ΚΡΙΤΗΡΙΑ ΑΞΙΟΛΟΓΗΣΗΣ ΜΕΘΟΔΩΝ ΚΑΙ ΕΡΓΑΛΕΙΩΝ ΤΟΥ ΕΥ

1) ΑΚΡΙΒΕΙΑ

σφάλματα που οφείλονται: λάθη δεδομένων.

2) στη διακριτοποίηση εξισώσεων

3) διακριτοποίηση πραγματικών αρ. με α.κ.υ. σ' πεπερασμένης ακρίβειας αριθμητικές πράξεις με α.κ.υ.

4) περιορισμός «πεπερασμένης» επανάληψης στις επαναληπτικές μεθόδους για την εύρεση αποτελέσματος.

5) σε δεδομένα ή σε ενδιάμεσα αποτελέσματα που δεν έχουν προβλεφθεί από τη λογική της μεθόδου επίλυσης.

2) ΤΑΧΥΤΗΤΑ

3) ΚΟΣΤΟΣ

βαθμός προσέγγισης + κόστος άμεσα συνυφασμένα.

ΥΛΙΚΟ ΚΑΙ ΑΡΧΙΤΕΚΤΟΝΙΚΗ

1) Αρχιτεκτ. Risc: αρχεία καταχωρητών, οργάνωση LOAD-STORE, εύρους χρήση pipel

2) ΙΕΡΑΡΧΙΚΗ ΟΡΓΑΝΩΣΗ ΣΥΣΤΗΜΑΤΟΣ ΑΠΟΘΗΚΕΥΣΗΣ

καταχωρητές, κρυφή μνήμη, κ. μνήμη, δευτερ. μνήμη/δίσκος.

3) παράλληλα συστήματα, superscalar επεξεργαστές, clusters κλπ.

ΣΚΟΠΟΣ ΤΟΥ ΕΥ: Ο σχεδιασμός, η ανάπτυξη και η χρήση αποδοτικών ~~[illegible]~~ εργαλείων που βοηθούν στην πρακτική επίλυση των μαθηματικών μοντέλων της επιστήμης και της τεχνολογίας.

ΥΠΟΛΟΓΙΣΤΙΚΟ ΜΟΝΤΕΛΟ ΜΕ ΙΕΡΑΡΧΙΑ ΜΝΗΜΗΣ

↳ 1 επεξεργαστή, καταχωρητές, κρυφή μνήμη, κύρια μνήμη.

* Ο επεξεργαστής προσπελαύνει στοιχεία μνήμης με load/store.
* κρυφή μνήμη: k στοιχεία, κύρια μνήμη M>>k στοιχεία.
* Χρόνος για εκτέλεση πράξης σε α.κ.υ. ενιαίος → $\boxed{\tau_{αρθ}}$
* $\boxed{\tau_{μεμ}}$: χρόνος προσκόμισης στοιχείου από κεντρική μνήμη σε καταχωρητή για επεξεργασία ΜΕ ΕΝΤΟΛΗ LOAD ΑΠΟΚΛΕΙΣΤΙΚΑ.
* $\boxed{\tau^{(0)}_{μεμ}}$: χρόνος προσκόμισης (με load) στοιχείου από κρυφή μνήμη σε καταχωρητή

$<< \tau_{μεμ}$

* ΠΡΟ ΤΗΣ ΕΝΑΡΞΗΣ ΕΚΤΕΛΕΣΗΣ ΤΟΥ ΠΡΟΓΡΑΜΜΑΤΟΣ ΟΛΑ ΤΑ ΣΤΟΙΧΕΙΑ ΕΙΣΟΔΟΥ ΒΡΙΣΚΟΝΤΑΙ ΣΕ ΚΕΝΤΡΙΚΗ ΜΝΗΜΗ.
* Η ΕΚΤΕΛΕΣΗ ΤΕΛΕΙΩΝΕΙ ΟΤΑΝ ΟΛΑ ΤΑ ΣΤΟΙΧΕΙΑ ΑΠΟΘΗΚΕΥΤΟΥΝ ΠΙΣΩ ΣΕ ΚΕΝΤΡΙΚΗ ΜΝΗΜΗ

Θεωρούμε $\tau_{μετ}^{(0)} = 0$.

ΜΕΤΡΗΤΕΣ

Ω: αριθμός $\overset{\text{πράξεων}}{\text{α.κ.υ}}$ (flops)

Φ: αριθμός μεταφορών μεταξύ κύριας μνήμης και καταχωρητών ή κρυφής μνήμης

Φ_{min} - ελάχιστο Φ αν διαθέτατε απεριόριστη μνήμη όλα τα επίπ[εδα]

ΘΕΩΡΗΜΑ: Αν ο αριθμός δεδομένων εισόδου είναι n και ο αλγόριθμος τα χρησιμοποιεί όλα για να υπολογίσει τουλάχιστον 1 αποτελ., τότε $\Phi_{min} \geq n+$

$\Downarrow$ n load

$\Downarrow$ τουλάχ. 1 store.

m αποτελέσματα $\Rightarrow$ $\Phi_{min} \geq n+m$.

Mflops/s: αριθμός πράξεων ανά μονάδα χρόνου.

ΚΟΣΤΟΣ T: $T = T_{αρθ} + T_{μετ} \Rightarrow$

$$T = \tau_{αρθ}\Omega + \tau_{μετ}\Phi \Rightarrow$$

$$T = T_{αρθ}\left(1 + \frac{\tau_{μετ}\Phi}{\tau_{αρθ}\Omega}\right) \Rightarrow$$

$$T = T_{αρθ}\left(1 + \mu\frac{\tau_{μετ}}{\tau_{αρθ}}\right)$$

$\mu: \frac{\Phi}{\Omega}$ (αριθμός μεταφορών / αριθμητική πράξη)

$\mu_{min} = \frac{\Phi_{min}}{\Omega}$

$$T = T_{αρθ}\left(1 + \mu\frac{\tau_{μετ}}{\tau_{αρθ}}\right) \geq T_{αρθ}\left(1 + \mu_{min}\frac{\tau_{μετ}}{\tau_{αρθ}}\right)$$

[illegible]

$E_{abs}(\hat{x}) = \|x-\hat{x}\|$
$E_{rel}(\hat{x}) = \dfrac{\|x-\hat{x}\|}{\|x\|}$ } με νόρμες.

«Απώλεια πληροφορίας» → από πολλά στοιχεία καταλήγουμε σε ένα βαθμωτό.

απόλυτο σφάλμα/στοιχείο: $|x_1-\hat{x}_1|, |x_2-\hat{x}_2|, \ldots$
σχετικό —//— —//— : $\dfrac{|x_1-\hat{x}_1|}{|x_1|}, \dfrac{|x_2-\hat{x}_2|}{|x_2|}, \ldots, \dfrac{|x_n-\hat{x}_n|}{|x_n|}$

Μέγιστο σφάλμα/στοιχείο: $\boxed{\max_i \dfrac{|x_i-\hat{x}_i|}{|x_i|}}$

Σύστημα α.κ.υ. $\mathcal{F} \subset \mathbb{R}$: ΣΤΟΙΧΕΙΑ ΤΗΣ ΜΟΡΦΗΣ $\boxed{y = \pm m \times \beta^{e-t}}$
β, t σταθερές του συστήματος.
m: «ακέραιο μέρος του y»
$\frac{m}{\beta^t} < 1$: «δεκαδικό μέρος του y» (ουρά του y)

β: βάση του ...
t: αριθμός ... για κωδι... που m ... βάση β
$(e_{min}, e_{ma...}$ εύρος εκθ...
$\mathcal{F}(\beta, t, e_{min}, e...$

ΠΕΠΕΡΑΣΜΕΝΟΥ ΜΕΓΕΘΟΥΣ.

- Τα στοιχεία του $\mathcal{F}$ φράσσονται κατ'απόλυτη τιμή από ένα μέγιστο $M = m_{max} \times \beta^{e_{max}-t}$ και ένα ελάχιστο $\mu = m_{min} \times \beta^{e_{min}-t}$
- Η τάξη μεγέθους ενός α.κ.υ. καθορίζεται από το e ενώ η «διακρίβ... του συστήματος από το t. («ακρίβεια» του $\mathcal{F}$).

ΑΝΤΙΣΤΟΙΧΙΑ $\mathbb{R}$ και F: $fl: \mathbb{R} \to F$

$fl(x) \in F$
$x \in \mathbb{R}$.

<u>Στρογγύλευση</u>

$x \in G$ και $x \notin F$.

«προς τον πλησιέστερο» → $fl(x)$ είναι ο α.κ.υ. $y \in F$ που βρίσκεται πλησιέστερα στον x: $y = fl(x)$ όπου $y = \arg\min_{y^* \in F} |y^*-x|$

Όταν x στο μέσο $(\overset{\in F}{y_-}, y_+)$ → ① στρογγύλευση «προς το ζυγό»: μεταξύ των y_- και y_+ επιλέγεται ο αριθμός εκείνος του οποίου το τελευταίο ψηφίο στην ουρά είναι ζυγός.

② «στρογγύλευση προς τον πλησιέστερο και μακρύτερα από το 0»

③ αποκοπή: $fl(x) = \text{sign}(x) \max(|y_-, y_+|)$

<u>Κανονικοποίηση & κρυμμένο bit</u>

1ο ψηφίο του m (στη βάση β) μη μηδενικό. Αν $\beta = 2$, $d_1 \neq 0 \Rightarrow d_1 = 1 \Rightarrow$ ΔΕΝ ΤΟ ΑΠΟΘΗΚΕΥ... (αλλά πρέπει να το λαμβάνουμε υπόψη)

$\boxed{\begin{array}{l} y = \pm \beta^e \times .d_1 d_2 \ldots d_t \\ 0 \leq d_i \leq \beta - 1,\ d_1 \neq 0 \end{array}}$ *

<u>ΙΔΙΟΤΗΤΕΣ ΣΥΣΤΗΜΑΤΟΣ Α.Κ.Υ</u>

$y \in F$ και $y \neq 0$ τότε $\beta^{e_{min}-1} \leq |y| \leq \beta^{e_{max}}(1-\beta^{-t})$

$z \in G$ και z_-, z_+ οι πλησιέστεροι α.κ.υ. που εσωκλείουν το z. Έστω ότι $z = \underset{t \text{ ψηφία}}{.x\ldots x}\, \eta \times \beta^e$ και $z_- = m \times \beta^{e-t}$, $z_+ = (m+1)\beta^{e-t} \Rightarrow \boxed{z_+ - z_- = \beta^{e-t}}$

(η: οποιαδήποτε ακολουθία ψηφίων που χρειάζονται για να αναπαραστ...)

→

→ Η απόσταση z από τον πλησιέστερο ... είναι ... στη μέση του διαστήματος. Επομένως, με στρογγύλευση προς τον πλησιέστερο

$|z - fl(z)| = \frac{\beta^{e-t}}{2}$ ← ΜΕΓΙΣΤΟ ΑΠΟΛΥΤΟ ΣΦΑΛΜΑ.

ΓΕΝΙΚΑ

$$|z - fl(z)| \le \overbrace{0.0 \ldots 0}^{t}\eta \times \beta^{e} \le \frac{\beta}{2} \cdot \beta^{-t-1} \cdot \beta^{e} \le \frac{\beta^{e-t}}{2}$$

$$\left|\frac{z - fl(z)}{z}\right| \le \frac{\frac{\beta^{e-t}}{2}}{\beta^{-1} \cdot \beta^{e}} = \frac{\beta^{1-t}}{2} = u$$

ΜΟΝΑΔΑ ΣΤΡΟΓΓΥΛΕΥΣΗΣ

$$\boxed{u = \frac{\beta^{1-t}}{2}}$$

$z \in G \Rightarrow fl(z) = z(1+\delta)$, για κάποιο $|\delta| < u$

Οι αποστάσεις μεταξύ διαδοχικών α.κ.υ. μεγαλώνουν κατά β καθώς οι αριθμοί περνούν από διάστημα που αντιστοιχεί στον εκθέτη β^e σε αυτό που αντιστοιχεί σε εκθέτη β^{e+1} (wobbling).

$$|m\beta^{e-t} - (m+1)\beta^{e-t}| = \beta^{e-t}$$

$$|m\beta^{e+1-t} - (m+1)\beta^{e+1-t}| = \beta^{e+1-t}$$

→ Το «έψιλον της μηχανής»

Για κάθε σύστημα α.κ.υ. ισχύει ότι υπάρχει ένας μικρός αριθμός ε για τον οποίο ισχύει ότι κάθε α.κ.υ. x $0 \le x \le \varepsilon$ ικανοποιεί τη σχέση $1 \tilde{+} x = 0$.

Ορισμός: Το «ε της μηχανής», ε_M, είναι η απόσταση από το 1 ως τον αμέσως μεγαλύτερο α.κ.υ. δηλ. $\varepsilon_M := \delta(1, 1^+)$

u: ελάχιστη δυνατή διακριτότητα του συστήματος α.κ.υ.

Γνωρίζουμε ότι το σύστημα α.κ.υ. χρησιμοποιεί t-1 δυαδικά ψηφία μετά την υποδιαστολή τότε το στοιχείο μετά το 1 θα είναι το $1 + 2^{1-t}$

Για $\beta = 2 \rightsquigarrow u = \frac{2^{1-t}}{2} = 2^{-t} \Rightarrow \varepsilon_M = 2u$ $(\varepsilon_M = 2^{1-t})$.

→ Το σύστημα F δεν είναι κλειστό ως προς τις συνηθισμένες πράξεις του $\mathbb{R}$.

Δηλ. υπάρχουν $x, y \in F$ τέτοια ώστε $x \odot y \notin F$.

$x, y \in F \rightsquigarrow x \tilde{\odot} y = fl(x \odot y) \in F$ ← υλοποίηση.

σα να εκτελείται η πράξη ακριβώς στο $\mathbb{R}$ & μετά να στρογγυλεύεται.

«ακριβής στρογγύλευση»

Πρέπει να χρησιμοποιήσουμε μερικά μόνο ακόμη bits και sticky bit.

ψηφίο προστασίας, ψηφίο στρογγύλευσης

→ ψηφίο προστασίας

→ χωρίς ψηφίο προστασίας χειρότερο διάδοχο: $fl(x \pm y) = (1+\alpha)x \pm (1+\beta)y$, $|\alpha|, |\beta| \le u$

Μετάδοση σφάλματος

• κανονικοποιημένοι αριθμοί, στρογγύλευση προς πλησιέστερο κ' προς ζυγό.

- $x, y \in F$, όχι υπερχείλιση ή υποχείλιση (δηλ. $x \odot y \in G$), τότε

$$\frac{|fl(x \odot y) - (x \odot y)|}{|x \odot y|} \leq u \quad , x \odot y \neq 0.$$

- $x, y \in F$ κ' $x \odot y \in G$, τότε $fl(x \odot y) = (x \odot y)(1+\delta)$, $|\delta| \leq u$

και $fl(x \odot y) = \frac{x \odot y}{1+\delta}$, $|\delta| \leq u$.

Αν $x \odot y \in F \Rightarrow \delta =$

Στοιχειώδεις πράξεις μητρώων κ' βαθμωτών

- $fl(a_{ij}) = a_{ij}(1+\delta_{ij})$, $|\delta_{ij}| \leq u$
- $|fl(A) - A| \leq |A| u$
- $fl(\beta \cdot A) = \beta \cdot A + E$, $|E| \leq u|\beta A|$
- $fl(A+B) = (A+B) + E$, $|E| \leq u|A+B|$

ΑΣΚΗΣΗ 3.3.6. $x, y \in F$, $x^2+y^2 \in G$, $x > 0 \in F$ $fl(\sqrt{x}) = \sqrt{x}(1+\delta)$, $|\delta| \leq u$

Σφάλμα για $z = \frac{x}{\sqrt{x^2+y^2}}$.

