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

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


*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.