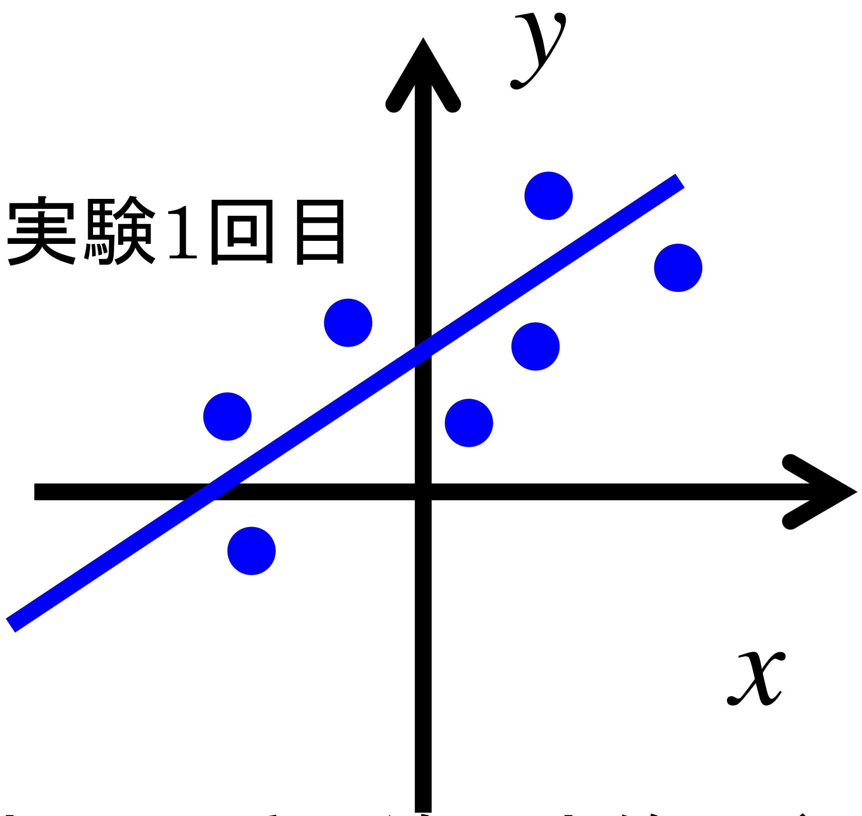
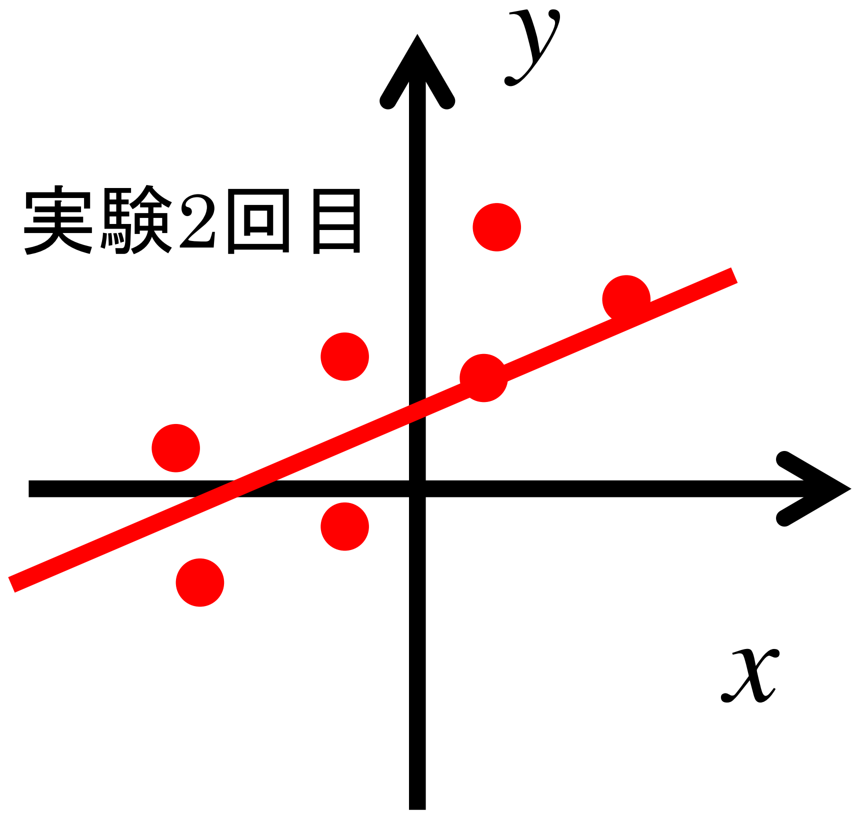
y=ax+bのベイズ推論

