Machine Learning for Complex Networks
SoSe 2024

Prof. Dr. Ingo Scholtes
Chair of Informatics XV
University of Würzburg

Exercise Sheet 07

Published: June 19, 2024
Due: June 27, 2024
Total points: 10

Please upload your solutions to WueCampus as a scanned document (image format or pdf), a typesetted PDF document, and/or as a `jupyter` notebook.

1. **Laplacian Embeddings, PCA, and SVD**

   This exercise guides you through relating Laplacian Embeddings, Principal Component Analysis (PCA), and Singular Value Decomposition (SVD) through pseudoinverses and incidence matrices.

   (a) An inverse of a real matrix does not always exists. A generalization of the matrix inverse is known as the *pseudoinverse*,also known as the *Moore-Penrose* inverse. One way to calculate the inverse of a real matrix M is to calculate its Singular Value Decomposition (SVD) $M = U\Sigma V^T$ before taking its inverse $M^+ = V\Sigma^+U^T$, where the pseudoinverse $\Sigma^+ = \Sigma^{-1}$ of the diagonal singular value matrix Σ calculated by simply taking the reciprocal of each its elements. **1P**

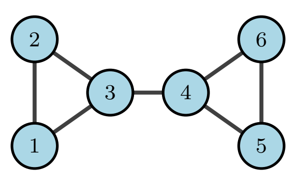
   Recall that the eigendecomposition of a real symmetric matrix L can be written as

   $$L = Q\Lambda Q^T = MM^T$$

   where $M = Q\sqrt{\Lambda}$. Show that $L = U\Sigma^T\Sigma U^T$. This is one way eigendecomposition and SVD are related for real symmetric matrices.

   (b) We can define $L^+ = \left(M^T\right)^+ M^+ = X^TX$. where $X = M^+$. This creates a (centered) "data matrix" X so L^+ can be interpreted as a covariance matrix. Now consider when L is the Laplacian of a simple graph G. Argue how the objective of Principal Component Analysis (PCA) of X is equivalent to that of spectral embedding of G via eigendecomposition of its Laplacian L (Laplacian Embedding). **3P**

   (c) Consider the following simple graph. **1P**

   Write down its (unnormalized) Laplacian L.

   (d) Calculate its "data matrix" X and show that PCA of X is equivalent to that of its spectral embedding via eigendecomposition of its Laplacian L (Laplacian Embedding). **2P**

   (e) Instead of constructing the Laplacian as $L = D - A$ where D and A are the degree and adjacency matrices, respectively, the Laplacian can be constructed via *incidence matrices*. The incidence matrix B describes the connectivity between nodes and edges. For the simple graph above the incidence matrix can be written as **1P**

$$B = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & -1 \end{pmatrix}$$

where its rows represent nodes and its columns represent edges. Note how every edge is *incident* to only 2 nodes, and the incidence matrix requires an *orientation*, labelling one incidence 1 and another -1. The choice of which incidence is labelled 1 or -1 is arbitrary, and is simply a matter of bookkeeping [1].

For the simple graph above, calculate $L = BB^T$.

(f) $L = BB^T$ holds more generally than for this example. Show that (generally, for simple graphs) the singular vectors of Singular Value Decompostion (SVD) of B are equivalent to the eigenvectors of eigendecomposition of L. 1P

(g) Show for the simple graph above that the singular vectors of Singular Value Decompostion (SVD) of B are equivalent to the eigenvectors of eigendecomposition of L. 1P

[1] For the interested reader, a simple graph is a 1-*simplicial complex* A simplicial complex is one generalization of a simple graph to describe higher-order connectivity e.g. between 3 or more nodes. The 0-*Hodge Laplacian* L_0 for simplicial complexes describes its pairwise node connectivity and can be constructed from its corresponding node-edge *incidence matrix* (also known as *boundary operator*) B_1. For simple graphs, the 0-Hodge Laplacian L_0 corresponds to the standard (unnormalized) graph Laplacian L, and $B_1 = B$.