ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

Εξέταση Προόδου (Δεκ. 2004): ΑΠΑΝΤΗΣΕΙΣ

I. *(25 β.)*

1. Πότε λέμε ότι ένα σύστημα αριθμητικής κινητής υποδιαστολής ικανοποιεί τη συνθήκη ακριβούς στρογγύλευσης;
2. Να επιχειρηματολογήσετε, με βάση τις τιμές των $\Omega, \Phi_{\min}$ για να δείξετε ότι η παραγοντοποίηση LU ενός τετραγωνικού μητρώου παρέχει τη δυνατότητα καλής απόδοσης σε συστήματα με ιεραρχική μνήμη (1-2 συνοπτικές προτάσεις αρκούν).

Απάντηση.

1. Ένα σύστημα α.κ.υ. ικανοποιεί τη συνθήκη ακριβούς στρογγύλευσης όταν το αποτέλεσμα οποιασδήποτε από τις στοιχειώδεις πράξεις της αριθμητικής στο σύστημα και σε δεδομένα x, y που είναι α.κ.υ. είναι ο α.κ.υ. που θα προέκυπτε αν εκτελούσαμε την πράξη με αριθμητική άπειρης ακρίβειας και μετά στρογγυλεύαμε. Ειδικότερα, $\mathrm{fl}(x \odot y) = x \tilde{\odot} y$ όπου $\tilde{\odot}$ είναι η υλοποίηση της πράξης $\odot$ στο σύστημα α.κ.υ.
2. Η παραγοντοποίηση $A = LU$ αφορά στον υπολογισμό των παραγόντων L, U από το A. Επομένως απαιτούνται τουλάχιστον $\Phi_{\min} = 2n^2$ load-store. Επίσης, είναι γνωστό ότι η παραγοντοποίηση LU απαιτεί περίπου $2n^3/3$ πράξεις α.κ.υ. Επομένως $\mu_{\min} = O(1/n)$ όπως και στα BLAS-3, επομένως φαίνεται ότι υπάρχει τοπικότητα (μένει να γίνουν οι κατάλληλες υλοποιήσεις!).

□

II. *(25 β.)* Να επιλέξετε ποια από τα παρακάτω επιστρέφουν λογικό αληθές (1) και ποια όχι (0) και να σχολιάσετε τις απαντήσεις σας. Θεωρούμε ότι χρησιμοποιείται αριθμητική κινητής υποδιαστολής IEEE διπλής ακρίβειας (με βαθμιαία υποκανονικοποίηση). Το `realmax` συμβολίζει το μέγιστο πεπερασμένο θετικό αριθμό κινητής υποδιαστολής, `realmin` το μικρότερο κανονικοποιημένο θετικό α.κ.υ. και $==$ είναι το σύμβολο λογικής ισότητας:

α) $\mathtt{realmax} + 1.0 == \infty$. β) $\mathtt{Inf/Inf} == \mathtt{NaN}$. γ) $\mathtt{realmin}/2 == 0$. δ) $0/0 == \mathtt{Inf}$.

Απάντηση.

α) Λάθος (0): Ο εκθέτης του `realmax` είναι 1023. Για να γίνει η πρόσθεση με το $y = 1.0$, θα πρέπει να διαιρέσουμε το y έτσι ώστε ο εκθέτης του να γίνει 1023. Κάθε διαίρεση με το 2 αντιστοιχεί σε ολίσθηση του πιο σημαντικού bit της ουράς του y κατά μια θέση προς τα δεξιά. Το μήκος της ουράς στην αριθμητική IEEE διπλής ακρίβειας είναι 52, επομένως μετά από 53 ολισθήσεις προς τα δεξιά, η ουρά θα μηδενιστεί. Επομένως $\mathtt{realmax}+1.0 == \mathtt{realmax}$.

β) Σωστό (1): Το αποτέλεσμα της διαίρεσης δεν ορίζεται.

γ) Λάθος (0): Το σύστημα υποστηρίζει βαθμιαία υποκανονικοποίηση. Το αποτέλεσμα της διαίρεσης θα είναι μη κανονικοποιημένος και μη μηδενικός αριθμός κ.υ.

δ) Λάθος (0): Το αποτέλεσμα της διαίρεσης $0/0$ δεν ορίζεται.

□

III. *(25 β.)* Χρησιμοποιώντας *εμπρός ανάλυση σφάλματος*, να δείξετε ότι ο παρακάτω αλγόριθμος

```
n = 6; s=0; for j = 1:n, s = s+ x(j)*x(j); end;
```

υπολογίζει το `s` με μικρό σχετικό σφάλμα. Ειδικότερα, να δείξετε όλα τα βήματα της ανάλυσης και να υπολογίσετε (μικρό) φράγμα για το εμπρός σχετικό σφάλμα συναρτήσει της μονάδας στρογγύλευσης, $\mathbf{u}$, του συστήματος. Μπορείτε να υποθέσετε ότι κάθε στοιχείο `x(j)` είναι αριθμός κινητής υποδιαστολής και ότι $\mathtt{x(j)} * \mathtt{x(j)} < \mathtt{Inf}$, όπου `Inf` είναι το «άπειρο» για το χρησιμοποιούμενο σύστημα α.κ.υ.

