

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Лекция 1. Байесовские и марковские сети

Д. П. Ветров¹ Д. А. Кропотов² А. А. Осокин¹

¹МГУ, ВМиК, каф. ММП

²ВЦ РАН

Курс «Графические модели»

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Основной предмет изучения

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Основные объекты изучения

- Аппарат графических моделей (байесовские и марковские сети)
- Аппарат байесовского вывода
- Некоторые методы дискретной оптимизации
- Методы структурного обучения

Примеры применения графических моделей

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Семантическая сегментация изображений
- Обработка видео
- Распознавание речи
- Анализ социальных сетей
- Трехмерная реконструкция
- Задачи трекинга
- Биоинформатика
- Органическая химия
- Декодирование зашумленных сообщений
- и многое другое

Примеры использования

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети



Оценка диспаратитетов и восстановление стерео

Примеры использования

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

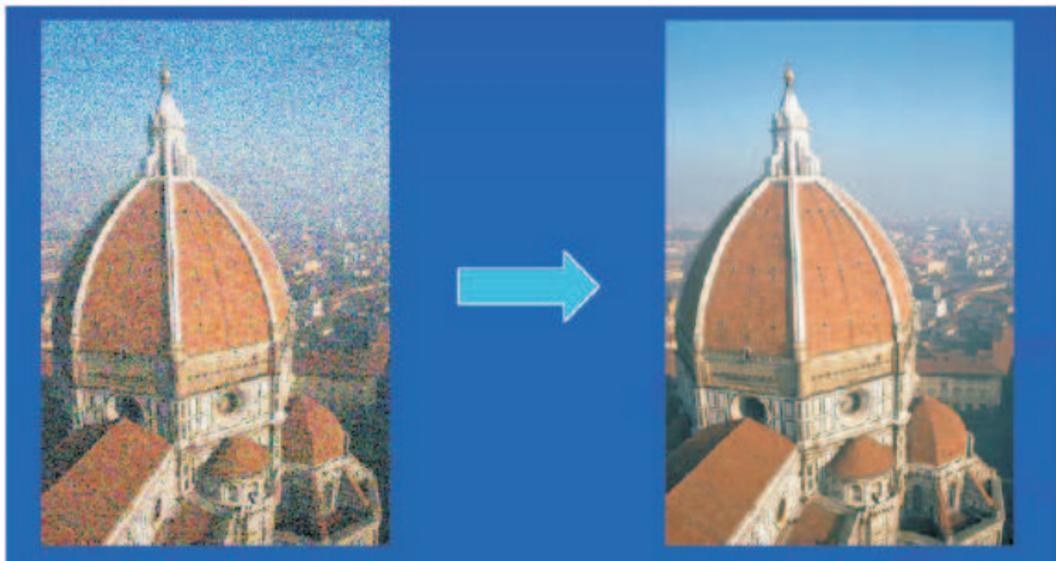
Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети



Фильтрация изображений

Примеры использования

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети



Создание коллажей

Примеры использования

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

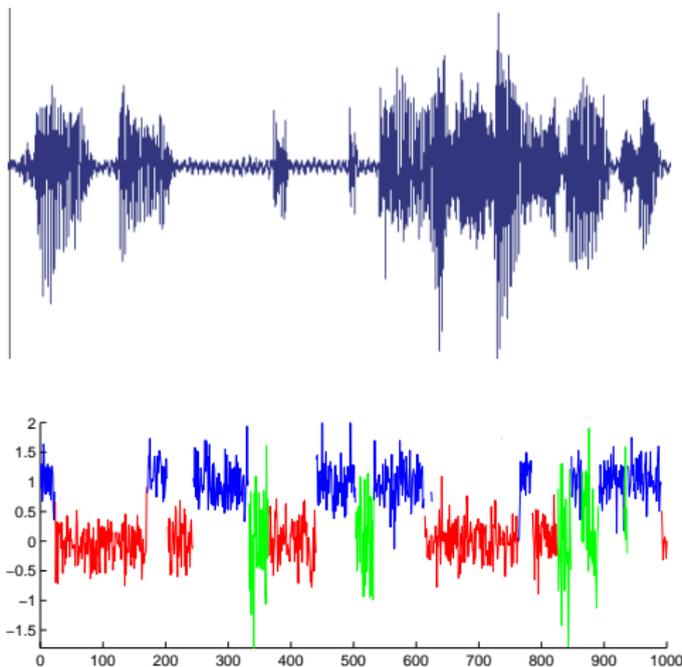
Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети



Обработка сигналов и распознавание речи

Примеры использования

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

```
u LITTLE-JACK-HORNER-SAT-IN-THE-CORNER-EATING-A-CHRISTMAS-PIE--HE-PUT-IN-H
v RIDE-A-COCK-HORSE-TO-BANBURY-CROSS-TO-SEE-A-FINE-LADY-UPON-A-WHITE-HORSE
matches: .*.***.*****.*.....*.....*.....*.....*
```

Восстановление зашумленных сообщений в ситуациях,
когда уровень шума превышает возможности
самокоррекции

Матричная нотация

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- При работе с многомерными величинами очень удобна матричная нотация, т.е. представление многих операций над векторами и числами в виде операций над матрицами
- Скалярное произведение двух векторов $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ принимает вид

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^d x_i y_i = \mathbf{x}^T \mathbf{y},$$

т.е. вектора трактуются как частные случаи матриц

- Квадратичная форма

$$\langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^d \sum_{j=1}^d x_i a_{ij} y_j = \mathbf{x}^T \mathbf{A} \mathbf{y}$$

- Матричная нотация облегчает математические выкладки и позволяет реализовать вычисления на ЭВМ более эффективно

Пример использования

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Предположим нам надо решить несовместную систему линейных уравнений $Ax \approx b$, $A \in \mathbb{R}^{m \times n}$
- Для этого будем минимизировать квадрат нормы невязки (система-то нерешаемая) $\|Ax - b\|^2 \rightarrow \min_x$
- Представляя норму в матричной виде, дифференцируя по вектору и приравнивая производную к нулю получаем известную формулу для псевдорешения СЛАУ

$$\begin{aligned}\|Ax - b\|^2 &= \langle Ax - b, Ax - b \rangle = (Ax - b)^T (Ax - b) = \\ &= (Ax)^T Ax - b^T Ax - (Ax)^T b + b^T b = \\ &= x^T A^T Ax - 2x^T A^T b + b^T b\end{aligned}$$

$$\frac{\partial}{\partial x} (x^T A^T Ax - 2x^T A^T b + b^T b) = 2A^T Ax - 2A^T b = 0$$

$$x = (A^T A)^{-1} A^T b$$

- Заметим, что если матрица A квадратная (число уравнений равно числу неизвестных) и невырожденная, то последняя формула переходит в формулу обычного решения СЛАУ $x = A^{-1}b$

