

## ΑΠΑΝΤΗΣΕΙΣ ΕΠΙΣΤΗΜΟΝΙΚΟΥ ΥΠΟΛΟΓΙΣΜΟΥ Ι Σεπτέμβριος 2004

Να διαβάσετε προσεκτικά τις εκφωνήσεις. Για πλήρη αξιολόγηση του γραπτού σας πρέπει να παρουσιάσετε όλο σας τον συλλογισμό και όλα τα ενδιάμεσα αποτελέσματα. Έχετε 3 ώρες. Οι αλγόριθμοι πρέπει να περιγράφονται με σαφήνεια, π.χ. όπως στις σημειώσεις ή με *MATLAB*.

**I.** (50 β.) Να απαντήσετε στα παρακάτω ερωτήματα (στις ερωτήσεις Σωστό/Λάθος πρέπει να δικαιολογήσετε τις απαντήσεις για να βαθμολογηθείτε).

1) **6β** Ποιά είναι τα κύρια κριτήρια με βάση τα οποία αξιολογούνται τα εργαλεία του Επιστημονικού Υπολογισμού;

*Απάντηση.* Ακρίβεια, Ταχύτητα, Κόστος.  $\square$

2) **8β** Έστω ότι επιθυμούμε να υπολογίσουμε την παραγοντοποίηση  $LU$  του μητρώου  $A$  όπου με το συμβολισμό της *MATLAB*  $A = [4, 3, 1; 6, 2, 10; 0, -2, -4]$ . Να δείξετε ποιος θα είναι ο πρώτος «οδηγός» (δηλ. το στοιχείο που θα είναι στη θέση (1,1) του  $U$ ) στις παρακάτω περιπτώσεις: α) Αν δεν χρησιμοποιήσουμε οδήγηση; β) Αν χρησιμοποιήσουμε μερική οδήγηση; γ) Αν χρησιμοποιήσουμε πλήρη οδήγηση;

*Απάντηση.* Και στις τρεις περιπτώσεις το  $U(1, 1)$  θα είναι ο οδηγός του πρώτου βήματος. Επομένως:

α) το στοιχείο στη θέση (1, 1) του  $A$ , δηλ. 4. β) το μέγιστο σε απόλυτη τιμή στοιχείο της πρώτης στήλης του  $A$ , δηλ. 6. γ) Το μέγιστο σε απόλυτη τιμή στοιχείο του  $A$ , δηλ. 10. (Διευκρινίσεις: Υπήρχε και άλλη παραλλαγή της άσκησης με διαφορετικούς αριθμούς που λυνόταν με τον ίδιο τρόπο).  $\square$

3) **8β** Σωστό ή Λάθος: Αν το συμμετρικό μητρώο  $A \in \mathbb{R}^{n \times n}$  είναι θετικά ορισμένο το ίδιο ισχύει και για το αντίστροφό του  $B = A^{-1}$ .

*Απάντηση.* Παρουσιάζουμε 3 εναλλακτικές λύσεις. I)  $(A^{-1})^T = (A^T)^{-1} = A^{-1}$  επομένως το αντίστροφο είναι και αυτό συμμετρικό. Είναι και θετικά ορισμένο αφού  $x^T A x > 0 \Rightarrow \forall y : (A^{-1}y)^T A (A^{-1}y) > 0 \Rightarrow y^T A^{-1} y > 0$  επομένως και το  $A^{-1}$  είναι θετικά ορισμένο. II) Εναλλακτικά, μπορείτε να επικαλεστείτε α) το ότι ένα μητρώο είναι σθο αν και μόνον αν έχει αποκλειστικά θετικές ιδιοτιμές. β) Αν οι ιδιοτιμές του  $A$  είναι:  $\Lambda(A) := \{\lambda | Ax = \lambda x\}$  τότε οι ιδιοτιμές του  $A^{-1}$  είναι  $\Lambda(A^{-1}) := \{\mu | \mu^{-1} \in \Lambda(A)\}$ . Επομένως  $\Lambda(A^{-1}) \in \mathbb{R}^+$ . III) Εναλλακτικά: Επειδή το  $A$  σθο τότε υπάρχει παραγοντοποίηση Chole-

sky  $A = LL^T$ . Τότε  $A^{-1} = L^{-T}L^{-1}$ . Τότε για κάθε διάνυσμα  $x \neq 0$   $x^T A^{-1} x = x^T L^{-T} L^{-1} x = (L^{-1}x)^T (L^{-1}x) = \|L^{-1}x\|_2 > 0$  (θετικότητα νορμών).  $\square$

4) **7β** Σωστό ή Λάθος: Το «έψιλον της μηχανής» για την α.κ.υ. IEEE είναι αριθμός κινητής υποδιαστολής.

