マルチメディア工学５
マルチメディアデータの解析
基礎数理：最小二乗法

佐藤 嘉伸

大阪大学　大学院医学系研究科
放射線統合医学講座
yoshi@image.med.osaka-u.ac.jp
http://www.image.med.osaka-u.ac.jp/member/yoshi/
講義ホームページ: 日本語ページ → 授業の資料 → マルチメディア工学

マルチメディア工学：講義計画

- イントロダクション
- コンピュータグラフィックス (Computer Graphics: **CG**)
- マルチメディアデータの解析

マルチメディア工学：講義計画

- イントロダクション
- コンピュータグラフィックス (Computer Graphics: CG)
- マルチメディアデータの解析
  - 基礎数理
    - 最小二乗法
    - 直交変換、直交関数展開
  - 代表的解析手法

マルチメディアデータの解析

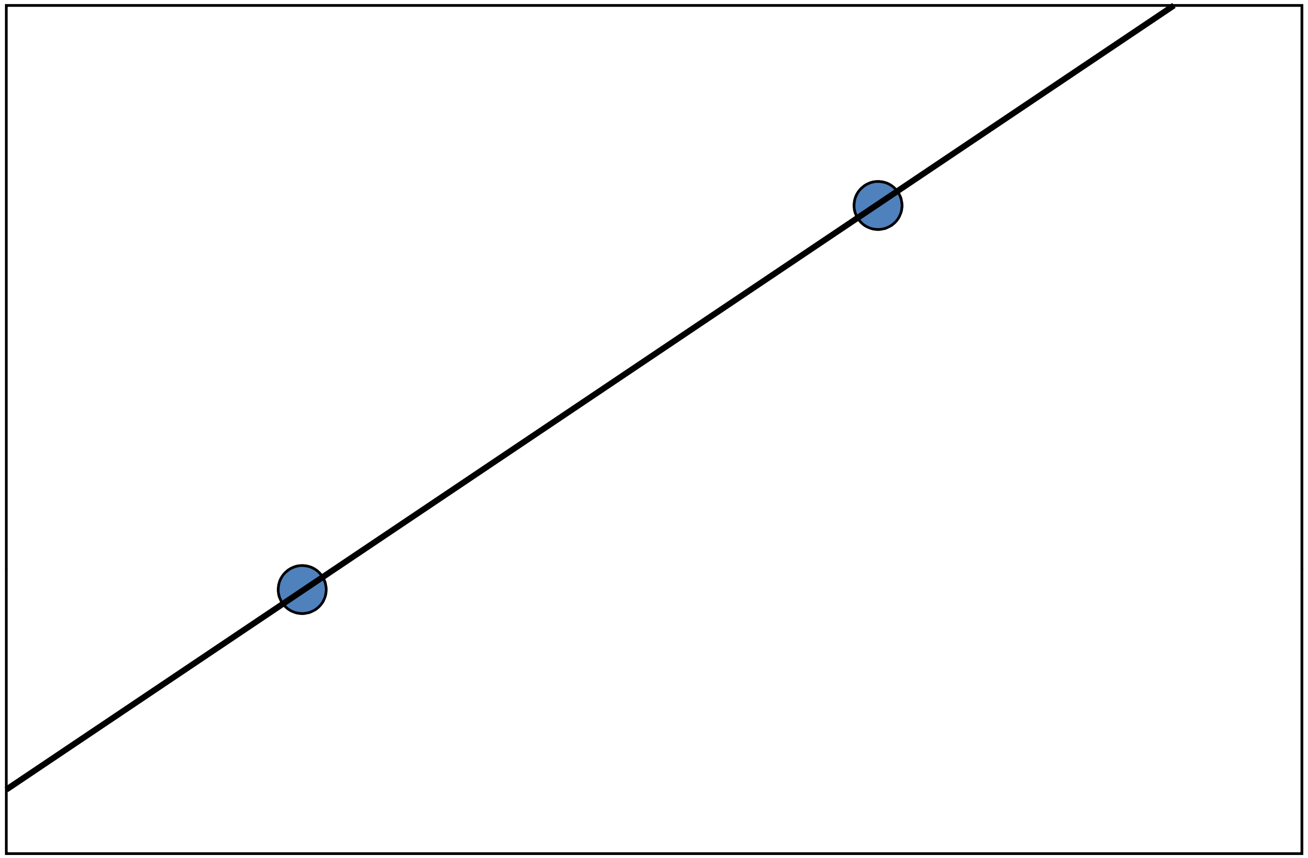
- 基礎数理
  - 最小二乗法
  - 直交変換、直交関数展開
- 代表的解析手法
  - データ圧縮：離散コサイン変換とＪＰＥＧ
  - データ表現：形状の主成分分析
  - （データ認識：隠れマルコフモデル）
    - 音声・言語を含む時系列データに対して有効
  - （音声の独立成分分析）

基礎数理：最小二乗法

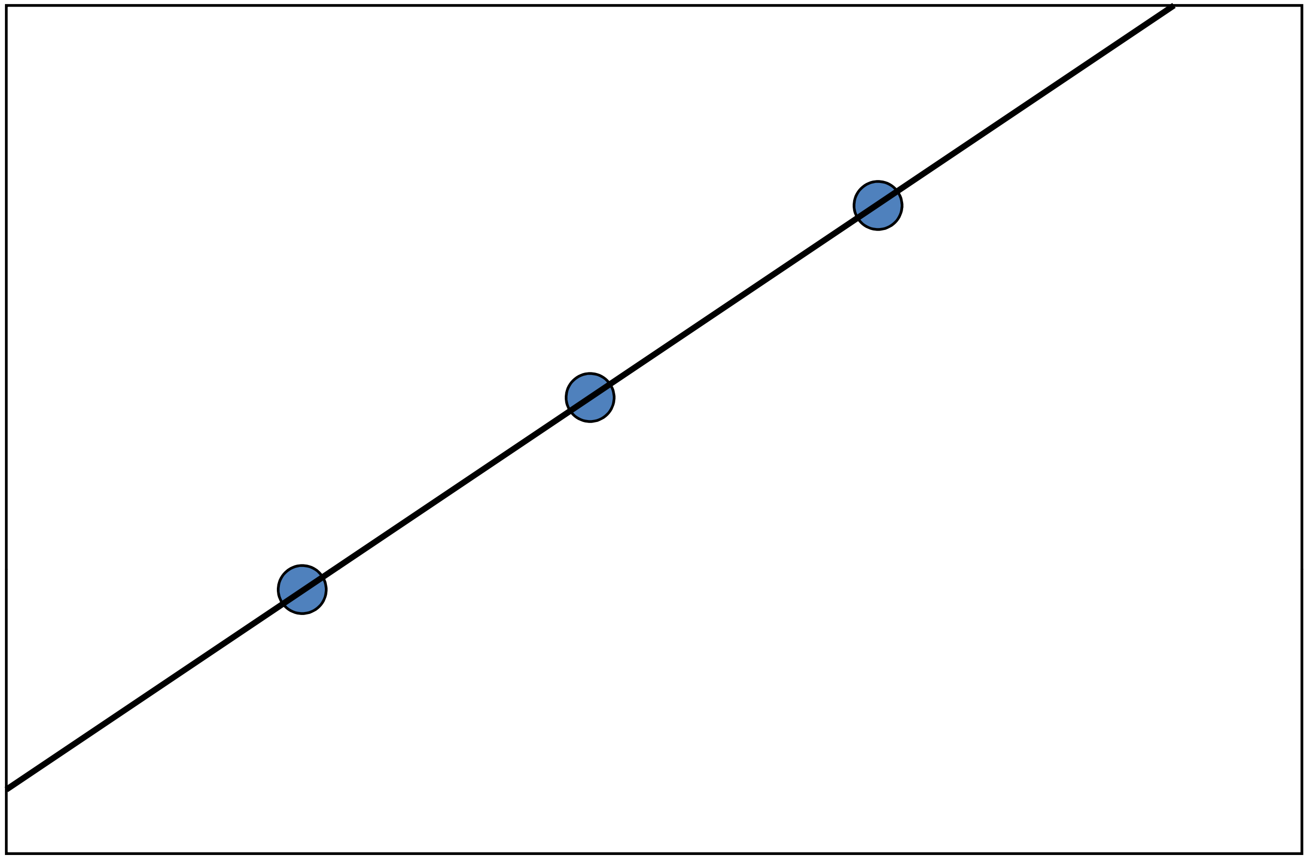
- **直線当てはめ**
  - **直線当てはめ** (Line Fitting)
  - **最小二乗法** (Least Squares) **の定式化**
  - **最尤推定法**(Maximum Likelihood Estimation)**としての最小二乗法**
- **最小二乗法の一般化**
  - **多項式当てはめ**
  - **線形最小二乗と非線形最小二乗** (Linear/Nonlinear Least Squares)

直線当てはめ

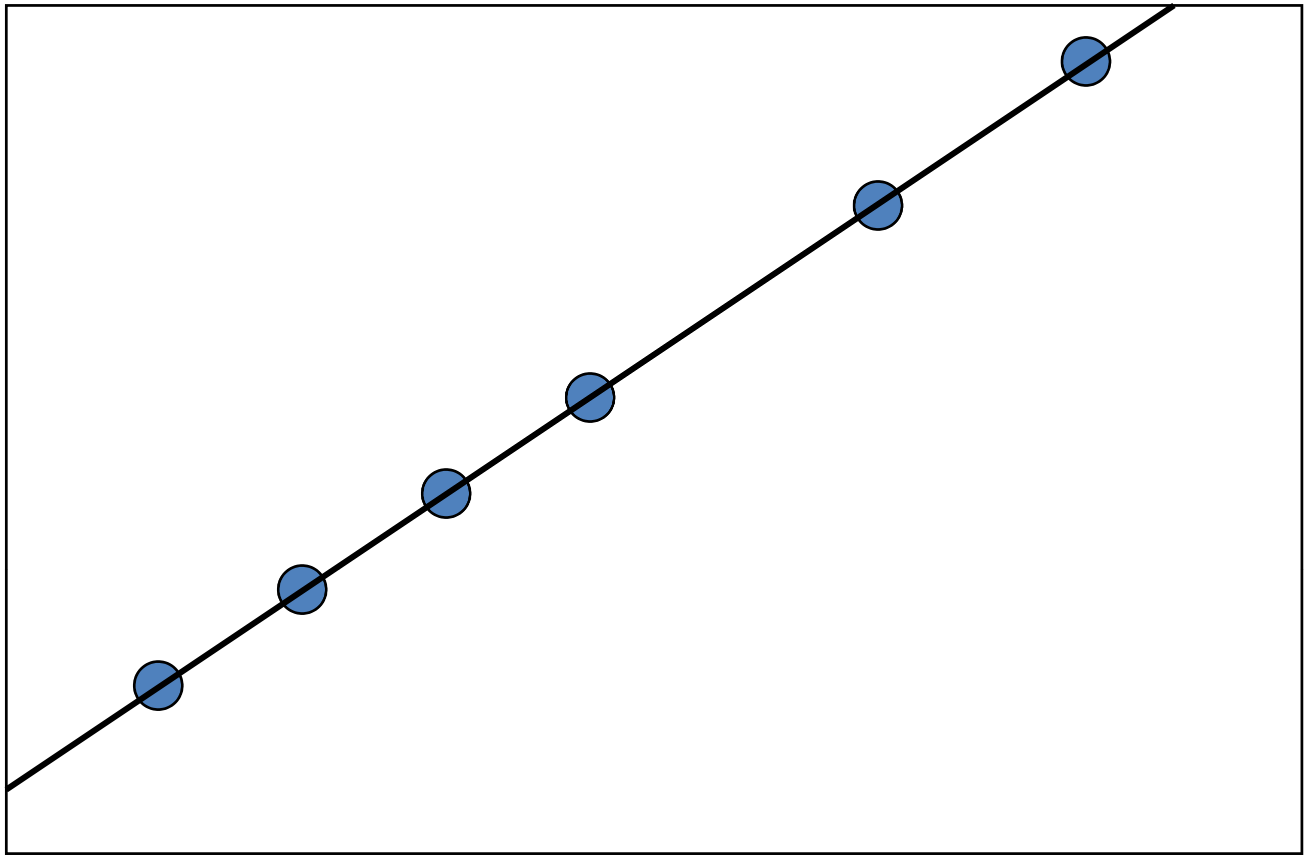
直線当てはめとは
(Line Fitting)