Задача условной оптимизации

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Пусть $f(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$ — гладкая функция. Предположим, что нам необходимо найти ее экстремум:

$$f(\mathbf{x}) \rightarrow \underset{\mathbf{x}}{\text{extr}}$$

Для того, чтобы найти экстремум (решить задачу безусловной оптимизации), достаточно проверить условие стационарности:

$$\nabla f(\mathbf{x}) = 0$$

Предположим, что нам необходимо найти экстремум функции при ограничениях:

$$f(\mathbf{x}) \rightarrow \underset{\mathbf{x}}{\text{extr}}$$

$$g(\mathbf{x}) = 0$$

Поверхность ограничения

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

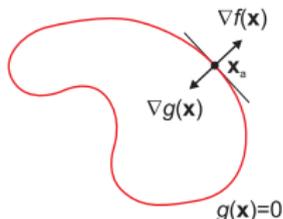
Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети
Марковские сети



Заметим, что $\nabla g(\mathbf{x})$ ортогонален поверхности ограничения $g(\mathbf{x}) = 0$. Пусть \mathbf{x} и $\mathbf{x} + \boldsymbol{\varepsilon}$ — две близкие точки поверхности. Тогда

$$g(\mathbf{x} + \boldsymbol{\varepsilon}) \simeq g(\mathbf{x}) + \boldsymbol{\varepsilon}^T \nabla g(\mathbf{x})$$

Т.к. $g(\mathbf{x} + \boldsymbol{\varepsilon}) = g(\mathbf{x})$, то $\boldsymbol{\varepsilon}^T \nabla g(\mathbf{x}) \simeq 0$. При стремлении $\|\boldsymbol{\varepsilon}\| \rightarrow 0$ получаем $\boldsymbol{\varepsilon}^T \nabla g(\mathbf{x}) = 0$. Т.к. $\boldsymbol{\varepsilon}$ параллелен поверхности $g(\mathbf{x}) = 0$, то $\nabla g(\mathbf{x})$ является нормалью к этой поверхности.

Функция Лагранжа

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Необходимым условием оптимальности является ортогональность $\nabla f(\mathbf{x})$ поверхности ограничения, т.е.:

$$\nabla f + \lambda \nabla g = 0 \quad (1)$$

Здесь $\lambda \neq 0$ — коэффициент Лагранжа. Он может быть любого знака.

Функция Лагранжа

$$L(\mathbf{x}, \lambda) \triangleq f(\mathbf{x}) + \lambda g(\mathbf{x})$$

Тогда

$$\nabla_{\mathbf{x}} L = 0 \quad \Rightarrow \quad \text{условие (1)}$$

$$\frac{\partial}{\partial \lambda} L = 0 \quad \Rightarrow \quad g(\mathbf{x}) = 0$$

Функция Лагранжа. Пример.

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

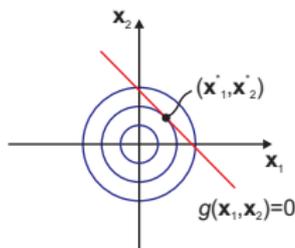
Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети



$$f(x_1, x_2) = 1 - x_1^2 - x_2^2 \rightarrow \max_{x_1, x_2}$$

$$g(x_1, x_2) = x_1 + x_2 - 1 = 0$$

Функция Лагранжа:

$$L(\mathbf{x}, \lambda) = 1 - x_1^2 - x_2^2 + \lambda(x_1 + x_2 - 1)$$

Условия стационарности:

$$-2x_1 + \lambda = 0$$

$$-2x_2 + \lambda = 0$$

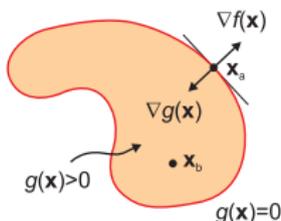
$$x_1 + x_2 - 1 = 0$$

Решение: $(x_1^*, x_2^*) = (\frac{1}{2}, \frac{1}{2})$, $\lambda = 1$.

Ограничение в виде неравенства

$$f(\mathbf{x}) \rightarrow \max_{\mathbf{x}}$$

$$g(\mathbf{x}) \geq 0$$



Решение	Ограничение	Условие стационарности
Внутри области $g(\mathbf{x}) > 0$	неактивно	$\nabla f(\mathbf{x}) = 0, \nabla_{\mathbf{x}} L = 0, \lambda = 0$
На границе $g(\mathbf{x}) = 0$	активно	$\nabla f(\mathbf{x}) = -\lambda \nabla g(\mathbf{x}),$ $\nabla_{\mathbf{x}, \lambda} L = 0, \lambda > 0$

Условие дополняющей нежесткости:

$$\lambda g(\mathbf{x}) = 0$$

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Теорема Каруша-Куна-Таккера

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Пусть $f_i : X \rightarrow \mathbb{R}$, $i = 0, 1, \dots, m$ — выпуклые функции, отображающие нормированное пространство X в прямую, $A \in X$ — выпуклое множество. Рассмотрим следующую задачу оптимизации:

$$f_0(\mathbf{x}) \rightarrow \min; \quad f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m, \quad \mathbf{x} \in A \quad (2)$$

Теорема

1. Если $\hat{\mathbf{x}} \in \text{absmin}(2)$ — решение задачи, то найдется ненулевой вектор множителей Лагранжа $\boldsymbol{\lambda} \in \mathbb{R}^{m+1}$ такой, что для функции Лагранжа $L(\mathbf{x}) = \sum_{i=0}^m \lambda_i f_i(\mathbf{x})$ выполняются условия:
 - a) стационарности $\min_{\mathbf{x} \in A} L(\mathbf{x}) = L(\hat{\mathbf{x}})$
 - b) дополняющей нежесткости $\lambda_i f_i(\hat{\mathbf{x}}) = 0$, $i = 1, \dots, m$
 - c) неотрицательности $\lambda_i \geq 0$
2. Если для допустимой точки $\hat{\mathbf{x}}$ выполняются условия а)–с) и $\lambda_0 \neq 0$, то $\hat{\mathbf{x}} \in \text{absmin}(2)$
3. Если для допустимой точки $\hat{\mathbf{x}}$ выполняются условия а)–с) и $\exists \bar{\mathbf{x}} \in A : f_i(\bar{\mathbf{x}}) < 0, i = 0, \dots, m$ (условие Слейтера), то $\hat{\mathbf{x}} \in \text{absmin}(2)$

План

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия**
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Краткое напоминание основных вероятностных понятий

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Нормальное распределение
Формула Байеса

Ликбез

Графические модели

Байесовские сети

Марковские сети

- $X : \Omega \rightarrow \mathbb{R}$ – случайная величина
- Вероятность попадания величины в интервал (a, b) равна

$$P(a \leq X \leq b) = \int_a^b p(x)dx,$$

где $p(x)$ – плотность распределения X ,

$$p(x) \geq 0, \quad \int_{-\infty}^{\infty} p(x)dx = 1$$

- Если поведение случайной величины определяется некоторым параметром, возникают условные плотности $p(x|\theta)$. Если рассматривать условную плотность как функцию от параметра

$$f(\theta) = p(x|\theta),$$

то принято говорить о т.н. функции правдоподобия

Нормальное распределение

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

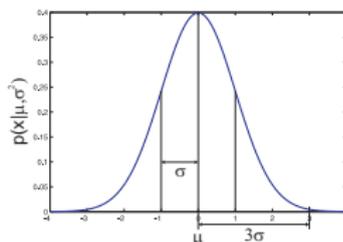
Графические
модели

Байесовские сети

Марковские сети

- Нормальное распределение играет важнейшую роль в математической статистике

$$X \sim \mathcal{N}(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$
$$\mu = \mathbb{E}X, \quad \sigma^2 = \mathbb{D}X \triangleq \mathbb{E}(X - \mathbb{E}X)^2$$