*Απάντηση.* Το έψιλον της μηχανής είναι η απόσταση του 1 από το αμέσως διαδοχικό μεγαλύτερο αριθμό κ.υ. Στην αριθμητική κινητής υποδιαστολής IEEE έχουμε ότι  $1.0 \dots 01$  όπου η ουρά έχει  $t = 23$  ή 52 δυαδικά ψηφία (και ένα κρυμμένο), και  $\epsilon = 2^{1-t}$ . Ο αριθμός αυτός μπορεί να αναπαρασταθεί ως  $1.0 \times 2^{-t}$  καθώς οι παραπάνω τιμές του  $t$  είναι μέσα στο διάστημα  $[e_{\min}, e_{\max}]$  των άκρων του εκθέτη.

*(Διευκρινίσεις: Σε παλαιότερες εκδόσεις του βιβλίου, δινόταν μεγαλύτερο βάρος στον ορισμό που έλεγε ότι έψιλον της μηχανής είναι ο ελάχιστος θετικός α.κ.υ. για τον οποίο ισχύει ότι  $\text{fl}(1 + \epsilon) > 1$ . Αν χρησιμοποιούσατε αυτόν τον ορισμό, δεν χρειαζόταν περαιτέρω απόδειξη.)*  $\square$

5) **7β** Στην αριθμητική επίλυση συστημάτων συνήθων διαφορικών εξισώσεων αρχικών τιμών  $\frac{du}{dt}(t) = F(u, t)$ , να περιγράψετε ένα μειονέκτημα και ένα πλεονέκτημα α) για τις άμεσες μεθόδους· β) για τις έμμεσες μεθόδους.

*Απάντηση.* α) Οι άμεσες μέθοδοι υλοποιούνται εύκολα (σε κάθε βήμα ο υπολογισμός της νέας προσέγγισης του  $U$  προκύπτει εύκολα από το  $F$  σε γνωστές τιμές του  $U(t)$ ) αλλά συνήθως επιβάλουν σημαντικό περιορισμό στο μέγεθος του βήματος  $\Delta t$  για να υπάρχει ευστάθεια. β) Οι έμμεσες μέθοδοι δεν επιβάλουν περιορισμό στο βήμα για ευστάθεια αλλά η υλοποίησή τους είναι πιο πολύπλοκη καθώς σε κάθε βήμα η τιμή του  $U$  δίδεται μόνον έμμεσα (π.χ. από τη λύση ενός γραμμικού συστήματος).

*Διευκρινίσεις: Προσοχή, η αστάθεια που μπορεί να εμφανιστεί αν το  $\Delta t$  είναι μεγάλο στις άμεσες μεθόδους δεν συνεπάγεται ότι οι άμεσες μέθοδοι είναι λιγότερο ακριβείς! Μάλιστα αν το  $\Delta t$  είναι αρκετά μικρό, οι άμεσες μέθοδοι μπορεί να είναι πιο ακριβείς. Επίσης, το  $\Delta t$  μπορεί να αλλάζει σε κάθε βήμα και στις έμμεσες μεθόδους, μόνον που αυτό έχει μεγαλύτερο κόστος γιατί μπορεί να μας αναγκάσει να ξαναπαραγοντοποιήσουμε το μητρώο. Επίσης, το μικρότερο κόστος των άμεσων μεθόδων έχει να κάνει με το κόστος ανά βήμα του υπολογισμού - καθώς «βηματίζουμε στο χρόνο» - και όχι με το συνολικό κόστος της μεθόδου σε*

σχέση με τις έμμεσες. Η απόφαση για μέθοδο είναι προτιμότερη είναι θέμα που απασχολεί τους ερευνητές ανάλογα με το πρόβλημα κ.λπ.  $\square$

6α) **7β** Να αποδείξετε ότι οι διανυσματικές νόρμες  $\|\cdot\|_2$  και  $\|\cdot\|_\infty$  είναι ισοδύναμες στο χώρο  $\mathbb{R}^n$ .  
6β) Το ίδιο αλλά για τις νόρμες  $\|\cdot\|_1$  και  $\|\cdot\|_\infty$ .

