1 問題設定

2 ベイズ推論

2.1 a と b の分布

2.2 ノイズ σ の推定

前のファイルを参照

2.3 モデル選択

$$y = ax + b \quad \text{か} \quad y = ax \text{ か}$$

たとえば …

フックの法則

$$f = kx + \cancel{f_0}$$

ニュートンの運動方程式

$$f = ma + \cancel{f_0}$$

など $y = ax$ 型の理論が選ばれることもある

$y = ax$ の理論を作る

$$\mathbb{X} = \{x_1, x_2, \cdots, x_N\}, \mathbb{Y} = \{y_1, y_2, \cdots, y_N\}$$

$$y = ax$$

$$a_0 = \frac{\overline{xy}}{\overline{x^2}} \quad (\text{MAP 推定})$$

$$p(a \mid \mathbb{X}, \mathbb{Y}) = \sqrt{\frac{N\overline{x^2}}{2\pi\sigma^2}} \exp\left(-\frac{N\overline{x^2}}{2\sigma^2}(a - a_0)^2\right)$$

$p(\sigma \mid \mathbb{X}, \mathbb{Y})$ はどうなるか？

$$
\begin{aligned}
E(a) &= \frac{1}{N}\sum_{i=1}^{N}(y_i - ax_i)^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\left\{y_i^2 - 2ax_iy_i + a^2x_i^2\right\} \\
&= \frac{1}{N}\sum_{i=1}^{N}y_i^2 - 2a\frac{1}{N}\sum_{i=1}^{N}x_iy_i + a^2\frac{1}{N}\sum_{i=1}^{N}x_i^2 \\
&= \overline{y^2} - 2a\overline{xy} + a^2\overline{x^2}
\end{aligned}
$$

$$
\frac{\partial E(a)}{\partial a} = -2\overline{xy} + 2a\overline{x^2} = 0
$$

よって

$$
a = \frac{\overline{xy}}{\overline{x^2}}
$$

のとき $E(a)$ が最小。平方完成すると

$$
\begin{aligned}
E(a) &= \overline{x^2}\left(a^2 - 2a\frac{\overline{xy}}{\overline{x^2}} + \left(\frac{\overline{xy}}{\overline{x^2}}\right)^2 - \left(\frac{\overline{xy}}{\overline{x^2}}\right)^2\right) + y^2 \\
&= \overline{x^2}\left(a - \frac{\overline{xy}}{\overline{x^2}}\right)^2 + \frac{(\overline{xy})^2}{\overline{x^2}} + \overline{y^2}
\end{aligned}
$$

$$
a_0 = \frac{\overline{xy}}{\overline{x^2}}
$$

$$
E(a_0) = \frac{(\overline{xy})^2}{\overline{x^2}} + y^2
$$

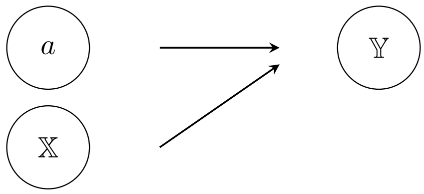
したがって $E(a)$ は次のように書き換えられる

$$
E(a) = E(a_0) + \overline{x^2}(a - a_0)^2
$$

a の分布

$$
\begin{aligned}
y_i &= ax_i + n_i \\
n_i &\sim N(0, \sigma^2) \\
p(n_i) &= \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{n_i^2}{2\sigma^2}\right) \\
\mathbb{X} &= \{x_1, x_2, \cdots, x_N\} \\
\mathbb{Y} &= \{y_1, y_2, \cdots, y_N\}
\end{aligned}
$$

$p(a \mid \mathbb{X}, \mathbb{Y})$ が知りたい



同時分布

$$p(\mathbb{X}, \mathbb{Y}, a)$$

$$\begin{aligned} p(\mathbb{X}, \mathbb{Y}, a) &= p(\mathbb{Y} \mid \mathbb{X}, a) p(\mathbb{X}) p(a) \\ p(\mathbb{Y} \mid \mathbb{X}, a) &= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - a x_i)^2}{2\sigma^2} \right) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{N}{2\sigma^2} \frac{1}{N} \sum_{i=1}^{N} (y_i - a_i)^2 \right) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{N}{2\sigma^2} E(a) \right) \\ E(a) &= E(a_0) + \overline{x^2}(a - a_0)^2 \end{aligned}$$

