

Předpověď výstupu s kategorickým modelem

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

- čtvercová tabulka modelu, tj. model $f(y_t|y_{t-1}\Theta)$
- obecná n -kroková predikce
- predikce s bodovými odhady ve tvaru maximum pravděpodobnosti nebo střední hodnota
- volba, zda predikce běží se známými nebo odhadovanými parametry

Data jsou v programu simulována.

Použitý model má tvar tabulky

$[u_t, y_{t-1}]$	$y_t = 1$	$y_t = 2$
1, 1	$\Theta_{1 11}$	$\Theta_{2 11}$
1, 2	$\Theta_{1 12}$	$\Theta_{2 12}$
2, 1	$\Theta_{1 21}$	$\Theta_{2 21}$
2, 2	$\Theta_{1 22}$	$\Theta_{2 22}$

kde y, u jsou výstup a vstup, $\Theta_{i|jk}$ jsou pravděpodobnosti.

Predikce se provádí mocněním tabulky modelu (jako čtvercové matice) a násobením počáteční hodnotou.

Předpoklady: Model s čtvercovou tabulkou.

Značení

- y - yt,
- u - ut,
- $\hat{\Theta}_t$ - thE - odhadnuté parametry.

Volitelné parametry

- np - počet kroků v predikci
- I_est - volba predikce se známými nebo odhadovanými parametry
- I_pred - volba bodových odhadů (maximum pravděpodobnosti nebo střední hodnota)

Doporučené experimenty

1. Volte různé druhy soustavy od málo neurčitých až po deterministické.
2. Zkuste různý počet kroků predikce np a porovnejte výsledky. Najděte hranici pro délku predikce, kdy ještě dostáváme přijatelné výsledky. Tato hranice je samozřejmě pro každý typ simulované soustavy jiná.
3. Porovnejte výsledky predikce se známými a odhadovanými parametry.
4. Upravte program tak, aby pracoval s off-line odhadnutými parametry. Porovnejte výsledky se známými, odhadovanými a odhadnutými parametry. Navrhněte kritérium pro porovnání.

Program

```
// Prediction with a "square" discrete model  $f(z(t)|z(t-1),th)$ 
//
[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 session

np=5; // prediction length
nd=500; // data sample length
I_est=2; // estimation: 1=yes, 2=no
I_pred=2; // prediction: 1=arg max, 2=expect.

T=[.99 .005 .005; // simulation model parameter
    .01 .98 .01
    .02 .01 .97];

// SIMULATION
z(1)=1; // initial output
for t=2:nd // data generation
    z(t)=(sum(rand(1,1,'unif')>(cumsum(T(z(t-1),:)))))+1;
end

// ESTIMATION AND PREDICTION
S=ones(T); // initial model statistics
vals=(1:max(size(T)))'; // values of output
for t=2:nd
    S(z(t-1),z(t))=S(z(t-1),z(t))+1; // statistics update
    // known or estimated parameters
    select I_est // estimation 1 = no, 2 = yes
    case 1 then
        Te=T;
    case 2 then
        Te=fnorm(S,2); // parameter point estimates
    end
    Tp=Te^np; // model for prediction
    // type of prediction point estimates
    select I_pred
    case 1 then
        [xxx yy]=max(Tp(z(t),:));
        yp(t+np)=yy;
    case 2 then
        yp(t+np)=Tp(z(t),:)*vals;
    end
end

// RESULTS
s=2:nd;
```

```
set(scf(1), 'position', [100 100 500 400])
plot(s, z(s), 'o', s, yp(s), '. ')
set(gca(), 'data_bounds', [1 nd .9 3.1]);
title 'Discrete data and their prediction'
legend('data', 'prediction');
```