ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

ΔΙΑΛΕΞΗ 14, 11/12/09

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

Σχετικά με την παραγοντοποίηση LU και τη λύση μέσω αυτής

- Στα παραπάνω δεν λάβαμε καθόλου υπόψη τον αλγόριθμο επίλυσης.
- Πρώτα τρέχαμε τον υπό μελέτη αλγόριθμο και μετά, μέσω του β_N, εξετάζαμε αν ήταν πίσω ευσταθής.
- Ιδανικά, θέλουμε να είμαστε βέβαιοι για την ευστάθειά ενός αλγορίθμου πριν τον τρέξουμε!
- Μπορούμε να το κάνουμε αυτό για την LU;

Σημαντικό:

- Για να μιλήσουμε για την LU πρέπει να προσδιορίσουμε ποιόν αλγόριθμο εννοούμε,
- ... τα αποτελέσματα εξαρτώνται από τη μέθοδο οδήγησης!

Αξίζει να σημειωθεί:

1. Γενικά πρέπει να αναφερόμαστε σε παραγοντοποίηση που υπάρχει υπό τον όρο ότι το μητρώο είναι αντιστρέψιμο
2. ... άρα είναι απαραίτητο να υπάρχει κάποια μορφή οδήγησης.
3. Στην πράξη, ο αλγόριθμος που χρησιμοποιείται περισσότερο είναι LU με μερική οδήγηση.
4. ... Ο αλγόριθμος είναι σχεδόν πάντα πίσω ευσταθής,
5. εκτός από ορισμένους «παθολογικούς» συνδυασμούς $[A, b]$.
6. ... όπως η `gfpp`.

Θεώρημα παραγοντοποίησης Αν A αντιστρέψιμο τότε υπάρχουν μητρώα μετάθεσης P_1, P_2, κάτω τριγωνικό L με τη μονάδα στη διαγώνιο και αντιστρέψιμο U ώστε $P_1AP_2 = LU$. Η παραγοντοποίηση είναι εφικτή και μόνο με ένα από τα P_1, P_2 (τα L, U εξαρτώνται από την επιλογή που κάνουμε για τα P_j).

GEPP: (Gaussian elimination with partial pivoting) Επιλέγουμε $P_2 = I$ και τον k παράγοντα του P_1 ώστε στο βήμα k να φέρνει το μέγιστο σε απόλυτη τιμή στοιχείο στις θέσεις $(k : n, k)$.

GECP: (Gaussian elimination with complete pivoting) Επιλέγουμε τους παράγοντες των P_2 και P_1 ώστε στο βήμα k να φέρνουν το μέγιστο σε απόλυτη τιμή στοιχείο στις θέσεις $(k : n, k : n)$.

Rook pivoting: Πρόσφατη μέθοδος οδήγησης που επιτυγχάνει περιορισμένο σφάλμα με μικρότερο κόστος από την GECP.

<u>Πώς συμπεριφέρεται η επίλυση με διάσπαση LU; (συνοπτικά)</u> Έστω ότι έγιναν οι μεταθέσεις και $A = LU$. Μπορεί να δειχτεί (βιβλίο) ότι τα υπολογισμένα $\hat{L}, \hat{U}$ είναι τ.ώ.

$$A + E = \hat{L}\hat{U}$$

όπου

$$|E| \leq \gamma_n |\hat{L}||\hat{U}|, \quad \||E|\| \leq \gamma_n \||\hat{L}|\|\||\hat{U}|\|$$

Λύνοντας $\hat{L}y = b$ έχουμε $(\hat{L} + \Delta L)\hat{y} = b$ όπου $|\Delta L| \leq \gamma_n |\hat{L}|$ και το ίδιο για το $\hat{U}$. Επίσης

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

Άρα

$$\begin{aligned} b &= (\hat{L}\hat{U} + \hat{L}\Delta U + \Delta L\hat{U} + \Delta L\Delta U)\hat{x} \\ &= (A + \underbrace{E + \hat{L}\Delta U + \Delta L\hat{U} + \Delta L\Delta U}_{\Delta A})\hat{x} \end{aligned}$$

Επομένως το υπολογισμένο x ικανοποιεί το $b = (A + \Delta A)\hat{x}$ όπου

$$\Delta A = E + \hat{L}\Delta U + \Delta L\hat{U} + \Delta L\Delta U.$$

Η πίσω ευστάθεια εξαρτάται από το μέγεθος του $\frac{\|\Delta A\|}{\|A\|}$

Για να φράξουμε το ΔA

$$\begin{aligned} |\Delta A| &\leq |E| + |\hat{L}||\Delta U| + |\Delta L||\hat{U}| + |\Delta L||\Delta U| \\ &\leq 3\gamma_n |\hat{L}||\hat{U}| + O(u^2) \\ \|\Delta A\| &\leq 3\gamma_n \||\hat{L}|\|\||\hat{U}|\| \end{aligned}$$

