

Rekurzív és egylépéses idősorelőjelzők pontosságának vizsgálata

Szakos Máté

Lichter Bertalan

Kivonat

A gyakorlatban használt előrejelző modellektől az esetek többségében egy előre ismert előrejelzési horizonton szeretnénk tudni az idősorok jövőbeli értékeit. Ez elsősorban abból adódik, hogy a legtöbb alkalmazásban a döntések fix jövőbeli értékekre épülnek, például egy napra, hétre előrejelzett értékekre. Ha megvizsgáljuk a legelterjedtebb idősorelőjelző modelleket, azt láthatjuk, hogy a legtöbb esetben rekurzív előrejelzést alkalmaznak, azaz a jövőbeli értékeket a múltbeli értékekből, pontonként számítják ki. Ez sokszor jó illeszkedést tesz lehetővé, azonban megjelenhet az úgynevezett drift jelenség. Ennek lényege, hogy a modell kis hibával előrejelzi a jövőbeli értékeket, azonban a hibák összeadódnak, így a jövőbeli előrejelzések pontossága csökkenhet. Ezzel szemben az egylépéses előrejelzés esetén a modell minden egyes jövőbeli értéket külön számít ki, így elkerülhető a drift jelenség. Utóbbi megközelítés hátránya azonban, hogy a legtöbb esetben rosszabb illeszkedést eredményez, mivel a függvény, amelyet megközelít, lényegesen bonyolultabb, mint a rekurzív előrejelzés esetén. Jelen házi feladat célja, hogy a két megközelítés pontosságát összehasonlítsa, és megvizsgálja, hogy melyik esetben érhető el jobb előrejelzési teljesítmény. Házi feladatunkban azt találtuk, hogy stacioner, periodikus idősorok esetén a rekurzív előrejelzés sokszor kevésbé pontos, mint az egylépéses előrejelzés. Az interferáló jel, illetve a fehér zaj mértéke nagy hatással van a pontosságra, azonban szinte minden esetben az egylépéses előrejelzés volt a pontosabb.

Motiváció

[1]

A vizsgált idősor

A vizsgált idősor egy szintetikus, periodikus jel, amelyet a következő képlettel definiálunk:

$$\begin{aligned}
x(t) &= \sin(a \cdot t + b) \cdot \sin(c \cdot t) + \rho(t) + \epsilon(t), \\
s.t. \quad \rho(t) &= \sum_{i=1}^N \nu_i \sin(2\pi \cdot i \cdot f_{\text{base}} \cdot t + \phi_i), \\
\epsilon(t) &\sim \mathcal{N}(0, \sigma^2)
\end{aligned}$$

A jel egy nemlineáris, periodikus determinisztikus rész, egy periodikus interferáló zaj, és egy fehér zaj összegéből áll. Az interferáló jelet úgy állítottuk össze, hogy egy adott f_{base} alaphfrekvenciát, valamint annak felharmonikusait tartalmazza. Ez jól közelíti a valós életben előforduló periodikus zajokat, például az elektromos hálózatokból származó zajokat. A fehér zaj pedig egy normális eloszlású véletlenszerű zaj, amely az érzékelő hibáját hivatott modellezni.

A periodikus jel a következő paraméterekkel definiálható:

- a : a periodikus jel egyik komponensének frekvenciája
- b : a periodikus jel komponensei közti fáziseltolás
- c : a periodikus jel másik komponensének frekvenciája

Az interferáló zaj a következő paraméterekkel definiálható:

- N : az interferáló zaj komponenseinek száma
- ν_i : az interferáló zaj komponenseinek amplitúdója
- f_{base} : az interferáló zaj komponenseinek alaphfrekvenciája
- ϕ_i : az interferáló zaj komponenseinek fáziseltolása

A fehér zaj a következő paraméterekkel definiálható:

- σ : a fehér zaj szórása

Megvalósítás

A kódot tartalmazó jegyzetfüzetet itt lehet elérni.

Kísérletek

Konklúzió

Irodalomgyűjtemény

- [1] S. B. Taieb, R. J. Hyndman, és mtsai., *Recursive and direct multi-step forecasting: the best of both worlds*, köt. 19. Department of Econometrics and Business Statistics, Monash Univ., 2012.