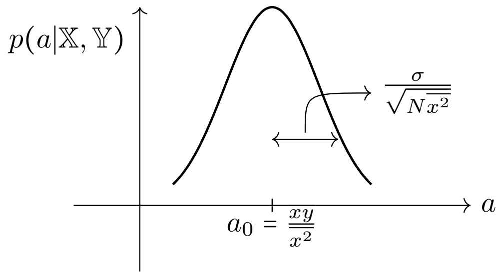
$$a_0 = \frac{\overline{xy}}{\overline{x^2}}$$

$$\begin{aligned} p(\mathbb{X}, \mathbb{Y}, a) &= p(\mathbb{Y} \mid \mathbb{X}, a) p(\mathbb{X}) p(a) \\ &= p(a \mid \mathbb{X}, \mathbb{Y}) p(\mathbb{X}, \mathbb{Y}) \\ \therefore p(a \mid \mathbb{X}, \mathbb{Y}) &\propto p(\mathbb{Y} \mid \mathbb{X}, a) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{N}{2\sigma^2} E(a) \right) \\ &= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{N}{2\sigma^2} E(a_0) \right) \exp\left(-\frac{N\overline{x^2}}{2\sigma^2} (a - a_0)^2 \right) \\ &\propto \exp\left(-\frac{N\overline{x^2}}{2\sigma^2} (a - a_0)^2 \right) \end{aligned}$$

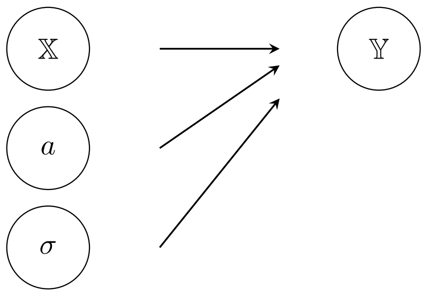
$\int da \; p(a \mid \mathbb{X}, \mathbb{Y}) = 1$ より

$$p(a \mid \mathbb{X}, \mathbb{Y}) = \sqrt{\frac{N\overline{x^2}}{2\pi\sigma^2}} \exp\left(-\frac{N\overline{x^2}}{2\sigma^2} (a - a_0)^2 \right)$$

となる



ノイズ σ の推定



生成モデル

$$
\begin{aligned}
&p(\mathbb{Y} \mid \mathbb{X}, a, \sigma) \\
&p(\sigma \mid \mathbb{X}, \mathbb{Y}) \to \sigma\text{を決めて} \\
&p(a \mid \mathbb{X}, \mathbb{Y}, \sigma)
\end{aligned}
$$

が得られる

$p(\mathbb{X}, \mathbb{Y}, a, \sigma)$ を考える

$$
\begin{aligned}
p(\mathbb{X}, \mathbb{Y}, \sigma) &= \int da \quad \underbrace{p(\mathbb{X}, \mathbb{Y}, a, \sigma)}_{p(\sigma)p(a)p(\mathbb{X})p(\mathbb{Y}|\mathbb{X},a,\sigma)} \\
&\propto \int da\ p(\sigma)p(a)p(\mathbb{X})p(\mathbb{Y} \mid \mathbb{X}, a, \sigma), \\
p(\mathbb{X}, \mathbb{Y}, \sigma) &= p(\sigma \mid \mathbb{X}, \mathbb{Y})p(\mathbb{X}, \mathbb{Y}) \\
\therefore p(\sigma \mid \mathbb{X}, \mathbb{Y}) &\propto \int da\ p(\mathbb{Y} \mid \mathbb{X}, a, \sigma) \\
&= \int da\ (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{N}{2\sigma^2}E(a)\right) \\
&= (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{N}{2\sigma^2}E(a_0)\right) \int da\ \exp\left(-\frac{N\overline{x^2}}{2\sigma^2}(a - a_0)^2\right) \\
&= (2\pi\sigma^2)^{-\frac{N-1}{2}} \exp\left(-\frac{N}{2\sigma^2}E(a_0)\right) \times \sqrt{\frac{1}{N\overline{x^2}}}
\end{aligned}
$$

MAP 推定

$$
\begin{aligned}
F(\sigma) &= -\log p(\sigma \mid \mathbb{X}, \mathbb{Y}) \\
&= \frac{N}{2\sigma^2} E(a_0) + \frac{N-1}{2} \log\left(2\pi\sigma^2\right) \\
\text{変数変換} \quad & S = \sigma^2 \\
F(S) &= \frac{N}{2S} E(a_0) + \frac{N-1}{2} \log(2\pi S) \\
\frac{\mathrm{d}F(S)}{\mathrm{d}S} &= -\frac{N}{2S^2} E(a_0) + \frac{N-1}{2S} = 0
\end{aligned}
$$

S の微分 = 0 より

$$
\begin{aligned}
S(N-1) &= NE(a_0) \\
S &= \frac{N}{N-1} E(a_0) \\
\sigma^2 &= \frac{N}{N-1} \frac{1}{N} \sum_{i=1}^{N} (y_i - a_0 x_i)^2 \\
&= \frac{1}{N-1} \sum_{i=1}^{N} (y_i - a_0 x_i)^2
\end{aligned}
$$

$y = ax$ の理論を作る

$$\mathbb{X} = \{x_1, x_2, \cdots, x_N\}, \mathbb{Y} = \{y_1, y_2, \cdots, y_N\}$$

$$y = ax$$

$$a_0 = \frac{\overline{xy}}{\overline{x^2}} \quad (\text{MAP 推定 })$$

$$p(a \mid \mathbb{X}, \mathbb{Y}) = \sqrt{\frac{N\overline{x^2}}{2\pi\sigma^2}} \exp\left(-\frac{N\overline{x^2}}{2\sigma^2} (a - a_0)^2\right)$$

$$p(\sigma \mid \mathbb{X}, \mathbb{Y}) ？？$$

σ は

$$
\begin{aligned}
\sigma^2 &= \frac{N}{N-1} E(a_0) \\
&= \frac{1}{N-1} \sum_{i=1}^{N} (y_i - a_0 x_i)^2
\end{aligned}
$$

と書ける

モデル選択

$$K=1 \qquad y=ax$$
$$K=2 \qquad y=ax+b$$

同時分布

$$p(\mathbb{X},\mathbb{Y},a,b,K=2)$$
$$p(\mathbb{X},\mathbb{Y},a,K=1)$$

$K=2$ を考える

$$
\begin{aligned}
&p(\mathbb{X},\mathbb{Y},a,b,K=2)\\
=&p(K=2)p(a,b\mid K=2)p(\mathbb{Y}\mid\mathbb{X},a,b)p(\mathbb{X})\\
&p(\mathbb{X},\mathbb{Y},K=2)\\
=&\int da\int db\ p(\mathbb{X},\mathbb{Y},a,b,K=2)\quad(\text{周辺化})\\
=&p(K=2\mid\mathbb{X},\mathbb{Y})p(\mathbb{X},\mathbb{Y})
\end{aligned}
$$

$$
\begin{aligned}
\therefore p(K=2\mid\mathbb{X},\mathbb{Y}) &\propto \int da\ \int db\ p(\mathbb{X},\mathbb{Y},a,b,K=2)\\
&\propto \int da\ \int db\ p(\mathbb{Y}\mid\mathbb{X},a,b)\\
&= \int da\ \int db\ (2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}\frac{1}{N}\sum_{i=1}^{N}\{y_i-(ax_i+b)\}^2\right)\\
&= \int da\ \int db\ (2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma}E(a,b)\right)\\
&= (2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}E(a_0,b_0)\right)\int da\ \exp\left(-\frac{N\overline{x^2}}{2\sigma^2}(a-a_0)^2\right)\int db\ \exp\left(-\frac{N}{2\sigma^2}(b-b_0)^2\right)\\
&= (2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}E(a_0,b_0)\right)\sqrt{\frac{2\pi\sigma^2}{N\overline{x^2}}}\times\sqrt{\frac{2\pi\sigma^2}{N}}
\end{aligned}
$$

$$
\begin{aligned}
F(K=2) &= -\log p(K=2\mid\mathbb{X},\mathbb{Y})\\
&= N\left\{\frac{1}{2\sigma^2}E(a_0,b_0)+\frac{\log N}{2N}+\frac{\log N}{2N}+O\left(\frac{1}{N}\right)\right\}\\
&= N\left\{\frac{1}{2\sigma^2}E(a_0,b_0)+\frac{\log N}{N}+O\left(\frac{1}{N}\right)\right\}
\end{aligned}
$$

$K=1$ を考える

$$\begin{aligned}
&p(\mathbb{X},\mathbb{Y},a,K=1)\\
=&p(K=1)p(a\mid K=1)p(\mathbb{Y}\mid\mathbb{X},a)p(\mathbb{X})\\
&p(\mathbb{X},\mathbb{Y},K=1)\\
=&\int da\ p(\mathbb{X},\mathbb{Y}a,K=1)\\
=&p(K=1\mid\mathbb{X},\mathbb{Y})p(\mathbb{X},\mathbb{Y})
\end{aligned}$$

$$\begin{aligned}
\therefore p(K=1\mid\mathbb{X},\mathbb{Y}) &\propto \int da\ p(\mathbb{X},\mathbb{Y},a,K=1)\\
&=\int da\ (2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}\frac{1}{N}\sum_{i=1}^{N}(y_i-ax_i)^2\right)\\
&=\int da\ (2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}E(a)\right)\\
&=(2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}E(a_0)\right)\int da\ \ \exp\left(-\frac{N\overline{x^2}}{2\sigma^2}(a-a_0)^2\right)\\
&=(2\pi\sigma^2)^{-\frac{N}{2}}\exp\left(-\frac{N}{2\sigma^2}E(a_0)\right)\times\sqrt{\frac{2\pi\sigma^2}{\overline{x^2}}}
\end{aligned}$$

$$\begin{aligned}
F(K=1)&=-\log p(K=1\mid\mathbb{X},\mathbb{Y})\\
&=N\left\{\frac{1}{2\sigma^2}E(a_0)+\frac{\log N}{2N}+O\left(\frac{1}{N}\right)\right\}
\end{aligned}$$