*Απάντηση.* Και στις δυο εκδοχές αρκεί να αποδείξετε ότι υπάρχουν σταθερές, έστω  $\kappa_1, \kappa_2$  για τις οποίες  $\kappa_1\|x\|_\alpha \leq \|x\|_\beta \leq \kappa_2\|x\|_\alpha$  όπου οι δείκτες  $\alpha, \beta$  είναι 2,  $\infty$  στην πρώτη εκδοχή και 1,  $\infty$  στη δεύτερη. α)  
Έστω ότι  $x \in \mathbb{R}^n$ . Τότε

$$\|x\|_2 = \left( \sum_{j=1}^n \xi_j^2 \right)^{1/2} \geq \max_j |\xi_j| = \|x\|_\infty.$$

Επίσης

$$\|x\|_2 = \left( \sum_{j=1}^n \xi_j^2 \right)^{1/2} \leq \left( \sum_{j=1}^n \max_j |\xi_j| \right)^{1/2} = \sqrt{n}\|x\|_\infty.$$

Δείξαμε επομένως ότι  $\forall x \in \mathbb{R}^n$  ισχύει ότι

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

επομένως οι νόρμες είναι ισοδύναμες με συντελεστές 1 και  $\sqrt{n}$ . β) Στη δεύτερη εκδοχή, όπως πριν είναι προφανές ότι  $\|x\|_\infty \leq \|x\|_1$ . Επίσης,

$$\|x\|_1 = \sum_{j=1}^n |\xi_j| \leq \sum_{j=1}^n \max_{j=1}^n |\xi_j| = n\|x\|_\infty,$$

επομένως οι νόρμες είναι ισοδύναμες με συντελεστές 1 και  $n$ .  $\square$

7) **7β** Σωστό ή Λάθος: Το γινόμενο τριδιαγώνιου μητρώου με κάτω διδιαγώνιο μητρώο είναι τριδιαγώνιο.

*Απάντηση.* Λάθος: Αποδεικνύεται με απλό (αντι)παράδειγμα τυχαίων μητρώων του τύπου που προαναφερθηκαν δείχνοντας ότι το στοιχείο στη θέση  $(3, 1)$  δεν θα είναι μηδενικό μια και θα περιέχει τον όρο  $\alpha_{32}\beta_{21}$  που γενικά δεν είναι 0.  $\square$

**II.** (25 β.) (Σημ.: Θεωρούμε ότι όλα τα στοιχεία εισόδου είναι αριθμοί κινητής υποδιαστολής.)

1. **12β** Να δείξετε ότι η μέθοδος Horner ( $s = \alpha_n$ ; for  $j = n - 1 : -1 : 0, s = sx + \alpha_j$ ; end) για τον υπολογισμό της τιμής ενός πολυωνύμου  $p(x) = \sum_{j=0}^n \alpha_j x^j$  που δίνεται μέσω δυναμομορφής είναι πίσω ευσταθής.

*Απάντηση.* Πίσω ευστάθεια σημαίνει ότι το αποτέλεσμα που υπολογίστηκε με τον κανόνα Horner από τα δεδομένα  $Q = [\alpha_0, \dots, \alpha_n, x]$ , έστω δηλ. το υπολογισμένο αποτέλεσμα  $S = H_{\text{prog}}(X)$  όπου  $H_{\text{prog}}$  η υλοποίηση του αλγόριθμου, μπορεί να υπολογιστεί με τον ίδιο αλγόριθμο σε αριθμητική άπειρης ακρίβειας αλλά με δεδομένα  $X_{\text{prog}}$  που μπορεί να είναι διαφορετικά. Το πόσο διαφορετικά καθορίζει και το βαθμό της πίσω ευστάθειας. Εδώ αρκεί να δείξουμε ότι υπάρχουν δεδομένα για τα οποία ισχύει αυτό που έχουν μικρή σχετική απόσταση από τα ακριβή δεδομένα. Χρησιμοποιώντας τους κανόνες που μάθαμε σχετικά με τη διάδοση του σφάλματος, μπορούμε να γράψουμε ότι:

$$\begin{aligned} s_n &= \alpha_n \\ s_{n-1} &= (\alpha_n x \langle 1 \rangle + \alpha_{n-1}) \langle 1 \rangle = \alpha_n x \langle 2 \rangle + \alpha_{n-1} \langle 1 \rangle \end{aligned}$$

