第 3 章　基于分数低阶统计量的信号处理理论与应用

3.1　概述

随机信号的统计量（statistics）主要包括均值、方差、相关、协方差等，它们为信号的分析和处理提供了丰富的信息，是分析处理随机信号的有效工具。随机信号统计量（统计矩）的整个分布范围可以从零阶一直延伸到∞阶。图 3.1 给出了统计矩的分布示意图。

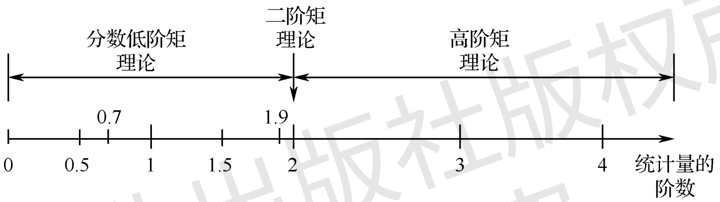


图 3.1　随机信号统计矩的分布示意图

图 3.1 中，一阶和二阶统计量（second-order statistics，SOS）是经典统计信号处理常用的统计量，例如，均值是常用的一阶统计量，而方差和相关等是常用的二阶统计量。高于二阶的统计量称为高阶矩或高阶统计量（higher-order statistics，HOS），一般取整数阶，即取 $[3,+\infty)$ 范围内的整数。而低于二阶的统计量称为分数低阶矩或分数低阶统计量，其范围为 $(0,2)$，可以取这个范围内的任意值，故称这个范围内的统计量为分数低阶量（fractional lower-order statistics，FLOS）。此外，文献中也给出了关于零阶统计量和负阶统计量的概念。

直到 20 世纪 80 年代中期，统计信号处理基本上是建立在二阶统计量基础上的，例如，对随机信号的均值、方差、相关函数和功率谱密度等的分析，以及基于信号二阶统计量的滤波、预测、检测与估值等。由于功率谱密度函数是相关函数的一维傅里叶变换，因此，功率谱也是建立在二阶统计量基础上的。众所周知，高斯分布是统计信号处理领域所普遍采用的描述随机信号和/或随机噪声的模型，这主要是由于中心极限定理决定了满足有限均值和方差的独立同分布随机变量之极限和为高斯分布，并且高斯概率密度函数可以完全由两个参数来描述，即均值和方差。这样，在统计信号处理领域采用高斯分布模型和基于二阶统计量的信号处理方法就成为顺理成章的事情。然而，基于二阶统计量的方法会受到信号噪声模型假设的限制，例如通常假设信号或噪声是二阶矩有限的，常用高斯分布来进行描述。

如果随机信号不是高斯分布的，其概率密度函数就不能仅由均值和方差这两个参数确定，高阶矩（HOM）或高阶统计量（HOS）可能比单独使用二阶统计量揭示出信号中更多的信息。严格地说，对于非高斯随机信号而言，通常需要利用其概率密度函数才能对其进行完整的刻画。但是在实际应用中，要获得随机信号的概率密度函数往往是比较困难的。不过幸

运的是，概率密度函数的特征往往可以由信号的统计量来描述。这样，在非高斯信号处理中，高阶矩或高阶统计量（特别是三阶和四阶统计量）受到了普遍的重视并得到了广泛的应用。

FLOS 是另一类分析处理非高斯信号的有力工具，是整个矩分布的另一个方面。我们知道，Alpha 稳定分布是广义的高斯分布，它比高斯分布具有更广泛的适用性。根据广义中心极限定理，Alpha 稳定分布是唯一的一类构成独立同分布随机变量之和的极限分布。若随机信号的特征指数为α，则只有阶数小于α的统计量是有界的。即若信号或噪声的特征指数满足$0<\alpha<2$（称为分数低阶 Alpha 稳定分布），则其二阶统计量和高阶统计量都是不存在的。在这种情况下，基于二阶统计量和基于高阶统计量的信号分析处理方法都会出现显著的性能退化，甚至会得出错误的结果。于是，分数低阶矩（FLOM）或分数低阶统计量成为非高斯 Alpha 稳定分布条件下信号分析处理的重要手段，自 20 世纪 90 年代中期以来，受到信号处理学术界的广泛关注。

另一方面，基于分数低阶统计量的分数低阶 Alpha 稳定分布信号处理不可避免地会对线性问题引入非线性，从而导致信号处理算法的复杂化。引入非线性的基本原因是必须在巴拿赫（Banach）空间或度量空间来解决线性估计问题，而不是在希尔伯特（Hilbert）空间。众所周知，由高斯过程所产生的线性空间是希尔伯特空间，而 Alpha 稳定分布过程的线性空间为巴拿赫空间（当$1\leqslant\alpha<2$时）或度量空间（当$0<\alpha<1$时）。对于线性估计问题来说，巴拿赫空间或度量空间均不具有与希尔伯特空间一样好的特性和结构。尽管如此，作为一种具有良好韧性的信号分析和处理工具，Alpha 稳定分布和分数低阶统计量理论及方法依然受到普遍的关注和研究，并在许多领域得到了广泛的应用。

本章重点介绍以下 4 个方面的问题：①分数低阶统计量的基本概念与基本理论；②Alpha 稳定分布随机变量的线性空间与线性理论问题；③Alpha 稳定分布下基于分数低阶统计量信号处理的应用问题；④Alpha 稳定分布下的核方法特别是核自适应滤波问题。

3.2　分数低阶统计量

随机信号的统计量包含了丰富的有关信号特性的信息。如图 3.1 所示，统计量的分布可以从零阶一直延伸到无穷阶。传统的信号处理方法通常只利用一阶和二阶统计量，后来伴随着信号处理理论的发展，出现了基于高阶统计量的信号处理技术，主要从三阶或四阶统计量中提取有用的信息。实际上，除了一阶、二阶和高阶统计量，还存在低于二阶的分数低阶统计量，其范围为$(0, 2)$，记为 FLOS。FLOS 是分析处理 Alpha 稳定分布随机信号的有力工具。

3.2.1　分数低阶矩与$S\alpha S$分布下的距离概念

1．分数低阶矩

随机变量X的二阶矩通常定义为$E[X^2]$。对于 Alpha 稳定分布随机变量，定义分数低阶矩（fractional lower order moment，FLOM）为$E[|X|^p]$，其中，$0<p<\alpha\leqslant 2$。

由第 2 章给出的定理 2.3 可知，对于 Alpha 稳定分布随机变量X，若其特征指数满足$0<\alpha<2$，则有

$$\begin{cases} E[|X|^p]=\infty, & 若\ p\geqslant\alpha \\ E[|X|^p]<\infty, & 若\ 0\leqslant p<\alpha \end{cases} \tag{3.1}$$

若满足$\alpha=2$，则有

$$E[|X|^p]<\infty,\quad p\geqslant 0 \tag{3.2}$$

显然，若 Alpha 稳定分布随机变量的特征指数为$0<\alpha<2$，则只有阶数小于α的统计量是有界的，其方差（或其他二阶统计量）是无界的。这样，基于二阶统计量有界假设的信号处理方法将会显著退化，甚至会导致错误的结果。

