

## Předpověď výstupu s regresním modelem

**Úloha:** Předpověď výstupu soustavy jako informace o soustavě nebo pro další použití.

- jednorozměrný vstup i výstup
- obecný řád modelu
- obecný počet kroků predikce

Jako data se využívá výsledku simulace ze souboru T12SimCat1 který se ukládá jako datový soubor dataT11.dat v adresáři `_data` a odkud je natažen. Nic ale nebrání tomu, vyrobit si vlastní data - buď přímo v souboru T11SimReg1 nebo si data uložit na disk a natáhnout si je tak, jak je to připravené. Odhad parametrů modelu se provádí v souboru T21EstReg1, a to v batchové (jednorázové) formě. Pro tyto bodové odhady parametrů se provádí predikce na celém časovém intervalu.

Použitý model má tvar

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \dots + a_n y_{t-n} + b_0 u_t + b_1 u_{t-1} + \dots + b_n u_{t-n} + k + e_t$$

kde  $y$ ,  $u$ ,  $e$  jsou výstup, vstup a porucha,  $a$ ,  $b$ ,  $k$  jsou bodové odhady parametrů.

Predikce se provádí postupným dosazováním dat nebo predikcí z minulých kroků do modelu. Tedy např. pro 2 řád modelu a predikci na 2 kroky platí

$$\begin{aligned}\hat{y}_t &= a_1 y_{t-1} + a_2 y_{t-2} \\ \hat{y}_{t+1} &= a_1 \hat{y}_t + a_2 y_{t-1} \\ \hat{y}_{t+2} &= a_1 \hat{y}_{t+1} + a_2 \hat{y}_{t+2}\end{aligned}$$

kde stříška označuje bodovou predikci výstupu.

**Předpoklady:**  $e \sim N(0, r)$ ,  $r$  konstantní; vstup je generován předem.

Poznámka

*Právě v jednoduchosti uvažované predikce se nejlépe ukazuje, jak je výhodné použití bodových odhadů parametrů modelu soustavy. Použití celé aposteriori hustoty pravděpodobnosti by vedlo na tak složitý algoritmus, že bychom se nevyhnuli pracnému numerickému řešení.*

**Značení**

- $y$  - yt,
- $u$  - ut,
- $a$ ,  $b$ ,  $k$  - a, b, k,
- $r$  - cv (parametry ze simulace),
- thE - odhadnuté parametry,
- ordE - řád odhadovaného modelu,

## Volitelné parametry

- np - počet kroků predikce

## Doporučené experimenty

1. Měňte délku predikce a kontrolujte její kvalitu v grafu.
2. Přidejte další kritérium kvality, kterým je průměrná kvadratická chyba predikce. Chybu predikce určíte jako rozdíl mezi hodnotami výstupu a jejich predikcí

$$ep_i = y_i - yp_i, \quad i = 1, 2, \dots, nd$$

Kritérium potom je

$$SE = \frac{1}{nd} \sum_{i=1}^{nd} ep_i^2$$

Průměr chyby predikce by měl být blízko nuly, kvadratická chyba  $SE$  by měla být co nejmenší.

3. Měňte parametry simulované soustavy (zejména rozptyl šumu) a sledujte kvalitu predikce. Soustavu lze měnit v simulaci, tj. ve souboru T11SimReg1.sce.
4. Podstatný vliv na kvalitu odhadu a tedy i predikce má volba vstupního signálu. Jestliže vstup je příliš hladký bez skoků, soustava není pro odhad dostatečně vybudena, parametry se určí nedostatečně přesně a predikce není dobrá.

## Program

```
// Prediction np-steps with scalar regression model of order n
//
[u,t,n]=file(); // find working directory
chdir(dirname(n(1))); // set working directory
clear("u","t","n") // clear auxiliary data
exec("ScIntro.sce",-1),mode(0) // intro to sesion

// definition of regression vector
deff('ps=genpsi(t,n,y,u)', 'ps=[y(t-(1:n)),u(t-(0:n)) 1]', 'c')

load _data/dataT11.dat Sim // results from simulation T11SimReg1.sce
yt=Sim.Cy.yt;
ut=Sim.Cy.ut;
th=Sim.Cy.th;
cv=Sim.Cy.cv;
ord=Sim.Cy.ord;
load _data/dataT21.dat Est // results from estimation T21EstReg1.sce
thE=Est.Cy.th;
ordE=Est.Cy.ord;

np=5; // no of prediction steps (1=prediction from the model)
```

```

nd=length(yt);
yp=zeros(1,nd);
for t=(ordE+1):(nd-np+1)      // time loop
    yi=yt(1:(t-1));          // old data (at time t)
    for j=0:(np-1)           // loop of prediction
        ps=genpsi(t+j,ordE,yi,ut); // construction of regression vector
        yi(t+j)=ps*thE;      // auxiliary prediction
    end
    yp(t+np-1)=yi(t+np-1);   // final prediction at t
end
Pre.Cy.np=np;
Pre.Cy.yp=yp;

// Results
s=(ordE+1):(nd-np);
plot(s,yt(s),s,yp(s))
title('Output and its prediction')
legend('output','prediction');

save _data/dataT31.dat Sim Est Pre

```