

ΕΙΣΑΓΩΓΗ ΣΤΗΝ

ΠΟΛΥΜΕΤΑΒΛΗΤΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

(Μέρος Α Μαθήματος «Πολυμεταβλητή Ανάλυση»)

1. ΕΙΣΑΓΩΓΗ

Σε αρκετές εφαρμογές (π.χ. σε βιολογικές, οικονομικές ή κοινωνικές επιστήμες) τα δεδομένα που συλλέγονται αφορούν περισσότερες από μια μεταβλητές. Για παράδειγμα, ο παρακάτω πίνακας δίνει τις μετρήσεις τριών μεταβλητών

X_1 : *Περίμετρος Κοιλίας*, X_2 : *Βάρος*, X_3 : *Ύψος*

σε 36 τυχαία επιλεγμένους άνδρες 20-29 ετών (36 στοιχεία – items – cases).

άτομο	Περίμετρος Κοιλίας (cm)	Βάρος (Kgr)	Ύψος (cm)
1	83	78,5	183,5
2	87,9	69,8	168,2
3	85,2	69,9	172
4	88,6	89,9	186,6
5	77,1	72,4	183,5
6	85,3	85,3	196,8
7	100	83,5	180,9
8	94,4	95,3	189,8
9	81,9	70,7	179,7
10	99,1	94,5	184,7
11	88,5	79,8	184,1
12	82,5	86,6	187,9
13	100,5	93,6	177,1
14	76,5	65,2	184,1
15	86,4	83,8	183,5
16	90,7	82,1	177,1
17	83,6	84,4	189,2
18	106,8	101,1	178,4

άτομο	Περίμετρος Κοιλίας (cm)	Βάρος (Kgr)	Ύψος (cm)
19	77,6	69	175,2
20	102,9	109,6	189,2
21	90,9	97,9	193
22	79,7	72,2	181,6
23	73,9	60,4	164,4
24	79,1	76,2	180,9
25	72,8	66,2	183,5
26	88,2	71,1	170,8
27	100,1	90,8	186,6
28	89,6	83,3	172
29	95,9	81,1	172,7
30	98,8	90,9	177,1
31	76,3	68,6	172
32	83,5	77,7	191,1
33	105	93,3	175,2
34	90,8	82,7	183,5
35	83,5	72,9	175,2
36	84,5	72,6	180,9

Μας ενδιαφέρει να αναλύσουμε και να μελετήσουμε παρόμοια πολυμεταβλητά δεδομένα. Συνήθως μας ενδιαφέρει:

- (1) Να *μειώσουμε* το πλήθος των μεταβλητών χωρίς να χαθεί πολύτιμη πληροφορία (αν π.χ. έχουμε 50 μεταβλητές μας ενδιαφέρει να κρατήσουμε τις κύριες συνιστώσες)
- (2) Να *ομαδοποιήσουμε* ή να *κατατάξουμε* τα δεδομένα σε ομοειδή υποσύνολα.
- (3) Να *διερευνήσουμε την εξάρτηση* μεταξύ των μεταβλητών.
- (4) Να κατασκευάσουμε υποδείγματα που θα μας βοηθήσουν στην *πρόβλεψη* ποσοτήτων
- (5) Να πραγματοποιήσουμε *ελέγχους υποθέσεων* που αφορούν παραμέτρους των δεδομένων

Είναι προφανές ότι η στατιστική ανάλυση πολυμεταβλητών δεδομένων ότι είναι υπολογιστικά επίπονη. Για το λόγο αυτό, η ανάλυση αυτή γίνεται με την χρήση H/Y μέσω κατάλληλου λογισμικού.

Οι τεχνικές πολυμεταβλητών δεδομένων εφαρμόζονται π.χ. στην *βιοστατιστική* και στην *φαρμακολογία*, στην *κοινωνιολογία* (κοινωνική στατιστική), στα *οικονομικά – χρηματοοικονομικά*, στην *εκπαίδευση*, στην *βιολογία*, στην *μελέτη του περιβάλλοντος*, στην *μετεωρολογία*, στην *γεωλογία*, στην *ψυχολογία* κ.α.

Στο πρώτο αυτό μέρος των σημειώσεων γίνεται αρχικά μια σύντομη επισκόπηση της θεωρίας πινάκων και στην συνέχεια πραγματοποιείται μια εισαγωγή στην πολυμεταβλητή κανονική κατανομή, καθώς και σε ελέγχους υποθέσεων που αφορούν τις παραμέτρους της.

2. ΠΟΛΥΜΕΤΑΒΛΗΤΑ ΔΕΔΟΜΕΝΑ

Πριν προχωρήσουμε στην στατιστική ανάλυση, θα πρέπει να βρούμε έναν τρόπο αναπαράστασης πολυμεταβλητών δεδομένων. Ας υποθέσουμε ότι έχουμε p το πλήθος μεταβλητές (variables), κάθε μια από τις οποίες καταγράφεται για n το πλήθος άτομα ή στοιχεία ή πειράματα (items, cases, individuals, experimental trials). Θα συμβολίζουμε με

$$x_{ij} = \text{μέτρηση της τιμής της } i \text{ μεταβλητής στο } j \text{ άτομο του δείγματος}$$

Επομένως το σύνολο των δεδομένων μπορεί να αναπαρασταθεί ως εξής:

	Άτομο 1	Άτομο 2	...	Άτομο j	...	Άτομο n
Μεταβλητή 1:	x_{11}	x_{12}	...	x_{1j}	...	x_{1n}
Μεταβλητή 2:	x_{21}	x_{22}	...	x_{2j}	...	x_{2n}
...	...	...	...	...	...	...
Μεταβλητή i :	x_{i1}	x_{i2}	...	x_{ij}	...	x_{in}
...	...	...	...	...	...	...
Μεταβλητή p :	x_{p1}	x_{p2}	...	x_{pj}	...	x_{pn}

Είναι πιο βολικό να αναπαραστήσουμε τα παραπάνω με τη μορφή πίνακα:

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1j} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2j} & \cdots & x_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{i1} & x_{i2} & \cdots & x_{ij} & \cdots & x_{in} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_{p1} & x_{p2} & \cdots & x_{pj} & \cdots & x_{pn} \end{bmatrix}$$

Χρησιμοποιούμε μικρό γράμμα x στα παραπάνω δεδομένα για να εκφράσουμε το γεγονός ότι πρόκειται για μετρήσεις «μετά» την πραγματοποίηση του πειράματος ή της δειγματοληψίας (δηλαδή πιο απλά πρόκειται για συγκεκριμένους αριθμούς). Όταν θέλουμε να συμβολίσουμε τα δεδομένα «πριν» την πραγματοποίηση του πειράματος τότε ως συνήθως χρησιμοποιούμε κεφαλαίο X (θα πρόκειται δηλαδή για τυχαία μεταβλητή).

Παράδειγμα 1. Ας θεωρήσουμε και πάλι ως παράδειγμα τα δεδομένα που αφορούν μετρήσεις από 36 άτομα. Για ευκολία ας περιοριστούμε στα 5 πρώτα άτομα. Τα δεδομένα αυτά αναπαρίστανται με τη μορφή πίνακα ως εξής ($p = 3$):

$$\mathbf{x} = \begin{bmatrix} 83 & 87.9 & 85.2 & 88.6 & 77.1 \\ 78.5 & 69.8 & 69.9 & 89.9 & 72.4 \\ 183.5 & 168.2 & 172 & 186.6 & 183.5 \end{bmatrix}$$

Θα συμβολίζουμε τα βασικά περιγραφικά μέτρα ενός πολυμεταβλητού δείγματος ως ακολούθως:

- Δειγματική μέση τιμή i μεταβλητής :

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij}, i = 1, 2, \dots, p$$

- Δειγματική διασπορά i μεταβλητής :

$$s_i^2 = s_{ii} = \frac{1}{n} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2, i = 1, 2, \dots, p$$

(μερικές φορές χρησιμοποιούμε $n - 1$ αντί για n στον παρονομαστή).

- Δειγματική συνδιασπορά ή συνδιακύμανση μεταξύ των μεταβλητών i, k :

$$s_{ik} = \frac{1}{n} \sum_{j=1}^n (x_{ij} - \bar{x}_i)(x_{kj} - \bar{x}_k), \quad i, k = 1, 2, \dots, p$$

Προφανώς $s_{ik} = s_{ki}$.

- Δειγματικός συντελεστής συσχέτισης (του Pearson) μεταξύ των μεταβλητών i, k :

$$r_{ik} = \frac{s_{ik}}{\sqrt{s_{ii}} \sqrt{s_{kk}}}, \quad i, k = 1, 2, \dots, p$$

Προφανώς $r_{ik} = r_{ki}$ και $r_{ii} = 1$. Ο δειγματικός συντελεστής συσχέτισης λαμβάνει τιμές μεταξύ των $-1, 1$ και μπορεί να θεωρηθεί ως ένα μέτρο της γραμμικής εξάρτησης μεταξύ των δύο μεταβλητών. Αν είναι θετικός τότε έχουμε θετική γραμμική εξάρτηση (όταν αυξάνεται η τιμή της μιας μεταβλητής τότε «τείνει να αυξάνεται» και η τιμή της άλλης), ενώ όταν είναι αρνητικός έχουμε αρνητική γραμμική εξάρτηση (όταν αυξάνεται η τιμή της μιας μεταβλητής τότε «τείνει να μειώνεται» η τιμή της άλλης). Όσο πιο κοντά στο 1 ή στο -1 είναι, τόσο ισχυρότερη είναι η θετική ή η αρνητική εξάρτηση αντίστοιχα.

Οι παραπάνω ποσότητες μπορούν για ευκολία να παρασταθούν με τη μορφή πίνακα:

- Διάνυσμα των δειγματικών μέσων τιμών των p μεταβλητών

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \vdots \\ \bar{x}_p \end{bmatrix}$$

- Πίνακας των δειγματικών συνδιασπορών ή συνδιακυμάνσεων μεταξύ των p μεταβλητών

$$\mathbf{s}_n = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{21} & s_{22} & \cdots & s_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1} & s_{p2} & \cdots & s_{pp} \end{bmatrix}$$

- Πίνακας των δειγματικών συντελεστών συσχέτισης μεταξύ των p μεταβλητών

$$\mathbf{r}_n = \begin{bmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{bmatrix}$$

Παράδειγμα 2. Ας θεωρήσουμε και πάλι ως παράδειγμα τα δεδομένα που αφορούν μετρήσεις από 36 άτομα. Το διάνυσμα των δειγματικών μέσων, και οι πίνακες των δειγματικών συνδιακυμάνσεων και συντελεστών συσχέτισης (ή συσχετίσεων) υπολογίζονται:

$$\bar{\mathbf{x}} = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \bar{x}_3 \end{bmatrix} = \begin{bmatrix} 88.086 \\ 81.172 \\ 180.486 \end{bmatrix}, \quad \mathbf{s}_n = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} = \begin{bmatrix} 81.58 & 84.27 & 4.98 \\ 84.27 & 125.94 & 38.88 \\ 4.98 & 38.88 & 52.56 \end{bmatrix}$$

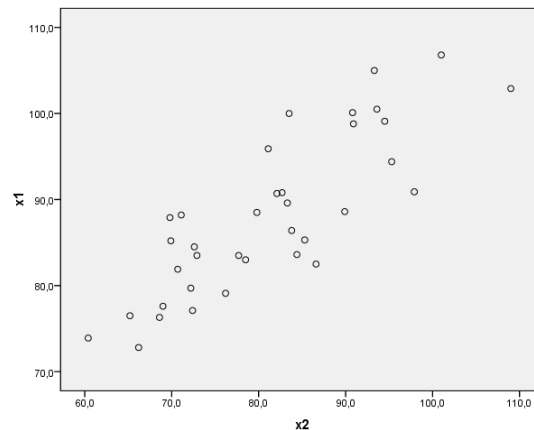
$$\mathbf{r}_n = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{21} & 1 & r_{23} \\ r_{31} & r_{32} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0.832 & 0.076 \\ 0.832 & 1 & 0.482 \\ 0.076 & 0.482 & 1 \end{bmatrix}.$$

Παρατηρούμε αρκετά υψηλή θετική (δειγματική) συσχέτιση μεταξύ των X_1 (Περίμετρος Κοιλίας) και X_2 (Βάρος) και επίσης υψηλή θετική (δειγματική) συσχέτιση μεταξύ των X_2 (Βάρος) και X_3 (Υψος). Πολύ μικρή δειγματική συσχέτιση (σχεδόν μηδενική) παρατηρείται μεταξύ των X_1 (Περίμετρος Κοιλίας) και X_3 (Υψος). Σε αυτή την φάση θα είχε νόημα να ελέγξουμε αν το ίδιο μπορεί να ισχύει σε όλο τον πληθυσμό από τον οποίο προήλθε το δείγμα (δηλαδή αν ισχύει για τις θεωρητικές συσχετίσεις ρ_{ij}).

Να σημειωθεί ότι, συνήθως, οι παραπάνω στατιστικές συναρτήσεις s_{ij} υπολογίζονται διαιρώντας με $n - 1$ αντί n (ώστε, όπως θα δούμε, να πρόκειται για αμερόληπτες εκτιμήτριες των σ_{ij}). Σε αυτή την περίπτωση ο αντίστοιχος πίνακας θα είναι

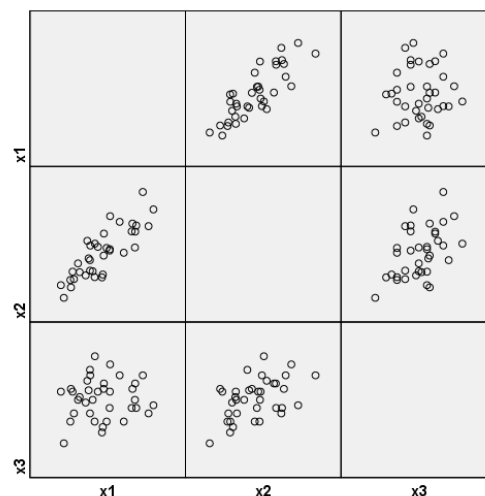
$$\mathbf{s}_n = \begin{bmatrix} 83.92 & 86.67 & 5.12 \\ 86.67 & 129.54 & 39.99 \\ 5.12 & 39.99 & 54.06 \end{bmatrix}.$$

Μπορούμε επίσης να χρησιμοποιήσουμε και κάποια γραφήματα για να πάρουμε μια πιο εποπτική εικόνα για την σχέση μεταξύ των μεταβλητών. Συνηθέστερα χρησιμοποιούμε τα διαγράμματα διασποράς (scatterplots). Το διάγραμμα διασποράς μεταξύ των μεταβλητών X_1 και X_2 είναι το γράφημα των σημείων $(x_{11}, x_{21}), (x_{12}, x_{22}), \dots, (x_{1n}, x_{2n})$.



Από το παραπάνω γράφημα μπορούμε να δούμε ότι οι μεταβλητές X_1 , X_2 παρουσιάζουν υψηλή θετική συσχέτιση (κάτι που φάνηκε και από τον δειγματικό συντελεστή συσχέτισης).

Με τον ίδιο τρόπο μπορούμε να κατασκευάσουμε όλα τα διαγράμματα διασποράς μεταξύ των μεταβλητών X_1 , X_2 , X_3 για να δούμε τη σχέση μεταξύ τους.



Το παραπάνω πολλαπλό διάγραμμα διασποράς (multiple scatterplot) επιβεβαιώνει τα συμπεράσματα που εξήχθησαν με βάση τον πίνακα των δειγματικών συσχετίσεων.

2. ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ ΑΠΟ ΤΗ ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ

Όπως διαφαίνεται και από τα παραπάνω, για να χειριστούμε πολυμεταβλητά δεδομένα πρέπει να είμαστε σε θέση να χειριστούμε διανύσματα και πίνακες. Για το λόγο αυτό στη συνέχεια γίνεται μια σύντομη επισκόπηση βασικών εννοιών από τη *Γραμμική Άλγεβρα*.

2.1. ΔΙΑΝΥΣΜΑΤΑ

(1) Διάνυσμα p συντεταγμένων καλείται μια διατεταγμένη συλλογή από p πραγματικούς αριθμούς $x_1, x_2, \dots, x_p$. Συνήθως συμβολίζεται χρησιμοποιώντας κάποιο παχύ γράμμα και παρίσταται σε μια στήλη ως εξής:

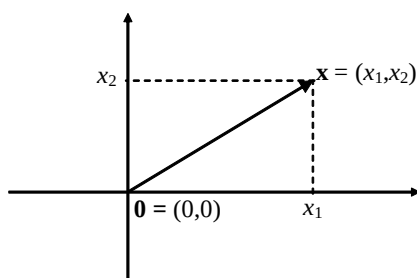
$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}, \text{ ή π.χ. } \mathbf{a} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} \text{ ή } \mathbf{b} = \begin{bmatrix} 3 \\ 2.2 \\ -4 \\ 1 \end{bmatrix} \text{ κ.ο.κ.}$$

Ορίζεται η πρόσθεση μεταξύ διανυσμάτων $\mathbf{x}, \mathbf{y}$ (ιδίου μεγέθους) και ο βαθμωτός πολλαπλασιασμός μεταξύ ενός πραγματικού αριθμού a και ενός διανύσματος $\mathbf{x}$:

$$\mathbf{z} = \mathbf{x} + \mathbf{y} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_p \end{bmatrix} = \begin{bmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_p + y_p \end{bmatrix}, \quad a\mathbf{x} = a \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = \begin{bmatrix} ax_1 \\ ax_2 \\ \vdots \\ ax_p \end{bmatrix}$$

Ο χώρος $V_p = R^p$ των διανυσμάτων (p συντεταγμένων) εφοδιασμένος με τις παραπάνω δύο πράξεις καλείται *διανυσματικός ή γραμμικός χώρος διάστασης p* .

(2) Κάθε στοιχείο (διάνυσμα) $\mathbf{x}$ του χώρου V_p μπορεί να θεωρηθεί ως ένα ευθύγραμμο τμήμα στο (υπερ)επίπεδο p διαστάσεων με αρχή το σημείο $\mathbf{0}$ και τέλος το σημείο $\mathbf{x}$. Αυτό μπορεί εύκολα να παρασταθεί όταν $p = 2$ (στο επίπεδο) ή όταν $p = 3$ (στο χώρο 3 διαστάσεων).



(3) Ο χώρος των γραμμικών συνδυασμών των διανυσμάτων $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ (με p συντεταγμένες) ο οποίος αποτελείται από όλα τα διανύσματα της μορφής

$$\mathbf{y} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_k\mathbf{x}_k, \quad a_1, a_2, \dots, a_k \in R$$

καλείται *γραμμικός υπόχωρος* που παράγεται από τα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$.

Για παράδειγμα, ο υπόχωρος που αποτελείται από όλα τα διανύσματα

$$\mathbf{y} = a_1 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad a_1, a_2 \in R$$

είναι το επίπεδο του R^3 που καθορίζεται από τα δύο παραπάνω διανύσματα.

(4) Ένα σύνολο από διανύσματα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ καλείται *γραμμικά εξαρτημένο* αν υπάρχουν πραγματικοί αριθμοί $a_1, a_2, \dots, a_k$, όχι όλοι ίσοι με 0, έτσι ώστε

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_k\mathbf{x}_k = \mathbf{0}$$

(δηλαδή κάποιο από τα $\mathbf{x}_i$ μπορεί να παρασταθεί ως γραμμικός συνδυασμός των υπολοίπων). Διαφορετικά, τα παραπάνω διανύσματα καλούνται *γραμμικά ανεξάρτητα*. Για παράδειγμα τα διανύσματα

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 9 \\ 4 \\ 2 \end{bmatrix}$$

είναι γραμμικά εξαρτημένα διότι $\mathbf{x}_1 + 2\mathbf{x}_2 - \mathbf{x}_3 = \mathbf{0}$. Ενώ για παράδειγμα τα διανύσματα

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

είναι γραμμικά ανεξάρτητα (διότι $a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + a_3\mathbf{x}_3 = \mathbf{0} \Leftrightarrow a_1 = a_2 = a_3 = 0$).

