Η πολυδιάστατη κανονική κατανομή

Ορίζουμε την τυπική πολυδιάστατη κανονική, σαν την κατανομή του τυχαίου διανύσματος $X = \left(X_1, \ldots, X_n\right)^T$, όπου $X_i \sim N(0,1)$ και όλα τα X_i μεταξύ τους ανεξάρτητα. Τότε

$$
\begin{aligned}
& f_X(x) = f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = N(x_1 \mid 0,1)\cdots N(x_n \mid 0,1) \\
& = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} x_i^2\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2} x^T x\right).
\end{aligned}
$$

Εμφανώς οι περιθώριες είναι όλες τυπικές κανονικές εφόσον

$$
\begin{aligned}
& f_{X_i}(x_i) = \int_{\mathbb{R}^{n-1}} f_X(x) \prod_{j \neq i} dx_j = N(x_i \mid 0,1) \int_{\mathbb{R}^{n-1}} \prod_{j \neq i} N\left(x_j \mid 0,1\right) dx_j \\
& = N(x_i \mid 0,1) \prod_{j \neq i} \int_{\mathbb{R}} N\left(x_j \mid 0,1\right) dx_j = N(x_i \mid 0,1).
\end{aligned}
$$

Η μέση τιμή του X είναι το μηδενικό διάνυσμα $\mathcal{O}^T = (0,\ldots,0)$

$$
\mu_X = \mathbb{E}(X) = \mathbb{E}\left(\left(X_1,\ldots,X_n\right)^T\right) = \left(\mathbb{E}(X_1),\ldots,\mathbb{E}(X_n)\right)^T = (0,\ldots,0)^T = \mathcal{O}.
$$

Ορίζουμε τον πίνακα συνδιασποράς του τυχαίου διανύσματος X, σαν τον πίνακα $\mathbb{V}(X)$ (εναλλακτικός συμβολισμός Σ_X) με $ij-$στοιχείο

$$
\left(\mathbb{V}(X)\right)_{ij} = Cov\left(X_i, X_j\right) = \sigma_{ij}.
$$

Επειδή $X - \mu \in \mathbb{R}^{n\times 1}$ και $(X-\mu)^T \in \mathbb{R}^{1\times n}$, το γινόμενο (outer product) του διανύσματος $X - \mu$ με το διάνυσμα $(X-\mu)^T$, θα έχει διάσταση $n \times n$

$$
(X-\mu)(X-\mu)^T = \left((X_i - \mu_i)\left(X_j - \mu_j\right)\right) \in \mathbb{R}^{n\times n}.
$$

Παίρνοντας την μέση τιμή έχουμε:

$$
\begin{aligned}
& \mathbb{V}(X) = \mathbb{E}\left[\left(\left(X_i - \mu_{X_i}\right)\left(X_j - \mu_{X_j}\right)\right)\right] \\
& = \left(\mathbb{E}\left[(X_i - \mu_i)\left(X_j - \mu_j\right)\right]\right) = \left(Cov\left(X_i, X_j\right)\right) = \left(\sigma_{ij}\right).
\end{aligned}
$$

Παρατηρούμε ότι ο πίνακας $\mathbb{V}(X)$ είναι συμμετρικός, με διαγώνια στοιχεία τις περιθώριες διασπορές $\sigma_{ii} = \sigma_i^2$.

Πιο συγκεκριμένα, για την τυπική πολυδιάστατη κανονική έχουμε

$$\mathbb{V}(X) = \mathbb{E}\left[XX^T\right] = \mathbb{E}\left[\left(X_i X_j\right)\right] = \left(\mathbb{E}\left[X_i X_j\right]\right) = \left(\delta_{ij}\right) = \mathbb{I}_n$$

Συμβολίζουμε την τυπική πολυδιάστατη κανονική ως

$$N_n\left(x \mid \mathcal{O}_n, \mathbb{I}_n\right) = \frac{1}{(2\pi)^{n/2}} \exp\left(-\frac{1}{2}x^T x\right) \Leftrightarrow X \sim N_n\left(\mathcal{O}_n, \mathbb{I}_n\right).$$

Θέτοντας για $A = \left(a_{ij}\right)$ με $\det(A) \neq 0$ και $A^{-1} = \left(b_{ij}\right)$

$$Y = T(X) = \mu + AX \Leftrightarrow X = T^{-1}(Y) = A^{-1}(Y - \mu),$$

ή ισοδύναμα σε συντεταγμένες, για $1 = 1, \ldots, n$

$$Y_i = T\left(X_1, \ldots, X_n\right) = \mu_i + \sum_{k=1}^{n} a_{ik} X_k \Leftrightarrow X_i = T^{-1}\left(Y_1, \ldots, Y_n\right) = \sum_{k=1}^{n} b_{ik}\left(Y_k - \mu_k\right).$$