όπου το σύμβολο  $\langle k \rangle := \prod_{j=1}^k (1 + \delta_j)$  για κάποια  $\delta_j$ . Συνεχίζοντας αναδρομικά

$$\begin{aligned} s_{n-2} &= s_{n-1} x \langle 2 \rangle + \alpha_{n-2} \langle 1 \rangle = \alpha_n x^2 \langle 4 \rangle + \alpha_{n-1} x \langle 3 \rangle + \alpha_{n-2} \langle 1 \rangle \\ s_{n-3} &= \alpha_n x^3 \langle 6 \rangle + \alpha_{n-1} x^2 \langle 5 \rangle + \alpha_{n-2} x \langle 3 \rangle + \alpha_{n-3} \langle 1 \rangle \\ &\vdots \\ &\vdots \\ s_0 &= \alpha_n \langle 2n \rangle x^n + \alpha_{n-1} \langle 2n - 1 \rangle x^{n-1} + \end{aligned}$$

$$\alpha_{n-2} \langle 2n-3 \rangle x^{n-2} + \dots + \alpha_1 \langle 3 \rangle + \alpha_0 \langle 1 \rangle$$

Επομένως αν χρησιμοποιούσαμε τον αλγόριθμο Horner με αριθμητική άπειρης ακρίβειας στο πολώνυμο με συντελεστές  $\tilde{\alpha}_0 = \alpha_0(1 + \delta_1^{(0)})$ ,  $\tilde{\alpha}_1 = \alpha_1(1 + \delta_1^{(1)})(1 + \delta_2^{(1)})(1 + \delta_3^{(1)})$ , ...,  $\tilde{\alpha}_{n-1} = \alpha_{n-1}(1 + \delta_1^{(n-1)})(1 + \delta_2^{(n-1)}) \dots (1 + \delta_{2n-1}^{(n-1)})$ ,  $\tilde{\alpha}_n = \alpha_n(1 + \delta_1^{(n)})(1 + \delta_2^{(n)}) \dots (1 + \delta_{2n}^{(n)})$ , θα είχαμε το επιθυμητό αποτέλεσμα. Στη συνέχεια, χρησιμοποιούμε ως διάνυσμα  $X_{\text{prog}}$  εκείνο που αποτελείται από τα στοιχεία  $X_{\text{prog}} = [\tilde{\alpha}_0, \dots, \tilde{\alpha}_n, x]$ . Να προσέξετε ότι το  $x$  παραμένει ως έχει (η πίσω ευστάθεια δεν αποδεικνύεται άμεσα μέσω αλλαγών στο  $x$ ). Επομένως, η μέθοδος είναι πίσω ευσταθής αν μπορούμε να αποδείξουμε ότι τα στοιχεία  $\tilde{\alpha}_j$  είναι αρκετά κοντά στα ακριβή στοιχεία  $\alpha_j$ . Έχουμε ότι για  $j = 0, \dots, n-1$  ισχύει

$$\frac{|\tilde{\alpha}_j - \alpha_j|}{|\alpha_j|} \leq \frac{|\alpha_j(1 + \theta_{2j-1}^{(j)}) - \alpha_j|}{|\alpha_j|} = |\theta_{2j-1}^{(j)}| \leq \gamma_{2j-1} = \frac{(2j-1)\mathbf{u}}{1 - (2j-1)\mathbf{u}}$$

όπου  $\mathbf{u}$  είναι η μονάδα στρογγύλευσης, καθώς επίσης ότι

$$\frac{|\tilde{\alpha}_n - \alpha_n|}{|\alpha_n|} \leq \frac{|\alpha_n(1 + \theta_{2n}^{(n)}) - \alpha_n|}{|\alpha_n|} = |\theta_{2n}^{(n)}| \leq \gamma_{2n} = \frac{(2n)\mathbf{u}}{1 - (2n)\mathbf{u}}$$

και τα αριστερά μέλη είναι μικροί αριθμοί (εφόσον το  $n$  δεν είναι πάρα πολύ μεγάλο).  $\square$

## 2. 13β Δίδεται ο απλός κώδικας για τον υπολογισμό της νόρμας-1 διανύσματος

```
s = abs(x(1)); for j = 2:length(x), s = s + abs(x(j)); end; M = s;
```

