A　最小二乗法

熱電対の較正曲線を作る上で必要な最小二乗法について説明する。第一部では、標準熱起電力表を用いて換算した仮の温度を参照データと比較するように書いてあるが、傾きがほぼ1の直線が得られる例では似た数値が現れて紛らわしいので、この解説では熱起電力と温度の関係を示すグラフを例に取る。小レポートでは文献値の温度 vs. 測定された温度で結果を示すこと。計算方法はどちらでも同じである。

実験の結果として表5のような測定値が得られたとしよう。熱起電力が与えられた時に温度が求められる関係式が欲しい。温度と起電力は一次式 $E = a \times T + b$ で関係づけられると仮定しよう。ここで a 及び b は測定系によって決る定数である。実験で求められた4点全てを通る直線があれば良いのだが、一般にはこうした線は存在しない。そこで4点との距離が最も小さい直線が最も尤もらしい線ということになる。この直線と水の融点、沸点、スズの融点、亜鉛の融点での実験値との差はそれぞれ $-0.018-(a\times0.0+b)$、$4.12-(a\times100.0+b)$、$9.34-(a\times232.0+b)$、$17.23-(a\times419.6+b)$ である。実験値と直線との差 (この差を「残差」と呼ぶ) はプラスにもマイナスにもなるので、これらの二乗の和を取ったものが最小になる a と b を求める。つまり

$$\begin{aligned} S &= \{-0.018 - (a \times 0.0 + b)\}^2 + \{4.12 - (a \times 100.0 + b)\}^2 && (1)\\ &+ \{9.34 - (a \times 232.0 + b)\}^2 + \{17.23 - (a \times 419.6 + b)\}^2 && (2)\end{aligned}$$

を最小にする a と b を求めるというのが、最小二乗法のアイディアである。これには

$$\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = 0 \tag{3}$$

を解けばいい。すなわち

$$\begin{aligned} \frac{\partial S}{\partial a} &= -2 \times \Big\{[-0.018 - (a \times 0.0 + b)] \times 0.0 + [4.12 - (a \times 100.0 + b)] \times 100.0 \\ &\quad +[9.34 - (a \times 232.0 + b)] \times 232.0 + [17.23 - (a \times 419.6 + b)] \times 419.6\Big\} \\ &= -2 \times \Big(9.809 \times 10^3 - 2.399 \times 10^5 a - 7.516 \times 10^2 b\Big) = 0 \end{aligned}$$

及び

$$\begin{aligned} \frac{\partial S}{\partial b} &= -2 \times \Big\{[-0.018 - (a \times 0.0 + b)] + [4.12 - (a \times 100.0 + b)] \\ &\quad +[9.34 - (a \times 232.0 + b)] + [17.23 - (a \times 419.6 + b)]\Big\} \\ &= -2 \times \Big(3.067 \times 10^1 - 7.516 \times 10^2 a - 4b\Big) = 0 \end{aligned}$$

表 5: 実験結果の例。

温度定点	真の温度 (°C)	熱起電力 (mV)
水の融点	0.0	-0.018
水の沸点	100.0	4.12
錫の融点	232.0	9.34
亜鉛の融点	419.6	17.23

を満す a と b を求めればいい。連立方程式を解いて、答は $a = 4.099 \times 10^{-2}$ 及び $b = -3.7 \times 10^{-2}$ となる。

上の手続きを実験のたびに繰り返すのは煩わしいので、手順を一般化しておこう。N 個の測定点 $x = x_1, x_2, \ldots x_N$ に対して測定値 $y = y_1, y_2, \ldots y_N$ を得たとしよう (上の例では $N = 4$ で $x_1 = 0.0$、$x_2 = 100.0$、$x_3 = 232.0$、$x_4 = 419.6$、$y_1 = -0.018$、$y_2 = 4.12$、$y_3 = 9.34$、$y_4 = 17.23$)。最小化すべき残差の二乗和は

$$S = \sum_{i=1}^{N} \{y_i - (ax_i + b)\}^2 \tag{4}$$

と書ける。これを微分すると

$$\begin{aligned} \frac{\partial S}{\partial a} &= \sum_{i=1}^{N} -2x_i\{y_i - (ax_i + b)\} \\ &= -\sum_{i=1}^{N} 2x_i y_i + 2a\sum_{i=1}^{N} x_i^2 + 2b\sum_{i=1}^{N} x_i = 0 \end{aligned}$$

及び

$$\begin{aligned} \frac{\partial S}{\partial b} &= \sum_{i=1}^{N} -2\{y_i - (ax_i + b)\} \\ &= -\sum_{i=1}^{N} 2y_i + 2a\sum_{i=1}^{N} x_i + 2b\sum_{i=1}^{N} 1 = 0 \text{。} \end{aligned}$$

故に a と b を求める連立方程式は

$$a\sum_{i=1}^{N} x_i^2 + b\sum_{i=1}^{N} x_i = \sum_{i=1}^{N} x_i y_i \tag{5}$$

$$a\sum_{i=1}^{N} x_i + b\sum_{i=1}^{N} 1 = \sum_{i=1}^{N} y_i \tag{6}$$

となる。この方程式中で必要な係数は表 6 のような表を作れば機械的に求められる。

問：実験結果を二次式 $y = ax^2 + bx + c$ にフィットするにはどんな式を解けばいいか。

表 6: 係数を求めるための表。

	x_i	x_i^2	y_i	$x_i y_i$
1	0.0	0.0	-0.018	0.0
2	100.0	1.000×10^4	4.12	4.12×10^2
3	232.0	5.382×10^4	9.34	2.167×10^3
4	419.6	1.761×10^5	17.23	7.230×10^3
合計	751.6	2.399×10^5	30.67	9.809×10^3