定理 3.1　分数低阶矩定理　设X为位置参数$a=0$的对称 Alpha 稳定分布（$S\alpha S$分布）随机变量，其分数低阶矩$E[|X|^p]$与分散系数γ之间满足

$$E[|X|^p]=\begin{cases}C(p,\alpha)\gamma^{p/\alpha}, & 0<p<\alpha\\ +\infty, & p\geqslant\alpha\end{cases} \tag{3.3}$$

式中，

$$C(p,\alpha)=\frac{2^{p+1}\Gamma\left(\dfrac{p+1}{2}\right)\Gamma(-p/\alpha)}{\alpha\sqrt{\pi}\Gamma(-p/2)} \tag{3.4}$$

式中，伽马函数$\Gamma(x)=\int_{-\infty}^{x}t^{x-1}\mathrm{e}^{-t}\mathrm{d}t$。$C(p,\alpha)$为$\alpha$和$p$的函数。

若X是$S\alpha S$分布随机变量，分散系数为$\gamma>0$，位置参数为$a=0$，则X的α范数定义为

$$\|X\|_\alpha=\begin{cases}\gamma^{1/\alpha}, & 1\leqslant\alpha\leqslant 2\\ \gamma, & 0<\alpha<1\end{cases} \tag{3.5}$$

即范数$\|X\|_\alpha$只与分散系数γ有关。

2．$S\alpha S$分布下的距离概念

若X和Y是联合$S\alpha S$分布的，则X和Y之间的距离定义为

$$d_\alpha(X,Y)=\|X-Y\|_\alpha \tag{3.6}$$

结合式（3.3）和式（3.5），可以得到

$$d_\alpha(X,Y)=\begin{cases}\left[\dfrac{E(|X-Y|^p)}{C(p,\alpha)}\right]^{1/p}, & 0<p<\alpha,\ 1\leqslant\alpha\leqslant 2\\ \left[\dfrac{E(|X-Y|^p)}{C(p,\alpha)}\right]^{\alpha/p}, & 0<p<\alpha,\ 0<\alpha<1\end{cases} \tag{3.7}$$

故当$0<p<\alpha$时，两个$S\alpha S$分布随机变量的距离等同于这两个随机变量之差的p阶矩。当$\alpha=2$时，距离等于这两个随机变量之差的方差的一半。另一方面，$S\alpha S$分布随机变量的所有分数低阶矩都是等价的，也就是说，对于任意$0<p$，$q<\alpha$，$S\alpha S$分布随机变量的p阶矩和q阶矩只相差一个与该随机变量独立的系数。

3.2.2　零阶矩与负阶矩

1．零阶矩

（1）对数矩的概念

零阶统计量（zero-order statistics，ZOS）的概念是 1997 年由 Gonzalez 等人基于对数“矩”提出的。随机变量X的分布具有代数型拖尾时，定义对数矩为

$$E[\lg|X|]<\infty \tag{3.8}$$

已经证明，Alpha 稳定分布过程的对数矩是有界的。由于通常以“二阶”过程来表示方差有

界的随机过程，故以“对数阶”过程来表示对数矩有界的随机过程。这样，Alpha 稳定分布过程属于对数阶过程。

（2）几何功率

定义 3.1　几何功率　设 X 为一对数阶随机变量，则 X 的几何功率定义为

$$S_0 = S_0(X) = \exp\{E[\lg|X|]\} \tag{3.9}$$

由于随机过程的功率属于二阶统计量范畴，而分数低阶 Alpha 稳定分布没有有限的二阶统计量，故不能用“功率”这个概念来描述和评价 Alpha 稳定分布过程。但是，用几何功率却可以描述和评价 Alpha 稳定分布过程的强度或“功率”指标，显示出一定的优越性。图 3.2 给出了两个随机信号分别用常规的“二阶功率”和几何功率来描述的情况。

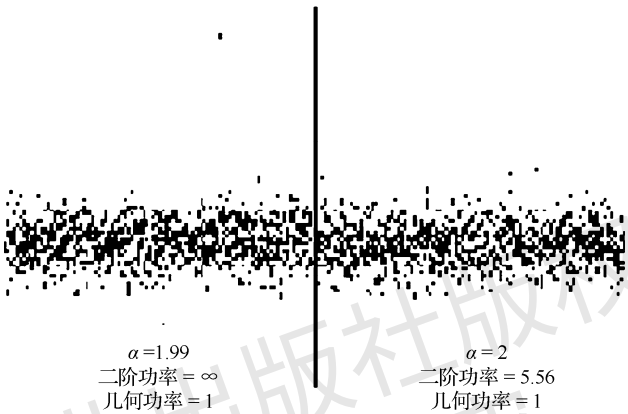


图 3.2　独立同分布过程二阶功率与几何功率的比较

在图 3.2 中，左半部分对应于 $\alpha=1.99$ 的 Alpha 稳定分布过程，右半部分对应于 $\alpha=2$ 的高斯分布过程。若用几何功率来测量这两个信号，所得到的结果都是 $S_0=1$，即二者具有相同的几何功率。事实上，二者确实具有相同的强度。然而，如果采用常规的二阶功率来测量，由于 Alpha 稳定分布的二阶统计量趋于无穷，故左半部分的二阶功率趋于无穷，这与真实情况相去甚远。这表明常规的二阶功率不适合描述分数低阶 Alpha 稳定分布过程，而几何功率则可以客观准确地描述。几何功率的这种特性提供了一种可以用于比较任意对数阶信号的通用框架。

（3）零阶矩

零阶矩（zero order moment，ZOM）的概念是根据几何功率与分数低阶统计量的关联而引出的。

定理 3.2　零阶矩定理　设 $S_p=\left[E[|X|^p]\right]^{1/p}$ 表示由随机变量 X 的分数低阶矩 $E[|X|^p]$ 定义的尺度参数。若对于足够小的 p 值，S_p 存在，则几何功率 S_0 与分数低阶矩尺度参数 S_p 的关联如下：

$$S_0=\lim_{p\to 0} S_p \tag{3.10}$$

定理 3.2 表明，几何功率与分数低阶统计量是零阶相关的。由此定义了零阶矩或零阶统计量的概念。零阶矩或零阶统计量对于估计强脉冲噪声下随机信号的均值或中值等参数，可以得到比分数低阶统计量更好的估计结果。

2．负阶矩

定理 3.3　负阶矩定理　设 X 是一个 $S\alpha S$ 分布随机变量，其位置参数 $a=0$，分散系数为 γ，其矩的统一表达式为

$$E(|X|^p)=C(p,\alpha)\gamma^{p/\alpha},\quad -1<p<\alpha \tag{3.11}$$

式中，$C(p,\alpha)$ 与式（3.4）的形式相同。

显然，定理 3.3 把分数低阶矩的范围由式（3.3）的 $0<p<\alpha$ 扩展到 $-1<p<\alpha$。实际上，负阶矩式（3.11）的定义式，对于 $p=0$ 也是成立的，这是另一种形式的零阶矩。可以认为，式（3.11）是一个扩展的分数低阶统计量。

