UNIVERZITA PALACKÉHO V OLOMOUCI

PŘÍRODOVĚDECKÁ FAKULTA

DIPLOMOVÁ PRÁCE

Regrese s chybami v proměnných

Katedra matematické analýzy a aplikací matematiky

Vedoucí diplomové práce: **Doc. RNDr. Eva Fišerová, Ph.D.**

Vypracovala: **Bc. Vendula Tichá**

Studijní program: N1103 Aplikovaná matematika

Studijní obor: Aplikace matematiky v ekonomii

Forma studia: prezenční

Rok odevzdání: 2014

BIBLIOGRAFICKÁ IDENTIFIKACE

Autor: Bc. Vendula Tichá

Název práce: Regrese s chybami v proměnných

Typ práce: Diplomová práce

Pracoviště: Katedra matematické analýzy a aplikací matematiky

Vedoucí práce: Doc. RNDr. Eva Fišerová, Ph.D.

Rok obhajoby práce: 2015

Abstrakt: Cílem diplomové práce je popis regresních modelů, které lze použít v situaci, kdy nejen vysvětlovaná proměnná, ale i vysvětlující proměnné jsou zatíženy chybami. Nejprve je popsáno modelování lineárního vztahu mezi proměnnými a následně modelování polynomiálního vztahu. V závěrečné části jsou uvedeny příklady a simulační studie.

Klíčová slova: ortogonální regrese, maximálně věrohodné odhady, odhady pomocí lineárních modelů s podmínkou typu II., polynomiální regrese, upravená metoda nejmenších čtverců.

Počet stran: 85

Počet příloh: 1

Jazyk: český

BIBLIOGRAPHICAL IDENTIFICATION

Author: Bc. Vendula Tichá

Title: Regression with errors in variables

Type of thesis: Postgraduate´s

Department: Department of Mathematical Analysis and Application of Mathematics

Supervisor: Doc. RNDr. Eva Fišerová, Ph.D.

The year of presentation: 2015

Abstract: The aim of this thesis is to describe regression models for a situation when both response and explanatory variables are measured with errors. Firstly, the modeling of linear relationship is described and the description of polynomial regression follows. Finally, some illustrative examples and simulation study are presented.

Key words: orthogonal regression, the maximum likelihood estimators, the estimates using linear models with type-II constraints, polynomial regression, the adjusted least squares estimator.
Number of pages: 85
Number of appendices: 1
Language: Czech

Prohlášení

Prohlašuji, že jsem diplomovou práci zpracovala samostatně pod vedením doc. RNDr. Evy Fišerové, Ph.D. a uvedla jsem všechny použité zdroje.

V Olomouci dne 10. 12. 2014 …………………………......

Tichá Vendula

Obsah

Poděkování

Děkuji vedoucí diplomové práce Doc. RNDr. Evě Fišerové, Ph.D. za trpělivost a čas, který mi věnovala při tvorbě této diplomové práce.

Úvod

Tématem diplomové práce je popis regresních metod, které lze použít v situaci, kdy nejen vysvětlovaná proměnná, ale i vysvětlující proměnné jsou zatíženy chybami. Práce je rozdělena na tři kapitoly, které jsou dále děleny na podkapitoly.

První kapitola je věnována modelování lineárního vztahu mezi proměnnými pomocí ortogonální regrese. Odvozeny jsou odhady parametrů jak pomocí metody maximální věrohodnosti, tak i pomocí lineárních modelů s podmínkou typu II. Popsány jsou i intervaly spolehlivosti a testování hypotéz.

Druhá kapitola se zabývá modelováním polynomiálního vztahu za pomoci polynomiální regrese a odhadem parametrů pomocí upravené metody nejmenších čtverců. Představena je také i modifikace upravené metody nejmenších čtverců.

Poslední kapitola je věnována příkladům, pro jejichž výpočet byl použit software R. V prvním příkladě zpracováváme reálný datový soubor „Gejzír Old Faithful v Národním parku v Yellowstone“. Ve druhém příkladě umělá data vygenerovaná pro předem daný lineární vztah. Následuje simulační studie intervalů spolehlivosti pro směrnici ortogonální regresní přímky a závěrečný příklad zabývající se polynomiální regresí.

1 Modelování lineárního vztahu mezi proměnnými

Lineární vztah mezi proměnnými budeme modelovat pomocí ortogonální regrese, kterou si podrobně rozebere v následující podkapitole. V další podkapitole se budeme zabývat odhadem parametrů pomocí metody maximální věrohodnosti pro stejné a rozdílné rozptyly. Dále se budeme zabývat odhadem parametrů pomocí lineárních modelů s podmínkou typu II, což je spolu se statistickou inferencí rozebráno v následujících podkapitolách. Při tvorbě této kapitoly byly použity zejména tyto zdroje [1, 2, 3, 4, 7, 16].

1.1 Ortogonální regrese

Ortogonální regrese, také známá jako metoda úplných nejmenších čtverců nebo regrese s chybami v proměnných, analyzuje vztah mezi proměnnými. V porovnání se standardní regresí, závislé a nezávislé proměnné jsou zatíženy chybou měření.

Pro formulaci ortogonální regrese v nejjednodušší podobě uvažujeme lineární vztah mezi dvěma proměnnými μ_i a ν_i (dáno *n* pozorování), tj.

$$\nu_i = \beta_0 + \beta_1 \mu_i, \quad i = 1, \dots, n, \tag{1}$$

kde β_0, β_1 jsou neznámé parametry, β_0 je absolutní člen a β_1 je směrnice ortogonální regresní přímky.

Měření bodů $(\mu_i, \nu_i)', i = 1, \dots, n,$ je narušené náhodnými chybami měření, proto místo nich pozorujeme body $(x_i, y_i)'$, kde

$$x_i = \mu_i + \varepsilon_{1i}, \quad y_i = \nu_i + \varepsilon_{2i}. \tag{2}$$

Nepozorované proměnné μ_i a ν_i se stanou bezchybnými hodnotami pozorovaných proměnných x_i, y_i a $\varepsilon_{1i}, \varepsilon_{2i}, i = 1, \dots, n$ představují náhodné chyby měření μ_i a ν_i. Dále budeme předpokládat, že náhodné chyby jsou nezávislé náhodné veličiny s nulovou střední hodnotou a rozptyly $var(\varepsilon_{1i}) = \sigma_1^2$ a $var(\varepsilon_{2i}) = \sigma_2^2$.

Pokud jsou čísla μ_i daná (v experimentu se nějakým způsobem nastavují), pak se model (1) nazývá funkční vztah. Pokud jsou μ_i nezávislé stejně rozdělené náhodné veličiny, pak se model (1) nazývá strukturální relace, podrobněji viz [1]. Dále se budeme věnovat modelům s funkčním vztahem, kdy předpokládáme, že rozptyly jsou si rovny, tj. $\sigma_1^2 = \sigma_2^2 = \sigma^2$.

Dosadíme-li z (2) do (1), dostaneme

$$y_i = \beta_0 + \beta_1 x_i + (\varepsilon_{2i} - \beta_1 \varepsilon_{1i}), \qquad i = 1, \dots, n.$$

Na rozdíl od klasického regresního modelu, kde hodnoty nezávislé proměnné nezávisejí na vektoru chyb, zde dostaneme vztah

$$cov(x_i, \varepsilon_{2i} - \beta_1 \varepsilon_{1i}) = cov(\mu_i + \varepsilon_{1i}, \varepsilon_{2i} - \beta_1 \varepsilon_{1i}) = -\beta_1 \sigma_1^2.$$

Tato hodnota je obecně nenulová, což nám komplikuje odvozování příslušných statistických metod.

Oproti obyčejné metodě nejmenších čtverců, ortogonální regrese minimalizuje součet čtverců vzdáleností pozorovaných bodů od regresní přímky. Vzdálenost bodu $A = [a_1, a_2]$ od přímky p dané obecnou rovnicí $p: ax + by + c = 0$ vypočteme pomocí vztahu

$$v(A, p) = \frac{|a \cdot a_1 + b \cdot a_2 + c|}{\sqrt{a^2 + b^2}}.$$

Potom čtverec vzdálenosti bodu (x_i, y_i) od přímky $y = \beta_0 + \beta_1 x$ je

$$\frac{(y_i - \beta_0 - \beta_1 x_i)^2}{1+\beta_1{}^2}.$$

Odhad parametrů β_0 a β_1 ortogonální regresní přímky představuje řešení minimalizační úlohy

$$\min_{\beta_0, \beta_1} \frac{\sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2}{1 + \beta_1{}^2}.$$

Součet čtverců činí

$$S(\beta_0, \beta_1) = \left(1+{\beta_1}^2\right)^{-1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2.$$

Položíme-li

$$c_i = x_i - \bar{x}, \qquad f_i = y_i - \bar{y}, \qquad A = \beta_0 - \bar{y} + \beta_1 \bar{x},$$

platí, že

$$\sum_{i=1}^{n} c_i = 0\,, \qquad \sum_{i=1}^{n} f_i = 0\,, \qquad \frac{1}{n}\sum_{i=1}^{n} c_i^2 = s_x^2\,, \quad \frac{1}{n}\sum_{i=1}^{n} f_i^2 = s_{y,}^2 \qquad \frac{1}{n}\sum_{i=1}^{n} c_i f_i = s_{xy},$$

kde s_x^2 a s_y^2 značí výběrové průměry a s_{xy} značí výběrovou kovarianci. Postupnými úpravami dostaneme

$$\begin{aligned}
S(\beta_0, \beta_1) &= \left(1+{\beta_1}^2\right)^{-1} \sum_{i=1}^{n} [y_i - \bar{y} - \beta_0 + \bar{y} - \beta_1 \bar{x} - \beta_1 (x_i - \bar{x})]^2 \\
&= \left(1+{\beta_1}^2\right)^{-1} \sum_{i=1}^{n} (f_i - A - \beta_1 c_i)^2 \\
&= \left(1+{\beta_1}^2\right)^{-1} \left(\sum_{i=1}^{n} f_i^2 - 2A \sum_{i=1}^{n} f_i - 2\beta_1 \sum_{i=1}^{n} c_i f_i + nA^2 + 2A\beta_1 \sum_{i=1}^{n} c_i + {\beta_1}^2 \sum_{i=1}^{n} c_i^2 \right) \\
&= \left(1+{\beta_1}^2\right)^{-1} \left(\sum_{i=1}^{n} f_i^2 + nA^2 + {\beta_1}^2 \sum_{i=1}^{n} c_i^2 - 2\beta_1 \sum_{i=1}^{n} c_i f_i \right) \\
&\geq \left(1+{\beta_1}^2\right)^{-1} \left(\sum_{i=1}^{n} f_i^2 + {\beta_1}^2 \sum_{i=1}^{n} c_i^2 - 2\beta_1 \sum_{i=1}^{n} c_i f_i \right) \\
&= n\left(1+{\beta_1}^2\right)^{-1} \left(s_x^2 {\beta_1}^2 - 2 s_{xy} \beta_1 + s_y^2 \right) = S(\beta_1),
\end{aligned}$$

přičemž rovnost $S(\beta_0, \beta_1) = S(\beta_1)$ nastává právě tehdy, je-li $A = 0,$ tj. pro $\beta_0 = \bar{y} - \beta_1 \bar{x}.$ Minimum funkce $S(\beta_1)$ najdeme pomocí derivací. Protože

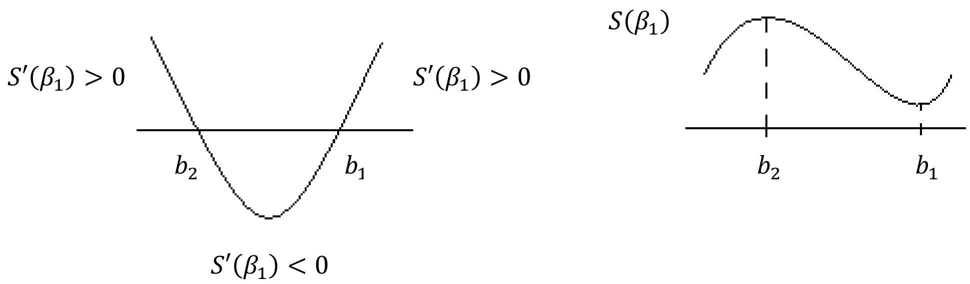
$$S'(\beta_1) = -n\left(1+{\beta_1}^2\right)^{-2} \beta_1 \left(s_x^2 {\beta_1}^2 - 2 s_{xy} \beta_1 + s_y^2 \right) + n\left(1+{\beta_1}^2\right)^{-1} \left(2 s_x^2 \beta_1 - 2 s_{xy} \right)$$

$$= -n\left(1+{\beta_1}^2\right)^{-2}\left[2s_x^2{\beta_1}^3 - 4s_{xy}{\beta_1}^2 + 2s_y^2\beta_1 - \left(1+{\beta_1}^2\right)2s_x^2\beta_1 + \left(1+{\beta_1}^2\right)2s_{xy}\right]$$
$$= -n\left(1+{\beta_1}^2\right)^{-2}\left[-2s_{xy}{\beta_1}^2 + 2s_y^2\beta_1 - 2s_x^2\beta_1 + 2s_{xy}\right]$$
$$= 2n\left(1+{\beta_1}^2\right)^{-2}\left[s_{xy}{\beta_1}^2 - \left(s_y^2 - s_x^2\right)\beta_1 - s_{xy}\right],$$

je derivace nulová v bodech

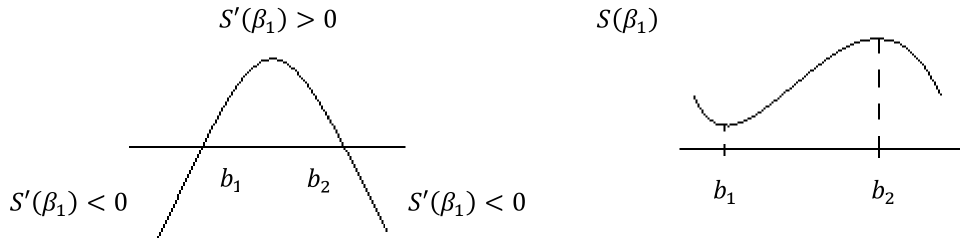
$$b_{1,2} = \frac{s_y^2 - s_x^2 \pm \sqrt{\left(s_y^2 - s_x^2\right)^2 + 4s_{xy}^2}}{2s_{xy}}.$$

Je-li $s_{xy} > 0$, pak $S'(\beta_1) > 0$ pro $\beta_1 < b_2$, $S'(\beta_1) < 0$ pro $b_2 < \beta_1 < b_1$ a $S'(\beta_1) > 0$ pro $\beta_1 > b_1$ viz obrázek 1. Jelikož je $S'(\beta_1)$ pro β_1 jdoucí k bodu b_1 zleva záporná a zprava kladná, původní funkce $S(\beta_1)$ nejprve klesá k bodu b_1 a následně roste. Proto $S(\beta_1)$ má minimum v bodě $\beta_1 = b_1$.



Obrázek 1: Vyšetření průběhu funkce $S(\beta_1)$ (vpravo) na základě její první derivace (vlevo) pro $s_{xy} > 0$.

Je-li $s_{xy} < 0$, potom $S'(\beta_1) < 0$ pro $\beta_1 < b_1$, $S'(\beta_1) > 0$ pro $b_1 < \beta_1 < b_2$ a $S'(\beta_1) < 0$ pro $b_2 < \beta_1$ viz obrázek 2. Jelikož je $S'(\beta_1)$ pro β_1 jdoucí k bodu b_1 zleva záporná a zprava kladná, původní funkce $S(\beta_1)$ nejprve klesá k bodu b_1 a následně roste. Proto $S(\beta_1)$ má minimum opět v bodě $\beta_1 = b_1$.



Obrázek 2: Vyšetření průběhu funkce $S(\beta_1)$ (vpravo) na základě její první derivace (vlevo) pro $s_{xy} < 0$.

Řešením jsme získali odhady

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \tag{3}$$

$$\hat{\beta}_1 = \frac{s_y^2 - s_x^2 + \sqrt{\left(s_y^2 - s_x^2\right)^2 + 4s_{xy}^2}}{2s_{xy}}, \tag{4}$$

kde $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ a $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$ jsou výběrové průměry, $s_x^2 = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2$ a $s_y^2 = \frac{1}{n}\sum_{i=1}^{n} y_i^2 - \bar{y}^2$ jsou výběrové rozptyly a $s_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}$ je výběrová kovariance.

1.2 Maximálně věrohodné odhady

Další možností, jak můžeme získat odhady je pomocí metody maximální věrohodnosti, kde předpokládáme, že $\sigma_1^2 = \sigma_2^2 = \sigma^2$ a náhodné chyby měření mají normální rozdělení. Potom platí, že

$$x_i \sim N(\mu_i, \sigma^2), \quad y_i \sim N(\beta_0 + \beta_1 \mu_i, \sigma^2).$$

Jelikož jsou veličiny $x_1, \dots, x_n, y_1, \dots, y_n$ nezávislé, je jejich sdružená hustota vyjádřena vztahem

$$f(x_1, \dots, x_n, y_1, \dots, y_n) = (2\pi)^{-n} \sigma^{-2n} exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left[(y_i - \beta_0 - \beta_1 \mu_i)^2 + (x_i - \mu_i)^2\right]\right\}.$$

Neznámé parametry tady jsou $\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2$ a počet neznámých parametrů je $n+3$, ale k dispozici máme jenom n dvojic (x_i, y_i). Parametry jsou přesto odhadnutelné, dostaneme $n+3$ normálních rovnic.

Věrohodnostní funkce je vyjádřena součinem sdružených hustot $f(x_1, \dots, x_n, y_1, \dots, y_n)$ jako

$$L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2) = \prod_{i=1}^{n} (2\pi)^{-n} \sigma^{-2n} exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} [(y_i - \beta_0 - \beta_1 \mu_i)^2 + (x_i - \mu_i)^2]\right\}.$$

Jelikož sdružená hustota obsahuje exponenciální funkci, použijeme pro výpočet přirozený logaritmus věrohodnostní funkce.

$$\begin{aligned} &ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2)] = \\ &\qquad ln[(2\pi)^{-n} \sigma^{-2n}] + ln\left(exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} [(y_i - \beta_0 - \beta_1 \mu_i)^2 + (x_i - \mu_i)^2]\right\}\right) \\ &\qquad = -n \cdot ln(2\pi) - n \cdot ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} [(y_i - \beta_0 - \beta_1 \mu_i)^2 + (x_i - \mu_i)^2]. \end{aligned}$$

Maximálně věrohodné odhady parametrů jsou takové odhady, které maximalizují pro daná pozorování hodnotu věrohodnostní funkce. Tyto odhady se nezmění, maximalizujeme-li namísto dané funkce její logaritmus. Odhady nalezneme tak, že derivujeme logaritmus věrohodnostní funkce podle neznámých parametrů a položíme je rovno nule, dostaneme tak soustavu rovnic:

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2)]}{\partial \beta_0}: \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 \mu_i) = 0, \tag{5}$$

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2)]}{\partial \beta_1}: \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 \mu_i)\mu_i = 0, \tag{6}$$

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2)]}{\partial \mu_i}: (y_i - \beta_0 - \beta_1 \mu_i)\beta_1 + (x_i - \mu_i) = 0, \quad 1 \le i \le n \tag{7}$$

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma^2)]}{\partial \sigma^2}: -\frac{n}{\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}[(y_i - \beta_0 - \beta_1\mu_i)^2 + (x_i - \mu_i)^2] = 0. \quad (8)$$

Ze vztahu (7) dostaneme:

$$x_i = \mu_i - (y_i - \beta_0 - \beta_1\mu_i)\beta_1.$$

Sečtením přes všechna *i* dostaneme:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu_i - \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1\mu_i)\beta_1.$$

Ze vztahu (5) víme, že druhý výraz na pravé straně rovnice je roven 0, takže dostaneme rovnost:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu_i .$$

Obě strany rovnice vynásobíme výrazem $\frac{1}{n}$ a dostaneme rovnost:

$$\bar{x} = \bar{\mu} .$$

Ze vztahu (5) dostaneme:

$$\sum_{i=1}^{n} y_i - n\beta_0 - \beta_1 \sum_{i=1}^{n} \mu_i = 0$$

$$-n\beta_0 = -\sum_{i=1}^{n} y_i + \beta_1 \sum_{i=1}^{n} \mu_i.$$

Obě strany rovnice vynásobíme výrazem $\left(-\frac{1}{n}\right)$ a dostaneme:

$$\beta_0 = \frac{1}{n}\sum_{i=1}^{n} y_i - \beta_1 \frac{1}{n}\sum_{i=1}^{n} \mu_i.$$

Užijeme rovnosti $\bar{x} = \bar{\mu}$ a dostaneme odhad parametru β_0 ve tvaru:

$$\hat{\beta}_0 = \bar{y} - \beta_1 \bar{x}.$$

Ze vztahu (7) dostaneme po úpravách:

$$y_i\beta_1 - \beta_0\beta_1 - {\beta_1}^2\mu_i + x_i - \mu_i = 0$$
$$\mu_i\left(1 + {\beta_1}^2\right) = x_i + y_i\beta_1 - \beta_0\beta_1 \,.$$

Obě strany rovnice vynásobíme výrazem $\frac{1}{\left(1+{\beta_1}^2\right)}$, dosadíme za $\beta_0 = \bar{y} - \beta_1\bar{x}$ a dostaneme:

$$\mu_i = \frac{x_i + y_i\beta_1 - \beta_1\bar{y} + {\beta_1}^2\bar{x}}{\left(1 + {\beta_1}^2\right)}$$

$$= \frac{x_i + \beta_1(y_i - \bar{y}) + {\beta_1}^2\bar{x} - {\beta_1}^2 x_i + {\beta_1}^2 x_i}{\left(1 + {\beta_1}^2\right)}$$

$$= \frac{x_i\left(1 + {\beta_1}^2\right)}{\left(1 + {\beta_1}^2\right)} + \frac{\beta_1[(y_i - \bar{y}) - \beta_1(x_i - \bar{x})]}{\left(1 + {\beta_1}^2\right)}$$

$$= x_i + \frac{\beta_1[(y_i - \bar{y}) - \beta_1(x_i - \bar{x})]}{\left(1 + {\beta_1}^2\right)}.$$

Získaný vztah pro μ_i nyní dosadíme do vztahu (6) :

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1\mu_i)\left\{x_i + \frac{\beta_1}{\left(1+{\beta_1}^2\right)}[(y_i-\bar{y}) - \beta_1(x_i-\bar{x})]\right\} = 0.$$

Roznásobením a s využitím vztahu (5) dostaneme:

$$\sum_{i=1}^{n} x_i(y_i - \beta_0 - \beta_1\mu_i) + \frac{\beta_1}{\left(1+{\beta_1}^2\right)}\sum_{i=1}^{n} y_i\,(y_i - \beta_0 - \beta_1\mu_i)$$
$$- \frac{{\beta_1}^2}{\left(1+{\beta_1}^2\right)}\sum_{i=1}^{n} x_i\,(y_i - \beta_0 - \beta_1\mu_i) = 0.$$

Roznásobíme:

$$\sum_{i=1}^{n} x_i y_i - \beta_0\sum_{i=1}^{n} x_i - \beta_1\sum_{i=1}^{n} x_i\,\mu_i + \frac{\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n} {y_i}^2 - \frac{\beta_0\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n} y_i - \frac{{\beta_1}^2}{1+{\beta_1}^2}\sum_{i=1}^{n} y_i\,\mu_i$$
$$- \frac{{\beta_1}^2}{1+{\beta_1}^2}\sum_{i=1}^{n} x_i y_i + \frac{\beta_0{\beta_1}^2}{1+{\beta_1}^2}\sum_{i=1}^{n} x_i + \frac{{\beta_1}^3}{1+{\beta_1}^2}\sum_{i=1}^{n} x_i\,\mu_i = 0.$$

Po úpravě dostaneme:

$$\frac{1}{1+{\beta_1}^2}\sum_{i=1}^{n} x_i y_i - \frac{\beta_0}{1+{\beta_1}^2}\sum_{i=1}^{n} x_i - \frac{\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n} x_i\,\mu_i + \frac{\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n} {y_i}^2$$
$$- \frac{\beta_0\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n} y_i - \frac{{\beta_1}^2}{1+{\beta_1}^2}\sum_{i=1}^{n} y_i\,\mu_i = 0.$$

Opět dosadíme za μ_i a po úpravě dostaneme:

$$\frac{1}{1+{\beta_1}^2}\sum_{i=1}^{n}x_iy_i-\frac{\beta_0}{1+{\beta_1}^2}\sum_{i=1}^{n}x_i+\frac{\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n}{y_i}^2-\frac{\beta_0\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n}y_i-\frac{\beta_1}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}{x_i}^2$$
$$-\frac{2{\beta_1}^2}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}x_iy_i-\frac{{\beta_1}^3}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}{y_i}^2+\frac{{\beta_1}^2}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i-\frac{{\beta_1}^3}{\left(1+{\beta_1}^2\right)^2}$$
$$\cdot\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2+\frac{{\beta_1}^3}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2-\frac{{\beta_1}^4}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i=0.$$

