ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι (24 Φεβρ. 2008, 12-3μμ) ΕΠΙΛΕΓΜΕΝΕΣ ΑΠΑΝΤΗΣΕΙΣ

1. α) Σ - Λ : Οι εντολές BLAS-2 μπορούν να υλοποιηθούν να έχουν καλύτερη επίδοση από τις BLAS-3.

   Απάντηση. Λάθος: Οι εντολές BLAS-3 έχουν μικρότερο ελάχιστο αριθμό μεταφορών ανά πράξη α.κ.υ. από τις πράξεις BLAS «μικρότερων κατηγοριών». Επομένως, υπό την προϋπόθεση ότι αναφερόμαστε σε υλοποιήσεις που έχουν γίνει με στόχο την επίτευξη του μικρότερου λόγου μεταφορών προς πράξεις για κάθε κατηγορία, οι πράξεις BLAS-3 θα έχουν καλύτερη επίδοση (μετρούμενου με βάση τα Mflops). □

   β) Σ - Λ : Ξεδίπλωμα βρόχου γενικά χρησιμοποιείται για να μειώσει το πλήθος πράξεων α.κ.υ.

   Απάντηση. ΛΑΘΟΣ το ξεδίπλωμα δεν επιφέρει αλλαγή του Ω, μόνον ο βρόχος εκτελείται λιγότερες φορές αλλά με περισσότερες εντολές σε κάθε επανάληψη. □

   γ) Έστω στη MATLAB οι εκφράσεις $M+20-10-M$, $M+20-M-10$, $M-10-M+20$. *Να* εξηγήσετε τις τιμές που υπολογίζονται αν το M αρχικοποιηθεί ως `realmax`.

   Απάντηση. Το `realmax` της α.κ.υ. διπλής ακρίβειας είναι της μορφής $1.*\times 2^{1023}$ επομένως η προσθαφαίρεση αριθμών σαν το 10 και 20 με αυτό δεν επιφέρει καμια αλλαγή λόγω της απαιτούμενης κανονικοποίησης και επακόλουθου μηδενισμού τους κατά την πρώτη φάση της διαδικασίας. Επομένως τα αποτελέσματα θα είναι $((M+20)-10)-M = (M-10)-M = M-M = 0$, $((M+20)-M)-10 = (M-M)-10 = -10$, $((M-10)-M)+20 = (M-M)+20 = 20$. □

   δ) Έστω αντιστρέψιμο $A\in\mathbb{R}^{n\times n}$ με μικρό δείκτη κατάστασης, $b\in\mathbb{R}^n$ και ο υπολογισμός $[L,U] =$ `lu`(A); $x = U\backslash(L\backslash b)$ (η MATLAB χρησιμοποιεί LAPACK). Ισχύει ή όχι ότι το εμπρός σφάλμα στο υπολογισμένο x δεν θα είναι μεγάλο;

   Απάντηση. Για την LU γενικού μητρώου δεν μπορεί να αποδειχτεί μικρή πίσω ευστάθεια, που είναι *απαραίτητη για να εγγυηθούμε μικρό εμπρός σφάλμα λόγω μικρού δείκτη κατάστασης, επομένως* ΔΕΝ ΙΣΧΥΕΙ. Βασιζόμαστε και στον γνωστό τύπο (εμπρός σφ.) < (πίσω σφ.) × (δείκτης κατ. A).) □

2. Μας δίδονται α.κ.υ. και ένας αλγόριθμος για να τους αθροίσουμε. Να εξηγήσετε ποιοί από τους παρακάτω ισχυρισμούς είναι σωστοί και ποιοί λάθος:

   α) Αν αλλάξουμε τον αλγόριθμο άθροισης, μπορεί να αλλάξουν το πίσω σφάλμα και το εμπρός σφάλμα.

   β) Αν γνωρίζουμε τους α.κ.υ. και τον αλγόριθμο άθροισης, μπορούμε να υπολογίσουμε το ***ακριβές*** εμπρός σφάλμα.

   γ) Αν οι αριθμοί είναι ομόσημοι, ενας καλός τρόπος άθροισης είναι από το μικρότερο προς το μεγαλύτερο.

   δ) Αν η απόλυτη τιμή του υπολογισμένου αθροίσματος είναι πολύ μικρότερη του μέσου όρου των απολύτων τιμών των στοιχείων που αθροίστηκαν, μπορούμε να υποθέσουμε με ασφάλεια ότι το σχετικό εμπρός σφάλμα στο άθροισμα θα είναι και αυτό μικρό.

   Απάντηση. α) ΣΩΣΤΟ, και τα δυο εξαρτώνται από τον αλγόριθμο και επομένως τη σειρά άθροισης (εξάλλου το πίσω σφάλμα μετρά τον «δείκτη κατάστασης του αλγορίθμου».) β) ΛΑΘΟΣ, το ακριβές σφάλμα δεν μπορεί να υπολογιστεί γενικά γιατί χρειαζόμαστε αριθμητική άπειρης ακρίβειας. γ) ΣΩΣΤΟ, γιατί τότε μειώνεται η πιθανότητα σφάλματος από την πρόσθεση αριθμών που διαφέρουν πάρα πολύ σε μέγεθος που θα είχε για συνέπεια μηδενισμό των μικρότερων λόγω κανονικοποίησης των εκθετών. Επίσης τα «δ» που συσσωρεύονται στη διάδοση του σφάλματος επιβαρύνουν περισσότερο τους μικρότερους όρους του αθροίσματος. δ) ΛΑΘΟΣ: Τυπικό παράδειγμα $(1+\delta_1)-(1-\delta_2) = \delta_1+\delta_2$ όπου τα δ_j είναι πολύ μικρά και περιέχουν κυρίως «θόρυβο» από προηγούμενες πράξεις. Κλασικό παράδειγμα που δημιουργείται πρόβλημα από καταστροφική απαλοιφή. □

3. Δίδονται τα στοιχεία $A\in\mathbb{R}^{10\times m}$ και $b\in\mathbb{R}^m$, $c\in\mathbb{R}^{10}$ και θέλουμε να υπολογίσουμε το $y\leftarrow c+Ab$. Το m δεν έχει κανέναν περιορισμό. α) Ποιό είναι το $\Phi_{\min}$ για την πράξη; β) Να δείξετε πώς μπορείτε να υλοποιήσετε τον πολλαπλασιασμό με $\Phi = \Phi_{\min}$ χρησιμοποιώντας κρυφή μνήμη και καταχωρητές $O(1)$ (δηλ. προσωρινή μνήμη άμεσης πρόσβασης μεγέθους ανεξάρτητου του m).

   Απάντηση. α) Με απλή καταμέτρηση των α.κ.υ. εισόδου/εξόδου που χρησιμοποιούνται στον υπολογισμό, έχουμε $10m$ για φόρτωση του A, $m+10$ για φόρτωση των c, b, και 10 για την αποθήκευση στο y, συνολικά δηλ. $\Phi_{\min} = 11m + 20$. β) Η σχετική ύλη υπάρχει και στις διαφάνειες. Συνοψίζουμε λέγοντας ότι η υλοποίηση μπορεί να κωδικοποιηθεί ως εξής, εφόσον διατίθεται χώρος για την αποθήκευση σε καταχωρητές και cache της τάξης του $O(1)$. Η μεταβλητή temp έχει αναφέρεται σε καταχωρητές μήκους 10.

1. LOAD c
2. for $j = 1 : m$
3. LOAD $b(j)$
4. for $i = 1 : 10$
5. LOAD $A(i, j)$ □
6. temp(i) = $c(i) + A(i, j) * b(j)$
7. end
8. end
9. STORE y = temp

4. α) Τι θα εμφανιστεί στην οθόνη αν εκτελέσετε τις παρακάτω εντολές σε περιβάλλον MATLAB και n=3:

```
for j=1:n, A = kron(ones(j,1),[1:j]), end
```

Απάντηση.

$A = 1$

$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$

$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ □

Υπενθυμίζουμε ότι η εντολή kron(A,B) επιστρέφει το γινόμενο Kronecker $A \otimes B$.

β) Να ενθέσετε (αιτιολογώντας, πάντα) σε επιπλέον κώδικα που να υπολογίζει όσο μπορείτε πιο αξιόπιστα (επιστρέφοντας σε κάποια μεταβλητή) τα Mflop/s των παραπάνω εντολών στο υπολογιστικό σας περιβάλλον. Μπορείτε να υποθέσετε ότι άν $A \in \mathbb{R}^{m_A \times n_A}$, $B \in \mathbb{R}^{m_B \times n_B}$ τότε το κόστος του kron(A,B) είναι $\Omega = m_A n_A m_B n_B$.

Απάντηση. Για συντομία συμβολίζουμε με $\mathbf{\Delta}$ τις εντολές `for j=1:n, A = kron(ones(j,1),[1:j]), end`. Προσέξτε ότι το Ω θα είναι $\sum_{j=1}^{n} j^2$. Μπορεί να υπολογιστεί από κλασικούς τύπους αθροισμάτων προόδων ή στο πρόγραμμα, συσσωρεύοντας τις πράξεις κάθε επανάληψης σε μεταβλητή. Τότε

```
% εκτέλεση για να αποφευχθεί «θόρυβος» από την αρχικοποίηση
tic; for j=1:itmax, Δ; end;
optime = toc/itmax; ops = 0;
for j=1:itmax, ops = ops+j*j; end; mflops = ops*1e-6/toc;
```
□

5. α) Είναι το μοντέλο διάδοσης του σφάλματος στον πολλαπλασιασμό κινητής υποδιαστολής, $x \tilde{\times} y = x \times y(1+\delta)$ όπου $|\delta| \le \mathbf{u}$, $\mathbf{u}$ η μονάδα στρογγύλευσης και x, y αριθμοί κινητής υποδιαστολής, άμεσο επακόλουθο της «αρχής ακριβούς στρογγύλευσης»; Αν ναι, να το δείξετε, αν όχι να εξηγήσετε γιατί.

Απάντηση. ΕΙΝΑΙ: Η αρχή προσδιορίζει ότι με τις παραπάνω συνθήκες, για τον πολλαπλασιασμό ισχύει ότι η πράξη που εκτελείται στη μηχανή έχει ως αποτέλεσμα την ποσότητα που θα υπολογιζόταν με αριθμητική άπειρης ακρίβειας (δηλ. το $x \times y$) με στρογγύλευση (υποθέτουμε προς το πλησιέστερο) μετά, επομένως το τελικό αποτέλεσμα θα είναι $x \times y(1+\delta)$ όπου $|\delta| \le \mathbf{u}$. □

β) Γνωρίζουμε ότι ο κλασικός δείκτης κατάστασης ενός μητρώου ως προς την επίλυση συστήματος $Ax = b$ ορίζεται ως $\kappa(A) := \|A\|\|A^{-1}\|$ για επιλεγμένη νόρμα. Να δείξετε ένα μητρώο 3×3 για το οποίο το $\kappa(A)$ είναι πάρα πολύ μεγάλο και το υπολογισμένο $\tilde{x}$ να έχει συγκριτικά πολύ μικρό σχετικό σφάλμα.

Απάντηση. Μπορείτε να διαλέξετε ένα διαγώνιο μητρώο A, με διαγώνιο $[1, 1, 1e-10]$, οπότε ο δείκτης κατάστασης είναι 1e10. Από την άλλη, αν λύσετε το σύστημα $Ax = b$, λόγω της διαγώνιας δομής του A, κάθε στοιχείο της λύσης x υπολογίζεται με μια διαίρεση, επομένως το άνω φράγμα για το σχετικό σφάλμα κάθε στοιχείου της υπολογισμένης λύσης $\tilde{x}$ θα είναι $\mathbf{u}$. □

6. α) Έστω ότι ένα μητρώο $H \in \mathbb{R}^{n\times n}$ έχει μηδενικά στις θέσεις που βρίσκονται κάτω από την πρώτη υποδιαγώνιο, δηλ. $(3:n,1),(4:n,2),\ldots,(n,n-1)$. Να δείξετε ότι (χωρίς οδήγηση και εφόσον υπάρχει) η παραγοντοποίηση LU του H κοστίζει $\Omega = \alpha n^2 + O(n)$. Επίσης να υπολογίσετε τον κυρίαρχο συντελεστή α.

Απάντηση. Προσέχουμε ότι σε κάθε βήμα $k = 1,\ldots,n-1$ της κλασικής απαλοιφής, χρειάζεται να απαλείψουμε μόνον ένα υποδιαγώνιο στοιχείο (στη θέση $(k+1,k)$). Επομένως το κόστος θα είναι $\Omega = \sum_{k=1}^{n-1}(1+\sum_{j=k+1}^{n} 2)$ επομένως $\Omega = n(n-1) + O(n)$ άρα $\alpha = 1$. Ο κώδικας μπορεί να είναι ο εξής (προαιρετικά):

```
for k=1:n-1
  H(k+1,k) = H(k+1,k)/H(k,k)
  for j=k+1:n
    H(k+1,k+1:n) = H(k+1,k+1:n) - H(k+1,k)*H(k+1,k+1:n)
  end
end
```
□

β) Δίδεται $A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 1 & -1 \\ 0 & 3 & -1 & 1 \\ 0 & 4 & 1 & 2 \end{pmatrix}$.

Να υπολογίσετε διάνυσμα Householder ώστε ο (ορθογώνιος) ανακλαστής P που παράγεται από το διάνυσμα, να μηδενίζει τη θέση $(4,2)$ του μητρώου PA καθώς επίσης και του $B = PAP^{\top}$. Επίσης να υπολογίσετε το B (να φέρετε σε πέρας όλες τις αριθμητικές πράξεις.) Προσοχή: Δεν χρειάζεται (δεν είναι εφικτό) να είναι 0 το στοιχείο στη θέση $(3,2)$.

Απάντηση. Σε MATLAB, u = [0; 0; A(3 : 4, 2)] + [0, 0, 1, 0]' * norm(A(3 : 4, 2)), επομένως $u = [0, 0, 8, 4]^{\top}$ και υπολογίζεται ότι

Επιστρέφει τη 2-νόρμα του A(3:4,2)

$$B = \begin{pmatrix} 1 & 1 & -2 & -1 \\ 0 & 2 & 0.2 & -1.4 \\ 0 & -5 & 1.88 & -1.16 \\ 0 & 0 & -1.16 & -0.88 \end{pmatrix}$$

□

γ) Για κάθε A, μπορεί να υπολογιστεί (π.χ. η συνάρτηση hess στη MATLAB) ορθογώνιο μητρώο Q ως γινόμενο ανακλαστών Householder, ώστε το $QAQ^{\top}$ να έχει μηδενικά κάτω από την υποδιαγώνιο. Ο υπολογισμός των Q και $QAQ^{\top}$ κοστίζουν συνολικά περί τις $5n^3$ πράξεις α.κ.υ. Έστω ότι χρειάζεται να υπολογίσετε τις λύσεις x_j, $j = 1,\ldots,s$ των s συστημάτων $(A-\omega_j I)x_j = b_j$ όπου $A \in \mathbb{R}^{n\times n}$ και τα ω_j είναι πραγματικοί αριθμοί τέτοιοι ώστε τα μητρώα $A-\omega_j I$ να είναι αντιστρέψιμα και I το ταυτοτικό μητρώο. Να περιγράψετε τα βασικά βήματα αλγορίθμου που επιτυγχάνει τη λύση των s συστημάτων με κόστος $\Omega \approx 5n^3 + O(sn^2)$ αντί για $O(sn^3)$ που θα στοίχιζε αν χρησιμοποιούσατε απευθείας LU.

Απάντηση. ΒΙΒΛΙΟ □

7. Δίδεται η διαφορική εξίσωση $u''(x) + 10^{-2}(20-u) = 0$ στο διάστημα $[0,10]$ με συνοριακές συνθήκες $u(0) = 40$, $u(10) = 200$ και θέλουμε να προσεγγίσουμε τη λύση με κεντρισμένες πεπερασμένες διαφορές και ακρίβεια τάξης $O(h^2)$, όπου h είναι η απόσταση μεταξύ των ισαπέχοντων κόμβων του πλέγματος που θα χρησιμοποιήσουμε στη διακριτοποίηση.

α) Να εξηγήσετε σύντομα γιατί συνήθως απαιτούμε από τη συνάρτηση $u(x)$ να έχει παραγώγους μέχρι και 4ης τάξης και αυτές να είναι συνεχείς στο διάστημα $[0,10]$.

Απάντηση. Από τη θεωρία γνωρίζουμε ότι η διακριτοποίηση βασίζεται στο συνδυασμό τιμών της συνάρτησης σε επιλεγμένους (γειτονικούς) κόμβους του πλέγματος και στα σχετικά αναπτύγματα Taylor. *Ειδικότερα, υπό την προϋπόθεση ότι η u διαθέτει τουλάχιστον 4 παραγώγους και συμβολίζοντας με u_j* την τιμή της συνάρτησης στον κόμβο j ενός φυσικά αριθμημένου πλέγματος, μπορούμε να γράψουμε

$$u_{j\pm1} = u_j \pm h u_j^{(1)} + \frac{h^2}{2}u_j^{(2)} \pm \frac{h^3}{6}u_j^{(3)} + \frac{h^4}{24}u^{(4)}(x_j + \theta_j^{\pm} h)$$

όπου $-1 < \theta_j^- < 0 < \theta_j^+ < 1$. Επομένως

$$u_{j-1} + u_{j+1} - 2u_j = h^2 u_j^{(2)} + \frac{h^4}{24}\left(u^{(4)}(\xi_j + \theta_j^+ h) + u^{(4)}(\xi_j + \theta_j^- h)\right)$$

Επομένως, το σφάλμα διακριτοποίησης της 2ης παραγώγου σε κάθε σημείο εξαρτάται άμεσα από την διακριτοποίηση (δηλ. το h) και τη διακύμανση της τιμής του $|u^{(4)}|$. Το h το επιλέγεται από εμάς, επομένως μπορούμε να το επιλέξουμε όσο μικρό θέλουμε (μόνος περιορισμός είναι το μέγεθος του προκύπτοντος συστήματος) για να πετύχουμε αποδεκτό σφάλμα. Όμως, παράλληλα, θα πρέπει να αποκλείσουμε την περίπτωση να γίνεται το h πολύ μεγάλο. Αυτό εξασφαλίζεται «αυτόματα» όταν η συνάρτηση $u^{(4)}$ είναι συνεχής στο κλειστό διάστημα ορισμού της, καθώς τότε, από γνωστό στοιχειώδες θεώρημα της Μαθηματικής Ανάλυσης, έπεται ότι το $|u^{(4)}|$ θα είναι φραγμένο σε όλο το διάστημα. □

β) Να υπολογίσετε μητρώο $A \in \mathbb{R}^{4\times 4}$ και δεξιό μέλος $b \in \mathbb{R}^4$ τέτοια ώστε το διάνυσμα y που ικανοποιεί το σύστημα $Ay = b$ να προσεγγίζει τη λύση u στους κόμβους.

Απάντηση. Διαμερίζουμε το διάστημα $[0, 10]$ σε 4 ισαπέχοντες εσωτερικούς κόμβους επομένως $h = 10/5 = 2$ και οι κόμβοι θα είναι $\xi_j = jh$ για $j = 1, \dots, 4$. Χρησιμοποιώντας κεντρισμένες πεπερασμένες *διαφορές 2ης τάξης για την προσέγγιση της 2ης παραγώγου θα έχουμε*

$$\frac{u(\xi_{j-1}) - 2u(x_j) + u(\xi_{j+1})}{h^2} + 20 \times 10^{-2} - 10^{-2}u(x_j) = 0$$

επομένως οι εξισώσεις σε κάθε σημείο καθορίζονται από τον τύπο

$$\frac{1}{h^2}U_{j-1} - (\frac{2}{h^2} + 10^{-2})U_j + \frac{1}{h^2}U_{j+1} = -20 \times 10^{-2}$$

που ξαναγράφουμε ως

$$-\frac{1}{4}U_{j-1} + (\frac{1}{2} + 10^{-2})U_j - \frac{1}{4}U_{j+1} = 20 \times 10^{-2}$$

Επομένως το σύστημα θα είναι

$$\begin{pmatrix} 0.51 & -0.25 & 0 & 0 \\ -0.25 & 0.51 & -0.25 & 0 \\ 0 & -0.25 & 0.51 & -0.25 \\ 0 & 0 & -0.25 & 0.51 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{pmatrix} = \begin{pmatrix} 10.2 \\ 0.2 \\ 0.2 \\ 50.2 \end{pmatrix}$$

□

γ) Έστω ότι η παραπάνω διαφορική εξίσωση τροποποιείται σε $u''(x) + 10^{-2}(20 - u) - (1 + x^2) = 0$. Ποιοί θα είναι τώρα οι νέοι παράγοντες A και b;

Απάντηση. Για να ληφθεί υπόψη ο νέος παράγοντας $1 + x^2$, διαφοροποιείται μόνον το δεξιό μέλος: $b = [15.2, 17.2, 37.2, 115.2]^\top$. □

δ) Στη συνέχεια, αλλάζουμε τη συνοριακή συνθήκη του αρχικού προβλήματος (δηλ. του μέρους (α)) από $u(0) = 40$ σε $u'(0) = -2$. Χρησιμοποιώντας κεντρισμένες πεπερασμένες διαφορές 2ης τάξης να γράψετε το νέο σύστημα που θα προκύψει, έστω $\hat{A}y = \hat{b}$. Προσοχή: Τα $\hat{A}, \hat{b}$ μπορεί να έχουν διαφορετικό μέγεθος από πριν.