(5) Κάθε σύνολο γραμμικά ανεξάρτητων διανυσμάτων $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ (διάστασης p) καλείται *βάση* του διανυσματικού χώρου V_p . Προφανώς η βάση ενός χώρου δεν είναι μοναδική. Αν $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ είναι μια βάση του V_p , τότε κάθε διάνυσμα $\mathbf{y}$ του V_p μπορεί να γραφεί με μοναδικό τρόπο ως γραμμικός συνδυασμός των $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$. Η συνήθης βάση του χώρου V_p είναι η

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_p = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

Χρησιμοποιώντας την παραπάνω βάση, κάθε διάνυσμα του χώρου V_p προφανώς γράφεται ως εξής:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_p\mathbf{e}_p.$$

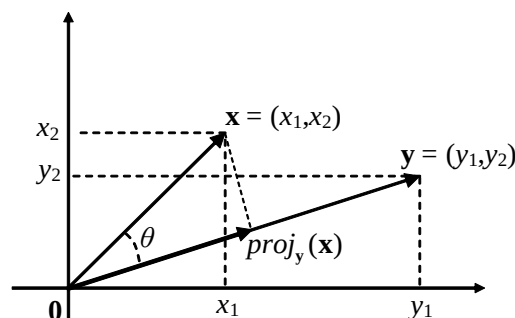
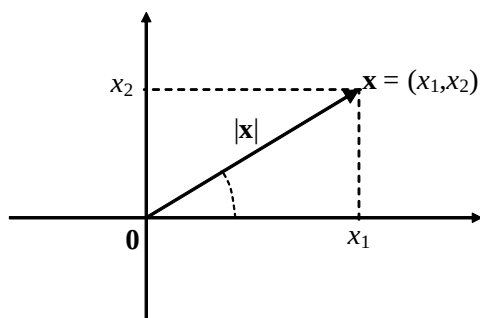
(6) Το *μήκος* $|\mathbf{x}|$ ενός διανύσματος $\mathbf{x}$ θα είναι,

$$|\mathbf{x}| = \sqrt{x_1^2 + x_2^2 + \dots + x_p^2} = \sqrt{\mathbf{x}'\mathbf{x}},$$

ενώ για τη *γωνία* θ μεταξύ δύο διανυσμάτων $\mathbf{x}, \mathbf{y}$ ισχύει

$$\cos(\theta) = \frac{x_1y_1 + x_2y_2 + \dots + x_py_p}{|\mathbf{x}| \cdot |\mathbf{y}|} = \frac{\mathbf{x}'\mathbf{y}}{\sqrt{\mathbf{x}'\mathbf{x}} \cdot \sqrt{\mathbf{y}'\mathbf{y}}}$$

(διότι: $\cos(\theta) = \cos(\theta_1 - \theta_2) = \cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2) = (y_1/|\mathbf{y}|)(x_1/|\mathbf{x}|) + (y_2/|\mathbf{y}|)(x_2/|\mathbf{x}|)$) όπου με $\mathbf{x}'\mathbf{y} = \langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + x_2y_2 + \dots + x_py_p$ συμβολίζουμε το *εσωτερικό γινόμενο* των $\mathbf{x}, \mathbf{y}$.



Επίσης, η προβολή του διανύσματος $\mathbf{x}$ πάνω στο διάνυσμα $\mathbf{y}$ (βλ. σχήμα) είναι

$$\text{proj}_{\mathbf{y}}(\mathbf{x}) = \frac{\mathbf{x}'\mathbf{y}}{\mathbf{y}'\mathbf{y}} \mathbf{y} = \cos(\theta) |\mathbf{x}| \cdot \frac{\mathbf{y}}{|\mathbf{y}|}.$$

Το παραπάνω προκύπτει από ότι η προβολή $\text{proj}_{\mathbf{y}}(\mathbf{x})$ έχει την ίδια διεύθυνση με το $\mathbf{y}$ (και άρα $\text{proj}_{\mathbf{y}}(\mathbf{x}) = a\mathbf{y}$ για κάποιο $a \in \mathbb{R}$) και μήκος $|\text{proj}_{\mathbf{y}}(\mathbf{x})| = |\mathbf{x}| \cos \theta$ (αν $\theta \leq \pi/2$). Από τα δύο προηγούμενα, θα είναι $a|\mathbf{y}| = |\mathbf{x}| \cos \theta \Rightarrow a = |\mathbf{x}| \cos \theta / |\mathbf{y}|$ (όμοια αποδ. για $\theta > \pi/2$).

(7) Αν $\mathbf{x}'\mathbf{y} = 0$ τότε $\cos(\theta) = 0$ και στην περίπτωση αυτή τα δύο διανύσματα $\mathbf{x}, \mathbf{y}$ καλούνται *κάθετα* ($\theta = \pm \pi/2$). Ένα σύνολο από ανά δύο κάθετα μεταξύ τους διανύσματα είναι γραμμικά ανεξάρτητα. Επίσης αν ένα $\mathbf{y}$ είναι κάθετο προς καθένα από τα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ τότε είναι κάθετο και προς όλο το γραμμικό υπόχωρο που παράγεται από τα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$.

(8) *Gram – Schmidt ορθογωνιοποίηση*: Αν τα διανύσματα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$ είναι γραμμικά ανεξάρτητα, τότε μπορούμε να βρούμε κάθετα ανά δύο διανύσματα $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$ που παράγουν τον ίδιο γραμμικό υπόχωρο με τα $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k$. Μάλιστα, μπορούμε αναδρομικά να πάρουμε

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{x}_1 \\ \mathbf{u}_2 &= \mathbf{x}_2 - \text{proj}_{\mathbf{u}_1}(\mathbf{x}_2) \\ \mathbf{u}_3 &= \mathbf{x}_3 - \text{proj}_{\mathbf{u}_1}(\mathbf{x}_3) - \text{proj}_{\mathbf{u}_2}(\mathbf{x}_3) \\ &\dots \\ \mathbf{u}_k &= \mathbf{x}_k - \text{proj}_{\mathbf{u}_1}(\mathbf{x}_k) - \dots - \text{proj}_{\mathbf{u}_{k-1}}(\mathbf{x}_k). \end{aligned}$$

Μπορούμε επίσης να πάρουμε ως ορθογώνια βάση την $\mathbf{z}_1 = \mathbf{u}_1 / \sqrt{\mathbf{u}_1'\mathbf{u}_1}$, $\mathbf{z}_2 = \mathbf{u}_2 / \sqrt{\mathbf{u}_2'\mathbf{u}_2}$, ... ώστε τα $\mathbf{z}_i$ να έχουν μήκος 1.

2.2. ΠΙΝΑΚΕΣ

Ένας πίνακας $\mathbf{A}$ διαστάσεων $p \times n$ αποτελείται από pn πραγματικούς αριθμούς τοποθετημένους σε p γραμμές και n στήλες. Η γενική του μορφή είναι

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{bmatrix} = [a_{ij}]_{p \times n}$$

Ένα διάνυσμα $\mathbf{x}$ μπορεί να θεωρηθεί ως ένας πίνακας $p \times 1$. Δύο πίνακες ίδιων διαστάσεων $\mathbf{A} = [a_{ij}]_{p \times n}$, $\mathbf{B} = [b_{ij}]_{p \times n}$ είναι ίσοι αν $a_{ij} = b_{ij}$ για κάθε i, j .

(1) Όμοια με τα διανύσματα, ορίζεται η πρόσθεση μεταξύ πινάκων $\mathbf{A}, \mathbf{B}$ (ίδιων διαστάσεων) και ο βαθμωτός πολλαπλασιασμός μεταξύ ενός πραγματικού αριθμού c και ενός πίνακα $\mathbf{A}$:

$$\mathbf{C} = \mathbf{A} + \mathbf{B} = [a_{ij}]_{p \times n} + [b_{ij}]_{p \times n} = [a_{ij} + b_{ij}]_{p \times n},$$

$$c\mathbf{A} = c[a_{ij}]_{p \times n} = [ca_{ij}]_{p \times n},$$

και $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-1)\mathbf{B}$.

Επίσης ορίζεται και ο πολλαπλασιασμός πινάκων $\mathbf{A} = [a_{ij}]_{p \times n}$ επί $\mathbf{B} = [b_{ij}]_{n \times m}$:

$$\mathbf{C} = \mathbf{AB} = [a_{ij}]_{p \times n} \cdot [b_{ij}]_{n \times m} = \left[\sum_{l=1}^n a_{il} b_{lj} \right]_{p \times m}$$

Για παράδειγμα,

$$\mathbf{AB} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}$$

Ο πολλαπλασιασμός πινάκων είναι συμβατός με το εσωτερικό γινόμενο διανυσμάτων που είδαμε παραπάνω. Ο *ανάστροφος* ενός πίνακα $\mathbf{A} = [a_{ij}]_{p \times n}$ θα συμβολίζεται με $\mathbf{A}^T$ ή με $\mathbf{A}'$ και προκύπτει μετατρέποντας τις γραμμές σε στήλες, δηλαδή $\mathbf{A}' = [a_{ji}]_{n \times p}$.

(2) Ένας πίνακας καλείται *τετραγωνικός* αν έχει τόσες γραμμές όσες και στήλες. *Μοναδιαίος* πίνακας $\mathbf{I}_n$ διάστασης n καλείται ο τετραγωνικός πίνακας $n \times n$ που έχει στην κύρια διαγώνιο n μονάδες και 0 σε όλες τις άλλες θέσεις, δηλαδή

$$\mathbf{I}_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Αν $\mathbf{A}$ είναι πίνακας $n \times n$ τότε $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$. *Μηδενικός* (συμβ. $\mathbf{0}$) καλείται ο πίνακας που έχει όλα τα στοιχεία του ίσα με 0. *Συμμετρικός* καλείται ένας τετραγωνικός πίνακας που είναι ίσος με τον ανάστροφό του. *Διαγώνιος* καλείται ένας τετραγωνικός πίνακας που έχει μη μηδενικά στοιχεία μόνο στην κύρια διαγώνιο του. Αν τα στοιχεία αυτά είναι τα $a_1, a_2, \dots, a_n$ τότε συμβολίζεται και με $\text{diag}\{a_1, a_2, \dots, a_n\}$, π.χ.

$$\text{diag}\{1,3,4\} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

(3) Αποδεικνύεται εύκολα ότι ισχύουν οι ακόλουθες ιδιότητες:

(i) $(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C})$

(ii) $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$

(iii) $c(\mathbf{A} + \mathbf{B}) = c\mathbf{A} + c\mathbf{B}$

(iv) $(c + d)\mathbf{A} = c\mathbf{A} + d\mathbf{A}$

(v) $(a\mathbf{A} + b\mathbf{B})' = a\mathbf{A}' + b\mathbf{B}'$

(vi) $c(\mathbf{AB}) = (c\mathbf{A})\mathbf{B}$

(vii) $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$

(viii) $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$

(ix) $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$

(x) $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

Ιδιαίτερη προσοχή θα πρέπει να δοθεί στο γεγονός ότι κατά τον πολλαπλασιασμό πινάκων δεν ισχύει η αντιμεταθετική ιδιότητα, δηλαδή, γενικά, $\mathbf{AB} \neq \mathbf{BA}$. Επίσης η ισότητα $\mathbf{AB} = \mathbf{0}$ δεν συνεπάγεται πάντοτε ότι $\mathbf{A} = \mathbf{0}$ ή $\mathbf{B} = \mathbf{0}$ όπως ισχύει στους πραγματικούς αριθμούς.

(4) Η *ορίζουσα* ενός τετραγωνικού πίνακα $\mathbf{A} = [a_{ij}]_{n \times n}$ συμβολίζεται με $\text{Det}(\mathbf{A})$ ή $|\mathbf{A}|$ και είναι ο αριθμός που ορίζεται αναδρομικά ως εξής:

$$|\mathbf{A}| = a_{11}, \text{ αν } n = 1,$$

$$|\mathbf{A}| = a_{11}|\mathbf{A}_{11}| - a_{12}|\mathbf{A}_{12}| + \dots + (-1)^{n+1}a_{1n}|\mathbf{A}_{1n}|$$

όπου $|\mathbf{A}_{ij}|$ είναι η ορίζουσα του πίνακα που προκύπτει από τον $\mathbf{A}$ αφού του αποκοπεί η i γραμμή και η j στήλη. Στην περίπτωση που $n = 2$ και $n = 3$ αποδεικνύεται εύκολα ότι

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12},$$

και

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{31}a_{22}a_{13} - a_{21}a_{12}a_{33} - a_{11}a_{32}a_{23}.$$

(5) Ο βαθμός (rank) των γραμμών (στηλών) ενός πίνακα είναι ο μέγιστος αριθμός από γραμμικά ανεξάρτητες γραμμές (στήλες), θεωρούμενες ως διανύσματα. Αποδεικνύεται ότι ο βαθμός των στηλών είναι ίσος με το βαθμό των γραμμών ενός πίνακα και για αυτό καλείται βαθμός του πίνακα.

(6) Αν $\mathbf{A}$ είναι ένας τετραγωνικός πίνακας $n \times n$, αποδεικνύεται ότι τα ακόλουθα είναι ισοδύναμα:

(i) $\mathbf{Ax} = \mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$ (δηλ. οι στήλες είναι γραμμικά ανεξάρτητες ή ισοδύναμα ο βαθμός του $\mathbf{A}$ είναι ίσος με n),

(ii) η ορίζουσα $|\mathbf{A}| \neq 0$,

(iii) Υπάρχει ένας πίνακας $\mathbf{B}$ διαστάσεων $n \times n$ τέτοιος ώστε $\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$.

Ο παραπάνω πίνακας $\mathbf{B}$ καλείται *αντίστροφος* του $\mathbf{A}$ και συμβολίζεται με $\mathbf{A}^{-1}$. Αποδεικνύεται ότι το i, j στοιχείο του πίνακα $\mathbf{A}^{-1}$ δίνεται από τον γενικό τύπο

$$(-1)^{i+j} \frac{|\mathbf{A}_{ji}|}{|\mathbf{A}|}.$$

Για $n = 2$ έχουμε απλά ότι

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} |\mathbf{A}_{11}| & -|\mathbf{A}_{21}| \\ -|\mathbf{A}_{12}| & |\mathbf{A}_{22}| \end{bmatrix} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}.$$

Όμοια, για $n = 3$ έχουμε

$$\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} |\mathbf{A}_{11}| & -|\mathbf{A}_{21}| & |\mathbf{A}_{31}| \\ -|\mathbf{A}_{12}| & |\mathbf{A}_{22}| & -|\mathbf{A}_{32}| \\ |\mathbf{A}_{13}| & -|\mathbf{A}_{23}| & |\mathbf{A}_{33}| \end{bmatrix} = \frac{1}{|\mathbf{A}|} \begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} \\ \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}.$$

Για τετραγωνικούς $k \times k$ πίνακες $\mathbf{A}, \mathbf{B}$ ισχύουν οι ακόλουθες ιδιότητες:

- (i) $(\mathbf{A}^{-1})' = (\mathbf{A}')^{-1}$ (αν $\mathbf{A}$ αντιστρέψιμος)
- (ii) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ (αν $\mathbf{A}, \mathbf{B}$ αντιστρέψιμοι)
- (iii) $|\mathbf{A}| = |\mathbf{A}'|$
- (iv) Αν όλα τα στοιχεία μιας γραμμής ή μιας στήλης του $\mathbf{A}$ είναι 0 τότε $|\mathbf{A}| = 0$.
- (iv) Αν δύο γραμμές (ή στήλες) του $\mathbf{A}$ είναι ίδιες τότε $|\mathbf{A}| = 0$.
- (v) Αν υπάρχει ο $\mathbf{A}^{-1}$ τότε $|\mathbf{A}^{-1}| = 1/|\mathbf{A}|$
- (vi) $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$
- (vii) $|c\mathbf{A}| = c^k |\mathbf{A}|$

(7) Το ίχνος ενός τετραγωνικού πίνακα $\mathbf{A} = [a_{ij}]_{n \times n}$ ορίζεται ως το άθροισμα των στοιχείων της διαγωνίου του

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n a_{ii}$$

Για τετραγωνικούς $n \times n$ πίνακες ισχύει ότι:

$$(i) \text{tr}(c\mathbf{A}) = c \text{tr}(\mathbf{A}), \quad (ii) \text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B}), \quad (iii) \text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}),$$

$$(iv) \text{tr}(\mathbf{B}^{-1}\mathbf{AB}) = \text{tr}(\mathbf{A}), \quad (v) \text{tr}(\mathbf{AA}') = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2$$

(8) Ένας τετραγωνικός πίνακας καλείται *ορθογώνιος* αν $\mathbf{AA}' = \mathbf{I}$ (δηλαδή οι γραμμές του θεωρούμενες ως διανύσματα είναι μεταξύ τους κάθετες με μοναδιαίο μήκος). Αποδεικνύεται ότι ο $\mathbf{A}$ είναι ορθογώνιος αν και μόνο αν $\mathbf{A}^{-1} = \mathbf{A}'$. Αν $\mathbf{AA}' = \mathbf{I}$ τότε και $\mathbf{A}'\mathbf{A} = \mathbf{I}$.

(9) Αν $\mathbf{A}$ είναι ένας τετραγωνικός $n \times n$ πίνακας, τότε οι αριθμοί $\lambda_1, \lambda_2, \dots, \lambda_n$ (πραγματικοί ή μιγαδικοί) που ικανοποιούν την πολυωνυμική εξίσωση

$$|\mathbf{A} - \lambda \mathbf{I}| = 0,$$

καλούνται *ιδιοτιμές* του $\mathbf{A}$. Η εξίσωση αυτή καλείται *χαρακτηριστική εξίσωση*. Π.χ. οι ιδιοτιμές λ_1, λ_2 του πίνακα

$$\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix}$$

είναι οι δύο ρίζες της εξίσωσης $|\mathbf{A} - \lambda \mathbf{I}| = 0$, δηλαδή

$$\begin{vmatrix} 2-\lambda & 3-0 \\ 5-0 & 4-\lambda \end{vmatrix} = \lambda^2 - 6\lambda - 7 = 0 \Leftrightarrow \lambda_1 = -1, \lambda_2 = 7.$$

(10) Αν λ είναι μια ιδιοτιμή του $\mathbf{A}$, τότε ένα διάνυσμα $\mathbf{x} \neq \mathbf{0}$ που ικανοποιεί την εξίσωση

$$\mathbf{Ax} = \lambda \mathbf{x},$$

καλείται *ιδιοδιάνυσμα* του $\mathbf{A}$ που αντιστοιχεί στην ιδιοτιμή λ . Π.χ. για την ιδιοτιμή $\lambda_1 = -1$ έχουμε το σύστημα των εξισώσεων

$$\mathbf{Ax} = \lambda_1 \mathbf{x} \Leftrightarrow \begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = -1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \Leftrightarrow \begin{cases} 2x_1 + 3x_2 = -x_1 \\ 5x_1 + 4x_2 = -x_2 \end{cases} \Leftrightarrow \begin{cases} 3x_1 + 3x_2 = 0 \\ 5x_1 + 5x_2 = 0 \end{cases}$$

Οι παραπάνω δύο εξισώσεις είναι ισοδύναμες και καταλήγουμε σε μια εξίσωση, την $x_1 + x_2 = 0$, η οποία έχει άπειρες λύσεις ως προς $\mathbf{x}$ και επομένως έχουμε άπειρες ιδιοτιμές της μορφής $\mathbf{x}_1 = [a, -a]$, $a \in \mathbb{R}$. Για παράδειγμα ένα ιδιοδιάνυσμα για την ιδιοτιμή $\lambda_1 = -1$ είναι το $\mathbf{x}_1 = [1, -1]$. Με τον ίδιο τρόπο, και για την λ_2 έχουμε άπειρες ιδιοτιμές της μορφής $\mathbf{x}_2 = [a, 5a/3]$, $a \in \mathbb{R}$. Συνήθως λαμβάνουμε τις μοναδιαίες ιδιοτιμές διαιρώντας τις με το μήκος τους:

$$\mathbf{e}_1 = \frac{\mathbf{x}_1}{\sqrt{\mathbf{x}_1' \mathbf{x}_1}} = \frac{1}{\sqrt{a^2 + (-a)^2}} \begin{bmatrix} a \\ -a \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix},$$

$$\mathbf{e}_2 = \frac{\mathbf{x}_2}{\sqrt{\mathbf{x}_2' \mathbf{x}_2}} = \frac{1}{\sqrt{a^2 + (5a/3)^2}} \begin{bmatrix} a \\ 5a/3 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{34} \\ 5/\sqrt{34} \end{bmatrix}.$$

Η διαδικασία εύρεσης των ιδιοτιμών και των ιδιοδιανυσμάτων ενός πίνακα είναι υπολογιστικά πολύ επίπονη (ιδιαίτερα για μεγάλους πίνακες) αλλά αυτό δεν αποτελεί πρόβλημα διότι στην πράξη οι υπολογισμοί γίνονται πάντοτε από Η/Υ.

