Επιστημονικός Υπολογισμός Ι
ΗΥ 343: ΔΙΑΛΕΞΗ 7

Ε. Γαλλόπουλος
Τμήμα Η/Υ & Πληροφορικής
Πανεπιστήμιο Πατρών

ΤΜΗΥΠ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΗΛΕΚΤΡΟΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ & ΠΛΗΡΟΦΟΡΙΚΗΣ

Πανεπιστήμιο Πατρών

Παράδειγμα

Υποθέτουμε ότι αν $0 < x \in F$ τότε η τετραγωνική ρίζα υπολογίζεται με $\tilde{\sqrt{x}} = \sqrt{x}(1+\delta)$ όπου $|\delta| \le u$.
Έστω $x, y \in F$ και $x^2 + y^2 \in G$. Να υπολογίσουμε το σφάλμα για το $z = \frac{x}{\sqrt{x^2+y^2}}$.

Πράξη σε α.κ.υ.	*Υπολογισθέν αποτέλεσμα*
$t_1 \leftarrow \text{fl}(x * x)$	$x^2(1+\delta_1)$
$t_2 \leftarrow \text{fl}(y * y)$	$y^2(1+\delta_2)$
$t_3 \leftarrow \text{fl}(t_1 + t_2)$	$(t_1+t_2)(1+\delta_3)$
$t_3 \leftarrow \text{fl}(\sqrt{t_3})$	$\sqrt{t_3}(1+\delta_4)$
$t_1 \leftarrow \text{fl}(\frac{x}{t_3})$	$\frac{x}{t_3}(1+\delta_5)$

Yap 2.4.1398 - [2004Root.dvi]

Με διαδοχική αντικατάσταση έχουμε:

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

όπου γνωρίζουμε για κάθε δ_j ότι είναι φραγμένο σε απόλυτη τιμή εκ των άνω από το u.

(no source specials found) 444,421pt Page: 75 (75th of 269)

Η έκφραση είναι μεν ενδεικτική του πώς επιδρούν στο τελικό αποτέλεσμα τα σφάλματα που γίνονται σε κάθε βήμα ...

Λέει ότι αν ακολουθήσουμε την αρχή ακριβούς στρογγύλευσης, στο τέλος θα υπολογίσουμε το t_1 ως αποτέλεσμα. Οι τιμές των δ_j δεν είναι γνωστές αλλά γνωρίζουμε ότι καθένα θα είναι φραγμένο σε απόλυτη τιμή από το **u.**

… αλλά θα τη θέλαμε πιο κατατοπιστική όσον αφορά στο τελικό σφάλμα που **εμπεριέχει** η προσέγγιση.

Το καλύτερο που μπορούμε να κάνουμε είναι να εξετάσουμε πόσο μακρυά μπορεί να είναι το t_1 από το z, δηλ. να φράξουμε το $|t_1-z|$ ή το $|t_1-z|/|z|$

Εμπρός σφάλμα

- **Στόχος μας είναι να υπολογίσουμε άνω φράγμα για το σφάλμα στα τελικά αποτελέσματα που είναι όσο πιο κοντά γίνεται στο ελάχιστο άνω φράγμα. Στην καλύτερη περίπτωση θα υπολογίσουμε το ελάχιστο άνω φράγμα του σφάλματος.**

Το πρόβλημα μοντελοποιείται ως μαθηματική συνάρτηση:

$$f : X \to Y$$

Η απεικόνιση από τα στοιχεία εισόδου στα υπολογισμένα αποτελέσματα μοντελοποιείται

$$f_{\text{prog}} : F \cap X \to F \cap Y$$

Αν συλλέξουμε

- τα δεδομένα σε διάνυσμα $x \in X$
- τα θεωρητικά αποτελέσματα σε διάνυσμα $f(x) \in Y$
- τα δεδομένα ως α.κ.υ. σε διάνυσμα $x^* \in F \cap X$
- τα υπολογισμένα αποτελέσματα σε διάνυσμα $f(x^*) \in F \cap Y$

Εμπρός σφάλμα

- **Στόχος μας είναι να υπολογίσουμε άνω φράγμα για το σφάλμα στα τελικά αποτελέσματα που είναι όσο πιο κοντά γίνεται στο ελάχιστο άνω φράγμα. Στην καλύτερη περίπτωση θα υπολογίσουμε το ελάχιστο άνω φράγμα του σφάλματος.**

