ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

ΔΙΑΛΕΞΗ 13, 27/11/09

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

Ανάλυση σφάλματος για την επίλυση συστήματος
Βήματα:

Κατάσταση προβλήματος Πόσο αλλάζει το αποτέλεσμα για μικρές αλλαγές στα στοιχεία εισόδου; Εφαρμογή θεωρίας διαταραχών (βλ. προηγούμενες διαλέξεις).

Κατάσταση αλγορίθμου προς τα πίσω ανάλυση σφάλματος

Κατάσταση προβλήματος «επίλυση συστήματος» Ρωτάμε: Πόσο θα αλλάξει η λύση x αν αλλάξουμε τα δεδομένα A, b;
Ανάλυση:

$$(A+\Delta A)\overbrace{(x+\Delta x)}^{\hat{x}} = b+\Delta b \quad (1)$$
$$Ax = b \quad (2)$$
$$(3)$$

επομένως λύνουμε ως προς Δx και έχουμε

$$\Delta x = A^{-1}(-\Delta A\hat{x}+\Delta b)$$

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

επομένως (Θεώρημα:)

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

Από τον καθοριστικό ρόλο της ποσοτητας $\kappa(A)$, έχει νόημα να την εκλάβουμε ως «δείκτη κατάστασης του προβλήματος» της επίλυσης συστήματος[1].
Προσέξτε ότι αυτός ο δείκτης «έλαβε υπόψη του και διαταραχές 2ης τάξης!»
ΕΡΩΤΗΣΗ: Πώς μπορούμε να συνδέσουμε το παραπάνω με μια πλήρη ανάλυση του σφάλματος·

[1] Δεν λειτουργεί πάντα καλά. Θα δούμε αργότερα.

- Αν δείχναμε ότι ένας συγκεκριμένος αλγόριθμος επίλυσης είναι πίσω ευσταθής, δηλ. ότι η υπολογισμένη λύση $\tilde{x}$ είναι η ακριβής λύση ενός παραπλήσιου συστήματος,

π.χ.

$$(A + \Delta A)\tilde{x} = b + \delta b$$

έτσι ώστε $\frac{\|\Delta A\|}{\|A\|}, \frac{\|\delta b\|}{\|b\|}$ να είναι μικρά, π.χ. και φραγμένα από κάποιο μικρό ϵ.
Τότε με βάση το παραπάνω φράγμα, μπορούμε να υπολογίσουμε άμεσα χρήσιμο άνω φράγμα για το εμπρός σφάλμα!

Παρατήρηση (για όσους ενδιαφέρονται!) Για να ισχύει το παραπάνω υποθέσαμε έμμεσα ότι

$$\begin{aligned} 0 &< 1 - \kappa(A)\frac{\|\Delta A\|}{\|A\|} \\ &< 1 - \|A^{-1}\|\|\Delta A\| \end{aligned}$$

δηλ. $\|A^{-1}\|\|\Delta A\| < 1$.

Λήμμα 1. *Έστω νόρμα μητρώου που ικανοποιεί τη σχέση $\|AB\| \leq \|A\|\|B\|$. Τότε αν $\|A\| < 1$ το μητρώο $I - A$ είναι αντιστρέψιμο και $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$ και $\|(I - A)^{-1}\| \leq 1/(1 - \|A\|)$.* □

Ένας τρόπος να θέσουμε το πρόβλημα
Έστω ότι με το συμβολισμό της «προς τα πίσω ανάλυσης» σφάλματος, θεωρήσουμε ότι

$$f([A, b]) = x, \quad f_{\text{prog}}([A, b]) = \hat{x}$$

για κάποιο $\hat{x}$.
Μπορούμε να ρωτήσουμε το εξής: Είναι δυνατόν να βρούμε $\hat{A}$ τέτοιο ώστε να είναι κοντά στο A και επίσης

$$f([\hat{A}, b]) = \hat{x}$$

Τότε θα μπορούσαμε να θέσουμε $X_{\text{prog}} := [\hat{A}, b]$ και $X := [A, b]$ οπότε θα ισχύει ότι

