2 Jádrové odhady

2.1 Lokálně polynomiální jádrové odhady

V tomto odstavci se budeme věnovat speciálnímu typu jádrových odhadů, který se nazývá *lokálně polynomiální*. Používá se především při odhadech regresní funkce. Hodnota neznámé regresní funkce v libovolném bodě plánu se získá tak, že proložíme dané body polynomem stupně p váženou metodou nejmenších čtverců. Ve speciálním případě pro $p = 0$, tj. prokládáním konstantou, obdržíme tzv. Nadarayovy – Watsonovy estimátory. Podobně, pro $p = 1$, prokládáme-li naměřená data přímkou, získáme lokálně lineární estimátory. Naším úkolem bude odvodit formální vzorec pro obecný stupeň polynomu.

Označme $\widehat{m}(x; p, h)$ odhad regresní funkce m v bodě x proložením polynomu stupně p váženou metodou nejmenších čtverců. Nechť tento polynom má tvar

$$P(u) = \beta_0 + \beta_1(u - x) + \cdots + \beta_p(u - x)^p.$$

Nechť K je nezáporné jádro na intervalu $[-1, 1]$, tj. $K(x) \geq 0,\ \forall x \in [-1, 1]$. Hodnotu $\widehat{m}(x; p, h)$ získáme váženou metodou nejmenších čtverců, tj. minimalizujeme funkcionál Φ v závislosti na vektoru parametrů $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)'$

$$\Phi = \sum_{i=0}^{T-1} \{Y_i - \beta_0 - \beta_1(x_i - x) - \cdots - \beta_p(x_i - x)^p\}^2 K_h(x_i - x).$$

Označme $\widehat{\boldsymbol{\beta}} = (\hat{\beta}_0, \ldots \hat{\beta}_p)'$ vektor, pro který Φ nabývá minimální hodnoty. Odhad regresní funkce v bodě x získaný výše popsanou metodou je tedy hodnota parametru $\hat{\beta}_0$, tj.

$$\widehat{m}(x; p, h) = \hat{\beta}_0.$$

Naším cílem je najít explicitní vyjádření pro $\hat{\beta}_0$. To je popsáno v níže uvedené větě. Nejprve však předchází pomocné tvrzení, které bude potřeba při důkazu této věty.

Nechť $g(\boldsymbol{\beta})$ je skalární funkce vektoru $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_p)'$. Derivací funkce g podle vektoru $\boldsymbol{\beta}$ rozumíme vektor

$$\frac{\partial g(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} = \left(\frac{\partial g(\boldsymbol{\beta})}{\partial \beta_0}, \ldots, \frac{\partial g(\boldsymbol{\beta})}{\partial \beta_p}\right)'.$$

Lemma 2.1.1. *Nechť $\boldsymbol{c} = (c_0, \ldots, c_p)'$ je vektor délky $p+1$ a $\boldsymbol{A} = (a_{ij})_{i,j=0}^{p}$ je symetrická čtvercová matice řádu $p + 1$. Pak*

1. $\dfrac{\partial \boldsymbol{c}'\boldsymbol{\beta}}{\partial \boldsymbol{\beta}} = \boldsymbol{c}$

2. $\dfrac{\partial \boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}}{\partial \boldsymbol{\beta}} = 2\boldsymbol{A}\boldsymbol{\beta}.$

Důkaz. Nejdřív dokážeme první vztah. Nechť $0 \le k \le p$, derivací výrazu $\boldsymbol{c}'\boldsymbol{\beta}$ podle β_k dostáváme k-tou složku vektoru $\frac{\partial \boldsymbol{c}'\boldsymbol{\beta}}{\partial \boldsymbol{\beta}}$. Rozepsáním dostáváme $\boldsymbol{c}'\boldsymbol{\beta} = \sum_{i=0}^{p} c_i\beta_i$, derivujeme-li tento součet podle β_k, obdržíme právě c_k.
Druhý vztah se dokáže podobně. Nejprve podrobněji rozepíšeme $\boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}$

$$\begin{aligned}\boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta} &= \sum_{i=0}^{p}\sum_{j=0}^{p} a_{ij}\beta_i\beta_j \\ &= \sum_{\substack{i=0\\ i\neq k}}^{p}\sum_{\substack{j=0\\ j\neq k}}^{p} a_{ij}\beta_i\beta_j + \sum_{\substack{i=0\\ i\neq k}}^{p} a_{ik}\beta_i\beta_k + \sum_{\substack{j=0\\ j\neq k}}^{p} a_{kj}\beta_k\beta_j + a_{kk}\beta_k^2\end{aligned}$$

a následně zderivujeme podle β_k

$$\frac{\partial \boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}}{\partial \beta_k} = \sum_{\substack{i=0\\ i\neq k}}^{p} a_{ik}\beta_i + \sum_{\substack{j=0\\ j\neq k}}^{p} a_{kj}\beta_j + 2a_{kk}\beta_k = \sum_{\substack{i=0\\ i\neq k}}^{p}(a_{ik}+a_{ki})\beta_i + 2a_{kk}\beta_k.$$

