ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΥΠΟΛΟΓΙΣΤΙΚΗ ΝΟΗΜΟΣΥΝΗ

(ΕΡΓΑΣΤΗΡΙΟ)

ΕΑΡΙΝΟ ΕΞΑΜΗΝΟ 2022-2023

2η Σειρά ασκήσεων

Αθανασίου Ελένη 19387004

Βαβαΐτη Κωνσταντίνα 18387257

Εργαστηριακή ομάδα

Ομάδα Γ

Αιγάλεω 04/04/2023

1. Σκοπός και περίληψη της άσκησης

Σκοπός αυτής της άσκησης είναι η μελέτη της μεθόδου βελτιστοποίησης ελαχίστων τετραγώνων μέσω δύο εφαρμογών. Στην πρώτη εφαρμογή βρίσκεται η βέλτιστη ευθεία που διαπερνά τα δεδομένα ενός ενισχυτή αφού πρώτα έχει γίνει η κατάλληλη επεξεργασία τους, σχεδιάζεται η γραφική παράσταση που απεικονίζει τις τιμές της ενίσχυσης(σε db) ως προς αυτές του λογαρίθμου της συχνότητας και τέλος, με την χρήση του προηγούμενου προγράμματος, υλοποιείται πρόγραμμα εύρεσης της αντίστοιχης συχνότητας μιας ενίσχυσης. Στην δεύτερη εφαρμογή κατασκευάζεται ένα γραμμικό μοντέλο που μπορεί να προβλέψει την σχετική απόδοση ενός επεξεργαστή σύμφωνα με τα χαρακτηριστικά του συστήματος, σχεδιάζεται η γραφική παράσταση που απεικονίζει τις τιμές της σχετικής απόδοσης και τις τιμές των προβλέψεων της σχετικής απόδοσης ως προς τον αριθμό του κάθε δεδομένου όπως και η γραφική παράσταση που απεικονίζει τις τιμές της σχετικής απόδοσης ως προς τις τιμές των προβλέψεων της σχετικής απόδοσης και τέλος, υλοποιείται τρισδιάστατη γραφική παράσταση που να απεικονίζει τις προβλέψεις του μοντέλου που κατασκευάστηκε προηγουμένως για την απόδοση ενός δεδομένου επεξεργαστή συναρτήσει της μέγιστης κεντρικής μνήμης και της μνήμης Cache.

2. Θεωρητικό Υπόβαθρο

✓ **Μέθοδος ελαχίστων τετραγώνων**

Τα ελάχιστα τετράγωνα είναι μια μέθοδος βελτιστοποίησης η οποία βρίσκει την καλύτερη ευθεία που διαπερνά τα πειραματικά δεδομένα, δηλαδή αυτή που απέχει την ελάχιστη απόσταση από αυτά.
Έστω ότι έχουμε ένα σύστημα με είσοδο x και έξοδο y:

Η εξίσωση που το περιγράφει είναι της μορφής: $y = a \cdot x + b$.
Για αυτό το σύστημα η υλοποίηση της μεθόδου των ελαχίστων τετραγώνων και γίνεται με τον υπολογισμό των παρακάτω εξισώσεων:

$$a = \frac{SS_{xy}}{SS_{xx}}, όπου\ SS_{xy} = \sum_{j=1}^{P} x_j y_j - \frac{\sum_{j=1}^{P} x_j \sum_{j=1}^{P} y_j}{P}, SS_{xx} = \sum_{j=1}^{P} (x_j)^2 - \frac{\left(\sum_{j=1}^{P} x_j\right)^2}{P}$$

και

$$b = \bar{y} - a\bar{x}, όπου\ \bar{y} = \frac{\sum_{j=1}^{P} y_j}{P}, \bar{x} = \frac{\sum_{j=1}^{P} x_j}{P}$$

<u>Μέσο απόλυτο σχετικό σφάλμα %</u>: $MARE\% = 100 \frac{\sum_{j=1}^{P} \frac{|y_j - \hat{y}_j|}{y_j}}{P}$

<u>Συντελεστής R^2</u>: $R^2 = 1 - \frac{SSE}{SST}, όπου\ SSE = \sum_{j=1}^{P} (y_j - \hat{y}_j)^2\ και\ SST = \sum_{j=1}^{P} (y_j - \bar{y})^2$

Με P συμβολίζεται το πλήθος των δεδομένων.

