Approaches to Brain Parcellation using Energy Statistics and Graph Partitioning

Felix Xiao

March 15, 2016

Contents

Chapter 1

Symmetric Nonnegative Matrix Factorization

In the previous chapter, we showed that the problem of finding the minimum ratio cut of a graph (with Laplacian matrix L, degree matrix D, and adjacency matrix A) can be formulated as minimizing

$$\mathrm{Tr}(R^T LR) \tag{1.1}$$

over the set, $\mathcal{R}$, of $n \times k$ matrices satisfying

1. $R^T R = I$
2. $R \geq 0$ (element-wise)
3. $RR^T u_n = u_n$ where u_n is a n-dimensional vector of all ones.

If the sizes of the components in the optimal ratio cut partition are perfectly balanced, which is equivalent to saying if the diagonal of the optimal ratioed assignment matrix RR^T has entries all equal to $\frac{k}{n}$, then

$$\begin{aligned}
\mathrm{Tr}(R^T DR) &= \sum_{i=1}^{n} [RR^T]_{ii} D_{ii} \\
&= \sum_{i=1}^{n} \frac{k}{n} D_{ii} \\
&= \frac{k}{n} \sum_{i,j} A_{ij}
\end{aligned}$$

is a constant that does not depend on R. The same is true if each vertex has the same degree $D_{ii} = d$, in which case

$$\begin{aligned}\mathrm{Tr}(R^T D R) &= \sum_{i=1}^{n} [RR^T]_{ii} D_{ii} \\ &= d \sum_{i=1}^{n} [RR^T]_{ii} \\ &= dk\end{aligned}$$

is also a constant that does not depend on R. In either case,

$$\operatorname*{argmin}_{R \in \mathcal{R}} \mathrm{Tr}(R^T L R) = \operatorname*{argmax}_{R \in \mathcal{R}} \mathrm{Tr}(R^T A R)$$

This equality may also hold even if neither condition is true, especially if they are approximately true.

Spectral k-partitioning drops the second and third constraints of $\mathcal{R}$ to derive a closed-form minimizer of 1.1, from which the original assignment matrix can be obtained by k-means. This chapter deals with an alternative relaxation of $\mathcal{R}$ that drops the first and third constraints.

1.1 Nonnegative Matrix Factorization

For an $n \times m$ matrix A, a nonnegative matrix factorization (NMF) is a pair of matrices $W \in \mathbb{R}^{n \times k}$ and $H \in \mathbb{R}^{m \times k}$ that minimizes $\|A - WH^T\|_F^2$ subject to elementwise nonnegativity: $H \geq 0$ and $W \geq 0$. Here, $\|X\|_F = \sqrt{\sum_{ij} X_{ij}}$ refers to the Frobenius norm.

For $n \times n$ symmetric matrices A, a *symmetric* NMF (SymNMF) is a matrix $H \in \mathbb{R}^{n \times k}$ that minimizes $\|A - HH^T\|_F^2$, and k is an arbitrary positive integer typically much smaller than n.

The following theorem from [Ding et al., 2005] illustrates the connection between SymNMF and graph partitioning.

Theorem 1.1.1 *Let A be a $n \times n$ symmetric matrix. Then*

$$\operatorname*{argmax}_{H^T H = I, H \geq 0} \mathrm{Tr}(H^T A H) = \operatorname*{argmin}_{H^T H = I, H \geq 0} \|A - HH^T\|_F^2$$

Proof.

$$\begin{aligned}\operatorname*{argmax}_{H^T H = I, H \geq 0} \mathrm{Tr}(H^T A H) &= \operatorname*{argmin}_{H^T H = I, H \geq 0} -2\,\mathrm{Tr}(H^T A H) \\ &= \operatorname*{argmin}_{H^T H = I, H \geq 0} \mathrm{Tr}(AA^T) - 2\,\mathrm{Tr}(H^T A H) + \|H^T H\|_F^2 \\ &= \operatorname*{argmin}_{H^T H = I, H \geq 0} \|A - HH^T\|_F^2\end{aligned}$$

If A is the adjacency matrix, then under the equal vertex degrees condition described earlier $\underset{H^T H=I, H\geq 0}{\operatorname{argmax}} \operatorname{Tr}(H^T A H) = \underset{H^T H=I, H\geq 0}{\operatorname{argmin}} \operatorname{Tr}(H^T L H)$. Hence an alternative approach to the minimum ratio-cut problem is to drop the $H^T H = I$ constraint and solve the SymNMF problem:

$$\begin{aligned} \min_{H\in\mathbb{R}^{n\times k}} \quad & \|A - HH^T\|_F^2 \\ \text{s.t.} \quad & H \geq 0 \end{aligned} \tag{1.2}$$

This relaxation has two key differences from the spectral relaxation .

- There is no closed-form solution, and the optimal value is found via an optimization algorithm, described in the next section.
- The optimal assignments are recovered directly from the largest entry in each row. There is no need for k-means.

1.2 An Alternating Nonnegative Least Squares Algorithm for SymNMF

[Kuang et al., 2015] re-formulates 1.2 as a non-symmetric NMF with a penalty on the difference between the two matrix factors:

$$\min_{W,H\geq 0} \|A - WH^T\|_F^2 + \alpha\|W - H\|_F^2 \tag{1.3}$$

where $W, H \in \mathbb{R}^{n\times k}$. The α parameter

The rationale for this to use known methods for solving the non-symmetric NMF and adapt them to the symmetric problem. One powerful framework for solving NMF is Alternating Nonnegative Least Squares (ANLS), which factors A into nonnegative W and H by fixing the H matrix and solving for W:

$$W \leftarrow \underset{W\geq 0}{\operatorname{argmin}} \|A - WH^T\|_F^2$$

and fixing this new matrix W and solving for H:

$$H \leftarrow \underset{H\geq 0}{\operatorname{argmin}} \|A - WH^T\|_F^2$$

and repeating the two steps until convergence. Both subproblems in the ANLS framework are convex, and the algorithm requires only an initial W to get started.

[Kuang et al., 2015] describes an algorithm for solving SymNMF that uses the ANLS framework. The objective function in 1.3 can be re-written as

$$\left\| \begin{bmatrix} W \\ \sqrt{\alpha} I_k \end{bmatrix} H^T - \begin{bmatrix} A \\ \sqrt{\alpha} W^T \end{bmatrix} \right\|_F^2 \tag{1.4}$$

Algorithm 1 ANLS algorithm for SymNMF

1: Initialize H
2: **repeat**
3: $\quad W \leftarrow H$
4: $\quad H \leftarrow \underset{H \geq 0}{\operatorname{argmin}} \left\| \begin{bmatrix} W \\ \sqrt{\alpha} I_k \end{bmatrix} H^T - \begin{bmatrix} A \\ \sqrt{\alpha} W^T \end{bmatrix} \right\|_F^2$
5: **until** convergence

with $\begin{bmatrix} W \\ \sqrt{\alpha} I_k \end{bmatrix}$ taking on the part of the fixed matrix and H the decision matrix. The ANLS algorithm for SymNMF is the following:

The α can be increased each iteration.

1.2.1 Block Principal Pivoting for Nonnegative Least Squares

[Kim and Park, 2011]

Bibliography

[Ding et al., 2005] Ding, C. H., He, X., and Simon, H. D. (2005). On the equivalence of nonnegative matrix factorization and spectral clustering. In *SDM*, volume 5, pages 606–610. SIAM.

[Kim and Park, 2011] Kim, J. and Park, H. (2011). Fast nonnegative matrix factorization: An active-set-like method and comparisons. *SIAM Journal on Scientific Computing*, 33(6):3261–3281.

[Kuang et al., 2015] Kuang, D., Yun, S., and Park, H. (2015). Symnmf: nonnegative low-rank approximation of a similarity matrix for graph clustering. *Journal of Global Optimization*, 62(3):545–574.