α) Να δείξετε ένα διάνυσμα με στοιχεία που είναι αριθμοί κινητής υποδιαστολής για το οποίο το σχετικό σφάλμα της υπολογισμένης νόρμας, έστω  $\tilde{M}$ , με τον παραπάνω κώδικα σε αριθμητική IEEE διπλής ακρίβειας να είναι  $\frac{|M - \tilde{M}|}{|\tilde{M}|} \approx 10^{-8}$ . (Υπόδειξη: Η διάσταση του διανύσματος μπορεί να είναι πολύ μεγάλη). β) Να περιγράψετε πώς θα μπορούσε να τροποποιηθεί ο κώδικας ώστε να βελτιωθεί το σφάλμα. γ) Να σχολιάσετε τις επιπτώσεις της τροποποίησης ως προς τα κριτήρια αξιολόγησης εργαλείων επιστημονικού υπολογισμού.

*Απάντηση.* Μπορούμε εύκολα να κατασκευάσουμε διάνυσμα που επιφέρει τέτοιο σφάλμα με βάση τη

σχέση  $\text{fl}(M + m) = M$  για πολλές περιπτώσεις θετικών στοιχείων  $M \gg m$ . Το πιο προφανές παράδειγμα μπορεί να κατασκευαστεί με βάση τη σχέση  $\text{fl}(1 + \epsilon/2) = 1$  (όπου  $\epsilon$  είναι το έψιλον της μηχανής) που επιφέρει σχετικό σφάλμα  $\frac{\epsilon/2}{1+\epsilon/2}$ . Επομένως αρκεί να λάβουμε ένα διάνυσμα μεγέθους  $n + 1$  με στοιχεία

$$[1, \epsilon/2, \dots, \epsilon/2]$$

οπότε η υπολογισμένη νόρμα θα είναι  $\tilde{M} = 1$  ενώ η ακριβής τιμή  $M = 1 + n\epsilon/2$ . Επομένως, το σχετικό σφάλμα θα είναι

$$\frac{|M - \tilde{M}|}{|M|} = \frac{n\epsilon/2}{1 + n\epsilon/2}$$

οπότε μπορούμε να επιλέξουμε το  $n$  ώστε να ισχύει ότι το σφάλμα είναι όσο θέλουμε. Π.χ. αν

$$\frac{n\epsilon/2}{1 + n\epsilon/2} = 10^{-8} \Rightarrow 10^{-8}(1 + n\epsilon/2) = n\epsilon/2$$

και λύνοντας ως προς  $n$  θα έχουμε

$$n = \left\lceil \frac{2 \times 10^{-8}}{\epsilon - 10^{-8}\epsilon} \right\rceil$$

ώστε να είναι και ακέραιος. Για αριθμητική IEEE διπλής ακρίβειας,  $\epsilon \approx 10^{-16}$  οπότε

$$n = \left\lceil \frac{2 \times 10^{-8}}{10^{-16} - 10^{-8}10^{-16}} \right\rceil \approx 2 \times 10^8$$

Για να βελτιώσουμε το σφάλμα στον υπολογισμό της νόρμας, συνίσταται, πριν από τον υπολογισμό της να επαναριθμήσουμε τα στοιχεία του διανύσματος ώστε οι απόλυτες τιμές του να είναι αύξουσες. Το αποτέλεσμα θα είναι πιο ακριβές αλλά το κόστος μεγαλύτερο καθώς πρέπει να εφαρμόσουμε αλγόριθμο

διάταξης  $n$  στοιχείων το κόστος των οποίων είναι τουλάχιστον  $O(n \log n)$ . Σε σύγκριση, το κόστος της άθροισης είναι μόνον  $O(n)$ . Μια υλοποίηση είναι η ακόλουθη:

```
function [nmx] = Smynorm(x);  
  
y = mysort(abs(x));  
  
s = y(1); for j = 2:length(y), s = s+ y(j); end; nmx = s;
```

όπου `mysort` είναι κάποιος αλγόριθμος διάταξης. Θα μπορούσε να είναι ο αλγόριθμος `sort` της `MATLAB` ή κάποιος άλλος αλγόριθμος. Σε κάθε περίπτωση, σε σχέση με τα κριτήρια του Επιστημονικού Υπολογισμού, βελτιώσαμε την ακρίβεια αλλά μειώσαμε την ταχύτητα του υπολογισμού.

