Grzegorz Mika

Testowanie regresji liniowej przeciwko wypukłej

6 kwietnia 2016

Spis treści

Literatura

[1] Fraser D.A.S., Massam H., *A Mixed Primal- Dual Bases Algorithm for Regression under Inequality Constraints. Application to Concave regression*, Scand J. Statist, **16** 65-74, 1989

[2] Meyer Mary C., *A test for linear vs convex regression function using shape- restricted regression*, Stanford University, Technical Report No. 2001-20, sierpień 2001

1. Wstęp

Rozważmy pewien zestaw danych $\{(x_i, y_i)\}_{i=1,2,\ldots,n}$ i spróbujmy dopasować pewną funkcję f do danych według modelu

$$y_i = f(x_i) + \varepsilon_i$$

gdzie zakładamy, że błędy ε_i są niezależnymi zmiennymi losowym o tym samym rozkładzie normalnym.

Najprostszym związkiem między obserwcjami x_i a odpowiedziami y_i jest zależność liniowa, możliwy jest jednak również inny związek między obserwacjami a odpowiedziami, co prowadzi do sformułowania hipotezy

$$H_0 \colon f(x) = ax + b \; vs. \; H_1 \colon f \in \mathcal{F}$$

gdzie $\mathcal{F}$ jest klasą funkcji wypukłych.

W niniejszej pracy postaramy się skonstruować odpowiedni do postawionego problemu test statystyczny. Zaproponowane zostanie rozwiązanie oparte o iloraz wiarogodniści w przypadku modelu regresji z ograniczeniami w postaci nierówności.

W pierwszym rozdziale zostaną omówione podstawowe własności stożków wypukłych traktowanych jako po podzbiór przestrzeni liniowej. Drugi rozdział będzie traktował o konstrukcji estymatora regresji wypukłej jako rzutu wektora danych na wielościan wypukły przy pomocy algorytmu baz prymalno- dualnych. W trzecim rozdziale zostanie wyznaczony rozkład szukanego testu w przypadku ze znaną wariancją błędu obserwacji.

Praca została napisana na podstawie [2], natomiast rozdział o algorytmie baz prymalno- dualnych został napisany w dużym stopniu na podstawie [1].

2. Stożki wypukłe

Poszukiwany test zostanie wyznaczony metodą rzutowania wektora danych na wielościan powstały w wyniku narzuconych ograniczeń liniowym. W tym rozdziale zostaną przedstwione podstawowe definicje i własności stożków i wielościanów wypukłych użyteczne w dalszych rozważaniach.

Definicja 1 (**Ortant**). *Ortantem w n- wymiarowej przestrzeni $\mathbb{R}^n$ nazywamy podzbiór powstały przez ograniczenie każdej ze współrzędnych do bycia nieujemną lub niedodatnią, czyli*

$$O = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \colon \epsilon_i x_i \geqslant 0, |\epsilon_i| = 1, i = 1, 2, \ldots, n\}$$

Definicja 2 (**Stożek wypukły**). *Niech V będzie przestrzenią wektorową. Stożkiem wypukłym nazywamy przecięcie skończonej ilości półprzestrzeni przestrzeni V.*

Rozważmy n- wymiarową przestrzeń wektorową V. Dowolną półprzestrzeń H przestrzeni V można wyrazić jako

$$H = \{(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \colon a_1 x_1 + a_2 x_2 + \ldots a_n x_n \geqslant b\}$$

gdzie $a_1, a_2, \ldots, a_n, b$ są pewnymi, ustalonymi liczbami rzeczywistymi.

