Using Quasi Renewal Processes to Investigate Feature Distributions in Markov Switching Models

JAD Aston† and JY Peng, Academia Sinica

DEK Martin, North Carolina State University

Abstract

Markov switching models are widely used in the analysis of nonlinear time series. However, evaluation of distributions such as the number of regime switches, the maximal length of regimes or other feature distributions related to the data can be problematic as they are functions of the entire state sequence. Direct computation is not possible given the exponential number of possible state sequences. Estimates can be made of these quantities by using surrogates such as the most probable sequence, but these yield no distributional information about the features, nor is it easy to evaluate the properties of the estimates themselves. Here it will be shown that by defining appropriate auxiliary processes, which will be shown to be quasi renewal processes, computation of exact distributions for many features, including those above, is possible. This computation is then not only possible, but very efficient given the almost linear algorithm that can be defined based on the renewal properties. These processes can be found for many commonly used Markov switching models, and will be demonstrated using Hamilton's classic Markov switching AR(4) model used to find recessions in GNP data. It will be shown that distributional information can be found for some of the features qualitatively mentioned in the original paper.

1 Preliminaries

The methods which will be examined here can be applied to Markov switching models [3] with the general form

$$\begin{aligned} y_t &\sim f(S_{t-r:t}, y_{1:t-1}), \\ P[S_t|S_{-r+1:t-1}] &= P[S_t|S_{t-1}], \quad t = 1, \ldots, n. \end{aligned} \tag{1}$$

The data y_t is distributed conditional on previous data and r previous switching states $S_{t-r}, \ldots, S_{t-1}$ in addition to the current state S_t (as well as other parameters which are implicitly assumed here).

The waiting time probability for a regime period (run) of length at least $k \geq 1$ will be defined as the following

$$P[W(k,1) \leq t] = P[S_{t_1-k+1} = s, \ldots, S_{t_1} = s, t_1 \leq t], \quad t_1 \geq k \tag{2}$$

where $W(k,1)$ is the waiting time for a regime period of length at least k, and s is the state of the regime period.

2 Distributions Associated with Regimes in Markov Switching Models

The non-negative integer-valued random variable $X_n(\Lambda)$ is *finite Markov chain imbeddable* if

- there exists a finite Markov chain $\{Z_t : t \in 0, \ldots, n\}$ defined on a finite state space Ω with initial probability vector ξ_0,
- there exists a finite partition $\{C_x : x = 0, 1, \ldots, l_n\}$ on the state space Ω, and
- for every $x = 0, 1, \ldots, l_n$, we have

$$\mathbb{P}(X_n(\Lambda) = x) = \mathbb{P}(Z_n \in C_x | \xi_0)$$

from [2].

A Markov chain $\{Z_t\}$ is developed such that there is a one-to-one correspondence between classes of its states and those of $\{S_t\}$. The Markov chain $\{Z_t\}$ facilitates the computation of waiting time distributions for runs, since an absorbing state for $\{Z_t\}$ can be defined to correspond to the occurrence of the run of interest. The desired probability may then be computed by determining the probability that $\{Z_t\}$ lies in the corresponding absorbing state, which is computed using basic properties of Markov chains [2].

Lemma 1 - Smoothed Transition Probabilities

For a Markov switching model with rth order lag dependence on $\{S_t\}$, the smoothed conditional probabilities of $\{S_t\}$ form an rth order Markov process:

$$P[S_t|S_{-r+1:t-1}, y_{1:n}] = P[S_t|S_{t-r:t-1}, y_{1:n}] \tag{3}$$

Theorem 2 - Waiting Time Distribution for Regimes in MS Models

Consider an r-th order MS Model as defined above. If $\{Z_t\}_{t=0}^{T^*}$ is a Markov chain such that, conditioned on $Y^{(T)}$, the desired regime has occurred by time t if and only if Z_t lies in a corresponding absorbing state $\Gamma \in S_Z$, then conditional on $Y^{(T)}$, the waiting time distributions for regimes may be computed through $\{Z_t\}$ by the equation

$$P[W(k,1) \leq t] = P[Z_t = \Gamma] = \psi_0 \left(\prod_{j=1}^{t} M_j \right) U(\Gamma), \tag{4}$$

where M_j is the transition probability matrix holding transitions from Z_{j-1} and Z_j, ψ_0 is a row vector holding the initial distribution of Z_0, and $U(\Gamma)$ is a column vector with a one in the location corresponding to absorbing state Γ, and zeros elsewhere.