- Έχουμε πίσω ευστάθεια όταν $\frac{\|\Delta A\|}{\|A\|} \approx \mathrm{O}(u)$.
- Αν

  $$3\gamma_n \||\hat{L}|\|\||\hat{U}|\| \approx O(u)\|A\|$$

  τότε

  $$\frac{\|\Delta A\|}{\|A\|} \approx O(u)$$

Η πίσω ευστάθεια καθορίζεται από το μέγεθος του $|||\hat{L}|||\,|||\hat{U}|||$

Παρατηρήσεις σχετικά με την $|||\hat{L}|||\,|||\hat{U}|||$ Χωρίς οδήγηση: Ο παράγοντας μπορεί να είναι πολύ μεγάλος γιατί δεν έχουμε τρόπο να φράξουμε τους όρους του $\hat{L}$ (θυμηθείτε: χωρίς οδήγηση μπορεί να προκύψουν πολύ μικρά στοιχεία στη διαγώνιο η διαίρεση με τα οποία θα έχει σαν αποτέλεσμα πολύ μεγάλα στοιχεία στο $\hat{L}$ κ.λπ. Τότε, εκτός από ειδικές περιπτώσεις δεν θα έχουμε πίσω ευστάθεια.

Παρατήρηση: Αν χρησιμοποιήσουμε μερική ή πλήρη οδήγηση, ο παράγοντας $|||\hat{L}|||$ δεν μπορεί να γίνει πολύ μεγάλος γιατί το μέγιστο στοιχείο κάθε στήλης είναι 1.
Με οδήγηση που εξασφαλίζει ότι $|\hat{\lambda}_{ij}| \leq 1$: Αν χρησιμοποιήσουμε μερική οδήγηση ή πλήρη τότε $|\hat{\lambda}_{ij}| \leq 1$ και αν χρησιμοποιήσουμε τις νόρμες $1, \infty$ τότε $|||\hat{L}||| \approx O(n)$.
Επομένως αρκεί να εξετάσουμε το μέγεθος των στοιχείων του $\hat{U}$.
Μπορούμε να χρησιμοποιήσουμε τον παράγοντα

$$\frac{\max_{i,j} |\hat{\eta}_{ij}|}{\max_{i,j} |\alpha_{ij}|}$$

ως ένδειξη του μεγέθους του

Παρατήρηση: Αφού σε κάθε βήμα υπολογίζουμε το

$$\alpha_{ij}^{(k)} = \alpha_{ij}^{(k-1)} - \frac{\alpha_{ik}^{(k-1)}}{\alpha_{kk}^{(k-1)}} \alpha_{kj}^{(k-1)}$$

από τη μερική οδήγηση

$$|\frac{\alpha_{ik}^{(k-1)}}{\alpha_{kk}^{(k-1)}}| \leq 1$$

επομένως

$$\begin{aligned} |\alpha_{ij}^{(k)}| &\leq |\alpha_{ij}^{(k-1)}| + |\frac{\alpha_{ik}^{(k-1)}}{\alpha_{kk}^{(k-1)}} \alpha_{kj}^{(k-1)}| \\ &\leq |\alpha_{ij}^{(k-1)}| + |\alpha_{kj}^{(k-1)}| \end{aligned}$$

Επαγωγικά[1] δείχνεται ότι

$$\max_{i,j} |\eta_{ij}| \leq 2^{n-1} \max_{i,j} |\alpha_{ij}|$$

- Με μερική οδήγηση ο παράγοντας είναι «φραγμένος» αλλά το φράγμα μπορεί να είναι πολύ μεγάλο.
- Δυστυχώς υπάρχουν μητρώα για τα οποία το άνω φράγμα είναι εφικτό.

[1]Κάνουμε την απλοποίηση να δείξουμε το φράγμα για τον ακριβή U αντι για τον $\hat{U}$.

- Αν χρησιμοποιήσουμε GEPP *συνήθως* ισχύει ότι και τα στοιχεία του U παραμένουν μικρά.

Συντελεστής αύξησης στην οδήγηση Ορίζεται ως $\rho := \frac{\max\{\alpha_0,\alpha_1,\ldots,\alpha_{n-1}\}}{\alpha_0}$ όπου $\alpha_k := \max_{i,j}|\alpha_{ij}^{(k)}|$ είναι τα στοιχεία με μέγιστη απόλυτη τιμή στις θέσεις $(k+1:n, k+1:n)$ του $L_k P_k \cdots L_1 P_1 A$.
Παράδειγμα 1.