3.2.3　共变的定义与性质

由于当 $\alpha<2$ 时，$S\alpha S$ 分布没有有限的二阶统计量，故其方差和协方差都是不存在的。为了解决 $S\alpha S$ 分布下随机变量的分析处理问题，Miller 和 Cambanis 提出了共变（covariation）的概念。共变这个量在 $S\alpha S$ 分布理论中的地位与协方差在高斯分布理论中的地位非常相似。

1．共变的定义与计算

定义 3.2　共变　对于联合 $S\alpha S$ 分布的随机变量 X 和 Y，其特征指数 $1<\alpha\leqslant 2$，则 X 与 Y 的共变定义为

$$[X,Y]_\alpha=\int_s xy^{\langle\alpha-1\rangle}\mu(\mathrm{d}s) \tag{3.12}$$

式中，s 表示单位圆，$\mu(\cdot)$ 表示 $S\alpha S$ 分布随机矢量 (X,Y) 的谱测度。式（3.12）中运算符 $\langle\cdot\rangle$ 的含义为

$$z^{\langle b\rangle}=|z|^b\,\mathrm{sgn}(z) \tag{3.13}$$

式中，z 为任意实变量，参数 $b\geqslant 0$，$\mathrm{sgn}(\cdot)$ 为符号函数。此外，定义联合 $S\alpha S$ 分布随机变量 X 与 Y 的共变系数为

$$\lambda_{XY}=\frac{[X,Y]_\alpha}{[Y,Y]_\alpha} \tag{3.14}$$

与高斯随机变量的协方差不同，除 $\alpha=2$ 以外，对于其他的 α 值，共变不存在对称性，即

$$[X,Y]_\alpha\neq[Y,X]_\alpha,\quad 1<\alpha<2 \tag{3.15}$$

由于谱测度 $\mu(\cdot)$ 不易得到，因此上述关于共变和共变系数的定义很难应用于实际问题。然而，由于共变、共变系数与分数低阶矩之间存在一定的联系，从而使共变成为具有实际应用价值的概念与算子。

定理 3.4　共变定理　设 X 和 Y 是联合 $S\alpha S$ 分布随机变量，特征指数满足 $1<\alpha\leqslant 2$。若 Y 的分散系数为 γ_y，则

$$[Y,Y]_\alpha=\|Y\|_\alpha^\alpha=\gamma_y \tag{3.16}$$

$$\lambda_{XY}=\frac{E[XY^{\langle p-1\rangle}]}{E[|Y|^p]},\qquad 1\leqslant p<\alpha \tag{3.17}$$

$$[X,Y]_\alpha=\frac{E[XY^{\langle p-1\rangle}]}{E[|Y|^p]}\gamma_y,\qquad 1\leqslant p<\alpha \tag{3.18}$$

定理 3.4 把共变的概念转化为对分数低阶矩 $E[|Y|^p]$ 和分数低阶协方差 $E[XY^{\langle p-1\rangle}]$ 的计算，从而使共变成为分数低阶信号处理中的一个有力工具。关于分数低阶协方差或分数低阶相关的概念，可参见 3.2.4 节的介绍。

2．共变的性质

共变具有一些十分重要的性质，对 Alpha 稳定分布信号处理有很重要的作用。

性质 3.1　若 X 和 Y 是联合 $S\alpha S$ 分布的，设 a 和 b 是任意实数，则共变 $[X,Y]_\alpha$ 对第一变元 X 是线性的，即

$$[aX_1+bX_2,Y]_\alpha = a[X_1,Y]_\alpha + b[X_2,Y]_\alpha \tag{3.19}$$

性质 3.2　当 $\alpha=2$，即 X 和 Y 是零均值联合高斯分布随机变量时，X 和 Y 的共变退化为协方差

$$[X,Y]_\alpha = E[XY] \tag{3.20}$$

性质 3.3　若 Y_1 和 Y_2 是独立的，设 a 和 b 是任意实数，并且 X、Y_1 和 Y_2 是联合 $S\alpha S$ 分布的，则共变具有对第二变元的伪线性。

$$[X,aY_1+bY_2]_\alpha = a^{\langle\alpha-1\rangle}[X,Y_1]_\alpha + b^{\langle\alpha-1\rangle}[X,Y_2]_\alpha \tag{3.21}$$

性质 3.4　若 X 和 Y 是独立 $S\alpha S$ 分布的，则

$$[X,Y]_\alpha = 0 \tag{3.22}$$

但是，反之通常是不成立的。

性质 3.5　对于联合 $S\alpha S$ 分布的随机变量 X 和 Y，存在 Cauchy-Schwartz 不等式

$$\left|[X,Y]_\alpha\right| \leqslant \|X\|_\alpha \|Y\|_\alpha^{\langle\alpha-1\rangle} \tag{3.23}$$

如果 X 和 Y 具有单位分散系数，则

$$\left|[X,Y]_\alpha\right| \leqslant 1 \tag{3.24}$$

定理 3.5　共变计算　设 U_i，$i=1,\cdots,n$ 为独立 $S\alpha S$ 分布随机变量，其分散系数为 γ_i，$i=1,\cdots,n$。若对于任意系数 $a_1,\cdots,a_n$ 和 $b_1,\cdots,b_n$，其中 $b_i\neq 0,\ i=1,\cdots,n$，有 $X=a_1U_1+\cdots+a_nU_n$ 和 $Y=b_1U_1+\cdots+b_nU_n$，则

$$[X,X]_\alpha = \gamma_1|a_1|^\alpha + \cdots + \gamma_n|a_n|^\alpha \tag{3.25}$$

$$[Y,Y]_\alpha = \gamma_1|b_1|^\alpha + \cdots + \gamma_n|b_n|^\alpha \tag{3.26}$$

$$[X,Y]_\alpha = \gamma_1 a_1 b_1^{\langle\alpha-1\rangle} + \cdots + \gamma_n a_n b_n^{\langle\alpha-1\rangle} \tag{3.27}$$

$$\lambda_{XY} = \frac{\gamma_1 a_1 b_1^{\langle\alpha-1\rangle} + \cdots + \gamma_n a_n b_n^{\langle\alpha-1\rangle}}{\gamma_1|b_1|^\alpha + \cdots + \gamma_n|b_n|^\alpha} \tag{3.28}$$

例 3.1　设 $\alpha=1.5$，$A=1$，$B=-2$，$[X,Y_1]_\alpha=2$，$[X,Y_2]_\alpha=3$。试求 $[X,AY_1+BY_2]_\alpha$。

解　利用共变的伪线性性质，有 $[X,AY_1+BY_2]_\alpha = A^{\langle\alpha-1\rangle}[X,Y_1]_\alpha + B^{\langle\alpha-1\rangle}[X,Y_2]_\alpha$。将其代入已知条件，有

$$[X,AY_1+BY_2]_\alpha = 2-3\sqrt{2}$$

例 3.2　设 U 为 $S\alpha S$ 分布随机变量，其分散系数为 $\gamma=1$。设 $X=a_1U$，$Y=U$，试求共变 $[X,Y]_\alpha$ 和共变系数 λ_{XY}。

