ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

ΔΙΑΛΕΞΗ 19, 9/01/10

Ε. ΓΑΛΛΟΠΟΥΛΟΣ
Τμήμα Μηχ. Η/Υ και Πληροφορικής
Παν/μιο Πατρών

Γραμμικό πρόβλημα ελαχίστων τετραγώνων (Κεφ. 6)

$$\boxed{\arg\min_{x\in\mathbb{R}^n} \|Ax - b\|_2}$$

... από εδώ και πέρα υπονοείται η ευκλείδεια νόρμα.

Θεώρημα 1. *Αν το A είναι πλήρους τάξης:*

$$\mathrm{rank(A) = min(m, n)}$$

τότε η λύση είναι μοναδική. □

Δυσκολία: Αν το A δεν είναι πλήρους τάξης, το x δεν είναι μοναδικό $\Rightarrow$ χρειάζεται επιπλέον θεωρία που να βοηθήσει στην επιλογή του «καλύτερου», π.χ. εκείνου με την ελάχιστη νόρμα $\|x\|$.

Τί κάνει η MATLAB;
Στη σελίδα `help` της MATLAB :

```
MATLAB Functions:  Arithmetic Operators + - *  &#92; ß
```

αναφέρεται σχετικά με την πράξη $A\backslash b$:

If A is not square, then Householder reflections are used to compute an orthogonal-triangular factorization. A*P = Q*R where P is a permutation, Q is orthogonal and R is upper triangular (see qr). The least squares solution X is computed with X = P*(R\(Q'*B)) If A is sparse, then MATLAB computes a least squares solution using the sparse qr factorization of A.

_ Βασικές μέθοδοι για προβλήματα του τύπου που επιλύει η LAPACK:

μέθοδος κανονικών εξισώσεων ταχύτερη, χρησιμοποιείται όταν $\kappa(A)$ μικρό

παραγοντοποίηση QR πιο συνηθισμένη μέθοδος, στοιχίζει περίπου διπλάσια της προηγούμενης

... οι παράγοντες Q, R χρησιμοποιούνται και σε άλλες εφαρμογές

SVD χρήσιμη αν το πρόβλημα έχει πολύ μεγάλο δείκτη κατάστασης αλλά κοστίζει περισσότερο

... οι παράγοντες U, V, Σ χρησιμοποιούνται και σε άλλες εφαρμογές

ισοδύναμο γραμμικό σύστημα

Μέθοδος κανονικών εξισώσεων (normal equations) - ΠΡΟΑΙΡΕΤΙΚΟ Αν το $A^\top A$ είναι σθο, θα μπορούσαμε να χρησιμοποιήσουμε τον παρακάτω αλγόριθμο:

Μέθοδος κανονικών εξισώσεων

$G = A'A$	(* υπολογισμός γινομένου αξιοποιώντας συμμετρία n^2m+ *)
$R = \mathtt{chol}(G)$	(* παραγοντοποίηση Cholesky, R άνω τριγωνικός παράγοντας $n^3/3+$ *)
$x = R \setminus (R' \setminus b)$	(* επίλυση με τριγωνικούς παράγοντες $2n^2+$*)

Κόστος: $n^2m + \frac{1}{3}n^3 + O(n^2)$ πράξεις α.κ.υ.
Βασικά εργαλεία

κανονικοποίηση κανονικές εξισώσεις, ημικανονικές εξισώσεις

ορθογωνιοποίηση

- παραγοντοποίηση QR
- διαδικασία Gram-Schmidt

SVD

ορθογώνιες προβολές

στοιχειώδεις μετασχηματισμοί

- Givens
- Householder

Παρατηρήσεις για την επίλυση μέσω κανονικών εξισώσεων Με βάση τα κριτήρια του ΕΥ:

- ταχεία μέθοδος
- ... αλλά μπορεί να εντείνει υπάρχοντα αριθμητικά προβλήματα

Βασική παρατήρηση: Πρόκειται για επίλυση του συστήματος $Bx = b$ όπου $B := A^\top A$. Όμως ο δείκτης κατάστασης του προβλήματος «επίλυση κανονικών εξισώσεων» $\kappa_2(B) = \kappa_2^2(A)$.
Γιατί: $A = U\Sigma V^\top \Rightarrow A^\top A = V\Sigma^\top \Sigma V^\top$ οπότε

$$\kappa_2(B) = \left(\frac{\sigma_{\max}(A)}{\sigma_{\max}(A)}\right)^{\mathbf{2}}$$

Δηλαδή, ακόμα και αν ο αλγόριθμος επίλυσης του συστήματος είναι πισω ευσταθής, το εμπρός σφάλμα μπορεί να είναι ανάλογο με το *τετράγωνο* του δείκτη κατάστασης του A.

