**A december 7-i szeminárium témája**
*Rövid összefoglaló*

Mivel egy véletlen vektor eloszlását meghatározza annak karakterisztikus függvénye, a normális eloszlású valószínűségi változó karakterisztikus függvényének viszonylag egyszerű alakjából sok egyszerű, de fontos következtetést lehet levonni. Érdemes ezt a karakterisztikus függvényt tömör mátrix és vektor jelölésekkel felírni. Jelölje egy $(\xi_1,\dots,\xi_k)$ $k$-dimenziós normális eloszlású véletlen vektor kovariancia mátrixát egy $\Sigma$ mátrix $(M_1,\dots,M_k)=(E\xi_1,\dots,E\xi_k)$ várható érték vektorát egy $M$ vektor, és jelöljük két

$$u=(u_1,\dots,u_k)\in R^k \quad \text{és} \quad v=(v_1,\dots,v_k)\in R^k$$

vektor $u_1v_1+\cdots+u_kv_k$ skalárszorzatát $(u,v)$-vel. Ezekkel a jelölésekkel a $(\xi_1,\dots,\xi_k)$ normális eloszlású véletlen vektor karakterisztikus függvénye a $t=(t_1,\dots,t_k)$ pontban

$$\varphi(t_1,\dots,t_k)=\varphi(t)=\exp\left\{-\frac{t\Sigma t^*}{2}+i(M,t)\right\}.$$

E formulából azonnal látható, hogy egy normális eloszlású valószínűségi vektor eloszlását meghatározza annak kovariancia mátrixa és várható érték vektora. Jegyezzük meg, hogy egy normális eloszlású vektor előállítható $(\eta_1,\dots,\eta_k)A+M$ alakban alkalmas $k\times k$ mátrix-szal, független standard normális $\eta_1,\dots,\eta_k$ valószínűségi változókkal és $M$ vektorral. Ekkor az $A$ mátrix és $M$ vektor meghatározza a normális eloszlású valószínűségi vektor eloszlását, mert $\Sigma=A^*A$. Ennek az állításnak a megfordítása viszont nem igaz, mert a $\Sigma=A^*A$ egyenlet nem határozza meg az $A$ mátrixot. (Lásd az előző gyakorlaton tárgyalt Lemma utáni megjegyzést.)

A fenti formula egyik egyszerű, de fontos következménye az, hogy ha $(\xi_1,\dots,\xi_k)$ normális eloszlású véletlen vektor, melynek kovarianciamátrixa egy $\Sigma_1$ mátrix, várható értéke egy $M^{(1)}=\left(M_l^{(1)},\dots,M_k^{(1)}\right)$ vektor, $(\eta_1,\dots,\eta_k)$ szintén normális eloszlású véletlen vektor, melynek kovarianciamátrixa egy $\Sigma_2$ mátrix, várható értéke egy $M^{(2)}=\left((M_l^{(2)},\dots,M_k^{(2)}\right)$ vektor, és a $(\xi_1,\dots,\xi_k)$ és $(\eta_1,\dots,\eta_k)$ véletlen vektorok függetlenek, akkor a $(\xi_1+\eta_1,\dots,\xi_k+\eta_k)$ véletlen vektor is normális eloszlású $\Sigma_1+\Sigma_2$ kovarianciamátrix-szal és $M^{(1)}+M^{(2)}$ várható értékkel. Ez az állítás, mely az egydimenziós normális eloszlású valószínűségi változók egyik fontos tulajdonságának többdimenziós általánosítása, azonnal kiolvasható abból a tényből, hogy a $(\xi_1+\eta_1,\dots,\xi_k+\eta_k)$ véletlen vektor karakterisztikus függvénye a $(\xi_1,\dots,\xi_k)$ és $(\eta_1,\dots,\eta_k)$ véletlen vektorok karakterisztikus függvényeinek a szorzata.

Továbbá a normális eloszlású valószínűségi változók karakterisztikus függvényének alakjából következik, hogy ha egy normális elozlású véletlen vektor bizonyos koordinátái korrelálatlanok akkor ezek a koordináták egyben függetlenek is. Hangsúlyozzuk, hogy ez az állítás tetszőleges, nem feltétlenül normális eloszlású vektorra nem érvényes. A függetlenségből mindig következik a korrelálatlanság, de megfordítva ez nem igaz. A korrelálatlanság a függetlenségnél jóval gyengébb tulajdonság.

