Προαιρετική Άσκηση 1

1.) ***Να περιγράψετε σε μια παράγραφο τα χαρακτηριστικά του συστήματος που χρησιμοποιείτε (υλικό και λογισμικό): Σύστημα, επεξεργαστή, επίπεδα μνήμης και μέγεθός τους.***

Το σύστημα που χρησιμοποιώ είναι ένα **laptop** με λειτουργικό σύστημα **Microsoft Windows 10 Pro Education (έκδοση 10.0.19045)**. Ο επεξεργαστής είναι ένας **AMD Athlon Gold 3150U με Radeon Graphics**, ο οποίος διαθέτει **2 πυρήνες και 4 λογικούς επεξεργαστές με συχνότητα 2.40 GHz**. Η μνήμη **RAM** του συστήματος είναι **4 GB**. Τα επίπεδα της **cache** μνήμης είναι 192 KB (L1), 1 MB (L2) και 4 MB (L3).

Σχόλιο: Το σύστημα αυτό είναι επαρκές για βασικές εργασίες σε περιβάλλον MATLAB, αν και ενδέχεται να υπάρξουν περιορισμοί σε πιο απαιτητικά έργα λόγω της χαμηλής μνήμης RAM.

2.) ***(Άσκηση A1.1.2 από GvL) Σε έναν συνήθη πολλαπλασιασμό μητρώων 2 × 2 της μορφής C = AB, υπάρχουν οκτώ πολλαπλασιασμοί: aij βjk , i, j, k = 1 , 2. Να φτιάξετε έναν πίνακα που να δείχνει τη σειρά με την οποία εκτελούνται οι πολλαπλασιασμοί αυτοί από τους αλγόριθμους ijk, jik, kij, ikj, jki, και kji πολλαπλασιασμού μητρώων (βλ. διάλεξη 4).***

Σε έναν συνήθη πολλαπλασιασμό δύο μητρώων 2×2, δηλαδή C=AB, κάθε στοιχείο του πίνακα C υπολογίζεται με βάση τα στοιχεία των πινάκων A και B. Ο πολλαπλασιασμός γίνεται ως εξής:

Για τα στοιχεία του C:

C11=A11B11+A12B21

C12=A11B12+A12B22

C21=A21B11+A22B21

C22=A21B12+A22B22

Αυτός ο πολλαπλασιασμός περιλαμβάνει οκτώ επιμέρους πολλαπλασιασμούς, που είναι οι:

A11B11, A12B21, A11B12, A12B22, A21B11, A22B21, A21B12, A22B22

Ο τρόπος με τον οποίο εκτελούνται αυτοί οι πολλαπλασιασμοί μπορεί να διαφέρει ανάλογα με τη σειρά πρόσβασης στους δείκτες των πινάκων από τους διαφορετικούς αλγόριθμους πολλαπλασιασμού. Οι αλγόριθμοι αυτοί είναι οι: ijk, jik, kij, ikj, jki, και kji. Κάθε αλγόριθμος καθορίζει τη σειρά με την οποία θα διατρέξουμε τα στοιχεία των πινάκων. Παρακάτω δίνεται ο ζητούμενος πίνακας:

Αλγόριθμος	1ος	2ος	3ος	4ος	5ος	6ος	7ος	8ος
ijk	A11B11	A12B21	A11B12	A12B22	A21B11	A22B21	A21B12	A22B22
jik	A11B11	A12B21	A21B11	A22B21	A11B12	A12B22	A21B12	A22B22
kij	A11B11	A11B12	A21B11	A21B12	A12B21	A12B22	A22B21	A22B22
ikj	A11B11	A11B12	A12B21	A12B22	A21B11	A21B12	A22B21	A22B22
jki	A11B11	A21B11	A12B21	A22B21	A11B12	A21B12	A12B22	A22B22
kji	A11B11	A21B11	A11B12	A21B12	A12B21	A22B21	A12B22	A22B22

3.) (Άσκηση A1.1.3 από GvL) Να γράψετε έναν αλγόριθμο για τον οποίο Ω = O (n^2) για τον υπολογισμό του μητρώου C = (xy⊤) k , όπου τα x και y είναι διανύσματα μεγέθους n και η μεταβλητή k είναι θετικός ακέραιος αριθμός.

