Univerzita Karlova v Praze
Matematicko-fyzikální fakulta

BAKALÁŘSKÁ PRÁCE

Miroslav Svoboda

Průběžné odhadování modelů závislostí diskrétních veličin na spojitých s aplikací na obchodování s futures

Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce:	Ing. Miroslav Kárný, DrSc.
Studijní program:	Matematika
Studijní obor:	Obecná matematika

Praha 2013

V prvom rade by som chcel poďakovať Ing. Miroslavi Kárnému, DrSc. za cenné rady a vytvorenie vhodných podmienok na spracovanie bakalárskej práce. Tiež by som chcel poďakovať svojej rodine za podporu pri štúdiu a písaní.

Tato práce bola podporená projektom GAČR 13-13502S.

Prohlašuji, že jsem tuto bakalářskou práci vypracoval(a) samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

V .... dne ....   Podpis autora

Název práce: Průběžné odhadování modelů závislostí diskrétních veličin na spojitých s aplikací na obchodování s futures

Autor: Miroslav Svoboda

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: Ing. Miroslav Kárný, DrSc., Ústav teorie informace a automatizace AV ČR, v.v.i

Abstrakt: Tato práce se zaobírá průběžným odhadováním závislosti modelů dis krétních veličin na náhodných veličinách, které mají diskrétní nebo spojité rozdělení. Využívá se na to Bayesův vzorec, popsaný v první kapitole, kde je přidaný předpoklad podmíněné nezávislosti, aby jej bylo možné používat dynamicky. Ve druhé kapitole je následně popsán aproximační algoritmus, pomocí kterého se průběžně aproximuje bayesovsky odhadnutá hustota náhodné veličiny. Celý tento postup je aplikován na speciální tvar modelu logistické regrese, popsaný ve třetí kapitole. Výsledek je dále ukázán na příkladech se simulovanými daty. Nakonec je model spolu s aproximačním mechanismem vyskoušen aplikovat na obchodování s futures.

Klíčová slova: Bayesovske odhadovani, rekurzivny algoritmus, logisticka regrese, futures

Title: Recursive estimation of models relating discrete-valued variables to continuous-valued ones applied to trading with futures

Author: Miroslav Svoboda

Department: Department of Probability and Mathematical Statistics

Supervisor: Ing. Miroslav Kárný, DrSc., Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Department of Adaptive Systems

Abstract: This bachelor thesis deals with recursive estimation of a dependence of the models with discrete variables on variables that are either discretely or continuously distributed. To this purpose Bayes formula, described in the first chapter, is used, to which an additional assumption of conditional independence is added so that it can be used dynamically. The second chapter describes an approximation algorithm, which is used for recursive approximation of the density of random variable that has been estimated by the Bayesian equation. The third chapter deals with the application of the whole model on a special form of logistic regression. Results are shown on the examples using simulated data. At last, the model along with approximation algorithm is applied on a trading with futures.

Keywords: Bayesian estimation, recursive algorithm, logistic regression, futures

Názov práce: Priebežné odhadovanie modelov závislostí diskrétnych veličín na spojitých s aplikáciou na obchodovanie s futures.

Autor: Miroslav Svoboda

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedúci bakalárskej práce: Ing. Miroslav Kárný, DrSc., Ústav teorie informace a automatizace AV ČR, v.v.i

Abstrakt: Táto práca sa zaoberá priebežným odhadovaním závislosti modelov diskrétnych veličín na náhodných veličinách ktoré majú diskrétne alebo spojité rozdelenie. Využíva sa na to Bayesov vzorec popísaný v prvej kapitole kde je pridaný predpoklad podmienenej nezávislosti aby ho bolo možné používat dinamicky. V druhej kapitole je následne popísaný aproximačný algoritmus, pomocou ktorého sa priebežne aproximuje bayesovsky odhadnutá hustota náhodnej veličiny. Celý tento postup je aplikovaný na špeciálny tvar modelu logistickej regresie popísanej v tretej kapitole. Výsledok je ukázaný na príkladoch so simulovanými datami. Nakoniec je model spolu s aproximačným algoritmom vyskúšaný aplikovať na obchodovanie s futures.

Kľúčové slová: Bayesovske odhadovanie, rekurzivny algoritmus, logisticka regresia, futures

Obsah

Použité značenie

θ	regresný vektor
Y	vysvetľovaná premenná
X	vysvetľujúca premenná
$Z = (X,Y)$	
$g_{\theta\|X}$	hustota získaná bayesovským odhadom
$D(g\|\|p)$	Kullback-Leiblerova divergencia
p	aproximovaná hustota
GiW	rodina rozdelen *Gauss-Invers-wishart* GiW
L, $(V = L^T L)$	horná trojholníkova matica, parameter určujúci hustotu GiW
ν	parmameter určujúci hustotu GiW
λ	parameter zabúdania
α	normalizačná konšntanta
Pr_t	cena futures v čase t
x_t	stav obchodníka
K_t	kapitál v čase t
Yd_t	výnos v čase t

Značenie ktoré má v hornej časti značku "$\sim$"značí vektor premenných bez "$\sim$", často aj s indexom dĺžky ($\widetilde{Z}_n = (Z_1, \ldots, Z_n)$).
Značenie ktoré má v hornej časti značku "."značí aproximáciu premenneh bez ".".

Úvod

Práca sa zaoberá odhadovaním závislostí diskrétnych veličín na veličinách ktoré môžu mať a spojité rozdelenie. To má veľké praktické využitie napríklad v medicíne pri výskyte chorôb ale aj mnohých ďalších. Konkrétne sa bude zaoberať odhadovaním modelu logistickej regresie, popisujúcim závislosť na vysvetľujúcich premenných ktoré môžu byť diskrétne alebo spojité. Vo všeobecnosti sa na riešenie tohto problému, odhadovania parametrov, využíva numerický postup, ktorý pri zvyšovaní sa počtu dát býva časovo príliš náročný alebo asymptotický výpočet na ktorého presnosť je spravidla potreba až veľký počet dát. Táto práca sa zaoberá metodikou, ktorá by sa dala používať priebežne s prichádzajúcimi datami. Na to sa dá vhodne použiť bayesovské odhadovanie parametrov popísané v prvej kapitole a algoritmus navrhnutý v [3], na ktorý táto práca priamo nadväzuje, popísaný v druhej kapitole. V tretej kapitole je aplikovaný na logistickú regresiu a následne testovaný na simulovaných datach a tiež na obchodovanie s futures. Na to bude využitý software matlab s pridaním balíkom Mixtools [9]. Obchodovaniu s futures sa zaoberali mnohé práce ako napr. [6]. Postup v tejto práci sa ale bude líšiť v spôsobe rozhodovania či v danom čase obchodovať. Uvažovanie o obchode sa bude rozlišovať ako diskrétna náhodná veličina, ktorej stavy sú len trojica predať, kúpiť a neobchodovať, v závislosti na predchádzajúcom vývoji cien.

Kapitola 1

Bayesovské odhadovanie

Základom bayesovského odhadovania je pristupovanie ku odhadovaného parametru θ ako ku náhodnej veličine závislej na pozorovaných datach, náhodnom výbere $\widetilde{X}_n = (X_1, \ldots, X_n)$ z nejakého rozdelenia s hustotou $f(x)$ závislou na vektorovom parametry θ. Základným nástrojom odhadu je Bayesova veta, vzorec, ktorý vyjadruje podmienenú pravdepodobnosť javu θ_0 na vektore $\widetilde{X}_n$, ktorá vychádza z predpokladu, že máme nejakú počiatočnú informáciu o odhadovanom parametre θ, nezávislú na výbere $\widetilde{X}_n$. Teda predpokladáme, že parameter má rozdelenie, o ktorom máme predstavu na základe minulosti alebo z iných dôvodov [2]. V tejto kapitole popíšeme základné princípi bayesovského odhadovania a pridáme predpoklad podmienenej nezávislosti, ktorý nám umožní prístup k dynamickému odhadovaniu náhodnej veličiny.

