

1 Definicje i twierdzenia z algebry liniowej.

Przy analizie wielowymiarowego modelu regresji liniowej w sposób intensywny korzysta się z pojęć oraz faktów z algebry liniowej. Dlatego przypomnimy kilka definicji oraz twierdzeń, z których będziemy korzystać w dalszej części wykładu.

1.1 Iloczyn skalarny i rzut ortogonalny.

Jeśli $\mathbf{x} = (x_1, x_2, \dots, x_n)$ i $\mathbf{y} = (y_1, y_2, \dots, y_n)$ są dwoma wektorami przestrzeni euklidesowej $\mathbb{R}^n$. Iloczyn skalarny dwóch wektorów $\mathbf{x}$ i $\mathbf{y}$ określa się jako

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Mówimy, że dwa wektory są ortogonalne, jeśli ich iloczyn skalarny jest równy zero. Jeśli θ jest kątem między wektorami $(x_1, x_2, \dots, x_n)$ i $(y_1, y_2, \dots, y_n)$, to

$$\cos(\theta) = \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}| |\mathbf{y}|} = \frac{x_1 y_1 + x_2 y_2 + \dots + x_n y_n}{\sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \sqrt{y_1^2 + y_2^2 + \dots + y_n^2}}$$

Kąt między dwoma wektorami ortogonalnymi wynosi 90° , tzn. że są one prostopadłe. Jednym z ważnych zastosowań iloczynu skalarnego jest wykorzystanie jego w definicji projekcji. Projekcja wektora $\mathbf{y}$ na inny wektor $\mathbf{x}$ tworzy nowy wektor, który ma ten sam kierunek jako wektor $\mathbf{x}$ i długość równą $|\mathbf{y}| \cos(\theta)$, gdzie $|\mathbf{y}|$ oznacza długość wektora $\mathbf{y}$, a θ jest kątem między wektorami $\mathbf{x}$ i $\mathbf{y}$. Będziemy oznaczać projekcję $\mathbf{y}$ na wektor $\mathbf{x}$ przez $\mathbf{P}_x \mathbf{y}$. Wektor $\mathbf{P}_x \mathbf{y}$ można wyrazić jako

$$\begin{aligned} \mathbf{P}_x \mathbf{y} &= |\mathbf{y}| \cos(\theta) \frac{\mathbf{x}}{|\mathbf{x}|} = |\mathbf{y}| \frac{\mathbf{x} \cdot \mathbf{y}}{|\mathbf{x}| |\mathbf{y}|} \frac{\mathbf{x}}{|\mathbf{x}|} \\ &= \frac{x_1 y_1 + x_2 y_2 + \dots + x_n y_n}{x_1^2 + x_2^2 + \dots + x_n^2} \mathbf{x} = \lambda \mathbf{x}, \end{aligned}$$

gdzie λ jest skalar

$$\lambda = \frac{x_1 y_1 + x_2 y_2 + \dots + x_n y_n}{x_1^2 + x_2^2 + \dots + x_n^2} = \frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{x} \cdot \mathbf{x}}.$$

Niech $\mathbf{x}$ i $\mathbf{y}$ będą dwoma wektorami w $\mathbb{R}^n$. Rozważmy wektor $\mathbf{e} = \lambda \mathbf{x} - \mathbf{y}$, dla którego $\lambda = (\mathbf{x} \cdot \mathbf{y})/(\mathbf{x} \cdot \mathbf{x})$. Wektor $\mathbf{e}$ jest prostopadły do wektora $\mathbf{x}$. Aby to zobaczyć, wystarczy po prostu obliczyć iloczyn skalarny $\mathbf{e} \cdot \mathbf{x}$.

$$\mathbf{e} \cdot \mathbf{x} = (\lambda \mathbf{x} - \mathbf{y}) \cdot \mathbf{x} = \lambda \mathbf{x} \cdot \mathbf{x} - \mathbf{x} \cdot \mathbf{y} = \left(\frac{\mathbf{x} \cdot \mathbf{y}}{\mathbf{x} \cdot \mathbf{x}} \right) \mathbf{x} \cdot \mathbf{x} - \mathbf{x} \cdot \mathbf{y} = 0$$

