Odhad hustoty pravděpodobnosti metodou maximální penalizované věrohodnosti

Marian Koutný

Ústav matematiky a statistiky, PřF MU, Brno

25. březen 2014

Úvod

- Nematematicky řečeno: Zvolená data budu prokládat nějakou křivkou (hustotou) pomoci metody zmíněné v nadpise.

ML

- Metoda maximální věrohodnosti se snaží vykreslit hustotu pravděpodobnosti v těch bodech, ve kterých máme dána pozorování. Výsledkem by pak bylo to, že výsledná hustota bude skoková právě v oněch bodech. Taková metoda by se pak ale ve skocích blížila nekonečné hodnotě. Pro takovou metodu bude Fischerova míra informace konvergovat k hodnotě nekonečno. Mým cílem tedy bude tuto nepříjemnost nějak modifikovat a to tím, že odečteme penalizující funkci od té věrohodnostní, abychom dostali pěknější tvar funkce hustoty. Při volbě Fischerovy míry informace jako penalizující funkce se pak v teoretické rovině odečte od vysokého čísla jiné vysoké číslo a výsledná hustota bude pak mít přirozenější tvar, jaký požadujeme.

ML

- Máme náhodný vektor $X = (X_1, ..., X_n)$ a známe jeho hustotu $f(x, \theta)$.
  Předpokládejme, že $\theta \in \Theta$ je otevřený interval v $\mathbb{R}$. Při pevné hodnotě θ je hustota $f(x, \theta)$ funkcí x. Ale pro libovolné pevné x můžeme hustotu $f(x, \theta)$ chápat jako funkci parametru θ. Tuto funkci budeme značit $L(x, \theta)$ a nazývá se *věrohodnostní funkce*.

ML

Definice

Pokud existuje $\theta^* \in \Theta$ takové, že pro všechny $\theta \in \Theta$ platí

$$L(X, \theta) \leq L(X, \theta^*).$$

pak řekneme, že θ^* *je odhad parametru* θ *získaný metodou maximální věrohodnosti*. Pokud je funkce $L(x, \theta)$ spojitě diferencovatelná, pak θ^* nutně musí být řešením rovnice

$$\frac{\partial L(X, \theta)}{\partial \theta} = 0 \qquad \textit{pro} \quad \theta = \theta^*.$$

Tato rovnice se nazývá *věrohodnostní rovnice*.

ML

Definice

Pokud položíme $\ln 0 = -\infty$, tak $L(X, \theta) \leq L(X, \theta^*)$ bude platit právě tehdy pro každé $\theta \in \Omega$, když pro každé $\theta \in \Omega$ bude

$$\ln L(X, \theta) \leq \ln L(X, \theta^*).$$

Věrohodnostní rovnici pak tedy můžeme přetransformovat do tvaru

$$\frac{\partial \ln L(X, \theta)}{\partial \theta} = \frac{\partial \ln f(X, \theta)}{\partial \theta} = 0 \qquad pro \quad \theta = \theta^*.$$

Fischerova míra informace

Definice

Fisherova míra informace o parametru θ je

$$J(\theta) = \int_{-\infty}^{\infty} \left(\frac{f'(x,\theta)}{f(x,\theta)} \right)^2 \mathrm{d}F(x,\theta).$$

$$w = w(f) = L - \psi(f),$$

kde $\psi(f)$ je penalizující funkce.

Fischerova míra informace

- Nyní je potřeba vybrat nějakou vhodnou penalizující funkci takovou, pro kterou bude dobře fungovat metoda maximální věrohodnosti. Ukazuje se, že jako vhodnou penalizující funkci je možné vzít Fisherovu míru informace. Tedy zvolíme

$$\psi(f) = J = \int_{-\infty}^{\infty} \left(\frac{f'^2}{f} \right) \mathrm{d}x.$$

