Applications of Stochastics — Exercise sheet 2

Gábor Pete
http://www.math.bme.hu/~gabor

January 16, 2020

Bonus exercises are marked with a star. They can be handed in for extra points.

The first exercise was needed in the analysis of $\nu_t := N_k(t)/t$, the expected ratio of degree k vertices at time t of the preferential attachment graph process.

▷ **Exercise 1.** Let $\beta, \gamma > 0$ be constants, and let $\nu_1, \nu_2, \dots$ be positive reals satisfying the recursion

$$\nu_{t+1} = \left(1 - \frac{\beta}{t}\right) \nu_t + \frac{\gamma}{t}$$

for all $t \geq T > \beta$, starting with some $\nu_T \geq 0$. Prove that $\lim_{t\to\infty} \nu_t = \frac{\gamma}{1+\beta}$. In more detail:

(a) Notice that, if $\nu_t < \frac{\gamma}{1+\beta}$, then $\nu_{t+1} > \nu_t$. Similarly, if $\nu_t > \frac{\gamma}{1+\beta}$, then $\nu_{t+1} < \nu_t$.

(b) Show that there can not exist an $\epsilon > 0$ such that $\nu_t < \frac{\gamma}{1+\beta} - \epsilon$ for all large enough $t > t_0(\epsilon)$. Similarly, there is no $\epsilon > 0$ such that $\nu_t > \frac{\gamma}{1+\beta} + \epsilon$ for all large enough $t > t_0(\epsilon)$.

(c) Show that $\nu_{t+1} - \nu_t \to 0$.

(d) Deduce the existence and value of the limit.

▷ **Exercise 2.** Recall that we defined the **clustering coefficient** of an undirected graph as

$$\mathsf{CC} := \frac{\#\text{ paths of length 2 with endpoints connected by an edge}}{\#\text{ paths of length 2}}.$$

With n vertices and $10n$ edges, find a graph with small CC, and another one with large CC.

Linear algebra brush-up:

▷ **Exercise 3.** For $u, v \in \mathbb{C}^n$ column vectors, define the inner product $(u, v) := u^T \overline{v}$, where $\overline{v}$ is coordinate-wise complex conjugation. Let A be a symmetric $n \times n$ real matrix.

(a) Show that $(v, u) = \overline{(u, v)}$, and $(Au, v) = (u, Av)$. Deduce that if $v \in \mathbb{C}^n$ is an eigenvector of A with eigenvalue λ, then $\lambda \in \mathbb{R}$.

(b) From the fundamental theorem of algebra we know that $\det(A - \lambda I)$ has a root $\lambda \in \mathbb{C}$. Recall that this implies that there exists a nonzero $v \in \mathbb{C}^n$ in the kernel of $A - \lambda I$, hence λ is an eigenvalue, with eigenvector v.

(c) Show that $v^\perp := \{u \in \mathbb{C}^n : (u, v) = 0\}$ is a linear subspace, and $Av^\perp \subseteq v^\perp$.

(d) Prove by induction that A has an orthonormal basis of eigenvectors $v_1, \dots, v_n \in \mathbb{C}^n$, with all real eigenvalues $\lambda_1, \dots, \lambda_n$.

(e) Show that A also has an orthonormal basis of eigenvectors $u_1, \dots, u_n \in \mathbb{R}^n$, with the same eigenvalues.

If G is an undirected simple graph on the vertex set $\{1, \dots, n\}$, then its adjacency matrix A is a real symmetric $n \times n$ matrix, and $P = D^{-1}A$ is the Markov transition matrix of the associated simple random walk, where D is the diagonal matrix formed by the degrees $\deg(i)$.

▷ **Exercise 4.**

(a) Show that $\|Pf\|_\infty \leq \|f\|_\infty$ holds for any $f : V \longrightarrow \mathbb{R}$ (considered as a row vector).

(b) Deduce from part (a) that all the eigenvalues of P have absolute value at most 1.

(c) Note that it is clear what $D^{-1/2}$ means. Show that $B = D^{-1/2}AD^{-1/2}$ is a symmetric matrix that is conjugate to P, hence has the same eigenvalues. Deduce that all the eigenvalues of B are real, are between 1 and -1, and that the vector $\left(\sqrt{\deg(i)}\right)_{1\leq i\leq n}$ is an eigenvector for the eigenvalue 1.

(d) Find a left eigenvector and a right eigenvector for P with eigenvalue 1.

Remark. Graph theorists prefer B to P because it is symmetric, and sometimes to A because it is normalized to have spectrum between 1 and -1.

▷ **Exercise 5.** Let P be the Markov transition matrix for the simple random walk on a finite undirected simple graph G. Write $-1 \leq \lambda_n \leq \cdots \leq \lambda_1 = 1$ for its eigenvalues (see the previous exercise).

(a) Show that $\lambda_2 < 1$ if and only if G is connected (the chain is irreducible), and this is precisely when P has a unique stationary distribution.

(b) Show that $\lambda_n > -1$ if and only if G is not bipartite. (Recall here the easy lemma that a graph is bipartite if and only if all cycles are even.)

(c) Let $\pi_t := \pi_0 P^t$ be the distribution of the random walker after t steps. Show that π_t converges coordinate-wise to the unique stationary distribution precisely when $\lambda_2 < 1$ and $\lambda_n > -1$.

Now back to general directed graphs and their associated Markov transition matrix P.

▷ **Exercise 6.** In Google's **PageRank**, the iteration $\overline{x}_{t+1} := \alpha\overline{x}_t P + (1-\alpha)\mathbf{1}$ is used, with some $\alpha \in (0,1)$. Show that, for any starting vector $\overline{x}_0$, the sequence $\overline{x}_t$ converges to $(1-\alpha)\mathbf{1}\,(I-\alpha P)^{-1}$.

(Hint: use the Banach fixed point theorem, with an appropriate notion of distance; see part (a) of Exercise 4. In order to have a strict contraction, don't forget to use that $\alpha < 1$. Also, note that part (b) implies that $I - \alpha P$ is invertible for any $\alpha \in (0,1)$.)

▷ **Exercise 7.** Consider the undirected graph on the vertex set $\{1,2,3,4\}$, where $1,2,3$ form a triangle, and 1 and 4 are also connected by an edge.

(a) Calculate the Eigenvector centrality of the four vertices.

(b) Calculate the PageRank scores, for several values of α.

You are welcome to use Mathematica or other software.

▷ **Exercise 8.** Let G be a directed graph on 3 vertices, where there is an undirected path through vertices $1,2,3$, plus a directed edge from 1 to 3. Let A be its adjacency matrix.

(a) Find the eigenvalues of A and an orthonormal basis of eigenvectors.

(b) Consider the iteration $\overline{x}_{t+1} := \overline{x}_t A$, with $\overline{x}_0 = \mathbf{1}$. Find a sequence of scalars c_t such that $c_t\overline{x}_t$ converges to a nonzero vector.