Uzupełnienia z algebry

Adam Ćmiel

2 października 2023

Spis treści

1 Przestrzenie wektorowe i wektory losowe **2**
1.1 Wybrane elementy teorii przestrzeni unitarnych 2
1.2 Wektory losowe i ich struktura kowariancyjna 15

2 Rozkład normalny na przestrzeni wektorowej **33**
2.1 Formy kwadratowe wektorów o rozkładzie normalnym 38
2.2 Rozkład warunkowy wektora losowego o rozkładzie normalnym 42
2.3 Funkcja gęstości rozkładu normalnego na przestrzeni wektorowej 45

Rozdział 1

Przestrzenie wektorowe i wektory losowe

Istotą każdego dobrze postawionego problemu statystycznego jest precyzyjne wyspecyfikowanie rodziny rozkładów prawdopodobieństwa na rozważanej przestrzeni prób. Analizując strukturę rozważanej rodziny rozkładów prawdopodobieństwa, warto wyróżnić te charakterystyki rozkładów, które mogą być opisane bez odwoływania się do przyjętego układu współrzędnych (tzw. aspekty geometryczne) i te, dla których nie jest to możliwe. Pomocna jest tu teoria skończenie wymiarowych przestrzeni wektorowych wyposażonych w iloczyn skalarny, której elementy omówimy w podrozdziale 1.1. Niniejszy rozdział w całości oparty jest na książce [2] (rozdziały 1 oraz 2).

1.1 Wybrane elementy teorii przestrzeni unitarnych

W statystycznej analizie wielowymiarowej typową pojedynczą obserwacją jest p-elementowy wektor z przestrzeni unitarnej $\mathbb{R}^p$. Próba prosta, czyli n-elementowy ciąg niezależnych obserwacji (wektorów losowych), tworzy macierz losową, której każdą realizację można potraktować jako reprezentację pewnego odwzorowania liniowego. Widać więc, że przestrzeń odwzorowań liniowych przestrzeni $\mathbb{R}^p$ w przestrzeń $\mathbb{R}^n$ jest naturalnym kandydatem na przestrzeń prób, dlatego znaczną część tego podrozdziału poświęcimy charakteryzacji przestrzeni odwzorowań liniowych określonych na skończenie wymiarowej przestrzeni unitarnej. Pominiemy znane z kursu algebry liniowej podstawowe pojęcia i fakty i przytoczymy jedynie mniej znane elementy teorii przestrzeni unitarnych, wykorzystywane w niniejszej pracy.

Niech $(V, (\cdot,\cdot))$ oraz $(W, [\cdot,\cdot])$ będą skończenie wymiarowymi rzeczywistymi przestrzeniami unitarnymi, natomiast niech $\mathcal{L}(V, W)$ będzie zbiorem wszystkich odwzorowań liniowych prowadzących z przestrzeni V do W.

Przez $\mathcal{R}(A)$ i $\mathcal{N}(A)$ będziemy oznaczać odpowiednio obraz i jądro odwzorowania A.

Ponieważ $\mathcal{L}(V, W)$ jest przestrzenią wektorową, wszystkie ogólne wyniki dotyczące przestrzeni liniowych odnoszą się również do przestrzeni $\mathcal{L}(V, W)$. Jednakże wektory w $\mathcal{L}(V, W)$, czyli liniowe transformacje, mają wiele interesujących własności, których nie mają inne wektory. Mając dwie transformacje liniowe $A \in \mathcal{L}(V_1, V_2)$ i $B \in \mathcal{L}(V_2, V_3)$ możemy zdefiniować złożenie $BA \in \mathcal{L}(V_1, V_3)$ wzorem $BA(x) = B(A(x))$, które nazywamy produktem (iloczynem) transformacji A i B. Szczególnie interesująca jest przestrzeń $\mathcal{L}(V, V)$ endomorfizmów przestrzeni V, gdzie oba złożenia AB i BA są możliwe, a gdy przekształcenia są odwracalne, możemy w zbiorze endomorfizmów $\mathcal{L}(V, V)$ zadać strukturę grupy. Drugi ważny przypadek to przestrzeń $\mathcal{L}(V, R)$ rzeczywistych funkcji (form) liniowych określonych na V.

Szczególną rolę wśród endomorfizmów $\mathcal{L}(V, V)$ odgrywają tzw. projekcje. Jeśli skończenie wymiarowa przestrzeń V jest sumą prostą $V = M \oplus N$ podprzestrzeni M oraz N, i jeśli $x = y + z$, gdzie $y \in M$ i $z \in N$, to jednoznacznie wyznaczony wektor y nazywamy projekcją (rzutem) wektora x na podprzestrzeń M wzdłuż N.

Można pokazać [2], że prawdziwe jest

Twierdzenie 1.1. *Funkcja P przypisująca dowolnemu wektorowi $x \in V$ jego rzut $y = Px$ na M wzdłuż N jest transformacją liniową spełniającą warunki*

(i) $\mathcal{R}(P) = M$, $\mathcal{N}(P) = N$,

(ii) $P^2 = P$.

Ale i odwrotnie, jeżeli endomorfizm $A \in \mathcal{L}(V, V)$ spełnia warunek idempotentności $A^2 = A$, to $\mathcal{R}(A) \oplus \mathcal{N}(A) = V$ i A jest projekcją na $\mathcal{R}(A)$ wzdłuż $\mathcal{N}(A)$.

Definiując w przestrzeni wektorowej iloczyn skalarny wektorów nadamy sens takim pojęciom geometrycznym jak długość, kąt czy ortogonalność.

Definicja 1.1. *Iloczynem skalarnym określonym na rzeczywistej przestrzeni wektorowej V nazywamy rzeczywistą funkcję $(\cdot, \cdot)$ określoną na $V \times V$ spełniającą warunki:*

(i) $(x, y) = (y, x)$ (symetria),

(ii) $(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y)$ (liniowość),

(iii) $(x, x) \geqslant 0$ oraz $(x, x) = 0 \Leftrightarrow x = 0$ (dodatnia określoność).

W przestrzeni wektorowej z iloczynem skalarnym (przestrzeni unitarnej) definiujemy normę wektora $\|x\| = (x, x)^{1/2}$ i odległość $\|x - y\|$ pomiędzy wektorami x i y. Zakładamy znajomość takich pojęć jak: ortogonalność wektorów, ortogonalność zbiorów wektorów, baza ortonormalna, procedura ortonormalizacji Grama-Schmidta. Z uwagi na dalsze zastosowania przypomnimy wersję twierdzenia o postaci formy liniowej (funkcjonału liniowego) na przestrzeni unitarnej V.

Twierdzenie 1.2. *Jeśli $(V, (\cdot, \cdot))$ jest skończenie wymiarową rzeczywistą przestrzenią unitarną i $f \in \mathcal{L}(V, R)$, to istnieje wektor $x_0 \in V$ taki, że $f(x) = (x_0, x)$ dla $x \in V$. Odwrotnie, funkcja postaci $(x_0, \cdot)$ jest funkcją liniową określoną na przestrzeni V dla każdego $x_0 \in V$.*

Dowód. Niech $x_1, \ldots, x_n$ będzie bazą ortonormalną przestrzeni V oraz niech $\alpha_i = f(x_i)$ dla $i = 1, \ldots, n$. Dla $x_0 = \sum \alpha_i x_i$ oczywistym jest, że $(x_0, x_j) = \alpha_j = f(x_j)$. Ponieważ dwie funkcje liniowe f oraz $(x_0, \cdot)$ przyjmują te same wartości dla elementów bazowych, są sobie równe. Stąd $f(x) = (x_0, x)$ dla $x \in V$. Implikacja w drugą stronę jest oczywista. □

Definicja 1.2. *Odwzorowaniem sprzężonym do danego odwzorowania $A \in \mathcal{L}(V, W)$ nazywamy jednoznacznie wyznaczone odwzorowanie $A' \in \mathcal{L}(W, V)$ takie, że dla dowolnych wektorów $v \in V$ i $w \in W$ spełniona jest równość $[w, Av] = (A'w, v)$.*

W dalszej części pracy wielokrotnie korzystali będziemy z następujących własności jądra i obrazu odwzorowania liniowego.

Twierdzenie 1.3. *Niech dane będzie odwzorowanie $A \in \mathcal{L}(V, W)$ oraz odwzorowanie do niego sprzężone $A' \in \mathcal{L}(W, V)$. Wówczas:*

(i) $\mathcal{R}(A) = (\mathcal{N}(A'))^{\perp}$ ($\perp$ *oznacza dopełnienie ortogonalne*),

(ii) $\mathcal{R}(A) = \mathcal{R}(AA')$,

(iii) $\mathcal{N}(A) = \mathcal{N}(A'A)$,

(iv) $\dim(\mathcal{R}(A)) = \dim(\mathcal{R}(A'))$.

Dowód. Równość (i) jest równoważna równości $(\mathcal{R}(A))^{\perp} = \mathcal{N}(A')$, a ta wynika z następującego ciągu równoważności:

$$w \in \mathcal{N}(A') \Leftrightarrow \ \forall v \in V : \ (v, A'w) = 0 \Leftrightarrow \ \forall v \in V : \ [Av, w] = 0 \Leftrightarrow w \in \mathcal{R}(A)^{\perp}.$$

Aby udowodnić część (ii), zauważmy, że inkluzja $\mathcal{R}(AA') \subseteq \mathcal{R}(A)$ jest oczywistą konsekwencją monotoniczności operacji brania obrazu, wystarczy więc wykazać, że prawdziwa jest inkluzja przeciwna. Jeśli $w \in \mathcal{R}(A)$, to istnieje $v \in V$ taki, że $w = Av$. Wektor v można zapisać w postaci $v = v_1 + v_2$, gdzie $v_1 \in \mathcal{R}(A')$ oraz $v_2 \in (\mathcal{R}(A'))^\perp$. Z udowodnionego już punktu (i) mamy $(\mathcal{R}(A'))^\perp = \mathcal{N}(A)$, więc $Av_2 = 0$. Ponieważ $v_1 \in \mathcal{R}(A')$, to istnieje $u \in W$ taki, że $v_1 = A'u$, stąd $w = Av = Av_1 = AA'u$, czyli $w \in \mathcal{R}(AA')$. Punkt (iii) udowodnimy, pokazując dwie inkluzje. Jeżeli $Av = 0$, to $A'Av = 0$, więc $\mathcal{N}(A) \subseteq \mathcal{N}(A'A)$. Z drugiej strony, jeśli $A'Av = 0$, to wówczas $0 = (v, A'Av) = [Av, Av]$, więc $Av = 0$, czyli $\mathcal{N}(A'A) \subseteq \mathcal{N}(A)$, co kończy dowód części (iii). Część (iv) wykażemy, korzystając ze znanego z kursu algebry faktu

$$\dim(V) = \dim(\mathcal{R}(A)) + \dim(\mathcal{N}(A))$$

oraz przedstawienia przestrzeni V w postaci sumy prostej $V = \mathcal{R}(A') \oplus \mathcal{R}(A')^\perp$. Z punktu (i) otrzymujemy $V = \mathcal{R}(A') \oplus \mathcal{N}(A)$, skąd wynika, że

$$\dim(V) = \dim(\mathcal{R}(A')) + \dim(\mathcal{N}(A)).$$

Zestawiając ze sobą powyższe równości, otrzymujemy $\dim(\mathcal{R}(A')) = \dim(\mathcal{R}(A))$. $\square$

Z kursu algebry wiadomo (por. tw. 1.1), że każdy idempotentny endomorfizm $P \in \mathcal{L}(V, V)$ (czyli spełniający warunek $P^2 = P$) jest projekcją na $\mathcal{R}(P)$ wzdłuż $\mathcal{N}(P)$, to znaczy $V = \mathcal{R}(P) \oplus \mathcal{N}(P)$ i każdy wektor $v \in V$ da się jednoznacznie przedstawić w postaci $v = v_1 + v_2$, $v_1 \in \mathcal{R}(P)$, $v_2 \in \mathcal{N}(P)$. Jednoznacznie wyznaczony wektor $v_1 \in \mathcal{R}(P)$ nazywamy projekcją (rzutem) wektora v na $\mathcal{R}(P)$ i oznaczamy $v_1 = Pv$. Jeśli dodatkowo endomorfizm P jest samosprzężony, czyli $P = P'$, to z punktu (i) twierdzenia 1.3 wynika, że $\mathcal{R}(P)^\perp = \mathcal{N}(P)$ i P jest wtedy ortogonalną projekcją na $\mathcal{R}(P)$.

Niezwykle użyteczne w charakteryzacji struktury dowolnych odwzorowań liniowych są szczególne odwzorowania liniowe zwane iloczynami zewnętrznymi.

Definicja 1.3. *Iloczynem zewnętrznym* (ang. *outer product*) *wektorów* $v \in (V, (\cdot,\cdot))$ *oraz* $w \in (W, [\cdot,\cdot])$ *nazywamy odwzorowanie liniowe* $w\Box v \in \mathcal{L}(V, W)$ *określone wzorem*

$$w\Box v \colon V \ni x \to (v, x)w \in W.$$

Z powyższej definicji wynika, że iloczyn zewnętrzny wektorów określony na przestrzeni V zależy od iloczynu skalarnego, w który wyposażona jest ta przestrzeń. Ponadto $\mathcal{R}(w\Box v) = \mathrm{span}\{w\}$ oraz $\mathcal{N}(w\Box v) = (\mathrm{span}\{v\})^\perp$. Pozostałe podstawowe własności iloczynu zewnętrznego sformułujemy w postaci poniższego twierdzenia.

Twierdzenie 1.4. *Niech dane będą wektory* $v_1, v_2 \in (V, (\cdot,\cdot))$ *oraz* $w_1, w_2 \in (W, [\cdot,\cdot])$. *Wówczas dla dowolnych stałych* α_1, α_2 *prawdziwe są następujące równości:*

(i) $w_1 \Box (\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 w_1 \Box v_1 + \alpha_2 w_1 \Box v_2$,

(ii) $(\alpha_1 w_1 + \alpha_2 w_2) \Box v_1 = \alpha_1 w_1 \Box v_1 + \alpha_2 w_2 \Box v_1$,

(iii) $(w_1 \Box v_1)' = v_1 \Box w_1$.

Dowód. Równości (i) i (ii) dowodzimy, wyznaczając wartości odpowiednich odwzorowań dla dowolnego $x \in V$. Na przykład, dla dowodu (i) mamy

$$\begin{aligned}(w_1 \Box (\alpha_1 v_1 + \alpha_2 v_2))x &= (\alpha_1 v_1 + \alpha_2 v_2, x)w_1 = \alpha_1 (v_1, x) w_1 + \alpha_2 (v_2, x) w_1 \\ &= \alpha_1 (w_1 \Box v_1)x + \alpha_2 (w_1 \Box v_2)x.\end{aligned}$$

Dla dowolnych wektorów $x \in V$ oraz $y \in W$ prawdziwy jest ciąg równości

$$[y, (w_1 \Box v_1)x] = [y, (v_1, x)w_1] = [y, w_1](v_1, x) = ([y, w_1]v_1, x) = ((v_1 \Box w_1)y, x),$$

co z kolei dowodzi, że odwzorowanie $v_1 \Box w_1 \in \mathcal{L}(W, V)$ jest odwzorowaniem sprzężonym do $w_1 \Box v_1$. $\Box$

Kolejna własność odwzorowań postaci $w \Box v$, z której w istotny sposób skorzystamy w dalszych rozważaniach, jest następująca. Każde odwzorowanie $A \in \mathcal{L}(V, W)$ można przedstawić jako kombinację liniową iloczynów zewnętrznych wektorów bazowych przestrzeni V oraz W. W szczególności, każde odwzorowanie jednowymiarowe jest iloczynem zewnętrznym pewnych niezerowych wektorów $v \in V$ oraz $w \in W$. Bardziej formalnie podaje to poniższe twierdzenie.

Twierdzenie 1.5. *Niech $\{v_1, v_2, \ldots, v_m\}$ i $\{w_1, w_2, \ldots, w_n\}$ będą bazami ortonormalnymi odpowiednio przestrzeni $(V, (\cdot, \cdot))$ oraz $(W, [\cdot, \cdot])$. Wówczas odwzorowania $w_i \Box v_j$, $i = 1, \ldots, n$, $j = 1, 2, \ldots, m$ tworzą bazę przestrzeni $\mathcal{L}(V, W)$.*

Dowód. Liczność zbioru $\{w_i \Box v_j : \ i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, m\}$ jest równa wymiarowi przestrzeni $\mathcal{L}(V, W)$, który wynosi mn (por. [2], str. 7), dlatego wystarczy wykazać, że jego elementy są liniowo niezależne. Załóżmy, że

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} w_i \Box v_j \equiv 0.$$

Obliczając wartość odwzorowania A dla każdego wektora bazowego v_k przestrzeni V, otrzymujemy

$$\sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} (w_i \Box v_j) v_k = \sum_{i=1}^{n} \sum_{j=1}^{m} \alpha_{ij} (v_j, v_k) w_i = \sum_{i=1}^{n} \alpha_{ik} w_i = 0.$$

Ponieważ wektory w_i są liniowo niezależne, to $\alpha_{ik} = 0$ dla $i = 1, 2, \ldots, n$ oraz dowolnego $k \in \{1, 2, \ldots, m\}$, co chcieliśmy pokazać. $\Box$

Każda n wymiarowa przestrzeń rzeczywista U jest izomorficzna z przestrzenią współrzędnych $\mathbb{R}^n$, w której jest określony naturalny iloczyn skalarny. Można więc, wykorzystując ten izomorfizm, zadać w U iloczyn skalarny tak, aby dowolna wybrana wcześniej baza przestrzeni U stała się bazą ortonormalną. W szczególności przestrzenią U może być $\mathcal{L}(V, W)$ z bazą iloczynów zewnętrznych jak w twierdzeniu 1.5. Wówczas, zapisując dowolne odwzorowania $A, B \in \mathcal{L}(V, W)$ w postaci

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} w_i \Box v_j \qquad \text{oraz} \qquad B = \sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij} w_i \Box v_j,$$

rozważana baza będzie ortonormalna, gdy określimy w $\mathcal{L}(V, W)$ iloczyn skalarny $\langle \cdot, \cdot \rangle$ wzorem

$$\langle A, B \rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} b_{ij}. \tag{1.1}$$

Fakt, że do konstrukcji bazy przestrzeni $\mathcal{L}(V, W)$ użyliśmy ortonormalnych baz przestrzeni V oraz W powoduje, że pomiędzy rozważanymi iloczynami skalarnymi zachodzi fundamentalny związek opisany w poniższym twierdzeniu.

Twierdzenie 1.6. *Niech dane będą przestrzenie wektorowe* $(V, (\cdot, \cdot))$, $(W, [\cdot, \cdot])$ *oraz* $(\mathcal{L}(V, W), \langle \cdot, \cdot \rangle)$. *Dla dowolnego odwzorowania* $A \in \mathcal{L}(V, W)$ *oraz wektorów* $v \in V$, $w \in W$ *zachodzi równość*

$$\langle A, w \Box v \rangle = [w, Av].$$

Dowód. Niech $\{v_1, v_2, \ldots, v_m\}$ i $\{w_1, w_2, \ldots, w_n\}$ będą bazami ortonormalnymi odpowiednich przestrzeni oraz niech a_{ij} będą współrzędnymi odwzorowania A w bazie $\{w_i \Box v_j : \ i = 1, 2, \ldots, n, \ j = 1, 2, \ldots, m\}$. Wtedy

$$[w_k, Av_l] = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} [w_k, (w_i \Box v_j) v_l] = \sum_{i=1}^{n} \sum_{j=1}^{m} a_{ij} [w_k, w_i](v_j, v_l) = a_{kl},$$

więc odwzorowanie A możemy przedstawić w postaci

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} [w_i, Av_j] w_i \Box v_j.$$

Co więcej, naturalnie spełniona jest równość

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} \langle A, w_i \Box v_j \rangle w_i \Box v_j,$$

ponieważ rozważana baza przestrzeni $\mathcal{L}(V, W)$ jest ortonormalna. Widzimy więc, że $[w_i, Av_j] = \langle A, w_i \Box v_j \rangle$ dla $i = 1, 2 \ldots, n$, $j = 1, 2, \ldots, m$, a stąd $[w, Av] = \langle A, w \Box v \rangle$ dla każdego $v \in V$ oraz $w \in W$, bowiem po obu stronach równości występują odwzorowania dwuliniowe. $\Box$

Wniosek 1.1. *W szczególności, przyjmując $A = w'\Box v'$ dla pewnych $w' \in W$, $v' \in V$, dla każdego $v \in V$ oraz każdego $w \in W$ zachodzi równość*

$$\langle w'\Box v', w\Box v\rangle = [w, (w'\Box v')v] = [w, (v', v)w'] = [w', w](v', v). \tag{1.2}$$

Ponadto odwzorowanie $\langle \cdot, \cdot \rangle$ jest jedynym iloczynem skalarnym w przestrzeni $\mathcal{L}(V, W)$ o tej własności. Załóżmy, że istnieje iloczyn skalarny $\langle \cdot, \cdot \rangle_1$ taki, że

$$\langle w'\Box v', w\Box v\rangle_1 = [w', w](v, v), \qquad v, v' \in V, \quad w, w' \in W.$$

Iloczyn skalarny jest funkcją dwuliniową, dlatego z równości

$$\langle w'\Box v', w\Box v\rangle_1 = \langle w'\Box v', w\Box v\rangle,$$

zachodzącej w szczególności dla elementów bazowych przestrzeni $\mathcal{L}(V, W)$, wynika, że dla dowolnych odwzorowań $A, B \in \mathcal{L}(V, W)$ zachodzi równość $\langle A, B\rangle_1 = \langle A, B\rangle$, a więc rozpatrywane iloczyny skalarne są równe.