Παράδειγμα 1. *Αν $\kappa_2(A) = 1e8$ στη λύση του συστήματος $Ax = b$ μπορεί να «χάσουμε» περί τα 8 ψηφία (τα μισά της διπλής ακρίβειας IEEE) ενώ στη λύση του $A^\top Ax = A^\top b$ θα τα χάσουμε όλα!* □

Αν ο δείκτης κατάστασης του A είναι σχετικά μεγάλος, τότε καλό είναι να αποφεύγουμε τη χρήση τους

Παραγοντοποίηση QR και ύπαρξη ΟΚ βάσης

Θεώρημα 2. *Έστω $A \in \mathbb{R}^{m\times n}$ και $m \geq n$. Τότε υπάρχει μητρώο $Q \in \mathbb{R}^{m\times n}$ με ορθοκανονικές (ΟΚ) στήλες και άνω τριγωνικό μητρώο $R \in \mathbb{R}^{n\times n}$ ώστε $A = QR$.*

- *Αν $m = n$ τότε το Q είναι μοναδιαίο·*
- *αν A αντιστρέψιμο ή έχει γραμμικά ανεξάρτητες στήλες τότε το R μπορεί να επιλεχθεί έτσι ώστε να έχει θετικά στοιχεία στη διαγώνιο. Τότε τα Q, R ειναι μοναδικά.*

□

Προσοχή: $Ax = Q(Rx)$ επομένως

Οι στήλες του Q είναι ΟΚ βάση του $\text{Range}(A) = \text{span}\{a_1, ..., a_n\}$.

Η παραπάνω παραγοντοποίηση επιστρέφει $Q \in \mathbb{R}^{m\times n}, R \in \mathbb{R}^{n\times n}$.

- Η παραπάνω παραγοντοποίηση πολλές φορές αποκαλείται «οικονομική διάσπαση QR» και στη MATLAB υπολογίζεται με την εντολή `qr(A,0)`

Αν $m > n$, θα μπορούσαμε να επεκτείνουμε τις ΟΚ στήλες του Q με $m - n$ ακόμα στήλες, έστω G ώστε $[Q, G]$ να είναι ΟΚ βάση για το $\mathbb{R}^m$.

Επίλυση ΑΓΑ.2 με QR

Έστω ότι γνωρίζουμε την οικονομική διάσπαση $A = QR$

$$\begin{aligned} \|Ax - b\| &= \|Q^\top Ax - Q^\top b\| \\ &= \|Rx - \hat{b}\| \end{aligned}$$

Τότε λύση του ΑΓΑ.2 είναι

αν $m \geq n$**:** η λύση x του $Rx = Q^\top b$, εφόσον R αντιστρέψιμο.

αν $m < n$**:** $R = [R_1, R_2]$ όπου $R_1 \in \mathbb{R}^{m\times m}$, οπότε επιλέγουμε $R_1 x_1 = \hat{b}_1$ και $x = [x_1; 0]$.

Προσοχή: και οι δυο παραπάνω περιπτώσεις απαιτούν από το R ή το R_1 να είναι αντιστρέψιμα. Αυτό δεν ισχύει πάντα στην απλή παραγοντοποίηση QR, π.χ. αν

$$\text{rank}(A) < \min(m, n)$$

Γενικά αν ισχύει η παραπάνω συνθήκη, τότε έχουμε

πρόβλημα ελαχίστων τετραγώνων ελλειπούς τάξης (rank deficient)

Τότε η λύση δεν είναι μοναδική.
Γενικά χρειάζεται:

$$\text{παραγοντοποίηση } QR \text{ με οδήγηση στήλης: } AE = QR$$

Κανονικές εξισώσεις και QR Αν γνωρίζουμε ότι $A = QR$ τότε $Ax = b \Rightarrow Rx = Q^\top b$ και μπορούμε να δοκιμάσουμε να λύσουμε ως προς x. Θα δούμε σε λίγο τις μεθόδους πιο αναλυτικά.
Αντί γι' αυτό θα μπορούσαμε να χρησιμοποιήσουμε την παραγοντοποίηση ως εξής:
Μέθοδος ημικανονικών (seminormal) εξισώσεων:

$$A^\top A = R^\top R \Rightarrow x = (R^\top R)^{-1} A^\top b$$

όπου το R προέρχεται από την παραγοντοποίηση QR. Παρατηρήστε ότι δεν χρησιμοποιείται το Q.
Ορθογώνια προβολή σε υπόχωρο
Ορισμός 1. *Έστω υπόχωρος $\mathcal{U}$ διάστασης n και βάση του $\{u_1, ..., u_n\}$. Τότε αν θέσουμε $U := [u_1, ..., u_n]$, το μητρώο ορθογώνιας προβολής επί του $\mathcal{U}$ είναι*

$$P_{\mathcal{U}} = U(U^\top U)^{-1} U^\top$$

□

Σύνδεση με τις κανονικές εξισώσεις Θυμηθείτε ότι η λύση του ΑΓΑ.2 από τις κανονικές εξισώσεις είναι

$$x = (A^\top A)^{-1} A^\top b$$

Το x που ελαχιστοποιεί το $\|Ax - b\|_2$ ελαχιστοποιεί και την απόσταση $\|y - b\|_2$ όπου το $y \in < a_1, ..., a_n >$ ανήκει στο χώρο που παράγεται από τις στήλες του A, δηλ.