A fenti állítás bizonyítása előtt lássunk egyszerű példát arra, hogy a korrelálatlanság gyengébb állítás mint a függetlenség. Egy lehetséges konstrukciót kapunk, ha véges sok, de legalább három értéket felvevő valószínűségi változókat tekintünk. Akkor a korrelálatlanság egy azonosság teljesülését követeli meg (azon kívül, hogy a valószínűségi változók által felvett értékek valószínűségeinek összege egy). Ezért némi számolással megválaszthatjuk a valószínűségeket úgy, hogy ez az azonosság teljesüljön, de a függetlenséget követelő (sokkal több) azonosság mindegyike ne teljesüljön. Talán tanulságosabb a következő példa. Legyen az $(\Omega, \mathcal{A}, P)$ valószínűségi mező a $[0,1]$ intervallum a szokásos Borel $\sigma$-algebrával és Lebesgue mértékkel. Legyen $\xi(x) = \cos(2\pi x)$, $\eta(x) = \xi(2x) = \cos(4\pi x)$. Mint azt analízisből tanultuk, a trigonometrikus függvények (teljes) ortogonális rendszert alkotnak, a $[0, 2\pi]$ intervallumban. Ez jelen esetben azt jelenti speciálisan, hogy a fent definiált (nulla várható értékű) $\xi$ és $\eta$ valószínűségi változók korrelálatlanok. Viszont $\xi$ és $\eta$ nem független. Sőt, a $\cos 2\alpha = 2\cos^2\alpha - 1$ azonosság alapján $\eta(x) = 2\xi(x)^2 - 1$, azaz az $\eta$ valószínűségi változó a $\xi$ valószínűségi változó determinisztikus függvénye.

Amikor normális eloszlású véletlen vektorok koordinátáinak korrelálatlansága és függetlensége közötti kapcsolatot vizsgáljuk, akkor tekintsük az egyszerű jelölés érdekében azt az esetet, amikor az első $j$ $\xi_u$, $1 \le u \le j$, és az utolsó $k-j$ $\xi_v$, $k-j < v \le k$ valószínűségi változók korrelálatlanok, azaz $E(\xi_u - E\xi_u)(\xi_v - E\xi_v) = 0$, ha $1 \le u \le j < v \le k$. Ekkor a $(\xi_1, \dots, \xi_k)$ mátrix $\Sigma$ kovariancia mátrixa a következő egyszerűbb szerkezettel rendelkezik:

$$\Sigma = \begin{pmatrix} \Sigma_1, \ 0 \\ 0, \ \Sigma_2 \end{pmatrix} \tag{+}$$

ahol $\Sigma_1$ a $\xi^{(1)} = (\xi_1, \dots, \xi_j)$, $\Sigma_2$ a $\xi^{(2)} = (\xi_{j+1}, \dots, \xi_k)$ véletlen vektor kovariancia mátrixa. Legyen $M^{(1)} = (E\xi_1, \dots, E\xi_j)$, $M^{(2)} = (E\xi_{j+1}, \dots, E\xi_k)$, $t^{(1)} = (t_1, \dots, t_j)$, $t^{(2)} = (t_{j+1}, \dots, t_k)$, ha $t = (t_1, \dots, t_k) \in R^k$. Innen a $(\xi_1, \dots, \xi_k)$ vektor karakterisztikus függvénye

$$\begin{aligned}
\varphi(t) &= \exp\left\{-\frac{t\Sigma t^*}{2} + i(M, t)\right\} \\
&= \exp\left\{-\frac{t^{(1)}\Sigma_1\left(t^{(1)}\right)^*}{2} + i\left(M^{(1)}, t^{(1)}\right)\right\} \exp\left\{-\frac{t^{(2)}\Sigma_2\left(t^{(2)}\right)^*}{2} + i\left(M^{(2)}, t^{(2)}\right)\right\} \\
&= \varphi_1\left(t^{(1)}\right)\varphi_2\left(t^{(2)}\right),
\end{aligned}$$

ahol $\varphi_1\left(t^{(1)}\right)$ a $\xi^{(1)}$ $\varphi_2\left(t^{(2)}\right)$ pedig a $\xi^{(2)}$ véletlen vektor karakterisztikus függvénye. A karakterisztikus függvény alábbi szorzatelőállításából következik a megfelelő komponensek függetlensége.

