

Лекция 8. Линейные динамические системы. Фильтр Калмана.

А. С. Коцушин¹ Д. П. Ветров² Д. А. Кропотов³
В. С. Коцушин¹ О. В. Барина¹

¹МГУ, ВМиК, лаб. КГ

²МГУ, ВМиК, каф. ММП

³ВЦ РАН

Спецкурс «Структурные методы анализа изображений
и сигналов»

План лекции

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Линейные динамические системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

План лекции

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Линейные динамические системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Пример - задача трекинга

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

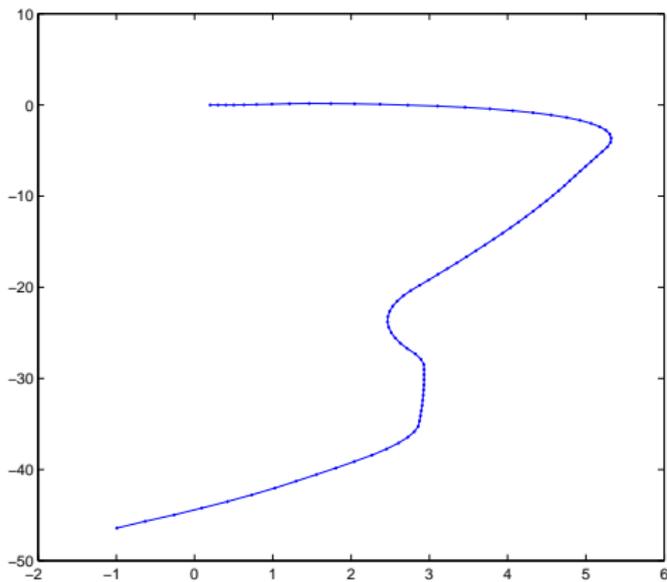
Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Пусть имеется траектория движения объекта во времени



Пример - задача трекинга

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

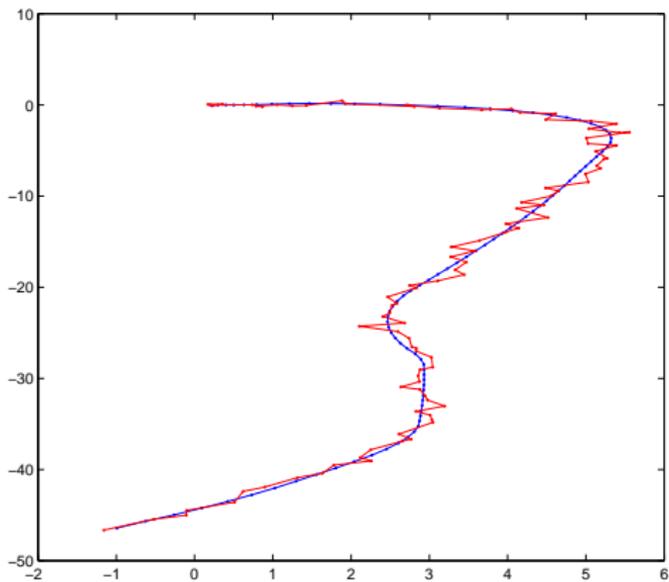
Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Координаты объекты измерены с некоторой погрешностью



Пример - задача трекинга

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

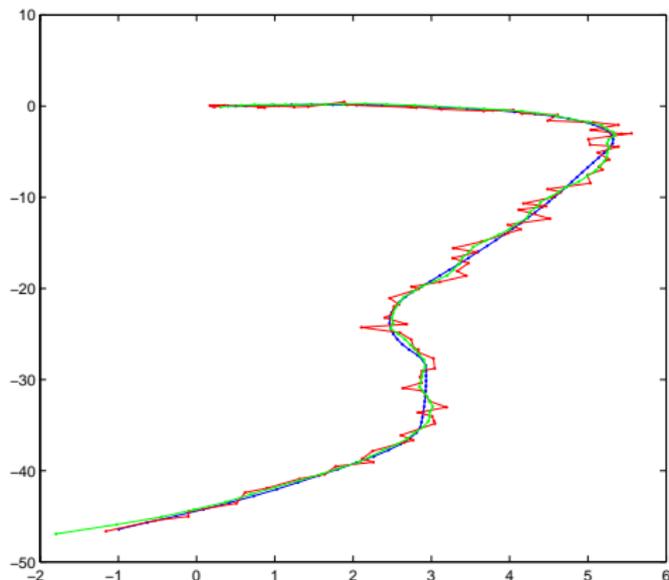
Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Задача - оценить истинные координаты объекта с использованием координат объекта в различные моменты времени



