

THE EMBEDDING PROBLEM FOR MARKOV MATRICES

MARTA CASANELLAS, JESÚS FERNÁNDEZ-SÁNCHEZ,
AND JORDI ROCA-LACOSTENA

Abstract: Characterizing whether a Markov process of discrete random variables has a homogeneous continuous-time realization is a hard problem. In practice, this problem reduces to deciding when a given Markov matrix can be written as the exponential of some rate matrix (a Markov generator). This is an old question known in the literature as the *embedding problem* [11], which has been solved only for matrices of size 2×2 or 3×3 . In this paper, we address this problem and related questions and obtain results along two different lines. First, for matrices of any size, we give a bound on the number of Markov generators in terms of the spectrum of the Markov matrix. Based on this, we establish a criterion for deciding whether a generic (distinct eigenvalues) Markov matrix is embeddable and propose an algorithm that lists all its Markov generators. Then, motivated and inspired by recent results on substitution models of DNA, we focus on the 4×4 case and completely solve the embedding problem for any Markov matrix. The solution in this case is more concise as the embeddability is given in terms of a single condition.

2020 Mathematics Subject Classification: 60J10, 60J27, 15B51, 15A16.

Key words: Markov matrix, Markov generator, embedding problem, rate identifiability.