Geometric Search: A New Approach for Fitting PARAFAC2 Models on GC-MS Data

Kuangda Tian, Lijun Wu, Shungeng Min, Rasmus Bro

www.elsevier.com/locate/talanta

PII: S0039-9140(18)30329-1
DOI: https://doi.org/10.1016/j.talanta.2018.03.088
Reference: TAL18522

To appear in: *Talanta*

Received date: 22 January 2018
Revised date: 25 March 2018
Accepted date: 26 March 2018

Cite this article as: Kuangda Tian, Lijun Wu, Shungeng Min and Rasmus Bro, Geometric Search: A New Approach for Fitting PARAFAC2 Models on GC-MS Data, *Talanta,* https://doi.org/10.1016/j.talanta.2018.03.088

Geometric Search: A New Approach for Fitting PARAFAC2 Models on GC-MS Data

Kuangda Tian [a, b], Lijun Wu [c], Shungeng Min [a*], Rasmus Bro[b*]

[a] Department of Applied Chemistry, College of Science, China Agricultural University, Bejing, 10093, China

[b] Department of Food Science, Faculty of Science, Copenhagen University, Copenhagen, DK-1958, Denmark

[c] China Tobacco Yunnan Industrial Co., Ltd., Kunming 650000, China

Shungeng Min, minsg@cau.edu.cn

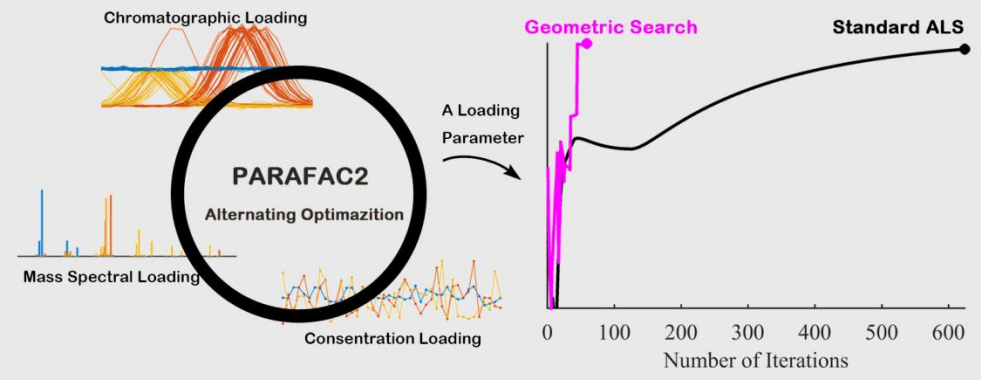
Rasmus Bro, rb@life.ku.dk;

*Corresponding authors:

Abstract

PARAFAC2 is a powerful decomposition method which is ideally suited for modeling gas chromatography-mass spectrometry (GC-MS) data. However, the most widely used fitting algorithms (alternating least squares, ALS) are very slow which hinders use of the model. In this paper, an iterative method called geometric search is proposed to fit the PARAFAC2 model. This method models the PARAFAC2 loading parameters as geometric sequences with offsets during the ALS iterations. It extrapolates the optimal parameters from prior iterations to accelerate ALS convergence process. The performance of this method was evaluated by simulated datasets and two GC-MS datasets of wine and tobacco samples. This geometric search method proved an efficient way to fit PARAFAC2 models, compared with a standard ALS algorithm and two widely used line search algorithms in terms of convergence speed and fitting quality.

Graphical abstract



Keywords: Geometric search; PARAFAC2; Line search; Alternating least squares (ALS); GC-MS