Вероятностная модель

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

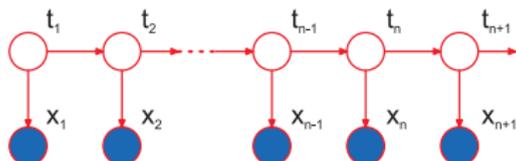
Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание



- $\mathbf{x}_1, \dots, \mathbf{x}_N$ — наблюдаемая последовательность
- $\mathbf{t}_1, \dots, \mathbf{t}_N$ — истинные параметры объекта
- $p(\mathbf{t}_n | \mathbf{t}_{n-1})$ — модель движения объекта
- $p(\mathbf{x}_n | \mathbf{t}_n)$ — модель шума

Пример спецификации модели

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

- $\mathbf{t}_n = \{\xi_1(n), \xi_2(n), \dot{\xi}_1(n), \dot{\xi}_2(n), \ddot{\xi}_1(n), \ddot{\xi}_2(n)\}$ - истинные координаты, скорости и ускорения объекта в момент времени n

- Модель движения объекта

$$\xi_i(n) = \xi_i(n-1) + \dot{\xi}_i(n-1)\Delta t + \frac{1}{2}\ddot{\xi}_i(n-1)(\Delta t)^2 + \varepsilon_1, \quad i = 1, 2$$

$$\dot{\xi}_i(n) = \dot{\xi}_i(n-1) + \ddot{\xi}_i(n-1)\Delta t + \varepsilon_2, \quad i = 1, 2$$

$$\ddot{\xi}_i(n) = \ddot{\xi}_i(n-1) + \varepsilon_3, \quad i = 1, 2$$

$$\varepsilon_i \sim \mathcal{N}(0, \sigma_i), \quad i = 1, 2, 3$$

- Модель сенсора (шума)

$$x_i(n) = \xi_i(n) + \nu_i, \quad i = 1, 2$$

$$\nu_i \sim \mathcal{N}(0, s_i), \quad i = 1, 2$$

Скрытая марковская модель

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

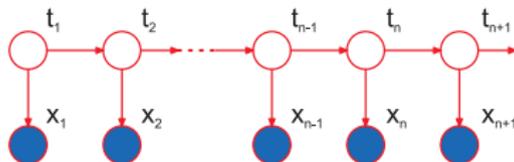
Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание



- Переменные x_n — произвольные
- Переменные t_n — **дискретные**, принимают значения из $\{1, \dots, K\}$
- Распределения $p(x_n|t_n)$ и $p(t_n|t_{n-1})$ — произвольные.

Наша вероятностная модель имеют такую же структуру.

Основное отличие - переменные t_n **непрерывные**.

Обучение и вывод в скрытой марковской модели

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$p(\mathbf{t}_n|X) = \frac{p(\mathbf{t}_n, X)}{p(X)} = \frac{p(\mathbf{t}_n, \mathbf{x}_1, \dots, \mathbf{x}_n)}{p(\mathbf{x}_1, \dots, \mathbf{x}_n)} \cdot \frac{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N | \mathbf{t}_n)}{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N | \mathbf{x}_1, \dots, \mathbf{x}_n)} = \hat{\alpha}(\mathbf{t}_n) \hat{\beta}(\mathbf{t}_n)$$

Формулы пересчета

$$c_n \hat{\alpha}(\mathbf{t}_n) = p(\mathbf{x}_n | \mathbf{t}_n) \sum_{\mathbf{t}_{n-1}} \hat{\alpha}(\mathbf{t}_{n-1}) p(\mathbf{t}_n | \mathbf{t}_{n-1})$$

$$c_{n+1} \hat{\beta}(\mathbf{t}_n) = \sum_{\mathbf{t}_{n+1}} \hat{\beta}(\mathbf{t}_{n+1}) p(\mathbf{x}_{n+1} | \mathbf{t}_{n+1}) p(\mathbf{t}_{n+1} | \mathbf{t}_n)$$

В случае непрерывных \mathbf{t}_n сумма заменяется на интеграл.

