AS in ML

2020-10-09

Видно ли? Слышно ли?

Слышно ли? ×2 ? !!

1. Линейная регрессия?
2. Норм-ое распр-ие. (Гаусс)
3. Норм-ое распр-ие (Херинг. Максвелл)

Модель: y - зависимая пер-ая

$$y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

предикторы / регрессоры

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad b = \begin{pmatrix} \vdots \end{pmatrix} \quad c = \begin{pmatrix} \vdots \end{pmatrix}$$

набл.

случ-ая сост-ая: $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$

ненабл.

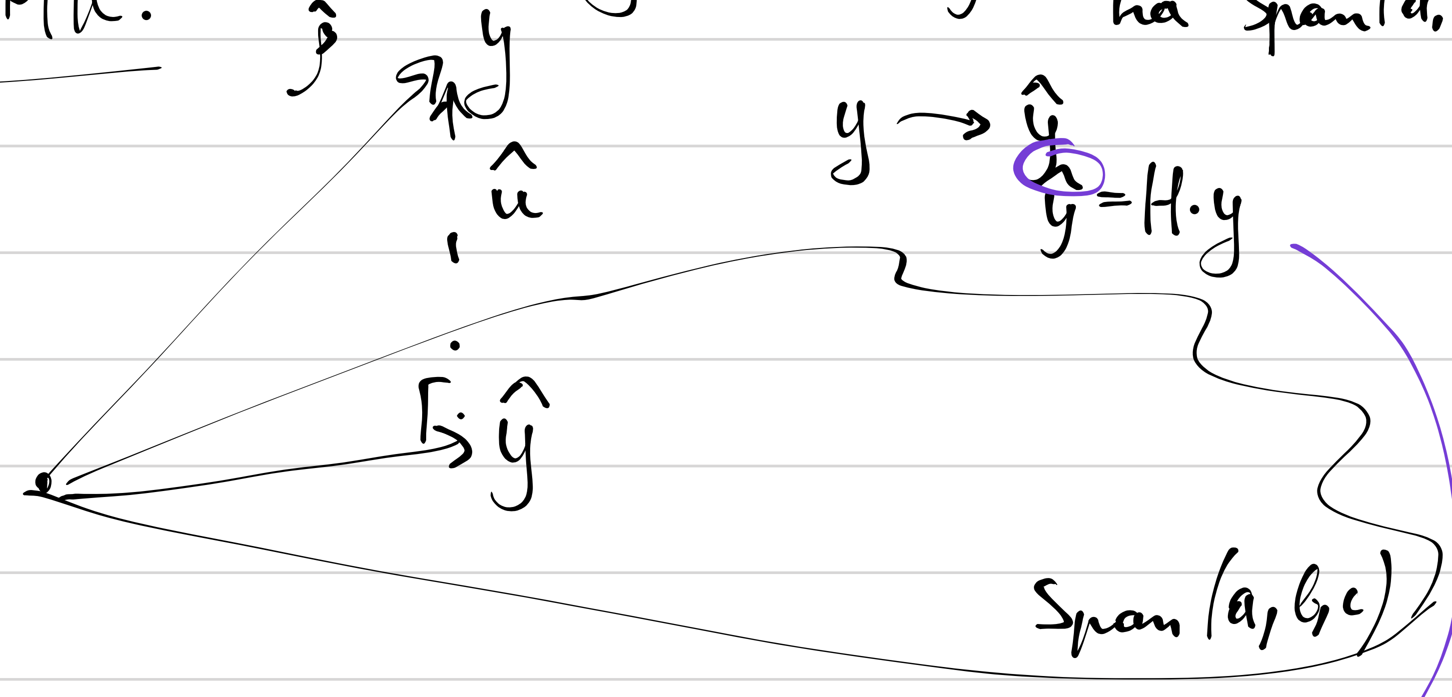
Скал: $y_i = \beta_1 \cdot a_i + \beta_2 \cdot b_i + \beta_3 \cdot c_i + u_i$

