Linear Algebra and its Applications 520 (2017) 306–334

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Minimising the largest mean first passage time of a Markov chain: The influence of directed graphs ☆

Jane Breen *, Steve Kirkland

Department of Mathematics, University of Manitoba, Winnipeg MB, R3T 2N2, Canada

ARTICLE INFO

Article history:
Received 30 November 2016
Accepted 26 January 2017
Available online 30 January 2017
Submitted by R. Brualdi

MSC:
15B51
05C50
60J10

Keywords:
Markov chain
Mean first passage time
Stochastic matrix
Directed graph
Hessenberg matrix

ABSTRACT

For a Markov chain described by an irreducible stochastic matrix A of order n, the mean first passage time $m_{i,j}$ measures the expected time for the Markov chain to reach state j for the first time given that the system begins in state i, thus quantifying the short-term behaviour of the chain. In this article, a lower bound for the maximum mean first passage time is found in terms of the stationary distribution vector of A, and some matrices for which equality is attained are produced. The main objective of this article is to characterise the directed graphs for which any stochastic matrix A respecting this directed graph attains equality in this lower bound, producing a class of Markov chains with optimal short-term behaviour.

☆ Funding: This work was supported in part by NSERC grant no. RGPIN-2014-06123 and by the University of Manitoba Graduate Fellowship.

* Corresponding author.
E-mail addresses: breenj3@myumanitoba.ca (J. Breen), Stephen.Kirkland@umanitoba.ca (S. Kirkland).

http://dx.doi.org/10.1016/j.laa.2017.01.034
0024-3795/© 2017 Published by Elsevier Inc.

1. Introduction

A stochastic matrix A is an entrywise nonnegative matrix whose rows sum to 1, i.e. $A\mathbb{1} = \mathbb{1}$, where $\mathbb{1}$ represents the vector of all ones. Stochastic matrices are central to the study of Markov chains, which are a type of probabilistic model describing dynamical systems which move between some finite number of states in discrete time-steps, where these transitions between states depend only on the present state occupied by the system. The connection with stochastic matrices is the following: given a Markov chain describing a system with a finite state space indexed by the integers $\{1, 2, \ldots, n\}$, construct a matrix such that the (i, j)th entry is the probability of the system transitioning from state i to state j in a single time step. This is a stochastic matrix, referred to as the *probability transition matrix*, or simply *transition matrix* of the chain. This transition matrix A wholly represents the Markov chain, in that given an initial vector u_0 describing the probabilities that the process is in one of the various states at time 0, the probability distribution across all states after k time-steps is the vector $u_0^\top A^k$, for $k \geq 1$.

Representing a Markov chain by a stochastic matrix in this way enables us to analyse the long-term behaviour of the modelled system using basic techniques from linear algebra. If the transition matrix A is *irreducible* – i.e. for any pair of indices i, j, there exists some $m \in \mathbb{N}$ such that the (i, j)th entry of A^m is positive – then by the Perron–Frobenius theorem, A must have a strictly positive left eigenvector $w = [w_1\, w_2 \cdots w_n]^\top$ corresponding to the eigenvalue 1. This eigenvector, when normalised so that the entries sum to 1 (thus producing a probability distribution) is referred to as the *stationary distribution vector* of the chain. It is an important quantity for the following reason: in the case that the transition matrix A is also *primitive* – i.e. there exists some $m \in \mathbb{N}$ for which every entry of A^m is positive – the iterates of the chain converge to $w^\top$ independent of the initial distribution. This is powerful in analysing the underlying system, since we can then say that the probability the Markov chain is in the ith state in the long term is the ith entry w_i of this eigenvector $w^\top$. Thus the long-term behaviour of the modelled system is summarised by a fundamental feature of the corresponding stochastic matrix.

The short-term behaviour of a system modelled by a Markov chain is considered as follows. Define $F_{i,j}$ to be a random variable representing the first passage time from state i to state j; i.e. the number of time steps elapsed (≥ 1) before the system reaches state j for the first time, given that it began in state i. The expected value of this random variable, then, is a key quantity of interest. It is referred to as the *mean first passage time from i to j*, denoted $m_{i,j}$. In the special case that $i = j$, $m_{i,i}$ is referred to as the *mean first return time* to state i. This facilitates the construction of the *matrix of mean first passage times* $M = [m_{i,j}]$ which is the unique solution (see [15, Section 6.1]) to the equation

$$M = A(M - M_{diag}) + J, \tag{1.1}$$