2 Sistemes sobredeterminats. Aproximació per mínims quadrats

S'anomena *sistema sobredeterminat* a un sistema lineal en el qual el nombre d'observacions m és major que el nombre n d'incògnites, és a dir, del tipus

$$Ax = b, \qquad A \in \mathcal{M}_{m\times n}, \quad m > n.$$

Aquests sistemes, típicament, no tenen solució. El que es busca en canvi es trobar aquella solució x tal que minimitzi $\|Ax - b\|_2$. Per aquest motiu es coneix com a *aproximació per mínims quadrats*. Una manera usual de resoldre'l és projectant el vector b sobre l'espai vectorial generat pels vectors columna de la matriu A:

$$A^\top Ax = A^\top b. \tag{1}$$

L'expressió (1) s'anomena *Equacions Normals*. Si la matriu inicial A té rang màxim (o sigui n), aleshores la matriu $A^\top A \in \mathcal{M}_{n\times n}$ és regular. Malgrat això, és fàcil trobar exemples de matrius A de rang màxim que en aritmètica de coma flotant de doble precisió proporcionen $A^\top A$ molt propera a singular.

Fins i tot en el cas de que $A^\top A$ sigui regular aplicar un mètode gaussià per a resoldre (1) no és tampoc el més convenient en general. Concretament, atès que $A^\top A$ és simètrica, seria raonable aplicar una descomposició de Choleski $A^\top A = LL^\top$, amb L matriu triangular inferior. La dificultat numèrica prové del fet que el *nombre de condició* de la matriu $A^\top A$ i el de la matriu L poden crèixer considerablement.

Un manera de preservar un mínim control sobre el nombre de condició $\mu_2(A) = \|A\|_2\|A^{-1}\|_2$ es basa en l'anomenada *descomposició QR* de la matriu A: escrivim $A = QR$, on Q és ortogonal (unitària si és complexa) i R triangular superior. Així,

$$A^\top Ax = A^\top b \Leftrightarrow (QR)^\top(QR)x = (QR)^\top b \Leftrightarrow R^\top Q^\top QRx = R^\top Q^\top b \Leftrightarrow Rx = Q^\top b$$

El sistema final a resoldre

$$Rx = Q^\top b$$

és triangular superior i els valors $x = (x_1, x_2, \ldots, x_n)$ s'obtenen per *substitució cap a endarrere*.

El mètode que haureu de programar és una versió modificada de l'algorisme d'ortogonalització de Gram-Schmidt i té l'avantatge respecte d'aquest que la matriu Q obtinguda és de mida $m \times n$ enlloc de $m \times m$ i el nombre d'operacions emprades.

Problema: Donada una matriu $A \in \mathcal{M}_{m,n}$, calculeu la seva descomposició QR, amb Q ortogonal ($Q^\top Q = \mathrm{I}$) i R triangular superior.

Algorisme:

[0]T. Lázaro, M. Ollé i J.R. Pacha. Departament de Matemàtica Aplicada I

Pas 0: Definiu $A_1 = A = \left(a_1^{(1)}\ a_2^{(1)}\ \cdots a_n^{(1)}\right)$ on $a_j^{(1)}$ indica la columna j-ésima de la matriu A_1.

Pas 1: Normalitzem la columna $a_1^{(1)}$:

$$r_{11} = \|a_1^{(1)}\|, \qquad q_1 = \frac{a_1^{(1)}}{r_{11}}.$$

Ortogonalitzem respecte a aquesta columna totes les columnes posteriors:

$$r_{1s} = q_1^\top a_s^{(1)}, \qquad a_s^{(2)} = a_s^{(1)} - r_{1s} q_1.$$

Obtenim així la matriu $A_2 = \left(q_1\ a_2^{(2)}\ a_3^{(2)}\ \cdots a_n^{(2)}\right)$.

Pas k: Suposem que tenim ara la matriu $A_k = \left(q_1\ q_2\ \cdots q_{k-1}\ a_k^{(k)}\ a_{k+1}^{(k)}\ \cdots a_n^{(k)}\right)$. Per construcció, les seves columnes satisfan

$$q_j^\top q_\ell = \delta_{j\ell}, \qquad q_j^\top a_s^{(k)} = 0, \qquad j, \ell = 1 \div k-1, \quad s = k \div n.$$

Normalitzem la columna k-ésima,

$$r_{kk} = \|a_k^{(k)}\|, \qquad q_k = \frac{a_k^{(k)}}{r_{kk}},$$

i ortogonalitzem respecte de q_k les columnes $a_{k+1}^{(k)}, \cdots, a_n^{(k)}$,

$$r_{ks} = q_k^\top a_s^{(k)}, \qquad a_s^{(k+1)} = a_s^{(k)} - r_{ks} q_k, \qquad s = k+1 \div n.$$

S'obté d'aquesta manera $A_{k+1} = \left(q_1\ q_2\ \cdots q_k\ a_{k+1}^{(k+1)}\ a_{k+2}^{(k+1)}\ \cdots a_n^{(k+1)}\right)$ que verifica $q_j^\top q_\ell = \delta_{j\ell}$ i $q_j^\top a_s^{(k)} = 0$ per a $j, \ell = 1 \div k$, $s = k+1 \div n$.

