Задачи по регуляризации

(1)

$$p(x)=\frac{1}{\sqrt{2\pi}\,\sigma}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$L_x(\mu,\sigma)=\prod p(x_i)$$

$$\mathcal{L}=\ln L_x=\sum \ln p(x_i)=n\ln\left(\frac{1}{\sqrt{2\pi}\,\sigma}\right)-\sum\frac{(x_i-\mu)^2}{2\sigma^2}$$

$$\frac{\partial\mathcal{L}}{\partial\mu}=0=-\sum\frac{x_i-\mu}{2\sigma^2}=0\Rightarrow\mu=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{x}$$

$$\frac{\partial\mathcal{L}}{\partial\sigma}=-\frac{n}{\sigma}+\frac{\sum(x_i-\bar{x})^2}{\sigma^3}=0$$

$$\Rightarrow\sigma^2=\frac{1}{n}\sum(x_i-\bar{x})^2$$ — выборочная дисперсия (смещённая, не $\frac{1}{n-1}$)

(2)

1) $$P_\lambda(n)=\frac{\lambda^n}{n!}e^{-\lambda}$$

$$P_\lambda(\lambda_0\mid n=m)\propto P(n=m\mid\lambda_0)\,P_\lambda(\lambda_0)=\frac{\lambda_0^m}{m!}e^{-\lambda_0}p_0$$

$$P(n=m)=\int_0^\infty P(n=m/\lambda)\,P_\lambda(\lambda)\,d\lambda=p_0\int P(n=m/\lambda)\,d\lambda$$

$$=p_0\int_0^\infty\frac{\lambda^m}{m!}e^{-\lambda}d\lambda=p_0\Rightarrow\boxed{P_\lambda(\lambda_0\mid n=m)=\frac{\lambda_0^m}{m!}e^{-\lambda_0}}$$

2) Делаем то же самое, но теперь считаем

$$p_\lambda(\lambda_0) = \frac{\lambda_0^m}{m!} e^{-\lambda_0}$$ (получили/улучшили априорное распределение)

$$\Rightarrow \quad p_\lambda(\lambda_0 \mid n=k) = \frac{P(n=k \mid \lambda_0)\, p_\lambda(\lambda_0)}{P(n=k)} =$$

$$= \frac{\frac{\lambda_0^k}{k!} e^{-\lambda_0} \frac{\lambda_0^m}{m!} e^{-\lambda_0}}{\int\limits_0^\infty \frac{\lambda^k}{k!} e^{-\lambda} \frac{\lambda^m}{m!} e^{-\lambda} d\lambda} = \frac{\lambda_0^{k+m} e^{-2\lambda_0}}{\int\limits_0^\infty \lambda^{k+m} e^{-2\lambda} d\lambda}$$

$$\int\limits_0^\infty \lambda^{k+m} e^{-2\lambda} d\lambda = \int\limits_0^\infty \frac{(2\lambda)^{k+m}}{(2)^{k+m}} e^{-2\lambda} \frac{d(2\lambda)}{2}$$

$$= \frac{1}{2^{k+m+1}} \int\limits_0^\infty t^{k+m} e^{-t} dt =$$

$$= \frac{\Gamma(k+m+1)}{2^{k+m+1}} = \frac{(k+m)!}{2^{k+m+1}}$$

$$\Rightarrow \quad \frac{\lambda_0^{k+m} e^{-2\lambda_0}}{(k+m)!} 2^{k+m}$$

$$= \frac{(2\lambda_0)^{k+m}}{(k+m)!} e^{-2\lambda_0}$$

$\Rightarrow$ новое

$$\boxed{p_\lambda(\lambda_0 \mid n=k) = \frac{(2\lambda_0)^{k+m}}{(k+m)!} e^{-2\lambda_0}}$$

(3)

1) $$p(\text{болеет} \mid \text{положит.}) = \frac{P(\text{положит} \mid \text{болеет})\, P(\text{болеет})}{P(п \mid б)P(б) + P(п \mid нб)P(нб)}$$

$$= \frac{0.99 \cdot 10^{-5}}{0.99 \cdot 10^{-5} + 0.01 \cdot (1 - 10^{-5})} = \frac{0.99}{0.99 + 10^{3} - 0.01} \simeq \underline{10^{-3}}$$

2) $$P(\text{болеет} \mid x2\ \text{положит}) = \frac{P(x2\ \text{положит} \mid \text{болеет})\, P(\text{болеет})}{P(x2\ \text{положит} \mid \text{болеет}) + P(x2\ \text{полож.} \mid \text{не б.})} =$$

$$= \frac{0.99 \cdot 0.99 \cdot 10^{-5}}{0.99^2 \cdot 10^{-5} + 0.01^2 \cdot (1 - 10^{-5})} = \frac{0.99^2}{0.99^2 + 10} \simeq \underline{0.1}$$

$$P(x3\ \text{положит} \mid б) \simeq \frac{0.99^3}{0.99^3 + 10^{-1}} = \frac{0.99^3}{0.1 + 0.99^3} = \frac{0.97}{1.07} \simeq \underline{0.9}$$

$\Rightarrow$ для уд-результата Пете нужно потратить 300 \$

(7)

$$L = \|Xw - y\|^2 \to \min_w, \quad \sum_i |w_i| < C$$

$$\mathcal{L} = (Xw - y)^T (Xw - y) - \lambda\left(\sum |w_i| - C\right) =$$

$$= w^T X^T X w - 2w^T X^T y + y^T y - \lambda\left(\sum |w_i| - c\right)$$

$$\sum |w_\alpha| = |w|^T \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = w^T \begin{pmatrix} \operatorname{sign} w_1 \\ \vdots \\ \operatorname{sign} w_n \end{pmatrix}$$

$$\Rightarrow \frac{\partial L}{\partial w} = 0 \Leftrightarrow 2X^T X w - 2X^T y - \lambda \begin{pmatrix} \operatorname{sign} w_1 \\ \vdots \\ \operatorname{sign} w_n \end{pmatrix} = 0$$

$$\frac{\partial L}{\partial \lambda} = 0 \Leftrightarrow \sum |w_\alpha| - C = 0 \Rightarrow \sum |w_\alpha| = C$$

$$\operatorname{sign} w_i = \frac{w_i}{|w_i|}; \quad C = w^T \begin{pmatrix} \operatorname{sign} w_1 \\ \vdots \\ \operatorname{sign} w_n \end{pmatrix}$$

$$\Rightarrow 2w^T X^T X w - 2w^T X^T y - \lambda C = 0$$

$$\Rightarrow \lambda = \frac{2}{C} w^T X^T (Xw - y) = \frac{2}{C}\left[(Xw-y)^T (Xw-y) + y^T (Xw-y)\right] =$$

$$= \frac{2}{C} \|Xw - y\|^2 + \frac{2}{C} y^T (Xw - y)$$

Задача эквивалентна Lasso-регрессии, т.к. везде используется метод мн-телей Лагранжа

(4)

$$p_0(\bar{x}) = C \exp\left(-\frac{\bar{x}^T A \bar{x}}{2}\right)$$

Изначально предполагаем, что $\bar{x}$ распределено как $p_0(\bar{x})$. После измерения $\bar{x}_1$ можем найти A: $p_A(\bar{x}_1) \to \max$ (max-likelihood)

Для этого предварительно нормируем $p_0(\bar{x})$

$$\int_{\mathbb{R}^n} p_0(\bar{x})\, d\bar{x} = C \int_{\mathbb{R}^n} \exp\left(-\frac{\bar{x}^T A \bar{x}}{2}\right) d\bar{x} =$$

$$\bar{x}^T A \bar{x} = (x_1, \dots, x_n)\begin{pmatrix} a_{11} & \dots & a_{1n} \\ - & - & - \\ a_{n1} & \dots & a_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_i a_{i1} \dots x_i a_{in})\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= x_i a_{ij} x_j$$

$$\exp\left(-\frac{\bar{x}^T A \bar{x}}{2}\right) = \exp\left(-\frac{\sum x_i^2 a_i}{2}\right) \exp\left(-\left.\frac{x_i a_{ij} x_j}{2}\right|_{i \neq j}\right)$$