Το πρόβλημα μοντελοποιείται ως μαθηματική συνάρτηση:

$$f : X \to Y$$

Η απεικόνιση από τα στοιχεία εισόδου στα υπολογισμένα αποτελέσματα μοντελοποιείται

$$f_{\text{prog}} : F \cap X \to F \cap Y$$

Αν συλλέξουμε

- τα δεδομένα σε διάνυσμα $x \in X$
- τα θεωρητικά αποτελέσματα σε διάνυσμα $f(x) \in Y$
- τα δεδομένα ως α.κ.υ. σε διάνυσμα $x^* \in F \cap X$
- τα υπολογισμένα αποτελέσματα σε διάνυσμα $f(x^*) \in F \cap Y$

θέλουμε να εκτιμήσουμε (άνω φράγμα) για

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

Αποκαλούνται **απόλυτο** (ή **σχετικό**) **εμπρός σφάλμα** του υπολογισμού της f στο x με την f_{prog}

- **Το σφάλμα θα εξαρτάται από (όπως θα περιμέναμε)**
  - **το μαθηματικό πρόβλημα (συνάρτηση f)**
  - **την υλοποίηση για τον υπολογισμό («αλγόριθμος» f_{prog})**
  - **το σύστημα α.κ.υ.**
  - **τα δεδομένα εισόδου (x και x^*)**

Εμπρός σφάλμα

- **Αν δεν γνωρίζουμε εκ των προτέρων τα δεδομένα εισόδου θα πρέπει να διερευνήσουμε το άνω φράγμα για όλο το χώρο X ή μέρος του:**

Γενικότερα: Θέλουμε να εκτιμήσουμε το ελάχιστο άνω φράγμα για όλες τις δυνατές τιμές του $x \in X$ των

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

Τα ελάχιστα άνω φράγματα αποκαλούνται **απόλυτο** (ή **σχετικό**) **εμπρός σφάλμα** στον υπολογισμό της f με την f_{prog}.

Στην πράξη, υπολογίζουμε εκτιμήσεις του ελάχιστου άνω φράγματος, δηλ. αρκούμαστε να φράξουμε όσο γίνεται καλύτερα το εμπρός σφάλμα.

ΠΡΟΣΟΧΗ: Για το σχετικό σφάλμα η μεγιστοποίηση πρέπει να γίνει υπό την προϋπόθεση ότι $\|f(x)\| \neq 0$.

Εμπρός ανάλυση σφάλματος

- Η διαδικασία που ακολουθήσαμε για να φράξουμε το εμπρός σφάλμα λέγεται <u>εμπρός ανάλυση σφάλματος</u>
- προσπαθήστε να την εφαρμόσετε σε απλές πράξεις, π.χ. _axpy, _dot
- Περιγράφεται απλά αλλά εφαρμόζεται δύσκολα «με το χέρι»
  - Έχουν γίνει πολλές προσπάθειες να «αυτοματοποιηθεί»
  - δεν έχει βρεθεί «νικητής»

- Υποθέτουμε ότι το πρόγραμμα που εκτελείται και που παράγει το $f(x)$ αποτελείται από μια σειρά από στοιχειώδεις αριθμητικές πράξεις.
- Κάθε πράξη παράγει ένα νέο αποτέλεσμα που εξαρτάται από τα δεδομένα εισόδου ή και από τις τιμές που έχουν ήδη υπολογισθεί, π.χ. αν $x = (\xi_1, \dots, \xi_N)$ και
$$a_1 = f_1(\xi_1, \dots, \xi_N), a_2 = f_2(a_1, \xi_1, \dots, \xi_N), \cdots, z = f_n(a_{n-1}, \cdots, a_1, \xi_1, \dots, \xi_N),$$
όπου $f_n \circ \cdots \circ f_1 \equiv f$ και κάθε $f_j(a_{j-1}, \cdots, a_1, x)$ αντιστοιχεί στην εκτέλεση μιας στοιχειώδους πράξης με δεδομένα ένα ή δύο από τα ορίσματα
$$\{a_{j-1}, \cdots, a_1, \xi_1, \dots, \xi_N\}.$$
- Εφόσον όμως εκτελούμε τις ανωτέρω πράξεις σε α.κ.υ., το αποτέλεσμα σε κάθε βήμα θα είναι διαφορετικό από αυτό που θα πέρναμε αν χρησιμοποιούσαμε τις κοινές πράξεις του $\mathbb{R}$, οπότε το υπολογισθέν αποτέλεσμα $\tilde{z}$ είναι προσέγγιση του z και το σφάλμα υπακούει τον κανόνα μετάδοσης που περιγράψαμε παραπάνω.
- Επομένως μπορούμε να παρακολουθήσουμε την εξέλιξη του σφάλματος από την βήμα προς βήμα με την εφαρμογή των f_j.