Απάντηση. Χρησιμοποιώντας τις γνωστές σχέσεις διάδοσης σφάλματος, και θέτοντας για ευκολία $\mathtt{x(j)} \equiv \xi_j$ και $\sigma = \sum_{j=1}^{6} x_j^2$ και $\hat{\sigma}$ το s που υπολογίζεται με α.κ.υ. από τον παραπάνω κώδικα, έχουμε:

$$\begin{aligned}
\hat{\sigma} &= (((((\xi_1^2 < 1 > +\xi_2^2 < 1 >) < 1 > +\xi_3^2 < 1 >) < 1 > +\xi_4^2 < 1 >) < 1 > +\xi_5^2 < 1 >) < 1 > +\xi_5^2 < 1 >) < 1 > \\
&= \xi_1^2 < 6 > +\xi_2^2 < 6 > +\xi_3^2 < 5 > +\xi_4^2 < 4 > +\xi_5^2 < 3 > +\xi_6^2 < 2 > \\
&= \xi_1^2(1+\theta_6) + \xi_2^2(1+\hat{\theta}_6) + \xi_3^2(1+\theta_5) + \xi_4^2(1+\theta_4) + \xi_5^2(1+\theta_3) + \xi_6^2(1+\theta_2)
\end{aligned}$$

Επομένως,

$$\begin{aligned}
|\hat{\sigma} - \sigma| &= |\xi_1^2\theta_6 + \xi_2^2\hat{\theta}_6 + \xi_3^2\theta_5 + \xi_4^2\theta_4 + \xi_5^2\theta_3 + \xi_6^2\theta_2| \\
&\le |\xi_1^2||\theta_6| + |\xi_2^2||\hat{\theta}_6| + |\xi_3^2||\theta_5||\xi_4^2||\theta_4| + |\xi_5^2||\theta_3| + |\xi_6^2||\theta_2| \\
&\le \xi_1^2\gamma_6 + \xi_2^2\gamma_6 + \xi_3^2\gamma_5 + \xi_4^2\gamma_4 + \xi_5^2\gamma_3 + \xi_6^2\gamma_2
\end{aligned}$$

όπου $\gamma_j = j\mathbf{u}/(1 - j\mathbf{u})$, επομένως, επειδή $\gamma_6 > \gamma_5, ..., \gamma_2$,

$$\frac{|\hat{\sigma} - \sigma|}{|\sigma|} \le \gamma_6 = 6\mathbf{u}/(1 - 6\mathbf{u}).$$

$\square$

V. *(25 β.)* (*Στο ερώτημα αυτό θα σας είναι χρήσιμος ο τύπος (Sherman-Morrison):*

$$(A + uv^\top)^{-1} = A^{-1} - A^{-1}\frac{uv^\top}{1 + v^\top A^{-1}u}A^{-1}$$

με κατάλληλα επιλεγμένα διανύσματα $u, v \in \mathbb{R}^n$.) Έστω ότι έχετε διαθέσιμους τους παράγοντες L και U της παραγοντοποίησης LU ενός γενικού μητρώου $A \in \mathbb{R}^{n\times n}$, δηλ. $A = LU$, και ότι γίνεται η εξής «ανανέωση»: Αντικαθίσταται η τελευταία στήλη του A με ένα άλλο διάνυσμα, έστω p. Αν γνωρίζετε ότι το νέο μητρώο $B = [A(:,1), ..., A(:, n-1), p]$ είναι επίσης αντιστρέψιμο, να περιγράψετε, με βήματα που αντιστοιχούν σε εντολές τύπου BLAS έναν αποδοτικό τρόπο επίλυσης του συστήματος $Bx = y$ που να κοστίζει $\Omega = \alpha_2 n^2 + \alpha_1 n + \alpha_0$ πράξεις, όπου οι συντελεστές α_1 και α_2 είναι σταθερές. Επίσης να βρείτε τους συντελεστές αυτούς. Για πλήρη βαθμό θα πρέπει να χρησιμοποιήσετε και εντολές τύπου BLAS-3 χωρίς να αυξήσετε ταυτόχρονα και το συνολικό κόστος α.κ.υ. Ω για το πρόγραμμα.

