Economics Letters 133 (2015) 14–18

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

A martingale decomposition of discrete Markov chains

Peter Reinhard Hansen *

European University Institute, Italy
CREATES, Denmark

HIGHLIGHTS

- We consider a multivariate time series given from a discrete Markov chain.
- Its martingale decomposition is derived, with all terms given in closed form.
- The decomposition is analogous to the Beveridge–Nelson decomposition.
- Decomposition has three terms: a persistent, a transitory, and a deterministic trend.
- The autocovariance structure across all terms is fully characterized.

ARTICLE INFO

Article history:
Received 24 October 2014
Received in revised form
7 April 2015
Accepted 26 April 2015
Available online 30 April 2015

JEL classification:
C10
C22
C58

Keywords:
Markov chain
Martingale
Beveridge–Nelson decomposition

ABSTRACT

We consider a multivariate time series whose increments are given from a homogeneous Markov chain. We show that the martingale component of this process can be extracted by a filtering method and establish the corresponding martingale decomposition in closed-form. This representation is useful for the analysis of time series that are confined to a grid, such as financial high frequency data.

1. Introduction

We consider a d-dimensional time series, $\{X_t\}$, whose increments, $\Delta X_t = X_t - X_{t-1}$, follow a homogeneous ergodic Markov chain with a countable state space. Thus, $X_t = X_0 + \sum_{j=1}^{t} \Delta X_j$, which makes X_t a (possibly non-stationary) Markov chain on a countable state space. We consider, $\mathrm{E}(X_{t+h}|\mathcal{F}_t)$, where $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots)$, is the natural filtration. The limit, as $h \to \infty$, is particularly interesting, because it leads to a martingale decomposition,

$$X_t = Y_t + \mu_t + U_t,$$

where μ_t is a linear deterministic trend, $\{Y_t, \mathcal{F}_t\}$ is a martingale with $Y_t = \lim_{h\to\infty} \mathrm{E}(X_{t+h} - \mu_{t+h}|\mathcal{F}_t)$, and U_t is a bounded stationary process. We derive closed-form expressions for all terms in the representation of X_t.

The martingale decomposition of finite Markov chains is akin to the Beveridge–Nelson decomposition for ARIMA processes, see Beveridge and Nelson (1981),[1] and the Granger representation for vector autoregressive processes, see Johansen (1991). The decomposition has many applications, as the long-run properties of X_t are governed by the persistent component, Y_t, while U_t characterizes the transitory component of X_t. In macro-econometrics Y_t and U_t are often called "trend" and "cycle", respectively, with Y_t being interpreted as the long run growth while U_t defines the fluctuations around the growth path, see, e.g. Low and Anderson (2008). A martingale decomposition of a stochastic discount process can be used to disentangle economic components with long term and short run impact on asset valuation, see Hansen (2012). For the broader

* Correspondence to: Department of Economics, European University Institute, Villa San Paolo, Via Della Piazzuola 43, 50133 FI Fiesole, Italy.
E-mail address: peter.hansen@eui.eu.

[1] The result, known as Beveridge–Nelson decomposition, appeared earlier in the statistics literature, e.g. Fuller (1976, Theorem 8.5.1). See Phillips and Solo (1992) for further discussion. The martingale decomposition is also key for the central limit theorem for stationary processes by Gordin (1969).

http://dx.doi.org/10.1016/j.econlet.2015.04.028
0165-1765/

concept of signal extraction of the "trend", see Harvey and Koopman (2002).

In the context with high-frequency financial data (which often are confined to a grid), Y_t and U_t may be labelled the efficient price and market microstructure noise, respectively. One could use the decomposition to estimate the quadratic variation of the latent efficient price Y_t, as in Large (2011) and Hansen and Horel (2009), and the framework could be adapted to study market information share, see e.g. Hasbrouck (1995). Markov processes are often used to approximate autoregressive processes in dynamic optimization problems, see Tauchen (1986) and Adda and Cooper (2000), and the decomposition could be used to compare the long-run properties of the approximating Markov process with those of the autoregressive process.