解　根据给定条件，X 和 Y 不是相互独立的。由定理 3.5，可得共变和共变系数为

$$[X,Y]_\alpha = \gamma a_1 b_1^{\langle\alpha-1\rangle} = 1\times a_1\times 1 = a_1$$

$$\lambda_{XY} = \frac{\gamma_1 a_1 b_1^{\langle \alpha-1 \rangle}}{\gamma |b_1|^{\alpha}} = a_1$$

3．复 $S\alpha S$ 分布随机变量的共变

若复变量 $X = X_1 + \mathrm{j}X_2$ 和 $Y = Y_1 + \mathrm{j}Y_2$ 服从联合 $S\alpha S$ 分布，则可定义 X 与 Y 的共变和共变系数。实际上，复变量的共变系数与实变量的共变系数具有相同的形式，并且实 $S\alpha S$ 分布共变的一些性质可以推广到复 $S\alpha S$ 分布共变的情况。例如，复变量共变系数的计算式与式(3.17）相同；关于第一变元的线性性质，复共变与实共变二者具有相同的形式；若复 $S\alpha S$ 分布随机变量 X 和 Y 是独立的，则 $[X,Y]_\alpha = 0$，也与实 $S\alpha S$ 分布的情况相同。

3.2.4 分数低阶相关与相位分数低阶矩算子

1．分数低阶相关

由定义 3.2 和定理 3.4 可知，尽管共变作为一种重要的分数低阶统计量，在信号处理中占有重要的地位，但其只适用于特征指数 $1 < \alpha \leqslant 2$ 的情形，并且只对单个变量进行 $z^{\langle p-1 \rangle} = |z|^{p-1}\,\mathrm{sgn}(z)$ 的幅度压缩处理。针对这些问题，提出了分数低阶相关（fractional lower-order correlation），又称为分数低阶协方差（fractional lower-order covariance）的概念。

需要说明的是，在统计信号处理中，相关函数（correlation function）与协方差函数（covariance function）的概念与计算是相似的，但二者之间有一定的差异。简单来说，二者的差异主要表现为是否去除随机信号的均值。不去除均值的是相关函数，去除均值的是协方差函数。在零均值条件下，则认为二者是相等的。鉴于此，本节不严格区分分数低阶相关函数与分数低阶协方差函数，以分数低阶相关的概念表示二者的运算。

（1）第一类分数低阶相关

第一类分数低阶相关表示为

$$R_{XY}^{C} = E[XY^{\langle p-1 \rangle}], \quad 1 \leqslant p < \alpha \tag{3.29}$$

显然，第一类分数低阶相关是非对称的，即 $R_{XY}^{C} \neq R_{YX}^{C}$。可以看出，第一类分数低阶相关与共变系数式（3.17）是非常相似的。

（2）第二类分数低阶相关

第二类分数低阶相关表示为

$$R_{XY}^{D} = E[X^{\langle a \rangle} Y^{\langle b \rangle}], \quad 0 \leqslant a < \alpha/2, \quad 0 \leqslant b < \alpha/2 \tag{3.30}$$

式中，a 和 b 为表示阶数的参量。当 $a = b$ 时，第二类分数低阶相关是对称的，即 $R_{XY}^{D} = R_{YX}^{D}$。

2．相位分数低阶协方差与相位分数低阶矩算子

对于任意两个服从联合 $S\alpha S$ 分布的随机变量 X 和 Y，若其特征指数满足 $0 < \alpha < 2$，且 p 指数满足 $0 < p < \alpha/2$，则相位分数低阶协方差定义为

$$R_{XY}^{P} = E[X^{\langle p \rangle} Y^{-\langle p \rangle}] \tag{3.31}$$

式中，对于复变量 z，有

$$z^{\langle p \rangle} = \begin{cases} |z|^{p+1}/z^{*}, & z \neq 0, \quad z \in \mathbb{C} \\ 0, & z = 0 \end{cases} \tag{3.32}$$

称为相位分数低阶矩（phased fractional lower-order moment，PFLOM）算子，简称为 PFLOM 算子。显然，PFLOM 算子是针对复变量 z 设计的。若 z 为实变量，则式（3.32）退化为式（3.13）。

对于脉冲噪声，PFLOM 算子具有很好的抑制能力。设复变量 $z = r\mathrm{e}^{\mathrm{j}\theta}$，易于得到 $z^{\langle p\rangle} = r^{p}\mathrm{e}^{\mathrm{j}\theta}$。这表明，PFLOM 算子仅对被处理信号噪声的幅度进行处理，而保留其相位不变。若设定 $p=1$，数据没有任何变化；而若设定 $p=0$，则从信号中删除了所有的幅度信息，仅保留下其相位信息。PFLOM 算子 p 指数变化对信号振幅的调节作用如图 3.3 所示。

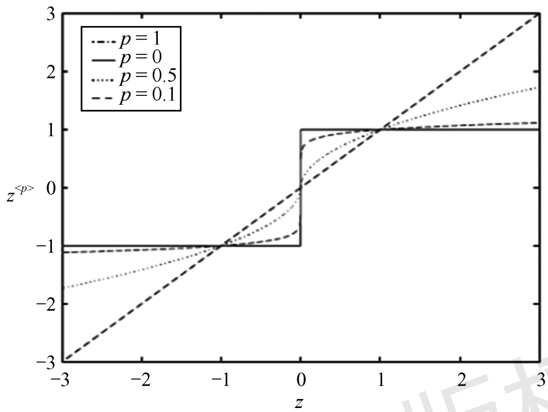


图 3.3　PFLOM 算子 p 指数变化对信号振幅的调节作用

由于 PFLOM 算子对信号振幅的抑制或调节作用，可以证明式（3.31）的相位分数低阶协方差在分数低阶 Alpha 稳定分布噪声环境下是有界的。

3.3　Alpha 稳定分布的线性理论

3.3.1　Alpha 稳定分布的线性空间

所谓线性空间，是定义了加法和数乘的空间。或者说，线性空间中的一个元素，可以由其他元素线性表出。

统计信号处理中的一个重要问题是给出一组观测样本 $\{X(t), t\in T\}$，从其张成的线性空间中找到未知随机变量 Y 的最佳估计。这就是所谓的随机过程的线性理论，包括线性估计、线性预测和线性滤波。

长期以来，二阶矩过程（特别是高斯过程）的线性理论得到了充分的发展。在二阶统计量存在的情况下，观测样本 $\{X(t), t\in T\}$ 的线性空间 $L\{X(t), t\in T\}$ 是希尔伯特空间。在最小均方误差（MMSE）准则下对未知变量 Y 的最佳线性估计是 Y 在 $L\{X(t), t\in T\}$ 上的正交投影。

然而，对于分数低阶 Alpha 稳定分布过程来说，其线性空间的问题要更为复杂。假设 $\{X(t), t\in T\}$ 为服从 $S\alpha S$ 分布的随机过程，研究表明，当 $1\leqslant\alpha<2$ 时，$S\alpha S$ 分布过程张成的线性空间是巴拿赫（Banach）空间；而当 $0<\alpha<1$ 时，$S\alpha S$ 分布过程张成的线性空间是度量（metric）空间。