Korzystając z tego przedstawienia możemy dowolny stożek wypukły P zapisać jako

$$K = \bigcap_{i=1}^{m} H_i,$$

gdzie

$$H_j = \{\mathbf{x} \in \mathbb{R}^n \colon \sum_{i=1}^{n} a_i^j x_j \geqslant b^j\}.$$

Stąd możemy zapisać, że

$$K = \left\{ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n \colon \left\{ \begin{array}{l} \sum_{i=1}^{n} a_i^1 x_i \geqslant b^1 \\ \sum_{i=1}^{n} a_i^2 x_i \geqslant b^2 \\ \vdots \\ \sum_{i=1}^{n} a_i^m x_i \geqslant b^m \end{array} \right. \right\}$$

co będziemy zapisywać skrótowo jako

$$K = \{\mathbf{x} \in \mathbb{R}^n \colon A\mathbf{x} \geqslant b\}$$

gdzie

$$A = \begin{bmatrix} a_1^1 & a_2^1 & \ldots & a_n^1 \\ a_1^2 & a_2^2 & \ldots & a_m^2 \\ \vdots & \vdots & \ldots & \vdots \\ a_1^n & a_2^n & \ldots & a_m^n \end{bmatrix} \in \mathbb{R}^{m \times n}, b^T = [b^1, b^2, \ldots, b^m] \in \mathbb{R}^m$$

Symbolem $\langle \cdot, \cdot \rangle$ będziemy oznaczać iloczyn skalarny w przestrzeni wektorowej V. Oznaczmy przez γ_i kolejne wiersze macierzy $-A$, załużmy ponadto, że

tworzą one układ wektorów liniowo niezależnych oraz że $m \leqslant n$. Wtedy stożek K możemy też zapisać w sposób

$$K = \{\mathbf{x} \in \mathbb{R}^n \colon \langle \mathbf{x}, \gamma_i \rangle \leqslant 0, i = 1, 2, \ldots, m\}$$

Uzupełniając zbiór wektorów $\{\gamma_i\}$ do bazy przestrzni $\mathbb{R}^n$ o wektory ortogonalne i definując bazę dualną złożoną z wektorów β_i w następujący sposób

$$\beta_i^T \gamma_j = \begin{cases} -1, & i = j \\ 0, & i \neq j \end{cases}$$

możemy zapisać równoważne przedstawienie stożka K

$$K = \{\mathbf{x} \in \mathbb{R}^n \colon \mathbf{x} = \sum_{i=1}^{m} b_i \beta_i + \sum_{i=m+1}^{n} c_i \beta_i, b_i \geqslant 0, c_i \in \mathbb{R}\}$$

Twierdzenie 1. *Przedsawienia*

$$K = \{\boldsymbol{x} \in \mathbb{R}^n \colon \langle \boldsymbol{x}, \gamma_i \rangle \leqslant 0, i = 1, 2, \ldots, m\}$$

$$K = \{\boldsymbol{x} \in \mathbb{R}^n \colon \boldsymbol{x} = \sum_{i=1}^{m} b_i \beta_i + \sum_{i=m+1}^{n} c_i \beta_i, b_i \geqslant 0, c_i \in \mathbb{R}\}$$

są równoważne.

Dowód. Wektory $\beta_i, \gamma_i, i = 1, 2, \ldots, n$ spełniają zależność $\beta_i^T \gamma_j = \begin{cases} -1, & i=j \\ 0, & i \neq j \end{cases}$. Oznaczając rzez B,C macierze, których kolumnami są odpowiednio wektory β_i, γ_i, związek ten możemy przedstawić jako $B^T C = -I$. Wyrażenia $\langle \mathbf{x}, \gamma_i \rangle, i = 1, 2, \ldots, m$ są pierwszymi m współrzędnymi $C^T \mathbf{x}$. Ze związku $B^T C = -I$ dostajemy, że $C^T \mathbf{x} = -B^{-1} \mathbf{x}$. Zatem wektor $\mathbf{x}$ wyrażony w bazie złożonej z wektorów β_i ma pierwsze m współrzędnych nieujemnych, co dowodzi równoważności przedstawień. □

Definicja 3 (**wymiar stożka**). *Wymiarem stożka C nazywamy wymiar najmniejszej przestrzeni liniowej zawierającej stożek C.*

3. Regresja wypukła

Podobnie jak w przypadku zwykłego estymatora regresji liniowej, który jest rzutem wektora danych na pewną mniej wymiarową podprzestrzeń, tak w przypadku estymatora regresji wypukłej jest on rzutem na pewen wielościan wypukły powstały w wyniku stosownych ograniczeń.