часто: $a_i = 1$

Вектор: $y = \beta_1 \cdot a + \beta_2 \cdot b + \beta_3 \cdot c + u$

$$X = \begin{bmatrix} a_1 & b_1 & c_1 \\ \vdots & \vdots & \vdots \\ a_n & b_n & c_n \end{bmatrix} \quad \beta = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix} \quad y = X \cdot \beta + u$$

МНК: $\min\limits_{\hat\beta} \|y-\hat y\|^2 \Rightarrow \hat y$ - проекция y на Span(a, b, c)



$y \to \hat y$ $\qquad \mathbb{R}^n$

$\hat y = H \cdot y$

Регрессия: $\hat y = \hat\beta_1 a + \hat\beta_2 b + \hat\beta_3 c \in \mathrm{Span}(a, b, c)$

$H = X(X^T X)^{-1} X^T$ hat-matrix

Модель: $y = \beta_1 \cdot a + \beta_2 \cdot b + \beta_3 c + u$

ошибки прогнозов/остатки

$\hat u = y - \hat y$ $\qquad \hat u \perp \mathrm{Span}(a, b, c)$

$\hat u \perp a, b, c$ - условия 1-го порядка $\Rightarrow \hat\beta_1, \hat\beta_2, \hat\beta_3$

$$\begin{cases} \langle \hat u, a \rangle = 0 \\ \langle \hat u, b \rangle = 0 \\ \langle \hat u, c \rangle = 0 \end{cases} \qquad X^T \cdot \hat u = 0$$

векторы $\qquad$ матрица

$$\begin{pmatrix} a_1 & \dots & a_n \\ b_1 & \dots & b_n \\ c_1 & \dots & c_n \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$X_{n \times k}$ $(X_{n \times 3})$

$X^T(y - \hat y) = 0$

$X^T(y - X \cdot \hat\beta) = 0$ (опасное желание сократить X)

$X^T y = X^T X \hat\beta$

если $X^T X$ обратима $\qquad \underset{k\times n}{X^T} \underset{n \times k}{X} \leftarrow (k \times k)$

$$\boxed{\hat\beta = (X^T X)^{-1} X^T y.}$$

$\hat y = X \cdot \hat\beta = \boxed{X(X^T X)^{-1} X^T y}$

$\hat y$ - проекция y на Span(столбцов X)

Тривиальная модель

$y_i = \beta_1 \cdot 1 + u_i$ $\qquad a = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$ $\qquad X = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}$

$$\hat\beta_1 = (X^TX)^{-1}X^Ty = (n)^{-1}\cdot\sum_{i=1}^{n} y_i = \frac{\sum y_i}{n} = \bar y$$

$X^T = (1\ 1 \dots 1)$ $\qquad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$

$\hat y_i = \hat\beta_1 = \bar y$ $\qquad$ МНК $\Rightarrow \hat\beta_1 = \bar y$

Почему расп-ие Гаусса
наз-ся расп-м Гаусса?

истина $\qquad$ пок-ия свид-ей

? $\longleftarrow$ ✓ ✓ ✓

Два лагеря:
→ берём среднее!
(берём медиану)
→ (Кеплер): нельзя брать среднее!

вы берёте среднее $\cancel{\Rightarrow}$ наблюдения равновозможны $\Rightarrow$

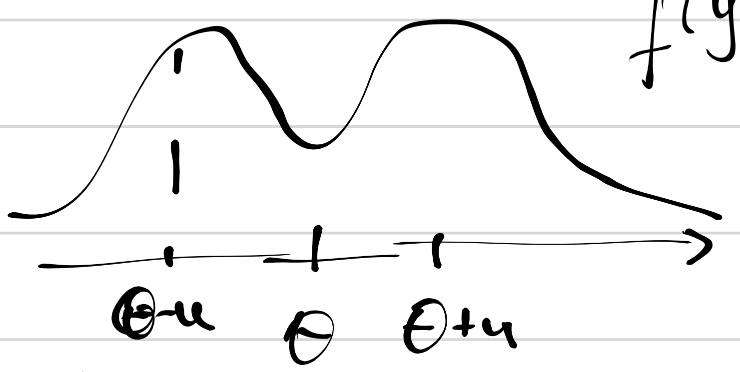
Гаусс

$\Rightarrow$ это явная ерунда!