Απάντηση. Με την αλλαγή αυτή δεν γνωρίζουμε πλέον το $u(0)$ αλλά την παράγωγο την οποία προσεγγίζουμε ως

$$\frac{U_1 - U_{-1}}{2h} \approx u'(0) = 40 \Rightarrow U_{-1} = U_1 - 160$$

θεωρώντας ότι U_{-1} είναι προσέγγιση του u στο -2. Επίσης, γράφουμε την εξίσωση για το σημείο 0, δηλ.

$$-\frac{1}{4}U_{-1} + (\frac{1}{2} + 10^{-2})U_0 - \frac{1}{4}U_1 = 20 \times 10^{-2}$$

οπότε

$$-\frac{1}{4}(U_1 - 160) + (\frac{1}{2} + 10^{-2})U_0 - \frac{1}{4}U_1 = 20 \times 10^{-2}$$

άρα επαυξάνουμε το αρχικό σύστημα ως εξής:

$$\begin{pmatrix} 0.51 & -0.5 & 0 & 0 & 0 \\ -0.25 & 0.51 & -0.25 & 0 & 0 \\ 0 & -0.25 & 0.51 & -0.25 & 0 \\ 0 & 0 & -0.25 & 0.51 & -0.25 \\ 0 & 0 & 0 & -0.25 & 0.51 \end{pmatrix} \begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \\ U_4 \end{pmatrix} \quad \begin{pmatrix} -39.8 \\ 0.2 \\ 0.2 \\ 0.2 \\ 50.2 \end{pmatrix}$$

□

8. *Έστω η διαφορική εξίσωση* $u'''(t) = -1000u(t) - 300u'(t) - 30u''(t)$ *με αρχικές τιμές* $u(0) = 1, u'(0) = 0, u''(0) = 1$. α) Να υπολογίσετε το $u(1.6)$ χρησιμοποιώντας εμπρός **Euler** και βήμα $h = 0.8$. (Προσοχή: Η εξίσωση είναι 3ης τάξης και είναι προτιμότερο να την μετατρέψετε σε γραμμικό σύστημα συνήθων διαφορικών εξισώσεων). β) Να εξηγήσετε αν με το παραπάνω βήμα *μπορεί να παρουσιαστεί αστάθεια αν συνεχίσετε την προσέγγιση για πολλά βήματα και αν ναι, να* υπολογίσετε άνω φράγμα για το βήμα h ώστε να αποφευχθεί η αστάθεια.

Απάντηση. α) Όπως προτείνεται μετατρέπουμε το παραπάνω σε σύστημα με την εισαγωγή βοηθητικών μεταβλητών (δείτε βιβλίο και διαφάνειες): $u_1(t) := u(t)$, $u'(t) := u_2(t)$, και $u''(t) := u_3(t)$ οπότε η διαφορική μετατρέπεται σε σύστημα 3 συνήθων διαφορικών, ως εξής

$$\frac{d}{dt}\begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix} = -\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1000 & 300 & 30 \end{pmatrix}\begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix}$$

ή για συντομία

$$\frac{d}{dt}\mathbf{u} = -A\mathbf{u}$$

όπου $\mathbf{u} := [u_1, u_2, u_3]^\top$ (παραλείπουμε το t το οποίο εννοείται). Εφαρμόζοντας εμπρός **Euler** με το βήμα $h = 0.8$ και $U(0) = [1, 0, 1]^\top$, για να υπολογίσουμε την τιμή στο $t = 2h$ έχουμε έχουμε ότι

$$U(2h) = (I - hA)((I - hA)U(0)) = [\mathbf{1.64}, -657.6, 17937]^\top.$$

Με παχειά γραφή έχουμε συμβολίσει το ζητούμενο, δηλ. την προσέγγιση στο $u(2h)$ με εμπρός **Euler**.

β) Προσέξτε ότι από τη διακύμανση των στοιχείων φαίνεται ότι μάλλον υπάρχει αστάθεια! Για να το επιβεβαιώσουμε, εξετάζουμε τη μέγιστη ιδιοτιμή του $I - hA$ για το βήμα h που χρησιμοποιήσαμε. Οι ιδιοτιμές του A είναι οι ρίζες του πολυωνύμου $1000 + 300\lambda + 30\lambda^2 + \lambda^3 = 0$, οπότε $\lambda_1 = \lambda_2 = \lambda_3 = -10$. Επομένως με $h = 0.8$ η φασματική ακτίνα του $I - hA$ θα είναι $7 = |1 - 0.8 \times 10|$ και θα έχουμε αστάθεια. Εδικότερα, το βήμα h πρέπει να επιλέγεται μικρότερο από $2/\max|\lambda_j| = 0.5$. □

γ) Γενικά στην **Euler** για την επίλυση ενός γραμμικού προβλήματος του τύπου $u' = -Au$, είναι σωστό ή λάθος ότι αν μειωθεί το βήμα στο μισό, τότε το μέγιστο ολικό σφάλμα διακριτοποίησης θα υποτετραπλασιαστεί.

Απάντηση. ΛΑΘΟΣ, το ολικό σφάλμα συμπεριφέρεται όπως το $O(1/h)$ άρα περιμένουμε να υποδιπλασιαστεί. □

δ) *Για καθένα από τα παρακάτω σχετικά με τις άμεσες μεθόδους* **Runge-Kutta** *τάξης 2 και πάνω για* την επίλυση της ΣΔΕ $u'(t) = f(t, u)$, να κυκλώσετε αν είναι σωστό ή λάθος:

(Σ - Λ) Προβλέπουν τη νέα τιμή συνδυάζοντας την προσέγγιση στο t_k με προσεγγίσεις της παραγώγου της u σε μια ή περισσότερες τιμές του t στο διάστημα $[t_k, t_{k+1})$.

Απάντηση. ΣΩΣΤΟ, οι μέθοδοι **RK** είναι μονοβηματικές και χρησιμοποιούν ως πληροφορία την προσέγγιση στο t_k με εκτιμήσεις της παραγώγου στο t_k και άλλα σημεία στο παραπάνω διάστημα. Ο γενικός τύπος είναι

$$U_{k+1} = U_k + h\sum_{i=1}^{s} b_i K_i$$

όπου

$$K_i = f(t_n + c_i h, U_n + h \sum_{j=1}^{s} a_{ij} K_j)$$

□

(Σ - Λ) Έχουν πιο εκτεταμένο χωρίο ευστάθειας από την πίσω Euler.

Απάντηση. ΣΩΣΤΟ, η πίσω ευστάθεια καθορίζεται από χωρία της μορφής

$$D := \{z : h\lambda \,|\, |p_n(z)| \leq 1\}, \quad p_n(z) = \sum_{j=0}^{s} \frac{z^j}{j!}$$

όπου το πολυώνυμο προκύπτει από την εφαρμογή της μεθόδου στο $u' = \lambda u$. □

Ονοματεπώνυμο: .... Α.Μ. ....· Έτος ...

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι *Σεπτέμβριος 2005*

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!!! Διαβάστε προσεκτικά τις εκφωνήσεις **(2 σελίδες)**. *Για πλήρη αξιολόγηση του γραπτού σας πρέπει να παρουσιάσετε όλο σας τον συλλογισμό και όλα τα ενδιάμεσα αποτελέσματα. Έχετε 3 ώρες. Οι αλγόριθμοι να περιγράφονται με σαφήνεια, π.χ. όπως στις σημειώσεις ή με MATLAB. Εάν κατά τη περιγραφή ενός αλγορίθμου απαιτηθεί η διάσπαση Cholesky, μπορείτε να χρησιμοποιήσετε κατευθείαν την εντολή της MATLAB,* `R = chol(A)` *όπου R είναι άνω τριγωνικό μητρώο τέτοιο ώστε* $R^* \cdot R = A$.

I. *(20 β.)*

1. Να περιγράψετε με συντομία τα τρία βασικά κριτήρια αξιολόγησης στον Επιστημονικό Υπολογισμό.
2. Στην προσπάθεια να μετρηθεί η επίδοση μιας συνάρτησης `flat` γραμμένης σε MATLAB, εκτελέστηκαν οι παρακάτω εντολές σε περιβάλλον MATLAB:

   ```
   tic; [x,ops]=flat; val=ops/toc/1e6; end;
   ```

   όπου x είναι κάποιο αποτέλεσμα που υπολογίζει η flat και ops είναι το πλήθος πράξεων α.κ.υ. της `flat`. Να εξηγήσετε τι μετρά το `val`.
3. Να αναφέρετε ένα λόγο για τον οποίο θα μπορούσε να επιστραφεί η τιμή `Inf` στο `val` όταν εκτελείται το παραπάνω σε έναν ταχύ Η/Υ (π.χ. ένα σύγχρονο Pentium).
4. Να εξηγήσετε πώς θα μπορούσατε να μετρήσετε το val με πιο αξιόπιστο τρόπο (χωρίς να αλλάξετε το περιεχόμενο της `flat`).

Απάντηση. 1) Η val μετρά το πλήθος των πράξεων α.κ.υ. ανά μονάδα χρόνου και επειδή διαιρούμε με το 1e6, έχουμε τα εκατομμύρια πράξεων α.κ.υ. ανά δευτερόλεπτο, που σημαίνει τα Mflops του αλγορίθμου.

2) Επειδή το πραγματικό διάστημα που μεσολαβεί μεταξύ της κλήσης του `tic` και του toc μπορεί να είναι μικρότερο της διακριτότητας του συστήματος α.κ.υ. και να επιστρέφεται toc=0 ή ένας αριθμός τόσο μικρός που να υπάρχει υπερχείλιση στη διαίρεση.

3) Για να αποφύγουμε το παραπάνω πρόβλημα και να μετρήσουμε αξιόπιστα τα Mflops, μπορούμε να εμφωλεύσουμε τον flat σε βρόχο, με κατάλληλα επιλεγμένο s, ως εξής:

```
tic; for j=1:s, [x,ops]=flat; end; val=s*ops/toc/1e6; end;
```

□

II. *(20 β.)*

1. Ποια είναι η «συνθήκη ακριβούς στρογγύλευσης»;
2. Να δείξετε ότι ο αλγόριθμος πολλαπλασιασμού δυο άνω τριγωνικών μητρώων $A, B \in \mathbb{R}^{2\times 2}$ είναι πίσω ευσταθής. Μπορείτε να θεωρήσετε ότι τα στοιχεία των A, B είναι α.κ.υ.
3. Έστω ότι υπολογίζετε με α.κ.υ την τιμή μιας (άγνωστης) ποσότητας x στο μοντέλο α.κ.υ. και διαπιστώνετε ότι είναι $\tilde{x}$. Έστω επίσης ότι είναι γνωστό ότι $|x-\tilde{x}|/|\tilde{x}| \leq \delta$ για κάποιο μικρό $\delta < 1$. Με βάση τα παραπάνω, να βρείτε, ως συνάρτηση του δ, ένα καλό άνω φράγμα για το σχετικό σφάλμα $|x-\tilde{x}|/|x|$.

III. *(20 β.)*

1. Να βρείτε ακριβώς ένα μητρώο μετάθεσης P για το οποίο ισχύει ότι το μητρώο $B := PA$, για οποιοδήποτε μητρώο $A \in \mathbb{R}^{n\times n}$, έχει για στοιχεία τα $\beta_{ij} = \alpha_{n+1-i,j}$, όπου, ως συνήθως, $\alpha_{i,j}$ συμβολίζει το στοιχείο στη θέση (i,j) του A. Τότε, αν το μητρώο L είναι άνω τριγωνικό, ποια θα είναι η δομή του μητρώου $C := PLP$;

2. Με βάση τα παραπάνω, να υποδείξετε έναν τρόπο για τον υπολογισμό της παραγοντοποίησης ενός μητρώου A ως $A = UL$, όπου U, L αντίστοιχα είναι άνω και κάτω τριγωνικά και το U έχει μονάδες στη διαγώνιο. Μπορείτε να υποθέσετε ότι δεν απαιτείται οδήγηση.

3. Να χρησιμοποιήσετε την παραπάνω ιδέα (πάντα χωρίς οδήγηση) για να λύσετε το γραμμικό σύστημα $Ax = e_1$, όπου e_1 είναι το διάνυσμα $[1,0,0]^\top$ και $A = [10,-1,-1;-1,8,-1;-1,-1,5]$ ώστε να εξοικονομήσετε περίπου n^2 πράξεις κατά τη λύση σε σχέση με την κλασική LU (πρέπει να δείξετε που οφείλεται αυτή η εξοικονόμιση), όπου βέβαια στην περίπτωσή μας $n = 3$.

Απάντηση. 1) Το μητρώο P είναι το αντιδιαγώνιο μητρώο που, σε MATLAB , ορίζεται ως `P = I(n:-1:1,:)` καθώς ο πολλαπλασιασμός PA έχει ως αποτέλεσμα την ανταλλαγή των γραμμών i και n+1-j. Επομένως, για άνω τριγωνικό μητρώο L, το μητρώο PLP θα έχει κάτω τριγωνική μορφή.

2) Το μητρώο $\hat{A}; = PAP$ το οποίο θα έχει ως στοιχεία τα $\hat{\alpha}_{ij} = \alpha_{n+1-i,n+1-j}$. Όμως ισχύει επίσης ότι $P^2 = I$. Επομένως, αν χρησιμοποιήσουμε LU στο $\hat{A}$, θα έχουμε $PAP = \hat{L}\hat{U}$. επομένως $A = P\hat{L}PP\hat{U}P$ και με βάση τη δράση του P, το $U := P\hat{L}P$ είναι άνω τριγωνικό με 1 στη διαγώνιο και το $L := P\hat{L}P$ είναι άνω τριγωνικό. Επομένως ο υπολογισμός μπορεί να γίνει ως εξής (π.χ. σε MATLAB):

```
[tL,tU] = lu(P*A*P); U = P*tL*P; L = P*tU*P;
```

3) Με βάση τα παραπάνω, θα χρησιμοποιήσουμε LU επί του $PAP = [5,-1,-1;-1,8,-1;-1,-1,10]$. Στο πρώτο βήμα:

$$L_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1/5 & 1 & 0 \\ 1/5 & 0 & 1 \end{pmatrix}, L_1 A = \begin{pmatrix} 5 & -1 & -1 \\ 0 & \frac{39}{5} & -6/5 \\ 0 & -6/5 & \frac{49}{5} \end{pmatrix},$$

$$L_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2/13 & 1 \end{pmatrix}, L_2(L_1 A) = \begin{pmatrix} 5 & -1 & -1 \\ 0 & \frac{39}{5} & -6/5 \\ 0 & 0 & \frac{125}{13} \end{pmatrix}$$

Επίσης

$$L_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1/5 & 1 & 0 \\ -1/5 & 0 & 1 \end{pmatrix}, L_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2/13 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1/5 & 1 & 0 \\ 1/5 & -2/13 & 1 \end{pmatrix} \begin{pmatrix} 5 & -1 & -1 \\ 0 & \frac{39}{5} & -6/5 \\ 0 & 0 & \frac{125}{13} \end{pmatrix}$$

Όπως είδαμε πριν, $A = UL$ όπου

$$U := \begin{pmatrix} 1 & -2/13 & -1/5 \\ 0 & 1 & -1/5 \\ 0 & 0 & 1 \end{pmatrix}, L := \begin{pmatrix} \frac{125}{13} & 0 & 0 \\ -6/5 & \frac{39}{5} & 0 \\ -1 & -1 & 5 \end{pmatrix}.$$

Επομένως, λύνουμε ως εξής:

$$Ax = e_1 \rightarrow U(Lx) = e_1$$

αλλά, μετά τον υπολογισμό των παραγόντων U, L (με το συνηθισμένο κόστος της LU καθώς όλα τα στοιχεία είναι διαθέσιμα από την παραγοντοποίηση LU του PAP), το βήμα επίλυσης του $Uy = e_1$ γίνεται χωρίς πράξεις, καθώς άμεσα λαμβάνουμε ότι $y = e_1$. Αυτό, υπό κανονικές συνθήκες, θα χρειαζόταν τη λύση ενός άνω τριγωνικού συστήματος, που θα στοίχιζε n^2 πράξεις. Το επόμενο βήμα απαιτεί τη λύση του κάτω τριγωνικού συστήματος $Lx = y = e_1$ με το συνηθισμένο κόστος $O(n^2)$ σε πράξεις. Για τα παραπάνω δεδομένα, η απάντηση θα είναι

$$x = L^{-1}e_1 = [\frac{13}{125}, \frac{2}{125}, \frac{3}{125},]^\top.$$

□

V. *(20 β.)*

1. Να περιγράψετε συνοπτικά τη μέθοδο κανονικών εξισώσεων για την επίλυση του προβλήματος ελαχίστων τετραγώνων $\min_{x\in\mathbb{R}^n} \|Ax - b\|_2$, όπου $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ και $m \geq n$. Μπορείτε να θεωρήσετε ότι οι στήλες του A είναι γραμμικά ανεξάρτητες.

   Απάντηση. (Βιβλίο) 1) Πολλαπλασιάζετε τα δυο μέλη της εξίσωσης $Ax = b$ με $A^\top$: $A^\top Ax = A^\top b$. 2) Παραγοντοποιείτε με Cholesky το $A^\top A = LL^\top$ όπου $L \in \mathbb{R}^{n\times n}$. Αυτό γίνεται λόγω της γραμμικής αννεξαρτησίας των στηλών του A γιατί τότε το A είναι συμμετρικό θετικά ορισμένο. 3) Επιλύετε τα τριγωνικά συστήματα: $Ly = (A^\top b)$ ως προς y και το $Lx = y$ ως προς x. □

2. Να μετρήσετε το κόστος της μεθόδου σε πράξεις α.κ.υ. ως συνάρτηση των m και n (μόνον οι κυρίαρχοι όροι της πολυπλοκότητας να υπολογιστούν ακριβώς, για τους υπόλοιπους αρκεί να χρησιμοποιήσετε τάξη μεγέθους).

   Απάντηση. (Βιβλίο) Βήμα (1): $n^2/2(2m-1)$. Βήμα (2): $n^3/3$. Βήμα (3): $2n^2$. Συνολικά: $mn^2 + n^3/3 + O(n^2)$. □

3. Χρησιμοποιώντας ένα παράδειγμα με συγκεκριμένο A να δείξετε ότι η μέθοδος μπορεί να δώσει μη ικανοποιητικά αριθμητικά αποτελέσματα.

   Απάντηση. (Βιβλίο) Δείτε για παράδειγμα το μητρώο $A = [1+\delta, 1; 1, 1]$ όπου $\delta^2 < \varepsilon$ της μηχανής, οπότε $A^\top A = [1, 1; 1, 1]$ που είναι μη αντιστρέψιμο. □

IV. *(20 β.)*
Έστω η συνήθης διαφορική εξίσωση (ΣΔΕ) $\frac{du}{dt}(t) = -Au(t)$ όπου $A = [3/2, -1; -1, 3/2]$, $u = [u_1(t), u_2(t)]^\top$ και οι συναρτήσεις u_1, u_2 είναι επιλεγμένες ώστε $u_1(0) = 2$, $u_2(0) = 1$.

1. Να χρησιμοποιήσετε την εμπρός Euler με σταθερό βήμα διακριτοποίησης $h = 2$ για να υπολογίσετε την αριθμητική προσέγγιση της λύσης στο χρονικό σημείο $t = 6$.

   Απάντηση. στην εμπρός Euler η παραπάνω ΣΔΕ προσεγγίζεται ως $(U^{(j+1)} - U^{(j)})/h = -AU^{(j)}$. Επομένως ισχύει $U^{(j+1)} = (I - Ah)U^{(j)}$ όπου συμβολίζουμε με $U^{(j)}$ την προσέγγιση της τιμής του

$u(t_j) = u(t_0 + jh)$ μέσω της μεθόδου (που μπορεί να είναι αξιόπιστη ή όχι). Σε αριθμητική άπειρης ακρίβειας έχουμε ότι $U \in \mathbb{R}^2$. Επομένως

$$\begin{pmatrix} U_1^{(j+1)} \\ U_2^{(j+1)} \end{pmatrix} = (I - Ah)U^{(j)} = \begin{pmatrix} U_1^{(j)} \\ U_2^{(j)} \end{pmatrix} - h \begin{pmatrix} 3/2 & -1 \\ -1 & 3/2 \end{pmatrix} \begin{pmatrix} U_1^{(j)} \\ U_2^{(j)} \end{pmatrix}$$

Με βάση το παραπάνω υπολογίζουμε τη λύση στα βήματα $t = 2.4.6$ εκκινώντας από το 0:

$$\begin{pmatrix} U_1^{(j+1)} \\ U_2^{(j+1)} \end{pmatrix} = \begin{pmatrix} 1 - h3/2 & h \\ h & 1 - h3/2 \end{pmatrix} \begin{pmatrix} U_1^{(j)} \\ U_2^{(j)} \end{pmatrix} = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} \begin{pmatrix} U_1^{(j)} \\ U_2^{(j)} \end{pmatrix}$$

και κάνοντας τις πράξεις προκύπτουν οι τιμές:

$$U^{(2)} = [-2, 2]^\top, U^{(4)} = [8, -8]^\top, U^{(6)} = [-32.32]^\top.$$

□

2. Είναι γνωστό ότι η ακριβής λύση του παραπάνω συστήματος των ΣΔΕ τείνει στο 0 καθώς το $t \to \infty$. Με βάση αυτό το στοιχείο, να σχολιάσετε τη συμπεριφορά της παραπάνω προσέγγισης. (Υπόδειξη: Για μια πλήρη εξήγηση, είναι χρήσιμο να υπολογίσετε τις ιδιοτιμές του A.)