Zbiór nad którym będziemy minimalizować kwadrat błędu powstaje w sposób następujący. Przypuśmy, że wartości x są różne między sobą i uporządkowane rosnąco oraz niech $\theta_i = f(x_i), i = 1, 2, \ldots, n$. Rozważając kawałkami liniowe przybliżenie funkcji regresji z węzłami w punktach x_i, wymóg wypukłości może zostać zapisany jako zbiór ograniczeń w postaci nierówności linowych następującej postaci:

$$\theta_i(x_{i+2} - x_{i+1}) - \theta_{i+1}(x_{i+2} - x_i) + \theta_{i+2}(x_{i+1} - x_i) \geqslant 0, i = 1, 2, \ldots, n-2$$

Zgodnie z definicją 1 możemy zbiór tych ograniczeń zapisać jako

$$K = \{G\theta \geqslant 0\}$$

gdzie G jest rzeczywistą macierzą wymiaru $(n-2) \times n$.

W tym momencie nasz problem znalezienia estymatora regresji wypukłej przyjmuje postać

$$\text{minimalizuj } ||y - \theta||^2 \text{ po } \theta \in K.$$

Niech $B = (e_1, e_2, \ldots, e_n)$ oznacza bazę kanoniczną przestrzeni $\mathbb{R}^n$. Oznaczmy przez $\gamma_i = -e_i^T G^T$, $i = 1, 2, \ldots, n-2$. Wtedy zbiór K możemy zapisać jako $K = \{\theta \in \mathbb{R}^n \colon -e_i^T G^T \theta \leqslant 0, i = 1, 2, \ldots, n-2\} = \{\theta \in \mathbb{R}^k \colon \gamma_i \circ \theta \leqslant 0, i = 1, 2, \ldots, n-2\}$.

Z określenia macierzy G oraz wektorów $\gamma_i, i = 1, 2, \ldots, n-2$, widać, że tworzą one układ wektorów liniowo niezależnych. Zatem zbiór $B'_\gamma = \{\gamma_i, i = 1, 2, \ldots, n-2\}$ można uzupełnić do bazy B_γ przestrzeni $\mathbb{R}^n$ o wektory γ_{n-1}, γ_n tak, żeby były one ortogonalne do wszytkich wektorów z bazy B'_γ. Łatwo sorawdzić, że warunek ten spełniają wektory $\gamma_{n-1} = \mathbf{1}$ oraz $\gamma_n = (x - \bar{x}\mathbf{1})$, gdzie $x = (x_1, x_2, \ldots, x_n)$, $\bar{x}$ oznacza wartość średnią, a norma $||\circ||$ jest normą.

Teraz możemy zdefiniować bazę $B_\beta = \{\beta_1, \beta_2, \ldots, \beta_n\}$ dualną do bazy B_γ w następujący sposób:

$$\beta_i^T \gamma_j = \begin{cases} -1, & i = j \\ 0, & i \neq j \end{cases}$$

Oznaczając przez A i H macierze, których kolumnami są odpowiednio wektory β_i i γ_i związek między nimi możemy wyrazić jako

$$A^T H = -I,$$

gdzie I oznacza macierz jednostkową.

Niech E oznacza podprzestrzeń przestrzeni $\mathbb{R}^n$ rozpinaną przez wektory β_{n-1}, β_n, natomiast $\mathcal{L}(K)$ oznacza przestrzeń rozpiętą przez wektory $\beta_i, i = 1, 2, \ldots, n-2$. Przestrzenie B oraz $\mathcal{L}(K)$ są do siebie ortogonalne, zatem wektor obserwacji y możemy zapisać jako sumę $y_E + z$, gdzie y_E i z są rzutami wektora y odpowiednio na podprzestrzeń E oraz $\mathcal{L}(K)$.

Prześledźmy powyższe rozważania na przykładzie dla przypadku czterowymiarowego i równoodległych punktów x_i odległych o 1.

