Kalmánův filtr – LETKF (Local Ensemble Transform Kalman Filter)

Úloha:

Máme diferenciální rovnici:

$$\frac{d\mathbf{x}}{dt} = \mathbf{F}(t, \mathbf{x}) \qquad (1)$$

a naměřené hodnoty y_j^o v časech tj , naměřené hodnoty obsahují chyby.

Úkolem je nalézt „optimální-nejlepší" trajektorii $x_a(t)$ splňující rovnici (1).

Co to je „optimální-nejlepší" trajektorie ?

Matematická formulace problému:

- Máme modelovou rovnici (model je přesný):

$$\frac{d\mathbf{x}}{dt} = F(t, \mathbf{x})$$

- V časech tj , j=1, .. , n máme následující informace:

$\mathbf{y}_j^O$ .. naměřené hodnoty

$H_j, \mathbf{R}_j$ .. popisují vztah mezi $\mathbf{y}_j^O$ a $\mathbf{x}(t_j)$

$$\mathbf{y}_j^O = H_j(\mathbf{x}(t_j)) + \boldsymbol{\varepsilon}_j$$

$\boldsymbol{\varepsilon}_j$ je chyba s Gaussovým rozdělením s průměrem **0**

a kovariancí $\mathbf{R}_j = \varepsilon_j (\varepsilon_j)^T$

$t_1 < t_2 < \dots < t_n$

Pravděpodobnost trajektorie :

$$\prod_{j=1}^{n} \exp\left(-\frac{1}{2}[\mathbf{y}_j^O - H_j(\mathbf{x}(t_j))]^T \mathbf{R}_j^{-1}[\mathbf{y}_j^O - H_j(\mathbf{x}(t_j))]\right)$$

Maximalizace pravděpodobnosti vede k minimalizace funkcionálu:

$$\mathbf{J}(\{\mathbf{x}(t_j)\}) = \sum_{j=1}^{n} [\mathbf{y}_j^O - H_j(\mathbf{x}(t_j))]^T \mathbf{R}_j^{-1}[\mathbf{y}_j^O - H_j(\mathbf{x}(t_j))]$$

$$\{\mathbf{x}(t_j)\} = [\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_n)]$$

kde $M_{t,t_j}(\mathbf{x}) : \mathbf{x}(t) \to \mathbf{x}(t_j)$.

Označme ,

pak $$J_t(\mathbf{x}) = \sum_{j=1}^{n} [\mathbf{y}_j^O - H_j M_{t,t_j}(\mathbf{x})]^T \mathbf{R}_j^{-1}[\mathbf{y}_j^O - H_j M_{t,t_j}(\mathbf{x})]$$

kde $\mathbf{x}$ je „počáteční“ podmínka v čase t

Lineární případ:

$$\mathbf{H}_j\mathbf{x} = \mathbf{H}_j(\mathbf{x}(t_j))$$

$$\mathbf{M}_{t,t'}\mathbf{x} = \mathbf{M}_{t,t'}(\mathbf{x})$$

Pokud je model diskrétní ve tvaru

$$\mathbf{x}_n = \mathbf{M}\mathbf{x}_{n-1}$$

pak

$$M_{k,m}(\mathbf{x}) = \mathbf{M}^{m-k}\mathbf{x} \quad .$$

Pro výpočet **x**(tn-1) se minimalizuje:

$$J_{t_{n-1}}(\mathbf{x}) = \sum_{j=1}^{n-1} [\mathbf{y}_j^O - \mathbf{H}_j\mathbf{M}_{t_{n=1},t_j}\mathbf{x}]^T \mathbf{R}_j^{-1} [\mathbf{y}_j^O - \mathbf{H}_j\mathbf{M}_{t_{n=1},t_j}\mathbf{x}] \quad . \qquad \textbf{(3)}$$

Předpokládejme, že v čase tn-1 je analýza $\bar{\mathbf{x}}^a_{n-1}$ a tomu odpovídá kovarianční matice $\mathbf{P}^a_{n-1}$.

Z hlediska pravděpodobnosti je

$\bar{\mathbf{x}}^a_{n-1}$ - průměr

$\mathbf{P}^a_{n-1}$ - kovariance pro Gaussovo rozdělení pravděpodobnosti

za předpokladu, že máme naměřené hodnoty v t1, … , tn-1 .

Lze tedy psát:

$$J_{t_{n-1}}(\mathbf{x}) = \sum_{j=1}^{n-1} [\mathbf{y}_j^O - \mathbf{H}_j\mathbf{M}_{t_{n=1},t_j}\mathbf{x}]^T \mathbf{R}_j^{-1} [\mathbf{y}_j^O - \mathbf{H}_j\mathbf{M}_{t_{n=1},t_j}\mathbf{x}] = [\mathbf{x} - \bar{\mathbf{x}}^a_{n-1}](P^a_{n-1})^{-1}[\mathbf{x} - \bar{\mathbf{x}}^a_{n-1}] + c \qquad \textbf{(4)}$$

Máme $\bar{\mathbf{x}}^a_{n-1}$ a P^a_{n-1}, chceme vypočítat $\bar{\mathbf{x}}^a_n$ a P^a_n.

Kalmánův filtr:

$$\bar{\mathbf{x}}^b_n = \mathbf{M}_{t_{n-1},t_n} \bar{\mathbf{x}}^a_{n-1} \qquad \textbf{(5)}$$

$$P^b_n = \mathbf{M}_{t_{n-1},t_n} P^a_{n-1} \mathbf{M}^T_{t_{n-1},t_n} \qquad \textbf{(6)}$$

1. krok - předpověď

$$J_{t_{n-1}} = \sum_{j=1}^{n-1} [\mathbf{y}^O_j - \mathbf{H}_j \mathbf{M}_{t_{n-1},t_j} \mathbf{x}_{t_{n-1}}]^T \mathbf{R}_j^{-1} [\mathbf{y}^O_j - \mathbf{H}_j \mathbf{M}_{t_{n-1},t_j} \mathbf{x}_{t_{n-1}}] =$$

$$= \sum_{j=1}^{n-1} [\mathbf{y}^O_j - \mathbf{H}_j \mathbf{M}_{t_{n-1},t_j} \mathbf{M}_{t_n,t_{n-1}} \mathbf{x}_{t_n}]^T \mathbf{R}_j^{-1} [\mathbf{y}^O_j - \mathbf{H}_j \mathbf{M}_{t_{n-1},t_j} \mathbf{M}_{t_n,t_{n-1}} \mathbf{x}_{t_n}] =$$

$$= \sum_{j=1}^{n-1} [\mathbf{y}^O_j - \mathbf{H}_j \mathbf{M}_{t_n,t_j} \mathbf{x}_{t_n}]^T \mathbf{R}_j^{-1} [\mathbf{y}^O_j - \mathbf{H}_j \mathbf{M}_{t_n,t_j} \mathbf{x}_{t_n}] =$$

$$= [x - \bar{\mathbf{x}}^b_n](P^b_n)^{-1}[x - \bar{\mathbf{x}}^b_n] + c$$

Pro x v čase tn :

$$J_{t_n}(\mathbf{x}) = [\mathbf{x} - \bar{\mathbf{x}}_n^b](\mathbf{P}_n^b)^{-1}[\mathbf{x} - \bar{\mathbf{x}}_n^b] + [\mathbf{y}_n^O - \mathbf{H}_n\mathbf{x}]^T\mathbf{R}_n^{-1}[\mathbf{y}_n^O - \mathbf{H}_n\mathbf{x}] + c \qquad (7)$$

a zároveň J vyjádříme ve tvaru:

$$J_{t_n}(\mathbf{x}) = [x - \bar{\mathbf{x}}_n^a](\mathbf{P}_n^a)^{-1}[x - \bar{\mathbf{x}}_n^a] + c'$$

Porovnáním mocnin u x dostaneme:

$$\mathbf{P}_n^a = [(\mathbf{P}_n^b)^{-1} + \mathbf{H}_n^T\mathbf{R}_n^{-1}\mathrm{H}_n]^{-1} \qquad (8)$$

$$\bar{\mathbf{x}}_n^a = \mathbf{P}_n^a[(\mathbf{P}_n^b)^{-1}\bar{\mathbf{x}}_n^b + \mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{y}_n^O] \qquad (9)$$

Rovnice (8) a (9) lze upravit:

$$\bar{\mathbf{x}}_n^a = \bar{\mathbf{x}}_n^b + \mathbf{P}_n^a\mathbf{H}_n^T\mathbf{R}_n^{-1}(\mathbf{y}_n^O - \mathbf{H}_n\bar{\mathbf{x}}_n^b) \qquad (10)$$

$$\mathbf{P}_n^a = (\mathbf{I} + \mathbf{P}_n^b\mathbf{H}_n^T\mathbf{R}_n^{-1}\mathbf{H}_n)^{-1}\mathbf{P}_n^b \qquad (11)$$

Shrnutí:

$$\left.\begin{aligned} \bar{\mathbf{x}}^b_n &= \mathbf{M}_{t_{n-1},t_n} \bar{\mathbf{x}}^a_{n-1} \\ \mathrm{P}^b_n &= \mathbf{M}_{t_{n-1},t_n} \mathbf{P}^a_{n-1} \mathbf{M}^T_{t_{n-1},t_n} \end{aligned}\right\} \textbf{1. Krok - předpověď}$$

$$\left.\begin{aligned} \bar{\mathbf{x}}^a_n &= \bar{\mathbf{x}}^b_n + \mathbf{P}^a_n \mathbf{H}^T_n \mathbf{R}^{-1}_n (\mathbf{y}^O_n - \mathbf{H}_n \bar{\mathbf{x}}^b_n) \\ \mathbf{P}^a_n &= (\mathbf{I} + \mathbf{P}^b_n \mathbf{H}^T_n \mathbf{R}^{-1}_n \mathbf{H}_n)^{-1} \mathbf{P}^b_n \end{aligned}\right\} \textbf{2. Krok - analýza}$$

Počáteční podmínky:

1. Nekonečná chyba

$P_1^b = \infty \quad \longrightarrow \quad (P_1^b) = 0$

To je rozumné, pokud je k dispozici hodně naměřených dat v čase t1 , aby analýza využívající pouze tato data byla dostatečně přesná.

3. Velká chyba

$P_1^b \ll \infty$

Označení – dimenze

m	… dimenze modelu , tj. dimenze x
l	… dimenze měření
k	… velikost ansámblu

Nelineární případ: M a H jsou nelineární

V čase tn-1 máme ansámbl počátečních podmínek stavového vektoru x v m dimensionálním prostoru

$$\left\{ x_{n-1}^{a(i)},\ i=1,\dots,k \right\}$$

Aplikujeme model a dostaneme k předpovědí, které se použijí jako předběžná pole

$$x_n^{b(i)} = \mathbf{M}_{t_{n-1},t_n}\left(x_{n-1}^{a(i)}\right)$$

a vypočteme výběrový průměr a kovarianci:

$$\overline{x}^{b} = \frac{1}{k}\sum_{i=1}^{k} x^{b(i)}$$

$$\mathbf{P}^b = (k-1)^{-1}\sum_{i=1}^{k}\left(\mathbf{x}^{b(i)} - \overline{\mathbf{x}}^{b}\right)\left(\mathbf{x}^{b(i)} - \overline{\mathbf{x}}^{b}\right)^T = (k-1)^{-1}\,\mathbf{X}^b\left(\mathbf{X}^b\right)^T$$

$$\mathbf{X}^b = \left(\mathbf{x}^{b(1)} - \overline{\mathbf{x}}^{b},\dots,\mathbf{x}^{b(k)} - \overline{\mathbf{x}}^{b}\right),\quad \dim\left(\mathbf{X}^b\right) = m\times k$$

Pro hodnosti matic platí

$$\operatorname{rank}\left(\mathbf{P}^{b}\right) = \operatorname{rank}\left(\mathbf{X}^{b}\right), \quad \dim\left(\mathbf{X}^{b}\right) \leq k - 1$$

protože

$$\sum_{i=1}^{k} \mathbf{X}^{b}_{i} = \mathbf{0}$$

Je třeba určit nejen $\overline{\mathbf{x}}^{a}$ a $\mathbf{P}^{a}$, ale i $\left\{ \mathbf{x}^{a(i)}_{n-1},\ i = 1,\ldots,k \right\}$

s průměrem a kovariancí:

$$\overline{\mathbf{x}}^{a} = \frac{1}{k}\sum_{i=1}^{k} \mathbf{x}^{a(i)}$$

$$\mathbf{P}^{a} = (k-1)^{-1} \sum_{i=1}^{k} (\mathbf{x}^{a(i)} - \overline{\mathbf{x}}^{a})(\mathbf{x}^{a(i)} - \overline{\mathbf{x}}^{a})^{T} = (k-1)^{-1} \mathbf{X}^{a}\left(\mathbf{X}^{a}\right)^{T}$$

$$\mathbf{X}^{a} = \left(\mathbf{x}^{a(1)} - \overline{\mathbf{x}}^{a}, \ldots, \mathbf{x}^{a(k)} - \overline{\mathbf{x}}^{a}\right), \quad \dim(\mathbf{X}^{a}) = m \times k$$

Jak vybrat $\left\{ \mathbf{x}^{a(i)}_{n-1},\ i=1,\dots,k \right\}$ na základě $\left\{ \mathbf{x}^{b(i)}_{n-1},\ i=1,\dots,k \right\}$?

Předpokládáme: $k < m,\ k < l$

Hledáme x minimalizující J:

$$J(\mathbf{x}) = (\mathbf{x} - \overline{\mathbf{x}}^{b})^{T} (\mathbf{P}^{b})^{-1} (\mathbf{x} - \overline{\mathbf{x}}^{b}) + (\mathbf{y}^{O} - H(\mathbf{x}))^{T} \mathbf{R}^{-1} (\mathbf{y}^{O} - H(\mathbf{x}))$$

$$\dim\left(\mathbf{P}^{b}\right) = m \times m, \quad h\left(\mathbf{P}^{b}\right) = k - 1$$

Pb je jednoznačné zobrazení $\left\{ \mathbf{x}^{b(i)}_{n-1},\ i=1,\dots,k \right\}$ na $\left\{ \mathbf{x}^{b(i)}_{n-1},\ i=1,\dots,k \right\}$,

protože $\mathbf{P}^{b} = (k-1)^{-1} \sum_{i=1}^{k} (\mathbf{x}^{b(i)} - \overline{\mathbf{x}}^{b})(\mathbf{x}^{b(i)} - \overline{\mathbf{x}}^{b})^{T} = (k-1)^{-1} \mathbf{X}^{b} \left(\mathbf{X}^{b}\right)^{T}$.

Definujeme se lineární prostor S,

$$S = \text{lin. prostor}\left\{\mathbf{x}^{b(i)}_{n-1},\ i = 1,\dots,k\right\}$$

ve kterém potřebujeme najít řešení (analyzovanou hodnotu).

Řešení je lineární kombinací sloupců matice $\mathbf{X}^b$.

Řešíme v prostoru $\tilde{S}$:

$$\mathbf{X}^b : \tilde{S} \to S \qquad \text{lineární zobrazení na}$$

$$w \in R^k \to \mathbf{X}^b w$$

$$x = \bar{x}^b + \mathbf{X}^b w$$

x je perturbovaná modelová stavová hodnota,

w má Gaussovo rozdělení s μ=0 a Pb=(k-1)-1 **I**, pak

x má Gaussovo rozdělení s μ=0 a Pb=(k-1)-1 **X**b(**X**b)T

$$J(\mathbf{x}) = (\mathbf{x} - \overline{\mathbf{x}}^b)^T (\mathbf{P}^b)^{-1} (\mathbf{x} - \overline{\mathbf{x}}^b) + (\mathbf{y}^O - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y}^O - H(\mathbf{x})) \qquad \textbf{(16)}$$

$$\tilde{J}(\mathbf{w}) = (k-1)\mathbf{w}^T\mathbf{w} + (\mathbf{y}^O - H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}))^T \mathbf{R}^{-1} (\mathbf{y}^O - H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w})) \qquad \textbf{(15)}$$

Jestliže $\overline{\mathbf{w}}^a$ minimalizuje $\tilde{J}$, pak $\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{X}^b\overline{\mathbf{w}}^a$ minimalizuje J.

Minimalizujeme:

$$\tilde{J}(\mathbf{w}) = (k-1)\mathbf{w}^T\mathbf{w} + (\mathbf{y}^O - H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}))^T \mathbf{R}^{-1} (\mathbf{y}^O - H(\overline{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}))$$

H je nelineární - co s tím?

Linearizace H např.:

Potřebujeme umět spočítat $H_j(\bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w})$ pro $\mathbf{w} \in \tilde{S}$

Postupujeme takto:

$$y^{b(i)} = H(\mathbf{x}^{b(i)})$$

$$\bar{y}^b = \frac{1}{k}\sum_{i=1}^{k} y^{b(i)} \quad , \quad \mathbf{Y}^b = \left(\mathbf{y}^{b(1)} - \bar{\mathbf{y}}^b, \dots, \mathbf{y}^{b(k)} - \bar{\mathbf{y}}^b\right)$$

$$H(\bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}) \approx \bar{\mathbf{y}}^b + \mathbf{Y}^b\mathbf{w}$$ ... linearizace

Původní funkcionál:

$$\tilde{J}(\mathbf{w}) = (k-1)\mathbf{w}^T\mathbf{w} + (\mathbf{y}^O - H(\bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}))^T\mathbf{R}^{-1}(\mathbf{y}^O - H(\bar{\mathbf{x}}^b + \mathbf{X}^b\mathbf{w}))$$

Po linearizaci:

$$\tilde{J}(\mathbf{w}) = (k-1)\mathbf{w}^T\mathbf{w} + (\mathbf{y}^O - \bar{\mathbf{y}}^b - \mathbf{Y}^b\mathbf{w})^T\mathbf{R}^{-1}(\mathbf{y}^O - \bar{\mathbf{y}}^b - \mathbf{Y}^b\mathbf{w})$$

Obdobně jako předtím (vztah 7) dostaneme:

$$\overline{\mathbf{w}}^a = \tilde{\mathbf{P}}^a (\mathbf{Y}^b)^T \mathbf{R}^{-1} (\mathbf{y}^O - \overline{\mathbf{y}}^b)$$

$$\tilde{\mathbf{P}}^a = ((k-1)\mathbf{I} + (\mathbf{Y}^b)^T \mathbf{R}^{-1} \mathbf{Y}^b)^{-1}$$

a v modelovém prostoru:

$$\overline{\mathbf{x}}^a = \overline{\mathbf{x}}^b + \mathbf{X}^b \overline{\mathbf{w}}^a$$

$$\mathbf{P}^a = \mathbf{X}^b \tilde{\mathbf{P}}^a (\mathbf{X}^b)^T$$

Potřebujeme vybrat ansámbl xa(i) s vlastnostmi $\overline{\mathbf{x}}^a, \mathbf{P}^a$.

Jak na to?

Vytvoříme ansámbl následujícím způsobem:

$$\mathbf{X}^a = \mathbf{X}^b \mathbf{W}^a \quad ,$$

kde

$$\mathbf{W}^a = [(k-1)\tilde{\mathbf{P}}^a]^{\frac{1}{2}} \qquad \text{tj.} \quad \tilde{\mathbf{P}}^a = (k-1)^{-1}\mathbf{W}^a(\mathbf{W}^a)^T$$

Vlastnosti „odmocniny" symetrické matice $\tilde{\mathbf{P}}^a$:

- zajišťuje, že suma sloupců matice $\mathbf{X}^a$ je nula
- matice $\mathbf{W}^a$ spojitě závisí na $\tilde{\mathbf{P}}^a$, což umožňuje lokální aplikaci algoritmu
- matice $\mathbf{W}^a$ minimalizuje průměrný čtverec vzdálenosti mezi $\mathbf{W}^a$ a identickou maticí

Algoritmus – lokální implementace

$m_{[g]}$ … dimenze modelu , tj. dimenze x

$l_{[g]}$ … dimenze měření

k … velikost ansámblu

$\{\mathbf{x}^{b(i)}_{[g]}, i=1,\ldots,k\}$ … ansámbl předběžných hodnot

$H_{[g]} \ (m_{[g]} \rightarrow \ l_{[g]})$ … operátor převádějící modelové hodnoty na stanice

$\mathbf{y}^{O}_{[g]}$ … vektor naměřených hodnot

$\mathbf{R}_{[g]} \ (l_{[g]} \times l_{[g]})$ … kovarianční matice naměřených hodnot

Krok 1:

$$\mathbf{y}^{b(i)}_{[g]} = H_{[g]}\left(\mathbf{x}^{b(i)}_{[g]}\right) \qquad \text{pro } i = 1,\ldots,k$$

$$\bar{\mathbf{y}}^{b}_{[g]} = \frac{1}{k}\sum_{i=1}^{k} \mathbf{y}^{b(i)}_{[g]}$$

$$\mathbf{Y}^{b}_{[g]} = \left(\mathbf{y}^{b(1)}_{[g]} - \bar{\mathbf{y}}^{b}_{[g]}, \ldots, \mathbf{y}^{b(k)}_{[g]} - \bar{\mathbf{y}}^{b}_{[g]}\right)^{T}$$

Krok 2:

$$\bar{\mathbf{x}}^{b}_{[g]} = \frac{1}{k}\sum_{i=1}^{k} \mathbf{x}^{b(i)}_{[g]}$$

$$\mathbf{X}^{b}_{[g]} = \left(\mathbf{x}^{b(1)}_{[g]} - \bar{\mathbf{x}}^{b}_{[g]}, \ldots, \mathbf{x}^{b(k)}_{[g]} - \bar{\mathbf{x}}^{b}_{[g]}\right)^{T}$$

Krok 3:

Pro uzlový bod se vyberou „odpovídající" řádky z matic-vektorů

$$\bar{\mathbf{x}}^{b}_{[g]},\ \mathbf{X}^{b}_{[g]},\ \bar{\mathbf{y}}^{b}_{[g]},\ \mathbf{Y}^{b}_{[g]}$$

Odpovídající znamená blízké k uzlovému bodu.
Vybere se m řádků, což určuje rozměry matic-vektorů:

$$\bar{\mathbf{x}}^{b},\ \mathbf{X}^{b},\ \bar{\mathbf{y}}^{b},\ \mathbf{Y}^{b}$$

Krok 4:

Vypočte se:

$$\mathbf{C} = (\mathbf{Y}^{b})^{T}\mathbf{R}^{-1} \qquad (k \times l)$$

$$\mathbf{R}\mathbf{C}^{T} = \mathbf{Y}^{b}$$

Krok 5:

Vypočte se:

$$\widetilde{\mathbf{P}}^{a} = \left[(k-1)\mathbf{I}/\rho + \mathbf{C}\mathbf{Y}^{b}\right]^{-1} \qquad (k \times k)$$

ρ uměle navyšuje kovarianci na diagonále – zvyšuje se vliv naměřených hodnot

Krok 6:

Vypočte se:

$$\mathbf{W}^{a} = [(k-1)\widetilde{\mathbf{P}}^{a}]^{\frac{1}{2}} \qquad \text{tj.} \quad \widetilde{\mathbf{P}}^{a} = (k-1)^{-1}\mathbf{W}^{a}(\mathbf{W}^{a})^{T}$$

Krok 7:

Vypočte se:

$$\overline{\mathbf{w}}^{a} = \tilde{\mathbf{P}}^{a}\mathbf{C}(\mathbf{y}^{O} - \overline{\mathbf{y}}) \qquad (k \times 1)$$

a přičte se ke každému sloupci matice $\mathbf{W}$.

Krok 8:

Vypočte se:

$$\mathbf{x}^{a(i)} = \overline{\mathbf{x}}^{b} + \mathbf{X}^{b}\mathbf{w}^{a(i)} \qquad \text{pro } i = 1,\dots,k$$

Krok 9:

Výsledky pro jednotlivé uzly se složí dohromady a dostane se ansámbl analyzovaných hodnot

$$\left\{\mathbf{x}^{a(i)}_{[g]}\right\} \qquad \text{pro } i = 1,\dots,k$$