我们知道，希尔伯特空间是一个内积空间，有度量和角度以及由此引申而来的正交性概念。此外，希尔伯特空间还是一个完备的空间，该空间上的所有柯西序列均收敛，因此数学分析中的大部分概念都可以无障碍地推广到希尔伯特空间中。希尔伯特空间为基于正交基上

多项式表示的傅里叶级数和傅里叶变换提供了一种有效的表述方式，而这也是泛函分析的核心概念之一。由此可见，在希尔伯特空间进行信号处理有许多便利条件。

作为完备的内积赋范空间，希尔伯特空间是巴拿赫空间的一个特例，或者说，巴拿赫空间是希尔伯特空间的推广。巴拿赫空间是完备的线性赋范空间，不过其范数却不一定是基于内积定义的。另一方面，巴拿赫空间又是引入范数的度量空间，而一般的度量空间则为仅定义了距离测度的空间。显然，巴拿赫空间特别是度量空间并不具备像希尔伯特空间那样具有好的特性和结构，因此在这两个空间进行信号分析与处理，有许多需要特别解决的问题。

3.3.2 Alpha 稳定分布的线性理论

随机信号的线性理论可以简要地归纳为，依据观测信号数据样本，从随机信号张成的线性空间中对未知变量进行最优估计、预测或滤波的问题。对于二阶随机过程而言，线性估计（或预测、滤波）的最优准则常采用最小均方误差（MMSE）准则，即

$$E[(\theta-\hat{\theta})^2]\to\min \tag{3.33}$$

式（3.33）表示，MMSE 准则使估计值 $\hat{\theta}$ 与真值 θ 的误差均方值达到最小，从而实现最优估计（或预测、滤波）。

对于分数低阶 Alpha 稳定分布，由于其没有有限的二阶统计量，因此最小均方误差准则不再成立。然而，由 Alpha 稳定分布的线性空间理论可知，Alpha 稳定分布所张成的线性空间是巴拿赫空间（$1\leqslant\alpha<2$）或度量空间（$0<\alpha<1$），前者可以定义范数的概念，且二者均可定义距离的概念。式（3.6）实际上定义了 Alpha 稳定分布的距离的概念，并可以进一步引申为线性估计（或预测、滤波）中真实值与估计值之间的距离，用来测度二者之间误差的大小

$$d_\alpha(\theta,\hat{\theta})=\left\|\theta-\hat{\theta}\right\|_\alpha=\|e\|_\alpha \tag{3.34}$$

式中，$e=\theta-\hat{\theta}$ 表示二者的误差。当距离 $d_\alpha(\theta,\hat{\theta})$ 达到最小时，$\hat{\theta}$ 为 θ 的最优估计。故 $d_\alpha(\theta,\hat{\theta})\to\min$ 可称为最小距离准则。

另一方面，由式（3.7）可知，Alpha 稳定分布距离的最小化对应于其分数低阶矩的最小化，而由式（3.5）和式（3.34），可推得 $d_\alpha(\theta,\hat{\theta})$ 的最小化对应于 Alpha 稳定分布分散系数的最小化，即

$$\min\{d_\alpha(\theta,\hat{\theta})\}=\min\{\|e\|_\alpha\}\Leftrightarrow\min\{E[|e|^p\}\Leftrightarrow\min\{\gamma_e\},\quad 0<\alpha<2 \tag{3.35}$$

式中，符号“$\Leftrightarrow$”表示等价于，γ_e 表示误差 e 的分散系数。由式（3.35）可见，Alpha 稳定分布的最小距离准则，等价于最小 α 范数准则，等价于最小 p 阶矩准则，也等价于最小分散系数（minimal distribution，MD）准则。

对于 Alpha 稳定分布过程来说，分散系数 γ 的概念与高斯过程方差的概念具有同样的地位和作用，均表示统计分布的分散程度，或表示接收数据的“功率”或“能量”方面的信息。分散系数越大，远离分布均值或中值的随机变量样本越多，概率密度函数的拖尾越厚。通过分散系数的最小化，可以使估计误差的平均幅度达到最小。因此，在 Alpha 稳定分布的线性滤波、估计和预测等问题中，常用 MD 准则替代二阶过程的 MMSE 准则，实现在 Alpha 稳定分布噪声条件下的统计滤波或参数估计等运算。

3.4　Alpha 稳定分布下的参数模型与估计

3.4.1　经典的线性参数模型

许多平稳随机信号 $x(n)$ 可以看作由白噪声 $w(n)$ 激励某一确定的线性系统 $h(n)$ 而得到的响应。这样，只要白噪声的参数确定了，对随机信号的研究就可以转化为对产生随机信号的线性系统的研究。这就是随机信号研究分析的参数模型法。一方面，由给定的平稳随机信号序列，可以为其建立相应的参数模型，即模型建立或模型参数估计问题；另一方面，利用给定白噪声和参数模型，可以产生出所需要的随机信号序列，即随机信号产生问题。来进行功率谱估计及其他应用问题。

1．MA 模型

广义平稳随机信号 $x(n)$ 的 MA（moving average，滑动平均）模型满足以下表达式：

$$x(n)=\sum_{k=0}^{q}b_k w(n-k) \tag{3.36}$$

式中，$w(n)$ 表示激励白噪声，b_k 为 MA 模型系数，q 为模型阶数。记式（3.36）所示模型为 $\mathrm{MA}(q)$。经过 z 变换和整理，可得 MA 系统的系统函数为

$$H(z)=\frac{X(z)}{W(z)}=\sum_{k=0}^{q}b_k z^{-k} \tag{3.37}$$

由于 $H(z)$ 在有限范围内没有极点，故称 MA 模型为全零点模型，且系统一定是稳定的。

2．AR 模型

广义平稳随机信号 $x(n)$ 的 AR（autoregressive，自回归）模型满足以下表达式：

$$x(n)=-\sum_{k=1}^{p}a_k x(n-k)+w(n) \tag{3.38}$$

式中，a_k 为 AR 模型的系数，p 为模型阶数，记为 $\mathrm{AR}(p)$。经过 z 变换和整理，可得 AR 系统的系统函数为

$$H(z)=\frac{1}{1+\sum\limits_{k=1}^{p}a_k z^{-k}} \tag{3.39}$$

可见，系统函数中只有极点，没有零点，故称该系统为全极点模型。由于系统存在极点，因此需要考虑系统的稳定性问题。

3．ARMA 模型

广义平稳随机信号 $x(n)$ 的 ARMA（autoregressive moving average，自回归滑动平均）模型满足以下表达式：

$$x(n)=\sum_{k=0}^{q}w(n-k)-\sum_{k=1}^{p}a_k x(n-k) \tag{3.40}$$

显然，ARMA 模型是 AR 模型与 MA 模型的结合，记为 $\mathrm{ARMA}(p,q)$，其系统函数为

$$H(z)=\frac{\sum_{k=0}^{q} b_k z^{-k}}{1+\sum_{k=1}^{p} a_k z^{-k}} \tag{3.41}$$