Jeśli S jest p -wymiarową podprzestrzenią przestrzeni wektorowej V , wówczas możliwe jest rzutowanie wektorów z przestrzeni V na S . Jeśli podprzestrzeń S ma bazę ortonormalną $(w_1, w_2, \dots, w_p)$, to dla dowolnego wektora $\mathbf{y}$ z przestrzeni V , rzut wektora $\mathbf{y}$ na podprzestrzeń S , $\mathbf{P}_S \mathbf{y}$, ma postać

$$\mathbf{P}_S \mathbf{y} = \sum_{i=1}^p (\mathbf{y} \cdot \mathbf{w}_i) \mathbf{w}_i.$$

Rzut wektora na przestrzeń liniową S jest w rzeczywistości liniowym przekształceniem wektora i może być reprezentowany przez macierz projekcji pomnożonej przez wektor. Macierz projekcyjna $\mathbf{P}$ jest macierzą kwadratową rozmiaru $n \times n$, która zadaje rzut z $\mathbb{R}^n$ na podprzestrzeń S . Kolumny macierzy $\mathbf{P}$ są rzutami wektorów bazowych, a S jest obrazem przekształcenia $\mathbf{P}$. Dla macierzy definiującej rzutu zachodzą następujące twierdzenia.

Twierdzenie. *Macierz kwadratowa $\mathbf{P}$ jest macierzą projekcji wtedy i tylko wtedy, gdy jest macierzą idempotentną, $\mathbf{P}^2 = \mathbf{P}$.*

Twierdzenie. *Niech $U = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k)$ będzie bazą ortonormalną dla podprzestrzeni W przestrzeni liniowej V . Macierz UU^T jest macierzą projekcji V na W . To znaczy, dla dowolnego wektora $\mathbf{v} \in V$ rzut $\mathbf{v}$ na W jest $\text{Proj}_W \mathbf{v} = UU^T \mathbf{v}$.*

Macierz UU^T nazywa się macierzą rzutowania dla podprzestrzeni W . Nie zależy ona od wyboru bazy ortonormalnej. Jeśli baza przestrzeni W nie jest ortonormalna, nadal możemy skonstruować macierz projekcji. Wynika to z następującego twierdzenia.

Twierdzenie. *Niech $\mathbf{A} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_k)$ będzie dowolną bazą podprzestrzeni W przestrzeni V . Macierz $\mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$ jest macierzą projekcji V na W . To znaczy, że dla dowolnego wektora $\mathbf{v} \in V$ projekcja $\mathbf{v}$ na W jest postaci*

$$\text{Proj}_W \mathbf{v} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{v}.$$

Następujący lemat jest istotnym elementem dowodów powyższych trzech twierdzeń.

Lemat. *Założmy, że $\mathbf{A}$ jest macierzą rozmiaru $n \times k$, której kolumny są liniowo niezależne. Wtedy macierz $\mathbf{A} \mathbf{A}^T$ jest odwracalna.*

Macierz projekcyjna odgrywa ważną rolę w analizie modelu wielowymiarowej regresji liniowej $Y = \mathbf{X}\boldsymbol{\beta} + \epsilon$. Macierz kwadratowa, $\mathbf{P} = \mathbf{X}(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}^T$, zdefiniowana jest przy użyciu macierzy planu. Można łatwo sprawdzić, że macierz $\mathbf{P}$ jest macierzą idempotentną:

$$\mathbf{P}^2 = \mathbf{X}(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}^T \mathbf{X}(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}^T = \mathbf{P}.$$

Zatem $\mathbf{P} = \mathbf{X}(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}^T$ jest macierzą projekcji. Dodatkowo, macierz

$$\mathbf{I} - \mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}^T$$

jest również macierzą idempotentną.

$$(\mathbf{I} - \mathbf{P})^2 = \mathbf{I} - 2\mathbf{P} + \mathbf{P}^2 = \mathbf{I} - 2\mathbf{P} + \mathbf{P} = \mathbf{I} - \mathbf{P}.$$

wynika stąd, że macierz $\mathbf{I} - \mathbf{P} = \mathbf{I} - \mathbf{X}(\mathbf{X}\mathbf{X}^T)^{-1} \mathbf{X}^T$ jest macierzą projekcji.