Ограничения на вероятностную модель

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

- Алгоритм вывода должен иметь линейную по N сложность.
- Формулы пересчета

$$c_n \hat{\alpha}(t_n) = p(\mathbf{x}_n | t_n) \int \hat{\alpha}(t_{n-1}) p(t_n | t_{n-1}) dt_{n-1}$$

$$c_{n+1} \hat{\beta}(t_n) = \int \hat{\beta}(t_{n+1}) p(\mathbf{x}_{n+1} | t_{n+1}) p(t_{n+1} | t_n) dt_{n+1}$$

Эти интегралы должны вычисляться аналитически и модель не должна усложняться при переходе от $\hat{\alpha}(t_{n-1})$ к $\hat{\alpha}(t_n)$.

- Пример усложнения модели. Пусть $p(\mathbf{x}_n | t_n)$ - смесь K гауссиан. Тогда если $\hat{\alpha}(t_1)$ гауссиана, то $\hat{\alpha}(t_2)$ - смесь из K гауссиан, $\hat{\alpha}(t_3)$ - смесь из K^2 гауссиан и т.д.
- Для решения задачи сегментации функция Беллмана должна вычисляться аналитически.

Линейная динамическая система (ЛДС)

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$p(\mathbf{t}_n | \mathbf{t}_{n-1}) = \mathcal{N}(\mathbf{t}_n | A\mathbf{t}_{n-1}, \Gamma)$$

$$p(\mathbf{x}_n | \mathbf{t}_n) = \mathcal{N}(\mathbf{x}_n | C\mathbf{t}_n, \Sigma)$$

$$p(\mathbf{t}_1) = \mathcal{N}(\mathbf{t}_1 | \boldsymbol{\mu}_0, V_0)$$

Эквивалентная формулировка

$$\mathbf{t}_n = A\mathbf{t}_{n-1} + \mathbf{w}_n$$

$$\mathbf{x}_n = C\mathbf{t}_n + \mathbf{v}_n$$

$$\mathbf{t}_1 = \boldsymbol{\mu}_0 + \mathbf{u}$$

$$\mathbf{w} \sim \mathcal{N}(\mathbf{w} | \mathbf{0}, \Gamma)$$

$$\mathbf{v} \sim \mathcal{N}(\mathbf{v} | \mathbf{0}, \Sigma)$$

$$\mathbf{u} \sim \mathcal{N}(\mathbf{u} | \mathbf{0}, V_0)$$

Параметры модели $\{A, \Gamma, C, \Sigma, \boldsymbol{\mu}_0, V_0\}$

Свойство ЛДС

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$p(X, T|\Theta) = p(t_1) \prod_{n=2}^N p(t_n|t_{n-1}) \prod_{n=1}^N p(x_n|t_n)$$

Все атомарные распределения представляют собой линейную гауссовскую модель. Поэтому совместное распределение $p(X, T|\Theta)$, а также все его маргинальные и условные распределения будут также гауссовскими. Наиболее вероятная последовательность T_* определяется по индивидуально наиболее вероятным значениям t_n :

$$t_n^* = \arg \max_{t_n} p(t_n|X, \Theta)$$

Вывод: если известны $p(t_n|X, \Theta)$, то аналог алгоритма Витерби для сегментации не требуется!

Многомерное нормальное распределение

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$\mathbf{x} \sim \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^d \det \boldsymbol{\Sigma}} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right)$$

Пусть \mathbf{x} состоит из двух групп переменных \mathbf{x}_a и \mathbf{x}_b , т.е.