ARMA 模型既有零点，又有极点，称为零极点模型。

例 3.3 试利用 MATLAB 编程，分别产生 AR(2)、MA(1)和 ARMA(1, 2)随机序列，并绘制波形图。

解 MATLAB 程序如下：

```
clear all; clear; clc;
whitenoise = randn(1000,1); % 产生1000点高斯分布白噪声
a1 = [1.0 -0.2 -0.1]; b1 = 1; y1 = filter(b1,a1,whitenoise); % 2阶AR模型与信号
a2 = 1; b2 = [1.0 0.1]; y2 = filter(b2,a2,whitenoise); % 1阶MA模型与信号
a3 = [1.0 -0.2]; b3 = [1.0 0.1 -0.2]; y3 = filter(b3,a3,whitenoise); % ARMA模型与信号
figure(1);
subplot(3,1,1); plot(y1); xlabel('样本数'/n); ylabel('幅度'); % 绘制曲线
subplot(3,1,2); plot(y2); xlabel('样本数'/n); ylabel('幅度'); % 绘制曲线
subplot(3,1,3); plot(y3); xlabel('样本数'/n); ylabel('幅度'); % 绘制曲线
```

运行上述程序，可以分别产生 AR(2)、MA(1)、ARMA(1, 2)数据，并绘制曲线。

4．AR 模型的参数估计

根据沃尔德分解定理，AR、MA 和 ARMA 三个模型是可以相互转换的。故在信号处理的参数建模领域，一般主要讨论 AR 模型的参数估计问题。

对式（3.38）所示的 AR 模型 $x(n)=-\sum_{k=1}^{p} a_k x(n-k)+w(n)$ 求取自相关函数，并利用自相关函数 $R_x(m)$ 的对称性，可以得到

$$R_x(m)=\begin{cases} -\sum_{k=1}^{p} a_k R_x(m-k), & m>0 \\ -\sum_{k=1}^{p} a_k R_x(m-k)+h(0)\sigma_w^2, & m=0 \\ R_x(-m), & m<0 \end{cases} \tag{3.42}$$

对于因果系统 $h(n)$，经整理可得到 Yule-Walker 方程的矩阵形式为

$$\begin{bmatrix} R_x(0) & R_x(-1) & \cdots & R_x(-p) \\ R_x(1) & R_x(0) & \cdots & R_x(1-p) \\ \vdots & \vdots & \ddots & \vdots \\ R_x(p) & R_x(p-1) & \cdots & R_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma_w^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{3.43}$$

在阶数 p 及自相关函数 $R_x(m)$ 已知或可估计的前提下，通过求解 Yule-Walker 方程，可以得到 AR 参数的估计。

3.4.2　基于共变的 AR 参数模型

在分数低阶 Alpha 稳定分布条件下，经典的 Yule-Walker 方程由于自相关函数不存在而失去意义，因此 AR、MA 和 ARMA 模型参数的估计需要寻找新的方法。为了解决这个问题，文献中提出了一种广义 Yule-Walker（GYW）方程。

1．广义 Yule-Walker 方程

设服从 $S\alpha S$ 分布的随机信号 $x(n)$ 的 p_a 阶 AR 模型为

$$x(n)=-\sum_{k=1}^{p_a}a_k x(n-k)+v(n) \tag{3.44}$$

式中，$v(n)$ 表示特征指数为 α 、分散系数为 γ 的 $S\alpha S$ 分布过程。对式（3.44）两边同时取条件期望，可以得到

$$E[x(n)|x(m)]=a_1E[x(n-1)|x(m)]+\cdots+a_{p_a}E[x(n-p_a)|x(m)]，\ n-p_a<m<n-1 \tag{3.45}$$

由于 $S\alpha S$ 分布噪声 $v(n)$ 与 $x(n)$ 互相独立，故 $E[v(n)|x(m)]=0$ 。根据条件期望的性质，可得

$$E[x(n+l)|x(n)]=\lambda(l)x(n) \tag{3.46}$$

式中，$\lambda(l)$ 是 $x(n+l)$ 与 $x(n)$ 的共变系数，且 $\lambda(0)=1$ 。将式（3.46）代入式（3.45），有

$$\lambda(n-m)x(m)=a_1\lambda(n-m-1)x(m)+\cdots+a_{p_a}\lambda(n-m-p_a)x(m),\ \ n-p_a\leqslant m\leqslant n-1 \tag{3.47}$$

由式（3.47）可以得到广义 Yule-Walker 方程的矩阵形式为

$$\begin{bmatrix}\lambda(0) & \lambda(-1) & \cdots & \lambda(1-p_a)\\ \lambda(1) & \lambda(0) & \cdots & \lambda(2-p_a)\\ \vdots & \vdots & \ddots & \vdots\\ \lambda(p_a-1) & \lambda(p_a-2) & \cdots & \lambda(0)\end{bmatrix}\begin{bmatrix}a_1\\ a_2\\ \vdots\\ a_{p_a}\end{bmatrix}=\begin{bmatrix}\lambda(1)\\ \lambda(2)\\ \vdots\\ \lambda(p_a)\end{bmatrix} \tag{3.48}$$

式中，$\boldsymbol{C}=\begin{bmatrix}\lambda(0) & \lambda(-1) & \cdots & \lambda(1-p_a)\\ \lambda(1) & \lambda(0) & \cdots & \lambda(2-p_a)\\ \vdots & \vdots & \ddots & \vdots\\ \lambda(p_a-1) & \lambda(p_a-2) & \cdots & \lambda(0)\end{bmatrix}$ 为共变矩阵，是托普利兹矩阵。定义共变系数矢量 $\boldsymbol{p}=[\lambda(1),\ \lambda(2),\ \cdots,\ \lambda(p_a)]^{\mathrm{T}}$ 和 AR 参数矢量 $\boldsymbol{a}=[a_1,a_2,\cdots,a_{p_a}]^{\mathrm{T}}$ ，有

$$\boldsymbol{Ca}=\boldsymbol{p} \tag{3.49}$$

若 $\boldsymbol{C}$ 是满秩的，则可以通过求解式（3.48）或式（3.49）而得到 AR 参数 $\boldsymbol{a}$ 的估值。在实际应用中，需要根据接收数据 $x(n)$ 来估计共变矩阵 $\boldsymbol{C}$ 和矢量 $\boldsymbol{p}$ ，因此需要估计共变系数 $\lambda(l)$ 。

2．共变系数的 FLOM 估计方法

为了求解广义 Yule-Walker 方程，需要依据接收数据对共变系数 $\lambda(l)$ 进行估计。根据式（3.17）给出的共变系数定义，可以得到共变系数估计的分数低阶矩（FLOM）法

$$\hat{\lambda}_{\mathrm{FLOM}}(X,Y)=\frac{\sum_{i=1}^{N}X_i|Y_i|^{p-1}\operatorname{sgn}(Y_i)}{\sum_{i=1}^{N}|Y_i|^{p}} \tag{3.50}$$

