Identifying metabotypes from tensor data

Authors: Viktor Skantze[1,2], Mikael Wallman[1], Ann-Sofie Sandberg[2], Rikard Landberg[2], Mats Jirstrand[1], Carl Brunius[2]

[1] The Fraunhofer-Chalmers Research Centre for Industrial Mathematics, Gothenburg, Sweden

[2] Department of Biology and Biological Engineering, Chalmers University of Technology, Gothenburg Sweden

Theme: Computational Metabolomics, Statistics & Bioinformatics

Sub-topic: Statistical methods and machine learning

Keywords: Metabotyping, Metabolism, Precision nutrition

Max length = 300 words, Deadline March 14, 2022.

Intro: Metabolic response to diet shows large individual variation, which warrants tailored dietary recommendation i.e., personalized nutrition (PN). A step towards PN is to tailor diet to groups of individuals with similar metabolic phenotype, so called metabotypes (i.e., clusters of individuals with similar metabolism). Metabotyping of high-dimensional data is commonly performed in matrix form using matrix decompositions (e.g., PCA). However, data from e.g., crossover studies can be conveniently organized in multi-dimensional form (i.e., as tensor data) and methods for detecting metabotypes in such data are still lacking.

Aim: We therefore aimed to develop and evaluate tools to identify potential metabotypes in high-dimensional tensor data.

Methods: We developed two methods: The first uses CANDECOMP/PARAFAC (CP) decomposition directly on tensor data where clustering was performed on individual's scores, whereas the second was developed specifically for time-resolved data and uses dynamic mode decomposition (DMD) to model metabolite dynamics, where clustering was performed on individual's dynamic state trajectories. We applied the methods to identify metabotypes in data from a crossover acute post-prandial dietary intervention study on 17 overweight males (BMI 25–30 kg/m^2, 41–67 y of age) undergoing three dietary interventions (pickled herring, baked herring and baked beef, measuring 79 metabolites (from GC-MS metabolomics) at 8 time points (0-7h).

Results and conclusions: Both methods identified two potential metabotype clusters, predominantly in amino acids after the meat diet. The clustering associated to baseline levels of creatinine, strengthening the plausibility of found metabotypes. The CP method is a general approach, not specific to time-resolved data, and provides better fit if the data is multilinear. Conversely, DMD is designed for time-resolved data, for which it often provides a better fit than CP. We concluded that both the CP and the DMD approach are well suited to identify metabotypes in tensor data from a wide variety of complex experimental designs.

Metabolomics Society: Abstract Evaluation Instructions

ALL abstracts must meet the following two criteria:

1. Topic is relevant to the Metabolomics Society conference
2. Abstract does not contain obvious marketing content

Abstracts that have been submitted for an oral presentation or are being judged for an award must also be evaluated using the attached rubric.

Notes on Using the Rubric

- Methodology: The level of methodological detail should be assessed relative to the type of study. A higher level of detail and depth is expected for methodological studies.
- Contribution to Discipline: The discipline is context-dependent and should be defined by the authors. For example, the results of a clinical metabolomics study may contribute to the specific disease discipline, broader clinical metabolomics, and/or even the field of metabolomics.
- Scientific Writing: The abstract structure should follow a logical sequence: Intro, Aim, Methods, Results, Conclusions. When assessing Scientific Writing, consider how well the ideas connect through the abstract. This criterion should also be assessed *relative to language barriers*.

Finally, if an abstract has been submitted for an oral presentation, please consider the following question and click the corresponding checkbox if it applies.

This abstract is of high general interest or relevance to the Metabolomics Society community as a whole. It should be considered for an oral presentation.

Metabolomics Society: Abstract Evaluation Rubric

1	2	3	4	5
1. Introduction and Aim				
Need and aim of study are missing, inaccurate, or misleading.	Need and aim of study stated but unclear.	Need and aim stated but not well connected or irrelevant information included.	Need and aim are concisely stated. Link between need and aim lacks clarity. All information is relevant.	Clearly and concisely states scientific need for study and aim/hypothesis. Need and aim are well connected. All information is highly relevant.
2. Methodology				
Methods are missing, inaccurate, or misleading.	Methods provided but connection to aim is unclear. Type and level of detail inappropriate relative to aim.	Methods provided but connection to aim is unclear. Type and level of detail appropriate relative to aim.	Methods articulated and connected to aim but lack clarity. Type and level of detail appropriate relative to aim.	Methods clearly articulated, well connected to aim of study, type and level of detail appropriate relative to aim.
3. Findings and Significance				
Findings and their significance are missing.	Findings presented but connection to aim unclear. Significance of findings not addressed.	Findings presented but connection to aim lacks clarity. Significance of findings also lacks clarity.	Findings presented and well connected to aim. Significance of findings explained. Lack of clarity in one of the above.	Findings clearly presented and well connected to aim. Significance of findings clearly explained relative to aim.
4. Contribution to Discipline				
Contribution to discipline missing or misleading.	Unclear attempt to connect study to discipline.	Connects study to discipline in terms of future use case. Unclear how study will advance the discipline.	States how study will advance discipline or address the specific discipline need but lacks clarity.	Clearly articulates how the study will advance the discipline or address the specific need in the discipline.
5. Scientific Writing				
Abstract lacks structure. Grammatical and spelling errors significantly impede understanding. Inappropriate verb tense used.	Partially structured abstract; ideas not all well connected. Grammatical and spelling errors present throughout, but do not impede understanding. Inappropriate verb tense used.	Abstract structured appropriately, with minor lack of connection of ideas. Few grammatical and/or spelling errors present, but do not impede understanding. Minor inconsistencies in verb tense usage.	Abstract well structured, with clear connection of ideas. Very minor grammatical or spelling errors present. Verb tense used appropriately throughout.	Excellent structure: all ideas well connected. Free of grammatical and spelling errors. Verb tense used appropriately throughout.