Προσοχή:Αρκεί να περιγράψετε με σαφήνεια τι εκτελείται σε κάθε βήμα, σε τι δεδομένα, σε ποιά κατηγορία BLAS ανήκει. Επίσης να αναλύσετε τα κόστη. Δεν χρειάζεται να χρησιμοποιήσετε την ακριβή ονοματολογία των BLAS. Για παράδειγμα, αν είχατε να περιγράψετε τα βήματα για τον υπολογισμό του $B = B + (Ax)y^\top$, όπου τα $x, y \in \mathbb{R}^n$ και $A, B \in \mathbb{R}^{n\times n}$, θα αρκούσε ο παρακάτω πίνακας (τα ονόματα `GEMV`, `GER` δεν είναι απαραίτητα):

	Υπολογισμός	Περιγραφή	Κατηγορία	
1.	$q = Ax$	Πολλαπλασιασμός μητρώο-διάνυσμα (ή `GEMV`)	BLAS-2	$n(2n-1)$
2.	$B = B + qy^\top$	Ανανέωση πρώτης τάξης (ή `GER`)	BLAS-2	$2n^2$
	Συνολικό κόστος			$\Omega = 4n^2 - n$

Απάντηση. Παρατηρούμε ότι $uv^\top = [0, ..., 0, p - A - (:, n)]$ (μητρώο τάξης 1) επομένως αρκεί να επιλέξουμε $u = p - A(:, n)$ και $v = e_n$.

	Υπολογισμός	Περιγραφή	Κατηγορία	Κόστος (όροι $O(n)$ και $O(n^2)$)
1.	$u = p - A(:, n)$	πράξη τύπου `sAXPY`	BLAS-1	n
2.	$L[z_u, z_b] = [u, b]$	Λύση κάτω τριγωνικού συστήματος με 1 στη διαγ. και 2 δεξιά μέλη (ή `TRSM`)	BLAS-3	$2n(n-1)$
3.	$U[z_u, z_b] = [z_u, z_b]$	Λύση άνω τριγωνικού με 2 δεξιά μέλη (ή `TRSM`)	BLAS-3	$2n^2$
4.	$x = z_b + (z_b(n)/(1 + z_u(n)))z_u$	`sAXPY`	BLAS-1	$2n$
	Συνολικό κόστος	$\Omega = 4n^2 + n + O(1)$	επομένως	$\alpha_2 = 4, \alpha_1 = 1$

όπου $z_u(n), z_b(n)$ είναι τα τελευταία στοιχεία των διανυσμάτων z_u και z_b αντίστοιχα (δηλ. βαθμωτοί). Προσέξτε ότι στο παραπάνω αξιοποιήσαμε το ότι $v = e_n$ οπότε η πράξη $v^\top[z_u, z_b]$ (διάνυσμα γραμμή επί μητρώο ή 2 εσωτερικά γινόμενα) είναι τετριμμένη (δεν χρειάζονται υπολογισμοί). Επίσης, αν εκτελέσουμε το 1ο βήμα με απευθείας κλήση στη BLAS-1 SAXPY, το κόστος θα είναι κατά n μεγαλύτερο, καθώς η κλασική SAXPY δεν ελέγχει αν ο βαθμωτός συντελεστής είναι ± 1 (μόνο τη μηδενική περίπτωση).

Ένας εναλλακτικός (και φθηνότερος!) τρόπος επίλυσης, χωρίς τη χρήση του τύπου, είναι ο ακόλουθος: Παρατηρούμε ότι $LU = A$ επομένως $LU(:,j) = A(:,j)$ για $j = 1, ..., n$. Επομένως, αν αλλάξουμε την τελευταία στήλη του A σε p, η παραγοντοποίηση LU του B θα έχει ως παράγοντες το ίδιο L και ως U το $\tilde{U} = [U(:,1), ..., U(:,n-1), g]$, όπου οι πρώτες $n-1$ στήλες είναι όπως πριν και η τελευταία ικανοποιεί τη σχέση $Lg = p$. Επομένως, μπορούμε να υπολογίσουμε το g με μια λύση με το γνωστό κάτω τριγωνικό L. Επομένως, μπορούμε να κάνουμε τα εξής βήματα για να λύσουμε το $Bx = y$: 1) Λύνουμε τα $L[y, g] = [z, p]$ με BLAS-3 και κόστος $2n(n-1)$ (λύση τριγωνικού με 2 δεξιά μέλη και μονάδα στη διαγώνιο). 2) Θέτουμε $\tilde{U} = [U(:,1), ..., U(:,n-1), g]$, οπότε $B = L\tilde{U}$, και λύνουμε $\tilde{U}x = y$ ως προς x με BLAS-1 (λύση τριγωνικού συστήματος, κόστος n^2.) Επομένως το συνολικό κόστος θα είναι $\Omega = 3n^2 - 2n$. $\square$