$t_1 \leftarrow fl(x * x)$ $\quad x^2(1+\delta_1)$

$t_2 \leftarrow fl(y * y)$ $\quad y^2(1+\delta_2)$

$t_3 \leftarrow fl(t_1 + t_2)$ $\quad (t_1+t_2)(1+\delta_3)$

$t_3 \leftarrow fl(\sqrt{t_3})$ $\quad \sqrt{t_3}(1+\delta_4)$

$t_1 \leftarrow fl\left(\frac{x}{t_3}\right)$ $\quad \frac{x}{t_3}(1+\delta_5)$

$$t_1 = \frac{x}{t_3}(1+\delta_5) = \frac{x}{\sqrt{t_3}(1+\delta_4)}(1+\delta_5) = \frac{x}{\sqrt{(t_1+t_2)(1+\delta_3)}\,(1+\delta_4)}(1+\delta_5) = \frac{x}{\sqrt{(x^2(1+\delta_1)+y^2(1+\delta_2))(1+\delta_3)}\,(1+\delta_4)}(1$$

$|\delta_i| \leq u$.

όχι ΚΑΤΑΣΚΕΥΗ ΣΕ ΤΕΛΙΚΟ ΑΠΟΛΥΤΟ ΚΑΙ ΣΧΕΤΙΚΟ ΣΦΑΛΜΑ ΠΟΥ ΕΜΠΕΡΙΕΧΕΙ Η ΠΡΟΣΕΓΓΙΣΗ ΤΟΥ z με το t_1.

$F(\beta, t, e_{min}, e_{max})$

$F(2, 24, -125, 128) \longrightarrow$ σύστημα απλής ακρίβειας.

$F(2, 53, -1021, 1024) \longrightarrow$ σύστημα διπλής ακρίβειας.

$1 \leq m \leq \frac{\beta}{\beta - 1}$. $\quad \boxed{m = d_0 + d_1 \cdot \beta^{-1} + \ldots + d_{t-1} \cdot \beta^{-(t-1)}}$

ουρά ενός α.κ.υ. $\quad$ hidden bit, $d_0 \neq 0$ ($\Rightarrow d_0 = 1$ για $\beta = 2$)

$\boxed{\mu A}$: $z+x*y$

$fl(z+x*y)=(z+x*y)(1+\delta)$, $|\delta|\leq u$. $\neq fl(z+fl(xy))=(z+xy(1+\delta_1))(1+\delta_2)$.

*** ΕΠΙΛΥΣΗ ΣΦΑΛΜΑΤΟΣ ΚΑΙ ΠΟΙΟΤΗΤΑΣ ΥΠΟΛΟΓΙΣΜΩΝ**

$f: U \subset \mathbb{R}^m \to \mathbb{R}^n$

$x \in U$: τα στοιχεία του π.ο. της f.

$f(x)$: τα στοιχεία της τιμής της συνάρτησης στο x χωρίς λάθη υπολογισμών

$x^* \in F$: πραγματικά στοιχεία εισόδου που χρησιμοποιούνται στην υλοποίηση.
Αν μόνο λάθη στρογγύλευσης $\Rightarrow x^*=fl(x)$.

$f(x^*)$

f_{prog}: υλοποίηση της f με πράξεις α.κ.υ. στο σύστημα F.

σχετικό σφάλμα: $\dfrac{\|f_{prog}(x^*)-f(x)\|}{\|f(x)\|}$

απόλυτο —"— : $\|f_{prog}(x^*)-f(x)\|$.

Αν το σφάλμα είναι μικρό της τάξης του ϵ της μηχανής ο αλγόριθμος f είναι ακριβής

ΕΜΠΡΟΣ ΣΦΑΛΜΑ

κατάσταση προβλήματος: μελέτη της ευαισθησίας των αποτελεσμάτων (της f) ως προς τις αναταράξεις των αποτελεσμάτων ή μόνο ως προς αυτές.

ανεξάρτητη από υλοποίηση του αλγορίθμου.

δείκτης κατάστασης προβλήματος: $\boxed{\dfrac{f'(x)\cdot x}{f(x)}}$ $cond(f;x^*)$

"Αριθμητικά ευσταθείς αλγ." - «εμπρός-πίσω ευστάθεια»: Για κάθε $x\in U$ υπάρχει κάποιο x^* κοντά στο x τέτοιο ώστε το $f_{prog}(x)$ να είναι κοντά στο $f(x^*)$. $\boxed{f_{prog}(x)\approx f(x^*)}$

- **Πίσω ευσταθής**: Έστω ότι υπάρχει x^* κοντά στο x τέτοιο ώστε $\boxed{f_{prog}(x)=f(x^*)}$

***ανάλυση σφάλματος προς τα εμπρός**: Ξεκινάμε από αρχικό σφάλμα στο $\|x^*-x\|$ και παρακολουθάμε διάδοση του από είσοδο προς έξοδο καθώς εκτελούνται οι πράξεις α.κ.υ.

↳ όχι πρακτικό η σύνθεση επί πέρας σφαλμάτων.

↳ η ανάλυση πρέπει να γίνει σε πολλές διαστάσεις ($f:\mathbb{R}^n\to\mathbb{R}^m$)

$P_n=\prod_{i=1}^{n}(1+\delta_i)$, $|\delta_i|\leq u$

$(1-u)^n\leq P_n\leq(1+u)^n$

$P_n=1+nu+O(u^2)$

Αν $|\delta_i|\leq u$ και $\rho_i=\pm1$ για $i=1{:}n$ και $nu<1$ τότε

$$\prod_{i=1}^{n}(1+\delta_i)^{\rho_i}=1+\theta_n,\quad |\theta_n|\leq\frac{nu}{1-nu}=:\gamma_n.$$

$\prod_{i=1}^{n}(1+\delta_i)^{\rho_i}=:\langle n\rangle$.

$\Rightarrow \langle n\rangle\times\langle k\rangle=\langle n+k\rangle$

$\langle n\rangle/\langle k\rangle=\langle n+k\rangle$

$\gamma_n=\dfrac{nu}{1-nu}=nu\left(\dfrac{1}{1-nu}\right)\leq nu\left(1+nu+(nu)^2+\dots\right)=nu+O(u^2)$

$\prod_{i=1}^{n}(1+\delta_i)\leq(1+u)^n<e^{nu}$

$$\|z^*-z\| = \|f_{prog}(x) - f(x)\| = \|f(x^*) - f(x)\|$$

αν υπο. τε πράξεις α.κ.υ. — ΕΥΣΤΑΘΕΙΑ

Εξετάζε τη διαφορά $\|f(x^*)-f(x)\|$. Αν το $\|x^*-x\|$ μικρό, τότε το πρόβλημα ανάγεται στο μαθηματικό πρόβλημα της ευαισθησίας του f σε [redacted] μικρές διαταράξεις των στοιχείων εισόδου. (ΟΧΙ ΑΝΑΛΥΣΗ ΣΦΑΛΜΑΤΟΣ ΕΝΔΙΑΜ. ΠΡΑΞΕΩΝ)

Παράδειγμα: $f(x_1,x_2,x_3) = (x_1+x_2)+x_3$ [redacted]

[redacted] $f_{prog}(x_1,x_2,x_3) = ((x_1+x_2)(1+\delta_1)+x_3)(1+\delta_2)$, $|\delta_i| \le u$

$f_{prog}(x_1,x_2,x_3) = f(\tilde{x}_1,\tilde{x}_2,\tilde{x}_3)$ όπου

$\tilde{x}_1 = x_1(1+\delta_1)(1+\delta_2)$

$\tilde{x}_2 = x_2(1+\delta_1)(1+\delta_2)$

$\tilde{x}_3 = x_3(1+\delta_2)$

$$|\tilde{x}_1 - x_1| = |x_1(1+\delta_1)(1+\delta_2) - x_1| = |x_1((1+\delta_1)(1+\delta_2)-1)| = |x_1(\cancel{1}+\delta_2+\delta_1+\delta_1\delta_2-\cancel{1})| = |x_1(\delta_1+\delta_2+\delta_1\delta_2)| \le |x_j|$$

$$|\tilde{x}_2 - x_2| = |x_2(\delta_1+\delta_2+\delta_1\delta_2)|$$

$$|\tilde{x}_3 - x_3| = |x_3(1+\delta_2)-x_3| = |x_3\delta_2| \le u|x_3|$$

$$\frac{|\tilde{x}_j - x_j|}{|x_j|} \le \rho_j u, \quad \rho_j = 3 \quad j=1,2, \quad \rho_3 = 1$$

$|f_{prog}(x_1,x_2,x_3) - f(x_1,x_2,x_3)| = |f(\tilde{x}_1,\tilde{x}_2,\tilde{x}_3) - f(x_1,x_2, \dots$

$\Rightarrow$ ΠΙΣΩ ΕΥΣΤΑΘΗΣ ΑΛΓΟΡΙΘΜΟΣ.

+ ΜΙΚΡΟΣ ΔΕΙΚΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΠΡΟΒΛΗΜΑΤΟΣ $\boxed{\frac{f'(x)x}{f(x)}}$

→ Κατάσταση αλγορίθμου: μελέτη της επίδρασης της α.κ.υ. σε συγκεκριμένη υλοποίηση του αλγορίθμου που αντιστοιχεί σε υπολογιστό. ($cond\, f_{prog}$)

Ορισμός:
$$\boxed{cond(f;x^*) = \lim_{\delta\to 0} \sup_{\|h\|=\delta} \left\{ \frac{\frac{\|f(x^*+h)-f(x^*)\|}{\|f(x^*)\|}}{\frac{\|h\|}{\|x^*\|}} \right\}}$$

σχετ. αλλαγή του $f(x^*)$

σχετ. αλλαγή του x^*

- Αν η απεικόνιση διαθέτει παράγωγο στο x^* → $cond(f;x^*) = \frac{\|x^*\|}{\|y^*\|} \left\|\frac{\partial f}{\partial x}\right\|_{x^*}$
- $\sup_{\|h\|=\delta} \frac{\|f(x^*+h)-f(x^*)\|}{\|h\|} = \sup_{\|h\|=\delta} \frac{\|f(h)\|}{h}$ ← ανεξάρτητο του x^* και του δ.

$\Rightarrow cond(f;x^*) = \frac{\|x^*\|}{\|y^*\|}\|f\|$.

$\|A\|_M = \sup_{x\in\mathbb{R}^n, x\neq 0} \frac{\|Ax\|}{\|x\|}$

$\|A\|_M = \sup_{x\in\mathbb{R}^n} \frac{\|Ax\|}{\|x\|}$ τέτοιο ώστε

→ Υπόθεση: [redacted] Αν ένας αλγορ. f υλοποιείται μέσω προγράμματος f_{prog} και χρησιμοποιεί διάνυσμα εισόδου x, τότε υπάρχει διάνυσμα εισόδου x_{prog} ώστε $f_{prog}(x) = f(x_{prog})$. ← για πίσω ευστάθεια

↳ **Ορισμός:** Τότε η κατάσταση υλοποίησης του αλγορ. ορίζεται ως η βέλτιστη (ελάχιστη) τιμή $cond(f_{prog})$ για την οποία ισχύει
$$\boxed{\frac{\|x - x_{prog}\|}{\|x\|} \le cond(f_{prog}) \cdot u}$$

$$\frac{\|f_{prog}(x^*) - f(x)\|}{\|f(x)\|} \le \frac{\|f(x^*_{prog}) - f(x^*)\|}{\|f(x)\|} + \frac{\|f(x^*) - f(x)\|}{\|f(x)\|}$$

$$\frac{\|f(x^*_{prog}) - f(x)\|}{\|f(x)\|} \le \frac{\|f(x^*_{prog}) - f(x^*)\|}{\|f(x^*)\|} \le cond(f;x^*)\frac{\|x^* - x^*_{prog}\|}{\|x^*\|} \le cond(f;x^*)\, cond(f_{prog})\, u$$

$$\frac{\|f(x^*) - f(x)\|}{\|f(x)\|} \le cond(f;x)\frac{\|x^*-x\|}{\|x\|} \le cond(f;x) E$$

$\Rightarrow$

$$\boxed{\frac{\|f_{prog}(x^*) - f(x)\|}{\|f(x)\|} \le cond(f;x^*)\, cond(f_{prog})\, u + cond(f;x) E}$$

$\boxed{\text{Προς τα εμπρός σφάλμα} < \text{δείκτης κατάστασης} \times \text{πίσω σφάλμα}}$

Υπολογισμός εσωτερικού γινομένου (DOI)

$f([x;y]) = x^T y \quad , x, y \in \mathbb{R}^n$.

Θέτουμε $X := [x;y] \in \mathbb{R}^{2n}$

$$\text{cond}(f;X) = \frac{\|X\|}{\|x^T y\|} \left\| \frac{\partial f}{\partial X}\Big|_{[x;y]} \right\|$$

$$\frac{\partial f}{\partial X}\Big|_{[x;y]} = [y;x] \in \mathbb{R}^{1\times 2n}$$

$$\text{cond}(f;X) = \frac{\|[x;y]\|}{|x^T y|} \|[y;x]\|$$

$\Rightarrow \text{cond}(f;X) \leq \dfrac{\|[x;y]\|^2}{|x^T y|}$. Πρόβλημα αν $\cos(x,y) \to 0$ $\Rightarrow$ μεγάλος δείκτης κατάστασης.

$\boxed{\text{ΝΟΡΜΕΣ}}$

Ανισότητα τριγωνική: $\|x+y\| \leq \|x\| + \|y\|$ / $\|x-y\| \geq \big|\|x\| - \|y\|\big|$

$\|ax\| = |a| \cdot \|x\|$.

Ανισότητα Holder: $|x^* y| \leq \|x\|_p \cdot \|y\|_q$ όπου $\frac{1}{p} + \frac{1}{q} = 1$ $\quad \frac{1}{2} + \frac{1}{2} = \quad \leadsto |x^* y| \leq \|x\|_2 \cdot \|y\|$

$\|x\|_2^2 = x^* x$

Αν $Q^* Q = I$ τότε $\|Qx\|_2 = \|x\|_2$

$\nabla \|x\|_2 = \dfrac{x}{\|x\|_2} \longleftarrow$ Η Ευκλείδια νόρμα είναι παραγωγίσιμη.

$\mu_a^\beta \|u\|_a \leq \|u\|_\beta \leq \mu_\beta^a \|u\|_a$

$a \backslash \beta$	1	2	∞
1	1	$\sqrt{n}$	n
2	1	1	$\sqrt{n}$
∞	1	1	1

Ορισμός: Έστω U γραμμικός χώρος και οι νόρμες $\|\ \|$ και $\|\ \|'$. Καλούνται ισοδύναμες αν υπάρχουν θετικές σταθερές μ_1, μ_2 τέτοιες ώστε $\mu_1 \|u\| \leq \|u\|' \leq \mu_2 \|u\|$.

Θεώρημα: ΟΛΕΣ ΟΙ ΝΟΡΜΕΣ ΣΤΟΝ $\mathbb{R}^n$ (αντ. στον $\mathbb{C}^n$) ΕΙΝΑΙ ΙΣΟΔΥΝΑΜΕΣ ΜΕΤΑΞΥ Τ

Δυϊκή νόρμα: $\|x\|_D = \max\limits_{z \neq 0} \dfrac{|z^* x|}{\|z\|}$ $\quad$ θετική ορισιμότ τριγων. ανισότ ομοιογ

δυϊκό διάνυσμα.

