Επιστημονικός Υπολογισμός Ι

Ενότητα **6** : Παραγοντοποίηση QR και Ελάχιστα Τετράγωνα

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
ΕΣΠΑ 2007-2013
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

’Αδειες Χρήσης

- Το παρόν εκπαιδευτικό υλικό υπόκειται σε άδειες χρήσης Creative Commons.
- Για εκπαιδευτικό υλικό, όπως εικόνες, που υπόκειται σε άλλου τύπου άδειας χρήσης, η άδεια χρήσης αναφέρεται ρητώς.

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.

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σκοπός Ενότητας

- Παραγοντοποίηση QR
- Στοιχειώδη μητρώα Householder (ανακλαστές)
- Ορθογώνιες προβολές
- Επίλυση προβλημάτων ελαχίστων τετραγώνων μέσω QR

Περιεχόμενα

Επίλυση ΑΓΑ.2 με QR

Αν $A \in \mathbb{R}^{m\times n}$ όπου $m > n$ τότε η παραγοντοποίηση QR μας επιστρέφει $Q \in \mathbb{R}^{m\times n}$ και $R \in \mathbb{R}^{n\times n}$. Αν το μητρώο A είναι πλήρους τάξης, τότε οι πρώτες n στήλες του Q (ας πούμε Q_1, είναι ΟΚ βάση για το διανυσματικό υποχώρο διάστασης n του $\mathbb{R}^m$ που παράγεται από τις στήλες του A Έστω ότι $Q = [Q_1, Q_2]$ Τότε

$$Q^\top A = R = \begin{pmatrix} R_1 \\ 0 \end{pmatrix}$$

είναι άνω τριγωνικό.

Ο τελεστής ΟΠ επί του χώρου που παράγεται από τις στήλες του A είναι ο

$$P = Q_1 \underbrace{\left(Q_1^\top Q_1\right)^{-1}}_{I} Q_1^\top$$

Η λύση του προβλήματος είναι $x = Pb = Q_1 Q_1^\top b$.

Αν χρησιμοποιήσουμε ευκλ. νόρμα

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

Έστω ότι $Q^\top b = \begin{pmatrix} c \\ d \end{pmatrix}$ και εφόσον $Rx = \begin{pmatrix} R_1 x \\ 0 \end{pmatrix}$ τότε

$$
\begin{aligned}
\|Ax - b\|_2^2 &= \|Q^\top Ax - Q^\top b\|_2^2 \\
&= \| \begin{pmatrix} R_1 x - c \\ d \end{pmatrix} \|_2 \\
&= \|R_1 x - c\|_2^2 + \|d\|_2^2
\end{aligned}
$$

Για να λύσουμε το ΑΓΑ.2 αρκεί να διαλέξουμε το x ώστε να μηδενίσει τον πρώτο όρο του δεξιού σκέλους. Σημειώστε πως δεν μπορούμε να ελέγξουμε τον άλλο όρο, ο οποιος δίνει το μέτρο του υπολοίπου (δηλ. σφάλμα) της προσέγγισης.

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. *)

$$u(j) = 1; u(j+1:m) = A(j+1:m, j)$$

$$b(j:m) = \mathrm{REFL.ROW}(b(j:m), u(j:m))$$

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.

Κώδικας 1: help polyfit

```
POLYFIT Fit polynomial to data.
P = POLYFIT(X,Y,N) finds coeffs of polynomial P(X) of
deg N that fits  data Y best in least-squares sense.P is
row vector of length N+1 with polynomial coeffs in
descending powers, P(1)*X^N +...+ P(N)*X + P(N+1)

[P,S] = POLYFIT(X,Y,N) returns pol coeffs P and
structure S for use with POLYVAL to obtain error
estimates for predictions.  S contains the R from
QR of Vandermonde matrix of X, the degrees of freedom
and norm of residuals. If data Y are random,
an estimate of the covariance matrix of P is
(Rinv*Rinv')*normr^2/df, where Rinv is inverse of R.

[P,S,MU] = POLYFIT(X,Y,N) finds coefs of polynomial in
XHAT = (X-MU(1))/MU(2) where MU(1) = MEAN(X) and
MU(2) = STD(X). Centering and scaling improves
numerical properties of polynomial and fitting
of both the polynomial and the fitting algorithm.
```

Τί γίνεται στη MATLAB

`help` για το $\backslash$ στη MATLAB R2011b

Backslash or matrix left division. If `A` is a square matrix, `A\B` is roughly the same as `inv(A)*B`, except it is computed in a different way. If `A` is an `n`-by-`n` matrix and `B` is a column vector with `n` components, or a matrix with several such columns, then `X = A\B` is the solution to the equation $AX = B$. A warning message is displayed if `A` is badly scaled or nearly singular. See the reference page for `mldivide` for more information.

If `A` is an `m`-by-`n` matrix with `m ~= n` and `B` is a column vector with `m` components, or a matrix with several such columns, then `X = A\B` is the solution in the least squares sense to the under- or overdetermined system of equations $AX = B$. The effective rank, `k`, of `A` is determined from the QR decomposition with pivoting. A solution `X` is computed that has at most `k` nonzero components per column. If `k < n`, this is usually not the same solution as `pinv(A)*B`, which is the least squares solution with the smallest norm $\|X\|$.

ΠΡΟΣΟΧΗ η QR υλοποιείται και με οδήγηση στηλών για να αντιμετωπίσει πιο αποτελεσματικά τις περιπτώσεις μητρώων μειωμένης τάξης (δηλ. γραμμικά εξαρτημένες στήλες).