*Διευκρίνιση: Αν επιχειρήσετε να χρησιμοποιήσετε τη συνάρτηση `MATLAB` `norm(x,1)` στο διάνυσμα πριν τη διάταξη, το αποτέλεσμα θα είναι πάλι 1. Ενδεχομένως υπολογισμένο πιο γρήγορα, καθώς πρόκειται για εσωτερική συνάρτηση της `MATLAB`, όμως πάντα με το σφάλμα που περιγράψαμε. Επομένως δεν είναι αποδεκτή ως απάντηση καθώς η βελτίωση που ζητούσαμε αφορούσε το σφάλμα που θα προέκυπτε από ένα διάνυσμα σαν αυτό που παρουσιάστηκε στο μέρος (α). □*

## II. (25 β.)

1. **12β** Έστω ένα συμμετρικό θετικά ορισμένο τριδιαγώνιο μητρώο  $A$  που ορίζεται από τα διανύσματα (μήκους  $n$ )  $d, e$  έτσι ώστε το  $d$  να είναι η διαγώνιος του  $A$  και τα στοιχεία 2 έως  $n$  του  $e$  να είναι στην υπερ- και υποδιαγώνιο του  $A$ . Να γράψετε (π.χ. σε `MATLAB`) τον αλγόριθμο που αντιστοιχεί στην παραγοντοποίηση Cholesky του  $A$ . Το πρόγραμμά σας πρέπει να δέχεται ως στοιχεία εισόδου τα  $d, e$  και να επιστρέφει τα διανύσματα  $g, h$  όπου το  $g$  περιέχει τη διαγώνιο και το  $h$  στις θέσεις 2 έως  $n$  την υποδιαγώνιο του παράγοντα Cholesky  $L$ . Να ονομάσετε το πρόγραμμά σας `CholTridFact`, οπότε αν χρησιμοποιείτε `MATLAB` η κλήση στη συνάρτηση θα είναι `[g,h] = CholTridFact(d,e)`.

*Απάντηση.* Πρόκειται για τον αλγόριθμο Cholesky εξειδικευμένο για τριδιαγώνια μητρώα:

```

function [g,h] = Cholesky(d,e)

n = length(d); g = zeros(n,1); h = zeros(n,1);

g(1) = sqrt(d(1));

for i = 2:n

    h(i) = e(i)/g(i-1);

    g(i) = sqrt(d(i)-h(i)*h(i));

end
□

```

2. **13β** Έστω ότι έχετε στη διάθεσή μας και ένα πρόγραμμα (π.χ. ως συνάρτηση MATLAB) που ονομάζουμε `CholTridSol(g,h,b)`, το οποίο, με είσοδο τα διανύσματα  $g, h$  που ορίζουν τον παράγοντα Cholesky  $L$  και ένα διάνυσμα  $b$ , επιστρέφει τη λύση  $x$  του συστήματος  $LL^T x = b$ . Να σχεδιάσετε αλγόριθμο για την αριθμητική επίλυση της διαφορικής εξίσωσης  $-\frac{d^2 u}{dx^2}(x) + u(x) = 2x \sin x - 2 \cos x$  στο διάστημα  $[0, \pi]$  όπου  $u(0) = u(\pi) = 0$ . Για τη λύση, να χρησιμοποιήσετε πλέγμα  $n$  ισοκαταναμημένων σημείων  $0 < x_1 < \dots < x_{n-2} < \pi$ . Συγκεκριμένα, να περιγράψετε (εξηγώντας τα βήματά σας) τον αλγόριθμο που α) δημιουργεί το πλέγμα και το διακριτό σύστημα που αντιστοιχεί στη διαφορική εξίσωση με κεντρισμένες πεπερασμένες διαφορές 2ης τάξης και β) επιστρέφει το διάνυσμα  $U$  που περιέχει στις θέσεις 2 έως  $n-1$  την προσέγγιση για τη λύση  $u(x)$  στους κόμβους του πλέγματος. Να χρησιμοποιήσετε, τις `CholTridFact` και `CholTridSol`. Να εξηγήσετε επίσης γιατί μπορείτε να τις χρησιμοποιήσετε και να αιτιολογήσετε τη χρήση τους συγκριτικά με την επίλυση μέσω  $LU$  (να υποθέσετε ότι τα κόστη των πράξεων  $\pm, \times, /, \sqrt{\phantom{x}}$  είναι ίδια).

