

现代信号处理

——高阶统计量信号处理方法

Signal Processing with Higher-order Statistics

授课教师： 吴嗣亮 教授

联系方式： 四号教学楼 106

专题内容

一、概述	1
二、高阶统计量的定义、性质和估计	5
(一) 高阶矩、高阶累积量及其谱	5
(二) 高阶累积量与高阶谱的性质	11
(三) 高阶累积量与高阶谱的估计	19
三、非最小相位系统的辨识	21
(一) 基本问题	21
(二) MA 系统的辨识	25
(三) ARMA 系统的辨识	35
四、谐波恢复	42
(一) 基本问题	42
(二) 谐波恢复的高阶累积量方法	43
五、空间窄带信号源的波达方向估计	46
(一) 基本问题	46
(二) 基于二阶统计量的 DOA 估计方法及其不足	47
(三) 基于高阶统计量的 DOA 估计方法	53

一、概述

高阶统计量(Higher-order Statistics)是指比二阶统计量更高阶的随机变量或随机过程的统计量。

二阶统计量有：

- 随机变量（矢量）：方差、协方差（相关矩）、二阶矩。
- 随机过程：自相关函数、功率谱、互相关函数、互功率谱、自协方差函数等。

高阶统计量有：

- 随机变量（矢量）：高阶矩(Higher-order Moment)，高阶累积量(Higher-order Cumulant)。
- 随机过程：高阶矩、高阶累积量、高阶谱(Higher-order Spectra, Polyspectra) 。

从统计学的角度，对正态分布的随机变量（矢量），用一阶和二阶统计量就可以完备地表示其统计特征。如对一个高斯分布的随机矢量，知道了其数学期望和协方差矩阵，就可以知道它的联合概率密度函数。对一个高斯随机过程，知道了均值和自相关函数（或自协方差函数），就可以知道它的概率结构，即知道它的整个统计特征。

但是，对不服从高斯分布的随机变量（矢量）或随机过程，一阶和二阶统计量不能完备地表示其统计特征。或者说，信息没有全部包含在一、二阶统计量中，更高阶的统计量中也包含了大量有用的信息。高阶统计量信号处理方法，就是从**非高斯信号的高阶统计量中提取信号的有用信息**，

特别是从一、二阶统计量中无法提取的信息的方法。从这个角度来说，高阶统计量方法不仅是对基于相关函数或功率谱的随机信号处理方法的重要补充，而且可以为二阶统计量方法无法解决的许多信号处理问题提供手段。可以毫不夸张地说，凡是使用功率谱或相关函数进行过分析与处理，而又未得到满意结果的任何问题，都值得重新试用高阶统计量方法。

高阶统计量的概念于 1889 年提出。

高阶统计量的研究始于六十年代初，主要是数学家和统计学家们在做基础理论的研究，以及针对光学、流体动力学、地球物理、信号处理等领域特定问题的应用研究。直到八十年代中、后期，在信号处理和系统理论领域才掀起了高阶统计量方法的研究热潮。标志性的事件有：

1. K. S. Lii , M. Rosenblatt “Deconvolution and Estimation of Transfer Function Phase and Coefficients for non-Gaussian Linear Processes”. Ann. Statistics, Vol.10, pp.1195-1208, 1982

首次用高阶统计量解决了非最小相位系统的盲辨识问题。

2. C. L. Nikias, M. R. Raghuveer 的综述文章 “Bispectrum Estimation: A Digital Signal Processing Framework” 在 Proc. IEEE 发表，1987 July.
3. 1989、1991、1993、1995、1997、1999 年举办了六届关于高阶统计量的信号处理专题研讨会（海军研究办公室，NSF， IEEE Control System Society, IEEE ASSP Society, IEEE Geoscience and Remote sensing Society）。
4. IEEE Trans. on AC 1990 年 1 月专辑。
5. IEEE Trans. on ASSP 1990 年 7 月专辑。

6. J. M. Mendel 的综述文章”Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory: Theoretical Results and Some Applications”. Proc. IEEE, 1991 (主要是关于非最小相位系统辨识)。
7. C. L. Nikias & A. P. Petropula 的专著 Higher-order Spectral Analysis: A Nonlinear Processing Framework, 由 Prentice-Hall 1993 出版。
8. Signal Processing, 1994 4 月专辑。
9. Circuits, Systems, and Signal Processing, 1994.6 月专辑。

高阶统计量方法已在雷达、声纳、通信、海洋学、电磁学、等离子体物理、结晶学、地球物理、生物医学、故障诊断、振动分析、流体动力学等领域的信号处理问题中获得应用。典型的信号处理应用包括系统辨识与时间序列分析建模、自适应估计与滤波、信号重构、信号检测、谐波恢复、图像处理、阵列信号处理、盲反卷积与盲均衡等。

在信号处理中使用高阶统计量的主要动机可以归纳成四点：

- 1、抑制未知功率谱的加性有色噪声的影响。
- 2、辨识非最小相位系统或重构非最小相位信号。

自相关函数或功率谱是相盲的，即不包含信号或系统的相位信息。仅当系统或信号是最小相位时，二阶统计量的方法才能获得正确的结果。相反，高阶统计量既包含了幅度信息，又保留了信号的相位信息，因而可以用来解决非最小相位系统的辨识或非最小相位信号的重构问题。

3、提取由于高斯性偏离带来的各种信息

对于非高斯信号，其高阶统计量中也包含了大量的信息。对模式识别、

信号检测、分类等问题，有可能从高阶统计量获得信号的显著分类特征。

4、检测和表征信号中的非线性以及辨识非线性系统。

如用来解决非线性引起的二次、三次相位耦合问题。

参考资料：

- 1、 张贤达,《时间序列分析—高阶统计量方法》,清华大学出版社,1996。
- 2、 沈凤麟等,《生物医学随机信号处理》(第9章),中国科学技术大学出版社,1999。
- 3、 J. M. Mendel. “Tutorial on Higher-order Statistics (Spectra) in Signal Processing and Systems Theory: Theoretical Results and Some Applications” . Proc. IEEE, Vol.79, pp.278-305, 1991
- 4、 C. L. Nikias & A. P . Petropulu. Higher-order Spectral Analysis: A Nonlinear Processing Framework. Prentice-Hall. 1993
- 5、 C. L. Nikias & J. M. Mendel. “Signal Processing with Higher-order Spectra” . IEEE Signal Processing Magazine, Vol.10, July, pp.10-37, 1993
- 6、 C. L. Nikias & M. R. Raghuveer. “Bispectrum Estimation: A Digital Signal Processing Framework” . Proc. IEEE, Vol.75, pp.869-891, 1987
- 7、 P. A. Delaney & D. O. Walsh. “A Bibliography of Higher-Order Spectra and Cumulants”. IEEE Signal Processing Magazine, Vol.11 July, pp. 61-70, 1994
- 8、 J. A. Cadzow. “Blind Deconvolution via Cumulant Extrema” . IEEE Signal Processing Magazine, Vol.13, No.3, pp.24-42, 1996
- 9、 www.ant.uni-bremen.edu.de/HOSHOME/

二、高阶统计量的定义、性质和估计

(一) 高阶矩、高阶累积量及其谱

(从随机变量→随机矢量→随机过程)

1、随机变量的特征函数与累积量

定义： 设随机变量 x 具有概率密度 $f(x)$ ，其特征函数定义为

$$\Phi(s) = \int_{-\infty}^{+\infty} f(x)e^{sx} dx = E\{e^{sx}\}$$

其中 s 为特征函数的参数。

(可看作 $f(x)$ 的拉普拉斯变换)

特征函数 $\Phi(s)$ 只是参数 s 的函数。对 $\Phi(s)$ 求 k 次导数，可得

$$\Phi^k(s) = E\{x^k e^{sx}\}$$

因此

$$\Phi^k(0) = E\{x^k\} = m_k$$

也就是说 $\Phi(s)$ 在原点的 k 阶导数等于 x 的 k 阶矩 m_k 。因此， $\Phi(s)$ 也称作矩生成函数（又叫第一特征函数）。

矩生成函数可以唯一地、完全地确定一个概率分布。这可由矩生成函数唯一性定理阐明：

定理： 设 $F(x)$ 和 $G(x)$ 是具有相同矩生成函数的分布函数，即：

$$\int_{-\infty}^{\infty} e^{sx} dF(x) = \int_{-\infty}^{\infty} e^{sx} dG(x)$$

则 $F(x) \equiv G(x)$

由矩生成函数可以定义随机变量 x 的累积量生成函数（又叫第二特征函数）及累积量。

定义：设随机变量 x 的矩生成函数为 $\Phi(s)$ ，则函数

$$\Psi(s) = \ln \Phi(s)$$

称为 x 的累积量生成函数，而 $\Psi(s)$ 在原点的 k 阶导数

$$c_k = \left. \frac{d^k \Psi(s)}{ds^k} \right|_{s=0}$$

称为 x 的 k 阶累积量。

如果将 $\Phi(s)$ 和 $\Psi(s)$ 展开成 Taylor 级数，根据以上定义，就会有

$$\Phi(s) = 1 + m_1 s + \frac{1}{2} m_2 s^2 + \cdots + \frac{1}{k!} m_k s^k + \cdots$$

$$\Psi(s) = c_1 s + \frac{1}{2} c_2 s^2 + \cdots + \frac{1}{k!} c_k s^k + \cdots$$

也就是说， x 的 k 阶矩和累积量分别是其矩生成函数和累积量生成函数的 Taylor 级数展开中 s^k 项的系数。

2、随机矢量的特征函数与累积量

定义：令 $\mathbf{x}=[x_1, x_2, \cdots, x_k]^T$ 是一随机矢量，且 $\mathbf{s}=[s_1, s_2, \cdots, s_k]^T$ ，则随机矢量 \mathbf{x} 的矩生成函数定义为

$$\Phi(s_1, s_2, \cdots, s_k) = E \left\{ e^{s_1 x_1 + s_2 x_2 + \cdots + s_k x_k} \right\} = E \left\{ e^{\mathbf{s}^T \mathbf{x}} \right\}$$

\mathbf{x} 的累积量生成函数定义为

$$\Psi(s_1, s_2, \dots, s_k) = \ln \Phi(s_1, s_2, \dots, s_k)$$

\mathbf{x} 的 (v_1, v_2, \dots, v_k) 阶矩和累积量分别定义为矩生成函数和累积量生成函数的

Taylor级数展开中 $s_1^{v_1} s_2^{v_2} \dots s_k^{v_k}$ 项的函数，即

$$m_{v_1, v_2, \dots, v_k} = \left. \frac{\partial^v \Phi(s_1, s_2, \dots, s_k)}{\partial s_1^{v_1} \partial s_2^{v_2} \dots \partial s_k^{v_k}} \right|_{s_1=s_2=\dots=s_k=0} = E\{x_1^{v_1} \dots x_k^{v_k}\}$$

$$c_{v_1, v_2, \dots, v_k} = \left. \frac{\partial^v \Psi(s_1, s_2, \dots, s_k)}{\partial s_1^{v_1} \partial s_2^{v_2} \dots \partial s_k^{v_k}} \right|_{s_1=s_2=\dots=s_k=0}$$

其中， $v=v_1+v_2+\dots+v_k$ 。

对 $v_1=v_2=\dots=v_k=1$ 的特殊情况，记随机矢量 \mathbf{x} 的矩和累积量分别为

$$m_k = m_{1,1,\dots,1} = \text{mom}(x_1, \dots, x_k)$$

$$c_k = c_{1,1,\dots,1} = \text{cum}(x_1, \dots, x_k)$$

我们下面将用它们来定义随机过程的高阶矩和累积量。

3、随机过程的高阶矩和高阶累积量

定义： 设 $\{x(n)\}$ 为 k 阶平稳随机过程，则该过程的 k 阶矩定义为

$$m_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) = \text{mom}\{x(n), x(n+\tau_1), \dots, x(n+\tau_{k-1})\}$$

而 k 阶累积量定义为

$$c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) = \text{cum}\{x(n), x(n+\tau_1), \dots, x(n+\tau_{k-1})\}$$

根据这一定义，平稳随机过程的 k 阶矩和 k 阶累积量实质上就是取 $x_1=x(n)$, $x_2=x(n+\tau_1)$, \dots , $x_k=x(n+\tau_{k-1})$ 之后的随机矢量 $[x(n), x(n+\tau_1), \dots$,

$x(n+\tau_{k-1})]^T$ 的 $v_1=v_2=\cdots=v_k=1$ 时的矩和累积量。由于 k 阶平稳性假设，随机过程的 k 阶矩和 k 阶累积量均只有 $k-1$ 个独立的变元，它们仅仅是滞后 $\tau_1, \tau_2, \cdots, \tau_{k-1}$ 的函数，而与时间 n 无关。

根据这一定义，随机过程 $\{x(n)\}$ 的

① 一阶矩和一阶累积量为：

$$m_{1x} = c_{1x} = E\{x(n)\} \quad (\text{均值})$$

② 二阶矩和二阶累积量为：

$$m_{2x}(\tau_1) = E\{x(n)x(n+\tau_1)\} = r_x(\tau_1) \quad (\text{自相关函数})$$

$$\begin{aligned} c_{2x}(\tau_1) &= m_{2x}(\tau_1) - m_{1x}^2 \\ &= \text{cov}\{x(n), x(n+\tau_1)\} \quad (\text{自协方差函数}) \end{aligned}$$

仅当随机过程有零均值时， $m_{2x}(\tau_1) = c_{2x}(\tau_1) = r_x(\tau_1)$

③ 三阶矩和三阶累积量为：

$$\begin{aligned} m_{3x}(\tau_1, \tau_2) &= E\{x(n)x(n+\tau_1)x(n+\tau_2)\} \\ c_{3x}(\tau_1, \tau_2) &= \text{cum}\{x(n), x(n+\tau_1), x(n+\tau_2)\} \\ &= m_{3x}(\tau_1, \tau_2) - m_{1x}[m_{2x}(\tau_1) + m_{2x}(\tau_2) + m_{2x}(\tau_1-\tau_2)] + 2m_{1x}^2 \\ &= E\{[x(n) - m_{1x}][x(n+\tau_1) - m_{1x}][x(n+\tau_2) - m_{1x}]\} \\ &\quad (\text{三阶中心矩}) \end{aligned}$$

仅当随机过程有零均值时， $m_{3x}(\tau_1, \tau_2) = c_{3x}(\tau_1, \tau_2)$

④ 四阶矩和四阶累积量

$$\begin{aligned} m_{4x}(\tau_1, \tau_2, \tau_3) &= E\{x(n)x(n+\tau_1)x(n+\tau_2)x(n+\tau_3)\} \\ c_{4x}(\tau_1, \tau_2, \tau_3) &= \text{cum}\{x(n), x(n+\tau_1), x(n+\tau_2), x(n+\tau_3)\} \end{aligned}$$

$$\begin{aligned}
&= m_{4x}(\tau_1, \tau_2, \tau_3) - m_{2x}(\tau_1)m_{2x}(\tau_3 - \tau_2) - m_{2x}(\tau_2)m_{2x}(\tau_3 - \tau_1) - m_{2x}(\tau_3)m_{2x}(\tau_2 - \tau_1) - \\
&\quad m_{1x}[m_{3x}(\tau_2 - \tau_1, \tau_3 - \tau_1) + m_{3x}(\tau_2, \tau_3) + m_{3x}(\tau_1, \tau_3) + m_{3x}(\tau_1, \tau_2)] - \\
&\quad (m_{1x})^2[m_{2x}(\tau_1) + m_{2x}(\tau_2) + m_{2x}(\tau_3) + m_{2x}(\tau_3 - \tau_1) + m_{2x}(\tau_3 - \tau_2) + m_{2x}(\tau_2 - \tau_1)] - \\
&\quad 6(m_{1x})^4
\end{aligned}$$