W dalszej części wykładu zobaczymy, jak tak określone macierze projekcji można wykorzystać do analizy modelu regresji liniowej oraz konstrukcji najlepszego liniowego estymatora nieobciążonego parametrów regresji liniowej.

1.2 Macierze idempotentne.

W tej części wykładu omawiamy własności macierzy idempotentnych oraz ich związek z modelem wielowymiarowej regresji liniowej. Najpierw przypomnimy definicję macierzy idempotentnej.

Definicja. Macierz symetryczna $\mathbf{A}$ rozmiaru $n \times n$ jest idempotentna, jeśli

$$\mathbf{A}^2 = \mathbf{A}.$$

Niech $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_k)$ będzie k -wymiarowym wektorem, a $\mathbf{A}$ macierzą rozmiaru $k \times k$. Rozpatrzy formę kwadratową $\boldsymbol{\alpha}^T \mathbf{A} \boldsymbol{\alpha}$ wektora $(\alpha_1, \alpha_2, \dots, \alpha_k)$. Gdy $\mathbf{A}$ jest macierzą idempotentną, odpowiadająca jej forma kwadratowa ma pewne szczególne właściwości. Formy kwadratowa z idempotentnymi macierzami pojawiają się często podczas analizy modelu wielowymiarowej regresji liniowej. Dlatego przedstawimy kilka własności macierzy idempotentnych.

Twierdzenie. Niech $\mathbf{A}$ będzie idempotentną macierzą rozmiaru $n \times n$ oraz rzędu p . Wtedy wartości własne macierzy $\mathbf{A}$ są równe 1 lub 0.

Dowód. Niech λ_i i v_i będą odpowiednio wartością własną i odpowiadającym jej unormowanemu wektorowi własnemu macierzy $\mathbf{A}$. Wtedy $\mathbf{A}v_i = \lambda_i v_i$ i $v_i^T \mathbf{A}v_i = \lambda_i v_i^T v_i = \lambda_i$. Z drugiej strony, skoro $\mathbf{A}^2 = \mathbf{A}$, możemy napisać

$$\lambda_i = v_i^T \mathbf{A}v_i = v_i^T \mathbf{A}^2 v_i = v_i^T \mathbf{A}^T \mathbf{A} v_i = (\mathbf{A}v_i)^T \mathbf{A}v_i = (\lambda_i v_i)^T (\lambda_i v_i) = \lambda_i^2.$$

Zatem $\lambda_i(\lambda_i - 1) = 0$, co daje $\lambda_i = 1$, albo $\lambda_i = 0$, i kończy dowód.

Ponieważ p wartości własnych macierzy $\mathbf{A}$ jest równa 1, a $n - p$ wartości własne jest równa zero, więc rząd macierzy idempotentnej $\mathbf{A}$ jest równa sumie niezerowych wartości własne. Rząd macierzy idempotentnej można wyrazić również w terminach jej śladu. Następujące twierdzenie podaje postać tej zależności.

Twierdzenie. Jeśli $\mathbf{A}$ jest macierzą idempotentną, to $\text{tr}(\mathbf{A}) = \text{rank}(\mathbf{A}) = p$.

1.3 Formy kwadratowe wektorów losowych.

Rozpatrzy formę kwadratową wektora losowego $\mathbf{y}^T = (y_1, y_2, \dots, y_n)$ postaci $\mathbf{y}^T \mathbf{A} \mathbf{y}$. Poniższe twierdzenie podaje postać wartości oczekiwanej i wariancji wektora losowego $\mathbf{y}$ o niezależnych składowych.

Twierdzenie. Niech $\mathbf{y}^T = (y_1, y_2, \dots, y_n)$ będzie wektorem losowym z wektorem wartości oczekiwanych $\boldsymbol{\mu}^T = (\mu_1, \mu_2, \dots, \mu_n)$ oraz takiej samej wariancji σ^2 każdej składowej wektora. Dodatkowo zakładamy, że składowe wektora $\mathbf{y}$, są niezależnymi zmiennymi losowymi. Niech $\mathbf{A}$ będzie macierzą rozmiaru $n \times n$. Wartość oczekiwana formy kwadratowej $\mathbf{y}^T \mathbf{A} \mathbf{y}$ jest postaci

$$E(\mathbf{y}^T \mathbf{A} \mathbf{y}) = \sigma^2 \text{tr}(\mathbf{A}) + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}.$$