1.1 Bayesov vzorec

Veta 1 (Bayesova veta)**.** *Nech $\theta \in \Theta$ je náhodný vektor s hustotou $g_\theta(\theta)$ vzhľadom ku σ - konečnej miere λ na $(\Theta,\mathcal{B}(\Theta))$ kde $\Theta \subset \mathbb{R}^k$ neprázdna borelovská množina a $\mathcal{B}(\Theta)$ je množina borelovských podmnožín Θ. Nech $\widetilde{X}_n = (X_1, \cdots, X_n)$ je náhodný vektor s podmienenou hustotou $f_{X|\theta}(\widetilde{X}_n|\theta)$ vzhľadom ku σ-konečnej miere ν_n na $(\mathbb{R}^n,\mathcal{B}^n)$ kde $\mathcal{B}^n$ je množina borelovských podmnožín $\mathbb{R}^n$. Potom pre hustotu vektoru θ podmienenú vektorom $\widetilde{X}_n$ platí*

$$g_{\theta|X}(\theta_0|\widetilde{X}_n) = \begin{cases} \frac{g_\theta(\theta_0) f_{X|\theta}(\widetilde{X}_n|\theta_0)}{\int_\Theta g_\theta(\theta) f_{X|\theta}(\widetilde{X}_n|\theta) d\lambda(\theta)} & ak \int_\Theta g_\theta(\theta) f_{X|\theta}(\widetilde{X}_n|\theta) d\lambda(\theta) \neq 0 \\ 0 & inak. \end{cases} \tag{1.1}$$

Dôkaz. Dôkaz tejto vety sa dá nájsť v [1] strana 57. ❑

Je vidno, že v menovateli sa nevyskytuje informácia o vektore θ. Je to len normovacia konštanta aby $g_{\theta|X}(\theta|\widetilde{X}_n)$ bola hustota. Preto sa zvykne používať skrátený zápis

$$g_{\theta|X}(\theta|\widetilde{X}_n) \propto g_\theta(\theta) \times f_{X|\theta}(\widetilde{X}_n|\theta).$$

$g_\theta(\theta)$ sa nazýva *apriórna* hustota náhodného vektoru θ, volí sa pred samotným vyhodnocovaním závislej hustoty. Je vnej zahrnutá subjektívna úvaha, skúsenosť z minulosti o rozdelení náhodného vektoru θ a nieje nijak závislá od náhodného vektoru $\widetilde{X}_n$.Podmienená hustota $g_{\theta|X}(\theta|\widetilde{X}_n)$ sa nazýva *aposteriórna* hustota a definuje podmienené rozdelenie parametru θ za podmienky $\widetilde{X}_n$. $f_{X|\theta}(\widetilde{X}_n|\theta)$ je hustota náhodného vektoru $\widetilde{X}_n$ s parametrom θ.

Poznámka 1. V texte práce budeme uvažovať mieru ν_n ako n-rozmernú Lebesgueovu mieru a λ ako k-rozmernú Lebesgueova mieru, prípadne sčítacie miery. Preto budeme miesto symbolu $d\lambda(\theta)$ písať len symbol $d\theta$ a pri použítí sčítacej miery nahradíme integrál sumou.

Princíp bayesovského odhadovania sa nám hodí ku priebežnému využívaniu nových informácií získaných z nových odpozorovaných dat alebo iných výsledkov. Prínos novej informácie však môže čiastočne alebo úplne zmeniť tvar aposteriórnej hustoty. Práca s aposteriórnymi hustotami môže mať veľkú časovú náročnosť a analytické riešenie nemusí vôbec existovať.

Definícia 1. *Systém hustôt $\mathcal{F} = \{g(\theta),\theta \in \Theta\}$ pre ktorý platí, že ak apriórna hustota $g_\theta \in \mathcal{F}$ a pre dostatočne veľké n aposteriórna hustota $g_{\theta|X}(\theta|\widetilde{X}_n)$ patria do $\mathcal{F}$ tak sa tento systém nazýva systém konjugovaný s hustotami $\{f(\widetilde{x}|\theta),\theta \in \Theta\}$. Kde $\widetilde{X}_n = (X_1,\dots,X_n)$ je náhodný výber z rozdelenia s hustotou $f(x|\theta)$.*

Ak konjugovaný systém hustôt obsahuje hustoty, s ktorými sa dá analyticky pracovať tak nemusí nastať problém zložitosti aposteriórnej hustoty.

1.2 Predpoklad podmienej nezávislosti

Nech je náhodný vektor $\widetilde{Z}_n = (Z_1,\dots,Z_n)$ kde $Z_i = (X_i,Y_i)$. X je vysvetľujúca premenná a Y je vysvetľovaná premenná. Predpokladajme podmienenú nezávislosť vektoru θ s veličinou X_{n+1} za podmienky $\widetilde{Z}_n$.

$$g_{\theta|Z,X}(\theta|\widetilde{Z}_n,X_{(n+1)}) = g_{\theta|Z}(\theta|\widetilde{Z}_n) \tag{1.2}$$

Poznámka 2. Vďalšom budeme $g_{\theta|Z}(\theta|\widetilde{Z}_n)$ značiť skráteným značením $g_n(\theta)$ a $g_\theta(\theta)$ budeme značiť $g_0(\theta)$.

Z predpokladu 1.2 dostávame, že posteriórna hustota má tvar

$$\begin{aligned} g_n(\theta_0) &= \frac{f(Y_n|\widetilde{Z}_{n-1},X_n,\theta_0)g_{n-1}(\theta_0|,X_n)}{\int_\Theta f(Y_n|\widetilde{Z}_{n-1},X_n,\theta)g_{n-1}(\theta|X_n)\,d\theta} \\ &= \frac{f(Y_n|\widetilde{Z}_{n-1},X_n,\theta_0)g_{n-1}(\theta_0)}{\int_\Theta f(Y_n|\widetilde{Z}_{n-1},X_n,\theta)g_{n-1}(\theta)\,d\theta} \end{aligned}$$

teda

$$g_n(\theta) \propto f(Y_n|\widetilde{Z}_{n-1},X_n,\theta) \times g_{n-1}(\theta_0|,X_n). \tag{1.3}$$

Dostali sme rekurentný vzťah. Postupným dosadením dostaneme

$$\begin{aligned} g_n(\theta_0) &= \frac{\prod_{i=1}^n f(Y_i|\widetilde{Z}_{i-1},X_i,\theta)g_0(\theta)}{\int_\Theta \prod_{i=1}^n f(Y_i|\widetilde{Z}_{i-1},X_i,\theta)g_0(\theta)\,d\theta} \\ g_n(\theta) &\propto L(\widetilde{Y}_n|\widetilde{Z}_{n-1},X_n,\theta) \times g_0(\theta) \end{aligned} \tag{1.4}$$

kde výraz v (1.4)

$$L(\widetilde{Y}_n|\widetilde{Z}_{n-1},X_n,\theta) = \prod_{i=1}^{n} f(Y_i|\widetilde{Z}_{i-1},X_i,\theta) \tag{1.5}$$

je vierohodnostná funkcia parametru θ [1].

Je vidieť, že za platnosti prepdokladu (1.2) sa dá pristupovať k náhodnej veličine θ dinamicky, teda pri získaní novej informácie Z_{i+1} je možné využiť už spočítanú hustotu závislú len na $\widetilde{Z}_i$.

V štatistike ide o prirodzený predpoklad, ktorý vlastne hovorí, že namerané vysvetlovacie premenné X_i nám bez vysvetľovanej premennej Y_i nedajú žiadne informácie o parametry θ.

Kapitola 2

Aproximácia Bayesovho odhadu

Ak aposteriórna hustota nemá vlastnosti, ktoré by boli vhodné pre priame použitie ako napríklad to, že sa nedá počítať analyticky, tak ju budeme aproximovať inou pravdepodobnostnou hustotou, ktorá by nespôsobovala tieto problémy. Budeme pritom využívať algoritmus navrhnutý a presne popísaný v [3].

2.1 Princíp algoritmu

Základom algoritmu je pre aposteriórnu pravdepodobnostnú hustotu $g(\theta)$ nájsť parametrickú pravdepodobnostnú hustotu $p(\theta) \in G$, ktorá by minimalizovala tzv . „Kullback-Leiblerovu divergenciu“(KLD) definovanú ako

$$D(g||p) = \int_{\Theta} g(\theta) \log \frac{g(\theta)}{p(\theta)} d\theta. \tag{2.1}$$

Rodinu pravdepodobnostných hustôt G pritom volíme podľa potrieb na prácu s výslednou hustotou. Teda

$$\dot{p} \in \arg\min\{D(g||p)|p \in G\} \tag{2.2}$$

Poznámka 3. Hodnota $D\left(g\|p\right)$ definovaná v 2.1 vyjadruje istú „podobnosť“ husoty p ku hustote g. Jedná sa o špeciálny prípad tzv. *f-divergencií*S bližšie popísaných v [5]. Vhodnosť použitia KLD je popísaná v [4].

