

適応型非線型回帰分析手法の開発およびソフトセンサーへの応用

○金子 弘昌、船津 公人

東京大学大学院工学系研究科(〒113-8656 東京都文京区本郷 7-3-1)

【緒言】

化学プラントにおいては、測定困難なプロセス変数を推定する手法としてソフトセンサーが広く用いられている。ソフトセンサーとは、オンラインで測定可能な変数と測定困難な変数の間で数値モデルを構築し、目的とした変数 y の値を推定する方法である。ソフトセンサーを用いることで、オンラインで精度良く y を推定できる。

しかしソフトセンサーは、プラントの運転状態の変化や触媒性能の変化、機器や配管への汚れ付着などによって予測精度が劣化してしまう。このソフトセンサーの劣化に対応するため、予測データに時間的に最も近いデータセットを用いて構築される **moving window (MW)** モデル、予測データと類似したデータセットを用いて構築される **just-in-time (JIT)** モデル、各プロセス変数の時間差分 (**time difference, TD**) の間で構築される **TD** モデルという適応型モデルが開発されている。あるプロセスのすべての状態において万能に使用可能な適応型モデルは存在せず、プラントの状態によって各モデルの精度の優劣は変化する。ソフトセンサーモデルの劣化の種類に基づき各適応型ソフトセンサー手法の特徴は整理され、 X と y の傾き、つまりプロセスの特性が急激に変化した場合に各適応型モデルは対応できないことが明らかになった [1]。

そこで本研究では傾きが急激に変化する劣化における予測精度の向上を目的とした。傾きの急激な変化とは傾きの非線型変化であることから非線型の回帰モデルの更新を目指す。しかし非線型回帰モデルを逐次更新するには多くの時間がかかってしまう。そこで非線型回帰モデルの一つである **support vector regression (SVR)** [2] モデルを効率的に更新する手法である **online SVR (OSVR)** 法 [3] をソフトセンサーに応用する。さらに、たとえ非線型回帰モデルを更新したとしても時間的にプロセス特性が変化する際の対応は困難と考え、時間を表現する変数を X に追加する。

本手法の有効性を確認するため、 X と y の関係が時間的に変化する場合を想定したシミュレーションデータおよびポリマー重合プラントを対象にした実際の運転データを用いた解析を行う。各データ解析において従来の適応型ソフトセンサー手法と予測性能を比較、検討する。

【手法】

OSVR 法とは、**SVR** モデルが満たすべき **Karush-Kuhn Tucker** 条件について、データが追加および削除された際にも成立するよう効率的に **SVR** モデルを更新する手法である。詳細なアルゴリズムは文献 [3] を参照されたい。**OSVR** モデルでは **SVR** モデルと同様に3つのパラメータを事前に決定する必要がある。今回はパラメータ決定の方法として、時間を基準にしてモデル構築用データを2つに分けた後に後半のデータの予測性能が高くなるよう決定する方法 (**OpV**) およびモデル構築用データを用いた **5-fold cross validation** により最適化された **SVR** モデルのパラメータを使用する方法 (**CrV**) を検討する。**OpV** はモデルの予測性能を考慮してパラメータを決定できる反面、更新回数が多いためパラメータの決定に時間がかかってしまう。一方 **CrV** では **SVR** モデル構築時と同じ比較的短時間でパラメータの決定が可能である。

さらに本研究では X に時間を表わす変数を追加する。時間変数により直近のプロセス特性の時間的変化をモデル化することで、将来の y の値を適切に予測できると考えられる。また時間変数を追加した非線型モデルを更新することで時間的に非線型なプロセス特性の変化に対応可能である。なお、時間変数は間隔尺度であり任意にゼロ点を設定できる。

【結果と考察】

X と y の間の関係が時間的に変化するシミュレーションデータを用いた解析を行った。比較を行ったモデルは、**TD** モデルとして **partial least squares (PLS)** 法により構築された **TD** モデル (**TDPLS**) および **SVR** 法を用いた非線型 **TD** モデル (**TDSVR**)、**JIT** モデルとして予測用データとユークリッド距離の近いデータのみで構築され

た PLS モデル (JITPLS)、MW モデルとして予測用データに時間的に近いデータのみを用いて構築された PLS モデル (MWPLS) および OSVR である。それぞれ時間変数を追加した場合についても検討を行った。

今回の X を 2 変数の一様乱数とした。y は X の線形結合で表わされるとし、X の 1 変数の係数を 1 に固定してもう 1 変数の係数を時間の sin 関数で変化させた。モデル構築用データ数とモデル検証用データ数はともに 100 である。JITPLS、MWPLS、OSVR におけるモデル構築用データ数を 50 とした。

モデル検証用データにおける決定係数 r_p^2 の値を表 1 に示す。 r_p^2 の値が 1 に近いほどモデルの予測性能は高い。表 1 より時間変数を追加しない場合はどの適応型モデルを用いても r_p^2 値が負になってしまい予測精度は低かった。時間変数を追加しても従来法では r_p^2 値は小さいままである。一方 OSVR と時間変数を組み合わせた提案手法により r_p^2 値がほぼ 1 となり予測精度の大幅な向上が達成された。また OpV と CrV の間でほとんど差異はなく CrV により短時間で適切なパラメータを選択可能であることを確認した。

図 1 にモデル検証用データの y の実測値と予測値のプロットを示す。従来法である時間変数を加える前の MWPLS および OSVR においては対角線から外れるデータが散見されたのに対し、提案手法を用いることでデータが対角線付近に固まり精度良く予測できていることが分かる。時間変数を X に追加して非線型モデルを更新することでプロセス特性の時間変化に適切に追従可能であることを確認した。

実プラントデータを用いた解析結果は当日報告する。

表 1. モデル検証用データにおける r_p^2

モデル	時間変数なし	時間変数あり
TDPLS	-6.1	-6.1
TDSVR	-9.9	-19.2
JITPLS	-6.0	-0.25
MWPLS	-1.3	0.147
OSVR, OpV	-1.1	0.995
OSVR, CrV	-1.5	0.997

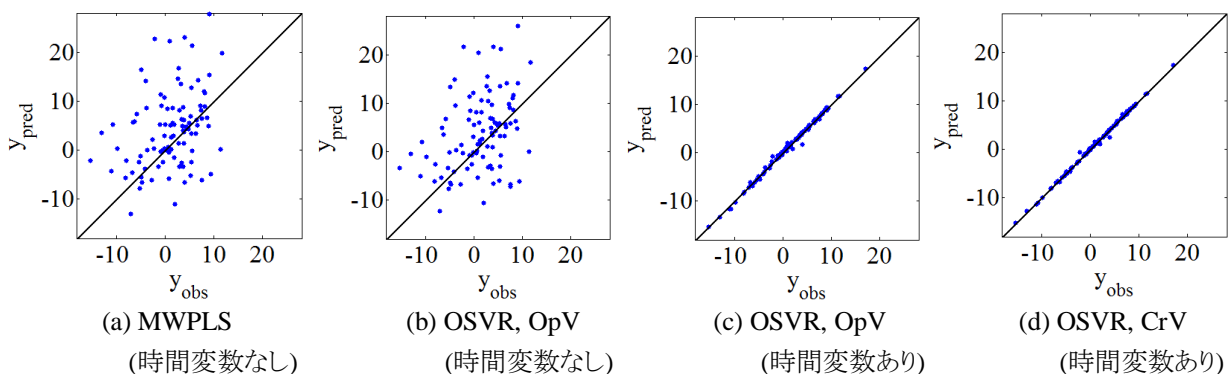


図 1. モデル検証用データにおける y の実測値と予測値の関係

【結言】

本研究では、時間的にプロセス特性が変化するモデル劣化に対して適切に対応するため時間変数と OSVR 法を組み合わせた新規なソフトセンサー手法を提案した。データ解析を通して提案手法によりプロセス特性の時間変化に適切に追従して精度良く予測できることを確認した。また CrV により短時間で OSVR パラメータを決定可能であることを示した。本手法による効率かつ安定的なプラント管理の実現を期待する。

【謝辞】

本研究の一部は金子弘昌が科研費の若手研究(B)(No. 24760629)および財団法人みずほ学術振興財団の助成を受けたものである。ここに謝意を示す。データ提供の面で多大なご支援をいただいた三井化学株式会社に感謝する。

【参考文献】

- [1] Kaneko H, Funatsu K. *AIChE J.*, in press.
- [2] Vapnik V.N. *The Nature of Statistical Learning Theory, second ed.*, Springer, New York (1999).
- [3] Ma J, Theliler J, Perkins S. *Neural Comput.* 2003;15:2683–2703.