Επιστημονικός Υπολογισμός Ι

Ενότητα **7** : Διαχείρηση Μητρώων Ειδικής Δομής

Ευστράτιος Γαλλόπουλος
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης
ΕΣΠΑ 2007-2013
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο «**Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Πατρών**» έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σκοπός Ενότητας

- Μητρώα ζώνης
- Αραιά μητρώα
- Βιβλιοθήκη LAPACK
- Στοιχεία επαναληπτικών μεθόδων για την επίλυση γραμμικών συστημάτων

Περιεχόμενα

Υπενθύμιση

Γενική συζήτηση περί διαφορετικών μεθόδων Γενικά αραιά μητρώα: Παραδείγματα 2 ;αραιών δομών; αποθήκευσης. Πολλαπλασιασμός μητρώου-διανύσματος για μητρώο σε μορφή CSR και CSC.

Στοιχεία επαναληπτικών μεθόδων (σε λίγες διαφάνειες!) Γενική δομή επαναληπτικών μεθόδων. Κλασικές μέθοδοι Jacobi, Gauss-Seidel και η επαναληπτική εκλέπτυνση ως μέθοδος της κατηγορίας αυτής. Τύπος Sherman-Morrison.. Γενική περιγραφή μεθόδων Krylov.

Επαναληπτικές μέθοδοι: Χονδρική κατηγοριοποίηση

Κλασικές	λύνουμε ως προς προσέγγιση του A και διορθώνουμε
Προβολής	προσεγγίζουμε τη λύση από υπόχωρους (Krylov)

Επαναληπτικές μέθοδοι Krylov

Θεώρημα Cayley-Hamilton Αν $p(z) = \det(A - zI) = z^n + \sum_{j=0}^{n-1} \gamma_j z^j$ τότε

$$p(A) = 0 \Rightarrow A^{-1} = -\frac{1}{\gamma_0}\left(A^{n-1} + \gamma_{n-1}A^{n-2} + \cdots \gamma_1\right)$$

Επαναληπτικές μέθοδοι Krylov

Θεώρημα Cayley-Hamilton Αν $p(z) = \det(A - zI) = z^n + \sum_{j=0}^{n-1} \gamma_j z^j$ τότε

$$\begin{aligned} p(A) &= 0 \Rightarrow A^{-1} = -\frac{1}{\gamma_0}\left(A^{n-1} + \gamma_{n-1}A^{n-2} + \cdots \gamma_1\right) \\ A^{-1}b &= [b, Ab, \cdots, A^{n-1}b]c \quad \text{για } c = -\frac{1}{\gamma_0}[\gamma_1, \cdots, \gamma_{n-1}, 1]^\top, \text{ μη πρακτικό} \end{aligned}$$

ΜΕΘΟΔΟΙ Krylov Αναζητούμε προσέγγιση

$$\begin{aligned} x^{(m)} &\in K_m(A; b) = \overbrace{\mathrm{span}\{b, Ab, \ldots, A^{m-1}b\}}^{\text{υπόχωρος Krylov}} \\ &= V_m y^{(m)} \text{ όπου } V_m \text{ ΟΚ βάση του } K_m(A; b) \end{aligned}$$

Θεώρημα Cayley-Hamilton Αν $p(z) = \det(A - zI) = z^n + \sum_{j=0}^{n-1} \gamma_j z^j$ τότε

$$\begin{aligned} p(A) &= 0 \Rightarrow A^{-1} = -\frac{1}{\gamma_0}\left(A^{n-1} + \gamma_{n-1}A^{n-2} + \cdots \gamma_1\right) \\ A^{-1}b &= [b, Ab, \cdots, A^{n-1}b]c \quad \text{για } c = -\frac{1}{\gamma_0}[\gamma_1, \cdots, \gamma_{n-1}, 1]^\top, \text{ μη πρακτικό} \end{aligned}$$