$$\min_x \|Ax - b\|_2 = \min_{y \in \text{span} < a_1, ..., a_n >} \|y - b\|_2.$$

Από τη θεωρία της Γραμμικής Άλγεβρας, το y είναι το διάνυσμα που προκύπτει από την ορθογώνια προβολή του b στο $\text{span} < a_1, ..., a_n >$. Σύμφωνα με τα παραπάνω:

$$y = A(A^\top A)^{-1} A^\top b$$

και επομένως

$$x = (A^\top A)^{-1} A^\top b,$$

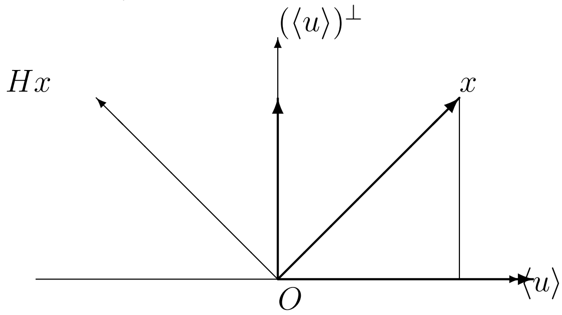
ακριβώς όπως προβλέπει και η λύση με τις κανονικές εξισώσεις!
Στοιχειώδη Ερμιτιανά μητρώα/Ανακλαστές/Μετασχηματισμοί Householder Έτσι ορίζονται τα μητρώα

$$H := E(u, u; 2/u^* u) = I - \frac{2}{u^* u} u u^*.$$

Ισχύει ότι[1]

[1] Αν τα στοιχεία είναι πραγματικα, το $*$ ισοδυναμεί με το ανάστροφο $\top$.

- $H = I - P_u - P_u$ όπου P_u ο τελεστής ορθ. προβολής στο u.
- $H^* = H$ και $H^*H = H^2 = I$.



Γεωμετρική ερμηνεία:

Η εφαρμογή του H στον x έχει σαν αποτέλεσμα την ανάκλαση του x ως προς τον υποχώρο $(\langle u\rangle)^\perp$.

Πώς επιλέγουμε το u; Γεωμετρική ερμηνεία στις 2δ

- Είναι εμφανές στις 2 διαστάσεις:
- Για να έχουμε $Hx = \|x\|e_1$ αρκεί να «ανακλαστεί» το x ως προς τη διχοτόμο μεταξύ των διανυσμάτων e_1 και του x.
- Είναι προφανές ότι η διχοτόμος είναι $u = x + \|x\|e_1$

Ιδιότητες και υπολογισμός ανακλαστών

Οι ανακλαστές

- Είναι συμμετρικά και ορθογώνια μητρώα.
- Ανανεώσεις 1ης τάξης του ταυτοτικού μητρώου.
- Χρησιμοποιούνται για να μηδενίσουμε επιλεγμένα διαδοχικά στοιχεία διανύσματος.

