最小 2 乗法

アギレ・エルナン

信州大学　電子情報システム工学部

数値計算

2023 年 5 月 29 日

最小2乗法
Least squares method

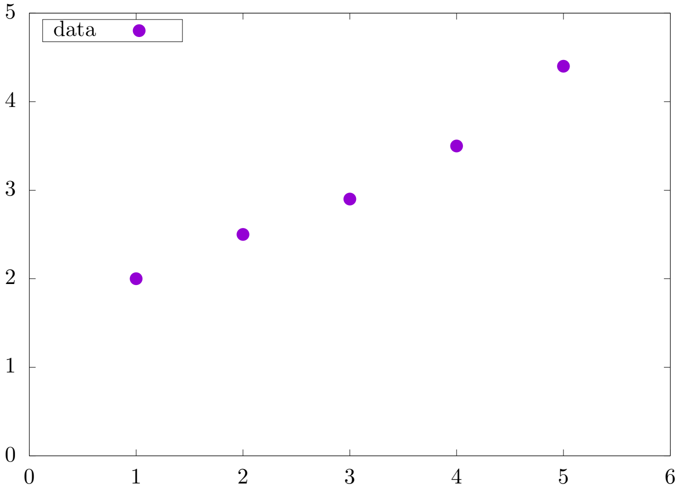
$$A^t Ax = A^t b$$

最小２乗法

データ点の近くを通るなめらかな関数を求める方法の一つとして、
最小２乗法がよく用いられる。

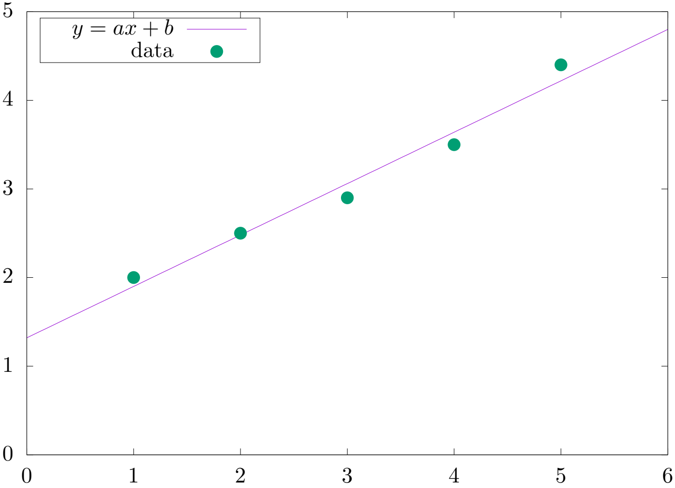
最小 2 乗法の導入 1/7

次の図における，点がほぼ直線状にみえる



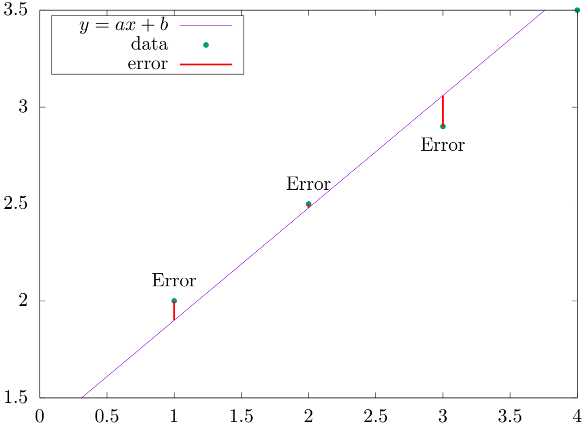
最小 2 乗法の導入 2/7

少しのずれがあっても，
各点を通る直線の計算ができる方法は**最小 2 乗法**という



最小 2 乗法の導入 3/7

- 最小 2 乗法は誤差 E を最小化して線の係数を見つけ出す
- 例えば，直線の方程式 $y = ax + b$ における与えられたデータから a, b 係数を見つけ出す



