非線形成分を考慮したMT 法（2 次元非線形 MT 法）の研究

Study of MT-Method to mitigate Non-linear Phonomena(2-Dimensional Non-linear MT-Method)

有限会社 増田技術事務所　増田 雪也

1. はじめに

MT 法は、「機械学習」の中の「分類問題」を扱うAI 手法である。近年急速に普及が進み、ロケットの打ち上げ判断アルゴリズム[1]の AI として採用されるなど、様々な分野での実績が多く報告されている。

分類問題を扱う同様な手法に「ディープラーニング」があるが、「膨大なデータ数」と「潤沢な CPU パワー」が必要となり、気軽に活用できる手法ではない。技術者が直面する分類問題では、データが豊富にあるケースは少なく、また、CPU パワーも潤沢に使用できるとは限らないため、ディープラーニングよりMT 法の方が活用場面が多いと考えられる。

MT 法の特徴は、項目（特徴量）のパターンを相関行列で特徴化し、マハラノビスの距離を求めることである。しかし、相関行列を作成する際、「項目間に非線形成分が存在すると、判別精度が悪化する」という問題点がある。

そこで本研究では、「非線形成分を考慮した T 法の研究」[2)3)]で報告した非線形成分を補正する方法を流用し、2 項目間（2 次元）の非線形成分を補正するMT 法（2 次元非線形 MT 法）を検討した。その結果、従来の MT 法に比べ、判別精度を向上させることができた。

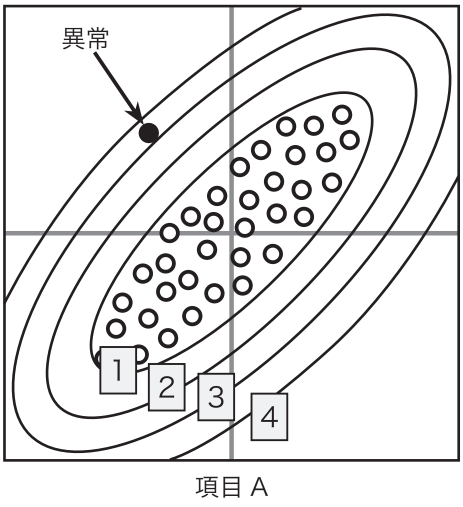


図１　マハラノビスの距離

Masuda Setsuya (info2qe@abox3.so-net.ne.jp)
Masuda Engineering Consultant Office,Inc.

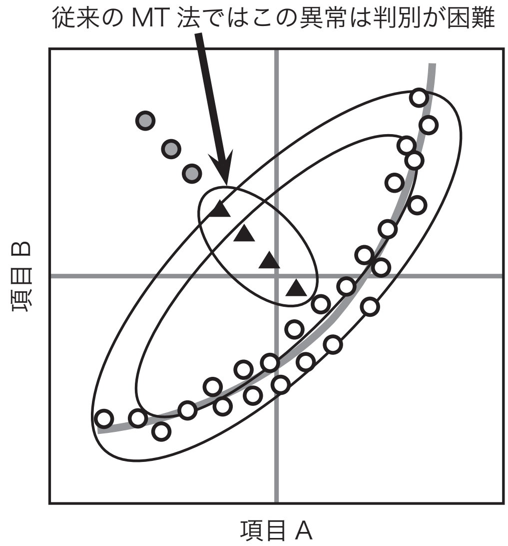


図２　項目間の非線形成分

2. 従来の MT 法の問題点

MT 法は、**図 1** に示すように、正常品（単位データ）の項目間の分布からマハラノビスの距離（以後、MD 値と呼称する）を算出する。項目間の分布は、非線形成分が無いことを前提としているため、**図 2** に示すように、項目間に非線形成分が存在すると、▲で示す異常品の MD 値は小さな値となり、【正常品】であると誤判別されてしまう。これが従来の MT 法の問題点である。

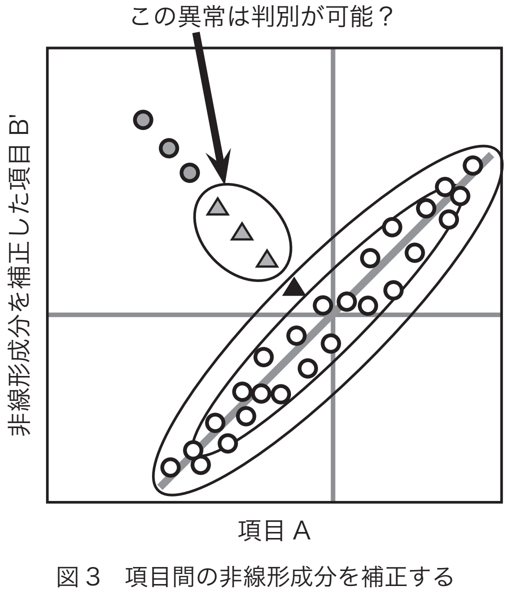


図３　項目間の非線形成分を補正する

表１　手順を説明するための数値データ

項目 A	2.69	4.62	8.85	12.88	14.62	23.46
項目 B	13.46	18.65	14.42	20.38	24.23	26.73

表２　数値データ（補正した項目 B'）

項目 A	2.69	4.62	8.85	12.88	14.62	23.46
補正した項目 B'	B_1'	B_2'	B_3'	B_4'	B_5'	B_6'