Podle předpokladu je $\boldsymbol{A}$ symetrická matice, tj. $a_{ij} = a_{ji}$, pro všechna $0 \le i, j \le p$, a tedy

$$\frac{\partial \boldsymbol{\beta}'\boldsymbol{A}\boldsymbol{\beta}}{\partial \beta_k} = 2\sum_{i=0}^{p} a_{ki}\beta_i,$$

což je k-tá složka vektoru $2\boldsymbol{A}\boldsymbol{\beta}$. □

V dalším budeme používat následující označení

$$\boldsymbol{Y} = \begin{pmatrix} Y_0 \\ \vdots \\ Y_{T-1}\end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} 1 & x_0 - x & \dots & (x_0-x)^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{T-1}-x & \dots & (x_{T-1}-x)^p\end{pmatrix},$$

$$\boldsymbol{W} = \begin{pmatrix} K_h(x_0-x) & 0 & \dots & 0 \\ 0 & K_h(x_1-x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & K_h(x_{T-1}-x)\end{pmatrix},$$

$$\boldsymbol{e}_1 = (1, 0 \dots 0)' \quad \text{vektor délky } p+1.$$

Věta 2.1.2. *Nechť je matice $\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}$ regulární. Pro odhad regresní funkce platí*

$$\widehat{m}(x; p, h) = \boldsymbol{e}_1'(\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}. \tag{3}$$

Důkaz. Hodnotu $\widehat{m}(x;p,h)$ chceme získat váženou metodou nejmenších čtverců, tj. minimalizací funkce

$$\Phi = \sum_{i=0}^{T-1} \{Y_i - \beta_0 - \beta_1(x_i - x) - \cdots - \beta_p(x_i - x)^p\}^2 K_h(x_i - x)$$

v závislosti na vektoru parametrů $\boldsymbol{\beta} = (\beta_0, \ \ldots, \ \beta_p)'$. Tuto funkci lze zapsat vektorově

$$\Phi = (\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{W}(\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\beta}).$$

Roznásobením dostáváme

$$\Phi = \boldsymbol{Y}'\boldsymbol{W}\boldsymbol{Y} - \boldsymbol{Y}'\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta}.$$

Nyní zderivujeme Φ podle $\boldsymbol{\beta}$ a výraz položíme roven nule. Protože matice $\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}$ je symetrická, můžeme při derivování použít předchozí lemma

$$\frac{\partial \Phi}{\partial \boldsymbol{\beta}} = \boldsymbol{0} - \boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y} - \boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{0}.$$

Po jednoduché úpravě dostáváme tzv. *vážený systém normálních rovnic*

$$\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}.$$

Podle předpokladu je matice $\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}$ regulární, a tedy řešení existuje a je rovno

$$\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}.$$

Nakonec stačí vyjádřit první složku $\hat{\beta}_0$, která je odhadem regresní funkce m, tj.

$$\widehat{m}(x;p,h) = \hat{\beta}_0 = \boldsymbol{e}_1'(\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}.$$

□

Příklad. Uvažujme speciální případ, kdy stupeň polynomu je $p = 0$, tj. naměřená data prokládáme lokálně konstantou. Jednotlivé matice jsou tvaru

$$\boldsymbol{X} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \boldsymbol{W} = \begin{pmatrix} K_h(x_0 - x) & 0 & \ldots & 0 \\ 0 & K_h(x_1 - x) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & K_h(x_{T-1} - x) \end{pmatrix}, \quad \boldsymbol{e}_1 = 1.$$

Spočítáme jejich součin $\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X}$

$$\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X} = \sum_{i=0}^{T-1} K_h(x_i - x).$$

Za předpokladu, že tento součet je nenulový, můžeme vypočítat inverzi

$$(\boldsymbol{X}'\boldsymbol{W}\boldsymbol{X})^{-1} = \frac{1}{\sum\limits_{i=0}^{T-1} K_h(x_i - x)}.$$

Nakonec vyjádříme součin $\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y}$

$$\boldsymbol{X}'\boldsymbol{W}\boldsymbol{Y} = \sum_{i=0}^{T-1} K_h(x_i - x)Y_i.$$

Dosazením do (3) dostáváme

$$\widehat{m}(x; 0, h) = \frac{\sum\limits_{i=0}^{T-1} K_h(x_i - x)Y_i}{\sum\limits_{i=0}^{T-1} K_h(x - x_i)}.$$

Takto sestrojené odhady regresní funkce se nazývají *Nadarayovy – Watsonovy odhady.*

Poznámka. V tomto odstavci jsme užili předpokladu, že jádro K je nezáporné. I když jádra vyšších řádů tuto podmínku nesplňují, používají se ke konstrukci lokálně polynomiálních odhadů pro jejich dobré asymptotické vlastnosti (viz [19]).

2.2 Statistické vlastnosti odhadů

Kvalitu jádrového odhadu lze lokálně popsat pomocí střední kvadratické chyby. Budeme se zabývat asymptotickým tvarem této chyby, neboť pro neparametrické odhady, narozdíl od odhadů parametrických, neexistuje nevychýlený odhad, tj. takový odhad, že $E\widehat{m} = m$ pro s.v. $x \in \mathbb{R}$ (Collomb 1976).

Věta 2.2.1. *Nechť $K \in S_{0\kappa}, E\boldsymbol{Y}^2 < \infty$ a nechť posloupnost vyhlazovacích parametrů $h = h_T,\ T = 1, 2, \ldots,$ splňuje podmínky: $h_T \to 0,\ Th_T \to \infty$ pro $T \to \infty$. Pak v každém bodě spojitosti funkce m platí*

$$\sum_{k=0}^{T-1} W_k(x)Y_k \xrightarrow{p} m(x),$$

kde W_k jsou váhové funkce odpovídající postupně odhadům $\widehat{m}_{PCH}$, $\widehat{m}_{NW}$, $\widehat{m}_{LL}$, $\widehat{m}_{GM}$.

Důkaz. Důkaz můžeme najít např. v [6]. $\square$

Poznámka. Uvedené odhady jsou tedy *konzistentními* odhady m.