最小 2 乗法の導入 4/7

未知数は x の代わりに直線のパラメータ a, b を未知数にすると

$$y = ax + b \qquad \rightarrow \qquad y_i = ax_i + b$$

$\therefore$ 与えられた点 (x_i,y_i) の誤差 e_i において

$$e_i = ax_i + b - y_i$$

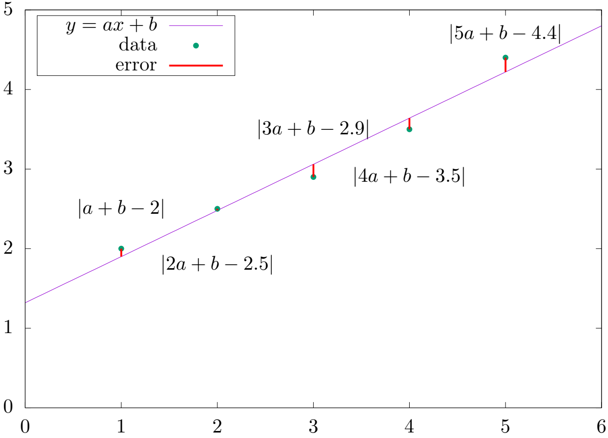
- 未知数は a, b です
- 与えられたデータから既知数は x_i, y_i です

最小２乗法の導入 5/7

符号が付いている誤差

$$E_0 = |a+b-2|+|2a+b-2.5|+|3a+b-2.9|+|4a+b-3.5|+|5a+b-4.4|$$

X	Y
1.0	2.0
2.0	2.5
3.0	2.9
4.0	3.5
5.0	4.4

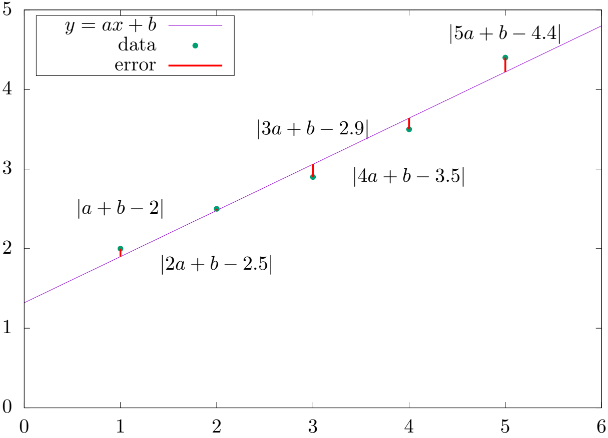


最小 2 乗法の導入 5/7

符号が付いていない誤差

$$E = (a + b - 2)^2 + (2a + b - 2.5)^2 + (3a + b - 2.9)^2 + (4a + b - 3.5)^2 + (5$$

X	Y
1.0	2.0
2.0	2.5
3.0	2.9
4.0	3.5
5.0	4.4



最小 2 乗法の導入 6/7

符号が付いていない誤差を小さくなるとともに，a, b を求める.

$$E = (a+b-2)^2+(2a+b-2.5)^2+(3a+b-2.9)^2+(4a+b-3.5)^2+(5a+b-4.4)^2$$

極値条件より $\dfrac{\partial E}{\partial a} = 0$, $\dfrac{\partial E}{\partial b} = 0$ と解けばよい

$$\begin{aligned}\frac{\partial E}{\partial a} &= 2(a+b-2)+4(2a+b-2.5)+6(3a+b-2.9)\\ &\quad +8(4a+b-3.5)+10(5a+b-4.4)=0\\ &\therefore \frac{\partial E}{\partial a}=0 \Rightarrow \underline{110a+30b=86.6}\end{aligned}$$

$$\begin{aligned}\frac{\partial E}{\partial b} &= 2(a+b-2)+2(2a+b-2.5)+2(3a+b-2.9)\\ &\quad +2(4a+b-3.5)+2(5a+b-4.4)=0\\ &\therefore \frac{\partial E}{\partial b}=0 \Rightarrow \underline{15a+5b=15.3}\end{aligned}$$

