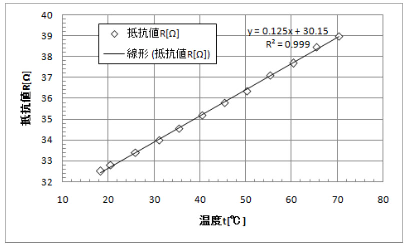
Excelを用いた実験データ処理2

1

近似式(回帰式)の導出の代表的な方法は最小二乗法である．Excelの近似曲線を用いると，この原理を知らなくても近似式を求めることはできる．機械技術者の強みは，その原理および特徴を理解しており，様々な分野に応用できることである．



今回の目標

(1) 最小二乗法の原理を理解する．
(2) Excelの表計算により近似式を導出する．

近似式の計算表

2

表 B3-1 最小二乗法による金属抵抗の温度変化の解析

番号	温度 t [°C]	抵抗値 R [Ω]	t^2 [°C²]	tR [°C·Ω]
1	18.3	32.50	334.89	594.750
2	20.5	32.81	420.25	672.605
3	25.9	33.39	670.81	864.801
4	31.0	33.99	961.00	1053.690
5	35.4	34.55	1253.16	1223.070
6	40.5	35.19	1640.25	1425.195
7	45.4	35.78	2061.16	1624.412
8	50.2	36.36	2520.04	1825.272
9	55.3	37.10	3058.09	2051.630
10	60.4	37.69	3648.16	2276.476
11	65.4	38.47	4277.16	2515.938
12	70.3	38.99	4942.09	2740.997
合計	518.6	426.82	25787.06	18868.836

標準偏差の計算表

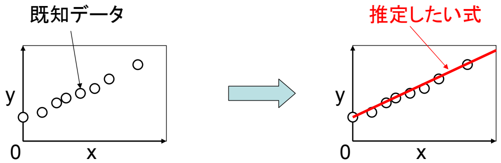
3

表 B3-2 残差の計算

番号	実測値 R [Ω]	最確値 $\bar{R}$ [Ω]	$\Delta = R - \bar{R}$ [Ω]	Δ^2 [Ω²]
1	32.50	32.44	0.06	3.6×10^{-3}
2	32.81	32.72	0.09	8.1
3	33.39	33.40	−0.01	0.1
4	33.99	34.03	−0.04	1.6
5	34.55	34.59	−0.04	1.6
6	35.19	35.22	−0.03	0.9
7	35.78	35.84	−0.06	3.6
8	36.36	36.44	−0.08	6.4
9	37.10	37.08	0.02	0.4
10	37.69	37.72	−0.03	0.9
11	38.47	38.34	0.13	16.9
12	38.99	38.96	0.03	0.9
合計				4.5×10^{-2}

最小二乗法(直線の場合)

4

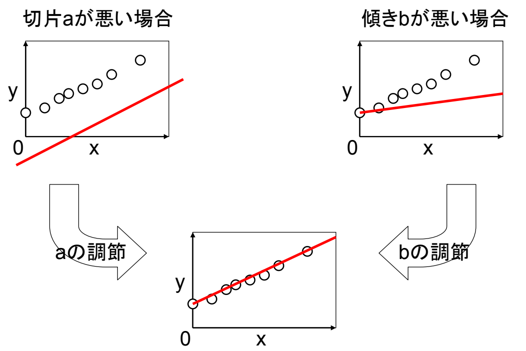


近似式	$y = a + bx$
既知量（データ）	$(x_i, y_i), (i = 1,2,\cdots,n)$
未知量（係数）	a, b

問題 データと近似式の誤差を最も小さくする時の係数a,bを求めなさい

係数a,bの調節

5



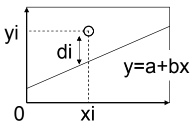
切片aと傾きbを調節して，誤差を小さくする

最適な係数a,bの条件

6

データと近似式の誤差

$$d_i = y_i - (a + bx_i)$$



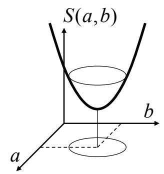
誤差の2乗和

$$S(a,b) = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (y_i - a - bx_i)^2$$

最適な係数a,bの条件

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

接線の傾きが零



(註) $\frac{\partial}{\partial a}, \frac{\partial}{\partial b}$ は偏微分分

最適な係数a,b

7

接線の傾き

$$\frac{\partial}{\partial a} S(a,b) = 2\sum(-1)(y_i - a - bx_i) = 2na + 2b\sum x_i - 2\sum y_i$$

$$\frac{\partial}{\partial b} S(a,b) = 2\sum(-x_i)(y_i - a - bx_i) = 2a\sum x_i + 2b\sum x_i^2 - 2\sum x_i y_i$$

