

Hipotezy liniowe

Rozważany poprzednio problem testowania hipotezy o równości wartości oczekiwanych w k populacjach o rozkładach $N(m_i, \sigma^2)$, $i=1, \dots, k$ jest szczególnym przypadkiem pewnego ogólniejszego problemu testowania:

Model: $X = \varphi + \varepsilon$

gdzie $X \in R^n$ jest wektorem obserwacji

$\varphi \in \Omega \subset R^n$ jest wektorem średnich, o którym wiadomo, że należy do pewnej **właściwej podprzestrzeni liniowej Ω przestrzeni R^n** , tzn. $\Omega \subset R^n$ i $\dim(\Omega) = k < n$

$\varepsilon \in R^n$ jest losowym wektorem błędów o rozkładzie $\varepsilon \sim N_n(0, \sigma^2 I)$

Uwaga: $X \sim N_n(\varphi, \sigma^2 I)$

Interesuje nas problem testowania hipotezy

$H_0: \varphi \in \omega \subset \Omega$ przeciwko $H_1: \varphi \in \Omega - \omega$,

gdzie ω jest pewną **właściwą podprzestrzenią liniową** przestrzeni Ω ,

czyli $\omega \subset \Omega$ i $\dim(\omega) = k - r < \dim(\Omega)$.

Uwaga. Zbiory ω i Ω są **podprzestrzeniami liniowymi**, czyli nie mogą to być zbiory ograniczone np. kule w R^n .

Przykłady

Przykład 1. Model regresji liniowej

$$x_i = \beta_0 + \beta_1 a_i + \varepsilon_i, \quad i=1, \dots, n$$

$$\varepsilon_i \quad \text{i.i.d.} \quad N(0, \sigma^2)$$

Inaczej

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \beta_0 \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} + \beta_1 \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

Jeżeli nie wszystkie a_i są sobie równe, to $\Omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \right\}$ $\dim(\Omega) = 2$.

Jeżeli testujemy hipotezę o braku wpływu zmiennej a (tu niecosowej) na X , to problem sprowadza się

do testowania: $H_0: \beta_1 = 0$ wobec alternatywy $H_0: \beta_1 \neq 0$, czyli $\omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \right\}$, $\dim(\omega) = 1$.

Przykład 2. Porównywanie k średnich

Rozważmy k niezależnych prób prostych $X_{11}, \dots, X_{1n_1}$
 $\vdots$ pochodzących odpowiednio z
 $X_{k1}, \dots, X_{kn_k}$

rozkładów $N(m_1, \sigma^2), \dots, N(m_k, \sigma^2)$. Można więc napisać

$$X_{ij} = m_i + \varepsilon_{ij}, i=1, \dots, k, j=1, \dots, n_i,$$

przy czym ε_{ij} są niezależnymi zmiennymi losowymi o identycznych rozkładach normalnych $N(0, \sigma^2)$.

Testowana jest hipoteza $H_0: m_1 = \dots = m_k$ wobec alternatywy $H_1: \sim H_0$

W zapisie macierzowym

$$\begin{bmatrix} X_{11} \\ \vdots \\ X_{1n_1} \\ X_{21} \\ \vdots \\ X_{2k_2} \\ \vdots \\ X_{k1} \\ \vdots \\ X_{kn_k} \end{bmatrix} = m_1 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + m_2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + m_k \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2k_2} \\ \vdots \\ \varepsilon_{k1} \\ \vdots \\ \varepsilon_{kn_k} \end{bmatrix}$$

$$X = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 + \dots + m_k \mathbf{a}_k + \boldsymbol{\varepsilon}.$$

$$X = A\mathbf{m} + \boldsymbol{\varepsilon}.$$

W tym problemie

$$\Omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\} \quad \dim \Omega = k, \quad \omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\}, \quad \dim \omega = 1$$

Inna parametryzacja w problemie porównywania k średnich (można pominąć)

Niech $m = \sum_{i=1}^k \frac{n_i}{n} m_i$ będzie ważoną średnią grupowych wartości oczekiwanych. Oznaczając $\alpha_i = m_i - m$,

$i=1, \dots, k$ odchylenie wartości oczekiwanej w i -tej grupie od m otrzymujemy ograniczenie $\sum_{i=1}^k \frac{n_i}{n} \alpha_i = 0$.

Po przeparametryzowaniu model można zapisać w postaci

$$X_{ij} = m + \alpha_i + \varepsilon_{ij}, i=1, \dots, k, j=1, \dots, n_i, \text{ przy czym } \sum_{i=1}^k \frac{n_i}{n} \alpha_i = 0.$$

Jest to tzw. parametryzacja z sigma ograniczeniami

Parametr m reprezentuje średni poziom (poziom odniesienia) dla wszystkich obserwowanych zmiennych. Parametr α_i jest odchyleniem wartości oczekiwanej w i tej grupie od poziomu odniesienia.

Problem testowania równości k średnich sprowadza się w tym przypadku do testowania hipotezy

$$H_0: \alpha_1 = \dots = \alpha_k \text{ wobec alternatywy } H_1: \sim H_0 \text{ przy czym } \sum_{i=1}^k \frac{n_i}{n} \alpha_i = 0 \text{ (parametr } m \text{ nas nie interesuje).}$$

Zadanie Przy tej parametryzacji znaleźć przestrzenie Ω i ω .

$$\Omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \right\}, \dim \Omega = k, \omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix} \right\}, \dim \omega = 1$$

Uwaga: pierwszy z $k+1$ wektorów jest sumą pozostałych.