Przykład 1.1. *Rozważmy przestrzenie $V = \mathbb{R}^n$ oraz $W = \mathbb{R}^m$ ze standardowymi iloczynami skalarnymi. Niech $\{e_1, e_2, \ldots, e_n\}$, $\{f_1, f_2, \ldots, f_m\}$ będą standardowymi bazami przestrzeni V oraz W. Niech $\mathcal{M}_{m\times n}$ oznacza przestrzeń macierzy rzeczywistych o m wierszach i n kolumnach. Każda macierz $M \in \mathcal{M}_{m\times n}$ definiuje odwzorowanie $A\colon V \to W$ takie, że kolumny macierzy M to współrzędne obrazów wektorów bazowych przestrzeni V w przestrzeni W. Wiemy, że zbiór $\{f_i\Box e_j, i = 1, 2\ldots, m, j = 1, 2\ldots n\}$ tworzy bazę przestrzeni $\mathcal{L}(V, W)$, więc odwzorowanie A możemy zapisać w postaci*

$$A = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} f_i\Box e_j,$$

gdzie a_{ij} to współrzędne odwzorowania A w rozważanej bazie. Wówczas

$$Ae_k = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}(f_i\Box e_j)e_k = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}e_j' e_k f_i = \sum_{i=1}^{m} a_{ik} f_i$$

dla $k = 1, 2, \ldots n$, zatem parametry a_{ik} są współrzędnymi obrazu wektora e_k przez odwzorowanie A w przestrzeni W, a stąd już wynika, że parametry a_{ij} są elementami macierzy M. Z racji tego dla dowolnych macierzy $M, N \in \mathcal{M}_{m\times n}$ postaci $M = [m_{ij}]$ oraz $N = [n_{ij}]$ przyjmijmy

$$\langle M, N\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} m_{ij} n_{ij}.$$

Niech $C = MN'$, gdzie przez N' oznaczamy transpozycję macierzy N, wtedy elementy diagonalne kwadratowej macierzy C mają postać

$$c_{ii} = \sum_{j=1}^{n} m_{ij} n_{ij},$$

czyli $\langle M,N\rangle = \sum_i c_{ii}$. *Iloczynem skalarnym macierzy* M *i* N *jest więc ślad macierzy* C*, który będziemy oznaczać przez* $\mathrm{tr}(C)$. *Oczywiście* $\langle M,N\rangle = \mathrm{tr}(MN') = \mathrm{tr}(N'M)$. *Przedstawmy macierz* $M \in \mathcal{M}_{m\times n}$ *w postaci wektora* $\tilde{M} \in \mathbb{R}^{mn}$ *powstałego przez „sklejenie" kolejnych kolumn macierzy* M *w jeden wektor. Tworząc w ten sam sposób wektor* $\tilde{N}$ *odpowiadający dowolnej macierzy* $N \in \mathcal{M}_{m\times n}$*, mamy* $\langle M,N\rangle = \tilde{M}'\tilde{N}$. *Inny sposób kodowania macierzy, zapewniający dla macierzy symetrycznych zgodność iloczynu skalarnego wprowadzonego w* $\mathcal{M}_{m\times n}$ *z klasycznym euklidesowym iloczynem skalarnym w* $\mathbb{R}^{mn}$*, opisany jest w rozdziale 8 monografii Barry* ([1]).

Powyższy przykład można uogólnić na przypadek dowolnych przestrzeni $(V,(\cdot,\cdot))$ oraz $(W,[\cdot,\cdot])$ z bazami ortonormalnymi $\{v_1,v_2,\ldots,v_m\} \subset V$, $\{w_1,w_2,\ldots,w_n\} \subset W$. Wynika z niego, że dla dowolnych odwzorowań $A,B \in \mathcal{L}(V,W)$ współczynniki a_{ij} i b_{ij} w zapisie

$$A = \sum_{i=1}^{n}\sum_{j=1}^{m} a_{ij} w_i \Box v_j \qquad \text{oraz} \qquad B = \sum_{i=1}^{n}\sum_{j=1}^{m} b_{ij} w_i \Box v_j,$$

zestawione w macierze $[A] = [a_{ij}]$, $[B] = [b_{ij}]$ stanowią reprezentacje macierzowe odwzorowań A i B w bazach ortonormalnych $\{v_1,\ldots,v_m\}$ i $\{w_1,\ldots,w_n\}$, a iloczyn skalarny zadany wzorem (1.1) spełnia równość $\langle A,B\rangle = \mathrm{tr}[A][B]'$. Do tego samego wniosku można dojść, korzystając z twierdzenia 1.6, ponieważ $a_{ij} = \langle A, w_i\Box v_j\rangle = [w_i, Av_j]$, gdzie $[w_i, Av_j]$ jest i-tym współczynnikiem wektora Av_j w bazie przestrzeni W. Jeśli zmienimy bazy w przestrzeniach V i W na nowe bazy ortonormalne $\{\tilde{v}_1,\ldots,\tilde{v}_m\}$ i $\{\tilde{w}_1,\ldots,\tilde{w}_n\}$, to macierze przejścia od baz starych do nowych będą macierzami ortonormalnymi $[\tilde{P}]$ i $[\tilde{Q}]$, natomiast przekształcenia $A,B \in \mathcal{L}(V,W)$ będą miały nowe reprezentacje macierzowe $[\tilde{A}] = [Q'AP]$ i $[\tilde{B}] = [Q'BP]$, przy czym $Q' = Q^{-1}$. Ponieważ $\mathrm{tr}[\tilde{A}][\tilde{B}]' = \mathrm{tr}[Q'AP][Q'BP]' = \mathrm{tr}[Q]'[A][P][P]'[B]'[Q] = \mathrm{tr}[Q]'[A][B]'[Q] = \mathrm{tr}[A][B]'$, to iloczyn skalarny $\langle A,B\rangle$ możemy wyznaczyć, korzystając z reprezentacji odwzorowań A i B w dowolnych bazach ortonormalnych przestrzeni $(V,(\cdot,\cdot))$ i $(W,[\cdot,\cdot])$. Postać odwzorowania liniowego wynikająca z twierdzenia 1.5 przyjmuje szczególną formę, gdy rozważane odwzorowanie określone na przestrzeni V jest samosprzężone. Poniżej podajemy wersję twierdzenia spektralnego sformułowaną w języku iloczynów zewnętrznych wektorów. Przez $\mathcal{S}(V)$ będziemy oznaczać zbiór wszystkich odwzorowań samosprzężonych określonych na przestrzeni V (względem zadanego na V iloczynu skalarnego).

Twierdzenie 1.7. (spektralne [2], str. 50) *Niech dane będzie odwzorowanie* $A \in \mathcal{S}(V)$. *Wówczas istnieje ortonormalna baza* $\{x_1,x_2,\ldots x_n\}$ *przestrzeni* V *taka, że odwzorowanie* A *ma następującą reprezentację:*

$$A = \sum_{i=1}^{n} \lambda_i x_i \Box x_i,$$

gdzie $\lambda_1, \lambda_2, \ldots, \lambda_n$ są wartościami własnymi odwzorowania A oraz $Ax_i = \lambda_i x_i$ dla $i = 1, 2, \ldots, n$.

Twierdzenie spektralne wyraża formalnie znany z kursu algebry fakt, że endomorfizm samosprzężony jest diagonalizowalny i bazą, w której ma on taką postać, jest baza zbudowana z ortonormalnych wektorów własnych tego endomorfizmu. Ponieważ wartości własne odwzorowania samosprzężonego nie muszą być różne, wybór bazy ortonormalnej pojawiającej się w reprezentacji spektralnej takiego odwzorowania (wynikającej z powyższego twierdzenia) nie jest jednoznaczny. Zagadnienie to zostało szczegółowo opisane w [3], gdzie podano również zmodyfikowaną wersję twierdzenia spektralnego, dla której taki problem nie istnieje. Wszystkie rezultaty uzyskane w oparciu o przyjętą w tej pracy wersję twierdzenia spektralnego nie zależą od wyboru wspomnianej bazy ortonormalnej.

Uzyskana reprezentacja odwzorowań samosprzężonych posłuży nam w pierwszej kolejności do wyrażenia warunku koniecznego i wystarczającego dodatniej określoności elementów przestrzeni $\mathcal{S}(V)$. Następnie wykorzystamy ją do konstrukcji pierwiastka oraz uogólnionej odwrotności dowolnego odwzorowania $A \in \mathcal{S}(V)$.

Twierdzenie 1.8. *Odwzorowanie $A \in \mathcal{S}(V)$ jest dodatnio określone wtedy i tylko wtedy, gdy wszystkie jego wartości własne są większe od zera.*

Dowód. Samosprzężone odwzorowanie A nazywamy dodatnio określonym w $(V, (\cdot,\cdot))$, gdy dla każdego niezerowego wektora $x \in V$ wartość iloczynu skalarnego (x, Ax) jest dodatnia. Korzystając z twierdzenia 1.7, zapiszmy odwzorowanie A w postaci

$$A = \sum_{i=1}^{n} \lambda_i x_i \Box x_i,$$

gdzie $\{x_1, x_2, \ldots x_n\}$ to odpowiednia baza ortonormalna przestrzeni $(V, (\cdot,\cdot))$. Dla dowolnego wektora $x \in V$ mamy

$$(x, Ax) = (x, \sum_{i=1}^{n} \lambda_i (x_i \Box x_i) x) = \sum_{i=1}^{n} \lambda_i (x, (x_i \Box x_i) x) = \sum_{i=1}^{n} \lambda_i (x, x_i)^2,$$

więc jeżeli $\lambda_i > 0$, $i = 1, 2, \ldots, n$, to $(x, Ax) > 0$. Załóżmy, że istnieje indeks k taki, że $\lambda_k \leqslant 0$. Wówczas dla wektora bazowego x_k zachodzi nierówność

$$(x_k, Ax_k) = \sum_{i=1}^{n} \lambda_i (x_k, x_i)^2 = \lambda_k \leqslant 0$$

i tym samym A nie jest dodatnio określone. W takim razie $(x, Ax) > 0$ dla wszystkich niezerowych $x \in V$ tylko wtedy, gdy wartości λ_i są dodatnie dla $i = 1, 2, \ldots, n$. □

Modyfikując powyższy dowód w oczywisty sposób, można wykazać analogiczną własność dla odwzorowań nieujemnie określonych.

Twierdzenie 1.9. *Niech $A \in \mathcal{S}(V)$ będzie odwzorowaniem nieujemnie określonym. Wówczas istnieje nieujemnie określone odwzorowanie $B \in \mathcal{S}(V)$, dla którego spełniona jest równość $B^2 = A$.*

Dowód. Zapiszmy odwzorowanie A w postaci spektralnej

$$A = \sum_{i=1}^{n} \lambda_i x_i \square x_i,$$

gdzie $\{x_1, x_2, \ldots, x_n\}$ jest bazą ortonormalną przestrzeni $(V, (\cdot, \cdot))$ i określmy odwzorowanie B w następujący sposób:

$$B = \sum_{i=1}^{n} \lambda_i^{1/2} x_i \square x_i.$$

Odwzorowanie B jest poprawnie zdefiniowane, ponieważ $\lambda_i \geqslant 0$ dla $i = 1, 2, \ldots n$. Dla dowolnego $x \in V$ zachodzą równości

$$(x_i \square x_i)(x_j \square x_j)x = (x_j, x)(x_i \square x_i)x_j = (x_j, x)(x_i, x_j)x_i = \delta_{ij}(x_i \square x_j)x,$$

więc

$$\begin{aligned} B^2 &= \left(\sum_{i=1}^{n} \lambda_i^{1/2} x_i \square x_i \right) \left(\sum_{i=j}^{n} \lambda_j^{1/2} x_j \square x_j \right) = \sum_{i=1}^{n} \sum_{i=j}^{n} \lambda_i^{1/2} \lambda_j^{1/2} (x_i \square x_i)(x_j \square x_j) \\ &= \sum_{i=1}^{n} \sum_{i=j}^{n} \lambda_i^{1/2} \lambda_j^{1/2} \delta_{ij} (x_i \square x_j) = \sum_{i=1}^{n} \lambda_i x_i \square x_i = A, \end{aligned}$$

czyli odwzorowanie B spełnia warunki dowodzonego twierdzenia. $\square$

Niech teraz $A \in \mathcal{S}(V)$ będzie odwzorowaniem rzędu $k \leqslant n$, gdzie n to wymiar przestrzeni V. Na mocy twierdzenia 1.7 istnieje baza ortonormalna przestrzeni $\mathcal{R}(A)$ złożona z wektorów własnych $\{x_1, x_2, \ldots, x_k\}$ odwzorowania A oraz odpowiednie stałe $\lambda_i \neq 0$, $i = 1, 2, \ldots, k$ takie, że

$$A = \sum_{i=1}^{k} \lambda_i x_i \square x_i.$$

Uogólnioną odwrotnością odwzorowania $A \in \mathcal{S}(V)$ będziemy nazywać odwzorowanie $A^- \in \mathcal{S}(V)$ dane wzorem

$$A^- = \sum_{i=1}^{k} \frac{1}{\lambda_i} x_i \square x_i. \tag{1.3}$$

Odwzorowanie A^- jest samosprzężone, co jest oczywistą konsekwencją tego, że iloczyn zewnętrzny $x\Box x$ jest samosprzężony (por. twierdzenie 1.4, (iii)). Gdy rząd odwzorowania A wynosi n, wszystkie wartości własne A są różne od zera. Wówczas, postępując podobnie jak w dowodzie twierdzenia 1.9, można pokazać, że A^- jest w rzeczywistości odwzorowaniem odwrotnym do A. Z własności iloczynów zewnętrznych otrzymujemy

$$AA^- = A^-A = \sum_{i=1}^{k} x_i \Box x_i \tag{1.4}$$

oraz $AA^-AA^- = AA^-$, co dowodzi, że AA^- jest endomorfizmem samosprzężonym i idempotentnym, więc jest ortogonalną projekcją na $\mathcal{R}(A)$. Ponadto $A^-AA^- = A^-$ i $AA^-A = A$, więc zdefiniowana powyżej uogólniona odwrotność jest tak zwaną uogólnioną wzajemną odwrotnością Moore'a Penrose'a oznaczaną zwykle przez A^+. Różne warianty uogólnionych odwrotności dla $A \in \mathcal{L}(V, W)$ można znaleźć w [6] (str. 42).

W ogólniejszej sytuacji, kiedy A jest określone między dwiema przestrzeniami wektorowymi, pojęcia wartości i wektorów własnych przestają mieć sens. Reprezentację przekształcenia A uzyskujemy, korzystając z twierdzenia 1.7 dla nieujemnie określonego odwzorowania $A'A$.
Pokażemy teraz, jak można wykorzystać twierdzenie spektralne do reprezentacji dowolnego przekształcenia liniowego. Niech $(V, (\cdot,\cdot))$ i $(W, [\cdot,\cdot])$ będą skończenie wymiarowymi rzeczywistymi przestrzeniami unitarnymi i $A \in \mathcal{L}(V, W)$. Z definicji przekształcenia sprzężonego wynika, że

$$A^{'}A \in \mathcal{L}(V, V)$$

jest endomorfizmem

- samosprzężonym, gdyż $\forall x, y \in V \ (y, A^{'}Ax) = [Ay, Ax] = (A^{'}Ay, x)$,
- nieujemnie określonym, gdyż $\forall x \in V \ (x, A^{'}Ax) = [Ax, Ax] \geqslant 0$.

Z twierdzenia 1.3 mamy $\mathcal{N}(A) = \mathcal{N}(A^{'}A)$. Ale

$$dim\mathcal{N}(A) + dim\mathcal{R}(A) = dimV = dim\mathcal{N}(A^{'}A) + \mathcal{R}(A^{'}A),$$

więc

$$dim\mathcal{R}(A) = dim\mathcal{R}(A^{'}A) = k.$$

Niech $\lambda_1 \geqslant \lambda_2, \ldots, \lambda_k$ będą niezerowymi wartościami własnymi $A^{'}A$. Jest ich dokładnie k licząc z krotnościami, gdyż wymiar przestrzeni własnej odpowiadającej wartości własnej $\lambda = 0$ jest równy wymiarowi przestrzeni $\mathcal{N}(A^{'}A)$. Z twierdzenia spektralnego otrzymujemy

$$A^{'}A = \sum_{1}^{k} \lambda_i v_i \Box v_i,$$

gdzie $\{v_1, \ldots, v_n\}$ jest ortonormalną bazą wektorów własnych endomorfizmu $A^{'}A$ przy czym $\{v_1, \ldots, v_k\}$ jest bazą $\mathcal{R}(A^{'}A)$ a $\{v_{k+1}, \ldots, v_n\}$ jest bazą w $\mathcal{N}(A^{'}A) = \mathcal{N}(A)$. Przy powyższych oznaczeniach niech $w_i = \frac{1}{\sqrt{\lambda_i}} Av_i$ dla $i = 1, \ldots, k$. Wówczas $\{w_1, \ldots, w_k\}$ jest bazą ortonormalną w $\mathcal{R}(A) \subset W$ i

$$A = \sum_{1}^{k} \sqrt{\lambda_i} w_i \Box v_i. \tag{1.5}$$

Ponieważ $dim\mathcal{R}(A) = k$, to $\{w_1, \ldots, w_k\}$ będzie bazą w $\mathcal{R}(A)$, gdy $\{w_1, \ldots, w_k\}$ będzie układem ortogonalnym.

$$[w_i, w_j] = \frac{1}{\sqrt{\lambda_i \lambda_j}}[Av_i, Av_j] = \frac{1}{\sqrt{\lambda_i \lambda_j}}[A^{'}Av_i, v_j] = \frac{1}{\sqrt{\lambda_i \lambda_j}}[\lambda_i v_i, v_j] = \frac{\sqrt{\lambda_i}}{\sqrt{\lambda_j}}(v_i, v_j) = \delta_{ij}.$$

Aby pokazać równość (1.5), wystarczy pokazać zgodność transformacji A i $\sum_1^k \sqrt{\lambda_i} w_i \Box v_i$ na wektorach bazowych $v_1, \ldots, v_n$. Oczywiście z definicji w_j mamy $Av_j = \sqrt{\lambda_j} w_j$, $j = 1, \ldots, k$ i $Av_j = 0$ dla $j = k+1, \ldots, n$. Z drugiej strony

$$\left(\sum_{i=1}^{k} \sqrt{\lambda_i} w_i \Box v_i\right) x_j = \sum_{i=1}^{k} \sqrt{\lambda_i} (w_i \Box v_i) x_j =$$
$$\sum_{i=1}^{k} \sqrt{\lambda_i} (v_i, v_j) w_i = \begin{cases} \sqrt{\lambda_j} w_j, & j = 1, \ldots, k \\ 0 & j = k+1, \ldots, n. \end{cases}$$

Konsekwencjami reprezentacji (1.7) przekształcenia A i własności iloczynu zewnętrznego są następujące równości:

(i) $A^{'} = \sum_1^k \sqrt{\lambda_i} v_i \Box w_i$,

(ii) $A^{'} w_i = \sqrt{\lambda_i} v_i$,

(iii) $AA^{'} = \sum_1^k \lambda_i w_i \Box w_i$.

W konsekwencji prawdziwe jest następujące twierdzenie o rozkładzie według wartości osobliwych (w skrócie SVD od ang. *Singular Value Decomposition*).

