Привет :) 2021-12-17

Алгоритм Метрополиса-Гастингса.

Байесовский подход.

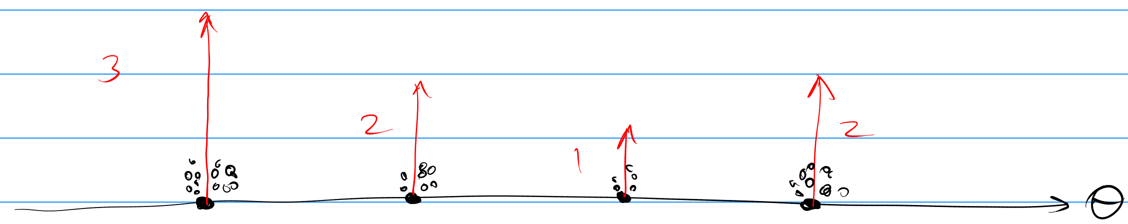
- ✓ модель: $(y|\theta) \sim$ модель
- ✓ данные y
- ✓ априорное мнение $\theta \sim$ prior, $p(\theta)$

„именное" распр-ие

→ ? апостер-ое $p(\theta|y)$

большая выборка, сходящаяся к апостер-му распределению.

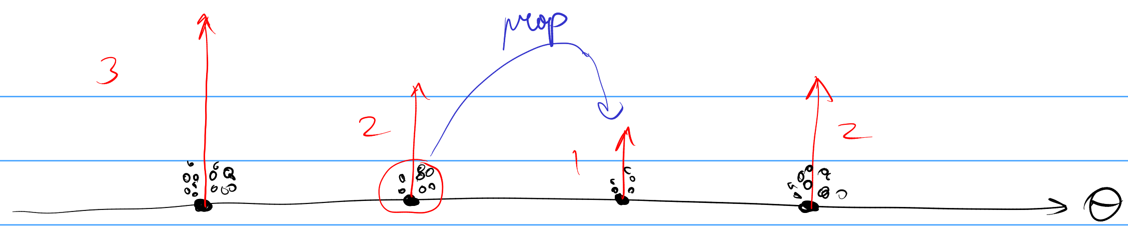
$\boxed{\text{Знаем } p(\theta|y) \text{ с точностью до константы!}}$



легко:

$$p(\theta|y) = \frac{p(\theta,y)}{p(y)} = \frac{p(y|\theta)\cdot p(\theta)}{p(y)} \propto \underbrace{p(y|\theta)}_{\text{модель ф. правд.}}\underbrace{p(\theta)}_{\text{апр. мнени.}}$$

найти $p(\theta|y)$ с точностью до константы!



Алг Метрополиса-Гастингса (со случайным блужданием)

1. Берем θ_0 любое из носителя $p(\theta)$.
[можно сгенерировать случай из априорного распр-ия]

$i = 0$

2. Находясь в точке θ_i генерируем предложение

$$\theta_{prop} = \theta_i + u$$

u генерируется незав от θ_i
типичный вар-т для непр-х пар-в $u \sim N(0; \delta^2)$

3. Рассчитаем [свертим формулу] вер-сть одобрения предложения. $P_{accept}(\theta_i \to \theta_{prop}) = a$

параметр алгоритма

4. $$\theta_{i+1} = \begin{cases} \theta_{prop} & \text{с вер-стью } a \\ \theta_i & \text{с вер-стью } 1-a. \end{cases}$$

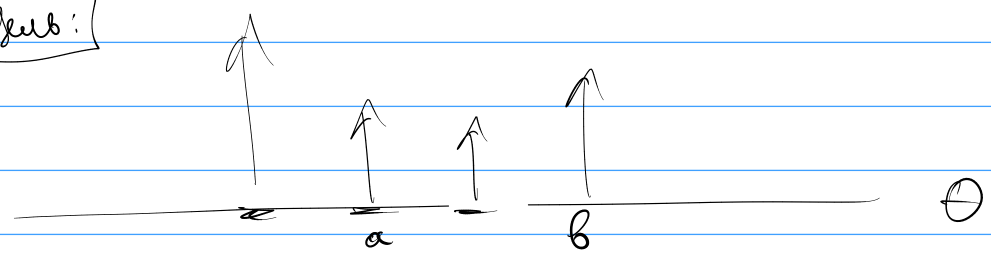
[не инт. сам по себе]

$i \leftarrow i+1$

$$a = \min\left\{1, \frac{post(\theta_{prop}) \cdot \boxed{prop(\theta_{prop} \to \theta_0)}}{post(\theta_i) \cdot \boxed{prop(\theta_i \to \theta_{prop})}}\right\}$$

это ровно 1, если u распределено симм-но около 0.