岡田真人

東京大学　大学院新領域創成科学研究科

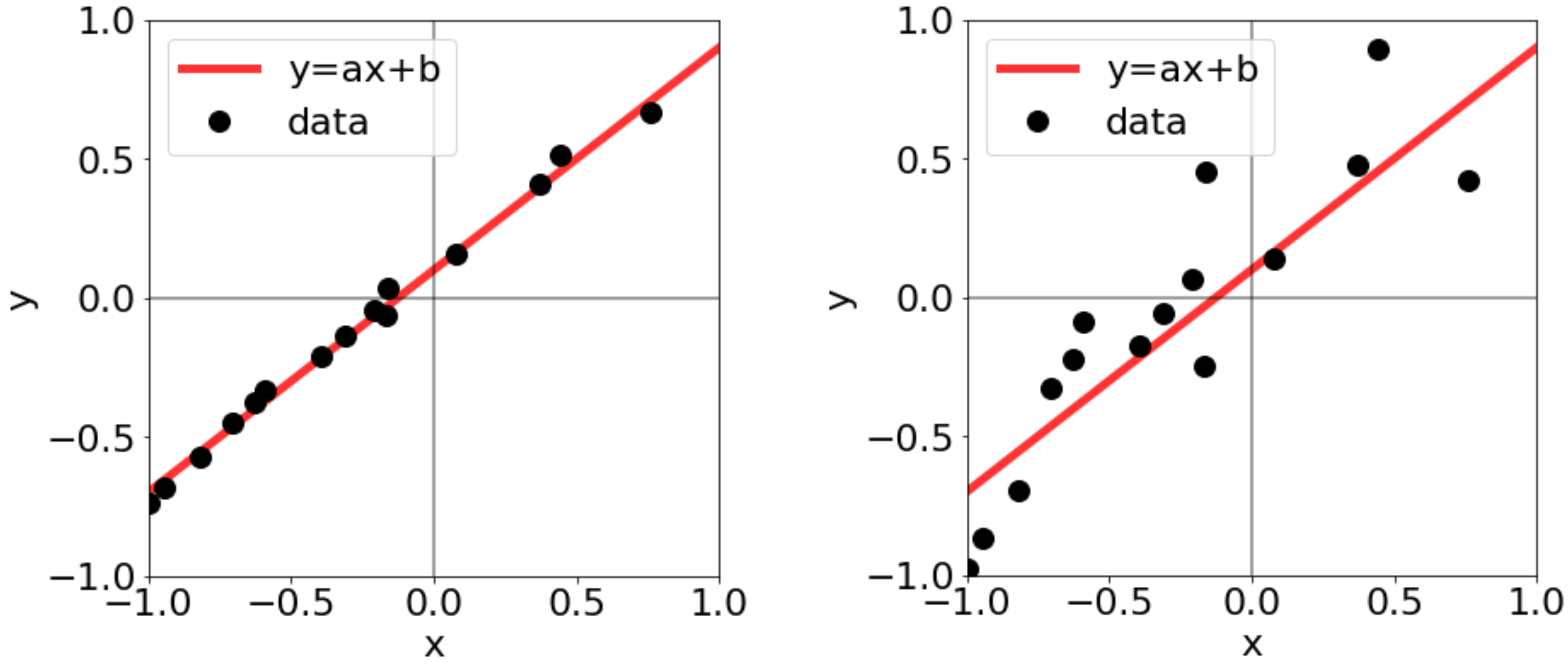
データのばらつきの評価
データの背後にある物理量の評価





傾き a： 系の線形応答、バネ定数、電気伝導度、誘電率、
実験複数回おこなって、a のばらつきを見る
これを1回の実験でもとめられないか → ベイズ推論

最小二乗法



この二つの違いを数学的に表現したい

傾き a と切片 b は同じだけど，ばらつきが違う

4.1 最小二乗法 (1/2)

