Επιστημονικός Υπολογισμός Ι
ΗΥ 343: ΔΙΑΛΕΞΗ 8

Ε. Γαλλόπουλος
Τμήμα Η/Υ & Πληροφορικής
Πανεπιστήμιο Πατρών

ΤΜΗΥΠ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ & ΠΛΗΡΟΦΟΡΙΚΗΣ

Πανεπιστήμιο Πατρών

Στόχος

Να εκτιμήσουμε (άνω φράγμα για) το απόλυτο ή σχετικό σφάλμα

$$\|f_{\text{prog}}(x^*) - f(x)\| \text{ ή } \frac{\|f_{\text{prog}}(x^*) - f(x)\|}{\|f(x)\|}.$$

Αν το σφάλμα είναι μικρό (π.χ. της τάξης του ϵ της μηχανής) ο αλγόριθμος f είναι ακριβής.

Ορισμός 1. *Το παραπάνω σφάλμα αποκαλείται προς τα εμπρός σφάλμα (απόλυτο ή σχετικό) (=forward error).* □

Αλλά η ακριβής τιμή της $f(x)$ είναι άγνωστη,

επομένως το καλύτερο που μπορούμε να έχουμε είναι άνω φράγμα για το εμπρός σφάλμα,

γι' αυτό Συχνά λέμε «εμπρός σφάλμα» εννοώντας «άνω φράγμα για το εμπρός σφάλμα».

Πρόκληση: Το άνω φράγμα να είναι όσο πιο σφιχτό γίνεται.

3

Πίσω ευστάθεια: Ορισμός και μετρική

- Ορισμός
  - Αν για τον «αλγόριθμο» f_{prog} που υπολογίζει τη συνάρτηση f στο x μπορούμε να βρούμε x_{prog} ώστε να ισχύουν
    1. το $\|x_{\text{prog}} - x\|$ να είναι (σχετικά) μικρό
    2. να ισχύει $f_{\text{prog}}(x) = f(x_{\text{prog}})$

τότε λέμε ότι ο «αλγόριθμος» f_{prog} είναι «πίσω ευσταθής».

Όσο πιο μικρή η απόσταση του x_{prog} από το x τόσο πιο πίσω ευσταθής είναι ο αλγόριθμος.

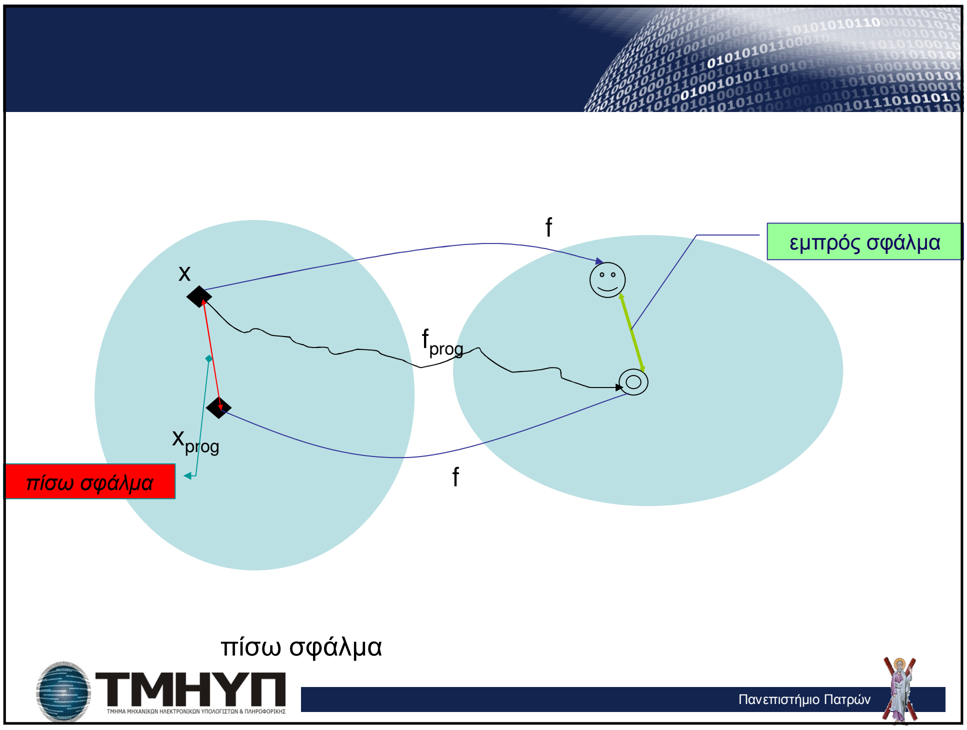
Μετρική: Η ελάχιστη δυνατή σχετική απόσταση θεωρείται ως το μέτρο της πίσω ευστάθεις.