当随机过程的均值 $m_{1x}=0$ 时,

$$\begin{aligned}
c_{4x}(\tau_1, \tau_2, \tau_3) &= m_{4x}(\tau_1, \tau_2, \tau_3) - m_{2x}(\tau_1)m_{2x}(\tau_3 - \tau_2) - m_{2x}(\tau_2)m_{2x}(\tau_3 - \tau_1) - \\
&\quad m_{2x}(\tau_3)m_{2x}(\tau_2 - \tau_1)
\end{aligned}$$

对 $k > 4$ 的任意 k 阶累积量, 也可以由 **M-C 公式** 用 k 阶及 k 阶以下的各阶矩表示出来。反过来任意 k 阶矩也可由 **C-M 公式** 用 k 阶及其以下的各阶累积量表示(1977 年 Brillinger 首先建立)。由于 M-C 公式和 C-M 公式很复杂, 而实际应用中一般很少用到高于 4 阶的统计量, 我们也就不再介绍。

从上面的讨论可以看出:

- ✧ 当随机过程的均值不为零时, 二阶以上的矩和累积量是不相同的。
- ✧ 如果随机过程的均值为零, 则三阶以下的矩和累积量相等, 而高于三阶的矩和累积量则一般均不等。

4、高斯过程的矩和累积量

根据定义, 可以证明平稳高斯过程 $\{x(n)\}$ 的矩和累积量为:

$$\begin{cases} m_{1x} = E\{x(n)\} \\ m_{2x}(\tau) = r_x(\tau) \\ m_{kx}(\tau_1, \dots, \tau_{k-1}) = \begin{cases} \equiv 0, & \text{若 } k \geq 3 \text{ 且为奇数} \\ \neq 0, & \text{若 } k \geq 4 \text{ 且为偶数} \end{cases} \end{cases}$$

$$\begin{cases} c_{1x} = E\{x(n)\} \\ c_{2x}(\tau) = \text{COV}\{x(n), x(n+\tau)\} = r_x(\tau) - m_{1x}^2 \\ c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) \equiv 0, \quad \text{若 } k \geq 3 \end{cases}$$

可见，高斯过程的高阶累积量恒等于零，即使对有色高斯过程也是这样。这是在信号处理中使用高阶累积量的主要动机之一。高阶矩不具备这一优点。

5、高阶矩谱和高阶累积量谱

大家知道，对平稳随机过程 $\{x(n)\}$ ，它的功率谱密度定义为其自相关函数的 Fourier 变换。与这相类似，我们可以分别由高阶矩和高阶累积量的多维 Fourier 变换定义高阶矩谱和高阶累积量谱。

定义： 设 $m_{kx}(\tau_1, \dots, \tau_{k-1})$ 绝对可和，即

$$\sum_{\tau_1=-\infty}^{+\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{+\infty} |m_{kx}(\tau_1, \dots, \tau_{k-1})| < \infty$$

则 k 阶矩谱定义为 k 阶矩的 $k-1$ 维 Fourier 变换，即

$$M_{kx}(\omega_1, \omega_2, \dots, \omega_{k-1}) = \sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{+\infty} m_{kx}(\tau_1, \dots, \tau_{k-1}) \exp[-j \sum_{i=1}^{k-1} \omega_i \tau_i]$$

定义： 设高阶累积量 $c_{kx}(\tau_1, \dots, \tau_{k-1})$ 绝对可和，即

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} |c_{kx}(\tau_1, \cdots, \tau_{k-1})| < \infty$$

则 k 阶累积量谱定义为 k 阶累积量的 $k-1$ 维 Fourier 变换, 即

$$S_{kx}(\omega_1, \cdots, \omega_{k-1}) = \sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{k-1}=-\infty}^{\infty} c_{kx}(\tau_1, \cdots, \tau_{k-1}) \cdot \exp[-j \sum_{i=1}^{k-1} \omega_i \tau_i]$$

习惯上, **高阶累积量谱**常简称为**高阶谱或多谱(Polyspectra)**。最常用的高阶谱是三阶谱(双谱, Bispectrum):

$$B_x(\omega_1, \omega_2) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} c_{3x}(\tau_1, \tau_2) e^{-j(\omega_1 \tau_1 + \omega_2 \tau_2)}$$

和四阶谱(三谱, Trispectrum):

$$T_x(\omega_1, \omega_2, \omega_3) = \sum_{\tau_1=-\infty}^{\infty} \sum_{\tau_2=-\infty}^{\infty} \sum_{\tau_3=-\infty}^{\infty} c_{4x}(\tau_1, \tau_2, \tau_3) e^{-j(\omega_1 \tau_1 + \omega_2 \tau_2 + \omega_3 \tau_3)}$$

显然, 对高斯随机过程, 其双谱、三谱以及更高阶的谱恒等于零。

(二) 高阶累积量与高阶谱的性质

1、高阶累积量的性质 (常用的六条)

CP1、若 $\lambda_i (i = 1, 2, \cdots, k)$ 是常数, 且 $x_i (i = 1, 2, \cdots, k)$ 为随机变量, 则

$$\text{cum}(\lambda_1 x_1, \lambda_2 x_2, \cdots, \lambda_k x_k) = \left(\prod_{i=1}^k \lambda_i \right) \cdot \text{cum}(x_1, x_2, \cdots, x_k)$$

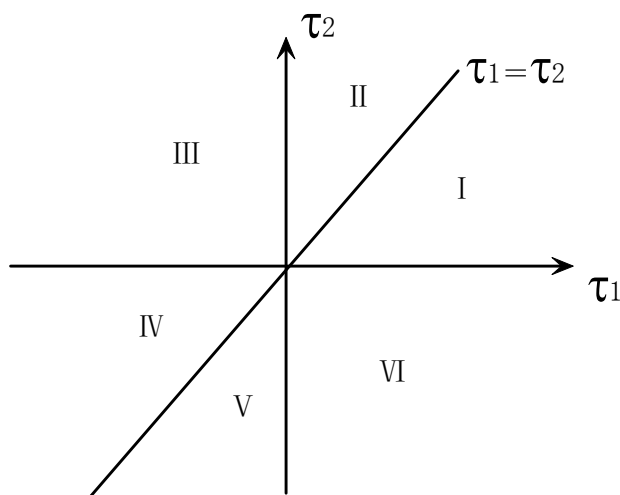
CP2、累积量相对于其变元是对称的, 即

$$\text{cum}(x_1, x_2, \cdots, x_k) = \text{cum}(x_{i_1}, x_{i_2}, \cdots, x_{i_k})$$

其中 (i_1, \dots, i_k) 是 $(1, \dots, k)$ 的任一个排列。

根据这一性质， k 阶累积量有 $k!$ 种对称形式。以三阶累积量为例，共有 $3! = 6$ 种对称形式：

$$\begin{aligned} c_{3x}(m, n) &= c_{3x}(n, m) = c_{3x}(-n, m-n) = c_{3x}(m-n, -m) \\ &= c_{3x}(m-n, -n) = c_{3x}(-m, n-m) \end{aligned}$$



三阶累积量的对称区域

如图所示，因此只要获得了这六个扇形区域中任意一个扇形内的三阶累积量，我们都可以求出全部三阶累积量序列。

CP3、累积量相对于其变元是加性的，即

$$\text{cum}(x_1 + y_1, x_2, \dots, x_k) = \text{cum}(x_1, x_2, \dots, x_k) + \text{cum}(y_1, x_2, \dots, x_k)$$

也就是和的累积量等于累积量之和（“累积量”由此得名）。

CP4、若随机变量 $\{x_i\}$ 与随机变量 $\{y_i\}$ 统计独立，则

$$\text{cum}(x_1 + y_1, \dots, x_k + y_k) = \text{cum}(x_1, \dots, x_k) + \text{cum}(y_1, \dots, y_k)$$

这一性质说明，两个统计独立随机过程之和的累积量等于各个随机过程累积量之和。因此，如果一个非高斯信号 $\{x(n)\}$ 的观测受到与之独立的加性有色高斯噪声 $\{e(n)\}$ 的污染，即

$$y(n) = x(n) + e(n)$$

那么观测过程的高阶累积量就等于非高斯信号的高阶累积量，即

$$\begin{aligned} c_{ky}(\tau_1, \tau_2, \dots, \tau_{k-1}) &= c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) + c_{ke}(\tau_1, \tau_2, \dots, \tau_{k-1}) \\ &= c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) \end{aligned}$$

也就是高阶累积量从理论上可以完全抑制高斯有色噪声的影响。仅此一点，就可以让我们对高阶累积量方法产生很大的兴趣。

这一重要的结论对高阶矩不成立。

CP5、若 k 个随机变量 $\{x_i\}$ 的一个子集同其它部分独立，则

$$\text{cum}(x_1, x_2, \dots, x_k) \equiv 0$$

由这个性质可说明，独立同分布(i.i.d.)随机序列的累积量为 δ 函数，即若 $\{w(n)\}$ 是独立同分布过程，则

$$c_{kw}(\tau_1, \dots, \tau_{k-1}) = \gamma_{kw} \delta(\tau_1) \dots \delta(\tau_{k-1})$$

因此，这种随机过程的高阶谱是多维平坦的。文献中把这种噪声叫做高阶白噪声。

对 *i.i.d* 随机过程，其高阶矩却不是 δ 函数。

CP6、若 α 是一常数，则

$$\text{cum}(\alpha + x_1, x_2, \dots, x_k) = \text{cum}(x_1, x_2, \dots, x_k)$$

高阶矩不具备这一性质。

根据前面的讨论，高阶矩和高阶累积量存在着一些重要的差别，归纳起来有三点：

① 高斯过程其高于二阶的累积量及其谱恒等于零。因而非零的高阶累积量或高阶谱提供了一种高斯性偏离程度的度量。高阶矩却不具备这一点（如下可进一步看清这一点）。

对零均值的平稳过程 $\{x(n)\}$ ，可以证明其三、四阶累积量满足

$$c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) = m_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1}) - E\{g(n)g(n+\tau_1)\dots g(n+\tau_{k-1})\} \\ (k=3,4)$$

其中 $\{g(n)\}$ 为与 $\{x(n)\}$ 具有相同的功率谱或相同的自相关函数的高斯过程（功率谱等价的高斯过程）。

从这个式中，可以直接看出，高阶累积量确实是反映了过程 $\{x(n)\}$ 的高斯性偏离度。

② 正如白噪声的自相关函数或协方差函数是脉冲函数，其功率谱是平坦的，（高阶）白噪声的高阶累积量是多维脉冲函数，而多谱是多维平坦的。高阶矩和高阶矩谱则不然。

③ 统计独立随机过程和的累积量等于各随机过程累积量的和，而高阶矩则不然。

正是由于这些差别，造成了在高阶统计量信号处理中，主要以高阶累积量和高阶谱而很少以高阶矩和高阶矩谱作为数学工具。

考虑计算复杂性，实际问题中主要用到三、四阶累积量以及双谱和三

谱。

对具有对称分布的随机过程，其三阶累积量恒等于零，因此在有些问题中，必须采用四阶以上的累积量。

2、高阶谱的性质

① 高阶谱一般为多维复函数，即它们具有幅度和相位：

$$s_{kx}(\omega_1, \dots, \omega_{k-1}) = |s_{kx}(\omega_1, \dots, \omega_{k-1})| \exp\{j\Phi_{kx}(\omega_1, \dots, \omega_{k-1})\}$$

这是与功率谱明显不同的一点。对随机过程 $\{x(n)\}$ 的功率谱 $P_x(\omega)$ ，我们知道， $P_x(\omega)$ 为实数且

$$P_x(\omega) \geq 0$$

因此功率谱不含相位信息。

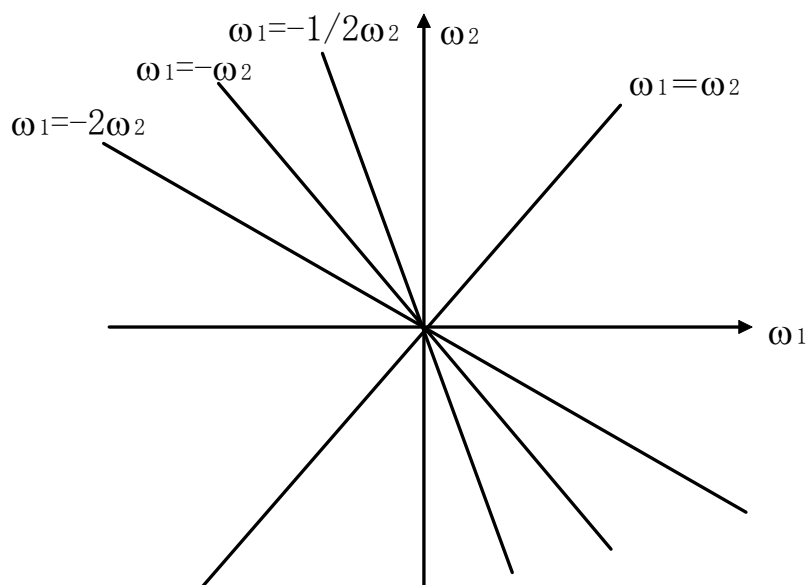
② 高阶谱存在丰富的对称性。

高阶谱的对称性是由于高阶累积量的对称性产生的。

对于双谱，有：

$$\begin{aligned} B_x(\omega_1, \omega_2) &= B_x(\omega_2, \omega_1) = B_x(-\omega_2, -\omega_1) = B_x(-\omega_1 - \omega_2, \omega_2) \\ &= B_x(\omega_1, -\omega_1 - \omega_2) = B_x(-\omega_1 - \omega_2, \omega_1) = B_x(\omega_2, -\omega_1 - \omega_2) \end{aligned}$$

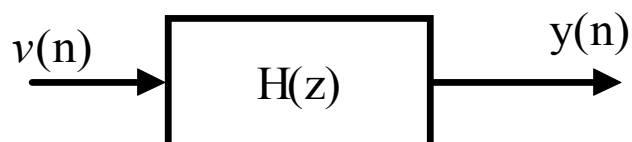
即存在如图所示的 12 个对称区域。



三阶累积量的对称区域

对于三谱，已有人证明存在 96 个对称区域。

③ 设 $H(z)$ 为有限维数、线性移不变、指数稳定的系统的传递函数，



若系统输入 $\{v(n)\}$ 为平稳非高斯随机过程，其 k 阶谱为 $S_{kv}(\omega_1, \omega_2, \dots, \omega_{k-1})$ ，

则系统输出 $\{y(n)\}$ 的 k 阶谱为

$$\begin{aligned}
 & S_{ky}(\omega_1, \omega_2, \dots, \omega_{k-1}) \\
 &= S_{kv}(\omega_1, \omega_2, \dots, \omega_{k-1}) \cdot H(e^{j\omega_1}) H(e^{j\omega_2}) \dots H(e^{j\omega_{k-1}}) \\
 & \quad \cdot H\left(\exp\left\{-j\sum_{i=1}^{k-1} \omega_i\right\}\right)
 \end{aligned}$$

写成多维 Z 变换形式, 有

$$\begin{aligned}
& S_{ky}(z_1, z_2, \dots, z_{k-1}) \\
&= S_{kv}(z_1, z_2, \dots, z_{k-1}) \cdot H(z_1)H(z_2) \cdots H(z_{k-1}) \cdot H\left(-\prod_{i=1}^{k-1} z_i^{-1}\right)
\end{aligned}$$

