

Aegridade analüüs

Praktikum nr. 6, 2016

Käsitletavad teemad: AR, MA ja ARMA tüüpi mudelite valimine aegrea kirjeldamiseks, sobivuse diagnostika

Sobiva mudeli valimise protseduur on järgmine:

1. Vaatleme aegrida, autokorrelatsioonifunktsiooni ja osaaautokorrelatsioonifunktsiooni, et teha kindlaks statsionaarsuse eelduse täidetud ja sobiva mudeli esialgne kuju. Mittestatsionaarse aegrea korral tuleb kindlaks teha sobiva teisenduse iseloom; seda vaatleme järgnevas praktikumis. Sobiva ARMA(p,q) tüüpi mudeli valikut abistavad faktid on järgmised:
 - AR(p) mudeli (ehk ARMA(p,0) mudeli) tunnuseks on see, et osaaautokorrelatsioonid võrduvad nulliga alates järgust $p + 1$. AR(1) mudeli puhul on lisakontrolliks teadmine, et sel juhul autokorrelatsioonid peavad kahanema nagu absoluutväärtuselt ühest väiksema arvu astmed.
 - MA(q) mudeli tunnuseks on autokorrelatsioonide võrdumine nulliga alates järgust $q + 1$. MA(1) mudeli puhul on lisakontrolliks osaaautokorrelatsioonide kahanemine nagu mingi absoluutväärtuselt ühest väiksema arvu astmed.
 - ARMA(1,q) mudeli tunnuseks on autokorrelatsioonide kahanemine vastavalt ühest väiksema arvuga korrutamisele alates järgust $q+1$ (st $q+1$ järku autokorrelatsioon on nagu mingi arv korda q -s autokorrelatsioon jne). ARMA(1,1) mudeli puhul peaks nii autokorrelatsioonid kui osaaautokorrelatsioonid kahane- ma alates järgust 2 vastavalt mingile absoluutväärtuselt ühest väiksema arvuga korrutamisele.
2. Sobitame tarkvara abil valitud tüüpi aegrea (st laseme hinnata vajalikud parameetrid).
3. Uurime leitud mudeli sobivust. Põhiliseks sobivuse kriteeriumiks on ühesammuliste prognoosivigade juhuslikkus; tüüpiliseks testiks on prognoosivigade autokorrelatsioonide nulliga võrdumise hüpoteesi kontroll. Selleks analüüsime prognoosivigade autokorrelatsioonikordajaid (kas jäävad etteantud veapiiridesse) ning leiame Ljung-Box statistiku põhjal tõenäosuse, et juhuslike prognoosivigade korral tekivad sellised autokorrelatsioonid, nagu andmete põhjal näha. Kui Ljung-Box testi põhjal saadud p -väärtused on alates mingist rühma suurusest kõik kuni umbes nihkeni 30 (mittese- soonse mudeli korral) piisavalt suured (näiteks suuremad kui 0.05), siis loeme mudeli adekvaatseks; vastasel korral sobitame uue mudeli ja kordame diagnostikat.

Praktikumi tegevused:

1. Leiame sobivad mudelid andmestikes `prax6.csv` ja `praktikum6.dat` olevatele and- mestikele, mis võivad vastata statsinaarsetele protsessidele. Selleks lugege kõigepealt andmed sisse ja moodustage vastavad aegread.
2. Mudeli sobitamiseks on käsk `arima` ning diagnostikaks käsk `tsdiag`. Siin on kasu- lik teada, et käsus `tsdiag` ei arvestata vähemalt praeguses R versioonis Ljung-Box statistiku väärtuste arvutamisel sobitatud mudeli parameetrite arvu, mistõttu leitud

p -väärtused on suuremad kui korrektse lähenemise korral. Samas suuremate autokorrelatsioonirühmade korral see erinevus kahaneb ning seetõttu on käesoleva kursuse raames lubatud otsuste tegemisel kasutada selle käsu tulemusi. Leidke neid käske kasutades vähimat arvu parameetreid sisaldavad sobivad mudelid kõikidele statsionaarsetele aegridadele. Käsule `arima` antakse mudeli kuju ette kolmest numbrist koosneva vektoriga `order`, millest esimene on autoregressiivse osa järk p , teine on AR, MA ja ARMA mudelite korral 0 ja kolmas on MA osa järk q . Parima mudeli kuju pange paberile kirja.

Märkus. R tarkvara korral tuleb arvestada, et väljundis näidatavad kordajad on täpselt sama märgiga, nagu nad mudeli kirjapanekul tekivad paremal pool võrdusmärgi ning et liige nimega "intercept" tähistab tegelikult rea keskväärtuse hinnangut.