Macierz ograniczeń G przybiera wtedy postać

$$G = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix}$$

Zatem stożek powstały z ograniczeń jest postaci

$$K = \{\theta \in \mathbb{R}^4 \colon \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \end{bmatrix} \theta \geqslant 0\}$$

Baza wektorów B_γ jest postaci

$$B_\gamma = ((-1, 2, -1, 0), (0, -1, 2, -1), (1, 1, 1, 1), (-3/2, -1/2, 1/2, 3/2))$$

Wektory β_i spełniające warunek $\langle \beta_i, \gamma_j \rangle = \begin{cases} -1, & i=j \\ 0, & i \neq j \end{cases}$ przybierają następująca postać

$$B_\beta = ((3, -4, -1, 2), (2, -1, -4, 3), (-1, -1, -1, -1), (3, 1, -1, -3))$$

Przestrzenie na które będziemy rzutować wektor danych przybierają postać

$$E = \{t_1(-1, -1, -1, 1) + t_2(3, 1, -1, -3), t_1, t_2 \in \mathbb{R}\}$$

$$\mathcal{L}(K) = \{t_1(3, -4, -1, 2) + t_2(2, -1, -4, 3), t_1, t_2 \in \mathbb{R}\}$$

Zadanie znalezienia rzutu wektora danych na stożek K sprowadza się w tym momencie do znalezienia rzutu jego składowych na stożek K. Wszytkie elementy podprzestrzeni E należą do stożka K, więc rzut wektora y na stożek K jest tym samym co jego rzut na podprzestrzeń E i wyraża się wzorem

$$y_E = X(X^T X)^{-1} X^T y,$$

gdzie macierz X jest podmacierzą macierzy A złożoną z pierwszych $n-2$ kolumn. Pozostaje zagadnienie znalezienia rzutu z na stożek K. Sprowadza się ono do znalezienia rzutu z na stożek

$$K' = K \cap \mathcal{L}(K) = \{\theta \in \mathbb{R}^n \colon \theta = \sum_{i=1}^{n-2} b_i \beta_i, b_i \geqslant 0\}.$$

W pracy [1] zostało pokazane, że przestrzeń $\mathcal{L}(K)$ może zostać podzielona na 2^{n-2} rozłącznych regionów w taki sposób, że każdy z nich może byś opisany jako nieujemny ortant w bazie $B_J = \{\beta_i, i \in J, \gamma_i, i \in L \setminus J\}$, gdzie J jest pewnym podzbiorem zbioru $L = \{1, 2, \ldots, n-2\}$. Zatem każdy element z należacy do $\mathcal{L}(K)$ może być przedstawiony w następujący sposób

$$z = \sum_{i \in J} b_i \beta_i + \sum_{i \in L \setminus J} c_i \gamma_i, \;\; b_i > 0, c_i \geqslant 0$$

Dla dowolnego zbioru $J \subset L$ B_J jest bazą przestrzeni $\mathcal{L}(K)$, ponadto $\beta_i, i \in J$ oraz $\gamma_i, i \in L \setminus J$ są wzajemnie ortogonalne, zatem rzutem z na K' jest wektor postaci

$$z_{K'} = \sum_{i \in J} b_i \beta_i, b_i > 0$$

Podsumowując, dowolny wektor y z przestrzeni $\mathbb{R}^n$ można przedstawić w następującej postaci

$$y = z + y_E = \sum_{i \in J} b_i \beta_i + \sum_{i \in L \setminus J} c_i \gamma_i + d_1 \gamma_{n-1} + d_2 \gamma_n, \ b_i > 0, c_i \geqslant 0, d_1, d_2 \in \mathbb{R}.$$

Wtedy rzut tego wektora na stożek

$$K = \{G\theta \geqslant 0\}$$

jest postaci

$$\hat{\theta} = \sum_{i \in J} b_i \beta_i + d_1 \gamma_{n-1} + d_2 \gamma_n.$$

Natomiast wektor błędu $\hat{\rho} = y - \hat{\theta}$ jest postaci

$$\hat{\rho} = \sum_{i \in L \setminus J} c_i \gamma_i.$$

4. Test statystyczny i jego rozkład

Na początek wprowadzimy kilka oznaczeń i udowodnimy trzy lematy z których skorzystamy w dalszej części rozważań.