Jeszcze inna parametryzacja w problemie porównywania k średnich

Niech $m = \frac{1}{k} \sum_{i=1}^k m_i$ będzie średnią arytmetyczną grupowych wartości oczekiwanych. Oznaczając

$\alpha_i = m_i - m$, $i=1, \dots, k$ odchylenie wartości oczekiwanej w i -tej grupie od m otrzymujemy ograniczenie

$$\sum_{i=1}^k \alpha_i = 0.$$

Po przeparametryzowaniu model można zapisać w postaci

$$X_{ij} = m + \alpha_i + \varepsilon_{ij}, i=1, \dots, k, j=1, \dots, n_i, \text{ przy czym } \sum_{i=1}^k \alpha_i = 0.$$

Jest to tzw. parametryzacja z sigma ograniczeniami

Parametr m reprezentuje średni poziom (poziom odniesienia) dla wszystkich obserwowanych zmiennych. Parametr α_i jest odchyleniem wartości oczekiwanej w i tej grupie od poziomu odniesienia.

Problem testowania równości k średnich sprowadza się w tym przypadku do testowania hipotezy

$$H_0: \alpha_1 = \dots = \alpha_k \text{ wobec alternatywy } H_1: \sim H_0 \text{ przy czym } \sum_{i=1}^k \alpha_i = 0 \text{ (parametr } m \text{ nas nie interesuje)}$$

Zadanie Przy tej parametryzacji znaleźć przestrzenie Ω i ω .

$$\Omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\} \quad \dim \Omega = k, \quad \omega = \text{span} \left\{ \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} \right\}, \quad \dim \omega = 1$$

Uwaga: pierwszy z $k+1$ wektorów jest sumą pozostałych.

I jeszcze inna parametryzacja. Przypuśćmy, że jedna z grup (dla ustalenia uwagi pierwsza) jest wyróżniona np. jest to **grupa kontrolna**. Przyjmując parametry

$$\beta_1 = m_1$$

$$\beta_i = m_i - m_1, \quad i=2, \dots, k$$

$$\begin{bmatrix} X_{11} \\ \vdots \\ X_{1n_1} \\ X_{21} \\ \vdots \\ X_{2k_2} \\ \vdots \\ X_{k1} \\ \vdots \\ X_{kn_k} \end{bmatrix} = \beta_1 \begin{bmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \beta_2 \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + \beta_k \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{bmatrix} + \begin{bmatrix} \varepsilon_{11} \\ \vdots \\ \varepsilon_{1n_1} \\ \varepsilon_{21} \\ \vdots \\ \varepsilon_{2k_2} \\ \vdots \\ \varepsilon_{k1} \\ \vdots \\ \varepsilon_{kn_k} \end{bmatrix}$$

Testowanie hipotezy o równości średnich sprowadza się do testowania hipotezy $H_0: \beta_i = 0, \quad i=2, \dots, k$.

Parametr $\beta_1 = m_1$ wyznaczający „poziom odniesienia” nas nie interesuje.

Parametryzacja z grupą kontrolną jest przykładem **parametryzacji oszczędnej**.

Zadanie 1. Linie lotnicze zainteresowane są porównaniem przydatności 2 wyświetlaczy dla kontrolerów ruchu lotniczego. Każdy wyświetlacz był sprawdzany w 3 symulowanych warunkach zagrożenia. Dwunastu wysoko wykwalifikowanych kontrolerów z podobnym doświadczeniem zawodowym zostało wybranych do badań. Kontrolerzy zostali losowo przypisani do klatek (wyświetlacz, warunki zagrożenia) po dwóch do każdej klatki. Mierzono czas w sekundach potrzebny do opanowania sytuacji awaryjnej.

Model strukturalny

$$X_{ijk} = m + \alpha_i + \beta_j + \varepsilon_{ijk} \quad i=1,2; j=1,2,3; k=1,2$$

$$\sum_i \alpha_i = 0,$$

$$\sum_j \beta_j = 0$$

$$\varepsilon_{ijk} \sim iid \quad N(0, \sigma^2)$$

Uwaga Statistica stosuje następującą parametryzację

Z sigma ograniczeń otrzymujemy $\alpha_2 = -\alpha_1$ i $\beta_3 = -\beta_1 - \beta_2$. Stąd

$$\begin{matrix} A \\ B \end{matrix} \left\{ \begin{matrix} \text{I} \\ \text{II} \\ \text{III} \\ \text{I} \\ \text{II} \\ \text{III} \end{matrix} \right\} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} m \\ \alpha_1 \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \varepsilon$$

Można także dla rozważanego przypadku zbudować ogólniejszy model uwzględniający **interakcję Wyświetlacz $\times$ Zagrożenie**, a mianowicie

$$X_{ijk} = m + \alpha_i + \beta_j + \gamma_{ij} + \varepsilon_{ijk} \quad i=1,2; j=1,2,3, \text{ przy czym}$$

$$\sum_i \alpha_i = 0, \sum_j \beta_j = 0, \sum_i \gamma_{ij} = 0 \text{ dla każdego } j, \sum_j \gamma_{ij} = 0 \text{ dla każdego } i.$$

