第5卷　第6期　　光学　精密工程　　Vol. 5, No. 6
1997年12月　　OPTICS AND PRECISION ENGINEERING　　December, 1997

平稳时间序列高斯性检验的积分多谱方法

丛玉良　　王树勋　　王宏志　　戴逸松
(吉林工业大学电子工程系　长春 130025)

摘要　利用积分多谱给出了一种平稳时间序列的高斯性检验方法, 并通过仿真实验进行了验证。

关键词　积分多谱　平稳时间序列　高斯性检验

1　引　言

在时间序列分析中, 必须首先确定一个平稳时间序列是否服从高斯分布。如果一个平稳时间序列是服从高斯分布的, 则其均值和二阶矩就包含了序列全部信息, 如果一个平稳时间序列是服从非高斯分布的, 则建模时必须用到三阶以上统计量。已有的利用双谱来进行检验的方法对服从非高斯对称分布的时间序列是无法识别的[1]。而利用三谱来进行检验的方法对某些服从非高斯分布但三谱值较小的时间序列也是无法识别的, 并且其计算速度较慢。本文利用积分双谱和积分三谱给出了一种平稳时间序列的高斯性检验方法, 可以较好的解决这些问题。

2　积分双谱和积分三谱定义

文献[2]给出积分双谱和积分三谱定义:

2.1　积分双谱(**Integrated Bispectrum**)

设 $\{x(n)\}$ 表示一长为 N 的平稳离散时间序列, 其均值为零, 定义:

$$y(n) = x^2(n) - E\{x^2(n)\} \tag{1}$$

则 $\{y(n)\}$ 与 $\{x(n)\}$ 间的交叉谱为 $S_{yx}(\omega)$:

收稿日期: 1997-09-28
修稿日期: 1997-10-15

$$S_{yx}(\omega)=\sum_{k=-\infty}^{\infty}E\{y(n)x(n+k)\}\exp(-j\omega k) \tag{2}$$

文献[2]已推出

$$S_{yx}(\omega)=\frac{1}{2\pi}\int_{-\pi}^{\pi}B_x(\omega,\omega_2)\mathrm{d}\omega_2=\frac{1}{2\pi}\int_{-\pi}^{\pi}B_x(\omega_1,\omega)\mathrm{d}\omega_1 \tag{3}$$

其中$B_x(\omega_1,\omega_2)$为序列$\{x(n)\}$的双谱。

因此将$S_{yx}(\omega)$称为积分双谱。积分双谱$S_{yx}(\omega)$可用以下方法进行估计。

假设数据长度N被分成K个互不重合的小段，每段长N_B，则$N=KN_B$。用$X^{(i)}(\omega)$表示第i段$\{x(n+(i-1)N_B),1\leq(n\leq N_B\}(i=1,2,...,K)$的离散傅立叶变换,即

$$X^{(i)}(\omega_n)=\sum_{l=0}^{N_R-1}x(l+1+(i-1)N_B)\exp(-j\omega_n l) \tag{4}$$

其中

$$\omega_n=\frac{2\pi}{N_B}m,\ m=0,1,\cdots,N_B-1 \tag{5}$$

同样用$Y^{(i)}(\omega_n)$表示第i段$\{x^2(n+(i-1)N_B)-E\{x^2(n)\}\},(1\leq n\geq N_B,i=1,2,\cdots,K)$的离散傅立叶变换,即

$$Y^{(i)}(\omega_n)=\sum_{l=0}^{N_B-1}\{x^2(l+1+(i-1)N_B)-\hat{\mu}_{2x}\}\exp(-j\omega_n l) \tag{6}$$

其中

$$\hat{\mu}_{2x}=\frac{1}{N}\sum_{n=1}^{N}x^2(n) \tag{7}$$

则第i段$y(n)$与$\{x(n)\}$交叉周期图为

$$\hat{S}_{yx}^{(i)}(\omega_n)=\frac{1}{N_B}X^{(i)}(\omega_n)[Y^{(i)}(\omega_n)]^* \tag{8}$$

