

GLAJENJE ČASOVNIH VRST

Časovna vrsta je zaporedje podatkov y_1, y_2, \dots izmerjenih v zaporednih časovnih trenutkih, med katerimi so običajno enaki časovni razmiki.

Glajenje je aproksimacija zbranih podatkov s krivuljo, ki vsebuje vse bistvene vzorce iz podatkov, ne vsebuje pa šuma in drugih mikrostruktur. Zglajena časovna vrsta bolj nazorno prikaže dinamiko opazovanega pojava. Če se bistveni vzorec pojava s časom ne spreminja, lahko zglajeno časovno vrsto uporabimo za napovedovanje prihodnosti.

1. Uvoz in predstavitev podatkov

- (a) Na internetu poiščite podatke za zaključne tečaje trgovanja z bitcoini na borzi Bitstamp od 15. oktobra 2013 do 15. januarja 2014 (npr. internetna stran: <http://www.quandl.com>). Podatke uvozite v R.

Datoteko s podatki oddajte na spletni učilnici skupaj z R-ovo datoteko.

- (b) Narišite graf časovne vrste.

2. Glajenje z drsečim povprečjem reda k

Glajenje z drsečim povprečjem reda k časovni vrsti y_1, \dots, y_T priredi glajene vrednosti $\hat{y}_{k+1}, \dots, \hat{y}_T$, ki so povprečja zadnjih k vrednosti

$$\hat{y}_{t+1} = \frac{y_t + \dots + y_{t-k+1}}{k}.$$

Napoved vrednosti y_{T+h} v trenutku T je enaka \hat{y}_{T+1} .

Prileganje krivulje dani časovni vrsti merimo s *srednjo kvadratno napako*

$$\text{MSE} = \frac{1}{T-k} \sum_{t=k}^{T-1} (y_{t+1} - \hat{y}_{t+1})^2.$$

- (a) Pripravite funkcijo $G(\text{vrsta}, k)$, ki časovni vrsti y_1, \dots, y_T priredi zglajene vrednosti $\hat{y}_{k+1}, \dots, \hat{y}_T$.
- (b) Dano časovno vrsto zgladite z drsečim povprečjem reda 7. Pripravite še napoved za naslednji dan.
- (c) Na graf časovne vrste iz (1b) dodajte še zglajeno vrsto.
- (d) Izračunajte srednjo kvadratno napako.

3. Eksponentno glajenje

Enostavno eksponentno glajenje je podano z začetno vrednostjo $\ell_1 = y_1$ in rekurzivno formulo

$$\ell_t = \alpha y_t + (1 - \alpha)\ell_{t-1},$$

kjer je $0 \leq \alpha \leq 1$. Glajena vrednost $\hat{y}_{t+1} = \ell_t$ je uteženo povprečje opazovanj y_1, \dots, y_t .

Napoved vrednosti y_{T+h} v trenutku T je enaka $\hat{y}_{T+1} = \ell_T$.

Srednja kvadratna napaka je dana z enačbo

$$\text{MSE} = \frac{1}{T-1} \sum_{t=1}^{T-1} (y_{t+1} - \hat{y}_{t+1})^2.$$

- (a) Pripravite funkcijo $\text{EG}(\text{vrsta}, \alpha)$, ki sprejme časovno vrsto y_1, \dots, y_T ter parameter α in vrne zglajene vrednosti ℓ_1, \dots, ℓ_T .
- (b) Izberite $0 \leq \alpha \leq 1$ ter zgladite časovno vrsto. Izračunajte napoved za naslednji dan. Na graf časovne vrste iz (1b) dodajte še zglajeno vrsto.
- (c) Določite vrednost α^* , pri kateri zglajena časovna vrsta najmanj odstopa od opazovanih vrednosti. Za kriterij prileganja uporabite srednjo kvadratno napako.
- (d) Ponovite nalogo (b) pri optimalni vrednosti α .