

# 第 6 回

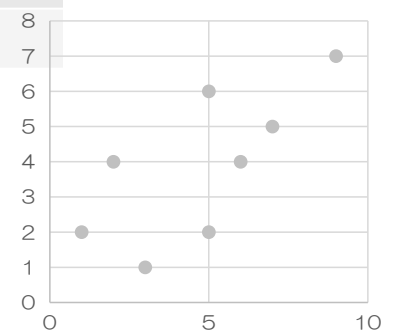
## 最小2乗法 Least Squares Method

### 最小2乗法によるモデルの当てはめ

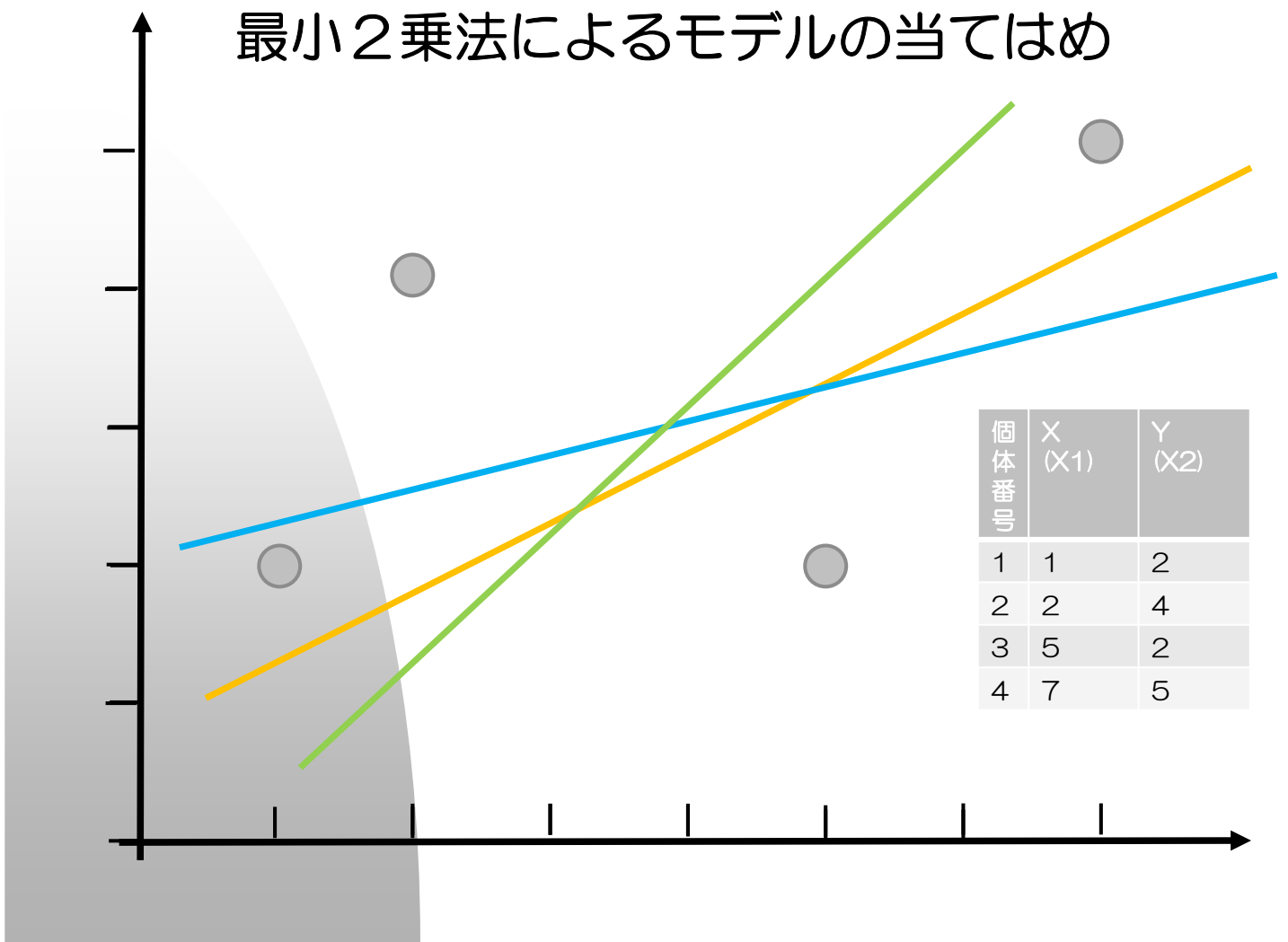
個体番号	X (X1)	Y (X2)
1	1	2
2	2	4
3	5	2
4	7	5
5	6	4
6	3	1
7	9	7
8	5	6