$$C_{L\setminus J} = \{y \in \mathbb{R}^n \colon y = \sum_{i\in L\setminus J} b_i\gamma_i + \sum_{i\in J} c_i\beta_i + d_1\gamma_{n-1} + d_2\gamma_n, b_i > 0, c_i \geqslant 0, d_1, d_2 \in \mathbb{R}\} \quad (1)$$

$$S_{L\setminus J} = \operatorname{span}\{\gamma_i, i \in L \setminus J\} \quad (2)$$

$$d = |L \setminus J| = n - 2 - |J| \quad (3)$$

Niech

$$A_{L\setminus J} \quad (4)$$

oznacza macierz wymiaru $(n-2)\times n$ taką, że pierwsze d wierszy to wektory $-\gamma_i, i \in L \setminus J$ natomiast pozostałe $n-2-d$ wierszy to wektory $-\beta_i, i \in J$.

Lemat 1. *Niech $Z = (Z_1, Z_2, \ldots, Z_n)^T \sim N_n(0, \boldsymbol{I})$ oraz niech A będzie rzeczywistą macierzą wymiaru $m \times n$. Wtedy rozkładem warunkowym $||Z||^2$ pod warunkiem $AZ \geqslant 0$ jest χ^2_n, o ile zbiór $\{Z\colon AZ \geqslant 0\}$ jest niepusty.*

Dowód. Chcemy pokazać, że $P(||Z||^2 \leqslant a | AZ \geqslant 0) = \chi^2_n(a)$. W tym celu zapiszmy wektor Z we współrzędnych biegunowych

$$\begin{aligned}
Z_1 &= r\cos\phi_1\cos\phi_2\cos\phi_3\ldots\cos\phi_{n-1}\\
Z_2 &= r\sin\phi_1\cos\phi_2\cos\phi_3\ldots\cos\phi_{n-1}\\
Z_3 &= r\sin\phi_2\cos\phi_3\ldots\cos\phi_{n-1}\\
Z_4 &= r\sin\phi_3\ldots\cos\phi_{n-1}\\
&\vdots\\
Z_n &= r\sin\phi_{n-1},
\end{aligned}$$

gdzie $r \in (0,\infty), \phi_i \in [0, 2\pi), i = 1, 2, \ldots, n-1$.
Wtedy

$$||Z||^2 = r^2$$

oraz

$$AZ \geqslant 0 \iff A \begin{bmatrix} \cos\phi_1\cos\phi_2\cos\phi_3\ldots\cos\phi_{n-1}\\ \sin\phi_1\cos\phi_2\cos\phi_3\ldots\cos\phi_{n-1}\\ \sin\phi_2\cos\phi_3\ldots\cos\phi_{n-1}\\ \sin\phi_3\ldots\cos\phi_{n-1}\\ \vdots\\ \sin\phi_{n-1}\end{bmatrix} \geqslant 0.$$

Widzimy zatem, że wartość $||Z||^2$ zależy jedynie od wartości r natomiast warunek $AZ \geqslant 0$ dotyczy jedynie kąta, który jest niezależny od promienia r, a zatem $P(||Z||^2 \leqslant a | AZ \geqslant 0) = \chi^2_n(a)$. □

Lemat 2. *Niech $Z = (Z_1, Z_2, \ldots, Z_n)^T \sim N(0, \boldsymbol{I})$ oraz niech $\hat{Z}$ będzie rzutem Z na przestrzeń liniową S wymiaru $d < n$. Ponadto niech A będzie rzeczywistą macierzą wymiaru $m \times n$ taką, że każdy jej wiersz jest ortogonalny do przestrzeni S. Wtedy rozkładem warunkowym $||\hat{Z}||^2$ pod warunkiem $AZ \geqslant 0$ jest χ^2_d, o ile zbiór $\{AZ \geqslant 0\}$ jest niepusty.*