$x_D^* x = \|x_D\|_D \|x\| = 1$

ΝΟΡΜΑ ΜΗΤΡΩΟΥ: Αν μια συνάρτηση $\|\cdot\| : \mathbb{R}^{n\times n} \to \mathbb{R}$ ικανοποιεί εις συνθήκες Α.1.1 κ ισχύει $\|A \cdot B\| \leq \|A\| \cdot \|B\|$.

$$\|A\|_M = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \iff \|A\|_M = \sup_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} \frac{\|Ax\|}{\|x\|}$$

ΝΟΡΜΑ ΜΗΤΡΩΟΥ ΠΑΡΑΓΟΜΕΝΗ ΑΠΟ ΝΟΡΜΑ ΔΙΑΝΥΣΜΑΤΟΣ.

$A \in \mathbb{R}^{m\times n}$

$$\|A\|_F = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}$$

$$\|A\|_1 = \max_{j=1:n} \sum_{i=1}^{n} |a_{i,j}|$$

$$\|A\|_2 = \sqrt{\lambda_{max}(A^* A)}$$

μέγιστη ιδιοτιμή

ΦΑΣΜΑΤΙΚΗ ΝΟΡΜ

$$\|A\|_\infty = \max_{i=1:n} \sum_{j=1}^{n} |a_{i,j}|$$

Αν $X^* X = I$, $Y^* Y = I \Rightarrow \|XAY\|$

* Για κάθε πεπερασμένο χώρο πεπερασμένης διάστασης, οι νόρμες μητρώων είναι ισοδύναμες

$\|A\|_p \leq J_{pq} \|A\|_q$

Ορισμός: Έστω $A \in \mathbb{R}^{m \times n}$ και οι φυσικοί αριθμοί $1 \le i_1 < i_2 < i_3 < \ldots < i_k \le m$ και $1 \le j_1 < j_2 < j_3 < \ldots < j_\ell \le n$. Τότε το μητρώο μεγέθους $k \times \ell$ του οποίου το (μ,ν) στοιχείο είναι $a_{i_\mu j_\nu}$ αποκαλείται υπομητρώο της A

- $k=\ell \Rightarrow i_1 = j_1, \ldots, i_k = j_k \Rightarrow$ ΚΥΡΙΟ ΜΗΤΡΩΟ
- $i_1 = 1, \ldots, i_k = k \Rightarrow$ ΑΡΧΙΚΟ ΜΗΤΡΩΟ.

ΑΝΑΛΥΣΗ ΣΦΑΛΜΑΤΟΣ ΓΙΑ DOT

Εμπρός σφάλμα < δείκτης κατάστασης × πίσω σ

α) Ανάλυση πίσω ευστάθειας αλγορίθμου

β) —//— δείκτη κατάστασης

3) άνω φράγμα για δείκτη κατάστασης

$f([x;y]) = x^T y$, $x, y \in \mathbb{R}^n$. Θέτουμε $X := [x;y] \in \mathbb{R}^{2n}$.

$$\mathrm{cond}(f;X) = \frac{\|X\|}{\|x^T y\|} \left\| \frac{\partial f}{\partial X}\Big|_{[x;y]} \right\| = \frac{\|X\|}{\|x^T y\|} \|[y;x]\| = \frac{\|[x;y]\|}{\|x^T y\|} \|[y;x]\| =$$

$$\le \frac{\|[x;y]\|^2}{|x^T y|}$$

α) Πίσω σφάλμα.

$s_n = x^T y \qquad x, y \in \mathbb{R}^n$

$\tilde{s}_1 = x_1 y_1 (1+\delta_1)$ $\qquad |\delta_i| \le u.$

$\tilde{s}_2 = fl(\tilde{s}_1 + fl(x_2 y_2)) = (x_1 y_1 (1+\delta_1) + x_2 y_2 (1+\delta_2))(1+\delta_3)$

$= x_1 y_1 (1+\delta_1)(1+\delta_3) + x_2 y_2 (1+\delta_2)(1+\delta_3).$

$\vdots$

$$\tilde{s}_n = x_1 y_1 \prod_{\substack{j=1 \\ j\ne 2}}^{n+1} (1+\delta_j) + x_2 y_2 \prod_{j=2}^{n+1} (1+\delta_j) + x_3 y_3 \prod_{j=3}^{n+1} (1+\delta_j) + \ldots + x_n y_n \prod_{j=n}^{n+1} (1+\delta_j)$$

$$\tilde{s}_n = x_1 y_1 (1+\theta_n) + x_2 y_2 (1+\hat{\theta}_n) + x_3 y_3 (1+\theta_{n-1}) + \ldots + x_n y_n (1+\theta_2)$$

[redacted] Υπολογισθέν DOT $\tilde{\hat{s}}_n$ είναι ακριβές εσωτ. γινόμενο για στοιχεία $x_1, \ldots, x_n, y_1(1+\theta_n), y_2(1+\hat{\theta}_n), y_3(1+\theta_{n-1}), \ldots, y_n(1+\theta_2)$.

$|\theta_j| \le \gamma_j = \frac{ju}{1-ju}$. ΑΠΟΔΕΙΞΑΜΕ ΟΤΙ Ο ΑΛΓΟΡΙΘΜΟΣ ΕΙΝΑΙ ΠΙΣΩ ΕΥΣΤΑΘΗΣ ΓΙΑΤΙ:

$$fl(x^T y) = x^T (y + \Delta y) = (x + \Delta x)^T y$$

$$|\Delta x| \le \gamma_n |x|$$

$$|\Delta y| \le \gamma_n |y|$$

ενημέρωση 1ης τάξης: $C \leftarrow \underset{n_1 \times n_2}{C} + \underset{n_1 \times 1}{a} \underset{1 \times n_2}{b^T}$

$\Omega = 2 n_1 n_2 \left(= \overset{\text{πολ/σμοί}}{n_1 n_2} + \overset{\text{προσθ.}}{n_1 \times n_2} \right).$

$\Phi_{min} = \underbrace{2 n_1 n_2}_{\text{load}} + n_1 + n_2.$ ← Όταν έχουμε αρκετούς καταχωρητές για την αποθήκευση των στοιχείων του a.

* Ο ρυθμός που οδηγεί στην καλύτερη υλοποίηση καθορίζεται απ' το μέγεθος της κρυφής μνήμης & των καταχωρητών.

* Η ελαχιστοποίηση έγινε εξετάζοντας μόνο για το [redacted] μικρότερο πρόβλημα.

* Αύξηση της απόδοσης εξαρτάται από βελτιώσεις που παρουσιάζουν αύξηση χρον. τοπικότητας

* Για να δείξουμε καλύτερη υλοποίηση πρέπει να εξετάσουμε και αυτή που βασίζεται σε εναλλακτική σειρά από βρόχους.

Σφάλμα υπολογισμού εξωτερικού γινομένου

$\boxed{A = xy^T}$

$fl(a_{ij}) = x_i y_j (1+\delta_{ij}), \quad |\delta_{ij}| \le u.$

από $fl(\beta A) = \beta A + E$, $|E| \le u|\beta A|$

$fl(A) = xy^T + E$ όπου $|E| \le |x||y|^T u$

→ Εμπρός-πίσω ευστάθεια, όχι πίσω σταθερή μπορεί εξωτερικού γινομένου
γιατί αν ήταν θα υπήρχαν $\Delta x, \Delta y$ τέτοια ώστε
$fl(A) = xy^T + E = (x+\Delta x)(y+\Delta y)^T$!! όχι
Τότε θα ίσχυε $E = (x+\Delta x)(y+\Delta y)^T - xy^T \Rightarrow$

$$E = \underbrace{[x \quad x+\Delta x]}_{1\times 2} \underbrace{\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{2\times 2} \underbrace{\begin{bmatrix} (y+\Delta y)^T \\ y^T \end{bmatrix}}_{2\times 1}.$$

τάξη n — βαθμωτός.

ΜΗΤΡΩΟ × ΔΙΑΝΥΣΜΑ (MV) — Gaxpy

$y \leftarrow y + Ax$, με y: $n_1 \times 1$, A: $n_1 \times n_3$, x: $n_3 \times 1$.

$y \leftarrow y + x^T A$, με y: $1 \times n_1$, x^T: $1 \times n_3$, A: $n_3 \times n_1$, $x \in n_3 \times 1$.

$$\begin{bmatrix} a_1 & a_2 & \dots & a_{n_3} \\ a_{21} & a_{22} & & \\ \vdots & & & \\ a_{n_1} & \dots & & a_{n_1 n_3} \end{bmatrix}_{n_1\times n_3} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_3} \end{bmatrix}_{n_3\times 1} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n_3}x_{n_3} \\ \vdots \\ \vdots \end{bmatrix}_{n_1\times 1}$$

$n_1 n_3$ πολλαπλ.
+
$(n_3 - 1)n_1$ προσθ.
+
n_1 προσθ.

$\Rightarrow \Omega = n_1 n_3 + n_1 n_3 - n_1 + n_1 \Rightarrow \Omega = 2n_1 n_3$

$\Phi_{min} = n_1 n_3 + 2n_1 + n_3$

Μορφή DOT: Τα στοιχεία του y υπολογ. από το εσωτ. γινόμενο των γραμμών του A με το x. (i, j) κ. γραμμές
$y_i \leftarrow y_i + a_{i:}^T x = y_i + \sum_{k=1}^{n_3} a_{ik}\xi_k, \quad i = 1, \dots, n_1.$

Μορφή SAXPY: Τα στοιχεία του y υπολογίζονται από το γρ. συνδυασμό των στηλών του A. (j, i) κ. στήλες

$y \leftarrow y + \sum_{k=1}^{n_3} a_{:k}\xi_k$

ΠΟΛΛΑΠΛΑΣΙΑΣΜΟΣ ΜΗΤΡΩΩΝ: $C \leftarrow C + AB$

C: $n_1 \times n_2$, A: $n_1 \times n_3$, B: $n_3 \times n_2$

$\Omega = 2n_1 n_2 n_3$

$\Phi_{min} = 2n_1 n_2 + n_1 n_3 + n_2 n_3.$

6 ΔΙΑΦΟΡΕΤΙΚΟΙ ΤΡΟΠΟΙ ΠΟΛ/ΣΜΟΥ ΠΙΝΑΚΩΝ.

ΟΛΕΣ ΟΙ ΔΙΑΤΑΞΕΙΣ ΤΟΥ ΚΟΜΒΟΥ ΕΧΟΥΝ ΤΟ ΙΔΙΟ ΣΦΑΛΜΑ.

$C = AB \Rightarrow c_i = Ab_i, \quad i = 1:n$

$\tilde{c}_i = (A + \Delta A)b_i, \quad |\Delta A_i| \le \gamma_n |A|$ (πίσω σταθερό ανά στήλη)

$fl(AB) = (A+\Delta A)B, \quad |\Delta A| \le \gamma_n |A||B||B^{-1}|$

$|C - \tilde{C}| = |AB - (A+\Delta A)B| \le |\Delta A B| \le |\Delta A| \cdot |B| \le \gamma_n |A||B||B^{-1}||B| = \gamma_n |A||B|$

↳ εξοικονόμηση πολ/σμών μέσω αναδρ. τύπων ΑΛΛΑ ΧΕΙΡΟΤΕΡΗ ΑΡΙΘΜΗΤΙΚΗ ΣΥΜΠΕΡΙΦ. ΑΠΟ ΤΗΝ ΚΛΑΣΣΙΚΗ ΜΕΘΟΔΟ.

↳ $T(n) = 7T(\frac{n}{2}) + 18(\frac{n}{2})^2$

$T(n) = \Theta(n^{\log_2 7}) = \Theta(n^{2.81})$

$n = 2^k$.

<u>Ανάλυση σφάλματος</u>: Θ. Brent: $A, B \in \mathbb{R}^{n \times n}$ με $n = 2^k$ και έστω $\hat{C} = AB$ υπολ. με τι Strassen ενώ από τη διάσταση $n_0 = 2^{r+1}$ και κάτω χρησιμοποιείται κλασσικός πολ/σμός. Τότε: $\|\hat{C} - C\| \le \left[\left(\frac{n}{n_0}\right)^{\log_2 12} (n_0^2 + 5n_0) - 5n \right] u \|A\| \cdot \|B\| + O(u^2)$.

ΑΣΚΗΣΕΙΣ

1) Πόσες πράξεις για $(I - uv^T)x$, $I \to n \times n$ ταυτοτικό, $x, u, v \in \mathbb{R}^n$.

$= x - u(v^T x)$

$v^T x \to 2n-1$ πράξεις
$u(v^T x) \to n$ πράξεις
$x - u(v^T x) \to n$ πράξεις

$\Big\} \Rightarrow 4n-1$ πράξεις. $= O(n)$

2) $p, q, x \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ με $a_{ij} = a_{ji} = p_i q_j$, $1 \le i \le j \le n$.

$\uparrow p_i$ $\uparrow \kappa_i$ $\uparrow \xi_i$

Πόσες πράξεις για γινόμενο Ax;

$A = A^T$ $\quad A^T x$;

ΤΕΤΡΑΓΩΝΙΚΑ ΜΗΤΡΩΑ

<u>Άνω Hessenberg</u> → σαν άνω τριγωνικό + 1 διαγώνιο κάτω από την κύρια διαγώνιο

$i > j+1 \Rightarrow a_{ij} = 0$.

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & & a_{1n} \\ a_{21} & a_{22} & \cdots & & a_{2n} \\ 0 & a_{32} & \cdots & & a_{3n} \\ 0 & 0 & a_{43} & \cdots & a_{4n} \\ \vdots & \vdots & 0 & & \\ 0 & 0 & \cdots & a_{n,n-1} & a_{nn} \end{bmatrix}$$

<u>Κάτω Hessenberg</u> → σαν κάτω τριγωνικό + 1 επιπλέον διαγώνιο πάνω από την κύρια διαγώνιο

$j > i+1 \Rightarrow a_{ij} = 0$

$$\begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & & 0 \\ a_{21} & a_{22} & a_{23} & 0 & \cdots & 0 \\ \vdots & & & a_{34} & & \\ \vdots & & & & & a_{n-1,n} \\ a_{n1} & & & & & a_{nn} \end{bmatrix}$$

Αν $C = AB$
A: κάτω Hessenberg
B: κάτω τριγωνικό.

$C = ?$ είναι κάτω Hessenb.

$a_{ij} = 0$ για $i < j$

Για να είναι το C κάτω Hessenberg αρκεί να δ.ό.τι για $j > i+1 \Rightarrow c_{ij} = 0$.

Έστω το στοιχείο c_{ij} όπου $j > i+1$. (λόγω κάτω Hessenberg)

$c_{ij} = A(i,:) * B(:,j) = A(i, 1:i+1) * B(1:i+1, j) = 0$

αφού $j > i+1$ και B κάτω τριγ.

c_{12}.

$$\begin{bmatrix} a_{11} & a_{12} & 0 & \cdots & 0 \\ a_{21} & a_{22} & a_{23} & \cdots & 0 \\ & & & & \\ & & & & \end{bmatrix} \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & a_{n3} \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} a_{11}^2 + a_{12}a_{21} & \\ & \end{bmatrix}$$

ΑΣΚΗΣΗ 4.7.1

$B = \underset{n\times n}{\hat{B}} + (xy^T)^P \quad x,y \in \mathbb{R}^n, \; p>0$

Η ύψωση βαθμωτού σε δύναμη ~~[illegible]~~ είναι ισοδύναμη με μια πράξη α.κ.υ. μ_{min} = ίσαν συναρτήσει του n.