3 Quasi-Renewal Processes and Multiple Occurrences

Define $W(k,m)$ to be the waiting time of the mth run of at least length k occurring in the data. The additional state $(\overbrace{s \ldots s}^{r}, -1) = C$, called the continuing state, will provide the initialisation for the mth run conditional on the $m-1$ previous occurrences. Thus, instead of an initial vector, an initial matrix will be used, defined as

$$\Psi_0 = (\psi_0 | \overbrace{\underline{0} | \ldots | \underline{0}}^{m-1})$$

where $\underline{0}$ refers to a row vector of zeros and ψ_0 is the initial vector in the previous subsection, and $(a|b) = \begin{pmatrix} a \\ b \end{pmatrix}$, with one row vector stacked upon another. Having setup the state space in this way, the following theorem results

Theorem 3 - Waiting Time Distribution for Multiple Regimes in MS Models

For a MS Model defined as above, the waiting time distribution for multiple occurrences,

$$P[W(k,i) \leq t] = \Psi_t(i, \Gamma), \quad i = 2, \ldots, m. \tag{5}$$

where the Markov transition, used in Eq (4), is replaced with the

$$\Psi_t = \Psi_{t-1} M_t, \tag{6}$$

$$\Psi_t(i, C) \leftarrow \Psi_t(i, C) + P[W(k, i-1) = t], \quad i = 2, \ldots m, \tag{7}$$

Proof.

Let $\phi_{0,t}^{(i)}$ be the initial distribution for the Markov chain associated with the i'th occurrence, $i = 2, 3, \ldots, r$ when the chain starts at time t. Since $i > 1$, the chain will start in one of the continuation states. Define $W\{\Lambda^i(k)\}$ to be the waiting time for the i'th occurrence of the pattern. Then

$$\begin{aligned} P[W\{\Lambda^i(k)\} \leq t] = & \sum_{0 \leq t_1 \leq t} P[W\{\Lambda^{i-1}(k)\} = t_1] \\ & \times P[W\{\Lambda^i(k)\} \leq t | W\{\Lambda^{i-1}(k)\} = t_1]. \end{aligned} \tag{8}$$

This can be simplified using a renewal argument. The second term in the summation in (8) is the same as $P[W\{\Lambda(k)\} \leq (t - t_1)]$ assuming that the chain starts at time t_1 (the chain is inhomogeneous, and thus the starting time is needed), as the probability of the i'th occurrence conditioned on the $(i-1)$'st can be calculated using the same probability structures as those of the first occurrence (but starting in one of the continuation states). Hence (8) equals

$$\sum_{0 \leq t_1 \leq t} P[W\{\Lambda^{i-1}(k)\} = t_1] P[W\{\Lambda(k)\} \leq (t - t_1)] \tag{9}$$

$$= \sum_{0 \leq t_1 \leq t} P[W\{\Lambda^{i-1}(k)\} = t_1] \phi_{0,t_1}^{(i)} \left(\prod_{j=t_1+1}^{t} M_j \right) U(\Gamma) \tag{10}$$

by Theorem 2, where the product is taken to be an identity matrix if $t_1 = t$.