$$A = \begin{bmatrix} \frac{1}{10000} & 1 \\ 1 & 1 \end{bmatrix}$$

χωρίς οδήγηση:

$$L_1 = \begin{bmatrix} 1 & 0 \\ -10000 & 1 \end{bmatrix}$$

$$L_1 A = \begin{bmatrix} \frac{1}{10000} & 1 \\ 0 & -9999 \end{bmatrix}.$$

Δηλαδή $\alpha_0 = 1$, $\alpha_1 = 9999$ *και* $\rho = 9999$.
Με μερική οδήγηση

$$P_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad L_1 = \begin{bmatrix} 1 & 0 \\ -\frac{1}{10000} & 1 \end{bmatrix} \quad L_1 P_1 A = \begin{bmatrix} 1 & 1 \\ 0 & \frac{9999}{10000} \end{bmatrix}$$

οπότε $\alpha_0 = 1$, $\alpha_1 = 0.9999$ *και* $\rho = 1$. □

Προβληματική LU λόγω μεγάλου ρ_n

$$A = \begin{pmatrix} 1 & 0 & 0 & 1 \\ -1 & 1 & 0 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & 1 \end{pmatrix}$$

Με μερική οδήγηση $\rho^{\text{MO}} = 8$:

$$U = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 4 \\ 0 & 0 & 0 & 8 \end{pmatrix}$$

Με πλήρη οδήγηση $\rho^{\Pi\text{O}} = 2$.

$$U = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & -2 & 1 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

Όταν το μητρώο είναι έχει την παραπάνω μορφή $\rho_n^{\text{MO}} = 2^{n-1}$.
Στις περιπτώσεις αυτές υπάρχει πρόβλημα (για μεγάλο n). Δοκιμάσαμε αυτό το μητρώο (συνάρτηση `gfpp` από το toolbox του N. Higham) και πήραμε τα εξής:

```
)) n=64;a=gfpp(n);x=ones(n,1);b=a*x;
)) norm(x-a\b) ans = 3.1623
)) norm(b-a*(a\b)) ans = 177.4458
)) [l,u,rpp]=gePP(a);
)) cond(a) =
1.0775e+005
)) rpp = 1.8558e+016
)) condest(a) = 1200
[l,u,p,q,rcp]=geCP(a); rcp=1
```

Επίδραση του συντελεστή Ο συντελεστής εμφανίζεται στην ανάλυση μέσω του μεγέθους του $|||\hat{U}|||$ στην ανάλυση της πίσω ευστάθειας της LU που ήδη παρουσιάσαμε.
Αξίζει να σημειωθεί ότι από νωρίς, ο Wilkinson είχε διατυπώσει το παρακάτω συναφές αποτέλεσμα.
Θεώρημα 1 (Wilkinson). *Έστω $A \in \mathbb{R}^{n\times n}$ και ότι υπολογίζουμε τη λύση $\hat{x}$ χρησιμοποιώντας GEPP. Τότε*

$$(A+\Delta A)\hat{x} = b, \quad \|\Delta A\|_\infty \leq 3n^2\gamma_n\rho_n^{MO}\|A\|_\infty$$

□

Το άνω φράγμα για το συντελεστή αύξησης της πλήρους οδήγησης είναι πολύ μικρότερο:

$$\begin{aligned}\rho_n^{\Pi O} &< \sqrt{n2^1 3^{1/2} 4^{1/3}\cdots n^{1/(n-1)}} \\ &< \sqrt{n n^{1+\frac{1}{2}+\cdots+\frac{1}{n-1}}} = \mathrm{O}(n^{\log n}))\end{aligned}$$

Για μητρώα με ειδική δομή[2] οι συντελεστής αύξησης $\rho_n^{\text{MO}}, \rho_n^{\Pi O}$, μπορεί να φραχτούν πολύ καλύτερα.

Υπενθύμιση και ανάλυση: Πώς προκύπτουν οι παράγοντες P, L, U στην $PA = LU$

- Γνωρίζουμε πώς αν το A υπακούει σε ορισμένες «σκληρές» συνθήκες, τότε μπορούμε με τοιχειώδεις μετασχηματισμούς Gauss να το μετατρέψουμε σε άνω τριγωνικό:

  $$L_{n-1}\cdots L_1 A = U \Rightarrow A = L_1^{-1}\cdots L_{n-1}^{-1}U$$

  όπου

  $$L_j = I - u_j e_j^\top \Rightarrow L_j^{-1} = I + u_j e_j^\top$$

- Μπορούμε να μεθοδεύσουμε και την $PA = LU$ ανάλογαα αλλά θέλει λίγη προσοχή (έχει ενδιαφέρον). Θα το δούμε για ένα παράδειγμα $A \in \mathbb{R}^{4\times 4}$

[2] Θα δούμε αργότερα.

- Θυμηθείτε: Μητρώο ανταλλαγής μητρώο που έχει προέλθει από το ταυτοτικό με μια ανταλλαγή δύο στηλών (ή γραμμών) του.
- Αν P είναι μητρώο ανταλλαγής τότε $P = P^\top = P^{-1}$
- Τα γινόμενα μητρώων ανταλλαγής είναι μητρώα μετάθεσης (όπως το P στην $PA = LU$).