Ennek az eredménynek a segítségével be fogjuk látni, hogy amennyiben $(\xi, \eta)$, $\xi = (\xi_1, \dots, \xi_k)$, $\eta = (\eta_1, \dots, \eta_l)$ $k + l$ dimenziós normális eloszlású valószínűségi változó, akkor létezik olyan $A$ $l \times k$ mátrix, melyre a $\zeta = \xi - \eta A$ és az $\eta$ vektor független. Ez azt jelenti, hogy a $\xi = \zeta + \eta A$ egyenlőség a $\xi$ véletlen vektort előállítja, mint az $\eta$ vektor egy lineáris transzformáltját plusz egy az $\eta$ véletlen vektortól független normális eloszlású véletlen vektor összgét. Mivel a $(\zeta,\eta)$ véletlen vektor a $(\xi,\eta)$ vektor lineáris transzformáltja, explicit képletben

$$(\zeta,\eta)=(\xi,\eta)\begin{pmatrix} I, & -A \\ 0, & I \end{pmatrix},$$

ezért a $(\zeta,\eta)$ véletlen vektor is normális eloszlású. Az előzőek alapján a $\zeta$ és $\eta$ vektorok függetlenségének biztosításához elég belátni, hogy az $A$ mátrix alkalmas megválasztásával elérhető az, hogy a $\zeta$ és $\eta$ vektor koordinátái korrelálatlanok, azaz a $(\zeta,\eta)$ vektor kovarianciamátrixa (+) alakú legyen.

Az állítás bizonyításának megértéséhez tekintsük először a legegyszerűbb esetet, azt amikor $k=l=1$. Ekkor a $\zeta=\xi-\dfrac{\mathrm{Cov}\,(\xi,\eta)}{\mathrm{Var}\,\eta}\eta$ valószínűségi változó teljesíti a $\mathrm{Cov}\,(\zeta,\eta)=0$ relációt, ezért $\eta$ és $\zeta$ független valószínűségi változók. A több dimenziós eset bizonyítása hasonló, de az ebben az esetben felmerülő lineáris algebrai feladat nehezebb.

Ha az $(\eta_1,\dots,\eta_l)$ véletlen vektor koordinátái korrelálatlanok (és ezért az együttes normális eloszlásuk miatt függetlenek is), akkor az előző formula természetes módosítása megoldja a problémát. Ekkor legyen $\zeta_p=\xi_p-\sum\limits_{q=1}^{l}\dfrac{\mathrm{Cov}\,(\xi_p,\eta_q)}{\mathrm{Var}\,\eta_q}\eta_q$, $1\le p\le k$, $1\le q\le l$. Egyszerű számolás adja, hogy ebben az esetben $\mathrm{Cov}\,(\zeta_p,\eta_q)=0$, $1\le p\le k$, $1\le q\le l$.

Az általános eset visszavezethető erre az esetre az $\eta$ vektor alkalmas lineáris transzformáltját alkalmazva. Jelölje $\Sigma_2$ az $\eta$ vektor kovariancia mátrixát. Megint fel fogjuk használni azt a lineáris algebrai eredményt mely szerint a (pozitív (szemi)definit) $\Sigma_2$ mátrix felírható $\Sigma_2=U\Lambda U^*$ alakban, ahol $U$ unitér $\Lambda$ pedig diagonális mátrix. Be fogjuk látni, hogy az $\bar\eta=(\bar\eta_1,\dots,\bar\eta_k)=(\eta_1,\dots,\eta_k)U=\eta U$ vektor kovarianciamátrixa a $\Lambda$ diagonális mátrix. Másrészt egy $\zeta$ vektor akkor és csak akkor független az $\eta$ vektortól, ha független a $\bar\eta=\eta U$ vektortól. Ezért a $\zeta$ vektort megválaszthatjuk mint $\zeta=\xi-\bar\eta A=\xi-\eta UA$ vektornak, ahol az $A$ mátrixot a diagonális kovarianciamátrixú $\bar\eta$ véletlen normális eloszlású vektor által meghatározott lineáris transzformáció határozza meg, azaz az $A=(a_{q,p})$ mátrix az az $l\times k$ méretű mátrix, melyre $a_{q,p}=\dfrac{\mathrm{Cov}\,(\xi_p,\bar\eta_q)}{\mathrm{Var}\,\bar\eta_q}$, $1\le p\le k$, $1\le q\le l$.

