Chemometrics and Intelligent Laboratory Systems 150 (2016) 9–22

Contents lists available at ScienceDirect

Chemometrics and Intelligent Laboratory Systems

journal homepage: www.elsevier.com/locate/chemolab

Quality prediction and quality-relevant monitoring with multilinear PLS for batch processes

Lijia Luo *, Shiyi Bao, Jianfeng Mao, Di Tang

College of Mechanical Engineering, Zhejiang University of Technology, Engineering Research Center of Process Equipment and Remanufacturing, Ministry of Education, Hangzhou, China

ARTICLE INFO

Article history:
Received 22 June 2015
Received in revised form 26 October 2015
Accepted 4 November 2015
Available online 10 November 2015

Keywords:
Tensor
Multilinear PLS
Quality-relevant monitoring
Quality prediction
Batch processes

ABSTRACT

The multilinear regression method is applied for quality prediction and quality-relevant monitoring in batch processes. Four multilinear partial least squares (PLS) models are investigated, including three higher-order PLS (HOPLS) models, termed as HOPLS-Tucker, HOPLS-RTucker and HOPLS-CP, and the *N*-way PLS (N-PLS) model. These multilinear PLS methods have two advantages as compared to the unfold-PLS method. Firstly, they retain the inherent three-way representation of batch data and avoid the disadvantages caused by data unfolding, resulting in more stable process models. Secondly, they summarize the main information on each mode of data and describe the three-way interactions between them, and therefore have better modeling accuracy and intuitive interpretability. Online quality prediction and quality-relevant monitoring methods are developed by combining multilinear PLS with the moving data window technique. These methods are tested in a fed-batch penicillin fermentation process. The results indicate that the multilinear PLS method has higher predictive accuracy, better anti-noise capability and monitoring performance than the unfold-PLS method.

1. Introduction

Batch process is widely used to produce low volume and high value-added products, including polymers, chemicals, pharmaceuticals and semiconductors, because of its high flexibility to adapt the rapidly changing market situations. The increasing international competition has aroused the demand for high quality products. Reliable process monitoring as well as precise quality prediction and control are therefore necessary to ensure the safe and profitable operation of batch processes. The process monitoring aims to warn abnormal operating conditions by timely detecting and diagnosing process faults, which is helpful for taking corrective actions to recover the normal production and enhance the process safety. The objective of quality prediction and control is to estimate the product quality from operating conditions in a fast and accurate way, and further to derive better operating conditions that can reach a high product quality. However, process monitoring and quality prediction are difficult for batch processes, because they suffer not only from those intractable nonlinear and time-varying process behaviors, but also from the unavailable online quality measurements, batch-to-batch variations, and so on.

In last decades, data-driven process monitoring and control methods have been widely accepted in various industrials, benefiting from the widespread application of automation instrument and data acquisition technologies. Data-driven methods build process models using operation data and rarely require the knowledge of process mechanism. Therefore, they are especially applicable for those complicated industrial processes, in which precise first-principle models are not available or difficult to be built. Different from continuous processes, operation data of a batch process are usually recorded in a three-way data array with three directions of batch, variable and sampling time. Bilinear analysis methods and multilinear analysis methods are often adopted to analyze this three-way data array. The bilinear analysis method is also known as the unfolding method, including multiway principal component analysis (MPCA) [1], multiway partial least squares (MPLS) [2], multiway independent component analysis (MICA) [3], and so on. These methods unfold the three-way data array to a matrix along the batch-wise or variable-wise direction firstly, and then build process models based on the unfolded data. However, multilinear analysis methods maintain the three-way representation of the data array and build process models using tensor decompositions, such as parallel factor analysis (PARAFAC) [4] and Tucker3 decomposition [5].