根据这一性质, 对于双谱有

$$\begin{aligned}
& B_y(\omega_1, \omega_2) \\
&= B_v(\omega_1, \omega_2) \cdot H(e^{j\omega_1})H(e^{j\omega_2}) \cdot H(e^{-j(\omega_1+\omega_2)}) \\
&= B_v(\omega_1, \omega_2) \cdot H(e^{j\omega_1})H(e^{j\omega_2}) \cdot H^*(e^{j(\omega_1+\omega_2)})
\end{aligned}$$

对于三谱有

$$\begin{aligned}
& T_y(\omega_1, \omega_2, \omega_3) \\
&= T_v(\omega_1, \omega_2, \omega_3) \cdot H(e^{j\omega_1})H(e^{j\omega_2})H(e^{j\omega_3})H^*(e^{j(\omega_1+\omega_2+\omega_3)})
\end{aligned}$$

我们知道, $\{v(n)\}$ 和 $\{y(n)\}$ 的功率谱之间存在以下关系:

$$\begin{aligned}
P_y(\omega) &= P_v(\omega) \cdot H(e^{j\omega}) \cdot H^*(e^{j\omega}) \\
&= P_v(\omega) \cdot |H(e^{j\omega})|^2
\end{aligned}$$

不难看出, 线性系统输入、输出过程高阶谱之间的关系是功率谱之间关系的直接扩展。同时, 我们也有再次看到了功率谱是相盲的, 系统输出的功率谱中不含有系统的相位信息。

功率谱和高阶谱存在以下主要的差别: **高阶谱和高阶累积量可反映系统的非线性。**

根据谱表示定理, 任何一个规则平稳随机过程 $\{x(n)\}$ 的功率谱 $P_x(\omega)$ 都可以分解为

$$P_x(\omega) = \sigma^2 \cdot H(e^{j\omega}) \cdot H^*(e^{j\omega})$$

其中 $H(z)$ 为有限维数、线性移不变、指数稳定的系统的传递函数。也就是说，从功率谱等价的角度，随机过程 $\{x(n)\}$ 总可以看成是方差为 σ^2 的零均值白噪声激励一个稳定的线性移不变系统 $H(z)$ 产生的有色随机过程。

一个非高斯平稳随机过程的高阶谱却一般并不总能表示为：

$$\begin{aligned} S_{kx}(\omega_1, \omega_2, \dots, \omega_{k-1}) \\ = \gamma_{kw} H(e^{j\omega_1}) H(e^{j\omega_2}) \dots H(e^{j\omega_{k-1}}) H\left(\exp\left\{-j \sum_{i=1}^{k-1} \omega_i\right\}\right) \end{aligned}$$

除非 $\{x(n)\}$ 确实是一个线性非高斯过程，即是由非高斯白噪声激励一个稳定的线性移不变系统产生的。

从这个角度来说，功率谱不能反映信号或系统的非线性，而高阶谱或高阶累积量却可以检测和表征信号或系统中的非线性。

定义随机过程 $\{x(n)\}$ 的双相干因子为

$$K_{3x}(\omega_1, \omega_2) = \frac{B_x(\omega_1, \omega_2)}{\sqrt{P_x(\omega_1)P_x(\omega_2)P_x(\omega_1 + \omega_2)}}$$

三相干因子为

$$K_{4x}(\omega_1, \omega_2, \omega_3) = \frac{T_x(\omega_1, \omega_2, \omega_3)}{\sqrt{P_x(\omega_1)P_x(\omega_2)P_x(\omega_3)P_x(\omega_1 + \omega_2 + \omega_3)}}$$

对线性过程， $K_{3x}(\omega_1, \omega_2)$ 和 $K_{4x}(\omega_1, \omega_2, \omega_3)$ 的幅度将恒等于一个常数，即与 $\omega_1, \omega_2, (\omega_3)$ 无关。

对于线性非高斯过程，高阶谱的一个直接的应用是重构其功率谱：

考虑一个线性非高斯信号 $\{x(n)\}$ ，其观测受到与之独立的未知加性有色

高斯噪声 $\{e(n)\}$ 的污染，即

$$y(n) = x(n) + e(n)$$

如果要估计 $\{x(n)\}$ 的功率谱 $P_x(\omega)$ ，可以先得到 $\{y(n)\}$ 的双谱 $B_y(\omega_1, \omega_2) = B_x(\omega_1, \omega_2)$ 或三谱 $T_y(\omega_1, \omega_2, \omega_3) = T_x(\omega_1, \omega_2, \omega_3)$ ，由

$$\begin{aligned} B_x(\omega, 0) &= \gamma_{3w} H(e^{j\omega}) \cdot H(1) \cdot H^*(e^{-j\omega}) \\ &= \gamma_{3w} \cdot H(1) \cdot \sigma^2 \cdot \left\{ \sigma^2 \cdot |H(e^{j\omega})|^2 \right\} \\ &= \gamma_{3w} \cdot H(1) / \sigma^2 \cdot P_x(\omega) \end{aligned}$$

可以获得

$$P_x(\omega) = B_x(\omega, 0) \cdot a$$

结果只相差一个比例常数 a 未知

由

$$T_x(\omega, 0, 0) = \gamma_{4w} H^2(1) / \sigma^2 \cdot P_{x(\omega)}$$

可以获得

$$P_{x(\omega)} = T_x(\omega, 0, 0) \cdot b$$

结果只相差一个比例常数 b 未知。

无疑，这比从 $\{y(n)\}$ 估计 $\{x(n)\}$ 的功率谱要好得多（观测噪声无影响）。

（三）高阶累积量与高阶谱的估计

1、高阶累积量的估计

在实际应用中，我们是无法知道随机过程各阶累积量的真值的。通常所有的仅是随机过程的一个样本 $\{x(n); n=1, 2, \dots, N\}$ 。对各态遍历的平稳随

机过程，我们可以用随机过程的时间平均代替集总平均来计算均值、方差、自相关函数等统计量。

同样，对各态遍历的平稳随机过程，其高阶累积量也可由时间平均代替集总平均来估计。例如，三阶累积量可估计为

$$\hat{c}_{3x}(\tau_1, \tau_2) = \frac{1}{N} \sum_{n=1}^N [x(n) - \hat{m}_{1x}] [x(n + \tau_1) - \hat{m}_{1x}] [x(n + \tau_2) - \hat{m}_{1x}]$$

其中
$$\hat{m}_{1x} = \frac{1}{N} \sum_{n=1}^N x(n)$$

$$x(n)=0 \quad (n>N, n\leq 0)$$

对 $\tau_1 \neq 0, \tau_2 \neq 0$ ，这是三阶累积量 $c_{3x}(\tau_1, \tau_2)$ 的有偏估计，但当 $N \rightarrow \infty$ 时，偏差 $\rightarrow 0$ ，即为渐近无偏的。

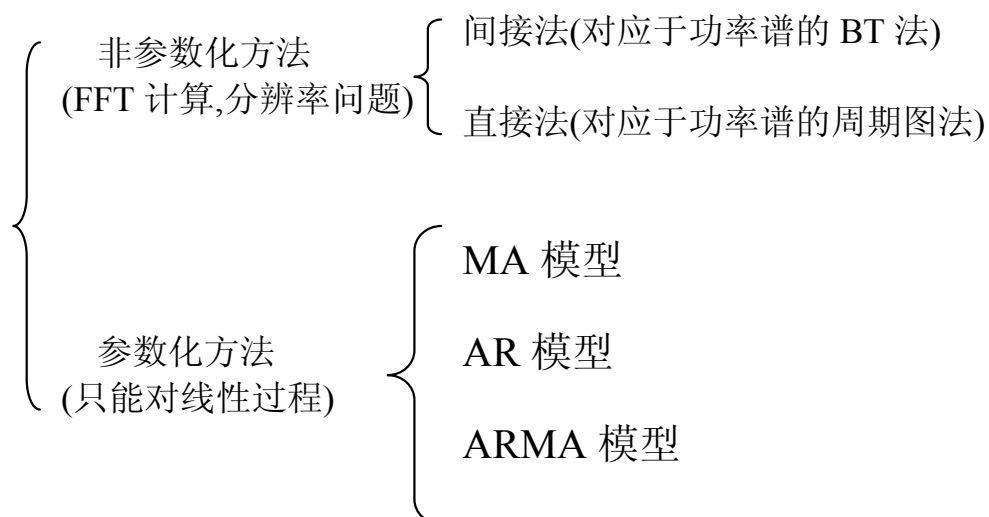
类似地，可以从 $x(n)$ 的一个样本获得及其任意 k 阶累积量的渐近无偏估计 $\hat{c}_{kx}(\tau_1, \dots, \tau_{k-1})$ 。 $\hat{c}_{kx}(\tau_1, \dots, \tau_{k-1})$ 以概率收敛于 $c_{kx}(\tau_1, \dots, \tau_{k-1})$ 的充分条件是： $x(n)$ 的前 $2k$ 阶累积量绝对可和，即：

$$\sum_{\tau_1=-\infty}^{\infty} \cdots \sum_{\tau_{m-1}=-\infty}^{\infty} |c_{mx}(\tau_1, \dots, \tau_{m-1})| < \infty \quad m = 1, 2, \dots, 2k$$

这时， $\hat{c}_{kx}(\tau_1, \dots, \tau_{k-1})$ 是 k 阶累积量的渐近无偏的一致估计。

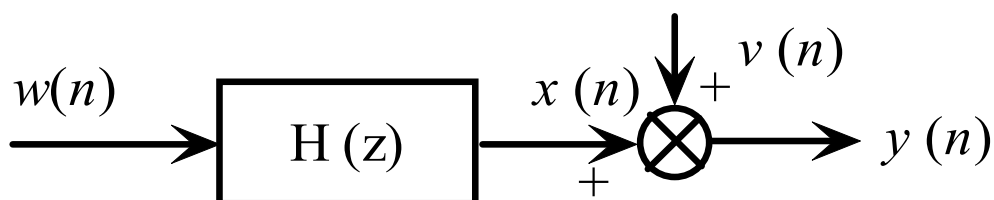
2、高阶谱的估计

同功率谱估计一样，高阶谱估计方法可分为



三、非最小相位系统的辨识

(一) 基本问题



这里我们要讨论的，也是信号处理和系统理论中经常遇到的基本问题是：一个指数稳定的线性移不变系统，它的传递函数为 $H(z)$,

$$H(z) = \sum_{i=-\infty}^{\infty} h(i)z^{-i} = B(z)/A(z) = \left[\sum_{j=0}^q b_j z^{-j} \right] / \left[\sum_{i=0}^p a_i z^{-i} \right]$$

$\{h(i)\}$ 为系统的单位脉冲响应。系统的输入 $\{w(n)\}$ 为独立同分布的 (i. i. d.)

平稳非高斯随机过程，且

$$E\{w(n)\} = 0$$

$$E\{w^2(n)\} = \sigma_w^2$$

$$c_{kw}(\tau_1, \dots, \tau_{k-1}) = \gamma_{kw} \cdot \delta(\tau_1) \cdots \delta(\tau_{k-1})$$

系统的输出为

$$x(n) = \sum_{i=-\infty}^{+\infty} h(i)w(n-i)$$

对系统输出的观测量为：

$$y(n) = x(n) + v(n)$$

$\{v(n)\}$ 为与 $\{w(n)\}$ 独立的加性观测噪声。要求只根据观测量 $\{y(n)\}$ 辨识系统的传递函数 $H(z)$ 。

由于以下三个方面的原因使这个问题较难解决：

1、功率谱等价现象引起的非最小相位系统辨识问题

我们知道，系统输出平稳过程 $\{x(n)\}$ 的功率谱为

$$P_x(\omega) = |H(e^{j\omega})|^2 \cdot \sigma_w^2$$

可见，系统输出的功率谱只反映了系统的幅频特性 $|H(e^{j\omega})|$ ，而不能反映系统的相频特性。

一个问题是，由系统输出的功率谱 $P_x(\omega)$ 能否唯一地确定 $H(z)$ ，或者说系统的幅频特性 $|H(e^{j\omega})|$ 是否唯一地对于传递函数 $H(z)$ 。这个问题的一般性答案是：

①如果已知 $H(z)$ 是最小相位的，或已知 $H(z)$ 是最大相位的，则 $|H(e^{j\omega})|$ 与 $H(z)$ 一一对应，由 $P_x(\omega)$ 可以唯一确定 $H(z)$ 。

②如果无任何先验知识，或 $H(z)$ 是非最小（大）相位的，则 $|H(e^{j\omega})|$ 与 $H(z)$ 不存在一一对应关系，由 $P_x(\omega)$ 不能唯一确定 $H(z)$ 。

一个简单的例子：对于 $0 < a < 1, 0 < b < 1$,

$$H_1(z)=(1-az^{-1})(1-bz^{-1}) \quad (\text{最小相位})$$

$$H_2(z)=(1-az)(1-bz) \quad (\text{最大相位})$$

$$H_3(z)=(1-az)(1-bz^{-1}) \quad (\text{混合相位})$$

$$H_4(z)=(1-az^{-1})(1-bz) \quad (\text{混合相位})$$

这四个系统具有相同的幅频特性，即对同一输入 $\{w(n)\}$ ，它们的输出将有相同的功率谱。

功率谱对应的系统模型的不唯一现象称为功率谱等价现象，也叫做模型的多重性。由于功率谱与自相关函数构成Fourier变换时，功率谱等价也叫做自相关等价。

由于这一现象，仅由系统输出的功率谱或自相关函数，将不可能辨识出非最小相位系统的传递函数，即使不存在观测噪声。

尽管功率谱等价的不同系统，它们在同一激励下的输出有完全相同的功率谱和自相关函数，但是，对同一非高斯分布的输入，它们的高阶谱或高阶累积量可能将明显不同。利用这一点，可以仅仅从系统的输出中辨识非最小相位系统的传递函数。

2、由自相关函数估计 MA 参数的困难

对一个 MA 系统

$$H(z) = \sum_{i=0}^q b_i z^{-i} \quad b_0 = 1$$

即使我们知道它是最小相位或最大相位系统，而且不考虑观测噪声，要从系统的输出中估计 MA 的参数也有困难。这要求解 $q+1$ 个变量的以下非线性方程组：

$$\begin{cases} \sigma_w^2(b_0^2 + b_1^2 + \cdots + b_q^2) = r_x(0) \\ \sigma_w^2(b_0b_1 + b_1b_2 + \cdots + b_{q-1}b_q) = r_x(1) \\ \vdots \\ \sigma_w^2b_0b_q = r_x(q) \end{cases}$$

这个方程组的求解并没有十分有效的方法。

对于 ARMA 系统的辨识，其 MA 部分的参数估计也存在同样的困难。

后面我们将看到，采用高阶累积量不仅可以线性地估计 MA 系统或者 ARMA 系统的 MA 部分的参数，而且可以获得 MA 参数的闭式解。

3、观测噪声的影响

对最小相位 MA 系统的辨识，即使观测噪声 $\{v(n)\}$ 为白噪声，由于

$$r_y(\tau) = r_x(\tau) + \sigma_v^2 \cdot \delta(\tau) \quad 0 \leq \tau \leq q$$

采用自相关函数估计 MA 参数时，方程组中未知参数的个数为 $q+2$ ，多于方程数，不能保证解的唯一性。如果忽略噪声的影响，即认为 $\sigma_v^2 = 0$ ，则估计所得的参数将是有偏的和非一致的。

对 AR 系统或 ARMA 系统 AR 部分的辨识，当观测噪声 $\{v(n)\}$ 为白噪声或 MA 噪声时，可以通过求解修正的 Yule-walker 方程，由自相关函数获得 AR 参数的无偏一致估计。但是，如果观测噪声 $\{v(n)\}$ 为 AR 或 ARMA 噪声，由自相关函数估计 AR 参数不仅是高度非线性的，而且估计结果是有偏的和非一致的。

