Institute for Advanced Studies
in Basic Sciences
Gava Zang, Zanjan, Iran

Calculation and Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Three-Component System

Ph.D. Thesis

Samira Beyramysoltan

Supervisors: Prof. H. Abdollahi
Prof. R. Rajkó

January 2014

Acknowledgment

It is not so vague that the success of any project depends largely on the encouragement and guidelines of many others. I take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this project.

I would like to start with the person who motivates me in doing this project, my supervisor Prof. Hamid Abdollahi. I would like to express my deep and sincere gratitude to him for the continuous support of my Ph.D. study and research, and for all I learned from his immense knowledge. Thanks him for questioning me about my ideas, helping me think rationally and even for hearing my problems. I wish to especially thank him for giving opportunity to cooperate with Prof. Róbert Rajkó. A special gratitude I give to Prof. Róbert Rajkó who accepted to supervise this project from the University of Szeged, Hungary. I enjoyed the personal discussion with him and the time I spent with him during nine months in Szeged. He has taught me how a person can succeed in achieving what seems impossible to begin with.

I hope the best success for my best academic family, Prof. Abdollahi and Prof. Rajkó whom taught me to constrain all of power I have for reducing the possible ambiguities in research and life ways.

I would also like to thank my committee members, Prof. Mojtaba Shamsipour, Prof. Taghi Khayamiyan, Dr. Mohamad Rafiee, and Dr. Mehdi Vasighi for serving as my committee members even at hardship. I also want to thank you for letting my defense be an enjoyable moment, and for your brilliant comments and suggestions, thanks to you.

My acknowledgement will never be complete without the special mention of Dr. Fariba Nazari for all support that she has extended to me throughout.

I also wish to acknowledge all the teachers I learnt from since my childhood, I would not have been here without their guidance, blessing and support. My special thanks to Prof. Behzad Haghighi, Dr. Mohsen Kompany-Zareh, for teaching me in my Ph. D. time. I feel privileged to be their student once, a feeling that every student of their will has.

I would like to acknowledge all of my friends for their moral support and motivation, which drives me to give my best. I find myself lucky to have friends like them in my life. Unfortunately, I cannot list all of them at this text, but they will always be safe in my heart.

Last but not the least; I would like to acknowledge the people who mean world to me, my parents, sisters, and brothers. They were always supporting me spiritually throughout my life and encouraging me with their best wishes. Special thanks to my Dad who enlightened my direction in science for first time and it is so gloomy day to me that he is not by the side of me now. Words cannot express how grateful I am to my Mum for all of the sacrifices that she has made on my behalf. Her prayer for me was what sustained me thus far.

Samira Beyramysoltan

IASBS

January 21, 2014

Abstract

The curve resolution (CR) methods attract substantial research affords aimed to discover knowledge of multicomponent systems. The objective of soft modeling, as special part of CR methods, is to resolve a second-order bilinear data matrix into its contributing matrices without prior knowledge about the chemical or physical model of system under study. The obtained results by soft modeling multivariate curve resolution methods often are not unique and are questionable because of rotational ambiguity. It means a range of feasible solutions equally fit experimental data. Due to rotational ambiguity, the accuracy of the result of soft methods was abolished by systematic error in case of quantitative analysis. Aiming to increase the accuracy, the range of solutions can be reduced by applying the useful knowledge of the studied chemical system as constraints.

It seems when there is no unique solution for the system, consideration of all feasible solutions (comparing with one solution that can be obtained with several multivariate curve resolution methods) can provide useful information about system and process under study.

Introduction to a novel grid search method

SMCR methods method can calculate the area of the feasible regions for each component, while they are not able to detect which solutions of these feasible regions are corresponding with each other. Therefore, it is valuable to modify pervious methods to develop one new method with this ability. In this study, firstly in chemometric literature, a novel grid search method is proposed. The method can provide the area of feasible solutions associated to each component and correspondence relations among solutions simultaneously.

This introduced method gives an opportunity to investigate constraints effect and to compare obtained feasible solutions by different approaches as well.

I. Investigation of the Equality Constraint Effect on the Reduction of Rotational Ambiguity in Three-Component System Using a Novel Grid search Method

Regarding to chemometric literature, a survey of useful constraints for the reduction of the rotational ambiguity is a big challenge for chemometrician. It is worth to study the effects of applying constraints on the reduction of rotational ambiguity, since it can help us to choose the useful constraints in order to impose in multivariate curve resolution methods for analyzing data sets. In this work, we have investigated the effect of equality constraint on decreasing of the rotational ambiguity. For calculation of all feasible solutions corresponding with known spectrum, a novel systematic grid search method based on Species-based Particle Swarm Optimization is proposed in a three-component system.

II. A Systematic Investigation on Meaning of Feasible Band Boundaries in Multivariate Curve Resolution of a Three-Component System

In the different studies by Abdollahi et al. and Rajko, different approaches were compared on the meaning of feasible band boundaries in MCR methods for two component system. It was considered that all approaches characterize the same feasible band boundaries in the case of a two-component system. With the same aim, to best of our knowledge firstly in the literature, grid search and MCR-Bands approaches are examined for the calculation of band boundaries of feasible solutions of a three component system. The MCR-Bands method based on exploring of the extreme values of the Signal Contribution Function (SCF) relative to the elements of the transformation matrix by a non-linear constrained optimization algorithm defines the estimated boundaries of the feasible solutions band. Aimed at evaluation of two methods in meaning of band boundaries, the SCF is calculated for all the explored feasible solutions by a grid search method based on Species Particle Swarm Optimization (SPSO) and results are displayed in appropriate mesh and contour plots. The extreme values of SCF are determined and compared with the results of MCR-Bands program.

III. Newer Developments on Analytical Self-Modeling Curve Resolution (SMCR)

Analytical SMCR methods resolve the data set to all feasible solutions using only non-negativity constraints. Lawton-Sylvester method was the first presented direct method to analyze a two-component system and it was generalized as Borgen plot for determining the feasible regions in three-component system. It seems to require the geometric view for considering curve resolution methods stopped general investigation on the Borgen work for 20 years. Then, Rajkó and Istevan revised and elucidated principles of existing theory in

SMCR methods and subsequently introduced computational geometry tools for developing an algorithm to draw Borgen plot in three-component system. Although, the analytical SMCR methods are definitely fast, but their inability in analysis of the data set with immoderate noise and in implementation of constraints limits their application. In this study, analytical SMCR methods will be described with simplest concepts and details of a developmental type of Borgen plot drawing algorithm will be given. Subsequently, for the first time in the literature, the equality and the unimodality are successfully implemented in the Lawton-Sylvester method. To end, a state-of-the-art procedure was proposed to impose equality constraint in the Borgen plot.

IV. Users' guide to *LSandBP*

Toolbox for the computation of the feasible solutions of two and three component systems using analytical SMCR methods.

A new Graphical User friendly Interface (GUI) is presented in MatLab computer program to introduce the abilities of analytical SMCR methods in resolving two-way data set to pure factors. Currently, this interface provides the direct computation of feasible regions and bands associated to components in two and three component systems using the Lawton-Sylvester and Borgen plot methods, respectively. The result can also be provided by implementing the equality and unimodality constraints in Lawton-Sylvester method and the equality constraint in Borgen plot. A prominent feature of the representation in this GUI is an interactive visual inspection of constraint effects on the feasible regions and bands.

V. Definition and Detection of Data based Uniqueness

Incorporation of some useful information may lead to drastically decreasing the ambiguity in the analysis of bilinear data sets by the curve resolution tasks. Some profiles with specific condition could be resolved uniquely under non-negativity constraint in the decomposition of a two-way data set. Uniqueness condition of minimal constraint systems (termed data based uniqueness) remains as a mysterious question in the literature. In this study, for the first in the literature, data based uniqueness is investigated in details and general procedure based on data set structure information is presented for detection of profiles in which recovered unambiguously in minimal constrained SMCR. Close inspection of local rank information and Borgen plot results lead to devise the comprehensive theorem which could be considered as a corner-stone of data based uniqueness detection.

CONTANTS

List of Figures

List of Schemes