Dowód. Po pierwsze zaobserwujemy, że

$$\mathbf{y}^T \mathbf{A} \mathbf{y} = (\mathbf{y} - \boldsymbol{\mu})^T \mathbf{A} (\mathbf{y} - \boldsymbol{\mu}) + 2\boldsymbol{\mu}^T \mathbf{A} (\mathbf{y} - \boldsymbol{\mu}) + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}.$$

Możemy więc napisać

$$\begin{aligned}
E(\mathbf{y}' \mathbf{A} \mathbf{y}) &= E[(\mathbf{y} - \boldsymbol{\mu})' \mathbf{A} (\mathbf{y} - \boldsymbol{\mu})] + 2E[\boldsymbol{\mu}' \mathbf{A} (\mathbf{y} - \boldsymbol{\mu})] + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu} \\
&= E\left[\sum_{i,j=1}^n a_{ij}(y_i - \mu_i)(y_j - \mu_j)\right] + 2\boldsymbol{\mu}' \mathbf{A} E(\mathbf{y} - \boldsymbol{\mu}) + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu} \\
&= \sum_{i=1}^n a_{ii}E(y_i - \mu_i)^2 + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu} = \sigma^2 \text{tr}(\mathbf{A}) + \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}.
\end{aligned}$$

Tak więc otrzymujemy tezę twierdzenia.

Postać wariancji formy kwadratowej $\mathbf{y}^T \mathbf{A} \mathbf{y}$ podaje następujące twierdzenie.

Twierdzenie. Niech $\mathbf{y}$ będzie n -wymiarowym wektorem losowym z wektorem wartości oczekiwanych $\boldsymbol{\mu}^T = (\mu_1, \mu_2, \dots, \mu_n)$ oraz takiej samej wariancji σ^2 każdej składowej wektora. Zakładamy dodatkowo, że składowe wektora $\mathbf{y}$, $(y_1, y_2, \dots, y_n)$ są niezależnymi zmiennymi losowymi. Niech $\mathbf{A}$ będzie macierzą symetryczną rozmiaru $n \times n$. Wariancja formy kwadratowej $\mathbf{y}^T \mathbf{A} \mathbf{y}$ jest równa

$$\begin{aligned}
\text{Var}(\mathbf{y}^T \mathbf{A} \mathbf{y}) &= (\boldsymbol{\mu}^{(4)} - 3\sigma^2) \mathbf{a}^T \mathbf{a} + \sigma^4 (2\text{tr}(\mathbf{A}^2) + [\text{tr}(\mathbf{A})]^2) \\
&\quad + 4\sigma^2 \boldsymbol{\mu}^T \mathbf{A}^2 \boldsymbol{\mu} + 4\boldsymbol{\mu}^{(3)} \mathbf{a}^T \mathbf{A} \boldsymbol{\mu}.
\end{aligned}$$

gdzie $\mu_i^{(4)} = E(y_i - \mu_i)^4$, $\mu_i^{(3)} = E(y_i - \mu_i)^3$, oraz $\mathbf{a}^T = (a_{11}, a_{22}, \dots, a_{nn})$.

Dowód. Niech $\mathbf{Z} = \mathbf{y} - \boldsymbol{\mu}$ oraz $\mathbf{b} = (b_1, b_2, \dots, b_n) = \boldsymbol{\mu} \mathbf{A}$ możemy napisać

$$\begin{aligned}
\mathbf{y}^T \mathbf{A} \mathbf{y} &= (\mathbf{y}^T - \boldsymbol{\mu}) \mathbf{A} (\mathbf{y} - \boldsymbol{\mu}) + 2\boldsymbol{\mu}^T \mathbf{A} (\mathbf{y} - \boldsymbol{\mu}) + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu} \\
&= \mathbf{Z}^T \mathbf{A} \mathbf{Z} + 2\mathbf{b} \mathbf{Z} + \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}.
\end{aligned}$$

Tak więc

$$\text{Var}(\mathbf{y}^T \mathbf{A} \mathbf{y}) = \text{Var}(\mathbf{Z}^T \mathbf{A} \mathbf{Z}) + 4\text{Var}(\mathbf{b} \mathbf{Z}) + 4\text{Cov}(\mathbf{Z}^T \mathbf{A} \mathbf{Z}, \mathbf{b} \mathbf{Z}).$$