Επαναληπτική εκλέπτυνση (*iterative refinement*)

<u>Ιδέα</u> Επαναληπτική εύρεση διόρθωση της λύσης με Newton επί $F(x) = Ax - b = 0$.
<u>Προσοχή:</u> Λόγω γραμμικότητας, $(F'(x))_{i,j} = \frac{\partial F_i}{\partial \xi_j}(x) = A$ και οι υψηλότερη παράγωγοι είναι 0.
Αν το Ιακωβιανό μητρώο $J(x) = [\frac{\partial F_i}{\partial \xi_j}(x)]_{i,j}$ και θέσουμε e για τη διόρθωση,

$$F(x^{(k-1)} + e) = 0 \Rightarrow F(x^{(k-1)}) + F(x^{(k-1)})e^{(k)} = 0 \Rightarrow \nabla F(x^{(k-1)})e^{(k)} = -F(x^{(k-1)})$$

οπότε θέτοντας $r^{(k-1)} = b - Ax^{(k-1)}$, για $k = 1, \ldots$

$$\begin{aligned} Ae^{(k)} &= r^{(k-1)} && \text{η ακριβής λύση θα έδινε την ορθή διόρθωση} \\ x^{(k)} &= x^{(k-1)} + e^{(k)} && \text{λόγω μειωμένης ακρίβειας (π.χ. α.κ.υ.) προσεγγιστική διόρθωση} \end{aligned}$$

(s) υπολογισμός L, U και επίλυση $x^{(0)} = U^{-1}(L^{-1}b)$
$k = 1$
repeat
  $r^{(k)} = b - Ax^{(k-1)}$
(s) $z^{(k)} = U^{-1}(L^{-1}r^{(k)})$ % (αξιοποιούμε την παραγοντοποίηση)
  $x^{(k)} = x^{(k-1)} + z^{(k)}$
  $k = k + 1$
until convergence

Παρατηρήσεις

σχετικά με την επαναληπτική εκλέπτυνση

- Αν υπολογίζαμε $A^{-1}r^{(k)}$ ακριβώς, τότε $x^{(j)}$ θα ήταν η λύση!
- Για ταχύτητα, πρέπει να κάνουμε χρήση της παραγοντοποίησης *LU* για τον *A*.
- Συνήθως: Επιζητούμε λύση $x^{(j)}$ που έχει μικρότερο εμπρός σφάλμα από τη $x^{(0)}$.
- Υπό ορισμένες συνθήκες, μερικά βήματα μπορεί να επιφέρουν μεγάλη βελτίωση στην ακρίβεια.
- Ενδιαφέρουσα ιδέα σε ορισμένα συστήματα (π.χ. GPU): Υλοποίηση βημάτων υψηλού κόστους (1, 5) σε μονή ακρίβεια.

TECHNICAL SPECIFICATIONS	TESLA K10[a]	TESLA K20	TESLA K20X
Peak double precision floating point performance (board)	0.19 teraflops	1.17 teraflops	1.31 teraflops
Peak single precision floating point performance (board)	4.58 teraflops	3.52 teraflops	3.95 teraflops

Παράδειγμα

Χρησιμοποιούμε $A = \mathrm{single}(\texttt{gfpp}(20))$, γνωστό x και $b = Ax$. Θέτουμε το υπολογισμένο $\tilde{x} = A^{-1}b$, και

$$\tilde{x}^{(d)} := \tilde{x} + A^{-1}(b - A\tilde{x})$$

Αποτελέσματα χωρίς και με εκλέπτυνση χρησιμοποιώντας διπλή ακρίβεια επιλεκτικά:

0 βήμ. εκλ. $\|x - \tilde{x}\| = 3.3462e-03$

1 βημ. εκλ. $\|x - \tilde{x}^{(1)}\| = 3.4757e-07$

2 βημ. εκλ. $\|x - \tilde{x}^{(2)}\| = 9.2238e-11$

3 βημ. εκλ. $\|x - \tilde{x}^{(3)}\| = 1.3467e-15$

ΠΑΡΑΤΗΡΗΣΗ Η εκλέπτυνση μείωσε το εμπρός σφάλμα.

Σύσταση για επίλυση ΑΓΑ.2 στη MATLAB

Μέσω Cholesky των κανονικών εξισώσεων

Κώδικας 2: help qr

```
% Example: The least squares approximate solution to A*x ...
    = b
% can be found with the Q-less qr decomposition and
% one step of iterative refinement:

 if issparse(A), R = qr(A); else R = triu(qr(A)); end
 x = R\(R'\(A'*b));
 r = b - A*x;
 e = R\(R'\(A'*r));
 x = x + e;
```

<u>ΠΡΟΣΟΧΗ:</u> 1) Το R υπολογίστηκε κατευθείαν από την QR και όχι από το $A^\top A$, 2) χωρίς χρήση Q, 3) με (ένα βήμα) επαναληπτικής εκλέπτυνσης για να μειώσουμε το σφάλμα που προκύπτει από τη χρήση του $R^\top R$.

Βιβλιογραφία I

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

Χρήση Έργου Τρίτων Ι

1. http://www.mathworks.com/ (βλ. σελ 8,9,13)
2. http://www.ozone3d.net/public/jegx/201311/nvidia-tesla-gpu-specifications.png (βλ. σελ 11)

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών - Ευστράτιος Γαλλόπουλος 2015

``Επιστημονικός Υπολογισμός Ι'', Έκδοση: 1.0, Πάτρα 2013-2014.

Διαθέσιμο από τη δικτυακή διεύθυνση: https://eclass.upatras.gr/courses/CEID1096/

Τέλος Ενότητας

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