Η ιακωβιανή ορίζουσα του αντίστροφου μετασχηματισμού $Jac\left(T^{-1}\right) = \frac{\partial\left(x_1, \ldots, x_n\right)}{\partial\left(y_1, \ldots, y_n\right)} = \frac{\partial x}{\partial y}$, είναι η ορίζουσα των μερικών παραγώγων

$$\frac{\partial x_i}{\partial y_j} = \frac{\partial}{\partial y_j}\left[\sum_{k=1}^{n} b_{ik}\left(y_k - \mu_k\right)\right] = b_{ij} \Rightarrow Jac\left(T^{-1}\right) = \det\left(A^{-1}\right) = \det(A)^{-1}.$$

Έτσι

$$f_Y(y) = f_X\left(T^{-1}(y)\right)\left|Jac\left(T^{-1}\right)\right| = \frac{1}{\left|\det(A)\right|} f_X\left(A^{-1}(y - \mu)\right)$$

$$= \frac{1}{(2\pi)^{n/2}} \frac{1}{\left|\det(A)\right|} \exp\left(-\frac{1}{2}\left(A^{-1}(y-\mu)\right)^T \left(A^{-1}(y-\mu)\right)\right).$$

Παρατηρούμε ότι

$$\left|\det(A)\right| = \sqrt{\det(A)^2} = \sqrt{\det(A)\det\left(A^T\right)} = \sqrt{\det\left(AA^T\right)},$$

και

$$\left(A^{-1}(y-\mu)\right)^T\left(A^{-1}(y-\mu)\right)=(y-\mu)^T A^{-T}A^{-1}(y-\mu)$$
$$=(y-\mu)^T\left(AA^T\right)^{-1}(y-\mu),$$

που δίνουν

Συμβολίζοντας με $\Sigma=\Sigma_Y=AA^T$ τον πίνακα συνδιασποράς, η πυκνότητα γίνεται

$$f_Y(y)=\frac{1}{\sqrt{\det(2\pi\Sigma)}}\exp\left(-\frac{1}{2}(y-\mu)^T\Sigma^{-1}(y-\mu)\right)=N_n(y\,|\,\mu,\Sigma)$$

ή ότι $Y\sim N_n(\mu,\Sigma)$. Ο πίνακας συνδιασποράς Σ είναι συμμετρικός και θετικά ορισμένος, το τελευταίο σημαίνει ότι εάν $Q(y)=(y-\mu)^T\Sigma^{-1}(y-\mu)$ τότε $\min_{y\in\mathbb{R}^n}(y-\mu)^T\Sigma^{-1}(y-\mu)=Q(\mu)=0$.

<u>Εφαρμογή</u>

Εάν $Y\sim N_n(\mu,\Sigma)$, και τα Y_i είναι μεταξύ τους ασυσχέτιστα, δηλαδή $Cov(Y_i,Y_j)=\begin{cases}\sigma_i^2, & j=i\\ 0, & j\neq i\end{cases}$ τότε τα Y_i είναι μεταξύ τους ανεξάρτητα

Ο πίνακας συνδιασποράς είναι:

$$\Sigma=diag\left(\sigma_1^2,\dots,\sigma_n^2\right)\Leftrightarrow\Sigma^{-1}=diag\left(\frac{1}{\sigma_1^2},\dots,\frac{1}{\sigma_n^2}\right)$$

$$(y-\mu)^T\Sigma^{-1}(y-\mu)=(y-\mu)^T diag\left(\frac{1}{\sigma_1^2},\dots,\frac{1}{\sigma_n^2}\right)(y-\mu)$$
$$\left(\frac{y_1-\mu_1}{\sigma_1^2},\dots,\frac{y_n-\mu_n}{\sigma_n^2}\right)(y-\mu)=\sum_{i=1}^{n}\left(\frac{y_i-\mu_i}{\sigma_i}\right)^2$$

Έτσι έχουμε

$$f_Y(y)=\frac{1}{\sqrt{\det(2\pi\Sigma)}}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2\right)$$

$$= \frac{1}{\left(\sigma_1\sqrt{2\pi}\right)\cdots\left(\sigma_n\sqrt{2\pi}\right)} \exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{y_1-\mu_1}{\sigma_1}\right)^2\right) = \prod_{i=1}^{n}\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{y_i-\mu_i}{\sigma_i}\right)^2\right)$$

$$= \prod_{i=1}^{n}\underbrace{\frac{1}{\sigma_i\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{y_i-\mu_i}{\sigma_i}\right)^2\right)}_{N\left(y_i|\mu_i,\sigma_i^2\right)} = \prod_{i=1}^{n} f_{Y_i}\left(y_i\right)$$

<u>Άσκηση</u>

Δίνεται ότι η τ.μ. (X,Y) ακολουθεί την δισδιάστατη κανονική κατανομή
$(X,Y) \sim N_2\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right)$.