X と Y の関係を知りたい。

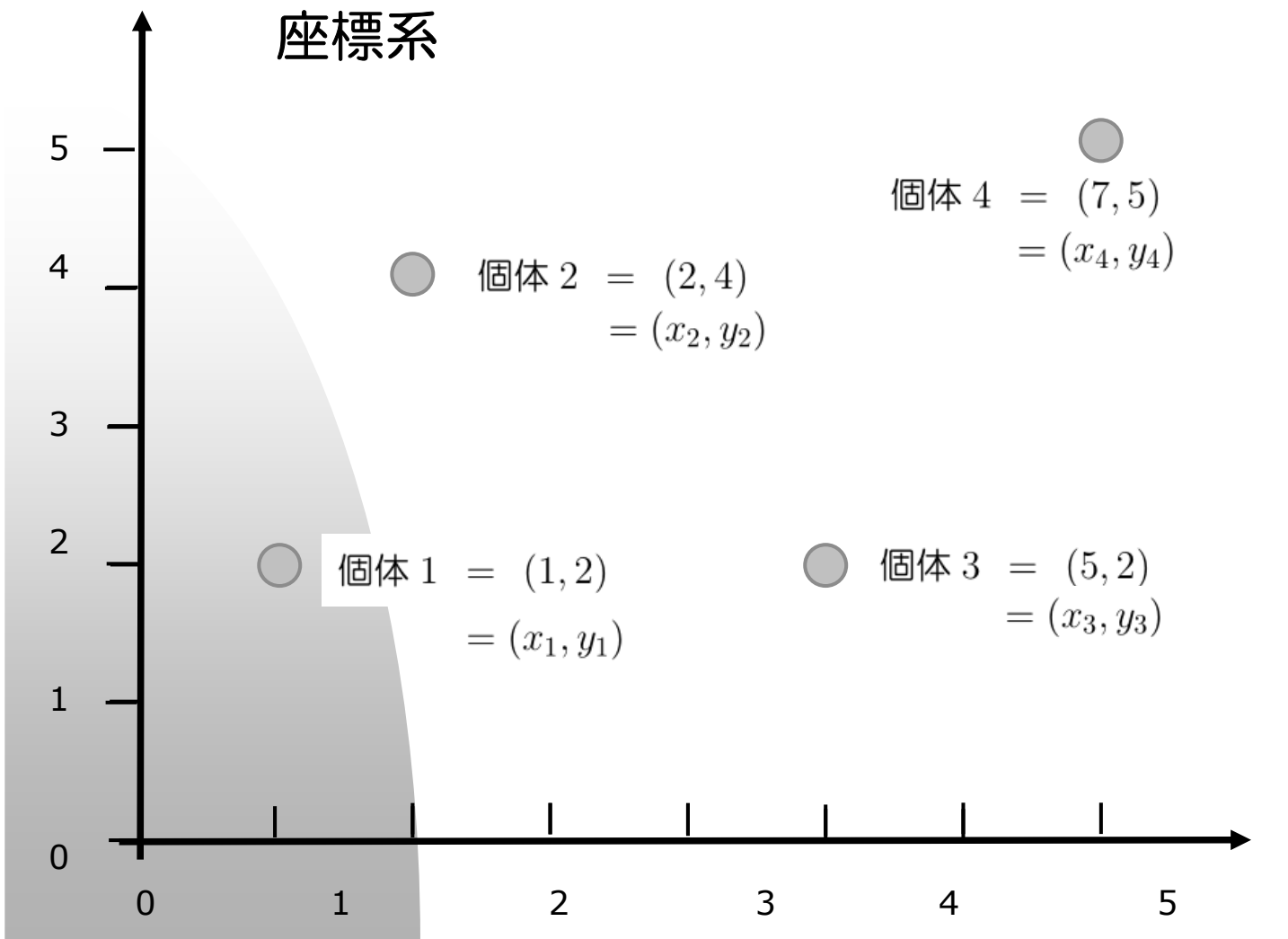
X で Y を予測したい。  
Y を X で説明したい。



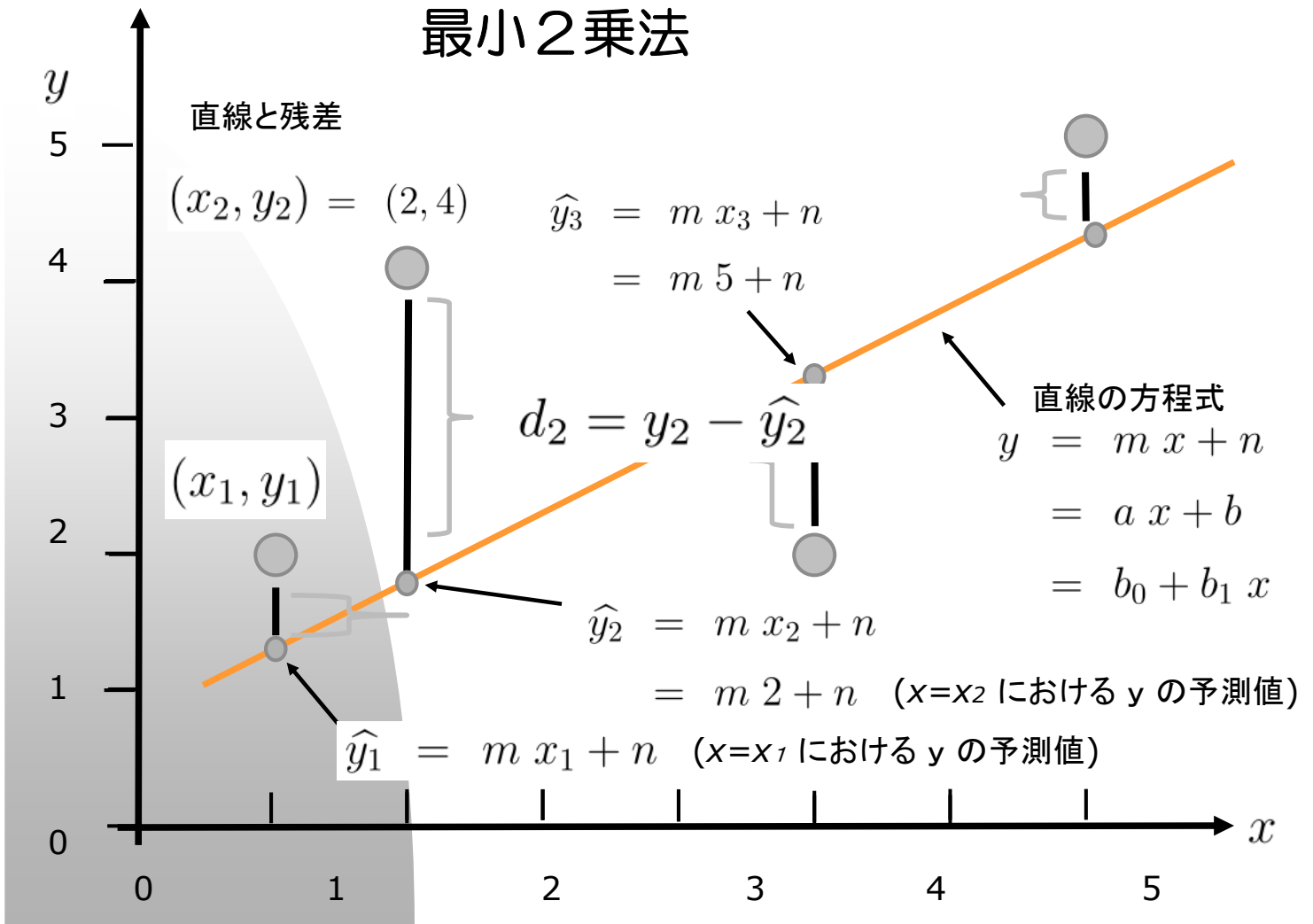
# 最小2乗法によるモデルの当てはめ



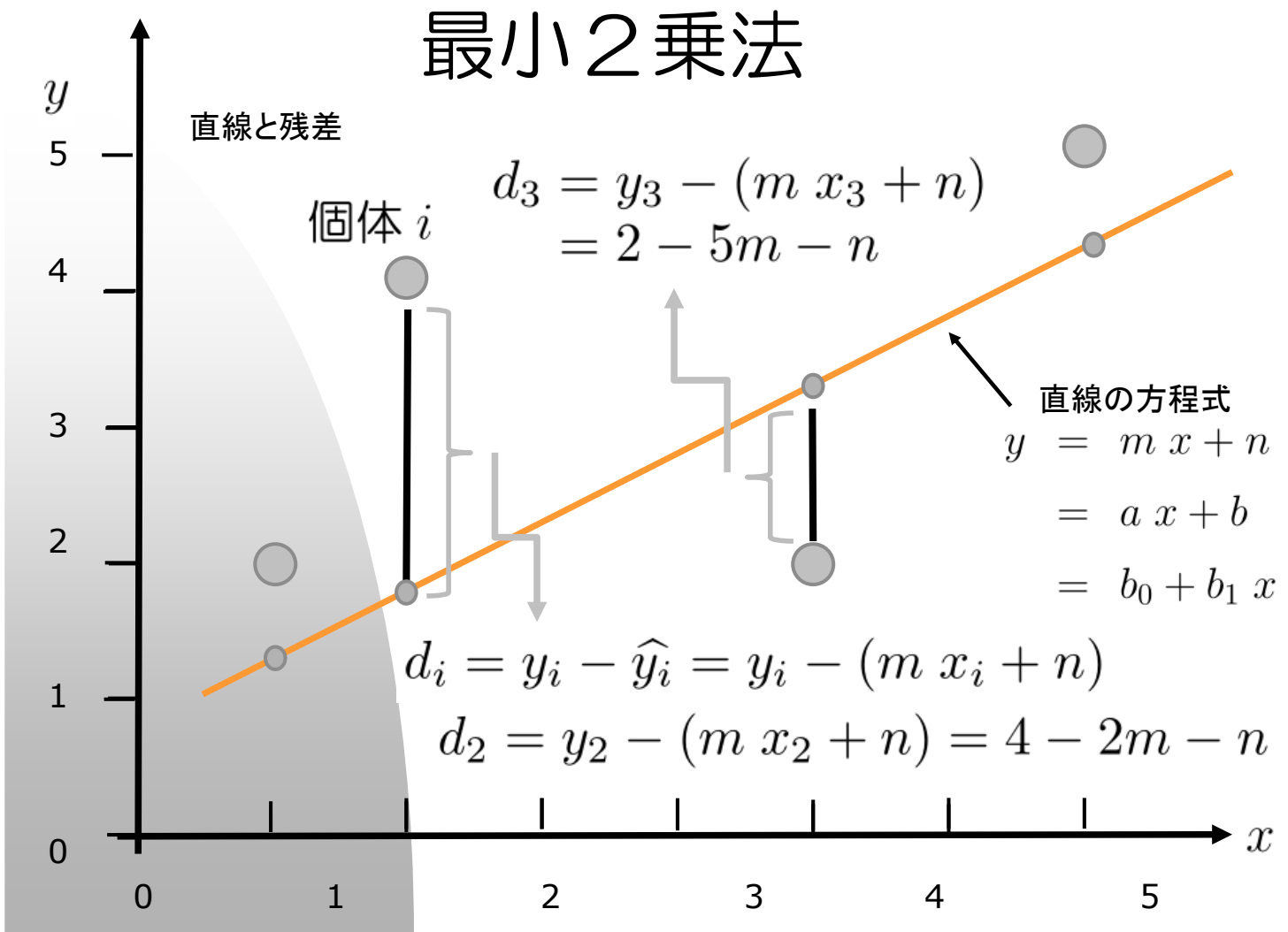
## 座標系



# 最小2乗法



# 最小2乗法



# 最小2乗法

残差の和

$$\begin{aligned}RS &= d_1 + d_2 + d_3 + d_4 \dots \\ &= (y_1 - \hat{y}_1) + (y_2 - \hat{y}_2) + (y_3 - \hat{y}_3) + (y_4 - \hat{y}_4) + \dots \\ &= (y_1 - (m x_1 + n)) + (y_2 - (m x_2 + n)) + \dots\end{aligned}$$

残差  $d_i = y_i - \hat{y}_i$  は正になったり負になったりする。

残差の2乗和

$$\begin{aligned}RSS &= d_1^2 + d_2^2 + d_3^2 + d_4^2 \dots \\ &= (y_1 - \hat{y}_1)^2 + (y_2 - \hat{y}_2)^2 + (y_3 - \hat{y}_3)^2 + (y_4 - \hat{y}_4)^2 + \dots \\ &= (y_1 - (m x_1 + n))^2 + (y_2 - (m x_2 + n))^2 + \dots \\ &= (2 - (1m + n))^2 + (4 - (2m + n))^2 + (2 - (5m + n))^2 + \dots\end{aligned}$$

