

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY



KVANTILOVÁ REGRESIA V EKONOMETRII

DIPLOMOVÁ PRÁCA

UNIVERZITA KOMENSKÉHO V BRATISLAVE
FAKULTA MATEMATIKY, FYZIKY A INFORMATIKY

KVANTILOVÁ REGRESIA V EKONOMETRII

DIPLOMOVÁ PRÁCA

Študijný program: Ekonomická a finančná matematika
Študijný odbor: 9.1.9. Aplikovaná matematika
Školiace pracovisko: Katedra aplikovanej matematiky a štatistiky
Vedúci práce: Doc. Mgr. Marián Grendár, PhD.



Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

ZADANIE ZÁVEREČNEJ PRÁCE

Meno a priezvisko študenta: Bc. Lucia Kubalová
Študijný program: ekonomická a finančná matematika (Jednoodborové štúdium, magisterský II. st., denná forma)
Študijný odbor: 9.1.9. aplikovaná matematika
Typ záverečnej práce: diplomová
Jazyk záverečnej práce: slovenský

Názov: Kvantilová regresia v ekonometrii / *Quantile Regression in Econometrics*
Cieľ: Kvantilová regresia (QR, z angl. Quantile Regression) sa dá chápať ako nástroj na prieskum podmieneného rozdelenia pravdepodobnosti. Diplomová práca sa bude venovať: 1) algoritmom na výpočet QR odhadov (simplexová metóda, metóda vnútorného bodu, MM algoritmus), 2) vlastnosti QR odhadov (asymptotická normalita), Waldov test, 3) aplikácii QR na mikroekonomické dáta.

Vedúci: doc. Mgr. Marián Grendár, CSc.
Katedra: FMFI.KAMŠ - Katedra aplikovanej matematiky a štatistiky
Vedúci katedry: prof. RNDr. Daniel Ševčovič, CSc.
Dátum zadania: 25.01.2013

Dátum schválenia: 04.02.2013
prof. RNDr. Daniel Ševčovič, CSc.
garant študijného programu

študent

vedúci práce

Pod'akovanie

Touto cestou sa chcem veľmi pekne poďakovať svojmu vedúcemu diplomovej práce Doc. Mgr. Mariánovi Grendárovi, PhD. za trpezlivosť, ochotu vždy pomôcť a za odborné rady, ktoré mi poskytol pri písaní diplomovej práce.

Abstrakt

KUBALOVÁ, Lucia: Kvantilová regresia v ekonometrii [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky, fyziky a informatiky, Katedra aplikovanej matematiky a štatistiky; školiteľ: Doc. Mgr. Marián Grendár, PhD., Bratislava, 2014, 82 s.

Diplomová práca sa zaoberá kvantilovou regresiou v ekonometrii. Hlavným cieľom práce je predstaviť teóriu kvantilovej regresie, odvodiť optimalizačný problém na odhad parametrov modelu a popísať algoritmy na jeho riešenie. V teórii sa zaoberáme základnými vlastnosťami odhadov parametrov ako sú napríklad vlastnosti ekvivariancie a asymptotická teória. Úloha na výpočet odhadov parametrov kvantilovej regresie sa dá previesť na úlohu lineárneho programovania a vyriešiť ju pomocou simplexovej metódy a metódy vnútorného bodu. Analyzujeme obe metódy a predstavíme si MM algoritmus, ktorý rieši zložitú úlohu kvantilovej regresie pomocou sekvencie jednoduchších úloh, ktoré vo väčšine prípadov, vedú k optimu základnej úlohy. Nakoniec práce analyzujeme Americké mzdové dáta pomocou kvantilovej regresie a podrobne opíšeme dosiahnuté výsledky.

Kľúčové slová: kvantil, asymptotika, simplexová metóda, metóda vnútorného bodu, MM algoritmus

Abstract

KUBALOVÁ, Lucia: Quantile Regression in Econometrics [Master Thesis], Comenius University in Bratislava, Faculty of Mathematics, Physics and Informatics, Department of Applied Mathematics and Statistics; Supervisor: Doc. Mgr. Marián Grendár, PhD., Bratislava, 2014, 82 p.

The master thesis deals with the quantile regression in econometrics. The main objective of the work is to present a basic theory of quantile regression, to define an optimization problem associated with the quantile regression and to introduce some algorithms for the solution of the quantile regression problem. We describe some basic characteristics of the quantile regression estimator, its equivariance properties and some basic asymptotic results. Computation of quantile regression estimators may be formulated as a linear programming problem and solved by the simplex algorithm or the interior point method. We analyse both of this algorithms and introduce an MM algorithm, which replaces the difficult optimalization problem of quantile regresion by a sequence of easier optimization problems, which in the most cases converges to a solution of the original problem. The last part of this thesis applies the quantile regression to an analysis of the American income data.

Keywords: quantile, asymptotic, simplex method, interior point method, MM algorithm

Obsah

Úvod	9
1 Kvantilová regresia	11
1.1 Základné pojmy	11
1.2 Optimalizačná úloha pre kvantilovú regresiu	13
1.3 Vlastnosti odhadu parametrov	15
1.3.1 Vlastnosti ekvivariancie	19
1.3.2 Asymptotická normalita odhadov	21
1.3.3 Waldov test	24
2 Úloha kvantilovej regresie a lineárne programovanie	25
2.1 Simplexová metóda	28
2.1.1 Simplexová metóda a úloha kvantilovej regresie	28
2.2 Metóda vnútorného bodu	30
2.2.1 Newtonova metóda	33
2.2.2 Frisch-Newton algoritmus	35
3 MM algoritmus	39
3.1 MM algoritmus a kvantilová regresia	41
3.2 Zovšeobecnená Newtonova metóda	44
3.2.1 Voľba parametrov a konvergenčné kritérium	49
3.2.2 Zhrenutie MM algoritmu	49
4 Americké mzdové dáta	53
4.1 Popis dát	53
4.2 Kvantilová regresia v systéme R	55
4.3 Odhad parametrov modelu	59
4.3.1 Angristove dáta a porovnanie vybraných premenných	62
Záver	67
Zoznam použitej literatúry	69

Príloha

72

Úvod

Z historického hľadiska pojem regresie pochádza z genetiky a bol vytvorený Francisom Galtonom v 19. storočí. Všimol si, že vysoký rodičia majú tendenciu mať nízke deti a nízky rodičia zase vysoké deti. Jeho priateľ K. Pearson zozbieral štatistické údaje o výške jednotlivých členov rodiny a zistil, že synovia nižších otcov majú tendenciu byť vyšší ako priemer výšky otcov a naopak.

V súčasnosti je pri pojme regresia najviac skloňovaná klasická metóda najmenších štvorcov. Myšlienka metódy najmenších štvorcov je v minimalizovaní súčtu štvorcov odchýlok (rezíduí).

Menej známa je l_1 regresia, ktorá je založená na minimalizácii súčtu absolútnych rezíduí. Často sa jej hovorí aj mediánová regresia. Je zaujímavé, že z historického hľadiska je práve l_1 regresia datovaná v strede 18. storočia R.J. Boškovicom, skoro pol storočia pred publikovaním práce o metóde najmenších štvorcov.

Jednoduchým rozšírením mediánovej regresie sa dostaneme ku kvantilovej regresii. Kvantilovou regresiou modelujeme podmienené kvantily rozdelenia, teda nie len podmienenú strednú hodnotu, ktorú nám ponúka štandardná regresia. Týmto dostávame celkovú informáciu o podmienenom rozdelení. Kvantilová regresia ponúka odhad podmienenej kvantilovej funkcie prostredníctvom minimalizácie po častiach lineárnej účelovej funkcie rezíduí.

V prvej kapitole diplomovej práce si predstavíme základné pojmy kvantilovej regresie a odvodíme optimalizačnú úlohu pre odhad parametrov kvantilovej regresie. Ukážeme si, ako sa dá táto úloha kvantilovej regresie previesť na úlohu lineárneho programovania [14]. Ďalej sa pozrieme na základné vlastnosti odhadov modelu, ich asymptotiku, podmienku optimality a Waldov test na testovanie hypotéz o parametroch modelu.

V druhej kapitole rozoberieme riešenie úlohy kvantilovej regresie práve pomocou teórie lineárneho programovania. Zavedieme si základné pojmy a budeme sa zaoberať dvoma algoritmami, a to simplexovou metódou pre kvantilovú regresiu a metódou vnútorného bodu, špeciálne Frisch-Newtonovým algoritmom.

V tretej kapitole si predstavíme MM algoritmus a jeho súvislosť s viac známym EM algoritmom. Ukážeme si, ako môžeme riešiť úlohu kvantilovej regresie pomocou MM algoritmu. Hlavná časť tejto kapitoly bude spracovaná podľa článku [11]. Na konci ka-

pitoly ilustrujeme prípad pre odhad jedného parametra pomocou MM algoritmu, ktorý naprogramujeme v systéme R.

V poslednej kapitole aplikujeme kvantilovú regresiu na Americké mzdové dáta pre rok 2012. Predstavíme si základné funkcie z knižnice `quantreg` v systéme R na odhad parametrov úlohy kvantilovej regresie. Podrobne si popíšeme výsledky a skúsime ich porovnať s inými rokmi.

1 Kvantilová regresia

Pojem kvantilová regresia zaviedli v roku 1978 Roger Koenker a Gilbert Bassett v článku [14]. V tejto práci predstavili myšlienku, ako získať odhad parametrov pre jednotlivé kvantily z minimalizácie súčtu zošikmených absolútnych odchýlok. Odvodili i podmienku optimality pre kvantilovú regresiu a ukázali asymptotickosť daných odhadov. V roku 2005 publikoval R. Koenker knihu *Quantile regression* s detailnou teóriou o kvantilovej regresii a metódami aplikovanými na príkladoch z ekonómie, financií, biológie či ekológie. Nasledujúca kapitola je spracovaná podľa prvých kapitol tejto knihy [13].

1.1 Základné pojmy

Definícia 1.1. *Nech X označuje reálnu náhodnú premennú, ktorá je charakterizovaná svojou distribučnou funkciou $F(x) = P(X \leq x)$. Nech $\tau \in (0, 1)$, potom*

$$F^{-1}(\tau) = \inf\{x \in \mathbb{R} : F(x) \geq \tau\} \quad (1)$$

sa nazýva kvantilová funkcia a hodnota tejto funkcie je τ -kvantil náhodnej premennej X .

Rovnako ako distribučná funkcia náhodnej premennej X , tak i kvantilová funkcia poskytuje kompletný obraz o príslušnej náhodnej premennej. Z definície (1) je zrejmé, že kvantil sa dá nájsť ako riešenie jednoduchého optimalizačného problému.

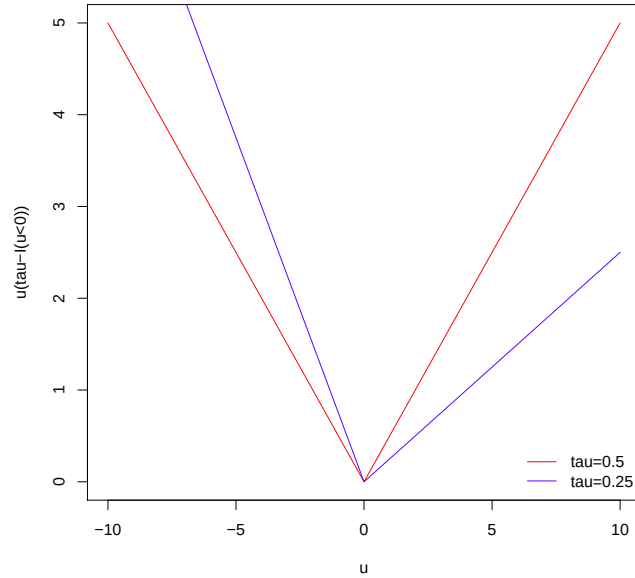
Definujeme po častiach lineárnu účelovú funkciu nazývanú *check function*:

$$\rho_\tau(u) = u(\tau - I(u < 0)), \quad \text{pre } \tau \in (0, 1), \quad (2)$$

kde $I(u < 0)$ je indikátorová funkcia, ktorej hodnota je 1, ak je argument splnený, inak je hodnota 0.

$$I(u \in A) = \begin{cases} 1 & u \in A \\ 0 & u \notin A \end{cases}$$

Sklon *check function* je teda $\tau - 1$ pre $u < 0$ a τ pre $u \geq 0$. Táto funkcia sa dá predstaviť ako *naklonená* absolútna hodnota, ktorej *naklonenie* závisí od voľby $\tau \in (0, 1)$. Pre najznámejší kvantil medián, kde $\tau = 0.5$, je *check function* presne absolútna hodnota.



Obr. 1: Check function pre rôzne hodnoty $\tau \in (0, 1)$.

Hlavná myšlienka, ktorú autori predstavili v článku [14] bola, že navrhli merať stratu pomocou takej stratovej funkcie, za ktorú zvolili práve *check function* $\rho_\tau(u)$ pre $\tau \in (0, 1)$. Minimalizovaním očakávanej straty dostávame:

$$\operatorname{argmin}_{\hat{x} \in \mathbb{R}} E\rho_\tau(X - \hat{x}) = (\tau - 1) \int_{-\infty}^{\hat{x}} (x - \hat{x}) dF(x) + \tau \int_{\hat{x}}^{\infty} (x - \hat{x}) dF(x). \quad (3)$$

Zderivujeme danú funkciu (3) podľa $\hat{x}$ a dostaneme:

$$0 = (1 - \tau) \int_{-\infty}^{\hat{x}} dF(x) - \tau \int_{\hat{x}}^{\infty} dF(x) = F(\hat{x}) - \tau. \quad (4)$$

Pokým je distribučná funkcia $F(x)$ monotónna, každý prvok množiny $\{x : F(x) = \tau\}$ minimalizuje očakávanú stratu (3). Ak existuje inverzná funkcia, tak riešenie je jednoznačné $\hat{x} = F^{-1}(\tau)$. V prípade, že inverzná funkcia neexistuje, tak riešením je interval. Z vlastnosti, že distribučná funkcia je spojitá zľava

$$\lim_{x \rightarrow a^-} F(x) = F(a),$$

vyplýva, že z daného intervalu vyberieme ten najmenší prvok.

Distribučnú funkciu F náhodnej premennej X v praxi často nepoznáme, a preto ju nahradíme empirickou distribučnou funkciou

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(X_i \leq x). \quad (5)$$

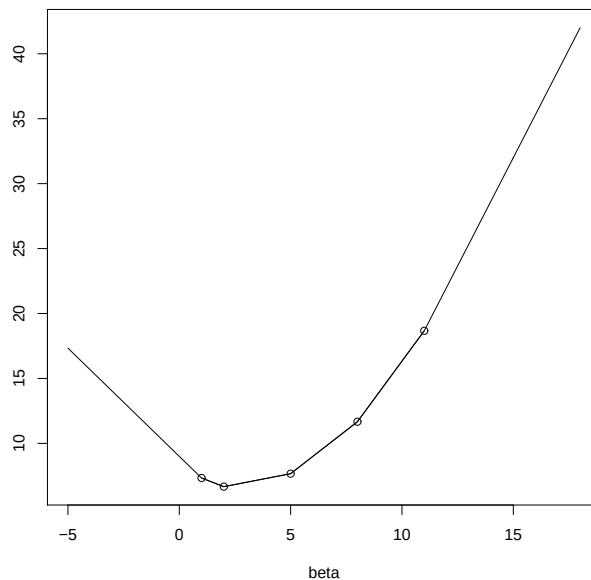
Potom môžeme prepísať funkciu očakávanej straty nasledovne:

$$E\rho_\tau(X - \hat{x}) = \int \rho_\tau(X - \hat{x})dF_n(x) = \frac{1}{n} \sum_{i=1}^n \rho_\tau(x_i - \hat{x}). \quad (6)$$

Úloha na nájdenie τ výberového kvantilu môže byť nakoniec preformulovaná :

$$\min_{\xi \in \mathbb{R}} \sum_{i=1}^n \rho_\tau(x_i - \xi). \quad (7)$$

Je zrejmé, že účelová funkcia, ktorú chceme minimalizovať je po častiach lineárna funkcia.



Obr. 2: Po častiach lineárna účelová funkcia pre voľbu $\tau = \frac{1}{3}$ v bodoch (1,2,5,8,11).

Zistili sme, že kvantily sa dajú vyjadriť ako riešenie jednoduchého optimalizačného problému (7). Prirodzene sa dá toto pozorovanie previesť na všeobecnejšiu metódu odhadov parametrov modelu pre podmienenú kvantilovú funkciu.

1.2 Optimalizačná úloha pre kvantilovú regresiu

Pod jednoduchým modelom lineárnej kvantilovej regresie budeme rozumieť

$$y = X\beta_\tau + \varepsilon_\tau, \quad (8)$$

kde

- $y_{n \times 1}$ je vektor závislých premenných,
- $X_{n \times p}$ sa nazýva *matica plánu*,
- β_τ rozmeru $p \times 1$ je vektor parametrov pre danú hodnotu $\tau \in (0, 1)$,
- ε_τ je vektor chýb, ktorých rozdelenie bližšie nebudeme špecifikovať.

Definícia 1.2. *Nech $(X, Y)^T$ je náhodný vektor a $F_{Y|X}(y|x)$ je jeho podmienená distribučná funkcia. Podmienená kvantilová funkcia pre $\tau \in (0, 1)$ je definovaná:*

$$Q_{Y|X}(\tau|x) = F_{Y|X}^{-1}(\tau|x) = \inf\{x \in \mathbb{R} : F_{Y|X}(y|x) \geq \tau\}.$$

Podmienená kvantilová funkcia pre náš model (8) má tvar:

$$Q_{Y|X}(\tau|x) = x^T \beta_\tau. \quad (9)$$

Odhad parametrov $\hat{\beta}_\tau$ podmienenej kvantilovej funkcie (9) dostaneme riešením optimalizačnej úlohy:

$$\min_{\beta \in \mathbb{R}^p} \sum_{i=1}^n \rho_\tau(y_i - x_i^T \beta). \quad (10)$$

Úloha kvantilovej regresie (10) sa dá riešiť viacerými spôsobmi. My ju budeme riešiť pomocou teórie lineárneho programovania a MM algoritmu.

Ak chceme pristúpiť k riešeniu prostredníctvom lineárneho programovania, musíme si úlohu (10) preformulovať nasledovne. Ako prvé si zavedieme $2n$ nových, umelých, premenných $\{u_i, v_i : 1, \dots, n\}$, ktoré nám budú reprezentovať kladné a záporné časti vektora rezíduí. Potom platí:

$$\begin{aligned} \min \sum_{i=1}^n \rho_\tau(y_i - x_i^T \beta) &= \min \sum_{i=1}^n (\tau u_i + (1 - \tau)v_i) \\ &= \min[\tau 1_n^T u + (1 - \tau)1_n^T v]. \end{aligned}$$

Úlohu (10) môžeme vyjadriť:

$$\min_{(\beta, u, v) \in \mathbb{R}^p \times \mathbb{R}_+^{2n}} \{\tau 1_n^T u + (1 - \tau)1_n^T v | X\beta + u - v = y\}, \quad (11)$$

kde 1_n označuje vektor samých jednotiek dĺžky n . Síce sme sa dostali k úlohe lineárneho programovania, ktorú už vieme riešiť, ale za to budeme platiť nárastom počtu premenných z p na $(p + 2n)$. Metódami, ktoré využijeme na riešenie danej úlohy (11) sa budeme zaoberať v ďalších kapitolách. Teraz si povieme niečo o vlastnostiach odhadov parametra modelu (8).

1.3 Vlastnosti odhadu parametrov

Metóda najmenších štvorcov je založená na minimalizácii súčtu štvorcov rezíduí. Deriváciou účelovej funkcie dostaneme normálne rovnice, ktorých riešením dostaneme odhad parametra $\hat{\beta}$. Tieto normálne rovnice vedú k jednoznačnému riešeniu, ak existuje inverzná matica k matici $X^T X$. Pre úlohu kvantilovej regresie sa pokúsime spraviť podobnú úvahu.

Označme účelovú funkciu kvantilovej regresie

$$R(\beta) = \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta), \quad (12)$$

ktorá, ako vieme, je po častiach lineárna a spojitá. To znamená, že je diferencovateľná, okrem bodov, kde rezíduá $y_i - x_i^T \beta = 0$. V týchto bodoch však existuje derivácia v smere.

Derivácia v smere $w \in \mathbb{R}^p$ účelovej funkcie (12) je:

$$\begin{aligned} \nabla R(\beta, w) &= \frac{d}{dt} R(\beta + tw)|_{t=0} \\ &= \frac{d}{dt} \sum_{i=1}^n (y_i - x_i^T \beta - x_i^T tw) [\tau - (I(y_i - x_i^T \beta - x_i^T tw) < 0)]|_{t=0} \\ &= - \sum_{i=1}^n \psi_{\tau}^*(y_i - x_i^T \beta, -x_i^T w) x_i^T w, \end{aligned}$$

kde

$$\psi_{\tau}^*(u, v) = \begin{cases} \tau - I(u < 0) & u \neq 0, \\ \tau - I(v < 0) & u = 0. \end{cases}$$

Ak platí, že v bode $\hat{\beta}$ je derivácia v smere pre všetky smery nezáporná (t.j., $\nabla R(\hat{\beta}, w) \geq 0$ pre všetky $w \in \mathbb{R}^p$ s normou $\|w\| = 1$), potom $\hat{\beta}$ je bod minima účelovej funkcie $R(\beta)$. Tento výsledok vyplýva z intuície, že bez ohľadu na to, do ktorého smeru sa rozhodneme pohnúť z daného bodu $\hat{\beta}$, tak hodnota účelovej funkcie môže vzrásť. A tým pádom sme v bode minima. Toto minimum však nemusí byť jediné.

Toto tvrdenie vyplýva z faktu, že úloha kvantilovej regresie sa dá previesť na úlohu lineárneho programovania (11). Z teórie lineárneho programovania vieme, že účelovú funkciu minimalizujeme na polyédri, čo je konvexná uzavretá množina [17]. Tiež je známe, že ak množina prípustných riešení obsahuje vrcholy a úloha má optimálne riešenie, tak optimum je jeden z daných vrcholov. To znamená, že úlohu môžeme obmedziť

iba na vrcholy a preveriť, v ktorom bude účelová funkcia neklesajúca pre všetky hrany, ktoré vychádzajú z daného vrcholu. Toto je myšlienka simplexovej metódy, ku ktorej sa vrátíme v nasledujúcej kapitole.

Teraz si predstavíme označenia, ktoré nám budú prospešné pre zavedenie základných vlastností odhadu parametra $\hat{\beta}_\tau$ pre úlohu kvantilovej regresie.

Označme množinu $\mathcal{N} = \{1, 2, \dots, n\}$ a jej p -prvkovú podmnožinu $\mathcal{H} \subset \mathcal{N}$. Nech h označuje prvky množiny $\mathcal{H}$. $X(h)$ označuje podmaticu matice X , s riadkami $\{x_i : i \in h\}$ a $y(h)$ označuje p -prvkový vektor s prvkami $\{y_i : i \in h\}$. Prvky h majú prislúchajúci komplement $\bar{h} = \mathcal{N} \setminus h$ a platí, že matica $X(\bar{h})$ je veľkosti $(n - p) \times p$, matica s riadkami $\{x_i : i \in \bar{h}\}$ a podobne $(n - p)$ -prvkový vektor $y(\bar{h})$ s prvkami $\{y_i : i \in \bar{h}\}$. Vďaka tomuto označeniu môžeme vyjadriť akékoľvek prípustné bázičné riešenie predchádzajúce bodmi $\{(x_i, y_i), i \in h\}$ ako

$$b(h) = X(h)^{-1}y(h). \quad (13)$$

Samozrejme za predpokladu, že matica $X(h)$ je regulárna. Je zjavné, že daných prípustných bázičných riešení je príliš mnoho, a to $\binom{n}{p} = O(n^p)$.