Αποδεικνύεται ότι αν ο $\mathbf{A}$ είναι συμμετρικός $k \times k$ πίνακας τότε μπορούν να βρεθούν k (μοναδικά) ιδιοδιανύσματα $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ τα οποία έχουν μοναδιαίο μήκος ($\mathbf{e}_i' \mathbf{e}_i = 1$) και είναι μεταξύ τους κάθετα, $\mathbf{e}_i' \mathbf{e}_j = 0$ (καλούνται ορθοκανονικά).

Στο εξής θα ασχοληθούμε κατά κύριο λόγο με *συμμετρικούς πίνακες*.

(11) *Φασματική αποσύνθεση ή παραγοντοποίηση (spectral decomposition)*. Αποδεικνύεται ότι αν ο $\mathbf{A}$ είναι ένας $k \times k$ συμμετρικός πίνακας, τότε γράφεται στη μορφή

$$\mathbf{A} = \lambda_1 \mathbf{e}_1 \mathbf{e}_1' + \lambda_2 \mathbf{e}_2 \mathbf{e}_2' + \dots + \lambda_k \mathbf{e}_k \mathbf{e}_k'$$

όπου $\lambda_1, \lambda_2, \dots, \lambda_k$ είναι οι ιδιοτιμές του $\mathbf{A}$ και $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ τα αντίστοιχα ορθοκανονικά ιδιοδιανύσματα (π.χ. είναι εύκολο να επαληθεύσουμε από την παραπάνω αναπαράσταση ότι $\mathbf{A} \mathbf{e}_i = \lambda_i \mathbf{e}_i$ και επομένως τα $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k$ είναι πράγματι ιδιοδιανύσματα του $\mathbf{A}$).

Αν συμβολίσουμε με $\mathbf{P}$ τον πίνακα που έχει ως στήλες τα παραπάνω k ιδιοδιανύσματα και με $\mathbf{\Lambda}$ τον διαγώνιο πίνακα που έχει στην κύρια διαγώνιο τις παραπάνω ιδιοτιμές (και 0 αλλού) τότε η παραπάνω αναπαράσταση του $\mathbf{A}$ (γνωστή και ως φασματική αποσύνθεση) γράφεται ως εξής

$$\mathbf{A} = \mathbf{P} \mathbf{\Lambda} \mathbf{P}' = \begin{bmatrix} \mathbf{e}_{11} & \mathbf{e}_{21} & \dots & \mathbf{e}_{k1} \\ \mathbf{e}_{12} & \mathbf{e}_{22} & & \mathbf{e}_{k2} \\ \vdots & & \ddots & \vdots \\ \mathbf{e}_{1k} & \mathbf{e}_{2k} & \dots & \mathbf{e}_{kk} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & \lambda_k \end{bmatrix} \begin{bmatrix} \mathbf{e}_{11} & \mathbf{e}_{12} & \dots & \mathbf{e}_{1k} \\ \mathbf{e}_{21} & \mathbf{e}_{22} & & \mathbf{e}_{2k} \\ \vdots & & \ddots & \vdots \\ \mathbf{e}_{k1} & \mathbf{e}_{k2} & \dots & \mathbf{e}_{kk} \end{bmatrix}$$

(όπου $\mathbf{e}_i = [e_{i1}, e_{i2}, \dots, e_{ik}]'$). Για παράδειγμα, είναι εύκολο να βρούμε ότι ο συμμετρικός πίνακας

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix},$$

έχει ιδιοτιμές τις $\lambda_1 = 3$ και $\lambda_2 = 1$ με αντίστοιχα (ορθοκανονικά) ιδιοδιανύσματα τα $\mathbf{e}_1 = [1/\sqrt{2}, 1/\sqrt{2}]'$ και $\mathbf{e}_2 = [-1/\sqrt{2}, 1/\sqrt{2}]'$ και να επαληθεύσουμε ότι πράγματι,

$$\mathbf{P} \mathbf{\Lambda} \mathbf{P}' = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{3}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \mathbf{A}.$$

(12) *Δυνάμεις πίνακα*. Η φασματική αποσύνθεση ενός πίνακα μας βοηθά να εκφράσουμε με απλό τρόπο τις δυνάμεις ενός συμμετρικού πίνακα $\mathbf{A}$. Παρατηρούμε ότι, αν m φυσικός αριθμός,

$$\mathbf{A}^m = (\mathbf{P}\mathbf{\Lambda}\mathbf{P}')^m = (\mathbf{P}\mathbf{\Lambda}\mathbf{P}')(\mathbf{P}\mathbf{\Lambda}\mathbf{P}')\dots(\mathbf{P}\mathbf{\Lambda}\mathbf{P}') = \mathbf{P}\mathbf{\Lambda}(\mathbf{P}'\mathbf{P})\mathbf{\Lambda}(\mathbf{P}'\dots\mathbf{P})\mathbf{\Lambda}\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}^m\mathbf{P}'$$

και ο $\mathbf{\Lambda}^m$ είναι ο διαγώνιος πίνακας που στην κύρια διαγώνιο έχει τις ιδιοτιμές υψωμένες στην δύναμη m . Επίσης αν ο $\mathbf{A}$ έχει ιδιοτιμές διαφορετικές του 0, τότε

$$\mathbf{A}^{-1} = (\mathbf{P}\mathbf{\Lambda}\mathbf{P}')^{-1} = (\mathbf{P}')^{-1}\mathbf{\Lambda}^{-1}\mathbf{P}^{-1} = (\mathbf{P}')^{-1}\mathbf{\Lambda}^{-1}\mathbf{P}^{-1} = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}'$$

και ο $\mathbf{\Lambda}^{-1}$ έχει στη διαγώνιο τις αντίστροφες τιμές των ιδιοτιμών. Από τα παραπάνω, αν ο συμμετρικός $\mathbf{A}$ έχει θετικές ιδιοτιμές τότε μπορούμε να ορίσουμε και τη ρίζα του $\mathbf{A}$ ως εξής:

$$\mathbf{A}^{1/2} = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'$$

(ο $\mathbf{\Lambda}^{1/2}$ έχει στη διαγώνιο τις ρίζες των ιδιοτιμών του $\mathbf{A}$). Ο παραπάνω πίνακας έχει τις ιδιότητες της ρίζας δηλαδή,

$$\mathbf{A}^{1/2}\mathbf{A}^{1/2} = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}'\mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}^{1/2}\mathbf{\Lambda}^{1/2}\mathbf{P}' = \mathbf{P}\mathbf{\Lambda}\mathbf{P}' = \mathbf{A}.$$

Γενικότερα μπορούμε να ορίσουμε $\mathbf{A}^b = \mathbf{P}\mathbf{\Lambda}^b\mathbf{P}'$, $b \in \mathbb{R}$ (αρκεί $\lambda_i > 0$, $i = 1, 2, \dots, k$).

(13) *Τετραγωνική μορφή πίνακα*. Τετραγωνική μορφή καλείται μία συνάρτηση k μεταβλητών $x_1, x_2, \dots, x_k$ η οποία γράφεται στη μορφή

$$Q(\mathbf{x}) = \sum_{i=1}^k \sum_{j=1}^k a_{ij} x_i x_j = \mathbf{x}'\mathbf{A}\mathbf{x}$$

όπου $\mathbf{x} = [x_1, x_2, \dots, x_k]'$, $\mathbf{A} = [a_{ij}]_{k \times k}$ και $a_{ij} = a_{ji}$ (δηλαδή ο $\mathbf{A}$ είναι συμμετρικός).

Π.χ. αν $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ τότε $Q(\mathbf{x}) = 2x_1^2 + 2x_1x_2 + 2x_2^2$

(14) Αν για ένα $k \times k$ συμμετρικό πίνακα $\mathbf{A}$ ισχύει ότι

$$\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0,$$

για κάθε $\mathbf{x} = [x_1, x_2, \dots, x_k]'$, τότε ο $\mathbf{A}$ καλείται *οριστικά μη-αρνητικός* (non negative definite). Αν ισχύει ότι $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ για κάθε $\mathbf{x} \neq \mathbf{0}$, τότε καλείται *οριστικά θετικός* (positive definite). Χρησιμοποιώντας την φασματική αποσύνθεση του πίνακα, αποδεικνύεται εύκολα ότι ο $\mathbf{A}$ είναι οριστικά μη αρνητικός (θετικός) αν και μόνο αν όλες οι ιδιοτιμές του είναι μη-αρνητικές (θετικές). Πράγματι,

$$\mathbf{x}'\mathbf{A}\mathbf{x} = \lambda_1 \mathbf{x}'\mathbf{e}_1 \mathbf{e}_1' \mathbf{x} + \lambda_2 \mathbf{x}'\mathbf{e}_2 \mathbf{e}_2' \mathbf{x} + \dots + \lambda_k \mathbf{x}'\mathbf{e}_k \mathbf{e}_k' \mathbf{x} = \lambda_1 (\mathbf{x}'\mathbf{e}_1)^2 + \lambda_2 (\mathbf{x}'\mathbf{e}_2)^2 + \dots + \lambda_k (\mathbf{x}'\mathbf{e}_k)^2,$$

το οποίο είναι μη-αρνητικό (θετικό) για κάθε $\mathbf{x}$ αν και μόνο αν οι ιδιοτιμές είναι μη-αρνητικές (θετικές).

Για παράδειγμα ο πίνακας $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$ που εξετάσαμε παραπάνω έχει θετικές ιδιοτιμές και επομένως είναι οριστικά θετικός.

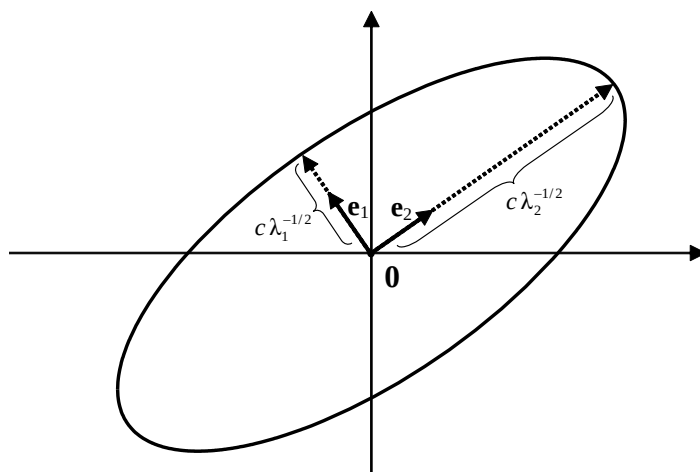
(15) Όπως θα δούμε, σε ορισμένες περιπτώσεις έχει ενδιαφέρον να δούμε ποιος είναι ο χώρος (γεωμετρικός τόπος) των $\mathbf{x} \in \mathbb{R}^k$ που ικανοποιούν την εξίσωση $\mathbf{x}'\mathbf{A}\mathbf{x} = c^2$, για κάποιο $k \times k$ συμμετρικό πίνακα $\mathbf{A}$ (και $c > 0$). Ας εξετάσουμε αρχικά το πρόβλημα αυτό στις δύο διαστάσεις ($k = 2$) που μπορούμε να έχουμε εποπτεία. Αν ο $\mathbf{A}$ είναι ο μοναδιαίος πίνακας τότε $\mathbf{x}'\mathbf{A}\mathbf{x} = \mathbf{x}'\mathbf{x} = x_1^2 + x_2^2$ και η εξίσωση $\mathbf{x}'\mathbf{A}\mathbf{x} = c^2$ είναι της μορφής

$$x_1^2 + x_2^2 = c^2.$$

Είναι γνωστό ότι τα σημεία (x_1, x_2) του επιπέδου που ικανοποιούν την παραπάνω εξίσωση βρίσκονται πάνω στην περιφέρεια ενός κύκλου με κέντρο το 0 και ακτίνα c . Γενικότερα, από την φασματική αποσύνθεση του πίνακα έχουμε ότι

$$\mathbf{x}'\mathbf{A}\mathbf{x} = c^2 \Leftrightarrow \lambda_1(\mathbf{x}'\mathbf{e}_1)^2 + \lambda_2(\mathbf{x}'\mathbf{e}_2)^2 = c^2 \Leftrightarrow \frac{y_1^2}{(c/\sqrt{\lambda_1})^2} + \frac{y_2^2}{(c/\sqrt{\lambda_2})^2} = 1,$$

όπου $y_i = \mathbf{x}'\mathbf{e}_i$, $i = 1, 2$, και είναι γνωστό ότι την παραπάνω εξίσωση ικανοποιούν τα σημεία (y_1, y_2) που βρίσκονται πάνω στην περιφέρεια μιας έλλειψης με ακτίνες $c/\sqrt{\lambda_1}, c/\sqrt{\lambda_2}$. Είναι εύκολο τώρα να δούμε ότι τα $\mathbf{x} = [\mathbf{e}_1, \mathbf{e}_2] [y_1, y_2]'$ που ικανοποιούν την παραπάνω βρίσκονται πάνω στην περιφέρεια μιας περιστραμμένης έλλειψης με κύριους άξονες τα διανύσματα $\mathbf{e}_1, \mathbf{e}_2$ και αντίστοιχες ακτίνες τις $c/\sqrt{\lambda_1}, c/\sqrt{\lambda_2}$ (βλέπε σχήμα).



Σε μεγαλύτερες τώρα διαστάσεις ($k > 2$) προκύπτει ανάλογο αποτέλεσμα, δηλαδή τα σημεία που ικανοποιούν την $\mathbf{x}'\mathbf{A}\mathbf{x} = c^2$ βρίσκονται πάνω στην (υπερ)επιφάνεια ενός (υπερ)ελλειψοειδούς με κύριους άξονες τα ιδιοδιανύσματα του πίνακα $\mathbf{A}$ και αντίστοιχες ακτίνες τις $c/\sqrt{\lambda_i}$, $i = 1, 2, 3, \dots, k$.

3. ΤΥΧΑΙΑ ΔΙΑΝΥΣΜΑΤΑ ΚΑΙ ΤΥΧΑΙΟΙ ΠΙΝΑΚΕΣ

Κατά την πολυμεταβλητή στατιστική ανάλυση, από κάθε τυχαία επιλεγμένο άτομο ενός πληθυσμού δεν καταγράφουμε μόνο μια τιμή X (όπως στην μονομεταβλητή στατιστική ανάλυση), αλλά περισσότερες τιμές, οι οποίες μπορούν να παρασταθούν με ένα διάνυσμα:

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix},$$

(π.χ. $X_1 = \text{βάρος}$, $X_2 = \text{ύψος}$, $X_3 = \text{περιφέρεια κοιλίας}$). Το διάνυσμα αυτό καλείται *τυχαίο διάνυσμα* και αποτελείται από p συντεταγμένες οι οποίες είναι τυχαίες μεταβλητές. Γενικά, οι τυχαίες μεταβλητές $X_1, X_2, \dots, X_p$ δεν είναι ανεξάρτητες διότι αφορούν το ίδιο άτομο του πληθυσμού, και έτσι γενικά τα $\text{Cov}(X_i, X_j)$ δεν είναι 0. Θεωρώντας k τυχαία διανύσματα ως στήλες ενός πίνακα ($p \times k$), θα έχουμε έναν πίνακα της μορφής

$$\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_k] = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & & X_{2k} \\ \vdots & & \ddots & \vdots \\ X_{p1} & X_{p2} & \dots & X_{pk} \end{bmatrix},$$

ο οποίος θα έχει ως στοιχεία του τυχαίες μεταβλητές. Έτσι, ένας πίνακας $\mathbf{X}$ (ή διάνυσμα) καλείται τυχαίος αν τα στοιχεία του είναι τυχαίες μεταβλητές, δηλαδή $\mathbf{X} = [X_{ij}]_{k \times k}$. Ορίζουμε τη μέση τιμή πίνακα (ή διανύσματος):

$$E(\mathbf{X}) = \begin{bmatrix} E(X_{11}) & E(X_{12}) & \dots & E(X_{1n}) \\ E(X_{21}) & E(X_{22}) & & E(X_{2k}) \\ \vdots & & \ddots & \vdots \\ E(X_{p1}) & E(X_{p2}) & \dots & E(X_{pk}) \end{bmatrix}$$

Προκύπτει πολύ εύκολα η επόμενη πρόταση.