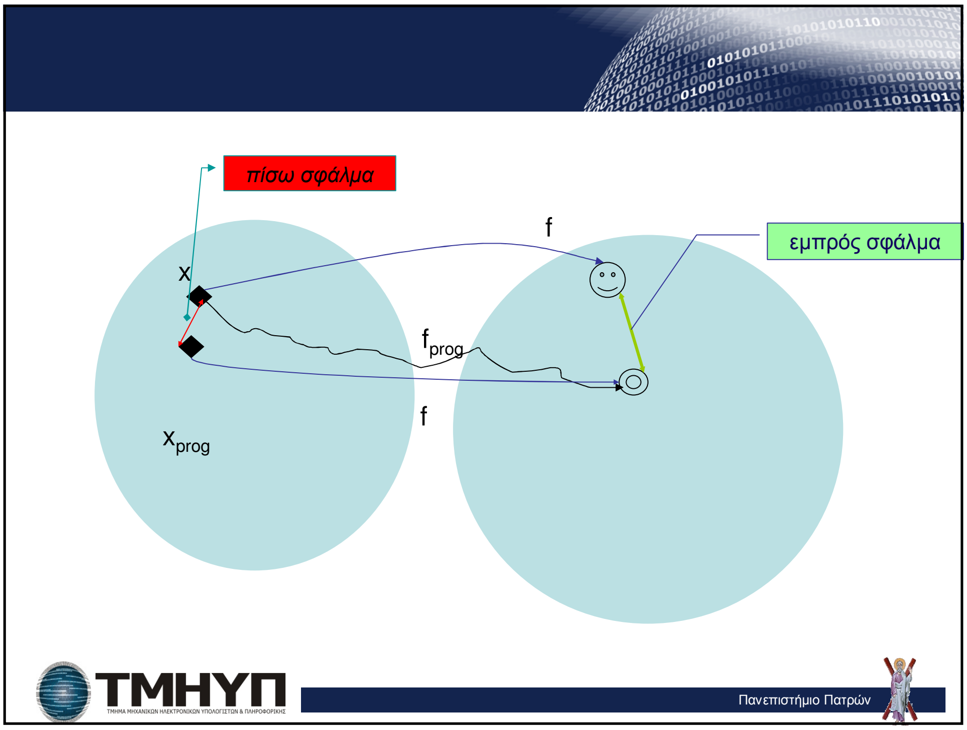
Ορισμός: Αν ο αλγόριθμος f_{prog} έχει δειχτεί πίσω ευσταθής, ο δείκτης κατάστασης (υλοποίησης) αλγορίθμου ορίζεται ως η ελάχιστη τιμή $\text{cond}(f_{\text{prog}})$ για την οποία

$$\frac{\|x - x_{\text{prog}}\|}{\|x\|} \leq \text{cond}(f_{\text{prog}})\mathbf{u}.$$

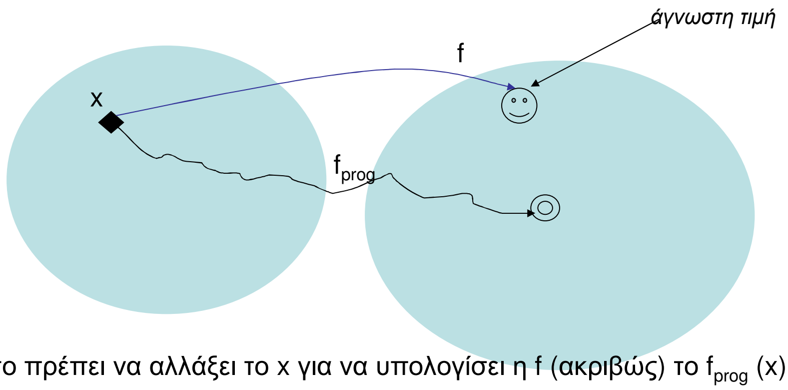
για όλα τα δυνατά x και αντίστοιχα x_{prog}. Το $\frac{\|x-x_{\text{prog}}\|}{\|x\|}$ αποκαλείται συχνά πίσω σφάλμα στο x.