Obliczając każdy z składników powyższej sumy oddzielnie otrzymujemy:

$$\begin{aligned}
(Z' \mathbf{A} \mathbf{Z})^2 &= \sum_{i,j} a_{ij} a_{lm} Z_i Z_j Z_l Z_m \\
E(Z' \mathbf{A} \mathbf{Z})^2 &= \sum_{i,j,l,m} a_{ij} a_{lm} E(Z_i Z_j Z_l Z_m)
\end{aligned}$$

Zauważmy, że

$$E(Z_i Z_j Z_l Z_k) = \begin{cases} \boldsymbol{\mu}^{(4)}, & \text{gdy } i = j = k = l, \\ \sigma^4, & \text{gdy } i = j, l = k \text{ lub } i = l, j = k \text{ lub } i = k, j = l, \\ 0, & \text{w pozostałych przypadkach.} \end{cases}$$

Tak więc mamy

$$\begin{aligned} E(Z'AZ)^2 &= \sum_{i,j,l,m} a_{ij}a_{lm}E(Z_iZ_jZ_lZ_m) \\ &= \mu^{(4)} \sum_{i=1}^n a_{ii}^2 + \sigma^4 \left(\sum_{i \neq k} a_{ii}a_{kk} + \sum_{i \neq j} a_{ij}^2 + \sum_{i \neq j} a_{ij}a_{ji} \right) \end{aligned}$$

Ponieważ założyliśmy, że $\mathbf{A}$ jest symetryczna więc $a_{ij} = a_{ji}$, a stąd otrzymujemy

$$\begin{aligned} &\sum_{i \neq j} a_{ij}^2 + \sum_{i \neq j} a_{ij}a_{ji} \\ &= 2 \sum_{i \neq j} a_{ij}^2 = 2 \sum_{i,j} a_{ij}^2 - 2 \sum_{i=j} a_{ij}^2 \\ &= 2tr(A^2) - 2 \sum_{i=1}^n a_{ii}^2 \\ &= 2tr(A^2) - 2a' a \end{aligned}$$

oraz

$$\begin{aligned} \sum_{i \neq k} a_{ii}a_{kk} &= \sum_{i,k} a_{ii}a_{kk} - \sum_{i=k} a_{ii}a_{kk} \\ &= [tr(A)]^2 - \sum_{i=1}^n a_{ii}^2 = [tr(A)]^2 - a' a. \end{aligned}$$

Możemy więc napisać

$$E(\mathbf{Z}^T \mathbf{A} \mathbf{Z})^2 = (\boldsymbol{\mu}^{(4)} - 3\sigma^4) \mathbf{a}^T \mathbf{a} + \sigma^4 \left(2tr(\mathbf{A}^2) + [tr(\mathbf{A})]^2 \right).$$

Dla $\text{Var}(\mathbf{bZ})$ mamy równość

$$\text{Var}(\mathbf{bZ}) = \mathbf{b} \text{Var}(\mathbf{Z}) \mathbf{b}^T = \mathbf{b} \mathbf{b}^T \sigma^2 = (\boldsymbol{\mu}^T \mathbf{A})(\boldsymbol{\mu}^T \mathbf{A})^T \sigma^2 = \boldsymbol{\mu} \mathbf{A}^2 \boldsymbol{\mu} \sigma^2.$$

Aby wyznaczyć $\text{Cov}(\mathbf{Z}^T \mathbf{A} \mathbf{Z}, \mathbf{bZ})$ zauważmy, że $E\mathbf{Z} = 0$ i stąd

$$\begin{aligned}
& \text{Cov}(Z'AZ, bZ) \\
&= \text{Cov}\left(\sum_{i,j} a_{ij} Z_i Z_j, \sum_k b_k Z_k\right) \\
&= \sum_{i,j,k} a_{ij} b_k \text{Cov}(Z_i Z_j, Z_k) \\
&= \sum_{i,j,k} a_{ij} b_k E[(Z_i Z_j - E(Z_i Z_j)) Z_k] \\
&= \sum_{i,j,k} a_{ij} b_k [E(Z_i Z_j Z_k) - E(Z_i Z_j) E Z_k] \\
&= \sum_{i,j,k} a_{ij} b_k [E(Z_i Z_j Z_k)] \quad (\text{since } E Z_k = 0).
\end{aligned}$$