Πρόταση 3.1. Αν $\mathbf{X}, \mathbf{Y}$ είναι τυχαίοι πίνακες και $\mathbf{A}, \mathbf{B}$ είναι πίνακες με στοιχεία πραγματικούς αριθμούς τότε

$$E(\mathbf{X} + \mathbf{Y}) = E(\mathbf{X}) + E(\mathbf{Y}), \quad E(\mathbf{A}\mathbf{X}) = \mathbf{A}E(\mathbf{X}), \quad E(\mathbf{A}\mathbf{X}\mathbf{B}) = \mathbf{A}E(\mathbf{X})\mathbf{B}.$$

Απόδειξη. Η πρώτη ισότητα είναι προφανής. Η δεύτερη προκύπτει ως εξής:

$$\begin{aligned} E(\mathbf{A}\mathbf{X}) &= E \left[\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & & a_{2p} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1k} \\ X_{21} & X_{22} & & X_{2k} \\ \vdots & & \ddots & \vdots \\ X_{p1} & X_{p2} & \dots & X_{pk} \end{bmatrix} \right] \\ &= E \left[\begin{bmatrix} a_{11}X_{11} + a_{12}X_{21} + \dots + a_{1p}X_{p1} & \dots & a_{11}X_{1k} + a_{12}X_{2k} + \dots + a_{1p}X_{pk} \\ \vdots & & \vdots \\ a_{m1}X_{11} + a_{m2}X_{21} + \dots + a_{mp}X_{p1} & \dots & a_{m1}X_{1k} + a_{m2}X_{2k} + \dots + a_{mp}X_{pk} \end{bmatrix} \right] \\ &= \begin{bmatrix} a_{11}E(X_{11}) + \dots + a_{1p}E(X_{p1}) & \dots & a_{11}E(X_{1k}) + \dots + a_{1p}E(X_{pk}) \\ \vdots & & \vdots \\ a_{m1}E(X_{11}) + \dots + a_{mp}E(X_{p1}) & \dots & a_{m1}E(X_{1k}) + \dots + a_{mp}E(X_{pk}) \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ a_{21} & a_{22} & & a_{2p} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} E(X_{11}) & E(X_{12}) & \dots & E(X_{1k}) \\ E(X_{21}) & E(X_{22}) & & E(X_{2k}) \\ \vdots & & \ddots & \vdots \\ E(X_{p1}) & E(X_{p2}) & \dots & E(X_{pk}) \end{bmatrix} \\ &= \mathbf{A} E(\mathbf{X}) \end{aligned}$$

Παρόμοια αποδεικνύεται και η τρίτη ισότητα. $\square$

Αν τώρα $\mathbf{X}$ είναι ένα τυχαίο διάνυσμα διαστάσεων $1 \times p$, η μέση του τιμή συνήθως συμβολίζεται με $\boldsymbol{\mu}$,

$$E(\mathbf{X}) = \begin{bmatrix} E(X_1) \\ E(X_2) \\ \vdots \\ E(X_p) \end{bmatrix} \equiv \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \boldsymbol{\mu},$$

και ο πίνακας των διασπορών – συνδιακυμάνσεων του είναι ο:

$$\begin{aligned} E((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})') &= E \begin{bmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \\ \vdots \\ X_p - \mu_p \end{bmatrix} \begin{bmatrix} X_1 - \mu_1 & X_2 - \mu_2 & \cdots & X_p - \mu_p \end{bmatrix} \\ &= E \begin{bmatrix} (X_1 - \mu_1)^2 & (X_1 - \mu_1)(X_2 - \mu_2) & \cdots & (X_1 - \mu_1)(X_p - \mu_p) \\ (X_2 - \mu_2)(X_1 - \mu_1) & (X_2 - \mu_2)^2 & \cdots & (X_2 - \mu_2)(X_p - \mu_p) \\ \vdots & \vdots & \ddots & \vdots \\ (X_p - \mu_p)(X_1 - \mu_1) & (X_p - \mu_p)(X_2 - \mu_2) & \cdots & (X_p - \mu_p)^2 \end{bmatrix} \\ &= \begin{bmatrix} V(X_1) & \text{Cov}(X_1, X_2) & \cdots & \text{Cov}(X_1, X_p) \\ \text{Cov}(X_2, X_1) & V(X_2) & \cdots & \text{Cov}(X_2, X_p) \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}(X_p, X_1) & \text{Cov}(X_p, X_2) & \cdots & V(X_p) \end{bmatrix} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix} \end{aligned}$$

Ο παραπάνω πίνακας συμβολίζεται με $\text{Cov}(\mathbf{X})$ ή με $\boldsymbol{\Sigma}$. Δηλαδή,

$$\text{Cov}(\mathbf{X}) = E((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})').$$

Ο παραπάνω πίνακας καλείται και *πίνακας των (πληθυσμιακών) διασπορών – συνδιακυμάνσεων* $\boldsymbol{\Sigma}$. Όμοια ορίζεται και ο πίνακας των (πληθυσμιακών) *συσχετίσεων*

$$\boldsymbol{\rho} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1p} \\ \rho_{21} & 1 & \cdots & \rho_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p1} & \rho_{p2} & \cdots & 1 \end{bmatrix}, \quad \rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_{ii}} \sqrt{\sigma_{jj}}}$$

Επίσης, αν συμβολίζουμε με $\mathbf{V}$ τον διαγώνιο πίνακα που στην κύρια διαγώνιο έχει τις διασπορές σ_{ii} , $i = 1, 2, \dots, p$ τότε είναι εύκολο να επαληθεύσουμε ότι

$$\boldsymbol{\rho} = \mathbf{V}^{-1/2} \boldsymbol{\Sigma} \mathbf{V}^{-1/2}, \quad \boldsymbol{\Sigma} = \mathbf{V}^{1/2} \boldsymbol{\rho} \mathbf{V}^{1/2}.$$

Ισχύουν οι παρακάτω ιδιότητες:

Πρόταση 3.2. Αν $\mathbf{X}$ είναι τυχαίο διάνυσμα και $\mathbf{c}$ ένα διάνυσμα πραγματικών αριθμών, τότε

$$E(\mathbf{c}'\mathbf{X}) = \mathbf{c}'E(\mathbf{X}) = \mathbf{c}'\boldsymbol{\mu}, \quad V(\mathbf{c}'\mathbf{X}) = \mathbf{c}'\text{Cov}(\mathbf{X})\mathbf{c} = \mathbf{c}'\boldsymbol{\Sigma}\mathbf{c}$$

Απόδειξη. Η πρώτη ισότητα είναι προφανής, η δεύτερη ισχύει διότι

$$V(\mathbf{c}'\mathbf{X}) = V \begin{bmatrix} c_1 & c_2 & \cdots & c_k \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_k \end{bmatrix} = V \left[\sum_{i=1}^k c_i X_i \right] = \text{Cov} \left(\sum_{i=1}^k c_i X_i, \sum_{j=1}^k c_j X_j \right)$$

$$= \sum_{i=1}^k \sum_{j=1}^k c_i c_j \text{Cov}(X_i, X_j) = \begin{bmatrix} c_1 & \cdots & c_k \end{bmatrix} \begin{bmatrix} \text{Cov}(X_1, X_1) & \cdots & \text{Cov}(X_1, X_k) \\ \vdots & & \vdots \\ \text{Cov}(X_k, X_1) & \cdots & \text{Cov}(X_k, X_k) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_k \end{bmatrix}$$

$$= \mathbf{c}' \text{Cov}(\mathbf{X}) \mathbf{c}. \quad \square$$

Από το παραπάνω προκύπτει ότι $\mathbf{c}' \Sigma \mathbf{c} = V(\mathbf{c}' \mathbf{X}) \geq 0$ για κάθε $\mathbf{c}$ και άρα ο Σ είναι ημιοριστικά θετικός. Γενικότερα, αν το διάνυσμα $\mathbf{c}$ αντικατασταθεί με πίνακα $\mathbf{C}$ τότε ισχύει το ακόλουθο.

Πρόταση 3.3. Αν $\mathbf{X}$ είναι τυχαίο διάνυσμα και $\mathbf{C}$ είναι ένας πίνακας πραγματικών αριθμών, τότε

$$E(\mathbf{CX}) = \mathbf{CE}(\mathbf{X}) = \mathbf{C}\boldsymbol{\mu}, \quad \text{Cov}(\mathbf{CX}) = \mathbf{CCov}(\mathbf{X})\mathbf{C}' = \mathbf{C}\Sigma\mathbf{C}'$$

Απόδειξη. Η πρώτη ισότητα προκύπτει εύκολα. Για τη δεύτερη, από τον ορισμό του πίνακα συνδιακυμάνσεων έχουμε ότι

$$\begin{aligned} \text{Cov}(\mathbf{CX}) &= E((\mathbf{CX} - E(\mathbf{CX}))(\mathbf{CX} - E(\mathbf{CX}))') = E((\mathbf{CX} - \mathbf{C}\boldsymbol{\mu})(\mathbf{CX} - \mathbf{C}\boldsymbol{\mu})') \\ &= E(\mathbf{C}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{C}(\mathbf{X} - \boldsymbol{\mu}))') = E(\mathbf{C}(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})'\mathbf{C}') = \mathbf{CE}((\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})')\mathbf{C}' \\ &= \mathbf{CCov}(\mathbf{X})\mathbf{C}'. \quad \square \end{aligned}$$

4. ΤΥΧΑΙΑ ΔΕΙΓΜΑΤΑ

Στην μονομεταβλητή στατιστική ανάλυση, ένα τυχαίο δείγμα (τ.δ.) μεγέθους n αποτελείται από τις μετρήσεις ενός μόνο χαρακτηριστικού Y (π.χ. ύψος) σε n τυχαία επιλεγμένα άτομα ενός πληθυσμού. Πιο συγκεκριμένα, θεωρούμε ότι ένα τ.δ. είναι ένα σύνολο ανεξάρτητων και ισόνομων τυχαίων μεταβλητών $Y_1, Y_2, \dots, Y_n$ (π.χ. Y_i = ύψος i -οστού ατόμου) που ακολουθούν την κατανομή F του πληθυσμού.

Στην πολυμεταβλητή στατιστική ανάλυση από κάθε άτομο καταγράφουμε περισσότερες από μία τιμές (π.χ. βάρος, ύψος, περιφέρεια μέσης). Μπορούμε επομένως να πούμε ότι τυχαίο δείγμα τώρα θα είναι ένα σύνολο ανεξάρτητων και ισόνομων τυχαίων διανυσμάτων $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ (π.χ. $\mathbf{X}_i$ = (βάρος, ύψος, περιφέρεια i ατόμου) που ακολουθούν μια πολυδιάστατη κατανομή F του πληθυσμού.

Ας υποθέσουμε ότι οι $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ έχουν την ίδια κατανομή με ένα τυχαίο διάνυσμα

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_p \end{bmatrix}$$

το οποίο ακολουθεί πολυδιάστατη κατανομή F , δηλαδή

$$P(\mathbf{X} \leq \mathbf{x}) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_p \leq x_p) = F(x_1, x_2, \dots, x_p) = F(\mathbf{x}), \quad \mathbf{x} \in R^p,$$

και ας συμβολίσουμε με $\boldsymbol{\mu}$ και Σ τη μέση τιμή και τον πίνακα συνδιασπορών του $\mathbf{X}$. Όπως και στην μονομεταβλητή περίπτωση, μας ενδιαφέρει να εκτιμήσουμε το μέσο $\boldsymbol{\mu}$ του πληθυσμού από το τυχαίο (πολυμεταβλητό) δείγμα $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$. Θα χρησιμοποιήσουμε τον δειγματικό μέσο των

$$\mathbf{X}_1 = \begin{bmatrix} X_{11} \\ \vdots \\ X_{p1} \end{bmatrix}, \mathbf{X}_2 = \begin{bmatrix} X_{12} \\ \vdots \\ X_{p2} \end{bmatrix}, \dots, \mathbf{X}_n = \begin{bmatrix} X_{1n} \\ \vdots \\ X_{pn} \end{bmatrix},$$

δηλαδή τον

$$\bar{\mathbf{X}} = \frac{1}{n}(\mathbf{X}_1 + \dots + \mathbf{X}_n) = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n X_{1i} \\ \vdots \\ \frac{1}{n} \sum_{i=1}^n X_{pi} \end{bmatrix}.$$

Θα αποδείξουμε την ακόλουθη πρόταση.

Πρόταση 4.1. Ο δειγματικός μέσος $\bar{\mathbf{X}}$ είναι αμερόληπτη εκτιμήτρια του πληθυσμιακού μέσου $\boldsymbol{\mu}$ με πίνακα συνδιασπορών $(1/n)\boldsymbol{\Sigma}$.

Απόδειξη. Θα έχουμε ότι

$$E(\bar{\mathbf{X}}) = \frac{1}{n}(E(\mathbf{X}_1) + \dots + E(\mathbf{X}_p)) = \frac{1}{n}(\boldsymbol{\mu} + \dots + \boldsymbol{\mu}) = \frac{n\boldsymbol{\mu}}{n} = \boldsymbol{\mu},$$

και

$$\begin{aligned} \text{Cov}(\bar{\mathbf{X}}) &= E((\bar{\mathbf{X}} - E(\bar{\mathbf{X}}))(\bar{\mathbf{X}} - E(\bar{\mathbf{X}}))') = E((\bar{\mathbf{X}} - \boldsymbol{\mu})(\bar{\mathbf{X}} - \boldsymbol{\mu})') = E\left(\left(\frac{1}{n} \sum_{i=1}^n \mathbf{X}_i - \boldsymbol{\mu}\right)\left(\frac{1}{n} \sum_{j=1}^n \mathbf{X}_j - \boldsymbol{\mu}\right)'\right) \\ &= \frac{1}{n^2} E\left(\sum_{i=1}^n (\mathbf{X}_i - \boldsymbol{\mu}) \sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})'\right) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E((\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})') \end{aligned}$$

Στο παραπάνω όμως άθροισμα, παρατηρούμε ότι όταν $i \neq j$, η μέση τιμή

$$E((\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})') = \mathbf{0},$$

διότι ο παραπάνω πίνακας αποτελείται από συνδιακυμάνσεις μεταξύ των συντεταγμένων της $\mathbf{X}_i$ και της $\mathbf{X}_j$, οι οποίες όμως είναι ανεξάρτητες. Άρα τελικά,

$$\text{Cov}(\bar{\mathbf{X}}) = \frac{1}{n^2} \sum_{i=1}^n E((\mathbf{X}_i - \boldsymbol{\mu})(\mathbf{X}_i - \boldsymbol{\mu})') = \frac{1}{n^2} \sum_{i=1}^n \text{Cov}(\mathbf{X}_i) = \frac{1}{n^2} n\boldsymbol{\Sigma} = \frac{\boldsymbol{\Sigma}}{n}. \quad \square$$

Το ακόλουθο λήμμα αποτελεί γενίκευση της γνωστής σχέσης, $E(Y^2) = V(Y) + E(Y)^2$, που ισχύει στην μονομεταβλητή περίπτωση.

Πρόταση 4.2. Αν $\mathbf{Y}$ είναι ένα τυχαίο διάνυσμα με μέση τιμή $\boldsymbol{\mu}$ τότε

$$E(\mathbf{Y}\mathbf{Y}') = \text{Cov}(\mathbf{Y}) + \boldsymbol{\mu}\boldsymbol{\mu}'.$$

Απόδειξη. Ισχύει ότι

$$\begin{aligned} E(\mathbf{Y}\mathbf{Y}') &= E\left(\begin{bmatrix} Y_1 \\ \vdots \\ Y_p \end{bmatrix} \begin{bmatrix} Y_1 & \dots & Y_p \end{bmatrix}\right) = \begin{bmatrix} E(Y_1 Y_1) & \dots & E(Y_1 Y_p) \\ \vdots & \ddots & \vdots \\ E(Y_p Y_1) & \dots & E(Y_p Y_p) \end{bmatrix} \\ &= \begin{bmatrix} V(Y_1) + \mu_1^2 & \dots & \text{Cov}(Y_1, Y_p) + \mu_1 \mu_p \\ \vdots & \ddots & \vdots \\ \text{Cov}(Y_p, Y_1) + \mu_p \mu_1 & \dots & V(Y_p) + \mu_p^2 \end{bmatrix} = \text{Cov}(\mathbf{Y}) + \boldsymbol{\mu}\boldsymbol{\mu}' \end{aligned}$$

□

Για τον πίνακα $\mathbf{S}$ των δειγματικών συνδιασπορών, ισχύει η ακόλουθη πρόταση.

Πρόταση 4.3. Ο πίνακας $n/(n-1)\mathbf{S}$ είναι αμερόληπτος εκτιμητής του πίνακα συνδιασπορών Σ , δηλαδή,

$$E\left(\frac{n}{n-1}\mathbf{S}\right) = \Sigma, \quad \frac{n}{n-1}\mathbf{S} = \frac{1}{n-1} \left[\sum_{i=1}^n (X_{ki} - \bar{X}_k)(X_{ji} - \bar{X}_j) \right]_{kj}$$

Απόδειξη. Παρατηρούμε ότι

$$\begin{aligned} n\mathbf{S} &= \begin{bmatrix} \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{1i} - \bar{X}_1) & \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) & \cdots & \sum_{i=1}^n (X_{1i} - \bar{X}_1)(X_{pi} - \bar{X}_p) \\ \sum_{i=1}^n (X_{2i} - \bar{X}_2)(X_{1i} - \bar{X}_1) & \sum_{i=1}^n (X_{2i} - \bar{X}_2)(X_{2i} - \bar{X}_2) & & \\ & & \ddots & \vdots \\ \sum_{i=1}^n (X_{pi} - \bar{X}_p)(X_{1i} - \bar{X}_1) & & & \sum_{i=1}^n (X_{pi} - \bar{X}_p)(X_{pi} - \bar{X}_p) \end{bmatrix} \\ &= \sum_{i=1}^n \begin{bmatrix} X_{1i} - \bar{X}_1 \\ X_{2i} - \bar{X}_2 \\ \vdots \\ X_{pi} - \bar{X}_p \end{bmatrix} \begin{bmatrix} X_{1i} - \bar{X}_1 & X_{2i} - \bar{X}_2 & \cdots & X_{pi} - \bar{X}_p \end{bmatrix} = \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})' \\ &= \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})\mathbf{X}_i' - \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})\bar{\mathbf{X}}' = \sum_{i=1}^n (\mathbf{X}_i\mathbf{X}_i' - \bar{\mathbf{X}}\mathbf{X}_i') - \mathbf{0}\bar{\mathbf{X}}' = \sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i' - n\bar{\mathbf{X}}\bar{\mathbf{X}}'. \end{aligned}$$

Επομένως,

$$\begin{aligned} E(\mathbf{S}) &= \frac{1}{n} \sum_{i=1}^n E(\mathbf{X}_i\mathbf{X}_i') - E(\bar{\mathbf{X}}\bar{\mathbf{X}}') = \frac{1}{n} \sum_{i=1}^n (\text{Cov}(\mathbf{X}_i) + \boldsymbol{\mu}\boldsymbol{\mu}') - (\text{Cov}(\bar{\mathbf{X}}) + \boldsymbol{\mu}\boldsymbol{\mu}') \\ &= \frac{1}{n} \sum_{i=1}^n (\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}') - \left(\frac{1}{n}\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}'\right) = \Sigma + \boldsymbol{\mu}\boldsymbol{\mu}' - \frac{1}{n}\Sigma - \boldsymbol{\mu}\boldsymbol{\mu}' = \frac{n-1}{n}\Sigma. \quad \square \end{aligned}$$

5. ΠΟΛΥΔΙΑΣΤΑΤΗ ΚΑΝΟΝΙΚΗ ΚΑΤΑΝΟΜΗ

Παραπάνω υποθέσαμε ότι ένα τυχαίο διάνυσμα $\mathbf{X} = [X_1, X_2, \dots, X_p]'$ ακολουθεί μια πολυδιάστατη κατανομή $F(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^p$, και αποδείξαμε κάποια γενικά αποτελέσματα για τον δειγματικό μέσο $\bar{\mathbf{X}}$ και το δειγματικό πίνακα συνδιασπορών $\mathbf{S}$. Για να προχωρήσουμε όμως σε περισσότερα και πιο εξειδικευμένα αποτελέσματα και μοντέλα θα πρέπει να γνωρίζουμε και την μορφή της F . Επειδή στα περισσότερα πρακτικά προβλήματα οι τ.μ. X_i ή οι δειγματικοί μέσοι μπορεί να θεωρηθεί ότι ακολουθούν προσεγγιστικά κανονική κατανομή (λόγω π.χ. του Κεντρικού Οριακού Θεωρήματος), είναι εύλογο να ξεκινήσουμε γενικεύοντας την έννοια της κανονικής κατανομής σε περισσότερες από μια διαστάσεις. Πιο συγκεκριμένα θα αναζητήσουμε μια πολυδιάστατη κατανομή F , έτσι ώστε όταν $\mathbf{X} \sim F$, οι συντεταγμένες της $X_1, X_2, \dots, X_p$ θα ακολουθούν κανονική κατανομή, και δεν θα είναι κατ' ανάγκη ανεξάρτητες (δηλαδή γενικά $\text{Cov}(X_i, X_j) \neq 0$). Ας συμβολίσουμε με $\boldsymbol{\mu}$ το διάνυσμα των μέσων τιμών και με $\Sigma = [\sigma_{ij}]_{p \times p}$ τον πίνακα συνδιασπορών αυτής της F . Ο πίνακας Σ υποθέτουμε ότι είναι συμμετρικός με (γνήσια) θετικές ιδιοτιμές.