最小 2 乗法の導入 7/7

問題は $Ax = b$ の形になり，ガウス消去あるいは LU 分解より，次の連立線形方程式を解けば a, b が分かる

$$110a + 30b = 86.6$$
$$15a + 5b = 15.3$$

問題

上記の連立方程式を lu_solve 関数で解きましょう

最小 2 乗法の導入 7/7

問題は $Ax = b$ の形になり，ガウス消去あるいは LU 分解より，次の連立線形方程式を解けば a, b が分かる

$$110a + 30b = 86.6$$
$$15a + 5b = 15.3$$

問題

上記の連立方程式を lu_solve 関数で解きましょう
答:

$$a \approx 0.58$$
$$b \approx 1.32$$

最小 2 乗法の一般化 1/8
連立方程式

直線あるいは曲線に限らず問題の性質から，いくつかの
基本になる関数 $f_1(x), f_2(x), \cdots, f_k(x)$の１次結合として，

$$y = a_1 f_1(x) + a_2 f_2(x) + \cdots + a_k f_k(x) \tag{4.23}$$

与えられたデータ点 (x_i, y_i) を上記の式に代入して

$$\left\{ \begin{array}{c} a_1 f_1(x_1) + a_2 f_2(x_1) + \cdots + a_k f_k(x_1) = y_1 \\ a_1 f_1(x_2) + a_2 f_2(x_2) + \cdots + a_k f_k(x_2) = y_2 \\ \vdots \\ a_1 f_1(x_n) + a_2 f_2(x_n) + \cdots + a_k f_k(x_n) = y_n \end{array} \right\} \tag{4.24}$$

最小 2 乗法の一般化 2/8
Ax=b ?

- $Ax = b$
- **未知数 x : $a_1, \cdots, a_k$ (k 基本関数の係数)**
- **与えられた点 (x_i, y_i) $(i = 1, \cdots, n)$ から係数行列 A とベクター b**

$$\begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_k(x_n) \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix}$$

- この連立方程式は必ずしも一意解をもつとは限らないし、解があって対応する式 (4.23) のグラフがデータ点を通るとしても、起伏がなめらかになっているとは限らない式 (4.23) がすべてのデータ点を通るようにするよりも, それらの点の近くを通って、全体として無理のないなめらかな曲線になっているほうがよい.

最小 2 乗法の一般化 3/8

各点 (x_i, y_i) の誤差

- そこで、連立方程式 (4.24) の等式がそれぞれ近似的に成立し, それぞれの誤差ができるだけ小さくなるように係数 $a_1, a_2, \cdots, a_k$ を定めようというのが最小 2 乗法の考え方である.

$$\text{点 } x_1 \text{の誤差} \quad \left| \sum_{i=1}^{k} f_i(x_1)a_i - y_1 \right|$$

$$\text{点 } x_2 \text{の誤差} \quad \left| \sum_{i=1}^{k} f_i(x_2)a_i - y_2 \right|$$

$$\vdots$$

$$\text{点 } x_n \text{の誤差} \quad \left| \sum_{i=1}^{k} f_i(x_n)a_i - y_n \right|$$

最小 2 乗法の一般化 4/8
誤差

符号が付いていない誤差

$$E = \left(\sum_{i=1}^{k} f_i(x_1)a_i - y_1\right)^2 + \left(\sum_{i=1}^{k} f_i(x_2)a_i - y_2\right)^2 + \cdots + \left(\sum_{i=1}^{k} f_i(x_n)a_i - y_n\right)^2$$

最小 2 乗法の一般化 5/8
誤差の最小化

E の 2 次の各係数は全て正だから，

$$\frac{\partial E}{a_1} = 0, \frac{\partial E}{a_2} = 0, \cdots, \frac{\partial E}{a_k} = 0$$

を連立して解けばよい

$f_i(x_j) = f_{ij}$ と書くと

$$\begin{aligned} E = &(f_{11}a_1 + f_{21}a_2 + \cdots + f_{k1}a_k - y_1)^2 \\ &+(f_{12}a_1 + f_{22}a_2 + \cdots + f_{k2}a_k - y_2)^2 \\ &+ \cdots + (f_{1n}a_1 + f_{2n}a_2 + \cdots + f_{kn}a_k - y_n)^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial E}{\partial a_1} = &2(f_{11}a_1 + f_{21}a_2 + \cdots + f_{k1}a_k - y_1)f_{11} \\ &+2(f_{12}a_1 + f_{22}a_2 + \cdots + f_{k2}a_k - y_2)f_{12} \\ &+ \cdots + 2(f_{1n}a_1 + f_{2n}a_2 + \cdots + f_{kn}a_k - y_n)f_{1n} = 0 \end{aligned}$$

最小 2 乗法の一般化 6/8
誤差の最小化

未知数 $a_1, a_2, \cdots, a_k$ **を含む項と既知数の項を分けると**

$$
\begin{aligned}
f_{11}(f_{11}a_1 + f_{21}a_2 + \cdots + f_{k1}a_k) + f_{12}(f_{12}a_1 + f_{22}a_2 + \cdots + f_{k2}a_k) \\
+ \cdots + f_{1n}(f_{1n}a_1 + f_{2n}a_2 + \cdots + f_{kn}a_k) \\
= f_{11}y_1 + f_{12}y_2 + \cdots + f_{1n}y_n
\end{aligned}
$$

となり，両辺を行列の積の形に書き表すと，

$$
[f_{11} f_{12} \cdots f_{1n}] \begin{bmatrix} f_{11}a_1 + f_{21}a_2 + \cdots + f_{k1}a_k \\ f_{12}a_1 + f_{22}a_2 + \cdots + f_{k2}a_k \\ \vdots \\ f_{1n}a_1 + f_{2n}a_2 + \cdots + f_{kn}a_k \end{bmatrix} = [f_{11} f_{12} \cdots f_{1n}] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}
$$

最小２乗法の一般化 7/8
誤差の最小化

$$\frac{\partial E}{\partial a_1} = 0, \frac{\partial E}{\partial a_2} = 0, \cdots, \frac{\partial E}{\partial a_k} = 0$$

からも同様な等式が得られる．それをまとめて書くと

$$\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{k1} & f_{k2} & \cdots & f_{kn} \end{bmatrix} \begin{bmatrix} f_{11}a_1 + f_{21}a_2 + \cdots + f_{k1}a_k \\ f_{12}a_1 + f_{22}a_2 + \cdots + f_{k2}a_k \\ \vdots \\ f_{1n}a_1 + f_{2n}a_2 + \cdots + f_{kn}a_k \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{k1} & f_{k2} & \cdots & f_{kn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

最小 2 乗法の一般化 8/8
誤差の最小化

左辺の第 2 行列を積で表すと

$$\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{k1} & f_{k2} & \cdots & f_{kn} \end{bmatrix} \begin{bmatrix} f_{11} & f_{21} & \cdots & f_{k1} \\ f_{12} & f_{22} & \cdots & f_{k2} \\ \vdots & \vdots & \ddots & \vdots \\ f_{1n} & f_{2n} & \cdots & f_{kn} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{k1} & f_{k2} & \cdots & f_{kn} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$A^T A x = A^T b \Rightarrow$ 最小 2 乗の正規方程式

上記の連立方程式を掃き出し法により，$a_1, a_2, \cdots, a_k$ を求められる

まとめ

- 与えられた点を $(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$ とする
- いくつかの基本になる関数 $f_1(x), f_2(x), \cdots, f_k(x)$ を指定する

$$y = a_1 f_1(x) + a_2 f_2(x) + \cdots + a_k f_k(x)$$

  とおく
- 最小 2 乗を適用して，未定係数 $a_1, a_2, \cdots, a_k$ を定めるには，正規方程式

$$A^T A x = A^T b$$

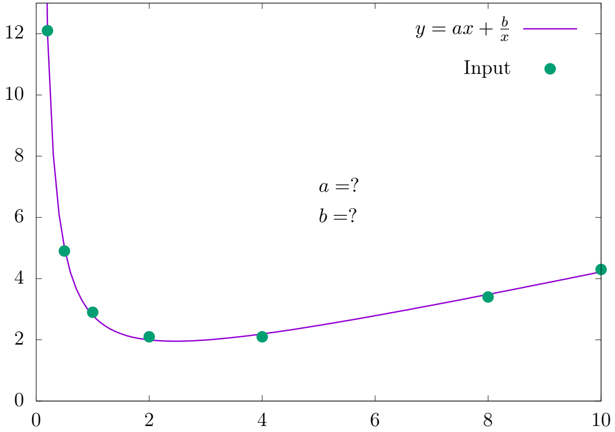
  より x を求めればよい．ここに

$$\mathbf{A} = \begin{bmatrix} f_1(x_1) & f_2(x_1) & \cdots & f_k(x_1) \\ f_1(x_2) & f_2(x_2) & \cdots & f_k(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ f_1(x_n) & f_2(x_n) & \cdots & f_k(x_n) \end{bmatrix}, \mathbf{x} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_l \end{bmatrix} \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

例題 1/6

- $y = ax + \dfrac{b}{x}$ の関係関数があることが分かっている
- 与えられた点 x, y を使って最小 2 乗法を適用して a, b を求め, y を x の式で表せ

X	Y
0.2	12.1
0.5	4.9
1.0	2.9
2.0	2.1
4.0	2.1
8.0	3.4
10.0	4.3



例題 2/6

基本になる関数

$y = ax + \dfrac{b}{x}$ ですから，

$$f_1(x) = x$$
$$f_2(x) = \frac{1}{x}$$

$$\left\{\begin{array}{ccccc} 0.2a & + & 5b & = & 12.1 \\ 0.5a & + & 2b & = & 4.9 \\ a & + & b & = & 2.9 \\ 2a & + & 0.5b & = & 2.9 \\ 4a & + & 0.25b & = & 2.9 \\ 8a & + & 0.125b & = & 2.9 \\ 10a & + & 0.1b & = & 2.9 \end{array}\right\} \Rightarrow \mathbf{A} = \begin{bmatrix} 0.2 & 5 \\ 0.5 & 2 \\ 1.0 & 1 \\ 2.0 & 0.5 \\ 4.0 & 0.25 \\ 8.0 & 0.125 \\ 10.0 & 0.1 \end{bmatrix} \mathbf{b} = \begin{bmatrix} 12.1 \\ 4.9 \\ 2.9 \\ 2.1 \\ 2.1 \\ 3.4 \\ 4.3 \end{bmatrix}$$

例題 3/6

基本になる関数

$$A^T A = \begin{bmatrix} 185.29 & 7 \\ 7 & 30.338125 \end{bmatrix}, A^T b = \begin{bmatrix} 90.57 \\ 75.63 \end{bmatrix}$$

となる
$A^T Ax = A^T b$ は $Ax = b$ の形になったので，掃き出すと

$$a = 0.3980930$$
$$b = 2.4010498$$

例題 4/6

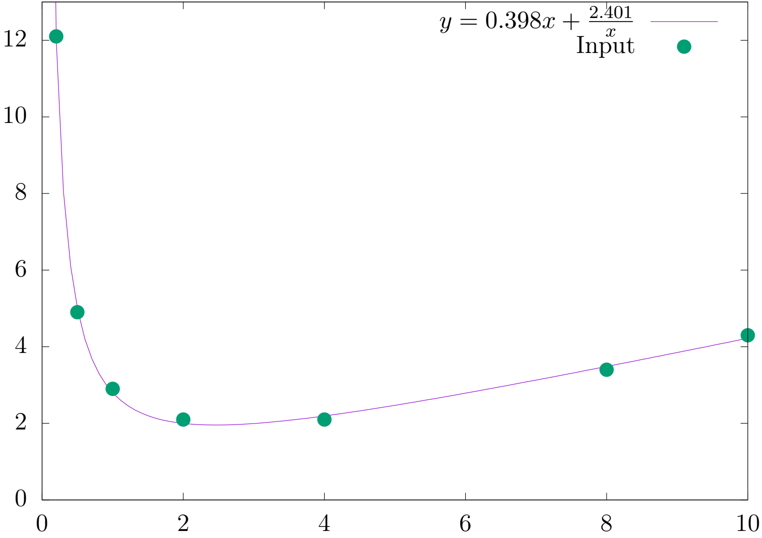
基本になる関数

$$A^T A = \begin{bmatrix} 185.29 & 7 \\ 7 & 30.338125 \end{bmatrix}, A^T b = \begin{bmatrix} 90.57 \\ 75.63 \end{bmatrix}$$

となる
$A^T Ax = A^T b$ は $Ax = b$ の形になったので，掃き出すと

$$a = 0.3980930$$
$$b = 2.4010498$$

例題 5/6



例題 6/6