$$E(a,b) = \frac{1}{N}\sum_{i=1}^{N}(y_i - (ax_i + b))^2$$

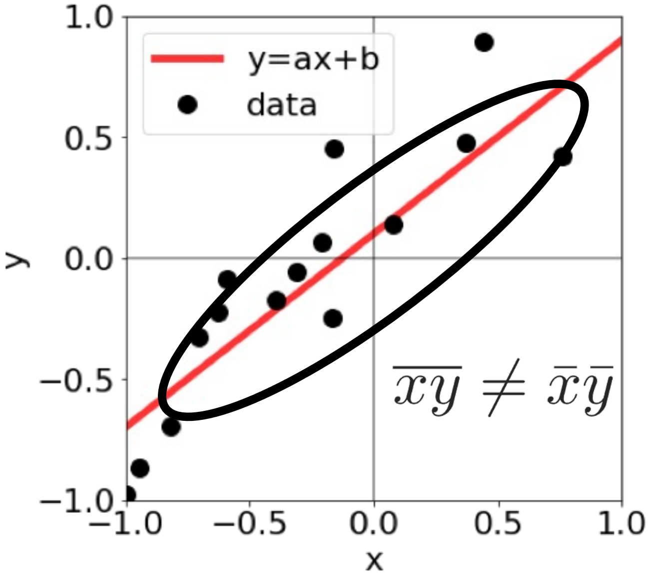
二乗誤差$E(a,b)$を最小にするようにパラメータをフィット（最小二乗法）

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i = 0$$ とする場合

$$E(a,b) = \frac{1}{2}\left(\overline{x^2}\left(a - \frac{\overline{xy}}{\overline{x^2}}\right)^2 + (b-\bar{y})^2 - \frac{\overline{xy}^2}{\overline{x^2}} - \bar{y}^2 + \overline{y^2}\right)$$

$$a_0 = \frac{\overline{xy}}{\overline{x^2}} \quad b_0 = \overline{y}$$

$$= \mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0, b_0) \geqq E(a_0, b_0)$$



平均: $\bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i$, 分散: $\overline{x^2} = \frac{1}{N}\sum_{i=1}^{N} x_i^2$

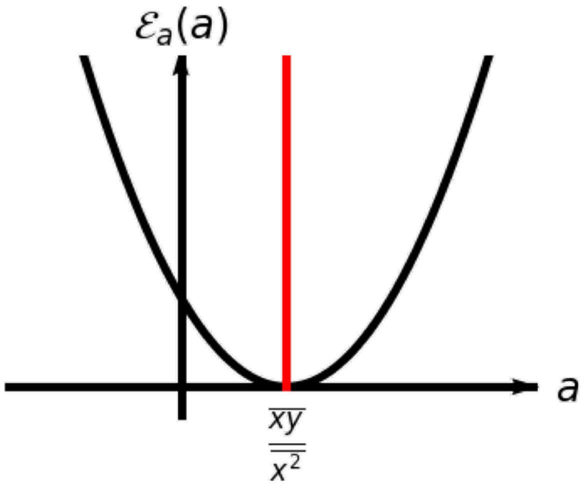
$$\overline{xy} = \frac{1}{N}\sum_{i=1}^{N} x_i y_i$$