Χρήση των μητρώων μετάθεσης Μπορούμε να αναπαραστήσουμε τις ανταλλαγές γραμμών με σκοπό την προσκόμιση μεγαλου στοιχείου στη θέση kk του μητρώου $A^{(k-1)}$ με πολλαπλασιασμό με μητρώο μετάθεσης, έστω P_k.
Μερική οδήγηση Στο βήμα k διαλέγουμε για οδηγό το μέγιστο, σε απόλυτη τιμή, στοιχείο από τις θέσεις $(k : n, k)$ του $A^{(k-1)}$. Αν το στοιχείο αυτό είναι στην γραμμή $\hat{k} \geq k$ τότε μεταθέτουμε τις γραμμές k και $\hat{k}$.
Επομένως

$$L_{n-1}P_{n-1} \cdots L_1 P_1 A = U,$$

όπου P_k είναι το μητρώο μετάθεσης που ανταλλάσει την γραμμή k με την γραμμή που περιέχει το μέγιστο στοιχείο από τις θέσεις $(k+1 : n)$ της k στήλης του μητρώου $L_{k-1}P_{k-1} \cdots L_1 P_1 A$.
Μορφή $PA = LU$ Παράδειγμα για $n = 4$: Από τις ιδιότητες των μητρώων ανταλλαγής

$$\begin{aligned} L_3 P_3 L_2 P_2 L_1 P_1 A &= U \\ L_3(P_3 L_2 P_3)(P_3 P_2 L_1 P_2 P_3)\underbrace{P_3 P_2 P_1}_{P} A &= U \\ \tilde{L}_3 \tilde{L}_2 \tilde{L}_1 P A &= U \end{aligned}$$

Επομένως

$$\begin{aligned} PA &= (\tilde{L}_3 \tilde{L}_2 \tilde{L}_1)^{-1} U \\ &= \tilde{L}_1^{-1} \tilde{L}_2^{-1} \tilde{L}_3^{-1} U \end{aligned}$$

Επίσης

$$\begin{aligned} P_3 L_2 P_3 &= P_3(I - u_2 e_2^T) P_3 \\ &= I - (P_3 u_2) e_2^T P_3 \\ &= I - (P_3 u_2)(P_3^T e_2)^T \\ &= I - (P_3 u_2)(P_3 e_2)^T \end{aligned}$$

αλλά το P_3 είναι το μητρώο ανταλλαγής που επιδρά μόνο στις γραμμές $j \geq 3$, άρα $P_3 e_2 = e_2$ άρα $P_3 L_2 P_3 = I - P_3 u_2 e_2^T$.

Ομοίως

$$\begin{aligned} P_3P_2L_1P_2P_3 &= P_3P_2(I - u_1e_1^T)P_2P_3 \\ &= I - (P_3P_2u_1)(P_3P_2e_1)^T \\ &= I - P_3P_2u_1e_1^T \end{aligned}$$

<u>Αποτέλεσμα</u> Αν χρησιμοποιήσουμε μερική οδήγηση για την παραγοντοποίηση του A τότε $PA = LU$ όπου $P = P_{n-1}\cdots P_1$ και L είναι κάτω τριγωνικό με μονάδα στη διαγώνιο και

$$L^{-1} = \tilde{L}_{n-1}\cdots\tilde{L}_1$$

όπου $\tilde{L}_{n-1} = L_{n-1} = I - u_{n-1}e_{n-1}^T$ και

$$\tilde{L}_k = P_{n-1}\cdots P_{k+1}L_kP_{k+1}\cdots P_{n-1}$$

επομένως

$$\tilde{L}_k = I - \underbrace{(P_{n-1}\cdots P_{k+1}u_k)}_{\tilde{u}_k}e_k^T$$

Το διάνυσμα $\tilde{u}_k$ είναι της μορφής

$$\tilde{u}_k = [\underbrace{0,\cdots,0,}_{k}*,*,\cdots*]$$

$$L = (\tilde{L}_{n-1}\cdots\tilde{L}_1)^{-1}$$

αρα

$$L = \tilde{L}_1^{-1}\tilde{L}_2^{-1}\cdots\tilde{L}_{n-1}$$

παρατηρούμε πως

$$\begin{aligned} \tilde{L}_k^{-1} &= (P_{n-1}\cdots P_{k+1}L_kP_{k+1}\cdots P_{n-1})^{-1} \\ &= P_{n-1}\cdots P_{k+1}L_k^{-1}P_{k+1}\cdots P_{n-1} \\ &= P_{n-1}\cdots P_{k+1}(I + u_ke_k^T)P_{k+1}\cdots P_{n-1} \\ &= I + \tilde{u}_ke_k^T \end{aligned}$$

Επίσης

$$L = (I + \tilde{u}_1e_1^T)(I + \tilde{u}_2e_2^T)\cdots(I + \tilde{u}_{n-1}e_{n-1}^T)$$

το οποίο είναι κάτω τριγωνικό! Π.χ. το $(I + \tilde{u}_1e_1^T)(I + \tilde{u}_2e_2^T)$ είναι ίσο με

$$L = I + \tilde{u}_1e_1^T + \tilde{u}_2e_2^T + \tilde{u}_1\underbrace{e_1^T\tilde{u}_2}_{0}e_2^T$$

Η στήλη k περιέχει από τη διαγώνιο και κάτω τους συντελεστές που υπολογίσαμε για την απαλοιφή των στοιχείων της στήλης.

