

Домашнее задание 1. Линейные динамические системы

Курс: Графические модели, весна 2016

Начало выполнения задания: 12 марта

Срок сдачи: **18 марта (пятница), 23:59.**

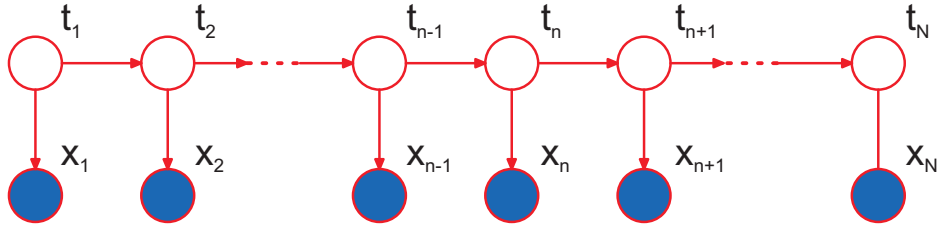


Рис. 1: Графическая модель для ЛДС

Линейной динамической системой (ЛДС) называется байесовская сеть, показанная на рис. 1, где $\mathbf{x}_n \in \mathbb{R}^d$, $\mathbf{t}_n \in \mathbb{R}^D$, и все атомарные распределения задаются линейной гауссовской моделью:

$$\begin{aligned} p(\mathbf{t}_n | \mathbf{t}_{n-1}) &= \mathcal{N}(\mathbf{t}_n | A\mathbf{t}_{n-1}, \Gamma), \\ p(\mathbf{x}_n | \mathbf{t}_n) &= \mathcal{N}(\mathbf{x}_n | C\mathbf{t}_n, \Sigma), \\ p(\mathbf{t}_1) &= \mathcal{N}(\mathbf{t}_1 | \boldsymbol{\mu}_0, V_0). \end{aligned}$$

Таким образом, совместное распределение в модели ЛДС задаётся как

$$p(X, T | \boldsymbol{\mu}_0, V_0, A, \Gamma, C, \Sigma) = \mathcal{N}(\mathbf{t}_1 | \boldsymbol{\mu}_0, V_0) \mathcal{N}(\mathbf{x}_1 | C\mathbf{t}_1, \Sigma) \prod_{n=2}^N \mathcal{N}(\mathbf{t}_n | A\mathbf{t}_{n-1}, \Gamma) \mathcal{N}(\mathbf{x}_n | C\mathbf{t}_n, \Sigma).$$

Задание

1. Рассмотрим задачу онлайн-фильтрации (фильтр Калмана), т.е. последовательное вычисление величин

$$p(\mathbf{t}_n | \mathbf{x}_1, \dots, \mathbf{x}_n) = \mathcal{N}(\mathbf{t}_n | \boldsymbol{\mu}_n, V_n)$$

для всех $n = 1, 2, \dots, N$. Пусть в момент времени n известно прогнозное распределение

$$p(\mathbf{t}_n | \mathbf{x}_1, \dots, \mathbf{x}_{n-1}) = \mathcal{N}(\mathbf{t}_n | \tilde{\boldsymbol{\mu}}_n, \tilde{V}_n).$$

Требуется вывести формулы фильтра Калмана для вычисления величин $\boldsymbol{\mu}_n, V_n$.

2. Рассмотрим задачу обучения параметров модели ЛДС $\Theta = (\boldsymbol{\mu}_0, V_0, A, \Gamma, C, \Sigma)$ с помощью метода максимального правдоподобия:

$$p(X | \Theta) \rightarrow \max_{\Theta}.$$

Эту задачу можно решать EM-алгоритмом, в котором на E-шаге при фиксированных параметрах Θ оцениваются распределения $p(\mathbf{t}_n | \mathbf{x}_1, \dots, \mathbf{x}_N)$ и $p([\mathbf{t}_n, \mathbf{t}_{n-1}] | \mathbf{x}_1, \dots, \mathbf{x}_N)$ для всех $n = 1, \dots, N$, а на M-шаге решается задача

$$\mathbb{E}_{p(T|X, \Theta)} \log p(X, T | \Theta) \rightarrow \max_{\Theta}.$$

Требуется вывести формулы пересчёта на M-шаге для параметров $C, \Sigma, \boldsymbol{\mu}_0, V_0$.

Полезные формулы

1. Тождество Вудбери:

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}.$$

Здесь $A \in \mathbb{R}^{n \times n}$, $C \in \mathbb{R}^{m \times m}$, $U \in \mathbb{R}^{n \times m}$, $V \in \mathbb{R}^{m \times n}$.

2. Пусть $p(\mathbf{x}) = \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \Sigma)$, $p(\mathbf{y}|\mathbf{x}) = \mathcal{N}(\mathbf{y}|A\mathbf{x}, \Gamma)$, $A \in \mathbb{R}^{m \times n}$. Тогда:

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{y}|A\boldsymbol{\mu}, \Gamma + A\Sigma A^T),$$

$$p(\mathbf{x}|\mathbf{y}) = \mathcal{N}(\mathbf{x}|P(A^T\Gamma^{-1}\mathbf{y} + \Sigma^{-1}\boldsymbol{\mu}), P), \quad P = (A^T\Gamma^{-1}A + \Sigma^{-1})^{-1}.$$

3. Производные следа: $\frac{\partial}{\partial A} \text{tr}(A^T B) = B$; $\frac{\partial}{\partial A} \text{tr}(A^T B A C) = B A C + B^T A C^T$.

4. Производная определителя: $\frac{\partial}{\partial A} \log \det A = A^{-T}$.