Twierdzenie 1.10. (SVD [2], str. 58) *Niech $A \in \mathcal{L}(V, W)$ będzie odwzorowaniem rzędu k. Wówczas istnieją zbiory ortonormalne $\{v_1, v_2, \ldots, v_k\} \subset V$ i $\{w_1, w_2, \ldots, w_k\} \subset W$ oraz dodatnie stałe $\lambda_1, \ldots, \lambda_k$ takie, że*

$$A = \sum_{i=1}^{k} \sqrt{\lambda_i} w_i \Box v_i.$$

Ponadto $\mathcal{N}(A) = (\mathrm{span}\{v_1, \ldots, v_k\})^{\perp}$, $\mathcal{R}(A) = \mathrm{span}\{w_1, \ldots, w_k\}$, $Av_i = \sqrt{\lambda_i} w_i$, $A'A = \sum_{i=1}^{k} \lambda_i v_i \Box v_i$, $AA' = \sum_{i=1}^{k} \lambda_i w_i \Box w_i$ oraz stałe $\lambda_1, \ldots, \lambda_k$ są dodatnimi wartościami własnymi każdego z odwzorowań $A'A$ i AA'.

Zdefiniujemy teraz odwzorowanie, które pełni istotną rolę przy opisie struktury kowariancyjnej wektorów losowych.

Definicja 1.4. *Niech dane będą przestrzenie* $(V_1, (\cdot,\cdot)_1)$, $(V_2, (\cdot,\cdot)_2)$, $(W_1, [\cdot,\cdot]_1)$ *oraz* $(W_2, [\cdot,\cdot]_2)$*. Iloczynem Kroneckera odwzorowań* $A \in \mathcal{L}(V_1, V_2)$ *i* $B \in \mathcal{L}(W_1, W_2)$ *nazywamy odwzorowanie* $B \otimes A\colon \mathcal{L}(V_1, W_1) \to \mathcal{L}(V_2, W_2)$ *zdefiniowane następująco:*

$$(B \otimes A)C = BCA', \qquad C \in \mathcal{L}(V_1, W_1). \tag{1.6}$$

Iloczyn Kroneckera ma tę szczególną własność, że przekształca iloczyn zewnętrzny wektorów z przestrzeni V_1 oraz W_1 w iloczyn zewnętrzny odpowiednich wektorów z przestrzeni V_2 i W_2. Tę i inne użyteczne własności iloczynu Kroneckera podaje poniższe twierdzenie.

Twierdzenie 1.11. *Przyjmijmy oznaczenia zgodne z definicją* 1.4*. Wówczas:*

(i) $(B \otimes A)(w_1 \Box v_1) = (Bw_1)\Box(Av_1)$ *dla dowolnych wektorów* $v_1 \in V_1$ *oraz* $w_1 \in W_1$,

(ii) $(B \otimes A)' = B' \otimes A'$.
Ponadto, jeżeli $V_1 = V_2 = V$ *oraz* $W_1 = W_2 = W$*, to:*

(iii) $(B_1 \otimes A_1)(B_2 \otimes A_2) = (B_1B_2) \otimes (A_1A_2)$ *dla dowolnych* $A_1, A_2 \in \mathcal{L}(V,V)$ *oraz* $B_1, B_2 \in \mathcal{L}(W,W)$,

(iv) *jeżeli* A_1^{-1} *oraz* B_1^{-1} *istnieją, to* $(B_1 \otimes A_1)^{-1} = B_1^{-1} \otimes A_1^{-1}$,

(v) *jeżeli* A_1 *oraz* B_1 *są projekcjami ortogonalnymi, to* $B_1 \otimes A_1$ *jest projekcją ortogonalną.*

W punktach (ii) *oraz* (v) *przyjmujemy, że na* $\mathcal{L}(V_i, W_i)$, $i = 1, 2$, *określone są iloczyny skalarne* $\langle\cdot,\cdot\rangle_i$ *zadane wzorem* (1.1).

Dowód. Dla dowodu części (i) ustalmy element $v \in V_2$ i wyznaczmy wartość odwzorowania $(B \otimes A)(w_1 \Box v_1)$ w tym punkcie. Otrzymujemy

$$\begin{aligned}(B \otimes A)(w_1 \Box v_1)v &= B(w_1 \Box v_1)A'v = B(v_1, A'v)_1 w_1 \\ &= (Av_1, v)_2 Bw_1 = (Bw_1 \Box Av_1)v.\end{aligned}$$

Dzięki dowolności $v \in V_2$ uzyskujemy żądaną równość. Aby dowieść (ii), należy pokazać, że dla każdego $C_1 \in \mathcal{L}(V_1, W_1)$ oraz $C_2 \in \mathcal{L}(V_2, W_2)$ odwzorowanie $A' \otimes B'$ spełnia równanie

$$\langle C_2, (B \otimes A)C_1\rangle_2 = \langle (B' \otimes A')C_2, C_1\rangle_1. \tag{1.7}$$

Z twierdzenia 1.5 wynika, że odwzorowania postaci $w_1 \Box v_1$ tworzą bazę przestrzeni $\mathcal{L}(V_1, W_1)$. Ograniczymy się zatem do sprawdzenia czy powyższa równość zachodzi dla $C_1 = w_1 \Box v_1$, gdzie $w_1 \in W_1$ oraz $v_1 \in V_1$ to dowolne ustalone elementy. Rozpisując lewą stronę równości (1.7) i korzystając z punktu (i) oraz twierdzenia 1.6, otrzymujemy

$$\begin{aligned}\langle C_2, (B \otimes A)(w_1 \Box v_1)\rangle_2 &= \langle C_2, Bw_1 \Box Av_1\rangle_2 = [C_2 Av_1, Bw_1]_2 = [B' C_2 Av_1, w_1]_1 \\ &= \langle B' C_2 A, w_1 \Box v_1\rangle_1 = \langle (B' \otimes A')C_2, w_1 \Box v_1\rangle_1,\end{aligned}$$

co kończy dowód tej części twierdzenia. Ustalmy następnie dowolne odwzorowanie $C \in \mathcal{L}(V, W)$. Ponieważ

$$\begin{aligned}(B_1 \otimes A_1)(B_2 \otimes A_2)C &= (B_1 \otimes A_1)(B_2 C A_2') = B_1 B_2 C A_2' A_1' = B_1 B_2 C (A_1 A_2)' \\ &= (B_1 B_2 \otimes A_1 A_2)C,\end{aligned}$$

to prawdziwa jest równość $(B_1 \otimes A_1)(B_2 \otimes A_2) = (B_1 B_2 \otimes A_1 A_2)$.

Części (iv) oraz (v) wynikają w prosty sposób z tego, co zostało już pokazane. Mianowicie załóżmy, że istnieją odwzorowania A_1^{-1} oraz B_1^{-1}, wtedy

$$(B_1^{-1} \otimes A_1^{-1})(B_1 \otimes A_1) = (B_1 \otimes A_1)(B_1^{-1} \otimes A_1^{-1}) = (B_1^{-1} B_1 \otimes A_1^{-1} A_1) = I \otimes I,$$

stąd $(B_1^{-1} \otimes A_1^{-1}) = (B_1 \otimes A_1)^{-1}$. Przyjmując natomiast, że A_1 oraz B_1 są projekcjami ortogonalnymi, otrzymujemy natychmiast równości

$$(B_1 \otimes A_1)' = B_1' \otimes A_1' = B_1 \otimes A_1$$

oraz

$$(B_1 \otimes A_1)(B_1 \otimes A_1) = B_1^2 \otimes A_1^2 = B_1 \otimes A_1,$$

co oznacza, że odwzorowanie $(B_1 \otimes A_1)$ jest projekcją ortogonalną w $\mathcal{L}(V, W)$. Tym samym dowód twierdzenia został zakończony. $\Box$

1.2 Wektory losowe i ich struktura kowariancyjna

Przestrzeń unitarna V jest przestrzenią metryczną, więc mają w niej sens takie pojęcia jak kula, zbiór otwarty i w konsekwencji borelowska $\sigma-$ algebra $\mathcal{B}(V)$. Warto zauważyć, że $\mathcal{B}(V)$ nie zależy od wyboru iloczynu skalarnego w przestrzeni V, gdyż dowolne dwa iloczyny skalarne $(\cdot,\cdot)_1$, $(\cdot,\cdot)_2$ określone na V są związane ze sobą poprzez dodatnio określone przekształcenie $A \in \mathcal{S}(V, (\cdot,\cdot)_1)$ wzorem $(x, y)_2 = (x, Ay)_1$ dla dowolnych $x, y \in V$ (por. dowód punktu (ii) twierdzenia 1.12, str. 17).

Niech dana będzie przestrzeń probabilistyczna $(\Omega, \mathcal{F}, \mathbb{P}_0)$ oraz skończenie wymiarowa przestrzeń wektorowa z iloczynem skalarnym $(V, (\cdot,\cdot))$.

Definicja 1.5. *Wektorem losowym o wartościach w V nazywamy borelowsko mierzalne odwzorowanie $X\colon \Omega \to V$. Indukowaną przez X miarę probabilistyczną określoną na $\mathcal{B}(V)$ wzorem*

$$Q(B) = \mathbb{P}_0(X^{-1}(B)), \quad B \in \mathcal{B}(V)$$

nazywamy rozkładem prawdopodobieństwa wektora X i oznaczamy $\mathcal{L}(X)$.

Każde odwzorowanie liniowe określone na skończenie wymiarowej przestrzeni unormowanej jest ciągłe, a więc borelowsko mierzalne. Z tego powodu wektor losowy poddany przekształceniu liniowemu jest wektorem losowym, a w przypadku funkcji rzeczywistych — zmienną losową. Mając to na uwadze, w dowodach twierdzeń przedstawionych w tym podrozdziale, nie będziemy uzasadniać mierzalności rozpatrywanych odwzorowań.

Definicja 1.6. *Niech $X \in V$ będzie wektorem losowym o rozkładzie Q. Niech f będzie rzeczywistą funkcją borelowsko mierzalną określoną na przestrzeni V. Jeżeli*

$$\int\limits_V |f(x)|Q(dx) < \infty,$$

to mówimy, że zmienna losowa $f(X)$ ma skończoną wartość oczekiwaną i oznaczamy ją przez $\mathbb{E}f(X)$, gdzie

$$\mathbb{E}f(X) = \int\limits_V f(x)Q(dx).$$

Podobnie jak w przypadku zmiennych losowych, rozkład wektora losowego może być równoważnie opisany za pomocą funkcji charakterystycznej (patrz [2], str. 77).

Definicja 1.7. *Funkcję Φ o wartościach zespolonych określoną na przestrzeni $(V, (\cdot, \cdot))$ wzorem*

$$\Phi(v) = \int\limits_V e^{i(v,x)}Q(dx)$$

nazywamy funkcją charakterystyczną wektora $X \in (V, (\cdot, \cdot))$.

Definicja 1.8. *Niech X będzie wektorem losowym w przestrzeni $(V, (\cdot, \cdot))$ oraz niech zmienna losowa (x, X) ma skończoną wartość oczekiwaną dla każdego wektora $x \in V$. Wektor $\mu \in V$ taki, że dla każdego $x \in V$ spełniona jest równość*

$$\mathbb{E}(x, X) = (x, \mu)$$

nazywamy wartością oczekiwaną wektora X i oznaczamy przez $\mathbb{E}(X)$.

Liniowość wartości oczekiwanej zmiennej losowej oraz iloczynu skalarnego gwarantuje, że funkcja $f(x) = \mathbb{E}(x, X)$ jest liniowa. Istnieje więc wektor $\mu \in V$ (dokładnie jeden) taki, że $f(x) = (x, \mu)$ dla każdego $x \in V$ (patrz [2], str 16).

Twierdzenie 1.12. *Niech $X \in (V, (\cdot,\cdot))$ będzie wektorem losowym o wartości oczekiwanej $\mu \in V$ oraz niech $(W, [\cdot,\cdot])$ będzie przestrzenią wektorową z iloczynem skalarnym. Wówczas:*

(i) *dla odwzorowania liniowego $A \in \mathcal{L}(V, W)$ oraz dowolnego wektora $w_0 \in W$ zachodzi równość*

$$\mathbb{E}(AX + w_0) = A\mu + w_0,$$

(ii) *dla każdej dwuliniowej funkcji $f\colon V \times V \to \mathbb{R}$ spełniona jest równość*

$$\mathbb{E}f(x, X) = f(x, \mu).$$

Dowód. Dla dowodu punktu (i) sprawdzimy, że $A\mu + w_0$ spełnia równanie definiujące wartość oczekiwaną wektora $AX+w_0$. Mianowicie, dla dowolnego $w \in W$ mamy

$$\begin{aligned}\mathbb{E}[w, AX + w_0] &= \mathbb{E}[w, AX] + [w, w_0] = \mathbb{E}(A'w, X) + [w, w_0]\\ &= (A'w, \mu) + [w, w_0] = [w, A\mu] + [w, w_0]\\ &= [w, A\mu + w_0].\end{aligned}$$

Punkt (ii) wynika z faktu, że dla każdej rzeczywistej funkcji dwuliniowej f określonej na $V \times V$ istnieje odwzorowanie $A \in \mathcal{L}(V, V)$ takie, że funkcję f możemy przedstawić w postaci $f(x, y) = (x, Ay)$, $(x, y) \in V \times V$ (patrz [2] str. 34). Dokładniej

$$\mathbb{E}f(x, X) = \mathbb{E}(x, AX) = (x, A\mu) = f(x, \mu),$$

gdzie druga równość jest konsekwencją tezy zawartej w punkcie (i). □

Z punktu (ii) powyższego twierdzenia wynika następujący wniosek.

Wniosek 1.2. *Wartość oczekiwana wektora losowego $X \in V$ nie zależy od iloczynu skalarnego określonego w przestrzeni V. Istotnie, niech $(\cdot,\cdot)_1$ będzie pewnym iloczynem skalarnym na V oraz niech $\mathbb{E}(x, X)_1 = (x, \mu)_1$. Niech $(\cdot,\cdot)_2$ będzie dowolnym innym iloczynem skalarnym na V. Rozważmy funkcję dwuliniową $f\colon V \times V \to \mathbb{R}$ określoną wzorem*

$$f(x, y) = (x, y)_2, \quad x, y \in V.$$

Wówczas z twierdzenia 1.12 *wynika, że $\mathbb{E}(x, X)_2 = \mathbb{E}f(x, X) = f(x, \mu) = (x, \mu)_2$, czyli μ jest wartością oczekiwaną wektora X w $(V, (\cdot,\cdot)_2)$.*

Załóżmy następnie, że wartość oczekiwana zmiennej losowej $(x, X)^2$ jest skończona dla każdego $x \in V$ i zdefiniujmy kowariancję wektora losowego X.

Definicja 1.9. *Nieujemnie określone odwzorowanie $\Sigma \in \mathcal{S}(V)$ nazywamy kowariancją wektora losowego X w przestrzeni $(V, (\cdot,\cdot))$ i oznaczamy przez* $\mathrm{Cov}(X)$*, gdy dla każdych dwóch elementów $x, y \in V$ spełniona jest równość*

$$\mathrm{cov}\{(x, X), (y, X)\} = (x, \Sigma y). \tag{1.8}$$

Aby uzasadnić istnienie odwzorowania Σ spełniającego powyższą definicję, wystarczy, zgodnie z faktem wykorzystanym w dowodzie twierdzenia 1.12, zauważyć, że funkcja $f(x, y) = \mathrm{cov}\{(x, X), (y, X)\}$ jest funkcją dwuliniową, więc istnieje odwzorowanie $\Sigma \in \mathcal{L}(V, V)$ takie, że dla wszystkich wektorów $x, y \in V$ zachodzi równość $f(x, y) = (x, \Sigma y)$. Odwzorowanie Σ jest samosprzężone, ponieważ $f(x, y) = f(y, x)$ dla dowolnych $x, y \in V$. Ponadto, ponieważ dla każdego $x \in V$ mamy

$$(x, \Sigma x) = \mathrm{cov}\{(x, X), (x, X)\} = \mathrm{var}(x, X) \geqslant 0,$$

odwzorowanie Σ jest nieujemnie określone. Rozważmy teraz dowolne inne odwzorowanie $\Sigma_1 \in \mathcal{S}(V)$ spełniające równanie (1.8), a więc takie, że

$$(x, \Sigma_1 y) = (x, \Sigma y)$$

lub równoważnie

$$(x, (\Sigma_1 - \Sigma)y) = 0$$

dla każdego $x \in V$ oraz każdego $y \in V$. W szczególności, przyjmując $x = (\Sigma_1 - \Sigma)y$, otrzymujemy

$$((\Sigma_1 - \Sigma)y, (\Sigma_1 - \Sigma)y) = 0,$$

stąd $(\Sigma_1 - \Sigma)y = 0$ dla każdego $y \in V$, wiec $\Sigma_1 = \Sigma$, zatem odwzorowanie Σ spełniające równanie (1.8) jest wyznaczone jednoznacznie.

Wykorzystując powyższe rozumowanie, można pokazać, że dla dowolnych odwzorowań samosprzężonych $A_1, A_2 \in \mathcal{S}(V)$ warunek $(x, A_1 x) = (x, A_2 x)$ spełniony dla każdego $x \in V$ oznacza, że $A_1 = A_2$ (aby to pokazać wystarczy skorzystać z równości $(x + y, A_1(x + y)) = (x + y, A_2(x + y))$, $x, y \in V$. (por. [2], str. 21)). W związku z tym prawdziwe jest następujące twierdzenie.

Twierdzenie 1.13. *Niech $X \in (V, (\cdot,\cdot))$ będzie wektorem losowym. Załóżmy, że kowariancja wektora X istnieje i jest równa* $\mathrm{Cov}(X) = \Sigma$*. Jeżeli samosprzężone odwzorowanie $\Sigma_1 \in \mathcal{S}(V)$ spełnia równanie*

$$\mathrm{var}(x, X) = (x, \Sigma_1 x) \tag{1.9}$$

dla każdego wektora $x \in V$, to $\Sigma = \Sigma_1$.

Następne twierdzenie pokazuje w jaki sposób kowariancja wektora losowego zależy od rozważanego iloczynu skalarnego.

Twierdzenie 1.14. *Niech $X \in (V, (\cdot,\cdot)_1)$ będzie wektorem losowym. Załóżmy, że kowariancja wektora X istnieje i jest równa* $\mathrm{Cov}(X) = \Sigma$. *Określmy iloczyn skalarny $(\cdot,\cdot)_2$ na V następująco*

$$(x, y)_2 = (x, Ay)_1, \qquad x, y \in V, \tag{1.10}$$

gdzie $A \in \mathcal{S}(V, (\cdot,\cdot)_1)$ jest odwzorowaniem dodatnio określonym. Wówczas kowariancja wektora X w przestrzeni $(V, (\cdot,\cdot)_2)$ jest postaci ΣA.

Dowód. Odwzorowanie ΣA jest liniowe oraz samosprzężone (względem $(\cdot,\cdot)_2$), ponieważ dla dowolnych wektorów $x, y \in V$ mamy $(x, \Sigma Ay)_2 = (x, A\Sigma Ay)_1 = (\Sigma Ax, y)_2$. Ponadto ΣA jest nieujemnie określone, ponieważ $(x, \Sigma Ax)_2 = (\Sigma Ax, Ax)_1 \geqslant 0$ dla niezerowych wektorów $x \in V$. Dla dowolnych wektorów $x, y \in V$ mamy

$$\begin{aligned}\mathrm{cov}\{(x, X)_2, (y, X)_2\} &= \mathrm{cov}\{(x, AX)_1, (y, AX)_1\} = \mathrm{cov}\{(Ax, X)_1, (Ay, X)_1\} \\ &= (Ax, \Sigma Ay)_1 = (x, A\Sigma Ay)_1 = (x, \Sigma Ay)_2,\end{aligned}$$

czyli odwzorowanie ΣA spełnia równanie definiujące kowariancję wektora X w przestrzeni $(V, (\cdot,\cdot)_2)$, co kończy dowód. □

Z twierdzenia 1.14 wynikają dwa istotne wnioski.

Wniosek 1.3. *Kowariancja wektora losowego o wartościach w przestrzeni V zależy od iloczynu skalarnego określonego w V. Dobierając odpowiednio iloczyn skalarny, można zapewnić brak korelacji pomiędzy poszczególnymi składowymi wektora o dodatnio określonej kowariancji. Mianowicie, niech Σ będzie kowariancją wektora X w przestrzeni $(V, (\cdot,\cdot)_1)$. Dobierając iloczyn skalarny $(\cdot,\cdot)_2$ tak, aby równanie* (1.10) *było spełnione dla $A = \Sigma^{-1}$ oraz ustalając dowolną bazę ortonormalną $\{x_1, x_2, \ldots, x_n\}$ w $(V, (\cdot,\cdot)_2)$, kowariancja składowych i-tej oraz j-tej wektora X w $(V, (\cdot,\cdot)_2)$ wynosi*

$$\mathrm{cov}\{(x_i, X)_2, (x_j, X)_2\} = (x_i, \Sigma Ax_j)_2 = (x_i, x_j)_2 = \delta_{ij}.$$

Wniosek 1.4. *Jeśli kowariancja wektora losowego istnieje dla pewnego iloczynu skalarnego zadanego na przestrzeni V, to istnieje dla wszystkich możliwych iloczynów skalarnych w V.*

Kolejne twierdzenie podaje jak zmienia się postać kowariancji wektora losowego poddanego przekształceniu liniowemu.