ΜΕΘΟΔΟΙ Krylov Αναζητούμε προσέγγιση

$$\begin{aligned} x^{(m)} &\in K_m(A;b) = \overbrace{\mathrm{span}\{b, Ab, \ldots, A^{m-1}b\}}^{\text{υπόχωρος Krylov}} \\ &= V_m y^{(m)} \text{ όπου } V_m \text{ ΟΚ βάση του } K_m(A;b) \end{aligned}$$

Για πλήρη καθορισμό του $x^{(m)}$ περιορίζουμε ώστε

$$r^{(m)} = b - Ax^{(m)} \perp K_m(A;b) \Rightarrow 0 = V_m^\top(b - AV_m y^{(m)})$$

Θεώρημα Cayley-Hamilton Αν $p(z) = \det(A - zI) = z^n + \sum_{j=0}^{n-1} \gamma_j z^j$ τότε

$$p(A) = 0 \Rightarrow A^{-1} = -\frac{1}{\gamma_0}\left(A^{n-1} + \gamma_{n-1}A^{n-2} + \cdots \gamma_1\right)$$

$$A^{-1}b = [b, Ab, \cdots, A^{n-1}b]c \quad \text{για } c = -\frac{1}{\gamma_0}[\gamma_1, \cdots, \gamma_{n-1}, 1]^\top, \ \text{μη πρακτικό}$$

ΜΕΘΟΔΟΙ Krylov Αναζητούμε προσέγγιση

$$\begin{aligned} x^{(m)} &\in K_m(A; b) = \overbrace{\mathrm{span}\{b, Ab, \ldots, A^{m-1}b\}}^{\text{υπόχωρος Krylov}} \\ &= V_m y^{(m)} \ \text{όπου } V_m \text{ ΟΚ βάση του } K_m(A; b) \end{aligned}$$

Για πλήρη καθορισμό του $x^{(m)}$ περιορίζουμε ώστε

$$r^{(m)} = b - Ax^{(m)} \perp K_m(A; b) \Rightarrow 0 = V_m^\top(b - AV_m y^{(m)})$$

Το αρχικό πρόβλημα $Ax = b$ προσεγγίζεται μέσω επίλυσης συστήματος $m \times m$ (αν αντιστρέψιμο):

$$\boxed{(V_m^\top A V_m)y^{(m)} = V_m^\top b \Rightarrow x^{(m)} = V_m y^{(m)}} \rightarrowtail$$

πολλά σημαντικά ζητήματα, π.χ. (Saa03)

Υπολογιστικά θέματα

Αποθήκευση αραιό μητρώο $\to$ φθηνή αποθήκευση

Δυσκολία προγραμματισμού χρειάζεται ειδικός που εξαρτάται από την αραιή δομή αποθήκευσης. Η MATLAB απλουστεύει γιατί αναφερόμαστε σε αραιά μητρώα ως πυκνά. π.χ. το `A(i,j)` εξακολουθεί να σημαίνει το στοιχείο στη θέση (i, j) του μητρώου A αν και εσωτερικά, υλοποιείται εντελώς διαφορετικά.

Μειωμένη απόδοση (MV) με αραιό $A \in \mathbb{R}^{n\times n}$ με 5 περίπου στοιχεία ανά γραμμή (π.χ. `gallery('poisson',N)` όπου $n = \mathtt{N}^2$). Τότε

$$\mu_{\text{min}}^{\text{sparse}} = 4/5$$

Αν ήταν πυκνό, η τιμή για MV θα ήταν $\mu_{\text{min}} = (n^2 + 3n)/2n^2 \approx 1/2$. Δηλαδή, η τοπικότητα είναι μικρότερη λόγω αραιότητας.