Dosadíme-li za $\beta_0=\bar{y}-\beta_1\bar{x}=\frac{1}{n}\sum_{i=1}^{n}y_i-\beta_1\cdot\frac{1}{n}\sum_{i=1}^{n}x_i$ dostaneme:

$$\frac{1}{1+{\beta_1}^2}\sum_{i=1}^{n}x_iy_i-\frac{1}{1+{\beta_1}^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i+\frac{\beta_1}{1+{\beta_1}^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2+\frac{\beta_1}{1+{\beta_1}^2}\sum_{i=1}^{n}{y_i}^2$$
$$-\frac{\beta_1}{1+{\beta_1}^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2+\frac{{\beta_1}^2}{1+{\beta_1}^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i-\frac{\beta_1}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}{x_i}^2-\frac{2{\beta_1}^2}{\left(1+{\beta_1}^2\right)^2}$$
$$\cdot\sum_{i=1}^{n}x_iy_i-\frac{{\beta_1}^3}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}{y_i}^2+\frac{{\beta_1}^2}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i-\frac{{\beta_1}^3}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2$$
$$+\frac{{\beta_1}^3}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2-\frac{{\beta_1}^4}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i=0.$$

Postupnými úpravami obdržíme:

$$\frac{1-{\beta_1}^2}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}x_iy_i-\frac{1-{\beta_1}^2}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i+\frac{\beta_1}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}x_i\right)^2$$
$$+\frac{\beta_1}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}{y_i}^2-\frac{\beta_1}{\left(1+{\beta_1}^2\right)^2}\cdot\frac{1}{n}\left(\sum_{i=1}^{n}y_i\right)^2-\frac{\beta_1}{\left(1+{\beta_1}^2\right)^2}\sum_{i=1}^{n}{x_i}^2=0.$$

Rovnici vynásobíme výrazem $\left(1+{\beta_1}^2\right)^2$:

$$\left(1-{\beta_1}^2\right)\sum_{i=1}^{n} x_i y_i - \left(1-{\beta_1}^2\right)\cdot\frac{1}{n}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + \beta_1\cdot\frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 + \beta_1\sum_{i=1}^{n} {y_i}^2$$
$$-\beta_1\cdot\frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2 - \beta_1\sum_{i=1}^{n} {x_i}^2 = 0.$$

Rovnici vynásobíme výrazem $\frac{n}{n}$:

$$n\left(1-{\beta_1}^2\right)\left[\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n^2}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right] - \beta_1 n\left[\frac{1}{n}\sum_{i=1}^{n} {x_i}^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2\right]$$
$$+\beta_1 n\left[\frac{1}{n}\sum_{i=1}^{n} {y_i}^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} y_i\right)^2\right] = 0.$$

Nyní využijeme vztahy pro výběrový rozptyl a výběrovou kovarianci:

$$s_x^2 = E(X^2) - [E(X)]^2 = \frac{1}{n}\sum_{i=1}^{n} {x_i}^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2$$
$$s_y^2 = E(Y^2) - [E(Y)]^2 = \frac{1}{n}\sum_{i=1}^{n} {y_i}^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} y_i\right)^2$$
$$s_{xy} = E(XY) - E(X)E(Y) = \frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n^2}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i,$$

které dosadíme do předchozí rovnice a získáme rovnici:

$$\left(1-{\beta_1}^2\right)s_{xy} - \beta_1 s_x^2 + \beta_1 s_y^2 = 0.$$

Tuto rovnici lze přepsat:

$$s_{xy} - {\beta_1}^2 s_{xy} + \beta_1\left(s_y^2 - s_x^2\right) = 0$$
$${\beta_1}^2 s_{xy} - \beta_1\left(s_y^2 - s_x^2\right) - s_{xy} = 0.$$

Získali jsme stejnou rovnici jako v případě ortogonální regrese. Věrohodnostní funkci maximalizuje pouze jeden kořen této kvadratické rovnice. (Předpokládáme, že $s_{xy} \neq 0$), a to

$$\hat{\beta}_1 = \frac{\left(s_y^2 - s_x^2\right) + \sqrt{\left(s_y^2 - s_x^2\right)^2 + 4{s_{xy}}^2}}{2s_{xy}}.$$

Odhady pomocí metody maximální věrohodnosti a pomocí ortogonální metody nejmenších čtverců jsou stejné v posuzovaném speciálním případě, tj. x_i a y_i jsou nezávislé normálně rozdělené náhodné veličiny se stejným rozptylem.

Pokud rozptyl σ^2 neznáme, můžeme jeho maximálně věrohodnostní odhad získat úpravou vztahu (8)

$$\frac{1}{2\sigma^4}\sum_{i=1}^{n}[(y_i - \beta_0 - \beta_1\mu_i)^2 + (x_i - \mu_i)^2] = \frac{n}{\sigma^2}.$$

Obě strany rovnice vynásobíme výrazem $\frac{\sigma^4}{n}$ a dostaneme výsledný odhad parametru σ^2 ve tvaru:

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}\left[\left(y_i - \hat{\beta}_0 - \hat{\beta}_1\hat{\mu}_i\right)^2 + (x_i - \hat{\mu}_i)^2\right]}{2n}, \tag{9}$$

kde odhad $\hat{\mu}_i, i = 1, \dots, n$, získáme po úpravě vztahu (7) ve tvaru

$$\hat{\mu}_i = \frac{x_i + \hat{\beta}_1 y_i - \hat{\beta}_0\hat{\beta}_1}{1 + \hat{\beta}_1^2}. \tag{10}$$

Odhad $\hat{\sigma}^2$ konverguje podle pravděpodobnosti k $\sigma^2/2$. Tato konkrétní nekonzistence nezpůsobuje žádné problémy, konzistentní odhad σ^2 je prostě $2n\hat{\sigma}^2/(n-2)$. Odhad $\hat{\mu}_i$ je také nekonzistentní [2]. Nakonec, odhad $v_i, i = 1, \dots, n$ je

$$\hat{v}_i = \hat{\beta}_0 + \hat{\beta}_1\hat{\mu}_i.$$

1.2.1 Maximálně věrohodné odhady pro různé rozptyly

Pokud jsou rozptyly σ_1^2 a σ_2^2 rozdílné, pak $\sigma_1^2 = \lambda\sigma_2^2$. Dále budeme předpokládat, že $\lambda > 0$ je známé. Potom platí, že

$$x_i \sim N(\mu_i, \lambda\sigma_2^2), \quad y_i \sim N(\beta_0 + \beta_1\mu_i, \sigma_2^2).$$

Jelikož jsou veličiny $x_1, \dots, x_n, y_1, \dots, y_n$ nezávislé, je jejich sdružená hustota vyjádřena

$$f(\boldsymbol{x}, \boldsymbol{y}) = (2\pi)^{-n}\lambda^{(-n/2)}\sigma_2^{-2n} exp\left\{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{n}\left[(y_i - \beta_0 - \beta_1\mu_i)^2 + \frac{(x_i - \mu_i)^2}{\lambda}\right]\right\}.$$

Věrohodnostní funkce má tvar

$$L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma_2^2) = \prod_{i=1}^{n}(2\pi)^{-n}\lambda^{(-n/2)}\sigma_2^{-2n} exp\left\{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{n}\left[(y_i - \beta_0 - \beta_1\mu_i)^2 + \frac{(x_i - \mu_i)^2}{\lambda}\right]\right\}.$$

Jelikož sdružená hustota opět obsahuje exponenciální funkci, tak pro výpočet použijeme přirozený logaritmus věrohodnostní funkce.

$$ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma_2^2)] = ln\left[(2\pi)^{-n}\lambda^{(-n/2)}\sigma_2^{-2n}\right] + ln\left(exp\left\{-\frac{1}{2\sigma_2^2}\sum_{i=1}^{n}\left[(y_i - \beta_0 - \beta_1\mu_i)^2 + \frac{(x_i - \mu_i)^2}{\lambda}\right]\right\}\right)$$
$$= -n \cdot ln(2\pi) - n \cdot ln\left(\sqrt{\lambda}\,\sigma_2^2\right) - \frac{1}{2\sigma_2^2}\sum_{i=1}^{n}\left[(y_i - \beta_0 - \beta_1\mu_i)^2 + \frac{(x_i - \mu_i)^2}{\lambda}\right].$$

Odhady nalezneme tak, že derivujeme logaritmus věrohodnostní funkce podle neznámých parametrů a položíme je rovno nule, dostaneme tak soustavu rovnic:

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma_2^2)]}{\partial\beta_0}: \sum_{i=1}^{n}(y_i - \beta_0 - \beta_1\mu_i) = 0, \tag{11}$$

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma_2^2)]}{\partial \beta_1}: \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 \mu_i)\mu_i = 0, \tag{12}$$

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma_2^2)]}{\partial \mu_i}: (y_i - \beta_0 - \beta_1 \mu_i)\beta_1 + \frac{(x_i - \mu_i)}{\lambda} = 0, \quad 1 \le i \le n \tag{13}$$

$$\frac{\partial ln[L(\beta_0, \beta_1, \mu_1, \dots, \mu_n, \sigma_2^2)]}{\partial \sigma^2}:$$

$$-\frac{n}{\sqrt{\lambda}\,\sigma_2^2} + \frac{1}{2\sigma_2^4} \sum_{i=1}^{n} \left[(y_i - \beta_0 - \beta_1 \mu_i)^2 + \frac{(x_i - \mu_i)^2}{\lambda} \right] = 0. \tag{14}$$

Ze vztahu (13) dostaneme:

$$x_i = \mu_i - \lambda (y_i - \beta_0 - \beta_1 \mu_i)\beta_1.$$

Sečtením přes všechna *i* dostaneme:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu_i - \lambda \beta_1 \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 \mu_i).$$

Ze vztahu (11) víme, že druhý výraz na pravé straně rovnice je roven 0, takže dostaneme:

$$\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} \mu_i.$$

Obě strany rovnice vynásobíme výrazem $\frac{1}{n}$ a dostaneme rovnost:

$$\bar{x} = \bar{\mu}.$$

Ze vztahu (11) dostaneme po úpravách stejný odhad parametru β_0 jako v předchozím případě ve tvaru:

$$\hat{\beta}_0 = \bar{y} - \beta_1 \bar{x}.$$

Ze vztahu (13) dostaneme po úpravách:

$$y_i\beta_1 - \beta_0\beta_1 - {\beta_1}^2\mu_i + \frac{x_i}{\lambda} - \frac{\mu_i}{\lambda} = 0$$

$$\mu_i\left(\frac{1}{\lambda} + {\beta_1}^2\right) = \frac{x_i}{\lambda} + y_i\beta_1 - \beta_0\beta_1.$$

Obě strany rovnice vynásobíme výrazem $\frac{1}{(1+{\beta_1}^2)}$, dosadíme za $\beta_0 = \bar{y} - \beta_1\bar{x}$ a dostaneme:

$$\mu_i = \frac{\frac{x_i}{\lambda} + y_i\beta_1 - \beta_1\bar{y} + {\beta_1}^2\bar{x}}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)}$$

$$= \frac{\frac{x_i}{\lambda} + \beta_1(y_i - \bar{y}) + {\beta_1}^2\bar{x} - {\beta_1}^2 x_i + {\beta_1}^2 x_i}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)}$$

$$= \frac{x_i\left(\frac{1}{\lambda} + {\beta_1}^2\right)}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)} + \frac{\beta_1[(y_i - \bar{y}) - \beta_1(x_i - \bar{x})]}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)}$$

$$= x_i + \frac{\beta_1[(y_i - \bar{y}) - \beta_1(x_i - \bar{x})]}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)}.$$

Získaný vztah pro μ_i nyní dosadíme do vztahu (12) :

$$\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1\mu_i)\left\{x_i + \frac{\beta_1}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)}[(y_i - \bar{y}) - \beta_1(x_i - \bar{x})]\right\} = 0.$$

Roznásobením a s využitím vztahu (11) dostaneme:

$$\sum_{i=1}^{n} x_i(y_i - \beta_0 - \beta_1\mu_i) + \frac{\beta_1}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)} \sum_{i=1}^{n} y_i\,(y_i - \beta_0 - \beta_1\mu_i)$$
$$- \frac{{\beta_1}^2}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)} \sum_{i=1}^{n} x_i\,(y_i - \beta_0 - \beta_1\mu_i) = 0.$$

Roznásobíme a po úpravě dostaneme:

$$\frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} x_i y_i - \frac{\frac{1}{\lambda}\beta_0}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} x_i - \frac{\frac{1}{\lambda}\beta_1}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} x_i\,\mu_i + \frac{\beta_1}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} {y_i}^2$$
$$- \frac{\beta_0\beta_1}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} y_i - \frac{{\beta_1}^2}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} y_i\,\mu_i = 0.$$

Opět dosadíme za μ_i a po úpravě dostaneme:

$$\frac{\frac{1}{\lambda}}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} x_i y_i - \frac{\frac{1}{\lambda}\beta_0}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} x_i + \frac{\beta_1}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} {y_i}^2 - \frac{\beta_0\beta_1}{\frac{1}{\lambda} + {\beta_1}^2} \sum_{i=1}^{n} y_i - \frac{\left(\frac{1}{\lambda}\right)^2 \beta_1}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2} \sum_{i=1}^{n} {x_i}^2$$
$$- \frac{\frac{2}{\lambda}{\beta_1}^2}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2} \sum_{i=1}^{n} x_i y_i - \frac{{\beta_1}^3}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2} \sum_{i=1}^{n} {y_i}^2 + \frac{\frac{1}{\lambda}{\beta_1}^2}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2} \cdot \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i - \frac{\frac{1}{\lambda}{\beta_1}^3}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2}$$
$$\cdot \frac{1}{n} \left(\sum_{i=1}^{n} x_i\right)^2 + \frac{{\beta_1}^3}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2} \cdot \frac{1}{n} \left(\sum_{i=1}^{n} y_i\right)^2 - \frac{{\beta_1}^4}{\left(\frac{1}{\lambda} + {\beta_1}^2\right)^2} \cdot \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i = 0.$$

Vztah upravujeme stejným způsobem, jako v případě kdy rozptyly byly stejné a dojdeme k rovnici:

$$n\left[\left(\frac{1}{\lambda}\right)^2 - \frac{1}{\lambda}{\beta_1}^2\right]\left[\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n^2}\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i\right] - \frac{1}{\lambda}\beta_1 n\left[\frac{1}{\lambda}\cdot\frac{1}{n}\sum_{i=1}^{n} {x_i}^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2\right]$$
$$+\frac{1}{\lambda}\beta_1 n\left[\frac{1}{n}\sum_{i=1}^{n} {y_i}^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} y_i\right)^2\right] = 0.$$

Využitím vztahů pro výběrový rozptyl a výběrovou kovarianci získáme rovnici:

$$\left(\frac{1}{\lambda} - {\beta_1}^2\right) s_{xy} - \frac{1}{\lambda}\beta_1 s_x^2 + \beta_1 s_y^2 = 0.$$

Tuto rovnici můžeme přepsat:

$$\frac{1}{\lambda} s_{xy} - {\beta_1}^2 s_{xy} - \frac{1}{\lambda}\beta_1 s_x^2 + \beta_1 s_y^2 = 0$$

rovnici vynásobíme λ

$$\lambda {\beta_1}^2 s_{xy} - \beta_1\left(\lambda s_y^2 - s_x^2\right) - s_{xy} = 0.$$

Věrohodnostní funkci maximalizuje pouze jeden kořen této kvadratické rovnice. (Předpokládáme, že $s_{xy} \neq 0$), a to

$$\hat{\beta}_1 = \frac{\lambda s_y^2 - s_x^2 + \sqrt{\left(\lambda s_y^2 - s_x^2\right)^2 + 4\lambda s_{xy}^2}}{2\lambda s_{xy}}.$$

1.3 Odhady pomocí lineárních modelů s podmínkou typu II

1.3.1 Lineární modely s podmínkou typu II

Uvažujeme regulární model nepřímého měření vektorového parametru s podmínkou typu II tvaru:

$$\boldsymbol{Y} = \boldsymbol{X\gamma} + \boldsymbol{\varepsilon}, \quad \boldsymbol{b} + \boldsymbol{B\gamma} + \boldsymbol{C\kappa} = \boldsymbol{o}, \; var(\boldsymbol{\varepsilon}) = \boldsymbol{\Sigma},$$

kde $\boldsymbol{\gamma} \in \mathbb{R}^{k_1}, \boldsymbol{\kappa} \in \mathbb{R}^{k_2}$, $\boldsymbol{\Sigma}$ je pozitivně definitní matice, $\boldsymbol{X}_{n\times k_1}$ je matice s hodností $h(\boldsymbol{X}) = k_1 \leq n$, $\boldsymbol{B}$ a $\boldsymbol{C}$ jsou matice typu $q \times k_1$ a $q \times k_2$ s hodnostmi $\mathrm{h}(\mathbf{C}) = \mathrm{k}_2 < q$, $\mathrm{h}(\mathbf{C}, \mathbf{B}) = \mathrm{q} < \mathrm{k}_1 + \mathrm{k}_2$ a $\boldsymbol{b} \in \mathbb{R}^q$ je daný vektor.

Užitím metody zobecněných nejmenších čtverců získáme nejlepší nestranné lineární odhady vektorových parametrů $\boldsymbol{\gamma}$ a $\boldsymbol{\kappa}$, kdy minimalizujeme Lagrangeovu funkci:

$$\Phi(\boldsymbol{\gamma}, \boldsymbol{\kappa}) = (\mathbf{Y} - \mathbf{X}\boldsymbol{\gamma})'\boldsymbol{\Sigma}^{-1}(\mathbf{Y} - \mathbf{X}\boldsymbol{\gamma}) - 2\boldsymbol{\lambda}'(\mathbf{b} + \mathbf{B}\boldsymbol{\gamma} + \mathbf{C}\boldsymbol{\kappa}),$$

kde $\boldsymbol{\lambda} \in \mathbb{R}^{\mathrm{q}}$ je vektor lagrangeových multiplikátorů.

Provedeme parciální derivaci funkce $\Phi(\boldsymbol{\gamma}, \boldsymbol{\kappa})$ [5] a položíme rovno nule

$$\frac{\partial\Phi(\boldsymbol{\gamma}, \boldsymbol{\kappa})}{\partial\boldsymbol{\gamma}} = \mathbf{2X'\Sigma^{-1}X}\boldsymbol{\gamma} - \mathbf{2X'\Sigma^{-1}Y} - \mathbf{2B'}\boldsymbol{\lambda} = \mathbf{0}$$

$$\frac{\partial\Phi(\boldsymbol{\gamma}, \boldsymbol{\kappa})}{\partial\boldsymbol{\kappa}} = -\mathbf{2C'}\boldsymbol{\lambda} = \mathbf{0}.$$

Vyjádříme si, že

$$\boldsymbol{\gamma} = (\mathbf{X'\Sigma^{-1}X})^{-1}[\mathbf{X'\Sigma^{-1}Y} + \mathbf{B'}\boldsymbol{\lambda}],$$

potom $\boldsymbol{\gamma}$ dosadíme do podmínky $\mathbf{b} + \mathbf{B}\boldsymbol{\gamma} + \mathbf{C}\boldsymbol{\kappa} = \mathbf{0}$ a řešíme systém podmínek:

$$\mathbf{b} + \mathbf{B}(\mathbf{X'\Sigma^{-1}X})^{-1}[\mathbf{X'\Sigma^{-1}Y} + \mathbf{B'}\boldsymbol{\lambda}] + \mathbf{C}\boldsymbol{\kappa} = \mathbf{0}$$

$$\mathbf{C'}\boldsymbol{\lambda} = \mathbf{0}.$$

Vektorově to můžeme zapsat

$$\begin{pmatrix} \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}' & \mathbf{C} \\ \mathbf{C}' & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\kappa} \end{pmatrix} = -\begin{pmatrix} \mathbf{b} + \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y} \\ \mathbf{o} \end{pmatrix}.$$

Označme

$$\begin{pmatrix} \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}' & \mathbf{C} \\ \mathbf{C}' & \mathbf{0} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} = \mathbb{Q}.$$

Potom

$$\begin{pmatrix} \boldsymbol{\lambda} \\ \boldsymbol{\kappa} \end{pmatrix} = -\begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{b} + \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y} \\ \mathbf{o} \end{pmatrix},$$

a tedy

$$\boldsymbol{\lambda} = -\mathbf{Q}_{11}(\mathbf{b} + \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y})$$
$$\widehat{\widehat{\boldsymbol{\kappa}}} = -\mathbf{Q}_{21}(\mathbf{b} + \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y}).$$

Nyní dosadíme $\boldsymbol{\lambda}$ do výrazu pro odhad vektorového parametru $\boldsymbol{\gamma}$ a dostaneme

$$\widehat{\widehat{\boldsymbol{\gamma}}} = (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}[\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y} - \mathbf{B}'\mathbf{Q}_{11}(\mathbf{b} + \mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y})].$$

Nyní spočítáme varianční matici odhadů $\widehat{\widehat{\boldsymbol{\gamma}}}$ a $\widehat{\widehat{\boldsymbol{\kappa}}}$ **,** která má tvar

$$\mathbf{var}\begin{pmatrix} \widehat{\widehat{\boldsymbol{\gamma}}} \\ \widehat{\widehat{\boldsymbol{\kappa}}} \end{pmatrix} = \begin{pmatrix} \mathrm{var}(\widehat{\widehat{\boldsymbol{\gamma}}}) & \mathrm{cov}(\widehat{\widehat{\boldsymbol{\gamma}}}, \widehat{\widehat{\boldsymbol{\kappa}}}) \\ \mathrm{cov}(\widehat{\widehat{\boldsymbol{\gamma}}}, \widehat{\widehat{\boldsymbol{\kappa}}}) & \mathrm{var}(\widehat{\widehat{\boldsymbol{\kappa}}}) \end{pmatrix}.$$

Jestliže

$$\begin{pmatrix} \widehat{\widehat{\boldsymbol{\gamma}}} \\ \widehat{\widehat{\boldsymbol{\kappa}}} \end{pmatrix} = \begin{pmatrix} -(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q}_{11} \\ -\mathbf{Q}_{21} \end{pmatrix} \mathbf{b} + \begin{pmatrix} \mathbf{I} - (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q}_{11}\mathbf{B} \\ -\mathbf{Q}_{21}\mathbf{B} \end{pmatrix} (\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{Y},$$

je hledaná varianční matice tvaru

$$\mathbf{var}\begin{pmatrix}\widehat{\boldsymbol{\gamma}}\\ \widehat{\boldsymbol{\kappa}}\end{pmatrix} = \begin{pmatrix}\mathbf{I}-(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}B}\\ -\mathbf{Q_{21}B}\end{pmatrix}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{var}(\mathbf{Y})\boldsymbol{\Sigma}^{-1}\mathbf{X}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$$
$$\times\,(\mathbf{I}-\mathbf{B}'\mathbf{Q_{11}}'\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\quad;\quad -\mathbf{B}'\mathbf{Q_{21}}').$$

Potom

$$\begin{aligned}\mathrm{var}(\widehat{\boldsymbol{\gamma}}) &= (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}-(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}}'\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\\ &\quad-(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}}\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\\ &\quad+(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}}\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}}'\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\\ &= (\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}-(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}}\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1},\end{aligned}$$

$$\mathrm{var}(\widehat{\boldsymbol{\kappa}}) = \mathbf{Q_{21}}\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{12}} = -\mathbf{Q_{22}}$$

a kovarianční matice je rovna

$$\begin{aligned}\mathrm{cov}(\widehat{\boldsymbol{\gamma}},\widehat{\boldsymbol{\kappa}}) &= -(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{12}}+(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{11}}'\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{12}}\\ &= -(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q_{12}}.\end{aligned}$$

Tvar blokové matice $\mathbb{Q}$ s neznámými maticemi $\mathbf{Q_{ij}}$ můžeme určit pomocí Rohdeho věty [5].