Σε μορφή πινάκων η εξίσωση που το περιγράφει είναι της μορφής: $A = (X' \cdot X)^{-1} \cdot X' \cdot Y$

Οι πίνακες X,Y και A θα πρέπει να έχουν την εξής μορφή:

$$\begin{array}{c} \xrightarrow{\text{Αριθμός μεταβλητής}} \\ \mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1N} & 1 \\ x_{21} & \cdots & x_{2N} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_{P1} & x_{P2} & x_{PN} & 1 \end{bmatrix} \end{array} \Bigg\downarrow \begin{array}{c}\text{Αριθμός} \\ \text{δεδομένου}\end{array} \quad \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{bmatrix} \Bigg\downarrow \begin{array}{c}\text{Αριθμός} \\ \text{δεδομένου}\end{array} \quad \mathbf{A} = \begin{bmatrix} a_1 \\ \vdots \\ a_N \\ b \end{bmatrix}$$

- ✓ **Δεδομένα 1ης Άσκησης-Συχνοτική απόκριση ενισχυτή**

Συλλέγονται δεδομένα ενίσχυσης-συχνότητας για έναν ενισχυτή. Τα δεδομένα αυτά βρίσκονται στο αρχείο ampldata.mat. Στην πρώτη στήλη του πίνακα dat βρίσκονται τα δεδομένα της συχνότητας (Hz) και στη δεύτερη στήλη βρίσκονται τα αντίστοιχα δεδομένα ενίσχυσης (dB). Για την περιοχή συχνοτήτων 1kHz-10MHz, η σχέση ανάμεσα στη συχνότητα και τη ενίσχυση δίνεται από μια εξίσωση της μορφής $y = a \cdot x + b$, όπου με y συμβολίζεται η ενίσχυση και με x ο δεκαδικός λογάριθμός της συχνότητας.

- ✓ **Δεδομένα 2ης Άσκησης-Πρόβλεψη απόδοσης CPU**

Στο αρχείο Machine_CPU.xlxs παρέχονται δεδομένα που αφορούν χαρακτηριστικά επεξεργαστών και την σχετική απόδοσή τους, όπως αυτή υπολογίζεται από benchmarks tests. Το γραμμικό μοντέλο που θα κατασκευαστεί θα είναι της μορφής:

$y = \left(\sum_{i=1}^{6} a_i x_i\right) + b$, όπου:

y: Η σχετική απόδοση του επεξεργαστή
x_1: Χρόνος κύκλου (ns)
x_2: Ελάχιστη κεντρική μνήμη (KB)
x_3: Μέγιστη κεντρική μνήμη (KB)
x_4: Μνήμη Cache (KB)
x_5: Ελάχιστος αριθμός καναλιών
x_6: Μέγιστος αριθμός καναλιών

Τα δεδομένα που δίνονται για τον υποθετικό επεξεργαστή που θα αξιοποιηθεί είναι:

Χρόνος κύκλου: 200ns
Ελάχιστη κεντρική μνήμη: 3000KB
Μέγιστη κεντρική μνήμη: 8000KB-16000KB
Μνήμη Cache: 32KB-128KB
Ελάχιστος αριθμός καναλιών: 6
Μέγιστος αριθμός καναλιών: 16

3. Πορεία Εργασίας

✓ 1η Άσκηση-Συχνότική απόκριση ενισχυτή

Κατασκευάζεται στο Matlab πρόγραμμα τύπου function το οποίο δέχεται σαν εισόδους τα δεδομένα για την ενίσχυση και την συχνότητα και επιστέφει σαν εξόδους:

1. Τους συντελεστές a και b
2. Τις προβλέψεις $\hat{y}$ για όλα τα διαθέσιμα δεδομένα συχνότητας σύμφωνα με το μοντέλο $\hat{y} = ax + b$
3. Το μέσο απόλυτο σχετικό σφάλμα % (MARE%) και το συντελεστή R^2

Οι συντελεστές το MARE% και το R^2 υπολογίζονται σύμφωνα με τους τύπους και τις μεθόδους που παρουσιάζονται στο θεωρητικό υπόβαθρο. Ο κώδικας αυτού του function παρουσιάζεται παρακάτω

```
function [a,b,yk, mare, R2] = fun(db, hz)

x = log10(hz);
p = length(x);

y = db;
SSxy = sum(x.*y)- (sum(x)*sum(y))/p;
SSxx = sum(x.*x)-(sum(x)^2)/p;

ybar = sum(y)/p;
xbar = sum(x)/p;

a = SSxy/SSxx;
b= ybar - a*xbar;

yk= a*x+ b;
mare = 100*(sum(abs(y-yk)./y)/p);

SSE = sum((y - yk).^2);
SST = sum((y-ybar).^2);

R2 = 1 - (SSE/SST);

end
```