Δοθέντος $x \in \mathbb{R}^n$ έχουμε

$$Hx = (I - 2\frac{uu^\top}{u^\top u})x = x - 2\frac{u^\top x}{u^\top u}u$$

Για να μηδενίσουμε όλα εκτός του 1ου στοιχείου του x,

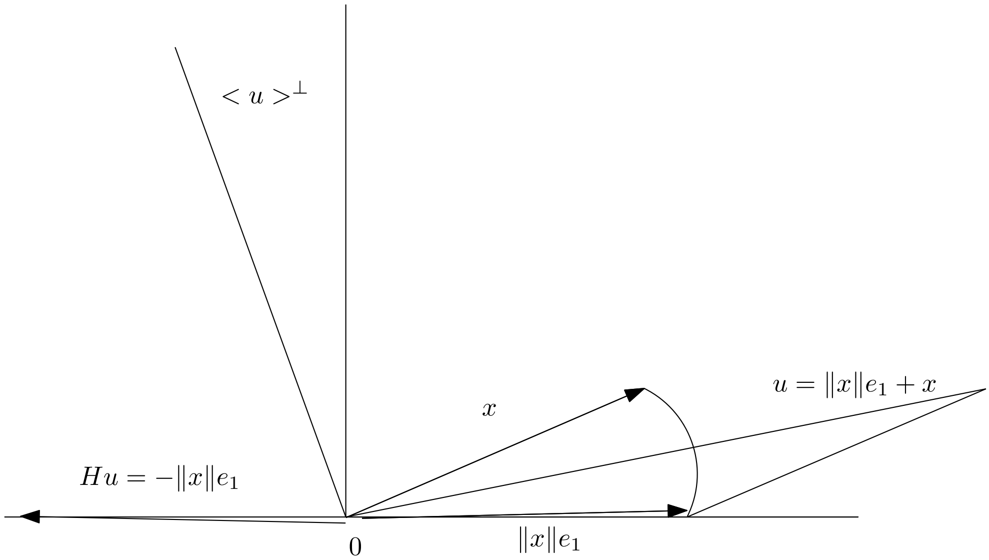
$$Hx \in \langle e_1\rangle \Rightarrow x \in \langle e_1, u\rangle$$

άρα $u = x + \alpha e_1$. Υπολογίζοντας τα $u^\top x, u^\top u$ έχουμε πως

$$Hx = (1 - 2\frac{x^\top x + \alpha\xi_1}{x^\top x + 2\alpha\xi_1 + \alpha^2})x - 2\alpha\frac{u^\top x}{u^\top u}e_1.$$

Θέτοντας $\alpha = \pm\|x\|_2$:

$$u = x \pm \|x\|_2 e_1 \Rightarrow Hx = \mp\|x\|_2 e_1$$



Υλοποίηση:

- Το πρόσημο επιλέγεται έτσι ώστε να αποφεύγεται πιθανό πρόβλημα *καταστροφικής απαλοιφής*:
$$u = x + \mathbf{sign}(\xi_1)\|x\|_2 e_1.$$

- το $H(u)$ είναι ανεξάρτητο της κλιμάκωσης του u:
$$H = I - \frac{2}{u^\top u}uu^\top = I - \frac{2}{(\alpha u)^\top(\alpha u)}(\alpha u)(\alpha u)^\top.$$
Το u λέγεται *διάνυσμα Householder*. Καθώς $H(u) = H(\alpha u)$, μπορούμε να κανονικοποιήσουμε ώστε $u(1) = 1$.

- Ο παρακάτω αλγόριθμος στοιχίζει $T^{\text{REFL}}_{\text{αρθ}} = 3n$ α.κ.υ.

```
function u = REFL (x)
επιστρέφει το διάνυσμα Householder u
ο ανακλαστής H(u) να ανακλά τα στοιχεία του x.
    n = length(x); μ = norm(x, 2); u = x
    if  μ ≠ 0
        β = x(1) + sign(x(1))μ
        u(2 : n) = u(2 : n)/β
    end
    u(1) = 1 (κανονικοποίηση)
    end
```

Αλγοριθμικά θέματα και υλοποίηση

- Οι ανακλαστές είναι στοιχειώδη μητρώα
- ... σπάνια τα κατασκευάζουμε ....
- ... τα διαχειριζόμαστε χρησιμοποιώντας την ειδική τους μορφή (παρόμοια με τα μητρώα Gauss)
- π.χ. $HA = A - \frac{2}{u^\top u} u(A^\top u)^\top$,
- Το μητρώο Q συνήθως δεν κατασκευάζεται παρά μόνον αν χρειζόμαστε την ορθογώνια βάση ...
- αλλά αποθηκεύουμε τα (κανονικοποιημένα) διανύσματα Householder στο κάτω τριγωνικό μέρος του A.
- και όταν χρειάζεται να πολλαπλασιάσουμε με το Q το κάνουμε «έξυπνα» με τα διανύσματα h_j
- Ο σχηματισμός του AH ή HA απαιτεί 1 MXV, και μια ανανέωση τάξης 1.
- Μερικό κόστος: $T_{\text{αρθ}}^{\text{REFL.(COL/ROW)}} = 4mn$ πράξεις α.κ.υ.

Παραγοντοποίηση QR: μέθοδος **Householder**

I Θέτουμε $A^{(0)} := A$.

II Για $k = 1, .., \hat{n}$ ($\hat{n} = n$ αν $m > n$, $\hat{n} = n - 1$ αν $m = n$):

1. Κατασκευή διανύσματος Householder, u_k, τέτοιου ώστε ο ανακλαστής να μηδενιζει τα υποδιαγώνια στοιχεία της στήλης k A. Το $u_k = [\overbrace{0, ..., 0}^{k-1}, 1, \star ..., \star]^\top$.
2. $A^{(k)} = A^{(k)} - \frac{2}{u_k^\top u_k} u_k (u_k^\top A^{(k-1)})$

Παρατηρήσεις 1.

