac{m \sum_{i=1}^m x_i y_i - \sum_{i=1}^m x_i \sum_{i=1}^m y_i}{\sqrt{m \sum_{i=1}^m x_i^2 - (\sum_{i=1}^m x_i)^2} \sqrt{m \sum_{i=1}^m y_i^2 - (\sum_{i=1}^m y_i)^2}}.$$

Error estàndard

Continuant amb l'exemple anterior, podem calcular les mesures d'error que hem introduït. La suma de residus al quadrat seria

$$S_r = (-0,5)^2 + 0^2 + 0,5^2 = 0,5.$$

Un cop calculat S_r , l'error estàndard s'obté fàcilment, però cal recordar que en aquest exemple $m = 3$,

$$s = s = \sqrt{\frac{S_r}{m-2}} = s = \sqrt{0,5} = 0,71.$$

Finalment, el coeficient de correlació r és

$$r = \sqrt{\frac{S_t - S_r}{S_t}} = \sqrt{\frac{2 - 0,5}{2}} = 0,87.$$

2.2. Linealització de relacions no lineals

La regressió lineal és una metodologia molt potent per descriure de la millor manera un conjunt de dades amb una línia. No obstant això, la seva aplicabilitat i eficiència depenen significativament del fet que la relació entre les variables dependent i independent sigui lineal (o gairebé). Per descomptat, existeixen múltiples casos de disciplines de la ciència i l'enginyeria en els quals les dades no presenten un comportament lineal, com ara l'estudi de poblacions, la radiació, etc. Com a primer pas intuïtiu en l'anàlisi de les dades n'hi ha prou a presentar-los en un gràfic i destriar si el model lineal és el més convenient.

Quan un comportament no lineal és observat, es pot fer ús de tècniques de regressió no lineal. No obstant això, en alguns casos, les relacions altament no lineals es poden transformar per expressar les dades de manera que siguin compatibles amb la regressió lineal. Així se'n simplifica molt el tractament. A continuació en presentem alguns exemples:

- **Model exponencial**

$$y = \alpha_1 \exp \beta_1 x,$$

en què α_1 i β_1 són constants. Aquest model s'empra en moltes disciplines d'enginyeria i ciència per simular valors que creixen (β_1 positiu) o decreixen (β_1 negatiu) a una velocitat directament proporcional a la seva pròpia magnitud. La relació entre les variables x i y és, en principi, clarament no lineal. No obstant això, aplicant una transformació bastant simple, prenent el logaritme natural, obtenim

$$\log(y) = \log(\alpha_1) + \beta_1 x,$$

que s'ajusta al model lineal.

Automòbils al món

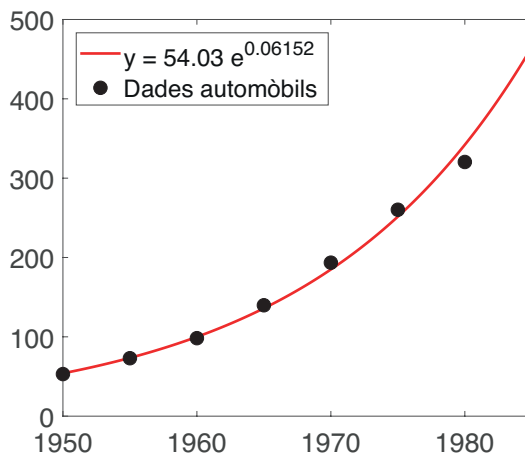
Les dades de la taula 4 descriuen el nombre d'automòbils que han operat al món en diferents anys començant pel 1950. A simple vista, sembla que la quantitat de cotxes ha seguit un creixement exponencial. Així doncs, cal ajustar un model exponencial per a aquestes dades o, el que és el mateix, determinar α_1 i β_1 de la descripció anterior.

Taula 4

Any	Automòbils ($\times 10^6$)
1950	53,05
1955	73,04
1960	98,31
1965	139,78
1970	193,48
1975	260,20
1980	320,39

Després de fer la transformació logarítmica i resoldre el sistema de les equacions normals obtenim: $c_0 = \log(\alpha_1) = 3,986$ i $c_1 = \beta_1 = 0,06152$. Desfent la transformació calculem $\alpha_1 = \exp(c_0) = 54,03$. En la figura 8 podem veure les dades i la forma de la corba exponencial de regressió.

Figura 8



• **Equació potència**

$$y = \alpha_2 x^{\beta_2},$$

en què α_1 i β_1 són constants. Aquesta relació s'utilitza habitualment per ajustar dades experimentals quan el model matemàtic és desconegut. Si $\beta_2 \neq 0$, l'equació és no lineal. De manera similar a l'anterior, però a partir de logaritmes de base 10, tenim

$$\log_{10}(y) = \log_{10}(\alpha_2) + \beta_2 \log_{10}(x),$$

del qual resulta en una recta amb pendent β_2 i una ordenada en l'origen $\log_{10}(\alpha_2)$. Observeu que es podria fer servir qualsevol base per al logaritme, tot i que la base 10 és la més habitual.

- **Equació de la taxa de creixement de saturació**

$$y = \alpha_3 \frac{x}{\beta_3 + x},$$

en què α_1 i β_1 són constants. L'equació descriu un creixement que s'atenua quan x es fa gran, la qual cosa representa un efecte de saturació. Per linealitzar aquesta equació només cal invertir-la:

$$\frac{1}{y} = \frac{1}{\alpha_3} + \frac{\beta_3}{\alpha_3} \frac{1}{x}.$$

Aquestes transformacions permeten utilitzar aquests models en el marc de la regressió lineal i calcular els coeficients. Una vegada calculats, es recupera el model original, amb la qual cosa es desfà la transformació i es pot fer servir per a propòsits de predicció.

