Εργασία 3

Χαράλαμπος Αναστασίου

Νοέμβριος 2024

1 Εισαγωγή

Ερωτήματα που απαντήθηκαν: 1, 2, 3, 4, 5, 6, 7

2 Απαντήσεις Ερωτημάτων

2.1 Ερώτημα 1

(GvL A3.1.2) Υποθέτουμε ότι το $L = I_n - N$ είναι μοναδιαίο κάτω τριγωνικό, με $N \in \mathbb{R}^{n\times n}$. Να δείξετε ότι:

$$L^{-1} = I_n + N + N^2 + \cdots + N^{n-1}.$$

Ποια είναι η τιμή του $\|L^{-1}\|_F$ αν έχουμε $N_{ij} = 1$ για κάθε $i > j$?

ΑΠΑΝΤΗΣΗ:

Για να αποδείξουμε ότι $L^{-1} = I_n + N + N^2 + \cdots + N^{n-1}$,
αρκεί να εξασφαλίσουμε ότι ισχύει:

$$LL^{-1} = I_n$$

Έχουμε:

$$LL^{-1} = (I_n - N)(I_n + N + N^2 + \cdots + N^{n-1}).$$

Κάνοντας τις πράξεις προκύπτει:

$$LL^{-1} = I_n + N + N^2 + \cdots + N^{n-1} - N - N^2 - N^3 - \cdots - N^n.$$

Μπορούμε να παρατηρήσουμε ότι οι όροι $N, N^2, \ldots, N^{n-1}$ και $-N, -N^2, \ldots, -N^{n-1}$ θα ακυρωθούν. Έτσι, απομένει:

$$I_n - N^n.$$

Αλλά $N^n = 0$ επειδή το N είναι αυστηρά κάτω τριγωνικό. Το γεγονός ότι είναι αυστηρά κάτω τριγωνικό προκύπτει από τον τύπο $L = I_n - N$ όπου το L είναι μοναδιαίο κάτω τριγωνικό μητρώο(επομένως προκειμένου να εξασφαλιστεί η μοναδιαία διαγώνιος του L πρέπει το N να έχει μηδενική διαγώνιο, άρα να είναι αυστηρά κάτω τριγωνικό). Έτσι, έχουμε:

$$LL^{-1} = I_n.$$

Απεδείχθη επομένως ότι

$$L^{-1} = I_n + N + N^2 + \cdots + N^{n-1}.$$

2.2 Ερώτημα 2

Δίνεται το συμμετρικό μητρώο 4×4:

$$\begin{pmatrix} a_{11} & a_{12} & | & a_{13} & a_{14} \\ a_{12} & a_{22} & | & a_{23} & a_{24} \\ a_{13} & a_{23} & | & a_{33} & a_{34} \\ a_{14} & a_{24} & | & a_{34} & a_{44} \end{pmatrix}$$

Να δείξετε την αποθήκευση του μητρώου στη μορφή RFP (rectangular full packed format).

ΑΠΑΝΤΗΣΗ:

Στη μορφή RFP, αποθηκεύουμε μόνο τις μοναδικές τιμές του συμμετρικού μητρώου, καθώς οι τιμές πάνω από την κύρια διαγώνιο είναι ίσες με τις αντίστοιχες τιμές κάτω από αυτήν. Θα αποθηκεύσουμε επομένως τα στοιχεία ως εξής:

Κύρια διαγώνιος: $a_{11}, a_{22}, a_{33}, a_{44}$

Στοιχεία πάνω από την κύρια διαγώνιο:

$$a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34}$$

Η αποθήκευση του μητρώου A στη μορφή RFP θα είναι:

$$\mathrm{RFP} = \{a_{11}, a_{22}, a_{33}, a_{44}, a_{12}, a_{13}, a_{14}, a_{23}, a_{24}, a_{34}\}$$

2.3 Ερώτημα 3

Να γράψετε συνάρτηση MATLAB po2pp που λαμβάνει ως είσοδο ένα συμμετρικό μητρώο που δίνεται σε κανονική αποθήκευση και επιστρέφει το μητρώο σε packed μορφή ως διάνυσμα. Ο κώδικας θα πρέπει να επιτρέπει στο χρήστη να επιλέξει κατά πόσον το packing θα είναι column ή row-major και κατά πόσον αποθηκεύεται το άνω ή το κάτω τριγωνικό τμήμα.