   Απάντηση. Παρατηρούμε ότι οι τιμές αυξάνουν (και θα συνεχίσουν να αυξάνουν) Αυτό είναι αντίθετο με ό,τι προβλέπεται για τη λύση της ΣΔΕ. (Προσέξτε επίσης από τα πρόσημα ότι έχουμε και ταλάντωση, που επίσης δεν συμβαίνει στη λύση της ΣΔΕ). Ο λόγος της αστάθειας εξηγείται ως εξής: Για να τείνει στο 0 η λύση θα πρέπει η απόλυτες τιμές των ιδιοτιμών του μητρώου να είναι μικρότερες του 1, δηλ. η φασματική ακτίνα του μητρώου να είναι φραγμένη αυστηρά από το 1. Οι ιδιοτιμές του μητρώου $I - hA$ θα είναι $1 - h\lambda(A)$ όπου με $\lambda(A)$ συμβολίζουμε τις ιδιοτιμές του A. Σημειώστε ότι είναι προτιμότερο να εξετάσουμε τις ιδιοτιμές ως συνάρτηση του h γιατί βοηθά και στην επιλογή βήματος με το οποίο δεν θα υπάρχει αστάθεια. Έχουμε επομένως ότι θα πρέπει $|1 - h\lambda(A)| < 1$ και επειδή οι ιδιοτιμές είναι πραγματικές (συμμετρικό μητρώο) θα πρέπει να έχουμε $2 > h/\lambda(A) > 0$. Επομένως το βήμα διακριτοποίησης πρέπει να ικανοποιεί $h < 2\lambda$. Για το συγκεκριμένο A οι ιδιοτιμές υπολογίζονται εύκολα (λύσεις του $(3/2 - \lambda)^2 - 1 = 0 \Rightarrow \lambda_1 = 1/2, \lambda_2 = 5/2$ επομένως θα πρέπει $h < 1$. □

3. Να εξηγήσετε με συντομία έναν τρόπο με τον οποίο θα μπορούσατε να υπολογίσετε τη λύση πιο αξιόπιστα.

   Απάντηση. Θα μπορούσαμε να χρησιμοποιήσουμε πίσω Euler ή να αλλάξουμε το βήμα h ώστε να ικανοποιούνται οι ανισότητες, δηλ. λαμβάνοντας h τέτοιο ώστε $h < 1$. □

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

Φεβρουάριος 2005

ΕΚΦΩΝΗΣΕΙΣ-ΑΠΑΝΤΗΣΕΙΣ-ΕΞΗΓΗΣΕΙΣ

Ι. *(25 β.)* Να απαντήσετε στα παρακάτω ερωτήματα (πρέπει να δικαιολογείτε τις απαντήσεις σας).

1. Να περιγράψετε, χρησιμοποιώντας MATLAB ή άλλη γλώσσα, τρεις διαφορετικούς αλγόριθμους υπολογισμού του $C = C + AB$ όπου $C \in \mathbb{R}^{m\times n}$, $A \in \mathbb{R}^{m\times p}$, $B \in \mathbb{R}^{p\times n}$ και τα μητρώα A, B, C είναι αποθηκευμένα, ως συνήθως, σε πίνακες κατάλληλων διαστάσεων ($m\times n, m\times p, p\times n$). Ο πρώτος θα πρέπει να βασίζεται σε DOT, ο δεύτερος σε SAXPY και ο τελευταίος σε ανανεώσεις 1ης τάξης. Η περιγραφή σας πρέπει να αναδεικνύει τις διαφορές των μεθόδων.

   Απάντηση.

   ```
   I) for i=1:m, for j=1:n, C(i,j) = C(i,j) + A(i,1:p)*B(1:p,j); end; end;
   II) for j=1:n, for k=1:p, C(1:m,j) = C(1:m,j) + A(1:m,k)*B(k,j); end; end;  □
   III) for j=1:p, C(1:m,1:n) = C(1:m,1:n) + A(1:m,j)*B(j,1:n); end;
   ```

2. Έστω ότι γνωρίζουμε την ακριβή λύση x^* ενός γραμμικού συστήματος $Ax = b$ και ότι λύνουμε ένα γραμμικό σύστημα μέσω διάσπασης LU με μερική οδήγηση. Έστω ότι η υπολογισμένη λύση είναι $\tilde{x}$ και διαπιστώνουμε α) μικρό δείκτη κατάστασης για το A (ως προς την επίλυση) αλλά β) εμπρός σχετικό σφάλμα $\|\tilde{x} - x^*\|/\|x^*\|$ κοντά στη μονάδα. Τι μπορούμε να συμπεράνουμε σχετικά με τον αλγόριθμο επίλυσης; (Υπόδειξη: Απαντήστε πολύ σύντομα και περιεκτικά.)

   Απάντηση. Όπως είναι γνωστό, το σχετικό εμπρός σφάλμα φράσσεται εκ των άνω από το γινόμενο του δείκτη κατάστασης επί το πίσω σφάλμα του αλγορίθμου. Επομένως, αν ο δείκτης κατάστασης είναι μικρός, κατ' ανάγκη θα πρέπει και το πίσω σφάλμα του αλγορίθμου επίλυσης να είναι μεγάλο. Επομένως είμαστε στην (αρκετά σπάνια) περίπτωση που το μητρώο κάνει τη διάσπαση LU με μερική οδήγηση να μην είναι πίσω ευσταθής. (Αυτό σημαίνει ότι ο παράγοντας U θα έχει μεγάλα στοιχεία σχετικά με τα στοιχεία του A. Προσοχή: Στόχος της ερώτησης αφορούσε και τον έλεγχο της κατανόησης του ότι «σχετικό σφάλμα κοντά στο 1» σημαίνει μεγάλη απώλεια ψηφίων, δηλ. μεγάλο σφάλμα σε σχέση με τον τρόπο που μετράμε το δείκτη κατάστασης και το πίσω σφάλμα.) □

ΙΙ. *(25 β.)* Έστω τα s υπερπροσδιορισμένα γραμμικά συστήματα $Ax_j = b_j$ όπου $A \in \mathbb{R}^{m\times n}$, $m > n$, $\mathrm{rank}(A) = n$ και $b_j \in \mathbb{R}^m$, $j = 1:s$.

~ 681-232

1. Να περιγράψετε με συντομία (κατά προτίμηση σε MATLAB) τη μέθοδο των κανονικών εξισώσεων για τον αποτελεσματικό υπολογισμό των λύσεων των παραπάνω συστημάτων και να εκτιμήσετε όσο μπορείτε καλύτερα το αριθμητικό κόστος **Ω**.

   Απάντηση. Θέτουμε $B = [b_1, ..., b_s] \in \mathbb{R}^{m\times s}$. Πολλαπλασιάζοντας τη σχέση $AX = B$ από α αριστερά με $A^\top$ λαμβάνουμε το $n\times n$ «σύστημα κανονικών εξισώσεων»:

$$A^\top AX = A^\top B$$

Το παραπάνω σύστημα κανονικών εξισώσεων έχει μοναδική λύση στη περίπτωσή μας αφού $\mathrm{rank}(A) = n$ (όλες οι στήλες του A είναι γραμμικά ανεξάρτητες), η οποία ισούται με:

$$X = (A^\top A)^{-1} A^\top B.$$

Ειδικότερα, το $X \in \mathbb{R}^{m\times s}$ περιέχει s στήλες, που θα είναι οι λύσεις των s συστημάτων. Το μητρώο $A^\top A$ είναι συμμετρικό και λόγω της γραμμικής ανεξαρτησίας των στηλών του A και θετικά ορισμένο. Επομένως, *μπορούμε να χρησιμοποιήσουμε διάσπαση* Cholesky. *Η εντολή που πραγματοποιεί τη διάσπαση αυτή είναι* η `R = chol(C)` όπου $C = A^\top A$ και R είναι άνω τριγωνικό μητρώο τέτοιο ώστε $R^\top \cdot R = A^\top A$. Η διάσπαση πραγματοποιείται μια φορά, ενώ ακολουθούν s εμπρός και πίσω αντικαταστάσεις με τους παράγοντες $R^\top$ και R. Για παράδειγμα, σε MATLAB μπορούμε να έχουμε:

```
G = A'*[A,B];   %% dhl. poll'ec pr'axeic BLAS-3
R = chol(G(:,1:n));
X = R \ (R' \ G(:,n+1:s));
```

Σημειώστε ότι λόγω της συμμετρίας, θα μπορούσαμε να μειώσουμε τις πράξεις α.κ.υ. στον υπολογισμό του $A^\top A$ στο μισό. Το αριθμητικό κόστος είναι το άθροισμα των εξής επιμέρους ποσοτήτων:

- Προσδιορισμός του $A^\top A$ και $A^\top B$. Για τον πρώτον όρο, αρκεί να υπολογίσουμε μόνο τα $\frac{n(n+1)}{2}$ στοιχεία του κάτω τριγωνικού τμήματος, οπότε για κάθε τέτοιο στοιχείο εκτελούμε ένα εσωτερικό γινόμενο μήκους m, και το συνολικό κόστος του προσδιορισμού αυτού θα είναι $\frac{n(n+1)(2m-1)}{2} + ns(2m-1)$.
- Διάσπαση Cholesky. Η διάσπαση απαιτεί τις μισές περίπου πράξεις από αυτές που απαιτεί η διάσπαση LU. Επομένως το κόστος είναι $\frac{n^3}{3} + O(n^2)$.
- Επίλυση με εμπρος και πίσω αντικατάσταση των s συστημάτων $R^T R x_j = A^T b_j$ για $j = 1:s$. Κάθε λύση με τριγωνικό σύστημα κοστίζει n^2 πράξεις επομένως συνολικά θα έχουμε $2sn^2$ πράξεις.

Το αριθμητικό κόστος του αλγορίθμου θα είναι:

$$\Omega = \frac{n^3}{3} + \frac{n(n+1)(2m-1)}{2} + sn(2m-1) + 2sn^2 + O(n^2) = \frac{n^3}{3} + mn^2 + 2mns + 2sn^2 + \text{όροι χαμηλότερης τάξης}$$

□

2. *Να αναδείξετε ένα (καταστροφικό) μειονέκτημα της μεθόδου χρησιμοποιώντας το μητρώο*

$$A = \begin{pmatrix} 1 & 1 \\ \delta & 0 \\ 0 & \delta \end{pmatrix}$$

Η στήλη του A είναι γραμμικά
ανεξάρτητες επομένως $A^T A$ είναι
ΘΟ και αντιστρέψιμο

και $0 < \delta \leq \sqrt{\varepsilon_M}$ (όπου ε_M είναι το έψιλον της μηχανής).

Απάντηση. Παρατηρείστε ότι το μητρώο A έχει τάξη 2, εφόσον $\delta > 0$, επομένως κάθε σύστημα $A^\top A x = A^\top b$ θα πρέπει να έχει μοναδική λύση *σε αριθμητική άπειρης ακρίβειας*. Όμως:

$$A^\top A = \begin{pmatrix} 1 & \delta & 0 \\ 1 & 0 & \delta \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ \delta & 0 \\ 0 & \delta \end{pmatrix} = \begin{pmatrix} 1+\delta^2 & 1 \\ 1 & 1+\delta^2 \end{pmatrix}$$

Εφόσον $\delta \leq \sqrt{\varepsilon_M}$, έχουμε οτι $\delta^2 \leq \varepsilon_M$, επομένως $fl(1+\delta^2) = 1$. Στη περίπτωση λοιπόν αυτή το μητρώο $A^\top A$ θα έχει δύο ίδιες στήλες, δηλαδή τάξη 1 και το μητρώο $A^\top A$ δεν θα είναι αντιστρέψιμο. (Αξίζει να σημειώσετε επίσης ότι ένα από τα μειονεκτήματα της μεθόδου κανονικών εξισώσεων για τη λύση προβλημάτων είναι ότι το εμπρός σφάλμα εξαρτάται από το δείκτη κατάστασης του $A^\top A$. Για παράδειγμα, αν $A = A^\top$ τότε $\kappa_2(A^\top A) = \kappa_2(A^2) = \kappa_2(A)^2$.) □

III. *(25 β.)*

1. Δίδεται το μιγαδικό μητρώο $G \in \mathbb{C}^{n\times n}$ για το οποίο γνωρίζουμε ότι μπορεί να γραφεί ως $G = iA + I$ όπου το $A \in \mathbb{R}^{n\times n}$ είναι πραγματικό, Σ.Θ.Ο. και τριδιαγώνιο (όπως στο 1ο μέρος). Έστω επίσης ότι πρέπει να λύσουμε το σύστημα $Gx = b$ όπου $b \in \mathbb{R}^n$ και $x \in \mathbb{C}^n$ με $x = x_R + ix_I$ όπου τα διανύσματα x_R, x_I είναι το πραγματικό και φανταστικό μέρος του x αντίστοιχα.

   α) Να δείξετε ότι το παραπάνω πρόβλημα είναι ισοδύναμο με τη λύση ενός πραγματικού γραμμικού συστήματος διπλάσιας διάστασης (δηλ. $2n \times 2n$), που μπορεί να γραφτεί ως $SX = B$ όπου:

$$S = \begin{pmatrix} \square & \square \\ \square & \square \end{pmatrix}, X = \begin{pmatrix} x_R \\ x_I \end{pmatrix}, \quad B = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

   όπου τα □ πρέπει να συμπληρωθούν από εσάς κατάλληλα ώστε να ισχύει η ισοδυναμία.

Απάντηση. $Gx = b \Rightarrow (iA+I)(x_R+ix_I) = b \Rightarrow x_R - Ax_I = b$ και $Ax_R + x_I = 0$ επομένως: □

$$S = \begin{pmatrix} I & -A \\ A & I \end{pmatrix}$$ σελ. 172

β) Να συμπληρώσετε τα □ στην παρακάτω Block LU διάσπαση του μητρώου S:

$$S = \begin{pmatrix} I & 0 \\ \square & I \end{pmatrix} \begin{pmatrix} \square & \square \\ 0 & \square \end{pmatrix}$$

Απάντηση. Θα έχουμε

$$S = \begin{pmatrix} I & 0 \\ A & I \end{pmatrix} \begin{pmatrix} I & -A \\ 0 & I-A^2 \end{pmatrix}$$ συμπλήρωμα Schur

Έστω ότι ο αριστερός πολλαπλασιαστής είναι $[I,0;L_{21},I]$ και ο δεξιός $[U_{11},U_{12};0,U_{22}]$. Το ζητούμενο προκύπτει εξισώνοντας την αριστερή με τη δεξιά πλευρά ως προς τα υπομητρώα που προκύπτουν από τους πολλαπλασιασμούς των ομαθών

$$\begin{aligned} I &= U_{11}, \quad -A = U_{12} \\ A &= L_{21}U_{11} \Rightarrow L_{21} = A \\ I &= L_{21}U_{12} + U_{22} \Rightarrow I = -A^2 + U_{22} \end{aligned}$$

Ο τελευταίος όρος είναι το συμπλήρωμα Schur του S για το συγκεκριμένο τεμαχισμό. □

β) Να περιγράψετε συνοπτικά τα βήματα για την επίλυση του συστήματος $SX = B$ χρησιμοποιώντας την προηγούμενη διάσπαση το S.

Απάντηση. Η διαδικασία θα είναι να λύσουμε πρώτα το

Μέθοδος LU

$$\begin{pmatrix} I & 0 \\ A & I \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

1. Υπολογισμός L, U

αξιοποιώντας την τριγωνική κατά ομαθούς μορφή. Θα έχουμε:

2. $Ly = b$

$$y_1 = b, Ay_1 + y_2 = 0 \Rightarrow y_2 = -Ab.$$

3. $Ux = y$

Μετά λύνουμε:

$$\begin{pmatrix} I & -A \\ 0 & I+A^2 \end{pmatrix} \begin{pmatrix} x_R \\ x_I \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} b \\ -Ab \end{pmatrix}$$

Επομένως

$$x_I = -(I+A^2)^{-1}Ab, \quad x_R = b + Ax_I = b - A(I+A^2)^{-1}Ab$$

και αν θέλουμε, με κάποιες απλοποιήσεις:

$$x_R = b - (I+A^2)^{-1}A^2 b = (I+A^2)^{-1}((I+A^2) - A^2)b = (I+A^2)^{-1}b.$$

□

IV. *(25 β.)* Έστω η συνάρτηση $y : \mathbb{R} \to \mathbb{R}$ της οποίας οι παράγωγοι μέχρι 4ης τάξης υπάρχουν και είναι συνεχείς στο διάστημα $[0,1]$. Η συνάρτηση ικανοποιεί τη συνήθη διαφορική εξίσωση:

$$y''(t) - 4t \cdot y'(t) = 16t$$

για $t \in (0,1)$. Θέλουμε να προσδιορίσουμε τη λύση της παραπάνω εξίσωσης, εφαρμόζοντας πλέγμα αποτελούμενο από 3 ισαπέχοντες εσωτερικούς κόμβους στο διάστημα $(0,1)$. Οι συνοριακές συνθήκες που διαθέτουμε είναι $y(0) = 1$ και $y(1) = 0$. Η διακριτοποίηση των διαφορικών τελεστών θα γίνει με χρήση κεντρισμένων πεπερασμένων διαφορών 2ης τάξης.

1. Να υπολογίσετε τους συντελεστές του μητρώου και του δεξιού μέλους που προκύπτουν από τη διακριτοποίηση της Δ.Ε.

Απάντηση. Στη περίπτωσή μας το διάστημα μεταξύ 2 διαδοχικών κόμβων θα είναι $h = 1/4 = 0.25$. Οι ζητούμενοι κόμβοι είναι οι $t_k = 0 + k \cdot h$, $k = 1,2,3$. Για τη διακριτοποίηση των διαφορικών τελεστών χρησιμοποιούμε τις προσεγγίσεις:

$$y'(t_k) \approx \frac{y_{k+1} - y_{k-1}}{2h} = 2(y_{k+1} - y_{k-1})$$

και:

$$y''(t_k) = y_k'' \approx \frac{y_{k+1} - 2y_k + y_{k-1}}{h^2} = 16(y_{k+1} - 2y_k + y_{k-1})$$

Οι διακριτοποιημένες εξισώσεις γράφονται, ως προς τους αγνώστους Y_1, Y_2, Y_3 ως εξής:

$$16t_k = 16(Y_{k+1} - 2Y_k + Y_{k-1}) - 8t_k(Y_{k+1} - Y_{k-1}), \qquad k = 1,2,3.$$

Επειδή $y(0) = Y_0 = 1, y(1) = 0 = Y_4$, οι συνοριακές συνθήκες επιδρούν μόνο στον 1ο όρο του δεξιού μέλους. Λαμβάνοντας υπόψη τις συνοριακές συνθήκες, προκύπτει το παρακάτω τριδιαγώνιο σύστημα μεγέθους 3:

$$\begin{pmatrix} -4 & 1.75 & 0 \\ 2.50 & -4 & 1.5 \\ 0 & 2.75 & -4 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} -1.75 \\ 1 \\ 1.5 \end{pmatrix}$$

□

2. Να γράψετε κώδικα MATLAB που να υλοποιεί την αρχικοποίηση του μητρώου και διανύσματος που προκύπτουν από τη παραπάνω διακριτοποίηση για *n* συνολικά εσωτερικούς κόμβους. Το μητρώο να επιστρέφεται σε 2-διάστατο πίνακα *A* και το διάνυσμα σε πίνακα-στήλη *F*. Ο κώδικάς θα πρέπει να είναι γραμμένος έτσι ώστε να επιτυγχάνεται καλή επίδοση κατά την εκτέλεσή του.

Απάντηση. Ο κώδικας ακολουθεί:

```
h = 1/(n+1); ih2 = 1/h^2; i2h = 1/(2*h);
tk = h*[1:n]';
a = -2*ih2*ones(n,1);
a_l = ih2*ones(n-1,1)+4*i2h*tk(2:n);
a_r = ih2*ones(n-1,1)-4*i2h*tk(1:n-1);
A = diag(a,0) + diag(a_l,-1) + diag(a_r,1);
F = 16*tk;
F(1) = F(1)-ih2-4*i2h*tk(1); % Sunoriak'h sunj'hkh
```

□

Ονοματεπώνυμο: .... Α.Μ. ....: Έτος ...

