Επιστημονικός Υπολογισμός Ι

Ενότητα **5** : Επίλυση Γραμμικών Συστημάτων

Ευστράτιος Γαλλόπουλος
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Άδειες 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.

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
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Σκοπός Ενότητας

- Επίλυση γραμμικών συστημάτων και εκμετάλλευση ιδιοτήτων του μητρώου
- Παραγοντοποίηση LU με μερική οδήγηση
- Επαναληπτική εκλέπτυνση
- Συμμετρικά θετικά ορισμένα μητρώα
- Παραγοντοποίηση Cholesky και μέθοδος Συζυγών Κλίσεων (CG)

Περιεχόμενα

Υπενθύμιση προηγούμενης διάλεξης

- Γενικό πρόβλημα
- Είδη/τύποι μητρώων
- Πολυπλοκότητα αντιστροφής
- Παραγοντοποίηση *LU* κατά πλοκάδες και συμπλήρωμα Schur
- Μεθοδολογία παραγοντοποίησης και πλεονεκτήματα
- Κλασικές παραγοντοποιήσεις
- Πολυπλοκότητα παραγοντοποίησης *LU* για διαφορετικά είδη μητρώων
- ... να αντιλαμβανόμαστε & να αξιοποιούμε τη δομή (όταν υπάρχει)

«Νέα» θέματα !!

Υπολογιστικά

1. Συμπεριφορά «κλασικών» αλγορίθμων στο μοντέλο ιεραρχικής μνήμης;
2. Τροποποιήσεις για καλύτερη επί/απόδοση υπό το νέο μοντέλο.
3. Υλοποιήσεις

Αριθμητικά

1. Συμπεριφορά «κλασικών» αλγορίθμων στο μοντέλο αριθμητικής.
2. Συμπεριφορά «νέων» αλγορίθμων στο μοντέλο αριθμητικής.

Επίλυση τριγωνικών συστημάτων

Έστω A κάτω τριγωνικό: Γράφουμε:

$$\lambda_{i,:}^T x = \beta_i, \quad 1 \le i \le n,$$

Υλοποίηση με BLAS-1 _DOT:

$$\xi_j \;=\; (\beta_j - \overbrace{\lambda_{j,1:j-1}^\top \xi_{1:j-1}}^{\text{DOT}})/\lambda_{jj},\; j = 1, \ldots, n.$$

Ανάλυση

- $\Omega = \sum_{j=1}^{n}(1 + (2j - 2)) = n^2$
- $\Phi_{\min} = \sum_{j=1}^{n} j + 2n$
- $\mu_{\min} = \frac{n+5}{2n} = \frac{1}{2} + O(\frac{1}{n})$

Οι αλγόριθμοι επίλυσης τριγωνικών συστημάτων με 1 δεξιό μέλος θεωρούνται BLAS-2.

Αν υπάρχουν s δεξιά μέλη:

- $\Omega = \sum_{j=1}^{n}(1 + (2j - 2)) = sn^2$
- $\Phi_{\min} = \sum_{j=1}^{n} j + 2ns$
- $\mu_{\min} = \frac{n+4s+1}{2ns} = \frac{1}{2s} + O(\frac{1}{n})$

οπότε για $s \gg 1$ δεξιά μέλη η επίλυση κατατάσσεται στα *BLAS-3*.

Εναλλακτική υλοποίηση

Έστω A κάτω τριγωνικός:

$$b = \lambda_{:,1}\xi_1 + \ldots + \lambda_{:,n}\xi_n.$$

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_j \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} \lambda_{11} \\ \vdots \\ \lambda_{j1} \\ \vdots \\ \lambda_{n1} \end{pmatrix} \xi_1 + \begin{pmatrix} 0 \\ \vdots \\ \lambda_{j2} \\ \vdots \\ \lambda_{n2} \end{pmatrix} \xi_2 + \ldots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \lambda_{nn} \end{pmatrix} \xi_n$$

$$\begin{pmatrix} \lambda_{11} & 0_{1,n-1} \\ \lambda_{2:n,1} & A_{2:n,2:n} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_{2:n} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_{2:n} \end{pmatrix}$$

Το ξ_1 υπολογίζεται από

$$\begin{aligned} \lambda_{11}\xi_1 &= \beta_1 \\ A_{2:n,2:n}\xi_{2:n} &= \beta_{2:n} - \lambda_{2:n,1}\xi_1 \end{aligned}$$