- Θεωρούμε ότι έχουμε αρκετή κρυφή μνήμη ή καταχωρητές ώστε
  - α) να φορτώσουμε όλα τα δεδομένα στη μνήμη
  - β) να εκτελέσουμε πράξεις αποθηκεύοντας τα ενδιάμεσα αποτελέσματα τους στους καταχωρητές.
  - γ) να εγγράψουμε τα αποτελέσματα στην κύρια μνήμη.

$\Phi_{min} = 2n^2 + 2n + 1.$

load n^2 για B
load $2n$ για x, y
load p, (ω)
store n^2 για B.

$O(n^2)$ θέσεις κρυφής μνήμης. !!!

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1} \begin{bmatrix} y_1 & y_2 & \dots & y_n \end{bmatrix}_{1\times n} = \rightarrow n\times 1$$

$$\begin{bmatrix} x_1y_1 & x_1y_2 & \dots & x_1y_n \\ x_2y_1 & x_2y_2 & \dots & \\ \vdots & \vdots & & \\ x_ny_1 & x_ny_2 & \dots & x_ny_n \end{bmatrix}$$

$$(xy^T)^P = \underbrace{(xy^T)(xy^T)\dots(xy^T)}_{p \text{ φορές}} =$$

$$= x\underbrace{(y^Tx)(y^Tx)\dots(y^Tx)}_{p-1 \text{ φορές}}y^T$$

Η πράξη $\psi = y^Tx \longrightarrow 2n-1$ πράξεις.

πράξη: $\psi^{p-1}x \longrightarrow 1 + n$ για πολ/σμό (βαθμωτός × διάνυσμα).
(1: για ύψωση)

~~αναγέννηση στάθμης.~~

$B = B + (\psi^{p-1}x)y^T \longrightarrow 2n^2$ πράξεις (n^2 για προσθ. + n^2 για πολ/σμούς)

$$\Longrightarrow \Omega = 2n^2 + n + \cancel{1} + 2n - \cancel{1} = 3n + 2n^2.$$

$$\mu_{min} = \frac{\phi_{min}}{\Omega} = \frac{2n^2 + 2n + 1}{3n + 2n^2}.$$

ΑΣΚΗΣΗ 4.7.2

$A, B \in \mathbb{R}^{n\times n}$ κάτω τριγωνικά. Αλγόριθμο που να εκμεταλλεύεται δομή.

$$\begin{bmatrix} a_{11} & 0 & 0 & \dots \\ a_{21} & a_{22} & & \\ a_{31} & & \ddots & 0 \\ a_{n1} & \dots & & a_{nn} \end{bmatrix} \begin{bmatrix} b_{11} & 0 & 0 & \dots \\ b_{21} & b_{22} & & \\ \vdots & & \ddots & 0 \\ b_{n1} & \dots & & b_{nn} \end{bmatrix}$$

κάτω τριγ.
$j > i \Rightarrow a_{ij} = 0$
$j > i \Rightarrow b_{ij} = 0$.

```
C = zeros(n);
for i=1:n
    for j=1:i
        C(i,j)=A(i,j:i)*B(j:i,j);
    end
end
```

βήμα i: i πολ/σμούς + i − 1 προσθ. = 2i − 1

$$\sum_{i=1}^{n}\sum_{j=1}^{i}(2(i-j+1)-1) = \sum_{i=1}^{n}\sum_{j=1}^{i}(2i-2j+2-1) = \sum_{i=1}^{n}\sum_{j=1}^{i}(2i-2j+1)$$

$k = i - j \Rightarrow j = i - k$.
$j = 1 \quad k = i - 1$
$j = i \quad k = 0$.

→ ΘΕΩΡΗΜΑ: Έστω το σύστημα $Ax=b$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^{m}$ τότε:

- το σύστημα έχει λύση
- η τάξη του επαυξημένου συστήματος $[A, b]$ ίδια με την τάξη του A $\mathrm{rank}(A,b) = \mathrm{rank}(A)$
- Για κάθε $y \in \mathbb{R}^m$ ώστε $y^T A = 0$ ισχύει ότι $y^T b = 0$.

ΙΣΟΔΥΝΑΜΕΣ!!!

$Ax=b$

$x_{γενική} = x_{ειδική} + x_{ομογενούς}$

$\uparrow$

$Ax=0$

① Ανάγουμε $Ax=b$ σε $Ux=c$.

② Μηδεν. ελεύθ. μεταβλητές κ' βρίσκ. ειδική λύση.

③ $Ax=0$ και διαδοχικά σε κάθε ελ. μεταβλ. την τιμή 1. και άλλες 0. (ομογενής λύση) → αντιστοιχούν σε στήλες που δεν περιέχουν οδηγό

$x_{ειδική}$: από λύση συστήματος όταν όλες τις ελεύθερες μεταβλητές δοθεί η τιμή 0.

r οδηγοί ⇒ r βασικές μεταβλ. r: τάξη του A

n − r ελευθ. μεταβλ.

στήλ.

→ Αν $A \in \mathbb{R}^{n\times n}$ αντιστρέψιμο, τότε το σύστημα $Ax=b$ έχει μοναδική λύση!!!

Κανόνας Cramer

Έστω $A \in \mathbb{R}^{n\times n}$ κ' $b \in \mathbb{R}^n$. Τα ξ_j $(1 \le j \le n)$ του διανύσματος $x \in \mathbb{R}^n$ λύνουν το $Ax=b$ ικανοποιούν τις σχέσεις: $\boxed{\xi_i = \frac{\det(A(i|b))}{\det(A)}}$, $i=1,\dots,n$.

$A(i|b)$: το μητρώο που λαμβάνουμε αντικαθιστώντας την i στήλη του A με το b.

→ ΥΠΕΡΒΟΛΙΚΟΣ ΑΡ. ΠΡΑΞΕΩΝ!

→ ΜΕΓΑΛΟ ΣΦΑΛΜΑ ΣΤΡΟΓΓΥΛΕΥΣΗΣ! } ⇒ ΟΧΙ ΧΡΗΣΗ CRAMER για $Ax=b$!!!

Εύρεση Αντιστρόφου

ΔΕΝ ΕΝΔΕΙΚΝΥΤΑΙ

↳ επιπλέον πράξεις

↳ καταστροφή χρήσιμων δομικών στοιχείων μητρώου. (π.χ. A αραιό → οικον. τρόπους αποθ. ενώ A^{-1} πυκνό → δαπανηρή)

Επιλογή Μεθόδου Επίλυσης

① Από: ΜΕΓΕΘΟΣ ΜΗΤΡΩΟΥ

② Από: ΔΟΜΗ ΜΗΤΡΩΟΥ

ΤΥΠΟΙ ΜΗΤΡΩΩΝ

Πυκνό → πλήρως αποθηκεύσιμο.

Πυκνό δομημένης μορφής → πυκνό μητρώο ειδικής μορφής που χαρακτηρίζεται από λιγότερα από n^2 στοιχεία

ΕΙΔΗ

↳ Συμμετρικό: $A=A^T \Rightarrow a_{ij}=a_{ji}$. $\frac{n(n+1)}{2}$ ή λιγότερα στοιχεία αποθηκεύσιμα.

↳ Ερμιτιανό: $A=A^* \Rightarrow a_{ij}=\overline{a_{ji}}$ / αυτοσυζυγές.

↳ Τριγωνικό: Άνω → $a_{ij}=0$ για $i>j$

Κάτω → $a_{ij}=0$ για $i<j$

→ Hessenberg: Άνω → $i>j+1 \Rightarrow a_{ij}=0$ (1 διαγώνιος κάτω απ' την κύρια διαγώνιο)

Κάτω → $j>i+1 \Rightarrow a_{ij}=0$ (1 διαγώνιος πάνω απ' την κύρια διαγώνιο)

↳ Toeplitz: μητρώα των οποίων τα στοιχεία διαγώνιο ίσα μεταξύ τους. $a_{ij} = \tau_{i-j}$, όλα τα στοιχεία γνωστά από $n+n-1=2n-1$ στοιχεία 1ης γραμμής + 1ης στήλης.

↳ Hankel: τα στοιχεία /αντι-διαγώνιο ίσα μεταξύ τους.

↳ Κυκλοτερές: Τα στοιχεία κάθε γραμμής προέρχονται από διαδοχική περιστροφή στοιχείων 1ης γραμμής. [redacted] 2ταθ. διαγωνίους χαρακτηρίζ. πλήρως από [redacted] n στοιχεία 1ης γραμμής.

↳ Vandermonde : $V(\xi_1,\dots,\xi_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \xi_0 & \xi_1 & \cdots & \xi_{n-1} \\ \xi_0^2 & \xi_1^2 & & \xi_{n-1}^2 \\ \vdots & & & \\ \xi_0^{n-1} & \xi_1^{n-1} & \cdots & \xi_{n-1}^{n-1} \end{pmatrix}$

Αραιά Μητρώα : λίγα μη μηδενικά στοιχεία σε σχέση με το μέγεθός τους (nnz)

↓

$nnz = O(n)$

«πυκνότητα» $= \dfrac{nnz}{n^2}$

Αναπαράσταση με μονοδιάστατους πίνακες ή linked lists.

Αραιοί δομημένα μητρώα

↳ ταινιακά μητρώα : $m \ll n$ ώστε $a_{ij} = 0$ όταν $|i-j| > m$.

↳ τριδιαγώνια $a_{ij} = 0$ όταν $|i-j| > 2$

↳ διαγώνια μητρώα $a_{ij} = 0$ όταν $i \neq j$

↳ ομαδοποιημένα μητρώα : μητρώο που προέρχεται από αντικατάσταση κάθε μηδενικού στοιχείου αραιού μητρώου με πυκνό μητρώο

↳ block diagonial

↳ block tridiagonial.

ΓΕΝΙΚΑ

Ορθογώνιο : [redacted] $A \cdot A^T = A^T A = I$.

Μοναδιαίο : $A \cdot A^* = A^* A = I$

Κανονικό : $A^T A = AA^T$ [redacted] αν $A \in \mathbb{R}^{m \times n}$

$A^* A = AA^*$ αν $A \in \mathbb{C}^{m \times n}$

ΔΙΑΓΩΝΙΑ ΚΥΡΙΑΡΧΟ : «ανά γραμμές» όταν $|a_{ii}| \geq \sum_{j=1 \,//\, j \neq i}^{n} |a_{ij}|$

«ανά στήλες» όταν $|a_{jj}| \geq \sum_{i=1 \,//\, i \neq j}^{n} |a_{ij}|$

θ. → Τα αυστηρά ΔΚ μητρώα είναι αντιστρέψιμα

θ. → Αν A αυστηρά ΔΚ τότε ∃ $A = LU$.

→ Σε αρ. άπειρης ακρίβειας : αν A είναι αυστηρά ΔΚ ανά στήλες τότε παραγοντοποιείται χωρίς οδήγηση

→ Συμμετρικό και ΔΚ → ΟΧΙ ΟΔΗΓΗΣΗ

Θετικά ορισμένο : $A \in \mathbb{R}^{n \times n}$, $x^T A x > 0$, $\forall\, x \neq 0 \in \mathbb{R}^n$.

* Αν $rank(A) = r$ τότε υπάρχουν $X \in \mathbb{R}^{m \times r}$, $Y \in \mathbb{R}^{r \times n}$ και αντιστρέψιμο $C \in \mathbb{R}^{r \times r}$ ώστε $A = XCY$

* Αν $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^n$ κ' $rank(A) = r$ τότε το A χαρακτηρίζεται από $mr + nr + r^2$ στοιχεία και ο πολλ/σμός Ab εκτελείται με $\Omega = 2r(n+m+r-1) - m$ πράξεις α.κ.υ.

Αριθμητική τάξη μητρώου : (r) $r = \min\{rank(B) : \|A - B\|_2 \leq \epsilon\}$

Στοιχειώδη μητρώα

↳ μητρώα της μορφής $E(u,v;\tau) = I_n - \tau u v^T$

$$\boxed{E(u,v;\tau)} = \underbrace{\begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix}}_{n\times n} - \tau \underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}}_{n\times 1} \underbrace{\begin{bmatrix} v_1 & \dots & v_n \end{bmatrix}}_{1\times n}$$

τάξη 1

→ μικρό κόστος αποθήκευσης και σε (από πράξ. μητρώου με διάνυσμα σε στοιχειώδη μητρώο)

$2n+1$ στοιχεία u, v, τ.

$E(u,v;\tau)x = (I - \tau u v^T)x = x - \tau u (\underbrace{v^T x}_{DOT})$

5 AXPY

$\Omega = 4n$.

ΙΔΙΟΤΗΤΕΣ

(1) $E(u,v;\rho)E(u,v;\tau) = (I-\rho u v^T)(I-\tau u v^T) =$

$= I - \tau u v^T - \rho u v^T + \rho u v^T \tau u v^T =$

$= I - (\rho+\tau) u v^T + \tau\rho u (v^T u) v^T =$

$= I - (\rho + \tau + \tau\rho(v^T u)) u v^T =$

$= \mathbf{E}(u, v; \rho+\tau-\tau\rho(v^T u))$

$E(u,v;\rho)E(u,v;\tau) = E(u,v;\rho+\tau-\tau\rho(v^T u))$

(2) Αν $\rho^{-1}+\tau^{-1} = v^T u \Rightarrow E(u,v;\rho)E(u,v;\tau)$ [redacted]

$= E(u, v; \rho+\tau-\tau\rho(\frac{1}{\rho}+\frac{1}{\tau})) = E(u,v;\rho+\tau-\tau-\rho) = E(u,v;0) = I_n.$

(3) $\det E(u,v;\rho) = 1 - \rho v^T u.$

(4) $\min_{u,v,\rho} \operatorname{rank}(E) = n-1 \quad (\operatorname{rank}(E) \geq n-1)$

Στοιχειώδη τριγωνικά μητρώα / ΣΜ Gauss.

u ώστε $e_j^T u = 0 \quad 1 \leq j \leq k$. $L_k(u) = E(u, e_k; 1) = I - u e_k^T$

$e_j^T u = 0 \Rightarrow u = [\underbrace{0, \dots 0}_{k \text{ μηδενικά}}, \eta_{k+1}, \dots \eta_n]^T$

$$L_k(u) = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} - \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \eta_{k+1} \\ \vdots \\ \eta_n \end{bmatrix} \begin{array}{c} \quad k \\ \begin{bmatrix} 0 & \dots & 1 & 0 & \dots & 0 \end{bmatrix} \end{array} =$$

$$= \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & \dots & 0 & \vdots & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & & & & & \\ 0 & \dots & \dots & 0 & 1 & 0 & \dots & 0 \\ \vdots & & & 0 & -\eta_{k+1} & 1 & & 0 \\ & & & 0 & -\eta_{k+1} & & \ddots & \\ & & & \vdots & \vdots & & & \\ 0 & \dots & \dots & 0 & -\eta_n & 0 & \dots & 1 \end{bmatrix}$$

$L(k+1:n,k)$. / u: Διάνυσμα Gauss!!!

$E(u, e_k; 1) E(u, e_k; -1)$

$\Rightarrow (L_k(u))^{-1} = L_k(-u)$

Στοιχειώδη αντιμεταθετικά μητρώα

$P_{ij} = E(e_i - e_j, e_i - e_j; 1) = I - (e_i - e_j)(e_i - e_j)^T$

διαφέρουν απ'το ταυτοτικό μόνο σε γραμμές κ' στήλες (i, j).

