Journal of Multivariate Analysis 101 (2010) 1263–1273

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

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

Generalized Levinson–Durbin sequences, binomial coefficients and autoregressive estimation

Paul Shaman

Department of Statistics, University of Pennsylvania, Philadelphia, PA 19104-6340, United States

ARTICLE INFO

Article history:
Received 9 March 2009
Available online 18 January 2010

AMS subject classifications:
62M10
05A10
11B65

Keywords:
Binomial coefficients
Generalized Levinson–Durbin sequence
Least squares estimator
Levinson–Durbin sequence
Partial correlations
Yule–Walker estimator

ABSTRACT

For a discrete time second-order stationary process, the Levinson–Durbin recursion is used to determine the coefficients of the best linear predictor of the observation at time $k+1$, given k previous observations, best in the sense of minimizing the mean square error. The coefficients determined by the recursion define a Levinson–Durbin sequence. We also define a generalized Levinson–Durbin sequence and note that binomial coefficients form a special case of a generalized Levinson–Durbin sequence. All generalized Levinson–Durbin sequences are shown to obey summation formulas which generalize formulas satisfied by binomial coefficients. Levinson–Durbin sequences arise in the construction of several autoregressive model coefficient estimators. The least squares autoregressive estimator does not give rise to a Levinson–Durbin sequence, but least squares fixed point processes, which yield least squares estimates of the coefficients unbiased to order $1/T$, where T is the sample length, can be combined to construct a Levinson–Durbin sequence. By contrast, analogous fixed point processes arising from the Yule–Walker estimator do not combine to construct a Levinson–Durbin sequence, although the Yule–Walker estimator itself does determine a Levinson–Durbin sequence. The least squares and Yule–Walker fixed point processes are further studied when the mean of the process is a polynomial time trend that is estimated by least squares.

1. Introduction

The Levinson–Durbin recursion has long been a fixture in time series analysis. It is commonly viewed in two contexts. One is that of prediction for a discrete time, second-order stationary process $\{y_t\}$ with known structure. Given $y_1, \ldots, y_k$, for any $k \geq 1$ the recursion determines the coefficients $\alpha_{j,k}$, $j = 1, \ldots, k$, of the best linear predictor of y_{k+1},

$$\hat{y}_{k+1} = -\alpha_{1,k} y_k - \cdots - \alpha_{k,k} y_1, \tag{1.1}$$

best in the sense of minimizing the mean square error. The recursion begins with specification of $\alpha_{1,1}$, and at the nth stage one obtains $\alpha_{1,n}, \ldots, \alpha_{n,n}$. The mean square error of the predictor is also specified at each step. Levinson [1] devised the recursion to give a simple procedure for construction of the best linear predictor when the structure of the process is known. His paper was reprinted as Appendix B to Wiener's monograph on time series [2]. Wiener's work had originally been issued in February 1942 as a classified government report. For some details of this history see [3]. The second context for the recursion is that of estimation of the coefficients of an autoregressive model of finite order, given data $y_1, \ldots, y_T$. The sample Yule–Walker equations are commonly used to construct an estimator of the coefficients. Bartlett [4, pp. 264–265], Daniels [5, p. 183] and

E-mail address: shaman@wharton.upenn.edu.

0047-259X/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2010.01.004

Durbin [6] all derived the recursion as a simple method of solving the sample Yule–Walker equations, which are linear in the coefficients.

The recursion determines a double sequence $\alpha_{j,n}$, $j = 1, \ldots, n$, $n = 1, 2, \ldots$. The sequence $\alpha_{n,n}$, $n = 1, 2, \ldots$, determines all of the values $\alpha_{j,n}$.

In this paper we study properties of sequences produced by the Levinson–Durbin recursion, and we further study generalization of such sequences.

Definition 1. $\alpha_{j,n}, j = 1, \ldots, n$, $n = 1, 2, \ldots$, is a *Levinson–Durbin sequence* if the coefficients, all real-valued, satisfy

$$\alpha_{j,n} = \alpha_{j,n-1} + \alpha_{n,n}\alpha_{n-j,n-1}, \quad j = 1, \ldots, n-1, \ n = 2, 3, \ldots, \tag{1.2}$$

and

$$\left|\alpha_{n,n}\right| < 1, \quad n = 1, 2, \ldots. \tag{1.3}$$

If (1.2) holds and the $\alpha_{n,n}$'s are not subject to (1.3), we say that the $\alpha_{j,n}$'s form a *generalized Levinson–Durbin sequence*.

