谱分析方法在仿真结果分析中的运用*

张金槐

(国防科技大学自动控制系 长沙 410073)

摘 要 文中给出了谱估计的分析, 并将它直接应用于仿真系统输出量的特性分析。对于仿真的可信性问题, 讨论了相容性检验方法, 特别是小子样现场试验下, 仿真与现场试验之间的一致性问题。

关键词 谱分析, 分布检测, 仿真可信性

分类号 O212.2, TP391.9

Spectrum Analysis and Its Application in Simulation

Zhang Jinhuai

(Department of Automatic Control, NUDT, Changsha, 410073)

Abstract In this paper, the distribution of spectrum estimation is qiven, and using this result, analysis of statistical characteristics of the output of simulation system is studied. And then, the confidence of system simulation and the consistency of the data between simulation and true system are discussed.

Key words spectrum analysis, distribution test, confidence of simulation

对随机信号采集后的处理、系统特性分析、时序建模和分析、仿真输出量的特性分析等工程技术问题, 谱分析是一个基本方法。近年来, 特别是关于武器系统仿真输出特性分析, 谱分析方法已成为一种时尚。对于平稳过程, 谱分析方法已日趋成熟, 然而, 对于短时序、低信噪比或者变化比较剧烈的系统, 则尚有不少工作可做。

从谱估计来说, 人们总是要求有较小的分辨误差和方差。在有限观测时间内, 减少分辨误差将导致方差的增大。因此, 一个良好的谱估计总是在两者之间寻求一种折衷, 如经典的 Fourier 谱估计。然而它存在明显的缺陷, 于是 Burg 提出最大熵谱估计。Box 指出了这种方法与自回归谱分析法是等价的。由于一个 ARMA 模型总可以用长自回归模型作逼近, 因此, 近代谱估计, 常采用 AR 建模 (高阶) 。在可操作性方面, 也有较好的算法。对于变化剧烈的系统, 常采用自适应谱估计方法[1]。

在运用谱分析方法的过程中, 给出谱分布是十分重要的。但是在容量较小的场合, 极限分布在应用上受到限制。因此, 寻求谱估计的确切分布有着重要的作用。本文将对谱估计及其分布作必要的论述, 并给出在仿真中的应用。

1 谱估计的分布

众所周知, Burg 提出的最大熵谱估计是一种分辨力高, 且适用于短时序的方法。在工程技术中被广泛应用, 但是其对于噪声干扰比较敏感, 因此, 在运用谱分析作置信估计或检测中, 还是经常应用窗谱估计。窗谱估计可表示为

$$\hat{S}_w(\omega) = \frac{1}{2\pi}\left[w_0^{(N)}\hat{r}_0 + 2\sum_{k=1}^{M_N} w_k^{(N)}\hat{r}_k\cos\omega k\right], \omega \in [-\pi, \pi]$$

其中, $\hat{r}_k$ ($k=0, 1, \cdots, M_N$) 为平稳时序 $\{x_t, t\in\mathbf{Z}\}$ ($\mathbf{Z}$ 为整数集) 的样本相关函数, N 为样本容量,

* 1997 年 10 月 23 日收稿
第一作者: 张金槐, 男, 1930 年生, 教授

M_N 常取 $\sqrt{N}$ 或 $\sqrt[3]{N}$。$\{w_k^{(N)}\}$ 为窗函数序列。

窗谱估计是渐近无偏的，且当 $\frac{1}{N}\int_{-\pi}^{\pi}[W_N(\omega)]^2 d\omega \to 0$ 时，$D[\hat{S}_w(\omega)] \to 0$，即它是一致估计。这里，$W_N(\omega) = \frac{1}{2\pi}\sum_{k=-N+1}^{N-1} w_k^{(N)} e^{-ik\omega}$。