最適な係数a,bは次の式より得られる．

$$na + b\sum x_i = \sum y_i$$

$$a\sum x_i + b\sum x_i^2 = \sum x_i y_i$$

行列を用いて，データ(既知量)と係数(未知量)を分離する

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

正規方程式

$$\hat{a} = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$\hat{b} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

近似式の計算

8

計算に必要な値

$$n, \quad \sum x_i, \quad \sum x_i^2, \quad \sum y_i, \quad \sum x_i y_i$$

係数a,bを求めるための表

i	x_i	y_i	x_i^2	$x_i y_i$
1	x_1	y_1	x_1^2	$x_1 y_1$
2	x_2	y_2	x_2^2	$x_2 y_2$
⋮	⋮	⋮	⋮	⋮
n	x_n	y_n	x_n^2	$x_n y_n$
合計	$\sum x_i$	$\sum y_i$	$\sum x_i^2$	$\sum x_i y_i$

$$\hat{a} = \frac{\sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

$$\hat{b} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}$$

近似式

$$y = \hat{a} + \hat{b}x$$

Excelによる近似式の計算(表)

9

	A	B	C	D	E
1	番号	温度t[℃]	抵抗値R[Ω]	t^2	tR
2	1	18.3	32.50	334.89	594.750
3	2	20.5	32.81	420.25	672.605
:	:	:	:	:	:
12	11	65.4	38.47	4277.16	2515.938
13	12	70.3	38.99	4942.09	2740.997
14	合計	518.6	426.8	25787.06	18868.836
15					
16	行列式=	40498.8			
17	近似式	a		b	
18	R=	30.15	+	0.125	t

Excelによる近似式の計算(式)

10

	A	B	C	D	E
1	番号	温度t[℃]	抵抗値R[Ω]	t^2	tR
2	1	18.3	32.50	=B2^2	=B2*C2
3	2	20.5	32.81	=B3^2	=B3*C3
:	:	:	:	:	:
12	11	65.4	38.47	=B12^2	=B12*C12
13	12	70.3	38.99	=B13^2	=B13*C13
14	合計	=SUM(B2:B13)	=SUM(C2:C13)	=SUM(D2:D13)	=SUM(E2:E13)
15					
16	行列式=	40498.8			
17	近似式	a		b	
18	R=	30.15	+	0.125	t

B16セル: =A13*D14-B14^2
B18セル: =(D14*C14-B14*E14)/B16
D18セル: =(A13*E14-B14*C14)/B16

Excelによる近似式の計算(手順)

11

数値の作成手順
(1) セルD2に式「=B2^2」を入力
(2) セルE2に式「=B2*C2」を入力
(3) セルD2,E2の式をセルD3〜E13にコピー
(4) セルB14に式「=sum(B2:B13)」を入力
(5) セルB14の式をセルC14〜E14にコピー
(6) セルB16に式「=A13*D14-B14^2 」を入力
(7) セルB18に式「=(D14*C14-B14*E14)/B16」を入力
(8) セルD18に式「=(A13*E14-B14*C14)/B16」を入力

手順(3),(5)のコピーは
数式を相対指定する

係数a,bの標準偏差

12

標準偏差

$$s(a) = \sqrt{\frac{\sum x_i^2}{n\sum x_i^2 - (\sum x_i)^2} \cdot \frac{\sum d_i^2}{n-2}}$$

$$s(b) = \sqrt{\frac{n}{n\sum x_i^2 - (\sum x_i)^2} \cdot \frac{\sum d_i^2}{n-2}}$$

計算に必要な値
(追加)
$\sum d_i^2$

係数a,bの標準偏差を求めるための表

i	y_i	$\hat{y}_i = a + bx_i$	$d_i = y_i - \hat{y}_i$	d_i^2
1	y_1	$\hat{y}_1$	d_1	d_1^2
2	y_2	$\hat{y}_2$	d_2	d_2^2
⋮	⋮	⋮	⋮	⋮
n	y_n	$\hat{y}_n$	d_n	d_n^2
合計	---	---	---	$\sum d_i^2$

Excelによる標準偏差の計算

13

	A	...	F	G	H
1	番号	...	最確値	誤差d	d^2
2	1	...	32.44	0.06	0.0031
3	2	...	32.72	0.09	0.0080
:	:	...	:	:	:
12	11	...	38.35	0.12	0.0146
13	12	...	38.96	0.03	0.0007
14	合計	...			0.0437
15		...			
16	行列式=	...		標準偏差	
17	近似式	...		s(a)=	0.0527
18	R=	...		s(b)=	0.0011

Excelによる標準偏差の計算

14

	A	...	F	G	H
1	番号	...	最確値	誤差d	d^2
2	1	...	=$B$18+$D$18*B2	=C2-F2	=G2^2
3	2	...	=$B$18+$D$18*B3	=C3-F3	=G3^2
:	:	...	:	:	:
12	11	...	=$B$18+$D$18*B12	=C12-F12	=G12^2
13	12	...	=$B$18+$D$18*B13	=C13-F13	=G13^2
14	合計	...			=SUM(H2:H13)
15		...			
16	行列式=	...		標準偏差	
17	近似式	...		s(a)=	0.0527
18	R=	...		s(b)=	0.0011

H17セル: =SQRT(D14*H14/B16/(A13-2))
H18セル: =SQRT(A13*H14/B16/(A13-2))

Excelによる標準偏差の計算(手順)

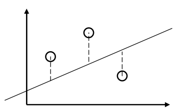
15

数値の作成手順
(1) セルF2に式「=$B$18+$D$18*B2」を入力
(2) セルG2に式「=C2-F2」を入力
(3) セルH2に式「=G2^2」を入力
(4) セルF2,G2,H2の式をセルF3〜H13にコピー
(5) セルH14に式「=sum(H2:H13)」を入力
(6) セルH17に式「=SQRT(D14*H14/B16/(A13-2))」を入力
(7) セルH18に式「=SQRT(A13*H14/B16/(A13-2))」を入力

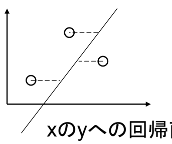
手順(1)の「$」により
手順(3)のコピーで
セルが絶対指定される.
「$」直後の記号を固定

回帰直線

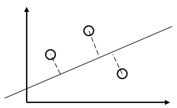
16



yのxへの回帰直線



xのyへの回帰直線



直交回帰直線

	近似式	x	y
yのxへの回帰直線	y=a+bx	正確	誤差あり
xのyへの回帰直線	x=a+by	誤差あり	正確
直交回帰直線	ax+by+c=0	誤差あり	誤差あり

n次多項式による近似

17

データをn次多項式で近似する

$$y = a_0 + a_1 x + \cdots + a_p x^p$$

既知量	$(x_i, y_i), (i = 1, 2, \cdots, n)$
未知量	$a_0, a_1, \cdots, a_p$

誤差の2乗和

$$S(a_0, a_1, \cdots, a_p) = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} \{y_i - (a_0 + a_1 x_i + \cdots + a_p x_i^p)\}^2$$

最適な係数ak,(k=0,1,...,p)の条件

$$\frac{\partial}{\partial a_k} S(a_0, a_1, \cdots, a_p) = -2\sum x_i^k \{y_i - (a_0 + a_1 x_i + \cdots + a_p x_i^p)\} \to 0$$

得られた関係式

$$a_0 \sum x_i^k + a_1 \sum x_i^{k+1} + \cdots + a_1 \sum x_i^{k+p} = \sum x_i^k y_i \quad \text{(k=0,1,...,p)}$$

正規方程式

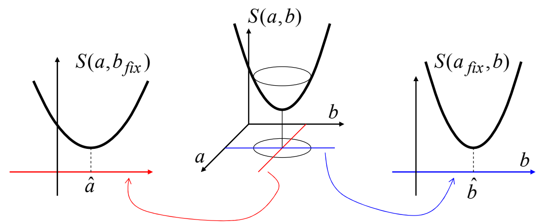
$$\begin{pmatrix} n & \sum x_i & \cdots & \sum x_i^p \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{p+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^p & \sum x_i^{p+1} & \cdots & \sum x_i^{2p} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^p y_i \end{pmatrix}$$

(用語)偏微分

18

$$\frac{\partial}{\partial a} S(a,b)$$

関数Sは変数a,bに依存する.
この関係において,bを固定してaのみを変化させたとき(aの1変数関数と考えたとき)の
Sの変動(傾き)に注視する



全微分 $dS(a,b) = \frac{\partial S(a,b)}{\partial a} da + \frac{\partial S(a,b)}{\partial b} db$