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι *Εξεταστική Φεβρουαρίου 2006*

ΑΠΑΝΤΗΣΕΙΣ

Ι. (≈ *40 β.*)

1. Πότε λέμε ότι ένα σύστημα αριθμητικής κινητής υποδιαστολής ικανοποιεί την αρχή ακριβούς στρογγύλευσης;

   Απάντηση. Θεωρία: Όταν το αποτέλεσμα μιας στοιχειώδους αριθμητικής πράξης στην α.κ.υ. (επί τέτοιων αριθμών) επιστρέφει το αποτέλεσμα που θα προέκυπτε αν η πράξη γινόταν σε ακριβή αριθμητική (άπειρης ακρίβειας) και μετά χρησιμοποιείτο στρογγύλευση. □

   Το υπολογισμένο αποτέλεσμα είναι ακριβώς το ίδιο με το αποτέλεσμα θα είχαμε αν χρησιμοποιούσαμε αριθμητική άπειρης ακρίβειας με στοιχεία εισόδου λίγο παραλλαγμένα

2. Να ορίσετε με σαφήνεια τι σημαίνει πίσω ευστάθεια στον υπολογισμό της τιμής μιας συνάρτησης $f : \mathbb{R}^n \to \mathbb{R}^m$ και να εξηγήσετε πώς αν αποδείξουμε την πίσω ευστάθεια ενός αλγορίθμου μπορούμε να αναγάγουμε την εύρεση της απόστασης του $f(x)$ από το υπολογισμένο $f_{\text{prog}}(x)$ σε ένα πρόβλημα διαταραχών.

   Απάντηση. Αν ο υπολογισμός είναι πίσω ευσταθής τότε μπορούμε να βρούμε διάνυσμα έστω $x_{\text{prog}} \in \mathbb{R}^n$, που είναι κοντά στο x, τέτοιο ώστε $f(x_{\text{prog}}) \equiv f_{\text{prog}}(x)$, δηλ. ο υπολογισμός της συνάρτησης με το συγκεκριμένο αλγόριθμο σε α.κ.υ. δίνει το ίδιο ακριβώς αποτέλεσμα με τον υπολογισμό της συνάρτησης στο x_{prog}, με αριθμητική άπειρης ακρίβειας. Αν δείξουμε κάτι τέτοιο, ισχύει ότι

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

   επομένως η απόσταση (το απόλυτο εμπρός σφάλμα) των δυο θα είναι ίση με την απόσταση της τιμής του f στο x από την τιμή του f σε ένα κοντινό σημείο x_{prog}, και η μελέτη της, εφόσον το x είναι κοντά στο x_{prog}, αντιστοιχεί σε πρόβλημα διαταραχών. □

3. Έστω το μιγαδικό μητρώο $A \in \mathbb{C}^{n\times n}$ και ότι πρέπει να υπολογίσουμε το γινόμενο AB, όπου $B \in \mathbb{C}^{n\times k}$. Στη συνέχεια υποθέτουμε ότι τα k και n είναι πολύ μεγαλύτερα του 1. α) Να υπολογίσετε το Ω (πλήθος πράξεων *πραγματικών* α.κ.υ.) και το $\Phi_{\min}$ (πλήθος ελάχιστων μεταφορών *πραγματικών α.κ.υ.*) μεταξύ μνήμης και επεξεργαστή/καταχωρητών/κρυφής μνήμης για τον υπολογισμό του C μέσω της παραπάνω έκφρασης. β) Να δείξετε (επιχειρηματολογόντας βάσει των τιμών για τα $\Phi_{\min}$ και Ω, γιατί οι παραπάνω πράξεις μπορούν να θεωρηθούν ως BLAS-3.

   Απάντηση. α) Από την εκφώνηση φαίνεται ότι μπορούμε να γράψουμε

   $$C = AB = (A_R B_R - A_I B_I) + i(A_R B_I + A_I B_R).$$

   όπου $A = A_R + iA_I$, $B = B_R + iB_I$, $A_R, A_I \in \mathbb{R}^{n\times n}$ και $B_R, B_I \in \mathbb{R}^{n\times k}$. Καθένας από τους παραπάνω πολλαπλασιασμούς στοιχίζει $nk(2n-1)$ πράξεις. Επομένως, το συνολικό κόστος θα είναι

   $$\Omega = 4nk(2n-1) + 2nk.$$

όπου ο τελευταίος όρος οφείλεται στο ότι πρέπει επίσης να προστεθούν τα ενδιάμεσα αποτελέσματα (δηλ. τα $A_R B_I + A_I B_R, A_R B_R - A_I B_I$. Συνολικά, θα είναι $\Omega = 8n^2k - 2nk$. Εναλλακτικά, μπορείτε να πείτε ότι ο πολλαπλασιασμός βαθμωτών μιγαδικών στοιχίζει 4 πολλαπλασιασμούς και 2 προσθέσεις πραγματικών (φαίνεται στον παραπάνω τύπο αν θεωρήσετε ότι όλα τα στοιχεία είναι βαθμωτοί) δηλ. 6 πράξεις πραγματικών, ενώ μια πρόσθεση 2 μιγαδικών στοιχίζει όσο 2 προσθέσεις πραγματικών. Ο πολλαπλασιασμός μιγαδικών μητρώων μεγέθους $n \times n$ με $n \times n$ στοιχίζει nk φορές n πολλαπλασιασμούς μιγαδικών και $nk - 1$ φορές προσθέσεις μιγαδικών. Επομένως το συνολικό κόστος θα είναι $nk6n + nk(2nk - 2) = 8n^2k - 2nk$, όπως και πριν. Αφού κάθε μεταφορά μιγαδικού ισοδυναμεί με 2 μεταφορές πραγματικών, το $\Phi_{\min} = 2(n^2 + 2nk) = 2n^2 + 4nk$.

β) Παρατηρούμε ότι το $\mu_{\min} = (2n^2 + 4nk)/(8n^2k - 2nk) \approx O(1/k)$. Επίσης γνωρίζουμε ότι $k \gg 1$. Επομένως, η πράξη παρουσιάζει τοπικότητα αντίστοιχη με εκείνη των πράξεων BLAS-3. □

4. Έστω ο βρόχος `for i=1:n, z(i) = a*x(i)+y(i); end`. Να τον ξαναγράψετε χρησιμοποιώντας ξετύλιγμα μήκους 3 και που να λειτουργεί σωστά ανεξάρτητα από την τιμή του n.

   Απάντηση. Θα θεωρήσουμε, όπως στη MATLAB, ότι η συνάρτηση $\text{rem}(x, y)$ επιστρέφει το υπόλοιπο της διαίρεσης δυο αριθμών x, y. Ο βρόχος θα είναι ως εξής:

```
m = rem(n,3); for j=1:m, z(i) = a*x(i)+y(j); end
for j=m+1:3:n
 z(i)  = a*x(i)+y(j);
 z(i+1) = a*x(i+1)+y(j+1);
 z(i+2) = a*x(i+2)+y(j+2);
end
```

□

5. Η συνάρτηση `rank` της MATLAB εκτιμά το πλήθος των ιδιαζουσών τιμών ενός μητρώου που είναι μεγαλύτερες ενός όρου `tol` που ορίζεται ως

$$\texttt{max(size(A))} * \texttt{norm(A)} * \texttt{eps}.$$

   Έστω ότι ο υπολογισμός αφορά ένα μητρώο A που δίνεται σε μορφή γινομένου $A = DQ$, όπου το $Q \in \mathbb{R}^{7\times 7}$ είναι ορθογώνιο και το D διαγώνιο με στοιχεία

$$D = \text{diag}[100, 1, 0.00001, 1e-10, 2048, 1e-32, 0]$$

   και ότι χρησιμοποιείται αριθμητική IEEE διπλής ακρίβειας. Να υπολογίσετε και να εξηγήσετε την τιμή που θα επιστρέψει το `rank`(A).

Απάντηση. Το D είναι διαγώνιο και το Q ορθογώνιο. Γράφοντας $A = ID(Q^\top)^\top$, όπου I είναι ταυτοτικό μητρώο, φαίνεται ότι το D είναι το διαγώνιο μητρώο που περιέχει τις ιδιάζουσες τιμές, το $Q^\top$ περιέχει τα δεξιά ιδιάζοντα διανύσματα, ενώ τα αριστερά ιδιάζοντα διανύσματα είναι τα διανύσματα της τυπικής βάσης $e_1, \ldots, e_7$. Επίσης, norm$(A) = 2048$ καθώς η συνάρτηση επιστρέφει τη 2-νόρμα του A που είναι η μέγιστη ιδιάζουσα τιμή του• εναλλακτικά, φαίνεται από το ότι η 2-νόρμα είναι αμετάβλητη αν πολλαπλασιάσουμε το μητρώο με ορθογώνιο μητρώο, επομένως $\|A\|_2 = \|AQ^\top\|_2 = \|D\|_2$. Επίσης `max(size(A))` $= 7$, $eps = \epsilon_M = 2^{-52}$ (το έψιλον της μηχανής για α.κ.υ. IEEE διπλής ακρίβειας). Επομένως

$$max(size(A)) * norm(A) * eps = 2^{10}(2^3 - 1)2^{-52} \approx 2^{-39} \approx 10^{-12}.$$

Υπάρχουν 5 τιμές του D που είναι μεγαλύτερες του 10^{-12}, επομένως η απάντηση είναι 5. □

II. ($\approx$ *30 β.*) Έστω το διάνυσμα $e = [1, 4, \texttt{zeros}(1, 6), 2, -2]^\top$ και ότι θέλουμε να υπολογίσουμε μητρώα $M \in \mathbb{R}^{10\times 10}$ και $H \in \mathbb{R}^{10\times 10}$ τα οποία έχουν τις παρακάτω ιδιότητες:

α) Το M είναι κάτω τριγωνικό με μονάδες στη διαγώνιο και τέτοιο ώστε $M\ e = \texttt{eye}(10, 1)$. β) Το H είναι ορθογώνιο και τέτοιο ώστε $H\ e = \gamma\ \texttt{eye}(10, 1)$ για κάποιο βαθμωτό γ.

1. Να υπολογίσετε τα στοιχεία τα μητρώα M και H καθώς και το γ για το συγκεκριμένο e. Προτείνεται να τα υπολογίσετε στη μορφή $I + \tau u v^\top$ όπου $u, v \in \mathbb{R}^{10}$ και τ βαθμωτός. Πρέπει βέβαια να υπολογίσετε τα τ, u, v για καθένα από τα M και H. Μερικά από αυτά μπορεί να είναι τετριμμένα, π.χ. $\tau = -1$, και να μην εξαρτώνται από το διάνυσμα e για το μερικό μηδενισμό του οποίου κατασκευάστηκαν.

   Απάντηση. Πρόκειται για το στοιχειώδες μητρώο Gauss και το το στοιχειώδη ανακλαστή Householder αντίστοιχα. Ειδικότερα, το $M = I - be_1^\top$ όπου $b = [0; e_{2:10}] = [0, \texttt{zeros}(1, 6), 2, -2]^\top$ θα είναι στη ζητούμενη μορφή. Το $H = I - 2uu^\top/u^\top u$ όπου $u = e + \|e\|_2 e_1$ και καθώς εύκολα φαίνεται ότι $\|e\|_2 = 5$, $u = [6, 4, \texttt{zeros}(1, 6), 2, -2]^\top$. Επομένως $H = I - \frac{2}{60}uu^\top$ που είναι στη ζητούμενη μορφή. Τέλος, επαληθεύεται άμεσα ότι $He = -5e_1$. □

2. Δίδεται το διάνυσμα $z \in \mathbb{R}^{10}$. Να δείξετε και να γράψετε πώς μπορεί να υπολογιστεί φθηνά, χρησιμοποιώντας αποκλειστικά πράξεις BLAS-1, το $(M\ H)^{-1}z$.

   Απάντηση.

   $$(M\ H)^{-1}z = H^{-1}M^{-1}z = H^\top M^{-1}z$$

   αλλά επίσης επαληθεύεται άμεσα ότι $M^{-1} = I + be_1^\top$ και $H^\top = H$, επομένως

   $$(M\ H)^{-1}z = H(I + be_1^\top)z = Hz + Hb(e_1^\top z) = Hz + Hb\zeta_1 = H * (z + \zeta_1 b),$$

   όπου ζ_1 είναι το πρώτο στοιχείο του z. Επομένως, το διάνυσμα $\hat{z} = z + \zeta_1 b$ μπορεί να κατασκευαστεί καλώντας μια πράξη `saxpy`. Αντικαθιστώντας την ειδική μορφή του H, θα έχουμε

   $$(I - \frac{1}{30}uu^\top)\hat{z} = \hat{z} - \frac{u^\top \hat{z}}{30}u.$$

   Συνεπάγεται ότι μπορεί να υπολογιστεί καλώντας μια φορά το `dot` για το $u^\top\hat{z}$ και μια πράξη `saxpy` για τα υπόλοιπα. □

3. Να υπολογίσετε τον (άνω τριγωνικό) παράγοντα R (και όχι το Q) της παραγοντοποίησης QR του μητρώου $A = \texttt{tril}(\texttt{ones}(4, 3))$ με βάση τα παραπάνω.

   Απάντηση. Θα χρησιμοποιήσουμε ανακλαστές (στοιχειώδεις μετασχηματισμούς Householder) που έχουν την ίδια μορφή με την προηγούμενη ερώτηση.

Ειδικότερα, αν θέσουμε τώρα $e = [1, 1, 1, 1]^T$, τότε $H_1 = I - 2u_1u_1^T/u_1^Tu_1$ όπου $u_1 = e + \|e\|_2 e_1 = [3, 1, 1, 1]^T$ και $H_1 e = -\|e\|_2 e_1$. Επομένως $H_1 e = -2e_1$ και $H_1 = I - \frac{2}{12} u_1 * u_1^T$. Επίσης,

$$H_1 A = \begin{pmatrix} -2.0000 & -1.5000 & -1.0000 \\ 0.0000 & 0.5000 & -0.3333 \\ 0.0000 & 0.5000 & 0.6667 \\ 0.0000 & 0.5000 & 0.6667 \end{pmatrix}.$$

Θέτουμε

$$H_2 = I - 2u_2u_2^T/u_2^Tu_2, \quad u_2 = 0.5[0, 1, 1, 1]^T + \sqrt{3}e_2 = [0, 0.5(1 + \sqrt{3}), 0.5, 0.5]^T.$$

και

$$H_2 H_1 A = \begin{pmatrix} -2.0000 & -1.5000 & -1.0000 \\ 0 & -0.8660 & -0.5774 \\ 0 & 0 & 0.5774 \\ 0 & 0 & 0.5774 \end{pmatrix}.$$

Στο τελευταίο βήμα θα έχουμε:

$$H_3 = I - 2u_3u_3^T/u_3^Tu_3, \quad u_3 = 0.5774[0, 0, 1, 1]^T + \sqrt{2}e_3 = [0, 0, 0.5774(1 + \sqrt{2}), 0.5774]^T.$$

$$R = H_3 H_2 H_1 A = \begin{pmatrix} -2.0000 & -1.5000 & -1.0000 \\ 0 & -0.8660 & -0.5774 \\ 0 & 0 & -0.8165 \\ 0 & 0 & 0 \end{pmatrix}.$$

□

III. ($\approx$ *30 β.*) Έστω η συνάρτηση $u : \mathbb{R} \to \mathbb{R}$ για την οποία γνωρίζουμε ότι οι παράγωγοι $u^{(j)}(x)$, $j = 1, ..., 4$ είναι συνεχείς για $x \in [-1, 1]$. Γνωρίζουμε επίσης ότι $u(-1) = 0$, $u(1) = 1$ *και ότι η u ικανοποιεί τη διαφορική εξίσωση*

$$-u^{(2)}(x) + \mu u^{(1)}(x) + u(x) = x^2 \quad \forall x \in (-1, 1).$$

Η ακριβής τιμή του βαθμωτού μ δεν είναι γνωστή αλλά εξαρτάται από το πρόβλημα. Θέλουμε να προσεγγίσουμε τη λύση αριθμητικά, λύνοντας γραμμικό σύστημα που προκύπτει από κεντρισμένες πεπερασμένες διαφορές δεύτερης τάξης για την προσέγγιση των παραγώγων.

1. Να διακριτοποιήσετε την εξίσωση χρησιμοποιώντας πλέγμα ισαπεχόντων κόμβων $x_0 = 0, x_1 = -1 + h, x_2 = -1 + 2h, ..., x_6 = 1$ στο διάστημα $(-1, 1)$, όπου h είναι η απόσταση μεταξύ διαδοχικών κόμβων (οι κόμβοι $x_0 = -1$, $x_6 = 1$ αντιστοιχούν στα άκρα του διαστήματος). Να γράψετε προσεκτικά το γραμμικό σύστημα $Aw = b$ που πρέπει να λυθεί. Τα στοιχεία των A και b πρέπει να είναι όσο πιο απλοποιημένα γίνεται (απλές αριθμητικές εκφράσεις, ενδεχομένως εξαρτόμενες από το μ.)

   Απάντηση. Το πλέγμα έχει 5 κόμβους (εκτός των ακραίων σημείων -1 και 1) επομένως θα έχουμε ότι $h = (x_6 - x_0)/(5 + 1) = 1/3$. Θα προκύψει ένα σύστημα 5 εξισώσεων με 5 αγνώστους, που θα είναι τιμές $U_1, ..., U_5$ που θα προσεγγίζουν τη συνάρτηση u στους κόμβους $x_1, ..., x_5$ του πλέγματος. Οι κεντρισμένες πεπερασμένες διαφορές δεύτερης τάξης για την προσέγγιση των παραγώγων υπάρχουν (λόγω των συνεχών παραγώγων ως και την 4η τάξη) και θα είναι (από τη θεωρία):

$$u^{(1)}(x_j) \approx \frac{U(x_{j+1}) - U(x_{j-1})}{2h}, \quad u^{(2)}(x_j) \approx \frac{U(x_{j+1}) - 2U(x_j) + U(x_{j-1})}{h^2}.$$

Προσέξτε επίσης ότι οι συντελεστές του u και των παραγώγων του στο αριστερό μέλος της διαφορικής εξίσωσης δεν εξαρτώνται από το x, επομένως, το μητρώο που θα προκύψει, θα είναι τριδιαγώνιο Toeplitz (άρα αρκεί να υπολογίσουμε τρεις όρους). Αντικαθιστώντας στην εξίσωση θα έχουμε:

$$-\frac{U(x_{j+1}) - 2U(x_j) + U(x_{j-1})}{h^2} + \mu\frac{U(x_{j+1}) - U(x_{j-1})}{2h} + U_j = x_j^2, \quad j = 1, \dots, 5.$$

Συγκεντρώνοντας τους συντελεστές θα έχουμε:

$$(-\frac{1}{h^2} - \frac{\mu}{2h})U(x_{j-1}) + (\frac{2}{h^2} + 1)U(x_j) + (-\frac{1}{h^2} + \frac{\mu}{2h})U(x_{j+1}) \quad = \quad (-1 + jh)^2,$$

δηλ.

$$(-9 - \frac{3\mu}{2})U(x_{j-1}) + 19U(x_j) + (-9 + \frac{3\mu}{2})U(x_{j+1}) \quad = \quad (-1 + j/3)^2.$$

Προσέξτε ότι οι εξισώσεις για τα σημεία στο σύνορο (x_1, x_5) θα έχουν τη μορφή:

$$19U(x_1) + (-9 + \frac{3\mu}{2})U(x_2) \quad = \quad 4/9$$

επειδή $U_0 = 0$, και

$$(-9 - \frac{3\mu}{2})U(x_4) + 19U(x_5) = 4/9 - (-9 + \frac{3\mu}{2}),$$

επειδή $U_6 = 1$. Επομένως, μπορούμε να γράψουμε το σύστημα $AU = b$, όπου:

$$\begin{aligned} A &= \text{toeplitz}([19, -9 - \frac{3\mu}{2}, 0, 0, 0], [19, -9 + \frac{3\mu}{2}, 0, 0, 0]), \\ b &= [4/9, 1/9, 0, 1/9, 4/9 - (-9 + \frac{3\mu}{2})]^T. \end{aligned}$$

Το A μπορείτε να το γράψετε και ως

$$A \quad = \quad \text{trid}_5[-9 - \frac{3\mu}{2}, 19, -9 + \frac{3\mu}{2}]$$

□

2. *Έστω ότι εφαρμόζετε απαλοιφή Gauss με μερική οδήγηση για την παραγοντοποίηση* του *παραπάνω* συστήματος. Να υπολογίσετε σε ποιο ή σε ποιά πραγματικά διαστήματα επιτρέπεται να βρίσκεται ο μ ώστε να *μη* χρειαστεί *εναλλαγή* γραμμών *στο πρώτο βήμα* της απαλοιφής (δηλ. εκείνο κατά το οποίο μηδενίζονται τα στοιχεία στις θέσεις 2 και κάτω της πρώτης στήλης του A).

Απάντηση. Για να μη χρειάζεται εναλλαγή γραμμών στο 1ο βήμα, αρκεί το στοιχείο στη θέση (1, 1) του μητρώου να είναι το μέγιστο στοιχείο, σε απόλυτη τιμή, της 1ης στήλης. Καθώς το μητρώο είναι τριδιαγώνιο, αρκεί

$$19 \geq |-9 - \frac{3\mu}{2}|$$

επομένως πρέπει να ισχύει

$$-19 \leq 9 + \frac{3\mu}{2} \leq 19 \Rightarrow \frac{-56}{3} \leq \mu \leq \frac{20}{3}.$$

□

3. *Έστω ότι ισχύουν οι παραπάνω περιορισμοί και ότι εκτελείται το πρώτο βήμα* χωρίς εναλλαγή γραμμών. Όπως γνωρίζετε από τη θεωρία, το αρχικό σύστημα θα μετατραπεί σε ένα ισοδύναμο της μορφής $A^{(1)}w = Mb$ όπου το M (μετασχηματισμός Gauss) είναι επιλεγμένο έτσι ώστε το $A^{(1)} = MA$ να είναι 0 στις θέσεις 2 και κάτω της πρώτης στήλης. Να δείξετε ότι το υπομητρώο στις θέσεις $(2:5, 2:5)$ του $A^{(1)}$ μπορεί να γραφτεί ως $G - \frac{1}{\sigma}e_1e_1^\top$ όπου $G \in \mathbb{R}^{4\times4}$, $e_1 = [1, 0, 0, 0]^\top$ και σ βαθμωτός (ενδεχομένως συνάρτηση του μ), και να δείξετε ποια είναι τα στοιχεία του G και το σ. (Προσοχή: Δεν χρειάζεται να υπολογίσετε τα στοιχεία του $A^{(1)}$.)