- *Η πληροφορία για τα κανονικοποιημένα διανύσματα u_k μπορεί να αποθηκευτεί στο κάτω τριγωνικό τμήμα του A.*
- *Στο παραπάνω το $Q \in \mathbb{R}^{m \times m}$ δεν υπολογίζεται άμεσα αλλά δίδεται μέσω των παραγόντων του ανακλαστών.*

$$\begin{bmatrix} \rho_{11} & \rho_{12} & \cdot & \rho_{1n} \\ \eta_2^{(1)} & \rho_{22} & \cdot & \rho_{2n} \\ \cdot & \eta_2^{(1)} & \cdot & \cdot \\ \cdot & \cdot & \cdot & \rho_{n,n} \\ \vdots & \cdots & \cdots & \eta_{n+1}^{(n)} \\ \eta_m^{(1)} & \eta_m^{(2)} & \cdot & \eta_m^{(n)} \end{bmatrix}$$

- *Αν το μητρώο είναι τετραγωνικό ($m = n$) χρησιμοποιούνται μόνο $n - 1$ ανακλαστές.*
- $\Omega = T_{\alpha\rho\vartheta} = 2n^2(m - n/3)$ *πράξεις α.κ.υ.*
- *Αν χρειάζεται το Q, πρέπει να πολλαπλασιάσυμε τους ανακλαστές. Αυτό επιφέρει περίπου $2n^2(m - n/3)$ επιπλέον πράξεις, άρα το κόστος διπλασιάζεται.*

□

Παράδειγμα (στιγμιότυπο)

$$A^{(2)} = H_2 H_1 A = \left(\begin{array}{cc|c|cc} x & x & x & x & x \\ 0 & x & x & x & x \\ \hline 0 & 0 & x & x & x \\ 0 & 0 & \star & x & x \\ 0 & 0 & \star & x & x \\ 0 & 0 & \star & x & x \end{array}\right)$$

διαλέγουμε τον u_3 ώστε να μηδενιστούν τα στοιχεία του $A^{(2)}$ σημειωμένα ως «$\star$». Τελικά

$$H_n ... H_1 A = R,$$

και επειδή οι ανακλαστές H_j είναι ορθογώνιοι συμμετρικοί,

$$\begin{aligned} A &= H_1 H_2 ... H_n R \\ &= QR \end{aligned}$$

όπου $Q \in \mathbb{R}^{m\times m}$ ορθογώνιο ($Q^\top Q = I$) και το $R \in \mathbb{R}^{m\times n}$άνω τριγωνικό.
Ορθογώνιοι μετασχηματισμοί και σφάλματα ακυ
Θάπρεπε να εξετάσουμε τα χαρακτηριστικά σε ακυ όλων των ακόλουθων πράξεων: α) Υπολογισμός διανύσματος Householder, β) Εφαρμογή ανακλαστή σε διάνυσμα και σε μητρώο, γ) παραγοντοποίηση QR, και επίλυση συστήματος με QR. Περιοριζόμαστε στα πιο σημαντικά αποτελέσματα:

- Έστω ανακλαστές $H_1^L, \cdots, H_k^L, H_1^R, \cdots, H_k^R$ και σημειώνουμε ως $\tilde{H}$ τις αναπαραστάσεις τους σε ακυ. Επίσης έστω $H^L = H_k^L \cdots H_1^L$ και $H^R = H_1^R \cdots H_k^R$. Τότε

  $$\mathrm{fl}(\tilde{H}_k^L \cdots \tilde{H}_1^L A \tilde{H}_1^R \cdots \tilde{H}_k^R) = H^L(A + E)H^R,$$

  όπου $\|E\|_2 = k\|A\|_2 O(u)$.

Συμπεράσματα Όλα τα βήματα της λύσης μέσω QR με Householder είναι πίσω ευσταθή:

1. Η εφαρμογή ανακλαστών Householder είναι πίσω ευσταθής υπολογισμός, επομένως αυτό ισχύει και για την αναγωγή του A σε άνω τριγωνικό R.

2. ... υπενθυμίζουμε ότι η επίλυση τριγωνικού συστήματος είναι πίσω ευσταθής.

QR και LU Σύγκριση επίλυσης $Ax = b$ μέσω LU σε σχέση με QR:

$$\textbf{LU}: PA = LU \Rightarrow x = U^{-1}(L^{-1}b). \ \textbf{QR}: A = QR \Rightarrow x = R^{-1}(Q^\top b).$$

Συγκρίνουμε για τετραγωνικά προβλήματα που μπορεί να λύσει και η LU.