由于高阶累积量固有的对高斯噪声的抑制能力，采用高阶累积量的系统辨识方法，可以自动抑制白色和有色高斯观测噪声对估计结果的影响。另外，由于对称分布随机过程的三阶累积量恒等于零，如果系统输入 $\{w(n)\}$

服从非对称分布，采用三阶累积量辨识方法，将自动抑制任何对称分布观测噪声的影响。

以上三方面的原因，使线性系统盲辨识的高阶累积量方法有重要的意义。

(二) MA 系统的辨识

考虑 q 阶 MA 系统

$$H(z) = \sum_{i=0}^q b(i)z^{-i}$$

其中， $b(0)=1$, $b(q) \neq 0$ 。

根据累积量的性质 CP1 和 CP3，可以证明，系统输出 $\{x(n)\}$ 的 k 阶累积量为

$$c_{kx}(\tau_1, \dots, \tau_{k-1}) = \gamma_{kw} \cdot \sum_{i=0}^q b(i)b(i+\tau_1)b(i+\tau_2) \cdots b(i+\tau_{k-1})$$

特别地

$$\begin{aligned} c_{3x}(\tau_1, \tau_2) &= \gamma_{3w} \sum_{i=0}^q b(i)b(i+\tau_1)b(i+\tau_2) \\ c_{4x}(\tau_1, \tau_2, \tau_3) &= \gamma_{4w} \sum_{i=0}^q b(i)b(i+\tau_1)b(i+\tau_2)b(i+\tau_3) \end{aligned}$$

由于

$$b(i) = 0 \quad (i > q, i < 0)$$

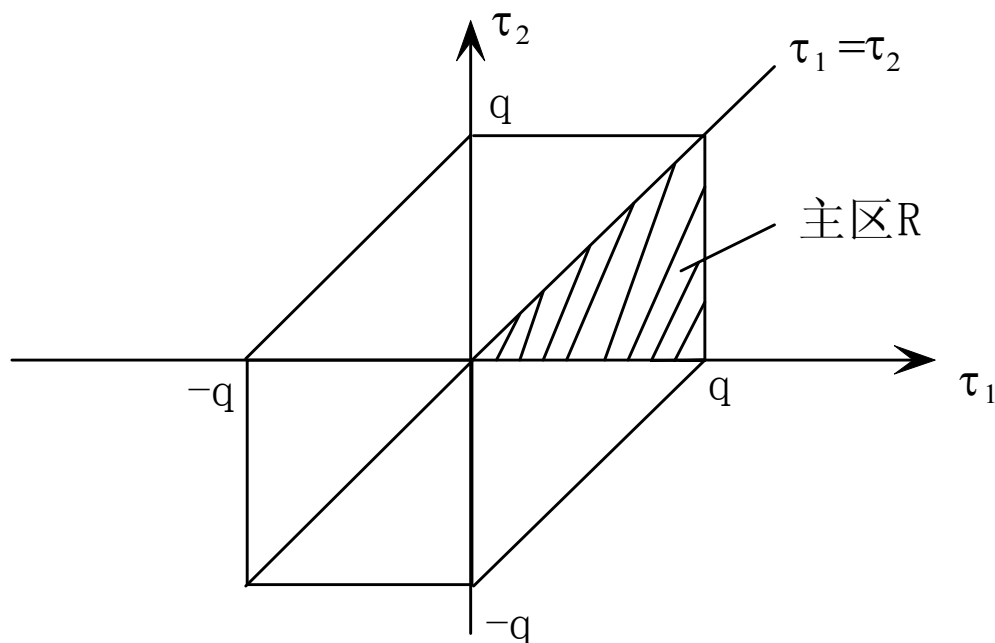
因此只要 $\tau_1, \dots, \tau_{k-1}$ 中的任何一个 τ_m 满足

$$|\tau_m| > q$$

则

$$c_{kx}(\tau_1, \dots, \tau_{k-1}) = 0$$

不难验证，对 q 阶 MA 系统， $c_{3x}(\tau_1, \tau_2) \neq 0$ 的区域如下图所示：



利用高阶累积量辨识 MA 系统的方法，可以分成三类：

$$\begin{cases} \text{闭式解} \\ \text{线性代数解} \\ \text{非线性优化} \end{cases}$$

1、闭式解

1) $C(q, n)$ 算法

Giannakis 提出

由

$$c_{kx}(\tau_1, \dots, \tau_{k-1}) = \gamma_{kw} \cdot \sum_{i=0}^q b(i)b(i+\tau_1)b(i+\tau_2)\cdots b(i+\tau_{k-1})$$

可得：

$$\begin{cases} c_{kx}(q, n, 0, \dots, 0) = \gamma_{kw} \cdot b(q) \cdot b(n) & (n = 1, 2, \dots, q) \\ c_{kx}(q, 0, 0, \dots, 0) = \gamma_{kw} \cdot b(q) \end{cases}$$

因此，MA 的参数可以计算为

$$b(n) = \frac{c_{kx}(q, n, 0, \dots, 0)}{c_{kx}(q, 0, \dots, 0)} \quad (n = 1, 2, \dots, q)$$

显然，如果观测噪声 $\{v(n)\}$ 服从高斯分布，不论是否为白色的，有

$$c_{ky}(\tau_1, \tau_2, \dots, \tau_{k-1}) = c_{kx}(\tau_1, \tau_2, \dots, \tau_{k-1})$$

因此，MA 的参数可以从观测量的累积量中**计算**。

如果系统输入 $\{w(n)\}$ 服从非对称分布，即 $\gamma_{3w} \neq 0$ ，则观测噪声也可以为任意对称分布的有色噪声（采用三阶累积量）。

优点：简单、自动抑制观测噪声的影响

缺点：对样本累积量的估计误差敏感，数值性能差，仅有理论意义。

2) RC 闭式递推算法

这是一类**既使用三阶累积量又使用自相关函数**的算法。

考虑 $\{v(n)\}$ 为高斯或对称分布的白噪声，方差为 σ_v^2 ，观测量 $\{y(n)\}$ 的自相关函数为

$$r_y(\tau) = \sigma_w^2 \sum_{i=0}^q b(i)b(i+\tau) + \sigma_v^2 \cdot \delta(\tau)$$

$\{y(n)\}$ 的三阶累积量的**对角切片**(slice)为：

$$c_{3y}(\tau, \tau) = c_{3x}(\tau, \tau) = \gamma_{3w} \sum_{i=0}^q b(i)b^2(i+\tau)$$

① 令 $\tau = \pm q$ ，可得

$$\begin{cases} r_y(q) = \sigma_w^2 b(q) \\ c_{3y}(q, q) = \gamma_{3w} b^2(q) \\ c_{3y}(-q, -q) = \gamma_{3w} b(q) \end{cases}$$

可以解出：

$$\begin{cases} b(q) = \frac{c_{3y}(q, q)}{c_{3y}(-q, -q)} \\ \sigma_w^2 = \frac{r_y(q) \cdot c_{3y}(-q, -q)}{c_{3y}(q, q)} \\ \gamma_{3w} = \frac{c_{3y}^2(-q, -q)}{c_{3y}(q, q)} \end{cases}$$

② 再令 $\tau = \pm(q-1)$ ，可得：

$$\begin{cases} c_{3y}(q-1, q-1) = \gamma_{3w} [b^2(q-1) + b(1)b^2(q)] \\ c_{3y}(1-q, 1-q) = \gamma_{3w} [b(q-1) + b(q) \cdot b^2(1)] \\ r_y(q-1) = \sigma_w^2 [b(q-1) + b(1)b(q)] \end{cases}$$

求解该方差组，可以获得：

$$\begin{cases} b(q-1) = \frac{1}{2} r_y(q-1) / \sigma_w^2 + \frac{b(q) \cdot c_{3y}(1-q, 1-q) - c_{3y}(q-1, q-1)}{2\gamma_{3w} [b(q) - r_y(q-1) / \sigma_w^2]} \\ b(1) = [r_y(q-1) / \sigma_w^2 - b(q-1)] / b(q) \end{cases}$$

③ 再依次令 $\tau = \pm(q-2) \cdot \pm(q-3), \dots, \pm[\frac{q}{2}]$ ，可以递推获得：

$$b(q-2), b(2); b(q-3), b(3); \dots; b(q - [\frac{q}{2}]) \cdot b([\frac{q}{2}])$$

缺点：

- ① 观测噪声只能为高斯白噪声或其它对称分布白噪声。
- ② 递推计算中存在 0/0 的隐患。
- ③ 不能扩展到四阶以上的累积量，因而要求 $\{w(n)\}$ 服从非对称分布。
- ④ 对样本累积量的估计误差敏感，存在误差传播，仅有理论意义。

Tugnait 利用三阶累积量切片 $c_{3y}(\tau, \tau + q)$ 与自相关函数，也提出了一种 RC 闭式递推算法。这种算法不存在 0/0 的隐患，但同样存在其它缺点。

3) 累积量闭式递推算法[张贤达等]

由

$$c_{ky}(\tau_1, \dots, \tau_{k-1}) = \gamma_{kw} \sum_{i=0}^q b(i)b(i + \tau_1) \cdots b(i + \tau_{k-1})$$

可以得到

$$\begin{aligned} c_{ky}(\tau, 0, \dots, 0) &= \gamma_{kw} \sum_{i=0}^q b^{k-1}(i)b(i + \tau) \\ &= \gamma_{kw} \sum_{i=0}^q b^{k-1}(i - \tau)b(i) \end{aligned}$$

而由 $C(q, n)$ 公式，有：

$$b(n) = \frac{c_{ky}(q, n, 0, \dots, 0)}{c_{ky}(q, 0, \dots, 0)}$$

代入整理后可得

$$\gamma_{kw} \sum_{i=0}^q b(i)c_{ky}^{k-1}(q, i - \tau, 0, \dots, 0) = c_{ky}(\tau, 0, \dots, 0) \cdot c_{ky}^{k-1}(q, 0, \dots, 0)$$

- ① 令 $\tau = -q$ ，可得

$$\gamma_{kw} = \frac{c_{ky}(-q, 0, \dots, 0) \cdot c_{ky}^{k-1}(q, 0, \dots, 0)}{c_{ky}^{k-1}(q, q, 0, \dots, 0)}$$

②令 $\tau = q$, 可得

$$\gamma_{kw} b(q) \cdot c_{ky}^{k-1}(q, 0, \dots, 0) = c_{ky}^k(q, 0, \dots, 0)$$

可解出

$$b(q) = \frac{c_{ky}(q, 0, \dots, 0)}{\gamma_{kw}}$$

③再令 $\tau = q - 1$, 可得

$$\gamma_{kw} b(q) \cdot c_{ky}^{k-1}(q, 1, 0, 0, \dots, 0) + \gamma_{kw} b(q-1) \cdot c_{ky}^{k-1}(q, 0, 0, \dots, 0) = c_{ky}(q-1, 0, \dots, 0) \cdot c_{ky}^{k-1}(q, 0, \dots, 0)$$

可解出

$$b(q-1) = \frac{c_{ky}(q-1, 0, \dots, 0)}{\gamma_{kw}} - \frac{b(q) \cdot c_{ky}^{k-1}(q, 1, 0, \dots, 0)}{c_{ky}^{k-1}(q, 0, \dots, 0)}$$

④ 依次再令 $\tau = q-2, q-3, \dots, 1$, 可以递推获得

$$b(q-m) = \frac{c_{ky}(q-m, 0, \dots, 0)}{\gamma_{kw}} - \frac{1}{c_{ky}^{k-1}(q, 0, \dots, 0)} \cdot \sum_{i=1}^m b(q-m+i) \cdot c_{ky}^{k-1}(q, i, 0, \dots, 0)$$

$$(m = 2, 3, \dots, q-1)$$

优点: 适用于有色高斯噪声, 如果 $\{w(n)\}$ 服从非对称分布, 用三阶累积量求解, 即取 $k = 3$, 也适用于任何对称分布的有色噪声情况。

缺点: 对样本累积量的估计误差敏感, 存在误差传播, 仅有理论意义。

2、线性代数解

闭式解虽然简单，但它们存在共同的缺点，即不存在对样本累积量估计误差的平滑，因而数值性能很差。引入线性代数解的目的，就是为了对样本累积量的估计误差进行平滑，降低估计结果对样本累积量估计误差的敏感性。

1) GM 算法

Giannakis和Mendel证明了以下GM方程：

$$\sum_{i=0}^q b^2(i) r_x(m-i) = (\sigma_w^2 / \gamma_{3w}) \sum_{i=0}^q b(i) c_{3x}(m-i, m-i) \quad -q \leq m \leq 2q$$

（三阶累积量对角切片元素与自相关函数之间的关系）

视该方程中 $b^2(1), \dots, b^2(q), \sigma_w^2 / \gamma_{3w}, \sigma_w^2 / \gamma_{3w} b(1), \dots, \sigma_w^2 / \gamma_{3w} b(q)$ 为 $2q+1$ 个独立变量，用线性最小二乘法求解这组 $(3q+1)$ 方程，求得 $2q+1$ 个变量的解，并进一步可求得 $b(1), b(2), \dots, b(q)$ 的估计。

对于四阶累积量情况，可以证明：

$$\sum_{i=0}^q b^3(i) r_x(m-i) = (\sigma_w^2 / \gamma_{4w}) \sum_{i=0}^q b(i) c_{4x}(m-i, m-i, m-i) \quad -q \leq m \leq 2q$$

（四阶累积量对角切片元素与自相关函数之间的关系）

因而可以类似地用线性最小二乘法求解，得到 $b(1), \dots, b(q)$ 的估计。

关于 GM 算法，有几个问题需要注意：

① 算法把 $b^2(i)$ 视作与 $b(i)$ 独立的变量，而实际上不是，因此这种估计只

能是**次优**的。

② 方程组可能**不存在唯一解**。因为矩阵方程中矩阵的秩有可能不等于 $2q+1$ 。

③ 算法**只适用于无加性观测噪声的特殊情况**。即使对加性白色观测噪声，GM 方程中也只有 $-q \leq m \leq -1$ 和 $q+1 \leq m \leq 2q$ 对应的 $2q$ 个方程不受噪声影响（不含 $r_x(0)$ 项的方程）。

2) 广义 GM 方程与 T 算法

Tugnait 证明了以下**广义 GM 方程**：

$$\sum_{i=0}^q b(i)b(i+k)r_x(m-i) = (\sigma_w^2 / \gamma_{3w}) \sum_{i=0}^q b(i)c_{3x}(m-i, m-i+k)$$

它将关于三阶累积量对角切片元素与自相关函数关系的 GM 方程推广到**三阶累积量任意次对角切片元素与自相关函数的关系**。

不难看出，取 $k=0$ ，广义 GM 方程退化成 GM 方程。

取 $k=q$ ，广义 GM 方程退化成

$$b(q) \cdot r_x(m) = (\sigma_w^2 / \gamma_{3w}) \cdot \sum_{i=0}^q b(i)c_{3x}(m-i, m-i+q)$$

或者

$$\sum_{i=1}^q b(i)c_{3x}(m-i, m-i+q) - [b(q) \cdot \gamma_{3w} / \sigma_w^2] r_x(m) = -c_{3x}(m, m+q)$$
$$-q \leq m \leq q$$

它们是关于 $b(1), b(2), \dots, b(q)$ ， $b(q) \cdot \gamma_{3w} / \sigma_w^2$ 共 $(q+1)$ 个独立变量的线性方程组。利用线性最小二乘法求解此方程，可获得 $b(1), b(2), \dots, b(q)$ 的估计。

这是**不考虑观测噪声时的 T 算法**。

考虑对称分布的加性白色观测噪声 $\{v(n)\}$ ，即：

$$y(n)=x(n)+v(n)$$

由于
$$\begin{cases} c_{3y}(\tau_1, \tau_2) = c_{3x}(\tau_1, \tau_2) \\ r_y(\tau) = r_x(\tau) + \sigma_v^2 \cdot \delta(\tau) \end{cases}$$