其中Y^*表示Y的复数共轭。因此可得积分双谱$S_{yx}(\omega)$的一致估计$\hat{S}_{yxN}(\omega_n)$:[2]

$$\hat{S}_{yxN}(\omega_n)=\frac{1}{K}\sum_{i=1}^{K}S_{yx}^{(i)}(\omega_n) \tag{9}$$

2.2 积分三谱(**Integrated Trispectrum**)

定义:

$$r(n)=x^3(n)-3x(n)E\{x^2(n)\}-E\{x^3(n)\} \tag{10}$$

则$r(n)$与$\{x(n)\}$间的交叉谱为$S_{rx}(\omega)$:

$$S_{rx}(\omega)=\sum_{k=-\infty}^{\infty}E\{r(n)x(n+k)\}\exp(-j\omega k) \tag{11}$$

文献[2]已推出

$$S_{rx}(\omega)=\frac{1}{(2\pi)^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}T_x(\omega,\omega_2,\omega_3)\mathrm{d}\omega_2\mathrm{d}\omega_3 \tag{12}$$

其中$T_x(\omega_1,\omega_2,\omega_3)$为序列$\{x(n)\}$的三谱。

因此将$S_{rx}(\omega)$称为积分三谱。积分三谱$S_{yx}(\omega)$可用以下方法进行估计。

用 $R^{(i)}(\omega_n)$ 表示第 i 段 $\{x^3(n+(i-1)N_B)-3x(n+(i-1)N_B)E\{x^2(n)\}-E\{x^3(n)\}\}$, $(1\leq n\leq N_B,\ i=1,2,\cdots,K)$ 的离散傅立叶变换, 即

$$R^{(i)}(\omega_n)=\sum_{l=0}^{N_D-1}[x^3(l+1+(i-1)N_B-3x(l+(i-1)N_B)\hat{\mu}_{2x}-\hat{\mu}_{3x}]\exp(-j\omega_n l) \tag{13}$$

其中

$$\hat{\mu}_{3x}=\frac{1}{N}\sum_{n=1}^{N}x^3(n) \tag{14}$$

则第 i 段 $\{r(n)\}$ 与 $\{x(n)\}$ 交叉周期图为

$$\hat{S}_{rx}^{(i)}(\omega_n)=\frac{1}{N_B}X^{(i)}(\omega_n)[R^{(i)}(\omega_n)]^* \tag{15}$$

其中 R^* 表示 R 的复数共轭。因此可得积分三谱 $S_{rx}(\omega)$ 的一致估计 $\hat{S}_{rxN}(\omega_n)$ [2]:

$$\hat{S}_{rxN}(\omega_n)=\frac{1}{K}\sum_{i=1}^{K}S_{rxN}^{(i)}(\omega_n) \tag{16}$$

3　平稳时间序列高斯性检验的基本原理

3.1　积分双谱检验统计量

文献[2]定义:

$$D_{3N}(\omega_n)=\frac{\hat{S}_{yxN}(\omega_n)}{[\frac{1}{K}S_{yy}(\omega_n)S_{xx}(\omega_n)]^{\frac{1}{2}}} \tag{17}$$

如果 $\{x(n)\}$ 是高斯序列, 则 $D_{3N}(\omega_n)$ 是具有独立实部和虚部的零均值、单位方差的渐近复高斯过程。当 $m\neq n(m,\ n=1,2,\cdots,\frac{N_B}{2}-1)$ 时 $D_{3N}(\omega_m)$ 与 $D_{3N}(\omega_n)$ 是渐近统计独立的[2]。因此当 N 很大时有

$$2|D_{3N}(\omega_n)|^2\sim\chi_2^2(0) \tag{18}$$

在式(17)中, $S_{yy}(\omega_n)$ 和 $S_{xx}(\omega_n)$ 可分别用其一致估计 $\hat{S}_{yyN}(\omega_n)$ 和 $\hat{S}_{xxN}(\omega_n)$ [2] 来表示, 由此可定义:

$$\hat{D}_{3N}(\omega_n)=\frac{\hat{D}_{yxN}(\omega_n)}{[\frac{1}{k}\hat{S}_{yyN}(\omega_n)\hat{S}_{xxN}(\omega_n)]^{\frac{1}{2}}} \tag{19}$$