Μητρώο	Μέθοδος	Πίσω σφάλμα	κ_2	εμπρός σφάλμα
`rand`(40, 40)	MATLAB	2.0563ε-016	1.1636ε+004	2.1566ε-013
	LU με μ.ο.	2.1504ε-016	1.1636ε+004	3.9386ε-013
	LU με π.ο.	1.7061ε-016	1.1636ε+004	4.0071ε-013
	QR	2.4437ε-016	1.1636ε+004	5.6660ε-013
`gfpp`(40)	MATLAB	1.2194ε-007	1.7810ε+001	1.7551ε-006
	LU με μ.ο.	2.2572ε-007	1.7810ε+001	4.2668ε-006
	LU με π.ο.	6.9208ε-017	1.7810ε+001	1.2790ε-015
	QR	1.6951ε-016	1.7810ε+001	3.3777ε-015
`hilb`(20)	MATLAB	1.0110ε-017	1.8458ε+018	2.2835ε+001
	LU με μ.ο.	8.6654ε-018	1.8458ε+018	1.4933ε+001
	LU με π.ο.	9.9313ε-018	1.8458ε+018	9.5113ε+001
	QR	2.4162ε-017	1.8458ε+018	1.5966ε+001

Επίλυση ΑΓΑ.2 με ανακλαστές **Householder**

Βήμα 1. Παραγοντοποίηση $A = QR$. Τα στοιχεία του A έχουν αντικατασταθεί με τα στοιχεία των Q και R.

Βήμα 2.

for $j = 1 : n$
(* Εφαρμογή του υπ. αριθμ. j ανακλαστή. *)
  $u(j) = 1; u(j+1 : m) = A(j+1 : m, j)$
  $b(j : m) = \text{REFL.ROW}(b(j : m), u(j : m))$
end

Βήμα 3. Επίλυση $R(1 : n, 1 : n)x = b(1 : n)$ με πίσω αντικατάσταση.

$\Omega = 2n^2(m - n/3) + O(mn + n^2)$ πράξεις α.κ.υ.

Η QR στη MATLAB

```
QR     Orthogonal-triangular decomposition.
    [Q,R] = QR(A), where A is m-by-n, produces an m-by-n upper triangular
    matrix R and an m-by-m unitary matrix Q so that A = Q*R.

    [Q,R] = QR(A,0) produces the "economy size" decomposition.
    If m>n, only the first n columns of Q and the first n rows of R are
    computed. If m<=n, this is the same as [Q,R] = QR(A).

    If A is full:
```

```
[Q,R,E] = QR(A) produces unitary Q, upper triangular R and a
permutation matrix E so that A*E = Q*R. The column permutation E is
chosen so that ABS(DIAG(R)) is decreasing.

X = QR(A) and X = QR(A,0) return the output of LAPACK's *GEQRF routin
TRIU(X) is the upper triangular factor R.
```

Στην LAPACK Περιοριζόμαστε στις περιπτώσεις που ο A είναι πυκνός και πλήρως αποθηκευμένος, και δεν έχει ιδιαίτερη δομή.

LAPACK:

Για QR: Όταν ο A είναι «πλήρους τάξεως» χρησιμοποιείται η ρουτίνα _GEQRF που υπολογίζει ορμαθοποιημένη QR.

- Ο Q αναπαρασταίνεται σαν γινόμενο στοιχειωδών ανακλαστών και δεν σχηματίζεται άμεσα.
- Η ρουτίνα _ORGQR (_ORMQR) υλοποιούν πολλαπλασιασμό από τα αριστερά (δεξιά) του Q ή $Q^\top$ με μητρώο A.

Για το ΑΓΑ.2:

με QR η LQ	_GELS
με πλήρη ορθογώνια παραγοντοποίηση	_GELSU
με διάσπαση SVD	_GELSS

Στις _GELSU, _GELSS δεν απαιτείται πλήρης τάξη από τον A.

Κλασικός αλγόριθμος Gram-Schmidt

Γεωμετρική περιγραφή: Έστω ότι έχουμε ήδη κατασκευάσει ορθοκανονική βάση $q_1, \cdots, q_{k-1}$ για τις στήλες $a_1, ..., a_{k-1}$ και ότι θέλουμε να υπολογίσουμε το ΟΚ διάνυσμα q_k που είναι ορθογώνιο ως προς τα $q_1, ..., q_{k-1}$. Αν αφαιρέσουμε από το a_k την ορθ. προβολή του επί τα $q_1, .., q_{k-1}$ προκύπτει διάνυσμα παράλληλο στο ζητούμενο q_k. Αν

$$P_j = \frac{q_j q_j^\top}{\underbrace{q_j^\top q^j}_{=1}}$$

συμβολίζει τον τελεστή (ορθογώνιας) προβολής επί του q_j τότε

$$\begin{aligned} \tilde{q}_k &= a_k - P_1 a_k - P_2 a_2 - \cdots - P_{k-1} a_k \\ q_k = \frac{\tilde{q}_k}{\|\tilde{q}_k\|_2} \end{aligned}$$