因此, 只要不取 $m=0$ 对应的广义 GM 方程, 由 $-q \leq m \leq -1$ 和 $1 \leq m \leq q$ 对应的 $2q$ 方程求解, 获得 $b(1), \dots, b(q)$, 观测噪声将不影响估计结果。

对加性白色观测噪声的情况, 还可以联合求解 $2q$ 个不受噪声影响的 GM 方程和 $2q$ 个不受噪声影响的广义 GM 方程, 来估计 $b(1), b(2), \dots, b(q)$ 。

广义 GM 方程可以推广到四阶累积量情况。

3) 累积量算法

无论是 T 算法还是 GM 算法, 由于其中用到了自相关函数, 因此它们不适用于加性有色噪声的情况。张贤达等人提出了一种仅使用累积量的 MA 系统辨识的线性代数方法。由于仅使用了累积量, 这种方法适用于加性有色噪声的情况。

与前面所介绍的累积量闭式递推算法一样, 这种算法也基于如下方程：

$$\gamma_{kw} \sum_{i=0}^q b(i) c_{ky}^{k-1}(q, i - \tau, 0, \dots, 0) = c_{ky}(\tau, 0, \dots, 0) \cdot c_{ky}^{k-1}(q, 0, \dots, 0)$$

取 $\tau = -q, -q+1, \dots, -1, 0, 1, \dots, q$, 由最小二乘法求解所得的联立方程组, 得到 $q+1$ 个未知量 $\gamma_{kw}, \gamma_{kw} b(1), \gamma_{kw} b(2), \dots, \gamma_{kw} b(q)$ 的估计值, 再由

$$b(i) = \frac{\gamma_{kw} b(i)}{\gamma_{kw}}$$

可得 $b(1), \dots, b(q)$ 的估计。

此外，香港理工大学的两位学者还提出了利用不同阶的累积量在有色噪声中盲辨识 MA 系统的两种方法：

Wei Li & Wan-Chi Siu. New Approaches Without Postprocessing to FIR System Identification Using Selected Order Cumulants. IEEE Trans. On SP, 2000, Vol.48, No.4: 1144~1153

[New approaches without postprocessing to FIR system identification using selected order cumulants.pdf](#)

其它类似思想的工作还有：

1. CARRION, M.C., RUIZ, D.P., GALLEG0, A., and MORENTE, J.A., FIR system identification using third- and fourth-order cumulants, Electron. Lett., 1995, 31, (8), pp. 612-614
2. Y. Xiao, M. Shadaydeh and Y. Tadokoro, Overdetermined C(q, k) formula using third and fourth order cumulants. Electron. Lett., 1996, 32, (6), pp. 601-603

3、非线性优化解

由于

$$\begin{aligned}
c_{3y}(\tau_1, \tau_2) &= c_{3x}(\tau_1, \tau_2) \\
&= \gamma_{3w} \sum_{i=0}^q b(i)b(i+\tau_1)b(i+\tau_2) \\
&= \sum_{i=0}^q [\gamma_{3w}^{\frac{1}{3}} b(i)] \cdot [\gamma_{3w}^{\frac{1}{3}} b(i+\tau_1)] [\gamma_{3w}^{\frac{1}{3}} b(i+\tau_2)] \\
&= \sum_{i=0}^q b'(i)b'(i+\tau_1)b'(i+\tau_2)
\end{aligned}$$

从观测样本中估计出 $\hat{c}_{3y}(\tau_1, \tau_2)$ 后，我们可以通过主区 R 内理论累积量与其样本估计值的最优逼近来估计 MA 系统的参数，即定义准则函数：

$$\begin{aligned}
&J[b'(0), b'(1), \dots, b'(q)] \\
&= \sum_R [\sum_{i=0}^q b'(i)b'(i+\tau_1)b'(i+\tau_2) - \hat{c}_{3y}(\tau_1, \tau_2)]^2
\end{aligned}$$

极小化 $J[b'(0), b'(1), \dots, b'(q)]$ 获得最优解 $[b'(0), b'(1), \dots, b'(q)]$ 后，可以求得：

$$b(i) = \frac{b'(i)}{b'(0)}$$

准则函数的极小化可以采用任何非线性优化算法，包括神经网络的方法。

同样，也可以由四阶累积量的非线性最优逼近来估计 MA 系统的参数。

(三)ARMA 系统的辨识

考虑 ARMA 系统

$$H(z) = \frac{\sum_{j=0}^q b(j)z^{-j}}{\sum_{i=0}^p a(i)z^{-i}}$$

假设它是因果的、指数稳定的，且没有零极点对消。 $\{v(n)\}$ 为高斯分布的有色噪声。显然，AR 系统是 ARMA 系统在 $q=0$ 时的特例。

与基于自相关函数的 ARMA 系统辨识方法一样，基于高阶累积量的 ARMA 辨识方法一般也**首先辨识 AR 部分的参数，然后再辨识 MA 部分的参数**。

1、AR 部分的辨识

可以证明，高阶累积量满足以下方程：

$$\sum_{i=0}^p a(i)c_{kx}(\tau_1 - i, \tau_2, \tau_3, \dots, \tau_{k-1}) = 0$$

$$(\tau_1 > q)$$

对 $\tau_3 = \tau_4 = \dots = \tau_{k-1} = 0$ 的特殊情况，有

$$\sum_{i=0}^p a(i)c_{kx}(\tau_1 - i, \tau_2, 0, \dots, 0) = 0$$

$$(\tau_1 > q)$$

这一方程与自相关函数的修正 **Yule-Walker** 方程很相似。

对 $\tau_1 = q+1, q+2, \dots, q+p+M$ ($M \geq 0$) 联立方程，可以获得**基于累积量的法方程**：

$$C(\tau_2) \cdot a = d(\tau_2)$$

其中， $C(\tau_2)$ 为 $(p+M) \times p$ 阶矩阵；

$$a = [a(1), \dots, a(p)]^T$$

显然，如果矩阵 $C(\tau_2)$ 的秩为：

$$\text{rank}\{C(\tau_2)\} = p$$

方程将存在唯一解，它一定对应 AR 参数的真值。但是，如果该矩阵的秩小于 p ，方程将存在无穷多组解，所得解将不对应 AR 参数的真值。

τ_2 取固定值时，如果 $C(\tau_2)$ 的秩等于 p ，则 τ_2 对应的累积量的一维切片称为**满秩切片**。

对 ARMA 系统，有人已经证明：

- 1) 不是任何一个一维切片都是满秩切片，即不是 τ_2 取任何值， $C(\tau_2)$ 的秩都为 p ；
- 2) 满秩的一维累积量切片可能不存在。

但是，对 AR 系统(即 $q=0$ 时)，目前还没有找到反例说明其存在不满秩的一维切片，因此，由一个一维切片是否一定能够估计 AR 系统的参数，仍然是一个没有解决的问题 (open problem)。

既然采用单个累积量切片不能保证可以辨识出 AR 参数，一个自然要问的问题是：如果同时采用 τ_2 取不同值对应的多个累积量切片，是否能够保证 AR 参数的可辨识性？

对这一问题的答案是肯定的。Mendel 和他的学生证明了使用 $\tau_2 = q - p, q - p + 1, \dots, q$ 共 $p+1$ 个累积量切片，可以保证 AR 参数的可辨识性，即方程

$$\begin{bmatrix} C(q-p) \\ C(q-p+1) \\ \vdots \\ C(q) \end{bmatrix} \cdot a = \begin{bmatrix} d(q-p) \\ d(q-p+1) \\ \vdots \\ d(q) \end{bmatrix}$$

一定存在唯一的解。

显然，使用更多的一维累积量切片(必须包括 $\tau_2 = q-p, q-p+1, \dots, q$ 共 $p+1$ 个切片)，也可保证 AR 参数的可辨识性。

实际应用中，我们可以利用这一结果，用 SVD 的方法来确定 p ，并由总体最小二乘技术(TLS)求解方程。

顺便指出，这里有一些的问题可以作进一步的研究：

① 首先，这一结果是对 $\tau_3 = \dots = \tau_{k-1} = 0$ 的特殊情况的。那么对 $\tau_3 = \dots = \tau_{k-1} \neq 0$ 时，是否可以由一个或多个一维切片辨识 AR 参数？更一般地，如何利用 k 取固定值时的方程

$$\sum_{i=0}^p a(i) c_{kx}(\tau_1 - i, \tau_2, \dots, \tau_{k-1}) = 0 \quad (\tau_1 > q)$$

来辨识 AR 参数？

② 组合 k 取不同值时的法方程，是否可以更简单地估计 AR 参数？

需要指出，**以上方法仅适用于因果系统的 AR 参数估计**，这是因为这里的法方程仅对因果系统成立。

除了以上方法外，AR 参数辨识还有**既适用于因果系统又适用于非因果系统**的双 $C(q,n)$ 法，以及穷举搜索、累积量匹配等寻优方法。

2、MA 部分的辨识

1) 残差时间序列法

在 AR 参数辨识出来后,可以利用已知的 AR 多项式,对观测量 $\{y(n)\}$ 进行滤波,构造所谓的“残差序列”:

$$\tilde{y}(n) = \sum_{i=0}^p a(i)y(n-i)$$

不难验证:

$$\tilde{y}(n) = \sum_{j=0}^q b(j)w(n-j) + \sum_{i=0}^p a(i)v(n-i)$$

可见,残差序列 $\{\tilde{y}(n)\}$ 是一个非高斯的 MA(q)过程,它在高斯有色噪声中被观测。因此,MA 参数可以由残差序列 $\{\tilde{y}(n)\}$ 的高阶累积量,采用我们前面介绍的任何一种不用到自相关函数的 MA 系统的辨识方法估计出来。

残差序列的累积量既可以实际产生出残差序列后根据定义计算,也可以不构造出残差序列而由 $\{\tilde{y}(n)\}$ 的累积量和 AR 参数直接计算出来:

$$c_{3\tilde{y}}(\tau_1, \tau_2) = \sum_{i_1=0}^p \sum_{i_2=0}^p \sum_{i_3=0}^p a(i_1)a(i_2)a(i_3)c_{3y}(\tau_1 + i_1 - i_2, \tau_2 + i_1 - i_3)$$

$$c_{4\tilde{y}}(\tau_1, \tau_2, \tau_3) = \sum_{i_1=0}^p \sum_{i_2=0}^p \sum_{i_3=0}^p \sum_{i_4=0}^p a(i_1)a(i_2)a(i_3)a(i_4)c_{4y}(\tau_1 + i_1 - i_2, \tau_2 + i_1 - i_3, \tau_3 + i_1 - i_4)$$

在计算量上,后一种方法将明显减少。

2) q 切片算法 (C(q, n)算法的推广)

对应因果ARMA系统,(Swami and Mendel)可以证明它的单位脉冲响应序列 $\{h(n)\}$ 可以计算为:

$$h(n) = \frac{\sum_{i=0}^p a(i)c_{ky}(q-i, n, 0, \dots, 0)}{\sum_{i=0}^p a(i)c_{ky}(q-i, 0, 0, \dots, 0)} \quad (n = 0, 1, \dots) \quad (1)$$

($p = 0$ 时退化为 $C(q, n)$ 公式)

因此，我们辨识出 AR 参数 $a(1), \dots, a(p)$ 后，就可以根据它计算出单位脉冲响应序列 $\{h(n)\}$ 。

另一方面，根据单位脉冲响应序列的定义，有

$$\sum_{i=0}^p a(i)h(n-i) = \sum_{i=0}^q b(i)\delta(n-i) = b(n) \quad (n = 1, \dots, q)$$

这样就得到了系统 MA 部分的参数 $b(1), b(2), \dots, b(q)$

这种算法中，由于用到了 $n=1, 2, \dots, q$ 对应的 q 个一维切片 $c_{ky}(q-i, n, 0, \dots, 0)$ ，因此而得名 q 切片算法。

注意当 $p = 0$ ，则系统为 q 阶 MA 系统时，上两式退化为 $C(q, n)$ 公式。

正如 $C(q, n)$ 算法没有多大实用意义一样， q 切片算法也不是一种实用的算法。但是，在这个算法的基础上，我们稍加改进，却可以获得一种很不错的 ARMA 参数估计算法。

记

$$f_k(q) = \sum_{i=0}^p a(i)c_{ky}(q-i, 0, \dots, 0)$$

将式 (1) 表示为

$$\sum_{i=0}^p a(i)c_{ky}(q-i, n, 0, \dots, 0) = f_k(q) \cdot h(n)$$

改进算法的基本思想是：取 $n = 0, 1, \dots, Q \geq q$ 对应的 $Q+1$ 方程与 AR 参数估计的法方程一起，联立求解 $a(1), \dots, a(p)$ 以及 $f_k(q) \cdot h(1), \dots, f_k(q) \cdot h(Q)$ 共

$p+Q+1$ 个变量的线性方程组，再由

$$h(n) = \frac{f_k(q) \cdot h(n)}{f_k(q)}$$

获得 $h(n)$ ($n = 1, 2, \dots, q$)，最后再由

$$b(n) = \sum_{i=0}^p a(i)h(n-i)$$

得到 MA 参数的估计。

小结：

在这一节，我们介绍了采用高阶累积量仅从线性系统的输出中辨识系统传递函数的一些典型方法。我们看到采用高阶累积量的方法：

- ① 可以辨识非最小相位系统。
- ② 可以线性地估计 MA 参数。
- ③ 可以有效地抑制观测噪声的影响。

这三个方面都是采用自相关函数的辨识方法难以实现的。

对系统辨识，很重要的一个内容是系统阶次的确定，即定阶问题。对于这个问题，已经提出不少采用高阶累积量的方法。由于时间关系，我们不再介绍。

四、谐波恢复

(一) 基本问题

在许多领域的信号处理中，会遇到这样一类信号，它由多个谐波分量叠加构成：

$$x(n) = \sum_{i=1}^P \alpha_i \exp\{j(\omega_i n + \phi_i)\}$$

关于这个信号的一般假设是：

- ① $\phi_i (i=1,2,\dots,P)$ 为在 $(-\pi,\pi]$ 上均匀分布的相互独立的随机变量。
- ② $\omega_i \neq \omega_j; \quad i \neq j$
- ③ α_i, ω_i 为未知的确定量。

对信号的观测受到噪声的污染：

$$y(n) = x(n) + v(n)$$

所要解决的基本问题是：根据一段观测数据 $\{y(n): n=1,2,\dots,N\}$ 估计信号的参数 $\{(\alpha_i, \omega_i, \phi_i): i=1,2,\dots,P\}$

这个问题的核心是**信号角频率 $\omega_i (i=1,2,\dots,P)$ 的估计**，特别是频率间隔很小的情况下，如何获得良好的频率估计。因此，这个问题有时也叫**高分辨频率估计问题**。

对这个问题的研究已经有近四十年的历史，特别是从八十年代初以来，提出了许多出色的方法，如 TLS-LP 法、MUSIC 法、ESPRIT 法等。但是，这些方法的基本假设是：观测噪声 $\{v(n)\}$ 为**白噪声或二阶统计特性已知的有色噪声**。未知统计特性有色观测噪声情况下，如何解决高分辨