Υλοποίηση με BLAS-1 _AXPY:

```
for j = 1 : n − 1
    ξ_j = β_j/λ_jj
    for i = j + 1 : n
        β_i = β_i − λ_ij ξ_j
    end
end
```

Η επίλυση τριγωνικών συστημάτων στο BLAS API

BLAS-2:	_ TRSV	επίλυση $Ax = b$
BLAS-3:	_ TRSM	επίλυση $AX = B$ (περισσότερα δεξιά μέλη)
	_ TRCON	εκτίμηση του δείκτη κατάστασης
	_ TRTRS	επίλυση $AX = B$ και $A^\top X = B$
	_ TRTRI	υπολογισμός A^{-1}

Παρατηρήσεις

Έστω το σύνολο των κάτω τριγωνικών μητρώων
$\mathcal{L} = \{A \in \mathbb{R}^{n\times n} | A \text{ κάτω τριγωνικό } \}$. Τότε

- Αν $A, B \in \mathcal{L}$ τότε $\rho A + \sigma B, A \cdot B, A^{-1} \in \mathcal{L}$.
- Αντίστοιχα και για τα άνω τριγωνικά μητρώα
- Οι πράξεις με τριγωνικά μητρώα κοστίζουν $O(n^2)$.

Παρατηρήσεις

Έστω το σύνολο των κάτω τριγωνικών μητρώων
$\mathcal{L} = \{A \in \mathbb{R}^{n \times n} | A \text{ κάτω τριγωνικό } \}$. Τότε

- Αν $A, B \in \mathcal{L}$ τότε $\rho A + \sigma B, A \cdot B, A^{-1} \in \mathcal{L}$.
- Αντίστοιχα και για τα άνω τριγωνικά μητρώα
- Οι πράξεις με τριγωνικά μητρώα κοστίζουν $O(n^2)$.
- $\Rightarrow$ Αξίζει να αξιοποιούμε την πληροφορία ότι το μητρώο είναι τριγωνικό καλώντας κατάλληλους αλγορίθμους.

Πίσω ανάλυση σφάλματος εμπρός αντικατάστασης

Έστω $n = 2$

$$\begin{pmatrix} \lambda_{11} & 0 \\ \lambda_{21} & \lambda_{22} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$$

1. $\xi_1 = \frac{\beta_1}{\lambda_{11}}$ $\quad \tilde{\xi}_1 = \frac{\beta_1}{\lambda_{11}(1+\delta_1)}$
2. $\xi_2 = \frac{\beta_2 - \lambda_{21}\xi_1}{\lambda_{22}}$ $\quad \tilde{\xi}_2 = \mathrm{fl}(\mathrm{fl}(\beta_2 - \mathrm{fl}(\lambda_{21}\tilde{\xi}_1))/\lambda_{22})$

Χρησιμοποιώντας τις μεταβλητές t_i για τα ενδιάμεσα αποτελέσματα,

$$\begin{aligned} t_1 = \mathrm{fl}(\lambda_{21}\tilde{\xi}_1) &= \lambda_{21}\tilde{\xi}_1(1+\delta_2) \\ t_2 = \mathrm{fl}(\beta_2 - t_1) &= (\beta_2 - t_1)/(1+\delta_3) \\ \tilde{\xi}_2 = \mathrm{fl}(t_2/\lambda_{22}) &= \frac{t_2}{\lambda_{22}(1+\delta_4)} \\ \Rightarrow \tilde{\xi}_2 &= \frac{\beta_2}{\lambda_{22}(1+\delta_4)(1+\delta_3)} - \frac{\frac{\lambda_{21}\beta_1}{\lambda_{11}(1+\delta_1)}(1+\delta_2)}{\lambda_{22}(1+\delta_4)(1+\delta_3)} \end{aligned}$$

Επομένως

$$\begin{aligned}\tilde{\xi}_1 &= \frac{\beta_1}{\lambda_{11}(1+\delta_1)} \\ \tilde{\xi}_2 &= \frac{\beta_2}{\lambda_{22}(1+\theta_2)} - \frac{\frac{\lambda_{21}\beta_1}{\lambda_{11}(1+\delta_1)}(1+\delta_2)}{\lambda_{22}(1+\theta_2)}\end{aligned}$$

επομένως

$$\begin{pmatrix}\lambda_{11}(1+\delta_1) & 0 \\ \lambda_{21}(1+\delta_2) & \lambda_{22}(1+\theta_2)\end{pmatrix}\begin{pmatrix}\tilde{\xi}_1 \\ \tilde{\xi}_2\end{pmatrix} = \begin{pmatrix}\beta_1 \\ \beta_2\end{pmatrix}$$