Ponieważ

$$E(Z_i Z_j Z_k) = \begin{cases} \mu^{(3)}, & \text{if } i = j = k; \\ 0, & \text{else.} \end{cases}$$

Stąd

$$\begin{aligned}
\text{Cov}(Z'AZ, bZ) &= \sum_{i=1}^n a_{ii} b_i \mu^{(3)} \\
&= \sum_{i=1}^n a_{ii} \mu' A_i \mu^{(3)} = \sum_{i=1}^n a_{ii} A_i' \mu \mu^{(3)} = a' A \mu \mu^{(3)}.
\end{aligned}$$

Łącząc powyższe równości otrzymujemy tezę twierdzenia.

1.4 Forma kwadratowa wektorów losowych o wielowymiarowym rozkładzie normalnym.

Rozkład formy kwadratowej $\mathbf{y}^T \mathbf{A} \mathbf{y}$ gdzie $\mathbf{y}$ jest wektorem losowym o wielowymiarowym rozkładzie normalnym odgrywa istotną rolę w analizie metody regresji liniowej. Podamy niektóre twierdzenia dotyczące rozkładu formy kwadratowej na podstawie średniej i kowariancji macierzy wektora normalnego $\mathbf{y}$, oraz postaci macierzy $\mathbf{A}$.

Twierdzenie. Niech $\mathbf{y}$ będzie n -wymiarowym wektorem normalnym $N(0, I)$. Niech $\mathbf{A}$ będzie idempotentną macierzą rzędu p . Forma kwadratowa $\mathbf{y}^T \mathbf{A} \mathbf{y}$ ma rozkład chi-kwadrat z p stopniami swobody.

Gdy w założeniach powyższego twierdzenia opuścimy założenie $E(\mathbf{y}) = 0$ rozkład formy kwadratowej $\mathbf{y}^T \mathbf{A} \mathbf{y}$ opisuje następujące twierdzenie

Twierdzenie. Niech $\mathbf{y}$ będzie n -wymiarowym wektorem normalnym $N(\mu, I)$. Niech $\mathbf{A}$ będzie idempotentną macierzą rzędu p . Forma kwadratowa $\mathbf{y}^T \mathbf{A} \mathbf{y}$ ma niecentralny rozkład chi-kwadrat z p stopniami swobody i parametrem nie-centralności $\lambda = \frac{1}{2} \mu^T \mathbf{A} \mu$.

W sytuacji ogólnej gdy wektor normalny $\mathbf{y}$ ma wielowymiarowy rozkład normalny z wektorem wartości oczekiwanych μ i macierzy kowariancji Σ rozkład formy kwadratowej tego wektora normalnego opisuje następujące twierdzenie.

Twierdzenie. Niech $\mathbf{y}$ będzie wektorem normalnym z wektorem wartości oczekiwanych $\boldsymbol{\mu}$ i macierzą kowariancji $\boldsymbol{\Sigma}$. Jeśli $\mathbf{A}\boldsymbol{\Sigma}$ jest idempotentną macierzą rzędu p . Forma kwadratowa $\mathbf{y}^T \mathbf{A} \mathbf{y}$ ma niecentralny rozkład chi-kwadrat z p stopniami swobody i parametrem nie-centralności $\lambda = \frac{1}{2} \boldsymbol{\mu}^T \mathbf{A} \boldsymbol{\mu}$.

Dowód. Po pierwsze, dla macierzy kowariancji $\boldsymbol{\Sigma}$ istnieje macierz ortogonalna $\boldsymbol{\Gamma}$ takie, że $\boldsymbol{\Sigma} = \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T$. Zdefiniujmy $\mathbf{Z} = \boldsymbol{\Gamma}^{-1}(\mathbf{y} - \boldsymbol{\mu})$. Tak określony wektor $\mathbf{Z}$ jest wektorem normalnym, dla którego $E(\mathbf{Z}) = 0$ i

$$\begin{aligned} \text{Cov}(\mathbf{Z}) = \text{Cov}(\boldsymbol{\Gamma}^{-1}(\mathbf{y} - \boldsymbol{\mu})) &= \boldsymbol{\Gamma}^{-1} \text{Cov}(\mathbf{y}) [\boldsymbol{\Gamma}]^{-1} = \boldsymbol{\Gamma}^{-1} \boldsymbol{\Sigma} [\boldsymbol{\Gamma}^T]^{-1} \\ &= \boldsymbol{\Gamma}^{-1} (\boldsymbol{\Gamma} \boldsymbol{\Gamma}^T) [\boldsymbol{\Gamma}^T]^{-1} = \mathbf{I}_p \end{aligned}$$