$$f_{\text{prog}}(X) = f(X_{\text{prog}}).$$

Επομένως, θα έχουμε πίσω ευστάθεια αν $i)$ μπορούμε να επιλέξουμε τέτοιο $\hat{A}$ που επίσης όμως $ii)$ να είναι κοντά στο A.
Παρατήρηση: Εντέλει, η πίσω ευστάθεια θα κριθεί από αυτήν την απόσταση.[2]
Πρακτικός «εκ των υστέρων» υπολογισμός πίσω σφάλματος
Θεώρημα 1 (Rigal+Gaches). *Έστω ότι υπολογίσαμε τη λύση $\hat{x}$ του συστήματος $Ax = b$ και ότι θέτουμε $r = b - A\hat{x}$. Τότε, το (εκ των υστέρων) πίσω σφάλμα*

$$\beta_N = \inf\{\omega | (A + \Delta A)\hat{x} = b + \Delta b, \|\Delta A\| \leq \omega\|A\|, \|\Delta b\| \leq \omega\|b\|\}$$

[2] Για ολοκληρωμένη αντιμετώπιση, μπορούμε να θέσουμε την απόσταση σε άπειρο αν δεν υπήρχε τέτοιο $\hat{A}$.

μπορεί να υπολογιστεί από τον παρακάτω τύπο

$$\beta_N = \frac{\|r\|}{\|A\|\|\hat{x}\| + \|b\|}$$

Επίσης, οι παράγοντες $\Delta A, \delta b$ για τους οποίους $(A + \Delta A)\hat{x} = b + \Delta b$ μπορούν να υπολογιστούν.
□
Παρατηρήσεις

- Το πίσω σφάλμα καθορίζεται από την υπολογισμένη λύση $\hat{x}$ και το υπόλοιπο r.
- Έχει ενδιαφέρον ότι θα μπορούσαμε να εξειδικεύσουμε τον παραπάνω ορισμό:
  - Αντί για νόρμες μπορούμε να χρησιμοποιήσουμε απόλυτες τιμές και ανισότητες στοιχείο προς στοιχείο (componentwise).
  - Το είδος των διαταραχών μπορεί να προκαθοριστεί. Για παράδειγμα, να δεχόμαστε ΔA που να έχουν την ίδια δομή με το A, κάτι που είναι πολύ σημαντικό όταν αναλύουμε την ευστάθεια προβλημάτων με ειδικά δεδομένα, π.χ. αραιά μητρώα, μητρώα Toeplitz, κ.λπ.

Απόδειξη για μια ελαφρώς πιο εύκολη περίπτωση *Θα προσπαθήσουμε να «αναθέσουμε» όλα τα σφάλματα του υπολογισμού στο A (καθόλου στο b).*
Αν θέσουμε $\hat{A} = A + \Delta A$ και $r := b - A\hat{x}$, τότε η σχέση (4) συνεπάγεται

$$\begin{aligned} b &= (A + \Delta A)\hat{x} \Rightarrow r + A\hat{x} = (A + \Delta A)\hat{x} \Rightarrow \\ r &= \Delta A\hat{x} \end{aligned}$$

Επομένως, απομένει να κατασκευάσουμε $\hat{A}$ που να ικανοποιεί την παραπάνω σχέση (4). Αυτό είναι εύκολο - για παράδειγμα δοκιμάστε το μητρώο 1ης τάξης

$$\Delta A := \frac{r\hat{x}^\top}{\hat{x}^\top\hat{x}}$$

Επίσης μπορείτε να αποδείξετε[3] ότι

$$\|\Delta A\|_2 = \frac{\|r\|_2}{\|\hat{x}\|_2}$$

και υπό την προϋπόθεση ότι χρησιμοποιούμε νόρμα 2,

$$\beta_N := \frac{\|\Delta A\|}{\|A\|} = \frac{\|r\|}{\|A\|\|\hat{x}\|}$$

Προσοχή: Επομένως το υπόλοιπο, r, καθορίζει το μέγεθος του πίσω σφάλματος.