Jenkins 和 Watts 在 $\{x_t, t\in \mathbf{Z}\}$ 为平稳正态的场合，给出了窗谱估计的分布，证明了 $r\hat{S}_w(\omega)/S(\omega)$ 渐近于 χ^2 分布，这里，$r = 2N\Big/\Big[\sum_{k=-M_N}^{M_N}(w_k^{(N)})^2\Big]$。这样，可以立刻得到 $S(\omega)$ 的置信估计，$S(\omega)$ 的置信区间为 $\left[\frac{r\hat{S}_w(\omega)}{\chi^2_{r,\alpha/2}}, \frac{r\hat{S}_w(\omega)}{\chi^2_{r,1-\alpha/2}}\right]$，置信概率为 $1-\alpha$。

如果需要检验两个平稳时序谱密度是否相等，那么可以立即构成一种一致性检验的准则。事实上，设 $(x_1^{(1)},\cdots,x_{N_1}^{(1)})$ 和 $(x_1^{(2)},\cdots,x_{N_2}^{(2)})$ 分别为来自谱密度为 $S^{(1)}(\omega)$ 和 $S^{(2)}(\omega)$ 的独立子样，设

$$H_0: S^{(1)}(\omega) = S^{(2)}(\omega),$$

则用窗谱估计可得 $\hat{S}_w^{(i)}(\omega)\,(i=1,2)$，于是

$$r_i\hat{S}_w^{(i)}(\omega)/S^{(i)}(\omega) \xrightarrow{\sim} \chi^2_{r_i}\,(i=1,2)$$

此处

$$r_i = 2N_i\Big/\Big[\sum_{-M_{N_i}}^{M_{N_i}}(w_k^{N_i})^2\Big]$$

于是

$$F = \frac{r_1\hat{S}_w^{(1)}(\omega)/S^{(1)}(\omega)/r_1}{r_2\hat{S}_w^{(2)}(\omega)/S^{(2)}(\omega)/r_2} \xrightarrow{\sim} F_{r_1,r_2},$$

即 F 为具有 (r_1, r_2) 个自由度的 F 变量。

令

$$P\left\{\frac{\hat{S}_w^{(1)}(\omega)}{\hat{S}_w^{(2)}(\omega)} > F_{r_1,r_2,\alpha}\,\middle|\, H_0\right\} = \alpha$$

于是，当 $\hat{S}^{(1)}(\omega)/\hat{S}^{(2)}(\omega) > F_{r_1,r_2,\alpha}$，则拒绝 H_0，即认为两个时序的谱不一致。反之，则采纳 H_0，此时的检验水平为 α。

具体实施时，总是取 $\omega\in[-\pi,\pi]$，$i=1,\cdots,m$，在每个 ω 处验证频谱的一致性。

必须提出，运用极限分布方法，对短时序并不总是适合的，为此，必须寻求谱估计的确切分布。

对于有限数据 $x_1,\cdots,x_N$，不妨设 N 为偶数。用 Fourier 变换表示它们：

$$x_t = a_0 + \sum_{p=1}^{\frac{N}{2}-1}\left[a_p\cos\left(p\frac{2\pi}{N}t\right) + b_p\sin\left(p\frac{2\pi}{N}t\right)\right] + \frac{a_N}{2}\cos\pi t,\quad t=1,\cdots,N.$$

其中：

$$a_0 = \bar{x},$$

$$\frac{a_N}{2} = \frac{1}{N}\sum_{t=1}^{N}(-1)^t x_t,$$

$$a_p = \frac{2}{N}\sum_{t=1}^{N}x_t\cos\left(\frac{2\pi}{N}pt\right),\qquad b_p = \frac{2}{N}\sum_{t=1}^{N}x_t\sin\left(\frac{2\pi}{N}pt\right)$$

记 $\omega_p = \frac{2\pi}{N}\cdot p$，$\left(p=1,2,\cdots,\frac{N}{2}\right)$，它为 x_t 的第 p 次谐波的频率。p 次谐波为 $a_p\cos\omega_p t + b_p\sin\omega_p t = R_p\cos(\omega_p t + \phi_p)$，此处 $R_p = \sqrt{a_p^2 + b_p^2}$，$\phi_p = \arctan\left(-\frac{b_p}{a_p}\right)$，记

$$I(\omega_p) = \frac{1}{N\pi}\left[\left(\sum_{t=1}^{N}x_t\cos\omega_p t\right)^2 + \left(\sum_{t=1}^{N}x_t\sin\omega_p t\right)^2\right] = \frac{N}{4\pi}(a_p^2 + b_p^2) = \frac{N}{4\pi}R_p^2$$