Η μέθοδος αποκαλείται προς τα εμπρός ανάλυση σφάλματος.

Σημαντική παρατήρηση

Έστω ότι $\odot$ είναι πολλαπλασιασμός ή διαίρεση. Τότε
$$x \tilde{\odot} y \equiv \mathrm{fl}(x \odot y) = (x \odot y)(1+\delta),$$
επομένως το αποτέλεσμα της στοιχειώδους πράξης σε α.κ.υ. $\tilde{\odot}$ μεταξύ των $x, y \in F$

... είναι ίσο με, και επομένως μπορεί να θεωρηθεί ότι είναι το ακριβές αποτέλεσμα της στοιχειώδους πράξης (στο $\mathbb{R}$) $\odot$ μεταξύ των $\tilde{x} = x, \tilde{y} = y(1+\delta)$.

ή ακόμα

... είναι ίδιο με το αποτέλεσμα της ακριβούς πράξης μεταξύ των $\tilde{x} = x(1+\delta), \tilde{y} = y$ ή μεταξύ των $\tilde{x} = x\sqrt{1+\delta}, \tilde{y} = y\sqrt{1+\delta}$, κλπ.

Αντίστοιχα στην πρόσθεση θα είχαμε:
$$x \tilde{+} y = (x+y)(1+\delta) = x(1+\delta) + y(1+\delta)$$
και το υπολογισμένο αποτέλεσμα είναι ίδιο με την ακριβή άθροιση των δεδομένων $\tilde{x} = x(1+\delta)$ και $\tilde{y} = y(1+\delta)$.

Γενίκευση

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 $O(\|x_{\text{prog}} - x\|)$.

Γενίκευση

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

μαθηματική θεωρία διαταραχών (perturbation theory)

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}.

Επηρεάζουν τα παρακάτω:

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

Μια «χονδρική» ανάλυση

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

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

επομένως

$$\frac{\|f(x_{\text{prog}}) - f(x)\|}{\|f(x)\|} \leq \frac{\|f'(x)\|}{\|f(x)\|}\|x_{\text{prog}} - x\| \leq \left(\frac{\|f'(x)\|}{\|f(x)\|}\|x\|\right)\frac{\|x_{\text{prog}} - x\|}{\|x\|}$$

(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 $\leq$ (ΔΕΙΚΤΗΣ ΚΑΤΑΣΤΑΣΗΣ ΠΡΒΛ) $\times$ (ΠΙΣΩ ΣΦΑΛΜΑ ΑΛΓΟΡΙΘΜΟΥ)

Κατάσταση (conditioning)

Κατάσταση προβλήματος: Η μελέτη της ευαισθησίας των αποτελεσμάτων ως προς διαταράξεις των στοιχείων εισόδου, και μόνο ως προς αυτές.

Κατάσταση (υλοποίησης) αλγορίθμου: Η μελέτη της επίδρασης της αριθμητικής πεπερασμένης ακρίβειας στην υλοποίηση του αλγορίθμου (μέσω ενός προγράμματος).

Δείκτης κατάστασης αλγορίθμου και πίσω σφάλμα

Ορισμός 2. *Ο δείκτης κατάστασης (της υλοποίησης) του αλγορίθμου/προγράμματος ορίζεται ως η ελάχιστη τιμή,* $\mathrm{cond}(f_{\mathrm{prog}})$, *για την οποία*

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

Το $\frac{\|x-x_{\mathrm{prog}}\|}{\|x\|}$ *αποκαλείται συχνά πίσω σφάλμ στο* x.
Ο δείκτης $\mathrm{cond}(f_{\mathrm{prog}})$ *δείχνει το πίσω σφάλμα για όλο το πεδίο ορισμού της* f. □

Ένας αλγόριθμος f_{prog} για τον υπολογισμό ενός προβλήματος f ονομάζεται πίσω ευσταθής αν ισχύουν οι συνθήκες που είδαμε, δηλ. να υπάρχουν δεδομένα x_{prog} τέτοια ώστε

- $\|x - x_{\mathrm{prog}}\|$ να είναι «μικρό»
- $f_{\mathrm{prog}}(x) = f(x_{\mathrm{prog}})$

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*

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

εφόσον το όριο υπάρχει. □

προσοχή στις νόρμες

Εμείς «απλοποιήσαμε» ως

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

προσοχή στην παράγωγο

[1] Λέγονται και σταθμισμένοι.