- Κάθε στοιχείο του νέου πίνακα απαιτεί n πολλαπλασιασμούς και n−1 προσθέσεις. Δηλαδή, για κάθε στοιχείο χρειαζόμαστε O(n) πράξεις.

- Δεδομένου ότι έχουμε n^2 στοιχεία στον πίνακα, η συνολική πολυπλοκότητα του πολλαπλασιασμού είναι $O(n^2)\cdot O(n)=O(n^3)$. Ωστόσο, στην περίπτωση του $C=xy^T$, μπορούμε να εκμεταλλευτούμε την ειδική δομή του πίνακα.

  Η δύναμη του πίνακα C μπορεί να υπολογιστεί πιο αποτελεσματικά ως εξής:

1. **Πρώτη Δύναμη**:
   $C=xy^T$

2. **Δεύτερη Δύναμη**:
   $C^2 = C \cdot C = (xy^T)(xy^T) = x(y^Tx)y^T$
   Εδώ, το y^Tx υπολογίζεται ως εσωτερικό γινόμενο του y και του x. Έτσι, το C^2 γίνεται:

   $C^2=x(y^Tx)y^T$

   Γενική Δύναμη:
   Ομοίως, για οποιοδήποτε k: $C^k=x(y^Tx)^{k-1}y^T$
   Αυτό σημαίνει ότι η εκτίμηση της k-οστής δύναμης του C εξαρτάται μόνο από το εσωτερικό γινόμενο y^Tx και είναι πολύ πιο αποτελεσματική από τον κλασικό υπολογισμό $C \cdot C \cdots C$ k φορές.

   Ο υπολογισμός y^Tx απαιτεί O(n) και είναι πολύ πιο γρήγορος από το $O(n^3)$ που θα απαιτούσε ο κλασικός πολλαπλασιασμός πινάκων.

Αλγόριθμος σε MATLAB:

```
function C = compute_matrix_power(x, y, k)

    % x: Διάνυσμα στήλη μεγέθους n
    % y: Διάνυσμα στήλη μεγέθους n
    % k: Θετικός ακέραιος

    % Υπολογισμός του εξωτερικού γινομένου xy^T ( O(n^2) )
    C = x * y';

    % Επαναληπτικός υπολογισμός της δύναμης του πίνακα C
    for i = 2:k
        C = C * (x * y');
     end
end
```

4.) ***(Άσκηση A1.1.4 από GvL) Έστω D = ABC, όπου $A \in R^{n1 \times n2}$, $B \in R^{n2 \times n3}$, και $C \in R^{n3 \times n4}$. Να συγκρίνετε το Ω ενός αλγόριθμου που υπολογίζει το D μέσω του τύπου D = (AB) C σε σχέση με το Ω ενός αλγόριθμου που - αξιοποιώντας την προσεταιριστική ιδιότητα για τους πολλαπλασιασμούς μητρώων - υπολογίζει το D χρησιμοποιώντας τον τύπο D = A (BC). Σε ποια περίπτωση η πρώτη διαδικασία είναι πιο αποδοτική από τη δεύτερη ως προς το Ω ;***

Υπολογισμός του D=(AB)C

Αρχικά υπολογίζουμε το AB:

- Το A έχει διαστάσεις n1×n2 και το B έχει διαστάσεις n2×n3.
- Ο αριθμός των πράξεων για το πολλαπλασιασμό AB είναι $n1 \cdot n2 \cdot n3$.

Στη συνέχεια υπολογίζουμε το (AB)C:

- Το αποτέλεσμα AB έχει διαστάσεις n1×n3, και το C έχει διαστάσεις n3×n4.
- Ο αριθμός των πράξεων για το πολλαπλασιασμό (AB)C είναι $n1 \cdot n3 \cdot n4$.