5.1. Πολυδιάστατη τυπική κανονική κατανομή.

Ας αρχίσουμε από την απλούστερη περίπτωση που οι συντεταγμένες του τυχαίου διανύσματος είναι ανεξάρτητες κανονικές $N(0,1)$. Αν σε αυτή την περίπτωση συμβολίσουμε το τ.δ. με $\mathbf{Z} = [Z_1, Z_2, \dots, Z_p]$, είναι προφανές ότι

$$F_{\mathbf{Z}}(\mathbf{z}) = P(\mathbf{Z} \leq \mathbf{z}) = P(Z_1 \leq z_1, Z_2 \leq z_2, \dots, Z_p \leq z_p) = P(Z_1 \leq z_1)P(Z_2 \leq z_2) \dots P(Z_p \leq z_p) \\ = \Phi(z_1)\Phi(z_2) \dots \Phi(z_p), \mathbf{z} \in \mathbb{R}.$$

Και η αντίστοιχη συνάρτηση πυκνότητας πιθανότητας θα είναι:

$$f_{\mathbf{Z}}(\mathbf{z}) = \frac{\partial^p}{\partial z_1 \dots \partial z_p} F_{\mathbf{Z}}(\mathbf{z}) = \frac{\partial}{\partial z_1} \Phi(z_1) \frac{\partial}{\partial z_2} \Phi(z_2) \dots \frac{\partial}{\partial z_p} \Phi(z_p) = \varphi(z_1)\varphi(z_2) \dots \varphi(z_p), \mathbf{z} \in \mathbb{R}.$$

Ως συνήθως, με Φ και φ συμβολίζουμε την συνάρτηση κατανομής και συνάρτηση πυκνότητας πιθανότητας της τυπικής κανονικής κατανομής,

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad \Phi(z) = \int_{-\infty}^z \varphi(t) dt, z \in \mathbb{R}$$

Επομένως,

$$f_{\mathbf{Z}}(\mathbf{z}) = \varphi(z_1)\varphi(z_2) \dots \varphi(z_p) = \frac{1}{(2\pi)^{p/2}} \exp\left(-\frac{z_1^2 + z_2^2 + \dots + z_p^2}{2}\right) = \frac{1}{(2\pi)^{p/2}} \exp\left(-\frac{\mathbf{z}'\mathbf{z}}{2}\right).$$

Αν μια τ.μ. $\mathbf{Z}$ ακολουθεί την παραπάνω $f_{\mathbf{Z}}$ τότε,

$$E(\mathbf{Z}) = \mathbf{0} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{Cov}(\mathbf{Z}) = \mathbf{I}_p = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix}$$

διότι, $E(Z_i) = 0$, $\text{Cov}(Z_i, Z_j) = 0$, $i \neq j$, $V(Z_i) = 1$. Καλούμε την παραπάνω κατανομή $f_{\mathbf{Z}}$, *πολυδιάστατη τυπική κανονική κατανομή*.

5.2. Πολυδιάστατη κανονική κατανομή $N_p(\mu, \Sigma)$

Θα επιχειρήσουμε τώρα να γενικεύσουμε την πολυδιάστατη τυπική κανονική κατανομή. Αν $\mathbf{Z} \sim$ πολυδιάστατη τυπική κανονική κατανομή, ορίζουμε τη νέα τ.μ.

$$\mathbf{X} = \mu + \Sigma^{1/2} \mathbf{Z} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_p \end{bmatrix} + \begin{bmatrix} \sigma_{11} & \dots & \sigma_{1p} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} & \dots & \sigma_{pp} \end{bmatrix}^{1/2} \cdot \begin{bmatrix} Z_1 \\ \vdots \\ Z_p \end{bmatrix},$$

όπου $\Sigma^{1/2}$ είναι η ρίζα του πίνακα Σ (βλ. προηγ. παράγραφο). Θα ισχύει ότι

$$E(\mathbf{X}) = \mu + \Sigma^{1/2} E(\mathbf{Z}) = \mu + \Sigma^{1/2} \cdot \mathbf{0} = \mu$$

και

$$\text{Cov}(\mathbf{X}) = E((\mathbf{X} - \mu)(\mathbf{X} - \mu)') = E((\Sigma^{1/2} \mathbf{Z})(\Sigma^{1/2} \mathbf{Z})') = E(\Sigma^{1/2} \mathbf{Z} \mathbf{Z}' (\Sigma^{1/2})') = \Sigma^{1/2} E(\mathbf{Z} \mathbf{Z}') \Sigma^{1/2} \\ = \Sigma^{1/2} \mathbf{I}_p \Sigma^{1/2} = \Sigma.$$

Σημειώνουμε ότι στα παραπάνω μπορούμε να θεωρήσουμε τον πίνακα $\Sigma^{1/2}$ διότι έχουμε υποθέσει ότι ο Σ είναι συμμετρικός με (γνήσια) θετικές ιδιοτιμές (δηλαδή είναι οριστικά θετικός).

Ας αναζητήσουμε τώρα την πολυδιάστατη σ.π.π. $f_{\mathbf{X}}$ της $\mathbf{X}$. Αυτό μπορεί να γίνει χρησιμοποιώντας το ακόλουθο γνωστό (γενικό) αποτέλεσμα:

Λήμμα 5.1. Αν ένα τ.δ. $\mathbf{Z} = [Z_1, Z_2, \dots, Z_p]'$ έχει σ.π.π. $f_{\mathbf{Z}}(\mathbf{z}) > 0$ για $\mathbf{z} \in A \subseteq \mathbb{R}^p$, και ο μετασχηματισμός $\mathbf{x} = \mathbf{g}(\mathbf{z})$ ($\mathbf{g}: A \rightarrow B \subseteq \mathbb{R}^p$) έχει αντίστροφο μετασχηματισμό τον $\mathbf{z} = \mathbf{h}(\mathbf{x})$ ($\mathbf{h} = \mathbf{g}^{-1}: B \rightarrow A$), τότε το τ.δ. $\mathbf{X} = \mathbf{g}(\mathbf{Z})$ έχει σ.π.π.

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{Z}}(\mathbf{h}(\mathbf{x})) \cdot \left| \frac{d\mathbf{h}}{d\mathbf{x}} \right|$$

όπου $|d\mathbf{h}/d\mathbf{x}|$ είναι το απόλυτο της ορίζουσας του $(p \times p)$ πίνακα *Jacobi* που ως i, j στοιχείο έχει την μερική παράγωγο $\partial h_i / \partial x_j$.

Θα εφαρμόσουμε το παραπάνω λήμμα στην περίπτωση που εξετάζουμε. Θα είναι

$$\mathbf{x} = \boldsymbol{\mu} + \Sigma^{1/2} \mathbf{z} = \mathbf{g}(\mathbf{z}) \Rightarrow \Sigma^{-1/2} \mathbf{x} = \Sigma^{-1/2} \boldsymbol{\mu} + \Sigma^{-1/2} \Sigma^{1/2} \mathbf{z} \Rightarrow \mathbf{z} = \Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu}) = \mathbf{h}(\mathbf{x})$$

και άρα

$$f_{\mathbf{X}}(\mathbf{x}) = f_{\mathbf{Z}}(\Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu})) \cdot \left| \frac{d}{d\mathbf{x}} \Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu}) \right|$$

Αλλά, αν $\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_p$ οι γραμμές του $\Sigma^{-1/2} = [b_{ij}]_{p \times p}$ τότε

$$\Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{b}'_1 \\ \vdots \\ \mathbf{b}'_p \end{bmatrix} (\mathbf{x} - \boldsymbol{\mu}) = \begin{bmatrix} \mathbf{b}'_1 (\mathbf{x} - \boldsymbol{\mu}) \\ \vdots \\ \mathbf{b}'_p (\mathbf{x} - \boldsymbol{\mu}) \end{bmatrix}$$

και άρα

$$\left| \frac{d}{d\mathbf{x}} \Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu}) \right| = \begin{vmatrix} \frac{\partial}{\partial x_1} \mathbf{b}'_1 (\mathbf{x} - \boldsymbol{\mu}) & \cdots & \frac{\partial}{\partial x_p} \mathbf{b}'_1 (\mathbf{x} - \boldsymbol{\mu}) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial x_1} \mathbf{b}'_p (\mathbf{x} - \boldsymbol{\mu}) & \cdots & \frac{\partial}{\partial x_p} \mathbf{b}'_p (\mathbf{x} - \boldsymbol{\mu}) \end{vmatrix} = \begin{vmatrix} b_{11} & \cdots & b_{1p} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pp} \end{vmatrix} = |\Sigma^{-1/2}| = |\Sigma|^{-1/2}$$

Τελικά λαμβάνουμε ότι

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &= f_{\mathbf{Z}}(\Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu})) |\Sigma|^{-1/2} = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left(-\frac{(\Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu}))' \Sigma^{-1/2} (\mathbf{x} - \boldsymbol{\mu})}{2}\right) \\ &= \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} \exp\left(-\frac{(\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}{2}\right), \mathbf{x} \in \mathbb{R}^p. \end{aligned}$$

Επομένως προκύπτει ο ακόλουθος ορισμός.

Ορισμός. Η p -διάστατη κατανομή με σ.π.π.

$$f(x_1, x_2, \dots, x_p) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})' \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})}, \mathbf{x} \in \mathbb{R}^p$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}, \boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & & \sigma_{2p} \\ \vdots & & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

όπου $\boldsymbol{\Sigma}$ είναι ένας συμμετρικός πίνακας (οριστικά θετικός) και $|\boldsymbol{\Sigma}|$ η ορίζουσά του, θα καλείται p -διάστατη κανονική και θα συμβολίζεται με $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

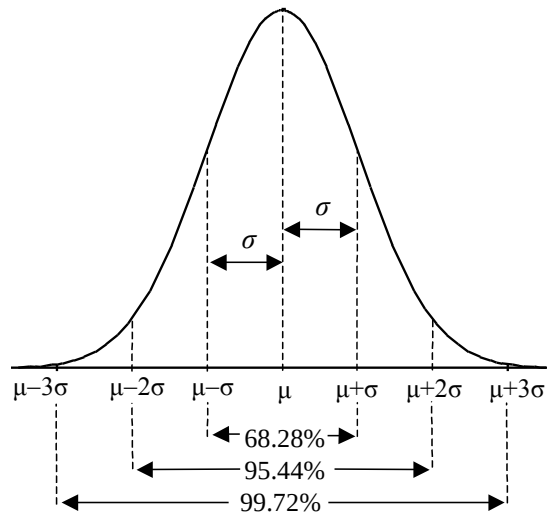
Όπως είδαμε παραπάνω, αν $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, τότε $E(\mathbf{X}) = \boldsymbol{\mu}$, $\text{Cov}(\mathbf{X}) = \boldsymbol{\Sigma}$. Αν $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \mathbf{I}_p$, τότε το τ.δ. $\mathbf{X}$ ακολουθεί την p -διάστατη τυπική κανονική κατανομή ($N_p(\mathbf{0}, \mathbf{I}_p)$).

Επίσης, είναι εύκολο να δούμε ότι κάθε μια από τις συντεταγμένες $X_1, X_2, \dots, X_p$ του τ.δ. $\mathbf{X}$ ακολουθεί κανονική κατανομή (ως γραμμικός συνδυασμός των ανεξάρτητων κανονικών τ.μ. $Z_1, Z_2, \dots, Z_p$) με μέσους $\mu_1, \mu_2, \dots, \mu_p$, διασπορές $\sigma_{11}, \sigma_{22}, \dots, \sigma_{pp}$ και συνδιασπορές $\text{Cov}(X_i, X_j) = \sigma_{ij}$.

Ας μελετήσουμε τώρα την μορφή της σ.π.π. της $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Στην τετριμμένη περίπτωση που $p = 1$, με $\boldsymbol{\mu} = [\mu]$, $\boldsymbol{\Sigma} = [\sigma^2]$, αυτή η σ.π.π. είναι η

$$f(x) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left(-\frac{(\mathbf{x}-\boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu})}{2}\right) = \frac{\exp\left(-\frac{1}{2}(x-\mu)\sigma^{-2}(x-\mu)\right)}{(2\pi)^{1/2} \sigma} = \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

Δηλαδή η μονοδιάστατη κανονική $N(\mu, \sigma^2)$, όπως ήταν αναμενόμενο. Είναι γνωστό ότι η σ.π.π. της $N(\mu, \sigma^2)$ έχει τη μορφή



Ας περάσουμε τώρα στην περίπτωση που $p = 2$. Θα ισχύει ότι

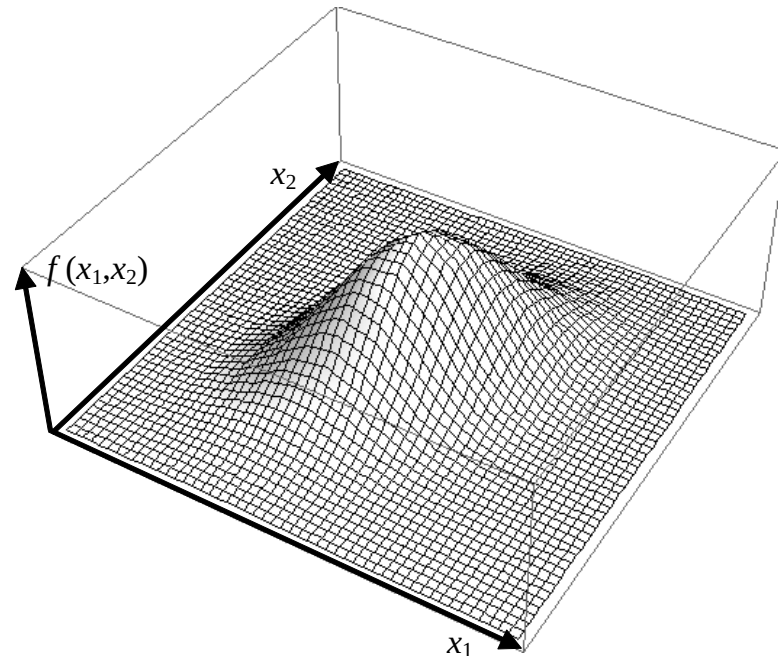
$$f(x_1, x_2) = \frac{1}{2\pi \begin{vmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{vmatrix}^{1/2}} \exp\left(-\frac{1}{2} \begin{bmatrix} x_1 - \mu_1 & x_2 - \mu_2 \end{bmatrix} \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix}^{-1} \begin{bmatrix} x_1 - \mu_1 \\ x_2 - \mu_2 \end{bmatrix}\right),$$

και μετά από αρκετές πράξεις καταλήγουμε στον τύπο

$$f(x_1, x_2) = \frac{1}{2\pi \sqrt{\sigma_{11}\sigma_{22}(1-\rho_{12}^2)}} e^{-\frac{1}{2(1-\rho_{12}^2)} \left[\left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \right)^2 + \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right)^2 - 2\rho_{12} \frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}} \frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}} \right]}$$

για $x_1, x_2 \in \mathbb{R}$, όπου $\rho_{12} = \sigma_{12} / \sqrt{\sigma_{11}\sigma_{22}}$ είναι ο συντελεστής συσχέτισης μεταξύ των X_1, X_2 .

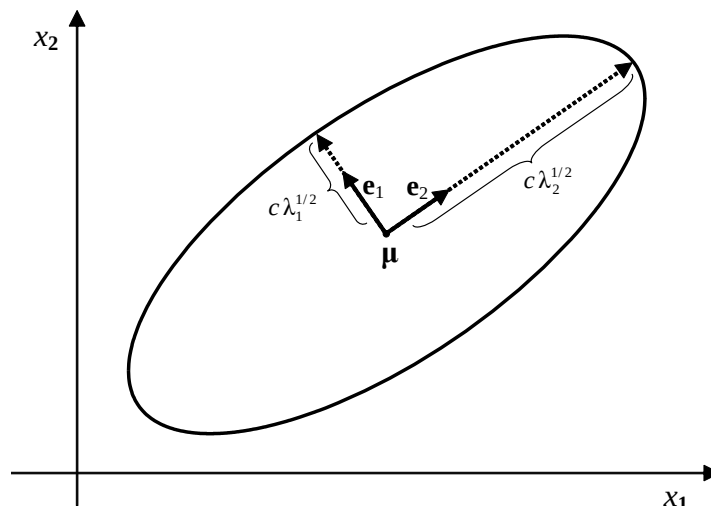
Η συγκεκριμένη σ.π.π. γίνεται πολύ πιο πολύπλοκη όταν είμαστε σε μεγαλύτερη διάσταση και επομένως είναι πολύ προτιμότερη η αναπαράσταση της μέσω πινάκων και διανυσμάτων. Όταν $p = 2$, η σ.π.π. μπορεί να παρασταθεί όπως στο ακόλουθο σχήμα (σε 3 διαστάσεις, στους δύο οριζόντιους άξονες δίνονται τα x_1, x_2 , ενώ στον κάθετο άξονα δίνεται η τιμή της $f(x_1, x_2)$).



Η σ.π.π. είναι σταθερή πάνω σε ελλείψεις και συγκεκριμένα,

$$f(\mathbf{x}) = c_0^2 \Leftrightarrow (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) = c^2 \Leftrightarrow \mathbf{y}' \boldsymbol{\Sigma}^{-1} \mathbf{y} = c^2,$$

άλλα έχουμε ήδη αναφέρει ότι τα σημεία $\mathbf{x} = (x_1, x_2)$ του επιπέδου που ικανοποιούν την παραπάνω βρίσκονται πάνω στην περιφέρεια μιας περιστραμμένης έλλειψης με κέντρο το $\boldsymbol{\mu}$, κύριους άξονες τα διανύσματα $\mathbf{e}_1, \mathbf{e}_2$ και αντίστοιχες ακτίνες τις $c\sqrt{\lambda_1}, c\sqrt{\lambda_2}$. Σχηματικά,



Με $\mathbf{e}_1, \mathbf{e}_2$ και λ_1, λ_2 συμβολίζουμε τα ιδιοδιανύσματα και τις ιδιοτιμές αντίστοιχα του πίνακα $\boldsymbol{\Sigma}$ (χρησιμοποιούμε και το γεγονός ότι αν $(\lambda_1, \mathbf{e}_1), (\lambda_2, \mathbf{e}_2)$ είναι τα ζεύγη των ιδιοτιμών – ιδιοδιανυσμάτων του πίνακα $\boldsymbol{\Sigma}$ τότε ο πίνακας $\boldsymbol{\Sigma}^{-1}$ έχει αντίστοιχα ζεύγη $(1/\lambda_1, \mathbf{e}_1), (1/\lambda_2, \mathbf{e}_2)$). Π.χ. για $p = 2$ και $\sigma_{11} = \sigma_{22}$ υπολογίζουμε ότι

$$(\lambda_1, \mathbf{e}_1) = (\sigma_{11} + \sigma_{12}, \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}), \quad (\lambda_2, \mathbf{e}_2) = (\sigma_{11} - \sigma_{12}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}),$$

Η μέγιστη τιμή (κορυφή) της $f(\mathbf{x})$ επιτυγχάνεται όταν $\mathbf{x} = \boldsymbol{\mu}$.

5.3. Βασικές προτάσεις για τ.δ. που ακολουθούν την $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$

Για την μονομεταβλητή κανονική γνωρίζουμε ότι αν $X \sim N(\mu, \sigma^2)$ τότε

$$\left(\frac{X - \mu}{\sigma} \right)^2 \sim \chi_1^2.$$

Ακολουθεί δηλαδή χι-τετράγωνο κατανομή με 1 β.ε. Ας δούμε το ανάλογο αποτέλεσμα στην πολυμεταβλητή περίπτωση.