Ο δείκτης $\text{cond}(f_{\text{prog}})$ δείχνει το πίσω σφάλμα για όλο το πεδίο ορισμού της f.


f
x
εμπρός σφάλμα
f_{prog}
x_{prog}
πίσω σφάλμα
f
πίσω σφάλμα
ΤΜΗΥΠ
Πανεπιστήμιο Πατρών


πίσω σφάλμα
f
εμπρός σφάλμα
x
f_{prog}
f
x_{prog}
ΤΜΗΥΠ
Πανεπιστήμιο Πατρών

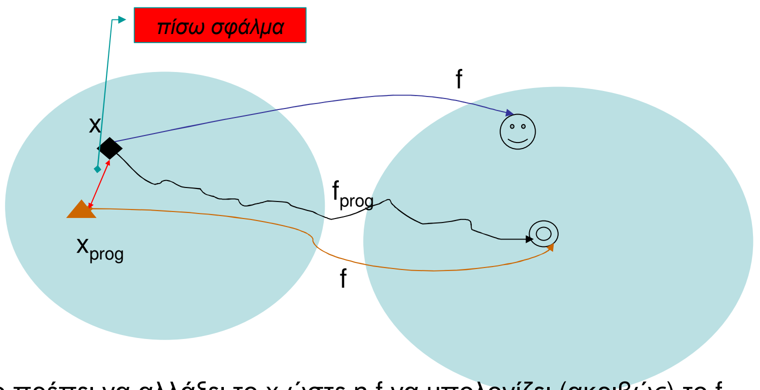
Το πρόβλημα: Να εκτιμήσουμε φράγμα για το εμπρός σφάλμα



Πόσο πρέπει να αλλάξει το x για να υπολογίσει η f (ακριβώς) το f_{prog} (x)?

$$f(x+?) = f_{prog}(x)$$

x_{prog}-x



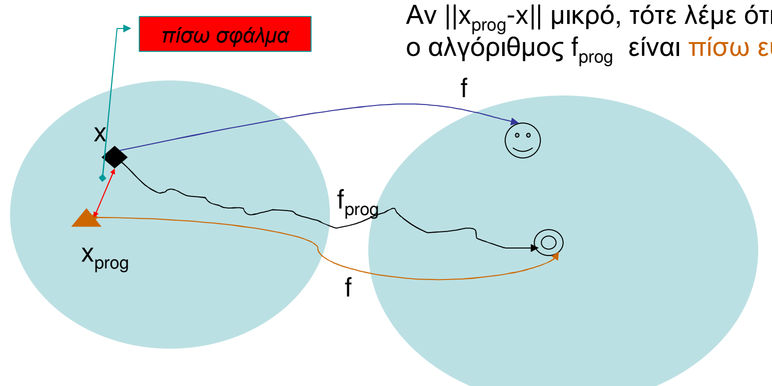
Πόσο πρέπει να αλλάξει το x ώστε η f να υπολογίζει (ακριβώς) το f_{prog} (x)?

$$f(x+?) = f_{prog}(x)$$

x_{prog}-x

ΤΜΗΥΠ
Πανεπιστήμιο Πατρών

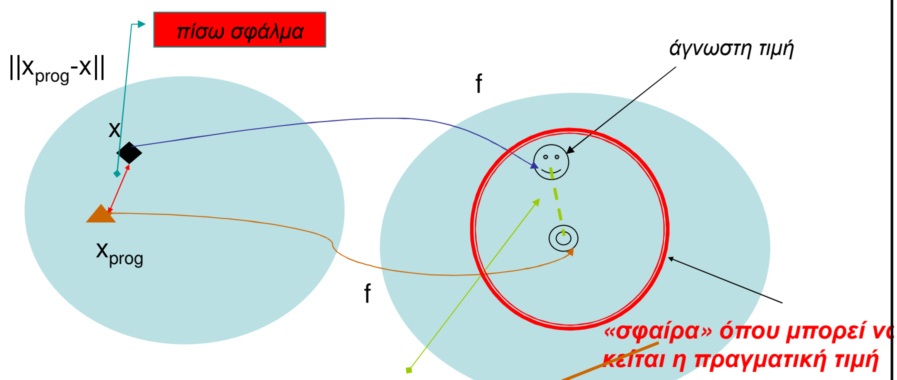
Αν $||x_{prog}-x||$ μικρό, τότε λέμε ότι ο αλγόριθμος f_{prog} είναι πίσω ευσταθής



Πόσο πρέπει να αλλάξει το x ώστε η f να υπολογίζει (ακριβώς) το f_{prog} (x)?