2.3. Regressió lineal polinòmica

Fins ara només hem estudiat la manera de fer servir la recta per descriure les dades. No obstant això, hi ha situacions en què, encara que la distribució dels punts presenti un patró clar, aquest no es pot ajustar mitjançant una línia recta. En aquests casos, una corba sol ser la millor opció. Així, a més de la linealització mitjançant transformacions presentada anteriorment, es pot optar per ajustar un polinomi a les dades, utilitzant l'anomenada *regressió polinòmica*. Aquest tipus de regressió és, en certa manera, una generalització del que es basa en la recta, tot i que continua essent lineal i un cas particular de la teoria d'aproximacions de funcions per mínims quadrats.

Així doncs, generalitzarem el cas de la recta en un polinomi de grau superior. Imaginem que ajustem un polinomi de segon ordre (quadràtic) així:

$$y = c_0 + c_1x + c_2x^2.$$

Com en el cas anterior, aquesta formulació és una especialització del problema de mínims quadrats presentat a l'apartat 1. Simplement hem de partir de $n = 2$, $\phi_0(x) = 1$, $\phi_1(x) = x$, $\phi_2(x) = x^2$ i $f(x) = y$. Ara la suma del quadrat dels residus és

$$S_r = \sum_{i=0}^m \left(y_i - c_0 - c_1x_i - c_2x_i^2 \right)^2,$$

en què podem observar que ara tenim un terme més per tractar de reduir les diferències individuals. Aplicant la solució ja estudiada basada en equacions normals, els coeficients c_0 , c_1 i c_2 es poden calcular fàcilment. No oblidem, però, que aquest tipus d'ajustament segueix formant part del context de regressió lineal, ja que les incògnites del problema són els coeficients, la relació dels quals és lineal.

Finalment, la regressió polinòmica es pot utilitzar amb polinomis de grau superior, n . Així doncs, tindriem:

$$y = c_0 + c_1x + c_2x^2 + \dots + c_nx^n.$$

De nou, mitjançant el sistema d'equacions normals ($n + 1$ equacions lineals) s'obtenen els coeficients $c_0, c_1, c_2, \dots, c_n$. L'error estàndard es formula així:

$$s = \sqrt{\frac{S_r}{m - (n + 1)}},$$

en què el denominador és $m - (n + 1)$ perquè $n + 1$ coeficients s'utilitzen per calcular S_r , amb la qual cosa hem perdut $n + 1$ graus de llibertat. El coeficient de determinació r^2 es pot calcular utilitzant l'equació (22).

2.4. Regressió lineal múltiple

Una altra extensió molt emprada que també s'emmarca en la regressió lineal apareix quan la variable y és una funció lineal de dues o més variables. Per exemple, y pot ser aquesta funció de x_1 i x_2 :

$$y = c_0 + c_1x_1 + c_2x_2.$$

Aquesta formulació és particularment útil quan s'ajusten dades experimentals en què la variable estudiada depèn de dues variables més. Per això, aquesta tècnica també se sol anomenar **variables explicatives**, ja que es tracta de determinar la influència de x_1 i x_2 en la variable y . En aquest cas concret bidimensional, la línia se substitueix per un pla.

Com en els casos anteriors, definim la suma de residus d'aquesta manera:

$$S_r = \sum_{i=0}^m (y_i - c_0 - c_1x_{1,i} - c_2x_{2,i})^2.$$

Els coeficients c_0 , c_1 i c_2 s'obtenen resolent el problema de mínims quadrats amb $n = 2$, $\phi_0(x) = 1$, $\phi_1(x) = x_1$, $\phi_2(x) = x_2$ i $f(x) = y$.

De nou, el cas bidimensional es pot estendre al cas de n dimensions, amb

$$y = c_0 + c_1x_1 + c_2x_2 + \dots + c_nx_n,$$

en què l'error estàndard es formula com a

$$s = \sqrt{\frac{S_r}{m - (n + 1)}},$$

i el coeficient de determinació es calcula amb l'equació (22).

A més dels casos en què una variable és linealment dependent de dues o més variables, la regressió lineal múltiple es pot fer servir en la derivació d'equacions de potències de la manera genèrica:

$$y = c_0x_1^{c_1}x_2^{c_2} \dots x_n^{c_n}.$$

Aquest tipus d'equacions són extremament útils per ajustar dades experimentals. Per fer servir la regressió lineal múltiple en aquest context, només cal aplicar una transformació logarítmica:

$$\log(y) = \log(c_0) + c_1 \log(x_1) + c_2 \log(x_2) + \dots + c_n \log(x_n).$$

En aquesta guia hem presentat diversos models i tècniques numèriques per aproximar funcions. La selecció del model adequat a partir d'un conjunt de dades donat no és una tasca trivial i, en molts casos, requereix una anàlisi exhaustiva, tant de les propietats del model com de la qualitat i la quantitat de les dades. En aquest sentit, particularment a l'apartat de regressió, hem donat algunes mesures de la bondat en l'ajustament, que es poden fer servir per comparar diferents models, tot i que s'han de valorar molts altres factors: robustesa, capacitat predictiva, facilitat d'ús, etc. En la pràctica, qualsevol coneixement previ o la intervenció d'experts en la matèria també són aspectes que s'han de tenir en compte.

Bibliografia

Chapra, Steven C. (2018). *Applied Numerical Methods with MATLAB for Engineers and Scientists, Fourth Edition*. Nova York: McGraw-Hill Education.

Dahlquist, Germund; Bjorck, Ake (2003). *Numerical Methods*. Dover Publications, Inc.

