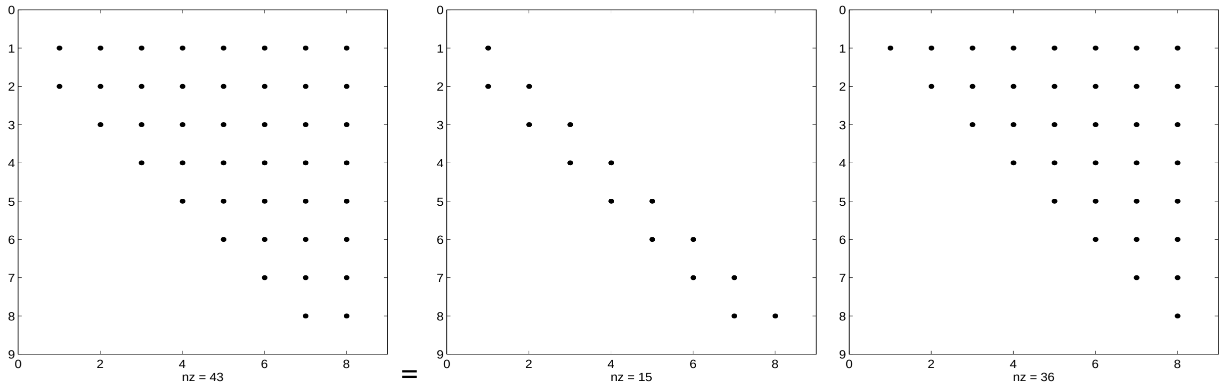
ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

# ΔΙΑΛΕΞΗ 20, 11/01/10

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

Μητρώα Hessenberg Παραγοντοποίηση $LU$ μητρώου άνω Hessenberg:




- Παραγοντοποίηση LU $T_{\text{αρθ}} = O(n^2)$ πράξεις.
- LU με μερική οδήγηση $\Rightarrow \rho \leq n$.

Αναγωγή σε μορφή Hessenberg Πολύ συχνά στις εφαρμογές χρησιμοποιούμε μετασχηματισμούς Householder για να υπολογίσουμε $Q$ το οποίο οδηγεί τον $A$ σε μορφή άνω Hessenberg.

$$Q^\top AQ = H$$

ΠΡΟΣΟΧΗ Επιλέγουμε τους ανακλαστές ώστε να μην καταστρέφουν τα μηδενικά που εισάγουν όταν εφαρμόζονται από τα δεξιά. Αυτό μπορεί να γίνει μόνον αν «χαλαρώσουμε» την απαίτηση για άνω τριγωνικότητα του $Q^\top A$ και επιτρέψουμε μιαν ακόμα υποδιαγώνιο, άρα μορφή Hessenberg.
Στο 1ο βήμα επιλέγουμε διάνυσμα Householder $u_1 = [0, \star, ..., \star]^\top$ για να μηδενίσουμε τα στοιχεία $A(3:n,1)$ αντί των $A(2:n,1)$ που κάνουμε στην τυπική $QR$.
Τότε,

$$(H_{u_1}A)H_{u_1}^\top = \left(\begin{array}{c|ccc} x & x & x & x \\ x & x & x & x \\ \hline 0 & x & x & x \\ 0 & x & x & x \end{array}\right) - \left(\begin{array}{c|ccc} x & x & x & x \\ x & x & x & x \\ \hline 0 & x & x & x \\ 0 & x & x & x \end{array}\right) \begin{pmatrix} 0 \\ x \\ x \\ x \end{pmatrix} \begin{pmatrix} 0 & x & x & x \end{pmatrix}$$

```
HESS Hessenberg form.  The Hessenberg form of a matrix is zero
below the first sub diagonal.  If the matrix is symmetric or
Hermitian, the form is tridiagonal.

[P,H] = HESS(A) produces a unitary matrix P and a Hessenberg
matrix H so that A = P*H*P' and P'*P = EYE(SIZE(P)). By itself,
HESS(A) returns H.
```