άρα

$$(L+\Delta L)\tilde{x} = b, \text{ όπου } |\Delta L| \leq \gamma_n |L|$$

Πίσω ανάλυση σφάλματος επίλυσης τριγωνικού συστήματος

Αποδεικνύεται:

το υπολογισθέν $x = \hat{x}$ ικανοποιεί τη σχέση

$$b = (L + \Delta L)\hat{x}$$

όπου τα στοιχεία του ΔL είναι φραγμένα

$$|\Delta L| \leq \gamma_n |L|$$

ΣΥΜΠΕΡΑΣΜΑ Η επίλυση τριγωνικών συστημάτων με τον αλγόριθμο βασισμένο σε DOT είναι προς τα πίσω ευσταθής. Μπορούμε να δείξουμε το ίδιο και για τον αλγόριθμο που βασίζεται σε _AXPY.

Νόρμες μητρώων

Μετρική μεγέθους των στοιχείων ενός μητρώου Επεκτείνει την έννοια που είχαμε για τα διανύσματα.

Υπενθυμίζουμε:

Νόρμα διανύσματος

Νόρμα διανύσματος $x \in \mathbb{C}^n$, $||x||$, είναι οποιαδήποτε συνάρτηση με πεδίο ορισμού το $\mathbb{C}^n$ και πεδίο τιμών το $\mathbb{R}$ που ικανοποιεί:

1. $||x|| > 0$ εκτός αν $x = 0$ οπότε $||x|| = 0$ (θετικά ορισμένη)·
2. $||x + y|| \leq ||x|| + ||y||$ (τριγωνική ανισότητα - υποαθροιστική)·
3. $||\alpha x|| = |\alpha| ||x||$ για $\alpha \in \mathbb{R}$ (απόλυτα ομογενής).

Νόρμα μητρώου

Βασική ιδέα Οποιαδήποτε συνάρτηση $\|\cdot\| : \mathbb{C}^{n\times n} \to \mathbb{R}$ που ικανοποιεί τις 3 συνθήκες ορισμού της νόρμας διανύσματος ονομάζεται νόρμα μητρώου.
Υποπολλαπλασιαστική νόρμα Αν επιπλέον ισχύει ότι

$$C = AB \Rightarrow \|C\| \leq \|A\|\|B\|$$

η νόρμα λέγεται υποπολλαπλασιαστική.

Ορισμός

Για οποιαδήποτε νόρμα διανύσματος $\|\cdot\|_p$, ορίζεται η επαγόμενη νόρμα μητρώου

$$\|A\|_p = \max_{x\neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

Προσοχή: Οι επαγόμενες νόρμες είναι υποπολλαπλασιαστικές.

Συνηθισμένες νόρμες μητρώων

Έστω $A \in \mathbb{C}^{m \times n}$.

- $\|A\|_1 = \max_j \sum_{i=1}^{m} |\alpha_{i,j}|$.
- $\|A\|_2 = \sqrt{\lambda_{\max}(A^* A)}$ όπου $\lambda_{\max}(\cdot)$ συμβολίζει τη μέγιστη ιδιοτιμή (φασματική νόρμα).
- $\|A\|_\infty = \max_i \sum_{j=1}^{n} |\alpha_{i,j}|$ (νόρμα μεγίστου)

Νόρμα Frobenius

$$\|A\|_{\mathrm{F}} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |\alpha_{i,j}|^2 \right)^{1/2}.$$

Σημαντικές ιδιότητες και παρατηρήσεις

1. $\|A\|_p = \||A|\|_p$ για $p = 1, \infty, \mathrm{F}$ αλλά όχι για την νόρμα-2.
2. Αν $X^* X = I_m$, $Y^* Y = I_n$ τότε $\|XAY\| = \|A\|$ για τις νόρμες-2 και -F.
3. Η νόρμα $\|A\| = \max_{i,j} |\alpha_{i,j}|$ δεν έχει την υποπολλαπλασιαστική ιδιότητα.

Ισοδυναμίες

Ισοδύναμες νόρμες Αν $\|\cdot\|_p$ και $\|\cdot\|_q$ είναι δύο νόρμες στον ίδιο δ.χ. V τότε λέγονται ισοδύναμες αν υπάρχουν θετικές σταθερές γ_1, γ_2 τ.ώ. για κάθε $x \in V$,

$$\gamma_1 \|x\|_p \leq \|x\|_q \leq \gamma_2 \|x\|_p$$

Σε δ.χ. πεπερασμένων διαστάσεων, όλες οι νόρμες είναι ισοδύναμες.