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_a \\ \mathbf{x}_b \end{pmatrix}, \quad \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_a \\ \boldsymbol{\mu}_b \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{aa} & \boldsymbol{\Sigma}_{ab} \\ \boldsymbol{\Sigma}_{ba} & \boldsymbol{\Sigma}_{bb} \end{pmatrix}, \quad \boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1} = \begin{pmatrix} \boldsymbol{\Lambda}_{aa} & \boldsymbol{\Lambda}_{ab} \\ \boldsymbol{\Lambda}_{ba} & \boldsymbol{\Lambda}_{bb} \end{pmatrix}$$

Тогда

$$p(\mathbf{x}_a) = \mathcal{N}(\mathbf{x}_a|\boldsymbol{\mu}_a, \boldsymbol{\Sigma}_{aa})$$

$$p(\mathbf{x}_a|\mathbf{x}_b) = \mathcal{N}(\mathbf{x}_a|\boldsymbol{\mu}_a - \boldsymbol{\Lambda}_{aa}^{-1}\boldsymbol{\Lambda}_{ab}(\mathbf{x}_b - \boldsymbol{\mu}_b), \boldsymbol{\Lambda}_{aa}^{-1})$$

План лекции

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Линейные динамические системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Вывод в ЛДС

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Для решения задачи сегментации, а также для EM-алгоритма обучения ЛДС важно уметь вычислять характеристики:

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

$$\hat{\beta}(t_n) = \frac{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N | t_n)}{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N | \mathbf{x}_1, \dots, \mathbf{x}_n)}$$

Для их вычисления можно применять алгоритм «вперед-назад». Проход вперед для ЛДС получил название фильтра Калмана, а проход назад - РТС уравнения (Rauch-Tung-Striebel).

Проход вперед (фильтр Калмана)

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Формула пересчета

$$c_n \hat{\alpha}(t_n) = p(\mathbf{x}_n | t_n) \int \hat{\alpha}(t_{n-1}) p(t_n | t_{n-1}) dt_{n-1}$$

Подставляем

$$c_n \mathcal{N}(t_n | \mu_n, V_n) = \mathcal{N}(\mathbf{x}_n | C t_n, \Sigma) \int \mathcal{N}(t_n | A t_{n-1}, \Gamma) \mathcal{N}(t_{n-1} | \mu_{n-1}, V_{n-1}) dt_{n-1}$$

Результат

$$\begin{aligned} \mu_n &= A \mu_{n-1} + K_n (\mathbf{x}_n - C A \mu_{n-1}) \\ V_n &= (I - K_n C) P_{n-1} \\ c_n &= \mathcal{N}(\mathbf{x}_n | C A \mu_{n-1}, C P_{n-1} C^T + \Sigma) \\ P_{n-1} &= A V_{n-1} A^T + \Gamma \\ K_n &= P_{n-1} C^T (C P_{n-1} C^T + \Sigma)^{-1} \end{aligned}$$

Фильтр Калмана

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Традиционно в фильтре Калмана рассматривают два этапа:
Прогноз. Оценивается распределение

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

$$\tilde{\boldsymbol{\mu}}_n = A\boldsymbol{\mu}_{n-1}$$

$$\tilde{\mathbf{V}}_n = P_{n-1}$$

$$\tilde{\mathbf{x}}_n = C\tilde{\boldsymbol{\mu}}_n$$

Коррекция.

$$\boldsymbol{\mu}_n = \tilde{\boldsymbol{\mu}}_n + K_n(\mathbf{x}_n - \tilde{\mathbf{x}}_n)$$

$$\mathbf{V}_n = (I - K_n C)\tilde{\mathbf{V}}_n$$

$$c_n = \mathcal{N}(\mathbf{x}_n | \tilde{\mathbf{x}}_n, C\tilde{\mathbf{V}}_n C^T + \Sigma)$$

