

Časové rady: domáca úloha 4

Termín odovzdania: 15. 11. 2023 (streda, skupina 1), 21. 11. 2023 (utorok, skupina 2)

- Úlohu vypracováva každý samostatne alebo v dvojici. Komunikácia ohľadom riešenia, zdieľanie kódu a pod. je zakázané. Pri odpísaní úlohy alebo jej časti je DÚ hodnotená 0 bodmi, a to aj pre tých, ktorí úlohu odpísali aj pre tých, ktorí ju dali odpísať.
- Posielajte vypracovanú úlohu v pdf formáte (súvislý, dobre čitateľný text doplnený grafmi, nie iba výstupy z R so stručnými poznámkami) a použitý kód ako samostatný súbor. Alternatíva: výstup z R markdownu v html alebo pdf, s uvedením kompletného kódu.
- Dáta priložte ako prílohu k mailu alebo vytvorte a priložte workfile (súbor s koncovkou `Rdata`). Cieľom je, aby som si v prípade použitia dát z nejakého R-kovského balíka nemusela inštalovať celý balík, ale len načítať dáta.
- Každý (resp. každá skupina) pracuje s inými dátami, rezervácia dát prebieha v google dokumente <http://bit.ly/3ta0eMt> v hárku ČR mEMM - DÚ4.

Príklad 1: ARIMA modelovanie (10 bodov).

Nájdite si vhodné dáta. Požiadavky na dáta:

- Nájdete pre ne vyhovujúci ARIMA model. Požiadavky na model: správny rád diferencovania, nekorelované rezíduá, stacionárny a invertovateľný model
- Dáta majú byť bez pravidelnej sezónnosti. Pravidelnú sezónnosť majú napríklad mesačné hodnoty teploty alebo prietoku riek, sezónne neočistené kvartálne hodnoty HDP, dáta Air-Passengers zo začiatku semestra a pod. - je to na grafe s priebehom jasne viditeľné. Ak počas hľadania dát na takéto dáta narazíte, odložte si ich, zídu sa v poslednej domácej úlohe.
- Nie z balíkov `astsa` a `datasets`, nie dáta použité na prednáške alebo na cvičení. Môžete použiť iné balíky, súčasťou ktorých sú dáta s časovými radmi, príklady k učebniciam, reálne dáta dostupné na internete.

V domácej úlohe:

- Z dát vynechajte niekoľko posledných hodnôt. Tieto nebudete používať pri hľadaní modelu, ale použite ich na zhodnotenie kvality predikcií.
- Vysvetlite, čo vyjadrujú vaše dáta, uveďte ich zdroj a znázornite ich priebeh.
- Zistite správny rád diferencovania - treba zobrať do úvahy trend a jednotkový koreň. Spíšte výsledky a doplňte ich výstupmi z R-ka tak, aby sa vaše tvrdenia dali skontrolovať bez spúšťania kódu.
- Nájdite vhodný ARIMA model pre vaše dáta (požiadavky v úvode zadania). Overte stacionaritu a invertovateľnosť. Váš komentár k modelu znovu doplňte výstupmi z R-ka tak, aby sa vaše tvrdenia dali skontrolovať bez spúšťania kódu.
- Spravte predikcie z vášho modelu (pre pôvodnú premennú, nie pre prípadné diferencie) a porovnajte ich so skutočnými hodnotami, ktoré ste na začiatku vynechali.
- Zhodnoťte váš model. Čo si o ňom myslíte?