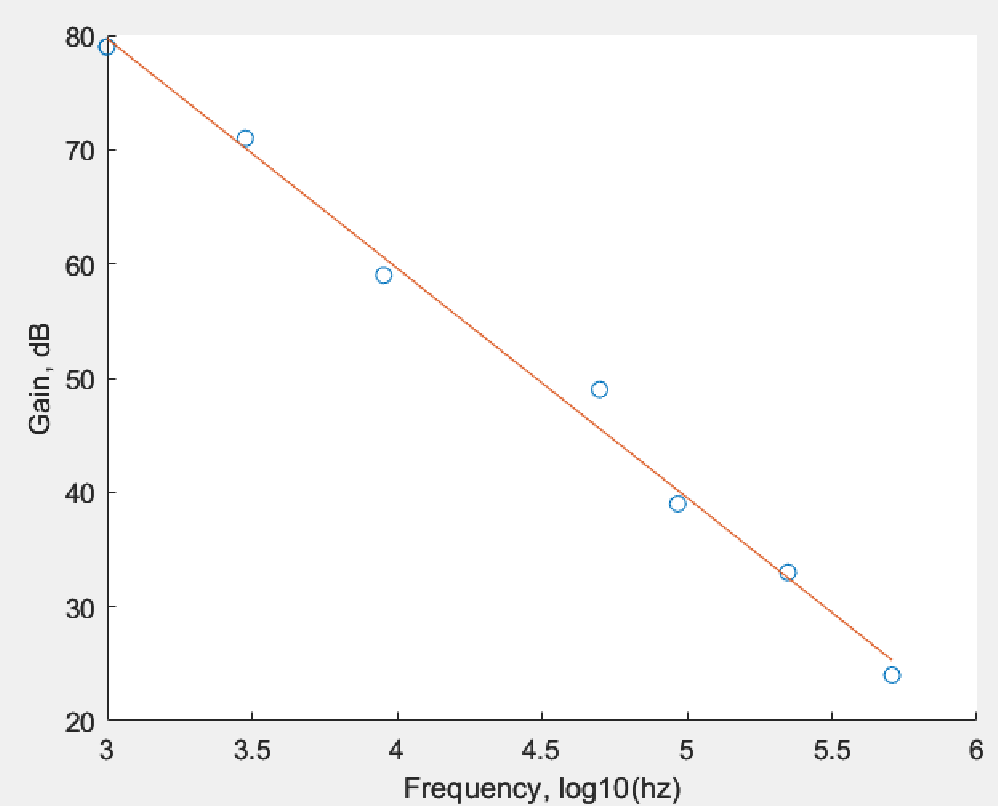
Φορτώνεται το αρχείο ampldata.mat στο Matlab και έπειτα εκτελείτε το παραπάνω function με είσοδο ενίσχυσης την δεύτερη στήλη του πίνακα dat, ο οποίος περιέχει τα δεδομένα αυτού του αρχείου, και είσοδο συχνότητας την πρώτη του στήλη. Τα αποτελέσματα που λήφθηκαν φαίνονται παρακάτω

```
>> load('ampldata.mat')
>> [a,b,yk, mare, R2] = fun(dat(:,2),dat(:,1))

a =

  -20.0973
```

```
b =

  140.0178


yk =

   79.7260
   70.1371
   60.5483
   45.5813
   40.1648
   32.5314
   25.3112


mare =

    3.0865


R2 =

    0.9924
```

Σχεδιάζεται γραφική παράσταση χρησιμοποιώντας σαν άξονες την ενίσχυση και το λογάριθμο της συχνότητας. Στη γραφική παράσταση απεικονίζονται τα δεδομένα ενίσχυσης-συχνότητας, καθώς και η βέλτιστη γραμμή που υπολογίστηκε με την μέθοδο των ελαχίστων τετραγώνων. Ο κώδικας υλοποίησης όπως και η γραφική παρουσιάζονται παρακάτω

```
>> scatter(log10(dat(:,1)),dat(:,2));
>> hold on
>> plot(log10(dat(:,1)),yk);
```



Εικόνα 1. Γραφική παράσταση ενίσχυσης ως προς το λογάριθμο της συχνότητας

Κατασκευάζεται πρόγραμμα τύπου function το οποίο δέχεται σαν εισόδους τους συντελεστές a και b που υπολογίστηκαν με την μέθοδο των ελαχίστων τετραγώνων καθώς και μια οποιαδήποτε τιμή ενίσχυσης (dB) και επιστρέφει σαν έξοδο την συχνότητα (Hz) στην οποία αντιστοιχεί αυτή η τιμή της ενίσχυσης. Ο κώδικας αυτού παρουσιάζεται παρακάτω

```
function hz = fun2(a,b,db)

hz = 10^((db - b)/a);

end
```