Parametryzacja Statistica

$$\begin{matrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \end{matrix}$$

$$\begin{matrix} A \\ B \end{matrix} \left\{ \begin{matrix} \text{I} \\ \text{II} \\ \text{III} \\ \text{I} \\ \text{II} \\ \text{III} \end{matrix} \right\} = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 0 & -1 & 0 \\ 1 & -1 & 0 & 1 & 0 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \begin{bmatrix} m \\ \alpha_1 \\ \beta_1 \\ \beta_2 \\ \gamma_{11} \\ \gamma_{12} \end{bmatrix} + \begin{bmatrix} \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \end{bmatrix} \varepsilon.$$

Po wyborze parametryzacji rozważany model liniowy można zapisać w postaci gaussowskiego modelu liniowego

$$\mathbf{X} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

$\mathbf{A}$ jest tu nielosową macierzą wymiaru (n, k) zwaną **macierzą planu eksperymentu** o kolumnach $\mathbf{a}_i$, $i=1, \dots, k$, przy czym dla uproszczenia założymy, że te kolumny są liniowo niezależne tzn. $\text{rz}(\mathbf{A}) = k$.

Wektor $\boldsymbol{\varphi} = E(\mathbf{X}) \in \Omega \subset R^n$ przy czym $\Omega = \text{span}\{\mathbf{a}_i; i=1, \dots, k\} = \text{Im}(\mathbf{A})$

Testować będziemy hipotezę liniową

$$H_0: \mathbf{H}\boldsymbol{\beta} = \mathbf{0}, \quad \mathbf{H} \text{ jest macierzą } (r, k) \quad r < k \text{ pełnego rzędu tzn. } \text{rz}(\mathbf{H}) = r,$$

która specyfikuje podprzestrzeń $\omega = \text{Im}(\mathbf{A}_{|_{\text{Ker}(\mathbf{H})}})$ ($\dim(\omega) = k - r$) będącą obrazem obciążenia przekształcenia $\mathbf{A}$ do jądra przekształcenia $\mathbf{H}$, przeciwko $H_1: \boldsymbol{\varphi} \in \Omega - \omega$.

Test hipotezy liniowej oparty na ilorazie wiarygodności (wariant I)

Na początku skonstruujemy estymator największej wiarygodności parametrów (φ, σ^2) w modelu liniowym

$$\mathbf{X} = \varphi + \boldsymbol{\varepsilon},$$

gdzie $\varphi \in M \subset R^n$ i $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ a M jest podprzestrzenią liniową przestrzeni R^n .

Funkcja wiarygodności jest postaci

$$L(\varphi, \sigma^2; \mathbf{X}) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \|\mathbf{X} - \varphi\|^2}$$

a jej logarytm

$$l(\varphi, \sigma^2; \mathbf{X}) = \ln L = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|\mathbf{X} - \varphi\|^2.$$

Oznaczmy przez P_M projekcję ortogonalną na podprzestrzeń $M \subset R^n$. Tym samym symbolem będziemy oznaczać reprezentację macierzową tej projekcji (rzutowania). Macierz P_M jest idempotentna ($P_M^2 = P_M$ - to gwarantuje, że jest to macierz rzutowania) i symetryczna ($P_M^T = P_M$ - ten warunek zapewnia, że projekcja jest ortogonalna). Stąd

$$l(\varphi, \sigma^2; \mathbf{X}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|\mathbf{X} - P_M \mathbf{X} + P_M \mathbf{X} - \varphi\|^2, \text{ ale}$$

$$\|\mathbf{X} - P_M \mathbf{X} + P_M \mathbf{X} - \varphi\|^2 = \langle \mathbf{X} - P_M \mathbf{X} + P_M \mathbf{X} - \varphi, \mathbf{X} - P_M \mathbf{X} + P_M \mathbf{X} - \varphi \rangle = \|\mathbf{X} - P_M \mathbf{X}\|^2 + \|P_M \mathbf{X} - \varphi\|^2$$

bo $\langle \mathbf{X} - P_M \mathbf{X}, P_M \mathbf{X} - \varphi \rangle = 0$ gdyż $\mathbf{X} - P_M \mathbf{X} \in M^\perp$ a $P_M \mathbf{X} - \varphi \in M$.

$$\text{Wobec tego } l(\varphi, \sigma^2; \mathbf{X}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|\mathbf{X} - P_M \mathbf{X}\|^2 - \frac{1}{2\sigma^2} \|P_M \mathbf{X} - \varphi\|^2$$

W powyższej funkcji wektor φ występuje tylko w jednym miejscu, więc łatwo widać, że $\forall \varphi$

$$l(\varphi, \sigma^2; \mathbf{X}) \leq -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|\mathbf{X} - P_M \mathbf{X}\|^2 \text{ i równość występuje tylko dla } \varphi = P_M \mathbf{X} \text{ więc}$$

$$\hat{\varphi} = P_M \mathbf{X} = ENW[\varphi]. \hat{\sigma}^2 = \arg \max l(\hat{\varphi}, \sigma^2; \mathbf{X}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \|\mathbf{X} - P_M \mathbf{X}\|^2 \text{ wyznaczamy}$$

szukając ekstremum funkcji różniczkowalnej jednej zmiennej. Oczywiście $\hat{\sigma}^2 = \frac{1}{n} \|\mathbf{X} - P_M \mathbf{X}\|^2$.