ΣΥΜΜΕΤΡΙΚΟ + ΟΡΘΟΓΩΝΙΟ

$A = A^{-1}$.

Στοιχειώδη μεταθετικά μητρώα ΓΕΝΙΚΕΥΣΗ.

$P_\sigma = [e_{\sigma(1)}, e_{\sigma(2)}, \dots, e_{\sigma(n)}]$ // Η μονάδα σα μοναδικό μη μηδενικό στοιχείο σε κάθε γραμμή κ' στήλη.

$$P_\sigma = \begin{pmatrix} e^T_{\sigma(1)} \\ e^T_{\sigma(2)} \\ \vdots \\ e^T_{\sigma(n)} \end{pmatrix} \qquad \sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ a_1 & a_2 & \dots & a_n \end{pmatrix}$$

- $PP^T = I \Rightarrow P$ ορθογώνιος
- P^T μεταθετικός.
- Αν Q μεταθετικός τότε PQ κ' QP μεταθετικοί.

$P_{ij}A \to$ αλλάζει τη γραμμή $i \leftrightarrow j$ του A

$AP_{ij} \to$ αλλάζει τη στήλη $i \leftrightarrow j$ του

Στοιχειώδη ερμιτιανά μητρώα

$$E\left(u, u; \frac{2}{u^*u}\right) = I - \frac{2}{u^*u} uu^* := H$$

$H^* = H$, $H^*H = H^2 = I$.

Απεικόνιση μητρώου $m \times n$ στη μνήμη που αποθηκεύεται σε διδιάστατο πίνακα. $m \times n$.

Το 1ο στοιχείο $\to b_0 + 1$ δ/νση μνήμης.

- ΚΑΤΑ ΣΤΗΛΕΣ : Στοιχείο στη θ. $(I, J) \to$ δ/νση $b_0 + (J-1)m + I$.
- ΚΑΤΑ ΓΡΑΜΜΕΣ : Στοιχείο στη θ. $(I, J) \to b_0 + (I-1)n + J$.

access stride/βήμα πρόσβασης $\to$ απόσταση δύο διαδοχικών δ/νσεων σταθερή.

ΣΥΣΤΗΜΑΤΑ ΜΕ ΤΡΙΓΩΝΙΚΑ ΜΗΤΡΩΑ

A: κάτω τριγωνικό κ' αντιστρέψιμο. $a_{ij} = 0, j > i$, $a_{jj} \neq 0$.

ΕΣΩΤΕΡΙΚΑ ΓΙΝΟΜΕΝΑ ΚΑΙ ΠΡΟΣΒΑΣΗ/ΓΡΑΜΜΕΣ

- $a_j^T x = b_j$ // a_j^T: γραμμές του A. $j = 1, \dots, n$.

$a_{11} x_1 = b_1$ $\qquad x_1 = \frac{b_1}{a_{11}}$

$a_{21} x_1 + a_{22} x_2 = b_2$ $\qquad x_2 = \frac{b_2 - a_{21} x_1}{a_{22}}$

$\vdots$

$a_{n1} x_1 + \dots + a_{nn-1} x_{n-1} + a_{nn} x_n = b_n$ $\qquad x_n = \frac{b_n - \sum_{k=1}^{n-1} a_{nk} x_k}{a_{nn}}$

$$x_j = \frac{b_j - \sum_{k=1}^{j-1} a_{jk} x_k}{a_{jj}}$$

A: άνω τριγωνικός.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ \vdots \\ b_n \end{bmatrix}$$

ΠΙΣΩ ΑΝΤΙΚΑΤΑΣΤΑΣΗ

$a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1$

$a_{22} x_2 + \dots + a_{2n} x_n = b_2$

$\vdots$

$a_{n-1n-1} x_{n-1} + a_{n-1n} x_n = b_{n-1}$

$a_{nn} x_n = b_n$

$x_n = \frac{b_n}{a_{nn}}$

$x_{n-1} = \frac{b_{n-1} - a_{n-1n} x_n}{a_{n-1n-1}}$

$\vdots$

$x_1 = \frac{b_1 - \sum_{j=2}^{n} a_{1j} x_j}{a_{11}}$

$$x_j = \frac{b_j - \sum_{k=j+1}^{n} a_{jk} x_k}{a_{jj}}$$

④ • Κατά στήλες πρόσβαση

[redacted] $\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n,1} \end{pmatrix} \xi_1 + \begin{pmatrix} 0 \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} \xi_2 + \dots + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ a_{nn} \end{pmatrix} \xi_n = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$

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

ΣΦΑΛΜΑ

Info: Έστω ότι υπολογίζουμε το $y = \dfrac{\left(c - \sum_{i=1}^{k-1} a_i b_i\right)}{b_k}$ με α.κ.υ. ως εξής:

```
s = c;
for i = 1:k-1
    s = s - a_i b_i
end
y = s / β_k ;
```

$\Longrightarrow \beta_k \hat{y}_k (1+\theta_k) = c - \sum_{i=1}^{k-1} a_i b_i (1+\theta_i)$

$|\theta_i| \le \gamma_i = \dfrac{iu}{1-iu}$

→ **ΘΕΩΡΗΜΑ:** Έστω τριγωνικό σύστημα $Ax=b$, $A \in \mathbb{R}^{n\times n}$ αντιστρέψιμο, λύνεται με αντικατάσταση

$(A+\Delta A)\hat{x} = b$, $|\Delta A| \le \gamma_n |A|$

↑ πίσω σταθερή!!!

→ **Αρχή της αναγωγής:** Αν $P \in \mathbb{R}^{m\times n}$ τότε $PAx = Pb$

Αν P, Q τετραγωνικά και ισχύει $PAQx = Pb$, τότε το $x = Qy$ λύνει το σύστημα

→ **ΘΕΩΡΗΜΑ:** Έστω μητρώο $A \in \mathbb{R}^{n\times n}$. Τότε υπάρχουν κάτω & άνω τριγωνικά L, U αντίστοιχα & μεταθετικό μητρώο P ώστε $LU = PA$, L αντιστρέψιμο.

→ **ΘΕΩΡΗΜΑ:** Αν τα n κύρια υπομητρώα $A(1:k,1:k)$, $k=1,\dots,n$ είναι αντιστρέψιμα τότε μπορούμε να επιλέξουμε $P=I$ ώστε $A=LU$. Αν κανονικοποιήσαμε επιλέγοντας τα στοιχεία στη διαγώνιο του L ίσα με τη μονάδα, τότε L, U, μοναδικά.

$A=LDU$: L, U στη διαγώνιο έχουν 1 και D διαγ. & αντιστρέψιμο ώστε $\det(A(1:k,1:k)) = \det(D(1:k,1:k))$.

Τότε L, D, U μοναδικοί.

* **Επίλυση με απαλοιφή Gauss:** Εφαρμόζουμε Gauss.MUL στο επαυξημένο μητρώο $[A,b]$ μέχρι απαλοιφής όλων των στοιχείων που βρίσκονται κάτω απ' την κύρια διαγώνιο. Μετά λύνουμε το σύστημα που αποτελείται από άνω τριγωνικό μητρώο & που έχει στο δεξιό μέλος μετασχηματισμένο b.
* **Επίλυση με παραγ. LU:** Εφαρμόζουμε Gauss.MUL μέχρι να λάβουμε τους παράγοντες L, U & συνεχίζουμε την επίλυση 2 τριγωνικών συστημάτων $Lc=b$, $Ux=c$.

ΠΑΡΑΓΟΝΤΟΠΟΙΗΣΗ ΜΕ ΟΜΑΡΘΟΥΣ

→ πράξεις τύπου BLAS-3. Αξιοποιούν ιεραρχία.

$\begin{pmatrix} A_{11} & A_{12} & A_{22} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & U_{22} & U_{23} \\ 0 & 0 & U_{33} \end{pmatrix}$

$A_{12} = L_{11} U_{12}$ → τ_{22}

$A_{22} = L_{22} U_{22} + L_{21} U_{12}$

$A_{32} = L_{31} U_{12} + L_{32} U_{22}$

$A_{11} \in \mathbb{R}^{(k-1)\beta \times (k-1)\beta}$

$A_{22} \in \mathbb{R}^{\beta\times\beta}$

$A_{33} \in \mathbb{R}^{(n-k\beta)\times(n-k\beta)}$

Αλγόριθμος από αριστερά και ανά στήλες (left-looking)

- Σε κάθε βήμα υπολογίζουμε β στήλες του L & του U.
- Στην αρχή βήματος k έχουμε υπολογίσει $(k-1)\beta$ στήλες του L & του U. → L_{11}, L_{21}, L_{31}

Επόμενο βήμα: $L_{22}, L_{32}, U_{12}, U_{22}$.

* Προς τα δεξιά ανά στήλες β γραμμές

- Σε κάθε βήμα, β διαδοχικές στήλες του **L** ή β διαδοχ. γραμμές του **U**

Γνωστά L_{11}, L_{21}, L_{31}, U_{11}, U_{12}, U_{13} $\longrightarrow$

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & I & 0 \\ L_{31} & 0 & I \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ 0 & \tilde{A}_{22} & \tilde{A}_{23} \\ 0 & \tilde{A}_{32} & \tilde{A}_{33} \end{pmatrix}$$

κ βήμα $\longrightarrow$

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} = \begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{23} & I \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} & U_{13} \\ & U_{22} & U_{23} \\ & 0 & \tilde{A}_{33} \end{pmatrix}$$

(1)

- $A_{22} = L_{21}U_{12} + L_{22}U_{22}$
  $A_{32} = L_{31}U_{12} + L_{32}U_{22}$

(2) $U_{23} = L_{22}^{-1}\tilde{A}_{23}$

(3) $\hat{A}_{33} = \tilde{A}_{33} - L_{32}U_{23}$

* ΕΞΑΡΤΑΤΑΙ ΑΠΟ ΕΠΙΛΟΓΗ Κ
  $\Longrightarrow$ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ !!!

* Τριγωνική Παραγοντοποίηση κατά ομάδες

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ L_{21} & I \end{pmatrix} \begin{pmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{pmatrix}$$

$L_{21} := A_{21}A_{11}^{-1}$
$U_{11} := A_{11}$
$U_{12} := A_{12}$
$U_{22} := A_{22} - A_{21}A_{11}^{-1}A_{12}$

* **ΜΕΡΙΚΗ ΟΔΗΓΗΣΗ**: στο βήμα κ διαλέγουμε για οδηγό, το μέγιστο στοιχείο που είναι στις θέσεις (κ:n, Αν το στοιχείο στη θέση $(\hat{k}, k)$, $\hat{k} \geq k$, τότε μεταθέτουμε τις γραμμές.
  $\Longrightarrow$ ΠΟΛΛΑΠΛΑΣΙΑΣΤΕΣ ΜΙΚΡΟΤΕΡΟΙ ΤΗΣ ΜΟΝΑΔΑΣ ΣΕ ΑΠΟΛΥΤΗ ΤΙΜΗ!

* **ΠΛΗΡΗΣ ΟΔΗΓΗΣΗ**: Στο βήμα κ διαλέγουμε για οδηγό το μέγιστο, σε απόλυτη τιμή στοιχείο ανάμεσα στις θέσεις $A_{k:n,k:n}$. Αν το στοιχείο είναι στη θέση (k_1, k_2) τότε μεταθέτουμε τις γραμμές k, k_1 και τις στήλες k, k_2.

$\boxed{PA = LU}$

L

$|l_{ij}| \leq 1$

$PAx = LUx = Pb$: Αρκεί να μεταθέσουμε τα στοιχεία του b και να λύσουμε ως προς L, U.

$Lz = Pb$
$Ux = z$

* Αειχ και ότι υπολογίζουμε ... με απαλοιφή Gauss με μερική οδήγηση.
Τότε $(A+\Delta A)\hat{x} = b$, $\|\Delta A\|_\infty \leq c(n^3)\rho_n^{\text{μο}} \|A\|_\infty u$

Η ΠΙΣΩ ΕΥΣΤΑΘΕΙΑ ΤΗΣ ΑΠΑΛΟΙΦΗΣ GAUSS ΜΕ ΜΕΡΙΚΗ ΟΔΗΓΗΣΗ ΕΞΑΡΤΑΤΑΙ ΑΠΟ ΤΟ ΜΕΓΕΘΟΣ ΤΟΥ ΣΥΝΤΕΛΕΣΤΗ ρ_n.

Για αντιστρ. μητρώα:

Λύση $Ax=b$.
① $A^TAx = A^Tb$.
② $x = (A^TA)^{-1}A^Tb$.

ΑΝ Α ΑΝΤΙΣΤΡΕΨΙΜΟ
A^TA συμμετρικό θετικά ορισμένο. (ΣΘΟ)

* Ένα συμμετρικό μητρώο είναι και ΣΘΟ $\Leftrightarrow$ ιδιοτιμές θετικές
* Αν οι στήλες του A γρ. ανεξάρτητες ($m \geq n$ και), $B = A^TA$ ΣΘΟ.
* Αν A συμμετρικό και δ.κ., αν τα διαγώνια στοιχεία όλα θετικά $\Rightarrow$ A ΣΘΟ.

• Αν A ΣΘΟ $\Rightarrow$ A αντιστρέψιμο
κάθε κύριο υπομητρώο ΣΘΟ $\Rightarrow$ αντιστρ. $\Rightarrow \exists$ L, U ώστε $A=LU$
όλα στοιχεία διαγωνίου θετικά
συμπλήρωμα Schur ΣΘΟ...

ΘΕΩΡΗΜΑ 1: Έστω $A \in \mathbb{R}^{n\times n}$ ΣΘΟ και έστω $A^{(k)} = L_k \dots L_1 A$ $(1 \leq k \leq n-1)$ τότε:

① Τα $A^{(k)}_{k:n,k:n}$ είναι ΣΘΟ
② Αν $a_{ij}^{(k)}$ στοιχεία του $A^{(k)}_{k:n,k:n}$ τότε $\max\limits_{k\leq i,j\leq n} |a_{ij}^{(k)}| \leq \max\limits_{k \leq i,j\leq n} |a_{ij}^{(k-1)}|$ $k=1,2,\dots,n-1$.
$\Downarrow$
αν εφαρμόσουμε LU χωρίς οδήγηση $\rho=1$.

ΘΕΩΡΗΜΑ 2: Αν όλα τα αρχικά κύρια υπομητρώα του $A \in \mathbb{R}^{n\times n}$ είναι αντιστρέψιμα τότε υπάρχουν μητρώα L, M, τα οποία είναι κάτω τριγωνικά με 1 στη διαγώνιο και D με θετικά στοιχεία, τέτοια ώστε $A = LDM^T$. L, D, M μοναδικά.

ΘΕΩΡΗΜΑ 3: Έστω το συμμετρικό θ.ο. μητρώο $A \in \mathbb{R}^{n\times n}$. Τότε υπάρχει μοναδικό άνω τριγωνικό μητρώο R με θετικά διαγώνια στοιχεία ώστε $A = R^TR$.

Τάξη$^{\text{Chol.}}$ $\approx \frac{n^3}{3}$ (μισό της LU).

Αλγ. Cholesky για ΣΘΟ πίσω σταθερός / χωρίς οδήγηση
$\|\Delta A\|_\infty \leq 3\gamma_n \||L|\|_\infty \||L^T|\|_\infty \leq 3\gamma_n \|A\|_\infty$

• $A = LDL^T$
D: διαγώνιο, L: κάτω τριγωνικό με 1 στη διαγ.
↑ ΓΙΑ ΕΥΣΤΑΘΕΙΑ ΧΡΕΙΑΖ. ΟΔΗΓΗΣΗ!..