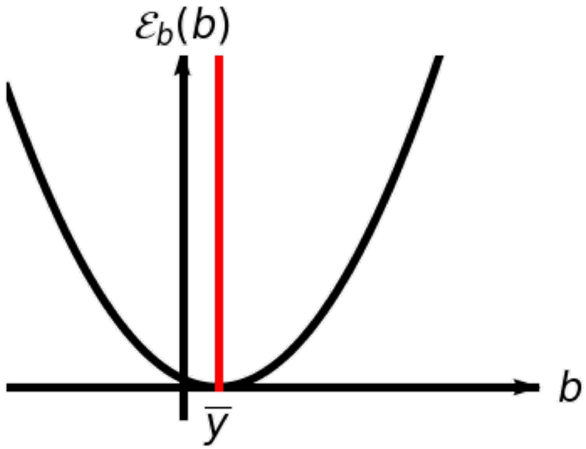
4.1 最小二乗法 (2/2)

$$E(a,b) = \frac{1}{N}\sum_{i=1}^{N}(y_i - (ax_i + b))^2$$

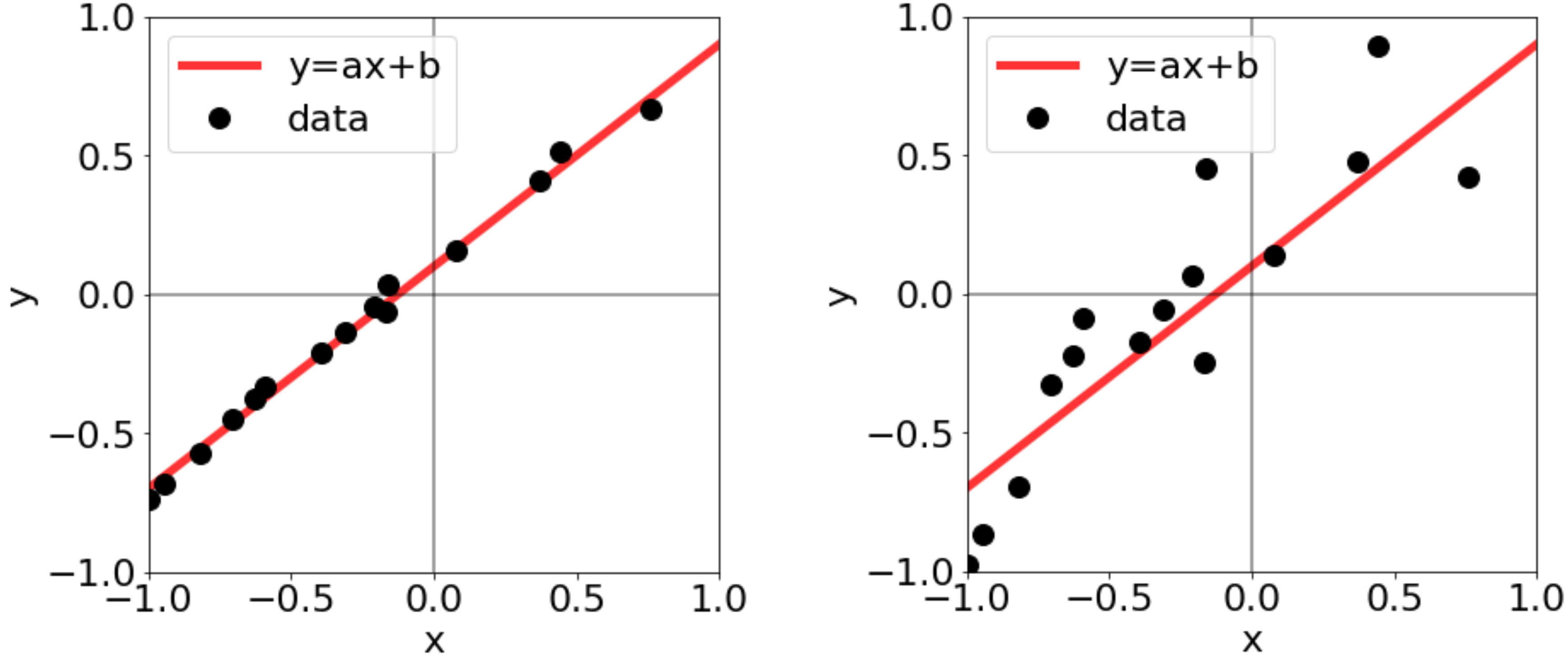
$$E(a,b) = \frac{1}{2}\left(\overline{x^2}\left(a - \frac{\overline{xy}}{\overline{x^2}}\right)^2 + (b - \bar{y})^2 - \frac{\overline{xy}^2}{\overline{x^2}} - \bar{y}^2 + \overline{y^2}\right)$$

$$E(a,b) = \mathcal{E}_a(a) + \mathcal{E}_b(b) + E(a_0,b_0) \geqq E(a_0,b_0)$$





4.2 自然科学的視点からのベイズ推論の解説 (1/2)



