

ΛΥΣΕΙΣ ΕΞΕΤΑΣΗΣ ΤΕΛΕΙΟΦΟΙΤΩΝ

Πέμπτη 13 Ιουλίου 2006

Θέμα 1 (20%)

Σας δίνεται το ακόλουθο πρόβλημα γραμμικού προγραμματισμού:

$$\begin{array}{ll} \max & 3x_1 + 2x_2 \\ \text{subject to} & x_1 + 3x_2 \leq 5 \\ & 4x_1 + 2x_2 \leq 11 \\ & x_1, x_2 \geq 0 \end{array}$$

1. Απεικονίστε σε γραφική παράσταση τους περιορισμούς και (εάν το κρίνετε σκόπιμα) και την αντικειμενική συνάρτηση.

Μετατρέπουμε τον πρώτο περιορισμό σε ισότητα

$$x_1 + 3x_2 = 5$$

και βρίσκουμε δυο ζεύγη σημείων που να ικανοποιούν την ισότητα αυτή:

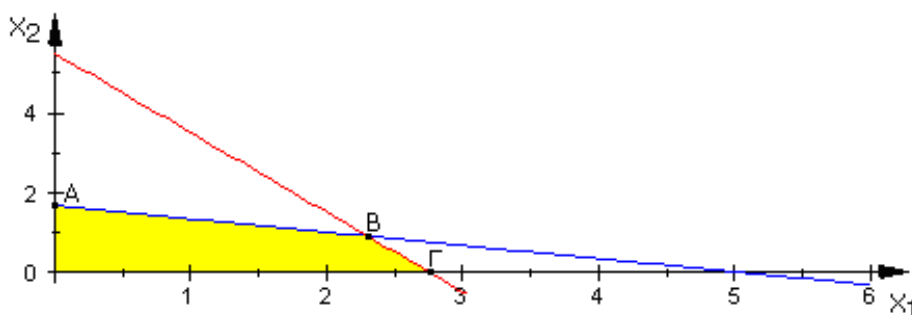
$$\begin{cases} x_2 = 0 \Rightarrow x_1 = 5 & \text{επομένως ένα σημείο είναι το } (5, 0) \\ x_1 = 0 \Rightarrow x_2 = \frac{5}{3} = 1.67 & \text{επομένως ένα δεύτερο σημείο είναι το } (0, 1.67) \end{cases}$$

Ενεργούμε ομοίως για τον δεύτερο περιορισμό:

$$4x_1 + 2x_2 = 11$$

$$\begin{cases} x_1 = 0 \Rightarrow x_2 = \frac{11}{2} = 5.5 & \text{επομένως ένα σημείο είναι το } (0, 5.5) \\ x_2 = 0 \Rightarrow x_1 = \frac{11}{4} = 2.75 & \text{επομένως ένα δεύτερο σημείο είναι το } (2.75, 0) \end{cases}$$

Απεικονίζουμε τώρα στην κατωτέρω γραφική παράσταση τους δυο αυτούς περιορισμούς εξετάζοντας μόνο το πρώτο τεταρτημόριο διότι οι τιμές του x_1 και x_2 πρέπει να είναι μη αρνητικές – προσέξτε ότι αντιστοιχούμε τη μεταβλητή x_1 στον άξονα x και την μεταβλητή x_2 στον άξονα y .



Η ισότητα που αντιστοιχεί στον πρώτο περιορισμό είναι η μπλε γραμμή ενώ εκείνη του δεύτερου περιορισμού είναι η κόκκινη γραμμή. Ο πρώτος περιορισμός επαληθεύεται κάτω από την μπλε γραμμή ενώ ο δεύτερος περιορισμός κάτω από την κόκκινη γραμμή. Επομένως, η εφικτή περιοχή (στην οποία επαληθεύονται και οι δυο περιορισμοί) είναι η κίτρινη περιοχή (δηλαδή ο κοινός τόπος των περιοχών στις οποίες επαληθεύονται και οι δυο περιορισμοί).

Όπως θα εξηγήσουμε στο επόμενο ερώτημα, δεν είναι απαραίτητο να χαράξουμε και την αντικειμενική συνάρτηση.

2. Βρείτε τη βέλτιστη τιμή της αντικειμενικής συνάρτησης.

Αναφερόμενοι στο γράφημα του προηγούμενου ερωτήματος, γνωρίζουμε από τη θεωρία ότι η βέλτιστη (δηλαδή η μέγιστη, εφόσον το πρόβλημα που εξετάζουμε είναι πρόβλημα μεγιστοποίησης) τιμή της αντικειμενικής συνάρτησης, οπωσδήποτε θα ευρίσκεται στο όριο της εφικτής περιοχής και μάλιστα θα ευρίσκεται σε κάποια από τις γωνίες της. Επομένως, δεδομένου ότι το πρόβλημα είναι πρόβλημα μεγιστοποίησης, η βέλτιστη τιμή θα ευρίσκεται είτε στο σημείο Α, είτε στο σημείο Β είτε στο σημείο Γ (αποκλείεται δηλαδή το σημείο τομής των αξόνων). Αρκεί λοιπόν να υπολογίσουμε την τιμή της αντικειμενικής συνάρτησης στα τρία αυτά σημεία και να επιλέξουμε τη μέγιστη εκ των τριών: αυτή θα είναι η βέλτιστη λύση του προβλήματος.

Στο σημείο Α προφανώς έχουμε $x_1 = 0$ επομένως από την ισότητα που αντιστοιχεί στον πρώτο περιορισμό έχουμε ήδη βρει ότι $x_2 = 1.67$. Με $x_1 = 0$ και $x_2 = 1.67$, η αντικειμενική συνάρτηση $Z = 3x_1 + 2x_2 = 3(0) + 2(1.67) = \mathbf{3.34}$.

Στο σημείο Γ προφανώς έχουμε $x_2 = 0$ επομένως από την ισότητα που αντιστοιχεί στον δεύτερο περιορισμό έχουμε ήδη βρει ότι $x_1 = 2.75$. Με $x_1 = 2.75$ και $x_2 = 0$, η αντικειμενική συνάρτηση $Z = 3x_1 + 2x_2 = 3(2.75) + 2(0) = \mathbf{8.25}$.

Αφήσαμε τελευταίο το σημείο Β διότι εκεί οι τιμές των x_1 και x_2 δεν μπορούν να προκύψουν “με το μάτι”. Σκεφτόμενοι ότι στο σημείο Β επαληθεύονται ως ισότητες και ο πρώτος και ο δεύτερος περιορισμός, αντιλαμβανόμαστε ότι για να βρούμε τις τιμές των x_1 και x_2 πρέπει να λύσουμε το ακόλουθο σύστημα (δύο εξισώσεων με δύο αγνώστους):

$$x_1 + 3x_2 = 5 \quad (1)$$

$$4x_1 + 2x_2 = 11 \quad (2)$$

$$(1) \Rightarrow x_1 = 5 - 3x_2 \quad (3)$$

$$(2), (3) \Rightarrow 4(5 - 3x_2) + 2x_2 = 11 \Rightarrow 20 - 12x_2 + 2x_2 = 11 \Rightarrow -12x_2 + 2x_2 = 11 - 20 \Rightarrow -10x_2 = -9 \Rightarrow 10x_2 = 9 \Rightarrow x_2 = 0.9 \quad (4)$$

$$(2), (4) \Rightarrow 4x_1 + 2(0.9) = 11 \Rightarrow 4x_1 + 1.8 = 11 \Rightarrow 4x_1 = 11 - 1.8 = 9.2 \Rightarrow x_1 = \frac{9.2}{4} = 2.3$$

Εναλλακτικά, μπορούμε να λύσουμε το σύστημα με ορίζουσες (πολύ γρήγορη λύση με μικρή πιθανότητα αριθμητικού λάθους λόγω της καλής οργάνωσης που επιβάλλουν οι ορίζουσες):

$$x_1 = \frac{\begin{vmatrix} 5 & 3 \\ 11 & 2 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix}} = \frac{5 \times 2 - 3 \times 11}{1 \times 2 - 3 \times 4} = \frac{10 - 33}{2 - 12} = \frac{23}{10} = 2.3$$

$$x_2 = \frac{\begin{vmatrix} 1 & 5 \\ 4 & 11 \end{vmatrix}}{\begin{vmatrix} 1 & 3 \\ 4 & 2 \end{vmatrix}} = \frac{1 \times 11 - 5 \times 4}{1 \times 2 - 3 \times 4} = \frac{11 - 20}{2 - 12} = \frac{9}{10} = 0.9$$

Με $x_1 = 2.3$ και $x_2 = 0.9$, η αντικειμενική συνάρτηση $Z = 3x_1 + 2x_2 = 3(2.3) + 2(0.9) = \mathbf{8.7}$.

Συγκρίνοντας τις τιμές της αντικειμενικής συνάρτησης στα σημεία Α (3.34), Β(8.7) και Γ (8.25), συμπεραίνουμε ότι η μέγιστη τιμή της αντικειμενικής συνάρτησης ευρίσκεται στο σημείο Β. Επομένως η βέλτιστη λύση του προβλήματος είναι

$x_1 = 2.3$ και $x_2 = 0.9$ με την αντικειμενική συνάρτηση $Z = 8.7$.

Αν και δεν απαιτείται στην εξέταση, στο σημείο αυτό λύνουμε (για εκπαιδευτικούς σκοπούς) το πρόβλημα με το πρόγραμμα LINGO εισάγοντας τον ακόλουθο κώδικα

```
! ANTIKEIMENIKH SYNARTHSH;  
max=3*x1+2*x2;  
! PERIORISMOI;  
x1+3*x2<=5;
```

$$4 \times x_1 + 2 \times x_2 \leq 11;$$

που μας δίνει την αναμενόμενη λύση

Global optimal solution found.

Objective value:

8.700000

Total solver iterations:

2

Variable	Value	Reduced Cost
X1	<u>2.300000</u>	0.000000
X2	<u>0.900000</u>	0.000000

Row	Slack or Surplus	Dual Price
1	8.700000	1.000000
2	0.000000	0.200000
3	0.000000	0.700000

όπου η βέλτιστη τιμή της αντικειμενικής συνάρτησης και οι αντίστοιχες τιμές των μεταβλητών απόφασης έχουν υπογραμμιστεί.

Θέμα 2 (30%)

Μια εταιρεία παραγωγής και πώλησης υπολογιστών (που ιδρύθηκε από αριστούχους αποφοίτους του τμήματος ΔΤΨΣ στο ΠΑΠΕΙ) λειτουργεί κέντρο συναρμολόγησης στην Αθήνα με ετήσια παραγωγή 30 000 μονάδων. Οι παραγόμενοι υπολογιστές μεταφέρονται σε περιφερειακά κέντρα διανομής που ευρίσκονται στη Ρώμη, το Παρίσι και τις Βρυξέλες.

Επειδή αναμένεται αύξηση στην ζήτηση υπολογιστών προς τη βόρεια Ευρώπη, η εταιρεία επιθυμεί να αυξήσει την ετήσια παραγωγή με την κατασκευή ενός ή περισσότερων νέων κέντρων συναρμολόγησης σε μια ή περισσότερες από τις εξής πόλεις: Βαρσοβία, Κοπεγχάγη, Άμστερνταμ και Βερολίνο. Το ετήσιο κόστος και η ετήσια παραγωγή για κάθε υποψήφια πόλη θα έχουν ως ακολούθως:

Πόλη	Ετήσιο κόστος (€)	Ετήσια παραγωγή (μονάδες)
Βαρσοβία	175 000	10 000
Κοπεγχάγη	300 000	20 000
Άμστερνταμ	375 000	30 000
Βερολίνο	500 000	40 000

Σημειωτέον ότι προβλέπονται οι ακόλουθες ετήσιες ζητήσεις σε κάθε ένα από τα τρία κέντρα διανομής:

Κέντρο διανομής	Ετήσια ζήτηση (μονάδες)
Ρώμη	30 000
Παρίσι	20 000
Βρυξέλες	20 000

Τα μεταφορικά κόστη (σε € ανά μονάδα) από κάθε υποψήφια πόλη (και το υφιστάμενο κέντρο συναρμολόγησης) προς κάθε κέντρο διανομής θα έχουν ως ακολούθως:

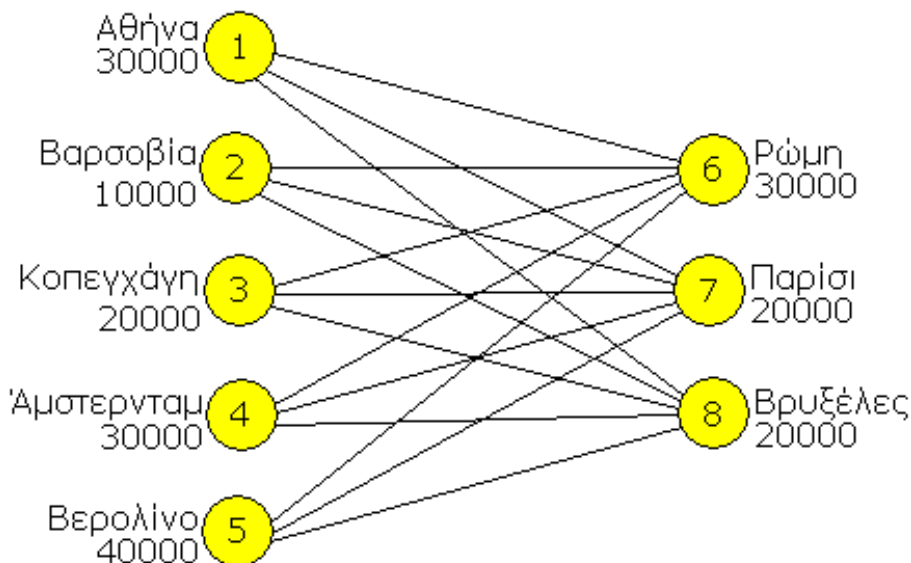
Πόλη	Κέντρο διανομής		
	Ρώμη	Παρίσι	Βρυξέλες
Βαρσοβία	5	2	3
Κοπεγχάγη	4	3	4
Άμστερνταμ	9	7	5
Βερολίνο	10	4	2
Αθήνα	8	4	3

Το πρόβλημα αυτό ανήκει στην κατηγορία των λεγομένων distribution design problems. Απαντήστε στις ακόλουθες ερωτήσεις:

1. Σχεδιάστε (στο κουτί της επόμενης σελίδας) την δικτυακή αναπαράσταση του προβλήματος (όπως θα κάνατε σε περίπτωση προβλήματος μεταφοράς).
2. Σας ζητείται (στο κουτί της επόμενης σελίδας) να ορίσετε το πρόβλημα ως πρόβλημα γραμμικού προγραμματισμού ώστε (με τη λύση του) αφενός μεν να επιλεγούν οι βέλτιστες τοποθεσίες για την εγκατάσταση των νέων κέντρων συναρμολόγησης (ενός ή περισσοτέρων) αφετέρου δε να υπολογιστούν οι βέλτιστες ποσότητες που θα μεταφέρονται ανά κλάδο. Δεν σας ζητείται να λύσετε το πρόβλημα, απλά να το ορίσετε μαθηματικά.

Οποσδήποτε να ορίζονται με σαφήνεια οι μεταβλητές απόφασης (**ΥΠΟΔΕΙΞΗ**: είναι δυαδικές), να φαίνεται καθαρά η αντικειμενική συνάρτηση (και το είδος του προβλήματος) και να είναι γραμμένοι με πληρότητα όλοι οι περιορισμοί.

Το πρόβλημα μπορεί να παρασταθεί δικτυακά σαν πρόβλημα μεταφοράς, όπου (α) κόμβοι προέλευσης είναι το υφιστάμενο κέντρο συναρμολόγησης (Αθήνα) και τα 4 υποψήφια κέντρα συναρμολόγησης (Βαρσοβία, Κοπεγχάγη, Άμστερνταμ και Βερολίνο) και (β) κόμβοι προορισμού είναι τα 3 κέντρα διανομής (Ρώμη, Παρίσι, Βρυξέλες).



Παρατηρούμε ότι το δίκτυο έχει 8 κόμβους, εκ των οποίων οι κόμβοι 1 έως και 5 είναι κόμβοι προέλευσης ενώ οι κόμβοι 6 έως και 8 είναι κόμβοι προορισμού (στους αριστερούς κόμβους, που είναι κόμβοι προέλευσης, σημειώνουμε την ετήσια παραγωγή, ενώ στους δεξιούς κόμβους, που είναι κόμβοι προορισμού, σημειώνουμε την ετήσια ζήτηση). Επίσης, το δίκτυο περιλαμβάνει $5 \times 3 = 15$ κλάδους (1-6, 1-7, 1-8, 2-6 ... 5-6, 5-7, 5-8) κάθε ένας εκ των οποίων χαρακτηρίζεται από το αντίστοιχο μοναδιαίο μεταφορικό κόστος (π.χ. ο κλάδος Βαρσοβία-Ρώμη, δηλαδή ο κλάδος 2-6, σύμφωνα με τον τελευταίο πίνακα της εκφώνησης του προβλήματος, χαρακτηρίζεται από κόστος 5 ευρώ ανά μονάδα μεταφερόμενου προϊόντος, δηλαδή υπολογιστή). Το μοναδιαίο αυτό κόστος μεταφοράς, για τον κλάδο $i-j$ το συμβολίζουμε c_{ij} .

Τονίζουμε ότι ενώ το διάγραμμα δικτύου είναι πανομοιότυπο με εκείνο ενός προβλήματος μεταφοράς, πρέπει να θυμόμαστε ότι ενώ στον κόμβο 1 (Αθήνα) υπάρχει σίγουρα κέντρο συναρμολόγησης, στους κόμβους 2 έως και 5 ενδέχεται να μην υπάρχουν παντού κέντρα συναρμολόγησης. Σε ποιους από τους κόμβους 2 έως και 5 συμφέρει να τοποθετηθεί ένα (ή περισσότερα) νέα κέντρα συναρμολόγησης, θα το δείξει η βέλτιστη λύση του προβλήματος που θα ορίσουμε ακολούθως.

Έχοντας ολοκληρώσει την απάντηση στο πρώτο ερώτημα, προχωρούμε τώρα στον ορισμό του προβλήματος. Καταρχήν θεωρούμε τις ακόλουθες μεταβλητές απόφασης:

d_i δυαδική μεταβλητή που ισούται (α) με ένα εάν στον κόμβο i υπάρχει νέο κέντρο συναρμολόγησης ή (β) με μηδέν εάν όχι

$i = 1, 2 \dots 5$

x_{ij} μεταφερόμενη ποσότητα υπολογιστών από τον κόμβο i στον κόμβο j

$i = 1, 2 \dots 5$ και $j = 6, 7, 8$.

Όσον αφορά τις μεταβλητές d_i , επειδή στον κόμβο 1 (Αθήνα) υπάρχει κέντρο συναρμολόγησης, η μεταβλητή d_1 οπωσδήποτε θα είναι ίση με ένα. Από την άλλη πλευρά, τουλάχιστον μια εκ των d_2, d_3, d_4 και d_5 θα πρέπει να ισούται με ένα (αλλιώς δεν μπορεί να ικανοποιηθεί η συνολική ζήτηση των κέντρων διανομής). Το ποιες εκ των d_2, d_3, d_4 και d_5 θα πρέπει να είναι ίσες με ένα ώστε η εταιρία να λειτουργεί με το ελάχιστο κόστος, θα μας το δώσει η βέλτιστη λύση του προβλήματος που θα ορίσουμε στο ερώτημα αυτό.

Με βάση τις ανωτέρω μεταβλητές απόφασης, γράφουμε την ακόλουθη αντικειμενική συνάρτηση που εκφράζει ελαχιστοποίηση του συνολικού κόστους που αποτελείται από το ετήσιο κόστος λειτουργίας (K_i για $i = 2, 3, 4$ και 5 δηλαδή για Βαρσοβία, Κοπεγχάγη, Άμστερνταμ και Βερολίνο αντίστοιχα) και του κόστους μεταφοράς (c_{ij} για όλους τους κλάδους $i-j$):

$$\begin{aligned} \min Z &= K_2 d_2 + K_3 d_3 + K_4 d_4 + K_5 d_5 + \{\text{ετήσιο κόστος παραγωγής για} \\ &\quad \text{Βαρσοβία, Κοπεγχάγη Άμστερνταμ} \\ &\quad \text{και Βερολίνο}\} \\ &\quad c_{16} x_{16} + c_{17} x_{17} + c_{18} x_{18} + \{\text{κόστος μεταφοράς από Αθήνα}\} \\ &\quad c_{26} x_{26} + c_{27} x_{27} + c_{28} x_{28} + \{\text{κόστος μεταφοράς από Βαρσοβία}\} \\ &\quad c_{36} x_{36} + c_{37} x_{37} + c_{38} x_{38} + \{\text{κόστος μεταφοράς από Κοπεγχάγη}\} \\ &\quad c_{46} x_{46} + c_{47} x_{47} + c_{48} x_{48} + \{\text{κόστος μεταφοράς από Άμστερνταμ}\} \\ &\quad c_{56} x_{56} + c_{57} x_{57} + c_{58} x_{58} \{\text{κόστος μεταφοράς από Βερολίνο}\} \\ &= 175000 d_2 + 300000 d_3 + 375000 d_4 + 500000 d_5 + \\ &\quad 8 x_{16} + 4 x_{17} + 3 x_{18} + \\ &\quad 5 x_{26} + 2 x_{27} + 3 x_{28} + \\ &\quad 4 x_{36} + 3 x_{37} + 4 x_{38} + \\ &\quad 9 x_{46} + 7 x_{47} + 5 x_{48} + \\ &\quad 10 x_{56} + 4 x_{57} + 2 x_{58} \end{aligned}$$

Σημειώνεται ότι μέσα σε {αγκύλες} γράφουμε σχόλια που δεν αποτελούν μέρος του ορισμού του προβλήματος. Παρατηρούμε ότι στην αντικειμενική συνάρτηση δεν έχει συμπεριληφθεί το ετήσιο κόστος παραγωγής του κέντρου συναρμολόγησης της Αθήνας το οποίο άλλωστε είναι άγνωστο – αυτό δεν μας πειράζει όμως γιατί το κόστος αυτό είναι δεδομένο αφού στην Αθήνα θα υπάρχει οπωσδήποτε κέντρο συναρμολόγησης.

Έχοντας γράψει την αντικειμενική συνάρτηση, προχωρούμε τώρα στο γράψιμο των περιορισμών. Όπως και σε ένα πρόβλημα μεταφοράς, ορίζουμε ότι η ποσότητα που φεύγει από κάθε κόμβο προορισμού δεν μπορεί να είναι μεγαλύτερη από την αντίστοιχη παραγωγή:

$$\begin{aligned} x_{16} + x_{17} + x_{18} &\leq 30000 d_1 && \{\text{Αθήνα, στην οποία υπάρχει οπωσδήποτε κέντρο} \\ &&& \text{συναρμολόγησης άρα το } d_1 \text{ πρέπει οπωσδήποτε να είναι} \\ &&& \text{ίσο με ένα – το βλέπουμε παρακάτω}\} \\ x_{26} + x_{27} + x_{28} &\leq 10000 d_2 && \{\text{Βαρσοβία}\} \\ x_{36} + x_{37} + x_{38} &\leq 20000 d_3 && \{\text{Κοπεγχάγη}\} \\ x_{46} + x_{47} + x_{48} &\leq 30000 d_4 && \{\text{Άμστερνταμ}\} \\ x_{56} + x_{57} + x_{58} &\leq 40000 d_5 && \{\text{Βερολίνο}\} \end{aligned}$$

Ομοίως, γράφουμε περιορισμούς που εξασφαλίζουν ότι ικανοποιείται η ζήτηση στα κέντρα διανομής:

$$\begin{aligned} x_{16} + x_{26} + x_{36} + x_{46} + x_{56} &\geq 30000 && \{\text{Ρώμη}\} \\ x_{17} + x_{27} + x_{37} + x_{47} + x_{57} &\geq 20000 && \{\text{Παρίσι}\} \\ x_{18} + x_{28} + x_{38} + x_{48} + x_{58} &\geq 20000 && \{\text{Βρυξέλες}\} \end{aligned}$$

Τέλος, όπως είχαμε πει στην αρχή, ορίζουμε ότι στον κόμβο 1 (Αθήνα) υπάρχει οπωσδήποτε κέντρο συναρμολόγησης:

$$d_1 = 1$$

Ο ορισμός του προβλήματος ολοκληρώνεται με τους περιορισμούς που ορίζουν τη φύση και μη αρνητικότητα των μεταβλητών απόφασης:

$$\begin{aligned} x_{ij} &\geq 0 \text{ για κάθε } i = 1, 2, 3, 4, 5 \text{ και } j = 6, 7, 8 \\ d_i &\geq 0 \text{ για κάθε } i = 1, 2, 3, 4, 5 \end{aligned}$$

x_{ij} ακέραιες
 d_i δυαδικές

Αν και δεν απαιτείται στην εξέταση, στο σημείο αυτό λύνουμε (για εκπαιδευτικούς σκοπούς) το πρόβλημα με το πρόγραμμα LINGO γράφοντας τον ακόλουθο κώδικα:

```
! ANTIKEIMENIKH SYNARTHSH ;

min=
175000*d2 + 300000*d3 + 375000*d4 + 500000*d5 +
8*x16 + 4*x17 + 3*x18+
5*x26 + 2*x27 + 3*x28+
4*x36 + 3*x37 + 4*x38+
9*x46 + 7*x47 + 5*x48+
10*x56 + 4*x57 + 2*x58 ;

! PERIORISMOI ;

x16 + x17 + x18 <= 30000*d1 ;
x26 + x27 + x28 <= 10000*d2 ;
x36 + x37 + x38 <= 20000*d3 ;
x46 + x47 + x48 <= 30000*d4 ;
x56 + x57 + x58 <= 40000*d5 ;

x16 + x26 + x36 + x46 + x56 >= 30000 ;
x17 + x27 + x37 + x47 + x57 >= 20000 ;
x18 + x28 + x38 + x48 + x58 >= 20000 ;

d1=1 ;

! METABLHTES APOFASHS ;

@GIN(x16); @GIN(x17); @GIN(x18);
@GIN(x26); @GIN(x27); @GIN(x28);
@GIN(x36); @GIN(x37); @GIN(x38);
@GIN(x46); @GIN(x47); @GIN(x48);
@GIN(x56); @GIN(x57); @GIN(x58);

@BIN(d1);
@BIN(d2);
@BIN(d3);
@BIN(d4);
@BIN(d5);
```

που μας δίνει την ακόλουθη βέλτιστη λύση ($Z = 86000$):

```
Global optimal solution found.
Objective value:                860000.0
Extended solver steps:          0
Total solver iterations:        201
```

Variable	Value	Reduced Cost
D2	0.000000	175000.0
D3	0.000000	300000.0
D4	0.000000	375000.0
D5	<u>1.000000</u>	500000.0
X16	<u>30000.00</u>	8.000000
X17	0.000000	4.000000
X18	0.000000	3.000000
X26	0.000000	5.000000
X27	0.000000	2.000000
X28	0.000000	3.000000
X36	0.000000	4.000000
X37	0.000000	3.000000

X38	0.000000	4.000000
X46	0.000000	9.000000
X47	0.000000	7.000000
X48	0.000000	5.000000
X56	0.000000	10.000000
X57	<u>20000.00</u>	4.000000
X58	<u>20000.00</u>	2.000000
D1	1.000000	0.000000

Row	Slack or Surplus	Dual Price
1	860000.0	-1.000000
2	0.000000	0.000000
3	0.000000	0.000000
4	0.000000	0.000000
5	0.000000	0.000000
6	0.000000	0.000000
7	0.000000	0.000000
8	0.000000	0.000000
9	0.000000	0.000000
10	0.000000	0.000000

σύμφωνα με την οποία, για να είναι η αντικειμενική συνάρτηση ελάχιστη, νέο εργοστάσιο συναρμολόγησης θα πρέπει να τοποθετηθεί μόνο στον κόμβο 5 (Βερολίνο, $d_5 = 1$ υπογραμμισμένο στην ανωτέρω λύση).

Από την λύση μπορείτε επίσης να δείτε ότι οι μεταφερόμενες ποσότητες (υπογραμμισμένες στην ανωτέρω λύση) θα προέρχονται από τον κόμβο 1 (Αθήνα, $x_{16} = 30000$) και τον κόμβο 5 (Βερολίνο, $x_{57} = 30000$ και $x_{58} = 30000$). Βλέπετε δηλαδή ότι για να πάρουμε τη βέλτιστη λύση, τοποθετήθηκε ένα μόνο κέντρο συναρμολόγησης (ώστε να ελαχιστοποιηθεί το ετήσιο κόστος) στον κόμβο 5 που έχει αρκετή παραγωγή ώστε να μπορεί να εξυπηρετήσει τους κόμβους 6 και 7.

Θέμα 3 (30%)

Διερευνάτε την επιρροή διαφόρων χαρακτηριστικών σπιτιών στην τιμή πώλησής του. Μοντέλο γραμμικής παλινδρόμησης δίνει τα ακόλουθα αποτελέσματα στο πρόγραμμα gretl (salepric=τιμή πώλησης σπιτιού, sqft=εμβαδόν σπιτιού, bedrms=αριθμός υπνοδωματίων, baths=αριθμός λουτρών, garage=αριθμός θέσεων στάθμευσης, age=ηλικία σπιτιού σε έτη).

Model 1: OLS estimates using the 224 observations 1-224
Dependent variable: salepric
Heteroskedasticity-robust standard errors, variant HC1

VARIABLE	COEFFICIENT	STDERROR	T STAT	P-VALUE
const	-745.438	114.430	-6.514	<0.00001 ***
sqft	0.235604	0.0296806	7.938	<0.00001 ***
bedrms	<u>-19.9784</u>	23.5758	-0.847	<u>0.39769</u>
baths	<u>11.0591</u>	20.0591	0.551	<u>0.58197</u>
garage	136.599	42.5104	3.213	0.00151 ***
age	9.76488	3.27117	2.985	0.00316 ***

Mean of dependent variable = 642.929
Standard deviation of dep. var. = 371.376
Sum of squared residuals = 3.84204e+006
Standard error of residuals = 132.755
Unadjusted R-squared = 0.875081
Adjusted R-squared = 0.872216
F-statistic (5, 218) = 72.913 (p-value < 0.00001)
Log-likelihood = -1409.83
Akaike information criterion (AIC) = 2831.65
Schwarz Bayesian criterion (BIC) = 2852.12
Hannan-Quinn criterion (HQC) = 2839.92

Excluding the constant, p-value was highest for variable 4 (baths)

1. Γράψτε την εξίσωση του μοντέλου και σχολιάστε το κατά πόσον εξηγεί τα δεδομένα.

Από την στήλη στην οποία εμφανίζονται οι τιμές των συντελεστών της πολλαπλής παλινδρόμησης (COEFFICIENT), αντιγράψουμε την ακόλουθη εξίσωση (στρογγυλεύοντας τις τιμές σε 4 σημαντικά ψηφία):

$$\text{salepric} = -745.4 + 0.2356 \text{ sqft} - 19.98 \text{ bedrms} \\ + 11.06 \text{ baths} + 136.6 \text{ garage} + 9.765 \text{ age}$$

Το στατιστικό μέγεθος που μας δείχνει πόσο ένα μοντέλο παλινδρόμησης εξηγεί την διασπορά των δεδομένων είναι ο λεγόμενος συντελεστής προσδιορισμού ή συντελεστής αποφασιστικότητας (coefficient of determination) δηλαδή το R^2 . Στην προκειμένη περίπτωση το R^2 (unadjusted) είναι 0.875 ή 87.5%, που αποτελεί μια ιδιαίτερα υψηλή τιμή (το R^2 κυμαίνεται από 0 έως 100%) και δείχνει μια πολύ καλή προσαρμογή του μοντέλου στα δεδομένα (λέμε ότι το μοντέλο “εξηγεί” το 87.5% της διασποράς των δεδομένων).

2. Σχολιάστε τα πρόσημα των συντελεστών των ανεξάρτητων μεταβλητών (σε συνάρτηση με τη στατιστική σημαντικότητα).

Η μεταβλητή **bedrms** εμφανίζεται στην εξίσωση της παλινδρόμησης με αρνητικό πρόσημο, γεγονός που θα σήμαινε ότι σπίτια με περισσότερες κρεβατοκάμαρες κοστίζουν λιγότερο. Αυτό δεν είναι λογικό αλλά ο συντελεστής της μεταβλητής αυτής (-19.9784) δεν είναι στατιστικά σημαντικός, όπως δείχνει η τιμή $p = 0.39769$ (στη δεξιά στήλη, P-VALUE) που είναι πολύ μεγαλύτερη από την οριακή τιμή 0.05 (κάτω από την οποία μια τιμή είναι στατιστικά σημαντική “σε επίπεδο εμπιστοσύνης 95%”). Επομένως, δεν μας προβληματίζει ιδιαίτερα, σκεφτόμαστε δε ότι ενδέχεται με ένα μεγαλύτερο ή καλύτερο δείγμα να μπορούσαμε να εκτιμούσαμε ένα καλύτερο μοντέλο παλινδρόμησης στο οποίο το πρόσημο της **bedrms** να είναι όπως το περιμέναμε.

Επί τη ευκαιρία, ούτε ο συντελεστής της μεταβλητής **baths** (11.0591) είναι στατιστικά σημαντικός ($p = 0.58197$, που είναι πολύ μεγαλύτερο του 0.05). Επομένως, σε μια προσπάθεια βελτίωσης του μοντέλου, θα αφαιρούσαμε τις δυο αυτές ανεξάρτητες μεταβλητές.

3. Με βάση τα ακόλουθα πρόσθετα στοιχεία, θεωρείτε ότι η πολυσυγγραμμικότητα είναι πρόβλημα; Αιτιολογείστε.

Variance Inflation Factors

Minimum possible value = 1.0

Values > 10.0 may indicate a collinearity problem

2)	sqft	<u>2.845</u>
3)	bedrms	<u>1.775</u>
4)	baths	<u>2.830</u>
5)	garage	<u>1.667</u>
6)	age	<u>1.186</u>

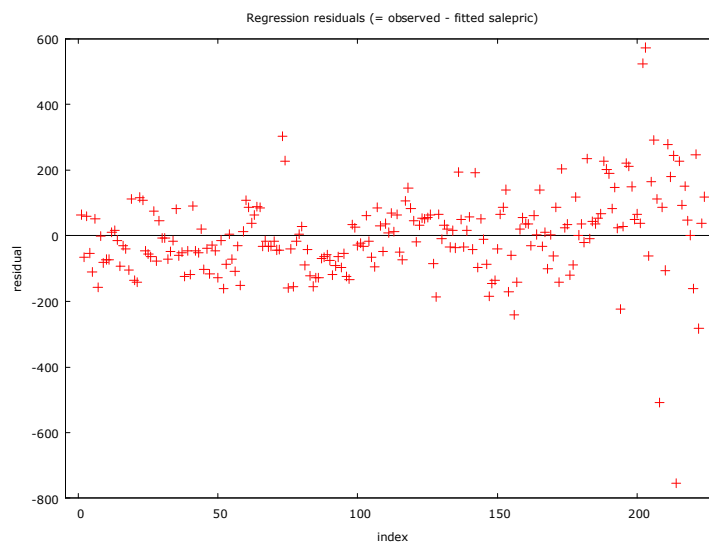
$VIF(j) = 1/(1 - R(j)^2)$, where $R(j)$ is the multiple correlation coefficient between variable j and the other independent variables

Η πολυσυγγραμμικότητα είναι ένα σοβαρό πρόβλημα στην πολλαπλή γραμμική παλινδρόμηση και πρέπει πάντα να εξετάζεται το κατά πόσον επηρεάζει σοβαρά τις τιμές των συντελεστών των ανεξάρτητων μεταβλητών. Αυτό γίνεται με τα VIFs (Variance Inflation Factors) για τα οποία έχουμε στο μάθημα παρουσιάσει τον εξής εμπειρικό κανόνα:

- Εάν κάποια VIF είναι από 1 έως 5, δεν έχουμε πρόβλημα πολυσυγγραμμικότητας.
- Εάν κάποια VIF είναι από 5 έως 10, έχουμε πρόβλημα αλλά δεν υπάρχει ανάγκη να αλλάξουμε το μοντέλο μας.
- Εάν κάποια VIF είναι πάνω από 10, έχουμε σοβαρό πρόβλημα πολυσυγγραμμικότητας και θα έπρεπε να ψάξουμε καλύτερο μοντέλο (π.χ. εξαιρώντας μια από τις ανεξάρτητες μεταβλητές που, λόγω υψηλής συσχέτισης με άλλη ανεξάρτητη μεταβλητή, δημιουργεί το πρόβλημα).

Βάσει των ανωτέρω, στο συγκεκριμένο μοντέλο δεν έχουμε πρόβλημα πολυσυγγραμμικότητας (όλες οι τιμές VIF, υπογραμμισμένες στην ανωτέρω εκτύπωση, είναι κάτω από 5).

4. Σχολιάστε το ακόλουθο διάγραμμα υπολοίπων.

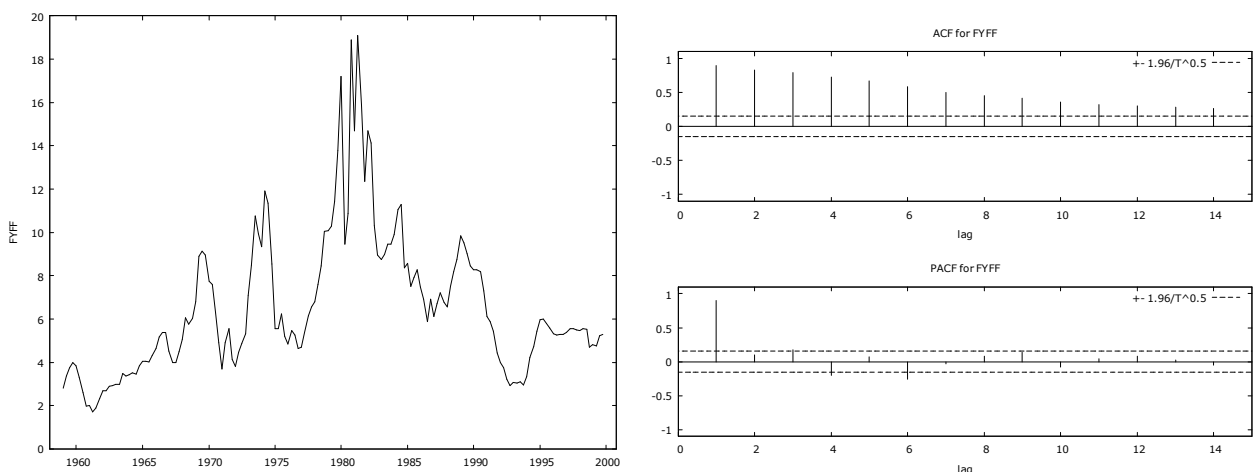


Εφόσον οι τιμές των υπολοίπων είναι διάσπαρτες γύρω από τον οριζόντιο άξονα του μηδενός σε μια ζώνη που πάνω-κάτω έχει το το ίδιο πλάτος, κρίνουμε ότι το διάγραμμα υπολοίπων είναι ικανοποιητικό και δεν μας δείχνει κάποιο σοβαρό πρόβλημα του μοντέλου.

Ανησυχούμε λιγάκι για το γεγονός ότι τα υπόλοιπα εμφανίζονται να ανοίγουν λίγο (σαν βεντάλια) στο δεξιό άκρο της γραφικής παράστασης αλλά εφόσον ο οριζόντιος άξονας δείχνει απλά τη σειρά των παρατηρήσεων (index) δεν μπορούμε να κρίνουμε κατά πόσον αυτός ο ετεροσκεδασμός είναι άξιος περαιτέρω μελέτης.

Θέμα 4 (20%)

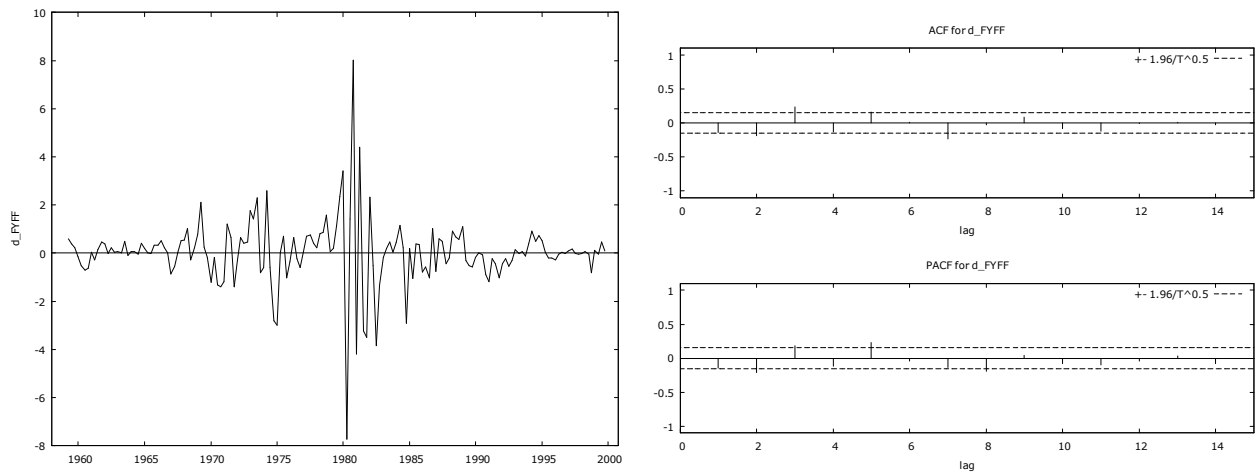
Σας δίνεται η κατωτέρω χρονοσειρά του ομοσπονδιακού επιτοκίου των ΗΠΑ μαζί με διαγράμματα των συναρτήσεων αυτοσυσχέτισης και μερικής αυτοσυσχέτισης.



1. Σχολιάστε το θέμα της στασιμότητάς της.

ΒΛΕΠΕ ΔΙΑΦΑΝΕΙΕΣ **ΣΤΑΣΙΜΟΤΗΤΑΣ** ΚΑΙ **ARIMA**
ΚΑΙ ΣΧΕΤΙΚΗ ΗΧΟΓΡΑΦΗΣΗ mp3 ΤΟΥ ΔΙΔΑΣΚΟΝΤΑ.

2. Κατωτέρω σας δίνονται οι πρώτες διαφορές της.



Ομοίως σχολιάστε το κατά πόσον είναι στάσιμες.

ΒΛΕΠΕ ΔΙΑΦΑΝΕΙΕΣ **ΣΤΑΣΙΜΟΤΗΤΑΣ** ΚΑΙ **ARIMA**
ΚΑΙ ΣΧΕΤΙΚΗ ΗΧΟΓΡΑΦΗΣΗ mp3 ΤΟΥ ΔΙΔΑΣΚΟΝΤΑ.