Εκτελείτε το παραπάνω πρόγραμμα για να υπολογιστεί για ποια συχνότητα η ενίσχυση είναι ίση με 0dB. Ο κώδικας υλοποίησης και τα αποτελέσματα παρουσιάζονται παρακάτω

```
>> hz = fun2(a,b,0)

hz =

   9.2684e+06
```

- **2η Άσκηση-Πρόβλεψη απόδοσης CPU**

Κατασκευάζεται πρόγραμμα τύπου function το οποίο δέχεται σαν εισόδους τα χαρακτηριστικά του επεξεργαστή και τη σχετική απόδοση και επιστέφει σαν εξόδους:

1. Τους συντελεστές $a_i, i = 1,2,\ldots,6$ και το σταθερό όρο b
2. Τις προβλέψεις $\hat{y}$ για όλα τα διαθέσιμα δεδομένα σύμφωνα με το μοντέλο $\hat{y} = \left(\sum_{i=1}^{6} a_i x_i\right) + b$
3. Το μέσο απόλυτο σχετικό σφάλμα % (MARA%) και τον συντελεστή R^2

Οι συντελεστές το MARE% και το R^2 υπολογίζονται σύμφωνα με τους τύπους και τις μεθόδους που παρουσιάζονται στο θεωρητικό υπόβαθρο. Ο κώδικας αυτού του function παρουσιάζεται παρακάτω

```
function [A,yk, mare, R2] = fun3(x,Y)

[r,c]=size(x);
x7 =  ones(r,1);

X = [x,x7];

A = inv(X'*X) * X'* Y;

yk = X*A;
p = length(X);
ybar = sum(Y)/p;

mare = 100*(sum(abs(Y-yk)./Y)/p);

SSE = sum((Y - yk).^2);
SST = sum((Y-ybar).^2);

R2 = 1 - (SSE/SST);

end
```

Φορτώνεται το αρχείο Machine_CPU.xlxs στο Matlab και έπειτα εκτελείτε το παραπάνω function με είσοδο τα χαρακτηριστικά των επεξεργαστών όπου είναι οι στήλες 4 έως 9 του πίνακα MachnineCPU και τις σχετικούς τους αποδόσεις όπου είναι η 10 στήλη του MachineCPU. Τα αποτελέσματα που λήφθηκαν φαίνονται παρακάτω

```
>> [A,yk, mare, R2] = fun3(MachineCPU(:,4:9),MachineCPU(:,10))

A =

    0.0489
    0.0153
    0.0056
    0.6414
   -0.2704
    1.4825
  -55.8939
```

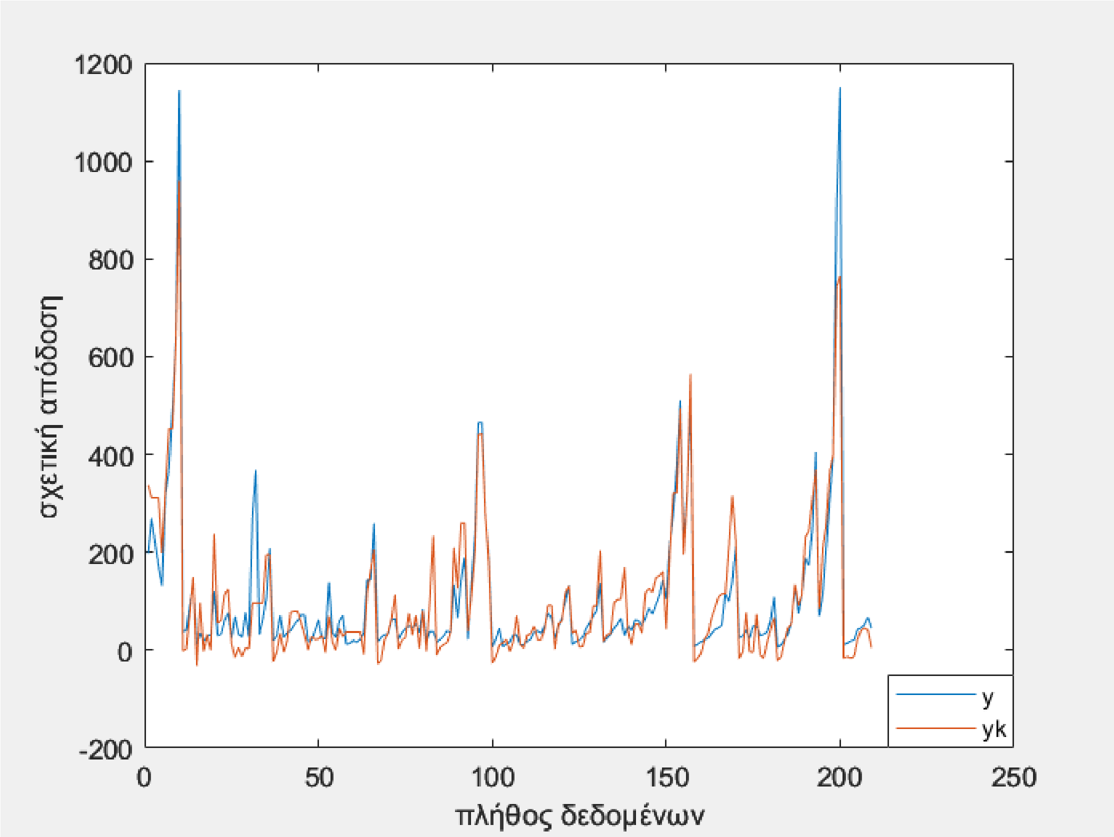
Οι τιμές του 209x1 διανύσματος yk παρουσιάζονται παρακάτω