7

## 最小2乗法

- $(m, n)$  の値を変えると RSS の値は変化する。
  - RSS は  $(m, n)$  の関数である。
  - その意味で  $RSS(m, n)$  と書ける。
  - $(m, n)$  は変数、もしくは引数と呼ばれる。
  - RSS は (スカラー値をとる) 2変数の関数である。
  - RSS は2乗和 (残差を2乗して足したもの) であるから非負の値である。
  - RSS の最小値を与える  $(m, n)$  の値を求める。
- 
- 関数の極値ではその微分が 0 となっている。
  - RSS の (偏) 微分が 0 となる  $(m, n)$  の値を求める。
  - RSS を  $(m, n)$  で (偏) 微分した (連立) 方程式を解く。

8

## 最小2乗法

$$\text{RSS} = (2 - (1m + n))^2 + (4 - (2m + n))^2 + (2 - (5m + n))^2 + (5 - (7m + n))^2$$

$$\begin{aligned}\frac{\partial \text{RSS}}{\partial m} &= \frac{\partial (2 - (1m + n))^2}{\partial m} + \frac{\partial (4 - (2m + n))^2}{\partial m} \\ &\quad + \frac{\partial (2 - (5m + n))^2}{\partial m} + \frac{\partial (5 - (7m + n))^2}{\partial m} \\ &= -2 \times (2 - (1m + n)) \times 1 - 2 \times (4 - (2m + n)) \times 2 \\ &\quad - 2 \times (2 - (5m + n)) \times 5 - 2 \times (5 - (7m + n)) \times 7 \\ &= -2 \times (79m + 15n - 55) = 0\end{aligned}$$

$$\begin{aligned}\frac{\partial \text{RSS}}{\partial n} &= \frac{\partial (2 - (1m + n))^2}{\partial n} + \frac{\partial (4 - (2m + n))^2}{\partial n} + \frac{\partial (2 - (5m + n))^2}{\partial n} + \frac{\partial (5 - (7m + n))^2}{\partial n} \\ &= -2 \times (2 - (1m + n)) - 2 \times (4 - (2m + n)) - 2 \times (2 - (5m + n)) - 2 \times (5 - (7m + n)) \\ &= -2 \times (15m + 4n - 13) = 0\end{aligned}$$