[3] Από αυτά που γνωρίζετε στη Γραμμική Άλγεβρα.

Για «φανατικούς» Μπορεί επίσης να αποδειχτεί ότι διαταραχές που επιτυγχάνουν το ελάχιστο πίσω σφάλμα είναι

$$\Delta A = \frac{\|A\|\|\hat{x}\|}{\|A\|\|\hat{x}\| + \|b\|} r z^T, \Delta b = \frac{\|b\|}{\|A\|\|\hat{x}\| + \|b\|} r$$

όπου το z είναι ένα διάνυσμα (λέγεται «δυϊκό») για το οποίο ισχύει ότι

$$z^T \hat{x} = \|z\|_D \|\hat{x}\| = 1, \quad \text{όπου} \quad \|z\|_D = \max_{y \neq 0} \frac{z^T y}{\|y\|}$$

Σύνοψη ανάλυσης και πρακτικά φράγματα Είδαμε ότι μικρές αλλαγές $\Delta A, \Delta b$ στα δεδομένα του προβλήματος «επίλυση γραμμικού συστήματος $Ax = b$» οδηγούν στο παρακάτω φράγμα για το εμπρός σχετικό σφάλμα:

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

Επίσης, αν εφαρμόσουμε έναν αλγόριθμο επίλυσης το (εκ των υστέρων) πίσω σφάλμα ορίζεται

$$\beta_N \quad = \quad \inf\{\omega | (A + \Delta A)\hat{x} = b + \Delta b, \|\Delta A\| \leq \omega \|A\|, \|\Delta b\| \leq \omega \|b\|\}$$

και

$$\beta_N = \frac{\|r\|}{\|A\|\|\hat{x}\| + \|b\|}$$

Επομένως, αν γνωρίζουμε τα $\kappa(A), \hat{x}, r$ μπορούμε να εκτιμήσουμε ένα άνω φράγμα για το εμπρός σφάλμα ως εξής:

- Υπολογίζουμε το $\kappa(A)$,
- υπολογίζουμε το $\beta_N = \frac{\|r\|}{\|A\|\|\hat{x}\| + \|b\|}$,
- οπότε, εφόσον $\beta_N \kappa(A) < 1$,

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{2\beta_N \kappa(A)}{1 - \kappa(A)\beta_N}$$

Παραδείγματα Δεξιό μέλος $b = A * \texttt{ones}(n, 1)$. Ως `ferr2` είναι το παραπάνω φράγμα. Λύση μέσω $\backslash$ σε MATLAB 7.5.

Μητρώο	n	$\kappa_2(A)$	β_N	$\|x - \tilde{x}\|_2 / \|x\|_2$	`ferr2`
H	10	1.6e+13	5.0804e-017	2.7571e-004	0.0016
V	10	1.5e+07	3.6797e-017	3.3080e-010	1.1181e-009
R	100	1.7737e+003	2.2492e-016	5.2216e-014	7.9788e-013
Rn	100	684	5.0267e-016	1.3761e-014	6.8694e-013
D	100	1.0e+10	0	0	0
G	60	26.8	0.0156	0.4714	1.4317

H : hilb(n) - μεγάλο $\kappa(A)$, μικρό πίσω σφάλμα

V : vander(linspace(0, 1, 10)) - μεγάλο $\kappa(A)$, μικρό πίσω σφάλμα

R : rand(n) - μέτριο $\kappa(A)$, μικρό πίσω σφάλμα

Rn : randn(n) - μέτριο $\kappa(A)$, μικρό πίσω σφάλμα

D : diag([1e − 10, ones(1, n − 1)]), μεγάλο $\kappa(A)$, 0 πίσω σφάλμα, κακή εκτίμηση σφάλματος

G : gfpp(n) - μικρό $\kappa(A)$, μεγάλο πίσω σφάλμα

Παρατηρήσεις Τα παραπάνω επαληθεύουν ότι:

- Η επίλυση, τουλάχιστον όπως γίνεται στη MATLAB , είναι πίσω ευσταθής.
- ... στη MATLAB συνήθως χρησιμοποιείται απαλοιφή Gauss με μερική οδήγηση.
- ... όμως το τι υλοποιεί ο τελεστής «\», δηλ. η συνάρτηση `mldivide`, εξαρτάται από το όρισμα. Δείτε τις εκτενείς πληροφορίες στη MATLAB (`doc mldivide`).
- Αν λοιπόν ο αλγόριθμος για τα δεδομένα είναι πίσω ευσταθής, τότε αν γνωρίζουμε το A με σχετική ακρίβεια t δεκαδικών ψηφίων (π.χ. στην αριθμητική διπλής ακρίβειας IEEE το μέγιστο t θα είναι περίπου $\log_{10} \mathbf{u} \approx 16$ και $\kappa(A) \approx 10^p$, τότε η λύση του $Ax = b$ με «πίσω ευσταθή αλγόριθμο» αναμένεται να έχει περί τα $16 - t$ σωστά ψηφία.

Προσοχή - Μερικές αδυναμίες των παραπάνω:

- Η δυσκολία υπολογισμού του δείκτη κατάστασης ($O(n^3)$), πιο ακριβό από το ίδιο το πρόβλημα!
- ... ανάγκη μεθόδων εκτίμησής του (π.χ. αλγόριθμος `condest` της MATLAB).
- Η χρήση διαφορετικών νορμών δεν επιφέρει σοβαρές αλλαγές στις τιμές (όλες οι νόρμες ισοδύναμες).
- Γενικά, μια αδυναμία των παραπάνω είναι ότι βασίζονται σε πολύ συνοπτικές πληροφορίες, όπως εκφράζονται από τις νόρμες.
- Οι κλασικοί δείκτες κατάστασης αλλάζουν αν κλιμακώσουμε τους όρους του μητρώου: $\kappa(AD^{-1}) \leq \kappa(A)\|D^{-1}\|\|D\|$
- ... μια καλύτερη μετρική αποκαλείται δείκτης κατάστασης κατά Skeel
- ... είναι $\kappa_{\text{Skeel}}(A) := \||A^{-1}|\ |A|\|_\infty$.
- Ισχύει ότι $\kappa_{\text{Skeel}}(A) \leq \kappa(A)$ για άλλους δείκτες.
- ... επίσης $\kappa_{\text{Skeel}}(D^{-1}A) = \||A^{-1}D||D^{-1}A|\| = \||A^{-1}||A|\| = \kappa_{\text{Skeel}}(A)$.

- … Πολύ πιο εκτενής πληροφορία μας παρέχεται με επέκταση των παραπάνω αναλύσεων και μετρικών «ανά στοιχείο» (componentwise).

- Στην πράξη, χρησιμοποιούνται φράγματα που προέρχονται από «ανά στοιχείο» ανάλυση.

Πρακτικά φράγματα σφάλματος στην LAPACK[4]

Εκτίμηση δείκτη κατάστασης με ειδικούς αλγόριθμους (Hager, Higham) κόστους μικρότερου από το σχηματισμό των A^{-1}, $\|A\|, \|A^{-1}\|$.

Δείκτης πίσω σφάλματος

$$\omega_{A,b}(\hat{x}) = \max_i \frac{|r|_i}{(|A||\hat{x}| + |b|)_i}$$

Δείκτης εμπρός σφάλματος

$$\frac{\|x-\hat{x}\|}{\|\hat{x}\|} \leq \frac{\||A^{-1}|(|A||\hat{x}| + |b|)\|_\infty}{\|\hat{x}\|_\infty}\omega_{A,b}(\hat{x})$$

αλλά ως ακόμα πιο ακριβής δείκτης για το εμπρός σφάλμα χρησιμοποιείται το

$$\texttt{ferr} = \frac{\||A^{-1}|(|\hat{r}| + \gamma_{n+1}(|A||\hat{x}| + |b|))\|_\infty}{\|\hat{x}\|_\infty}$$