για τις νόρμες διανύσματος

	1	2	∞
1	1	$\sqrt{n}$	n
2	1	1	$\sqrt{n}$
∞	1	1	1

για τις νόρμες μητρώου (παραδείγματα)

- $\|A\|_2 \leq \|A\|_F \leq \sqrt{n}\|A\|_2$
- $\frac{1}{\sqrt{n}}\|A\|_\infty \leq \|A\|_2 \leq \sqrt{m}\|A\|_\infty$
- $\frac{1}{\sqrt{m}}\|A\|_1 \leq \|A\|_2 \leq \sqrt{n}\|A\|_1$.

Στον πίνακα για τις νόρμες διανυσμάτων, στη γραμμή που αντιστοιχεί στη νόρμα 2, οι αριθμοί $1, 1, \sqrt{n}$ σημαίνουν

$$\|x\|_2 \leq \|x\|_1, \|x\|_2 \leq \|x\|_2, \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

Αντίστοιχος πίνακας μπορεί να συνταχθεί για νόρμες μητρώων.

Κατάσταση προβλήματος

ΠΡΟΒΛΗΜΑ Να λυθεί το κάτω τριγωνικό σύστημα $Lx = b$ ως προς x
Πώς επιδρούν μικρές διαταραχές;

$$\begin{aligned} (L+\Delta L)(x+\Delta x) &= b+\Delta b \\ Lx &= b \end{aligned}$$

Κατάσταση προβλήματος

ΠΡΟΒΛΗΜΑ Να λυθεί το κάτω τριγωνικό σύστημα $Lx = b$ ως προς x
Πώς επιδρούν μικρές διαταραχές;

$$
\begin{aligned}
(L+\Delta L)(x+\Delta x) &= b+\Delta b\\
Lx &= b\\
L\Delta x + \Delta L x + \Delta L \Delta x &= \Delta b\\
\Delta x &= L^{-1}\Delta b - L^{-1}\Delta L x - L^{-1}\Delta L \Delta x\\
\|\Delta x\| &\leq \|L^{-1}\| \left(\|\Delta L\| \|x\| + \|\Delta L\| \|\Delta x\| + \|\Delta b\|\right)\\
\frac{\|\Delta x\|}{\|x\|} &\leq \kappa(L)\left(\frac{\|\Delta L\|}{\|L\|} + \frac{\|\Delta L\|}{\|L\|}\frac{\|\Delta x\|}{\|x\|} + \frac{\|\Delta b\|}{\|L\|\|x\|}\right)
\end{aligned}
$$

επομένως

$$
\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(L)}{1-\kappa(L)\frac{\|\Delta L\|}{\|L\|}}\left(\frac{\|\Delta L\|}{\|L\|} + \frac{\|\Delta b\|}{\|b\|}\right).
$$

Παρατηρήσεις

Εμπρός σφάλμα από την πίσω ανάλυση

Έστω ότι επιλύουμε $Lx = b$ και ότι η υπολογισμένη λύση με εμπρός αντικατάσταση είναι $\hat{x}$. Χρησιμοποιούμε μία από τις νόρμες-1, $-\infty$, $-\mathrm{F}$:

- από την πίσω ανάλυση, ισχύει $(L + \Delta L)\hat{x} = b$, όπου

$$|\Delta L| \leq \gamma_n |L|$$

- .. επίσης, (για τις νόρμες-1, $-\infty$, $-\mathrm{F}$), $\|\Delta L\| \leq \gamma_n \|L\|$
- Το εμπρός σφάλμα φράσσεται ως εξής,

$$\frac{\|\hat{x} - x\|}{\|x\|} \leq \frac{\kappa(L)}{1 - \kappa(L)\frac{\|\Delta L\|}{\|L\|}} \frac{\|\Delta L\|}{\|L\|}$$

και $\kappa(L)\frac{\|\Delta L\|}{\|L\|} \ll 1$, τότε το επόμενο φράγμα είναι πιο εύληπτο (και αγνοεί όρους τετραγωνικής τάξης)

$$\frac{\|\hat{x} - x\|}{\|x\|} \leq \kappa(L)\gamma_n.$$

Βιβλιογραφία Ι

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

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