Věta 1 (Rohdeho věta): Nechť

$$\mathbf{D} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{B}' & \mathbf{C}\end{pmatrix}$$

je pozitivně definitní matice. Potom

$$\mathbf{D}^{-1} = \begin{pmatrix}\mathbf{D}^{11} & \mathbf{D}^{12}\\ \mathbf{D}^{21} & \mathbf{D}^{22}\end{pmatrix},$$

kde

$$\begin{aligned}\mathbf{D}^{11} &= \mathbf{A}^{-1}+\mathbf{A}^{-1}\mathbf{B}(\mathbf{C}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{A}^{-1}\\ \mathbf{D}^{12} &= -\,\mathbf{A}^{-1}\mathbf{B}(\mathbf{C}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\\ \mathbf{D}^{21} &= -\,(\mathbf{C}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{B}'\mathbf{A}^{-1}\\ \mathbf{D}^{22} &= (\mathbf{C}-\mathbf{B}'\mathbf{A}^{-1}\mathbf{B})^{-1}.\end{aligned}$$

Potom v našem případě platí

$$\mathbf{Q_{11}} = [\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1} - [\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C}$$
$$\times (\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C})^{-1}\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}$$

$$\mathbf{Q_{12}} = -[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\,\mathbf{C}(-\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C})^{-1}$$

$$\mathbf{Q_{21}} = (\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C})^{-1}\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}$$

$$\mathbf{Q_{22}} = (\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C})^{-1}.$$

Uvažujeme-li regulární model

$$\boldsymbol{Y} = \boldsymbol{X\gamma} + \boldsymbol{\varepsilon},\ \ \boldsymbol{b} + \boldsymbol{B\gamma} + \boldsymbol{C\kappa} = \boldsymbol{o},\ var(\boldsymbol{\varepsilon}) = \sigma^2\mathbf{V},$$

kde V je známá pozitivně definitní matice a σ^2 je neznámý parametr. Nestranný odhad parametru σ^2 je daný vztahem

$$\hat{\sigma}^2 = \frac{\boldsymbol{e'V^{-1}e}}{n - k_1 - k_2 + q}\ , \tag{15}$$

kde $\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\widehat{\boldsymbol{\gamma}}}$ je reziduální vektor, n je počet pozorování, k_1 počet neznámých parametrů $\boldsymbol{\gamma}$, k_2 počet neznámých parametrů $\boldsymbol{\kappa}$, q počet podmínek.

Ukážeme, že uvedený odhad je skutečně nestranným odhadem σ^2. Nejdříve si spočítáme střední hodnotu kvadratické formy $\boldsymbol{e'V^{-1}e}$. Při výpočtu užijeme vlastnosti stopy matice [5]. Stopa matice je rovna součtu prvků na hlavní diagonále čtvercové matice: $Tr = \sum_{i=1}^{n} a_{ii}$.

Uvažujeme čtvercové matice *A* a *B* typu $n \times n$ a $c, d \in R^1$. Potom platí

a) $Tr(cA \pm dB) = cTr(A) \pm dTr(B)$,
b) $Tr(A') = Tr(A)$,
c) $Tr(AB) = Tr(BA)$,
d) $Tr(AA') = Tr(A'A)$,
e) $Tr(a) = a$.

Pravidla c) a d) platí i pro obdélníkové matice A typu $m \times n$ a B typu $n \times m$.

Potom

$$\boldsymbol{E}(\boldsymbol{e'V^{-1}e}) = \boldsymbol{E}[Tr(\boldsymbol{e'V^{-1}e})] = \boldsymbol{E}[Tr(\boldsymbol{V^{-1}ee'})] = Tr(\boldsymbol{V^{-1}}\boldsymbol{E}[\boldsymbol{ee'}]).$$

Protože $\boldsymbol{E}(\boldsymbol{e}) = 0$, platí $var(\boldsymbol{e}) = \boldsymbol{E}[\boldsymbol{ee'}]$, a proto můžeme psát

$$\begin{aligned}
\boldsymbol{Tr}(\boldsymbol{V^{-1}}\boldsymbol{E}[\boldsymbol{ee'}]) &= \boldsymbol{Tr}\left(\boldsymbol{V^{-1}}var(\boldsymbol{e})\right) \\
&= \sigma^2\boldsymbol{Tr}\{\boldsymbol{V^{-1}}[\boldsymbol{V} - \mathbf{X}(\mathbf{X'V^{-1}X})^{-1}\mathbf{X'} + \mathbf{X}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'Q_{11}B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{X'}]\} \\
&= \sigma^2\boldsymbol{Tr}\{\boldsymbol{I_n} - \boldsymbol{V^{-1}}\mathbf{X}(\mathbf{X'V^{-1}X})^{-1}\mathbf{X'} + \boldsymbol{V^{-1}}\mathbf{X}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}\{[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1} \\
&\quad -[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\mathbf{C}(\mathbf{C'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\mathbf{C})^{-1} \\
&\quad \times \mathbf{C'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\}\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{X'}\} \\
&= \sigma^2\boldsymbol{Tr}(\boldsymbol{I_n}) - \sigma^2\boldsymbol{Tr}[\mathbf{X'V^{-1}X}(\mathbf{X'V^{-1}X})^{-1}] \\
&\quad + \sigma^2\boldsymbol{Tr}\{\mathbf{X'V^{-1}X}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\} \\
&\quad - \sigma^2\boldsymbol{Tr}\{\mathbf{X'V^{-1}X}(\mathbf{X'\Sigma^{-1}X})^{-1}[\mathbf{B}(\mathbf{X'\Sigma^{-1}X})^{-1}\mathbf{B'}][\mathbf{B}(\mathbf{X'\Sigma^{-1}X})^{-1}\mathbf{B'}]^{-1} \\
&\quad \times \mathbf{C}(\mathbf{C'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\mathbf{C})^{-1}\mathbf{C'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\} \\
&= \sigma^2\boldsymbol{Tr}(\boldsymbol{I_n}) - \sigma^2\boldsymbol{Tr}\left(\boldsymbol{I_{k_1}}\right) + \sigma^2\boldsymbol{Tr}\{\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\} \\
&\quad - \sigma^2\boldsymbol{Tr}\{\mathbf{C'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\mathbf{C}(\mathbf{C'}[\mathbf{B}(\mathbf{X'V^{-1}X})^{-1}\mathbf{B'}]^{-1}\mathbf{C})^{-1}\} \\
&= \sigma^2(n - k_1) + \sigma^2\boldsymbol{Tr}\left(\boldsymbol{I_q}\right) - \sigma^2\boldsymbol{Tr}\left(\boldsymbol{I_{k_2}}\right) = \sigma^2(n - k_1 - k_2 + q).
\end{aligned}$$

Tedy:

$$\begin{aligned}
\boldsymbol{E}(\hat{\sigma}^2) = \boldsymbol{E}\left[\frac{\boldsymbol{e'V^{-1}e}}{n - k_1 - k_2 + q}\right] &= \frac{1}{n - k_1 - k_2 + q}\boldsymbol{E}(\boldsymbol{e'V^{-1}e}) \\
&= \frac{1}{n - k_1 - k_2 + q}(n - k_1 - k_2 + q)\sigma^2 = \sigma^2.
\end{aligned}$$

□

Za předpokladu normality náhodných chyb měření má odhad parametru σ^2 rozdělení

$$\hat{\sigma}^2 \sim \frac{\sigma^2}{n-k_1-k_2+q}\chi^2_{n-k_1-k_2+q}.$$

Hledáme rozdělení kvadratické formy $\boldsymbol{e}'\boldsymbol{V}^{-1}\boldsymbol{e}$. Reziduální vektor je dán

$$\boldsymbol{e} = \boldsymbol{Y} - \boldsymbol{X}\widehat{\boldsymbol{\gamma}} = \boldsymbol{Y} - \boldsymbol{X}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}[\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y} - \mathbf{B}'\mathbf{Q}_{11}(\mathbf{b} + \mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{Y})]$$

a rozdělení reziduálního vektoru je

$$\boldsymbol{e} \sim \mathbb{N}_{\boldsymbol{n}}(\mathbf{0}, \sigma^2[\boldsymbol{V} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}' + \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q}_{11}\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}']).$$

Protože matice $[\boldsymbol{V} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}' + \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}'\mathbf{Q}_{11}\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}']\mathbf{V}^{-1}$ je idempotentní, její hodnost je rovna stopě:

$$\begin{gathered}\boldsymbol{h}\{\boldsymbol{I_n} - \boldsymbol{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}' + \boldsymbol{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{B}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}' \\ -\boldsymbol{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C}(\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C})^{-1} \\ \times\, \mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\} =\end{gathered}$$

$$\begin{aligned}= {}& \boldsymbol{Tr}(\boldsymbol{I_n}) - \boldsymbol{Tr}(\boldsymbol{I_{k_1}}) + \boldsymbol{Tr}\{\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\} \\ & - \boldsymbol{Tr}\{\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C}(\mathbf{C}'[\mathbf{B}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{B}']^{-1}\mathbf{C})^{-1}\}\end{aligned}$$

$$= n - k_1 - k_2 + q.$$

Kvadratická forma $\boldsymbol{e}'\boldsymbol{V}^{-1}\boldsymbol{e}$ má rozdělení $\sigma^2\chi^2$ s = $n - k_1 - k_2 + q$ stupni volnosti. Tudíž

$$\hat{\sigma}^2 \sim \frac{\sigma^2}{n-k_1-k_2+q}\chi^2_{n-k_1-k_2+q}.$$

Lemma 1: Uvažujeme lineární statistický model

$$\boldsymbol{Y} = \boldsymbol{\gamma} + \varepsilon, \quad var(\varepsilon) = \sigma^2 \mathbf{I}_n\,,$$

kde $\boldsymbol{Y}$ je n-rozměrný vektor pozorování a $\boldsymbol{\gamma}$ je n-rozměrný neznámý vektor, tak že γ splňuje

$$\boldsymbol{b} + \mathbf{B}\boldsymbol{\gamma} + \mathbf{C}\boldsymbol{\kappa} = 0,$$

kde $\boldsymbol{b}$ je daný n-rozměrný vektor, $\mathbf{B}$ a $\mathbf{C}$ jsou dané matice řádu $q \times n$ a $q \times k$ a $\boldsymbol{\kappa}$ je neznámý k-rozměrný vektor. Hodnosti matic $\mathbf{B}$ a $\mathbf{C}$ jsou $\mathrm{h}(\mathbf{C}) = k$, $\mathrm{h}(\mathbf{B}, \mathbf{C}) = q$ a $k < q < n + k$. Potom BLUE (nejlepší nestranné lineární odhady) $\boldsymbol{\gamma}$ a $\boldsymbol{\kappa}$ jsou

$$\begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\boldsymbol{\kappa}} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n - \mathbf{B}'\mathbf{Q}_{11}\mathbf{B} \\ -\mathbf{Q}_{21}\mathbf{B} \end{pmatrix} \boldsymbol{Y} + \begin{pmatrix} -\mathbf{B}'\mathbf{Q}_{11} \\ -\mathbf{Q}_{21} \end{pmatrix} \boldsymbol{b}$$

a kovarianční matice tohoto odhadu je

$$var \begin{pmatrix} \hat{\boldsymbol{\gamma}} \\ \hat{\boldsymbol{\kappa}} \end{pmatrix} = \begin{pmatrix} \mathbf{I}_n - \mathbf{B}'\mathbf{Q}_{11}\mathbf{B} & -\mathbf{B}'\mathbf{Q}_{12} \\ -\mathbf{Q}_{21}\mathbf{B} & -\mathbf{Q}_{22} \end{pmatrix} \sigma^2,$$

kde

$$\begin{pmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{pmatrix} = \begin{pmatrix} \mathbf{BB}' & \mathbf{C} \\ \mathbf{C}' & \mathbf{0} \end{pmatrix}^{-1}.$$

1.3.2 Ortogonální regrese pomocí lineárních modelů

Předpokládejme $(n \times 2)$-rozměrnou datovou matici $(\boldsymbol{x}, \boldsymbol{y})$. Pokud bezchybná hodnota $\boldsymbol{x}$ je $\boldsymbol{\mu} = (\mu_1, \dots, \mu_n)'$ a bezchybná hodnota $\boldsymbol{y}$ je $\boldsymbol{\nu} = (\nu_1, \dots, \nu_n)'$, potom lineární vztah mezi těmito bezchybnými hodnotami je

$$\boldsymbol{\nu} = \beta_0 \mathbf{1}_n + \beta_1 \boldsymbol{\mu}. \tag{16}$$

Předpokládejme, že pozorování $\boldsymbol{x}$ a $\boldsymbol{y}$ jsou nezávislé, se směrodatnou odchylkou $\sigma > 0$. Odpovídající model může být vyjádřen v maticovém tvaru

$$\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} = \begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix} + \varepsilon, \quad var(\varepsilon) = \sigma^2 \mathbf{I}_{2n},$$

kde $\boldsymbol{\mu}, \boldsymbol{\nu}$ splňují vztah (16). Protože ortogonální regresní přímka (16) navíc obsahuje neznámé parametry β_0 a β_1, lze ji chápat jako nelineární podmínku typu II. Nelinearita je způsobena výrazem $\beta_1 \boldsymbol{\mu}$. Pokud jsou derivace druhého a vyšších řádů zanedbatelné, můžeme model (16) linearizovat použitím Taylorova rozvoje v bodech $\boldsymbol{\mu}^{(\mathbf{0})}, \boldsymbol{\nu}^{(\mathbf{0})}, \beta_0^{(0)}$ a $\beta_1^{(0)}$, které splňují systém rovnic (16). Takto získáme lineární model

$$\begin{pmatrix} \boldsymbol{x} - \boldsymbol{\mu}^{(\mathbf{0})} \\ \boldsymbol{y} - \boldsymbol{\nu}^{(\mathbf{0})} \end{pmatrix} = \begin{pmatrix} \Delta\boldsymbol{\mu} \\ \Delta\boldsymbol{\nu} \end{pmatrix} + \boldsymbol{\varepsilon}, \qquad \Delta\boldsymbol{\mu} = \boldsymbol{\mu} - \boldsymbol{\mu}^{(\mathbf{0})}, \qquad \Delta\boldsymbol{\nu} = \boldsymbol{\nu} - \boldsymbol{\nu}^{(\mathbf{0})},$$

kde $\Delta\mu$ a $\Delta\nu$ vyhovují vztahu

$$\Delta\beta_0 \mathbf{1}_{\boldsymbol{n}} + \Delta\beta_1 \boldsymbol{\mu}^{(\mathbf{0})} + \beta_1^{(0)}\, \Delta\boldsymbol{\mu} = \Delta\boldsymbol{\nu},$$

$\Delta\beta_0 = \beta_0 - {\beta_0}^{(0)}$ a $\Delta\beta_1 = \beta_1 - \beta_1^{(0)}$. Předchozí rovnici můžeme přepsat jako

$$\left(\beta_1^{(0)} \mathbf{I_n}, -\mathbf{I_n}\right) \begin{pmatrix} \Delta\boldsymbol{\mu} \\ \Delta\boldsymbol{\nu} \end{pmatrix} + \left(\mathbf{1}_{\boldsymbol{n}}, \boldsymbol{\mu}^{(\mathbf{0})}\right) \begin{pmatrix} \Delta\beta_0 \\ \Delta\beta_1 \end{pmatrix} = \mathbf{0}.$$

Využitím **Lemma 1**, substitucí

$$\boldsymbol{Y} \rightarrow \begin{pmatrix} \boldsymbol{x} - \boldsymbol{\mu}^{(\mathbf{0})} \\ \boldsymbol{y} - \boldsymbol{\nu}^{(0)} \end{pmatrix}, \quad \boldsymbol{\gamma} \rightarrow \begin{pmatrix} \Delta\boldsymbol{\mu} \\ \Delta\boldsymbol{\nu} \end{pmatrix}, \quad \boldsymbol{\kappa} \rightarrow \begin{pmatrix} \Delta\beta_0 \\ \Delta\beta_1 \end{pmatrix}$$

$$\boldsymbol{b} \rightarrow \mathbf{0}, \quad \boldsymbol{B} \rightarrow \left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right), \quad \boldsymbol{C} \rightarrow \left(\mathbf{1}_{\boldsymbol{n}}, \boldsymbol{\mu}^{(\mathbf{0})}\right),$$

a vztahů

$$\widehat{\boldsymbol{\mu}} = \widehat{\Delta\boldsymbol{\mu}} + \boldsymbol{\mu}^{(\mathbf{0})}, \quad \widehat{\boldsymbol{\nu}} = \widehat{\Delta\boldsymbol{\nu}} + \boldsymbol{\nu}^{(\mathbf{0})}, \quad \hat{\beta}_0 = \widehat{\Delta\beta}_0 + \beta_0^{(0)}, \quad \hat{\beta}_1 = \widehat{\Delta\beta}_1 + \beta_1^{(0)}$$

získáme lokálně nejlepší lineární nestranné odhady vektorů $\boldsymbol{\mu}\,, \boldsymbol{\nu}$ a parametrů β_0 a β_1 v linearizovaném modelu.

Vyjádříme si, že

$$\widehat{\boldsymbol{\gamma}} = (\mathbf{I}_{\boldsymbol{n}} - \mathbf{B}'\mathbf{Q_{11}}\mathbf{B})\boldsymbol{Y}.$$

Dosadíme za substituci

$$\begin{pmatrix} \widehat{\Delta\boldsymbol{\mu}} \\ \widehat{\Delta\boldsymbol{\nu}} \end{pmatrix} = \left[\mathbf{I}_{\mathbf{2}\boldsymbol{n}} - \begin{pmatrix} \beta_1^{(0)}\mathbf{I_n} \\ -\mathbf{I_n} \end{pmatrix} \mathbf{Q_{11}} \left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right)\right] \begin{pmatrix} \boldsymbol{x} - \boldsymbol{\mu}^{(\mathbf{0})} \\ \boldsymbol{y} - \boldsymbol{\nu}^{(0)} \end{pmatrix}.$$

$\mathbf{Q_{11}}$ vypočítáme jako:

$$\mathbf{Q_{11}} = (\mathbf{BB'})^{-\mathbf{1}} - (\mathbf{BB'})^{-\mathbf{1}}\mathbf{C}[\mathbf{C}'(\mathbf{BB'})^{-\mathbf{1}}\mathbf{C}]^{-\mathbf{1}}\mathbf{C}'(\mathbf{BB'})^{-\mathbf{1}},$$

kde

$$\mathbf{BB}' = \left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right)\begin{pmatrix} \beta_1^{(0)}\mathbf{I_n} \\ -\mathbf{I_n} \end{pmatrix} = \left[\beta_1^{(0)}\right]^2\mathbf{I_n} + \mathbf{I_n} = \mathbf{I_n}\left(\left[\beta_1^{(0)}\right]^2 + 1\right).$$

Potom

$$\mathbf{Q_{11}} = \frac{1}{\left[\beta_1^{(0)}\right]^2 + 1}\mathbf{I_n} - \frac{1}{\left[\beta_1^{(0)}\right]^2 + 1}\mathbf{I_n}\left(\mathbf{1}_{\boldsymbol{n}}, \boldsymbol{\mu}^{(\mathbf{0})}\right)\left[\begin{pmatrix} \mathbf{1}_{\boldsymbol{n}}{}' \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]' \end{pmatrix}\frac{1}{\left[\beta_1^{(0)}\right]^2 + 1}\mathbf{I_n}\left(\mathbf{1}_{\boldsymbol{n}}, \boldsymbol{\mu}^{(\mathbf{0})}\right)\right]^{-1}$$

$$\cdot\begin{pmatrix} \mathbf{1}_{\boldsymbol{n}}{}' \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]' \end{pmatrix}\frac{1}{\left[\beta_1^{(0)}\right]^2 + 1}\mathbf{I_n}$$

$$= \frac{1}{\left[\beta_1^{(0)}\right]^2+1}\mathbf{I_n} - \frac{1}{\left[\beta_1^{(0)}\right]^2+1}\left(\mathbf{1},\boldsymbol{\mu}^{(\mathbf{0})}\right)\left(\left[\beta_1^{(0)}\right]^2+1\right)\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})} \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\mathbf{1} & \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\boldsymbol{\mu}^{(\mathbf{0})} \end{pmatrix}^{-1}$$

$$\times \begin{pmatrix} \mathbf{1}' \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]' \end{pmatrix}$$

$$= \frac{1}{\left[\beta_1^{(0)}\right]^2+1}\left[\mathbf{I_n} - \left(\mathbf{1},\boldsymbol{\mu}^{(\mathbf{0})}\right)\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})} \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\mathbf{1} & \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\boldsymbol{\mu}^{(\mathbf{0})} \end{pmatrix}^{-1}\begin{pmatrix} \mathbf{1}' \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]' \end{pmatrix}\right].$$

Označíme

$$\mathbf{M}^{(\mathbf{0})} = \mathbf{I}_{\boldsymbol{n}} - \left(\mathbf{1},\boldsymbol{\mu}^{(\mathbf{0})}\right)\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})} \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\mathbf{1} & \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\boldsymbol{\mu}^{(\mathbf{0})} \end{pmatrix}^{-1}\begin{pmatrix} \mathbf{1}' \\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]' \end{pmatrix} \tag{17}$$

a dostaneme

$$\begin{pmatrix} \hat{\boldsymbol{\mu}} - \boldsymbol{\mu}^{(\mathbf{0})} \\ \hat{\boldsymbol{\nu}} - \boldsymbol{\nu}^{(\mathbf{0})} \end{pmatrix} = \begin{pmatrix} \boldsymbol{x} - \boldsymbol{\mu}^{(\mathbf{0})} \\ \boldsymbol{y} - \boldsymbol{\nu}^{(\mathbf{0})} \end{pmatrix} - \begin{pmatrix} \beta_1^{(0)}\mathbf{I_n} \\ -\mathbf{I_n} \end{pmatrix}\frac{1}{\left[\beta_1^{(0)}\right]^2+1}\mathbf{M}^{(\mathbf{0})}\left[\beta_1^{(0)}\left(\boldsymbol{x}-\boldsymbol{\mu}^{(\mathbf{0})}\right) - \boldsymbol{y} - \boldsymbol{\nu}^{(\mathbf{0})}\right].$$

Potom

$$\hat{\boldsymbol{\mu}} = \boldsymbol{x} + \frac{\beta_1^{(0)}}{\left[\beta_1^{(0)}\right]^2+1}\mathbf{M}^{(\mathbf{0})}\left[\boldsymbol{y} - \boldsymbol{\nu}^{(\mathbf{0})} - \beta_1^{(0)}\left(x - \boldsymbol{\mu}^{(\mathbf{0})}\right)\right], \tag{18}$$

a

$$\hat{\boldsymbol{\nu}} = \boldsymbol{y} - \frac{1}{\left[\beta_1^{(0)}\right]^2+1}\mathbf{M}^{(\mathbf{0})}\left[\boldsymbol{y} - \boldsymbol{\nu}^{(\mathbf{0})} - \beta_1^{(0)}\left(x - \boldsymbol{\mu}^{(\mathbf{0})}\right)\right]. \tag{19}$$

Vyjádříme si, že

$$\hat{\boldsymbol{\kappa}} = (-\mathbf{Q_{21}B})\boldsymbol{Y}.$$

Dosadíme za substituci

$$\begin{pmatrix} \widehat{\Delta\beta_0} \\ \widehat{\Delta\beta_1} \end{pmatrix} = -\mathbf{Q_{21}}\left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right)\begin{pmatrix} \boldsymbol{x} - \boldsymbol{\mu}^{(\mathbf{0})} \\ \boldsymbol{y} - \boldsymbol{\nu}^{(\mathbf{0})} \end{pmatrix},$$

kde

$$\mathbf{Q_{21}} = [\mathbf{C}'(\mathbf{BB'})^{-1}\mathbf{C}]^{-1}\mathbf{C}'(\mathbf{BB'})^{-1}$$

$$= \left[\begin{pmatrix} \mathbf{1}_n{}' \\ [\boldsymbol{\mu}^{(0)}]' \end{pmatrix} \frac{1}{\left[\beta_1^{(0)}\right]^2 + 1} \mathbf{I_n}\left(\mathbf{1}_n, \boldsymbol{\mu}^{(0)}\right)\right]^{-1} \begin{pmatrix} \mathbf{1}_n{}' \\ [\boldsymbol{\mu}^{(0)}]' \end{pmatrix} \frac{1}{\left[\beta_1^{(0)}\right]^2 + 1} \mathbf{I_n}$$

$$= \left(\left[\beta_1^{(0)}\right]^2 + 1\right) \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(0)} \\ [\boldsymbol{\mu}^{(0)}]'\mathbf{1} & [\boldsymbol{\mu}^{(0)}]'\boldsymbol{\mu}^{(0)} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{1}' \\ [\boldsymbol{\mu}^{(0)}]' \end{pmatrix} \frac{1}{\left[\beta_1^{(0)}\right]^2 + 1}$$

$$= \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(0)} \\ [\boldsymbol{\mu}^{(0)}]'\mathbf{1} & [\boldsymbol{\mu}^{(0)}]'\boldsymbol{\mu}^{(0)} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{1}' \\ [\boldsymbol{\mu}^{(0)}]' \end{pmatrix}.$$

