Two Methods for Estimating a Markov Transition Matrix from Subsampled Data

Samantha Massengill

July 7, 2011

1 Introduction

This work was done at the Autonomous Newtorks Research Group at the University of Southern California under the supervision of Professor Bhaskar Krishnamachari[1]. In this work, we use adaptive algorithms to estimate 2-state Markovian transition probabilities from a subsampled sequence when these probabilities are not known *a priori*. A maximum likelihood estimation (MLE) problem is set up, and solved using two different methods: the expectation maximization (EM) algorithm, and an exhaustive search through a finite number of possible transition matrices.

2 Maximum Likelihood Estimation

2.1 Expectation Maximization (EM) Algorithm

The EM algorithm is a learning technique that consists of two steps: the E-step and the M-step. In the E-step, an expected value of the log likelihood function constructs a complete set of data (a Markov chain) using an initial estimate of the transition matrix. In the M-step, a global maximum for this expected value of the log likelihood function is chosen and a new estimate of the transition matrix is found. The E-step and M-step are repeated until convergence and this limit is the maximum likelihood estimate of the transition matrix. C. Sherlaw-Johnson, et. al., [1] outlines how to implement the EM algorithm for the specific case of estimating a Markov transition matrix.

The MATLAB function `EM` takes the inputs `iter_max`, `P`, and `Y`, where `iter_max` is the number of iterations for the EM algorithm, `P` is the initial estimate for the transition matrix, and `Y` is a row vector representing a subsampled 2-state Markov chain, where 1s and 2s are states and 0s are unknown. The function returns a new estimate for `P` after `iter_max` iterations of the EM algorithm. This function calls the function `Omnt` which inputs `Y` and returns `m`, `n`, and `t` vectors, which define the Markov chain in terms of all combinations of going from state m to state n in t time slots.

2.2 Brute-Force Maximization: An Exhaustive Search

A brute-force method is used to exhaustively maximize the log likelihood function: $log(g(Y|P)) = \sum_i \sum_j \sum_t O_{ijt} log(P^t)_{ij}$. Although this is too difficult to maximize analytically, a finite number of transition matrices are searched through, depending on how many (user-specifiable) values are used for the probabilities.

The MATLAB function `exhaust` takes as input `Y`, a row vector representing a subsampled 2-state Markov chain, again with 1s, 2s, and 0s. It returns an estimate for `P` that maximizes the log likelihood function. This function also calls the function `Omnt`.

References

[1] C. Sherlaw-Johnson, S. Gallivan, J. Burridge (1995) Estimating a Markov Transition Matrix from Observational Data, *Journal of the Operational Research Society* **46**, 405-410.

[1] http://anrg.usc.edu/