Δείξτε ότι

- $X \sim N(0,1)$ και $Y \sim N(0,1)$
- $[X \mid Y=y] \sim N\left(\rho y, 1-\rho^2\right)$ και $[Y \mid X=x] \sim N\left(\rho x, 1-\rho^2\right)$
- $M_{X,Y}(t,s) = \exp\left\{\frac{1}{2}\left(t^2+2\rho ts+s^2\right)\right\}$
- Η παλινδρόμηση του X στο Y είναι $\hat{X} = \mathbb{E}[X \mid Y=y] = \rho y$
- Η γραμμική παλινδρόμηση του X στο Y ταυτίζεται με την παλινδρόμηση του X στο Y

$$f_{X,Y}(x,y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left(x^2-2\rho xy+y^2\right)\right\}$$

$$f_X(x) = \frac{1}{2\pi\sqrt{1-\rho^2}}\int_{\mathbb{R}}\exp\left\{-\frac{1}{2\left(1-\rho^2\right)}\left(x^2-2\rho xy+y^2\right)\right\}dy$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2}{2\left(1-\rho^2\right)}\right\}\int_{\mathbb{R}}\exp\left\{-\frac{\left(-2\rho xy+y^2\right)}{2\left(1-\rho^2\right)}\right\}dy$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2}{2\left(1-\rho^2\right)}\right\}\int_{\mathbb{R}}\exp\left\{-\frac{\left((y-\rho x)^2-\rho^2x^2\right)}{2\left(1-\rho^2\right)}\right\}dy$$

$$= \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{x^2}{2\left(1-\rho^2\right)}+\frac{\rho^2x^2}{2\left(1-\rho^2\right)}\right\}\int_{\mathbb{R}}\exp\left\{-\frac{(y-\rho x)^2}{2\left(1-\rho^2\right)}\right\}dy$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\left(1-\rho^2\right)}} \exp\left\{-\frac{\left(y-\rho x\right)^2}{2\left(1-\rho^2\right)}\right\} dy$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} = N\left(x \mid 0,1\right)$$

Ενώ από συμμετρία έχουμε ότι $f_Y\left(y\right) = N\left(y \mid 0,1\right)$

$$f_{X|Y}\left(x \mid y\right) = \frac{f_{X,Y}\left(x,y\right)}{f_Y\left(y\right)} \propto f_{X,Y}\left(x,y\right) \propto \exp\left\{-\frac{\left(x^2 - 2\rho xy\right)}{2\left(1-\rho^2\right)}\right\}$$

$$\propto \exp\left\{-\frac{\left(x-\rho y\right)^2}{2\left(1-\rho^2\right)}\right\} \propto N\left(x \mid \rho y, 1-\rho^2\right) \Rightarrow f_{X|Y}\left(x \mid y\right) = N\left(x \mid \rho y, 1-\rho^2\right)$$

Ενώ από συμμετρία έχουμε ότι $f_{Y|X}\left(y \mid x\right) = N\left(y \mid \rho x, 1-\rho^2\right)$.

Για την δισδιάστατη ροπογεννήτρια έχουμε:

$$M_{X,Y}\left(t,s\right) = \mathbb{E}\left[e^{tX+sY}\right] = \mathbb{E}\left\{\mathbb{E}\left[e^{tX+sY} \mid Y\right]\right\} = \mathbb{E}\left\{e^{sY}\mathbb{E}\left[e^{tX} \mid Y\right]\right\}$$

$$= \mathbb{E}\left\{e^{sY} M_{X|Y}\left(t\right)\right\}$$

Γνωρίζουμε ότι $\left[X \mid Y = y\right] \sim N\left(\rho y, 1-\rho^2\right)$

Έτσι η δεσμευμένη ροπογεννήτρια για $Y = y$ είναι

$$\mathbb{E}\left[e^{tX} \mid Y = y\right] = \exp\left\{\rho y t + \left(1-\rho^2\right)t^2/2\right\} \Rightarrow M_{X|Y}\left(t\right) = \exp\left\{\rho t Y + \frac{1}{2}\left(1-\rho^2\right)t^2\right\}$$

Έτσι

$$M_{X,Y}\left(t,s\right) = \mathbb{E}\left\{e^{sY} M_{X|Y}\left(t\right)\right\} = \mathbb{E}\left\{\exp\left\{\left(s+\rho t\right)Y + \left(1-\rho^2\right)t^2/2\right\}\right\}$$

$$= \exp\left\{\left(1-\rho^2\right)t^2/2\right\} \mathbb{E}\left\{\exp\left\{\left(s+\rho t\right)Y\right\}\right\}$$

Η περιθώρια ως προς Y είναι τυπική κανονική έτσι

$$\mathbb{E}\left\{\exp\left\{\left(s+\rho t\right)Y\right\}\right\} = M_Y\left(s+\rho t\right) = \exp\left\{\frac{1}{2}\left(s+\rho t\right)^2\right\}$$