所以有

$$\hat{S}_{3N}=2\sum_{m=1}^{\frac{N_B}{2}-1}|\hat{D}_{3N}(\omega_n)|^2\sim\chi_{2P}^2(0) \tag{20}$$

其中 $P=N_B/2-1$, P 为总的检验点数。当 $\hat{S}_{3N}$ 大于某一常数 λ 时, (λ 由显著水平 α 决定), 则判决平稳时间序列 $\{x(n)\}$ 服从非高斯分布, 否则判决平稳时间序列 $\{x(n)\}$ 服从高斯分布。

3.2　积分三谱检验统计量

同3.1，文献[2]定义：

$$\hat{S}_{4N} = 2\sum_{m=1}^{\frac{N_B}{2}-1} \left| \hat{D}_{4N}(\omega_n) \right|^2 \tag{21}$$

其中

$$\hat{D}_{4N}(\omega_n) = \frac{\hat{S}_{rxN}(\omega_n)}{\left[\frac{1}{K}\hat{S}_{rrN}(\omega_n)\hat{S}_{xxN}(\omega_n)\right]^{\frac{1}{2}}} \tag{22}$$

如果$\{x(n)\}$是高斯序列，当N很大时有

$$2\left|\hat{D}_{4N}(\omega_n)\right|^2 \sim \chi_2^2(0) \tag{23}$$

$$\hat{S}_{4N} \sim \chi_{2P}^2(0) \tag{24}$$

因此当$\hat{S}_{4N}$大于某一常数λ时，(λ由显著水平α决定)，则判决平稳时间序列$\{x(n)\}$服从非高斯分布，否则判决平稳时间序列$\{x(n)\}$服从高斯分布。

3.3 积分双谱、三谱混合型检验统计量

由文献[3]知，当N很大时(即$K\to\infty$和$N_B\to\infty$)有

$$\mathrm{cov}\{\mathrm{Re}\{\hat{S}_{yx}(\omega_n)\},\ \mathrm{Im}\{\hat{S}_{yx}(\omega_n)\}\} = \frac{1}{2K}\mathrm{lm}\{S_{yx}^2(\omega_n)\} + O(N^{-1}) \tag{25}$$

$$\mathrm{cov}\{\mathrm{Re}\{\hat{S}_{rx}(\omega_n)\},\ \mathrm{Im}\{\hat{S}_{rx}(\omega_n)\}\} + \frac{1}{2K}\mathrm{lm}\{S_{rx}^2(\omega_n)\} + O(N^{-1}) \tag{26}$$

$$\mathrm{cov}\{\hat{S}_{yx}(\omega_n), \hat{S}_{rx}(\omega_n)\} = \frac{1}{K}S_{yr}(\omega_n)S_{xx}(\omega_n) + O(N^{-1}) \tag{27}$$

$$\mathrm{cov}\{\hat{S}_{yx}(\omega_n), \hat{S}_{rx}^*(\omega_n)\} = -\frac{1}{K}S_{yx}(\omega_n)S_{rx}(\omega_n) + O(N^{-1}) \tag{28}$$

如果$\{x(n)\}$是零均值高斯序列，则有$S_{yx}(\omega)=0$，$S_{rx}(\omega)=0$，$S_{yr}(\omega)=0$，因此上述所有协方差函数均趋于零，所以可定义新的检验统计量

$$\hat{T}_{34}(\omega_n) = 2\left|\hat{D}_{3N}(\omega_n)\right|^2 + 2\left|\hat{D}_{4N}(\omega_n)\right|^2 \tag{29}$$

$$\hat{T}_{34}(\omega_n) \sim \chi_4^2(0) \tag{30}$$

$$T = 2\sum_{m=1}^{\frac{N_B}{2}-1}\hat{T}_{34N}(\omega_n) \sim \chi_{4P}^2(0) \tag{31}$$

