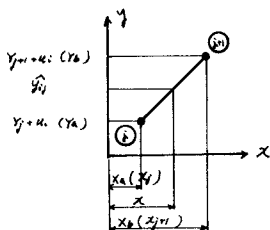


[1] まえがき

曲線回帰分析に於いては、従来 ①直交多項式の応用 ②重回帰分析の応用、の二つの手法が用いられているが、前者については、汎用性の点で、後者については、Multi-colinearity の点で実用に支障を感じる。本文は、既往理論の難点を解決すべく、回帰曲線を、有限長の直線要素でモデル化したものであり、既往理論に比しての得失は、構造力学に於ける有限要素法のそれに等しい。

[2] 正規方程式の誘導



ここで、プロックとは、回帰方程式の勾配が等しく、定数項のみが異なるグループ(片)と定義する。図に於て u_i を各プロックの別個因子として

$$x_j < x \leq x_{j+1} \quad \dots (1)$$

の各要素の推定値を

$$u_j = u_i + \alpha_i + \beta_j x \quad \dots (2)$$

と置き、以後、適宜、サフィックスを略して記せば、

$$\beta = \frac{y_b - y_a}{x_b - x_a}, \quad \alpha = y_b - \frac{y_b - y_a}{x_b - x_a} x_b \quad \dots (3)$$

である故

$$t = \frac{1}{x_b - x_a}, \quad \theta = \frac{1}{x} \quad \dots (4)$$

とおけば

$$\beta = t(y_b - y_a), \quad \alpha = \frac{\theta y_b - \theta_a y_a}{\theta_b - \theta_a} \triangleq g_b y_b - g_a y_a \quad \dots (5)$$

である。ここで、 m を各プロック、各要素の観測数とし

$$x = [x_1 \dots x_m]^T, \quad y = [y_1 \dots y_m]^T, \quad J = [1 \dots 1]^T \quad \dots (6)$$

とおけば、回帰要素からの残差平方和は

$$\begin{aligned} e &= [y - (u + \alpha)J - \beta x]^T [y - (u + \alpha)J - \beta x] \\ &= [y_b^T J^T J + 2t g_b J^T x + t^2 x^T x] y_b^2 \\ &\quad + [y_a^T J^T J + 2t g_a J^T x + t^2 x^T x] y_a^2 \\ &\quad - 2[J^T J g_a y_b + (g_a + g_b) J^T x + t^2 x^T x] y_a y_b \\ &\quad + 2[t x^T x + g_b J^T J] y_b u - 2[t x^T x + g_a J^T J] y_a u \\ &\quad - 2[t x^T y + g_a J^T y] y_b + 2[t x^T y + g_b J^T y] y_a \end{aligned}$$

$$+ u^T J^T J - 2u J^T y + y^T y \quad \dots (7)$$

となり、これを

$$e = c y_b^2 + d x^2 - 2f y_b y_a + 2g y_b - 2h y_a - p u^2 + 2g y_b u - 2r y_a u - 2s u + w \quad \dots (8)$$

と書く。プロック数を M とし、全プロックについて集計し、

$$\bar{e} = \sum_{i=1}^M e_i, \quad \bar{c} = \sum_{i=1}^M c_i, \quad \bar{d} = \sum_{i=1}^M d_i \quad \dots (9)$$

とおき、サフィックスを適宜復活して記せば、

$$\begin{aligned} \bar{e} &= \bar{w} + \bar{c}_i y_i^2 + \bar{d}_i x_i^2 - 2\bar{f}_i y_i y_a + 2\bar{g}_i y_i - 2\bar{h}_i y_a \\ &\quad + (p_i u_i^2 + \dots p_{mi} u_m^2) + 2(\bar{g}_{ij} + \dots \bar{g}_{mi} u_m) y_i + \\ &\quad - 2(r_i u_i + \dots r_{mi} u_m) y_a - 2(\bar{s}_{ij} u_i + \dots \bar{s}_{mi} u_m) \quad \dots (10) \end{aligned}$$

全要素について集計し

$$\begin{aligned} \bar{e} &\triangleq \sum_{i=1}^M (\bar{e} - \bar{w}) \\ &= (-c_i y_i^2 + \dots c_n y_n^2) + (d_i x_i^2 + \dots d_n x_n^2) - 2(f_i y_i y_a + \dots f_n y_n y_a) \\ &\quad + 2(g_i y_a + \dots g_n y_n) - 2(h_i y_i + \dots h_n y_n) \\ &\quad + [(p_{i1} + \dots p_{in}) u_i^2 + \dots (p_{m1} + \dots p_{mn}) u_m^2] \\ &\quad + 2[(g_{i1} u_i + \dots g_{in} u_n) y_a + \dots (g_{m1} u_i + \dots g_{mn} u_n) y_n] \\ &\quad - 2[(r_{i1} u_i + \dots r_{in} u_n) y_a + \dots (r_{m1} u_i + \dots r_{mn} u_n) y_n] \\ &\quad - 2[(s_{i1} + \dots s_{in}) u_i + \dots (s_{m1} + \dots s_{mn})] \quad \dots (11) \end{aligned}$$

残差平方和最小の条件から

$$\begin{aligned} \frac{1}{2} \frac{\partial \bar{e}}{\partial y_i} &= d_i y_i - f_i y_a - [r_{i1} \dots r_{in}] [u_1 \dots u_n]^T - h_i \\ \frac{1}{2} \frac{\partial \bar{e}}{\partial y_n} &= (c_n + d_n) y_n - (f_{n1} y_a + \dots f_{nn} y_n) + g_{n1} - h_n \\ &\quad + [g_{n1} \dots g_{nn}] [u_1 \dots u_n]^T - [r_{n1} \dots r_{nn}] [u_1 \dots u_n]^T \\ \frac{1}{2} \frac{\partial \bar{e}}{\partial y_{n+1}} &= c_n y_{n+1} - f_n y_n + g_n + [g_{n1} \dots g_{nn}] [u_1 \dots u_n]^T \\ \frac{1}{2} \frac{\partial \bar{e}}{\partial u_k} &= (p_{k1} + \dots p_{kn}) u_k + [g_{k1} + \dots g_{kn}] [y_a \dots y_n]^T \\ &\quad - [r_{k1} \dots r_{kn}] [y_a \dots y_n]^T - (s_{k1} + \dots s_{kn}) \quad \dots (12) \end{aligned}$$

$$= 0,$$

$$\bar{p}_i = (p_{i1} + \dots p_{in}), \quad \bar{s}_i = (s_{i1} + \dots s_{in}) \quad \dots (13)$$

$$C = \begin{bmatrix} d_1 & -f_1 \\ -f_1 & (c_1+d_2) & -f_2 \\ & -f_2 & (c_2+d_3) & -f_3 \\ & & -f_3 & c_n \end{bmatrix}, \quad T = \begin{bmatrix} r_{11} & r_{12} \\ r_{12} & r_{22} \\ r_{1n} & r_{2n} \\ 0 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \\ y_n \\ y_{n+1} \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\ g_1 \\ g_{n+1} \\ g_n \end{bmatrix}, \quad \tilde{P} = \begin{bmatrix} \tilde{p}_1 & \dots & \tilde{p}_n \\ \tilde{p}_1 & \dots & \tilde{p}_n \end{bmatrix}, \quad \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_n \end{bmatrix} \dots (14)$$

とおけば、正規方程式は

$$\begin{bmatrix} C & -T \\ -T^T & \tilde{P} \end{bmatrix} \begin{bmatrix} Y \\ U \end{bmatrix} = \begin{bmatrix} g \\ \delta \end{bmatrix} \dots (15)$$

となる。このままでは、係数行列は singular となり、別定数因子の比率のみが求まり、一般に不定となるため、 $u_1 \sim u_m$ の内、少なくとも一箇の変数を固定しなければならぬ。これは、構造力学に於ける変形法の、支条件の導入に相当する。ここで、例えば、 $u_k=0$ の時、係数行列、定数ベクトルを下記のように修正する。但し、斜線部はゼロ要素である。これから、連立方程式を解けばよい。

$$\begin{bmatrix} C & -T & Y & -g \\ -T^T & \tilde{P} & U & \delta \end{bmatrix} \begin{bmatrix} Y \\ U \end{bmatrix} = \begin{bmatrix} g \\ \delta \end{bmatrix} \dots (16)$$

[3] 数値計算例

文献 P 200 に掲げられた、3 台の化学製品製造装置に於ける、材料成分(x)と製品濃度(y)の測定結果(表-1)を例に採る。

装置 1		装置 2		装置 3	
x(%)	y(%)	x(%)	y(%)	x(%)	y(%)
51.45	16.25	48.71	11.60	52.58	15.80
47.35	19.70	53.12	15.08	46.45	9.80
53.23	17.69	45.95	7.75	51.56	14.78
45.98	11.60	50.70	10.85	53.67	16.50
53.85	19.00	46.75	9.44	47.90	10.06
48.35	14.13	52.41	10.20	51.65	13.28
47.00	14.80	47.95	10.30	48.09	11.63
50.77	12.78	50.70	11.84	51.69	13.88
48.60	16.10	48.75	8.45	49.70	13.10
52.10	18.75	47.40	11.05	48.70	13.10
46.70	13.02	48.93	9.76	52.49	16.78
52.10	17.60	54.11	13.15	47.60	11.12
50.69	16.90	50.00	12.80	50.18	11.55
48.52	15.58	52.14	13.29	51.06	15.30
50.30	15.80	53.12	12.60		
49.55	14.55				

要素数を 1 とする時、正規方程式は

$$\begin{bmatrix} 11.6518 & 10.6117 & 8.1615 & 7.4630 & 6.6390 \\ 10.6117 & 12.1248 & 7.8385 & 7.5370 & 7.4610 \\ 8.1615 & 7.8385 & 16.0 & 0 & 0 \\ 7.4630 & 7.5370 & 0 & 15.0 & 0 \\ 6.6390 & 7.4610 & 0 & 0 & 14.0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 290.590 \\ 314.850 \\ 247.500 \\ 171.460 \\ 186.480 \end{bmatrix} \dots (17)$$

となる。ここで、各装置に差がないと仮定し、 $u_1=u_2=u_3=0$ なる条件を代入すれば $[y_1, y_2] = [6.35776, 20.2050]$ が得られる。各装置に差があると仮定し、 $u_1=0$ なる条件を代入して解けば、

$$[y_1, y_2, u_1, u_2, u_3] = [8.13009, 23.1098, 0, -2.22624, -2.68622]$$

となる。この例に示す如く、境界条件が導入されるのは、逆行列演算の直前であり、境界条件の値を変えて、同様のデ-タを連続して演算する事も容易である。

[4] 参考文献

小林隆一 「相関・回帰分析入門」(QC テキストシリーズ NO 22) 日科技連