$$\text{Wykażemy, że rzeczywiście } (\hat{\varphi}, \hat{\sigma}^2) = (\mathbf{X} - P_M \mathbf{X}, \frac{1}{n} \|\mathbf{X} - P_M \mathbf{X}\|^2) = ENW(\varphi, \sigma^2).$$

$$\begin{aligned} l(\hat{\varphi}, \hat{\sigma}^2; \mathbf{X}) - l(\varphi, \sigma^2; \mathbf{X}) &= -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \hat{\sigma}^2 - \frac{1}{2\hat{\sigma}^2} \|\mathbf{X} - P_M \mathbf{X}\|^2 \\ &\quad - (-\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \|\mathbf{X} - P_M \mathbf{X}\|^2 - \frac{1}{2\sigma^2} \|P_M \mathbf{X} - \varphi\|^2) \\ &= \frac{n}{2} (-\ln \frac{\hat{\sigma}^2}{\sigma^2} - 1 + \frac{\hat{\sigma}^2}{\sigma^2}) + \frac{1}{2\sigma^2} \|P_M \mathbf{X} - \varphi\|^2 \end{aligned}$$

Z nierówności $\ln x \leq x - 1$, która jest konsekwencją wklęsłości funkcji $\ln x$ otrzymujemy nierówność $-\ln x - 1 + x \geq 0$, która jest prawdziwa dla wszystkich $x > 0$ a równość zachodzi jedynie dla $x = 1$.

Widać więc że $l(\hat{\varphi}, \hat{\sigma}^2; \mathbf{X}) - l(\varphi, \sigma^2; \mathbf{X})$ jest sumą dwóch nieujemnych składników

$\frac{n}{2}(-\ln \frac{\hat{\sigma}^2}{\sigma^2} - 1 + \frac{\hat{\sigma}^2}{\sigma^2})$ i $\frac{1}{2\sigma^2} \|P_M \mathbf{X} - \varphi\|^2$ więc $l(\hat{\varphi}, \hat{\sigma}^2; \mathbf{X}) - l(\varphi, \sigma^2; \mathbf{X}) \geq 0$, a wartość 0 różnica ta osiąga, gdy oba składniki równocześnie się zerują, czyli $\varphi = P_M \mathbf{X}$ a $\frac{\hat{\sigma}^2}{\sigma^2} = 1$.

Konstrukcja testu hipotezy liniowej oparta na ilorazie wiarygodności

$$C = \{\mathbf{X} : \lambda(\mathbf{X}) = \frac{L(\hat{\varphi}, \hat{\sigma}^2; \mathbf{X})}{L(\varphi, \sigma^2; \mathbf{X})} > c_1\} = \{\mathbf{X} : \left(\frac{\hat{\sigma}^2}{\sigma^2}\right)^{\frac{n}{2}} > c_1\} = \{\mathbf{X} : \frac{\hat{\sigma}^2}{\sigma^2} > c_2\} = \{\mathbf{X} : \frac{n-k}{r} \frac{\hat{\sigma}^2 - \sigma^2}{\sigma^2} > c\}$$

Oznaczając

$$RSS_{\omega} = \|\mathbf{X} - P_{\omega} \mathbf{X}\|^2 = \|\dot{\mathbf{e}}\|^2 \quad \text{reszkowa suma kwadratów przy prawdziwości } H_0$$

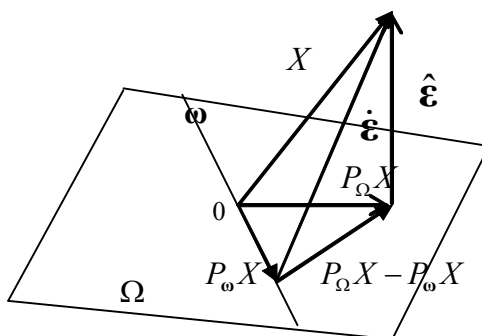
(w problemie k -prób całkowita suma kwadratów SS_{total})

$$RSS_{\Omega} = \|\mathbf{X} - P_{\Omega} \mathbf{X}\|^2 = \|\hat{\mathbf{e}}\|^2 \quad \text{reszkowa suma kwadratów w modelu bez ograniczeń}$$

(w problemie k -prób suma kwadratów wewnątrz grup SS_{within})

otrzymujemy
$$C = \{\mathbf{X} : \frac{n-k}{r} \frac{RSS_{\omega} - RSS_{\Omega}}{RSS_{\Omega}} > c\}.$$

Interpretacja geometryczna



$$RSS_{\omega} - RSS_{\Omega} = \|\dot{\mathbf{e}}\|^2 - \|\hat{\mathbf{e}}\|^2 = \|P_{\Omega} X - P_{\omega} X\|^2$$

pomiędzy

(w problemie k -prób suma kwadratów grupami $SS_{between}$)

Postać kanoniczna

Niech $\Omega \subset R^n$ $\dim(\Omega) = k$

$\omega \subset \Omega$ $\dim(\omega) = k-r$

Wprowadźmy w R^n nową bazę ortonormalną $(\mathbf{e}'_1, \dots, \mathbf{e}'_n)$ tak aby

$(\mathbf{e}'_1, \dots, \mathbf{e}'_{k-r})$ baza w ω

$(\mathbf{e}'_1, \dots, \mathbf{e}'_{k-r}, \mathbf{e}'_{k-r+1}, \dots, \mathbf{e}'_k)$ baza w Ω