337.1856	29.9601	178.4158	148.9076	207.7342
311.9490	-5.4257	-26.9108	151.3840	256.2650
311.9490	69.8200	-12.3990	160.5425	367.3739
311.9490	16.4353	9.8694	43.2799	397.9590
199.0872	0.0152	15.2455	199.6246	743.7266
332.3273	44.5864	21.9460	321.1515	764.2514
452.3584	27.6748	-2.4814	320.9072	-16.6308
452.3584	37.3355	17.7308	494.1409	-12.8076
630.6429	37.3355	70.6015	195.2761	-16.6308
959.4871	37.3355	11.6340	321.0538	-12.8076
-1.6506	37.3355	4.1712	563.9897	21.6172
2.1678	37.3355	30.9138	-23.8987	40.6014
75.4741	-8.8993	33.4794	-16.7303	45.5831
150.1418	120.0397	49.0299	-6.9555	40.4822
-31.7999	164.6108	21.1841	22.4448	5.1345
97.2068	205.6605	21.1841	33.5876	
-2.2922	-27.9971	41.8119	64.1728	
32.3257	-21.6928	91.7585	86.4583	
-0.1046	21.5822	91.5142	108.7439	
237.5943	34.5936	2.0431	115.3422	
55.2183	62.8259	53.5600	115.3422	
61.3396	113.6978	59.6330	209.2734	
114.0639	2.4816	116.5479	315.4191	
124.6858	21.1786	131.9415	221.4878	
12.6447	25.0270	34.7045	-16.9930	
-14.5451	75.1695	40.2759	-2.8853	
5.7073	30.8875	6.0775	77.2250	
-12.8886	70.6489	7.5033	-3.4579	
5.1340	3.7831	34.9200	-4.2690	
2.7016	81.0360	35.4608	72.9854	
96.3164	-2.3759	90.2943	-9.5370	
96.3164	95.5390	90.2943	-16.7731	
96.5806	233.8247	203.3733	15.1164	
96.5806	-10.6419	17.2094	42.5332	
195.3841	5.8901	32.6030	66.6057	
195.7750	11.9002	32.3326	-21.4593	
-23.8714	14.5948	96.8815	-12.9230	
-3.5998	38.3268	103.7672	18.2864	
33.9783	209.7193	103.7672	45.6091	
-4.1406	126.4100	169.1432	54.3277	
18.1450	260.1234	39.3201	134.1921	
78.0087	260.1234	11.0199	90.9217	
79.4912	38.3268	52.9063	111.7078	
79.4912	114.0122	52.9063	231.8560	
55.5042	203.2994	34.9471	243.5118	
33.2187	440.6452	116.9696	307.5263	
0.7042	442.8081	126.4992	369.0513	
28.0402	279.4178	118.3224	85.6118	
21.8482				
22.9986				

```
mare =

   76.4075


R2 =

    0.8649
```

Σχεδιάζεται γραφική παράσταση των πραγματικών τιμών y για όλα τα διαθέσιμα δεδομένα και των προβλέψεων $\hat{y}$ για κάθε δεδομένο ως προς τον αριθμό του κάθε δεδομένου. Ο κώδικας υλοποίησης όπως και η γραφική παρουσιάζονται παρακάτω