Ezért az állítást bebizonyítjuk, ha megmutatjuk, hogy az $\bar\eta=\eta U$ véletlen vektor kovariancia mátrixa a diagonális $\Lambda$ mátrix. Ez következik az alábbi viszonylag egyszerű, de egyébként is hasznos állításból. Legyen a $\xi=(\xi_1,\dots,\xi_l)$ $l$-dimenziós véletlen vektor, melynek kovariancia mátrixa $\Sigma$, $\eta=(\eta_1,\dots,\eta_l)=\xi A$, ahol $A=(a_{p,q})$ $l\times l$ méretű mátrix. Ekkor az $\eta$ vektor kovariancia mátrixa az $A^*\Sigma A$ mátrix. Valóban az $\eta$ kovariancia mátrixának $p$-edik sorában és $q$-ik oszlopában álló elem $\mathrm{Cov}\,(\eta_p,\eta_q)=\sum\limits_{u=1}^{l}\sum\limits_{v=1}^{l}a_{u,p}\mathrm{Cov}\,(\xi_u,\xi_v)a_{v,q}$, $1\le p,q\le l$, és ez az $A^*\Sigma A$ mátrix megfelelő eleme.

Egy többdimenziós normális eloszlású valószínűségi vektornak (jelölje a dimenziót a $k$ szám) szám akkor és csak akkor van sűrűségfüggvénye, ha ennek a vektornak a $\Sigma$

kovarianciamátrixa invertálható. Megjegyezzük, hogy az a tény, hogy a $\Sigma$ kovariancia mátrix invertálható ekvivalens azzal, hogy pozitív definit, és nem pusztán pozitív szemidefinit. Ezt legegyszerűbben onnan látható, hogy $\Sigma = U\Lambda U^*$ alakban írható egy $U$ unitér és egy $\Lambda$ nem-negatív elemeket tartalmazó diagonális mátrix-szal. A $\Sigma$ mátrixnak mind az invertálhatósága mind a (szigorúan) pozitív definit volta azzal ekvivalens, hogy a fenti $\Lambda$ diagonáli mátrix nem tartalmaz a diagonálisban nullát. Ha a normális eloszlású véletlen vektor kovarianciamátrixa egy invertálható $\Sigma$ mátrix várható értéke egy $M$ vektor akkor e vektor sűrűségfüggvényét a

$$f(x_1,\dots,x_k) = \frac{1}{\sqrt{(2\pi)^k \operatorname{Det}\Sigma}} \exp\left\{-\frac{(x-M)\Sigma^{-1}(x-M)^*}{2}\right\} \qquad (++)$$

képlet adja meg, ahol $x = (x_1,\dots,x_k)$, $\Sigma^{-1}$ a $\Sigma$ mátrix inverzét, $\operatorname{Det}\Sigma$ pedig a $\Sigma$ mátrix determinánsát jelöli. Érdemes ezt a képletet összehasonlítani az egydimenziós normális sűrűségfüggvénnyel. Az egydimenziós esetben az exponensben szereplő $-\frac{x^2}{\sigma^2}$ kifejezésnek a többdimenziós esetben az $-\frac{x\Sigma x^*}{2}$ kvadratikus forma felel meg, a $\frac{1}{\sqrt{2\pi}\sigma}$ normáló faktornak pedig az $\frac{1}{\sqrt{(2\pi)^k \operatorname{Det}\Sigma}}$ kifejezés felel meg.