Αναγωγή Ο αλγόριθμος ορίζεται αναδρομικά για τετράγωνο μητρώο $A \in \mathbb{R}^{n\times n}$: Έστω ότι πριν την έναρξη του βήματος $k = 1, ..., n-2$ έχουμε ήδη μηδενίσει τα στοιχεία στις θέσεις $(3 : n, 1)$, $(4 : n, 2)$, ..., $(k + 1 : n, k - 1)$ του $A$ με ανακλαστές $H_1, ..., H_{k-1}$:

$$A^{(k)} = H_{k-1} \cdots H_1 A H_1 \cdots H_{k-1}.$$

Τότε στο βήμα $k$ υπολογίζεται ο ανακλαστής $H_k$ που μηδενίζει τα στοιχεία $(k + 2 : n, k)$ του $A^{(k)}$. Παρατηρούμε πως ο πολλαπλασιασμός $(H_k A^{(k)})H_k$ αφήνει τα ήδη εισαχθέντα μηδενικά ανέπαφα, οπότε στο τέλος θα έχουμε πως $A^{(n-2)}$ είναι άνω Hessenberg.
Σχέδιο αλγορίθμου κόστους $T_{\text{αρθ}} \approx \frac{10}{3}n^3$:

**for** $k = 1 : n - 2$
  $H = A$; $u = \text{zeros}(n, 1)$
  $[u(k + 1 : n)] = \text{REFL}(A(k + 1 : n, k))$
  $A(k + 1 : n, k : n) = \text{REFL.ROW}(A(k + 1 : n, k : n), u(k + 1, n))$
  $A(1 : n, k + 1 : n) = \text{REFL.COL}(A(1 : n, k + 1 : n), u(k + 1, n))$
**end**

Εφαρμογές Η αναγωγή σε μορφή Hessenberg με ορθογώνιο μετασχηματισμό έχει πολλές εφαρμογές.
Παράδειγμα Πολλές φορές δίδονται $A$ και $s$ διανύσματα $b_j$ και βαθμωτοί $\omega_j$ και πρέπει να υπολογίσουμε τα $x_j$:

$$(A - \omega_j I)x_j = b_j, j = 1, ..., s$$

(* υπολ. $H$ και (παραγοντοποιημένο) $Q$ *)

$$Q^\top A Q = H$$
$$\Downarrow$$
$$Q^\top (A - \omega_j I)Q = \underbrace{H - \omega_j I}_{\text{άνω Hessenberg}}$$

άρα αντί να λύσουμε συστήματα $(A - \omega I)x_j = b_j$ λύνουμε τα ισοδύναμα συστήματα

$$\underbrace{H - \omega_j I}_{Q^\top (A-\omega_j I)Q}\ Q^\top x_j = Q^\top b_j$$

και έχουμε τα εξής βήματα: _

**Αναγωγή** Υπολογισμός του $H = Q^\top A Q$ και του $\hat{B} = Q^\top [b_1, ..., b_s]$. Αυτό γίνεται μόνο μια φορά και κοστίζει $T_{\text{αρθ}} = \mathrm{O}(n^3)$.

**Επίλυση** $(H - \omega_j I)\hat{x}_j = \hat{b}_j, j = 1, ..., s$. Κόστος $s$O$(n^2)$.

**Κατασκευή** $x_j = Q\hat{x}_j, j = 1, ...,$. Κόστος $s$O$(n^2)$.

Η διαδικασία είναι δυνατή επειδή α) το $H$ Hessenberg $\Rightarrow H - \alpha I$ Hessenberg, β)η παραγοντοποίηση και Hessenberg κοστίζει $\Omega =$ O$(n^2)$, και γ) το $Q$ είναι ορθογώνιο.
Αν εφαρμόζαμε $LU$, για κάθε $\omega_j$ θα έπρεπε να υπολογίσουμε την παραγοντοποίηση $L(\omega_j)U(\omega_j) = A - \omega_j I$, άρα συνολικό κόστος $sO(n^3)$. Άλλα θέματα σχετικά με ΑΓΑ.2

- Householder κατά ορμαθούς για υλοποιήσεις μέσω BLAS-3
- $QR$ με οδήγηση που αποκαλύπτει την τάξη (rank-revealing)
- Περιστροφές Givens
- Αξιοποίηση αραιής μορφής και sparse QR

Πρόβλημα ιδιοτιμών και ιδιοδιανυσμάτων (Algebraic Eigenvalue Problem)

Δίνεται•μητρώο $A \in \mathbb{R}^{n\times n}$ και αναζητούμε βαθμωτούς $\lambda$ (ιδιοτιμές) και τα αντίστοιχα διανύσματα $x \neq 0$ (ιδιοδιάνυσμα) τ.ώ. $Ax = \lambda x$.

