

研究発表 GMDH

中市新太

富山県立大学

u020025@st.pu-toyama.ac.jp

June 14, 2023

GMDH

GMDH についての論文を読み，理解を深める．

GMDH のプログラム

(未完成) 論文を読みながらプログラムを作る．

前提条件

図 1 の表の奇数部をトレーニングデータ，偶数部をチェックデータとし，GMDH によって Y を予測していく。

No.		X_1	X_2	X_3	X_4	Y
トレーニングデータ	1	1	-9	81	-729	-656
	3	1	-7	49	-343	-300
	5	1	-5	25	-125	-104
	7	1	-3	9	-27	-20
	9	1	-1	1	-1	0
	11	1	1	1	1	4
	13	1	3	9	27	40
	15	1	5	25	125	156
	17	1	7	49	343	400
19	1	9	81	729	820	
チェックデータ	2	1	-8	64	-515	-455
	4	1	-6	36	-216	-185
	6	1	-4	16	-64	-51
	8	1	-2	4	-8	-5
	10	1	0	0	0	1
	12	1	2	4	8	15
	14	1	4	16	64	85
	16	1	6	36	216	259
	18	1	8	64	512	585
20	1	10	100	1000	1111	

図 1: 前提条件

アルゴリズムの流れ

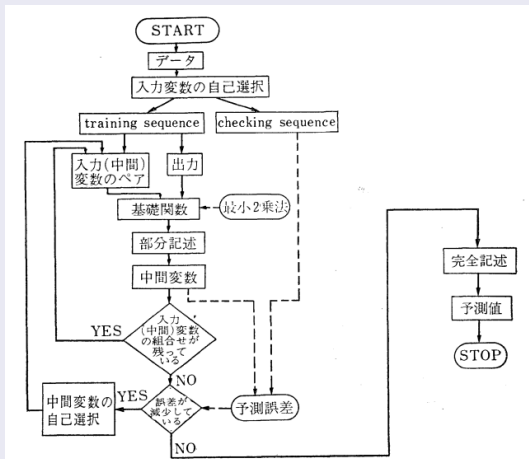


図 2: 前提条件

中間変数の作成

入力変数 $x_k = x_i, x_j$ に対して中間変数 z_k を作る.

$i = i + 1$ であり, $0 < i, j < N (=10)$ の範囲をとる. 入力変数の多項式は (1) 式とする.

$$z_k = a_0 + a_1 x_i + a_2 x_j + a_3 x_i x_j + a_4 x_i^2 + a_5 x_j^2 \quad (1)$$

(1) 式右辺の $a_l (l = 0, 1, 2, 3, 4, 5)$ は実データに基づく二乗平均誤差を最小にするよう最小二乗法を用いて求める.

二乗平均誤差の式は,

$$e_k^2 = \overline{(y - z_k)^2} \quad (2)$$

最小二乗法

e^2 を偏微分し、右辺を 0 とする方程式として極値を求める。

$$\begin{pmatrix} K_t & K_t \sum_{i=1}^t x_i & K_t \sum_{i=1}^t y_i & K_t \sum_{i=1}^t x_i y_i & K_t \sum_{i=1}^t x_i^2 & K_t \sum_{i=1}^t y_i^2 \\ K_t \sum_{i=1}^t x_i & K_t \sum_{i=1}^t x_i^2 & K_t \sum_{i=1}^t x_i y_i & K_t \sum_{i=1}^t x_i^2 y_i & K_t \sum_{i=1}^t x_i^3 & K_t \sum_{i=1}^t x_i^2 y_i \\ K_t \sum_{i=1}^t y_i & K_t \sum_{i=1}^t x_i y_i & K_t \sum_{i=1}^t y_i^2 & K_t \sum_{i=1}^t x_i y_i^2 & K_t \sum_{i=1}^t x_i^2 y_i & K_t \sum_{i=1}^t y_i^3 \\ K_t \sum_{i=1}^t x_i y_i & K_t \sum_{i=1}^t x_i^2 y_i & K_t \sum_{i=1}^t x_i y_i^2 & K_t \sum_{i=1}^t x_i^3 y_i & K_t \sum_{i=1}^t x_i^2 y_i^2 & K_t \sum_{i=1}^t x_i y_i^3 \\ K_t \sum_{i=1}^t x_i^2 & K_t \sum_{i=1}^t x_i^3 & K_t \sum_{i=1}^t x_i^2 y_i & K_t \sum_{i=1}^t x_i^3 y_i & K_t \sum_{i=1}^t x_i^4 & K_t \sum_{i=1}^t x_i^2 y_i^2 \\ K_t \sum_{i=1}^t y_i^2 & K_t \sum_{i=1}^t x_i y_i^2 & K_t \sum_{i=1}^t y_i^3 & K_t \sum_{i=1}^t x_i y_i^3 & K_t \sum_{i=1}^t x_i^2 y_i^2 & K_t \sum_{i=1}^t y_i^4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{pmatrix} = \begin{pmatrix} K_t \sum_{i=1}^t z_i \\ K_t \sum_{i=1}^t z_i x_i \\ K_t \sum_{i=1}^t z_i y_i \\ K_t \sum_{i=1}^t z_i x_i y_i \\ K_t \sum_{i=1}^t z_i x_i^2 \\ K_t \sum_{i=1}^t z_i y_i^2 \end{pmatrix}$$

図 3: 方程式

これを解くと、 a の値が予測でき、 $f(x_i, y_i) = z_i$ が求まる

繰り返し

$E_k < E_{k-1}$ のとき, e_m の小さい順に N 個対応する中間変数を選ぶ
前項で求めた中間変数を使用して次の中間変数を作成し, 同様に二乗平均
誤差を求める.

$$v_k = a_0 + a_1 z_i + a_2 z_j + a_3 z_i z_j + a_4 z_i^2 + a_5 z_j^2 \quad (3)$$

$$w_k = a_0 + a_1 v_i + a_2 v_j + a_3 v_i v_j + a_4 v_i^2 + a_5 v_j^2 \quad (4)$$

これを繰り返し, $E_k \geq E_{k-1}$ ならば繰り返しを止める.

今後の展開

GMDH についてより深く知る
プログラムを完成させ、結果を出力する
python での最小二乗法を勉強する.