9

## 最小2乗解

正規方程式

$$\frac{\partial \text{RSS}}{\partial m} = 0$$

$$\frac{\partial \text{RSS}}{\partial n} = 0$$

$$y = mx + n$$

$$= ax + b$$

$$= b_0 + b_1 x$$

$$79m + 15n = 55$$

$$15m + 4n = 13$$

$$m = 0.27, \quad n = 2.22$$

$$\hat{y} = 0.27x + 2.22$$

$$y = \hat{y} + e$$

10

## 最小2乗解

$$\text{RSS} = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N (y_i - (m x_i + n))^2$$

$$\frac{\partial \text{RSS}}{\partial m} = \sum_{i=1}^N 2 \times (y_i - \hat{y}_i) \frac{\partial (-y_i)}{\partial m}$$

$$= \sum_{i=1}^N 2 \times (y_i - \hat{y}_i) \times (-x_i)$$

$$= -2 \times \sum_{i=1}^N (y_i - (m x_i + n)) \times x_i$$

$$= -2 \times \left[ \sum_{i=1}^N y_i x_i - m \sum_{i=1}^N x_i^2 - n \sum_{i=1}^N x_i \right]$$

$$\left( \sum_{i=1}^N x_i^2 \right) m + \left( \sum_{i=1}^N x_i \right) n = \sum_{i=1}^N y_i x_i$$

11

## 最小2乗解

$$\text{RSS} = \sum_{i=1}^N (y_i - \hat{y}_i)^2 = \sum_{i=1}^N (y_i - (m x_i + n))^2$$

$$\frac{\partial \text{RSS}}{\partial n} = \sum_{i=1}^N 2 \times (y_i - \hat{y}_i) \frac{\partial (-y_i)}{\partial n}$$

$$= \sum_{i=1}^N 2 \times (y_i - \hat{y}_i) \times (-1)$$

$$= -2 \times \sum_{i=1}^N (y_i - (m x_i + n))$$

$$= -2 \times \left[ \sum_{i=1}^N y_i - m \sum_{i=1}^N x_i - n N \right]$$

$$\left( \sum_{i=1}^N x_i \right) m + N n = \sum_{i=1}^N y_i$$

12

# 正規方程式

$$\begin{aligned}\hat{y} &= m x + n \\ &= a x + b \\ &= b_0 + b_1 x\end{aligned}$$

$$\left(\sum_{i=1}^N x_i^2\right) m + \left(\sum_{i=1}^N x_i\right) n = \sum_{i=1}^N y_i x_i$$

$$\left(\sum_{i=1}^N x_i\right) m + N n = \sum_{i=1}^N y_i$$

個体番号	X (X1)	Y (X2)
1	1	2
2	2	4
3	5	2
4	7	5

$$79m + 15n = 55$$

$$15m + 4n = 13$$

13

# 最小2乗法

- データ

$$(y_i, \mathbf{x}_i), i = 1, 2, \dots, N,$$

- モデル

$$y_i = f(x_i | \beta) + e_i, \quad i = 1, 2, \dots, N$$

$f(x)$  は  $x$  を与えたときにスカラー値を出力する関数

モデルの予測値  $\hat{y}_i$

$\beta$  はそのパラメタ

$e_i$  はモデルとデータのズレ

14

# モデル

$$y_i = \hat{y}_i + e_i$$

$$\text{where } \hat{y}_i = f(\mathbf{x}_i), \quad i = 1, 2, \dots, N,$$

単回帰モデル

$$\hat{y}_i = f(\mathbf{x}_i) = \beta_0 + \beta_1 x_i, \quad i = 1, 2, \dots, N,$$

重回帰モデル

$$\hat{y}_i = f(\mathbf{x}_i) = \beta_0 + \sum_{j=1}^q x_{ij} \beta_j, \quad i = 1, 2, \dots, N,$$

非線形回帰モデル

$$\hat{y}_i = \beta_0 + \log(x_i - \beta_1), \quad \hat{y}_i = \beta_0 + \frac{1}{1 + \exp(\beta_1 x_i - \beta_2)}$$