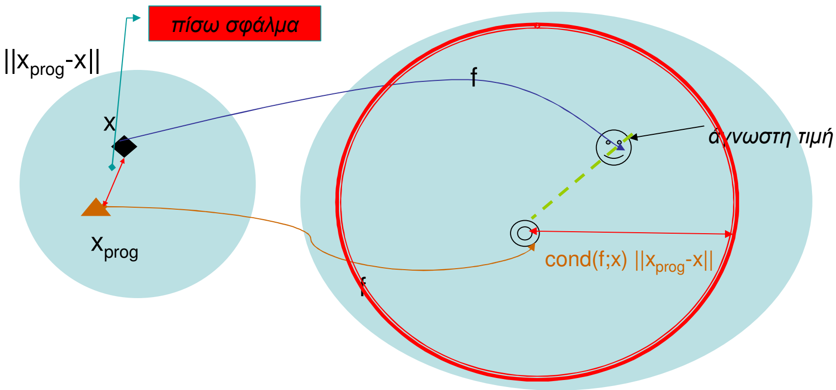
$f(x+?) = f_{prog}(x)$

$x_{prog}-x$

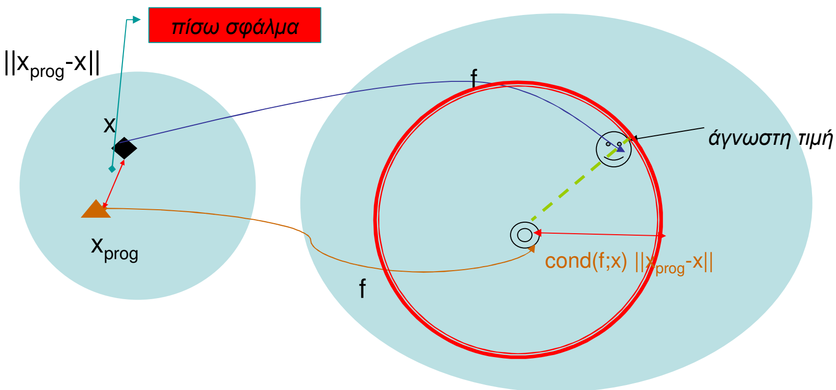


Η εκτίμηση του εμπρός σφάλματος γίνεται «μαθηματικά» βάσει
- του «δείκτη κατάστασης του προβλήματος» (δηλ. της f)
- χρησιμοποιώντας το πίσω σφάλμα.

$||f_{prog}(x) - f(x)|| \leq$ **cond(f;x)** $||x_{prog}-x||$



Αν η συνάρτηση είναι πολύ ευαίσθητη, το cond(f;x) είναι μεγάλο (π.χ. αν $\frac{df}{dx}(x)$ μεγάλο) επομένως το f(x) μπορεί να απέχει πολύ από το $f(x_{prog})$

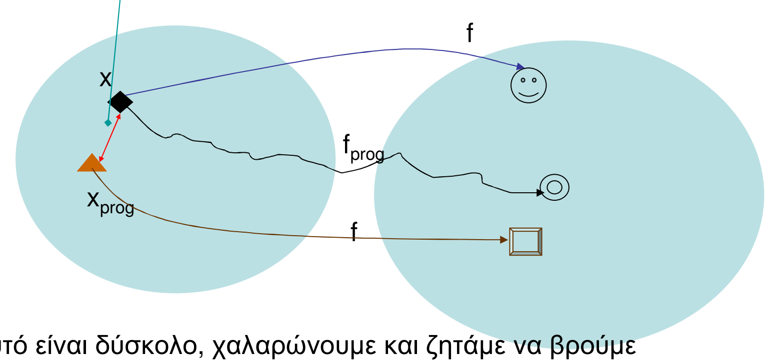


Για δεδομένο cond(f;x), η $||f(x_{prog})-f(x)||$ θα κυμαίνεται ανάλογα με το πίσω σφάλμα $||x_{prog}-x||$

ΤΜΗΥΠ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ & ΠΛΗΡΟΦΟΡΙΚΗΣ
Πανεπιστήμιο Πατρών

Αν η πίσω ανάλυση είναι δύσκολη – χαλαρώνουμε!

