

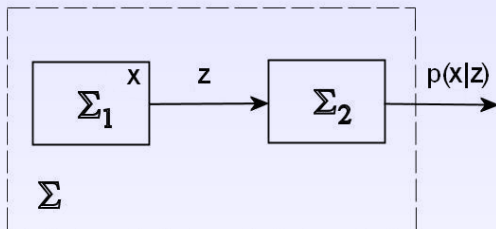
# Metody nelineárního odhadu

Jindřich Duník

Katedra kybernetiky & Výzkumné centrum Data - Algoritmy - Rozhodování  
Fakulta aplikovaných věd  
Západočeská univerzita v Plzni

21. dubna 2005

- 1 Úvod do problematiky odhadu stavu
- 2 Lokální metody odhadu stavu nelineárních stochastických systémů
- 3 Cíle pro budoucí práci



## Úloha odhadu stavu

*Systém odhadu*  $\Sigma$  je tvořen dvěma subsystemy  $\Sigma_1$  a  $\Sigma_2$ .

Subsystem  $\Sigma_1$  představuje *systém*, jehož neměřitelný stav je subsystemem  $\Sigma_2$  odhadován na základě měřených dat. Subsystem  $\Sigma_2$  je nazýván *estimátorem*.

Úloha odhadu stavu nachází uplatnění v mnoha oborech, mezi které patří teorie řízení, detekce poruch, informační fúze, navigace, zpracování řeči apod.

## Popis systému

$$\mathbf{x}_{k+1} = \mathbf{f}_k(\mathbf{x}_k, \mathbf{u}_k) + \mathbf{w}_k,$$

$$\mathbf{z}_k = \mathbf{h}_k(\mathbf{x}_k) + \mathbf{v}_k,$$

kde

- $\mathbf{x}_k$  - neměřitelný stav systému,
- $\mathbf{u}_k$  - vstup systému,
- $\mathbf{z}_k$  - měřitelný výstup systému,
- $\mathbf{f}_k(\cdot)$ ,  $\mathbf{h}_k(\cdot)$  - známé funkce,
- $\mathbf{w}_k$ ,  $\mathbf{v}_k$  - stavový šum a šum v rovnici měření.

## Řešení úlohy odhadu stavu

Pro úplné řešení problému estimace je nutné nalézt hustotu pravděpodobnosti stavu  $\mathbf{x}_l$  podmíněnou měřeními

$$\mathbf{z}^k = [\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k].$$

$$p(\mathbf{x}_l | \mathbf{z}^k) = ?$$

Úlohu odhadu lze rozdělit podle vztahu časových okamžiků  $k$  a  $l$  na

- úlohu predikce, kdy  $l > k$ ,
- úlohu filtrace, kdy  $l = k$
- a úlohu vyhlazování (retrodikce), kdy  $l < k$ .

# Rekurzivní vztahy pro výpočet podmíněných hustot pravděpodobnosti

## m-kroková predikce

$$p(\mathbf{x}_{k+m}|\mathbf{z}^k) = \int p(\mathbf{x}_{k+m}|\mathbf{x}_{k+m-1})p(\mathbf{x}_{k+m-1}|\mathbf{z}^k)d\mathbf{x}_{k+m-1}$$

## filtrace

$$p(\mathbf{x}_k|\mathbf{z}^k) = \frac{p(\mathbf{x}_k|\mathbf{z}^{k-1})p(\mathbf{z}_k|\mathbf{x}_k)}{p(\mathbf{z}_k|\mathbf{z}^{k-1})}$$

## m-krokové vyhlazování

$$p(\mathbf{x}_{k-m}|\mathbf{z}^k) = p(\mathbf{x}_{k-m}|\mathbf{z}^{k-m}) \int \frac{p(\mathbf{x}_{k-m+1}|\mathbf{z}^k)}{p(\mathbf{x}_{k-m+1}|\mathbf{z}^{k-m})} \times \\ \times p(\mathbf{x}_{k-m+1}|\mathbf{x}_{k-m})d\mathbf{x}_{k-m+1}$$

## Řešení rekurzivních vztahů

### Exaktní řešení

Exaktní řešení rekurzivních vztahů pro výpočet podmíněných hustot pravděpodobnosti je možné jen v několika případech, např. pro lineární systém, kde stavový šum, šum v rovnici měření a počáteční podmínka jsou dány normálním rozložením. Řešení rekurzivních vztahů pak vede na lineární estimační algoritmy (např. na **Kalmanův filtr**).