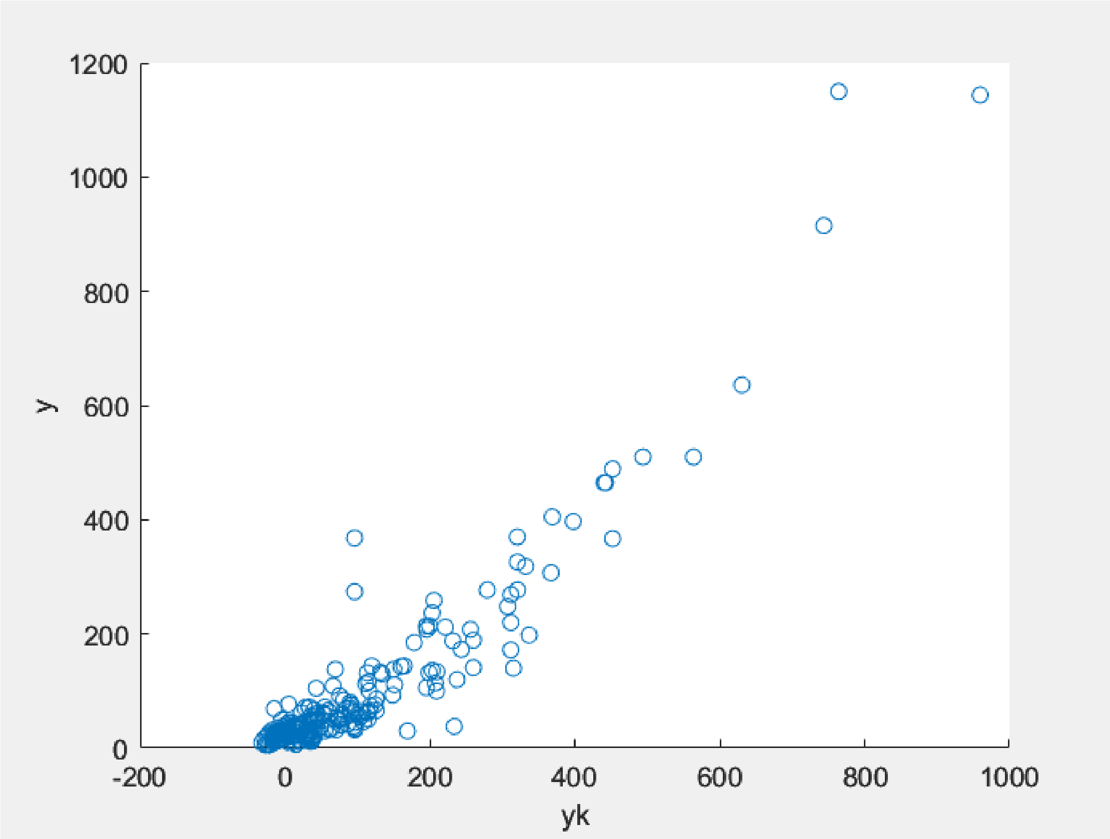
```
>> plot(MachineCPU(:,1),MachineCPU(:,10));
>> hold on
>> plot(MachineCPU(:,1),yk);
```



Εικόνα 2.Γραφική παράσταση σχετικής απόδοσης ως προς πλήθος δεδομένων

Σχεδιάζεται γραφική παράσταση των πραγματικών τιμών y συναρτήσει των προβλέψεων του μοντέλου $\hat{y}$. Ο κώδικας υλοποίησης όπως και η γραφική παρουσιάζονται παρακάτω