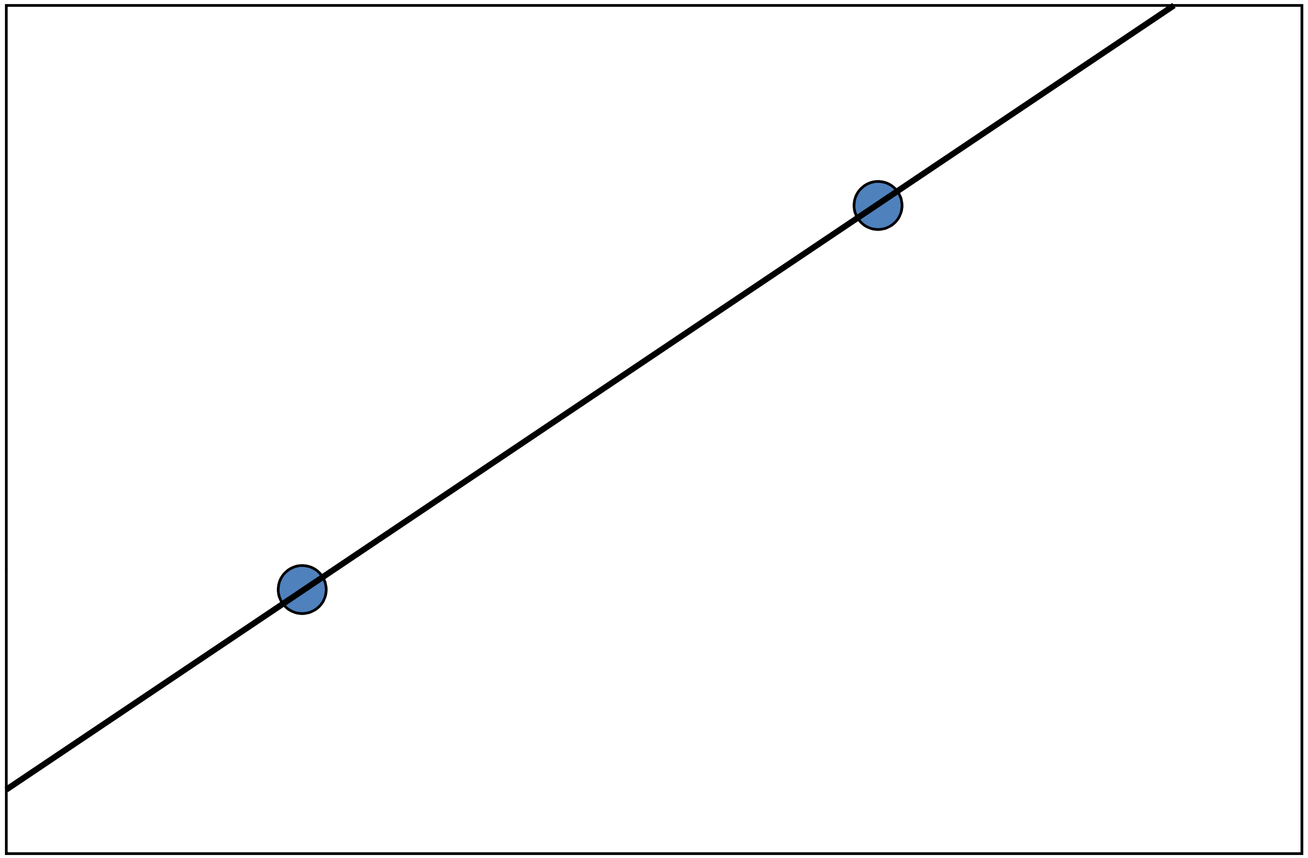
直線当てはめとは



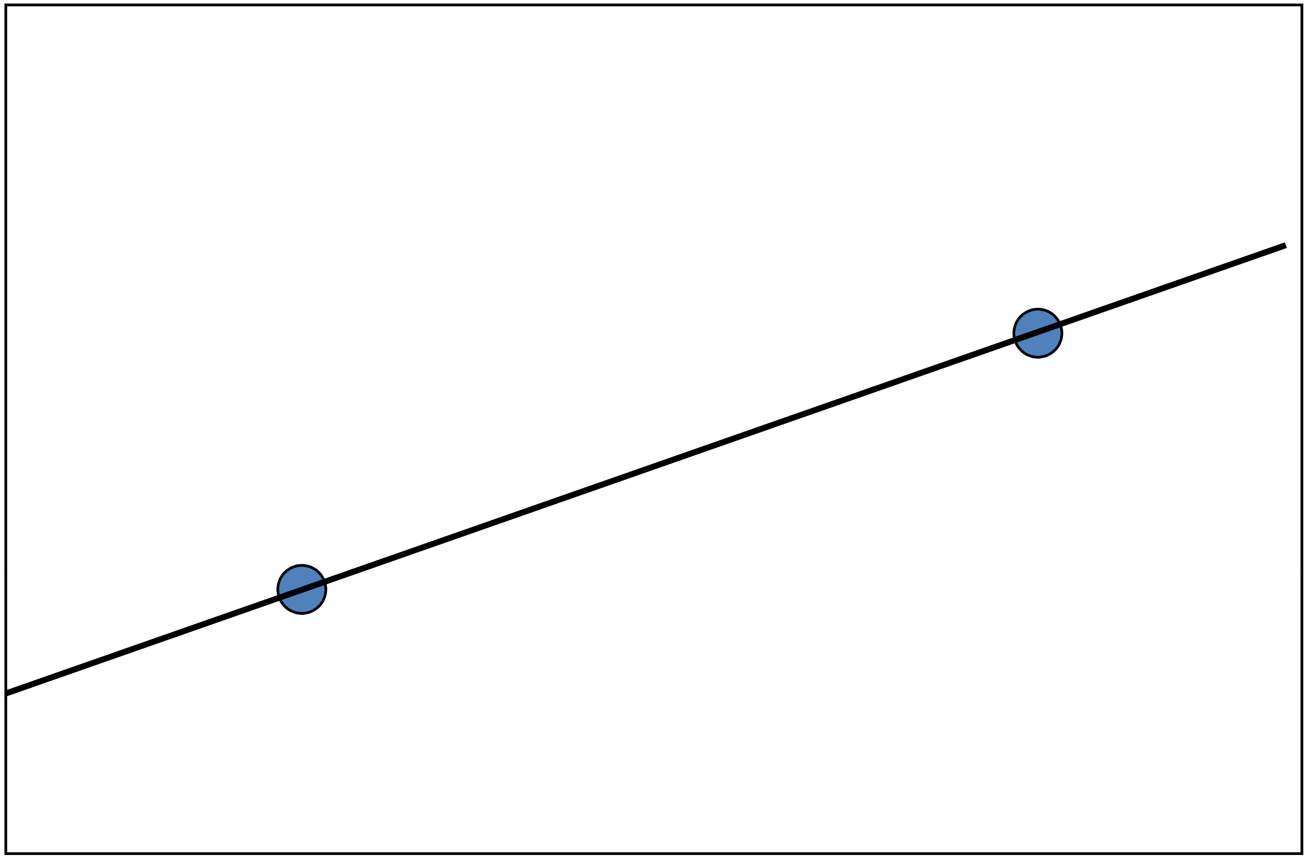
直線当てはめとは



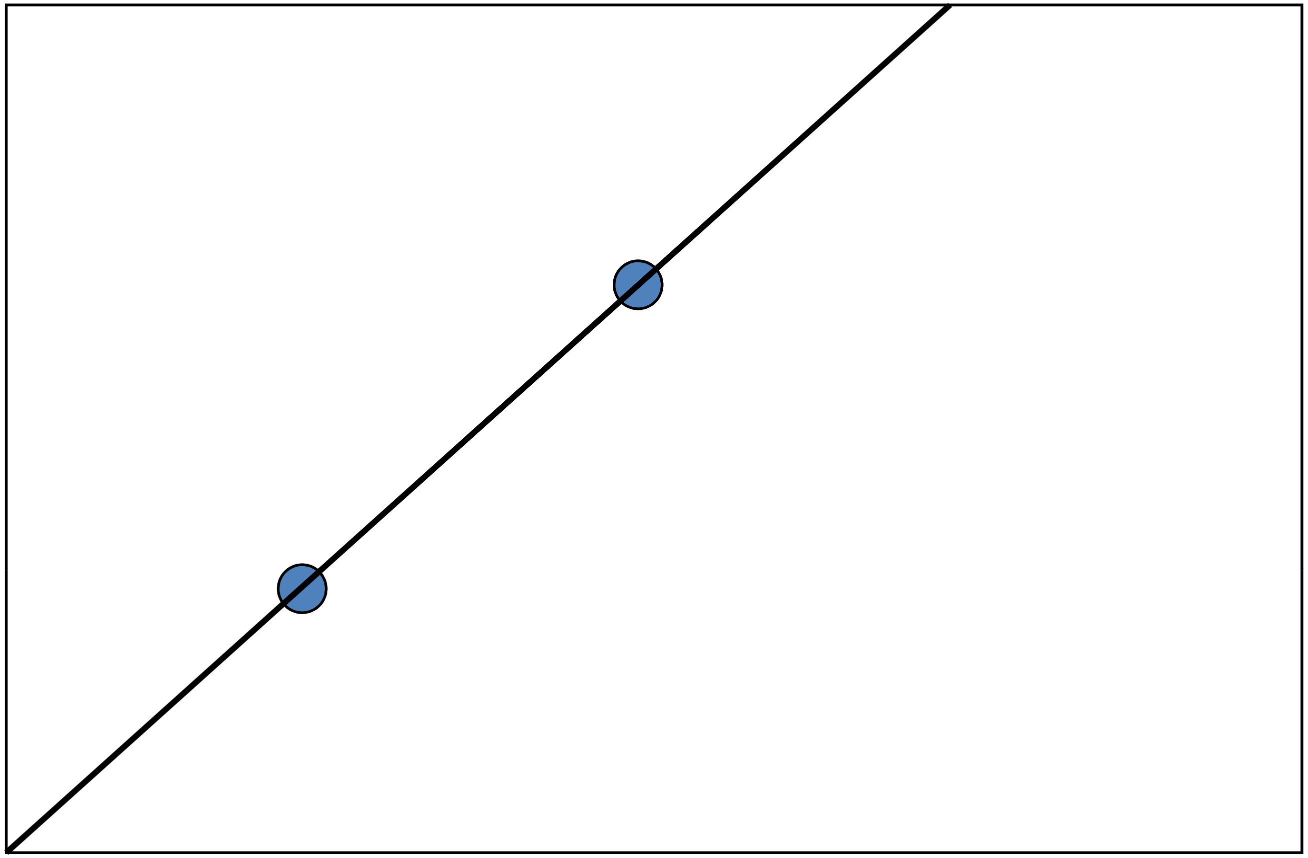
直線当てはめとは



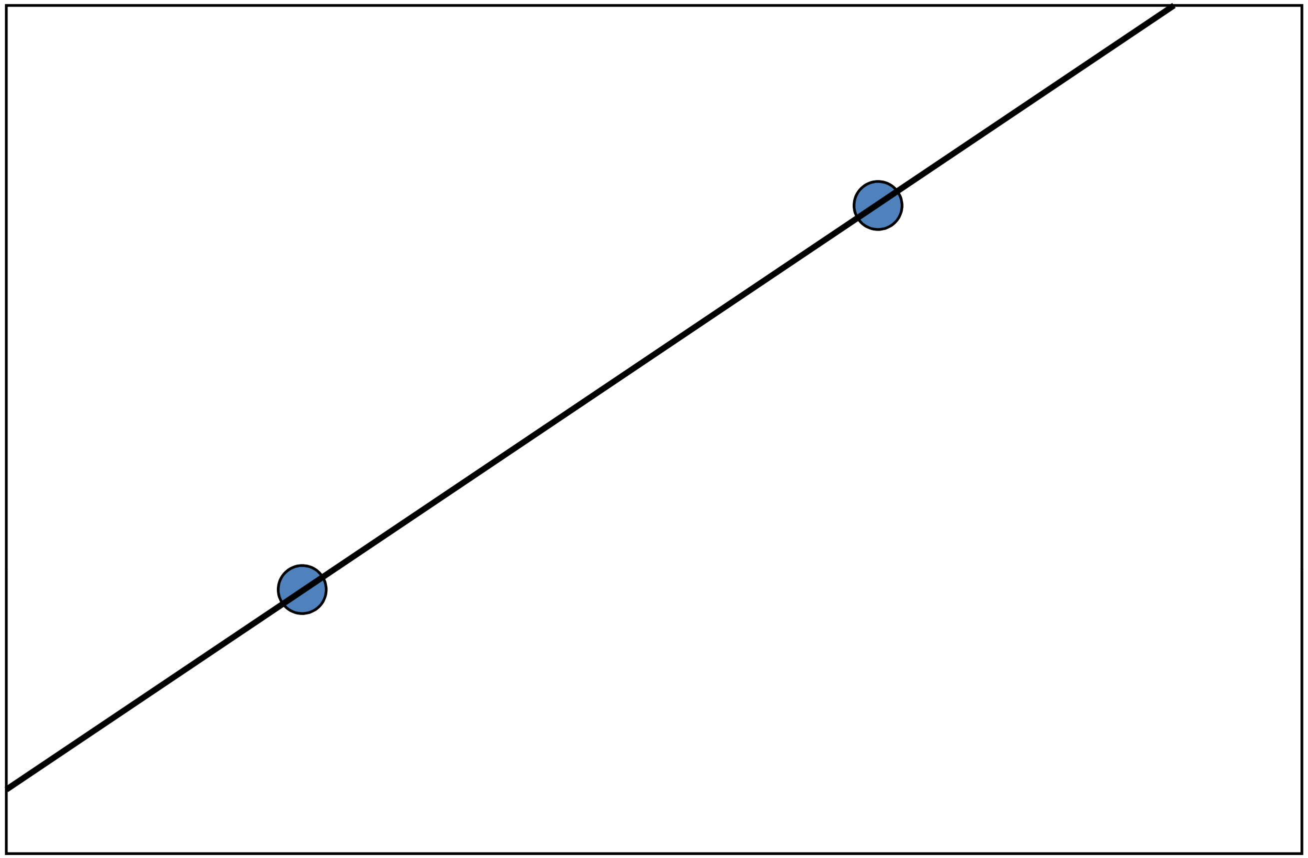
直線当てはめとは



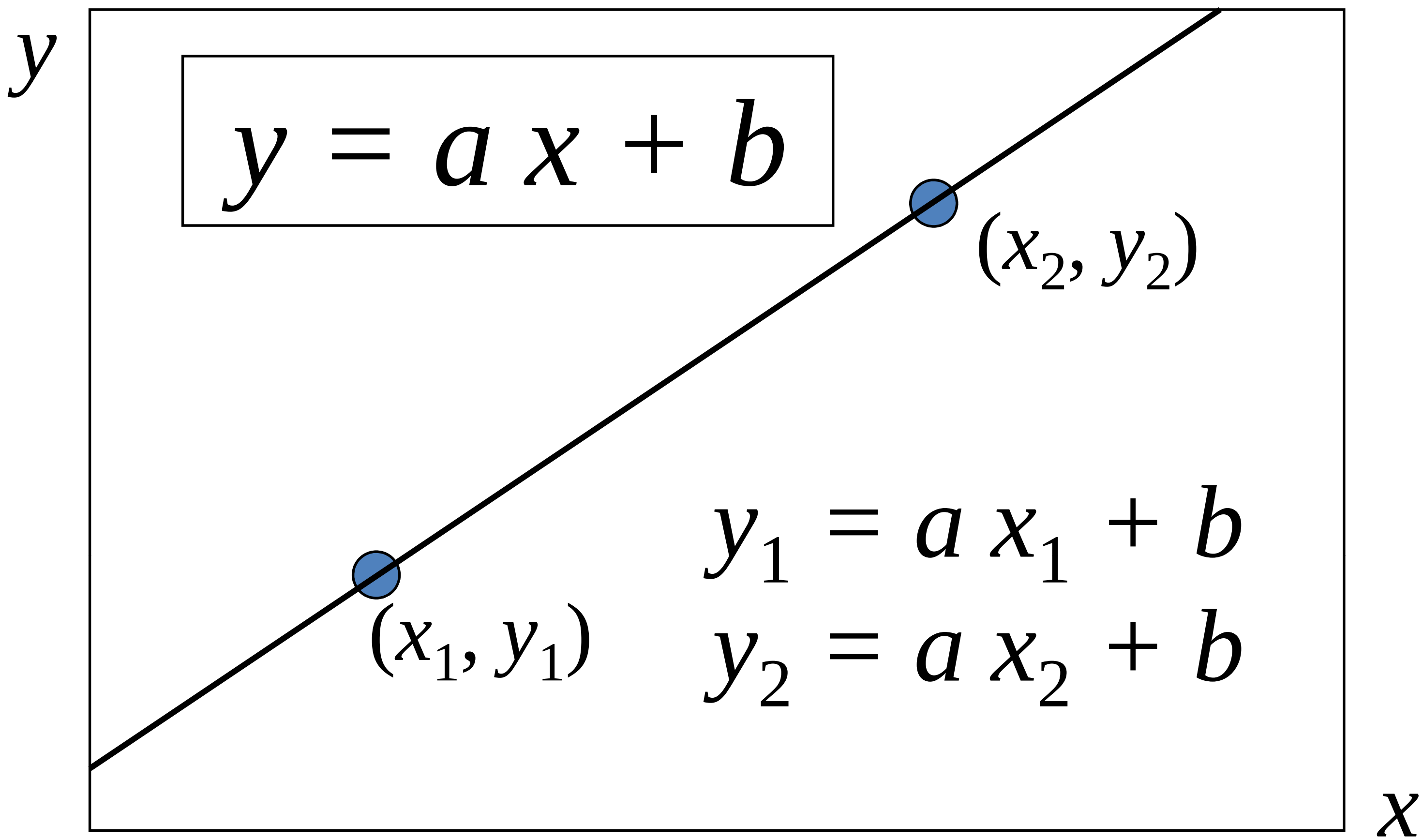
直線当てはめとは



直線当てはめとは



直線当てはめとは



未知数：a, b

2点の座標値から直線を求める

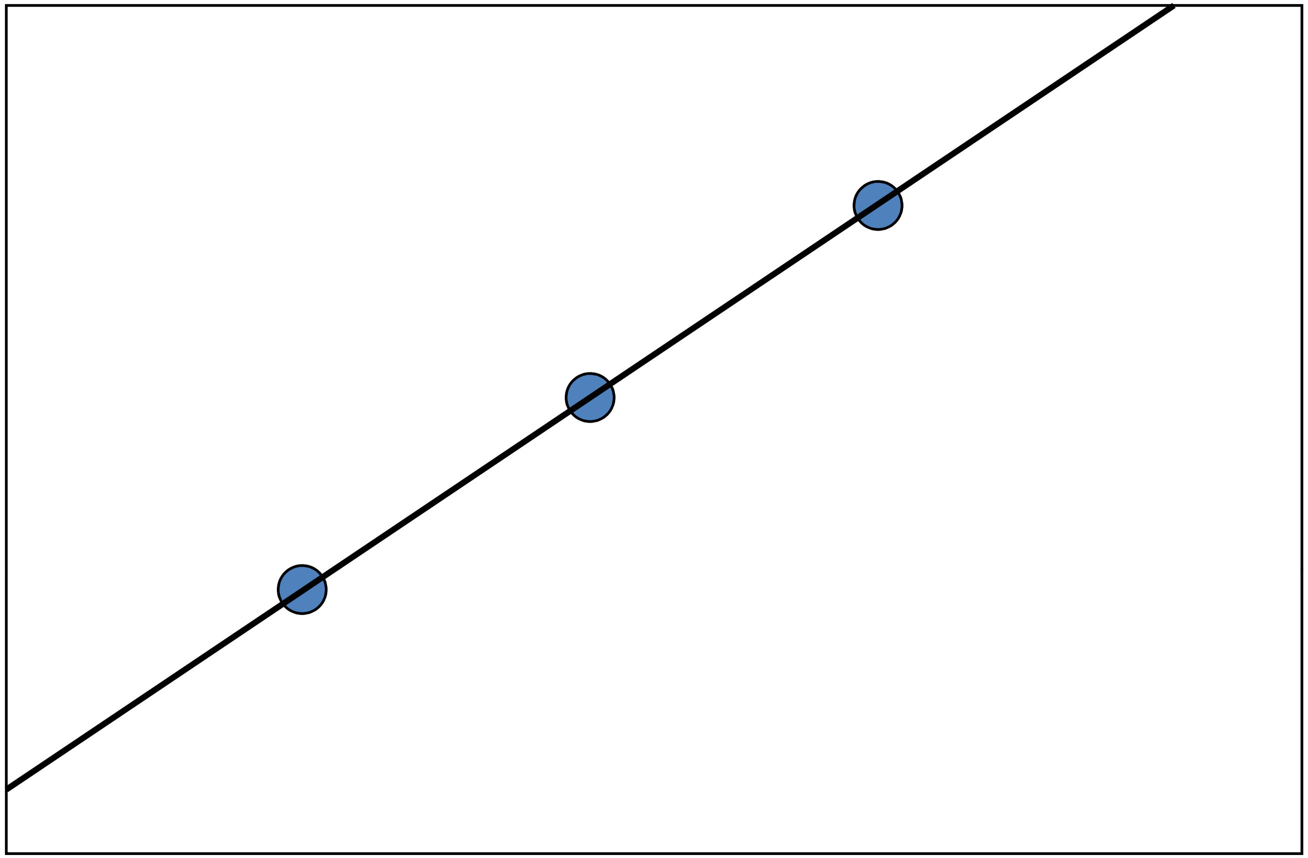
$$y_1 = a\,x_1 + b$$
$$y_2 = a\,x_2 + b$$

未知数：a, b

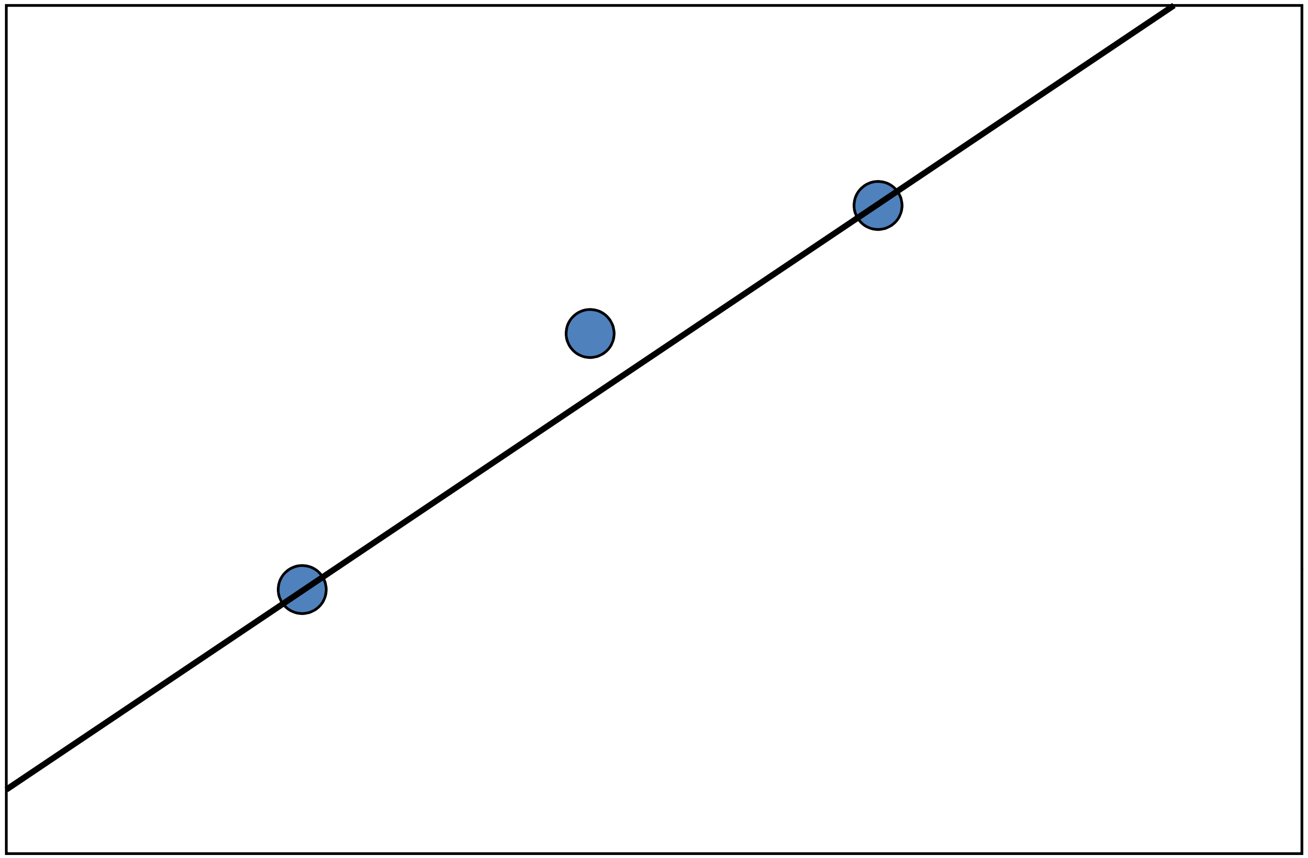
$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

