

LABORATION 4

732G52 - HT2025

Intro

Innehåll:

- ARIMA-modeller

Denna laboration är till för er övning på kursmaterialet. Det finns ingen obligatorisk inläming för den. Uppgifterna utgår från R och Rstudio. Om du behöver repetition i R:

- Titta på kurshemsidan för 732G33: här.
- Cheat sheets här

Material:

- Time Series for Data Science: se R-kod, dataset mm: länk
- Forecasting: Principles and Practice (FPP), tredje upplagan: länk
- Repo för kursen: länk
- Om man behöver repetition på vanlig linjär regression, se kursen Linjära modeller 1, titta tex här: länk

Uppgifter

1. Simulera data från ARMA-modeller: Gå igenom koden som finns här: länk
Tanken är att ni ska bekanta er mer ARMA-familjen genom att simula-
ra tidserier med kända parametervärden och sedan studera tidseriegrafen,
den teoretiska autokorrelationen och den teoretiska partiella autocorrela-
tionen.
2. FPP3 kap 9: ARIMA models
 - (a) Gå igenom och återskapa koden i 9.1 (ej unit root test)
 - (b) Gå igenom och återskapa koden i 9.5
 - (c) Gå igenom och återskapa koden i 9.7 (ej Plotting the characteristic
roots)
 - (d) Gå igenom och återskapa koden i 9.9
 - (e) Gå igenom och återskapa koden i 9.10
 - Här finns kod relaterad till kap 9: länk1, länk2
3. FPP3 9.11 Exercises: Gör uppgifterna:
 - 6, 7, 9, 11, 15
4. Ladda in datamaterialet `hare` eller `color` som finns i paketet `TSA`. Ta
reda på vad data handlar om och gör en lämplig tidserieplot. Vad har
data för egenskaper? Avgör om tidserien är stationär eller ej. Om inte ge-
nomför lämpliga differanser. Studera SAC och SPAC och välj en lämplig
ARIMA-modell för tidserien. Anpassa den valda modellen. Plotta anpas-
sade värden tillsammans med data. Studera residualerna. Gör prognos 5
tidssteg framåt.
5. Ni ska studera antalet nya lastbilsregistreringar, se filen "NyLastbilregistreringar.csv".
som finns här: länk
 - (a) Plotta först data. Eftersom det är månadsdata med tydlig säsongsva-
riation så differentierar du serien för månad och sparar den i en ko-
lumn som du kallar `diff12`.
 - (b) Plotta nu `diff12` samt skatta acf, dvs beräkna SAC. Dessa båda grafer
tyder på att du behöver differentiera `diff12` en gång för trend. Gör
det och kalla denna diffade serie `diff12_1`.
 - (c) Plotta denna serie samt beräkna SAC. Visst ser den stationär ut!
Beräkna även SPAC på `diff12_1` och bestäm vilken ARIMA modell
som är lämplig.
 - (d) Skatta den valda modellen och skapa lämpliga residualplottar.

- (e) Är du nöjd med modellen? (Det är svårt att få riktigt bra residualer). Om inte, testa några olika modeller. När du är nöjd, ta fram prognoser för 12 månader fram i tiden. Korrigera nu dessa prognoser så att du får prognoser för antal nylastbilsregistreringar (alltså den odiffrade serien) . Kommentera.
6. Turist-data: utgå från filen "turisterTurkiet.csv". Beskrivningen av data finns i "tu_foreign_visitors_93-04.txt". Filerna ligger här: [länk](#)
- Logaritmiera serien och plotta den.
 - Målet är nu att anpassa en lämplig SARIMA-modell. Studera SAP och SPAC, och testa olika diffar om ni anser att det behövs. Använd kunskapen från dessa undersökningar för att välja en lämplig SARIMA-modell.
 - Anpassa modellen och studera residualerna. Är ni nöjd? Om inte försök att justera modellen. När ni är nöjda, gör prognos 12 månader framåt.
7. Det finns flera alternativ till paketet **fable** om man arbetar med ARIMA/SARIMA-modeller i R. Några exempel: I base R så finns funktionerna **ar()** och **arima()**. I paket **forecast** finns det funktioner för ARIMA-modellering. Den andra upplagan av boken FPP använder paketet **forecast**.