$(\mathbf{e}'_1, \dots, \mathbf{e}'_{k-r}, \mathbf{e}'_{k-r+1}, \dots, \mathbf{e}'_k, \mathbf{e}'_{k+1}, \dots, \mathbf{e}'_n)$ baza w R^n

Niech $\mathbf{P}$ będzie macierzą przejścia od **starej bazy kanonicznej** do nowej bazy $(\mathbf{e}'_1, \dots, \mathbf{e}'_n)$

Oznacza to, że $\mathbf{e}'_j = \sum_{i=1}^n P_{ij} \mathbf{e}_i$. Macierz $\mathbf{P}$ jest oczywiście macierzą ortogonalną. Związek pomiędzy

współrzednymi $\mathbf{x}$ wektora w starej bazie i jego współrzednymi $\mathbf{x}'$ w nowej bazie jest postaci $\mathbf{x} = \mathbf{P}\mathbf{x}'$ lub równoważnie $\mathbf{x}' = \mathbf{P}^T \mathbf{x}$.

Ponieważ w bazie kanonicznej możemy utożsamiać wektor z jego zestawem współrzednych w bazie kanonicznej kolumny macierzy $\mathbf{P}$ (wiersze $\mathbf{P}^T$) są zestawione z wektorów nowej bazy $(\mathbf{e}'_1, \dots, \mathbf{e}'_n)$

Każdy wektor leżący w przestrzeni Ω (m.in. φ i $\hat{\varphi} = P_\Omega X$) rozpinanej także przez pierwszych

k wektorów nowej bazy $(\mathbf{e}'_1, \dots, \mathbf{e}'_{k-r}, \mathbf{e}'_{k-r+1}, \dots, \mathbf{e}'_k)$ może być zapisany w postaci $\sum_{i=1}^k \gamma_i \mathbf{e}'_i$. Oczywiście

wektor $\omega \ni \hat{\varphi} = P_\omega X = \sum_{i=1}^{k-r} \gamma_i \mathbf{e}'_i$ jest liniową kombinacją pierwszych $k-r$ wektorów nowej bazy.

Przechodząc do nowej bazy zapiszemy nasz model w nowych współrzednych

$$\mathbf{P}^T \mathbf{X} = \mathbf{P}^T \varphi + \mathbf{P}^T \varepsilon$$

Oznaczając $\mathbf{Y} = \mathbf{P}^T \mathbf{X}$, $\boldsymbol{\gamma} = \mathbf{P}^T \varphi$, $\boldsymbol{\eta} = \mathbf{P}^T \varepsilon$ otrzymujemy równoważny model w postaci kanonicznej

$$\begin{bmatrix} Y_1 \\ \vdots \\ Y_{k-r} \\ Y_{k-r+1} \\ \vdots \\ Y_k \\ Y_{k+1} \\ \vdots \\ Y_n \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_{k-r} \\ \gamma_{k-r+1} \\ \vdots \\ \gamma_k \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_{k-r} \\ \eta_{k-r+1} \\ \vdots \\ \eta_k \\ \eta_{k+1} \\ \vdots \\ \eta_n \end{bmatrix} \Leftrightarrow \mathbf{Y} = \boldsymbol{\gamma} + \boldsymbol{\eta}$$

lub równoważnie

$$\begin{bmatrix} \mathbf{Y} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \mathbf{I}_{k-r} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_r \end{bmatrix} \\ \mathbf{O} \end{bmatrix} \begin{bmatrix} \gamma_1 \\ \vdots \\ \gamma_k \end{bmatrix} + \begin{bmatrix} \boldsymbol{\eta} \end{bmatrix} \quad \text{przy czym} \quad \boldsymbol{\eta} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

Dla modelu w postaci kanonicznej hipoteza H_0 przybiera postać

$$H_0 : \gamma_{k-r+1} = 0, \dots, \gamma_k = 0$$

Ponieważ przekształcenie ortogonalne zachowuje normę więc

$$RSS_{\Omega} = \|\hat{\epsilon}\|^2 = \|\mathbf{X} - P_{\Omega}\mathbf{X}\|^2 = \|\mathbf{Y} - P_{\Omega}\mathbf{Y}\|^2 = \sum_{i=k+1}^n Y_i^2 = Y_{k+1}^2 + \dots + Y_n^2$$

reszkowa suma kwadratów (wewnątrzgrupowa) w modelu bezwarunkowym (SS_{within}).

$$RSS_{\omega} = \|\hat{\epsilon}\|^2 = \|\mathbf{X} - P_{\omega}\mathbf{X}\|^2 = \|\mathbf{Y} - P_{\omega}\mathbf{Y}\|^2 = \sum_{i=k-r+1}^n Y_i^2 = Y_{k-r+1}^2 + \dots + Y_k^2 + Y_{k+1}^2 + \dots + Y_n^2 \quad (SS_{total})$$

reszkowa suma kwadratów w modelu warunkowym (całkowita suma kwadratów)

$$RSS_{\omega} - RSS_{\Omega} = \sum_{i=k-r+1}^k Y_i^2 = Y_{k-r+1}^2 + \dots + Y_k^2 \quad \text{suma kwadratów pomiędzy grupami } (SS_{between})$$

Ponieważ $\mathbf{Y} \sim N(\boldsymbol{\gamma}, \sigma^2 \mathbf{I})$ więc otrzymujemy wniosek