15

## モデルの当てはめ

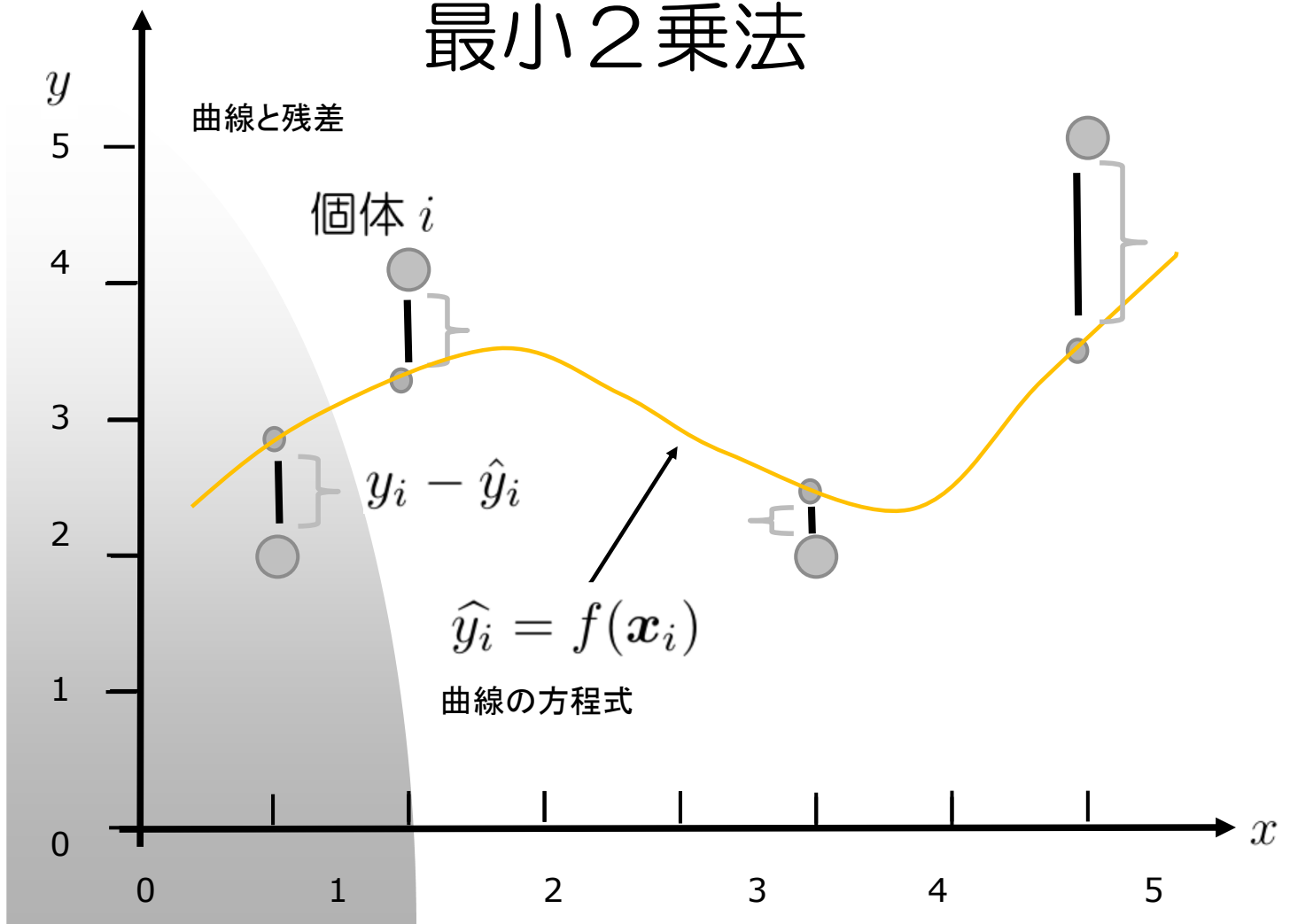
残差の2乗和が最小となるようにモデルのパラメタ  $\beta$  を求める。

$$\text{RSS} = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

16



# 最小2乗法



## 重回帰分析

$q$  個の説明変数を用いた重回帰モデルを

$$y_i = \hat{y}_i + e_i, \quad i = 1, 2, \dots, N$$

$$\text{where } \hat{y}_i = \beta_0 + \sum_{j=1}^q x_{ij}\beta_j$$

とする。誤差項  $e_i$  に関して分布を仮定する場合は、通常、正規分布

$$e_i \sim N(0, \sigma^2), \quad i = 1, 2, \dots, N, \quad iid$$

を仮定する場合が多い。  $\mathcal{E}(e_i) = 0$

このモデルのパラメタは  $q+1 \times 1$  のベクトル  $\beta$  であるが、誤差項に正規分布を仮定した場合には誤差分散  $\sigma^2$  もパラメタとして取り扱う。

切片の取り扱い

$$x_{i0} = 1$$

$$y_i = \sum_{j=0}^q x_{ij}\beta_j + \varepsilon_{ij}$$

## 重回帰分析の行列表記

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ &= \hat{\mathbf{y}} + \boldsymbol{\varepsilon} \end{aligned}$$

$$\mathbf{y} \quad n \times 1$$

基準変数

$$\mathbf{X} \quad n \times (q + 1)$$

説明変数 + 1

$$\boldsymbol{\beta} \quad (q + 1) \times 1$$

回帰係数 (切片を含む)

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta} \quad n \times 1$$

モデルの予測値

$$\begin{pmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1q} \\ 1 & x_{21} & x_{22} & \cdots & x_{2q} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nq} \end{bmatrix} \times \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \vdots \\ \beta_q \end{pmatrix}$$

## RSS の行列表記

$$\begin{aligned} \text{RSS} &= \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= \|\mathbf{y} - \hat{\mathbf{y}}\|^2 \\ &= (\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}}) \\ &= \text{tr}((\mathbf{y} - \hat{\mathbf{y}})'(\mathbf{y} - \hat{\mathbf{y}})) = \text{tr}((\mathbf{y} - \hat{\mathbf{y}})(\mathbf{y} - \hat{\mathbf{y}})') \end{aligned}$$

$$\hat{\mathbf{y}} = \mathbf{X}\boldsymbol{\beta}$$

$$\begin{aligned} &= \text{tr}((\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})) \\ &= \text{tr}(\mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \\ &= \text{tr}(\mathbf{y}'\mathbf{y}) - 2\text{tr}(\mathbf{y}'\mathbf{X}\boldsymbol{\beta}) + \text{tr}(\boldsymbol{\beta}'\mathbf{X}'\mathbf{X}\boldsymbol{\beta}) \end{aligned}$$

# 正規方程式

$$\begin{aligned} \frac{\partial \text{RSS}}{\partial \beta_k} &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \times \frac{\partial (y_i - \hat{y}_i)}{\partial \beta_k} \\ &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \times \frac{\partial \sum_{j=0}^q x_{ij} \beta_j}{\partial \beta_k} \\ &= 2 \sum_{i=1}^n (y_i - \hat{y}_i) \times x_{ik} \\ &= 2 \sum_{i=1}^n (y_i - \sum_{j=0}^q x_{ij} \beta_j) \times x_{ik} \\ &= 2 \sum_{i=1}^n y_i x_{ik} - 2 \sum_{i=1}^n (\sum_{j=0}^q x_{ij} \beta_j) x_{ik} \\ &= 2 \sum_{i=1}^n y_i x_{ik} - 2 \sum_{i=1}^n (\sum_{j=0}^q x_{ij} x_{ik} \beta_j) \\ &= 2 \sum_{i=1}^n y_i x_{ik} - 2 \sum_{j=0}^q (\sum_{i=1}^n x_{ij} x_{ik}) \beta_j \end{aligned}$$

$\frac{\partial \text{RSS}}{\partial \beta_k}$  を要素として持つベクトル

$$\frac{\partial \text{RSS}}{\partial \beta} = \mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\beta$$

ベクトルで書いた正規方程式

$$\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\beta = \mathbf{0}$$

正規方程式の解は

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y}$$

$\sum_{i=1}^n y_i x_{ik}$  は  $\mathbf{X}'\mathbf{y}$  の第  $k$  要素

$\sum_{i=1}^n x_{ij} x_{ik}$  は  $\mathbf{X}'\mathbf{X}$  の第  $jk$  要素

# RSS の行列表記

$$\begin{aligned} \text{RSS} &= \text{tr}((\mathbf{y} - \mathbf{X}\beta)'(\mathbf{y} - \mathbf{X}\beta)) \\ &= \text{tr}(\mathbf{y}'\mathbf{y} - 2\mathbf{y}'\mathbf{X}\beta + \beta'\mathbf{X}'\mathbf{X}\beta) \\ &= \text{tr}(\mathbf{y}'\mathbf{y}) - 2\text{tr}(\mathbf{y}'\mathbf{X}\beta) + \text{tr}(\beta'\mathbf{X}'\mathbf{X}\beta) \end{aligned}$$

$$\frac{\partial \text{RSS}}{\partial \beta} = 2 \frac{\partial \text{tr}(\mathbf{y}'\mathbf{X}\beta)}{\partial \beta} + \frac{\partial \text{tr}(\beta'\mathbf{X}'\mathbf{X}\beta)}{\partial \beta}$$

$$\frac{\partial \text{tr}(\mathbf{A}\mathbf{X})}{\partial \mathbf{X}} = \mathbf{A}', \quad \mathbf{X} \text{ が非対称の場合}$$

$$\mathbf{X} \rightarrow \beta \quad \text{and} \quad \mathbf{A} \rightarrow \mathbf{y}'\mathbf{X},$$

$$\frac{\partial \text{tr}(\mathbf{X}'\mathbf{A}\mathbf{X}\mathbf{B})}{\partial \mathbf{X}} = \mathbf{A}\mathbf{X}\mathbf{B} + \mathbf{A}'\mathbf{X}\mathbf{B}', \quad \mathbf{X} \text{ が非対称の場合}$$

$$\beta \rightarrow \mathbf{X} \quad \text{and} \quad \mathbf{X}'\mathbf{X} \rightarrow \mathbf{A} \quad \text{and} \quad \mathbf{I} \rightarrow \mathbf{B}$$

$$\begin{aligned} &= -2(\mathbf{y}'\mathbf{X})' + \mathbf{X}'\mathbf{X}\beta + (\mathbf{X}'\mathbf{X})'\beta \\ &= -2\mathbf{X}'\mathbf{y} + 2\mathbf{X}'\mathbf{X}\beta = -2(\mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\beta) \end{aligned}$$

# スカラー値を返す関数の行列による微分

$$\begin{aligned}\frac{\partial \operatorname{tr}(AX)}{\partial X} &= A', \quad X \text{ が非対称の場合} \\ &= A + A' - \operatorname{diag}(A), \quad X \text{ が対称の場合}\end{aligned}$$

$$\begin{aligned}\frac{\partial \operatorname{tr}(X'AXB)}{\partial X} &= AXB + A'XB', \quad X \text{ が非対称の場合} \\ &= AXB + BXA + A'XB' + B'XA' - \operatorname{diag}(AXB) - \operatorname{diag}(BXA), \\ &\quad X \text{ が対称の場合}\end{aligned}$$

$$\begin{aligned}\frac{\partial \operatorname{tr}(XAXB)}{\partial X} &= B'X'A' + A'X'B', \quad X \text{ が非対称の場合} \\ &= B'X'A' + A'X'B' + BXA + AXB - \operatorname{diag}(BXA) - \operatorname{diag}(AXB), \\ &\quad X \text{ が対称の場合}\end{aligned}$$

$$\begin{aligned}\frac{\partial \operatorname{tr}(X^{-1}A)}{\partial X} &= -(X^{-1}AX^{-1})', \quad X \text{ が非対称の場合} \\ &= -(X^{-1}AX^{-1})' - X^{-1}AX^{-1} + \operatorname{diag}(X^{-1}A'X^{-1}), \quad X \text{ が対称の場合}\end{aligned}$$

$$\begin{aligned}\frac{\partial |X|^a}{\partial X} &= a|X|^a(X^{-1})', \quad X \text{ が非対称の場合} \\ &= a|X|^a(2X^{-1} - \operatorname{diag}(X^{-1})), \quad X \text{ が対称の場合}\end{aligned}$$

23

## 最小2乗法の別解

- 残差の2乗和の分解公式を使った解法

$$\operatorname{tr}[(Y - X\beta)'](Y - X\beta) = \operatorname{tr}((\beta - \hat{\beta})X'X(\beta - \hat{\beta})) + \operatorname{tr}[(Y - X\hat{\beta})'](Y - X\hat{\beta})$$

上式右辺を最小とする  $\beta$  を求めたい。  
上式右辺の第2項は  $\beta$  を含まない。  
上式右辺の第1項を最小とする  $\beta$  を求めれば良い。  
 $X'X$  は正定値行列であるからこの項は負にはならない。  
したがって、最小値は  $\beta = \hat{\beta}$  のときである。

24

## 平方完成関係の公式

$$\|Y - XB\|^2 = \text{tr}((B - \hat{B})'X'X(B - \hat{B})) - \text{tr}(\hat{B}'X'X\hat{B}) + \text{tr}(Y'Y)$$

$$\hat{\beta} = (X'X)^{-1}X'Y$$

$$(y - X\beta)'Q^{-1}(y - X\beta) = (y - X\hat{\beta})'Q^{-1}(y - X\hat{\beta}) + (\beta - \hat{\beta})'X'Q^{-1}X(\beta - \hat{\beta})$$

$$\hat{\beta} = (X'Q^{-1}X)^{-1}X'Q^{-1}y$$

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 - \frac{b^2 - 4ac}{4a}$$

$$x'Ax - 2b'x = (x - d)'A(x - d) - b'Ab \quad d = A^{-1}b$$

$$\text{tr}(X'AX) - 2\text{tr}(B'X) = \text{tr}((X - D)'A(X - D)) - \text{tr}(B'AB)$$

$$D = A^{-1}B$$

$$(x - v_1)'W_1(x - v_1) + (x - v_2)'W_2(x - v_2) = (x - \tilde{x})'(W_1 + W_2)(x - \tilde{x})$$

$$+ (v_1 - v_2)'(W_1^{-1} + W_2^{-1})^{-1}(v_1 - v_2)$$

$$(v_1 - v_2)'W_1(W_1 + W_2)^{-1}W_2(v_1 - v_2)$$