Twierdzenie 1.15. *Niech $X \in (V, (\cdot,\cdot))$ będzie wektorem losowym. Załóżmy, że kowariancja wektora X istnieje i jest równa* $\mathrm{Cov}(X) = \Sigma$. *Niech* *dane będzie odwzorowanie $A \in \mathcal{L}(V, W)$, gdzie $(W, [\cdot,\cdot])$ to pewna przestrzeń unitarna. Wówczas dla dowolnego wektora $w_0 \in W$ zachodzi równość*

$$\mathrm{Cov}(AX + w_0) = A\Sigma A'.$$

Dowód. Odwzorowanie $A\Sigma A'$ jest samosprzężone, więc, mając na uwadze twierdzenie 1.13, wystarczy sprawdzić, że $\mathrm{var}[w, AX + w_0] = [w, A\Sigma A' w]$ dla każdego $w \in W$. Ponieważ

$$\begin{aligned}\mathrm{var}[w, AX + w_0] &= \mathrm{var}([w, AX] + [w, w_0]) = \mathrm{var}[w, AX] = \mathrm{var}(A'w, X)\\ &= (A'w, \Sigma A'w) = [w, A\Sigma A'w],\end{aligned}$$

to $\mathrm{Cov(AX + w_0)} = A\Sigma A'$. □

Gdy transformacja $\mathrm{Cov}(X) = \Sigma$ jest osobliwa, wektor losowy X przyjmuje wartości w przestrzeni $\mathcal{R}(\Sigma) + \mu$, czyli przesuniętej o wektor μ podprzestrzeni $\mathcal{R}(\Sigma) \subset V$. Prawdziwe jest bowiem następujące twierdzenie.

Twierdzenie 1.16. *Niech $X \in (V, (\cdot,\cdot))$ będzie wektorem losowym o wartości oczekiwanej $\mathbb{E}(X) = \mu$ i kowariancji $\mathrm{Cov}(X) = \Sigma$. Wówczas $\mathbb{P}(X \in \mathcal{R}(\Sigma) + \mu) = 1$.*

Dowód. Zauważmy, że równość $\mathbb{P}(X \in \mathcal{R}(\Sigma) + \mu) = 1$ zachodzi wtedy i tylko wtedy, gdy $\mathbb{P}(X - \mu \in \mathcal{R}(\Sigma)) = 1$. Niech $Y = X - \mu$, wówczas $\mathrm{Cov}(Y) = \mathrm{Cov}(X) = \Sigma$ (patrz twierdzenie 1.15), $\mathbb{E}(Y) = 0$ i wystarczy pokazać, że $\mathbb{P}(Y \in \mathcal{R}(\Sigma)) = 1$. Jeżeli transformacja Σ jest nieosobliwa, to $\mathcal{R}(\Sigma) = V$ i teza twierdzenia jest prawdziwa. Załóżmy, że Σ jest transformacją osobliwą, a więc jądro $\mathcal{N}(\Sigma)$ ma wymiar $k > 0$. Niech $\{x_1, \ldots, x_k\}$ będzie bazą ortonormalną podprzestrzeni $\mathcal{N}(\Sigma)$. Dowolny wektor $x \in V$ nie należy do podprzestrzeni $\mathcal{R}(\Sigma)$, gdy istnieje wektor bazowy x_i przestrzeni $\mathcal{N}(\Sigma)$ taki, że $(x_i, x) \neq 0$ (wynika to z faktu, że $\mathcal{R}(\Sigma)$ i $\mathcal{N}(\Sigma)$ są ortogonalne i $\mathcal{R}(\Sigma) \oplus \mathcal{N}(\Sigma) = V$). Wobec tego

$$\mathbb{P}(Y \notin \mathcal{R}(\Sigma)) = \mathbb{P}((x_i, Y) \neq 0 \text{ dla pewnego } i \in \{1, \ldots, k\}) \leqslant \sum_{i=1}^{k} \mathbb{P}((x_i, Y) \neq 0).$$

Zmienna losowa (x_i, Y), $i = 1, \ldots, k$, ma zerową wartość oczekiwaną i zerową wariancję, gdyż $\mathrm{var}(x_i, Y) = (x_i, \Sigma x_i)$ i $x_i \in \mathcal{N}(\Sigma)$, więc $\mathbb{P}((x_i, Y) \neq 0) = 0$. Wynika z tego, że $\mathbb{P}(Y \notin \mathcal{R}(\Sigma)) = 0$. □

Przejdźmy do pojęcia niezależności wektorów losowych oraz zdefiniowania kowariancji dwóch wektorów losowych. W tym celu rozważmy wektory X_1 oraz X_2 określone na tej samej przestrzeni probabilistycznej $(\Omega, \mathcal{F}, \mathbb{P}_0)$ o wartościach w $(V_i, (\cdot,\cdot)_i)$, $i = 1, 2$.

Definicja 1.10. *Wektory losowe X_1, X_2 są niezależne, gdy dla dowolnych zbiorów borelowskich $B_i \in \mathcal{B}(V_i)$, $i = 1, 2$, zachodzi równość*

$$\mathbb{P}_0(X_1 \in B_1,\ X_2 \in B_2) = \mathbb{P}_0(X_1 \in B_1)\mathbb{P}_0(X_2 \in B_2).$$

Zdefiniujmy na przestrzeni $V_1 \oplus V_2 = \{\{v_1, v_2\} : \ v_1 \in V_1, v_2 \in V_2\}$ iloczyn skalarny $[\cdot, \cdot]$ w następujący sposób:

$$[\{v_1, v_2\}, \{v_1', v_2'\}] = (v_1, v_1')_1 + (v_2, v_2')_2, \quad \{v_1, v_2\}, \{v_1', v_2'\} \in V_1 \oplus V_2. \tag{1.11}$$

Wówczas $\{X_1, X_2\} \in (V_1 \oplus V_2, [\cdot, \cdot])$ jest wektorem losowym (patrz [2], str. 77) oraz prawdziwe jest następujące twierdzenie.

Twierdzenie 1.17. *Niech $X_1 \in (V_1, (\cdot, \cdot)_1)$ oraz $X_2 \in (V_2, (\cdot, \cdot)_2)$ będą wektorami losowymi określonymi na tej samej przestrzeni probabilistycznej. Wówczas wektory X_1 i X_2 są niezależne wtedy i tylko wtedy, gdy*

$$\Phi_{\{X_1,\ X_2\}}(\{v_1, v_2\}) = \Phi_{X_1}(v_1)\Phi_{X_2}(v_2)$$

dla wszystkich wektorów $v_1 \in V_1$, $v_2 \in V_2$, gdzie Φ_{X_i} i $\Phi_{\{X_1, X_2\}}$ oznaczają funkcje charakterystyczne wektorów X_i oraz $\{X_1, X_2\} \in V_1 \oplus V_2$.

Dowód twierdzenia można znaleźć w [2] (rozdział 2, str. 78).

Załóżmy, że wektory X_1 i X_2 są niezależne. Dla dowolnych $v_1 \in V_1$, $v_2 \in V_2$ funkcje $f_{v_1} \colon V_1 \ni x \to (v_1, x)_1 \in \mathbb{R}$ oraz $g_{v_2} \colon V_2 \ni y \to (v_2, y)_2 \in \mathbb{R}$ są ciągłe, a więc borelowsko mierzalne. Wobec tego dla dowolnych zbiorów borelowskich B_1, B_2 w przestrzeni $\mathbb{R}$ zachodzą równości

$$\begin{aligned}
\mathbb{P}_0((v_1, X_1)_1 \in B_1,\ (v_2, X_2)_2 \in B_2) &= \mathbb{P}_0(f_{v_1}(X_1) \in B_1,\ g_{v_2}(X_2) \in B_2) \\
&= \mathbb{P}_0(X_1 \in f_{v_1}^{-1}(B_1),\ X_2 \in g_{v_2}^{-1}(B_2)) \\
&= \mathbb{P}_0(X_1 \in f_{v_1}^{-1}(B_1))\mathbb{P}_0(X_2 \in g_{v_2}^{-1}(B_2)) \\
&= \mathbb{P}_0((v_1, X_1)_1 \in B_1)\mathbb{P}_0((v_2, X_2)_2 \in B_2),
\end{aligned}$$

zatem niezależność wektorów X_1 i X_2 implikuje niezależność zmiennych losowych $(v_1, X_1)_1$, $(v_2, X_2)_2$ dla każdych $v_1 \in V_1$, $v_2 \in V_2$.

Załóżmy teraz, że dla wszystkich $v_1 \in V_1$, $v_2 \in V_2$ zmienne losowe $(v_1, X_1)_1$ oraz $(v_2, X_2)_2$ są niezależne. Wtedy dla dowolnego $t \in \mathbb{R}$ prawdziwa jest równość

$$\Phi_{(v_1, X_1)_1 + (v_2, X_2)_2}(t) = \Phi_{(v_1, X_1)_1}(t)\Phi_{(v_2, X_2)_2}(t),$$

stąd

$$\begin{aligned}
\Phi_{\{X_1, X_2\}}(\{tv_1, tv_2\}) &= \mathbb{E}\exp[i(tv_1, X_1)_1 + i(tv_2, X_2)_2] \\
&= \mathbb{E}\exp[it((v_1, X_1)_1 + (v_2, X_2)_2)] \\
&= \Phi_{(v_1, X_1)_1 + (v_2, X_2)_2}(t) \\
&= \Phi_{(v_1, X_1)_1}(t)\Phi_{(v_2, X_2)_2}(t) = \Phi_{X_1}(tv_1)\Phi_{X_2}(tv_2),
\end{aligned}$$

więc z twierdzenia 1.17 wynika, że wektory losowe X_1 i X_2 są niezależne.

Wniosek 1.5. *Wektory losowe* $X_1 \in (V_1, (\cdot,\cdot)_1)$ *oraz* $X_2 \in (V_2, (\cdot,\cdot))$ *określone na tej samej przestrzeni probabilistycznej są niezależne wtedy i tylko wtedy, gdy zmienne* $(v_1, X_1)_1$ *i* $(v_2, X_2)_2$ *są niezależne dla wszystkich* $v_1 \in V_1$, $v_2 \in V_2$.

Definicja 1.11. *Niech* $X_1 \in (V_1, (\cdot,\cdot)_1)$ *oraz* $X_2 \in (V_2, (\cdot,\cdot)_2)$ *będą wektorami losowymi określonymi na tej samej przestrzeni probabilistycznej. Niech kowariancje obu wektorów istnieją i* $\mathrm{Cov}(X_1) = \Sigma_{11}$ *oraz* $\mathrm{Cov}(X_2) = \Sigma_{22}$. *Kowariancją wektorów* X_1 *i* X_2 *nazywamy odwzorowanie liniowe* $\Sigma_{12} \in \mathcal{L}(V_2, V_1)$ *takie, że dla każdej pary elementów* $v_1 \in V_1, v_2 \in V_2$ *spełniona jest równość*

$$\mathrm{cov}\{(v_1, X_1)_1, (v_2, X_2)_2\} = (v_1, \Sigma_{12} v_2)_1. \tag{1.12}$$

Rozumując podobnie jak w przypadku kowariancji wektora losowego, można pokazać, że tak zdefiniowane odwzorowanie Σ_{12} istnieje i jest wyznaczone jednoznacznie. Naturalnie, wektory X_1 oraz X_2 nazywamy nieskorelowanymi, gdy $\Sigma_{12} = 0$. Jak pokazano wcześniej, gdy wektory X_1, X_2 są niezależne, to zmienne $(v_1, X_1)_1$ oraz $(v_2, X_2)_2$ są niezależne i co za tym idzie nieskorelowane dla dowolnych $v_1 \in V_1$, $v_2 \in V_2$, czyli

$$\mathrm{cov}\{(v_1, X_1)_1, (v_2, X_2)_2\} = 0$$

dla wszystkich $v_1 \in V_1$, $v_2 \in V_2$, a więc niezależne wektory losowe są nieskorelowane. Ponadto, dla dowolnych innych iloczynów skalarnych $(\cdot,\cdot)_3$, $(\cdot,\cdot)_4$ określonych odpowiednio na V_1 oraz V_2 istnieją dodatnio określone odwzorowania $A_i \in \mathcal{S}(V_i, (\cdot,\cdot)_i)$, $i = 1, 2$ takie, że

$$\begin{aligned}\mathrm{cov}\{(v_1, X_1)_3, (v_2, X_2)_4\} &= \mathrm{cov}\{(v_1, A_1 X_1)_1, (v_2, A_2 X_2)_2\}\\ &= (A_1 v_1, \Sigma_{12} A_2 v_2)_1 = 0,\end{aligned}$$

gdy $\Sigma_{12} = 0$, stąd brak korelacji wektorów X_1 oraz X_2 zachowuje się bez względu na rozważane iloczyny skalarne.

Poniższe twierdzenie określa związek między odwzorowaniami Σ_{11}, Σ_{22}, Σ_{12} i kowariancją wektora losowego $\{X_1, X_2\}$ w przestrzeni $(V_1 \oplus V_2, [\cdot,\cdot])$.

Twierdzenie 1.18. *Niech dane będą wektory losowe* $X_1 \in (V_1, (\cdot,\cdot)_1)$, $X_2 \in (V_2, (\cdot,\cdot)_2)$. *Załóżmy, że kowariancje obu wektorów istnieją i* $\mathrm{Cov}(X_1) = \Sigma_{11}$, $\mathrm{Cov}(X_2) = \Sigma_{22}$ *oraz* $\mathrm{Cov}\{X_1, X_2\} = \Sigma_{12}$. *Niech* Σ *oznacza kowariancję wektora* $\{X_1, X_2\}$ *w przestrzeni* $(V_1 \oplus V_2, [\cdot,\cdot])$. *Zdefiniujmy odwzorowanie* $A \in \mathcal{L}(V_1 \oplus V_2, V_1 \oplus V_2)$ *następująco:*

$$A\{v_1, v_2\} = \{\Sigma_{11} v_1 + \Sigma_{12} v_2, \Sigma'_{12} v_1 + \Sigma_{22} v_2\}.$$

Wtedy odwzorowania A *i* Σ *są równe.*

Dowód. Łatwo pokazać, że odwzorowanie A jest samosprzężone, dlatego aby wykazać, że $A = \Sigma$, wystarczy sprawdzić, że dla każdego elementu $\{v_1, v_2\} \in V_1 \oplus V_2$ zachodzi równość

$$[\{v_1, v_2\}, \Sigma\{v_1, v_2\}] = [\{v_1, v_2\}, A\{v_1, v_2\}]$$

(por. komentarz poprzedzający twierdzenie 1.13, str. 18). Równość tę otrzymujemy, wykonując następujące elementarne przekształcenia:

$$\begin{aligned}
[\{v_1, v_2\}, \Sigma\{v_1, v_2\}] &= \mathrm{var}[\{v_1, v_2\}, \{X_1, X_2\}] = \mathrm{var}\{(v_1, X_1)_1 + (v_2, X_2)_2\} \\
&= \mathrm{var}(v_1, X_1)_1 + \mathrm{var}(v_2, X_2)_2 + 2\mathrm{cov}\{(v_1, X_1)_1, (v_2, X_2)_2\} \\
&= (v_1, \Sigma_{11} v_1)_1 + (v_2, \Sigma_{22} v_2)_2 + 2(v_1, \Sigma_{12} v_2)_1 \\
&= [\{v_1, v_2\}, \{\Sigma_{11} v_1 + \Sigma_{12} v_2, \Sigma'_{12} v_1 + \Sigma_{22} v_2\}] \\
&= [\{v_1, v_2\}, A\{v_1, v_2\}],
\end{aligned}$$

gdzie $\{v_1, v_2\}$ to dowolny ustalony element przestrzeni $V_1 \oplus V_2$. $\square$

Zapisując odwzorowanie A zdefiniowane w tezie powyższego twierdzenia w postaci

$$A\{v_1, v_2\} = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix} \{v_1, v_2\} = \{\Sigma_{11} v_1 + \Sigma_{12} v_2, \Sigma'_{12} v_1 + \Sigma_{22} v_2\},$$

możemy w wygodny sposób przedstawić kowariancję wektora $\{X_1, X_2\}$ jako

$$\mathrm{Cov}(\{X_1, X_2\}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix}.$$

Użyteczne w analizie modeli liniowych są fakty zawarte w poniższym twierdzeniu.

Twierdzenie 1.19. *Niech dane będą wektory losowe $X_1 \in (V, (\cdot, \cdot)_1)$, $X_2 \in (V, (\cdot, \cdot)_2)$. Niech wektor $\{X_1, X_2\} \in (V_1 \oplus V_2, [\cdot, \cdot])$ ma kowariancję postaci*

$$\mathrm{Cov}(\{X_1, X_2\}) = \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma'_{12} & \Sigma_{22} \end{pmatrix},$$

gdzie $[\cdot, \cdot]$ jest zadany wzorem (1.11). Wówczas $\mathcal{N}(\Sigma_{22}) \subseteq \mathcal{N}(\Sigma_{12})$ i $\Sigma_{12} = \Sigma_{12}\Sigma_{22}^{-}\Sigma_{22}$. Ponadto wektory $X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2$ oraz X_2 są nieskorelowane oraz kowariancja wektora $X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2$ w przestrzeni $(V_1, (\cdot, \cdot)_1)$ jest postaci

$$\mathrm{Cov}(X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma'_{12}.$$

Dowód. Najpierw pokażemy, że dla każdego $v_2 \in \mathcal{N}(\Sigma_{22})$ zachodzi równość $\Sigma_{12}v_2 = 0$. Ustalmy $v_1 \in V_1$ i stałą $\alpha \in \mathbb{R}$. Odwzorowanie Σ jest nieujemnie określone, dlatego mamy

$$\begin{aligned} 0 &\leqslant [\{v_1, \alpha v_2\}, \Sigma\{v_1, \alpha v_2\}] = [\{v_1, \alpha v_2\}, \{\Sigma_{11}v_1 + \alpha\Sigma_{12}v_2, \Sigma'_{12}v_1\}] \\ &= (v_1, \Sigma_{11}v_1)_1 + \alpha(v_1, \Sigma_{12}v_2)_1 + \alpha(v_2, \Sigma'_{12}v_1)_2 \\ &= (v_1, \Sigma_{11}v_1)_1 + 2\alpha(v_1, \Sigma_{12}v_2)_1. \end{aligned}$$

Powyższa nierówność jest prawdziwa dla każdego $v_1 \in V_1$ i dowolnej stałej α, więc $(v_1, \Sigma_{12}v_2)_1 = 0$, a stąd $\Sigma_{12}v_2 = 0$. Dla dowodu dalszej części tezy twierdzenia zauważmy, że równość $\Sigma_{12} = \Sigma_{12}\Sigma_{22}^-\Sigma_{22}$ jest równoważna równości $\Sigma_{12}(I - \Sigma_{22}^-\Sigma_{22}) = 0$, która jest prawdziwa, ponieważ $I - \Sigma_{22}^-\Sigma_{22}$ jest projekcją ortogonalną na $\mathcal{N}(\Sigma_{22}) \subseteq \mathcal{N}(\Sigma_{12})$ (por. str. 13). Wykonując standardowe przekształcenia, kowariancję

$$\mathrm{cov}\{(v_1, X_1 - \Sigma_{12}\Sigma_{22}^-X_2)_1, (v_2, X_2)_2\} = \mathrm{cov}\{(v_1, X_1)_1 - (v_1, \Sigma_{12}\Sigma_{22}^-X_2)_1, (v_2, X_2)_2\}$$

można zapisać w postaci $(v_1, (\Sigma_{12} - \Sigma_{12}\Sigma_{22}^-\Sigma_{22})v_2) = 0$, bo $\Sigma_{12} = \Sigma_{12}\Sigma_{22}^-\Sigma_{22}$, co dowodzi, że wektory $X_1 - \Sigma_{12}\Sigma_{22}^-X_2$ oraz X_2 są nieskorelowane. Aby wykazać ostatnią część tezy twierdzenia, wystarczy sprawdzić czy dla dowolnego $v_1 \in V_1$ zachodzi równość

$$\mathrm{var}(v_1, X_1 - \Sigma_{12}\Sigma_{22}^-X_2)_1 = (v_1, (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^-\Sigma'_{12})v_1)_1,$$

ponieważ odwzorowanie $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^-\Sigma'_{12}$ jest samosprzężone i prawdziwe jest twierdzenie 1.13. Przekształcając lewą stronę powyższej równości, otrzymujemy

$$\begin{aligned} \mathrm{var}(v_1, X_1 - \Sigma_{12}\Sigma_{22}^-X_2)_1 &= (v_1, \Sigma_{11}v_1)_1 + \mathrm{var}(v_1, \Sigma_{12}\Sigma_{22}^-X_2)_1 \\ &\quad - 2\mathrm{cov}\Big\{(v_1, X_1)_1, (v_1, \Sigma_{12}\Sigma_{22}^-X_2)_1\Big\} \\ &= (v_1, \Sigma_{11}v_1)_1 + (\Sigma_{22}^-\Sigma'_{12}v_1, \Sigma_{22}\Sigma_{22}^-\Sigma'_{12}v_1)_2 \\ &\quad - 2(v_1, \Sigma_{12}\Sigma_{22}^-\Sigma'_{12}v_1)_1 \\ &= (v_1, (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^-\Sigma'_{12})v_1)_1, \end{aligned}$$

gdzie ostatnia równość zachodzi, ponieważ $\Sigma_{22}^-\Sigma_{22}\Sigma_{22}^- = \Sigma_{22}^-$ (por. str. 13) i

$$(\Sigma_{22}^-\Sigma'_{12}v_1, \Sigma_{22}\Sigma_{22}^-\Sigma'_{12}v_1)_2 = (v_1, \Sigma_{12}\Sigma_{22}^-\Sigma_{22}\Sigma_{22}^-\Sigma'_{12}v_1)_1 = (v_1, \Sigma_{12}\Sigma_{22}^-\Sigma'_{12}v_1)_1.$$

□

Omówimy teraz pewne własności wektorów losowych o rozkładach sferycznych i słabo sferycznych. Niech V będzie przestrzenią unitarną, przez $\mathcal{O}(V)$ będziemy oznaczać grupę transformacji ortogonalnych (względem iloczynu skalarnego zadanego na V) określonych na przestrzeni V.