Τελικά έχουμε:

$$M_{X,Y}(t,s) = \exp\left\{\frac{1}{2}\left(1-\rho^2\right)t^2\right\}\exp\left\{\frac{1}{2}\left(s+\rho t\right)^2\right\} = \exp\left\{\frac{1}{2}\left(t^2+2\rho ts+s^2\right)\right\}$$

$$\widehat{X} = \mathbb{E}\left[X \mid Y=y\right] = \int_{\mathbb{R}} x\, N\left(x \mid \rho y, 1-\rho^2\right) = \rho y$$

$$\widehat{\alpha} = \mathbb{E}\left[X\right] - \widehat{\beta}\,\mathbb{E}\left[Y\right] = 0,\ \widehat{\beta} = \rho_{X,Y}\frac{\sigma_X}{\sigma_Y} = \rho \ \Rightarrow\ \widehat{X} = \widehat{\alpha} + \widehat{\beta} y = \rho y\,.$$

Ορισμός: Ένας πίνακας $\Sigma \in \mathbb{R}^{n\times n}$ λέμε ότι είναι **θετικά ορισμένος** εάν είναι συμμετρικός και η τετραγωνική μορφή[1] $Q_\Sigma(x) = x^T \Sigma\, x$, όπου $x = \left(x_1,\ldots,x_n\right)^T$, είναι θετική για κάθε $x \in \mathbb{R}_*^{n\times 1} = \mathbb{R}^{n\times 1} \setminus \left\{\mathcal{O}_n\right\}$

1. $\Sigma = \Sigma^T$.
2. $Q_\Sigma(x) = x^T \Sigma\, x = \sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j > 0,\ \forall x \in \mathbb{R}_*^{n\times 1}$.

Στην περίπτωση θετικά ημιορισμένου πίνακα ζητάμε $Q_\Sigma(x) \geq 0,\ \forall x \in \mathbb{R}_*^{n\times 1}$.

Στον πιο πάνω ορισμό το 2 είναι ισοδύναμο με το να ζητήσουμε η συνάρτηση $Q_\Sigma(x_1,\ldots,x_n)$ να έχει ολικό ελάχιστο στο $\mathcal{O}_n$, ή ότι,
$0 = Q_\Sigma(\mathcal{O}_n) < Q_\Sigma(x_1,\ldots,x_n),\ \forall x \in \mathbb{R}_*^{n\times 1}$ (να είναι δηλαδή γνησίως κυρτή).

Οι θετικά ορισμένοι και ημιορισμένοι πίνακες είναι πολύ σημαντικοί στην στατιστική εφόσον αντιστοιχούν σε **πίνακες συνδιασποράς**.

Από την γραμμική άλγεβρα γνωρίζουμε ότι **εάν ο πίνακας M είναι συμμετρικός, τότε όλες οι ιδιοτιμές του είναι πραγματικές**, δηλαδή[2] $spec(M) \subset \mathbb{R}$, με $spec(M) = \left\{\lambda \in C : \det\left(M - \lambda \mathbb{I}\right) = 0\right\}$. Έστω ιδιοτιμή $\lambda \in spec(M)$ με αντίστοιχο ιδιοδιάνυσμα v, τότε $M\,v = \lambda\, v$ που σημαίνει ότι εάν ο M είναι και θετικά ορισμένος $Q_M(v) > 0$ ή ότι

$$Q_M(v) = v^T M\, v = \lambda v^T v = \lambda \left\|v\right\|^2 > 0 \ \Rightarrow\ \lambda > 0\,.$$

[1] The purely quadratic form associated with M
[2] Με spec συμβολίζουμε το spectrum – φάσμα, δηλαδή το σύνολο ιδιοτιμών ενός πίνακα.

Έτσι όλες οι ιδιοτιμές ενός συμμετρικού και θετικά ορισμένου πίνακα M είναι θετικές, δηλαδή $spec(M) \subset \mathbb{R}^+$, με αποτέλεσμα ο M να είναι και αντιστρέψιμος εφόσον $\det(M) = \prod_{\lambda \in spec(M)} \lambda > 0$.

Έστω ότι ο M είναι συμμετρικός. Εάν υπάρχουν n διαφορετικά ζεύγη (λ_i, v_i) τέτοια ώστε $M v_i = \lambda_i v_i$, τότε $M[v_1, \dots, v_n] = [v_1, \dots, v_n] D$ όπου $D = (\lambda_i \delta_{ij})$. Ο πίνακας $P = [v_1, \dots, v_n]$ είναι **ορθογώνιος**, δηλαδή $P^{-1} = P^T \Leftrightarrow P = P^{-T}$ και

$D = P^T M P$.

Επειδή $PP^T = \mathbb{I}$ θα έχουμε και $\det(P)^2 = 1 \Leftrightarrow |\det(P)| = 1$. Εάν επιπροσθέτως ο M είναι και θετικά ορισμένος, θα έχουμε $\det(P) = 1$.