Κόστος MV $O(\mathtt{nnz}(A))$

Μέθοδος συζυγών κλίσεων για ΣΘΟ ορισμένα μητρώα

Θεώρηση βελτιστοποίησης Αν A ΣΘΟ τότε η λύση του συστήματος επιλύει και ένα πρόβλημα βελτιστοποίησης:

$$Ax = b \Leftrightarrow x = \arg\min_{y \in \mathbb{R}^n} \frac{1}{2} y^\top A y - b^\top y$$

ειδικότερα, επιζητείται το x που επιτυγχάνει την ελαχιστοποίηση της τετραγωνικής μορφής $\phi(y) = \frac{1}{2}(Ay, y) - (b, y)$.

Ιδιότητες Κυρτή συνάρτηση, το τοπικό ελάχιστο είναι και ολικό ελάχιστο.

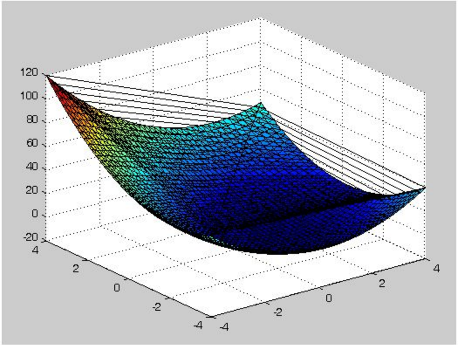
Γενικός αλγόριθμος εύρεσης ελαχίστου Εκκινώντας από $x^{(0)}$, σε κάθε βήμα $k = 1, \ldots$ «διορθώνουμε» την προσέγγιση $x^{(k+1)} \leftarrow x^{(k)} + d^{(k)}$ ώστε $\phi(x^{(k+1)}) < \phi(x^{(k)})$. Όταν δεν υπάρχει τέτοια διόρθωση, έχει υπολογιστεί η λύση. Εναλλακτικά, όταν το $\phi(x^{(k+1)})$ είναι αρκετά μικρό τερματίζουν οι επαναλήψεις.

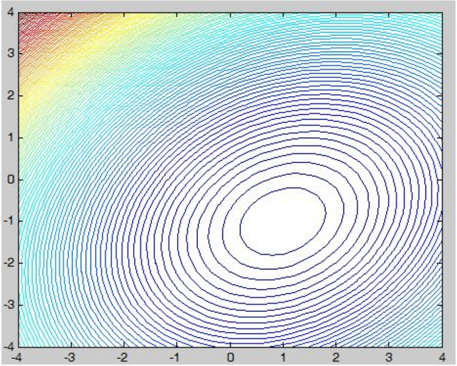
Επιλογές Στη CG $d^{(k)} = x^{(k)} + \alpha_k p^{(k)}$ όπου το διάνυσμα κατεύθυνσης $p^{(k)}$ και ο συντελεστής α_k επιλέγονται με ειδικό τρόπο.

Για περισσότερα: βλ. καλό λήμμα Wikipedia και ελάτε στον Επιστημονικό II για εμβάθυνση στο θέμα.

Οπτικοποίηση

Οπτικοποίηση της $\phi(y)$και των ισοϋψών της

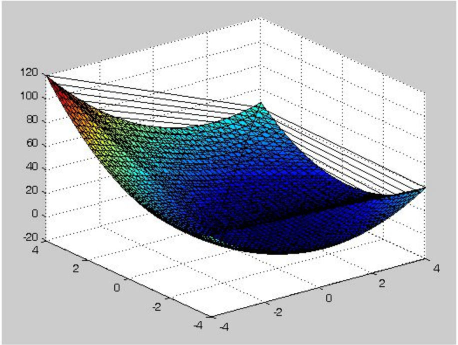


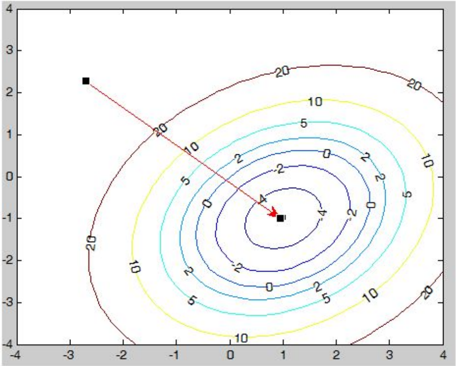


$$A = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}, b = \begin{pmatrix} 5 & -5 \end{pmatrix} \Rightarrow x = \begin{pmatrix} 1 & -1 \end{pmatrix} \Rightarrow \phi(x) = -5$$