- Из центральной предельной теоремы следует, что сумма независимых случайных величин с ограниченной дисперсией стремится к нормальному распределению
- На практике многие случайные величины можно считать приближенно нормальными

Многомерное нормальное распределение

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Многомерное нормальное распределение имеет вид

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

где $\mu = \mathbb{E}X$, $\Sigma = \mathbb{E}(X - \mu)(X - \mu)^T$ — вектор математических ожиданий каждой из n компонент и матрица ковариаций соответственно

- Матрица ковариаций показывает, насколько сильно связаны (коррелируют) компоненты многомерного нормального распределения

$$\Sigma_{ij} = \mathbb{E}(X_i - \mu_i)(X_j - \mu_j) = \text{Cov}(X_i, X_j)$$

- Если мы поделим ковариацию на корень из произведений дисперсий, то получим коэффициент корреляции

$$\rho(X_i, X_j) \triangleq \frac{\text{Cov}(X_i, X_j)}{\sqrt{\mathbb{D}X_i \mathbb{D}X_j}} \in [-1, 1]$$

Особенности нормального распределения

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

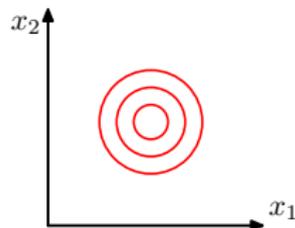
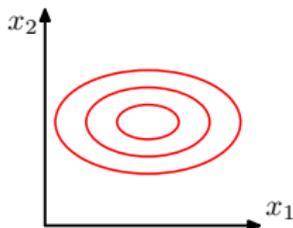
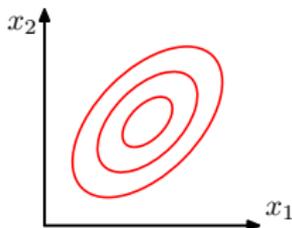
Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Нормальное распределение **полностью задается** первыми двумя моментами (мат. ожидание и матрица ковариаций/дисперсия)
- Матрица ковариаций неотрицательно определена, причем на диагоналях стоят дисперсии соответствующих компонент
- Нормальное распределение имеет очень легкие хвосты: большие отклонения от мат. ожидания практически невозможны. Это обстоятельство нужно учитывать при приближении произвольных случайных величин нормальными



Основная задача мат. статистики

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Нормальное распределение
Формула Байеса

Ликбез

Графические модели

Байесовские сети

Марковские сети

- Распределение случайной величины X известно с точностью до параметра θ
- Имеется выборка значений величины X , $\mathbf{x} = (x_1, \dots, x_n)$
- Требуется оценить значение θ
- Метод максимального правдоподобия

$$\hat{\theta}_{ML} = \arg \max f(\theta) = \arg \max p(\mathbf{x}|\theta) = \arg \max \prod_{i=1}^n p(x_i|\theta)$$

- Можно показать, что ММП (не путать с одноименной кафедрой) является асимптотически оптимальным при $n \rightarrow \infty$
- Обычно максимизируют не само правдоподобие, а его логарифм, т.к. это вычислительно проще (произведение плотностей по всем объектам переходит в сумму логарифмов плотностей)

Пример использования

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Пусть имеется выборка из нормального распределения $\mathcal{N}(x|\mu, \sigma^2)$ с неизвестными мат. ожиданием и дисперсией
- Выписываем логарифм функции правдоподобия

$$L(X|\mu, \sigma) = - \sum_{i=1}^n \frac{(x_i - \mu)^2}{2\sigma^2} - n \log \sigma - \frac{n}{2} \log(2\pi) \rightarrow \max_{\mu, \sigma}$$

$$\frac{\partial L}{\partial \mu} = - \sum_{i=1}^n \frac{(x_i - \mu)}{\sigma^2} = 0$$

$$\frac{\partial L}{\partial \sigma} = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^3} - \frac{n}{\sigma} = 0$$

$$\mu_{ML} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\sigma_{ML}^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 =$$

$$= \frac{1}{n} \sum_{i=1}^n \left(x_i - \frac{1}{n} \sum_{j=1}^n x_j \right)^2$$

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Нормальное распределение
Формула Байеса

Ликбез

Графические модели

Байесовские сети

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия**
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Условная вероятность

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Пусть X и Y — случайные величины с плотностями $p(x)$ и $p(y)$ соответственно
- В общем случае их совместная плотность $p(x, y) \neq p(x)p(y)$. Если это равенство выполняется, величины называют **независимыми**
- Условной плотностью называется величина

$$p(x|y) = \frac{p(x, y)}{p(y)}$$

- Смысл: как факт $Y = y$ влияет на распределение X . Заметим, что $\int p(x|y)dx \equiv 1$, но $\int p(x|y)dy$ не обязан равняться единице, т.к. относительно y это не плотность, а **функция правдоподобия**
- Очевидная система тождеств $p(x|y)p(y) = p(x, y) = p(y|x)p(x)$ позволяет легко переходить от $p(x|y)$ к $p(y|x)$

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}$$

Правило суммирования вероятностей

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Все операции над вероятностями базируются на применении всего двух правил
- Правило суммирования: Пусть A_1, \dots, A_k взаимоисключающие события, одно из которых **всегда происходит**. Тогда

$$P(A_i \cup A_j) = P(A_i) + P(A_j) \quad \sum_{i=1}^k P(A_i) = 1$$

- Очевидное следствие (формула полной вероятности): $\forall B$ верно $\sum_{i=1}^k P(A_i|B) = 1$, откуда

$$\sum_{i=1}^k \frac{P(B|A_i)P(A_i)}{P(B)} = 1 \quad P(B) = \sum_{i=1}^k P(B|A_i)P(A_i)$$

- В интегральной форме

$$p(b) = \int p(b, a) da = \int p(b|a)p(a) da$$

Условное и маргинальное распределения

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

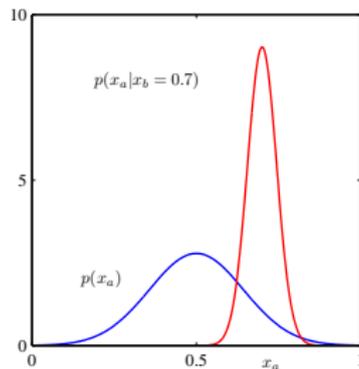
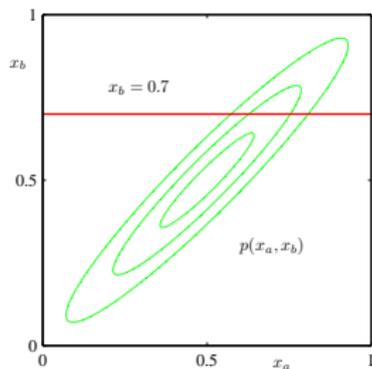
Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Плотность распределения интересующей нас компоненты x_a многомерной случайной величины (x_a, x_b) можно получить двумя способами в зависимости от имеющейся информации
- Если нам неизвестны значения остальных компонент, мы **маргинализуем** плотность по ним:
$$p(x_a) = \int p(x_a, x_b) dx_b$$
- Если значения других компонент нам известны, то мы **обуславливаем** плотность по ним: $p(x_a|x_b) = \frac{p(x_a, x_b)}{p(x_b)}$