Dowód. Niech $v_1, v_2, \ldots, v_n$ będą wzjamnie ortonormalnymi wektorami w $\mathbb{R}^n$ takimi, że wektory $v_1, v_2, \ldots, v_d$ rozpinają przestrzeń S. Wektor Z możemy zapisać jako $Z = \sum_{i=1}^n a_i v_i$, gdzie $a_i = \langle v_i, z \rangle$. Stąd $a_i, i = 1, 2, \ldots, n$ są niezależnymi zmiennymi losowymi o standardowym rozkładzie normalnym oraz $\hat{Z} = \sum_{i=1}^d a_i v_i$. Wtedy $||\hat{Z}||^2 = a_1^2 + a_2^2 + \cdots + a_d^2$ co ma gęstość χ^2_d. Niech teraz V oznacza macierz taką, której poszczególne kolumny są kolejno wektorami $v_1, v_2, \ldots, v_n$ oraz niech $a = (a_1, a_2, \ldots, a_n)^T$. Macierz V możemy zapisać jako $V = [V_1|V_2]$, gdzie V_1 jest macierzą wymiaru $n \times d$, oznaczmy też przez a^1 wektor $(a_1, a_2, \ldots, a_d)^T$ a przez a^2 wektor $(a_{d+1}, \ldots, a_n)^T$. Wtedy $Z = Va = V_1 a^1 + V_2 a^2$ a warunek $AZ \geqslant 0$ możemy zapisać jako $AV_1 a^1 + AV_2 a^2 \geqslant 0$. Zauważmy, że z założeń oraz konstrukcji macierzy V dostajemy, że $AV_1 = 0$ oraz a^1, a^2 są niezależne. Zatem warunek $AZ \geqslant 0$ nie wpływa na gęstość $||\hat{Z}||^2$. □

Lemat 3. *Niech $y \in C_J$ dla pewngo zbioru $J \subset L = \{1, 2, \ldots, n-2\}$ oraz niech $a, b \in \mathbb{R}$. Wtedy $y' = y + a\gamma_{n-1} + b\gamma_n \in C_J$ oraz wektory błędów $\rho = y - \hat{\theta}$ i $\rho' = y' - \hat{\theta}'$ są sobie równe.*

Dowód. Jeśli $y \in C_J$ to y możemy zapisać jako $y = \sum_{i \in J} b_i \beta_i + \sum_{i \in L \setminus J} c_i \gamma_i + d_1 \gamma_{n-1} + d_2 \gamma_n,\ b_i > 0, c_i \geqslant 0, d_1, d_2 \in \mathbb{R}$. Wtedy $y' = \sum_{i \in J} b_i \beta_i + \sum_{i \in L \setminus J} c_i \gamma_i + (d_1 + a)\gamma_{n-1} + (d_2 + b)\gamma_n,\ b_i > 0, c_i \geqslant 0, d_1, d_2 \in \mathbb{R}$. Oczywiście $d_1 + a, d_2 + b \in \mathbb{R}$ zatem $y' \in C_J$.

Wektor ρ jest postaci $\rho = \sum_{i \in L \setminus J} c_i \gamma_i$. Z postaci wktora y' widzimy jednak, że $\rho' = \sum_{i \in L \setminus J} c_i \gamma_i = \rho$. □

Lemat 4. *Wektory losowe $y - \hat{\theta}$ i $\hat{\theta} - \hat{y}$ są niezależne.*

Dowód. Zauważmy, że $\langle y - \hat{\theta}, \hat{y} - \hat{\theta} \rangle = \langle y - \hat{\theta}, \hat{y} \rangle + \langle y - \hat{\theta}, \hat{\theta} \rangle = 0 + 0 = 0$. □

Dla ułatwienia rozważań założymy, że wariancja w zaproponowanym modelu σ^2 jest znana.

Teraz możemy przystąpić do wyliczania rozkładu testu opartego o iloraz wiarogodności hipotezy

$$H_0 \colon f(x) = ax + b \ vs. \ H_1 \colon f \in \mathcal{F}$$

gdzie $\mathcal{F}$ jest klasą funkcji wypukłych.