<u>Άσκηση</u>
Να βρεθούν οι ελάχιστες συνθήκες κάτω από τις οποίες ο 2×2 συμμετρικός πίνακας $M = (m_{ij})$ είναι θετικά ορισμένος.

$$Q_M(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} m_{11} & m_{12} \\ m_{12} & m_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = m_{11} x^2 + 2 m_{12} xy + m_{22} y^2$$
$$= m_{11} \left\{ \left(x + \frac{m_{12}}{m_{11}} y \right)^2 + \frac{m_{22}}{m_{11}} y^2 - \left(\frac{m_{12}}{m_{11}} y \right)^2 \right\} = m_{11} \left(x + \frac{m_{12}}{m_{11}} y \right)^2 + \left(m_{22} - \frac{m_{12}^2}{m_{11}} \right) y^2$$

$$Q_M(x, y) > 0,\ \forall (x, y) \in \mathbb{R}_*^2 \Leftrightarrow \left\{ \begin{matrix} m_{11} > 0 \\ m_{22} - \dfrac{m_{12}^2}{m_{11}} > 0 \end{matrix} \right\} \Leftrightarrow \left\{ \begin{matrix} m_{11} > 0 \\ \det(M) > 0 \end{matrix} \right\}.$$

<u>Άσκηση</u>
Δίνεται δ.τ.μ. $X = (X_1, \dots, X_n)^T \in \mathbb{R}^{n \times 1}$. Εάν $M = (m_{ij}) \in \mathbb{R}^{n \times n}$ ορίζουμε την δ.τ.μ. $Y = MX$. Δείξτε ότι $\mathbb{E}[Y] = M\, \mathbb{E}[X]$

$$\mathbb{E}[Y] = \mathbb{E}[M X] = \mathbb{E}\left[\left(\sum_j m_{ij} X_j \right) \right] = \left(\sum_j m_{ij} \mathbb{E}[X_j] \right) = M\, \mathbb{E}[X].$$

Εάν ο M είναι θετικά ορισμένος τότε μπορεί να παραγοντοποιηθεί σαν $M = L L^T$ όπου L είναι πίνακας κάτω τριγωνικός, με θετικά διαγώνια στοιχεία. **Ο πίνακας L καλείται Cholesky παράγοντας του** M είτε τετραγωνική ρίζα του M. Τότε ο

M^{-1} έχει Cholesky παράγοντα L^{-1}, εφόσον $M^{-1} = \left(L^{-1}\right)^T L^{-1}$, και για αυτό το λόγο είναι και αυτός θετικά ορισμένος.

Είναι προφανές ότι εάν $M = L\,L^T$ τότε ο M είναι συμμετρικός και θετικά ορισμένος, πράγματι:

$$M^T = \left(L\,L^T\right)^T = L\,L^T = M\ ,$$
$$Q_M\left(x\right) = x^T L\,L^T x = \left(L^T x\right)^T L^T x = \left\|L^T x\right\|^2 > 0,\ \forall x \in \mathbb{R}^{n\times 1*}.$$

Επίσης $M^{-1} = \left(LL^T\right)^{-1} = L^{-T}L^{-1}$ διότι $MM^{-1} = \left(L\,L^T\right)\left(L^{-T}L^{-1}\right) = L\left(L^T L^{-T}\right)L^{-1} = L\,L^{-1} = \mathbb{I}$. Παρόμοια και $M^{-1}M = \mathbb{I}$.

Τα επόμενα έως το τέλος του PDF είναι εκτός

Παράδειγμα

Να αποδειχθεί ότι ο πίνακας

$$M = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix},$$

είναι θετικά ορισμένος.

Αρκεί να δείξουμε ότι $M = L\,L^T$ με $L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$ και $l_{ii} > 0$ για $i = 1,2,3$.

$$L\,L^T = \begin{bmatrix} \left(l_{11}^2\right) & \left(l_{11}l_{21}\right) & \left(l_{11}l_{31}\right) \\ l_{21}l_{11} & \left(l_{21}^2 + l_{22}^2\right) & \left(l_{21}l_{31} + l_{22}l_{32}\right) \\ l_{31}l_{11} & l_{31}l_{21} + l_{32}l_{22} & \left(l_{31}^2 + l_{32}^2 + l_{33}^2\right) \end{bmatrix} = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} \Rightarrow$$

Λύνοντας τις εξισώσεις στις παρενθέσεις παίρνουμε: $L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix}$.

Άσκηση

Να βρεθεί ο αντίστροφος του πίνακα $L = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix}$ με $l_{ii} > 0$ για $i = 1,2,3$. Στη συνέχεια να αντιστραφεί ο πίνακας L του προηγούμενου παραδείγματος.