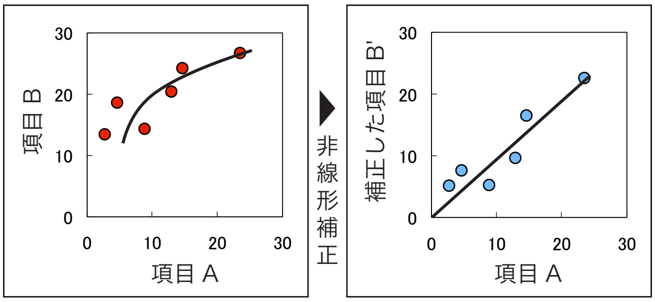


図４　非線形成分を補正する数値モデル

そこで本研究では、項目間の分布におけるこの非線形成分を**図３**に示すように補正できれば、MT 法の判別精度を向上させることができると考えた。次に、この非線形成分を補正する方法を述べる。

3. 非線形成分を考慮した MT 法

「非線形成分が強い場合、判別精度が下がる」という MT 法の問題点を解決するために、「非線形成分を考慮した T 法の研究」[2)3)] で報告した非線形成分を補正する方法を用いた。

3.1 非線形補正の手順

本研究で用いた非線形補正の手順は、以下の通りである。

1）項目 A と項目 B を２次の多項式で近似する

2）近似した多項式を用いて補正した項目 B' を求める

表１の数値データを用いて、非線形補正する手順を具体的に説明する。「非線形補正する」とは、**表２**に示すように「補正した項目 B' を求める」ことである。非線形補正のイメージをグラフにプロットすると、**図４**のようになる。ばらつき成分はそのままで、非線形成分のみを補正するのである。

手順 1)「項目 A と項目 B を２次の多項式で近似する」について述べる。まずは**表３**に示すように、項目 B の２乗の値 B^2 を求める。次に MS-Excel の「LINEST」関数を用いて、項目 A の多項式を求めると、

表３　数値データ（項目値を２乗）

項目 A	2.69	4.62	8.85	12.88	14.62	23.46
項目 B	13.46	18.65	14.42	20.38	24.23	26.73
項目 B^2	181.21	347.97	208.03	415.53	587.13	714.53

表４　数値データ（補正して求まった項目 B'）

項目 A	2.69	4.62	8.85	12.88	14.62	23.46
項目 B	13.46	18.65	14.42	20.38	24.23	26.73
補正した項目 B'	5.27	7.68	5.29	9.74	16.53	22.60

$$A = 0.104B^2 - 2.877B + 25.141$$

となる。

この多項式が、**図４**の左のグラフ上にある近似曲線である。

次に手順 2)「近似した多項式を用いて補正した項目 B' を求める」について述べる。

■補正した項目 B_1' の求め方

先ほど求めた多項式に「項目 B_1=13.46」を代入する。

$$\begin{aligned} A &= 0.104B_1^2 - 2.877B_1 + 25.141 \\ &= 0.104 \times 13.46^2 - 2.877 \times 13.46 + 25.141 \\ &= 5.27 \end{aligned}$$

ここで求めた A の値を補正した項目 B_1' とする。つまり、

$$B_1' = A$$

よって、

$$B_1' = 5.27$$

となる。

■補正した項目 B_2' の求め方

補正した項目 B_2' も同様に求める。先ほど求めた多項式に「項目 B_2=18.65」を代入する。

$$\begin{aligned} A &= 0.104B_2^2 - 2.877B_2 + 25.141 \\ &= 0.104 \times 18.65^2 - 2.877 \times 18.65 + 25.141 \\ &= 7.68 \end{aligned}$$