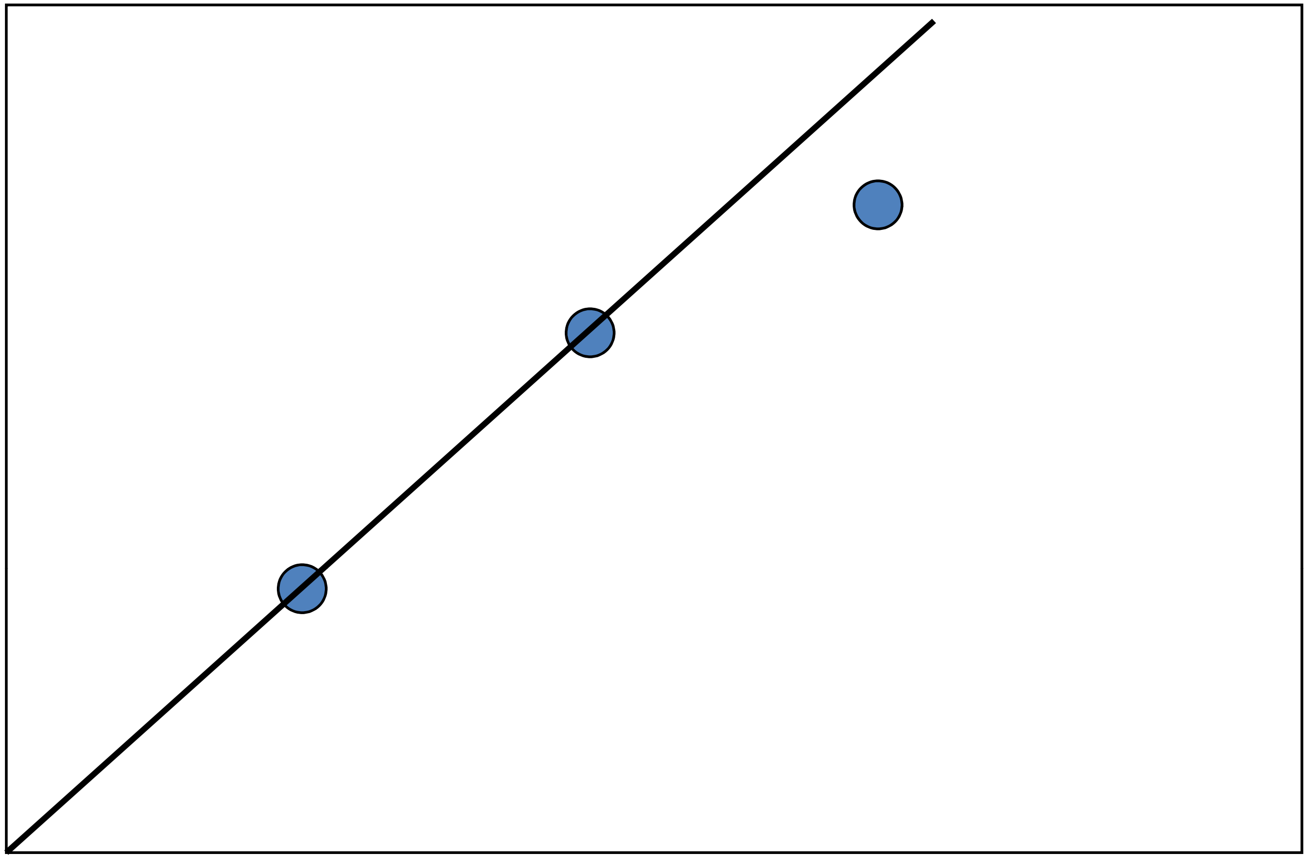
直線当てはめとは



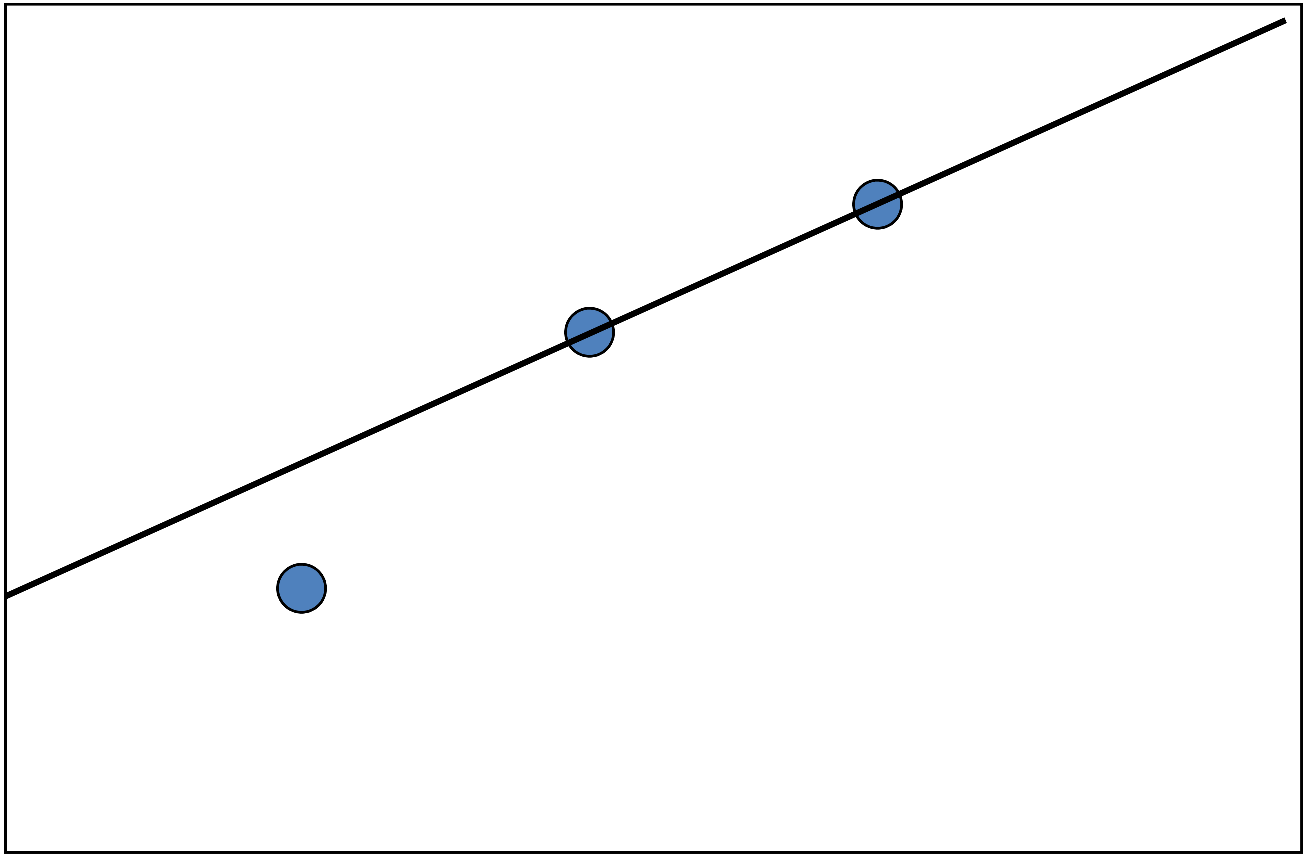
直線当てはめとは



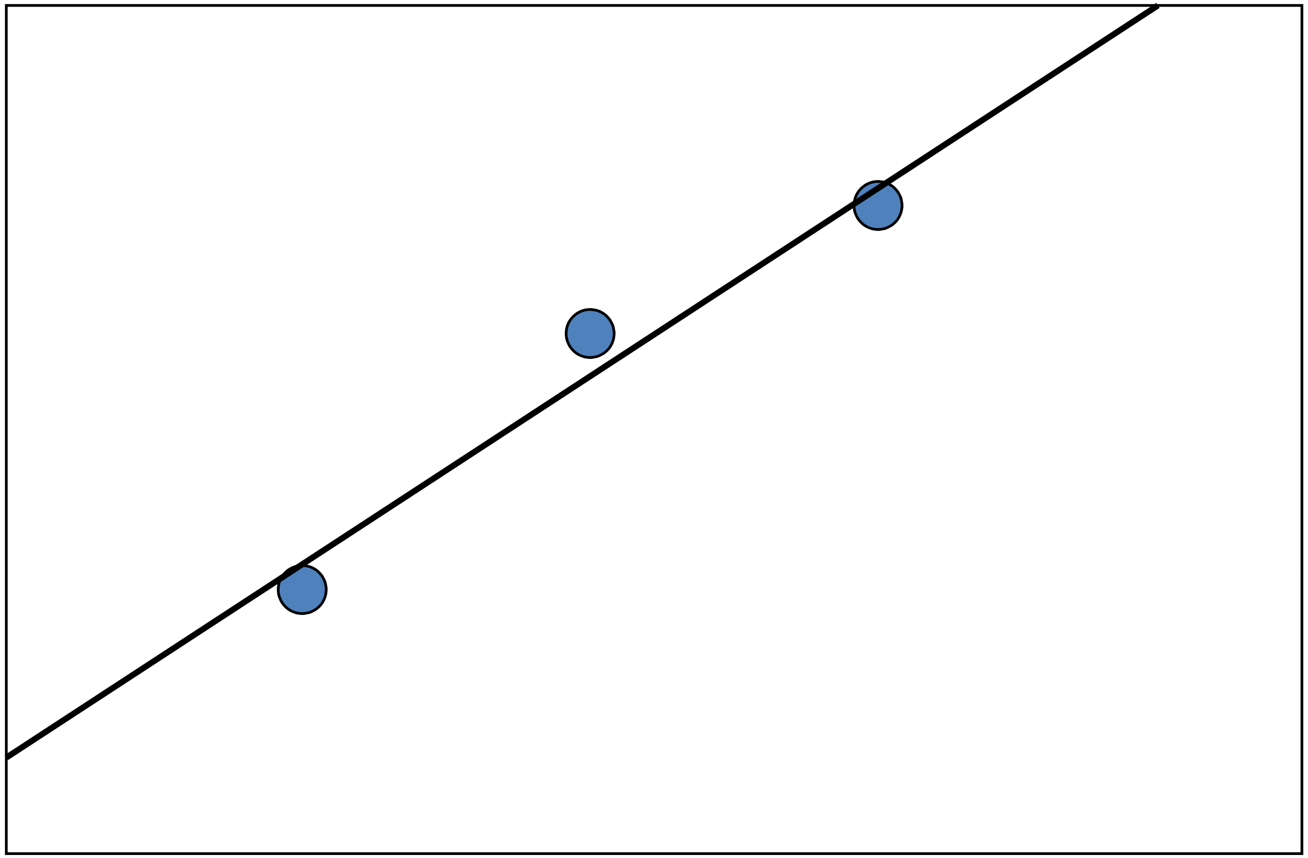
直線当てはめとは



直線当てはめとは

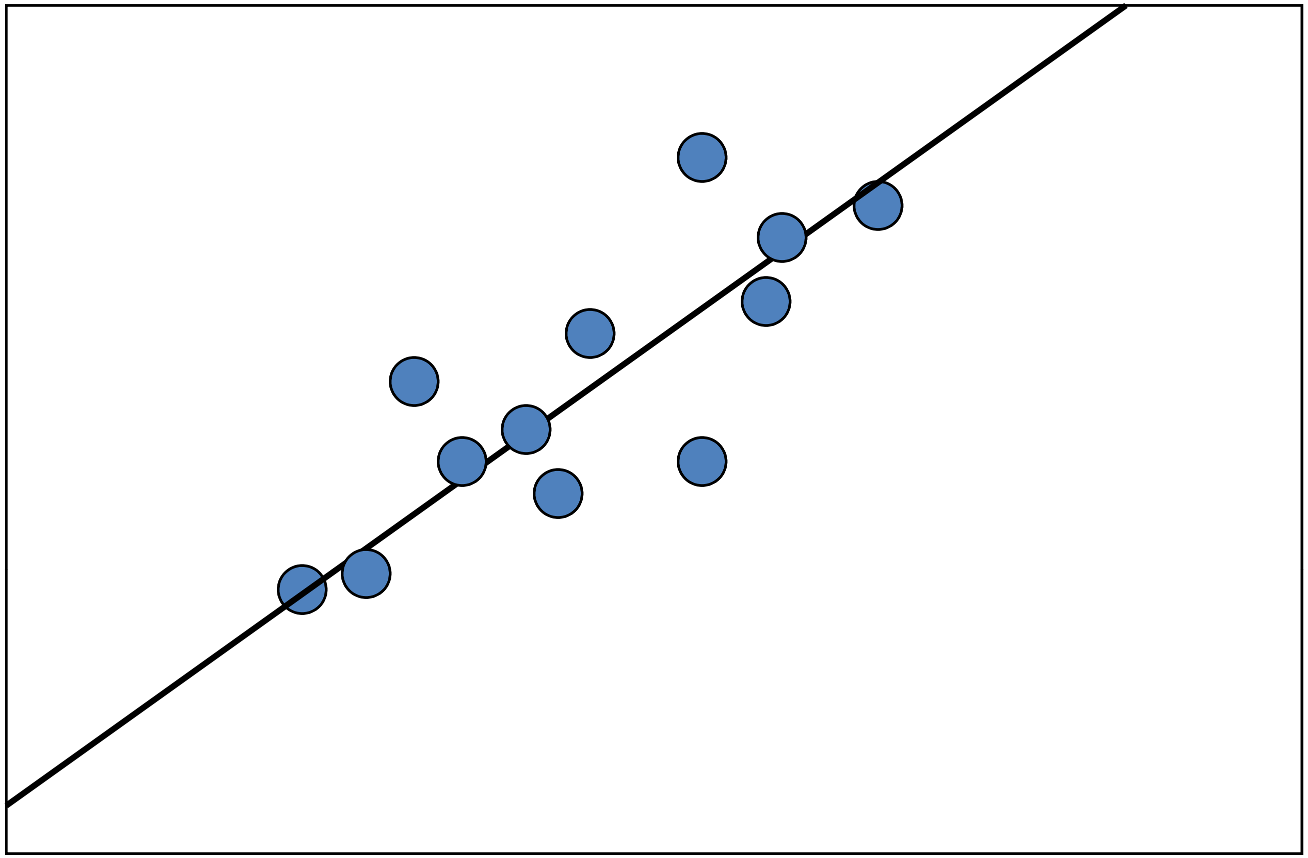


直線当てはめとは



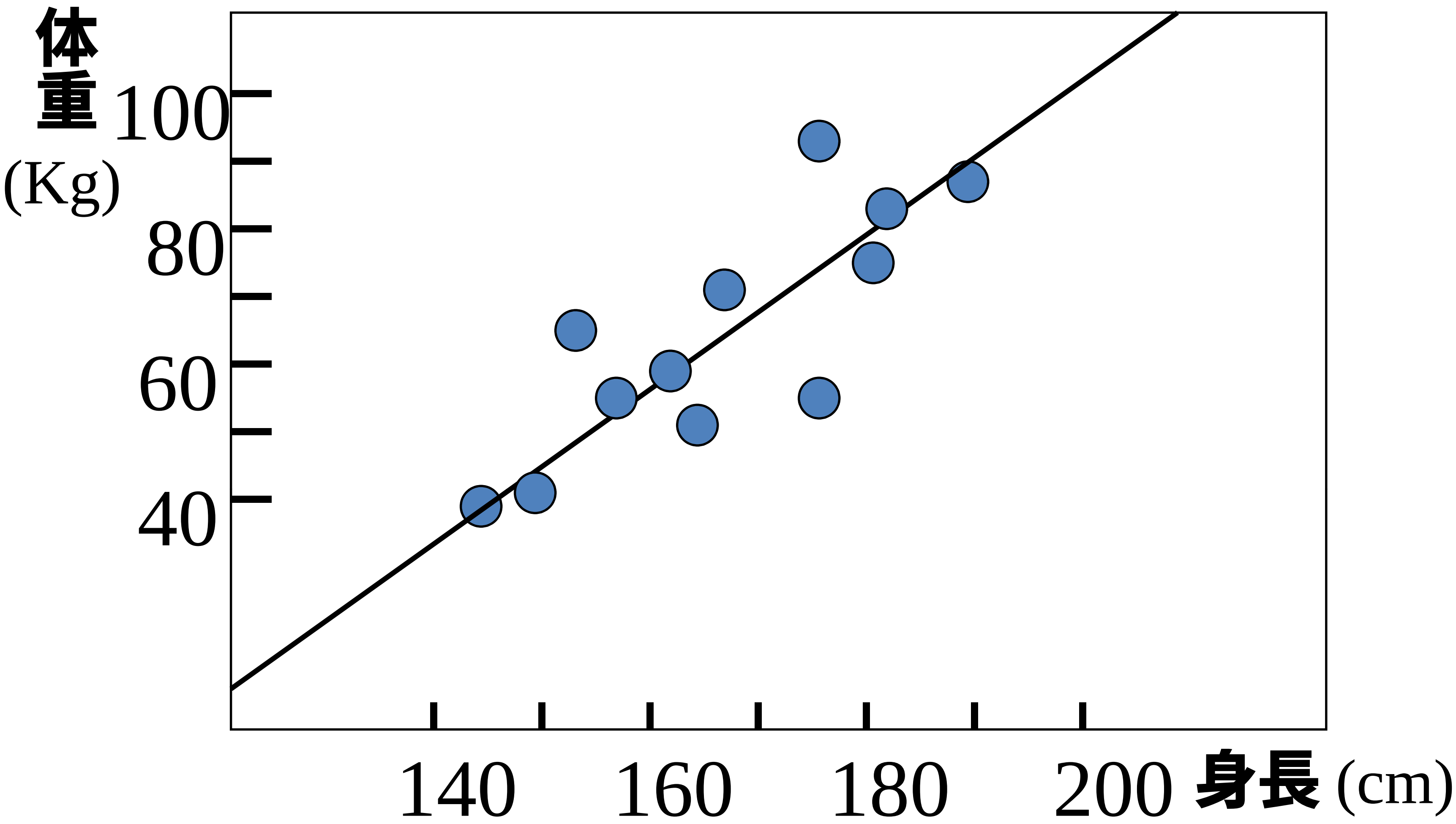
点の数が３以上になれば、直線当てはめを行うためには、何らかの妥協が必要である。

直線当てはめとは

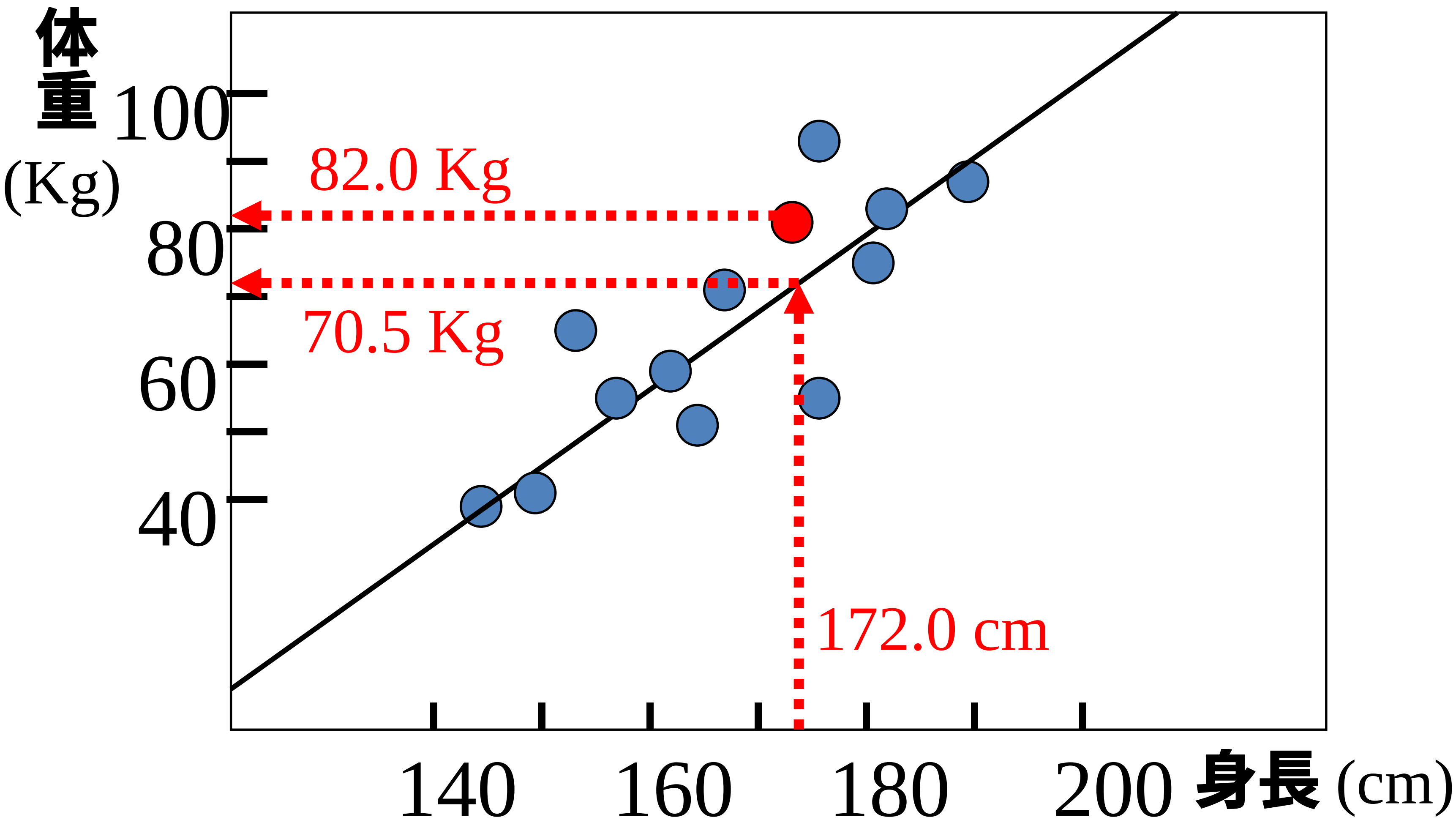


点の数が多くなって、妥当な直線当てはめを行うための基準が必要である。

直線当てはめの応用

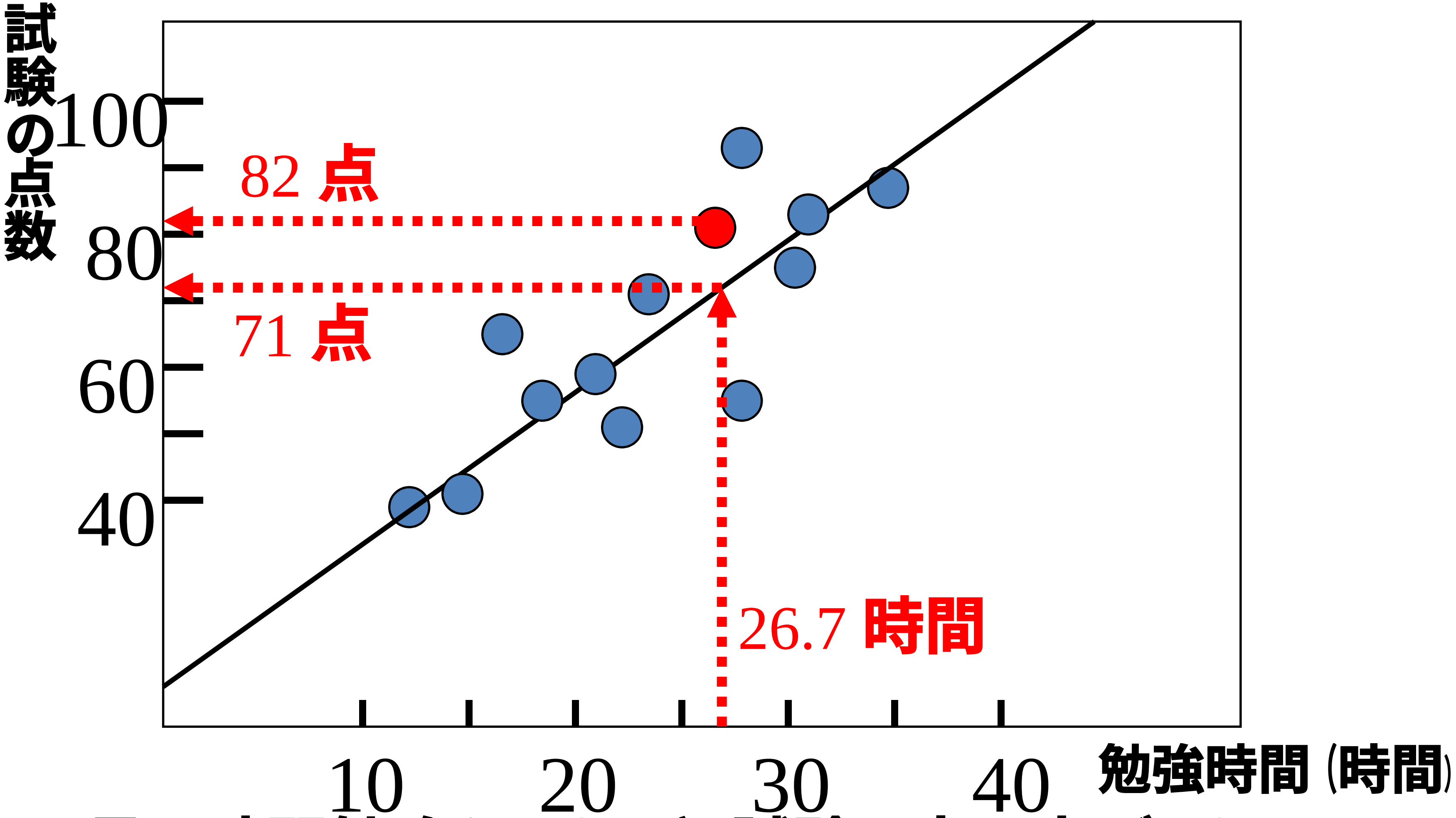


直線当てはめの応用



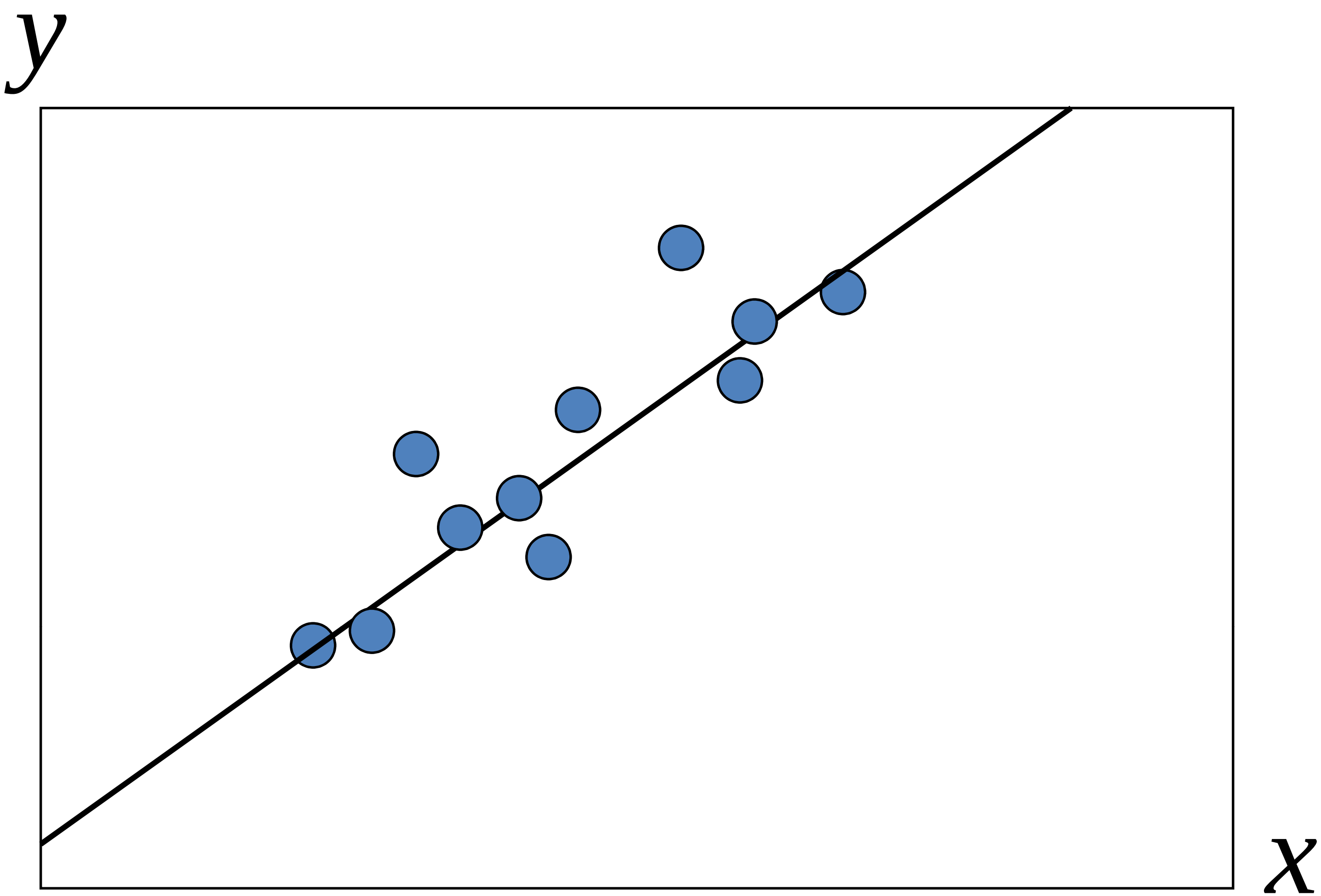
自分の身長から平均的体重を知ることができる。

直線当てはめの応用

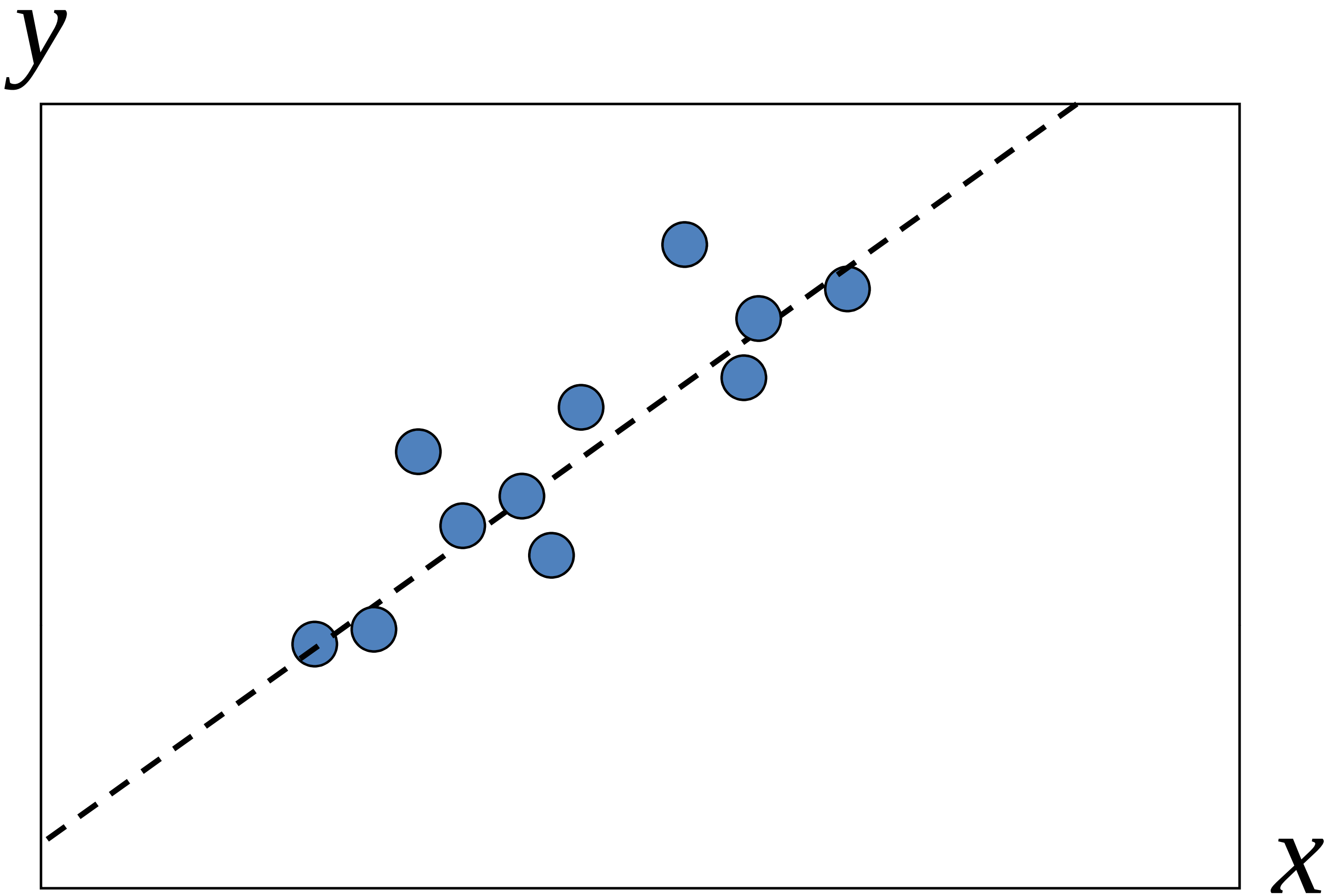


長い時間勉強をすれば、試験で良い点がとれる。

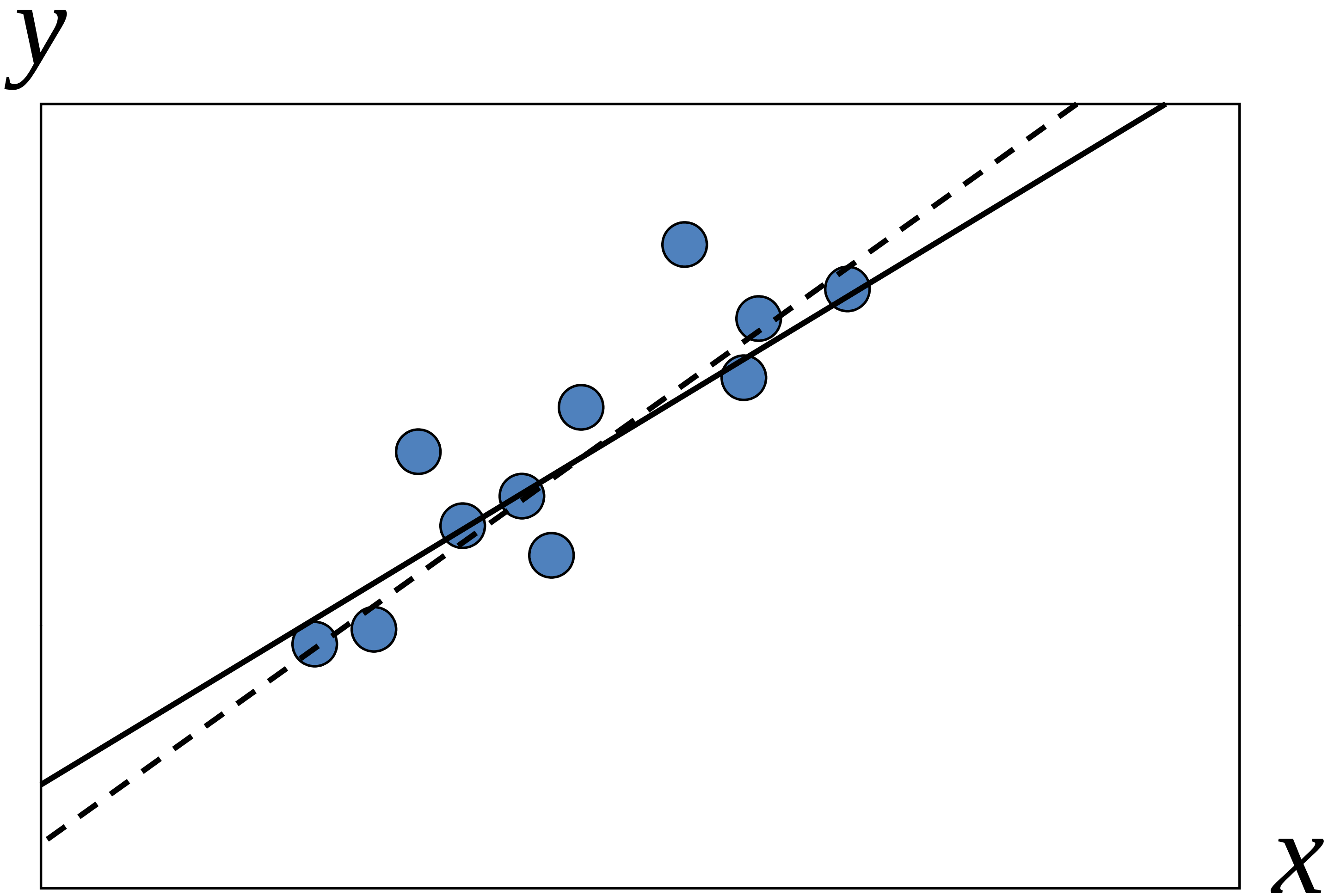
最小二乗(Least Squares)基準



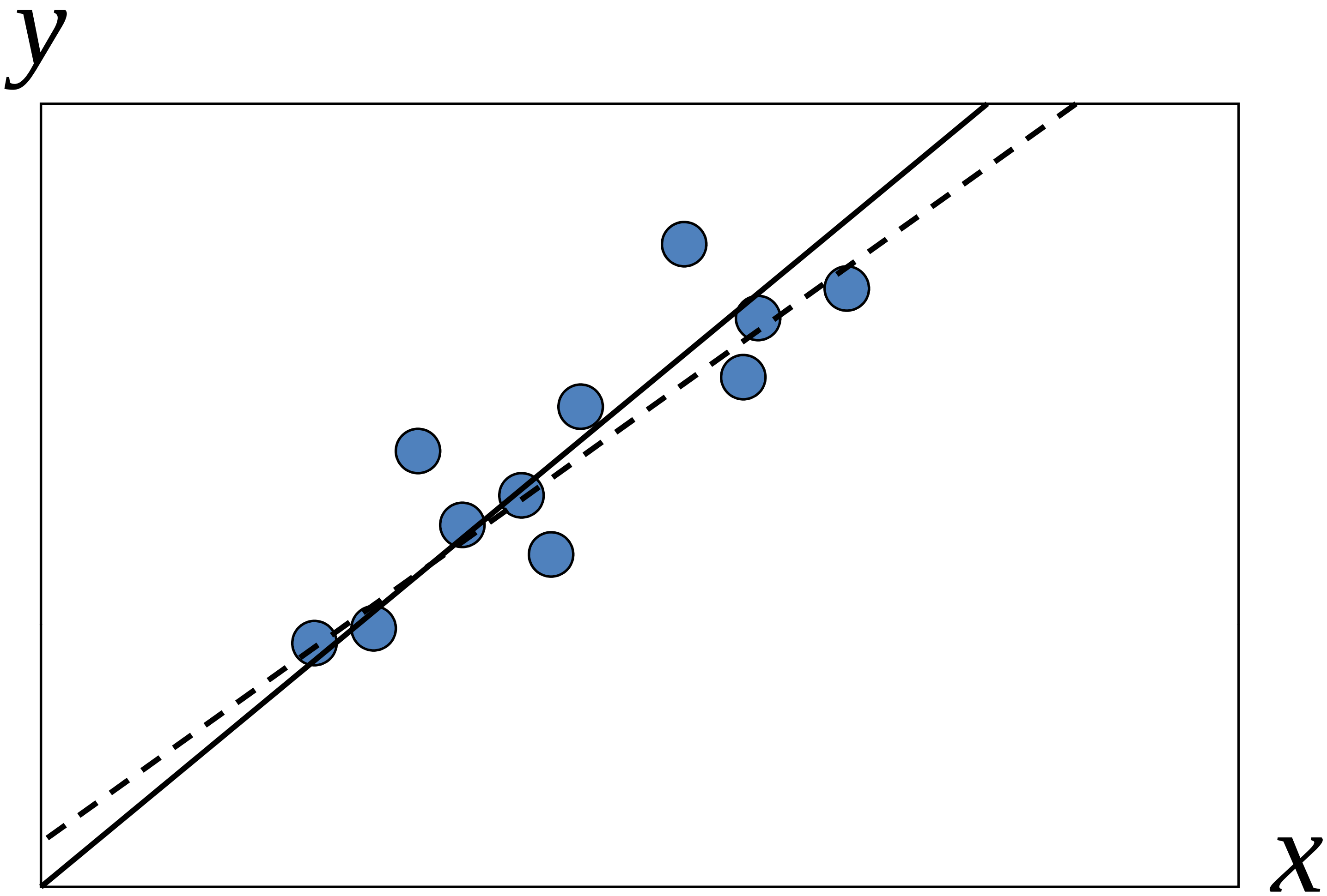
最小二乗基準



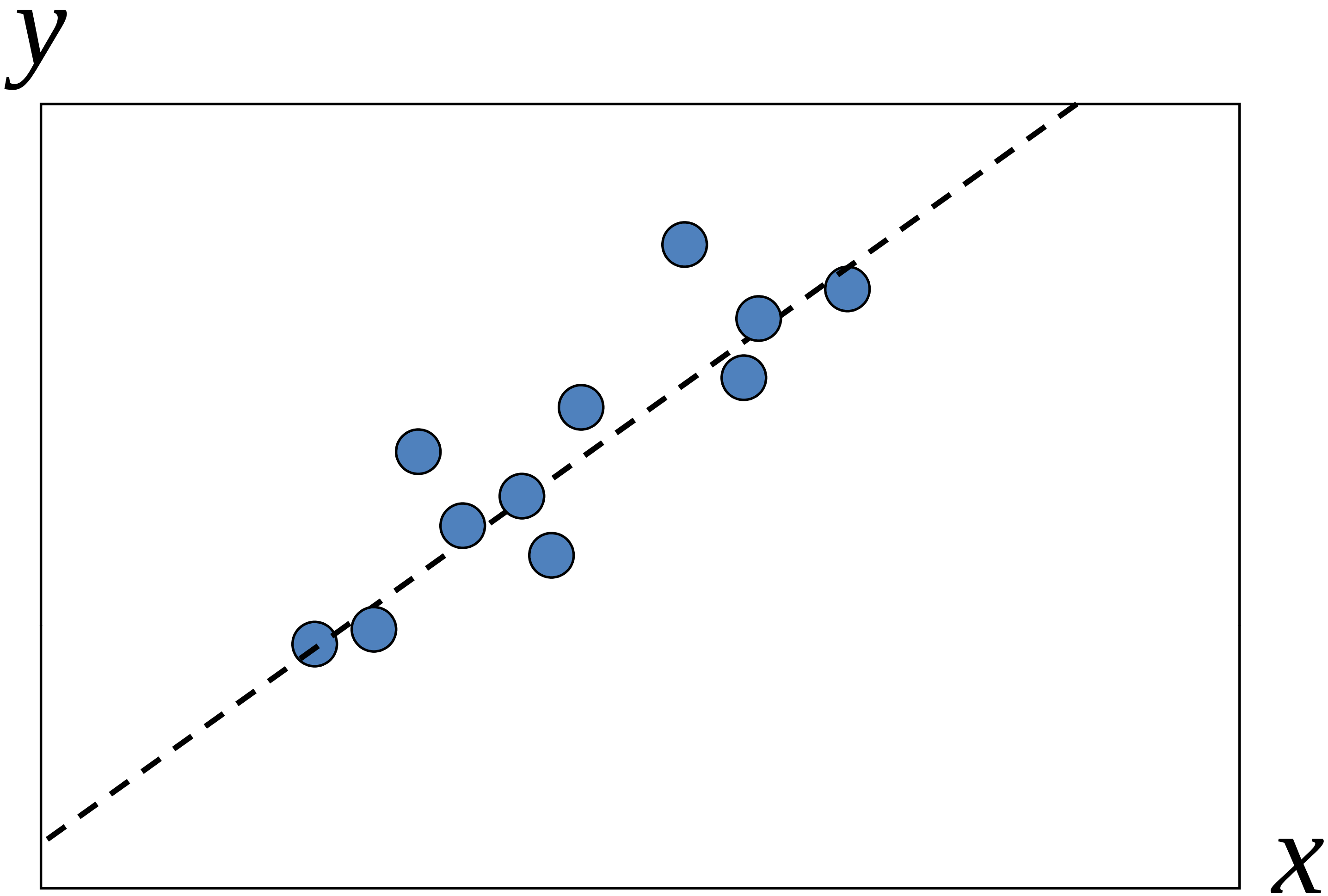
最小二乗基準



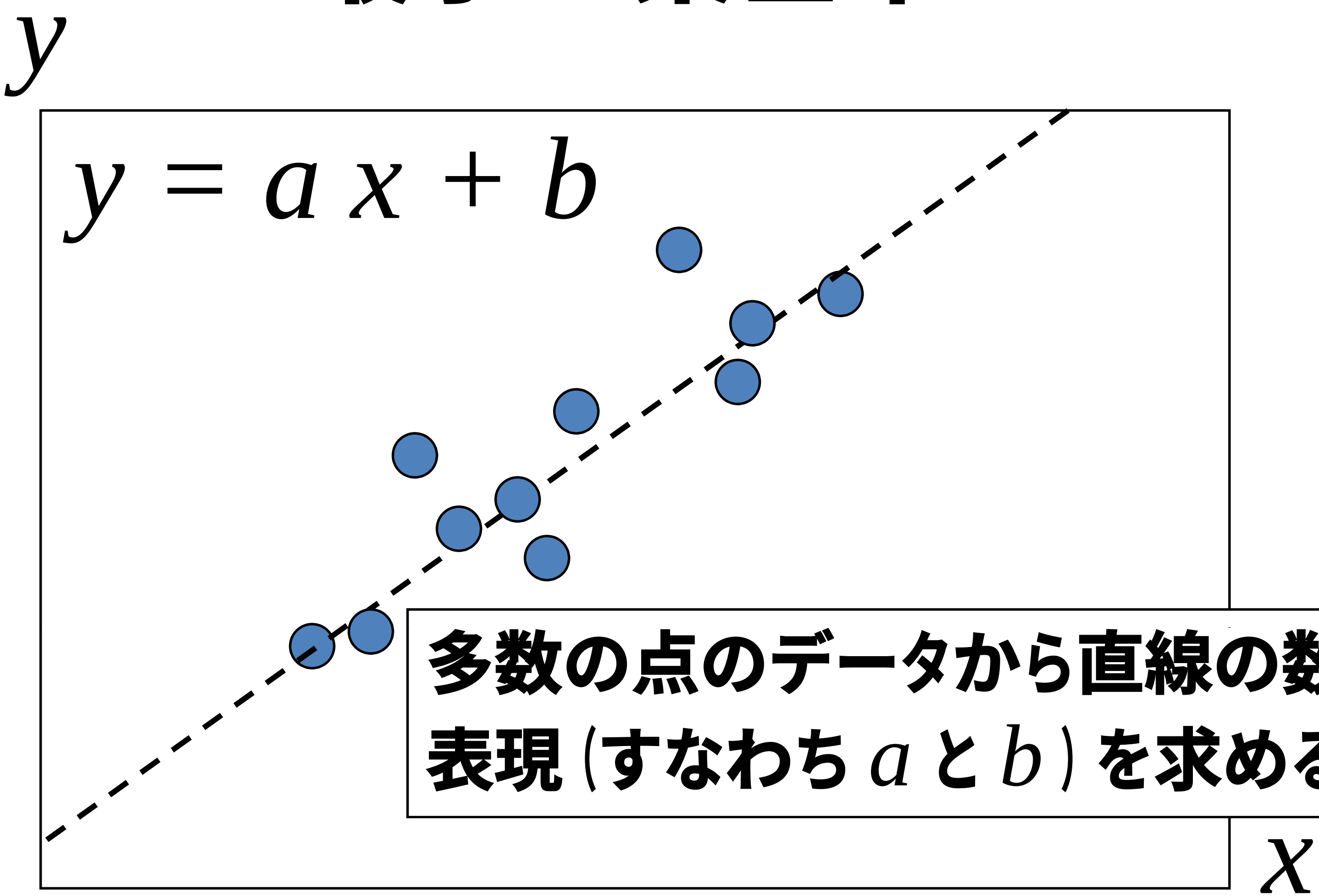
最小二乗基準



最小二乗基準

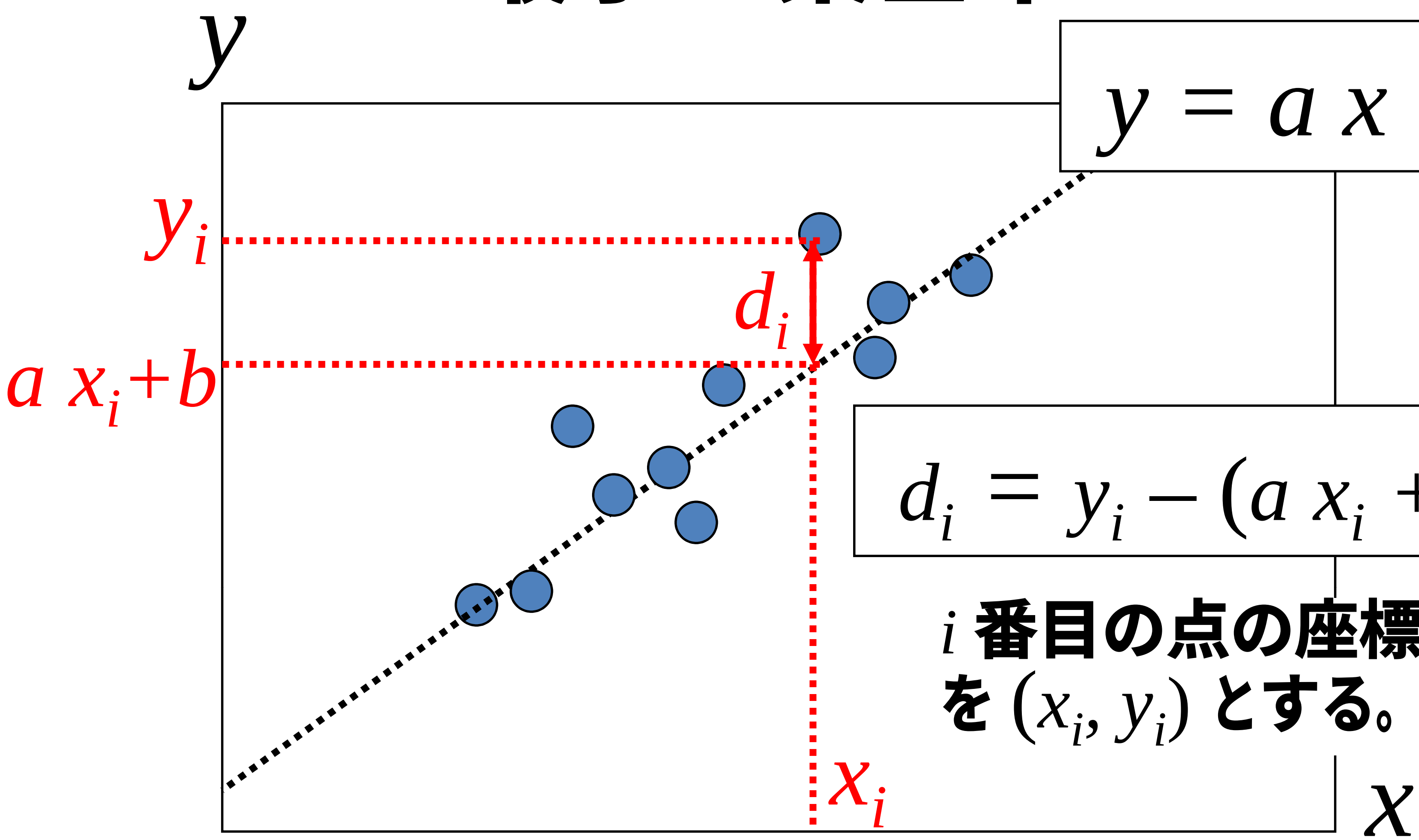


最小二乗基準



多数の点のデータから直線の数式表現（すなわち a と b）を求める。

最小二乗基準



$$y = a\,x + b$$

$$d_i = y_i - (a\,x_i + b)$$

i 番目の点の座標値を (x_i, y_i) とする。

$\sum d_i^2 = \sum\{\, y_i - (a\,x_i + b)\}^2$ を最小にする a, b を求める。

最小二乗基準

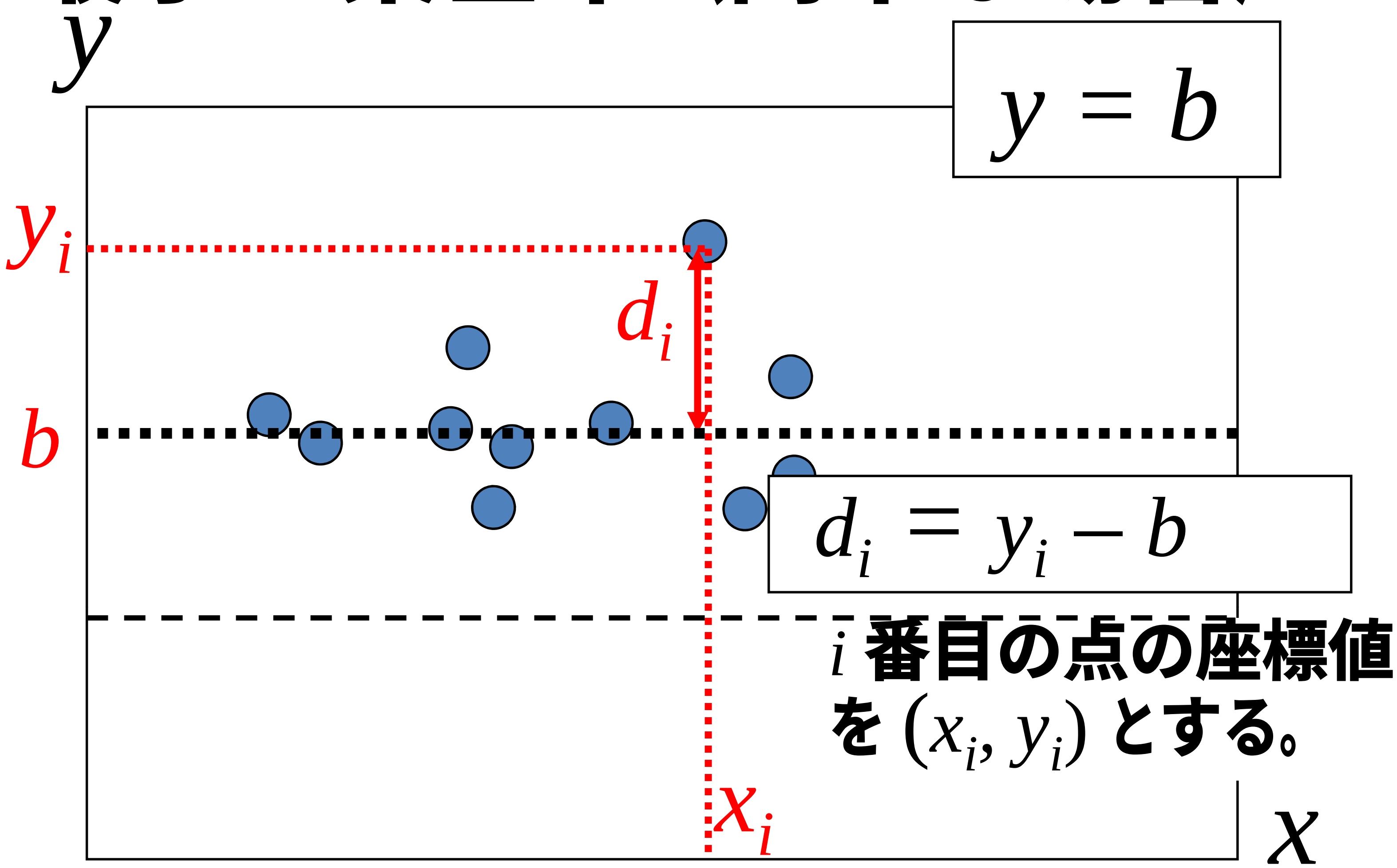
- すべての点が、直線 $y = a\,x + b$ の上になるべくのっている。

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \{ y_i - (a\,x_i + b) \}^2$$ を最小にする a, b を求める。