Иллюстрация фильтра Калмана

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

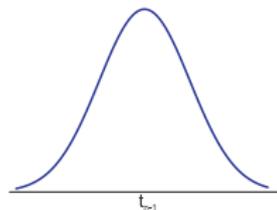
Кропотов

Линейные
динамические
системы

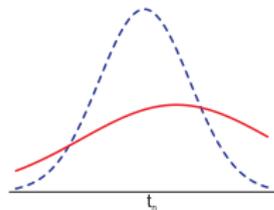
Вывод в ЛДС

Обучение в ЛДС

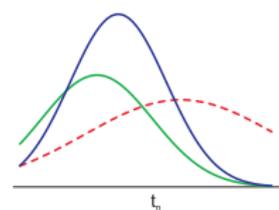
Второе задание



(a)



(b)



(c)

— $p(t_{n-1} | \mathbf{x}_1, \dots, \mathbf{x}_{n-1})$

— $p(t_n | \mathbf{x}_1, \dots, \mathbf{x}_{n-1})$

— $p(t_n | \mathbf{x}_1, \dots, \mathbf{x}_n)$

Фильтр Калмана. Начальное приближение.

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Значение $\hat{\alpha}(t_1)$ в начальный момент времени вычисляется из условия:

$$c_1 \hat{\alpha}(t_1) = p(t_1) p(\mathbf{x}_1 | t_1)$$

Производя свертку двух гауссиан, получаем:

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_0 + K_1(\mathbf{x}_1 - C\boldsymbol{\mu}_0)$$

$$V_1 = (I - K_1 C)V_0$$

$$c_1 = \mathcal{N}(\mathbf{x}_1 | C\boldsymbol{\mu}_0, CV_0 C^T + \Sigma)$$

$$K_1 = V_0 C^T (CV_0 C^T + \Sigma)^{-1}$$

Проход назад

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$\gamma(\mathbf{t}_n) = p(\mathbf{t}_n | X, \Theta) = \hat{\alpha}(\mathbf{t}_n) \hat{\beta}(\mathbf{t}_n) = \mathcal{N}(\mathbf{t}_n | \hat{\boldsymbol{\mu}}_n, \hat{V}_n)$$

В отличие от $\gamma(\mathbf{t}_n)$ $\hat{\beta}(\mathbf{t}_n)$ не является маргинальным распределением:

$$\hat{\beta}(\mathbf{t}_n) = \frac{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N | \mathbf{t}_n)}{p(\mathbf{x}_{n+1}, \dots, \mathbf{x}_N | \mathbf{x}_1, \dots, \mathbf{x}_n)}$$

Поэтому формулы для обратного прохода удобнее записывать в терминах $\gamma(\mathbf{t}_n)$

Формулы для обратного прохода

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Формула пересчета

$$c_{n+1}\hat{\beta}(t_n) = \int \hat{\beta}(t_{n+1})p(x_{n+1}|t_{n+1})p(t_{n+1}|t_n)dt_{n+1}$$

$$\hat{\mu}_n = \mu_n + J_n(\hat{\mu}_{n+1} - A\mu_n)$$

$$\hat{V}_n = V_n + J_n(\hat{V}_{n+1} - P_n)J_n^T$$

$$J_n = V_n A^T (P_n)^{-1}$$

Распределение для пары переменных

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Для EM-алгоритма обучения понадобятся также величины

$$\begin{aligned}\xi(\mathbf{t}_{n-1}, \mathbf{t}_n) &= p(\mathbf{t}_{n-1}, \mathbf{t}_n | X, \Theta) = \\ &= (c_n)^{-1} \hat{\alpha}(\mathbf{t}_{n-1}) p(\mathbf{x}_n | \mathbf{t}_n) p(\mathbf{t}_n | \mathbf{t}_{n-1}) \hat{\beta}(\mathbf{t}_n) = \\ &= \frac{\mathcal{N}(\mathbf{t}_{n-1} | \boldsymbol{\mu}_{n-1}, V_{n-1}) \mathcal{N}(\mathbf{t}_n | A\mathbf{t}_{n-1}, \Gamma) \mathcal{N}(\mathbf{x}_n | C\mathbf{t}_n, \Sigma) \mathcal{N}(\mathbf{t}_n | \hat{\boldsymbol{\mu}}_n, \hat{V}_n)}{c_n \mathcal{N}(\mathbf{t}_n | \boldsymbol{\mu}_n, V_n)} = \\ &= \mathcal{N}(\mathbf{t}_{n-1}, \mathbf{t}_n | [\gamma(\mathbf{t}_{n-1}), \gamma(\mathbf{t}_n)]^T, J_{n-1} \hat{V}_n)\end{aligned}$$