### Aproximativní řešení

- Lokální metody
- Globální metody

## Řešení rekurzivních vztahů

### Exaktní řešení

Exaktní řešení rekurzivních vztahů pro výpočet podmíněných hustot pravděpodobnosti je možné jen v několika případech, např. pro lineární systém, kde stavový šum, šum v rovnici měření a počáteční podmínka jsou dány normálním rozložením. Řešení rekurzivních vztahů pak vede na lineární estimační algoritmy (např. na **Kalmanův filtr**).

### Aproximativní řešení

- Lokální metody
- Globální metody



## Vlastnosti lokálních metod

Lokální metody jsou založeny na myšlence využití techniky Kalmanova filtru i v oblasti nelineárních systémů. Aby bylo možné tuto myšlenku použít, je nutné popis nelineárního systému vhodným způsobem aproximovat.

### Výhody a nevýhody

- Za výhodu lze považovat analytické řešení rekurzivních vztahů.
- Naopak hlavní nevýhodu lze spatřovat v lokální platnosti získaných odhadů.

### Transformace náhodné veličiny

- Nechť je známa  $\bar{\mathbf{x}} = E[\mathbf{x}]$  a  $\mathbf{P}_x = \text{cov}[\mathbf{x}]$  náhodné veličiny  $\mathbf{x}$  a nechť je dána nelineární funkce  $\mathbf{y} = \mathbf{g}(\mathbf{x})$ .
- Cílem je nalézt
  - $\bar{\mathbf{y}} = E[\mathbf{y}]$ ,
  - $\mathbf{P}_y = E[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})]$ ,
  - $\mathbf{P}_{xy} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})]$ .

## Vlastnosti lokálních metod

Lokální metody jsou založeny na myšlence využití techniky Kalmanova filtru i v oblasti nelineárních systémů. Aby bylo možné tuto myšlenku použít, je nutné popis nelineárního systému vhodným způsobem aproximovat.

## Výhody a nevýhody

- Za výhodu lze považovat analytické řešení rekurzivních vztahů.
- Naopak hlavní nevýhodu lze spatřovat v lokální platnosti získaných odhadů.

## Transformace náhodné veličiny

- Nechť je známa  $\bar{\mathbf{x}} = E[\mathbf{x}]$  a  $\mathbf{P}_x = \text{cov}[\mathbf{x}]$  náhodné veličiny  $\mathbf{x}$  a nechť je dána nelineární funkce  $\mathbf{y} = \mathbf{g}(\mathbf{x})$ .
- Cílem je nalézt
  - $\bar{\mathbf{y}} = E[\mathbf{y}]$ ,
  - $\mathbf{P}_y = E[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})]$ ,
  - $\mathbf{P}_{xy} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})]$ .

## Vlastnosti lokálních metod

Lokální metody jsou založeny na myšlence využití techniky Kalmanova filtru i v oblasti nelineárních systémů. Aby bylo možné tuto myšlenku použít, je nutné popis nelineárního systému vhodným způsobem aproximovat.

### Výhody a nevýhody

- Za výhodu lze považovat analytické řešení rekurzivních vztahů.
- Naopak hlavní nevýhodu lze spatřovat v lokální platnosti získaných odhadů.

### Transformace náhodné veličiny

- Nechť je známa  $\bar{\mathbf{x}} = E[\mathbf{x}]$  a  $\mathbf{P}_x = \text{cov}[\mathbf{x}]$  náhodné veličiny  $\mathbf{x}$  a nechť je dána nelineární funkce  $\mathbf{y} = \mathbf{g}(\mathbf{x})$ .
- Cílem je nalézt
  - $\bar{\mathbf{y}} = E[\mathbf{y}]$ ,
  - $\mathbf{P}_y = E[(\mathbf{y} - \bar{\mathbf{y}})(\mathbf{y} - \bar{\mathbf{y}})]$ ,
  - $\mathbf{P}_{xy} = E[(\mathbf{x} - \bar{\mathbf{x}})(\mathbf{y} - \bar{\mathbf{y}})]$ .

## Taylorův rozvoj prvního řádu (standardní přístup, 1970)

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) \approx \mathbf{g}(\bar{\mathbf{x}}) + \mathbf{G}(\mathbf{x} - \bar{\mathbf{x}}),$$

$$\text{kde } \mathbf{G} = \left. \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}}.$$

## Charakteristiky náhodné proměnné $\mathbf{y}$

- $\bar{\mathbf{y}}_A = \mathbf{g}(\bar{\mathbf{x}})$
- $\mathbf{P}_{y,A} = \mathbf{G}\mathbf{P}_x\mathbf{G}^T$
- $\mathbf{P}_{xy,A} = \mathbf{P}_x\mathbf{G}^T$

## Lokální filtr

Na linearizaci nelineárních funkcí v popisu systému za pomoci Taylorova rozvoje prvního řádu je založen např. **rozšířený Kalmanův filtr**.

## Stirlingova polynomiální interpolace prvního řádu (nový přístup, 2000)

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) \approx \mathbf{g}(\bar{\mathbf{x}}) + \frac{1}{2h} \left( \sum_{i=1}^{n_x} \Delta x_i \eta_i \right) \mathbf{g}(\bar{\mathbf{x}}),$$

kde  $\eta_i \mathbf{g}(\mathbf{x}) = \mathbf{g}(\bar{\mathbf{x}} + h\mathbf{s}_i) - \mathbf{g}(\bar{\mathbf{x}} - h\mathbf{s}_i)$ ,  $\mathbf{s}_i$  je  $i$ -tý sloupec  $\mathbf{S}_x$ ,  $\mathbf{P}_x = \mathbf{S}_x \mathbf{S}_x^T$ .

Charakteristiky náhodné proměnné  $\mathbf{y}$ 

- $\bar{\mathbf{y}}_A = \mathbf{g}(\bar{\mathbf{x}})$
- $\mathbf{P}_{y,A} = \frac{1}{4h^2} \sum_{i=1}^{n_x} (\mathbf{g}(\bar{\mathbf{x}} + h\mathbf{s}_i) - \mathbf{g}(\bar{\mathbf{x}} - h\mathbf{s}_i)) (\mathbf{g}(\bar{\mathbf{x}} + h\mathbf{s}_i) - \mathbf{g}(\bar{\mathbf{x}} - h\mathbf{s}_i))^T$
- $\mathbf{P}_{xy,A} = \frac{1}{2h} \sum_{i=1}^{n_x} \mathbf{s}_i (\mathbf{g}(\bar{\mathbf{x}} + h\mathbf{s}_i) - \mathbf{g}(\bar{\mathbf{x}} - h\mathbf{s}_i))$

## Lokální filtr

Na aproximaci nelineárních funkcí v popisu systému za pomoci Stirlingovy interpolace prvního řádu je založen **diferenční lokální filtr prvního řádu**.

## Transformace charakteristických bodů (nový přístup, 2000)

Náhodná veličina  $\mathbf{x}$  je aproximována množinou bodů

$$\mathcal{X}_0 = \bar{\mathbf{x}}, \mathcal{W}_0 = \frac{\kappa}{n_x + \kappa}$$

$$\mathcal{X}_i = \bar{\mathbf{x}} + \left( \sqrt{(n_x + \kappa) \mathbf{P}_x} \right)_i, i = 1, 2, \dots, n_x,$$

$$\mathcal{X}_j = \bar{\mathbf{x}} - \left( \sqrt{(n_x + \kappa) \mathbf{P}_x} \right)_{j-n_x}, j = (n_x + 1), (n_x + 2), \dots, 2n_x,$$

$$\text{kde } \mathcal{W}_i = \mathcal{W}_j = \frac{1}{2(n_x + \kappa)}, \forall i, j.$$

Množina bodů pak může být transformována skrze nezměněnou nelineární funkci

$$\mathcal{Y}_i = \mathbf{g}(\mathcal{X}_i), \forall i.$$

## Transformace charakteristických bodů

### Charakteristiky náhodné proměnné $\mathbf{y}$

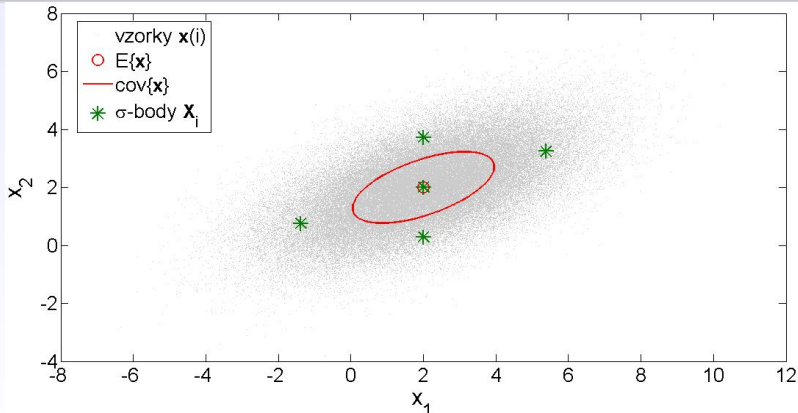
- $\bar{\mathbf{y}}_A = \sum_{i=0}^{2n_x} \mathcal{W}_i \mathcal{Y}_i$
- $\mathbf{P}_{y,A} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{Y}_i - \bar{\mathbf{y}}_A)(\mathcal{Y}_i - \bar{\mathbf{y}}_A)^T$
- $\mathbf{P}_{xy,A} = \sum_{i=0}^{2n_x} \mathcal{W}_i (\mathcal{X}_i - \bar{\mathbf{x}})(\mathcal{Y}_i - \bar{\mathbf{y}}_A)^T$

### Lokální filtr

Na aproximaci popisu náhodné veličiny množinou charakteristických bodů jsou založeny **transformační lokální filtry**.

## Transformace náhodné proměnné skrze nelineární funkci

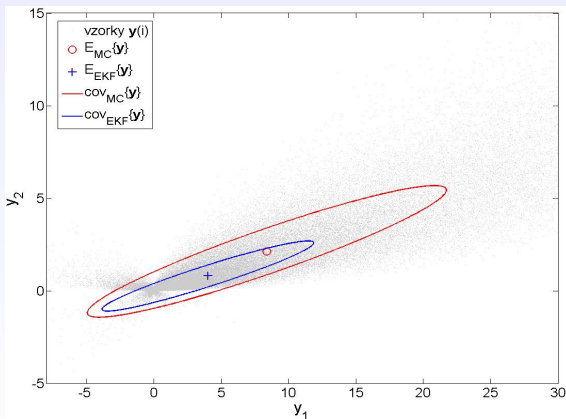
$$\text{Nechť } \mathbf{x} \sim \mathcal{N}\left\{\mathbf{x} : \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 3.8 & 1.4 \\ 1.4 & 1.5 \end{bmatrix}\right\} \text{ a } \mathbf{y} = \mathbf{g}(\mathbf{x}) = \begin{bmatrix} 0.5x_1x_2^2 \\ 0.1x_1^2x_2 \end{bmatrix}.$$





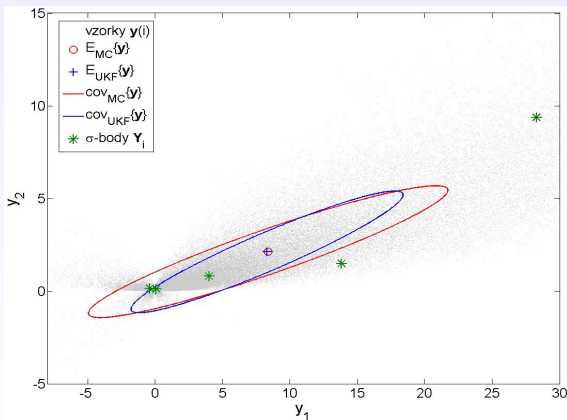
## Transformace náhodné proměnné skrze nelineární funkci

- Aproximace nelineární funkce  $g(\cdot)$  za pomoci Taylorova rozvoje 1. řádu.



## Transformace náhodné proměnné skrze nelineární funkci

- Aproximace popisu náhodné proměnné  $\mathbf{x}$  za pomoci charakteristických bodů.



## Víceřadová predikce a vyhlazování

- Návrh metod víceřadové predikce (tj. pro  $p(\mathbf{x}_l | \mathbf{z}^k)$ ,  $l > k$ ) založených na Stirlingově polynomiální interpolaci a transformaci charakteristických bodů.
- Návrh metod vyhlazování ( $p(\mathbf{x}_l | \mathbf{z}^k)$ ,  $l < k$ ) založených na Stirlingově polynomiální interpolaci a transformaci charakteristických bodů, kdy
  - čas. okamžik  $l$  je fixován a čas. okamžik  $k$  se mění,  
$$\bar{\mathbf{x}}_k = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{x}_0 \end{bmatrix}; \text{ p. p. estim. } \hat{\mathbf{x}}'_0 = \begin{bmatrix} \hat{\mathbf{x}}'_0 \\ \hat{\mathbf{x}}'_0 \end{bmatrix}, \bar{\mathbf{P}}'_0 = \begin{bmatrix} \mathbf{P}'_0 & \mathbf{P}'_0 \\ \mathbf{P}'_0 & \mathbf{P}'_0 \end{bmatrix},$$
  - fixován je jen rozdíl časových okamžiků  $k - l$ ,  
navrhovaný algoritmus vychází z Rauch-Tung-Striebelova vyhlazovače, používaného v oblasti lineárních gaussovských systémů,
  - čas. okamžik  $k$  je fixován a cílem je vyhladit odhady ve všech předchozích časových okamžicích.

## Využití estimačních metod v dopravních úlohách

- Spolupráce na řešení dopravní problematiky v rámci výzkumného centra DAR.
- Cílem je optimalizovat provoz na dané křižovatce. Pro řízení provozu je potřebné znát aktuální délky kolon vozidel v každém rameni. Ty nejsou přímo měřitelné a je nutné je odhadovat.
- Měřená data pochází z dopravních detektorů umístěných pod povrchem vozovky, jenž udávají intenzitu provozu v daném rameni.