Άλλοι τρόποι να βελτιώσουμε την αριθμητική συμπεριφορά Κλιμάκωση/ισοστάθμιση[3] (scaling/equilib
Κατασκευάζουμε διαγώνια P_1, P_2 ώστε ώστε $P_1 A P_2 P_2^{-1} x = P_1 b$ να έχει καλύτερες ιδιότητες ως προς το σφάλμα επίλυσης από το $Ax = b$· π.χ. $\kappa(P_1 A P_2) < \kappa(A)$.
Επαναληπτική εκλέπτυνση (=*iterative improvement/refinement*): Η ιδέα προέρχεται από την εφαρμογή Newton στο $f(x) \equiv Ax - b = 0$.

- Η ιδέα οφείλεται στον Wilkinson και υπεισέρχεται (και μεταμφιεσμένη) σε πάμπολλες εφαρμογές.
- Αν λύναμε $A^{-1} r_k$ ακριβώς, τότε $x^{(j)}$ θα ήταν η ακριβής λύση!
- Για ταχύτητα, πρέπει να κάνουμε χρήση της παραγοντοποίησης LU για τον A.
- Συνήθως: Επιζητούμε τον υπολογισμό του βήματος (4) με υψηλότερη ακρίβεια από το x_0.
- Μερικά βήματα μπορεί να επιφέρουν μεγάλη βελτίωση στην ακρίβεια.
- Πρόσφατα προτάθηκε για χρήση σε συστήματα όπως Cell, GPU ως εξής: Να τρέξουν ειδικά τα βήματα (1, 5) σε πολύ γρήγορη μονή ακρίβεια και τα υπόλοιπα σε διπλή.
- Κάτω από ορισμένες συνθήκες μπορεί να έχουμε εμπρός σφάλμα φραγμένο φραγμένο (περίπου) από μικρό πολλαπλάσιο του $\mathbf{u}_{32}$ είτε από $\kappa(A)\mathbf{u}_{64}$.

Σχόλια: Επαναληπτική εκλέπτυνση (παλαιά τεχνική που πάλι προσελκύει)

- Στο επίκεντρο του ενδιαφέροντος λόγω της νέας γενιάς επεξεργαστών[4] (Cell, GPUs)
- Ένας πυρήνας STI Cell BE (συνεργασία Sony, Toshiba, IBM, Sony) έχει μέγιστη επίδοση 25 Gflop/s για μονή ακρίβεια αλλά 1.8 Gflop/s για διπλή
- Στόχος να συνδυάσουμε την επίδοση των 32-bits με την ακρίβεια των 64ων στη λύση γραμμικών συστημάτων.
- Η θεωρία ανάλυσης σφάλματος δείχνει πως κάτι τέτοιο είναι εφικτό, κάτω από ορισμένες συνθήκες.

1.	(s)	$x_0 = A^{-1} b$ /* apojhke'uontai oi par'agontec L, U
2.		$k = 1$
3.		**repeat**
4.		$r_k = b - A x_{k-1}$
5.	(s)	$z_k = A^{-1} r_k$ /* qrhsimopoio'untai oi par'agontec L, U
6.		$x_k = x_{k-1} + z_k$
7.		$k = k + 1$
8.		**until** convergence

[4] www.netlib.org/lapack/lawnspdf/lawn175.pdf

ΣΥΣΤΑΣΗ: Δείτε και την εργασία (Buttari, et al.) που αναρτήθηκε μέσω ημερολογίου.
Παράδειγμα
Χρησιμοποιούμε $A = \text{single}(\texttt{gfpp}(20))$, τυχαίο 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:

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

- Κάτω από ορισμένες συνθήκες μπορεί να έχουμε εμπρός σφάλμα φραγμένο φραγμένο (περίπου) από μικρό πολλαπλάσιο του $\mathbf{u}_{32}$ είτε από $\kappa(A)\mathbf{u}_{64}$.

ΠΑΡΑΤΗΡΟΥΜΕ Η εκλέπτυνση αύξησε την ακρίβεια - στη θεωρία, η βελτίωση αναδεικνύεται μέσω της μείωσης του πίσω σφάλματος.