План лекции

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Линейные динамические системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

EM-алгоритм. Разложение логарифма правдоподобия

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Требуется найти максимум правдоподобия в вероятностной модели со скрытыми переменными:

$$p(X|\Theta) = \int p(X, T|\Theta) dT \rightarrow \max_{\Theta} \Leftrightarrow \log p(X|\Theta) \rightarrow \max_{\Theta}$$

$$p(X, T|\Theta) = p(T|X, \Theta)p(X|\Theta) \Rightarrow$$

$$\log p(X, T|\Theta) = \log p(T|X, \Theta) + \log p(X|\Theta)$$

$q(T)$ — произвольное распределение.

$$\log p(X|\Theta) = \int \log p(X|\Theta) q(T) dT =$$

$$\int [\log p(X, T|\Theta) - \log p(T|X, \Theta)] q(T) dT =$$

$$\int \log p(X, T|\Theta) q(T) dT - \int \log p(T|X, \Theta) q(T) dT =$$

$$\int \log \frac{p(X, T|\Theta)}{q(T)} q(T) dT - \int \log \frac{p(T|X, \Theta)}{q(T)} q(T) dT$$

Нижняя оценка для логарифма правдоподобия

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$\log p(X|\Theta) = \underbrace{\int \log \frac{p(X, T|\Theta)}{q(T)} q(T) dT}_{l(q, \Theta)} - \underbrace{\int \log \frac{p(T|X, \Theta)}{q(T)} q(T) dT}_{KL(q||p) \geq 0}$$

Дивергенция Кульбака-Лейблера $KL(q||p)$ определяет расстояние между вероятностными распределениями

- $KL(q||p) = - \int q(x) \log(p(x)/q(x)) dx$
- $KL(q||p) \geq 0$ и $KL(q||p) = 0 \Leftrightarrow q \equiv p$.
- $KL(q||p) \neq KL(p||q)$

Тогда $l(q, \Theta)$ является нижней оценкой правдоподобия $\log p(X|\Theta)$:

$$\log p(X|\Theta) \geq l(q, \Theta) \text{ и равенство } \Leftrightarrow q(T) = p(T|X, \Theta)$$

Идея EM-алгоритма

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$\log p(X|\Theta) = l(q, \Theta) + KL(q||p)$$

Итерационная схема. Фиксируем некоторое значение Θ_{old} . Приближим в точке Θ_{old} правдоподобие с помощью его нижней оценки:

$$q(T) = p(T|X, \Theta_{old})$$

$$\log p(X|\Theta) \geq l(q, \Theta) = \int \log p(X, T|\Theta) p(T|X, \Theta_{old}) dT - \int \log p(T|X, \Theta_{old}) p(T|X, \Theta_{old}) dT$$

Найдем новое значение Θ с помощью максимизации нижней оценки:

$$l(q, \Theta) \rightarrow \max_{\Theta} \Leftrightarrow \mathbb{E}_{T|X, \Theta_{old}} \log p(X, T|\Theta) \rightarrow \max_{\Theta}$$

