# MCMC для dummies

Винни-Пух

14 февраля 2017 г.

# Предисловие

План:

1. Энтропия/Джини

Распределения максимизирующие энтропию? что-то про ROC кривые до кучи?

2. Одиноко стоящий дуб

Типичное заданичие: Вырастить дерево согласно такому-то критерию. Сюда борьбу с NA. Сюда же регуляризацию? Или отдельно?

3. Логит-модель

Логистическое распределение? Перевод y=0/1 в y=-1/1. Максимум правдоподобия в минимум штрафа? Предельные эффекты?

4. Мини-мими-лес

Типичное: Два-три дерева. По ним построить прогноз/оценить важность переменных. Что еще?

5. Регуляризация.

Общая идея. Парадокс James-Stein. Для среднего, для регрессии, для дерева. L1 и L2.

6. Про кросс-валидацию?

Как это делать руками? Какие тут теоретические задачи?

Упр: Дано одно-два-три дерева. И 5 наблюдений. Посчитать кросс-валидационную ошибку.

Упр: На наборе данных в 5 наблюдений подобрать параметр жесткости с помощью кросс-валидации.

7. Несколько практических упражнений.

Упр: сделайте с дефолтными параметрами и ответьте на все подробности про алгоритм
тут решения в python/R.

Упр: Нарисуйте дерево номер 5.

Из теории:

- определения
- табличка с параметрами xgboost, rforest
- несколько практик подбора параметров

```r
library("knitr") # грамотное программирование
library("tikzDevice") # сохранение графиков в формате tikz

library("tidyverse") # Хэдли на нашей стороне
library("xtable")

theme_set(theme_bw()) # чёрно-белая тема для графиков
```

# Глава 1

# Неразобранные :)

**1.1** Шахматный конь начинает в клетке A1. Каждый свой ход он выбирает равновероятно из возможных. Какова вероятность того, что через много-много ходов он окажется в клетке H8? Сколько в среднем длится путь от клетки A1 до клетки A1?

**1.2** Случайные величины $X_i$ независимы и одинаково распределены с табличкой

| $X$ | 1 | 2 | 6 |
|---|---|---|---|
| $\mathbb{P}()$ | $\beta$ | $2\beta$ | $1-3\beta$ |

Известно, что $X_1 = 1$, $X_2 = 2$, $X_3 = 2$, $X_4 = 4$.

1. Найдите оценку $\hat{\beta}$ методом моментов
2. Найдите оценку $\hat{\beta}$ методом максимального правдоподобия
3. Предположим, что $\beta$ равномерно на отрезке $[0; 1/3]$. Найдите апостериорную условную функцию плотности $\beta$ с учётом полученных наблюдений. С какой функцией она совпадает?
4. Предположим, что $\beta$ имеет функцию плотности $f(t) = 18t$ на отрезке $[0; 1/3]$. Найдите апостериорную функцию плотности $\beta$.

**1.3** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для биномиального распределения $Bin(n, p)$ из равновероятного на множестве $\{0, 1, 2, \ldots, n\}$

**1.4** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для биномиального распределения $Bin(n, p)$ из симметричного случайного блуждания на $\mathbb{Z}$

**1.5** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для геометрического распределения $Geom(p)$ из симметричного случайного блуждания на $\mathbb{Z}$

**1.6** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для пуассоновского распределения $Pois(\lambda)$ из симметричного случайного блуждания на $\mathbb{Z}$

**1.7** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для функции плотности $\pi(x) \sim \exp(-x^2)(3 + x^2 + \cos x)$ из нормального $N(0, 1)$. Из нормального $N(0, \sigma^2)$

**1.8** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для функции плотности $\pi(x) \sim \exp(-x^2)(3 + x^2 + \cos x)$ из случайного блуждания $X_{t+1} = X_t + \varepsilon_t$, где $\varepsilon_t \sim N(0, 1)$. Вариант с $N(0, \sigma^2)$

**1.9** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для стандартного нормального распределения $N(0, 1)$ из случайного блуждания $X_{t+1} = X_t + \varepsilon_t$, где $\varepsilon_t \sim U[-1, 1]$

**1.10** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для двумерного нормального распределения $N(0, A)$, $A = \begin{pmatrix} 4 & -1 \\ -1 & 2 \end{pmatrix}$ из случайного блуждания $X_{t+1,i} = X_{t,i} + \varepsilon_{t,i}$, где $\varepsilon_{t,i} \sim U[-1, 1]$.

**1.11** Используя алгоритм Метрополиса-Хастингса сгенерите выборку для двумерного распределения с функцией плотности $p(x, y) = \exp(-4x^2 - 6y^2 + 2x - y + xy)$, $x > 0, y > 0$ из случайного блуждания $X_{t+1,i} = X_{t,i} + \varepsilon_{t,i}$, где $\varepsilon_{t,i} \sim U[-1, 1]$.

**1.12** Рассмотрим модель

$$y = X\beta + u,$$

где $u_i$ независимы и $\mathcal{N}(0; \sigma^2)$.

Метод гребневой регрессии предполагает минимизацию функции

$$Q(\beta) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{k} \beta_j^2.$$

Рассмотрим байесовский подход к регрессии. Предположим, что априорное распределение имеет вид $\sigma^2 \sim InvGamma(a, b)$, $\beta_j | \sigma^2 \sim \mathcal{N}(0; c)$.

При каких $a$, $b$ и $c$ апостериорная мода $\hat{\beta}_{MAP}$ совпадёт с $\hat{\beta}_{Ridge}$?

**1.13** Рассмотрим модель

$$y = X\beta + u,$$

где $u_i$ независимы и $\mathcal{N}(0; \sigma^2)$.

Метод LASSO предполагает минимизацию функции

$$Q(\beta) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \lambda \sum_{j=1}^{k} |\beta_j|.$$

Рассмотрим байесовский подход к регрессии. Предположим, что априорное распределение имеет вид $\sigma^2 \sim InvGamma(a, b)$, $\beta_j | \sigma^2 \sim DoubleExp(c)$.

При каких $a$, $b$ и $c$ апостериорная мода $\hat{\beta}_{MAP}$ совпадёт с $\hat{\beta}_{LASSO}$?

Глава 2

# Решения и ответы к избранным задачам

1.1.

1.2.

1.3.

1.4.

1.5.

1.6.

1.7.

1.8.

1.9.

1.10.

1.11.

1.12.

1.13.

# Список обозначений

# Оглавление


1 Неразобранные :) 5

2 Решения и ответы к избранным задачам 7