Προβλήματα της οδήγησης

- Η οδήγηση μπορεί να επιφέρει επιπλέον ανεπιθύμητα κόστη (συγκρίσεις, μεταφορές, ...)
- Πιο σημαντική «παράπλευρη απώλεια» ότι συνήθως καταστρέφει τη δομή του μητρώου, π.χ. χάνονται συμμετρία, τριγωνικότητα, τριδιαγωνιότητα, κ.λπ.
- Για παράδειγμα, μπορεί $A = A^\top$ αλλά αν P είναι μη τετριμμένο μητρώο μετάθεσης (δηλ. $P \neq I$) τότε $PA \neq (PA)^\top$.

Πότε μπορούμε να αποφύγουμε την οδήγηση;

Για ορισμένες κατηγορίες μητρώων μπορούμε να αποφύγουμε την οδήγηση. α) π.χ. επειδή το διαγώνιο στοιχείο συμβαίνει και είναι το μέγιστο της στήλης, ή β) γιατί χωρίς να χρησιμοποιήσουμε οδήγηση, ο συντελεστής αύξησης είναι μικρός (π.χ. 1 ή 2).

Συμμετρικά θετικά ορισμένα (ΣΘΟ): υπάρχει παραγοντοποίηση $A = LL^\top$ (Cholesky). ΠΡΟΣΟΧΗ: Το L δεν είχει κατ' ανάγκη 1 στη διαγώνιο.

Αυστηρά διαγώνια κυρίαρχα: υπάρχει παραγοντοποίηση $A = LU$ αν ισχύει αυστηρή ΔΚ ανά στήλη, δηλ. $|\alpha_{jj}| \geq \sum_{i=1,i\neq j}^{n} |\alpha_{ij}|, j = 1, ..., n$.

Συμμετρικά ΔΚ: υπάρχει παραγοντοποίηση $A = LU$ αν ισχύει αυστηρή ΔΚ.

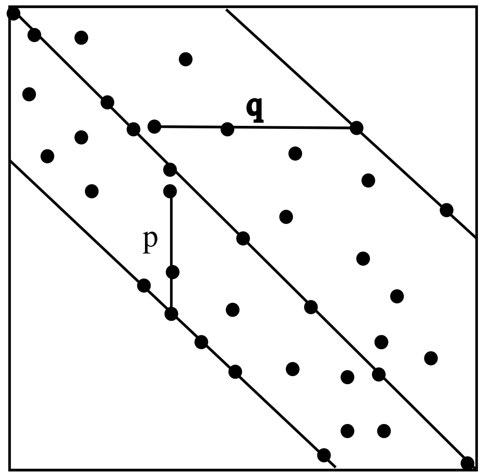
Μητρώα με ειδική δομή (Κεφ. 7.1)
Ειδική δομή: Γενικά, εννοούμε, μητρώα που μπορούμε να «περιγράψουμε με λιγότερα» από n^2 στοιχεία. Για να είμαστε πιο σαφείς χρειάζεται εξειδίκευση. Παραδείγματα: Συμμετρικά θετικά ορισμένα, Toeplitz, Hankel, Vandermonde, Cauchy, μητρώα χαμηλής τάξης, αραιά μητρώα, π.χ. μητρώα ζώνης, διαγώνια κατά ορμαθούς, τριδιαγώνια κατά ορμαθούς, ...
Παρατήρηση: Η διαχείριση μητρώων με ειδική δομή είναι σημαντική, γιατί:

1. προκύπτουν πολύ συχνά άμεσα ή έμμεσα στις εφαρμογές,
2. προσφέρονται για σχεδιασμό μεθόδων διαχείρισής τους με λιγότερο κόστος από τα γενικά μητρώα
3. πολλοί αλγόριθμοι (π.χ. εύρεσης ιδιοτιμών, ιδιαζουσών τιμών) αναγάγουν πρώτα το μητρώο σε τέτοια μορφή.
4. ... δείτε π.χ. την αναφορά lawn92.ps της LAPACK
   `My Documents\Download\L\LAPACKJava\lawn92.ps`

Παραδείγματα (εντός ύλης!)

- Αλγόριθμοι «διαχείρισης» τριγωνικών μητρώων με $\mathrm{O}(n^2)$ πράξεις α.κ.υ. (βλ. κεφ. 5)
- Αλγόριθμοι διαχείρισης Hessenberg μητρώων με $\mathrm{O}(n^2)$ πράξεις α.κ.υ.
- Αλγόριθμοι διαχείρισης τριδιαγώνιων μητρώων με $\mathrm{O}(n)$ πράξεις α.κ.υ.
- Αλγόριθμοι διαχείρισης μητρώων ζώνης (banded matrices) με $\mathrm{O}(nb)$ πράξεις α.κ.υ. όπου το b εξαρτάται από το «εύρος της ζώνης».