Правило произведения вероятностей

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Правило произведения гласит, что любую совместную плотность всегда можно разбить на множители

$$p(a, b) = p(a|b)p(b) \quad P(A, B) = P(A|B)P(B)$$

- Аналогично для многомерных совместных распределений

$$p(a_1, \dots, a_n) =$$

$$p(a_1|a_2, \dots, a_n)p(a_2|a_3, \dots, a_n) \dots p(a_{n-1}|a_n)p(a_n)$$

- Можно показать (Jaunes, 1995), что правила суммирования и произведения вероятностей являются единственными возможными операциями, позволяющими рассматривать вероятности как промежуточную ступень между истиной и ложью

Априорные и апостериорные суждения

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Нормальное распределение
Формула Байеса

Ликбез

Графические модели

Байесовские сети

Марковские сети

- Предположим, мы пытаемся изучить некоторое явление
- У нас имеются некоторые знания, полученные до (лат. a priori) наблюдений/эксперимента. Это может быть опыт прошлых наблюдений, какие-то модельные гипотезы, ожидания
- В процессе наблюдений эти знания подвергаются постепенному уточнению. После (лат. a posteriori) наблюдений/эксперимента у нас формируются новые знания о явлении
- Будем считать, что мы пытаемся оценить неизвестное значение величины θ посредством наблюдений некоторых ее косвенных характеристик $x|\theta$

Формула Байеса

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Нормальное
распределение
Формула Байеса

Ликбез

Графические
модели

Байесовские сети

Марковские сети

- Знаменитая формула Байеса (1763 г.) устанавливает правила, по которым происходит преобразование знаний в процессе наблюдений
- Обозначим априорные знания о величине θ за $p(\theta)$
- В процессе наблюдений мы получаем серию значений $\mathbf{x} = (x_1, \dots, x_n)$. При разных θ наблюдение выборки \mathbf{x} более или менее вероятно и определяется значением правдоподобия $p(\mathbf{x}|\theta)$
- За счет наблюдений наши представления о значении θ меняются согласно формуле Байеса

$$p(\theta|\mathbf{x}) = \frac{p(\mathbf{x}|\theta)p(\theta)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\theta)p(\theta)}{\int p(\mathbf{x}|\theta)p(\theta)d\theta}$$

- Заметим, что знаменатель не зависит от θ и нужен исключительно для нормировки апостериорной плотности

Точечные оценки при использовании метода Байеса

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Нормальное распределение
Формула Байеса

Ликбез

Графические модели

Байесовские сети

Марковские сети

- Хотя строгое применение байесовского оценивания подразумевает работу со всем апостериорным распределением, на практике его часто заменяют на точечную оценку, взятую в точке максимума апостериорной плотности

$$\hat{\theta}_{MP} = \arg \max P(\theta|\mathbf{x}) = \arg \max P(\mathbf{x}|\theta)P(\theta) = \\ \arg \max (\log P(\mathbf{x}|\theta) + \log P(\theta))$$

- Это фактически регуляризация метода максимального правдоподобия, обеспечивающая компромисс между априорными ограничениями и наблюдениями

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Условная независимость случайных величин

Графические модели

Байесовские сети

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Условная независимость случайных величин

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Условная независимость случайных величин

Графические модели

Байесовские сети

Марковские сети

- Случайные величины x и y называются условно независимыми от z , если

$$p(x, y|z) = p(x|z)p(y|z)$$

- Другими словами вся информация о взаимозависимостях между x и y содержится в z
- Заметим, что из безусловной независимости не следует условная и наоборот
- Основное свойство условно независимых случайных величин

$$p(z|x, y) = \frac{p(x, y|z)p(z)}{p(x, y)} = \frac{p(x|z)p(y|z)p(z)}{p(x, y)} =$$

$$\frac{p(x|z)p(z)p(y|z)p(z)}{p(x, y)p(z)} = \frac{p(z|x)p(z|y)}{p(z)p(x)p(y)p(x, y)} = \frac{1}{Z} \frac{p(z|x)p(z|y)}{p(z)}$$

Пример

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

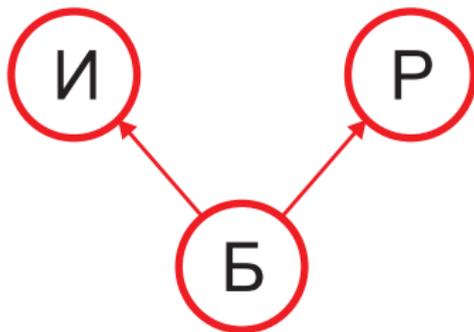
Условная независимость случайных величин

Графические модели

Байесовские сети

Марковские сети

- Рассмотрим следующую гипотетическую ситуацию: римские легионы во главе с императором атакуют вторгшихся варваров
- События «гибель императора» и «уничтожение Рима» не являются независимыми
- Однако, если нам дополнительно известен исход битвы с варварами, эти два события становятся независимыми
- В самом деле, если легионы битву проиграли, то судьба Рима мало зависит от того, был ли император убит в сражении



План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клик
 - Пример использования
 - Связь с байесовскими сетями

Классическая задача машинного обучения

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

- Задачу машинного обучения можно трактовать как восстановление неизвестных зависимостей между наблюдаемыми переменными X и скрытыми (латентными) переменными T . В случае обучения с учителем такое восстановление производится по обучающей выборке Y
- В классических задачах машинного обучения предполагается, что обучающая выборка сформирована из **однородных и независимых** объектов $Y = \{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^n$
- До недавнего времени вероятностные методы обработки данных ограничивались только таким простейшим случаем, а изложение каждого метода начиналось со слов «Предположим, что нам дана выборка из независимых одинаково распределенных случайных величин...»