- Zmienne losowe $R_1, \dots, R_r$ oraz R_0 są niezależne oraz
- $\frac{RSS_{\omega} - RSS_{\Omega}}{\sigma^2}$ ma rozkład $\chi^2_{r, \delta}$ przy

$$\text{czyli } \delta^2 = \frac{1}{\sigma^2} \sum_{i=k-r+1}^k \gamma_i^2 = \frac{1}{\sigma^2} \|\boldsymbol{\gamma} - P_{\omega}\boldsymbol{\gamma}\|^2 = \frac{1}{\sigma^2} \|\boldsymbol{\varphi} - P_{\omega}\boldsymbol{\varphi}\|^2$$

- $\frac{RSS_{\Omega}}{\sigma^2}$ ma rozkład χ^2_{n-k}

Stąd statystyka testowa $F = \frac{n-k}{r} \frac{RSS_{\omega} - RSS_{\Omega}}{RSS_{\Omega}}$ ma niecentralny rozkład Snedecora-Fishera $F_{r, n-k, \delta}$ a

przy prawdziwości H_0 rozkład centralny $F_{r, n-k}$.

Podsumowaniem jest następująca tabelka ANOVA

Źródło zmienności	Suma kwadratów SS	Stopnie swobody df	Średni kwadrat MS	Iloraz F
Pomiędzy grupami Between	$RSS_{\omega} - RSS_{\Omega}$	r	$\frac{1}{r} (RSS_{\omega} - RSS_{\Omega})$	$F = \frac{n-k}{r} \frac{RSS_{\omega} - RSS_{\Omega}}{RSS_{\Omega}}$
Wewnątrz grup Within	RSS_{Ω}	$n-k$	$\frac{1}{n-k} RSS_{\Omega}$	
Ogółem Total	RSS_{ω}	$n-k+r$		

Test hipotezy liniowej oparty na ilorazie wiarygodności (wariant II)

Rozważmy gaussowski model liniowy $(\mathbf{X}, \mathbf{A}\boldsymbol{\beta}, \sigma^2 \mathbf{I})$ z oszczędną parametryzacją tzn.

$$\mathbf{X} = \mathbf{A}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

$$\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$$

przy czym macierz $\mathbf{A}$ wymiaru (n, k) planu eksperymentu jest tu nielosową macierzą o liniowo niezależnych kolumnach $\mathbf{a}_i$, $i=1, \dots, k$ tzn. $\text{rz}(\mathbf{A})=k$.

Rodzina rozkładów na przestrzeni prób R^n jest parametryzowana wektorowym parametrem $\theta = (\boldsymbol{\beta}, \sigma^2)$ ze zbioru $\Theta = \{\theta = (\boldsymbol{\beta}, \sigma^2) \in R^k \times R_+\}$. Interesować nas będzie hipoteza liniowa $H_0: \mathbf{H}\boldsymbol{\beta} = \mathbf{0}$, dotycząca parametru $\boldsymbol{\beta}$. Parametr σ^2 pełni rolę parametru zakłócającego. Hipoteza zerowa specyfikuje podzbiór $\Theta_0 = \{\theta = (\boldsymbol{\beta}, \sigma^2) \in R^k \times R_+ : \mathbf{H}\boldsymbol{\beta} = \mathbf{0}\} \subset \Theta$ przestrzeni parametrów.

Z punktu widzenia interesującego nas parametru $\boldsymbol{\beta}$ zbiór Θ specyfikuje $\boldsymbol{\beta} \in R^k$, co odpowiada specyfikacji k wymiarowej podprzestrzeni $\Omega = \text{span}\{\mathbf{a}_i; i=1, \dots, k\} = \text{Im}(\mathbf{A})$ w której leży wektor $E(\mathbf{X})$. Z kolei zbiór Θ_0 specyfikuje kolejną podprzestrzeń $\omega = \text{Im}(\mathbf{A}|_{\text{Ker} \mathbf{H}}) \subset \Omega$

Przy założeniu $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \sigma^2 \mathbf{I})$ funkcja wiarygodności ma postać

$$L(\boldsymbol{\beta}, \sigma^2; \mathbf{X}) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}(\mathbf{X} - \mathbf{A}\boldsymbol{\beta})^T(\mathbf{X} - \mathbf{A}\boldsymbol{\beta})}$$

a jej logarytm

$$l(\boldsymbol{\beta}, \sigma^2; \mathbf{X}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\mathbf{X} - \mathbf{A}\boldsymbol{\beta})^T(\mathbf{X} - \mathbf{A}\boldsymbol{\beta}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} S(\boldsymbol{\beta}), \text{ gdzie}$$

$S(\boldsymbol{\beta}) = (\mathbf{X} - \mathbf{A}\boldsymbol{\beta})^T(\mathbf{X} - \mathbf{A}\boldsymbol{\beta})$. Wiadomo, że (bezwarunkowymi) estymatorami NW $(\boldsymbol{\beta}, \sigma^2)$ są odpowiednio $\hat{\boldsymbol{\beta}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{X}$ i $\hat{\sigma}^2 = \frac{1}{n} \langle \mathbf{X} - \mathbf{A}\hat{\boldsymbol{\beta}}, \mathbf{X} - \mathbf{A}\hat{\boldsymbol{\beta}} \rangle$.

$$\text{Stąd } L(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2; \mathbf{X}) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}.$$