Гаусс.

набл-ия: $y_1 \dots y_n$

1. Незав-мы, один-во распр-ы.
2. $f(y_i)$ закон распр-ия y_i симм отн-но θ



3. $\hat{\theta} = \bar{y}$ (МНК)
4. $\hat{\theta}$ - это оценка метода максим-го правдоподобия.

[техн]: плотность дифф-ма.

При параметре равном $\bar{y}$ вер-сть данных измерений максим-а.

$\boxed{\text{И что?}}$

MaxLik

$$\max_{\theta} \ln f(y_1 \dots y_n)$$

по предп 1.

$$\max_{\theta} \sum_{i=1}^{n} \ln f(y_i|\theta) \leftarrow \ell(y_1 \dots y_n|\theta)$$

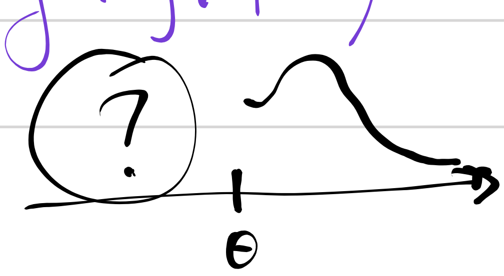
$$\frac{\partial \ell}{\partial \theta} = \sum_{i=1}^{n} \frac{f'_{\theta}(y_i|\theta)}{f(y_i|\theta)} = 0 \Rightarrow \hat{\theta}$$

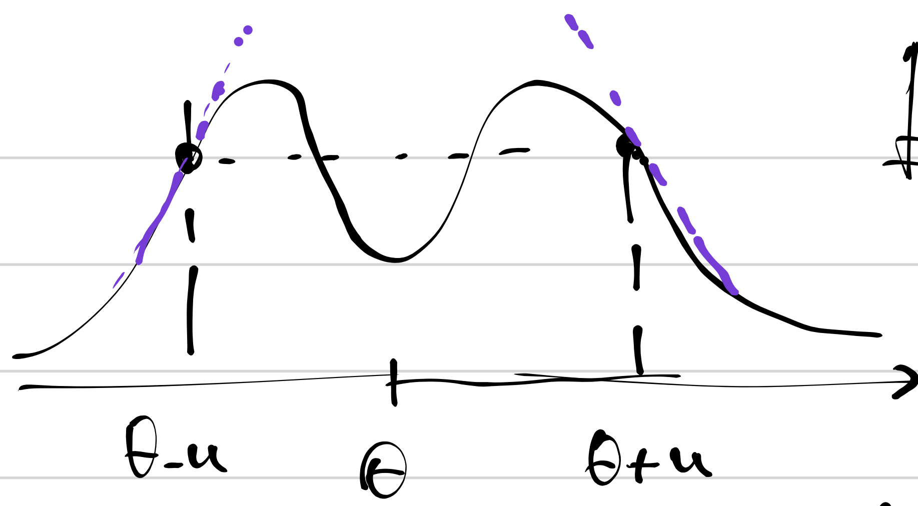
мы знаем! $\hat{\theta} = \bar{y}$

$$g(y_i|\theta) = \frac{f'_{\theta}(y_i|\theta)}{f(y_i|\theta)}$$

f - симм. отн-но θ

?