Αν $||x_{prog}-x||$ και $||f(x_{prog})-f_{prog}(x)||$ μικρό, λέμε ότι ο αλγόριθμος f_{prog} είναι ευσταθής



Αν αυτό είναι δύσκολο, χαλαρώνουμε και ζητάμε να βρούμε πόσο πρέπει να αλλάξει το x ώστε η f να υπολογίζει τιμή κοντά στο $f_{prog}(x)$

$$f(x+?) \approx f_{prog}(x)$$

$x_{prog}-x$

Δυο βασικές μεθοδολογίες εκτίμησης ελάχιστου άνω φράγματος εμπρός σφάλματος

1. Προς τα εμπρός ανάλυση
   - «παρακολούθηση» του σφάλματος σε κάθε βήμα του αλγορίθμου
2. Προς τα πίσω ανάλυση σφάλματος
   - εκτίμηση «πίσω σφάλματος» ή «δείκτη κατάσταση αλγορίθμου»
     - πόσο πρέπει να αλλάξουμε τα δεδομένα ώστε αν λύσουμε ακριβώς, να πάρουμε τα ίδια αποτελέσματα με αυτά που υπολογίστηκαν σε πεπερασμένη ακρίβεια;
   - εκτίμηση «δείκτη κατάστασης προβλήματος»
     - πόσο αλλάζουν τα αποτελέσματα αν αλλάξουμε λίγο τα δεδομένα (ανάλογα με το πίσω σφάλμα)

Πίσω ανάλυση σφάλματος

Χρησιμοποιώντας την παραπάνω ορολογία να δούμε τι θα γινόταν αν μπορούσαμε να δείξουμε ότι

> το υπολογισμένο αποτέλεσμα $f_{\text{prog}}(x)$ είναι το αποτέλεσμα των υπολογισμών με αριθμητική άπειρης ακρίβειας που απαιτούνται για την f μόνον που τα δεδομένα εισόδου μπορεί να είναι λίγο διαφορετικά.

Δηλαδή αν υπάρχουν δεδομένα x_{prog} τέτοια ώστε

1. το $\|x_{\text{prog}} - x\|$ να είναι (σχετικά) μικρό
2. να ισχύει ακριβώς $f_{\text{prog}}(x) = f(x_{\text{prog}})$

Να δούμε πώς αυτό βοηθά στην εκτίμηση του εμπρός σφάλματος:

$$\|f_{\text{prog}}(x) - f(x)\| = \|f(x_{\text{prog}}) - f(x)\|$$

> $\Rightarrow$ η εκτίμηση του εμπρός σφάλματος μπορεί να γίνει διερευνώντας πόσο μεταβάλλεται η τιμή $f(x)$ αν αλλάξουμε το x σε x_{prog}.

μαθηματική θεωρία διαταραχών (perturbation theory)

Επηρεάζουν τα παρακάτω:

- Η απόσταση $\|x_{\text{prog}} - x\|$.
- Η ευαισθησία της f σε μεταβολές μεγέθους $O(\|x_{\text{prog}} - x\|)$.

Παράδειγμα

- **Αν μια βιβλιοθήκη επιστρέφει για κάθε 0<x $\in$ F αντί για τον ακριβή λογάριθμο y=ln(x) την τιμή**

$$\tilde{y} = (1+e)\ln(x)$$

ποιός είναι ο δείκτης κατάστασης του αλγορίθμου;

- **Αξιοποιούμε το γεγονός ότι ισχύει ακριβώς ότι**

$$\tilde{y} = (1+e)\ln(x) = \ln(x^{1+e})$$

Επομένως θέτουμε $x_{\text{prog}} := x^{1+e}$

άρα

$$\begin{aligned} \left|\frac{x_{\text{prog}} - x}{x}\right| &= \left|\frac{x^{1+e} - x}{x}\right| = |x^e - 1| \approx |e\ln(x)| \\ &\leq 5\ln(x)\mathbf{u} \\ \text{cond}(ln; x) &\leq 5\ln(x)\mathbf{u} \end{aligned}$$

Υπενθυμίζουμε τη «χονδρική» ανάλυση

Αν θεωρήσουμε ότι η συνάρτηση f μπορεί να αναπαρασταθεί αρκετά αξιόπιστα από τους γραμμικούς όρους της σειράς Taylor

$$f(x_{\rm prog}) - f(x) \;\approx\; f'(x)(x_{\rm prog} - x)$$

επομένως

$$\begin{aligned}\frac{\|f(x_{\rm prog}) - f(x)\|}{\|f(x)\|} &\le \frac{\|f'(x)\|}{\|f(x)\|}\|x_{\rm prog} - x\| \\ &\le \left(\frac{\|f'(x)\|}{\|f(x)\|}\|x\|\right)\frac{\|x_{\rm prog} - x\|}{\|x\|}\end{aligned}$$

(1) εμπρός σχετικό σφάλμα της f στο x υπολογισμένη με f_{prog}

(2) «δείκτης κατάστασης» της f στο x εξαρτάται μόνον από (f, x)

(3) εξαρτάται από (f_{prog}, x)

Αν γνωρίζουμε τιμές ή σφικτά άνω φράγματα για τους όρους δεξιά, μπορούμε να φράξουμε το εμπρός σχετικό σφάλμα!