S týmto označením sa vrátíme naspäť k smerovej derivácii, aby sme odvodili podmienku optimality pre úlohu kvantilovej regresie (10). Teraz môžeme upriamiť pozornosť na kandidáta na riešenie z (13). Pre niektoré h sa môže stať, že matica $X(h)$ bude singularná. Tento fakt nám však nemusí robiť starosti, bude nám stačiť, keď sa obmedzíme len na také h , pre ktoré bude príslušná matica $X(h)$ regulárna.

Z predchádzajúceho pozorovania vieme, že podmienka optimality pre (13) si vyžaduje overenie nezápornosti smerovej derivácie vo všetkých smeroch. Platí:

$$\nabla R(b(h), w) = - \sum_{i=1}^n \psi_\tau^*(y_i - x_i^T b(h), -x_i^T w) x_i^T w.$$

Preparametrizujme smer $v = X(h)w$ a dostaneme:

$$\nabla R(b(h), X(h)^{-1}v) = - \sum_{i=1}^n \psi_\tau^*(y_i - x_i^T b(h), -x_i^T X(h)^{-1}v) x_i^T X(h)^{-1}v.$$

Optimalitu máme zaručenú vtedy a len vtedy, keď:

$$0 \leq - \sum_{i=1}^n \psi_\tau^*(y_i - x_i^T b(h), -x_i^T X(h)^{-1}v) x_i^T X(h)^{-1}v,$$

pre všetky smery $v \in \mathbb{R}^p$. Môžeme si všimnúť, že vhodným usporiadaním prvkov h dostaneme $e_i^T = x_i^T X(h)^{-1}$ pre $i \in h$, kde vektor e_i je i -ty bázový vektor. S týmto poznatkom dostávame:

$$\begin{aligned} 0 &\leq - \sum_{i \in h} \psi_\tau^*(0, -v_i) v_i - \sum_{i \in \bar{h}} \psi_\tau^*(y_i - x_i^T b(h), -x_i^T X(h)^{-1} v) x_i^T X(h)^{-1} v \\ &= - \sum_{i \in h} \psi_\tau^*(0, -v_i) v_i - \xi^T v \\ &= - \sum_{i \in h} (\tau - I(-v_i < 0)) v_i - \xi^T v \end{aligned} \quad (14)$$

kde

$$\xi^T = \sum_{i \in \bar{h}} \psi_\tau^*(y_i - x_i^T b(h), -x_i^T X(h)^{-1} v) x_i^T X(h)^{-1} v$$

Definícia 1.3. *Hovoríme, že pozorovania (y, X) sú v netriviálnej pozícii, ak pre nejaké h platí:*

$$y_i - x_i^T b(h) \neq 0 \quad \forall i \notin h.$$

Veta 1.4. *Ak sú pozorovania (y, X) v netriviálnej pozícii, potom existuje riešenie pre úlohu kvantilovej regresie (10) v tvare $b(h) = X(h)^{-1} y(h)$ vtedy a len vtedy, keď pre nejaké h platí:*

$$-\tau 1_p \leq \xi(h) \leq (1 - \tau) 1_p, \quad (15)$$

kde $\xi^T(h) = \sum_{i \in \bar{h}} \psi_\tau(y_i - x_i^T b(h)) x_i^T X(h)^{-1}$ a $\psi_\tau = \tau - I(u < 0)$. Navyše platí, že riešenie $b(h)$ je jediné vtedy a len vtedy, keď máme ostré nerovnosti v (15). Inak platí, že množina riešení je konvexný obal niektorých prípustných riešení v tvare $b(h)$.

Dôkaz. Môžeme si všimnúť, že priestor smerov $v \in \mathbb{R}^p$ je lineárnym obalom smerov $v = \pm e_i$, kde $(i = 1, \dots, p)$. To teda znamená, že podmienka pre smerovú deriváciu platí pre všetky smery $v \in \mathbb{R}^p$ len vtedy, keď platí pre $2p$ všeobecné smery $v = \pm e_i$.

Ak zoberieme za smer $v = e_i$, $i = 1, \dots, p$ z nerovnosti (14) dostávame:

$$-(\tau - 1) - \xi_i(h) \geq 0 \quad i = 1, \dots, p.$$

Pre $v = -e_i$ máme:

$$\tau + \xi_i(h) \geq 0 \quad i = 1, \dots, p.$$

Kombinácia týchto dvoch nerovností a z faktu, že (y, X) sú v netriviálnej pozícii platí:

$$-\tau 1_p \leq \xi(h) \leq (1 - \tau) 1_p.$$

Týmto sme dokázali prvú časť tvrdenia. Druhú časť o jednoznačnosti riešenia $b(h)$ dokážeme nasledovne jednoduchou úvahou. Vieme, že účelová funkcia (12) je konvexná, to znamená, že jednoznačné riešenie dosahuje vtedy, ak pre smerovú deriváciu platí:

$$\nabla R(b(h), w) > 0, \quad \forall w \neq 0_p.$$

To znamená, že do ktoréhokoľvek smeru sa pohnem, hodnotu účelovej funkcie len zvýšim. Ďalej už pokračujeme tak ako na začiatku tohto dôkazu, len všade namiesto nerovnosti bude ostrá nerovnosť. $\square$

Veta 1.5. *Nech P , N a Z označujú počet kladných, záporných a nulových prvkov vektora rezíduí $y - X\hat{\beta}_\tau$. Ak matica plánu obsahuje intercept (t.j., ak existuje taký vektor $\alpha \in \mathbb{R}^p$, že $X\alpha = 1_n$), potom platí:*

$$N \leq n\tau \leq N + Z$$

a

$$P \leq n(1 - \tau) \leq P + Z.$$

Dôkaz. Vieme, že $\hat{\beta}_\tau$ je optimálne riešenie, ak je splnená podmienka

$$-\sum_{i=1}^n \psi_\tau^*(y_i - x_i^T \hat{\beta}_\tau, -x_i^T w) x_i^T w \geq 0$$

pre všetky smery $w \in \mathbb{R}^p$. Potom musí platiť táto podmienka aj pre $\alpha \in \mathbb{R}^p$ spĺňajúce $X\alpha = 1_n$. Tým dostávame:

$$-\sum_{i=1}^n \psi_\tau^*(y_i - x_i^T \hat{\beta}_\tau, -1) \geq 0,$$

čo vedie k nerovnosti:

$$\tau P - (1 - \tau)N - (1 - \tau)Z \leq 0. \quad (16)$$

Podobne, ak za smer zvolíme $w = -\alpha$ získame nerovnosť:

$$-\tau P + (1 - \tau)N - \tau Z \leq 0. \quad (17)$$

Využijeme fakt, že súčet $P + N + Z = n$. Do nerovnice (16) dosadíme postupne $N = n - P - Z$ a $P = n - N - Z$. Z prvej substitúcie priamo získame nerovnosť $P \leq n(1 - \tau)$ a z druhej substitúcie $\tau n \leq N + Z$.

Podobne do druhej nerovnosti (17) postupne dosadíme $Z = n - N - P$ a $N = n - P - Z$ a dostaneme ďalšie dve nerovnosti $N \leq \tau n$ a $n(1 - \tau) \leq P + Z$. Spojením týchto štyroch výsledkov je dôkaz hotový. $\square$

Dôsledok 1.6. Ak $Z=p$, potom podiel záporných rezíduí je približne τ

$$\frac{N}{n} \leq \tau \leq \frac{N + p}{n},$$

a podiel kladných rezíduí je približne $(1 - \tau)$

$$\frac{P}{n} \leq 1 - \tau \leq \frac{P + p}{n}.$$

1.3.1 Vlastnosti ekvivariancie

Predpokladajme, že máme model, v ktorom je premenná *teplota* meraná vo Fahrenheitoch a my chceme prejsť k inej jednotke, ako sú napríklad stupne Celzia. Očakávame, že táto zmena v jednotkách nebude mať zásadný vplyv na náš odhad parametrov modelu. Označme odhad parametra modelu založený na pozorovaní (y, X) pre pevné $\tau \in (0, 1)$ ako $\hat{\beta}(\tau; y, X)$. V nasledujúcej vete (Koenker a Bassett, 1978) sú zhromaždené štyri základné vlastnosti ekvivariancie pre $\hat{\beta}(\tau; y, X)$.

Veta 1.7. Nech matica A veľkosti $p \times p$ je regulárna, $\gamma \in \mathbb{R}^p$ a $a > 0$. Potom pre každé $\tau \in (0, 1)$ platí:

- (i) $\hat{\beta}(\tau; ay, X) = a\hat{\beta}(\tau; y, X)$
- (ii) $\hat{\beta}(\tau; -ay, X) = -a\hat{\beta}(1 - \tau; y, X)$
- (iii) $\hat{\beta}(\tau; y + X\gamma, X) = \hat{\beta}(\tau; y, X) + \gamma$
- (iv) $\hat{\beta}(\tau; y, XA) = A^{-1}\hat{\beta}(\tau; y, X)$.

Dôkaz. Účelovú funkciu (12) môžeme prepísať nasledovne:

$$R(\beta; \tau, y, X) = \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta) = \sum_{i=1}^n \left[\tau - \frac{1}{2} + \frac{1}{2} \operatorname{sgn}(y_i - x_i^T \beta) \right] (y_i - x_i^T \beta),$$

kde funkcia

$$\operatorname{sgn}(x) = \begin{cases} -1 & x \leq 0 \\ 0 & x = 0 \\ 1 & x \geq 0. \end{cases}$$

Teraz si môžeme všimnúť, že platí:

- (i) $\lambda R(\beta; \tau, y, X) = R(\lambda\beta; \tau, \lambda y, X), \quad \lambda \in \langle 0, \infty \rangle$
- (ii) $-\lambda R(\beta; \tau, y, X) = R(\lambda\beta; 1 - \tau, \lambda y, X), \quad \lambda \in (-\infty, 0)$
- (iii) $R(\beta; \tau, y, X) = R(\beta + \gamma; \tau, y + X\gamma, X), \quad \gamma \in \mathbb{R}^p$
- (iv) $R(\beta; \tau, y, X) = R(A^{-1}\beta; \tau, y, XA), \quad |A_{p \times p}| \neq 0.$

A týmto je veta dokázaná. □

Poznámka: Všimnime si, že vlastnosti (i) a (ii) sú *ekvivariancia škály*, vlastnosť (iii) je *ekvivariancia posunu* a vlastnosť (iv) je *ekvivariancia preparametrizovania* matice plánu.

Za predpokladu, že matica plánu obsahuje intercept, náš príklad so zmenou jednotiek teploty sa dá ľahko pretransformovať z parametra $\hat{\beta}(\tau; y, X)$ vo Fahrenheitoch na $\frac{5}{9}(\hat{\beta}(\tau; y, X) - 32\gamma)$ v stupňoch Celzia. V predchádzajúcom predstavuje γ podmienku interceptu ($X\gamma = 1_n$). Pre voľbu γ je typické zobrať prvý bázičný vektor e_1 .

Ďalšou užitočnou vlastnosťou, ktorou disponuje kvantilová regresia je *ekvivariancia na monotónnu transformáciu*. Nech $h(\cdot)$ je neklesajúca funkcia na $\mathbb{R}$. Potom pre náhodnú premennú Y platí:

$$Q_{h(Y)}(\tau) = h(Q_Y(\tau)).$$

Táto vlastnosť nám hovorí len to, že kvantily transformovanej náhodnej premennej $h(Y)$ sú jednoducho transformované kvantily pôvodnej náhodnej premennej Y . Táto vlastnosť vyplýva z tvrdenia, že pre monotónnu funkciu $h(\cdot)$ platí:

$$P(Y \leq y) = P(h(Y) \leq h(y)).$$

Avšak, pre strednú hodnotu táto vlastnosť neplatí:

$$Eh(Y) \neq h(E(Y)).$$

1.3.2 Asymptotická normalita odhadov

Veta 1.8. (Koenker a Bassett, 1978)

Uvažujme lineárny model

$$y_i = x_i^T \beta + u_i \quad i = 1, 2, \dots, n$$

s nezávislými, rovnako rozdelenými chybami $\{u_i\}$, ktoré majú spoločnú distribučnú funkciu F a pozitívnu hustotu f pre $F^{-1}(\tau)$. Potom hustota odhadu $\hat{\beta}_\tau$ má tvar:

$$g(b) = \sum_{h \in \mathcal{H}} P\{\xi_h(b) \in \mathcal{C}\} |X(h)| \prod_{i \in h} f(x_i^T(b - \beta_\tau) + F^{-1}(\tau)), \quad (18)$$

kde $\xi_h(b) = \sum_{i \in h} \psi_\tau(y_i - x_i^T b) x_i^T X(h)^{-1}$ a množina $\mathcal{C}$ označuje kocku $[\tau - 1, \tau]^p$.

Dôkaz. Z vety (1.4) platí, že $\hat{\beta}_\tau = b(h) \equiv (X(h))^{-1}y(h)$ vtedy a len vtedy, keď $\xi_h(b(h)) \in \mathcal{C}$. Nech $B(b, \delta) = b + [-\delta/2, \delta/2]^p$ pre každé $b \in \mathbb{R}^p$. $B(b, \delta)$ je kocka so stredom b s hranami o dĺžke δ . Platí:

$$\begin{aligned} P\{\hat{\beta}_\tau \in B(b, \delta)\} &= \sum_{h \in \mathcal{H}} P\{b(h) \in B(b, \delta), \xi_h(b(h)) \in \mathcal{C}\} \\ &= \sum_{h \in \mathcal{H}} EI(b(h) \in B(b, \delta)) P\{\xi_h(b(h)) \in \mathcal{C} | y(h)\}. \end{aligned}$$

Vydelíme obe strany objemom $B(b, \delta) = \delta^p$ a pošleme $\delta \rightarrow 0$, aby sme dostali tvar hustoty:

$$g(b) = \lim_{\delta \rightarrow 0} \frac{P\{\hat{\beta}_\tau \in B(b, \delta)\}}{\delta^p} = \lim_{\delta \rightarrow 0} \sum_{h \in \mathcal{H}} \frac{EI(b(h) \in B(b, \delta))}{\delta^p} P\{\xi_h(b(h)) \in \mathcal{C} | y(h)\}.$$

Prvý člen sumy ide k združenej hustote vektora $X(h)^{-1}y(h)$, ktorú môžeme prepísať podľa:

$$f_{y(h)}(y) = \prod_{i \in h} f(y_i - x_i^T \beta),$$

na

$$\begin{aligned} f_{X(h)^{-1}y(h)}(b) &= |X(h)| \prod_{i \in h} f((X(h)b)_i - x_i^T \beta_\tau + F^{-1}(\tau)) \\ &= |X(h)| \prod_{i \in h} f(x_i^T(b - \beta_\tau) + F^{-1}(\tau)). \end{aligned}$$

Môžeme si všimnúť, že pre také h , kde $X(h)$ je singulárna, k hustote sa nič nepridáva.

Druhý člen, a to podmienená pravdepodobnosť ide k $P\{\xi_h(b) \in \mathcal{C}\}$.

Spojením oboch častí výsledkov dostaneme danú hustotu $g(b)$. □

Avšak z praktického hľadiska je sčítancov pre danú hustotu $\binom{n}{p}$, čo nie je veľmi šikovné v aplikáciách. Preto sa musíme vydať cestou, ktorou sa vybrali v teórii najmenších štvorcov a spraviť asymptotickú aproximáciu pre teóriu kvantilovej regrese. Predpokladajme, že $y_1, y_2, \dots, y_n$ sú rovnako rozdelené a nezávislé náhodné premenné so spoločnou distribučnou funkciou F . Predpokladajme ešte, že F má spojitú hustotu f v blízkosti $\xi_\tau = F^{-1}(\tau)$ s $f(\xi_\tau) > 0$. Účelová funkcia pre τ -výberový kvantil má tvar:

$$\hat{\xi}_\tau \equiv \inf_{\xi} \left(\xi \in \mathbb{R} \mid n^{-1} \sum_{i=1}^n \rho_\tau(y_i - \xi) = \min! \right).$$

Daná suma je súčet konvexných funkcií, teda množina je tiež konvexná. Príslušný gradient má tvar:

$$g_n(\xi) = n^{-1} \sum_{i=1}^n (I(y_i < \xi) - \tau),$$

a je monotónny v ξ . Samozrejme, pre niektoré kvantily platí, že $y_i = \xi$. V tomto prípade potrebujeme prejsť k smerovej derivácii, ktorú sme už rozoberali v predchádzajúcich častiach. Toto však nie je dôležité pre nasledujúcu úvahu.

Funkcia g_n je monotónne rastúca v ξ a z monotónnosti vyplýva, že $\hat{\xi}_\tau$ je väčšie ako ξ vtedy a len vtedy, keď $g_n(\xi) < 0$ a teda:

$$\begin{aligned} P\{\sqrt{n}(\hat{\xi}_\tau - \xi_\tau) > \delta\} &= P\{g_n(\xi_\tau + \delta/\sqrt{n}) < 0\} \\ &= P\{n^{-1} \sum (I(y_i < \xi_\tau + \delta/\sqrt{n}) - \tau) < 0\} \end{aligned}$$

Môžeme si všimnúť, že sčítance nadobúdajú hodnotu $(1 - \tau)$ s pravdepodobnosťou $F(\xi_\tau + \delta/\sqrt{n})$ a $-\tau$ s pravdepodobnosťou $1 - F(\xi_\tau + \delta/\sqrt{n})$. Dostali sme alternatívne rozdelenie. Môžeme použiť DeMoivreovu-Laplaceovu centrálnu limitnú vetu a popísať asymptotické správanie $\hat{\xi}_\tau$.

Stredná hodnota má tvar:

$$E(g_n(\xi_\tau + \delta/\sqrt{n})) = (F(\xi_\tau + \delta/\sqrt{n}) - \tau) \rightarrow_{n \rightarrow \infty} f(\xi_\tau)\delta/\sqrt{n}$$

a disperzia:

$$D(g_n(\xi_\tau + \delta/\sqrt{n})) = F(\xi_\tau + \delta/\sqrt{n})(1 - F(\xi_\tau + \delta/\sqrt{n}))/n \rightarrow_{n \rightarrow \infty} \tau(1 - \tau)/n.$$

Označme množinu $\omega^2 = \tau(1 - \tau)/f^2(\xi_\tau)$ a použitím DeMoivreovej-Laplaceovej centrálnej limitnej vety dostávame:

$$P\{\sqrt{n}(\hat{\xi}_\tau - \xi_\tau) > \delta\} = P\left\{\frac{g_n(\xi_\tau + \delta/\sqrt{n}) - f(\xi_\tau)\delta/\sqrt{n}}{\sqrt{\tau(1 - \tau)/n}} < -\omega^{-1}\delta\right\} \\ \rightarrow 1 - \Phi(\omega^{-1}\delta),$$

kde $\Phi(x)$ je distribučná funkcia normovaného normálneho rozdelenia. Z toho dostávame asymptotické rozdelenie pre $\hat{\xi}_\tau$

$$\sqrt{n}(\hat{\xi}_\tau - \xi_\tau) \rightsquigarrow \mathcal{N}(0, \omega^2).$$

Rozšírme predchádzajúcu úvahu na združené rozdelenie pre niekoľko kvantilov. Označme $\hat{\zeta}_n = (\hat{\xi}_{\tau_1}, \hat{\xi}_{\tau_2}, \dots, \hat{\xi}_{\tau_m})$ a $\zeta_n = (\xi_{\tau_1}, \xi_{\tau_2}, \dots, \xi_{\tau_m})$ a dostávame:

$$\sqrt{n}(\hat{\zeta}_\tau - \zeta_\tau) \rightsquigarrow \mathcal{N}(0, \Omega),$$

kde Ω je kovariančná matica $\Omega(\tau_1, \dots, \tau_m, F)$ veľkosti $m \times m$ s prvkami $\omega_{ij} = \frac{\tau_i(1 - \tau_j)}{(f(F^{-1}(\tau_i))f(F^{-1}(\tau_j)))}$, pre $i \leq j$. Analógia medzi výberovými kvantilmi v *one-sample* modeli a regresnými kvantilmi v lineárnom modeli je veľmi podobná v ich asymptotickom správaní. Nasledujúca veta nám to explicitne uvádza.

Veta 1.9. *Predpokladajme klasický lineárny regresný model*

$$y_i = x_i^T \beta + u_i,$$

kde chyby $\{u_i\}$ sú nezávislé, rovnako rozdelené a majú spoločnú distribučnú funkciu F s príslušnou hustotou f , pre ktorú platí $f(F^{-1}(\tau_i)) > 0$ pre $i = 1, \dots, m$. Predpokladajme, že $Q_n \equiv n^{-1} \sum x_i x_i^T$ konverguje ku kladne definitnej matici Q_0 . Potom združené asymptotické rozdelenie odhadov kvantilovej regresie $\hat{\zeta}_n = (\hat{\beta}_n(\tau_1)^T, \dots, \hat{\beta}_n(\tau_m)^T)^T$ má tvar:

$$\sqrt{n}(\hat{\zeta}_n - \zeta) = (\sqrt{n}(\hat{\beta}_n(\tau_j) - \beta(\tau_j)))_{j=1}^m \rightsquigarrow \mathcal{N}(0, \Omega \otimes Q_0^{-1}).$$

Dôkaz. Formálny dôkaz môžeme nájsť v ([14]). □

Pre prípad, keď chyby merania nie sú nezávislé a rovnako rozdelené je asymptotická teória pre β_τ trochu komplikovanejšia. Avšak platí [13] :

$$\sqrt{n}(\hat{\beta}_\tau - \beta_\tau) \rightsquigarrow \mathcal{N}(0, \tau(1 - \tau)H_n^{-1}J_nH_n^{-1}),$$

kde

$$J_n(\tau) = n^{-1} \sum_{i=1}^n x_i x_i^T$$

a

$$H_n(\tau) = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n x_i x_i^T f_i(\xi_i(\tau)).$$

1.3.3 Waldov test

Klasická teória lineárnej regresie predpokladá, že podmienené kvantilové funkcie odozvy y , zodpovedajúce premenným x , sú navzájom rovnobežné. To znamená, že odhad koeficientov pre sklon je rovnaký na všetkých kvantiloch. V praxi však sú odhady sklonu kvantilovej regresie rôzne rozhádzané. Jeden zo základných problémov kvantilovej regresie zahŕňa i testovanie rovnosti pre parametre sklonu naprieč jednotlivých kvantilov.

Uvažujme zovšeobecnenú lineárnu hypotézu pre vektor parametrov $\zeta = (\beta(\tau_1)^T, \dots, \beta(\tau_m)^T)^T$ v tvare:

$$H_0 : R\zeta = r$$

s testovacou štatistikou

$$T_n = n(R\hat{\zeta} - r)^T [RV^{-1}R^T]^{-1} (R\hat{\zeta} - r),$$

kde V_n je $mp \times mp$ matica s ij blokmi

$$V_n(\tau_i, \tau_j) = [\tau_i(1 - \tau_j)] H_n(\tau_i)^{-1} J_n(\tau_i, \tau_j) H_n(\tau_j)^{-1}.$$

Matice $H_n(\tau)$ a $J_n(\tau)$ sme definovali v predchádzajúcej časti.

Testovacia štatistika T_n je asymptoticky χ_q^2 rozdelená za platnosti hypotézy H_0 , kde počet stupňov voľnosti q je hodnota matice R .

Táto formulácia nám dáva do rúk širokú škálu testov, od jednoduchých testov pre samostatné koeficienty kvantilovej regresie, až po zložité, ktoré zahŕňajú mnoho premenných na rozličných kvantiloch.

V ďalších dvoch kapitolách sa vrátíme k optimalizačnej úlohe kvantilovej regresie (10) a predstavíme metódy na jej riešenie.

2 Úloha kvantilovej regresie a lineárne programovanie

Predpokladáme, že čitateľ je oboznámený so všeobecnou teóriou lineárneho programovania. Základné pojmy lineárneho programovania môžeme nájsť v literatúre [17].