Задачи со структурными ограничениями

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

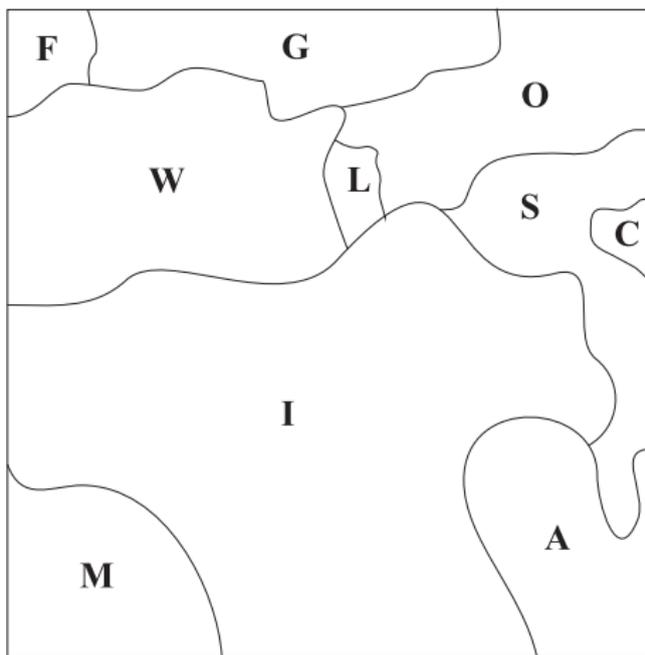
Байесовские сети

Марковские сети

- Во многих задачах взаимосвязи между наблюдаемыми и скрытыми переменными носят сложный характер
- В частности, между отдельными переменными существуют вероятностные зависимости
- Факт зависимости переменных друг от друга удобно отображать с помощью неориентированного графа (марковской сети)
- Если связи между переменными причинно-следственные, то их удобно отображать в виде ориентированных графов (байесовских сетей)
- Основным средством работы с графическими моделями служит аппарат теории вероятностей, в ее байесовской интерпретации

Простой пример

Задача о раскраске областей на плоскости так, чтобы никакие соседние не были окрашены в одинаковый цвет



Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Задачи со
структурными
ограничениями

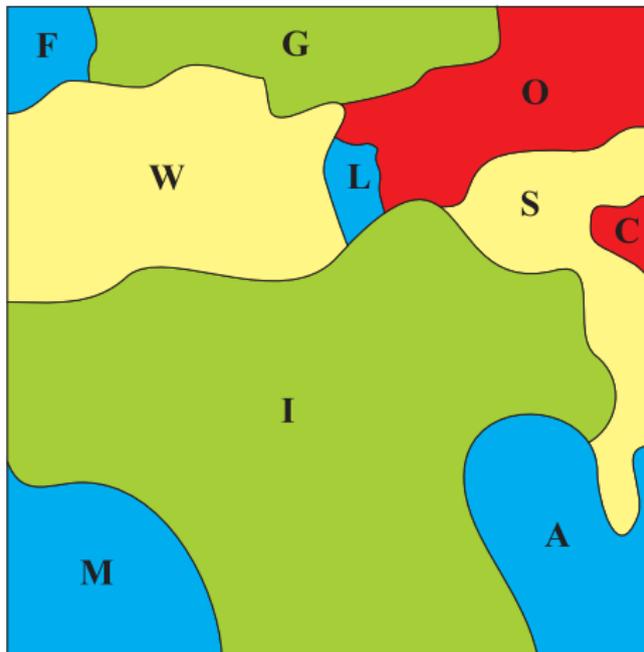
Основные
проблемы в
анализе
графических
моделей

Байесовские сети

Марковские сети

Простой пример

Задача о раскраске областей на плоскости так, чтобы никакие соседние не были окрашены в одинаковый цвет



Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Задачи со
структурными
ограничениями

Основные
проблемы в
анализе
графических
моделей

Байесовские сети

Марковские сети

Простой пример

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

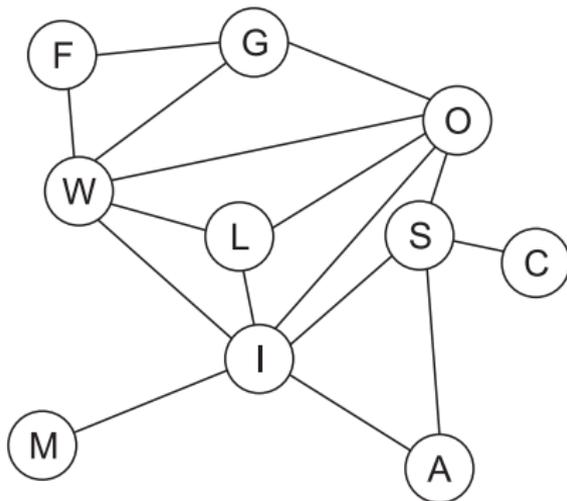
Задачи со
структурными
ограничениями

Основные
проблемы в
анализе
графических
моделей

Байесовские сети

Марковские сети

Такая задача легко формулируется в терминах графической модели, в которой каждая вершина графа может находиться в одном из четырех состояний



Вопрос залу: почему четырех?

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Графические модели

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

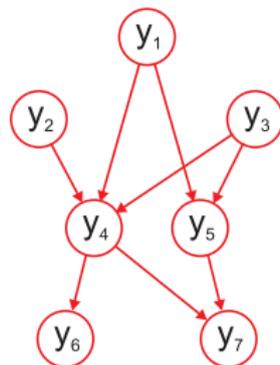
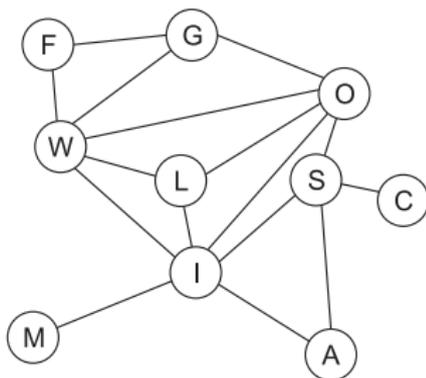
Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

- Графическая модель представляет собой ориентированный или неориентированный граф
- Вершины графа соответствуют переменным
- Ребра графа соответствуют вероятностным отношениям, определяющим непосредственные зависимости



Главные задачи в анализе графических моделей

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

Обозначим совокупность наблюдаемых переменных X , а ненаблюдаемых переменных T . Основными задачами в анализе графических моделей являются

- Подсчет условного распределения на значения отдельной скрытой переменной $p(t_i|X)$ —?
- Нахождение наиболее вероятной конфигурации скрытых переменных $p(T|X) \rightarrow \max_T$
- Оценка адекватности выбранной графической модели данным $p(X)$ —?

Трудности, возникающие при использовании графических моделей

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Задачи со структурными ограничениями

Основные проблемы в анализе графических моделей

Байесовские сети

Марковские сети

- Не во всех случаях существуют строгие алгоритмы вывода и обучения графических моделей
- Даже там, где они существуют, их применение может оказаться невозможно из-за высоких вычислительных требований и требований к памяти
- В настоящее время в мире активно разрабатываются приближенные эффективные методы обучения и принятия решения в графических моделях
 - Monte Carlo Markov chains
 - Variational bounds
 - Expectation propagation
 - Loopy belief propagation
 - Tree reweighted message passing
 - и др.