ΕΜΠΡΟΣ ΣΦΑΛΜΑ ≤ (ΔΕΙΚΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΠΡΒΛ) × (ΠΙΣΩ ΣΦΑΛΜΑ ΑΛΓΟΡΙΘΜΟΥ)

- **Βασικά μας ενδιαφέρει μια ποσότητα (η μικρότερη δυνατή) που να εκφράζει τη «διαστολή» της αλλαγής στα στοιχεία της εισόδου μέγεθος της εισόδου ώστε να επικαλύπτει το μέγεθος της αλλαγής στα στοιχεία εξόδου!**
- **Να βρούμε Κ τέτοιο ώστε**

σχετική αλλαγή τάξης u

$$\frac{\|\Delta f\|}{\|f\|} \le K\frac{\|\Delta x\|}{\|x\|}$$

- Το K θα λέγεται δείκτης κατάστασης του προβληματος

- Αν μπορούμε να βρούμε τέτοιο K, το πρόβλημα αποκαλείται καλά τοποθετημένο (**well-posed**)
- αν όχι (δηλ. K=∞) τότε κακά τοποθετημένο (**ill-posed**)

- Το μέγεθος του K μετρά την κατάσταση,
  - μεγάλο K σημαίνει «κακή κατάσταση» και το πρόβλημα είναι δύσκολο να λυθεί
  - συνήθως σημαίνει ότι είναι «κοντά» σε «μη επιλύσιμο» πρόβλημα
  - Ο ακριβής ορισμός και ο υπολογισμός του K δεν είναι εύκολη υπόθεση!

Ενδιαφέρον παράδειγμα

- Υπολογισμός του $f(x) = \sqrt{1+x^2} - 1$ για μικρές τιμές του **x**
- Ο υπολογισμός δημιουργεί προβλήματα γιατί έχουμε καταστροφική απαλοιφή – αναδεικνύονται τα σφάλματα στρογγύλευσης μετά την αφαίρεση
- ... και όμως, ο δείκτης κατάστασης του προβλήματος είναι μικρός

$$\mathrm{cond}(f;x) = \frac{x}{\sqrt{1+x^2}} \frac{x}{\sqrt{1+x^2}-1}$$

$$\lim_{x\to 0} \mathrm{cond}(f;x) = 2$$

- Ειδικότερα, για μικρές τιμές του x
  - αν χρησιμοποιήσουμε τον παραπάνω υπολογισμό, το εμπρός σχετικό σφάλμα μπορεί να γίνει πολύ μεγάλο (≈ 1)
  - Μπορούμε να αποφύγουμε την αφαίρεση με την ισοδύναμη έκφραση!

$$g(x) = \frac{x^2}{\sqrt{1+x^2}+1}$$

  - Εύκολα φαίνεται ότι $f \equiv g$
  - Ο δείκτης κατάστασης προβλήματος παραμένει ίδιος!

ΤΜΗΥΠ

Πανεπιστήμιο Πατρών

Πώς θα δείχναμε πίσω ευστάθεια;

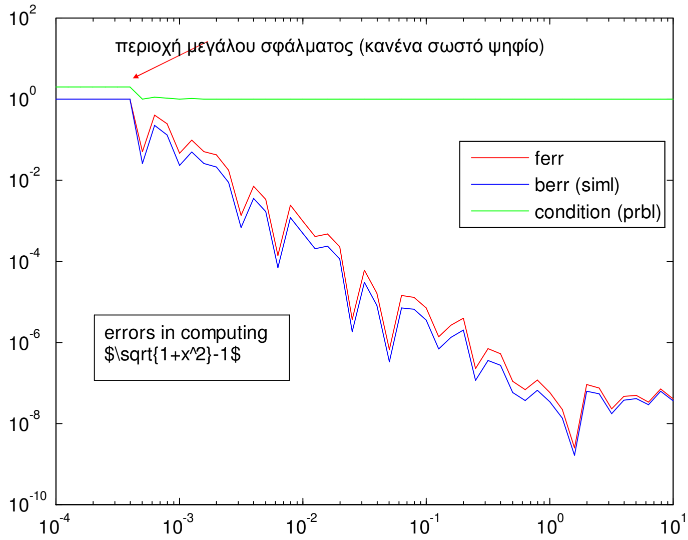
- Πρέπει να βρούμε το κατάλληλο x_{prog}
- Δεν είναι εύκολο με τη μέθοδο διάδοσης των δ (οι πράξεις με σταθερές δυσκολεύουν τα πράγματα).
- Θα το κάνουμε αριθμητικά:
  - Για ένα υπολογισμένο $z=f_{prog}(x)$ θα χρησιμοποιήσουμε ως x_{prog} το