Цель:



распр-ие $\theta_i \xrightarrow{dist} p(\theta|y)$

необходимое условие

! Правильный ответ портить не надо! $x \neq a$

$i \to \infty$

легко посчитать

$$\frac{post(a)}{post(b)} = \frac{P(\theta_{i+1}=a|y)}{P(\theta_{i+1}=b|y)} =$$

prop(x→a) — вер-сть предложения перехода из x в a

$$= \frac{\sum_{x\neq a} P(\theta_i=x|y)\cdot prop(x\to a)\cdot accept(x\to a) + \sum_{x\neq a} P(\theta_i=a|y)\cdot prop(a\to x)\cdot(1-accept) + P(\theta_i=a|y)\cdot prop(a\to a)}{\sum_{x\neq b} P(\theta_i=x|y)\cdot prop(x\to b)\cdot accept(x\to b) + \sum_{x\neq b} P(\theta_i=b|y)\cdot prop(b\to x)\cdot(1-accept(b\to x)) + P(\theta_i=b)\cdot prop(b\to b)}$$

accept(x→a) — вер-сть одобрения предложения из x→a

необходимое

$$\frac{post(a)}{post(b)} = \frac{\sum_{x\neq a} post(x)\cdot prop(x\to a)\cdot acc(x\to a) + \sum_{x\neq a} post(a)\cdot prop(a\to x)\cdot(1-acc(a\to x)) + post(a)\cdot prop(a\to a)}{\ldots}$$

...

A условие:

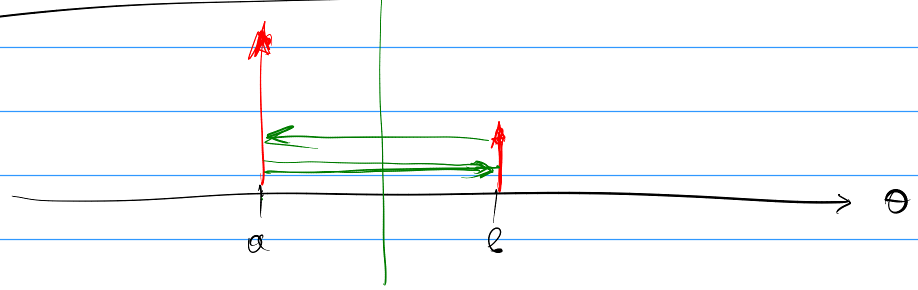
[достаточное для необходимого]

$$\boxed{post(x)\cdot prop(x\to a)\cdot accept(x\to a) = post(a)\cdot prop(a\to x)\cdot accept(a\to x)} \quad \forall x, a$$

$$\frac{post(a)}{post(b)} = \frac{post(a)\cdot prop(a\to a) + \sum_{x\neq a} post(a)\cdot prop(a\to x)}{—//—} = \frac{post(a)\cdot \overbrace{\sum_x prop(a\to x)}^{=1}}{post(b)\cdot\sum_x prop(b\to x)}$$

$$post(b)\cdot prop(b\to a)\cdot accept(b\to a) = post(a)\cdot prop(a\to b)\cdot accept(a\to b) \quad \forall a, b$$

Бухгалтерский баланс



Чтобы побыстрее обойти весь носитель $p(\theta|y)$ мы хотим $accept(a\to b)\uparrow$

(2) $\theta_{prop} = \theta_i + u$ u генер. кусков от θ_i

типичное: $u\sim N(0;\sigma^2)$

если закон распр. u симметр-н около 0, то

$$prop(a\to b) = prop(b\to a) \qquad [!]$$

и балансовое ур-ие проще:

$$\underset{0,2}{post(a)}\cdot \underset{1}{accept(a\to b)} = \underset{0,5}{post(b)}\cdot \underset{2/5}{accept(b\to a)}$$

$$accept(a\to b) = accept(b\to a)\cdot \frac{post(b)}{post(a)}$$

$$\boxed{accept(a\to b) = \begin{cases} 1, & \text{если } post(b)/post(a) \ge 1 \\ \dfrac{post(b)}{post(a)} & \text{если } \dfrac{post(b)}{post(a)} < 1 \end{cases}}$$

Про AIC

!! ☺ все модели ошибочны!

реальность

Модель A (θ)

Модель Б (θ)

Модель C (θ)

$KL(truth \,\|\, Mod\, A)$?

$KL(truth \,\|\, Mod\, B)$?

$KL(truth \,\|\, Mod\, C)$?

некоторые утвержд.

$$KL(truth \,\|\, Mod\, B) - KL(truth \,\|\, Mod\, A) \approx$$

$$\approx \frac{AIC_B - AIC_A}{2}$$

$$AIC_A = 2k_A - 2\ln L_A$$

k_A - число пар-в в модели A

$\ln L_A$ - логарифм правдоп-ия в модели A.

Generalized Additive Models: Introduction with R

↳ в аппендиксе!

Simon Wood