

多変量解析手法を用いた異常検出

本研究では、従来から提案されている方法よりも優れた、多変量解析手法を用いた異常検出方法を提案することを目的としている。従来法では、正常運転時に測定データが分布する領域を求めておいて、逐次測定される運転データがその領域から逸脱しているかどうかに着目する。このため、変数間の関係が変化したとしても、測定データが求めた領域内に存在する限り、その変化を検出することはできない。変数間の関係の変化そのものを監視するためには、主成分が変数間の関係を表現するという特徴を利用して、逐次主成分を計算し、主成分の変化を監視すればよい。これが、提案する第1の異常検出方法の基本概念である。主成分の決定に利用するデータを移動させながら、逐次主成分分析を行うため、この方法を移動主成分分析 (MPCA; Moving Principal Component Analysis) による異常検出と呼ぶ。また、データの分布が運転状態を代表していることに着目すると、基準となる正常運転時のデータの分布と各時刻における運転状態を表すデータの分布との違いを定量的に評価する指標が導入できれば、主成分そのものを利用することなく、その指標に基づく異常検出が可能となる。これが、提案する第2の異常検出方法の基本概念である。データの分布の違いを定量的に評価するために非類似度 (Dissimilarity Index) を導入するため、この方法を DISSIM と呼ぶ。

異常検出は通常、予め設定した管理限界を監視指標が超えるか否かによって判断される。主成分分析を利用する従来法では、主成分で張られる空間とその補空間とを区別し、主成分得点から計算される T^2 と残差 (予測誤差) から計算される Q という2つの指標が同時に監視される。一方、MPCA では、正常運転状態を表す基準主成分と現時刻の運転状態を表す主成分の違いを監視指標として導入する。ただし、複数の主成分に関して、その主成分得点の分散が概ね等しくなる場合には、正常運転時においても主成分が激しく変化するため、複数の主成分が張る線形部分空間の違いを監視指標とする方がよい。また、DISSIM では、非類似度が唯一の監視指標となる。管理限界の設定に際して指標が従う確率分布を仮定する場合も多いが、本研究では、正常運転時における

誤検出の確率が1%となるように、運転データに基づいて管理限界を定める。

提案する異常検出方法の有効性を検証するために、Tennessee Eastman プロセス^[1]への適用を試みた。シミュレータには予め20種類の外乱あるいは異常が設定されており、これらの外乱あるいは異常を検出することを目的とする。異常検出方法を評価するために、異常発生後の特定時間内に指標が管理限界を超える割合として定義される信頼度を用いる。誤検出の確率が1%となるように管理限界を設定したため、信頼度が大きいほど良い異常検出方法と判断できる。10回の異なるシミュレーションを実行し、その平均値を求めた結果を Table 1 に示す。この表において、cMSPC は従来法を表し、各ケース番号 (IDV) は論文で使用されている番号と一致している。また、指標の添え字は採用した主成分数であり、 A_{1-11} は第1から第11主成分で張られる空間を監視することを意味する。この表に示した以外の多くのケースについても、提案する異常検出方法の信頼度は従来法と同等かそれ以上であった。この結果から、提案法が非常に優れていることがわかる。ただし、提案法の異常検出性能は、指標を計算する際に用いるサンプル数 (窓幅) や採用する主成分数に依存することに注意する必要がある。Table 1 に示した結果は静的監視のものであるが、過去の測定値も監視変数として利用する動的監視を行うことにより、さらに異常検出性能を向上させることが可能である。また、本研究で提案した手法は、異常検出のみならず異常診断にも利用することができる。

Table 1 Comparison of monitoring methods by reliability (%).

Case (IDV)		3	9	11	14	16
Method	Index					
cMSPC	T_{11}^2	1.7	3.1	0.8	6.2	9.0
	Q_{11}	18.9	14.9	20.4	70.8	21.9
MPCA	A_{1-11}	41.7	41.0	71.4	97.5	43.3
DISSIM	D	58.1	28.7	49.4	92.5	60.9

[1] Downs, J. J. and E. F. Vogel, *Computers chem. Engng*, **17**, 245-255 (1993)