A New Spectral Analysis in Time Series Data

Hoonja Lee
Pyongtaek University, Department of Information Science
111 Yongi-Dong
Pyongtaek City, 450-701, S. Korea
esther@ptuniv.ac.kr

Robert V. Foutz
Virginia Polytechnic Institute and State University, Department of Statistics
Blacksburg, Virginia, 24060, USA
B2948@vt.edu

1. Introduction

The classical Fourier analysis are used to detect periodic trends that are of the sinusoidal shape in time series data. The basic idea of the Fourier series is that a time series data $x(t)$ can be decomposed into a sum of sines and cosines:

$$x(t) = \sum_{k=0}^{\infty} a_k \cos\frac{2\pi kt}{T} + b_k \sin\frac{2\pi kt}{T}. \tag{1}$$

This composition shows that $x(t)$ is a sum of sinusoidal shapes at frequencies $\lambda_k = 2\pi k/T$ for k =0, 1, … In addition, the variability in $x(t)$ as measured by $\int_0^T |x(t)|^2 dt$ partitions into the sum of the variabilities of the sinusoidal shapes. The Fourier analysis is to treat the partition as an analysis of variance (ANOVA) for identifying sinusoidal periodicities in a time series data $x(t)$.

This article describes an ANOVA technique that is better suited for identifying non-sinusoidal periodic components. In addition, the new technique can be used to investigate the shapes of these periodic components.

2. A New Spectral Analysis

Let L^2 be the Hilbert space with the inner product $< x, y >_L = \frac{1}{2\pi}\int_0^{2\pi} x(t)\overline{y(t)}\,dt$ and let $\lfloor b \rfloor$ be the largest integer no larger than b, for example, $\lfloor 3.5 \rfloor$ =3 and $\lfloor -3.5 \rfloor$ =-4.

The complex-valued functions

$$f_{k,n,j}(t) = \exp\left(i\frac{\pi}{2}\left\lfloor \frac{2kt}{\pi} + \frac{2}{\pi} - \frac{j-1}{n} + 1 \right\rfloor\right) \quad \text{and} \quad f_{-k,n,j}(t) = \exp\left(i\frac{\pi}{2}\left\lfloor -\frac{2kt}{\pi} - \frac{2}{\pi} + \frac{j-1}{n} \right\rfloor\right)$$

are step functions that takes the values –1, 0, and 1. The functions $f_{k,n,j}$ and $f_{-k,n,j}$ are conjugates of each other. Construct $B_{k,n}$ to be the subspace of L^2 spanned by $f_{k,n,1}, \ldots, f_{k,n,n}$. The direct sum $L_{m,n} = B_{-m,n} + \ldots + B_{m,n}$ is the subspace of the of L^2 spanned by the functions $\{f_{k,n,j}(t)\}$. For $x(t)$, $y(t)$ in $L_{m,n}$, define the inner product $< x, y >_G = \int_0^1 < x_u, y_u >_L du,$ where $x_u(t) = \sum_{k=-m}^{m} x_k(t + \frac{4u\boldsymbol{p}}{2k})$ for the time shift parameter $4u\boldsymbol{p} / 2k$.

The central result follows.

THEOREM1. The subspaces $B_{-m,n}, \ldots, B_{m,n}$ of $L_{m,n}$ are mutually orthogonal with respect to the inner product $< x, y >_G$. Furthermore, each function $x(t)$ in L^2 has the unique new spectral representation

$$x(t) = \lim_{m \to \infty} \lim_{n \to \infty} \sum_{k=-m}^{m} P_G(x_{m,n} | B_{k,n})(t) \tag{2}$$

where $x_{m,n}(t)$ is the projection $P_L(x | L_{m,n})(t)$ of $x(t)$ onto $L_{m,n}$ with respect to the inner product $< x, y >_L$ and where $P_G(x_{m,n} | B_{k,n})(t)$ is the projection of $x_{m,n}(t)$ onto $B_{k,n}$ with respect to the inner product $< x, y >_G$. The component $P_G(x_{m,n} | B_{k,n})(t)$ in (2) is periodic with period $2\boldsymbol{p} / k$ and it may take non-sinusoidal shapes.

The squared norm of this series approximation is

$$\left| \sum_{k=-m}^{m} P_G(x_{m,n} | B_{k,n})(t) \right|_G^2 = \sum_{k=-m}^{m} \left| P_G(x_{m,n} | B_{k,n})(t) \right|_L^2 . \tag{3}$$

The ANOVA partition in (3) is the basis for the new spectral analysis.

REFERENCE

Lee, H and Foutz, R. V. (1997) A New Spectral Model for Binary or Categorical-valued Time Series Data, *Journal of Statistical Computation and Simulation*, 58, 217-235.

Foutz, R. V. and Lee, H. (2000) Adaptive Fourier Series and the Analysis of Periodicities in Time Series Data, *J. Time Ser Anal*, 21, 649-662.

Stoffer, D. S. (1991) Walsh-Fourier Analysis and its Statistical Application, *J. Am., Stat. Assoc*, 86, 461-485.

Koopmans, L.H. (1974) *The Spectral Analysis of Time Series*, New York: Academic Press.