n 個の点があり、i 番目の点の座標値を (x_i, y_i) とする。

最小二乗法の定式化

最小二乗基準（簡単な場合）



$\sum d_i^2 = \sum\{y_i - b\}^2$ を最小にする b を求める。
（未知数は b のみ）

最小二乗基準 (簡単な場合)

- すべての点が、直線 $y = b$ の上になるべくのっている。

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \{ y_i - b \}^2$$ を最小にする b を求める。

n 個の点があり、i 番目の点の座標値を (x_i, y_i) とする。

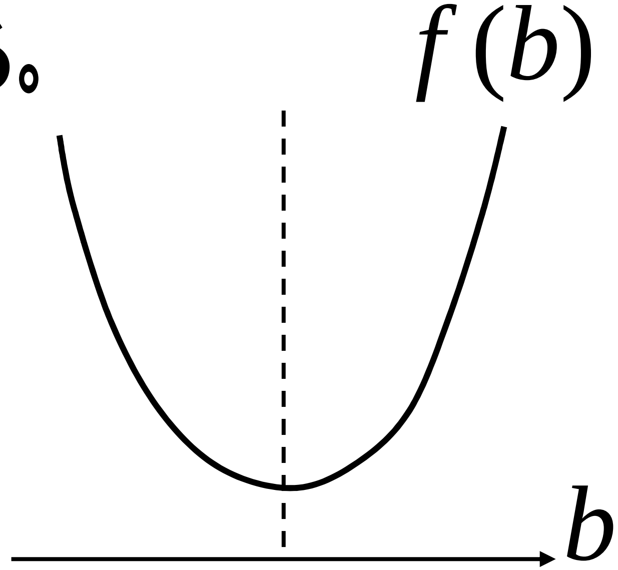
最小二乗基準 (簡単な場合)

$$f(b) = \sum_{i=1}^{n} \{ y_i - b \}^2$$ を最小にする b を求める。

$f(b)$ を最小にする b を求める。

$$\frac{\mathrm{d}}{\mathrm{d}b} f(b) = 0$$

b に関する微分が 0 になる。

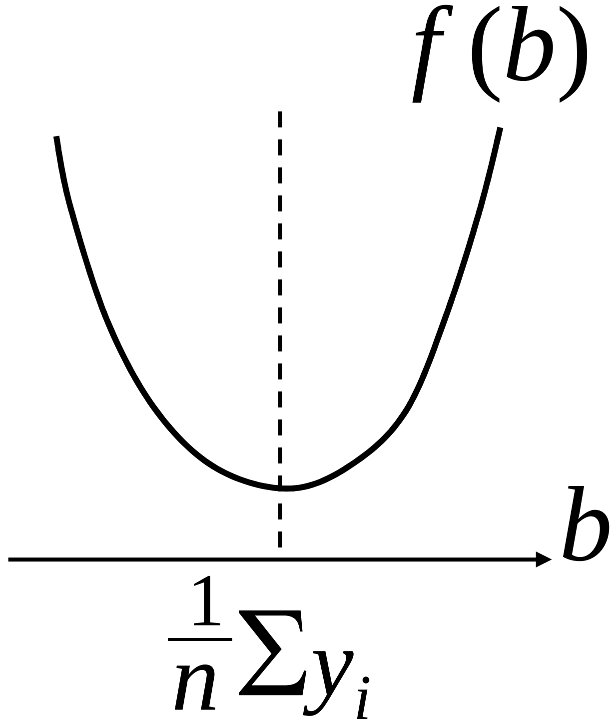


最小二乗基準（簡単な場合）

$$f(b) = \sum_{i=1}^{n}\{y_i - b\}^2$$

$$= \sum\{y_i{}^2 - 2y_ib + b^2\}$$

$$= \sum y_i{}^2 - 2b\sum y_i + nb^2$$

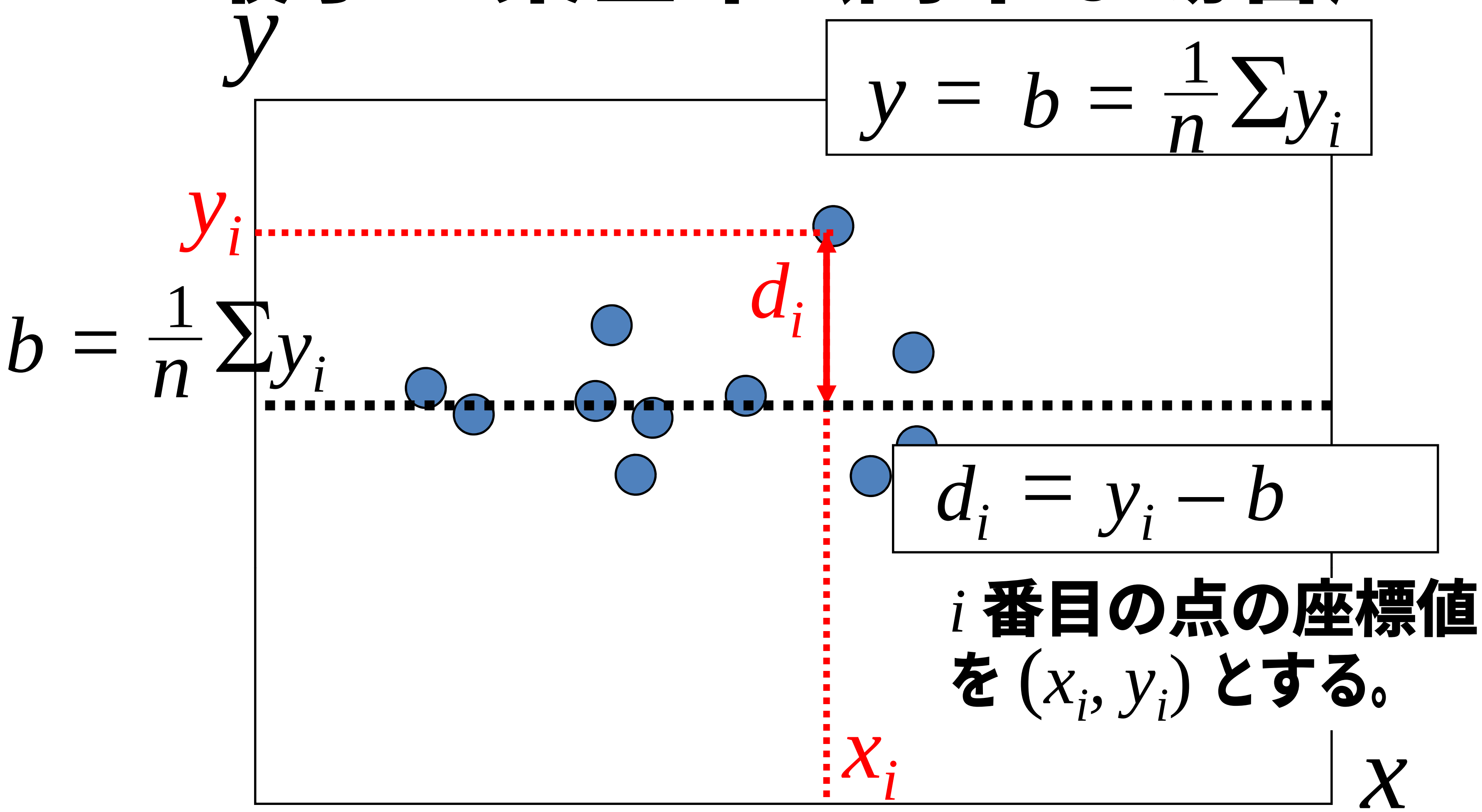


最小値をとる条件

$$\frac{\mathrm{d}}{\mathrm{d}b}f(b) = 2nb - 2\sum y_i = 2(nb - \sum y_i) = 0$$

$$b = \frac{1}{n}\sum y_i \quad (b \text{ は } y_i \text{ の平均値})$$

最小二乗基準（簡単な場合）



$\sum d_i^2 = \sum \{ y_i - b\}^2$ を最小にする b は $\frac{1}{n}\sum y_i$

最小二乗基準

$$f(a,b) = \sum_{i=1}^{n} \{ y_i - (a\,x_i + b) \}^2$$ を最小にする a, b を求める。

$f(a,b)$ を最小にする a, b を求める。

$$\frac{\partial}{\partial a} f(a,b) = 0 \qquad \frac{\partial}{\partial b} f(a,b) = 0$$

