

Odhad s regresním modelem

Úloha: Odhad parametrů modelu pro další použití.

- jednorozměrný vstup i výstup
- obecný řád modelu

Jako data se využívá výsledku simulace ze souboru T11SimReg1 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.

Vlastní odhad se provádí postupným přepočtem statistik V a κ v časové smyčce a teprve po jejím ukončení jsou spočteny bodové odhady parametrů modelu ve tvaru

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

kde y , u , e jsou výstup, vstup a porucha, a , b , k jsou parametry.

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

Poznámka

Teorie bayesovského odhadování je založena na úplném popisu parametrů pomocí aposteriorní hustoty pravděpodobnosti. Využití této informace v podobě celé hustoty pravděpodobnosti je ale značně komplikované. Z této hustoty (tedy přímo z napočítaných statistik odhadu) je možno spočítat bodové odhady parametrů. Jejich využití je jednoduché a velmi často postačující.

Značení

1. y - yt
2. u - ut,
3. a, b, k - a,b,k (parametry)
4. r - cv (rozptyl šumu)
5. thE - odhadnuté parametry
6. ordE - řád odhadovaného modelu

Volitelné parametry

- I_dataSim - data: simulovaná, reálná,
- !!! ostatní parametry se přebírají ze simulace.

Doporučené experimenty

1. Jeden ze základních předpokladů dobrého odhadu jsou data, která skutečně vypovídají o vlastnostech soustavy.
Zkuste měnit počet dat, vstupující do odhadu a porovnejte kvalitu odhadu pomocí připravené jednokrokové predikce (simulace z odhadnutého modelu).
2. Do odhadu je možno zavést tzv. exponenciální zapomínání. To způsobí, že vliv starších měřených dat se potlačuje a tím se zdůrazní význam aktuálních dat. To má význam, jestliže se vlastnosti soustavy pomalu mění nebo jestliže v minulosti byla změřena některá data chybová. Historie se pomalu zapomíná. Faktor zapomínání φ se volí blízký jedné (0.99, 0.995, 0.999) se do rovnic pro přepočtení statistik zavede takto

$$\begin{aligned}V_t &= \varphi V_{t-1} + \Psi \Psi' \\ \kappa_t &= \varphi \kappa_{t-1} + 1\end{aligned}$$

Zkuste použít různé hodnoty zapomínání při odhadu a porovnejte kvalitu odhadu v jednotlivých případech.

3. Zkuste odhadovat s modelem různého řádu a sledujte vliv na kvalitu odhadu.

Program

```
// Estimation of 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 session

deff('ps=genpsi(t,n,y,u)', 'ps=[y(t-(1:n)),u(t-(0:n)) 1]', 'c')

I_dataSim=2; // get data: 1=simulated, 2=real
ordE=2; // model order (valid for real data)

select I_dataSim
case 1
load _data/dataT11.dat Sim // load results of task T11
yt=Sim.Cy.yt; // or use your own data yt, ut
ut=Sim.Cy.ut; // and define order of the model
ord=Sim.Cy.ord;
case 2
load _data/dataReg1.dat yt ut // from file _data/dataReg1.dat
end

nd=length(yt);
if I_dataSim==1,
ordE=Sim.Cy.ord;
th=Sim.Cy.th;
end
```

```

nPsi=2*ordE+3;                // length of regression vector

V=1e-8*eye(nPsi,nPsi);       // comput. of information matrix
for t=(ordE+1):nd
    Ps=[yt(t) genpsi(t,ordE,yt,ut)]';
    V=V+Ps*Ps';
end
thE=inv(V(2:$,2:$))*V(2:$,1); // point estimates of parameters
Est.Cy.ord=ordE;
Est.Cy.V=V;
Est.Cy.th=thE;

// results
if I_dataSim==1
    disp('Simulated parameters')
    disp(th)
end
disp('Estimated parameters')
disp(thE')

if I_dataSim==1
    if length(th)==length(thE)
        bar([th' thE])
        title('Simulated and estimated parameters')
    end
else
    bar(thE)
    title('Estimated parameters')
end

save _data/dataT21.dat Est

```