План

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа
Пример
использования

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Совместное распределение переменных

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

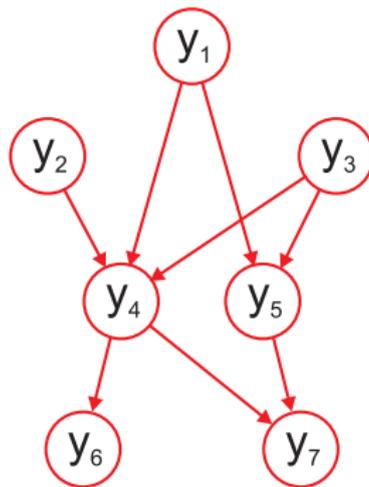
Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа

Пример
использования

Марковские сети



Совместное распределение системы переменных задается выражением

$$p(Y) = p(y_1, y_2, y_3, y_4, y_5, y_6, y_7) = \\ p(y_1)p(y_2)p(y_3)p(y_4|y_1, y_2, y_3)p(y_5|y_1, y_3)p(y_6|y_4)p(y_7|y_4, y_5).$$

Совместное и условные распределения

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

- В общем случае совместное распределение для ориентированного графа с n вершинами

$$p(Y) = \prod_{i=1}^n p(y_i | \text{pa}_i),$$

где pa_i — множество вершин-родителей y_i

- Обычно предполагается, что атомарные условные распределения $p(y_i | \text{pa}_i)$ известны
- Зная атомарные распределения, мы можем рассчитать (хотя бы теоретически) любые условные вероятности одних подмножеств переменных по другим подмножествам переменных

Вычисление условных распределений I

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

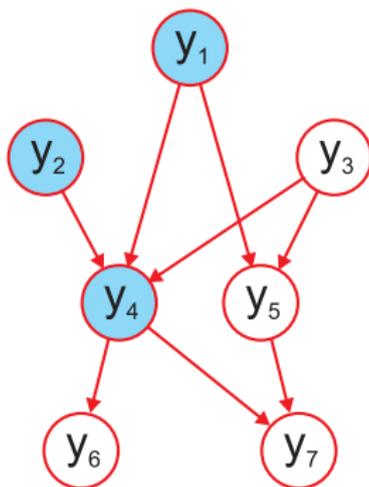
Факторизация байесовских сетей

Гри элементарных графа

Пример использования

Марковские сети

- Вернемся к иллюстрации графической модели из семи переменных
- Пусть нам необходимо найти распределение (y_5, y_7) при заданных значениях y_1, y_2, y_4 и неизвестных y_3, y_6



Вычисление условных распределений II

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

- По определению условной вероятности

$$p(y_5, y_7 | y_1, y_2, y_4) = \frac{p(y_1, y_2, y_4, y_5, y_7)}{p(y_1, y_2, y_4)}$$

- Расписываем знаменатель

$$p(y_1, y_2, y_4) = p(y_1)p(y_2)p(y_4 | y_1, y_2) = \{Sum\ rule\}$$

$$p(y_1)p(y_2) \int p(y_4 | y_1, y_2, y_3)p(y_3)dy_3$$

- Аналогично числитель

$$p(y_1, y_2, y_4, y_5, y_7) = p(y_1)p(y_2)p(y_4 | y_1, y_2)p(y_5 | y_1)p(y_7 | y_5, y_4) = p(y_1) \times p(y_2) \left(\int p(y_4 | y_1, y_2, y_3)p(y_3)dy_3 \right) \left(\int p(y_5 | y_1, y_3)p(y_3)dy_3 \right) p(y_7 | y_5, y_4)$$

- Для взятия возникающих интегралов обычно пользуются различными аппроксимационными методами
- Таким образом, условное распределение выражено через известные атомарные распределения вида $p(y_i | pa_i)$

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Особенности использования байесовских сетей

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

- По смыслу построения байесовские сети не могут содержать ориентированные циклы, т.к. это будет нарушать правило умножения вероятностей
- Главным достоинством графических моделей является относительно простое выделение условно-независимых величин, которое облегчает дальнейший анализ, позволяя значительно уменьшить количество факторов, влияющих на данную переменную
- В байесовских сетях сделать это несколько сложнее, чем в марковских

Граф 1

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

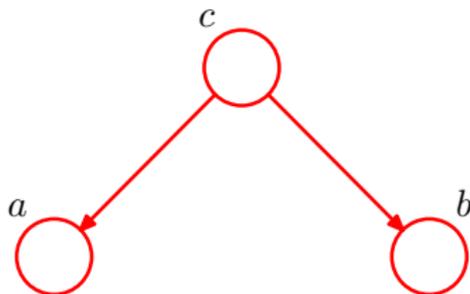
Байесовские сети

Факторизация
байесовских
сетей

Гри
элементарных
графа

Пример
использования

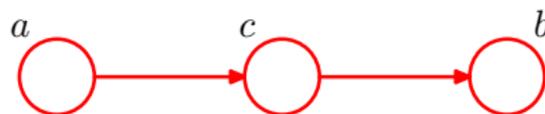
Марковские сети



- Аналогия: Рим (a), император (b) и варвары (c)
- Переменные a и b независимы при заданном c
- Возможна маргинализация (исключение переменной)

$$p(a, b) = \int p(a|c)p(b|c)p(c)dc \neq p(a)p(b)$$

Граф 2



- Аналогия: хорошая работа (a), премия (c), яхта (b)
- Переменные a и b независимы при заданном c
- Возможна маргинализация (исключение переменной)

$$p(a, b) = p(a) \int p(b|c)p(c|a)dc \neq p(a)p(b)$$

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа

Пример
использования

Марковские сети

Граф 3

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

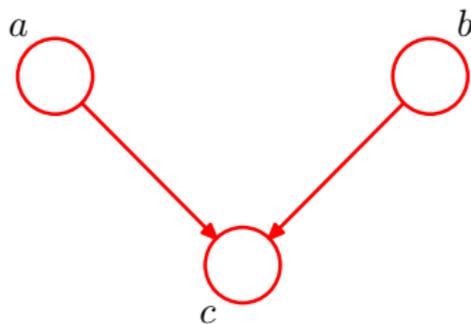
Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа

Пример
использования

Марковские сети



- Аналогия: вор (a), землетрясение (b) и сигнализация (c)
- Переменные a и b независимы, т.е. $p(a, b) = p(a)p(b)$, но не условно независимы!
- Зависимость $p(c|a, b)$ не может быть выражена через $p(c|a)$ и $p(c|b)$, хотя обратное верно

$$p(c|a) = \int p(c|a, b)p(b)db$$

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Три элементарных графа

Пример использования

Марковские сети

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Пример нестрогих вероятностных рассуждений I

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

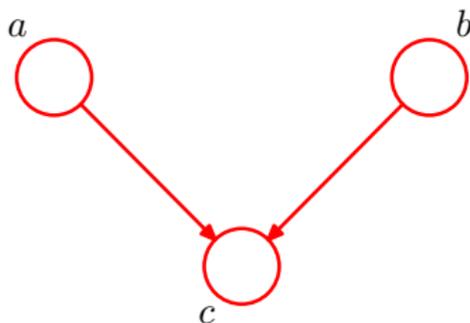
Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа

Пример
использования

Марковские сети



- Рассмотрим последний граф подробнее. Введем обозначения событий: «сигнализация сработала/не сработала» ($s/\neg s$), «вор есть/вора нет» ($v/\neg v$) и «землетрясение произошло/не произошло» ($z/\neg z$)
- Пусть $p(s|v, \neg z) = p(s|v, z) = 1$, $p(s|\neg v, z) = 0.1$, $p(s|\neg v, \neg z) = 0$, $p(v) = 2 \times 10^{-4}$, $p(z) = 10^{-2}$.
Графическая модель полностью определена