Pas n: Després de n iteracions obtenim $A_{n+1} = (q_1\ q_2\ \cdots q_n)$, matriu ortogonal satisfent que

$$A = QR, \qquad Q = A_{n+1}, \qquad R = (r_{ks}).$$

Recordeu que si A és regular i totes les $r_{kk} > 0$ – com al nostre cas – la descomposició QR és única.

Exercici 2.1 *Programeu la descomposició QR d'una matriu A seguint aquest mètode. Useu-la per a calcular A^{-1}, $\mu_2(A)$ i $\mu_\infty(A)$, on $\mu_p(A) = \|A\|_p\ \|A^{-1}\|_p$ és el* número de condició *(en norma sub-p) de la matriu A.*

Exercici 2.2 *Les xifres oficials corresponents al padró municipal de Catalunya (vegeu la pàgina web de l'Institut d'Estadística de Catalunya IDESCAT) entre els anys 2000 i 2010 són aquestes:*

2000	2001	2002	2003	2004	2005
6.261.999	6.361.365	6.506.440	6.704.146	6.813.319	6.995.206

2006	2007	2008	2009	2010
7.134.697	7.210.508	7.364.078	7.475.420	7.512.381

Aleshores,

(*i*) *Busqueu el polinomi del grau escaient que millor aproximi aquesta taula en norma sub-*2*. Justifiqueu l'elecció mostrant una fita de l'error d'aquest i de la resta amb els que heu provat.*

(*ii*) *Useu-lo per a extrapolar la població de Catalunya l'any 2011 i compareu-la amb la xifra oficial (7.535.251). Quina seria la predicció pel 2012 ?*

(*iii*) *Feu una gràfica on apareguin els punts de la taula (representants per una rodona) i les gràfiques de les tres millors aproximacions.*

(*iv*) *Considereu el cas de $p_{10}(x)$, polinomi de grau* 10 *obtingut per mínims quadrats. Representeu-lo juntament amb els punts de la taula (aquests amb una rodona). Què observeu? Comenteu-lo.*

Comentari 2.3 *Numèricament és millor que considereu la taula anterior amb abscisses* $0, 1, \ldots, 10$ *enlloc de* $2000, 2001, \ldots, 2010$.

Un segon mètode, equivalent per a resoldre un sistema lineal sobredeterminat i per a trobar la projecció ortogonal d'una funció f sobre un espai de funcions $\mathcal{F}$, és trobar una base de polinomis ortogonals. En el cas de l'exercici 2.2, les abscisses són equidistants i, llavors, podem fer servir els anomenats *polinomis de Gram*, $P_{j,m}(t)$. Aquesta família de polinomis és ortogonal respecte de la taula $t_k = k$, $k = \div m$. Es poden obtenir a través de la següent recurrència:

$$\begin{aligned} P_{0,m}(t) &= 1, \qquad P_{1,m}(t) = 1 - \frac{2t}{m}, \\ (j+1)(m-j)P_{j+1,m}(t) &= (2j+1)(m-2t)P_{j,m}(t) - j(m+j+1)P_{j-1,m}(t), \qquad j = 1 \div m-1. \end{aligned}$$

En el cas que les nostres $m+1$ abscisses x_k estiguessin equidistribuïdes a l'interval $[a, b]$, caldria considerar el canvi de variable

$$x = a + \frac{b-a}{m} t \Leftrightarrow t = \frac{x-a}{b-a} m$$

i treballar amb els polinomis ortogonals $\tilde{P}_{j,m}(x) = P_{j,m}\left(\frac{x-a}{b-a} m\right)$

Exercici 2.4 (Continuació de l'exer. 2.2) *Considereu novament el problema de buscar un model per a la població catalana a la darrera dècada.*

(*i*) *Amb l'ajut d'un manipulador algebraic (Maple, Matlab, Octave, ...) calculeu els polinomis de Gram $P_{j,m}(t)$ associats a les abscisses $0, 1, \ldots, 10$ (que correspon al nostre cas, vegeu Comentari 2.3). Guardeu-los en un programa `C` o `C++`, conservant únicament els seus coeficients.*

(*ii*) *Plantegeu la resolució del problema 2.2 fent servir aquesta família de polinomis i responeu a les mateixes preguntes. Compareu els resultats obtinguts en ambdós casos. Recordeu que si necessiteu avaluar un polinomi en un cert punt una manera eficient de fer-ho és fent servir la Regla de Horner.*

Bibliografia

[1] A. Aubanell, A. Benseny, A. Delshams. *Eines bàsiques de càlcul numèric*. Universitat Autònoma de Barcelona, Barcelona, 1991.

[2] C. Bonet, À. Jorba, T.M-Seara, J. Masdemont, M. Ollé, A. Susin, M. Valencia. *Càlcul numèric*. Aula Teòrica 23, Edicions UPC, Barcelona, 1994.

[3] G. Dahlquist, A. Björck. *Numerical methods*. Prentice Hall Inc., New Jersey, 1974.

[4] A. Quarteroni, F. Saleri. *Scientific Computing with MATLAB and Octave, Second Edition*. Springer-Verlag, Berlín, Heidelberg, 2006.

[5] J. Stoer, R. Bulirsch. *Introduction to numerical analysis* (third edition). Texts in Applied Mathematics, Springer-Verlag, New York - Berlin - Heidelberg, 2002.