Definicja 1.12. *Mówimy, że wektor losowy $X \in (V, (\cdot,\cdot))$ o rozkładzie Q ma rozkład sferyczny lub inaczej ortogonalnie niezmienniczy, gdy dla dowolnego zbioru $B \in \mathcal{B}(V)$ oraz dla każdego $\Gamma \in \mathcal{O}(V)$ zachodzi równość $Q(B) = Q(\Gamma B)$, czyli $\mathcal{L}(X) = \mathcal{L}(\Gamma X)$ dla dowolnego $\Gamma \in \mathcal{O}(V)$.*

Twierdzenie 1.20. *Niech $X \in (V, (\cdot,\cdot))$ będzie wektorem losowym o rozkładzie sferycznym. Niech $\|\cdot\|$ oznacza normę pochodzącą od iloczynu skalarnego określonego na V oraz niech $x_0 \in V$ będzie wektorem takim, że $\|x_0\| = 1$. Wówczas $\mathcal{L}((x,X)) = \mathcal{L}(\|x\|(x_0,X))$ dla dowolnego $x \in V$.*

Dowód. Dla $x = 0$ twierdzenie jest oczywiście prawdziwe. Dla $x \neq 0$ wybierzmy takie odwzorowanie $\Gamma \in \mathcal{O}(V)$, aby zachodziła zależność $\Gamma x_0 = \frac{x}{\|x\|}$. Taki dobór odwzorowania Γ jest możliwy, gdyż wektory x_0 i $\frac{x}{\|x\|}$ mają taką samą normę, równą jeden. Wtedy

$$\begin{aligned}\mathcal{L}((x,X)) &= \mathcal{L}\left(\|x\|\left(\frac{x}{\|x\|}, X\right)\right) = \mathcal{L}(\|x\|(\Gamma x_0, X)) \\ &= \mathcal{L}(\|x\|(x_0, \Gamma' X)) = \mathcal{L}(\|x\|(x_0, X)),\end{aligned}$$

gdzie ostatnia równość wynika z założenia o niezmienniczości rozkładu wektora X względem transformacji ortogonalnych oraz faktu, że jeśli $\Gamma \in \mathcal{O}(V)$ to również $\Gamma' \in \mathcal{O}(V)$. □

Twierdzenie 1.20 pokazuje, że dla sferycznego rozkładu $\mathcal{L}(X)$ rozkład zmiennej losowej (x,X) jest taki sam jak rozkład zmiennej $\| x \| (x_0, X)$, stąd, znając rozkład zmiennej (x_0, X) dla pewnego niezerowego wektora $x_0 \in V$, możemy wyznaczyć rozkład (x,X) dla dowolnego $x \in V$.

Twierdzenie 1.21. *Niech $X \in (V, (\cdot,\cdot))$ będzie wektorem losowym o rozkładzie sferycznym. Niech $\|\cdot\|$ oznacza normę pochodzącą od iloczynu skalarnego określonego na V oraz niech $x_0 \in V$, $\|x_0\| = 1$. Wówczas*

(i) $\Phi_X(x) = \mathbb{E}e^{i(x,X)} = \Phi(\|x\|x_0)$,

(ii) *jeśli* $\mathbb{E}(X)$ *istnieje, to* $\mathbb{E}(X) = 0$,

(iii) *jeśli* $\mathrm{Cov}(X)$ *istnieje, to wówczas* $\mathrm{Cov}(X) = \sigma^2 I$, *gdzie* $\sigma^2 = \mathrm{var}(x_0, X)$.

Dowód. Własność (i) wynika wprost z twierdzenia 1.20 oraz zależności

$$\mathbb{E}e^{i(x,X)} = \mathbb{E}e^{i\|x\|(x_0,X)} = \mathbb{E}e^{i(\|x\|x_0,X)} = \Phi(\|x\|x_0).$$

Dla dowodu (ii) przyjmijmy $\mathbb{E}(X) = \mu$. Ponieważ $\mathcal{L}(X) = \mathcal{L}(\Gamma X)$, to dla dla dowolnego $\Gamma \in \mathcal{O}(V)$ mamy

$$\mu = \mathbb{E}(X) = \mathbb{E}(\Gamma X) = \Gamma\mathbb{E}(X) = \Gamma\mu.$$

Jedynym wektorem μ spełniającym równość $\mu = \Gamma\mu$ dla każdego $\Gamma \in \mathcal{O}(V)$ jest $\mu = 0$, co dowodzi drugiej części twierdzenia. Na mocy twierdzenia 1.20 otrzymujemy ciąg równości

$$\mathrm{var}(x, X) = \mathrm{var}\,\{\|x\|(x_0, X)\} = \|x\|^2 \mathrm{var}(x_0, X) = \sigma^2(x, x) = (x, \sigma^2 I x),$$

stąd oraz z twierdzenia 1.13 wynika, że $\mathrm{Cov}(X) = \sigma^2 I$, co kończy dowód twierdzenia. $\square$

Wykazaliśmy, że jeśli kowariancja wektora losowego X o rozkładzie sferycznym istnieje, to jest postaci $\mathrm{Cov}(X) = \sigma^2 I$ dla pewnego $\sigma^2 \geqslant 0$. Oczywiście istnieją inne rozkłady o kowariancji postaci $\sigma^2 I$.

Definicja 1.13. *Wektor losowy $X \in (V, (\cdot,\cdot))$ ma rozkład słabo sferyczny, gdy*

$$\mathrm{Cov}(X) = \sigma^2 I \quad \textit{dla pewnego} \quad \sigma^2 > 0.$$

Twierdzenie 1.22. *Niech X będzie wektorem losowym o wartościach w przestrzeni $(V, (\cdot,\cdot))$. Załóżmy, że kowariancja wektora X istnieje. Następujące zależności są równoważne:*

(i) $\mathrm{Cov}(X) = \sigma^2 I$ *dla pewnego* $\sigma^2 \geqslant 0$,

(ii) $\mathrm{Cov}(X) = \mathrm{Cov}(\Gamma X)$ *dla każdego* $\Gamma \in \mathcal{O}(V)$.

Dowód. Jeśli $\mathrm{Cov}(X) = \sigma^2 I$, to z twierdzenia 1.15 wynika, że dla każdego $\Gamma \in \mathcal{O}(V)$ prawdziwa jest równość $\mathrm{Cov}(X) = \mathrm{Cov}(\Gamma X)$. Dla dowodu implikacji odwrotnej załóżmy, że $\Sigma = \mathrm{Cov}(X)$ oraz zachodzi równość (ii). Z twierdzenia 1.15 wnioskujemy, że Σ musi spełniać równość $\Sigma = \Gamma\Sigma\Gamma'$ dla dowolnego $\Gamma \in \mathcal{O}(V)$, stąd dla każdego $x \in V$ otrzymujemy

$$(x, \Sigma x) = (x, \Gamma\Sigma\Gamma' x) = (\Gamma' x, \Sigma\Gamma' x).$$

Rozważmy wektor $x \in V$ taki, że $\|x\| = 1$. Zauważmy, że wówczas $\Gamma' x$ może być dowolnym wektorem z przestrzeni V o normie równej jeden, ponieważ Γ' może być dowolnym elementem grupy przekształceń $\mathcal{O}(V)$. Wynika stąd, że dla dowolnych $x, y \in V$ takich, że $\|x\| = \|y\| = 1$ zachodzi równość

$$(x, \Sigma x) = (y, \Sigma y).$$

Korzystając z twierdzenia spektralnego, zapiszmy $\Sigma = \sum_1^n \lambda_i x_i \Box x_i$, gdzie $\{x_1, \ldots, x_n\}$ jest odpowiednią bazą ortonormalną w $(V, (\cdot,\cdot))$ oraz wybierzmy $x = x_j$ oraz $y = x_k$. Wówczas otrzymujemy

$$\lambda_j = (x_j, \Sigma x_j) = (x_k, \Sigma x_k) = \lambda_k$$

dla dowolnych (spośród rozważanych) indeksów j, k. Ustalając $\sigma^2 = \lambda_1$, mamy

$$\Sigma = \sum_{i=1}^n \sigma^2 x_i \Box x_i = \sigma^2 \sum_{i=1}^n x_i \Box x_i = \sigma^2 I.$$

Ponieważ odwzorowanie Σ jest nieujemnie określone, to $\sigma^2 \geqslant 0$, co kończy dowód. $\square$

Twierdzenie 1.22 orzeka, że rozkład słabo sferyczny jest jedynym rozkładem (poza rozkładem jednopunktowym), którego struktura kowariancyjna jest niezmiennicza względem transformacji ortogonalnych.

Wyznaczymy teraz pewne charakterystyki iloczynów zewnętrznych wektorów losowych.

Lemat 1.1. *Niech dane będą wektory losowe $X_i \in (V_i, (\cdot,\cdot)_i)$, $i = 1, 2$, określone na tej samej przestrzeni probabilistycznej o wartościach oczekiwanych μ_1 oraz μ_2. Załóżmy, że kowariancje wektorów X_1, X_2 istnieją i oznaczmy je przez Σ_{11} i Σ_{22}. Niech* $\operatorname{Cov}\{X_1, X_2\} = \Sigma_{12}$. *Wówczas zachodzi równość*

$$\mathbb{E}(X_1 \square X_2) = \Sigma_{12} + \mu_1 \square \mu_2. \tag{1.13}$$

Ponadto, jeżeli $V_1 = V_2$ oraz $(\cdot,\cdot)_1 = (\cdot,\cdot)_2$, to

$$\mathbb{E}(X_1, X_2)_1 = \langle I, \Sigma_{12}\rangle + (\mu_1, \mu_2)_1, \tag{1.14}$$

gdzie $\langle\cdot,\cdot\rangle$ jest iloczynem skalarnym określonym na $\mathcal{L}(V_2, V_1)$ zgodnie ze wzorem (1.1).

Dowód. Wektor losowy $X_1 \square X_2$ przyjmuje wartości w przestrzeni odwzorowań liniowych $\mathcal{L}(V_2, V_1)$. Zgodnie z definicją wartości oczekiwanej wektora losowego, aby wykazać prawdziwość pierwszej części tezy twierdzenia, wystarczy sprawdzić, że dla dowolnego odwzorowania $A \in (\mathcal{L}(V_2, V_1), \langle\cdot,\cdot\rangle)$ zachodzi równość

$$\mathbb{E}\langle A, X_1 \square X_2\rangle = \langle A, \Sigma_{12}\rangle + \langle A, \mu_1 \square \mu_2\rangle. \tag{1.15}$$

Na mocy twierdzenia 1.5 dowolne odwzorowanie $A \in \mathcal{L}(V_2, V_1)$ jest odpowiednią kombinacją liniową iloczynów zewnętrznych wektorów bazowych przestrzeni V_1 oraz V_2, a więc odwzorowań postaci $v_1 \square v_2$, gdzie $v_i \in V_i$. Ponadto wyrażenia znajdujące się po obu stronach równości (1.15) są liniowe ze względu na A, dlatego wystarczy sprawdzić, że równość (1.15) zachodzi dla dowolnego odwzorowania postaci $v_1 \square v_2$, $v_i \in V_i$. Z własności iloczynu skalarnego $\langle\cdot,\cdot\rangle$, przytoczonych w twierdzeniu 1.6 oraz następującym po nim wniosku 1.1, wynika, iż prawdzie są następujące równości:

$$\begin{aligned}
\mathbb{E}\langle v_1 \square v_2, X_1 \square X_2\rangle &= \mathbb{E}(v_1, X_1)_1 (v_2, X_2)_2 \\
&= \operatorname{cov}\{(v_1, X_1)_1, (v_2, X_2)_2\} + \mathbb{E}(v_1, X_1)_1 \mathbb{E}(v_2, X_2)_2 \\
&= (v_1, \Sigma_{12} v_2)_1 + (v_1, \mu_1)_1 (v_2, \mu_2)_2 \\
&= \langle v_1 \square v_2, \Sigma_{12}\rangle + \langle v_1 \square v_2, \mu_1 \square \mu_2\rangle = \langle v_1 \square v_2, \Sigma_{12} + \mu_1 \square \mu_2\rangle.
\end{aligned}$$

Dla dowodu równości (1.14) wystarczy zauważyć, że na mocy twierdzenia 1.6 mamy

$$(X_1, X_2)_1 = \langle I, X_1 \Box X_2 \rangle,$$

a stąd, korzystając ponownie z twierdzenia 1.6 i udowodnionej już równości (1.13), otrzymujemy

$$\mathbb{E}(X_1, X_2)_1 = \mathbb{E}\langle I, X_1 \Box X_2 \rangle = \langle I, \Sigma_{12} \rangle + \langle I, \mu_1 \Box \mu_2 \rangle = \langle I, \Sigma_{12} \rangle + (\mu_1, \mu_2)_1.$$

□

Twierdzenie 1.23. *Niech X będzie wektorem losowym w przestrzeni $(V, (\cdot, \cdot))$ o rozkładzie sferycznym takim, że wartość oczekiwana zmiennej losowej $\|X\|^4$ jest skończona. Niech v_1, v_2 będą dowolnymi wektorami ortonormalnymi w przestrzeni $(V, (\cdot, \cdot))$. Wówczas*

$$\mathrm{Cov}(X \Box X) = (c_1 - c_2) I \otimes I + c_2 T,$$

gdzie $c_1 = \mathrm{var}(v_1, X)^2$, $c_2 = \mathrm{cov}\{(v_1, X)^2, (v_2, X)^2\}$, natomiast odwzorowanie T jest określone na $\mathcal{S}(V)$ w następujący sposób:

$$T(A) = \langle I, A \rangle I, \qquad A \in \mathcal{S}(V),$$

gdzie iloczyn skalarny $\langle \cdot, \cdot \rangle$ jest zadany zgodnie ze wzorem (1.1).

Dowód twierdzenia można znaleźć w [2] (rozdział 2, str. 96).

Niech $X_1, X_2, \ldots, X_n$ będzie ciągiem niezależnych wektorów losowych w przestrzeni $\mathbb{R}^p$ (ze standardowym iloczynem skalarnym) o tym samym rozkładzie Q_0, czyli $\mathcal{L}(X_i) = Q_0$, $i = 1, \ldots, n$. W statystyce ciąg ten traktujemy jako n-elementową próbę prostą z populacji o rozkładzie Q_0. Ciąg $\{X_1, X_2, \ldots, X_n\}$ jest losowym wektorem w sumie prostej $\underbrace{\mathbb{R}^p \oplus \cdots \oplus \mathbb{R}^p}_{n}$. Możemy go przedstawić w postaci

$$Y = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} \in \mathbb{R}^{np}$$

i potraktować jako wektor losowy w przestrzeni $\mathbb{R}^{np}$ ze standardowym iloczynem skalarnym. W zapisie tym nie widać jednak niezależności poszczególnych wektorów X_i. Możemy także utworzyć macierz losową

$$X = \begin{bmatrix} X_1' \\ X_2' \\ \vdots \\ X_n' \end{bmatrix} \in \mathcal{M}_{n \times p}.$$

Tym sposobem macierz X jest wektorem losowym w przestrzeni macierzy $\mathcal{M}_{n\times p}$ z iloczynem skalarnym zadanym wzorem $\langle A, B\rangle = \mathrm{tr}AB' = \mathrm{tr}B'A = \mathrm{tr}A'B = \mathrm{tr}BA'$ (rozważanym w przykładzie 1.1). Wiersze macierzy X są niezależnymi wektorami losowymi i każdy z nich ma rozkład Q_0. Załóżmy, że $\mathbb{E}(X_i) = \mu$ oraz $\mathrm{Cov}(X_i) = \Sigma$ dla $i = 1, 2, \ldots, n$ (założenie o równości rozkładów wektorów X_i nie jest konieczne). Wartość oczekiwaną i kowariancję wektora $X \in \mathcal{M}_{n\times p}$ podaje następujące twierdzenie.

Twierdzenie 1.24. *Zgodnie z przyjętymi oznaczeniami zachodzą następujące równości:*

(i) $\mathbb{E}(X) = e\mu'$,

(ii) $\mathrm{Cov}(X) = I_n \otimes \Sigma$,

gdzie $e \in \mathbb{R}^n$ oznacza wektor złożony z n jedynek.

Dowód. Ustalmy dowolną macierz $A \in \mathcal{M}_{n\times p}$. Oznaczmy przez $a_1', \ldots, a_n' \in \mathbb{R}^p$ wiersze macierzy A. Wówczas

$$\mathbb{E}\langle A, X\rangle = \mathbb{E}(\mathrm{tr}(AX')) = \mathbb{E}\left(\sum_{i=1}^{n} a_i' X_i\right) = \sum_{i=1}^{n} a_i'\mu = \mathrm{tr}(A(e\mu')') = \langle A, e\mu'\rangle.$$

Zgodnie z definicją 1.8 macierz $e\mu' \in \mathcal{M}_{n\times p}$ o n identycznych wierszach postaci μ' jest wartością oczekiwaną macierzy X.