Δείκτης κατάστασης

- **Για μια συνάρτηση, είναι ένας αριθμός που αναδεικνύει την αλλαγή που υφίσταται η συνάρτηση αν αλλάξουν λίγο οι τιμές εισόδου**
- **Αλλαγή στην τιμή <= δείκτης κατάστασης προβλήματος * αλλαγή στην τιμή εισόδου**

Προς τα πίσω ανάλυση

Μοχλός: Έστω ότι έχουμε αποδείξει πως υπάρχει x_{prog} «κοντά» στο x ώστε αν το αποτέλεσμα που υπολογίστηκε με πράξεις α.κ.υ. είναι $z_{\text{prog}} = f_{\text{prog}}(x)$ τότε $z_{\text{prog}} = f(x_{\text{prog}})$. Τότε

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

Ιδέα Αντί να εκτιμήσουμε απευθείας το προς τα εμπρός σφάλμα, εκτιμούμε τη διαφορά

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

Αν το $\|x_{\text{prog}} - x\|$ είναι μικρό, τότε το πρόβλημα ανάγεται στο μαθηματικό (και όχι αριθμητικό) πρόβλημα της ευαισθησίας του f σε μικρές διαταράξεις (= perturbations) των στοιχείων εισόδου.

> *«Ρίξαμε το σφάλμα προς τα πίσω» και το αναγάγαμε σε σφάλμα που οφείλεται στα στοιχεία εισόδου. Από εκεί και πέρα δεν υπάρχει λόγος να εξετάζουμε το σφάλμα που οφείλεται στις ενδιάμεσες πράξεις!*

Ιστορικά σχόλια

- Η παραπάνω ιδέα είναι το κλειδί της *προς τα πίσω ανάλυσης σφάλματος* (= backward error analysis).
- Θεμελιώδης τεχνική που οφείλεται κυρίως στον Βρεταννό μαθηματικό James Hardy W-ilkinson (1919-86).
- Παραπλήσιες ιδέες προτάθηκαν και από τους Wallace Givens (1954), von Neumann και Goldstine(1947), και Turing (1948).

Παράδειγμα

Έστω $f(x_1, x_2, x_3) = (x_1 + x_2) + x_3$ όπου τα $x_j \in F$. Τότε

$$f_{\text{prog}}(x_1, x_2, x_3) = ((x_1 + x_2)(1 + \delta_1) + x_3)(1 + \delta_2)$$

άρα

$$f_{\text{prog}}(x_1, x_2, x_3) = x_1(1 + \delta_1)(1 + \delta_2) + x_2(1 + \delta_1)(1 + \delta_2) + x_3(1 + \delta_2)$$

όπου $|\delta_j| \leq u$ για $j = 1, 2$.

Επομένως

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

όπου

$$\begin{aligned}\tilde{x}_1 &= x_1(1+\delta_1)(1+\delta_2)\\ \tilde{x}_2 &= x_2(1+\delta_1)(1+\delta_2)\\ \tilde{x}_3 &= x_3(1+\delta_2)\end{aligned}$$

Επομένως $|\tilde{x}_j - x_j| = |x_j(\delta_1 + \delta_2 + \delta_1\delta_2)|$ για $j = 1, 2$ και $|\tilde{x}_3 - x_3| = |x_3\delta_2|$.
άρα

$$\begin{aligned}|\tilde{x}_j - x_j| &\le 3u|x_j|, j = 1, 2\\ |\tilde{x}_3 - x_3| &\le u|x_3|.\end{aligned}$$

Επομένως

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

όπου

$$\frac{|\tilde{x}_j - x_j|}{|x_j|} \le \rho_j u, \rho_1 = \rho_2 = 3, \rho_3 = 1.$$

Ο αλγόριθμος υπολογισμού είναι «προς τα πίσω ευσταθής»