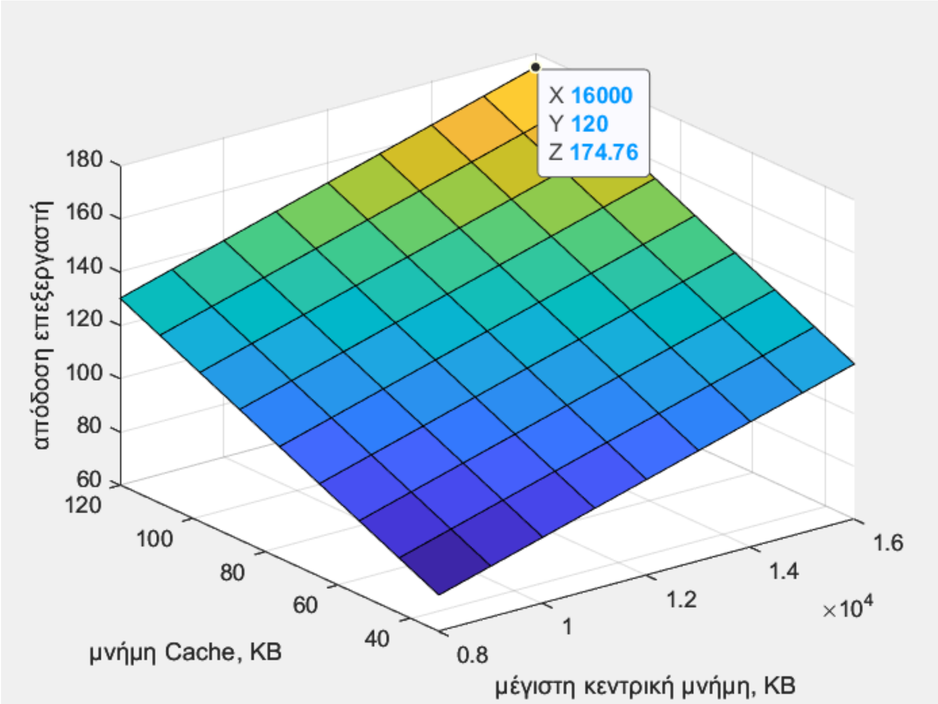
```
>> scatter(yk,MachineCPU(:,10))
```



Εικόνα 3. Γραφική παράσταση των πραγματικών τιμών y ως προς αυτών των προβλέψεων του μοντέλου yk

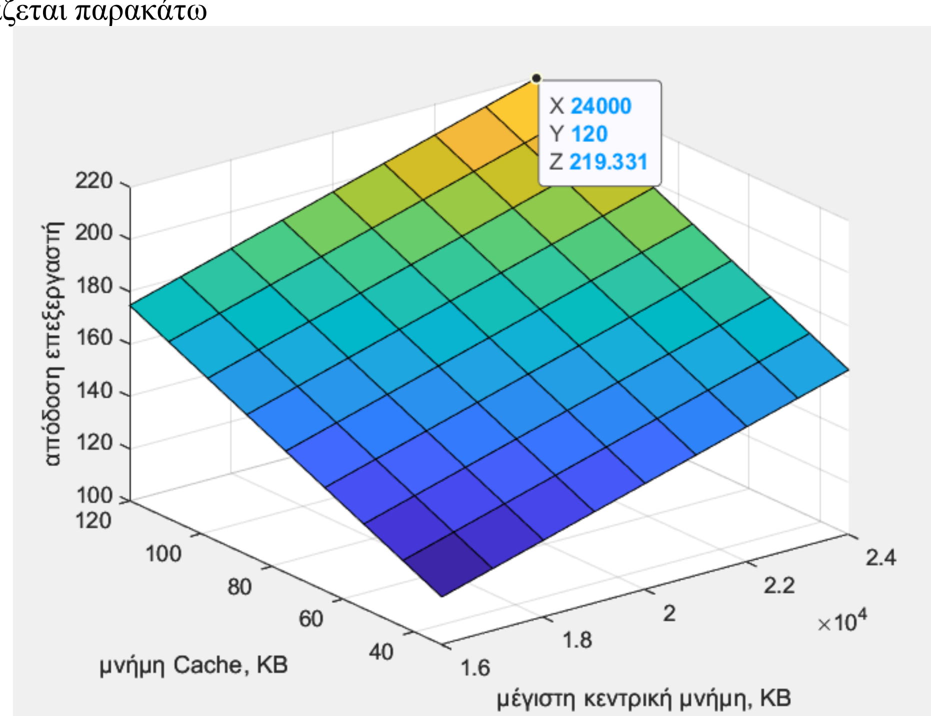
Κατασκευάζεται τρισδιάστατη γραφική παράσταση που απεικονίζει τις προβλέψεις του μοντέλου που κατασκευάστηκε νωρίτερα για την απόδοση ενός υποτιθέμενου επεξεργαστή, του οποίου οι προδιαγραφές αναφέρονται στο θεωρητικό υπόβαθρο, συναρτήσει της μέγιστης κεντρικής μνήμης και της μνήμης Cache. Ο κώδικας υλοποίησης όπως και η γραφική παρουσιάζονται παρακάτω

```
>> x3=[8000:1000:16000];
>> x4=[32:11:128];
>> [xx3,xx4]=meshgrid(x3,x4);
>> z=200*A(1,1)+3000*A(2,1)+xx3*A(3,1)+xx4*A(4,1)+6*A(6,1)+A(7,1);
>> surf(xx3,xx4,z)
```