Μητρώα ζώνης
Ορισμός 1. *Έστω τετραγωνικός πίνακας με στοιχεία α_{ij} και έστω ότι p, q είναι ακέραιοι τέτοιοι ώστε $\alpha_{ij} = 0$ αν $j > i + q$ ή $i > j + p$.*
Το αντίστοιχο μητρώο αποκαλείται μητρώο ζώνης με εύρος άνω ζώνης q, εύρος p κάτω ζώνης και συνολικό εύρος $p + q + 1$. Θα το συμβολίζουμε συχνά ως $A(p|q)$.
Αν $p = q$ (π.χ. το μητρώο είναι συμμετρικό), τότε λέμε και ότι το μητρώο έχει ημιεύρος (= semibandwidth*) p.* $\square$



Παραδείγματα

- Αν $p = q = 1 \rightarrow$ τριδιαγώνιο: τα μη μηδενικά στοιχεία του βρίσκονται στη διαγώνιο, στην υπερδιαγώνιο και στην υποδιαγώνιο.

- Αν $p = 1, q = 0$ (αντ. $p = 0, q = 1$) $\rightarrow$ κάτω (άνω) διδιαγώνιο.

- Τα τριγωνικά και τα Hessenberg μητρώα έχουν μορφή ζώνης άνω ή κάτω της διαγωνίου: π.χ. $p = 1$ για άνω Hessenberg και $p = 0$ για άνω τριγωνικό.

Αν A είναι μητρώο ζώνης άνω εύρους q και κάτω εύρους p, ο μέγιστος αριθμός μη μηδενικών στοιχείων του μητρώου είναι

$$\mathrm{nnz} := n + (p+q)n - \frac{p(p+1) + q(q+1)}{2}$$

επομένως για $p, q \ll n$ ο αριθμός μη μηδενικών στοιχείων είναι $\mathrm{O}(n)$ και το μητρώο χωρά σε $p + q + 1$ μονοδιάστατους πίνακες (στήλες) μήκους n.

Ισχύει ότι $e_i^T A(p|q) e_j = 0$ αν $i > j + p$ ή αν $j > i + q$.
Θεώρημα 2. *Έστω μητρώα ζώνης* $A(p_1|q_1), B(p_2|q_2) \in \mathbb{R}^{n \times n}$. *Τότε*

1. *Το μητρώο* $C = A(p_1|q_1) \pm B(p_2|q_2)$ *έχει μέγιστο άνω εύρος* $\max(q_1, q_2)$ *και μέγιστο κάτω εύρος* $\max(p_1, p_2)$.

2. *Το μητρώο* $C = A(p_1|q_1)B(p_2|q_2)$ *έχει μέγιστο άνω εύρος* $\min(q_1 + q_2, n)$ *και μέγιστο κάτω εύρος* $\min(p_1 + p_2, n)$.

□

Παρατηρήσεις Τα μητρώα για τα οποία $\mathrm{nnz} = \mathrm{O}(n)$ λέγονται **αραιά** και έχουν πολλά σημαντικά χαρακτηριστικά.

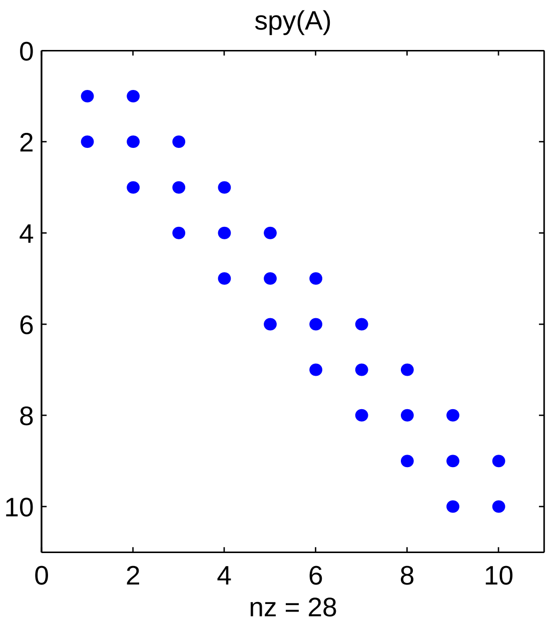
- Μπορούν να αποθηκευτούν σε μικρότερο χώρο
- Οι πράξεις με αυτά μπορούν να γραφτούν με τρόπο που λαμβάνει υπόψη του τη γνωστή μηδενική δομή των μητρώων.
- Αν τελέσουμε πράξεις μεταξύ αραιών μητρώων, το αποτέλεσμα μπορεί να μην είναι αραιό. Π.χ αν $A(1|1) \in \mathbb{R}^{n\times n}$ τότε το A^{n-1} έχει άνω και κάτω εύρος n.