- Ιδιοτιμές του $A$ είναι οι $n$ ρίζες του χαρακτηριστικού πολυώνυμου

$$\det(A - \lambda I) = 0$$

- Φάσμα ονομάζεται το σύνολο των ιδιοτιμών, συχνά συμβολίζεται $\sigma(A)$ ή $\Lambda(A)$.

Προσοχή: Ως ρίζες πολυωνύμου, γενικά δεν υπάρχουν «πεπερασμένες μεθόδοι» υπολογισμού ιδιοτιμών, δηλ. μέθοδοι που με αριθμητική άπειρης ακρίβειας υπολογίζουν ακριβώς τις ιδιοτιμές σε πεπερασμένο αριθμό πράξεων).

> *... the computation of eigenvalues from the characteristic polynomial is one of the best known stupidities of numerical analysis. Good numerical analysis turns it the other way round: the real matrix A is directly reduced, first to Hessenberg form, then by a sequence of orthogonal transformations to the real Schur form .... [Hairer,Nørsett,Wanner, Solving Ordinary Differential Equations I: Nonstiff Problems]*

Παρατηρήσεις

- «Στην καρδιά» μυριάδων εφαρμογών του Επιστημονικού Υπολογισμού (π.χ. Διαφορικές, Υπολογιστική Στατιστική)
- Πρόσφατα «διάσημο» στο ευρύ κοινό από την εφαρμογή του στη μέθοδο PageRank της Google...

- .... το ιδιοδιάνυσμα των $25.000.000.000 δολλαρίων!!
- Πολλές μέθοδοι υπολογισμού, με ειδικά χαρακτηριστικά
- .... πεδίο εφαρμοσιμότητας ανάλογα με τα χαρακτηριστικά του προβλήματος
- ... πραγματική συμμετρία, μέγεθος, αραιή δομή, πυκνή δομή, ....
- και το ζητούμενο: Ιδιοτιμές (όλες, μερικές, ... ), Ιδιοδιανύσματα (όλα, μερικά, ... )
- Περισσότερα στον «Επιστημονικό Υπολογισμό ΙΙ»...

Σχέση με SVD: Αν έχουμε αλγορίθμους για υπολογισμό ιδιο/τιμών-διανυσμάτων, εφαρμόζονται στον υπολογισμό του SVD:

$$A \;=\; U\Sigma V^\top \Rightarrow A^\top A = V\Sigma^2 V^\top, \;\; AA^\top = U\Sigma^2 U^\top$$

επομένως μπορούμε να εφαρμόσουμε τον αλγόριθμο υπολογισμού ιδιοτιμών στο $A^\top A$ για να υπολογίσουμε τα $V, \Sigma$ ή στο $AA^\top$ για υπολογισμό των $U, \Sigma$. Προσοχή: Ο πολλαπλασιασμός του $A$ με το ανάστροφο ενέχει κινδύνους (όπως και με τη χρήση «κανονικών εξισώσεων» ....
Αριστερά ιδιοδιανύσματα Ποιά είναι η σχέση των ιδιοτιμών/διανυσμάτων του $A$ με αυτά του $A^*$;

- Αν $Ax = \lambda x$ τότε $x^*A^* = \bar{\lambda}x^*$επομένως το $\bar{\lambda}$ είναι ιδιοτιμή του $A^*$
- Επομένως οι ιδιοτιμές του $A^*$ είναι οι συζυγείς των ιδιοτιμών του $A$.

Γενικά δεν υπάρχει απλή σχέση μεταξύ των ιδιοδιανυσμάτων των $A, A^*$.

- Αν $\bar{\lambda}$ είναι ιδιοτιμή του $A^*$ τότε θα υπάρχει ιδιοδιάνυσμα $y$ ώστε $A^*y = \bar{\lambda}y$
- επομένως $y^*A = \lambda y^*$

το $y^*$ αποκαλείται αριστερό ιδιοδιάνυσμα του $A$ που αντιστοιχεί στο $\lambda$.