$I(\omega_p)$ 就是熟知的周期图。

设 $\{x_t, t\in\mathbf{Z}\}$ 为平稳时序，且 $x_t \sim N(\mu,\sigma^2)$，则有

$$E[a_p] = E[b_p] = 0,\ p \neq N/2$$

$$\mathrm{Var}[a_p] = \mathrm{Var}[b_p] = \frac{2\sigma^2}{N}, p \neq N/2, \quad \mathrm{Cov}(a_p, b_p) = 0$$

于是 $\frac{a_p}{2\sigma^2/N}$ 和 $\frac{b_p}{2\sigma^2/N}$ 均为 $N(0,1)$ 变量，且它们为独立的，因此

$$\frac{a_p^2 + b_p^2}{2\sigma^2/N} = \frac{2\pi}{\sigma^2} I(\omega_p) \sim \chi_2^2,$$

于是 $\mathrm{E}[I(\omega)] \to S(\omega)$ $(N \to \infty)$，但 $\mathrm{Var}[I(\omega)] = \frac{\sigma^4}{\pi^2} \to 0, N \to \infty$。因此，周期图不是一致估计。为此作平滑处理，将谱估计改写成周期图的平均，即

$$\hat{S}(\omega) = \frac{1}{m}\sum_{j=1}^{m} I(\omega), \omega = \frac{2\pi}{N} j, j = 1, \cdots, m。$$

这里，j 在 m 个顺序整数内，使 $\omega, \cdots, \omega_n$ 对 ω 是对称的。若取 m 为奇数，$m = 2m^* + 1$，则 ω_p 处的谱估计取作 $\hat{S}(\omega_p) = \frac{1}{m}\sum_{j=-m^*}^{m^*} I(\omega_p + \frac{2\pi}{N} j)$。

由于 $\frac{2\pi}{\sigma^2} I(\omega_p) \sim \chi_2^2$，于是 $m\left(\frac{2\pi}{\sigma^2}\right)\hat{S}(\omega_p) \sim \chi_{2m}^2$。这样，就确定了平滑周期图的确切分布，$m$ 的选取考虑在分辨力与方差之间进行折衷，工程实践中，常取 $m \leq \frac{N}{40}$。

2 仿真系统输出量的统计分析

本节论述：①仿真动态输出与真实试验结果的相容性问题；②运用仿真模型作预报及其置信度分析。

关于①，可以用前部分给出的谱估计的分布作相容性检验。这里不赘述。下面介绍一种新方法，它特别地适用于真实试验次数特别小的场合。假定仿真系统动态输出为平稳过程，由输出时序所作出的谱估计记 $\hat{S}^{(1)}(\omega)$；而真实试验所得到的谱估计为 $\hat{S}^{(2)}(\omega)$。现要判定两个谱估计是否属于同一总体。

对于第 i 次仿真的动态输出时序 $(x_1^{(i)}, \cdots, x_N^{(i)})$，可以计算出谱估计，记作 $\hat{S}_i^{(1)}(\omega)$，$i = 1, \cdots, M$. M 是仿真的次数，于是 $\hat{S}_1^{(1)}(\omega), \cdots, \hat{S}_M^{(1)}(\omega)$ 就是仿真所获得的谱估计子样。对于真实试验，如武器系统的靶场试验，在同一状态下，设共进行了 S 次，得时序 $(x_1^{(j)}, \cdots, x_N^{(j)})$，$j = 1, \cdots, S$，于是可算得 $\hat{S}_j^{(2)}(\omega)$，$j = 1, \cdots, S$，因此 $\hat{S}_1^{(2)}(\omega), \cdots, \hat{S}_S^{(2)}(\omega)$ 就是现场试验所获得的谱估计子样。这样，运用非参数检验法，可以检验两子样是否属于同一总体，如秩检验方法等。但是，要特别指出的是现场试验（如导弹的全程试验）次数特少，如 $S = 1$ 或 2，此时的相容性检验是人们关心的。这里，对于最极端的情况，如 $S = 1$，进行相容性检验。

