Topic Modeling with SVD and NMF

April 25, 2020

0.1 Non-negative Matrix Factorization

In this section, we will briefly introduce the non-negative matrix factorization problem and algorithms in solving this problem.

0.1.1 Non-negative Matrix Factorization Problem

The non-negative matrix factorization (NMF for short) problem can be expressed formally as the following:

> **NMF problem:** Given a non-negative matrix $A \in R^{m \times n}$ and a positive integer $k < minm, n$, find non-negative matrices $W \in R^{m \times k}$ and $H \in R^{k \times n}$ to minimize the cost function

$$f(W, H) = \frac{1}{2}\|A - WH\|_F^2 \tag{1}$$

The product WH is called a non-negative matrix factorization of A. Through NMF, we get a low-rank (at most rank k) approximation, WH, of the original matrix A.

There are two note-worthy properties about NMF that we would like to mention at this moment. First is that, NMF can be written column by column as

$$a_i = Wh_i, \quad i \in 1, 2, ..., n \tag{2}$$

Where a_i is the i-th column vector of matrix A, and h_i is the corresponding column vector of matrix H. That is to say, each data vector a_i is approximated by a linear combination of the columns of W, weighted by the components of h_i. Therefore, W can be regarded as a basis that is optimized for the linear approximation of data A. Usually, the parameter k is small compare to the dimension of A, so NMF will be a good approximation if there exists latent structures in vectors of data matrix A.

Second property of NMF is that, negative entries are not allowed in W and H. For this reason, NMF is intuitively a part-based representation, since the whole is formed by combining the parts. These two features make NMF a good method for text classification, we will discuss this with details in later sections.

0.1.2 Basic Algorithms for Solving NMF Problems

Though NMF problems has a simple form, it has been proven that finding the exact solution of these problems in general is NP-hard.