Niech $\hat{y}$ oznacza estymatorem regresji liniowej, czyli rzut wektora danych y na przestrzeń $\text{span}\{\gamma_{n-1}, \gamma_n\}$. Ponadto oznaczmy przez $R_0 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ oraz $R_1 = \sum_{i=1}^n (y_i - \hat{\theta}_i)^2$. Wtedy poszukiwany test przyjmuje postać

$$M = \frac{R_0 - R_1}{\sigma^2}.$$

W celu znalezienia rozkładu testu M, gdy prawdziwa jest hipoteza zerowa potrzebne będzie znalezienie wymiaru modelu i liczby stopni swobody błędu

dla modelu regresji wypukłej. Z postaci rzutu $\hat{\theta}$ wektora y można przypuszczać, że wymiar modelu wynosi $n - d$ oraz liczba stopni swobody błędu wynosi d. Jednak zbiór J jest losowy, różne wartości wektora błędu ε mogą umieścić wektor danych y w różnych zbiorach C_J. Wyliczenie rozkładu testu zaczniemy w następujący sposób

$$P(M \leqslant a) = \sum_{J \in \mathcal{P}(L)} P(M \leqslant a, y \in C_J = \sum_{J \in \mathcal{P}(L)} P(M \leqslant a | y \in C_J) P(y \in C_J),$$

gdzie $\mathcal{P}(L)$ oznacza zbiór potęgowy zbioru L.

Z Lematu 3. możemy bez straty ogólności założyć, że $f(x) = 0$ i $y = \varepsilon$. Dla dowolnej realizacji wektora $\varepsilon \in \mathbb{R}^n$ oznaczmy przez $L \setminus J$ taki zbiór indeksów, że $\varepsilon \in C_{L \setminus J}$. Niech $\hat{\varepsilon}$ będzie rzutem wektora ε na przestrzeń $S_{L \setminus J}$. Zauważmy, że macierz $A_{L \setminus J}$ można zapisać jako $[A^1 | A^2]$, gdzie macierz A^1 jest wymiaru $d \times n$. Zatem kolumny macierzy A^1 rozpinają $S_{L \setminus J}$, natomiast kolumny macierzy A^2 są ortogonalne do przestrzeni $S_{L \setminus J}$. Dodatkowo, gdy $\varepsilon \in C_J$, zachodzi $A^1 \varepsilon \geqslant 0$ oraz $A^2 \varepsilon \geqslant 0$. Stąd na mocy Lematu 2. dostajemy, że rozkładem warunkowym $\frac{||\hat{\varepsilon}||^2}{\sigma^2}$ przy zadanym J jest χ^2_d. Jako że $R_1 = ||\hat{\rho}||^2$ a przy założeniu prawdziwości hipotezy zerowej $||\hat{\rho}||^2$ jest równa $\frac{||\hat{\varepsilon}||^2}{\sigma^2}$ przy ustalonym zbiorze J możemy napisać następujący wniosek

Wniosek 1. *Jeśli hipoteza zerowa $\theta \in span\{\gamma_{n-1}, \gamma n\}$ jest prawdziwa to rozkładem warunkowym $\frac{R_1}{\sigma^2}$ przy ustalonym $y \in C_J$ jest χ^2_d, gdzie $d = |L \setminus J|$.*

Zmienna losowa $\frac{R_0}{\sigma^2}$ ma oczywiście rozkład χ^2_{n-2}. Niech D będzie zmienną losową reprezentującą liczność zbioru $L \setminus J$. Z Wniosku 1. mamy, że rozkładem warunkowym $\frac{R_1}{\sigma^2}$ pod warunkiem $D = d$ jest χ^2_d. Z Lematu 4. dostajemy zatem, że rozkładem warunkowym M jest χ^2_{n-d-2} pod warunkiem $D = d$. Stąd możemy zapisać następujący wniosek

Wniosek 2. *Przy założeniu prawdziwości hipotezy zerowej postawionego problemu mamy*

$$P(M \leqslant a) = \sum_{d=0}^{n-2} P(\chi^2_{n-d-2} \leqslant a) P(D = d),$$

gdzie $\chi^2_0 \equiv 0$.

Wartości prawdopodobieństw $P(D = d), d = 0, 1, \ldots, n-2$ jest wyliczane na podstawie względnych objętości zbiorów $C_J, J \in \mathcal{P}(L)$. Prawdopodobieństwo, że $y \in C_J$, gdy hipoteza zerowa jest prawdziwa, jest równoważne prawdopodobieństwu, że wektor losowy o n- wymiarowym standardowym rozkładzie normalny wpada do zbioru C_J.

Kropka nie oznacza końca zdania.
Ona daje możliwość coraz to lepszej kontynuacji.