The paper is organized as follows: We establish an expression for the filtered process within the Markov chain framework, in Section 2, which leads to the martingale decomposition. Concluding remarks with discussion of various extensions are given in Section 3, and all proofs are given in the Appendix.

2. Theoretical framework

In this section we show how the observed process, $X_0, X_1, \ldots, X_n$, can be filtered in a Markov chain framework, using the natural filtration $\mathcal{F}_t = \sigma(X_t, X_{t-1}, \ldots)$. This leads to a martingale decomposition for X_t that is useful for a number of things.

Initially we seek the filtered price, $\mathrm{E}(X_{t+h}|\mathcal{F}_t)$, and we use the limit, as $h \to \infty$, to define the process,

$$Y_t = \lim_{h\to\infty} \mathrm{E}(X_{t+h} - \mu_{t+h}|\mathcal{F}_t),$$

where $\mu_t = t\mu$ with $\mu = \mathrm{E}(\Delta X_t)$. We will show that $\{Y_t, \mathcal{F}_t\}$ is a martingale, in fact, Y_t is the martingale component of X_t that, in turn, reveals a martingale representation theorem for finite Markov processes.

Note that the one step increments of $\mathrm{E}(X_{t+h} - \mu_{t+h}|\mathcal{F}_t)$ are, in general, autocorrelated at all order (including those lower than h), however all autocorrelations vanish as $h \to \infty$ and the martingale property of Y emerges. This filtering argument can be applied to any I(1) process for which $\mathrm{E}(\Delta X_{t+h}|\mathcal{F}_t) \overset{\text{a.s.}}{\to} \mathrm{E}(\Delta X_t)$ as $h \to \infty$, and this is the basic principle that Beveridge and Nelson (1981) used to extract the (stochastic) trend component of ARIMA processes.

2.1. Notation and assumptions

In this section we review the Markov terminology and present our notation that largely follows that in Brémaud (1999, Chapter 6). The following assumption is the only assumption we need to make.

Assumption 1. The increments $\{\Delta X_t\}_{t=1}^n$ are ergodic and distributed as a homogeneous Markov chain of order $k < \infty$, with $S < \infty$ states.

The assumption that S is finite can be dispensed with, which we detail in Section 3. For now we will assume S to be finite because it greatly simplifies the exposition. The transition matrix for price increments is denoted by P. For a Markov chain of order k with S basic states, P will be an $S^k \times S^k$ matrix. We use $\pi \in \mathbb{R}^{S^k}$ to denote the stationary distribution associated with P, which is uniquely defined by $\pi' P = \pi'$. The *fundamental matrix* is defined by[2]

$$Z = (I - P + \Pi)^{-1},$$

where $\Pi = \iota\pi'$ is a square matrix and $\iota = (1, \ldots, 1)'$, (so all rows of Π are simply π'). We use e_r to denote the r-th unit vector, so that $e_r' A$ is the r-th row of a matrix A of proper dimensions.

Let $\{x_1, \ldots, x_S\}$ be the support for ΔX_t, with $x_s \in \mathbb{R}^d$. We will index the possible realizations for the k-tuple, $\Delta\mathcal{X}_t = (\Delta X_{t-k+1}, \ldots, \Delta X_t)$, by $\mathbf{x}_s$, $s = 1, \ldots, S^k$, which includes all the perturbations, $(x_{i_1}, \ldots, x_{i_k})$, $i_1, \ldots, i_k = 1, \ldots, S$. The transition matrix, P, is given by

$$P_{r,s} = \Pr(\Delta\mathcal{X}_{t+1} = \mathbf{x}_s|\Delta\mathcal{X}_t = \mathbf{x}_r).$$

This matrix will be sparse when $k > 1$, because at most S transitions from any state have non-zero probability, regardless of the order of the Markov chain.