Pod úlohou lineárneho programovania chápeme optimalizovanie lineárnej účelovej funkcie pri lineárnych obmedzeniach. Označme maticu $A \in \mathbb{R}^{m \times n}$, vektory $c, x \in \mathbb{R}^n$ a vektory $b, y \in \mathbb{R}^m$.

Berman vo svojej knihe [3] uvádza, že sústava lineárnych nerovností sa dá previesť na sústavu lineárnych rovností na vhodne definovaných konvexných kužeľoch. Konvexný kužeľ definujeme ako množinu $K \subseteq \mathbb{R}^n$, ktorá je uzavretá na:

- násobenie nezáporným skalárom, $aK \subseteq K$ pre $a \geq 0$
- konvexnú kombináciu, $aK + (1 - a)K \subseteq K$ pre $a \in (0, 1)$.

Bude užitočné zavedenie primárnej úlohy lineárneho programovania vo všeobecnejšom tvare:

$$\min_x \{c^T x \mid Ax - b \in T, x \in S\}, \quad (19)$$

kde množiny T a S sú ľubovoľné uzavreté konvexné kužele.

K danej primárnej úlohe (19) zodpovedá duálna:

$$\max_y \{b^T y \mid c - A^T y \in S^*, y \in T^*\}, \quad (20)$$

kde množiny T^* a S^* sú príslušné duálne množiny k T a S . Pričom duálna množina pre konvexný kužeľ K je definovaná:

$$K^* = \{y \in \mathbb{R}^n \mid x^T y \geq 0, x \in K\}.$$

To znamená, že napríklad, ak $K = \mathbb{R}^n$, tak $K^* = 0_n$ alebo ak $K = \mathbb{R}_+^n$, tak $K^* = \mathbb{R}_+^n$ a ich karteziánsky súčin.

Veta 2.1. (*Veta o dualite*). *Majme dvojicu navzájom duálnych úloh. Potom platí jedno z nasledujúcich tvrdení:*

1. Obe úlohy majú optimálne riešenie a hodnoty účelových funkcií sa v danom optimálnom riešení rovnajú.
2. Jedna z úloh nemá prípustné riešenie a druhá úloha má neohraničené prípustné riešenie.
3. Ani jedna z daných úloh nemá prípustné riešenie.

Dôkaz. Dôkaz môžeme nájsť v [17]. □

Veta 2.2. (*Veta o komplementarite*). Prípustné riešenie x primárnej úlohy a y duálnej úlohy sú optimálne riešenia vtedy a len vtedy, keď platia nasledujúce podmienky komplementarity:

$$\begin{aligned} y_i \left(\sum_{j=1}^n a_{ij} x_j - b_i \right) &= 0, \quad \forall i = 1, \dots, m \\ \left(\sum_{i=1}^m a_{ij} y_i - c_j \right) x_j &= 0, \quad \forall j = 1, \dots, n \end{aligned}$$

Dôkaz. Dôkaz môžeme nájsť v [17]. □

V predchádzajúcej kapitole sme sa dozvedeli, že úloha lineárnej kvantilovej regresie sa dá previesť na úlohu lineárneho programovania (11). Vráťme sa k úlohe lineárnej kvantilovej regresie:

$$\min_{b \in \mathbb{R}^p} \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T b),$$

ktorú môžeme teda prepísať na úlohu:

$$\min_{(b, u, v) \in \mathbb{R}^p \times \mathbb{R}_+^{2n}} \{ \tau 1_n^T u + (1 - \tau) 1_n^T v \mid y - Xb = u - v, b \in \mathbb{R}^p, (u, v) \in \mathbb{R}_+^{2n} \}. \quad (21)$$

Môžeme si všimnúť, že úloha (21) má tvar Bermanovej úlohy z (19), kde:

- $c = (0_p^T, \tau 1_n^T, (1 - \tau) 1_n^T)^T$
- $x = (b^T, u^T, v^T)^T$
- $A = [X, I, -I]$
- $b = y$

- $T = 0_n$
- $S = \mathbb{R}^p \times \mathbb{R}_+^{2n}$.

K primárnej úlohe (21) existuje duálna úloha, ktorá má tvar:

$$\max_{d \in \mathbb{R}^n} \{y^T d \mid X^T d = 0, d \in [\tau - 1, \tau]^n\}. \quad (22)$$

Množinu obmedzení sme získali pomocou:

$$\begin{pmatrix} 0_p \\ \tau 1_n \\ (1 - \tau)1_n \end{pmatrix} - \begin{pmatrix} X^T \\ I_n \\ -I_n \end{pmatrix} d \in 0_p \times \mathbb{R}_+^{2n}.$$

Preparametrizujeme duálnu úlohu pre $a = d + (1 - \tau)1_n$ a dostaneme nový problém:

$$\max_{a \in \mathbb{R}^n} \{y^T a \mid X^T a = (1 - \tau)X^T 1_n, a \in [0, 1]^n\}. \quad (23)$$

Vidíme, že sme dostali duálnu úlohu, ktorej vektor riešení a leží v jednotkovej n -rozmernej kocke.

Ako sme už avizovali v prvej kapitole, stačí sa pozerať na vrcholy množiny obmedzení, pretože ak existuje optimálne riešenie danej úlohy, tak týmto riešením je jeden z vrcholov. Tieto vrcholy sme indexovali pomocou bázičných prvkov h množiny $\mathcal{H}$ a vyjadrili sme ich ako:

$$b(h) = X(h)^{-1}y(h).$$

Pre dané prvky platí:

$$\begin{aligned} u(h) &= v(h) = 0 \\ u(\bar{h}) &= (y - Xb(h))^+ \\ v(\bar{h}) &= (y - Xb(h))^- \end{aligned}$$

Z vety o dualite [17] vyplýva, že pre optimálne riešenie $\{(\hat{b}, \hat{u}, \hat{v}), \hat{d}\}$ musí platiť:

$$\tau 1_n^T \hat{u} + (1 - \tau) 1_n^T \hat{v} = y^T \hat{d}.$$

Optimálne $\hat{d} = (\hat{d}(h), \hat{d}(\bar{h}))$ dostaneme postupne tak, že pre pozorovania $i \in \bar{h}$ z podmienok komplementarity platí:

$$\hat{d}_i = \begin{cases} \tau & \hat{u}_i > 0 \\ \tau - 1 & \hat{v}_i > 0 \end{cases}$$

2.1 Simplexová metóda

a ak $i \in h$, tak $\hat{d}(h)$ určíme nasledovne z obmedzenia $X^T d = 0$:

$$X^T d = [X^T(h), X^T(\bar{h})] \begin{bmatrix} d(h) \\ d(\bar{h}) \end{bmatrix} = X^T(h)d(h) + X^T(\bar{h})d(\bar{h}) = 0$$

$$\hat{d}(h) = -[X(h)^T]^{-1} X(\bar{h})^T \hat{d}(\bar{h}).$$

Riešenie $\hat{d}(h)$ duálnej úlohy pre bázičné riešenie h zodpovedá vektoru $-\xi(h)$ z vety (15). Vidíme, že z obmedzenia $\hat{d}(h) \in [\tau - 1, \tau]^p$ je splnená i podmienka optimality.

2.1 Simplexová metóda

Simplexovú metódu predstavil Dantzing v roku 1947 [17]. Je to základná metóda, ktorá sa používa na riešenie úloh lineárneho programovania. Predpokladá, že množina prípustných riešení obsahuje vrcholy. Za predpokladu, že účelovú funkciu minimalizujeme, odštartujeme z niektorého z prípustného vrcholu a nájdeme, pokiaľ existuje, takú hranu, ktoré vedie do menších hodnôt účelovej funkcie. Ak bola daná hrana úsečkou, tak sa dostaneme do ďalšieho bodu. Ďalej iteratívne pokračujeme, pokiaľ už neexistuje pri danom vrchole hrana, ktorá by nás doviedla do menších hodnôt účelovej funkcie. Potom sme v optimálnom riešení a algoritmus končí. Môže sa však stať, že pri danom vrchole existuje hrana, ktorá by nás mohla doviesť do menších hodnôt účelovej funkcie, ale táto hrana nie je úsečka, ale polpriamka. Potom dostávame neohraničené riešenie.

2.1.1 Simplexová metóda a úloha kvantilovej regresie

V roku 1974 predstavili Barrodale a Roberts v článku [2] modifikovaný simplexov algoritmus na výpočet parametrov mediánovej regresie. Neskôr v článku [15] Koenker a d'Orey tento algoritmus zovšeobecnil na odhad parametrov pre všetky kvantily.

V niektorých úlohách lineárneho programovania nie je jednoduché určiť prípustný počiatočný vrchol. Avšak pre prípad kvantilovej regresie zvolíme za prípustný počiatočný vrchol ktorúkoľvek množinu prvkov h , ktorá spĺňa regularitu matice $X(h)$. Označme danú množinu prvkov pre prípustný počiatočný vrchol h_0 .

Pre modifikovaný simplexov algoritmus využijeme poznatky, ktoré sme nadobudli v

2.1 Simplexová metóda

prvej kapitole. Uvažujme účelovú funkciu kvantilovej regresie

$$R(\beta) = \sum_{i=1}^n \rho_{\tau}(y_i - x_i^T \beta),$$

a jej smerovú deriváciu v bode $b(h_0) = X(h_0)^{-1}y(h_0)$ a v príslušnom smere δ

$$\nabla R(b(h_0), \delta) = - \sum_{i=1}^n \psi_{\tau}^*(y_i - x_i^T b(h_0), -x_i^T \delta) x_i^T \delta.$$

V prvej kapitole sme si ukázali, že pre optimálne riešenie platí, nezápornosť smerovej derivácie vo všetkých smeroch (t.j., $\nabla R(\hat{b}, \delta) \geq 0$ pre všetky $\delta \in \mathbb{R}^p$).

Z dôvodu, že sa už nachádzame v prípustnom vrchole pre prvky h_0 , môžeme zúžiť výber smeru na smer po hrane, ktorá vychádza z daného vrcholu. Túto hranu môžeme vyjadriť parametricky nasledovne:

$$\eta(a, h_0, \delta_j) = b(h_0) + a\delta_j,$$

kde δ_j predstavuje j -ty stĺpec matice $X(h)^{-1}$ a a je skalár, ktorý má za úlohu kontrolovať, kde na hrane sa nachádzame.

V predchádzajúcej časti sme sa dozvedeli, že vektor riešení duálnej úlohy $\hat{d}(h)$ zodpovedá vektoru $-\xi(h)$ z vety (1.4)

$$\hat{d}(h) = -\xi(h) = -[X(h)^{-1}]^T \sum_{i \in \bar{h}} x_i \psi_{\tau}(y_i - x_i^T b(h)),$$

kde

$$\psi_{\tau} = \tau - I(u < 0).$$

Nech $d(h)$ má tvar:

$$d(h) = -[X(h)^{-1}]^T \sum_{i \in \bar{h}} x_i \psi_{\tau}(y_i - x_i^T b(h)).$$

Ak δ_j predstavuje j -ty stĺpec matice $X(h)^{-1}$, tak pre $i \in h$ platí, že $x_i^T \delta_j = 0$ ak $j \neq i$.

Pre $j = i$ je $x_i^T \delta_j = 1$, tým pádom pre takýto smer δ_j platí :

$$\begin{aligned} \nabla R(b(h_0), \delta_j) &= - \sum_{i=1}^n \psi_{\tau}(y_i - x_i^T b(h_0)) x_i^T \delta_j \\ &= - \sum_{i \in h} \psi_{\tau}(y_i - x_i^T b(h_0)) x_i^T \delta_j - \sum_{i \in \bar{h}} \psi_{\tau}(y_i - x_i^T b(h_0)) x_i^T \delta_j \\ &= 1 - \tau + d_j(h) \end{aligned}$$

2.2 Metóda vnútorného bodu

a podobne pre $-\delta_j$:

$$\nabla R(b(h_0), -\delta_j) = \tau - d_j(h).$$

Ak sú tieto smerové derivácie nezáporné pre všetky $j = 1, \dots, p$, tak potom máme splnenú podmienku optimality pre úlohu kvantilovej regresie $d(h) \in [\tau - 1, \tau]^p$.

Ak nie sú smerové derivácie nezáporné, potom určite existuje hrana, ktorou poklesne hodnota účelovej funkcie $R(b)$, a tak sa dostaneme do nového vrcholu. Avšak takýchto hrán môže byť viac. Je veľa teórií, akým pravidlom sa riadiť pri výbere správnej hrany. My pôjdeme po tej, ktorá je najstrmšia, teda taká, v ktorej bude smerová derivácia najnegatívnejšia. Toto je základná myšlienka modifikovanej simplexovej metódy pre úlohu kvantilovej regresie, ktorá sa využíva i v knižnici `quantreg` v štatistickom programe R. Podrobnejšie o tejto metóde sa dá nájsť v článku [15]. Môžeme si ešte všimnúť, že tento algoritmus sme rozoberali pre fixné $\tau \in (0, 1)$. Obyčajne však chceme odhad parametrov pre jednotlivé kvantily. Preto je praktické použiť parametrické programovanie. Môžeme sa o ňom viac dočítať v [13], [17].

2.2 Metóda vnútorného bodu

Nasledujúca časť je voľne spracovaná podľa [18], [12] a [13].

Simplexová metóda sa veľmi často využíva v praxi. Podarilo sa však zostrojiť taký problém lineárneho programovania, v ktorom algoritmus navštívi všetky vrcholy, kým sa dostane do optima. To znamená, že pre veľké úlohy môže daná metóda konvergovať dlho a numericky nemusí byť presná. Práve z toho dôvodu odštartoval výskum iných alternatívnych algoritmov. Metóda, ktorá sa ujala a začala sa používať sa volá metóda vnútorného bodu. Na rozdiel od simplexovej metódy neprebíha po hranici množiny prípustných riešení, ale vo vnútri tejto množiny.

Štandardnú primárnu úlohu lineárneho programovania definujeme v tvare

$$(P) \quad \min_x \{c^T x \mid Ax = b, x \geq 0\}.$$

Množinu prípustných riešení primárnej úlohy lineárneho programovania (P) označujeme:

$$\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}.$$

Podobne definujeme množinu vnútorný bodov:

$$\mathcal{P}^0 = \{x \in \mathcal{P} | x > 0\}$$

a množina optimálnych riešení má tvar:

$$\mathcal{P}^* = \{x \in \mathcal{P} | c^T x^* \leq c^T x, \forall x \in \mathcal{P}\}.$$

K danej primárnej úlohe (P) existuje duálna úloha v tvare:

$$(D) \quad \max_{y,z} \{b^T y | A^T y + z = c, s \geq 0\}.$$

Množinu prípustných riešení označujeme:

$$\mathcal{D} = \{(y, z) \in \mathbb{R}^m \times \mathbb{R}^n | A^T y + z = c, s \geq 0\},$$

množina vnútorných bodov duálnej úlohy má tvar:

$$\mathcal{D}^0 = \{(y, z) \in \mathcal{D} | s > 0\}$$

a množina optimálnych riešení:

$$\mathcal{D}^* = \{(y^*, z^*) \in \mathcal{D} | b^T y^* \geq b^T y, \quad \forall (y, z) \in \mathcal{D}\}.$$

Metóda vnútorného bodu si vyžaduje splnenie dvoch predpokladov:

- **Predpoklad 1:** Matica $A_{m \times n}$ má plnú hodnotu.
- **Predpoklad 2:** Množiny vnútorných bodov $\mathcal{P}^0$ a $\mathcal{D}^0$ sú neprázdne.

Dôležitým pojmom teórie metódy vnútorného bodu je pojem logaritmická bariérová funkcia, ktorú definujeme nasledovne.

Definícia 2.3. Pod pojmom primárna logaritmická bariérová funkcia rozumieme funkciu

$$B(x, \mu) = c^T x - \mu \sum_{i=1}^n \ln x_i,$$

kde $\mu > 0$.

2.2 Metóda vnútorného bodu

Namiesto pôvodnej primárnej úlohy lineárneho programovania (P) budeme mať novú úlohu, ktorá má tvar:

$$(P_\mu) \quad \min_{x \in \mathbb{R}_{++}^n} \{B(x, \mu) | Ax = b\}.$$

To znamená, že bariérová funkcia je funkciou, ktorú budeme minimalizovať namiesto našej účelovej funkcie $c^T x$. Môžeme si všimnúť, že sme prešli ku nelineárnej účelovej funkcii, tým pádom sa táto nová úloha bude riešiť pomocou nástrojov nelineárneho programovania. Prešli sme k novej triede transformačných bariérových úloh parametrizovaných $\mu > 0$. [12] Tiež si však môžeme všimnúť, že z tvaru účelovej funkcie $B(x, \mu) = c^T x - \mu \sum_{i=1}^n \ln x_i$ sa riešenie tejto novej úlohy pre postupnosť $\mu \rightarrow 0$, blíži k riešeniu našej pôvodnej úlohy. Bariéra $\mu \sum_{i=1}^n \ln x_i$ nám zabezpečuje, že sa budeme nachádzať vo vnútri množiny prípustných riešení. Dôvodom je fakt, že keď sa budeme blížiť k hranici, čiže $x \rightarrow 0$, tak hodnota bariéry bude $\lim_{x \rightarrow 0^+} \mu \sum_{i=1}^n \ln x_i = -\infty$. Tým pádom hodnota novej účelovej funkcie $\lim_{x \rightarrow 0^+} B(x, \mu) = \infty$.

Za bariéru sa vo všeobecnosti môžu vziať aj iné funkcia ako je prirodzený logaritmus. Prirodzený logaritmus berieme z toho dôvodu, lebo sa ukázal byť užitočný v praxi. Dôvodom je i to, že nová primárna úloha (P_μ) je rýdzokonvexná [18] a pri zachovaní primárno-duálnej symetrie ju môžeme aplikovať aj na primárnu i duálnu úlohu. [6]

Definícia 2.4. (*Centrálna trajektória*). Nech $x(\mu)$ je optimálne riešenie pre $\mu > 0$ úlohy (P_μ) . Množinu $\{x(\mu) | \mu > 0\}$ sa nazýva primárna centrálna trajektória.

Nech $(y(\mu), s(\mu))$ je optimálne riešenie (D_μ) . Množinu riešení $\{(y(\mu), s(\mu)) | \mu > 0\}$ nazývame duálna centrálna trajektória.

Nakoniec primárno-duálna centrálna trajektória je množina $\{(x(\mu), y(\mu), z(\mu)) | \mu > 0\}$.

Viac o centrálnej trajektórii sa môžeme dozvedieť v práci [18].

Metódy vnútorného bodu rozdeľujeme na primárne, duálne a primárno-duálne, podľa toho, či iterácia algoritmu prebieha v priestore primárnej, duálnej úloha, alebo v priestore kartézkeho súčinu oboch. Súčasné metódy sa zaoberajú primárno-duálnym riešením úlohy. [8]

Prechod na nové úlohy (P_μ) a (D_μ) nám dáva možnosť použiť Lagrangeovu metódu pre fixný parameter $\mu > 0$. V tomto prípade sú Lagrangeove podmienky nutné i postačujúce podmienky optimality. [8]

2.2.1 Newtonova metóda

Uvažujme úlohu, v ktorej účelová funkcia $f(x)$ je C^2 hladká. Predpokladajme, že bod $x^k \in \mathbb{R}^n$ je približným riešením našej úlohy. Aproximujme účelovú funkciu $f(x)$ Taylorovým polynómom druhého stupňa v okolí bodu x^k :

$$f(x) \approx f_k + \nabla f_k(x - x^k) + \frac{1}{2}(x - x^k)^T \nabla^2 f_k(x - x^k).$$

Účelovú funkciu $f(x)$ sme aproximovali kvadratickou funkciou, ktorú vieme ľahko optimalizovať. Nové približné riešenie $x^{k+1} \in \mathbb{R}^n$ našej úlohy bude práve optimum danej kvadratickej funkcie a má tvar:

$$x^{k+1} = x^k - (\nabla^2 f_k)^{-1} \nabla f_k,$$

kde predpokladáme, že Hessova matica je regulárna.

Nájdenie optimálneho riešenia úloh (P) a (D) je ekvivalentné s hľadaním optima pre úlohu (P_μ) a (D_μ) pre $\mu \rightarrow 0$.

Nech $x^k > 0$ je približné riešenie primárnej úlohy (P_μ) . Gradient a Hessova matica funkcie $B(x, \mu)$ majú tvar:

$$\nabla B = c - X^{-1}1_n$$

$$\nabla^2 B = \mu X^{-2},$$

kde matica $X = \text{diag}(x)$. Nové približné riešenie v Newtonovej metóde označme $x^{k+1} = x^k + p_k$. V každom iteračnom kroku chceme nájsť také $p_k = x^{k+1} - x^k$, ktoré minimalizuje hodnotu účelovej funkcie $B(x, \mu)$, čo nám zaručuje, že sa budeme približovať k optimálnemu riešeniu. Aproximujme účelovú funkciu pre úlohu (P_μ) Taylorovým rozvojom druhého stupňa v okolí bodu $x^k = x$ a dostávame:

$$B(x + p, \mu) \approx B(x, \mu) + p^T \nabla B(x, \mu) + \frac{1}{2} p^T \nabla^2 B(x, \mu) p.$$

Ohraničenie $A(x + p) = b$ sa nám zmení na $Ap = 0$. V každom kroku budeme teda riešiť úlohu:

$$\min_p \{c^T p - \mu p^T X^{-1} 1_n + \frac{1}{2} \mu p^T X^{-2} p \mid Ap = 0\}. \quad (24)$$

2.2 Metóda vnútorného bodu

Riešením tohto problému sa pohneme z daného bodu x v optimálnom smere p .

Na danú úlohu (24) sa môžeme pozeráť ako na úlohu o viazanom extréme a príslušná Lagrangeova funkcia má tvar:

$$L(p, y) = c^T p - \mu p^T X^{-1} 1_n + \frac{1}{2} \mu p^T X^{-2} p - y^T A p$$

Zderivujeme Lagrangeovu funkciu podľa jednotlivých premenných a dostaneme :

$$\begin{aligned} \frac{\partial L}{\partial p} &= c - \mu X^{-1} 1_n + \mu X^{-2} p - A^T y = 0 \\ \frac{\partial L}{\partial y} &= -A p = 0. \end{aligned}$$

Prvú rovnicu vynásobíme maticou AX^2 zľava a druhú vynásobíme konštantou μ :

$$\begin{aligned} AX^2 c - \mu AX 1_n + \mu A p - AX^2 A^T y &= 0 \\ \mu A p &= 0. \end{aligned}$$

Vzájomným odčítaním daných rovníc, čím eliminujeme premennú p dostávame:

$$AX^2 c - \mu AX 1_n = AX^2 A^T y. \quad (25)$$

Dostali sme systém normálnej rovníc, ktoré dostaneme metódou najmenších štvorcov optimalizovaním úlohy:

$$\min_y \|X A^T y - X c - \mu 1_n\|^2.$$

Vyriešením systému týchto rovníc dostaneme príslušné y , pomocou ktorého potom vypočítame optimálny Newtonov smer p . Tento prístup je primárny typ algoritmu vnútorného bodu.

