Comparison of interpolation algorithms for handling of scattering signal before PARAFAC processing of fluorescence images

Ivan Krylov Timur Labutin

Lomonosov Moscow State University, Department of Chemistry

PARAFAC is an attractive method in fluorescence spectroscopy because its mathematical formulation closely mathes the underlying physical model for the fluorescence signal intensity of a mixture of fluorophores under certain, typically valid, assumptions. Unfortunately, excitation-emission matrices (EEMs) typically contain a scattering component, which doesn't adhere to PARAFAC assumpltions and therefore must be taken care of before the analysis. Interpolation of the scattering signal region is a popular method of dealing with the scattering signal. It's computationally cheap and easy to understand. Unlike setting the scatter region to missing, it doesn't give rise to local minima in the PARAFAC solution [1].

This work takes a multitude of surface interpolation methods and compares their performance on two model datasets consisting of simulated fluorescence and scattering signals. Compared performance metrics include root-mean-squared error of recovered intensity values and the Tucker's congruence coefficient between loadings estimated by PARAFAC and their *a priori* values. The methods compared include monotone piecewise bicubic interpolation of each row or column of the EEM, multilevel B-splines, Whittaker interpolation with different combinations of difference orders, the LOESS algorithm and Kriging. One of the datasets is constructed from orthogonal scores to be easier to fit and the other has scores as a sample from multivariate normal distribution with non-zero correlations to simulate the kind of harder-to-fit data that comes from real-world samples.

The trends in recovery of the original fluorescence signal and in the PARAFAC estimation of the ground truth loadings are basically the same. The results show that Whittaker interpolation, LOESS and Kriging provide the best results, subject to the need to cross-validate for optimal parameter values (Whittaker & LOESS) and high computational complexity (Kriging).

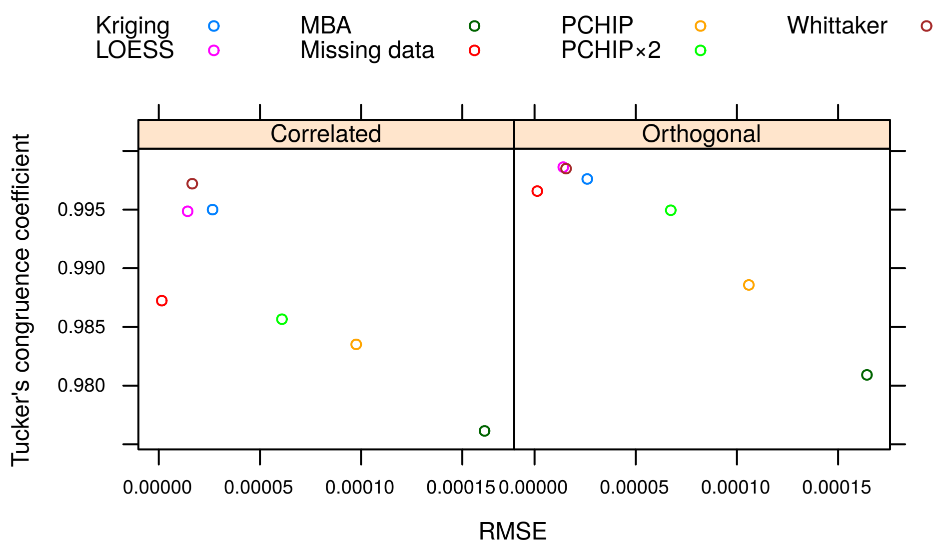

Figure 1. Root-mean-square error of original fluorescence signal recovery and Tucker's congruence coefficient between ground truth loadings and their PARAFAC estimation for different interpolation methods.

[1] S. Elcoroaristizabal, R. Bro, J.A. García, L. Alonso, PARAFAC models of fluorescence data with scattering: A comparative study, Chemometrics and Intelligent Laboratory Systems. 142 (2015) 124–130. https://doi.org/10.1016/j.chemolab.2015.01.017.