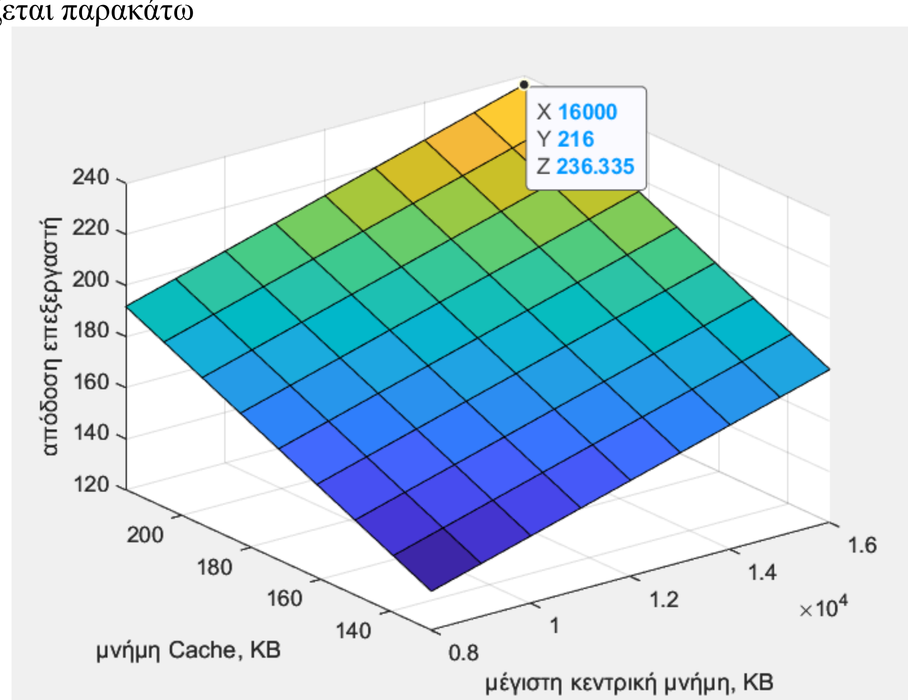
Εικόνα 4.Τριστιάστατη γραφική παράσταση της απόδοσης του επεξεργαστη, της μέγιστης κεντρικής μνήμης και της μνήμης Cache

Αν αυξηθεί η μέγιστη κεντρική μνήμη κατά 8000KB η γραφική που δημιουργείται παρουσιάζεται παρακάτω



Εικόνα 5.Τριστιάστατη γραφική παράσταση της απόδοσης του επεξεργαστη, της μέγιστης κεντρικής μνήμης και της μνήμης Cache αν αυξηθεί η μέγιστη κεντρική μνήμη

Αν αυξηθεί η μέγιστη η μνήμη Cache κατά 96KB η γραφική που δημιουργείται παρουσιάζεται παρακάτω



Εικόνα 6.Τριστιάστατη γραφική παράσταση της απόδοσης του επεξεργαστη, της μέγιστης κεντρικής μνήμης και της μνήμης Cache αν αυξηθεί η μνήμη Cache

Η διαδικασία αυτή εκτελέστηκε με σκοπό να κριθεί ποια από τις δύο μνήμες έχει μεγαλύτερη επίδραση στην απόδοση του επεξεργαστή.

4. Συμπεράσματα

Στην 1η Άσκηση φαίνεται ότι οι τιμές των προβλέψεων(yk) είναι πολύ κοντά στις πραγματικές τιμές (y). Αυτό δείχνει ότι το μοντέλο που χρησιμοποιήθηκε είναι αποτελεσματικό όπως επιβεβαιώνεται και από:

1. το χαμηλό ποσοστό % του mare
2. την υψηλή τιμή του R^2.
3. την γραφική παράσταση "Εικόνα 1" καθώς τα πειραματικά δεδομένα έχουν μικρή απόκλιση από την βέλτιστη γραμμή.

Στην 2η Άσκηση υπάρχει σχετικά μεγάλη απόκλιση μεταξύ των προβλέψεων και των πραγματικών τιμών και αυτό φαίνεται από

1. τις γραφικές παραστάσεις των εικόνων 2 και 3
2. το mare το οποίο είναι πολύ μεγάλο

Ωστόσο τα θετικά στο μοντέλο αυτό είναι

1. το R^2 το οποίο είναι πολύ κοντά στην μέγιστη τιμή του
2. υπάρχει συνέπεια ως προς τις δύο γραφικές της εικόνας 2

Έτσι εντοπίζεται μια ασυμφωνία ως προς την αξιολόγηση του μοντέλου μας με βάση αυτούς τους δύο δείκτες.

Με βάση τα συμπεράσματα που λήφθηκαν μέσω της διαδικασίας αξιολόγησης της επίδρασης της μέγιστης κεντρικής μνήμης και της μνήμης Cashe, κρίνεται πως είναι προτιμότερη η αύξηση της δεύτερης. Η επιλογή αύξησης αυτής της μνήμης προκαλεί μεγαλύτερη αύξηση της απόδοση του επεξεργαστή όπως φαίνεται και στις γραφικές παραστάσεις των εικόνων 5 και 6.

5. Βιβλιογραφία

- ✓ *Σημειώσεις του μαθήματος*