Οπτικοποίηση

Οπτικοποίηση της $\phi(y)$και των ισοϋψών της





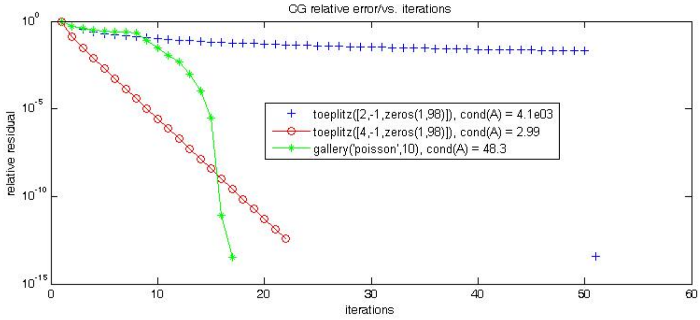
$$A = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}, b = \begin{pmatrix} 5 & -5 \end{pmatrix} \Rightarrow x = \begin{pmatrix} 1 & -1 \end{pmatrix} \Rightarrow \phi(x) = -5$$

```
function [x,flag,relres,iter,resvec] = ...
    pcg(A,b,tol,maxit,M1,M2,x0,varargin)
%PCG   Preconditioned Conjugate Gradients Method.
%    X = PCG(A,B) attempts to solve system A*X=B
%    A must be symmetric and positive definite
%    X = PCG(AFUN,B) accepts function handle AFUN instead.
%    AFUN(X) accepts vector X and returns A*X
%    In all of the following one can replace A by AFUN.
%    TOL specifies the tolerance of the method.
%    MAXIT specifies the max iterations.
%    X0 specifies initial guess.
%    [] then PCG uses the default, an all zero vector.
```

Μία απλή εκδοχή

```
function [x, relres, iter, flag] = cg(A,x,b,max_it,tol)
flag = 0;iter = 0; alpha = 0.0;beta  = 0.0;
bnrm2 = 0.0; relres = [0];rho   = 0.0;[n,n] = size(A);
p= zeros(n,1);q=p;r=p; bnrm2=norm(b);
if  ( bnrm2 == 0.0 ), bnrm2 = 1.0; end
r = b - A*x; relres = [norm( r )/bnrm2];
if ( relres(end) < tol ) return, end
for iter = 1:max_it
   rho = (r'*r); % 1 DOT
 if ( iter > 1 ), beta = rho/rho_1; p = r + beta*p; % 1 ...
      SAXPY
   else, p = r; end
   q = A*p; alpha = rho / (p'*q ); % 1 MV + 1 DOT
   x = x + alpha * p; r = r - alpha*q; % 2 SAXPY
   relres = [relres,norm( r )/bnrm2]; % 1 DOT
   if ( relres(end) <= tol ), break, end
   rho_1 = rho;
end
if ( relres(end) > tol ) flag = 1; end;
```

Παραδείγματα



Σχήμα: Οπτικοποίηση σχετικού καταλοίπου $\|b - Ax_k\|_2 / \|b\|_2$ για $k = 1, \dots$ ως τη σύγκλιση (σχετικό κατάλοιπο $< 10^{-12}$.) Ο ρυθμός σύγκλισης φαίνεται να επηρεάζεται από το δείκτη κατάστασης (πάλι!!!!) Υπάρχουν και άλλοι παράγοντες (ΕΥ-ΙΙ).