Πρόταση 5.1. Αν $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, τότε

$$(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \sim \chi_p^2$$

Απόδειξη. Μπορούμε να θεωρήσουμε ότι $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2} \mathbf{Z}$, με $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I})$. Θα ισχύει ότι

$$(\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) = (\boldsymbol{\Sigma}^{1/2} \mathbf{Z})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\Sigma}^{1/2} \mathbf{Z}) = \mathbf{Z}' \boldsymbol{\Sigma}^{1/2} \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}^{1/2} \mathbf{Z} = \mathbf{Z}' \mathbf{Z} = Z_1^2 + Z_2^2 + \dots + Z_p^2$$

και είναι γνωστό ότι το άθροισμα των τετραγώνων p ανεξάρτητων $N(0,1)$ ακολουθεί κατανομή χι-τετράγωνο με p βαθμούς ελευθερίας. $\square$

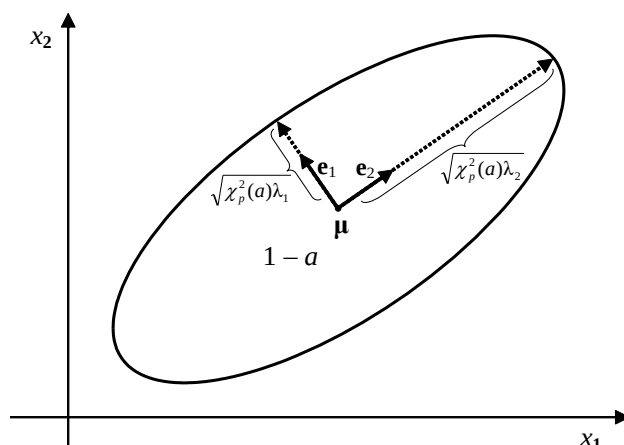
Από την παραπάνω πρόταση προκύπτει άμεσα ότι,

$$P((\mathbf{X} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{X} - \boldsymbol{\mu}) \leq \chi_p^2(a)) = 1 - a$$

όπου $\chi_p^2(a)$ είναι το άνω a -σημείο της χ_p^2 . Άρα η πολυδιάστατη κανονική κατανομή $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ κατανέμει στο σύνολο των σημείων $\mathbf{x} \in \mathbb{R}^p$ που ικανοποιούν την ανισότητα

$$(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \leq \chi_p^2(a)$$

πιθανότητα ίση με $1 - a$. Όπως έχουμε παρατηρήσει παραπάνω, πρόκειται για το εσωτερικό ελλειψοειδούς με κέντρο το $\boldsymbol{\mu}$, κύριους άξονες τα διανύσματα $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$ και ακτίνες $\sqrt{\chi_p^2(a)\lambda_1}, \sqrt{\chi_p^2(a)\lambda_2}, \dots, \sqrt{\chi_p^2(a)\lambda_p}$. Για $p = 2$ σχηματικά:



Π.χ. αν $p = 2$, και $\sigma_{11} = \sigma_{22} = 1$, $\sigma_{12} = 1/2$, $a = 0.05$, τότε βρήκαμε παραπάνω ότι

$$(\lambda_1, \mathbf{e}_1) = (\sigma_{11} + \sigma_{12}, \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}), \quad (\lambda_2, \mathbf{e}_2) = (\sigma_{11} - \sigma_{12}, \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}),$$

και επομένως, στο εσωτερικό ελλείψεως με κέντρο το $\boldsymbol{\mu}$, κύριους άξονες τα διανύσματα $\mathbf{e}_1$, $\mathbf{e}_2$ και ακτίνες

$$\sqrt{\chi_p^2(a)\lambda_1} = \sqrt{9.21 \cdot 1.5} = 3.716, \quad \sqrt{\chi_p^2(a)\lambda_2} = \sqrt{9.21 \cdot 0.5} = 2.146$$

κατανέμεται πιθανότητα $1 - a = 95\%$. Αν π.χ. $\sigma_{12} = 0$ τότε $\sqrt{\chi_p^2(a)\lambda_1} = \sqrt{\chi_p^2(a)\lambda_2} = \sqrt{9.21 \cdot 1} = 3.034$ (πρόκειται για κύκλο).

Πρόταση 5.2. Αν $\mathbf{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ τότε η τυχαία μεταβλητή

$$\mathbf{a}'\mathbf{X} = a_1 X_1 + \dots + a_p X_p \sim N(\mathbf{a}'\boldsymbol{\mu}, \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}).$$

Απόδειξη. Μπορούμε να θεωρήσουμε ότι $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{Z}$, με $\mathbf{Z} \sim N_p(\mathbf{0}, \mathbf{I})$. Ισχύει ότι

$$\mathbf{a}'\mathbf{X} = \mathbf{a}'(\boldsymbol{\mu} + \boldsymbol{\Sigma}^{1/2}\mathbf{Z}) = \mathbf{a}'\boldsymbol{\mu} + \mathbf{a}'\boldsymbol{\Sigma}^{1/2}\mathbf{Z}$$

και επομένως η τ.μ. $\mathbf{a}'\mathbf{X}$ γράφεται ως γραμμικός συνδυασμός ανεξάρτητων κανονικών τ.μ. Γνωρίζουμε όμως ότι ένας γραμμικός συνδυασμός ανεξάρτητων κανονικών τ.μ. ακολουθεί και πάλι κανονική κατανομή. Η μέση τιμή και η διασπορά της θα είναι

$$E(\mathbf{a}'\mathbf{X}) = \mathbf{a}'E(\mathbf{X}) = \mathbf{a}'\boldsymbol{\mu}, \quad V(\mathbf{a}'\mathbf{X}) = \mathbf{a}'\text{Cov}(\mathbf{X})\mathbf{a} = \mathbf{a}'\boldsymbol{\Sigma}\mathbf{a}. \quad \square$$

Επίσης γενικότερα αποδεικνύεται η ακόλουθη πρόταση.

Πρόταση 5.3. Αν $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, και $\mathbf{A}$ πίνακας $q \times p$ τότε

$$\mathbf{A}\mathbf{X} + \mathbf{d} \sim N_q(\mathbf{A}\boldsymbol{\mu} + \mathbf{d}, \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}').$$

Πριν περάσουμε σε αποτελέσματα που εμπλέκουν υποπίνακες, αξίζει να αναφέρουμε το ακόλουθο αποτέλεσμα σύμφωνα με το οποίο αν ένας πίνακας $\mathbf{A}$ ($k \times n$) απαρτίζεται από τους υποπίνακες

$$\mathbf{A}_{11} (k_1 \times n_1), \mathbf{A}_{12} (k_1 \times n_2), \mathbf{A}_{21} (k_2 \times n_1), \mathbf{A}_{22} (k_2 \times n_2)$$

ώστε $k_1 + k_2 = k$, $n_1 + n_2 = n$, και ένας πίνακας $\mathbf{B}$ ($n \times m$) απαρτίζεται από τους υποπίνακες

$$\mathbf{B}_{11} (n_1 \times m_1), \mathbf{B}_{12} (n_1 \times m_2), \mathbf{B}_{21} (n_2 \times m_1), \mathbf{B}_{22} (n_2 \times m_2)$$

ώστε $m_1 + m_2 = m$, τότε το γινόμενο $\mathbf{A}\mathbf{B}$ γράφεται

$$\mathbf{A}\mathbf{B} = \left[\begin{array}{c|c} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} \end{array} \right] \left[\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array} \right] = \left[\begin{array}{c|c} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \hline \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{array} \right],$$

όπως αντίστοιχα θα ίσχυε αν τα $\mathbf{A}_{ij}$, $\mathbf{B}_{ij}$ ήταν αριθμοί και όχι πίνακες. Η απόδειξη του παραπάνω γίνεται με τον προφανή τρόπο χρησιμοποιώντας τον ορισμό του γινομένου πινάκων. Ως ειδική περίπτωση του παραπάνω, ισχύει ότι

$$\mathbf{A}\mathbf{B} = [\mathbf{A}_1 \mid \mathbf{A}_2] \left[\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array} \right] = [\mathbf{A}_1\mathbf{B}_{11} + \mathbf{A}_2\mathbf{B}_{21} \mid \mathbf{A}_1\mathbf{B}_{12} + \mathbf{A}_2\mathbf{B}_{22}].$$

Η επόμενη πρόταση αφορά την κατανομή ενός οποιουδήποτε υποδιανύσματος ενός κανονικού τυχαίου διανύσματος.

Πρόταση 5.4. Αν $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, τότε κάθε υποδιάνυσμα $\mathbf{X}_1$ του $\mathbf{X}$ ακολουθεί πολυδιάστατη κανονική κατανομή. Αν (χωρίς βλάβη της γενικότητας), το $\mathbf{X}_1$ αποτελείται από τις πρώτες q συντεταγμένες του $\mathbf{X}$, και

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_q \\ X_{q+1} \\ \vdots \\ X_p \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_q \\ \mu_{q+1} \\ \vdots \\ \mu_p \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12} \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22} \end{bmatrix}$$

τότε το τυχαίο διάνυσμα $\mathbf{X}_1 \sim N_q(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$ ($\boldsymbol{\Sigma}_{11}$ = άνω αριστερά υποπίνακας $p \times p$ του $\boldsymbol{\Sigma}$).

Απόδειξη. Προκύπτει άμεσα από την Πρόταση 5.3. θέτοντας $\mathbf{A} = [\mathbf{I}_q \mid \mathbf{0}]$ όπου $\mathbf{I}_q$ ο μοναδιαίος $q \times q$ πίνακας, $\mathbf{0}$ ο μηδενικός $q \times (p - q)$ πίνακας, και $\mathbf{d} = \mathbf{0}$. $\square$

Πρόταση 5.5. Αν $\mathbf{X}_1 \sim N_p(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_{11})$, $\mathbf{X}_2 \sim N_q(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_{22})$ είναι ανεξάρτητα τυχαία διανύσματα τότε το τυχαίο διάνυσμα

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim N(\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix})$$

Απόδειξη. Προκύπτει εύκολα από την (από κοινού) σ.π.π. του $\mathbf{X}$ η οποία γράφεται ως γινόμενο των σ.π.π. των $\mathbf{X}_1, \mathbf{X}_2$.

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}_1, \mathbf{x}_2) &= f_{\mathbf{X}_1}(\mathbf{x}_1) f_{\mathbf{X}_2}(\mathbf{x}_2) = \frac{\exp(-\frac{1}{2}(\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1))}{(2\pi)^{p/2} |\boldsymbol{\Sigma}_{11}|^{1/2}} \frac{\exp(-\frac{1}{2}(\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2))}{(2\pi)^{q/2} |\boldsymbol{\Sigma}_{22}|^{1/2}} \\ &= \frac{\exp(-\frac{1}{2}((\mathbf{x}_1 - \boldsymbol{\mu}_1)' \boldsymbol{\Sigma}_{11}^{-1}(\mathbf{x}_1 - \boldsymbol{\mu}_1) + (\mathbf{x}_2 - \boldsymbol{\mu}_2)' \boldsymbol{\Sigma}_{22}^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)))}{(2\pi)^{(p+q)/2} |\boldsymbol{\Sigma}_{11} \boldsymbol{\Sigma}_{22}|^{1/2}} \\ &= \frac{\exp(-\frac{1}{2}[(\mathbf{x}_1 - \boldsymbol{\mu}_1)' \mid (\mathbf{x}_2 - \boldsymbol{\mu}_2)'] \begin{bmatrix} \boldsymbol{\Sigma}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 - \boldsymbol{\mu}_1 \\ \mathbf{x}_2 - \boldsymbol{\mu}_2 \end{bmatrix})}{(2\pi)^{(p+q)/2} |\boldsymbol{\Sigma}|^{1/2}} \\ &= \frac{\exp(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}))}{(2\pi)^{(p+q)/2} |\boldsymbol{\Sigma}|^{1/2}}. \quad \square \end{aligned}$$

Η παραπάνω γενικεύεται με τον προφανή τρόπο και για περισσότερα των δύο ανεξάρτητα τυχαία διανύσματα.

Η ακόλουθη πρόταση αφορά την κατανομή του αθροίσματος ανεξάρτητων κανονικών τυχαίων διανυσμάτων.

Πρόταση 5.6. Αν $\mathbf{X}_1, \mathbf{X}_2$ είναι ανεξάρτητα τ.δ. με $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma}_{ii})$, τότε

$$a_1 \mathbf{X}_1 + a_2 \mathbf{X}_2 \sim N_p(a_1 \boldsymbol{\mu}_1 + a_2 \boldsymbol{\mu}_2, a_1^2 \boldsymbol{\Sigma}_{11} + a_2^2 \boldsymbol{\Sigma}_{22})$$

Απόδειξη. Από την Πρόταση 5.5, το τυχαίο διάνυσμα

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} \sim N(\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix})$$

και επειδή

$$a_1 \mathbf{X}_1 + a_2 \mathbf{X}_2 = \begin{bmatrix} a_1 \mathbf{I}_p & a_2 \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \end{bmatrix} = \mathbf{A} \mathbf{X},$$

από την Πρόταση 5.3. προκύπτει ότι

$$\mathbf{A} \mathbf{X} \sim N_p(\mathbf{A} \boldsymbol{\mu}, \mathbf{A} \boldsymbol{\Sigma} \mathbf{A}') = N_p(a_1 \boldsymbol{\mu}_1 + a_2 \boldsymbol{\mu}_2, a_1^2 \boldsymbol{\Sigma}_{11} + a_2^2 \boldsymbol{\Sigma}_{22})$$

διότι

$$\mathbf{A} \boldsymbol{\Sigma} \mathbf{A}' = \begin{bmatrix} a_1 \mathbf{I}_p & a_2 \mathbf{I}_p \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} a_1 \mathbf{I}_p \\ a_2 \mathbf{I}_p \end{bmatrix} = a_1 \boldsymbol{\Sigma}_{11} a_1 + a_2 \boldsymbol{\Sigma}_{11} a_2 = a_1^2 \boldsymbol{\Sigma}_{11} + a_2^2 \boldsymbol{\Sigma}_{22}. \quad \square$$

Η παραπάνω γενικεύεται με τον προφανή τρόπο και για άθροισμα περισσότερων των δύο ανεξάρτητων τυχαίων διανυσμάτων. Ιδιαίτερα, για τ.δ. με τον ίδιο πίνακα διασποράς λαμβάνεται η επόμενη Πρόταση.

Πρόταση 5.7. Αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ είναι ανεξάρτητα τ.δ. με $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}_i, \boldsymbol{\Sigma})$, τότε

$$\sum_{i=1}^n a_i \mathbf{X}_i \sim N_p\left(\sum_{i=1}^n a_i \boldsymbol{\mu}_i, \sum_{i=1}^n a_i^2 \boldsymbol{\Sigma}\right).$$

5.4. Εκτιμήτριες μέγιστης πιθανοφάνειας των παραμέτρων $\boldsymbol{\mu}, \boldsymbol{\Sigma}$.

Θεωρούμε τυχαίο δείγμα από p -διάστατες παρατηρήσεις $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ που προέρχονται από την $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Ισχύει η επόμενη πρόταση.

Πρόταση 5.8. Οι εκτιμήτριες μέγιστης πιθανοφάνειας των $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ αντίστοιχα είναι

$$\hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}, \quad \hat{\boldsymbol{\Sigma}} = \mathbf{S} = [s_{ik}]_{i,k}$$

Απόδειξη (για την ε.μ.π. του $\boldsymbol{\mu}$). Η πιθανοφάνεια του δείγματος (των ανεξάρτητων τ.δ.) θα είναι

$$L(\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \prod_{i=1}^n f(\mathbf{x}_i) = \prod_{i=1}^n \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} e^{-\frac{1}{2}(\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu})} = \frac{\exp(-\frac{1}{2} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}))}{(2\pi)^{np/2} |\boldsymbol{\Sigma}|^{n/2}},$$

Η παραπάνω συνάρτηση μεγιστοποιείται ως προς $\boldsymbol{\mu}$ όταν ελαχιστοποιείται η ποσότητα στον εκθέτη,

$$g(\boldsymbol{\mu}) = \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1}(\mathbf{x}_i - \boldsymbol{\mu}).$$

Για να βρούμε το μέγιστο της παραπάνω ως προς $\boldsymbol{\mu}$, αρκεί να πάρουμε τις μερικές παραγώγους ως προς $\mu_1, \dots, \mu_p$ και να τις εξισώσουμε με το 0. Από τις ρίζες αυτής της εξίσωσης προκύπτουν τα πιθανά ελάχιστα. Ισοδύναμα, χρησιμοποιώντας συμβολισμούς πινάκων, αρκεί να πάρουμε την γενικευμένη παράγωγο (ανάδελτα) ως προς $\boldsymbol{\mu}$,

$$\frac{d}{d\boldsymbol{\mu}}[g(\boldsymbol{\mu})]' = \frac{d}{d\boldsymbol{\mu}} g(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial}{\partial \mu_1} \\ \vdots \\ \frac{\partial}{\partial \mu_p} \end{bmatrix} g(\boldsymbol{\mu}) = \begin{bmatrix} \frac{\partial}{\partial \mu_1} g(\boldsymbol{\mu}) \\ \vdots \\ \frac{\partial}{\partial \mu_p} g(\boldsymbol{\mu}) \end{bmatrix}$$

(το ανάδελτα $\nabla = \frac{d}{d\mathbf{x}} = [\frac{\partial}{\partial x_1} \quad \cdots \quad \frac{\partial}{\partial x_p}]'$ καλείται και διαφορικός διανυσματικός τελεστής) και να την εξισώσουμε με το διάνυσμα $\mathbf{0}$. Είναι πολύ εύκολο να αποδείξουμε ότι γενικά ισχύει ότι $\frac{d}{d\mathbf{x}} \mathbf{x}'\mathbf{A}\mathbf{x} = 2\mathbf{A}\mathbf{x}$ (όταν $\mathbf{A}$ συμμετρικός πίνακας) και $\frac{d}{d\mathbf{x}} \mathbf{a}'\mathbf{x} = \frac{d}{d\mathbf{x}} \mathbf{x}'\mathbf{a} = \mathbf{a}$. Επομένως

$$\begin{aligned} \frac{d}{d\boldsymbol{\mu}} g(\boldsymbol{\mu}) &= \sum_{i=1}^n \frac{d}{d\boldsymbol{\mu}} (\mathbf{x}_i - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) = \sum_{i=1}^n \frac{d}{d\boldsymbol{\mu}} (\mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \mathbf{x}_i - \mathbf{x}_i' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \mathbf{x}_i + \boldsymbol{\mu}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}), \\ &= \sum_{i=1}^n (0 - \boldsymbol{\Sigma}^{-1} \mathbf{x}_i - \boldsymbol{\Sigma}^{-1} \mathbf{x}_i + 2\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}) = -\sum_{i=1}^n 2\boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) = -2\boldsymbol{\Sigma}^{-1} \sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu}) \end{aligned}$$

και το παραπάνω θα είναι ίσο με $\mathbf{0}$ αν και μόνο αν

$$\sum_{i=1}^n (\mathbf{x}_i - \boldsymbol{\mu}) = \mathbf{0} \Leftrightarrow \sum_{i=1}^n \mathbf{x}_i = n\boldsymbol{\mu} \Leftrightarrow \hat{\boldsymbol{\mu}} = \bar{\mathbf{X}}.$$

Παρατηρούμε επίσης ότι ο πίνακας Hesse $[\frac{\partial^2}{\partial \mu_i \partial \mu_j} g(\boldsymbol{\mu})]_{i,j}$ των μερικών παραγώγων 2^{ης} τάξης γράφεται με μορφή πινάκων

$$\frac{d^2}{d\boldsymbol{\mu}^2} g(\boldsymbol{\mu}) = \frac{d}{d\boldsymbol{\mu}} \left(\frac{d}{d\boldsymbol{\mu}} g(\boldsymbol{\mu}) \right)' = \frac{d}{d\boldsymbol{\mu}} \left(-\sum_{i=1}^n 2\boldsymbol{\Sigma}^{-1} (\mathbf{x}_i - \boldsymbol{\mu}) \right)' = 2 \sum_{i=1}^n \frac{d}{d\boldsymbol{\mu}} (\boldsymbol{\mu} - \mathbf{x}_i)' \boldsymbol{\Sigma}^{-1} = 2n\boldsymbol{\Sigma}^{-1}$$

ο οποίος είναι οριστικά θετικός και επομένως στο $\bar{\mathbf{X}}$ έχουμε ελάχιστο. Άρα η ε.μ.π. του $\boldsymbol{\mu}$ θα είναι το $\bar{\mathbf{X}}$. $\square$

Για την εύρεση της ε.μ.π. του $\boldsymbol{\Sigma}$ η απόδειξη είναι πιο σύνθετη. Αποδεικνύεται ότι η πιθανοφάνεια $L(\bar{\mathbf{X}}, \boldsymbol{\Sigma})$ μεγιστοποιείται όταν

$$\hat{\boldsymbol{\Sigma}} = \mathbf{S} = \frac{1}{n} \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})' = \left[\frac{1}{n} \sum_{j=1}^n (X_{ij} - \bar{X}_i)(X_{kj} - \bar{X}_k) \right]_{i,k} = [s_{ik}]_{i,k},$$

η οποία αποτελεί και την ε.μ.π. του $\boldsymbol{\Sigma}$.

