

Περιεχόμενα

Πρόλογος.....	3
Περίληψη.....	4
Κεφάλαιο 1- Εισαγωγή.....	5
Κεφάλαιο 2- Θεωρητικό Μέρος.....	9
2.1 Εισαγωγή στην Επιχειρησιακή Έρευνα.....	11
2.2 Μαθηματικός Προγραμματισμός (Mathematical Programming)	13
2.3 Ακέραιος Προγραμματισμός (Integer Programming).....	16
2.4 Πολυκριτηριακός Προγραμματισμός (Multiple Objective /Multiobjective Programming).....	19
2.5 Ανάλυση Παλινδρόμησης (Regression Analysis).....	25
2.6 Απλή Γραμμική Παλινδρόμηση (Simple Linear Regression).....	31
2.7 Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression).....	35
2.8 Επιλογή Μεταβλητών (Variable Selection)	41
2.9 Πολυκριτηριακή Παλινδρόμηση (Multiple Criteria Regression).....	50
Κεφάλαιο 3 - Μελέτη περίπτωσης και σχεδιασμός υπολογιστικού πειράματος.....	54
3.1 Μελέτη περίπτωσης.....	55
3.2 Σχεδιασμός υπολογιστικού πειράματος.....	56
3.3 Εφαρμογή μεθόδων εντός δείγματος.....	57
3.4 Εφαρμογή μεθόδων εκτός δείγματος.....	58
Κεφάλαιο 4 - Ανάπτυξη του μοντέλου.....	59
4.1 Το απλό μοντέλο παλινδρόμησης με ελαχιστοποίηση του SAE	60
4.2 Το απλό μοντέλο παλινδρόμησης με ελαχιστοποίηση του MAE.....	61
4.3 Το μοντέλο παλινδρόμησης με ελαχιστοποίηση του SAE και επιλογή μεταβλητών (variable selection).....	62
4.4 Το πολυκριτηριακό μοντέλο παλινδρόμησης και variable selection.....	62
4.5 Η γλώσσα GAMS	63

Κεφάλαιο 5 - Συζήτηση αποτελεσμάτων	67
5.1 Ανάλυση εντός Δείγματος: Η βελτίωση της προσαρμογής ως συνάρτηση του αριθμού μεταβλητών.....	68
5.2 Ανάλυση εντός Δείγματος: Καθορισμός βέλτιστου αριθμού μεταβλητών με βάση τα πληροφοριακά κριτήρια	69
5.3 Ανάλυση εκτός Δείγματος-Καθορισμός βέλτιστου αριθμού μεταβλητών.....	77
5.4 Σύγκριση μονοκριτηριακού μοντέλου με το πολυκριτηριακό μοντέλο. Ποιο είναι καλύτερο;	79
5.5 Ποιο είναι καλύτερο κριτήριο βελτιστοποίησης η ελαχιστοποίηση του SAE ή του MAE;.....	82
5.6 Ποια είναι καλύτερη προσέγγιση για την επιλογή μεταβλητών: Η βηματική ανάλυση παλινδρόμησης (stepwise regression) ή η ανάλυση παλινδρόμησης με Μαθηματικό Προγραμματισμό για το καθορισμό του βέλτιστου αριθμού μεταβλητών.	83
Κεφάλαιο 6 - Συμπεράσματα	87
6.1 Εντός Δείγματος.....	88
6.1.1 Καθορισμός βέλτιστου αριθμού μεταβλητών με βάση τα πληροφοριακά κριτήρια.....	88
6.2 Εκτός Δείγματος.....	88
6.2.1 Καθορισμός βέλτιστου αριθμού μεταβλητών	88
6.2.2 Σύγκριση μονοκριτηριακού μοντέλου με το πολυκριτηριακό μοντέλο. Ποιο είναι καλύτερο;	89
6.2.3 Ποιο μοντέλο δίνει καλύτερα αποτελέσματα, όσον αφορά το SAE ή το MAE;	89
6.2.4 Ποια είναι καλύτερη προσέγγιση για την επιλογή μεταβλητών: Η βηματική ανάλυση παλινδρόμησης (stepwise regression) ή η ανάλυση παλινδρόμησης με Μαθηματικό Προγραμματισμό για το καθορισμό του βέλτιστου αριθμού μεταβλητών.	90
Βιβλιογραφία.....	91
Παράρτημα Ι.....	93
Παράρτημα ΙΙ.....	99
Παράρτημα ΙΙΙ.....	110

Πρόλογος

Αντικείμενο της παρούσας διπλωματικής εργασίας είναι η πολυκριτηριακή ανάλυση παλινδρόμησης και η επιλογή μεταβλητών κατά τη διαδικασία αυτή. Ουσιαστικά, θα μελετήσουμε τη συμπεριφορά της πολυκριτηριακής παλινδρόμησης χρησιμοποιώντας ένα συγκεκριμένο δείγμα 250 παρατηρήσεων όπου ανεξάρτητες μεταβλητές είναι διάφορες σωματομετρικές μετρήσεις κι εξαρτημένη μεταβλητή το ποσοστό λίπους. Επιπλέον θα επιχειρήσουμε να ελαχιστοποιήσουμε τον αριθμό των μεταβλητών που χρειάζονται για τη δημιουργία ενός προβλεπτικού μοντέλου για το συγκεκριμένο δείγμα, το οποίο μελετάμε. Η δημιουργία προβλεπτικών μοντέλων είναι ένας τομέας, ο οποίος απασχολεί πάσης φύσεως τομέα της επιστήμης, καθώς δίνει τη δυνατότητα προσέγγισης μελλοντικών αποτελεσμάτων. Συνεπώς, μας δίνουν τη δυνατότητα πρόβλεψης μεταβλητών γρήγορα και αποτελεσματικά.

Η διπλωματική εργασία πραγματοποιήθηκε στο υπολογιστικό εργαστήριο Βιομηχανικής & Ενεργειακής Οικονομίας. Ο επιβλέπων καθηγητής αυτής της διπλωματικής εργασίας ήταν ο επίκουρος καθηγητής Γ. Μαυρωτάς.

Περίληψη

Σκοπός αυτής της διπλωματικής είναι να χρησιμοποιήσουμε τον μαθηματικό προγραμματισμό στην ανάλυση παλινδρόμησης. Για το λόγο αυτό χρησιμοποιούμε περισσότερα από ένα κριτήρια βελτιστοποίησης και επιχειρούμε να πραγματοποιήσουμε συγχρόνως κι επιλογή μεταβλητών δηλαδή να επιλέξουμε από τις ανεξάρτητες μεταβλητές εκείνες που είναι απαραίτητες. Για τον σκοπό της διπλωματικής βασιστήκαμε σε ένα σετ δεδομένων το οποίο το βρήκαμε στη βιβλιογραφία. Στη συνέχεια από το σετ αυτό, δημιουργήσαμε 10 διαφορετικά σετ δεδομένων με λιγότερες παρατηρήσεις με τη μέθοδο bootstrapping. Δηλαδή κρατήσαμε το 10% των παρατηρήσεων εκτός δείγματος (out of sample) δηλαδή δε το λάβαμε υπ' όψιν μας στη κατασκευή του μοντέλου πρόβλεψης, ώστε να εφαρμόσουμε το μοντέλο που προέκυπτε και να δούμε κατά πόσο προσαρμόζεται αυτό στις παρατηρήσεις που έμειναν «εκτός δείγματος». Τις υπόλοιπες παρατηρήσεις (90%), τις χρησιμοποιήσαμε για να κατασκευάσουμε αυτό το μοντέλο. Με τον τρόπο αυτό κατασκευάστηκαν μοντέλα παλινδρόμησης με διαφορετικά χαρακτηριστικά για να μπορούμε στη συνέχεια να τα συγκρίνουμε μεταξύ τους. Για τη δημιουργία κι επίλυση των μοντέλων χρησιμοποιήσαμε τη γλώσσα μοντελοποίησης GAMS. Τα μοντέλα αυτά είχαν ως αντικειμενικές συναρτήσεις (1) την ελαχιστοποίηση του αθροίσματος των απόλυτων σφαλμάτων και (2) την ελαχιστοποίηση του μέγιστου σφάλματος. Συγχρόνως μελετήσαμε και την ταυτόχρονη εφαρμογή των δύο κριτηρίων δημιουργώντας ένα μοντέλο πολυκριτηριακής ανάλυσης παλινδρόμησης.

Επιχειρήσαμε επίσης να μελετήσουμε το πρόβλημα της επιλογής των κατάλληλων μεταβλητών κάνοντας συγκρίσεις και με άλλες καθιερωμένες μεθόδους όπως η βηματική παλινδρόμηση (stepwise regression).

Από την μελέτη των αποτελεσμάτων στην προσαρμογή των μοντέλων στα δεδομένα εκτός δείγματος βγήκαν χρήσιμα συμπεράσματα. Προέκυψε ότι ο μαθηματικός προγραμματισμός μπορεί πολύ ικανοποιητικά να χρησιμοποιηθεί στην ανάλυση παλινδρόμησης. Τα πολυκριτηριακά μοντέλα έδωσαν καλύτερα αποτελέσματα από τα μονοκριτηριακά. Το κριτήριο της ελαχιστοποίησης του απόλυτου μέγιστου σφάλματος έδωσε καλύτερα αποτελέσματα από το κριτήριο της ελαχιστοποίησης του αθροίσματος των απόλυτων σφαλμάτων. Η επιλογή μεταβλητών με τον μαθηματικό προγραμματισμό δεν δίνει τα ίδια αποτελέσματα με την βηματική παλινδρόμηση.

Κεφάλαιο 1

Εισαγωγή

Τα τελευταία χρόνια έχει αρχίσει να απασχολεί ιδιαίτερα η διαχείριση της αβεβαιότητας ως προς τις παραμέτρους ενός μοντέλου ξεφεύγοντας από τις απλές μορφές ανάλυσης ευαισθησίας που μπορεί να προσφέρει και ο Μαθηματικός Προγραμματισμός. Στη στατιστική, η ανάλυση παλινδρόμησης αποτελείται από τεχνικές για την μοντελοποίηση των σχέσεων μεταξύ της εξαρτημένης μεταβλητής (ονομάζεται επίσης μεταβλητή απόκρισης) και μια ή περισσότερες ανεξάρτητες μεταβλητές (επίσης γνωστή ως επεξηγηματικές μεταβλητές ή μεταβλητές πρόβλεψης). Στην παλινδρόμηση, η εξαρτημένη μεταβλητή διαμορφώνεται ως συνάρτηση των ανεξάρτητων μεταβλητών στις οποίες αντιστοιχούν οι παράμετροι παλινδρόμησης (συντελεστές). Θεωρούμε ότι η ανάλυση παλινδρόμησης είναι μια τεχνική, η οποία εξετάζει την αλληλοεξάρτηση μεταξύ των μεταβλητών του συγκεκριμένου συνόλου δεδομένων. Η συνηθισμένη πολυκριτηριακή γραμμική παλινδρόμηση εφαρμόζεται σε εάν σύνολο δεδομένων όπου χρησιμοποιούμε ένα σταθερό αριθμό επεξηγηματικών (ή αλλιώς ανεξάρτητων) μεταβλητών και ένα κριτήριο βελτιστοποίησης (συνήθως χρησιμοποιούμε τη συνάρτηση ελαχιστοποίησης των σφαλμάτων όπως το άθροισμα των τετραγώνων των σφαλμάτων). Μερικές φορές, σκόπιμα χωρίζουμε ένα σύνολο N των υπάρχουσων επεξηγηματικών (ή αλλιώς ανεξάρτητων) μεταβλητών και να τις απομονώσουμε από αυτές το καλύτερο υποσύνολο M των μεταβλητών ($M \subset N$). Αυτό το πρόβλημα το συναντάμε συχνά στη στατιστική βιβλιογραφία σαν ένα πρόβλημα **επιλογής μεταβλητών** και η αιτία πίσω από αυτό είναι η αρχή της μετριοπάθειας, η οποία ορίζει ότι χρησιμοποιούμε τον ελάχιστο αριθμό επεξηγηματικών μεταβλητών, οι οποίες περιγράφουν επαρκώς τα δεδομένα μας.

Η πολυκριτηριακή ανάλυση αποφάσεων αποτελεί ένα εξελιγμένο πεδίο της επιχειρησιακής έρευνας, ο οποίος τις τελευταίες τρεις δεκαετίες έχει γνωρίσει ιδιαίτερη άνθηση τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Βασικό ρόλο στην ανάπτυξη και διάδοση της πολυκριτηριακής ανάλυσης αποτέλεσε η απλή διαπίστωση ότι η επίλυση πολύπλοκων και ιδιαίτερα σημαντικών προβλημάτων λήψης αποφάσεων δεν είναι δυνατό να πραγματοποιείται μέσω μιας μονόπλευρης και μονοδιάστατης ανάλυσης.

Το πρόβλημα της επιλογής μεταβλητών έχει πολλές πρακτικές εφαρμογές σε διάφορους τομείς, όπως κοινωνικές επιστήμες, οικονομία, μετεωρολογικά φαινόμενα και πολλές άλλες. Ανά διαστήματα έχουν δημοσιευτεί αρκετά άρθρα σχετικά με την Ανάλυση Παλινδρόμησης αλλά και για το πρόβλημα της Επιλογής Μεταβλητών. Μια σημαντική προσπάθεια έγινε από την ομάδα Narula et Al. το 2003, οι οποίοι παρουσίασαν μια σειρά από διαδικασίες οι οποίες θα παρουσιάζουν τη δυνατότητα να βρίσκουν μοντέλα χρησιμοποιώντας όσο λιγότερες μεταβλητές είναι δυνατόν και κάποια κριτήρια ώστε να είναι η δυνατή η επιλογή του μοντέλου. Ένα άλλο σημαντικό άρθρο είναι του Bruce Ratner το οποίο δημοσιεύθηκε το 2003 ο οποίος μελέτησε τις μεθόδους επιλογής μεταβλητών στην Παλινδρόμηση. Παρ' όλα αυτά η ανάλυση Παλινδρόμησης έχει χρησιμοποιηθεί και για τη

μελέτη οικολογικών προβλημάτων όπως έκαναν οι Mark J. Whittingham, et Al το 2006. Ένα άλλο ενδιαφέρον άρθρο δημοσιεύθηκε από τον Mark W. Watson το 2004, στο οποίο εξηγούσε πως χρησιμοποιούν τα πληροφοριακά κριτήρια για την πρόβλεψη του καιρού, κατασκευάζοντας μοντέλα πρόβλεψης. Κατά την αναζήτηση μας βιβλιογραφικών πηγών σχετικά με την ανάλυση παλινδρόμησης βρήκαμε το βιβλίο με τίτλο “Applied Regression Analysis” των Drapper και Smith. Στο βιβλίο αυτό εφαρμόζεται η μέθοδο παλινδρόμησης σε κάθε περίπτωση συνάμα, παρουσιάζονται και όλες οι μορφές που μπορεί να πάρει η μέθοδος παλινδρόμησης. Ουσιαστικά έχει γίνει μια προσπάθεια συγκέντρωσης όλων των μορφών Παλινδρόμησης. Ένα άλλο ενδιαφέρον το οποίο βρήκαμε κατά την αναζήτηση μας για τις βιβλιογραφικές μας πηγές, είναι <<Introduction to Linear Regression Analysis>> των Montgomer et Al. Το βιβλίο αυτό επικεντρώνεται στην ανάλυση της γραμμικής παλινδρόμησης, αναλύει κάθε περίπτωση ξεχωριστά. Ο στόχος του βιβλίου είναι να κάνει μια εισαγωγή σε όλες τις μορφές της ανάλυσης παλινδρόμησης και τα εφαρμόζει σε διάφορα δείγματα.

Στη συγκεκριμένη διπλωματική χρησιμοποιούμε ως μελέτη περίπτωσης ένα σύνολο δεδομένων από τη βιβλιογραφία. Τα δεδομένα αυτά προέρχονται από μετρήσεις που έγιναν σε ένα πληθυσμό με σκοπό τη μέτρηση του ποσοστού του λίπους τους. Το μέγεθος του δείγματος ανέρχεται στις 250 παρατηρήσεις όπου μετράμε το ποσοστό λίπους σε διάφορα σημεία του σώματος ώστε να υπολογίσουμε το λίπος που υπάρχει στο σώμα του κάθε ανθρώπου.

Το θέμα αυτής της διπλωματικής εργασίας χωρίζεται σε δύο μέρη. Το πρώτο μέρος αφορά τη πολυκριτηριακή ανάλυση παλινδρόμησης και το δεύτερο μέρος επικεντρώνεται στην επιλογή μεταβλητών. Ουσιαστικά, δημιουργούμε δύο ομάδες μετρήσεων. Από το σύνολο των μετρήσεων τις 25 τις απομονώνουμε και τις υπόλοιπες τις θεωρούμε μια δεύτερη ομάδα. Την ομάδα με τις 25 μετρήσεις την ονομάζουμε εκτός δείγματος παρατηρήσεις. Έτσι, λοιπόν εφαρμόζουμε μεθόδους ώστε να δημιουργήσουμε προβλεπτικά μοντέλα και ταυτόχρονα να προσδιορίσουμε τον αριθμό των μεταβλητών. Τα αποτελέσματα ποικίλουν, καθώς έχουμε δύο κατηγορίες. Η μια κατηγορία αναφέρεται στις εκτός δείγματος παρατηρήσεις και η άλλη στις εντός δείγματος παρατηρήσεις.

Η παρούσα διπλωματική εργασία δομείται από τα εξής κεφάλαια:

Στο **κεφάλαιο 1^ο** εξηγούμε το σκοπό αυτής της εργασίας και διάφορες βιβλιογραφικές αναφορές που έχουν γίνει ανά διαστήματα στο συγκεκριμένο ερευνητικό τομέα.

Στο **κεφάλαιο 2^ο** είναι αφιερωμένο στο θεωρητικό υπόβαθρο που χρειάζεται να γνωρίζει κάποιος ώστε να μπορέσει να κατανοήσει τους υπολογισμούς που έχουμε κάνει στα παρακάτω κεφάλαια.

Στο **κεφάλαιο 3^ο** μελετάμε το πρόβλημα μας και σχεδιάζουμε την ανάλυση του.

Στο *κεφάλαιο 4^ο* εξηγούμε τη διαδικασία ανάλυσης που ακολουθούμε στη δική μας περίπτωση.

Στο *κεφάλαιο 5^ο* εξηγούμε τον τρόπο ανάπτυξης των μοντέλων μας καθώς και τη γλώσσα προγραμματισμού που χρησιμοποιούμε στην εργασία αυτή.

Στο *κεφάλαιο 6^ο* αναλύουμε τα αποτελέσματα των μοντέλων μας και καταλήγουμε σε κάποια συμπεράσματα τα οποία προκύπτουν από αυτή την διπλωματική εργασία.

Κεφάλαιο 2

Θεωρητικό Μέρος

Σκοπός αυτού του κεφαλαίου είναι να αναπτύξουμε και να υπενθυμίσουμε έννοιες και όρους που θα χρησιμοποιήσουμε στη παρούσα διπλωματική εργασία. Αρχικά, αναλύουμε το Μαθηματικό Προγραμματισμό (MP) και πως αυτός αναλύεται. Στη συνέχεια, εστιάζουμε στα είδη Παλινδρόμησης (Regression) και στις μεθόδους που χρησιμοποιούμε για την Επιλογή Μεταβλητών (Variable Selection). Τέλος, αναλύουμε την πολλαπλή κριτηριακή παλινδρόμηση (Multiple Criteria Regression) στην οποία επικεντρώνεται ένα μεγάλο μέρος της διπλωματικής αυτής εργασίας.

Περιεχόμενα

2.1	<i>Εισαγωγή στην Επιχειρησιακή Έρευνα.....</i>	11
2.2	<i>Μαθηματικός Προγραμματισμός.....</i>	13
2.3	<i>Ακέραιος Προγραμματισμός (Integer Programming).....</i>	16
2.4	<i>Πολυκριτηριακός Προγραμματισμός (Multiple Programming).....</i>	19
2.5	<i>Ανάλυση Παλινδρόμησης (Regression Analysis).....</i>	25
2.5.1	Στόχοι της Ανάλυσης Παλινδρόμησης	30
2.5.2	Τα βήματα για την ανάλυση παλινδρόμησης	30
2.6	<i>Απλή Γραμμική Παλινδρόμηση (Simple Linear Regression).....</i>	31
2.7	<i>Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression).....</i>	35
2.8	<i>Επιλογή Μεταβλητών (Variable Selection)</i>	41
2.8.1	Ιεραρχία Μοντέλων	43
2.8.2	Βηματικές Διαδικασίες (Stepwise Procedures).....	44
2.8.3	Criterion-based procedures.....	46
2.8.4	Περιληπτικά.....	50
2.9	<i>Πολυκριτηριακή Παλινδρόμηση (Multiple Criteria Regression).....</i>	50

2.1 Εισαγωγή στην Επιχειρησιακή Έρευνα

Η Επιχειρησιακή Έρευνα, Operational (ή Operations) Research είναι επιστημονικός κλάδος που ασχολείται με την ανάπτυξη μαθηματικών μοντέλων για την περιγραφή συστημάτων και διαδικασιών με κύριο σκοπό την αριστοποίηση τους και τη λήψη αποφάσεων. Ο όρος «Επιχείρηση» (Operation) έχει την έννοια της διαδικασίας, λειτουργίας και όχι της εταιρίας. Μετά τον Β' Παγκόσμιο Πόλεμο καθιερώθηκε ως νέο επιστημονικό πεδίο και αναπτύχθηκε ραγδαία κυρίως στις ΗΠΑ. Πρώτες εφαρμογές σε βιομηχανία και διοίκηση τη δεκαετία του 1950. Μάθημα στα πανεπιστήμια πριν από το 1960. Τις δεκαετίες του 1950 και 1960 αναπτύχθηκαν οι περισσότεροι αλγόριθμοι και μέθοδοι που χρησιμοποιούνται ακόμα και σήμερα. Ραγδαία ήταν η ανάπτυξη με την πρόοδο των Η/Υ.

Καθορισμός του προβλήματος

- ❖ (Ορισμός του συστήματος, καθορισμός στόχων)

Διατύπωση του μοντέλου

- ❖ Καθορισμός μεταβλητών απόφασης (=άγνωστοι), εύρεση δεδομένων για τον καθορισμό παραμέτρων

(=γνωστοί), διατύπωση μαθηματικών σχέσεων

- ❖ Επίλυση του μοντέλου

Αλγόριθμος επίλυσης

Έλεγχος αποτελεσμάτων

- ❖ Ανάλυση ευαισθησίας, αβεβαιότητες

Υλοποίηση της τελικής λύσης

ΚΑΤΗΓΟΡΙΕΣ ΜΕΘΟΔΩΝ ΕΠΙΧΕΙΡΗΣΙΑΚΗΣ ΕΡΕΥΝΑΣ

- Μαθηματικός Προγραμματισμός

Γραμμικός Προγραμματισμός (Linear Programming)

Ακέραιος Προγραμματισμός (Integer Programming)

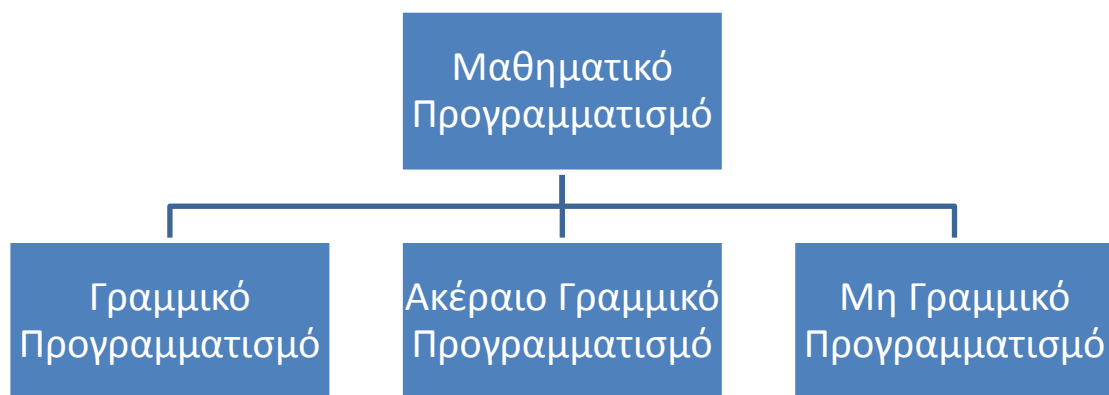
Μικτός Ακέραιος Γραμμικός Προγραμματισμός (Mixed Integer LP)

Μη-Γραμμικός Προγραμματισμός, Πολυκριτηριακός Προγραμματισμός,

Δυναμικός Προγραμματισμός

- Δένδρα αποφάσεων (Decision trees)
- Πολυκριτηριακή Ανάλυση (Multiple Criteria Decision)

- Ανάλυση (Analysis)
- Ανάλυση δικτύων (Network flows, PERT, CPM)
- Διαχείριση αποθεμάτων (Inventory control, EOQ)
- Ανάλυση γραμμών αναμονής (Queuing theory, simulation)
- Θεωρία παιγνίων (Game theory)
- Προσομοίωση (Simulation)
- Ευρεστικές τεχνικές (Heuristics)



Σχήμα 2.1: Ταξινόμηση Κατηγοριών Γραμμικού Προγραμματισμού

Τα Στάδια της Μεθόδου

Η αντιμετώπιση ενός προβλήματος όπου πρέπει να ληφθεί η βέλτιστη απόφαση περιλαμβάνει συνήθως τα εξής στάδια:

1. Αναγνώριση και περιγραφή του προβλήματος. αιτίες του προβλήματος καθορισμός στόχων με αντικειμενικό τρόπο οποιοδήποτε λάθος θα οδηγήσει σε αποτυχία τα επόμενα στάδια.
2. Καθορισμός των παραμέτρων του προβλήματος. Ποιοι παράγοντες επηρεάζουν τη λύση. Πως μεταβάλλονται.
3. Εντοπισμός των περιορισμών του προβλήματος.
4. Αναζήτηση λύσεων και επιλογή της Βέλτιστης λύσης. Βέλτιστη λύση, με βάση των αντικειμενικό στόχο που θέσαμε στο βήμα 1.
5. Δοκιμή και υλοποίηση-εφαρμογή της Βέλτιστης λύσης.

Το Μοντέλο του Προβλήματος

Τα τρία πρώτα στάδια συχνά αναφέρονται και σαν μοντελοποίηση του προβλήματος. Το μαθηματικό μοντέλο ενός προβλήματος περιλαμβάνει:

1. Τις μεταβλητές (μεταβάλλουμε για να πετύχουμε το στόχο)
2. Τις παραμέτρους (Τεχνολογικοί συντελεστές)

η τιμή ενός προϊόντος, η ταχύτητα ή ο χρόνος μεταφοράς του, η αναλογία ανάμιξης δυο υλικών, η αξία μιας ανθρωποώρας, κλπ.

3. Τους περιορισμούς (ή συνθήκες) - μορφή ανισοεξισώσεων.
4. Τον αντικειμενικό στόχο (ή αντικειμενική συνάρτηση, ΑΣ) δεν είναι πάντα μοναδικός αλλά μπορεί να αποτελείται από επί μέρους στόχους.

Ο ορισμός σύμφωνα με την Βρετανική Εταιρεία Επιχειρησιακής Έρευνας ορίζεται ως:

« Εφαρμογή επιστημονικών μεθόδων από διεπιστημονικές ομάδες για την αντιμετώπιση σύνθετων προβλημάτων ελέγχου οργανωμένων συστημάτων ανθρώπων και μηχανών με στόχο την εύρεση λύσεων που εξυπηρετούν τους συνολικούς σκοπούς της επιχείρησης με τον καλύτερο τρόπο. »

2.2 Μαθηματικός Προγραμματισμός (Mathematical Programming)

Ο Μαθηματικός Προγραμματισμός είναι ίσως η πιο ανεπτυγμένη και περισσότερο χρησιμοποιημένη τεχνική λήψης αποφάσεων στην οικονομία ή τη διοίκηση επιχειρήσεων. Έχει ως αντικείμενο την άριστη κατανομή περιορισμένων πόρων μεταξύ διάφορων ανταγωνιστικών δραστηριοτήτων κάτω από συνθήκες βεβαιότητας . Ένα μοντέλο Μαθηματικού Προγραμματισμού αποτελείται από μια ‘‘αντικειμενική συνάρτηση’’ η οποία εκφράζει το στόχο που θέλουμε να μεγιστοποιήσουμε ή να ελαχιστοποιήσουμε και από ένα σύνολο περιορισμών (δυναμικότητας, διαθεσιμότητας πόρων, τεχνολογίας, κλπ.), που εκφράζουν τους περιορισμούς του περιβάλλοντος μέσα στο οποίο πρέπει να κινηθούμε . Γενικό χαρακτηριστικό των μοντέλων Μαθηματικού Προγραμματισμού είναι ότι και η αντικειμενική συνάρτηση και οι περιορισμοί εκφράζονται ως μαθηματικές συναρτήσεις. (Γ.Πραστάκος,2003)

Τα βασικά χαρακτηριστικά ενός προβλήματος Μαθηματικού Προγραμματισμού είναι τα ακόλουθα:

Μεταβλητές απόφασης: εκφράζουν ουσιαστικά τους αγνώστους του προβλήματος και είναι οι μεταβλητές που ελέγχει ο αποφασίζων, δηλ.

εκείνες των οποίων τις τιμές μπορεί να καθορίσει. Το σύνολο των μεταβλητών απόφασης αποτελεί ουσιαστικά το αντικείμενο της διαδικασίας λήψης απόφασης. Η διαδικασία αριστοποίησης αποσκοπεί στο να βρεθούν οι τιμές εκείνες για τις μεταβλητές απόφασης οι οποίες βελτιστοποιούν την αντικειμενική συνάρτηση.

Αντικειμενική συνάρτηση: αποτελεί τη μαθηματική σχέση των μεταβλητών απόφασης που εκφράζει το κριτήριο βελτιστοποίησης. Επιδιώκεται είτε η ελαχιστοποίηση είτε η μεγιστοποίησή της τιμής της αντικειμενικής συνάρτησης. Στα προβλήματα Πολυκριτηριακού Μαθηματικού Προγραμματισμού υπάρχουν περισσότερες από μία αντικειμενικές συναρτήσεις (κριτήρια απόφασης), γι αυτό και τα προβλήματα αυτά αναφέρονται και ως προβλήματα διανυσματικής βελτιστοποίησης (vector optimization).

Περιορισμοί: είναι οι μαθηματικές σχέσεις που καθορίζουν τις τιμές που μπορούν να πάρουν οι μεταβλητές απόφασης στη διαδικασία της βελτιστοποίησης. Καθορίζουν δηλαδή το πεδίο ορισμού (εφικτό χωρίο) του προβλήματος. Οι περιορισμοί μπορεί να είναι ισότητες ή ανισότητες.

Παράμετροι: είναι τα εξωγενώς οριζόμενα (εκτός του ελέγχου του αποφασίζοντα) μεγέθη του προβλήματος. Πρόκειται ουσιαστικά για τους γνωστούς όρους του προβλήματος οι οποίοι έχουν σταθερή τιμή στη διαδικασία βελτιστοποίησης. Συνήθως είναι συντελεστές των μεταβλητών απόφασης ή εκφράζουν ποσότητες απαραίτητες στη διαμόρφωση των περιορισμών (π.χ. την απαιτούμενη ζήτηση μιας δραστηριότητας) (Williams, 1985; Winston and Venkataramanan, 2003; Σίσκος, 1998).

Ο Μαθηματικός Προγραμματισμός περιλαμβάνει μια σειρά από τεχνικές ανάλογα με τις προϋποθέσεις που ισχύουν στο κάθε πρόβλημα. Ο Γραμμικός Προγραμματισμός είναι ο πιο γνωστός κλάδος Μαθηματικού Προγραμματισμού. Προϋποθέτει ότι οι συναρτήσεις στόχου και περιορισμών είναι γραμμικές και οι μεταβλητές αποφάσεων μπορούν να πάρουν όχι μόνο ακέραιες, αλλά και δεκαδικές τιμές. Άλλος επίσης γνωστός και ιδιαίτερα σημαντικός κλάδος του Μαθηματικού Προγραμματισμού είναι ο Ακέραιος Προγραμματισμός, ο οποίος εφαρμόζεται όταν οι μεταβλητές αυτές αναπαριστούν αποφάσεις "λογικής" και όχι φυσικά μεγέθη. Συνηθισμένο παράδειγμα προβλημάτων Ακέραιου Προγραμματισμού είναι τα προβλήματα επενδύσεων, όπου η μεταβλητή για μια επένδυση μπορεί να πάρει μόνο δύο τιμές: "1" εάν αναληφθεί η επένδυση και "0" εάν όχι. Ένας τρίτος γνωστός κλάδος Μαθηματικού Προγραμματισμού είναι ο Μη Γραμμικός Προγραμματισμός, όπου μερικές από τις συναρτήσεις του προβλήματος (αντικειμενική συνάρτηση ή/και περιορισμοί) είναι μη γραμμικές. Τέλος, ο Δυναμικός Προγραμματισμός είναι ένας κλάδος του Μαθηματικού Προγραμματισμού που προσφέρει εναλλακτικό, ακόμα και κάτω από συνθήκες αβεβαιότητας (Γ.Πραστάκος,2003).

Όταν οι μαθηματικές σχέσεις που περιγράφουν το πρόβλημα (αντικειμενικές συναρτήσεις και περιορισμοί) είναι γραμμικές ως προς τις μεταβλητές απόφασης τότε το πρόβλημα χαρακτηρίζεται ως πρόβλημα Γραμμικού Προγραμματισμού (Linear Programming), ενώ αν είναι μη γραμμικές χαρακτηρίζεται ως πρόβλημα Μη Γραμμικού Προγραμματισμού (Non Linear Programming). Τα προβλήματα Γραμμικού Προγραμματισμού αποτελούν τη συντριπτική πλειοψηφία των προβλημάτων Μαθηματικών Προγραμματισμού κυρίως λόγω των συγκεκριμένων χαρακτηριστικών τους και την ευκολία επίλυσης τους. Με τη μέθοδο Simplex και τις παραλλαγές της να κυριαρχούν στην επίλυση τέτοιου είδους προβλημάτων εδώ και 60 περίπου χρόνια, προβλήματα Γραμμικού Προγραμματισμού με χιλιάδες μεταβλητές απόφασης και περιορισμούς επιλύονται σήμερα σε δευτερόλεπτα. Αντίθετα η επίλυση προβλημάτων Μη Γραμμικού Προγραμματισμού είναι πιο δύσκολη υπόθεση και καταλήγει σε τοπικά βέλτιστα τα οποία δεν είναι πάντα και ολικά βέλτιστα. Για τους λόγους αυτούς επιδιώκεται στις περισσότερες περιπτώσεις τα πραγματικά προβλήματα να μοντελοποιούνται ως προβλήματα Γραμμικού Προγραμματισμού καταφεύγοντας αρκετές φορές σε προσεγγίσεις μη γραμμικών συστημάτων με γραμμικές σχέσεις.

Μία άλλη ταξινόμηση είναι ανάλογα με το είδος των μεταβλητών απόφασης αν δηλαδή είναι συνεχείς μεταβλητές ή ακέραιες. Τα προβλήματα που έχουν μόνο συνεχείς μεταβλητές είναι πιο εύκολο να λυθούν σε σχέση με αυτά που έχουν ακέραιες μεταβλητές. Αυτό οφείλεται στο γεγονός ότι το εφικτό χωρίο σε ένα πρόβλημα με ακέραιες μεταβλητές παρουσιάζει ασυνέχειες δυσκολεύοντας έτσι κατά πολύ τη διαδικασία επίλυσης. Από την άλλη μεριά όμως η δυνατότητα χρήσης ακεραίων μεταβλητών δίνει τη δυνατότητα μιας πιο ρεαλιστικής μοντελοποίησης της πραγματικότητας και επίσης επεκτείνει σημαντικά το πεδίο εφαρμογής του Μαθηματικού Προγραμματισμού και σε προβλήματα που έχουν συνδυαστικό χαρακτήρα. (συνδυαστική βελτιστοποίηση) τα οποία χωρίς τη χρήση ακεραίων μεταβλητών θα ήταν αδύνατο να λυθούν. Στο 95% των περιπτώσεων οι ακέραιες μεταβλητές που συναντώνται σε μοντέλα Μαθηματικού Προγραμματισμού είναι δυαδικές μεταβλητές δηλαδή παίρνουν τιμή 0 ή 1. Αν ένα μοντέλο Μαθηματικού Προγραμματισμού έχει αποκλειστικά ακέραιες μεταβλητές χαρακτηρίζεται ως μοντέλο Ακέραιου Προγραμματισμού (Integer Programming). Αν έχει και συνεχείς και ακέραιες μεταβλητές χαρακτηρίζεται ως μοντέλο Μικτού Ακέραιου Προγραμματισμού (Mixed Integer Programming). Η επίλυση προβλημάτων Ακέραιου και Μικτού Ακέραιου Προγραμματισμού γίνεται συνήθως με τη μεθοδολογία “διακλάδωση και οριοθέτηση” (branch and bound), μια τεχνική συστηματικής εξερεύνησης του πεδίου των δυνατών λύσεων. Σε κάποιες περιπτώσεις οι παράμετροι ενός μοντέλου Μαθηματικού

Προγραμματισμού μπορεί να μην εκφράζονται με πραγματικούς αριθμούς αλλά με κατανομές πιθανότητας ή με ασαφείς αριθμούς απεικονίζοντας έτσι την αβεβαιότητα ως προς την τιμή τους. Τότε το πρόβλημα ανάγεται αντίστοιχα σε πρόβλημα Στοχαστικού Προγραμματισμού (Stochastic Programming) ή Ασαφούς Προγραμματισμού (Fuzzy Programming). Τα τελευταία χρόνια μάλιστα έχει αρχίσει να απασχολεί ιδιαίτερα η διαχείριση της αβεβαιότητας ως προς τις παραμέτρους ενός μοντέλου ξεφεύγοντας από τις απλές μορφές ανάλυσης ευαισθησίας που μπορεί να προσφέρει και ο Μαθηματικός Προγραμματισμός.

Τέλος όταν υπάρχουν περισσότερες από μία αντικειμενικές συναρτήσεις (όπως και στην προκειμένη περίπτωση), το πρόβλημα χαρακτηρίζεται ως πρόβλημα Πολυκριτηριακού Μαθηματικού Προγραμματισμού (Multiobjective Programming, Multicriteria Programming). Ο όρος πολυκριτηριακή βελτιστοποίηση είναι ταυτόσημος με τον όρο διανυσματική βελτιστοποίηση (vector optimization) σε αντιδιαστολή με την μονοδιάστατη βελτιστοποίηση (scalar optimization) που πραγματεύεται ο συμβατικός Μαθηματικός Προγραμματισμός. Ο Πολυκριτηριακός Μαθηματικός Προγραμματισμός έκανε την εμφάνισή του τη δεκαετία του 70 ενώ αναπτύχθηκε κυρίως τις δύο τελευταίες δεκαετίες, όταν η θεώρηση περισσοτέρων από μιας αντικειμενικών συναρτήσεων άρχισε να καθορίζει ένα πιο ρεαλιστικό πλαίσιο μοντελοποίησης των πολύπλοκων προβλημάτων μανάτζμεντ (Cohon, 1978; Steuer, 1989, Evans, 1984). Η πολλαπλότητα των κριτηρίων στη σύγχρονη λήψη αποφάσεων όπου πλέον μαζί με το συνηθισμένο κριτήριο βελτιστοποίησης (το οικονομικό) εξετάζονται και άλλα κριτήρια (περιβαλλοντικά, κοινωνικά κλπ) καθιέρωσε το Πολυκριτηριακό Μαθηματικό Προγραμματισμό ως ένα σύγχρονο, πλήρες εργαλείο στη διαδικασία λήψης αποφάσεων.

2.3 Ακέραιος Προγραμματισμός (Integer Programming)

Μία από τις προϋποθέσεις εφαρμογής του συνεχούς γραμμικού προγραμματισμού είναι η διαιρετότητα των μεταβλητών απόφασης. Σε ένα κλασσικό (συνεχές) γραμμικό πρόγραμμα, οι μεταβλητές μπορούν να πάρουν οποιαδήποτε πραγματική τιμή. Υπάρχει, όμως ένας σημαντικός αριθμός προβλημάτων γραμμικού προγραμματισμού στα οποία όλες οι μεταβλητές, ή μερικές από αυτές, υποχρεούνται να πάρουν μόνο ακέραιες τιμές. Τέτοιες μεταβλητές είναι, για παράδειγμα, αυτές που δηλώνουν αριθμό εργατών, αριθμό εργοστασιακών μονάδων, αποφάσεις χρηματοδότησης ή μη χρηματοδότησης ενός έργου, κ.λ.π. Τα προβλήματα του γραμμικού προγραμματισμού στα οποία όλες ανεξαιρέτα οι μεταβλητές απόφασης περιορίζονται να πάρουν ακέραιες τιμές, εμπίπτουν στο πεδίο του ακέραιου γραμμικού προγραμματισμού. Εκείνα, στα οποία ο περιορισμός ακεραιότητας δεν ισχύει για όλες τις

μεταβλητές, αλλά για μερικές από αυτές, ονομάζονται προβλήματα μικτού ακέραιου προγραμματισμού.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η περίπτωση που οι μεταβλητές είναι περιορισμένες να παίρνουν τιμές 0 ή 1. Το πρόβλημα αυτό καλείται πρόβλημα δυαδικού ακέραιου προγραμματισμού. Μια σημαντική χρήση μιας δυαδικής μεταβλητής είναι να κωδικοποιήσουμε μια απόφαση μεταξύ δύο εναλλακτικών που θα πρέπει να ληφθεί στο πρόβλημα. Η τιμή που θα πάρει η μεταβλητή απόφασης κατά την επίλυση δείχνει ποια απόφαση πρέπει να επιλεγεί ώστε να βελτιστοποιηθεί η αντικειμενική συνάρτηση.

Μία απλή μέθοδος επίτευξης ακέραιης λύσης σε ένα πρόβλημα ακέραιου γραμμικού προγραμματισμού συνίσταται στο να επιλυθεί με τη συνηθισμένη μέθοδο Simplex για συνεχή γραμμικό προγραμματισμό και στη συνέχεια, να στρογγυλευθούν στον πλησιέστερο ακέραιο οι τιμές των μεταβλητών οι οποίες παίρνουν ρητές τιμές. Μία τέτοια διαδικασία είναι πολύ επικίνδυνη, όσο απλή κι αν φαίνεται, γιατί μπορεί να καταλήξει είτε σε υποβέλτιστες λύσεις, κατώτερες δηλαδή της πραγματικά βέλτιστης ακέραιης λύσης, είτε σε λύσεις μη πραγματοποιήσιμες, που παραβιάζουν δηλαδή τουλάχιστον έναν από τους περιορισμούς του προβλήματος. Το πρόβλημα του γραμμικού προγραμματισμού που προκύπτει από ένα πρόβλημα ακέραιου ή μικτού ακέραιου προγραμματισμού αν αφαιρέσουμε τη συνθήκη οι μεταβλητές να είναι ακέραιες, ονομάζεται το γραμμικό πρόβλημα χαλάρωσης του ακέραιου προγραμματισμού. Στην πραγματικότητα, ο χώρος των εφικτών λύσεων για ένα πρόβλημα ακέραιου προγραμματισμού περικλείεται μέσα στο χώρο των εφικτών λύσεων του αντίστοιχου γραμμικού προβλήματος χαλάρωσης. Αυτό συνεπάγεται ότι η βέλτιστη τιμή για το γραμμικό πρόβλημα χαλάρωσης θα είναι μεγαλύτερη ή ίση της αντίστοιχης βέλτιστης τιμής για το πρόβλημα του ακέραιου προγραμματισμού.

Σε περιπτώσεις που οι μεταβλητές απόφασης ενός γραμμικού προβλήματος είναι φραγμένες, παίρνουν δηλαδή περιορισμένο αριθμό ακέραιων τιμών, οι ιδεώδεις μέθοδοι επίλυσης του είναι οι μέθοδοι κλάδου και φράγματος (branch and bound methods) οι οποίοι στηρίζονται σε μια έμμεση απαρίθμηση των δυνατών ακέραιων λύσεων που επιδέχεται το πρόβλημα. Φυσικά, σήμερα υπάρχουν αρκετές μέθοδοι μικτού ακέραιου και ακέραιου γραμμικού προγραμματισμού. Παρόλο όμως, που αρκετή ανθρώπινη προσπάθεια έχει αφιερωθεί με σκοπό να κατασκευαστούν αποτελεσματικοί αλγόριθμοι για τη λύση των προβλημάτων αυτών, εν τούτοις δε μπορεί να πει κανείς ότι κάθε πρόβλημα ακέραιου προγραμματισμού έχει βρει τη λύση του. Πάντως, ενδιαφέρον είναι το γεγονός, ότι τα προβλήματα του ακέραιου προγραμματισμού που παρουσιάζονται σαν συνέπεια ενός πραγματικού προβλήματος σχεδόν πάντα λύνονται.

Είναι γεγονός πάντως, ότι σε προβλήματα όπου οι μεταβλητές ξεπερνούν τις 100, το μέγεθος του προβλήματος ξεπερνά συνήθως τα όρια και η πιθανότητα να το λύσει κανείς σε ένα λογικό χρόνο στον υπολογιστή είναι πολύ μικρή.

Ο ακέραιος προγραμματισμός έχει σημαντικές πρακτικές εφαρμογές και αποτελεί ένα ισχυρό πλαίσιο μοντελοποίησης που παρέχει μεγάλη ευελιξία στην έκφραση διακριτών προβλημάτων βελτιστοποίησης. Παρόλα αυτά στον αντίποδα έχουμε το γεγονός ότι ο ακέραιος προγραμματισμός είναι ένα πολύ πιο δύσκολο πρόβλημα από ότι ο γραμμικός προγραμματισμός.

Στο Μικτό Ακέραιο Προγραμματισμό (mixed integer programming-IP) οι μεταβλητές ενός γνήσιου μη κενού υποσυνόλου του συνόλου των μεταβλητών παίρνουν ακέραιες τιμές και οι υπόλοιπες μεταβλητές είναι συνεχείς. Διακρίνουμε τις περιπτώσεις του μικτού ακέραιου προγραμματισμού και του μικτού ακέραιου μη γραμμικού προγραμματισμού.

Ένα πρόβλημα μικτού ακέραιου προγραμματισμού (MIP) που αφορά μεγιστοποίηση διατυπώνεται ειδικότερα ως εξής:

Μεγιστοποίησε

$$z_{MIP} = \sum_{j \in I} c_j x_j + \sum_{j \in C} c_j x_j \quad (2.1)$$

Με περιορισμούς

$$\sum_{j \in I} a_{ij} x_j + \sum_{j \in C} a_{ij} x_j \leq, \geq \text{ ή } = b_i \quad i = 1, \dots, m \quad (2.2)$$

$$l_j \leq x_j \leq u_j \quad j \in N \quad (2.3)$$

$$x_j \in Z \quad j \in I$$

$$x_j \in R \quad j \in C$$

I είναι το σύνολο των ακέραιων μεταβλητών, C είναι το σύνολο των συνεχών μεταβλητών και $N = I \cup C$. Τα κάτω και άνω όρια l_i και u_i μπορούν να πάρουν τιμές $\pm\infty$.

$\sum_{j \in N} c_j x_j$ είναι η αντικειμενική συνάρτηση της οποίας ζητείται η βελτιστοποίηση (μεγιστοποίηση ή ελαχιστοποίηση). Μια εφικτή λύση είναι κάθε διάνυσμα x που ικανοποιεί όλους τους περιορισμούς που ορίζει το πρόβλημα MIP και ως εφικτή περιοχή (feasible region) του προβλήματος ορίζεται το σύνολο όλως των εφικτών λύσεων. Μια εφικτή λύση όπου η αντικειμενική συνάρτηση παίρνει τη μέγιστη δυνατή τιμή (για προβλήματα μεγιστοποίησης) ονομάζεται βέλτιστη λύση.

2.4 Πολυκριτηριακός Προγραμματισμός (Multiple Objective /Multiobjective Programming)

Η πολυκριτηριακή ανάλυση αποφάσεων αποτελεί ένα εξελιγμένο πεδίο της επιχειρησιακής έρευνας, ο οποίος τις τελευταίες τρεις δεκαετίες έχει γνωρίσει ιδιαίτερη άνθηση τόσο σε θεωρητικό όσο και σε πρακτικό επίπεδο. Βασικό ρόλο στην ανάπτυξη και διάδοση της πολυκριτηριακής ανάλυσης αποτέλεσε η απλή διαπίστωση ότι η επίλυση πολύπλοκων και ιδιαίτερα σημαντικών προβλημάτων λήψης αποφάσεων δεν είναι δυνατό να πραγματοποιείται μέσω μιας μονόπλευρης και μονοδιάστατης ανάλυσης. Κατά την προσπάθεια, όμως, εξέτασης όλων των παραμέτρων ενός προβλήματος και των κριτηρίων-παραγόντων που επηρεάζουν τη λήψη της κατάλληλης απόφασης, γεννιάται ένα ιδιαίτερα σημαντικό πρόβλημα, το οποίο ορισμένες φορές αποθαρρύνει τους αποφασίζοντες και αναλυτές από την υιοθέτηση αυτής της πιο ρεαλιστικής προσέγγισης. Το πρόβλημα αυτό αναφέρεται στον τρόπο με τον οποίο μπορεί να πραγματοποιηθεί η σύνθεση όλων των παραμέτρων ώστε να επιτευχθεί η λήψη ορθολογικών αποφάσεων. Η αντιμετώπιση του προβλήματος αυτού αποτελεί το βασικό αντικείμενο της πολυκριτηριακής ανάλυσης αποφάσεων. Η κύρια όμως ειδοποιός διαφορά της πολυκριτηριακής ανάλυσης από άλλες εναλλακτικές προσεγγίσεις, δεν είναι η απλή σύνθεση όλων των παραμέτρων ενός προβλήματος. Αυτή πραγματοποιείται και μέσω άλλων μεθοδολογικών προσεγγίσεων. Το βασικό χαρακτηριστικό γνώρισμα της πολυκριτηριακής ανάλυσης είναι η πραγματοποίηση της αναγκαίας σύνθεσης υπό το πρίσμα της πολιτικής λήψης των αποφάσεων και του συστήματος προτιμήσεων και αξιών, το οποίο συνειδητά ή ασυνείδητα χρησιμοποιεί ο αποφασίζων.

Το χαρακτηριστικό αυτό έχει ιδιαίτερη σημασία στο χώρο της λήψης αποφάσεων. Όπως είναι κατανοητό, το αποτέλεσμα της οποίας ανάλυσης πραγματοποιείται με σκοπό την αντιμετώπιση ενός προβλήματος λήψης αποφάσεων, έχει ως τελικό αποδέκτη τον ίδιο τον αποφασίζοντα. Συνεπώς, η ανάπτυξη υποδειγμάτων λήψης αποφάσεων μέσω μεθοδολογικών προσεγγίσεων που δεν είναι σε θέση να ενσωματώσουν τον αποφασίζοντα και τις προτιμήσεις του στη διαδικασία ανάπτυξης των υποδειγμάτων αυτών, ουσιαστικά προσδίδουν στον αποφασίζοντα έναν παθητικό ρόλο, ο οποίος περιορίζεται στην παρακολούθηση και εφαρμογή των αποτελεσμάτων μαθηματικών υποδειγμάτων. Υπό το πρίσμα των παρατηρήσεων αυτών, η πολυκριτηριακή ανάλυση έχει δώσει ιδιαίτερο ενδιαφέρον στην έρευνα θεμάτων που σχετίζονται με την ανάλυση, μαθηματική μοντελοποίηση και αναπαράσταση των προτιμήσεων που διέπουν την πολιτική λήψης αποφάσεων από τη πλευρά του εκάστοτε αποφασίζοντα. Απώτερος στόχος είναι η παροχή των απαραίτητων πληροφοριών για την υποστήριξη της διαδικασίας λήψης των αποφάσεων, συμβάλλοντας στον εντοπισμό των βασικών χαρακτηριστικών του εξεταζόμενου

προβλήματος καθώς και των ιδιαιτεροτήτων των διαθέσιμων εναλλακτικών λύσεων. Οι μεθοδολογικές εξελίξεις που έχουν πραγματοποιηθεί κατά τη διάρκεια της ιστορικής πορείας του χώρου της πολυκριτηριακής ανάλυσης, καλύπτουν, όπως θα παρουσιαστεί στη συνέχεια, όλα τα είδη των προβλημάτων λήψης αποφάσεων. Στις παραγράφους που ακολουθούν παρουσιάζονται αναλυτικότερα οι κύριες μεθοδολογικές προσεγγίσεις της πολυκριτηριακής ανάλυσης.

Ως πρώτη τεκμηριωμένη προσπάθεια επιστημονικής αντιμετώπισης του προβλήματος της σύνθεσης πολλαπλών κριτηρίων μπορεί να θεωρηθεί η εργασία του Pareto (1896), ο οποίος έθεσε τις απαραίτητες αξιωματικές βάσεις, εισάγοντας παράλληλα μια εκ των πλέον βασικών εννοιών της σύγχρονης πολυκριτηριακής ανάλυσης, την έννοια της αποτελεσματικότητας (efficiency). Μεταπολεμικά, ο Koopmans (1951) επέκτεινε την έννοια της αποτελεσματικότητας του Pareto εισάγοντας την έννοια του αποτελεσματικού συνόλου, δηλαδή του συνόλου των εναλλακτικών δραστηριοτήτων οι οποίες δεν κυριαρχούνται από καμία άλλη εναλλακτική δραστηριότητα (non-dominated set of alternatives).

Κατά την ίδια περίπου χρονική περίοδο (1940–1950) οι Von Neumann και Morgenstern (1944) αναπτύσσουν τη θεωρία χρησιμότητας, η οποία αποτελεί τη βάση ενός από τα κυριότερα μεθοδολογικά ρεύματα της πολυκριτηριακής ανάλυσης αποφάσεων. Στη δεκαετία του 1960 όλες οι προαναφερθείσες «προκαταρκτικές» ερευνητικές εργασίες αποτέλεσαν το έναυσμα για την πραγματοποίηση περαιτέρω έρευνας από τους Charnes και Cooper (1961) όσον αφορά τη σύνδεση της θεωρίας του γραμμικού προγραμματισμού και της πολυκριτηριακής ανάλυσης (προγραμματισμός στόχων – goal programming), καθώς και από τον Fishburn (1965) όσον αφορά την επέκταση της θεωρίας χρησιμότητας σε προβλήματα λήψης αποφάσεων υπό καθεστώς πολλαπλών κριτηρίων. Περί τα τέλη της δεκαετίας του 1960, η πολυκριτηριακή ανάλυση άρχισε να απασχολεί και τους Ευρωπαίους επιχειρησιακούς ερευνητές. Πρωτοπόρος μεταξύ αυτών υπήρξε ο Roy (1968), ο οποίος ανέπτυξε τη θεωρία των σχέσεων υπεροχής (outranking relations) και θεωρείται ο ιδρυτής της «Ευρωπαϊκής σχολής» της πολυκριτηριακής ανάλυσης.

Τις επόμενες δύο δεκαετίες (1970–1990), η πολυκριτηριακή ανάλυση αναπτύχθηκε ραγδαία σε θεωρητικό επίπεδο αλλά και σε θέματα πρακτικών εφαρμογών για την αντιμετώπιση διαφόρων πολύπλοκων πραγματικών προβλημάτων λήψης αποφάσεων. Προς την κατεύθυνση αυτή σημαντική υπήρξε η συμβολή της πληροφορικής και της επιστήμης των υπολογιστών. Η ταχύτατη τεχνολογική πρόοδος που συντελέστηκε στους χώρους αυτούς, κυρίως κατά τις τελευταίες δύο δεκαετίες, έδωσε τα απαραίτητα μέσα για την υλοποίηση των μεθοδολογικών εξελίξεων της πολυκριτηριακής ανάλυσης σε ολοκληρωμένα πληροφορικά συστήματα (πολυκριτηριακή συστήματα

υποστήριξης αποφάσεων), τα οποία παράλληλα συνέβαλλαν και στην προώθηση των πρακτικών εφαρμογών της πολυκριτηριακής ανάλυσης.

Η εισαγωγή πολλών αντικειμενικών συναρτήσεων σ' ένα πρόβλημα ΓΠ δημιουργεί το πρόβλημα της γραμμικής διανυσματικής μεγιστοποίησης (linear vector maximum problem). Με τον όρο μεγιστοποίηση εννοείται γενικότερα η αριστοποίηση αφού και η ελαχιστοποίηση μπορεί εύκολα να μετατραπεί σε μεγιστοποίηση με αλλαγή προσήμου. Η ανάλυση των προβλημάτων αυτών αποτελεί το αντικείμενο του ΠΚΓΠ. Το πρόβλημα του ΠΚΓΠ μαθηματικά ορίζεται ως εξής:

$$\max \{c^1 x = z_1 \} \quad (2.4)$$

$$\max \{c^2 x = z_2 \} \quad (2.5)$$

$$\max \{c^p x = z_p \} \quad (2.6)$$

$$\text{s.t. } x \in S = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0, b \in \mathbb{R}^m \} \quad (2.7)$$

S: το εφικτό χωρίο των περιορισμών

n: ο αριθμός των μεταβλητών

m: ο αριθμός των περιορισμών

p: ο αριθμός των αντικειμενικών συναρτήσεων

c^i : το διάνυσμα γραμμής των συντελεστών της 'i' αντικειμενικής συνάρτησης

z^i : η τιμή της 'i' αντικειμενικής συνάρτησης

A : η μήτρα ($m \times n$) των τεχνολογικών συντελεστών

b : το διάνυσμα ($m \times 1$) των σταθερών όρων (δεξί σκέλος περιορισμών)

x : το διάνυσμα ($n \times 1$) των μεταβλητών απόφασης

Οι μεταβλητές απόφασης εκφράζουν τα μεγέθη εκείνα του προβλήματος για τα οποία πρέπει να υπολογιστούν οι τιμές τους. Τα ήδη γνωστά μεγέθη του προβλήματος (μήτρα A, διανύσματα b και c_i) ονομάζονται και παράμετροι του προβλήματος. Οι αντικειμενικές συναρτήσεις είναι συναρτήσεις των μεταβλητών απόφασης, των οποίων επιδιώκεται η βελτιστοποίηση (μεγιστοποίηση ή ελαχιστοποίηση). Οι περιορισμοί είναι οι διάφορες σχέσεις (ισότητες ή ανισότητες ' \leq ' και ' \geq ') που πρέπει να πληρούν οι μεταβλητές απόφασης σύμφωνα με το πρόβλημα και οριοθετούν το εφικτό χωρίο S. Με τον όρο λύση ενός προβλήματος ΠΚΓΠ ή ΓΠ, εννοείται ο συνδυασμός των τιμών που λαμβάνουν οι μεταβλητές απόφασης. Αν η λύση αυτή ανήκει στο S (οι τιμές των μεταβλητών απόφασης ικανοποιούν τους περιορισμούς) τότε πρόκειται για εφικτή λύση του προβλήματος αλλιώς χαρακτηρίζεται ως μη εφικτή λύση.

Σε αντίθεση με τον ΓΠ όπου η διαδικασία επίλυσης επικεντρώνεται στην εξέταση του χώρου των μεταβλητών απόφασης R_n , στον ΠΚΓΠ έχει ιδιαίτερη σημασία ο χώρος των αντικειμενικών συναρτήσεων R_p . Αυτό συμβαίνει διότι η διαδικασία επίλυσης στα προβλήματα του ΠΚΓΠ αφορά την εξέταση των τιμών των αντικειμενικών συναρτήσεων που προκύπτουν από κάθε λύση. Συνεπώς, είναι αρκετά βοηθητική η απεικόνιση του προβλήματος στο χώρο των αντικειμενικών συναρτήσεων R_p αντί του χώρου των μεταβλητών απόφασης R_n . Εξ' άλλου, το p , δηλαδή ο αριθμός των αντικειμενικών συναρτήσεων, είναι συνήθως πολύ μικρότερο του n που εκφράζει το πλήθος των μεταβλητών απόφασης, κι έτσι η γραφική απεικόνιση της κατάστασης του προβλήματος είναι πιο εύκολη. Το βασικό εργαλείο επίλυσης των προβλημάτων ΓΠ είναι η μέθοδος Simplex που ανέπτυξε ο Dantzig το 1947. Από τη θεωρία της μεθόδου Simplex προκύπτει ότι τα ακραία σημεία αντιστοιχούν σε λύσεις που έχουν m βασικές και n μη βασικές μεταβλητές (m είναι ο αριθμός των περιορισμών και n ο αριθμός των μεταβλητών).

Κατά γενικό κανόνα, ως βασικές μεταβλητές (basic variables) ορίζονται οι μη μηδενικές μεταβλητές της λύσης και ως μη βασικές (non basic variables) οι μεταβλητές που παίρνουν την τιμή μηδέν στη λύση. Ικανή λύση (efficient, non-dominated solution): Μία λύση x' ενός προβλήματος ΠΚΓΠ λέγεται ικανή (ή κατά Pareto άριστη, ή αποτελεσματική) αν και μόνο αν $x' \in S$ και δεν υπάρχει άλλη λύση $x \in S$ τέτοια ώστε $c_i x \geq c_i x'$ για κάθε $i=1, 2, \dots, p$ και $c_i x > c_i x'$ για τουλάχιστον ένα i . Κάθε ικανή λύση αντιστοιχεί σε ένα μη βελτιώσιμο διάνυσμα στο χώρο των αντικειμενικών συναρτήσεων, με την εννοια ότι δεν μπορούμε να βελτιώσουμε την τιμή μιάς αντικειμενικής συναρτήσεως χωρίς να χειροτερεύσουμε τουλάχιστον μία από τις άλλες. Ο όρος "ικανή λύση" χρησιμοποιείται επίσης για να εκφράσει και το αντίστοιχο διάνυσμα τιμών των κριτηρίων $z'=(c_1x', c_2x', \dots, c_px')$ στο χώρο των αντικειμενικών συναρτήσεων R_p . Όταν υπάρχει λύση $x \in S$ τέτοια ώστε $c_i x \geq c_i x'$ για $i=1, 2, \dots, p$ με τουλάχιστον μία αυστηρή ανισότητα τότε η λύση x υπερτερεί της x' και η x' κυριαρχείται από την x . Αντικειμενικός σκοπός της επίλυσης των προβλημάτων ΠΚΓΠ είναι, αρχικά, ο προσδιορισμός των ικανών λύσεων και στη συνέχεια η υποστήριξη του αποφασίζοντα ώστε να επιλέξει εκείνη την ικανή λύση που εκφράζει καλύτερα τις προτιμήσεις του. Απαραίτητη προϋπόθεση για να είναι μία λύση ενός προβλήματος ΠΚΓΠ αποδεκτή είναι να πρόκειται για ικανή λύση.

Τελική ή σχετικά βέλτιστη λύση (final or best compromise solution): Η ικανή εκείνη λύση που τελικά επιλέγει από το σύνολο των ικανών λύσεων ο αποφασίζων, λέγεται τελική ή σχετικά βέλτιστη λύση. Ο όρος «σχετικά» σημαίνει ακριβώς ότι αποτελεί υποκειμενική επιλογή του αποφασίζοντα σε αντίθεση με τη βέλτιστη λύση σε ένα πρόβλημα ΓΠ που είναι αντικειμενικά προσδιοριζόμενη.

Ικανή ακραία λύση (efficient extreme solution): Μια ικανή λύση ενός προβλήματος ΠΚΓΠ που αντιστοιχεί σ'ένα ακραίο σημείο του εφικτού χωρίου του προβλήματος (κορυφή) λέγεται ικανή ακραία λύση. Μία ακραία λύση αντιστοιχεί σε ακραίο σημείο και στο χώρο των αντικειμενικών συναρτήσεων. Οι ακραίες λύσεις που αντιστοιχούν σε γειτονικές κορυφές στο εφικτό σύνολο ονομάζονται γειτονικές ακραίες λύσεις (adjacent extreme solutions).

Κάθε ικανή ακραία λύση έχει τουλάχιστον μια γειτονική ικανή ακραία λύση. Από αυτή την ιδιότητα προκύπτουν και οι παρακάτω αλγόριθμοι παραγωγής του συνόλου των ακραίων ικανών λύσεων.

Διάνυσμα τιμών των κριτηρίων (criterion vector): Κάθε λύση $\mathbf{x} = (x_1, x_2, \dots, x_n)$ που ανήκει στο χώρο των μεταβλητών απόφασης \mathbb{R}^n απαικονίζεται μέσω των αντικειμενικών συναρτήσεων στο χώρο των κριτηρίων \mathbb{R}^p σ' ένα διάνυσμα (σημείο) $\mathbf{z} = (z_1, z_2, \dots, z_p)$ όπου $z_1 = \mathbf{c}^1 \mathbf{x}$, $z_2 = \mathbf{c}^2 \mathbf{x}$ κλπ. Το διάνυσμα \mathbf{z} ορίζεται ως διάνυσμα τιμών των κριτηρίων.

Ιδεώδες σημείο (ideal point): Ιδεώδες σημείο (ή ιδεώδες διάνυσμα, ideal vector) στο χώρο των αντικειμενικών συναρτήσεων \mathbb{R}^p , ορίζεται το σημείο εκείνο που έχει ως συντεταγμένες τα άριστα των αντικειμενικών συναρτήσεων όπως αυτά προκύπτουν από τις μεμονωμένες αριστοποιήσεις. Συνήθως πρόκειται για σημείο που αντιστοιχεί σε μη εφικτή λύση (δεν υπάρχει εφικτή λύση που να αριστοποιεί συγχρόνως όλες τις αντικειμενικές συναρτήσεις). Συχνά αναφέρεται και ως σημείο ουτοπίας (utopia point). Το ιδεώδες σημείο χρησιμοποιείται συνήθως ως σημείο αναφοράς για να διερευνηθεί η προσέγγιση σ' αυτό των ικανών λύσεων.

Εναλλακτικές άριστες λύσεις (alternative optima): Όταν η άριστη τιμή z^* μιας αντικειμενικής συνάρτησης προκύπτει από διαφορετικές λύσεις $\mathbf{x}^{1*}, \mathbf{x}^{2*}, \dots, \mathbf{x}^{k*}$ τότε λέμε ότι η συγκεκριμένη αντικειμενική συνάρτηση έχει k εναλλακτικές άριστες λύσεις.

Λεξικογραφική αριστοποίηση (lexicographic optimization): Ο όρος αναφέρεται σε μια κατά στάδια αριστοποίηση προβλημάτων ΠΚΓΠ. Στο πρώτο στάδιο αριστοποιείται η πρώτη αντικειμενική συνάρτηση. Αν υπάρχουν εναλλακτικά άριστα αναζητείται μεταξύ αυτών αυτό που αριστοποιεί τη δεύτερη αντικειμενική συνάρτηση (δεύτερο στάδιο). Με σταθερή δηλαδή την άριστη τιμή της πρώτης αντικειμενικής συνάρτησης επιδιώκεται η αριστοποίηση της δεύτερης. Στη συνέχεια με σταθερές τις τιμές της πρώτης και της δεύτερης αντικειμενικής συνάρτησης επιδιώκεται η βελτιστοποίηση της τρίτης κ.ο.κ. μέχρι την τελευταία αντικειμενική συνάρτηση. Το αποτέλεσμα της λεξικογραφικής αριστοποίησης για κάποια αντικειμενική συνάρτηση είναι η λύση εκείνη που αριστοποιεί τη συγκεκριμένη αντικειμενική συνάρτηση και παρουσιάζει τις καλύτερες δυνατές τιμές για τις υπόλοιπες αντικειμενικές συναρτήσεις με κάποια σειρά προτεραιότητας. Πρέπει να σημειωθεί ότι αν δεν υπάρχουν εναλλακτικά άριστες λύσεις για μία

αντικειμενική συνάρτηση τότε το αποτέλεσμα της λεξικογραφικής αριστοποίησης ταυτίζεται με αυτό της απλής αριστοποίησης.

Πίνακας πληρωμών ή πίνακας τιμών (payoff table): Είναι ο τετραγωνικός πίνακας που αποτελείται από τόσες γραμμές και στήλες όσες και οι αντικειμενικές συναρτήσεις του προβλήματος ΠΚΓΠ. Σε κάθε στήλη αριστοποιείται μεμονωμένα μια αντικειμενική συνάρτηση και οι τιμές των αντικειμενικών συναρτήσεων που προκύπτουν (τόσο αυτής που αριστοποιήθηκε όσο και των υπολοίπων) τοποθετούνται στις σειρές του πίνακα. Στη διαγώνιο του πίνακα τιμών μπορούμε να διαβάσουμε τις συντεταγμένες του ιδεώδους σημείου.

Συντελεστής αντιστάθμισης (trade off): Με τον όρο αντιστάθμιση κριτηρίων ορίζεται το κατά πόσο πρέπει να χειροτερεύσει ένα κριτήριο ώστε να βελτιωθεί κάποιο άλλο. Μεταξύ δύο ικανών λύσεων, ο συντελεστής αντιστάθμισης μας δείχνει πόσο πρέπει να θυσιάσουμε από ένα κριτήριο προκειμένου να κερδίσουμε σε κάποιο άλλο και να μεταβούμε από την μία ικανή λύση στην άλλη.

Σταθμισμένο άθροισμα αντικειμενικών συναρτήσεων (weighted sum of objective functions): Με τον όρο αυτό εννοείται η σύνθεση των αντικειμενικών συναρτήσεων, μέσω του σταθμισμένου αθροίσματός τους, ώστε να δημιουργηθεί μία γενικευμένη αντικειμενική συνάρτηση. Το αντίστοιχο πρόβλημα σταθμισμένου αθροίσματος του προβλήματος ΠΚΓΠ που περιγράφεται στην (2.1) ορίζεται ως εξής:

$$\max \{ \lambda^T c^i x, i:=1..p, \mid x \in S \} \quad (2.8)$$

Πρόκειται δηλαδή για πρόβλημα ΓΠ, όπου οι *συντελεστές στάθμισης* λ_i (ή απλά «βάρη») των αντικειμενικών συναρτήσεων παίρνουν τιμές στο χώρο των συντελεστών στάθμισης, Λ που ορίζεται ως εξής:

$$\lambda_i \in \Lambda = \{ \lambda_i \in R^p \mid \lambda_i \in (0,1), \sum \lambda_i = 1 \} \quad (2.9)$$

Πρόβλημα min-max (min-max problem): Στον ΠΚΓΠ με τον όρο min-max πρόβλημα εννοείται το πρόβλημα της ελαχιστοποίησης της μέγιστης απόκλισης από κάποιο σημείο-στόχο. Στον ΠΚΓΠ το σημείο-στόχος ανήκει στον χώρο των p διαστάσεων (R^p , όπου p ο αριθμός των αντικειμενικών συναρτήσεων). Η ελαχιστοποίηση της μέγιστης απόκλισης από το σημείο-στόχο, σημαίνει την εύρεση της εφικτής λύσης που αντιστοιχεί στο σημείο εκείνο του R^p , του οποίου η μεγαλύτερη απόκλιση (ως προς τις p διαστάσεις) από το σημείο-στόχο αποκτά την ελάχιστη τιμή της. Το σημείο-στόχος μπορεί να είναι το ιδεώδες σημείο του προβλήματος ΠΚΓΠ.

Οικογένεια L_p αποστάσεων (Family of L_p -metrics): Για να εξαχθούν συμπεράσματα για το πόσο απέχουν μεταξύ τους οι διάφορες λύσεις στο χώρο R^k των αντικειμενικών συναρτήσεων, πρέπει να ορισθούν οι αποστάσεις μεταξύ σημείων (διανυσμάτων) στο χώρο R^k . Αυτό γίνεται μέσω των L_p αποστάσεων. Για δύο διανύσματα $x=(x_1, x_2, \dots, x_k)$ και $y=(y_1, y_2, \dots, y_k)$ ορίζεται:

$$\|\mathbf{x} - \mathbf{y}\|_p = \left[\sum_{i=1}^k |x_i - y_i|^p \right]^{1/p} \quad (2.10)$$

Συνήθως χρησιμοποιούνται $p=1$ (Γραμμική απόσταση), $p=2$ (Ευκλείδεια απόσταση) και $p=\infty$. Όσο μεγαλύτερο είναι το p τόσο μεγαλύτερη βαρύτητα δίδεται στις μεγαλύτερες αποκλίσεις. Για $p=\infty$ η απόσταση ορίζεται ως εξής:

$$\|\mathbf{x} - \mathbf{y}\|_\infty = \max\{|x_1 - y_1|, |x_2 - y_2|, \dots, |x_k - y_k|\} \quad (2.11)$$

και ονομάζεται απόσταση Tchebycheff. Η απόσταση Tchebycheff χρησιμοποιείται ευρέως στον ΠΚΓΠ διότι δεν καταστρέφει τη γραμμικότητα του προβλήματος. Η επιδίωξη της ελαχιστοποίησης της απόστασης Tchebycheff μιας λύσης από κάποιο σημείο-στόχο είναι ουσιαστικά η επίλυση του προβλήματος min-max.

2.5 Ανάλυση Παλινδρόμησης (Regression Analysis)

Με την ανάλυση παλινδρόμησης (regression analysis) εξετάζουμε τη σχέση μεταξύ δύο ή περισσότερων μεταβλητών με σκοπό την πρόβλεψη των τιμών της μιας, μέσω των τιμών της άλλης (ή των άλλων). Σε κάθε πρόβλημα παλινδρόμησης διακρίνουμε δύο είδη μεταβλητών: τις ανεξάρτητες ή ελεγχόμενες ή επεξηγηματικές (independent, predictor, casual, input, explanatory variables) και τις εξαρτημένες ή απόκρισης (dependent, response variables). Σε πειραματικές έρευνες, ανεξάρτητη μεταβλητή X είναι εκείνη την οποία μπορούμε να ελέγξουμε, δηλαδή, να καθορίσουμε τις τιμές της (π.χ. το ύψος της διαφημιστικής δαπάνης ενός προϊόντος, ο αριθμός των λειτουργούντων ταμείων σε ένα υποκατάστημα τραπεζής, η ποσότητα λιπάσματος, η θερμοκρασία επεξεργασίας ενός προϊόντος). Εξαρτημένη μεταβλητή Y είναι εκείνη στην οποία αντανακλάται το αποτέλεσμα των μεταβολών στις ανεξάρτητες μεταβλητές (π.χ. η ζήτηση ενός προϊόντος, ο χρόνος αναμονής των πελατών ενός υποκαταστήματος τραπεζής, η απόδοση μιας καλλιέργειας, η αντοχή ενός υλικού).

Σε μη πειραματικές έρευνες (δειγματοληψίες) η διάκριση μεταξύ ανεξάρτητων και εξαρτημένων μεταβλητών δεν είναι πάντοτε σαφής γιατί καμία μεταβλητή δεν είναι ελεγχόμενη αλλά όλες είναι τυχαίες (π.χ. το ύψος και το βάρος των φοιτητών, οι ώρες μελέτης των φοιτητών ενός πανεπιστημιακού τμήματος και η απόδοση τους σε ένα τεστ, οι εβδομάδες εμπειρίας ενός εργάτη σε μια επιχείρηση και ο αριθμός των ελαττωματικών προϊόντων που παράγει, η κατάταξη δέκα προϊόντων από έναν κριτή και η κατάταξη των ιδίων προϊόντων από έναν άλλο κριτή, ο αριθμός των πωλήσεων μουσικών CD σε μια περιοχή και ο αριθμός των νέων στην ίδια περιοχή). Μερικά παραδείγματα όπως όταν θέλει ένας μεσίτης να συσχετίσει τη τιμή πώλησης ενός σπιτιού με συγκεκριμένα χαρακτηριστικά του κτιρίου και τους φόρους για το συγκεκριμένο

κτίριο, όταν θέλουμε να εξετάσουμε τη κατανάλωση τσιγάρων σε σχέση με την ηλικία, τη μόρφωση, το εισόδημα και τη τιμή τους. Στο παράδειγμα εκτίμησης της τιμής των ακινήτων η μεταβλητή απόκρισης είναι η τιμή του σπιτιού και η μεταβλητή πρόβλεψης είναι τα χαρακτηριστικά του κτιρίου και οι φόροι που καταβλήθηκαν για το κτίριο. Στο παράδειγμα κατανάλωσης τσιγάρων η μεταβλητή απόκρισης είναι η κατανάλωση τσιγάρων (με βάση τον αριθμό των πακέτων των τσιγάρων που πωλούνται σε ένα κράτος σε κατά κεφαλήν βάση κατά τη διάρκεια ενός έτους) και οι μεταβλητές πρόβλεψης είναι οι διάφορες κοινωνικοοικονομικές και δημογραφικές μεταβλητές.

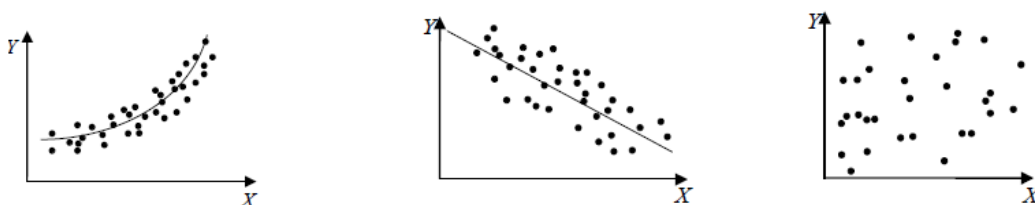
Η ανάλυση παλινδρόμησης απασχολεί το τομέα των μαθηματικών και ειδικότερα της στατιστικής εδώ και διακόσια χρόνια περίπου. Η παλιότερη μορφή της γραμμική παλινδρόμησης ήταν η μέθοδος των ελαχίστων τετραγώνων, η οποία δημοσιεύθηκε από τον Legendre το 1805 και από το Gauss το 1809. Ο όρος “ελάχιστα τετράγωνα” πρόεκυψε από τον Legendre. Ο Legendre και ο Gauss εφάρμοσαν αυτή τη μέθοδο για να προσδιορίσουν από τις αστρονομικές προβλέψεις τις τροχιές των σωμάτων σε σχέση με τον ήλιο. Ο Euler εργάστηκε στο ίδιο πρόβλημα το 1748 χωρίς όμως να έχει επιτυχία. Ο Gauss δημοσίευσε μια πιο βελτιωμένη θεωρία για τη μέθοδο των ελαχίστων τετραγώνων το 1821, συμπεριλαμβάνοντας μια έκδοση η οποία είναι πολύ γνωστή σήμερα ως το θεώρημα του Gauss-Markov. Το θεώρημα αυτό, αποτελεί θεμελιώδες θεώρημα στο τομέα γραμμικών μοντέλων.

Ας θεωρήσουμε δύο μεταβλητές X , Y . Αν οι μεταβλητές αυτές συνδέονται με μια σχέση της μορφής $Y = f(X)$ μέσω της οποίας για κάθε τιμή της X μπορούμε να προβλέψουμε ακριβώς την τιμή της Y , δηλαδή, αν οι τιμές της Y δεν υπόκεινται σε σφάλματα, τότε λέμε ότι οι δύο μεταβλητές συνδέονται με τη συναρτησιακή-προσδιοριστική (deterministic) σχέση $Y = f(X)$. Για παράδειγμα, το ρεύμα που καταναλώνει μια οικογένεια σε ένα δίμηνο και το ποσό που πληρώνει για την κατανάλωση αυτή συνδέονται με συναρτησιακή-προσδιοριστική σχέση. Επίσης, το ποσό που καταθέτει κάποιος στο Ταμειυτήριο και ο τόκος που παίρνει για το ποσό αυτό, συνδέονται με συναρτησιακή-προσδιοριστική σχέση. Σε αυτές τις περιπτώσεις τα σημεία του διαγράμματος διασποράς βρίσκονται όλα πάνω στην καμπύλη που έχει εξίσωση $Y = f(X)$ και όσες φορές και αν επαναλάβουμε το πείραμα θέτοντας το X στο ίδιο επίπεδο $X = x_i$, θα παίρνουμε πάντα την ίδια τιμή για το Y .

Οι μη προσδιοριστικές σχέσεις μεταξύ μεταβλητών ονομάζονται στοχαστικές – στατιστικές (stochastic, probabilistic) σχέσεις. Στην περίπτωση αυτή, αν επαναλάβουμε το πείραμα πολλές φορές θέτοντας το X στο ίδιο επίπεδο $X = x_i$ τότε στην τιμή x_i της X δεν αντιστοιχεί μια μόνο τιμή y_i της Y αλλά, γενικά, αντιστοιχεί ένα πλήθος διαφορετικών τιμών της Y . Για παράδειγμα, αν X είναι η τιμή ενός προϊόντος και Y είναι η ζήτησή του, η Y βρίσκεται σε στοχαστική σχέση-εξάρτηση από

τη X , γιατί η ζήτηση ενός προϊόντος επηρεάζεται και από άλλους παράγοντες όπως είναι το ύψος του εισοδήματος των καταναλωτών, οι τιμές ομοειδών προϊόντων, οι καταναλωτικές συνήθειες, κ.ά.

Σε μια στοχαστική σχέση το διάγραμμα διασποράς είναι, γενικά, ένα νέφος σημείων το οποίο πολλές φορές καθορίζει μια ιδεατή γραμμή η οποία δίνει μια πρώτη εικόνα της σχέσης που συνδέει τις δύο μεταβλητές. Η σχέση μάλιστα μεταξύ των δύο μεταβλητών είναι τόσο περισσότερο ισχυρή όσο πιο κοντά στην ιδεατή γραμμή βρίσκονται τα σημεία του διαγράμματος διασποράς. Στο πρώτο από τα παρακάτω σχήματα έχουμε το διάγραμμα διασποράς μιας ισχυρής σχέσης στην οποία όταν αυξάνουν οι τιμές της X αυξάνουν γενικά και οι τιμές της Y , ενώ στο δεύτερο σχήμα έχουμε μια λιγότερο ισχυρή σχέση στην οποία όταν αυξάνουν οι τιμές της X ελαττώνονται γενικά και οι τιμές της Y .



Σχήμα 2.2: Διαγράμματα διασποράς

Τέλος, στην περίπτωση του τρίτου σχήματος δε φαίνεται να υπάρχει κάποια σχέση μεταξύ των X και Y . Γενικά, δύο μεταβλητές που συνδέονται είτε με συναρτησιακή-προσδιοριστική σχέση είτε με στοχαστική σχέση λέγονται «εξαρτημένες». Αν υπάρχει εξάρτηση μεταξύ δύο μεταβλητών, τότε μπορούμε τη μια από αυτές να τη χαρακτηρίσουμε ως «αιτία» και την άλλη ως «αποτέλεσμα». Αυτό όμως, μόνο στην περίπτωση που η εξάρτηση οφείλεται σε σχέση αιτιότητας των δύο μεταβλητών και όχι σε μια απλή συμμεταβολή η οποία μπορεί να οφείλεται σε εξάρτηση των δύο μεταβλητών από μια τρίτη μεταβλητή. Αν, για παράδειγμα, X είναι το ετήσιο εισόδημα μιας οικογένειας και Y , Z είναι τα ποσά που ξοδεύει η οικογένεια αυτή σε ένα έτος για κρέας και για αγορά λογοτεχνικών βιβλίων, τότε: αν διαπιστώσουμε σε ένα σύνολο οικογενειών σχέση μεταξύ των X και Y (ή μεταξύ των X και Z) δεχόμαστε ότι υπάρχει εξάρτηση μεταξύ των δύο μεταβλητών και τότε μπορούμε να χαρακτηρίσουμε τη X ως «αιτία» και την Y (ή τη Z) ως «αποτέλεσμα». Αν όμως διαπιστωθεί σχέση μεταξύ των Y και Z (που είναι πολύ πιθανό, αφού και οι δύο μεταβάλλονται με το ετήσιο εισόδημα X) ασφαλώς θα πρόκειται για «νόθα» εξάρτηση.

Για να περιγράψουμε τη στοχαστική εξάρτηση δύο μεταβλητών X και Y προσπαθούμε να βρούμε, όπως και στην προσδιοριστική εξάρτηση, μια σχέση μεταξύ των X και Y η οποία όμως τώρα δε θα δίνει ακριβή αλλά προσεγγιστική μόνο εικόνα της εξάρτησης των X και Y και

τα σημεία του διαγράμματος διασποράς των X και Y δε θα βρίσκονται πάνω, αλλά, γύρω από μια καμπύλη.

Θεωρούμε Y τη μεταβλητή απόκρισης και τη μεταβλητή πρόβλεψης X_1, X_2, \dots, X_p όπου p είναι ο αριθμός που δηλώνει τον αριθμό των μεταβλητών πρόβλεψης. Η σχέση μεταξύ των Y και των X_1, X_2, \dots, X_p μεταβλητών μπορεί να προσδιοριστεί από το μοντέλο παλινδρόμησης

$$Y = f (X_1, X_2, \dots, X_p) + \varepsilon , \quad (2.12)$$

Όπου ε είναι το τυχαίο σφάλμα το οποίο αντιπροσωπεύει την απόκλιση από τη προσεγγιστική τιμή. Ουσιαστικά δηλώνει την αποτυχία της συνάρτησης αυτής ώστε το αποτέλεσμα αυτής να είναι κοντά στην πραγματική τιμή. Η συνάρτηση αυτή περιγράφει τη σχέση μεταξύ Y και των X_1, X_2, \dots, X_p μεταβλητών.

Οι μεταβλητές πρόβλεψης ή αλλιώς επεξηγηματικές έχουν και άλλες ονομασίες όπως ανεξάρτητες μεταβλητές, συμπεαράγοντες, παλινδρομούσα μεταβλητή, παράγοντες και φορείς. Η ονομασία ανεξάρτητη μεταβλητή, παρ'όλο που χρησιμοποιείται συχνά, δεν απικονίζει ακριβώς την πραγματικότητα, επειδή στη πράξη οι μεταβλητές πρόβλεψης είναι σπάνια ανεξάρτητες μεταξύ τους.

Υπάρχουν τρία είδη παλινδρόμησης. Η **πρώτη είναι η απλή γραμμική παλινδρόμηση**. Αυτό το είδος παλινδρόμησης χρησιμοποιείται για να βρούμε τη σχέση μεταξύ δύο μεταβλητών, οι οποίες μεταβάλλονται γραμμικά. Μια από αυτές είναι η εξαρτημένη μεταβλητή y και η ανεξάρτητη μεταβλητή x . Για παράδειγμα, η απλή γραμμική παλινδρόμηση μπορεί να προβλέψει τη σχέση μεταξύ της μουικής δύναμης (y) και της μουικής μάζα σώματος (x). Συνήθως η συνάρτηση του απλού γραμμικού μοντέλου έχει την ακόλουθη μορφή

$$y = \beta_0 + \beta_1 x + \varepsilon \quad (2.13)$$

όπου y είναι η εξαρτημένη μεταβλητή, β_0 είναι το σημείο τομής με τον άξονα των x , β_1 είναι η κλίση της ευθείας ή κλίση της καμπύλης της παλινδρόμησης, x είναι η ανεξάρτητη μεταβλητή και ε είναι το τυχαίο σφάλμα. Συνήθως θεωρείται ότι το σφάλμα ε ακολουθεί κανονική κατανομή με $E(\varepsilon)=0$ και σταθερή διακύμανση $Var(\varepsilon) = \sigma^2$ στην απλή γραμμική παλινδρόμηση.

Το **δεύτερο είδος παλινδρόμησης είναι η πολλαπλή γραμμική παλινδρόμηση**, η οποία είναι μια γραμμική παλινδρόμηση με μια εξαρτημένη μεταβλητή και περισσότερες από μια ανεξάρτητες

μεταβλητές. Η πολλαπλή γραμμική παλινδρόμηση υποθέτει ότι η μεταβλητή απόκρισης είναι μια γραμμική συνάρτηση των παραμέτρων του μοντέλου και υπάρχουν περισσότερες από μια ανεξάρτητες μεταβλητές στο μοντέλο. Η γενική μορφή της πολλαπλής γραμμικής παλινδρόμησης έχει ως εξής:

$$y = \beta_0 + \beta_1 X_1 + \dots + \beta_p x_p + \varepsilon \quad (2.14)$$

όπου y είναι η εξαρτημένη μεταβλητή $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ είναι συντελεστές παλινδρόμησης και x_1, x_2, \dots, x_n είναι ανεξάρτητες μεταβλητές του μοντέλου. Στη κλασική παλινδρόμηση συνήθως θεωρείται ότι ο όρος σφάλμα ακολουθεί κανονική κατανομή με $E(\varepsilon) = 0$ και η τυπική απόκλιση είναι $\text{Var}(\varepsilon) = \sigma^2$.

Με την απλή γραμμική παλινδρόμηση διερευνούμε τη γραμμική σχέση μεταξύ μιας εξαρτημένης μεταβλητής και μιας ανεξάρτητης μεταβλητής, ενώ με τη πολλαπλή γραμμική παλινδρόμηση εστιάζουμε στη γραμμική σχέση μεταξύ μιας εξαρτημένης μεταβλητής σε περισσότερες από μια ανεξάρτητες μεταβλητές. Η πολλαπλή γραμμική παλινδρόμηση περιλαμβάνει περισσότερα παραμέτρους από την απλή γραμμική παλινδρόμηση, όπως τη συγγραμμικότητα, την ετεροσκεδάση, τη γραφική απεικόνιση της διάγνωσης παλινδρόμησης, την ανίχνευση παλινδρόμησης σε ακραίες τιμές και στις παρατηρήσεις που επηρεάζουν την παλινδρόμηση.

Το **τρίτο είδος παλινδρόμησης είναι η μη γραμμική παλινδρόμηση**, η οποία υποθέτει ότι η σχέση μεταξύ της εξαρτημένης μεταβλητής και της ανεξάρτητης μεταβλητής είναι μη γραμμική στις παραμέτρους παλινδρόμησης. Ένα παράδειγμα μη γραμμικής παλινδρόμησης είναι το ακολουθο

$$y = \frac{\alpha}{1+e^{\beta t}} + \varepsilon, \quad (2.15)$$

όπου y είναι η ανάπτυξη ενός συγκεκριμένου οργανισμού συναρτήσει του χρόνου t , α και β είναι οι παράμετροι του μοντέλου, και ε είναι το τυχαίο σφάλμα. Το μοντέλο μη γραμμικής παλινδρόμησης είναι πιο περίπλοκο από το γραμμικό μοντέλο παλινδρόμησης όσον αφορά την εκτίμηση των παραμέτρων του μοντέλου, επιλογή μοντέλου, μοντέλο πρόβλεψης, επιλογή μεταβλητών, την ανίχνευση της ακραίας τιμής, ή το προσδιορισμό των καθοριστικών παρατηρήσεων (Julian J. Faraway, 2002).

2.5.1 Στόχοι της Ανάλυσης Παλινδρόμησης

Η ανάλυση παλινδρόμησης είναι μια από τις πιο εύχρηστες στατιστικές μεθόδους στη πράξη. Μπορεί να εφαρμοστεί σε πολλούς επιστημονικούς τομείς συμπεριλαμβανομένου της ιατρικής, βιολογίας, γεωργίας, οικονομικών, μηχανικής, κοινωνιολογίας, γεωλογίας, κ.ά. Ο σκοπός της ανάλυσης παλινδρόμησης έχει τρεις στόχους:

- 1) Προσδιορισμός της σχέσης μεταξύ της μεταβλητής απόκρισης y και των μεταβλητών παλινδρόμησης x_1, x_2, \dots, x_n .
- 2) Η πρόβλεψη της μεταβλητής y βασίζεται σε ένα σύνολο τιμών των x_1, x_2, \dots, x_n .
- 3) Διαλέγει τις μεταβλητές x_1, x_2, \dots, x_n για να προσδιορίσει ποιες από αυτές είναι πιο σημαντικές από τις υπόλοιπες ώστε να αντιπροσωπεύει τη μεταβλητή απόκρισης y όσο το δυνατόν πιο αποτελεσματικά και με τη περισσότερη ακρίβεια (Julian J. Faraway, 2002)

Πότε χρησιμοποιούμε την ανάλυση παλινδρόμησης

Με την ανάλυση παλινδρόμησης μπορούμε να επιτύχουμε πολλούς σημαντικούς στόχους.

1. Πρόβλεψη μελλοντικών παρατηρήσεων
2. Εκτίμηση των μεταβολών των, ή τη σχέση μεταξύ των, μεταβλητών απόκρισης.
3. Μια γενική περιγραφή της δομής των δεδομένων. (Samprit Chatterjee, Ali S.Hadi, 2006)

2.5.2 Τα βήματα για την ανάλυση παλινδρόμησης

- Ορισμός του προβλήματος
- Επιλογή των δυναμικών σχετικών μεταβλητών
- Συλλογή δεδομένων
- Προδιαγραφή του μοντέλου
- Επιλογή της μεθόδου προσαρμογής
- Προσαρμογή Μοντέλου
- Επικύρωση και κριτική μοντέλου

Χρησιμοποίηση του επιλεγμένου μοντέλου για τη λύση του προβλήματος που μας έχει τεθεί . (Samprit Chatterjee, Ali S.Hadi,2006)

2.6 Απλή Γραμμική Παλινδρόμηση (Simple Linear Regression)

Ο όρος παλινδρόμηση και οι μέθοδοι για την διερεύνηση της σχέσης μεταξύ των δύο μεταβλητών χρονολογείται ότι ξεκίνησε γύρω στα 100 χρόνια πριν. Η πρώτη αναφορά έγινε από τον Francis Galton το 1908, το διάσημο Βρετανό βιολόγο, όταν ασχολήθηκε με τη μελέτη της κληρονομικότητας. Μια από τις παρατηρήσεις του ήταν ότι τα παιδιά με ψηλούς γονείς, είναι ψηλότερα από το μέσο όρο, αλλά όχι τόσο ψηλά όσο οι γονείς τους. Αυτή η παλινδρόμηση “προς τη μετριότητα” έδωσε σ’ αυτές τις στατιστικές μεθόδους το όνομα τους. Ο όρος παλινδρόμηση και η εξέλιξη του, περιγράφει κυρίως τη στατιστική σχέση μεταξύ των μεταβλητών.

Μια μέθοδος που χρησιμοποιείται για την περιγραφή της στοχαστικής εξάρτησης δύο μεταβλητών είναι η μέθοδος των ελαχίστων τετραγώνων και αυτή θα εφαρμόσουμε στη συνέχεια για να μελετήσουμε την πιο απλή μορφή στοχαστικής εξάρτησης, τη γραμμική.

1. **Μέθοδος Ελαχίστων Τετραγώνων (Method of Least Squares):** με την μέθοδο αυτή προσδιορίζονται οι συντελεστές α και β από τους τύπους:

$$\hat{\beta} = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \quad (2.16)$$

$$\text{και } \hat{\alpha} = \bar{y} - \hat{\beta} \cdot \bar{x} \quad (2.17)$$

η ευθεία ελαχίστων τετραγώνων θα είναι η :

$$\hat{y} = \hat{\alpha} + \hat{\beta} \cdot x \quad (2.18)$$

Το $\hat{\alpha}$ είναι η τεταγμένη του σημείου στο οποίο η ευθεία τέμνει τον άξονα $y'y$ ενώ το $\hat{\beta}$, που είναι ο συντελεστής διεύθυνσης της ευθείας, εκφράζει την μεταβολή της μεταβλητής Y όταν η μεταβλητή X μεταβληθεί κατά μια μονάδα.

2. **Συντελεστής Γραμμικής Συσχέτισης (Linear Correlation Coefficient):** είναι το μέτρο που εκφράζει την συγκέντρωση των σημείων ενός διαγράμματος διασποράς γύρω από την ευθεία παλινδρόμησης.

Αν X και Y δύο μεταβλητές μεγέθους n τότε ο συντελεστής γραμμικής συσχέτισης είναι ο εξής:

$$r(X,Y) = r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (2.19)$$

Εάν οι μέσες τιμές δεν είναι ακέραιοι αριθμοί, τότε ο συντελεστής γραμμικής συσχέτισης r δίνεται από τον τύπο:

$$r = \frac{n \sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i) \cdot (\sum_{i=1}^n y_i)}{\sqrt{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \cdot \sqrt{n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2}} \quad (2.20)$$

Ιδιότητες του r :

- i. Εάν $0 < r < 1$ τότε οι X και Y είναι θετικά γραμμικά συσχετισμένες.
- ii. Εάν $-1 < r < 0$ τότε οι X και Y είναι αρνητικά γραμμικά συσχετισμένες.
- iii. Εάν $r = 1$ τότε έχουμε τέλεια θετική γραμμική συσχέτιση και όλα τα σημεία βρίσκονται πάνω στην ευθεία $y = \alpha + \beta \cdot x$ και $\beta > 0$. αντίστοιχα αν $r = -1$ και $\beta < 0$.
- iv. Εάν $r = 0$ τότε δεν υπάρχει γραμμική συσχέτιση μεταξύ των μεταβλητών X και Y . Οπότε λέμε ότι είναι γραμμικά ασυσχέτιστες.

Το εργαλείο ανάλυσης "Συσχέτιση" μετρά τη σχέση μεταξύ δύο συνόλων δεδομένων, τα οποία έχουν κλιμάκωση, ώστε να είναι ανεξάρτητα από μονάδα μετρήσεως. Ο υπολογισμός της συσχέτισης πληθυσμού επιστρέφει τη συνδιακύμανση των δύο συνόλων δεδομένων, διαιρούμενη δια του γινομένου των τυπικών τους αποκλίσεων, βάσει των ακόλουθων τύπων.

Μπορούμε να χρησιμοποιήσουμε το εργαλείο ανάλυσης "Συσχέτιση" για να καθορίσουμε κατά πόσο δύο περιοχές δεδομένων

μεταβάλλονται μαζί — δηλαδή, κατά πόσο μεγάλες τιμές του ενός συνόλου σχετίζονται με μεγάλες τιμές του άλλου (θετική συσχέτιση), κατά πόσο μικρές τιμές του ενός συνόλου σχετίζονται με μεγάλες τιμές του άλλου (αρνητική συσχέτιση) ή κατά πόσο οι τιμές και των δύο συνόλων είναι άσχετες μεταξύ τους (σχεδόν μηδενική συσχέτιση).

3. **Συνδιακύμανση (Covariance of the two variables)**: αποτελεί ένα μέτρο της σχέσης μεταξύ δύο περιοχών δεδομένων. Το εργαλείο ανάλυσης "Συνδιακύμανση" αποδίδει το μέσο όρο του γινομένου των αποκλίσεων των σημείων δεδομένων από τις αντίστοιχες μέσες τιμές τους, βάσει του παρακάτω τύπου.

$$\text{cov}(X, Y) = \frac{1}{v} \sum_{i=1}^v (x_i - \bar{x})(y_i - \bar{y}) \quad (2.21)$$

Μπορεί να χρησιμοποιηθεί το εργαλείο "Συνδιακύμανση" για να καθορισθεί κατά πόσο δύο περιοχές δεδομένων μεταβάλλονται μαζί — δηλαδή, κατά πόσο μεγάλες τιμές του ενός συνόλου σχετίζονται με μεγάλες τιμές του άλλου (θετική συνδιακύμανση), κατά πόσο μικρές τιμές του ενός συνόλου σχετίζονται με μεγάλες τιμές του άλλου (αρνητική συνδιακύμανση) ή κατά πόσο οι τιμές και των δύο συνόλων είναι άσχετες μεταξύ τους (σχεδόν μηδενική συνδιακύμανση).

4. **Ανάλυση Διασποράς (Analysis of Variance) ANOVA**: Εξετάζει τη σχέση της εξαρτημένης μεταβλητής με την ανεξάρτητη, υπολογίζοντας στην ουσία το αν η μεταβλητότητα των τιμών της εξαρτημένης μεταβλητής Y εξηγείται από την ανεξάρτητη μεταβλητή X . Η ανάλυση της διασποράς για το απλό γραμμικό μοντέλο μπορεί να παρουσιαστεί ως εξής:

Πίνακας 2.1: Ανάλυση Διάσπορας

Πηγή Μεταβλητότητας	Βαθμοί Ελευθερίας	Άθροισμα Τετραγώνων	Μέσο Τετράγωνο	F-test
Παλινδρόμηση	1	$SSR = \sum_{i=1}^v (\hat{y}_i - \bar{y})^2$	$MSR = \frac{SSR}{1}$	$F = \frac{MSR}{MSE}$
Σφάλματα	$v - 2$	$SSE = \sum_{i=1}^v (y_i - \hat{y}_i)^2$	$MSE = S^2 = \frac{SSE}{v - 2}$	
Ολική	$v - 1$	$SST = \sum_{i=1}^v (y_i - \bar{y})^2$		

Σε κάθε άθροισμα τετραγώνων αντιστοιχούν ορισμένοι βαθμοί ελευθερίας, που ισοδυναμούν με το πλήθος των ανεξαρτήτων συναρτήσεων των y_i , οι οποίοι απαιτούνται για τον υπολογισμό του εν λόγω αθροίσματος.

Οπότε το SST έχει $n - 1$ βαθμούς ελευθερίας διότι:

$$\sum_{i=1}^n (y_i - \bar{y}) = 0 \quad (2.22)$$

Το SSR έχει έναν βαθμό ελευθερίας, διότι μπορεί να υπολογισθεί από μία συνάρτηση των y_i την $\hat{\beta}$ δοθέντος ότι:

$$SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2 = \hat{\beta}^2 \cdot \sum_{i=1}^n (x_i - \bar{x})^2 \quad (2.23)$$

Ενώ το SSE έχει $n - 2$ βαθμούς ελευθερίας διότι:

$$\sum_{i=1}^n (y_i - \hat{y}_i) = 0 \text{ και } \sum_{i=1}^n x_i \cdot (y_i - \hat{y}_i) = 0 \quad (2.24)$$

Μια θεμελιώδης ιδιότητα που χρησιμοποιούμε για τη απλή αλλά και τη πολλαπλή παλινδρόμηση είναι η ακόλουθη

$$SST = SSE + SSR \quad (2.25)$$

Τα εργαλεία ανάλυσης Anova προσφέρουν διαφορετικά είδη ανάλυση της διακύμανσης. Το εργαλείο που χρησιμοποιήσετε εξαρτάται από τον αριθμό των συντελεστών και τον αριθμό των δειγμάτων που έχουμε από τους πληθυσμούς τους οποίους θέλουμε να ελέγξουμε.

Ανάλυση διακύμανσης κατά ένα παράγοντα Αυτό το εργαλείο εκτελεί μια απλή ανάλυση διακύμανσης, επαληθεύοντας την υπόθεση ότι οι μέσες τιμές δύο ή περισσότερων δειγμάτων είναι ίσες (εφόσον λαμβάνονται από πληθυσμούς με την ίδια μέση τιμή). Η τεχνική αυτή επεκτείνεται στις δοκιμές δύο μέσων τιμών, όπως ο έλεγχος t.

Ανάλυση διακύμανσης δύο παραγόντων με αλληλεπίδραση . Αυτό το εργαλείο ανάλυσης εκτελεί μια παραλλαγή, δύο παραγόντων με αναπαραγωγή, της ανάλυσης διακύμανσης ενός παράγοντα, που περιλαμβάνει περισσότερα από ένα δείγματα για κάθε ομάδα δεδομένων.

Ανάλυση διακύμανσης δύο παραγόντων χωρίς αλληλεπίδραση . Αυτό το εργαλείο ανάλυσης εκτελεί μια ανάλυση διακύμανσης δύο παραγόντων, η οποία δεν περιλαμβάνει περισσότερες από μία δειγματοληψίες ανά ομάδα, κάνοντας δοκιμή της υπόθεσης ότι οι μέσες τιμές δύο ή περισσότερων δειγμάτων είναι ίσες (εφόσον λαμβάνονται από πληθυσμούς με την ίδια μέση τιμή). Η τεχνική αυτή επεκτείνεται σε δοκιμές για δύο μέσες τιμές, όπως ο έλεγχος t.

5. **Στατιστικοί Έλεγχοι F:** Οι στατιστικές συναρτήσεις SSR και SSE είναι ανεξάρτητες μεταξύ τους και τα μέσα τετράγωνα ακολουθούν την X^2 κατανομή με βαθμούς ελευθερίας τους αντίστοιχους των αθροισμάτων τετραγώνων. Επομένως η συνάρτηση F ακολουθεί την $F_{1,n-2}$ κατανομή και μπορεί να χρησιμοποιηθεί, για να ελέγξουμε την σημαντικότητα της παλινδρόμησης, δηλαδή της υπόθεσης $H_0 : \beta=0$ έναντι της $H_a : \beta \neq 0$. Ουσιαστικά ελέγχουμε την ισχύ της υπόθεσης, ότι τα δεδομένα μας μπορούν να περιγραφούν ικανοποιητικά από το γραμμικό μοντέλο.

6. **Συντελεστής Προσδιορισμού R^2 :** είναι το ηλίκο

$$R^2 = SSR/SST \rightarrow 0 \quad (2.26)$$

Με τον οποίο μπορούμε να ελέγξουμε την αξία του απλού γραμμικού μοντέλου, το οποίο προσαρμόζουμε στα δεδομένα. Ο συντελεστής προσδιορισμού εκφράζει το ποσοστό της μεταβλητότητας της μεταβλητής Y που εξηγείται από την μεταβλητή X. Όσο πιο κοντά βρίσκεται η τιμή του R^2 στην μονάδα, τόσο πιο ισχυρή γίνεται η γραμμική σχέση εξάρτησης των μεταβλητών Y και X.

2.7 Πολλαπλή Γραμμική Παλινδρόμηση (Multiple Linear Regression)

Στην ενότητα αυτή θα μελετήσουμε τη πολλαπλή γραμμική παλινδρόμηση. Ο γενικός σκοπός της πολλαπλής γραμμικής παλινδρόμησης είναι η αναζήτηση για τη γραμμική σχέση μεταξύ μιας εξαρτημένης μεταβλητής και πολλών ανεξάρτητων μεταβλητών. Πολλαπλή παλινδρόμηση επιτρέπει στους ερευνητές να εξετάσουν την επίδραση της πάνω από μία ανεξάρτητες μεταβλητές σχετικά με την απάντηση την ίδια στιγμή. Σε αρκετά όμως προβλήματα

η μεταβλητή απόκρισης Y μπορεί να θεωρηθεί ότι επηρεάζεται από περισσότερες από μια ερμηνευτικές μεταβλητές, έστω X_1, X_2, \dots, X_{p-1} . Μπορούμε και πάλι να χρησιμοποιήσουμε ένα γραμμικό μοντέλο για να διερευνήσουμε την εξάρτηση της Y από τις X_1, X_2, \dots, X_{p-1} . Το μοντέλο αυτό, το οποίο αποτελεί γενίκευση του μοντέλου $Y = b_0 + b_1 X_1 + \varepsilon$ που μελετήσαμε στην προηγούμενη ενότητα, θα έχει τη μορφή

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_{p-1} X_{p-1} + \varepsilon, \quad (2.27)$$

Για κάποιες παραμέτρους b_0, b_1, \dots, b_{p-1} . Για να διερευνηθεί η σχέση μεταξύ της Y και των X_1, X_2, \dots, X_{p-1} λαμβάνεται δείγμα μεγέθους n και για κάθε άτομο του δείγματος καταγράφονται οι τιμές των συγκεκριμένων μεταβλητών. Π.χ. για το i -άτομο του δείγματος καταγράφονται οι τιμές $(Y_i, X_{i1}, \dots, X_{i,p-1})$, $i=1, 2, \dots, n$. Εφαρμόζουμε λοιπόν το μοντέλο

$$Y_i = b_0 + b_1 X_{i1} + b_2 X_{i2} + \dots + b_{p-1} X_{i,p-1} + \varepsilon_i, \quad i=1, 2, \dots, n \quad (2.28)$$

Όπου τα σφάλματα $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ θεωρούνται ανεξάρτητες τιμές μεταβλητών από την $N(0,1)$ ενώ οι ερμηνευτικές μεταβλητές X_1, X_2, \dots, X_{p-1} , όπως και στο απλό γραμμικό μοντέλο, δεν θεωρούνται τυχαίες. Το παραπάνω μοντέλο γράφεται με τη βοήθεια πινάκων στην απλούστερη μορφή:

$$Y = Xb + \varepsilon \quad (2.29)$$

Όπου

$$Y = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \quad X = \begin{bmatrix} 1 & X_{11} & \dots & X_{1,p-1} \\ 1 & X_{21} & \dots & X_{2,p-1} \\ \vdots & \vdots & \dots & \vdots \\ 1 & X_{n1} & \dots & X_{n,p-1} \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.30)$$

Ο πίνακας X καλείται και πίνακας σχεδιασμού. Στο πολλαπλό μοντέλο θεωρούμε ότι τα σημεία δεν βρίσκονται «κοντά» σε μια ευθεία αλλά «κοντά» σε ένα (υπερ)επίπεδο στις p διαστάσεις.

Μέθοδοι Ελαχίστων Τετραγώνων

Όπως είπαμε και παραπάνω δεχόμαστε ότι το μοντέλο πολλαπλή γραμμικής παλινδρόμησης εκφράζεται από της παρακάτω εξίσωση, με τη μορφή μήτρας

$$Y = X b + \varepsilon, \quad (2.31)$$

Όπου

$$X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & \dots & X_{2k} \\ \dots & \dots & \vdots & \vdots \\ X_{n1} & X_{n2} & \dots & X_{nk} \end{bmatrix} \quad b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-1} \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.32)$$

Η εκτίμηση των ελαχίστων τετραγώνων μπορεί να γίνει μέσω της αρχής των ελαχίστων τετραγώνων:

$$b = \operatorname{argmin}_b [(y - Xb)'(y - Xb)], \quad (2.33)$$

όπου $b' = (b_0, b_1, \dots, b_{k-1})'$, ένα k -διάστατο διάνυσμα των εκτιμήσεων των συντελεστών παλινδρόμησης.

Εκτίμηση των παραμέτρων b_0, b_1, \dots, b_{p-1} και σ^2

Το τυχαίο διάνυσμα ε αποτελείται από n ανεξάρτητες $N(0, \sigma^2)$ τυχαίες μεταβλητές και επομένως θα έχει από κοινού $\sigma.p.p.$ $N(0, \sigma^2 I_n)$, δηλαδή ακολουθεί μία πολυδιάστατη κανονική κατανομή όπου I_n είναι ο μοναδιαίος πίνακας διάστασης n . Ως συνέπεια το τυχαίο διάνυσμα $Y = [Y_1, Y_2, \dots, Y_n]^T$ θα ακολουθεί και αυτό πολυδιάστατη κανονική $N(Xb, \sigma^2 I_n)$

Ερμηνεύοντας τη συνολική μεταβλητότητα του μοντέλου

Όπως ακριβώς και στο απλό γραμμικό μοντέλο, η δειγματική διασπορά των παρατηρήσεων Y_i αποδεικνύεται ότι χωρίζεται σε δύο αθροίσματα τα οποία συμβολίζονται και πάλι με SST, SSE και SSR αντίστοιχα. Το SST εκφράζει τη συνολική παρατηρούμενη μεταβλητότητα των Y_i , το SSR εκφράζει τη μεταβλητότητα των

προσαρμοσμένων τιμών ενώ το SSE εκφράζει τη μεταβλητότητα των Y_i σε σχέση με τις αντίστοιχες προσαρμοσμένες τιμές. Η μεταβλητότητα του SSR ερμηνεύεται από το μοντέλο ενώ του SSE όχι. Το ηλίκο (συντελεστής προσδιορισμού)

$$R^2 = \frac{SSR}{SST} = \frac{SST - SSE}{SST} \quad (2.34)$$

μπορεί να θεωρηθεί ότι εκφράζει το ποσοστό της μεταβλητότητας των Y_i που ερμηνεύεται από το μοντέλο. Είδαμε παραπάνω ότι,

$$\frac{SSE}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \sim \chi_{n-p}^2, \quad (2.35)$$

Επίσης, αν $b_1 = b_2 = \dots = b_{p-1} = 0$ τότε αποδεικνύεται ότι

$$\frac{SSE}{\sigma^2} \sim \chi_{p-1}^2 \quad \text{και} \quad \frac{SST}{\sigma^2} \sim \chi_{n-1}^2, \quad (2.36)$$

(διαφορετικά ακολουθούν κάποιες μη-κεντρικές κατανομές χι-τετράγωνο). Επομένως αν $b_1 = b_2 = \dots = b_{p-1} = 0$ τότε

$$F = \frac{\frac{SSR}{\sigma^2}}{\frac{p-1}{\frac{SSE}{\sigma^2}}} = \frac{\frac{SSR}{(p-1)}}{\frac{SSE}{(n-p)}} \sim F_{p-1, n-p} \quad (2.37)$$

(διότι SSR και SSE είναι ανεξάρτητες). Από το παραπάνω μπορούμε να κατασκευάσουμε έναν έλεγχο για την υπόθεση $H_0: b_1 = b_2 = \dots = b_{p-1} = 0$ (δηλ. ότι η Y δεν εξαρτάται από καμία από τις X_1, X_2, \dots, X_{p-1}). Θα απορρίπτεται η H_0 όταν η παραπάνω στατιστική συνάρτηση λαμβάνει (ασυνήθιστα υπό την H_0) μεγάλες τιμές, δηλαδή όταν (ε.σ. α)

$F = \frac{SSR/(p-1)}{SSE/(n-p)} > F_{p-1, n-p}(\alpha)$: άνω α -σημείο της κατανομής F με $p-1$ και $n-p$ β.ε. με αντίστοιχο p -value:

$$p\text{-value} = 1 - F_{F_{p-1, n-p}}\left(\frac{\frac{SSR}{p-1}}{\frac{SSE}{n-p}}\right) \quad (2.38)$$

(όπου $F_{F_{p-1, n-p}}$, είναι η σ.κ. της κατανομής $F_{p-1, n-p}$). Αντίθετα από το απλό, στο πολλαπλό μοντέλο, ο παραπάνω έλεγχος μέσω της F είναι διαφορετικός από τους έλεγχους που είδαμε στην προηγούμενη παράγραφο χρησιμοποιώντας τις στατιστικές συναρτήσεις T_i , $i = 1, 2, \dots, p-1$ (διότι οι έλεγχοι εκείνοι αφορούν τις υποθέσεις $H_0: b_1 = 0$, $H_0: b_2 = 0, \dots, H_0: b_{p-1} = 0$ ξεχωριστά). Όλες οι παραπάνω ποσότητες συνοψίζονται και πάλι στον πίνακα ανάλυσης διασποράς (ANOVA):

Πίνακας 2: Ανάλυση Διασποράς

Πηγή Μεταβλητότητας	Βαθμοί Ελευθερίας	Άθροισμα Τετραγώνων	Μέσο τετράγωνο	F-test
Παλινδρόμηση	$p-1$	$SSR = \sum (\hat{Y}_i - \bar{Y})^2$	$MSR = \frac{SSR}{p-1}$	$\frac{MSR}{MSE}$
Σφάλματα	$n - p$	$SSE = \sum (Y_i - \hat{Y}_i)^2$	$MSE = \frac{SSE}{n-p} = S^2$	
Ολική	$n - 1$	$SST = \sum (Y_i - \bar{Y})^2$		

Σύγκριση μοντέλων

Αρκετές φορές συμβαίνει να έχουμε καταγράψει τις τιμές αρκετών ανεξάρτητων μεταβλητών X_1, X_2, \dots, X_{p-1} και θέλουμε να εξετάσουμε ποιές από αυτές επηρεάζουν την μεταβλητή απόκρισης Y . Ένας τρόπος (που δεν είναι πάντα ο συντομότερος) είναι να εφαρμόσουμε όλα τα δυνατά μοντέλα και να επιλέξουμε αυτό που δίνει τα «καλύτερα» αποτελέσματα. Αν π.χ. έχουμε καταγράψει τρεις μεταβλητές X_1, X_2, X_3 τότε μπορούμε να εξετάσουμε τα μοντέλα:

$$Y = b_0 + b_1 X_1 + \varepsilon,$$

$$Y = b_0 + b_1 X_3 + \varepsilon,$$

$$Y = b_0 + b_1 X_1 + b_2 X_2 + \varepsilon,$$

$$Y = b_0 + b_1 X_2 + b_2 X_3 + \varepsilon,$$

$$Y = b_0 + b_1 X_2 + \varepsilon,$$

$$Y = b_0 + b_1 X_1 + b_2 X_3 + \varepsilon,$$

$$Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \varepsilon,$$

και να θεωρήσουμε καλύτερο αυτό που δίνει το μεγαλύτερο συντελεστή προσδιορισμού R^2 . Γρήγορα όμως διαπιστώνουμε ότι το μοντέλο με το μεγαλύτερο $R^2 = 1 - SSE/SST$ δεν είναι πάντοτε το καλύτερο. Αυτό συμβαίνει διότι όσο προσθέτουμε ανεξάρτητες μεταβλητές στο μοντέλο (όποιες και αν είναι αυτές) το R^2 αυξάνεται (ή παραμένει σταθερό). Πράγματι, όταν προσθέτουμε ανεξάρτητες μεταβλητές το SSE μειώνεται ή μένει σταθερό αφού

$$SSE = \min_{b_0, b_1, \dots, b_{k-1}} \sum_{i=1}^n (Y_i - b_0 - b_1 X_{i1} - \dots - b_{p-1} X_{i,p-1})^2 \quad (2.39)$$

(ενώ το SST παραμένει πάντοτε σταθερό) και η παραπάνω ελαχιστοποίηση γίνεται σε μεγαλύτερο χώρο (περισσότερα b_i). Π.χ. έστω ότι η Y επηρεάζεται από τις X_1, X_2 ενώ δεν επηρεάζεται καθόλου από την X_3 οπότε καλύτερο μοντέλο θα έπρεπε να είναι το

$Y = b_0 + b_1 X_1 + b_2 X_2 + \varepsilon$. Προσθέτοντας όμως την X_3 σε αυτό το μοντέλο προκύπτει το μοντέλο $Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \varepsilon$, το οποίο, σύμφωνα με τα παραπάνω, θα έχει μεγαλύτερο ή ίσο R^2 από το

$Y = b_0 + b_1 X_1 + b_2 X_2 + \varepsilon$. Επομένως το R^2 δεν δείχνει πάντοτε το «καλύτερο» μοντέλο. Αντί του R^2 προτείνεται η χρήση του «προσαρμοσμένου» R^2 (R^2 adjusted). Έτσι, καλύτερο θα θεωρείται το μοντέλο με το μεγαλύτερο

$$R^2 = 1 - \frac{SSE(n-p)}{SST/(n-1)} \quad (2.40)$$

Το προσαρμοσμένο R^2 «δείχνει» ως καλύτερο το μοντέλο που έχει το μικρότερο $S^2 = SSE/(n-p)$ (το SST είναι σταθερό σε όλα τα μοντέλα). Ο δείκτης αυτός δεν αυξάνεται πάντοτε όταν αυξάνονται οι ανεξάρτητες μεταβλητές. Αξίζει να αναφερθεί ότι για την εύρεση του καλύτερου μοντέλου έχουν προταθεί και άλλοι δείκτες (με διάφορες αιτιολογήσεις).

Πολυσυγγραμμικότητα

Στην πολλαπλή παλινδρόμηση είναι δυνατό κάποιες από τις ανεξάρτητες μεταβλητές X_1, X_2, \dots, X_{p-1} να είναι γραμμικά εξαρτημένες με συνέπεια ο πίνακας πληροφορίας $X^T X$ να μην αντιστρέφεται (η ορίζουσα του είναι 0). Αυτό το πρόβλημα είναι γνωστό ως πρόβλημα πολυσυγγραμμικότητας. Ένας απλός τρόπος αντιμετώπισης του είναι η αφαίρεση κάποιων ανεξάρτητων μεταβλητών από το μοντέλο (χάνοντας όμως «πληροφορία»). Ακόμη και όταν η ορίζουσα του $X^T X$ δεν είναι ακριβώς 0 αλλά «κοντά» στο 0 (ασθενής πολυσυγγραμμικότητα) παρουσιάζεται πρόβλημα. Σε αυτή την περίπτωση μπορεί να εμφανιστούν σφάλματα στρογγύλευσης κατά την αντιστροφή του $X^T X$ (με συνέπεια οι εκτιμήσεις που παίρνουμε να μην είναι αξιόπιστες). Αυτό συνήθως αντιμετωπίζεται ως ένα βαθμό από το ίδιο το πακέτο το οποίο πριν προβεί στους υπολογισμούς των εκτιμήσεων τυποποιεί όλες τις μεταβλητές (π.χ. Ωστε να παίρνουν τιμές στο $(-1,1)$) και κατά την εμφάνιση των αποτελεσμάτων τις επαναφέρει στην αρχική κλίμακα. Η ορίζουσα του $X^T X$ είναι κοντά στο 0 όταν υπάρχει ισχυρή συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών.

Μία ακόμη «παρενέργεια» της συγκεκριμένης κατάστασης είναι ότι μπορεί ορισμένες μεταβλητές να φαίνονται «σημαντικές» (με αντίστοιχο p-value στο t-τεστ κοντά στο 0) σε κάποιο μοντέλο, ενώ παύουν να είναι σημαντικές όταν στο μοντέλο προσθέσουμε και άλλες ανεξάρτητες μεταβλητές. Για παράδειγμα μπορεί στο μοντέλο $Y = b_0 + b_1 X_1 + b_2 X_2 + \varepsilon$, η X_2 να είναι σημαντική (απορρίπτουμε ότι $b_2 \neq 0$) ενώ στο μεγαλύτερο μοντέλο $Y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 + \varepsilon$, να μην είναι πια σημαντική (ενώ η X_3 που προσθέσαμε να είναι). Χονδρικά, αυτό μπορεί να συμβαίνει διότι η X_3 είναι αυτή που επηρεάζει την Y αλλά όταν εφαρμόζουμε το πρώτο μοντέλο (στο οποίο απουσιάζει η X_3), η X_2 φαίνεται σημαντική διότι «μοιάζει» στην X_3 . Όταν εντοπίσουμε ομάδα ή ομάδες από ισχυρά συσχετισμένες μεταβλητές (π.χ. από τον πίνακα συσχετίσεων των X_i) τότε θα πρέπει στο βέλτιστο μοντέλο να κρατήσουμε μία από κάθε ομάδα. Σε αρκετές περιπτώσεις αυτό δεν είναι εύκολο οπότε είναι ανάγκη να χρησιμοποιήσουμε άλλες μεθόδους (π.χ. principal components regression ή ridge regression).

2.8 Επιλογή Μεταβλητών (Variable Selection)

Στις ιατρικές μελέτες, οι αναλυτές χρησιμοποιούν ποσοτικές μεθόδους για να διαγνώσουν και να αξιολογήσουν το αποτέλεσμα για τη θεραπεία. Η παλινδρόμηση ελαχίστων τετραγώνων (Least Squares regression) είναι η μέθοδος που χρησιμοποιούν πιο συχνά, ωστόσο είναι πολύ ευαίσθητη στις ακραίες τιμές. Υπάρχουν πολύ τρόποι για να

αντιμετωπίσουμε αυτό το πρόβλημα. Για παράδειγμα, μπορεί κανείς να εντοπίσει τις ακραίες τιμές, να τις απορρίψει και εφαρμόσει το μοντέλο με τις υπόλοιπες μεταβλητές, ή κάποιος μπορεί να χρησιμοποιήσει μια πιο αξιόπιστη μέθοδο από αυτή των ελαχίστων τετραγώνων. Συνήθως η πρώτη εναλλακτική είναι μη αποδεκτή επειδή οι παρατηρήσεις είναι έγκυρες και συνεπώς δε πρέπει να απορρίπτεται (Subhash C. Narula et Al, 1999).

Η επιλογή μεταβλητών (Variable Selection) χρησιμοποιείται για την πρόβλεψη του “καλύτερου” υποσυνόλου ανεξάρτητων μεταβλητών. Με αυτή τη μέθοδο εξηγούμε τα δεδομένα με τον απλούστερο τρόπο - οι περιττές προβλέψεις αφαιρούνται. Η αρχή της μετριοπάθειας (Law of parsimony) ορίζει ότι μεταξύ των διαφόρων πιθανών εξηγήσεων για μια περίπτωση, η απλούστερη είναι και η καλύτερη. Η εφαρμογή της στην ανάλυση παλινδρόμησης αποδύκνυει ότι το μικρότερο μοντέλο που ταιριάζει με τα δεδομένα είναι και το καλύτερο. Επιπλέον, περιττοί προγνωστικοί παράγοντες θα προσθέσουν περισσότερα εμπόδια ώστε να προσδιορίσουμε άλλες ποσότητες οι οποίες μας ενδιαφέρουν. Με αποτέλεσμα οι βαθμοί ελευθερίας να αυξηθούν. Η συγγραμικότητα (collinearity) δημιουργείται όταν έχουμε πολλές μεταβλητές και προσπαθούμε να κάνουμε την ίδια δουλειά. Εάν το μοντέλο πρόκειται να χρησιμοποιηθεί για πρόβλεψη μεταβλητών, μπορούμε να ζώσουμε χρόνο ή/και λεφτά όταν δε λαμβάνουμε υπόψη μας τους περιττούς προγνωστικούς παράγοντες. Από την άλλη πλευρά, με τη μέθοδο αυτή καθίσταται δυνατό ο προσδιορισμός των ακραίων τιμών και των καθοριστικών σημείων (δηλαδή των σημείων που επηρεάζουν περισσότερο), ή τουλάχιστον να αποκλείσουμε κάποια προσωρινά. Επίσης αυτή η μέθοδος μας δίνει τη δυνατότητα να προσθέσουμε οποιοδήποτε μετασχηματισμό μεταβλητών ο οποίος φαίνεται να είναι κατάλληλος για τη συγκεκριμένη περίπτωση. Ένας αναλυτής παλινδρόμησης συνήθως δεν γνωρίζει ποια μέθοδο να χρησιμοποιήσει για να διαλέξει το καλύτερο υποσύνολο μεταβλητών πρόβλεψης χρησιμοποιώντας κάποια συγκεκριμένα κριτήρια.

Ιστορικά, όταν υπάρχουν πολλές μεταβλητές πρόβλεψης, ένα ή περισσότερα υποσύνολα με λιγότερες μεταβλητές πρόβλεψης παράγονται χρησιμοποιώντας τη μέθοδο της επιλογής του αναλυτή. Λαμβάνοντας δεδομένα της μορφής $\{y_i, x_{1i}, \dots, x_{ki}\}, i=1, \dots, n$, το υποσύνολο επιλογής μεταβλητών προκύπτει από την επιλογή ενός υποσυνόλου M του N (M υποσυνόλου του N) όπου $N=\{1, \dots, k\}$ είναι περιεχόμενο ενός συνόλου μεταβλητών πρόβλεψης $\{X_1, \dots, X_k\}$ έτσι ώστε κάποια μέτρα της ισχύος το επεξηγηματικού μοντέλου να μεγιστοποιείται (Julian J. Faraway, 2002).

Το κυριότερο κριτήριο της επιλογής υποσυνόλου είναι το κριτήριο της μετριοπάθειας (parsimony): εάν 3 μεταβλητές μπορούν να εξηγήσουν ή να ικανοποιήσουν τη μεταβλητή Y , γιατί να χρησιμοποιήσουμε 4; σύμφωνα με τις σημειώσεις του Mandel (1989).

Μερικούς από τους λόγους για τη χρήση μόνο ενός υποσυνόλου διαθέσιμων μεταβλητών πρόβλεψης (σύμφωνα με το Miler, 1984) είναι :

- Να εκτιμήσουμε ή να προβλέψουμε με μικρό κόστος με τη μείωση του αριθμού των μεταβλητών από τα δεδομένα που πρόκειται να συλλεχθούν
- Να προβλέψουμε με περισσότερη ακρίβεια καταργώντας ανούσιες μεταβλητές
- Να περιγράψουμε μια σειρά πολυμεταβλητών δεδομένων με μετριοτήτα (parsimony)
- Να εκτιμήσουμε τους συντελεστές παλινδρόμησης με μικρότερα σφάλματα (ειδικά όταν μερικοί από τους προγνωστικούς παράγοντες σχετίζονται πολύ μεταξύ τους),(Burak Eksioglu, Riza Demirer, Ismail Capar,2005).

2.8.1 Ιεραρχία Μοντέλων

Μερικά μοντέλα έχουν φυσική ιεραρχία. Για παράδειγμα, στα πολυωνυμικά μοντέλα, x^2 είναι όρος υψηλότερης τάξης από το x . Όταν επιλέγουμε μεταβλητές, είναι σημαντικό να λαμβάνουμε υπόψη μας την ιεραρχία. Οι όροι μικρότερης τάξης δε πρέπει να αφαιρούνται από το μοντέλο πριν από τον όρο υψηλότερης τάξης της ίδιας μεταβλητής. Υπάρχουν δύο κοινές καταστάσεις όπου αυτή η κατάσταση προκύπτει:

❖ *Πολυωνυμικά μοντέλα.* Θεωρούμε ένα μοντέλο

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon \quad (2.41)$$

Υποθέτουμε ότι το μοντέλο αυτό ταιριάζει στη περίπτωση μας και βρίσκουμε από τη παλινδρόμηση ότι ο όρος x δεν είναι σημαντικός αλλά ο όρος x^2 . Εάν, τότε εμείς απορρίψουμε τον όρο x τότε προκύπτει ένα νέο μοντέλο της παρακάτω μορφής:

$$y = \beta_0 + \beta_2 x^2 + \varepsilon \quad (2.42)$$

αλλά εάν υποθέσουμε μια αλλαγή κλίμακας $x \rightarrow x + a$, τότε το μοντέλο που θα προκύψει είναι

$$y = \beta_0 + \beta_2 a^2 + 2\beta_2 a x + \beta_2 x^2 + \varepsilon \quad (2.43)$$

Όπως παρατηρούμε ο πρώτος όρος x τώρα εμφανίζεται πάλι. Η κλίμακα αλλάζει και δε θα έπρεπε να δημιουργεί καμία σημαντική αλλαγή στο μοντέλο αλλά σε αυτή τη περίπτωση ένας επιπλέον όρος έχει προστεθεί. Αυτό δεν είναι καλό. Με αυτό αποδεικνύεται γιατί δε πρέπει να απαλείφουμε όρους χαμηλότερης τάξης όταν στην εξίσωση μας υπάρχουν και άλλοι με υψηλότερο βαθμό. Χωρίς όμως αυτό να σημαίνει ότι η επιλογή κλίμακας είναι η εξαρτημένη μεταβλητή. Η απομάκρυνση του πρώτου όρου σημαίνει ότι η προβλεπόμενη λύση είναι συμμετρική και έχει μια βέλτιστη λύση στο $x=0$. Συχνά, αυτή η λύση δεν έχει νόημα και δε θεωρείται σωστή. Αυτή η υπόθεση θα είχε νόημα στα πλαίσια αυτού του προβλήματος, μόνο όταν θα θέλαμε να προσδιορίσουμε τον τελευταίο όρο.

❖ *Μοντέλα με αλληλεπιδράσεις.* Θεωρούμε ένα δεύτερο μοντέλο επιφάνειας:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \beta_{12} x_2 x_1 \quad (2.44)$$

Δε θα μπορούσαμε να αφαιρέσουμε τους όρους αλληλεπίδρασης $x_1 x_2$ χωρίς ταυτόχρονα να μην έχουμε αφαιρέσει και τους όρους x_1^2 και x_2^2 . Εάν όμως απομακρύνουμε και τους τέσσερις όρους τότε η εξίσωση του μοντέλου μας θα μετατραπεί σε γραμμική. Εάν αφαιρέσουμε τους όρους $x_1 x_2$ τότε η νέα εξίσωση θα αντιπροσωπεύει μια επιφάνεια η οποία είναι ευθυγραμμισμένη με τους άξονες συντεταγμένων. Το οποίο είναι δύσκολο να ερμηνευθεί και ισχύει σε πολύ συγκεκριμένες περιπτώσεις. Κάθε μεταβολή του διαστήματος πρόβλεψης θα επανέφερε τον όρο αλληλεπίδρασης και όπως σε κάθε πολυώνυμο δεν είναι επιθυμητό το μοντέλο να εξαρτάται από μια συγκεκριμένη βάση για τη πρόβλεψη της.

2.8.2 Βηματικές Διαδικασίες (Stepwise Procedures)

Σταδιακή διαγραφή μεταβλητών (Backward Elimination)

Αυτή είναι η πιο απλή μέθοδος από όλες τις διαδικασίες επιλογής μεταβλητών και μπορούν πολύ εύκολα να εφαρμοστούν χωρίς κανένα ειδικό πρόγραμμα. Σε περιπτώσεις όπου έχουμε περίπλοκη ιεραρχία, σταδιακή διαγραφή μεταβλητών (backward elimination) μπορεί να χρησιμοποιηθεί όταν λαμβάνουμε υπ' όψη μας ποιες μεταβλητές είναι επιλεγμένες για να αφαιρεθούν. Η μέθοδος αυτή λειτουργεί ως εξής:

1. Ξεκινάμε με το σύνολο των ανεξάρτητων μεταβλητών στο μοντέλο.

2. Αφαιρούμε την ανεξάρτητη μεταβλητή η οποία έχει μεγαλύτερο p από το α_{crit} .
3. Υπολογίζουμε ξανά με τις νέες μεταβλητές.
4. Σταματάμε την παραπάνω διαδικασία που περιγράφουμε όταν οι τιμές του p είναι μικρότερη του α_{crit} .

Όπου p οι δυναμικοί προβλεπτικοί παράγοντες. Το κρίσιμο $\alpha_{(crit)}$ μπορεί να μην είναι το καθιερωμένο 5%. Εάν η επίδοση στην πρόβλεψη είναι ο πρωταρχικός στόχος, τότε ένα κατώφλι απόρριψης 15%-20% δρα καλύτερα. Σε κάθε περίπτωση πάντως μέθοδοι που έχουν σχεδιαστεί για βέλτιστη πρόβλεψη πρέπει να προτιμούνται.

2.8.2.1 Σταδιακή προσθήκη μεταβλητών (Forward Selection)

Σε αυτή τη μέθοδο ακολουθείται η αντίστροφη πορεία από αυτή στη διαγραφή μεταβλητών.

- Ξεκινάμε το μοντέλο χωρίς να υπάρχει κάποια μεταβλητή.
- Για όλες τις ανεξάρτητες μεταβλητές που δεν τις έχουμε εισάγει στο μοντέλο μας, ελέγχουμε τη τιμή του p κάθε μια από αυτές εάν τις εισάγουμε στο μοντέλο μας. Επιλέγουμε μια από αυτές και πιο συγκεκριμένα εκείνη τη μεταβλητή η οποία παρουσιάζει τη χαμηλότερη τιμή p με μικρότερη τιμή από το α_{crit} .
- Συνεχίζουμε τη διαδικασία μέχρι το σημείο εκείνο όπου δεν μπορούμε να προσθέσουμε άλλη προβλέψιμη μεταβλητή στο μοντέλο μας.

2.8.2.2 Βηματική Παλινδρόμηση (Stepwise Regression)

Είναι ο συνδυασμός των δύο παραπάνω μεθόδων δηλαδή της σταδιακής διαγραφής μεταβλητών (backward elimination) και της σταδιακής προσθήκης μεταβλητών (forward selection) . Αυτή η μέθοδος απευθύνεται σε περιπτώσεις όπου οι μεταβλητές προστίθενται και αφαιρούνται στο αρχικό στάδιο της μεθόδου και θέλουμε να αποφασίσουμε για αυτές αργότερα. Σε κάθε στάδιο, μια μεταβλητή μπορεί να προστεθεί ή να αφαιρεθεί και υπάρχουν αρκετές παραλλαγές για το πώς ακριβώς γίνεται αυτή η διαδικασία. Η διαδικασία της βηματικής παλινδρόμησης είναι σχετικά φθηνή από υπολογιστική άποψη όμως έχει κάποια αρνητικά σημεία.

1. Εξαιτίας της προσθήκης ή αφαίρεσης μεταβλητών (μόνο μια κάθε φορά), είναι δυνατό να χαθεί η βέλτιστη λύση.
2. Οι p -τιμές που χρησιμοποιούνται δεν πρέπει να αντιμετωπίζεται πολύ κυριολεκτικά. Υπάρχει τόσες πολλές πολλές δοκιμές που

μπορούν να επικυρώσουν ότι η ισχύς τους είναι αμφίβολη. Η άρση των λιγότερο σημαντικών προβλέψεων τείνει να αυξήσει τη σημασία των υπόλοιπων προβλέψεων. Το αποτέλεσμα αυτό οδηγεί στην υπερεκτίμηση της σημασίας των υπόλοιπων προβλέψεων.

3. Οι διαδικασίες δεν συνδέονται άμεσα με τους τελικούς στόχους της πρόβλεψης ή εξήγησης και έτσι μπορεί να βοηθήσει στην επίλυση του προβλήματος που μας ενδιαφέρει. Με κάθε μέθοδο επιλογής μεταβλητών, είναι σημαντικό να έχουμε κατά νου ότι μοντέλο επιλογής δεν μπορεί να διαχωριστεί από το βασικό σκοπό της έρευνας. Η επιλογή μεταβλητών (variable selection) τείνει να ενισχύσει τη στατιστική σημαντικότητα των μεταβλητών που μένουν στο μοντέλο. Οι μεταβλητές που πέφτουν μπορεί ακόμα να σχετίζονται με την απάντηση. Θα ήταν λάθος να πούμε ότι αυτές οι μεταβλητές δεν έχουν καμία σχέση με την απάντηση, είναι ακριβώς ότι δεν παρέχουν καμία συμπληρωματική επεξηγηματική ισχύ πέραν των μεταβλητών εκείνων που έχουν ήδη συμπεριληφθεί στο μοντέλο.
- Η Σταδιακή επιλογή μεταβλητών τείνει να πάρει τα μοντέλα που είναι μικρότερα από το επιθυμητό για λόγους πρόβλεψης. Για να δώσουμε ένα απλό παράδειγμα, ας θεωρήσουμε την απλή παλινδρόμηση με μία μόνο μεταβλητή πρόβλεψης. Ας υποθέσουμε ότι η κλίση για αυτό το προγνωστικό δεν είναι αρκετά στατιστικά σημαντική. Εμείς μπορεί να μην έχουμε αρκετά στοιχεία για να πούμε ότι είναι σχετικές με το y το οποίο όμως θα ήταν καλύτερα να το χρησιμοποιήσει για προγνωστικούς σκοπούς.

2.8.3 Μέθοδοι βασιζόμενη σε κριτήρια πληροφορίας (Criterion-based procedures)

Εάν σε μία ανάλυση παλινδρόμησης υπάρχουν συνολικά p προγνωστικοί παράγοντες (ανεξάρτητες μεταβλητές), τότε θα υπάρχουν 2^p πιθανά μοντέλα. Εμείς προσαρμόζουμε όλα αυτά τα μοντέλα και διαλέγουμε το καλύτερο σύμφωνα με κάποια κριτήρια. Υπάρχουν αποτελεσματικές μέθοδοι όπως η μεθοδολογία “διακλάδωση και οριοθέτηση” (branch and bound) η οποία αποφεύγει την δοκιμή όλων των δυνατών μοντέλων για να καταλήξει στο προτιμότερο. Μερικά κριτήρια είναι:

- ✓ Το Akaike Κριτήριο Πληροφοριών (Akaike Information Criterion/AIC)
- ✓ Το Bayes Κριτήριο Πληροφοριών (Bayes Information Criterion/BIC)

είναι μερικά από τα πιο συνηθισμένα κριτήρια.

Το Akaike Κριτήριο Πληροφοριών δημιουργήθηκε από τον Hirotosugu Akaike, με το όνομα ‘‘το κριτήριο πληροφοριών’’ (information criterion) και εκδόθηκε από τον Akaike το 1974. Το κριτήριο αυτό στηρίζεται στην ιδέα της εντροπίας των πληροφοριών, στη πραγματικότητα προσφέρεται ως ένα σχετικό μέτρο για τα στοιχεία που χάνονται σε ένα πραγματικό μοντέλο. Θα μπορούσαμε να πούμε ότι το κριτήριο αυτό περιγράφει τη σχέση μεταξύ της ακρίβειας και της πολυπλοκότητας του μοντέλου. Όμως το κριτήριο αυτό δε μπορεί να περιγράψει εάν το μοντέλο που μελετάμε ταιριάζει ικανοποιητικά με τα δεδομένα μας.

Γενικά ισχύει,

$$AIC = n \times \log\text{-likelihood} + 2 \times \text{number of parameters} \quad (2.45)$$

Όταν εφαρμόζουμε τη παραπάνω εξίσωση σε Γκάους ή κανονικά μοντέλα τότε αυτή μετατρέπεται στην εξής μορφή:

$$AIC \approx n \times \log(\text{SSE}) + 2p \quad (2.46)$$

Το πρώτο μέρος $2 \log(\text{SSE})$ μας δείχνει την ακρίβεια της προσαρμογής του μοντέλου. Η σταθερά 2 συχνά έχει το ρόλο της παραμέτρου οσον αφορά της πολυπλοκότητα της παραμέτρου. Όσο πιο μικρή τιμή παρουσιάζει η μεταβλητή AIC τόσο καλύτερο είναι το μοντέλο.

Στη παρούσα εργασία υπολογίζουμε το AIC μέσω του τύπου

$$AIC = n \times \ln(\text{SSE}/n) + 2 \times (p+1) \quad (2.47)$$

Όμως στη θέση της μεταβλητής SSE μπορούμε να βάλουμε το SAE, οπότε ο τύπος γίνεται

$$AIC = n \times 2 \times \ln(\text{SAE}/n) + 2 \times (p+1) \quad (2.48)$$

Εμείς θα υπολογίσουμε και τις δύο παραπάνω μορφές.

Παρατηρώντας ότι το Κριτήριο Πληροφορίας Akaike (AIC) δεν έκανε ακριβείς εκτιμήσεις αλλά αντιθέτως υπερεκτιμούσε τα μοντέλα όταν το μέγεθος του δείγματος ήταν σχετικά μικρό, οι Hurvich και Tsai (1989) πρότειναν μια διαφορετική μορφή, την οποία την ονόμασαν AIC_c , το οποίο δίνεται από την παρακάτω εξίσωση

$$AIC_c \approx n \times \log(SSE) + \frac{n(n+p+1)}{n-p-3} \quad (2.49)$$

Με βάση την εξίσωση του Bayes, ο Schwarz (1978) ανεπτυξε ένα άλλο κριτήριο, το οποίο το ονόμασε ως το κριτήριο πληροφορίας του Μπάιες (Bayesian Information Criterion, BIC) επειδή είναι κριτήριο πληροφοριών με βάση τη Bayesian Μέθοδο (ή αλλιώς SIC επειδή είναι κριτήριο πληροφοριών του Schwarz και SBC καθώς είναι κριτήριο των Schwarz-Bayesian). Δημιουργήθηκε από τον Gideon E. Schwarz, ουσιαστικά είναι ένα κριτήριο για την επιλογή μοντέλων μεταξύ μιας ομάδας παραμετρικών μοντέλων με διαφορετικούς αριθμούς παραμέτρων. Διαλέγοντας ένα μοντέλο για τη βελτιστοποίηση του BIC είναι μια μορφή κανονικοποίησης.

Η τιμή του κριτηρίου BIC δίνεται ως,

$$BIC = -2 \log \text{ - likelihood} + p \log n \quad (2.50)$$

Για τα μοντέλα γραμμικής παλινδρόμησης, ο όρος $-2 \log \text{ - likelihood}$ (γνωστό και ως απόκλιση είναι $n \log(SSE/n)$). Θέλουμε να ελαχιστοποιήσουμε το AIC ή το BIC. Τα μεγάλα μοντέλα ταιριάζουν καλύτερα και έτσι έχουμε μικρότερο SSR αλλά χρησιμοποιούμε περισσότερες παραμέτρους. Συνεπώς η καλύτερη επιλογή μοντέλου θα εξαρτάται από το μέγεθος του μοντέλου. Το κριτήριο BIC ταιριάζει καλύτερα σε μικρά μοντέλα σε αντίθεση με το AIC, το οποίο χρησιμοποιείται για μεγάλα μοντέλα. Τα κριτήρια AIC και BIC μπορούν να χρησιμοποιηθούν ως κριτήρια επιλογής για πολλά διαφορετικά μοντέλα.

Η σειρά απόρριψης μεταβλητών είναι ίδια με την μέθοδο σταδιακής διαγραφής μεταβλητών (backward elimination). Η μόνη διαφορά είναι ότι ο πληθυσμός μεταβλητών δε μεταβάλλεται.

- ✓ Ο συντελεστής προσδιορισμού R^2 . Ή αλλιώς γράφεται $R^2 = 1 - SSR/SST$. Πρόσθετοντας μια μεταβλητή στο μοντέλο μπορούμε να

μειώσουμε το SSR και έτσι μόνο αυξάνεται το R^2 , έτσι το R^2 δεν μπορεί να θεωρηθεί καλό κριτήριο γιατί σε κάθε πιθανή αλλαγή θα προτιμάμε πάντα το μεγαλύτερο.

$$R_a^2 = 1 - \frac{RSS/(n-p)}{TSS/(n-p)} = 1 - \frac{n-1}{n-p}(1 - R^2) = 1 - \widehat{\sigma}_{model}^2 / \widehat{\sigma}_{null}^2 \quad (2.51)$$

- ✓ Το προβλεπόμενο άθροισμα τετραγώνων των υπολοίπων (Predicted Residual Sum of Squares /PRESS) το οποίο ορίζεται από το $\sum_i \widehat{\varepsilon}_{(i)}^2$ όπου $\widehat{\varepsilon}_{(i)}$ είναι the residuals χρησιμοποιώντας την υπόθεση ότι το i ταιριάζει στο μοντέλο μας. Επιλέγουμε το μοντέλο που παρουσιάζει το μικρότερο PRESS. Τείνει να πάρει τα μεγαλύτερα μοντέλα (το οποίο μπορεί να είναι επιθυμητό εάν επιτυγχάνει το αποτέλεσμα που θέλουμε) (Julian J. Faraway, 2002).

Γενικά ισχύει ότι :

- i. Το C_p είναι εύκολο να υπολογισθεί.
- ii. Είναι στενά συνδεδεμένα το R_a^2 και το AIC.
- iii. Για το πλήρες μοντέλο $C_p = p$.
- iv. Εάν η μεταβλητή ταιριάζει με το μοντέλο πρόβλεψης τότε $E(SSR_p) = (n-p) \sigma^2$ και τότε $E(C_p) \approx p$. Ένα μοντέλο με κακή πρόβλεψη τότε θα έχει C_p πολύ μεγαλύτερο p .

Το C_p μπορεί να εκτιμηθεί από το τύπο

$$C_p = \frac{SSR_p}{\widehat{\sigma}^2} + 2p - n \quad (2.52)$$

Οι McQuarrie και Tsai (1998) κατέληξαν σε κάποια γενικά συμπεράσματα όσον αφορά τη χρησιμοποίηση κριτηρίων για την επιλογή μοντέλων. Υπέδειξαν ότι το AIC κριτήριο και το C_p λειτουργούν καλύτερα σε δείγματα μετρίου μεγέθους, το AIC_c κριτήριο είναι η πιο αποτελεσματική επιλογή για μικρά δείγματα, ενώ το BIC κριτήριο είναι το καταλληλότερο για μεγάλα δείγματα (Xin Yan, Xiao Gang Su, 2009).

2.8.4 Περιληπτικά

Η επιλογή μεταβλητών (variable selection) είναι ένα μέσο για έναν σκοπό και όχι αυτοσκοπός. Ο στόχος είναι να οικοδομήσουμε ένα μοντέλο το οποίο προβλέπει καλά ή εξηγεί τις σχέσεις με τα δεδομένα. Η αυτόματη επιλογή μεταβλητών δεν εγγυάται συνέπεια με αυτούς τους στόχους. Οι μέθοδοι αυτοί μπορούν να χρησιμοποιηθούν ως οδηγοί. Η σταδιακή (Stepwise) μέθοδος χρησιμοποιεί μια περιορισμένη μηχανή μέσω του χώρου των δυνητικών μοντέλων και χρησιμοποιεί μια αμφίβολη υπόθεση δοκιμών βασισμένη στη μέθοδο για την επιλογή μεταξύ των μοντέλων. Το Criterion-based procedures συνήθως συνεπάγονται μια ευρύτερη αναζήτηση και η σύγκριση μοντέλων σε ένα καλύτερο τρόπο. Για το λόγο αυτό, συνίσταται να χρησιμοποιείτε αυτή η μέθοδος. Δέχεται τη δυνατότητα ότι πολλά μοντέλα μπορούν να προταθούν τα οποία ταιριάζουν εξίσου με άλλα. Εάν συμβεί αυτό, εξετάζει:

- (a) Εάν τα μοντέλα αυτά έχουν τις ίδιες ποιοτικές συνέπειες;
- (b) Εάν έχουν τις ίδιες προβλέψεις;
- (c) Ποιο είναι το κόστος της μέτρησης της πρόβλεψης;
- (d) Ποια μέθοδο μας δίνει τη καλύτερη διάγνωση;

Εάν βρούμε μοντέλα τα οποία μοιάζουν πολύ μεταξύ τους αλλά καταλήγουν σε διαφορετικά αποτελέσματα τότε είναι σαφές ότι τα δεδομένα, τα οποία έχουμε, δε μπορούν να απαντήσουν στα ερωτήματα που μας ενδιαφέρουν. Πρέπει να το έχουμε υπόψη μας πολύ συχνά, καθώς είναι πολύ συχνό φαινόμενο να παρουσιάζονται μοντέλα τα οποία να μας δίνουν τέτοιου είδους αποτελέσματα (Norman R. Draper & Harry Smith, 1998).

2.9 Πολυκριτηριακή Παλινδρόμηση (Multiple Criteria Regression)

Οι άγνωστοι παράμετροι σε ένα μοντέλο πολλαπλής γραμμικής παλινδρόμησης (multiple linear regression) μπορούν να εκτιμηθούν χρησιμοποιώντας ένα από τα κριτήρια, όπως η ελαχιστοποίηση του μέσου όρου των τετραγώνων των σφαλμάτων (minimization of the sum of squared errors/minimum SSE), η ελαχιστοποίηση του μέσου όρου των απόλυτων σφάλματων (minimization of the sum of absolute

errors/minimumSAE), και η ελαχιστοποίηση των μεγίστων απόλυτων σφαλμάτων (minimization of the maximum absolute error/minimumMAE). Το κριτήριο των ελαχίστων τετραγώνων, εδώ και χρόνια, συνεχίζει να είναι το πιο γνωστό. Δυστυχώς κάποιες φορές η επιλογή του κριτηρίου δεν είναι πάντα ξεκάθαρη από στατιστικής, πρακτικής ή άλλων θεωρήσεων. Σε συγκεκριμένες συνθήκες, ίσως είναι καταλληλότερο να χρησιμοποιήσουμε πολλαπλά κριτήρια (multiple criteria) από το να χρησιμοποιήσουμε ένα μόνο κριτήριο για να εκτιμήσουμε τις άγνωστες παραμέτρους ενός πολλαπλού γραμμικού μοντέλου.

Το ελάχιστο άθροισμα των τετραγώνων των σφαλμάτων (minimumSSE), το ελάχιστο άθροισμα των απόλυτων σφαλμάτων MSAE και η ελαχιστοποίηση του μεγίστου των σφαλμάτων (minimumMAE) είναι πιθανά κριτήρια για την εκτίμηση των άγνωστων παραμέτρων ενός πολλαπλού γραμμικού μοντέλου παλινδρόμησης. Το MSSE ή κριτήριο ελαχίστων τετραγώνων συνεχίζει να είναι το πιο εύχρηστο όλων. Η θεωρία του είναι πολύ απλή, καλά ανεπτυγμένη και τεκμηριωμένη, και είναι άμεσα διαθέσιμη σε προγράμματα υπολογιστών. Το διαμορφωμένο μοντέλο παλινδρόμησης διέρχεται μέσα από το μέσο όρο όλων των μεταβλητών. Ευτυχώς, η διαδικασία είναι πολύ ευαίσθητη στις ακραίες τιμές και οι ακραίες τιμές είναι δύσκολο να εντοπιστούν σε προβλήματα παλινδρόμησης. Η επίδραση των ακραίων τιμών που σημειώνονται κατά τις τιμές των μεταβλητών πρόβλεψης μπορεί να επιφέρει διαταραχές. Επίσης, δεν είναι ικανοποιητικό κριτήριο ούτε όταν μία ή περισσότερες υποθέσεις δεν ικανοποιούνται ούτε όταν η δευτεροβάθμια συνάρτηση σφαλμάτων που προσαρμόζουμε δεν είναι αποδεκτό μέτρο των αποκλίσεων. Η MSAE παλινδρόμηση είναι λιγότερο ευαίσθητη με τις ακραίες τιμές από τα ελάχιστα τετράγωνα παλινδρόμησης. Παρέχει ένα καλύτερο σημείο εκκίνησης σε σχέση με τη παλινδρόμηση ελαχίστων τετραγώνων για ορισμένες αυστηρές διαδικασίες παλινδρόμησης. Ωστόσο, σε αντίθεση με τις περισσότερες άλλες αποτελεσματικές διαδικασίες παλινδρόμησης που απαιτούν μια "σταθερά," η παλινδρόμηση MSAE δεν απαιτεί κάποια. Το πρόβλημα υπολογισμού εκτιμήσεων των MSAE μπορεί να διατυπωθεί και να λυθεί σαν ένα πρόβλημα γραμμικού προγραμματισμού.

Μια σειρά από πολύ αποδοτικούς και αποτελεσματικούς αλγόριθμους είναι διαθέσιμοι για την επίλυση των προβλημάτων πολλαπλής γραμμικής παλινδρόμησης MSAE, Coleman and Li, 1992 T.F. Coleman and Y. Li, A globally and quadratically convergent affine scaling method for linear II problems, *Mathematical Programming* 56 (1992), pp. 189–222. Coleman and Li, 1992, Madsen and Nielsen, 1993, Ruzinsky and Olsen, 1989, Zhang, 1993 and Portnoy and Koenker, 1997 μεταξύ άλλων.

Η ελαχιστοποίηση του μεγίστου των σφαλμάτων (MMAE), στη παλινδρόμηση είναι πολύ ευαίσθητη με αποτέλεσμα να μπορεί να

προλαμβάνει τα χειρότερα. Τα αποτελέσματα των κριτηρίων της μέγιστης πιθανοφάνειας των παραμέτρων του μοντέλου, ενός πολλαπλού γραμμικού μοντέλου παλινδρόμησης, εάν τα σφάλματα είναι ανεξάρτητα και ακολουθούν μια ομοιόμορφη κατανομή με ένα ευρύ φάσμα. Το πρόβλημα υπολογισμού εκτίμησεις των ΜΜΑΕ μπορεί να διατυπωθεί και να λυθεί σε πρόβλημα γραμμικού προγραμματισμού. Μερικά αποτελεσματικά προγράμματα υπολογιστών, για παράδειγμα Armstrong and Kung (1979) and Barrodale and Phillips 1974 and Barrodale and Phillips, 1975, μπορούν να χρησιμοποιηθούν για γραμμική διαμόρφωση προγραμματισμού αποτελεσματικά και αποδοτικά.

Είναι σαφές ότι κάθε ένα από τα κριτήρια αυτά είναι η καλύτερη δυνατή κάτω από ορισμένες συνθήκες διαφορετικά, το κριτήριο μπορεί να έχει ή να μην έχει καλή απόδοση. Συνεπώς, φαίνεται λογικό να μην εξαρτάται από οποιοδήποτε κριτήριο ανά πάσα στιγμή. Σύμφωνα με όσα είπε ο Hogg (1974). Όπως γνωρίζουμε, στην πράξη τα περισσότερα μοντέλα θα ταιριάζουν ακριβώς στη πραγματική κατάσταση. Μια ρεαλιστικότερη προσέγγιση θα ήταν να αναζητήσει τις στατιστικές διαδικασίες, καλό για μια ευρεία κατηγορία των πιθανών γενεσιουργών μοντέλα, τα οποία όμως δεν είναι απαραίτητα καλύτερα για κάθε μία από αυτές. Στα περισσότερα πρακτικά προβλήματα η φύση της κατανομής των σφαλμάτων είναι σπανίως (κι αν είναι) γνωστή. Τα σφάλματα μπορεί να μην προέρχονται από μια μοναδική κατανομή. Ακραίες τιμές μπορεί να υπάρχουν αλλά είναι δύσκολο να εντοπιστούν. Η επιλογή της συνάρτησης σφάλματος μπορεί να μην είναι σαφής από στατιστική και πρακτική άποψη.

Η χρήση του πολλαπλού κριτηρίου (multiple criteria) αντί του απλού κριτηρίου, για την εκτίμηση άγνωστων παραμέτρων σε ένα μοντέλο πολλαπλής γραμμικής παλινδρόμησης είναι κατάλληλο από οποιοδήποτε

Η φύση του σφάλματος κατανομής είναι άγνωστη

Τα λάθη δε προκύπτουν από μια ενιαία κατανομή

Οι ακραίες τιμές εμφανίζονται και είναι δύσκολο να ανιχνευθούν.

Η επιλογή μιας συνάρτησης απωλειών δεν είναι σαφής.

Η διαδικασία επίλυσης είναι διαδραστική και απαιτεί τη λήψη αποφάσεων.

Αυτό έχει σημαντικές συνέπειες για λήψη αποφάσεων στο τομέα διοίκησης καθώς ο άνθρωπος που πρέπει να πάρει αποφάσεις δεν είναι ένας παθητικός θεατής, αλλά συμμετέχει ενεργά στη διαμόρφωση, τη διαδικασία λύσης, και η επιλογή του τελικού διαμορφωμένου μοντέλου.

Αυτό μπορεί επίσης να βελτιώσει το διαμορφωμένο μοντέλο. Ενδιαφέροντα άρθρα πάνω σε αυτό το θέμα είναι των Dielman, 2005, Andre et al., 2000, Andre et al., 2003, Elian et al., 2000 and Narula et al., 1999 για MSAE παλινδρόμηση.

Συνοπτικά, το κριτήριο της πολλαπλής γραμμικής παλινδρόμησης συνδυάζει την εκτίμηση των παραμέτρων και την απόφαση επιλογής μεταβλητών. Σε αντίθεση με το μοναδικό μοντέλο προβλέπει από το απλό κριτήριο παλινδρόμησης, το πολλαπλό κριτήριο παλινδρόμησης προβλέπει ένα σύνολο υποψηφίων μοντέλων σύμφωνα με την επιλογή κριτηρίων. Άλλωστε, το κριτήριο πολλαπλής παλινδρόμησης είναι κατάλληλο όποτε η κατανομή σφάλματος είναι άγνωστη ή προκύπτει από τις διάφορες κατανομές, οι ακραίες τιμές είναι παρών και γενικά είναι δύσκολο να ανιχνευθούν, οπότε δεν είναι δύσκολο να αποφασίσουν για μια απλή συνάρτηση σφάλματος. Η πολυκριτηριακή επιλογή μεταβλητών μπορεί να θεωρηθεί μια επέκταση της πολυκριτηριακής γραμμικής παλινδρόμησης επειδή δεν προσμετρά μόνο τους συντελεστές παλινδρόμησης αλλά επίσης και τις μεταβλητές τις οποίες εισάγουμε στο μοντέλο (model selection) (Subhash C.Narula, John F.Wellington, 2006).

Κεφάλαιο 3

Μελέτη περίπτωσης και σχεδιασμός υπολογιστικού πειράματος

Σε αυτό το κεφάλαιο θέλουμε να αναλύσουμε την εφαρμογή των μοντέλων και να παρουσιάσουμε τα νέα δεδομένα που προκύπτουν από την επεξεργασία των δεδομένων μας. Εξηγούμε τη διαδικασία επιλογής των 10 συνόλων δεδομένων που έχουμε και στη συνέχεια αναλύουμε τη πορεία που ακολουθήσαμε για την επεξεργασία αυτών.

3.1 Μελέτη περίπτωσης

Στο κεφάλαιο αυτό, θα εφαρμόσουμε τις μεθόδους και τα παραπάνω μοντέλα που αναλύσαμε σε ένα σύνολο σωματομετρικών δεδομένων. Η συγκεκριμένη βάση δεδομένων υπάρχει διαθέσιμη στο διαδίκτυο στην ηλεκτρονική διεύθυνση <http://lib.stat.cmu.edu/datasets/bodyfat> και χρησιμοποιήθηκε και στην εργασία των Μαυρωτά και Φλώριου με τίτλο “A combined approach for multiple criteria regression and variable selection using multi-objective programming”.

Σε αυτή τη βάση δεδομένων υπάρχουν 250 παρατηρήσεις πειραματικών μετρήσεων λίπους του σώματος (% λίπος ως προς το βάρος). Παράλληλα, κατάγραφοσαν η ηλικία και 13 σωματομετρικές παράμετροι που είναι οι εξής:

3.1 Πίνακας Παραμέτρων

Παράμετροι	parameters
Ηλικία (χρόνια)	Age (years)
Βάρος (κιλά)	Weight (kg)
Ύψος (εκατοστά)	Height (cm)
Δείκτης μάζας σώματος	BMI
Περιφέρεια Λαιμού (εκατοστά)	Neck (cm)
Περιφέρεια Στήθους (εκατοστά)	Chest (cm)
Περιφέρεια Κοιλιάς (εκατοστά)	Abdom (cm)
Περιφέρεια Ισχίου (εκατοστά)	Hip (cm)
Περιφέρεια Μηρού (εκατοστά)	Thigh (cm)
Περιφέρεια Γονάτου (εκατοστά)	Knee (cm)
Περιφέρεια Αστράγαλου (εκατοστά)	Ankle (cm)
Περιφέρεια δικεφάλου (εκατοστά)	Biceps (cm)
Περιφέρεια Καρπού (εκατοστά)	Wrist (cm)
Περιφέρεια Πήχης (εκατοστά)	Forarm (cm)

Στο παράρτημα I φαίνεται και η βάση δεδομένων, στην οποία βασιστήκαμε για την εκπόνηση αυτής της εργασίας.

Πρέπει να σημειωθεί ότι αρχικά υπήρχαν 252 παρατηρήσεις αλλά οι 2 αφαιρέθηκαν από το δείγμα γιατί υπήρχαν αμφιβολίες για την

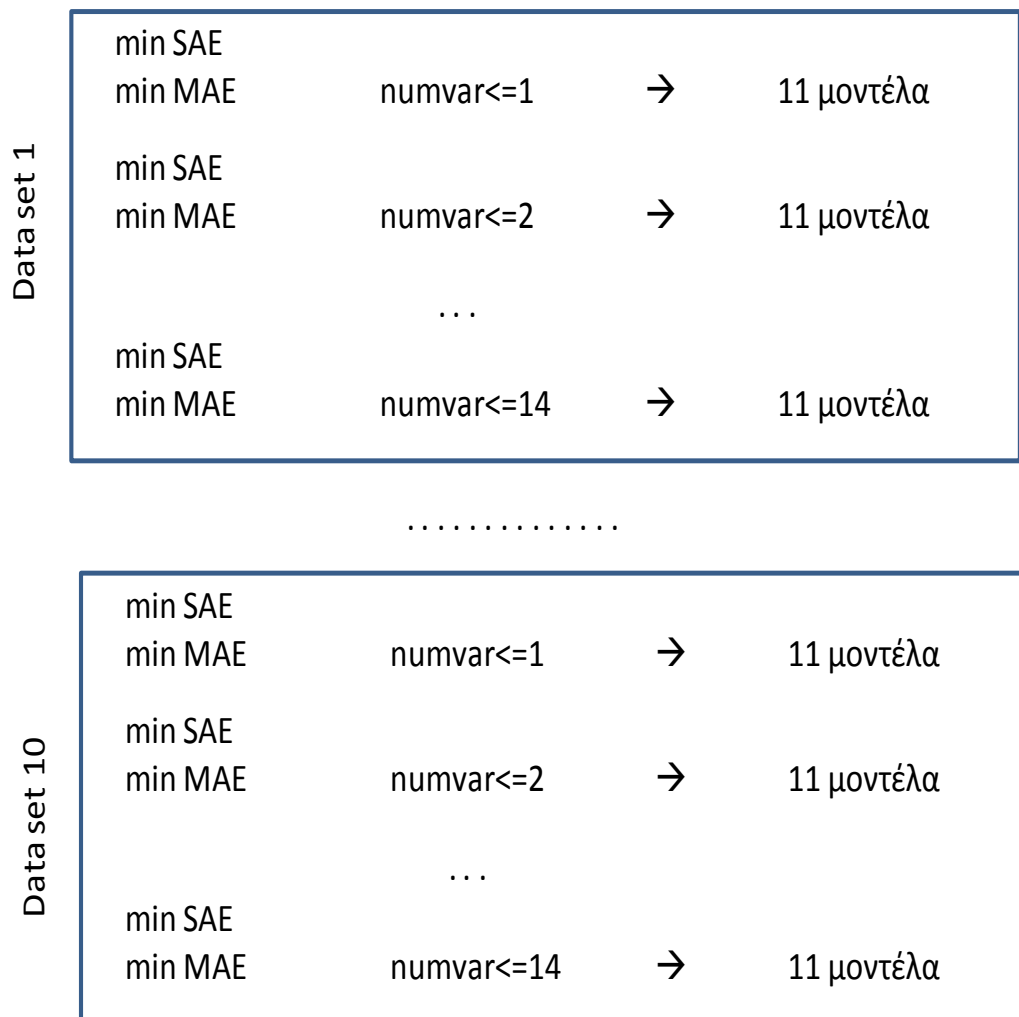
ορθότητά τους. Το μοντέλο παλινδρόμησης που αναπτύχθηκε επιχειρεί να συνδέσει το ποσοστό λίπους με τις 13 αυτές σωματομετρικές παράμετροι μαζί με την ηλικία. Λαμβάνουμε δηλαδή ως εξαρτημένη μεταβλητή το ποσοστό λίπους και ως ανεξάρτητες μεταβλητές τις 14 αυτές παραμέτρους.

3.2 Σχεδιασμός υπολογιστικού πειράματος

Στην συγκεκριμένη εργασία ακολουθήσαμε την πρακτική της μέτρησης της προβλεπτικής ικανότητας του μοντέλου εντός κι εκτός δείγματος (out of sample). Για το δεύτερο, εξάγαμε το μοντέλο από ένα υποσύνολο παρατηρήσεων και το εφαρμόζαμε για να διαπιστώσουμε την επάρκειά τους στις υπόλοιπες παρατηρήσεις. Το αρχικό υποσύνολο στην περίπτωση μας (υποσύνολο εξαγωγής μοντέλου) περιελάμβανε το 90% των παρατηρήσεων (225 παρατηρήσεις) και το υποσύνολο δοκιμής το 10% (25 παρατηρήσεις). Για να εξάγουμε πιο εύρωστα συμπεράσματα χρησιμοποιήσαμε την μέθοδο bootstrapping για την δημιουργία 10 σετ δεδομένων. Κάθε δηλαδή σετ είχε 225 διαφορετικές παρατηρήσεις ως υποσύνολο εξαγωγής του μοντέλου. Η επιλογή των 225 παρατηρήσεων ήταν εντελώς τυχαία και έγινε χρησιμοποιώντας την εντολή RANDOM στο πρόγραμμα EXCEL. Εφαρμόζοντας την εντολή αυτή 25 φορές απομονώσαμε τις 25 παρατηρήσεις και τις θεωρήσαμε ως εκτός δείγματος μετρήσεις (out of sample). Έτσι δημιουργήσαμε μια νέα βάση δεδομένων με 225 παρατηρήσεις. Όμως για να έχουμε αξιόπιστα αποτελέσματα και να μπορέσουμε να καταλήξουμε σε κάποια σωστά συμπεράσματα δεν ήταν επαρκή μόνο αυτή η μια βάση δεδομένων. Για το λόγο αυτό, δημιουργήσαμε και άλλες βάσεις δεδομένων με τον ίδιο τρόπο που περιγράψαμε παραπάνω, φτάνοντας στη δημιουργία συνολικά 10 σετ δεδομένων.

Για κάθε ένα από τα 10 σετ δεδομένων εφαρμόζεται το πολυκριτηριακό μοντέλο παλινδρόμησης με κριτήρια ελαχιστοποίησης το min SAE και min MAE μεταβάλλοντας συγχρόνως και τον μέγιστο αριθμό των μεταβλητών που περιλαμβάνονται στο μοντέλο (από 1-14). Θέτοντας λοιπόν ένα άνω όριο στις συμμετέχουσες μεταβλητές στο μοντέλο επιλύουμε το πολυκριτηριακού μοντέλου παλινδρόμησης παράγοντας 11 μοντέλα παλινδρόμησης (11 κατά pareto άριστες λύσεις). Να σημειωθεί εδώ ότι το πολυκριτηριακό μοντέλο παλινδρόμησης δεν παράγει μόνο μία βέλτιστη λύση δηλαδή ένα μοντέλο παλινδρόμησης αλλά ένα σύνολο 11 Pareto βέλτιστων λύσεων (pareto βέλτιστα μοντέλα παλινδρόμησης). Τα δύο ακραία μοντέλα, το 1 και το 11 αντιστοιχούν στα μονοκριτηριακά μοντέλα όπου στο μεν 1^ο επιδιώκεται η ελαχιστοποίηση του SAE ενώ στο 11^ο επιδιώκεται η ελαχιστοποίηση του MAE.

Ο αριθμός των συνολικών μοντέλων που παράγονται είναι $11 \times 14 = 151$ για κάθε data set. Η διαδικασία φαίνεται παραστατικά στο ακόλουθο σχήμα:



Σχήμα 3.1: Η υπολογιστική διαδικασία παραγωγής των μοντέλων παλινδρόμησης

3.3 Εφαρμογή μεθόδων εντός δείγματος

Η πρώτη κατηγορία υπολογιστικών πειραμάτων ήταν αυτά που πραγματοποιήσαμε μόνο με τις 225 παρατηρήσεις (εντός δείγματος). Για κάθε ένα από τα 10 σετ δεδομένων εξάγονται 14 μοντέλα με βάση την ελαχιστοποίηση του Αθροίσματος των Απόλυτων Σφαλμάτων (Sum of Absolute Errors, SAE) και θέτοντας ως ανώτερο όριο ανεξάρτητων μεταβλητών σε κάθε μοντέλο από 1 έως 14. Σε αυτά τα 14 μοντέλα υπολογίζονται τα κριτήρια AIC και BIC προκειμένου να διαπιστώσουμε ποιο από τα 14 μοντέλα είναι καλύτερο (παρουσιάζει το ελάχιστο AIC και BIC). Μάλιστα ο υπολογισμός του AIC και του BIC γίνεται με βάση το Άθροισμα των Απόλυτων Σφαλμάτων (SAE) αλλά και το Άθροισμα των Τετραγώνων των Σφαλμάτων (Sum of Square Errors, SSE). Από τα 10 σετ δεδομένων θα βγάλουμε συμπέρασμα για το πόσες μεταβλητές

προτείνεται να έχει τελικά το μοντέλο με βάση τα κριτήρια AIC και BIC.

3.4 Εφαρμογή μεθόδων εκτός δείγματος

Η εφαρμογή της μεθόδου γινόταν στο σύνολο των 225 παρατηρήσεων προκειμένου να εξαχθούν τα κατάλληλα μοντέλα παλινδρόμησης. Στη συνέχεια αυτά τα μοντέλα τα εφαρμόζαμε στις 25 παρατηρήσεις εκτός δείγματος για να δούμε ποιο μοντέλο περιγράφει καλύτερα τα δεδομένα κι έτσι να κάνουμε μια συγκριτική ανάλυση των εξαγόμενων μοντέλων. Η επάρκεια της προσαρμογής εκτός δείγματος γίνεται μετρώντας το SAE το MAE και το SSE που προκύπτει όταν εφαρμόζεται το μοντέλο στις 25 παρατηρήσεις που είναι εκτός δείγματος. Με βάση αυτό θα μελετήσουμε:

A) Ποιο είναι καλύτερο κριτήριο στα μοντέλα παλινδρόμησης: Η ελαχιστοποίηση του SAE ή του MAE;

B) Ποιο μοντέλο δίνει καλύτερα αποτελέσματα: αυτό που προκύπτει από μονοκριτηριακή ανάλυση παλινδρόμησης (τα δύο ακραία μοντέλα δηλαδή που αντιστοιχούν στο \min SAE και στο \min MAE) ή αυτό που προκύπτει από την πολυκριτηριακή ανάλυση παλινδρόμησης (κάποιο από τα 9 ενδιάμεσα μοντέλα)

Γ) Ποιος είναι ο βέλτιστος αριθμός μεταβλητών από το 1 ως το 14

Δ) Ποια είναι καλύτερη προσέγγιση για την επιλογή μεταβλητών: Η βηματική ανάλυση παλινδρόμησης (stepwise regression) ή η ανάλυση παλινδρόμησης με Μαθηματικό Προγραμματισμό για την επιλογή μεταβλητών.

Η βηματική ανάλυση παλινδρόμησης πραγματοποιήθηκε με τη βοήθεια ενός πρόσθετου υπολογιστικού εργαλείου που διαθέτει το πρόγραμμα Microsoft Excel και ονομάζεται Winstat. Αρχικά, δημιουργούμε ένα φύλλο σε ένα αρχείο όπου έχουμε τη βάση δεδομένων μας. Από τη γραμμή εργαλείων διαλέγουμε την εφαρμογή Regression Analysis και στη συνέχεια ορίζουμε τα κελιά στα οποία θα εφαρμοστεί η μέθοδος που διαλέξαμε καθώς και τις ανεξάρτητες/εξαρτημένες μεταβλητές. Το υποπρόγραμμα αυτό επεξεργάζεται τα δεδομένα και στη συνέχεια εξάγει τα αποτελέσματα σε ένα νέο φύλλο excel. Η διαδικασία που περιγράφουμε παραπάνω εκτελέστηκε άλλες εννέα φορές καθώς τα σύνολα δεδομένων μας είναι δέκα. Το πρόγραμμα αυτό κάνει ταυτόχρονα και variable selection, δηλαδή διαλέγει τον καλύτερο αριθμό των μεταβλητών ως άνω όριο.

Κεφάλαιο 4

Ανάπτυξη του μοντέλου

Σε αυτό το κεφάλαιο προσπαθήσαμε να αναπτύξουμε τα μοντέλα που χρησιμοποιούμε σε αυτή την εργασία. Πιο αναλυτικά, γράφουμε τις εξισώσεις και τους περιορισμούς για τα μοντέλα παλινδρόμησης τα οποία αφορούν για την ελαχιστοποίηση του αθροίσματος των σφαλμάτων (SAE), την ελαχιστοποίηση του μεγίστου των απόλυτων σφαλμάτων (MAE), η επιλογή μεταβλητών και το πολυκριτηριακό μοντέλο.

Για την ανάπτυξη του μοντέλου μας χρησιμοποιήσαμε διάφορα μοντέλα μικτού ακέραιου προγραμματισμού (Mixed Integer Programming). Παρακάτω θα αναλύσουμε τα χαρακτηριστικά αυτών των μοντέλων τα οποία υλοποιήθηκαν και επιλύθηκαν με τη γλώσσα μοντελοποίησης GAMS.

4.1 Το απλό μοντέλο παλινδρόμησης με ελαχιστοποίηση του SAE

Υπάρχουν πολλές συναρτήσεις σφαλμάτων παρ' όλο που στη βιβλιογραφία κυριαρχεί η συνάρτηση ελαχίστων τετραγώνων της παλινδρόμησης. Μια συνάρτηση είναι το άθροισμα των απόλυτων σφαλμάτων (Sum of Absolute Errors, SAE) η οποία είναι λιγότερο ευαίσθητη σε σχέση με το άθροισμα των τετραγώνων των σφαλμάτων (Sum of Squared Errors, SSE) για τις ακραίες τιμές αλλά παρουσιάζει ένα πολύ καλό μοντέλο. Το μοντέλο επιλογής μεταβλητών (variable selection) το οποίο ελαχιστοποιεί το SAE στο μοντέλο στο Μεικτό Ακέραιο Προγραμματισμό (MILP) είναι παρόμοιο με τους στόχους του προγραμματισμού. Συνεπώς, η αντικειμενική συνάρτηση για το απλό μοντέλο παλινδρόμησης με ελαχιστοποίηση του SAE θα είναι η εξής:

$$\mathbf{min} \sum_{i=1}^n (u_i + v_i) \quad (4.1)$$

με το παρακάτω περιορισμό

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_m X_{mi} + u_i - v_i, \quad i=1 \dots n \quad (4.2)$$

Όπου n ο αριθμός των παρατηρήσεων

m ο αριθμός των ανεξάρτητων μεταβλητών

i είναι ο δείκτης των παρατηρήσεων

j είναι ο δείκτης των ανεξαρτητών μεταβλητών (AGE, WEIGHT, HEIGHT, BMI, NECK, CHEST, ABDOM, HIP, THIGH, KNEE, ANKLE, BICEPS, FORARM, WRIST)

Y_i είναι η εξαρτημένη μεταβλητή της i -παρατήρησης

X_{ij} είναι η τιμή της i -παρατήρησης στην j -επεξηγηματική μεταβλητή

u_i είναι η μεταβλητή απόφασης που εκφράζει την υποεκτίμηση της παρατήρησης i (συνεχής θετική μεταβλητή)

v_i είναι η μεταβλητή απόφασης που εκφράζει την υπερεκτίμηση της παρατήρησης i (συνεχής θετική μεταβλητή)

b_0 είναι η μεταβλητή απόφασης που εκφράζει τον σταθερό όρο (ελεύθερη συνεχή μεταβλητή)

b_j είναι η μεταβλητή απόφασης που εκφράζει τον συντελεστή για την j - ανεξάρτητη μεταβλητή (ελεύθερη συνεχής μεταβλητή)

4.2 Το απλό μοντέλο παλινδρόμησης με ελαχιστοποίηση του MAE

Επειδή θέλαμε να αποφύγουμε το χειρότερο σενάριο εκτίμησης, μια εναλλακτική συνάρτηση σφαλμάτων είναι η ελαχιστοποίηση του μέγιστου απόλυτου σφάλματος (Maximum Absolute Error, MAE) . Σε αυτή τη περίπτωση, αναζητήσαμε ένα μοντέλο το οποίο θα προσπαθεί να ελαχιστοποιήσει όχι το μέσο όρο σφαλμάτων αλλά το μέγιστο σφάλμα μεταξύ των παρατηρήσεων. Ως εκ τούτου ακόμα και στο χειρότερο σενάριο (το οποίο είναι να έχουμε πολύ μεγάλη απόκλιση από το μοντέλο το οποίο παρατηρούμε) το σφάλμα θα είναι το μικρότερο δυνατό. Η αντικειμενική συνάρτηση για απλό μοντέλο παλινδρόμησης με ελαχιστοποίηση του MAE θα είναι η εξής:

$$\mathbf{min} \mathbf{dmax} \quad (4.3)$$

Με περιορισμούς τις παρακάτω συναρτήσεις:

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_m X_{mi} + u_i - v_i, \quad i=1 \dots n \quad (4.4)$$

$$\begin{aligned} dmax &\geq u_i & i = 1 \dots n \\ dmax &\geq v_i & i = 1 \dots n \end{aligned} \quad (4.5)$$

όπου

dmax είναι η μεταβλητή απόφασης η οποία εκφράζει το μέγιστο απόλυτο σφάλμα (απόκλιση) μεταξύ των παρατηρήσεων.

4.3 Το μοντέλο παλινδρόμησης με ελαχιστοποίηση του SAE και επιλογή μεταβλητών (variable selection)

Για να καταφέρουμε να επιτύχουμε και την ελαχιστοποίηση των σφαλμάτων SAE αλλά και την καλύτερη επιλογή μεταβλητών, κατασκευάσαμε ένα μοντέλο το οποίο να ικανοποιεί και τα δύο παραπάνω κριτήρια. Έτσι λοιπόν, η αντικειμενική συνάρτηση για το μοντέλο παλινδρόμησης με ελαχιστοποίηση του SAE και την καλύτερη επιλογή μεταβλητών (variable selection) θα έχει τη παρακάτω μορφή:

$$\min \sum_{i=1}^n (u_i + v_i) \quad (4.6)$$

με το παρακάτω περιορισμό

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_m X_{mi} + u_i - v_i, \quad i=1 \dots n \quad (4.7)$$

και

$$\sum b_j \leq \text{upperbound} \quad (4.8)$$

4.4 Το πολυκριτηριακό μοντέλο παλινδρόμησης και variable selection

Σε αυτό το μοντέλο θα έχουμε 3 αντικειμενικές συναρτήσεις χρησιμοποιώντας παράλληλα και την αρχή της μετριοπάθειας (concept of parsimony). Ταυτόχρονα, θα επιλέγουμε και το κατάλληλο αριθμό μεταβλητών.

Δηλαδή θα ισχύει

$$\text{Max parsimony} \rightarrow \min \text{NVAR} = d_1 + d_2 + \dots + d_m, \quad (4.9)$$

Οι συναρτήσεις που θα περιγράφουν τις 3 αντικειμενικές συναρτήσεις θα έχουν τη παρακάτω μορφή:

$$\min \sum_{i=1}^n (u_i + v_i) \quad (4.10)$$

$$\min d_{\max} \quad (4.11)$$

$$\min \sum_{j=1}^m d_j \quad (4.12)$$

st

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_m X_{mi} + u_i - v_i \quad , \quad i=1 \dots n \quad (4.13)$$

$$\begin{aligned} \max u_i & \quad i = 1 \dots n \\ \max v_i & \quad i = 1 \dots n \end{aligned} \quad (4.14)$$

Όπου

d_j είναι η δυαδική μεταβλητή απόφασης η οποία εκπροσωπεί την j -ανεξάρτητη μεταβλητή του μοντέλου.

4.5 Η γλώσσα GAMS

Η γλώσσα που χρησιμοποιήσαμε για το μοντέλο μας είναι η γλώσσα GAMS. Παρακάτω αναλύουμε το μοντέλο GAMS. Πιο συγκεκριμένα εξηγούμε το σχεδιασμό και τον τρόπο λειτουργίας του κώδικα GAMS. Επιπλέον, εξηγεί πως δομείται ένα μοντέλο GAMS και τι σχόλια χρησιμοποιούμε για τη γλώσσα GAMS.

Το Γενικό Αλγεβρικό Σύστημα Μοντελοποίησης GAMS (General Algebraic Modeling System) σχεδιάζεται για την ανάλυση γραμμικών, μη γραμμικών εφαρμογών αλλά και μικτών προβλημάτων βελτιστοποίησης ακέραιων αριθμών. Το σύστημα είναι ιδιαίτερα χρήσιμο για μεγάλα και πολυσύνθετα προβλήματα, ενώ επιτρέπει στο χρήστη να επικεντρωθεί στο πρόβλημα του μοντέλου με το να καταστήσει την οργάνωσή του απλή. Ο χρήστης μπορεί να αλλάξει τη διατύπωση γρήγορα και εύκολα μετατρέποντας ένα γραμμικό πρόβλημα σε μη γραμμικό χωρίς μεγάλη δυσκολία. Η γλώσσα, που χρησιμοποιεί το GAMS, είναι τυπικά παρόμοια με τις συνήθως χρησιμοποιημένες γλώσσες προγραμματισμού, καθιστώντας την οικία σε οποιονδήποτε, που έχει κάποια επαφή με τον προγραμματισμό.

Χρησιμοποιώντας το GAMS, τα στοιχεία εισάγονται μόνο μια φορά με τη γνωστή μορφή καταλόγων και πινάκων. Όλοι οι περιορισμοί του προβλήματος εισάγονται σε μια δήλωση και το GAMS παράγει αυτόματα περιορισμό για κάθε εξίσωση και αφήνει το χρήστη να κάνει τις εξαιρέσεις σε περιπτώσεις όπου η γενικότητα δεν επιδιώκεται.

Ο σχεδιασμός στο GAMS έχει ενσωματώσει τις έννοιες, που προέρχονται από τη θεωρία βάσεων δεδομένων και το μαθηματικό προγραμματισμό και προσπαθεί να συγχωνεύσει αυτές τις ιδέες να ανταποκριθούν στις ανάγκες των σχεδιασμών των μοντέλων. Η σχετική θεωρία βάσεων δεδομένων παρέχει ένα δομημένο πλαίσιο για τις γενικές ικανότητες οργάνωσης και μετασχηματισμού των στοιχείων του μοντέλου και σε συνδυασμό με το μαθηματικό προγραμματισμό που προσφέρει ποικίλες μεθόδους βοηθούν στην επίλυση δύσκολων προβλημάτων. (Rosenthal 2008).

Ο κώδικας GAMS είναι σχεδιασμένος ώστε να:

- ✓ Παρέχει μια αλγεβρικά βασισμένη και υψηλού επιπέδου γλώσσα για την παρουσίαση μεγάλων και πολύπλοκων μοντέλων.
- ✓ Επιτρέπει αλλαγές στο μοντέλο σχεδιασμού με απλότητα και ασφάλεια.
- ✓ Δηλώνονται σαφώς οι αλγεβρικών σχέσεις.
- ✓ Παρέχει ένα περιβάλλον, όπου ο χρήστης να μπορεί να αναπτύξει το μοντέλο του με ένα μικρό σύνολο δεδομένων και στη συνέχεια την επέκτασή του σε ένα ευρύτερο και ορθό πλαίσιο.
- ✓ Επιτρέπει τη χρήση περισσότερων μεταβλητών, εξισώσεων, ονόματα δεικτών, σχολίων και ορισμών δεδομένων, τα οποία συνοδεύονται από υπολογισμούς δημιουργώντας ένα τεκμηριωμένο και αυτόνομο αρχείο.
- ✓ Ενημερώνεται με τις νεότερες και αποτελεσματικότερες εκδόσεις επιλυτών.
- ✓ Αυτοματοποιεί τη μοντελοποίηση με τους υπολογισμούς δεδομένων, την ορθή διόρθωση των δηλώσεων, τον έλεγχο των λαθών, την διασύνδεση με επιλυτές και την αποθήκευση λύσεων.
- ✓ Επιτρέπει τη φορητότητα του μοντέλου σε διαφορετικούς υπολογιστές.
- ✓ Μετατρέπει εύκολα το μοντέλο από γραμμικό σε μη γραμμικό.
- ✓ Διευκολύνει την εισαγωγή και εξαγωγή δεδομένων από και προς διαφορετικά πακέτα υπολογιστών.
- ✓ Επιτρέπει τη χρήση από άτομα ή ομάδες διαφορετικής εμπειρίας.
- ✓ Παρέχει πρότυπα μοντέλα, τα οποία βοηθούν το χρήστη, μέσω βιβλιοθήκης πληροφοριών. (Mc Carl 2008).

4.5.1 Δομή ενός μοντέλου στο GAMS

Σε αυτή τη παράγραφο εξηγούνται τα βασικά συστατικά που αποτελούν τη δομή οποιουδήποτε μοντέλου του GAMS, τα οποία παρουσιάζονται συνοπτικά στον πίνακα(...).

Τα βασικά συστατικά του συστήματος GAMS

Inputs (Είσοδος Δεδομένων):

• Sets – (Σύνολα)

Declaration - (Δήλωση τους)

Assignment of members - (Καθορισμός των μελών τους)

• Data (Parameters, Tables, Scalars) – Δεδομένα (Παράμετροι, Πίνακες, Πίνακες στοιχείων)

Declaration - (Δήλωση τους)

Assignment of values - (Καθορισμός των τιμών τους)

- Variables - (Μεταβλητές)

Declaration - (Δήλωση τους)

Assignment of type - (Καθορισμός του τύπου τους)

- Assignment of bounds and/or initial values (optional) - (Καθορισμός των ορίων ή των αρχικών τιμών)

- Equations – (Εξισώσεις)

Declaration - (Δήλωση τους)

Definition - (Ορισμός τους)

- Model and Solve statements – (Δήλωση Μοντέλου και Επίλυσης)

- Display statement (optional) – (προαιρετική Εμφάνιση δήλωσης)

Outputs (Εξοδος Αποτελεσμάτων):

- Echo Print – (Αποτύπωση προγράμματος)

- Symbol Reference Maps – (Χάρτες Αναφοράς Συμβόλων)

- Equation Listings – (Λίστα Εξισώσεων)

- Status Reports – (Αναφορά Κατάστασης)

- Results – (Αποτελέσματα)

Σε
το

αυτό

σημείο, είναι απαραίτητο να επισημανθούν κάποιες απαραίτητες παρατηρήσεις που αποτελούν και απαράβατους κανόνες του προγράμματος GAMS:

- Ένα μοντέλο στο GAMS είναι μια συλλογή των δηλώσεων στη γλώσσα GAMS. Κάθε οντότητα (εξαρτημένη ή ανεξάρτητη, μεταβλητή ή σταθερή) στο μοντέλο δεν μπορεί να χρησιμοποιηθεί αν δεν έχει δηλωθεί προηγουμένως.
- Οι οντότητες στο GAMS μπορούν να δηλωθούν σχεδόν με οποιοδήποτε τρόπο επιθυμεί ο χρήστης. Έτσι, επιτρέπονται οι δηλώσεις σε πολλαπλές γραμμές, οι κενές γραμμές μεταξύ των δηλώσεων όπως και οι πολλαπλές δηλώσεις ανά γραμμή.
- Η ολοκλήρωση κάθε δήλωσης πρέπει να συνοδεύεται με το σύμβολο ελληνικού ερωτηματικού «;» ('semicolon'). Ο μεταγλωττιστής ('compiler') GAMS δεν διακρίνει κεφαλαία και πεζά γράμματα, έτσι είναι και τα δύο είδη αποδεκτά.
- Τα επεξηγηματικά σχόλια είναι χρήσιμα για την τεκμηρίωση των μαθηματικών μοντέλων. Είναι καλύτερο να ενσωματώνονται μέσα στο ίδιο το μοντέλο παρά να παρουσιάζονται ξεχωριστά. Υπάρχουν δύο τρόποι να παρεμβληθεί η επεξήγηση μέσα σε μια εφαρμογή του GAMS. Καταρχήν, οποιαδήποτε γραμμή που αρχίζει με έναν αστερίσκο (*) στη πρώτη στήλη λαμβάνεται ως γραμμή

σχολίου από το μεταγλωττιστή GAMS. Δεύτερον, ίσως το σημαντικότερο, τα σχόλια μπορούν να παρεμβληθούν μετά από τις δηλώσεις των οντοτήτων του GAMS.

- Η δημιουργία οντοτήτων στο GAMS περιλαμβάνει δύο βήματα: μια δήλωση (declaration) και μια ανάθεση ή καθορισμός (assignment). Η δήλωση περιλαμβάνει την ύπαρξη της οντότητας στο πρόγραμμα δίνοντας ένα όνομα. Η ανάθεση ή καθορισμός δίνει μια συγκεκριμένη τιμή ή μια μορφή. Στην περίπτωση των εξισώσεων, πρέπει να γίνεται δήλωση και ανάθεση σε ξεχωριστές δηλώσεις στο GAMS. Για όλες τις άλλες οντότητες του GAMS, ωστόσο, υπάρχει η επιλογή των δηλώσεων και αναθέσεων στην ίδια δήλωση ή χωριστά.
- Τα ονόματα που δίνονται στις οντότητες του μοντέλου πρέπει να αρχίζουν με γράμμα και μπορούν να ακολουθηθούν μέχρι 31 οποιοδήποτε χαρακτήρες ή ψηφία.
- Όλες οι γραμμές δεν είναι μέρος της γλώσσας GAMS. Δύο ειδικά σύμβολα, ο αστερίσκος «*» και το σύμβολο δολαρίου «\$» μπορούν να χρησιμοποιηθούν στην πρώτη θέση σε μια γραμμή για να δείξει μια μη-γλωσσική γραμμή.

4.5.2 Σχόλια

- Ένα σχόλιο (comment) είναι ένα επεξηγηματικό κείμενο, που δεν υποβάλλεται σε επεξεργασία και δεν διατηρείται από τον υπολογιστή. Υπάρχουν τρεις τρόποι να συμπεριληφθούν τα σχόλια σε ένα πρόγραμμα GAMS, η επιλογή μεταξύ των οποίων είναι ένα θέμα μεμονωμένης προτίμησης ή χρησιμότητας.
- Ο πρώτος, αναφέρθηκε ήδη παραπάνω, πρόκειται να αρχίσει μια γραμμή με έναν αστερίσκο "*" στην πρώτη θέση χαρακτήρα. Οι υπόλοιποι χαρακτήρες στη γραμμή αγνοούνται αλλά τυπώνονται στο αρχείο output.
- Δεύτερον, τα σχόλια μπορούν να παρεμβληθούν μετά από τις δηλώσεις των οντοτήτων του GAMS.
- Ο τρίτος τρόπος πρόκειται να χρησιμοποιήσει τους ειδικούς οριοθέτες (blocks), που αναγκάζουν το GAMS να αγνοήσει ένα ολόκληρο τμήμα του προγράμματος. Το σύμβολο δολαρίου '\$' πρέπει να είναι στην πρώτη θέση. Η αρχή ενός σχολίου πρέπει να ξεκινάει στην πρώτη γραμμή με τη φράση '\$ontext' και να καταλήγει με τη φράση '\$offtext'.

Κεφάλαιο 5

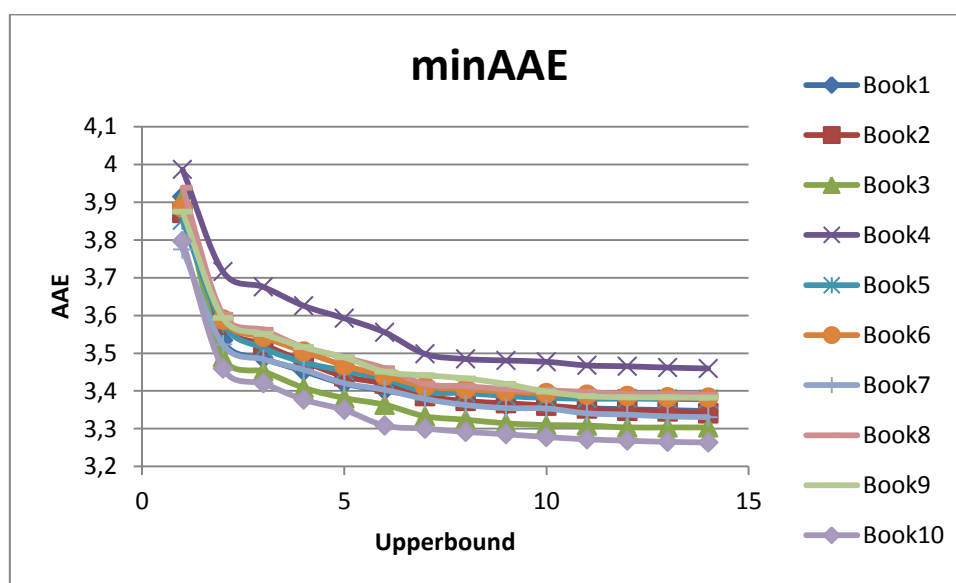
Συζήτηση αποτελεσμάτων

Σε αυτό το κεφάλαιο αναλύουμε τα αποτελέσματα από τους υπολογισμούς που κάναμε στο προηγούμενο κεφάλαιο. Ουσιαστικά, περιγράφουμε το σκεπτικό ανάλυσης των αποτελεσμάτων και τα βήματα τα οποία ακολουθήσαμε για να καταλήξουμε σε κάποια συμπεράσματα, τα οποία θα τα καταγράψουμε στο επόμενο κεφάλαιο.

Για να καταφέρουμε να καταλήξουμε σε κάποια συμπεράσματα συγκεντρώσαμε τα αποτελέσματα μας σε πίνακες, τους οποίους παραθέτουμε παρακάτω αναλυτικά. Αρχικά, θα μελετήσουμε τα αποτελέσματα που προκύπτουν από την ανάλυση των παρατηρήσεων ως προς το ελάχιστο SAE, MAE και στη συνέχεια θα μελετήσουμε τις παρατηρήσεις τις οποίες θεωρήσαμε ως εκτός δείγματος (out of sample) και στη συνέχεια θα μελετήσουμε τις υπόλοιπες παρατηρήσεις.

5.1 Ανάλυση εντός Δείγματος: Η βελτίωση της προσαρμογής ως συνάρτηση του αριθμού μεταβλητών

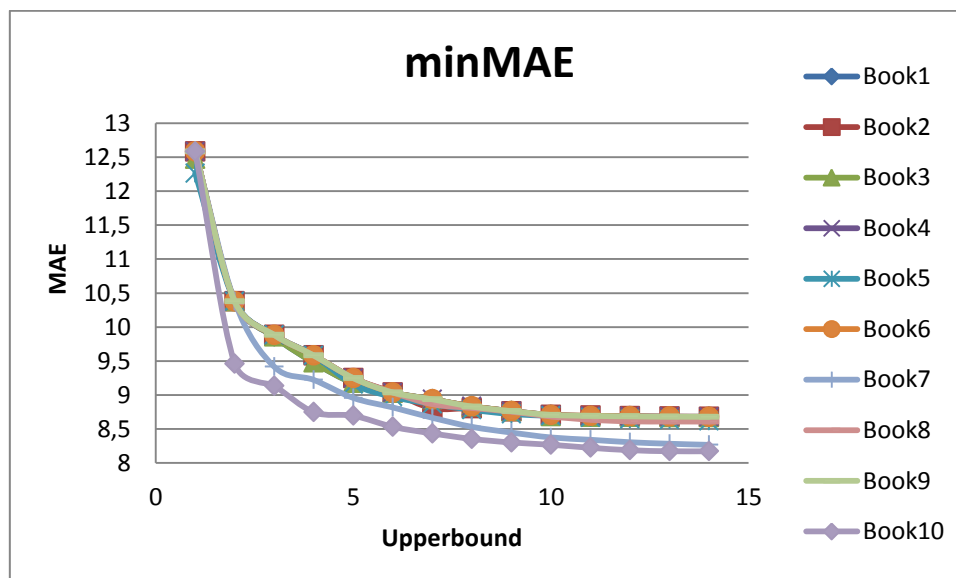
Για να μπορέσουμε να καταλάβουμε πως μεταβάλλεται ο μέγιστος αριθμός μεταβλητών (upperbound) συναρτήσει των σφαλμάτων, στο κάθε μοντέλο, δημιουργήσαμε δύο γραφήματα. Το ένα γράφημα αναφέρεται για τη μελέτη ελαχιστοποίησης του μέσου των απόλυτων σφαλμάτων (minAAE) και το άλλο γράφημα αναφέρεται για τη μελέτη του μέγιστου απόλυτου σφάλματος (minMAE).



Γράφημα 5.1

Σε αυτό το γράφημα μπορούμε εύκολα να παρατηρήσουμε ότι η συνάρτηση έχει μια μορφή η οποία μοιάζει με τη συνάρτηση υπερβολής. Επιπλέον, συγκρίνοντας το κάθε σύνολο δεδομένων με τα υπόλοιπα μπορούμε εύκολα να διαπιστώσουμε ότι δε παρουσιάζουν μεγάλη διαφορά όσον αφορά τα σφάλματα. Τέλος, καταλήγουμε στο συμπέρασμα ότι όσο πιο μεγάλο αριθμό μεταβλητών έχουμε τόσο πιο μικρό είναι το σφάλμα μας, το οποίο είναι και το αναμενόμενο.

Θα πρέπει να αναφέρουμε ότι η κλίμακα του άξονα y είναι μικρή και αυτό οφείλεται στο γεγονός ότι κάθε τιμή του άξονα y έχει διαιρεθεί με το άθροισμα των παρατηρήσεων, δηλαδή με 225 ώστε να έχουμε μια μέση τιμή του σφάλματος ανά μεταβλητή. Ουσιαστικά, στον άξονα y είναι οι τιμές του αθροίσματος των απόλυτων σφαλμάτων (SAE) διαιρεμένες με το άθροισμα αυτών που είναι 225.



Γράφημα 5.2

Σε αυτό το γράφημα όπως αναφέραμε και παραπάνω, παρουσιάζεται η ελαχιστοποίηση του μέγιστου απόλυτου σφάλματος (MAE). Είναι πασιφανές ότι η μορφή της συνάρτησης είναι υπερβολή και τα σφάλματα για κάθε σύνολο δεδομένων δε παρουσιάζουν μεγάλη απόκλιση μεταξύ τους. Με άλλα λόγια, για κάθε όριο μεταβλητών (upperbound) το συγκεκριμένο σφάλμα παρουσιάζει σχεδόν σταθερή τιμή.

5.2 Ανάλυση εντός Δείγματος: Καθορισμός βέλτιστου αριθμού μεταβλητών με βάση τα πληροφοριακά κριτήρια

Για να μπορέσουμε να καταλήξουμε σε κάποια αποτελέσματα για το Κριτήριο Πληροφορίας Akaike και το Κριτήριο Πληροφορίας Bayes, δημιουργήσαμε συγκεντρωτικούς πίνακες τους οποίους και θα αναλύσουμε παρακάτω.

Πίνακας 5.1α : Συγκεντρωτικά Αποτελέσματα των Κριτηρίων Πληροφοριών

Book1																	
																minimum	model
	SSE	5409,636	4332,32	4275,051	4170,122	4112,513	4108,379	4112,422	4182,331	4057,602	4075,368	4079,141	4088,526	4111,895	4155,409	4057,602	9
	SAE	880,7057	795,3465	784,7535	776,8362	769,4451	765,4143	761,4839	758,5015	756,3935	755,2311	754,6595	754,2894	753,9157	753,6653	753,6653	14
SSE	AIC	719,4632	671,4956	670,5014	666,91	665,7801	667,5538	669,7751	675,5678	670,7556	673,7386	675,9468	678,4639	681,7462	686,1148	665,7801	5
	BIC	726,2954	681,7439	684,1659	683,9905	686,2767	691,4665	697,1039	706,3127	704,9166	711,3157	716,94	722,8732	729,5717	737,3563	681,7439	2
	AICc	719,6451	671,7696	670,8868	667,4262	666,4467	668,391	670,8031	676,8073	672,2273	675,4637	677,9468	680,7605	684,3616	689,0713	666,4467	5
SAE	AICsae	618,0804	574,2049	570,1711	567,6081	565,3061	564,9426	564,6259	564,86	565,6076	566,9156	568,5748	570,354	572,1311	573,9816	564,6259	7
	BICsae	624,9126	584,4532	583,8355	584,6886	585,8027	588,8553	591,9547	595,6049	599,7686	604,4927	609,568	614,7634	619,9565	625,2231	583,8355	3
	AICsae	618,2622	574,4789	570,5565	568,1243	565,9728	565,7798	565,6539	566,0994	567,0793	568,6407	570,5748	572,6507	574,7465	576,9381	565,6539	7
Book2																	
																minimum	model
	SSE	548470,6	432905,1	625112,5	184881,7	1210525	1162647	542621,4	595589,3	645090,1	351660,3	354341,1	233617	209016,2	213894,2	184881,7	4
	SAE	10680,84	9330,375	10925,04	5637,226	15620,58	15102,6	8909,492	9374,64	9686,618	7364,521	7373,82	5832,669	5419,416	5670,342	5419,416	13
SSE	AIC	1758,727	1707,489	1792,157	1520,058	1944,855	1937,775	1768,315	1791,271	1811,235	1676,722	1680,431	1588,701	1565,665	1572,856	1520,058	4
	BIC	1765,56	1717,737	1805,821	1537,139	1965,351	1961,687	1795,644	1822,016	1845,396	1714,299	1721,424	1633,11	1613,49	1624,097	1537,139	4
	AICc	1758,909	1707,763	1792,542	1520,575	1945,521	1938,612	1769,343	1792,511	1812,707	1678,447	1682,431	1590,998	1568,28	1575,812	1520,575	4
SAE	AICsae	1741,048	1682,219	1755,221	1459,471	1920,11	1906,935	1671,447	1696,348	1713,08	1591,748	1594,316	1490,808	1459,739	1482,107	1459,471	13
	BICsae	1747,88	1692,467	1768,885	1476,552	1940,606	1930,847	1698,776	1727,093	1747,241	1629,325	1635,309	1535,218	1507,565	1533,348	1476,552	4
	AICsae	1741,23	1682,493	1755,606	1459,987	1920,776	1907,772	1672,475	1697,588	1714,552	1593,473	1596,316	1493,105	1462,355	1485,063	1459,987	4
Book3																	
																minimum	model
	SSE	5320,268	4234,342	4182,501	4053,505	3987,179	3982,911	3872,205	3872,144	3912,801	3869,88	3861,919	3921,938	3919,77	3912,781	3861,919	11
	SAE	879,6903	785,3671	777,0051	767,0825	761,0273	757,0693	750,1026	748,1198	745,939	744,9128	744,5856	743,7073	743,7489	743,9823	743,7073	12
SSE	AIC	715,7152	666,3486	665,577	660,5283	658,8162	660,5753	656,2328	658,2292	662,5794	662,0977	663,6343	669,1042	670,9797	672,5782	656,2328	7
	BIC	722,5474	676,5969	679,2414	677,6088	679,3128	684,488	683,5616	688,9741	696,7404	699,6748	704,6275	713,5135	718,8052	723,8197	676,5969	2
	AICc	715,897	666,6226	665,9623	661,0444	659,4829	661,4125	657,2608	659,4686	664,0511	663,8228	665,6343	671,4008	673,5951	675,5348	657,2608	7
SAE	AICsae	617,5613	568,5229	565,706	561,9223	560,356	560,0095	557,8493	558,6582	559,3445	560,725	562,5273	563,9962	566,0214	568,1626	557,8493	7
	BICsae	624,3935	578,7712	579,3704	579,0028	580,8526	583,9222	585,1781	589,4031	593,5055	598,3021	603,5205	608,4055	613,8468	619,4041	578,7712	2
	AICsae	617,7431	568,7969	566,0913	562,4384	561,0226	560,8467	558,8773	559,8977	560,8162	562,4502	564,5273	566,2929	568,6368	571,1191	558,8773	7

Πίνακας 5.1 β: Συγκεντρωτικά Αποτελέσματα των Πληροφοριακών Κριτηρίων

Book4																		
																	minimum	model
	SSE	5612,713	4659,914	4669,64	4450,357	4392,838	4381,413	4334,816	4293,911	4300,588	4331,52	4285,915	4282,565	4255,949	4335,362	4255,949		13
	SAE	896,9219	836,3221	827,0706	815,8636	808,4042	800,1142	787,0889	784,0431	783,2034	782,4004	780,4999	779,9705	779,1302	778,6517	778,6517		14
SSE	AIC	727,755	687,8967	690,3658	681,5438	680,6168	682,0309	681,6252	681,4919	683,8415	687,454	687,0725	688,8966	689,4938	695,6535	680,6168		5
	BIC	734,5872	698,145	704,0302	698,6243	701,1134	705,9436	708,954	712,2368	718,0025	725,0311	728,0657	733,3059	737,3193	746,895	698,145		2
	AICc	727,9368	688,1706	690,7511	682,06	681,2835	682,8681	682,6532	682,7313	685,3132	689,1792	689,0725	691,1932	692,1092	698,61	681,2835		5
SAE	AICsae	626,2908	596,8111	593,8053	589,666	587,5328	584,8943	579,5083	579,7636	581,2814	582,8198	583,7254	585,4201	586,935	588,6586	579,5083		7
	BICsae	633,123	607,0594	607,4697	606,7465	608,0294	608,807	606,8371	610,5085	615,4424	620,3969	624,7186	629,8294	634,7604	639,9001	606,7465		4
	AICsae	626,4726	597,085	594,1907	590,1822	588,1995	585,7316	580,5364	581,003	582,7531	584,5449	585,7254	587,7167	589,5504	591,6151	580,5364		7
Book5																		
																	minimum	model
	SSE	5225,023	4415,745	4310,403	4248,304	4172,071	4171,361	4042,561	4065,514	3998,121	4081,709	4022,826	4041,566	4035,903	4041,219	3998,121		9
	SAE	866,6127	805,0201	790,6221	782,0854	777,0665	772,1856	764,499	763,7537	761,974	760,9568	760,645	760,6804	760,4796	760,3259	760,3259		14
SSE	AIC	711,6507	675,7871	672,3544	671,0893	669,0152	670,9769	665,92	669,1939	667,4329	674,0884	672,8189	675,8646	677,5491	679,8453	665,92		7
	BIC	718,4829	686,0354	686,0188	688,1698	689,5118	694,8896	693,2488	699,9388	701,5939	711,6655	713,8121	720,2739	725,3745	731,0868	686,0188		3
	AICc	711,8325	676,061	672,7397	671,6054	669,6819	671,8141	666,948	670,4333	668,9046	675,8135	674,8189	678,1612	680,1645	682,8018	666,948		7
SAE	AICsae	610,8213	579,6451	573,5239	570,6386	569,7415	568,906	566,4041	567,9652	568,9154	570,3143	572,1299	574,1508	576,032	577,9411	566,4041		7
	BICsae	617,6535	589,8934	587,1883	587,7191	590,2381	592,8187	593,7329	598,7101	603,0764	607,8914	613,1231	618,5601	623,8574	629,1826	587,1883		3
	AICsae	611,0031	579,919	573,9092	571,1547	570,4082	569,7432	567,4322	569,2047	570,3871	572,0394	574,1299	576,4475	578,6474	580,8976	567,4322		7
Book6																		
																	minimum	model
	SSE	5348,387	4462,898	4434,291	4337,491	4219,032	4216,569	4123,867	4119,616	4111,334	4130,989	4152,209	4158,828	4136,252	4165,998	4111,334		9
	SAE	874,2032	807,7454	797,2441	788,4724	780,2213	774,1326	767,7916	766,1793	764,7924	763,7314	762,763	762,01	761,3391	761,2153	761,2153		14
SSE	AIC	716,9012	678,177	678,7301	675,764	671,5336	673,4023	670,4004	672,1683	673,7155	676,7886	679,9415	682,2998	683,0751	686,6874	670,4004		7
	BIC	723,7334	688,4253	692,3945	692,8445	692,0302	697,315	697,7292	702,9132	707,8765	714,3657	720,9347	726,7091	730,9005	737,9289	688,4253		2
	AICc	717,083	678,4509	679,1154	676,2801	672,2003	674,2395	671,4285	673,4078	675,1872	678,5138	681,9415	684,5965	685,6905	689,6439	671,4285		7
SAE	AICSAE	614,7456	581,1659	577,2772	574,2986	571,5648	570,0392	568,3381	569,3921	570,5768	571,9521	573,3811	574,9367	576,5403	578,4671	568,3381		7
	BICsae	621,5778	591,4142	590,9416	591,3791	592,0614	593,952	595,6669	600,137	604,7378	609,5292	614,3743	619,346	624,3657	629,7086	590,9416		3
	AICsae	614,9274	581,4399	577,6626	574,8148	572,2314	570,8765	569,3661	570,6316	572,0485	573,6772	575,3811	577,2333	579,1557	581,4236	569,3661		7

Πίνακας 5.1 γ: Συγκεντρικά Αποτελέσματα των Πληροφοριακών Κριτηρίων

Book7																	
																minimum	model
	SSE	5178,736	4303,151	4266,08	4123,133	4081,396	4075,151	3955,739	4007,567	3921,208	3905,099	3962,883	3961,129	3936,11	3945,295	3905,099	10
	SAE	849,3238	792,6697	783,9248	777,7245	769,6603	765,7171	760,2287	756,6614	754,8328	754,5654	752,0827	751,1347	750,0258	749,7338	749,7338	14
SSE	AIC	709,6486	669,9755	670,0288	664,3603	664,0711	665,7266	661,035	665,9638	663,0623	664,1361	669,441	671,3414	671,9158	674,4402	661,035	7
	BIC	716,4808	680,2238	683,6932	681,4408	684,5677	689,6393	688,3638	696,7087	697,2233	701,7132	710,4342	715,7507	719,7412	725,6817	680,2238	2
	AICc	709,8304	670,2495	670,4141	664,8764	664,7378	666,5638	662,0631	667,2033	664,534	665,8612	671,441	673,638	674,5312	677,3967	662,0631	7
SAE	AICsae	601,753	572,6878	569,6957	568,1224	565,432	565,1206	563,8835	563,7669	564,6781	566,5187	567,0357	568,468	569,8033	571,628	563,7669	8
	BICsae	608,5852	582,9361	583,3601	585,2029	585,9286	589,0333	591,2123	594,5118	598,8391	604,0958	608,0289	612,8774	617,6287	622,8695	582,9361	5
	AICsae	601,9349	572,9618	570,081	568,6385	566,0987	565,9578	564,9115	565,0064	566,1498	568,2438	569,0357	570,7647	572,4187	574,5845	564,9115	7
Book8																	
																minimum	model
	SSE	5413,096	4476,052	4360,413	4244,171	4191,398	4224,316	4029,144	4011,128	4021,863	4031,706	4046,751	4028,712	4023,484	4075,099	4011,128	8
	SAE	886,1383	810,5613	801,4796	791,0867	785,1052	778,6763	769,1865	767,9533	766,2186	765,3919	764,7131	763,5179	763,4864	762,7505	762,7505	14
SSE	AIC	719,6071	678,8392	674,9499	670,8703	670,0551	673,8153	665,172	666,1636	668,765	671,315	674,1531	675,1478	676,8557	681,7238	665,172	7
	BIC	726,4393	689,0875	688,6143	687,9508	690,5517	697,728	692,5008	696,9086	702,9261	708,8921	715,1463	719,5571	724,6811	732,9653	687,9508	4
	AICc	719,789	679,1131	675,3352	671,3865	670,7217	674,6525	666,2	667,4031	670,2367	673,0401	676,1531	677,4445	679,4711	684,6803	666,2	7
SAE	AICsae	620,8477	582,7319	579,6616	575,7883	574,3728	572,6728	569,1549	570,4328	571,4152	572,9294	574,5302	575,8263	577,8077	579,3738	569,1549	7
	BICsae	627,6799	592,9802	593,326	592,8688	594,8694	596,5855	596,4837	601,1777	605,5762	610,5065	615,5234	620,2356	625,6331	630,6153	592,8688	4
	AICsae	621,0295	583,0059	580,0469	576,3044	575,0394	573,51	570,1829	571,6723	572,8869	574,6545	576,5302	578,1229	580,4231	582,3303	570,1829	7
Book9																	
																minimum	model
	SSE	5437,006	4487,076	4426,052	4251,491	4239,075	4220,755	4144,445	4190,152	4237,057	4132,943	4151,2	4165,906	4167,836	4117,907	4117,907	14
	SAE	871,7641	808,6892	799,1699	791,132	785,1341	776,5379	774,3604	772,5716	769,5416	765,056	762,3804	761,7052	761,6181	761,3378	761,3378	14
SSE	AIC	720,5988	679,3926	678,3116	671,258	672,6	673,6255	671,5203	675,9882	680,4929	676,895	679,8868	682,6824	684,7867	684,075	671,258	4
	BIC	727,431	689,6409	691,976	688,3385	693,0966	697,5382	698,8491	706,7331	714,6539	714,4722	720,88	727,0917	732,6121	735,3165	688,3385	4
	AICc	720,7806	679,6666	678,697	671,7742	673,2667	674,4627	672,5484	677,2276	681,9646	678,6202	681,8868	684,9791	687,402	687,0315	671,7742	4
SAE	AICsae	613,4883	581,6914	578,3629	575,814	574,3894	571,4353	572,1717	573,1309	573,3626	572,7319	573,1553	574,7566	576,7052	578,5395	571,4353	6
	BICsae	620,3205	591,9397	592,0273	592,8945	594,886	595,348	599,5005	603,8758	607,5236	610,309	614,1486	619,1659	624,5306	629,7811	591,9397	2
	AICsae	613,6701	581,9654	578,7482	576,3301	575,056	572,2725	573,1997	574,3704	574,8343	574,457	575,1553	577,0533	579,3205	581,4961	572,2725	6

Πίνακας 5.1 δ: Συγκεντρωτικά Αποτελέσματα των Πληροφοριακών Κριτηρίων

Book10																	
																minimum	model
	SSE	5120	4136,645	3997,001	3890,817	3886,264	3893,786	3817,932	3843,343	3807,326	3894,123	3821,022	3827,069	3830,633	3843,707	3807,326	9
	SAE	854,1801	778,5197	769,5754	759,896	753,7993	744,4176	742,6263	740,6934	739,3248	737,5611	736,5909	735,7032	734,963	734,6049	734,6049	14
SSE	AIC	707,0821	661,0965	655,3698	651,3117	653,0482	655,4833	653,0568	656,5494	656,4309	663,5028	661,2389	663,5947	665,8041	668,5708	651,3117	4
	BIC	713,9143	671,3448	669,0342	668,3922	673,5448	679,396	680,3856	687,2943	690,592	701,0799	702,2321	708,004	713,6295	719,8123	668,3922	4
	AICc	707,2639	661,3705	655,7551	651,8278	653,7149	656,3205	654,0849	657,7889	657,9026	665,2279	663,2389	665,8913	668,4195	671,5273	651,8278	4
SAE	AICsae	604,3188	564,5823	561,3824	557,6865	556,0616	552,4258	553,3416	554,1689	555,3366	556,2618	557,6695	559,1269	560,6739	562,4546	552,4258	6
	BICsae	611,151	574,8306	575,0468	574,7671	576,5582	576,3385	580,6705	584,9138	589,4976	593,8389	598,6627	603,5362	608,4993	613,6961	574,7671	4
	AICsae	604,5006	564,8562	561,7677	558,2027	556,7283	553,263	554,3697	555,4083	556,8083	557,9869	559,6695	561,4235	563,2893	565,4111	553,263	6

Στους παραπάνω πίνακες έχουμε συγκεντρώσει τα σφάλματα των 14 μοντέλων που προκύπτουν από τη σύγκριση των μετρήσεων μας με το μοντέλο που έχει προκύψει από το πρόγραμμα GAMS. Με κίτρινο παρουσιάζουμε τα μοντέλα που παρουσιάζουν τη μικρότερη τιμή σφαλμάτων. Στις γραμμές AIC έως AICsae έχουμε κάνει υπολογισμούς σύμφωνα το Akaike Κριτήριο. Στις γραμμές SAE,SSE κάνουμε τους υπολογισμούς που αναλύσαμε στο κεφάλαιο 3. Στη στήλη model παρουσιάζουμε ποιο μοντέλο είναι αυτό που παρουσιάζει την ελάχιστη τιμή σφαλμάτων.

Θα πρέπει να αναφέρουμε ότι όπου γράφουμε AIC_{SAE} εννοούμε το τύπο υπολογισμού 2.48 του AIC με το SAE και όπου γράφουμε AIC έχει υπολογιστεί με βάση το τύπο 2.47 που περιλαμβάνει το SSE.

Συνοπτικά έχουμε το πίνακα

Πίνακας 5.2

		Book1	Book2	Book3	Book4	Book5	Book6	Book7	Book8	Book9	Book10
		model	model	model	model	model	model	model	model	model	model
	SSE	10	4	11	13	9	9	10	8	14	9
	SAD	14	13	12	14	14	14	14	14	14	14
SSE	AIC	5	4	7	5	7	7	7	7	4	4
	BIC	2	4	2	2	3	2	2	4	4	4
	AICc	5	4	7	5	7	7	7	7	4	4
SAD	AICsad	8	4	7	7	7	7	8	7	6	6
	BICsad	3	4	2	4	3	3	5	4	2	4
	AICcsad	7	4	7	7	7	7	7	7	6	6

Στον οποίο πίνακα παρουσιάζουμε συνοπτικά τα μοντέλα που παρουσιάζουν τα μοντέλα με την ελάχιστη τιμή στη κάθε μια μεταβλητή. Για να μπορέσουμε να καταλήξουμε ποιο μοντέλο παρουσιάζεται περισσότερες στις μεταβλητές που μας ενδιαφέρουν κατασκευάσαμε τους παρακάτω πίνακες.

Για τη μεταβλητή AIC έχουμε:

Πίνακας 5.3

	Αριθμός μεταβλητών	Συχνότητα εμφάνισης
AIC	1	0
	2	0
	3	0
	4	3
	5	2
	6	0
	7	5
	8	0
	9	0
	10	0
	11	0
	12	0
	13	0
	14	0

Όπου στη πρώτη στήλη έχουμε τα μοντέλα και στη δεύτερη στήλη παρουσιάζουμε τη συχνότητα εμφάνισης του μοντέλου αυτού. Για παράδειγμα, τα μοντέλα με 4 μεταβλητές παρουσιάζονται 3 φορές στο σύνολο των αποτελεσμάτων για την AIC μεταβλητή ενώ τα μοντέλα με 14 μεταβλητές δε παρουσιάζονται καμία φορά. Παρατηρώντας το παραπάνω πίνακα καταλήγουμε ότι τα μοντέλα που παρουσιάζουν το ελάχιστο σφάλμα είναι αυτά που χρειάζονται 4,5 και 7 μεταβλητές.

Εκείνο όμως το μοντέλο το οποίο παρουσιάζει τη μεγαλύτερη συχνότητα εμφάνισης είναι το μοντέλο με 7 μεταβλητές.

Για τη μεταβλητή BIC έχουμε:

Πίνακας 5.4

	Αριθμός μεταβλητών	Συχνότητα εμφάνισης
BIC	1	0
	2	5
	3	1
	4	4
	5	0
	6	0
	7	0
	8	0
	9	0
	10	0
	11	0
	12	0
	13	0
	14	0

Όπως και στη μεταβλητή AIC έτσι και σε αυτή τη περίπτωση, στη πρώτη στήλη έχουμε τα μοντέλα με τον αριθμό των μεταβλητών κατά αύξουσα σειρά και στη δεύτερη στήλη έχουμε τη συχνότητα εμφάνισης για το κάθε μοντέλο. Τα μοντέλα με αριθμό μεταβλητών 2,3 και 4 παρουσιάζονται αρκετές φορές. Το μοντέλο που παρουσιάζει τη μεγαλύτερη συχνότητα εμφάνισης είναι το μοντέλο με 2 μεταβλητές.

Για τη μεταβλητή AIC_{sae} έχουμε:

Πίνακας 5.5

	Αριθμός μεταβλητών	Συχνότητα εμφάνισης
AIC_{sae}	1	0
	2	0
	3	0
	4	1
	5	0
	6	2
	7	5
	8	2
	9	0
	10	0
	11	0
	12	0
	13	0
	14	0

Και σε αυτή τη περίπτωση έχουμε στη μια στήλη τα μοντέλα με τον αριθμό μεταβλητών κατά αύξουσα σειρά και στη δεύτερη στήλη εμφανίζεται ο αριθμός επανάληψης του κάθε μοντέλου στο σύνολο των μοντέλων που μελετάμε. Παρατηρώντας το παρακάτω πίνακα τα μοντέλα με 4,6,7 και 8 μεταβλητές παρουσιάζονται στη συγκεκριμένη περίπτωση, όπου με μεγαλύτερη εμφάνιση παρουσιάζεται το μοντέλο με 7 μεταβλητές.

Για το BIC_{sae} έχουμε:

Πίνακας 5.6

	Αριθμός μεταβλητών	Συχνότητα εμφάνισης
BIC_{sae}	1	0
	2	2
	3	3
	4	4
	5	1
	6	0
	7	0
	8	0
	9	0
	10	0
	11	0
	12	0
	13	0
	14	0

Όμοια για το BIC_{sae} ισχύει ότι και για τους υπόλοιπους πίνακες όσον αφορά το τι μας δείχνει ο συγκεκριμένος πίνακας. Τα μοντέλα με αριθμό μεταβλητών 2,3,4,5 παρουσιάζονται τουλάχιστον μια φορά. Συνεπώς, το μοντέλο το οποίο έχει τη μεγαλύτερη συχνότητα εμφάνισης είναι το 4. Με άλλα λόγια, το μοντέλο το οποίο παρουσιάζει τις περισσότερες φορές ελάχιστο σφάλμα είναι το μοντέλο 4.

Στη συνέχεια θα αναλύσουμε τις παρατηρήσεις του εκτός δείγματος, δηλαδή το ‘out of sample’.

5.3 Ανάλυση εκτός Δείγματος-Καθορισμός βέλτιστου αριθμού μεταβλητών

Στη συνέχεια προχωρήσαμε στην ανάλυση εκτός δείγματος. Χρησιμοποιήσαμε δηλαδή τις 225 παρατηρήσεις (90%) για την εξαγωγή των μοντέλων και στη συνέχεια εφαρμόσαμε τα μοντέλα αυτά στις 25 παρατηρήσεις (10%). Για τις 25 αυτές παρατηρήσεις υπολογίσαμε το SAE και MAE, προσεγγίζει καλύτερα τα δεδομένα.

Για να διαπιστώσουμε ποιος είναι ο μέγιστος αριθμός μεταβλητών που πρέπει να χρησιμοποιήσουμε στην ανάλυση μας (upperbound), στον οποίο παρουσιάζεται το ελάχιστο SAE και MAE, παρατηρήσαμε σε κάθε σύνολο δεδομένων ποιο μοντέλο παρουσιάζει το ελάχιστο SAE και

ΜΑΕ, αντίστοιχα. Έτσι λοιπόν προέκυψε ο παρακάτω συγκεντρωτικός πίνακας.

Πίνακας 5.7: Το μοντέλο που παρουσιάζει το ελάχιστο SAE και MAE σε κάθε σύνολο δεδομένων

	SAE	MAE
ds1	1	2
ds2	2	3
ds3	1	8
ds4	2	5
ds5	8	3
ds6	2	6
ds7	6	2
ds8	2	6
ds9	8	2
ds10	9	2

Όπως φαίνεται από το πίνακα, υπάρχουν μοντέλα τα οποία επαναλαμβάνονται πολλές φορές, οπότε κατασκευάσαμε ένα πίνακα όπου θα συγκεντρώνει τα μοντέλα και θα φαίνεται πόσες φορές επαναλαμβάνεται κάθε ένα από αυτό.

Πίνακας 5.8: Πίνακας Συχνοτήτων

upper-bound	SAE	MAE
1	2	0
2	4	4
3	0	2
4	0	0
5	0	1
6	1	2
7	0	0
8	2	1
9	1	0
10	0	0
11	0	0
12	0	0
13	0	0
14	0	0

Αθροίζοντας οριζοντίως τη κάθε γραμμή του παραπάνω πίνακα καταλήγουμε ότι το μοντέλο που παρουσιάζει το ελάχιστο SAE και MAE είναι το μοντέλο 2 δηλαδή αυτό που έχει δύο ανεξάρτητες μεταβλητές οι οποίες στην περίπτωσή μας είναι ABDOM, WEIGHT.

5.4 Σύγκριση μονοκριτηριακού μοντέλου με το πολυκριτηριακό μοντέλο. Ποιο είναι καλύτερο;

Αρχικά, παρατηρήσαμε στο κάθε σύνολο δεδομένων ποιο μοντέλο από τα 11, για κάθε upperbound, παρουσιάζει το ελάχιστο SAE. Στο παρακάτω πίνακα φαίνεται ποιο μοντέλο παρουσιάζει την ελάχιστη τιμή, για κάθε σύνολο δεδομένων (dataset) και για κάθε τιμή της μεταβλητής μέγιστου αριθμού μεταβλητών (upperbound).

Πίνακας 5.9: Πίνακας με τα μοντέλα (1 έως 11) που παρουσιάζουν το καλύτερο SAE out of sample

Πίνακας με τα μοντέλα (1 έως 11) που παρουσιάζουν το καλύτερο SAE out of sample										
	ds1	ds2	ds3	ds4	ds5	ds6	ds7	ds8	ds9	ds10
1	1	10	1	2	1	1	5	1	10	1
2	7	10	1	3	11	1	10	9	11	9
3	7	9	2	10	5	11	9	10	11	11
4	4	4	7	6	10	8	6	9	11	10
5	10	6	11	11	11	11	9	9	10	10
upper bound	10	10	11	11	9	10	10	1	10	11
7	10	10	11	11	10	11	11	6	7	11
8	8	10	11	10	6	10	11	6	7	10
9	8	10	11	11	10	10	10	6	7	10
10	9	8	11	9	7	10	10	5	10	10
11	9	9	11	9	6	10	10	6	11	10
12	9	2	11	10	10	11	10	7	11	10
13	8	8	11	10	9	10	10	7	10	11
14	9	6	11	10	6	8	10	7	11	10

Στη συνέχεια μετρήσαμε τη συχνότητα εμφάνισης του κάθε μοντέλου στο παραπάνω πίνακα. Ωστε να μπορέσουμε να συμπεράνουμε ποιο μοντέλο εμφανίζεται περισσότερο.

Πίνακας 5.10: Πίνακας συχνότητων των μοντέλων

model	συχνότητα εμφάνισης
1	9
2	3
3	1
4	2
5	3
6	11
7	10
8	7
9	16
10	46
11	32

Στο παραπάνω πίνακα βλέπουμε ότι τα μοντέλα 1 και 11, τα οποία είναι και τα μονοκριτηριακά μας μοντέλα δεν εμφανίζονται τόσο συχνά όσο τα πολυκριτηριακά μας μοντέλα, γεγονός το οποίο επιβεβαιώνει ότι το πολυκριτηριακό μοντέλο παρουσιάζει σαφώς καλύτερα αποτελέσματα από το μονοκριτηριακό.

Όμοια μελετήσαμε και το MAE. Δηλαδή, κατασκευάσαμε ένα πίνακα όπου με μπλέ χρώμα παρουσιάζουμε τα σύνολα δεδομένων, με πράσινο χρώμα είναι η μεταβλητή, η οποία εκφράζει το μέγιστο αριθμό μεταβλητών που ορίζουμε κάθε φορά στο πρόγραμμα μας να χρησιμοποιήσει (upper bound) και στα λευκά κελιά υπάρχουν τα μοντέλα τα οποία παρουσιάζουν το καλύτερο MAE στις παρατηρήσεις εκτός δείγματος (out of sample). Παρακάτω παραθέτουμε το πίνακα με τις τιμές των MAE.

Πίνακας 5.11: Πίνακας με τα μοντέλα (1 έως 11) που παρουσιάζουν το καλύτερο MAE out of sample

Πίνακας με τα μοντέλα (1 έως 11) που παρουσιάζουν το καλύτερο MAE out of sample											
	ds1	ds2	ds3	ds4	ds5	ds6	ds7	ds8	ds9	ds10	
upper bound	1	10	11	3	11	1	10	11	5	11	
	2	8	10	9	9	8	9	11	9	10	8
	3	8	11	9	2	8	7	4	10	5	9
	4	10	5	2	10	9	6	1	6	10	7
	5	2	9	3	11	8	11	10	11	10	8
	6	1	11	4	11	5	11	1	11	11	9
	7	9	11	4	11	6	11	11	9	9	10
	8	4	10	11	9	6	11	11	8	8	5
	9	8	10	11	9	6	11	5	10	7	9
	10	7	11	2	9	7	10	3	9	5	9
	11	9	11	11	9	4	5	3	9	11	9
	12	4	11	11	10	2	11	3	9	11	7
	13	8	11	11	10	2	11	1	8	11	9
	14	2	11	11	9	4	11	3	10	11	9

Όπως παρατηρούμε στο παραπάνω πίνακα μεγαλύτερη συχνότητα παρουσιάζουν τα πολυκριτηριακά μοντέλα από τα μονοκριτηριακά. Αυτό γίνεται πιο εμφανή με το παρακάτω πίνακα όπου παρουσιάζουμε τη συχνότητα εμφάνισης του κάθε μοντέλου.

Πίνακας 5.12: Πίνακας συχνότητων των μοντέλων

model	συχνότητα εμφάνισης
1	5
2	7
3	6
4	7
5	8
6	5
7	6
8	12
9	26
10	19
11	39

Όπως βλέπουμε όλα τα μοντέλα εκτός από το 1 και το 11 (μονοκριτηριακά μοντέλα) παρουσιάζουν την υψηλότερη συχνότητα εμφάνισης. Συνεπώς καταλήγουμε στο συμπέρασμα ότι με τα πολυκριτηριακά μοντέλα έχουμε καλύτερα αποτελέσματα.

5.5 Ποιο είναι καλύτερο κριτήριο βελτιστοποίησης η ελαχιστοποίηση του SAE ή του MAE;

Για να μπορέσουμε να απαντήσουμε σε αυτό το ερώτημα παρατηρήσαμε πρώτα πόσες φορές το MAE model δίνει καλύτερα αποτελέσματα ως προς το SAE του εκτός δείγματος (out of sample). Στη πραγματικότητα συγκρίνουμε δηλαδή τα ακραία σημεία του Pareto set, δηλαδή το minSAE (το μοντέλο 1) και minMAE (το μοντέλο 11) για κάθε upper bound σε κάθε σύνολο δεδομένων.

Πίνακας 5.12: Πόσες φορές το MAE model δίνει καλύτερα αποτελέσματα ως προς το SAE του out of sample

ds1	ds2	ds3	ds4	ds5	ds6	ds7	ds8	ds9	ds10
12	6	10	11	4	10	13	3	11	12

Όπως παρατηρούμε το MAE μοντέλο παρουσιάζει καλύτερα ή ίδια αποτελέσματα ως προς το SAE μοντέλο, για τις εκτός δείγματος παρατηρήσεις 92 φορές. Δηλαδή σε μορφή ποσοστού παρουσιάζει στο 66% των περιπτώσεων (=92/140) καλύτερα αποτελέσματα.

Μελετώντας επίσης, το SAE μοντέλο σε σχέση με το MAE μοντέλο, δημιουργείται ο παρακάτω πίνακας. Σε αυτή τη περίπτωση συγκρίνουμε πάλι τα ακραία σημεία του Pareto set αλλά αυτή τη φορά συγκρίνουμε το SAE μοντέλο σε σχέση με το MAE μοντέλο.

Πίνακας 5.12: Πόσες φορές το SAE model δίνει καλύτερα αποτελέσματα ως προς το MAE του out of sample

ds1	ds2	ds3	ds4	ds5	ds6	ds7	ds8	ds9	ds10
10	1	5	4	13	3	8	6	5	10

Όπως και στο προηγούμενο πίνακα, έτσι και σε αυτό μετράμε πόσες φορές το SAE μοντέλο παρουσιάζει καλύτερα ή ίδια αποτελέσματα από το MAE μοντέλο και βρίσκουμε ότι παρουσιάζει 65 φορές καλύτερα αποτελέσματα. Μετατρέποντας αυτόν τον αριθμό σε ποσοστό βρίσκουμε ότι παρουσιάζει στο 46% των περιπτώσεων (=65/140) καλύτερα αποτελέσματα.

Συνεπώς, καταλήγουμε ότι το MAE μοντέλο υπερέχει από το SAE μοντέλο, καθότι δίνει καλύτερα αποτελέσματα σε μεγαλύτερο ποσοστό.

5.6 Ποια είναι καλύτερη προσέγγιση για την επιλογή μεταβλητών: Η βηματική ανάλυση παλινδρόμησης (stepwise regression) ή η ανάλυση παλινδρόμησης με Μαθηματικό Προγραμματισμό για το καθορισμό του βέλτιστου αριθμού μεταβλητών.

Για να συγκρίνουμε αυτές τις δύο μεθόδους, καταστρώσαμε ένα πίνακα όπου παρουσιάζουμε τον αριθμό και τις μεταβλητές τις οποίες χρησιμοποιεί. Για παράδειγμα, σε όλα τα σύνολα δεδομένων με τη μέθοδο προγραμματισμού χρησιμοποιούμε σταθερά 4 μεταβλητές. Οι 4 αυτές μεταβλητές σε κάθε μοντέλο ποικίλουν ανάλογα το μοντέλο, όμως οι 4 πιο συχνά εμφανιζόμενες μεταβλητές είναι WEIGHT, ABDOM, FORARM, WRIST. Ενώ, με τη βηματική παλινδρόμηση υπάρχουν περιπτώσεις που χρησιμοποιεί και 5 μεταβλητές όπως στο σύνολο δεδομένων 5,6 και 8.

Οι μεταβλητές που επιλέγονται σε κάθε σύνολο δεδομένων είναι ίδιες εάν εξαιρέσουμε το σύνολο δεδομένων 5,6,7 και 9.

Τα παραπάνω σχόλια φαίνονται και στους παρακάτω πίνακες.

Πίνακας 5.14: Σύγκριση μεθόδων (Μαθηματικό Προγραμματισμό και Βηματική Παλινδρόμηση)

Τα μοντέλα min SAD των 10 datasets																	
#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST
1	776,7414	10,9205	-46,0661	0	-0,3294	0	0	0	0	1,043	0	0	0	0	0	0,4234	-0,9459
2	782,5179	11,5990	-40,8392	0	-0,3035	0	0	0	0	1,0347	0	0	0	0	0	0,4286	-1,0000
3	766,8880	10,5882	-46,3177	0	-0,3784	0	0	-0,4150	0	1,1164	0	0	0	0	0	0,3004	0
4	815,717	11,1541	-41,1957	0	-0,3092	0	0	0	0	1,0314	0	0	0	0	0	0,4238	-1,2384
5	782,0458	15,3506	-6,0678	0,1025	0	0	0	-0,5073	0	0,788	0	0	0	0	0	0	-1,8148
6	788,4953	15,2471	7,9652	0	0	-0,0933	0	0	0	0,9007	-0,2659	0	0	0	0	0	-1,6127
7	777,6178	10,8803	-42,5792	0	-0,2927	0	0	-0,4549	0	1,0134	0	0	0	0	0	0,3151	0
8	790,9585	10,7958	-42,5169	0	-0,3172	0	0	-0,4647	0	1,0435	0	0	0	0	0	0,2958	0
9	791,0802	11,714	-41,1613	0	-0,3058	0	0	0	0	1,0246	0	0	0	0	0	0,3727	-1,1237
10	759,7776	11,1803	-38,6633	0	-0,3037	0	0	0	0	1,0055	0	0	0	0	0	0,2459	-0,9772

Πίνακας 5.15: Σύγκριση μεθόδων (Μαθηματικό Προγραμματισμό και Βηματική Παλινδρόμηση)

Stepwise Regression																	
	1	2	3	4	5	6	7	8	9	10							
Constant	-35,4613	Constant -32,3357	Constant -35,2151	Constant -34,8725	Constant -24,1568	Constant -24,1568	Constant -32,8874	Constant -30,2307	Constant -33,9754	Constant -30,0114							
Weight	-0,32112	Weight -0,30544	Weight -0,32917	Weight -0,30668	Age 0,064412	Age 0,064412	Weight -0,29648	Weight -0,27713	Weight -0,2975	Weight -0,28882							
Abdom	1,013495	Abdom 1,014914	Abdom 1,020819	Abdom 0,980534	Weight -0,18708	Weight -0,18708	Abdom 0,97688	Neck -0,48051	Abdom 0,996306	Abdom 0,980636							
Forearm	0,446353	Forearm 0,50521	Forearm 0,40657	Forearm 0,551236	Abdom 0,874018	Abdom 0,874018	Forearm 0,452098	Abdom 1,027548	Forearm 0,449953	Forearm 0,410688							
Wrist	-1,42507	Wrist -1,76255	Wrist -1,37142	Wrist -1,51738	Forearm 0,468311	Forearm 0,468311	Wrist -1,48697	Forearm 0,559352	Wrist -1,51512	Wrist -1,64296							
					Wrist -2,11675	Wrist -2,11675		Wrist -1,1514									

Τέλος, από κάθε σύνολο δεδομένων (το ονομάζουμε Book) και για εκείνο το upperbound για το οποίο είχαμε εφαρμόσει την ανάλυση παλινδρόμησης, πήραμε τα αποτελέσματα των μεταβλητών SSE,SAE,MAE για τρία διαφορετικά μοντέλα. Το πρώτο μοντέλο είναι το πρώτο μοντέλο το οποίο προκύπτει από το πρόγραμμα, το δεύτερο μοντέλο είναι το μοντέλο το οποίο εκπροσωπεί όλες τις ελάχιστες τιμές για κάθε μεταβλητή και για το συγκεκριμένο σύνολο δεδομένων και το τρίτο μοντέλο είναι τα αποτελέσματα από την ανάλυση παλινδρόμησης. Παρακάτω φαίνεται όλα όσα περιγράφουμε σε αυτή τη παράγραφο.

Book1				Book2			
	1	min	SR		1	min	SR
SAE	100,3757	98,1539	94,13984	SAE	122,5151	96,7843	97,74653
MAE	9,570649	9,4320	9,070056	MAE	13,59976	8,5772	9,19713
SSE	567,8996	544,5510	506,2224	SSE	930,0265	511,1843	502,8941
Book3				Book4			
	1	min	SR		1	min	SR
SAE	111,368	104,1390	100,8833	SAE	59,73258	59,3365	60,14438
MAE	8,805825	8,6939	8,902688	MAE	7,331347	6,9546	6,722695
SSE	679,0162	602,1905	590,3103	SSE	219,3056	217,6524	221,5838
Book5				Book6			
	1	min	SR		1	min	SR
SAE	92,8970	85,8332	93,50364	SAE	92,5509	84,0773	86,62664
MAE	8,7229	8,0634	8,153339	MAE	6,91148	6,7376	8,432709
SSE	480,7172	424,6672	476,2061	SSE	433,7127	362,2941	405,7942
Book7				Book8			
	1	min	SR		1	min	SR
SAE	98,27063	94,3104	95,40376	SAE	86,6240	79,9062	86,62399
MAE	9,659515	9,6595	10,48899	MAE	10,0212	8,7457	10,0212
SSE	541,2168	511,9950	534,8878	SSE	437,4744	389,2189	437,4744
Book9				Book10			
	1	min	SR		1	min	SR
SAE	83,51268	71,2337	83,28262	SAE	80,87465	76,9082	80,48529
MAE	7,956354	7,0884	7,793134	MAE	7,446682	6,9809	7,310684
SSE	373,2101	319,8470	370,7133	SSE	356,089	335,9245	348,4113

Για να μπορέσουμε να δούμε πόσες φορές η ανάλυση παλινδρόμησης παρουσιάζει καλύτερα αποτελέσματα σε σχέση με το μοντέλο 1 και το μοντέλο min, κατασκευάσαμε ένα πίνακα όπου φαίνεται η συχνότητα εμφάνισης σε κάθε ένα από αυτά

SAE	7
MAE	6
SSE	8

Πίνακας 5.18: Συχνότητα Εμφάνισης του Stepwise Regression καλύτερο από το μοντέλο min

SAE	2
MAE	2
SSE	3

Συνεπώς, από το πίνακα 5.17 καταλήγουμε ότι το μοντέλο Ανάλυσης Παλινδρόμησης (SR) παρουσιάζεται περισσότερες φορές ως το καλύτερο σε σχέση με το μονοκριτηριακό μοντέλο 1. Δηλαδή το μοντέλο Ανάλυσης Παλινδρόμησης (SR) παρουσιάζει αρκετές φορές χαμηλότερη τιμή σε σχέση με το μονοκριτηριακό μοντέλο 1. Επίσης, από το πίνακα 5.18 διαπιστώνουμε ότι το πολυκριτηριακό μοντέλο (Min) παρουσιάζεται περισσότερες φορές σε σχέση με το μοντέλο Ανάλυσης Παλινδρόμησης (SR). Δηλαδή το πολυκριτηριακό μοντέλο αρκετές φορές παρουσιάζει ελάχιστη τιμή σε σχέση με την ανάλυση παλινδρόμησης. Επομένως, το μοντέλο Ανάλυσης Παλινδρόμησης (SR) είναι καλύτερο από το μονόκριτηριακο μοντέλο αλλά όχι και από το πολυκριτηριακό μοντέλο.

Κεφάλαιο 6

Συμπεράσματα

Μετά από όλη την ανάλυση που κάναμε στο παραπάνω κεφάλαιο για τη συγκεκριμένη βάση δεδομένων, καταλήγουμε στα παρακάτω συμπεράσματα.

6.1 Εντός Δείγματος

6.1.1 Καθορισμός βέλτιστου αριθμού μεταβλητών με βάση τα πληροφοριακά κριτήρια

Για να καταφέρουμε να καθορίσουμε το βέλτιστο αριθμό μεταβλητών με βάση τα πληροφοριακά κριτήρια, θα βασιστούμε στα πινακάκια που δημιουργήσαμε κατά την ανάλυση και μελέτη των αποτελεσμάτων μας. Σύμφωνα, λοιπόν, με το πίνακα 5.2 συμπεραίνουμε ότι χρειαζόμαστε 6 έως 8 μεταβλητές όταν υπολογίζουμε με βάση το AICSAD, ενώ όταν υπολογίζουμε με βάση το BICSAD χρειαζόμαστε πολύ περισσότερες μεταβλητές για τον υπολογισμό μας. Παρατηρώντας τα πληροφοριακά κριτήρια (AIC και BIC) τα οποία είναι υπολογισμένα με βάση το SSE, δε παρατηρούμε καμία διαφορά. Και σε αυτή τη περίπτωση ισχύει ότι στη περίπτωση του AIC χρειαζόμαστε λιγότερες μεταβλητές σε σχέση με το πληροφοριακό κριτήριο BIC. Επιπλέον, παρατηρούμε ότι όσον αφορά τη τιμή των πληροφοριακών κριτηρίων βλέπουμε ότι σταθερά η τιμή του BIC δεν υπερβαίνει τη τιμή του AIC. Συνεπώς, βλέπουμε ότι τα αποτελέσματά μας συμφωνούν με τη θεωρία των πληροφοριακών κριτηρίων που βρήκαμε από τη βιβλιογραφία και επιβεβαιώνεται λοιπόν ότι ακολουθούν την αρχή της μετριοπάθειας.

6.2 Εκτός Δείγματος

6.2.1 Καθορισμός βέλτιστου αριθμού μεταβλητών

Ένα σημαντικό συμπέρασμα στο οποίο μπορούμε να οδηγηθούμε μέσα από αυτή τη διπλωματική εργασία είναι ο καθορισμός του βέλτιστου αριθμού μεταβλητών. Αυτό το συμπέρασμα μας δίνει τη δυνατότητα να δούμε πόσο αποτελεσματική είναι η μέθοδος η οποία ακολουθήσαμε για τη μελέτη του δείγματος, λαμβάνοντας ταυτόχρονα και την εγκυρότητα των αποτελεσμάτων. Για να καταλήξουμε στο συμπέρασμα αυτό, αρκεί να παρατηρήσουμε προσεκτικά τους πίνακες 5.7 και 5.8. Έτσι καταλήγουμε ότι ο βέλτιστος αριθμός μεταβλητών ώστε να έχουμε το ελάχιστο SAE και MAE είναι το μοντέλο στο οποίο ορίζουμε ως άνω όριο μεταβλητών τον αριθμό 2. Το συμπέρασμα προκύπτει από το γεγονός ότι σε αυτή τη περίπτωση παρατηρούμε να παρουσιάζονται περισσότερες φορές ελάχιστες τιμές για το SAE και το MAE.

6.2.2 Σύγκριση μονοκριτηριακού μοντέλου με το πολυκριτηριακό μοντέλο. Ποιο είναι καλύτερο;

Απο τη βιβλιογραφία είναι εύκολο να καταλάβουμε, πως τα μοντέλα υπολογισμού που έχουν αναπτυχθεί και που μπορούμε να χρησιμοποιήσουμε είναι πολλά. Εμείς χρησιμοποιήσαμε μερικά απο αυτά. Θέλοντας, λοιπόν να καταλήξουμε ποιο μοντέλο είναι καλύτερο μεταξύ του μονοκριτηριακού και του πολυκριτηριακού βασιστήκαμε στο πίνακα 5.12, τον οποίο κατασκευάσαμε στο προηγούμενο κεφάλαιο ώστε να έχουμε συγκεντρωμένα τα αποτελέσματα απο τη μελέτη των συγκεκριμένων μοντέλων προσαρμογής. Συνεπώς, παρατηρώντας το πίνακα 5.12 βλέπουμε ότι τα ενδιάμεσα μοντέλα, δηλαδή όλα τα μοντέλα εκτός από το 1 και το 14 εμφανίζονται περισσότερες φορές. Συνεπώς καταλήγουμε ότι, τα πολυκριτηριακά μοντέλα (2 έως 13) είναι καλύτερα από τα μονοκριτηριακά, καθώς εμφανίζουν περισσότερες φορές από τα μονοκριτηριακά, ελάχιστη τιμή στις μεταβλητές SAE και MAE.

6.2.3 Ποιο μοντέλο δίνει καλύτερα αποτελέσματα, όσον αφορά το SAE ή το MAE;

Μελετώντας τους πίνακες που προέκυψαν από την επεξεργασία των αποτελεσμάτων καταλήγουμε ότι το MAE μοντέλο παρουσιάζει 92 φορές καλύτερα αποτελέσματα ως προς το SAE μοντέλο, για τις εκτός δείγματος παρατηρήσεις (δηλαδή σε ποσοστό 66% καλύτερα αποτελέσματα) ενώ το SAE μοντέλο παρουσιάζει 65 φορές καλύτερα αποτελέσματα από το MAE μοντέλο (δηλαδή σε ποσοστό 46%). Συνεπώς, το MAE μοντέλο παρουσιάζει καλύτερα αποτελέσματα από το SAE μοντέλο, καθώς παρουσιάζει μεγαλύτερο ποσοστό με τα πιο αξιόπιστα αποτελέσματα Τέλος, είναι αναμενόμενο όσο ο αριθμός μεταβλητών αυξάνεται, μειώνεται το SAE του μοντέλου. Όμοια συμβαίνει και για το MAE μοντέλο, δηλαδή όσο ο αριθμός μεταβλητών αυξάνεται τόσο το MAE του μοντέλου μειώνεται.

6.2.4 Ποια είναι καλύτερη προσέγγιση για την επιλογή μεταβλητών: Η βηματική ανάλυση παλινδρόμησης (stepwise regression) ή η ανάλυση παλινδρόμησης με Μαθηματικό Προγραμματισμό για το καθορισμό του βέλτιστου αριθμού μεταβλητών.

Από το πίνακα 5.17 συμπεραίνουμε ότι το μοντέλο Ανάλυσης Παλινδρόμησης (SR) παρουσιάζεται περισσότερες φορές ως το καλύτερο σε σχέση με το μονοκριτηριακό μοντέλο 1. Δηλαδή το μοντέλο Ανάλυσης Παλινδρόμησης (SR) παρουσιάζει αρκετές φορές χαμηλότερη τιμή σε σχέση με το μονοκριτηριακό μοντέλο 1. Επίσης, από το πίνακα 5.18 συμπεραίνουμε ότι το πολυκριτηριακό μοντέλο (Min) παρουσιάζεται περισσότερες φορές σε σχέση με το μοντέλο Ανάλυσης Παλινδρόμησης (SR). Δηλαδή το πολυκριτηριακό μοντέλο αρκετές φορές παρουσιάζει ελάχιστη τιμή σε σχέση με την ανάλυση παλινδρόμησης. Καταλήγουμε λοιπόν, ότι το μοντέλο Ανάλυσης Παλινδρόμησης (SR) είναι καλύτερο από το μονόκριτηριακό μοντέλο αλλά όχι και από το πολυκριτηριακό μοντέλο.

Βιβλιογραφία

Πραστάκος Γ., Β' Έκδοση, *Διοικητική Επιστήμη- Λήψη Επιχειρηματικών αποφάσεων στην Κοινωνία της Πληροφορικής*

Andre CDS, Narula SC, Elian SN, Tavares RA. An overview of the variables selection methods for the minimum sum of absolute errors regression. *Statistics in Medicine* 2003; 22: 2101-2111

Broersen, P. M. T. (1986). Subset regression with stepwise directed search. *Applied Statistics*, 35(2), 168–177.

Burak Eksioglu, Riza Demirer, Ismail Capar, 2005, Subset selection in multiple linear regression: a new mathematical programming approach, *Computers and Industrial engineering* 49, 155-167

D'Urso P, Santoro A. Goodness of fit and variable selection in the fuzzy multiple linear regression, *Fuzzy Sets and Systems*, 2006; 157 (19): 2627-2647

Erwin Kalvelagen, 2004, *A regression solver for GAMS*

George Mavrotas, Kostas Florios A combined approach for multiple criteria regression and variable selection using multi-objective programming, *Technical Report*

Grechanovsky, E. (1987). Stepwise regression procedures: overview, problems, results, and suggestions. *Annals of New York Academy of Sciences*, 491, 197–232.

Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics*, 32, 1–49.

John Wiley and Sons NY Birkes και Dodge (1993). *Alternative Methods of Regression*, New York

Julian J. Faraway, 2002. *Practical Regression and Anova using R*

J.J.Glen, 2003, An iterative mixed integer programming method for classification accuracy maximizing discriminant analysis, *Computers & Operations Research* 30, 181-198

Kashid, D. N., & Kulakarni, S. R. (2002). A more general criterion for subset selection in multiple linear regression. *Communications in Statistics*

Mandel J. Some thoughts on variable selection in multiple regression. *Journal of Quality Technology* 1989; 21(1): 2-6.

Michael J. Brusco and Stephanie Stahl, *Statistics and Computing*, Branch-and-Bound Applications in Combinatorial Data Analysis 2005, Part III, 187-202
<http://www.springerlink.com/content/1v2162vnj66h24n0/>

- Miller, A. J. (1984). Selection of subsets of regression variables (with discussion). *Journal of the Royal Statistical Society*, 147, 389–425
- Narula, S. C. and Wellington, J. F. &The minimum sum of absolute errors regression: A state of the art survey', *International Statistical Review*, 50, 317-326 (1982)
- Narula, S. C. The minimum sum of absolute errors regression', *Journal of Quality Technology*, 19, 37-45(1987).
- Norman R. Draper & Harry Smith, 1998, *Applied Regression Analysis*, Third Edition
- Samprit Chatterjee, Ali S.Hadi . (2006) Fourth Edition, *Regression Analysis by Example*
- Stock James and Watson Mark, 2005, *An Emprical Comparison of Methods for Forecasting Using many Predictors*
- Subhash C.Narula et Al, 1994, Multivariate multiple linear regression based on the minimum sum of absolute errors criterion, *European Journal of Operational Research* 73, 70-75
- Subhash C.Narula, John F.Wellington, 2006.Multiple Criteria Linear Regression. *European Journal of Operational Research* 181 (2007) 767-772
- Subhash C. Narula et Al, 1999, The minimum sum of absolute errors regression: A robust alternative to the least squares regression, *Statist.Med.* 18, 1401-1417
- Thompson, M. L. (1978). *Selection of variables in multiple regression: Part I. A review and evaluation. International Statistical Review*, 46, 1–49.
- Thompson, M. L. (1978). Selection of variables in multiple regression: Part II. Chosen procedures, computations and examples. *International Statistical Review*, 46, 129–146.
- Xin Yan, Xiao Gang Su, 2009, *Linear Regression Analysis-Theory and Methods*, 31(5), 795–811.

Παράρτημα Ι

Μετρήση	%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist
1	12,3	23	70,0	172,1	23,7	36,2	93,1	85,2	94,5	59	37,3	21,9	32	27,4	17,1
2	6,1	22	78,7	183,5	23,4	38,5	93,6	83	98,7	58,7	37,3	23,4	30,5	28,9	18,2
3	25,3	22	69,9	168,3	24,7	34	95,8	87,9	92,2	59,6	38,9	24	28,8	25,2	16,6
4	10,4	26	83,9	183,5	24,9	37,4	101,8	86,4	101,2	60,1	37,3	22,8	32,4	29,4	18,2
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8
7	19,2	26	82,2	177,2	26,2	36,4	105,1	90,7	100,3	58,4	38,3	22,9	31,9	27,8	17,7
8	12,4	25	79,9	184,2	23,6	37,8	99,6	88,5	97,1	60	39,4	23,2	30,5	29	18,8
9	4,1	25	86,7	188,0	24,6	38,1	100,9	82,5	99,9	62,9	38,3	23,8	35,9	31,1	18,2
10	11,7	23	90,0	186,7	25,8	42,1	99,6	88,6	104,1	63,1	41,7	25	35,6	30	19,2
11	7,1	26	84,6	189,2	23,6	38,5	101,5	83,6	98,2	59,7	39,7	25,2	32,8	29,4	18,5
12	7,8	27	98,1	193,0	26,3	39,4	103,6	90,9	107,7	66,2	39,2	25,9	37,2	30,2	19
13	20,8	32	81,9	176,5	26,3	38,4	102	91,6	103,9	63,4	38,3	21,5	32,5	28,6	17,7
14	21,2	30	93,2	181,0	28,5	39,4	104,1	101,8	108,6	66	41,5	23,7	36,9	31,6	18,8
15	22,1	35	85,2	176,5	27,4	40,5	101,3	96,4	100,1	69	39	23,1	36,1	30,5	18,2
16	20,9	35	73,9	167,6	26,3	36,4	99,1	92,8	99,2	63,1	38,7	21,7	31,1	26,4	16,9
17	29	34	88,9	180,3	27,3	38,9	101,9	96,4	105,2	64,8	40,8	23,1	36,2	30,8	17,3
18	22,9	32	95,0	180,3	29,2	42,1	107,6	97,5	107	66,9	40	24,4	38,2	31,6	19,3
19	16	28	83,4	172,1	28,2	38	106,8	89,6	102,4	64,2	38,7	22,9	37,2	30,5	18,5
20	16,5	33	96,1	186,7	27,6	40	106,2	100,5	109	65,8	40,6	24	37,1	30,1	18,2
21	19,1	28	81,3	172,7	27,3	39,1	103,3	95,9	104,9	63,5	38	22,1	32,5	30,3	18,4
22	15,2	28	91,0	177,2	29,1	41,3	111,4	98,8	104,8	63,4	40,6	24,6	33	32,8	19,9
23	15,6	31	63,7	173,4	21,2	33,9	86	76,4	94,6	57,4	35,3	22,2	27,9	25,9	16,7
24	17,7	32	67,5	177,8	21,4	35,5	86,7	80	93,4	54,9	36,2	22,1	29,8	26,7	17,1
25	14	28	68,7	172,1	23,2	34,5	90,2	76,3	95,8	58,4	35,5	22,9	31,1	28	17,6
26	3,7	27	72,3	181,6	21,9	35,7	89,6	79,7	96,5	55	36,7	22,5	29,9	28,2	17,7
27	7,9	34	59,7	171,5	20,3	36,2	88,6	74,6	85,3	51,7	34,7	21,4	28,7	27	16,5
28	22,9	31	67,2	171,5	22,9	38,8	97,4	88,7	94,7	57,5	36	21	29,2	26,6	17
29	3,7	27	60,5	164,5	22,4	36,4	93,5	73,9	88,5	50,1	34,5	21,3	30,5	27,9	17,2
30	8,8	29	73,0	175,3	23,8	36,7	97,4	83,5	98,7	58,9	35,3	22,6	30,1	26,7	17,6
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4
32	5,7	29	72,8	181,0	22,2	37,3	93,5	84,5	100,6	58,5	38,8	21,5	30,1	26,4	17,9
33	11,8	27	76,3	181,0	23,3	38,1	93	79,1	94,5	57,3	36,2	24,5	29	30	18,8
34	21,3	41	99,2	180,3	30,5	39,8	111,7	100,5	108,3	67,1	44,2	25,2	37,5	31,5	18,7
35	32,3	41	112,3	186,7	32,2	42,1	117	115,6	116,1	71,2	43,3	26,3	37,3	31,7	19,7
36	40,1	49	87,1	165,1	32	38,4	118,5	113,1	113,8	61,9	38,3	21,9	32	29,8	17
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19
38	28,4	50	89,3	173,4	29,7	42,1	105,6	98,8	104,8	66	41,5	24,7	33,2	30,5	19,4
39	35,2	46	164,9	183,5	48,9	51,2	136,2	148,1	147,7	87,3	49,1	29,6	45	29	21,4
40	32,6	50	92,2	170,2	31,8	40,2	114,8	108,1	102,5	61,3	41,1	24,7	34,1	31	18,3
41	34,5	45	119,3	174,6	39,1	43,2	128,3	126,2	125,6	72,5	39,6	26,6	36,4	32,7	21,4
42	32,9	44	93,1	174,9	29,9	36,6	106	104,3	115,5	70,6	42,5	23,7	33,6	28,7	17,4
43	31,6	48	98,5	177,8	31,2	37,3	113,3	111,2	114,1	67,7	40,9	25	36,7	29,8	18,4
44	32	41	96,2	181,6	29,2	41,5	106,6	104,3	106	65	40,2	23	35,8	31,5	18,8
45	7,7	39	56,9	172,7	19,1	31,5	85,1	76	88,2	50	34,7	21	26,1	23,1	16,1
46	13,9	43	74,6	186,1	21,3	35,7	96,6	81,5	97,2	58,4	38,2	23,4	29,7	27,4	18,3
47	10,8	40	60,6	171,5	20,6	33,6	88,2	73,7	88,5	53,3	34,5	22,5	27,9	26,2	17,3
48	5,6	39	67,4	181,0	20,6	34,6	89,8	79,5	92,7	52,7	37,5	21,9	28,8	26,8	17,9
49	13,6	45	61,6	174,0	20,4	32,8	92,3	83,4	90,4	52	35,8	20,6	28,8	25,5	16,3
50	4	47	57,885	169,545	20,2	34	83,4	70,4	87,2	50,6	34,4	21,9	26,8	25,8	16,8
51	10,2	47	71,8	183,5	21,3	34,9	90,2	86,7	98,3	52,6	37,2	22,4	26	25,8	17,3
52	6,6	40	63,2	175,3	20,6	34,3	89,2	77,9	91	51,4	34,9	21	26,7	26,1	17,2

53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9
54	6,3	49	69,3	186,7	19,9	35,1	93,3	79,6	91,6	52,6	37,6	22,6	38,5	27,4	18,5
55	3,9	42	61,9	171,5	21,1	37,8	87,6	77,6	88,6	51,9	34,9	22,5	27,7	27,5	18,5
56	22,6	54	89,9	182,9	26,9	39,9	107,6	100	99,6	57,2	38	22	35,9	30,2	18,9
57	20,4	58	82,4	172,7	27,6	39,1	100	99,8	102,5	62,1	39,6	22,5	33,1	28,3	18,5
58	28	62	91,4	176,5	29,3	40,5	111,5	104,2	105,8	61,8	39,8	22,7	37,7	30,9	19,2
59	31,5	54	91,9	179,7	28,4	40,5	115,4	105,3	97	59,1	38	22,5	31,6	28,8	18,2
60	24,6	61	81,6	167,0	29,2	38,4	104,8	98,3	99,6	60,6	37,7	22,9	34,5	29,6	18,5
61	26,1	62	98,1	186,1	28,2	41,4	112,3	104,8	103,1	61,6	40,9	23,1	36,2	31,8	20,2
62	29,8	56	81,2	174,0	26,8	35,6	102,9	94,7	100,8	60,9	38	22,1	32,5	29,8	18,3
63	30,7	54	87,7	178,4	27,6	38	107,6	102,4	99,4	61	39,4	23,6	32,7	29,9	19,1
64	25,8	61	80,8	170,2	27,9	37,4	105,3	99,7	99,7	60,8	40,1	22,7	33,6	29	18,8
65	32,3	57	93,3	177,8	29,5	40,1	105,3	105,5	108,3	65	41,2	24,7	35,3	31,1	18,4
66	30	55	83,3	171,5	28,3	40,9	103	100,3	104,2	64,8	40,2	22,7	34,8	30,1	18,7
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4
68	13,8	55	70,3	181,6	21,3	36,9	95,4	86,6	91,8	54,3	35,4	21,5	32,8	27,4	18,7
69	6,3	54	70,5	175,9	22,8	37,5	89,3	78,4	96,1	56	37,4	22,4	32,6	28,1	18,1
70	12,9	55	71,2	181,6	21,6	36,3	94,4	84,6	94,3	51,2	37,4	21,6	27,3	27,1	17,3
71	24,3	62	76,0	181,6	23,1	35,5	97,6	91,5	98,5	56,6	38,6	22,4	31,5	27,3	18,6
72	8,8	55	66,6	174,6	21,9	38,7	88,5	82,8	95,5	58,9	37,6	21,6	30,3	27,3	18,3
73	8,5	56	73,0	187,3	20,8	36,4	93,6	82,9	96,3	52,9	37,5	23,1	29,7	27,3	18,2
74	13,5	55	56,8	162,6	21,5	33,2	87,7	76	88,6	50,9	35,4	19,1	29,3	25,7	16,9
75	11,8	61	64,9	167,0	23,3	36,5	93,4	83,3	93	55,5	35,2	20,9	29,4	27	16,8
76	18,5	61	67,3	171,5	22,9	36	91,6	81,8	94,8	54,5	37	21,4	29,3	27	18,3
77	8,8	57	73,8	176,5	23,7	38,7	91,6	78,8	94,3	56,7	39,7	24,2	30,2	29,2	18,1
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8
79	21,5	81	73,2	178,4	23	37,8	96,4	95,4	99,3	53,5	37,5	21,5	31,4	26,8	18,3
80	18,8	66	77,7	175,9	25,1	37,4	102,7	98,6	100,2	56,5	39,3	22,7	30,3	28,7	19
81	31,4	67	74,3	172,1	25,1	38,4	97,7	95,8	97,1	54,8	38,2	23,7	29,4	27,2	19
82	26,8	64	68,2	170,8	23,4	38,1	97,1	89	96,9	54,8	38	22	29,9	25,2	17,7
83	18,4	64	86,4	184,8	25,3	39,3	103,1	97,8	99,6	58,9	39	23	34,3	29,6	19
84	27	70	77,5	177,8	24,5	38,7	101,8	94,9	95	56	36,5	24,1	31,2	27,3	19,2
85	27	72	76,3	175,9	24,7	38,5	101,4	99,8	96,2	56,3	36,6	22	29,7	26,3	18
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2
87	14,9	72	71,6	170,8	24,6	37,7	97,5	88,1	96,9	57,2	37,7	21,8	32,6	28	18,8
88	23,1	64	72,6	167,0	26	36,5	104,3	90,9	93,8	57,8	39,5	23,3	29,2	28,4	18,1
89	8,3	46	80,2	184,2	23,7	38	97,3	86	99,3	61	38,4	23,8	30,2	29,3	18,8
90	14,1	48	79,9	185,4	23,3	36,7	96,7	86,5	98,3	60,4	39,9	24,4	28,8	29,6	18,7
91	20,5	46	80,4	177,8	25,4	37,2	99,7	95,6	102,2	58,3	38,2	22,5	29,1	27,7	17,7
92	18,2	44	81,6	176,5	26,2	39,2	101,9	93,2	100,6	58,9	39,7	23,1	31,4	28,4	18,8
93	8,5	47	75,0	179,1	23,4	37,5	97,2	83,1	95,4	56,9	38,3	22,1	30,1	28,2	18,4
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8
96	17,4	53	101,9	197,5	26,1	41,1	113,2	99,2	107,5	61,7	42,3	23,2	32,9	30,8	20,4
97	9,6	38	85,7	186,1	24,8	37,5	99,1	91,6	102,4	60,6	39,4	22,9	31,6	30,1	18,5
98	11,3	50	73,8	168,9	25,9	38,7	99,4	86,7	96,2	62,1	39,3	23,3	30,6	27,8	18,2
99	17,8	46	71,1	173,4	23,7	35,9	95,1	88,2	92,8	54,7	37,3	21,9	31,6	27,5	18,2
100	22,2	47	89,4	182,9	26,7	40	107,5	94	103,7	62,7	39	22,3	35,3	30,9	18,3
101	21,2	49	90,1	186,7	25,9	40,1	106,5	95	101,7	59	39,4	22,3	32,2	31	18,6
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17
103	20,1	41	78,4	181,0	24	36,3	96,7	89,2	98,3	60	38,4	23,2	31	29,2	18,4
104	22,3	49	89,3	187,3	25,5	40,7	103,5	95,5	101,6	59,1	39,8	25,4	31	30,3	19,7
105	25,4	43	80,4	175,9	26	39,6	104	98,6	99,5	59,5	36,1	22	30,1	27,2	17,7

106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8
107	19,3	43	90,9	186,7	26	38,6	105,2	102,8	103,6	61,2	39,3	23,5	30,5	28,5	18,1
108	18,3	52	92,3	188,6	26	42	110	101,6	100,7	55,8	38,7	23,4	35,1	29,6	19,1
109	17,3	43	88,1	191,8	24	38,5	110,1	88,7	102,1	57,5	40	24,8	35,1	30,7	19,2
110	21,4	40	76,5	175,9	24,7	34,2	97,8	92,3	100,6	57,5	36,8	22,8	32,1	26	17,3
111	19,7	43	77,5	174,0	25,6	37,2	96,3	90,6	99,3	61,9	38	22,3	33,3	28,2	18,1
112	28	43	83,2	177,8	26,3	37,1	108	105	103	63,7	40	23,6	33,5	27,8	17,4
113	22,1	47	80,9	177,8	25,6	40,2	99,7	95	98,6	62,3	38,1	23,9	35,3	31,1	19,8
114	21,3	42	74,0	178,4	23,3	35,3	93,5	89,6	99,8	61,5	37,8	21,9	30,7	27,6	17,4
115	26,7	48	79,6	182,2	24	38	100,7	92,4	97,5	59,3	38,1	21,8	31,8	27,3	17,5
116	16,7	40	71,7	175,9	23,4	36,3	97	86,6	92,6	55,9	36,3	22,1	29,8	26,3	17,3
117	20,1	48	80,5	184,8	23,6	36,8	96	90	99,7	58,8	38,4	22,8	29,9	28	18,1
118	13,9	51	81,3	182,9	24,3	41	99,2	90	96,4	56,8	38,8	23,3	33,4	29,8	19,5
119	25,8	40	86,7	188,0	24,6	38,3	95,4	92,4	104,3	64,6	41,1	24,8	33,6	29,5	18,5
120	18,1	44	85,1	183,5	25,3	38	101,8	87,5	101	58,5	39,2	24,5	32,1	28,6	18
121	27,9	52	93,8	189,2	26,2	40,8	104,3	99,2	104,1	58,5	39,3	24,6	33,9	31,2	19,5
122	25,3	44	84,1	181,6	25,5	39,5	99,2	98,1	101,4	57,1	40,5	23,2	33	29,6	18,4
123	14,7	40	72,8	174,6	23,9	36,9	99,3	83,3	97,5	60,5	38,7	22,6	34,4	28	17,6
124	16	47	68,8	169,5	23,9	36,9	94	86,1	95,2	58,1	36,5	22,1	30,6	27,5	17,6
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18
126	17,5	46	75,8	170,2	26,2	36,6	101	89,9	100	60,7	36	21,9	35,6	30,2	17,6
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2
128	17,4	43	69,1	172,1	23,4	37,5	95,9	78	93,2	53,5	35,8	20,8	33,9	28,2	17,4
129	20,8	40	87,3	186,1	25,2	39,8	103,9	93,5	99,5	61,7	39	21,8	33,3	29,6	18,1
130	14,9	42	75,0	177,2	23,9	38,3	96,2	87	97,8	57,4	36,9	22,2	31,6	27,8	17,7
131	18,1	49	78,0	181,6	23,7	35,5	97,8	90,1	95,8	57	38,7	23,2	27,5	26,5	17,6
132	22,7	40	77,7	179,1	24,3	36,3	94,6	90,3	99,1	60,3	38,5	23	31,2	28,4	17,1
133	23,6	47	89,4	186,1	25,8	37,8	103,6	99,8	103,2	61,2	38,1	22,6	33,5	28,6	17,9
134	26,1	50	71,3	169,5	24,8	37,8	100,4	89,4	92,3	56,1	35,6	20,5	33,6	29,3	17,3
135	24,4	41	76,4	176,5	24,5	36,5	98,4	87,2	98,4	56	36,9	23	34	29,8	18,1
136	27,1	44	84,4	177,2	26,8	37,8	104,6	101,1	102,1	58,9	37,9	22,7	30,9	28,8	17,6
137	21,8	39	75,7	179,7	23,5	37	92,9	86,1	95,6	58,8	36,1	22,4	32,7	28,3	17,1
138	29,4	43	85,2	188,0	24,1	37,7	97,8	98,6	100,6	63,6	39,2	23,8	34,3	28,4	17,7
139	22,4	40	76,4	181,0	23,3	34,3	98,3	88,5	98,3	58,1	38,4	22,5	31,7	27,4	17,6
140	20,4	49	96,6	190,5	26,6	40,8	104,7	106,6	107,7	66,5	42,5	24,5	35,5	29,8	18,7
141	24,9	40	80,2	180,3	24,6	37,4	98,6	93,1	101,6	59,1	39,6	21,6	30,8	27,9	16,6
142	18,3	40	78,7	176,5	25,3	36,5	99,5	93	99,3	60,4	38,2	22	32	28,5	17,8
143	23,3	52	75,8	172,1	25,6	37,5	102,7	91	98,9	57,1	36,7	22,3	31,6	27,5	17,9
144	9,4	23	72,5	183,5	21,6	35,5	92,1	77,1	93,9	56,1	36,1	22,7	30,5	27,2	18,2
145	10,3	23	85,4	196,9	22,1	38	96,6	85,3	102,5	59,1	37,6	23,2	31,8	29,7	18,3
146	14,2	24	70,8	179,7	21,9	35,7	92,7	81,9	95,3	56,4	36,5	22	33,5	28,3	17,3
147	19,2	24	94,7	184,8	27,7	39,2	102	99,1	110,1	71,2	43,5	25,2	36,1	30,3	18,7
148	29,6	25	93,8	177,2	29,8	40,9	110,9	100,5	106,2	68,4	40,8	24,6	33,3	29,7	18,4
149	5,3	25	65,3	184,2	19,3	35,2	92,3	76,5	92,1	51,9	35,7	22	25,8	25,2	16,9
150	25,2	26	101,2	178,4	31,8	40,6	114,1	106,8	113,9	67,6	42,7	24,7	36	30,4	18,4
151	9,4	26	69,1	175,3	22,5	35,4	92,9	77,6	93,5	56,9	35,9	20,4	31,6	29	17,8
152	19,6	26	109,8	189,2	30,7	41,8	108,3	102,9	114,4	72,9	43,5	25,1	38,5	33,8	19,6
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4
154	16,5	27	71,2	170,8	24,4	37,9	94	88,2	95,2	56,8	37,4	22,8	30,6	28,3	17,9
155	21	27	90,9	186,7	26,1	38,2	101,1	100,1	105	62,1	40	24,9	33,7	29,2	19,4
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7
157	31,2	28	93,4	175,3	30,4	38,5	105,6	105	106,4	68,6	40	25,2	35,2	30,7	19,1
158	10	28	82,9	183,5	24,6	37	98,5	90,8	102,5	60,8	38,5	25	31,6	28	18,6

159	12,5	30	62,0	174,6	20,3	35,9	88,7	76,6	89,8	50,1	34,8	21,8	27	34,9	16,9
160	22,5	31	80,5	181,6	24,4	36,2	101,1	92,4	99,3	59,4	39	24,6	30,1	28,2	18,2
161	9,4	31	68,7	183,5	20,4	35	94	81,2	91,5	52,5	36,6	21	27	26,3	16,5
162	14,6	33	89,0	185,4	25,9	38,5	103,8	95,6	105,1	61,4	40,6	25	31,3	29,2	19,1
163	13	33	83,6	174,6	24,4	40,7	98,9	92,1	103,5	64	37,3	23,5	33,5	30,6	19,7
164	15,1	34	63,6	179,1	19,8	36	89,2	83,4	89,6	52,4	35,6	20,4	28,3	26,2	16,5
165	27,3	34	99,3	182,9	29,7	39,5	111,4	106	108,8	63,8	42	23,4	34	31,2	18,5
166	19,2	35	98,5	187,3	28,1	40,5	107,5	95,1	104,5	64,8	41,3	25,6	36,4	33,7	19,4
167	21,8	35	75,5	172,7	25,3	38,5	99,1	90,4	95,6	55,5	34,2	21,9	30,2	28,7	17,7
168	20,3	35	102,0	183,5	30,3	43,9	108,2	100,4	106,8	63,3	41,7	24,6	37,2	33,1	19,8
169	34,3	35	103,6	176,5	33,3	40,4	114,9	115,9	111,9	74,4	40,6	24	36,1	31,8	18,8
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4
171	3,0	35,0	69,1	172,1	23,4	37,0	92,2	81,9	92,8	54,7	36,2	22,1	30,4	27,4	17,7
172	20,5	35	80,5	180,3	24,8	38,4	100,5	90,3	98,7	57,8	37,3	22,4	31	28,7	17,7
173	16,9	36	80,0	181,6	24,3	38,7	98,2	90,3	99,9	59,2	37,7	21,5	32,4	28,4	17,8
174	25,3	36	102,9	182,2	31	41,5	115,3	108,8	114,4	69,2	42,4	24	35,4	21	20,1
175	9,9	37	65,9	175,9	21,3	36	96,8	79,4	89,2	50,3	34,8	22,2	31	26,9	16,9
176	13,1	37	68,6	170,2	23,7	35,3	92,6	83,2	96,4	60	38,1	22	31,5	26,6	16,7
177	29,9	37	109,5	181,6	33,2	42,1	119,2	110,3	113,9	69,8	42,6	24,8	34,4	29,5	18,4
178	22,5	38	85,0	175,9	27,5	38	102,7	92,7	101,9	64,7	39,5	24,7	34,8	30,3	18,1
179	16,9	39	106,6	189,2	29,8	42,8	109,5	104,5	109,9	69,5	43,1	25,8	39,1	32,5	19,9
180	26,6	39	99,5	188,6	28	40	108,5	104,6	109,8	68,1	42,8	24,1	35,6	29	19
181	11,5	40	66,2	170,8	22,7	35,5	95,5	83,6	91,6	54,1	36,2	21,8	31,4	28,3	17,2
182	12,1	40	72,3	177,2	23	35,3	92,3	86,8	96,1	58	39,4	22,7	30	26,4	17,4
183	17,5	40	77,4	188,6	21,8	37,7	98,9	90,4	95,5	55,4	38,9	22,4	30,5	28,9	17,7
184	8,6	40	76,0	181,6	23,1	39,4	89,5	83,7	98,1	57,3	39,7	22,6	32,9	29,3	18,2
185	23,6	41	105,7	188,6	29,7	41,9	117,5	109,3	108,8	67,7	41,3	24,7	37,2	31,8	20
186	20,4	41	95,6	182,9	28,6	38,5	107,4	98,9	104,1	63,5	39,8	23,5	36,4	30,4	19,1
187	20,5	41	91,8	184,2	27	40,8	109,2	98	101,8	62,8	41,3	24,8	36,6	32,4	18,8
188	24,4	41	84,0	173,4	28	38	103,4	101,2	103,1	61,5	40,4	22,9	33,4	29,2	18,5
189	11,4	41	69,5	175,9	22,5	36,4	91,4	80,6	92,3	54,3	36,3	21,8	29,6	27,3	17,9
190	38,1	42	110,9	193,0	29,8	41,8	115,2	113,7	112,4	68,5	45	25,5	37,1	31,2	19,9
191	15,9	42	87,8	179,1	27,4	40,7	104,9	94,1	102,7	60,6	38,6	24,7	34	30,1	18,7
192	24,7	42	102,0	189,9	28,3	38,5	106,7	105,7	111,8	65,3	43,3	26	33,7	29,9	18,5
193	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1
194	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8
195	22	42	70,9	175,3	23,1	35,5	97,8	86	96,2	57,7	38,6	24	31,2	27,3	17,4
196	17,7	42	76,3	181,6	23,1	36,5	92	89,7	101	62,3	38	22,3	30,8	27,8	16,9
197	6,6	42	75,9	184,8	22,3	37,6	94	78	99	57,5	40	22,5	30,6	30	18,5
198	23,6	43	77,5	171,5	26,4	37,4	103,7	89,7	94,2	58,5	39	24,1	33,8	28,8	18,8
199	12,2	43	80,9	178,4	25,4	37,8	102,7	89,2	99,2	60,2	39,2	23,8	31,7	28,4	18,6
200	22,1	43	68,1	175,9	22	35,2	91,1	85,7	96,9	55,5	35,7	22	29,4	26,6	17,4
201	28,7	43	91,0	181,6	27,6	37,9	107,2	103,1	105,5	68,8	38,3	23,7	32,1	28,9	18,7
202	6	44	83,5	188,0	23,7	37,9	100,8	89,1	102,6	60,6	39	24	32,9	29,2	18,4
203	34,8	44	101,2	177,2	32,3	40,9	121,6	113,9	107,1	63,5	40,3	21,8	34,8	30,7	17,4
204	16,6	44	94,8	185,4	27,6	41,9	105,6	96,3	102	63,3	39,8	24,1	37,3	23,1	19,4
205	32,9	44	75,4	166,4	27,2	39,1	100,6	93,9	100,1	58,9	37,6	21,4	33,1	29,5	17,3
206	32,8	47	88,5	184,2	26,1	40,2	102,7	101,3	101,7	60,7	39,4	23,3	36,7	31,6	18,4
207	9,6	47	72,9	178,4	22,9	36	99,8	83,9	91,8	53	36,2	22,5	31,4	27,5	17,7
208	10,8	47	72,5	179,7	22,5	34,5	92,9	84,4	94	56	38,2	22,6	29	26,2	17,6
209	7,1	49	63,8	172,7	21,4	35,8	91,2	79,4	89	51,1	35	21,7	30,9	28,8	17,4
210	27,2	49	98,2	189,2	27,4	40,2	115,6	104	109	63,7	40,3	23,2	36,8	31	18,9
211	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6

212	18,7	50	88,4	179,7	27,4	39	103,7	97,6	104,2	60	40,9	25,5	32,7	30	19
213	19,5	50	78,4	185,4	22,8	37,4	98,7	87,6	96,1	57,1	38,1	21,8	28,6	26,7	18
214	47,5	51	99,4	162,6	37,6	41,2	119,8	122,1	112,8	62,5	36,9	23,6	34,7	29,1	18,4
215	13,6	51	67,8	177,2	21,6	34,8	92,8	81,1	96,3	53,8	36,5	21,5	31,3	26,3	17,8
216	7,5	51	70,1	177,8	22,2	36,9	93,3	81,5	94,4	54,7	39	22,6	27,5	25,9	18,6
217	24,5	52	90,5	182,2	27,2	39,4	106,8	100	105	63,9	39,2	22,9	35,7	30,4	19,2
218	15	53	70,1	175,9	22,7	37,6	93,9	88,7	94,5	53,7	36,2	22	28,5	25,7	17,1
219	12,4	54	69,6	179,1	24,5	38,5	99	91,8	96,2	57,7	38,1	23,9	31,4	29,9	18,9
220	26	54	104,4	183,5	31	42,5	119,9	110,4	105,5	64,2	42,7	27	38,4	32	19,6
221	11,5	54	73,4	171,5	25	37,4	94,2	87,6	95,6	59,7	40,2	23,4	27,9	27	17,8
222	5,2	55	64,6	170,8	22,2	35,2	92,7	82,8	91,9	54,4	35,2	22,5	29,4	26,8	17
223	10,9	55	81,6	174,6	26,8	41,1	106,9	95,3	98,2	57,4	37,1	21,8	34,1	31,1	19,2
224	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8
225	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8
226	25,2	55	90,1	188,6	25,3	38,3	105,3	96,7	106,6	64	42,6	23,4	33,2	30	18,4
227	14,9	56	79,2	176,5	25,4	38,1	104	89,4	98,4	58,4	37,4	22,5	34,6	30,1	18,8
228	17	56	76,2	174,0	25,2	37,4	98,6	93	97	55,4	38,8	23,2	32,4	29,7	19
229	10,6	57	67,1	167,0	24,1	35,2	99,6	86,4	90,1	53	35	21,3	31,7	27,3	16,9
230	16,1	57	82,7	182,2	24,9	39,4	103,4	96,7	100,7	59,3	38,6	22,8	31,8	29,1	19
231	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18
232	26,7	58	73,4	170,8	25,2	35,1	94,9	94,9	100,2	56,8	35,9	21	27,8	26,1	17,6
233	25,8	60	71,6	171,5	24,1	40,4	97,2	93,3	94	54,3	35,7	21	31,3	28,7	18,3
234	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3
235	24,8	62	86,9	183,5	25,8	40,6	104	98,2	101,1	59,3	40,3	23	32,6	28,5	19
236	27,3	63	99,5	176,5	31,9	40,2	117,6	113,8	111,8	63,4	41,1	22,3	35,1	29,6	18,5
237	12,4	64	70,5	176,5	22,6	37,9	95,8	82,8	94,5	61,2	39,1	22,3	29,8	28,9	18,3
238	29,9	65	86,1	167,0	30,9	40,8	106,4	100,5	100,5	59,2	38,1	24	35,9	30,5	19,1
239	17	65	57,9	167,0	20,8	34,7	93	79,7	87,6	50,7	33,4	20,1	28,5	24,8	16,5
240	35	65	101,9	173,4	33,9	38,8	119,6	118	114,3	61,3	42,1	23,4	34,9	30,1	19,4
241	30,4	66	106,3	182,9	31,8	41,4	119,7	109	109,1	63,7	42,4	24,6	35,6	30,7	19,5
242	32,6	67	103,4	184,8	30,3	41,3	115,8	113,4	109,8	65,6	46	25,4	35,3	29,8	19,5
243	29	67	90,6	174,0	29,9	40,7	118,3	106,1	101,6	58,2	38,8	24,1	32,1	29,3	18,5
244	15,2	68	70,6	175,9	22,8	36,3	97,4	84,3	94,4	54,3	37,5	22,6	29,2	27,3	18,5
245	30,2	69	97,8	179,1	30,5	40,8	113,7	107,6	110	63,3	44	22,6	37,5	32,6	18,8
246	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5
247	33,6	72	91,3	177,2	29,1	40,9	108,5	105	104,5	59,6	40,8	23,2	35,2	28,6	20,1
248	29,3	72	84,8	167,6	30,2	38,9	111,1	111,5	101,7	60,3	37,3	21,5	31,3	27,2	18
249	26	72	86,6	179,1	27	38,9	108,3	101,3	97,8	56	41,6	22,7	30,5	29,4	19,8
250	31,9	74	94,2	177,8	29,8	40,8	112,4	108,5	107,1	59,3	42,2	24,6	33,7	30	20,9

Παράρτημα ΙΙ

Κώδικας GAMS

\$ontext

This is a model for the multiobjective variable selection

\$offtext

\$inlinecom []

\$eolcom //

\$STitle Example specific model definitions

SETS

K criteria /1*2/

I observations /1*225/

J parameters (14 explanatory variables) /1*14/

*/AGE, WEIGHT, HEIGHT, BMI, NECK, CHEST, ABDOM, HIP, THIGH, KNEE, ANKLE, BICEPS,
FORARM, WRIST/

;

TABLE X(I,J) the value of the J-th explanatory variable in the I-th observation

\$include "c:\gams\book_2.txt" ;

PARAMETERS

Y(I) the dependent variable

/

\$include "c:\gams\book_2_dep.v.txt" ;

/

M(J) big M for B(J) (upper bound)

dir(k) direction of the objective functions

/

1 -1

2 -1

/;

scalar

upbnd upperbound for number of variables /7/

Y_AVG average value of Y

SST sum of absolute deviation of Y

Rsqr R square

;

POSITIVE VARIABLES

U(I) underestimation in observation I

V(I) overestimation in observation I

Z(K) value of the objective functions

BINARY VARIABLES

D(J) indicates if the explanatory variable J will be used in the model (Dj=1) or not (Dj=0)

FREE VARIABLES

B0 decision variable for constant term

B(J) decision variable for parameter Bj

SAD sum of absolute deviations

MAXDEV maximum deviation

;

EQUATIONS

$$*Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + b_m X_{mi} + u_i - v_i \quad i=1\dots n$$

* becomes with the insertion of binary variables

$$*Y_i = b_0 + \hat{a}_1 b_1 X_{1i} + \hat{a}_2 b_2 X_{2i} + \dots + \hat{a}_m b_m X_{mi} + u_i - v_i \quad i=1\dots n$$

$$* b_j \leq M \hat{a}_j$$

$$* b_j \geq -M \hat{a}_j$$

LINK_Y_XJ(I) equation that links Y_i with X_{ij} (regressors)

AUXP1(J) 1st equation for the bilinear terms connecting \hat{A}_j with D_j

AUXP2(J) 2nd equation for the bilinear terms connecting \hat{A}_j with D_j

NUMVAR calculate the number of variables in the model

MINSAD sum of absolute deviations (to be minimized)

MAXDEVU(I) maximum negative deviation (underestimation)

MAXDEVV(I) maximum positive deviation (overestimation)

EQ_OBJ1 definition of the objective function 1

EQ_OBJ2 definition of the objective function 2

;

$$* Y_i = b_0 + (z_{1p} - z_{1m}) X_{1i} + (z_{2p} - z_{2m}) X_{2i} + \dots + (z_{mp} - z_{mm}) X_{mi} + u_i - v_i$$

$$\text{LINK_Y_XJ(I).. } Y(I) = b_0 + \text{SUM}(J, B(J) * X(I, J)) + U(I) - V(I);$$

$$\text{AUXP1(J).. } B(J) - M(J) * D(J) = 0;$$

$$\text{AUXP2(J).. } B(J) + M(J) * D(J) = 0 \quad ;$$

$$\text{NUMVAR.. } \text{SUM}(J, D(J)) = \text{upbnd};$$

$$\text{MINSAD.. } \text{SAD} = \text{SUM}(I, (U(I) + V(I)));$$

$$\text{MAXDEVU(I).. } \text{MAXDEV} = G = U(I);$$

MAXDEVV(I).. MAXDEV =G= V(I);

EQ_OBJ1.. SAD=e=Z('1');

EQ_OBJ2.. MAXDEV=e=Z('2');

*upper bound for B(J) is derived from slope(Y,Xj)*2

loop (J, M(J)=3);

MODEL VAR_SEL /ALL/ ;

\$STitle eps-constraint method

Set k1(k) the first element of k, km1(k) all but the first elements of k;

k1(k)\$ (ord(k)=1) = yes; km1(k)=yes; km1(k1) = no;

Set kk(k) active objective function in constraint allobj

Parameter

rhs(k) right hand side of the constrained obj functions in eps-constraint

maxobj(k) maximum value from the payoff table

minobj(k) minimum value from the payoff table

scalar

earlyex number of early exits from the loops

elapsed_time elapsed time for payoff and e-constraint

pareto_num number of Pareto optimal solutions

start start time

finish finish time

numloop number of loops in e-constraint

;

Variables

a_objval auxiliary variable for the objective function

obj auxiliary variable during the construction of the payoff table

Positive Variables

sl(k) slack or surplus variables for the eps-constraints

Equations

con_obj(k) constrained objective functions

augm_obj augmented objective function to avoid weakly efficient solutions

allobj all the objective functions in one expression;

con_obj(km1).. z(km1) - dir(km1)*sl(km1) =e= rhs(km1);

* We optimize the first objective function and put the others as constraints

* the second term is for avoiding weakly efficient points

augm_obj..

sum(k1,dir(k1)*z(k1))+1e-3*sum(km1,sl(km1)/(maxobj(km1)-minobj(km1))) =e= a_objval;

allobj.. sum(kk, dir(kk)*z(kk)) =e= obj;

*Model mod_payoff / objcost, objco2, objes, defcap, defdem, allobj / ;

*Model mod_epsmethod / objcost, objco2, objes, defcap, defdem, con_obj, augm_obj / ;

Model mod_payoff / VAR_SEL, allobj / ;

Model mod_epsmethod / VAR_SEL, con_obj, augm_obj / ;

option limrow=0, limcol=0, solprint=off, solvelink=2;


```
option optcr=0.001;
```

```
$ontext
```

```
mod_payoff.optfile=1;
```

```
mod_epsmethod.optfile=1;
```

```
file opt /"C:\Documents and Settings\user\My Documents\gamsdir\projdir\cplex.opt"/;
```

```
*file opt /"C:\windows\gamsdir\cplex.opt"/;
```

```
put opt;
```

```
put 'mipstart 0' /
```

```
put 'mipemphasis 2' /
```

```
putclose opt;
```

```
$offtext
```

```
Parameter
```

```
    payoff(k,k) payoff tables entries;
```

```
Alias(k,kp);
```

```
start=jnow;
```

```
* Generate payoff table applying lexicographic optimization
```

```
loop(kp,
```

```
    kk(kp)=yes;
```

```
    repeat
```

```
        solve mod_payoff using mip maximizing obj;
```

```
        payoff(kp,kk) = z.l(kk);
```

```
        z.fx(kk) = z.l(kk); // freeze the value of the last objective optimized
```

```
        kk(k+1) = kk(k); // cycle through the objective functions
```

```

until kk(kp);

kk(kp) = no;

* release the fixed values of the objective functions for the new iteration

z.up(k) = inf; z.lo(k) =-inf;

);

if (mod_payoff.modelstat<>1 and mod_payoff.modelstat<>8, abort 'no optimal solution for
mod_payoff');

display payoff;

minobj(k)=smin(kp,payoff(kp,k));

maxobj(k)=smax(kp,payoff(kp,k));

option optcr=0.000;

*$set fname p.scr

File fx / c:\windows\gamsdir\sad_mad_225_14.out /;

File rhsx / c:\windows\gamsdir\rhs_sad_mad_225_14.out /;

fx.pw=1000;

*File fx solution points from epsi-method / "%gams.scrdir%%fname%" /;

*FILE fx /D:\research\gams\posix-gams-epscom\econ4.out/;

$if not set gridpoints $set gridpoints 10

Set g grid points /g0*g%gridpoints%/

grid(k,g) grid

Parameter

gridrhs(k,g) rhs of eps-constraint at grid point

maxg(k) maximum point in grid for objective k

posg(k) grid position of objective

```

```

firstOffMax, lastZero some counters

numk(k) ordinal value of k starting with 1

numg(g) ordinal value of g starting with 0;

lastZero=1; loop(km1, numk(km1)=lastZero; lastZero=lastZero+1); numg(g) = ord(g)-1;

grid(km1,g) = yes; // Here we could define different grid intervals for different objectives

maxg(km1) = smax(grid(km1,g), numg(g)); //used if different number of grid points are set
for objective km1

gridrhs(grid(km1,g))$(dir(km1)=-1) = maxobj(km1) - numg(g)/maxg(km1)*(maxobj(km1)-
minobj(km1));

gridrhs(grid(km1,g))$(dir(km1)=1) = minobj(km1) + numg(g)/maxg(km1)*(maxobj(km1)-
minobj(km1));

display gridrhs;

* Walk the grid points and take shortcuts if the model becomes infeasible

posg(km1) = 0;

earlyex=0;

pareto_num=0;

numloop=0;

put fx ' #   SAD   MAD   BO   AGE  WEIGHT  HEIGHT   BMI   NECK   CHEST
ABDOM   HIP   THIGH   KNEE  ANKLE  BICEPS  FORARM  WRIST' / ;

repeat

  numloop=numloop+1;

  rhs(km1) = sum(grid(km1,g)$(numg(g)=posg(km1)), gridrhs(km1,g));

  loop((km1,g)$(numg(g)=posg(km1)),put rhsx gridrhs(km1,g):10:4);

  solve mod_epsmethod maximizing a_objval using mip;

  if (mod_epsmethod.modelstat<>1 and mod_epsmethod.modelstat<>8, // not optimal is in
this case infeasible

```

```

    put rhsx 'INF'/;

    lastZero = 0; loop(km1$(posg(km1)>0 and lastZero=0), lastZero=numk(km1)); //find the
first non-zero posg

    posg(km1$(numk(km1)<=lastZero) = maxg(km1); // skip all solves for more demanding
values of rhs(km1)

    earlyex=earlyex+1

else

    put rhsx /;

    pareto_num:=pareto_num+1;

    put fx pareto_num:4:0;

    loop(k, put fx z.l(k):10:4);

    put B0.L:10:4;

    loop(j, put B.l(J):10:4);

*   loop(j,if (B.L(J)=1, put ord(J):4:0, ' ',X.l(J):6:3));

    put /;

    );

* Proceed forward in the grid

    firstOffMax = 0;

    loop(km1$(posg(km1)<maxg(km1) and firstOffMax=0), posg(km1)=posg(km1)+1;
firstOffMax=numk(km1));

    posg(km1$(numk(km1)<firstOffMax) = 0;

until sum(km1$(posg(km1)=maxg(km1)),1)=card(km1) and firstOffMax=0;

finish=jnow;

putclose rhsx;

put fx 'Number of solves in e-constraint: ', numloop:0:0/;

```

```
put fx 'Number of early exits from the loop: ', earlyex:0:0/;  
*elapsed_time=mod_payoff.resusd+mod_epsmethod.resusd;  
elapsed_time=(finish-start)*86400;  
put fx 'Elapsed time: ',elapsed_time:7:2, ' seconds' / ;  
putclose fx; // close the point file
```

Παράρτημα ΙΙΙ

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 22.70 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST												
1	784,4779	11,5102	-56,9904	0.0000	-0,3871	0.0000	0.0000	0.0000	0.0000	1,0614	0.0000	0.0000	0.0000	0.0000	0.0000	0,3163	0.0000												
2	784,8670	11,3451	-56,2215	0.0000	-0,3789	0.0000	0.0000	0.0000	0.0000	1,0494	0.0000	0.0000	0.0000	0.0000	0.0000	0,3013	0.0000												
3	785,9545	11,1800	-54,9590	0.0000	-0,3701	0.0000	0.0000	0.0000	0.0000	1,0447	0.0000	0.0000	0.0000	0.0000	0.0000	0,2469	0.0000												
4	787,5322	11,0149	-54,0430	0.0000	-0,3587	0.0000	0.0000	0.0000	0.0000	1,0346	0.0000	0.0000	0.0000	0.0000	0.0000	0,2131	0.0000												
5	789,8902	10,8498	-52,9000	0.0000	-0,3482	0.0000	0.0000	0.0000	0.0000	1,0174	0.0000	0.0000	0.0000	0.0000	0.0000	0,1964	0.0000												
6	794,1875	10,6846	-51,0897	0.0000	-0,3643	0.0000	0.0000	0.0000	0.0000	1,0392	0.0000	0.0000	0.0000	0.0000	0.0000	0,1081	0.0000												
7	794,8283	10,5195	-33,1594	0.0000	-0,3155	0.0000	0.0000	0.0000	0.0000	1,0046	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0,8357												
8	795,4667	10,3544	-32,6941	0.0000	-0,3170	0.0000	0.0000	0.0000	0.0000	1,0048	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0,8552												
9	796,2784	10,1893	-34,0283	0.0000	-0,3240	0.0000	0.0000	0.0000	0.0000	1,0096	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0,7792												
10	798,6794	10,0242	-37,8566	0.0000	-0,3610	0.0000	0.0000	0.0000	0.0000	1,0584	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	-0,6611												
11	827,1368	9,8590	-47,3439	-0,0985	-0,4556	0.0000	0.0000	0.0000	0.0000	1,1615	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000												
%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11				
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	14,14742	14,05327	14,06431	13,98366	13,93634	14,23522	15,43519	15,49387	15,40149	15,26516	14,48648		
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	13,70597	13,67669	13,74019	13,76989	13,79792	13,73374	14,7822	14,79269	14,65198	14,16427	13,05709		
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,939674	2,984002	3,056922	3,104339	3,251951	3,259829	4,521698	4,562832	4,436484	3,763256	4,35481		
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	16,278	16,15798	16,16584	16,06364	15,9795	16,36169	16,13514	16,16499	16,21142	16,39252	15,09391		
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	13,93199	13,89854	13,9644	13,99023	14,01259	13,98033	14,50261	14,5029	14,41447	14,03053	14,88383		
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	10,73814	10,70957	10,91313	10,95042	10,98401	11,63183	14,07678	14,16347	14,00342	13,73231	11,90234		
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	13,21461	13,13568	13,05766	13,07056	12,94689	12,8354	13,2237	13,24518	13,171	12,86811	12,11162		
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	19,18222	19,04944	18,96703	18,84985	18,74463	18,71477	18,36983	18,37132	18,36453	18,35493	17,99146		
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	15,03845	14,9595	14,89885	14,83288	14,80285	14,63725	15,16517	15,17757	15,07651	14,72524	12,97041		
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	18,45079	18,3262	18,23099	18,11676	18,02384	17,91954	18,77291	18,79943	18,67447	18,43029	18,94078		
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	10,25802	10,24788	10,29622	10,32023	10,38198	10,31311	11,36766	11,38772	11,2545	10,71051	11,40988		
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	15,88351	15,77993	15,69937	15,59108	15,53048	15,43819	15,1253	15,12875	15,11643	14,98858	15,58713		
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	23,14143	22,99685	23,0535	22,98839	22,85229	23,19092	23,41581	23,42846	23,40765	23,5712	22,68385		
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	17,47312	17,37348	17,36883	17,30663	17,24291	17,31529	17,30957	17,31554	17,2908	17,21025	17,21492		
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	18,41004	18,32878	18,42777	18,42827	18,37959	18,61203	20,16947	20,20335	20,05057	19,80092	18,84723		
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	15,93449	15,89431	15,88848	15,90493	15,91549	15,54493	15,8481	15,8225	15,70205	15,16897	16,44909		
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	14,10255	14,01766	14,00054	13,92714	13,89168	14,0042	14,88495	14,92456	14,83402	14,61008	13,45033		
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20,73312	20,63825	20,76632	20,77412	20,70126	21,01401	21,10606	21,10272	21,08843	21,13056	19,43586		
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	13,83493	13,77602	13,81218	13,78997	13,7839	13,87775	15,23207	15,27021	15,12927	14,76388	14,2799		
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	24,23954	24,09157	24,02323	23,93263	23,79325	23,62287	23,35676	23,334	23,28537	23,2275	24,07757		
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	22,02798	21,89951	21,8626	21,78967	21,679	21,59406	21,25411	21,23537	21,209	21,1608	21,55419		
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	23,33277	23,16557	23,07207	22,93678	22,77866	22,7478	22,39114	22,38658	22,37282	22,45524	23,48837		
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	17,62954	17,52825	17,52948	17,46757	17,40123	17,50041	17,8229	17,83751	17,78635	17,67956	15,70047		
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	19,15398	19,02397	18,91443	18,79475	18,69376	18,53893	19,56565	19,5931	19,44408	19,15994	18,77872		
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	25,53039	25,36971	25,39145	25,31479	25,15352	25,35116	26,11729	26,13197	26,03742	26,08446	28,33139		
																	SAE	100,1420	100,9353	100,9122	101,3257	101,7347	101,9972	97,7289	97,7577	98,5443	101,1334	101,1334	97,7289
																	MAE	8,9705	9,0718	9,0705	9,1324	9,1988	9,0996	8,7771	8,7625	8,8137	8,9204	8,9204	8,7625
																	SSE	582,6459	587,402	587,3318	589,7452	590,6602	598,9271	538,6701	538,3486	546,151	576,53	577,2345	538,3486

1° σελ δεδομένων για Upperbound 3

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 12,10seconds

				-35,4613					-0,32112					1,013495					0,446353	-1,42506917								
				Constant					Weight					Abdom					Forearm	Wrist								
#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST											
1	776,7414	10,9205	-46,0661	0	-0,3294	0	0	0	0	1,043	0	0	0	0	0	0,4234	-0,9459											
2	776,7496	10,7763	-46,5972	0	-0,3331	0	0	0	0	1,0451	0	0	0	0	0	0,438	-0,9339											
3	776,7641	10,632	-46,7702	0	-0,3368	0	0	0	0	1,0499	0	0	0	0	0	0,4583	-0,9651											
4	777,202	10,4878	-43,5242	0	-0,3257	0	0	0	0	1,0359	0	0	0	0	0	0,4296	-1,0723											
5	778,1484	10,3435	-40,8801	0	-0,3161	0	0	0	0	1,02	0	0	0	0	0	0,3813	-1,1012											
6	779,3481	10,1993	-39,0541	0	-0,3109	0	0	0	0	1,018	0	0	0	0	0	0,3942	-1,2368											
7	780,6581	10,055	-37,409	0	-0,3139	0	0	0	0	1,0276	0	0	0	0	0	0,3903	-1,3589											
8	782,0348	9,9108	-36,0803	0	-0,3129	0	0	0	0	1,0297	0	0	0	0	0	0,3894	-1,4489											
9	783,6284	9,7665	-34,9185	0	-0,3098	0	0	0	0	1,0278	0	0	0	0	0	0,3901	-1,5221											
10	785,3096	9,6223	-33,7566	0	-0,3066	0	0	0	0	1,026	0	0	0	0	0	0,3907	-1,5952											
11	787,6673	9,4781	-33,2036	0	-0,3108	0	0	0	0	1,0372	0	0	0	0	0	0,4056	-1,6938											
%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11	Stepwise Regression		
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	13,9572	13,9501	13,9407	14,1725	14,3668	14,3965	14,4770	14,4958	14,4761	14,4700	14,4008	15,15709
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	13,9740	13,9408	13,8956	14,1180	14,2912	14,3208	14,2853	14,2570	14,2259	14,2112	14,0681	14,76221
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,7071	2,6764	2,5986	2,9411	3,2908	3,2957	3,2138	3,1720	3,1456	3,1325	2,8987	3,978638
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	14,4342	14,4503	14,3976	14,4290	14,5530	14,3516	14,2574	14,1345	13,9901	13,8597	13,6518	14,80204
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	13,5540	13,5286	13,4639	13,6178	13,7751	13,7157	13,6097	13,5264	13,4483	13,3866	13,1845	14,04458
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	10,9483	10,8797	10,8232	11,2868	11,7196	11,8292	12,0377	12,1463	12,1964	12,2595	12,2170	12,65522
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	12,9101	12,9275	12,9189	13,0312	13,1092	13,0830	13,0045	12,9366	12,8681	12,8142	12,6624	13,68366
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,2256	18,2544	18,2422	18,2107	18,1866	18,0592	17,9383	17,8169	17,6885	17,5760	17,4090	18,44348
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	15,0830	15,0903	15,0847	15,2026	15,2609	15,2660	15,2046	15,1514	15,0959	15,0562	14,9275	15,80222
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	18,9626	18,9759	19,0106	19,1354	19,1465	19,2237	19,2645	19,2694	19,2490	19,2444	19,2133	19,88389
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	10,3627	10,3357	10,2855	10,5343	10,7488	10,7699	10,7138	10,6778	10,6461	10,6296	10,4583	11,30275
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,9027	14,9318	14,9101	14,9087	14,9228	14,7948	14,6425	14,5112	14,3845	14,2730	14,0726	15,2
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	22,2340	22,2238	22,1891	22,2305	22,2909	22,1819	22,1779	22,1133	22,0098	21,9229	21,8123	22,48703
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	97,1	56,3	38,8	23	29,5	27,9	18,6	14	16,5487	16,5539	16,5149	16,5583	16,6244	16,5059	16,4031	16,2975	16,1837	16,0856	15,9042	16,86695
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	18,9188	18,8755	18,8537	19,1130	19,2948	19,3798	19,4935	19,5407	19,5462	19,5680	19,5342	19,91745
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	16,1238	16,1102	16,0717	16,1621	16,2003	16,1793	16,0294	15,9323	15,8559	15,7974	15,6091	16,42731
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	13,9276	13,9271	13,9156	14,0989	14,2430	14,2528	14,2656	14,2487	14,2098	14,1852	14,0817	14,91758
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	19,5353	19,5059	19,4255	19,4843	19,6080	19,4458	19,3651	19,2589	19,1322	19,0225	18,8298	19,58647
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	14,1751	14,1512	14,1323	14,3892	14,5785	14,6483	14,6999	14,7189	14,7143	14,7248	14,6440	15,29399
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	23,9105	23,9263	23,9263	23,8639	23,7581	23,6781	23,5772	23,4720	23,3574	23,2609	23,1409	23,89634
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	21,3044	21,3161	21,2927	21,2467	21,1979	21,0752	20,9466	20,8218	20,6919	20,5795	20,4152	21,28305
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	22,6035	22,6336	22,6403	22,5704	22,4804	22,3733	22,2889	22,1833	22,0596	21,9527	21,8348	22,71008
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	17,0294	17,0289	17,0004	17,0862	17,1680	17,0983	17,0468	16,9767	16,8903	16,8196	16,6770	17,52994
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	20,0090	20,0220	20,0713	20,2090	20,2017	20,3211	20,3903	20,4178	20,4166	20,4314	20,4330	21,01712
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	25,6655	25,6517	25,6593	25,7414	25,7491	25,7673	25,8522	25,8566	25,8115	25,7838	25,7730	26,16659
SAE	100,375712	100,4801	100,7146	99,18598	98,15391	98,18555	98,53438	98,99286	99,58207	99,99429	99,99429	98,1539	94,13984															
MAE	9,5706492	9,571086	9,599582	9,513833	9,432	9,501736	9,55322	9,623262	9,709706	9,780449	9,780449	9,4320	9,070056															
SSE	567,8995627	568,9871	571,435	555,2778	544,8461	544,551	548,8602	553,4338	558,8779	563,0683	577,5955	544,5510	506,2224															

1° σετ δεδομένων για Upperbound 4

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 13,82 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST												
1	769,3108	11,3949	-38,7249	0,054	-0,2491	0	0	0	0,9628	0	0	0	0	0	0	0,5043	-1,5565												
2	769,3890	11,1730	-39,6308	0,0572	-0,2569	0	0	0	0,9664	0	0	0	0	0	0	0,5044	-1,4964												
3	769,5839	10,9511	-40,3557	0,0556	-0,2669	0	0	0	0,9769	0	0	0	0	0	0	0,5088	-1,4703												
4	769,8756	10,7291	-40,4968	0,054	-0,2691	0	0	0	0,9768	0	0	0	0	0	0	0,5303	-1,4826												
5	770,4518	10,5072	-39,2896	0,0477	-0,2749	0	0	0	0,9839	0	0	0	0	0	0	0,4949	-1,4923												
6	771,8180	10,2853	-35,7448	0,0482	-0,2654	0	0	0	0,9737	0	0	0	0	0	0	0,466	-1,6329												
7	773,8606	10,0633	-33,3433	0,0449	-0,2519	0	0	0	0,9528	0	0	0	0	0	0	0,443	-1,6762												
8	776,4179	9,8414	-32,5835	0,0404	-0,2669	0	0	0	0,9818	0	0	0	0	0	0	0,4544	-1,8132												
9	780,3146	9,6195	-32,6739	0,0312	-0,2775	0	0	0	0,9994	0	0	0	0	0	0	0,4685	-1,8591												
10	782,1181	9,3976	-40,6621	0	-0,3373	0	0	0	1,0166	0	0	0,2621	0	0	0	0,4144	-1,6152												
11	802,7699	9,1756	-50,9690	0	-0,3577	0	0	0	0,9616	0	0	0,6844	0	0	0	0,5232	-1,7222												
%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11				
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	14,2639	14,3486	14,3355	14,3186	14,3410	14,5393	14,5702	14,5249	14,3389	13,9917	13,1557		
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	14,8369	14,8212	14,7043	14,6760	14,5721	14,7204	14,8277	14,5759	14,3057	14,5874	15,5440		
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,4936	2,4675	2,3708	2,3849	2,4546	2,6935	3,0209	2,6922	2,5042	3,4283	4,3055		
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	14,5279	14,7860	14,8133	14,7428	14,6598	14,6221	14,4760	14,1894	13,7936	13,6574	13,0218		
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	13,1503	13,1303	13,0614	13,0476	13,0447	13,0982	13,2219	12,9655	12,8120	13,5424	14,0593		
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	11,7319	11,7873	11,7304	11,6458	11,7966	12,2595	12,4303	12,4352	12,1913	11,7105	10,5699		
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	13,4474	13,5189	13,4644	13,4804	13,3375	13,3692	13,3838	13,1280	12,8685	12,7991	13,0827		
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,3010	18,4316	18,4253	18,4165	18,2581	18,1422	18,0092	17,7675	17,5090	17,8609	18,4085		
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	16,2925	16,3527	16,2633	16,2603	16,0430	16,0889	16,0731	15,8031	15,4763	15,4136	16,3578		
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	18,9958	18,9658	18,9393	18,9872	18,9400	19,0136	19,0218	19,0368	19,0127	19,6840	20,6755		
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	10,2048	10,1483	10,0606	10,0729	10,0927	10,2503	10,4619	10,2401	10,1146	10,7587	11,3537		
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,4971	14,5899	14,5815	14,5984	14,4956	14,3923	14,3515	14,0820	13,8876	14,7829	15,7985		
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	22,3617	22,5033	22,5102	22,4322	22,3746	22,3701	22,2059	22,1027	21,8432	21,7479	21,1908		
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	16,3757	16,4744	16,4573	16,4318	16,3561	16,3146	16,2638	16,0285	15,7960	16,4506	17,0825		
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	19,5472	19,5352	19,4696	19,4232	19,4313	19,6637	19,7102	19,7084	19,5486	19,7216	19,9933		
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	15,8513	15,7505	15,6492	15,6769	15,5839	15,5570	15,6771	15,4080	15,3149	16,4316	18,0397		
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	14,5714	14,6523	14,6089	14,5967	14,5218	14,6529	14,6676	14,5144	14,2624	14,1427	14,1825		
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20,0634	20,2233	20,1823	20,0651	19,9519	19,9428	19,8154	19,5266	19,1174	19,0180	18,7985		
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	14,6035	14,5943	14,5276	14,5206	14,5172	14,7243	14,8319	14,7431	14,5868	15,2423	16,2663		
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	23,5670	23,5811	23,5639	23,5709	23,4424	23,2946	23,1684	23,0428	22,9463	23,3911	23,8076		
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	21,0601	21,1216	21,1017	21,0864	20,9561	20,8187	20,7027	20,4928	20,3095	21,0090	21,8367		
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	21,9591	22,0308	22,0576	22,0693	21,9865	21,8388	21,6866	21,5971	21,5115	22,0253	22,1853		
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	18,0108	18,1474	18,0959	18,0435	17,8686	17,8928	17,7969	17,5416	17,1650	16,9431	17,1853		
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	20,6517	20,6200	20,5713	20,6164	20,5061	20,6040	20,5883	20,6102	20,5379	20,2056	20,1861		
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	24,4331	24,3780	24,4121	24,4164	24,5450	24,5873	24,5454	24,7333	24,8735	26,7334	28,1610		
																	SAE	100,5609	100,6187	101,0476	101,0816	101,2898	100,6382	100,2032	101,3417	102,4551	96,59156	96,59156	96,5916
																	MAE	8,5891938	8,452564	8,504094	8,556474	8,731368	8,707187	8,803134	9,058424	9,434985	9,656871	9,656871	8,4526
																	SSE	560,2066393	561,7362	567,2493	566,3421	570,1443	563,5774	557,2161	573,9323	589,9202	530,5819	474,9033	474,9033

1° σελ δεδομένων για Upperbound 5

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 14.08 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST
1	765,2965	15,4207	-29,3940	0,0504	-0,2238	0	0,294	0	0	0,8811	0	0	0	0	0	0,2216	-1,7145
2	765,3931	14,7819	-30,3782	0,0513	-0,2296	0	0,2701	0	0	0,8885	0	0	0	0	0	0,2682	-1,7153
3	765,5431	14,1432	-32,3167	0,053	-0,2422	0	0,262	0	0	0,8967	0	0	0	0	0	0,3143	-1,6614
4	765,7965	13,5044	-32,1198	0,0519	-0,2428	0	0,2455	0	0	0,8975	0	0	0	0	0	0,3657	-1,7312
5	766,0647	12,8656	-42,9308	0,0637	-0,297	0	0	0	0	0,9409	0	0,2465	0	0	0	0,4391	-1,7085
6	766,1152	12,2269	-44,0166	0,0621	-0,3018	0	0	0	0	0,9433	0	0,2315	0	0	0	0,4979	-1,682
7	766,1889	11,5881	-44,8561	0,06	-0,3033	0	0	0	0	0,9425	0	0,2149	0	0	0	0,5594	-1,6668
8	766,3579	10,9493	-45,4395	0,0574	-0,3064	0	0	0	0	0,9358	0	0,2051	0	0	0	0,5585	-1,5524
9	767,1128	10,3106	-45,8846	0,0597	-0,3338	0	0	0	0	0,9487	0	0,2403	0	0	0	0,5525	-1,5802
10	774,6921	9,6718	-41,1629	0,0412	-0,3173	0	0	0	0	0,9646	0	0,1667	0	0	0	0,5231	-1,6874
11	824,2479	9,0330	-49,2602	0	-0,3469	0	0	0	0	0,9741	0	-0,175	0,9402	0	0	0,6389	-2,085

%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11		
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	14,4712	14,4721	14,4462	14,4238	13,6606	13,6280	13,5599	13,5281	13,6603	13,5659	13,4323
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	15,0399	14,9966	14,8682	14,8398	14,7598	14,6616	14,5605	14,4493	14,3241	14,0153	16,1211
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	3,0641	2,9993	2,8592	2,8571	2,6331	2,5642	2,5025	2,5973	2,5824	2,4724	4,4038
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	14,3337	14,3562	14,4653	14,2970	13,9895	13,9572	13,8507	13,9435	14,1452	13,5175	12,9174
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	13,4101	13,3439	13,2236	13,1568	12,9253	12,8585	12,7856	12,7851	12,6401	12,5246	14,1356
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	13,0931	12,9385	12,7149	12,5774	12,2822	11,9707	11,6169	11,4784	11,8119	11,6276	10,2410
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	14,0042	13,9979	14,0220	14,0094	14,1639	14,1656	14,1425	14,0986	14,2342	13,4805	12,6050
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,2108	18,2424	18,3315	18,2660	18,1314	18,1789	18,1837	18,1781	18,2200	17,5999	18,5998
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	16,1356	16,1800	16,1973	16,2188	16,3125	16,3239	16,3187	16,2203	16,1938	15,5556	16,8808
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	19,0018	19,0386	19,0235	19,1088	19,6111	19,6226	19,6188	19,4677	19,5785	19,2257	20,9858
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	10,3058	10,2736	10,1423	10,1604	10,3403	10,2747	10,2142	10,1956	10,1173	10,0043	11,3723
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,4512	14,4782	14,5497	14,5138	13,9065	14,0072	14,0806	14,1558	14,1034	13,7518	16,3947
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	22,9442	22,8681	22,8389	22,6420	21,9071	21,7992	21,6307	21,5448	21,6247	21,2447	21,2959
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	16,0701	16,0863	16,1067	16,0319	16,2535	16,2268	16,1621	16,1645	16,1935	15,7375	17,2900
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	20,0303	19,9559	19,7924	19,7335	20,3389	20,1200	19,8731	19,6440	19,7923	19,4332	20,0130
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	15,9683	15,9282	15,8112	15,8233	16,6549	16,6066	16,5701	16,4722	16,3016	15,8827	17,8728
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	14,3604	14,4006	14,4037	14,4095	14,8835	14,8297	14,7431	14,6709	14,8356	14,3083	14,1497
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	21,0401	20,8974	20,8202	20,5370	19,6051	19,4339	19,2058	19,1449	19,1320	18,6003	18,8914
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	14,6465	14,6380	14,5378	14,5556	15,8643	15,7057	15,5260	15,3727	15,5849	15,0492	16,1903
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	23,7362	23,7174	23,7163	23,6614	23,9107	23,9126	23,8873	23,7630	23,7190	23,2100	23,5104
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	20,8173	20,8246	20,8436	20,7698	20,6647	20,6829	20,6707	20,6247	20,5132	20,1234	22,2558
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	21,9160	21,9340	21,9962	21,9458	22,3912	22,4118	22,3884	22,3204	22,4282	21,8687	21,7416
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	18,6278	18,5828	18,5858	18,4597	17,7698	17,7187	17,6252	17,5527	17,5947	16,9627	17,5329
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	20,8239	20,8589	20,8417	20,9390	21,3488	21,3570	21,3520	21,1381	21,2463	20,7829	19,9730
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	24,5229	24,4800	24,3719	24,3382	25,3455	25,2096	25,0377	24,8432	25,0177	24,9022	28,5413

SAE	98,5635	98,67709	99,0525472	99,10393	94,56818	95,46331	96,49553	97,37719	96,94688	99,9514	92,28965	92,2897
MAE	8,153476	8,16205	8,2621847	8,244378	8,830196	8,881282	8,974799	9,047335	9,005338	9,637251	9,067104	8,1535
SSE	546,8742	547,8614	553,5111528	553,0095	507,2079	512,4504	519,6624	526,4332	523,0723	565,9448	488,5274	488,5274

1° σελ δεδομένων για Upperbound 6

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 15,15 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST
1	758,5071	12,6425	7,6911	0,0748	0	-0,0806	0	0	0	0,866	-0,3684	0,3263	0	-0,475	0	0,3556	-2,2002
2	758,6135	12,2615	4,9745	0,0759	0	-0,0604	0	0	0	0,8613	-0,3911	0,3692	0	-0,4956	0	0,3763	-2,2521
3	759,0064	11,8805	5,4436	0,0758	0	-0,0609	0	0	0	0,8608	-0,3935	0,3648	0	-0,5171	0	0,3913	-2,2393
4	759,1454	11,0533	-1,8381	0,0923	0	0	0	-0,3025	0	0,8843	-0,3767	0,3899	-0,3761	0	0	0,5976	-2,3037
5	759,1454	11,0533	-1,8381	0,0923	0	0	0	-0,3025	0	0,8843	-0,3767	0,3899	-0,3761	0	0	0,5976	-2,3037
6	759,5923	10,6891	-17,0202	0,0755	-0,1372	0	0	0	0	0,9368	-0,3578	0,4561	0	-0,4105	0	0,4609	-2,102
7	760,2177	10,3567	-23,1408	0,0674	-0,1688	0	0	0	0	0,9388	-0,32	0,4187	0	-0,2858	0	0,4659	-1,8685
8	762,3762	9,9758	-30,6166	0,0515	-0,2156	0	0	-0,1252	0	0,9676	-0,2824	0,3913	0	0	0	0,5369	-1,6792
9	770,9939	9,5948	-32,6956	0,029	-0,2981	0	0	-0,2226	0	1,0057	0	0,1599	0	-0,3195	0	0,3999	-1,3246
10	778,7507	9,2139	-32,9836	0	-0,3412	0	0,2728	-0,3981	0	0,9892	0	0	0,4392	-0,4693	0	0,3422	-1,0928
11	856,6021	8,8329	-39,1646	0	-0,3065	-0,1048	0	-0,208	0	0,9697	0	-0,3551	1,4417	0	0	0,587	-1,7196

%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11	
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	13,807219	13,550086	13,602504	13,92915	13,92915	13,379412	13,309559	13,05911	13,608272	13,28245	12,90544
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	13,302102	13,349918	13,296823	14,69909	14,69909	13,483859	13,566414	13,937	13,632554	13,79362	16,19348
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	1,656431	1,716914	1,7061965	3,08908	3,08908	1,7515252	2,0644508	2,80446	2,3883096	3,191369	4,326602
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	13,301472	12,947608	13,011508	13,5328	13,5328	13,201309	13,385804	13,43406	14,307254	14,55189	13,95873
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	7,060375	6,86743	6,6249675	12,28112	12,28112	7,8448884	9,3156136	12,48876	9,2597532	8,056566	13,83585
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	13,904553	13,728352	13,74325	13,47027	13,47027	13,214257	12,653207	11,7669	12,688009	12,44639	9,629162
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	14,723924	14,641026	14,658771	15,35969	15,35969	14,535363	14,512943	14,65178	12,832609	12,61177	12,62096
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	17,742547	17,549248	17,592781	17,75617	17,75617	17,699178	17,791624	18,0045	17,942491	18,83536	19,96118
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	16,000704	16,024246	16,033731	16,94123	16,94123	16,128506	16,051732	16,07115	14,962386	14,92121	17,13656
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	19,382572	19,426798	19,425473	19,81015	19,81015	19,7891	19,678579	20,02251	18,890334	19,3559	21,1252
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	9,773369	9,990546	10,006999	10,68946	10,68946	10,250541	10,159893	10,18791	10,818356	11,32288	10,47722
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12,057756	11,768534	11,778609	14,45273	14,45273	11,937794	12,504224	14,31629	14,70532	17,12755	19,48942
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	22,83719	22,53655	22,542375	22,35746	22,35746	22,330345	22,22033	21,9831	21,606654	21,95942	22,26168
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14,282323	14,183472	14,18822	14,2725	14,2725	14,771029	15,010208	15,22749	15,814112	16,18044	17,67823
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	20,622562	20,731698	20,690718	20,89884	20,89884	20,988091	20,561704	20,19968	19,754177	19,55747	19,45854
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	15,113461	15,329534	15,254222	16,36346	16,36346	15,7384	15,937429	16,74153	15,511052	15,89045	17,85725
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	14,409677	14,433258	14,464966	15,35489	15,35489	14,851657	14,686287	14,55686	14,591024	14,21335	13,5516
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	20,251646	19,914324	19,921549	18,89137	18,89137	19,405856	19,261783	18,60601	20,17388	20,38309	18,3809
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	15,534629	15,798636	15,775644	16,31791	16,31791	16,483888	16,177261	16,15085	15,359928	15,4402	15,67388
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	23,67035	23,61025	23,61802	23,6965	23,6965	23,77724	23,680061	23,83833	23,999931	24,73483	24,15818
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	19,900423	19,817442	19,8032	20,65728	20,65728	20,107006	20,309944	20,95109	20,310441	21,23746	23,59561
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	21,177397	21,036998	21,026131	21,79918	21,79918	21,630896	21,842554	22,50718	21,826668	21,99164	22,11443
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	12,77154	12,26782	12,051995	18,12505	18,12505	12,87202	14,173442	17,25938	13,681334	12,81702	19,05478
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	22,421775	22,49122	22,516638	22,92042	22,92042	22,513558	22,079632	21,73929	20,412722	19,65855	18,76254
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	24,088495	24,1425	24,072133	24,56366	24,56366	25,070319	25,064384	26,14158	26,046534	28,09165	30,26585

SAE	112,7552	112,4167	113,0884	97,3425	97,3425	107,0428	102,7647	93,9571	101,5205	98,6836	98,6836	93,9571
MAE	13,8285	14,3322	14,5480	8,4750	8,4750	13,7280	12,4266	9,3406	12,9187	13,7830	13,7830	8,4750
SSE	729,210761	733,6486163	744,18354	497,276	497,276	664,16388	612,58893	498,6682	617,63026	621,2723	518,9593	497,2760

1^ο σετ δεδομένων για Upperbound 8

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 9,70 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST												
1	756,4067	11,4202	9,9387	0,0654	0	-0,1004	0	-0,2758	0	0,87	-0,4143	0,3968	0	-0,4296	0	0,3626	-1,5831												
2	756,455	11,1545	9,1187	0,0671	0	-0,0892	0	-0,2262	0	0,8753	-0,4313	0,3978	0	-0,4408	0	0,372	-1,6965												
3	756,5439	10,8888	8,5384	0,0683	0	-0,0789	0	-0,1971	0	0,8817	-0,4445	0,3989	0	-0,4562	0	0,3845	-1,7954												
4	756,7124	10,6231	9,1593	0,0646	0	-0,0794	0	-0,2532	0	0,8964	-0,4515	0,3858	0	-0,4526	0	0,3708	-1,6765												
5	757,0053	10,3574	12,1382	0,0641	0	-0,0962	0	-0,3556	0	0,8935	-0,4578	0,4194	0	-0,4838	0	0,3515	-1,4419												
6	758,5022	10,0917	15,8534	0,0619	0	-0,0995	0	-0,4271	0	0,9359	-0,3985	0,2829	0	-0,5366	0	0,3262	-1,4619												
7	762,135	9,826	-24,0342	0,0471	-0,223	0	0	-0,3168	0	1,0003	-0,1762	0,3173	0	-0,4194	0	0,3958	-1,3554												
8	768,618	9,5604	-34,1814	0,0371	-0,2885	0	0	-0,2231	0	0,9839	0	0,1268	0,1573	-0,3438	0	0,4269	-1,4244												
9	774,5428	9,2947	-31,2906	0,0202	-0,3126	0	0,2722	-0,3819	0	0,9567	0	0	0,4346	-0,4992	0	0,368	-1,2279												
10	785,6728	9,029	-33,9696	0	-0,3451	0	0,3559	-0,4774	-0,0965	1,0142	0	0	0,5481	-0,5301	0	0,4985	-0,9626												
11	833,5847	8,7633	-41,4178	0	-0,3717	-0,109	0	-0,215	0	0,9938	0	-0,2964	1,3252	0	0,2873	0,3966	-1,5475												
%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11				
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	13,39795	13,45986	13,50232	13,46695	13,44433	13,77496	13,28669	13,49245	13,29413	12,93933	13,0175		
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	13,11797	13,11978	13,13285	13,05908	12,91808	12,82077	13,48146	14,08422	13,92996	13,51713	14,78294		
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,502524	2,373392	2,271737	2,312499	2,550807	2,065377	2,274448	2,563166	2,946992	3,096202	3,329782		
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	14,22372	14,04573	13,89582	14,09726	14,64634	15,03527	14,4114	14,29107	14,59855	14,70436	13,2901		
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	7,54001	7,41712	7,248838	7,268955	6,892295	6,123413	7,980036	9,145442	7,526347	7,038317	13,37182		
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	13,74295	13,7442	13,75851	13,7371	14,01627	14,32034	13,12297	12,52546	12,57269	11,52729	9,506722		
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	15,2846	15,18762	15,08885	14,96201	15,27329	14,44077	13,50622	12,84306	12,71794	12,5016	13,24681		
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,51139	18,35069	18,22694	18,30402	18,69949	18,64602	18,22529	18,18463	18,82996	19,03601	19,98741		
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	15,9022	15,93378	15,96151	15,82731	15,79709	15,45425	15,32377	15,42999	15,2497	14,9267	15,94232		
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	19,54176	19,54187	19,55958	19,45895	19,51923	19,01263	19,22173	19,10329	19,18053	19,6418	20,57271		
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	10,11825	10,09123	10,11059	10,12704	10,24004	10,06507	10,9015	10,97216	11,15899	11,63606	10,09082		
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,58358	14,0698	13,70824	14,15069	15,04589	15,40277	14,79853	15,0095	16,78747	18,42375	18,70417		
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	22,96852	22,95978	22,93849	23,02969	23,21295	23,44182	22,13317	21,62264	21,95055	21,24618	22,04734		
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	14,39455	14,34111	14,29416	14,29348	14,37895	14,34918	15,42195	16,06173	16,08005	16,1366	16,72179		
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	20,32108	20,40953	20,50491	20,41154	20,3953	20,20586	20,37083	19,93432	19,60421	19,15604	18,2125		
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	15,68043	15,56484	15,48659	15,4126	15,48205	14,61363	15,72395	15,80847	15,63708	16,16524	17,23214		
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	14,59065	14,58563	14,6025	14,56371	14,74043	14,62955	14,88443	14,6901	14,34277	14,47859	13,13642		
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20,24881	20,16722	20,08514	20,14228	20,38829	20,78292	19,66693	19,62987	20,38726	20,11378			
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	15,88159	15,89652	15,9429	15,86683	16,05166	15,45781	16,13026	15,64703	15,4136	15,73292	15,53458		
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	24,37412	24,19587	24,08439	24,13776	24,50085	24,34917	24,42363	24,05599	24,55222	25,22875	23,94064		
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	20,65779	20,54529	20,46769	20,6066	20,81727	20,74811	20,64966	20,61046	21,05464	21,2675	22,92011		
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	21,9601	21,80383	21,67941	21,72761	22,05321	21,67004	22,01971	21,70109	21,66984	22,32678	21,99025		
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	13,6826	13,42091	13,1408	13,21485	13,19769	12,78718	12,94527	13,74038	12,76659	12,26674	19,01828		
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	22,04908	22,14575	22,23828	22,01327	21,96368	21,40665	21,05125	20,35506	19,75073	19,87838	18,64671		
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	25,42423	25,22349	25,12763	25,37963	25,84415	25,78359	26,5805	26,30535	27,36542	29,05984	29,7845		
																	SAE	105,1299	106,2108	107,3809	106,4361	105,6862	108,9048	100,1771	99,3606	100,5490	99,5517	99,5517	99,3606
																	MAE	12,9174	13,1791	13,4592	13,3852	13,4023	13,8128	13,6547	12,8596	13,8334	14,3333	14,3333	12,8596
																	SSE	629,9861	646,5173	660,5225	655,428	643,0836	703,5187	616,6393	601,4277	637,9245	625,1161	530,4406	530,4406

1ο σετ δεδομένων για Upperbound 9

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 8,54 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST												
1	755,1783	11,2363	15,3268	0,079	0	-0,1262	0	-0,2851	0	0,8753	-0,359	0,3878	0	-0,5503	-0,1531	0,3776	-1,5419												
2	755,1892	10,9835	14,5767	0,0794	0	-0,1183	0	-0,2745	0	0,8783	-0,3744	0,3986	0	-0,5487	-0,1513	0,3932	-1,5986												
3	755,2001	10,7306	13,8266	0,0798	0	-0,1105	0	-0,2638	0	0,8813	-0,3898	0,4094	0	-0,5471	-0,1495	0,4089	-1,6552												
4	755,2128	10,4778	13,4226	0,0812	0	-0,1039	0	-0,2585	0	0,8833	-0,3997	0,4187	0	-0,5574	-0,1475	0,4279	-1,7186												
5	755,7422	10,2249	14,6039	0,0623	0	-0,1179	0	-0,3481	0	0,8883	-0,3902	0,4033	0	-0,5469	-0,1621	0,4548	-1,468												
6	757,3009	9,9721	-12,781	0,0462	-0,1732	0	0,2419	-0,4653	0	0,9724	-0,3463	0,3776	0	-0,4655	0	0,3459	-1,2017												
7	763,6767	9,7192	-22,811	0,0327	-0,2507	0	0,1505	-0,3541	0	1,0035	-0,1581	0,2391	0	-0,3935	0	0,3633	-1,2465												
8	768,8412	9,4664	-32,574	0,0343	-0,3025	0	0,1596	-0,3331	0	0,9521	0	0,1148	0,2656	-0,4919	0	0,3731	-1,1932												
9	776,5104	9,2135	-24,267	0	-0,2959	0	0,4078	-0,4532	-0,0609	1,0012	-0,0953	0	0,4418	-0,5029	0	0,3928	-1,0918												
10	788,682	8,9607	-38,6932	0	-0,3794	0	0,341	-0,49	-0,0898	1,0264	0	0	0,647	-0,5067	0,0887	0,4598	-0,9365												
11	832,6441	8,7078	-33,8769	0	-0,3468	-0,1178	0	-0,2626	-0,1583	1,09	0	-0,3066	1,0962	0	0,3799	0,5347	-1,3784												
%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11				
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	13,59007	13,54793	13,49653	13,49486	13,25698	13,12361	13,43308	13,2269	13,39023	12,79965	13,51629		
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	13,02141	13,02505	13,01807	13,04283	12,89498	12,92467	13,14081	13,76769	13,44579	13,47553	13,75676		
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,129817	2,107906	2,075422	2,050342	2,412712	2,475551	2,312581	2,67047	3,137524	2,95561	2,451602		
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	15,38921	15,23289	15,06745	14,96412	14,99004	14,68176	14,54441	14,60329	14,75029	14,45742	13,68474		
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	5,748065	5,752613	5,746908	5,628863	5,859013	7,1808	8,13717	7,43979	7,384455	7,308497	12,60087		
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	14,71992	14,64541	14,56106	14,55663	14,16084	13,48503	13,10479	12,72215	12,36889	11,19374	8,839218		
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	14,83815	14,813	14,77945	14,78206	14,72909	14,36662	13,44697	12,9236	12,84733	12,66375	13,04503		
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,54473	18,45709	18,36065	18,32105	18,45818	18,6699	18,29345	18,57098	18,93237	19,27141	19,73734		
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	16,00782	16,02784	16,03826	16,09441	15,82613	15,25462	14,92574	15,24188	14,879	14,84152	15,17615		
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	19,20072	19,27275	19,30561	19,37134	19,44443	19,37397	19,05736	18,9806	19,50173	19,74878	20,27098		
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	9,583353	9,608509	9,622653	9,679754	9,873614	10,68801	10,64255	11,08152	11,33462	11,59952	10,20688		
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,56678	14,39124	14,20624	14,07581	15,04375	15,84592	15,37802	15,86478	17,82863	18,59042	19,38592		
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	23,38855	23,31856	23,23971	23,19283	22,99016	22,9005	22,31473	21,71563	22,08391	21,23891	21,26833		
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	14,73215	14,66889	14,59588	14,56114	14,57863	14,75286	15,24917	15,98732	15,75609	16,11855	16,0766		
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	20,80769	20,84725	20,87653	20,95024	20,62229	20,28456	19,90582	19,72444	19,38549	18,92171	17,18791		
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	14,67935	14,70468	14,71987	14,74186	15,07875	15,5197	15,29896	15,61532	15,6713	16,35905	16,58111		
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	14,86495	14,86198	14,84908	14,89823	14,71104	14,69024	14,53602	14,58345	14,30538	14,3538	13,57752		
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20,98427	20,83697	20,6806	20,58841	20,30591	20,13592	19,79495	20,12082	20,26478	19,85655	19,53187		
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	15,54661	15,6153	15,67348	15,77463	15,6076	16,06396	15,50686	15,60224	15,42204	15,94982	15,29102		
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	24,33723	24,26827	24,19034	24,17506	24,50644	24,8364	24,41021	24,42526	24,93393	25,28151	24,1225		
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	20,34154	20,30969	20,26828	20,24543	20,49101	20,98124	20,60978	20,65685	21,25959	21,46475	22,25518		
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	21,88975	21,83119	21,76401	21,72435	22,01865	22,33586	22,08699	21,68068	22,16214	22,38639	22,34168		
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	12,65932	12,56818	12,46813	12,30879	12,25997	13,07675	13,36034	12,54482	12,73833	12,65616	18,86325		
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	21,83899	21,90632	21,96484	22,07605	21,85984	21,33331	20,75381	20,03805	20,05372	19,72549	19,21814		
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	25,06405	25,05608	25,03799	25,04462	25,61383	26,98172	26,58172	26,65566	28,44413	29,50539	29,88738		
																	SAE	112,1465	112,1365	112,1796	112,4970	109,0652	101,2614	101,8714	101,6808	98,8948	99,7141	99,7141	98,8948
																	MAE	13,9407	14,0318	14,1319	14,2912	14,3400	13,5232	13,2397	14,0552	13,8617	13,9438	13,9438	13,2397
																	SSE	720,0191	720,2862	721,9206	725,1875	698,23	620,5398	627,07	642,6343	627,3267	612,1476	540,1411	540,1411

1^o σελ δεδομένων για Upperbound 10

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 8,20 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST
1	754,6832	11,2462	13,571	0,0813	0	-0,1131	0	-0,3044	0	0,8857	-0,3599	0,4143	-0,1492	-0,5031	-0,1461	0,447	-1,5365
2	754,6863	10,9777	-50,8475	0,073	-0,3639	0,2194	1,0304	-0,2907	0	0,8777	-0,3525	0,4632	0	-0,5912	0	0,3247	-1,5129
3	754,7151	10,7349	-11,0061	0,071	-0,141	0	0,3773	-0,2837	0	0,8847	-0,3719	0,4463	0	-0,5324	-0,1878	0,387	-1,3577
4	754,732	10,4793	-8,3645	0,0822	-0,131	0	0,3887	-0,3205	0	0,8743	-0,3685	0,454	0	-0,581	-0,1842	0,4139	-1,4921
5	755,0598	10,2237	-7,2782	0,0716	-0,1213	0	0,3929	-0,347	0	0,867	-0,3925	0,4366	0	-0,5279	-0,1275	0,4041	-1,4462
6	756,3984	9,968	-46,1687	0,0649	-0,3623	0,1909	0,9304	-0,5012	0	0,9262	-0,3088	0,3868	0	-0,5083	0	0,3155	-1,2252
7	760,7164	9,7124	-19,6322	0,0397	-0,2276	0	0,3069	-0,4295	0	0,9376	-0,2223	0,2469	0,2173	-0,5322	0	0,3149	-1,1402
8	766,2266	9,4568	-24,5558	0,0283	-0,2708	0	0,2769	-0,4705	0	0,9726	-0,1013	0,1036	0,3373	-0,5634	0	0,3575	-1,181
9	775,7654	9,2012	-48,291	0	-0,4311	0,1248	0,7941	-0,3756	-0,0577	0,9909	-0,0738	0	0,4093	-0,4032	0	0,3671	-1,1916
10	788,0538	8,9455	-44,2635	0	-0,4158	0,0844	0,7655	-0,4754	-0,1183	0,9937	0	-0,0887	0,598	-0,5293	0	0,4812	-1,0547
11	832,9969	8,6899	-30,0973	0	-0,3369	-0,1397	0	-0,2519	-0,1789	1,0872	-0,0944	-0,2011	1,1317	0	0,4555	0,4949	-1,2838

%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11		
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	13,6654	12,7086	13,01129	13,06885	13,0811	12,87747	13,10963	13,16409	13,44285	13,16263	13,2454
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	12,98338	13,29869	13,09431	13,12786	13,12446	13,40226	13,24901	13,54342	13,91851	13,78467	13,07254
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,024144	2,407873	2,433406	2,287596	2,640221	2,74967	2,765702	2,655853	3,56731	3,363819	2,268705
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	15,33944	13,71026	15,07677	15,13468	14,8199	14,62708	14,76906	14,75889	14,42287	14,72426	13,39071
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	6,204603	6,008266	6,146322	5,418702	6,213594	6,834318	6,740567	6,544108	8,814669	7,131708	12,30566
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	14,54323	13,28104	14,0086	14,31671	14,01141	13,45522	13,28349	12,96526	12,16657	11,5232	8,58615
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	15,03571	14,81391	14,64434	14,72479	14,94911	14,47023	13,89867	13,03647	12,91364	12,47028	13,71843
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,51963	18,4205	18,40861	18,46047	18,52128	18,79583	18,84975	18,90442	18,76251	19,08868	19,83943
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	16,05026	15,85254	15,93042	16,09314	15,98959	15,64113	15,29199	15,17426	15,09231	15,05077	14,89653
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	19,25131	19,40456	19,27828	19,29458	19,409	19,23025	19,24428	19,19322	19,60854	19,59834	20,4526
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	9,634962	11,271	10,0295	10,03405	10,27933	11,35055	10,88835	11,09688	11,74886	11,8799	9,93492
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,57009	13,98635	14,53565	14,56063	14,94188	15,9224	16,20956	16,82797	17,40253	18,77742	19,25773
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	23,30593	22,68522	23,14074	23,12799	23,00727	22,90584	22,58595	22,29827	22,07347	21,4348	21,22303
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	14,61275	14,67945	14,9025	14,78616	14,63709	14,90847	15,2785	15,66702	15,98602	16,10432	15,59627
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	20,71708	20,74694	20,88019	21,02146	20,71503	20,53658	20,00938	19,85411	19,60619	19,06961	16,85472
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	14,764	15,57881	15,03731	14,84288	15,20035	15,44163	15,52109	15,4741	15,90493	15,8677	16,68514
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	14,98679	14,84953	14,78583	14,97342	14,83821	14,979	14,55855	14,4786	14,39234	14,46639	13,40768
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20,69186	20,47917	20,67494	20,76228	20,49184	20,63387	20,48991	20,48726	20,21496	20,26723	19,31362
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	15,47195	16,52357	15,76556	15,9176	15,94581	16,399	15,88097	15,65892	15,6147	15,53991	15,55315
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	24,48167	24,8279	24,47643	24,49739	24,52451	24,98957	24,74069	24,83629	24,77271	25,15296	24,17635
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	20,33414	20,37199	20,43737	20,30949	20,48404	20,87385	20,95337	21,10491	21,23172	21,24703	22,14779
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	22,09988	21,96107	22,10968	21,96211	22,02318	22,21132	22,00859	21,83787	22,22497	22,17823	22,57463
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	13,09336	11,94151	12,49189	12,18853	12,7171	13,17968	12,54695	12,17093	13,80091	12,71648	18,84091
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	22,18392	22,02372	21,68408	21,86197	21,86008	21,32988	20,64373	20,12923	20,15182	19,83868	19,48468
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	24,77331	25,70007	25,44632	25,32801	25,53306	26,66092	27,01397	27,57112	28,27028	28,91939	30,16808

SAE	111,3473	104,8678	108,3696	110,5345	107,2469	100,3569	101,4293	101,9942	94,9967	99,6279	99,6279	94,9967
MAE	13,5066	14,6585	14,1081	14,4115	13,8829	13,4203	14,0530	14,4291	12,7991	13,8835	13,8835	12,7991
SSE	700,1741	656,7536	684,0334	706,399	663,2162	602,3205	640,8105	659,3659	569,5259	617,2003	540,5376	540,5376

1° σελ δεδομένων για Upperbound 11

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 7,45 seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST												
1	753,5843	10,6218	-38,4015	0,0865	-0,2675	0,204	0,9971	-0,3186	0	0,8566	-0,3307	0,3857	-0,2743	-0,493	-0,1913	0,5139	-1,6705												
2	753,7439	10,4277	-36,6843	0,0819	-0,2651	0,1887	0,9312	-0,3124	0	0,8776	-0,3465	0,3981	-0,2105	-0,5176	-0,1767	0,4594	-1,6433												
3	754,1256	10,2337	-30,1148	0,085	-0,2295	0,161	0,8458	-0,381	0	0,8779	-0,3579	0,3897	-0,1751	-0,5005	-0,1799	0,4337	-1,5965												
4	755,345	10,0396	-39,3971	0,0756	-0,3092	0,171	0,8539	-0,4556	0	0,915	-0,3337	0,3987	-0,0347	-0,5208	-0,0484	0,3482	-1,3785												
5	757,0659	9,8456	-46,8197	0,0557	-0,3749	0,1823	0,9851	-0,4903	0	0,9007	-0,2764	0,3545	0,1077	-0,5494	-0,0681	0,3187	-1,1264												
6	760,212	9,6515	-54,2817	0,041	-0,4411	0,1869	0,9927	-0,4828	-0,013	0,9398	-0,2156	0,2936	0,1743	-0,4843	0	0,31	-1,0531												
7	764,685	9,4575	-39,966	0,0296	-0,3688	0,1095	0,7188	-0,4653	-0,0292	0,9706	-0,1263	0,1096	0,2638	-0,5243	0	0,3497	-1,1957												
8	771,6404	9,2634	-44,0787	0,0232	-0,3857	0,105	0,7216	-0,4267	-0,0728	0,9778	-0,082	0,0349	0,4144	-0,4688	0	0,4512	-1,2611												
9	779,946	9,0693	-43,1378	0,0047	-0,4033	0,104	0,8155	-0,4886	-0,1286	1,0161	-0,0795	-0,0168	0,4829	-0,4883	0	0,5027	-1,1151												
10	790,7538	8,8753	-41	0	-0,4043	0,0617	0,6014	-0,5162	-0,1518	1,0632	-0,0433	-0,0744	0,6684	-0,5553	0,1504	0,4551	-1,065												
11	829,8569	8,6812	-29,2835	-0,0061	-0,3525	-0,1103	0,088	-0,2655	-0,2178	1,1253	-0,1494	-0,1127	0,91	0	0,5502	0,4875	-1,2871												
%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11				
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	13,32112	13,18018	13,23514	12,96885	12,73479	12,74467	13,26448	13,20647	13,16387	13,24094	13,37189		
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	13,82776	13,62525	13,74963	13,46274	13,52934	13,37789	13,54723	13,87165	13,56781	13,41187	12,24088		
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	2,62902	2,448742	2,648857	2,552512	3,145383	3,058791	2,841723	3,073466	3,161913	2,803834	1,829786		
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	14,62278	14,52164	14,88696	14,70659	14,80938	14,56917	14,72311	14,54708	14,64237	14,50524	12,886		
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	6,78633	6,472575	6,624459	6,518637	6,53199	7,293972	6,952631	7,656245	7,352988	6,758832	11,98938		
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	13,83259	13,8402	14,1265	13,71532	13,42756	12,83698	12,88402	12,06671	11,56888	11,23073	8,527099		
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	14,59203	14,38075	14,39773	14,43788	14,31354	14,06259	13,2137	13,00305	12,72741	12,49101	14,01938		
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18,12534	18,04875	18,20244	18,60678	18,87323	18,90911	18,82205	18,87648	18,95442	19,30051	19,48717		
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	16,44703	16,19254	16,26956	15,85732	15,69494	15,31427	15,17268	15,38766	15,02417	14,74879	14,21683		
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	19,09318	19,05382	19,03535	19,15503	19,31154	19,35182	19,18961	19,51827	19,67673	19,76483	20,3788		
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	10,939	10,81779	10,81044	11,07353	11,41535	11,43174	11,37605	11,55385	11,75897	12,01842	10,06247		
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14,31464	14,1359	14,73789	15,3859	16,28663	16,63096	16,90057	17,49687	18,55842	18,96701	18,74866		
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	22,93324	22,96044	23,13268	23,00743	22,82222	22,53685	22,3892	21,89525	21,61528	21,40308	20,89435		
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	14,73366	14,68395	14,73502	14,81778	15,20895	15,30727	15,52043	15,80556	15,75142	15,87806	14,82465		
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	21,0238	21,04052	21,13186	20,76833	20,57353	20,09002	19,84218	19,60085	19,20694	18,83047	16,46012		
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	14,88981	14,82849	14,82152	15,16864	15,55807	15,66566	15,30375	15,74069	15,78221	15,98342	16,46499		
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	15,20733	15,05055	15,15419	15,04716	14,88579	14,67206	14,58566	14,60559	14,53749	14,56516	13,42533		
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20,50625	20,5096	20,70661	20,69758	20,77433	20,48684	20,57479	20,28221	20,07987	20,1618	18,66684		
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	15,6364	15,70771	15,82858	16,28354	16,2945	16,17814	15,66453	15,73796	15,63635	16,05293	15,61998		
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	24,50435	24,43815	24,42477	24,75936	25,06602	25,06112	24,8556	24,87343	25,17648	25,35067	24,24109		
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	20,20901	20,20052	20,38977	20,69304	20,92367	20,98154	20,96543	21,12699	21,21571	21,45705	21,60759		
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	21,83688	21,77232	21,76441	22,00325	22,24098	22,38199	21,97876	22,07229	22,35857	22,29338	22,74028		
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	13,15683	12,72057	13,14476	13,02323	12,73163	13,27576	12,78233	13,36335	13,08625	12,48937	18,4333		
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	22,2931	22,08233	21,88593	21,51097	21,07114	20,75789	20,37568	20,33575	20,26536	19,97207	20,09257		
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	24,22085	24,60326	24,92116	26,10428	26,86735	27,4139	27,45374	27,97103	28,91063	29,89866	30,07379		
																	SAE	107,8641	108,1479	107,3190	103,4656	100,1203	97,0094	100,2707	97,0085	98,4515	101,8075	101,8075	97,0085
																	MAE	13,4432	13,8794	13,4552	13,5768	13,8684	13,3242	13,8177	13,2367	13,5137	14,1106	14,1106	13,2367
																	SSE	654,2033	671,3005	652,6881	627,8675	611,2426	587,2144	630,2193	592,7702	600,3333	624,008	547,6177	547,6177

1^o σελ δεδομένων για Upperbound 13

Number of solves in e-constraint: 11
 Number of early exits from the loop: 0
 Elapsed time: 3.06seconds

#	SAD	MAD	BO	AGE	WEIGHT	HEIGHT	BMI	NECK	CHEST	ABDOM	HIP	THIGH	KNEE	ANKLE	BICEPS	FORARM	WRIST
1	753,0614	11,0041	-36,3035	0,0833	-0,2479	0,2207	1,1146	-0,2795	-0,0598	0,8643	-0,3785	0,3973	-0,3178	-0,4913	-0,1929	0,5467	-1,7208
2	753,0943	10,7717	-35,9182	0,0841	-0,2469	0,2201	1,1078	-0,2956	-0,0569	0,8645	-0,3724	0,3904	-0,3346	-0,4839	-0,1946	0,5642	-1,7243
3	753,2999	10,5393	-32,8027	0,0866	-0,2349	0,2011	1,0431	-0,3202	-0,0488	0,8663	-0,3721	0,3952	-0,3003	-0,5076	-0,1976	0,5548	-1,7214
4	753,8557	10,3069	-23,8163	0,0856	-0,1991	0,1583	0,9397	-0,4118	-0,0415	0,8731	-0,3873	0,3905	-0,2285	-0,5526	-0,202	0,4887	-1,5769
5	754,9955	10,0745	-33,0166	0,079	-0,2642	0,1665	0,9062	-0,4638	-0,0335	0,9047	-0,3562	0,3922	-0,0917	-0,5431	-0,1216	0,4299	-1,4187
6	757,0885	9,8421	-48,4445	0,0546	-0,3853	0,1858	0,9929	-0,483	0,0067	0,8982	-0,2698	0,3548	0,1106	-0,5372	-0,067	0,3091	-1,1184
7	760,8407	9,6097	-46,4874	0,0436	-0,3844	0,1673	0,953	-0,529	-0,047	0,9434	-0,2028	0,2512	0,1875	-0,5667	-0,0814	0,4227	-1,1143
8	767,0581	9,3773	-37,4393	0,0303	-0,3448	0,0973	0,7559	-0,4836	-0,0647	0,9576	-0,1095	0,071	0,3192	-0,53	-0,0712	0,4734	-1,2201
9	776,4426	9,1449	-45,0779	0,0109	-0,4079	0,1191	0,8553	-0,4817	-0,1185	1,0025	-0,0804	0,015	0,4162	-0,485	-0,0285	0,5085	-1,1275
10	788,4677	8,9125	-42,0003	-0,0067	-0,4059	0,0611	0,561	-0,4669	-0,1204	1,0573	-0,0372	-0,0771	0,6402	-0,4894	0,1106	0,4406	-1,1235
11	830,5698	8,6801	-28,667	-0,0063	-0,3474	-0,1126	0,104	-0,2727	-0,2108	1,1229	-0,1702	-0,1205	0,9438	0,028	0,5499	0,4969	-1,3363

%BF	Age	Weight	Height	BMI	Neck	Chest	Abdom	Hip	Thigh	Knee	Ankle	Biceps	Forearm	Wrist	1	2	3	4	5	6	7	8	9	10	11		
53	8	51	62,3	172,1	21,1	36,5	89,7	82	89,1	49,3	33,7	21,4	29,6	26	16,9	1	13	13,43512	13,39525	13,34728	13,06184	12,71499	12,77132	13,12265	13,14224	13,30237	13,31056
95	9	47	83,6	189,2	23,4	37,3	99,6	88,8	101,4	57,4	39,6	24,6	30,3	27,9	17,8	2	14	13,67372	13,62985	13,47934	13,46799	13,55463	13,55784	13,70221	13,63013	13,66253	12,27714
153	10,1	27	66,3	183,5	19,7	34,1	88,5	72,8	91,1	53,6	36,8	23,8	27,8	26,3	17,4	3	3	2,617742	2,546435	2,59306	2,605655	3,189092	3,03456	2,971916	3,169453	3,077061	1,851389
248	11	70	60,9	170,2	21,1	34,9	89,2	83,6	88,8	49,6	34,8	21,5	25,6	25,7	18,5	4	14	14,49727	14,61911	15,08496	14,9687	14,77094	14,8856	14,86377	14,70738	14,38156	12,78603
31	11,9	32	82,6	187,3	23,6	38,7	100,5	88,7	99,8	57,5	38,7	33,9	32,5	27,7	18,4	5	7	6,775409	6,42344	5,820843	6,083105	6,705257	6,269499	6,717358	7,350356	7,711022	12,25789
226	12,5	55	57,4	169,5	20	33,4	88,8	78,2	87,5	50,8	33	19,7	25,3	22	15,8	6	14	13,59635	13,75979	14,17497	13,78714	13,40948	12,86419	12,48045	11,73677	11,34786	8,421594
125	13,8	50	73,1	168,9	25,6	37,7	98,9	84,1	94	58,5	36,6	23,5	34,4	29,2	18	7	15	14,66737	14,62639	14,52985	14,40633	14,32433	13,66138	13,0629	12,82355	12,45752	14,10093
227	14,8	55	77,0	173,4	25,6	37,2	101,7	91,1	97,1	56,6	38,5	22,6	33,4	29,3	18,8	8	18	17,92856	17,9922	18,20452	18,48766	18,87031	18,86888	18,83175	18,8868	19,14934	19,53481
233	15,4	58	79,7	181,6	24,2	38	100,2	88,1	97,8	57,1	38,9	23,6	30,9	29,6	18	9	16	16,38166	16,37698	16,21694	16,01078	15,68775	15,53022	15,41836	15,13752	14,8343	14,29331
170	16,5	35	78,4	176,5	25,2	37,6	99,1	90,8	98,1	60,1	39,1	23,4	32,5	29,8	17,4	10	19	19,18569	19,16381	19,0408	19,15174	19,31168	19,38742	19,3265	19,60986	19,84232	20,47234
156	17,3	28	77,9	191,1	21,6	35,6	92,1	83,5	98,3	57,3	37,8	21,7	32,2	27,7	17,7	11	11	11,08851	10,98094	10,85099	11,05703	11,40708	11,59135	11,36605	11,72788	11,96332	9,937968
106	18	43	75,1	174,0	24,8	31,1	93,1	87,3	96,6	54,7	39	24,8	31	29,4	18,8	12	14	14,36467	14,45178	15,09415	15,60672	16,23995	17,14252	17,55054	18,30816	18,60977	18,8734
236	18,6	62	76,6	171,5	26,1	38,3	104,7	95,6	93,7	54,4	37,1	22,7	30,3	26,3	18,3	13	23	22,73047	22,79079	23,06219	22,94593	22,8247	22,34426	22,06378	21,66442	21,53578	20,99421
213	19,5	49	76,4	182,2	23	38,3	98,3	89,7	99,1	56,3	38,8	23	29,5	27,9	18,6	14	14	14,45654	14,46482	14,47142	14,70611	15,24453	15,38921	15,59506	15,74774	16,04157	14,76012
102	20,4	48	78,9	182,9	23,6	37	99,1	92	98,3	59,3	38,4	22,4	27,9	26,2	17	15	20	20,89189	20,95014	20,98303	20,83155	20,57017	20,25911	19,75074	19,36328	19,06021	16,45225
6	20,9	24	95,5	189,9	26,5	39	104,5	94,4	107,8	66	42	25,6	35,7	30,6	18,8	16	15	14,89417	14,79289	14,59028	14,97648	15,59554	15,56469	15,38326	15,72597	16,14978	16,50629
67	21,5	54	68,8	179,7	21,3	35,6	90	83,9	93,9	55	36,1	21,7	29,6	27,4	17,4	17	15	15,29762	15,30546	15,2756	15,20115	14,8489	14,90815	14,64963	14,61848	14,45376	13,31987
78	22,2	69	80,7	174,0	26,7	38,7	102	95	98,3	55	38,3	21,8	30,8	25,7	18,8	18	20	20,30798	20,3873	20,73043	20,69561	20,75987	20,5262	20,44364	20,1075	20,04662	18,65417
195	22,8	42	73,9	184,8	21,6	35,4	92,2	85,6	96,5	60,2	38,9	22,4	31,7	27,1	17,1	19	16	15,60702	15,65685	15,69217	16,09615	16,285	16,1074	15,51007	15,6265	15,95509	15,57975
37	24,2	40	91,8	177,8	29,1	38,5	106,5	100,9	106,2	63,5	39,9	22,6	35,1	30,6	19	20	25	24,56063	24,54596	24,60129	24,79331	25,04216	25,24917	24,96467	25,15252	25,18832	24,22015
94	24,9	46	87,4	182,2	26,3	38	106,6	97,5	100,6	58,9	40,5	24,5	33,3	29,6	19,1	21	20	20,02053	20,03831	20,20465	20,54289	20,93348	20,96812	20,97206	21,11806	21,53743	21,73237
196	25,5	42	81,7	173,4	27,2	38,5	101,6	96,6	100,6	61,1	38,4	24,1	32,9	29,8	18,8	22	21	21,94734	21,88974	21,85799	22,01587	22,24744	22,30399	22,02985	22,35399	22,30906	22,75275
86	26,6	67	75,8	171,5	26	36,5	98,9	89,7	96,2	54,7	37,8	33,7	32,4	27,7	18,2	23	13	13,18055	12,98274	12,71307	12,82887	12,83973	12,42929	12,78577	13,12171	12,9794	18,8078
127	27,2	42	80,6	174,6	26,4	38,9	98,7	92,1	98,5	60,7	36,8	22,2	33,8	30,3	17,2	24	23	22,72387	22,61441	22,21819	21,7879	21,02394	20,89286	20,44764	20,37472	19,93505	20,09382
5	28,7	24	83,6	181,0	25,6	34,4	97,3	100	101,9	63,2	42,2	24	32,2	27,7	17,7	25	24	24,21852	24,33179	24,82437	25,83383	26,8501	27,53417	27,53847	28,55222	29,74589	30,16464

SAE	107,0745	106,9483	107,8833	108,6880	104,7756	99,8112	98,1764	99,3383	97,5458	99,2589	99,2589	97,5458
MAE	13,5613	13,4194	13,6173	13,8869	13,7711	13,7603	14,1707	13,8142	13,4783	13,6206	13,4194	13,4194
SSE	652,7002	646,9368	659,2484	675,6725	641,7162	606,236	617,8177	628,2419	597,3655	598,4866	547,059	547,0590

1° σετ δεδομένων για Upperbound 14