Διαδοχικά, εφόσον $\rho_{jj} = \|\tilde{q}_j\|_2$ έχουμε

$$\begin{aligned}
q_1 &= a_1/\rho_{11}, \\
q_2 &= (I - P_1)a_2/\rho_{22} \\
... &= ... \\
q_k &= (I - P_{k-1} - \cdots - P_2 - P_1)a_k/\rho_{kk} \\
... &= ... \\
q_n &= (I - P_1 - \cdots - P_{n-1})a_n/\rho_{nn}
\end{aligned}$$

Εκ κατασκευής τα διανύσματα q_k είναι μεταξύ τους ορθοκανονικά.
Σημειώστε ότι δεν θέλουμε να υπολογίσουμε τους τελεστές $I - P_j$ αλλά μόνον την εφαρμογή τους σε διάνυσμα

$$(I - P_j)x = \underbrace{x - q_j(\overbrace{q_j^\top x}^{\text{DOT}})}_{\text{SAXPY}}$$

Από αυτά προκύπτει ο κλασικός αλγόριθμος Gram-Schmidt:

function$[Q, R] = \text{CGS}(A)$
$\rho_{11} = \|a_1\|$ $q_1 = a_1/\rho_{11}$
for $k = 1 : n$
  for $i = 1 : k - 1$
    $\rho_{ik} = q_i^\top a_k$
  end
  $q_k = a_k - \sum_{i=1}^{k-1} \rho_{ik} q_i$
  $\rho_{kk} = \|q_k\|_2$
  $q_k = q_k/\rho_{kk}$
end

Παρατηρήσεις

- κόστος $T^{\text{CGS}}_{\text{αρθ}} = 2mn^2 + O(mn)$.
- Κάθε βήμα υλοποιείται με πράξεις DOT και sAXPY.
- Σε κάθε βήμα του εξωτερικού βρόχου ο CGS υπολογίζει βαθμωτούς ρ_{jk} $(1 \leq j \leq k)$ ώστε

$$a_k = \sum_{j=1}^{k} \rho_{jk} q_k, \quad 1 \leq k \leq n,$$

δηλαδή την «οικονομική» παραγοντοποίηση $A = Q_1 R_1$ όπου $Q_1 = [q_1, ..., q_n]$ και

$$R_1 = \begin{bmatrix} \rho_{11} & \cdots & \rho_{1n} \\ 0 & \ddots & \vdots \\ \ddots & \ddots & \vdots \\ 0 & \cdots & \rho_{nn} \end{bmatrix}$$

- Η GS υπολογίζει ορθοκανονική βάση μόνο για το χώρο στηλών του A. Το Q της πλήρους QR (π.χ. από Householder) περιέχει ΟΚ βάση για όλο το $\Re^m$, δηλαδή το χώρο στηλών και το ορθογώνιο συμπληρωμά του (δηλ. τον αριστερό μηδενόχωρο, βλ. Strang.
- Η GS στοιχίζει περίπου $2mn^2$ και υπολογίζει το Q_1 και το R. Η Householder κοστίζει περίπου $4mn^2$.
- Αν δεν χρειάζεται το ίδιο το Q_1 (π.χ. στα ελάχιστα τετράγωνα), η QR με Householder είναι ταχύτερη.
- Η GS αποκτά περισσότερο ενδιαφέρον όταν το A είναι πολύ αραιό.

Αριθμητικά προβλήματα της κλασικής GS Ο αλγόριθμος χρησιμοποιεί τον τελεστή $(I - P_1)$ επί του a_1, μετά υπολογίζει τον P_2 βασιζόμενος στο αποτέλεσμα της εφαρμογής του $(I - P_1)a_1$, μετά υπολογίζει το $(I - P_1 - P_2)a_2$ κ.ό.κ. Όπως είδαμε αυτές οι πράξεις έχουν κινδύνους.

- Το q_2 μπορεί να μην είναι ορθογώνιο στο q_1.
- Το q_3 δεν θα είναι πια ορθογώνιο στα q_2, q_1, κλπ.