Иллюстрация EM-алгоритма

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

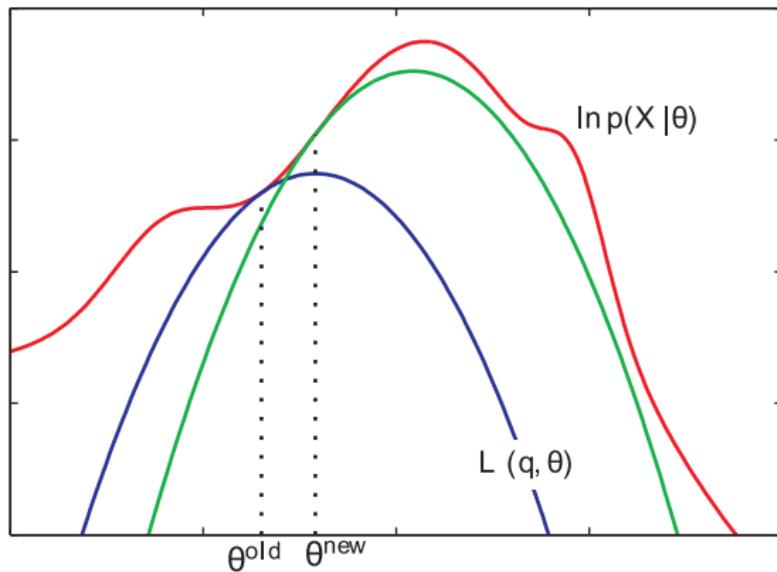


Схема EM-алгоритма

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

- **Е-шаг.** Фиксируется значение параметров Θ_{old} .
Оценивается апостериорное распределение на скрытые переменные $p(T|X, \Theta_{old})$, и полное правдоподобие усредняется по полученному распределению:

$$\mathbb{E}_{T|X, \Theta_{old}} \log p(X, T|\Theta) = \int \log p(X, T|\Theta) p(T|X, \Theta_{old}) dT$$

- **М-шаг.** Фиксируется апостериорное распределение $p(T|X, \Theta_{old})$, и производится поиск новых значений параметров Θ_{new} :

$$\Theta_{new} = \arg \max_{\Theta} \mathbb{E}_{T|X, \Theta_{old}} \log p(X, T|\Theta)$$

- Шаги Е и М повторяются до сходимости.

Максимизация апостериорного распределения

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Задача

$$p(\Theta|X) \rightarrow \max_{\Theta} \Leftrightarrow F = \log p(X|\Theta) + \log p(\Theta) \rightarrow \max_{\Theta}$$

Справедливо разложение

$$F = L(q, \theta) + \log p(\Theta) + KL(q||p) \geq L(q, \Theta) + \log p(\Theta)$$

E-шаг остается без изменений.

Модификация M-шага:

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

EM-алгоритм для ЛДС

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Задача - поиск значений параметров $\{A, \Gamma, C, \Sigma, \mu_0, V_0\}$ по методу максимального правдоподобия

Логарифм полного правдоподобия

$$\log p(X, T | \Theta) = \log p(t_1 | \mu_0, V_0) + \sum_{n=2}^N \log p(t_n | t_{n-1}, A, \Gamma) + \sum_{n=1}^N \log p(x_n | t_n, C, \Sigma)$$

Для вычисления нижней оценки

$$Q(\Theta, \Theta_{old}) = \mathbb{E}_{T|X, \Theta_{old}} \log p(X, T | \Theta)$$

достаточно знать следующие величины:

$$\mathbb{E} t_n = \hat{\mu}_n$$

$$\mathbb{E} t_n t_{n-1}^T = J_{n-1} \hat{V}_n + \hat{\mu} \hat{\mu}_{n-1}^T$$

$$\mathbb{E} t_n t_n^T = \hat{V}_n + \hat{\mu}_n \hat{\mu}_n^T$$

M-шаг. Формулы для μ_0 и V_0

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$Q(\Theta, \Theta_{old}) = -\frac{1}{2} \log \det V_0 - \mathbb{E}_{T|X, \Theta_{old}} \left[\frac{1}{2} (\mathbf{t}_1 - \boldsymbol{\mu}_0)^T V_0^{-1} (\mathbf{t}_1 - \boldsymbol{\mu}_0) \right] + \text{const}$$

Здесь const не зависит от μ_0 и V_0 .