Συνολικός αριθμός πράξεων για τον αλγόριθμο D=(AB)C: $n1 \cdot n2 \cdot n3 + n1 \cdot n3 \cdot n4$

2. Υπολογισμός του D=A(BC)

Αρχικά υπολογίζουμε το BC:

- Το B έχει διαστάσεις n2×n3 και το C έχει διαστάσεις n3×n4.
- Ο αριθμός των πράξεων για το πολλαπλασιασμό BC είναι $n2 \cdot n3 \cdot n4$.

Στη συνέχεια υπολογίζουμε το A(BC):

- Το αποτέλεσμα BC έχει διαστάσεις n2×n4, και το A έχει διαστάσεις n1×n2.
- Ο αριθμός των πράξεων για το πολλαπλασιασμό A(BC) είναι $n1 \cdot n2 \cdot n4$.

Συνολικός αριθμός πράξεων για τον αλγόριθμο D=A(BC): $n2 \cdot n3 \cdot n4 + n1 \cdot n2 \cdot n4$

Σύγκριση

Για να είναι η πρώτη διαδικασία πιο αποδοτική από την δεύτερη ως προς το Ω, αναζητούμε τις συνθήκες υπό τις οποίες ισχύει:

$$n1 \cdot n2 \cdot n3 + n1 \cdot n3 \cdot n4 < n2 \cdot n3 \cdot n4 + n1 \cdot n2 \cdot n4$$

Σημειώνεται ότι δεν λήφθηκαν στα παραπάνω υπόψιν οι προσθέσεις, οι οποίες ανήκουν προφανώς στο Ω, λόγω του ότι είναι περίπου ίδιο το πλήθος τους με τους πολλαπλασιασμούς με διαφορά +-1.

- Όταν όλες οι μεταβλητές είναι ίσες, η παραπάνω ανισότητα τείνει να μετατραπεί σε ισότητα (το δοκίμασα σε πρόχειρο με συνδυασμό τιμών). Οι αλλαγές στα n1 και n3 έχουν άμεση επίδραση στον αριθμό των πράξεων που απαιτούνται στην αριστερή πλευρά της ανισότητας, κάνοντάς τα πιο ευαίσθητα στην ανισότητα σε σύγκριση με τις άλλες παραμέτρους (με άλλα λόγια, αν αυξήσω σημαντικά τα n1 και n3 σε σχέση με τα n2 και n4, η ανισότητα παύει να ισχύει). Επομένως, επιθυμώ μεγάλα n2 και n4, ενώ παράλληλα συγκρατώ τα n1 και n3 από το να πάρουν μεγάλες (συγκριτικά με τα n2,n4) τιμές.

5.) (Άσκηση A1.5.3 από GvL) Δίνεται $A \in R^{n \times n}$ που είναι αποθηκευμένο σε διάταξη ανά στήλη και ότι m = m1M και n = n1N. Θεωρούμε το A ως ένα σύνθετο μητρώο M επί N με μπλοκ m 1 επί n1. Να υλοποιήσετε έναν αλγόριθμο storeInBlocks(A, m1,n1,M,N) σε MATLAB για την αποθήκευση του A σε ένα διάνυσμα A.block(1:mn) με την ιδιότητα ότι κάθε μπλοκ A_{IJ} αποθηκεύεται συνεχόμενα στη διάταξη ανά στήλη. Εξηγήστε τα βήματα του αλγορίθμου και να δείξετε ότι εκτελεί σωστά τη διεργασία χρησιμοποιώντας τα δεδομένα:

```
1     m1 = 4; M = 2;
2     m=m1*M; n=m; N=M; n1=m1;
3     A1 = reshape ( [ 1:m*n ] ,m, n);
4     A1_block = storeInBlocks (A1, m1, n1 ,M, N);
5
6     m1=4; M=2; n1=3; N=3;
7     m = m1*M; n = n1*N;
8     A2 = reshape ( [ 1 : n*m] ,m, n) ;
9     A2_block = storeInBlocks (A2, m1, n1 ,M,N) ;
```

Επιβεβαιώστε ότι οι τιμές στα A?_block είναι σωστές!

