Journée des jeunes chercheurs en Traitement du signal et de l'image 2019
GRAISyHM TSI 2019
Deuxième édition : 17 juin 2019, Amiens, France

Non-Negative Matrix Factorization with Missing Entries: A Random Projection Based Approach

Farouk Yahaya*, Matthieu Puigt, Gilles Delmaire, and Gilles Roussel
Univ. Littoral Côte d'Opale, LISIC – EA 4491, F-62228 Calais, France
e-mail: firstname.lastname@univ-littoral.fr

Domaine: Traitement du signal

1 Abstract

Non-negative Matrix Factorization (NMF) is a low-rank approximation tool which is very popular in signal processing, in image processing, and in machine learning [1]. It consists of factorizing a non-negative matrix by two non-negative matrices. While being extremely general, this problem finds many applications, including environmental data processing [2]. Unfortunately, classical NMF techniques are not well-suited to process very large data matrices. To solve such an issue, NMF has been recently combined with random projections (see, e.g., [3] and the references within). The latter is a distance-preserving dimension reduction technique based on randomized linear algebra [4].

However, random projections cannot be applied in the case of missing entries in the matrix to factorize, which occurs in many actual problems with large data matrices, e.g., mobile sensor calibration [5]. Our contribution to solve this issue lies in proposing a novel framework to apply random projections in a weighted NMF, where the weight models the confidence in the data (or the absence of confidence in the case of missing data) [6]. We experimentally show the proposed framework to significantly speed-up state-of-the-art NMF methods under some mild conditions. In particular, the proposed strategy is particularly efficient when combined with Nesterov gradient or alternating least squares.

Références

[1] N. Gillis, The why and how of nonnegative matrix factorization, *Regularization, Optimization, Kernels, and Support Vector Machines*, Chapman and Hall/CRC, pp. 257–291, 2014.

[2] M. Puigt, G. Delmaire, and G. Roussel, Environmental signal processing : new trends and applications, *in Proc. ESANN'17*, pp. 205–214, 2017.

[3] F. Yahaya, M. Puigt, G. Delmaire, and G. Roussel, Faster-than-fast NMF using random projections and Nesterov iterations, *in Proc. iTWIST*, 2018.

[4] N. Halko, P.-G. Martinsson, and J. A. Tropp, Finding structure with randomness : Probabilistic algorithms for constructing approximate matrix decompositions, *SIAM review*, vol. 53, no. 2, pp. 217–288, 2011.

[5] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, Informed nonnegative matrix factorization methods for mobile sensor network calibration, *IEEE Transactions on Signal and Information Processing over Networks*, vol. 4, no. 4, pp. 667–682, 2018.

[6] Y. D. Kim and S. Choi, Weighted nonnegative matrix factorization, *In Proc. ICASSP'09*, pp. 1541–1544, 2009.

*F. Yahaya gratefully acknowledges the Région Hauts-de-France to partly fund his Ph.D. fellowship. Experiments presented in this communication were carried out using the CALCULCO computing platform, supported by SCoSI/ULCO.