a, b に関する偏微分が 0 になる。

最小二乗基準

$f(a,b) = \sum_{i=1}^{n} \{ y_i - (a x_i + b) \}^2$ を最小にする a, b を求める。

$f(a,b)$ を最小にする a, b を求める。

$$\frac{\partial}{\partial a} f(a,b) = 0 \qquad \frac{\partial}{\partial b} f(a,b) = 0$$

a, b に関する偏微分が 0 になる。

最小二乗基準

$f(a,b) = \sum_{i=1}^{n} \{ y_i - (a\, x_i + b) \}^2$ を最小にする a, b を求める。

$f(a,b)$ を最小にする a, b を求める。

$$\frac{\partial}{\partial a} f(a,b) = 0 \qquad \frac{\partial}{\partial b} f(a,b) = 0$$

a, b に関する偏微分が 0 になる。

演習問題7

A. $f(a,b) = \sum_{i=1}^{n} \{ y_i - (a\,x_i + b)\}^2$ をa, b に関して偏微分せよ。

B. $f(a,b)$のa, b に関する偏微分を、それぞれ $f_a(a,b)$, $f_b(a,b)$ とすると、$f(a,b)$ を最小にする条件は、以下で与えられる。

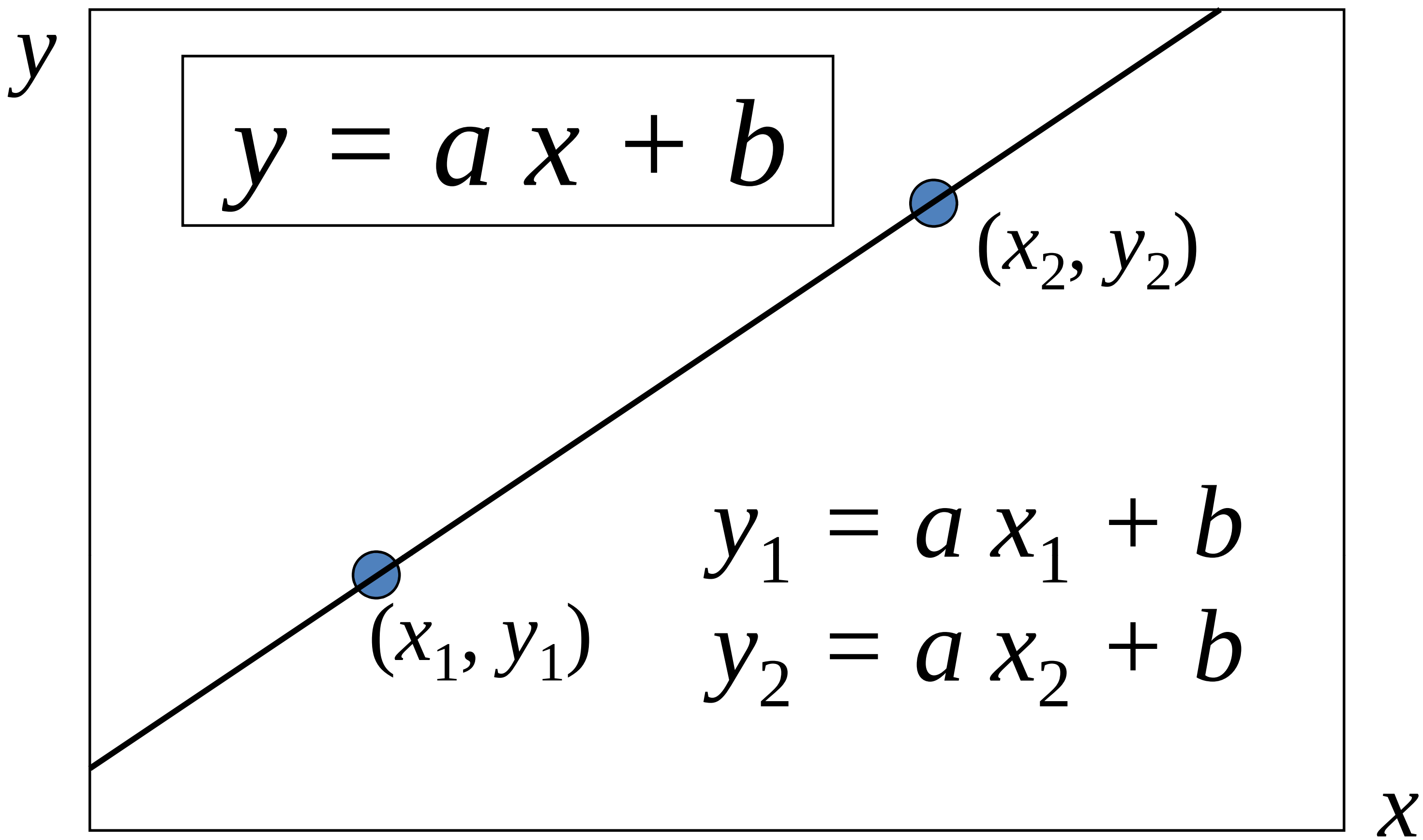
$$f_a(a,b) = 0$$
$$f_b(a,b) = 0$$

この2つの条件式から、未知数を a, b とする2元連立1次方程式が得られる。この連立方程式を行列を用いて表現せよ。

最小二乗法の別の定式化

- 一般逆行列 (擬逆行列) による定式化
  - 最小二乗基準を偏微分して導かれた最小二乗条件を、行列の式変形だけで導く。

直線当てはめ: 2点の場合



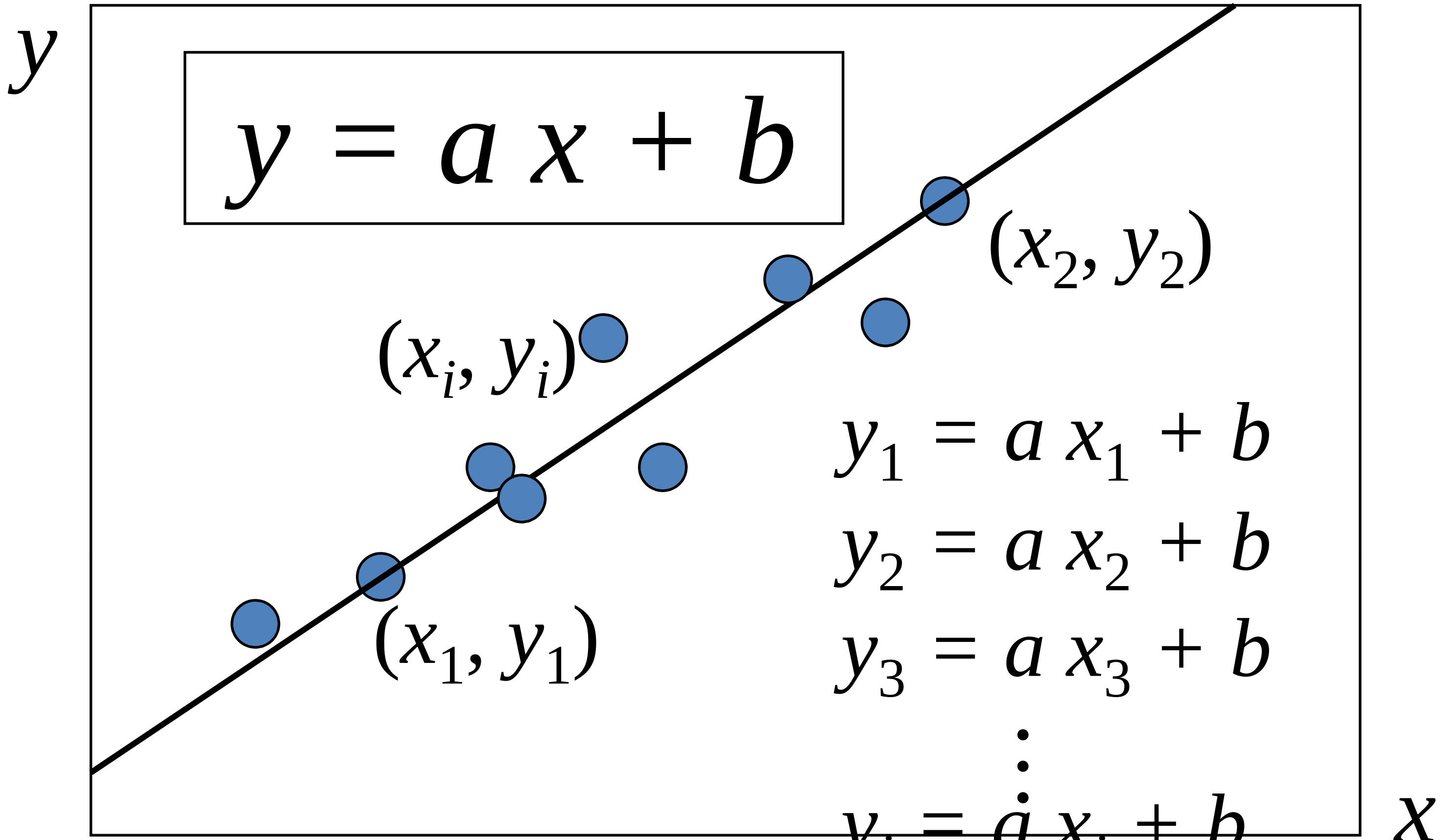
直線当てはめの定式化
2点の場合

$$y_1 = a\,x_1 + b$$
$$y_2 = a\,x_2 + b$$

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

直線当てはめ：n点の場合



$$
\begin{aligned}
y_1 &= a\,x_1 + b \\
y_2 &= a\,x_2 + b \\
y_3 &= a\,x_3 + b \\
&\vdots \\
y_i &= a\,x_i + b \\
&\vdots \\
y_n &= a\,x_n + b
\end{aligned}
$$

最小二乗法の異なる定式化
一般逆行列 (擬逆行列)
(Generalized/Pseudo Inverse)

$$
\begin{aligned}
y_1 &= a\,x_1 + b \\
y_2 &= a\,x_2 + b \\
y_3 &= a\,x_3 + b \\
&\vdots \\
y_i &= a\,x_i + b \\
&\vdots \\
y_n &= a\,x_n + b
\end{aligned}
\quad \Rightarrow \quad
\begin{aligned}
a\,x_1 + b &= y_1 \\
a\,x_2 + b &= y_2 \\
a\,x_3 + b &= y_3 \\
&\vdots \\
a\,x_i + b &= y_i \\
&\vdots \\
a\,x_n + b &= y_n
\end{aligned}
\quad \Rightarrow \quad
\begin{pmatrix}
x_1 & 1 \\
x_2 & 1 \\
x_3 & 1 \\
\vdots & \\
x_i & 1 \\
\vdots & \\
x_n & 1
\end{pmatrix}
\begin{pmatrix} a \\ b \end{pmatrix}
=
\begin{pmatrix}
y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_i \\ \vdots \\ y_n
\end{pmatrix}
$$

最小二乗法の異なる定式化
一般逆行列（擬逆行列）

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$\Downarrow$$

$$\begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ 1 & 1 & \dots & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\mathrm{n}} \end{pmatrix}$$

最小二乗法の異なる定式化
一般逆行列（擬逆行列）

$$
\begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ 1 & 1 & \ldots & 1 \end{pmatrix}
\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \\ x_n & 1 \end{pmatrix}
\begin{pmatrix} a \\ b \end{pmatrix}
=
\begin{pmatrix} x_1 & x_2 & \ldots & x_n \\ 1 & 1 & \ldots & 1 \end{pmatrix}
\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{\mathrm{n}} \end{pmatrix}
$$

$$
\begin{pmatrix} \sum x_i{}^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}
\begin{pmatrix} a \\ b \end{pmatrix}
=
\begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}
$$

最小二乗法の異なる定式化
一般逆行列（擬逆行列）

$$
\begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}
$$

$$
\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{pmatrix}^{-1} \begin{pmatrix} \sum x_i y_i \\ \sum y_i \end{pmatrix}
$$

最小二乗法による直線当てはめ
補足：用語について

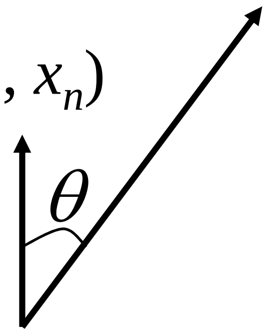
- 最小二乗法による直線当てはめは、「回帰分析 (Regression Analysis)」とも呼ばれる。
- 直線あてはめを行うことは、「回帰」と呼ばれる。
- 直線の傾き a は、「回帰係数」と呼ばれる。
- 当てはめた直線は、「回帰直線」とも呼ばれる。
- 誤差(Error)は、「残差(Residual)」とも呼ばれる。
- データ点の分布を示したグラフは、「散布図」と呼ばれる。

相関係数(Correlation Coefficient) R, 決定係数 R^{2}値

$y = (y_1, y_2, \ldots, y_n)$

- **相関係数R とは？**

$x = (x_1, x_2, \ldots, x_n)$



$$R = \cos\theta = \frac{x \cdot y}{|x| \cdot |y|}$$

$$-1 \leqq R \leqq 1$$

ベクトルの内積

$$x \cdot y = \Sigma x_i y_i = x_1 y_1 + x_2 y_2 + \ldots. + x_n y_n$$
$$= |x| \cdot |y| \cos\theta$$

$R = -1$　**負の (完全な) 相関**
$R = 0$　**無相関 (関連性なし)**
$R = 1$　**正の (完全な) 相関**

ベクトル大きさ

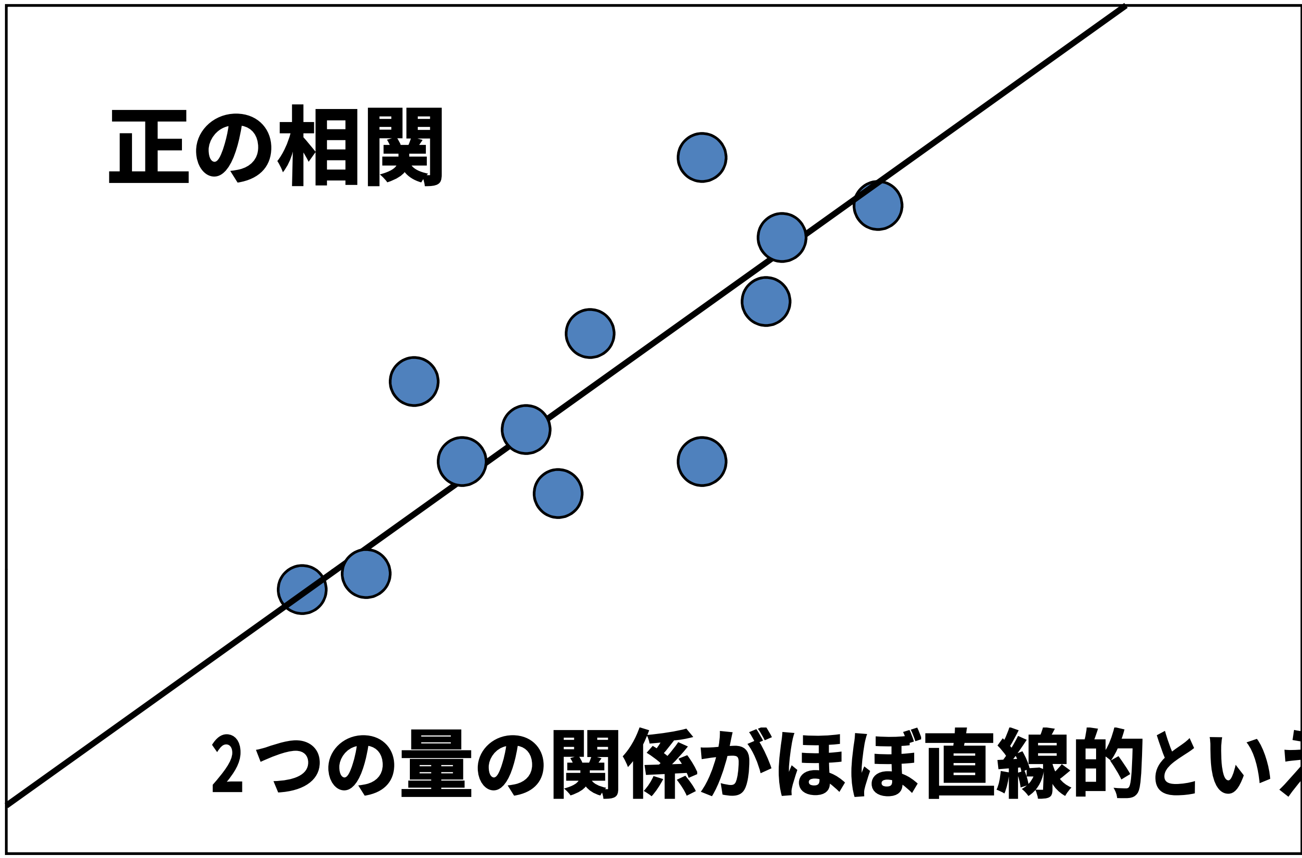
$$|x|^2 = \Sigma x_i^2 = x_1^2 + x_2^2 + \ldots. + x_n^2$$

- **決定係数** (直線当てはめの信頼性) **は、相関係数の二乗** R^2

$$-1 \leqq R \leqq 1$$

$R = 0$　**全く信頼できない。** $|R| = 1$　**完全に信頼できる。**

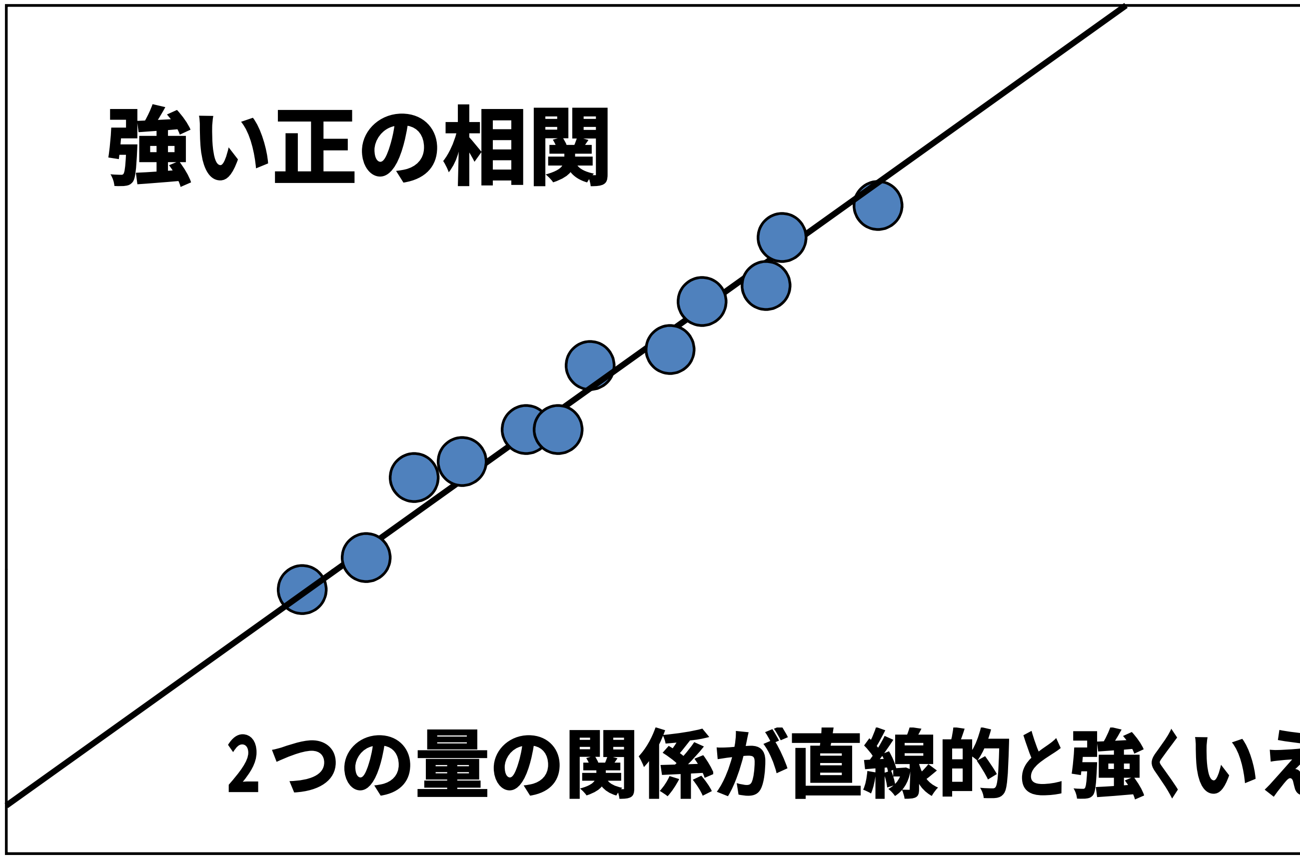
相関係数 R　決定係数 R^2 値



2つの量の関係がほぼ直線的といえる。

$R = 0.8$, $R^2 = 0.64$ くらい

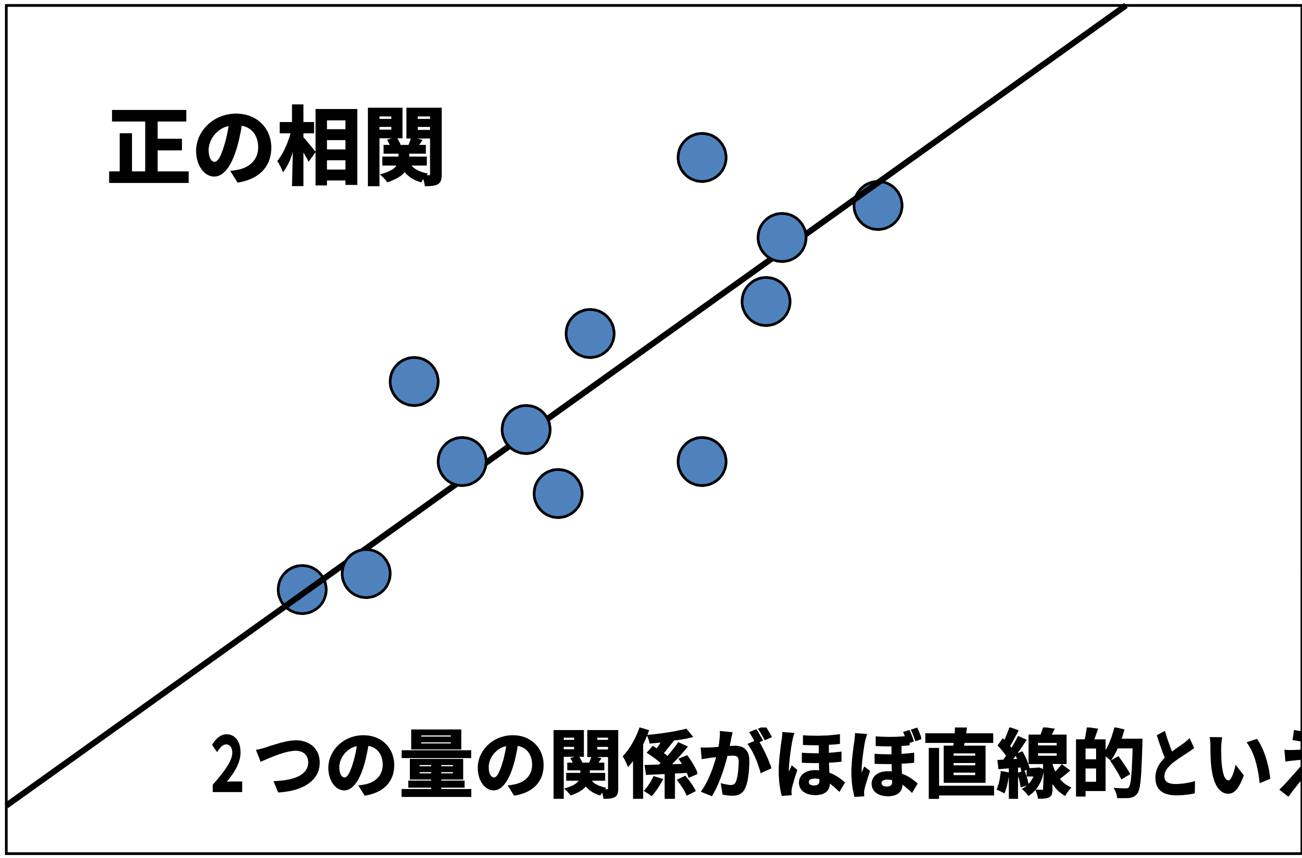
相関係数 R　決定係数 R^2 値



2つの量の関係が直線的と強くいえる。

$R = 0.95$, $R^2 = 0.9$ くらい

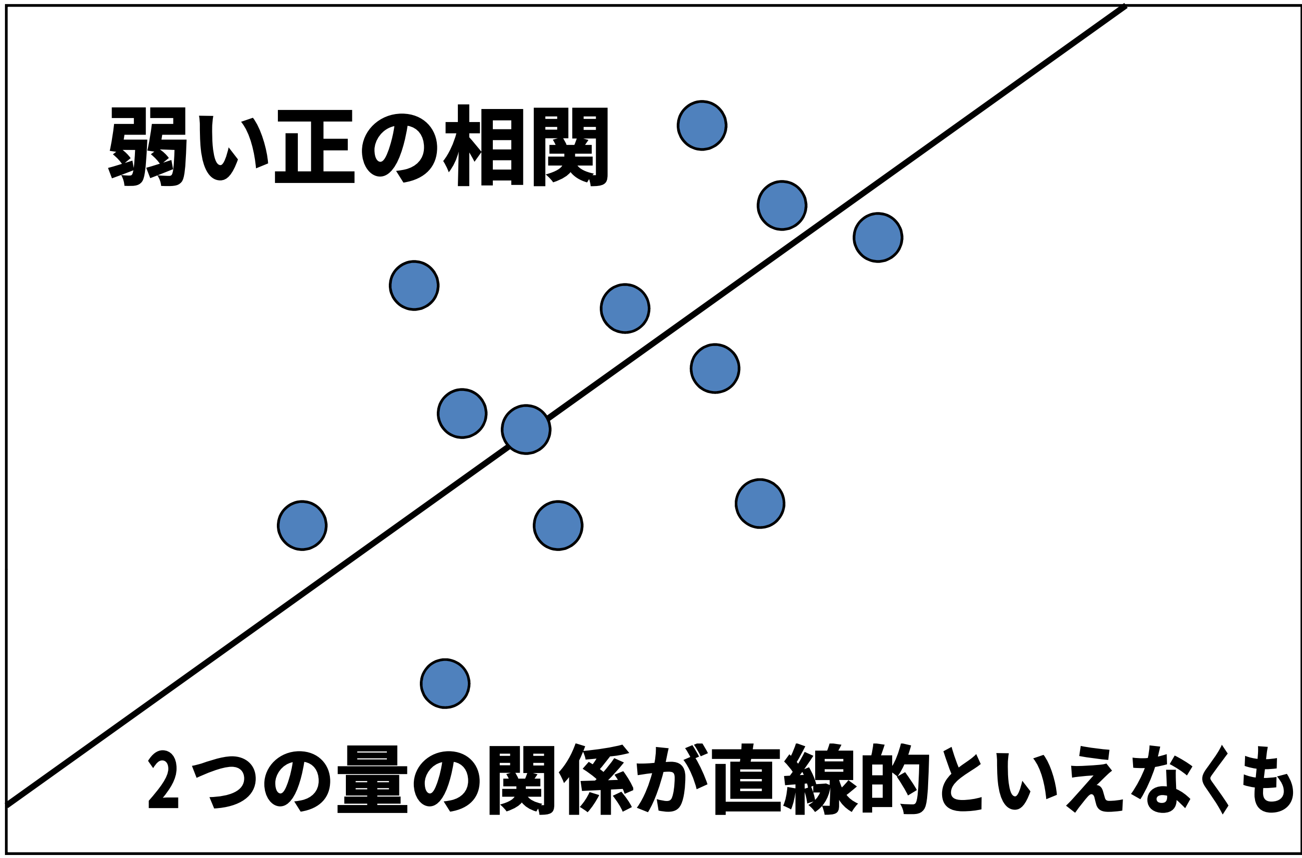
相関係数 R　決定係数 R^2 値



2つの量の関係がほぼ直線的といえる。

$R = 0.8$, $R^2 = 0.64$ くらい

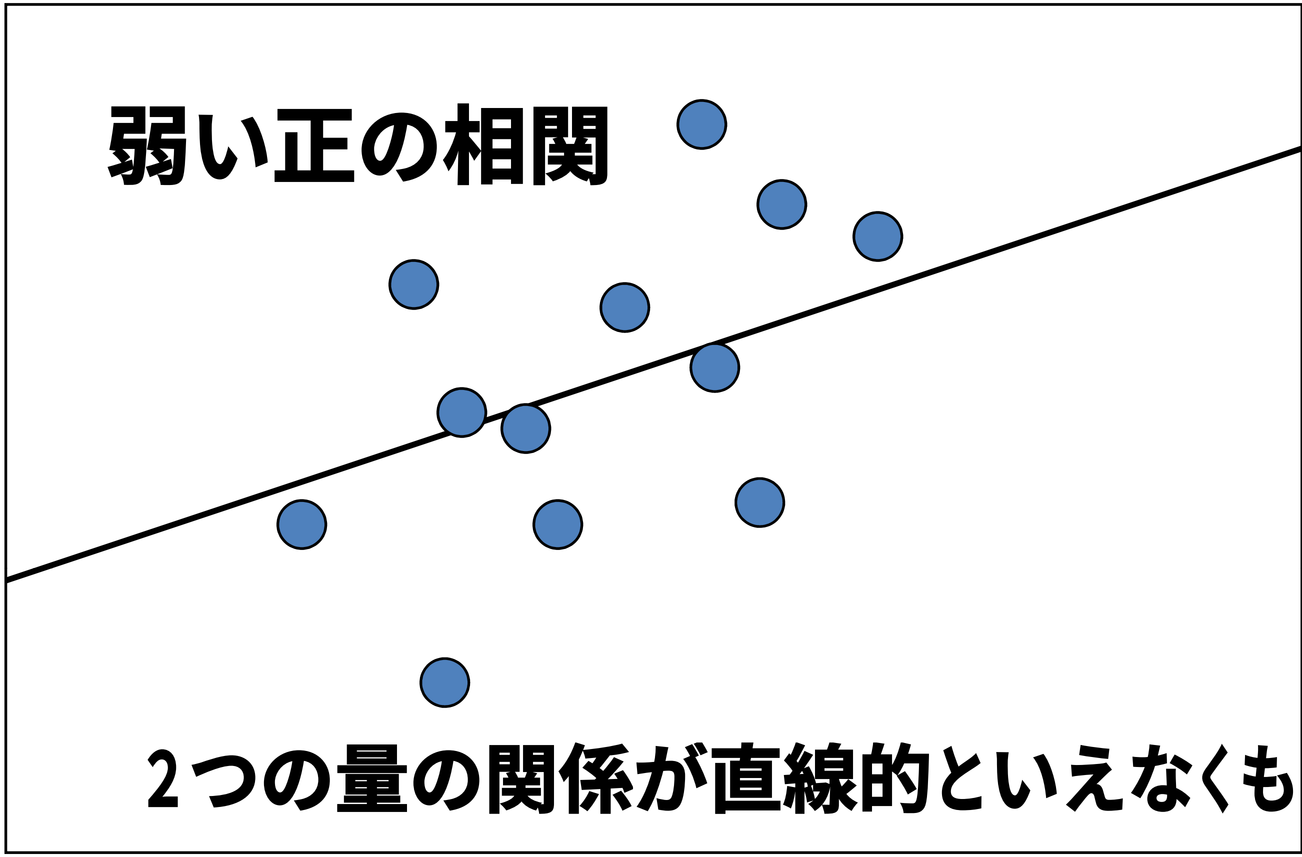
相関係数 R　決定係数 R^2 値



2つの量の関係が直線的といえなくもない。

$R = 0.5, R^2 = 0.25$ くらい

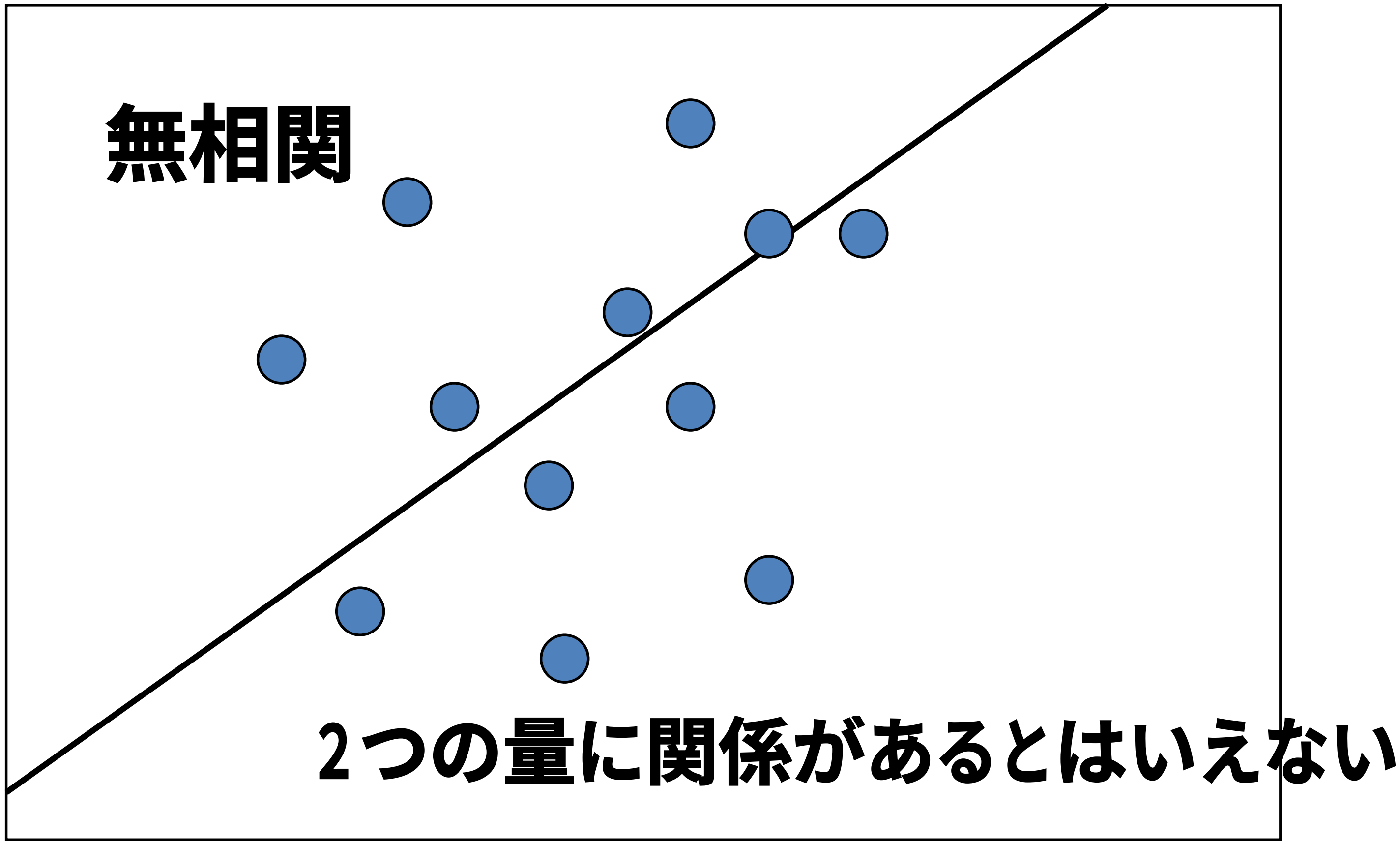
相関係数R　決定係数R^2値



2つの量の関係が直線的といえなくもない。

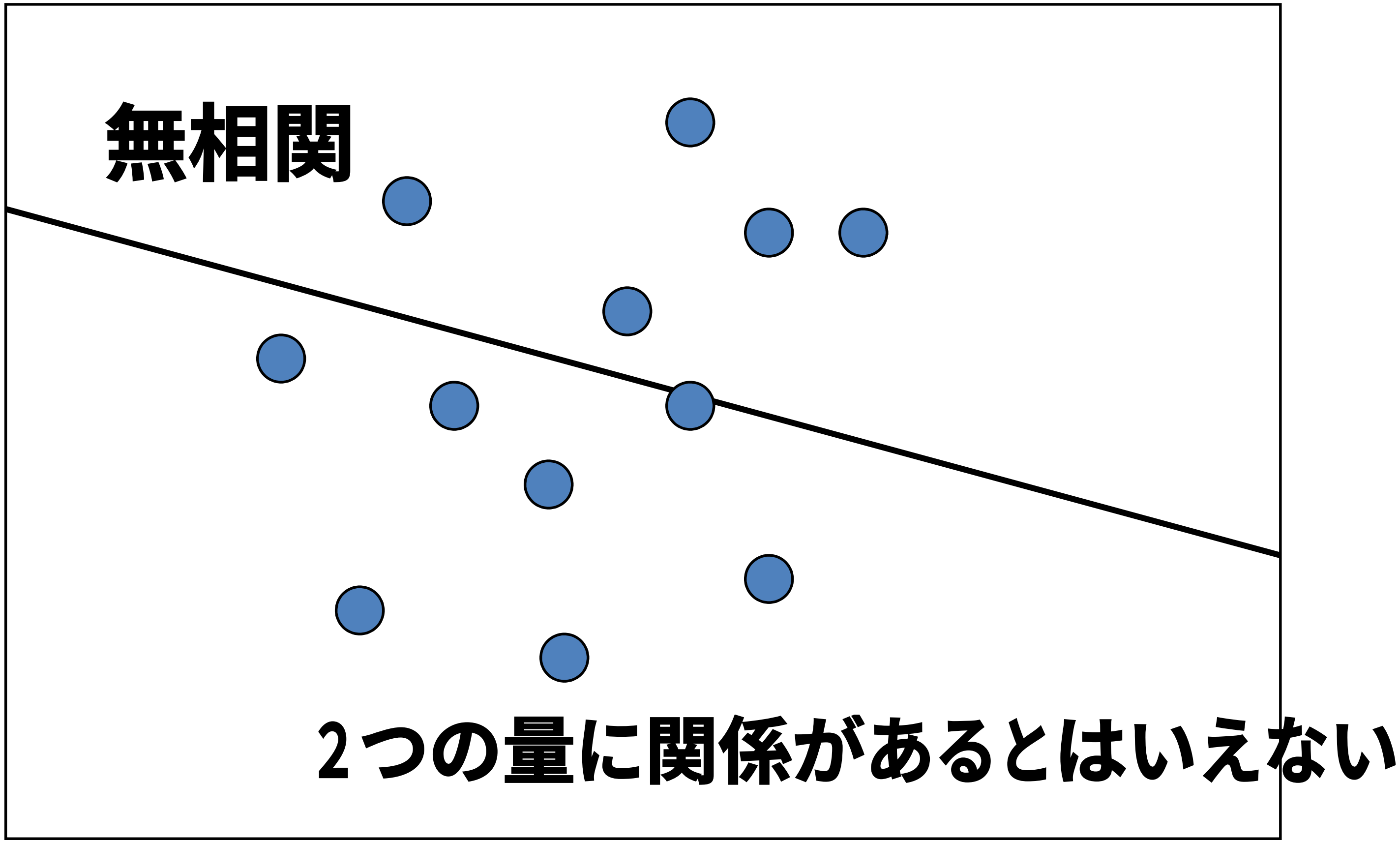
$R = 0.5$, $R^2 = 0.25$ くらい

相関係数 R　決定係数 R^2 値



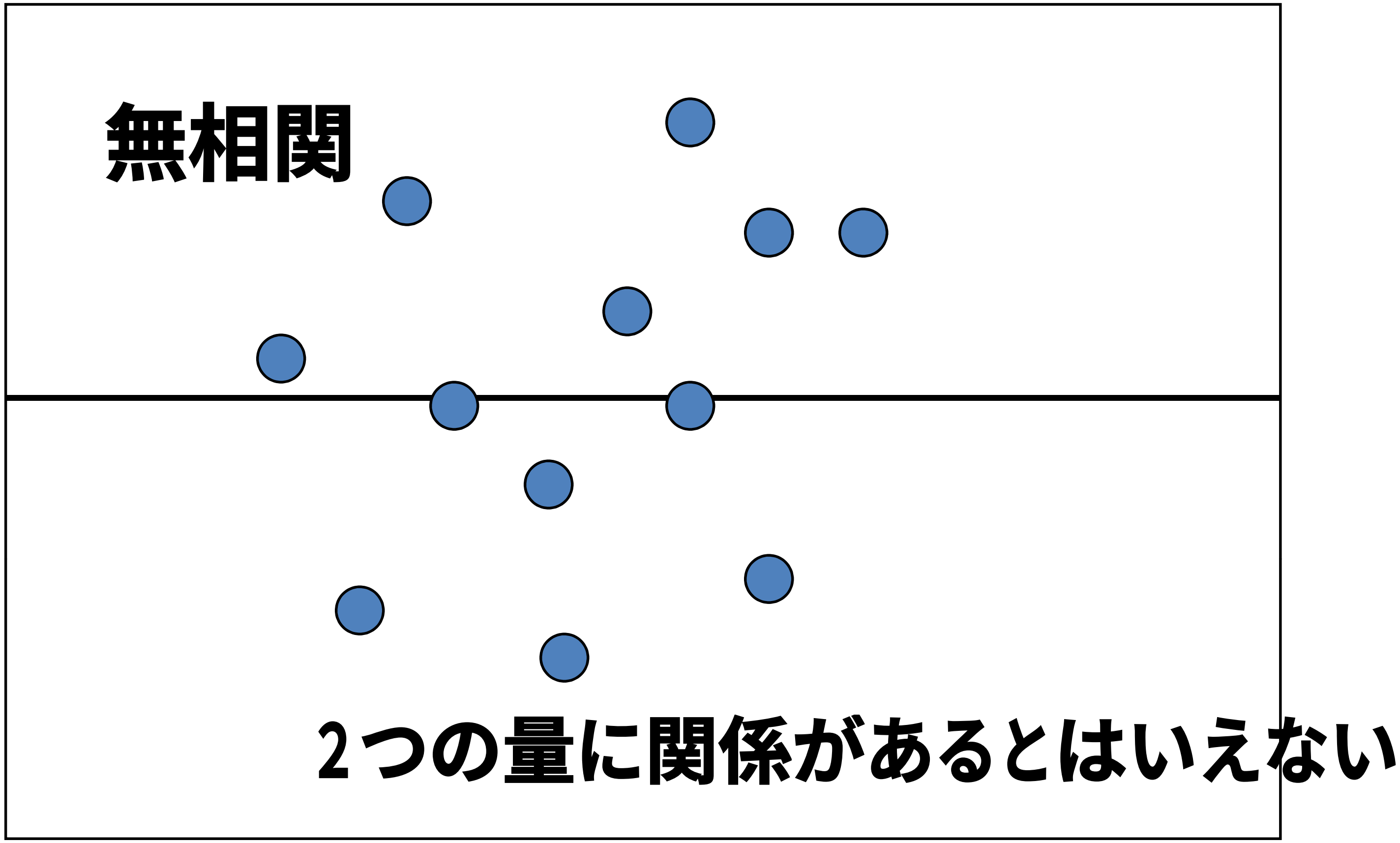
$$R = 0, R^2 = 0$$

相関係数 R　決定係数 R^2 値



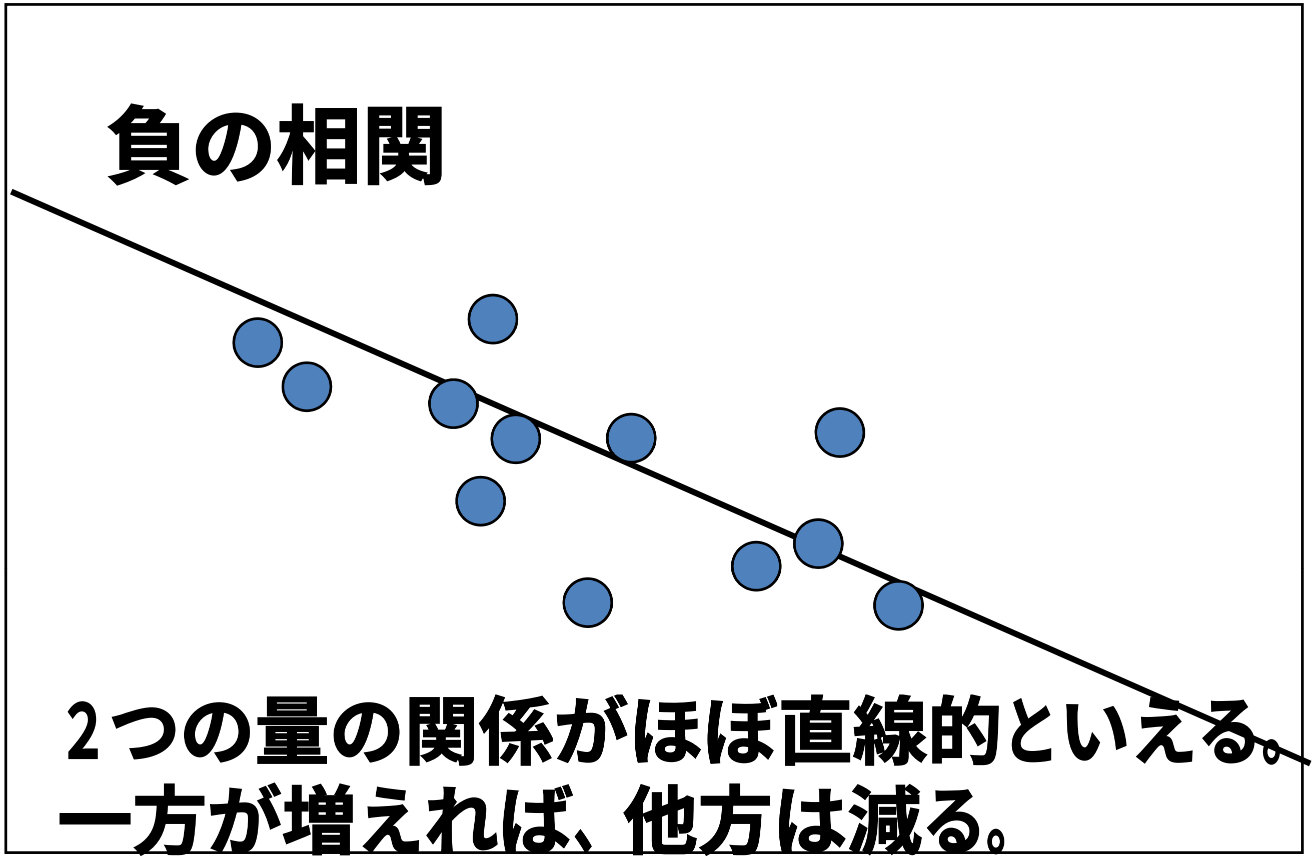
$R = 0, R^2 = 0$

相関係数R　決定係数R^2値



$R = 0, R^2 = 0$

相関係数 R　決定係数 R^2 値



$R = -0.8,\ R^2 = 0.64$ くらい

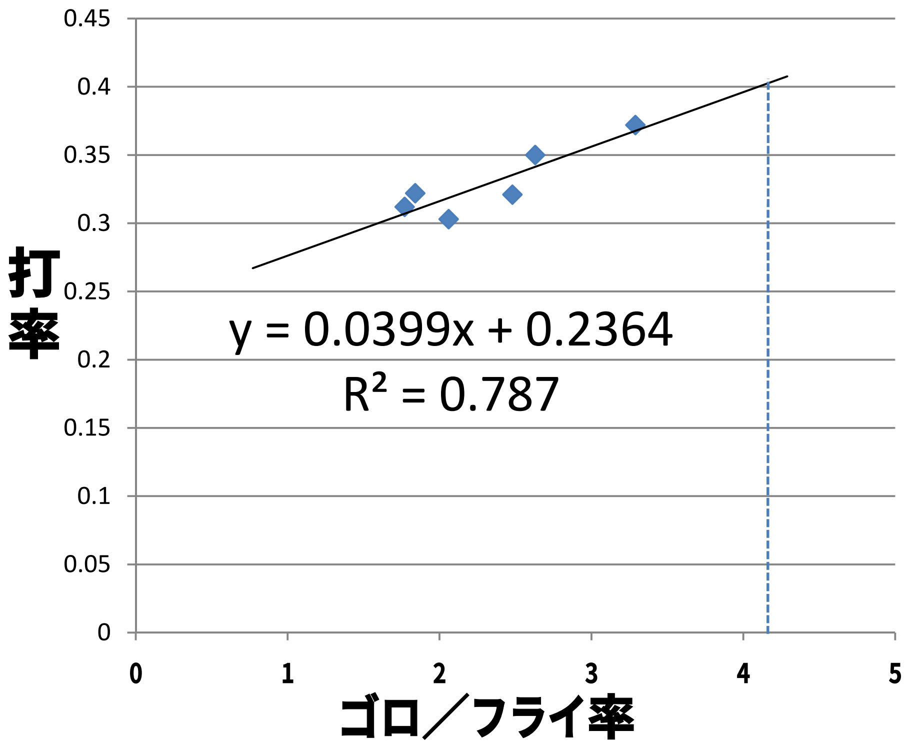
Excel デモ：イチロー選手の打撃傾向の分析

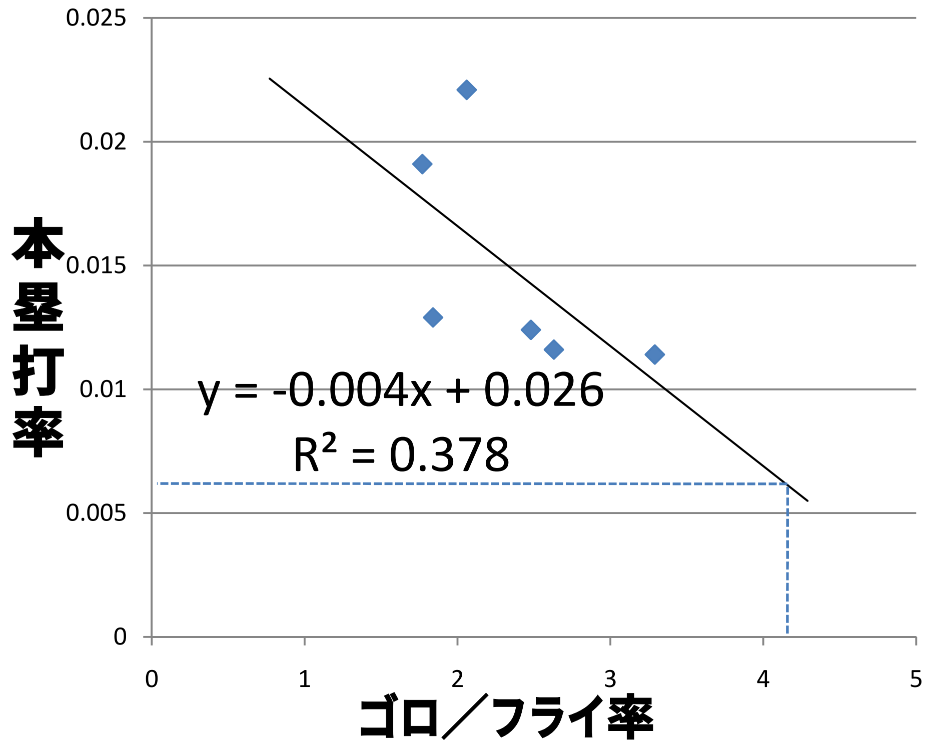
（ＥＳＰＮ ホームページ http://sports.espn.go.com/mlb/index より）

年	ゴロ／フライ率 （ゴロ数／フライ数）	打率 （ヒット数／打数）	本塁打率 （本塁打数／打数）
２００１	２．６３	０．３５０	０．０１１６（ ８／６９２）
２００２	２．４８	０．３２１	０．０１２４（ ８／６４７）
２００３	１．７７	０．３１２	０．０１９１（１３／６７９）
２００４	３．２９	０．３７２	０．０１１４（ ８／７０４）
２００５	２．０６	０．３０３	０．０２２１（１５／６７９）
２００６	１．８４	０．３２２	０．０１２９（ ９／６９５）

- ゴロ／フライ率を横軸、打率を縦軸にとり、直線当てはめを行う。（直線の式も表示する。）
- この直線当てはめの決定係数R^2を表示して、信頼度を確認する。
- ゴロ／フライ率を横軸、本塁打率を縦軸にとり、直線当てはめを行う。（直線の式も表示する。）
- 打率４割を達成する場合のゴロ／フライ率を予測する。
- そのときの本塁打数を予想する。（７００打数とする）

Excel デモ：イチロー選手の打撃傾向の分析
（ESPN　ホームページより）





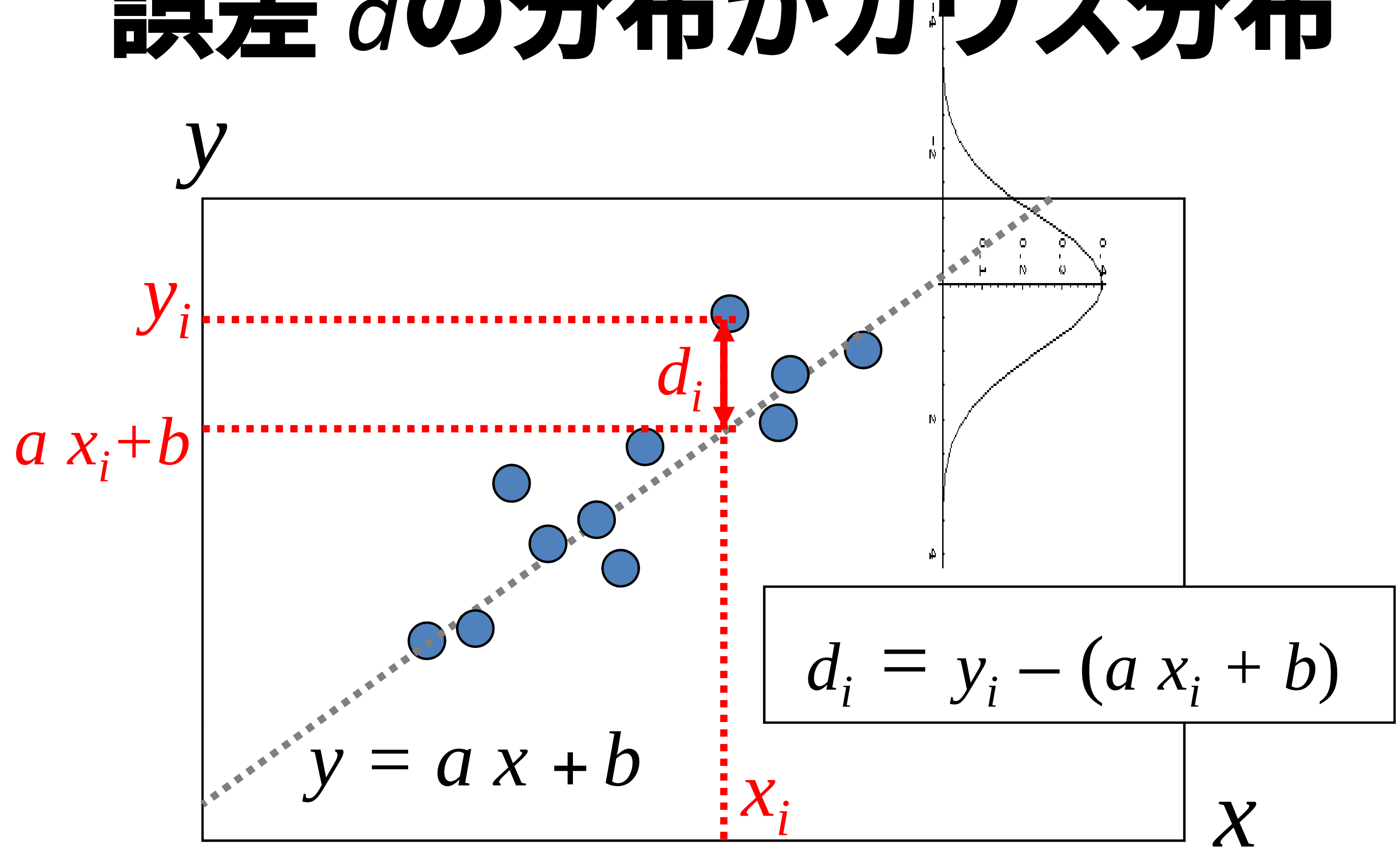
- ゴロ／フライ率を横軸、打率を縦軸にとり、直線当てはめを行う。（直線の式も表示する。）
- この直線当てはめの決定係数R^2を表示して、信頼度を確認する。
- ゴロ／フライ率を横軸、本塁打率を縦軸にとり、直線当てはめを行う。（直線の式も表示する。）
- 打率４割を達成する場合のゴロ／フライ率を予測する。
- そのときの本塁打数を予想する。（７００打数とする）

最尤推定法
(Maximum Likelihood Estimation: MLE)
としての最小二乗法

最尤推定法(Maximum Likelihood Estimation: MLE)としての最小二乗法

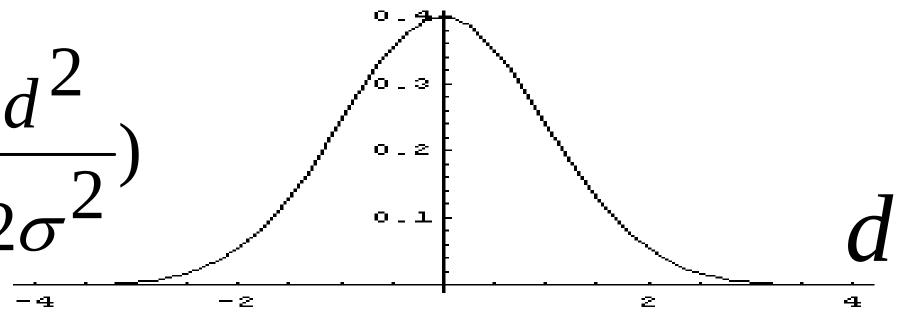
- なぜ最小“二乗”か？
  - $\sum_i \{y_i - (a\,x_i + b)\}^2$（2乗和）ではなく、たとえば、$\sum_i |\,y_i - (a\,x_i + b)|$（絶対値和）でもよいのではないか？
- 理由：
  1. <u>処理の利便性</u>：線形最小2乗の場合、処理が簡便である。（線形方程式に帰着できる。）
  2. <u>理論上の妥当性</u>：2乗和最小化による推定は、誤差($y_i - (a\,x_i + b)$)の分布がガウス分布（正規分布）であることを仮定した場合の最尤推定である。

誤差 d の分布がガウス分布

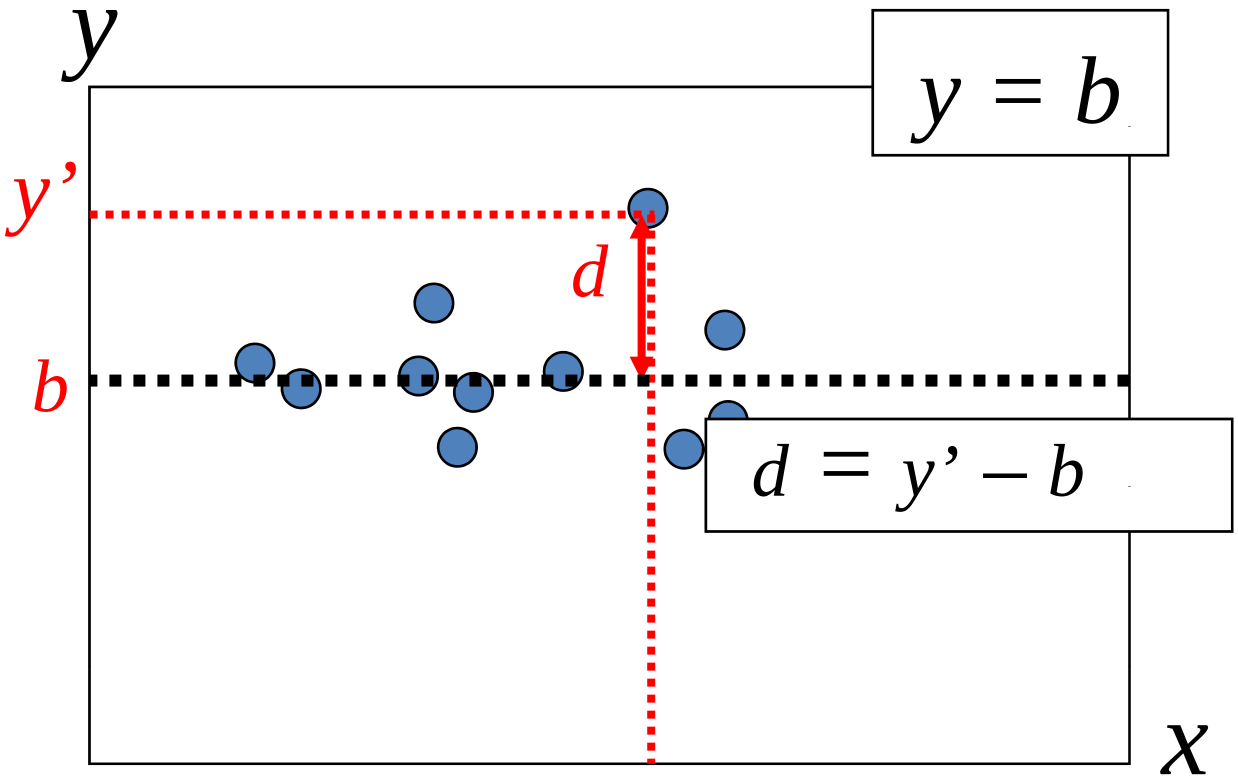


ガウス分布

$$Gauss(d;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{d^2}{2\sigma^2}\right)$$

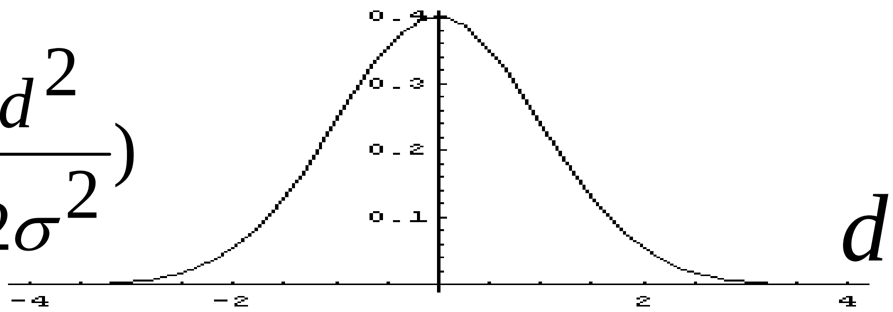


誤差 *d* がガウス分布 (簡単な場合)

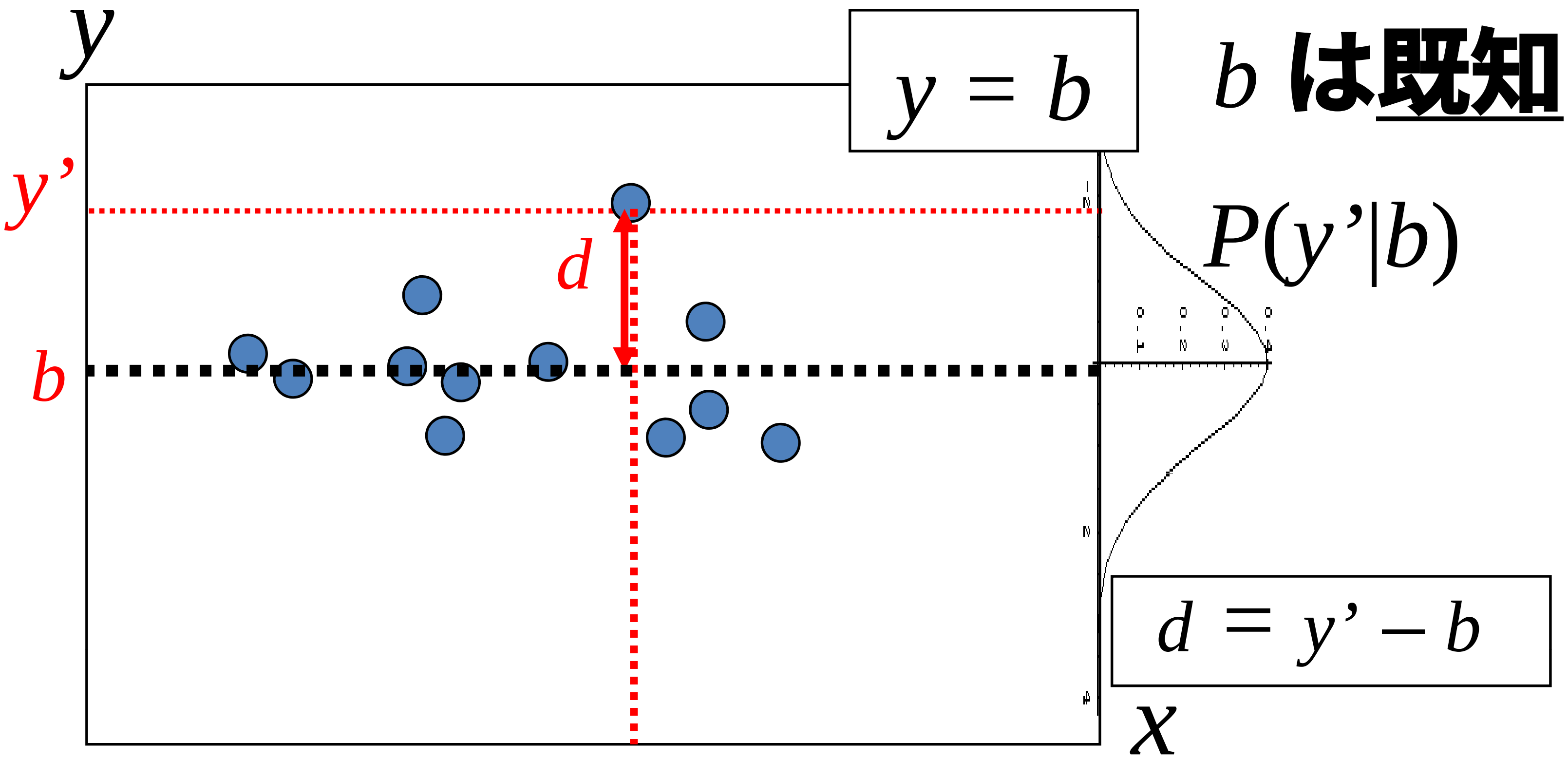


ガウス分布

$$Gauss(d;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{d^2}{2\sigma^2}\right)$$



yの計測値 (誤差 d) の確率分布

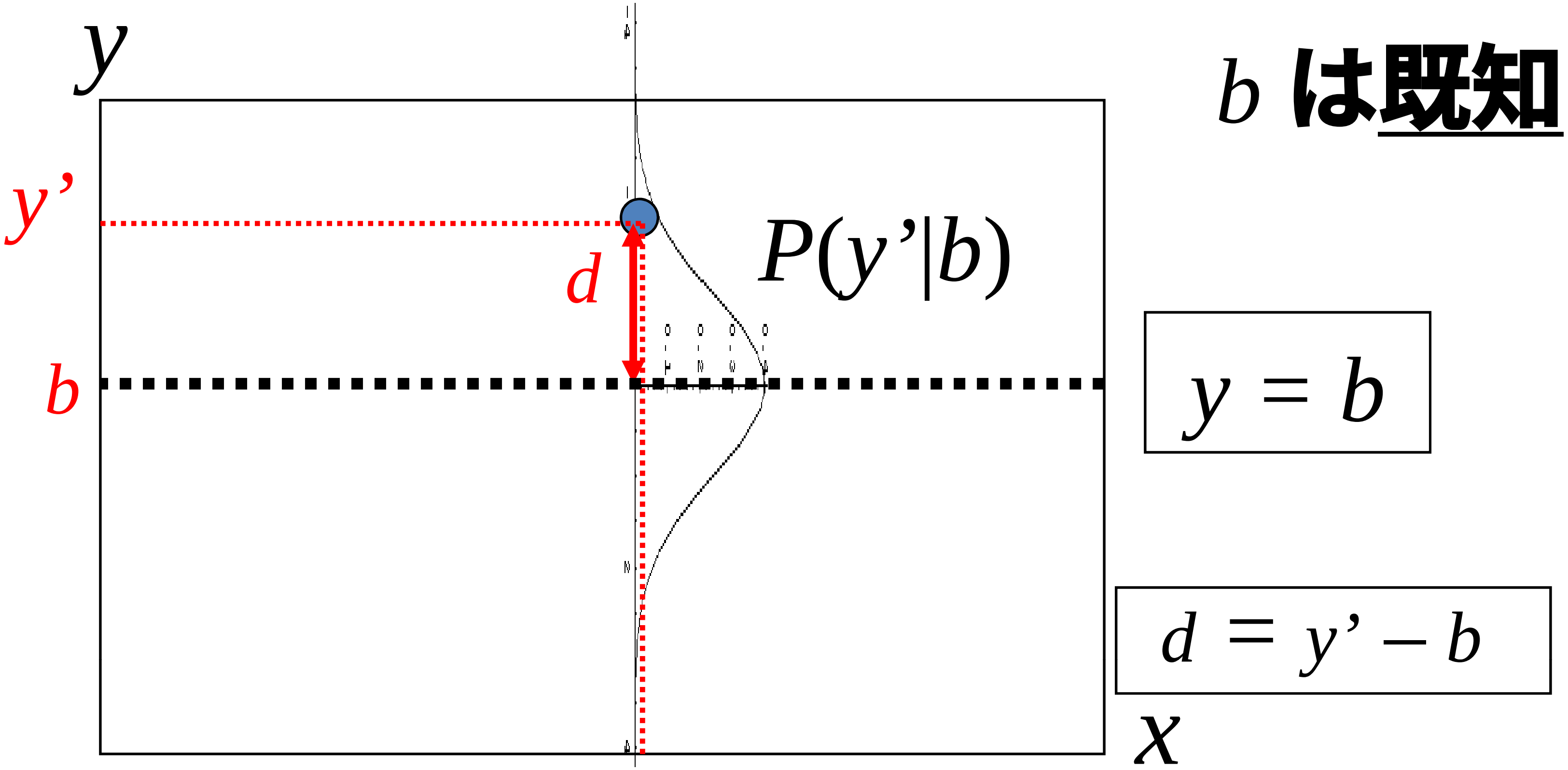


真値$y=b$**の場合の計測値** y' **の確率分布** $P(y'|b)$

$y = b$ **が確定している場合に、**y'**がとりうる値の確率分布。誤差は、** $d=y'-b$

yの計測値 (誤差 d) の確率分布

b は<u>既知</u>

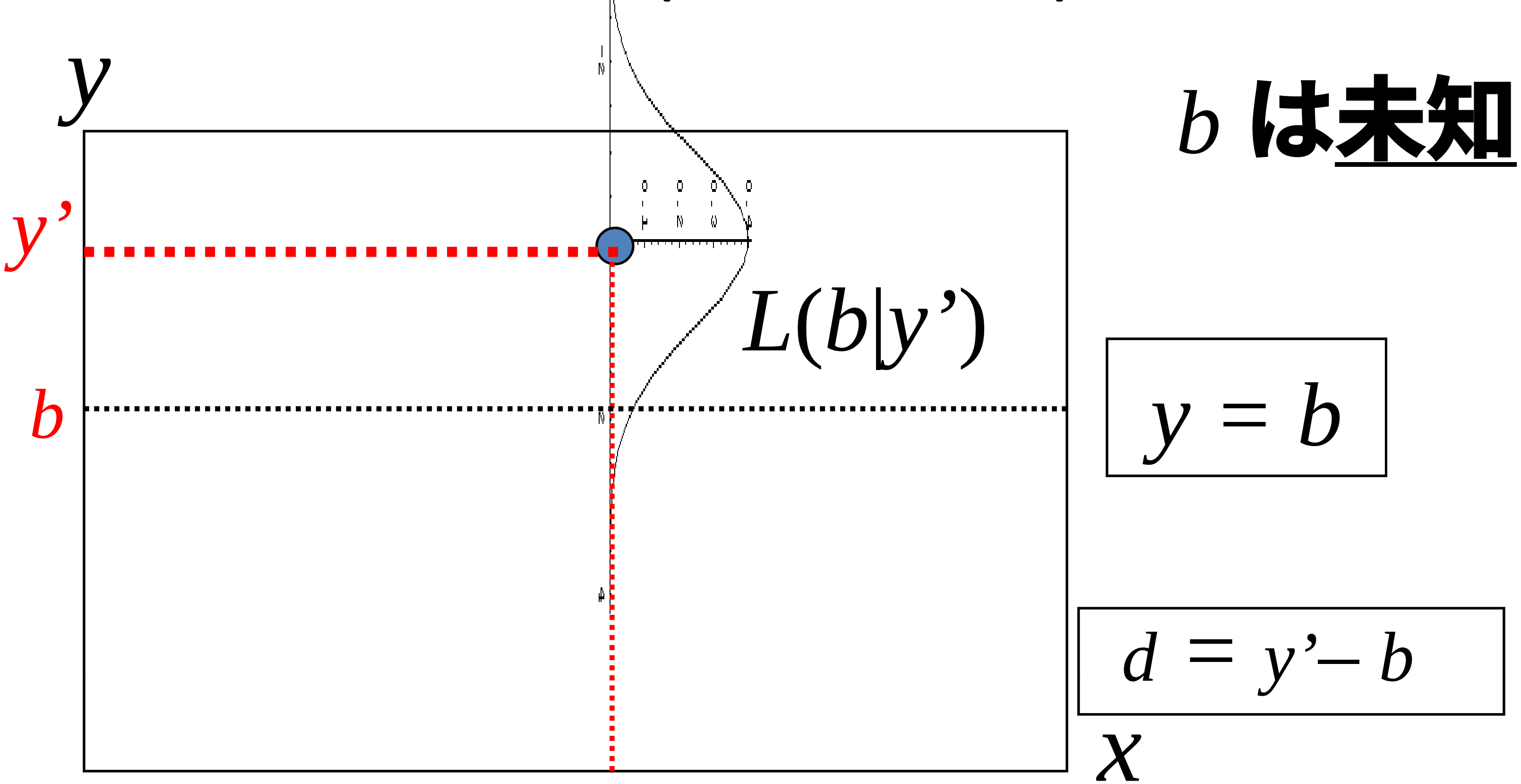


確率分布 $P(y'|b)$

真値 $y = b$ が確定している場合に、計測値y'がとりうる値の確率分布。誤差 $d=y'-b$ はガウス分布と仮定。

尤度(Likelihood)

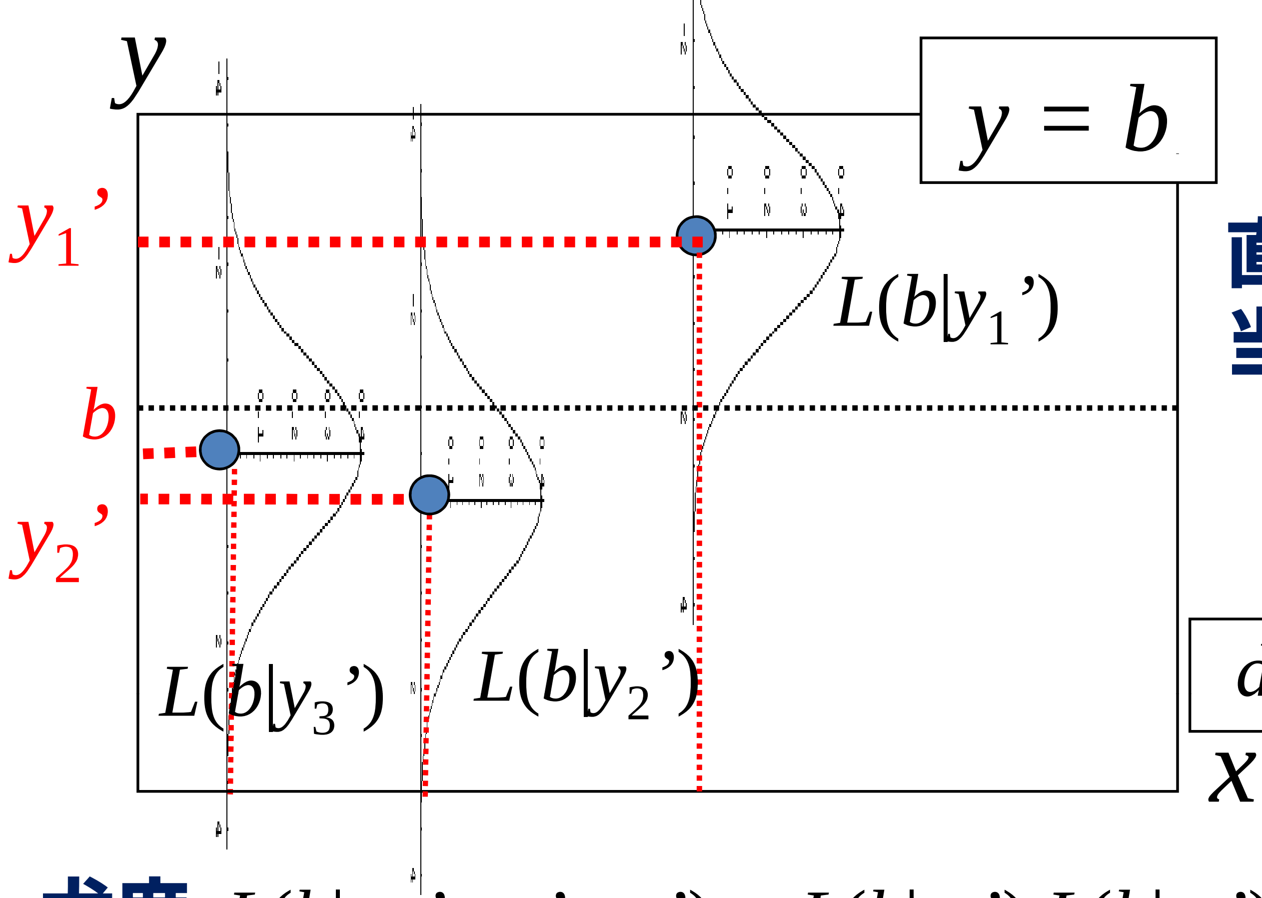
b は<u>未知</u>



尤度 $L(b|y')$　：　尤 (もっとも) らしさ

1つの y 座標の計測値 y' が与えられた場合に、真値が $y = b$ であろうと予測される確率 ($y=b$の尤らしさ)

尤度(Likelihood)



b は<u>未知</u>

直線 $y=b$ を当てはめる場合

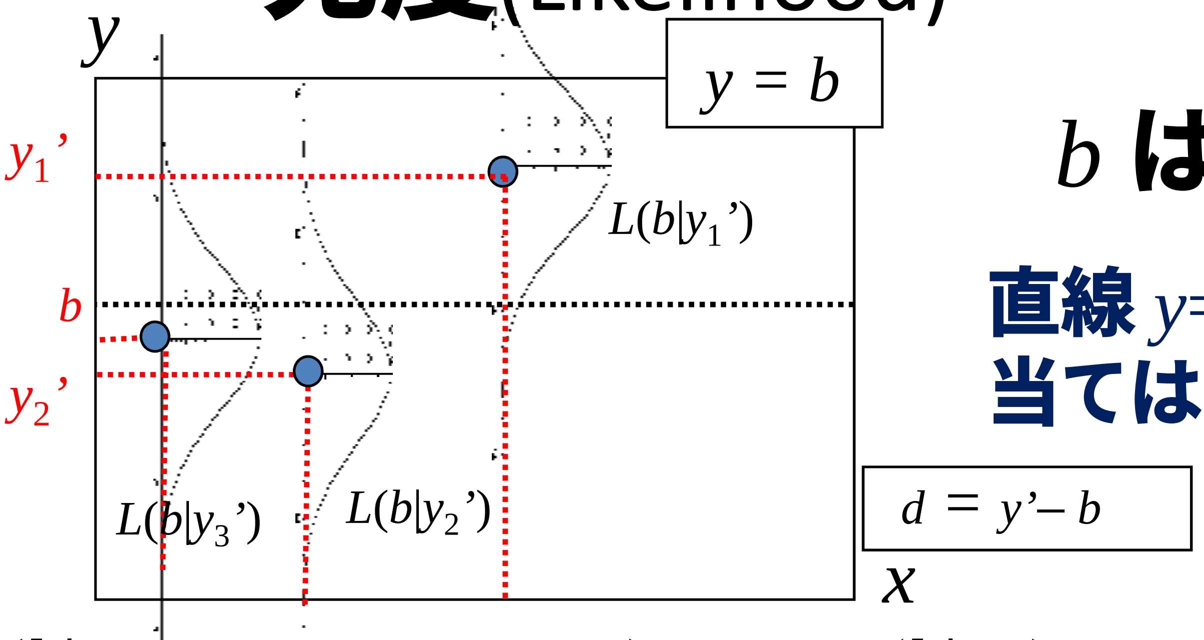
$d = y' - b$

尤度 $L(b|\, y_1', y_2', y_3') = L(b|y_1')\, L(b|y_2')\, L(b|y_3')$

$L(b|\, y_1', y_2', y_3', \ldots\ldots) = \Pi_i\, L(b|y_i')$

$= \Pi_i\, \mathrm{Gauss}(b - y_i')$

尤度(Likelihood)



b は未知

直線 $y=b$ を当てはめる場合

尤度 $L(b|\, y_1', y_2', y_3', \ldots\ldots) = \Pi_i\, L(b|y_i') = \Pi_i\, \mathrm{Gauss}(b-y_i')$

$$Gauss(d;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{d^2}{2\sigma^2}\right)$$ **より**

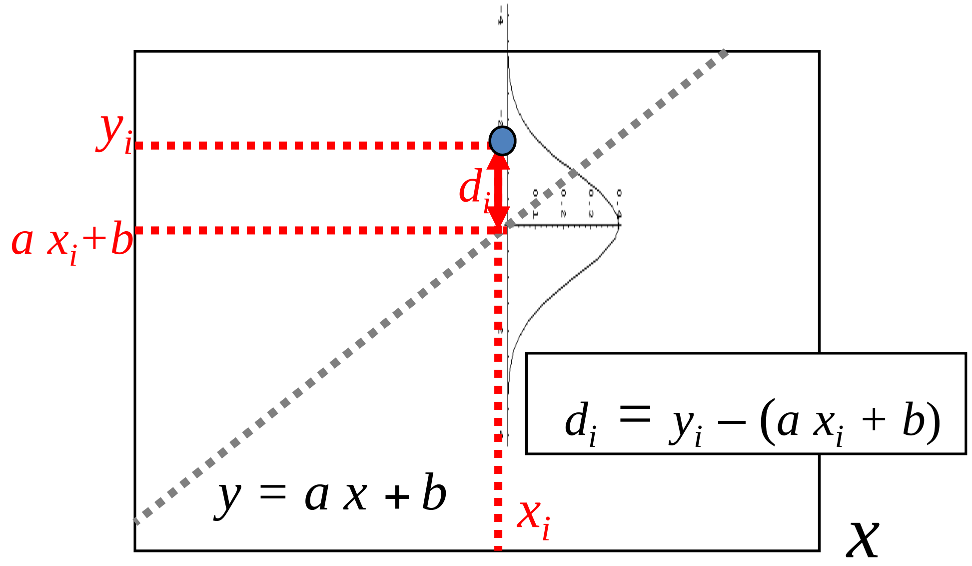
$$\log[\Pi_i\, \mathrm{Gauss}(b-y_i')] \propto \log[\Pi_i \exp[-(b-y_i')^2/(2\sigma^2)]]$$
$$= \log[\exp[-\Sigma_i (b-y_i')^2/(2\sigma^2)]]$$
$$= -\Sigma_i (b-y_i')^2/(2\sigma^2) \propto -\Sigma_i (b-y_i')^2$$

対数尤度最大 ＝ 二乗誤差最小

尤度(Likelihood)

直線 $y=ax+b$ を当てはめる場合

a, b **は未知**



$$Gauss(d;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{d^2}{2\sigma^2}\right)$$

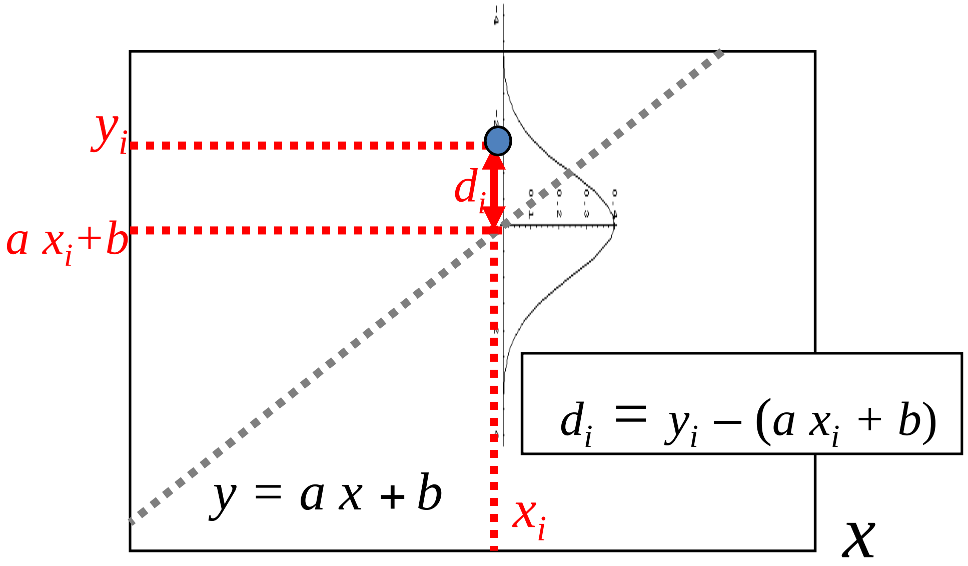
対数尤度

$$\log[\Pi_i \, \mathrm{Gauss}((ax_i+b) - y_i' \,;\sigma)]$$
$$\propto \log[\Pi_i \exp[-((ax_i+b) - y_i')^2/(2\sigma^2)]]$$
$$\propto \log[\exp[-\Sigma_i ((ax_i+b) - y_i')^2/(2\sigma^2)]]$$
$$= -\Sigma_i ((ax_i+b) - y_i')^2/(2\sigma^2) \propto -\Sigma_i ((ax_i+b) - y_i')^2$$

対数尤度最大 ＝ 二乗誤差最小

尤度(Likelihood)

各データ (x_i, y_i)の誤差分布標準偏差 σ_i が既知で、データ毎に異なる場合



$$Gauss(d;\sigma) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{d^2}{2\sigma^2}\right)$$

対数尤度

$$\log[\Pi_i\, \mathrm{Gauss}((ax_i+b) - y_i' \,;\sigma_i)]$$

$$\propto \log[\Pi_i \exp[-((ax_i+b) - y_i')^2 / (2\sigma_i^2)]]$$

$$\propto \log[\exp[-\Sigma_i ((ax_i+b) - y_i')^2 / (2\sigma_i^2)]]$$

$$= -\Sigma_i ((ax_i+b) - y_i')^2 / (2\sigma_i^2) \propto -\Sigma_i [(1/\sigma_i)((ax_i+b) - y_i')]^2$$

各データ毎に標準偏差逆数 $1/\sigma_i$ を掛け算した重み付き最小二乗和となる。

最小二乗法の統計学的意義づけ

- 最小二乗法は、誤差の分布が（平均ゼロの）ガウス関数に従うと仮定した場合の、最尤推定法と等価である。
  - よって、統計的推定の観点から確かな根拠があると言える。しかし、実際には、誤差の分布がガウス関数に従うかどうか検証がなされる場合はほとんどなく、数値計算上の利便性により、最小二乗法が用いられているのが実情である。
- 各計測データの誤差分布のガウス関数標準偏差σ_i（ばらつきの大きさ）が既知の場合、重み付き最小二乗法となる。重みは、$1/\sigma_i$ となる。

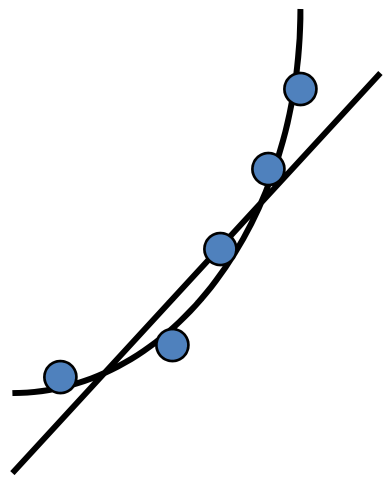
多項式当てはめ

直線当てはめから多項式当てはめへ

- 2つの量 x, y の関係がいつでも直線的になるとは限らない。
- 2つの量 x, y の関係を表現する数式として、様々な関係が考えられる。

$$y = ax + b$$

- **誤差** $d_i = y_i - (a\,x_i + b)$



$$y = ax^2 + bx + c$$

- **誤差** $d_i = y_i - (a\,x_i^2 + b\,x_i + c)$

最小二乗基準

- すべての点が、2次多項式 $y = ax^2 + bx + c$ の上になるべくのっている。

$$\sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \{y_i - (a\,x_i^2 + b\,x_i + c)\}^2$$ を最小にする a, b, c を求める。

n 個の点があり、i 番目の点の座標値を (x_i, y_i) とする。

最小二乗基準

$f(a,b,c) = \sum_{i=1}^{n}\{y_i - (a\,x_i^2 + b\,x_i + c)\}^2$ を最小にする a, b, c を求める。

$f(a,b,c)$ を最小にする a,b,c を求める。

↓

$$\frac{\partial}{\partial a}f(a,b,c) = 0 \quad \frac{\partial}{\partial b}f(a,b,c) = 0 \quad \frac{\partial}{\partial c}f(a,b,c) = 0$$

a, b, c に関する偏微分が 0 になる。

↓

線形連立方程式に帰着

線形最小二乗と
非線形最小二乗

最小二乗法の一般化

- 二乗和

推定すべきパラメータ： a_k

$$\sum_i \{y_i - (a_1 x_i + a_2)\}^2$$

$$\sum_i \{y_i - (a_1 x_i^2 + a_2 x_i + a_3)\}^2$$

- 一般化

$y = f(x;a)$ の形で記述できる (xi, yi) の測定値があると仮定。

$$\sum_i \{y_i - f(x_i;a)\}^2$$

直線当てはめの場合： $a = (a_1, a_2)$

２次多項式当てはめの場合： $a = (a_1, a_2, a_3)$

最小二乗法の一般化

- 二乗和

推定すべきパラメータ：a_k

$$\sum_i \{y_i - (a_1 x_i + a_2)\}^2$$

$$\sum_i \{y_i - (a_1 x_i^2 + a_2 x_i + a_3)\}^2$$

- 線形最小二乗法の一般化

$$\sum_i \{y_i - f(x_i;a)\}^2,$$ ただし $f(x;a) = \sum_k a_k f_k(x)$

直線当てはめの場合：　$f_1 = x$, $f_2 = 1$, $a = (a_1, a_2)$

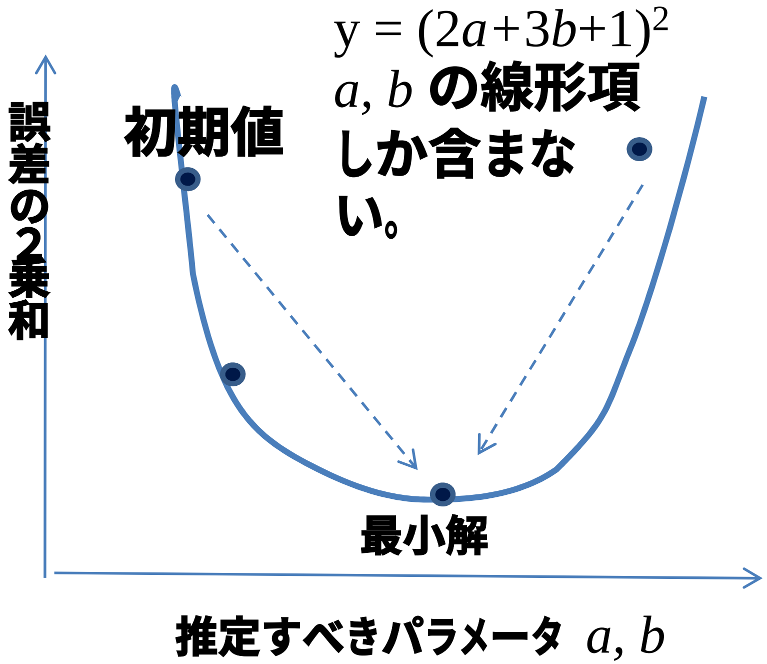
2次多項式当てはめの場合：　$f_1 = x^2$, $f_2 = x$, $f_3 = 1$, $a = (a_1, a_2, a_3)$

$f(x;\boldsymbol{a})$ に、推定すべきパラメータ a_k の1次項しか含まれない場合、すなわち、$f(x;\boldsymbol{a}) = \sum_k a_k f_k(x)$ とかける場合、線形最小二乗となり、推定すべき各パラメータで偏微分することにより得られる連立1次方程式を解くことにより、パラメータの推定が行える。そうでない場合、非線形となる。

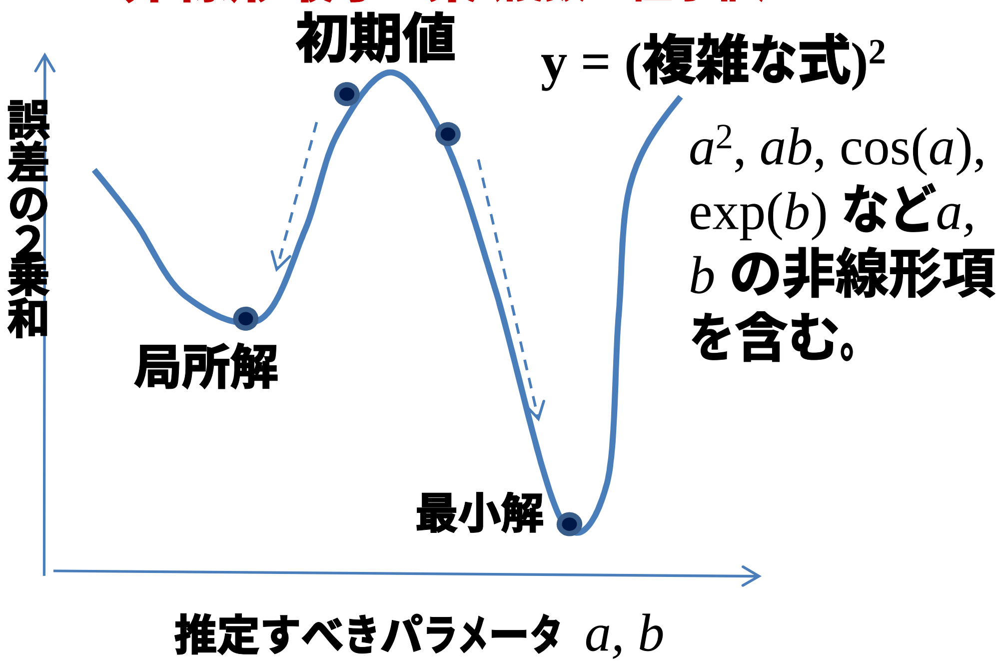
線形／非線形最小二乗

- 局所解 (極小解, Local Minimum) の問題
  - 局所解とは？
    - 誤差の2乗和を最小にする解を求める場合、線形最小二乗では、唯一の最小解を求めることができる。しかし、非線形最小二乗では、極小にする解しか求めることができない (極小値をとる解は一般に複数存在する) 。

線形最小2乗 (必ず1つの極小値)



非線形最小2乗 (複数の極小値)



最小二乗法のまとめ

- 最小二乗法は、誤差分布がガウス関数に従うときの最尤推定法である。
- 剛体位置合わせやカメラキャリブレーションも最小二乗問題として定式化される。
- “線形”最小二乗問題として定式化される場合は、連立一次方程式に帰着され、多くの場合、最小解を安定に求めることができる。線形計算に関する数値計算パッケージを利用して解を求める。
- “非線形”最小二乗問題として定式化される場合、最小解が求められる保証はないが、良い初期値を与えることにより実用的には良い解を求めることができる。Levenberg-Marquardt法は、代表的な非線形最小二乗法の数値計算アルゴリズムである。