Απάντηση. Προκύπτει άμεσα ότι το ζητούμενο μητρώο θα είναι το

$$A_{2:5,2:5} - A_{2:5,1}A_{1,1}^{-1}A_{1,2:5},$$

δηλ. $G = A_{2:5,2:5}$. Όμως επειδή το μητρώο είναι τριδιαγώνιο,

$$A_{2:5,1} = -(9 + \frac{3\mu}{2})e_1, \quad A_{1,2:5} = (-9 + \frac{3\mu}{2})e_1^\top$$

και $A_{1,1} = 19$. Επομένως, το ζητούμενο θα είναι

$$A_{2:5,2:5} - \frac{1}{19}(9 + \frac{3\mu}{2})(9 - \frac{3\mu}{2})e_1e_1^\top$$

επομένως

$$\sigma = 19/(81 - \frac{9}{4}\mu^2)$$

□

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι *ΑΠΑΝΤΗΣΕΙΣ Εξεταστική Σεπτεμβρίου 2006*

I.

1. Να αναφέρετε ένα παράδειγμα πράξης BLAS-3. Τι πλεονέκτημα έχουμε όταν υλοποιούμε αλγορίθμους με πράξεις BLAS-3 αντί με εναλλακτικές υλοποιήσεις με άλλες πράξεις BLAS;

   Απάντηση. Πολλαπλασιασμός μητρώων. Μεγαλύτερη τοπικότητα, εφόσον βέβαια αξιοποιηθεί από την υλοποίηση. □

2. Για κάθε έναν από τους παρακάτω υπολογισμούς να αναφέρετε τον εξέχοντα λόγο σφάλματος κατά τη διάρκεια των παρακάτω υπολογισμών (Οι επιλογές σας είναι μεταξύ σφάλματος στρογγύλευσης και σφάλματος διακριτοποίησης): α) Το σφάλμα κατά τον υπολογισμό του αθροίσματος 100 αριθμών κινητής υποδιαστολής. β) Το σφάλμα που υπεισέρχεται όταν προσεγγίζουμε την παράγωγο $\frac{df}{dx}$ της συνάρτησης $f(x)$ στο ξ ως $f(\xi+h)-f(\xi))/h$ (μπορείτε να υποθέσετε ότι τα h και h/ξ είναι μικρά και ότι οι τιμές $h, \xi+h, f(\xi+h), f(\xi)$ δίδονται ακριβώς ως αριθμοί κινητής υποδιαστολής). γ) Το σφάλμα που υπεισέρχεται όταν αναπαριστούμε το π *με τον πλησιέστερο αριθμό κινητής υποδιαστολής.*

   Απάντηση. α) ΣΤΡΟΓΓΥΛΕΥΣΗΣ, προφανώς δεν διακριτοποιούμε καμια συνεχή συνάρτηση, τελεστή, κ.λπ. β) ΔΙΑΚΡΙΤΟΠΟΙΗΣΗΣ, μια και πρόκειται για σφάλμα που προέρχεται σχεδόν αποκλειστικά από την προσέγγιση της παραγώγου. γ) ΣΤΡΟΓΓΥΛΕΥΣΗΣ, μια αναφερόμαστε σε σφάλμα που προέρχεται αποκλειστικά από την απεικόνιση ενός πραγματικού αριθμού επί του συστήματος των α.κ.υ. □

3. Έστω το μητρώο $A = [4, -8, 1; 6, 5, 7; 0, -10, -3]$. Ποιό θα είναι το στοιχείο «οδηγός» στο πρώτο βήμα α) αν δεν χρησιμοποιηθεί οδήγηση; β) Αν χρησιμοποιηθεί μερική οδήγηση; γ) Αν χρησιμοποιηθεί πλήρης οδήγηση;

   Απάντηση. α) 4 (το στοιχείο στη θέση (1,1)), β) 6 (το μέγιστο σε απόλυτη τιμή στοιχείο της 1ης στήλης), γ) -10 (το μέγιστο σε απόλυτη τιμή στοιχείο όλου του μητρώου). □

4. Για κάθε ένα από τα παρακάτω μητρώα να εξηγήσετε αν έχουν καλό ή κακό δείκτη κατάστασης: α) $[10^{10}, 0; 0, 10^{-10}]$, β) $[10^{10}, 0; 0, 10^{10}]$. γ) $[10^{-10}, 0; 0, 10^{-10}]$. δ) $[1, 2; 2, 4]$.

   Απάντηση. Δείκτης κατάστασης $\boxed{\kappa(A) := \|A\|\|A^{-1}\|}$ *για όποια νόρμα διαλέξουμε. α)* ΚΑΚΟΣ: *γιατί* $A^{-1} = [10^{-10}, 0; 0, 10^{10}]$ και σε όλες τις νορμες $(1, 2\ \infty)$ έχουμε $\kappa(A) = 10^{10}/10^{-10} = 10^{20}$. β) ΑΡΙΣΤΟΣ. Καλό (τέλειο) γιατί $A = 10^{10}I$ όπου I ο ταυτοτικό μητρώο και επομένως $\kappa(A) = 1$. γ) ΑΡΙΣΤΟΣ: Ομοίως με πριν. δ) ΚΑΚΙΣΤΟΣ: μητρώο συμμετρικό και ο δείκτης κατάστασης είναι η μέγιστη προς ελάχιστη ιδιοτιμή. Προφανώς το μητρώο είναι μη αντιστρέψιμο, και $\lambda_{min} = 0$, επομένως $\kappa(A) = \infty$). □

5. Έστω ότι για ένα μητρώο A γνωρίζετε ότι οι παραγοντοποιήσεις LU και QR είναι και οι δυο εφικτές. Με βάση τα κριτήρια του επιστημονικού υπολογισμού, να αναφέρετε έναν λόγο για τον οποίον συνήθως προτιμάται η LU για επίλυση τετραγωνικού συστήματος $Ax = b$ και έναν λόγο που θα μπορούσε να καταστήσει τη χρήση της QR πιο επιθυμητή *(αναφερόμαστε πάντα σε τετραγωνικό σύστημα).*

   Απάντηση. Η LU είναι φθηνότερη (εκτελείται ταχύτερα, μικρότερο κόστος) ενώ η QR (π.χ. με Householder) είναι πίσω ευσταθής ανεξαρτήτως των δεδομένων. □ + [illegible]

6. Έστω ο βρόχος `for i=1:n, z(i) = a*x(i)+y(i); end`. Να τον ξαναγράψετε χρησιμοποιώντας ξετύλιγμα μήκους 3 και που να λειτουργεί σωστά ανεξάρτητα από την τιμή του n (θεωρούμε ότι είναι πάντα θετικός ακέραιος).

   Απάντηση. Το θέμα είναι να λειτουργεί σωστά για τιμές του n μικρότερες ή όχι κατ' ανάγκη πολλαπλάσια του 3. Ένας τρόπος είναι να υπολογίσουμε το υπόλοιπο της διαίρεσης του n με το 3, π.χ. με τη συνάρτηση `rem` της MATLAB και να γράψουμε:

```
m = rem(n,3); mp1 = m+1;
for i=1:m, z(i) = a*x(i)+y(i); end
for i = mp1:3:n /* προσέξτε. το n − m είναι πολλαπλάσιο του 3
    z(i) = a*x(i)+y(i);
    z(i+1) = a*x(i+1)+y(i+1);
    z(i+2) = a*x(i+2)+y(i+2);
end
```

□

7. Έστω ότι γνωρίζετε ότι σας δίδεται ένα πρόγραμμα (π.χ. η συνάρτηση MATLAB `myfun`), για το οποίο γνωρίζετε ότι για n δεδομένα εισόδου, έχει πολυπλοκότητα $O(n^2)$ αριθμητικών πράξεων κινητής υποδιαστολής *αλλά* δεν γνωρίζετε τι ακριβώς υπολογισμούς εκτελεί. Μπορείτε όμως να τρέξετε το πρόγραμμα και να χρησιμοποιήσετε χρονομετρητές (π.χ. `tic`, `toc`) για να μετρήσετε τον χρόνο που αναλώνει. Με βάση τα παραπάνω, ποιός είναι ο ελάχιστος αριθμός μετρήσεων που χρειάζεται για να εκτιμήσετε την πολυπλοκότητά του (δηλ. να εκτιμήσετε τους παράγοντες $\alpha_2, \alpha_1, \alpha_0$ στην έκφραση πολυπλοκότητας $\Omega = \alpha_2 n^2 + \alpha_1 n + \alpha_0$). Να περιγράψετε *συνοπτικά αλλά ξεκάθαρα* πώς θα ενεργούσατε για να εκτιμήσετε την πολυπλοκότητά του (π.χ. σε μορφή $\alpha_2 n^2 + \alpha_1 n + \alpha_0$ με γνωστούς παράγοντες $\alpha_2, \alpha_1, \alpha_0$).

Απάντηση. Για να εκτιμήσω τους 3 παράγοντες χρειάζομαι τουλάχιστον 3 μετρήσεις με το χρονομετρητή. Συνήθως όμως χρειάζονται πολύ περισσότερες για να έχω μια καλή εκτίμηση. Ο τρόπος είναι να προσπαθήσω να υπολογίσω τους παράγοντες χρησιμοποιώντας ελάχιστα τετράγωνα και λύνοντας ένα πρόβλημα του τύπου $Va = b$, όπου κάθε γραμμή του $V \in \mathbb{R}^{n\times 3}$ περιέχει στοιχεία $[1, \nu_i, \nu_i^2]$ όπου ν_i είναι κάποια τιμή για το n και η αντίστοιχη θέση του b έχει τη χρονομέτρηση για της συνάρτησης για την τιμή $n = \nu_i$. Το διάνυσμα $a = [\alpha_0, \alpha_1, \alpha_2]^\top$. □

II. *Δίνονται διανύσματα $x, y \in \mathbb{R}^4$. Υποθέτουμε ότι όλα τα στοιχεία τους είναι μη αρνητικοί αριθμοί κινητής υποδιαστολής.* Έστω ο υπολογισμός

```
s=0; for i=1:4, s=s+x(i)*y(i); end
```

1. *Να δείξετε ότι ο αλγόριθμος υπολογισμού είναι πίσω σταθερός.*

   Απάντηση. Σύμφωνα με τα στοιχεία που γνωρίζουμε για την διάδοση σφάλματος θα έχουμε:

$$\begin{aligned} \mathrm{fl}(s) &= (((x(1)y(1)(1+\delta_1) + x(2)y(2)(1+\delta_2))(1+\delta_3) \\ &\quad + x(3)y(3)(1+\delta_4))(1+\delta_5) + x(4)y(4)(1+\delta_6))(1+\delta_7) \\ &= x(1)y(1)(1+\delta_1)(1+\delta_3)(1+\delta_5)(1+\delta_7) \\ &\quad + x(2)y(2)(1+\delta_2)(1+\delta_3)(1+\delta_5)(1+\delta_7) \\ &\quad + x(3)y(3)(1+\delta_4)(1+\delta_5)(1+\delta_7) + x(4)y(4)(1+\delta_6)(1+\delta_7) \\ &= x(1)y(1)(1+\theta_4) + x(2)y(2)(1+\hat{\theta}_4) + x(3)y(3)(1+\theta_3) + x(4)y(4)(1+\theta_2) \end{aligned}$$

   όπου ως συνήθως $|\delta_j| \leq u$ και $|\theta_j| \leq \gamma_j := ju/(1-ju)$ όπου u είναι η μονάδα στρογγύλευσης. Επομένως (1η συνθήκη πίσω ευστάθειας) θα μπορούσαμε να έχουμε τα ίδια αποτελέσματα χρησιμοποιώντας για είσοδο τα διανύσματα x και

$$\tilde{y} := [y(1)(1+\theta_4), y(2)(1+\hat{\theta}_4), y(3)(1+\theta_3), y(4)(1+\theta_2)]$$

   και επειδή $|\theta_j| \leq \gamma_j := ju/(1-ju)$ θα έχουμε ότι $\tilde{y}$ είναι κοντά στο y (2η συνθήκη πίσω ευστάθειας). Άρα ο αλγόριθμος πληροί και τις δυο συνθήκες για πίσω ευστάθεια. ΠΡΟΣΟΧΗ: Μερικοί (λανθασμένα) δεν χρησιμοποίησαν απόλυτες τιμές, π.χ. έφραζαν από τα δεξιά χρησιμοποιώντας μόνον το θ, κ.λπ. □

2. Να δείξετε ότι το ΣΧΕΤΙΚΟ εμπρός σφάλμα θα είναι μικρό.

   Απάντηση. Συνεχίζοντας τα παραπάνω, έχουμε ότι το σχετικό εμπρός σφάλμα είναι φραγμένο ως εξής:

$$\begin{aligned} \frac{|\mathrm{fl}(s)-s|}{|s|} &= \frac{|x(1)y(1)(1+\theta_4) + x(2)y(2)(1+\hat{\theta}_4) + x(3)y(3)(1+\theta_3) + x(4)y(4)(1+\theta_2) - s|}{|s|} \\ &= \frac{|x(1)y(1)\theta_4 + x(2)y(2)\hat{\theta}_4 + x(3)y(3)\theta_3 + x(4)y(4)\theta_2|}{|s|} \\ &\leq \frac{|x(1)y(1)||\theta_4| + |x(2)y(2)||\hat{\theta}_4| + |x(3)y(3)||\theta_3| + |x(4)y(4)||\theta_2|}{s} \\ &\leq \gamma_4 \frac{|x(1)y(1)| + |x(2)y(2)| + |x(3)y(3)| + |x(4)y(4)|}{s} \end{aligned}$$

αλλά επειδή είναι όλα μη αρνητικά θα έχουμε ότι

$$|s| = s = |x(1)y(1)| + |x(2)y(2)| + |x(3)y(3)| + |x(4)y(4)|$$

$$\frac{|\mathrm{fl}(s) - s|}{|s|} \leq \gamma_4.$$

το οποίο είναι πολύ μικρό. □

3. *Να σχολιάσετε τον ισχυρισμό: «Αν δεν ισχύει η προϋπόθεση ότι όλα τα στοιχεία των x, y είναι μη αρνητικά, δεν μπορούμε να είμαστε βέβαιοι για κάποιο από τα προηγούμενα (δηλ. την πίσω ευστάθεια, το μικρό εμπρός σφάλμα, ή και τα δύο)».*

   Απάντηση. Τότε δεν μπορούμε να είμαστε βέβαιοι για μικρό εμπρός σχετικό σφάλμα καθώς δεν υπάρχει βεβαιότητα για το μέγεθος του s που βρίσκεται στον παρονομαστή, μπορεί δηλ. το $|s|$ να είναι πολύ μικρό σε σύγκριση με το $|\mathrm{fl}(s) - s|$. □

4. Αν ο αλγόριθμος υλοποιηθεί σε σύστημα που διαθέτει την εντολή FMA, να εξηγήσετε συνοπτικά (χωρίς πλήρη ανάλυση σφάλματος) γιατί η χρήση της FMA μπορεί να επιδράσει ευνοϊκά στην ακρίβεια του παραπάνω υπολογισμού.

   Απάντηση. Η εντολή FMA υλοποιεί την πράξη $s = c + a \times b$ με ένα μόνο σφάλμα στρογγύλευσης, δηλ.

   $$\mathrm{fl}(s) = (c + ab)(1 + \delta) \quad \text{αντί για} \quad \mathrm{fl}(s) = (c + ab(1 + \delta_1))(1 + \delta_2)$$

   επομένως μπορεί να επιφέρει μικρότερο σφάλμα στον παραπάνω βρόχου του οποιου κάθε επανάληψη είναι μια FMA. □

III. Έστω ότι έχουμε εφαρμόσει τον αλγόριθμο παραγοντοποίησης QR στο μητρώο $A \in \mathbb{R}^{3\times 3}$ που επιστρέφει στη θέση του A μητρώο με στοιχεία

$$\begin{pmatrix} 3 & -2 & 1 \\ 2 & 1 & 4 \\ 1 & -1 & 2 \end{pmatrix}$$

1. Να υπολογίσετε το αρχικό μητρώο A.

2. Να χρησιμοποιήσετε τα παραπάνω (απαραίτητο για την πλήρη βαθμολόγηση) για να λύσετε το σύστημα $Ax = b$ όπου $b = \frac{1}{3}[-5, -16, 4]^\top$.

Απάντηση. Ως γνωστό, το R της QR είναι το μητρώο

$$\begin{pmatrix} 3 & -2 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 2 \end{pmatrix}$$

ενώ τα διανύσματα Householder θα είναι

$$u_1 = [1, 2, 1]^\top, u_2 = [0, 1, -1].$$

Οι αντίστοιχοι μετασχηματισμοί οι

$$\begin{aligned} H_1 &= I - \frac{2u_1u_1^\top}{u_1^\top u_1} = \frac{1}{3}\begin{pmatrix} 2 & -2 & -1 \\ -2 & -1 & -2 \\ -1 & -2 & 2 \end{pmatrix} \\ H_2 &= I - \frac{2u_2u_2^\top}{u_2^\top u_2} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}. \end{aligned}$$

Προσέξτε: Είναι συμμετρικοί και ορθογώνιοι (ανακλαστές) και ισχύει ότι

$$\boxed{H_2 H_1 A = R \Rightarrow A = H_1 H_2 R} = \begin{pmatrix} 2 & -5/3 & -2 \\ -2 & 2/3 & -4 \\ -1 & 4/3 & 1 \end{pmatrix}.$$

Τέλος για να λύσουμε το σύστημα $Ax = b$ χρησιμοποιούμε ότι

$$\boxed{H_2 H_1 Ax = Rx = H_2 H_1 b} = [2, 5, 2]^\top$$

$Q = H_1 H_2 \dots H_k$

$A = QR$ οπότε $Ax = b \Rightarrow$

$QRx = b \Rightarrow Q^T Q R x = Q^T b \Rightarrow$

$Rx = Q^T b$

επομένως λύνουμε ως προς R με πίσω αντικατάσταση.

$$Rx = [2, 5, 2]^\top \Rightarrow x = [1, 1, 1]^\top.$$

Προσέξτε ότι ένας εναλλακτικός τρόπος θα ήταν να πολλαπλασιάσετε με $Q^\top$ και να λύσετε το σύστημα με συντελεστή R: $x = Q^\top b$. Το ενδιαφέρον όμως είναι ότι δεν χρειάζεται να υπολογίσετε το Q. Επίσης, στ' αλήθεια δεν χρειάζεται να υπολογίσετε τα H_1 και H_2, ενώ όλα τα παραπάνω μπορούν να προκύψουν μέσω των διανυσμάτω u_1, u_2 και χρησιμοποιώντας τη δομή των H_j για φθηνό πολλαπλασιασμό $H_1 H_2 A$.

□

IV. Έστω η διαφορική εξίσωση (αρχικών τιμών) $\frac{du}{dt}(t) = -Au(t)$ όπου $A = [2, -1; -1, 2]$, $u = [u_1(t), u_2(t)]^\top$ και οι συναρτήσεις u_1, u_2 είναι επιλεγμένες ώστε $u_1(0) = 2$, $u_2(0) = 1$.

1. Να χρησιμοποιήσετε την εμπρός μέθοδο Euler με σταθερό βήμα $\Delta t = 0.5$ για να υπολογίσετε αριθμητική προσέγγιση της λύσης στο σημείο $T = 2.0$.

   Απάντηση. Για ευκολία συμβολίζω το Δt με h. Στην εμπρός μέθοδο Euler εφαρμόζουμε τον τύπο

   $$\boxed{U(t+h) - U(t) = -hAU(t) \Rightarrow u(t+h) = (I - hA)U(t)}$$

   Επίσης, $I - hA = \frac{1}{2}[0, 1; 1, 0]$ επομένως εύκολα υπολογίζουμε:

   $$U(2) \quad = \quad (I - hA)((I - hA)((I - hA)((I - hA)U(0)))) = \frac{1}{16}[2, 1]^\top.$$

   ΠΡΟΣΟΧΗ: Ορισμένοι έγραψαν τον τύπο στη μορφή $U(t+h) = U(t)(I - hA)$. Η έκφραση αυτή δεν είναι έγκυρη! □

2. Να χρησιμοποιήσετε την ίδια μέθοδο αλλά με βήμα $\Delta t = 1.0$ για να υπολογίσετε αριθμητική προσέγγιση της λύσης στο σημείο $T = 2.0$.