Пример нестрогих вероятностных рассуждений II

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа

Пример
использования

Марковские сети

Допустим, мы получили сигнал тревоги, т.е. имеет место событие s . Необходимо оценить вероятность того, что в квартире вор $p(v|s)$

$$p(v|s) = \frac{1}{Z}p(s|v)p(v) = \frac{p(s|v)p(v)}{p(s|v)p(v) + p(s|\neg v)p(\neg v)}$$

$$p(s|\neg v) = p(s|\neg v, z)p(z) + p(s|\neg v, \neg z)p(\neg z) = 10^{-3}$$

$$p(s|v) = 1$$

$$p(v|s) \approx \frac{1}{6}, \quad p(\neg v|s) \approx \frac{5}{6}, \quad Z \approx 1.2 \times 10^{-3}$$

Пример нестрогих вероятностных рассуждений III

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Факторизация байесовских сетей

Гри элементарных графа

Пример использования

Марковские сети

- Пусть теперь дополнительно стало известно, что произошло землетрясение (событие z). Как изменится вероятность того, что в квартире вор $p(v|s, z)$?

$$p(v|s, z) = \frac{1}{Z} p(s|v, z) p(v|z), \quad p(v|z) = p(v)$$

$$Z = p(s|v, z) p(v|z) + p(s|\neg v, z) p(\neg v|z) =$$

$$1 \times 2 \times 10^{-4} + 0.1 \times (1 - 2 \times 10^{-4}) = 0.1002$$

$$p(v|s, z) = 0.002, \quad p(\neg v|s, z) = 0.998$$

- Заметим, что события z и v перестали быть независимыми, и добавление сведений о значении z меняет знания о значении v . Это называется эффектом оправдания (explaining away)

Примеры байесовских сетей

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Факторизация
байесовских
сетей

Три
элементарных
графа

Пример
использования

Марковские сети

- Скрытые марковские модели
- Фильтр Калмана
- Экспертные системы
- Вероятностный РСА
- Смеси экспертов
- Факторный анализ
- и др.

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клика

Пример использования

Связь с байесовскими сетями

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клика
 - Пример использования
 - Связь с байесовскими сетями

Неориентированные графические модели

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

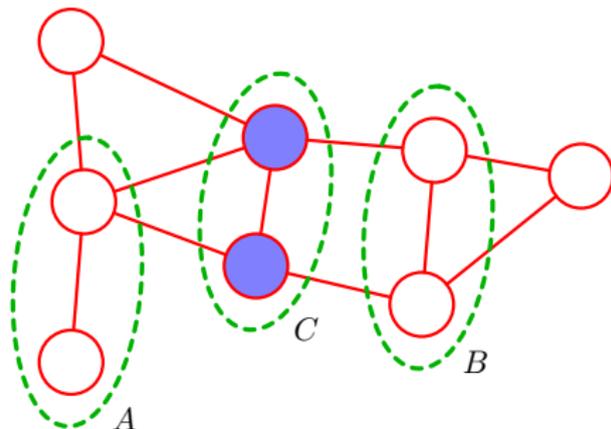
Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

- При использовании ориентированных графов определение условной независимости не очень просто
- В марковских сетях это проще. На рисунке A и B независимы при условии C
- Ребра графа связывают переменные, между которыми существуют непосредственные (а не опосредованные) зависимости



Факторизация в марковских сетях

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

- Пусть y_i и y_j независимы при условии, что все остальные переменные нам известны, т.е.
$$p(y_i, y_j | Y_{\{i,j\}}) = p(y_i | Y_{\{i,j\}})p(y_j | Y_{\{i,j\}})$$

- Это означает, что y_i и y_j не соединены ребром (иначе не было бы условной независимости)

- Запишем совместное распределение и применим правило умножения вероятностей

$$p(Y) = p(y_i, y_j | Y_{\{i,j\}})p(Y_{\{i,j\}}) = p(y_i | Y_{\{i,j\}})p(y_j | Y_{\{i,j\}})p(Y_{\{i,j\}})$$

- Таким образом, переменные, не соединенные ребрами, **входят в разные множители** совместного распределения

Потенциалы марковской сети

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Потенциалы и
энергия клик

Пример
использования

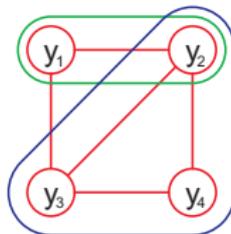
Связь с
байесовскими
сетями

- В общем виде совместное распределение значений элементов сети записывается с помощью неотрицательных потенциальных функций, определенных на максимальных кликах

$$p(Y) = \frac{1}{Z} \prod_C \psi_C(Y_C), \quad Z = \sum_Y \prod_C \psi_C(Y_C), \quad \psi_C(Y_C) \geq 0$$

- На рисунке синяя клика является максимальной, а зеленая — нет. Совместное распределение имеет вид

$$p(Y) = \frac{1}{Z} \psi_1(y_1, y_2, y_3) \psi_2(y_2, y_3, y_4)$$



Потенциальная функция не обязана иметь вероятностную природу, но чем она больше, тем более вероятны соответствующие значения переменных. Обычно потенциальные функции задаются пользователем исходя из априорных предпочтений тех или иных конфигураций переменных. Реже — настраиваются по данным

Энергетическая нотация

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик

Пример использования
Связь с байесовскими сетями

- Иногда удобно ввести обозначение $\psi_C(Y_C) = \exp(-E_C(Y_C))$, где $E_C(Y_C)$ имеет смысл энергии
- Тогда задача нахождения наиболее вероятного состояния системы сводится к задаче минимизации полной энергии системы

$$\arg \max p(Y) = \arg \max \frac{1}{Z} \prod_C \psi_C(Y_C) =$$

$$\arg \max \exp \left(- \sum_C E_C(Y_C) \right) = \arg \min \sum_C E_C(Y_C)$$

- Заметим, что в отличие от байесовских сетей для полного задания графической модели необходимо знать (или конструктивно уметь подсчитывать) нормировочную константу Z

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

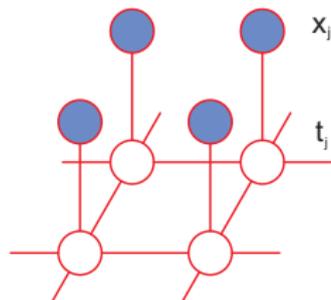
Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клик
 - Пример использования
 - Связь с байесовскими сетями

Фильтрация изображений



- Рассмотрим задачу фильтрации изображения. Пусть $x_i \in \{-1, 1\}$ — наблюдаемые пиксели бинарного изображения, а $t_i \in \{-1, 1\}$ — истинные значения пикселей
- Введем энергию системы

$$E(X, Y) = h \sum_i t_i - \beta \sum_{(i,j) \in E} t_i t_j - \eta \sum_i t_i x_i,$$

где $h \in \mathbb{R}$ позволяет отразить априорные предпочтения в пользу того или иного цвета (например, указать, что желтый цвет встречается чаще, чем синий), $\beta > 0$ выражает степень зависимости между соседними пикселями, а $\eta > 0$ показывает интенсивность шума