Συντελεστής αύξησης: $\rho_n := \dfrac{\max\{a_0, a_1, \dots, a_{n-1}\}}{a_0}$

$a_k := \max_{i,j} |a_{ij}^{(k)}|$ τα στοιχεία με μέγιστη απόλυτη τιμή στις θ. $(k+1:n, k+1:n)$ του υπομητρώου $L_k P_k L_{k-1} P_{k-1} \dots L_1 P_1 A$.

ΟΣΟ ΜΙΚΡΟΤΕΡΟΣ ΤΟΣΟ ΤΟ ΚΑΛΥΤΕΡΟ!!!

ΕΞΑΡΤΑΤΑΙ ΑΠ'ΤΗ ΔΟΜΗ ΤΟΥ ΜΗΤΡΩΟΥ $\rightarrow \rho_n^{MO} \leq n$ (Hessenberg)

$\rightarrow \rho_n^{MO} \leq 2$ (τριδιαγώνια).

ΣΦΑΛΜΑ ΜΕΘΟΔΩΝ ΕΠΙΛΥΣΗΣ

$(A+\Delta A)(x+\Delta x) = b + \Delta b$ $\qquad A(\epsilon) = A + \epsilon E$, $A := A(0)$.

Ορισμός: Έστω βαθμωτός ϵ και μητρώο $A(\epsilon) \in \mathbb{R}^{m \times n}$ με στοιχεία της μορφής $a_{ij}(\epsilon)$.

Αν οι συναρτήσεις $a_{ij}(\epsilon)$ παραγωγίσιμες ως προς ϵ τότε $A^{(1)}(\epsilon)$ μητρώο με στοιχεία $a^{(1)}_{ij}(\epsilon) = \frac{d}{d\epsilon}$

$\left.\begin{aligned} A(\epsilon) &\in \mathbb{R}^{m\times n} \\ B(\epsilon) &\in \mathbb{R}^{n\times s}\end{aligned}\right| \Rightarrow$ $\boxed{\begin{aligned} &\frac{d}{d\epsilon}[A(\epsilon)B(\epsilon)] = \frac{d}{d\epsilon}[A(\epsilon)]B(\epsilon) + A(\epsilon)\frac{d}{d\epsilon}[B(\epsilon)] \\ &\frac{d}{d\epsilon}[A(\epsilon)^{-1}] = -A(\epsilon)^{-1}\frac{d}{d\epsilon}[A(\epsilon)]A(\epsilon)^{-1}\end{aligned}}$ ΠΡΟΣΟΧΗ ΣΕ ΣΕΙΡΑ!!!

Έστω $(A+\epsilon E)x(\epsilon) = b + \epsilon e$, $x := x(0)$.

$$\frac{d}{d\epsilon}\left[(A+\epsilon E)x(\epsilon)\right] = \frac{d}{d\epsilon}(b + \epsilon e) \Rightarrow$$

$$\frac{d}{d\epsilon}[A+\epsilon E]x(\epsilon) + (A+\epsilon E)\frac{d}{d\epsilon}x(\epsilon) = e \Rightarrow$$

$$E x(\epsilon) + \text{[redacted]} (A+\epsilon E)x^{(1)}(\epsilon) = e \Rightarrow$$

για $\epsilon = 0 \rightsquigarrow E x(0) + A x^{(1)}(0) = e \Rightarrow$

$$A x^{(1)}(0) = e - Ex \Rightarrow$$

$$x^{(1)}(0) = A^{-1}(e - Ex).$$

Από σειρά Taylor: $x(\epsilon) = x(0) + \epsilon x^{(1)}(0) + O(\epsilon^2)$

$$\frac{\|x(\epsilon) - x\|}{\|x\|} \leq |\epsilon| \|A^{-1}\| \left[\frac{\|e\|}{\|b\|} + \|E\|\right]$$

$$\Rightarrow \leq k(A)(\rho_A + \rho_B) + O(\epsilon^2)$$

$\left.\begin{aligned}\rho_A &= |\epsilon|\frac{\|E\|}{\|A\|} \\ \rho_B &= |\epsilon|\frac{\|e\|}{\|b\|}\end{aligned}\right\}$ Μετρούνε σχετικό σφάλμα στα στοιχεία εισόδου.

$$\frac{1}{k_p(A)} = \min_{\substack{A+\Delta A \\ \text{αντιστρ.}}} \frac{\|\Delta A\|_p}{\|A\|_p} \qquad k_2(A) = \frac{\sigma_{max}}{\sigma_{min}} = \frac{\sigma_1}{\sigma_n}$$

σχετική απόσταση στην νόρμα p του A από τους μη αντιστρ. πίνακες.

* Αν $\|A\| < 1$, τότε το $I - A$ αντιστρέψιμο και $\|(I-A)^{-1}\| \leq \dfrac{1}{1 - \|A\|}$
* Αν A αντιστρέψιμο και $\|A^{-1}E\| < 1$ τότε $A+E$ είναι αντιστρέψιμο.
* ΘΕΩΡΗΜΑ: Αν $\|A^{-1}\| \|\Delta A\| < 1$ & $(A+\Delta A)y = b + \Delta b$ τότε το $A + \Delta A$ αντιστρέψιμο και $\dfrac{\|y - x\|}{\|x\|} \leq \dfrac{k(A)}{1 - k(A)\frac{\|\Delta A\|}{\|A\|}}\left(\dfrac{\|\Delta A\|}{\|A\|} + \dfrac{\|\Delta b\|}{\|b\|}\right)$

Εφόσον $\|A^{-1}\Delta A\| \leq \|A^{-1}\| \|\Delta A\| < 1$, $I + A^{-1}\Delta A$ [redacted] αντιστρέψιμο και $\|I + A^{-1}\Delta A\| \leq \dfrac{1}{1 - \|A^{-1}\| \|\Delta A\|}$ ...

* $\dfrac{\|y - x\|}{\|x\|} \leq \dfrac{2\epsilon k(A)}{1 - \epsilon k(A)}$

→ Αν $\epsilon k(A) < \frac{1}{2}$ → $\dfrac{\|y - x\|}{\|x\|} \leq 4\epsilon k(A)$

$\boxed{\|r\|_2 = \|b - Ax\|_2 \to \min}$ ΠΡΟΒΛΗΜΑ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Έστω $A \in \mathbb{R}^{m\times n}$ ... $Q \in \mathbb{R}^{m\times n}$ με ορθοκανονικές στήλες και άνω τριγωνικός $R \in \mathbb{R}^{n\times n}$ ώστε $A = QR$ ... τότε ο R ...

Αν A ... ο R ... θετικά ...

... QR ...

* Τα διανύσματα $q_1, \dots, q_n$ είναι ορθοκανονικά όταν $\boxed{q_i^T q_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases}}$

$\downarrow$

στήλες του Q

* Έστω διανυσματικός χώρος $S \subset \mathbb{R}^n$. Τότε τα διανύσματα που είναι ορθογώνια στο S αποτελούν το ορθογώνιο συμπλήρωμα του S.
* $P: \mathbb{R}^n \to \mathbb{R}^n$ ώστε $P^2 = P$ ← τελεστής προβολής!
* Αν P τελεστής προβολής → $(I - P)$ τελεστής προβολής

- S υπόχωρος του $\mathbb{R}^n$, P τελεστής ορθογώνιας προβολής

(1) Αν $x \in S$, $Px = x$

(2) $x = \underbrace{Px}_{S} + \underbrace{(I-P)x}_{T}$

(3) $P^2 = P$

(4) $P^T = P$

... Αν οι διανυσματικοί χώροι $S, T \subset V$ είναι συμπληρωματικοί, τότε κάθε διάνυσμα $x \in V$ είναι άθροισμα των μοναδικών συνιστωσών του $x = s + t$, $s \in S$, $t \in T$

Έστω μη μηδενικό $u \in \mathbb{R}^m$ και $\boxed{P = \frac{uu^T}{u^T u}} \in \mathbb{R}^{m\times m}$

ΟΡΘΟΚΑΝΟΝΙΚΟΠΟΙΗΣΗ → ορθοκανονική βάση $[q_1, q_2, \dots, q_n]$

$A = [a_1 \ \dots \ a_n]$

① $q_1 = \frac{a_1}{\|a_1\|}$

② $\tilde{q}_2 = a_2 - (q_1^T a_2) q_1 \qquad q_2 = \frac{\tilde{q}_2}{\|\tilde{q}_2\|}$

③ $\tilde{q}_3 = a_3 - (q_2^T a_3) q_2 - (q_1^T a_3) q_1 \qquad q_3 = \frac{\tilde{q}_3}{\|\tilde{q}_3\|}$

k-ιο: $\tilde{q}_k = a_k - (q_{k-1}^T a_k) q_{k-1} - \dots - (q_1^T a_k) q_1 \qquad q_k = \frac{\tilde{q}_k}{\|\tilde{q}_k\|}$

$H = H^{-1}$

$u^T \quad H^2 = I$

$Hx = \left(I - \frac{2}{u^T u} u u^T\right) x$

$Hx \in \langle e_1 \rangle$
$x \in \langle e_1 \rangle$

$u = x + a e_1$

$\bullet\ u = x + \text{sign}(\xi_1)\|x\|_2 e_1$: ΤΟ ΠΡΟΣΗΜΟ ΕΠΙΛΕΓΕΤΑΙ ΓΙΑ ΝΑ ΑΠΟΦΕΥΓΕΤΑΙ ΚΑΤΑΣΤΡΟΦΙΚΗ ΑΠΑΛΟΙΦΗ.

$\rightarrow HA = \left(I - \frac{2uu^T}{u^T u}\right)A = A - \frac{2}{u^T u} u(u^T A) = A - \frac{2}{u^T u} u(A^T u)^T$

$\rightarrow AH = A\left(I - \frac{2uu^T}{u^T u}\right) = A - \frac{2}{u^T u}(Au)u^T.$

* QR : ΜΕΘΟΔΟΣ HOUSEHOLDER *

$H = H^{-1}$ ($H^T H = I$ και $H = H^T$, $H^2 = I$)

$H_n H_{n-1} \dots H_1 A = R \Rightarrow A = (H_n H_{n-1} \dots H_1)^{-1} R \Rightarrow$

$A = H_1^{-1} \dots H_{n-1}^{-1} H_n^{-1} R \Rightarrow$

$A = \underbrace{H_1 \dots H_{n-1} H_n}_{Q} R = QR$

$*\ A = Q_1 R_1$

$A^T A = (Q_1 R_1)^T (Q_1 R_1) = R_1^T Q_1^T Q_1 R_1 = R_1^T R_1$

$\bullet\ (R^T A^T) A = R$

*** Επίλυση $\min_x \|Ax-b\|_2$ με $A=QR$ ***

$$A=QR \Rightarrow Q^TA=R=\begin{pmatrix} R_1 \\ 0 \end{pmatrix} \to n\times n$$

$$Q^Tb=\begin{pmatrix} c \\ d \end{pmatrix}$$

$$\|Ax-b\|_2^2=\|Q^TAx-Q^Tb\|_2^2=\|R_1x-c\|_2^2+\|d\|_2^2$$

Αφού A αντιστρ. $\Rightarrow$ R αντιστρ. $\Rightarrow$ $x=R_1^{-1}c$ $\Rightarrow$ $\min_{x\in\mathbb{R}^n}\|Ax-b\|_2=\|d\|_2$.

ΘΕΩΡΗΜΑ: Αν $A\in\mathbb{R}^{m\times n}$, $b\in\mathbb{R}^m$ και $A^T(b-Ax)=0$, τότε ισχύει $\|b-Ax\|_2\le\|b-Ay\|_2$ $\forall y\in\mathbb{R}^n$

$Ax \to$ ορθογ. προβολή του b στο Range(A).

$Pb=Ax$.

$b-Pb \perp \text{span}\{a_1,\dots,a_n\}$

$A^T(b-Pb)=0$

$*\; P=A(A^TA)^{-1}A^T$: τελεστής ορθογ. προβολής επί του υποχώρου που παράγεται από τις στήλες του A.

$Ax=Pb \Rightarrow$

$Ax=A(A^TA)^{-1}A^Tb \Rightarrow$

$x=A^{-1}A(A^TA)^{-1}A^Tb \Rightarrow$

$x=(A^TA)^{-1}A^Tb. \Rightarrow$

Αφού A γρ. ανεξάρτητες στήλες $\Rightarrow A^TA$ σ.θ.ο.κ. αντιστρέψιμο.

$$\boxed{A^TAx=A^Tb} \leftarrow$$ «ΣΥΣΤΗΜΑ ΚΑΝΟΝΙΚΩΝ ΕΞΙΣΩΣΕΩΝ»

$x=\xi_1+i\xi_2=|x|e^{i\phi}$

$xe^{i\phi}=(\xi_1+i\xi_2)(\cos\phi+i\sin\phi)=$

$=\xi_1\cos\phi+i\xi_1\sin\phi+i\xi_2\cos\phi+i^2\xi_2\sin\phi=$

$=\xi_1\cos\phi+i\xi_1\sin\phi+i\xi_2\cos\phi-\xi_2\sin\phi=$

$=(\xi_1\cos\phi-\xi_2\sin\phi)+i(\xi_1\sin\phi+\xi_2\cos\phi)=$

$=(c\xi_1-s\xi_2)+i(\xi_1 s+\xi_2 c)$

$$G=\begin{pmatrix} c & -s \\ s & c \end{pmatrix}$$

Αν $c=\dfrac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}}$, $s=\dfrac{-\xi_2}{\sqrt{\xi_1^2+\xi_2^2}}$

$$\begin{pmatrix} \frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}} & \frac{\xi_2}{\sqrt{\xi_1^2+\xi_2^2}} \\ \frac{-\xi_2}{\sqrt{\xi_1^2+\xi_2^2}} & \frac{\xi_1}{\sqrt{\xi_1^2+\xi_2^2}} \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix}=$$

$$\begin{pmatrix} \frac{\xi_1^2}{\sqrt{\xi_1^2+\xi_2^2}}+\frac{\xi_2^2}{\sqrt{\xi_1^2+\xi_2^2}}=\sqrt{\xi_1^2+\xi_2^2} \\ 0 \end{pmatrix}$$

*** Αναγωγή σε μορφή άνω Hessenberg
↳ του A

$Q^T A Q = H$.