$f(y) = f(y|\theta)$

$f(\theta-u) = f(\theta+u)$

$f'(\theta-u) = -f'(\theta+u)$

Условие первого порядка:

$$\sum_{i=1}^{n} \frac{f'_\theta(y_i|\theta)}{f(y_i|\theta)} = 0.$$

$\hat{\theta} = \bar{y}$

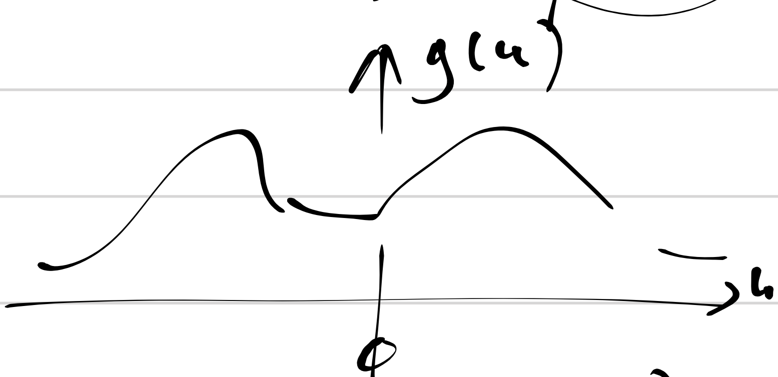
тождество при $\forall y_i$

$$\sum_{i=1}^{n} \frac{f'_\theta(y_i|\bar{y})}{f(y_i|\bar{y})} \equiv 0$$ при $\forall y_i$

n свед-лей: $\to n-1$ $\quad +1$ $\quad \theta+1$

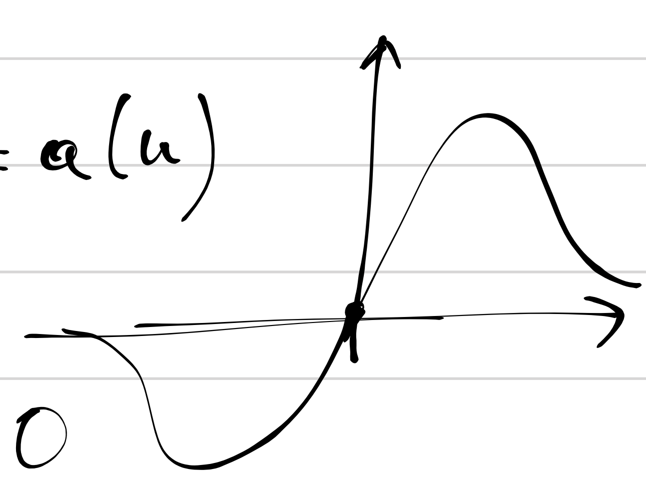
$\to 1$ $\quad -(n+1)$ $\quad \theta-(n-1)$

$g(u) = f(\theta+u)$



$$\sum_{i=1}^{n} \frac{g'(u_i)}{g(u_i)} = 0$$ (при $\forall \sum u_i = 0$)

