Kartkówka nr 4 z wykładu
– dzisiaj pojawią się zagadnienia

Poprzednim razem

- Metoda największej wiarygodności
- Metoda najmniejszych kwadratów
  - Funkcja regresji E(Y|X)

1

Regresja krzywoliniowa - przykład

- Syt. W wyniku pomiaru otrzymaliśmy n par punktów (X_i,Y_i) i=1,…,n. Zakładamy, że pomiary są niezależne, a ich dokładność może być różna (tzn. odchylenia standardowe Y_i mogą być różne). Funkcję regresji przybliżamy funkcją $y=\sum_{j=1}^{m}\theta_j f_j(x)$, gdzie $f_j(x)$ to dowolne funkcje np. wielomiany.
- Zatem chcemy minimalizować

$$Q^2=\sum_{i=1}^{n}\frac{1}{\sigma_i^2}(y_i-\sum_{j=1}^{m}\theta_j f_j(x_i))^2$$

$$\frac{\partial Q^2}{\partial\theta_k}=\sum_{i=1}^{n}\frac{2}{\sigma_i^2}(y_i-\sum_{j=1}^{m}\theta_j f_j(x_i))(-f_k(x_i))$$

$$\sum_{i=1}^{n}\frac{2}{\sigma_i^2}(y_i-\sum_{j=1}^{m}\theta_j f_j(x_i))(-f_k(x_i))=0 \leftarrow$$

Układ m równań do rozwiązania (k=1,…,m)

RPiS 2024/2025 2

2

Regresja krzywoliniowa - przykład

- Zapis macierzowy

$$B_{ij}=\delta_{i,j}\frac{1}{(\sigma_i)^2}=\begin{pmatrix}(\sigma_1)^{-2} & 0 & \cdots & 0\\ 0 & (\sigma_2)^{-2} & & \vdots\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & \cdots & (\sigma_n)^{-2}\end{pmatrix} \qquad A_{ij}=\begin{pmatrix}f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1)\\ f_1(x_2) & f_2(x_2) & & \vdots\\ \vdots & & \ddots & \vdots\\ f_1(x_n) & \cdots & \cdots & f_m(x_n)\end{pmatrix}$$

$$Q^2=(\vec{Y}-A\vec{\theta})^T B(\vec{Y}-A\vec{\theta}) \qquad \vec{Y}=\begin{pmatrix}Y_1\\ \vdots\\ Y_n\end{pmatrix} \qquad \vec{\theta}=\begin{pmatrix}\theta_1\\ \vdots\\ \theta_m\end{pmatrix}$$

$$\frac{\partial Q^2}{\partial\theta_k}=\left[-2A^T B(\vec{Y}-A\vec{\theta})\right]_k$$

$$\frac{\partial Q^2}{\partial\theta_k}=0 \qquad \rightarrow \qquad A^T B(\vec{Y}-A\vec{\theta})=0$$

$$A^T B\vec{Y}=A^T BA\vec{\theta} \quad /(A^T BA)^{-1}$$

$$\boxed{\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}}$$

RPiS 2024/2025 3

3

Regresja krzywoliniowa – przykład liniowy

$$\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}$$

- Wniosek: liniowy związek Y i Θ.
- Współczynniki Θ łatwe do wyliczenia (odwracanie macierzy)
- Przykład: regresja liniowa (m=2, $f_1(x_i)=x_i$, $f_2(x_i)=1$), dodatkowo zakładamy $\sigma_1=\sigma_2=\ldots=\sigma_n\equiv\sigma_y$.

RPiS 2024/2025 4

4

Przykład: regresja liniowa zwyczajna

- regresja liniowa ($f_1(x_i)=x_i$, $f_2(x_i)=1$); zakładamy $\sigma_1=\sigma_2=\ldots=\sigma_n\equiv\sigma_y$.

$$f_1(x_i)=x_i \quad f_2(x_i)=1$$

$$y_i=\theta_1 x_i+\theta_2 \qquad \vec{Y}=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix} \qquad \vec{\theta}=\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}$$

$$\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}$$

$$A_{ij}=\begin{pmatrix}f_1(x_1) & f_2(x_1) & \cdots & f_m(x_1)\\ f_1(x_2) & f_2(x_2) & & \vdots\\ \vdots & & \ddots & \vdots\\ f_1(x_n) & \cdots & \cdots & f_m(x_n)\end{pmatrix} \rightarrow \underset{(n\times 2)}{A_{ij}}=\begin{pmatrix}x_1,1\\ x_2,1\\ \vdots,\vdots\\ x_n,1\end{pmatrix}$$

$$B_{ij}=\delta_{i,j}\frac{1}{(\sigma_i)^2}=\begin{pmatrix}(\sigma_1)^{-2} & 0 & \cdots & 0\\ 0 & (\sigma_2)^{-2} & & \vdots\\ \vdots & & \ddots & \vdots\\ 0 & \cdots & \cdots & (\sigma_n)^{-2}\end{pmatrix} \rightarrow \underset{(n\times n)}{B_{ij}}=\frac{1}{\sigma_y^2}\begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & \vdots\\ 0 & 0 & \ddots & 0\\ 0 & \ldots & 0 & 1\end{pmatrix}$$