Βήματα του αλγόριθμου

Αρχικοποίηση:

1.) Δημιουργούμε ένα διάνυσμα A_block το οποίο θα έχει μέγεθος mxn.

2.) Για κάθε μπλοκ (i,j) όπου i κυμαίνεται από 0 έως M−1 και j από 0 έως N−1, υπολογίζουμε τους δείκτες για το κάθε μπλοκ Aij ως εξής:

block_row = i * m1

block_col = j * n1

(Αφού προηγουμένως έχω κάνει εφαρμογή της reshape πάνω στη μήτρα A), προκειμένου να εξάγω το κάθε block διαστάσεων m1 x n1 γράφω την παρακάτω γραμμή κώδικα:

```
block = A(block_row + 1:block_row + m1,block_col + 1:block_col + n1);
```

Στη συνέχεια, εφόσον έχω πλέον εξάγει το κατάλληλων διαστάσεων block, οφείλω να το αποθηκεύσω κατά στήλες στις σωστές θέσεις σε ένα διάνυσμα στήλης A_block (τελικό διάνυσμα εξόδου). Αυτό μπορεί να επιτευχθεί με τη χρήση ενός δείκτη ο οποίος κάθε φορά θα ανανεώνεται με βάση την νέα αρχική θέση του κάθε block.

```
A_block(index:index + m1*n1 - 1)
```

Σε αυτές τις θέσεις λοιπόν αποθηκεύω το κάθε block κατά στήλες, με την εντολή block(:), η οποία θα μετατρέψει το μητρώο/block διαστάσεων m1 x n1 σε διάνυσμα στήλη με τις αρχικές στήλες 'stacked' τη μία πάνω στην άλλη. Η τελική εντολή άρα είναι:

```
A_block(index:index + m1*n1 - 1) = block(:);
```

Ενιαίος Κώδικας Συνάρτησης:

```
function A_block = storeInBlocks(A, m1, n1, M, N)

   [m, n] = size(A);
   A_block = zeros(m1 * n1 * M * N, 1);
   index = 1;                    % δείκτης για το διάνυσμα εξόδου A_block

   for i = 0:M-1
      for j = 0:N-1
         block_row = i * m1;
         block_col = j * n1;

         block = A(block_row+1:block_row+m1, block_col+1:block_col+n1);
         A_block(index:index + m1*n1 - 1) = block(:);
         index = index + m1 * n1;          % ανανέωση του δείκτη για το A_block

      end
   end
end
```

6.) Αναθέστε τα τελευταία 4 ψηφία του ΑΜ σας στη μεταβλητή ΑΜ σε περιβάλλον MATLAB και στη συνέχεια εκτελέστε τις παρακάτω εντολές:

```
1       rng (AM) ;
2      p = randperm (14) ; id = p ( 1 : 4 ) ;
```

Στη συνέχεια δείτε την εικόνα/πίνακα με τίτλο "Μετρήσεις Ω και ρυθμού εκτέλεσης" του σετ 3 των διαλέξεων και αριθμώντας ως γραμμή υπ' αριθμ. 1 την δεύτερη (dot product), να επιλέξετε τις συναρτήσεις που αντιστοιχούν στις γραμμές που αντιστοιχούν στο id . Στη συνέχεια, ό,τι ζητάται στο ερώτημα αφορά αποκλειστικά αυτό το υποσύνολο συναρτήσεων.

α) Να χρονομετρήσετε με αξιόπιστο τρόπο τις κλήσεις για n=200 και να υπολογίσετε τα Gflops/s κάνοντας την υπόθεση εργασίας ότι τα Ω είναι ίδια με αυτά που εμφανίζονται στον πίνακα.

β) Βρείτε στη βιβλιογραφία (και να αναφέρετε ακριβώς από πού το βρήκατε), αν υπάρχει, την υπολογιστική πολυπλοκότητα για την συγκεκριμένη πράξη (π.χ. για το DOT είναι 2n.) Σχολιάστε τα αποτελέσματα (αν γνωρίζετε τη θεωρητική τιμή του Ω για γενικό n, θα ήταν καλό να το αναφέρετε.)