ここで求めた A の値を補正した項目 B_2' とする。つまり、

$$B_2' = A$$

よって、

$$B_2' = 7.68$$

となる。

このようにして、B_1' ～ B_6' まで求めると、**表４**のようになり、**図４**のイメージで示した補正となる。

3.2 非線形補正の手順【図解】

非線形補正する手順をより分かりやすくイメージできるように、手順の図解を**図５**に示す。

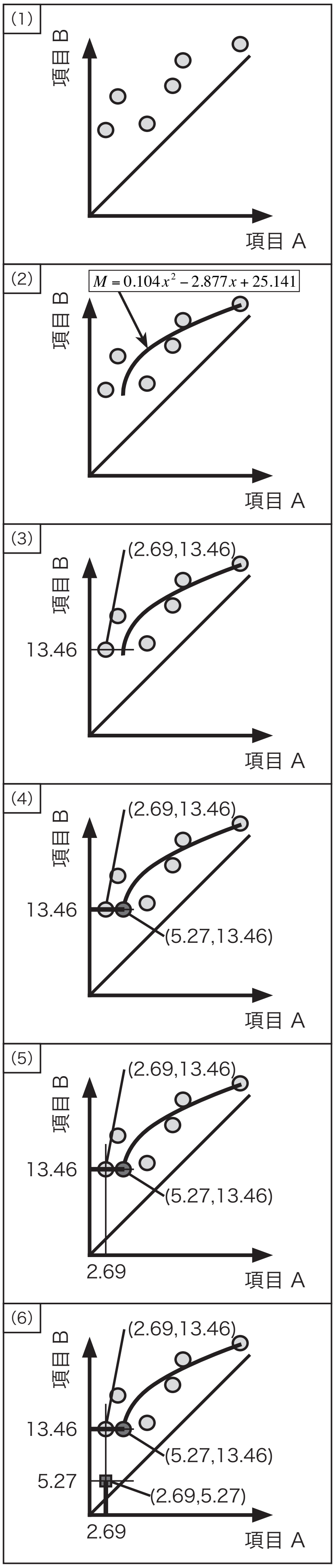


図 5　【図解】非線形成分を補正する手順

（1）**表 1** で示したデータをプロットする。

（2）MS-Excel の「LINEST」関数を用いて多項式を求める。

（3）$(A_1, B_1)=(2.69, 13.46)$ についての補正値 B_1' をこれから求めることにする。まずは $B=13.46$ の位置に水平にラインを描く。

（4）「$B=13.46$ のライン」と「先ほど求めた多項式」との交点座標を求めると (5.27,13.46) となる。

（5）$A=2.69$ の位置に垂直にラインを描く。

（6）$A=2.69$ のライン上の $B=5.27$ の座標 (2.69,5.27) が補正した項目 B_1' となる。

以上のような図解の手順で B_1' ～ B_6' を求めていく。

3.3 2 項目間の MD 値を算出する手順

図 6 に 2 項目間（項目 A と項目 B）のマハラノビスの距離を算出する手順を示す。なお、この正常品（単位データ）の項目間には非線形成分が存在している。**図 6** の左のグラフは、項目 A を横軸に、項目 B を縦軸にプロットしたものである。ここで、「3.1 非線形成分を補正する手順」にて説明した方法で、項目 B を非線形補正する。非線形補正した項目 B を項目 B' とする。**図 6** の右のグラフは、項目 A と項目 B' をプロットしたものである。正常品（単位データ）の非線形成分が補正され、線形に近い特性になっていることがわかる。

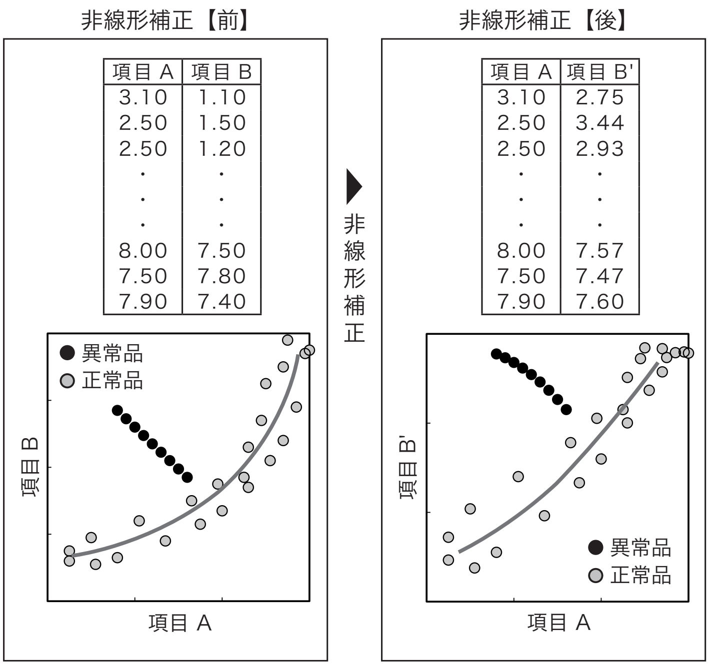
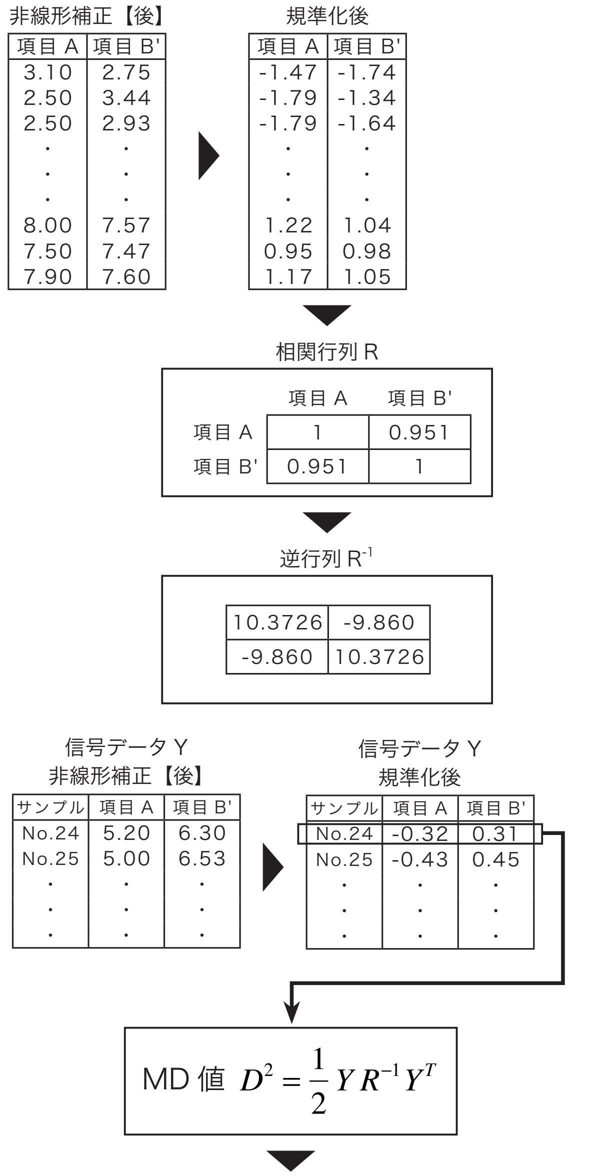


図 6　項目 B の非線形成分を補正する



$$D^2 = \frac{1}{2}\begin{bmatrix}-0.32 & 0.31\end{bmatrix} \times \begin{bmatrix}10.373 & -9.860 \\ -9.860 & 10.373\end{bmatrix} \times \begin{bmatrix}-0.32 \\ 0.31\end{bmatrix} = 2.033$$

図７　マハラノビスの距離を求める手順

図 7 にこの 2 項目間の MD 値を求める手順を示す。非線形補正した項目 A と項目 B' を規準化した後、2 次の相関行列 R を作成し、その逆行列 R^{-1} を求める。MT 法の数式により、信号データ Y のサンプル No.24 の項目 A と項目 B' の MD 値は、「2.033」と算出される。

図 8 にサンプル No.24 の「MD 値マトリクス」を示す。先ほど求めた項目 A と項目 B' の MD 値は 2.033 なので、該当する部分に MD 値を記入する。

次に、項目 B を横軸に、項目 A を非線形補正した項目 A' を縦軸にする。同様に、この項目間の MD 値 (=1.923) を求め、MD 値マトリクスに記入する。このようにして、全項目間について非線形補正処理と MD 値の算出を行い、MD 値マトリクスを完成させる。なお、項目 A 同士の組み合わせについても非線形補正するが、これは同じ数値同士なので非線形補正しても数値は変わらない。よって、求める MD 値は 1 次元のマハラノビスの距離となる。

	項目 A'	項目 B'	項目 C'	項目 D'	項目 E'	項目 F'
項目 A		2.033				
項目 B						
項目 C						
項目 D						
項目 E						
項目 F						

図８　サンプル No.24 の項目 A と項目 B' の MD 値

	項目A'	項目B'	項目C'	項目D'	項目E'	項目F'
項目 A	0.052	2.033	0.068	0.281	0.469	0.267
項目 B	1.923	0.007	0.233	0.319	0.138	0.156
項目 C	1.242	0.674	0.430	1.195	0.685	0.552
項目 D	0.406	0.527	0.711	0.406	0.407	0.481
項目 E	1.012	0.343	0.313	0.463	0.141	0.264
項目 F	1.329	1.013	0.175	0.342	0.562	0.156

図９　サンプル No.24 の MD 値マトリクス

	項目A'	項目B'	項目C'	項目D'	項目E'	項目F'
項目 A	0.094	3.913	0.22	0.791	0.389	0.819
項目 B	3.281	5E-07	0.058	0.443	0.019	0.910
項目 C	1.274	0.822	0.118	0.273	0.611	0.742
項目 D	0.875	0.399	0.973	0.395	0.749	0.407
項目 E	2.719	1.347	1.003	1.051	0.926	1.304
項目 F	1.007	0.683	0.378	0.385	1.354	0.085

図 10　サンプル No.25 の MD 値マトリクス

図 9 にサンプル No.24 の MD 値マトリクス、**図 10** にサンプル No.25 の MD 値マトリクスを示す。

3.4 サンプル毎の MD 値を算出する手順

項目数が 6 つの場合、このように各サンプルで 6 × 6 ＝ 36 個の MD 値が求まり、それが MD 値マトリクスとなる。本手法では、この MD 値マトリクスの中から、一番大きな MD 値をそのサンプルの代表 MD 値とする。サンプル No.24 では、項目 A と項目 B'（項目 B を非線形補正したもの）の組合せで最大値 =2.033 となり、これを代表 MD 値とする。サンプル No.25 では、最大値 =3.913 が代表 MD 値となる。

4. 非線形成分を考慮した MT 法の事例

非線形成分を補正する有効性を検証するため、システム開発ソフトウェア「LabVIEW」を用いて専用

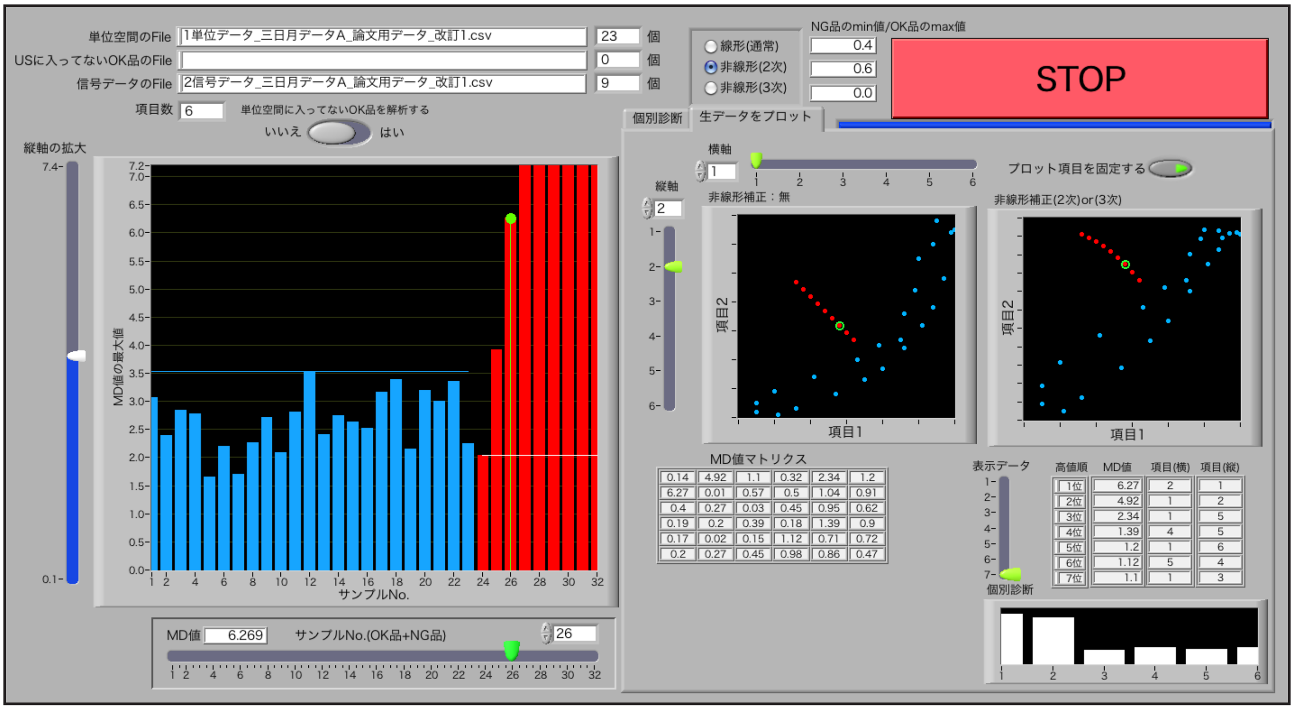


図 11　専用解析ソフト（LabVIEW によりプログラミング）

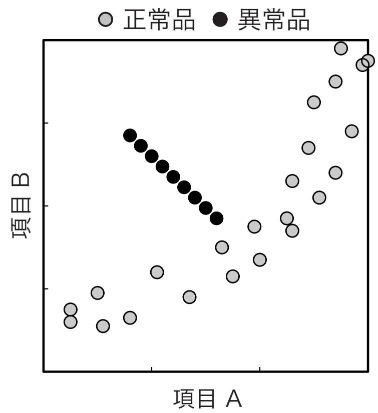
解析ソフトを作成した。**図 11** に専用解析ソフトの画面を示す。単位データおよび信号データは、CSV 形式のファイルで読み込み、MD 値を算出する。このソフトを使って実際の事例で解析した結果を以下に２つ紹介する。

4.1 事例 1「三日月データ」

事例 1 は、非線形補正の有効性を視覚的に分かりやすく検証するために、三日月型のダミーデータを作成した。**図 12** に示すように、全 6 項目の内、項目 A と項目 B をプロットすると、正常品は三日月型の分布となるようなデータとした。異常品は、三日月型の円弧から法線方向に伸びるように分布させた。残りの項目 C ～ F は、ランダムな分布となるようにデータを作成した。

図 13 に MD 値による全判別結果を示す。灰色の棒グラフは正常品、黒の棒グラフは異常品の MD 値を示している。非線形成分を補正した MT 法（右）の方が判別精度が高いことがわかる。

図 14 に判別結果の詳細を示す。非線形補正をしない MT 法（左）では、三日月型の分布の近くに位置している異常品 No.24 ～ 28 は、正常品の MD 値の max 値よりも MD 値が低く、判別不能となっている。異常品 No.29 でようやく判別可能領域となる。よって、異常品 No.24 ～ 28 の合計 5 個は判別不能という結果となった。



項目 A と B の関係において
非線形成分を意図的に作り出した

項目 A	項目 B	項目 C	項目 D	項目 E	項目 F
3.10	1.10	2.29	1.16	0.13	6.27
2.50	1.50	3.19	1.05	2.40	5.84
2.50	1.20	2.89	5.36	7.27	3.76
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
8.00	7.50	4.62	7.73	5.37	6.58
7.50	7.80	0.19	3.43	5.44	5.09
7.90	7.40	1.74	7.93	6.43	4.98

項目 C ～ F についてはランダムなデータとした

図 12　三日月データ

一方、非線形補正した MT 法（右）では、三日月型の分布は非線形補正することで線形な分布に補正されており、正常品にもっとも近い異常品 No.24 は

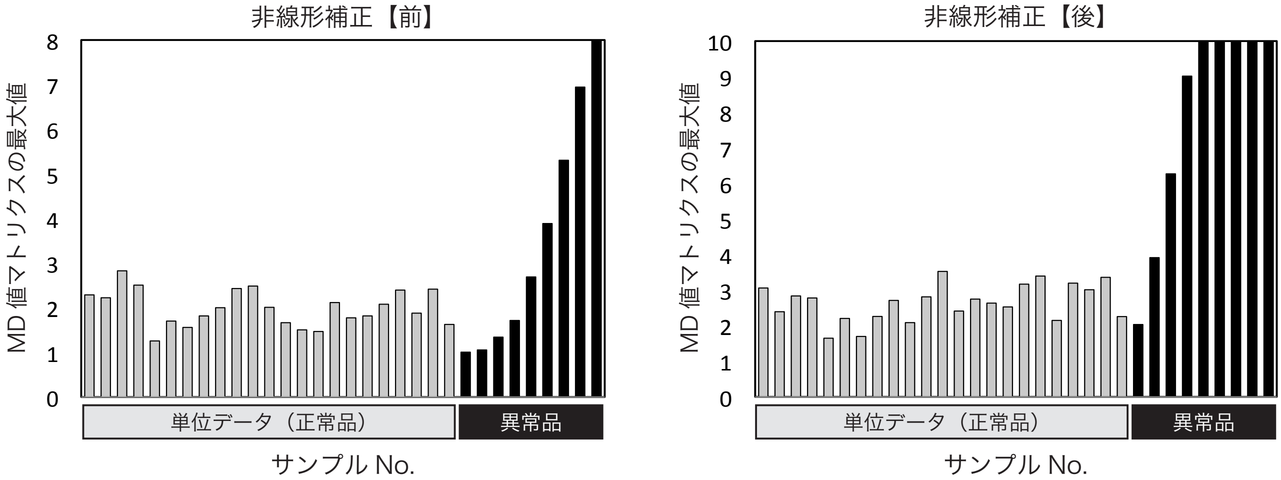


図 13　「三日月データ」の判別結果

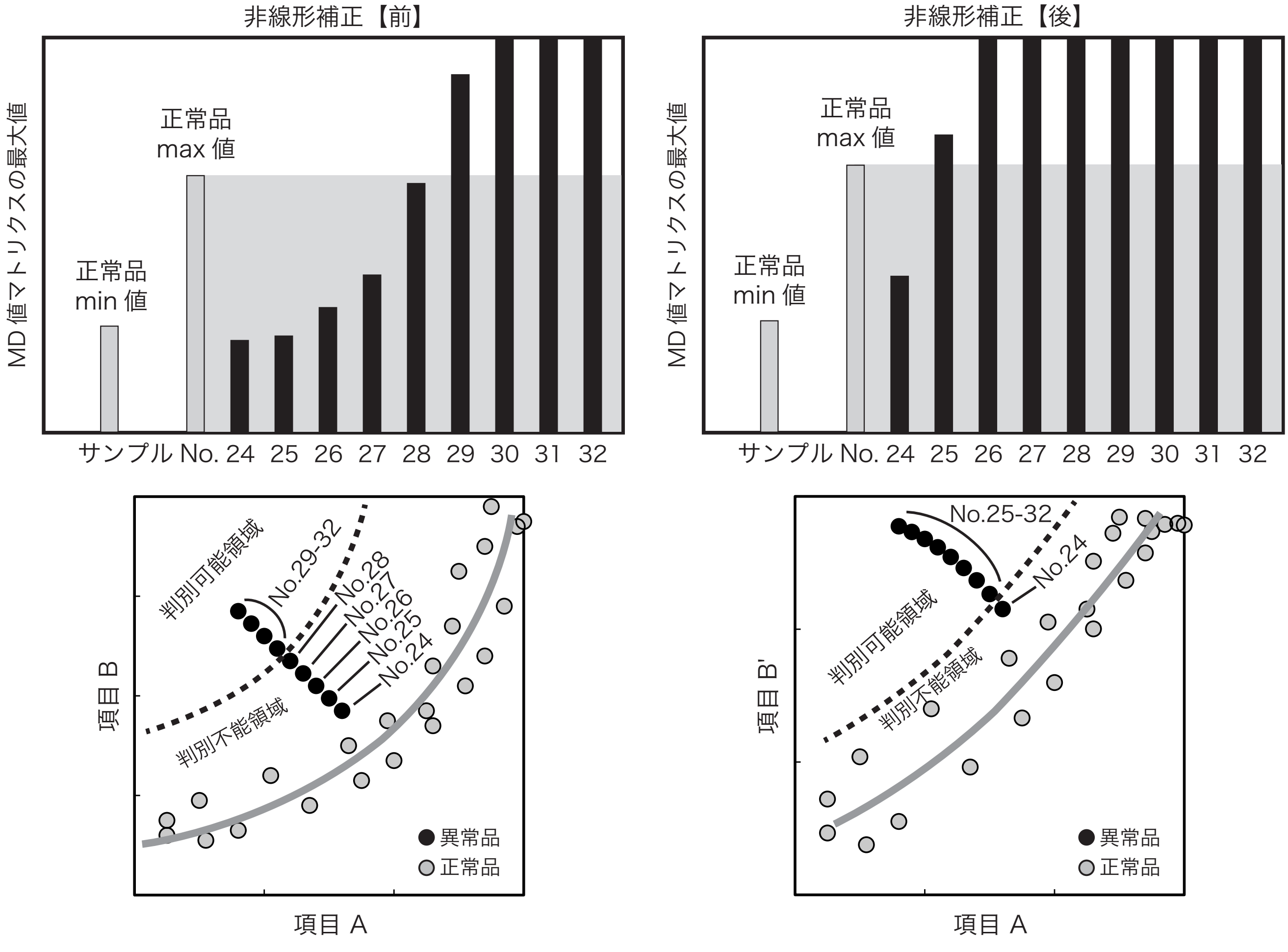


図 14　三日月データの判別結果（詳細）

判別不能であるが、異常品 No.25 以降では判別可能領域となっている。

以上まとめると、事例 1「三日月データ」では、非線形成分を補正したことにより、判別精度を向上させることが可能となった。

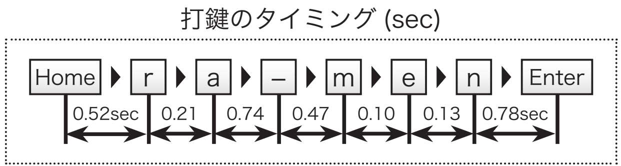


図 15　打鍵間隔の時間を項目に設定

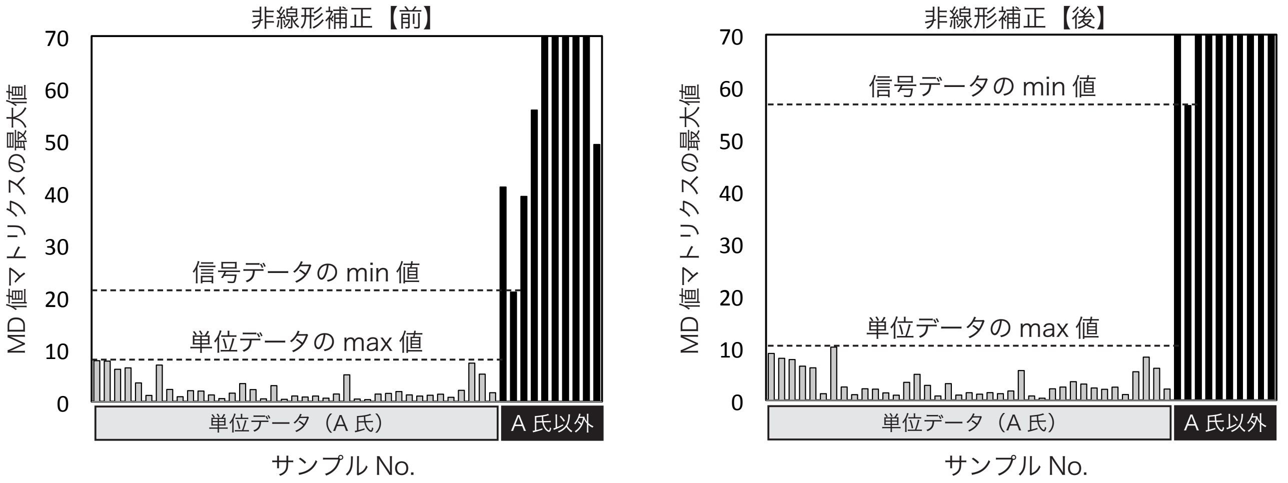


図 16　「キーボードの打鍵パターンから個人を判別したデータ」の判別結果

4.2 事例 2「キーボードの打鍵パターンから個人を判別したデータ」

事例 2 は、QES2008（第 16 回品質工学研究発表大会）で発表した「MT システム教育用教材の開発」[4] にて検討した事例を解析した。PC キーボードの打鍵タイミングから、個人を判別する事例である。**図 15** に示すように、「Home・r・a・-・m・e・n・Enter」の打鍵間隔の時間を 7 つの項目に設定し、各個人のパターンを解析 / 判別した事例である。**図 16** に、A 氏を単位空間として、A 氏以外の打鍵を判別できるかを試した結果を示す。灰色の棒グラフは A 氏、黒の棒グラフは A 氏以外である。非線形補正しない MT 法（左）では、判別可能ではあるが、単位データ（A 氏）の max 値と信号データ（A 氏以外）の min 値の差は小さい。一方、非線形補正した MT 法（右）では、単位データ（A 氏）の max 値と信号データ（A 氏以外）の min 値の差が大きく、完全に判別が可能となっている。

以上まとめると、事例 2「キーボードの打鍵パターンから個人を判別したデータ」においても、非線形成分を補正したことにより、判別精度を向上させることができた。

5. その他の検討

5.1 原因診断のやり方

通常の MT 法では直交表を用いて原因診断を行うが、本手法では**図 17** に示すように、MD 値マトリクスの各項目の平均値を計算することで原因診断を行う。直交表を用いないため、簡単に原因の診断ができるメリットがある。また、全ての 2 項目間の組合せで原因診断を行うため、項目間に交互作用が存在しても正しい原因診断結果が得られるというメリットもある。

	項目A'	項目B'	項目C'	項目D'	項目E'	項目F'	
項目 A	0.052	2.033	0.068	0.281	0.469	0.267	→ 項目 A の平均値を算出
項目 B	1.923	0.007	0.233	0.319	0.138	0.156	→ 項目 B の平均値を算出
項目 C	1.242	0.674	0.430	1.195	0.685	0.552	→ 項目 C の平均値を算出
項目 D	0.406	0.527	0.711	0.406	0.407	0.481	→ 項目 D の平均値を算出
項目 E	1.012	0.343	0.313	0.463	0.141	0.264	→ 項目 E の平均値を算出
項目 F	1.329	1.013	0.175	0.342	0.562	0.156	→ 項目 F の平均値を算出

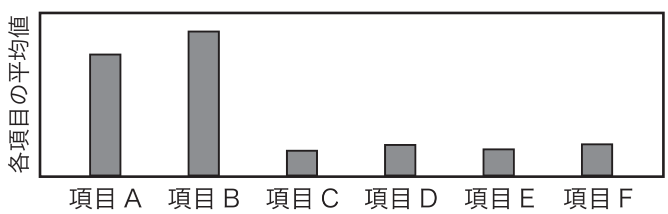


図 17　原因診断のやり方

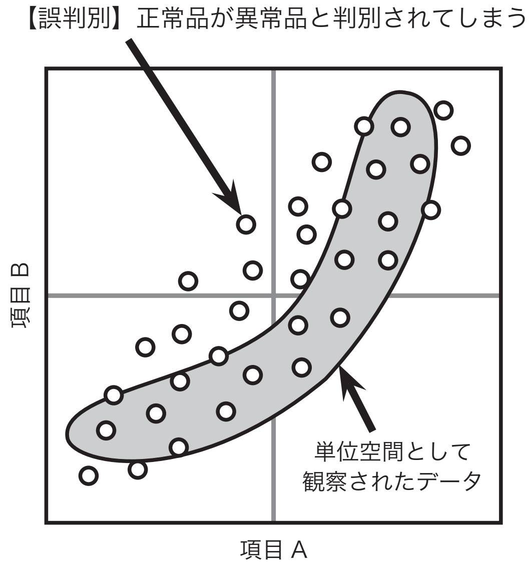


図 18　判別精度が悪化するケース

5.2 判別精度が悪化するケース

非線形成分が弱い場合は、非線形補正する効果は低く、この方法を導入するメリットは小さい。しかし、非線形成分が弱いケースに適用したとしても、判別精度が悪化することは無いため、どのような事例でも安心して非線形補正することができる。

ただし、**図 18** に示すように、本来は非線形成分が小さい特性にも関わらず、観測されたデータが偶然にも非線形成分を持っていた場合は、逆に判別精度を悪化させてしまう。そのような場合には、「非線形補正する」「非線形補正しない」の両方を試し、判別精度の高い方を採用すれば良い。

5.3 n 次元のパターン認識

本手法では、全２項目間（２次元）のパターン認識を行っているため、「項目数が単位データ数よりも少ない場合でも適用可能」という副次的なメリットがある。このこと自体はメリットであるが、通常の MT 法では最大 n 次元のパターンを認識できるのに対して、本手法では最大２次元のパターンしか認識できない。よって、２次元より大きなパターンを持つ場合は、判別精度が落ちるというデメリットがある。

6. さいごに

今回の研究では、MT 法の判別精度を向上させるため、非線形成分を補正する方法を検討した。その結果、判別精度を向上させること可能となった。

非線形成分が存在しない場合でも気軽に便利に使える手法であり、ぜひいろいろな事例に活用していただきたい。

謝 辞

非線形 MT 法の閃きのきっかけ及び数式に関する有益な助言をいただいた日本電産サンキョー（株）の中西徹氏及び長野県品質工学研究会のメンバーに感謝の意を表する。

■参考文献

1) 手島昌一：MT 法の特質と大規模機器監視への適用 , 標準化と品質管理 ,Vol.70,No.7,pp.16-18,2017
2) 増田雪也：非線形成分を考慮した T 法の研究 , 第 17 回品質工学会研究発表大会論文集 ,(2009),pp.422-425
3) 増田雪也：非線形成分を考慮した T 法の研究 , 増田技術事務所技報 ,Vol.1,2008 (http://www002.upp.so-net.ne.jp/sbux/pdf/paper_QE_TMethod_nonlinear.pdf)
4) 増田雪也：MT システム教育用教材の開発 , 第 16 回品質工学会研究発表大会論文集 ,(2008),pp.298-301