频率估计问题，长期没有很好的办法。

引入高阶统计量作为数学工具，不仅可以解决高斯有色观测噪声情况下的高分辨频率估计问题，最近几年我国学者还解决了非高斯有色观测噪声情况下的谐波恢复问题。

(二) 谐波恢复的高阶累积量方法

1、谐波信号的累积量

对前面介绍的 P 个分量的复谐波信号 $\{x(n)\}$ ，可以证明

$$c_{3x}(\tau_1, \tau_2) \equiv 0 \quad (\text{三阶累积量自动抑制谐波信号})$$

$$\begin{aligned} c_{4x}(\tau_1, \tau_2, \tau_3) &= cum(x^*(n), x^*(n + \tau_1), x(n + \tau_2), x(n + \tau_3)) \\ &= -\sum_{i=1}^P \alpha_i^4 \exp\{j\omega_i(-\tau_1 + \tau_2 + \tau_3)\} \end{aligned}$$

而且，对于复谐波信号，除了这种形式的四阶累积量定义，其它形式的定义将得到恒等于零的四阶累积量，即自动抑制谐波信号：

$$\begin{aligned} & cum(x(n), x(n + \tau_1), x(n + \tau_2), x(n + \tau_3)) \\ &= cum(x^*(n), x(n + \tau_1), x(n + \tau_2), x(n + \tau_3)) \\ &= cum(x^*(n), x^*(n + \tau_1), x^*(n + \tau_2), x(n + \tau_3)) \\ &= cum(x^*(n), x^*(n + \tau_1), x^*(n + \tau_2), x^*(n + \tau_3)) \\ &\equiv 0 \end{aligned}$$

由此定义，复谐波信号的四阶累积量的对角切片为：

$$c_{4x}(\tau, \tau, \tau) = -\sum_{i=1}^P \alpha_i^4 \exp\{j\omega_i \tau\}$$

它由 P 个谐波分量叠加构成，且频率对应于信号中谐波分量的频率。

2、高斯分布有色噪声中的谐波恢复方法

如果观测噪声 $\{v(n)\}$ 服从高斯分布，由于

$$\begin{aligned} c_{4y}(\tau, \tau, \tau) &= c_{4x}(\tau, \tau, \tau) \\ &= -\sum_{i=1}^P \alpha_i^4 \exp\{j\omega_i \tau\} \end{aligned}$$

因此，可以从观测量的四阶累积量对角切片中直接采用 TLS-LP、MUSIC、ESPRIT 等法，估计频率 $\omega_i (i=1, 2, \dots, P)$ 。

3、非对称分布线性非高斯噪声中的谐波恢复方法

（三阶累积量+自相关函数，**利用三阶累积量自动抑制谐波信号这一特性**）

如果观测噪声 $\{v(n)\}$ 是非对称分布的线性 ARMA 非高斯过程，其模型传递函数为

$$H(z) = \frac{\sum_{j=0}^q b(j)z^{-j}}{\sum_{i=0}^p a(i)z^{-i}} = \frac{B(z)}{A(z)}$$

由于

$$\begin{aligned} c_{3y}(\tau_1, \tau_2) &= c_{3x}(\tau_1, \tau_2) + c_{3v}(\tau_1, \tau_2) \\ &= c_{3v}(\tau_1, \tau_2) \end{aligned}$$

因此可以由观测量的三阶累积量中采用前面介绍的系统辨识方法，估计出噪声模型的 AR 部分。

用 $A(z)$ 对观测量 $\{y(n)\}$ 进行滤波处理，得到

$$\begin{aligned}\tilde{y}(n) &= A(z) \cdot y(n) = A(z)x(n) + A(z)v(n) \\ &= \tilde{x}(n) + e(n) \\ &\quad \downarrow \quad \downarrow \\ &\text{滤波后的} \quad \text{滤波后的} \\ &\text{谐波信号} \quad \text{观测噪声} \quad (MA(q)\text{过程})\end{aligned}$$

因此

$$r_{\tilde{y}}(\tau) = r_{\tilde{x}}(\tau) + r_e(\tau)$$

当 $\tau > q$ 时

$$r_{\tilde{y}}(\tau) = r_{\tilde{x}}(\tau)$$

从 $r_{\tilde{y}}(\tau)$ ($\tau > q$) 中，可以直接采用 TLS-LP、MUSIC、ESPRIT 等方法估计频率。

4、对称分布线性非高斯噪声中的谐波恢复方法

(能够抑制谐波信号的四阶累积量+自相关函数)

如果观测噪声是对称分布的线性 ARMA 非高斯过程，可以采用能够抑制谐波信号的四阶累积量，辨识噪声模型的 AR 部分，再用它对观测量进行滤波处理，最后从滤波结果的自相关函数中直接采用 TLS-LP、MUSIC、ESPRIT 等方法估计频率。

小结：

对谐波恢复问题，既可以利用高阶累积量对高斯噪声的抑制能力，又可以利用高阶累积量对谐波信号的抑制作用，灵活地解决有色噪声背景下

的信号参数估计问题。

五、空间窄带信号源的波达方向估计

(一) 基本问题

设空间有 P 个窄带信号 $\{s_1(t), s_2(t), \dots, s_P(t)\}$ 照射到一个有 M 个传感器的阵列上，第 p 个信号的波达方向(DOA)参数为 θ_p ，则阵列的输出为：

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t)$$

其中， $\mathbf{r}(t) = [r_1(t), r_2(t), \dots, r_M(t)]^T$ 为阵列输出向量；

$\mathbf{A} = [\mathbf{a}(\theta_1), \mathbf{a}(\theta_2), \dots, \mathbf{a}(\theta_P)]$ 为方向矩阵；

$\mathbf{a}(\theta)$ 为该阵列对波达方向参数为 θ 的单位信号的响应向量（称为方向向量或导向矢量），其在所有感兴趣的波达方向参数集 Θ 上的集合 $\mathcal{A} = \{\mathbf{a}(\theta) | \theta \in \Theta\}$ 通常称为该阵列的阵列流形（Array Manifold）；

$\mathbf{s}(t) = [s_1(t), s_2(t), \dots, s_P(t)]^T$ 为信号向量；

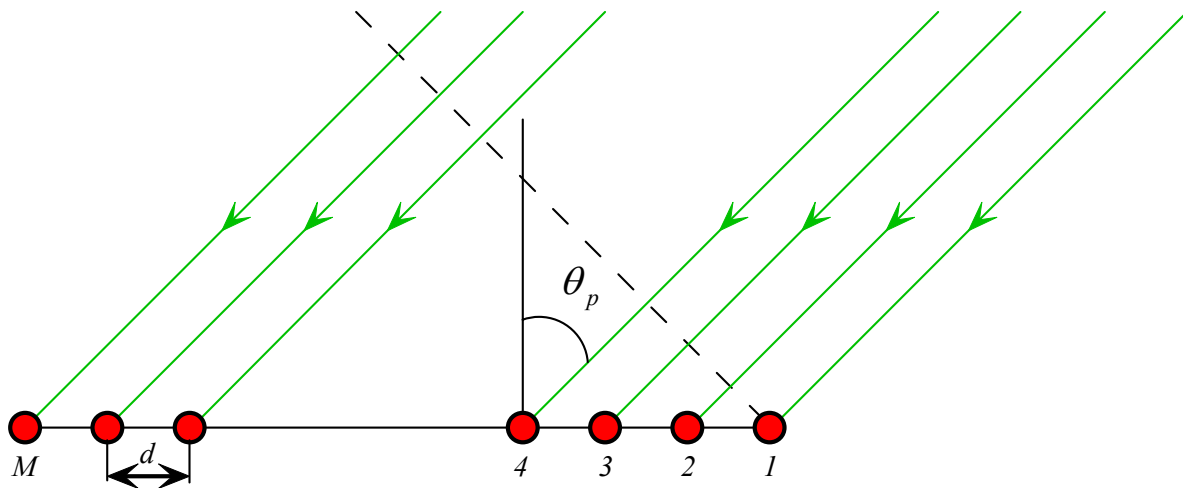
$\mathbf{n}(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T$ 为加性噪声向量。

通常，假设阵列流形是无模糊的，即对任意 $P \leq M$ 个不同的 θ_p ， $\mathbf{a}(\theta_p)$ （ $p = 1, 2, \dots, P$ ）是一组线性无关的向量。等效地有：

$$\text{rank}\{\mathbf{A}\} = P$$

在雷达、无线通讯、声纳、射电天文学等领域中一个重要的问题，就是根据阵列在有限个时刻 $t = 1, 2, \dots, N$ 的输出（快拍，snapshot） $\{\mathbf{r}(t)$ ；

$t = 1, 2, \dots, N$ }, 估计各空间信号源的到达方向参数 $\{\theta_p; p = 1, 2, \dots, P\}$ 。



对于最简单的**均匀线阵**(ULA), 若阵元间距为 d , 各传感器响应特性相同, 且 $s_1(t), s_2(t), \dots, s_P(t)$ 表示第一个传感器接收到的各信号源信号分量, 信号源载波波长为 λ , 则其方向向量为:

$$\mathbf{a}(\theta) = \left[1, e^{-j\frac{2\pi}{\lambda}d \sin \theta}, \dots, e^{-j\frac{2\pi}{\lambda}(M-1)d \sin \theta} \right]^T$$

(二) 基于二阶统计量的 DOA 估计方法及其不足

1. 基本假设

- (1) 各传感器输出中的噪声 $n_m(t)$ 为零均值复平稳随机过程, **具有相同的噪声功率 σ^2** , **且相互之间统计独立或不相关**, 即 $\mathbf{n}(t)$ 是**空间白的**, 其协方差矩阵为:

$$E\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \sigma^2 I$$

- (2) 各空间信号 $s_p(t)$ 为零均值复平稳随机过程，**相互之间不相干**，即信号向量的协方差矩阵 $E\{s(t)s^H(t)\} = \Sigma_{ss}$ 为非奇异矩阵。
- (3) 信号与噪声**统计独立**。

2. 基本的MUSIC (Multiple Signal Classification)法

在以上假设条件下，阵列输出的协方差矩阵为：

$$\mathbf{R} = E\{\mathbf{r}(t)\mathbf{r}^H(t)\} = \mathbf{A}\Sigma_{ss}\mathbf{A}^H + \sigma^2\mathbf{I}$$

对其作特征分解，可得：

$$\mathbf{R} = \mathbf{U} \cdot \mathbf{\Lambda} \cdot \mathbf{U}^H = \sum_{i=1}^M \lambda_i \mathbf{u}_i \mathbf{u}_i^H$$

其中， $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M]$ 为由特征向量组成的酉矩阵；

$\mathbf{\Lambda} = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_M\}$ 为由特征值构成的对角阵。

由于 \mathbf{R} 为满秩 Hermitian 矩阵，其特征值为正值。若将 \mathbf{R} 的特征值按大小依次排列，不难证明有：

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_P > \lambda_{P+1} = \lambda_{P+2} = \dots = \lambda_M = \sigma^2$$

且有：

$$\text{span}\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_P\} = \text{range}\{\mathbf{A}\} \quad (\text{信号子空间})$$

$$\text{span}\{\mathbf{u}_{P+1}, \mathbf{u}_{P+2}, \dots, \mathbf{u}_M\} \perp \text{range}\{\mathbf{A}\} \quad (\text{噪声子空间})$$

记

$$\mathbf{U}_s = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_P]$$

$$\mathbf{U}_n = [\mathbf{u}_{P+1}, \mathbf{u}_{P+2}, \dots, \mathbf{u}_M]$$

则有

$$\mathbf{U}_n^H \cdot \mathbf{A} = 0$$

即对 $\mathbf{a}(\theta)$ ，有

$$\mathbf{U}_n^H \mathbf{a}(\theta) = 0, \quad \theta \in \{\theta_1, \theta_2, \dots, \theta_p\}$$

定义 MUSIC 空间谱为：

$$\begin{aligned} P_{MUSIC}(\theta) &= \frac{\mathbf{a}^H(\theta) \mathbf{a}(\theta)}{\mathbf{a}^H(\theta) \mathbf{U}_n \mathbf{U}_n^H \mathbf{a}(\theta)} \\ &= \frac{\mathbf{a}^H(\theta) \mathbf{a}(\theta)}{\mathbf{a}^H(\theta) (\mathbf{I} - \mathbf{U}_S \mathbf{U}_S^H) \mathbf{a}(\theta)} \end{aligned}$$

由其 P 个谱峰对应的 θ 得到信号源的 DOA。

显然，对 MUSIC 法，需要阵列流形 $\{\mathbf{a}(\theta) | \theta \in \Theta\}$ 是完全已知的。

3. 基本的 ESPRIT 法

不失一般性，考虑传感器阵列为一维线阵的情况。

假设传感器阵列可分成部分重叠或完全不重叠的两个 L 个阵元的子阵，子阵 1 的输出为：

$$\mathbf{x}(t) = \mathbf{A}_1 \mathbf{s}(t) + \mathbf{n}_1(t)$$

子阵 2 的输出为：

$$\mathbf{y}(t) = \mathbf{A}_2 \mathbf{s}(t) + \mathbf{n}_2(t)$$

子阵 1 中的第 l 个传感器和子阵 2 中的第 l 个传感器具有完全相同的响应特性，只是空间上平移了一个已知的固定位移 Δ 。若 $s_1(t), s_2(t), \dots, s_p(t)$ 表示子阵 1 中第一个传感器接收到的各信号源信号分量，则显然有：

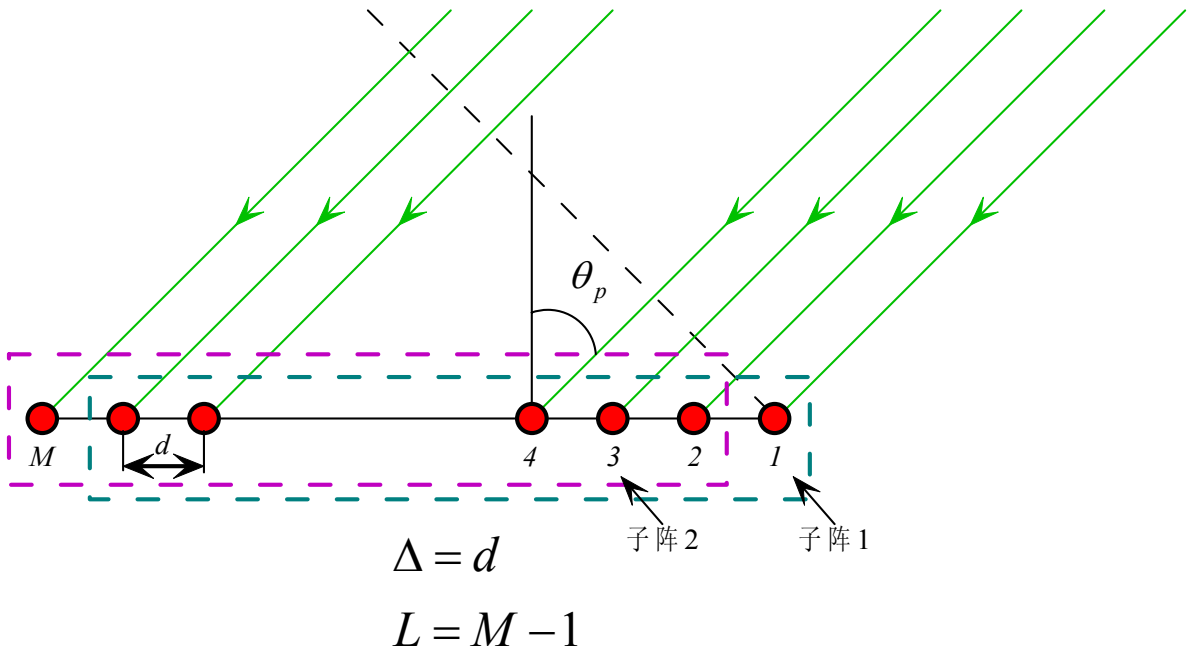
$$\mathbf{A}_2 = \mathbf{A}_1 \cdot \text{diag} \left\{ e^{-j\frac{2\pi}{\lambda} \Delta \sin \theta_1}, e^{-j\frac{2\pi}{\lambda} \Delta \sin \theta_2}, \dots, e^{-j\frac{2\pi}{\lambda} \Delta \sin \theta_P} \right\}$$

$$= \mathbf{A}_1 \cdot \Phi$$

其中，

$$\Phi = \text{diag} \left\{ e^{-j\frac{2\pi}{\lambda} \Delta \sin \theta_1}, e^{-j\frac{2\pi}{\lambda} \Delta \sin \theta_2}, \dots, e^{-j\frac{2\pi}{\lambda} \Delta \sin \theta_P} \right\}$$

