연세경제연구 제Ⅳ권, 제2호
Yonsei Economic Studies, Vol. Ⅳ, No. 2 (September 1997)

The Cholesky Decomposition and the Maximum Entropy Probability Density of a Stationary Process

Seongbaek Yi

Burg's maximum entropy spectrum [1] equals the spectrum of an autoregressive process. The purpose of this correspondence is to present a short and elementary proof of Burg's theorem and its generalization to a probability density function using the Cholesky decomposition.

I. Introduction

Burg's maximum entropy spectrum is based on the following theorem. If the first $p+1$ terms of the autocovariance function (ACVF) of a stationary process are given, the spectrum maximizing entropy rate is equal to that of the Gaussian autoregressive process of order p, AR(p), [1]~[3], [5]~[8]. Burg's theorem was generalized to the problem of selecting the maximum entropy probability density subject to the autocovariance constraints [2], [3]. The purpose of this correspondence is to present an elementary proof of Burg's theorem and its generalization using the Cholesky decomposition. The proof is simpler and shorter than any other existing one.

Lecturer, Department of Applied Statistics, Yonsei University, Seoul, 120-749, Korea.

Ⅱ. The Cholesky Decomposition

Let $\{\sigma(j)\}$ be the ACVF of a stationary process $\{X_t\}$. Let Σ_k be a $k \times k$ symmetric Toeplitz matrix, whose (i, j) element is $\sigma(i-j)$. Since $\{\sigma(j)\}$ is a positive-definite sequence, the system of the Yule-Walker equations

$$\sigma(j) = \varnothing_{k,1}\,\sigma(j-1) + \cdots + \varnothing_{k,k}\,\sigma(j-k), \qquad j = 1, \cdots, k$$

has the unique solution. Define the kth prediction error by

$$v_k = -\sum_{j=0}^{k} \varnothing_{k,j}\,\sigma(-j),\ \varnothing_{k,0} = -1.$$

Let Φ_k be a $k \times k$ lower triangular matrix whose (i, j) element is $\varnothing_{i-1,j-1}$ for $i \geq j$, and let V_k be a $k \times k$ diagonal matrix whose ith diagonal element is v_{i-1}. Then, it is known [4], [7] that for $k = 1, 2, \cdots$, the Cholesky decomposition of the inverse matrix of Σ_k is

$$\Phi_k \Sigma_k \Phi_k' = V_k.$$

Ⅲ. The Maximum Entropy Probability Density

Burg's spectrum was generalized to the maximum entropy probability density as follows [3].

Theorem For any $n(>p)$, let $X_1, \cdots, X_n$ be a regular second-order stationary process with covariance matrix Σ_n. Assume that the first $p+1$ autocovariances are given. Then, the random vector $X_n = (X_1, \cdots, X_n)'$ has the maximum entropy if and only if the process is from a Gaussian AR(p) model.

Proof Using Jensen's inequality, one can show [2] that X_n has the maximum entropy if and only if it is normally distributed. In this case, the maximum entropy is $0.5\ln\{(2\pi e)^n |\Sigma_n|\}$. Thus, the problem becomes to maximize $|\Sigma_n|$ subject to the $p+1$ autocovariance constraints. The Cholesky decomposition in Section Ⅱ implies that

$$|\Sigma_n| = |V_n| = \prod_{j=0}^{n-1} v_j.$$

Since the first $p+1$ autocovariances are given, $v_0, \cdots, v_p$, are fixed. Thus, it is left to maximize $\prod_{j=p+1}^{n-1} v_j$. Because the process is regular, v_j is positive for each j. The Yule-Walker equations imply that for $j = 0, 1, \cdots,$

$$v_{j+1} = (1-\phi_{j+1,j+1}^2)\, v_j$$

where $v_0 = \sigma(0)$. Thus, the maximum of $|\Sigma_n|$ is attained if and only if $\phi_{j,j} = 0$ for $j > p$. Then, the Levinson-Durbin algorithm implies that the process $\{X_1, \cdots, X_n\}$ is from the Gaussian AR(p) model with AR coefficients $\phi_{p,1}, \cdots, \phi_{p,p}$ and white-noise variance v_p.

Q. E. D.

❖ *REFERENCES* ❖

1. Burg, J. P., "Maximum Entropy Spectral Analysis," in *The Proceedings of the 37th Meet. Soc. Exploration Geophy.*, 1967.
2. Choi, B. S. and T. M. Cover, "An Information-Theoretic Proof of Burg's Maximum Entropy Spectrum," *Proceedings of IEEE*, 72, 1984, pp. 1094 ~ 1095.
3. Choi, B. S., "On the Relation between the Maximum Entropy Probability Density Function and the Autoregressive Model," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-34, 1986, pp. 1659 ~ 1661.
4. ______, "On the Block LU Decomposition of Block Toeplitz Matrices Using the Vector Yule-Walker Equations," *Communications in Statistics-Theory and Method*, 19, pp. 2815 ~ 2827.
5. Edward, J. A. and M. M. Fitelson, "Notes on Maximum-Entropy Processing," *IEEE Transactions on Information Theory*, IT-19, 1973, pp. 232 ~ 234.
6. Grandell, J., Hamrud, M. and P. Toll, "A Remark on the Correspondence between the Maximum Entropy Method and the Autoregressive Model," *IEEE Transactions on Information Theory*, IT-26, 1980, pp. 750 ~ 751.
7. Smylie, D. G., Clarke, G. K. C. and T. J. Ulrych, "Analysis of Irregularities in the Earth's Rotation," *Meth. Comp. Phys.*, 13, 1973, pp. 391 ~ 430.
8. Van den Bos, A., "Alternative Interpretation of Maximum Entropy Spectral Analysis," *IEEE Transactions on Information Theory*, IT-17, 1971, pp. 493 ~ 494.