Τα παραπάνω συνάδουν και με την μονομεταβλητή περίπτωση όπου $\hat{\mu} = \bar{X}$ και $\hat{\sigma}^2 = S^2 = \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2$ (μερικές φορές ως S^2 και $\mathbf{S}$ λαμβάνονται τα προηγούμενα α-θροίσματα διά $n-1$ αντί δια n , ώστε να λαμβάνονται αμερόληπτες εκτιμήσεις των σ^2 και $\boldsymbol{\Sigma}$ αντίστοιχα)

5.5. Η πολυδιάστατη κατανομή των στατιστικών συναρτήσεων $\bar{\mathbf{X}}, \mathbf{S}$

Στην μονομεταβλητή περίπτωση γνωρίζουμε για τις κατανομές των σ.σ. $\bar{X}$ και S^2 ότι

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right), \quad \frac{(n-1)S^2}{\sigma^2} = \sum_{j=1}^n \left(\frac{X_j - \bar{X}}{\sigma}\right)^2 \sim \chi_{n-1}^2.$$

Μας ενδιαφέρει να δούμε τι ισχύει στην πολυμεταβλητή περίπτωση. Εύκολα διαπιστώνουμε ότι αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ τότε

$$\bar{\mathbf{X}} \sim N(\boldsymbol{\mu}, \frac{1}{n} \boldsymbol{\Sigma})$$

το οποίο προκύπτει άμεσα από την Πρόταση 5.7 για $a_i = 1/n$. Εξάλλου έχουμε παραπάνω αποδείξει (βλ. Πρόταση 4.1) ότι, γενικά, $E(\bar{\mathbf{X}}) = \boldsymbol{\mu}$, $\text{Cov}(\bar{\mathbf{X}}) = \boldsymbol{\Sigma}/n$.

Απομένει να δούμε τώρα ποια κατανομή ακολουθεί ο τυχαίος πίνακας $\mathbf{S}$. Για τον σκοπό αυτό αρχικά δίνουμε τον ακόλουθο ορισμό.

Ορισμός. Αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_m$ είναι ανεξάρτητα τ.δ. με $\mathbf{X}_i \sim N_p(\mathbf{0}, \boldsymbol{\Sigma})$, τότε η (πολυδιάστατη από κοινού) κατανομή του τυχαίου πίνακα

$$\sum_{i=1}^m \mathbf{X}_i \mathbf{X}_i' = \sum_{i=1}^m \begin{bmatrix} X_{1i} \\ \vdots \\ X_{pi} \end{bmatrix} \begin{bmatrix} X_{1i} & \dots & X_{pi} \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m X_{1i}^2 & \sum_{i=1}^m X_{1i} X_{2i} & \dots & \sum_{i=1}^m X_{1i} X_{pi} \\ \sum_{i=1}^m X_{2i} X_{1i} & \sum_{i=1}^m X_{2i}^2 & & \sum_{i=1}^m X_{2i} X_{pi} \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^m X_{pi} X_{1i} & \sum_{i=1}^m X_{pi} X_{2i} & \dots & \sum_{i=1}^m X_{pi}^2 \end{bmatrix}$$

καλείται **κατανομή Wishart** με m βαθμούς ελευθερίας και πίνακα $\boldsymbol{\Sigma}$ και θα συμβολίζεται με $W(m, \boldsymbol{\Sigma})$. (Αν συμβολίσουμε με $\mathbf{X} = [\mathbf{X}_1 \mid \dots \mid \mathbf{X}_m]$ τον δειγματικό $p \times m$ πίνακα τότε το παραπάνω άθροισμα απλά είναι το $\mathbf{X} \mathbf{X}'$).

Είναι εύκολο να δούμε ότι αν $\mathbf{W} \sim W(m, \boldsymbol{\Sigma})$ τότε

$$E(\mathbf{W}) = \sum_{i=1}^m E(\mathbf{X}_i \mathbf{X}_i') = \sum_{i=1}^m E((\mathbf{X}_i - \mathbf{0})(\mathbf{X}_i - \mathbf{0})') \stackrel{E(\mathbf{X}_i) = \mathbf{0}}{=} \sum_{i=1}^m \text{Cov}(\mathbf{X}_i) = m \boldsymbol{\Sigma}.$$

Αποδεικνύεται επίσης ότι η (πολυδιάστατη) συνάρτηση πυκνότητας πιθανότητας της κατανομής Wishart $W(m, \boldsymbol{\Sigma})$ έχει την μορφή

$$f_W(\mathbf{x}) = \frac{|\mathbf{x}|^{\frac{m-p-1}{2}} \exp(-\frac{1}{2} \text{tr}(\mathbf{x} \boldsymbol{\Sigma}^{-1}))}{2^{\frac{pm}{2}} \pi^{\frac{p(p-1)}{4}} |\boldsymbol{\Sigma}|^{\frac{m}{2}} \prod_{i=1}^p \Gamma(\frac{1}{2}(m-i+1))}, m > p, \quad \mathbf{x} = \begin{bmatrix} x_{11} & \dots & x_{1p} \\ \vdots & & \vdots \\ x_{p1} & \dots & x_{pp} \end{bmatrix}$$

για όλα τα $\mathbf{x}$ ώστε ο πίνακας $\mathbf{x}$ να είναι οριστικά θετικός. Αν $p = 1$ και $\boldsymbol{\Sigma} = 1$ τότε η κατανομή $W(m, \boldsymbol{\Sigma})$ είναι η κατανομή χι-τετράγωνο με m β.ε.

Από τα παραπάνω προκύπτει ότι αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, τότε

$$\sum_{j=1}^n (\mathbf{X}_j - \boldsymbol{\mu})(\mathbf{X}_j - \boldsymbol{\mu})' \sim W(n, \boldsymbol{\Sigma}).$$

Αν όμως στην θέση του $\boldsymbol{\mu}$ πάρουμε την εκτίμησή του, $\bar{\mathbf{X}}$, τότε όπως και στην μονομεταβλητή περίπτωση, χάνεται ένας β.ε. Συγκεκριμένα ισχύει η ακόλουθη πρόταση.

Πρόταση 5.9. Αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ είναι ανεξάρτητα τ.δ. με $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, τότε ο τυχαίος πίνακας

$$n\mathbf{S} = \sum_{j=1}^n (\mathbf{X}_j - \bar{\mathbf{X}})(\mathbf{X}_j - \bar{\mathbf{X}})'$$

ακολουθεί την κατανομή Wishart με $n-1$ β.ε., και επίσης τα $\bar{\mathbf{X}}, \mathbf{S}$ είναι ανεξάρτητα.

Για την κατανομή Wishart ισχύουν οι ακόλουθες ιδιότητες.

Πρόταση 5.10. (a) Αν οι τυχαίοι πίνακες $\mathbf{W}_1 \sim W(m_1, \boldsymbol{\Sigma})$, $\mathbf{W}_2 \sim W(m_2, \boldsymbol{\Sigma})$, και είναι ανεξάρτητοι, τότε $\mathbf{W}_1 + \mathbf{W}_2 \sim W(m_1 + m_2, \boldsymbol{\Sigma})$.

(b) Αν ο τυχαίος πίνακας $\mathbf{W} \sim W(m, \mathbf{\Sigma})$, τότε ο πίνακας $\mathbf{CWC}' \sim W(m, \mathbf{C}\mathbf{\Sigma}\mathbf{C}')$,

Απόδειξη. Προκύπτουν άμεσα από τον ορισμό της κατανομής Wishart. $\square$

5.6. έλεγχοι υποθέσεων για το μέσο μ .

Στην μονομεταβλητή περίπτωση είναι γνωστό ότι αν έχουμε τ.δ. $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ και θέλουμε να ελέγξουμε την υπόθεση

$$H_0: \mu = \mu_0 \text{ έναντι της } H_1: \mu \neq \mu_0,$$

χρησιμοποιούμε τη στατιστική συνάρτηση

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}, \text{ όπου } S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Όταν ισχύει η H_0 , η T ακολουθεί κατανομή student (t) με $n-1$ β.ε. ενώ όταν ισχύει η H_1 τότε λαμβάνει ακραίες τιμές (μεγάλες αρνητικές ή θετικές). Επομένως απορρίπτουμε την H_0 όταν η $|T|$ λαμβάνει μεγάλες τιμές, δηλαδή όταν $|T| > c$ (Η συγκεκριμένη περιοχή απόρριψης της H_0 προκύπτει και από το κριτήριο του γενικευμένου λόγου πιθανοφανειών). Για να έχουμε πιθανότητα σφάλματος τύπου I (λανθασμένη απόρριψη της H_0) ίση με α τότε θα πρέπει

$$P(I) = \alpha \Leftrightarrow P(|T| > c | H_0) = \alpha \Leftrightarrow c = t_{n-1}(\alpha/2).$$

Επομένως θα απορρίπτεται η $H_0: \mu = \mu_0$ σε επίπεδο σημαντικότητας (ε.σ.) α όταν

$$|T| = \frac{|\bar{X} - \mu_0|}{S/\sqrt{n}} > t_{n-1}(\alpha/2) \text{ ή ισοδύναμα, } T^2 = \frac{(\bar{X} - \mu_0)^2}{S^2/n} > t_{n-1}^2(\alpha/2)$$

Παρατηρούμε ότι η τ.μ. T^2 γράφεται και ως εξής

$$T^2 = \frac{\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}\right)^2 / 1}{\frac{(n-1)S^2}{\sigma^2} / (n-1)} = \frac{Y_1 / 1}{Y_2 / (n-1)}$$

όπου οι Y_1, Y_2 αποδεικνύεται ότι είναι ανεξάρτητες τ.μ. που ακολουθούν χι-τετράγωνο κατανομή με 1 και $n-1$ β.ε. αντίστοιχα. Επομένως η τ.μ. T^2 θα ακολουθεί κατανομή Snedecor (F ή F-Ratio) με 1 και $n-1$ β.ε. (από τον ορισμό της κατανομής F). Συνεπώς, ισοδύναμα απορρίπτουμε την H_0 σε ε.σ. α όταν

$$T^2 > F_{1,n-1}(\alpha).$$

όπου με $F_{1,n-1}(\alpha)$ συμβολίζεται το άνω α -σημείο της κατανομής F-Ratio με 1, $n-1$ β.ε (από τα παραπάνω προκύπτει και ότι $t_{n-1}^2(\alpha/2) = F_{1,n-1}(\alpha)$).

Ας εξετάσουμε τώρα πως μπορούμε να γενικεύσουμε τα προηγούμενα στην πολυμεταβλητή περίπτωση. Παρατηρούμε ότι η παραπάνω σ.σ. T^2 γράφεται

$$T^2 = (\bar{X} - \mu_0) \left(\frac{S^2}{n} \right)^{-1} (\bar{X} - \mu_0), \quad S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

Και όπως είδαμε ακολουθεί κατανομή $F_{1,n-1}$. Αν έχουμε τα ανεξάρτητα τυχαία διανύσματα $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n \sim N_p(\boldsymbol{\mu}, \mathbf{\Sigma})$ στην θέση των $X_1, X_2, \dots, X_n \sim N(\mu, \sigma^2)$ τότε η παραπάνω T^2 γενικεύεται ως εξής

$$T_p^2 = (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \left(\frac{\mathbf{S}}{n-1} \right)^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0), \quad \mathbf{S} = \frac{1}{n} \sum_{i=1}^n (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})',$$

($T^2 = T_1^2$). Η παραπάνω στατιστική συνάρτηση καλείται T^2 του Hotelling, και αποδεικνύεται η επόμενη πρόταση,

Πρόταση 5.11. Αν ισχύει η $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$, η τ.μ.

$$\frac{(n-p)}{(n-1)p} T_p^2 = \frac{n-p}{p} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) \sim F_{p, n-p},$$

(δηλαδή ακολουθεί κατανομή Snedecor (F) με $p, n-p$ β.ε.)

Αν δεν ισχύει η H_0 τότε η παραπάνω τ.μ. λαμβάνει μεγάλες θετικές τιμές. Επομένως θα απορρίπτεται η $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$, έναντι της $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, όταν η παραπάνω λαμβάνει μεγάλες τιμές, μεγαλύτερες από ένα c . Προκειμένου να επιτύχουμε $P(I) = \alpha$ θα πρέπει,

$$P\left(\frac{(n-p)}{(n-1)p} T_p^2 > c\right) = \alpha \Leftrightarrow c = F_{p, n-p}(\alpha)$$

και τελικώς θα απορρίπτεται η $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$, σε ε.σ. α όταν

$$\frac{(n-p)}{(n-1)p} T_p^2 = \frac{n-p}{p} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > F_{p, n-p}(\alpha).$$

Η παραπάνω περιοχή απόρριψης μπορεί να θεωρηθεί ως γενίκευση της μονομεταβλητής περίπτωσης (για $p = 1$ προκύπτει το γνωστό t-τεστ), αλλά προκύπτει και ανεξάρτητα χρησιμοποιώντας το κριτήριο του γενικευμένου λόγου πιθανοφαινιών. Συγκεκριμένα, σύμφωνα με το κριτήριο αυτό, η περιοχή απόρριψης της H_0 θα είναι της μορφής

$$\frac{\max_{\Sigma} \prod_{i=1}^n f_{N(\boldsymbol{\mu}_0, \Sigma)}(\mathbf{X}_i)}{\max_{\boldsymbol{\mu}, \Sigma} \prod_{i=1}^n f_{N(\boldsymbol{\mu}, \Sigma)}(\mathbf{X}_i)} < c_\alpha$$

η οποία αποδεικνύεται ότι είναι ισοδύναμη με την $(n-p)/((n-1)p) T_p^2 > c$ που χρησιμοποιήσαμε παραπάνω.

Παράδειγμα 3. Ας θεωρήσουμε και πάλι το αρχικό παράδειγμα όπου X_1 : *Περίμετρος Κοιλίας* και X_2 : *Βάρος* και X_3 : *Ύψος*, και έστω ότι θέλουμε να ελέγξουμε αν

$$H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0 = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} 90 \\ 80 \\ 180 \end{bmatrix}.$$

με $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, σε ε.σ. $\alpha = 5\%$ (εδώ $p = 3$). Από το δείγμα μεγέθους $n = 36$ έχουμε υπολογίσει

$$\bar{\mathbf{x}} = \begin{bmatrix} 88.086 \\ 81.172 \\ 180.486 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} 81.58 & 84.27 & 4.98 \\ 84.27 & 125.94 & 38.88 \\ 4.98 & 38.88 & 52.56 \end{bmatrix}$$

και επομένως, η τιμή του T^2 του Hotelling από το δείγμα θα είναι

$$\begin{aligned}
t_p^2 &= (\bar{\mathbf{x}} - \boldsymbol{\mu}_0)' \left(\frac{\mathbf{S}}{n-1} \right)^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}_0) \\
&= (36-1) \begin{bmatrix} -1.914 & 1.172 & 0.486 \end{bmatrix} \begin{bmatrix} 81.58 & 84.27 & 4.98 \\ 84.27 & 125.94 & 38.88 \\ 4.98 & 38.88 & 52.56 \end{bmatrix}^{-1} \begin{bmatrix} -1.914 \\ 1.172 \\ 0.486 \end{bmatrix}, \\
&= \dots \approx 17.4
\end{aligned}$$

και άρα

$$\frac{(n-p)}{(n-1)p} t_p^2 \approx 5.467 > F_{p,n-p}(a) \approx 2.891.$$

από όπου απορρίπτεται η H_0 σε ε.σ. 5%. Το αντίστοιχο p-value είναι

$$p\text{-value} = P\left(\frac{(n-p)}{(n-1)p} T_p^2 > 5.467 \mid H_0\right) = 1 - F_{F_{p,n-p}}(5.467) \approx 0.0036$$

το οποίο είναι αρκετά μικρό, (μικρότερο του a) και επομένως θα απορρίπταμε την H_0 ακόμη και για $a = 1\%$.

Αξίζει να παρατηρήσουμε ότι αν ο παραπάνω έλεγχος γίνονταν ξεχωριστά για κάθε μ_i , δηλαδή κάναμε τους 3 ελέγχους

- (α) $H_{0\alpha}: \mu_1 = 90, H_{1\alpha}: \mu_1 \neq 90,$
- (β) $H_{0\beta}: \mu_2 = 80, H_{1\beta}: \mu_2 \neq 80$
- (γ) $H_{0\gamma}: \mu_3 = 180, H_{1\gamma}: \mu_3 \neq 180$

τον καθένα σε ε.σ. a (μέσω τριών διαφορετικών t-test), τότε η (συνολική) πιθανότητα σφάλματος τύπου I θα ήταν μεγαλύτερη από a . Πράγματι, αν συμβολίσουμε με A_1, A_2, A_3 τα ενδεχόμενα απόρριψης των υποθέσεων (α), (β), (γ) τότε

$$\begin{aligned}
P(I) &= P(A_1 \cup A_2 \cup A_3 \mid H_{0\alpha}, H_{0\beta}, H_{0\gamma}) = 1 - P(A_1^c \cap A_2^c \cap A_3^c \mid H_{0\alpha}, H_{0\beta}, H_{0\gamma}) \\
&\approx 1 - P(A_1^c \mid H_{0\alpha}) P(A_2^c \mid H_{0\beta}) P(A_3^c \mid H_{0\gamma}) = 1 - (1-a)^3 = 3a - 3a^2 + a^3 \approx 3a > a.
\end{aligned}$$

(π.χ. για $a = 5\%$, $1 - (1-a)^3 = 14.26\%$). Σημειώνεται ότι παραπάνω προσεγγίσαμε την πιθανότητα της τομής των ενδεχομένων A_i^c , με το γινόμενο των πιθανοτήτων των ενδεχομένων A_i^c . Οι δύο αυτές ποσότητες είναι ακριβώς ίσες όταν τα A_1^c, A_2^c, A_3^c είναι ανεξάρτητα ενδεχόμενα (υπό τις H_0), γεγονός που δεν ισχύει γενικά.

Επομένως, αν θέλουμε να διατηρήσουμε την πιθανότητα σφάλματος τύπου I ίση με a δεν θα πρέπει να γίνουν ξεχωριστά οι έλεγχοι, αλλά «ταυτόχρονα», δηλαδή χρησιμοποιώντας την κρίσιμη περιοχή μέσω του T^2 του Hotelling που είδαμε παραπάνω ή να μειώσουμε το ε.σ. στα τρία διαφορετικά t-test (π.χ. αντί a παίρνουμε $a/3$ – μέθοδος Bonferroni), αλλά και πάλι δεν θα είμαστε σίγουροι για το ακριβές $P(I)$.