$$\frac{g'(u)}{g(u)} = a(u)$$



$\forall n$

$(n-1)\cdot a(1) - a(n-1) = 0$

$a(n-1) = (n-1)\cdot a(1)$ $\quad \forall n$

$a(q) = q\cdot a(1)$ $\quad \forall q \in \mathbb{Q}$

$\boxed{a(z) = z\cdot a(1)}$ $\quad \forall z \in \mathbb{R}$

$a(z) = k\cdot z$

$$\frac{g'(u)}{g(u)} = \left(\ln g(u)\right)' = a(u) = c\cdot u$$

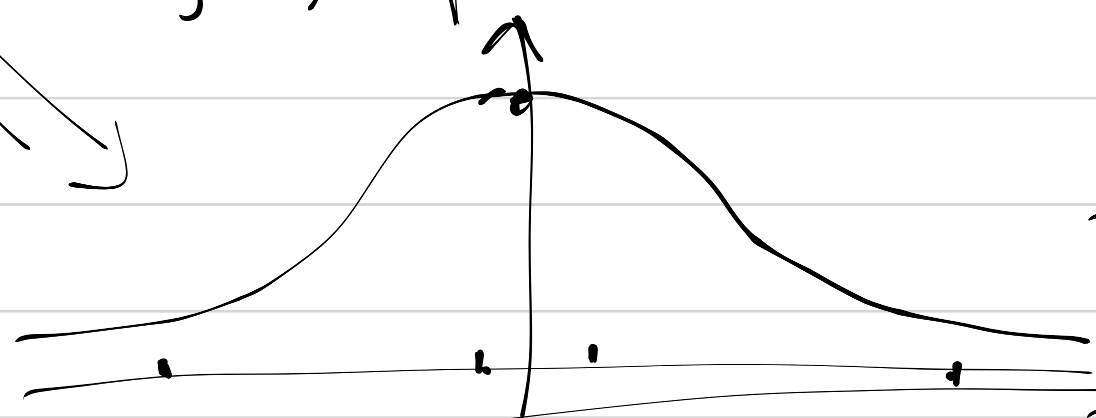
$$\ln g(u) = c\cdot u^2 + d$$

из ничего

$$g(u) = \exp\left(cu^2 + d\right)$$

$g(u)$ – ф-ция плотности! $c < 0$

$$\int_{-\infty}^{\infty} g(u)\,du = 1$$



$$c = -\frac{1}{2\sigma^2}$$

$U \sim N(0;\sigma^2)$

$$g(u) = \frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-u^2/2\sigma^2}$$

из · текста $\Rightarrow$ норм. плотн. (Гаусс)

Аксиоматика Хершела-Максвелла

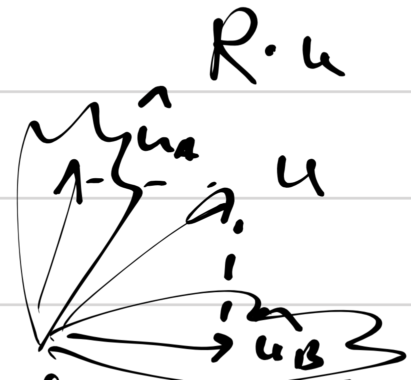
$\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$ – вектор скоростей моле-ул газа спички.

Rot Inv: закон распр-ия и не

1) зависит от поворота

$R\cdot u \sim u$ R – поворот

2) Ind Ort Proj:



$L_A \subset \mathbb{R}^3$ $L_A \perp L_B$

$L_B \subset \mathbb{R}^3$

$\hat u_B$ и $\hat u_A$ независимы.

[Макс]: $f(u)$ — сущ-т и дифф-ма.

$\Rightarrow$?

$$u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}$$

незав орт проекций $\Rightarrow$

$\Rightarrow u_1$ и u_2 и u_3 незав-мы!

$$f(u) = f_1(u_1)\cdot f_2(u_2)\cdot f_3(u_3) = f(u_1)\cdot f(u_2)\cdot f(u_3)$$

φ

Rot Inv

$$f(u_1, u_2, u_3) = g(u_1^2 + u_2^2 + u_3^2) = \tilde{g}(u_1^2)\cdot\tilde{g}(u_2^2)\cdot\tilde{g}(u_3^2)$$

$$g(u_1^2 + 0 + 0) = \tilde{g}(u_1^2)\cdot\tilde{g}(0)\cdot\tilde{g}(0)$$

$$\boxed{g(u_1^2 + u_2^2 + u_3^2) = c\cdot g(u_1^2)\cdot g(u_2^2)\cdot g(u_3^2)}$$

(интуиция) $g(t) = \exp(at + b)$

$$f(u) = \exp(a\cdot(u_1^2 + u_2^2 + u_3^2) + b)$$

$u_1, u_2, u_3 \sim$ нез $N(0;\sigma^2)$

$$f(u) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^3 \cdot \exp\left(-\frac{u_1^2 + u_2^2 + u_3^2}{2\sigma^2}\right)$$

$$a = -\frac{1}{2\sigma^2}$$

формально:

$$g'(u_1^2 + u_2^2 + u_3^2) = c\cdot g'(u_1^2)\cdot g(u_2^2)\cdot g(u_3^2)$$

$u_1^2 = 0 \quad u_3^2 = 0$

$$\boxed{g'(u_2^2) = a\cdot g(u_2^2)}$$ (!)