Μερικά στοιχεία για τη CG

- Κόστος: (1 MV + 3 DOT + 3 SAXPY) ανά επανάληψη. Έστω ότι εκτελούνται Iter επαναλήψεις.
- Αν Iter $\leq n$, σε ακριβή αριθμητική η CG παράγει την ακριβή λύση.
- Αν το μητρώο έχει ειδική δομή, κάθε MV μπορεί να εκτελεστεί με λιγότερες από $2n^2$ πράξεις και η μέθοδος να είναι οικονομικότερη της Cholesky.
- Συνήθως το A, αραιό, οπότε $\Omega \approx \mathrm{Iter} \times (2\mathtt{nnz}(\mathtt{A}) + 12n)$
- Παλαιά μέθοδος (1958) που προτάθηκε ως εναλλακτική της απαλοιφής Gauss ...
- θεωρήθηκε μη πρακτική: κάθε βήμα απαιτεί $2n^2 + O(n)$ πράξεις για γενικά μητρώα.
- Αναθεώρηση (αρχές 1970): παρατηρήθηκε ότι κάθε επανάληψη, συνήθως βελτιώνει την προσέγγιση στη λύση
- Μεγάλο ενδιαφέρον γιατί δεν απαιτείται ρητή γνώση του A (matrix-free), αλλά αρκεί να υπάρχει τρόπος να υπολογιστεί το Ax για κάθε διάνυσμα x.
- Εφαρμόζεται και για να λύσουμε τις κανονικές εξισώσεις γιατί συνήθως $A^\top A$ είναι ΣΘΟ και δεν χρειάζεται να κατασκευάσουμε το $A^\top A$ (γιατί;)

Πρόβλημα ιδιοτιμών και ιδιοδιανυσμάτων (Algebraic Eigenvalue Problem)

Για μητρώο $A \in \mathbb{R}^{n \times n}$

- Αναζητούμε βαθμωτούς λ (ιδιοτιμές) και αντίστοιχα διανύσματα $x \neq 0$ (ιδιοδιάνυσμα) τ.ώ. $Ax = \lambda x$.
- Αναζητούμε μητρώο X και διαγώνιο Λ ώστε $A = X\Lambda X^{-1}$ (δεν υπάρχει πάντα τέτοιο X, π.χ. το $A = \begin{pmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{pmatrix}$ δεν διαγωνιοποιείται.
- Αναζητούμε ορθομοναδιαίο Q ώστε $A = QTQ^*$ όπου T άνω τριγωνικό (παραγοντοποίηση Schur)
- Αναζητούμε U, Σ, V ώστε $A = U\Sigma V^\top$ και U, V ορθογώνια, $\Sigma \geq 0$ διαγώνιο (SVD)
- Γρήγορη εισαγωγή στη Wikipedia

Για το πρόβλημα ιδιοτιμών

- 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: Γενικά δεν υπάρχουν «πεπερασμένες μεθόδοι» υπολογισμού ιδιοτιμών, δηλ. μέθοδοι που με αριθμητική άπειρης ακρίβειας υπολογίζουν ακριβώς τις ιδιοτιμές (λύσεις πολυωνύμων βαθμού 5 και άνω) με τύπους εκφρασμένους με ριζικά, σε πεπερασμένο αριθμό πράξεων (Abel, Ruffini 1824, Galois posth. 1846)

> *... 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)*

Ίσως ο μεγαλύτερος και πιο διάσημος υπολογισμός με μητρώα στον κόσμο

αφορά υπολογισμό ιδιοδιανύσματος

12 CLEVE'S CORNER

THE WORLD'S LARGEST MATRIX COMPUTATION

Google's PageRank is an eigenvector of a matrix of order 2.7 billion.

BY CLEVE MOLER

One of the reasons why Google is such an effective search engine is the PageRank™ algorithm, developed by Google's founders, Larry Page and Sergey Brin, when they were graduate students at Stanford University. PageRank is determined entirely by the link structure of the Web. It is recomputed about once a month and does not involve any of the actual content of Web pages or of any individual query. Then, for any particular query, Google finds the pages on the Web that match that query and lists those pages in the order of their PageRank.