Κλείνοντας να αναφέρουμε ότι, με το ίδιο τρόπο θα ελέγχαμε π.χ. την $H_0: \mu_1 = 90, \mu_2 = 80$ έναντι της $H_1: \mu_1 \neq 90$ ή $\mu_2 \neq 80$, θεωρώντας μόνο τις δύο πρώτες μεταβλητές της $\mathbf{X}$, επειδή όπως γνωρίζουμε,

$$\mathbf{X}' = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim N_2(\boldsymbol{\mu}' = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \boldsymbol{\Sigma}' = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix})$$

Σε αυτή την περίπτωση απορρίπτουμε την $H_0: \boldsymbol{\mu}' = \boldsymbol{\mu}'_0$ όταν

$$\frac{(n-p)}{(n-1)p} T_p^2 \sim F_{p,n-p}$$

όπου τώρα $p = 2$ και το T^2 του Hotelling υπολογίζεται με βάση τις δύο πρώτες μεταβλητές του δείγματος.

5.7. Περιοχές εμπιστοσύνης για το μέσο μ .

Υπενθυμίζεται ότι ένα διάστημα εμπιστοσύνης για το μ συντελεστού $1 - \alpha$ στην μονομεταβλητή περίπτωση είναι

$$P\left(\bar{X} - \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2) \leq \mu \leq \bar{X} + \frac{S}{\sqrt{n}} t_{n-1}(\alpha/2)\right) = 1 - \alpha$$

το οποίο μπορεί να θεωρηθεί ότι περιλαμβάνει όλες τις τιμές του μ_0 για τις οποίες δεν απορρίπτεται η υπόθεση $H_0: \mu = \mu_0$ σε ε.σ. α .

Από το γεγονός ότι στην πολυμεταβλητή περίπτωση (με $\mathbf{X}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$) ισχύει (βλ. Πρόταση 5.11),

$$\frac{(n-p)}{(n-1)p} T_p^2 = \frac{n-p}{p} (\bar{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim F_{p, n-p},$$

θα ισχύει άμεσα ότι

$$P((\bar{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \frac{p}{n-p} F_{p, n-p}(\alpha)) = 1 - \alpha.$$

από όπου προκύπτει μια περιοχή εμπιστοσύνης για το $\boldsymbol{\mu}$ συντελεστού $1 - \alpha$. Αυτή την φορά δεν πρόκειται για διάστημα αλλά για περιοχή διότι πρόκειται για ένα υποσύνολο του R^p . Δεδομένων λοιπόν των δειγματικών $\bar{\mathbf{x}}, \mathbf{s}$, η περιοχή αυτή θα αποτελείται από όλα τα σημεία $\boldsymbol{\mu} \in R^p$ τα οποία ικανοποιούν την ανίσωση

$$(\bar{\mathbf{x}} - \boldsymbol{\mu})' \mathbf{s}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \leq \frac{p}{n-p} F_{p, n-p}(\alpha).$$

Γνωρίζουμε ότι τα σημεία που ικανοποιούν την παραπάνω βρίσκονται στο εσωτερικό ελλειψοειδών με κέντρο το $\bar{\mathbf{x}}$, κύριους άξονες τα διανύσματα $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_p$ και αντίστοιχες ακτίνες τις $c\sqrt{\lambda_1}, c\sqrt{\lambda_2}, \dots, c\sqrt{\lambda_n}$ όπου τώρα $c^2 = (p/(n-p)) F_{p, n-p}(\alpha)$. Σε αυτή την περίπτωση, με $(\lambda_i, \mathbf{e}_i)$, $i = 1, 2, \dots, p$, συμβολίζουμε τα ζεύγη ιδιοτιμών – ιδιοδιανυσμάτων του δειγματικού πίνακα $\mathbf{s}$.

Παράδειγμα 4. Συνεχίζουμε το παράδειγμα που εξετάσαμε και παραπάνω όπου X_1 : *Περίμετρος Κοιλίας* και X_2 : *Βάρος* και X_3 : *Ύψος*, και έστω ότι θέλουμε να κατασκευάσουμε περιοχή εμπιστοσύνης σ.ε. $1 - \alpha = 95\%$ για το ζεύγος (μ_1, μ_2) . Η περιοχή αυτή αποτελείται από όλα τα σημεία $\boldsymbol{\mu}$ που ικανοποιούν την

$$(\bar{\mathbf{x}} - \boldsymbol{\mu})' \mathbf{s}^{-1} (\bar{\mathbf{x}} - \boldsymbol{\mu}) \leq \frac{p}{n-p} F_{p, n-p}(\alpha)$$

όπου στα $\bar{\mathbf{x}}, \mathbf{s}$ παραπάνω έχουμε κρατήσει μόνο τα στοιχεία που αφορούν τις δύο πρώτες μεταβλητές. Η περιοχή αυτή θα αποτελείται από όλα τα (μ_1, μ_2) που ικανοποιούν την :

$$\left(\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \right)' \begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix}^{-1} \left(\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \right) \leq \frac{2}{n-2} F_{2, n-2}(0.05),$$

δηλαδή την

$$\left(\begin{bmatrix} 88.086 \\ 81.172 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \right)' \begin{bmatrix} 81.58 & 84.27 \\ 84.27 & 125.94 \end{bmatrix}^{-1} \left(\begin{bmatrix} 88.086 \\ 81.172 \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \right) \leq \frac{2}{36-2} 3.276 = 0.1927.$$

ή

$$97.589 + 0.03969 \mu_1^2 - \mu_1 (2.681 + 0.05312 \mu_2) + 0.50492 \mu_2 + 0.02571 \mu_2^2 \leq 0.1927.$$

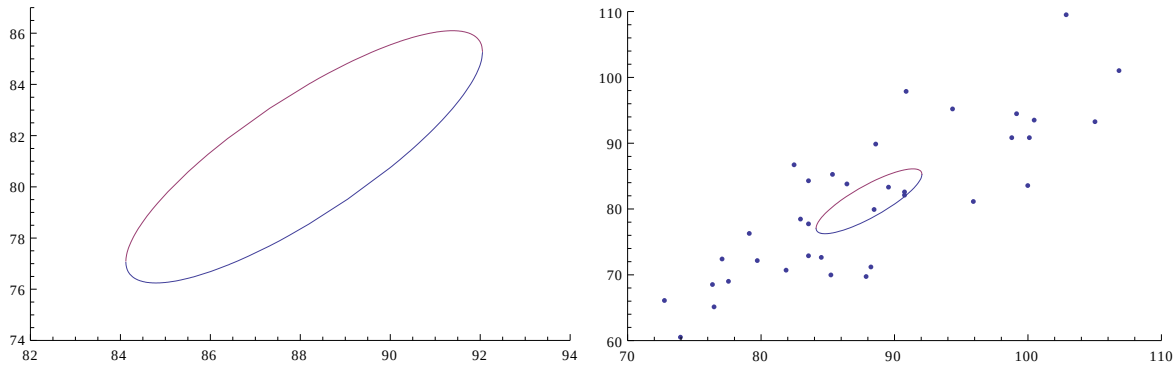
Υπολογίζεται ότι οι ιδιοτιμές – ιδιοδιανύσματα του πίνακα s είναι

$$\lambda_1 = 190.9, \mathbf{e}_1 = \begin{bmatrix} 0.6105 \\ 0.7920 \end{bmatrix}, \quad \lambda_2 = 16.62, \mathbf{e}_2 = \begin{bmatrix} -0.6105 \\ 0.7920 \end{bmatrix}$$

και επομένως το (μ_1, μ_2) βρίσκεται (με συντελεστή εμπιστοσύνης $1 - \alpha = 95\%$) στο εσωτερικό της έλλειψης με κέντρο το $\bar{\mathbf{x}} = (88.086, 81.172)$, με άξονες τα διανύσματα $(0.6105, 0.792)$, $(-0.792, 0.6105)$ και ακτίνες

$$\sqrt{0.1927 \cdot 190.9} \approx 6.065, \quad \sqrt{0.1927 \cdot 16.62} \approx 1.7896,$$

αντίστοιχα. Στο πρώτο σχήμα που ακολουθεί βλέπουμε την μορφή της περιοχής εμπιστοσύνης, ενώ στο δεύτερο βλέπουμε την ίδια περιοχή μαζί με τις $n = 36$ παρατηρήσεις.



Όπως αντίστοιχα παρατηρήσαμε και στους ελέγχους υποθέσεων για το μ , εάν κατασκευάζαμε ξεχωριστά δ.ε. (α_1, β_1) και (α_2, β_2) για τα μ_1, μ_2 αντίστοιχα, τότε το $(\alpha_1, \beta_1) \times (\alpha_2, \beta_2)$ θα περιελάμβανε το (μ_1, μ_2) με πιθανότητα μικρότερη του $1 - \alpha$ (περίπου $(1 - \alpha)^2$). Τα δ.ε. αυτά (95%) είναι

$$(\alpha_1, \beta_1) = (\bar{x}_1 - \sqrt{\frac{n}{n-1} S_{11}} t_{n-1}(\frac{\alpha}{2}), \bar{x}_1 + \sqrt{\frac{n}{n-1} S_{11}} t_{n-1}(\frac{\alpha}{2})) = (84.98, 91.18)$$

$$(\alpha_2, \beta_2) = (\bar{x}_2 - \sqrt{\frac{n}{n-1} S_{22}} t_{n-1}(\frac{\alpha}{2}), \bar{x}_2 + \sqrt{\frac{n}{n-1} S_{22}} t_{n-1}(\frac{\alpha}{2})) = (77.32, 85.02).$$

ορίζοντας ως περιοχή εμπιστοσύνης το ορθογώνιο $(\alpha_1, \beta_1) \times (\alpha_2, \beta_2)$, ενώ όπως γνωρίζουμε η πιθανότητα στην κανονική κατανομή μειώνεται όσο απομακρυνόμαστε από το μέσο πάνω σε ελλείψεις (ισοπίθανες ή ισοϋψείς καμπύλες). Συνεπώς η παραπάνω έλλειψη ως περιοχή εμπιστοσύνης συντελεστού $1 - \alpha$ είναι ορθότερη και ακριβέστερη.

5.8. Οριακά Θεωρήματα

Τα γνωστά οριακά θεωρήματα που ισχύουν στην μονομεταβλητή περίπτωση γενικεύονται σε p διαστάσεις.

Θεώρημα 5.12. (Νόμος Μεγάλων Αριθμών) Αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ είναι ανεξάρτητα και ισόνομα τ.δ. με $E(\mathbf{X}_i) = \mu$, $\text{Cov}(\mathbf{X}_i) = \Sigma < \infty$, τότε

$$(α) \bar{\mathbf{X}} \rightarrow \boldsymbol{\mu} \text{ όταν } n \rightarrow \infty \quad (β) \mathbf{S} \rightarrow \boldsymbol{\Sigma} \text{ όταν } n \rightarrow \infty$$

(κατά πιθανότητα, δηλ. $P(|\bar{\mathbf{X}} - \boldsymbol{\mu}| > \boldsymbol{\varepsilon}) \rightarrow 0$ για κάθε $\boldsymbol{\varepsilon} > \mathbf{0}$).

Αξίζει να σημειώσουμε ότι το (β) προκύπτει από το (α) παρατηρώντας ότι το ik στοιχείο του πίνακα $\mathbf{S}$ είναι

$$\begin{aligned} s_{ik} &= \frac{1}{n} \sum_{j=1}^n (X_{ij} - \bar{X}_i)(X_{kj} - \bar{X}_k) = \frac{1}{n} \sum_{j=1}^n (X_{ij} - \mu_i)(X_{kj} - \mu_k) + (\bar{X}_i - \mu_i)(\bar{X}_k - \mu_k) \\ &= \frac{1}{n} \sum_{j=1}^n Y_j + (\bar{X}_i - \mu_i)(\bar{X}_k - \mu_k) \end{aligned}$$

το οποίο (χρησιμοποιώντας το (α)) συγκλίνει στο $E(Y_j) + 0 = \sigma_{ik}$.

Θεώρημα 5.13. (Κεντρικό Οριακό Θεώρημα) Αν $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ είναι ανεξάρτητα και ισόνομα (p -διάστατα) τ.δ. με $E(\mathbf{X}_i) = \boldsymbol{\mu}$, $\text{Cov}(\mathbf{X}_i) = \boldsymbol{\Sigma} < \infty$, τότε

$$\sqrt{n}(\bar{\mathbf{X}} - \boldsymbol{\mu}) \rightarrow N_p(\mathbf{0}, \boldsymbol{\Sigma}) \text{ όταν } n \rightarrow \infty$$

(κατά κατανομή).

Χρησιμοποιώντας τα παραπάνω οριακά θεωρήματα μπορούμε να κατασκευάσουμε (προσεγγιστικούς) ελέγχους υποθέσεων για μεγάλα δείγματα, χωρίς να υποθέτουμε ότι οι αρχικές παρατηρήσεις $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ προέρχονται από κανονική κατανομή. Συγκεκριμένα έχουμε τα ακόλουθα.

Αν $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, τότε $\bar{\mathbf{X}} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$ και από την Πρόταση 5.1 (με $\bar{\mathbf{X}}$ στην θέση του $\mathbf{X}$, και $\boldsymbol{\Sigma}/n$ στη θέση του $\boldsymbol{\Sigma}$) προκύπτει ότι,

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_p^2.$$

Η παραπάνω εξακολουθεί να ισχύει (προσεγγιστικά) ακόμη και όταν τα $\mathbf{X}_i$ δεν είναι κανονικά, διότι από το Κ.Ο.Θ. προκύπτει ότι για μεγάλο n , $\bar{\mathbf{X}} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}/n)$ προσεγγιστικά.

Επίσης, για μεγάλο n , από τον Νόμο των Μεγάλων Αριθμών, $\mathbf{S} \approx \boldsymbol{\Sigma}$ και επομένως μπορούμε στις παραπάνω σχέσεις να πάρουμε $\mathbf{S}$ στη θέση του άγνωστου $\boldsymbol{\Sigma}$ και να πούμε ότι

$$n(\bar{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_p^2$$

προσεγγιστικά, για μεγάλο n ($> p$). Οι παραπάνω προσεγγιστικές εκφράσεις μπορούν να μας βοηθήσουν να κατασκευάσουμε ελέγχους και περιοχές εμπιστοσύνης για το $\boldsymbol{\mu}$.

Συγκεκριμένα, αν τα ανεξάρτητα και ισόνομα τ.δ. $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ προέρχονται από μια p -διάστατη κατανομή με (άγνωστο) μέσο $\boldsymbol{\mu}$, θα απορρίπτεται η $H_0: \boldsymbol{\mu} = \boldsymbol{\mu}_0$, έναντι της $H_1: \boldsymbol{\mu} \neq \boldsymbol{\mu}_0$, σε ε.σ. α (προσεγγιστικά για μεγάλο n) όταν

$$T_p^2 \approx n(\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > \chi_p^2(\alpha).$$

Συγκριτικά υπενθυμίζεται ότι αν $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ τότε απορρίπτουμε την H_0 σε ε.σ. α (ακριβώς) όταν

$$\frac{(n-p)}{(n-1)p} T_p^2 = \frac{n-p}{p} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0)' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}_0) > F_{p, n-p}(\alpha)$$

από όπου παρεμπιπτόντως συμπεραίνουμε ότι για μεγάλο n θα πρέπει,

$$\chi_p^2 \approx \frac{np}{n-p} F_{p, n-p}(a) \approx p F_{p, n-p}(a).$$

Τέλος, από το γεγονός ότι $n(\bar{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \sim \chi_p^2$ προσεγγιστικά, προκύπτει η α-κόλουθη προσεγγιστική περιοχή εμπιστοσύνης συντελεστού $1 - a$ για το $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} \in R^p : (\bar{\mathbf{X}} - \boldsymbol{\mu})' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \boldsymbol{\mu}) \leq \frac{1}{n} \chi_p^2(a).$$

ανεξαρτήτως της κατανομής από την οποία προέρχονται τα $\mathbf{X}_i$.

Παράδειγμα. Θεωρούμε n ερωτώμενους σχετικά με την πρόθεσή τους να ψηφίσουν ένα από k συγκεκριμένα κόμματα. Αν συμβολίσουμε με $\mathbf{X}_i$ την απάντηση του i ερωτώμενου, τότε θα είναι $\mathbf{X}_i' = (0, \dots, 0, 1, 0, \dots, 0)$ (όπου το 1 είναι στην j -θέση) αν ο ερωτώμενος απάντησε ότι θα ψηφίσει το $j \in \{1, 2, \dots, k\}$ κόμμα ($\mathbf{X}_i = \mathbf{0}$ αν δεν επιλέξει κάποιο από αυτά τα k κόμματα). Προφανώς η τ.μ. $\mathbf{X}_i$ ακολουθεί την πολυωνυμική κατανομή με παραμέτρους $m = 1$ και $\mathbf{q} = (q_1, q_2, \dots, q_k)$. Είναι εύκολο να διαπιστώσουμε ότι

$$E(\mathbf{X}_i) = \mathbf{q}, \quad \text{Cov}(\mathbf{X}_i) = \begin{bmatrix} q_1(1-q_1) & -q_1q_2 & \cdots & -q_1q_k \\ -q_2q_1 & q_2(1-q_2) & & -q_2q_k \\ \vdots & & \ddots & \vdots \\ -q_kq_1 & -q_kq_2 & \cdots & q_k(1-q_k) \end{bmatrix} = \boldsymbol{\Sigma}.$$

Άρα από ΚΟΘ, θα ισχύει για μεγάλο n προσεγγιστικά ότι

$$\sqrt{n}(\bar{\mathbf{X}} - \mathbf{q}) = \frac{1}{\sqrt{n}} \left(\sum_{i=1}^n \mathbf{X}_i - n\mathbf{q} \right) \sim N_k(\mathbf{0}, \boldsymbol{\Sigma}).$$

Το τυχαίο διάνυσμα $\sum_{i=1}^n \mathbf{X}_i$ ως γνωστό ακολουθεί πολυωνυμική κατανομή με παραμέτρους n και $\mathbf{q} = (q_1, q_2, \dots, q_k)$ (πλήθος επιτυχιών $1^{00}, 2^{00}, \dots, k$ -οστού είδους σε n δοκιμές). Από το παραπάνω προκύπτει ότι η πολυδιάστατη αυτή κατανομή ακολουθεί προσεγγιστικά (για μεγάλο n) την πολυμεταβλητή κανονική με μέση τιμή $n\mathbf{q}$ και πίνακα διασποράς $n\boldsymbol{\Sigma}$. Αν τώρα π.χ. $k = 2$ και θέλουμε να ελέγξουμε αν $q_1 = 20\%$ και $q_2 = 15\%$ τότε θα πρέπει να εργαστούμε όπως περιγράφεται παραπάνω χρησιμοποιώντας την κρίσιμη περιοχή

$$n(\bar{\mathbf{X}} - \mathbf{q}_0)' \mathbf{S}^{-1} (\bar{\mathbf{X}} - \mathbf{q}_0) > \chi_2^2(a).$$

Επίσης αν θέλουμε να εκτιμήσουμε τα ποσοστά των δύο κομμάτων τότε μπορούμε να κατασκευάσουμε μια περιοχή εμπιστοσύνης για το (q_1, q_2) η οποία θα είναι μια έλλειψη με κέντρο το $(\bar{x}_1, \bar{x}_2)$ και ακτίνες που προκύπτουν από τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα $\mathbf{S}$.