$$x_{\text{prog}} := \sqrt{(z+1)^2 - 1}$$

Ο υπολογισμός πρέπει να γίνει σε άπειρη ακρίβεια …
… θα προσομοιώσουμε με διπλή ακρίβεια

ΤΜΗΥΠ

Πανεπιστήμιο Πατρών



Πείραμα (MATLAB 7.5) με x=single(10).^[-4:0.1:1]

Χρήσιμες έννοιες: Νόρμα μητρώου

Ορισμός 4. *Έστω $A \in \mathbb{R}^{m\times n}$, οι διανυσματικές νόρμες $\|\cdot\|'$ και $\|\cdot\|$ για τους δ.χ. $\mathbb{R}^m$ και $\mathbb{R}^n$ και έστω συνάρτηση $\|\cdot\|_M : \mathbb{R}^{m\times n} \to \mathbb{R}$ που ορίζεται ως*

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

λέγεται νόρμα μητρώου παραγόμενη απο τις παραπάνω διανυσματικές νόρμες. □

Ισοδύναμα, μπορούμε να γράψουμε ότι

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

"sup" *είναι το ελάχιστο άνω φράγμα, Για πεπερασμένα σύνολα συμπίπτει με το μέγιστο* (max)

Έστω $A \in \mathbb{R}^{n\times n}$.

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

Επίσης η νόρμα του Frobenius που για γενικά μητρώα $A \in \mathbb{R}^{m\times n}$ ορίζεται ως εξής:

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

Σημειώνουμε ότι η φασματική νόρμα και η νόρμα του Frobenius ικανοποιούν το εξής: Αν $X^*X = I, Y^*Y = I$, τότε $\|XAY\| = \|A\|$.

Σχέσεις ισοδυναμίας

Όπως και στην περίπτωση της νόρμας διανύσματος, για κάθε συγκεκριμένο διανυσματικό χώρο πεπερασμένης διάστασης, οι νόρμες μητρώων είναι ισοδύναμες.
Για παράδειγμα ισχύει ότι

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

όπου οι τιμές του ζ_{pq} δίδονται στον παρακάτω πίνακα:

$p \setminus q$	1	2	∞	F
1	1	$\sqrt{m}$	m	$\sqrt{m}$
2	$\sqrt{n}$	1	$\sqrt{m}$	1
∞	n	$\sqrt{n}$	1	$\sqrt{n}$
F	$\sqrt{n}$	$\sqrt{\text{rank}(A)}$	$\sqrt{m}$	1

Η συνάρτηση **norm** στη MATLAB είναι εσωτερική συνάρτηση που υπολογίζει τη νόρμα μητρώων και διανυσμάτων.

- Η ίδια συνάρτηση ισχύει ανεξάρτητα από τον τύπο του ορίσματος (διάνυσμα ή μητρώο) παρόλο που οι ορισμοί είναι διαφορετικοί
- Το είδος της νόρμας επιλέγεται από παράμετρο εισόδου
- Για πολύ μεγάλα ορίσματα χρησιμοποιείται η **normest** (πιο προχωρημένο μάθημα)

```
NORM   Matrix or vector norm.
    For matrices...
      NORM(X) is the largest singular value of X, max(svd(X)).
      NORM(X,2) is the same as NORM(X).
      NORM(X,1) is the 1-norm of X, the largest column sum,
                     = max(sum(abs(X))).
      NORM(X,inf) is the infinity norm of X, the largest row sum,
                     = max(sum(abs(X'))).
      NORM(X,'fro') is the Frobenius norm, sqrt(sum(diag(X'*X))).
      NORM(X,P) is available for matrix X only if P is 1, 2, inf or 'fro'.
```