For the prediction problem the recursion (1.2) is used together with (1.3) and the $-\alpha_{n,n}$'s defined to be the partial correlations of the process being predicted. In the context of autoregressive estimation, the Yule–Walker estimator uses (1.2) and (1.3) and defines the $-\alpha_{n,n}$'s to be the sample partial correlations. Other estimators (the Burg and Kay procedures, mentioned below) employ (1.2) and (1.3) and define the $\alpha_{n,n}$'s differently. A generalized Levinson–Durbin sequence allows arbitrary specification of the $\alpha_{n,n}$'s. If (1.3) does hold, the sequence of $-\alpha_{n,n}$'s forms the partial correlation function for some second-order stationary process.

Levinson–Durbin sequences arise, e.g., from (i) Yule–Walker and tapered Yule–Walker estimation of the coefficients of an autoregressive process, (ii) fixed point models arising in least squares estimation of the autoregressive process coefficients, (iii) estimation of the autoregressive process coefficients by Burg's method and (iv) estimation of the autoregressive process coefficients by Kay's method [7].

If $\alpha_{n,n} = 1$ for each n, then (1.2) generates the binomial coefficients. Taking into account the symmetric structure of binomial coefficients, we see that (1.2) is simply an expression of Pascal's triangle if $\alpha_{n,n} = 1$ for each n. The binomial coefficients also arise as the limit of a sequence of fixed point models determined by least squares estimation of autoregressive process coefficients, as noted in [8]. This will be discussed in Section 3.

The Yule–Walker estimator of the coefficients of an autoregressive process of known finite order p is determined by a Levinson–Durbin recursion which defines a sequence for which $\alpha_{j,n} = \alpha_{j,p}$, $j = 1, \ldots, p$, and $\alpha_{j,n} = 0, j = p+1, \ldots, n$, for all $n > p$. The values $-\alpha_{n,n}$, $n = 1, \ldots, p$, are the Yule–Walker sample partial correlations. For sample length T the order $1/T$ bias of the Yule–Walker estimator has been discussed in [9,10]. For each value of p numerical calculations show that there is a unique autoregressive process of order p (unique up to scale) for which the order $1/T$ bias of the Yule–Walker estimator is 0. This process is called a fixed point process because it is given by the fixed point of a contraction mapping. This result may be extended to the case where a polynomial trend in time is estimated by least squares and the Yule–Walker estimator is subsequently calculated from the trend residuals. The Yule–Walker fixed point processes differ according to the autoregressive order p and the degree of the estimated polynomial trend, and they can be determined numerically by iterating the contraction mappings. Although the Yule–Walker estimator itself yields a Levinson–Durbin sequence, it is interesting that the Yule–Walker fixed point processes for a given degree of estimated polynomial trend do not combine to form a Levinson–Durbin sequence. These comments also hold for the tapered Yule–Walker estimator, with the proviso that the fixed point processes depend upon the specific data taper chosen. The tapered Yule–Walker estimator is considered in [11,12].

The order $1/T$ bias of the least squares estimator of the coefficients of an autoregressive process of known finite order p has been derived in [9,10,13]. The bias expression is linear in the autoregressive parameters and defines a contraction mapping. A fixed point process which is unique up to scale and for which the least squares estimator is unbiased to order $1/T$ can be derived analytically for each autoregressive order p and degree of estimated polynomial trend in time. Moreover, for each degree of estimated polynomial trend, the fixed point processes form a sequence of projections from an infinite order fixed point process. In contrast to the Yule–Walker situation, the least squares estimator does not yield a Levinson–Durbin sequence, but the least squares fixed point processes for a given degree of estimated polynomial trend do combine to form a Levinson–Durbin sequence.

The Burg and Kay estimators both generate Levinson–Durbin sequences. Burg's algorithm determines the $\alpha_{n,n}$'s by minimizing a sequence of sums of squares of forward and backward one-step prediction errors. The remaining $\alpha_{j,n}$ values are then determined from (1.2). For a description of the Burg estimator see, for example, [14, pp. 147–8]. Kay's estimator [7] of the autoregressive coefficients is a recursive maximum likelihood procedure. The parameter $\alpha_{n,n}$ is estimated at the nth stage by maximizing a partial Gaussian likelihood and then (1.2) is applied to determine $\alpha_{1,n}, \ldots, \alpha_{n-1,n}$.

This paper is organized as follows. In Section 2 some properties of generalized Levinson–Durbin sequences are presented. These results generalize relations satisfied by binomial coefficients. It is also noted that the Levinson–Durbin sequences define minimum phase filters. Least squares estimation bias and least squares fixed point processes are described in Section 3. Section 4 is devoted to Yule–Walker estimation bias and fixed point processes. Concluding discussion appears in Section 5, and proofs are in Section 6.