Definícia 2. *Hustotu $\dot{p}$ definovanú v (2.2) budeme nazývať projekciou pravdepodobnostnej hustoty g na rodinu G.*

Poznámka 4. $\dot{p}$ aj ďalšie hustoty v texte budú mať špeciálny tvar. Preto sa za parameter θ budeme považovať dvojicu $\theta = (\phi,\sigma)$ a parametrický priestor $\Theta = \Phi \times (0,\infty)$ kde $\Phi \subset \mathbb{R}^{k-1}$

V samotnom algoritme je ako rodina G uvažovaná rodina „Gauss-inverz-Wishartových“ pravdepodobnostných hustôt. Túto rodinu budeme značiť GiW.

$$f_{GiW}(\phi,\sigma|L,\nu) = \alpha_{L,v}\sigma^{-\nu}e^{-\frac{[1,-\phi]V[1,-\phi]^T}{2\sigma^2}} \tag{2.3}$$

kde pravý horný index T označuje transponovanie vektoru alebo matice, V je pozitívne definitná matica, $V = L^T L$, L je dolná trojuholníková matica s kladnou diagonálou a $\alpha_{L,v}$ je konštanta zaisťujúca jednotkový integrál.

Poznámka 5. L je Choleskyho odmocnina, teda rozklad $V = L^T L$ je jednoznačný,

2.2 Aproximačný algoritmus

Nech v čase $n+1$ nameráme hodnoty náhodného vektoru $Z_{n+1} = (X_{n+1},Y_{n+1})$, kde Y_{n+1} má za podmienky ($\widetilde{Z}_n$, X_{n+1}, θ) hustotu $f(Y_{n+1}|\widetilde{Z}_n,X_{n+1},\theta)$(v ďalšom budeme značiť f_{n+1}). Postupujeme nasledovne:

1. **krok**

$$\dot{g}_{n+1} \propto f_{n+1} \times p_n.$$

2. **krok**

$$\dot{p}_{n+1} = arg \min \{D(\dot{g}_{n+1} \| p) | p \in G\}$$

3. **krok** Zadefinujeme

$$\begin{aligned} p_\lambda &\propto \dot{p}_{n+1}^{\lambda} p_n^{1-\lambda}, \quad \lambda \in (0,1) \\ R_{n+1} &= \frac{\dot{p}_{n+1}}{p_n} \end{aligned}$$

Nájsť λ_{n+1} aby splňovalo

$$\int \dot{g}_{n+1} \log R_{n+1} = \int p_{\lambda_{n+1}} \log R_{n+1}$$

ak existuje, inak

$$\lambda_{n+1} = \begin{cases} 0 & \text{ak } \int \dot{g}_{n+1} \log R_{n+1} \leq 0 \\ 1 & \text{inak} \end{cases}$$

4. **krok**

$$p_{n+1} = p_{\lambda_{n+1}}$$

Problém algoritmu, ktorý nastáva v kroku 1 je, že do bayesovho vzorca nevkladáme pôvodnú vierohodnostnú funkciu (1.5) ale len novú informačnú hustotu a ako apriórnu husotu používame odhadnutú hustotu p_{n+1} a tak pri každom prechode z n do $n+1$ vzniká chyba z novej projekcie v kroku 2. Preto je nutné istú informáciu o parametre θ potlačiť. To sa deje v kroku 3.

2.3 Špecializácia algoritmu pre prípad aproximujúcich GiW hustôt

Uvažujme regresný model, ktorý popisuje závislosť náhodnej veličiny Y_{n+1} na vektore X_{n+1} za podmienky $\theta = (\beta,\phi)$, kde θ je vektor regresných koeficientov a σ^2 určuje rozptyl.

Poznámka 6. Ak nebude definované inak tak v ďalšom budeme matice uvažovať ako

$$\mathbf{M_n} = \begin{pmatrix} M_{n,y} & M_{n,yx} \\ M_{n,yx}^T & M_{n,x} \end{pmatrix}$$

kde $M_{n,y}$ je skalár a $M_{n,yx}$ je vektor o rozmere zhodnom s rozmerom vektora X.

Uvažujme, že máme dosiaľ pravdepodobnostnú hustotu $p_n \in GiW$. Získame novú informáciu $Z_{n+1} = (X_{n+1},Y_{n+1})$.

Krok 1: Podľa Bayesovho vzorca vypočítame posteriórnu hustotu $\dot{g}_{n+1}(\theta)$ z p_n, aproximovanej hustoty parametru θ

$$\dot{g}_{n+1}(\theta) \propto f(Y_{n+1}|\widetilde{Z}_n,X_{n+1},\theta) \times p_n(\theta|X_{n+1}).$$

Krok 2: Hľadanie projekcie na rodinu GiW je ekvivalentá úloha s hľadaním parametrov L_{n+1} a ν_{n+1} ktoré presne definujú hustotu 2.3. Preto $\dot{p}_{n+1} = p_{L_{n+1},\nu_{n+1}}$.

$$p_{L_{n+1},\nu_{n+1}} \in \arg\min\{D(\dot{g}_{n+1}||p)|p \in GiW\}$$

Veta 2 (najlepšia projekcia na GiW)**.** *Pre pravdepodobnostnú funkciu $p_{L_{n+1},\nu_{n+1}}$, ktorá minimalizuje $\{D(\dot{g}_{n+1}||p)|p \in GiW\}$ platí*

$$\begin{aligned}
L_{n+1,x} &= (U_{n+1,x}^{-1})^T \\
L_{n+1,yx} &= -U_{n+1,y}^{-1}(U_{n+1,x}^{-1})^T U_{n+1,yx} \\
L_{n+1,y} &= \frac{\sqrt{\nu_{n+1}}}{U_{n+1,y}} \\
\log(\frac{\nu_{n+1}}{2}) - \Psi(\frac{\nu_{n+1}}{2}) &= a_{n+1} + \log(B_{t,y}) \qquad (2.4)
\end{aligned}$$

kde

$$\begin{aligned}
a_{n+1} &= \alpha_{n+1} \iint_{\Phi\times(0,\infty)} \log(\sigma^2)\sigma f_{n+1}(\phi,\sigma)p_n(\phi,\sigma)d\phi d\sigma^2 \qquad (2.5) \\
B_{n+1} &= \alpha_{n+1} \iint_{\Phi\times(0,\infty)} \frac{[1,-\phi]^T[1,-\phi]}{\sigma} f_{n+1}(\phi,\sigma)p_n(\phi,\sigma)d\phi d\sigma^2 \qquad (2.6) \\
&= \begin{pmatrix} B_{n+1,y} & B_{n+1,yx} \\ B_{n+1,yx}^T & B_{n+1,x} \end{pmatrix} \\
&= \begin{pmatrix} U_{n+1,y} & 0 \\ U_{n+1,yx}^T & U_{n+1,x} \end{pmatrix} \begin{pmatrix} U_{n+1,y} & U_{n+1,yx} \\ 0 & U_{n+1,x} \end{pmatrix} \\
\alpha_{n+1} &= \iint_{\Phi\times(0,\infty)} \sigma f_{n+1}(\phi,\sigma)p_n(\phi,\sigma)d\phi d\sigma^2 \\
\Psi(x) &= \frac{\partial\Gamma(x)}{\partial x} = \frac{\partial}{\partial x}\int_0^\infty y^{x-1}e^{-y}dy
\end{aligned}$$

a riešenie rovnice (2.4) vždy existuje.

Dôkaz. odvodenie a dôkaz vety v [3] veta 7 a 8.

❑

Krok 3: Zadefinujeme p_λ a R_{n+1} predpisom

$$\begin{aligned}
p_\lambda &\propto \dot{p}_{n+1}^\lambda p_n^{1-\lambda}, \quad pre \quad \lambda \in [0,1] \\
R_{n+1} &= \frac{\dot{p}_{n+1}}{p_n}.
\end{aligned}$$

Poznámka 7. p_λ je pravdepodobnostná hustota, ktorá určitú informáciu z novej pravdepodobnostnej hustoty vynechá. V ďalšom dokážeme, že existuje $\lambda_0 \in [0,1]$,že p_{λ_0} aproximuje pravdepodobnostnú hustotu g lepšie ako $\dot{p}_{n+1}$. Funkciu $p_{\lambda_{n+1}}$ nazveme zabúdacou funkciou.