Приравнивая производные по μ_0 и V_0 к нулю, получаем:

$$\boldsymbol{\mu}_0^{new} = \mathbb{E} \mathbf{t}_1$$

$$V_0^{new} = \mathbb{E} \mathbf{t}_1 \mathbf{t}_1^T - \mathbb{E} \mathbf{t}_1 \mathbb{E} \mathbf{t}_1^T$$

M-шаг. Формулы для A и Γ

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$Q(\Theta, \Theta_{old}) = -\frac{N-1}{2} \log \det \Gamma - \mathbb{E}_{T|X, \Theta_{old}} \left[\frac{1}{2} \sum_{n=2}^N (\mathbf{t}_n - A\mathbf{t}_{n-1})^T \Gamma^{-1} (\mathbf{t}_n - A\mathbf{t}_{n-1}) \right] + \text{const}$$

Здесь const не зависит от A и Γ.

Приравнявая производные по A и Γ к нулю, получаем:

$$A^{new} = \left(\sum_{n=2}^N \mathbb{E} \mathbf{t}_n \mathbf{t}_{n-1}^T \right) \left(\sum_{n=2}^N \mathbb{E} \mathbf{t}_{n-1} \mathbf{t}_{n-1}^T \right)^{-1}$$

$$\Gamma^{new} = \frac{1}{N-1} \sum_{n=2}^N \left\{ \mathbb{E} \mathbf{t}_n \mathbf{t}_n^T - A^{new} \mathbb{E} \mathbf{t}_{n-1} \mathbf{t}_n^T - \mathbb{E} \mathbf{t}_n \mathbf{t}_{n-1}^T (A^{new})^T + A^{new} \mathbb{E} \mathbf{t}_{n-1} \mathbf{t}_{n-1}^T (A^{new})^T \right\}$$

M-шаг. Формулы для C и Σ

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

$$Q(\Theta, \Theta_{old}) = -\frac{N}{2} \log \det \Sigma -$$

$$\mathbb{E}_{T|X, \Theta_{old}} \left[\frac{1}{2} \sum_{n=1}^N (\mathbf{x}_n - C\mathbf{t}_n)^T \Sigma^{-1} (\mathbf{x}_n - C\mathbf{t}_n) \right] + \text{const}$$

Здесь const не зависит от C и Σ .

Приравнявая производные по C и Σ к нулю, получаем:

$$C^{new} = \left(\sum_{n=1}^N \mathbf{x}_n \mathbb{E} \mathbf{t}_n^T \right) \left(\sum_{n=1}^N \mathbb{E} \mathbf{t}_n \mathbf{t}_n^T \right)^{-1}$$

$$\Sigma^{new} = \frac{1}{N} \sum_{n=1}^N \left\{ \mathbb{E} \mathbf{x}_n \mathbf{x}_n^T - C^{new} \mathbb{E} \mathbf{t}_n \mathbf{x}_n^T -$$

$$\mathbf{x}_n \mathbb{E} \mathbf{t}_n^T (C^{new})^T + C^{new} \mathbb{E} \mathbf{t}_n \mathbf{t}_n^T (C^{new})^T \right\}$$

План лекции

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Линейные динамические системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Первый вариант

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Скрытая марковская модель первого порядка. Вероятности генерации данных $p(\mathbf{x}_n | t_n)$ — многомерные нормальные распределения.

- Реализовать алгоритм Витерби для сегментации с помощью СММ
- Реализовать EM-алгоритм обучения СММ с известным количеством скрытых состояний
- Реализовать алгоритм генерации выборки из модели СММ
- Среда реализации — MATLAB, эффективная реализация всячески приветствуется.

Второй вариант

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Линейная динамическая система.

- Реализовать фильтр Калмана и РТС уравнения
- Реализовать EM-алгоритм обучения ЛДС с известной размерностью скрытой компоненты. При этом предусмотреть возможность задания части параметров пользователем.
- Реализовать алгоритм генерации выборки из модели ЛДС, а также алгоритм генерации траектории движения объекта в многомерном пространстве.
- Среда реализации — MATLAB, эффективная реализация всячески приветствуется.

Лекция 8.
Линейные
динамические
системы. Фильтр
Калмана.

Кропотов

Линейные
динамические
системы

Вывод в ЛДС

Обучение в ЛДС

Второе задание

Срок сдачи задания - 2 декабря.