Z własności iloczynu Kroneckera wynika, że odwzorowanie $(I_n \otimes \Sigma)$ jest samosprzężone, dlatego dla dowodu punktu drugiego, na mocy twierdzenia 1.13, wystarczy pokazać, że dla każdej macierzy $A \in \mathcal{M}_{n\times p}$ zachodzi równość

$$\mathrm{var}\langle A, X\rangle = \langle A, (I_n \otimes \Sigma)A\rangle.$$

Ponieważ

$$\begin{aligned}\mathrm{var}\langle A, X\rangle &= \mathrm{var}\left(\sum_{i=1}^{n} a_i' X_i\right) = \sum_{i=1}^{n} \mathrm{var}(a_i' X_i) + \sum_{i\neq j} \mathrm{cov}\{a_i' X_i, a_j' X_j\} \\ &= \sum_{i=1}^{n} a_i'\Sigma a_i = \mathrm{tr}\left(A\Sigma A'\right) = \mathrm{tr}\left(A(A\Sigma)'\right) = \langle A, A\Sigma\rangle = \langle A, (I_n \otimes \Sigma)A\rangle,\end{aligned}$$

gdzie trzecia równość zachodzi dlatego, że wektory X_i, X_j są nieskorelowane dla $i \neq j$, dowód twierdzenia został zakończony. □

Zauważmy, że w dowodzie części (ii) powyższego twierdzenia skorzystaliśmy jedynie z faktu, że wiersze macierzy X są nieskorelowane. Okazuje się, że macierz losowa $X \in \mathcal{M}_{n\times p}$ ma kowariancję postaci $I_n \otimes \Sigma$ wtedy i tylko wtedy, gdy jej wiersze są nieskorelowane i kowariancja każdego z nich jest postaci Σ. Mianowicie, znając postać kowariancji macierzy X, można w prosty sposób wyznaczyć kowariancję między jej wierszami lub elementami. Wystarczy dobrać odpowiednie macierze i skorzystać z definicji 1.9. Załóżmy, że $\mathrm{Cov}(X) = I_n \otimes \Sigma$ i interesuje nas kowariancja wierszy X_i' oraz X_j'. Dla dowolnych wektorów $x_1, x_2 \in \mathbb{R}^p$ utwórzmy macierze A, której i-ty wiersz to x_1', a pozostałe elementy są zerami oraz macierz B, której j-ty wiersz to x_2', a pozostałe elementy są zerami. Wówczas

$$\mathrm{cov}\{x_1'X_i, x_2'X_j\} = \mathrm{cov}\{\langle A, X\rangle, \langle B, X\rangle\} = \langle A, (I_n \otimes \Sigma)B\rangle = \mathrm{tr}(A\Sigma B') = \delta_{ij} x_1' \Sigma x_2.$$

Traktując macierze jako reprezentacje odpowiednich przekształceń liniowych, wyprowadzimy z twierdzenia 1.24 regułę transformacji kowariancji macierzy X poprzez iloczyn Kroneckera. Jeśli $A \in \mathcal{M}_{n\times n}$ i $B \in \mathcal{M}_{p\times p}$, to

$$(A \otimes B)X = AXB^{'}.$$

Z twierdzenia 1.15 wynika, że

$$\mathrm{Cov}((A \otimes B)X) = (A \otimes B)\mathrm{Cov(X)}(A \otimes B)^{'}.$$

W szczególności, jeśli $\mathrm{Cov}(X) = I_n \otimes \Sigma$, to

$$\mathrm{Cov}((A \otimes B)X) = (A \otimes B)(I_n \otimes \Sigma)(A \otimes B)^{'} = (AA^{'}) \otimes (B\Sigma B^{'}).$$

Jeśli dodatkowo założymy, że macierz A jest macierzą ortogonalną, to

$$\mathrm{Cov}((A \otimes B)X) = I_n \otimes (B\Sigma B^{'})$$

oraz dla $B = I_p$ otrzymujemy

$$\mathrm{Cov}((A \otimes I_p)X) = I_n \otimes \Sigma. \tag{1.16}$$

Z równości (1.16) wynika, że brak korelacji pomiędzy wierszami macierzy losowej X zachowuje się przy przekształceniu tej macierzy przez odwzorowanie ortogonalne postaci $A \otimes I_p$.

Przykład 1.2. *Niech X_1, X_2 będą niezależnymi zmiennymi losowymi o standardowym rozkładzie normalnym. Rozważmy wektor $X = (X_1, X_2)' \in \mathbb{R}^2$ oraz symetryczną macierz*

$$S = XX' = \begin{bmatrix} X_1^2 & X_1X_2 \\ X_1X_2 & X_2^2 \end{bmatrix}.$$

Korzystając z wprowadzonych w tym paragrafie narzędzi, obliczymy najpierw kowariancję wektora $Y = (X_1^2, X_1X_2, X_2^2)'$ w przestrzeni $\mathbb{R}^3$ ze standardowym iloczynem skalarnym. Następnie wyznaczymy kowariancję macierzy S w przestrzeni $(\mathcal{S}_2, \langle\cdot,\cdot\rangle)$ macierzy symetrycznych wymiaru 2×2 ($(\mathcal{S}_2, \langle\cdot,\cdot\rangle) \subset (\mathcal{M}_{2\times 2}, \langle\cdot,\cdot\rangle)$).

Niech $x_0 = (\alpha, \beta, \gamma)' \in \mathbb{R}^3$ będzie dowolnym ustalonym wektorem. Wówczas

$$\begin{aligned}
\operatorname{var}(x_0'Y) &= \operatorname{var}(\alpha X_1^2 + \beta X_1 X_2 + \gamma X_2^2) \\
&= \alpha^2 \operatorname{var}(X_1^2) + \beta^2 \operatorname{var}(X_1 X_2) + \gamma^2 \operatorname{var}(X_2^2) \\
&+ 2\alpha\beta \operatorname{cov}(X_1^2, X_1 X_2) + 2\beta\gamma \operatorname{cov}(X_2^2, X_1 X_2) \\
&= 2\alpha^2 + \beta^2 + 2\gamma^2,
\end{aligned}$$

gdzie trzecia równość zachodzi, ponieważ zmienne X_1^2 i X_2^2 mają rozkład chi-kwadrat z jednym stopniem swobody, a stąd $\operatorname{var}(X_i^2) = 2$ oraz $\mathbb{E}(X_i^2) = 1$ dla $i = 1, 2$, więc $\operatorname{var}(X_1 X_2) = \mathbb{E}(X_1^2)\mathbb{E}(X_2^2) - (\mathbb{E}(X_1))^2(\mathbb{E}(X_2))^2 = 1$, zaś pozostałe wyrazy zerują się. Możemy zatem zapisać

$$\operatorname{var}(x_0'Y) = [\alpha, \beta, \gamma] \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = x_0' \Sigma x_0,$$

gdzie $\Sigma = diag(2, 1, 2)$. Oczywiście $\Sigma = \Sigma'$, więc możemy zastosować twierdzenie 1.13, na mocy którego $\operatorname{Cov}(Y) = \Sigma$. Przejdźmy następnie do wyznaczenia kowariancji macierzy S w przestrzeni $(\mathcal{S}_2, \langle \cdot, \cdot \rangle)$. W tym celu zauważmy najpierw, że dla dowolnego wektora $x \in \mathbb{R}^2$ postaci $x = (\alpha, \beta)'$ zachodzi równość

$$\begin{aligned}
(X \square X)x = X'xX = (\alpha X_1 + \beta X_2) \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} &= \begin{bmatrix} \alpha X_1^2 + \beta X_1 X_2 \\ \alpha X_1 X_2 + \beta X_2^2 \end{bmatrix} \\
= \begin{bmatrix} X_1^2 & X_1 X_2 \\ X_1 X_2 & X_2^2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} &= Sx,
\end{aligned}$$

stąd $S = X \square X$. Jak wiemy, przy pewnych warunkach, twierdzenie 1.23 podaje postać kowariancji odwzorowania $X \square X$. By móc zastosować je w tym przypadku musimy się upewnić, że spełnione są jego założenia. Wektor X ma rozkład normalny o wartości oczekiwanej $\mathbb{E}(X) = (0, 0)'$ oraz kowariancji $\operatorname{Cov}(X) = I_2$. Dla każdego odwzorowania ortogonalnego $\Gamma \in \mathcal{O}(\mathbb{R}^2)$ wektor ΓX ma rozkład normalny o wartości oczekiwanej $\mathbb{E}(\Gamma X) = \Gamma \mathbb{E}(X) = (0, 0)' = \mathbb{E}(X)$ oraz kowariancji $\operatorname{Cov}(\Gamma X) = \operatorname{Cov}(X)$ (por. twierdzenie 1.22). W takim razie rozkład wektora X jest sferyczny. Korzystając ponownie z wartości momentów zmiennej o rozkładzie chi-kwadrat, sprawdzamy, że wartość oczekiwana zmiennej $\|X\|^4$ jest skończona i wynosi $\mathbb{E}\|X\|^4 = \mathbb{E}(X_1^2 + X_2^2)^2 = 8$. Przyjmując $x_1 = (1, 0)'$ oraz $x_2 = (0, 1)'$, otrzymujemy

$$\begin{aligned}
c_1 &= \operatorname{var}(x_1'X)^2 = \operatorname{var}(X_1^2) = 2, \\
c_2 &= \operatorname{cov}\{(x_1'X)^2, (x_2'X)^2\} = \operatorname{cov}\{X_1^2, X_2^2\} = 0
\end{aligned}$$

oraz wobec tego

$$\mathrm{Cov}(S) = \mathrm{Cov}(X \Box X) = 2 I_2 \otimes I_2.$$

Wynik ten można uzyskać także bezpośrednio z definicji kowariancji wektora losowego. Wystarczy pokazać, że dla dowolnej macierzy $A = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}$ prawdziwa jest równość $\mathrm{var}\langle A, S\rangle = \langle A, 2(I_2 \otimes I_2)A\rangle$. Dobierając odpowiednie macierze symetryczne, można uzyskać kowariancje pomiędzy elementami macierzy S. Na przykład, wybierając macierze symetryczne $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ i $B = \frac{1}{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, otrzymujemy

$$\mathrm{cov}\{X_1^2, X_1X_2\} = \mathrm{cov}\{\langle A, S\rangle, \langle B, S\rangle\} = \langle A, 2(I_2 \otimes I_2)B\rangle = \langle A, 2B\rangle = 0.$$

Rozdział 2

Rozkład normalny na przestrzeni wektorowej

Bieżący rozdział w większości oparty jest na książce [2] (rozdział 3). Wyjątek stanowi podrozdział 2.1, w którym wykorzystywane są również fakty podane w [4] (podrozdział 1.6).

Jasnym jest, że dla niezależnych zmiennych losowych $X_1, X_2 \ldots, X_n$ o rozkładach normalnych na $\mathbb{R}$ z parametrami μ_i i σ_i^2 (co w skrócie będziemy oznaczali $\mathcal{L}(X_i) = N(\mu_i, \sigma_i^2)$) oraz dowolnych stałych $\alpha_1, \ldots, \alpha_n$ zmienna losowa $Y = \sum \alpha_i X_i$ ma rozkład $N(\sum_{i=1}^n \alpha_i \mu_i, \sum_{i=1}^n \alpha_i^2 \sigma_i^2)$. Ta własność rozkładu normalnego odgrywa kluczową rolę w procesie definiowania rozkładu normalnego na przestrzeni wektorowej. Mianowicie gwarantuje, że istnieją wektory losowe spełniające definicję tego rozkładu sformułowaną poniżej.

Niech $(V, (\cdot, \cdot))$ w dalszym ciągu oznacza skończenie wymiarową rzeczywistą przestrzeń unitarną.

Definicja 2.1. *Wektor losowy $X \in (V, (\cdot, \cdot))$ ma rozkład normalny, gdy dla każdego wektora $x \in V$ zmienna losowa $(x, X) \in \mathbb{R}$ ma rozkład normalny.*

Rozkład jednopunktowy na $\mathbb{R}$ traktujemy jak (zdegenerowany) rozkład normalny. W ten sposób definiujemy rozkład normalny wektora w przestrzeni unitarnej, nie wymagając przy tym, aby wektor ten miał nieosobliwą kowariancję. Zauważmy, że na mocy twierdzenia 1.16 nośnik rozkładu wektora o osobliwej kowariancji Σ ma wymiar mniejszy niż wymiar przestrzeni V. Istnieją zatem niezerowe wektory należące do $\mathcal{N}(\Sigma)$, a dla każdego $x \in \mathcal{N}(\Sigma)$ mamy $\mathrm{var}(x, X) = (x, \Sigma x) = 0$ i $\mathbb{P}((x, X) = (x, \mathbb{E}(X))) = 1$.

Stosowany w dalszej części pracy zapis $\mathcal{L}(X) = N(\mu, \Sigma)$ będzie oznaczał, że wektor losowy ma rozkład normalny z wartością oczekiwaną μ oraz kowariancją Σ przy pewnym iloczynie skalarnym zadanym na przestrzeni wartości wektora X. Każdorazowo jasnym będzie jaki iloczyn skalarny jest rozważany.

Aby wykazać, że istnieje wektor losowy o rozkładzie normalnym na V, oznaczmy przez $\{x_1, x_2, \ldots, \ x_n\}$ ortonormalną bazę przestrzeni $(V, (\cdot, \cdot))$. Niech $X_1, X_2, \ldots, X_n$ będą niezależnymi zmiennymi losowymi o standardowym rozkładzie normalnym na $\mathbb{R}$. Wówczas wektor postaci

$$Y = \sum_{i=1}^{n} x_i X_i \tag{2.1}$$

jest wektorem losowym w przestrzeni $(V, (\cdot, \cdot))$ oraz dla każdego $x \in V$ mamy

$$(x, Y) = \sum_{i=1}^{n} (x, x_i) X_i.$$

Liniowa kombinacja niezależnych zmiennych X_i ma rozkład normalny, dlatego dla każdego $x \in V$ zmienna (x, Y) ma rozkład normalny, co chcieliśmy pokazać. Co więcej, ponieważ dla każdego $x \in V$ zachodzą równości

$$\mathbb{E}(x, Y) = \mathbb{E}\left(\sum_{i=1}^{n} (x_i, x) X_i\right) = \sum_{i=1}^{n} (x_i, x) \mathbb{E}(X_i) = 0$$

oraz

$$\operatorname{var}(x, Y) = \operatorname{var}\left(\sum_{i=1}^{n} (x_i, x) X_i\right) = \sum_{i=1}^{n} (x_i, x)^2 \operatorname{var}(X_i) = \sum_{i=1}^{n} (x_i, x)^2 = (x, x) = (x, Ix),$$

to skonstruowany wektor Y ma rozkład normalny o wartości oczekiwanej $\mu = 0 \in V$ i kowariancji postaci $\operatorname{Cov}(Y) = I \in \mathcal{S}(V)$, co wynika z twierdzenia 1.13.

Twierdzenie 2.1. *Niech X będzie wektorem losowym o rozkładzie normalnym w przestrzeni $(V, (\cdot, \cdot))$. Niech W będzie przestrzenią wektorową z iloczynem skalarnym $[\cdot, \cdot]$. Wówczas dla dowolnego odwzorowania $A \in \mathcal{L}(V, W)$ oraz elementu $w_0 \in W$ wektor losowy $AX + w_0$ ma rozkład normalny w przestrzeni $(W, [\cdot, \cdot])$.*

Dowód. Dla dowolnego wektora $w \in W$ zachodzi równość

$$[w, AX + w_0] = [w, AX] + [w, w_0] = (A'w, X) + [w, w_0].$$

Ponieważ zmienna $(A'w, X)$ ma rozkład normalny oraz składnik $[w, w_0]$ jest stałą, to zmienna $(A'w, X) + [w, w_0]$ ma rozkład normalny na $\mathbb{R}$, co kończy dowód. $\square$

Niech wektor X ma rozkład normalny w przestrzeni $(V, (\cdot, \cdot))$. Wtedy dla dowolnych elementów $x_1, x_2 \in V$ istnieje kowariancja zmiennych (x_1, X) i (x_2, X), a więc istnieje $\operatorname{Cov}(X) = \Sigma$. Na mocy twierdzenia 1.9 odwzorowanie Σ można przedstawić w postaci $\Sigma = AA = AA'$, gdzie $A \in \mathcal{S}(V)$

jest pewnym odwzorowaniem nieujemnie określonym . Zgodnie z twierdzeniem 1.15 dla dowolnego wektora $Y \in V$ oraz $A \in \mathcal{S}(V)$ zachodzi równość $\mathrm{Cov}(AY+\mu) = A\mathrm{Cov}(Y)A'$. W szczególności, gdy $\mathrm{Cov}(Y) = I$, to $\mathrm{Cov}(AY+\mu) = AA' = \Sigma$. Z tego powodu oraz z powyższego twierdzenia wynika, że dla każdego wektora X o rozkładzie normalnym, wartości oczekiwanej równej $\mu \in V$ oraz kowariancji $\mathrm{Cov}(X) = \Sigma$ istnieje odwzorowanie $A \in \mathcal{L}(V,V)$ (na przykład $A = \Sigma^{1/2}$) takie, że $X = AY + \mu$, gdzie Y ma rozkład normalny z wartością oczekiwaną równą zeru oraz kowariancją $\mathrm{Cov}(Y) = I$. Dzięki temu opisana wcześniej konstrukcja ma istotne znaczenie w statystycznej analizie wielowymiarowej, gdyż pozwala traktować statystyki oparte na r-elementowej próbie prostej z rozkładu normalnego $N(\mu,\Sigma)$ na przestrzeni V jako funkcje wielu niezależnych zmiennych losowych o rozkładach $N(0,1)$.

Twierdzenie 2.2. *Niech $X_1 \in (V_1, (\cdot,\cdot)_1)$ oraz $X_2 \in (V_2, (\cdot,\cdot)_2)$ będą wektorami losowymi określonymi na tej samej przestrzeni probabilistycznej o łącznym rozkładzie normalnym na $(V_1 \oplus V_2, [\cdot,\cdot])$, gdzie iloczyn skalarny $[\cdot,\cdot]$ jest zadany wzorem* (1.11). *Wektory X_1 i X_2 są niezależne wtedy i tylko wtedy, gdy są nieskorelowane.*

Dowód. Jasnym jest, że wektory niezależne są nieskorelowane. Dla dowodu implikacji przeciwnej załóżmy, że wektory losowe X_1 oraz X_2 są nieskorelowane. Wtedy dla dowolnych wektorów $v_1 \in V_1$, $v_2 \in V_2$ zachodzi równość $\mathrm{cov}\{(v_1,X_1)_1, (v_2,X_2)_2\} = 0$, a więc zmienne losowe $(v_1,X_1)_1$ oraz $(v_2,X_2)_2$ są nieskorelowane. Ponieważ łączny rozkład wektorów X_1, X_2 jest rozkładem normalnym, to wektor $\{(v_1,X_1)_1, (v_2,X_2)_2\}$ ma rozkład normalny na $\mathbb{R}^2$, gdyż dla dowolnych stałych α_1, α_2 zachodzi równość

$$\alpha_1(v_1,X_1)_1 + \alpha_2(v_2,X_2)_2 = [\{\alpha_1 v_1, \alpha_2 v_2\}, \{X_1, X_2\}].$$

Brak korelacji zmiennych $(v_1,X_1)_1, (v_2,X_2)_2$ implikuje więc ich niezależność (dla dowolnych $v_1 \in V_1, v_2 \in V_2$). Wobec tego wektory X_1 i X_2 są niezależne (por. wniosek 1.5). □

Przykład 2.1. *Rozważmy przestrzeń $\mathbb{R}^p$ ze standardowym iloczynem skalarnym. Niech $X_1, X_2, \ldots, X_n$ będą niezależnymi wektorami losowymi o wartościach w $\mathbb{R}^p$. Załóżmy, że $\mathcal{L}(X_i) = N(\mu,\Sigma)$ dla $i = 1,2,\ldots,n$. Niech X będzie macierzą o wierszach $X_1', X_2', \ldots, X_n'$. Pokażemy, że macierz X ma rozkład normalny w rozważanej w przykładzie* 1.1 *przestrzeni macierzy $(\mathcal{M}_{n\times p}, \langle\cdot,\cdot\rangle)$. W tym celu weźmy dowolną macierz $A \in \mathcal{M}_{n\times p}$ i oznaczmy jej wiersze przez $a_1', a_2', \ldots, a_n'$, gdzie $a_i \in \mathbb{R}^p$. Wtedy*

$$\langle A, X\rangle = \mathrm{tr}(AX') = \sum_{i=1}^{n} a_i' X_i. \tag{2.2}$$

Ponieważ z przyjętych założeń wynika, że zmienne $a_i'X_i$ mają rozkłady normalne na $\mathbb{R}$ oraz (*zgodnie z wnioskiem* 1.5) *są niezależne, to ich kombinacja* *liniowa jest zmienną losową o rozkładzie normalnym. Zmienna $\langle A, X\rangle$ ma więc rozkład normalny dla każdej macierzy $A \in \mathcal{M}_{n\times p}$ i co za tym idzie*

$$\mathcal{L}(X) = N(e\mu', I_n \otimes \Sigma).$$

Odpowiednie momenty rozkładu zostały wyznaczone w twierdzeniu 1.24.

Wiersze macierzy $X \in \mathcal{M}_{n\times p}$ o kowariancji postaci $I_n \otimes \Sigma$ są nieskorelowane. Jeżeli dodatkowo $\mathcal{L}(X) = N(\mu, I_n \otimes \Sigma)$, to łączny rozkład każdych dwóch wierszy (lub dowolnej innej liczby) jest normalny (co wynika z równości (2.2)) i na mocy twierdzenia 2.2 wiersze macierzy X są niezależne. Biorąc pod uwagę powyższy przykład oraz komentarz znajdujący się na stronie 29, możemy sformułować następujący wniosek. Macierz losowa $X \in \mathcal{M}_{n\times p}$ ma rozkład normalny o kowariancji $I_n \otimes \Sigma$ wtedy i tylko wtedy, gdy jej wiersze są niezależnymi wektorami losowymi o rozkładach normalnych na $\mathbb{R}^p$ z kowariancją Σ.

Twierdzenie 2.3. *Niech wektor losowy $X \in (V, (\cdot,\cdot))$ ma rozkład normalny o wartości oczekiwanej $\mu \in V$ oraz kowariancji Σ. Przedstawmy odwzorowanie Σ w postaci spektralnej wynikającej z twierdzenia* 1.7

$$\Sigma = \sum_{i=1}^{n} \lambda_i x_i \Box x_i,$$

gdzie $\{x_1, x_2, \dots, x_n\}$ to odpowiednia baza ortonormalna przestrzeni $(V, (\cdot,\cdot))$. Przez X_i oznaczmy zmienne (x_i, X) dla $i = 1, 2, \dots, n$. Wówczas $X_1, X_2, \dots, X_n$ są niezależnymi zmiennymi losowymi o rozkładach normalnych. Ponadto $\mathbb{E}(X_i) = (x_i, \mu)$ oraz $\operatorname{var}(X_i) = \lambda_i$ dla $i = 1, 2, \dots, n$.

