Global estimation of Hidden Markov Model parameters via interval arithmetic

Tiago de M. Montanher and Walter F. Mascarenhas

January 10, 2014

Abstract

Hidden Markov Models are important tools in statistics and applied mathematics, with applications in speech recognition, physics, mathematical finance and biology. The Hidden Markov Models we consider here are formed by two discrete time and finite state stochastic process. The first process is a Markov chain (A, π) and is not observable directly. Instead, we observe a second process B which is driven by the hidden process. For instance, a Markov chain is a simple Hidden Markov Model in which the observed process and the hidden process are the same. These models have received much attention in the literature in the past forty years, and [1] presents a good didactic overview of this subject. From a historical perspective, the seminal paper by Rabiner [5] provides a good motivation for this subject.

In order to extract conclusions from a Hidden Markov Models we must estimate the parameters defining the hidden process (A, π) and the observed process B. In this article we present efficient global optimization techniques to estimate these parameters by maximum likelihood and compare our estimates with the ones obtained by the local likelihood maximization methods already described in the literature. Usually, this estimation problem is solved by local methods, like the Baum-Welch algorithm [6]. These methods are efficient, however they only find local maximizers and do not estimate the distance from the resulting parameters to global optima. Our work aims to improve this situation in practice.

In order to compute the global maximizers efficiently, we introduce an optimized version of Interval arithmetic, which is particularly appropriate for the estimation of Hidden Markov Models (and of more general probabilistic models.) Interval arithmetic is a natural tool to develop global optimization algorithms. It provides rigorous bounds for computation of complicated functions and allows an extension of branch and bound techniques of combinatorial optimization to the continuous case. The practical realization of interval arithmetic requires the ability to change rounding modes on computers or at least to simulate such changes. Our approach to interval arithmetic allows us to reduce significantly the cost of these changes in rounding mode.

The general theory of global optimization using interval analysis is presented, for instance, in [3], [2] and [7]. Here we develop a global optimization algorithm based on interval branch and bound framework to solve the estimation problem for Hidden Markov Models. Our algorithm starts with a local Baum-Welch method, which provides a warm lower bound for the problem. We then derive KKT conditions to obtain a nonlinear system of equations that can be solved without handling interval

Newton methods. Our algorithm is able, in a successful execution, to find a box with prescribed width which rigorously contains at least one feasible point x^* for the problem and such that x^* is a $\epsilon-$-global maximum. This algorithm is a modification of that described in [3] with additional cut-off tests due to conditions of Lagrange multipliers for this specific problem. We also apply the tests described in [4] to ensure existence of a feasible point.

As stated in [5], the objective function for this problem can be evaluated by the so called backward and forward recursions. In fact we can use only one of these recursions or we can combine both to evaluate function and its derivatives. These three formulations are equivalent using exact arithmetic. However they will usually be different in interval arithmetic due to the lack the distributivity law. In order to accelerate the convergence of upper bound of the global maximum we implement and compare interval extensions for the forward, backward and forward-backward equations and their respective derivatives of first order. Our implementation is optimized by taking into account the fact that functions are polynomials with positive coefficients and our variables always lies in the interval $[0, 1]$. Hence a novel positive interval arithmetic is introduced and compared with traditional implementations. Moreover, our implementation handles the underflow problems which arise frequently in the estimation problem for Hidden Markov models. We present the results of numerical experiments illustrating the effectiveness of our approach.

References

[1] Olivier Cappé, Eric Moulines, and Tobias Ryden. *Inference in Hidden Markov Models (Springer Series in Statistics)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005.

[2] E. R. Hansen and W. G. Walster. *Global Optimization Using Interval Analysis*. Marcel Dekker and Sun Microsystems, New York, second edition, 2004.

[3] R. B. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer Academic, Dordrecht, first edition edition, 1996.

[4] R. E. Moore. A test for existence of solutions to nonlinear systems. *SIAM Journal of Numerical Analysis*, 14(4):611–615, 1977.

[5] Lawrence R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, pages 257–286, 1989.

[6] S.E. Levinson; L.R. Rabiner and M.M. Sondhi. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *Bell Syst. Tech. J.*, 62:1035–1074, 1983.

[7] H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization*. Horwood Limited, Chichester, first edition edition, 1988.