其中$P = N_B/2-1$，P为总的检验点数. 当T大于某一常数T_0时，(T_0由显著水平α决定)，则判决平稳时间序列$\{x(n)\}$服从非高斯分布，否则判决平稳时间序列$\{x(n)\}$服从高斯分布。

当自由度n很大时，由文献[4]知，$\left(\frac{\chi_n^2(0)}{n}\right)^{\frac{1}{3}}$渐近服从均值为$\left(1-\frac{2}{9n}\right)$方差为$\frac{2}{9n}$的高斯分布。因此检验统计量$T$可进一步化简为

$$Z = \left(\frac{1}{18P}\right)^{-\frac{1}{2}}\left[\left(\frac{T}{4P}\right)^{\frac{1}{3}} - 1 + \frac{1}{18P}\right] \sim N(0,1) \tag{32}$$

其中 $N(0,1)$ 表示零均值单位方差的高斯分布。

这种新定义的检验统计量将积分双谱、积分三谱结合起来，称为积分双谱、三谱混合型检验统计量，可用于平稳时间序列的高斯性检验。当 Z 大于某一常数 Z_0 时，(Z_0 由显著水平α决定)，则判决平稳时间序列 $\{x(n)\}$ 服从非高斯分布，否则判决平稳时间序列 $\{x(n)\}$ 服从高斯分布。

为同已有的积分双谱、积分三谱方法进行比较，将检验统计量 $\hat{S}_{3N}$、$\hat{S}_{4N}$ 也分别化简为

$$Z_3 = \left(\frac{1}{9P}\right)^{-\frac{1}{2}}\left[\left(\frac{\hat{S}_{3N}}{2P}\right)^{\frac{1}{3}} - 1 + \frac{1}{9P}\right] \sim N(0,1) \tag{33}$$

$$Z_4 = \left(\frac{1}{9P}\right)^{-\frac{1}{2}}\left[\left(\frac{\hat{S}_{4N}}{2P}\right)^{\frac{1}{3}} - 1 + \frac{1}{9P}\right] \sim N(0,1) \tag{34}$$