Κατάταξη κατά αύξουσα σειρά χρόνου:

chol(A) → 0.000652 < x=A\y → 0.000864 < matrix multiply → 0.000930 < cond(A) → 0.008211

Κατάταξη κατά φθίνουσα σειρά Gflops/s:

matrix multiply → 17.21 > x=A\y → 7.21 > chol(A) → 4.91 > cond(A) → 3.29

Πολυπλοκότητες Πράξεων:

x = A\y

Για γενικές πυκνές μήτρες, αναμένεται χρόνος πολυπλοκότητας $O(n^3)$ για τις άμεσες μεθόδους, ενώ οι επαναληπτικές μέθοδοι μπορούν να είναι πιο γρήγορες ανάλογα με την αραιότητα και την κατάσταση της μήτρας.

chol(A) → $(1/3)*n^3 + O(n^2)$

matrix multiply

Naive Method: $O(n^3)$
Strassen's Algorithm: $O(n^{2.81})$
Coppersmith-Winograd Algorithm: $O(n^{2.376})$

cond(A) → $O(n^3)$

Για τον υπολογισμό του δείκτη κατάστασης απαιτείται ο υπολογισμός της L2 νόρμας ($O(n^2)$) και η αντιστροφή του μητρώου ($O(n^3)$).

Σχόλια:

- *Ερμηνεία με βάση τους χρόνους εκτέλεσης:*
  Η παραγοντοποίηση Cholesky είναι η πιο γρήγορη από όλες τις μεθόδους. Αυτή η μέθοδος είναι εξαιρετικά αποδοτική για θετικά ορισμένες μήτρες, κάτι που φαίνεται από τον χαμηλό χρόνο εκτέλεσης. Η επίλυση του συστήματος γραμμικών εξισώσεων είναι ελαφρώς πιο αργή από την παραγοντοποίηση, αλλά παραμένει σε ικανοποιητικό επίπεδο. Αυτό υποδηλώνει (προσωπική μου εκτίμηση) ότι η μήτρα A έχει καλές ιδιότητες. Ο υπολογισμός του γινομένου μητρώων , αν και είναι γρηγορότερος από τον υπολογισμό του δείκτη κατάστασης, είναι λίγο πιο αργός από τις προηγούμενες μεθόδους. Η απόδοση είναι επαρκής και δείχνει την αποδοτικότητα των αλγορίθμων για πολλαπλασιασμό μητρώων. Τέλος, ο υπολογισμός του δείκτη κατάστασης είναι ο πιο αργός και αυτό είναι αναμενόμενο, καθώς περιλαμβάνει την αντιστροφή της μήτρας και τον υπολογισμό της L2 νόρμας.

- *Ερμηνεία με βάση τα Gflops/s:*

Ο υπολογισμός του γινομένου μητρώων είναι ο πιο αποδοτικός, δείχνοντας την ικανότητα των αλγορίθμων να εκμεταλλεύονται την υπολογιστική ισχύ. Η λύση του συστήματος γραμμικών εξισώσεων είναι επίσης αποδοτική, με ικανοποιητικά Gflops/s, επιβεβαιώνοντας την αποτελεσματικότητα της μεθόδου. Παρά την υψηλή ταχύτητα χρόνου εκτέλεσης, τα Gflops/s της παραγοντοποίησης Cholesky είναι χαμηλότερα από άλλες μεθόδους. Αυτό μπορεί να οφείλεται στη φύση των υπολογισμών που γίνονται. Τέλος, ο υπολογισμός του δείκτη κατάστασης έχει την πιο χαμηλή απόδοση σε Gflops/s, γεγονός που δείχνει τη δυσκολία και τον χρόνο που απαιτείται για τις διαδικασίες που εμπλέκονται.

Βιβλιογραφικές Πηγές:

Asher-Greif ‘‘Εισαγωγή στις Αριθμητικές Μεθόδους’’ σελ.202 για Cholesky
και σελ.221 για δείκτη κατάστασης

https://en.wikipedia.org/wiki/Computational_complexity_of_mathematical_operations