KL-NMF as best suited NMF variant for the analysis of mass spectrometry imaging data

Melanie Nijs (*), Tina Smets, Etienne Waelkens, Bart De Moor
KU Leuven, ESAT/STADIUS, Kasteelpark Arenberg 10, 3001 Leuven, Belgium,
melanie.nijs@esat.kuleuven.be

Aim:
Over the last few decades, non-negative matrix factorization (NMF) has proven as a very effective approach for the analysis of biomedical data. Within the field of mass spectrometry imaging (MSI) data analysis, it offers intuitive results as it allows the visualization of the spatial abundance map for each spectral basis factor. Given the Poisson distributed nature of MSI data, we compare NMF to its Kullback-Leibler (KL-NMF) variant and probabilistic latent semantic analysis (PLSA). Latent Dirichlet allocation (LDA) is added as it improves PLSA in text mining. This work offers a detailed comparison of these different matrix factorization methods for MSI data analysis.

Methods:
The models investigated are non-negative matrix factorization (NMF), Kullback-Leibler non-negative matrix factorization (KL-NMF), probabilistic latent semantic analysis (PLSA), and latent Dirichlet allocation (LDA). The experiments have been conducted on seven different samples: three mice pancreas sections and four human lymph node sections, all obtained using MALDI-TOF MS.

Results and discussion:
The work shows that each method relies on a different framework which has implications on the exactness and interpretability of the results. KL-NMF, which is not typically used for MSI data analysis so far, improves both NMF and PLSA. PLSA, while too overfitting in text mining, improves LDA in the setting of MALDI-MSI. Also, we proved that KL-NMF and PLSA algorithms reach a different local optimum in the case of MSI data analysis. Additionally, the molecular results of the human-lymph-node data have been thoroughly analyzed for better assessment of the methods under investigation.

Conclusions:
This work offers a thorough comparison of multiple non-negative matrix factorization related methods for MSI. We aim to provide fellow researchers in the field of MSI a clear understanding of the mathematical implications using each of these analysis techniques, which might affect the correctness and interpretation of the results. (Nijs et al., 2021)

References:

Nijs, M., Smets, T., Waelkens, E., De Moor, B., 2021. A mathematical comparison of non-negative matrix factorization related methods with practical implications for the analysis of mass spectrometry imaging data. Rapid Commun. Mass Spectrom. 35, e9181. https://doi.org/10.1002/RCM.9181