Podobnie wyznaczamy warunkowe estymatory NW

$$(\dot{\boldsymbol{\beta}}, \dot{\sigma}^2) = \arg \max \left\{ -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} S(\boldsymbol{\beta}) \right\} \text{ przy warunku } \mathbf{H}\boldsymbol{\beta} = \mathbf{0}.$$

Funkcja Lagrange'a dla tego problemu po przeskalowaniu mnożników może być zapisana w postaci

$$\Phi(\boldsymbol{\beta}, \sigma^2; \boldsymbol{\lambda}) = -\frac{n}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} \{ \langle \mathbf{X} - \mathbf{A}\boldsymbol{\beta}, \mathbf{X} - \mathbf{A}\boldsymbol{\beta} \rangle + 2\langle \boldsymbol{\lambda}, \mathbf{H}\boldsymbol{\beta} \rangle \}$$

Problem wyznaczenia sprowadza się do wyznaczenia $\dot{\boldsymbol{\beta}} = \arg \min \langle \mathbf{X} - \mathbf{A}\boldsymbol{\beta}, \mathbf{X} - \mathbf{A}\boldsymbol{\beta} \rangle$ przy warunku

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{0} \text{ a następnie (jak poprzednio) } \dot{\sigma}^2 = \frac{1}{n} \langle \mathbf{X} - \mathbf{A}\dot{\boldsymbol{\beta}}, \mathbf{X} - \mathbf{A}\dot{\boldsymbol{\beta}} \rangle$$

Funkcja Lagrange'a dla tego ostatniego problemu jest postaci

$$\Phi(\boldsymbol{\beta}; \boldsymbol{\lambda}) = \langle \mathbf{X} - \mathbf{A}\boldsymbol{\beta}, \mathbf{X} - \mathbf{A}\boldsymbol{\beta} \rangle + 2\langle \boldsymbol{\lambda}, \mathbf{H}\boldsymbol{\beta} \rangle = \langle \mathbf{X}, \mathbf{X} \rangle - 2\langle \mathbf{A}^T \mathbf{X}, \boldsymbol{\beta} \rangle + \langle \mathbf{A}^T \mathbf{A} \boldsymbol{\beta}, \boldsymbol{\beta} \rangle + 2\langle \mathbf{H}^T \boldsymbol{\lambda}, \boldsymbol{\beta} \rangle$$

Wyznaczając pochodną Frecheta mamy dekompozycję

$$\Phi(\beta + \Delta\beta; \lambda + \Delta\lambda) - \Phi(\beta; \lambda) = 2 \left[\mathbf{A}^T \mathbf{A} \beta + \mathbf{H}^T \lambda - \mathbf{A}^T \mathbf{X}, \mathbf{H} \beta \right] \begin{bmatrix} \Delta\beta \\ \Delta\lambda \end{bmatrix} + \text{reszta}$$

Warunek znikania pochodnej $2 \left[\mathbf{A}^T \mathbf{A} \beta + \mathbf{H}^T \lambda - \mathbf{A}^T \mathbf{X}, \mathbf{H} \beta \right] = [\mathbf{0}, \mathbf{0}]$ prowadzi do układu równań

$$\begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{H}^T \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \beta \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T \mathbf{X} \\ \mathbf{0} \end{bmatrix}$$

Można pokazać (Silvey str. 235), przez konstrukcję macierzy odwrotnej do macierzy w postaci blokowej, że jeżeli $\mathbf{A}^T \mathbf{A}$ jest nieosobliwa (równoważnie $\text{rz}(\mathbf{A})=k$) i macierz $\mathbf{H}$ o wymiarach (k, m)

$m < k$ (jest pełnego rzędu tzn. $\text{rz}(\mathbf{H})=m$), to macierz $\begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{H}^T \\ \mathbf{H} & \mathbf{0} \end{bmatrix}$ jest nieosobliwa i

$$(*) \begin{bmatrix} \hat{\beta} \\ \hat{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^T \mathbf{A} & \mathbf{H}^T \\ \mathbf{H} & \mathbf{0} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{A}^T \mathbf{X} \\ \mathbf{0} \end{bmatrix} \text{ i w konsekwencji } L(\hat{\beta}, \hat{\sigma}^2; \mathbf{X}) = (2\pi\hat{\sigma}^2)^{-\frac{n}{2}} e^{-\frac{n}{2}}$$

Uzupełnienie. Konstrukcja macierzy odwrotnej

Z definicji $\begin{bmatrix} \Sigma & \mathbf{H}^T \\ \mathbf{H} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{P} & \mathbf{Q}^T \\ \mathbf{Q} & \mathbf{R} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$ otrzymujemy układ równań

1. $\Sigma \mathbf{P} + \mathbf{H}^T \mathbf{Q} = \mathbf{I}$
2. $\Sigma \mathbf{Q}^T + \mathbf{H}^T \mathbf{R} = \mathbf{0}$
3. $\mathbf{H} \mathbf{P} = \mathbf{0}$
4. $\mathbf{H} \mathbf{Q}^T = \mathbf{I}$