Az, hogy a $\Sigma$ mátrix nem invertálható, ekvivalens azzal, hogy $\operatorname{Det}\Sigma = 0$. Ekkor a $\Sigma = A^*A$ egyenletet kielégítő $A$ mátrixra, $\operatorname{Det} A = 0$, azaz az $A$ mátrix sem invertálható, tehát létezik olyan $x = (x_1,\dots,x_k)$ vektor, melyre $xA = 0$. Egy $\Sigma$ kovarianciamátrixú és $M$ várható értékű véletlen vektor eloszlása megegyezik egy $(\bar\xi_1,\dots,\bar\xi_k) = (\eta_1,\dots,\eta_k)A + M$ vektor eloszlásával, ahol $\Sigma = A^*A$, és $\eta_1,\dots,\eta_k$ független, standard normális eloszlású valószínűségi változók. Ha a $\Sigma$ és ezért az $A$ mátrix nem invertálható, akkor létezik olyan $(x_1,\dots,x_k)$ vektor, melyre $\sum\limits_{p=1}^{k} x_p(\bar\xi_p - M_p) = 0$ egy valószínűséggel. Ez azt jelenti, hogy a $(\bar\xi_1,\dots,\bar\xi_k)$, illetve a vele azonos eloszlású $(\xi_1,\dots,\xi_k)$ véletlen vektor egy valószínűséggel egy hipersíkra van koncentrálva, ezért nincs sűrűségfüggvénye.

Ha a normális eloszlású valószínűségi változó $\Sigma$ kovarianciamátrixa invertálható, akkor annak sűrűségfüggvényét kiszámíthtatjuk egyszerű integráltranszformáció segítségével.

Egy független standard normális eloszlású valószínűségi vektor sűrűségfüggvénye $\varphi(y) = \varphi(y_1,\dots,y_k) = \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{(y,y)}{2}\right\}$, ahol $y = (y_1,\dots,y_k)$, egy $\Sigma$ kovarianciamátrixú és $M = (M_1,\dots,M_k)$ várható értékű normális eloszlású valószínűségi vektort definiálhatunk $\xi = (\xi_1,\dots,\xi_k) = \eta A + M$ alakban, ahol $A^*A = \Sigma$. Alkalmazva az $x = yA + M$ transzformációt $x = (x_1,\dots,x_k)$ jelöléssel kapjuk, hogy tetszőleges mérhető $B \subset R^k$ halmazra

$$\begin{aligned} P(\xi \in B) &= P\left(\eta \in (B-M)A^{-1}\right) = \int_{(y_1,\dots,y_k)\in(B-M)A^{-1}} \varphi(y)\,dy \\ &= \frac{1}{\det A} \int_{(x_1,\dots,x_k)\in B} \varphi\left((x-M)A^{-1}\right)\,dx \end{aligned}$$

alakú, ahol $\mathrm{Det}\, A$ az $x = yA + M$ leképezés Jacobian-ja.

E formulából kiolvasható, hogy a vizsgált normális eloszlású valószínűségi változó sűrűségfüggvénye a $\frac{1}{\det A}\varphi\left((x-M)A^{-1}\right)$ függvény. Annak érdekében, hogy bebizonyítsuk a (++) formulát vegyük észre, hogy mivel $\Sigma = A^* A$, ezért $\mathrm{Det}\,\Sigma = (\mathrm{Det}\, A)^2$, továbbá

$$\begin{aligned}
\varphi\left((x-M)A^{-1}\right) &= \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{((x-M)A^{-1},(x-M)A^{-1})}{2}\right\} \\
&= \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{(x-M)A^{-1}\left(A^{-1}\right)^*(x-M)^*}{2}\right\} \\
&= \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{(x-M)(A^*A)^{-1}(x-M)^*}{2}\right\} \\
&= \frac{1}{(2\pi)^{k/2}} \exp\left\{-\frac{(x-M)\Sigma^{-1}(x-M)^*}{2}\right\}.
\end{aligned}$$

A több-dimenziós normális eloszlás eloszlás tárgyalását két nem kötelező házi feladat megfogalmzazásával fejezzük meg.

Egyrészt megjegyezzük, hogy egy több-dimenziós normális eloszlású valószínűségi vektor koordinátái normális eloszlásúak. Viszont ez az állítás nem fordítható meg. Létezik olyan többdimenziós valószínűségi vektor, melynek koordinátái normális eloszlásúak, de a véletlen vektor nem normális eloszlású. Egy ilyen vektorra ad példát a következő nem kötelező házifeladatként megfogalmazott példa.