Define

$$\begin{aligned} \psi_{0,t_1}^{(i)} &= P[W\{\Lambda^{i-1}(k)\} = t_1] \phi_{0,t_1}^{(i)} & (11) \\ &= (P[W\{\Lambda^{i-1}(k)\} \leq t_1] - P[W\{\Lambda^{i-1}(k)\} \leq t_1 - 1]) \phi_{0,t_1}^{(i)} & (12) \end{aligned}$$

Then (10) equals

$$\sum_{0 \leq t_1 \leq t} \psi_{0,t_1}^{(i)} \prod_{j=t_1+1}^{t} M_j U(\Gamma) \tag{13}$$

which when expanded out equals

$$(\overbrace{\ldots(}^{t} \psi_{0,0}^{(i)} M_1 + \psi_{0,1}^{(i)}) M_2 \ldots + \psi_{0,t-2}^{(i)}) M_{t-1} + \psi_{0,t-1}^{(i)}) M_t + \psi_{0,t}^{(i)}) U(\Gamma) \tag{14}$$

where the probability of seeing the $(i-1)$'st occurrence at time 0 is defined to be 0 and hence $\psi_{0,0}^{(i)} = 0$. If $\psi_{0,0}^{(1)} \equiv \psi_0$ and $\psi_{0,t}^{(1)} \equiv 0, t > 0$ then the algorithm given in (6)-(7) iteratively calculates the set of sums and products of each bracket given in (14) for all $i = 1, \ldots, r$ occurrences simultaneously in i before the multiplication with $U(\Gamma)$.

□

4 Feature and Change-Point Distributions

Many distributions can then be found directly from the waiting time distributions.

4.1 Number of Regimes in Data

The distribution of the number of regimes $R(i)$ in the data is given by

$$R(i) = P[W(k,i) \leq n] - P[W(k, i+1) \leq n] \text{ for } i = 0, \ldots, n/2. \tag{15}$$

4.2 Longest Regime Length

The distribution of the length $L(k)$ of the longest regime in the data is given by

$$L(k) = P[W(k,1) \leq n] - P[W(k+1,1) \leq n] \text{ for } k = 0, \ldots, n. \tag{16}$$

4.3 General Change Points

It has been shown that HMM's and Regime Switching Models are very flexible models for change points [1]. The distribution of the i'th change-point to a regime of length at least k is

$$\text{CP}_i(t) = P[W(k,i) = t - k + 1] \tag{17}$$

The probability of a similarly defined change point at time t is thus

$$\text{CP}(t) = \sum_i P[W(k,i) = t - k + 1] \tag{18}$$

and the probability of a similarly defined change point within an interval

$$\text{CP}(t_1 : t_2) = \sum_i \sum_{t=t_1}^{t_2} P[W(1,i) = t - k + 1 | W(1, i-1) < t_1 - k + 1] \tag{19}$$

5 Analysis of GNP Data

The logged differenced quarterly GNP data (1951.II to 1984.IV) from [3] is used as an example. This is modelled with a Markov Switching AR(4) model with parameter values estimated by ML.

i	Peak (t_1)	$P[W(2,i) = t_1+1]$	$P[W(2,i) < t_1+1]$	RSP$(t_1)^a$	Trough (t_2)	$P[W_e(2,i) = t_2]$	$P[W_e(2,i) < t_2]$	RSP$_e(t_2)^b$
1	1953.III	0.46	0.46	0.47	1954.II	0.72	0.18	0.73
2	1957.III	0.0036	0.99	0.0092	1958.II	0.13	0.87	0.18
3	1960.II	0.66	0.23	0.83	1961.I	0.034	0.92	0.042
4	1969.IV	0.20	0.73	0.33	1970.IV	0.44	0.52	0.71
5	1973.IV	0.065	0.56	0.10	1975.I	0.48	0.36	0.80
6	1980.I	0.0088	0.92	0.019	1980.III	0.20	0.74	0.42
7	1981.III	0.012	0.90	0.023	1982.IV	0.36	0.57	0.73

^{a}Probability that a change point (CP) into a recessionary state has occurred at that time $= \sum_i P[W(2,i) = t_1+1]$.

^{b}Probability that a CP out of a recessionary state has occurred at that time $= \sum_i P[W_e(2,i) = t_2]$.

Table 1: Dating of the US business cycle peaks and troughs as determined by the NBER, along with their associated probabilities of occurring at or before each time according to Hamilton's AR(4) mean switching model.