For notational reasons it is convenient to introduce the sequence $\{s_t\}$ that is defined by $\Delta\mathcal{X}_t = \mathbf{x}_{s_t}$, so that s_t denotes the observed state at time t. We also define the matrix $f \in \mathbb{R}^{S^k \times d}$ whose s-th row, denoted $f_s = e_s' f$, is the realization of $\Delta X'$ in state s. It follows that $\Delta X_t = f' e_{s_t}$ and that the expected value of the increments is given by $\mu = \mathrm{E}(\Delta X_t) = f'\pi \in \mathbb{R}^d$.

The auxiliary vector process, e_{s_t}, is such that $\mathrm{E}(e_{s_{t+1}}|\mathcal{F}_t) = P' e_{s_t}$, so that e_{s_t} can be expressed as a vector autoregressive process of order one with martingale difference innovations, see e.g. Hamilton (1994, p. 679).

2.2. Markov chain Filtering

The filtered process $\mathrm{E}(X_{t+h}|\mathcal{F}_t)$, is simple to compute in the Markov setting, because $\mathrm{E}(X_{t+h}|\mathcal{F}_t) = \mathrm{E}(X_{t+h}|\Delta\mathcal{X}_t)$ and $X_{t+h} = X_t + \sum_{j=1}^h \Delta X_{t+j}$ with $\mathrm{E}(\Delta X_{t+1}'|\Delta\mathcal{X}_t = \mathbf{x}_r) = \sum_{s=1}^{S^k} P_{r,s} f_s = e_r' Pf$. More generally we have $\mathrm{E}(\Delta X_{t+h}'|\Delta\mathcal{X}_t) = e_{s_t}' P^h f$, which shows that

$$\mathrm{E}(X_{t+h}'|\Delta\mathcal{X}_t) = X_t' + e_{s_t}' \sum_{j=1}^h P^j f.$$

After subtracting the deterministic trend, μ_{t+h}, we let $h \to \infty$ and define

$$Y_t = \lim_{h\to\infty} \mathrm{E}(X_{t+h} - \mu_{t+h}|\mathcal{F}_t),$$

which we label the filtered process of X_t. The process, Y_t is well defined and adapted to the filtration $\mathcal{F}_t$. We are now ready to formulate our main result.

Theorem 1. *The process and $\{Y_t, \mathcal{F}_t\}$ is a martingale with initial value, $Y_0 = X_0 + f'(Z' - I)e_{s_0}$ and its increments are given by $\Delta Y_t' = e_{s_t}' Zf - e_{s_{t-1}}' PZf$. Moreover, we have*

$$X_t = Y_t + \mu_t + U_t, \tag{1}$$

where $U_t' = e_{s_t}'(I - Z)f$ is a bounded, stationary, and ergodic process with mean zero.

All terms of the expression are given in closed-form, analogous to the Granger representation theorem by Hansen (2005).

It can be shown that ΔY_t is a Markov process with S^{k+1} possible states values. Analogous to P and f, let Q and g denote the transition matrix for ΔY_t and its matrix of state values, respectively. The martingale property dictates that $Qg = 0 \in \mathbb{R}^{S^{k+1}\times d}$. Note that ΔY_t is typically conditionally heterogeneous, as Q is not a matrix of rank one, which would be the structure corresponding to the case where ΔY_t is independent and identically distributed.

The autocovariance structure of the terms in the martingale decomposition is stated next.

Theorem 2. *We have* $\mathrm{var}(\Delta Y_t) = f'Z'(\Lambda_\pi - P'\Lambda_\pi P)Zf$ *where* $\Lambda_\pi = \mathrm{diag}(\pi_1, \ldots, \pi_{S^k})$ *and*

$$\begin{aligned}\mathrm{cov}(U_t, U_{t+j}) &= f'(I - Z)'\Lambda_\pi P^{|j|}(I - Z)f \\ &= f'Z'P'\Lambda_\pi P(P^{|j|} - \Pi)Zf,\end{aligned}$$

[2] The matrix, $I - P + \Pi$, is invertible since the largest eigenvalue of $P - \Pi$ is less than one under Assumption 1.