Dowód. Rozważmy przestrzeń $\mathbb{R}^n$ ze standardowym iloczynem skalarnym. Oznaczmy przez $Y \in \mathbb{R}^n$ wektor losowy postaci $Y = (X_1, X_2, \dots, X_n)'$. Pokażemy najpierw, że Y ma rozkład normalny. Dla dowolnego wektora $y = (y_1, y_2, \dots, y_n)' \in \mathbb{R}^n$ mamy

$$y'Y = \sum_{i=1}^{n} y_i X_i = \sum_{i=1}^{n} y_i (x_i, X) = (\sum_{i=1}^{n} y_i x_i, X). \tag{2.3}$$

Wektor X ma rozkład normalny, stąd zmienna losowa $(\sum_i y_i x_i, X) = y'Y$ ma rozkład normalny dla dowolnego wektora $y \in \mathbb{R}^n$, co chcieliśmy pokazać. Z równości

$$\operatorname{cov}\{X_i, X_j\} = \operatorname{cov}\{(x_i, X), (x_j, X)\} = (x_i, \Sigma x_j) = \sum_{k=1}^{n} \lambda_k (x_i, (x_k \Box x_k) x_j) = \lambda_i \delta_{ij} \tag{2.4}$$

i twierdzenia 2.2 wynika zatem, że zmienne $X_1, X_2, \dots, X_n$ są niezależne. W szczególności, oznaczając i-ty element standardowej bazy przestrzeni $\mathbb{R}^n$

przez e_i i przyjmując $y = e_i$, z równości (2.3) wynika, że zmienna $e_i'Y = X_i$ ma rozkład normalny na $\mathbb{R}$ dla $i = 1, 2, \ldots, n$. Korzystając z (2.4), otrzymujemy

$$\mathrm{var}(X_i) = \mathrm{cov}\{(x_i, X), (x_i, X)\} = (x_i, \Sigma x_i) = \lambda_i.$$

Następnie, wprost z definicji wartości oczekiwanej wektora losowego, mamy

$$\mathbb{E}(X_i) = \mathbb{E}(x_i, X) = (x_i, \mu),$$

co kończy dowód twierdzenia. □

Wniosek 2.1. *Niech wektor X spełnia założenia powyższego twierdzenia. Przyjmijmy, że $\Sigma = I$. Wówczas, dla dowolnej bazy ortonormalnej $\{x_1, x_2, \ldots, x_n\}$ przestrzeni $(V, (\cdot, \cdot))$ zachodzi równość $\Sigma = \sum\limits_{i=1}^{n} x_i \Box x_i$. Z twierdzenia 2.3 wynika zatem, że zmienne losowe (x_i, X) są niezależne dla dowolnej ortonormalnej bazy $\{x_1, \ldots, x_n\}$ w $(V, (\cdot, \cdot))$. Ponadto $\mathcal{L}(X_i) = N((x_i, \mu), 1)$ dla $i = 1, 2, \ldots, n$.*

Przedstawioną w tym paragrafie technikę definiowania rozkładu normalnego na dowolnej skończenie wymiarowej przestrzeni unitarnej można zastosować dla innych rozkładów prawdopodobieństwa, o ile tylko dla niezależnych zmiennych losowych o danym rozkładzie na $\mathbb{R}$ zachodzi twierdzenie o dodawaniu. Poniższy przykład ilustruje wykorzystanie tej procedury do zdefiniowania rozkładu Cauchy'ego na przestrzeni wektorowej.

Przykład 2.2. *Niech Z będzie zmienną losową o rozkładzie Cauchy'ego z parametrami (a, λ), gdzie $a \in \mathbb{R}$ oraz $\lambda > 0$. Funkcja charakterystyczna zmiennej Z ma wówczas postać*

$$\Phi_Z(t) = \exp\left[ait - \lambda\,|t|\right], \qquad t \in \mathbb{R}.$$

Niech $Z_1, Z_2, \ldots, Z_n$ będą niezależnymi zmiennymi losowymi o rozkładach Cauchy'ego z parametrami (a_i, λ_i) dla $i = 1, 2, \ldots, n$. Ustalmy dowolne stałe $\alpha_1, \alpha_2, \ldots, \alpha_n$. Korzystając z własności funkcji charakterystycznych, otrzymujemy funkcję charakterystyczną zmiennej $Z = \sum_i \alpha_i Z_i$ postaci

$$\Phi_Z(t) = \prod_{i=1}^{n} \Phi_{Z_i}(\alpha_i t) = \exp\left[it\sum_{i=1}^{n}\alpha_i a_i - |t|\sum_{i=1}^{n}|\alpha_i|\lambda_i\right], \qquad t \in \mathbb{R},$$

czyli zmienna losowa Z ma rozkład Cauchy'ego o parametrach $\sum_i \alpha_i a_i$ oraz $\sum_i |\alpha_i|\lambda_i$. Wykorzystując ten fakt, możemy przystąpić do zdefiniowania rozkładu Cauchy'ego na przestrzeni wektorowej tak, jak zostało to zrobione w przypadku rozkładu normalnego. Mianowicie, niech $(V, (\cdot, \cdot))$ będzie skończenie wymiarową rzeczywistą przestrzenią wektorową, wówczas wektor losowy $X \in V$ ma rozkład Cauchy'ego, gdy dla każdego $x \in V$ zmienna losowa $(x, X) \in \mathbb{R}$ ma rozkład Cauchy'ego.

Niech $\{x_1, x_2, \ldots, x_n\}$ będzie ortonormalną bazą przestrzeni $(V, (\cdot,\cdot))$ oraz niech $X_1, X_2, \ldots, X_n$, $X_i \in \mathbb{R}$, będą niezależnymi zmiennymi losowymi o rozkładzie Cauchy'ego z parametrami $(0,1)$. Wówczas wektor $X = \sum X_i x_i$ ma rozkład Cauchy'ego, ponieważ dla dowolnego wektora $x \in V$ mamy

$$(x, \sum_{i=1}^{n} X_i x_i) = \sum_{i=1}^{n} (x, x_i) X_i,$$

a zgodnie z tym, co zostało przedstawione powyżej, kombinacja liniowa niezależnych zmiennych losowych o rozkładzie Cauchy'ego ma rozkład Cauchy'ego. Ostatni etap uogólnienia można przeprowadzić na kilka sposobów. Ich opis można znaleźć w specjalistycznej literaturze, na przykład w [5].

2.1 Formy kwadratowe wektorów o rozkładzie normalnym

W bieżącym podrozdziale wyznaczymy, przy pewnych założeniach, rozkład formy kwadratowej (X, AX), gdzie $X \in (V, (\cdot,\cdot))$ jest wektorem losowym o rozkładzie normalnym $N(\mu, \Sigma)$, natomiast $A \in \mathcal{S}(V)$.

Definicja 2.2. *Niech X będzie wektorem losowym o rozkładzie normalnym $N(\mu, I)$ na przestrzeni $\mathbb{R}^n$ ze standardowym iloczynem skalarnym. Rozkład zmiennej losowej $Y = \|X\|^2$ nazywamy niecentralnym rozkładem chi-kwadrat o n stopniach swobody i parametrze niecentralności $\delta^2 = \|\mu\|^2$ i piszemy $\mathcal{L}(Y) = \chi^2_n(\delta^2)$.*

Wiadomo, że rozkład (centralny) *chi-kwadrat* o n stopniach swobody, oznaczany jako χ^2_n, to rozkład sumy kwadratów n niezależnych zmiennych losowych o rozkładzie $N(0,1)$. Z wniosku 2.1 wynika, że składowe X_i wektora X występującego w definicji 2.2 są niezależnymi zmiennymi losowymi o rozkładach $\mathcal{L}(X_i) = N(\mu_i, 1)$, gdzie $\mu = (\mu_1, \mu_2, \ldots, \mu_n)' \in \mathbb{R}^n$. Centralny rozkład chi-kwadrat jest więc szczególnym przypadkiem rozkładu niecentralnego, mianowicie

$$\chi^2_n = \chi^2_n(0).$$

Przekształcając wektor X ($\mathcal{L}(X) = N(\mu, I)$) przez transformację ortogonalną o macierzy Q, której pierwszym wierszem jest $\frac{\mu'}{\|\mu\|}$ oraz zauważając, iż

$$Y = \|X\|^2 = X'X = (QX)'QX,$$

gdzie $\mathcal{L}(QX) = N(\lambda, I)$, $\lambda = Q\mu = (\|\mu\|, 0, \ldots, 0)' \in \mathbb{R}^n$, możemy pokazać, że

$$Y = U^2 + W, \tag{2.5}$$

a zatem

$$\chi^2_n(\delta^2) = \mathcal{L}(U^2 + W), \tag{2.6}$$

gdzie U i W są niezależnymi zmiennymi losowymi o rozkładach $\mathcal{L}(U) = N(\delta, 1)$ oraz $\mathcal{L}(W) = \chi^2_{n-1}$, $\delta = \|\mu\|$. Mianowicie, oznaczając $QX = V = (V_1, V_2, \ldots, V_n)'$, mamy

$$Y = V'V = \sum_{i=1}^{n} V_i^2 = V_1^2 + \sum_{i=2}^{n} V_i^2,$$

gdzie $\mathcal{L}(V_1) = N(\delta, 1)$, $\mathcal{L}(V_i) = N(0, 1)$, $i = 2, 3, \ldots, n$ i wystarczy wybrać $U = V_1$, $W = \sum_{i=2}^{n} V_i^2$. Reprezentacja ta pozwala po pewnych przekształceniach (zob. [4], str. 35) uzyskać funkcję gęstości rozkładu $\chi^2_n(\delta^2)$ w następującej postaci:

$$f(x) = \sum_{k=0}^{\infty} p(k) g_{n+2k}(x), \quad x > 0, \tag{2.7}$$

gdzie $p(k) = \mathbb{P}(K = k)$ dla zmiennej losowej K o rozkładzie Poissona $\mathcal{P}\left(\frac{\delta^2}{2}\right)$, natomiast $g_{n+2k}(x)$ jest gęstością centralnego rozkładu χ^2_{n+2k}. Reprezentacja (2.5) pozwala także w prosty sposób wykazać twierdzenie o dodawaniu dla niecentralnego rozkładu chi-kwadrat.

Twierdzenie 2.4. *Niech X_1 oraz X_2 będą niezależnymi zmiennymi losowymi, przy czym $\mathcal{L}(X_i) = \chi^2_{n_i}(\delta_i^2)$, $i = 1, 2$. Wówczas $\mathcal{L}(X_1 + X_2) = \chi^2_{n_1+n_2}(\delta_1^2 + \delta_2^2)$.*

Dowód. Korzystając z równości (2.6), mamy $X_1 + X_2 = U_1^2 + U_2^2 + W_1 + W_2$, gdzie U_1, U_2, W_1, W_2 to niezależne zmienne losowe, $\mathcal{L}(U_i) = N(\delta_i, 1)$ oraz $\mathcal{L}(W_i) = \chi^2_{n_i-1}$, $i = 1, 2$. Korzystając ponownie z (2.6), możemy napisać $U_1^2 + U_2^2 = U_3^2 + W_3$, gdzie zmienne U_3 i W_3 są niezależne i $\mathcal{L}(U_3) = N(\sqrt{\delta_1^2 + \delta_2^2}, 1)$, $\mathcal{L}(W_3) = \chi^2_1$. Zastosowanie znanego twierdzenia o dodawaniu dla centralnego rozkładu χ^2 do wyznaczenia rozkładu zmiennej $W_1 + W_2 + W_3$ oraz wykorzystanie równości (2.6) daje tezę twierdzenia. $\square$

Przejdźmy do wyznaczenia rozkładu formy kwadratowej (X, AX), $X \in (V, (\cdot, \cdot))$, $A \in \mathcal{S}(V)$. Zaczniemy od szczególnego przypadku, gdy $\mathcal{L}(X) = N(\mu, I)$. Korzystając z twierdzenia spektralnego 1.7, zapiszmy $A = \sum_1^n \lambda_i x_i \Box x_i$, gdzie $\{x_1, x_2, \ldots, x_n\}$ jest odpowiednią bazą ortonormalną w $(V, (\cdot, \cdot))$. Wtedy

$$(X, AX) = (X, (\sum_{i=1}^{n} \lambda_i x_i \Box x_i) X) = \sum_{i=1}^{n} \lambda_i (x_i, X)^2,$$

a więc (X, AX) jest zmienną losową. Ponieważ $\mathrm{Cov}(X) = I$, to na mocy twierdzenia 2.3 i następującego po nim wniosku 2.1 zmienne $X_i = (x_i, X)$,

$i = 1, \ldots, n$, są niezależne oraz $\mathcal{L}(X_i) = N((x_i, \mu), 1)$. W szczególności, gdy A jest ortogonalną projekcją rzędu k na $\mathcal{R}(A)$, to $A = \sum_1^k x_i \Box x_i$, gdzie $\{x_1, \ldots, x_k\}$ jest bazą ortonormalną przestrzeni $\mathcal{R}(A)$, więc

$$(X, AX) = \sum_{i=1}^{k} (x_i, X)^2.$$

Twierdzenie 2.5. *Niech wektor losowy $X \in (V, (\cdot, \cdot))$ ma rozkład normalny $N(\mu, I)$. Jeśli $A \in \mathcal{S}(V)$ jest ortogonalną projekcją rzędu k, to $\mathcal{L}((X, AX)) = \chi_k^2((\mu, A\mu))$.*

Dowód. Niech $\{x_1, x_2, \ldots, x_k\}$ będzie bazą ortonormalną przestrzeni $\mathcal{R}(A)$. Zgodnie z komentarzem poprzedzającym dowodzone twierdzenie, wystarczy pokazać, że zachodzi równość

$$\mathcal{L}\left(\sum_{i=1}^{k} (x_i, X)^2\right) = \chi_k^2\left((\mu, A\mu)\right).$$

Ponieważ $\mathcal{L}(x_i, X) = N((x_i, \mu), 1)$, to $\mathcal{L}(x_i, X)^2 = \chi_1^2((x_i, \mu)^2)$. Ponadto zmienne (x_i, X) są niezależne, więc z twierdzenia 2.4 wynika, że

$$\mathcal{L}\left(\sum_{i=1}^{k} (x_i, X)^2\right) = \chi_k^2\left(\sum_{i=1}^{k} (x_i, \mu)^2\right).$$

Aby zakończyć dowód, wystarczy już tylko zauważyć, że

$$\sum_{i=1}^{k} (x_i, \mu)^2 = \sum_{i=1}^{k} (\mu, (x_i \Box x_i)\mu) = (\mu, A\mu).$$

□

Twierdzenie 2.6. *Niech $X \in (V, (\cdot, \cdot))$ będzie wektorem losowym o rozkładzie normalnym $N(\mu, \Sigma)$, gdzie $\mu \in \mathcal{R}(\Sigma)$. Niech $B = \Sigma^{1/2}$ oraz niech dane będzie odwzorowanie $A \in \mathcal{S}(V)$. Jeśli BAB jest projekcją ortogonalną rzędu k, to wówczas*

$$\mathcal{L}((X, AX)) = \chi_k^2((\mu, A\mu)).$$

Dowód. Jeśli $\mu \in \mathcal{R}(\Sigma)$, to $\mu \in \mathcal{R}(B)$, ponieważ $\mathcal{R}(\Sigma) = \mathcal{R}(B)$, więc $\mu = B\tau$ dla pewnego $\tau \in V$. Niech $Y \in (V, (\cdot, \cdot))$ będzie wektorem losowym takim, że $\mathcal{L}(Y) = N(\tau, I)$. Wówczas $\mathcal{L}(X) = \mathcal{L}(BY)$, czyli $\mathcal{L}\left((X, AX)\right) = \mathcal{L}\left((BY, ABY)\right) = \mathcal{L}\left((Y, BABY)\right)$. Odwzorowanie BAB jest projekcją ortogonalną rzędu k, więc z twierdzenia 2.5 wynika, że $\mathcal{L}(Y, BABY) = \chi_k^2((\tau, BAB\tau)) = \chi_k^2((B\tau, AB\tau)) = \chi_k^2((\mu, A\mu))$. □

Przedstawimy następnie warunki wystarczające niezależności form kwadratowych wektora losowego o rozkładzie normalnym. Załóżmy, że $X \in (V, (\cdot,\cdot))$ ma rozkład $N(\mu, \Sigma)$ i rozważmy formy kwadratowe (X, A_iX) dla $A_i \in \mathcal{S}(V)$, $i = 1, 2$. Przyjmując $B = \Sigma^{1/2}$, mamy $\mathcal{L}(X) = \mathcal{L}(BY + \mu)$ dla pewnego $Y \in (V, (\cdot,\cdot))$, $\mathcal{L}(Y) = N(0, I)$, dlatego

$$L(X, A_iX) = \mathcal{L}(BY+\mu, A_i(BY+\mu)) = \mathcal{L}((Y, BA_iBY)+2(BA_i\mu, Y)+(\mu, A_i\mu)).$$

W związku z tym badanie niezależności form (X, A_iX) sprowadza się do badania niezależności funkcji $Z_i = (Y, C_iY) + (x_i, Y)$, gdzie $C_i \in \mathcal{S}(V)$, $x_i \in V$ (stałe $(\mu, A_i\mu)$ nie mają wpływu na niezależność form (X, A_iX)).

Lemat 2.1. *Dla dowolnych przekształceń $A_i \in \mathcal{S}(V, (\cdot,\cdot))$, $i = 1, 2$, następujące warunki są równoważne:*

(i) $A_1A_2 = 0$,

(ii) $\mathcal{R}(A_1) \perp \mathcal{R}(A_2)$.

Dowód. Załóżmy, że $A_1A_2 = 0$, wówczas $\mathcal{R}(A_2) \subseteq \mathcal{N}(A_1)$. Ponieważ $\mathcal{R}(A_1) \perp \mathcal{N}(A_1)$, co wynika z twierdzenia 1.3, to $\mathcal{R}(A_2) \perp \mathcal{R}(A_1)$. Z drugiej strony jeśli $\mathcal{R}(A_2) \perp \mathcal{R}(A_1)$, to $\mathcal{R}(A_2) \subseteq (\mathcal{R}(A_1))^{\perp} = \mathcal{N}(A_1)$, więc $A_1A_2x = 0$ dla każdego $x \in V$. □

Twierdzenie 2.7. *Niech wektor $Y \in (V, (\cdot,\cdot))$ ma rozkład normalny $N(0, I)$ oraz niech $Z_i = (Y, A_iY) + (x_i, Y)$, gdzie $A_i \in \mathcal{S}(V)$, $x_i \in V$, $i = 1, 2$. Jeżeli $A_1A_2 = 0$, $A_1x_2 = 0$, $A_2x_1 = 0$ oraz $(x_1, x_2) = 0$, to Z_1 i Z_2 są niezależnymi zmiennymi losowymi.*

Dowód. Wykażemy, że zmienne Z_1 i Z_2 są funkcjami dwóch niezależnych wektorów losowych. Przez $[\cdot,\cdot]$ będziemy każdorazowo rozumieć standardowy iloczyn skalarny wprowadzony na odpowiedniej sumie prostej przestrzeni unitarnych (wzorem (1.11)). Niech P_i będą projekcjami ortogonalnymi na $\mathcal{R}(A_i)$. Ponieważ $\mathcal{R}(A_1) \perp \mathcal{N}(A_1)$, to dowolny wektor $v \in V$ można przestawić w postaci $v = v_1 + v_2$, $v_1 \in \mathcal{R}(A_1)$, $v_2 \in \mathcal{N}(A_1)$, dlatego dla dowolnego $v \in V$ mamy $A_1v = A_1v_1$ oraz $P_1A_1P_1v = P_1A_1v_1 = A_1v_1$, czyli $P_1A_1P_1 = A_1$. Podobnie pokazuje się, że $P_2A_2P_2 = A_2$. Wobec tego zmienne $Z_i = (P_iY, A_iP_iY) + (x_i, Y)$ są funkcjami wektorów $\{P_iY, (x_i, Y)\} \in (V \oplus \mathbb{R}, [\cdot,\cdot])$, $i = 1, 2$. Dla dowolnego wektora $\{\{y_1, \alpha_1\}, \{y_2, \alpha_2\}\} \in (V \oplus R \oplus V \oplus R, [\cdot,\cdot])$ zmienna losowa

$$(y_1, P_1Y)+\alpha_1(x_1, Y)+(y_2, P_2Y)+\alpha_2(x_2, Y) = (P_1y_1+\alpha_1x_1+P_2y_2+\alpha_2x_2, Y)$$

ma rozkład normalny, co dowodzi, że łączny rozkład wektorów $\{P_1Y, (x_1, Y)\}$ oraz $\{P_2Y, (x_2, Y)\}$ jest rozkładem normalnym. Wystarczy już tylko pokazać, że wektory $\{P_1Y, (x_1, Y)\}$ i $\{P_2Y, (x_2, Y)\}$ są nieskorelowane, wówczas na mocy twierdzenia 2.2 są one niezależne. Dla dowolnych $y_1, y_2 \in V$,

$\alpha_1, \alpha_2 \in \mathbb{R}$, wykonując elementarne przekształcenia, otrzymujemy

$$\operatorname{cov}\{[\{y_1, \alpha_1\}, \{P_1 Y, (x_1, Y)\}], [\{y_2, \alpha_2\}, \{P_2 Y, (x_2, Y)\}]\} = \\ = (y_1, P_1 P_2 y_2) + \alpha_1 (P_2 x_1, y_2) + \alpha_2 (P_1 x_2, y_1) + \alpha_1 \alpha_2 (x_1, x_2) = 0,$$

ponieważ na mocy lematu 2.1 mamy $\mathcal{R}(A_1) \perp \mathcal{R}(A_2)$, więc $P_1 P_2 = 0$, natomiast pozostałe składniki są zerami, co wynika wprost z przyjętych założeń. $\square$

Powyższe twierdzenie można uogólnić na większą liczbę form kwadratowych, przyjmując, że założenia twierdzenia 2.7 mają być spełnione dla każdych dwóch przekształceń spośród wszystkich rozpatrywanych form kwadratowych.