5.1 Regime Variability Plot

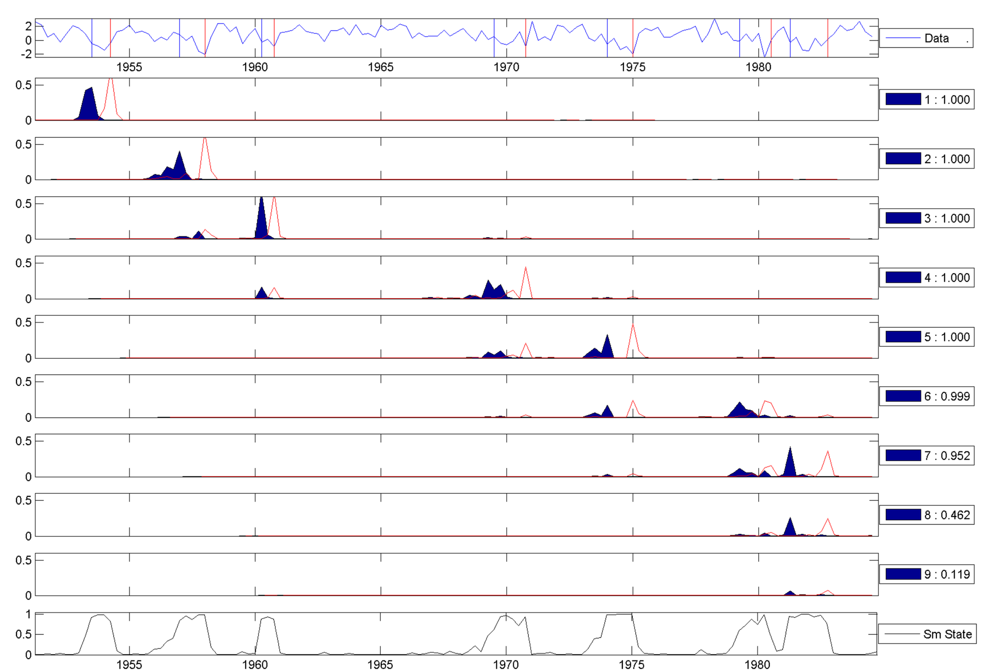

Figure 1: Regime Variability Plot for the Hamilton (1989) data.

5.2 Change Point Probabilities

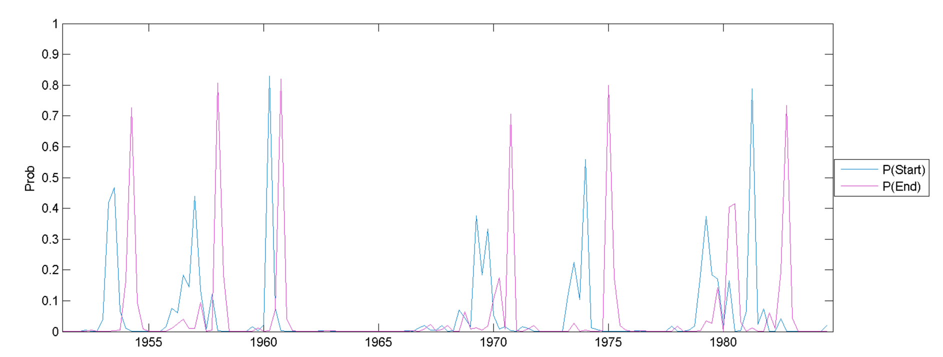

Figure 2: Change Point Probabilities for the Hamilton (1989) data.

References

[1] S. Chib. Estimation and comparison of multiple change-point models. *Journal of Econometrics*, 86:221–241, 1998.

[2] J. C. Fu and M. V. Koutras. Distribution theory of runs: A Markov chain approach. *Journal of the American Statistical Association*, 89:1050–1058, 1994.

[3] James D Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57:357–384, 1989.

†John Aston, Institute of Statistical Science, Academia Sinica,
128 Academia Road, Sec 2, Taipei 115, Taiwan
email : jaston@stat.sinica.edu.tw