在真实试验只作一次的场合下，由试验的动态结果的时序 $(x_1, \cdots, x_N)$，可算得谱估计 $\hat{S}_1^{(2)}(\omega)$。仿真子样 $(\hat{S}_1^{(1)}(\omega), \cdots, \hat{S}_M^{(1)}(\omega))$ 的容量 M 较大，于是将两个子样混合排序（由小到大），得

$$\hat{S}_{(1)}(\omega) \leq \hat{S}_{(2)}(\omega) \leq \cdots \leq \hat{S}_{(j-1)}(\omega) \leq \hat{S}_j^{(*)}(\omega) \leq \hat{S}_{(j+1)}(\omega) \leq \cdots \leq \hat{S}_{(M+1)}(\omega)$$

其中 $S_j^{(*)}(\omega) = \hat{S}_1^{(2)}(\omega)$，引入统计假设

H_0: $S_1^{(2)}(\omega)$ 与仿真结果的谱估计具有同一总体。

此时，在 H_0 为真时，$S_j^{(*)}(\omega)$ 的分布函数为

$$F_j(x) = \frac{(M+1)!}{(j-1)!(M+1-j)!}\int_0^{F(x)} t^{j-1}(1-t)^{n+1-j}\,\mathrm{d}t$$

其中 $F(x)$ 为总体分布，取 α 为某个小的概率值，它使

$$P\{S_{(j)}^*(\omega) < L_1\} + P\{S_j^*(\omega) > L_2\} = \alpha$$

可以给出检验规律如下：给定显著水平 α，使

$$P\{S_{(j)}^*(\omega) < L_1\} = F_j(L_1) = \alpha/2,$$

$$P\{S_{(j)}^*(\omega) > L_2\} = 1 - F_j(L_2) = \alpha/2,$$

查不完全 Beta 函数表，可获得 L_1, L_2，于是

当 $S^*_{(j)}(\omega) < L_1$ 或 $S^*_{(j)}(\omega) > L_2$，则拒绝 H_0;

当 $L_1 \le S^*_j(\omega) \le L_2$，则采纳 H_0。

这种检验是在 $F(x)$ 为已知时进行的。如果 $F(x)$ 为未知，则由 $(\hat{S}_1^{(1)}(\omega), \cdots, \hat{S}_M^{(1)}(\omega))$ 作抽样分布 $F_M(x)$，以代替 $F(x)$。

当验证了仿真结果与真实试验结果相容之后，我们要运用仿真输出所建立的模型作预报，并作置信度分析。

假定仿真输出时序，已建立的模型为 ARMA(p,q) 模型为

$$\hat{\Phi}(B)\hat{Z}_t = \hat{\theta}(B)\epsilon$$

其中 $\tilde{Z}_t = x_t - \bar{x}, \hat{\Phi}(B) = 1 - \hat{\Phi}B - \cdots - \hat{\Phi}_p B^q, \hat{\theta}(B) = 1 - \hat{\theta}B - \cdots - \hat{\theta}_l B^q$。$B$ 为迟后一步算子，于是

$$\tilde{Z}_t = \hat{\Phi}^{-1}(B)\hat{\theta}(B)\epsilon_t = \sum_{j=0}^{\infty} G_j B^j \epsilon_t = \sum_{j=0}^{\infty} G_j \epsilon_{-j}$$