Για παράδειγμα

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$$

Έστω δ τέτοιο ώστε

$$\mathrm{fl}(1 + \delta^2) = 1.$$

Π.χ. στο Toshiba 320CDT (Intel Pentium) έστω $\delta = 7.4506e - 009$. Τότε

$$Q_{\text{"ΓΣ}} = \begin{bmatrix} 1 & 0 & 0 \\ \delta & -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2} \end{bmatrix}$$

Παρατηρούμε πως $q_2^\top q_3 \approx 0.5$, επομένως οι στήλες του Q δεν είναι ΟΚ. Μάλιστα έχουμε

$$\|I - Q_{\text{CGS}}^\top Q_{\text{CGS}}\|_F \approx 0.7071$$

Αξίζει να σημειωθεί ότι αν και οι στήλες του Q_{CGS} δεν είναι ΟΚ, ισχύει ότι

$$\|A - Q_{\text{CGS}}^\top R_{\text{CGS}}\| \approx 0.$$

Εναλλακτική υλοποίηση Πως μπορούμε να αποφύγουμε πράξεις του τύπου $I - P_1 - P_2 - ...$; Στο πρώτο βήμα δεν γίνεται, αλλά από εκεί και πέρα έχουμε την ιδιότητα:

$$q_k^\top q_j = 0 \Leftrightarrow P_k P_j = 0$$

Άρα ισχύουν σχέσεις του τύπου

$$I - P_1 - P_2 = (I - P_2)(I - P_1)$$

Ο νέος αλγόριθμος υπολογίζει με τη σειρά

$$\begin{aligned}
q_1 &= a_1/\rho_{11}, \\
q_2 &= (I - P_1)a_2/\rho_{22} \\
... &= ... \\
q_k &= (I - P_{k-1})\cdots(I - P_2)(I - P_1)a_k/\rho_{kk} \\
... &= ... \\
q_n &= (I - P_{n-1})\cdots(I - P_1)a_n/\rho_{nn}
\end{aligned}$$

Τότε το παραπάνω είναι μαθηματικά ταυτόσημο με το μηχανισμό της κλασικής Gram-Schmidt.

Τροποποιημένος αλγόριθμος Gram-Schmidt

function $[Q, R] = \text{MGS}(A)$
$Q = A$
for $k = 1 : n$
  $\rho_{kk} = \|q_k\|_2$
  $q_k = q_k/\rho_{kk}$
  for $j = k + 1 : n$
    $\rho_{kj} = q_k^\top q_j$
    $q_j = q_j - \rho_{kj} q_k$
  end
end

- Αριθμητικό κόστος ίδιο με την κλασική GS.
- Τα εσωτερικά γινόμενα χρησιμοποιούν τις εν μέρει ορθοκανονικοποιημένες στήλες.
- Η μέθοδος υπολογίζει q_j που είναι πολύ πιο ορθογώνια μεταξύ τους από αυτά της κλασικής GS.
- Τα στοιχεία ρ_{ij} υπολογίζονται ανά γραμμή.
- Η MGS στο προηγούμενο παράδειγμα:

$$\|I - Q_{\text{MGS}}^\top Q_{\text{MGS}}\|_F \approx 6 \times 10^{-9},$$

- Αποδεικνύεται ότι για την MGS και QR με Householder ισχύουν

$$\begin{aligned} \|I - Q_{\mathrm{MGS}}^{\top} Q_{\mathrm{MGS}}\|_2 &\approx \mathbf{u}\kappa_2(A) \\ \|I - Q_{\mathrm{H}}^{\top} Q_{\mathrm{H}}\|_2 &\approx \mathbf{u} \end{aligned}$$

Σχόλιο: Μπορούμε να βελτιώσουμε ακόμα περισσότερο την ορθογωνιότητα του Q που προκύπτει από την MGS επαναλαμβάνοντας τη διαδικασία στο ίδιο το Q_{MGS} (στην πράξη αυτό ενσωματώνεται στη διαδικασία). Κάνουμε δηλαδή «επανορθογωνιοποίηση». Μπορεί να δειχτεί ότι το νέο Q μπορεί να είναι πολύ πιό ορθογώνιο! 2.7531ε+003

Παράδειγμα 2.

$A = \texttt{rand}(100)$, $\kappa_2(A) = 2.71e03$

	$\|A - QR\|_2$	$\|I - Q^{\top}Q\|_2$
Householder	$1.4024e-014$	$2.0119e-015$
CGS	$4.2237e-015$	$4.9100e-012$
MGS	$3.3213e-015$	$5.1946e-014$
MGS+MGS	$3.5175e-015$	$3.5175e-015$

$A = \texttt{hilb}(10)$, $\kappa_2(A) = 1.61e013$

	$\|A - QR\|_2$	$\|I - Q^{\top}Q\|_2$
Householder	$8.1619e-016$	$8.4938e-016$
CGS	$8.9079e-017$	2.6042
MGS	$6.9567e-017$	$1.4336e-004$
MGS+MGS	$6.9567e-017$	$5.9498e-016$

□