Veta 3. *Nech $q(\theta)$ je pravdepodobnostná hustota náhodnej veličiny θ, pre ktorú platí, že $D(q||\dot{p}_{n+1}) < \infty$, $D(q||p_n) < \infty$ 2.1. Potom existuje $\lambda \in [0,1]$ taká, že $D(q||p_\lambda) \leq D(q||\dot{p}_{n+1})$.*

Dôkaz. Postupujeme ako v [3] a spočítame prvú a druhú deriváciu funkcie $h(\lambda) := D(q||p_\lambda)$, definujeme $J(\lambda) := \int_\Theta \dot{p}_{n+1}^\lambda(\theta) p_n^{1-\lambda}(\theta) d\theta$

$$\begin{aligned}
h(\lambda) &= \int_\Theta q(\theta) \log(\frac{q(\theta)}{p_\lambda(\theta)}) d\theta \\
&= \int_\Theta q(\theta) \log(q(\theta)) d\theta - \int_\Theta q(\theta) \log(p_\lambda(\theta)) d\theta \\
&= \int_\Theta q(\theta) \log(q(\theta)) d\theta - \lambda \int_\Theta q(\theta) \log(\dot{p}_{n+1}(\theta)) - \\
&\quad -(1-\lambda) \int_\Theta q(\theta) \log(p_n(\theta)) d\theta + \log(J(\lambda)) \\
h'(\lambda) &= \frac{J'(\lambda)}{J(\lambda)} - \int_\Theta q(\theta) log(R_{n+1}(\theta)) d\theta \\
&= \int_\Theta p_\lambda(\theta) \log(R_{n+1}(\theta)) d\theta - \int_\Theta q(\theta) \log(R_{n+1}(\theta)) d\theta \\
h''(\lambda) &= \int_\Theta p_\lambda(\theta) \log^2(R_{n+1}(\theta)) d\theta - (\int_\Theta p_\lambda(\theta) \log(R_{n+1}(\theta)) d\theta)^2 \\
&= var[\log(R_{n+1}(\theta))] \qquad \textit{(vzhľadom ku pravdepodobnostnej hustote } p_\lambda\textit{)} \\
&\geq 0
\end{aligned}$$

kde symbolom θ v poslednom výraze je myslená náhodná veličina s hustotou p_λ Je vidieť, že pre $\lambda = 1$ je znenie vety triviálne. Ak ale exituje λ tak aby bolo splnené

$$\int_\Theta q(\theta) \log(R_{n+1}) d\theta = \int_\Theta p_\lambda(\theta) \log(R_{n+1}(\theta)) d\theta. \tag{2.7}$$

dostaneme ostré minimum funkcie (2.1). Tým sme dokázali tvrdenie vety.

❑

Poznámka 8. Veta 3 je obecná a neplatí len pre zobrazovanie na rodinu GiW. Je tu uvedená aby bolo vidno, že priebežne nezachovávame neaproximovanú hustotu g_{n+1}. Pre úplnosť algoritmu položíme miesto g_{n+1} najlepšie aproximujúcu hutotu $\dot{g}_{n+1}$. Jedná sa o heuristicky odvodený krok!

Veta 4. *Nech*

$$\begin{aligned}\Delta\nu &= \nu_{n-1} - \nu_n\\ \Delta V &= V_{n-1} - V_n.\end{aligned}$$

$\lambda_{n+1} \in [0,1]$, *ktorá minimalizuje* $D(\dot{g}_{n+1}\|p_{\lambda_{n+1}})$ *je riešením rovnice*

$$\begin{aligned}\Delta\nu a_{n+1} + tr(U_{n+1}^T U_{n+1}\Delta V) &= \Delta\nu(\log(\frac{L_{\lambda_{n+1}}^2}{2}y) - \Psi(\frac{\nu_{\lambda_{n+1}}}{2}))\\ &\quad + tr(L_{\lambda_{n+1}}^{-1}(L_{\lambda_{n+1}}^{-1})^T\Delta V)\end{aligned}$$

Ak riešenie rovnice existuje, inak

$$\lambda_{n+1} = \begin{cases} 0 & ak\ \Delta\nu a_{n+1} + tr(U_{n+1}^T U_{n+1}\Delta V) + 2\log\left(\frac{J(L_n,\nu_n)}{J(L_{n+1},\nu_{n+1})}\right) \leq 0\\ 1 & inak\end{cases}. \quad (2.8)$$

kde $L_{\lambda_{n+1}}$ *je dolná trojholníková matica, pre ktorú platí*

$$\begin{aligned}L_{\lambda_{n+1}}^T L_{\lambda_{n+1}} &= \lambda_{n+1}\left[L_{n+1}^T, \sqrt{\frac{1}{\lambda_{n+1}} - 1}L_n^T\right]\left[L_{n+1}^T, \sqrt{\frac{1}{\lambda_{n+1}} - 1}L_n^T\right]^T\\ \nu_{\lambda_{n+1}} &= \lambda_{n+1}\nu_{n+1} + (1-\lambda_{n+1})\nu_n\end{aligned}$$

a $J(L,\nu)$ *je normalizačná konštanta z hustoty* $p_{L,\nu}$ *definovanej v (2.3)* $a_{n+1}, \Psi(x), L, U$ *sú definované v 2*

Kôli zúplneniu z poznámky 8 nemusí λ, ktoré rieši vzťah (2.8) vo vete 4 odpovedať vhodnému parametru, ktorý by riešilo vzťah (2.7) pre hľadanie najlepšej projekcie hustoty g_{n+1}. Konvergenciu by to mohlo spomaliť a v prípade nevhodných dat, kedy by vychádzalo λ blízsko 0 aj úplne narušiť. Preto sa vo výpočtoch budeme pokúšať parameter λ aj obmedziť z dola.

Krok 4: V poslednom kroku položíme $p_{n+1} := p_{\lambda_{n+1}}$.

Kapitola 3

Logistická regresia

Logistická regresia sa snaží nájsť závislosť, dichomatickej náhodnej veličiny Y na pozorovaných datach $X = (X_{(1)}, \ldots, X_{(k)})$ [7]. Jej obecnejšou verziou je multinomická logistická regresia, kde Y môže nadobúdať konečne mnoho hodnôt.

V tejto práci sa budeme zaoberať špeciálnym tvarom multinomickej logistickej regesie a odhadovaním jej parametrov. Na to sa vo všeobecnosti dajú používať viaceré metódy. Vtejto kapitole zhrnieme nedostatky ich použitia a následne sa pokusíme implementovať algoritmus popísaný v kapitole 2.

Poznámka 9. Jednoduchá logistická regresia je špeciálnym prípadom multinomickej logistickej regresie. Teda závery práce sú priamo aplikovateľné na jednoduchý model.

3.1 Model

Všeobecný model multinomickej logistickej regresie vyzerá ako odhadovanie logaritmu pomeru šancí medzi jednotlivými stavmy náhodnej veličiny Y a pivotným stavom v závislosti na vektore X.tj.

$$\log\left(\frac{P[Y = j|\phi, X]}{P[Y = piv|\phi, X]}\right) = \phi \mathbf{X}^T$$

Budeme pracovať s upraveným modelom, kde do exponentu pridáme druhú mocninu. Zadefinujeme pravdepodobnosti jednotlivých hodnôt, ktoré môže nadobúdať Y v závislosti na X a ϕ predpisom

$$\begin{aligned} P[Y = j|X,\phi,\sigma] &= \frac{e^{-\frac{(j-(\phi\mathbf{X}^T))^2}{\sigma^2}}}{\sum_{i=-I}^{I} e^{-\frac{(i-(\phi\mathbf{X}^T))^2}{\sigma^2}}} \quad j = -I, \ldots, I \qquad (3.1) \\ kde \quad \phi &= (\phi_0, \ldots, \phi_n) \\ a\ \mathbf{X} &= (1, X). \end{aligned}$$

Tento nie úplne štandartný model bol zvolený pre podobnosť čitateľa s čitateľmi pravdepodobnostných hustôt z rodiny GiW (2.3). Zjednoduší to hľadanie projekcie na rodinu GiW.

Poznámka 10. Podmienenú pravdepodobnosť $P[Y = j|X,\phi,\sigma]$ budeme značiť skráteným značením $P[Y = j]$

Existujú viaceré spôsoby odhadu parametrou ϕ a σ. Bežný prístup odhadu býva použitím *Maximálnej vierohodnosti* [1] strana 146.

3.2 Odhad parametrov ϕ, σ metódou maximálnej vierohodnosti

Majme model (3.1) a náhodný výber $\widetilde{Z} = (Z_1, \ldots, Z_n)\,, \quad Z_i = (\mathbf{X}_i, Y_i)$ veľkosti n.

Bez ujmy na obecnosti môžeme zoradiť veličiny nadobudnuté v jednotlivých datach do $2I+1$ blokov kde v l-tom bloku budú dáta Z_i pre $i = m_{l-1}, \ldots, m_l$, kde $Y_i = l$, pre $l = -I-1, \ldots, I$ a platí, že $\{m\}_{l=-I-1}^{I}$ je neklesajúca postupnosť a $m_{-I-1} = 0, \quad m_I = n$. Potom vierohodnostná funkcia vyzerá

$$L_m(\phi, \sigma) = \prod_{i=-I}^{I} \prod_{l=m_{i-1}}^{m_i} \frac{e^{-\frac{(i-\phi\mathbf{X}_l^T)^2}{\sigma^2}}}{\sum_{j=-I}^{I} e^{-\frac{(j-\phi\mathbf{X}_l^T)^2}{\sigma^2}}}, \tag{3.2}$$

ktorá v logaritmickom tvare vyzerá

$$l_m(\phi, \sigma) = -\left(\sum_{i=-I}^{I} \sum_{j=m_{i-1}}^{m_i} \frac{(i-\phi\mathbf{X}_j^T)^2}{\sigma^2} + \sum_{i=-I}^{I} \sum_{j=m_{i-1}}^{m_i} \left(\log \sum_{l=-I}^{I} e^{\frac{-(l-\phi\mathbf{X}_j^T)^2}{\sigma^2}} \right) \right). \tag{3.3}$$

Hľadanie parametru ϕ, ktorý maximalizuje (3.2) a (3.3) sa ale nedá riešiť analyticky. Najbežnejšia metóda hľadania koreňa je „Newton-Raphsonova“ iteračná metóda popísaná v [8] Pre model (3.1) je postup ekvivalentný.

Táto metóda aj ďalšie numerické metódy majú však spoločný problém a to je, že pri veľkom počte dát sa stávaju nepoužitelné kôli časovej náročnosti numerických výpočtov. Ďalšou nevýhodou použitia tejto metódy je aj fakt, že odhad je bodový a nemáme bližšiu predstavu o presnosti parametrov.

3.3 Problémy odhadovania Bayesovou metodou

Odhadujeme závislosť Vierohodnostná funkcia parametru $\theta = (\phi, \sigma)$ má tvar (3.2). Po dosadení do vzorca (1.4) s apriórnou hustotou $g_0(\theta)$ dostaneme vzťah pre aposteriórnu hustotu

$$g_n(\phi_0, \sigma_0) = \frac{\prod_{i=-I-1}^{I} \prod_{l=m_{i-1}}^{m_i} \frac{e^{-\frac{(i-\phi\mathbf{X}_l^T)^2}{\sigma^2}}}{\sum_{j=-I-1}^{I} e^{-\frac{(j-\phi\mathbf{X}_l^T)^2}{\sigma^2}}} g_0(\theta)}{\int\int_{\phi\times(0,\infty)} \prod_{i=-I-1}^{I} \prod_{l=m_{i-1}}^{m_i} \frac{e^{-\frac{(i-\phi\mathbf{X}_l^T)^2}{\sigma^2}}}{\sum_{j=-I-1}^{I} e^{-\frac{(j-\phi\mathbf{X}_l^T)^2}{\sigma^2}}} g_0(\theta) d\phi d\sigma}.$$

Je vidieť, že aposteriórna hustota g_n(3.2) nieje veľmi vhodná na prácu a počítanie pravdepodobnosti hodnôt parametrov. Preto sa v praxi bayesovské odhadovanie na odhadovanie parametrov v logistickej regresii veľmi nepoužíva. My sa budeme snažiť posteriórnu hustotu aproximovať inou zvolenou hustotou, ktorej vlastnosti by boli „priaznivé“. Využijeme pritom algoritmu z kapitoly 2

3.4 Aproximácia bayesovho odhadu v logistickej regresii

Po zvolení príslušného apriórneho rozdelenia z rodiny GiW používame algoritmus z kapitoly 2. V kroku 3 vyjadríme premenné a (2.5) a B (2.6) z pravdepodobnostnej hustoty závislej na parametroch ϕ, σ

$$f_{n+1}(Y_{n+1}|X_{n+1},\phi,\sigma) = \frac{e^{\frac{-(Y_{n+1}-\phi X_{n+1}^T)^2}{\sigma^2}}}{\sum_{l=-I}^{I} e^{\frac{-(Y_{n+1}-\phi X_{n+1}^T)^2}{\sigma^2}}} \quad kde \quad Y_{n+1} \in \{-I, \ldots, I\} \tag{3.4}$$

Veta 5. *Parametre a, B definované (2.5), (2.6) majú pre model 3.1 v $n+1$ kroku tvar*

$$a_{n+1} = \alpha_{n+1} \iint_{\Phi\times(0,\infty)} \log(\sigma^2)\sigma \frac{p_{GiW}(\phi,\sigma|Y_{n+1},X_{n+1})}{\sum_{l=-k}^{k} e^{\frac{-(l-\phi X^T)^2}{\sigma^2}}} d\phi d\sigma^2 \tag{3.5}$$

$$B_{n+1} = \alpha_{n+1} \iint_{\Phi\times(0,\infty)} \frac{[1,-\phi]^T[1,-\phi]}{\sigma^2} \frac{p_{GiW}(\phi,\sigma|Y_{n+1},X_{n+1})}{\sum_{l=-k}^{k} e^{\frac{-(l-\phi X^T)^2}{\sigma^2}}} d\phi d\sigma^2 \tag{3.6}$$

kde

$$p_{GiW}(\phi,\sigma|Y_{n+1},X_{n+1}) = p_n(\phi,\sigma) e^{\frac{-(Y_{n+1}-\phi X_{n+1}^T)^2}{\sigma^2}} \tag{3.7}$$

a p_{GiW} patri do rodiny GiW.

Dôkaz. uvažujme hustotu p_n z rodiny GiW

$$\begin{aligned} p_n(\phi,\sigma) &= p_{L,\nu}(\phi,\sigma) \\ &= \alpha_{L,v}\sigma^{-\nu} e^{-\frac{[1,-\phi]V[1,-\phi]^T}{2\sigma^2}} \end{aligned} \tag{3.8}$$

pre nejaké ν, L. Exponenciálu v 3.7 prepíšeme do vhodného tvaru

$$\begin{aligned} e^{\frac{-(Y_{n+1}-\phi X_{n+1}^T)^2}{\sigma^2}} &= e^{\frac{-[1,-\phi]\psi^T\psi[1,-\phi]}{\sigma^2}} \\ \psi &= (Y_{n+1},X_{n+1}) \end{aligned} \tag{3.9}$$

Následne po vynásobení (3.8) s (3.9) dostaneme novú hustotu

$$p_{GiW}(\phi,\sigma|Y_{n+1},X_{n+1}) = \alpha^*_{L,v}\sigma^{-\nu} e^{-\frac{[1,-\phi](V+\frac{\psi^T\psi}{2})[1,-\phi]^T}{2\sigma^2}}$$

ktorá patrí do rodiny GiW.

Dosadením (3.4) do (2.5) dostaneme

$$\begin{aligned} a_{n+1} &= \alpha_{n+1} \iint_{\phi\times(0,\infty)} \log(\sigma^2)\sigma f_{n+1}(Y_{n+1}|X_{n+1},\phi,r) p_n(\phi,\sigma) d\phi d\sigma^2 \\ &= \alpha_{n+1} \iint_{\Phi\times(0,\infty)} \log(\sigma^2)\sigma p_n(\phi,\sigma) \frac{e^{\frac{-(Y_{n+1}-\phi X_{n+1}^T)^2}{\sigma^2}}}{\sum_{l=-k}^{k} e^{\frac{-(Y_{n+1}-\phi X_{n+1}^T)^2}{\sigma^2}}} d\phi d\sigma^2 \end{aligned}$$

Z čoho okamžite plynie tvrdenie vety pre parameter a_{n+1}. Tvrdenie pre parameter B_{n+1} sa dokáže obdobne.

❑

Kapitola 4

Futures

Futures je špecialna forma forwardov, finančný derivát, zmluva o budúcej kúpe alebo predaji určitého podkladového aktíva za dohodnutú cenu. Dve strany sa dohodnú, že zobchodujú medzi sebou určitý balík podkladového aktiva za určitú cenu, pričom predávajúci je v tzv. *krátkej pozícii* a kupujúci je v tzv. *dlhej pozícii.* Futures sú obchodované na burze a teda môžu byť predmetom záujmu tretej strany, špekulatnov, ktorý sa snažia maximalizovať svoj zisk na základe predikovania finančného derivátu. Pre zvyšovanie výnosu je dobre obchodovať s viacerými typmi podkladového aktíva. V tejto práci sa budeme ale pozerať len na jedno podkladové aktívum a vývoj jeho ceny. Implementovanie modelu vtejto práci sa bude pozerať na zmenu ceny finančného derivátu ako na diskrétnu náhodnú veličinu ktorej stavy udávajú či sa cena zvýšila alebo znížila o poplatok za uzatvorenie obchodu. Paramatre vstupujúce do tohoto modelu sa môžu časom meniť, pretože faktorov, ktoré ovplivňujú ceny je v skutočnosti oveľa viac než je možné namodelovať.

4.1 Obchodovanie s futures

Majme časové okamihy $\widetilde{T} = \{1, \ldots, T\}, T < \infty$ a pre každé $t \in \widetilde{T}$ máme odpozorované hodnoty $\widetilde{Pr}_t = \{Pr_1, \ldots, Pr_t\}$, kde Pr_i je cena v čase i.

Obchodník sa v každom čase t môže na základe vývoja ceny ktorý už nastal $\widetilde{Pr}_t$, rozhodnúť koľko akcií nakúpi alebo predá u_t, jeho stav v čase t je x_t, pričom ho obmedzuje rozhranie množstva ktoré môže držať $x_t \in [\underline{x}, \overline{x}]$. Akcie obchodníka sa uskutočnia v časoch $1 \leq t_{k-1} < t_k \leq T$ kde $k = 1, \ldots, k^*$.

Poznámka 11. Platí, že $|u_t| > 0$ ak $t \in \{t_k | k = 1, \ldots, k^*\}$ a $u_t = 0$ inak

Poznámka 12. Je zrejmé, že

$$x_t = x_{t-1} + u_t \tag{4.1}$$

Definícia 3. *Definujeme kapitál v čase t ako*

$$K_t = K_{t-1} - Pr_t u_t - c|u_t| \tag{4.2}$$

kde $c \geq 0$ sú náklady na obchod.

Definícia 4. *Definujeme výnosovú funkciu v čase $t \in \widetilde{T}$ predpisom*

$$Yd_t = K_t - K_0 + y_t x_t - y_0 x_0 \tag{4.3}$$

Cieľom obchodovania je maximalizovať strednú hodnotu výnosovej funkcie v čase T.

Veta 6 (Bang-Bang). *Maximum výnosu v čase T je daný maximom výnosu v čase jednotlivých akcií, pre ktoré platí, že*

$$\begin{aligned} x_{t-1}+u_t &\in \{\underline{x},\overline{x}\}, \qquad t\in\widetilde{T}\backslash\{T\} && (4.4)\\ &= 0, \qquad t=T && (4.5)\end{aligned}$$

Tvrdenie 7. *Výnosovú funkciu (4.3) v čase T môžeme podľa (4.2), (4.1) a vety 6 zapísať ako*

$$\begin{aligned} Yd_T &= \sum_{k=1}^{k^*}\left[\sum_{t=t_{k-1}}^{t_k}[(Pr_t-Pr_{t-1})x_{t-1}-c|u_t|]\right] && (4.6)\\ &= \sum_{k=1}^{k^*}\left[(Pr_{t_k}-Pr_{t_{k-1}})x_{t_{k-1}}-c|u_{t_k}|\right] && (4.7)\end{aligned}$$

Presné odvodenie sa dá nájsť v [6].

4.2 Obchodovanie ako diskrétna veličina

Na to aby sme mohli vôbec uvažovať či sa cena zmení dostatočne musíme pridať nasledujúci predpoklad.

Predpoklad 1. *Predpokladajme, že $\exists\xi_1,\xi_2\in(0,1)$ také, že $\xi_1+\xi_2\geq 1$ a pre každé $t\in\widetilde{T}\backslash T$ pri známom Pr_t platí*

$$\begin{aligned} E\left[Pr_{t+1}|\widetilde{Pr_t}\right] < Pr_t - c &\Leftrightarrow P[Pr_{t+1}-Pr_t<-c|\widetilde{Pr_t}]>\xi_1,\\ E\left[Pr_{t+1}|\widetilde{Pr_t}\right] > c + Pr_t &\Leftrightarrow P[Pr_{t+1}-Pr_t>c|\widetilde{Pr_t}]\quad >\xi_2\end{aligned}$$

Splnenie predpoklad 1 obmedzuje stratu v prípade vyžčích strát a ako výnosov. Nemôže sa teda stať, že v prípade omnho častejších výnosov kôli málo častým ale vysokým stratám budem stredná hodnota zisku záporná.

Tvrdenie 8. *Nech $P[Pr_{t+1}-Pr_t\notin\{-c,c\}|\widetilde{Pr_t}]=0$ potom náhodná veličina $Y_{t+1}=\frac{sgn(Pr_{t+1}-Prt-c)+sgn(Pr_{t+1}-Prt+c)}{2}$ má diskrétne rozdenlenie na $\{-1,0,1\}$ a platí, že*

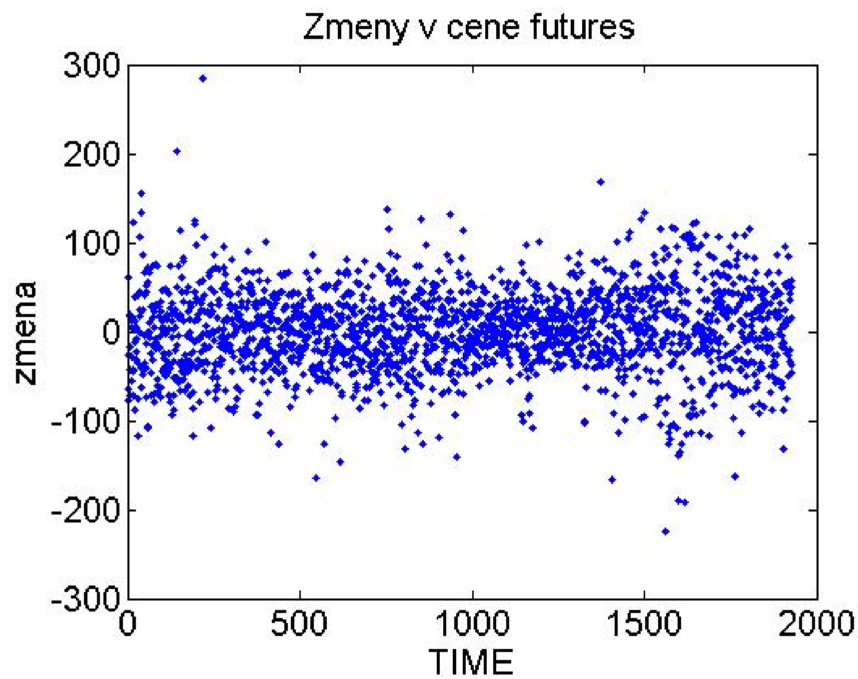
$$\begin{aligned} P[Y_{t+1}=1|\widetilde{Pr_t}] &= P[Pr_{t+1}-Prt>c|\widetilde{Pr_t}] && (4.8)\\ P[Y_{t+1}=0|\widetilde{Pr_t}] &= P[Pr_{t+1}-Prt\in(-c,c)|\widetilde{Pr_t}] && (4.9)\\ P[Y_{t+1}=-1|\widetilde{Pr_t}] &= P[Pr_{t+1}-Prt<-c|\widetilde{Pr_t}] && (4.10)\\ & && (4.11)\end{aligned}$$

Tvrdenie 9. *Predchádzajúce tvrdenia 8, 6 hovoria, že za predpokladu 1 a vedomosti o historickom vývoji ceny môžeme očakávaný výnos v nasledujúcom čase zapísať ako*

$$\begin{aligned}
E[Yd_{t+1}|Yd_t,Pr_t,x_t=\underline{x}] &= Yd_t + (\overline{x}-\underline{x})\Big(E[Pr_{t+1}|\widetilde{Pr}_t] - Pr_t - c\Big), \\
&\qquad\quad ak \quad P[Y_{t+1}=1|\widetilde{Pr}_t] > \xi_2 \\
&= Yd_t, \qquad inak \\
E[Yd_{t+1}|Yd_t,Pr_t,x_t=\overline{x}] &= Yd_t + (\overline{x}-\underline{x}()\Big(E[Pr_{t+1}|\widetilde{Pr}_t] - Pr_t - c\Big), \\
&\qquad\quad ak \quad P[Y_{t+1}=-1|\widetilde{Pr}_t] > \xi_1 \\
&= Yd_t, \qquad inak
\end{aligned}$$

4.3 Obchodovanie na reánych datach

Obchodovanie bolo vyskúšané na vzorke 3921 dát z cien finančného derivátu naviazaného na podkladové aktívum *Austrálskeho doláru*. Bolo vychádzaná z cien na konci dňa. Postupvalo sa rozhodovacím mechanizmom zavedeným v tvrdní 9, kde sa odhadovala závislosť na diskrétnej veličiny Y definovanej v tvrdení 8 na historických cenách. Keďže obchodovanie z futures nieje primárnym cieľom tejto práce tak bolo vopred predpokladané splnenie predpokladu 1 a hodnoty ξ_1 a ξ_2 boli experimentálne vyberané. Ukázalo sa, že vývoj ceny finančného derivátu je príliš tažko predikovateľný, teda nebola nájdená vhodná stratégia na obchodovanie. V nasledujúcom grafe je vidno ako sa menia ceny za 1 časové obdobie (pre prehľadnosť, je v grafe zaznamenaných len 2000 záznamov).



Kapitola 5

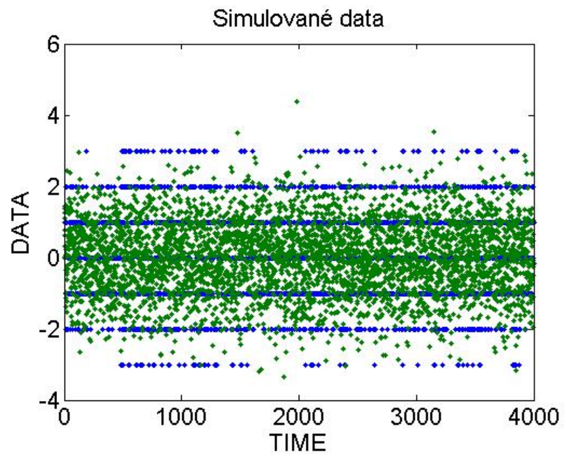
Implementácia

5.1 Príklad 1

Na testovanie bol zvolený 3 parametrový model, s rozdelením, ktoré môže nadobúdať 7 stavov $-3, \ldots, 3$, definovaný ako

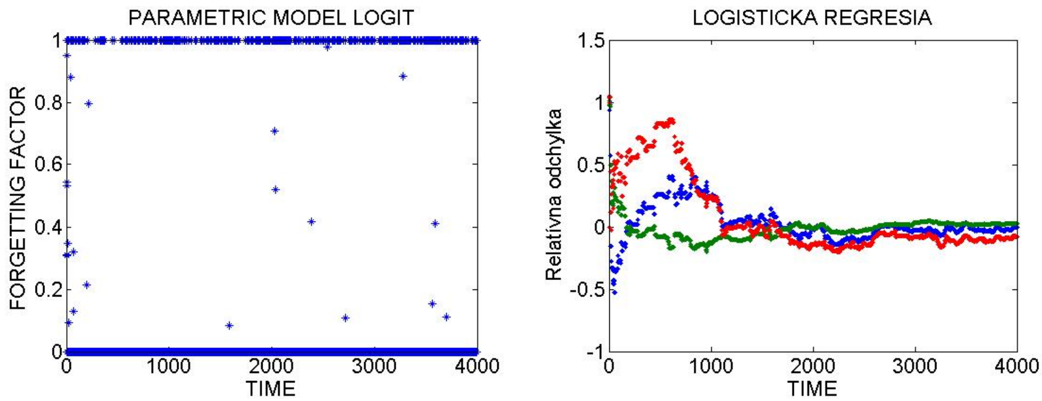
$$\begin{aligned} P[Y_j = k|\theta, X_j] &= \frac{exp\{-(k-\theta X_j)^2\}}{\sum_{i=-3}^{3} exp\{-(i-\theta X_j)^2\}}, k = -3, \ldots, 3 \\ \theta &= (-0.4, 0.5, 0.2) \\ X_j &= (Y_{j-2}, Y_{j-1}, W_j) \\ W_j &\sim N(0,1) \end{aligned}$$

Dinamicky bolo simulovaných 4000 dat (obrazok 5.1), na ktorý bol následne použitý algoritmus.



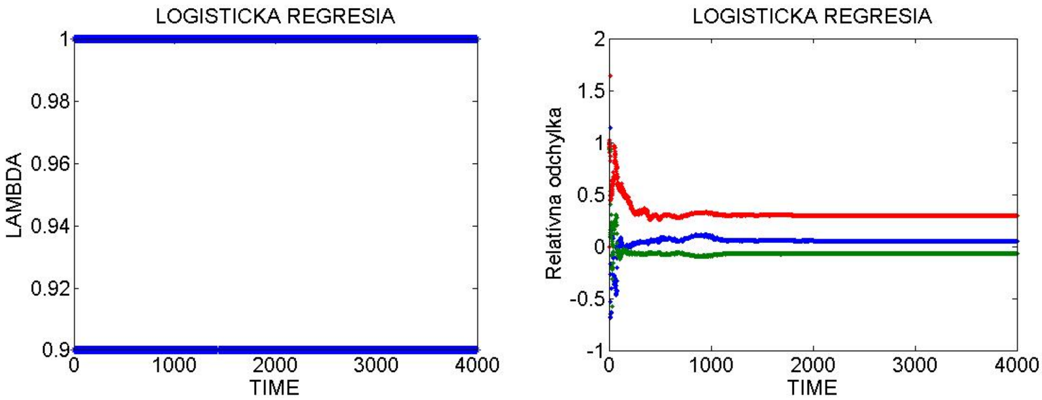
Obr. 5.1: 1. príklad, simulované data

Pre odhad parametrov (3.5),(3.6) boli použité metódy Monte Carlo s počtom simulácii 400. Pre hľadanie parametru λ 2.7 je použitá metóda dosadenia s krokom 0.01. σ je považovaný za známi parameter $\sigma = 1$. Apriórne rozdelenie bolo zvolené z rodiny GiW kde $\nu = 2$ a L je matica identity. Na obrazku 5.2 na ľavo je vidieť aké boli vypočítané hodnoty λ v čase a na obrázku 5.2 v pravo je vývoj relatívnej odchylky odhadu parametru θod skutočnej hodnoty. Odhad parametru θ v čase i je volený ako jeho stredná hodnota za predpodkladu, že parameter má hustotu p_i definovanú v kroku 1. Relatívna odchylka je definovaná ako $\frac{\theta - \mathsf{E}[\theta|p_i]}{\theta}$



Obr. 5.2: 1. príklad, odhadovanie bez ohraničenia

Pre problém s vkladaním aproximujúcej hustoty do vzorca pre hľadanie λ spomenutý v poznámke 8 sa pokusíme obmedziť parameter λ konštantou $\lambda_{min} = 0{,}9$ z dola.



Obr. 5.3: 1. príklad, odhadovanie pre ohranicene λ

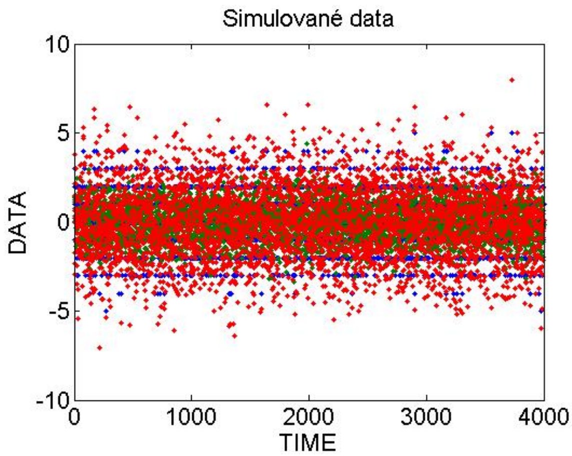
Porovnaním obrázkov 5.2 a 5.3 v pravo je vidieť, že ohraničenie parametru λ z dola spočiatku zlepšovalo odhad parametru θ ale postupne sa s prichádzajúcimi datami odhad parametru prestal meniť, respektíve sa menil príliš pomaly, zatiaľ čo pre neohraničené λ bol odhad presnejší.

5.2 Priklad 2

Na testovanie bol zvolený dvojparametrový model s rozdelením, ktoré môže nadobúdať 11 stavov, $-5,\ldots,5$.

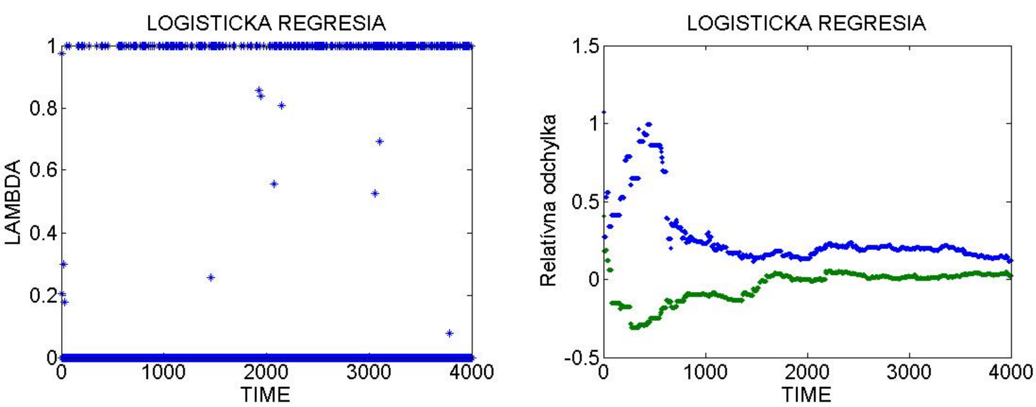
$$
\begin{aligned}
P[Y_j = k|\theta, X_j] &= \frac{exp\{-(k-\theta X_j)^2\}}{\sum_{i=-5}^{5} exp\{-(i-\theta X_j)^2\}}, k = -5,\ldots,5 \\
\theta &= (0.4,0.6) \\
X_j &= \left(W_j^1,W_j^2\right) \\
W_j^1 &\sim N\,(0,1) \\
W_j^2 &\sim N\,(0,4)
\end{aligned}
$$

Dinamicky bolo simulovaných 4000 dat (obrazok 5.2), na ktorý bol následne použitý algoritmus.



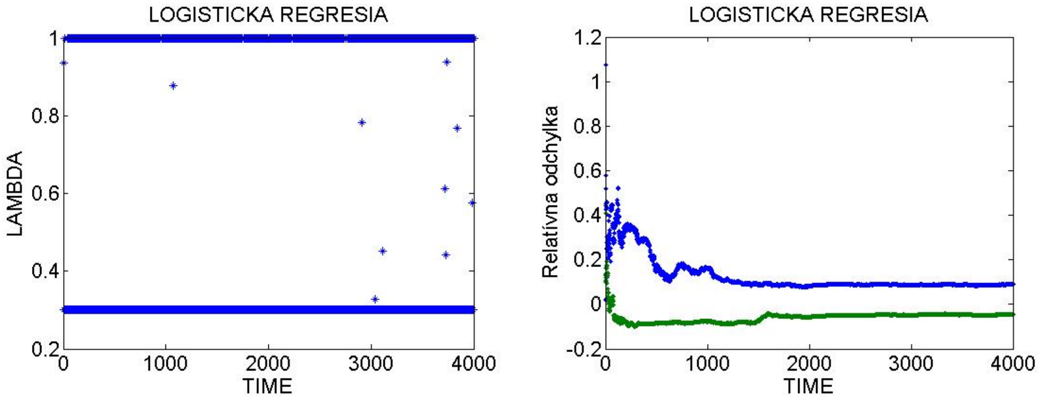
Obr. 5.4: 2. príklad, simulované data

Pre odhad parametrov 3.5,3.6 boli použité metódy Monte Carlo s počtom simulácii 400. Pre hľadanie parametru λ 2.7 je použitá metóda dosadenia s krokom 0.01. σ je považovaný za známi parameter $\sigma = 1$. Apriórne rozdelenie bolo zvolené z rodiny GiW kde $\nu = 2$ a L je matica identity.



Obr. 5.5: 2. príklad, odhadovanie bez ohraničenia

Pre problém s vkladaním aproximujúcej hustoty do vzorca pre hľadanie λ spomenutý v poznámke 8 sa pokusíme obmedziť parameter λ konštantou $\lambda_{min} = 0{,}3$ z dola.



Obr. 5.6: 2. príklad, odhadovanie pre ohranicene λ

Porovnaním obrázkov 5.5 a 5.6 v pravo je vidieť, že ohraničenie parametru λ spočiatku zrýchlilo konvergenciu k správnemu odhadu, neskôr sa ale ustálilo na nepresnej hodnote. Ohadnutie bez ohraničeného λ sa ku koncu začalo viac blížiť skutočnej hodnote.

Záver

V kapitole 3 bol implementovaný dinamický algoritmus popísaný v kapitole 2 na model logistickej regresie. Algoritmus je možné priebežne využívať na odhadovanie parametrov v logistickej regresii, ako je ukázané na praktických príkladoch. Tie v kapitole 6 ukazujú, že definovanie dolného ohraničenia parametru λ výrazne ovplivňuje rýchlosť konvergencie ako aj ju samotnú. V kapitole 4 sa ukázalo, že takéto predikovanie, nieje vhodné na obchodovanie s futures. Cena týchto finančných derivátov je príliš tažko predikovateľná. Možnosť zlepšiť obchodovaciu stratégiu by bola neuvažovať pravdepodobnosť zmeny ceny derivátu ale z historických dat, presne spočítať, v ktorých časových okamihoch by bolo vhodné nakupovať alebo predávať s maximálnym ziskom, čo by bola odhadovaná premenná. Na zlepšenie algoritmu je možné nahradiť heuristické dosadzovanie aproximujúcej hustoty do vzorca pre hľadanie parametru λ (2.8) nejakým iným presnejším tvrdením, čo by mohlo byť tiež predmetom ďalšej práce.

Zoznam použitej literatúry

[1] ANDĚL, J. (2007): *Základy matematické statistiky*
Matfyzpress, vydavatelství Matematicko-fyzikální fakulty Univerzity Karlovy v Praze.

[2] HUŠKOVÁ,M. (1985) *Bayesovské metody (skripta)*
Praha: Univerzita Karlova v Praze.

[3] KÁRNÝ M. (2014) *Aproximate Bayesian recursive estimation*
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic

[4] KULLBACK S. and LEIBLER R. A. (1951) *On information and sufficiency*
The George Washington University and Washington, D.C.

[5] FERDINAND ÖSTERREICHER (2002) *Csiszár's f-divergences-basic properties*
Institute of Mathematics, University of Salzburg, Austria.

[6] KÁRNÝ M., ŠINDELÁŘ J., PÍRKO Š., ZEMAN J. (2010) *Adaptively Optimized Trading with Futures*
Department of Adaptive Systems, Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

[7] ZVÁRA K. (2008) *Regrese*
Matfyzpress, vydavatelství Matematicko-fyzikální fakulty Univerzity Karlovy v Praze.

[8] CZEPIEL SCOTT A. (2002) *Maximum Likelihood Estimation of Logistic Regression Models: Theory and Implementation*

Použitý software:

[9] NEDOMA P., KÁRNÝ M., GUY T. V., NAGY I.,BOHM J. (2003) *Mixtools*
Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic.

Zoznam obrázkov