ΑΠΑΝΤΗΣΗ:

Δες την απάντηση στον κώδικα του .mlx αρχείου

2.4 Ερώτημα 4

(GvL A4.2.10) Έστω ότι $A = I + uu^T$ όπου $A \in \mathbb{R}^{n \times n}$ και $\|u\|_2 = 1$. Να γράψετε ακριβείς τύπους για τη διαγώνιο και την υποδιαγώνιο του παράγοντα Cholesky του A.

ΑΠΑΝΤΗΣΗ:

Αρχικά θα εξετάσω γιατί το μητρώο είναι θετικά ορισμένο και άρα μπορώ να εφαρμόσω Cholesky (ΔΕΝ ΖΗΤΕΙΤΑΙ - ΜΟΝΟ ΓΙΑ ΠΡΟΣΩΠΙΚΗ ΚΑΤΑΝΟΗΣΗ).

Σύμφωνα με το Θεώρημα 4.2.3 του GvL, ένα μητρώο A είναι θετικά ορισμένο αν και μόνο αν το συμμετρικό $T = \frac{A+A^T}{2}$ έχει θετικές ιδιοτιμές.

Κάνοντας τις πράξεις διαπιστώνω ότι ισχύει:

$$T = \frac{A + A^T}{2} = A$$

Το μητρώο I είναι το μοναδιαίο, το οποίο είναι θετικά ορισμένο, καθώς έχει μόνο ιδιοτιμές ίσες με 1.

Το μητρώο uu^T είναι θετικά ημι-ορισμένο (positive semi-definite). Αυτό σημαίνει ότι για κάθε μη μηδενικό διάνυσμα x, ισχύει ότι:

$$x^T(uu^T)x \geq 0$$

Πράγματι, η τιμή του $x^T(uu^T)x$ είναι:

$$x^T(uu^T)x = (u^T x)(u^T x) = (u^T x)^2 \geq 0$$

Για να εξετάσουμε αν το $I + uu^T$ είναι θετικά ορισμένο, θα πρέπει να δούμε τις ιδιοτιμές του:

Οι ιδιοτιμές του $I + uu^T$ προκύπτουν από τις ιδιοτιμές του uu^T, αλλά προστίθεται 1 (δηλαδή η ιδιοτιμή του μοναδιαίου μητρώου I). Το uu^T είναι τάξης-1 και έχει μία ιδιοτιμή ίση με 1(η μοναδική ιδιοτιμή ισούται με τη νόρμα-2 του u) και $n-1$ ιδιοτιμές ίσες με 0. Άρα οι ιδιοτιμές του $I + uu^T$ θα είναι:

- Μία ιδιοτιμή $1 + 1 = 2$ (που αντιστοιχεί στην ιδιοτιμή του uu^T που είναι 1).

- $n - 1$ ιδιοτιμές $1 + 0 = 1$ (που αντιστοιχούν στις μηδενικές ιδιοτιμές του uu^T).

Άρα οι ιδιοτιμές του $I + uu^T$ είναι 2 και 1 (με πολλαπλότητα $n - 1$).

Επειδή όλες οι ιδιοτιμές του $I+uu^T$ είναι θετικές (2 και οι $n-1$ ιδιοτιμές 1), το μητρώο $I + uu^T$ είναι θετικά ορισμένο(και μάλιστα συμμετρικά θετικά ορισμένο).