*Απάντηση.* Διακριτοποιούμε το πλέγμα με κόμβους στις θέσεις  $x_j = j * h, j = 0, \dots, n-1$  όπου  $h = \pi/(n-1)$ . Με κεντρισμένες πεπερασμένες διαφορές 2ης τάξης, η εξίσωση για την προσέγγιση  $U_j$  της λύσης  $u$  σε κάθε κόμβο  $x_j$  έχει τη μορφή:

$$\begin{aligned}
 -\frac{U_{j-1} - 2U_j + U_{j+1}}{h^2} + U_j &= 2x_j \sin(x_j) - 2 \cos(x_j) \Rightarrow \\
 -\frac{1}{h^2}U_{j+1} + \left(\frac{2}{h^2} + 1\right)U_j - \frac{1}{h^2}U_{j-1} &= 2j \frac{\pi}{n-1} \sin\left(j \frac{\pi}{n-1}\right) - 2 \cos\left(j \frac{\pi}{n-1}\right)
 \end{aligned}$$



Επομένως το σύστημα που προκύπτει είναι τριδιαγώνιο της μορφής

$$AU = b$$

όπου το  $A$  χαρακτηρίζεται πλήρως από τα διανύσματα  $d = (1+2/h^2)[1, \dots, 1]$  και  $e = -[1, \dots, 1]/h^2$ .

Το σύστημα αυτό είναι συμμετρικό και θετικά ορισμένο. Αυτό επαληθεύεται άμεσα από το θεώρημα

Gerschgorin: Οι ιδιοτιμές είναι πραγματικές (λόγω συμμετρίας) και βρίσκονται στην ένωση των

διαστημάτων που ορίζονται από:

$$\begin{aligned} |z - (\frac{2}{h^2} + 1)| &\leq \frac{2}{h^2} \\ (\frac{2}{h^2} + 1) - \frac{2}{h^2} &\leq z \leq (\frac{2}{h^2} + 1) + \frac{2}{h^2} \end{aligned}$$

και

$$\begin{aligned} |z - (\frac{2}{h^2} + 1)| &\leq \frac{1}{h^2} \\ (\frac{2}{h^2} + 1) - \frac{1}{h^2} &\leq z \leq (\frac{2}{h^2} + 1) + \frac{1}{h^2} \end{aligned}$$

Επομένως σε κάθε περίπτωση οι ιδιοτιμές είναι θετικές και επομένως το μητρώο σθο. Τέλος, το

δεξιό μέλος έχει για στοιχεία:

$$b_j = 2j \frac{\pi}{n-1} \sin(j \frac{\pi}{n-1}) - 2 \cos(j \frac{\pi}{n-1})$$

Επομένως μπορούμε να χρησιμοποιήσουμε παραγοντοποίηση Cholesky και μπορούμε να προχωρή-

σουμε ως εξής:

```

n1 = n-1; nn = 1:n-1;
h = pi/n1; onv = ones(n1,1);
d = (1+2/h^2)*ones(n1,1); e = -ones(n1,1)/h^2;
[g,h] = CholTridFact(d,e);
b = 2*nn*h*sin(nn*h)-2*cos(nn*h);
Ut = CholTridSol(g,h,b);
U(2:n-1) = Ut(1:n-2); U(1) = 0; U(n) = 0;

```

Κάτω από τις συνθήκες που αναφέραμε, το κόστος των υπολογισμών που αφορούν στην παραγοντοποίηση και επίλυση μέσω Cholesky (περίπου  $4n$  για την παραγοντοποίηση και περίπου  $6n$  για τις αντικαταστάσεις) είναι πολύ μικρότερο από εκείνο της κλασικής LU ( $O(n^3)$ ). Το κόστος είναι ελαφρώς μεγαλύτερο σε σύγκριση με την LU για τριδιαγώνια μητρώα που απαιτεί συνολικά περίπου  $8n$  πράξεις. (Σημ.: μπορούμε να μειώσουμε το κόστος κατά  $n$  αν χρησιμοποιήσουμε την παραπλήσια παραγοντοποίηση  $A = LDL^T$  με το  $L$  κάτω διδιαγώνιο με μονάδες στη διαγώνιο.)  $\square$