Z (1) mamy $\mathbf{P} + \Sigma^{-1} \mathbf{H}^T \mathbf{Q} = \Sigma^{-1}$. Mnożąc przez $\mathbf{H}$ i uwzględniając (3) mamy $\mathbf{H} \Sigma^{-1} \mathbf{H}^T \mathbf{Q} = \mathbf{H} \Sigma^{-1}$. Ponieważ $\mathbf{H}$ jest pełnego rzędu macierz $\mathbf{H} \Sigma^{-1} \mathbf{H}^T$ jest nieosobliwa więc $\mathbf{Q} = (\mathbf{H} \Sigma^{-1} \mathbf{H}^T)^{-1} \mathbf{H} \Sigma^{-1}$ i w konsekwencji $\mathbf{P} = \Sigma^{-1} - \Sigma^{-1} \mathbf{H}^T (\mathbf{H} \Sigma^{-1} \mathbf{H}^T)^{-1} \mathbf{H} \Sigma^{-1}$. Podobnie z (2) $\mathbf{Q}^T + \Sigma^{-1} \mathbf{H}^T \mathbf{R} = \mathbf{0}$, mnożąc przez $\mathbf{H}$ po uwzględnieniu (4) $\mathbf{I} + \mathbf{H} \Sigma^{-1} \mathbf{H}^T \mathbf{R} = \mathbf{0}$, stąd $\mathbf{R} = -(\mathbf{H} \Sigma^{-1} \mathbf{H}^T)^{-1}$.

Konstrukcja testu hipotezy liniowej

$$C = \{ \mathbf{X} : \lambda(\mathbf{X}) = \frac{L(\hat{\beta}, \hat{\sigma}^2; \mathbf{X})}{L(\hat{\beta}, \hat{\sigma}^2; \mathbf{X})} > c_1 \} = \{ \mathbf{X} : \left(\frac{\hat{\sigma}^2}{\hat{\sigma}^2} \right)^{\frac{n}{2}} > c_1 \} = \{ \mathbf{X} : \frac{\hat{\sigma}^2}{\hat{\sigma}^2} > c_2 \} = \{ \mathbf{X} : \frac{n-k}{r} \frac{\hat{\sigma}^2 - \hat{\sigma}^2}{\hat{\sigma}^2} > c \}$$

Oznaczając

$$RSS_{\omega} = \langle \mathbf{X} - \mathbf{A} \hat{\beta}, \mathbf{X} - \mathbf{A} \hat{\beta} \rangle = \|\mathbf{X} - \mathbf{A} \hat{\beta}\|^2 = \|\hat{\epsilon}\|^2 \quad \text{resztkowa suma kwadratów przy prawdziwości } H_0$$

(w problemie k -prób całkowita suma kwadratów)

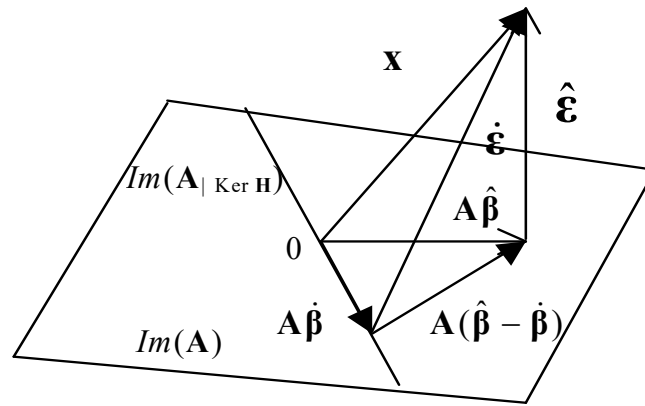
$$RSS_{\Omega} = \langle \mathbf{X} - \mathbf{A} \hat{\beta}, \mathbf{X} - \mathbf{A} \hat{\beta} \rangle = \|\mathbf{X} - \mathbf{A} \hat{\beta}\|^2 = \|\hat{\epsilon}\|^2 \quad \text{resztkowa suma kwadratów w modelu bez ograniczeń}$$

(w problemie k -prób suma kwadratów wewnątrz grup)

otrzymujemy

$$C = \left\{ \mathbf{X} : \frac{n-k}{r} \frac{RSS_{\omega} - RSS_{\Omega}}{RSS_{\Omega}} > c \right\}.$$

Interpretacja geometryczna



$$RSS_{\omega} - RSS_{\Omega} = \|\dot{\boldsymbol{\varepsilon}}\|^2 - \|\hat{\boldsymbol{\varepsilon}}\|^2 = \|\mathbf{A}(\dot{\boldsymbol{\beta}} - \hat{\boldsymbol{\beta}})\|^2 \quad (\text{w problemie } k\text{-prób suma kwadratów pomiędzy grupami})$$

Uwaga: Łącząc oba podejścia jesteśmy w stanie wypisać jawnie reprezentację macierzową projekcji

Jeśli $\Omega = \text{Im } \mathbf{A}$ a $\omega = \text{Im } \mathbf{A}_{|_{\text{Ker } \mathbf{H}}}$ i $\mathbf{A}$ jest pełnego rzędu (ma liniowo niezależne kolumny), to

$$P_{\Omega} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

$$P_{\omega} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T - \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{H}^T (\mathbf{H}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{H}^T)^{-1} \mathbf{H}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T.$$

$$\hat{\phi} = P_{\Omega} \mathbf{X} = \mathbf{A} \hat{\boldsymbol{\beta}} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{X} \Rightarrow P_{\Omega} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

$$\dot{\phi} = P_{\omega} \mathbf{X} = \mathbf{A} \dot{\boldsymbol{\beta}} \quad (\text{podstawiając za } \dot{\boldsymbol{\beta}} \text{ rozwiązanie układu (*) otrzymujemy postać } P_{\omega})$$