Lecture 17

8.4 Weighted path graphs

Taken from [10, Lecture 3]

As alluded to at the end of the previous section, we now analyze weighted path graphs. To that end, we prove the following:

Theorem 6 (Fiedler). *Let $P = (V, E, w)$ be a weighted path graph on n vertices, let L_P have eigenvalues $0 = \lambda_1 < \lambda_2 \leq \cdots \leq \lambda_n$, and let φ_k be an eigenvector with eigenvalue λ_k. Then φ_k changes sign $k-1$ times.*

We will need to first prove a few lemmas in order to prove Theorem 6. The first of these is Sylvester's Law of Inertia:

Theorem 7 (Sylvester's Law of Intertia). *Let A be any symmetric matrix and let B be any non-singular matrix (that is, B has no zero singular values). Then, the matrix BAB^T has the same number of positive, negative and zero eigenvalues as A.*

Proof. We first recall three facts from linear algebra.

1. The first is that BAB^{-1} has the same eigenvalues as A, since:

$$A\varphi = \lambda\varphi \Longleftrightarrow BAB^{-1}(B\varphi) = \lambda(B\varphi).$$

2. The second fact is that $\text{rank}(A) = \text{rank}(BAB)$.

3. The third is that every nonsingular matrix B can be written $B = QR$, where Q is an orthonormal matrix (meaning $Q^T Q = QQ^T = I$) and R is an upper-triangular matrix with positive diagonals (this is the so-called QR factorization).

We are going to begin by using a slight variation of the last fact, and write $B = RQ$. Now, since $Q^T = Q^{-1}$, by the first fact we know that QAQ^T has exactly the same eigenvalues as A. Define

$$\forall\, t \in [0, 1], \quad R_t = tR + (1-t)I,$$

and consider the family of matrices

$$\forall\, t \in [0,1], \quad M_t = R_t Q A Q^T R_t^T.$$

At $t = 0$ we have $R_0 = I$ and so $M_0 = QAQ^T$ has the same eigenvalues as A. For $t = 1$ we have $M_1 = BAB^T$. Since all of the matrices M_t are symmetric, they all have real eigenvalues (by the Spectral Theorem). Additionally, the eigenvalues of a symmetric matrix are continuous functions of the entries of the matrix. Therefore, if the number of positive, negative, or zero eigenvalues of BAB^T differs from that of A, then there must be some t for which M_t has more zero eigenvalues than does A. But the matrices R_t are upper triangular with positive diagonal entries, and hence are non-singular (since the eigenvalues of R_t *are* the diagonal entries). Thus the rank of M_t must equal the rank of A, which means this cannot happen. □

Fiedler's Theorem will follow from an analysis of the eigenvalues of tri-diagonal matrices with zero row-sums. These may be viewed as Laplacians of weighted path graphs in which some edges are allowed to have negative weights.

Lemma 2. *Let M be an $n \times n$ symmetric matrix such that*

$$M\mathbf{1} = \mathbf{0}.$$

Then:

$$M = \sum_{i \neq j} -M_{ij} L_{G_{i,j}}. \tag{53}$$

Proof. Equation (53) is an equality between two matrices. Let A denote the right hand side matrix. On the off diagonal it is clear that both M (the LHS) and A (the RHS) are equal. Notice as well that the right hand side satisfies:

$$A\mathbf{1} = \sum_{i \neq j} -M_{ij} L_{G_{i,j}} \mathbf{1} = \sum_{i \neq j} -M_{ij} \mathbf{0} = \mathbf{0}.$$

Thus $M\mathbf{1} = \mathbf{0}$ and $A\mathbf{1} = \mathbf{0}$. Notice that these are sets of n equations and n unknowns (i.e., the n diagonal entries), which have unique solutions. Since the off diagonal entries of M and A are identical, the n equations are the same, and thus the solutions are as well, meaning that the diagonal of M and A are the same. □

Lemma 3. *Let M be a symmetric tri-diagonal matrix with $2q$ positive off-diagonal entries such that*

$$M\mathbf{1} = \mathbf{0}.$$

Then M has q negative eigenvalues.

Proof. By Lemma 2 and the fact that M is symmetric and tri-diagonal, we may write:

$$M = \sum_{i=2}^{n} -M_{i-1,i} L_{G_{i-1,i}}.$$

Thus for $v \in \mathbb{R}^n$,

$$v^T M v = \sum_{i=2}^{n} -M_{i-1,i} \left(v[i-1] - v[i] \right)^2.$$

Now we perform a change variables that will diagonalize the matrix M. Let $\delta[1] = v[1]$ and set $\delta[i] = v[i] - v[i-1]$ for $i \geq 2$, so that:

$$v[i] = \delta[1] + \delta[2] + \cdots + \delta[i].$$

Notice that if we define the lower triangular matrix T as:

$$T = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix},$$

then

$$v = T\delta.$$

By Sylvester's Law of Inertia (Theorem 7), we know that

$$A = T^T M T$$

has the same number of positive, negative and zero eigenvalues as M. On the other hand,

$$\delta^T A \delta = \delta^T T^T M T \delta = v^T M v = \sum_{i=2}^{n} -M_{i-1,i} \delta[i]^2.$$

Thus A has one zero eigenvalue (with eigenvector $\delta[1] = 1$, $\delta[j] = 0$ for all $j \geq 2$) and a negative eigenvalue $-M_{i-1,i}$ for each $M_{i-1,i} > 0$ (with eigenvector $\delta[i] = 1$, $\delta[j] = 0$ for all $j \neq i$), of which there are q. □

Proof of Theorem 6. We consider the case when φ_k has no zero entries. The proof for the general case may be obtained by splitting the graph by removing the vertices with zero entries. For simplicity, we also assume that λ_k has multiplicity 1.

Recall we wish to show that φ_k changes sign $k-1$ times. This is equivalent to showing that:

$$\#\{i = 1, \ldots, n-1 : \varphi_k[i]\varphi_k[i+1] < 0\} = k-1.$$

Let V_k denote the diagonal matrix with φ_k on the diagonal. Consider the matrix:

$$M = V_k^T(L_P - \lambda_k I)V_k.$$

The inner matrix $L_P - \lambda_k I$ has one zero eigenvalue and $k-1$ negative eigenvalues derived from the eigenvalues and eigenvectors of L_P. So, by Sylvester's Law of Inertia (Theorem 7), M has $k-1$ negative eigenvalues, one zero eigenvalue, and $n-k$ postitive eigenvalues.

We are now going to use Lemma 3. The matrix M is clearly symmetric and tri-diagonal, and additionally:

$$M\mathbf{1} = V_k^T(L_P - \lambda_k I)V_k\mathbf{1} = V_k^T(L_P - \lambda_k I)\varphi_k = V_k^T\mathbf{0} = \mathbf{0}.$$

Thus we can apply Lemma 3 to M. We note additionally that

$$M_{i,i+1} = -w(i, i+1)\varphi_k[i]\varphi_k[i+1],$$

and thus we see that $M_{i,i+1}$ is positive precisely when $\varphi_k[i]\varphi_k[i+1] < 0$. Since M has $k-1$ negative eigenvalues, by Lemma 3 it must have $k-1$ positive entries on the upper diagonal, which means that $\varphi_k[i]\varphi_k[i+1] < 0$ must occur for exactly $k-1$ indices. □

References

[1] Bernhard Schölkopf and Alexander J. Smola. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. Adaptive Computation and Machine Learning. The MIT Press, 2002.

[2] Afonso S. Bandeira. Ten lectures and forty-two open problems in the mathematics of data science. MIT course *Topics in Mathematics of Data Science*, 2015.

[3] Jon Shlens. A tutorial on principal component analysis. arXiv:1404.1100, 2014.

[4] Karl Pearson. On lines and planes of closest fit to systems of points in space. *Philosophical Magazine, Series 6*, 2(11):559–572, 1901.

[5] V. A. Marchenko and L. A. Pastur. Distribution of eigenvalues in certain sets of random matrices. *Mat. Sb. (N.S.)*, 72(114):507–536, 1967.

[6] J. Baik, G. Ben-Arous, and S. Péché. Phase transition of the largest eigenvalue for nonnull complex sample covariance matrices. *The Annals of Probability*, 33(5):1643–1697, 2005.

[7] Debashis Paul. Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, pages 1617–1642, 2007.

[8] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning*. Springer-Verlag New York, 2nd edition, 2009.

[9] Athanasios Tsanas and Angeliki Xifara. Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools. *Energy and Buildings*, 49:560–567, 2012.

[10] Daniel A. Spielman. Spectral graph theory. *Yale Course Notes*, Fall, 2009.