Ο L είναι αντιστρέψιμος εφόσον $\det(L) = l_{11}l_{22}l_{33} > 0$ [3]. Έστω $L^{-1} = (r_{ij})$, τότε από την εξίσωση $(l_{ij})(r_{ij}) = \mathbb{I}$ έχουμε

$$L^{-1} = (r_{ij}) = \begin{bmatrix} \dfrac{1}{l_{11}} & 0 & 0 \\ -\dfrac{l_{21}}{l_{11}l_{22}} & \dfrac{1}{l_{22}} & 0 \\ \dfrac{l_{32}l_{21}}{l_{11}l_{22}l_{33}} - \dfrac{l_{31}}{l_{11}l_{33}} & -\dfrac{l_{32}}{l_{22}l_{33}} & \dfrac{1}{l_{33}} \end{bmatrix}$$

Αντικαθιστώντας $L = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix}$ παίρνουμε $L^{-1} = \begin{bmatrix} \dfrac{1}{5} & 0 & 0 \\ -\dfrac{1}{5} & \dfrac{1}{3} & 0 \\ \dfrac{2}{15} & -\dfrac{1}{9} & \dfrac{1}{3} \end{bmatrix}$.

<u>Πρόταση</u>

Δίνεται η δ.τ.μ. $X = (X_1, \ldots, X_n)^T$ όπου τα X_i είναι ανεξάρτητα και $X_i \sim N(0, \sigma_i^2)$ για $1 \le i \le n$. Η ροπογεννήτρια της δ.τ.μ. $Y = \mu + BX$ όπου $\mu \in \mathbb{R}^{n\times 1}$ και $B \in \mathbb{R}^{n\times n}$ αντιστρέψιμος πίνακας (δηλαδή $\det(B) \ne 0$) είναι

$$M_Y(t) = \exp\left(t^T\mu + \frac{1}{2}t^T\left(BDB^T\right)t\right).$$

$$M_Y(t) = \mathbb{E}\left[\exp\left(t^T Y\right)\right] = \exp\left(t^T\mu\right)\mathbb{E}\left[\exp\left(t^T BX\right)\right] = \exp\left(t^T\mu\right)M_X\left(B^T t\right)$$

$$M_X(u) = \mathbb{E}\left[\exp\left(u^T X\right)\right] = \mathbb{E}\left[\exp\left(\sum u_i X_i\right)\right] = \prod \mathbb{E}\left[\exp\left(u_i X_i\right)\right]$$
$$= \prod M_{X_i}(u_i) = \prod \exp\left(\frac{1}{2}\sigma_i^2 u_i^2\right) = \exp\left(\frac{1}{2}\sum \sigma_i^2 u_i^2\right)$$

Ορίζουμε τον διαγώνιο πίνακα $D = \left(\sigma_i^2\delta_{ij}\right)$, τότε, $u^T D u = \sum \sigma_i^2 u_i^2$. Έτσι η ροπογεννήτρια του X μπορεί να αναπαρασταθεί σαν

$$M_X(u) = \exp\left(\frac{1}{2}u^T D u\right)$$

[3] Είναι πολύ εύκολο να δείξουμε ότι οι ιδιοτιμές του L είναι τα διαγώνια στοιχεία του.

Για τη ροπογεννήτρια της Y έχουμε

$$M_Y(t) = \exp\left(t^T \mu\right) M_X\left(B^T t\right) = \exp\left(t^T \mu\right) \exp\left(\frac{1}{2}\left(B^T t\right)^T D\left(B^T t\right)\right)$$
$$= \exp\left(t^T \mu + \frac{1}{2} t^T \left(BDB^T\right) t\right).$$

Παρατηρούμε ότι:

$$\Sigma_X = \mathbb{E}\left[XX^T\right] = \mathbb{E}\left[\left(X_i X_j\right)\right] = \left(\mathbb{E}\left[X_i X_j\right]\right) = \left(\sigma_i^2 \delta_{ij}\right) = D$$

$$\Sigma_Y = \mathbb{E}\left[(Y-\mu)(Y-\mu)^T\right] = \mathbb{E}\left[(BX)(BX)^T\right] = B\mathbb{E}\left[XX^T\right]B^T = BDB^T$$

<u>Άσκηση</u>
Δείξτε ότι ο Σ_Y είναι θετικά ορισμένος

Πράγματι

$$\left(BDB^T\right)^T = \left(B^T\right)^T D^T B^T = BDB^T$$

Εάν $t = B^{-T} u$ τότε η τετραγωνική μορφή $Q_{BDB^T}(t) = t^T BDB^T t$ γίνεται
$Q_{BDB^T}\left(B^{-T}u\right) = \left(B^{-T}u\right)^T BDB^T \left(B^{-T}u\right) = u^T D u = \sum_i \sigma_i^2 u_i^2 > 0$.

