1 Overlapping Nonnegative Matrix tri-Factorization

This is a proposal algorithm that aims to solve the following problem:

$$\mathscr{F} = min\,||X - U\sum_{i=1}^{k} I^{(i)} S V^{(i)}||_F^2$$

$$s.t.\; U, S, V^{(1)}, \ldots, V^{(k)} \geq 0$$

where $X \in \mathbb{R}^{n\times m}$ a data matrix, $U \in \mathbb{R}^{n\times k}$ an matrix containing rows clusters, $I^{(i)} \in \mathbb{R}^{k\times k}$ an indicator matrix, $S \in \mathbb{R}^{k\times l}$ a block matrix, $V^{(i)} \in \mathbb{R}^{l\times m} \forall i \in \{1,\ldots,k\}$ matrices containing columns clusters for each row cluster, and $||\cdot||_F^2$ denotes the frobenius norm.

Consider the lagrange function:

$$\mathscr{L} = \mathscr{F} - tr(\Lambda_1 U^T) - tr(\Lambda_2 S^T) - \sum_{i=1}^{k} tr(A_i {V^{(i)}}^T)$$

where Λ_1, Λ_2 and $A_i \forall i \in \{1,\ldots,k\}$ are lagrange multipliers: $\Lambda_1 \in \mathbb{R}^{n\times k}$, $\Lambda_1 \in \mathbb{R}^{k\times l}$, and $A_i \in \mathbb{R}^{k\times l} \forall i \in \{1,\ldots,k\}$.

KKT Conditions:

- $\frac{\partial \mathscr{L}}{\partial U} = 0$
- $\frac{\partial \mathscr{L}}{\partial S} = 0$
- $\frac{\partial \mathscr{L}}{\partial V^{(i)}} = 0, \forall i \in \{1,\ldots,k\}$
- $\Lambda_1 \odot U = 0$
- $\Lambda_2 \odot S = 0$
- $A_i \odot V^{(i)} = 0, \forall i \in \{1,\ldots,k\}$

where $\odot$ denotes the element-wise multiplication.

Given that:

$$\frac{\partial}{\partial B} tr\left[(RBC+D)(RBC+D)^T\right] = 2R^T(RBC+D)C^T = 2R^t rBCC^T + 2R^T DC^T$$

It is possible to calculate $\frac{\partial \mathscr{L}}{\partial U}$, considering $R = -I, B = U, C = \sum I^{(i)} S V^{(i)}, D = X$:

$$\frac{\partial \mathscr{L}}{\partial U} = 2U\sum_{i=1}^{k} I^{(i)} S V^{(i)} \sum_{j=1}^{k} {V^{(j)}}^T S^T I^{(j)} - 2X\sum_{i=1}^{k} {V^{(i)}}^T S^T I^{(i)} - \Lambda_1$$

For $\frac{\partial \mathscr{L}}{\partial V}$, consider $R = U \sum I^{(i)} S, B = V^{(i)}, C = -I, D = X$:

$$\frac{\partial \mathscr{L}}{\partial V^{(j)}} = 2S^T I^{(j)} U^T U I^{(j)} S V^{(j)} - 2S^T I^{(j)} U^T X - A_j, \forall j \in \{1, \ldots, k\}$$

For $\frac{\partial \mathscr{L}}{\partial S}$, consider that $\frac{\partial (AXBX^T C)}{\partial X} = A^T C^T X B^T + CAXB$:

$$\frac{\partial \mathscr{L}}{\partial S} = \frac{\partial}{\partial S} tr(X - U \sum_{i=1}^{k} I^{(i)} S V^{(i)})(U \sum_{j=1}^{k} I^{(j)} S V^{(j)})^T - tr(\Lambda_2 S^T)$$

$$= \frac{\partial}{\partial S} tr(X^T X) - 2\frac{\partial}{\partial S} tr(XU \sum_{i=1}^{k} I^{(i)} S V^{(i)}) + tr(\sum_{j=1}^{k} \sum_{i=1}^{k} U I^{(i)} S V^{(i)} V^{(j)^T} S^T I^{(j)} U^T) - tr(\Lambda_2 S^T)$$

$$= -2 \sum_{i=1}^{k} I^{(i)} U^T X^T V^{(i)^T} + \sum_{i=1}^{k} \sum_{j=1}^{k} [I^{(i)} U^T U I^{(j)} S V^{(j)} V^{(i)^T} + I^{(j)} U^T U I^{(i)} S V^{(i)} V^{(j)^T}]$$

$$= -2 \sum_{i=1}^{k} I^{(i)} U^T X^T V^{(i)^T} + \sum_{i=1}^{k} \sum_{j=1}^{k} I^{(i)} U^T U I^{(j)} S V^{(j)} V^{(i)^T}$$