Fast Rank-1 NMF for Missing Data with KL Divergence
Kazu Ghalamkari¹,², Mahito Sugiyama¹,²
¹National Institute of Informatics @ Tokyo, Japan ²The Graduate University for Advanced Studies, SOKENDAI

The 25th International Conference on Artificial Intelligence and Statistics (AISTATS 2022), March 28–30, 2022

Introduction

• We found the formula of the best rank-1 factor sharing NMF w.r.t. minimizing KL divergence.
• We developed a non-gradient-based rank-1 NMF with missing values based on the formula.

Rank-1 Non-negative Multiple Matrix Factorization

For given X, Y and Z, rank-1 NMMF (non-negative multiple matrix factorization) finds a, b, w and h to minimize D(X, w ⊙ h) + αD(Y, a ⊙ h) + βD(Z, w ⊙ b).

Main Result

The best rank-1 NMMF of X and Y is given as

\[ w_i = \sqrt{\frac{S(X)}{S(X)+bS(Z)}} \left( \sum_{j} X_{ij} + \sum_{m} bZ_{im} \right), \quad a_m = \sum_{n} Y_{nj} \sqrt{\frac{S(Y)}{S(Y)+aS(X)}}, \]
\[ h_j = \sqrt{\frac{S(X)}{S(X)+aS(Y)}} \left( \sum_{i} X_{ij} + \sum_{n} aY_{nj} \right), \quad b_m = \sum_{n} Z_{im} \sqrt{\frac{S(X)}{S(Y)+aS(X)}}, \]

S(X) is the sum of all elements of X.

Faster Rank-1 NMF with Missing Values

• NMMF can be viewed as a special case of NMF with missing values.

A1GM: Proposed Method for Rank-1 NMF with Missing Values

A1GM is a non-gradient-based method.

Step 1: Increase the number of missing values.
Step 2: Gather missing values in the bottom right.
Step 3: Use the formula of rank-1 NMMF and repermutate.

Experiments on Real Data

Our method is compared with gradient-based KL-WNMF.
- Relative runtime < 1 means A1GM is faster than KL-WNMF.
- Relative error > 1 means worse reconstruction error of A1GM than KL-WNMF.