Definiáljuk a következő $(\Omega, \mathcal{B}, \mathbf{P})$ valószínűségi mezőt: $\Omega = [0,1]$, $\mathcal{B}$ a Borel $\sigma$-algebra $[0,1]$-en, és $\mathbf{P}$ a Lebesgue mérték. Definiáljuk a következő $\xi$ és $\eta$ valószínűségi változókat ezen a mezőn: $\xi(x) = \Phi^{-1}(x)$,

$$\eta(x) = \begin{cases} \xi(1-x) & \text{ha } 0 \le x < \frac{1}{2} \\ \xi\left(x - \frac{1}{2}\right) & \text{ha } \frac{1}{2} \le x \le 1 \end{cases}.$$

Lássuk be, hogy $\xi$ és $\eta$ normális eloszlású valószínűségi változók, de a $(\xi, \eta)$ vektor nem normális eloszlású valószínűségi vektor.

*Segítség:* Egy lehetséges magyarázat arra, hogy a $(\xi, \eta)$ vektor nem normális az, hogy $P(\xi + \eta = 0) = \frac{1}{2}$.

Végül az utolsó feladat.

Legyen $A$ tetszőleges $k \times l$ téglalap mátrix, (azaz nem tesszük fel, hogy $k = l$), $\eta_1, \dots, \eta_l$ független, standard normális eloszlású valószínűségi változók. Ekkor az $(\xi_1, \dots, \xi_k) = (\eta_1, \dots, \eta_l)A$ véletlen vektor nulla várható értékű normális eloszlású valószínűségi változó.

*Segítség:* Írjuk fel a $(\xi_1, \dots, \xi_k)$ véletlen vektor karakterisztikus függvényét. Lássuk be, hogy ez egy $\Sigma = A^*A$ kovarianciamátrixú nulla várható értékű normális eloszlású véletlen vektor. Speciálisan, $A^*A$ pozitív szemidefinit mátrix az általános (téglalapmátrixokra is érvényes) esetben.

## Kiegészítés a december 7-i szemináriumhoz

Egy többdimenziós normális eloszlás sűrűségfüggvényének kiszámítása során kiderült, hogy jobban meg kell értenünk egy a többváltozós függvény transzformáltjának az integrálját kifejező formulát. Részletesebben megfogalmazva a következő kérdésről van szó: Legyen $\mathbf{T}(x_1, \dots, x_k) = (y_1, \dots, y_k) \in R^k$ az $R^k$ euklideszi tér egy "szép" transzformációja önmagába, és legyen $f(y_1, \dots, y_k)$ $k$-változós függvény a képtéren. Definiáljuk ennek az $f(y_1, \dots, y_k)$ függvénynek a $\mathbf{T}$ transzformáció szerinti $g(x_1, \dots, x_k)$ ősképét a $g(x_1, \dots, x_k) = f(y_1, \dots, y_k)$, ha $(y_1, \dots, y_k) = \mathbf{T}(x_1, \dots, x_k)$ képlet segítségével. Azt akarjuk megérteni, hogy hogyen lehet az $f(y_1, \dots, y_k)$ függvény integrálját azaz az

$$\int_{R^k} f(y_1, \dots, y_k)\, dy_1 \dots dy_k$$

kifejezést az eredeti $(x_1, \dots, x_k)$ téren a $g(x_1, \dots, x_k)$ alkalmas függvényének az integráljaként kiszámítani.

Az eredmény megfogalmazása érdekében definiáljuk az $(y_1, \dots, y_k) = \mathbf{T}(x_1, \dots, x_k)$ Jacobian-ját. Koordinátánként kiírva az $R^k$ $k$-dimenziós térnek a $k$-dimenziós térbe való a $\mathbf{T}$ leképezését a következő jelölést alkalmazhatjuk: $y_j = \mathbf{T}_j(x_1, \dots, x_k)$, $j = 1, \dots, k$. Tekintsük azt az $R^k$ téren definiált $A(x_1, \dots, x_k)$ $k \times k$ méretű mátrixot, melynek a $j$-ik sorának és $i$-ik oszlopában szereplő eleme a $\dfrac{\partial \mathbf{T}_j(x_1, \dots, x_k)}{\partial x_i}$ parciális derivált az $(x_1, \dots, x_k)$ pontban. A $\mathbf{T}$ transzformáció $\mathcal{J}(\mathbf{T})(x_1, \dots, x_k)$ Jacobian-ja az $(x_1, \dots, x_k)$ pontban az $A(x_1, \dots, x_k)$ mátrix determinánsának az abszolut értéke. Mielőtt az alább tárgyalandó formulát megfogalmaznánk, értsük meg az $A(x_1, \dots, x_k)$ mátrix és $\mathcal{J}(\mathbf{T})(x_1, \dots, x_k)$ Jacobian szemléletes tartalmát.