式中，$(X_1,Y_1),\cdots,(X_n,Y_n)$ 为独立观测数据。若考虑计算效率，可以选择 $p=1$ ，则

$$\hat{\lambda}_{\mathrm{FLOM}}(X,Y)=\frac{\sum_{i=1}^{N}X_i\,\mathrm{sgn}(Y_i)}{\sum_{i=1}^{N}|Y_i|} \tag{3.51}$$

无论是对于相互独立观测数据还是 AR 过程的输出序列，基于 FLOM 的共变系数估计都是一致估计。分析表明，广义 Yule-Walker 方程中的共变系数可以由 FLOM 法来估计，进而得到 AR 参数的一致估计。

3.4.3 参数模型的最小 p 范数估计

1．AR 模型参数的最小 p 范数（LPN）估计

由分数低阶矩（FLOM）的定义可以得到在最小分散系数（MD）准则下模型参数辨识的最优方法：求 $S\alpha S$ 分布随机变量的 p 阶分数低阶矩的最小化，便得到 MD 意义下的最优参数集。特别地，对 p_a 阶线性 AR 模型估计问题，可以得到 AR 参数的最小 p 范数（least p-norm，LPN）估计

$$\hat{\boldsymbol{a}}_{\mathrm{LPN}}=\arg\min_{\boldsymbol{a}}\sum_{n=p_a}^{N}|x(n)-a_1x(n-1)-\cdots-a_{p_a}x(n-p_a)|^p,1\leqslant p\leqslant 2 \tag{3.52}$$

这里讨论的 LPN 方法限于$1\leqslant p\leqslant 2$，因为对于 $p<1$，误差函数不再是凸函数，且收敛算法复杂。而且，对于$\alpha<1$，$S\alpha S$ 分布信号或噪声的脉冲性过强，通常不用作信号处理问题的统计模型。LPN 估计是一致收敛的。

2．AR 参数的最小绝对偏差（LAD）估计

在式（3.52）中，若选取 $p=1$，则 AR 参数 $\boldsymbol{a}=[a_1,\ a_2,\ \cdots,\ a_{p_a}]^{\mathrm{T}}$ 的估计称为最小绝对偏差（LAD）估计

$$\hat{\boldsymbol{a}}_{\mathrm{LAD}}=\arg\min_{\boldsymbol{a}}\sum_{n=p_a}^{N}|x(n)-a_1x(n-1)-\cdots-a_{p_a}x(n-p_a)| \tag{3.53}$$

通常，LAD 估计是一致估计。

3.5 基于分数低阶统计量的自适应滤波

自适应滤波是信号处理中一个重要的研究领域，并在许多领域得到广泛的应用。本节介绍几种典型的基于分数低阶统计量的自适应滤波算法。

3.5.1 最小平均 p 范数（LMP）自适应滤波器

1．经典的最小均方（LMS）自适应滤波

考虑一个以 $\boldsymbol{x}(n)=[x(n),\ x(n-1),\ \cdots,\ x(n-M)]^{\mathrm{T}}$ 为输入信号的有限脉冲响应（FIR）横向自适应滤波器，其权矢量为 $\boldsymbol{w}(n)=[w_0(n),\ w_1(n),\ \cdots,\ w_M(n)]^{\mathrm{T}}$，期望响应（或称为参考信号）为 $d(n)$，滤波器的输出为 $y(n)$。

经典的 LMS（lease mean square）自适应滤波器是基于最小均方误差（MMSE）准则的算法，其期望响应 $d(n)$ 与滤波器输出 $y(n)$ 之差

$$e(n) = d(n) - y(n) \tag{3.54}$$

的均方值定义为 $\xi = E[e^2(n)]$。式（3.54）中，$y(n) = \boldsymbol{x}^{\mathrm{T}}(n)\boldsymbol{w}(n)$。为了使均方误差达到最小，使用梯度下降法在性能表面上搜索，并在梯度搜索中利用误差信号 $e(n)$ 的瞬时值替代其均方值 $E[e^2(n)]$，可以得到 LMS 算法权系数矢量的递推迭代公式为

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + 2\mu e(n)\boldsymbol{x}(n) \tag{3.55}$$

通常，只要满足迭代步长 μ 的收敛条件，LMS 自适应滤波器就能够收敛到最优维纳解。

2．最小平均 *p* 范数（LMP）自适应滤波

在 Alpha 稳定分布条件下，LMS 算法性能退化。采用最小分散系数（MD）准则替代 MMSE 准则，并用误差信号 $e(n)$ 的 p 范数取代均方误差函数。由分数低阶矩理论，只要满足 $0 < p < \alpha$，$S\alpha S$ 过程的 p 范数等价于其 p 阶矩。这样，构建最小平均 p 范数（least mean p norm，LMP）自适应滤波器，其代价函数写为

$$J = E[|e(n)|^p] \tag{3.56}$$

利用梯度技术，并以误差信号的瞬时值替代其统计平均，得到梯度估计为

$$\hat{\nabla}(n) = \frac{\partial J}{\partial \boldsymbol{w}(n)} = p|e(n)|^{p-1}\operatorname{sgn}[e(n)][-\boldsymbol{x}(n)], \quad 1 \leqslant p < \alpha \leqslant 2 \tag{3.57}$$

这样，可以得到 LMP 自适应滤波算法为

$$\begin{gathered} y(n) = \boldsymbol{w}^{\mathrm{T}}(n)\boldsymbol{x}(n) \\ e(n) = d(n) - y(n) \\ \boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu p\boldsymbol{x}(n)|e(n)|^{p-1}\operatorname{sgn}[e(n)] \end{gathered} \tag{3.58}$$

若取 $p = 1$，则 LMP 算法就退化成极性自适应滤波算法，命名为最小平均绝对偏差（LMAD）算法。其迭代表达式为

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu p\boldsymbol{x}(n)\operatorname{sgn}[e(n)] \tag{3.59}$$

与经典的 LMS 算法相比，LMP 算法对于脉冲噪声有较强的韧性。但是在强脉冲噪声情况下，LMP 算法并不理想。为了提高算法的稳定性和收敛速度，文献中又提出了针对 LMP 和 LMAD 算法的归一化方法，分别称为 NLMP 和 NLMAD 算法。二者的自适应迭代公式分别由式（3.60）和式（3.61）给出

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu p\boldsymbol{x}(n)\frac{|e(n)|^{p-1}\operatorname{sgn}[e(n)]}{\|x(n)\|_p^p + \lambda} \tag{3.60}$$

$$\boldsymbol{w}(n+1) = \boldsymbol{w}(n) + \mu p\boldsymbol{x}(n)\frac{\operatorname{sgn}[e(n)]}{\|\boldsymbol{x}(n)\|_1 + \lambda} \tag{3.61}$$

式中，λ 是为了避免出现分母为零而补充的一个较小的常量。

LMP 算法除了 FIR 结构，还存在对应的格形结构算法。近年来，LMP 算法已经广泛应用于非高斯自适应信号处理的各个应用领域。

3．自适应滤波器的计算机仿真

本节以 LMP 算法和 LMS 算法为例进行计算机仿真，使读者进一步体会以分数低阶统计量理论为基础的 FIR 线性自适应滤波的特性。

例 3.4　试设计一个 10 阶 FIR 自适应滤波器，期望信号是一个相对带宽为 0.1 的低通信号，利用 MATLAB 编程实现 LMP 和 LMS 自适应滤波器，并观察两种滤波器的性能。