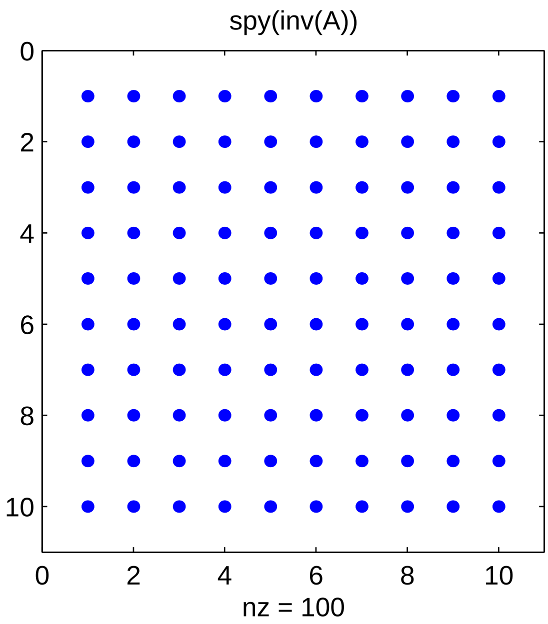
ΠΡΟΣΟΧΗ - 1) Το αντίστροφο αραιού μητρώου δεν είναι αραιό!

ΠΡΟΣΟΧΗ - 2) Αν $PA = LU$, οι παράγοντες L, U μπορεί να έχουν περισσότερα μη μηδενικά από τον A.

Έστω το τριδιαγώνιο μητρώο

```
A=toeplitz([2 -1 zeros(1,8)])
```





Τότε

$$\Downarrow$$

ένας ακόμα λόγος να μην αντιστρέφουμε!!

Επίσης όμως:

```
A =
    1.0952    0.9650    0.4745    0.0026    0.3095
    0.3971    1.0000         0         0         0
    0.3861         0    1.0000         0         0
    0.2550         0         0    1.0000         0
    0.9670         0         0         0    1.0000
>> [L,U]=lu(A)
```

```
L =
    1.0000         0         0         0         0
    0.3626   -0.7630   -0.4917   -0.0027    1.0000
    0.3525    0.3993    1.0000         0         0
    0.2328    0.2637    0.0000    1.0000         0
    0.8829    1.0000         0         0         0
U =
    1.0952    0.9650    0.4745    0.0026    0.3095
         0   -0.8520   -0.4190   -0.0023    0.7267
         0         0    1.0000    0.0000   -0.3993
         0         0         0    1.0000   -0.2637
         0         0         0         0    0.2452
```

και έχουμε παράγοντες με πολύ περισσότερα μη μηδενικά στοιχεία από το A. Προσέξτε: Εδώ, με μια απλή αναδιάταξη των γραμμών και στηλών, προκύπτουν πολύ πιο «αραιοί» παράγοντες:

```
>> B=A([2:5,1],[2:5,1])
B =
    1.0000         0         0         0    0.3971
         0    1.0000         0         0    0.3861
         0         0    1.0000         0    0.2550
         0         0         0    1.0000    0.9670
    0.9650    0.4745    0.0026    0.3095    1.0952
>> [L,U]=lu(B)
L =
    1.0000         0         0         0         0
         0    1.0000         0         0         0
         0         0    1.0000         0         0
         0         0         0    1.0000         0
    0.9650    0.4745    0.0026    0.3095    1.0000
U =
    1.0000         0         0         0    0.3971
         0    1.0000         0         0    0.3861
         0         0    1.0000         0    0.2550
         0         0         0    1.0000    0.9670
         0         0         0         0    0.2288
```

ΣΗΜΑΝΤΙΚΟ ΘΕΜΑ

Για τυχόν μητρώο με πολλά μηδενικά, πώς να αναδιατάξουμε τα στοιχεία του ώστε οι παράγοντες L, U να είναι «αρκετά αραιοί».

$$\Downarrow$$

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛ. ΙΙ (+ γραφοθεωρία)

Δομικά χαρακτηριστικά παραγοντοποίησης $PA = LU$ μητρώου ζώνης

Θεώρημα 3. *Έστω $A \in \mathbb{R}^{n\times n}$ αντιστρέψιμο και $A = A(p|q)$ και ότι χρησιμοποιείται απαλοιφή* Gauss *με μερική οδήγηση για τον υπολογισμό της παραγοντοποίησης $PA = LU$. Τότε το άνω τριγωνικό μητρώο U έχει ημιεύρος $p+q$ και το κάτω τριγωνικό μητρώο L έχει κατά μέγιστο $p+1$ μη μηδενικά στοιχεία ανά στήλη.* $\square$

Παράδειγμα: Επιβεβαιώστε τα παραπάνω βλέποντας σε μεγέθυνση για κάποιο n το

```
A=diag(rand(n,1))+diag(rand(n-1,1),-1)+diag(rand(n-1,1),1);
[L,U,P]=lu(A); subplot(1,3,1);spy(A);
subplot(1,3,2); spy(L);subplot(1,3,3);spy(U);
```

Στην παρακάτω εικόνα, $n = 20$: