Applied spectral graph analysis—Exercises

Dario Fasino

This document contains warm-up and starred exercises for the course "Applied spectral graph analysis", held within the Rome-Moscow school on Matrix Methods and Applied Linear Algebra, September 2012. To obtain a positive evaluation students should submit the solution of starred exercises (as many as possible) by the end of September 2012 to the email adress `dario.fasino@uniud.it`.

Hereafter the following notations are used: X_{ij} is the (i,j)-entry of the matrix X; e_i is the ith canonical vector (i.e., the ith column of the identity matrix); e is a vector having all entries equal to 1; when a graph is considered, n denotes its size, and the graph is assumed to be (strongly) connected.

1 Warm-up exercises

1. If a node has no outlinks then its hub index is zero. Conversely, if a node has no inlinks then its authority index is zero. Why?

   Hint: note that if $\deg^+(i) = 0$ *then the ith row/column of the hub matrix is zero.*

2. Let L be a Laplacian matrix, and let L^+ denote its Moore-Penrose inverse. Prove that $L^+ = (L + \frac{1}{n}ee^T)^{-1} - \frac{1}{n}ee^T$.

3. The Randić index of a graph $\mathcal{G} = (V, E)$ whose degree sequence is $d_i = \deg(i)$, for $i = 1, \dots, n$, is the topological network invariant defined as

$$R(\alpha) = \sum_{(i,j)\in E} (d_i d_j)^\alpha \qquad (\alpha \neq 0).$$

   (a) Compute $R(\alpha)$ for a starred star graph having center degree d_c and star degree d_s.
   Answer: $R(a) = d_s^\alpha(d_s - 1 + d_c^{\alpha+1})$.

   (b) Prove the following inequality between $R(\alpha)$ and the spectral radius of the adjacency matrix for a general graph:

$$R(\alpha) \le \rho(A)\Big(\sum_{i=1}^n d_i^{2\alpha}\Big).$$

   Hint: let $v_i = d_i^\alpha$. *Expand* $(v^TAv)/(v^Tv) \le \rho(A)$.

2 Starred exercises

2.1 What is an irreducible matrix?

The matrix $A \in \mathbb{R}^{n\times n}$ is said to be *reducible* if there is a permutation matrix P such that the matrix $B = PAP^T$ is in (lower) block triangular form:

$$B = PAP^T = \begin{pmatrix} B_{11} & B_{12} \\ O & B_{22} \end{pmatrix},$$

diagonal blocks B_{11}, B_{22} being square matrices. An *irreducible* matrix is a matrix that is not reducible.

The *graph associated to* A is the digraph $\mathcal{G}(A) = (V, E)$ such that $V = \{1, \dots, n\}$ and

$$(i,j) \in E \iff A_{ij} \neq 0.$$

Thus, if the entries of A belong to the set $\{0, 1\}$ then A is the adjacency matrix of $\mathcal{G}(A)$. Recall that a (di-)graph is *strongly connected* when any two nodes are connected by a walk. Prove the following claim:

Theorem 2.1. *A matrix A is irreducible if and only if $\mathcal{G}(A)$ is strongly connected.*

Hint: Firstly, note that we can safely assume that the entries of A belong to $\{0, 1\}$. Remind the relationship between entries of A^k and walks of length k in $\mathcal{G}(A)$; and that the powers of a block triangular matrix are still block triangular.

2.2 A closeness centrality index

Let $\mathcal{G}$ be a graph and let L be its Laplacian matrix. The *resistance distance* between nodes i and j is given by

$$R(i, j) = (e_i - e_j)^T L^+ (e_i - e_j).$$

Define the *current-flow closeness centrality* of node i as

$$C(i) = \frac{1}{n} \sum_{j=1}^{n} R(i, j).$$

1. Prove that the resistance distance is really a distance.
2. Prove that
$$C(i) = (L^+)_{ii} + \frac{1}{n} \operatorname{trace}(L^+).$$

2.3 HITS scores on a modified network

Suppose that in the digraph $\mathcal{G}$ node j has no inlinks ($\deg^-(j) = 0$). Let i ($i \neq j$) be a node having positive hub score ($h_i > 0$). We modify $\mathcal{G}$ by adding the oriented edge (i, j). Prove the strict inequality

$$\forall \ell \neq i, \quad \frac{\bar{h}_\ell}{h_\ell} < \frac{\bar{h}_i}{h_i},$$

where $\bar{h}_i$ is the updated hub score of node i, which shows that, for any fixed normalization, the hub score of node i gets the largest relative increase.

Hint: find a relationship between original and updated hub matrix; prove that the spectral radius of the updated hub matrix is strictly greater than the original.

2.4 A workload balancing process

To every node of a graph we allocate an initial *workload* $x_i(0) \in \mathbb{R}$, $i = 1, \ldots, n$. The workload is then updated at discrete time instants acording to the following rule: At every time step, every node sends a fixed fraction $0 < \alpha < 1$ of its current workload to its neighbors (and receives their workload fractions):

$$x_i(t+1) = x_i(t) + \alpha \sum_{j : i \sim j} (x_j(t) - x_i(t)).$$

1. Prove that $\sum_{i=1}^{n} x_i(t)$ is constant over time.
2. Describe the stationary workload distributions.
3. Find a necessary and sufficient condition (in terms of the Laplacian spectrum) so that any initial workload distribution evolves toward a stationary distribution.

`Last update: September 17, 2012`