4　实例验证

为验证这种新定义的积分双谱、三谱混合型检验统计量，考虑以下几种平稳时间序列。

序列A，由MA(1)模型产生，即

$$x(n) = e(n) + 0.9e(n-1)$$

序列B，由AR(2)模型产生，即

$$x(n) = 0.3x(n-1) - 0.4x(n-2) + e(n)$$

序列C，由以下模型产生，即

$$x(n) = 0.4x(n-1) + 0.2x(n-1)e(n-1) + e(n)$$

序列D，由以下模型产生，即

$$x(n) = 0.4x(n-1) + 0.6x(n-1)e(n-1) + e(n)$$

以上序列中 $e(n)$ 为独立地满足 $N(0,1)$ 分布的随机变量。

序列E，由以下MA(2)模型产生，即

$$x(n) = 0.3e(n) - 0.4e(n-1) + 0.8e(n-2)$$

其中 $e(n)$ 为独立地满足均值为零的指数分布的随机变量。

序列F，由以下MA(1)模型产生，即

$$x(n) = e(n) + 0.4e(n-1) + 0.2e(n-2)$$

其中 $e(n)$ 为独立地在 $[-0.5, 0.5]$ 区间满足均匀分布的随机变量。

序列G，由以下模型产生，即

$$x(n) = \sin(2\pi \times 0.01n + \varphi)$$

其中相位 φ 为独立地在 $[-\pi, \pi]$ 区间服从均匀分布的随机变量。

序列H，由以下模型产生，即

$$x(n) = -0.9|x(n-1)|e(n)$$

其中 $e(n)$ 为独立地满足 $N(0,1)$ 分布的随机变量。

序列I，由以下模型产生，即

$$x(n) = -0.7|x(n-1)| + e(n)$$

其中 $e(n)$ 为独立地在 $[-0.5, 0.5]$ 区间满足均匀分布的随机变量。

在上述几个模型中，采用10次 Monte Carlo 试验，每次试验 $x(n)$ 的记录长度 $N = 1024$，$N_B = 128$，$K = 8$，$P = 63$，根据式(32)、(33)、(34) 计算出九个序列的 Z 值列于表1．查正态分布表，对于显著水平 $\alpha = 0.01$，知 $Z_0 = 2.33$。由表1可知，用积分双谱、三谱混合型检验统计量计算，结果表明序列 A、B 的 $Z < Z_0$，服从高斯分布，序列 C、D、E、F、G、H、I 的 $Z > Z_0$，服从非高斯分布。用积分双谱检验统计量计算，结果表明序列 A、B、F、G 的 $Z < Z_0$，服从高斯分布，序列 C、D、E、H、I 的 $Z > Z_0$，服从非高斯分布。用积分三谱检验统计量计算，结果表明序列 A、B、H、I 的 $Z < Z_0$，服从高斯分布，序列 C、D、E、F、G 的 $Z > Z_0$，服从非高斯分布。而实际序列 A、B 是服从高斯分布的，序列 C、D、E、F、G、H、I 是服从非高斯分布的。由此可见用积分双谱、三谱混合型检验统计量进行检验，比用积分双谱、积分三谱检验统计量检验的准确程度要高。

Table 1 The result of test

series	$Z3$	$Z4$	Z	$Z0$
A	0.407	0.521	0.626	2.33
B	–0.985	0.536	–0.312	2.33
C	12.131	3.433	11.807	2.33
D	18.682	14.067	23.314	2.33
E	6.992	6.252	9.341	2.33
F	–1.131	9.121	6.916	2.33
G	–0.525	22.159	20.046	2.33
H	20.056	1.683	18.571	2.33
I	18.498	0.024	16.468	2.33

5 结　　论

本文给出了用积分双谱、三谱混合型检验统计量识别平稳时间序列高斯性的基本方法，实例结果表明，对于用积分双谱、积分三谱检验统计量以及常用的双谱、三谱检验统计量无法识别的非高斯序列，该方法都可以进行识别。而且因为积分多谱估计是采用交叉谱而不是采用多谱的积分，其计算速度很快，所以该方法是切实有效的。

参 考 文 献

1 Giannakis G, Tsatsanis M. Time-domain tests for Gaussianity and time-reversibility. IEEE Trans Signal Processing, 1994, 42: 3460～3472

2 Tugnait J K. Detection of non-Gaussian signals using integrated polyspectrum. IEEE Trans Signal Processing, 1994, 42: 3137～3149

3 Brillinger D R. Time Series Data Analysis and Theory. New York: Holt, Rinehart and Winston, 1975

4 Kendall M G, Stuart A, The Advanced Theory of Statistics. New York: Macmillan, 1977, 1. 1, 4th ed.

Integrated Polyspectrum Method for Testing Gaussianity of Stationary Time Series

CONG Yu-Liang, WANG Shu-Xun, WANG Hong-Zhi, DAI Yi-Song
(*Dept. of Electronic Engineering, Jilin University of Technology*, *Changchun* 130025)

Abstract

In this paper, the method for test Gaussianity of a stationary time series via integrated polyspectrum is given. This method is also proved by the simulation experiments.

Key words: Integrated polyspectrum, Stationary time series, Test for Gaussianity

丛玉良　女, 1993年毕业于吉林工业大学电子系, 获硕士学位, 现任讲师, 在职博士, 主要从事微弱信号处理, 高阶累积量方面的研究工作, 已在国内外发表论文四十余篇。