2.2 Rozkład warunkowy wektora losowego o rozkładzie normalnym

Rozważmy wektory losowe $X_1 \in (V_1, (\cdot,\cdot)_1)$ oraz $X_2 \in (V_2, (\cdot,\cdot)_2)$. Niech wektor $\{X_1, X_2\} \in (V_1 \oplus V_2, [\cdot,\cdot])$ ma rozkład normalny o wartości oczekiwanej $\{\mu_1, \mu_2\}$ oraz kowariancji postaci

$$\operatorname{Cov}(\{X_1, X_2\}) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix}, \tag{2.8}$$

gdzie $[\cdot,\cdot]$ jest zadany wzorem (1.11). Wówczas wektory X_1 oraz X_2 mają rozkłady normalne. Ponieważ

$$\mathbb{E}[\{v_1, v_2\}, \{X_1, X_2\}] = [\{v_1, v_2\}, \{\mu_1, \mu_2\}] = (v_1, \mu_1)_1 + (v_2, \mu_2)_2$$

i dla $v_2 = 0 \in V_2$ oraz dowolnego $v_1 \in V_1$ mamy

$$\mathbb{E}[\{v_1, 0\}, \{X_1, X_2\}] = \mathbb{E}(v_1, X_1)_1 = (v_1, \mu_1)_1,$$

to $\mathbb{E}(X_1) = \mu_1 \in V_1$. Wyznaczmy teraz kowariancję wektora X_1. Na mocy twierdzenia 1.18 dla dowolnych wektorów $\{v_1, v_2\}, \{v_1', v_2'\} \in V_1 \oplus V_2$ zachodzą równości

$$\operatorname{cov}\{[\{v_1, v_2\}, \{X_1, X_2\}], [\{v_1', v_2'\}, \{X_1, X_2\}]\} = \\ = [\{v_1, v_2\}, \{\Sigma_{11} v_1' + \Sigma_{12} v_2', \Sigma_{12}' v_1' + \Sigma_{22} v_2'\}].$$

W szczególności, dla $v_2 = v_2' = 0 \in V_2$, mamy

$$\operatorname{cov}\{(v_1, X_1)_1, (v_1', X_1)_1\} = [\{v_1, 0\}, \{\Sigma_{11} v_1', \Sigma_{12}' v_1'\}] = (v_1, \Sigma_{11} v_1')_1,$$

stąd oraz z faktu, że nieujemna określoność odwzorowania Σ implikuje nieujemną określoność odwzorowań Σ_{11} oraz Σ_{22}, wynika, iż $\operatorname{Cov}(X_1) = \Sigma_{11}$. Analogicznie wyznacza się momenty wektora X_2.

Twierdzenie 2.8. *Niech dane będą wektory losowe* $X_1 \in (V_1, (\cdot,\cdot)_1)$, $X_2 \in (V_2, (\cdot,\cdot)_2)$ *oraz wektor* $\{X_1, X_2\} \in (V_1 \oplus V_2, [\cdot,\cdot])$ *o rozkładzie normalnym z wartością oczekiwaną* $\{\mu_1, \mu_2\}$ *oraz kowariancją* Σ *postaci* (2.8). *Wówczas rozkład warunkowy wektora losowego* X_1, *przy* $X_2 = v_2 \in V_2$, *jest rozkładem normalnym o wartości oczekiwanej*

$$\mathbb{E}(X_1|X_2 = v_2) = \mu_1 + \Sigma_{12}\Sigma_{22}^{-}(v_2 - \mu_2)$$

oraz kowariancji postaci

$$\mathrm{Cov}(X_1|X_2 = v_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{12}',$$

gdzie Σ_{22}^{-} *oznacza uogólnioną odwrotność odwzorowania* Σ_{22}.

Dowód. W dowodzie twierdzenia wykorzystamy postać funkcji charakterystycznej wektora o rozkładzie normalnym, dlatego zauważmy, że jeżeli wektor $Y \in (V, (\cdot,\cdot))$ ma rozkład normalny o parametrach μ_Y oraz Σ_Y, to dla każdego $v \in V$ zmienna (v, Y) ma rozkład normalny o wartości oczekiwanej (v, μ_Y) oraz wariancji równej $(v, \Sigma_Y v)$. Wynika stąd, że funkcja charakterystyczna zmiennej (v, Y) ma postać

$$\Phi_{(v,Y)}(t) = \mathbb{E}\left(\exp[it(v, Y)]\right) = \exp\left[it(v, \mu_Y) - \frac{1}{2}t^2(v, \Sigma_Y v)\right], \qquad t \in \mathbb{R}.$$

Przyjmując $t = 1$, uzyskujemy postać funkcji charakterystycznej wektora Y

$$\Phi_Y(v) = \mathbb{E}\left(\exp[i(v, Y)]\right) = \exp\left[i(v, \mu_Y) - \frac{1}{2}(v, \Sigma_Y v)\right], \qquad v \in V. \quad (2.9)$$

Rozważmy wektory $X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2 \in V_1$ oraz $X_2 \in V_2$. Dla dowolnych elementów $v_1 \in V_1$, $v_2 \in V_2$ zachodzi równość

$$[\{v_1, v_2\}, \{X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2, X_2\}] = [\{v_1, v_2 - \Sigma_{22}^{-}\Sigma_{12}'v_1\}, \{X_1, X_2\}].$$

Ponieważ wektor $\{X_1, X_2\}$ ma rozkład normalny, to z powyższego równania wynika, że dla każdego $v_1 \in V_1$ zmienna losowa $(v_1, X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2)_1$ ma rozkład normalny (wystarczy przyjąć $v_2 = 0$), czyli wektor losowy $X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2$ ma rozkład normalny w przestrzeni V_1. Ponadto, z powyższego równania wnioskujemy, że wektory $X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2$ oraz X_2 mają łączny rozkład normalny. Zgodnie z twierdzeniem 1.19 wektory te są nieskorelowane, stąd na mocy twierdzenia 2.2 są niezależne. Z twierdzenia 1.19 otrzymujemy również następującą postać kowariancji wektora $X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2$:

$$\mathrm{Cov}(X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2) = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{12}'.$$

Natomiast, z własności wartości oczekiwanej wynika, że

$$\mathbb{E}(X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2) = \mu_1 - \Sigma_{12}\Sigma_{22}^{-}\mu_2.$$

Wykorzystując to, co zostało do tej pory stwierdzone, rozważmy następującą funkcję charakterystyczną:

$$
\begin{aligned}
\Phi_{X_1|X_2=v_2}(v_1) &= \mathbb{E}\left(\exp\left[i(v_1, X_1)_1\right] | X_2 = v_2\right) \\
&= \mathbb{E}\left(\exp\left[i(v_1, X_1)_1 - i(v_1, \Sigma_{12}\Sigma_{22}^{-}X_2)_1 + i(v_1, \Sigma_{12}\Sigma_{22}^{-}X_2)_1\right] | X_2 = v_2\right) \\
&= \exp\left[i(v_1, \Sigma_{12}\Sigma_{22}^{-}v_2)_1\right] \mathbb{E}\left(\exp\left[i(v_1, X_1 - \Sigma_{12}\Sigma_{22}^{-}X_2)_1\right]\right) \\
&= \exp\left[i(v_1, \Sigma_{12,}\Sigma_{22}^{-}v_2)_1\right] \times \\
&\qquad \times \exp\left[i(v_1, \mu_1 - \Sigma_{12}\Sigma_{22}^{-}\mu_2)_1 - \frac{1}{2}\left(v_1, (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{12}')v_1\right)_1\right] \\
&= \exp\left[i\left(v_1, \mu_1 + \Sigma_{12}\Sigma_{22}^{-}(v_2 - \mu_2)\right)_1 - \frac{1}{2}\left(v_1, (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-}\Sigma_{12}')v_1\right)_1\right],
\end{aligned}
$$

gdzie $v_1 \in V_1$. Ponieważ funkcja charakterystyczna jednoznacznie wyznacza rozkład, otrzymujemy tezę twierdzenia.

□

W następującym przykładzie wykorzystamy powyższe twierdzenie do wyznaczenia warunkowego rozkładu wybranej liczby początkowych kolumn macierzy losowej skonstruowanej w przykładzie 2.1 względem pozostałych kolumn tej macierzy.

Przykład 2.3. *Rozważmy przestrzeń $\mathbb{R}^m$ ze standardowym iloczynem skalarnym. Niech $X_1, X_2, \dots, X_n$ będą niezależnymi wektorami losowymi o rozkładzie normalnym na $\mathbb{R}^m$ z wartością oczekiwaną $\mu \in \mathbb{R}^m$ oraz kowariancją Σ, gdzie Σ jest dodatnio określone. Wówczas, zgodnie z tym, co zostało pokazane w przykładzie* 2.1*, macierz X o wierszach $X_1', X_2', \dots, X_n'$ ma rozkład normalny w przestrzeni $(\mathcal{M}_{n\times m}, \langle\cdot,\cdot\rangle)$ z parametrami $e\mu'$ oraz $I_n \otimes \Sigma$, gdzie $e \in \mathbb{R}^n$ jest wektorem złożonym z n jedynek oraz $\langle A, B\rangle = \mathrm{tr}(AB')$ dla dowolnych macierzy $A, B \in \mathcal{M}_{n\times m}$.*

W celu wyznaczenia warunkowego rozkładu pierwszych p kolumn macierzy X względem ostatnich q kolumn, $p + q = m$, podzielmy X na macierze $W_1 \in (\mathcal{M}_{n\times p}, \langle\cdot,\cdot\rangle_1)$ oraz $W_2 \in (\mathcal{M}_{n\times q}, \langle\cdot,\cdot\rangle_2)$ złożone odpowiednio z pierwszych p i ostatnich q kolumn macierzy X (iloczyny skalarne $\langle\cdot,\cdot\rangle_1$ oraz $\langle\cdot,\cdot\rangle_2$ definiujemy analogicznie jak w przypadku przestrzeni $\mathcal{M}_{n\times m}$). W ten sam sposób podzielmy macierz $e\mu'$ na $e\mu_1' \in \mathcal{M}_{n\times p}$ oraz $e\mu_2' \in \mathcal{M}_{n\times q}$. Macierz X możemy rozpatrywać jako wektor $\{W_1, W_2\}$ o rozkładzie normalnym w przestrzeni $(\mathcal{M}_{n\times p} \oplus \mathcal{M}_{n\times q}, [\cdot,\cdot])$ z wartością oczekiwaną $\{e\mu_1', e\mu_2'\}$, gdzie iloczyn skalarny $[\cdot,\cdot]$ jest zadany zgodnie z równaniem (1.11)*. Aby wyznaczyć kowariancję wektora $\{W_1, W_2\}$, zauważmy, że dzieląc macierz X dokonaliśmy podziału każdego z wektorów X_i na wektory $X_{ip} \in \mathbb{R}^p$ oraz $X_{iq} \in \mathbb{R}^q$, których transpozycje stanowią wiersze macierzy W_1 i W_2. Z tego, że kowariancja wektora X_i istnieje wynika, że istnieją kowariancje* $\mathrm{Cov}(X_{ip}) = \Sigma_{11}$,

$\mathrm{Cov}(X_{iq}) = \Sigma_{22}$ *oraz* $\mathrm{Cov}\{X_{ip}, X_{iq}\} = \Sigma_{12}$ *i na mocy twierdzenia* 1.18 *odwzorowanie* Σ *możemy zapisać w postaci*

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}' & \Sigma_{22} \end{pmatrix}.$$

Utożsamiając dowolną macierz $A \in \mathcal{M}_{n\times m}$ *z odpowiednim wektorem* $\{A_1, A_2\}$ *w przestrzeni* $\mathcal{M}_{n\times p} \oplus \mathcal{M}_{n\times q}$*, można pokazać, że zachodzi równość*

$$(I_n \otimes \Sigma)A = \begin{pmatrix} I \otimes \Sigma_{11} & I \otimes \Sigma_{12} \\ I \otimes \Sigma_{12}' & I \otimes \Sigma_{22} \end{pmatrix} \{A_1, A_2\} = S\{A_1, A_2\}.$$

Wobec tego $\mathrm{Cov}(\{W_1, W_2\}) = S$ *i na mocy twierdzenia* 2.8 *warunkowy rozkład wektora* W_1*, przy* $W_2 = w_2 \in \mathcal{M}_{n\times q}$*, jest rozkładem normalnym o wartości oczekiwanej*

$$\begin{aligned} \mathbb{E}(W_1 | W_2 = w_2) &= e\mu_1' + (I \otimes \Sigma_{12})(I \otimes \Sigma_{22})^{-1}(w_2 - e\mu_2') \\ &= e\mu_1' + (I \otimes \Sigma_{12}\Sigma_{22}^{-1})(w_2 - e\mu_2') \\ &= e\mu_1' + (w_2 - e\mu_2')(\Sigma_{22}^{-1}\Sigma_{12}') \end{aligned}$$

oraz kowariancji postaci

$$\begin{aligned} \mathrm{Cov}(W_1 | W_2 = w_2) &= I \otimes \Sigma_{11} - (I \otimes \Sigma_{12})(I \otimes \Sigma_{22})^{-1}(I \otimes \Sigma_{12}') \\ &= I \otimes (\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{12}'). \end{aligned}$$

W obu przypadkach skorzystaliśmy z faktu, że $(I \otimes \Sigma_{22})^{-1} = I \otimes \Sigma_{22}^{-1}$ *oraz z pozostałych własności iloczynu Kroneckera odwzorowań.*

Aby wyznaczyć rozkłady macierzy W_1 *oraz* W_2*, wystarczy ustalić macierze blokowe* $A = [I_p, 0] \in \mathcal{M}_{p\times m}$*,* $B = [0, I_q] \in \mathcal{M}_{q\times m}$ *i zauważyć, że*

$$\mathcal{L}(W_1) = \mathcal{L}(XA') = \mathcal{L}((I \otimes A)X) = N(e\mu_1', I \otimes \ \Sigma_{11}),$$

$$\mathcal{L}(W_2) = \mathcal{L}(XB') = \mathcal{L}((I \otimes B)X) = N(e\mu_2', I \otimes \ \Sigma_{22}).$$

2.3 Funkcja gęstości rozkładu normalnego na przestrzeni wektorowej

Rozkład normalny wektora losowego $X \in (V, (\cdot,\cdot))$ został zdefiniowany przez postulowanie normalności rozkładu wszystkich zmiennych losowych będących formami liniowymi wektora X. Dzięki temu założenie o dodatniej określoności kowariancji wektora X nie było konieczne. W bieżącym rozdziale wyznaczymy gęstość rozkładu normalnego, w przypadku nieosobliwej kowariancji. Oczywiście musimy wcześniej zdefiniować miarę dominującą w przestrzeni V. W tym celu rozpoczniemy od zadania na dowolnej skończenie wymiarowej przestrzeni unitarnej miary niezmienniczej ze względu na translacje i skończonej na zbiorach zwartych. Następnie sformułujemy twierdzenie podające postać gęstości rozkładu normalnego względem zdefiniowanej miary.

Wiadomo, że dowolna n-wymiarowa rzeczywista przestrzeń wektorowa V jest izomorficzna z przestrzenią $\mathbb{R}^n$. Potraktujmy obie przestrzenie jako unitarne i rozpatrzmy izometrię T, która przypisuje wektorom e_i bazy kanonicznej w $\mathbb{R}^n$ wektory v_i bazy ortonormalnej w $(V, (\cdot, \cdot))$, czyli

$$T(e_i) = v_i, \quad i = 1, \dots, n. \tag{2.10}$$

Wykorzystując własności odwzorowania T, możemy zadać na V miarę za pomocą miary Lebesgue'a w $\mathbb{R}^n$. Niech $\mathcal{B}(V)$ oznacza σ-ciało podzbiorów borelowskich przestrzeni V. Odwzorowanie T^{-1} jest liniowe, a więc również ciągłe, stąd dla każdego zbioru $B \in \mathcal{B}(V)$ mamy $T^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$. Oznaczmy przez l miarę Lebesgue'a na $\mathbb{R}^n$. Miarą Lebesgue'a na V będziemy nazywać miarę ν_0 określoną wzorem

$$\nu_0(B) = l\left(T^{-1}(B)\right), \quad B \in \mathcal{B}(V).$$

Dzięki własnościom miary Lebesgue'a na $\mathbb{R}^n$ (która jest niezmiennicza ze względu na translacje i skończona na zbiorach zwartych) dla dowolnego zbioru $B \in \mathcal{B}(V)$, dowolnego elementu $v \in V$ oraz zbioru $B+v = \{x+v : x \in B\}$ zachodzą następujące równości:

$$\nu_0(B+v) = l(T^{-1}(B+v)) = l\left(T^{-1}(B) + T^{-1}(v)\right) = l\left(T^{-1}(B)\right) = \nu_0(B).$$

Ponadto, miara ν_0 każdego zbioru zwartego w V jest skończona, ponieważ zbiór zwarty K w przestrzeni unormowanej jest domknięty, czyli $K \in \mathcal{B}(V)$, stąd $T^{-1}(K) \in \mathcal{B}(\mathbb{R}^n)$ i $T^{-1}(K)$ jest zwarty (bo odwzorowanie T^{-1} jest ciągłe). Wynika stąd, że miara ν_0 jest niezmiennicza ze względu na translacje i skończona na zbiorach zwartych w V.

Następujące twierdzenie podaje postać gęstości rozkładu normalnego względem miary ν_0, w przypadku gdy gęstość ta istnieje.

Twierdzenie 2.9. ([2], str. 122) *Niech wektor losowy $X \in (V, (\cdot, \cdot))$ ma rozkład normalny o wartości oczekiwanej μ oraz kowariancji Σ. Jeżeli odwzorowanie Σ jest dodatnio określone, to gęstość rozkładu wektora X względem miary ν_0 ma postać*

$$p(v) = (2\pi)^{-n/2}(\det(\Sigma))^{-1/2} \exp\left[-\frac{1}{2}\left(v-\mu, \Sigma^{-1}(v-\mu)\right)\right], \quad v \in V,$$

gdzie n to wymiar przestrzeni V.

Bibliografia

[1] Barra J. R., *Matematyczne podstawy statystyki*, PWN, Warszawa, 1982.

[2] Eaton M. L., *Multivariate statistics. A Vector Space Approach*, John Wiley & Sons, New York, 1983.

[3] Halmos P. R., *Finite Dimensional Vector Spaces. Undergraduate Texts in Mathematics*, Springer Verlag, New York, 1958.

[4] Krzyśko M., *Statystyka matematyczna*, Wydawnictwo Naukowe UAM, Poznań, 1996.

[5] Oja H., *Multivariate Nonparametric Methods with R. An Approach Based on Spatial Signs and Ranks*, Springer, New York, 2010.

[6] Rao C. R., *Modele liniowe statystyki matematycznej*, PWN, Warszawa, 1982.

[7] Watson G. W., A note on maximum likelihood, *Sankhya*, 26A, 303-304, 1964.

[8] Wei W. W. S., *Time Series Analysis: Univariate and Multivariate Methods*, Pearson Addison Wesley, Boston, 2006.