如对 ULA，一种子阵的划分情况如下图。



$$\mathbf{R}_{xx} = E\{\mathbf{x}(t)\mathbf{x}^H(t)\} = \mathbf{A}_1 \mathbf{\Sigma}_{SS} \mathbf{A}_1^H + \sigma^2 \mathbf{I}$$

$$\mathbf{R}_{xy} = E\{\mathbf{x}(t)\mathbf{y}^H(t)\} = \mathbf{A}_1 \mathbf{\Sigma}_{SS} \Phi^H \mathbf{A}_1^H + \sigma^2 \mathbf{Z}$$

其中，矩阵 \mathbf{Z} 与两个子阵阵元重叠情况有关，为已知的矩阵。若两个子阵没有阵元重叠，则 $\mathbf{Z}=\mathbf{0}$ ；否则，对应阵元重叠的位置， \mathbf{Z} 的元素为 1。

如对上图的子阵划分，有：

$$\mathbf{Z} = \begin{bmatrix} 0 & & & 0 \\ 1 & 0 & & \\ & \ddots & \ddots & \\ 0 & & 1 & 0 \end{bmatrix}$$

对 \mathbf{R}_{xx} 进行特征分解，由其 $L-P$ 个最小特征值可获得 σ^2 ；再计算：

$$\mathbf{C}_{xx} = \mathbf{R}_{xx} - \sigma^2 \mathbf{I} = \mathbf{A}_1 \Sigma_{SS} \mathbf{A}_1^H$$

$$\mathbf{C}_{xy} = \mathbf{R}_{xy} - \sigma^2 \mathbf{Z} = \mathbf{A}_1 \Sigma_{SS} \Phi^H \mathbf{A}_1^H$$

最后由矩阵束 $\{\mathbf{C}_{xx}, \mathbf{C}_{xy}\}$ 的 P 个广义特征值可得到

$e^{-j\frac{2\pi}{\lambda}\Delta \sin \theta_1}, e^{-j\frac{2\pi}{\lambda}\Delta \sin \theta_2}, \dots, e^{-j\frac{2\pi}{\lambda}\Delta \sin \theta_P}$ ，从而得到信号源的 DOA $\{\theta_p; p=1,2,\dots,P\}$ 。

显然，对 ESPRIT 法，**并不需要两个子阵的流形完全已知，只需保证成对的传感器具有相同的响应特性。**

4. 存在的不足

(1) 要求传感器噪声是空间白噪声且噪声功率相同，或者是协方差矩阵已知的空间色噪声(空间相关噪声)，否则，无法通过阵列协方差矩阵的特征分解分割出信号子空间和噪声子空间，即使**传感器间噪声统计独立但噪声功率不同也不行。**

(2) 要求空间信号源不相干，否则协方差矩阵 $E\{s(t)s^H(t)\} = \Sigma_{ss}$ 为奇异矩阵；虽然提出了一些处理相干信号的措施，如空间平滑、前后向平均等，但要求阵列具有特殊的结构**如均匀线阵**或者会造成**有效孔径减小**。在实际应用中，**由于多路径现象或转发干扰**，阵列接收到的信号之间高度相关和

相干的情况是经常存在的。

(3) MUSIC 法要求通过阵列校准, 获得完全的阵列流形信息, 且对校准误差敏感; ESPRIT 法虽不需完全的阵列流形信息, 但也需成对地校准传感器, 使其具有相同的响应特性。

基于高阶累计量的 DOA 估计方法则可以克服以上不足。

参考文献

1. M. C. Dogan, J. M. Mendel. Applications of Cumulants to Array Processing—Part I: Aperture Extension and Array Calibration. IEEE Trans. on SP, Vol.43, No.5, pp.1200~1216, May 1995
2. M. C. Dogan, J. M. Mendel. Applications of Cumulants to Array Processing—Part II: Non-Gaussian Noise Suppression. IEEE Trans. on SP, Vol.43, No.7, pp.1663~1676, July 1995
3. E. Gonen, J. M. Mendel. Applications of Cumulants to Array Processing—Part III: Blind Beamforming for Coherent Signals. IEEE Trans. on SP, Vol.45, No.9, pp.2252~2264, Sep. 1997
4. E. Gonen, J. M. Mendel, M. C. Dogan. Applications of Cumulants to Array Processing—Part IV: Direction Finding in Coherent Signals Case. IEEE Trans. on SP, Vol.45, No.9, pp.2265~2276, Sep. 1997
5. B. Porat and B. Friedlander. Direction Finding Algorithms Based on Higher Order Statistics. IEEE Trans. SP, Vol. 39, No.9, pp.2016~2023, Sep. 1991
6. J.-F. Cardoso and E. Moulines. Asymptotic Performance Analysis of Direction-Finding Algorithms Based on Fourth-Order Cumulants. IEEE Trans. SP, Vol. 43, No.1, pp.214~224, Jan. 1991,
7. R. Roy and T. Kailath. ESPRIT-Estimation of Signal Parameters Via Rotational Invariance Techniques. IEEE Trans. ASSP, Vol. 37, No.7, pp.984~995, July 1989

8. A. Swami and J. M. Mendel. Cumulant-Based Approach to the Harmonic Retrieval and Related Problems. IEEE Trans. SP, Vol. 39, No.5, pp.1099–1109, May 1991

(三) 基于高阶统计量的 DOA 估计方法

1. 基于高阶累计量的 MUSIC 法

利用如下性质，抑制高斯分布的空间相关噪声：

如果 $\{z_i\}_{i=1}^n$ 是与 $\{x_i\}_{i=1}^n$ 统计独立的高斯随机变量且 $n > 2$ ，则

$$\text{cum}(x_1 + z_1, x_2 + z_2, \dots, x_n + z_n) = \text{cum}(x_1, x_2, \dots, x_n)$$

设空间有 P 个窄带信号 $\{s_1(t), s_2(t), \dots, s_P(t)\}$ 照射到一个有 M 个传感器的阵列上，第 p 信号的波达方向(DOA)参数为 θ_p ，则阵列的输出为：

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \sum_{p=1}^P \mathbf{a}(\theta_p) s_p(t) + \mathbf{n}(t)$$

假设：

- (1) $\mathbf{n}(t)$ 是零均值的复高斯噪声，具有任意的协方差矩阵 $E\{\mathbf{n}(t)\mathbf{n}^H(t)\}$ 。
- (2) 各空间信号 $s_p(t)$ 为零均值非高斯分布复平稳随机过程，相互之间统计独立。
- (3) 信号与噪声统计独立。

定义阵列输出四阶累计量为：

$$q_{ik}^{jl} = \text{cum}(r_i(t), r_j^*(t), r_k(t), r_l^*(t)) \quad 1 \leq i, j, k, l \leq M$$

其中， $r_i(t)$ 为第 i 个传感器的输出，即 $\mathbf{r}(t)$ 的第 i 个元素，有：

$$r_i(t) = \sum_{p=1}^P a_i(\theta_p) s_i(t) + n_i(t) = x_i(t) + n_i(t)$$

根据累积量的性质 CP4，可得：

$$q_{ik}^{jl} = \text{cum}(x_i(t), x_j^*(t), x_k(t), x_l^*(t)) + \text{cum}(n_i(t), n_j^*(t), n_k(t), n_l^*(t))$$

由于假设噪声具有高斯分布，故有：

$$q_{ik}^{jl} = \text{cum}(x_i(t), x_j^*(t), x_k(t), x_l^*(t))$$

即噪声不影响信号的四阶累积量。

根据定义，阵列输出四阶累计量总共有 M^4 个元素。

方法1：四阶累积量对角切片矩阵（Pan and Nikias）

R. Pan and C. L. Nikias, “Harmonic decomposition methods in cumulant domains,” in *Proc. ICASSP '88*, New York, NY, pp. 2356-2359, Apr. 1988.

取阵列输出四阶累计量的如下子集：

$$q_{ii}^{jj} = \text{cum}(r_i(t), r_i^*(t), r_j(t), r_j^*(t)) \quad 1 \leq i, j \leq M$$

构成 $M \times M$ 矩阵：

$$\mathbf{D} = \begin{bmatrix} q_{11}^{11} & q_{11}^{12} & \cdots & q_{11}^{1M} \\ q_{22}^{21} & q_{22}^{22} & \cdots & q_{22}^{2M} \\ \vdots & \vdots & \ddots & \vdots \\ q_{MM}^{M1} & q_{MM}^{M2} & \cdots & q_{MM}^{MM} \end{bmatrix}$$

可以证明：

$$\mathbf{D} = \mathbf{A} \cdot \mathbf{\Gamma} \cdot \mathbf{A}^H$$

其中， $\mathbf{\Gamma} = \text{diag}\{\gamma_1, \gamma_2, \cdots, \gamma_P\}$,

$$\gamma_p = \text{cum}(s_p(t), s_p^*(t), s_p(t), s_p^*(t)) \neq 0 \quad (p=1,2,\dots,P)$$

故可对矩阵 \mathbf{D} 作特征分解，由其 P 个非零的特征值对应的特征向量得到信号子空间的一组基，由 $M-P$ 个零特征值对应的特征向量得到噪声子空间的一组基，从而可由 MUSIC 空间谱的谱峰得到 DOA 估计。

方法2：压缩四阶协方差矩阵 (J.-F. Cardoso)

J.F.Cardoso, “Higher-Order Narrow-Band Array Processing”, Proc. Int. Sig. Proc. Work. On HOS, pp121-130, 1991

J.F. Cardoso and E. Moulines, “Asymptotic Performance Analysis of Direction-Finding Algorithms Based on Fourth-Order Cumulants”, IEEE Trans. on SP, Vol. 43, No. 1, Jan. 1995

取：

$$\begin{aligned} c(i, j) &= \sum_{k=1}^M \text{cum}(r_i(t), r_j^*(t), r_k(t), r_k^*(t)) \\ &= \sum_{k=1}^M \text{cum}(x_i(t), x_j^*(t), x_k(t), x_k^*(t)) \\ &= \sum_{k=1}^M q_{ik}^{jk} \end{aligned}$$

构成 $M \times M$ 矩阵

$$\mathbf{C} = \begin{bmatrix} c(1,1) & c(1,2) & \cdots & c(1,M) \\ c(2,1) & c(2,2) & \cdots & c(2,M) \\ \vdots & \vdots & \ddots & \vdots \\ c(M,1) & c(M,2) & \cdots & c(M,M) \end{bmatrix}$$

可以证明：

$$\mathbf{C} = \mathbf{A} \cdot \mathbf{K}_S \cdot \mathbf{A}^H$$

其中， $\mathbf{K}_S = \text{diag}\{\gamma_1 |\mathbf{a}(\theta_1)|^2, \gamma_2 |\mathbf{a}(\theta_2)|^2, \dots, \gamma_P |\mathbf{a}(\theta_P)|^2\}$ 。故可对矩阵 \mathbf{C}

作特征分解，由其 P 个非零的特征值对应的特征向量得到信号子空间的一组基，由 $M-P$ 个零特征值对应的特征向量得到噪声子空间的一组基，从而可由 MUSIC 空间谱的谱峰得到 DOA 估计。

除以上两种 MUSIC 法之外，还有其它利用阵列输出四阶累计量全部 M^4 个元素的一些 MUSIC 法，包括可用于相关信号源的方法。如：

B. Porat and B. Friedlander, “Direction finding algorithms based on high-order statistics,” *IEEE Tran. Signal Processing*, vol. 39, pp.2016-2025, Jan. 1991.

2. Virtual-ESPRIT 算法(VESPA)——高斯噪声情况(M. C. Dogan, J. M. Mendel)

不失一般性，考虑传感器阵列为一维线阵的情况。

设空间有 P 个窄带信号 $\{s_1(t), s_2(t), \dots, s_P(t)\}$ 照射到一个有 M 个传感器的阵列上，第 p 信号的波达方向(DOA)参数为 θ_p ，则阵列的输出为：

$$\mathbf{r}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \sum_{p=1}^P \mathbf{a}(\theta_p) s_p(t) + \mathbf{n}(t)$$

假设：

- (1) $\mathbf{n}(t)$ 是零均值的复高斯噪声，具有任意的协方差矩阵 $E\{\mathbf{n}(t)\mathbf{n}^H(t)\}$ 。
- (2) 各空间信号 $s_p(t)$ 为零均值非高斯分布复平稳随机过程，相互之间统计独立。
- (3) 信号与噪声统计独立。
- (4) 阵列中有两个传感器具有完全相同的响应特性（称这两个传感器为传感器对），相距已知的间隔 Δ ；除此之外，**阵列中所有其它传感器的位**

置和响应特性均未知。

记传感器对接收到的信号为 $r_1(t)$ 和 $r_2(t)$ ，构造如下两个 $M \times M$ 的四阶累积量矩阵：

$$\mathbf{C}_1 = cum(r_1(t), r_1^*(t), \mathbf{r}(t), \mathbf{r}^H(t))$$

$$\mathbf{C}_2 = cum(r_2(t), r_1^*(t), \mathbf{r}(t), \mathbf{r}^H(t))$$

不难证明，这两个矩阵的第 (i, j) 个元素分别为：

$$\mathbf{C}_1(i, j) = cum(r_1(t), r_1^*(t), r_i(t), r_j^*(t))$$

$$= \sum_{p=1}^P \gamma_p |a_1(\theta_p)|^2 a_i(\theta_p) a_j^*(\theta_p)$$

$$\mathbf{C}_2(i, j) = cum(r_2(t), r_1^*(t), r_i(t), r_j^*(t))$$

$$= \sum_{p=1}^P \gamma_p |a_1(\theta_p)|^2 e^{-j \frac{2\pi}{\lambda} \Delta \sin \theta_p} a_i(\theta_p) a_j^*(\theta_p)$$

其中， $a_i(\theta_p)$ 表示 $\mathbf{a}(\theta_p)$ 的第 i 个元素；

$$\gamma_p = cum(s_p(t), s_p^*(t), s_p(t), s_p^*(t)) \neq 0 \quad (p = 1, 2, \dots, P)$$

因此，这两个矩阵可分别表示为：

$$\mathbf{C}_1 = \mathbf{A} \cdot \mathbf{D} \cdot \mathbf{A}^H$$

$$\mathbf{C}_2 = \mathbf{A} \cdot \Phi \cdot \mathbf{D} \cdot \mathbf{A}^H$$

其中，

$$\mathbf{D} = diag\left\{\gamma_1 |a_1(\theta_1)|^2, \gamma_2 |a_1(\theta_2)|^2, \dots, \gamma_P |a_1(\theta_P)|^2\right\}$$

$$\Phi = diag\left\{e^{-j \frac{2\pi}{\lambda} \Delta \sin \theta_1}, e^{-j \frac{2\pi}{\lambda} \Delta \sin \theta_2}, \dots, e^{-j \frac{2\pi}{\lambda} \Delta \sin \theta_P}\right\}$$

显然，若阵列流形是无模糊的，即 $\text{rank}(\mathbf{A}) = P$ ，当各信号源均有非零的四阶累积量且第一个传感器对入射信号均有非零响应时，有

$$\text{rank}(\mathbf{D}) = \text{rank}(\mathbf{C}_1) = \text{rank}(\mathbf{C}_2) = P$$