   Απάντηση. Ομοίως, $I - hA = [-1, 1; 1, -1]$, άρα

   $$U(2) \quad = \quad (I - hA)((I - hA)U(0)) = [2, -2; -2, 2][2, 1]^\top = [2, -2]^\top.$$

   □

3. Είναι γνωστό ότι η ακριβής λύση του παραπάνω συστήματος των ΔΕ τείνει στο 0 καθώς το $t \to \infty$. Με βάση αυτό το στοιχείο, να σχολιάσετε τη συμπεριφορά των προσεγγίσεων που λάβατε χρησιμοποιώντας το IV.1 και το IV.2. (Υπόδειξη: Για μια πλήρη εξήγηση, είναι χρήσιμο να υπολογίσετε τις ιδιοτιμές του A.)

   Απάντηση. Προφανώς, υπάρχει πρόβλημα όταν $h = 1.0$. Ειδικότερα, οι τιμές της λύσης μεγαλώνουν (και μάλιστα αλλάζουν πρόσημο). Επομένως υπάρχει σοβαρή ένδειξη ότι έχουμε αστάθεια. Αυτό εξηγείται από την φασματική ακτίνα του $I - hA$. Για να υπάρχει ευστάθεια πρέπει η φασματική ακτίνα του $I - hA$ να μην είναι μεγαλύτερη του 1. Οι ιδιοτιμές του A είναι $\lambda = \{1, 3\}$, επομένως του $I - hA$ θα είναι $\{1 - h, 1 - 3h\}$. Για δεδομένο h, για ευστάθεια απαιτείται $\max\{|1-h|, |1-3h|\} < 1$. Επομένως, στην περίπτωση του βήματος 0.5 έχουμε ευστάθεια, ενώ όταν $h = 1.0$ άρα έχουμε αστάθεια. □

4. Να εφαρμόσετε ένα βήμα πίσω Euler για να βρείτε τη λύση στο σημείο $T = 2.0$ κατευθείαν με βήμα $\Delta t = 2.0$.

   Απάντηση. Στην πίσω μέθοδο Euler εφαρμόζουμε τον τύπο

   $$\begin{aligned} \boxed{U(t+h) - U(t) \quad = \quad -hAU(t+h) \Rightarrow (I + hA)U(t+h) = U(t)} \\ \begin{pmatrix} 5 & -2 \\ -2 & 5 \end{pmatrix} \begin{pmatrix} U_1(2) \\ U_2(2) \end{pmatrix} \quad = \quad \begin{pmatrix} 2 \\ 1 \end{pmatrix} \Rightarrow U(2) = \begin{pmatrix} 0.5714 \\ 0.4286 \end{pmatrix} \end{aligned}$$

   □

Ονοματεπώνυμο: .... Α.Μ. ....· Έτος ...

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι

Εξέταση Χειμερινού εξαμήνου κάτω από τον καλοκαιρινό ήλιο (4/7/07)

ΕΠΙΛΕΓΜΕΝΕΣ ΑΠΑΝΤΗΣΕΙΣ

1) Όλοι οι βαθμοί κατατίθενται στη Γραμματεία. 2) Παρακαλείστε να κλείσετε βιβλία, σημειώσεις και κινητά. 3) Επιπλέον της κόλλας πρέπει να επιστρέψετε τα θέματα καθώς και όλα τα πρόχειρα που θα δείχνουν την προσπάθειά σας, καθώς 4) για πλήρη βαθμό πρέπει να παρουσιάσετε όλο το συλλογισμό σας, όλα τα βήματα που κάνετε καθώς και τα ενδιάμεσα αποτελέσματα. *Να διαβάσετε προσεκτικά τις εκφωνήσεις. Έχετε 3 ώρες. Οι αλγόριθμοι πρέπει να περιγράφονται με σαφήνεια, π.χ. όπως στις σημειώσεις ή με MATLAB .*

ΚΑΛΗ ΕΠΙΤΥΧΙΑ!!!

1. **β.** $30 = 5 \times 6$

(α΄) Να εξηγήσετε ποιά απλοποίηση γίνεται στο υπολογιστικό μοντέλο που χρησιμοποιήσαμε στο μάθημα όταν λέμε ότι ο χρόνος έκτέλεσης Ω αριθμητικών πράξεων για δεδομένα που βρίσκονται ήδη στους καταχωρητές (και δεν μας ενδιαφέρει ο χρόνος των μεταφορών) είναι $T_{\alpha\rho\theta} = \tau_{\alpha\rho\theta}\Omega$, όπου $\tau_{\alpha\rho\theta}$ είναι ο χρόνος εκτέλεσης μιας αριθμητικής πράξης.

Απάντηση. Ο χρόνος εκτέλεσης όλων των α.κ.υ. θεωρείται ότι είναι ίδιος ενώ είναι γνωστό, π.χ. *ότι ο χρόνος για την εκτέλεση μιας διαίρεσης μπορεί να είναι πολύ μεγαλύτερος από το χρόνο που χρειάζεται η άθροιση.* □

(β΄) Ο παρακάτω κώδικας εκτελείται σε περιβάλλον MATLAB που τρέχει σε PC στο υπολογιστικό του τμήματος ή στο φορητό σας. Τι αποτέλεσμα προκύπτει στο **y** αν αρχικοποιήσετε **x=realmin**; Να δικαιολογήσετε πλήρως την απάντησή σας.

```
y=x/2; if (y==0), y=2*x, else y=y+1; end
```

Απάντηση. Τα παραπάνω συστήματα χρησιμοποιούν α.κ.υ. διπλής ακρίβειας IEEE και υποστηρίζει την υποκανονικοποίηση, επομένως το `y = x/2` όταν `x = realmin` θα επιστρέψει μη μηδενικό αριθμό. Επομένως, στη συνέχεια θα εκτελεστεί η επιλογή `y = y+1`. Επειδή όμως το `y = realmin/2`, που είναι πολύ μικρότερο του έψιλον της μηχανής, το *αποτέλεσμα θα είναι* `y = 1`. □

(γ΄) Να εξηγήσετε με ένα παράδειγμα γιατί ο παρακάτω ισχυρισμός μπορεί να είναι σωστός: Υπάρχουν μητρώα, έστω $A \in \mathbb{R}^{n\times n}$ ένα από αυτά, για τα οποία η επίλυση του γραμμικού συστήματος $Ax = b$ να κοστίζει $O(n)$ αριθμητικές πράξεις ενώ η επίλυση του $Bx = b$, όπου $B := A^n$ να στοιχίζει $O(n^3)$ πράξεις. Προσοχή: Θεωρούμε ότι το B είναι γνωστό (οπότε δεν συνυπολογίζουμε το κόστος του υπολογισμού του).

Απάντηση. Αν το A είναι τριδιαγώνιο, η λύση του $Ax = b$ κοστίζει $\Omega = O(n)$ ενώ το B *θα είναι πυκνό και η λύση $O(n^3)$. Θεωρούμε βέβαια ότι όταν χρησιμοποιούμε το B δεν* γνωρίζουμε ότι προήλθε από το A^n, γιατί τότε θα μπορούσαμε να λύσουμε διαδοχικά τα $x^{(0)} := b, Ax^{(j)} = x^{(j-1)}, j = 1, \ldots, n$ και $x := x^{(n)}$. Επειδή κάθε λύση γίνεται σε $O(n)$, χρειάζονται $O(n^2)$ πράξεις. Καλύτερα ακόμα αν χρησιμοποιούσαμε LU του A (στην ειδική αυτή περίπτωση μόνον για βελτίωση της σταθεράς). □

(δ΄) Γνωρίζουμε ότι η πίσω μέθοδος Euler για την επίλυση της ΔΕ $\frac{du}{dt} = Au$, όπου $A \in \mathbb{R}^{n\times n}$ και $u \in \mathbb{R}^n$, είναι ευσταθής για κάθε επιλογή βήματος Δt, σε αντίθεση με την εμπρός μέθοδο Euler. Να εξηγήσετε γιατί παρόλη την ευστάθεια, δεν ενδείκνυται να χρησιμοποιήσουμε πολύ μεγάλο βήμα.

Απάντηση. Στην πίσω Euler για το παραπάνω πρόβλημα δεν παρουσιάζεται αστάθεια μια και η φασματική ακτίνα του μητρώου $(I - hA)^{-1}$ θα είναι μικρότερη του 1 για οποιαδήποτε επιλογή του $\Delta t > 0$. Όμως, το σφάλμα διακριτοποίησης για το σχήμα είναι $O(\Delta t)$ (το ίδιο και στην εμπρός Euler), επομένως, ένα μεγάλο βήμα θα οδηγούσε σε ανάλογα μεγάλο σφάλμα διακριποίησης. Σημειώνουμε πως στην εμπρός Euler δεν τίθεται τέτοιο θέμα γιατί αν λάβουμε μεγάλο Δt δεν θα προλάβουμε καν να δούμε μεγάλο σφάλμα, η μέθοδος θα παράγει πολύ γρήγορα σκουπίδια λόγω αστάθειας. □

(ε΄) Έστω ότι γνωρίζετε ότι υπάρχει διάνυσμα $z \neq 0$ τέτοιο ώστε $Az = 0$. Να εξηγήσετε πολύ σύντομα γιατί τότε θα είναι πολύ δύσκολο να υπολογίσετε λύση του συστήματος $Ax = b$ για κάποιο άλλο b.

Απάντηση. Αφού υπάρχει τέτοιο μη μηδενικό διάνυσμα, οι στήλες του A θα είναι γραμμικά εξαρτημένες, επομένως η τάξη του μητρώου θα είναι μικρότερη του n και επομένως το μητρώο δεν θα είναι αντιστρέψιμο. □

2. **β.** $30 = 6 \times 5$

(α΄) Να εξηγήσετε με συντομία τι είναι η πίσω ανάλυση σφάλματος και με ποιόν τρόπο μπορεί να βοηθήσει στην εκτίμηση του εμπρός σφάλματος υπολογισμών σε συστήματα αριθμητικής κινητής υποδιαστολής.

Απάντηση. Βιβλίο (εν. 3.5) □ σελ. 78

(β΄) Να δείξετε ότι ο αλγόριθμος Horner για τον υπολογισμό της τιμής ενός πολυωνύμου είναι πίσω ευσταθής.

<u>Υπενθύμιση:</u> *Ο αλγόριθμος* Horner *υπολογίζει την τιμή ενός πολυωνύμου* $p(x) := \sum_{j=0}^{n} \alpha_j x^j$ *αναδρομικά. Για παράδειγμα, στην περίπτωση του* $n = 3$, *αντιστοιχεί στον υπολογισμό*

$$((\alpha_3 x + \alpha_2)x + \alpha_1)x + \alpha_0$$

Απάντηση. Βιβλίο (εν. 3.5) □

(γ΄) Να δείξετε με σαφήνεια πώς θα βελτιωνόταν η πίσω ευστάθεια αν η πλατφόρμα σας είχε εντολή FMA.

Απάντηση. Ακολουθεί άμεσα από τη συζήτηση στο βιβλίο σχετικά με τις ιδιότητες της FMA (ένα σφάλμα ανά πράξη τύπου $\gamma \leftarrow \gamma + \alpha\beta$. □

(δ΄) Έστω ότι γνωρίζετε εκ των προτέρων ότι οι συντελεστές του πολυωνύμου είναι όλοι μη αρνητικοί. Να εξηγήσετε (αν υπήρχε η επιλογή) αν θα ήταν καλύτερα να υπολογίσετε την τιμή του για αρνητική ή για θετική τιμή του x.

Απάντηση. Για θετικό διότι μειώνεται η πιθανότητα καταστροφικής απαλοιφής, δηλ. το αποτέλεσμα να είναι πολύ μικρό ενώ οι προστιθέμενοι παράγοντες μεγάλοι σε απόλυτη τιμή, κάτι που οδηγεί σε πολύ μεγάλο σχετικό δείκτη κατάστασης. □

(ε΄) Δίδονται $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, όπου γνωρίζουμε ότι $n > 1, m \geq 1$ αλλά δεν ισχύει κατ΄ ανάγκη ότι $m = n$. Επιπλέον, όλα τα στοιχεία των x, y είναι αριθμοί κινητής υποδιαστολής. Να αιτιολογήσετε πλήρως τον ισχυρισμό ότι οποιοσδήποτε αλγόριθμος υπολογισμού του $xy^\top$ δεν μπορεί γενικά να αποδειχτεί ότι είναι πίσω ευσταθής.

Απάντηση. Η απάντηση στο βιβλίο, εν. 4.2.2 (Ένας αλγόριθμος υπολογισμού εξωτερικού γινομένου δεν μπορεί να είναι πίσω ευσταθής γιατί δεν υπάρχουν αρκετά στοιχεία στην είσοδο για να ρίξουμε τα σφάλματα κατά τις πράξεις) □

(ϛ΄) Υπάρχει περίπτωση να μπορέσουμε να αποδείξουμε πίσω ευστάθεια για γενικά δεδομένα αλλά ειδικές τιμές των m, n (εντός των παραπάνω περιορισμών).

Απάντηση. Αν οι διαστάσεις το επιτρέπουν, π.χ. αν $m = 1$ οπότε είναι προφανές, μια και πρόκειται για πολλαπλασιασμό διανύσματος με βαθμωτό και γενικότερα αν $mn \leq m + n$, π.χ. $m = 2, n = 2$. □

3. β. $24 = 3 \times 8$ Έστω ότι έχει χρησιμοποιηθεί παραγοντοποίηση QR ενός μητρώου $A \in \mathbb{R}^{4\times 3}$ και ότι στο τέλος, τα στοιχεία του A έχουν αντικατασταθεί με τα παρακάτω:

$$\begin{pmatrix} 8 & 1 & 6 \\ 1 & 5 & 7 \\ 1 & 2 & 2 \\ 1 & 2 & 2 \end{pmatrix}$$

α) Έστω ότι $b = [5, -15, 0, 30]^\top$. Με βάση μόνον αυτά τα στοιχεία, να υπολογίσετε τη λύση του $\mathrm{argmin}_{x\in\mathbb{R}^3} \|b - Ax\|_2$ για το αρχικό (και άγνωστο προς το παρόν) A. β) Να υπολογίσετε *το αρχικό μητρώο A χωρίς να υπολογίσετε άμεσα το Q. γ) Να υπολογίσετε το Q. Προσοχή:* Για πλήρη βαθμό θα πρέπει να λύσετε το (α) όπως ζητάται από την εκφώνηση, δηλ. χωρίς να χρησιμοποιήσετε το αποτέλεσμα του (β). Ομοίως, το (β) πρέπει να λυθεί πριν το (γ).

Απάντηση. Προσέξτε ότι η διάσταση ήταν $A \in \mathbb{R}^{4\times 3}$, επομένως χρειάστηκαν 3 ανακλαστές για να άνω τριγωνιοποιήσουν το μητρώο. Στις θέσεις του A επιστρέφονται, στο μεν άνω τριγωνικό μέρος το άνω τριγωνικό τμήμα του (άνω τριγωνικού) $R \in \mathbb{R}^{4\times 3}$, ενώ στο κάτω τριγωνικό υπάρχει η πληροφορία για την ανασύνθεση των διανυσμάτων Householder. Ειδικότερα:

$$u_1 = [1, 1, 1, 1]^\top, u_2 = [0, 1, 2, 2]^\top, u_3 = [0, 0, 1, 2]$$

ενώ οι ανακλαστές θα είναι της μορφής $H_j = I - 2\frac{u_j u_j^\top}{u_j^\top u_j}$ που είναι ορθογώνιοι και συμμετρικοί (στη συνέχεια, αυτές οι ιδιότητες χρησιμοποιούνται χωρίς να αναφερόμαστε ειδικά). Τότε επειδή η ευκλείδεια νόρμα διατηρείται μετά από ορθογώνιους μετασχηματισμούς, όπως οι ανακλαστές H_j, ισχύει ότι

$$\|b - Ax\| = \|H_3 H_2 H_1 b - Rx\|.$$

Προσέξτε τώρα ότι για τους υπολογισμούς των $H_3 H_2 H_1 b$ μπορούμε να αξιοποιήσουμε την ειδική μορφή των H_j, καθώς για οποιοδήποτε διάνυσμα, έστω g, $H_j g = g - \frac{2}{u_j^\top u_j} u_j (u_j^\top g)$. Ακολουθώντας αυτή τη μέθοδο (έτσι δεν κατασκευάζουμε το A και το Q, αλλά ούτε τα H_j *και όλοι οι υπολογισμοί γίνονται με πρόσβαση στα u_j), τα βήματα εδώ είναι τα ακόλουθα:*

$$\hat{b} := H_3(H_2(H_1 b)) = [-5, -215/9, -202/9, -64/9]^\top$$

Χρησιμοποιώντας πίσω αντικατάσταση, λύνουμε το $\boxed{Rx = \hat{b}.}$ Προσέξτε ότι η τελευταία γραμμή του R είναι όλο 0, και το τελικό σφάλμα θα είναι ο τελευταίος όρος του $\hat{b}$, δηλ. $\min_{x\in\mathbb{R}^3} \|\hat{b} - Rx\| = 64/9$. Όσο για το βέλτιστο x θα είναι η ακριβής λύση του άνω τριγωνικού (τετραγωνικού) $R_{1:3,1:3} x = \hat{b}_{1:3}$, που είναι $x = [257/40, 164/15, -101/9]^\top$.

Για τον υπολογισμό του A χρησιμοποιούμε πάλι τις ιδιότητες των H_j ανά στήλη του R, δηλ. το ότι

$$A_{:,j} = H_1(H_2(H_3 R_{:,j})) = H_1(H_2(R_{:,j} - \frac{2}{u_1^\top u_1} u_1^\top R_{:,j})), \quad j = 1, 2, 3$$

κ.ο.κ. οπότε

$$A = \begin{pmatrix} 4 & 7/9 & 283/90 \\ -4 & 11/4 & 83/30 \\ -4 & -22/9 & -397/90 \\ -4 & -22/9 & -649/90 \end{pmatrix}.$$

Τέλος υπολογίζουμε ανά στήλη και πάλι

$$Q_{:,j} = H_1(H_2(H_3(:,j)))$$

χρησιμοποιώντας το ότι

$$H_3(:,j) = e_j - \frac{2}{u_3^\top u_3} u_3 u_{3,j} = e_j - \frac{2}{5} u_3 u_{3,j}$$

και έχουμε

$$Q = \begin{pmatrix} 1/2 & 1/18 & -11/90 & -77/90 \\ -1/2 & 5/6 & -1/30 & -7/30 \\ -1/2 & -7/18 & 59/90 & -37/90 \\ -1/2 & -7/18 & -67/90 & -19/90 \end{pmatrix}.$$

ΠΡΟΣΟΧΗ: Στα παραπάνω η σειρά με την οποία γίνονται οι πράξεις έχει σημασία και βοηθά στην ταχύτερη επίλυση. Προσέξτε επίσης ότι η εκφώνηση δεν ζητά πουθενά τον άμεσο υπολογισμό των H_j. □

4. β. $24 = 3 \times 8$ Έστω ότι έχει χρησιμοποιηθεί παραγοντοποίηση QR ενός μητρώου $A \in \mathbb{R}^{4\times 3}$ και ότι στο τέλος, τα στοιχεία του A έχουν αντικατασταθεί με τα παρακάτω:

$$\begin{pmatrix} 5 & -15 & 5 \\ -2 & 15 & 0 \\ 2 & 2 & 15 \\ -1 & -1 & -3 \end{pmatrix}$$

α) Έστω ότι $b = [5, -15, 0, 30]^\top$. Με βάση μόνον αυτά τα στοιχεία, να υπολογίσετε τη λύση του $\operatorname{argmin}_{x\in\mathbb{R}^3} \|b - Ax\|_2$ για το αρχικό (και άγνωστο προς το παρόν) A. β) Να υπολογίσετε το αρχικό μητρώο A χωρίς να υπολογίσετε άμεσα το Q. γ) Να υπολογίσετε το Q. Προσοχή: Για πλήρη βαθμό θα πρέπει να λύσετε το (α) όπως ζητάται από την εκφώνηση, δηλ. χωρίς να χρησιμοποιήσετε το αποτέλεσμα του (β). Ομοίως, το (β) πρέπει να λυθεί πριν το (γ).