Potom

$$\begin{pmatrix} \hat{\beta}_0 - \beta_0^{(0)} \\ \hat{\beta}_1 - \beta_1^{(0)} \end{pmatrix} = -\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(0)} \\ [\boldsymbol{\mu}^{(0)}]'\mathbf{1} & [\boldsymbol{\mu}^{(0)}]'\boldsymbol{\mu}^{(0)} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{1}' \\ [\boldsymbol{\mu}^{(0)}]' \end{pmatrix} \left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right) \begin{pmatrix} \boldsymbol{x} - \boldsymbol{\mu}^{(0)} \\ \boldsymbol{y} - \boldsymbol{\nu}^{(0)} \end{pmatrix}$$

$$\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} \beta_0^{(0)} \\ \beta_1^{(0)} \end{pmatrix} + \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(0)} \\ [\boldsymbol{\mu}^{(0)}]'\mathbf{1} & [\boldsymbol{\mu}^{(0)}]'\boldsymbol{\mu}^{(0)} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{1}'\left[\boldsymbol{y} - \boldsymbol{\nu}^{(0)} - \beta_1^{(0)}\left(x - \boldsymbol{\mu}^{(0)}\right)\right] \\ [\boldsymbol{\mu}^{(0)}]'\left[\boldsymbol{y} - \boldsymbol{\nu}^{(0)} - \beta_1^{(0)}\left(x - \boldsymbol{\mu}^{(0)}\right)\right] \end{pmatrix}. \quad (20)$$

Varianční matice odhadů $\widehat{\boldsymbol{\mu}}$ a $\widehat{\boldsymbol{\nu}}$ určíme z varianční matice $\widehat{\boldsymbol{\gamma}}$:

$$var[\widehat{\boldsymbol{\gamma}}] = (\mathbf{I}_n - \mathbf{B}'\mathbf{Q_{11}}\mathbf{B})\sigma^2.$$

Dosadíme za substituci

$$var\begin{pmatrix} \widehat{\Delta\boldsymbol{\mu}} \\ \widehat{\Delta\boldsymbol{\nu}} \end{pmatrix} = \sigma^2\mathbf{I}_{2n} - \begin{pmatrix} \beta_1^{(0)}\mathbf{I_n} \\ -\mathbf{I_n} \end{pmatrix} \frac{1}{\left[\beta_1^{(0)}\right]^2 + 1} \mathbf{M}^{(0)} \left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right)\sigma^2.$$

Potom

$$var[\widehat{\boldsymbol{\mu}}] = \sigma^2\mathbf{I}_n - \frac{\left[\beta_1^{(0)}\right]^2\sigma^2}{\left[\beta_1^{(0)}\right]^2 + 1}\mathbf{M}^{(0)} \quad (21)$$

a

$$var[\hat{\boldsymbol{\nu}}] = \sigma^2 \mathbf{I}_{\boldsymbol{n}} - \frac{\sigma^2}{\left[\beta_1^{(0)}\right]^2 + 1} \mathbf{M}^{(\mathbf{0})}. \tag{22}$$

Kovarianční matice odhadů $\hat{\boldsymbol{\mu}}$ a $\hat{\boldsymbol{\nu}}$ je dána

$$cov[\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\nu}}] = \frac{\beta_1^{(0)} \sigma^2}{\left[\beta_1^{(0)}\right]^2 + 1} \mathbf{M}^{(\mathbf{0})}. \tag{23}$$

Varianční matice odhadu $\left(\hat{\beta}_{0,}\hat{\beta}_1\right)'$ určíme z varianční matice $\hat{\boldsymbol{\kappa}}$:

$$var[\hat{\boldsymbol{\kappa}}] = -\mathbf{Q}_{\mathbf{22}} \sigma^2,$$

kde

$$\mathbf{Q}_{\mathbf{22}} = -[\mathbf{C}'(\mathbf{B}\mathbf{B}')^{-\mathbf{1}}\mathbf{C}]^{-\mathbf{1}}.$$

Dosadíme za substituci

$$var\begin{pmatrix} \widehat{\Delta\beta_0} \\ \widehat{\Delta\beta_1} \end{pmatrix} = [\mathbf{C}'(\mathbf{B}\mathbf{B}')^{-\mathbf{1}}\mathbf{C}]^{-\mathbf{1}} \sigma^2$$

$$var\begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} = \sigma^2 \left(\left[\beta_1^{(0)}\right]^2 + 1 \right) \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})} \\ [\boldsymbol{\mu}^{(\mathbf{0})}]'\mathbf{1} & [\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} \end{pmatrix}^{-1}$$

a kovarianční matice odhadů $(\hat{\boldsymbol{\mu}}', \hat{\boldsymbol{\nu}}')'$ a $\left(\hat{\beta}_{0,}\hat{\beta}_1\right)'$ určíme z kovarianční matice $(\hat{\boldsymbol{\gamma}}', \hat{\boldsymbol{\kappa}}')'$

$$cov(\hat{\boldsymbol{\gamma}}', \hat{\boldsymbol{\kappa}}')' = -\mathbf{B}'\mathbf{Q}_{\mathbf{12}},$$

kde

$$\mathbf{Q}_{\mathbf{12}} = (\mathbf{B}\mathbf{B}')^{-\mathbf{1}}\mathbf{C}[\mathbf{C}'(\mathbf{B}\mathbf{B}')^{-\mathbf{1}}\mathbf{C}]^{-\mathbf{1}}.$$

Dosadíme za substituci

$$cov\left[\begin{pmatrix}\hat{\beta}_0\\ \hat{\beta}_1\end{pmatrix},\begin{pmatrix}\hat{\boldsymbol{\mu}}\\ \hat{\boldsymbol{\nu}}\end{pmatrix}\right] = -\sigma^2\left(\beta_1^{(0)}\mathbf{I_n}, -\mathbf{I_n}\right)\frac{1}{\left[\beta_1^{(0)}\right]^2+1}\mathbf{I_n}\left(\mathbf{1}_{\boldsymbol{n}}, \boldsymbol{\mu}^{(\mathbf{0})}\right)$$

$$\times\left(\left[\beta_1^{(0)}\right]^2+1\right)\mathbf{I_n}\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}\\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\mathbf{1} & \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\boldsymbol{\mu}^{(\mathbf{0})}\end{pmatrix}^{-1}$$

$$= -\sigma^2\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}\\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\mathbf{1} & \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\boldsymbol{\mu}^{(\mathbf{0})}\end{pmatrix}^{-1}\begin{pmatrix}\beta_1^{(0)}\mathbf{1}' & -\mathbf{1}'\\ \beta_1^{(0)}\left[\boldsymbol{\mu}^{(\mathbf{0})}\right]' & -\left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\end{pmatrix}.$$

Odhady dané rovnicemi (18)-(20) závisí na přibližných hodnotách neznámých parametrů $\beta_0, \beta_1, \boldsymbol{\mu}$ a $\boldsymbol{\nu}$, proto je třeba je řešit iteračním způsobem.

Rozptyl σ^2 je obvykle neznámý a je třeba ho odhadnout. Vztah pro odhad $\hat{\sigma}^2$ můžeme odvodit pomocí vzorce (15), kde za $\boldsymbol{e}$ dosadíme $\begin{pmatrix}\boldsymbol{x}-\hat{\boldsymbol{\mu}}\\ \boldsymbol{y}-\hat{\boldsymbol{\nu}}\end{pmatrix}$, za $\boldsymbol{V}^{-\mathbf{1}}$ dosadíme matici $\mathbf{I_{2n}} = \begin{pmatrix}\mathrm{I_n} & 0\\ 0 & \mathrm{I_n}\end{pmatrix}$ typu $2n \times 2n$, počet pozorování je $2n$, $k_1 = 2n$ (neznámé parametry jsou $\mu_1, \dots, \mu_n$ a $\nu_1, \dots, \nu_n$), $k_2 = 2$ (neznámé parametry jsou β_0, β_1) a $q = n$, protože máme n podmínek. Potom je odhad $\hat{\sigma}^2$ dán:

$$\hat{\sigma}^2 = \frac{\left((\boldsymbol{x}-\hat{\boldsymbol{\mu}})', (\boldsymbol{y}-\hat{\boldsymbol{\nu}})'\right)'\begin{pmatrix}\mathrm{I_n} & 0\\ 0 & \mathrm{I_n}\end{pmatrix}\begin{pmatrix}\boldsymbol{x}-\hat{\boldsymbol{\mu}}\\ \boldsymbol{y}-\hat{\boldsymbol{\nu}}\end{pmatrix}}{2n-2n-2+n}$$

$$= \frac{(\boldsymbol{x}-\hat{\boldsymbol{\mu}})'(\boldsymbol{x}-\hat{\boldsymbol{\mu}}) + (\boldsymbol{y}-\hat{\boldsymbol{\nu}})'(\boldsymbol{y}-\hat{\boldsymbol{\nu}})}{n-2}. \tag{24}$$

Tento odhad je nestranný a nezávislý na odhadech $\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\nu}}, \hat{\beta}_0$ a $\hat{\beta}_1$. Takže odhad může být použit jako aproximace pro odhad varianční matice $\hat{\beta}_0$ a $\hat{\beta}_1$:

$$\widehat{var}\left[\hat{\boldsymbol{\beta}}\right] = \hat{\sigma}^2\left(\left[\beta_1^{(0)}\right]^2+1\right)\begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}\\ \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\mathbf{1} & \left[\boldsymbol{\mu}^{(\mathbf{0})}\right]'\boldsymbol{\mu}^{(\mathbf{0})}\end{pmatrix}^{-1}.$$

Iterační algoritmus pro odhad ortogonální regresní přímky lze popsat ve čtyřech hlavních krocích:

1. krok: určíme počáteční hodnoty odhadu koeficientů $\beta_0^{(0)}, \beta_1^{(0)}$ regresní přímky a hodnoty $\boldsymbol{\mu}^{(\mathbf{0})}, \boldsymbol{\nu}^{(\mathbf{0})}$. Například:

$$\beta_0^{(0)} = \frac{x_j y_i - x_i y_j}{x_j - x_i}, \qquad \beta_1^{(0)} = \frac{y_j - y_i}{x_j - x_i}, \tag{25}$$

kde $x_i = min\{x_k : k = 1, \dots, n\}$, $x_j = max\{x_k : k = 1, \dots, n\}$ a y_i, y_j jsou odpovídající souřadnice y. Potom položíme $\boldsymbol{\mu}^{(\mathbf{0})} = \boldsymbol{x}$ a hodnotu $\boldsymbol{\nu}^{(\mathbf{0})}$ získáme z rovnice

$$\boldsymbol{\nu}^{(\mathbf{0})} = \beta_0^{(0)} \mathbf{1}_n + \beta_1^{(0)} \boldsymbol{\mu}^{(\mathbf{0})}.$$

2. krok: spočítáme odhady $\hat{\beta}_0$,$\hat{\beta}_1$, $\hat{\boldsymbol{\mu}}$ a $\hat{\boldsymbol{\nu}}$ pomocí rovnic (17)-(20).

3. krok: aktualizujeme počáteční hodnoty podle schématu

$$\boldsymbol{\nu}^{(\mathbf{0})} = \hat{\boldsymbol{\nu}} + \left(\hat{\beta}_1 - \beta_1^{(0)}\right)\left(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}^{(\mathbf{0})}\right), \quad \boldsymbol{\mu}^{(\mathbf{0})} = \hat{\boldsymbol{\mu}}, \quad \beta_0^{(0)} = \hat{\beta}_0, \quad \beta_1^{(0)} = \hat{\beta}_1.$$

4. krok: opakujeme kroky 2 a 3 dokud odhady konvergují, tj. dokud změny v odhadech každé iterace jsou menší než nějaká předem stanovená tolerance.

Odhady získané z tohoto iteračního algoritmu konvergují obvykle velmi rychle a také zachovávají předepsanou podmínku (1). Lineární modely tak představují alternativní metodu pro ortogonální regresi. Řešení je pouze přibližné z důvodu linearizace modelu.

1.4 Statistické inference

Statistické inference většinou provádíme pro směrnici ortogonální regresní přímky. Uvažujeme náhodný vektor $(\boldsymbol{x}', \boldsymbol{y}')'$ s normálním rozdělením. Za tohoto předpokladu je nejlepší nestranný lineární odhad parametru β také normálně rozdělený. Označme $\beta_1 = \operatorname{tg} \theta$

a $\hat{\beta}_1 = \text{tg}\,\hat{\theta}$. V článku [11] je navržen $100(1-\alpha)\%$ interval spolehlivosti pro parametr β_1 jako interval s mezemi $\text{tg}(\hat{\theta} - \Phi_1)$ a $\text{tg}(\hat{\theta} + \Phi_1)$, kde Φ_1 je definováno jako

$$\Phi_1 = \frac{1}{2}\arcsin\sqrt{\frac{4t_{n-2}^2(1-\alpha/2)\left(s_x^2 s_y^2 - s_{xy}^2\right)}{(n-2)\left[\left(s_x^2 - s_y^2\right)^2 + 4s_{xy}^2\right]}} \tag{26}$$

Zde, $t_{n-2}(1-\alpha/2)$ představuje $(1-\alpha/2)$-kvantil Studentova rozdělení o $(n-2)$ stupních volnosti.

Použitím metody hlavních komponent [10] lze odvodit jiné meze intervalu spolehlivosti pro β_1 ve tvaru $\text{tg}(\hat{\theta} - \Phi_2)$ a $\text{tg}(\hat{\theta} + \Phi_2)$, kde Φ_2 je

$$\Phi_2 = \arcsin\sqrt{\frac{\chi_1^2(1-\alpha)}{(n-1)\left[\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} - 2\right]}} \tag{27}$$

kde $\lambda_1 > \lambda_2$ jsou vlastní čísla výběrové varianční matice náhodného vektoru (x, y) a $\chi_1^2(1-\alpha)$ představuje $(1-\alpha)$-kvantil χ^2- rozdělení s jedním stupněm volnosti.

Podívejme se nyní podrobněji, jak vypadají vlastní čísla λ_1, λ_2 výběrové kovarianční matice (x, y), která má tvar

$$\begin{pmatrix} s_x^2 & s_{xy} \\ s_{yx} & s_y^2 \end{pmatrix}.$$

Vlastní čísla spočítáme jako $|A - \lambda I| = 0$. Výběrová kovarianční matice je symetrická a platí $s_{xy} = s_{yx}$. Potom

$$\begin{vmatrix} s_x^2 - \lambda & s_{xy} \\ s_{yx} & s_y^2 - \lambda \end{vmatrix} = (s_x^2 - \lambda)\left(s_y^2 - \lambda\right) - s_{xy}^2 \text{ položíme rovno 0.}$$

Výraz upravíme pro výpočet kořenů kvadratické rovnice:

$$\lambda^2 - \lambda\left(s_x^2 - s_y^2\right) + \left(s_x^2 s_y^2 - s_{xy}^2\right) = 0.$$

Potom vlastní čísla výběrové kovarianční matice (x, y) jsou:

$$\lambda_1 = \frac{s_x^2 + s_y^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}}{2}$$
$$\lambda_2 = \frac{s_x^2 + s_y^2 - \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}}{2}.$$

Platí, že $\lambda_1 > \lambda_2$. Takže největším vlastním číslem výběrové kovarianční matice je λ_1. Ke každému vlastnímu číslu λ_i, existuje alespoň jedno nenulové řešení soustavy $\boldsymbol{Az} = \lambda_i \boldsymbol{z}$. Řešení $\boldsymbol{z}$ nazveme vlastním vektorem:

$$\begin{pmatrix} s_x^2 & s_{xy} \\ s_{yx} & s_y^2 \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \frac{s_x^2 + s_y^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}}{2} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}.$$

Získáme soustavu dvou rovnic o dvou neznámých z_1 a z_2:

$$s_x^2 z_1 + s_{xy} z_2 = \frac{s_x^2 + s_y^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}}{2} z_1,$$
$$s_{xy} z_1 + s_y^2 z_2 = \frac{s_x^2 + s_y^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}}{2} z_2.$$

Upravíme:

$$2 s_{xy} z_2 = \left(s_y^2 - s_x^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}\right) z_1,$$

$$2 s_y^2 z_2 = \left(s_x^2 + s_y^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)} - s_{xy}\right) z_1.$$

Z první rovnice si vyjádříme z_2:

$$s_{xy}z_2 = \frac{\left(s_y^2 - s_x^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}\right)}{2} z_1.$$

Vlastní vektor (první hlavní komponenta) odpovídající vlastnímu číslu λ_1 je

$$z_1 = \left(s_{xy}\,, \frac{\left(s_y^2 - s_x^2 + \sqrt{\left(s_x^2 + s_y^2\right)^2 - 4\left(s_x^2 s_y^2 - s_{xy}^2\right)}\right)}{2} \right).$$

Vidíme, že u vlastního čísla λ_1 dochází ke shodě se vztahem (4).

V případě malých výběrů [10] použijeme Φ_3 místo Φ_2, kde

$$\Phi_3 = \arcsin\sqrt{\frac{4F_{1,n-2}(1-\alpha)}{(n-2)\left[\frac{\lambda_1}{\lambda_2} + \frac{\lambda_2}{\lambda_1} - 2\right]}}\,. \tag{28}$$

Zde, $F_{1,n-2}(1-\alpha)$ představuje $(1-\alpha)$-kvantil Fisherova rozdělení o $(1, n-2)$ stupních volnosti.

Chceme-li testovat hypotézu, že $\beta_1 = \theta = 0$, můžeme použít statistiku, která je odvozena v [11]

$$T = \sqrt{(n-2)\sin^2\left(2\hat{\theta}\right)\frac{\frac{1}{4}\left(s_x^2 - s_y^2\right)^2 + s_{xy}^2}{\left(s_x^2 s_y^2 - s_{xy}^2\right)}},$$

která má za platnosti nulové hypotézy Studentovo rozdělení o $(n-2)$ stupních volnosti.

Pro případ malých výběrů je v [9] navržena statistika

$$T = \frac{(n-2)r^2}{1-r^2},$$

kde

$$r^2 = \frac{(\lambda_1 - \lambda_2)^2}{(\lambda_1 + \lambda_2)^2},$$

která má za platnosti nulové hypotézy Fisherovo rozdělení o $(1, n-2)$ stupních volnosti.

Dále předpokládáme, že σ^2 je neznámé. Můžeme ho odhadnout ze vztahu (24) nezávisle na $\widehat{\boldsymbol{\beta}}$**.** Podle normality je rozdělení odhadu $\hat{\sigma}^2$ dáno:

$$\hat{\sigma}^2 \sim \frac{\sigma^2}{n-2}\chi^2_{n-2}.$$

Pomocí lineárních modelů můžeme použít testovací statistika pro ověření hypotézy $H_0\colon \beta_0 = 0$ proti $H_1\colon \beta_0 \neq 0$ ve tvaru

$$T_1 = \frac{\dfrac{\hat{\beta}_0 - \beta_0}{\sqrt{var\hat{\beta}_0}}}{\dfrac{\dfrac{\hat{\sigma}\sqrt{n-2}}{\sigma}}{\sqrt{n-2}}} = \frac{\dfrac{\hat{\beta}_0}{\sqrt{\dfrac{\sigma^2\left(\left[\beta_1^{(0)}\right]^2 + 1\right)[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})}}{n[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} - [\mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}]^2}}}}{\dfrac{\hat{\sigma}}{\sigma}}$$

$$= \frac{\hat{\beta}_0\sqrt{n[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} - [\mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}]^2}}{\hat{\sigma}\sqrt{\left(\left[\beta_1^{(0)}\right]^2 + 1\right)[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})}}},$$

která má za platnosti nulové hypotézy Studentovo rozdělení o $(n-2)$ stupni volnosti. Nulovou hypotézu zamítneme, pokud $|t_1| \geq t_{n-2}(1 - \alpha/2)$ (t_1 je realizace testovací statistiky T_1), nebo ekvivalentně, pokud p-hodnota $\leq \alpha$. $100(1-\alpha)\%$ interval spolehlivosti pro parametr β_0 je tvaru

$$\hat{\beta}_0 \pm \hat{\sigma} t_{n-2}(1-\alpha/2) \frac{\sqrt{\left(\left[\beta_1^{(0)}\right]^2 + 1\right)[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})}}}{\sqrt{n[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} - [\mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}]^2}}. \tag{29}$$

Obdobně vyjádříme $H_0: \beta_1 = 0$ jako

$$T_2 = \frac{\frac{\hat{\beta}_1 - \beta_1}{\sqrt{var\hat{\beta}_1}}}{\frac{\frac{\hat{\sigma}\sqrt{n-2}}{\sigma}}{\sqrt{n-2}}} = \frac{\frac{\hat{\beta}_1}{\sqrt{\frac{\sigma^2 n\left(\left[\beta_1^{(0)}\right]^2 + 1\right)}{n[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} - [\mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}]^2}}}}{\frac{\hat{\sigma}}{\sigma}}$$

$$= \frac{\hat{\beta}_1\sqrt{n[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} - [\mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}]^2}}{\hat{\sigma}\sqrt{n\left(\left[\beta_1^{(0)}\right]^2 + 1\right)}}$$

a T_2 má za platnosti nulové hypotézy Studentovo rozdělení o $(n-2)$ stupni volnosti. Tudíž meze $100(1-\alpha)\%$ intervalu spolehlivosti pro parametr β_1 jsou

$$\hat{\beta}_1 \pm \hat{\sigma} t_{n-2}(1-\alpha/2) \frac{\sqrt{n\left(\left[\beta_1^{(0)}\right]^2 + 1\right)}}{\sqrt{n[\boldsymbol{\mu}^{(\mathbf{0})}]'\boldsymbol{\mu}^{(\mathbf{0})} - [\mathbf{1}'\boldsymbol{\mu}^{(\mathbf{0})}]^2}}. \tag{30}$$

Na závěr uvažujeme nulovou hypotézu $H_0: \boldsymbol{\beta} = \mathbf{0}$ proti $H_1: \boldsymbol{\beta} \neq \mathbf{0}$. K ověření použijeme statistiku

$$F = \frac{\dfrac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{2 \cdot var\widehat{\boldsymbol{\beta}}}}{\dfrac{\dfrac{\hat{\sigma}^2(n-2)}{\sigma^2(n-2)}}{}} = \frac{\dfrac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{2 \cdot \sigma^2 \left(\left[\beta_1^{(0)}\right]^2 + 1\right) \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(0)} \\ [\boldsymbol{\mu}^{(0)}]'\mathbf{1} & [\boldsymbol{\mu}^{(0)}]'\boldsymbol{\mu}^{(0)} \end{pmatrix}^{-1}}}{\dfrac{\hat{\sigma}^2}{\sigma^2}}$$

$$= \frac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{2 \cdot \hat{\sigma}^2 \left(\left[\beta_1^{(0)}\right]^2 + 1\right) \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}^{(0)} \\ [\boldsymbol{\mu}^{(0)}]'\mathbf{1} & [\boldsymbol{\mu}^{(0)}]'\boldsymbol{\mu}^{(0)} \end{pmatrix}^{-1}}$$

$$= \frac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})}{2 \cdot \widehat{var}(\widehat{\boldsymbol{\beta}})} = \frac{1}{2}\widehat{\boldsymbol{\beta}}'\left[\widehat{var}(\widehat{\boldsymbol{\beta}})\right]^{-1}\widehat{\boldsymbol{\beta}},$$

která má za platnosti nulové hypotézy Fisherovo rozdělení o $(2, n-2)$ stupních volnosti. Nulovou hypotézu zamítáme, pokud $f \geq F_{2,n-2}(1-\alpha)$ (f je realizace testovací statistiky F), nebo ekvivalentně, pokud p-hodnota $\leq \alpha$.

Testovací statistiky T_1, T_2, F a intervaly spolehlivosti pro β_0, β_1 závisí na skutečných hodnotách $\beta_1^{(0)}$ a $\boldsymbol{\mu}^{(\mathbf{0})}$. V praxi se skutečné hodnoty nahrazují přibližnými hodnotami, např. odhadnutými hodnotami.