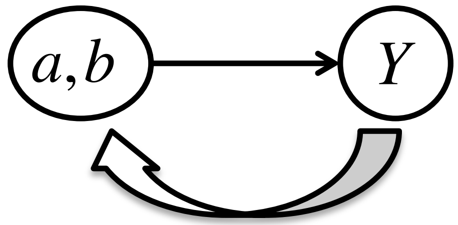
この二つの違いを数学的に表現したい

傾き a と切片 b は同じだけど，ばらつきが違う

4.2 自然科学的視点からの ベイズ推論の解説 (2/2)

$$p(Y,a,b) = \underline{p(Y \mid a,b)\,p(a,b) = p(a,b \mid Y)\,p(Y)}$$

生成（因果律）



＜ベイズの定理＞

$$p(a,b \mid Y) = \frac{p(Y \mid a,b)\,p(a,b)}{p(Y)} \propto \exp\left(-nE(a,b)\right)p(a,b)$$

$p(a,b \mid Y)$: 事後確率。データが与えられたもとでの，パラメータの確率.

$p(a,b)$: 事前確率。あらかじめ設定しておく必要がある。
これまで蓄積されてきた科学的知見

4.3 $p(a,b|Y)$の推定 (1/3)
最小二乗法では手で計算できる

$$y_i = ax_i + b + n_i$$

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

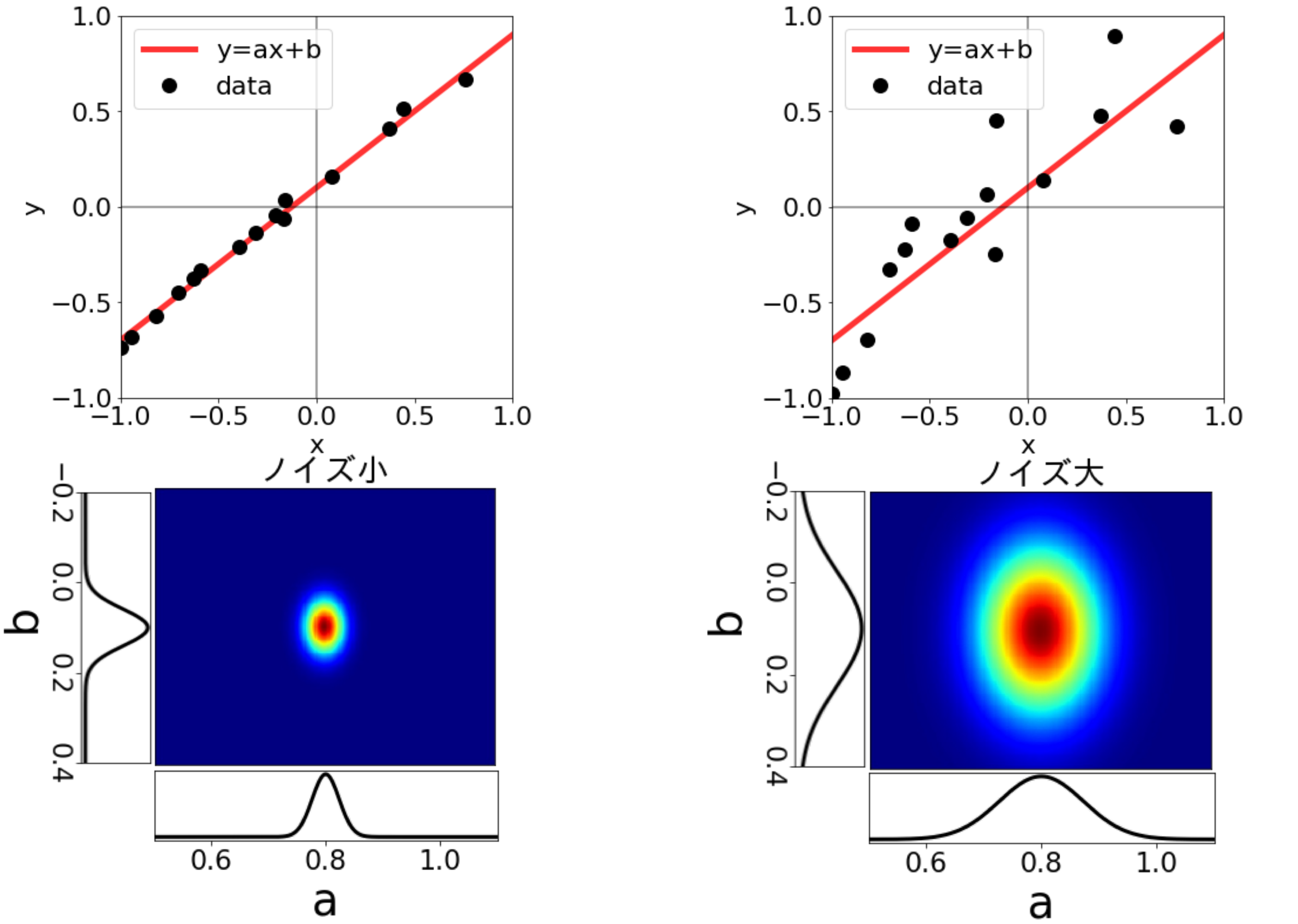
$$p(n_i) = p(y_i|a,b) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(y_i - (ax_i + b))^2}{2\sigma^2}\right)$$

$$\begin{aligned}
p(Y|a,b) &= \prod_{i=1}^{N} p(y_i|a,b) \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left(-\frac{\sum_{i=1}^{N}(y_i - (ax_i + b))^2}{2\sigma^2}\right) \\
&= \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^N \exp\left(-\frac{N}{\sigma^2}E(a,b)\right)
\end{aligned}$$

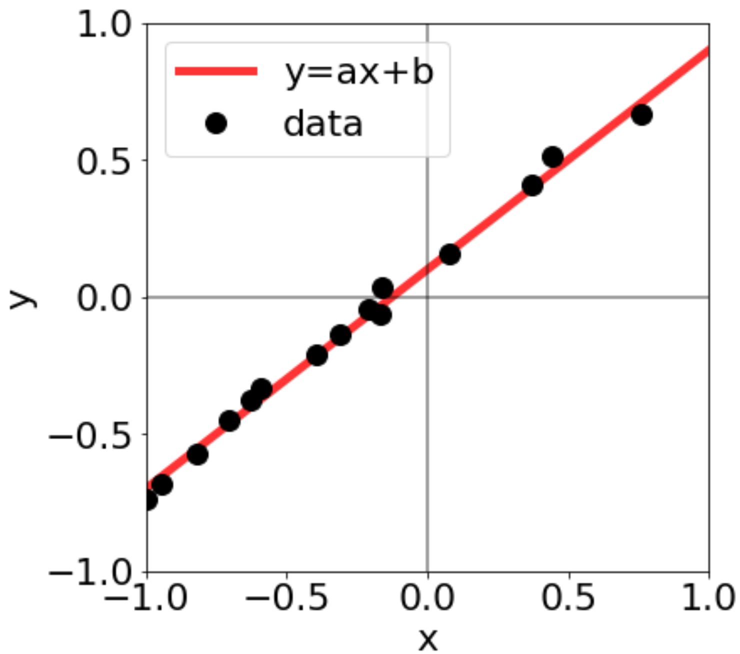
4.3 $p(a,b|Y)$の推定 (2/3) 最小二乗法では手で計算できる

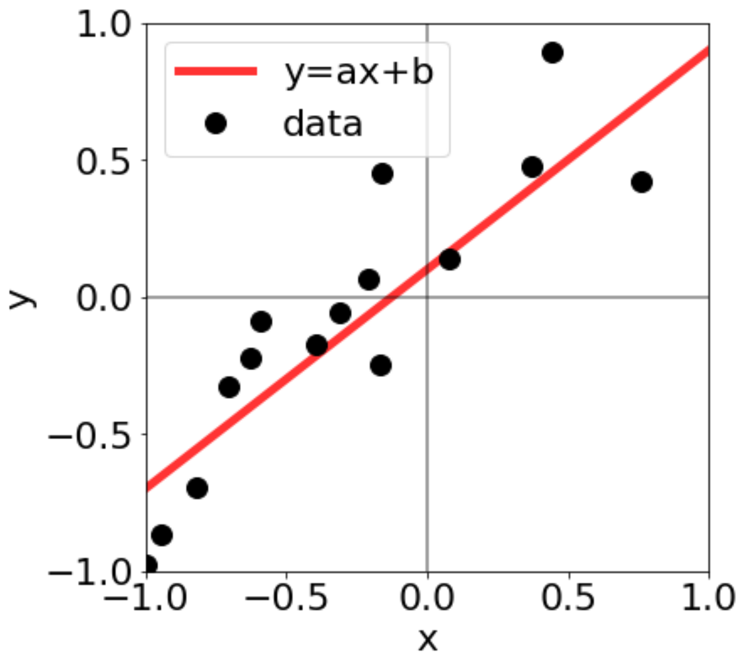
$$
\begin{aligned}
p(a,b|Y) &= \frac{p(Y|a,b)p(a,b)}{p(Y)} \propto p(Y|a,b) \\
&= \exp\left\{-\frac{N}{\sigma^2}\Big(\mathcal{E}_a(a)+\mathcal{E}_b(b)+E(a_0,b_0)\Big)\right\} \\
&\propto \exp\left\{-\frac{N}{\sigma^2}\Big(\mathcal{E}_a(a)+\mathcal{E}_b(b)\Big)\right\} \\
&= \exp\left\{-\frac{N\overline{x^2}}{2\sigma^2}(a-a_0)^2+\frac{N}{2\sigma^2}(b-b_0)^2\right\}
\end{aligned}
$$

4.3 $p(a,b|Y)$の推定 (3/3)
最小二乗法では手で計算できる



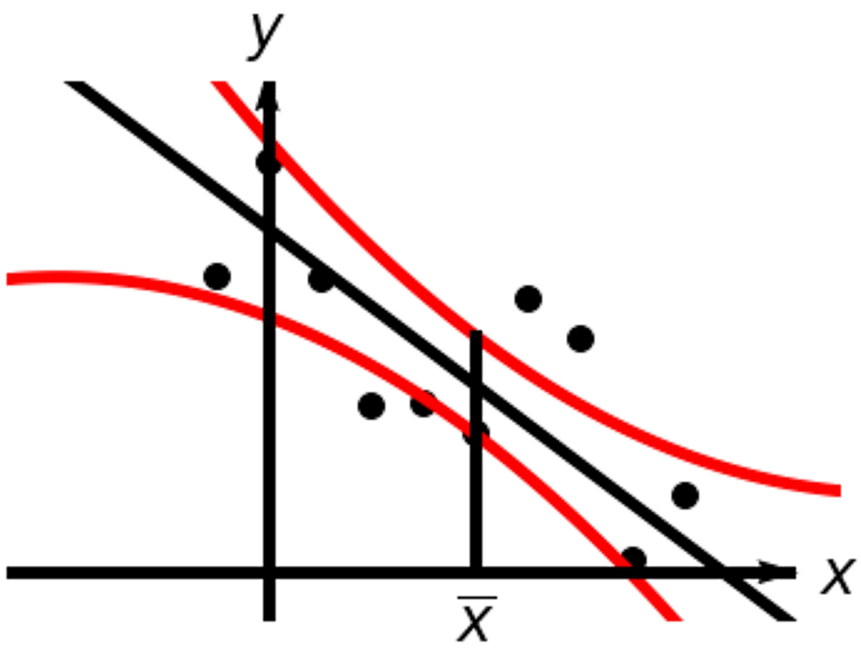
4.4 ノイズ分散の推定
ノイズ分散は誤差に比例





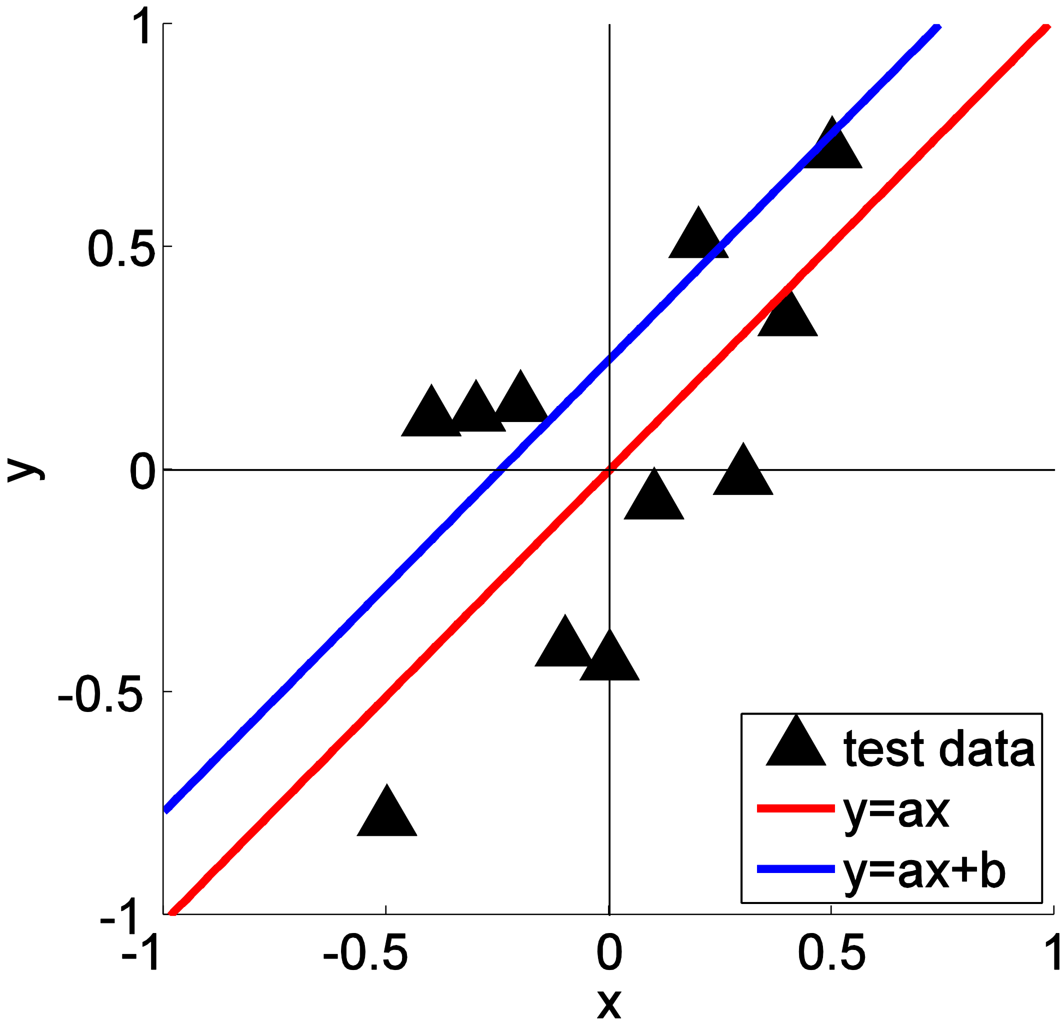
$$\sigma^2 = \frac{NE(a_0, b_0)}{N-2} = \frac{1}{N-2}\sum_{i=1}^{N}\{y_i - (a_0 x_i + b_0)\}^2$$

4.5 予測分布$p(y'|Y)$の推定
予測分布の分散は2次関数



$$\sigma_{y'}^2 = \frac{N}{N-2}E(a_0, b_0) + \frac{1}{(N-2)}\frac{x'^2}{\overline{x^2}}E(a_0, b_0) + \frac{1}{N-2}E(a_0, b_0)$$

4.6 ベイズ的モデル選択 $y=ax$か$y=ax+b$か?



	訓練誤差
$y=ax$	0.43
$y=ax+b$	0.34

	汎化誤差
$y=ax$	0.33
$y=ax+b$	0.40