RPiS 2024/2025 5

5

Przykład: regresja liniowa zwyczajna

$$\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}$$

$$\begin{pmatrix}\theta_1\\ \theta_2\end{pmatrix}=\left[\begin{pmatrix}x_1 & x_2\cdots x_n\\ 1 & 1\ \cdots 1\end{pmatrix}\begin{pmatrix}1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1\end{pmatrix}\frac{1}{\sigma_y^2}\begin{pmatrix}x_1,1\\ x_2,1\\ \vdots,\vdots\\ x_n,1\end{pmatrix}\right]^{-1}\begin{pmatrix}x_1\ldots x_n\\ 1 \quad 1\end{pmatrix}\frac{1}{\sigma_y^2}\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

W macierzy do odwracania: macierz jednostkową pomijamy

$$\begin{pmatrix}x_1 & x_2\ldots x_n\\ 1 & 1 \quad 1\end{pmatrix}\begin{pmatrix}x_1\,1\\ x_2\,1\\ \vdots\,1\\ x_n\,1\end{pmatrix}=\begin{pmatrix}\sum x_i^2 & \sum x_i\\ \sum x_i & n\end{pmatrix} \qquad \begin{bmatrix}a & b\\ c & d\end{bmatrix}^{-1}=\frac{1}{ad-bc}\begin{bmatrix}d & -b\\ -c & a\end{bmatrix}$$

$$(A^T BA)^{-1}=\frac{\sigma_y^2}{\det A^T 1A}\begin{pmatrix}n & -\sum x_i\\ -\sum x_i & \sum x_i^2\end{pmatrix}=\frac{\sigma_y^2}{n\sum x_i^2-(\sum x_i)^2}\begin{pmatrix}n & -\sum x_i\\ -\sum x_i & \sum x_i^2\end{pmatrix}$$

RPiS 2024/2025 6

6

Przykład: regresja liniowa zwyczajna

$$\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}$$

Zatem $(A^T BA)^{-1}A^T B=\frac{\sigma_y^2}{n\sum x_i^2-(\sum x_i)^2}\begin{pmatrix} n & -\sum x_i \\ -\sum x_i & \sum x_i^2 \end{pmatrix}\begin{pmatrix} x_1 \dots x_n \\ 1 \dots 1 \end{pmatrix}\frac{1}{\sigma_y^2}\begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix}=$

$$=\frac{1}{n\sum x_i^2-(\sum x_i)^2}\begin{pmatrix} nx_1-\sum x_i & , & \dots & , & nx_n-\sum x_i \\ -x_1\sum x_i+\sum x_i^2, & & \dots & , & -x_n\sum x_i+\sum x_i^2 \end{pmatrix}$$

i ostatecznie

$$\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}=\frac{1}{n\sum x_i^2-(\sum x_i)^2}\begin{pmatrix} nx_1-\sum x_i & , & \dots & , & nx_n-\sum x_i \\ -x_1\sum x_i+\sum x_i^2, & & \dots & , & -x_n\sum x_i+\sum x_i^2 \end{pmatrix}\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}=$$

$$=\frac{1}{n\sum x_i^2-(\sum x_i)^2}\begin{pmatrix} nx_1y_1-y_1\sum x_i+nx_2y_2-y_2\sum x_i+ & \dots & nx_ny_n-y_n\sum x_i \\ -x_1y_1\sum x_i+y_1\sum x_i^2-x_2y_2\sum x_i+y_2\sum x_i^2- & \dots & x_ny_n\sum x_i+y_n\sum x_i^2 \end{pmatrix}=$$

$$=\frac{1}{n\sum x_i^2-(\sum x_i)^2}\begin{pmatrix} n\sum x_iy_i-\sum x_i\sum y_i \\ -\sum x_i\sum x_iy_i+\sum x_i^2\sum y_i \end{pmatrix}$$

7

Regresja krzywoliniowa – przykład liniowy

$$\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}$$

- Wniosek: liniowy związek Y i Θ.
- Współczynniki Θ łatwe do wyliczenia (odwracanie macierzy)
- Przykład: regresja liniowa (m=2, $f_1(x_i)=x_i$, $f_2(x_i)=1$), dodatkowo zakładamy $\sigma_1=\sigma_2=\dots=\sigma_n\equiv\sigma_y$.