Απάντηση. Προσέξτε ότι η διάσταση ήταν $A \in \mathbb{R}^{4\times 3}$, επομένως χρειάστηκαν 3 ανακλαστές για να άνω τριγωνιοποιήσουν το μητρώο. Στις θέσεις του A επιστρέφονται, στο μεν άνω τριγωνικό μέρος το άνω τριγωνικό τμήμα του (άνω τριγωνικού) $R \in \mathbb{R}^{4\times 3}$, ενώ στο κάτω τριγωνικό υπάρχει η πληροφορία για την ανασύνθεση των διανυσμάτων Householder. Ειδικότερα:

$$u_1 = [1, -2, 2, -1]^\top, u_2 = [0, 1, 2, -2]^\top, u_3 = [0, 0, 1, -3]$$

ενώ οι ανακλαστές θα είναι της μορφής $H_j = I - 2\frac{u_j u_j^\top}{u_j^\top u_j}$ που είναι ορθογώνιοι και συμμετρικοί (στη συνέχεια, αυτές οι ιδιότητες χρησιμοποιούνται χωρίς να αναφερόμαστε ειδικά). Τότε επειδή η ευκλείδεια νόρμα διατηρείται μετά από ορθογώνιους μετασχηματισμούς, όπως οι ανακλαστές H_j, ισχύει ότι

$$\|b - Ax\| = \|H_3 H_2 H_1 b - Rx\|.$$

Προσέξτε τώρα ότι για τους υπολογισμούς των $H_3H_2H_1b$ μπορούμε να αξιοποιήσουμε την ειδική μορφή των H_j, καθώς για οποιοδήποτε διάνυσμα, έστω g, $H_j g = g - \frac{2}{u_j^\top u_j} u_j(u_j^\top g)$. Ακολουθώντας αυτή τη μέθοδο (έτσι δεν κατασκευάζουμε το A και το Q, αλλά ούτε τα H_j και όλοι οι υπολογισμοί γίνονται με πρόσβαση στα u_j), τα βήματα εδώ είναι τα ακόλουθα:

$$\hat{b} := H_3(H_2(H_1 b)) = [4, 3, 33, 6]^\top$$

Χρησιμοποιώντας πίσω αντικατάσταση, λύνουμε το $Rx = \hat{b}$. Προσέξτε ότι η τελευταία γραμμή του R είναι όλο 0, και το τελικό σφάλμα θα είναι ο τελευταίος όρος του $\hat{b}$, δηλ. $\min_{x\in\mathbb{R}^3} \|\hat{b} - Rx\| = 6$. Όσο για το βέλτιστο x θα είναι η ακριβής λύση του άνω τριγωνικού (τετραγωνικού) $R_{1:3,1:3}x = \hat{b}_{1:3}$, που είναι $x = [-4/5, 1/5, 11/5]^\top$.

Για τον υπολογισμό του A χρησιμοποιούμε πάλι τις ιδιότητες των H_j ανά στήλη του R, δηλ. το ότι

$$A_{:,j} = H_1(H_2(H_3 R_{:,j})) = H_1(H_2(R_{:,j} - \frac{2}{u_1^\top u_1} u_1^\top R_{:,j})), \quad j = 1, 2, 3$$

κ.ο.κ. οπότε

$$A = \begin{pmatrix} 4 & -3 & 4 \\ 2 & -14 & -3 \\ -2 & 14 & 0 \\ 1 & -7 & 15 \end{pmatrix}.$$

Τέλος υπολογίζουμε ανά στήλη και πάλι

$$Q_{:,j} = H_1(H_2(H_3(:, j)))$$

και έχουμε

$$Q = \begin{pmatrix} 4/5 & 3/5 & 0 & 0 \\ 2/5 & -8/15 & -1/3 & -2/3 \\ -2/5 & 8/15 & 2/15 & -11/15 \\ 1/5 & -4/15 & 14/15 & -2/15 \end{pmatrix}.$$

ΠΡΟΣΟΧΗ: Στα παραπάνω η σειρά με την οποία γίνονται οι πράξεις έχει σημασία και βοηθά στην ταχύτερη επίλυση. Προσέξτε επίσης ότι η εκφώνηση δεν ζητά πουθενά τον άμεσο υπολογισμό των H_j. □

5. β. 16

Ενδιαφέρει η επίλυση της ΣΔΕ (συνοριακό πρόβλημα 2 σημείων) με την εξής μορφή

$$-\frac{d^2}{dx^2}u(x) + 2\frac{d}{dx}u(x) + x^2 u(x) \;=\; x$$

στο διάστημα $x \in [1, 2]$. Γνωρίζουμε ότι $u(1) = 0$, $u(2) = 1$. Να διακριτοποιήσετε με πεπερασμένες κεντρισμένες διαφορές χρησιμοποιώντας $n = 4$ εσωτερικά και ισαπέχοντα σημεία στο παραπανω διάστημα και να γράψετε το σύστημα των εξισώσεων που προκύπτει, δηλ.. $AU = F$. Η απάντησή σας πρέπει να περιέχει την τελική μορφή για τα στοιχεία των A, F, (αριθμούς και όχι μεταβλητές!).

Απάντηση. Το πλέγμα θα περιέχει τα εσωτερικά σημεία $\{x_j = 1 + jh | j = 1, ..., 4\}$ όπου $h = (2-1)/(4+1) = 1/5$ Αντικαθιστούμε τις προσεγγίσεις

$$-\frac{d^2}{dx^2}u(x) \approx \frac{-U_{i-1} + 2U_i - U_{i+1}}{h^2}, \quad \frac{d}{dx}u(x) \approx \frac{U_{i+1} - U_{i-1}}{2h},$$

οπότε λαμβάνουμε τις παρακάτω εξισώσεις

$$-(\frac{1}{h} + \frac{1}{h^2})U_{i-1} + (\frac{2}{h^2} + (1+ih)^2)U_i + (\frac{1}{h} - \frac{1}{h^2})U_{i+1} = 1 + ih$$

για $i = 1, ..., 4$. Προσέξτε ότι είναι προτιμότερο (για δική σας ευκολία στις πράξεις, όχι χάριν ορθότητας) να αφήσετε το $1/h^2$ στον παρονομαστή, καθώς $1/h^2 = 25$.

Από τις συνοριακές συνθήκες $U_0 = u_0 = 0, U_5 = u(1) = 1$, θα έχουμε για τις ακραίες εξισώσεις ότι

$$\frac{1286}{25}U_1 - 20U_2 = x_1 + 30U_0 = \frac{6}{5}$$

και

$$-30U_3 + \frac{1331}{25}U_4 = x_4 + 20U_5 = \frac{109}{5}.$$

Εντέλει τα ζητούμενα στοιχεία είναι τα ακόλουθα:

$$A = \begin{pmatrix} \frac{1286}{25} & -20 & 0 & 0 \\ -30 & \frac{1299}{25} & -20 & 0 \\ 0 & -30 & \frac{1314}{25} & -20 \\ 0 & 0 & -30 & \frac{1331}{25} \end{pmatrix}, \quad F = \begin{pmatrix} \frac{6}{5} \\ \frac{7}{5} \\ \frac{8}{5} \\ \frac{109}{5} \end{pmatrix}$$

□

6. β. 16

Ενδιαφέρει η επίλυση της ΣΔΕ (συνοριακό πρόβλημα 2 σημείων) με την εξής μορφή

$$-x^2\frac{d^2}{dx^2}u(x) + 2x\frac{d}{dx}u(x) + u(x) \quad = \quad x$$

στο διάστημα $x \in [1, 2]$. Γνωρίζουμε ότι $u(1) = 0$, $u(2) = 1$. Να διακριτοποιήσετε με πεπερασμένες κεντρισμένες διαφορές χρησιμοποιώντας $n = 4$ εσωτερικά και ισαπέχοντα σημεία στο παραπανω διάστημα και να γράψετε το σύστημα των εξισώσεων που προκύπτει, δηλ.. $AU = F$. Η απάντησή σας πρέπει να περιέχει την τελική μορφή για τα στοιχεία των A, F, (αριθμούς και όχι μεταβλητές!).

Απάντηση. Το πλέγμα θα περιέχει τα εσωτερικά σημεία $\{x_j = 1 + jh | j = 1, ..., 4\}$ όπου $h = (2-1)/(4+1) = 1/5$ Αντικαθιστούμε τις προσεγγίσεις

$$-\frac{d^2}{dx^2}u(x) \approx \frac{-U_{i-1} + 2U_i - U_{i+1}}{h^2}, \quad \frac{d}{dx}u(x) \approx \frac{U_{i+1} - U_{i-1}}{2h},$$

οπότε προκύπτουν οι παρακάτω εξισώσεις

$$-(\frac{x_i}{h} + \frac{x_i^2}{h^2})U_{i-1} + (\frac{2x_i^2}{h^2} + 1)U_i + (\frac{x_i}{h} - \frac{x_i^2}{h^2})U_{i+1} = x_i$$

για $i = 1, \ldots, 4$. Προσέξτε ότι είναι προτιμότερο (για δική σας ευκολία στις πράξεις, όχι χάριν ορθότητας) να αφήσετε το $1/h^2$ στον παρονομαστή, καθώς $1/h^2 = 25$.

Από τις συνοριακές συνθήκες $U_0 = u_0 = 0, U_5 = u(1) = 1$, για τις ακραίες εξισώσεις θα έχουμε ότι

$$73U_1 - 30U_2 = 6/5 + 42U_0 = 6/5$$

και

$$-90U_3 + 163U_4 = 9/5 + 72U_5 = 9/5 + 72 = 369/5.$$

Εντέλει τα ζητούμενα στοιχεία είναι τα ακόλουθα:

$$A = \begin{pmatrix} 73 & -30 & 0 & 0 \\ -56 & 99 & -42 & 0 \\ 0 & -72 & 129 & -56 \\ 0 & 0 & -90 & 163 \end{pmatrix}, \quad F = \begin{pmatrix} \frac{6}{5} \\ \frac{7}{5} \\ \frac{8}{5} \\ \frac{369}{5} \end{pmatrix}$$

□

ΕΠΙΣΤΗΜΟΝΙΚΟΣ ΥΠΟΛΟΓΙΣΜΟΣ Ι (24 Φεβρ. 2008, 12-3μμ) ΕΠΙΛΕΓΜΕΝΕΣ ΑΠΑΝΤΗΣΕΙΣ

1) α) Σ - Λ : Οι εντολές BLAS-2 μπορούν να υλοποιηθούν να έχουν καλύτερη επίδοση από τις BLAS-3.

Απάντηση. Λάθος: Οι εντολές BLAS-3 έχουν μικρότερο ελάχιστο αριθμό μεταφορών ανά πράξη α.κ.υ. από τις πράξεις BLAS «μικρότερων κατηγοριών». Επομένως, υπό την προϋπόθεση ότι αναφερόμαστε σε υλοποιήσεις που έχουν γίνει με στόχο την επίτευξη του μικρότερου λόγου μεταφορών προς πράξεις για κάθε κατηγορία, οι πράξεις BLAS-3 θα έχουν καλύτερη επίδοση (μετρούμενου με βάση τα Mflops). □

β) Σ - Λ : Ξεδίπλωμα βρόχου γενικά χρησιμοποιείται για να μειώσει το πλήθος πράξεων α.κ.υ.

Απάντηση. ΛΑΘΟΣ το ξεδίπλωμα δεν επιφέρει αλλαγή του Ω, μόνον ο βρόχος εκτελείται λιγότερες φορές αλλά με περισσότερες εντολές σε κάθε επανάληψη. □

γ) Έστω στη MATLAB οι εκφράσεις $M + 20 - 10 - M$, $M + 20 - M - 10$, $M - 10 - M + 20$. Να εξηγήσετε τις τιμές που υπολογίζονται αν το M αρχικοποιηθεί ως `realmax`.

σελ 13
Set 08

Απάντηση. Το `realmax` της α.κ.υ. διπλής ακρίβειας είναι της μορφής $1.* \times 2^{1023}$ επομένως η προσθαφαίρεση αριθμών σαν το 10 και 20 με αυτό δεν επιφέρει καμια αλλαγή λόγω της απαιτούμενης κανονικοποίησης και επακόλουθου μηδενισμού τους κατά την πρώτη φάση της διαδικασίας. Επομένως τα αποτελέσματα θα είναι $((M + 20) - 10) - M \doteq (M - 10) - M = M - M = 0$, $((M + 20) - M) - 10 = (M - M) - 10 = -10$, $((M - 10) - M) + 20 = (M - M) + 20 = 20$. □

δ) Έστω αντιστρέψιμο $A \in \mathbb{R}^{n\times n}$ με μικρό δείκτη κατάστασης, $b \in \mathbb{R}^n$ και ο υπολογισμός `[L, U] = lu(A); x = U\(L\b)` (η MATLAB χρησιμοποιεί LAPACK). Ισχύει ή όχι ότι το εμπρός σφάλμα στο υπολογισμένο x δεν θα είναι μεγάλο;

Απάντηση. Για την LU γενικού μητρώου δεν μπορεί να αποδειχτεί μικρή πίσω ευστάθεια, που είναι απαραίτητη για να εγγυηθούμε μικρό εμπρός σφάλμα λόγω μικρού δείκτη κατάστασης, επομένως ΔΕΝ ΙΣΧΥΕΙ. Βασιζόμαστε και στον γνωστό τύπο (εμπρός σφ.) < (πίσω σφ.) × (δείκτης κατ. A). □

2. Μας δίδονται α.κ.υ. και ένας αλγόριθμος για να τους αθροίσουμε. Να εξηγήσετε ποιοί από τους παρακάτω ισχυρισμούς είναι σωστοί και ποιοί λάθος:

α) Αν αλλάξουμε τον αλγόριθμο άθροισης, μπορεί να αλλάξουν το πίσω σφάλμα και το εμπρός σφάλμα.

β) Αν γνωρίζουμε τους α.κ.υ. και τον αλγόριθμο άθροισης, μπορούμε να υπολογίσουμε το *ακριβές* εμπρός σφάλμα.

γ) Αν οι αριθμοί είναι ομόσημοι, ενας καλός τρόπος άθροισης είναι από το μικρότερο προς το μεγαλύτερο.

δ) Αν η απόλυτη τιμή του υπολογισμένου αθροίσματος είναι πολύ μικρότερη του μέσου όρου των απολύτων τιμών των στοιχείων που αθροίστηκαν, μπορούμε να υποθέσουμε με ασφάλεια ότι το σχετικό εμπρός σφάλμα στο άθροισμα θα είναι και αυτό μικρό. $|sum| \ll \frac{|a_1|+|a_2|+\dots+|a_m|}{m}$

Λάθος. Θα πρέπει ο όρος $\frac{|a_1|+|a_2|+\dots+|a_m|}{m}$ να είναι φραγμένος.

Απάντηση. α) ΣΩΣΤΟ, και τα δυο εξαρτώνται από τον αλγόριθμο και επομένως τη σειρά άθροισης (εξάλλου το πίσω σφάλμα μετρά τον «δείκτη κατάστασης του αλγορίθμου».) β) ΛΑΘΟΣ, το ακριβές σφάλμα δεν μπορεί να υπολογιστεί γενικά γιατί χρειαζόμαστε αριθμητική άπειρης ακρίβειας. γ) ΣΩΣΤΟ, γιατί τότε μειώνεται η πιθανότητα σφάλματος από την πρόσθεση αριθμών που διαφέρουν πάρα πολύ σε μέγεθος που θα είχε για συνέπεια μηδενισμό των μικρότερων λόγω κανονικοποίησης των εκθετών. Επίσης τα «δ» που συσσωρεύονται στη διάδοση του σφάλματος επιβαρύνουν περισσότερο τους μικρότερούς όρους του αθροίσματος. δ) ΛΑΘΟΣ : Τυπικό παράδειγμα $(1+\delta_1)-(1-\delta_2) = \delta_1+\delta_2$ όπου τα δ_i είναι πολύ μικρά και περιέχουν κυρίως «θόρυβο» από προηγούμενες πράξεις. Κλασικό παράδειγμα που δημιουργείται πρόβλημα από καταστροφική απαλοιφή. □

μόνο στο β υπολογίζεται το εμπρός σφάλμα.

3. Δίδονται τα στοιχεία $A \in \mathbb{R}^{10\times m}$ και $b \in \mathbb{R}^m$, $c \in \mathbb{R}^{10}$ και θέλουμε να υπολογίσουμε το $y \leftarrow c + Ab$. Το m δεν έχει κανένα περιορισμό. α) Ποιό είναι το $\Phi_{\min}$ για την πράξη; β) Να δείξετε πώς μπορείτε να υλοποιήσετε τον πολλαπλασιασμό με $\Phi = \Phi_{\min}$ χρησιμοποιώντας κρυφή μνήμη και καταχωρητές $O(1)$ (δηλ. προσωρινή μνήμη άμεσης προσβασης μεγέθους ανεξάρτητου του m).

Απάντηση. α) Με απλή καταμέτρηση των α.κ.υ. εισόδου/εξόδου που χρησιμοποιούνται στον υπολογισμό, έχουμε $10m$ για φόρτωση του A, $m + 10$ για φόρτωση των c, b, και 10 για την αποθήκευση

$y_{10\times1} \leftarrow c_{10\times1} + A_{10\times m} \cdot b_{m\times1}$

35

$\Omega = 10\cdot(2m-1)+10$
$= 20m - 10 + 10$
$= 20m$

$\Phi = 10 + 10m + m + 10 = 20 + 11m$

MATLAB

Αριστερή διαίρεση

$A/B = A * B^{-1}$

Δεξιά διαίρεση

$B \backslash A = B^{-1} * A$

$[L,U] = lu(A)$

$x = U \backslash \underbrace{(L \backslash b)}_{y}$

αυτή η εντολή
συνενώνει την
εμπρός και την
πίσω αντικατάσταση
από το βήμα 2
και 3 της LU.
Πιο συγκεκριμένα:

Εμπρός
Αντικατάσταση: $L y = b \Rightarrow y = L^{-1} b$ $(L \backslash b)$

Πίσω
Αντικατάσταση $U x = y \Rightarrow x = U^{-1} y$ $(U \backslash y)$

στο y, συνολικά δηλ. $\Phi_{\min} = 11m + 20$. β) Η σχετική ύλη υπάρχει και στις διαφάνειες. Συνοψίζουμε λέγοντας ότι η υλοποίηση μπορεί να κωδικοποιηθεί ως εξής, εφόσον διατίθεται χώρος για την αποθήκευση σε καταχωρητές και cache της τάξης του $O(1)$. Η μεταβλητή temp έχει αναφέρεται σε καταχωρητές μήκους 10.

```
LOAD c
for j = 1 : m
  LOAD b(j)
  for i = 1 : 10
    LOAD A(i, j)
    temp(i) = c(i) + A(i, j) * b(j)
  end
end
STORE y = temp
```
□

LOAD C
for J=1:m
LOAD b(J)
for i=1:10
LOAD A(i,J)
$y(i) \leftarrow c(i) + A(i,J)\, b(J)$
STORE y(i)
end
end

4. α) Τι θα εμφανιστεί στην οθόνη αν εκτελέσετε τις παρακάτω εντολές σε περιβάλλον MATLAB και n=3:
`for j=1:n, A = kron(ones(j,1),[1:j]), end`

Απάντηση.

$A = 1$

$A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$

$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}$ □

Υπενθυμίζουμε ότι η εντολή kron(A,B) επιστρέφει το γινόμενο Kronecker $A \otimes B$.

β) Να ενθέσετε (αιτιολογώντας, πάντα) σε επιπλέον κώδικα που να υπολογίζει όσο μπορείτε πιο αξιόπιστα (επιστρέφοντας σε κάποια μεταβλητή) τα Mflop/s των παραπάνω εντολών στο υπολογιστικό σας περιβάλλον. Μπορείτε να υποθέσετε ότι άν $A \in \mathbb{R}^{m_A \times n_A}$, $B \in \mathbb{R}^{m_B \times n_B}$ τότε το κόστος του kron(A,B) είναι $\Omega = m_A n_A m_B n_B$.

Απάντηση. Για συντομία συμβολίζουμε με Δ τις εντολές `for j=1:n, A = kron(ones(j,1),[1:j]), end`. Προσέξτε ότι το Ω θα είναι $\sum_{j=1}^{n} j^2$. Μπορεί να υπολογιστεί από κλασικούς τύπους αθροισμάτων προόδων ή στο πρόγραμμα, συσσωρεύοντας τις πράξεις κάθε επανάληψης σε μεταβλητή. Τότε

```
% εκτέλεση για να αποφευχθεί «θόρυβος» από την αρχικοποίηση
tic; for j=1:itmax, Δ; end;
optime = toc/itmax; ops = 0;
for j=1:itmax, ops = ops+j*j;  end; mflops = ops*1e-6/toc;
```
□

5. α) Είναι το μοντέλο διάδοσης του σφάλματος στον πολλαπλασιασμό κινητής υποδιαστολής. $x\tilde{\times}y = x \times y(1+\delta)$ όπου $|\delta| \le \mathbf{u}$, $\mathbf{u}$ η μονάδα στρογγύλευσης και x, y αριθμοί κινητής υποδιαστολής, άμεσο επακόλουθο της «αρχής ακριβούς στρογγύλευσης»; Αν ναι, να το δείξετε, αν όχι να εξηγήσετε γιατί.

Απάντηση. ΕΙΝΑΙ: Η αρχή προσδιορίζει ότι με τις παραπάνω συνθήκες, για τον πολλαπλασιασμό ισχύει ότι η πράξη που εκτελείται στη μηχανή έχει ως αποτέλεσμα την ποσότητα που θα υπολογιζόταν με αριθμητική άπειρης ακρίβειας (δηλ. το $x \times y$) με στρογγύλευση (υποθέτουμε προς το πλησιέστερο) μετά, επομένως το τελικό αποτέλεσμα θα είναι $x \times y(1+\delta)$ όπου $|\delta| \le \mathbf{u}$. □

β) Γνωρίζουμε ότι ο κλασικός δείκτης κατάστασης ενός μητρώου ως προς την επίλυση συστήματος $Ax = b$ ορίζεται ως $\kappa(A) := \|A\|\|A^{-1}\|$ για επιλεγμένη νόρμα. Να δείξετε ένα μητρώο 3×3 για το οποίο το $\kappa(A)$ είναι πάρα πολύ μεγάλο και το υπολογισμένο $\tilde{x}$ να έχει συγκριτικά πολύ μικρό σχετικό σφάλμα.