式中 G_j 为 Green 函数。由此，在未来时刻 $t+l$，有

$$\tilde{Z}_{t+l} = \sum_{j=0}^{\infty} G_j \epsilon_{+l-j} = G_0\epsilon_{+l} + G_1\epsilon_{+l-1} + \cdots + G_{l-1}\epsilon_{+1} + \sum_{j=0}^{\infty} G_{l+j}\epsilon_{-j}$$

记 $t+l$ 时刻的预测估计为 $\tilde{Z}_t(l)$，且将它表示为 $\epsilon, \epsilon_{-1}, \cdots$ 的加权和，即

$\tilde{Z}_t(l) = \sum_{j=0}^{\infty} G^*_{l+j}\epsilon_{t-j}$。权系数序列 $\{G^*_{l+j}\}$ 的选择，使 $E[(\tilde{Z}_{t+l} - \tilde{Z}_t(l))^2] = \min$，易知

$$\tilde{Z}_t(l) = E[\tilde{Z}_{t+l} | \tilde{Z}_t, \tilde{Z}_{t-l}, \cdots] \qquad (*)$$

可知，$\tilde{Z}_t(l) = \sum_{j=0}^{\infty} G_{l+j}\epsilon_{-j}$。因此，在实现预报计算时，只需确定条件数学期望就可以了。

对预报的置信度分析，在 $\{\epsilon\}$ 为零均值正态白噪声下，可知

$$\Delta\tilde{Z}_{t+l} = \tilde{Z}_{t+l} - \tilde{Z}_t(l) = \sum_{j=0}^{l-1} G_j\epsilon_{+l-j}, \Delta Z_{t+l} \sim N\left(0, \sqrt{1 + \sum_{j=1}^{l-1} G_j^2}\right)$$

于是，在置信度为 $1-\alpha$ 之下，$\tilde{Z}_{t+l}$ 的置信区间是

$$\left[\tilde{Z}_t(l) - u_\alpha\sqrt{1 + \sum_{j=1}^{l-1} G_j^2}\,\sigma_\epsilon, \tilde{Z}_t(l) + u_\alpha\sqrt{1 + \sum_{j=1}^{l-1} G_j^2}\,\sigma_\epsilon\right]$$

例如，α 与 u_α 的关系由正态表查得，如 $\alpha = 0.05$ 时，$u_\alpha = 1.96$.

对于非平稳时序的处理，这是工程实践感兴趣的问题，一些学者研究了瞬时谱分析方法，目前还有待进一步研究[2]。近年来，工程应用中，对于一般非平稳时序，常用下列模型来描述:

$$x_t = \sum_{j=0}^{s} A_j e^{k_j t} + \eta_t$$

其中 η 是离散时间的 ARMA(p,q) 模型表示的时序，即 η_t 满足: $\Phi(B)\eta_t = \theta(B)\epsilon$。$k_j$ 可以为复数。这样，常将所获得的时序 $\{x_t\}$ 用指数函数逼近，例如用 L. S 方法，使 A_j, k_j 的估计 $\hat{A}_j, \hat{k}_j$ 满足 $\sum_t \left|x_t - \sum \hat{A}_j e^{\hat{k}_j t}\right|^2 = \min$。这样，可得 $\hat{\eta}_t = x_t - \sum_{j=0}^{s} \hat{A}_j e^{\hat{k}_j t}$。对时序 $\{\hat{\eta}_t\}$ 建立 ARMA(p,q) 模型，它的谱估计就是

$$\hat{S}_\eta(\omega) = \frac{\sigma^2}{2\pi}\frac{|1 - \hat{\theta}_1 e^{j\omega} - \cdots - \hat{\theta}_q e^{jq\omega}|^2}{|1 - \varphi_1 e^{j\omega} - \cdots - \varphi_p e^{jp\omega}|^2}, \quad \omega \in [-\pi, \pi]$$

于是，非平稳时序建模及其分析，可以近似地化作平稳时序的建模和分析问题。

参考文献

1 张金槐. 时序频谱的自适应估计. 计算技术与自动化，1996，10

2 Loynes R M. On the concept of the spectrum for non-stationay processes. J. R. Statist. Soc. 1968, 30: 1～30