$$Y=\theta_1 f_1(X)+\theta_2 f_2(X)=\theta_1 X+\theta_2$$

$$\rightarrow \boxed{\begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix}=\frac{1}{n\sum_{i=1}^n x_i^2-\left(\sum_{i=1}^n x_i\right)^2}\begin{pmatrix} n\sum_{i=1}^n x_iy_i-\sum_{i=1}^n x_i\sum_{i=1}^n y_i \\ \sum_{i=1}^n x_i^2\sum_{i=1}^n y_i-\sum_{i=1}^n x_i\sum_{i=1}^n x_iy_i \end{pmatrix}}$$

- Jest to **regresja zwyczajna**.
- Istnieją też inne odmiany regresji: **klasyczna** (nic nie wiemy o σ_i), **ważona** (różne σ_i), **efektywna** (uwzględnia także niepewność x_i).

8

Prawo przenoszenia błędów

- Dla liniowych związków pomiędzy Y i Θ znajdujemy macierz kowariancji C(Θ) (ćwiczenia)

$$\vec{\theta}=(A^T BA)^{-1}A^T B\vec{Y}\equiv T\vec{Y}$$

$$\theta_i=T_{ij}Y_j$$

$$E(\theta_i)=T_{ij}\,E(Y_j)$$

$$\boxed{C_{ij}(\vec{\theta})\equiv Cov(\theta_i,\theta_j)=\sum_{k=1}^n\sum_{l=1}^n T_{ik}T_{jl}Cov(Y_k,Y_l)}$$

$$C(\vec{\theta})=TC(\vec{Y})T^T$$

9

Prawo przenoszenia błędów

- W naszym przykładzie

$$C(\vec{Y})=B^{-1}\quad\rightarrow$$

$$C(\vec{\theta})=(A^T BA)^{-1}A^T BB^{-1}\left[(A^T BA)^{-1}A^T B\right]^T=$$

$$=(A^T BA)^{-1}A^T\left[A^{-1}B^{-1}(A^T)^{-1}A^T B\right]^T=$$

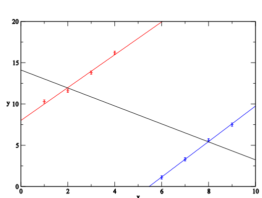
$$=(A^T BA)^{-1}A^T(A^{-1})^T=(A^T BA)^{-1}A^T(A^T)^{-1}=(A^T BA)^{-1}$$

$$C_{ij}(\vec{\theta})=\begin{pmatrix} var(\theta_1) & cov(\theta_1,\theta_2) \\ cov(\theta_1,\theta_2) & var(\theta_2) \end{pmatrix}=$$

$$=\sigma_y^2\left[A^T A\right]^{-1}=\frac{\sigma_y^2}{n\sum_{i=1}^n x_i^2-\left(\sum_{i=1}^n x_i\right)^2}\begin{pmatrix} n & -\sum_{i=1}^n x_i \\ -\sum_{i=1}^n x_i & \sum_{i=1}^n x_i^2 \end{pmatrix}$$

10

Paradoks Simpsona



y=1.99x+8.0

y=2.15x-11.75

y= -1.0883x+14.117

- Paradoks Simpsona – połączenie różnych grup może zmienić interpretację wyników. Prowadzi to do praktycznych problemów np. związanych z wyborem kuracji osoby chorej. Rozwiązanie problemu prowadzi do sieci Bayesowskich i teorii grafów i polega na prześledzeniu zależności pomiędzy zmiennymi X, Y i innymi zmiennymi, które mogą mieć wpływ na zmienne X i Y.

11

Miara jakości dopasowania – „test Chi-kwadrat"

Zał. σ_i mają rozkład normalny

$$M=\sum_{i=1}^n\frac{(y_i-\sum_{j=1}^p\theta_j f_j(x_i))^2}{\sigma_i^2}$$

- E(M)=n-p – dla rozkładu Chi-kwadrat o n-p stopniach swobody.
- Gdy $M>\chi_\gamma^{2,n-p}$ (n- liczba pomiarów, p liczba parametrów) to na poziomie ufności γ odrzucamy wynik dopasowania.
- Dla regresji liniowej mamy (p=2)

$$M=\sum_{i=1}^n\frac{(y_i-ax_i-b)^2}{\sigma_i^2}$$

- W praktyce sprawdzamy czy M/(n-2)≈1 – prosta dobrze przybliża.
- M/(n-2)<1 prosta, ale za duże σ_i; M/(n-2)>1 inna krzywa, za małe σ_i lub niektóre pomiary odstają od prostej.

12