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями

Разметка областей I

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

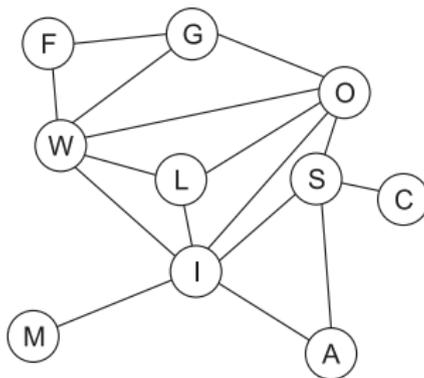
Марковские сети

Потенциалы и
энергия клик

Пример
использования

Связь с
байесовскими
сетями

Вернемся к примеру со странами



Совместная плотность задается формулой

$$p(X) = \frac{1}{Z} \psi_1(F, G, W) \psi_2(G, O, W) \psi_3(W, O, L, I) \psi_4(I, S, O) \times \\ \times \psi_5(S, I, A) \psi_6(S, C) \psi_7(M, I)$$

Разметка областей II

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

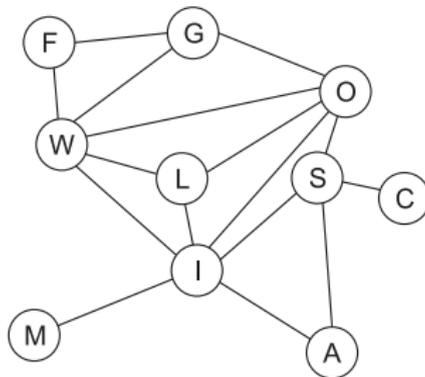
Байесовские сети

Марковские сети

Потенциалы и
энергия клик

Пример
использования

Связь с
байесовскими
сетями



- Предположим, что переменные могут принимать одно из четырех значений $\{red, yellow, blue, green\}$
- Требование несовпадающих цветов областей эквивалентно условию равенства нулю потенциала, если хотя бы два его аргумента имеют одинаковое значение, например $\psi_5(red, blue, red) = 0$

Разметка областей III

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

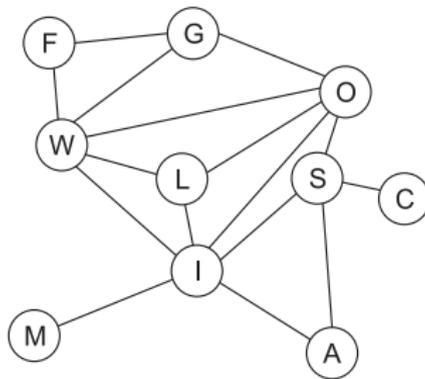
Байесовские сети

Марковские сети

Потенциалы и энергия клик

Пример использования

Связь с байесовскими сетями



Мы можем снизить число нежелаемых цветовых переходов (например, из желтого в красный) снизив соответствующие значения потенциалов $\psi_7(\text{red}, \text{yellow})$, $\psi_7(\text{yellow}, \text{red})$, $\psi_6(\text{red}, \text{yellow})$ и $\psi_6(\text{yellow}, \text{red})$

Разметка областей IV

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

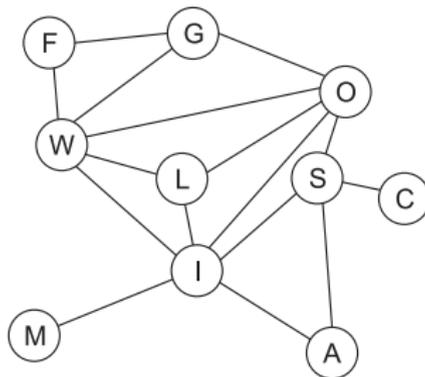
Байесовские сети

Марковские сети

Потенциалы и
энергия клик

Пример
использования

Связь с
байесовскими
сетями



Мы можем искусственно способствовать окраске отдельных регионов в выбранные цвета, вводя индивидуальные множители, например

$$\psi_1(F, G, W) = \phi_1(F, G, W)\phi_2(F)\phi_3(W).$$

Теперь можно увеличить значение $\phi_2(\text{blue})$ и $\phi_3(\text{green})$, чтобы получить политическую карту, привычную российскому глазу

Примеры марковских сетей

Лекция 1.

Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Потенциалы и
энергия клик

Пример
использования

Связь с
байесовскими
сетями

- Изображения
- Социальные сети
- Случайные поля
- Карты сайтов
- и др.

План

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик
Пример использования

Связь с байесовскими сетями

- 1 Содержание и задачи курса
- 2 Полезные сведения из линейной алгебры и теории оптимизации
- 3 Основные вероятностные понятия
 - Нормальное распределение
 - Формула Байеса
- 4 Ликбез
 - Условная независимость случайных величин
- 5 Графические модели
 - Задачи со структурными ограничениями
 - Основные проблемы в анализе графических моделей
- 6 Байесовские сети
 - Факторизация байесовских сетей
 - Три элементарных графа
 - Пример использования
- 7 Марковские сети
 - Потенциалы и энергия клик
 - Пример использования
 - Связь с байесовскими сетями

Марковские vs. Байесовские сети

Лекция 1.
Байесовские и
марковские сети

Ветров

Содержание и
задачи курса

Полезные
сведения из
линейной
алгебры и теории
оптимизации

Основные
вероятностные
понятия

Ликбез

Графические
модели

Байесовские сети

Марковские сети

Потенциалы и
энергия клик

Пример
использования

Связь с
байесовскими
сетями

Сходства и различия двух типов графических моделей

Свойство	Марковские сети	Байесовские сети
Форма	Произв. потенциалов	Произв. потенциалов
Потенциалы	Произвольные	Усл. вероятности
Циклы	Разрешены	Запрещены
Нормировка	$Z = ?$	$Z = 1$
Условная нез-ть	Легко проверяема	Сложнее
Полнота	нет	нет
Анализ	MCMC, BP, и т.д..	Сводит к марковским

Существующее положение дел

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клик
Пример использования

Связь с байесовскими сетями

- На сегодняшний день существуют эффективные алгоритмы (sum-product, max-product) анализа ациклических графов (деревьев), решающие все три основные задачи анализа графических моделей
- Частным случаем деревьев являются графы-цепочки, характеризующие, например, сигналы во времени
- В случае наличия циклов сложность точных алгоритмов резко возрастает
- Для анализа графов с циклами в основном используются приближенные методы (loopу BP, EP, MCMC)
- В некоторых частных случаях для анализа циклических сетей существуют эффективные точные алгоритмы, например, разрезы графов (лекция 3)

Сведение байесовских сетей к марковским

Лекция 1.

Байесовские и марковские сети

Ветров

Содержание и задачи курса

Полезные сведения из линейной алгебры и теории оптимизации

Основные вероятностные понятия

Ликбез

Графические модели

Байесовские сети

Марковские сети

Потенциалы и энергия клика

Пример использования

Связь с байесовскими сетями

- Наиболее разработаны в настоящее время методы анализа марковских сетей
- Байесовскую сеть можно легко свести к марковской с потерей информации, переживнив всех родителей (морализация)
- Заметим, что в приведенном примере все полезные свойства оказались потеряны, и мы получили клику, в которой все зависят от всех