Τώρα όσον αφορά τα ζητούμενα του ερωτήματος:

Έχουμε ότι: $A_{ij} = I_{ij} + u_i u_j$ και επομένως:

$$A_{ii} = 1 + u_i^2 \quad \text{και} \quad A_{ij} = u_i u_j \quad \text{για} \quad i \neq j$$

Για τη διαγώνιο του L:

$$l_{ii} = \sqrt{1 + u_i^2 - \sum_{k=1}^{i-1} l_{ik}^2}$$

Για την υποδιαγώνιο του L:

$$l_{i+1,i} = u_i u_{i+1} \quad \text{για} \quad i \neq j$$

2.5 Ερώτημα 5

Έστω ότι $D = \text{diag}(d_1, \ldots, d_n)$, με $d_i > 0$ για κάθε i. Να γράψετε έναν αποδοτικό αλγόριθμο για τον υπολογισμό του μεγαλύτερου στοιχείου του μητρώου $(D + CC^T)^{-1}$, όπου $C \in \mathbb{R}^{n\times r}$. Υπόδειξη: Να χρησιμοποιήσετε τον τύπο Sherman-Morrison-Woodbury.

ΑΠΑΝΤΗΣΗ:

Ο τύπος Sherman-Morrison-Woodbury είναι:

$$(B + UV^T)^{-1} = B^{-1} - B^{-1}U\left(I + V^T B^{-1} U\right)^{-1} V^T B^{-1}$$

Για το πρόβλημά μας, έχουμε το $B = D$, $U = C$, και $V = C$, άρα ο τύπος γίνεται:

$$(D + CC^T)^{-1} = D^{-1} - D^{-1}C\left(I + C^T D^{-1} C\right)^{-1} C^T D^{-1}$$

Δεδομένου ότι το D είναι διαγώνιο, μπορούμε να υπολογίσουμε το αντίστροφο του D εύκολα ως το διαγώνιο μητρώο

$$D^{-1} = \text{diag}\left(\frac{1}{d_1}, \frac{1}{d_2}, \ldots, \frac{1}{d_n}\right)$$

Υπολογισμός του $I + C^T D^{-1} C$: Αυτό είναι ένα $r \times r$ μητρώο, το οποίο μπορεί να υπολογιστεί ως το γινόμενο

$$C^T D^{-1} C$$

και προσθέτοντας την μοναδιαία μήτρα I.

Στη συνέχεια, εφόσον το $I + C^T D^{-1} C$ είναι μικρότερο μητρώο (διαστάσεων $r \times r$), ο υπολογισμός του αντίστροφου είναι πιο αποδοτικός και μπορεί να γίνει με χρήση της εντολής inv() με σημαντικά μικρότερο κόστος σε μνήμη και χρόνο.

Στο αρχείο .mlx υπάρχει ο αντίστοιχος αλγόριθμος ο οποίος υλοποιεί τα παραπάνω, καθώς και επιπλέον σύγκριση του χρόνου μεταξύ της άμεσης αντιστροφής με την απευθείας χρήση της inv() και της αντιστροφής και εύρεσης του μέγιστου στοιχείου με χρήση της μεθόδου Sherman-Morrison-Woodbury. Παρατηρούμε ότι η δεύτερη μέθοδος είναι σημαντικά γρηγορότερη (για τη σύγκριση των χρόνων έπρεπε κανονικά να μην συμπεριλάβω την εύρεση του μέγιστου στοιχείου αλλά λόγω του ότι πρόκειται για γρήγορη πράξη την συμπεριέλαβα στις μετρήσεις).

2.6 Ερώτημα 6

(GvL A4.3.1) Να αναπτύξετε μια εκδοχή του Αλγόριθμου 4.3.1 με την υπόθεση ότι το μητρώο A αποθηκεύεται σε μορφή ζώνης. (Δείτε την Ενότητα §1.2.5.)

ΑΠΑΝΤΗΣΗ: Δες το .mlx αρχείο.