Σχέση ιδιοτιμών/ιδιοδιανυσμάτων

- Αν υπολογίσουμε μια ιδιοτιμή, τα αντίστοιχα ιδιοδιανύσματα μπορούν να υπολογιστούν από τις σχέσεις $(A-\lambda I)x = 0$ και $y^*(A-\lambda I) = 0$. (... τα μητρώα δεν είναι αντιστρέψιμα!)
- Αν υπολογίσουμε ένα ιδιοδιάνυσμα, $x$, τότε μια αντίστοιχη ιδιοτιμή μπορεί να υπολογιστεί ως $\lambda = \frac{x^\top Ax}{x^\top x}$.

Λογισμικό Βασικά βήματα:

1. αναγωγή του $A$ σε απλούστερη[1] μορφή (π.χ. άνω Hessenberg) με ορθογώνιους μετασχηματισμούς (νέο σύστημα αναφοράς)
2. επίλυση προβλήματος ιδιοτιμών για το συμπιεσμένο μητρώο
3. (αν χρειάζεται) μετασχηματισμός της λύσης στο σύστημα αναφοράς του αρχικού μητρώου

**LAPACK** Δείτε `http://www.netlib.org/lapack/lug/node70.html`

**MATLAB `eig`:** Αλγόριθμοι LAPACK: Βασισμένοι στον αλγόριθμο $QR$ κατάλληλοι για μητρώα μέτριου μεγέθους χωρίς ιδιαίτερη δομή. Υπολογίζουν α) όλες τις ιδιοτιμές, β) αν επιλέξουμε, όλα τα ιδιοδιανύσματα.

**MATLAB `svd`:** Όπως η `eig` για το SVD.

**MATLAB `eigs`:** Αλγόριθμοι βασισμένοι σε προχωρημένες επαναληπτικές μεθόδους (Implicitly Restarted Arnoldi), κατάλληλες για μεγάλα αραιά μητρώα. Υπολογίζουν επιλεγμένες ιδιοτιμές και ιδιοδιανύσματα.

**MATLAB `svds`:** Όπως η `eigs` για το SVD.

Σκιαγράφηση αλγορίθμου $QR$ για ιδιοτιμές[2]

*Πιο σημαντικός «αλγόριθμος» εύρεσης ιδιοτιμών που βασίζεται σε επανειλημμένες διασπάσεις $QR$ - αποτελεί τον κύριο τρόπο υπολογισμού όλων των (ιδιοτιμών, ιδιοδιανυσμάτων) γενικών μητρώων.*

Βασική ιδέα: 1) $A = QR, B := RQ \Rightarrow B = Q^\top AQ \Rightarrow \Lambda(A) = \Lambda(B)$

2)
```
A^(0) := A, k=0
repeat
    k = k+1
    [Q, R] = qr(A^(k-1))
    A^(k) = mtimes(R, Q)
until convergence
```

- Παρατηρήθηκε ότι $A^{(k)} \to T$, σχεδόν άνω τριγωνικό[3], όμοιο του $A$
- Το $T$ είναι σχεδόν άνω τριγωνικό, όμοιο με το $A$: οι ιδιοτιμές «αποκαλύπτονται» στη διαγώνιο ή εμπεριέχονται (ως ζεύγη συζυγών μιγαδικών) σε μικρά ($2 \times 2$) μητρώα στη διαγώνιο του $T$
- Χρειάζονται πολλά για να καταστεί πρακτικός αλγόριθμος: **I.** Μείωση κόστους επανάληψης, **II.** Διερεύνηση σύγκλισης (συγκλίνει; Πόσο γρήγορα; $\to$ ΕΥΠ).

[1] συμπυκνωμένη - condensed
[3] Μορφή Schur

Κρίσιμο στοιχείο υλοποίησης → μείωση κόστους επανάληψης

*Το κόστος κάθε επανάληψης φαίνεται μεγάλο:* $\approx \frac{4}{3}n^3 + n^3$

*... επειδή το $A$ δεν έχει ιδιαίτερη δομή*

**ΙΔΕΑ** πριν ξεκινήσουν οι επαναλήψεις, μετασχηματίζουμε το $A$ σε άνω Hessenberg με ομοιότητα

**...** γίνεται άπαξ, κόστος $\approx \frac{10}{3}n^3$

**ΚΛΕΙΔΙ** Έστω ότι $H$ άνω Hessenberg και ότι

$$[Q, R] = \mathtt{qr}(H), \hat{H} = RQ,$$

Τότε το $\hat{H}$ παραμένει άνω Hessenberg.

**ΕΠΙΣΗΣ** 1) Η παραγοντοποίηση $QR$ μητρώου Hessenberg κοστίζει $O(n^2)$. Το $Q$ επιστρέφεται σε παραγοντοποιημένη μορφή ($n-1$ περιστροφές Givens). 2) Τότε ο υπολογισμός $RQ$ πραγματοποιείται με $O(n^2)$ πράξεις.

*Κάθε επανάληψη της $QR$ μπορεί να γίνει με $\Omega = O(n^2)$ πράξεις. Αν $A = A^\top$ τότε $\Omega = O(n)$.*

Μέθοδος Δύναμης (Power Method)

Απλή, επαναληπτική μέθοδος για τη βαθμιαία προσέγγιση (της κατεύθυνσης) του κυρίαρχου ιδιοδιανύσματος

... και για (συμπτωματικό) υπολογισμό της κυρίαρχης ιδιοτιμής

Προσοχή: Επειδή $Ax = \lambda x \Rightarrow A(x/\gamma) = \lambda(x/\gamma)$ για $\gamma \neq 0$, για να έχουμε μοναδικότητα χρειάζεται τουλάχιστον κανονικοποίηση, π.χ. το $x$ να ικανοποιεί $\|x\|_2 = 1$.

*Σε πάρα πολλές περιπτώσεις, αυτό που ενδιαφέρει σε ένα ιδιοδιάνυσμα είναι η κατεύθυνση και όχι το μέτρο του.*

Έστω $A$ με ιδιοτιμές $\sigma(A) = \{\lambda_j\}$, όπου

$$\underbrace{|\lambda_1|}_{\text{κυρίαρχη ιδιοτιμή}} > |\lambda_2| \geq |\lambda_3| > \cdots \geq |\lambda_n|$$

και με γ.α. ιδιοδιανύσματα $u_1, ..., u_n$

*Αλγόριθμος.* [*Μέθοδος δύναμης*]

Εκκίνηση: Τυχαίο διάνυσμα $x := x_0$
**for** $k = 1, ...$
$x \leftarrow Ax$
**end**

$$\begin{aligned} x &= \xi_1 u_1 + \cdots + \xi_n u_n \\ A^k x &= \xi_1 A^k u_1 + \cdots + \xi_n A^k u_n \\ &= \xi_1 \lambda_1^k u_1 + \cdots + \xi_n \lambda_n^k u_n \end{aligned}$$

$$\frac{1}{\lambda_1^k} A^k x = \xi_1 u_1 + \xi_2 (\frac{\lambda_2}{\lambda_1})^k u_2 + \cdots + \xi_n (\frac{\lambda_n}{\lambda_1})^k u_n$$

Αν $\xi_1 \neq 0$, $A^k x$ τείνει να γίνει παράλληλο με το $u_1$. Η ταχύτητα με την οποία γίνεται αυτό (όταν γίνεται, έχουμε σύγκλιση) εξαρτάται σε μεγάλο βαθμό από την ποσότητα $\frac{\lambda_2}{\lambda_1}$
.... όσο μικρότερη, τόσο καλύτερα ... προβλήματα όμως αν $|\lambda_2| = |\lambda_1|$.)

Παρατηρήσεις

- Αν $\xi_1 \neq 0$ και $[u_1]_j \neq 0$ τότε
$$\frac{\xi_j^{(k+1)}}{\xi_j^{(k)}} = \lambda_1 + \mathrm{O}(|\frac{\lambda_2}{\lambda_1}|^k)$$

Προς αποφυγή υπερ/υποχείλισης:
Εκκίνηση: Τυχαίο διάνυσμα $x := x_0$
**for** $k = 1, ...$
$t \leftarrow Ax$
$x \leftarrow \frac{1}{\gamma_j} t$
**end**
με κατάλληλα επιλεγμένα $\gamma_j$ π.χ. $\gamma_j := \|u\|_\infty$
Προσοχή: Αν γνωρίζουμε ότι $\lambda_1 = 1 > \lambda_2 \geq \cdots \geq \lambda_n$, δεν χρειάζεται κανονικοποίηση! Αυτή η ιδιότητα υπάρχει στο μητρώο Google που χρησιμοποιείται για το PageRank...