Druhý prístup, ktorú bude pre nás prospešnejší bude primárno-duálny typ algoritmu. Zaoberajme sa teraz riešením duálnej úlohy (D) . Zavedením bariérovej funkcie $B(y, z, \mu) = b^T y - \mu \sum_{i=1}^n \ln z_i$ dostávame nový problém v tvare:

$$(D_\mu) \quad \max\{B(y, z, \mu) \mid A^T y + z = c, z > 0\}.$$

Ako sme už opomenuli skôr, môžeme využiť Lagrangeovu funkciu pre úlohu (D_μ) , ktorá má tvar:

$$L_D(y, z, x) = b^T y + \mu \sum_{i=1}^n \ln z_i - x^T (A^T y + z - c).$$

Parciálnymi deriváciami dostaneme systém rovníc:

$$\begin{aligned} Ax &= b \quad x > 0 \\ A^T y + z &= c, \quad z > 0 \\ Xz &= \mu 1_n. \end{aligned}$$

Môžeme si všimnúť, že $Ax = b$ predstavuje primárnu prípustnosť, $A^T y + z = c$ predstavuje duálnu prípustnosť a $Xz = \mu 1_n$ je μ -komplementarita. [8] Rovnaký systém rovníc by sme dostali tým istým spôsobom pre primárnu úlohu. Prepíšme:

$$F(x(\mu), y(\mu), z(\mu)) = \begin{cases} Ax - b = 0 & x > 0 \\ A^T y + z - c = 0, & z > 0 \\ Xz - \mu 1_n = 0 \end{cases}$$

To znamená, že systém rovníc môžeme zapísať ako $F(x, y, z) = 0_{2n+m}$, kde $x > 0$ a $z > 0$. Na danú funkciu $F(x, y, z) : \mathbb{R}^{2n+m} \rightarrow \mathbb{R}^{2n+m}$ aplikujeme Newtonovu metódu a dostaneme sústavu pre výpočet optimálneho smeru v každej iterácii:

$$\begin{pmatrix} Z & 0 & X \\ A & 0 & 0 \\ 0 & A^T & I \end{pmatrix} \begin{pmatrix} p_x \\ p_y \\ p_z \end{pmatrix} = \begin{pmatrix} \mu 1_n - Xz \\ b - Ax \\ c - A^T y - z \end{pmatrix},$$

kde $Z = \text{diag}(z)$. Danú sústavu môžeme explicitne vyriešiť [19]:

$$\begin{aligned} p_y &= (AZ^{-1}XA^T)^{-1}[AZ^{-1}X(c - \mu X^{-1}1_n - A^T y) + b - Ax] \\ p_x &= XZ^{-1}[A^T p_y + \mu X^{-1}1_n - c + A^T y] \\ p_z &= -A^T p_y + c - A^T y - z. \end{aligned}$$

Pre dané rovnice však nemáme splnené podmienky $x > 0$ a $z > 0$ pre každú iteráciu, a preto sa pristupuje ku regulácii kroku. Viac o výbere kroku v práci [8].

2.2.2 Frisch-Newton algoritmus

Tento algoritmus predstavili v roku 1997 Koenker a Portnoy v článku [16]. Je určený na počítanie odhadov parametrov pre kvantilovú regresiu.

Zaoberáme sa úlohou kvantilovej regresie (10), ktorú vieme previesť na úlohu lineárneho

2.2 Metóda vnútorného bodu

programovania:

$$(P) \min_{(b,u,v) \in \mathbb{R}^p \times \mathbb{R}_+^{2n}} \{\tau 1_n^T u + (1 - \tau) 1_n^T v \mid y - Xb = u - v, b \in \mathbb{R}^p, (u, v) \in \mathbb{R}_+^{2n}\},$$

s príslušnou duálnou úlohou v tvare:

$$(D) \max_{d \in \mathbb{R}^n} \{y^T d \mid X^T d = 0, d \in [\tau - 1, \tau]^n\},$$

ktorú môže pomocou substitúcie $a = d + 1 - \tau$ previesť na úlohu ohraňovanú na jednotkovej n - rozmernej kocke:

$$(D) \max_{a \in \mathbb{R}^n} \{y^T a \mid X^T a = (1 - \tau) X^T 1_n, a \in [0, 1]^n\}.$$

Tento tvar duálnej úlohy zodpovedá tvaru úlohy lineárneho programovania, na ktorý môžeme využiť metódu vnútorného bodu.

Zavedme doplnkovú premennú $s \in \mathbb{R}^n$, ktorá spĺňa:

$$a + s = 1_n.$$

Týmto získame bariérovú funkciu:

$$B(a, s, \mu) = y^T a + \mu \sum_{i=1}^n (\ln a_i + \ln s_i),$$

ktorú maximalizujeme vzhľadom na obmedzenia $X^T a = (1 - \tau) X^T 1_n$ a $a + s = 1_n$. Pokračujeme podobne ako v predchádzajúcej časti pre primárny typ algoritmu, vypočítame gradient a Hessovu maticu bariérovej funkcie:

$$\nabla B = y - \mu(A^{-1} + S^{-1})1_n$$

$$\nabla^2 B = -\mu(A^{-2} + S^{-2}).$$

Newtonom krok δ_a dostaneme riešením úlohy v tvare:

$$\max_{\delta_a} \{y^T \delta_a + \mu \delta_a^T (A^{-1} + S^{-1})1_n - \frac{1}{2} \mu \delta_a^T (A^{-2} + S^{-2}) \delta_a \mid X^T \delta_a = 0\},$$

kde $A = \text{diag}(a)$ a $S = \text{diag}(s)$. Použitím Lagrangeovej funkcie, kde $b \in \mathbb{R}^p$ je Lagrangeov multiplikátor, zderivovaním podľa jednotlivých premenných nakoniec dostaneme systém rovníc:

$$\begin{aligned} y + \mu(A^{-1} + S^{-1})1_n - \mu(A^{-2} + S^{-2})\delta_a - Xb &= 0 \\ X^T \delta_a &= 0. \end{aligned}$$

2.2 Metóda vnútorného bodu

Označme $W = (A^{-2} + S^{-2})^{-1}$, prenásobme prvú rovnicu zľava maticou $X^T W$ a použitím $X^T \delta_a = 0$ môžeme explicitne vyjadriť b nasledovne:

$$b = (X^T W X)^{-1} X^T W (y + \mu(A^{-1} + S^{-1})1_n).$$

V každom kroku vypočítame Lagrangeov multiplikátor a potom Newtonov smer δ_a , čím sa dostaneme k novému bodu, ktoré sa blíži k optimu.

Teraz sa pozrieme na primárno-duálny typ algoritmu pre úlohu kvantilovej regrese.

K duálnej úlohe môžeme zostrojiť Lagrangeovu funkciu, ktorá má tvar:

$$L(a, s, b, u, \mu) = B(a, s, \mu) - b^T (X^T a - (1 - \tau)X^T 1_n) - u^T (a + s - 1_n).$$

Použijeme $v = \mu A^{-1}$ a dostaneme sústavu rovníc:

$$X^T a = (1 - \tau)X^T 1_n$$

$$a + s = 1_n$$

$$Xb + u - v = y$$

$$US1_n = \mu 1_n$$

$$AV1_n = \mu 1_n,$$

kde matice U, S, V sú diagonálne matice príslušných vektorov u, s, v . Na danú sústavu rovníc použijeme Newtonovu metódu a Newtonov smer získame, ako riešenie úlohy:

$$\begin{pmatrix} X^T & 0 & 0 & 0 & 0 \\ I & I & 0 & 0 & 0 \\ 0 & 0 & I & -I & X \\ 0 & U & S & 0 & 0 \\ V & 0 & 0 & A & 0 \end{pmatrix} \begin{pmatrix} \delta_a \\ \delta_s \\ \delta_u \\ \delta_v \\ \delta_b \end{pmatrix} = \begin{pmatrix} (1 - \tau)X^T 1_n - X^T a \\ 1_n - a - s \\ y - Xb - u + v \\ \mu 1_n - US1_n \\ \mu 1_n - AV1_n \end{pmatrix}.$$

Nakoniec môžeme danú úlohu upraviť na tvar:

$$\delta_b = (X^T W X)^{-1} ((1 - \tau)X^T 1_n - X^T a - X^T W \varphi(\mu))$$

$$\delta_a = W(X\delta_b + \varphi(\mu))$$

$$\delta_s = -\delta_a$$

$$\delta_u = \mu S^{-1}1_n - U1_n + S^{-1}U\delta_a$$

$$\delta_v = \mu A^{-1}1_n - V1_n + A^{-1}V\delta_s,$$

2.2 Metóda vnútorného bodu

kde funkcia $\varphi(\mu) = y - Xb + \mu(A^{-1} - S^{-1})1_n$ a matica $W = (S^{-1}U + A^{-1}V)^{-1}$.

Modifikovaný simplexov algoritmus a Fisch-Newton algoritmus na výpočet odhadu parametrov kvantilovej regresie nájdeme v balíčku `quantreg` v štatistickom programe R. Na začiatku sme spomenuli, že metóda vnútorného bodu je efektívnejšia pre rozsiahlejšie úlohy. Samozrejme, závisí to od počtu parametrov a i veľkosti dát. Napríklad, pre štyri parametre v dátovom súbore, začne byť Frisch-Newton algoritmus efektívnejší, keď máme okolo 5000 dát, pre 16 parametrov je efektívnejší už okolo 2000 dát. [13]

3 MM algoritmus

Aplikovaná štatistika sa zaoberá množstvom problémov, ktoré zahŕňajú optimalizáciu funkcií, ako sú napríklad funkcia vierohodnosti alebo funkcia sumy štvorcov rezíduí. Na výpočet odhadov parametrov sa často používa metóda maximálnej vierohodnosti. V špeciálnych prípadoch, napríklad keď chýbajú niektoré dáta, sa namiesto metódy maximálnej vierohodnosti používa EM algoritmus (Expectation–maximization algorithm), ktorý v roku 1977 zaviedli Dempster, Laird a Rubin. Veľa štatistikov používa daný algoritmus ako čiernu skrinku, do ktorej vložia údaje a dostanú výsledok. Avšak podstata tohto algoritmu nie je náročná na pochopenie. EM algoritmus prebieha v dvoch krokoch, ktoré sa iteratívne opakujú. V prvom kroku, v E-kroku, zostrojíme funkciu a to podmienenú strednú hodnotu z logaritmu funkcie vierohodnosti. V druhom kroku, v M-kroku, maximalizujeme túto funkciu z E-kroku.

Menej známy je však fakt, že EM algoritmus je špeciálny prípad všeobecnejšieho algoritmu, ktorý sa nazýva MM algoritmus. MM algoritmus tiež prebieha v dvoch krokoch. V prvom kroku vytvoríme náhradnú funkciu, ktorú v nasledujúcom kroku optimalizujeme. V prípade minimalizácie, skratka MM znamená z angličtiny majorize-minimize. Ak maximalizujeme, MM značí minorize-maximize. Hlavným pozitívom tohto algoritmu je, že namiesto zložitého optimalizačného problému nám ponúka sled jednoduchších optimalizačných problémov, ktoré vo viacerých prípadoch konvergujú k riešeniu nášho počiatočného problému.

Hlavná myšlienka MM algoritmu je nasledujúca. Predstavme si, že chceme minimalizovať funkciu reálnej premennej $L(\beta) : \mathbb{R}^p \rightarrow \mathbb{R}$. Nech $\beta^k \in \mathbb{R}^p$ označuje k -ty iteračný krok. Ako sme si už spomenuli a z názvu MM algoritmu vyplýva, že bude prebiehať v dvoch krokoch. Najprv si vytvoríme náhradnú funkciu $Q(\beta|\beta^k) : \mathbb{R}^p \times \mathbb{R}^p \rightarrow \mathbb{R}$, ktorá bude majorizovať funkciu $L(\beta)$ v bode $\beta^k \in \mathbb{R}^p$. Táto náhradná funkcia musí spĺňať nasledujúce vlastnosti:

$$Q(\beta^k|\beta^k) = L(\beta^k), \quad (26)$$

$$Q(\beta|\beta^k) \geq L(\beta) \quad \forall \beta \in \mathbb{R}^p. \quad (27)$$

Inými slovami, rovina $\beta \mapsto Q(\beta|\beta^k)$ sa nachádza nad rovinou $L(\beta)$ a majú spoločný dotykový bod $\beta = \beta^k$.

V druhom kroku MM algoritmu minimalizujeme náhradnú funkciu $Q(\beta|\beta^k)$. Bod minima je $\beta^{k+1} \in \mathbb{R}^p$. Pre ďalšiu iteráciu $\beta^{k+1} \in \mathbb{R}^p$ z vlastností (26) a (27) platí:

$$L(\beta^{k+1}) \leq Q(\beta^{k+1}|\beta^k) \leq Q(\beta^k|\beta^k) = L(\beta^k). \quad (28)$$

Z daných vlastností vidíme nasledujúcu nerovnosť $L(\beta^{k+1}) \leq L(\beta^k)$, čo znamená, že nová iterácia sa približuje k minimu danej funkcie $L(\beta)$. Rovnosť nastáva iba vtedy, keď nastane rovnosť v (28), teda $Q(\beta^{k+1}|\beta^k) = Q(\beta^k|\beta^k)$.

Pre druhý prípad MM algoritmu (minorize-maximize) hovoríme, že funkcia $Q(\beta|\beta^k)$ minorizuje účelovú funkciu $L(\beta)$ v bode $\beta^k \in \mathbb{R}^p$, ak funkcia $-Q(\beta|\beta^k)$ majorizuje funkciu $-L(\beta)$ v bode $\beta^k \in \mathbb{R}^p$. V druhom kroku minorize-maximize MM algoritmu maximalizujeme náhradnú funkciu z prvého kroku.

Je zrejmé, že MM algoritmus sa dá dobre využiť, keď potrebujeme optimalizovať funkciu, na ktorú nemôžeme použiť štandardné optimalizačné nástroje. Vidíme, že myšlienka MM algoritmu nie je vôbec komplikovaná. Avšak problém môže prísť s faktom, že treba nájsť náhradnú funkciu $Q(\beta|\beta^k)$, ktorá spĺňa dané podmienky (26) a (27). Od tejto funkcie očakávame hlavne to, aby bola ľahko optimalizovateľná. Vo všeobecnosti môžeme majorizujúce alebo minorizujúce funkcie získať z rôznych nerovností. Používajú sa napríklad:

- Jensenova nerovnosť
- Cauchy–Schwarzova nerovnosť
- Nerovnosti aritmetického a geometrického priemeru.

Pre prípad viacrozmernej účelovej funkcie postupujeme tak, že vytvoríme náhradnú funkciu v každej premennej a tú optimalizujeme.

Ako sme už načrtli skôr, EM algoritmus je špeciálny prípad MM algoritmu. Ak za účelovú funkciu $L(\beta)$ zvolíme logaritmus funkcie vierohodnosti napozorovaných dát a funkcia $Q(\beta|\beta^k)$ je funkcia vytvorená v E-kroku, tak funkcia, ktorá bude minorizovať funkciu $L(\beta)$ je podľa [20] :

$$Q^*(\beta|\beta^k) = Q(\beta|\beta^k) + L(\beta^k) - Q(\beta^k|\beta^k). \quad (29)$$

Maximalizáciou danej minorizujúcej funkcie (29) prejdeme k maximalizácii funkcie $Q(\beta|\beta^k)$, čo je práve myšlienka EM algoritmu.

Výhoda MM algoritmu je v numerickej stabilite. Dá sa využiť aj pre prípad viazaného extrému. Nevýhoda MM algoritmu je v dlhej konvergencii.

V nasledujúcej podkapitole sa vrátíme ku problému kvantilovej regresii a využitiu MM algoritmu na jeho riešenie.

3.1 MM algoritmus a kvantilová regresia

Nasledujúca podkapitola je spracovaná podľa článku [11].

V prvej kapitole sme zaviedli optimalizačnú úlohu kvantilovej regresie (10) a previedli sme ju na úlohu lineárneho programovania. Daný problém sme riešili pomocou simplexovej metódy a metódy vnútorného bodu. Obe tieto metódy sú implementované v systéme R. Zavedením nových premenných nám však vznikol rozsiahlejší problém. MM algoritmus má za úlohu majorizovanie pôvodnej účelovej funkcie kvadratickou funkciou a následné minimalizovanie danej kvadratickej funkcie.

Zaoberajme sa zovšeobecneným problémom kvantilovej regresie a označme:

$$L(\beta) = \sum_{i=1}^n \rho_{\tau}(y_i - f_i(\beta)). \quad (30)$$

Pre úlohu lineárnej kvantilovej regresie platí $f_i(\beta) = x_i^T \beta$.

Predpokladajme, že každá funkcia $f_i(\beta) : \mathbb{R}^p \rightarrow \mathbb{R}$ je spojitá diferencovateľná. Zavedieme si ešte jeden technický predpoklad:

$$\lim_{\|\beta\| \rightarrow \infty} \sum_{i=1}^n f_i(\beta)^2 = \infty. \quad (31)$$

Pre jednoduchosť si ešte zadefinujeme i -te rezíduum $r_i = r_i(\beta) = y_i - f_i(\beta)$.

Minimalizovanie po častiach lineárnej účelovej funkcie $L(\beta)$ klasickými optimalizačnými nástrojmi je komplikované, pretože môže dosahovať viacnásobné minimá. Problémom je i fakt, že funkcia $\rho_{\tau}(r)$ nie je diferencovateľná, ak $r = 0$. Tento nedostatok však odstránime tým, že účelovú funkciu aproximujeme vhodnejšou funkciou, na ktorú potom použijeme MM algoritmus. Definujeme si novú funkciu pre $\epsilon > 0$:

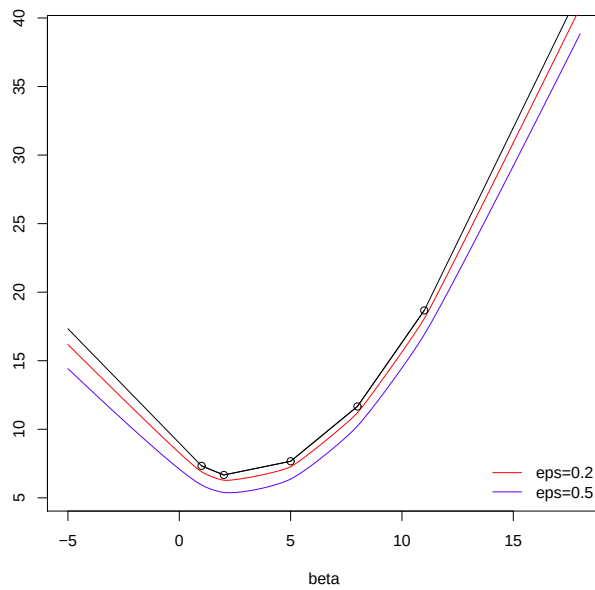
$$\rho_{\tau}^{\epsilon}(r_i) = \rho_{\tau}(r_i) - \frac{\epsilon}{2} \ln(\epsilon + |r_i|) \quad i = 1, \dots, n. \quad (32)$$

Nová funkcia, ktorá aproximuje účelovú funkciu $L(\beta)$ má tvar:

$$L_{\epsilon}(\beta) = \sum_{i=1}^n \rho_{\tau}^{\epsilon}(r_i). \quad (33)$$

Voľba konštanty $\epsilon > 0$ ovplyvňuje, ako ďaleko je aproximovaná funkcia (33) od našej pôvodnej účelovej funkcie (30). Z predpisu náhradnej funkcie (33) je zjavné, že čím je $\epsilon > 0$ menšie, tým sa daná aproximujúca funkcia blíži k funkcii (30).

Nasledujúci obrázok (3) nám znázorňuje jednorozmerný príklad kvantilovej regresie. Uvažujeme, že $n = 5$ a $p = 1$. Nech kvantil $\tau = \frac{1}{3}$ a vektor závislých premenných je $y = (1, 2, 5, 8, 11)$. Na obrázku môžeme vidieť po častiach lineárnu účelovú funkciu z obrázku 2 a aproximujúce funkcie (33) pre dve rôzne voľby parametra $\epsilon > 0$.



Obr. 3: Horná krivka predstavuje účelovú funkciu pre $\tau = \frac{1}{3}$. Dolné farebné krivky sú aproximácie účelovej funkcie pre rôznu voľbu parametra $\epsilon > 0$.

Z predchádzajúcej podkapitoly vieme, že MM algoritmus prebieha v dvoch krokoch. V prvom kroku volíme náhradnú funkciu, ktorá nám má majorizovať príslušnú funkciu (33). Nájsť však vhodnú náhradnú funkciu nie je triviálne. Autori článku [11] navrhli nasledovné riešenie.

Veta 3.1. *Funkcia*

$$\zeta_\tau^\epsilon(r_i|r_i^k) = \frac{1}{4} \left[\frac{r_i^2}{\epsilon + |r_i^k|} + (4\tau - 2)r_i + c_i \right] \quad (34)$$

majorizuje funkciu $\rho_\tau^\epsilon(r_i)$ z (32) v bode $r_i = r_i^k = r_i(\beta^k)$, kde $c_i \in \mathbb{R}$ je konštanta, ktorá spĺňa rovnosť $\zeta_\tau^\epsilon(r_i^k|r_i^k) = \rho_\tau^\epsilon(r_i^k)$.

Dôkaz. Z definície majorizujúcej funkcie vieme, že musí spĺňať podmienky (26) a (27).

Podmienka (26) vyplýva priamo z definície. Teraz overíme vlastnosť (27).

Potrebuje dokázať, že rozdiel $\zeta_\tau^\epsilon(r_i|r_i^k) - \rho_\tau^\epsilon(r_i) \geq 0$.

$$\zeta_\tau^\epsilon(r_i|r_i^k) - \rho_\tau^\epsilon(r_i) = \frac{1}{4} \left[\frac{r_i^2}{\epsilon + |r_i^k|} + (4\tau - 2)r_i + c_i \right] - \rho_\tau(r_i) - \frac{\epsilon}{2} \ln(\epsilon + |r_i|).$$

Jediný členy, kde sa nám nevyskytuje r_i v tvare párnej mocniny alebo normy je člen $\frac{1}{4}(4\tau - 2)r_i - \rho_\tau(r_i)$.

Počítajme:

$$\frac{1}{4}(4\tau - 2)r_i - \rho_\tau(r_i) = \frac{1}{2}(2\tau - 1)r_i - \begin{cases} r_i\tau & r_i \geq 0 \\ r_i(\tau - 1) & r_i < 0 \end{cases} = \begin{cases} -\frac{1}{2}r_i & r_i \geq 0 \\ \frac{1}{2}r_i & r_i < 0 \end{cases} = -\frac{1}{2}|r_i|$$

Z výpočtu nám vyplýva, že rozdiel $\zeta_\tau^\epsilon(r_i|r_i^k) - \rho_\tau^\epsilon(r_i)$ je symetrická funkcia podľa 0.

Tento poznatok nám dovoľuje obmedziť naše vyšetrovanie len na interval, kde $r_i \geq 0$.

Zderivujeme daný rozdiel na príslušnom intervale:

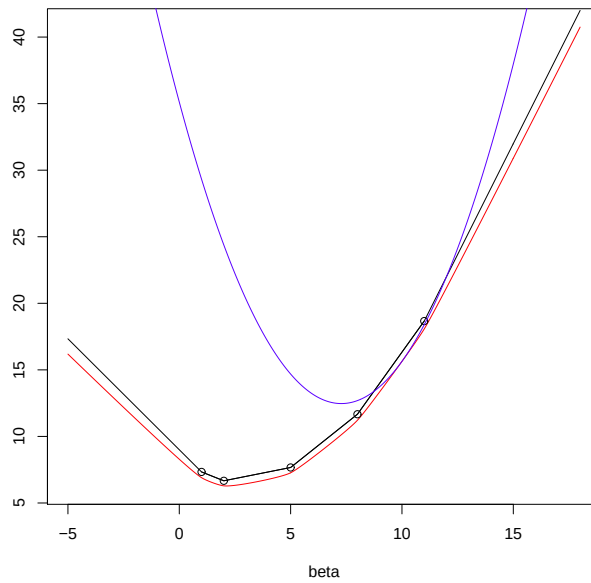
$$\begin{aligned} \frac{d}{dr_i} [\zeta_\tau^\epsilon(r_i|r_i^k) - \rho_\tau^\epsilon(r_i)] &= \frac{1}{4} \left[\frac{2r_i}{\epsilon + |r_i^k|} + (4\tau - 2) \right] - \tau - \frac{\epsilon}{2\epsilon + r_i} \\ &= \frac{r_i}{2(\epsilon + |r_i^k|)} - \frac{1}{2} + \frac{\epsilon}{2} \frac{1}{\epsilon + r_i} \\ &= \frac{r_i}{2(\epsilon + |r_i^k|)} - \frac{r_i}{2(\epsilon + r_i)} \\ &= \frac{r_i}{2} \frac{r_i - |r_i^k|}{(\epsilon + |r_i^k|)(\epsilon + r_i)} \end{aligned}$$

Z derivácie funkcie $\zeta_\tau^\epsilon(r_i|r_i^k) - \rho_\tau^\epsilon(r_i)$ vidieť, že na intervale $(0, r_i^k)$ je klesajúca, v bode r_i^k dosahuje lokálne minimum a na intervale (r_i^k, ∞) je rastúca. Z vlastnosti (26) platí $\zeta_\tau^\epsilon(r_i^k|r_i^k) - \rho_\tau^\epsilon(r_i^k) = 0$ a týmto je vlastnosť (27) dokázaná. $\square$

Náhradná funkcia k funkcii (33) má tvar:

$$Q_\epsilon(\beta|\beta^k) = \sum_{i=1}^n \zeta_\tau^\epsilon(r_i|r_i^k). \quad (35)$$

Vráťme sa znova k nášmu príkladu, ktorý je znázornený na obrázku (2). Nasledujúci obrázok (4) zobrazuje účelovú funkciu s príslušnou aproximačnou funkciou pre $\epsilon = 0.2$ a náhradnú majorizujúcu funkciu v bode $\beta^0 = 10$.



Obr. 4: Čierna krivka predstavuje účelovú funkciu $L(\beta)$ pre $\tau = \frac{1}{3}$. Červená krivka je jej aproximácia $L_\epsilon(\beta)$ pre voľbu parametra $\epsilon = 0.2$ a modrá krivka predstavuje majorizujúcu funkciu k $L_\epsilon(\beta)$ v bode $\beta^0 = 10$.

Novú optimalizačnú úlohu môžeme zapísať:

$$\min_{\beta \in \mathbb{R}^p} Q_\epsilon(\beta|\beta^k). \quad (36)$$

Môžeme pristúpiť k druhému kroku MM algoritmu. Na minimalizovanie funkcie (35) použijeme Newtonovu metódu.

3.2 Zovšeobecnená Newtonova metóda

Uvažujme úlohu (36). Predpokladajme, že $\beta^k \in \mathbb{R}^p$ je približným riešením tejto úlohy. V okolí tohto bodu β^k aproximujeme účelovú funkciu (35) Taylorovým polynómom

druhého rádu.

$$Q_\epsilon(\beta + \beta^k | \beta^k) \approx Q_\epsilon(\beta^k | \beta^k) + \nabla Q_\epsilon(\beta^k | \beta^k)^T (\beta - \beta^k) + \frac{1}{2} (\beta - \beta^k)^T \nabla^2 Q_\epsilon(\beta^k | \beta^k) (\beta - \beta^k) = F(\beta)$$

Nové približné riešenie $\beta^{k+1} \in \mathbb{R}^p$ v Newtonovej metóde pre danú úlohu (36) definujeme ako minimum kvadratickej funkcie $F(\beta)$. Deriváciu funkcie $F(\beta)$

$$\nabla F(\beta) = \nabla Q_\epsilon(\beta^k | \beta^k) + \nabla^2 Q_\epsilon(\beta^k | \beta^k) (\beta - \beta^k) = 0_{p \times 1}$$

upravíme na tvar

$$\nabla^2 Q_\epsilon(\beta^k | \beta^k) (\beta - \beta^k) = -\nabla Q_\epsilon(\beta^k | \beta^k).$$

Nový iteračný bod, ak predpokladáme že Hessova matica je kladne definitná, je

$$\beta^{k+1} = \beta^k - [\nabla^2 Q_\epsilon(\beta^k | \beta^k)]^{-1} \nabla Q_\epsilon(\beta^k | \beta^k).$$

Smer Newtonovej metódy je teda $\Delta_\epsilon^k = -[\nabla^2 Q_\epsilon(\beta^k | \beta^k)]^{-1} \nabla Q_\epsilon(\beta^k | \beta^k)$ a krok $\lambda_k = 1$.

Dopočítajme príslušný gradient a Hessovu maticu pre konkrétnu účelovú funkciu (35).

Gradient má tvar:

$$\nabla Q_\epsilon(\beta | \beta^k) = \frac{1}{2} \sum_{i=1}^n \left[\frac{r_i}{\epsilon + |r_i^k|} + 2\tau - 1 \right] \nabla r_i(\beta). \quad (37)$$

Aproximujeme $dr_i(\beta) = -df_i(\beta)$. Táto aproximácia je presná v prípade lineárnej kvantilovej regresie. Potom môžeme prepísať Hessovu maticu z (37)

$$\nabla Q_\epsilon(\beta | \beta^k) = \frac{1}{2} v_\epsilon(\beta) \nabla f(\beta), \quad (38)$$

kde vektor

$$v_\epsilon(\beta) = \left(1 - 2\tau - \frac{r_1(\beta)}{\epsilon + |r_1(\beta^k)|}, \dots, 1 - 2\tau - \frac{r_n(\beta)}{\epsilon + |r_n(\beta^k)|} \right)^T.$$

Hessova matica má tvar:

$$\nabla^2 Q_\epsilon(\beta | \beta^k) = \frac{1}{2} \sum_{i=1}^n \left[\frac{1}{\epsilon + |r_i^k|} \nabla r_i(\beta)^T \nabla r_i(\beta) + \left(\frac{r_i}{\epsilon + |r_i^k|} + 2\tau - 1 \right) \nabla^2 r_i(\beta) \right] \quad (39)$$

$$\approx \frac{1}{2} \sum_{i=1}^n \left[\frac{1}{\epsilon + |r_i^k|} \nabla r_i(\beta)^T \nabla r_i(\beta) \right]. \quad (40)$$

Využijeme aproximáciu použitú pre gradient $dr_i(\beta) = -df_i(\beta)$ a Hessovu maticu môžeme maticovo prepísať na:

$$\nabla^2 Q_\epsilon(\beta|\beta^k) \approx \frac{1}{2} \nabla f(\beta)^T W_\epsilon(\beta^k) \nabla f(\beta), \quad (41)$$

kde $\nabla f(\beta)$ je $n \times p$ matica prvých derivácií a $W_\epsilon(\beta^k)$ je diagonálna matica veľkosti $n \times n$, ktorej i -ty diagonálny prvok je $[\epsilon + |r_i^k(\beta)|]^{-1}$.

Newtonov spádový smer má teda tvar:

$$\Delta_\epsilon^k = -[\nabla f(\beta^k)^T W_\epsilon(\beta^k) \nabla f(\beta^k)]^{-1} \nabla f(\beta^k)^T v_\epsilon(\beta^k)^T. \quad (42)$$

V prípade lineárnej funkcie $f_i(\beta) = x_i^T \beta$ platí, že $\nabla f(\beta) = X$, a preto $\beta^{k+1} = \beta^k + \Delta_\epsilon^k$ je dobrý nový iteračný bod, ktorý nám zaručuje, že dosiahneme minimum. V prípade nelineárnej funkcie kvantilovej regresie nie je jasné, že $\beta^k + \Delta_\epsilon^k$ bude znižovať hodnotu náhradnej funkcie $Q_\epsilon(\beta|\beta^k)$. V dôsledku toho je spochybnený krok $\lambda_k = 1$ a musíme prejsť k modifikovanému Newtonovmu algoritmu a krok skrátiť. Nový iteračný bod vyzerá nasledovne:

$$\beta^{k+1} = \beta^k + \lambda^k \Delta_\epsilon^k, \quad (43)$$

kde krok volíme

$$\lambda^k = \max\{2^{-\nu} : Q_\epsilon(\beta^k + 2^{-\nu} \Delta_\epsilon^k|\beta^k) < Q_\epsilon(\beta^k|\beta^k), \nu \in \mathbb{N}\}. \quad (44)$$

Ak $\Delta_\epsilon^k = 0$, sme v stacionárnom bode. Vtedy položíme $\lambda^k = 1$.

Algoritmus modifikovanej Newtonovej metódy:

Vstup:

Zvolíme štartovací bod $\beta^0 \in \mathbb{R}^p$, tolerančnú konštantu $\epsilon > 0$ a nastavíme počítadlo iterácií $k = 0$.

Výstup:

1. Výpočet novej iterácie β^{k+1} pomocou (43), kde príslušný krok λ^k a smer Δ_ϵ^k volíme podľa rovníc (44) a (42).
2. Nastavíme počítadlo na $k = k + 1$ a ak nemáme splnené konvergentné kritérium, tak sa vrátime na 1. krok.

Ako voliť β^0 , ϵ a konvergentné kritérium rozoberieme v nasledujúcej podkapitole. Teraz nás ale zaujíma, či algoritmus (43) zaručuje minimalizovanie aproximujúcej účelovej funkcie $L_\epsilon(\beta)$ a ako blízko je získané minimum od počiatočnej účelovej funkcie $L(\beta)$. Na prvú otázku sa dá ľahko odpovedať. Tento algoritmus nedokáže nájsť minimum pre všetky problémy. Avšak pre úlohu lineárnej kvantilovej regresie za určitého predpokladu nám bude Newtonova metóda konvergovať.

Tvrdenie 3.2. *Nech matica plánu X má plnú hodnotu. Potom pre úlohu lineárnej kvantilovej regresie platí, že algoritmus (43) konverguje k jednoznačnému riešeniu úlohy (36).*

Dôkaz môže čitateľ nájsť v dodatku článku [11].

Nelineárna kvantilová regresia nám však môže priniesť problémy. Hlavným problémom je fakt, že účelová funkcia $L_\epsilon(\beta)$ nemusí byť konvexná. V dôsledku toho nám algoritmus (43) môže nájsť iba lokálne minimum. Z praktického hľadiska sa tento problém rieši sekvenciou viacerých úloh s rôznymi štartovacími bodmi β^0 . Z výsledkov týchto úloh sa vyberie tá hodnota, v ktorej nadobúda funkcia $L_\epsilon(\beta)$ najnižšiu funkčnú hodnotu. Tento bod potom prehlásime za naše globálne minimum.

V ďalšom tvrdení sa obmedzíme na výber voľby parametra z kompaktnej podmnožiny Ω množiny parametrov modelu.

Tvrdenie 3.3. *Nech β^0 je ľubovoľný bod z $\mathbb{R}^p$. Kompaktná množina*

$$\Omega = \{\beta \in \mathbb{R}^p : L_1(\beta) \leq L(\beta^0) + n\}$$

obsahuje všetky β , ktoré spĺňajú nerovnosť $L(\beta) \leq L(\beta^0)$ alebo tiež $L_\epsilon(\beta) \leq L_\epsilon(\beta^0)$ pre všetky $\epsilon \in (0, 1)$. Funkcia $L_1(\beta)$ je aproximujúca účelová funkcia pre $\epsilon = 1$.

Množina Ω obsahuje všetky iterácie z MM algoritmu, ktorý za počiatočný krok volí práve β^0 . Predovšetkým však množina Ω obsahuje aj všetky minimá funkcií $L(\beta)$ a $L_\epsilon(\beta)$ pre voľbu $\epsilon \in (0, 1)$.

Dôkaz. Množina Ω musí byť kompaktná, pretože predpoklad (31) nám zaručuje, že $\{\beta : L_\epsilon(\beta) \leq c\}$ je kompaktná podmnožina $\mathbb{R}^p$.

Z (33) pre každé $\epsilon \in (0, 1)$ platí:

$$L_1(\beta) \leq L_\epsilon(\beta) \leq L(\beta) + n.$$

Potom pre každé β spĺňajúce $L_\epsilon(\beta) \leq L_\epsilon(\beta^0)$ platí:

$$L_1(\beta) \leq L_\epsilon(\beta) \leq L_\epsilon(\beta^0) \leq L(\beta^0) + n.$$

To znamená, že skutočne $\beta \in \Omega$.

□

Dané tvrdenie je pre nás prospešné skôr z teoretického hľadiska. Môžeme si však všimnúť, že sme sa obmedzili pri výbere ϵ len na interval $(0, 1)$.

Druhá otázka, ktorú by bolo prospešné zodpovedať sa týka toho, ako blízko je získané minimum od optima počiatočnej účelovej funkcie $L(\beta)$. Odpoveď na túto otázku nám dá nasledujúce tvrdenie.

Tvrdenie 3.4. *Nech $d = 1 + \sup\{|r_i(\beta)| : \beta \in \Omega, 1 \leq i \leq n\}$. Ak $\epsilon d < 1$, potom platí*

$$|L(\beta) - L_\epsilon(\beta)| \leq -\frac{\epsilon n}{2} \ln \epsilon, \quad \forall \beta \in \Omega. \quad (45)$$

Ak postupne $\hat{\beta}$ a $\hat{\beta}_\epsilon$ minimalizujú $L(\beta)$ a $L_\epsilon(\beta)$, potom

$$L(\hat{\beta}_\epsilon) - L(\hat{\beta}) \leq -\epsilon n \ln \epsilon. \quad (46)$$

Dôkaz. Nakoľko množina Ω je kompaktná a funkcia $f_i(\beta)$ je spojitá je d konečné číslo. Ďalej platí, ak $\epsilon + \|r\| \leq 1$ a vieme že $\epsilon \in (0, 1)$, potom $|\ln(\epsilon + \|r\|)| \leq -\ln \epsilon$. Ak $\epsilon + \|r\| > 1$, potom z predpokladu $\epsilon < d^{-1}$ vyplýva $|\ln(\epsilon + \|r\|)| < \ln d < -\ln \epsilon$. Potom

$$|L(\beta) - L_\epsilon(\beta)| = \frac{\epsilon}{2} \left| \sum_{i=1}^n \ln(\epsilon + |r_i|) \right| \leq \frac{\epsilon}{2} \sum_{i=1}^n |\ln(\epsilon + |r_i|)| \leq -\frac{\epsilon n}{2} \ln \epsilon.$$

A druhá nerovnosť (46) vyplýva z:

$$L(\hat{\beta}_\epsilon) - L(\hat{\beta}) \leq L(\hat{\beta}_\epsilon) - L_\epsilon(\hat{\beta}_\epsilon) + L_\epsilon(\hat{\beta}) - L(\hat{\beta}) \leq |L(\hat{\beta}_\epsilon) - L_\epsilon(\hat{\beta}_\epsilon)| + |L_\epsilon(\hat{\beta}) - L(\hat{\beta})| \leq -\epsilon n \ln \epsilon.$$

□

Stále je nám ale neznáme, ako získame optimálne $\hat{\beta} \in \mathbb{R}^p$. Odpoveď nájdeme v nasledujúcom tvrdení.

Tvrdenie 3.5. *Nech $\hat{\beta}_\epsilon$ minimalizuje funkciu $L_\epsilon(\beta)$. Ak má funkcia $L(\beta)$ minimum $\hat{\beta}$, potom platí*

$$\lim_{\epsilon \rightarrow 0} \hat{\beta}_\epsilon = \hat{\beta}.$$

Dôkaz. Nech ϵ_k je postupnosť klesajúca k 0 s $\lim_{k \rightarrow \infty} \hat{\beta}_{\epsilon_k} = \beta^*$. Z (32) a (33) vyplýva, že $\lim_{k \rightarrow \infty} L_{\epsilon_k}(\beta) = L(\beta)$ pre všetky β . Tak limita z nerovnosti $L_{\epsilon_k}(\beta_{\epsilon_k}) \leq L_{\epsilon_k}(\beta)$ vedie pre $\beta \in \mathbb{R}^p$ k nerovnosti $L(\beta^*) \leq L(\beta)$. To znamená, že ak má funkcia $L(\beta)$ minimum, tak ho dosahuje práve v bode β^* . $\square$

3.2.1 Voľba parametrov a konvergenčné kritérium

V tejto podkapitole objasníme výber konvergenčného kritéria, ktorý budeme používať pre modifikovaný Newtonov algoritmus, voľbu $\epsilon \in (0, 1)$ a počiatočnú štartovaciu hodnotu β^0 .

Za konvergenčné kritérium pre MM algoritmus vyberieme:

$$Q_\epsilon(\beta^k | \beta^k) - Q_\epsilon(\beta^{k+1} | \beta^k) < v. \quad (47)$$

To znamená, že rozdiel funkčných hodnôt v dvoch po sebe idúcich iteračných bodoch musí byť menší ako konštanta $v \in \mathbb{R}$. Výber konštanty v a ϵ je navzájom podmienený, a to tak, že $\epsilon n |\ln \epsilon| = v$.

Výber štartovacej hodnoty β^0 určíme jednoduchým spôsobom. MM algoritmus začneme blízko hodnoty, ktorú dostaneme ako odhad parametra nášho modelu pomocou metódy najmenších štvorcov.

3.2.2 Zhrnutie MM algoritmu

MM algoritmus sme naprogramovali v systéme R. Pri programovaní, sme si všimli nezrovnalosti v článku [11]. Na začiatku si autori zadefinovali vektor rezíduí r . Neskôr, pri definovaní aproximačnej a majorizujúcej funkcie s týmto vektorom pracovali ako s jednorozmernou premennou. Tým sme dostali problém pre výpočet c z vety (34). Z článku vyplýva, že je to konštanta. Aby sa majorizujúca funkcia dotýkala aproximujúcej funkcie postupne v daných bodoch β^k pre jednotlivé k , c musí byť zaručene vektor rôznych konštánt. V tejto práci je veta (34) v opravenej verzii, pomocou ktorej je zrejmý i výpočet zložiek vektora konštánt $c \in \mathbb{R}^n$.

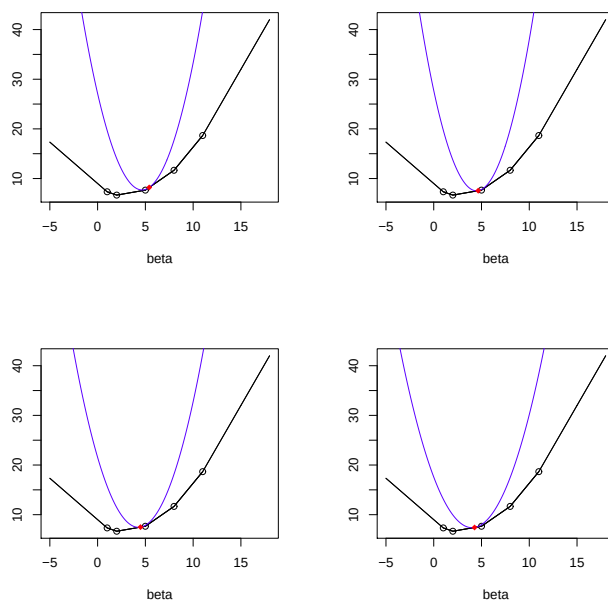
Najprv sme naprogramovali MM algoritmus pre odhad jedného parametra. V niektorých prípadoch nedokonvergoval MM algoritmus k optimu riešenie, ale zastal v jeho blízkosti.

Pre viacrozmerný prípad sme mali problém s odhadom interceptu. Použili sme dáta `engel`, ktoré sú dostupné v knižnici `quantreg` v systéme R. Odhad parametrov (bez interceptu) bol porovnateľný pri danej presnosti s výsledkami modifikovanej simplexovej metódy pre úlohu kvantilovej regresie. Intercept sa však líšil a to hlavne na najspodnejších ($\tau = 0.1$) a najvrchnejších ($\tau = 0.9$) kvantilochoch. Autori článku [11] sa o žiadnych podobných problémoch nezmieňujú.

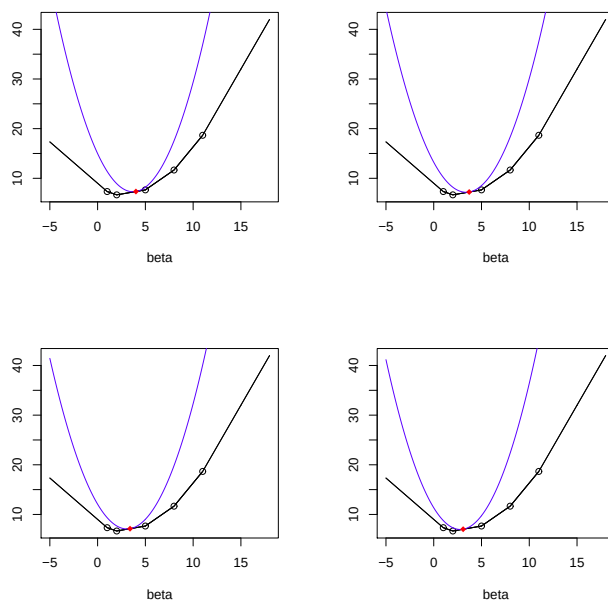
MM algoritmus, ktorý sme naprogramovali, bol výrazne pomalší ako alternatívne algoritmy, t.j. modifikovaná simplexová metóda a Frisch-Newton algoritmus. Je však prospešné zoznámiť sa s MM algoritmom, pretože má veľkú výhodu v jednoduchosti a pre zložité optimalizačné úlohy je to šikovný nástroj na ich riešenie.

Naposledy sa vráťme k nášmu príkladu, ktorý bol znázornený na obrázku (2). Nasledujúca sekvencia obrázkov znázorňuje iterácie MM algoritmu. Za počiatočnú hodnotu sme zvolili odhad parametra β pomocou metódy najmenších štvorcov $\beta^0 = 5.4$. Za konštantu *upsilon* sme zobrali číslo 0.01 a k tomu sme vypočítali príslušné $\epsilon = 0.00024$. Pre toto ϵ je aproximačná funkcia $L_\epsilon(\beta)$ takmer totožná s účelovou funkciou $L(\beta)$, preto nie sú viditeľné na grafe obe, lebo pri danej mierke splývajú.

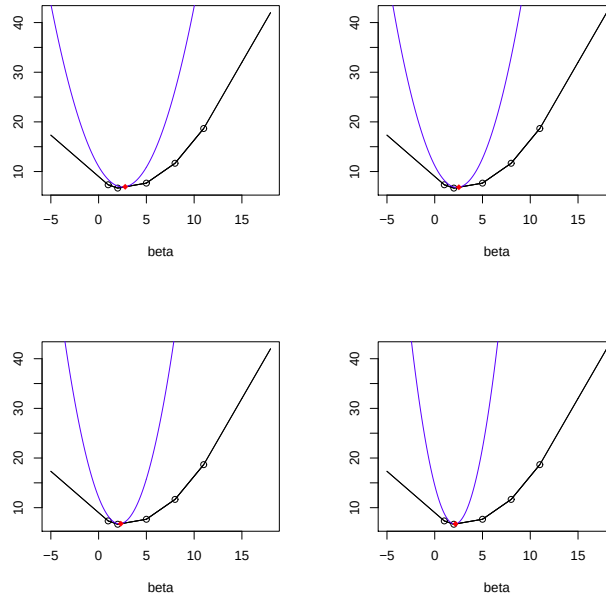
Spolu sme dostali 13 iterácií, pokým sme sa dostali k minimu $L_\epsilon(\beta)$. Nasledujúce obrázky nám postupne zobrazujú dané iterácie. Jednotlivé grafy nám znázorňujú účelovú funkciu $L(\beta) \doteq L_\epsilon(\beta)$ čiernou farbou a modrou farbou náhradnú majorizačnú funkciu k danej funkcii $L_\epsilon(\beta)$, ktorú majorizuje v červenom bode β^k , ktorý sme dostali ako minimum z predchádzajúcej iterácie.



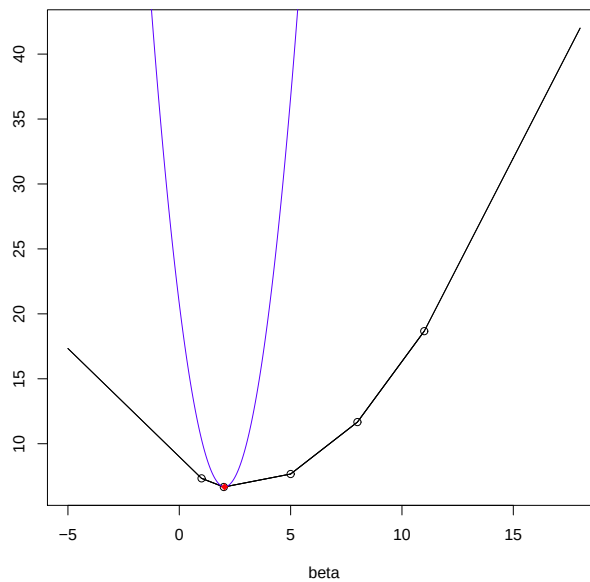
Obr. 5: Čierna krivka predstavuje funkciu $L(\beta) \doteq L_\epsilon(\beta)$. Modrá krivka je náhradná funkcia k účelovej v červenom bode β^k . Ide o 1.-4. iteráciu.



Obr. 6: Čierna krivka predstavuje funkciu $L(\beta) \doteq L_\epsilon(\beta)$. Modrá krivka je náhradná funkcia k účelovej v červenom bode β^k . Ide o 5.-8. iteráciu.



Obr. 7: Čierna krivka predstavuje funkciu $L(\beta) \doteq L_\epsilon(\beta)$. Modrá krivka je náhradná funkcia k účelovej v červenom bode β^k . Ide o 9.-12. iteráciu.



Obr. 8: Čierna krivka predstavuje funkciu $L(\beta) \doteq L_\epsilon(\beta)$. Modrá krivka je náhradná funkcia k účelovej v červenom bode β^k . Ide o poslednú 13. iteráciu, pričom sme dosiahli minimum.

4 Americké mzdové dáta

V tejto kapitole aplikujeme kvantilovú regresiu na mikroekonomické dáta. Dátový súbor, ktorý použijeme je voľne dostupný ¹. Dáta sú získané na základe prieskumov po celých Spojených štátoch amerických. Dátové súbory sú rozsiahle, obsahujú dáta okolo 2 miliónov respondentov. Dáta sa dajú získať ako pre domácnosti, tak i pre jedincov. Taktiež obsahujú dostatočné množstvo premenných od základných údajov, ako sú pohlavie, vek, rasa, až po údaje, ako príjem či zdravotné poistenie. My si vyberieme premenné podľa článku [1] od Angrista a spol. V tomto článku sa autor zaoberal americkými mzdovými dátami pre roky 1980, 1990 a 2000. My budeme analyzovať rok 2012. Avšak v modeli spravíme drobnú obmenu, vid [5].

4.1 Popis dát

Úpravou dátového súboru pre vybrané hodnoty premenných, ktoré navrhol Angrist, nám ostalo 94 027 údajov o jedincoch. Premenné, ktoré budeme zahŕňať do modelu sú:

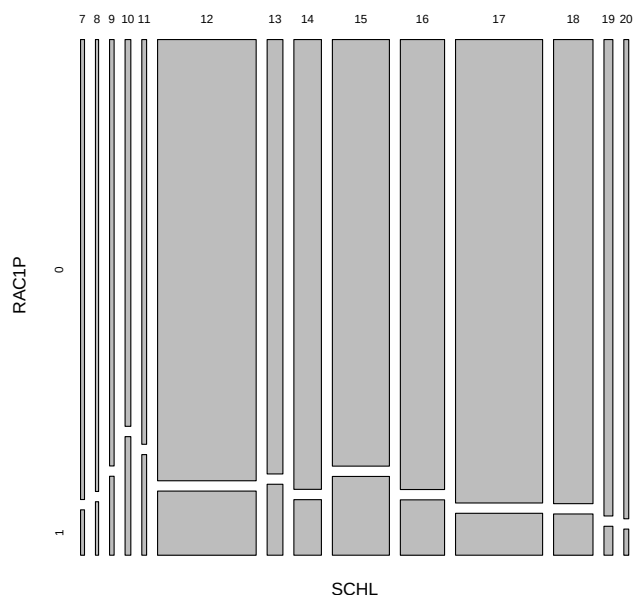
WAGP	Logaritmus týždenných príjmov vzhľadom k roku 1989 v dolároch.
AGEP	Vek v rokoch, s rozsahom 40-49.
RAC1P	Binárna premenná vyjadrujúca farbu pleti, 0 pre belochov a 1 pre černochoch.
SCHL	Dosiahnutý stupeň vzdelania.
EXP	Potenciálna skúsenosť vypočítaná ako $(AGEP - SCHL - 6)$.
EXP2	Druhá mocnina EXP.

Potenciálna skúsenosť EXP by mala byť vypočítaná ako rozdiel veku a počtu rokov štúdia a odrátania 6. Dáta však neobsahovali počet rokov štúdia jednotlivca a z toho dôvodu sme spravili danú aproximáciu.

¹Dáta sú dostupné na stránke <http://www.census.gov/>.

Dosiahnutý stupeň vzdelania nadobúda nasledujúce hodnoty:

- 7 Piaty až ôsmy stupeň vzdelania.
- 8, 9, 10 9., 10., 11. stupeň vzdelania.
- 11 12. stupeň vzdelania ale bez diplomu.
- 12 Stredná škola s maturitou.
- 13 Rozvoj všeobecného vzdelávania (GED) alebo náhradný diplom.
- 14 Vysoká škola, ale štúdium kratšie ako jeden rok.
- 15 Vysoká škola, viac ako jeden rok, avšak bez diplomu.
- 16 Associate's degree, obdoba diplomu z nadstavbového štúdia.
- 17 Dosiahnutý bakalár.
- 18 Dosiahnutý magister.
- 19 Professional degree².
- 20 Dosiahnutý doktorandský titul.



Obr. 9: Mozaikový graf.

²Professional degree je stupeň štúdia, ktorého absolventi sú pripravený na úlohu poradcov pre konkrétne profesie s dôrazom na vedomosti a praktické analýzy cez teóriu a výskum.

Dátový súbor obsahuje 94 027 dát, pričom belochoch je 83 272 a černochoch 10 755. Z obrázku 9 môžeme vidieť zastúpenie belochoch a černochoch v jednotlivých úrovniach vzdelania. Pomer belochoch voči černochochom je výrazne väčší pre vyššie úrovne vzdelania ako sú 17., 18., 19., 20., ale taktiež v nižšej úrovni 7. Černosi majú väčšie zastúpenie v nižších úrovniach štúdia, ale i v 13. a 15. úrovni. Je vidieť, že najatraktívnejšia úroveň štúdia je 12. a 17., teda stredná škola s maturitou a bakalár. Taktiež pomerná časť absolvuje 15., 16. a 18. úroveň.

Model, ktorý použil Angrist na svoje dáta má tvar:

$$WAGP = \beta_0 + \beta_1 RAC1P + \beta_2 EXP + \beta_3 EXP2 + \beta_4 SCHL + \varepsilon. \quad (48)$$

Premennú SCHL bral autor ako kvantitatívnu premennú a jej hodnoty vysvetlil ako počet rokov štúdia. Avšak je to iba jeho aproximácia, pretože v skutočnosti je premenná SCHL kategorická premenná, to znamená, že tvorcovia dát si mohli namiesto škály (7 – 20) zobrať úplne odlišnú, napríklad (107 – 120). V tomto prípade by aproximácia, ktorú Angrist použil bola úplne scestná. V [5] je kategorickosť danej premennej zohľadnená v modeli, ktorý má tvar:

$$\begin{aligned} WAGP = & \beta_0 + \beta_1 RAC1P + \beta_2 EXP + \beta_3 EXP2 + \beta_4 SCHL(7) + \beta_5 SCHL(8) + \beta_6 SCHL(9) \\ & + \beta_7 SCHL(10) + \beta_8 SCHL(11) + \beta_9 SCHL(12) + \beta_{10} SCHL(13) + \beta_{11} SCHL(14) \\ & + \beta_{12} SCHL(15) + \beta_{13} SCHL(16) + \beta_{14} SCHL(17) + \beta_{15} SCHL(18) + \beta_{16} SCHL(19) \\ & + \beta_{17} SCHL(20) + \varepsilon. \end{aligned} \quad (49)$$

V ďalšej podkapitole si predstavíme základné funkcie, ktoré sa používajú v systéme R na prácu s kvantilovou regresiou.

4.2 Kvantilová regresia v systéme R

Ak chceme pracovať s kvantilovou regresiou v systéme R, tak musíme ako prvé nainštalovať balíček `quantreg`. Najdôležitejšia funkcia ktorú budeme z daného balíčka používať bude funkcia `rq`. Táto funkcia sa využíva na odhad parametrov lineárneho modelu.

```
rq(formula, tau, data, method,...)
```

Argumenty tejto funkcie sú:

- formula** Predstavuje náš model, ktorého parametre chceme vypočítať, zadaný v klasickom tvare $y \sim x_1 + x_2 \dots$.
- tau** Hodnota kvantilu, pre ktorý chceme daný odhad.
 Ak je $\tau \in (0, 1)$ jedno číslo, dostaneme odhad pre dané τ .
 Ak je $\tau \in (0, 1)$ vektor, dostaneme odhady pre každú jeho zložku.
 Pre $\tau \notin (0, 1)$ dostaneme celý kvantilový proces.
- data** Dátový súbor, odkiaľ máme dáta.
- method** Volíme metódu, ktorú využijeme na odhad parametrov modelu. Na výber máme simplexovu metódu "**br**", Frisch-Newton algoritmus "**fn**" a modifikovaný Frisch-Newton algoritmus "**pfn**" pre extrémne veľké dátové súbory.

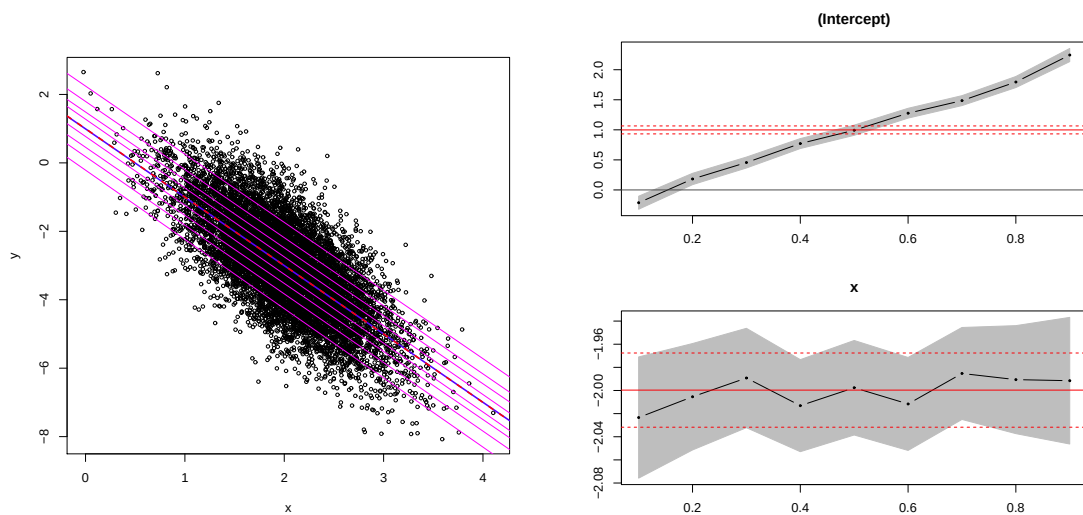
Na výstupe tejto funkcie dostaneme odhady daných parametrov modelu pre dané τ . Použitím funkcie `summary()` dostaneme i iné informácie o danom odhade, ako sú napríklad konfidenčné intervaly, štandardné odchýlky, t -štatistiky, p -hodnoty... . Keď použijeme `plot(summary(rq))`, dostaneme grafické znázornenie odhadov parametrov pre jednotlivé kvantily s príslušnými intervalmi spoľahlivosti a taktiež odhad pre podmienenú strednú hodnotu, tiež s príslušnými intervalmi spoľahlivosti. Z daného obrázku sa dajú vyčítať informácie o charaktere dát.

Zoberme do úvahy nasledujúce príklady.

1. Symetrické, homoskedastické dáta

Nech vektor x a ϵ sú z normálneho rozdelenia $x \sim N(2, 0.5)$ a $\epsilon \sim N(0, 1)$. Nech $\beta = (1, -2)$ a $y = 1 - 2x + \epsilon$. Nagenerujeme si 10 000 pozorovaní y a x a spravíme odhad parametrov β pre jednotlivé kvantily. Na obrázku 10 je znázornený bodový graf nagenovaných dát x verzus y a bodmi preložené priamky pre kvantilovú regresiu, kde jednotlivé kvantily sú $\{0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, 0.9\}$. Tieto priamky sú znázornené purpurovou farbou. Najspodnejšia priamka je pre $\tau = 0.1$ a pokračuje postupne až ku najvrchnejšej pre $\tau = 0.9$. Na danom obrázku je ešte znázornená modrá priamka pre medián a červená prerušovaná čiara pre podmienenú strednú hodnotu, ktorú sme dostali pomocou metódy najmenších štvorcov. Môžeme si všimnúť, že jednotlivé priamky pre

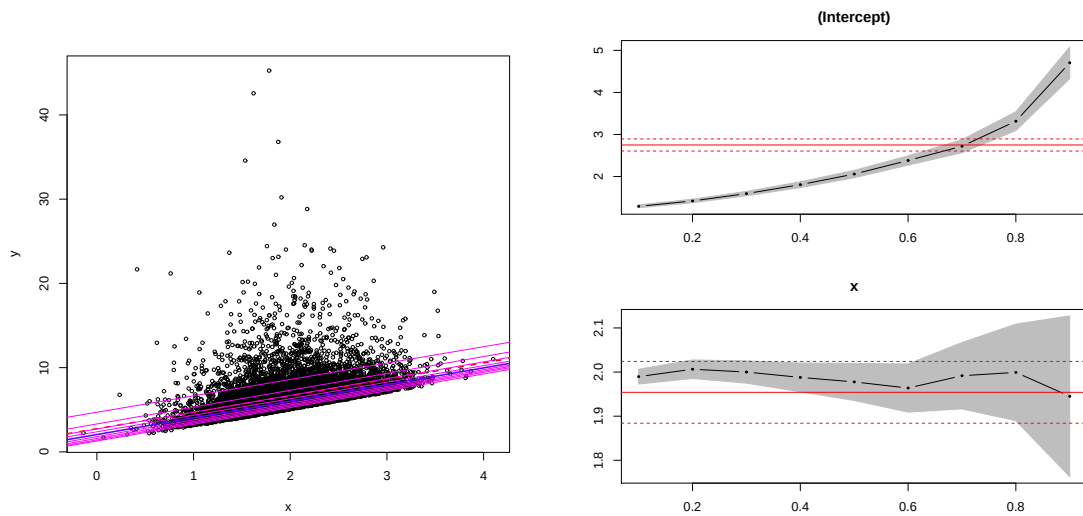
rôzne kvantily sú rovnobežné a medzery medzi nimi sú totožné. Rovnobežnosť priamok vyplýva z homoskedasticity dát a symetria rozdelenia vyplýva z toho, že mediánová priamka je totožná s priamkou podmienenej strednej hodnoty. Symetriu dát môžeme vidieť i na odhade interceptu pre jednotlivé kvantily. Ak je táto priamka približne pod uhlom 45° , tak sú dáta symetrické. Z odhadu pre sklon β_1 môžeme vidieť, že odhady pre jednotlivé kvantily sú totožné pre odhad podmienenej strednej hodnoty. Je to práve preto, že chyby ϵ sú nezávislé, rovnako rozdelené a z normálneho rozdelenia. Hovorí nám to o homoskedasticite.



Obr. 10: Odhad parametrov pre symetrické rozdelenie.

2. Asymetrické, homoskedastické dáta

Zoberieme si tie isté premenné ako v **1.** príklade, len ϵ budeme generovať s lognormálneho rozdelenia $\epsilon \sim LN(0, 1)$. Dostaneme asymetrické rozdelenie pre y . Priamky pre jednotlivé kvantily sú znova rovnobežné, avšak medzery medzi nimi sa s rastom kvantilov zväčšujú. Toto nám presne hovorí o asymetrickosti a z obrázku 11 môžeme vidieť, že je to ľavostranné asymetrické rozdelenie. Môžeme si tiež všimnúť, že aj odhad strednej hodnoty sa nám vzdialil oproti mediánu. Tiež sa dá vypožorovať, že odhad pre intercept nie je na 45° priamke, čo nám tiež indikuje, že dáta nie sú symetrické. Odhad pre sklon však zase vyšiel podobný odhadu podmienenej strednej hodnoty, to znamená, že dáta sú homoskedastické.

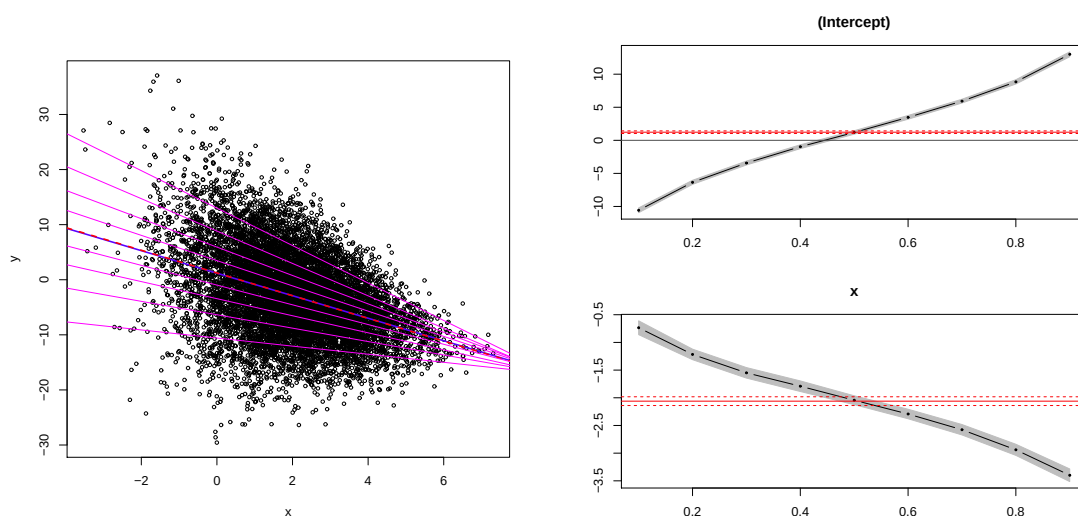


Obr. 11: Odhad parametrov pre asymetrické rozdelenie.

3. Symetrické, heteroskedastické dáta

Ako tretí príklad si zoberieme heteroskedastické dáta. Premenná x a ϵ sú rovnaké ako v ako pre 1. príklad, ale y budeme modelovať pomocou $y = 1 - 2x + (9 - x)\epsilon$. Obrázok 12 znázorňuje priamky pre jednotlivé kvantily, ktoré však už nie sú rovnobežné. Avšak mediánový odhad splýva s odhadom podmienenej strednej hodnoty, tak môžeme usúdiť že máme symetrické dáta. Vidíme to i z odhadu interceptu, ktorý leží na 45 stupňovej priamke. Sklon sa rapídne mení pre jednotlivé kvantili, čo nám potvrdzuje heteroskedasticitu dát.

Kvantilová regresia sa dá teda výborne využiť ako nástroj na preskúmanie dát. Na odhad parametrov pre nelineárnu kvantilovú regresiu sa využíva funkcia `nlrq()`. V nasledujúcej časti si odhadneme parametre nášho modelu v systéme R pomocou vyššie uvedenej funkcie `rq()`.



Obr. 12: Odhad parametrov pre heteroskedastické dáta.

4.3 Odhad parametrov modelu

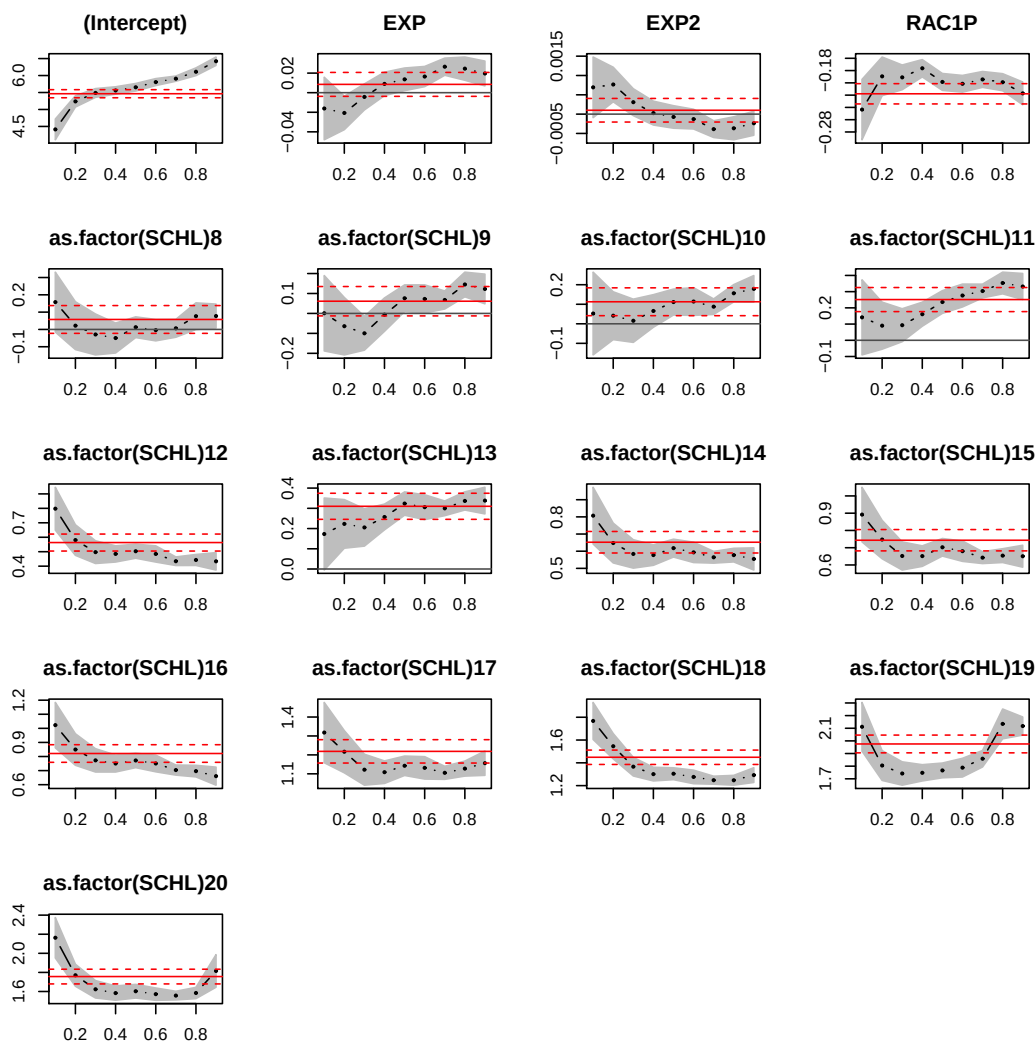
Odhad parametrov pre model (49) pomocou kvantilovej regresie, použitím algoritmu Frich-Newton získame nasledovne:

```
library(quantreg)
tau=seq(0.1,0.9, by=0.1 )
fit1=rq(WAGP~ EXP+EXP2+RAC1P+as.factor(SCHL), tau=tau, method="fn")
plot(summary(fit1))
```

Hodnoty parametrov pre jednotlivé vysvetľujúce premenné v závislosti na vybraných kvantiloch môžeme vidieť znázornené na obrázku 13 čiernou prerušovanou čiarou. Sivá oblasť predstavuje príslušný 90- percentný interval spoľahlivosti. Červené čiary sú hodnoty odhadu parametrov získané metódou najmenších štvorcov. Červené prerušované čiary sú príslušné 90- percentné intervaly spoľahlivosti pre dané odhady strednej hodnoty.

Intercept modelu (49) β_0 môžeme interpretovať ako odhad podmienenej kvantilovej funkcie pre rozdelenie logaritmu miezd WAGP belocha, ktorý má priemerné skúsenosti a dosiahol 5, 6 alebo 7 stupeň vzdelania. Všetky ostatné odhady parametrov modelu sú teda relatívne voči interceptu.

Skúsenosti EXP vstupujú do modelu kvadraticky. Obrázok 14 znázorňuje, aký vplyv

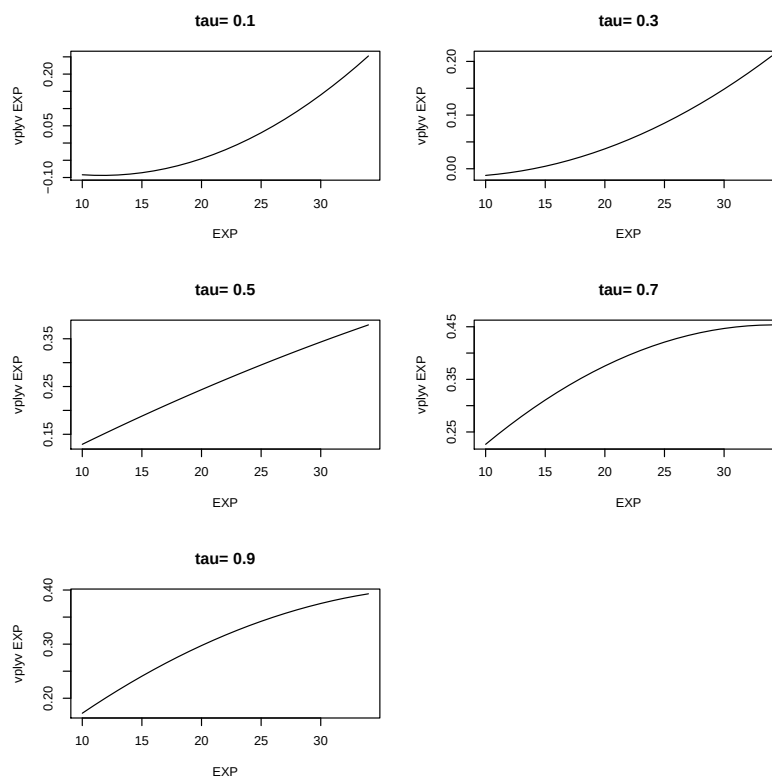


Obr. 13: Odhad parametrov modelu (49)

pre hodnoty $\tau = (0.1, 0.2, \dots, 0.9)$.

na WAGP majú skúsenosti EXP pre vybraté kvantily. Môžeme si všimnúť, že pre nižšie kvantily je funkcia konvexná a pre vyššie kvantily sa stáva viac konkávna. Pre všetky kvantily sa s rastúcimi skúsenosťami zvyšuje vplyv na WAGP. U nižších kvantilov s rastom skúseností rastie vplyv na WAGP oveľa pomalšie. Pre vyššie kvantily je vplyv oveľa väčší a s rastom skúseností rastie i rýchlejšie. Najviac zarábajúci občania, t.j. 0.9 kvantil, majú menší vplyv s rastom skúseností ako 0.7 kvantil.

Samostatná príslušnosť čiernej rasy má za príčinu pokles mzdy oproti belochom. Odhadom pre podmienenú strednú hodnotu metódou najmenších štvorcov je to skoro o 0.22. Môžeme si všimnúť, že černosi, ktorý majú nízku mzdu, síce na úrovni do 0.15 kvan-



Obr. 14: Vplyv skúseností na logaritmus týždenných miezd pre 5 rôznych kvantilov.

tilu majú rozdiel **WAGP** oproti belochom očividne vyšší, cez 0.24. Najbližšie ku mzdám belochov to majú černosí, ktorí majú **WAGP** okolo 0.4 kvantilu a tí zarábajú menej o 0.176. Môžeme si ešte všimnúť, že horný 10 percent zarábajúcich černochoch zarábajú aj tak o 0.22 **WAGP** menej ako belosi.

Všetky úrovne vzdelania nebudeme interpretovať, vyberieme si len niektoré, ktorých odhady sa zdajú byť najzaujímavejšie. Môžeme si všimnúť, že 12, 14, 15, 16, 17 a 18 úroveň majú podobný charakter. Tí, ktorí zarábajú najmenej, t.j. na úrovni 0.1 kvantilu majú najvyšší vplyv na **WAGP**. Pre ostatné kvantily sú odhady pomerne rovnaké, ich hodnota je pod hodnotou získanou metódou najmenších štvorcov.

Americkým občanom sa oplatí chodiť do školy. Čím vyšší level absolvujú, tým viac ich zárobky vzrastú oproti interceptu. Tí čo absolvujú maturitu, majú na úrovni 0.1 kvantilu vyššiu mzdu o 0.8. Ten čo skončí bakalára zvýši na tom istom kvantile mzdu o 1.3 a magister o 1.7 voči referenčnému občanovi. Pre ostatné kvantily platí, že s každým ďalším absolvovaným stupňom vzrastie **WAGP** približne o 0.1. Najväčší skok pre tieto

dané kvantily je z úrovne 16 na úroveň 17, kde sa z 0.7 dostanem na 1.1 WAGP.

13 úroveň má oproti predchádzajúcim úrovniam vzdelania skôr rastúci trend. Aj keď oproti 12 úrovni, čo je maturita, má na všetkých kvantiloch nižší nárast mzdy. Maximálny nárast dosahujú mediánoví občania a horných 20 percent zarábajúcich WAGP a to pod 0.4. Pričom na úrovni 12 dosahujú najnižší nárast horných 30 percent zarábajúcich a to je tesne nad 0.4.

19 úroveň má zaujímavý trend. Spočiatku taktiež klesá, pričom na 0.1 kvantile má mzdu vyššiu o 2.1 a od 0.2 – 0.7 má pomerne konštantné rozloženie a mzdu vyššiu oproti interceptu približne o 1.8 WAGP. Avšak pre zvyšné horné kvantily 0.8 a 0.9 začnú mzdy rásť a znova sa dostanú na úroveň 2.1.

Doktorandský titul má podobný charakter ako úroveň 19. Avšak okrem dolných 10 percent zarábajúcich WAGP, ktorí zvyšujú svoj príjem o 2.2, majú všetky ostatné kvantily menší vzrast príjmu ako v 19 úrovni.

Teraz sa pozrieme na roky, ktoré analyzoval Angrist. Budeme postupovať podľa [5] a skúsime sa pozrieť, ako sa vyvíjali niektoré parametre v priebehu rokov.

4.3.1 Angristove dáta a porovnanie vybraných premenných

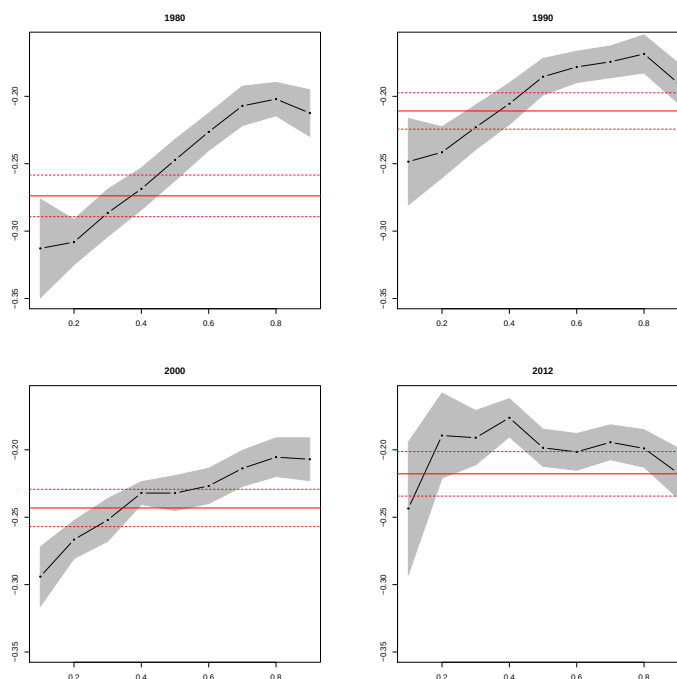
V prílohe A sú obrázky A.18, A.19 a A.20, na ktorých sú znázornené odhady koeficientov pre jednotlivé kvantily. Postupne zobrazujú rok 1980, 1990 a 2000. Vyberieme si niektoré zaujímavé premenné aby sme ich navzájom porovnali s našimi výsledkami.

Premenné, ktoré vstupujú do modelu majú odlišné značenie ako naše a niektoré úrovne vzdelania sa nezhodujú z našimi úrovňami. Nebudeme vypisovať jednotlivé úrovne vzdelania, sú tam iba malé zmeny a v našom porovnaní to nebude prekážať. Dajú sa však nájsť na stránke <<http://www.census.gov/>>, kde sú tieto dáta voľne dostupné.

Označenie premenných je nasledovné:

- logwk** Logaritmus týždenných príjmov vzhľadom k roku 1989 v dolároch, v našom prípade **WAGP**.
- age** Vek v rokoch, ohraňovaný pre osoby 40-49 ročné, v našom prípade **AGEP**.
- black** Binárna premenná vyjadrujúca farbu pleti, 0 pre belochov a 1 pre černochoch, v našom prípade **RAC1P**.
- educ** Dosiahnutý stupeň vzdelania, v našom prípade **SCHL**.
- exper** Potenciálna skúsenosť vypočítaná ako $(age - educ - 6)$, v našom prípade **EXP**.
- educ2** Druhá mocnina **educ**, v našom prípade **EXP2**.

Nasledujúci obrázok 15 nám znázorňuje vývoj **RAC1P** pre dané roky.



Obr. 15: Porovnanie vývoja premennej **RAC1P** podľa jednotlivých rokov 1980, 1990, 2000 a 2012.

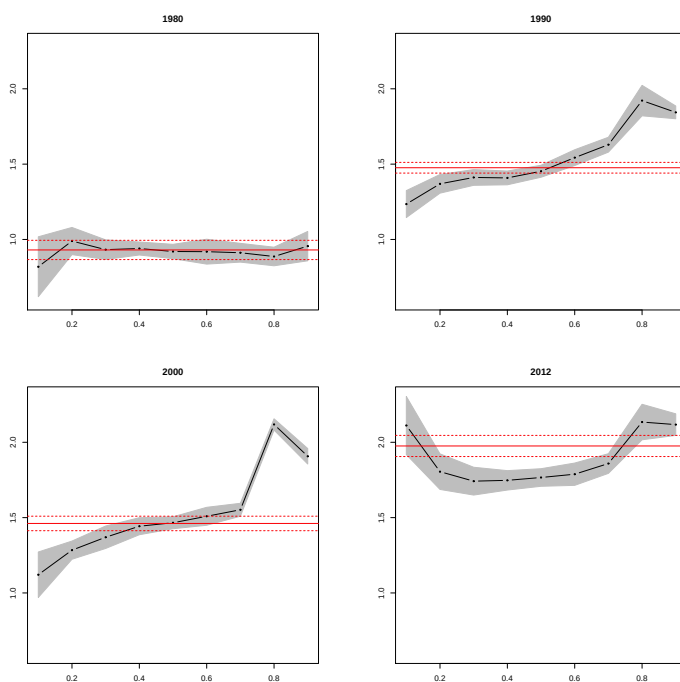
Ako prvé si môžeme všimnúť, že z roku 1980 na rok 1990 sa pre všetky kvantily posunul odhad parametrov smerom nahor. To znamená, že rozdiel v príjmoch medzi belochmi a černochochmi sa za 10 rokov zmenšil. V priemere sa tento rozdiel zmenšil z -0.27 na -0.21, teda o 0.06. Ani tvar vývoja podľa jednotlivých kvantilov sa rapídne nezmenil. Iba tí, ktorí najviac zarábali, to znamená horných 10 percent, ostali skoro na tom istom, ako

boli i v roku 1980.

Zmena znova nastalo o desať rokov. V roku 2000 sa prepadli a priblížili sa k roku 1980. Horných 40 percent WAGP sa dostali presne na rovnakú úroveň ako pred dvadsiatimi rokmi. Zvyšných 60 percent tiež poklesli oproti roku 1990, ale nie až na úroveň, v ktorej sa nachádzali v roku 1980. V priemere poklesli z -0.21 na -0.24, teda o 0.03.

Za ďalších 12 rokov sa znova situácia zlepšila. V priemere zrástol rozdiel z -0.24 na -0.22, čo je zlepšenie o 0.02. Nastala však veľká zmena v tvare. Kým doteraz mal tvar skôr rastúci trend, v roku 2012 to už tak nebolo. 10 percent najviac zarábajúcich černochoch sa prepadlo na úroveň -0.22, čo je pokles skoro o 0.02 WAGP. Naopak 0.4 kvantil sa vyšvihol o 0.04 WAGP a ich rozdiel oproti belochom bol na úrovni 0.17, čo je porovnateľné s 0.8 kvantilom v roku 1990. Ešte väčší progres mali na 0.2 kvantile, keď za 12 rokov sa ich rozdiel v príjmoch oproti belochom zmenšil o 0.08.

Obrázok 16 nám znázorňuje vývoj 19. úrovne vzdelania pre príslušné roky.



Obr. 16: Porovnanie vývoja premennej `as.factor(SCHL)19` podľa jednotlivých rokov 1980, 1990, 2000 a 2012.

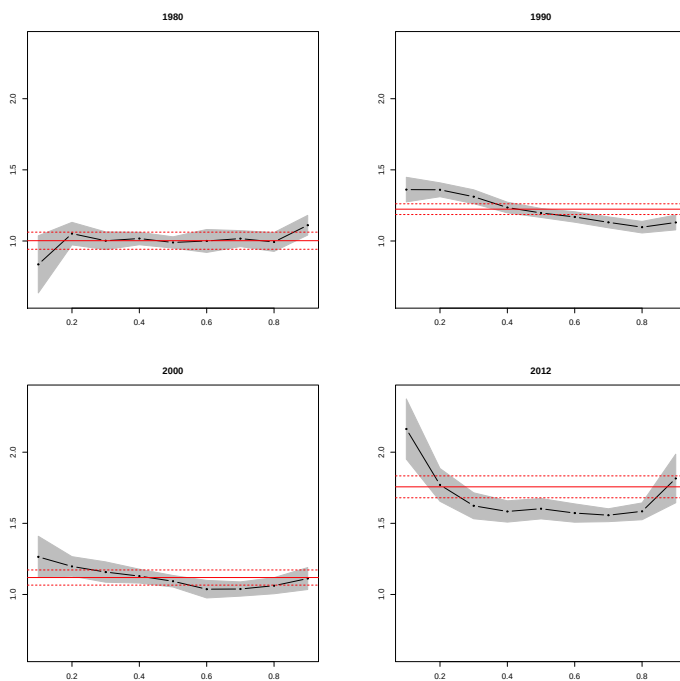
V roku 1980 bol efekt na mzdu skoro rovnaký ako pre málo zarábajúcich, tak pre tých, čo zarábali viac. O 10 rokov však nastal značný rozdiel. Vplyv na mzdu vzrástol a dostali sme celkom odlišný tvar. Je vidieť rastúci trend skoro na všetkých kvantiloch

okrem 0.9 a rapídne zvýšenie na 0.8 kvantile, kde efekt vzrástol o vyše 1 WAGP.

V roku 2000 bol tvar veľmi podobný ako v roku 1990 a zmena v efekte na WAGP nie je skoro žiadna až na 0.8 kvantile, kde sa vplyv zase zvýšil a to približne o 0.2.

V roku 2012 efekt na mzdu vzrástol a zmena v tvare je znova rapídna. O rastúcom trende sa už nedá ani hovoriť. Efekt mzdy pre najmenej zarábajúcich a to na úrovni 0.1 kvantilu vzrástol o 1, podobne ako v roku 1990 pre 0.8 kvantil. Ostatné kvantili mierne vzrástli, len 0.8 ostal na rovnakej úrovni.

Posledná premenná ktorú si porovnáme bude 20. úroveň vzdelania, teda dosiahnutý doktorandský titul. Na obrázku 17 máme znázornený vývoj 20. úrovne vzdelania pre dané roky.



Obr. 17: Porovnanie vývoja premennej `as.factor(SCHL)20` podľa jednotlivých rokov 1980, 1990, 2000 a 2012.

V roku 1980 bolo rozloženie na jednotlivých kvantiloch pomerne konštantné okolo hodnoty 1 WAGP. Len dolných 10 percent zarábajúcich bolo na úrovni tesne nad 0.8 a horných 10 percent na úrovni 1.1.

V roku 1990 sa tvar odhadov pre jednotlivé kvantili mierne zmenil, nabral klesajúci trend skoro na všetkých kvantiloch. Vplyv na WAGP sa zvýšil pre všetky kvantily v prie-

mere o viac ako 0.2. Avšak najväčší skok sa stal pre 0.1 kvantil a to o 0.55. 10 percent najviac zarábajúcich zostalo skoro na rovnakej úrovni, ako boli pred desiatimi rokmi. V roku 2000 nastal pokles na všetkých kvantiloch v priemere o vyše 1 WAGP. Ani tvar sa výrazne nezmenil. Zaujímavejší je však rok 2012. V priemere sa vplyv na mzdy zvýšil o 0.65. Na 0.3 - 0.8 kvantile je vplyv rovnaký okolo 1.6 WAGP. Na chvostoch je však vplyv výrazne vyšší. 10 percent najmenej zarábajúcich má vplyv dokonca skoro 2.2. Zaujímavé je, že v porovnaní s 19. úrovňou vzdelania je vplyv na mzdu pre 20 percent najviac zarábajúcich oveľa menší od roku 1990.

Záver

V diplomovej práci sme sa zaoberali kvantilovou regresiou v ekonometrii. V prvej kapitole sme si predstavili základnú teóriu kvantilovej regresie, odvodili sme si optimalizačnú úlohu kvantilovej regresie na odhad parametrov modelu. Pomocou smerovej derivácie sme zaviedli podmienku optimality pre danú úlohu. Zaoberali sme sa vlastnosťami odhadov, ako sú ekvivariancia škály, ekvivariancia posunu a ekvivariancia preparametrizovania matice plánu. Na záver tejto kapitoly sme si ukázali asymptotickú normalitu odhadov, ktorá je úzko spätá s teóriou jednorozmerných výberových kvantilov. Pomocou Waldovho testu sme naznačili, ako sa testujú hypotézy o parametroch modelu pre kvantilovú regresiu.

Druhá kapitola sa zaoberala riešením optimalizačnej úlohy kvantilovej regresie pomocou teórie lineárneho programovania. Úloha lineárnej kvantilovej regresie sa dá, za pomoci nových premenných, previesť na úlohy lineárneho programovania. Na začiatku kapitoly sme rozobrali základnú teóriu lineárneho programovania. Predstavili sme si modifikovanú simplexovú metódu na riešenie úlohy kvantilovej regresie, ktorá využíva podmienku optimality z prvej kapitoly. Pre veľké dátové súbory je však modifikovaný simplexov algoritmus nepraktický, a preto sme predstavili i alternatívny Frisch-Newtonov algoritmus pre odhad parametrov kvantilovej regresie pomocou teórie vnútorného bodu.

V tretej kapitole sme sa zaoberali MM algoritmom. Ukázali sme, že známejší EM algoritmus je špeciálny prípad MM algoritmu. MM algoritmus sme využili na odhad parametrov modelu pre úlohu kvantilovej regresie. Naprogramovali sme ho v systéme R. Pri programovaní sme narazili na malé nezrovnalosti v článku [11], z ktorého sme vychádzali. Tieto nezrovnalosti sú popísané v časti (3.2.2). V niektorých prípadoch nám nedokonvergoval MM algoritmus k optimu danej úlohy, ale zostal len v jeho blízkosti. Na záver tejto kapitoly sme ilustrovali fungovanie MM algoritmus na jednoduchom príklade.

Posledná, štvrtá kapitola analyzuje Americké mzdové dáta pre kvantilovú regresiu z roku 2012. V tejto kapitole sme ukázali ako sa pracuje s kvantilovou regresiou v systéme R. Ukázali sme, čo všetko nám môže kvantilová regresia prezradiť o charaktere dát. Môžeme zhrnúť, že kvantilová regresia je veľmi prospešný nástroj na prieskum dát.

Podrobne sme opísali význam odhadov parametrov kvantilovej regresie pre jednotlivé premenné. Na záver sme porovnali naše výsledky s výsledkami z predchádzajúcich rokov, aby sme videli, ako sa dané premenná vyvíjali v priebehu rokov.

Teória kvantilovej regresie sa od doby, čo ju v roku 1978 predstavil Koenker a Bassett výrazne rozšírila. Účelom tejto práce bolo priniesť čitateľovi základne poznatky z teórie kvantilovej regresie a predstaviť algoritmy na získanie odhadov modelu pre jednotlivé kvantily.

Zoznam použitej literatúry

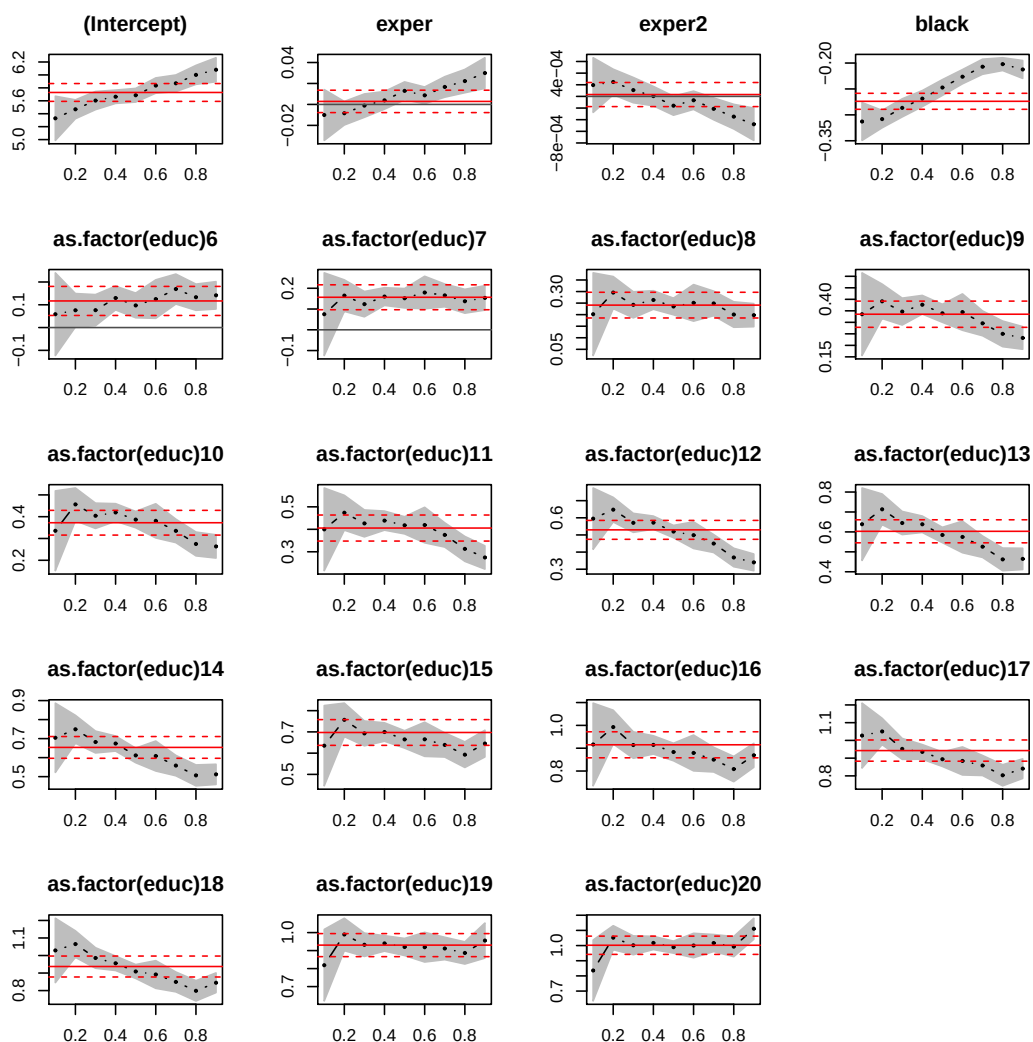
- [1] ANGRIST, J. - CHERNOZHUKOV, V. - FERNÁNDEZ-VAL, I.: *Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure*, Econometrica Volume 74, Issue 2, pages 539–563, March 2006 , dostupné na internete (17.4.2014): <http://economics.mit.edu/files/29>
- [2] BARRODALE, I. - ROBERTS, F.: *Solution of an Overdetermined System of Equations in l_1 Norm*, Communications of the ACM, 17, 319-320, 1974
- [3] BERMAN, A. - BEN-ISRAEL, A.: *Cones, Matrices and Mathematical Programming*, Springer, New York, 1973
- [4] DAVINO, C. - FURNO, M. - VISTOCCO, D.: *Quantile Regression: Theory and Applications*, Wiley, Chichester, 2013
- [5] GRENDÁR, M.: *Quantile regression*, 2014, učebné slajdy
- [6] HALICKÁ, M.: *Dvadsať rokov moderných metód vnútorného bodu*, Pokroky matematiky, fyziky a astronómie, Vol. 49, Iss. 3, 244-257, 2004, Dostupné na internete (10.4.2014): <http://www.dml.cz/handle/10338.dmlcz/141233>
- [7] HAMALA, M. - TRNOVSKÁ, M.: *Nelineárne programovanie*, Epos, Bratislava, 2013
- [8] HASONOVÁ, Z.: *Algoritmy primárno-duálnych metód vnútorného bodu na riešenie úloh lineárneho programovania* [Diplomová práca], Univerzita Komenského, Fakulta matematiky, fyziky a informatiky, Bratislava, 2007, dostupné na internete (12.4.2014): <http://www.iam.fmph.uniba.sk/studium/efm/diplomovky/2007/hasonova/diplomovka.pdf>
- [9] HUNTER, D. R.: *On the Geometry of EM algorithms*, Penn State Department of Statistics Technical Report 0303, February 2003, dostupné na internete (27.3.2014): <http://sites.stat.psu.edu/~dhunter/papers/tr0303.pdf>

- [10] HUNTER, D. R., - LANGE, K.: *A Tutorial on MM Algorithms*, The American Statistician, Vol. 58, No. 1, February 2004, dostupné na internete (27.3.2014): <http://med.stanford.edu/tanglab/courses/gen245/Reading/HunterLangeMMTutorial.pdf>
- [11] HUNTER, D. R. - LANGE, K.: *Quantile Regression via an MM Algorithm*, Journal of Computational and Graphical Statistics Vol. 9, No. 1 (Mar., 2000), 60-77, dostupné na internete (20.2.2014): <http://www.sites.stat.psu.edu/dhunter/papers/qr.pdf>
- [12] KABÁT, M.: *Metódy vnútorného bodu v lineárnom programovaní a ich aplikácie vo financiách* [Diplomová práca], Univerzita Komenského, Fakulta matematiky, fyziky a informatiky, Bratislava, 2013, dostupné na internete (10.4.2014): <http://www.iam.fmph.uniba.sk/studium/efm/diplomovky/2013/kabat/diplomovka.pdf>
- [13] KOENKER, R.: *Quantile Regression*, Cambridge University Press, Cambridge, 2005
- [14] KOENKER, R. - BASSETT, G.: *Regression Quantiles*, Econometrica, Vol. 46, No.1., pp. 33-50, (Jan., 1978), dostupné na internete (29.3.2014): <http://www.stanford.edu/doubleh/otherpapers/koenker.pdf>
- [15] KOENKER, R. - D'OREY, V.: *Computing Regression Quantiles*, Applied Statistics, 36, 383-393, 1987
- [16] KOENKER, R. - PORTNOY, S.: *The Gaussian Hare and the Laplacian Tortoise: Computability of Squared-Error Versus Absolute-Error Estimators*, Statistical Science, 12, 279-300, 1997
- [17] PLESNÍK, J. - DUPAČOVÁ, J. - VLACH, M.: *Lineárne programovanie*, Alfa, Bratislava, 1990
- [18] POLIAK, M.: *Centrálna trajektória v lineárnom programovaní* [Bakalárska práca], Univerzita Komenského, Fakulta matematiky, fyziky a informatiky, Bratislava, 2012, dostupné na internete (10.4.2014): <http://www.iam.fmph.uniba.sk/efm/bakalarky/2012/poliak/bakalarka.pdf>

- [19] ROOS, C. - TERLAKY, T. - VIAL, J. P.: *Interior point methods for linear optimization*, Springer, New York, 2005
- [20] WU, T. T. - LANGE, K.: *The MM Alternative to EM*, Statistical Science Vol. 25, No. 4, 492-505, 2010, dostupné na internete (27.3.2014): <http://arxiv.org/pdf/1104.2203.pdf>

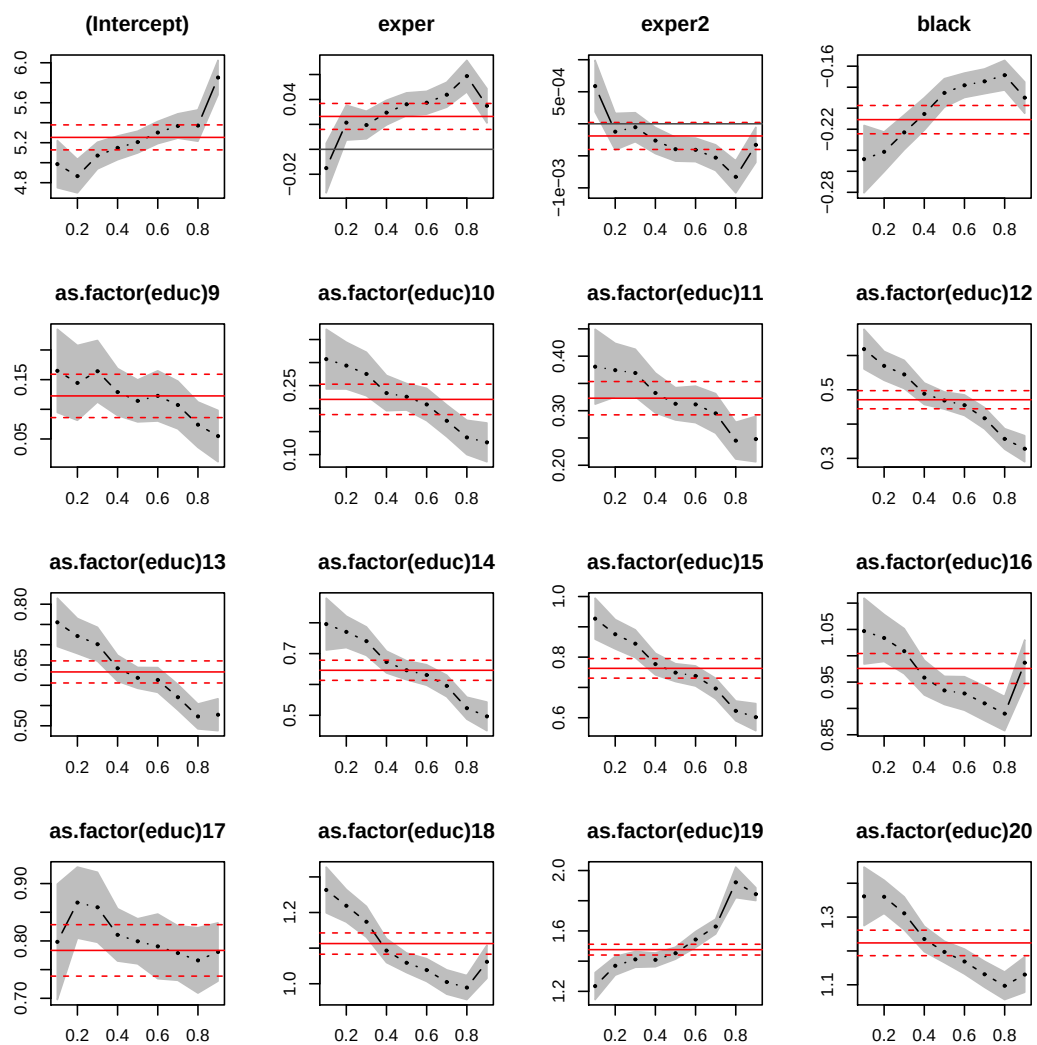
Príloha

Odhady parametrov modelu pre rok 1980



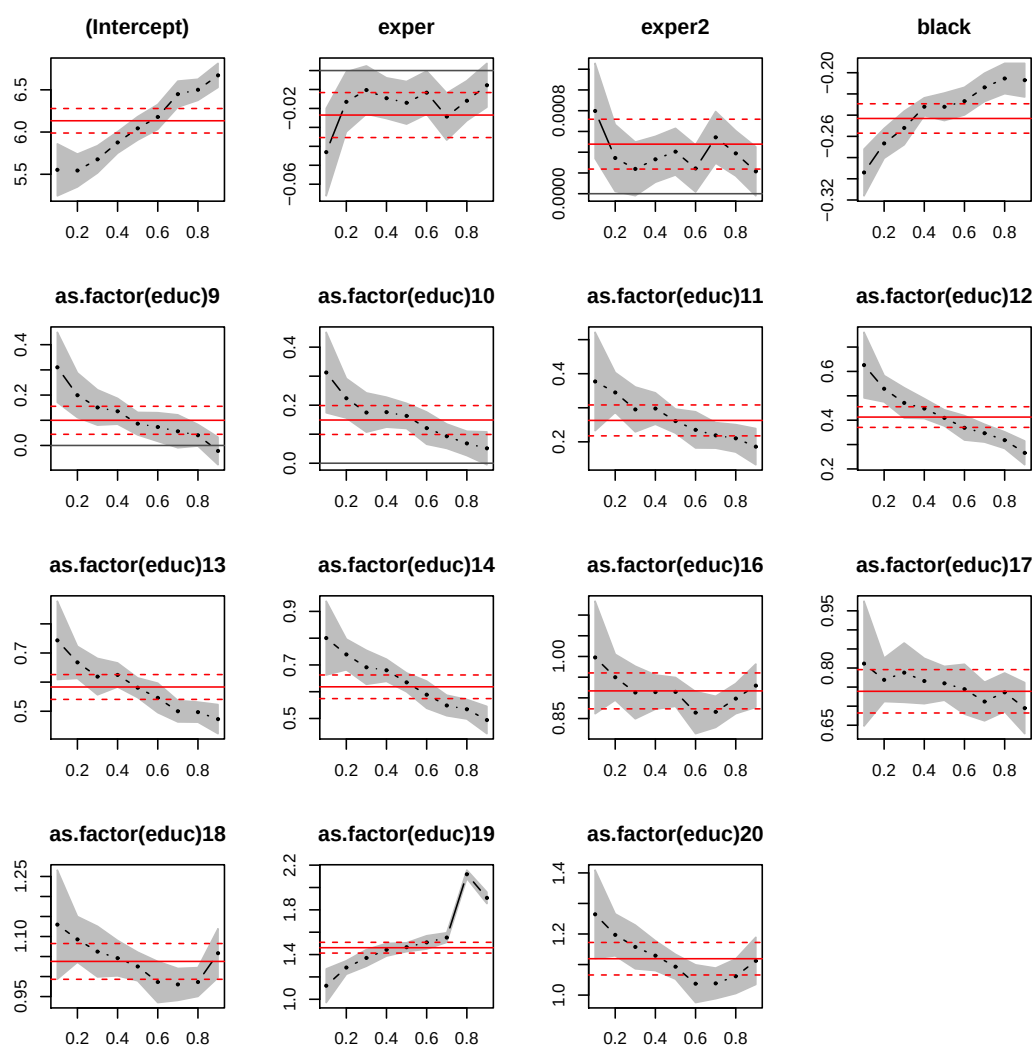
Obr. A.18: Odhad parametrov modelu (49) pre rok 1980.

Odhady parametrov modelu pre rok 1990



Obr. A.19: Odhad parametrov modelu (49) pre rok 1990.

Odhady parametrov modelu pre 2000



Obr. A.20: Odhad parametrov modelu (49) pre rok 2000.

Príkazy v systéme R na odhad parametrov modelu (49) pre dané roky

```

#načítanie dát z textového súboru 2012.txt
mydata<-read.table("C:\\Users\\Lucia\\Desktop\\Dipl_data\\2012.txt",header=T,sep="\t")
attach(mydata)
dim(mydata)

library(vcd)
crosstab <- table(RAC1P , SCHL)
setEPS()
postscript("mosaic.eps")
#mozaikový diagram pre zastúpenie rasy v jednotlivých úrovniach vzdelania
mosaicplot(crosstab,sort=c(2,1),dir=c("h","v"),main = " ")
dev.off()

#balíček na použitie kvantilovej regresie
library(quantreg)
#EXP pre vybrané kvantily
setEPS()
postscript("v2.eps")
par(mfrow=c(3,2))
tau=c(0.1,0.3,0.5,0.7,0.9)
for (i in 1:length(tau)){
fit=rq(WAGP~EXP+EXP2+RAC1P+as.factor(SCHL), tau=tau[i], data=mydata,method="fn")
x<-sort( unique(EXP, fromLast = TRUE))
plot(x,fit$coef[2]*x+fit$coef[3]*x^2,type="l",xlab="EXP",ylab="vplyv EXP",
      main=paste("tau=",tau[i]))
}
dev.off()

EXP<-EXP-mean(EXP)
EXP2<-EXP2-mean(EXP2)
#nastavenie kvantilov tau\in(0.1, 0.2, ..., 0.9)
tau=seq(0.1,0.9, by=0.1 )
#funkcia pre odhad parametrov kvantilovou regresiou
fit1=rq(WAGP~EXP+EXP2+RAC1P+as.factor(SCHL), tau=tau, data=mydata,method="fn")
summary(fit1)

#odhady jednotlivých parametrov pre tau
setEPS()
postscript("v1.eps")
plot(summary(fit1))
dev.off()
#-----
#Angristove dáta pre rok 1980
library(foreign)
mydata80<-read.dta("C://Users//Lucia//Desktop//data//census80.dta")
attach(mydata80)
exper<-exper-mean(exper)

```

```

exper2<-exper2-mean(exper2)
tau=seq(0.1,0.9, by=0.1 )
fit2=rq(logwk*perwt~exper+exper2+black+as.factor(educ), tau=tau, data=mydata80)
summary(fit2)

#odhady jednotlivých parametrov pre tau
setEPS()
postscript("v3.eps")
plot(summary(fit2))
dev.off()

#-----
#Angristove dáta pre rok 1990
mydata90<-read.dta("C://Users//Lucia//Desktop//data//census90.dta")
attach(mydata90)

tau=seq(0.1,0.9, by=0.1 )
exper<-exper-mean(exper)
exper2<-exper2-mean(exper2)
fit3=rq(logwk*perwt~exper+exper2+black+as.factor(educ),
tau=tau, data=mydata90,method="fn")
summary(fit3)

#odhady jednotlivých parametrov pre tau
setEPS()
postscript("v4.eps")
plot(summary(fit3))
dev.off()

#-----
#Angristove dáta pre rok 2000
mydata00<-read.dta("C://Users//Lucia//Desktop//data//census00.dta")
attach(mydata00)

tau=seq(0.1,0.9, by=0.1 )
exper<-exper-mean(exper)
exper2<-exper2-mean(exper2)
fit4=rq(logwk~exper+exper2+black+as.factor(educ), tau=tau, data=mydata00)
summary(fit4)

#odhady jednotlivých parametrov pre tau
setEPS()
postscript("v5.eps")
plot(summary(fit4))
dev.off()

#-----
#porovnanie RAC1P pre jednotlivé roky
setEPS()
postscript("v9.eps")
plot(summary(fit1),parm=c("RAC1P"),main="2012",ylim=c(-0.35, -0.16))

```

```
dev.off()

setEPS()
postscript("v6.eps")
plot(summary(fit2), parm=c("black"), main="1980", ylim=c(-0.35, -0.16))
dev.off()

setEPS()
postscript("v7.eps")
plot(summary(fit3), parm=c("black"), main="1990", ylim=c(-0.35, -0.16))
dev.off()

setEPS()
postscript("v8.eps")
plot(summary(fit4), parm=c("black"), main="2000", ylim=c(-0.35, -0.16))
dev.off()

#-----
#porovnanie (SCHL)19 pre jednotlivé roky
setEPS()
postscript("v13.eps")
plot(summary(fit1), parm=c("as.factor(SCHL)19"), main="2012", ylim=c(0.6, 2.3))
dev.off()

setEPS()
postscript("v10.eps")
plot(summary(fit2), parm=c("as.factor(educ)19"), main="1980", ylim=c(0.6, 2.3))
dev.off()

setEPS()
postscript("v11.eps")
plot(summary(fit3), parm=c("as.factor(educ)19"), main="1990", ylim=c(0.6, 2.3))
dev.off()

setEPS()
postscript("v12.eps")
plot(summary(fit4), parm=c("as.factor(educ)19"), main="2000", ylim=c(0.6, 2.3))
dev.off()

#-----
#porovnanie (SCHL)20 pre jednotlivé roky
setEPS()
postscript("v17.eps")
plot(summary(fit1), parm=c("as.factor(SCHL)20"), main="2012", ylim=c(0.6, 2.4))
dev.off()

setEPS()
postscript("v14.eps")
plot(summary(fit2), parm=c("as.factor(educ)20"), main="1980", ylim=c(0.6, 2.4))
dev.off()
```

```
setEPS()  
postscript("v15.eps")  
plot(summary(fit3),parm=c("as.factor(educ)20"),main="1990",ylim=c(0.6, 2.4))  
dev.off()
```

```
setEPS()  
postscript("v16.eps")  
plot(summary(fit4),parm=c("as.factor(educ)20"),main="2000",ylim=c(0.6, 2.4))  
dev.off()
```

```
#-----
```

Príkazy v systéme R na vykreslenie obrázkov 10, 11 a 12

```

#symetrické, homoskedastické dáta
eps<-rnorm(10000,mean=0,sd=1)
x<-rnorm(10000, mean = 2, sd = 0.5)
y=1-2*x+eps
plot(density(y))
library(quantreg)
tau=seq(0.1,0.9,by=0.1)
fit=rq(y~x,tau=tau)
fit1=lm(y~x)
setEPS()
postscript("pr1.eps")
plot(x, y, cex = 0.25, type = "n")
points(x, y, cex = 0.6)
abline(rq(y ~ x, tau = 0.5), col = "blue",lwd=2)
abline(lm(y ~ x), lty = 2, col = "red",lwd=2)
taus <- c(0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, 0.9)
for (i in 1:length(taus)) {
  abline(rq(y ~ x, tau = taus[i]), col="magenta")
}
dev.off()
setEPS()
postscript("pr2.eps")
plot(summary(fit))
dev.off()

#-----
#asymetrické, homoskedastické dáta
eps<-rlnorm(10000, meanlog = 0, sdlog = 1)
x<-rnorm(10000, mean = 2, sd = 0.5)
y=1+2*x+eps
plot(density(y))
library(quantreg)
tau=seq(0.1,0.9,by=0.1)
fit=rq(y~x,tau=tau)
fit1=lm(y~x)
setEPS()
postscript("pr5.eps")
plot(x, y, cex = 0.25, type = "n")
points(x, y, cex = 0.6)
abline(rq(y ~ x, tau = 0.5), col = "blue",lwd=2)
abline(lm(y ~ x), lty = 2, col = "red",lwd=2)
taus <- c(0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, 0.9)
for (i in 1:length(taus)) {
  abline(rq(y ~ x, tau = taus[i]), col="magenta")
}
dev.off()
setEPS()

```

```
postscript("pr6.eps")
plot(summary(fit))
dev.off()
#-----
#symetrické, heteroskedastické dáta
eps<-rnorm(10000,mean=0,sd=1)
x<-rnorm(10000, mean = 2, sd = 1.5)
y=1-2*x+(9-x)*eps
plot(density(y))
library(quantreg)
tau=seq(0.1,0.9,by=0.1)
fit=rq(y~x,tau=tau)
fit1=lm(y~x)
setEPS()
postscript("pr3.eps")
  plot(x, y, cex = 0.25, type = "n")
points(x, y, cex = 0.6)
abline(rq(y ~ x, tau = 0.5), col = "blue",lwd=2)
abline(lm(y ~ x), lty = 2, col = "red",lwd=2)
  taus <- c(0.1, 0.2, 0.3, 0.4, 0.6, 0.7, 0.8, 0.9)
for (i in 1:length(taus)) {
  abline(rq(y ~ x, tau = taus[i]), col="magenta")
}
dev.off()
setEPS()
postscript("pr4.eps")
plot(summary(fit))
dev.off()
#-----
```


MM algoritmus

```

# majorizačná funkcia (viacrozmerna)
fun1 <- function(tau,eps,beta,y,x,betak) {
  r=y-x**beta
  rr=y-x**betak
  for (i in 1:length(r))
    v[i]=ifelse(r[i]<0,1,0)
  for (l in 1:length(r))
    vv[l]=ifelse(rr[l]<0,1,0)
  rr1 <- abs(rr)
  #vypočet vektora konstant c
  c=4*((rr*(tau-vv)-eps/2*(log(eps+abs(rr)))))-(rr^2/(eps+rr1)+(4*tau-2)*rr)
  sum(1/4*(r^2/(eps+rr1)+(4*tau-2)*r+c))
}

#f majorizačná funkcia(jednorozmerna)
fun2 <- function(tau,eps,beta,y,x,betak) {
  r=y-x*beta
  for (i in 1:length(r))
    v[i]=ifelse(r[i]<0,1,0)
  rr=y-x*betak
  for (l in 1:length(r))
    vv[l]=ifelse(rr[l]<0,1,0)
  rr1 <- abs(rr)
  c=4*((rr*(tau-vv)-eps/2*(log(eps+(abs(rr))))))-(rr^2/(eps+rr1)+(4*tau-2)*rr)
  sum(1/4*(r^2/(eps+rr1)+(4*tau-2)*r+c))
}

#MM algoritmus
MM<-function(qtau,ni,y,X){
  x<-cbind(rep(1,length(y)),X)
  eps<-uniroot(function(epsilon) epsilon*length(y)*abs(log(epsilon))-ni , lower =
    10^(-20), upper = 0.9, tol = 1e-9)$root
  vysledok<-matrix(data=NA,nrow=length(qtau),ncol=dim(x)[2])
  #pre ubohy model(model len s interceptom)
  if (dim(x)[2]==1){
    for(t in 1:length(qtau)) {
      tau<-qtau[t]
      betak<-lm(formula=y~-1+x)$coeff #startovacia hodnota
      betakk<-optim(betak,fn=fun2,tau=tau,eps=eps,x=x,y=y,betak=betak,method ="Brent"
        ,lower = -10^6, upper = 10^6)$par
      while ((fun2(tau,eps,betak,y,x,betak)-fun2(tau,eps,betakk,y,x,betak))>ni)
      {
        betak<-betakk
        solve<-optim(betak,fn=fun2,tau=tau,eps=eps,x=x,y=y,betak=betak,method ="
          Brent",lower = -10^6, upper = 10^6)
        betakk<-solve$par
      }
      vysledok[t,]<-betakk
    }
  }
}

```

```

    }

    }

    #pre viac parametrov vratane interceptu
else {
  for(t in 1:length(qtau)) {
    tau<-qtau[t]
    betak<-lm(formula=y~-1+x)$coeff #pociatocna hodnota beta0
    betakk<-optim(betak,fn=fun1,tau=tau,eps=eps,x=x,y=y,betak=betak)$par
    while ((fun1(tau,eps,betak,y,x,betak)-fun1(tau=tau,eps,betakk,y,x,betak))>ni)
    {
      betak<-betakk
      solve<-optim(betak,fn=fun1,tau=tau,eps=eps,x=x,y=y,betak=betak)
      betakk<-solve$par
    }
    vysledok[t,]<-betakk
  }

  }

rownames(vysledok, do.NULL = TRUE)
rownames(vysledok) <- c(qtau)
t(vysledok)

}

```