2.7 Ερώτημα 7

A4.3.9 Έστω ότι το συμμετρικό και θετικά ορισμένο μητρώο $A \in \mathbb{R}^{n\times n}$ έχει «δομή βέλους», δηλαδή:

$$A = \begin{pmatrix} \times & \times & \times & \times & \times \\ \times & \times & 0 & 0 & 0 \\ \times & 0 & \times & 0 & 0 \\ \times & 0 & 0 & \times & 0 \\ \times & 0 & 0 & 0 & \times \end{pmatrix}$$

(α) Να δείξετε πώς μπορούμε να λύσουμε το γραμμικό σύστημα $Ax = b$ με $O(n)$ flop χρησιμοποιώντας τον τύπο Sherman-Morrison-Woodbury.

(β) Να προσδιορίσετε ένα μητρώο μετάθεσης P τέτοιο ώστε $PAP^T = GG^T$, το οποίο να μπορούμε να υπολογίσουμε με $O(n)$.

ΑΠΑΝΤΗΣΗ:

Το συγκεκριμένο μητρώο $A \in \mathbb{R}^{n\times n}$ έχει δομή «βέλους». Παρατηρώ ότι μπορώ να το διακρίνω σε δύο συστατικά στοιχεία: ένα διαγώνιο υπομητρώο $A(2:n, 2:n)$ και ένα μητρώο τάξης-1 (rank-1):

$$A = D + uv^T$$

όπου:

- D είναι ένας διαγώνιος πίνακας που περιέχει τις διαγώνιες τιμές του A, εκτός από την πρώτη στήλη και γραμμή.

- u και v είναι διανύσματα που αντιπροσωπεύουν την πρώτη στήλη και γραμμή αντίστοιχα.

Η μέθοδος Sherman-Morrison-Woodbury μας επιτρέπει να υπολογίσουμε το αντίστροφο του πίνακα A ως εξής:

$$A^{-1} = D^{-1} - D^{-1}u\left(I + v^T D^{-1}u\right)^{-1} v^T D^{-1}$$

Για να λύσουμε το σύστημα $Ax = b$ με $O(n)$ flops, ακολουθούμε τα εξής βήματα:

1. Επειδή το D είναι διαγώνιο μητρώο, ο υπολογισμός του $D^{-1}b$ γίνεται σε $O(n)$ flops.

2. Υπολογίζουμε τους όρους $D^{-1}u$ και $v^T D^{-1}$, οι οποίοι έχουν κόστος $O(n)$ καθένας, καθώς περιλαμβάνουν πράξεις με διαγώνιους πίνακες.

3. Υπολογίζουμε τον όρο $\left(I + v^T D^{-1}u\right)^{-1}$. Εντός της παρένθεσης έχουμε μια πράξη εσωτερικού γινομένου μεταξύ του $v^T D^{-1}$ και του u που απαιτεί $O(n)$ flops και την πρόσθεση με το ταυτοτικό μητρώο, άρα σύνολο $O(n)$ flops. Ο υπολογισμός τώρα του αντιστρόφου αυτής της έκφρασης παίρνει $O(n)$ flops λόγω της ειδικής δομής (πρόκειται για μητρώο τάξης 1 προστιθέμενο στον μοναδιαίο πίνακα). Αυτό μπορώ να το δω και αν εφαρμόσω τη μέθοδο Sherman-Morrison-Woodbury αντικαθιστώντας όπου $B = I$, είναι φανερό ότι οι πράξεις είναι O(n).

Συνεπώς, ο συνολικός αριθμός των πράξεων για την επίλυση του συστήματος $Ax = b$ είναι $O(n)$, λόγω της δομής του μητρώου A.

2.8 Ερώτημα 8

(A4.5.3) Πώς θα λύνατε ένα σύστημα της μορφής:

$$\begin{bmatrix} D_1 & F_1 \\ E_1 & D_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

όπου τα D_1 και D_2 είναι διαγώνια ενώ τα F_1 και E_1 είναι τριδιαγώνια; Υπόδειξη: Να χρησιμοποιήσετε τη μετάθεση τέλειας αναδιάταξης.

ΑΠΑΝΤΗΣΗ: