Vzorkovací schémata v částicovém filtru

David Coufal
KPMS MFF UK, ÚI AV ČR
david.coufal@cs.cas.cz

ROBUST 2016
Kurzovní, Jeseníky, 13. září 2016

Agenda

- filtrační úloha
- částicový filtr
- vzorkovací schémata
- vlastnosti vzorkovacích schémat

Signál

- nechť $\{X_t, t = 0, 1, \dots\}$ je Markovský řetězec se spojitou množinou stavů $X_t \in \mathbb{R}^{n_x}$, $n_x \in \mathbb{N}$.

- pravděpodobnostní chování MŘ je dáno specifikací počátečního rozdělení X_0 pomocí příslušné hustoty p_0

- množinou přechodových jader $K_{t-1}(dx_t|x_{t-1})$
  s hustotami $K_{t-1}(x_t|x_{t-1})$ pro $t = 1, \dots T$.

Pozorování

- observační proces $\{Y_t, t = 1, 2 \ldots\}$, $Y_t \in \mathbb{R}^{n_y}$
  určen pomocí funkce h_t na základě hodnot signálu
  a dále se předpokládá zašumění aditivním šumem
  formálně je observační proces definován:

$$Y_t = h_t(X_t) + V_t, \; t > 0$$

V_t je šum, tj. náhodná veličina o známé hustotě g_t.

Úloha filtrace

- evoluce signálu a observačního procesu:

$$\begin{array}{ccccccccc} X_0 & \to & X_1 & \to & \ldots & \to & X_t & \to & \ldots \\ & & \downarrow & & \downarrow & & \downarrow & & \\ & & y_1 & & \ldots & & y_t & & \end{array}$$

- filtrační úloha spočívá ve stanovení podmíněného rozdělení $P_{X_t|Y_{1:t}}$ aktuálního stavu signálu X_t na základě pozorováné historie $Y_1 = y_1, \ldots, Y_t = y_t$

- speciálně pouze stanovení momentů tohoto podmíněného rozdělení $\int f(X_t)\, dP_{X_t|Y_{1:t}}$

1D - Gaussovský lineární proces

- 1D - Gaussovský lineární signál/observační proces

$$X_t = aX_{t-1} + b + cW_t, \quad Y_t = hX_t + gV_t, \quad t \geq 1.$$

- a, b, c, h, g reálné parametry, $c, g > 0$

- $X_0, W_1, V_1, W_2, V_2, \ldots$ 1D - normálně rozdělené n.v.

- $X_0 \sim \mathcal{N}(\mu_0, \sigma_0^2)$, ostatní i.i.d. $\sim \mathcal{N}(0, 1)$

- $\{X_t, t \geq 0\}$ Gaussovský Markovský řetězec

1D - Gaussovský proces, filtrační úloha

- filtrační rozdělení - podmíněná rozdělení: $P_{X_t|Y_{1:t}}$
- rozdělení jsou normální, tj. určená odpovídajícími momenty (int. charakteristikami):

$$
\begin{aligned}
\hat{\mu}_t &= \mathbb{E}[X_t \,|\, Y_{1:t}] \\
\hat{\sigma}_t^2 &= \mathbb{E}[(X_t - \mu_t)^2 \,|\, Y_{1:t}]
\end{aligned}
$$

- analytické řešení ve formě rekurenčních rovnic

1D - Kalmánův filtr

- 1D - Kalmánův filtr

$$
\begin{aligned}
\hat{\mu}_t &= \frac{g^2}{v^2} b + \frac{c^2}{v^2} h y_t + \frac{g^2}{v^2} a \frac{\hat{\sigma}^2_{t-1} a h (y_t - b h) + v^2 \hat{\mu}_{t-1}}{h^2 \hat{\sigma}^2_{t-1} a^2 + v^2}, \\
\hat{\sigma}^2_t &= \frac{g^4}{v^2} a^2 \frac{\hat{\sigma}^2_{t-1}}{h^2 \hat{\sigma}^2_{t-1} a^2 + v^2} + \frac{g^2 c^2}{v^2}
\end{aligned}
$$

pro $v^2 = c^2 h^2 + g^2$, $\hat{\mu}_0 = \mu_0$, $\hat{\sigma}_0 = \sigma_0$

Filtrační rozdělení

- sdružené rozdělení signálu a pozorování má tvar

$$p(x_{0:t}, y_{1:t}) = p_0(x_0) \prod_{k=1}^{t} g_k(y_k|x_k) K_{k-1}(x_k|x_{k-1}).$$

kde $g_k(y_k|x_k) = g_k(y_t - h_t(x_t))$.

- filtrační hustota - odpovídá podmíněné hustotou $p(x_{0:t}|y_{1:t})$

$$p(x_t|y_{1:t}) = \frac{\int p(x_{0:t}, y_{1:t})\, dx_{0:t-1}}{\int p(x_{0:t}, y_{1:t})\, dx_{0:t}}$$

- filtrační rovnice

$$\begin{aligned} p(x_t|y_{1:t-1}) &= \int K_{t-1}(x_t|x_{t-1})\, p(x_{t-1}|y_{1:t-1})\, dx_{t-1}, \\ p(x_t|y_{1:t}) &= \frac{g_t(y_t|x_t)\, p(x_t|y_{1:t-1})}{\int g_t(y_t|x_t)\, p(x_t|y_{1:t-1})\, dx_t}, \quad t \in \mathbb{N} \end{aligned}$$

Částicový filtr

- sekvenčně generuje empirické míry approximující filtrační rozdělení

$$\widehat{\pi}_t^n(dx) \approx P_{X_t|Y_{1:t}} = p(x_t|y_{1:t})\,dx_t$$

- aproximující empirická míra má tvar váženého součtu Diracových měr založeném na nosiči $\{w(\overline{x}_t^i), \overline{x}_t^i\}_{i=1}^n$

$$\widehat{\pi}_t^n(dx_t) = \sum_{i=1}^n w(\overline{x}_t^i)\,\delta_{\overline{x}_t^i}(dx_t), \quad \sum_{i=1}^n w(\overline{x}_t^i) = 1$$

- resampling: uniformě vážená verze empirické míry založená na nosiči $\{1/n, x_t^i\}_{i=1}^n$, který vzniká z $\{w(\overline{x}_t^i), \overline{x}_t^i\}_{i=1}^n$

$$\pi_t^n(dx_t) = \frac{1}{n}\sum_{i=1}^n \delta_{x_t^i}(dx_t)$$

Částicový filtr - algoritmus

- **0. deklarace**:
  n - počet vzorků,
  T - výpočetní horizont,
  $p_0(x_0)$ - hustota počátečního rozdělení $\pi_0 \sim X_0$,
  $K_{t-1}(x_t|x_{t-1})$, $t = 1, \dots T$, - podmíněné přechodové hustoty

- **1. inicializace**:
  $t = 0$,
  vzorkování $\{\overline{x}_0^i \sim p_0(dx_0)\}_{i=1}^n$,
  sestavení emp. míry $\widehat{\pi}_0^n(dx_0) = \frac{1}{n}\sum_{i=1}^n \delta_{\overline{x}_0^i}(dx_0)$,
  přiřazení $\pi_0^n(dx_0) = \widehat{\pi}_0^n(dx_0)$, i.e., $\{x_0^i = \overline{x}_0^i\}_{i=1}^n$.

- **2. importance sampling**:
  $t = t + 1$,
  vzorkování $\{\overline{x}_t^i \sim K_{t-1}(dx_t | x_{t-1}^i)\}_{i=1}^n$,
  pro $i = 1{:}n$ výpočet vah

$$\widetilde{w}(\overline{x}_t^i) = \frac{g_t(y_t - h_t(\overline{x}_t^i))}{\sum_i g_t(y_t - h_t(\overline{x}_t^i))},$$

  sestavení emp. míry $\widehat{\pi}_t^n(dx_t) = \sum_{i=1}^n \widetilde{w}(\overline{x}_t^i)\, \delta_{\overline{x}_t^i}(dx_t)$.

- **3. resampling**
  převzorkování $\{1/n, x_t^i\}_{i=1}^n$ z $\{w(\overline{x}_t^i), \overline{x}_t^i\}_{i=1}^n$
  a sestavení emp. míry $\pi_t^n(dx_t) = \frac{1}{n} \sum_{i=1}^n \delta_{x_t^i}(dx_t)$.

- **4.** pokud $t = T$ konec, jinak návrat do kroku 2.

Konvergence částicového filtru

Věta. Nechť π_t^n, $t \geq 1$ je posloupnost empirických měr generovaných částicovým filtrem a π_t, $t \geq 0$ je posloupnost příslušných teoretických distribucí. Pak, pro všechna $t \geq 1$ a $f \in \mathcal{C}_b$ platí

$$\mathbb{E}[|\pi_t^n f - \pi_t f|^2] \leq \frac{c_t^2 ||f||_\infty^2}{n}.$$

Důkaz. Doucet et. al. *SMC Methods in Practice*, Springer 2001.

Multinomiální resampling

- výchozí stav $\{w_i, \overline{x}_i\}_{i=1}^n$, odpovídá empirické míře $\widehat{\pi}_t^n(dx)$
- výběr z vracením z $\{\overline{x}_i\}_{i=1}^n$ s pravděpodobnostmi $\{p_i = w_i\}$ získáme nový soubor $\{x_i\}_{i=1}^n$, konstruujeme

$$\pi_t^n(dx) = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}(dx)$$

- hodnoty duplikací $(N_1, \ldots, N_n) \sim \mathcal{M}(n, w_1, \ldots, w_n)$
- vlastnosti pro $f \in \mathcal{C}_b$

$$\mathbb{E}\left[\int f(x)\, \pi_t^n(dx)\right] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n f(x_i)\right] = \sum_{i=1}^n w_i f(x_i) = \int f(x)\, \widehat{\pi}_t^n(dx)$$

$$Var\left[\frac{1}{n}\sum_{i=1}^n f(x_i)\right] = \frac{1}{n}\left\{\sum_{i=1}^n w_i f^2(x_i) - \left(\sum_{i=1}^n w_i f(x_i)\right)^2\right\}$$

Residuální resampling

- výchozí stav $\{w_i, \overline{x}_i\}_{i=1}^n$, odpovídá empirické míře $\widehat{\pi}_t^n(dx)$

  1. $R = \sum_{i=1}^n \lfloor n w_i \rfloor$
  2. $\overline{w}_i = \frac{n w_i - \lfloor n w_i \rfloor}{n - R}$
  3. $(\overline{N}_1, \ldots, \overline{N}_n) \sim \mathcal{M}(n - R, \overline{w}_1, \ldots, \overline{w}_n)$
  4. $N_i = \lfloor n w_i \rfloor + \overline{N}_i \Rightarrow \{x_i\}_{i=1}^n$

- vlastnosti pro $f \in \mathcal{C}_b$

$$
\begin{aligned}
\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n f(x_i)\right] &= \int f(x)\,\widehat{\pi}_t^n(dx) \\
Var\left[\frac{1}{n}\sum_{i=1}^n f(x_i)\right] &= \frac{1}{n}\left\{\sum_{i=1}^n w_i f^2(x_i) - \left(\sum_{i=1}^n \frac{\lfloor n w_i \rfloor}{n} f^2(x_i) - \frac{n-R}{n}\sum_{i=1}^n \overline{w}_i f(x_i)\right)\right\}
\end{aligned}
$$

Stratifikovaný resampling

- výchozí stav $\{w_i, \overline{x}_i\}_{i=1}^n$, odpovídá empirické míře $\widehat{\pi}_t^n(dx)$

  1. $D_{inv}(u) = i$ jestliže $u \in (\sum_{j=1}^{i-1} w_j, \sum_{j=1}^{i} w_j]$
  2. $(0,1] = (0,1/n] \cup \cdots \cup (\{n-1\}/n, 1]$
  3. $U^i \sim R(\{n-1\}/n, 1]$
  4. $x_i = x_{D_{inv}(U_i)}$ pro $i = 1, \ldots, n$

- vlastnosti pro $f \in \mathcal{C}_b$

$$\mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n f(x_i)\right] = \int f(x)\,\widehat{\pi}_t^n(dx)$$

$$Var\left[\frac{1}{n}\sum_{i=1}^n f(x_i)\right] = \frac{1}{n}\left\{\sum_{i=1}^n w_i f^2(x_i) - \left(\sum_{i=1}^n n\int_{(i-1)/n}^{i/n} f(x_{D^{inv}(u)}\,du\right)^2\right\}$$

Závěrečné poznámky

- reziduání a stratifikovaný sampling mají menší rozptyl než standardní multinomiální verze
- konvergence filtru pro všechny tři verze
- jádrové odhady filtračních hustot?
- konvegence jádrových odhadů hustot filtračních rozdělení *(D. Coufal. Kernel methods in particle filtering, to appear in Kybernetika)*
- rychlost konvergence v závislosti na verzi resamplingu

Děkuji za pozornost !!!