

Identifying a finite graph by its random walk

Heinrich v. Weizsäcker, Kaiserslautern

In the following we illustrate by two examples and simplify the statement of the main result of a joint paper with Peter Scheffel [3].

Let $\Gamma_0 = (S, E_0)$ and $\Gamma_1 = (S, E_1)$ be two connected directed graphs with the same vertex set S and the two sets E_0, E_1 of edges. Suppose we observe $n + 1$ points X_0, \dots, X_n of S which are produced by the random walk on one of the two graphs. We want to infer which graph was used and the mathematical goal is to compute the asymptotic behaviour of the error probabilities. As soon as the walk makes a step which is impossible for one of the two graphs one knows it was the other graph. On the other hand, if all previous steps were possible for both models then the number of competing possibilities becomes important.

This is a particular case of the following problem. Let π_0 and π_1 be two finite irreducible Markov transition matrices on the same state space. Fix an initial distribution μ and let $P_i^{(n)}$ denote the law on S^{n+1} of the Markov chain with initial measure μ and transition matrix π_i . Clearly, the two laws $P_0^{(n)}$ and $P_1^{(n)}$ become more and more singular to each other. The following result determines the exponential rate at which the overlap $2 - \|P_0^{(n)} - P_1^{(n)}\|$ (which can also be described as the sum of the error probabilities of the natural likelihood test) converges to zero. For every matrix A the symbol $\rho(A)$ denotes its spectral radius.

Theorem 1 *The laws $P_0^{(n)}$ and $P_1^{(n)}$ become singular at the rate*

$$r = \lim_{n \rightarrow \infty} \frac{1}{n} \log(2 - \|P_0^{(n)} - P_1^{(n)}\|) = \max_{I \in \mathcal{S}} \inf_{0 < t < 1} \log \rho(\pi_{t,I})$$

where $\pi_{t,I}$ is the elementwise logarithmically convex combination

$$(\pi_0(i, j)^{1-t} \pi_1(i, j)^t)_{i, j \in I}$$

and \mathcal{S} denotes the system of all subsets I of S which are maximal with respect to the property that I can be completely covered by a single path which has positive probability under both models.

The proof uses an extension of the Large Deviation Theorem for the empirical pair distribution of ergodic Markov chains (cf. e.g. [2]). This

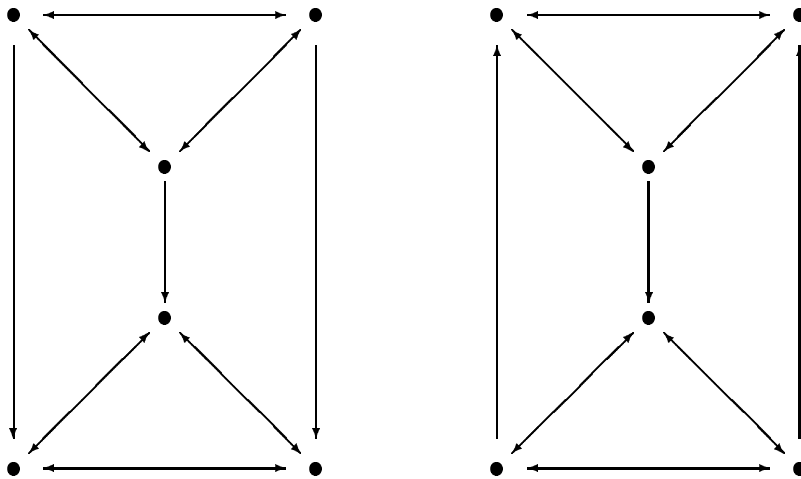
extension is needed for those cases in which some of the zero entries of one matrix are positive in the other matrix. This is always true in the graph problem. The above definition of the system \mathcal{S} is much simpler than in [3].

If both π_0 and π_1 are strictly positive, more generally if one can get from every point in S to every other point by transitions which are possible for both π_0 and π_1 , then the only element of the system \mathcal{S} is the set S itself and the result simplifies accordingly. As an example for $|S| = 3$ consider the three matrices

$$\pi_0 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \quad \pi_1 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{3}{4} & 0 \end{pmatrix}, \quad \pi_2 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \\ \frac{2}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

The first matrix differs from π_1 strongly in a single row and from π_2 in two rows but not so strongly. Intuitively it is not clear which pair of Markov chains is better separated asymptotically. A numerical calculation of the corresponding spectral radii show that the rate of separation is given by $r \approx -.0115$ when comparing π_0 and π_1 , and by $r \approx -.0096$ when comparing π_0 and π_2 . This shows that the matrix π_0 is, empirically, more easily separated from π_1 than from π_2 .

Now let us consider the following two directed graphs Γ_0 (left) and Γ_1 (right).



The only difference between the two graphs consists in the direction of the two long vertical edges. The random walk on these graphs leads to the

transition matrices

$$\pi_0 = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}, \quad \pi_1 = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

and hence we have

$$\pi_t = \begin{pmatrix} 0 & \frac{3^{t-1}}{2^t} & \frac{3^{t-1}}{2^t} & 0 & 0 & 0 \\ \frac{3^{t-1}}{2^t} & 0 & \frac{3^{t-1}}{2^t} & 0 & 0 & 0 \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{2^{t-1}}{3^t} & 0 & \frac{2^{t-1}}{3^t} \\ 0 & 0 & 0 & \frac{2^{t-1}}{3^t} & \frac{2^{t-1}}{3^t} & 0 \end{pmatrix}.$$

Consider a finite path which visits first all vertices in the upper part, then descends via the middle edge and finally visits all lower vertices. Such a path has positive probability under both models, provided the starting point has positive weight for the initial distribution μ . Therefore in this case the family \mathcal{S} contains (only) the full set S . Thus one has to compute the spectral radius of π_t and then pass to the infimum over t . Due to the block structure of π_t , its spectral radius is the maximum of the spectral radii of the upper left 3×3 -submatrix and the similar lower right submatrix. It is easily seen that the upper left spectral radius is increasing in t and the lower right spectral radius decreases in t . Therefore the infimum in t of the maximum of these two functions is attained at that value of t at which the two values coincide.

If however the initial distribution is concentrated on the lower triangle then the family \mathcal{S} contains (only) the set which consists of these three lower points. In this case one has to consider only the lower right submatrix of π_t . Due to the monotonicity mentioned above the rate r in this case is given by the logarithm of the spectral radius of the lower right submatrix of π_1 .

In the general graph problem the set \mathcal{S} typically is of a more complex structure. It is easy to construct examples in which it contains two different sets which then will automatically have nontrivial relative complements with respect to each other.

References

- [1] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis on the sum of observations. *Annals of Mathematical Statistics*, 23:493–507, 1952.
- [2] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Jones and Bartlett Publishers, Boston etc., 1993.
- [3] P. Scheffel and H.v. Weizsäcker. On Risk Rates and Large Deviations in Finite Parameter Markov Chain Experiments. *Math. Methods of Statistics*, 3:293–312, 1997.