Απάντηση. Μπορείτε να διαλέξετε ένα διαγώνιο μητρώο A, με διαγώνιο $[1, 1, 1e-10]$, οπότε ο δείκτης κατάστασης είναι 1e10. Από την άλλη, αν λύσετε το σύστημα $Ax = b$, λόγω της διαγώνιας δομής του A, κάθε στοιχείο της λύσης x υπολογίζεται με μια διαίρεση, επομένως το άνω φράγμα για το σχετικό σφάλμα κάθε στοιχείου τής υπολογισμένης λύσης $\tilde{x}$ θα είναι $\mathbf{u}$. □

6. α) Έστω ότι ένα μητρώο $H \in \mathbb{R}^{n\times n}$ έχει μηδενικά στις θέσεις που βρίσκονται κάτω από την πρώτη υποδιαγώνιο, δηλ. $(3:n,1),(4:n,2),...,(n,n-1)$. Να δείξετε ότι (χωρίς οδήγηση και εφόσον υπάρχει) η παραγοντοποίηση LU του H κοστίζει $\Omega = \alpha n^2 + O(n)$. Επίσης να υπολογίσετε τον κυρίαρχο συντελεστή α.

Απάντηση. Προσέχουμε ότι σε κάθε βήμα $k = 1, ..., n-1$ της κλασικής απαλοιφής, χρειάζεται να απαλείψουμε μόνον ένα υποδιαγώνιο στοιχείο (στη θέση $(k+1,k)$). Επομένως το κόστος θα είναι $\Omega = \sum_{k=1}^{n-1}(1+\sum_{j=k+1}^{n} 2)$ επομένως $\Omega = n(n-1)+O(n)$ άρα $\alpha = 1$. Ο κώδικας μπορεί να είναι ο εξής (προαιρετικά):

```
for k=1:n-1
  H(k+1,k) = H(k+1,k)/H(k,k)
  for j=k+1:n
    H(k+1,k+1:n) = H(k+1,k+1:n) - H(k+1,k)*H(k+1,k+1:n)
  end
end
```

□

β) Δίδεται $A = \begin{pmatrix} 1 & 1 & 2 & 1 \\ 0 & 2 & 1 & -1 \\ 0 & 3 & -1 & 1 \\ 0 & 4 & 1 & 2 \end{pmatrix}$.

Να υπολογίσετε διάνυσμα Householder ώστε ο (ορθογώνιος) ανακλαστής P που παράγεται από το διάνυσμα, να μηδενίζει τη θέση (4,2) του μητρώου PA καθώς επίσης και του $B = PAP^\top$. Επίσης να υπολογίσετε το B (να φέρετε σε πέρας όλες τις αριθμητικές πράξεις.) Προσοχή: Δεν χρειάζεται (δεν είναι εφικτό) να είναι 0 το στοιχείο στη θέση (3,2).

Απάντηση. Σε MATLAB, $\mathtt{u} = [0;0;\mathtt{A}(3:4,2)] + [0,0,1,0]' * \mathtt{norm}(\mathtt{A}(3:4,2))$, επομένως $u = [0,0,8,4]^\top$ και υπολογίζεται ότι

$$B = \begin{pmatrix} 1 & 1 & -2 & -1 \\ 0 & 2 & 0.2 & -1.4 \\ 0 & -5 & 1.88 & -1.16 \\ 0 & 0 & -1.16 & -0.88 \end{pmatrix}$$

□

γ) Για κάθε A, μπορεί να υπολογιστεί (π.χ. η συνάρτηση `hess` στη MATLAB) ορθογώνιο μητρώο Q ως γινόμενο ανακλαστών Householder, ώστε το $QAQ^\top$ να έχει μηδενικά κάτω από την υποδιαγώνιο. Ο υπολογισμός των Q και $QAQ^\top$ κοστίζουν συνολικά περί τις $5n^3$ πράξεις α.κ.υ. Έστω ότι χρειάζεται να υπολογίσετε τις λύσεις $x_j, j = 1, ..., s$ των s συστημάτων $(A - \omega_j I)x_j = b_j$ όπου $A \in \mathbb{R}^{n\times n}$ και τα ω_j είναι πραγματικοί αριθμοί τέτοιοι ώστε τα μητρώα $A - \omega_j I$ να είναι αντιστρέψιμα και I το ταυτοτικό μητρώο. Να περιγράψετε τα βασικά βήματα αλγορίθμου που επιτυγχάνει τη λύση των s συστημάτων με κόστος $\Omega \approx 5n^3 + O(sn^2)$ αντί για $O(sn^3)$ που θα στοίχιζε αν χρησιμοποιούσατε απευθείας LU.

Απάντηση. ΒΙΒΛΙΟ □

7. Δίδεται η διαφορική εξίσωση $u''(x)+10^{-2}(20-u) = 0$ στο διάστημα $[0,10]$ με συνοριακές συνθήκες $u(0) = 40, u(10) = 200$ και θέλουμε να προσεγγίσουμε τη λύση με κεντρισμένες πεπερασμένες διαφορές και ακρίβεια τάξης $O(h^2)$, όπου h είναι η απόσταση μεταξύ των ισαπέχοντων κόμβων του πλέγματος που θα χρησιμοποιήσουμε στη διακριτοποίηση.

α) Να εξηγήσετε σύντομα γιατί συνήθως απαιτούμε από τη συνάρτηση $u(x)$ να έχει παραγώγους μέχρι και 4ης τάξης και αυτές να είναι συνεχείς στο διάστημα $[0,10]$.

Απάντηση. Από τη θεωρία γνωρίζουμε ότι η διακριτοποίηση βασίζεται στο συνδυασμό τιμών της συνάρτησης σε επιλεγμένους (γειτονικούς) κομβους του πλέγματος και στα σχετικά αναπτύγματα Taylor. Ειδικότερα, υπό την προϋπόθεση ότι η u διαθέτει τουλάχιστον 4 παραγώγους και συμβολίζοντας με u_j την τιμή της συνάρτησης στον κόμβο j ενός φυσικά αριθμημένου πλέγματος, μπορούμε να γράψουμε

$$u_{j\pm1} = u_j \pm hu_j^{(1)} + \frac{h^2}{2}u_j^{(2)} \pm \frac{h^3}{6}u_j^{(3)} + \frac{h^4}{24}u^{(4)}(x_i + \theta_i^{\pm}h)$$

όπου $-1 < \theta_i^- < 0 < \theta_i^+ < 1$. Επομένως

$$\boxed{u_{j-1} + u_{j+1} - 2u_j \;=\; h^2 u_j^{(2)} + \frac{h^4}{24}\left(u^{(4)}(\xi_j + \theta_i^+ h) + u^{(4)}(\xi_j + \theta_i^- h)\right)}$$

Επομένως, το σφάλμα διακριτοποίησης της 2ης παραγώγου σε κάθε σημείο εξαρτάται άμεσα από την διακριτοποίηση (δηλ. το h) και τη διακύμανση της τιμής του $|u^{(4)}|$. Το h το επιλέγεται από εμάς, επομένως μπορούμε να το επιλέξουμε όσο μικρό θέλουμε (μόνος περιορισμός είναι το μέγεθος του προκύπτοντος συστήματος) για να πετύχουμε αποδεκτό σφάλμα. Όμως, παράλληλα, θα πρέπει να αποκλείσουμε την περίπτωση να γίνεται το h πολύ μεγάλο. Αυτό εξασφαλίζεται «αυτόματα» όταν η συνάρτηση $u^{(4)}$ είναι συνεχής στο κλειστό διάστημα ορισμού της, καθώς τότε, από γνωστό στοιχειώδες θεώρημα της Μαθηματικής Ανάλυσης, έπεται ότι το $|u^{(4)}|$ θα είναι φραγμένο σε όλο το διάστημα. □

β) Να υπολογίσετε μητρώο $A \in \mathbb{R}^{4\times 4}$ και δεξιό μέλος $b \in \mathbb{R}^4$ τέτοια ώστε το διάνυσμα g που ικανοποιεί το σύστημα $Ag = b$ να προσεγγίζει τη λύση u στους κόμβους.

Απάντηση. Διαμερίζουμε το διάστημα $[0,10]$ σε 4 ισαπέχοντες εσωτερικούς κόμβους επομένως $h = 10/5 = 2$ και οι κόμβοι θα είναι $\xi_j = jh$ για $j = 1, ..., 4$. Χρησιμοποιώντας κεντρισμένες πεπερασμένες διαφορές 2ης τάξης για την προσέγγιση της 2ης παραγώγου θα έχουμε

$$\frac{u(\xi_{j-1}) - 2u(x_j) + u(\xi_{j+1})}{h^2} + 20\times 10^{-2} - 10^{-2}u(x_j) = 0$$

επομένως οι εξισώσεις σε κάθε σημείο καθορίζονται από τον τύπο

$$\frac{1}{h^2}U_{j-1} - \left(\frac{2}{h^2} + 10^{-2}\right)U_j + \frac{1}{h^2}U_{j+1} = -20\times 10^{-2}$$

που ξαναγράφουμε ως

$$-\frac{1}{4}U_{j-1} + \left(\frac{1}{2} + 10^{-2}\right)U_j - \frac{1}{4}U_{j+1} = 20\times 10^{-2}$$

Επομένως το σύστημα θα είναι

$$\begin{pmatrix} 0.51 & -0.25 & 0 & 0 \\ -0.25 & 0.51 & -0.25 & 0 \\ 0 & -0.25 & 0.51 & -0.25 \\ 0 & 0 & -0.25 & 0.51 \end{pmatrix} \begin{pmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{pmatrix} = \begin{pmatrix} 10.2 \\ 0.2 \\ 0.2 \\ 50.2 \end{pmatrix}$$

□

γ) Έστω ότι η παραπάνω διαφορική εξίσωση τροποποιείται σε $u''(x) + 10^{-2}(20 - u) - (1 + x^2) = 0$. Ποιοί θα είναι τώρα οι νέοι παράγοντες A και b;

Απάντηση. Για να ληφθεί υπόψη ο νέος παράγοντας $\boxed{1 + x^2}$, διαφοροποιείται μόνον το δεξιό μέλος: $b = [15.2, 17.2, 37.2, 115.2]^\top$. □

δ) Στη συνέχεια, αλλάζουμε τη συνοριακή συνθήκη του αρχικού προβλήματος (δηλ. του μέρους (α)) από $u(0) = 40$ σε $u'(0) = 40$. Χρησιμοποιώντας κεντρισμένες πεπερασμένες διαφορές 2ης τάξης να γράψετε το νέο σύστημα που θα προκύψει, έστω $\hat{A}\hat{g} = \hat{b}$. Προσοχή: Τα $\hat{A}, \hat{b}$ μπορεί να έχουν διαφορετικό μέγεθος από πριν.

Απάντηση. Με την αλλαγή αυτή δεν γνωρίζουμε πλέον το $u(0)$ αλλά την παράγωγο την οποία προσεγγίζουμε ως

$$\boxed{\frac{U_1 - U_{-1}}{2h} \approx u'(0) = 40 \Rightarrow U_{-1} = U_1 - 160}$$

θεωρώντας ότι U_{-1} είναι προσέγγιση του u στο -2. Επίσης, γράφουμε την εξίσωση για το σημείο 0, δηλ.

$$\boxed{-\frac{1}{4}U_{-1} + \left(\frac{1}{2} + 10^{-2}\right)U_0 - \frac{1}{4}U_1 = 20\times 10^{-2}}$$

οπότε

$$\boxed{-\frac{1}{4}(U_1 - 160) + (\frac{1}{2} + 10^{-2})U_0 - \frac{1}{4}U_1 = 20 \times 10^{-2}}$$

άρα επαυξάνουμε το αρχικό σύστημα ως εξής:

$$\begin{pmatrix} 0.51 & -0.5 & 0 & 0 & 0 \\ -0.25 & 0.51 & -0.25 & 0 & 0 \\ 0 & -0.25 & 0.51 & -0.25 & 0 \\ 0 & 0 & -0.25 & 0.51 & -0.25 \\ 0 & 0 & 0 & -0.25 & 0.51 \end{pmatrix} \begin{pmatrix} U_0 \\ U_1 \\ U_2 \\ U_3 \\ U_4 \end{pmatrix} = \begin{pmatrix} -39.8 \\ 0.2 \\ 0.2 \\ 0.2 \\ 50.2 \end{pmatrix}$$

□

8. Έστω η διαφορική εξίσωση $\boxed{u'''(t) = -1000u(t) - 300u'(t) - 30u''(t)}$ με αρχικές τιμές $u(0) = 1, u'(0) = 0, u''(0) = 1$. α) Να υπολογίσετε το $u(1.6)$ χρησιμοποιώντας εμπρός Euler και βήμα $h = 0.8$. (Προσοχή: Η εξίσωση είναι 3ης τάξης και είναι προτιμότερο να την μετατρέψετε σε γραμμικό σύστημα συνήθων διαφορικών εξισώσεων). β) Να εξηγήσετε αν με το παραπάνω βήμα μπορεί να παρουσιαστεί αστάθεια αν συνεχίσετε την προσέγγιση για πολλά βήματα και αν ναι, να υπολογίσετε άνω φράγμα για το βήμα h ώστε να αποφευχθεί η αστάθεια.

Απάντηση. α) Όπως προτείνεται μετατρέπουμε το παραπάνω σε σύστημα με την εισαγωγή βοηθητικών μεταβλητών (δείτε βιβλίο και διαφάνειες): $u_1(t) := u(t)$, $u'(t) := u_2(t)$, και $u''(t) := u_3(t)$ οπότε η διαφορική μετατρέπεται σε σύστημα 3 σύνηθων διαφορικών, ως εξής

$$\frac{d}{dt}\begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix} = -\begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -1 \\ 1000 & 300 & 30 \end{pmatrix} \begin{pmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{pmatrix}$$

ή για συντομία

$$\boxed{\frac{d}{dt}\mathbf{u} = -A\mathbf{u}}$$

όπου $\mathbf{u} := [u_1, u_2, u_3]^\top$ (παραλείπουμε το t το οποίο εννοείται). Εφαρμόζοντας εμπρός Euler με το βήμα $h = 0.8$ και $U(0) = [1, 0, 1]^\top$, ~~για να υπολογίσουμε~~ την τιμή στο $t = 2h$ έχουμε έχουμε ότι

$$\boxed{U(2h) = (I - hA)((I - hA)U(0)) = [\mathbf{1.64}, -657.6, 17937]^\top.}$$

Με παχειά γραφή έχουμε συμβολίσει το ζητούμενο, δηλ. την προσέγγιση στο $u(2h)$ με εμπρός Euler.

β) Προσέξτε ότι από τη διακύμανση των στοιχείων φαίνεται ότι μάλλον υπάρχει αστάθεια! Για να το επιβεβαιώσουμε, εξετάζουμε τη μέγιστη ιδιοτιμή του $I - hA$ για το βήμα h που χρησιμοποιήσαμε. Οι ιδιοτιμές του A είναι οι ρίζες του πολυωνύμου $1000 + 300\lambda + 30\lambda^2 + \lambda^3 = 0$, οπότε $\lambda_1 = \lambda_2 = \lambda_3 = -10$. Επομένως με $h = 0.8$ η φασματική ακτίνα του $I - hA$ θα είναι $7 = |1 - 0.8 \times 10|$ και θα έχουμε αστάθεια. Εδικότερα, το βήμα h πρέπει να επιλέγεται μικρότερο από $2/\max|\lambda_j| = 0.5$. □

γ) Γενικά στην Euler για την επίλυση ενός γραμμικού προβλήματος του τύπου $u' = -Au$, είναι σωστό ή λάθος ότι αν μειωθεί το βήμα στο μισό, τότε το μέγιστο ολικό σφάλμα διακριτοποίησης θα υποτετραπλασιαστεί.

Απάντηση. ΛΑΘΟΣ, το ολικό σφάλμα συμπεριφέρεται όπως το $O(1/h)$ άρα περιμένουμε να υποδιπλασιαστεί. □

δ) Για καθένα από τα παρακάτω σχετικά με τις άμεσες μεθόδους Runge-Kutta τάξης 2 και πάνω για την επίλυση της ΣΔΕ $u'(t) = f(t, u)$, να κυκλώσετε αν είναι σωστό ή λάθος:

(Σ - Λ) Προβλέπουν τη νέα τιμή συνδυάζοντας την προσέγγιση στο t_k με προσεγγίσεις της παραγώγου της u σε μια ή περισσότερες τιμές του t στο διάστημα $[t_k, t_{k+1})$.

Απάντηση. ΣΩΣΤΟ, οι μέθοδοι RK είναι μονοβηματικές και χρησιμοποιούν ως πληροφορία την προσέγγιση στο t_k με εκτιμήσεις της παραγώγου στο t_k και αλλα σημεια στο παραπάνω διάστημα. Ο γενικός τύπος είναι

$$\boxed{U_{n+1} = U_k + h\sum_{i=1}^{s} b_i K_i}$$

Ρίζες εξίσωσης τρίτου βαθμού

$d^3+30d^2+300d+1000=0$. Μαντεύουμε μια ρίζα, π.χ. το −10 και μετά εφαρμόζουμε Horner, ως εξής:

1	30	300	1000	−10
	−10	−200	−1000	
1	20	100	0	

$\downarrow$

$(d-(-10))(d^2+20d+100)=0 \Rightarrow (d+10)(d^2+20d+100)=0$

Στηριζόμαστε πάνω στον τρόπο λύσης που υπάρχει στις Διαφάνειες σελ. 5 (Set22). Πιο συγκεκριμένα:

Ξέρουμε ότι προκύπτει το σύστημα Σ.Δ.Ε.

$\left\{\begin{aligned} n_1&=u_1=u, \\ n_2&=u_2=u', \\ n_3&=u_3=u'' \end{aligned}\right\}$

$n_1'=u_1'=n_2=u_2 \Rightarrow u_1'=u_2$

$n_2'=u_2'=n_3=u_3 \Rightarrow u_2'=u_3$

$n_3'=u_3'=\underline{f(t,u,u',u'')}$ και μάλιστα αν δούμε από τη σελ. 5-6, διαπιστώνουμε ότι η λύση της άσκησης που υπάρχει περιέχει τους συντελεστές της τελευταίας διαφορικής που δίνεται.

Άρα

$$u_1'=\frac{du_1}{d(t)}=0u_1+u_2+0u_3$$

$$u_2'=\frac{du_2}{d(t)}=0u_1+0\cdot u_2+u_3$$

$$u_3'=\frac{du_3}{d(t)}=-1000u_1-300u_2-30u_3$$

$$\Rightarrow \frac{d}{dt}\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}=\underbrace{\begin{bmatrix}0&1&0\\0&0&1\\-1000&-300&-30\end{bmatrix}}_{A}\underbrace{\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}}_{u} \Rightarrow$$

$$\frac{du}{dt}=-Au \quad \text{όπου } u=\begin{bmatrix}u_1\\u_2\\u_3\end{bmatrix}$$

$u''(0)=u_3(0)=1$

Εμείς ξέρουμε τις αρχικές τιμές, δηλ. $u_1(0)=1$, $u'(0)=u_2(0)=0$, $u''(0)=u_3(0)=1$. Οπότε έχουμε τη ΣΔΕ $\boxed{\frac{du(t)}{dt}=-Au(t)}$ όπου $A=[0\,1\,0;\,0\,0\,1;\,1000\;300\;30]$ και

Άρα το $u=\begin{bmatrix}1\\0\\1\end{bmatrix}$.

$u=[u_1\,u_2\,u_3]^T=[1\,0\,1]^T$ και $u_1(0)=1$, $u_2(0)=0$, $u_3(0)=1$

Τότε σύμφωνα με τη λύση της άσκησης σελ. 3 (Σεπτεμβρίου 2005) έχουμε:

$$\begin{bmatrix}U_1^{(j+1)}\\U_2^{(j+1)}\\U_3^{(j+1)}\end{bmatrix}=(I-Ah)U^{(j)}=\left(\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}-h\begin{bmatrix}0&-1&0\\0&0&-1\\1000&300&30\end{bmatrix}\right)\begin{bmatrix}U_1^{(j)}\\U_2^{(j)}\\U_3^{(j)}\end{bmatrix}=\left(\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}-\begin{bmatrix}0&-0{,}8&0\\0&0&-0{,}8\\1000\cdot0{,}8&300\cdot0{,}8&30\cdot0{,}8\end{bmatrix}\right)\begin{bmatrix}U_1^{(j)}\\U_2^{(j)}\\U_3^{(j)}\end{bmatrix}$$

$$\begin{bmatrix}1&+0.8&0\\0&1&+0.8\\-800&-240&+23\end{bmatrix}\begin{bmatrix}U_1^{(j)}\\U_2^{(j)}\\U_3^{(j)}\end{bmatrix}$$

για $t=0$: $\begin{bmatrix}1&+0.8&0\\0&1&+0.8\\-800&-240&-23\end{bmatrix}\begin{bmatrix}1\\0\\1\end{bmatrix}=\underset{3\times1}{\begin{bmatrix}1\\+0.8\\-823\end{bmatrix}}$

$\begin{array}{r}0{,}8\\30\cdot\\\hline 24\end{array}$

για $t=1.6$ το $u(1.6)$ έχουμε: $U^{1.6}=\begin{bmatrix}1&+0.8&0\\0&1&+0.8\\ \text{[illegible]}\end{bmatrix}\begin{bmatrix}1\\+0.8\\ \text{[illegible]}\end{bmatrix}=$