Tak więc $\mathbf{Z} \sim N(0, \mathbf{I})$. Co więcej, ponieważ $\mathbf{y} = \boldsymbol{\Gamma} \mathbf{Z} + \boldsymbol{\mu}$ mamy

$$\mathbf{y}^T \mathbf{A} \mathbf{y} = (\boldsymbol{\Gamma} \mathbf{Z} + \boldsymbol{\mu})^T \mathbf{A} (\boldsymbol{\Gamma} \mathbf{Z} + \boldsymbol{\mu}) = (\mathbf{Z}^T + [\boldsymbol{\Gamma}^T]^{-1})^T (\boldsymbol{\Gamma}^T \mathbf{A} \boldsymbol{\Gamma}) (\mathbf{Z} + [\boldsymbol{\Gamma}^T]^{-1} \boldsymbol{\mu}) = \mathbf{V}^T \mathbf{B} \mathbf{V},$$

gdzie $\mathbf{V} = \mathbf{Z}^T + [\boldsymbol{\Gamma}^T]^{-1} \boldsymbol{\mu} \sim N([\boldsymbol{\Gamma}^T]^{-1} \boldsymbol{\mu}, \mathbf{I}_p)$. Teraz musimy pokazać, że $\mathbf{B}$ jest macierzą idempotentną. Zauważmy, że

$$\mathbf{B}^2 = (\boldsymbol{\Gamma}^T \mathbf{A} \boldsymbol{\Gamma}) (\boldsymbol{\Gamma}^T \mathbf{A} \boldsymbol{\Gamma}) = \boldsymbol{\Gamma}^T (\mathbf{A} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T \mathbf{A}) \boldsymbol{\Gamma}.$$

Ponieważ $\mathbf{A} \boldsymbol{\Gamma}$ jest idempotentna, możemy napisać

$$\mathbf{A} \boldsymbol{\Gamma} = \mathbf{A} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T = \mathbf{A} \boldsymbol{\Gamma} \mathbf{A} \boldsymbol{\Gamma} = (\mathbf{A} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T \mathbf{A}) \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T = (\mathbf{A} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T \mathbf{A}) \boldsymbol{\Gamma}.$$

Zauważ, że $\boldsymbol{\Gamma}$ nie jest singularna więc

$$\mathbf{A} = \mathbf{A} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T \mathbf{A}.$$

Zatem

$$\mathbf{B}^2 = \boldsymbol{\Gamma}^T (\mathbf{A} \boldsymbol{\Gamma} \boldsymbol{\Gamma}^T \mathbf{A}) \boldsymbol{\Gamma} = \boldsymbol{\Gamma}^T \mathbf{A} \boldsymbol{\Gamma} = \mathbf{B},$$

więc $\mathbf{B}$ jest macierzą idempotentną. To pokazuje, że $\mathbf{V}^T \mathbf{B} \mathbf{V}$ ma rozkład chi-kwadrat z p stopniami swobody. Obliczmy wartość parametru niecentralności

$$\begin{aligned} \lambda &= \frac{1}{2} (\boldsymbol{\Gamma}'^{-1} \boldsymbol{\mu})' \mathbf{B} (\boldsymbol{\Gamma}'^{-1} \boldsymbol{\mu}) \\ &= \frac{1}{2} \boldsymbol{\mu}' \boldsymbol{\Gamma}'^{-1} (\boldsymbol{\Gamma}' \mathbf{A} \boldsymbol{\Gamma}) \boldsymbol{\Gamma}'^{-1} \boldsymbol{\mu} = \frac{1}{2} \boldsymbol{\mu}' \mathbf{A} \boldsymbol{\mu}. \end{aligned}$$

To kończy dowód.