Imagine surfing the Web, going from page to page by randomly choosing an outgoing link from one page to get to the next. This can lead to dead ends at pages with no outgoing links, or cycles around cliques of interconnected pages. So, a certain fraction of the time, simply choose a random page from anywhere on the Web. This theoretical random walk of the Web is a *Markov chain* or *Markov*

It tells us that the largest eigenvalue of A is equal to one and that the corresponding eigenvector, which satisfies the equation

$$x = Ax,$$

exists and is unique to within a scaling factor. When this scaling factor is chosen so that

$$\sum_i x_i = 1$$

then x is the state vector of the Markov chain. The elements of x are Google's PageRank.

If the matrix were small enough to fit in MATLAB, one way to compute the eigenvector x would be to start with a good approximate solution, such as the PageRanks from the previous month, and simply repeat the assignment statement

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»...

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

$A^{(0)} := A$, k=0

repeat

  k = k+1

  $[Q, R] = \texttt{qr}(A^{(k-1)})$

  $A^{(k)} = \texttt{mtimes}(R, Q)$

until convergence

- Παρατηρήθηκε ότι $A^{(k)} \to T$, σχεδόν άνω τριγωνικό[1], όμοιο του A
- Το T είναι σχεδόν άνω τριγωνικό, όμοιο με το A: οι ιδιοτιμές «αποκαλύπονται» στη διαγώνιο ή εμπεριέχονται (ως ζεύγη συζυγών μιγαδικών) σε μικρά (2×2) μητρώα στη διαγώνιο του T
- Χρειάζονται πολλά για να καταστεί πρακτικός αλγόριθμος: **I.** Μείωση κόστους επανάληψης, **II.** Διερεύνηση σύγκλισης (συγκλίνει; Πόσο γρήγορα; $\to$ ΕΥΙΙ).

[1] Μορφή Schur

Απλή, επαναληπτική μέθοδος για τη βαθμιαία προσέγγιση (της κατεύθυνσης) του κυρίαρχου ιδιοδιανύσματος, ιδιαίτερα βολική για μεγάλα αραιά μητρώα.

... 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 περισσότερο σε ένα ιδιοδιάνυσμα είναι η κατεύθυνση παρά 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 γραμμικά ανεξάρτητα ιδιοδιανύσματα $u_1, \ldots, u_n$

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

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

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

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 από την ποσότητα $\frac{\lambda_2}{\lambda_1}$: όσο μικρότερη, τόσο καλύτερα ... (όμως τι γίνεται όταν $|\lambda_2| = |\lambda_1|$;)

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:

Require: Εκκίνηση: Τυχαίο διάνυσμα $x := x_0$
  for $k = 1, \ldots$ **do**
    $t \leftarrow Ax$
    $x \leftarrow \frac{1}{\gamma_j} t$
  end for

με κατάλληλα επιλεγμένα γ_j π.χ. $\gamma_j := ||t||_\infty$

Λογισμικό

Βασικά βήματα (θεωρία και βιβλιογραφία στο βασικό σύγγραμμα (GV12)):

1. αναγωγή του A σε απλούστερη 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.

Βιβλιογραφία I

G.H. Golub and C.F. Van Loan.
Matrix Computations.
The Johns Hopkins University Press, Baltimore, 3d edition, 2012.

Y. Saad.
Iterative Methods for Sparse Linear Systems.
SIAM, Philadelphia, 2003.

Ε. Γαλλόπουλος.
Επιστημονικός Υπολογισμός Ι.
Πανεπιστήμιο Πατρών, 2008.

Χρήση Έργου Τρίτων I

1. http://www.mathworks.com/ (βλ. σελ 11)
2. http://www.mathworks.com/company/newsletters/articles/the-world-s-largest-matrix-computation.html (βλ. σελ 17)

Σημείωμα Αναφοράς

Τέλος Ενότητας

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης