



ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΙΓΑΙΟΥ

Τμήμα Στατιστικής και

Αναλογιστικών - Χρηματοοικονομικών Μαθηματικών

Στατιστική και Ανάλυση Δεδομένων

**Αλγεβρική Στατιστική και Πειραματικοί**

**Σχεδιασμοί με Εφαρμογές**

Μεταπτυχιακή Διατριβή

Μαρία Τρέντου

Σάμος 2020

### **Τριμελής Επιτροπή**

Καραγρηγορίου Αλέξανδρος,

Καθηγητής, Τμήμα Στατιστικής και Αναλογιστικών - Χρηματοοικονομικών Μαθηματικών,  
Πανεπιστήμιο Αιγαίου (Επιβλέπων Καθηγητής)

Τατάκης Χρήστος,

Επισκέπτης Καθηγητής, Τμήμα Μαθηματικών, Πανεπιστήμιο Ιωαννίνων

Ρακιτζής Αθανάσιος,

Επίκουρος Καθηγητής, Τμήμα Στατιστικής και Αναλογιστικών - Χρηματοοικονομικών  
Μαθηματικών, Πανεπιστήμιο Αιγαίου

# Περιεχόμενα

<b>Ευχαριστίες</b>	<b>3</b>
<b>Περίληψη</b>	<b>5</b>
<b>1 Στατιστική και Πειραματικοί Σχεδιασμοί</b>	<b>6</b>
1.1 Εισαγωγικές έννοιες . . . . .	6
1.1.1 Γραμμική Παλινδρόμηση . . . . .	12
1.2 Πειραματικοί σχεδιασμοί . . . . .	22
1.2.1 Τι είναι πειραματικό σχέδιο . . . . .	22
1.2.2 Πως αναλύεται ένα σχέδιο με χρήση στατιστικής; . . . . .	23
<b>2 Βάσεις Gröbner</b>	<b>36</b>
2.1 Εισαγωγικές Έννοιες . . . . .	36
2.2 Μονωνυμικές Διατάξεις . . . . .	40
2.2.1 Λεξικογραφική Διάταξη . . . . .	41
2.2.2 Βαθμωτή Λεξικογραφική Διάταξη . . . . .	42
2.2.3 Αντίστροφη Βαθμωτή Λεξικογραφική Διάταξη . . . . .	43
2.2.4 Διάταξη Βάρους . . . . .	44
2.3 Διάρθρωση Πολυωνύμων . . . . .	46

2.4	Βάσεις Gröbner και αλγόριθμος Buchberger . . . . .	48
2.5	Ελαχιστική και ανάγωγες βάσεις Gröbner . . . . .	53
2.5.1	Καθολική Βάση Gröbner . . . . .	55
2.6	Εφαρμογές των βάσεων Gröbner . . . . .	56
2.6.1	Επίλυση συστημάτων . . . . .	56
2.6.2	Μελέτη του διανυσματικού χώρου πηλίκου $K[x_1, \dots, x_n]/I$ . . . . .	58
<b>3</b>	<b>Βάσεις Gröbner και Πειραματικοί Σχεδιασμοί</b>	<b>60</b>
3.1	Εφαρμογή 1: $2^2$ Πλήρες Παραγοντικό Πείραμα . . . . .	67
3.2	Εφαρμογή 2: $3^3$ Πλήρες παραγοντικό πείραμα . . . . .	70
	<b>Συμπεράσματα</b>	<b>82</b>
	<b>Παράρτημα</b>	<b>84</b>
	<b>Βιβλιογραφία</b>	<b>89</b>

# Ευχαριστίες

Σε αυτό το σημείο θα ήθελα να ευχαριστήσω όλους όσους με βοήθησαν στην εκπόνηση της παρούσας μεταπτυχιακής διατριβής.

Αρχικά θα ήθελα να ευχαριστήσω ιδιαίτερω τον επιβλέποντα καθηγητή μου, Καθηγητή κ. Καραγρηγορίου Αλέξανδρο για την στήριξη του καθ' όλη την διάρκεια του μεταπτυχιακού προγράμματος καθώς και την αμέριστη βοήθεια που μου προσέφερε κατά την εκπόνηση της μεταπτυχιακής μου διατριβής. Χωρίς την βοήθεια, την στήριξη και τις πολύ σημαντικές συμβουλές του θεωρώ πως θα ήταν τρομερά δύσκολο το έργο μου.

Θα ήθελα να ευχαριστήσω επίσης τον Επισκέπτη Καθηγητή και μέλος της τριμελούς επιτροπής, κ. Τατάκη Χρήστο για τις εύστοχες παρατηρήσεις και τα σχόλια του καθώς και για την σημαντική βοήθειά του ως προς την κατανόηση του θέματος, πραγματικά η συμβολή του στην περάτωση της διατριβής ήταν ουσιώδης. Ακόμα, θα ήθελα να ευχαριστήσω και το τρίτο μέλος της επιτροπής, τον Επίκουρο Καθηγητή κ. Ρακιτζή Αθανάσιο αφενός για την συμμετοχή του στην τριμελή επιτροπή, αφετέρου για τις γνώσεις που μας προσέφερε κατά την διάρκεια του μεταπτυχιακού προγράμματος.

Φυσικά, ένα μεγάλο ευχαριστώ οφείλω στους γονείς μου, Βασίλη και Άννα, αφού χωρίς την βοήθεια και την στήριξή τους δεν θα είχα καταφέρει να

πραγματοποιήσω τα όνειρά μου. Παράλληλα θα ήθελα να ευχαριστήσω και τον αδερφό μου Θανάση για την υποστήριξή του, αλλά κυρίως, θα ήθελα να ευχαριστήσω την αδερφή μου Πασχαλινή, μιας και ήταν δίπλα μου καθ' όλη την διάρκεια των σπουδών μου, με βοηθούσε, με στήριζε συνεχώς και με ενθάρρυνε να κάνω πάντα ένα βήμα παραπάνω.

Τέλος, θα ήθελα να πω ένα μεγάλο ευχαριστώ στον Δημήτρη μου για την κατανόησή του όλο αυτό το διάστημα, για την συνεχή στήριξή του, γιατί πάντα πίστευε σε εμένα και τις δυνατότητές μου και μου το έδειχνε με κάθε τρόπο.

# Περίληψη

Η παρούσα διπλωματική εργασία εκπονήθηκε στο Τμήμα Στατιστικής και Αναλογιστικών - Χρηματοοικονομικών Μαθηματικών του Πανεπιστημίου Αιγαίου, στα πλαίσια του Προγράμματος Μεταπτυχιακών Σπουδών "Στατιστική και Ανάλυση Δεδομένων". Σκοπός της είναι η ανάλυση Πειραματικών Σχεδίων με χρήση των βάσεων Gröbner.

Η εργασία αυτή χωρίζεται σε τρία κεφάλαια. Στο πρώτο κεφάλαιο παρουσιάζεται το κομμάτι των Πειραματικών Σχεδιασμών και πιο συγκεκριμένα η στατιστική ανάλυση αυτών. Επίσης γίνεται και μια εισαγωγή στο γραμμικό πρότυπο, βάσει του οποίου γίνεται η ανάλυση των Παραγοντικών Σχεδιασμών.

Στο δεύτερο κεφάλαιο γίνεται μια εισαγωγή στη θεωρία των βάσεων Gröbner. Οι βάσεις αυτές έχουν κάποιες ιδιαίτερες ιδιότητες οι οποίες τις καθιστούν κατάλληλες για την επίλυση στατιστικών προβλημάτων και όχι μόνο.

Τέλος, στο τρίτο κεφάλαιο, συναντάται η εφαρμογή των βάσεων Gröbner στην ανάλυση των Πειραματικών Σχεδιασμών.

# Κεφάλαιο 1

## Στατιστική και Πειραματικοί Σχεδιασμοί

### 1.1 Εισαγωγικές έννοιες

Στατιστική είναι ο κλάδος των Μαθηματικών που εξάγει χρήσιμες πληροφορίες σχετικά τους κινδύνους που συνοδεύουν αποφάσεις όπου εμπεριέχεται αβεβαιότητα και μας δίνει την δυνατότητα να κατανοήσουμε και να επιταχύνουμε την διαδικασία λήψης ορθών αποφάσεων.

**Ορισμός 1.1.1.** Στατιστική είναι η επιστήμη που αναπτύσει μεθόδους:

1. Σχεδιασμού Πειράματος
2. Συλλογής Δεδομένων
3. Επεξεργασίας και Παρουσίασης
4. Ανάλυσης
5. Ερμηνείας - Συμπερασματολογίας



Η Στατιστική χωρίζεται σε δύο μεγάλους κλάδους :

I. Την Περιγραφική Στατιστική και

II. Την Επαγωγική ή Αναλυτική Στατιστική.

Η περιγραφική στατιστική πραγματεύεται μεθόδους συλλογής, περιγραφής και παρουσίασης δεδομένων, όπως για παράδειγμα οι πίνακες συχνοτήτων, γραφήματα και διαγράμματα. Η επαγωγική στατιστική, χρησιμοποιεί τεχνικές όπου με βάση τα δεδομένα ενός δείγματος που έχει επιλεγεί με υποδειγματικό τρόπο (με κάποια εκ των μεθόδων δειγματοληψίας) από ένα πληθυσμό, εξάγει συμπεράσματα για ολόκληρο τον πληθυσμό. Στη διαδικασία αυτή απαραίτητη είναι η δειγματοληψία, δηλαδή η διαδικασία επιλογής ενός αντιπροσωπευτικού δείγματος από το σύνολο του πληθυσμού.

Στην επαγωγική στατιστική διακρίνονται δυο κατηγορίες, η στατιστική εκτίμηση και ο στατιστικός έλεγχος υποθέσεων. Η πρώτη χωρίζεται σε σημειακή εκτίμηση, δηλαδή εκτίμηση της άγνωστης παραμέτρου του πληθυσμού βασισμένη σε μία μόνο ποσότητα που βασίζεται στα δεδομένα και σε εκτίμηση με χρήση διαστημάτων, δηλαδή εκτίμηση της ίδιας άγνωστης παραμέτρου του πληθυσμού βασισμένη σε ένα διάστημα το οποίο προέκυψε από το δείγμα. Το διάστημα εμπιστοσύνης ( $\Delta.E$ ) σχετίζεται με έναν συντελεστή τον οποίο ονομάζουμε συντελεστή εμπιστοσύνης του διαστήματος, παίρνει τιμές στο διάστημα  $[0, 1]$  και δίνει την πιθανότητα η πραγματική τιμή της άγνωστης παραμέτρου να βρίσκεται εντός του  $\Delta.E$ . [4]

Στον στατιστικό έλεγχο υποθέσεων ελέγχουμε, χρησιμοποιώντας τα δεδομένα ενός δείγματος, αν ισχύει ή όχι η υπόθεση που έχουμε κάνει για την άγνωστη παράμετρο του πληθυσμού. Στη διαδικασία αυτή είναι απαραίτητο να κάνουμε δύο βασικές υποθέσεις: α) Την εναλλακτική υπόθεση  $H_1$ , που περιγράφει

τον ισχυρισμό που επιθυμούμε να διερευνηθεί για την άγνωστη παράμετρο του πληθυσμού και

β) Τη μηδενική υπόθεση  $H_0$ , που είναι η άρνηση της εναλλακτικής υπόθεσης. Σημειώνεται ότι η μηδενική υπόθεση αποτελεί μια καθολικά αποδεκτή κατάσταση (π.χ. ο κατηγορούμενος είναι αθώος) και η εναλλακτική αποτελεί την κατάσταση που μπορεί να γίνει αποδεκτή μόνο εφόσον υπάρχουν επαρκή αποδεικτικά στοιχεία (ένοχος ο κατηγορούμενος) Πέραν του ελέγχου υποθέσεων, ή ακριβέστερα, συμπληρωματικά με τον έλεγχο υποθέσεων, υπάρχει και ο έλεγχος σημαντικότητας, όπου προσδιορίζεται το μέγεθος των ενδείξεων που δίνουν τα δεδομένα κατά την μηδενική υπόθεση. Αυτό γίνεται με την βοήθεια του παρατηρούμενου επιπέδου σημαντικότητας, του  $p$ -value.

**Ορισμός 1.1.2.** Παρατηρούμενο επίπεδο σημαντικότητας ( $p$ -value) είναι η πιθανότητα να παρατηρήσουμε τιμή της στατιστικής συνάρτησης ελέγχου ίδια με αυτή που παρατηρήθηκε ή ακόμα πιο ακραία προς την  $H_0$ .

Ένα αποτέλεσμα κρίνεται στατιστικά σημαντικό αν η πιθανότητα να πραγματοποιηθεί όταν ισχύει η  $H_0$  είναι πολύ μικρή. Το ότι πραγματοποιήθηκε δίνει σημαντικές ενδείξεις κατά της  $H_0$ . Άρα, όσο πιο μικρή είναι η τιμή του  $p$ -value, τόσο ισχυρότερες ενδείξεις δίνουν τα πειραματικά δεδομένα κατά της  $H_0$ .

Στη στατιστική, απαραίτητη κρίνεται η χρήση των τυχαίων μεταβλητών (τ.μ.), δηλαδή εννοιών που να προσδίδουν αριθμητικές τιμές σε ένα χαρακτηριστικό ή μια ιδιότητα ενός προσώπου, αντικειμένου ή κατάστασης, που μελετάμε και ενδέχεται να μεταβάλλονται τυχαία ή να παίρνουν διάφορες τιμές σε διαφορετικές καταστάσεις. Οι τυχαίες μεταβλητές έχουν διάφορους διαχωρισμούς, όμως κάποιοι εκ των πιο σημαντικών είναι ο διαχωρισμός τους σε εξαρτημένες

και ανεξάρτητες, αλλά και ο διαχωρισμός σε ποσοτικές και ποιοτικές. Εξαρτημένες είναι οι μεταβλητές οι οποίες θεωρούμε πως μπορούν να επηρεαστούν από άλλους παράγοντες ή μεταβλητές ενώ ανεξάρτητες μεταβλητές είναι οι παράγοντες που θεωρούμε ότι μπορούν να επηρεάσουν την εξαρτημένη μεταβλητή. Επιπλέον, ποσοτικές είναι οι μεταβλητές που παίρνουν τιμές από ένα αριθμητικό σύνολο, ενώ ποιοτικές είναι αυτές που δεν μπορούν να πάρουν αριθμητικές τιμές, για παράδειγμα, στην ερώτηση ποιο είναι το φύλο σου η απάντηση είναι αρσενικό ή θηλυκό. Οι ποσοτικές μεταβλητές χωρίζονται περαιτέρω σε διακριτές, δηλαδή αυτές που παίρνουν τιμές σε πεπερασμένο ή αριθμησιμο σύνολο και συνεχείς, αυτές δηλαδή που παίρνουν οποιαδήποτε τιμή σε οποιοδήποτε διάστημα του  $\mathbb{R}$ . Στην πρώτη περίπτωση μια τ.μ.  $X$  παίρνει τιμές  $X_1, X_2, \dots, X_k$ , ενώ στην δεύτερη μια τ.μ.  $X$  παίρνει τιμές σε οποιοδήποτε υποσύνολο του  $\mathbb{R}$ .

Για τον υπολογισμό των πιθανοτήτων μιας τ.μ.  $X$  είναι απαραίτητη η εύρεση της συνάρτησης κατανομής, η οποία ορίζεται παρακάτω. [3]

**Ορισμός 1.1.3.** Έστω τ.μ.  $X : \Omega \rightarrow \mathbb{R}$  και  $\Omega$  ο δειγματικός χώρος. Η συνάρτηση  $F_X(x) : \mathbb{R} \rightarrow [0, 1], F_X(x) = P(X \leq x)$  καλείται συνάρτηση κατανομής.

**Σχόλιο 1:** Δειγματικός χώρος  $\Omega$  ενός πειράματος καλείται το σύνολο των δυνατών αποτελεσμάτων που μπορούν να εμφανιστούν σε μια εκτέλεσή του. Για παράδειγμα η ρίψη ενός νομίσματος έχει πιθανά αποτελέσματα  $\{K, \Gamma\}$ , όπου  $K$  το αποτέλεσμα της ρίψης να είναι κορώνα και  $\Gamma$  το αποτέλεσμα της ρίψης να είναι γράμματα.

**Σχόλιο 2:** Η συνάρτηση κατανομής έχει τις παρακάτω ιδιότητες,

a) Είναι αύξουσα συνάρτηση.

b)  $\lim_{x \rightarrow \infty} F_X(x) = 1$

c)  $\lim_{x \rightarrow -\infty} F_X(x) = 0$

d) Η  $F_X(x)$  είναι δεξιά συνεχής.

Απόδειξη: (βλέπε Παράρτημα)

Αν η τ.μ. είναι διακριτή, τότε ορίζουμε τη συνάρτηση πιθανότητας, η οποία δίνεται από την σχέση:

$$f_X(x_i) = P(X = x_i), x_i \in \{x_1, x_2, \dots\} = S_x.$$

Η παραπάνω συνάρτηση ικανοποιεί τις ιδιότητες:

a)  $f_X(x_i) \geq 0 \forall i = 1, 2, \dots$

b)  $\sum_{x_i \in S_x} f(x_i) = 1$ . Μάλιστα, ισχύει ότι  $F_X(x) = \sum f_X(x_i)$ .

**Σχόλιο:** Το  $S_x$  μπορεί να είναι πεπερασμένο σύνολο ή αριθμήσιμο σύνολο.

Αν η τ.μ.  $X$  είναι συνεχής, και  $B \subseteq \mathbb{R}$ , τότε ορίζουμε τη συνάρτηση πυκνότητας πιθανότητας  $f_X(x)$ , η οποία δίνεται από την σχέση:

$$P(X \in B) = \int_B f_X(t) dt, t \in \mathbb{R}.$$

**Σχόλιο:** Το  $B$  μπορεί να γραφεί ως ένωση πεπερασμένου ή αριθμήσιμου πλήθους διαστημάτων του  $\mathbb{R}$ .

Η παραπάνω συνάρτηση ικανοποιεί τις ιδιότητες:

a)  $f_X(x) \geq 0, x \in \mathbb{R}$

b)  $\int_{-\infty}^{\infty} f_X(x) = 1$ . Σημειώνεται ότι  $F'_X(x) = f_X(x)$ .

Βασικό εργαλείο της στατιστικής συμπεραματολογίας αποτελούν και τα περιγραφικά μέτρα τα οποία ταξινομούνται σε δυο κατηγορίες, τα μέτρα θέσης και τα μέτρα διασποράς. Τα μέτρα θέσης σχετίζονται με τη θέση των τιμών στο χώρο. Το πιο διαδεδομένο μέτρο θέσης είναι ο αριθμητικός μέσος  $\mu$  ή αλλιώς η μέση τιμή της τ.μ.  $X$ .

**Ορισμός 1.1.4.** Η μέση τιμή μιας διακριτής τυχαίας μεταβλητής  $X = \{x_1, x_2, \dots, x_n, \dots\}$ , με συνάρτηση πιθανότητας  $f_X(x)$ , δίνεται από την σχέση:

$$\mu = E(X) = \sum x_i f_X(x_i).$$

Αντίστοιχα, η μέση τιμή μιας συνεχούς τ.μ.  $X$  με συνάρτηση πυκνότητας πιθανότητας  $f_X(x)$  δίνεται από την σχέση:

$$\mu = E(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Μερικές βασικές ιδιότητες της μέσης τιμής είναι:

1. Αν η τ.μ.  $X$  είναι η σταθερά  $c$  τότε  $E(X) = c$ .
2. Αν  $X$  τ.μ. και  $c$  σταθερά, τότε  $E(cX) = cE(X)$ .
3. Αν  $X$  και  $Y$  είναι δυο τ.μ. τότε  $E(X + Y) = E(X) + E(Y)$ .
4. Αν  $X$  και  $Y$  είναι ανεξάρτητες τ.μ. τότε  $E(XY) = E(X)E(Y)$ .

**Ορισμός 1.1.5.** Ο δειγματικός μέσος μιας τυχαίας μεταβλητής  $X = (x_1, \dots, x_n)$  δίνεται από τη σχέση:

$$\bar{X} = \frac{\sum_{k=1}^n x_k}{n}$$

Από την άλλη, τα μέτρα διασποράς καθορίζουν τον τρόπο με τον οποίο μεταβάλλονται οι τιμές της μεταβλητής, πως διασπείρονται δηλαδή στο χώρο. Το πιο διαδεδομένο μέτρο διασποράς είναι η διακύμανση, δηλαδή η μέση τετραγωνική απόσταση (απόκλιση) των τιμών από τη μέση τιμή.

**Ορισμός 1.1.6.** Η διακύμανση (ή διασπορά) μιας τ.μ.  $X$  με μέση τιμή  $\mu$  δίνεται από την σχέση

$$Var(X) = \sigma^2 = E(X - \mu)^2 = \begin{cases} \sum (x_i - \mu)^2 f_X(x_i), & (a) \\ \int (x - \mu)^2 f_X(x) dx, & (b) \end{cases}$$

όπου  $a$  είναι η διακριτή περίπτωση και  $b$  είναι η συνεχής περίπτωση. Η αντίστοιχη δειγματική διακύμανση δίνεται από την σχέση:

$$s^2 = \frac{\sum_{k=1}^n (x_k - \bar{X})^2}{n}$$

Κάποιες βασικές ιδιότητες της διακύμανσης είναι οι:

1. Αν η τ.μ.  $X$  είναι η σταθερά  $c$  τότε  $Var(c) = 0$ .
2. Αν  $X$  τ.μ. και  $c$  σταθερά, τότε  $Var(X + c) = Var(X)$  και  $Var(cX) = c^2 Var(X)$ .

**Σχόλιο:** Η θετική ρίζα της διακύμανσης ονομάζεται τυπική απόκλιση και υπολογίζεται ως εξής

$$\sigma_X = \sqrt{Var(X)}.$$

Ένα μέτρο βαθμού συσχέτισης μεταξύ δυο μεταβλητών  $X$  και  $Y$  είναι η συνδιακύμανση  $Cov(X, Y)$ .

**Ορισμός 1.1.7.** Η συνδιακύμανση  $\sigma_{XY}$  ή  $Cov(X, Y)$  είναι το μέτρο του βαθμού συσχέτισης δυο μεταβλητών  $X$  και  $Y$  και υπολογίζεται με την σχέση

$$\begin{aligned} Cov(X, Y) &= E[(X - E(X))(Y - E(Y))] = \\ &E(XY - YE(X) - XE(Y) + E(X)E(Y)) = E(XY) - E(X)E(Y). \end{aligned}$$

**Σχόλιο:** Αν  $X$  και  $Y$  ανεξάρτητες τ.μ. τότε  $E(XY) = E(X)E(Y) \Rightarrow Cov(XY) = 0$ .

### 1.1.1 Γραμμική Παλινδρόμηση

Σε αρκετές στατιστικές εφαρμογές συναντάται το πρόβλημα της μελέτης της σχέσης δύο ή περισσότερων τυχαίων μεταβλητών. Ένα παραδείγμα τέτοιας

σχέσης έχουμε στη μελέτη του ύψους και του βάρους μιας ομάδας ανθρώπων. Στόχος μας είναι να αποφασίσουμε αν υπάρχει μια τέτοια σχέση και στη συνέχεια να προσδιορίσουμε τη σχέση αυτή με βάση ορισμένες παρατηρήσεις. Ένας από τους κύριους λόγους που η μελέτη αυτή είναι σημαντική, κυρίως σε εφαρμογές που έχουν σχέση με επιχειρήσεις και με την οικονομία, είναι ότι οι σχέσεις αυτές χρησιμοποιούνται συχνά για προβλέψεις. Το πρώτο βήμα για να πραγματοποιηθεί η μελέτη αυτή είναι ο προσδιορισμός μιας μαθηματικής εξίσωσης (μοντέλου) που περιγράφει τη φύση της σχέσης που υφίσταται μεταξύ των υπό μελέτη μεταβλητών. [7] Στην περίπτωση που δυο μεταβλητές  $X$  και  $Y$  σχετίζονται με γραμμικό τρόπο, τότε με βάση ένα σύνολο  $n$  παρατηρήσεων  $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ , αυτό μπορεί να αποτυπωθεί με το μοντέλο της απλής γραμμικής παλινδρόμησης που είναι της μορφής:

$$Y_i = \beta_0 + \beta_1 * X_i + \varepsilon_i, i = 1, 2, \dots, n,$$

όπου,  $Y_i$ : είναι οι τιμές της εξαρτημένης μεταβλητής, γνωστή και ως μεταβλητή απόκρισης, η οποία είναι τυχαία μεταβλητή

$X_i$ : είναι οι τιμές της ανεξάρτητης μεταβλητής

$\beta_0, \beta_1$ : οι άγνωστες παράμετροι του μοντέλου που καλούνται συντελεστές παλινδρόμησης

$\varepsilon_i$ : τα σφάλματα του μοντέλου τα οποία είναι τυχαίες μεταβλητές και  $n$  το πλήθος των παρατηρήσεων τα οποία δηλώνουν την απόκλιση από την πραγματική τιμή.

Για να μπορούμε να χρησιμοποιήσουμε το γραμμικό μοντέλο κάνουμε κάποιες υποθέσεις για τα σφάλματα οι οποίες αν δεν πληρούνται δεν μπορούμε να εφαρμόσουμε γραμμικό πρότυπο. Αυτές είναι οι εξής:

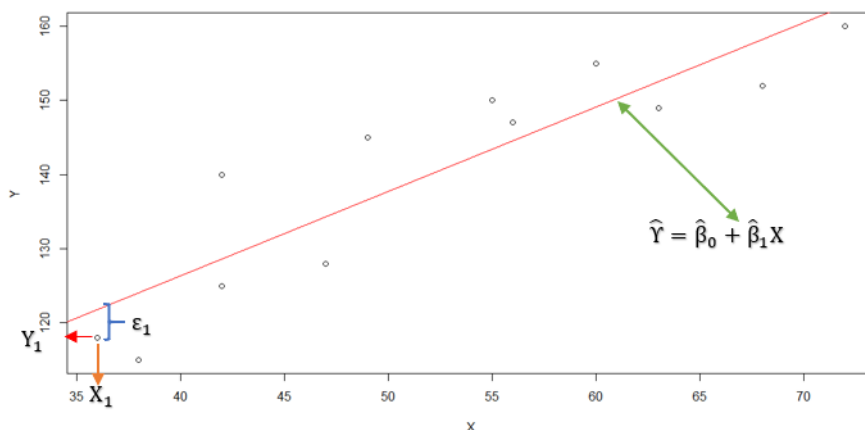
1.  $E(\varepsilon_i) = 0$  και  $Var(\varepsilon_i) = \sigma^2 = c$ , όπου  $c$  σταθερά.

2. Είναι ανά δύο ασυσχέτιστες τ.μ., δηλαδή  $Cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$

3. Τα σφάλματα ακολουθούν κανονική κατανομή.

**Σχόλιο:** Μια ερμηνεία των παραμέτρων  $\beta_0, \beta_1$  του γραμμικού προτύπου θα μπορούσε να δωθεί διαγραμματικά. Δηλαδή, το  $\beta_0$  αποτελεί το σημείο τομής της ευθείας που περιγράφει το γραμμικό μοντέλο με τον άξονα των  $Y$  και το  $\beta_1$  αποτελεί τον συντελεστή διεύθυνσης (κλίση) της ευθείας, αντιπροσωπεύει δηλαδή την αύξηση ή την μείωση στην αναμενόμενη τιμή της  $Y$ , κάτι που φαίνεται και από το Γράφημα 1. [2]

Το διάγραμμα διασποράς ή σκεδασμογράφημα επιτρέπει να αποκομίσει ο ερευνητής μια πρώτη εικόνα των δεδομένων και κατά πόσον οι τιμές  $(x, y)$  φαίνεται να μπορούν να περιγραφούν από μια ευθεία.



Γράφημα 1: Γράφημα παρατηρήσεων  $(X_i, Y_i)$

Από το Γράφημα 1 που αφορά σε γραφική παράσταση 12 σημείων  $(X_i, Y_i)$ ,  $i = 1, \dots, 12$ , παρατηρούμε πως τα σημεία εμφανίζουν μια γραμμική σχέση, καθώς κινούνται γύρω από την κόκκινη ευθεία με γραμμικό τρόπο και συνεπώς



κρίνεται σκόπιμη η χρήση του γραμμικού μοντέλου όπως θα διαπιστωθεί αργότερα, ο βαθμός της γραμμικής σχέσης μπορεί να υπολογισθεί από το συντελεστή συσχέτισης  $R^2$ , ο οποίος και θα αναλυθεί παρακάτω.

Η εκτιμώμενη ευθεία παλινδρόμησης δίνεται από την σχέση:

$$\hat{Y} = E(\hat{Y}_i) = \hat{\beta}_0 + \hat{\beta}_1 * X_i.$$

**Παρατήρηση:** Τα  $\varepsilon_i$  απουσιάζουν από την εκτιμώμενη ευθεία παλινδρόμησης γιατί  $E(\varepsilon_i) = 0$  ενώ τα  $\hat{\beta}_0$  και  $\hat{\beta}_1$  είναι οι εκτιμήτριες των άγνωστων παραμέτρων του μοντέλου.

Οι εκτιμήσεις των παραμέτρων  $\beta_0$  και  $\beta_1$  προκύπτουν από την ελαχιστοποίηση με χρήση της μεθόδου ελαχίστων τετραγώνων της ποσότητας:

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 * X_i)^2.$$

Μετά τη χρήση της μεθόδου προκύπτουν οι εκτιμήσεις των παραμέτρων:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

και

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 * \bar{X}$$

(βλέπε Παράρτημα).

### 1.1.1.1 Ανάλυση Διακύμανσης

Σημαντικό σημείο της ανάλυσης παλινδρόμησης αποτελεί ο πίνακας ανάλυσης διακύμανσης. Ο πίνακας απονα ή αλλιώς πίνακας ανάλυσης διακύμανσης (διασποράς) μας επιτρέπει να αναλύσουμε τη συνολική (ολική) μεταβλητότητα

των δεδομένων που δίνεται από τον τύπο  $SST = \sum_{i=1}^n (Y_i - \bar{Y})^2$  που έχει  $n - 1$  βαθμούς ελευθερίας (DF) αφού από το σύνολο των  $n$  όρων χάνουμε 1 DF λόγω της μιας εκτίμησης, του  $\bar{Y}$ , που υπεισέρχεται στον τύπο. Για την εύρεση του πίνακα ανομα υπολογίζουμε, το άθροισμα τετραγώνων της παλινδρόμησης  $SSR = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$  και το άθροισμα τετραγώνων των υπολοίπων  $SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$ . Ο πίνακας ανομα έχει την μορφή

Source	SS	DF	MS	F-test
Regression	SSE	1	MSB	$F = \frac{MSB}{MSE}$
Residuals	SSR	n-2	MSE	
Total	SST	n-1		

Σημειώνεται ότι  $SST = SSR + SSE$ , όπως προκύπτει με την βοήθεια αναπτύγματος τετραγώνων, ενώ ο αντίστοιχος τύπος ισχύει και για τους βαθμούς ελευθερίας:  $1 + (n - 2) = n - 1$ . Σημειώνεται ότι οι βαθμοί ελευθερίας για το SSR είναι n-2 αφού χάνονται 2 βαθμοί ελευθερίας λόγω των εκτιμήσεων των συντελεστών  $\hat{\beta}_0$  και  $\hat{\beta}_1$  που υπεισέρχονται στο  $\hat{Y}$ .

Ο υπολογισμός των αθροισμάτων των τετραγώνων του πίνακα ανάλυσης διακύμανσης γίνεται με την χρήση των παρακάτω τύπων.

$$SSR = \hat{\beta}_1^2 \sum (X_i - \bar{X})^2 = \hat{\beta}_1^2 \left[ \sum X_i^2 - \frac{(\sum X_i)^2}{n} \right],$$

$$SST = \sum (Y_i - \bar{Y})^2 = \left[ \sum Y_i^2 - \frac{(\sum Y_i)^2}{n} \right] \text{ και}$$

$$SSE = SST - SSR.$$

Σημειώνεται ότι αν το μοντέλο ήταν τέλειο, τότε το SSE θα ήταν 0. Παρατηρούμε ότι  $SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum \hat{e}_i^2$ , όπου  $\hat{e}_i = Y_i - \hat{Y}_i$ . Συνεπώς προκύπτει

ότι ένας αμερόληπτος εκτιμητής για τη διασπορά των σφαλμάτων είναι το μέσο άθροισμα τετραγώνων των υπολοίπων, δηλαδή το MSE:

$$s^2 = \hat{\sigma}^2 = MSE.$$

Το SSE μπορεί να ερμηνευθεί ως η ποσότητα μεταβλητότητας της  $Y$  που δεν ερμηνεύει το γραμμικό μοντέλο.

### 1.1.1.2 Συντελεστής Προσδιορισμού

Η αναλογία της συνολικής διασποράς η οποία και ερμηνεύεται από την παλινδρόμηση, καλείται συντελεστής προσδιορισμού. Το μέτρο αυτό συμβολίζεται με  $R^2$  και εκφράζει το ποσοστό της διασποράς της εξαρτημένης μεταβλητής  $Y$  που ερμηνεύεται από την ανεξάρτητη μεταβλητή  $X$ . Οπότε το ποσοστό της μεταβλητότητας που εξηγείται από την παλινδρόμηση, δίνεται από τον τύπο:

$$R^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST}$$

**Παρατήρηση:** Το  $R^2$  λαμβάνει τιμές μεταξύ του 0 και του 1.

Ένα μέτρο για τον βαθμό συσχέτισης δυο τυχαίων μεταβλητών αποτελεί ο συντελεστής συσχέτισης  $\rho$ .

**Ορισμός 1.1.8.** Το κλάσμα με αριθμητή την συνδιακύμανση των τ.μ.  $X$  και  $Y$ , και παρανομαστή το γινόμενο των τυπικών αποκλίσεων των  $X$  και  $Y$  ονομάζεται συντελεστής συσχέτισης  $\rho$ . Ως εκτιμητής του  $\rho$  χρησιμοποιείται ο συντελεστής συσχέτισης του δείγματος, που ορίζεται ως:

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{\sqrt{\sum(X - \bar{X})^2} \sqrt{\sum(Y - \bar{Y})^2}}$$

**Σχόλιο:** Ο συντελεστής συσχέτισης του δείγματος υπολογίζεται και με την σχέση  $r = \sqrt{R^2}$ , όμως οι ερμηνείες των δυο αυτών ποσοτήτων ( $r$  και  $R^2$ ) διαφέρουν σημαντικά καθώς το  $r$  μας δίνει τον βαθμό γραμμικής συσχέτισης των τ.μ. ενώ το  $R^2$  μας δίνει το ποσοστό της μεταβλητότητας που εξηγείται από το εκάστοτε πρότυπο παλινδρόμησης.

Οι τιμές του συντελεστή συσχέτισης του δείγματος  $r$  κυμαίνονται από -1 έως 1, όπου το -1 σημαίνει πλήρως αρνητική γραμμική συσχέτιση και το 1 σημαίνει πως έχουμε πλήρως θετική γραμμική συσχέτιση. Η τιμή 0 είναι αυτή που μας δείχνει πως δεν υπάρχει γραμμική συσχέτιση μεταξύ δυο μεταβλητών.

### 1.1.1.3 Εκτίμηση με διάστημα

Με την μέθοδο ελαχίστων τετραγώνων υπολογίσαμε τους σημειακούς εκτιμητές των παραμέτρων  $\beta_0$  και  $\beta_1$ . Πέρα από τους σημειακούς εκτιμητές αυτών μπορούμε να υπολογίσουμε ένα διάστημα εμπιστοσύνης για τον κάθε συντελεστή. Ένας μνημονικός κανόνας για τον υπολογισμό διαστήματος εμπιστοσύνης είναι ο παρακάτω συλλογισμός.[2]

Εκτιμητής  $\pm \frac{\alpha}{2}$  ποσοστιαίο σημείο επί τη ρίζα της διασποράς. Αποδεικνύεται μάλιστα ότι  $\hat{\beta}_1 \sim N(\beta_1, \frac{\hat{\sigma}^2}{\sum(X_i - \bar{X})^2})$  και  $\hat{\beta}_0 \sim N(\beta_0, \hat{\sigma}^2(\frac{1}{n} + \frac{\bar{X}^2}{\sum(X_i - \bar{X})^2}))$

[2]. Τότε,  $\frac{\hat{\beta}_i - \beta_i}{std\hat{\beta}_i} \sim N(0, 1)$  και αν η τυπική απόκλιση πρέπει να εκτιμηθεί, τότε η ποσότητα ακολουθεί  $t$  κατανομή με κατάλληλους βαθμούς ελευθερίας, συνεπώς, ο μνημονικός κανόνας δίνει:

$$\hat{\beta}_i \pm t_{\frac{\alpha}{2}, n-1} * std\hat{\beta}_i,$$

όπου  $std$  είναι η θετική ρίζα της διασποράς.

Για την καλύτερη κατανόηση της στατιστικής ανάλυσης ενός γραμμικού μον-

τέλου, θα προχωρήσουμε σε ένα παράδειγμα.

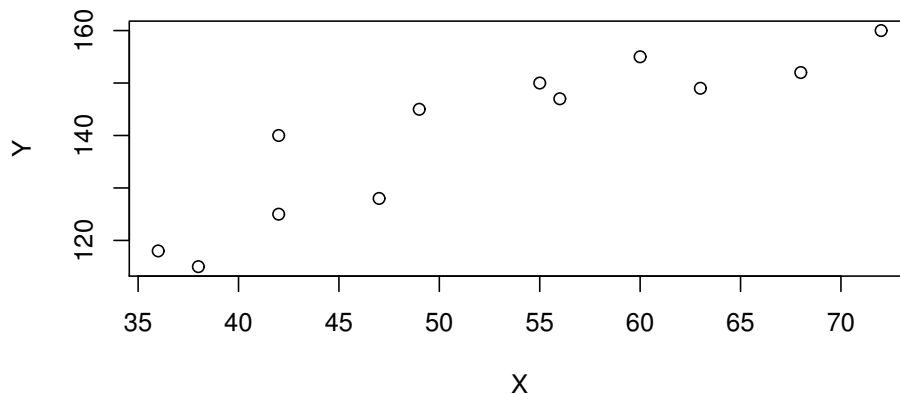
**Παράδειγμα 1.** (Δημητρακοπούλου, 2017) Για τη διερεύνηση της σχέσης μεταξύ της ηλικίας ( $X$ ) και την αρτηριακής πίεσης ( $Y$ ) σε γυναίκες δίνονται οι παρακάτω μετρήσεις:

$X_i$  : 36 38 42 42 47 49 55 56 60 63 68 72

$Y_i$  : 118 115 125 140 128 145 150 147 155 149 152 160

και ζητείται η εφαρμογή κατάλληλου γραμμικού μοντέλου παλινδρόμησης.

Με βάση τους παραπάνω τύπους για τα  $\hat{\beta}_0$  και  $\hat{\beta}_1$  έχουμε ότι η εκτιμώμενη ευθεία παλινδρόμησης είναι η:



Γράφημα 2: Διάγραμμα Σημείων

$$\hat{Y} = 80.778 + 1.138 * X.$$

Αυτό πρακτικά σημαίνει πως για κάθε ετήσια αύξηση στη ηλικία της γυναίκας κατά μέσο όρο η πίεση του αίματός της αυξάνει κατά 1.138 μονάδες.

Στη συνέχεια εργαζόμαστε ως ακολούθως για την εύρεση του πίνακα  $ANOVA$ .

$$SSR = \hat{\beta}_1^2 \sum (X_i - \bar{X})^2 = \hat{\beta}_1^2 [\sum X_i^2 - \frac{(\sum X_i)^2}{n}] = 2008.182$$

$$SST = \sum (Y_i - \bar{Y})^2 = [\sum Y_i^2 - \frac{(\sum Y_i)^2}{n}] = 2500.667$$

$$SSE = SST - SSR = 2500.667 - 2008.182 = 492.485.$$

Άρα ο πίνακας ανάλυσης διακύμανσης είναι ο:

Source	SS	DF	MS	F-test
Regression	2008.182	1	2008.182	$F = 40.777$
Residuals	492.485	10	49.2485	
Total	2500.667	11		

Παρατηρούμε ότι η ολική μεταβλητότητα ισούται με 2500.667 από την οποία το μοντέλο γραμμικής παλινδρόμησης μπορεί να δικαιολογήσει (επεξηγήσει) ένα αρκετά μεγάλο μέρος, ίσο με 2008.182. Παρά ταύτα, ένα τμήμα της μεταβλητότητας ίσο με 492.485 δεν έχει μπορέσει να εξηγηθεί και αποτελεί το σφάλμα. Για να εντοπίσουμε αυτό το ποσοστό της μεταβλητότητας που εξηγείται από το πρότυπο έχουμε με βάση τον συντελεστή προσδιορισμού

$$R^2 = \frac{SSR}{SST} = 0.8031 = 80.31\%.$$

Εν συνεχεία υπολογίζουμε διαστήματα εμπιστοσύνης για την εκτίμηση των παραμέτρων. Για να συμβεί αυτό χρειάζεται να υπολογίσουμε τον αμερόληπτο εκτιμητή της διασποράς των σφαλμάτων που στην προκειμένη περίπτωση είναι ο

$$s^2 = \hat{\sigma}^2 = MSE = 49.2485$$

Συνεχίζουμε υπολογίζοντας ένα 95% διάστημα εμπιστοσύνης για τα  $\beta_0$  και  $\beta_1$  ως εξής: Ένα 95% δ.ε. για το  $\beta_0$ :

$$\hat{\beta}_0 \pm t_{n-2, \frac{\alpha}{2}} \sqrt{\hat{\sigma}^2 \left( \frac{1}{n} + \frac{\bar{X}^2}{\sum (X_i - \bar{X})^2} \right)} = 80.778 \pm 2.228 * \sqrt{49.2485 \left( \frac{1}{12} + \frac{\left( \frac{628}{12} \right)^2}{1550.667} \right)} = 80.778 \pm 21.262 = [59.516, 102.04].$$

Άρα, η πραγματική τιμή του  $\beta_0$  βρίσκεται μεταξύ των τιμών 59.516 και 102.04 με βεβαιότητα 95%.

Με ακριβώς τον ίδιο τρόπο έχουμε και διάστημα εμπιστοσύνης για το  $\beta_1$ , το οποίο είναι

$$\hat{\beta}_1 \pm t_{n-2, \frac{\alpha}{2}} \sqrt{\frac{\hat{\sigma}^2}{\sum (X_i - \bar{X})^2}} = 1.138 \pm 2.228 \sqrt{\frac{49.2485}{1550.667}} = 1.138 \pm 0.397 = [0.741, 1.535]$$

Από το παραπάνω διάστημα εμπιστοσύνης μπορούμε να συμπεράνουμε με βεβαιότητα 95% πως με κάθε ετήσια αύξηση στην ηλικία μιας γυναίκας, η πίεση θα ανεβαίνει από 0.741 μέχρι και 1.535 μονάδες.

Στην ενότητα αυτή αναφέραμε χρήσιμες πληροφορίες που αφορούν στο απλό γραμμικό πρότυπο. Αν ωστόσο θεωρήσουμε και άλλη ανεξάρτητη μεταβλητή η οποία επηρεάζει τη μεταβλητή απόκρισης, τότε μπορούμε να επεκτείνουμε το πρότυπό μας χρησιμοποιώντας την πολλαπλή γραμμική παλινδρόμηση που αποτελεί επέκταση του απλού μοντέλου αλλά δεν θα αναλυθεί περαιτέρω.

Αφού κάναμε μια εισαγωγή στην Ανάλυση Διακύμανσης και τα Μοντέλα Παλινδρόμησης, προχωράμε σε μια εισαγωγή στους Πειραματικούς Σχεδιασμούς τους οποίους και θα αναλύσουμε τόσο με χρήση στατιστικών μεθόδων, όσο και με τη χρήση της Αλγεβρικής Στατιστικής.

## **1.2 Πειραματικοί σχεδιασμοί**

### **1.2.1 Τι είναι πειραματικό σχέδιο**

Η λέξη πείραμα για τους περισσότερους ανθρώπους είναι συνυφασμένη με τις θετικές επιστήμες. Ωστόσο, πειράματα μπορούν να συμβούν και εκτός ελεγχόμενων εργαστηριακών χώρων, για παράδειγμα μπορούν να γίνουν κοινωνικά πειράματα. Στην στατιστική τα δεδομένα συλλέγονται είτε μέσω ερευνών είτε μέσω πειραμάτων. Ο σχεδιασμός και η ανάλυση των πειραμάτων αυτών ονομάζεται πειραματικός σχεδιασμός. Εφαρμογές πειραματικών σχεδιασμών βρίσκουμε σε πληθώρα επιστημονικών πεδίων όπως η γεωργία, η βιολογία και η ιατρική.

Σε μια πειραματική μελέτη έχουμε μεταβλητές ενδιαφέροντος τις οποίες και ονομάζουμε κύριους παράγοντες και μια μεταβλητή την οποία οι παράγοντες επηρεάζουν και την ονομάζουμε μεταβλητή απόκρισης. Για να γίνουν κατανοητές οι έννοιες θεωρούμε μια καλλιεργήσιμη έκταση για την οποία ζητείται να υπολογισθεί η απόδοσή της. Αυτή επηρεάζεται από τις καιρικές συνθήκες που επικρατούν, τη συχνότητα του ποτίσματος, τη ρίψη ή μη λιπάσματος κ.α., έχουμε δηλαδή ως μεταβλητή απόκρισης την απόδοση ενώ οι υπόλοιπες μεταβλητές αποτελούν τους παράγοντες. Η μεταβλητή απόκρισης αναφέρεται και ως εξαρτημένη μεταβλητή ή ως output (εκροή) ενώ οι κύριοι παράγοντες ως ανεξάρτητες μεταβλητές ή input (εισοδή). [14]

Τα τρία πιο ευρέως χρησιμοποιούμενα πειραματικά σχέδια είναι το εντελώς (πλήρως) τυχαίοποιημένο σχέδιο, το σχέδιο πλήρων ομάδων και ο παραγοντικός σχεδιασμός. Στην εργασία αυτή θα ασχοληθούμε κατά βάση με τον τρίτο από τους σχεδιασμούς, δηλαδή με παραγοντικούς σχεδιασμούς.



Η υπολογιστική διαδικασία μέσω της οποίας μελετάμε έναν πειραματικό σχεδιασμό είναι η ανάλυση διακύμανσης, μέσω της οποίας και εξετάζουμε αν οι παράγοντες έχουν στατιστικά σημαντική επίδραση στη μεταβλητή απόκρισης.

### **1.2.2 Πως αναλύεται ένα σχέδιο με χρήση στατιστικής;**

Για την εκτέλεση ενός πειράματος απαιτείται η συλλογή δεδομένων τα οποία και πρέπει να αναλυθούν ώστε να εξαχθούν συμπεράσματα σχετικά με το υπό εξέταση φαινόμενο. Το γεγονός ότι είναι πρακτικά αδύνατον να σχεδιαστεί πείραμα που να καλύπτει το σύνολο ενός πληθυσμού δηλαδή το σύνολο όλων των δυνατών περιπτώσεων για το υπό εξέταση φαινόμενο, υποχρεώνει τον ερευνητή να μελετήσει, με τη χρήση στατιστικών μεθόδων, ένα (τυχαία επιλεγμένο) υποσύνολο των δυνατών περιπτώσεων με αποτέλεσμα να είναι αναπόφευκτη η ύπαρξη πειραματικού σφάλματος. Τα αποτελέσματα που προκύπτουν από την εφαρμογή στατιστικών μεθόδων επιτρέπουν στον ερευνητή να απαντήσει στα ζητούμενα ερωτήματα και να εξαχθούν τα σχετικά συμπεράσματα.

Η εκτέλεση πειραμάτων αφορά στη μελέτη της επίδρασης παραγόντων σε κάποιο αποτέλεσμα, στην απόκριση ενός συστήματος ή μίας διεργασίας. Υπάρχουν κύριοι παράγοντες ελεγχόμενοι και μη ελεγχόμενοι, που επηρεάζουν τη διεργασία, ενώ, φυσικά, κατά τη διεργασία απαιτούνται εισροές που παράγουν τουλάχιστον μία εκροή, που είναι το αποτέλεσμα. [14]

Οι στόχοι ενός πειράματος είναι:

1. Εύρεση παραγόντων που επηρεάζουν περισσότερο την απόκριση  $y$
2. Προσδιορισμός των επιπέδων των πιο σημαντικών  $x$  (ανεξάρτητες μεταβλητές) έτσι ώστε το  $y$  να είναι όσο το δυνατόν πιο κοντά στην επιθυμητή

τιμή

3. Προσδιορισμός των επιπέδων (στάθμες-τιμές) του  $x$  ώστε η διασπορά του  $y$  να είναι μικρή

4. Προσδιορισμός επιπέδων των πιο σημαντικών  $x$  ώστε οι επιδράσεις από μη ελεγχόμενους παράγοντες στην απόκριση  $y$  να είναι μικρές.

Έστω μια μεταβλητή απόκρισης  $y$  και δυο παράγοντες  $X_1$  και  $X_2$  με τον πρώτο να έχει  $I$  επίπεδα και τον δεύτερο να έχει  $J$  επίπεδα. Η συνήθης μορφή του μοντέλου πειραματικού διπαραγοντικού σχεδίου είναι η παρακάτω:

$$Y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}, i = 1, \dots, I, j = 1, \dots, J, k = 1, \dots, n.$$

όπου  $\mu$  είναι ο γενικός μέσος όρος των επεμβάσεων  $\alpha$  και  $\beta$ ,

$a_i$  η επίδραση του  $i$  επιπέδου του πρώτου παράγοντα,

$b_j$  η επίδραση του  $j$  επιπέδου του δεύτερου παράγοντα,

$(ab)_{ij}$  η αλληλεπίδραση των επιπέδων  $i$  και  $j$  των δυο παραγόντων και

$e_{ijk}$  το σφάλμα που αντιπροσωπεύει την απόκλιση από τη μέση τιμή. Το  $n$  είναι το πλήθος των επαναλήψεων του πειράματος.

**Παρατήρηση:** Η βασική προϋπόθεση στους πειραματικούς σχεδιασμούς ταυτίζεται με εκείνη της γραμμικής παλινδρόμησης: Τα  $e_{ijk}$  θεωρούνται ανεξάρτητες τυχαίες μεταβλητές που ακολουθούν την κανονική κατανομή και προέρχονται από πληθυσμό με σταθερή διασπορά. Με άλλα λόγια προϋποτίθεται η ανεξαρτησία, η κανονικότητα και η ομοσκεδαστικότητα (σταθερότητα της διασποράς) των σφαλμάτων.[6]

Για να γίνει κατανοητή η διαδικασία θα χρησιμοποιηθεί ως σημείο αναφοράς ένα παράδειγμα με δυο παράγοντες με δυο επίπεδα έκαστος που είναι γνωστό ως  $2^2$  παραγοντικό πείραμα. Γενικά το  $2^k$  παραγοντικό πείραμα αφορά

στη μελέτη  $k$  παραγόντων με δύο επίπεδα έκαστος με το ένα επίπεδο να χαρακτηρίζεται ως "χαμηλό" και να συμβολίζεται με "-" και το άλλο να χαρακτηρίζεται ως "υψηλό" και να συμβολίζεται με "+". Η ύπαρξη 2 επιπέδων ανά παράγοντα επιτρέπει στον ερευνητή να εντοπίσει τους δυνατούς συνδυασμούς συνθηκών ("θεραπειών" ή "επεμβάσεων" - "treatments") του πειράματος. Έτσι για το απλό παράδειγμα του  $2^2$  παραγοντικού πειράματος για τους παράγοντες, έστω A και B, έχουμε 4 συνολικά θεραπείες, ήτοι

- (1) όταν και οι 2 παράγοντες είναι στο χαμηλό επίπεδο (A="-", B="-")
- (2) όταν και οι 2 παράγοντες είναι στο υψηλό επίπεδο (A="+", B="+")
- (3) όταν ο A είναι στο χαμηλό και ο B στο υψηλό επίπεδο (A="-", B="+") και
- (4) όταν ο B είναι στο χαμηλό και ο A στο υψηλό επίπεδο (A="+", B="-").

Ο συμβολισμός της κάθε θεραπείας γίνεται χρησιμοποιώντας τα σύμβολα εκείνων μόνο των παραγόντων που έχουν τοποθετηθεί στο "υψηλό" επίπεδο. Έτσι η δεύτερη περίπτωση πιο πάνω συμβολίζεται με " $ab$ " αφού και οι 2 παράγοντες είναι τοποθετημένοι στο υψηλό επίπεδο. Το σύμβολο " $a$ " αντιπροσωπεύει την 4η περίπτωση πιο πάνω, αφού στην περίπτωση αυτή μόνο ο παράγοντας A είναι στο "υψηλό" επίπεδο. Η θεραπεία όπου όλοι οι εμπλεκόμενοι παράγοντες είναι στο "χαμηλό" επίπεδο συμβολίζεται με " $(1)$ ". Τονίζεται ότι τα σύμβολα  $(1)$ ,  $a$ ,  $b$  και  $ab$  αντιπροσωπεύουν, επιπρόσθετα, τα άθροισμα των επαναλήψεων των αντιστοίχων θεραπειών. Ένα τυπικό παράδειγμα ενός  $2^2$  παραγοντικού πειράματος με παράγοντες A και B και 4 επαναλήψεις (I, II, III και IV) ανά θεραπεία, δίνονται στον πίνακα 1.

A	B	Treatment	I	II	III	IV	Sum
-	-	(1)	18.2	18.9	12.9	14.4	64.4
+	-	a	27.2	24	22.4	22.5	96.1
-	+	b	15.9	14.5	15.1	14.2	59.7
+	+	ab	41	43.9	36.3	39.9	161.1

Πίνακας 1: Δεδομένα

Σημειώνεται ότι οι συμβολισμοί των θεραπειών γενικεύονται εύκολα για οποιονδήποτε αριθμό παραγόντων. Έτσι για την περίπτωση π.χ. ενός  $2^5$  παραγοντικού για τους παράγοντες A,B,C,D και E η θεραπεία "acd" αφορά στη θεραπεία όπου μόνο οι παράγοντες A, C και D είναι στο "υψηλό" επίπεδο και οι υπόλοιποι δύο, στο "χαμηλό". Για τον υπολογισμό των απλών επιδράσεων του κάθε παράγοντα σε κάθε επίπεδο του άλλου, εφαρμόζεται απλή λογική και υπολογίζεται ο μέσος όρος των επαναλήψεων, ανά περίπτωση. Έτσι για την απλή επίδραση του παράγοντα A στις περιπτώσεις που ο παράγοντας B είναι στο "υψηλό" επίπεδο επικεντρωνόμαστε στις 2 τελευταίες γραμμές του Πίνακα 1, όπου ο B εμφανίζεται στο επίπεδο "+". Στη συνέχεια υπολογίζεται ο μέσος όρος των παρατηρήσεων όταν ο A είναι στο "+" (τελευταία γραμμή του Πίνακα 1) καθώς και αυτός όταν ο A είναι στο "-" (προτελευταία γραμμή του Πίνακα 1). Προφανώς αν οι 2 μέσοι όροι είναι ίσοι μπορούμε να θεωρήσουμε ότι δεν επιδρά ο παράγοντας A (όταν ο B είναι στο "+"). Όσο μεγαλύτερη είναι η διαφορά των δύο μέσων τόσο ισχυρότερη είναι η επίδραση του A (όταν ο B είναι στο "+"). Με αντίστοιχο τρόπο μπορεί να προσδιορισθεί η απλή επίδραση του A όταν ο B είναι στο "-". Οι σχετικοί τύποι, όπου  $n$  ο αριθμός των επαναλήψεων ανά θεραπεία, είναι

$$\frac{(ab-b)}{n} \text{ και } \frac{(a-1)}{n}.$$

Αντιστοίχως οι απλές επιδράσεις του παράγοντα Β σε κάθε επίπεδο του Α δίνονται από τις αντίστοιχες σχέσεις:

$$\frac{(ab-a)}{n} \text{ και } \frac{(b-1)}{n}.$$

Ο προσδιορισμός της κύριας επίδρασης του παράγοντα Α βασίζεται στον μέσο όρο των μέσων όρων των απλών επιδράσεων του Α και δίδεται από τον παρακάτω τύπο:

$$A = \frac{1}{2^{k-1}n} \cdot [(ab - b) + (a - (1))].$$

Ομοίως, συνδυάζοντας τις δύο απλές επιδράσεις του Β, η κύρια επίδραση του παράγοντα Β δίδεται από τον τύπο:

$$B = \frac{1}{2^{k-1}n} \cdot [(ab - a) + (b - (1))].$$

Για την αλληλεπίδραση των παραγόντων Α και Β λαμβάνονται υπόψη και οι 4 γραμμές του Πίνακα 1 και υπολογίζεται ο μέσος όρος των παρατηρήσεων όταν οι δύο παράγοντες είναι στο ίδιο επίπεδο (πρώτη και τελευταία γραμμή του Πίνακα 1) καθώς και ο μέσος όρος όταν οι παράγοντες είναι σε διαφορετικό επίπεδο (δεύτερη και τρίτη γραμμή του Πίνακα 1). Αν δεν υπάρχει αλληλεπίδραση οι δύο αυτοί μέσοι όροι αναμένονται να είναι σχεδόν ίσοι. Όσο μεγαλύτερη είναι η διαφορά των συγκεκριμένων μέσων όρων, τόσο ισχυρότερη είναι η αλληλεπίδραση των δύο παραγόντων. Ο σχετικός τύπος είναι:

$$AB = \frac{1}{2^{k-1}n} \cdot [(ab - b) - (a - (1))].$$

Η διερεύνηση της (στατιστικής) σημαντικότητας των επιδράσεων (συμπεριλαμβανομένης της αλληλεπίδρασης) που προκύπτουν από τους πιο πάνω

τύπους μπορεί να γίνει είτε μέσω της ανάλυσης της ολικής μεταβλητότητας (analysis of variance, ANOVA) είτε μέσω της πολλαπλής γραμμικής παλινδρόμησης. Το παραπάνω μοντέλο μπορεί πράγματι να θεωρηθεί ως πολλαπλό γραμμικό μοντέλο παλινδρόμησης με τρεις μεταβλητές  $X_1$  και  $X_2$ , που αφορούν τους δυο παράγοντες A και B καθώς και μια ακόμα μεταβλητή που συμβολίζεται με  $X_1X_2$  και αφορά στην αλληλεπίδραση των A και B. Όταν ένας παράγοντας βρίσκεται στο "υψηλό" επίπεδο η αντίστοιχη μεταβλητή τίθεται ίση με "+1" ενώ αν βρίσκεται στο "χαμηλό" επίπεδο η αντίστοιχη μεταβλητή τίθεται ίση με "-1". Η μεταβλητή  $X_1X_2$  που αντιστοιχεί στην αλληλεπίδραση προκύπτει από τους προφανείς πολλαπλασιασμούς των τιμών των αντιστοιχών μεταβλητών  $X_1$  και  $X_2$ . Η μεταβλητή απόκρισης  $Y$  (που λαμβάνει την τιμή  $y$ ) σε κάθε συνδυασμό των  $X_1$  και  $X_2$  ισούται με το άθροισμα των παρατηρήσεων των επαναλήψεων του συγκεκριμένου συνδυασμού. Το μοντέλο πολλαπλής παλινδρόμησης που προκύπτει είναι το εξής:

$$y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2 + \varepsilon.$$

Χρησιμοποιώντας και πάλι ως σημείο αναφοράς το παράδειγμα του Πίνακα 1, παρατηρούμε ότι για εφαρμογή της τεχνικής της πολλαπλής παλινδρόμησης τα δεδομένα του Πίνακα 1 παίρνουν τη μορφή του Πίνακα 2, όπου η στήλη *Sum* αντιπροσωπεύει το άθροισμα των 4 επαναλήψεων ανά θεραπεία:

A	B	Tr	Sum	$X_1$	$X_2$	$X_1X_2$
-	-	(1)	64.4	-1	-1	+1
+	-	a	96.1	+1	-1	-1
-	+	b	59.7	-1	+1	-1
+	+	ab	161.1	+1	+1	+1

## Πίνακας 2: Δεδομένα για παλινδρόμηση

Σημειώνεται ότι οι συντελεστές παλινδρόμησης  $\beta_1, \beta_2, \beta_3$  του πιο πάνω μοντέλου παλινδρόμησης συνδέονται με τις επιδράσεις A, B και την αλληλεπίδραση AB μέσω των παρακάτω τύπων, όπως θα διαφανεί και από το παράδειγμα που αναπτύσσεται κατωτέρω:

$$\hat{\beta}_1 = b_1 = \frac{A}{2}, \hat{\beta}_2 = b_2 = \frac{B}{2} \text{ και } \hat{\beta}_3 = b_3 = \frac{AB}{2}.$$

Σχετικά με την ανάλυση της ολικής μεταβλητότητας ή ανάλυση διασποράς υπενθυμίζεται ότι η συνολική μεταβλητότητα που εμπεριέχεται σε ένα σύνολο δεδομένων  $Y_1, \dots, Y_n$  δίδεται από τον τύπο  $SST = \sum(Y_i - \bar{Y})^2$ . Η εφαρμογή συγκεκριμένου μοντέλου παλινδρόμησης όπως αναφέρθηκε και στην προηγούμενη ενότητα οδηγεί στην ανάλυση της ολικής διασποράς  $SST$ , στον όρο  $SSR$  που αφορά στο τμήμα της ολικής διασποράς που μπορεί να εξηγηθεί ή δικαιολογηθεί από τη χρήση του συγκεκριμένου μοντέλου και τον όρο  $SSE$  που αναφέρεται ως άθροισμα τετραγώνων των σφαλμάτων και αφορά στο τμήμα της ολικής διασποράς που δεν μπορεί να εξηγηθεί από το προτεινόμενο μοντέλο. Ως εκ τούτου, όπως είχε αναφερθεί στην προηγούμενη ενότητα, ισχύει γενικώς η σχέση

$$SST = SSR + SSE.$$

Στην περίπτωση του  $2^2$  παραγοντικού πειράματος η ποσότητα  $SSR$  αποτελείται από τρεις όρους τους  $SS_A, SS_B$  και  $SS_{AB}$  που αντιστοιχούν στους 3 όρους που εμπλέκονται στο προτεινόμενο μοντέλο πολλαπλής παλινδρόμησης και αφορούν στο τμήμα της ολικής μεταβλητότητας που εξηγείται από τον παράγοντα A, τον παράγοντα B και την αλληλεπίδραση AB, αντιστοίχως. Κατά

συνέπεια, ισχύει η σχέση

$$SSR = SS_A + SS_B + SS_{AB}.$$

Όσο μεγαλύτερο είναι ένα συγκεκριμένο εκ των τριών πιο πάνω αθροισμάτων τετραγώνων, τόσο μεγαλύτερη είναι η επίδραση του αντιστοίχου παράγοντα (συμπεριλαμβανομένης της αλληλεπίδρασης) στο υπό εξέταση φαινόμενο, δηλ. στη μεταβλητή απόκρισης. Ακριβώς όπως συνέβη και στην περίπτωση των συντελεστών παλινδρόμησης, έτσι και εδώ υπάρχουν σχέσεις, που θα επιβεβαιωθούν στο παράδειγμα κατωτέρω, που συνδέουν τα άθροισμα τετραγώνων της ανάλυσης διασποράς με τις θεραπείες (1),  $a$ ,  $b$  και  $ab$  οι οποίες αντιπροσωπεύουν τα αθροίσματα των επαναλήψεων των αντίστοιχων θεραπειών. Έτσι για τα  $SS_A$ ,  $SS_B$  και  $SS_{AB}$  που αφορούν στα αθροίσματα τετραγώνων του παράγοντα A, του παράγοντα B και της αλληλεπίδρασης AB αντίστοιχα, ισχύουν οι ακόλουθοι τύποι:

$$SS_A = \frac{(-1) + a - b + ab)^2}{4 * n},$$

$$SS_B = \frac{(-1) - a + b + ab)^2}{4 * n},$$

και

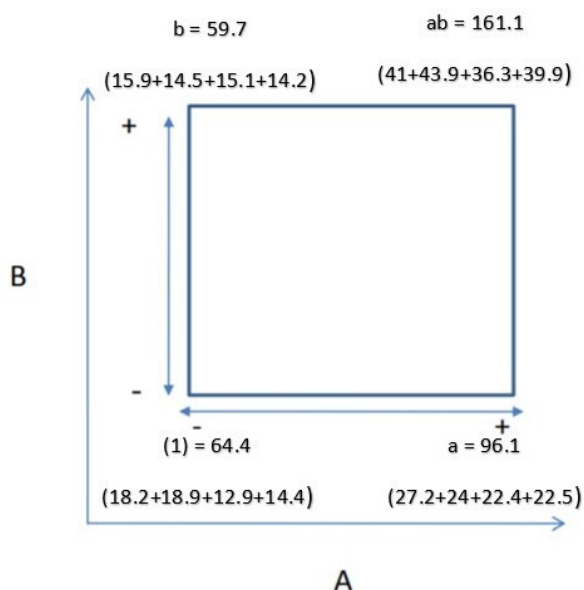
$$SS_{AB} = \frac{((1) - a - b + ab)^2}{4 * n}.$$

**Παράδειγμα 2.** (Montgomery, 2013) Ένας δρομολογητής χρησιμοποιείται για την κοπή πλακέτας με εγκοπή. Το επίπεδο των κραδασμών στην επιφάνεια της πλακέτας καθώς κόβεται θεωρείται ως μια σημαντική πηγή διακύμανσης των εγκοπών. Δυο παράγοντες A και B θεωρείται ότι επηρεάζουν τη δόνηση, το μέγεθος (A) και η ταχύτητα κοπής (B). Επιλέγονται δύο επίπεδα ανά παράγοντα, ήτοι δυο μεγέθη (υψηλό "+" και χαμηλό "-") και δυο ταχύτητες κοπής



(υψηλή "+" και χαμηλή "-") επιλέγονται τέσσερις πλάκες (I, II, III, IV) για κάθε συνδυασμό θεραπειών ( (1), a, b, και ab). Η μεταβλητή απόκριση μετρά τη δόνηση που προκύπτει ως αποτέλεσμα τριών επιταχυνσιόμετρων σε κάθε πλάκα. Έστω ότι τα δεδομένα είναι αυτά που δίνονται στον Πίνακα 1.

Με βάση την παραπάνω περιγραφή έχουμε ένα  $2^2$  παραγοντικό πειραματικό σχεδιασμό με τους συνδυασμούς των επεμβάσεων (θεραπείες, "Tr") να δίνονται σχηματικά στο Σχήμα 1



Σχήμα 1: Συνδυασμοί επεμβάσεων

**Σχόλιο:** Το Σχήμα 1 αποτελεί μια απεικόνιση της σχέσης των επεμβάσεων με τα επίπεδα των παραγόντων A και B. Παρατηρούμε δηλαδή πως αν βρισκόμαστε στο χαμηλό επίπεδο ("-") του παράγοντα A και στο χαμηλό επίπεδο του παράγοντα B, τότε η αντίστοιχη επέμβαση λαμβάνει την τιμή 64.4, ενώ αν βρισκόμαστε στο χαμηλό επίπεδο του παράγοντα B και στο υψηλό ("+")

επίπεδο του A, τότε η επίδραση λαμβάνει την τιμή 96.1. Αυτό σημαίνει πως κατά την μετάβαση από το χαμηλό επίπεδο του A στο υψηλό θα μεταβληθεί η τιμή της μεταβλητής απόκρισης κατά 26.7 μονάδες, παραμένοντας αμετάβλητος ο παράγοντας B. Αντίστοιχα, παρατηρούμε πως αν ο παράγοντας A παραμείνει στο χαμηλό επίπεδο, κατά την μετάβαση του B από το χαμηλό στο υψηλό θα έχουμε μείωση της τιμής της μεταβλητής απόκρισης κατά 4.7 μονάδες.

Οι απλές επιδράσεις του παράγοντα A σε κάθε επίπεδο του B με βάση τους πιο πάνω τύπους ισούνται με:

$$\frac{(ab-b)}{n} = \frac{161.1-59.7}{4} = 25.35$$

$$\frac{(a-(1))}{n} = \frac{96.1-64.4}{4} = 7.925.$$

Ομοίως, οι απλές επιδράσεις του παράγοντα B ισούνται με:

$$\frac{(ab-a)}{n} = \frac{161.1-96.1}{4} = 16.25$$

$$\frac{(b-(1))}{n} = \frac{59.7-64.4}{4} = -1.175.$$

Με χρήση των σχετικών τύπων, πιο πάνω, οι κύριες επιδράσεις των A και B ισούνται με:

$$A = \frac{1}{8}[(161.1 - 59.7) + (96.1 - 64.4)] = \frac{133.1}{8} = 16.64$$

και

$$B = \frac{1}{8}[(161.1 - 96.1) + (59.7 - 64.4)] = \frac{60.3}{8} = 7.54.$$

Τέλος η αλληλεπίδραση των A και B ισούται με:

$$AB = \frac{1}{8}[(161.1 - 59.7) - (96.1 - 64.4)] = \frac{69.7}{8} = 8.7.$$

Με βάση τα ανωτέρω φαίνεται ότι οι όλες οι τιμές απέχουν από το μηδέν (χωρίς ακόμα να είναι φανερό αν αυτή η διαφορά είναι στατιστικώς σημαντική). Σε οποιαδήποτε περίπτωση η επίδραση του παράγοντα A είναι σημαντικά μεγαλύτερη από αυτές των B και AB.

Η εφαρμογή του αντίστοιχου μοντέλου πολλαπλής παλινδρόμησης με χρήση της γλώσσας προγραμματισμού R δίνει τα εξής αποτελέσματα:

Coefficients:				
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	23.8312	0.6112	38.991	5.22e-14 ***
A	8.3187	0.6112	13.611	1.17e-08 ***
B	3.7687	0.6112	6.166	4.83e-05 ***
A:B	4.3562	0.6112	7.127	1.20e-05 ***

Πίνακας 3: Εκτιμητές παραμέτρων

Με βάση τον Πίνακα 3, το εκτιμώμενο μοντέλο είναι το:

$$y = 23.83 + 8.32x_1 + 3.77x_2 + 4.35x_1x_2$$

απ' όπου μπορεί εύκολα να διαπιστωθεί η σχέση των συντελεστών παλινδρόμησης με τις επιδράσεις των παραγόντων, που υπολογίσθηκαν νωρίτερα, αφού,

$$\hat{\beta}_1 = b_1 \equiv \frac{A}{2} = \frac{16.64}{2} = 8.32$$

$$\hat{\beta}_2 = b_2 \equiv \frac{B}{2} = \frac{7.54}{2} = 3.77$$

$$\hat{\beta}_3 = b_3 \equiv \frac{AB}{2} = \frac{8.7}{2} = 4.35.$$

Οι παραπάνω τύποι καταδεικνύουν ότι οι επιδράσεις του A, B και AB ταυτίζονται με το διπλάσιο των αντίστοιχων συντελεστών του πολλαπλού μοντέλου παλινδρόμησης.

Οι πολύ μικρές τιμές των p-values (τελευταία στήλη του Πίνακα 3) που αφορούν στη διερεύνηση της υπόθεσης ότι η αντίστοιχη επίδραση είναι ίση με το μηδέν, δηλώνουν ότι τόσο οι 2 παράγοντες όσο και η αλληλεπίδραση έχουν στατιστικώς σημαντική συνεισφορά στη μεταβλητή απόκρισης, ήτοι τη δόνηση των πλακών. Αντίστοιχα είναι και τα πορίσματα της ανάλυσης διασποράς η οποία με χρήση της R, αποτυπώνεται στον Πίνακα 4:

Analysis of Variance Table					
Response: Y					
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	1107.23	1107.23	185.252	1.175e-08 ***
B	1	227.26	227.26	38.023	4.826e-05 ***
A:B	1	303.63	303.63	50.801	1.201e-05 ***
Residuals	12	71.72	5.98		
Total	15	1709.84	-	-	-

Πίνακας 4: Πίνακας ANOVA

Με βάση τα αποτελέσματα του πιο πάνω Πίνακα 4, η ολική μεταβλητότητα ισούται με  $\sum(Y_i - \bar{Y})^2 = 1709.84$  ενώ για τα αθροίσματα τετραγώνων έχουμε

$$SS_A = 1107.23, SS_B = 227.26, SS_{AB} = 303.63.$$

Έτσι, από τη συνολική μεταβλητότητα, 1709.84, οι παράγοντες A, B καθώς και η αλληλεπίδρασή τους καταφέρνουν να εξηγήσουν ή ερμηνεύσουν ένα μεγάλο μέρος ίσο με  $1107.23 + 227.26 + 303.63 = 1638.12$ . Η υπόλοιπη μεταβλητότητα ίση με  $1709.84 - 1638.12 = 71.72$  δεν έχει μπορέσει να εξηγηθεί από το προτεινόμενο μοντέλο και αντιπροσωπεύει το σφάλμα (residuals) που συμβολίζεται με SSE (sum of squared errors). Το γεγονός ότι το μεγαλύτερο ποσοστό της μεταβλητότητας εξηγείται από το μοντέλο, περίπου 96%, δείχνει ότι και οι δυο παράγοντες, αλλά και η αλληλεπίδραση παίζουν πολύ σημαντικό ρόλο στο υπό εξέταση φαινόμενο. Σημειώνεται ότι το ποσοστό

της μεταβλητότητας που επεξηγείται από το προτεινόμενο μοντέλο πολλαπλής παλινδρόμησης αν και συμβολίζεται, όπως και στην προηγούμενη ενότητα, με  $R^2$  αναφέρεται ως συντελεστής προσδιορισμού (coefficient of determination) για να ξεχωρίζει από το συντελεστή συσχέτισεως που αναφέρεται αποκλειστικά, στο μοντέλο (απλής) γραμμικής παλινδρόμησης.

Τα αποτελέσματα του Πίνακα 4 και κυρίως οι F-values (εξαιρετικά μεγάλες τιμές) και τα p-values (εξαιρετικά μικρές τιμές και οι οποίες ταυτίζονται, προφανώς με τις αντίστοιχες τιμές του Πίνακα 3) επιβεβαιώνουν ότι τόσο οι κύριες επιδράσεις A και B όσο και η αλληλεπίδρασή τους είναι στατιστικά σημαντικές και συνεπώς επηρεάζουν τη δόνηση της πλακέτας. Οπότε, από το μοντέλο δεν είναι δυνατόν να παραληφθεί κάποιος παράγοντας και συνεπώς, αυτό θα έχει τη μορφή:

$$y = \mu + a_i + b_j + (ab)_{ij} + \varepsilon.$$

Με αντίστοιχη διεργασία μπορεί να αναλυθεί ένα πειραματικό σχέδιο  $3^3$  και γενικότερα οποιοδήποτε παραγοντικό σχέδιο, ή και κλάσμα αυτού. Βέβαια, για να αναλυθεί ένα κλασματικό παραγοντικό σχέδιο πρέπει να γίνει και διαχωρισμός των επιδράσεων με σκοπό την επιλογή του τμήματος που θα αναλύσουμε.

**Παρατήρηση:** Το αντίστοιχο μοντέλο για ένα παραγοντικό σχεδιασμό  $3^3$  είναι το:

$$Y_{ijkl} = \mu + a_i + b_j + (ab)_{ij} + c_k + (ac)_{ik} + (bc)_{jk} + (abc)_{ijk} + e_{ijkl},$$

όπου εδώ οι κύριες επιδράσεις έχουν 2 βαθμούς ελευθερίας, οι διπλές αλληλεπιδράσεις έχουν 4 βαθμούς ελευθερίας και οι τριπλές αλληλεπιδράσεις  $2^k = 2^3$  βαθμούς ελευθερίας. [6]

# Κεφάλαιο 2

## Βάσεις Gröbner

### 2.1 Εισαγωγικές Έννοιες

Στο κεφάλαιο αυτό γίνεται μια εισαγωγή στην Υπολογιστική Άλγεβρα και ειδικότερα στη θεωρία των βάσεων Gröbner. Η εφαρμογή αυτών είναι ευρεία στον κλάδο της Αλγεβρικής Στατιστικής, όπως θα δούμε λεπτομερώς στο Κεφάλαιο 3.

Οι βάσεις Gröbner εισήχθησαν το 1965 από τον Αυστριακό καθηγητή Bruno Buchberger μέσω της διδακτορικής του διατριβής με τίτλο "An Algorithm for Finding the Basis Elements of the Residue Class Ring Modulo a Zero-dimensional Polynomial Ideal", όπου επιβλέπωντας καθηγητής του ήταν ο επίσης Αυστριακός Dr. Wolfgang Gröbner. [10]

**Ορισμός 2.1.1.** Ένα σύνολο  $R$ , μη κενό, εφοδιασμένο με δυο διμελείς πράξεις,  $+$  :  $R \times R \rightarrow R$  και  $\cdot$  :  $R \times R \rightarrow R$  με

a) το σύνολο  $(R, +)$  είναι αβελιανή ομάδα

b)  $a \cdot (b \cdot c) = (a \cdot b) \cdot c, \forall a, b, c \in R$

c)  $a \cdot (b + c) = a \cdot b + a \cdot c$  και  $(a + b) \cdot c = a \cdot c + b \cdot c, \forall a, b, c \in R$ ,  
ονομάζεται δακτύλιος.

Για παράδειγμα, το σύνολο των πραγματικών αριθμών,  $\mathbb{R}$  με πράξεις την πρόσθεση και τον πολλαπλασιασμό,  $(\mathbb{R}, +, \cdot)$  είναι δακτύλιος, αφού

- Το  $(\mathbb{R}, +)$  είναι αβελιανή ομάδα,  $\forall a, b \in \mathbb{R}$  ισχύει ότι  $a + b = b + a$ ,
- $\forall a, b, c \in \mathbb{R}$  ισχύει ότι  $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ ,
- $\forall a, b, c \in \mathbb{R}$  ισχύει ότι  $a \cdot (b + c) = a \cdot b + a \cdot c$  και  $\forall a, b, c \in \mathbb{R}$  ισχύει ότι  $(a + b) \cdot c = a \cdot c + b \cdot c$ .

**Ορισμός 2.1.2.** Ένας δακτύλιος θα καλείται μεταθετικός δακτύλιος όταν  $a \cdot b = b \cdot a, \forall a, b \in R$ .

Για παράδειγμα, ο  $(\mathbb{R}, +, \cdot)$  είναι μεταθετικός δακτύλιος, αφού είναι δακτύλιος όπως έχουμε δει και για κάθε  $a, b \in \mathbb{R}$  έχουμε  $a \cdot b = b \cdot a$ .

**Ορισμός 2.1.3.** Ορίζουμε έναν δακτύλιο  $R$  ως πολυωνυμικό δακτύλιο του  $x$  και τον συμβολίζουμε με  $R[x]$ , το σύνολο όλων των πολυωνύμων του  $x$  με συντελεστές στοιχεία του δακτυλίου  $R$ . Μπορούμε ακόμα να ορίσουμε τον πολυωνυμικό δακτύλιο ως προς  $y$  με συντελεστές στον  $R[x]$  και τότε συμβολίζουμε με  $R[x, y]$ . (Οι πράξεις του πολυωνυμικού δακτυλίου ορίζονται στις σχέσεις 2.1 και 2.2)

Μπορούμε φυσικά να επεκτείνουμε τον ορισμό του πολυωνυμικού δακτυλίου και για  $n$  το πλήθος μεταβλητές,  $x_1, \dots, x_n$ .

Θεωρούμε  $K[x_1, \dots, x_n]$  πολυωνυμικό δακτύλιο, όπου  $K$ -διανυσματικός χώρος και  $T^n = \{x_1^{a_1} \dots x_n^{a_n}, a_i \in \mathbb{N}, i = 1, \dots, n\}$ . Τότε τα στοιχεία της μορφής  $\lambda \cdot x_1^{a_1} \dots x_n^{a_n}, \lambda \in K$  καλούνται μονώνυμα.

**Ορισμός 2.1.4.** Έστω μονώνυμο  $x^u = x_1^{u_1} \cdot \dots \cdot x_n^{u_n}$ . Καλούμε βαθμό του μονωνύμου τον  $\deg(x^u) = u_1 + u_2 + \dots + u_n$ ,  $u = (u_1, \dots, u_n) \in \mathbb{N}^n$ .

Έστω πολυωνυμικός δακτύλιος  $K[x_1, \dots, x_n]$  και έστω μονώνυμο  $x^u$ ,  $u = (u_1, \dots, u_n) \in \mathbb{N}^n$ . Έστω πολυώνυμα  $f = \sum_{u \in \mathbb{N}^n} c_u \cdot x^u$  και  $g = \sum_{u \in \mathbb{N}^n} d_u \cdot x^u$ . Τότε ορίζουμε την πρόσθεση και τον πολλαπλασιασμό πολυωνύμων ως εξής:

$$f + g = \sum_{u \in \mathbb{N}^n} (c_u + d_u) \cdot x^u \quad (2.1)$$

και

$$f \cdot g = \sum_{k \in \mathbb{N}^n} \left( \sum_{a+b=k} c_a \cdot d_b \right) \cdot x^k. \quad (2.2)$$

**Ορισμός 2.1.5.** Έστω  $I$  ένα υποσύνολο ενός δακτυλίου  $R$ . Το  $I$  ονομάζεται ιδεώδες του δακτυλίου αν:

- a)  $I \neq \emptyset$
- b)  $a + b \in I$ ,  $\forall a, b \in I$
- c)  $k \cdot a \in I$  και  $a \cdot k \in I$ ,  $\forall a \in I$ ,  $\forall k \in R$

Συμβολισμός:  $I \triangleleft R$ .

**Σχόλιο:** Στην περίπτωση όπου ο δακτύλιος είναι μεταθετικός, τότε λόγω του ορισμού του, η συνθήκη (c) του παραπάνω ορισμού γίνεται:

$$k \cdot a \in I, \forall a \in I, \forall k \in R$$

Για παράδειγμα, αν θεωρήσουμε τον  $\mathbb{Z}$  και  $n \in \mathbb{Z}$  τότε όλα τα ιδεώδη αυτού είναι της μορφής:

$n\mathbb{Z} = \{nx | x \in \mathbb{Z}\} = \langle n \rangle$ , όπου  $\langle 0 \rangle = \{0\}$ ,  $\langle 2 \rangle$  είναι οι άρτιοι αριθμοί,  $\langle 3 \rangle$  τα πολλαπλάσια του 3 κ.ο.κ. Σημειώνεται ότι  $a \in n\mathbb{Z}$  αν και μόνο αν το  $n$  διαιρεί το  $a$ .



Ένα ακόμα παράδειγμα αποτελεί ο  $\mathbb{Z}_{12}$ , όπου όλα τα ιδεώδη του είναι τα:

$$\langle 0 \rangle = \{\bar{0}\},$$

$$\langle 1 \rangle = \{\bar{0}, \bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}, \bar{7}, \bar{8}, \bar{9}, \bar{10}, \bar{11}\} = \mathbb{Z}_{12} = \langle 5 \rangle = \langle 7 \rangle = \langle 11 \rangle,$$

$$\langle 2 \rangle = \{\bar{0}, \bar{2}, \bar{4}, \bar{6}, \bar{8}, \bar{10}\} = \langle 10 \rangle,$$

$$\langle 3 \rangle = \{\bar{0}, \bar{3}, \bar{6}, \bar{9}\} = \langle 9 \rangle,$$

$$\langle 4 \rangle = \{\bar{0}, \bar{4}, \bar{8}\} = \langle 8 \rangle,$$

$$\langle 6 \rangle = \{\bar{0}, \bar{6}\}.$$

**Ορισμός 2.1.6.** Έστω  $I$  ιδεώδες του  $R$  και  $\Lambda \subseteq I$  με  $I = \langle \Lambda \rangle$ . Τότε το  $\Lambda$  ονομάζεται βάση του  $I$ .

**Σχόλιο:** Οι βάσεις Gröbner αποτελούν σύνολα γεννητόρων για ένα ιδεώδες με εξαιρετικές ιδιότητες. Βασική προϋπόθεση είναι ότι τα σύνολα αυτά είναι πεπερασμένα, ώστε ο αλγόριθμος υπολογισμού τους να είναι πεπερασμένος. Αυτό μας εξασφαλίζει το Θεώρημα Βάσης του Hilbert [5].

**Ορισμός 2.1.7.** Έστω  $R$  δακτύλιος. Ο  $R$  θα είναι δακτύλιος της Noether, αν κάθε ιδεώδες  $I$  του  $R$  είναι πεπερασμένα παραγόμενα, αν δηλαδή  $\exists f_1, \dots, f_s \in R : I = \langle f_1, \dots, f_s \rangle, \forall I \triangleleft R$ .

Ένα παράδειγμα ενός δακτυλίου Noether είναι το  $\mathbb{Z}$ , αφού όλα τα ιδεώδη του είναι της μορφής  $n\mathbb{Z} = \langle n \rangle$ , συνεπώς είναι πεπερασμένα παραγόμενα.

**Θεώρημα 2.1.1.** (Θεώρημα βάσης Hilbert) Αν ο δακτύλιος  $R$  είναι δακτύλιος της Noether, τότε και ο πολυωνυμικός δακτύλιος  $R[x]$  είναι δακτύλιος της Noether.

## 2.2 Μονωνυμικές Διατάξεις

Για την περαιτέρω ανάπτυξη της θεωρίας και ιδίως αυτής των βάσεων Gröbner, απαραίτητο εργαλείο είναι οι μονωνυμικές διατάξεις. Αυτές είναι άπειρες, αλλά περαιτέρω αναπτύσσουμε ενδεικτικά κάποιες, τις περισσότερο οικείες, δηλαδή, τη λεξικογραφική, τη βαθμωτή λεξικογραφική, την αντίστροφη βαθμωτή λεξικογραφική και τέλος τη διάταξη βάρους. Πριν την παρουσίαση των μονωνυμικών διατάξεων, αξίζει να σχολιαστεί η σημαντικότητα αυτών. Σε ένα οποιοδήποτε σύνολο αριθμών αν κληθούμε να διατάξουμε τα στοιχεία του κατά αύξουσα σειρά, γνωρίζουμε πως αυτό μπορεί να επιτευχθεί. Για παράδειγμα, αν θεωρήσουμε τους αριθμούς  $-2$ ,  $e$ ,  $\pi$ ,  $\sqrt{2}$ ,  $\frac{5}{3}$  και  $3.9$  γνωρίζουμε πως διατάσσονται κατά αύξουσα σειρά ως εξής:  $-2 < \sqrt{2} < \frac{5}{3} < e < \pi < 3.9$ . Ακόμα και αν βρισκόμαστε σε έναν πολυωνυμικό δακτύλιο  $R[x]$ , μπορούμε να διατάξουμε τα μονώνυμά του με βάση τον βαθμό του εκάστοτε μονωνύμου, δηλαδή:  $1 < x < x^2 < x^3 < \dots x^n < \dots$ . Τι συμβαίνει όμως όταν βρισκόμαστε σε πολυωνυμικό δακτύλιο με δύο ή και περισσότερες μεταβλητές; Αν λοιπόν έχουμε τα πολυώνυμα  $x^3y^3$  και  $x^2y^4$  του  $K[x, y]$ , τότε ποιο από τα δυο είναι μεγαλύτερο; Όπως θα δούμε και παρακάτω την απάντηση αυτή μας τη δίνουν οι μονωνυμικές διατάξεις.

**Ορισμός 2.2.1.** Μερική διάταξη ενός συνόλου  $X$  είναι μια σχέση  $\prec$  η οποία ικανοποιεί τα εξής:

- a) δεν ισχύει  $a \prec a$ ,  $\forall a \in X$
- b) αν  $a \prec b$  και  $b \prec c$  τότε  $a \prec c$ ,  $\forall a, b, c \in X$ .

Παράδειγμα μερικής διάταξης αποτελεί το  $\mathbb{N}$  με την διάταξη  $<$ , καθώς για κάθε φυσικό αριθμό  $a$  δεν ισχύει  $a < a$  και αν  $a < b$ , με  $a, b \in \mathbb{N}$  και  $b < c$

με  $b, c \in \mathbb{N}$  ισχύει ότι  $a < c$ .

**Ορισμός 2.2.2.** Ολική διάταξη ενός συνόλου  $X$  είναι μια μερική διάταξη  $\preceq$  για την οποία ισχύει ότι για οποιαδήποτε  $a, b \in X$  με  $a \neq b$  είτε  $a \prec b$  είτε  $b \prec a$ .

Η σχέση  $\leq$  στο  $\mathbb{Q}$  αποτελεί σχέση ολικής διάταξης, αφού είναι διάταξη και ικανοποιεί και την ιδιότητα της τριχοτομίας ( $a, b \in \mathbb{Q}$  με  $a \neq b$  είτε  $a < b$  είτε  $b < a$ ).

**Ορισμός 2.2.3.** Μια μονωνυμική διάταξη στο  $T^n$  είναι μια ολική διάταξη  $<$  για την οποία ισχύουν:

- a) Αν  $1 \neq x^a$  τότε  $1 < x^a, \forall x^a \in T^n$
- b) Αν  $x^a, x^b \in T^n$  με  $x^a < x^b$  τότε  $x^a x^c < x^b x^c, \forall x^c \in T^n$ .

Παραδείγματα μονωνυμικών διατάξεων παρουσιάζονται παρακάτω.

Για να οριστεί μια μονωνυμική διάταξη, πρέπει να ορίζουμε εξ αρχής μια διάταξη μεταξύ των μεταβλητών του δακτυλίου. Στα παρακάτω δεχόμαστε την σύμβαση  $x_1 > x_2 > \dots > x_n$ , η οποία αν αλλάξει θα αναφέρεται.

### 2.2.1 Λεξικογραφική Διάταξη

Η λεξικογραφική διάταξη  $<_{lex}$  στο  $T^n$  με  $x_1 > x_2 > \dots > x_n$  ορίζεται ως εξής: Έστω  $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n) \in \mathbb{N}^n$ , έχουμε:  $x^a <_{lex} x^b \Leftrightarrow$  οι πρώτες συνιστώσες από τα αριστερά,  $a_i, b_i$  που είναι διαφορετικές, ικανοποιούν την σχέση  $a_i < b_i$ .

**Σχόλιο:** Στη λεξικογραφική μεγαλύτερο είναι το μονώνυμο με τη μεγαλύτερη μεταβλητή (ως προς τη διάταξη μεταβλητών που έχουμε αρχικά οριστεί). Αν έχουν το ίδιο πλήθος αυτής, τότε ελέγχουμε την επόμενη μεταβλητή κ.ο.κ.

**Παράδειγμα 3.** Έστω πολυωνυμικός δακτύλιος  $K[x_1, x_2, x_3]$  με  $x_1 > x_2 > x_3$ . Θέλουμε να διατάξουμε με βάση την λεξικογραφική διάταξη τα μονώνυμα:  $x^{a_1} = x_1^3, x^{a_2} = x_1x_2^2, x^{a_3} = x_1x_2x_3, x^{a_4} = x_2^7x_3^{10}$ . Παρατηρούμε ότι  $a_1 = (3, 0, 0), a_2 = (1, 2, 0), a_3 = (1, 1, 1), a_4 = (0, 7, 10)$ . Με βάση τον ορισμό, αρχικά θα πρέπει να συγκρίνουμε την πρώτη συνιστώσα του κάθε  $a_i$ . Έτσι, παρατηρούμε πως επειδή το πρώτο μονώνυμο έχει την μεγαλύτερη πρώτη συνιστώσα, είναι ίση με 3, έναντι των άλλων τριών που είναι κατά σειρά 1, 1 και 0. Άρα το μεγαλύτερο είναι το  $x_1^3$ . Συνεχίζουμε, κοιτώντας την πρώτη συνιστώσα των υπόλοιπων μονωνύμων και παρατηρούμε πως είναι είτε 1 είτε 0, οπότε, αφού σε δύο από τα τρία εναπομείναντα έχουμε συνιστώσα ίση με 1, συγκρίνουμε αυτά τα δύο με βάση την δεύτερη συνιστώσα, συνεπώς το δεύτερο είναι το αμέσως μεγαλύτερο. Οπότε, μέχρι τώρα έχουμε  $x_1x_2^2 <_{lex} x_1^3$ . Για να εντοπίσουμε το αμέσως επόμενο, συγκρίνουμε τα δυο εναπομείναντα με βάση την πρώτη συνιστώσα και έτσι έχουμε συνολικά ότι  $x_2^7x_3^{10} <_{lex} x_1x_2x_3 <_{lex} x_1x_2^2 <_{lex} x_1^3$ .

### 2.2.2 Βαθμωτή Λεξικογραφική Διάταξη

Η βαθμωτή λεξικογραφική διάταξη  $<_{deglex}$  στο  $T^n$  με  $x_1 > x_2 > \dots > x_n$  ορίζεται ως εξής: Έστω  $a = (a_1, \dots, a_n), b = (b_1, \dots, b_n) \in \mathbb{N}^n$ , έχουμε:

$$x^a <_{deglex} x^b \Leftrightarrow \sum_{i=1}^n a_i < \sum_{i=1}^n b_i.$$

Αν τώρα προκύψει αθροιστικά ίδιος βαθμός, δηλαδή  $\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$  τότε  $x^a <_{deglex} x^b \Leftrightarrow x^a <_{lex} x^b$ .

**Παράδειγμα 3** Έστω πολυωνυμικός δακτύλιος  $K[x_1, x_2, x_3]$  με  $x_1 > x_2 > x_3$ . Θα διατάξουμε τα μονώνυμα  $x^{a_1} = x_1^3, x^{a_2} = x_1x_2^2, x^{a_3} = x_1x_2x_3, x^{a_4} = x_2^7x_3^{10}$

με τη βαθμωτή λεξικογραφική διάταξη. Παρατηρούμε ότι  $a_1 = (3, 0, 0)$ ,  $a_2 = (1, 2, 0)$ ,  $a_3 = (1, 1, 1)$ ,  $a_4 = (0, 7, 10)$ . Οπότε, το τέταρτο μονώνυμο είναι το μεγαλύτερο με βάση τη βαθμωτή λεξικογραφική διάταξη, αφού έχει το μεγαλύτερο βαθμό. Τα τρία πρώτα όμως έχουν το ίδιο άθροισμα βαθμών, οπότε η σύγκρισή τους γίνεται με λεξικογραφική διάταξη. Έχουμε ότι

$$x_1x_2x_3 <_{deglex} x_1x_2^2 <_{deglex} x_1^3 <_{deglex} x_2^7x_3^{10}.$$

### 2.2.3 Αντίστροφη Βαθμωτή Λεξικογραφική Διάταξη

Η αντίστροφη βαθμωτή λεξικογραφική διάταξη  $<_{degrevlex}$  στο  $T^n$  με  $x_1 > x_2 > \dots > x_n$  ορίζεται ως εξής: Έστω  $a = (a_1, \dots, a_n)$ ,  $b = (b_1, \dots, b_n) \in \mathbb{N}^n$ , έχουμε:

$$x^a <_{degrevlex} x^b \Leftrightarrow \sum_{i=1}^n a_i < \sum_{i=1}^n b_i.$$

Αν τώρα προκύψει αθροιστικά ίδιος βαθμός, δηλαδή

$$\sum_{i=1}^n a_i = \sum_{i=1}^n b_i$$

τότε

$$x^a <_{degrevlex} x^b \Leftrightarrow$$

οι πρώτες διαφορετικές συνιστώσες από τα δεξιά  $a_i, b_i$  ικανοποιούν την σχέση  $a_i > b_i$ .

**Παράδειγμα 3** Έστω πολυωνυμικός δακτύλιος  $K[x_1, x_2, x_3]$  με  $x_1 > x_2 > x_3$ . Θα διατάξουμε τα μονώνυμα  $x^{a_1} = x_1^3$ ,  $x^{a_2} = x_1x_2^2$ ,  $x^{a_3} = x_1x_2x_3$ ,  $x^{a_4} = x_2^7x_3^{10}$  με αντίστροφη βαθμωτή λεξικογραφική διάταξη. Παρατηρούμε ότι  $a_1 =$

$(3, 0, 0)$ ,  $a_2 = (1, 2, 0)$ ,  $a_3 = (1, 1, 1)$ ,  $a_4 = (0, 7, 10)$ , οπότε, με βάση την αντίστροφη βαθμωτή λεξικογραφική διάταξη, αφού το τέταρτο μονώνυμο έχει τον μεγαλύτερο βαθμό, είναι μεγαλύτερο από τα υπόλοιπα. Από εκεί και πέρα όμως, επειδή τα εναπομείναντα έχουν τον ίδιο βαθμό, αυτό που θα διατάξει τα μονώνυμα είναι το πλήθος των λιγότερων  $x_3$ . Παρατηρούμε πως τα δύο πρώτα δεν έχουν καθόλου  $x_3$ , συνεπώς προηγούνται στην διάταξη. Από τα δύο αυτά όμως, το μεγαλύτερο είναι αυτό που έχει τα λιγότερα  $x_2$ , δηλαδή το πρώτο. Οπότε, με βάση την αντίστροφη λεξικογραφική διάταξη έχουμε ότι

$$x_1x_2x_3 <_{deg_{revlex}} x_1x_2^2 <_{deg_{revlex}} x_1^3 <_{deg_{revlex}} x_2^7x_3^{10}.$$

## 2.2.4 Διάταξη Βάρους

**Ορισμός 2.2.4.** Βαθμός βάρους  $w = (w_1, \dots, w_n) \in \mathbb{R}^n$  ενός μονωνύμου  $x^u = x_1^{u_1} \cdots x_n^{u_n}$  είναι ο

$$deg_w(x^u) = u_1w_1 + \dots + u_nw_n, \quad u = (u_1, \dots, u_n) \in \mathbb{N}^n.$$

**Ορισμός 2.2.5.** Έστω διάνυσμα  $w = (w_1, \dots, w_n) \in \mathbb{R}_{\geq 0}^n$  και  $\prec$  μονωνυμική διάταξη. Τότε, η διάταξη βάρους ( $<_{(w, \prec)}$ ) στο  $T^n$  με  $x_1 > x_2 > \dots > x_n$  ορίζεται ως εξής:

$$x^a <_{(w, \prec)} x^b \Leftrightarrow deg_w(x^a) < deg_w(x^b).$$

Αν

$$deg_w(x^a) = deg_w(x^b), \text{ τότε } x^a <_{(w, \prec)} x^b \Leftrightarrow x^a \prec x^b,$$

ως προς την μονωνυμική διάταξη που έχει επιλεχθεί εξ αρχής,  $\prec$ .

**Παράδειγμα 3** Έστω πολυωνυμικός δακτύλιος  $K[x_1, x_2, x_3]$  με  $x_1 > x_2 > x_3$  και διάνυσμα  $w = (1, 3, 2)$ . Θα διατάξουμε τα μονώνυμα  $x_1^3, x_1x_2^2, x_1x_2x_3, x_2^7x_3^{10}$  με διάταξη βάρους, υπεράνω της λεξικογραφικής διάταξης. Πριν τα διατάξουμε όμως είναι απαραίτητη η εύρεση των βαθμών βάρους, οπότε έχουμε:

- $deg_w(x_1^3) = 1 \cdot 3 + 3 \cdot 0 + 2 \cdot 0 = 3$
- $deg_w(x_1x_2^2) = 1 \cdot 1 + 3 \cdot 2 + 2 \cdot 0 = 7$
- $deg_w(x_1x_2x_3) = 1 \cdot 1 + 3 \cdot 1 + 2 \cdot 1 = 6$
- $deg_w(x_2^7x_3^{10}) = 1 \cdot 0 + 3 \cdot 7 + 2 \cdot 10 = 41$ .

Με βάση λοιπόν την διάταξη βάρους, μεγαλύτερο μονώνυμο είναι αυτό με τον μεγαλύτερο βαθμό, εκτός και αν έχουμε δυο ίσους βαθμούς, οπότε και χρησιμοποιείται η λεξικογραφική διάταξη:

$$x_1^3 <_{wlex} x_1x_2x_3 <_{wlex} x_1x_2^2 <_{wlex} x_2^7x_3^{10}.$$

**Σχόλιο:** Αποδεικνύεται ότι υπάρχουν άπειρες το πλήθος μονωνυμικές διατάξεις, αφού από τις υπάρχουσες (γνωστές) μπορούν να προκύψουν και άλλες. Μάλιστα, με βάση την διάταξη βάρους μπορούμε πλέον να κατασκευάζουμε μονωνυμικές διατάξεις, ανάλογα με το πρόβλημα που επιθυμούμε να επιλύσουμε. Αυτή η κατασκευή είναι πολύ σημαντική στην επίλυση προβλημάτων στη θεωρία των βάσεων Gröbner. Ειδικότερα, σε ερευνητικά προβλήματα και στην προσπάθεια απόδειξης ότι ένα σύνολο αποτελεί βάση Gröbner για ένα ιδεώδες, το πρόβλημα ανάγεται στην εύρεση κατάλληλης μονωνυμικής διάταξης στον πολυωνυμικό δακτύλιο που δουλεύουμε. Βασικό εργαλείο στην προσέγγιση αυτή είναι η διάταξη βάρους.

## 2.3 Διαίρεση Πολυωνύμων

Θεωρούμε μια μονωνυμική διάταξη  $<$  με  $x_1 > \dots > x_n$  στον  $K[x_1, \dots, x_n]$ . Κάθε μη μηδενικό πολυώνυμο, έχοντας διαγράψει τους μηδενικούς όρους του  $f$  και διατάσσοντας τους υπόλοιπους όρους του ως προς την  $<$ ,  $f \in K[x_1, \dots, x_n]$  έχει την μορφή:  $f = c_1x^{a_1} + \dots + c_kx^{a_k}$ ,  $a_i \in \mathbb{N}^n$ ,  $c_i \neq 0$  με  $c_1x^{a_1} > \dots > c_kx^{a_k}$ . Τότε ο όρος  $c_1x^{a_1}$  καλείται αρχικός όρος του  $f$  και συμβολίζεται με  $lt(f)$ . Το  $c_1$  καλείται αρχικός συντελεστής του  $f$  και συμβολίζεται με  $lc(f)$ . Τέλος, το μονώνυμο  $x^{a_1}$  καλείται αρχικό μονώνυμο και συμβολίζεται με  $lm(f)$ .

**Ορισμός 2.3.1.** Θεωρούμε πολυώνυμα  $f, g, h \in K[x_1, \dots, x_n]$ ,  $g \neq 0$ . Θα λέμε ότι το  $f$  ανάγεται στο  $h$  μόδιο  $g$  σε ένα βήμα και θα συμβολίζουμε με  $f \xrightarrow{g} h$  αν και μόνο αν

- a) το αρχικό μονώνυμο  $lm(g)$  διαιρεί ένα μη μηδενικό όρο  $X$  του  $f$ .
- b)  $h = f - \frac{X}{lt(g)}g$ .

**Ορισμός 2.3.2.** Έστω  $f, h, f_1, \dots, f_s \in K[x_1, \dots, x_n]$ ,  $f_i \neq 0$ ,  $\forall i : 1 \leq i \leq s$  και  $F = \{f_1, \dots, f_s\}$ . Θα λέμε ότι το  $f$  ανάγεται στο  $h$  μόδιο  $F$  και θα συμβολίζουμε με  $f \xrightarrow{F}_+ h$  αν και μόνο αν υπάρχει ακολουθία δεικτών  $\{i_1, \dots, i_t\} \subseteq \{1, \dots, s\}$  και ακολουθία πολυωνύμων  $\{h_1, \dots, h_{t-1}\} \subseteq K[x_1, \dots, x_n]$  τέτοιο ώστε

$$f \xrightarrow{f_{i_1}} h_1 \xrightarrow{f_{i_2}} h_2 \dots \xrightarrow{f_{i_{t-1}}} h_{t-1} \xrightarrow{f_{i_t}} h.$$

**Ορισμός 2.3.3.** Ένα πολυώνυμο  $p$  καλείται ανάγωγο μόδιο  $F = \{f_1, \dots, f_s\}$  αν  $p = 0$  ή αν για κάθε  $i \in \{1, \dots, s\}$  το αρχικό μονώνυμο  $lm(f_i)$  δεν διαιρεί κάποιον από τους όρους του  $p$ .

**Παράδειγμα 4.** Θεωρούμε  $f = x_1x_2^2x_3^2 + x_1x_2 - x_2x_3 \in S = \mathbb{Q}[x_1, x_2, x_3]$  και  $F = \{f_1 = x_1 - x_2^2, f_2 = x_2 - x_3^3, f_3 = x_3^2 - 1, f_1, f_2, f_3 \in S\}$ . Θεωρούμε



την λεξικογραφική διάταξη με  $x_1 > x_2 > x_3$ . Αρχικά ελέγχουμε αν έχουμε διατάξει σωστά τα πολυώνυμα με βάση την διάταξη, όπως και είναι σωστά διατεταγμένα. Ακόμα, έχουμε  $lm(f_1) = x_1$ ,  $lm(f_2) = x_2$ ,  $lm(f_3) = x_3^2$ , τα οποία διαιρούν τουλάχιστον έναν όρο του  $f$ . Συνεπώς ξεκινάμε την διαδικασία αναγωγής με οποιοδήποτε από τα πολυώνυμα  $f_1, f_2, f_3$ , έστω το  $f_3$ . Έχουμε:

$$f \xrightarrow{f_3} f - \frac{x_1 x_2^2 x_3^2}{x_3^2} (x_3^2 - 1) = x_1 x_2^2 + x_1 x_2 - x_2 x_3 = u_1.$$

Το  $lm(f_3)$  δεν διαιρεί κανέναν όρο του  $u_1$ , οπότε συνεχίζουμε την διαδικασία με το  $f_1$ .

$$u_1 \xrightarrow{f_1} x_1 x_2^2 + x_1 x_2 - x_2 x_3 - \frac{x_1 x_2^2}{x_1} (x_1 - x_2^2) = x_1 x_2 + x_2^4 - x_2 x_3 = u_2.$$

Το  $lm(f_1)$  διαιρεί τον πρώτο όρο του  $u_2$ , οπότε συνεχίζουμε με το  $f_1$ .

$$u_2 \xrightarrow{f_1} x_1 x_2 + x_2^4 - x_2 x_3 - \frac{x_1 x_2}{x_1} (x_1 - x_2^2) = x_2^4 + x_2^3 - x_2 x_3 = u_3.$$

Το  $lm(f_1)$  δεν διαιρεί κανέναν όρο του  $u_3$ , οπότε συνεχίζουμε με το  $f_2$ .

$$u_3 \xrightarrow{f_2} x_2^4 + x_2^3 - x_2 x_3 - \frac{x_2^4}{x_2} (x_2 - x_3^3) = x_2^3 x_3^3 + x_2^3 - x_2 x_3$$

$$\xrightarrow{f_2} x_2^3 + x_2^2 x_3^6 - x_2 x_3 \xrightarrow{f_2} x_2^2 x_3^3 + x_2^2 x_3^6 - x_2 x_3$$

$$\xrightarrow{f_2} x_2^2 x_3^3 + x_2 x_3^9 - x_2 x_3 \xrightarrow{f_2} 2x_2 x_3^9 - x_2 x_3 \xrightarrow{f_2} -x_2 x_3 + 2x_3^{12} \xrightarrow{f_2} 2x_3^{12} - x_3^4 = u_4.$$

Το  $f_2$  αλλά και το  $f_1$  δεν διαιρούν κανέναν όρο του  $u_4$ , συνεπώς συνεχίζουμε την διαδικασία με το  $f_3$ .

$$u_4 \xrightarrow{f_3} 2x_3^{12} - x_3^4 - \frac{2x_3^{12}}{x_3^2} (x_3^2 - 1) \xrightarrow{f_3} 2x_3^{10} - x_3^4 \xrightarrow{f_3} 2x_3^8 + x_3^4 \xrightarrow{f_3} 2x_3^6 - x_3^4$$

$$\xrightarrow{f_3} x_3^4 \xrightarrow{f_3} x_3^2 \xrightarrow{f_3} 1 = u,$$

που είναι ανάγωγο ως προς τα  $f_1, f_2, f_3$ . Συνεπώς, το υπόλοιπο της διαίρεσης του  $f$  με τα στοιχεία του  $F$  είναι το  $u = 3$ .

**Θεώρημα 2.3.1.** (Ευκλείδειος αλγόριθμος στο  $K[x]$ ) Έστω σώμα  $K$  και πολυώνυμο  $\phi_1(x), \phi_2(x) \in K[x], \phi_2(x) \neq 0$ . Υπάρχουν μοναδικά πολυώνυμο  $\pi(x), v(x) \in K[x]$  τέτοια ώστε:

$$\phi_1(x) = \phi_2(x)\pi(x) + v(x), \deg(v(x)) < \deg(\pi(x)) \text{ ή } v(x) = 0.$$

**Θεώρημα 2.3.2.** (Ευκλείδειος αλγόριθμος στο  $K[x_1, \dots, x_n]$ ) Έστω πολυωνυμικός δακτύλιος  $K[x_1, \dots, x_n]$ , όπου  $K$  σώμα εφοδιασμένο με μια μονωνυμική διάταξη  $<$  και έστω το σύνολο πολυωνύμων  $F = \{\phi_1, \dots, \phi_s\} \subseteq K[x_1, \dots, x_n]$ . Κάθε πολυώνυμο  $\phi \in K[x_1, \dots, x_n]$  μπορεί να γραφεί ως:

$$\phi = \pi_1\phi_1 + \dots + \pi_s\phi_s + v, \pi_i, v \in K[x_1, \dots, x_n] \text{ και}$$

το πολυώνυμο  $v = 0$  ή το  $v$  γράφεται ως γραμμικός συνδυασμός μονωνύμων με συντελεστές στο  $K$  για τα οποία δεν υπάρχει  $i \in \{1, \dots, s\}$  τέτοιο ώστε ο αντίστοιχος αρχικός όρος  $lt(\phi_i)$  να διαιρεί κάποιο από τα μονώνυμα αυτά.

## 2.4 Βάσεις Gröbner και αλγόριθμος Buchberger

Μια βασική έννοια για την θεωρία των βάσεων Gröbner είναι αυτή του αρχικού ιδεώδους.

**Ορισμός 2.4.1.** Έστω  $I$  ιδεώδες του πολυωνυμικού δακτυλίου  $K[x_1, \dots, x_n]$  που είναι εφοδιασμένος με μια μονωνυμική διάταξη  $<$ . Αρχικό ιδεώδες του  $I$  ως προς τη διάταξη  $<$  είναι το μονωνυμικό ιδεώδες:

$$in_{<}(I) = \langle lt(f), f \in I \rangle.$$

**Ορισμός 2.4.2.** Θεωρούμε ιδεώδες  $I$  του πολυωνυμικού δακτυλίου  $K[x_1, \dots, x_n]$  εφοδιασμένου με μια μονωνυμική διάταξη  $<$ . Ένα σύνολο μη μηδενικών

πολυωνύμων  $\{f_1, \dots, f_s\} \subseteq I$  ονομάζεται βάση Gröbner του  $I$  αν

$$\langle lt(f_1), \dots, lt(f_s) \rangle = in_{<}(I).$$

Για παράδειγμα, αν θεωρήσουμε  $F = \{x_2^2 - x_1, x_2\}$  στον  $K[x_1, x_2]$  με διάταξη  $<_{lex}$  με  $x_2 > x_1$  τότε το  $F$  δεν είναι βάση Gröbner για το  $I = \langle F \rangle$  αφού:

$$x_1 = x_2x_2 - 1(x_2^2 - x_1) \in I,$$

επομένως  $x_1 \in in_{<}(I)$ , αλλά

$$x_1 \notin \langle x_2^2, x_2 \rangle = \langle x_2 \rangle = \langle lt(f), \forall f \in F \rangle.$$

**Ορισμός 2.4.3.** Θεωρούμε  $I$  ιδεώδες του πολυωνυμικού δακτυλίου  $K[x_1, \dots, x_n]$  και μια μονωνυμική διάταξη  $<$ . Το σύνολο των μη μηδενικών πολυωνύμων  $\{g_1, \dots, g_t\} \subseteq I$  θα ονομάζεται βάση Gröbner του  $I$  αν  $\forall f \in I, f \neq 0, \exists i \in \{1, \dots, t\}$  τέτοιο ώστε  $lm(g_i) | lm(f)$ .

**Πρόταση 2.4.1.** Κάθε μη μηδενικό ιδεώδες  $I$  του  $K[x_1, \dots, x_n]$  έχει βάση Gröbner.

**Πρόταση 2.4.2.** Έστω  $I$  ιδεώδες του  $K[x_1, \dots, x_n]$  και  $G = \{g_1, \dots, g_t\}$  βάση Gröbner του  $I$  ως προς μια μονωνυμική διάταξη  $<$ . Το σύνολο  $\{g_1, \dots, g_t\}$  είναι ένα σύνολο γεννητόρων του  $I$ .

**Σχόλιο:** Μια σημαντική ιδιότητα των βάσεων Gröbner είναι το μοναδικό υπόλοιπο που αφήνουν μόδιο ενός συνόλου  $G$ , κάτι που γενικά δεν ισχύει.

**Θεώρημα 2.4.1.** Θεωρούμε  $G = \{g_1, \dots, g_t\} \subseteq K[x_1, \dots, x_n]$  ένα σύνολο μη μηδενικών πολυωνύμων. Το  $G$  είναι βάση Gröbner αν και μόνο αν το υπόλοιπο της διαίρεσης μόδιο  $G$  οποιουδήποτε  $f \in K[x_1, \dots, x_n]$  είναι μοναδικό.

Για τον αλγόριθμο εύρεσης βάσεως Gröbner είναι αναγκαία η χρήση των  $S$ - πολυωνύμων.

**Ορισμός 2.4.4.** Θεωρούμε  $f, g$  μη μηδενικά πολυώνυμα του  $K[x_1, \dots, x_n]$  και έστω  $L = \epsilon.κ.π.(lm(f), lm(g))$ . Τότε, το πολυώνυμο

$$S(f, g) = \frac{L}{u(f)}f - \frac{L}{u(g)}g$$

καλείται  $S$ -πολυώνυμο των  $f$  και  $g$ .

**Σχόλιο:** Ο ορισμός των βάσεων αυτών πρακτικά είναι μη εφαρμόσιμος καθώς στην ουσία είναι αδύνατον να ελέγξεις τα άπειρα πολυώνυμα που ανήκουν στο ιδεώδες αν διαιρούνται ή όχι από τα  $lm(g_i)$ . Άρα υπάρχει ανάγκη κάποιας περισσότερο αλγοριθμικής διαδικασίας. Συνέπεια αυτού, οδηγούμαστε στα  $S$ - πολυώνυμα.

**Παράδειγμα 5.** Έστω τα πολυώνυμα  $f = 2x^3y^2 + xy^5z$  και  $g = xy + xy^2z$  στον  $\mathbb{Q}[x, y, z]$  ως προς την λεξικογραφική διάταξη με  $x > y > z$ . Τότε για την εύρεση του  $S$ -πολυωνύμου έχουμε:  $lm(f) = x^3y^2$  και  $lm(g) = xy^2z$  συνεπώς,  $L = x^3y^2z$ . Οπότε:

$$\begin{aligned} S(f, g) &= \frac{x^3y^2z}{2x^3y^2z}(2x^3y^2 + xy^5z) - \frac{x^3y^2z}{xy^2z}(xy + xy^2z) = \\ &= \frac{1}{2}(2x^3y^2 + xy^5z) - x^2(xy + xy^2z) = -x^3y^2z + x^3y^2 - x^3y + \frac{1}{2}xy^5z. \end{aligned}$$

**Σχόλιο:** Τα  $S$ - πολυώνυμα είναι αυτά που μας δίνουν την αλγοριθμική διαδικασία που ψάχνουμε για την εύρεση των βάσεων Gröbner.

**Θεώρημα 2.4.2.** (Buchberger) Θεωρούμε  $G = \{g_1, \dots, g_t\}$  σύνολο μη μηδενικών πολυωνύμων του  $K[x_1, \dots, x_n]$ . Τότε το  $G$  είναι βάση Gröbner για το ιδεώδες  $I = \langle g_1, \dots, g_t \rangle$  αν και μόνο αν  $\forall i \neq j$  ισχύει ότι

$$S(g_i, g_j) \xrightarrow{G} 0.$$

**Παράδειγμα 6.** Έστω  $G = \{f_1 = x^2 - y^3, f_2 = y^2 - z^3, f_3 = z^2 - w^3\}$  στον  $\mathbb{Q}[x, y, z, w]$ , με  $x > y > z > w$ . Θεωρούμε την λεξικογραφική διάταξη. Θα δείξουμε ότι η  $G$  είναι βάση Gröbner. Πράγματι,  $lm(f_1) = x^2$ ,  $lm(f_2) = y^2$ ,  $lm(f_3) = z^2$ . Έχουμε:

$$S(f_1, f_2) = \frac{L}{lt(f_1)}f_1 - \frac{L}{lt(f_2)}f_2 = \frac{x^2y^2}{x^2}(x^2 - y^3) - \frac{x^2y^2}{y^2}(y^2 - z^3) = x^2z^3 - y^5.$$

Κάνουμε αναγωγή σε μόνιο  $G$ .

$$S(f_1, f_2) \xrightarrow{f_1} x^2z^3 - y^5 - \frac{x^2z^3}{x^2}(x^2 - y^3) = -y^3(y^2 - z^3) \xrightarrow{f_2} 0.$$

Συνεχίζουμε την ίδια διαδικασία για  $f_1, f_3$  και  $f_2, f_3$  και έχουμε

$$S(f_1, f_3) = \frac{L}{lt(f_1)}f_1 - \frac{L}{lt(f_3)}f_3 = \frac{x^2z^2}{x^2}(x^2 - y^3) - \frac{x^2z^2}{z^2}(z^2 - w^3) = x^2w^3 - y^3z^2.$$

Κάνουμε αναγωγή σε μόνιο  $G$ .

$$S(f_1, f_3) \xrightarrow{f_1} x^2w^3 - y^3z^2 - \frac{x^2w^3}{x^2}(x^2 - y^3) = -y^3(z^2 - w^3) \xrightarrow{f_3} 0.$$

Ομοίως,

$$S(f_2, f_3) = \frac{L}{lt(f_2)}f_2 - \frac{L}{lt(f_3)}f_3 = \frac{y^2z^2}{y^2}(y^2 - z^3) - \frac{y^2z^2}{z^2}(z^2 - w^3) = y^2w^3 - z^5.$$

και μόνιο  $G$ :

$$S(f_2, f_3) \xrightarrow{f_2} y^2w^3 - z^5 - \frac{y^2w^3}{y^2}(y^2 - z^3) = -z^3(z^2 - w^3) \xrightarrow{f_3} 0.$$

Αφού λοιπόν για κάθε  $i, j$  με  $i \neq j$  δείξαμε ότι  $S(f_i, f_j) \xrightarrow{G} 0$ , τότε  $G$  βάση Gröbner.

Συνεχίζουμε με τον αλγόριθμο του Buchberger για την εύρεση βάσης Gröbner ενός ιδεώδους.

**Αλγόριθμος του Buchberger:**

**Είσοδος:**  $F = \{f_1, \dots, f_s\} \subseteq K[x_1, \dots, x_n]$ ,  $f_i \neq 0$ ,  $1 \leq i \leq s$ .

**Έξοδος:**  $G = \{g_1, \dots, g_t\}$  βάση Gröbner του ιδεώδους  $\langle f_1, \dots, f_s \rangle$ .

**Αρχή:**  $G = F$ ,  $\mathcal{G} = \{f_i, f_j\} : f_i \neq f_j \in G$ . Όσο  $\mathcal{G} \neq \emptyset$  επανέλαβε

Διάλεξε οποιοδήποτε  $\{f, g\} \in \mathcal{G}$  και θέσε  $\mathcal{G} = \mathcal{G} - \{\{f, g\}\}$

$S(f, g) \xrightarrow{G}_+ h$ , όπου  $h$  ανάγωγο μόδιο  $G$ .

Αν  $h \neq 0$  τότε

$\mathcal{G} = \mathcal{G} \cup \{\{u, h\} | u \in G\}$ ,  $G = G \cup \{h\}$ .

**Σχόλιο:** Το πεπερασμένο των βημάτων του παραπάνω αλγόριθμου εξασφαλίζεται από το Θεώρημα βάσης Hilbert. Η παρακάτω πρόταση απλουστεύει αρκετά τους υπολογισμούς μας στη διαδικασία εύρεσης βάσεων των Gröbner.

**Πρόταση 2.4.3.** Θεωρούμε  $f, g \in K[x_1, \dots, x_n]$ . Αν  $\mu.κ.δ.(lt(f), lt(g)) = 1$  τότε:

$$S(f, g) \xrightarrow{f, g}_+ 0.$$

**Παράδειγμα 6** Με βάση την παραπάνω πρόταση παρατηρούμε πως πράγματι το σύνολο  $G = \{f_1 = x^2 - y^3, f_2 = y^2 - z^3, f_3 = z^2 - w^3\}$  αποτελεί βάση Gröbner ως προς την λεξικογραφική διάταξη, αφού  $lt(f_1) = x^2, lt(f_2) = y^2, lt(f_3) = z^2$  και ο Μ.Κ.Δ. των τριών αυτών μονωνύμων είναι το 1, συνεπώς άμεσα από την Πρόταση 2.4.3. εξάγεται το συμπέρασμα.

**Παράδειγμα 7.** Θεωρούμε το ιδεώδες  $I = \langle x^2y + z, xz + y \rangle \subseteq \mathbb{Q}[x, y, z]$ . Αναζητούμε μια βάση Gröbner ως προς την λεξικογραφική διάταξη, με  $x <$

$y < z$ . Με βάση τον αλγόριθμο του Buchberger έχουμε: Θέτουμε  $G_1 = \{f_1 = x^2y + z, f_2 = xz + y\}$ ,  $\mathcal{G}_1 = \{\{f_1, f_2\}\}$ . Τότε,

$$S(f_1, f_2) = \frac{xz}{z}(x^2y + z) - \frac{xz}{xz}(xz + y) = x^3y - y = f_3,$$

που είναι ανάγωγο μόδιο  $G_1$ .

Άρα έχουμε  $G_2 = \{f_1 = x^2y + z, f_2 = xz + y, f_3 = x^3y - y\}$ ,  $\mathcal{G}_2 = \{\{f_1, f_3\}, \{f_2, f_3\}\}$ . Συνεπώς,

$$S(f_1, f_3) = \frac{x^3yz}{z}(x^2y + z) - \frac{x^3yz}{x^3y}(x^3y - y) = x^5y^2 + yz$$

Κάνουμε αναγωγή σε μόδιο  $G$ ,

$$S(f_1, f_3) \xrightarrow{f_1} x^5y^2 + yz - \frac{yz}{z}(x^2y + z) = x^2y(x^3y - y) \xrightarrow{f_3} 0.$$

Συνεχίζουμε με όμοιο τρόπο,

$$S(f_2, f_3) = \frac{x^3yz}{xz}(xz + y) - \frac{x^3yz}{x^3y}(x^3y - y) = y(x^2y + z) \xrightarrow{f_1} 0.$$

Άρα το  $G_2$  είναι μια βάση Gröbner του ιδεώδους  $I$ , αφού όλα τα  $S$ -πολύνομα μηδενίζονται μέσω αυτής.

**Θεώρημα 2.4.3.** Έστω σύνολο  $F = \{f_1, \dots, f_s\}$ ,  $f_i \neq 0$ . Ο αλγόριθμος Buchberger δίνει μια βάση Gröbner του ιδεώδους  $I = \langle f_1, \dots, f_s \rangle$ .

## 2.5 Ελαχιστική και ανάγωγες βάσεις Gröbner

Μέχρι εδώ έχουν αναλυθεί τα βασικά στοιχεία μιας βάσης Gröbner, χωρίς ωστόσο να σχολιασθεί η μοναδικότητα ή μη της βάσης αυτής. Η μελέτη αυτής οδηγεί στις έννοιες των ελαχιστικών και ανάγωγων βάσεων Gröbner.

Οι ελαχιστικές είναι άπειρες και οι ανάγωγες είναι μοναδική ως προς την εκάστοτε μονωνυμική διάταξη. Αν αλλάξει η διάταξη, ενδεχομένως να αλλάξει και η ανάγωγη.

**Ορισμός 2.5.1.** Θεωρούμε μια βάση Gröbner  $G = \{g_1, \dots, g_t\}$ . Η  $G$  θα καλείται ελαχιστική αν  $\forall i \in \{1, \dots, t\}$ ,  $lc(g_i) = 1$  και για οποιοδήποτε ζεύγος  $i \neq j$  το αρχικό μονώνυμο  $lm(g_i)$  δεν διαιρεί το αρχικό μονώνυμο  $lm(g_j)$ , με  $i, j \in \{1, \dots, t\}$ .

**Αλγόριθμος εύρεσης ελαχιστικής βάσης Gröbner:**

**Είσοδος:**  $F = \{f_1, \dots, f_s\} \subseteq K[x_1, \dots, x_n]$  βάση Gröbner του ιδεώδους  $I$ .

**Έξοδος:**  $G = \{g_1, \dots, g_t\}$ , μια ελαχιστική βάση του  $I$ .

**Αρχή:**  $G = \{g_1 = \frac{f_1}{lc(f_1)}, \dots, g_s = \frac{f_s}{lc(f_s)}\}$ .

Αν υπάρχουν  $i \neq j$  με  $lm(g_i) | lm(g_j)$  τότε αφαιρέσε το  $g_j$  από το  $G$ .

**Παράδειγμα 8.** Έστω  $I = \langle x^2y + z, xz + y \rangle \subseteq \mathbb{Q}[x, y, z]$  και έστω η λεξικογραφική διάταξη με  $x < y < z$ . Τότε, από το Παράδειγμα 7 έχουμε ότι μια βάση Gröbner για το ιδεώδες είναι η  $G = \{f_1 = x^2y + z, f_2 = xz + y, f_3 = x^3y - y\}$ . Η βάση αυτή δεν είναι ελαχιστική, καθώς  $lm(f_1) = z | xz = lm(f_2)$ . Συνεπώς, αφαιρούμε το  $f_2$  από την βάση Gröbner  $G$ . Έτσι προκύπτει η βάση  $G_1 = \{f_1 = x^2y + z, f_3 = x^3y - y\}$ , η οποία είναι ελαχιστική καθώς  $lm(f_1) = z \nmid x^3y = lm(f_3)$  και  $lc(f_1) = lc(f_3) = 1$ .

**Ορισμός 2.5.2.** Μια βάση Gröbner  $G = \{g_1, \dots, g_t\}$  καλείται ανάγωγη, αν  $\forall i \in \{1, \dots, t\}$ , ισχύει ότι  $lc(g_i) = 1$  και το πολυώνυμο  $g_i$  είναι ανάγωγο μόδιο  $G - \{g_j\}$ .

**Παράδειγμα 8.** (συνέχεια) Η βάση  $G_1$  αποτελεί και ανάγωγη βάση Gröbner καθώς το πολυώνυμο  $f_1$  είναι ανάγωγο μόδιο  $G_1 - \{f_3\}$ .



**Θεώρημα 2.5.1.** Έστω ιδεώδες  $I$ , με  $I \neq 0$  στον  $K[x_1, \dots, x_n]$ , εφοδιασμένος με μια μονωνυμική διάταξη  $<$ . Τότε το ιδεώδες  $I$  έχει μοναδική ανάγωγη βάση Gröbner ως προς την  $<$ .

### 2.5.1 Καθολική Βάση Gröbner

Στα ανωτέρω είδαμε ότι ένα σύνολο το οποίο είναι βάση Gröbner ως προς μια συγκεκριμένη μονωνυμική διάταξη, μπορεί να μην είναι βάση Gröbner του ιδεώδους του ως προς διαφορετική μονωνυμική διάταξη. Υπάρχει όμως η δυνατότητα να βρεθεί σύνολο πολυωνύμων το οποίο να αποτελεί βάση Gröbner ως προς οποιαδήποτε μονωνυμική διάταξη. Η βάση αυτή ονομάζεται καθολική βάση Gröbner και εισήχθει από τους V. Weispfenning και N. Schwartz την δεκαετία 1980-90.[5]

**Ορισμός 2.5.3.** Ως καθολική βάση Gröbner ενός ιδεώδους ορίζεται ως η ένωση όλων των ανάγωγων βάσεων Gröbner του ιδεώδους ως προς οποιαδήποτε μονωνυμική διάταξη. Συμβολισμός:  $\mathcal{U}_I$ .

Ο παραπάνω ορισμός πρακτικά δηλώνει ότι ένα σύνολο  $G = \{g_1, \dots, g_t\} \subseteq I$  του  $K[x_1, \dots, x_n]$  είναι καθολική βάση Gröbner του  $I$  αν και μόνο αν είναι βάση Gröbner ως προς οποιαδήποτε μονωνυμική διάταξη.

Αξίζει να σημειωθεί ότι οι καθολικές βάσεις Gröbner αποτελούν σημαντικό κομμάτι της Υπολογιστικής Άλγεβρας, ιδιαίτερος σε ερευνητικό επίπεδο. Αυτό οφείλεται στο γεγονός ότι η βάσεις Gröbner συνοδεύονται από την διάταξη και όπως έχουμε προαναφέρει, αν αλλάξει η διάταξη, δεν έχουμε απαραίτητα βάση Gröbner. Γενικά ο υπολογισμός μιας καθολικής βάσης Gröbner ενός ιδεώδους είναι εξαιρετικά δύσκολος υπολογισμός. Απόρροια

αυτού, ότι υπάρχουν πολύ λίγα ιδεώδη (κλάσεις ιδεωδών) των οποίων γνωρίζουμε την καθολική βάση Gröbner.

## 2.6 Εφαρμογές των βάσεων Gröbner

### 2.6.1 Επίλυση συστημάτων

Θεωρούμε σώμα  $K$  και  $K[x_1, \dots, x_n]$  τον αντίστοιχο πολυωνυμικό δακτύλιο. Γνωρίζουμε πως τα στοιχεία του  $K[x_1, \dots, x_n]$  θα είναι της μορφής

$$f(x_1, \dots, x_n) = \sum_a c_a x_1^{a_1} \cdots x_n^{a_n}, a = (a_1, \dots, a_n) \in \mathbb{Z}_{\geq 0}^n.$$

Παρακάτω θα αναλύσουμε τον τρόπο με τον οποίο μέσω των βάσεων Gröbner μπορεί να επιλυθεί ένα σύστημα πολυωνυμικών εξισώσεων της μορφής:

$$f_1(x_1, \dots, x_n) = 0$$

...

$$f_s(x_1, \dots, x_n) = 0.$$

Πριν προχωρήσουμε σε αυτό ωστόσο θα αναφέρουμε κάποιες βασικές έννοιες.

**Ορισμός 2.6.1.** Θεωρούμε  $K$  σώμα και ιδεώδες  $I = \langle f_1, \dots, f_s \rangle \subset K[x_1, \dots, x_n]$ .

Ως συσχετική ποικιλότητα του ιδεώδους  $I$  ορίζεται το σύνολο σημείων:

$$V(I) = \{(a_1, \dots, a_n) \in K : f_i(a_1, \dots, a_n) = 0, \forall i = 1, \dots, s\}.$$

Για παράδειγμα, αν  $I = \langle x^2 - 1, y^2 - 1 \rangle \subseteq \mathbb{Q}[x, y]$ , τότε:

$$V_{\mathbb{Q}}(I) = \{(-1, -1), (-1, 1), (1, -1), (1, 1)\}.$$

**Θεώρημα 2.6.1.** (Nullstellensatz) Έστω  $K$  αλγεβρικά κλειστό σώμα και έστω ιδεώδες  $I \subseteq K[x_1, \dots, x_n]$ . Τότε:

$$V_K(I) = \emptyset \Leftrightarrow I = K[x_1, \dots, x_n].$$

**Σχόλιο 1:** Το  $K$  λέγεται αλγεβρικά κλειστό αν για κάθε πολυώνυμο  $f \in K[x]$  η εξίσωση  $f = 0$  έχει λύση στο  $K$ . Συμβολισμός  $\bar{K}$ .

**Σχόλιο 2:** Αν θεωρήσουμε το σύστημα πολυωνυμικών εξισώσεων:

$$\begin{cases} f_1 = 0 \\ f_2 = 0 \\ \dots \\ f_s = 0 \end{cases}$$

τότε, το σύστημα δεν έχει λύση με βάση το θεώρημα Nullstellensatz αν και μόνον αν  $1 \in \langle f_1, \dots, f_s \rangle \Leftrightarrow 1 \xrightarrow{G}_+ 0$ , με  $G$  την ανάγωγη βάση Gröbner του  $I$ .

**Θεώρημα 2.6.2.** Έστω ιδεώδες  $I = \langle f_1, \dots, f_s \rangle \subseteq K[x_1, \dots, x_n]$  και  $G = \{g_1, \dots, g_k\}$  η ανάγωγη βάση Gröbner ως προς τη διάταξη  $\prec$ . Τα εξής είναι ισοδύναμα:

1.  $V_{\bar{K}}(I) < \infty$
2.  $\forall i = 1, \dots, n, \exists j \in \{1, \dots, k\} : lm(g_j) = x_i^v, \mu\epsilon v \in \mathbb{N}$ .
3.  $dim(K[x_1, \dots, x_n]/I) < \infty$ .

Συνοπτικά, για την επίλυση ενός συστήματος πολυωνυμικών εξισώσεων αρκεί να βρούμε την ανάγωγη βάση Gröbner του ιδεώδους, να επιβεβαιώσουμε ότι ως προς την βάση αυτή δεν ισχύει  $1 \xrightarrow{G}_+ 0$  ούτως ώστε να υπάρχει λύση.

### 2.6.2 Μελέτη του διανυσματικού χώρου πηλίκου $K[x_1, \dots, x_n]/I$

**Ορισμός 2.6.2.** Έστω δακτύλιος  $R = (R, +, \cdot)$  και  $I$  ιδεώδες του  $R$ . Ο δακτύλιος  $R/I$  εφοδιασμένος με τις πράξεις:

$$+ : R/I \times R/I \rightarrow R/I, (r + I) + (s + I) \rightarrow (r + s) + I$$

και

$$\cdot : R/I \times R/I \rightarrow R/I, (r + I) \cdot (s + I) \rightarrow (r \cdot s) + I$$

ονομάζεται δακτύλιος πηλίκου του  $R$  ως προς το  $I$ . Το μηδενικό και το μοναδιαίο στοιχείο του δακτυλίου  $R/I$  ορίζονται αντίστοιχα ως  $0_{R/I} = 0_R + I$ ,  $1_{R/I} = 1_R + I$ .

Για παράδειγμα, αν θεωρήσουμε  $R = \mathbb{Z}$  και  $I = 3\mathbb{Z}$  τότε ο δακτύλιος  $\mathbb{Z}/3\mathbb{Z}$  είναι δακτύλιος πηλίκου, αφού:

$$\mathbb{Z}/3\mathbb{Z} = \{r + 3\mathbb{Z}, r \in \mathbb{Z}\} = \{0 + 3\mathbb{Z}, 1 + 3\mathbb{Z}, 2 + 3\mathbb{Z}\}.$$

Μάλιστα, το μηδενικό στοιχείο του είναι το  $0 + 3\mathbb{Z}$  αφού για  $x = 0, 1, 2$   
 $(0 + 3\mathbb{Z}) + (x + 3\mathbb{Z}) = x + 3\mathbb{Z}$ . Ακόμα, έχουμε  $(1 + 3\mathbb{Z}) + (2 + 3\mathbb{Z}) = 0 + 3\mathbb{Z}$ ,  
 $(2 + 3\mathbb{Z}) + (2 + 3\mathbb{Z}) = 1 + 3\mathbb{Z}$  και το μοναδιαίο στοιχείο είναι το  $1 + 3\mathbb{Z}$  αφού,  
 $(1 + 3\mathbb{Z}) \cdot (x + 3\mathbb{Z}) = x + 3\mathbb{Z}$ . Ακόμα, έχουμε  $(0 + 3\mathbb{Z}) \cdot (1 + 3\mathbb{Z}) = 0 + 3\mathbb{Z}$ ,  
 $(2 + 3\mathbb{Z}) \cdot (2 + 3\mathbb{Z}) = 1 + 3\mathbb{Z}$

Για τα παρακάτω θα θεωρήσουμε πως βρισκόμαστε σε έναν πολυωνυμικό δακτύλιο  $R = K[x_1, \dots, x_n]$ ,  $I$  θα είναι ένα ιδεώδες αυτού. Επίσης, το υπόλοιπο της διαίρεσης κάθε  $f \in R$  με μια βάση Gröbner  $G$  του  $I$  θα το συμβολίζουμε με  $N_G(f)$ .

**Θεώρημα 2.6.3.** Έστω πολυώνυμα  $f, g \in R$ . Τότε

$$f + I = g + I \Leftrightarrow N_G(f) = N_G(g).$$

**Σχόλιο:** Από το παραπάνω θεώρημα εξάγεται το συμπέρασμα πως τα στοιχεία του χώρου πηλίκου  $R/I$  θα είναι της μορφής:

$$N_G(f) + I, f \in R.$$

**Θεώρημα 2.6.4.** Έστω ο  $K$ -διανυσματικός χώρος  $K[x_1, \dots, x_n]/I$ . Το σύνολο

$$B = \{x^a + I, x^a \notin \text{in}_<(I)\}$$

είναι μια βάση του.

**Παράδειγμα 9.** Έστω  $I = \langle f_1 = x_1^2 - x_1x_2 - 1, f_2 = x_2^2 - x_2 \rangle$  ιδεώδες του πολυωνυμικού δακτυλίου  $\mathbb{Q}[x_1, x_2]$  εφοδιασμένο με την λεξικογραφική διάταξη με  $x_1 > x_2$ . Για να υπολογιστεί το αρχικό ιδεώδες θα πρέπει να βρεθεί μια βάση Gröbner του ιδεώδους. Αφού ο μ.κ.δ.  $(\text{lt}(f_1), \text{lt}(f_2))=1$ , έχουμε πως μια βάση Gröbner του  $I$  είναι η  $G = \{f_1, f_2\}$ , οπότε, το αρχικό ιδεώδες είναι το  $\text{in}_<(I) = \langle x_1^2, x_2^2 \rangle$ . Παρατηρούμε λοιπόν πως τα μονώνυμα που δεν ανήκουν στο αρχικό ιδεώδες είναι τα  $\{1, x_1, x_2, x_1x_2\}$ . Οπότε, με βάση το παραπάνω θεώρημα, το σύνολο

$$\{1 + I, x_1 + I, x_2 + I, x_1x_2 + I\}$$

αποτελεί μια βάση για τον  $\mathbb{Q}[x_1x_2]/I$  και συνεπώς  $\dim_{\mathbb{Q}}(\mathbb{Q}[x_1, x_2]/I) = 4$

## Κεφάλαιο 3

# Βάσεις Gröbner και Πειραματικοί Σχεδιασμοί

Στο κεφάλαιο αυτό παρουσιάζεται μια αλγεβρική προσέγγιση για την μελέτη ενός πειραματικού σχεδιασμού  $D$ . Δεδομένου δηλαδή ενός πειραματικού σχεδιασμού αναζητάμε το μοντέλο που προσδιορίζει τον σχεδιασμό αυτόν. Αυτό που προτάθηκε από τους Pistone και Wynn (1996) ήταν να θεωρούμε κάθε σχεδιασμό ως συσχετική ποικιλότητα, ως ένα σύνολο λύσεων πολυωνυμικών εξισώσεων. Ως ιδεώδες του σχεδιασμού θα θεωρούμε το σύνολο των πολυωνύμων που ρίζες αυτών είναι τα σημεία του σχεδιασμού. Η προσέγγιση αυτή αξιοποιεί τις βάσεις Gröbner για τον χαρακτηρισμό του ιδεώδους του  $D$  και του χώρου πηλίκου. Για να χρησιμοποιηθεί η μέθοδος αυτή θα πρέπει να απαντηθούν τα εξής ερωτήματα:

1. Ποιες κατηγορίες πολυωνυμικών μοντέλων προσδιορίζουν ένα πειραματικό σχεδιασμό;
2. Είναι ένα δεδομένο μοντέλο αναγνωρίσιμο από ένα συγκεκριμένο σχέδιο;

3. Ποιες είναι οι συνθήκες που πρέπει να πληροί το μοντέλο για να ισχύει η θεωρία;

Για τα παρακάτω, θα θεωρούμε  $D$  έναν πειραματικό σχεδιασμό,  $I(D)$  το ιδεώδες αυτού και  $K[x_1, \dots, x_n]$  πολυωνυμικός δακτύλιος, όπου  $K$  τυχαίο σώμα.

**Ορισμός 3.0.1.** 1. Ως πειραματικό σχέδιο  $D$  ορίζουμε ένα πεπερασμένο σύνολο διακριτών σημείων στον  $K^d$ .

2. Θεωρούμε το ιδεώδες του  $D$ ,  $I(D)$  ως το ιδεώδες που παράγεται από το  $D$ , δηλαδή το σύνολο όλων των πολυωνύμων στο  $K[x] = K[x_1, \dots, x_d]$  των οποίων τα μηδενικά περιλαμβάνουν τα σημεία σχεδιασμού.

**Θεώρημα 3.0.1.** Έστω πειραματικός σχεδιασμός  $D$ . Τότε,

(i) Το  $I(D)$  είναι η τομή των ιδεωδών του σχεδιασμού που παράγονται από τα σημεία του,

(ii) Το  $D$  είναι συσχετική ποικιλότητα του  $I(D)$ , δηλαδή σύνολο λύσεων πολυωνυμικών εξισώσεων,

(iii) Το  $I(D)$  είναι ρίζα ιδεώδους.

Για παράδειγμα, αν θεωρήσουμε τον  $2^2$  πειραματικό σχεδιασμό με  $\{0, 1\}$  τότε:

$$V(D) = \{x_1^2 - x_1, x_2^2 - x_2\}.$$

**Θεώρημα 3.0.2.** Έστω μονωνυμική διάταξη  $\tau$  στον  $K[x]$  και  $I$  ιδεώδες του  $K[x]$ . Τότε:

• Κάθε συνάρτηση  $f \in K[x]$  ανάγεται μόδιο  $I$  σε ένα μοναδικό πολυώνυμο  $r(x)$  που είναι γραμμικός συνδυασμός των μονωνύμων στο συμπλήρωμα του

αρχικού ιδεώδους του  $I$ .

• Τα στοιχεία του συνόλου  $A = \{x^a : x^a \notin \langle lt_\tau(I) \rangle\}$  είναι γραμμικώς ανεξάρτητα μόδιο  $I$ , δηλαδή,

$$\sum_{x \in A} c_a x^a = 0, \quad c_a \in K$$

αν και μόνον αν  $c_a = 0, \forall a$ .

**Σχόλιο:** Έστω  $p(x)$  πολυώνυμο στον  $K[x]$  και  $d(x)$  διαιρέτης αυτού. Τότε έχουμε από τον αλγόριθμο της διαίρεσης:

$$p(x) = q(x)d(x) + r(x),$$

όπου  $r(x)$  θα είναι το υπόλοιπο της διαίρεσης. Τότε, για  $d(x) = 0$  ορίζεται η συσχετική ποικιλότητα,  $K[x] = \{p(x)\}$  είναι δακτύλιος πολυωνύμων,

$$I = \{q(x) : d(x) = 0 \Rightarrow q(x) = 0\}$$

είναι το ιδεώδες της συσχετικής ποικιλότητας και  $\{r(x)\} = K[x]/I$ , είναι ο δακτύλιος πηλίκο. Για τον προσδιορισμό των πολυωνύμων  $r(x)$  απαιτείται η χρήση των βάσεων Gröbner. Μάλιστα, για να είναι μοναδικό το υπόλοιπο της διαίρεσης,  $r(x)$ , απαιτείται η ύπαρξη της ανάγωγης βάσης Gröbner του ιδεώδους.

**Ορισμός 3.0.2.** Έστω  $D$  πειραματικός σχεδιασμός και  $\tau$  μονωνυμική διάταξη. Η μονωνυμική βάση του συνόλου  $K[x_1, \dots, x_n]/I(D)$  είναι το σύνολο

$$Est_\tau(D) = \{x^a : x^a \notin \langle lt(g), g \in I(D) \rangle\},$$

δηλαδή η μονωνυμική βάση περιέχει στοιχεία της μορφής  $x^a$  τα οποία δεν διαιρούνται από κανέναν αρχικό όρο των στοιχείων της βάσης Gröbner του  $I(D)$ .



Για παράδειγμα, αν θεωρήσουμε τον  $2^2$  πειραματικό σχεδιασμό με  $\{0, 1\}$  και  $<_{lex}$  τότε:

$$I(D) = \langle x_1^2 - x_1, x_2^2 - x_2 \rangle.$$

Μια βάση Gröbner αυτού είναι το σύνολο  $G = \{x_1^2 - x_1, x_2^2 - x_2\}$  και  $lt(G) = \{x_1^2, x_2^2\}$ . Συνεπώς, το σύνολο  $Est_{<_{lex}} = \{1, x_1, x_2, x_1x_2\}$ .

Μπορούμε να θεωρήσουμε το  $Est_\tau(D)$  ως μια οικογένεια  $\{x^a : x^a \in L\}$  όπου  $L$  είναι το σύνολο των εκθετών των στοιχείων στο  $Est_\tau(D)$ .

**Θεώρημα 3.0.3.** Το σύνολο  $Est_\tau(D)$  περιέχει τόσα στοιχεία, όσα τα σημεία του πειραματικού σχεδιασμού.

**Ορισμός 3.0.3.** Έστω  $\tau$  μονωνυμική διάταξη και ας θεωρήσουμε έναν σχεδιασμό σημείων  $D = \{a_i \in K^d : i = 1, \dots, n\}$ . Έστω  $L$  το σύνολο των εκθετών στο  $Est_\tau(D)$ . Καλούμε πίνακα σχεδιασμού τον πίνακα:

$$X = [a(i)^a]_{i=1, \dots, n}, \quad a \in L.$$

**Θεώρημα 3.0.4.** 1. Ο  $X$  δεν είναι μοναδικός.

2. Έστω το  $d$ -διάστατο διάνυσμα  $e_i = (0, 0, \dots, 1, 0, \dots, 0)$ , όπου στην  $i$ -οστή θέση είναι το 1 και στις υπόλοιπες  $d - 1$  θέσεις είναι το μηδέν. Για κάθε  $i = 1, \dots, d$  υπάρχει διάνυσμα  $c(i) \in K^d$  τέτοιο ώστε  $Xc(i) = e_i$  και τα πολυώνυμα  $\sum_{a \in L} c(i)_a x^a$  είναι η δείκτηρια συνάρτηση του σχεδιασμού στο σημείο  $a(i)$ , δηλαδή,

$$\sum_{a \in L} c(i)_a x^a = \begin{cases} 1, & x = a(i) \\ 0, & x \neq a(i), \quad x \in D \end{cases}$$

3. Αν  $f : D \rightarrow K$  απεικόνιση και  $y = (f(a(1)), \dots, f(a(n))) \in K$  είναι το διάνυσμα των αποκρίσεων, τότε

$$f(x) = \sum_{a \in L} c(i)_a x^a,$$

όπου το διάνυσμα των παραμέτρων  $c = [c_a : a \in L] = X^{-1}y$  και  $y = (f(a(1)), \dots, f(a(n))) \in K$

**Θεώρημα 3.0.5.** Έστω  $D$  ένας πειραματικός σχεδιασμός και  $\tau$  διάταξη. Το μοντέλο

$$f(x) = \sum_{x^a \in Est_\tau(D)} \theta_a x^a$$

είναι αναγνωρίσιμο από τα σημεία του σχεδιασμού  $a(i)$  και για τις παρατηρήσεις  $y_i, \forall a(i) \in D$ , το γραμμικό σύστημα των εξισώσεων

$$y_i = \sum_{x^a \in Est_\tau(D)} \theta_a x^a(a(i))$$

έχει μια και μοναδική λύση ως προς το  $\theta_a$ .

Η μέθοδος που προτάθηκε για την μελέτη ενός πειραματικού σχεδιασμού με χρήση των βάσεων Gröbner συνοπτικά είναι:

1. Επιλέγουμε έναν σχεδιασμό  $D$ .
2. Επιλέγουμε μονωνυμική διάταξη  $\tau$ .
3. Υπολογίζουμε μια βάση Gröbner για το  $I(D)$ .
4. Ο δακτύλιος πηλίκου  $K[x_1, \dots, x_n]/I(D)$  παράγεται από μονώνυμα της μορφής:  $x^a, a \in L$ . Αυτά είναι όλα τα μονώνυμα που δεν διαιρούνται από κανένα αρχικό όρο της βάσης Gröbner και μάλιστα  $|L| = |D|$ . Το σύνολο  $L$  έχει την ιδιότητα του αρχικού ιδεώδους: αν  $a \in L$  τότε  $b \in L, 0 \leq b \leq a$ .
5. Τότε κάθε συνάρτηση  $y(x)$  στο  $D$  προσεγγίζεται κατά μοναδικό τρόπο από το πολυώνυμο:

$$f(x) = \sum_{a \in L} \theta_a x^a$$

που είναι τέτοιο ώστε  $y(x) = f(x), \forall x \in D$ .

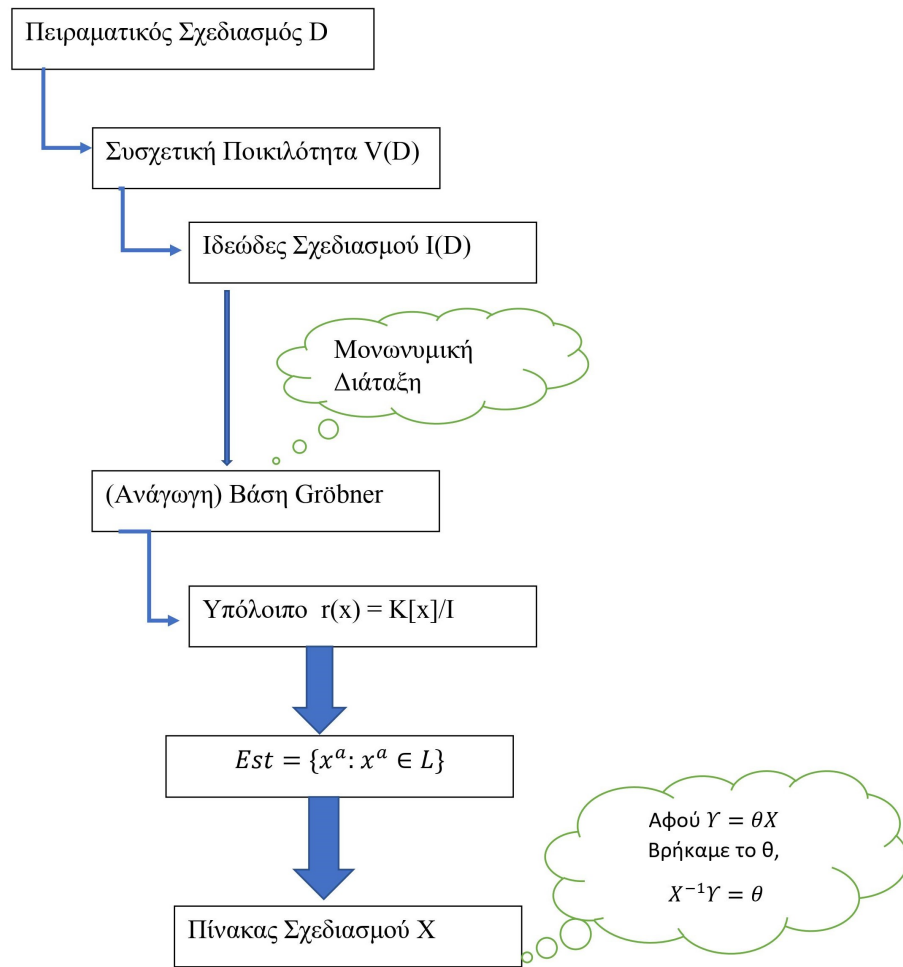
6. Το πλήθος των στοιχείων του  $D$  και των στοιχείων της βάσης  $L$  ταυτίζονται.

7. Ο πίνακας σχεδιασμού  $X$  που αποτελείται από τα σημεία του σχεδιασμού είναι πλήρους τάξεως και έχει την μορφή:  $X = \{x^a\}_{x \in D, a \in L}$ .

8. Για να βρεθούν οι εκτιμητές των συντελεστών έχουμε σε μορφή πινάκων:

$$\theta = X^{-1}y.$$

Για διευκόλυνση του αναγνώστη τα βήματα της μεθόδου αποτυπώνονται στο Σχεδιάγραμμα 1:



Σχεδιάγραμμα 1: Βήματα Μεθόδου

Στις ενότητες του παρόντος κεφαλαίου θα μελετηθούν δυο πλήρη παραγοντικά πειράματα, ένα  $2^2$  και ένα  $3^3$  εφαρμόζοντας κατάλληλα τις μεθόδους των βάσεων Gröbner. Εν συνεχεία, για λόγους πληρότητας θα γίνει σύγκριση των στατιστικών μεθόδων ανάλυσης των Πειραματικών Σχεδιασμών και της προτεινόμενης μεθόδου των βάσεων Gröbner.

### 3.1 Εφαρμογή 1: $2^2$ Πλήρες Παραγοντικό Πείραμα

Επιστρέφουμε στο **Παράδειγμα 2** του Κεφαλαίου 1 για να δούμε μια διαφορετική προσέγγισή του με χρήση της θεωρίας των βάσεων Gröbner. Τα δεδομένα του παραδείγματος δίνονται στον **Πίνακα 3**

A	B	Tr	Sum	$X_1$	$X_2$	$X_1X_2$
-	-	(1)	64.4	-1	-1	+1
+	-	a	96.1	+1	-1	-1
-	+	b	59.7	-1	+1	-1
+	+	ab	161.1	+1	+1	+1

Πίνακας 5: Δεδομένα  $2^2$  παραγοντικού πειράματος

Για τα επόμενα θεωρούμε την  $<_{lex}$ . Παρατηρούμε πως τα  $X_1, X_2$  λαμβάνουν τις τιμές  $\pm 1$ , δηλαδή οι τιμές τους είναι οι λύσεις των εξισώσεων

$$x_1^2 - 1 = 0 \text{ και } x_2^2 - 1 = 0.$$

Γνωρίζουμε ωστόσο ότι ένα πολυώνυμο δυο μεταβλητών διαιρείται διαδοχικά από τα  $x_1^2 - 1$  και  $x_2^2 - 1$  όταν είναι της μορφής

$$p(x_1, x_2) = s_1(x_1, x_2)x_1^2 + s_2(x_1, x_2)x_2^2 + r(x_1, x_2),$$

με  $r(x_1, x_2)$  το υπόλοιπο της διαίρεσης και  $s_1(x_1, x_2)$ ,  $s_2(x_1, x_2)$  πολυώνυμα. Επίσης, το  $r(x_1, x_2)$  θα έχει την μορφή

$$r(x_1, x_2) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2,$$

αφού από την Ευκλείδεια διαίρεση γνωρίζουμε πως το υπόλοιπο πρέπει να είναι μικρότερου βαθμού από τον διαιρέτη. Για την αξιοποίηση της θεωρίας των βάσεων Gröbner θα πρέπει να θεωρήσουμε ένα σώμα, στην προκειμένη περίπτωση το  $\mathbb{R}[x_1, x_2]$  και ένα ιδεώδες, το  $I = \langle f_1 = x_1^2 - 1, f_2 = x_2^2 - 1 \rangle$  και έστω  $x_1 > x_2$ . Αναζητούμε μια βάση Gröbner για το ιδεώδες.

Θέτουμε  $G_1 = \{f_1, f_2\}$  και έχουμε

$$S(f_1, f_2) = \frac{x_1^2 x_2^2}{x_1^2} (x_1^2 - 1) - \frac{x_1^2 x_2^2}{x_1^2} (x_2^2 - 1) = x_1^2 - x_2^2.$$

Κάνουμε αναγωγή σε μόνιο και έχουμε:

$$S(f_1, f_2) \xrightarrow{f_1} x_1^2 - x_2^2 - \frac{x_1^2}{x_1^2} (x_1^2 - 1) = -(x_2^2 - 1) \xrightarrow{f_2} 0.$$

Συνεπώς το  $G_1 = \{f_1, f_2\}$  είναι μια βάση Gröbner του ιδεώδους. Συνεπώς, το  $Est_{<lex}(D) = \{1, x_1, x_2, x_1 x_2\} = \{x^a\}_{a \in L}$ . Θεωρούμε μοντέλα της μορφής

$$\eta(x) = \sum \theta_a x^a + \varepsilon(x),$$

όπου το  $\sum \theta_a x^a$  είναι πεπερασμένο,  $x^a = x_1^{a_1} x_2^{a_2}$  και  $\varepsilon(x)$  τυχαία μεταβλητή με  $\varepsilon(x) \sim N(0, \sigma^2)$ .

Εμείς θα ασχοληθούμε με το μη στοχαστικό τμήμα του μοντέλου προκειμένου να προσδιορίσουμε το εκτιμώμενο μοντέλο του σχεδιασμού, δηλαδή  $\sum \theta_a x^a = y(x)$ . Το μοντέλο σε μορφή πινάκων θα είναι το  $Y = X\theta$ , όπου  $X$  ο πίνακας σχεδιασμού που είναι πλήρους τάξεως και  $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$  το διάνυσμα των παραμέτρων του μοντέλου. Με βάση λοιπόν το  $Est_{<lex}(D)$ , ο πίνακας σχεδιασμού είναι ο

$$X = \begin{pmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$$

**Πίνακας 6:** Πίνακας σχεδιασμού

Πράγματι, ο  $X$  είναι πλήρους τάξεως, αφού  $\det(X) = 16 \neq 0$ . Κατά συνέπεια, ο  $X$  αντιστρέφεται και μάλιστα

$$X^{-1} = \frac{1}{16} \begin{pmatrix} 4 & 4 & 4 & 4 \\ -4 & 4 & -4 & 4 \\ -4 & -4 & 4 & 4 \\ 4 & -4 & -4 & 4 \end{pmatrix}$$

**Πίνακας 7:** Αντίστροφος του πίνακα σχεδιασμού

Με βάση το **Θεώρημα 3.0.4** έχουμε ότι  $\theta = X^{-1}Y$ , συνεπώς έχουμε:

$$\hat{\theta} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_3 \\ \hat{\theta}_4 \end{pmatrix} = \frac{1}{16} \begin{pmatrix} 4 & 4 & 4 & 4 \\ -4 & 4 & -4 & 4 \\ -4 & -4 & 4 & 4 \\ 4 & -4 & -4 & 4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{pmatrix} \Leftrightarrow \begin{cases} \hat{\theta}_1 = \frac{1}{4}(y_1 + y_2 + y_3 + y_4) = \frac{1}{4}((1) + a + b + ab) = 23.8312 \\ \hat{\theta}_2 = \frac{1}{4}(-y_1 + y_2 - y_3 + y_4) = \frac{1}{4}(-1) + a - b + ab = 8.3187 \\ \hat{\theta}_3 = \frac{1}{4}(-y_1 - y_2 + y_3 + y_4) = \frac{1}{4}(-1) - a + b + ab = 3.7687 \\ \hat{\theta}_4 = \frac{1}{4}(y_1 - y_2 - y_3 + y_4) = \frac{1}{4}((1) - a - b + ab) = 4.3562, \end{cases}$$

απ' όπου διαπιστώνεται ότι  $y_1 = (1), y_2 = a, y_3 = b$  και  $y_4 = ab$  όπως αυτά είχαν ορισθεί και υπολογισθεί στην Ενότητα 1.2, Παράδειγμα 2. Άρα το εκτιμώμενο μοντέλο είναι το:

$$y = 23.8312 + 8.3187x_1 + 3.7687x_2 + 4.3562x_1x_2.$$

Παρατηρούμε πως πράγματι το εκτιμώμενο μοντέλο που προκύπτει με χρήση της παραπάνω αλγεβρικής προσέγγισης ταυτίζεται με αυτό της στατιστικής ανάλυσης ενός πειραματικού σχεδιασμού (Βλέπε Κεφάλαιο 1, Ενότητα 1.2).

### **3.2 Εφαρμογή 2: $3^3$ Πλήρες παραγοντικό πείραμα**

Το Example 9.3 (Montgomery, 2013) αφορά σε ένα πλήρες παραγοντικό πείραμα για τη μελέτη της επίδρασης τριών παραγόντων με τρία επίπεδα έκαστος. Συγκεκριμένα, τρεις διαφορετικοί τύποι φυαλών ( $X_1$ ) και τρεις διαφορετικοί τύποι ραφιών ( $X_2$ ), ενώ χρησιμοποιήθηκαν τρεις εργαζόμενοι ( $X_3$ ). Το πείραμα επαναλήφθηκε δυο φορές (rep I, rep II) και τα αποτελέσματα φαίνονται στον Πίνακα 7, όπου οι τιμές των παραγόντων είναι κωδικοποιημένες με τις τιμές -1, 0, 1.



$X_1$	$X_2$	$X_3$	<i>rep I</i>	<i>rep II</i>
-1	-1	-1	3.45	3.36
0	-1	-1	4.07	3.52
1	-1	-1	4.20	3.68
-1	0	-1	4.80	4.40
0	0	-1	4.52	4.44
1	0	-1	4.96	4.39
-1	1	-1	4.08	3.65
0	1	-1	4.30	4.04
1	1	-1	4.17	3.88
-1	-1	0	4.14	4.19
0	-1	0	4.38	4.26
1	-1	0	4.26	4.37
-1	0	0	5.22	4.70
0	0	0	5.15	4.65
1	0	0	5.17	4.75
-1	1	0	3.94	4.08
0	1	0	4.53	4.08
1	1	0	4.86	4.48

$X_1$	$X_2$	$X_3$	<i>rep I</i>	<i>rep II</i>
-1	-1	1	5.80	5.23
0	-1	1	5.48	4.85
1	-1	1	5.67	5.58
-1	0	1	6.21	5.88
0	0	1	6.25	6.20
1	0	1	6.03	4.37
-1	1	1	5.14	4.49
0	1	1	4.99	4.59
1	1	1	4.85	4.90

**Πίνακας 8:** Δεδομένα  $3^3$  παραγοντικού πειράματος

Εφαρμόζουμε τη διαδικασία για την εύρεση βάσης Gröbner και στη συνέχεια θα βρούμε το σύνολο  $Est$  για τον σχεδιασμό. Παρατηρούμε πως οι τιμές των παραγόντων αποτελούν λύσεις των εξισώσεων  $x_1^3 - x_1 = 0$ ,  $x_2^3 - x_2 = 0$ ,  $x_3^3 - x_3 = 0$ . Επίσης, για να διαιρείται ένα πολυώνυμο διαδοχικά από τα  $x_1^3 - x_1$ ,  $x_2^3 - x_2$ ,  $x_3^3 - x_3$ , θα πρέπει να είναι της μορφής:

$$p(x_1, x_2, x_3) = s_1(x_1, x_2, x_3)x_1^3 + s_2(x_1, x_2, x_3)x_2^3 + s_3(x_1, x_2, x_3)x_3^3 + r(x_1, x_2, x_3).$$

Θεωρούμε  $I = \langle x_1^3 - x_1, x_2^3 - x_2, x_3^3 - x_3 \rangle$  και έστω διάταξη  $<_{deglex}$ . Το αρχικό ιδεώδες αυτού είναι το (με οποιαδήποτε από τις διατάξεις λεξικογραφική, βαθμωτή λεξικογραφική και αντίστροφη βαθμωτή λεξικογραφική) το  $lt(I) = \langle x_1^3, x_2^3, x_3^3 \rangle$ . Μια βάση Gröbner του αρχικού ιδεώδους είναι η  $G = \{x_1^3, x_2^3, x_3^3\}$  η οποία είναι και καθολική βάση. Άρα, το σύνολο  $Est$ , το οποίο αποτελείται από τους γνήσιους διαιρέτες των στοιχείων της βάσης [15], είναι το

$$Est_{<deglex}(D) = \{1, x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3, x_2x_3, x_1x_2x_3, x_1^2x_2, x_1^2x_3, x_1x_2^2, x_1x_3^2, x_2^2x_3, x_2x_3^2, x_1^2x_2^2, x_1^2x_3^2, x_2^2x_3^2, x_1^2x_2x_3, x_1x_2^2x_3, x_1x_2x_3^2, x_1^2x_2^2x_3, x_1^2x_2x_3^2, x_1x_2^2x_3^2, x_1^2x_2^2x_3^2\}.$$

Θεωρούμε μοντέλα της μορφής

$$\eta(x) = \sum \theta_a x^a + \varepsilon(x),$$

όπου το  $\sum \theta_a x^a$  είναι πεπερασμένο,  $x^a = x_1^{a_1} x_2^{a_2} x_3^{a_3}$  και  $\varepsilon(x)$  τυχαία μεταβλητή με  $\varepsilon(x) \sim N(0, \sigma^2)$ .

Όπως και προηγουμένως, ασχολούμεθα με το μη στοχαστικό τμήμα του μοντέλου προκειμένου να προσδιορίσουμε το εκτιμώμενο μοντέλο του σχεδιασμού, δηλαδή  $\sum \theta_a x^a = y(x)$ . Το μοντέλο σε μορφή πινάκων θα είναι το  $Y = X\Theta$ , όπου  $X$  ο πίνακας σχεδιασμού που είναι πλήρους τάξεως. Με βάση και το  $Est_{<deglex}(D)$ , ο πίνακας σχεδιασμού  $X$  είναι:

1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	-1	-1	-1	1	1	1	1	1	1	-1	-1	-1	1
1	0	-1	-1	0	1	1	0	0	1	0	0	0	0	0	-1	-1	0	0	1	0	0	0	0	0	0
1	1	-1	-1	1	1	1	-1	1	1	1	-1	-1	1	1	-1	-1	1	1	1	1	-1	-1	-1	-1	1
1	-1	0	-1	1	0	1	0	1	0	0	0	-1	0	-1	0	0	0	1	0	0	0	0	0	0	0
1	0	0	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	-1	1	0	1	0	-1	0	0	0	-1	0	1	0	0	0	1	0	0	0	0	0	0	0
1	-1	1	-1	1	1	1	-1	1	-1	1	1	-1	-1	-1	-1	1	1	1	1	-1	1	-1	-1	-1	1
1	0	1	-1	0	1	1	0	0	-1	0	0	0	0	0	0	-1	1	0	0	1	0	0	0	0	0
1	1	1	-1	1	1	1	1	-1	-1	-1	1	-1	1	1	-1	1	1	1	1	-1	-1	1	-1	1	1
1	-1	-1	0	1	1	0	1	0	0	0	-1	0	-1	0	0	0	1	0	0	0	0	0	0	0	0
1	0	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	-1	0	1	1	0	-1	0	0	0	-1	0	1	0	0	0	1	0	0	0	0	0	0	0	0
1	-1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	-1	1	0	1	1	0	-1	0	0	0	1	0	-1	0	0	0	1	0	0	0	0	0	0	0	0
1	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	1	0	1	1	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0	0
1	-1	-1	1	1	1	1	1	-1	-1	1	-1	1	-1	-1	1	-1	1	1	1	-1	-1	1	1	-1	-1
1	0	-1	1	0	1	1	0	0	-1	0	0	0	0	0	0	1	-1	0	0	1	0	0	0	0	0
1	1	-1	1	1	1	1	-1	1	-1	-1	-1	1	1	1	1	-1	1	1	1	-1	1	-1	1	-1	1
1	-1	0	1	1	0	1	0	-1	0	0	0	1	0	-1	0	0	0	1	0	0	0	0	0	0	0
1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	1	1	0	1	0	1	0	0	0	1	0	1	0	0	0	1	0	0	0	0	0	0	0
1	-1	1	1	1	1	1	-1	-1	1	-1	1	1	-1	-1	1	1	1	1	1	1	-1	-1	1	1	-1
1	0	1	1	0	1	1	0	0	1	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

**Πίνακας 9**

Παρατηρούμε ότι κάθε στήλη του  $X$  αντιστοιχεί (με την ίδια σειρά) στα στοιχεία του  $Est_{<deglex}(D)$ . Ειδικότερα παρατηρούμε ότι η πρώτη στήλη του  $X$  έχει μόνο 1 και αντιστοιχεί στον σταθερό όρο του μοντέλου, ενώ οι επόμενες τρεις στήλες είναι οι στήλες  $X_1, X_2$  και  $X_3$  του Πίνακα 8 των δεδομένων. Οι υπόλοιπες 23 στήλες προκύπτουν με προφανείς πολλαπλασιασμούς των στηλών 2, 3 και 4 (των  $X_1, X_2$  και  $X_3$ ) ώστε να προκύψουν τα τελευταία 23 στοιχεία



ωστόσο εντοπίζονται αλληλεπιδράσεις του με τους άλλους δυο παράγοντες. Ακόμα, αξίζει να σχολιασθεί πως αν θεωρηθούν σταθεροί οι παράγοντες και οι αλληλεπιδράσεις τους τότε η μεταβλητή απόκρισης λαμβάνει την τιμή 4.9, για κάθε μοναδιαία μεταβολή στον τύπο των ραφιών, αν παραμένουν σταθεροί οι υπόλοιποι παράγοντες καθώς και οι αλληλεπιδράσεις θα παρατηρείται μείωση κατά 0.0075 στην τιμή της μεταβλητής απόκρισης και με κάθε μεταβολή στον εργαζόμενο, παραμένοντας σταθεροί οι υπόλοιποι παράγοντες και οι αλληλεπιδράσεις αυτών θα υπάρχει αύξηση κατά 0.87 στην τιμή της μεταβλητής απόκρισης.

Για σκοπούς σύγκρισης ακολουθεί το Output (σε MINITAB) της ανάλυσης παλινδρόμησης για την Εφαρμογή 2, όπου έχουν συμπεριληφθεί 6 παράγοντες, ήτοι οι  $X_1, X_2, X_3$  και τα τετράγωνά τους ( $X_{11}, X_{22}, X_{33}$ ) καθώς και όλες οι δυνατές αλληλεπιδράσεις των 6 προηγούμενων παραγόντων. Συνολικά ζητήθηκε από το μοντέλο η μελέτη 54 κύριων επιδράσεων και αλληλεπιδράσεων. Το μοντέλο επέστρεψε με την αξιολόγηση μόνο 27 (συμπεριλαμβάνεται και ο σταθερός όρος) των ίδιων που έχουν ήδη εντοπισθεί με την τεχνική των βάσεων Gröbner. Τα αποτελέσματα ταυτίζονται με αυτά που βρέθηκαν νωρίτερα, όπως φαίνεται και στο Output 1 (με προφανείς αποκλίσεις λόγω στρογγυλοποιήσεων).

### Regression Analysis: Y versus X1; X2; X3; X11; X22; X33

The following terms cannot be estimated and were removed:  
 X1\*X11; X2\*X22; X3\*X33; X1\*X2\*X11; X1\*X2\*X22; X1\*X3\*X11; X1\*X3\*X33; X1\*X11\*X22; X1\*X11\*X33; X2\*X3\*X22; X2\*X3\*X33;  
 X2\*X11\*X22; X2\*X22\*X33; X3\*X11\*X33; X3\*X22\*X33; X1\*X2\*X3\*X11; X1\*X2\*X3\*X22; X1\*X2\*X3\*X33; X1\*X2\*X11\*X22;  
 X1\*X2\*X11\*X33; X1\*X2\*X22\*X33; X1\*X3\*X11\*X22; X1\*X3\*X11\*X33; X1\*X3\*X22\*X33; X1\*X11\*X22\*X33; X2\*X3\*X11\*X22;  
 X2\*X3\*X11\*X33; X2\*X3\*X22\*X33; X2\*X11\*X22\*X33; X3\*X11\*X22\*X33; X1\*X2\*X3\*X11\*X22; X1\*X2\*X3\*X11\*X33;  
 X1\*X2\*X3\*X22\*X33; X1\*X2\*X11\*X22\*X33; X1\*X3\*X11\*X22\*X33; X2\*X3\*X11\*X22\*X33; X1\*X2\*X3\*X11\*X22\*X33

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	26	24,1651	0,92943	7,51	0,000
X1	1	0,0000	0,00000	0,00	1,000
X2	1	0,0002	0,00022	0,00	0,966
X3	1	3,0450	3,04502	24,60	0,000
X11	1	0,0048	0,00480	0,04	0,845
X22	1	0,4602	0,46021	3,72	0,064
X33	1	0,2730	0,27301	2,21	0,149
X1*X2	1	0,1300	0,13005	1,05	0,314
X1*X3	1	0,4232	0,42320	3,42	0,075
X1*X22	1	0,1094	0,10935	0,88	0,356
X1*X33	1	0,0988	0,09882	0,80	0,379
X2*X3	1	0,2812	0,28125	2,27	0,143
X2*X11	1	0,0068	0,00682	0,07	0,792
X2*X33	1	0,0001	0,00015	0,00	0,972
X3*X11	1	0,3851	0,38507	3,11	0,089
X3*X22	1	0,3750	0,37500	3,03	0,093
X11*X22	1	0,0060	0,00605	0,05	0,827
X11*X33	1	0,0709	0,07094	0,57	0,456
X22*X33	1	0,0722	0,07220	0,58	0,452
X1*X2*X3	1	0,0354	0,03541	0,21	0,648
X1*X2*X33	1	0,1740	0,17400	1,41	0,246
X1*X3*X22	1	0,1441	0,14410	1,16	0,290
X1*X22*X33	1	0,0171	0,01712	0,14	0,713
X2*X3*X11	1	0,0204	0,02042	0,16	0,688
X2*X11*X33	1	0,0518	0,05176	0,42	0,523
X3*X11*X22	1	0,6019	0,60192	4,86	0,036
X11*X22*X33	1	0,0658	0,06576	0,53	0,472
Error	27	3,3416	0,12376		
Total	53	27,5067			

#### Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0,351797	87,85%	76,15%	51,41%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	4,900	0,249	19,70	0,000	
X1	0,000	0,176	0,00	1,000	9,00
X2	-0,007	0,176	-0,04	0,966	9,00
X3	0,873	0,176	4,96	0,000	9,00
X11	0,060	0,305	0,20	0,845	9,00
X22	-0,588	0,305	-1,93	0,064	9,00
X33	0,452	0,305	1,49	0,149	9,00
X1*X2	0,127	0,124	1,03	0,314	3,00
X1*X3	-0,230	0,124	-1,85	0,075	3,00
X1*X22	0,202	0,215	0,94	0,356	9,00
X1*X33	-0,192	0,215	-0,89	0,379	9,00
X2*X3	-0,187	0,124	-1,51	0,143	3,00
X2*X11	0,057	0,215	0,27	0,792	9,00
X2*X33	0,007	0,215	0,03	0,972	9,00
X3*X11	-0,380	0,215	-1,76	0,089	9,00
X3*X22	-0,375	0,215	-1,74	0,093	9,00
X11*X22	-0,082	0,373	-0,22	0,827	15,00
X11*X33	-0,282	0,373	-0,76	0,456	15,00
X22*X33	-0,285	0,373	-0,76	0,452	15,00
X1*X2*X3	0,0406	0,0879	0,46	0,648	1,00
X1*X2*X33	-0,181	0,152	-1,19	0,246	3,00
X1*X3*X22	0,164	0,152	1,08	0,290	3,00
X1*X22*X33	0,098	0,264	0,37	0,713	9,00
X2*X3*X11	-0,062	0,152	-0,41	0,688	3,00
X2*X11*X33	-0,171	0,264	-0,65	0,523	9,00
X3*X11*X22	0,582	0,264	2,21	0,036	9,00
X11*X22*X33	0,333	0,457	0,73	0,472	19,00

#### Regression Equation

$$Y = 4,900 + 0,000 X1 - 0,007 X2 + 0,873 X3 + 0,060 X11 - 0,588 X22 + 0,452 X33 + 0,127 X1*X2 - 0,230 X1*X3 + 0,202 X1*X22 - 0,192 X1*X33 - 0,187 X2*X3 + 0,057 X2*X11 + 0,007 X2*X33 - 0,380 X3*X11 - 0,375 X3*X22 - 0,082 X11*X22 - 0,282 X11*X33 - 0,285 X22*X33 + 0,0406 X1*X2*X3 - 0,181 X1*X2*X33 + 0,164 X1*X3*X22 + 0,098 X1*X22*X33 - 0,062 X2*X3*X11 - 0,171 X2*X11*X33 + 0,582 X3*X11*X22 + 0,333 X11*X22*X33$$

#### Fits and Diagnostics for Unusual Observations

Obs	Y	Fit	Resid	Std Resid
24	6,030	5,200	0,830	3,34 R
51	4,370	5,200	-0,830	-3,34 R

R Large residual

Για σκοπούς αντιπαραβολής μελετήθηκε επιπρόσθετα και το κλασικό μοντέλο των τριών παραγόντων ( $X_1$ ,  $X_2$  και  $X_3$ ) και όλων των αλληλεπιδράσεων τους. Τα αποτελέσματα δίνουν τους 8 εκτιμητές (συμπεριλαμβάνεται και ο σταθερός όρος), ενώ δείχνουν ότι υπάρχει lack - of - fit με 19 βαθμούς ελευθερίας, που σημαίνει ότι υπάρχει η δυνατότητα εκτίμησης 19 επιπλέον παραγόντων (που γνωρίζουμε από τις βάσεις Gröbner ότι δεν είναι οποιοδήποτε, αλλά είναι εκείνα τα στοιχεία της βάσης που δεν συμπεριλήφθησαν στο κλασικό μοντέλο που μελετήθηκε). Όπως είναι προφανές, στις περιπτώσεις αυτές το συνολικό σφάλμα αυξάνεται αφού στο <<Pure Error>> που έτσι και αλλιώς δεν θα μπορούσε ποτέ να εξαληφθεί, προστίθεται και το σφάλμα λόγω <<lack - of - fit>> το οποίο όμως θα μπορούσε να αποφευχθεί αν το πλήρες μοντέλο όπως υπεδείχθη από την τεχνική των βάσεων Gröbner είχε προτιμηθεί. Τα παραπάνω επιβεβαιώνονται από το Output 2.

#### Regression Analysis: Y versus X1; X2; X3

##### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	7	15,9678	2,2811	9,09	0,000
X1	1	0,0910	0,0910	0,36	0,550
X2	1	0,0576	0,0576	0,23	0,634
X3	1	14,1878	14,1878	56,56	0,000
X1*X2	1	0,0012	0,0012	0,00	0,945
X1*X3	1	0,3480	0,3480	1,39	0,245
X2*X3	1	1,2558	1,2558	5,01	0,030
X1*X2*X3	1	0,0264	0,0264	0,11	0,747
Error	46	11,5389	0,2508		
Lack-of-Fit	19	8,1973	0,4314	3,49	0,002
Pure Error	27	3,3416	0,1238		
Total	53	27,5067			

##### Model Summary

S	R-sq	R-sq(Adj)	R-sq(Pred)
0,500844	58,05%	51,67%	45,29%

##### Coefficients

Term	Coeff	SE Coef	T-Value	P-Value	VIF
Constant	4,6598	0,0662	68,37	0,000	
X1	0,0503	0,0835	0,60	0,550	1,00
X2	-0,0400	0,0835	-0,48	0,634	1,00
X3	0,6278	0,0835	7,52	0,000	1,00
X1*X2	0,007	0,102	0,07	0,945	1,00
X1*X3	-0,120	0,102	-1,18	0,245	1,00
X2*X3	-0,229	0,102	-2,24	0,030	1,00
X1*X2*X3	0,041	0,125	0,32	0,747	1,00

##### Regression Equation

$$Y = 4,6598 + 0,0503 X1 - 0,0400 X2 + 0,6278 X3 + 0,007 X1 * X2 - 0,120 X1 * X3 - 0,229 X2 * X3 + 0,041 X1 * X2 * X3$$

## Output 2



Για λόγους πληρότητας δίδεται κατωτέρω ένα  $2^3$  πλήρες παραγοντικό πείραμα που περιλαμβάνει μόνο τους 2 πρώτους παράγοντες ( $X_1$  και  $X_2$ ) από τους παράγοντες της ενότητας αυτής (βλέπε Πίνακα 8). Όπως αναμένεται με χρήση των βάσεων Gröbner, το σύνολο  $Est_\tau(D) = \{1, x_1, x_2, x_1^2, x_2^2, x_1x_2, x_1x_2^2, x_1^2x_2, x_1^2x_2^2\}$  με τον αντίστροφο του πίνακα σχεδιασμού  $X$ , όπου

$$X = \begin{pmatrix} 1 & -1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 \\ 1 & 0 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & -1 & 1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & -1 & 1 & 1 & 1 & -1 & -1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

να δίνεται ως εξής:

$$X^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & -1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & -1 & 1/2 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & -1 & 0 & 0 & 1/2 & 0 \\ 1/4 & 0 & -1/4 & 0 & 0 & 0 & -1/4 & 0 & 1/4 \\ -1/4 & 0 & 1/4 & 1/2 & 0 & -1/2 & -1/4 & 0 & 1/4 \\ -1/4 & 1/2 & -1/4 & 0 & 0 & 0 & 1/4 & -1/2 & 1/4 \\ 1/4 & -1/2 & 1/4 & -1/2 & 1 & -1/2 & 1/4 & -1/2 & 1/4 \end{pmatrix}$$

Το τελικό (πλήρες) μοντέλο του παραπάνω πειραματικού σχεδιασμού δίνεται στο Output 3. Ας σημειωθεί ότι στο μοντέλο παλινδρόμησης ζητήθηκε η μελέτη τεσσάρων παραγόντων ( $X_1, X_2, X_1^2, X_2^2$ ) καθώς και όλες οι αλληλεπιδράσεις τους ανά δυο. Από τους συνολικά 15 όρους (μαζί με τον σταθερό όρο) το μοντέλο επέστρεψε τους 9 όρους, που δεν είναι άλλοι από αυτούς που υπέδειξε η βάση Gröbner

#### Regression Analysis: C21 versus Z1; Z2; Z11; Z22

The following terms cannot be estimated and were removed:  
Z1\*Z1; Z2\*Z2; Z11\*Z11; Z22\*Z22; Z1\*Z11; Z2\*Z22

#### Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	9	2,76416	0,340022	4,45	0,020
Z1	1	0,00562	0,005625	0,07	0,795
Z2	1	0,14062	0,140625	1,80	0,213
Z11	1	0,09307	0,093075	0,42	0,532
Z22	1	0,33001	0,330008	4,22	0,070
Z1*Z2	1	0,07031	0,070312	0,90	0,368
Z1*Z22	1	0,04950	0,049504	0,63	0,447
Z2*Z11	1	0,00700	0,007004	0,09	0,772
Z11*Z22	1	0,09753	0,097535	1,25	0,293
Error	9	0,70445	0,078272		
Total	17	3,46863			

#### Model Summary

S	R-sq	R-sq(Adj)	R-sq(Pred)
0,279772	79,81%	61,86%	19,23%

#### Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VF
Constant	4,480	0,199	22,55	0,000	
Z1	0,037	0,140	0,27	0,795	3,00
Z2	0,187	0,140	1,34	0,213	3,00
Z11	0,157	0,242	0,65	0,532	3,00
Z22	-0,497	0,242	-2,05	0,070	3,00
Z1*Z2	-0,0938	0,0989	-0,95	0,368	1,00
Z1*Z22	0,136	0,171	0,80	0,447	3,00
Z2*Z11	-0,051	0,171	-0,30	0,772	3,00
Z11*Z22	-0,331	0,297	-1,12	0,293	5,00

#### Regression Equation

$$C21 = 4,480 + 0,037 Z1 + 0,187 Z2 + 0,157 Z11 - 0,497 Z22 - 0,0938 Z1*Z2 + 0,136 Z1*Z22 - 0,051 Z2*Z11 - 0,331 Z11*Z22$$

### Output 3

Οποιαδήποτε απόπειρα περιορισμού του μοντέλου οδηγεί σε lack-of-fit (Output 4), όπως συνέβη και στο  $3^3$  πείραμα νωρίτερα.

Regression Analysis: C21 versus Z1; Z2

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	5	2.63013	0.52603	7.35	0.002
Z1	1	0.19763	0.19763	2.76	0.122
Z2	1	0.28213	0.28213	3.94	0.070
Z1*Z1	1	0.01604	0.01604	0.22	0.644
Z2*Z2	1	2.06401	2.06401	28.85	0.000
Z1*Z2	1	0.07031	0.07031	0.98	0.341
Error	12	0.85849	0.07154		
Lack-of-Fit	3	0.15404	0.05135	0.66	0.599
Pure Error	9	0.70445	0.07827		
Total	17	3.48863			

Model Summary

S	R-sq	R-sq(Adj)	R-sq(Pred)
0.267472	75.39%	65.14%	45.11%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	4.627	0.141	32.82	0.000	
Z1	0.1283	0.0772	1.66	0.122	1.00
Z2	0.1533	0.0772	1.99	0.070	1.00
Z1*Z1	-0.063	0.134	-0.47	0.644	1.00
Z2*Z2	-0.718	0.134	-5.37	0.000	1.00
Z1*Z2	-0.0938	0.0946	-0.99	0.341	1.00

Regression Equation

$$C21 = 4.627 + 0.1283 Z1 + 0.1533 Z2 - 0.063 Z1^2 - 0.718 Z2^2 - 0.0938 Z1^*Z2$$

## Output 4

Σε όλες τις περιπτώσεις παρατηρούμε ότι η ανάλυση παλινδρόμησης προχωράει ένα βήμα παραπάνω από την μέθοδο των βάσεων Gröbner καθώς δίνει και το βαθμό σημαντικότητας των εμπλεκόμενων παραγόντων, αφού αξιοποιώντας συμπληρωματικά τα τυπικά σφάλματα των εκτιμητών (SE Coeff) επιτρέπει την διατήρηση των παραγόντων με p-value μικρότερο του συνήθους 5% και την απομάκρυνση των παραγόντων με p-value μεγαλύτερο ή ίσο του 5% (άλλες συνήθεις τιμές είναι το 1% και το 10%).

## Συμπεράσματα

Συνοπτικά, αναφέρουμε ότι η βάση Gröbner καταλήγει να περιλαμβάνει τόσα στοιχεία όσες είναι όλες οι δυνατές θεραπείες του πειραματικού σχεδιασμού (4 για το  $2^2$ , 9 για το  $2^3$  και 27 για το  $3^3$ ). Ο αριθμός των στοιχείων του  $Est_\tau$  που προκύπτει από την βάση Gröbner καθορίζει τον μέγιστο αριθμό παραγόντων που μπορούν να αναγνωρισθούν ή εναλλακτικά των εκτιμητών που μπορούν να προσδιορισθούν. Έτσι, στο Παράδειγμα την ενότητας 3.1 εκτιμήθηκαν οι τέσσερις παράμετροι ( τα  $\theta_0, \theta_1, \theta_2, \theta_3$ ) του μοντέλου

$$y = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_1 x_2.$$

Ακόμα, στο παράδειγμα της Ενότητας 3.2 έχουν εκτιμηθεί 27 συνολικά ποσότητες που αντιστοιχούν στα 27 στοιχεία του  $Est_\tau$  της αντίστοιχης βάσης Gröbner. Με άλλα λόγια η διαδικασία των βάσεων Gröbner εξαντλεί όλες τις περιπτώσεις και "επιστρέφει" πλήρη μοντέλα, δηλαδή μοντέλα χωρίς "lack-of-fit", αλλά με τη μέγιστη δυνατή πολυπλοκότητα. Εναλλακτικά, θα μπορούσε να ειπωθεί ότι η τεχνική των βάσεων Gröbner επιχειρεί να αξιοποιήσει (και τα καταφέρνει) όλη την υπάρχουσα (στα δεδομένα) πληροφορία για να κατασκευάσει το πληρέστερο μοντέλο που περιγράφει ιδανικότερα το υπό μελέτη φαινόμενο. Στην περίπτωση παραγόντων με δυο επίπεδα το μοντέλο που προκύπτει

είναι αυτό που συνήθως μελετάται στην πράξη, αλλά στην περίπτωση παραγόντων με τρία και άνω επίπεδα συνηθίζεται να μελετώνται στην πράξη πολύ απλούστερα μοντέλα αυτών που προκύπτουν μέσω των βάσεων Gröbner. Αν και ένα πληρέστερο μοντέλο δεν είναι συνήθως εύχρηστο, ιδίως όταν ο σκοπός της μελέτης είναι η πρόβλεψη (τότε ένα απλό μοντέλο προτιμάται), στις περιπτώσεις που σκοπός της μελέτης είναι η περιγραφή όσο το δυνατόν ακριβέστερα του υπό εξέταση φαινομένου, η τεχνική των βάσεων Gröbner αποτελεί την ιδανική λύση αφού εντοπίζει το πλήρες μοντέλο, όσο πολύπλοκο και αν είναι. Σημειώνεται ότι μοντέλα όπως αυτό που προέκυψε στην Ενότητα 3.2 δεν είναι εύκολο να φανταστεί κανείς και να τα μελετήσει εκτός και αν έχει οδηγό τις βάσεις Gröbner. Ωστόσο, παρατηρούμε και μια σημαντική αδυναμία της μεθόδου των βάσεων Gröbner σε σχέση με τους πειραματικούς σχεδιασμούς. Αυτή η αδυναμία έγκειται στο ότι δεν προσφέρει κάποιο κριτήριο για τη διερεύνηση της στατιστικής σημαντικότητας των παραμέτρων του μοντέλου και συνεπώς δεν προσφέρει την ζητούμενη απλοποίηση του μοντέλου, κυρίως σε θέματα πρόβλεψης.

# Παράρτημα

## Ιδιότητες Συνάρτησης Κατανομής

a) Έστω  $x, y \in \mathbb{R}$  με  $x < y$ . Θα πρέπει να ισχύει ότι  $F_X(x) \leq F_X(y)$ . Όμως,  $\{\omega \in \Omega | X(\omega) \leq x\} \subseteq \{\omega \in \Omega | X(\omega) \leq y\}$ . Οπότε,  $P\{\omega \in \Omega | X(\omega) \leq x\} \leq P\{\omega \in \Omega | X(\omega) \leq y\} \Rightarrow P\{X \leq x\} \leq P\{X \leq y\} \Rightarrow F_X(x) \leq F_X(y)$ . Άρα η συνάρτηση κατανομής είναι αύξουσα.

b) Έστω μια αύξουσα ακολουθία  $\{x_n\}_{n \geq 1}$ , με  $\lim_{n \rightarrow \infty} x_n = \infty$ . Αρκεί να δείξουμε ότι  $\lim_{x \rightarrow \infty} F_X(x) = 1$ . Αυτό προκύπτει άμεσα, αφού η ακολουθία ενδεχομένων  $\{A_n\}_{n \geq 1}$ , όπου  $A_n = \{X \leq x_n\}$  είναι αύξουσα με  $\lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} \{X \leq x_n\} = \{X < \infty\}$ . Επομένως,  $\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P(A_n) = P(\lim_{n \rightarrow \infty} A_n) = P(X < \infty) = P(\Omega) = 1$ .

c) Όμοια με (b).

d) Για κάθε φθίνουσα ακολουθία πραγματικών αριθμών  $\{x_n\}_{n \geq 1}$ , με  $\lim_{n \rightarrow \infty} x_n = x$  ισχύει ότι  $\lim_{n \rightarrow \infty} F_X(x_n) = F_X(x)$ . Αφού η ακολουθία  $\{x_n\}_{n \geq 1}$  είναι φθίνουσα, η αντίστοιχη ακολουθία ενδεχομένων  $\{A_n\}_{n \geq 1}$  με  $A_n = \{X \leq x_n\}$  θα είναι φθίνουσα. Επιπλέον, ισχύει ότι  $\lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} \{X \leq x_n\} = \{X \leq x\}$ , επομένως  $\lim_{n \rightarrow \infty} F_X(x_n) = \lim_{n \rightarrow \infty} P(A_n) = P(\lim_{n \rightarrow \infty} A_n) = P(X \leq x) = F_X(x)$ .

## Μερικές Σημαντικές Κατανομές

### Κανονική Κατανομή

Η κανονική κατανομή είναι μια από τις σημαντικότερες κατανομές. Αρχικά μελετήθηκε από τον De Moivre (1667 - 1754) και τον Laplace (1749 - 1827). Ωστόσο, ο Gauss (1777 - 1855) παρατήρησε πως οι κατανομές των σφαλμάτων μπορούν να προσεγγιστούν από μια συνεχή καμπύλη, την κανονική καμπύλη των σφαλμάτων.

**Ορισμός 3.2.1.** Η συνεχής κατανομή με σ.π.π.

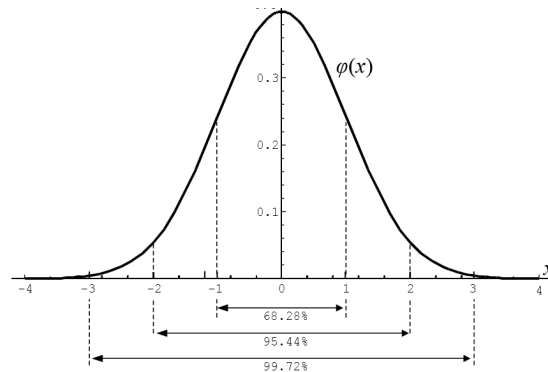
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, x \in \mathbb{R}, \mu \in \mathbb{R}, \sigma > 0,$$

ονομάζεται κανονική κατανομή με παραμέτρους  $\mu$ ,  $\sigma^2$  και συμβολίζεται με  $N(\mu, \sigma^2)$ .

Μια ιδιαίτερη κατηγορία κανονικής κατανομής αποτελεί η τυπική κανονική κατανομή η οποία προκύπτει από την κανονική κατανομή για τιμές  $\mu = 0, \sigma^2 = 1$ . Η σ.π.π. τότε γίνεται

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}.$$

Η παραπάνω συνάρτηση σχηματικά είναι η εξής:



Σχήμα 2: Τυπική Κανονική Κατανομή (Μπούτσικας, 2003)

Έχουμε δηλαδή ότι αν  $X \sim N(0, 1)$  τότε:

$$P(-3 \leq X \leq 3) \simeq 0.9972$$

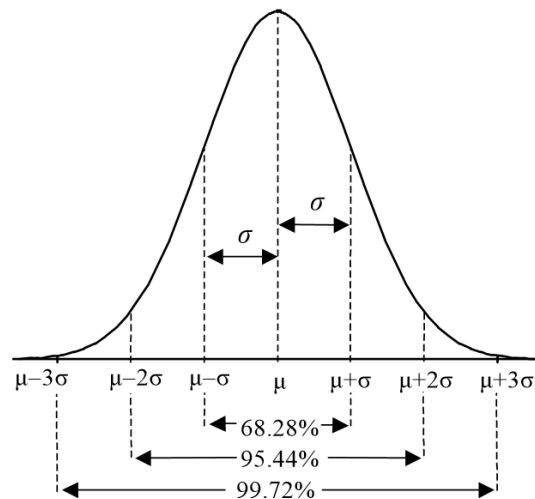
$$P(-2 \leq X \leq 2) \simeq 0.9544$$

$$P(-1 \leq X \leq 1) \simeq 0.6828$$

Παρατηρούμε λοιπόν πως πρόκειται για συμμετρική κατανομή, όπου μάλιστα ισχύει ότι  $\Phi(-x) = 1 - \Phi(x)$ ,  $x \in \mathbb{R}$ . Η χρησιμότητα της τυπικής κανονικής κατανομής έγκειται στο γεγονός ότι με κατάλληλη τροποποίηση μιας κανονικής κατανομής μπορούμε να αναχθούμε σε τυπική κανονική κατανομή, η οποία και είναι διαχειρίσιμη. Αυτό συμβάνει χάρη στην παρακάτω πρόταση.

**Πρόταση 3.2.1.** Αν η τ.μ.  $X \sim N(\mu, \sigma^2)$ , τότε η τ.μ.  $Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$ .

Συνεπώς, με χρήση της τυπικής κανονικής κατανομής έχουμε για την κανονική κατανομή



Σχήμα 3: Κανονική Κατανομή (Μπούτσικας, 2003)



Με βάση λοιπόν το Σχήμα 3 έχουμε ότι μια τ.μ.  $X \sim N(\mu, \sigma^2)$  λαμβάνει τιμές μεταξύ των τιμών  $\mu - 3\sigma$  και  $\mu + 3\sigma$  με πιθανότητα σχεδόν 1, με πιθανότητα 0.95 λαμβάνει τιμές μεταξύ των  $\mu - 2\sigma$  και  $\mu + 2\sigma$  και με πιθανότητα 0.65 λαμβάνει τιμές μεταξύ των  $\mu - \sigma$  και  $\mu + \sigma$ .

### $\chi^2$ Κατανομή

**Ορισμός 3.2.2.** Έστω  $Z_1, \dots, Z_n$  ανεξάρτητες και ισόνομες τ.μ. με  $Z_i \sim N(0, 1), \forall i = 1, \dots, n$ . Η κατανομή της τ.μ.

$$X = \sum_{i=1}^n Z_i^2$$

ονομάζεται  $\chi^2$  - τετράγωνο κατανομή με  $n$  βαθμούς ελευθερίας. Συμβολισμός:  $\chi_n^2$

### t Κατανομή

**Ορισμός 3.2.3.** Έστω  $Z \sim N(0, 1), Y \sim \chi_n^2$  ανεξάρτητες και ισόνομες τ.μ., τότε η κατανομή της τ.μ.

$$X = \frac{Z}{\sqrt{\frac{Y}{n}}}$$

ονομάζεται t κατανομή με  $n$  βαθμούς ελευθερίας. Συμβολισμός:  $t_n$ .

### F Κατανομή

**Ορισμός 3.2.4.** Έστω  $Y_1 \sim \chi_{n_1}^2, Y_2 \sim \chi_{n_2}^2$  ανεξάρτητες και ισόνομες τ.μ, τότε η κατανομή της τ.μ.

$$X = \frac{\frac{Y_1}{n_1}}{\frac{Y_2}{n_2}}$$

ονομάζεται F με  $n_1$  και  $n_2$  βαθμούς ελευθερίας. Συμβολισμός:  $F_{n_1, n_2}$ .

## Μέθοδος Ελαχίστων Τετραγώνων

Οι εκτιμητές των παραμέτρων  $\beta_0, \beta_1$  υπολογίζονται με την παρακάτω διαδικασία. Θεωρούμε την ποσότητα

$$Q(\beta_0, \beta_1) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 * X_i)^2$$

Παραγωγίζουμε την ποσότητα αυτήν ως προς  $\beta_0, \beta_1$  προκειμένου να την ελαχιστοποιήσουμε. Έχουμε:

$$\frac{\partial(\beta_0, \beta_1)}{\partial\beta_0} = \sum_{i=1}^n 2(y_i - (\beta_0 + \beta_1 x_i))(-1) = -2\left(\sum_{i=1}^n y_i - n\beta_0 - \beta_1 \sum_{i=1}^n x_i\right)$$

και

$$\frac{\partial(\beta_0, \beta_1)}{\partial\beta_1} = \sum_{i=1}^n 2(y_i - (\beta_0 + \beta_1 x_i))(-x_i) = -2\left(\sum_{i=1}^n x_i y_i - \beta_0 \sum_{i=1}^n x_i - \beta_1 \sum_{i=1}^n x_i^2\right)$$

Θέτουμε τις μερικές παραγώγους ίσες με το 0 και προκύπτει το σύστημα των κανονικών εξισώσεων,

$$\sum_{i=1}^n y_i - n\beta_0 - \beta_1 \sum_{i=1}^n x_i = 0 \Leftrightarrow n\beta_0 + \left(\sum_{i=1}^n x_i\right)\beta_1 = \sum_{i=1}^n y_i$$

$$\sum_{i=1}^n x_i y_i - \beta_0 \sum_{i=1}^n x_i - \beta_1 \sum_{i=1}^n x_i^2 = 0 \Leftrightarrow \left(\sum_{i=1}^n x_i\right)\beta_0 + \left(\sum_{i=1}^n x_i^2\right)\beta_1 = \sum_{i=1}^n x_i y_i.$$

Λύνουμε ως προς τα  $\beta_0, \beta_1$  και έχουμε

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

και

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 * \bar{X}.$$

### **Αλγόριθμος Διάρθρωσης Πολυωνύμων:**

**Είσοδος:**  $f, f_1, \dots, f_s \in K[x_1, \dots, x_n]$   $f_i \neq 0, \forall i : 1 \leq i \leq s$ .

**Έξοδος:**  $u_1, \dots, u_s$  τέτοια ώστε  $f = u_1 f_1 + \dots + u_s f_s + r$ , όπου  $r$  ανάγωγο μόνιο  $\{f_1, \dots, f_s\}$  και  $\max(\text{lm}(u_1)\text{lm}(f_1), \dots, \text{lm}(u_s)\text{lm}(f_s), \text{lm}(r)) = \text{lm}(f)$ .

**Αρχή:**  $u_1 = 0, \dots, u_s = 0, r = 0, h = f$ . **Όσο**  $h \neq 0$  **επανέλαβε:**

Αν υπάρχει  $i$  τέτοιο ώστε  $\text{lm}(f_i) | \text{lm}(h)$ , τότε διάλεξε το μικρότερο  $i$  με  $\text{lm}(f_i) | \text{lm}(h)$

και

$$u_i = u_i + \frac{\text{lt}(h)}{\text{lt}(f_i)}, \quad h = h - \frac{\text{lt}(h)}{\text{lt}(f_i)} f_i.$$

**Αλλιώς,**  $r = r + \text{lt}(h), \quad h = h - \text{lt}(h)$ .

# Βιβλιογραφία

## Ελληνική Βιβλιογραφία

- [1] Αμπατζή, Ε. (2011). Βάσεις Gröbner και Πολύτοπο Κατάστασης, Τμήμα Μαθηματικών, Πανεπιστήμιο Ιωαννίνων.
- [2] Δημητρακοπούλου, Θ. (2017). Μοντέλα Παλινδρόμησης, Τμήμα Στατιστικής και Αναλογιστικών - Χρηματοοικονομικών Μαθηματικών, Πανεπιστήμιο Αιγαίου.
- [3] Δημητράκος, Θ. (2017). Πιθανότητες, Τμήμα Μαθηματικών, Πανεπιστήμιο Αιγαίου.
- [4] Ηλιόπουλος, Γ. (2013). Βασικές Μέθοδοι Εκτίμησης Παραμέτρων με σημείο και με διάστημα. *Εκδόσεις ΑΘ. ΣΤΑΜΟΥΛΗΣ*, Β έκδοση. ISBN: 978-960-351-908-9.
- [5] Θωμά, Α., Τατάκης, Χ. (2018). Σημειώσεις στις βάσεις Gröbner, Τμήμα Μαθηματικών, Πανεπιστήμιο Ιωαννίνων.

- [6] Κατσιλέρος, Α. (2019). Υπολογιστική Στατιστική, Τμήμα Στατιστικής και Αναλογιστικών - Χρηματοοικονομικών Μαθηματικών, Πανεπιστήμιο Αιγαίου.
- [7] Κούτρας, Μ. (2011). Ανάλυση Παλινδρόμησης, Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης, Πανεπιστήμιο Πειραιώς.
- [8] Μπούτσικας, Μ. (2003). Σημειώσεις Στατιστικής ΙΙΙ, Τμήμα Οικονομικής Επιστήμης, Πανεπιστήμιο Πειραιώς.
- [9] Χατζησπύρος, Σ. (2013). Εισαγωγή στην Στατιστική, Τμήμα Στατιστικής και Αναλογιστικών - Χρηματοοικονομικών Μαθηματικών, Πανεπιστήμιο Αιγαίου.

### **Ξενόγλωσση Βιβλιογραφία**

- [10] Buchberger, B. (2005). Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *Journal of Symbolic Computation*.
- [11] Cohen A. M., Bucchianico A. Di, Riccomagno, E. (2010) Replications with Gröbner Bases. *Springer Link*. ISSN: 1389-2355.
- [12] Cox, D., Little, J., O'Shea, D. (2007). Ideals, Varieties, and Algorithms, Third edition *Springer*. ISBN-10: 0-387-35650-9.
- [13] Giglio, B., Riccomagno, E., R., Wynn, H-P. (2000). Gröbner basis strategies in regression. *Journal of Applied Statistics*, 27:7, 923-938.

- [14] Montgomery, D. (2013). Design and analysis of experiments, *John Wiley & Sons, Inc.*, Eighth edition. ISBN 978-1-118-14692-7.
- [15] Pistone, G., Riccomagno, E., Wynn, H-P.(2001).Algebraic Statistics Computational Commutative Algebra in Statistics, *CHAPMAN & HALL/CRC*. ISBN 1-58488-204-2.
- [16] Pistone, G., Wynn, H-P.(1996) Generalized confounding with Gröbner bases. *Biometrika*.
- [17] Sullivant, S. (2010). Algebraic Statistics. North Carolina State University.