$Q^T Q = I$ Q: γινόμενο μετασχ. Hauseholer

Στο βήμα κ υπολογίζεται ο ανακλαστής H_k που μηδενίζει τα στοιχεία $(k+2:n, \cdot)$
του $A^{(k)}$ $\longrightarrow A^{(n-2)}$ άνω Hessenberg.

$$A^{(k)} = \underbrace{H_{k-1} \cdots H_1}_{Q} A \underbrace{H_1 \cdots H_{k-1}}_{Q^T}$$

$$A^{(n-2)} = H_{n-3} \cdots H_2 A H_1 \cdots H_{n-3}.$$

- Σφάλμα sAXPY

$z \leftsquigarrow y + ax;$

$y \leftsquigarrow z$

$$fl(\zeta_i) = fl(\eta_i + fl(a\xi_i)) =$$

$$= (\eta_i + a\xi_i(1+\delta_{1,i}))(1+\delta_{2,i}) =$$

$$= \eta_i(1+\delta_{2,i}) + a\xi_i(1+\delta_{1,u})(1+\delta_{2,i}) =$$

$$= \eta_i(1+\delta_{2,i}) + a\xi_i(1+\theta_{2,i})$$

$|\delta_{i,j}| \leq u$

$$|\theta_{2,i}| \leq \gamma_2 = \frac{2u}{1-2u} \leq 2u + O(u^2)$$

$$|\zeta_i - \tilde{\zeta}_i| = |\eta_i + a\xi_i - \eta_i(1+\delta_{2,i}) - a\xi_i(1+\theta_{2,i})| =$$

$$= |\eta_i(\cancel{1}-\cancel{1}-\delta_{2,i}) + a\xi_i(\cancel{1}-\cancel{1}-\theta_{2,i})| =$$

$$= |\eta_i\delta_{2,i} + a\xi_i\theta_{2,i}|$$

$$\leq |\eta_i||\delta_{2,i}| + |a\xi_i||\theta_{2,i}|$$

$$\leq |\eta_i|u + \gamma_2|a\xi_i| \leq |\eta_i|\cdot u + 2u|a\xi_i| + O(u^2).$$

$$|z - \tilde{z}| \leq u(|y| + 2|a||x|) + O(u^2).$$

- Βρίσκουμε πίσω σφάλμα - cond(f_{prog}) - κατάσταση αλγορίθμου
- Βρίσκουμε δείκτη κατάστασης του προβλήματος.

$$fl(\zeta_i) = \underbrace{\eta_i(1+\delta_{2,i})}_{} + a\underbrace{\xi_i(1+\theta_{2,i})}_{},$$

Άρα, το υπολογισμένο $\tilde{z}$ μπορεί να θεωρηθεί σαν το ακριβές sAXPY

$\tilde{z} = \tilde{y} + a\tilde{x}$, όπου $\tilde{y} = y + \Delta y$

και $\tilde{x} = x + \Theta x$ όπου $\Delta, \Theta \rightarrow$

Πίνακες

$$\Delta = \begin{pmatrix} \delta_{2,1} & 0 & \dots & 0 \\ 0 & \delta_{2,2} & 0 \dots & 0 \\ & & \ddots & 0 \\ & \bigcirc & & \delta_{2,n} \end{pmatrix} \quad \text{και} \quad \Theta = \begin{pmatrix} \vartheta_{2,1} & 0 & \dots & 0 \\ 0 & \vartheta_{2,2} & 0 \dots & 0 \\ & & \ddots & 0 \\ & \bigcirc & & \vartheta_{2,n} \end{pmatrix}$$

$\|\Delta\| \le u$

$\|\Theta\| \le \gamma_2$. για οποιεσδήποτε γνωστές νόρμες.

$\tilde{z} = y + \Delta y + a(x + \Theta x)$

$= z + \Delta y + a\Theta x$

$\|\tilde{z} - z\| \le \|\Delta\|\|y\| + |a|\|$ ...

• Θεωρούμε συνάρτηση $f : \mathbb{R}^{2n+1} \longrightarrow \mathbb{R}^n$,

$$f([x;y;a]) = y + ax.$$

$$f([x+\Theta x; y+\Delta y; a]) = f_{prog}([x;y;a])$$

$X = [x;y;a]$

$\tilde{X} = [\tilde{x};\tilde{y};a] = [x+\Theta x; y+\Delta y; a]$

[redacted]

$\|\tilde{X} - X\| \le \|[\Theta x; \Delta y; 0]\|$

$= \left\| \begin{pmatrix} \Theta & & \\ & \Delta & \\ & & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ a \end{pmatrix} \right\| \le$

$\left\| \begin{pmatrix} \Theta & & \\ & \Delta & \\ & & 0 \end{pmatrix} \right\| \cdot \|X\|$

$$\frac{\|\tilde{X} - X\|}{\|X\|} \le \left\| \begin{pmatrix} \Theta & & \\ & \Delta & \\ & & 0 \end{pmatrix} \right\| < 2{,}04u.$$

Επομένως, μπορούμε να πούμε ότι $\text{cond}(f_{prg}) < 2{,}04$

✱ Θα βρούμε το δείκτη κατάστασης του προβλήματος

$$\text{cond}(f;x) = \frac{\|f'(x)\|}{\|z\|}\|X\|$$

$f'(X)$: Ιακωβιανό Μητρώο.

$f'([x;y;a]) = [aI; I, x] \in \mathbb{R}^{n \times 2n+1}$. $\quad f : \mathbb{R}^{2n+1} \to \mathbb{R}^n \Rightarrow f' \in \mathbb{R}^{n \times (2n+1)}$

$\|f'(x)\|_\infty = \max_i \{ 1 + |a| + |\xi_i| \} = 1 + |a| + \|x\|_\infty$

$\|X\|_\infty = \max\{\|x\|_\infty, \|y\|_\infty, |a|\}$

• Το σφάλμα σχετικό σφάλμα μεγάλο αν $\|z\|$ μικρή.

• Σφάλμα εξωτερικού γινομένου

$C = ab^T \quad , a, b \in \mathbb{R}^n$

$fl(\gamma_{ij}) = a_i b_j (1+\delta_{ij}) \Longrightarrow fl(\gamma_{ij}) = a_i b_j + a_i b_j \delta_{ij}$

$fl(C) = ab^T + E$ όπου

$e_{ij} = a_i b_j \delta_{ij}$

→ Αν είχαμε πίσω ευστάθεια θα υπήρχαν $\tilde{a}, \tilde{b}$ κοντά στα a, b ώστε $fl(C) = \tilde{a}\tilde{b}^T = (a+\Delta a)(b+\Delta b)^T$

Τότε θα ίσχυε $E = \underbrace{a\Delta b^T + \Delta a \cdot b^T + \Delta a \cdot \Delta b^T}_{\text{τάξη} \leq 2} = [a \quad \Delta a]\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}[$

Ισχύει όμως: $\boxed{rank(XY) \leq min(rank(X), rank(Y))}$

$\boxed{rank(X+Y) \leq rank(X) + rank(Y)}$

$\Longrightarrow rank(E) \leq rank\left[\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}\right] = 2.$

ΑΛΛΑ $rank(E) = n \Longrightarrow$ ΑΤΟΠΟ.

$\Longrightarrow$ Η ανανέωση $1^{ης}$ τάξης δεν είναι πίσω ευσταθής

ΕΜΠΡΟΣ ΣΦΑΛΜΑ: $|fl(\gamma_{ij}) - a_i b_j| = |a_i b_j \delta_{ij}| \leq |a_i b_j|$

$$\frac{|fl(\gamma_{ij}) - a_i b_j|}{|a_i b_j|} \leq u.$$

- Σφάλμα Πολ/σμού **MV** (μητρώου x διανύσματα)

$y = Ax$

- Κατά γραμμές : $\eta_i = a_{i,:}^T x \quad , i=1:m.$
- Κατά στήλες : $\eta = \sum_{i=1}^{n} \xi_i a_{:,i}$

$\longrightarrow \tilde{\eta}_i = (a_{i,:} + \Delta a_{i,:})^T x \quad , |\Delta a_{i,:}| \leq \gamma_n |a_{i,:}|$

<u>...ίσω σφάλμα</u> $\tilde{y} = (A + \Delta A) x \qquad |\Delta A| \leq \gamma_n |A|$

<u>...εμπρός σφάλμα</u> : $|y - \tilde{y}| \leq \gamma_n |A| \cdot |x|$

- Σφάλμα πολ/σμού MM (μητρώο x μητρώο)

$C = AB \Longrightarrow c_i = A b_i \quad , i=1:n$

$\tilde{c}_i = (A + \Delta A_i) b_i \quad , |\Delta A_i| \leq \gamma_n |A|$

$fl(AB) = (A + \Delta A) B \quad , |\Delta A| \leq \gamma_n |A| \cdot |B| |B^{-1}|$

$|C - \tilde{C}| = |AB - (A + \Delta A) B| = |\Delta A \cdot B| \leq |\Delta A| \cdot |B|$

$\leq \gamma_n |A| |B| |B^{-1}| |B|$

$= \gamma_n |A| |B|.$

$\Rightarrow$ Μεγάλο σχετικό σφάλμα αν $|AB| \ll |A| \cdot |B|$

• Αναδρομική άθροιση

```
S=0;
for i=1:n
    S=S+ξi
end
```

<u>Πίσω σφάλμα</u>: $fl(S) = ((\xi_1+\xi_2)(1+\delta_1)+\xi_3)(1+\delta_2)+\dots+\xi_n)(1+\delta_{n-1}) \Rightarrow$

$$fl(S) = \xi_1(1+\theta_{n-1})+\xi_2(1+\hat{\theta}_{n-1})+\xi_3(1+\theta_{n-2})+\dots+\xi_n(1+\theta_1)$$

$$|\theta_j| \le \gamma_j = \frac{j\cdot u}{1-j\cdot u}$$

$\Rightarrow$ Το πίσω σφάλμα εξαρτάται από το n.

• Ενθετική άθροιση

1. Ταξινόμηση του S κατ' απόλυτη τιμή σε αύξουσα σειρά
   $L := \xi_1 \le \xi_2 \le \xi_3 \le \dots \le \xi_n$
2. Διαγραφή του $\xi_1+\xi_2$ από την L και ένθεση στην L στην κατάλληλη θέση ώστε να παραμείνει μονοτονική. Επαναλ. του L αν περιέχει 2 ή περισ. στοιχεία.

ΑΝΑΛΥΣΗ ΣΦΑΛΜΑΤΟΣ ΓΙΑ DOT

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$S_n = x^T y$.

1º Βήμα : $\tilde{S}_1 = fl(x_1 y_1) = x_1 y_1 (1+\delta_1)$

$|\delta_i| \le u$

$$\tilde{S}_2 = fl(\tilde{S}_1 + fl(x_2 y_2)) = \left(x_1 y_1 (1+\delta_1) + x_2 y_2 (1+\delta_2)\right)(1+\delta_3)$$

$$\Rightarrow \tilde{S}_2 = x_1 y_1 (1+\theta_2) + x_2 y_2 (1+\hat{\theta}_2) \Rightarrow$$

$|\theta_2|, |\hat{\theta}_2| \le \gamma_2$.

$$\tilde{S}_3 = \left((x_1 y_1 (1+\delta_1) + x_2 y_2 (1+\delta_2))(1+\delta_3) + x_3 y_3 (1+\delta_4)\right)(1+\delta_5)$$

$$\tilde{S}_3 = x_1 y_1 (1+\theta_3) + x_2 y_2 (1+\hat{\theta}_3) + x_3 y_3 (1+\theta_2)$$

$\vdots$

$$\Rightarrow \tilde{S}_n = x_1 y_1 \prod_{j=1}^{n+1} (1+\delta_j) + x_2 y_2 \prod_{j=2}^{n+1} (1+\delta_j) + x_3 y_3 \prod_{j=3}^{n+1} (1+\delta_j) + \dots + x_n y_n \prod_{j=n}^{n+1} (1+\delta_j) \Rightarrow$$

$$\tilde{S}_n = x_1 y_1 (1+\theta_n) + x_2 y_2 (1+\hat{\theta}_n) + x_3 y_3 (1+\theta_{n-1}) + \dots + x_n y_n (1+\theta_2)$$

$$|\tilde{S}_n - S_n| \le |x_1 y_1 (1+\theta_n) - x_1 y_1 + x_2 y_2 (1+\hat{\theta}_n) - x_2 y_2 + \dots| =$$

$$= |x_1 y_1 \theta_n + x_2 y_2 \hat{\theta}_n + \dots| \le \gamma_n |x|^T |y|$$

- Αν για όλα τα στοιχεία $x_i y_i \ge 0$ $\quad |x|^T |y| = |x^T y| = |S_n|$

$$\frac{|\tilde{S}_n - S_n|}{|S_n|} \le \gamma_n.$$

ΘΕΤΟΥΜΕ $cond(f_{prob}) = \frac{\gamma_n}{u}$.

$\Rightarrow$ ΥΠΟΛΟΓΙΣΜΟΣ ΕΣ. ΓΙΝΟΜΕΝΟΥ ΠΙΣΩ ΕΥΣΤΑΘΗΣ!!!

→ **Κατάσταση υπολογισμού του DOT**

$$f([x;y]) = x^T y \quad , x, y \in \mathbb{R}^n.$$

$$X := [x;y] \in \mathbb{R}^{2n}$$

$$\mathrm{cond}(f;X) = \frac{\|f'(X)\|}{\|f([x;y])\|} \cdot \|X\| = \left\| \frac{\partial f}{\partial X}\Big|_{[x;y]} \right\| \cdot \frac{\|X\|}{\|f([x;y])\|}$$

$\frac{\partial f}{\partial X}\Big|_{[x;y]}$: Ιακωβιανό μητρώο της f.

$$f : \mathbb{R}^{2n} \to \mathbb{R} \quad \rightsquigarrow \quad \frac{\partial f}{\partial X} \in \mathbb{R}^{1 \times 2n}$$

$$\frac{\partial f}{\partial X}\Big|_{[x;y]} = [y;x]^T$$

$$\Rightarrow \mathrm{cond}(f;X) = \frac{\|[x;y]\|}{\|x^T y\|} \|[y;x]^T\| \Rightarrow$$

$$\mathrm{cond}(f;X) = \frac{\|[x;y]\| \, \|[y;x]^T\|}{|x^T y|} \Rightarrow$$

$$\mathrm{cond}(f;X) = \frac{\|[x;y]\|^2}{|x^T y|}$$

Εμπρός σφάλμα με προς τα εμπρός ανάλυση

$$|fl(x^T y) - x^T y| = |x_1 y_1 \theta_n + x_2 y_2 \hat{\theta}_n + \ldots + x_n y_n \theta_2| \leq \gamma_n \sum_{i=1}^{n} |x_i| \cdot |y_i| = \gamma_n |x|^T |y|$$

$$\leq n u |x|^T |y| + O(u^2)$$

- Αν $|x^T y| \ll |x|^T |y|$ άνω φράγμα για σχ. σφάλμα μπορεί να είναι μεγάλο επομένως καμία εξασφάλιση όσον αφορά το μέγεθος σχετικού σφάλματος.
  Αν $x_i y_i \geq 0 \ \forall i$ τότε το σχετικό σφάλμα μικρό.

ΣΘΟ ΜΗΤΡΩΑ

* Ένα συμμετρικό μητρώο είναι και ΣΘΟ αν όλες οι ιδιοτιμές του είναι θετικές.
* $B:=A^TA$. Αν οι στήλη του A γραμμικά ανεξάρτητες τότε το B είναι ΣΘΟ. (όταν $m \geq n$).
* A συμμετρικό και δ.κ. $\Rightarrow$ Αν τα διαγώνια στοιχεία θετικά τότε το A είναι ΣΘΟ.

Ιδιότητες ΣΘΟ μητρώων

Αν A ΣΘΟ:
- A αντιστρέψιμο
- κάθε κύριο υπομητρώο ΣΘΟ $\Rightarrow$ αντιστρέψιμο $\Rightarrow$ υπάρχουν L, U ώστε $A=LU$.
- Όλα τα στοιχεία της διαγωνίου θετικά
- Το συμπλήρωμα του Schur $S=A_{22}-A_{21}A_{11}^{-1}A_{12}$ είναι ΣΘΟ

ΘΕΩΡΗΜΑ: Αν $A \in \mathbb{R}^{n\times n}$ ΣΘΟ και έστω $A^{(k)}=L_kL_{k-1}\cdots L_1A$, το τελικό μητρώο που συναντάμε μετά την εφαρμογή των k πρώτων μετασχηματισμών Gauss στον A $(1 \leq k \leq n-1)$. Τότε:

1. $A^{(k)}_{k:n,k:n}$ είναι ΣΘΟ
2. Αν $a_{ij}^{(k)}$ στοιχεία του $A^{(k)}_{k:n,k:n}$ τότε

$$\max_{k\leq i,j\leq n} |a_{ij}^{(k)}| \leq \max_{k\leq i,j\leq n} |a_{ij}^{(k-1)}| \quad , k=1,\dots,n-1$$

$Ax=b$

Μέθοδος κανονικών εξισώσεων.