Pás spolehlivosti pro regresní přímku $\nu_0 = \beta_0 + \beta_1\mu_0$ je systém intervalů spolehlivosti pro $\beta_0 + \beta_1\mu_0$ v pevně zvoleném bodě μ_0, nebo ekvivalentně systém intervalů spolehlivosti pro ν_0 v pevně zvoleném bodě μ_0. Pro konstrukci intervalu spolehlivosti pro $\beta_0 + \beta_1\mu_0$ použijeme statistiku

$$T = \frac{\hat{\beta}_0 + \hat{\beta}_1\mu_0 - \beta_0 - \beta_1\mu_0}{\sqrt{(1, \mu_0)\widehat{var}(\widehat{\boldsymbol{\beta}})\begin{pmatrix} 1 \\ \mu_0 \end{pmatrix}}},$$

která má Studentovo rozdělení o $(n-2)$ stupních volnosti. Tedy, $100(1-\alpha)\%$-interval spolehlivosti pro $\beta_0 + \beta_1\mu_0$ v pevně zvoleném bodě μ_0 je tvaru

$$\hat{\beta}_0 + \hat{\beta}_1\mu_0 \pm \sqrt{(1,\mu_0)\widehat{var}(\widehat{\boldsymbol{\beta}})\begin{pmatrix}1\\ \mu_0\end{pmatrix}}\, t_{n-2}(1-\alpha/2). \qquad (31)$$

Daný interval spolehlivosti platí pouze pro jedno konkrétně zvolené μ_0. Pro intervaly zahrnující všechny možné hodnoty μ_0 je nutné vytvořit sdružené intervaly pro $\beta_0 + \beta_1\mu_0$. Ty jsou založeny na statistice

$$F = \frac{1}{2}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})'\left[\widehat{var}(\widehat{\boldsymbol{\beta}})\right]^{-1}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}),$$

která má Fisherovo rozdělení o $(2, n-2)$ stupních volnosti. Odtud dostaneme pravděpodobnostní rovnici

$$P\left\{\frac{1}{2}\widehat{\boldsymbol{\beta}}'\left[\widehat{var}(\widehat{\boldsymbol{\beta}})\right]^{-1}\widehat{\boldsymbol{\beta}} \le F_{2,n-2}(1-\alpha)\right\} = 1-\alpha. \qquad (32)$$

Dále budeme potřebovat Schefféého větu [6]: Nechť T je pozitivně definitní matice typu $n \times n$. Potom pro libovolný pevně zvolený vektor $v \in \mathbb{R}^n$ a libovolné kladné číslo c platí ekvivalence

$$\forall h \in \mathbb{R}^n \colon |\boldsymbol{v}'\boldsymbol{h}| \le c\sqrt{\boldsymbol{h}'\boldsymbol{T}\boldsymbol{h}} \iff \sqrt{\boldsymbol{v}'\boldsymbol{T}^{-1}\boldsymbol{v}} \le c.$$

Sdružené intervaly spolehlivosti konstruujeme tak, že aplikujeme Schefféého větu na pravděpodobnostní rovnici (32). Pomocí substituce

$$\boldsymbol{h} = \boldsymbol{a}, \qquad \boldsymbol{T} = \widehat{Var}(\widehat{\boldsymbol{\beta}})^{-1}, \qquad \boldsymbol{v} = \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}, \qquad c^2 = F_{2,n-2}(1-\alpha)$$

můžeme pravděpodobnostní rovnici (32) přepsat do tvaru

$$P\left\{\forall \boldsymbol{a} \in \mathbb{R}^2 \colon \left|\boldsymbol{a}'(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})\right| \le \sqrt{2F_{2,n-2}(1-\alpha)}\sqrt{\boldsymbol{a}'\widehat{var}(\widehat{\boldsymbol{\beta}})\boldsymbol{a}}\right\} = 1-\alpha.$$

Použitím $\boldsymbol{a} = (1, \mu_0)'$ pro všechna μ_0 získáme

$$P\left\{\forall \mu_0 \in \mathbb{R}^1 \colon \left|\hat{\beta}_0 + \hat{\beta}_1\mu_0 - \beta_0 - \beta_1\mu_0\right| \leq \sqrt{2F_{2,n-2}(1-\alpha)}\sqrt{(1, \mu_0)\widehat{var}\left(\widehat{\boldsymbol{\beta}}\right)\binom{1}{\mu_0}}\right\} \geq 1 - \alpha,$$

tedy meze sdruženého $100(1-\alpha)\%$ intervalu spolehlivosti pro $\beta_0 + \beta_1\mu_0$ v bodě μ_0 jsou dány vztahy

$$\hat{\beta}_0 + \hat{\beta}_1\mu_0 \pm \sqrt{(1, \mu_0)\widehat{var}\left(\widehat{\boldsymbol{\beta}}\right)\binom{1}{\mu_0}}\sqrt{2F_{2,n-2}(1-\alpha)}.$$

Šířka pásu spolehlivosti charakterizuje přesnost stanovení regresní přímky. Pás je nejužší v bodě $\mu_0 = \bar{\mu}$, šířka pásu se zvyšuje, když absolutní hodnota $|\mu_0 - \bar{\mu}|$ roste. Šířka intervalu spolehlivosti se snižuje v případě, že se zvyšuje velikost výběru. Sdružené intervaly spolehlivosti jsou širší než intervaly spolehlivosti konstruované pro každý bod μ_0 jednotlivě.

2 Modelování polynomiálního vztahu

Polynomiální vztah mezi proměnnými budeme modelovat pomocí polynomiální regrese, která bude rozebrána v následující podkapitole. Dále se budeme zabývat odhadem parametrů pomocí upravené metody nejmenších čtverců a modifikací upravené metody nejmenších čtverců. V poslední podkapitole si ukážeme situaci, kdy je známá varianční matice $\mathbf{\Omega}$ a neznámý jednotkový rozptyl. Při tvorbě této kapitoly byly použity zejména tyto zdroje [8, 14].

2.1 Polynomiální regrese

Polynomiální funkční vztah je nejpřirozenějším rozšířením lineárního modelu na model nelineární a je dán rovnicemi

$$\begin{gathered} \nu_i = \beta_0 + \beta_1\mu_i + \beta_2\mu_i^2 + \cdots + \beta_k\mu_i^k = \boldsymbol{\varsigma_i}'\boldsymbol{\beta}, \\ y_i = \nu_i + \varepsilon_{2i}, \qquad x_i = \mu_i + \varepsilon_{1i}\,, \end{gathered} \tag{33}$$

kde $\boldsymbol{\varsigma}_{\boldsymbol{i}}' = \left(1, \mu_i, \mu_i^2, \dots, \mu_i^k\right)$, $i = 1, .., n$ a $\boldsymbol{\beta} = (\beta_0, \beta_1, \dots, \beta_k)'$. Předpokládáme, že náhodné chyby $(\varepsilon_{1i}, \varepsilon_{2i}), i = 1, \dots, n$ jsou nezávislé a stejně rozdělené dvojce náhodných veličin s nulovou střední hodnotou a s kovarianční maticí

$$\mathbf{\Omega} = \begin{pmatrix} \sigma_{\varepsilon_{1i}}^2 & \sigma_{\varepsilon_{1i}\varepsilon_{2i}} \\ \sigma_{\varepsilon_{1i}\varepsilon_{2i}} & \sigma_{\varepsilon_{2i}}^2 \end{pmatrix}.$$

Dále předpokládáme, že

$$(\varepsilon_{1i}, \varepsilon_{2i}) \sim N(0, \mathbf{\Omega}). \tag{34}$$

Aby parametry $\beta_0, \dots, \beta_k$ byly odhadnutelné je nutná znalost $\mathbf{\Omega}$ nebo alespoň některých prvků $\mathbf{\Omega}$.

Předpokládáme, že všechny momenty $E({\varepsilon_{1i}}^l), l = 1, \dots, 2k$ a $E({\varepsilon_{1i}}^l \varepsilon_{2i}), l = 1, \dots, k$ jsou známé. S předpokladem (34) je nutná pouze znalost $\sigma^2_{\varepsilon_{1i}} = E({\varepsilon_{1i}}^2)$ a $\sigma_{\varepsilon_{1i}\varepsilon_{2i}} = E(\varepsilon_{1i}\varepsilon_{2i})$. Rozptyl $\sigma^2_{\varepsilon_{2i}}$ nemusí být znám, kromě alternativního odhadu navrženého v části 2.2.

2.2 Upravená metoda nejmenších čtverců

Pokud jsou hodnoty $\mu_i, i = 1, \dots, n$, známé, můžeme k odhadu regresních parametrů $\boldsymbol{\beta}$ použít obyčejnou metodu nejmenších čtverců (OLS), získáme tak odhad ve tvaru

$$\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{\varsigma}\boldsymbol{\varsigma}')^{-\mathbf{1}}\, \boldsymbol{\varsigma}\boldsymbol{y},$$

kde $\boldsymbol{\varsigma} = (\boldsymbol{\varsigma_1}, \dots, \boldsymbol{\varsigma_n})$.

Dosadíme-li, dostaneme

$$\widehat{\boldsymbol{\beta}}_{OLS} = \left[\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_n \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1^k & \mu_2^k & \cdots & \mu_n^k \end{pmatrix}\begin{pmatrix} 1 & \mu_1 & \cdots & \mu_1^k \\ 1 & \mu_2 & \cdots & \mu_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mu_n & \cdots & \mu_n^k \end{pmatrix}\right]^{-1}\left[\begin{pmatrix} 1 & 1 & \cdots & 1 \\ \mu_1 & \mu_2 & \cdots & \mu_n \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1^k & \mu_2^k & \cdots & \mu_n^k \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}\right]$$

$$= \begin{pmatrix} n & \sum_{i=1}^n \mu_i & \sum_{i=1}^n \mu_i^2 & \cdots & \sum_{i=1}^n \mu_i^k \\ \sum_{i=1}^n \mu_i & \sum_{i=1}^n \mu_i^2 & \sum_{i=1}^n \mu_i^3 & \cdots & \sum_{i=1}^n \mu_i^{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^n \mu_i^k & \sum_{i=1}^n \mu_i^{k+1} & \sum_{i=1}^n \mu_i^{k+2} & \cdots & \sum_{i=1}^n \mu_i^{2k} \end{pmatrix}^{-1}\begin{pmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n \mu_i y_i \\ \vdots \\ \sum_{i=1}^n \mu_i^k y_i \end{pmatrix}$$

$$= \left(\sum_{i=1}^n \boldsymbol{\varsigma_i}\boldsymbol{\varsigma_i'}\right)^{-1}\left(\sum_{i=1}^n \boldsymbol{\varsigma_i} y_i\right).$$

Označíme-li

$$\overline{\boldsymbol{\varsigma}\boldsymbol{\varsigma}'} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\varsigma_i}\boldsymbol{\varsigma_i'}\,, \quad \overline{\boldsymbol{\varsigma}\boldsymbol{y}} = \frac{1}{n}\sum_{i=1}^n \boldsymbol{\varsigma_i} y_i,$$

můžeme odhad $\boldsymbol{\beta}$ pomocí obyčejné metody nejmenších čtverců přepsat jako

$$\widehat{\boldsymbol{\beta}}_{OLS} = (\overline{\boldsymbol{\varsigma\varsigma'}})^{-1}\,\overline{\boldsymbol{\varsigma y}}\,. \tag{35}$$

Upravená metoda nejmenších čtverců (ALS) [8] pro regresní proměnné x_i zatížené chybami je vytvořena nahrazením neznámé matice $\overline{\boldsymbol{\varsigma\varsigma'}}$ a neznámého vektoru $\overline{\boldsymbol{\varsigma y}}$ vhodnými nestrannými odhady. Takovou maticí $\overline{\boldsymbol{H}}$ a vektorem $\overline{\boldsymbol{h}}$, že platí $E(\overline{\boldsymbol{H}}) = \overline{\boldsymbol{\varsigma\varsigma'}}$ a $E(\overline{\boldsymbol{h}}) = E(\overline{\boldsymbol{\varsigma y}}) = \overline{\boldsymbol{\varsigma\nu}}$, kde $\overline{\boldsymbol{\varsigma\nu}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\varsigma_i}\nu_i$.

Nejprve budeme hledat matici $\overline{\boldsymbol{H}}$. Uvažujeme matici $\boldsymbol{\varsigma_i\varsigma_i}'$ typu $(k+1)\times(k+1)$ a $\boldsymbol{\varsigma_i} = \left(1, \mu_i, \mu_i^2, \dots, \mu_i^k\right)'$. Pro snadnější zápis budeme řádky a sloupce matice $\boldsymbol{\varsigma_i\varsigma_i}'$ uvažovat v mezích $p, q = 0, \dots, k$. Prvek (p, q) matice $\boldsymbol{\varsigma_i\varsigma_i}'$ je roven $\sum_{i=1}^{n}\mu_i^{p+q}$. Hledáme tedy proměnné t_r, $r = 0, \dots, k$, které lze počítat z pozorovaných proměnných x_i , takové, že platí $E\big(t_r(x_i\,)\big) = \mu_i^r$. Zřejmě existují jednoznačně definované polynomy t_r proměnné x_i r-tého stupně, kde $r = 0, \dots, k$, tak že platí $E\big(t_r(x_i\,)\big) = \mu_i^r$. Tyto polynomy lze nalézt pomocí rovnice (33), kde

$$E(x_i^r) = E[(\mu_i + \varepsilon_{1i})^r] = \sum_{j=0}^{r}\binom{r}{j}{\mu_i}^j E\left({\varepsilon_{1i}}^{r-j}\right) = \sum_{j=0}^{r} c_{rj}{\mu_i}^j$$

je polynom proměnné μ_i s koeficienty $c_{rj} = \binom{r}{j}\,E\left({\varepsilon_{1i}}^{r-j}\right)$. Pokud nahradíme $E(x_i^r)$ proměnnou x_i^r a ${\mu_i}^j$ novou proměnnou $t_j(x_i)$, dostaneme systém rovnic

$$x_i^r = \sum_{j=0}^{r} c_{rj}t_j\,(x_i) \tag{36}$$

$r = 0, \dots, k$. Odtud získáme $t_j(x_i)$ ve tvaru

$$t_r(x_i) = \sum_{j=0}^{r} a_{rj}{x_i}^j, \tag{37}$$

kde a_{rj} jsou funkcemi $E({\varepsilon_{1i}}^l), l = 0, \ldots, r$. Ukážeme si, jak najdeme prvních pět $t_r(x_i)$:

$r = 0$: $\;{x_i}^0 = 1 \cdot E(\varepsilon_{1i}^0)t_0(x_i) \Longrightarrow t_0(x_i) = 1,$

$$r = 1:\; x_i = \left[\binom{1}{0} E(\varepsilon_{1i})t_0(x_i) + \binom{1}{1} E(\varepsilon_{1i}^0)t_1(x_i)\right] \Longrightarrow t_1(x_i) = x_i,$$

$$\begin{aligned} r = 2:\; {x_i}^2 &= \left[\binom{2}{0} E(\varepsilon_{1i}^2)t_0(x_i) + \binom{2}{1} E(\varepsilon_{1i})t_1(x_i) + \binom{2}{2} E(\varepsilon_{1i}^0)t_2(x_i)\right] \\ &= \sigma_{\varepsilon_{1i}}^2 + \; t_2(x_i) \;\Longrightarrow\; t_2(x_i) = {x_i}^2 - \sigma_{\varepsilon_{1i}}^2, \end{aligned}$$

$$\begin{aligned} r = 3:\; {x_i}^3 &= \left[\binom{3}{0} E(\varepsilon_{1i}^3)t_0(x_i) + \binom{3}{1} E(\varepsilon_{1i}^2)t_1(x_i) + \binom{3}{2} E(\varepsilon_{1i})t_2(x_i) + \binom{3}{3} E(\varepsilon_{1i}^0)t_3(x_i)\right] \\ &= E(\varepsilon_{1i}^3)t_0(x_i) + 3\sigma_{\varepsilon_{1i}}^2 t_1(x_i) + t_3(x_i) \;\Longrightarrow\; t_3(x_i) = {x_i}^3 - 3x_i\sigma_{\varepsilon_{1i}}^2 - E({\varepsilon_{1i}}^3) \end{aligned}$$

$$\begin{aligned} r = 4:\; {x_i}^4 &= \left[\binom{4}{0} E(\varepsilon_{1i}^4)t_0(x_i) + \binom{4}{1} E(\varepsilon_{1i}^3)t_1(x_i) + \binom{4}{2} E({\varepsilon_{1i}}^2)t_2(x_i) + \binom{4}{3} E(\varepsilon_{1i})t_3(x_i) \right. \\ &\quad \left. + \binom{4}{4} E(\varepsilon_{1i}^0)t_4(x_i)\right] \\ &= E(\varepsilon_{1i}^4)t_0(x_i) + 4E(\varepsilon_{1i}^3)t_1(x_i) + 6E({\varepsilon_{1i}}^2)t_2(x_i) + t_4(x_i) \\ &\Longrightarrow t_4(x_i) \; = {x_i}^4 - 6{x_i}^2\sigma_{\varepsilon_{1i}}^2 - 4x_iE({\varepsilon_{1i}}^3) - E({\varepsilon_{1i}}^4) + 6\sigma_{\varepsilon_{1i}}^4. \end{aligned}$$

Nyní, nechť $\boldsymbol{H}(x_i)$ je matice typu $(k+1) \times (k+1)$, jejíž prvek (p, q) je roven $t_{p+q}(x_i), p, q = 0, \ldots, k$. Matice $\boldsymbol{H}(x_i)$ má tedy tvar

$$\boldsymbol{H}(x_i) \; = \begin{pmatrix} t_0(x_i) & t_1(x_i) & t_2(x_i) & \cdots & t_k(x_i) \\ t_1(x_i) & t_2(x_i) & t_3(x_i) & \cdots & t_{(k+1)}(x_i) \\ t_2(x_i) & t_3(x_i) & t_4(x_i) & \cdots & t_{(k+2)}(x_i) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_k(x_i) & t_{(k+1)}(x_i) & t_{(k+2)}(x_i) & \cdots & t_{2k}(x_i) \end{pmatrix}.$$

Potom

$$E\big(\boldsymbol{H}(x_i)\big) \; = \begin{pmatrix} E\big(t_0(x_i)\big) & E\big(t_1(x_i)\big) & E\big(t_2(x_i)\big) & \cdots & E\big(t_k(x_i)\big) \\ E\big(t_1(x_i)\big) & E\big(t_2(x_i)\big) & E\big(t_3(x_i)\big) & \cdots & E\left(t_{(k+1)}(x_i)\right) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E\big(t_k(x_i)\big) & E\left(t_{(k+1)}(x_i)\right) & E\left(t_{(k+2)}(x_i)\right) & \cdots & E\big(t_{2k}(x_i)\big) \end{pmatrix}$$

$$= \begin{pmatrix} 1 & \mu_i & \mu_i^2 & \cdots & \mu_i^k \\ \mu_i & \mu_i^2 & \mu_i^3 & \cdots & \mu_i^{k+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_i^k & \mu_i^{k+1} & \mu_i^{k+2} & \cdots & \mu_i^{2k} \end{pmatrix} = \begin{pmatrix} 1 \\ \mu_i \\ \vdots \\ \mu_i^k \end{pmatrix} (1 \quad \mu_i \quad \dots \quad \mu_i^k) = \boldsymbol{\varsigma_i} \boldsymbol{\varsigma_i}'$$

a položíme-li

$$\overline{\boldsymbol{H}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{H}(x_i),$$

potom $E(\overline{\boldsymbol{H}}) = \overline{\boldsymbol{\varsigma\varsigma'}}$.

Nyní budeme hledat vektor $\overline{\boldsymbol{h}}$. Uvažujeme vektor $\boldsymbol{\varsigma_i} \nu_i$ typu $(k+1) \times 1$, který má prvky $\sum_{i=1}^{n} {\mu_i}^r \nu_i, r = 0, \dots, k$. Potom

$${\mu_i}^r \nu_i = E(t_r(x_i)\nu_i) = E\big(t_r(x_i)(y_i - \varepsilon_{2i})\big) = E(t_r(x_i)y_i) - E(t_r(x_i)\varepsilon_{2i}).$$

Užitím vztahu (37) dostaneme

$$E(t_r(x_i)\varepsilon_{2i}) = E\left\{\sum_{j=0}^{r} a_{rj} {x_i}^j \varepsilon_{2i}\right\} = E\left\{\sum_{j=0}^{r} a_{rj} (\mu_i + \varepsilon_{1i})^j \varepsilon_{2i}\right\}$$

$$= E\left\{\sum_{j=0}^{r} a_{rj} \sum_{s=j}^{r} \binom{s}{j} {\varepsilon_{1i}}^{s-j} {\mu_i}^j \varepsilon_{2i}\right\} = \sum_{j=0}^{r} {\mu_i}^j \sum_{s=j}^{r} a_{rs} \binom{s}{j} E\big({\varepsilon_{1i}}^{s-j} \varepsilon_{2i}\big)$$

$$= \sum_{j=0}^{r} b_{rj} {\mu_i}^j$$

s koeficienty

$$b_{rj} = \sum_{s=j}^{r} a_{rs} \binom{s}{j} E\big({\varepsilon_{1i}}^{s-j} \varepsilon_{2i}\big), \tag{38}$$

které závisí pouze na $E({\varepsilon_{1i}}^l)$ (pomocí koeficientů $a_{rs} = \binom{r}{s} E({\varepsilon_{1i}}^{r-s})$) a $E({\varepsilon_{1i}}^l \varepsilon_{2i})$, $l = 0, \dots, r$. Lze je vypočítat, jakmile jsou známy tyto momenty a a_{rs} byly stanoveny jako v rovnici (37). Přirozený odhad $E(t_r(x_i)\varepsilon_{2i})$, proto je $\hat{E}(t_r(x_i)\varepsilon_{2i}) = \sum_{j=0}^{r} b_{rj} t_j(x_i)$.

Proto, pokud definujeme

$$h_r = t_r(x_i)y_i - \hat{E}(t_r(x_i)\varepsilon_{2i}) = t_r(x_i)y_i - \sum_{j=0}^{r} b_{rj} t_j(x_i), \qquad (39)$$

potom $E(h_r) = E(t_r(x_i)y_i) - E(t_r(x_i)\varepsilon_{2i}) = {\mu_i}^r \nu_i$.

Nyní nechť $\boldsymbol{h} = \boldsymbol{h}(x_i, y_i) = (h_0, h_1, \dots, h_k)'$. Potom

$$E(\boldsymbol{h}) = \begin{pmatrix} E(h_0) \\ E(h_1) \\ \vdots \\ E(h_k) \end{pmatrix} = \begin{pmatrix} {\mu_i}^0 \nu_i \\ {\mu_i}^1 \nu_i \\ \vdots \\ {\mu_i}^k \nu_i \end{pmatrix} = \begin{pmatrix} \nu_i \\ \mu_i \nu_i \\ \vdots \\ {\mu_i}^k \nu_i \end{pmatrix} = \boldsymbol{\varsigma}_i \nu_i,$$

a proto $E\left(\overline{\boldsymbol{h}}\right) = \overline{\boldsymbol{\varsigma\nu}} = E(\overline{\boldsymbol{\varsigma y}})$, kde

$$\overline{\boldsymbol{h}} = \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{h}(x_i, y_i).$$

První tři složky vektoru $\boldsymbol{h}$ spočítáme pomocí vztahu (39):

$$\begin{aligned} r = 0\colon\ h_0 &= t_0(x_i)y_i - \sum_{j=0}^{0} b_{rj} t_j(x_i) = 1 \cdot y_i - \binom{0}{0} E(\varepsilon_{1i}^0) \binom{0}{0} E(\varepsilon_{1i}^0 \varepsilon_{2i}) \cdot 1 \\ &= y_i - E(\varepsilon_{2i}) = y_i \end{aligned}$$

$$r=1\colon\ h_1=t_1(x_i)y_i-\sum_{j=0}^{1}b_{rj}t_j(x_i)=x_iy_i-b_{01}t_0(x_i)-b_{11}t_1(x_i)=x_iy_i-$$

$$\sum_{s=0}^{1}\binom{r}{s}E({\varepsilon_{1i}}^{r-s})\binom{s}{0}E({\varepsilon_{1i}}^{s-0}\varepsilon_{2i})\cdot 1-\sum_{s=1}^{1}\binom{r}{s}E({\varepsilon_{1i}}^{r-s})\binom{s}{1}E({\varepsilon_{1i}}^{s-1}\varepsilon_{2i})x_i$$

$$=x_iy_i-\binom{1}{0}E(\varepsilon_{1i})\binom{0}{0}E(\varepsilon_{1i}^0\varepsilon_{2i})-\binom{1}{1}E(\varepsilon_{1i}^0)\binom{1}{0}E(\varepsilon_{1i}\varepsilon_{2i})$$

$$-\binom{1}{1}E(\varepsilon_{1i}^0)\binom{1}{1}E(\varepsilon_{1i}^0\varepsilon_{2i})x_i$$

$$=x_iy_i-E(\varepsilon_{1i}\varepsilon_{2i})=x_iy_i-\sigma_{\varepsilon_{1i}\varepsilon_{2i}}$$

$$r=2\colon\ h_2=t_2(x_i)y_i-\sum_{j=0}^{2}b_{rj}t_j(x_i)=\left({x_i}^2-\sigma^2_{\varepsilon_{1i}}\right)y_i-b_{02}t_0(x_i)-b_{12}t_1(x_i)-b_{22}t_2(x_i)$$

$$=\left({x_i}^2-\sigma^2_{\varepsilon_{1i}}\right)y_i-\sum_{s=0}^{2}\binom{r}{s}E({\varepsilon_{1i}}^{r-s})\binom{s}{0}E({\varepsilon_{1i}}^{s-0}\varepsilon_{2i})\cdot 1$$

$$-\sum_{s=1}^{2}\binom{r}{s}E({\varepsilon_{1i}}^{r-s})\binom{s}{1}E({\varepsilon_{1i}}^{s-1}\varepsilon_{2i})x_i$$

$$-\sum_{s=2}^{2}\binom{r}{s}E({\varepsilon_{1i}}^{r-s})\binom{s}{2}E({\varepsilon_{1i}}^{s-2}\varepsilon_{2i})\left({x_i}^2-\sigma^2_{\varepsilon_{1i}}\right)$$

$$=\left({x_i}^2-\sigma^2_{\varepsilon_{1i}}\right)y_i-\binom{2}{0}E(\varepsilon_{1i}^2)\binom{0}{0}E(\varepsilon_{1i}^0\varepsilon_{2i})-\binom{2}{1}E(\varepsilon_{1i})\binom{1}{0}E(\varepsilon_{1i}\varepsilon_{2i})$$

$$-\binom{2}{2}E(\varepsilon_{1i}^0)\binom{2}{0}E(\varepsilon_{1i}^2\varepsilon_{2i})-\binom{2}{1}E(\varepsilon_{1i})\binom{1}{1}E(\varepsilon_{1i}^0\varepsilon_{2i})x_i$$

$$-\binom{2}{2}E(\varepsilon_{1i}^0)\binom{2}{1}E(\varepsilon_{1i}\varepsilon_{2i})x_i-\binom{2}{2}E(\varepsilon_{1i}^0)\binom{2}{2}E(\varepsilon_{1i}^0\varepsilon_{2i})x_i$$

$$-\binom{2}{2}E(\varepsilon_{1i}^0)\binom{2}{2}E(\varepsilon_{1i}^0\varepsilon_{2i})\left({x_i}^2-\sigma^2_{\varepsilon_{1i}}\right)$$

$$=\left({x_i}^2-\sigma^2_{\varepsilon_{1i}}\right)y_i-E({\varepsilon_{1i}}^2\varepsilon_{2i})-2\sigma_{\varepsilon_{1i}\varepsilon_{2i}}x_i.$$

Ve speciálním případě nezávislosti mezi ε_{1i} a ε_{2i} ve výrazu $\hat{E}(t_r(x_i)\varepsilon_{2i})$ zmizí a $\overline{\boldsymbol{h}}$ redukuje $\overline{ty}$, kde $t=\left(t_0(x_i),t_1(x_i),\dots,t_k(x_i)\right)'$.