Κάνοντας τον αντίστροφο μετασχηματισμό $u = B^T t = \left(\sum_j b_{ji} t_j\right)$ έχουμε ότι

$$Q_{BDB^T}(t) = \sum_i \sigma_i^2 \left(\sum_j b_{ji} t_j\right)^2 > 0.$$

<u>Πρόταση</u>
Εάν γνωρίζουμε ότι $M_Y(t) = \exp\left(t^T \mu + \frac{1}{2} t^T \Sigma\, t\right)$ τότε υπάρχει δ.τ.μ.
$X = \left(X_1, \ldots, X_n\right)^T \in \mathbb{R}^{n\times 1}$ όπου τα X_i είναι ανεξάρτητα και $X_i \sim N\left(0, \sigma_i^2\right)$ για $1 \le i \le n$ και $X = P^T(Y-\mu) \in \mathbb{R}^{n\times 1}$, όπου P είναι ο πίνακας που διαγωνιοποιεί τον Σ, δηλαδή $D = P^T \Sigma\, P$ με $D = \left(\sigma_i^2 \delta_{ij}\right)$.

$$M_X(u) = \mathbb{E}\left[\exp\left(u^T X\right)\right] = \mathbb{E}\left[\exp\left(u^T P^T (Y-\mu)\right)\right] = \exp\left(-u^T P^T \mu\right)\mathbb{E}\left[\exp\left(u^T P^T Y\right)\right]$$

$$= \exp\left(-u^T P^T \mu\right) M_Y\left(Pu\right)$$

Επειδή $M_Y\left(t\right) = \exp\left(t^T \mu + \frac{1}{2} t^T \Sigma\, t\right)$, θέτοντας $t = Pu$ έχουμε

$$M_Y\left(Pu\right) = \exp\left(u^T P^T \mu + \frac{1}{2} u^T P^T \Sigma\, Pu\right) = \exp\left(u^T P^T \mu\right) \exp\left(\frac{1}{2} u^T P^T \Sigma\, Pu\right)$$

$$\Leftrightarrow \exp\left(-u^T P^T \mu\right) M_Y\left(Pu\right) = \exp\left(\frac{1}{2} u^T P^T \Sigma\, Pu\right)$$

Που δίνει

$$M_X\left(u\right) = \exp\left(\frac{1}{2} u^T P^T \Sigma\, Pu\right) = \exp\left(\frac{1}{2} u^T D u\right) = \exp\left(\frac{1}{2} \sum \sigma_i^2 u_i^2\right)$$

$$= \prod \exp\left(\frac{1}{2} \sigma_i^2 u_i^2\right) = \prod M_{X_i}\left(u_i\right).$$

Δηλαδή η δ.τ.μ. $X = \left(X_1, \ldots, X_n\right)^T \in \mathbb{R}^{n\times 1}$ είναι τέτοια ώστε τα X_i είναι ανεξάρτητα και $X_i \sim N\left(0, \sigma_i^2\right)$.

<u>Παράδειγμα</u>

Δίνεται δ.τ.μ. $X = \left(X_1, \ldots, X_n\right)^T \in \mathbb{R}^{n\times 1}$ όπου τα X_i είναι ανεξάρτητα και $X_i \sim N\left(0, \sigma_i^2\right)$ για $1 \le i \le n$, επίσης θέτουμε $D = \left(\sigma_i^2 \delta_{ij}\right)$.

1. Να βρεθεί η πυκνότητα f_Y της δ.τ.μ. $Y = BX + \mu$ για $\det\left(B\right) \neq 0$.

2. Δείξτε ότι $f_Y\left(y\right) < f_Y\left(\mu\right) = \dfrac{1}{\sqrt{\det\left(2\pi\Sigma\right)}}$, $\forall\, y \in \mathbb{R}^n \setminus \left\{\mu\right\}$.

1. Εάν $Y = T\left(X\right) = BX + \mu$ τότε $X = T^{-1}\left(Y\right) = B^{-1}\left(Y - \mu\right)$. Η Ιακωβιανή του T^{-1} είναι

$$Jac\left(T^{-1}\right) = \det\left(\frac{\partial x}{\partial y}\right) = \det\left(B^{-1}\right) = \frac{1}{\det\left(B\right)}$$

$$f_X\left(x\right) = \prod_i f_{X_i}\left(x_i\right) = \prod_i N\left(x_i \mid 0, \sigma_i^2\right) = \prod_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2\sigma_i^2} x_i^2\right)$$

$$= \frac{1}{(2\pi)^{n/2}\sqrt{\sigma_1^2\cdots\sigma_n^2}} \exp\left(-\frac{1}{2}\sum_i \frac{1}{\sigma_i^2} x_i^2\right) = \frac{1}{\sqrt{\det(2\pi D)}} \exp\left(-\frac{1}{2} x^T D^{-1} x\right).$$

Όμως $f_Y(y) = f_X\left(T^{-1}(y)\right)\left|Jac\left(T^{-1}\right)\right|$ που δίνει

$$f_Y(y) = f_X\left(B^{-1}(Y-\mu)\right)\left|\frac{1}{\det(B)}\right|$$

$$= \frac{1}{\sqrt{\det(2\pi D)\det(B)^2}} \exp\left(-\frac{1}{2}(Y-\mu)^T B^{-T} D^{-1} B^{-1}(Y-\mu)\right)$$

$$= \frac{1}{\sqrt{\det\left(2\pi BDB^T\right)}} \exp\left(-\frac{1}{2}(Y-\mu)^T \left(BDB^T\right)^{-1}(Y-\mu)\right)$$