故通过求矩阵束 $\{\mathbf{C}_1, \mathbf{C}_2\}$ 的 P 个广义特征值就可得到

$e^{-j\frac{2\pi}{\lambda}\Delta\sin\theta_1}, e^{-j\frac{2\pi}{\lambda}\Delta\sin\theta_2}, \dots, e^{-j\frac{2\pi}{\lambda}\Delta\sin\theta_P}$ ，从而得到信号源的 DOA $\{\theta_p; p=1,2,\dots,P\}$ 。

显然，相对基于二阶统计量的 ESPRIT 法，Virtual-ESPRIT 算法具有如下优点：

- (1) 能自动抑制空间有色高斯噪声的影响；
- (2) 仅需保证一对传感器具有相同的响应特性。

3. Virtual-ESPRIT 算法(VESPA)——非高斯噪声情况(M. C. Dogan, J. M. Mendel)

假设：

- (1) $\mathbf{n}(t)$ 是零均值的非复高斯噪声，具有任意的未知的统计特征。
- (2) 各空间信号 $s_p(t)$ 为零均值非高斯分布复平稳随机过程，相互之间统计独立。
- (3) 信号与噪声统计独立。
- (4) 阵列中有一传感器对，相距已知的间隔 Δ ；除此之外，阵列中所有其它传感器的位置和响应特性均未知。
- (5) 传感器对的两个传感器中，至少有一个传感器的噪声分量与其它传感

器的噪声分量统计独立。

不失一般性，设传感器对中的第一个传感器输出 $r_1(t)$ 中的噪声分量与其它传感器的噪声分量统计独立。

根据高阶累积量的性质 CP4、CP5，

$$\begin{aligned}\mathbf{C}_1(i, j) &= cum(r_1(t), r_1^*(t), r_i(t), r_j^*(t)) \\ &= cum(x_1(t) + n_1(t), x_1^*(t) + n_1^*(t), x_i(t) + n_i(t), x_j^*(t) + n_j^*(t)) \\ &= cum(x_1(t), x_1^*(t), x_i(t), x_j^*(t)) + cum(n_1(t), n_1^*(t), n_i(t), n_j^*(t)) \\ &= cum(x_1(t), x_1^*(t), x_i(t), x_j^*(t)) + \gamma_{e1} \delta(i-1, j-1)\end{aligned}$$

$$\begin{aligned}\mathbf{C}_2(i, j) &= cum(r_2(t), r_1^*(t), r_i(t), r_j^*(t)) \\ &= cum(x_2(t) + n_2(t), x_1^*(t) + n_1^*(t), x_i(t) + n_i(t), x_j^*(t) + n_j^*(t)) \\ &= cum(x_2(t), x_1^*(t), x_i(t), x_j^*(t)) + cum(n_2(t), n_1^*(t), n_i(t), n_j^*(t)) \\ &= cum(x_2(t), x_1^*(t), x_i(t), x_j^*(t))\end{aligned}$$

其中， $\gamma_{e1} = cum(n_1(t), n_1^*(t), n_1(t), n_1^*(t))$ 。

因此，有：

$$\mathbf{C}_1 = \mathbf{A} \cdot \mathbf{D} \cdot \mathbf{A}^H + \gamma_{e1} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

$$\mathbf{C}_2 = \mathbf{A} \cdot \Phi \cdot \mathbf{D} \cdot \mathbf{A}^H$$

由这两个矩阵，不难得到信号源的 DOA。

4. Extended VESPA 算法(EVESPA)——相干信号情况(E. Gonen, J. M. Mendel, M. C. Dogan)

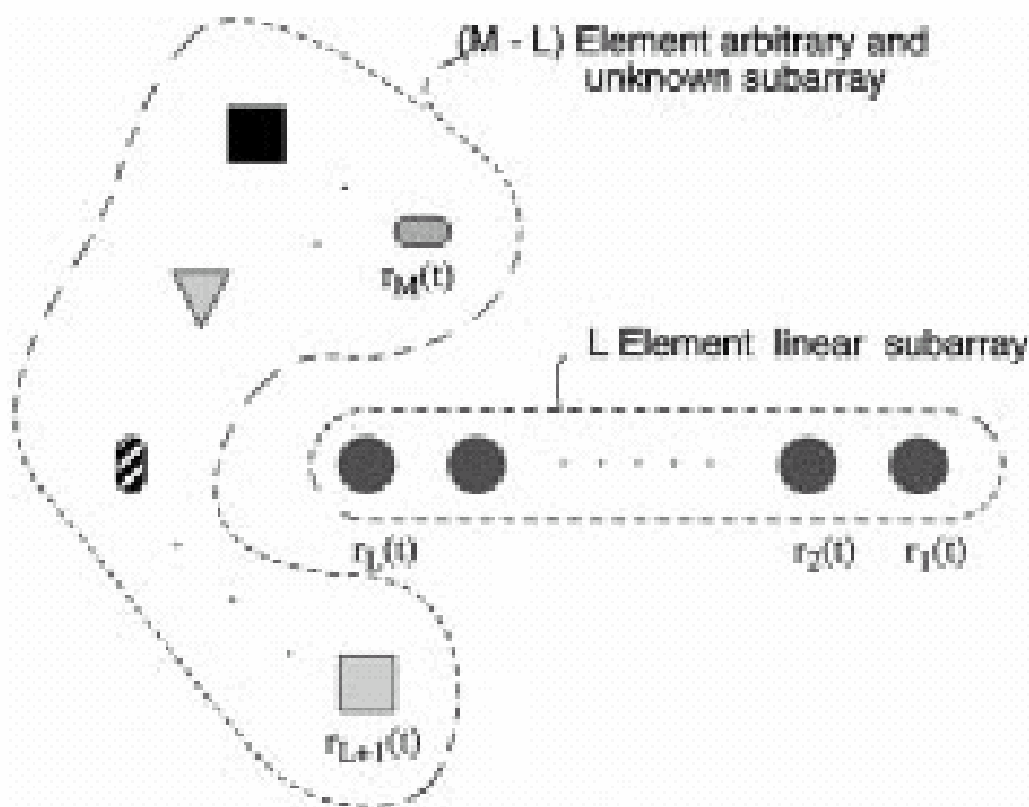
假设：

(1) 空间有 G 个相互之间统计独立的信号源 $u_1(t), u_2(t), \dots, u_G(t)$ ，它们具有非零的四阶累积量。

(2) 信号源 $u_i(t)$ 通过 p_i 条不同的路径从方向 $\{\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,p_i}\}$ 入射到阵列上，对应的入射信号 $\{s_{i,1}(t), s_{i,2}(t), \dots, s_{i,p_i}(t)\}$ 与 $u_i(t)$ 波形相同，只有幅度和延迟（相位）的差别，即与 $u_i(t)$ 相干。入射到阵列上的信号总数

$$\text{为 } \sum_{i=1}^G p_i = P$$

(3) 阵列共有 M 个传感器，其中的 L 个传感器（不失一般性，设为第 1~ L 个传感器）构成一维均匀线阵(ULA)，其它 $M-L$ 个传感器的位置和响应特性均未知。



(4) 阵列流形是无模糊的。

(5) $G < M$ 。

(6) $\mathbf{n}(t)$ 是零均值的高斯分布、对称分布或它们的混合的复噪声，具有任意的未知的统计特征。

(7) 信号与噪声统计独立。

则阵列的输出可表示为：

$$\begin{aligned}\mathbf{r}(t) &= \mathbf{A}\mathbf{s}(t) + \mathbf{n}(t) = \mathbf{A}\mathbf{Q}\mathbf{u}(t) + \mathbf{n}(t) \\ &= \mathbf{B}\mathbf{u}(t) + \mathbf{n}(t)\end{aligned}$$

其中，

$$\mathbf{s}(t) = \begin{bmatrix} \mathbf{s}_1(t) \\ \mathbf{s}_2(t) \\ \vdots \\ \mathbf{s}_G(t) \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{c}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{c}_G \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_G(t) \end{bmatrix} \triangleq \mathbf{Q}\mathbf{u}(t)$$

$$\mathbf{s}_i(t) = \begin{bmatrix} s_{i,1}(t) \\ s_{i,2}(t) \\ \vdots \\ s_{i,p_i}(t) \end{bmatrix}, \quad \mathbf{c}_i = \begin{bmatrix} c_{i,1} \\ c_{i,2} \\ \vdots \\ c_{i,p_i} \end{bmatrix}$$

$$\mathbf{B} \triangleq \mathbf{A}\mathbf{Q}$$

B的列称为广义导向矢量。

构造如下两个 $M \times M$ 的四阶累积量矩阵：

$$\mathbf{C}_1 = cum(r_1(t), r_1^*(t), \mathbf{r}(t), \mathbf{r}^H(t))$$

$$\mathbf{C}_2 = cum(r_2(t), r_1^*(t), \mathbf{r}(t), \mathbf{r}^H(t))$$

根据CP1、CP3、CP5和CP6，不难证明，

$$\mathbf{C}_1 = \mathbf{B} \cdot \mathbf{\Lambda} \cdot \mathbf{B}^H$$

$$\mathbf{C}_2 = \mathbf{B} \cdot \mathbf{D} \cdot \mathbf{\Lambda} \cdot \mathbf{B}^H$$

其中，

$$\mathbf{\Lambda} = \text{diag} \left\{ \gamma_{4,u_1} |B(1,1)|^2, \gamma_{4,u_2} |B(1,2)|^2, \dots, \gamma_{4,u_G} |B(1,G)|^2 \right\}$$

$$\mathbf{D} = \text{diag} \left\{ \frac{B(2,1)}{B(1,1)}, \frac{B(2,2)}{B(1,2)}, \dots, \frac{B(2,G)}{B(1,G)} \right\}$$

$$\gamma_{4,u_i} = \text{cum}(u_i(t), u_i^*(t), u_i(t), u_i^*(t))$$

假设传感器1和2对每一组信号的总响应均不为零，即对信号源 $u_i(t)$ 通过 p_i 条不同的路径入射到传感器1和2的总响应均不为零，则矩阵 \mathbf{B} 前两行的各元素均不等于零，从而有：

$$\text{rank} \{ \mathbf{C}_1 \} = \text{rank} \{ \mathbf{C}_2 \} = \text{rank} \{ \mathbf{\Lambda} \} = \text{rank} \{ \mathbf{D} \} = G$$

采用 TLS-ESPRIT 法，可从 \mathbf{C}_1 和 \mathbf{C}_2 中估计出矩阵 \mathbf{B} 的各列 $\{ \mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_G \}$ 。具体过程可见：

E. Gonen, J. M. Mendel. [Applications of cumulants to array processing. III Blind beamforming for coherent signals](#)中的附录。

将方向矩阵分块表示为：

$$\mathbf{A} = [\mathbf{A}_1 \quad \mathbf{A}_2 \quad \dots \quad \mathbf{A}_G]$$

其中，

$$\mathbf{A}_i = [\mathbf{a}(\theta_{i,1}) \quad \mathbf{a}(\theta_{i,2}) \quad \dots \quad \mathbf{a}(\theta_{i,p_i})]$$

显然，

$$\mathbf{b}_i = \mathbf{A}_i \cdot \mathbf{c}_i$$

记 \mathbf{b}_i 和 \mathbf{A}_i 的前 L 行分别为 $\mathbf{b}_{L,i}$ 和 $\mathbf{A}_{L,i}$ 。显然 $\mathbf{A}_{L,i}$ 是阵列中 L 个阵元的均匀线性子阵对信号源 $u_i(t)$ 通过 p_i 条不同的路径从方向 $\{ \theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,p_i} \}$ 入射到阵列上的入射信号 $\{ s_{i,1}(t), s_{i,2}(t), \dots, s_{i,p_i}(t) \}$ 的方向矩阵，且

$$\mathbf{b}_{L,i} = \mathbf{A}_{L,i} \cdot \mathbf{c}_i$$

用另一种方式表示，有：

$$\begin{aligned} b_{L,i}(n) &= \sum_{k=1}^{p_i} c_i(k) e^{j2\pi(n-1)\frac{\Delta}{\lambda} \sin(\theta_{i,k})} \\ &= \sum_{k=1}^{p_i} c_i(k) e^{j(n-1)\omega_{i,k}} \quad (n=1,2,\dots,L) \end{aligned}$$

其中， $b_{L,i}(n)$ 为 $\mathbf{b}_{L,i}$ 的第 n 个元素， $c_i(k)$ 为 \mathbf{c}_i 的第 k 个元素， $\omega_{i,k} = \frac{2\pi}{\lambda} \Delta \sin \theta_{i,k}$

为信号 $s_{i,k}(t)$ 的空间角频率。

显然， $\{b_{L,i}(n); n=1,2,\dots,L\}$ 可看成是角频率为 $\omega_{i,1}, \omega_{i,2}, \dots, \omega_{i,p_i}$ 的 p_i 个谐波分量叠加信号的一个观测样本，采用任何一种谐波恢复的方法可由其得到 $\omega_{i,1}, \omega_{i,2}, \dots, \omega_{i,p_i}$ 的估计，从而得到入射信号 $\{s_{i,1}(t), s_{i,2}(t), \dots, s_{i,p_i}(t)\}$ 的方向 $\{\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,p_i}\}$ 的估计结果。

综上所述，EVESPA算法的步骤是：

- (1) 由阵列输出估计四阶累积量矩阵 \mathbf{C}_1 和 \mathbf{C}_2 ；
- (2) 由 \mathbf{C}_1 和 \mathbf{C}_2 估计出 $\{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_G\}$ ，从而得到 $\{\mathbf{b}_{L,1}, \mathbf{b}_{L,2}, \dots, \mathbf{b}_{L,G}\}$ ；
- (3) 对 $i=1,2,\dots,G$ ，分别
 - a) 由 $\mathbf{b}_{L,i}$ 的元素 $\{b_{L,i}(n); n=1,2,\dots,L\}$ ，采用某种谐波恢复方法得到 $\omega_{i,1}, \omega_{i,2}, \dots, \omega_{i,p_i}$ 的估计；
 - b) 由 $\omega_{i,1}, \omega_{i,2}, \dots, \omega_{i,p_i}$ 得到 $\{\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,p_i}\}$ 的估计。

对谐波恢复问题，样本长度为 L 时，若采用前后向平均技术，可以估计的最大谐波分量数为 $\frac{2}{3}L$ 。因此，EVESPA算法能够估计 $\frac{2}{3}L$ 个相干信号的DOA。

另一方面，矩阵 B 的维数为 $M \times G$ ，故EVESPA算法能够分辨的最大独立信号源数为 $M-1$ 。

因而，EVESPA算法可分辨信号的总数为

$$\frac{2}{3}L(M-1) > M$$

当 $L=M$ 时，即整个阵列为ULA时，EVESPA算法可分辨信号的总数达到最大，为 $\frac{2}{3}M(M-1)$ 。

可见，**EVESPA算法不仅能够有效解决相干信号的DOA估计，而且不会造成有效孔径减小。**

小结：

(1) 对窄带信号源的波达方向估计问题，可以利用高阶累积量对高斯噪声的抑制作用，或利用存在独立子集的多个随机变量的高阶累积量恒为零，解决空间有色噪声（空间相关噪声）背景下的DOA估计问题。

(2) 基于高阶累积量的VESPA法还使任意阵列校准问题大大简化，它只需校准一对传感器，使其具有相同的响应特性，而阵列中所有其它传感器的位置和响应特性均不需知道，只要保证阵列流形无模糊。

(3) 基于高阶累积量的EVESPA法在能够分辨相干信号的同时，还不会造成有效孔径减小。