Časové rady: domáca úloha 2

Termín odovzdania: 18. 11. 2024

Odovzdávanie:

- Úlohu rieši každý samostatne. Za odpísané úlohy je nula bodov - pre tých, ktorí úlohu odpísali aj pre tých, ktorí ju dali odpísať.
- Odovzdávajú sa:
  - výpočty na papieri[1]
  - numerické výpočty mailom na adresu **beata.ulohy@gmail.com**, s predmetom **CR 2024 - DU2 - meno** - v rovnakej forme ako v prvej domácej úlohe
- Ak použijete nejaký kód, výpočet, tvrdenie a pod., ktoré nájdete na internete, treba ho citovať - uviesť na konci riešenia odkaz a odvolať sa naň na príslušnom mieste v texte. V prípade použitia umelej inteligencie treba v maili okrem úlohy uviesť aj kompletný priebeh jej použitia - všetky vaše otázky a všetky odpovede UI (v pdf alebo txt formáte) a informáciu o tom, akú konkrétnu UI ste použili.

Príklad 1. Zrekonštruujeme príklad zo slajdov o ARMA modeloch pre výmenný kurz s ČR a dopníme ho ďalšími analýzami. Budeme pritom používať dáta uložené v `ARMAprednaska.Rdata`, ktoré načítame pomocou `load`. Súbor je na webe pri prednáške o ARMA modeloch. **Modelujeme diferencie logaritmov** výmenného kurzu.

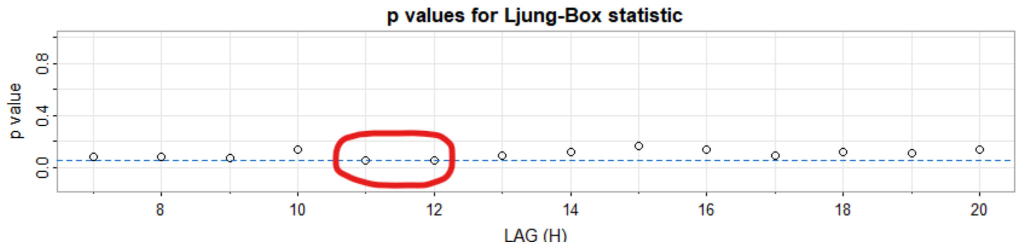
(a) [**1 bod**] Overte tvrdenia na str. 5 o vhodnosti modelov - teda, že AR a MA s rádom menším ako 5 nevyhovujú, ale AR(5), MA(5) a zmiešaný model ARMA(3,1) vyhovujú. Pod "vyhovujú"rozumieme výsledok Ljung-Boxovho testu pre rezíduá, pričom maximálny uvažovaný lag berieme ten, ktorý počíta funkcia `sarima` pri kreslení výstupu. Vysvetlite, čo sa testuje a prečo získaný výstup zodpovedá vyhovujúcemu, resp. nevyhovujúcemu modelu.

(b) [**1 bod**] Slajdy boli vytvorené v staršej verzii balíka `astsa` a získavanie BIC pomocou postupu v slajdoch na str. 5 nefunguje. Objekt `model` obsahuje model odhadnutý funkciou `sarima`. Ako získať hodnotu BIC v tejto verzii (vo výstupe vidíme, že ide o verziu 2.1) balíka?

```
> model$fit$BIC
NULL
>
> baliky <- installed.packages()
> colnames(baliky)
 [1] "Package"               "LibPath"               "Version"               "Priority"
 [5] "Depends"               "Imports"               "LinkingTo"             "Suggests"
 [9] "Enhances"              "License"               "License_is_FOSS"       "License_restricts_use"
[13] "OS_type"               "MD5sum"                "NeedsCompilation"      "Built"
> baliky["astsa", "Version"]
[1] "2.1"
```

(c) [**1 bod**] Pri odhadnutí modelu ARMA(4,2) vychádza Ljung-Boxov test pre lagy 11 a 12 vizuálne na vyznačenej 5% hranici. Zistite presnú číslednú hodnotu týchto p-hodnôt.

[1] Odovzdať ich môžete na prednáške. Ak na prednáške nebudete, pošlite do termínu odovzdania mail s oskenovaným alebo odfoteným riešením. Následne odovzdajte toto (už ďalej neupravované) riešenie na neskoršej prednáške alebo na cvičení, môžete ho aj po niekom poslať. Opravovať ho budem potom, ako ho dostanem na papieri.



(d) **[2 body]** Zapíšte rovnicu pre získaný model ARMA(5, 3). Spravte preň podobný obrázok, ako je na str. 45 v slajdoch o ARMA modeloch, ktorý naznačuje, že by sa mal rád AR aj MA zmenšiť. Interpretujte tento obrázok: Ak by ste ho dostali ako informáciu o odhadovanom modeli, čo by ste navrhli ako nasledujúci krok a prečo?

(e) **[3 body]** Vynechajte z dát posledných 10 hodnôt. Pre zostávajúce dáta nájdite ľubovoľný model s dobrými rezíduami (na základe Ljung-Boxovho testu z výsupu). Ukážte, že je stacionárny a invertovateľný. Spravte predikcie pre 10 vynechaných hodnôt a znázornite ich spolu s intervalmi spoľahlivosti a skutočnými hodnotami.

Príklad 2. Budeme sa zaoberať otázkou o súčte autokovariancií autoregresného modelu, t.j. o hodnote sumy $\sum_{k=-\infty}^{\infty} \gamma(k)$, kde $\gamma(k) = \mathrm{cov}(x_t, x_{t-k})$ pričom $x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + \varepsilon_t$ je stacionárny AR proces. Otázka sa riešila na *Stack Overflow*, odkiaľ preberieme niektoré postupy:
`https://stats.stackexchange.com/questions/371792/sum-of-autocovariances-for-arp-model`

(a) **[10 bodov]** Hľadanú sumu najskôr numericky ilustrujeme na príklade AR(2) procesu. Zvoľte si nejaký stacionárny AR(2) proces spolu s disperziou bieleho šumu a odvoďte jeho disperziu. Môžete ju počítať pomocou sústavy lineárnych rovníc z prednášky pre konkrétne hodnoty parametrov alebo odvodiť všeobecný predpis pre disperziu AR(2) procesu. Ak budete využívať sústavu rovníc z prednášky, napíšte aj jej odvodenie spolu so slovným zdôvodnením všetkých krokov výpočtu.

Pomocou funkcie `ARMAacf` vypočítajte hodnoty autokovariancií pre dostatočne veľa lagov. Vyjadrite čiastočný súčet $\sum_{k=-N}^{N} \gamma(k)$ pomocou disperzie $\gamma(0)$ a autokovariancií $\rho(k)$ pre $k = 1, \ldots, N$. Vypočítajte numericky pomocou tohto vyjadrenia čiastočné súčty pre váš proces pre $N = 1, 2, \ldots, N_{max}$ a znázornite ich graficky. Maximálnu hodnotu N_{max} zvoľte tak, aby bola viditeľná konvergencia čiastočných súčtov. (Ostatné výsledky z prednášky môžete používať bez dôkazov.)

(b) **[3 body]** V riešení na *Stack Overflow* sa využije vlastnosť

$$\sum_{k=0}^{\infty} \psi_k = \frac{1}{1 - \sum_{i=1}^{p} \phi_i},$$

kde kde $x_t = \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$ je Woldova reprezentácia nášho procesu, pričom vieme, že $\psi_0 = 1$. Dokážte toto tvrdenie. Na inšpiráciu sa môže hodiť nasledovná časť diskusie:

> @Greenparker yep, that's right. And $\psi(z)$ is defined to be the reciprocal of the $\phi(z)$ (complex) polynomial. In other words $(\psi_0 + \psi_1 z^1 + \psi_2 z^2 + \cdots)(1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p) = 1$. As long as the AR(p) model is causal, we won't divide by zero if we move the $\phi(B)$ to the other side in the second equation.
> – Taylor Oct 15, 2018 at 23:09

> Thanks, that clears things up. Regarding (2), the closed form expression is available for AR(1). I am intuitively certain there is a closed form expression for (2) for the general AR(p). I know the closed form expression for (2) for a VAR(1) model, and there must be a connection between an AR(p) and a p-variate VAR(1).
> – Greenparker Oct 16, 2018 at 13:32

(c) [**4 body**] Teraz môžeme spraviť riešenie pôvodného zadania. Pozrieme sa pritom na riešenie uvedené na webe:

$$
\begin{aligned}
\sum_{k=-\infty}^{\infty} \gamma(k) &\overset{(1)}{=} \sum_{k=-\infty}^{\infty} \operatorname{cov}\left(\sum_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \sum_{j=0}^{\infty} \psi_j \epsilon_{t+k-j}\right) \\
&\overset{(2)}{=} \sum_{k=-\infty}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \operatorname{cov}(\epsilon_{t-i}, \epsilon_{t+k-j}) \\
&\overset{(3)}{=} \sigma^2 \sum_{k=-\infty}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j 1(t-i=t+k-j) \\
&= \sigma^2 \sum_{k=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_{k+i} \\
&= \sigma^2 \sum_{i=0}^{\infty} \sum_{k=-\infty}^{\infty} \psi_i \psi_{k+i} \\
&\overset{(4)}{=} \sigma^2 \sum_{i=0}^{\infty} \sum_{k=-i}^{\infty} \psi_i \psi_{k+i} \\
&\overset{(5)}{=} \sigma^2 \left(\sum_{k=0}^{\infty} \psi_k\right)^2 \\
&\overset{(6)}{=} \sigma^2 \left(\frac{1}{1-\sum_{i=1}^{p} \phi_i}\right)^2 .
\end{aligned}
$$

- Odvoďte kroky (1), (2), (3), vysvetlite pritom aj značenie vyznačené modrou farbou, teda čo znamená $1(t-i-t+k-j)$.
- V nasledujúcich riadkoch vystupujú členy ψ_k aj so zápornými indexami (vyznačené žltou farbou), hoci koeficienty Woldovej reprezentácie sú definované len pre nezáporné indexy. Ako autor chápe tieto koeficienty so zápornými indexami? Spravte prechod na rovnosť (4) takým spôsobom, že v ňom tieto ψ_k so zápornými indexami nebudú vystupovať a zdôvodnite správnosť vášho postupu.
- Odvoďte rovnosť (5).
- Rovnosť (6) je pomocné tvrdenie z predchádzajúceho bodu domácej úlohy, ktorým sa dôkaz končí. Tvrdenie je teda dokázané. Doplňte do vášho grafu čiastočných súčtov $\sum_{k=-N}^{N} \gamma(k)$ hodnotu limity, ku ktorej tieto čiastočné súčty konvergujú.

Vo všetkých dôkazových úlohách sa hodnotí nielen správnosť napísaných vzťahov, ale aj ich slovné zdôvodnenie - jeho správnosť a úplnosť. Výsledkom má byť matematický text, ktorý jednotlivé kroky vysvetľuje, nielen postupnosť rovností, šípiek a pod.