Θέτοντας $\Sigma = BDB^T$ τον πίνακα συνδιασποράς του Y παίρνουμε τελικά

$$f_Y(y) = \frac{1}{\sqrt{\det(2\pi\Sigma)}} \exp\left(-\frac{1}{2}(Y-\mu)^T \Sigma^{-1}(Y-\mu)\right).$$

2. $Q_{\Sigma^{-1}}(y-\mu) = (y-\mu)^T \Sigma^{-1}(y-\mu)$

Επειδή εάν Σ είναι θετικά ορισμένος τότε και Σ^{-1} θετικά ορισμένος έχουμε ότι η τετραγωνική μορφή $Q_{\Sigma^{-1}}(y-\mu)$ είναι αυστηρά κυρτή με ελάχιστο 0 στο $y=\mu$. Δηλαδή

$$0 = Q_{\Sigma^{-1}}\left(0^T\right) < Q_{\Sigma^{-1}}(y-\mu),\ \forall\, y \in \mathbb{R}^n \setminus \{\mu\} \Leftrightarrow -\frac{1}{2} Q_{\Sigma^{-1}}(y-\mu) < 0,\ \forall\, y \in \mathbb{R}^n \setminus \{\mu\}$$

$$\Leftrightarrow \exp\left(-\frac{1}{2} Q_{\Sigma^{-1}}(y-\mu)\right) < 1,\ \forall\, y \in \mathbb{R}^n \setminus \{\mu\}$$

$$\Leftrightarrow f_Y(y) < f_Y(\mu) = \frac{1}{\sqrt{\det(2\pi\Sigma)}},\ \forall\, y \in \mathbb{R}^n \setminus \{\mu\}.$$

<u>Παρατήρηση</u>

Η συμβολική αναπαράσταση των προηγουμένων είναι

Εάν τα X_i είναι ανεξάρτητα και $X_i \sim N\left(0, \sigma_i^2\right)$ για $1 \le i \le n$, με $D = \left(\sigma_i^2 \delta_{ij}\right)$ τότε γράφουμε

$$X \sim N_n\left(0^T, D\right).$$

Επειδή $Y = BX + \mu$ με $\det(B) \neq 0$ γράφουμε

$$Y = BX + \mu \sim B \cdot N_n\left(0^T, D\right) + \mu \sim N_n\left(\mu, BDB^T\right).$$

Τότε είναι εμφανές ότι

$$Y \sim N_n\left(\mu, BDB^T\right) \Leftrightarrow Y - \mu \sim N_n\left(0^T, BDB^T\right)$$
$$\Leftrightarrow X = B^{-1}(Y-\mu) \sim N_n\left(0^T, B^{-1}\left(BDB^T\right)B^{-T}\right) = N_n\left(0^T, \left(B^{-1}B\right)D\left(B^{-1}B\right)^T\right)$$
$$= N_n\left(0^T, D\right).$$

Δηλαδή

$$X \sim N_n\left(0^T, D\right) \Leftrightarrow Y = BX + \mu \sim N_n\left(\mu, BDB^T\right).$$

Εάν $A \in \mathbb{R}^{n\times n}$ με $\det(A) \neq 0$ τότε η δ.τ.μ. $Z = AY + \mu'$ έχει κατανομή

$$Z = AY + \mu' \sim N_n\left(A\mu + \mu', A\left(BDB^T\right)A^T\right) = N_n\left(A\mu + \mu', (AB)D(AB)^T\right).$$

<u>Παράδειγμα</u>
Δίνεται ότι τα X_i είναι ανεξάρτητα με $X_i \sim N(0,1)$ για $1 \le i \le n$. Να βρεθεί η κατανομή της δ.τ.μ. $Z = \left(Z_1, \ldots, Z_n\right)^T \in \mathbb{R}^{n\times 1}$, όπου $Z_i = \sum_{j=1}^{n} a_{ij} X_j + b_i, 1 \le i \le n$ με $A = \left(a_{ij}\right) \in \mathbb{R}^{n\times n}$ αντιστρέψιμο πίνακα, καθώς και ο πίνακας συνδιασποράς Σ_Z.

$$Z = AX + b \sim A \cdot N_n\left(0^T, \mathbb{I}\right) + b = N_n\left(b, AA^T\right)$$

$$\Sigma_Z = \mathbb{E}\left[(Z-b)(Z-b)^T\right] = \mathbb{E}\left[(AX)(AX)^T\right] = A\mathbb{E}\left[XX^T\right]A^T = A\mathbb{I}_n A^T = AA^T.$$