Adva egy $(x_1, \dots, x_k)$ pont az $R^k$ $k$-dimenziós térben, és kis $h_1, \dots, h_k$ számok, akkor

$$\begin{aligned} \mathbf{T}(x_1 + h_1, \dots, x_k + h_k) - \mathbf{T}(x_1, \dots, x_k) &= (h_1, \dots, h_k) A(x_1, \dots, x_k) \\ &\quad + \text{elhanyagolhatóan kis hiba}, \end{aligned} \tag{1}$$

azaz az $A(x_1, \dots, x_k)$ mátrix a számegyenesen értelmezett függvények deriváltjának természetes általánosítása. Ennek következménye, hogy a $h_1 \cdots h_k$ térfogatú $[x_1, x_1 + h_1] \times \cdots \times [x_k, x_k + _k]$ téglatest képe a $\mathbf{T}$ transzformáció hatására jó közelítéssel egy $\mathcal{J}(\mathbf{T})(x_1, \dots, x_k) h_1 \cdots h_k$ térfogatú tartomány. Ugyanis az (1) formula alapján ez a tartomány jó közelítéssel a

$$T(x_1, \dots, x_k) + ([0, h_1] \times \cdots \times [0, h_k])\, A(x_1, \dots, x_k)$$

a $T(x_1, \dots, x_k)$ pontba eltolt $h_1 \dots h_k \mathcal{J}(\mathbf{T})(x_1, \dots, x_k)$ térfogatú parallelepipedon.

Ez a geometriai kép, illetve a többdimenziós integrál definiciója (felosztjuk a $\mathbf{B} \subset R^k$ teret kis átmérőjű diszjunkt tartományok uniójára, ezen tartományok térfogatát megszorozzuk az integrálandó függvény értékével e tartomány valamelyik pontjában, majd összegezünk) sugallja az alábbi formulát, melyet ennek a heurisztikus érvelésnek a finomításával (enyhe feltételek kikötése mellett) be lehet bizonyítani.

$$\begin{aligned} &\int f(y_1,\dots,y_k)\,dy_1\dots dy_k \\ &= \int \frac{g(x_1,\dots,x_k)}{\sum\limits_{(z_1,\dots,z_k)\colon\ \mathbf{T}(z_1,\dots,z_k)=\mathbf{T}(x_1,\dots,x_k)} \dfrac{1}{\mathcal{J}(\mathbf{T})(z_1,\dots,z_k)}}\,dx_1\dots,x_k. \end{aligned} \tag{2}$$

Valójában a többdimenziós normális eloszlás sűrűségfüggvény kiszámításában ennek az eredménynek csak egy speciális egyszerű esetét használtuk. Egy lineáris transzformációt $\mathbf{T}x = xB + \bar{M}$ alakú lineáris transzformációt tekintettünk. (Esetünkben egy $B = A^{-1}$ alakú mátrixot, egy alkalmas mátrix inverzét tekintettük.) Ebben az esetben a (2) formula sokkal egyszerűbb alakú. Ekkor a Jacobian nem függ az $(x_1,\dots,x_k)$ ponttól, hanem $|\mathrm{Det}\,B|$-vel egyenlő. Az $f(y_1,\dots,y_k)$ függvény $(x_1,\dots,x_k)$ ősképe is nagyon egyszerűen kiszámolható, $g(x_1,\dots,x_k) = f((x_1,\dots,x_k)B + \bar{M})$. Továbbá mivel a fenti $T$ transzformációnak egyetlen ősképe van, a (2) formula jobboldalán a nevező sokkal egyszerűbb ebben az esetben. Ez a

$$\frac{1}{\sum\limits_{(z_1,\dots,z_k)\colon\ \mathbf{T}(z_1,\dots,z_k)=\mathbf{T}(x_1,\dots,x_k)} \dfrac{1}{\mathcal{J}(\mathbf{T})(z_1,\dots,z_k)}} = \mathcal{J}(\mathbf{T})(x_1,\dots,x_k) = |\mathrm{Det}\,B|$$

kifejezés.