100

Χρήσιμες έννοιες: Ιακωβιανό μητρώο

Ορισμός 5. *Έστω η συνάρτηση $g : \mathbb{R}^n \to \mathbb{R}^m$ για την οποία γνωρίζουμε ότι κάθε στοιχείο της $g_i(x), i = 1, ..., m$ είναι συνεχώς παραγωγήσιμο στο $x \in \mathbb{R}^n$. Τότε ονομάζουμε Ιακωβιανό μητρώο ή απλά Ιακωβιανή της g στο x το $m \times n$ μητρώο με όρους*

$$\frac{\partial g_i}{\partial x_j}, i = 1 : m, j = 1 : n.$$

□

Συμβολίζεται ως $g^{'}(x) \in \mathbb{R}^{m\times n}$ ή $g^{'}(x)_{ij} = \frac{\partial g_i}{\partial x_j}(x)$, ή $g^{'}(x) = J(x) = \nabla g(x)^T$.
Χρησιμοποιώντας το Ιακωβιανό μητρώο μπορούμε να επιχειρήσουμε μια γραμμική προσέγγιση της $g(x + h)$ χρησιμοποιώντας τη συνάρτηση $g(x) + J(x)h$.

Παράδειγμα 8. *Η Ιακωβιανή των*

$f(\xi_1, \xi_2, \xi_3, \xi_4) = \xi_1\xi_3 + \xi_2\xi_4$ *είναι* $[\xi_3, \xi_4, \xi_1, \xi_2]$

$f(\xi_1, \xi_2) = e^{\xi_1\xi_2} + \xi_1$ *είναι* $[\xi_2 e^{\xi_1\xi_2} + 1, \xi_1 e^{\xi_1\xi_2}]$

$f(\xi_1, \xi_2) = \begin{pmatrix} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 \end{pmatrix}$ *είναι* $\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}$ □

$f(\xi_1, \xi_2) = \alpha_{11}\xi_1^2 + 2\alpha_{12}\xi_1\xi_2 + \alpha_{22}\xi_2^2$ *είναι* $2[\alpha_{11}\xi_1 + \alpha_{12}\xi_2, \alpha_{12}\xi_1 + \alpha_{12}\xi_2]$

Μερικές περιπτώσεις για το δείκτη κατάστασης προβλήματος

$$f : \mathbb{R} \to \mathbb{R} : \Rightarrow \text{cond}(f;x) := \frac{|f'(x)|}{|f(x)|}|x|$$

$$f : \mathbb{R}^n \to \mathbb{R}^m : \Rightarrow \text{cond}(f_i;x_j) := |\frac{\partial f_i}{\partial x_j}(x)x_j|\frac{1}{|f_i(x)|}$$

$$\hookrightarrow \text{cond}(f;x) := \|\text{cond}(f_i;x_j)\|$$

$$f : \mathbb{R}^n \to \mathbb{R}^m : \Rightarrow \text{cond}(f;x) := \frac{\|f'(x)\|}{\|f(x)\|}\|x\|$$

ΤΜΗΥΠ

Πανεπιστήμιο Πατρών

Παραδείγματα

$f(x)$	cond(f;x)
x	1
x^n	n
$\exp(x)$	x
$\cos(x)$	$\lvert x\tan(x)\rvert$

ΤΜΗΥΠ

Πανεπιστήμιο Πατρών

Κλασική θεωρία κατάστασης - δείκτης κατάστασης προβλήματος

Ορισμός 3 (Rice' 66). *Έστω γραμμικοί νορμισμένοι*[1] *χώροι* X, Y *που διαθέτουν νόρμα και έστω η απεικόνιση* $f : \Omega \subset X \to Y$, *όπου* Ω *είναι ανοικτό χωρίο. Έστω* x^* *σταθερό και η τιμή* $y^* := f(x^*)$. *Υποθέτουμε ότι τα* x^*, y^* *δεν είναι τα μηδενικά στοιχεία των* X, Y. *Ο ασυμπτωτικός σχετικός δείκτης κατάστασης της απεικόνισης* f *στο* x^* *ως προς μικρές αλλαγές του* x^* *ορίζεται ως*

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

εφόσον το όριο υπάρχει.

προσοχή στις νόρμες

Εμείς «απλοποιήσαμε» ως

$$\operatorname{cond}(f;x) := \frac{\|f'(x)\|}{\|f(x)\|}\,\|x\|$$

προσοχή στην παράγωγο (Ιακωβιανό μητρώο της f στο x)

[1] Λέγονται και σταθμισμένοι.

18