Pro veličinu J určitě platí, že

$$0 \leq J < \infty$$

Fischerova míra informace

- Protože platí

$$\mathrm{d}F = f\,\mathrm{d}x,$$

tak po úpravách Fisherovy informace

$$J = \int_{-\infty}^{\infty} \left(\frac{f'}{f}\right)^2 \mathrm{d}F.$$

je

$$J = \int_{-\infty}^{\infty} \left(\frac{f'}{f}\right)^2 f\,\mathrm{d}x = \int_{-\infty}^{\infty} \left(\frac{f'^2}{f^2}\right) f\,\mathrm{d}x = \int_{-\infty}^{\infty} \left(\frac{f'^2}{f}\right) \mathrm{d}x.$$

Funkce γ

- Pokud dále zvolíme substituci $f = \gamma^2$, tak obdržíme

$$\int_{-\infty}^{\infty} \left(\frac{f'^2}{f} \right) \mathrm{d}x = \int_{-\infty}^{\infty} \left(\frac{\left(\gamma^2\right)'^2}{\gamma^2} \right) \mathrm{d}x = \int_{-\infty}^{\infty} \left(\frac{(2\gamma\gamma')^2}{\gamma^2} \right) \mathrm{d}x$$

což po úpravě dává

$$\psi(f) = J = 4 \int_{-\infty}^{\infty} \gamma'^2 \, \mathrm{d}x.$$

Výhody funkce γ

- $f \geq 0 \Rightarrow \gamma \in \mathbb{R}$
- Musí samozřejmě platit, že $\int_{-\infty}^{\infty} \gamma^2 \mathrm{d}x = 1$
- Z toho plyne, že tato funkce integrovatelná v kvadrátu v Lebesgueově smyslu ($\int_{-\infty}^{\infty} \gamma^2 \mathrm{d}x < \infty$) a je tedy $\gamma \in L^2(\mathbb{R})$.
- Podle teorie o Hilbertových prostorech je γ limitou (součtem) své vlastní Fourierovy řady vzhledem k nějakému ortonormálnímu systému
- Ovšem do penalizující funkce budeme brát jak první derivaci funkce γ, tak i druhou, neboť průběh funkce (její zakřivenost) je závislý i na její druhé derivaci. Takže je za vhodné vzít trochu obecnější penalizující funkci, která závisí jak na první, tak na druhé derivaci.

$$\psi(f) = 4\alpha \int_{-\infty}^{\infty} \gamma'^2 \mathrm{d}x + \beta \int_{-\infty}^{\infty} \gamma''^2 \mathrm{d}x \qquad \alpha, \beta \geq 0,\ \alpha + \beta > 0$$

Hermitovské polynomy

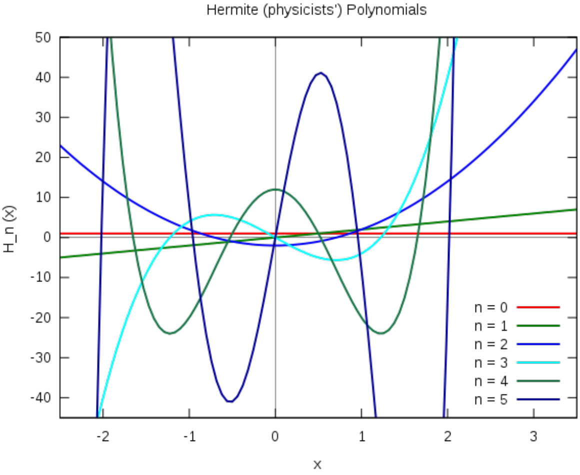
Definice

Posloupnost funkcí

$$H_n(x) = (-1)^n e^{x^2} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2} \quad n \geq 0$$

se nazývá *hermitovské polynomy*.

Hermitovské polynomy



Obrázek : Hermitovské polynomy

Hermitovské funkce

Definice

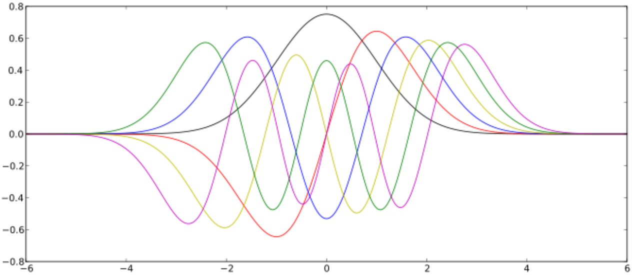
Posloupnost

$$\phi_n(x) = e^{\frac{-x^2}{2}}(2^n n!\sqrt{\pi})^{\frac{-1}{2}} H_n(x) = (-1)^n e^{\frac{x^2}{2}}(2^n n!\sqrt{\pi})^{\frac{-1}{2}} \frac{\mathrm{d}^n}{\mathrm{d}x^n} e^{-x^2}$$

se nazývá *hermitovské funkce*.

$$\begin{aligned}
\phi_0(x) &= \pi^{-\frac{1}{4}} e^{-\frac{x^2}{2}} \\
\phi_1(x) &= \sqrt{2}\pi^{-\frac{1}{4}} x e^{-\frac{x^2}{2}} \\
\phi_2(x) &= \sqrt{2}\pi^{-\frac{1}{4}} (2x^2 - 1) e^{-\frac{x^2}{2}} \\
&\vdots
\end{aligned}$$

Hermitovské funkce



Obrázek : Hermitovské funkce

Hermitovské funkce

Věta

Posloupnost hermitovských funkcí $\phi_n(x) = e^{\frac{-x^2}{2}}(2^n n!\sqrt{\pi})^{\frac{-1}{2}} H_n(x)$ *tvoří ortonormální systém v prostoru* $L^2(\mathbb{R})$.

Věta

Nechť $\{\phi_k\}_{k=0}^{\infty}$ *je ortonormální posloupnost hermitovských funkcí v prostoru* $L^2(\mathbb{R})$ *a nechť* γ *je reálná funkce. Pak podle předcházející teorie platí, že*

$$\gamma = \gamma(x) = \sum_{k=0}^{\infty} c_k \phi_k(x)$$

Derivace Hermitovské funkce

Hermitovské funkce $\phi_n(x) = e^{\frac{-x^2}{2}} (2^n n! \sqrt{\pi})^{\frac{-1}{2}} H_n(x)$ jsou spojitě diferencovatelné a jejich $m-$tá derivace je

$$\phi_n^{(m)}(x) = \sum_{k=0}^{m} \binom{m}{k} (-1)^k 2^{\frac{m-k}{2}} \sqrt{\frac{n!}{(n-m+k)!}} \quad \phi_{n-m+k}(x) H_k(x),$$

Speciálně pak první derivace je

$$\phi_n'(x) = \sqrt{2n}\ \phi_{n-1}(x) - x\ \phi_n(x).$$

Druhá derivace je

$$\phi_n''(x) = 2\sqrt{n(n-1)}\ \phi_{n-2}(x) - 2\sqrt{2n}\ \phi_{n-1}(x)x + (x^2-1)\ \phi_n(x).$$

Integrál Hermitovské funkce

Věta

Pro integrály hermitovských funkcí platí

$$\int_{-\infty}^{\infty} x\phi_m\phi_n \mathrm{d}x = \sqrt{\frac{m}{2}}\delta_{m,n+1} + \sqrt{\frac{n}{2}}\delta_{m+1,n}$$

$$\begin{aligned}\int_{-\infty}^{\infty} x^2\phi_m\phi_n \mathrm{d}x &= \frac{1}{2}\sqrt{m(m-1)}\delta_{m,n+2} + \frac{1}{2}\sqrt{n(n-1)}\delta_{m+2,n} + \\ &+ \left(m+\frac{1}{2}\right)\delta_{m,n}\end{aligned}$$

Odvození výrazu $\int_{-\infty}^{\infty} \gamma'^2 \mathrm{d}x$

$$\int_{-\infty}^{\infty} \gamma'^2 \mathrm{d}x = \int_{-\infty}^{\infty} \left(\sum_{m=0}^{\infty} c_m \phi_m \right)'^2 \mathrm{d}x = \int_{-\infty}^{\infty} \left(\sum_{m=0}^{\infty} c_m \phi_m' \right)^2 \mathrm{d}x$$

$$= \int \left(\sum_{m=0}^{\infty} c_m \phi_m' \right) \left(\sum_{n=0}^{\infty} c_n \phi_n' \right) \mathrm{d}x = \sum_{m,n=0}^{\infty} c_m c_n \int_{-\infty}^{\infty} \phi_m'(x) \phi_n'(x) \mathrm{d}x$$

$$= \sum_{m,n=0}^{\infty} c_m c_n \int \left(\sqrt{2m}\phi_{m-1}(x) - x\phi_m(x) \right) \left(\sqrt{2n}\phi_{n-1}(x) - x\phi_n(x) \right)$$

Odvození výrazu $\int_{-\infty}^{\infty} \gamma'^2 \mathrm{d}x$

Projedou se všechny kombinace m a n a výjde, že

- $m = n : \sum_{m=0}^{\infty} c_m^2 \left(m + \frac{1}{2}\right)$
- $m = n + 2 : \sum_{n=0}^{\infty} c_{n+2} c_n \left[-\frac{1}{2}\sqrt{(n+2)(n+1)}\right]$
- $m = n - 2 : \sum_{m=0}^{\infty} c_m c_{m+2} \left[-\frac{1}{2}\sqrt{(m+2)(m+1)}\right]$
- 0 jinak

Po sečtení máme pak

$$\int_{-\infty}^{\infty} \gamma'^2 \mathrm{d}x = \sum_{m=0}^{\infty} \left(m + \frac{1}{2}\right) c_m^2 - \sum_{m=0}^{\infty} \sqrt{(m+2)(m+1)}\, c_m c_{m+2}.$$

Odvození výrazu $\int_{-\infty}^{\infty} \gamma''^2 \mathrm{d}x$

Nyní stejně jako v předcházejícím případě odvodíme vzorec pro $\int_{-\infty}^{\infty} \gamma''^2 \mathrm{d}x$. Celkem tedy po sečtení všech alternativ můžeme psát, že

$$\begin{aligned}\int_{-\infty}^{\infty} \gamma''^2 \mathrm{d}x \quad = \quad & \frac{3}{4}\sum_{m=0}^{\infty}\left(2m^2+2m+1\right)c_m^2 - \\ & -\sum_{m=0}^{\infty}(2m+3)\sqrt{(m+1)(m+2)}c_m c_{m+2} \\ & +\frac{1}{2}\sum_{m=0}^{\infty}\sqrt{(m+1)(m+2)(m+3)(m+4)}c_m c_{m+4}\end{aligned}$$

Řešení soustavy

Tudíž funkce $w(\alpha, \beta, f)$ je maximalizována řešením soustavy rovnic

$$2\sum_{i=1}^{N} \phi_k \left(\sum_{m=0}^{\infty} c_m \phi_m\right)^{-1} - 4\alpha \left[2kc_k - \sqrt{(k+1)(k+2)}c_{k+2} - \right.$$
$$\left. - \sqrt{(k-1)k}c_{k-2}\right] - \beta \left[\frac{1}{2}\sqrt{(k-3)(k-2)(k-1)}c_{k-4} + \right.$$
$$+\frac{1}{2}\sqrt{(k+4)(k+3)(k+2)(k+1)}c_{k+4} + 3k^2 c_k$$
$$\left. +3kc_k - (2k+3)\sqrt{(k+1)(k+2)}c_{k+2} - (2k-1)\sqrt{(k-1)k}\; c_{k-2}\right]$$
$$-2\lambda c_k = 0,$$

$$\sum_{m=0}^{\infty} c_m^2 = 1$$

Maticové řešení

$$2\sum_{i=1}^{N}\phi_k(x_i)\left(\sum_{m=0}^{\infty}c_m\phi_m(x_i)\right)^{-1}$$

je vektor pravých stran. Bude se jednat tedy o jakousi iterační metodu. Pro případ, kdy $\alpha = \frac{1}{4}$, $\beta = 0$, $N = 1$ a $k \leq 3$ bude soustava vypadat zhruba takto

$$\begin{pmatrix} p & 0 & f & 0 \\ 0 & q & 0 & e \\ f & 0 & r & 0 \\ 0 & e & 0 & s \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Data

- Sledovaná data budou logaritmy jednodenních výnosnosti akciových kurzů (indexů)
- Výsledky se budou porovnávat s B-S vzorcem apod.

Už je konec !