$\boxed{A^TAx=A^Tb.}$

Αν A αντιστρέψιμο $\Rightarrow A^TA$ ΣΘΟ. Μπορούμε να δείξουμε ότι όλα τα υπομητρώα που χρειάζονται αντιστροφή είναι αντιστρέψιμα

$$A=\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}=\begin{pmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ 0 & S \end{pmatrix}$$

$S:=A_{22}-A_{21}A_{11}^{-1}A_{12}$.

① A^TA ΣΘΟ.

② Τα μητρώα α) A_{11} β) S είναι ΣΘΟ.

* $Ax=b$ *

$$Ax=b \Rightarrow U\Sigma V^T x = b \Rightarrow x = (U\Sigma V^T)^{-1} b \Rightarrow$$

$$x = (V^T)^{-1}\Sigma^{-1}U^{-1}b \Rightarrow$$

$$x = V\Sigma^{-1}U^T b.$$

$U^TU = I$

$V^TV = I$

Σ^{-1}: διαγώνιο μητρώο με στοιχεία $\frac{1}{\sigma_j}$

$$\boxed{x = \sum_{j=1}^{n} \frac{u_j^T b}{\sigma_j} v_j}$$

ΤΥΠΟΣ ΓΙΑ ΣΥΝΙΣΤΩΣΕΣ ΤΟΥ x ως προς βάση V.

$$\kappa_2(A) = \frac{\sigma_{max}}{\sigma_{min}}$$

Επίλυση τριγωνικών συστημάτων

$Ax=b$ • όπου A κάτω τριγωνικό.

$$\begin{bmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ a_{31} & & \ddots & \\ \vdots & & & \\ a_{3n} & & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

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

• Υλοποίηση με BLAS-1-DOT:

$$\xi_j = (\beta_j - \overbrace{a_{j,1:j-1}^T \xi_{1:j-1}}^{DOT})/a_{jj} \quad , j=1,2,\dots,n.$$

$a_{11}\xi_1 + a_{12}\xi_2 + \dots + a_{1n}\xi_n = \beta_1 \Rightarrow$

$\xi_1 = \frac{\beta_1 - a_{12}\xi_2 - \dots - a_{1n}\xi_n}{a_{11}}$

$\xi_1 = \frac{\beta_1}{a_{11}}$

ξ_2

• Υλοποίηση με BLAS-1-SAXPY:

```
for j=1:n-1.
    ξj = βj / ajj
    for i=j+1:n
        βi = βi - aij ξj
    end
end
```

j=1

$\xi_1 = \frac{b_1}{a_{11}}$

i=2

Έστω ότι υπολογίσαμε λύση $\hat{x}$ του $Ax=b$ και ότι

$(A+\Delta A)\hat{x} = b + \Delta b$ ενώ $r = b - A\hat{x}$.

Το "εκ των υστέρων πίσω σφάλμα" είναι: $\beta_N = \dfrac{\|r\|}{\|A\|\|\hat{x}\|+\|b\|} = \dfrac{\|b-Ax\|}{\|A\|\cdot\|\hat{x}\|+}$

ΚΕΦΑΛΑΙΟ 8

Συνήθεις ΔΕ: 1 ανεξάρτητη μεταβλητή – συνήθως ο χρόνος.

ΔΕ μερικού τύπου: χωροεξαρτώμενες ή και χρονοεξαρτώμενες.
(περισσότερες ανεξάρτητες μεταβλητές).

$$L\left(u(z), \frac{\partial u}{\partial z_1}(z), \frac{\partial u}{\partial z_2}(z), \ldots, \frac{\partial^2 u}{\partial z_1^2}(z), \ldots, \frac{\partial^p u}{\partial z_s^p}(z)\right) = 0 \quad , z \in \Omega \subset \mathbb{R}^s$$

- ↳ γραμμικές
- ↳ μη γραμμικές
- ↳ ημιγραμμική → L γραμμικός ως προς τα ορίσματα μέγιστης τάξης

ΠΕΠΕΡΑΣΜΕΝΕΣ ΔΙΑΦΟΡΕΣ

• Διακριτοποίηση χωρίου $\Omega \to \Omega_h$.

$$u(x_j \pm h) = u(x_j) \pm h u^{(1)}(x_j) + \frac{h^2}{2} u^{(2)}(x_j) \pm \frac{h^3}{6} u^{(3)}(x_j) + \frac{h^4}{24} u^{(4)}(x_j + \theta_j^{\pm} h)$$

$$-1 < \theta_j^- < 0 < \theta_j^+ < 1$$

① $$u(x_j - h) + u(x_j + h) - 2u(x_j) = h^2 u^{(2)}(x_j) + \frac{h^4}{24}\left(u^{(4)}(x_j + \theta_j^+ h) + u^{(4)}(x_j + \theta_j^- h)\right)$$

② $$u(x_j + h) - u(x_j - h) = 2h u^{(1)}(x_j) + \frac{h^3}{6}\left(u^{(3)}(x_j + \eta_j^+ h) + u^{(3)}(x_j + \eta_j^- h)\right)$$

όπου $-1 < \eta_j^- < 0 < \eta_j^+ < 1$

Θεώρημα ενδιάμεσης τιμής: Αν f συνεχής στο $[\alpha, \beta]$ και $f(\alpha) < \lambda < f(\beta)$ τότε υπάρχει $x \in (\alpha, \beta)$ ώστε $f(x) = \lambda$.

$\lambda = \frac{f(\alpha) + f(\beta)}{2}$ και εφαρμόζουμε το θεώρημα στις $u^{(3)}(x)$ και $u^{(4)}(x)$.

$f(\alpha) < \frac{f(\alpha) + f(\beta)}{2} < f(\beta)$.

- $f = u^{4}(x)$.

$x_j + \theta_j h$.

$$\frac{u^{(4)}(x_j + \theta_j^+ h) + u^{(4)}(x_j + \theta_j^- h)}{2} = u^{(4)}(x_j + \theta_j h)$$

$|\theta_j| \le \max\{\theta_j^+, \theta_j^-\} < 1$

$$\frac{u^{(3)}(x_j + \eta_j^+ h) + u^{(3)}(x_j + \eta_j^- h)}{2} = u^{(3)}(x_j + \eta_j h)$$

$|\eta_j| \le \max\{\eta_j^+, \eta_j^-\} < 1$

$$\frac{u(x_j - h) + u(x_j + h) - 2u(x_j)}{h^2} = u_j^{(2)} + \frac{h^4}{24}\left(u^{4}(x_j + \theta_j h)\right),$$

$$\frac{u(x_j + h) - u(x_j - h)}{2h} = u_j^{(1)} + \frac{h^3}{6} u^{(3)}(x_j + \eta_j h)$$

Αν υποθέσουμε ότι $u^{(3)}, u^{(4)}$ συνεχείς στο διάστημα ορισμού $[\alpha, b]$, τα $h^2 u^{(3)}(x_j + \eta_j h)$ και $h^2 u^{(4)}(x_j + \theta_j h)$ θα τείνουν στο 0 όπως και το h^2.

3 — Τιμές της συνάρτησης στα άκρα του διαστήματος

— Τιμές της παραγώγου της συνάρτησης στα άκρα του διαστήματος

*Οι συνοριακές συνθήκες επιδρούν στην ύπαρξη και μοναδικότητα της λύσης της ΔΕ.

Η ποιότητα κάθε προσέγγισης εξαρτάται:

1. Από το h
2. μέγεθος παραγώγων $u^{(3)}, u^{(4)}$ γύρω από το x_i

* SOS

Θεωρούμε ότι επιβάλλεται να γράψουμε 1 ακόμη εξίσωση στο x_0 (για $x_i = x_0$)

π.χ. $-\frac{U_{j-1}-U_{j+1}+2U_j}{h^2} + b_j\frac{U_{j+1}-U_{j-1}}{2h} + c_jU_j = d_j \quad , j=1,\dots,n.$

$$(1)\quad \frac{-U_{-1}+U_1+2U_0}{h^2}+b_0\frac{U_1-U_{-1}}{2h}+c_0U_0=d_0$$

όπου $U_{-1}=u(x_0-h)=u(x_{-1})$ $\qquad u_x(x_0)=g$.

- Διακριτοποιούμε τη συνοριακή συνθήκη

$$\frac{U_1-U_{-1}}{2h}=g \Rightarrow \boxed{U_{-1}=U_1-2hg}.$$

- Απαλοίφουμε τον όρο U_{-1} αντικαθιστώντας με $U_{-1}=U_1-2hg$ από (1)

$n+1$ εξισώσεις → σύστημα ως προς $[U_0,U_1,U_2,\dots,U_n]$: $A_{ND}U=F$

$A_{ND}=\text{trid}_{n+1}[\gamma_j,\alpha_j,\beta_j]$.

$\|U-u\|$, να συμπεριφέρει αμέσως ...

προκύπτουν από τη διακριτοποίηση και τη χρήση εξισώσεων διαφορ. για να προσεγγίσουμε το αρχικό πρόβλημα.

$\|U-u\|$

$*$ $\boxed{}$ $*$

→ Αν ένα μητρώο έχει όλες τις ιδιοτιμές μη μηδενικές τότε αντιστρέψιμο $\Rightarrow$ π.χ. LU με μερική οδήγηση

→ Αν ένα μητρώο έχει όλες του τις ιδιοτιμές θετικές πραγματικές τότε είναι ΣΘΟ $\Rightarrow$ π.χ. Cholesky.

LU για τριδιαγ. συστήματα → κόστος $O(n)$.

- Από τις ιδιοτιμές ενός συμμετρικού μητρώου A ο δείκτης κατάστασης $\boxed{K_2(A) = \dfrac{|\lambda_{max}(A)|}{|\lambda_{min}(A)|}}$

ΕΠΙΛΥΣΗ ΠΡΟΒΛΗΜΑΤΟΣ ΑΡΧΙΚΩΝ ΤΙΜΩΝ ΣΔΕ 1ης τάξης:

$\dfrac{du}{dt} = f(t,u)$, $u(t_0)=c$. Θέλουμε τιμές του u στο $[t_0,T]$.

t: «χρόνος»

$u, f \in \mathbb{R}^n$

$c \in \mathbb{R}^n$ σταθερά

1. Αν η f εξαρτάται μόνο από το t και όχι από το u τότε ολοκλήρωση f: $u(t) =$ [redacted] $u(t_0) + \int_{t_0}^{t} f(\tau)d\tau$.
   υπολογιστική ολοκλήρωση → μία τιμή της $u(t)$
   $\neq$ επίλυση ΣΔΕ → προσέγγιση συνάρτησης (της u)
   - μέσω τιμών στο $[t_0,t]$: $[u(t_0), u(t_1), \ldots$
   - μπορεί μόνο το $u(t)$.

Συνθήκη Lipschitz: $\|f(t,u)-f(t,v)\| \leq \lambda \|u-v\|$, $\forall u, v \in \mathbb{R}^n, t \geq t_0$ $\lambda > 0$ ανεξ. u,

↑

Αν ισχύει, η ΣΔΕ έχει μοναδική λύση.

Παράδειγμα

$\frac{du}{dt} = e^{-t}, u(0) = c.$

$u(t) = c + \int_0^t e^{-t} = u(0) + 1 - e^{-t}.$

$\frac{du}{dt} = -u \quad , u(0) = c \Longrightarrow u(t) = e^{-t} u(0)$

[redacted]

ΜΟΡΦΗ ΣΔΕ: $\frac{d^3 u}{dt^3} = f(t, u, u', u'')$ π.χ. $u(0) = \gamma_1, u'(0) = \gamma_2, u''(0) = \gamma_3.$

Θέτουμε:
1) $\eta_1 = u$
2) $\eta_2 = u'$
3) $\eta_3 = u''$

$\Rightarrow$ ΣΥΣΤΗΜΑ

$\eta_1' = \eta_2 \qquad \eta_1(0) = \gamma_1$

$\eta_2' = \eta_3 \qquad \eta_2(0) = \gamma_2$

$\eta_3' = f(t, \eta_1, \eta_2, \eta_3) \qquad \eta_3(0) = \gamma_3$

Αν π.χ. $f(t, u, u', u'') = -2u - 3u' - 4u''$

$\rightsquigarrow \frac{d\eta}{dt} = A\eta$ όπου $\eta = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \end{pmatrix}$

$\frac{d\eta_1}{dt} = \eta_2.$

$\frac{d\eta_2}{dt} = \eta_3.$

$\frac{d\eta_3}{dt} = -2\eta_1 - 3\eta_2 - 4\eta_3.$

$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -3 & 4 \end{pmatrix}$

Μέθοδος Euler $\longrightarrow$ Ρητός Euler : $\frac{d}{dt} u(t) = \frac{u(t+\Delta t) - u(t)}{\Delta t} + O(\Delta t)$

$\Delta t_k = t_{k+1} - t_k$: μήκος βήματος.

$\boxed{U_{k+1} = U_k + \Delta t_k f(t_k, U_k)}$

ΕΥΧΡΗΣΤΗ

ΑΛΛΑ.... ΕΥΘΡΑΥΣΤΗ.

Έστω $u'(t)=a\,u(t)$, $u(0)=c$ για σταθερό a.

Τότε η λύση είναι $u(t)=e^{at}c$.

- $a>0 \Rightarrow$ η θεωρητική λύση μεγαλώνει $u(t_0)<u(t_1)<\dots$ και $\lim_{t\to\infty} u(t)=\infty \longrightarrow$ ΑΣΤΑΘΗΣ ΣΥΝΑΡΤΗΣΗ
- $a<0 \Rightarrow \lim_{t\to\infty} u(t)=0 \longrightarrow$ ΕΥΣΤΑΘΗΣ ΣΥΝΑΡΤΗΣΗ.

ΠΙΣΩ EULER :

$$\frac{d}{dt}u(t)=\frac{u(t)-u(t-\Delta t)}{\Delta t}+O(\Delta t) \Rightarrow$$

↓
προηγούμενη συναρτήσει επόμενης.

$$\frac{U(t)-U(t-\Delta t)}{\Delta t}=f(t,U(t)) \qquad U(t_0)=c.$$

$$U(t-\Delta t)=U(t)-\Delta t\, f(t,U(t)).$$

*Πίσω Euler για ευστάθεια δε θέτει κανένα περιορισμό στο μέγεθος του Δt.

*Μόνος περιορισμός $\longrightarrow$ αποδεκτό σφάλμα αποκοπής.

$$U_{n+1} = U_n + h \sum_{i=1}^{s} b_i k_i$$

όπου $k_i = f\left(t_n + c_i h, U_n + h \sum_{j=1}^{s} a_{ij} k_j\right)$

Πίνακας "Butcher":

$$\begin{array}{c|cccc} c_1 & a_{11} & a_{12} & \dots & a_{1s} \\ c_2 & \vdots & & & \vdots \\ \vdots & & & & \\ c_s & a_{s1} & a_{s2} & \dots & a_{ss} \\ \hline & b_1 & b_2 & \dots & b_s \end{array} = \begin{array}{c|c} c & A \\ \hline & b^T \end{array}$$

- **$s \leq 4$:** άμεσες μέθοδοι τάξης s.