Odhad $\boldsymbol{\beta}$ metodou ALS zkonstruujeme nahrazením $\overline{\boldsymbol{\varsigma\varsigma'}}$ a $\overline{\boldsymbol{\varsigma y}}$ v rovnici (35) maticí $\overline{\boldsymbol{H}}$ a vektorem $\overline{\boldsymbol{h}}$:

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = (\overline{\boldsymbol{H}})^{-1}\overline{\boldsymbol{h}}\,. \tag{40}$$

V případě kvadratické závislosti ($k = 2$), je matice $\overline{\boldsymbol{H}}$ s prvními pěti $t_r(x_i)$ a vektor $\overline{\boldsymbol{h}}$ s prvními třemi h_r:

$$\overline{\boldsymbol{H}} = \begin{pmatrix} 1 & \bar{x} & \overline{x^2} - \sigma^2_{\varepsilon_{1i}} \\ \bar{x} & \overline{x^2} - \sigma^2_{\varepsilon_{1i}} & \overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}} \\ \overline{x^2} - \sigma^2_{\varepsilon_{1i}} & \overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}} & \overline{x^4} - 6\overline{x^2}\sigma^2_{\varepsilon_{1i}} + 3\sigma^4_{\varepsilon_{1i}} \end{pmatrix},$$

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix} \quad \text{a} \quad \overline{\boldsymbol{h}} = \begin{pmatrix} \bar{y} \\ \overline{xy} - \sigma_{\varepsilon_{1i}\varepsilon_{2i}} \\ \overline{x^2y} - \bar{y}\sigma^2_{\varepsilon_{1i}} - 2\sigma_{\varepsilon_{1i}\varepsilon_{2i}}\bar{x} \end{pmatrix}.$$

Odhady parametrů β_0, β_1 a β_2 potom získáme řešením soustavy rovnic $\overline{\boldsymbol{H}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = \overline{\boldsymbol{h}}$:

$$\begin{aligned} \hat{\beta}_0 + \bar{x}\hat{\beta}_1 + \left(\overline{x^2} - \sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_2 &= \bar{y}\,, \\ \bar{x}\hat{\beta}_0 + \left(\overline{x^2} - \sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_1 + \left(\overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_2 &= \overline{xy} - \sigma_{\varepsilon_{1i}\varepsilon_{2i}}\,, \\ \left(\overline{x^2} - \sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_0 + \left(\overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_1 + \left(\overline{x^4} - 6\overline{x^2}\sigma^2_{\varepsilon_{1i}} + 3\sigma^4_{\varepsilon_{1i}}\right)\hat{\beta}_2 &= \overline{x^2y} - \bar{y}\sigma^2_{\varepsilon_{1i}} - 2\sigma_{\varepsilon_{1i}\varepsilon_{2i}}\bar{x}\,. \end{aligned}$$

Odhad $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$ nalezneme řešením rovnice (40), tj. $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = \overline{\boldsymbol{H}}^{-\boldsymbol{1}}\overline{\boldsymbol{h}}$.

Podobně můžeme zkonstruovat také odhad $\sigma^2_{\varepsilon_{2i}}$. Obyčejná metoda nejmenších čtverců (OLS) odhadu $\sigma^2_{\varepsilon_{2i}}$ v modelu (33) se známými μ_i je dána

$$\begin{aligned} \hat{\sigma}^2_{\varepsilon_{2i},OLS} &= \left(\boldsymbol{y} - \boldsymbol{\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS}\right)'\left(\boldsymbol{y} - \boldsymbol{\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS}\right) \\ &= \boldsymbol{y}'\boldsymbol{y} - \widehat{\boldsymbol{\beta}}'_{OLS}\,\boldsymbol{\varsigma y} - \boldsymbol{y}'\boldsymbol{\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS} + \widehat{\boldsymbol{\beta}}'_{OLS}\,\boldsymbol{\varsigma\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS} \\ &= \boldsymbol{y}'\boldsymbol{y} - 2\boldsymbol{y}'\boldsymbol{\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS} + \boldsymbol{y}'\boldsymbol{\varsigma}'(\boldsymbol{\varsigma\varsigma}')^{-\boldsymbol{1}}\boldsymbol{\varsigma\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS} \\ &= \boldsymbol{y}'\boldsymbol{y} - \boldsymbol{y}'\boldsymbol{\varsigma}'\widehat{\boldsymbol{\beta}}_{OLS}\,. \end{aligned}$$

Dosadíme-li, dostaneme

$$\hat{\sigma}^2_{\varepsilon_{2i},OLS} = \left[(y_1 \quad y_2 \quad \cdots \quad y_n) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right] - \left[(y_1 \quad y_2 \quad \cdots \quad y_n) \begin{pmatrix} 1 & \mu_1 & \cdots & \mu_1^k \\ 1 & \mu_2 & \cdots & \mu_2^k \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \mu_n & \cdots & \mu_n^k \end{pmatrix} \widehat{\boldsymbol{\beta}}_{OLS} \right]$$

$$= \left(\sum_{i=1}^{n} y_i^2 \right) - \left(\sum_{i=1}^{n} y_i \quad \sum_{i=1}^{n} y_i \, \mu_i \quad \cdots \quad \sum_{i=1}^{n} y_i \mu_i^k \right) \widehat{\boldsymbol{\beta}}_{OLS}$$

$$= \sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i' \, \boldsymbol{\varsigma}_i' \right) \widehat{\boldsymbol{\beta}}_{OLS}.$$

Označíme-li

$$\overline{\boldsymbol{y}^2} = \frac{1}{n} \sum_{i=1}^{n} y_i^2 \,, \qquad \overline{\boldsymbol{y}'\boldsymbol{\varsigma}'} = \frac{1}{n} \sum_{i=1}^{n} y_i' \, \boldsymbol{\varsigma}_i',$$

dostaneme

$$\hat{\sigma}^2_{\varepsilon_{2i},OLS} = \overline{\boldsymbol{y}^2} - \overline{\boldsymbol{y}'\boldsymbol{\varsigma}'} \widehat{\boldsymbol{\beta}}_{OLS}.$$

Nahradíme-li $\overline{\boldsymbol{y}'\boldsymbol{\varsigma}'}$ vektorem $\overline{\boldsymbol{h}}'$ a odhad $\widehat{\boldsymbol{\beta}}_{OLS}$ odhadem $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$ získáme odhad parametru $\sigma^2_{\varepsilon_2}$ metodou ALS ve tvaru

$$\hat{\sigma}^2_{\varepsilon_{2i}} = \overline{\boldsymbol{y}^2} - \overline{\boldsymbol{h}}' \widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = \overline{\boldsymbol{y}^2} - \overline{\boldsymbol{h}}' \overline{\boldsymbol{H}}^{-1} \overline{\boldsymbol{h}}. \tag{41}$$

Za obecných předpokladů jsou odhady $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$ a $\hat{\sigma}^2_{\varepsilon_{2i}}$ metodou ALS konzistentní. Ve skutečnosti, pouze předpokládáme, že vyšší momenty ε_{1i} a ε_{2i} (konkrétně $E({\varepsilon_{1i}}^{4k})$ a $E({\varepsilon_{1i}}^{2k}{\varepsilon_{2i}}^2)$) existují, že limity vyšších momentů μ_i, tj. $\lim(\overline{\mu^r})$, existuje pro $r = 0, \ldots, 4k,$ a že $\lim(\overline{\boldsymbol{H}})$ je nesingulární. Za podobných předpokladů je odhad $\widehat{\boldsymbol{\beta}}$ asymptoticky normální, tj. $n^{1/2}(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \to N(0, \Sigma)$. Varianční matice aproximačního normálního rozdělení lze odhadnout vztahem

$$\hat{\Sigma} = \frac{1}{n} \overline{\boldsymbol{H}}^{-1} \hat{V} \overline{\boldsymbol{H}}^{-1},$$

kde

$$\hat{V} = \frac{1}{n}\sum_{i=1}^{n}\left(\boldsymbol{H}(x_i)\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} - \boldsymbol{h}(x_i, y_i)\right)\left(\boldsymbol{H}(x_i)\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} - \boldsymbol{h}(x_i, y_i)\right)'.$$

2.2.1 Upravená metoda nejmenších čtverců za předpokladu normality

Pokud vektor náhodných chyb $(\varepsilon_{1i}, \varepsilon_{2i})$ má normální rozdělení, musí být známy pouze $\sigma_{\varepsilon_{1i}\varepsilon_{2i}}$ a $\sigma^2_{\varepsilon_{1i}}$. Potom platí

$$t_{r+1}(x_i) = x_i t_r(x_i) - \sigma^2_{\varepsilon_{1i}} r t_{r-1}\,(x_i) \tag{42}$$

pro $r = 0,1,2,...$ pro $t_0(x_i) = t_{-1}(x_i) = 1$

a

$$\hat{E}(t_r(x_i)\varepsilon_{2i}) = \sigma_{\varepsilon_{1i}\varepsilon_{2i}} r t_{r-1}(x_i). \tag{43}$$

Důkazy těchto tvrzení lze nalézt v [8].

Rovnici (40) pro odhad $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$ můžeme psát jako

$$\overline{\boldsymbol{H}}\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = \overline{ty} - \sigma_{\varepsilon_{1i}\varepsilon_{2i}}\bar{p}, \tag{44}$$

kde $p = \left(0, t_0(x_i), 2t_1(x_i), \dots, kt_{k-1}(x_i)\right)'$.

Pro kvadratický funkční vztah $(k = 2)$ rovnice (44) máme

$$\overline{\boldsymbol{H}} = \begin{pmatrix} 1 & \bar{x} & \overline{x^2} - \sigma^2_{\varepsilon_{1i}} \\ \bar{x} & \overline{x^2} - \sigma^2_{\varepsilon_{1i}} & \overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}} \\ \overline{x^2} - \sigma^2_{\varepsilon_{1i}} & \overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}} & \overline{x^4} - 6\overline{x^2}\sigma^2_{\varepsilon_{1i}} + 3\sigma^4_{\varepsilon_{1i}} \end{pmatrix},$$

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}, \quad \bar{t} = \begin{pmatrix} \overline{t_0} \\ \overline{t_1} \\ \overline{t_2} \end{pmatrix} = \begin{pmatrix} 1 \\ \bar{x} \\ \overline{x^2} - \sigma^2_{\varepsilon_{1i}} \end{pmatrix} \quad \text{a} \quad p = \left(0, t_0(x_i), 2t_1(x_i)\right)' = (0, 1, 2\bar{x})'.$$

Potom dostaneme soustavu rovnic

$$\hat{\beta}_0 + \bar{x}\hat{\beta}_1 + \left(\overline{x^2} - \sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_2 = \bar{y}\,,$$
$$\bar{x}\hat{\beta}_0 + \left(\overline{x^2} - \sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_1 + \left(\overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_2 = \overline{xy} - \sigma_{\varepsilon_{1i}\varepsilon_{2i}}\,,$$
$$\left(\overline{x^2} - \sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_0 + \left(\overline{x^3} - 3\bar{x}\sigma^2_{\varepsilon_{1i}}\right)\hat{\beta}_1 + \left(\overline{x^4} - 6\overline{x^2}\sigma^2_{\varepsilon_{1i}} + 3\sigma^4_{\varepsilon_{1i}}\right)\hat{\beta}_2 = \overline{x^2y} - \bar{y}\sigma^2_{\varepsilon_{1i}} - 2\sigma_{\varepsilon_{1i}\varepsilon_{2i}}\bar{x}\,.$$

2.3 Modifikace upravené metody nejmenších čtverců

Odhad metodou ALS je konzistentní a asymptoticky normální. Pro náhodné výběry s malým rozsahem může vést k velké chybě odhadu, a to zejména v případě, že šum $\sigma^2_{\varepsilon_{1i}}/\sigma^2_{\varepsilon_{2i}}$ je větší než 0,1. Modifikace metody ALS redukuje rozptyl odhadu, aniž by došlo k vychýlení [14].

Definujeme vektor $t_i = \left(t_0(x_i), \dots, t_k(x_i)\right)'$ a matici $V_i = t_i t_i' - \boldsymbol{H}(x_i)$. Potom odhad $\boldsymbol{\beta}$ modifikovanou metodou ALS, je dán jako řešení:

$$(\overline{tt'} - a\bar{V})\widehat{\boldsymbol{\beta}}_{\boldsymbol{MALS}} = \bar{h},$$

kde

$$a = \begin{cases} 1 - \dfrac{(k+4)}{n} & \text{pokud} \quad \rho > 1 + \dfrac{1}{n} \\ \rho\dfrac{n-(k+4)}{n+1} & \text{pokud} \quad \rho \le 1 + \dfrac{1}{n} \end{cases}$$

a ρ je nejmenší kladný kořen rovnice

$$det\left[\begin{pmatrix} \overline{y^2} & \overline{yt'} \\ \overline{ty} & \overline{tt'} \end{pmatrix} - \rho\begin{pmatrix} 0 & 0' \\ 0 & \bar{V} \end{pmatrix}\right] = 0,$$

kde

$$\overline{tt'} = \frac{1}{n}\sum_{i=1}^{n} t_i\, t_i' , \qquad \overline{yt'} = \frac{1}{n}\sum_{i=1}^{n} y_i\, t_i' , \qquad \bar{V} = \frac{1}{n}\sum_{i=1}^{n} t_i\, t_i' - \boldsymbol{H}(x_i).$$

2.4 Známá varianční matice Ω a neznámý jednotkový rozptyl

Předpokládáme, že $(\varepsilon_{1i}, \varepsilon_{2i}) \sim N(0, \boldsymbol{\Omega})$, kde

$$\boldsymbol{\Omega} = \kappa \mathrm{W}$$

se známou maticí W a neznámým parametrem κ. Řešíme rovnice (40), (41) a $\widehat{\boldsymbol{\Omega}} = \hat{\kappa}\mathrm{W}$ současně pro $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$, $\widehat{\boldsymbol{\Omega}}$ a $\hat{\kappa}$.

Iterační algoritmus: začneme počáteční hodnotou κ_0 pro κ, např. $\kappa_0 = 1$. Jestliže κ_m je hodnota κ po m-té iteraci, spočítáme $\boldsymbol{\Omega}_m = \kappa_m W$. K výpočtu $\overline{H_m}$ a $\overline{h_m}$ použijeme $\hat{\sigma}^2_{\varepsilon_{1i}m}$ a $\hat{\sigma}_{\varepsilon_{1i}\varepsilon_{2i}m}$ ze vztahů (42), (43) a odvození m-té iterace pro $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$ a $\hat{\sigma}^2_{\varepsilon_{2i}}$, tj. $\widehat{\boldsymbol{\beta}_m}$ a $\hat{\sigma}^2_{\varepsilon_{2i}m}$ z rovnic (40) a (41). Další iterace pro κ z $\kappa_{m+1} = {\hat{\sigma}^2_{\varepsilon_{2i}m}}/{W^2_{\varepsilon_{2i}}}$, kde $W^2_{\varepsilon_{2i}}$ je prvek z matice W, který odpovídá $\sigma^2_{\varepsilon_{2i}}$. Iterační cyklus opakujeme do dosažení konvergence.

3 Praktická část

Nyní teorii o lineární a polynomiální regresi s chybami v proměnných aplikujeme na příkladech, které budeme řešit pomocí softwaru R. Všechny potřebné kódy jsou uvedené na přiloženém CD. První příklad zpracovává data o gejzíru OldFaithful v Národním parku v Yellowstone. Ve druhém příkladě uvažujeme umělá vygenerovaná pro předem daný lineární vztah mezi proměnnými. Potom si ukážeme simulační studie intervalů spolehlivosti pro směrnici přímky a ve třetím příkladě na umělá data odpovídající polynomiálnímu vztahu aplikujeme metodu z kapitoly 2.

3.1 Lineární vztah mezi proměnnými

Příklad 1: Zkoumáme závislost mezi délkou trvání erupce (eruptions) a čekáním na následující erupci (waiting) gejzíru OldFaithful v Národním parku v Yellowstone, Wyoming, USA. Soubor se skládá celkem z 272 pozorování, která jsou měřena v minutách. V softwaru R je uložen pod názvem faithful.

Datovou matici si označíme jako A, kterou si v softwaru R načteme příkazem

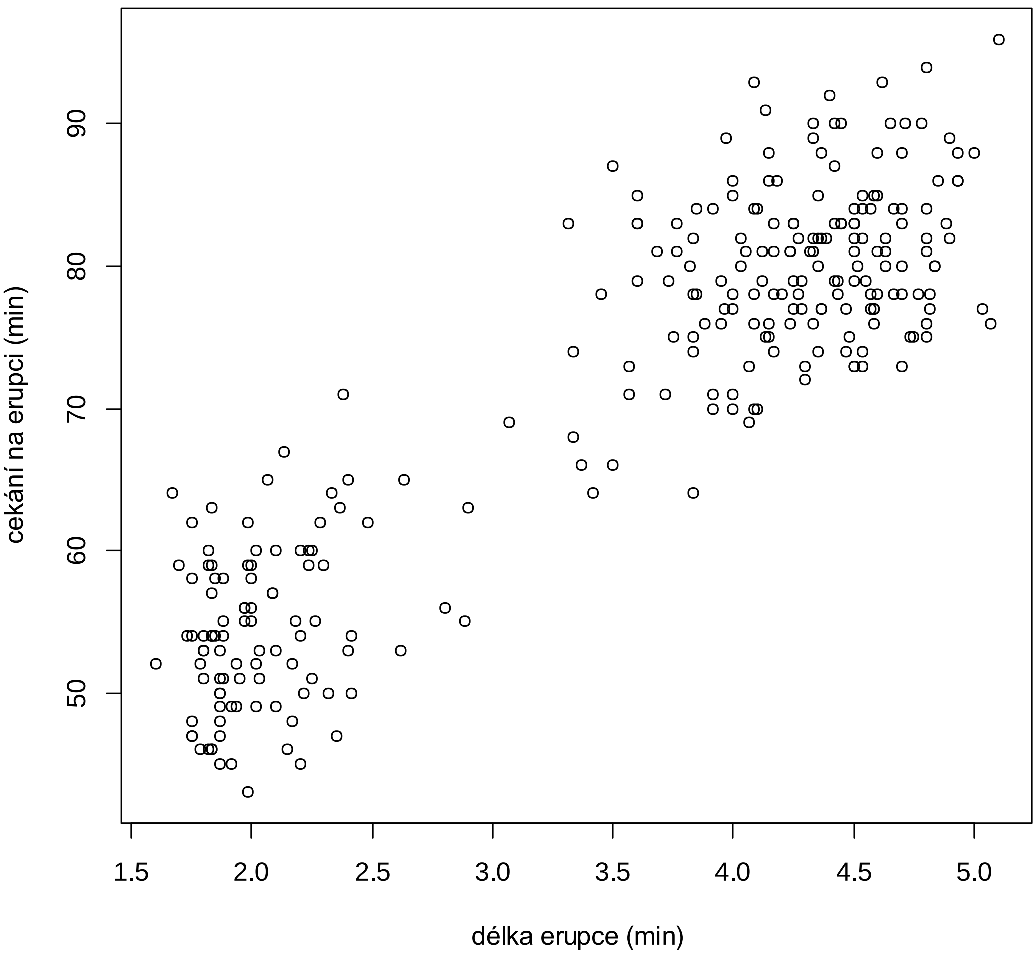
```
> A=faithful
```

Trvání erupce si označíme jako veličinu x a čekání na další erupci si označíme jako veličinu y.

```
>x=A[,1]
>y=A[,2]
```

Pro lepší představu o rozložení dat si je vykreslíme příkazem

```
>plot(y~x, xlab="délka erupce (min)",ylab="cekání na erupci (min)")
```



Obrázek 3: Závislost doby čekání na další erupci a délky předchozí erupce.

Z obrázku 3 je zřejmá lineární závislost mezi proměnnými. Zároveň se zdá, že data nejsou zatížena odlehlým pozorováním.

Analýzu dat začneme výpočtem výběrových charakteristik:

počet pozorování *n*

```
> n=length(x)
 [1] 272
```

průměrná doba trvání erupce $\bar{x}$ v minutách

```
>xpruh=mean(x)
 [1] 3.487783
```

průměrné čekání na další erupci $\bar{y}$ v minutách

```
>ypruh=mean(y)
 [1] 70.89706
```

V softwaru R můžeme výběrové rozptyly s_x^2 a s_y^2 vypočítat pomocí příkazu var(x) a var(y) a výběrovou kovarianci s_{xy} pomocí příkazu cov(x,y), které jsou nestranné, takže je vypočítáme takto:

výběrový rozptyl s_x^2

```
> x2 = x^2
> sx2 = mean(x2)-xpruh^2
[1] 1.297939
```

výběrový rozptyl s_y^2

```
> y2 = y^2
> sy2 = mean(y2)-ypruh^2
[1] 184.1438
```

výběrová kovariance s_{xy}

```
>xy=x*y
>sxy=mean(xy)-xpruh*ypruh
[1] 13.92642
```

Získané výsledky pro číselné charakteristiky nyní přehledně shrneme do tabulky 1:

číselné charakteristiky	trvání erupce	čekání na erupci
výběrové průměry	$\bar{x} = 3.487783$	$\bar{y} = 70.89706$
směrodatné odchylky	$s_x = 1.139271$	$s_y = 13.56996$
výběrová kovariance	$s_{xy} = 13.92642$	

Tabulka 1: Číselné charakteristiky pro datový soubor OldFaithful.

Jako S si označíme výběrovou kovarianční matici, která je tvaru

$$S = \begin{pmatrix} 1.297939 & 13.92642 \\ 13.92642 & 184.1438 \end{pmatrix}.$$

Nyní si ukážeme, jak pomocí softwaru R a vztahů (3) a (4) vypočítáme odhady parametrů v modelu ortogonální regresní přímky (1), kde ν_i je přesná hodnota doby do další erupce a μ_i je přesná hodnota doby erupce. Tyto odhady označíme jako beta

```
> beta=c(0,0)
```

Směrnici ortogonální regresní přímky vypočítáme pomocí vztahu (4), tj.

```
>beta[2]=(S[2,2]-S[1,1]+sqrt((S[2,2]-S[1,1])^2+4*S[1,2]^2))/(2*S[1,2])
```

a absolutní člen ortogonální regresní přímky potom vypočítáme pomocí vztahu (3)

```
> beta[1]=ypruh-beta[2]*xpruh.
```

Výsledné odhady parametrů ortogonální regresní přímky jsou

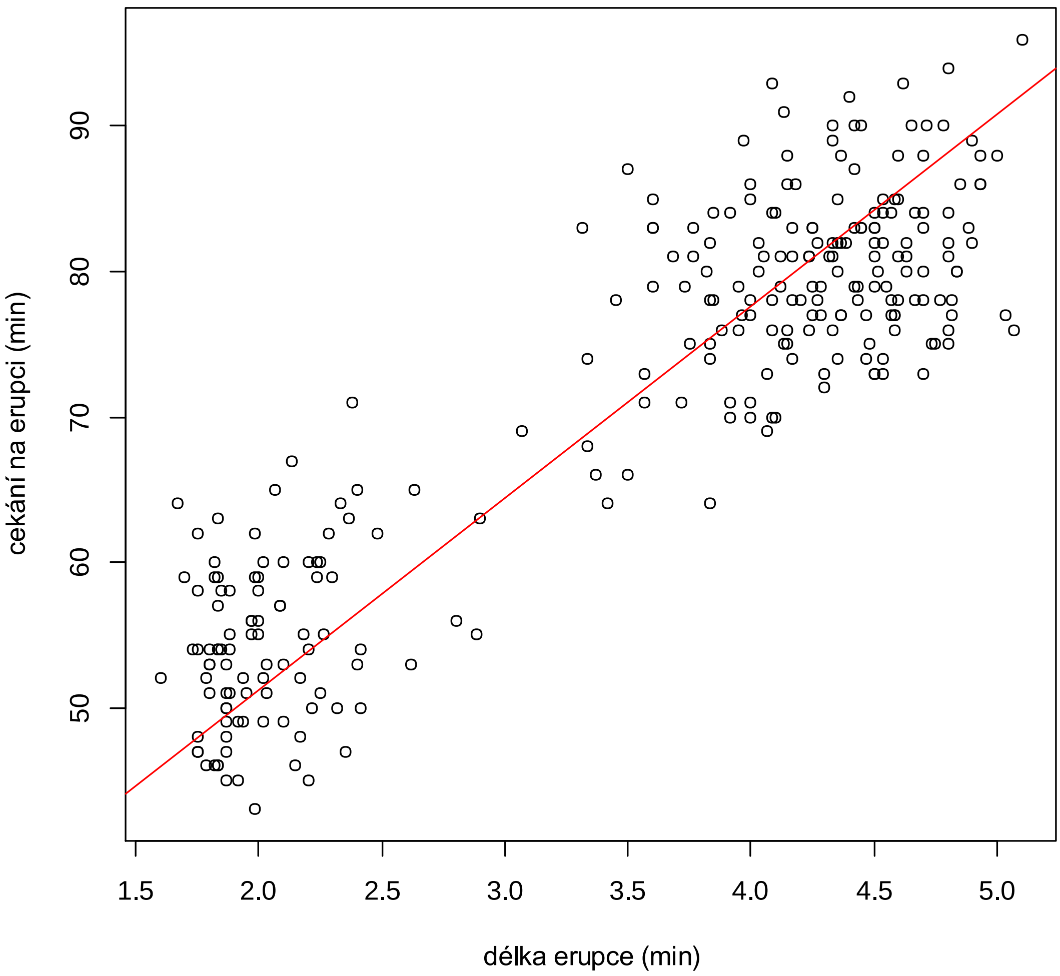
$$\hat{\beta}_0 = 24.84035 \quad \text{a} \quad \hat{\beta}_1 = 13.20515$$

a odhadnutý model tedy vypadá takto

$$\nu_i = 24.84035 + 13.20515\mu_i.$$

Nyní můžeme k našim datům přikreslit odhadnutou regresní přímku pomocí příkazu

```
>plot(y~x, xlab="délka erupce (min)",ylab="cekání na erupci (min)")
>abline(beta[1],beta[2],col='red')
```



Obrázek 4: Odhadnutá regresní přímka vztahu doby čekání na další erupci a délky předchozí erupce.

Než určíme odhad reziduálního rozptylu $\hat{\sigma}^2$ musíme nejdříve určit odhad přesné hodnoty doby erupce $\hat{\mu}_i$, který vypočítáme pomocí vzorce (10)

>mmu=((x+beta[2]*y-beta[1]*beta[2])/(1+beta[2]^2))

Potom maximálně věrohodný odhad $\hat{\sigma}^2$ určíme pomocí vzorce (9)

> rez=(y-beta[1]-beta[2]*mmu)^2+(x-mmu)^2

> sigma22 = sum(rez)/(2*n)

a výsledný odhad je

$$\hat{\sigma}^2 = 0.1216594.$$

Odhad $\hat{\sigma}^2$ není nestranný. Nestranný odhad určíme jako $\frac{2n\hat{\sigma}^2}{(n-2)} = 0.2451212.$

Další možností jak můžeme vypočítat parametry ortogonální regresní přímky je pomocí lineárních modelů. V tomto případě probíhá výpočet iteračně, kdy volíme za počáteční hodnoty koeficientů $\beta_0^{(0)} = 31.88571$ a $\beta_1^{(0)} = 12.57143$ určené pomocí vztahů (25). Za počáteční hodnoty $\boldsymbol{\mu}^{(\mathbf{0})}$ volíme $\boldsymbol{x}$ a hodnotu $\boldsymbol{\nu}^{(\mathbf{0})}$ získáme z rovnice $\boldsymbol{\nu}^{(\mathbf{0})} = \beta_0^{(0)}\mathbf{1}_{\boldsymbol{n}} + \beta_1^{(0)}\boldsymbol{\mu}^{(\mathbf{0})}$. Za ukončovací kritérium volíme druhou mocninu euklidovské normy $\left\|\hat{\beta}^{(i)} - \beta^{(i0)}\right\|^{(2)} < 10^{-9}$. Po 10 iteracích získáme stejné výsledky odhadů

$$\hat{\beta}_0 = 24.84035 \quad \text{a} \quad \hat{\beta}_1 = 13.20515$$

a odhady $\hat{\mu}, \hat{\nu}$, které jsou také shodné s výsledky z předchozí metody.

Nestranný odhad $\hat{\sigma}^2$ vypočítáme pomocí vztahu (24)

$$\hat{\sigma}^2 = 0.2451212.$$

S využitím odhadu $\hat{\sigma}^2$ můžeme odhadnout i varianční matici odhadu $\widehat{\boldsymbol{\beta}}$ ve tvaru

$$\widehat{var}(\widehat{\boldsymbol{\beta}}) = \begin{pmatrix} 1.6392912 & -0.4246955 \\ -0.4246955 & 0.1217666 \end{pmatrix}.$$

Nízké hodnoty rozptylu $\widehat{var}(\widehat{\beta_1})\ a\ \widehat{var}(\widehat{\beta_2})$ naznačují, že jsou parametry významně nenulové, a tedy existuje lineární vztah mezi délkou trvání erupce a čekáním na další erupci, což bylo patrné již při prvotním vykreslení dat.

Nyní si vypočítáme 95% intervaly spolehlivosti pro směrnici ortogonální regresní přímky pomocí vztahu (26), nebo použitím metody hlavních komponent pomocí vztahu (27) a pomocí lineárních modelů použitím vztahu (30).

Nejprve uvažujeme vztah (26):

```
>f1=(1/2)*asin(sqrt((4*(qt(0.975,n-2)^2)*(sx2*sy2-sxy^2))/((n-2)*
((sx2-sy2)^2+4*sxy^2))))
> dolni_mez1=tan(atan(beta[2])-f1)
> horni_mez1=tan(atan(beta[2])+f1)
> interval1=c(dolni_mez1, horni_mez1)
> interval1
[1] 12.48390 14.01429
```

Nyní pro výpočet intervalu spolehlivosti pomocí metody hlavních komponent využijeme vztahu (27), kdy si nejdříve vypočítáme vlastní čísla výběrové kovarianční matice (x, y) pomocí příkazu eigen:

```
> lambda=eigen(S)
> lambda1=lambda$values[1]
> lambda1
[1] 185.1984
> lambda2=lambda$values[2]
> lambda2
[1] 0.2433189
> f2=asin(sqrt(qchisq(0.95,1)/((n-1)*((lambda1/lambda2)+(lambda2/lambda1)-2))))
> dolni_mez2=tan(atan(beta[2])-f2)
> horni_mez2=tan(atan(beta[2])+f2)
> interval2=c(dolni_mez2, horni_mez2)
> interval2
[1] 12.48822 14.00886.
```

Interval spolehlivosti pomocí vztahu (30):

```
> dolnimez2=betaodhad[2]-(qt(0.975,n-2)*(sqrt(var_beta[2,2])))
> hornimez2=betaodhad[2]+(qt(0.975,n-2)*(sqrt(var_beta[2,2])))
> interval3=c(dolnimez2, hornimez2)
```

```
> interval3
[1] 12.51814 13.89216
```

Získané výsledky pro 95% intervaly spolehlivosti pro směrnici ortogonální regresní přímky nyní přehledně shrneme do následující tabulky 2:

vztah	95% intervaly spolehlivosti	délka intervalu
(26)	$\langle 12.48390; 14.01429 \rangle$	1.53039
(27)	$\langle 12.48822; 14.00886 \rangle$	1.52064
(30)	$\langle 12.51814; 13.89216 \rangle$	1.374021

Tabulka 2: Srovnání 95% intervalů spolehlivosti.

Z tabulky 2 vidíme, že intervaly spolehlivosti získané pomocí vztahů (26) a (27) se téměř překrývají. Jako nejužší vychází interval spolehlivosti spočítaný pomocí vztahu (30), tedy pomocí lineárních modelů.

Pro získanou regresní přímku určíme 95% pás spolehlivosti kolem regresní přímky (v grafu vykreslíme černou barvou) a 95% pás spolehlivosti pro regresní přímku (v grafu znázorníme zelenou barvou).

Označíme si mi0 generování čísel z intervalu $\langle -15, 15 \rangle$ s krokem 0,2. Nejprve si určíme 95% pás spolehlivosti kolem regresní přímky:

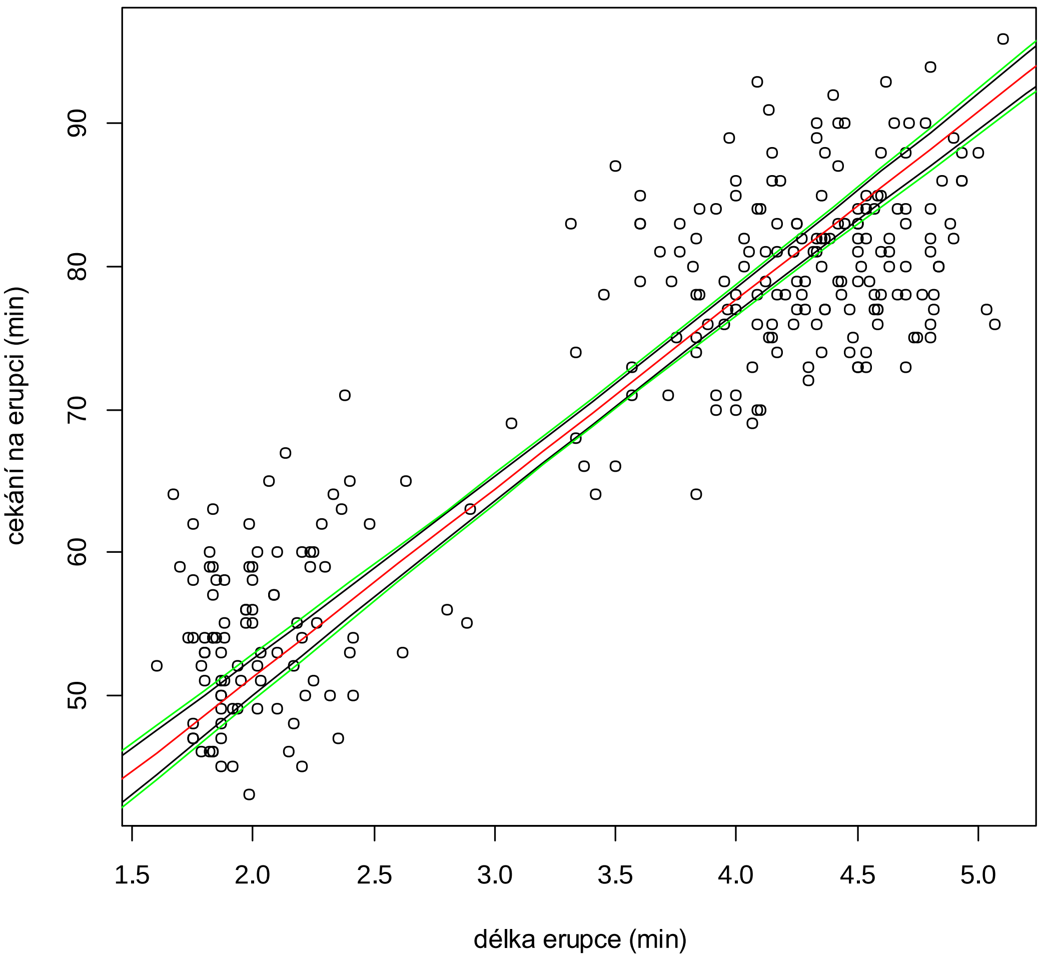
```
> mi0=seq(-15,15,0.2)
>dolni_meze=rep(0,length(mi0))
>horni_meze=rep(0,length(mi0))
>odhad_primky=rep(0,length(mi0))
>for(i in 1:length(mi0))
{dolni_meze[i]=obeta[1]+obeta[2]*mi0[i]-
sqrt(t(c(1,mi0[i]))%*%var_beta%*%c(1,mi0[i]))*qt(0.975,length(mmu)-2)
 horni_meze[i]=obeta[1]+obeta[2]*mi0[i]+sqrt(t(c(1,mi0[i]))%*%var_beta%*%
c(1,mi0[i]))*qt(0.975,length(mmu)-2)
odhad_primky[i]=obeta[1]+obeta[2]*mi0[i]}
```

Pás spolehlivosti pro regresní přímku určíme:

```
> dolni_meze2=rep(0,length(mi0))
> horni_meze2=rep(0,length(mi0))
>for(i in 1:length(mi0))
{dolni_meze2[i]=obeta[1]+obeta[2]*mi0[i]-
sqrt(t(c(1,mi0[i]))%*%var_beta%*%c(1,mi0[i]))*sqrt(2*qf(0.95,2,length(mmu)-2))
horni_meze2[i]=obeta[1]+obeta[2]*mi0[i]+sqrt(t(c(1,mi0[i]))%*%var_beta%*%
c(1,mi0[i]))*sqrt(2*qf(0.95,2,length(mmu)-2))}
```

Pásy spolehlivosti vykreslíme pomocí příkazů

```
>plot(y~x, xlab="délka erupce (min)",ylab="cekání na erupci (min)")
> lines(odhad_primky~mi0, type='l', col='red', xlim=c(-10,10), ylim=c(5,15))
> lines(dolni_meze~mi0, type='l')
> lines(horni_meze~mi0, type='l')
> lines(dolni_meze2~mi0, type='l',col='green')
> lines(horni_meze2~mi0, type='l',col='green')
```



Obrázek 5: 95% pás spolehlivosti kolem regresní přímky (černě) a 95% pás spolehlivosti pro regresní přímku (zeleně).

Z grafu vidíme, že jak pás spolehlivosti kolem regresní přímky a pás spolehlivosti pro regresní přímku téměř splývají. Přímka je odhadnuta velmi přesně, proto jsou pásy spolehlivosti úzké.

Pro ověření kvality odhadů získaných pomocí ortogonální regrese budeme uvažovat umělá data odpovídající zvolené regresní přímce $y = 10 + 5x$.

Příklad 2: Generujte dvourozměrná data rozsahu 100 odpovídající lineární závislosti $y = 10 + 5x$, jestliže předpokládáme, že chyby měření v obou proměnných jsou 2.5 cm, jsou nezávislé a mají normální rozdělení. Dále předpokládejme, že proměnná x nabývá hodnot v intervalu $\langle -10, 10 \rangle$. Na základě získaného datového souboru odhadněte ortogonální regresní přímku.

Nejprve si vygenerujeme 100 hodnot z rovnoměrného rozdělení v rozmezí od -10 do 10, které si označíme jako mii.

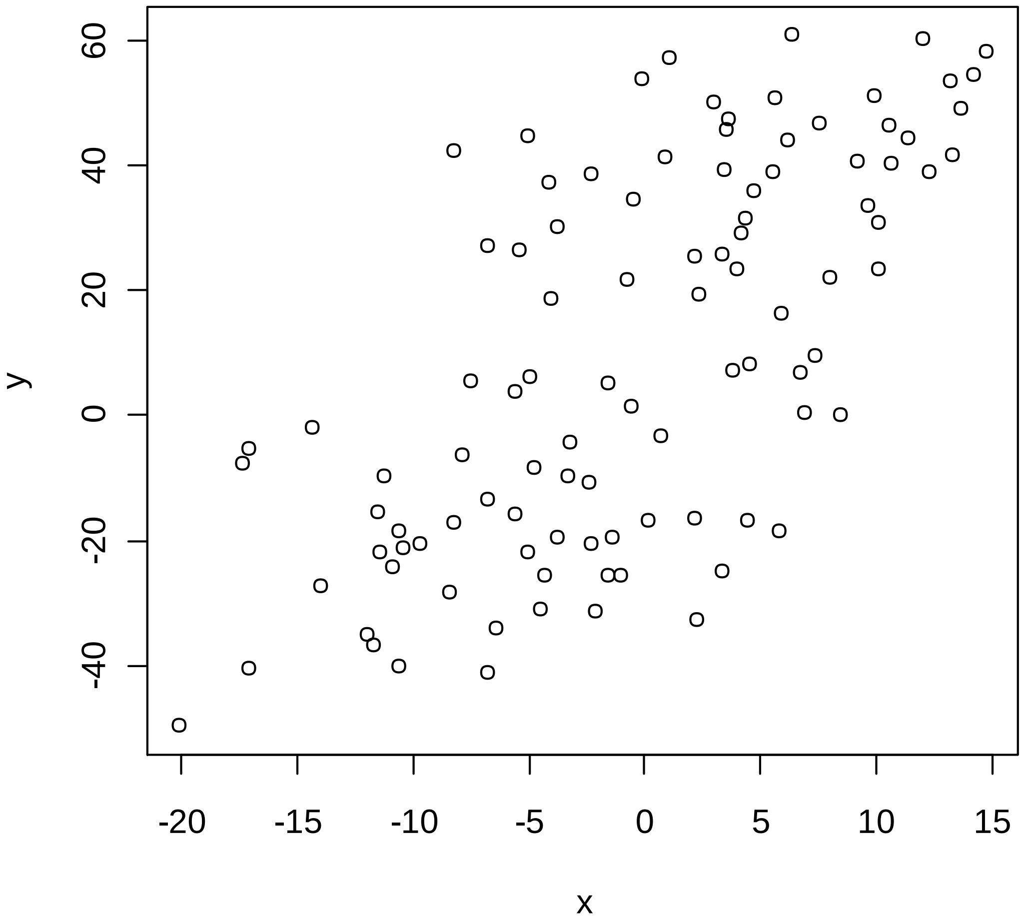
```
>mii=runif(100,min=-10, max=10)
> sigma=2.5
> sigma2=sigma^2
> beta=c(10,5)
```

Napozorované hodnoty x a y generujeme z normálního rozdělení s parametry $x \sim N_{100}(\mu_i, 2.5^2)$ a $y \sim N_{100}(10 + 5\mu_i, 2.5^2)$.

```
> x=rnorm(length(mii),mii,sigma2)
> y=rnorm(length(mii),beta[1]+beta[2]*mii,sigma2).
```

Pro představu o rozložení dat si je vykreslíme příkazem

```
>plot(y~x)
```



Obrázek 6: Generované hodnoty z příkladu 2.

Stejným způsobem jako v předchozím případě si nejdříve vypočítáme výběrové charakteristiky, z kterých můžeme určit výběrovou kovarianční matici ve tvaru

$$S = \begin{pmatrix} 65.56094 & 167.07104 \\ 167.07104 & 928.6703 \end{pmatrix}.$$

Následně odhady parametrů ortogonální regresní přímky jsou

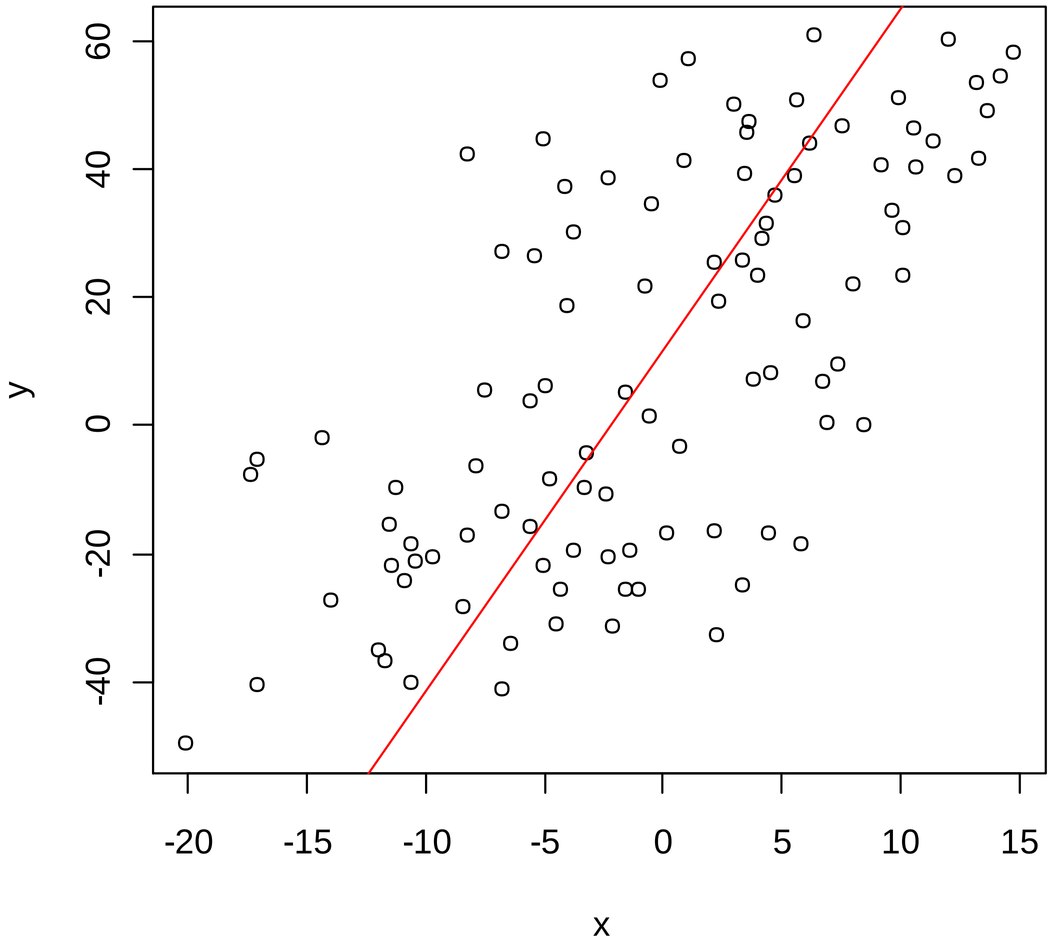
$$\hat{\beta}_0 = 11.75963 \text{ a } \hat{\beta}_1 = 5.352935.$$

Výsledný odhadnutý model je tvaru

$$\nu_i = 11.75963 + 5.352935\mu_i.$$

Nyní k vygenerovaným hodnotám můžeme přikreslit odhadnutou regresní přímku

```
>plot(y~x)
>abline(obeta[1],obeta[2],col='red')
```



Obrázek 7: Odhadnutá regresní přímka z příkladu 2.

Odhady $\widehat{\mu_i}$ a $\widehat{\nu_i}$ určíme pomocí příkazů

```
>mmu=((x+obeta[2]*y-obeta[1]*obeta[2])/(1+obeta[2]^2))
>nnu=obeta[1]+obeta[2]*mmu
```

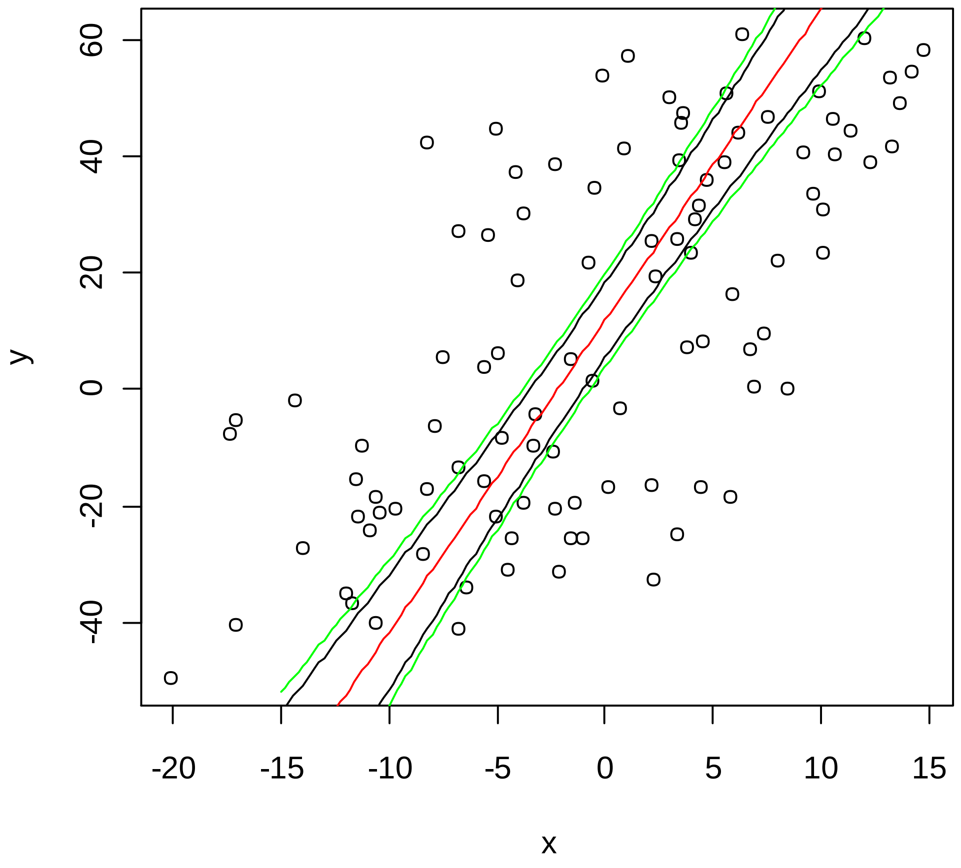
Pomocí vztahu (24) si ještě vypočítáme nestranný odhad $\hat{\sigma}^2$, který je roven

$$\hat{\sigma}^2 = 35.05085$$

a s využitím odhadu $\hat{\sigma}^2$ můžeme odhadnout i varianční matici odhadu $\widehat{\boldsymbol{\beta}}$ ve tvaru

$$\widehat{var}(\widehat{\boldsymbol{\beta}}) = \begin{pmatrix} 10.42671491 & 0.07207282 \\ 0.07207282 & 0.15853876 \end{pmatrix}.$$

Stejným způsobem jako v předchozím případě určíme pro získanou regresní přímku 95% pás spolehlivosti kolem regresní přímky (v grafu vykreslíme černou barvou) a 95% pás spolehlivosti pro regresní přímku (v grafu znázorníme zelenou barvou).



Obrázek 8: 95% pás spolehlivosti kolem regresní přímky (černě) a 95% pás spolehlivosti pro regresní přímku (zeleně) z příkladu 2.

Z obrázku 8 vidíme, že jak pás spolehlivosti kolem regresní přímky, tak i pás spolehlivosti pro regresní přímku jsou si velmi blízké.

3.1.2 Simulační studie

V této části si pomocí simulační studie porovnáme jednotlivé intervaly spolehlivosti pro směrnici ortogonální regresní přímky počítané pomocí vztahů (26), (27) a (30). Cílem studie je zjistit, jaká je empirická hladina spolehlivosti pro jednotlivé typy intervalů spolehlivosti, a jak se bude měnit v závislosti na velikosti rozsahu náhodného výběru, na přesnosti měření a předepsané hladině spolehlivosti.

Provedeme 1000 simulací, kdy budeme postupně uvažovat rozsah náhodného výběru velikosti 50, 100 a 500. Mezi bezchybnými veličinami budeme uvažovat lineární závislost $\nu = 10 + 5\mu_i$. Nepozorované přesné hodnoty μ_i budeme generovat z rovnoměrného rozdělení na intervalu $\langle -10, 10\rangle$. Následně pro různé volby směrodatné odchylky σ generujeme hodnoty x a y z normálního rozdělení s parametry $x \sim N(\mu_i, \sigma^2)$ a $y \sim N(\beta_0 + \beta_1\mu_i, \sigma^2)$.

Následující tabulky srovnávají jednotlivé intervaly spolehlivosti pro různé volby hladiny spolehlivosti $1 - \alpha$. V prvním řádku máme průměrný odhad parametru $\widehat{\beta_1}$ a na druhém řádku je empirická hladina spolehlivosti, tj. v kolika % případů interval spolehlivosti pokrývá skutečnou hodnotu $\beta_1 = 5$.

Nejprve uvažujeme vztah (26):

α=0.01

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =5.33381 98.8%	$\widehat{\beta_1}$ =5.07253 98.9%	$\widehat{\beta_1} = 5.13199$ 98.8%
0.5	$\widehat{\beta_1}$ =5.00811 99.3%	$\widehat{\beta_1}$ =4.998567 99%	$\widehat{\beta_1}$ =4.988177 99%
0.1	$\widehat{\beta_1}$ =4.998891 98.8%	$\widehat{\beta_1}$ =5.000531 98.9%	$\widehat{\beta_1}$ =4.999751 99%

α=0.05

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =5.134173 95.9%	$\widehat{\beta_1}$ =4.916848 95.3%	$\widehat{\beta_1}$ =4.949546 94.6%
0.5	$\widehat{\beta_1}$ =4.972889 96.1%	$\widehat{\beta_1}$ =4.998364 94.3%	$\widehat{\beta_1}$ =5.006847 94.5%
0.1	$\widehat{\beta_1}$ =5.000184 94.4%	$\widehat{\beta_1}$ =4.999658 94.2%	$\widehat{\beta_1}$ =4.999479 95.1%

α=0.1

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =4.595274 90.1%	$\widehat{\beta_1}$ =4.739466 91%	$\widehat{\beta_1}$ =5.015269 90%
0.5	$\widehat{\beta_1}$ =4.988523 89.2%	$\widehat{\beta_1}$ =4.97888 89.4%	$\widehat{\beta_1}$ =4.995379 90.2%
0.1	$\widehat{\beta_1}$ =5.000484 90.2%	$\widehat{\beta_1}$ =4.999978 90.2%	$\widehat{\beta_1}$ =4.999632 89.9%

Pro výpočet dalšího intervalu spolehlivosti využijeme vztahu (27)

α=0.01

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =5.061327 98.4%	$\widehat{\beta_1}$ =4.700248 99.2%	$\widehat{\beta_1}$ =4.838162 99%
0.5	$\widehat{\beta_1}$ =5.016673 99%	$\widehat{\beta_1}$ =4.984538 98.9%	$\widehat{\beta_1}$ =5.000249 99.2%
0.1	$\widehat{\beta_1}$ =4.999694 98%	$\widehat{\beta_1}$ =5.00038 98.2%	$\widehat{\beta_1}$ =5.000218 99.4%

α=0.05

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =4.642826 95.2%	$\widehat{\beta_1}$ =5.176025 94.5%	$\widehat{\beta_1}$ =4.952039 94.5%
0.5	$\widehat{\beta_1}$ =5.011174 94.4%	$\widehat{\beta_1}$ =4.976801 93.2%	$\widehat{\beta_1}$ =5.009371 95.3%
0.1	$\widehat{\beta_1}$ =5.000311 92.9%	$\widehat{\beta_1}$ =5.000736 94.2%	$\widehat{\beta_1}$ =4.999922 95.1%

α=0.1

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =4.852341 87.1%	$\widehat{\beta_1}$ =4.82751 88.8%	$\widehat{\beta_1}$ =4.983921 90.2%
0.5	$\widehat{\beta_1}$ =4.9609 90.3%	$\widehat{\beta_1}$ =5.010326 88.5%	$\widehat{\beta_1}$ =4.994891 91.2%
0.1	$\widehat{\beta_1}$ =4.999625 89.2%	$\widehat{\beta_1}$ =5.000162 88.6%	$\widehat{\beta_1}$ =4.999473 89.9%

Interval spolehlivosti pomocí vztahu (30):

α=0.01

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =5.074531 99.3%	$\widehat{\beta_1}$ =4.827902 98.3%	$\widehat{\beta_1}$ =4.877762 98.9%
0.5	$\widehat{\beta_1} = 4.990249$ 98.7%	$\widehat{\beta_1} = 4.990992$ 98.9%	$\widehat{\beta_1}$ =5.004379 99%
0.1	$\widehat{\beta_1}$ =5.000158 99.3%	$\widehat{\beta_1}$ =5.000057 99.4%	$\widehat{\beta_1}$ =4.999932 98.7%

α=0.05

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =4.8717 93.9%	$\widehat{\beta_1}$ =5.349996 93%	$\widehat{\beta_1}$ =5.103249 92.7%
0.5	$\widehat{\beta_1}$ =4.960301 95.4%	$\widehat{\beta_1}$ =4.989452 94.4%	$\widehat{\beta_1}$ =4.993051 94.1%
0.1	$\widehat{\beta_1}$ =5.001277 95.2%	$\widehat{\beta_1}$ =4.999839 94.3%	$\widehat{\beta_1}$ =4.99935 94.9%

α=0.1

σ/n	50	100	500
1.5	$\widehat{\beta_1}$ =4.88196 87.7%	$\widehat{\beta_1}$ =5.339145 85.9%	$\widehat{\beta_1}$ =5.022136 87.9%
0.5	$\widehat{\beta_1}$ =5.020662 90.6%	$\widehat{\beta_1}$ =4.976398 89.3%	$\widehat{\beta_1}$ =4.987802 90.3%
0.1	$\widehat{\beta_1}$ =4.999946 88.8%	$\widehat{\beta_1}$ =5.000476 90.6%	$\widehat{\beta_1}$ =5.00021 89.5%

Empirická hladina spolehlivosti odpovídá zvolené hladině spolehlivosti $1-\alpha$. Hodnoty občas kolísají, takže s větším počtem pozorování je dosažená empirická hladina spolehlivosti horší. Mělo být zvoleno více simulací.

3.2 Polynomiální vztah mezi proměnnými

Příklad 3: Generujte trojrozměrná data rozsahu 100 odpovídající kvadratické závislosti $y = 15 + 8x + 5x^2$, jestliže předpokládáme, že chyby měření v obou proměnných jsou 0.1, jsou nezávislé a mají normální rozdělení. Dále předpokládejme, že proměnná x nabývá hodnot v intervalu $\langle -5, 5 \rangle$. Na základě získaného datového souboru odhadněte parametr $\boldsymbol{\beta}$ upravenou metodou nejmenších čtverců.

Nejprve si vygenerujeme 100 hodnot rovnoměrného rozdělení v rozmezí od -5 do 5, které si označíme jako mii.

```
> mii=runif(100,min=-5,max=5)
> sigma1=0.1
> sigma2=0.1
> sigma12=0
> beta=c(15,8,5)
```

Z normálního rozdělení si vygenerujeme chyby ε_1 a ε_2 s parametry $\varepsilon_1 \sim N(0, 0.1^2)$ a $\varepsilon_2 \sim N(0, 0.1^2)$.

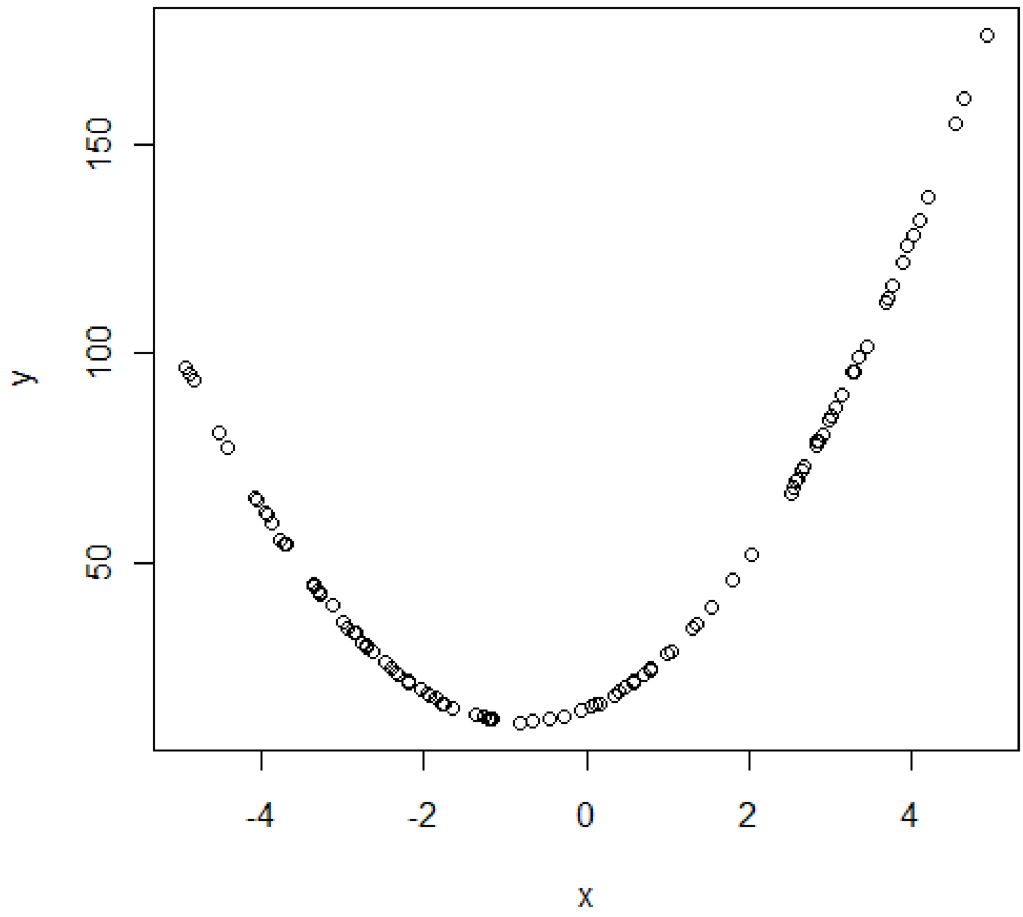
```
> chyba1=rnorm(length(mii),0, sigma1^2)
> chyba2=rnorm(length(mii),0, sigma2^2)
```

Hodnoty *x* a *y* určíme jako

```
> x=mii+chyba1
> y=beta[1]+beta[2]*mii+beta[3]*(mii^2)+chyba2
```

Data si vykreslíme příkzem

```
> plot(y~x)
```



Obrázek 9: Generované hodnoty z příkladu 3.

Nyní si vypočítáme jednotlivé průměry, které si potom dosadíme do matice $\bar{H}$ a vektoru $\bar{h}$.

$\bar{x}$:

```
> xpruh=mean(x)
 [1] -0.1593831
```

$\bar{y}$:

```
> ypruh=mean(y)
 [1] 54.44258
```

$\overline{y^2}$:

```
> y2=y^2
> y2pruh=mean(y2)
 [1] 4546.733
```

$\overline{x^2}$:

```
> x2=x^2
> x2pruh=mean(x2)
 [1] 8.143372
```

$\overline{x^3}$:

```
> x3=x^3
> x3pruh=mean(x3)
 [1] -1.710247
```

$\overline{x^4}$:

```
> x4=x^4
> x4pruh=mean(x4)
 [1] 110.204
```

$\overline{xy}$:

```
> xy=x*y
> xypruh=mean(xy)
 [1] 54.26573
```

$\overline{x^2y}$:

```
> x2y=x^2*y
> x2ypruh=mean(x2y)
 [1] 659.2846
```

Odhad β metodou ALS zkonstruujeme pomocí rovnice (40). Nejdříve si nadefinujeme matici $\overline{H}$:

```
>H=matrix(c(1,xpruh,x2pruh-sigma1^2,xpruh,x2pruh-sigma1^2,x3pruh-3*xpruh*sigma1^2,
  x2pruh-sigma1^2,x3pruh-3*xpruh*sigma1^2,x4pruh6*x2pruh*sigma1^2+3*(sigma1^2)^2),
  nrow=3,ncol=3)
```

$$\overline{H} = \begin{pmatrix} 1 & -0.1593831 & 8.133372 \\ -0.1593831 & 8.133372 & -1.705465 \\ 8.133372 & -1.705465 & 109.715688 \end{pmatrix}.$$

Dále si vypočítáme vektor $\bar{h}$:

```
> h=c(0,0,0)
> h[1]=ypruh
> h[2]=xypruh-sigma12
> h[3]=x2ypruh-(ypruh*sigma1^2)-(2*sigma12*xpruh)
```

$$\bar{h} = \begin{pmatrix} 54.44258 \\ 54.26573 \\ 658.74017 \end{pmatrix}.$$

Odhady parametrů $\widehat{\boldsymbol{\beta}}_{\boldsymbol{ALS}}$ získáme výpočtem

```
> obeta=solve(H)%*%h
> obeta
         [,1]
[1,] 14.79234
[2,]  8.017028
[3,]  5.032109
```

Odhadnutý model tedy vypadá takto

$$\nu_i = 14.79234 + 8.017028\mu_i + 5.032109\mu_i{}^2.$$

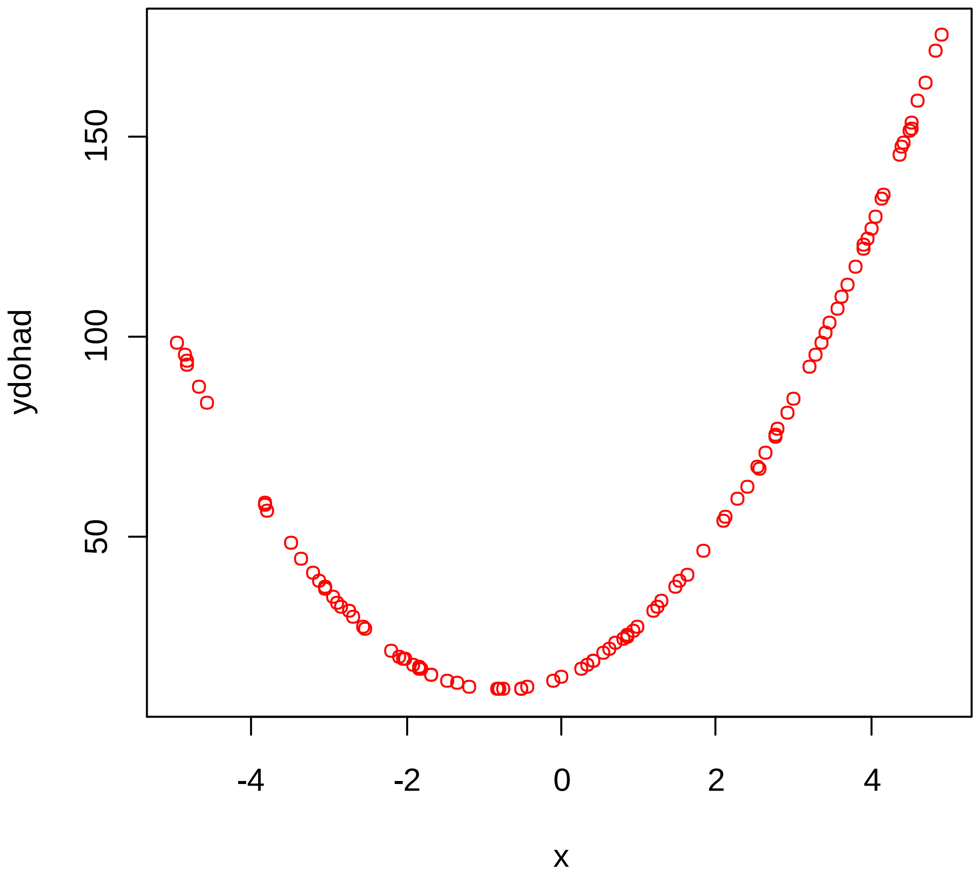
Odhad parametru $\sigma^2_{\varepsilon_2}$ metodou ALS určíme pomocí rovnice (41)

```
> odhadsigma=y2pruh-(t(h)%*%solve(H)%*%h)
        [,1]
[1,] -8.502166
```

$$\widehat{\sigma^2_{\varepsilon_2}} = -8.502166.$$

Odhadnutý model vykreslíme

```
> plot((obeta[1]+obeta[2]*mii+obeta[3]*mii^2)~x, ylab="ydohad",col="red")
```



Obrázek 10: Odhadnutý model z příkladu 3.

Závěr

V práci byla podrobně rozebrána ortogonální regrese, odhady parametrů pomocí metody maximální věrohodnosti pro stejné i rozdílné rozptyly. Připomněli jsme si teorii lineárních modelů s podmínkou typu II, kterou jsme následně využili k odhadu parametrů a stanovení intervalů spolehlivosti. Zabývali jsme se polynomiální regresí, odhadem parametrů pomocí upravené metody nejmenších čtverců.

Odhady parametrů ortogonální regresní přímky počítané pomocí vztahů (3) a (4) dávají stejné výsledky jako počítané pomocí lineárních modelů. Intervaly spolehlivosti pro směrnici ortogonální regresní přímky získané pomocí vztahů (26) a (27) se téměř překrývají a jako nejužší vychází interval spolehlivosti počítaný pomocí vztahu (30), tedy pomocí lineárních modelů. Empirická hladina spolehlivosti odpovídá zvolené hladině spolehlivosti $1 - \alpha$, takže čím vyšší je empirická hladina spolehlivosti, tím jsou intervaly spolehlivosti širší.

Díky této práci jsem si prohloubila znalosti z regresní analýzy a naučila se pracovat se softwarem R. Nejtěžší ze všeho pro mě bylo nastudování a pochopení teorie psané anglickým jazykem, a proto si cením trpělivosti a velké pomoci své vedoucí.

Použitá literatura

[1] Anděl, J., *Statistické metody,* 2. Přepracované vydání, Praha: MATFYZPRESS, 1998.

[2] Donevská S., Fišerová E., Hron K., *On the Equivalence between Orthogonal Regression and Linear Model with Type-II Constraints,* Mathematica **50** , 19-27, 2011.

[3] Fišerová E., Hron K., *Total least squares solution for compositionalusing linear models*, Journal of Applied Statistics **37** , No. 7, 1137–1152, 2010.

[4] Fišerová E., Hron K., *Statistical Inference in Orthogonal Regression for Three-part Compositional Data Using Linear Model with Type-II Constraints,* Communications In Statistics-Theory and Methods, Communications in Statistics – Theory and Methods **41** , No. 13-14, 2367-2385, 2012.

[5] Fišerová, E., Lineární statictické modely, 1. vydání, Olomouc, 2013.

[6] Fišerová, E., Pravděpodobnost a statistika 3 (texty k přednáškám), 2012.

[7] Fuller, W. A., Measurement Error Models, wiley, New York, 1987.

[8] Cheng, Ch.-L.Schneeweiss, H., Polynomial regression with errors in the variables, J. R. Statist. Soc. B, **60**, Par 1 , 189-199, 1998.

[9] Jackson, J. D., Dunlevy, J. A., Orthogonal least squares and the interchangeability of alternative proxy variables in the social sciences, J. Roy. Statist. Soc. D (the statisticion), 37(1): 7-14, 1988.

[10] Jolicoeur, P., Interval estimation of the slope of the major axis of a bivariate normal distribution in the case of a small sample, Biometrics 24: 679-682,1968.

[11] Kendall, M. G., Stuart, A.: The Advanced Theory of Statistics, Vol. 2. *Charles Griffin*, London, 1967.

[12] Kubáčková, L., Foundations of Experimental Data Analysis, CRC press, Boca Raton – Ann Arbor – London – Tokyo, 1992.

[13] Leng, L., Zhang, T., Kleinman, L., Zhu, W., Ordinary Least Square Regression, Orthogonal Regression, Geometric Mean Regression and their Applications in Aerosol Science, Jornal of physics: conference Series 78, 2007.

[14] Schneeweiss, H., Nittner, T., Estimating a polynomial regression with measurement errors in the structural and in the functional case – a comparison, sonderforchungsbereich 386, Paper 197, 2000.

[15] vzdálenost bodu od přímky [online], dostupné z: http://www.matematika.cz/vzdalenost-bod-primka [citováno 10. 11. 2013]

[16] Zvára, K., Koeficient determinace v regresi s chybami v obou proměnných [online], dostupné z: http://www.statspol.cz/robust/1994_zvara_94.pdf [citováno 20. 11. 2013]

Příloha (na přiloženém CD)

Skripty v programu R se všemi výpočty:	Příklad 1
	Příklad 2
	Příklad 3
	Simulace intervalu (26)
	Simulace intervalu (27)
	Simulace intervalu (30)