So far, bilinear methods, especially MPCA and MPLS, have attracted more research attention than multilinear methods in process control applications. Lots of MPCA/MPLS-based process monitoring and quality prediction methods have been developed. For example, Lee et al. [6] proposed a multiway kernel PCA method for monitoring nonlinear batch processes. Lu and Gao [7] and Zhao et al. [8] presented two phase-based PLS methods for the quality prediction of multiphase batch processes, which take the multiphase feature of batch processes into account. Yu [9] proposed a multi-way Gaussian mixture model based adaptive kernel partial least squares (MGMM-AKPLS) method to predict quality variables in nonlinear batch processes. As mentioned above, the MPCA/MPLS-based methods need to unfold the three-way

* Corresponding author. Tel.: +86 0571 88320349.
E-mail address: lijialuo@zjut.edu.cn (L. Luo).

http://dx.doi.org/10.1016/j.chemolab.2015.11.004
0169-7439/

batch data before building process models. However, two potential disadvantages are caused by the data unfolding operation. Firstly, the unfolded data may have the "large variable number but small sample size" problem. For instance, when a three-way array $\underline{\boldsymbol{X}}(10\times10\times100)$ containing 10 batches, 10 variables and 100 time slices is unfolded to a matrix $\boldsymbol{X}(10\times1000)$, only ten samples are available for PCA or PLS to compute the 1000-dimensional loading vectors, probably resulting in an unreliable estimate of model parameters [10]. Secondly, MPLS and MPCA fail to offer an explicit description of the three-way interactions in the batch dataset, because the three-way data structure is destroyed and the information from two directions is convolved together after data unfolding. These two disadvantages may reduce the monitoring performance, prediction ability and interpretability of MPCA/MPLS models.

In recent years, researchers have started to analyze batch data with multilinear (i.e., tensor) analysis methods, where batch data are represented in their natural three-way form as tensors. For example, Meng et al. [4] and Louwerse and Smilde [5] applied PARAFAC and Tucker decomposition for batch process monitoring, respectively. Luo et al. [11] proposed a generalized Tucker2 (GTucker2) model for monitoring uneven-length batch processes. Since the batch data array is inherently three-way, using multilinear analysis methods can naturally avoid the disadvantages caused by data unfolding. Multilinear models commonly have much fewer parameters than MPCA/MPLS models, because batch data are compressed in three directions instead of two. These multilinear models are therefore expected to be more stable [5]. Besides, multilinear models can summarize all main effects and interactions in batch data, because they extract the main information on each mode of data and describe the relations between them [12]. Thus, multilinear models have better modeling accuracy and intuitive interpretability. Particularly, two multilinear regression algorithms have been proposed, i.e., higher-order partial least squares (HOPLS) [10] and N-way partial least squares (N-PLS) [13]. HOPLS and N-PLS both have some advantages as compared to unfold-PLS (MPLS), including robustness to noise, stabilized solution, increased predictability and intuitive interpretability [10,13]. These multilinear regression methods may have good application prospects in batch processes.

In this paper, the multilinear regression method is applied for quality prediction and quality-relevant monitoring in batch processes. Four multilinear PLS models, including HOPLS-Tucker (i.e., HOPLS), HOPLS-RTucker, HOPLS-CP and N-PLS, are investigated. Especially, HOPLS-RTucker and HOPLS-CP are two new multilinear PLS models, which are derived by imposing extra restrictions on HOPLS-Tucker. HOPLS-RTucker aims to handle the multi-way data with a specific (or free) mode, on which the information remain unchanged. HOPLS-CP replaces the Tucker decomposition used in HOPLS-Tucker with the CP decomposition (i.e., canonical decomposition/parallel factor analysis) [14] to simplify the model and improve the computational efficiency. Online quality prediction and quality-relevant monitoring methods are then proposed by combining multilinear PLS with the moving data window technique. Monitoring statistics named T^2, Q_x and Q_y are constructed for fault detection, and the contribution plot is applied for fault diagnosis. These methods are tested in a fed-batch penicillin fermentation process. The results show that multilinear PLS outperforms MPLS (i.e., unfold-PLS) in terms of higher predictive accuracy, better anti-noise capability, higher fault detection rates and lower false alarm rate.

2. Notation and preliminaries

2.1. Notation and definitions

Tensors (multi-way arrays), matrices and vectors are denoted by underlined boldface capital letters, boldface capital letters and boldface lowercase letters respectively, e.g., $\underline{\boldsymbol{X}}$, $\boldsymbol{X}$ and $\boldsymbol{x}$. The order of a tensor is the number of ways or modes. The ith entry of a vector $\boldsymbol{x}$ is x_i. The ith column of a matrix $\boldsymbol{X}$ is $\boldsymbol{x}_i$ and the element (i, j) is x_{ij}. The element $(i_1, i_2, \ldots, i_N)$ of an Nth-order tensor $\underline{\boldsymbol{X}}\in\Re^{I_1\times I_2\times\cdots\times I_N}$ is $x_{i_1 i_2\cdots i_N}$. $\boldsymbol{X}^{(n)}$ and $\boldsymbol{x}^{(n)}$ denote the nth factor matrix and factor vector in a sequence, respectively. $\underline{\boldsymbol{X}}_{(n)}$ is the mode-n matricization of a tensor $\underline{\boldsymbol{X}}$ [14]. Symbols "$\circ$", "$\otimes$" and "$\oplus$" denote the vector outer product, Kronecker product and Khatri–Rao product, respectively. The superscript "+" denotes the Moore–Penrose pseudoinverse.

The norm of an Nth-order tensor $\underline{\boldsymbol{X}}\in\Re^{I_1\times I_2\times\cdots\times I_N}$ is $\|\underline{\boldsymbol{X}}\| = \sqrt{\sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\cdots\sum_{i_N=1}^{I_N} x_{i_1 i_2\cdots i_N}^2}$. The n-mode product of a tensor $\underline{\boldsymbol{X}}\in\Re^{I_1\times\cdots\times I_n\times\cdots\times I_N}$ with a matrix $\boldsymbol{A}\in\Re^{J\times I_n}$ or with a vector $\boldsymbol{t}\in\Re^{I_n}$ is $\underline{\boldsymbol{Y}}=\underline{\boldsymbol{X}}\times_n\boldsymbol{A}\in\Re^{I_1\times\cdots\times I_{n-1}\times J\times I_{n+1}\times\cdots\times I_N}$ with $y_{i_1\cdots i_{n-1} j i_{n+1}\cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1\cdots i_n\cdots i_N} a_{j i_n}$ or $\underline{\boldsymbol{Y}}=\underline{\boldsymbol{X}}\times_n\boldsymbol{t}\in\Re^{I_1\times\cdots\times I_{n-1}\times I_{n+1}\times\cdots\times I_N}$ with $y_{i_1\cdots i_{n-1} i_{n+1}\cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1\cdots i_n\cdots i_N} t_{i_n}$ [14,15]. For two tensors $\underline{\boldsymbol{X}}\in\Re^{I_1\times\cdots\times I_n\times\cdots\times I_N}$ and $\underline{\boldsymbol{Y}}\in\Re^{J_1\times\cdots\times I_n\times\cdots\times J_M}$ with the same size on the nth-mode, their n-mode cross-covariance is $\underline{\boldsymbol{Z}}=\langle\underline{\boldsymbol{X}},\underline{\boldsymbol{Y}}\rangle_{\{n;n\}}\in\Re^{I_1\times\cdots\times I_{n-1}\times I_{n+1}\times\cdots\times I_N\times J_1\times\cdots\times J_{n-1}\times J_{n+1}\times\cdots\times J_M}$ with $z_{i_1\cdots i_{n-1} i_{n+1}\cdots i_N j_1\cdots j_{n-1} j_{n+1}\cdots j_M} = \sum_{i_n=1}^{I_n} x_{i_1\cdots i_n\cdots i_N} y_{j_1\cdots i_n\cdots j_M}$.

The Tucker decomposition of a tensor $\underline{\boldsymbol{X}}\in\Re^{I_1\times\cdots\times I_n\times\cdots\times I_N}$ is concisely expressed as [14,15]

$$\underline{\boldsymbol{X}}\approx\underline{\boldsymbol{G}}\times_1\boldsymbol{A}^{(1)}\times_2\cdots\times_N\boldsymbol{A}^{(N)}\equiv\left[\!\left[\underline{\boldsymbol{G}};\boldsymbol{A}^{(1)},\ldots,\boldsymbol{A}^{(N)}\right]\!\right]$$

where $\underline{\boldsymbol{G}}\in\Re^{J_1\times\cdots\times J_n\times\cdots\times J_N}$ is the core tensor, and $\boldsymbol{A}^{(n)}\in\Re^{J_n\times I_n}$ denotes factor matrices. The canonical decomposition/parallel factor analysis (CP) of a tensor $\underline{\boldsymbol{X}}\in\Re^{I_1\times\cdots\times I_n\times\cdots\times I_N}$ is concisely expressed as [14,15]

$$\underline{\boldsymbol{X}}\approx\sum_{r=1}^{R}\boldsymbol{a}_r^{(1)}\circ\boldsymbol{a}_r^{(2)}\circ\cdots\circ\boldsymbol{a}_r^{(N)}\equiv\left[\!\left[\boldsymbol{A}^{(1)},\ldots,\boldsymbol{A}^{(N)}\right]\!\right]$$

where $\boldsymbol{a}_r^{(n)}\in\Re^{I_n}$ is the rth-column vector of matrix $\boldsymbol{A}^{(n)}$.

2.2. Bilinear PLS and unfold-PLS

The partial least squares (PLS) is a well-known multivariate regression method for two-way data [16]. The unfold-PLS, also known as MPLS, is an extended version of PLS to deal with three-way data [2]. The relation between PLS and MPLS is that MPLS performs ordinary PLS on two matrices unfolded from three-way data arrays. Given two three-way data arrays $\underline{\boldsymbol{X}}(I\times J\times K)$ and $\underline{\boldsymbol{Y}}(I\times M\times N)$, they are unfolded into two matrices $\boldsymbol{X}(I\times JK)$ and $\boldsymbol{Y}(I\times MN)$. MPLS decomposes matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ as

$$\begin{aligned}\boldsymbol{X}&=\sum_{r=1}^{R}\boldsymbol{t}_r\boldsymbol{p}_r^T+\boldsymbol{E}=\boldsymbol{T}\boldsymbol{P}^T+\boldsymbol{E}\\ \boldsymbol{Y}&=\sum_{r=1}^{R}\boldsymbol{t}_r\boldsymbol{q}_r^T+\boldsymbol{F}=\boldsymbol{T}\boldsymbol{Q}^T+\boldsymbol{F}\end{aligned}\tag{1}$$

where $\boldsymbol{t}_r$ $(I\times1)$ are score vectors, $\boldsymbol{p}(JK\times1)$ and $\boldsymbol{q}(MN\times1)$ are loading vectors, and R is the number of score vectors. $\boldsymbol{E}(I\times JK)$ and $\boldsymbol{F}(I\times MN)$ are residual matrices. $\boldsymbol{T}=[\boldsymbol{t}_1,\ldots,\boldsymbol{t}_R]$ is the score matrix, $\boldsymbol{P}=[\boldsymbol{p}_1,\ldots,\boldsymbol{p}_R]$ and $\boldsymbol{Q}=[\boldsymbol{q}_1,\ldots,\boldsymbol{q}_R]$ are loading matrices. MPLS has the same properties as PLS. More detailed descriptions about MPLS can be found in ref [2].

3. Multilinear PLS

3.1. A brief review of multilinear PLS methods

The conventional bilinear PLS cannot deal with multi-way data directly. The unfold-PLS is applicable for multi-way data by using the