解　MATLAB 程序代码如下：

```
    clc; clear; close all;
    L = 500; N =10; alpha = 1.5;        % 序列长度 L，滤波器阶数 N，Alpha=1.5
    p = alpha-2;  % LMP 算法的 p 参数
    tem = randn(1,L); [b,a] = butter(3,0.2); s = filter(b,a,tem); clear tem b a;
%产生期望信号
    v = (genta(alpha,L))’;              % 产生 Alpha 稳定分布噪声
    s = s/std(s); x = s+0.5*v;          % 产生观测信号并调整信号强度
    w_lmp(1:N+1,1:L) = 0; w_lms(1:N+1,1:L) = 0; mu = 0.003; LMP 和 LMS 滤波器设置
    for n = N+1:L-1
        y_lmp(n) = x(n-N:n)* w_lmp(:,n); err_lmp(n)=s(n)-y_lmp(n);  % LMP 滤波
        w_lmp(:,n+1) = w_lmp(:,n)+mu*abs(err_lmp(n))^(p-1)*sign(err_lmp(n)* x(n-N:n)’;
        y_lms(n)= x(n-N:n)* w_lms(:,n); err_lms(n)=s(n)-y_lms(n);  % LMS 滤波
        w_lms(:,n+1) = w_lms(:,n)+mu* err_lms(n)*x(n-N:n)’;
    end
    figure,
    subplot(2,2,1); plot(w_lmp’,’linewidth’,2); grid on;
    xlabel(‘(a) 样本下标’); ylabel(‘LMP 权系数幅值’);
    subplot(2,2,2); plot(w_lms’,’linewidth’,2); grid on;
    xlabel(‘(a) 样本下标’); ylabel(‘LMS 权系数幅值’);
    subplot(2,2,3); plot(s,’r’,’linewidth’,2); hold on; plot(x,’linewidth’,2);
    plot(y_lmp,’k’,’linewidth’,2); grid on;
    legend(‘理想信号’,’观察信号’,’lmp 滤波’,’Orientation’,’horizontal’);
    xlabel(‘(c) 样本下标’); ylabel(‘信号幅值’);
    subplot(2,2,4); plot(s,’r’,’linewidth’,2); hold on; plot(x,’linewidth’,2);
    plot(y_lms,’m’,’linewidth’,2); grid on;
    legend(‘理想信号’,’观察信号’,’lms 滤波’,’Orientation’,’horizontal’);
    xlabel(‘(d) 样本下标’); ylabel(‘信号幅值’);
```

通过运行例 3.4 的程序，可以观察到，在 Alpha 稳定分布噪声条件下，LMS 算法的性能受到较大影响，而 LMP 算法受到的影响较小，能够较好地收敛。

3.5.2　递推最小平均 *p* 范数自适应算法

递推最小平均 p 范数（简称为 RLMP）自适应算法的思路是，采用滑动窗方法来自适应调整滤波器的系数，即在一个大小为 L 的窗口内使得误差 $e(n)$ 的平均 p 范数最小。RLMP 自适应滤波器的代价函数为

$$J(n)=\sum_{k=n-L+1}^{n}|e(k)|^{p}=\sum_{k=n-L+1}^{n}|d(k)-\boldsymbol{w}^{\mathrm{T}}(n)\boldsymbol{x}(k)|^{p},\quad 0\leqslant p<2 \tag{3.62}$$

式中，$\boldsymbol{x}(k)$ 表示输入信号最近的 L 个样本。求取 $J(n)$ 相对于权系数矢量 $\boldsymbol{w}(n)$ 的梯度并令其为零，可以得到

$$\sum_{k=n-L+1}^{n}u(k)\boldsymbol{x}(k)\boldsymbol{x}^{\mathrm{T}}(k)\boldsymbol{w}(n)=\sum_{k=n-L+1}^{n}u(k)d(k)\boldsymbol{x}(k) \tag{3.63}$$

式中，$u(k)=|e(k)|^{p-2}$。进一步定义 $\boldsymbol{r}(n)=\sum_{k=n-L+1}^{n}u(k)d(k)\boldsymbol{x}(k)$ 和 $\boldsymbol{R}(n)=\sum_{k=n-L+1}^{n}u(k)\boldsymbol{x}(k)\boldsymbol{x}^{\mathrm{T}}(k)$，可得权矢量表达式为

$$\boldsymbol{w}^{*}(n)=\boldsymbol{R}^{-1}(n)\boldsymbol{r}(n) \tag{3.64}$$

上式是通过在大小为 L 的窗口内求误差 p 范数最小化得到的。注意到，由于 $\boldsymbol{R}(n)$ 和 $\boldsymbol{r}(n)$ 都是 $\boldsymbol{w}^*(n)$ 的函数，因此 $\boldsymbol{w}^*(n)$ 不能由式（3.64）直接求解得到。文献中给出了迭代求解 $\boldsymbol{w}^*(n)$ 的方法，其基本思路是通过逐步迭代求得每一时刻的最优权矢量，并以上一时刻的最优权矢量作为下一时刻权矢量的初值。

3.6　Alpha 稳定分布噪声下的核自适应滤波

3.6.1　核方法的概念

核方法（kernel method，KM）是近年来得到广泛重视和快速发展的一类模式识别算法，是解决非线性模式分析问题的一种有效途径。核方法的核心思想是：通过某种非线性映射将原始数据映射到适当的高维特征空间，再利用通用的线性学习器在这个新的空间对数据进行分析和处理。

与以往的非线性信号处理方法不同，核方法处理非线性问题具有坚实的数学基础，且成功应用于诸如支持向量机（SVM）、核主分量分析（KPCA）、核 Fisher 判别分析（KFDA）等。特别地，核方法在非线性自适应信号处理领域开辟了一个新领域。

核方法的理论可以追溯到 1909 年 Mercer 定理的提出。在 Mercer 定理的基础上，Aronszjn 于 1950 年提出了再生核希尔伯特空间（RKHS）的概念和理论。1995 年，Cortes 等人依据 RKHS 提出了非线性 SVM 并将其应用于手写字符识别问题。之后，核方法又被成功推广到 Fisher 判别分析和核主分量分析等领域。

近年来，核方法与自适应滤波技术深度结合，开创性地形成了一个核自适应滤波（kernel adaptive filtering）领域。其中，最早将核方法用于自适应滤波的是 Frieb 等人，他们于 1999 年提出了核 Adline 算法，把 Adline 推广到非线性领域。2008 年，Principe 团队的 Liu 等人提出了核最小均方（kernel LMS，KLMS）自适应滤波算法，得到广泛的重视，并在非线性信道均衡、信道辨识和信号预测等方面有很好的效果。之后，文献中报道了诸多 KLMS 算法的改进算法，包括归一化核 LMS 算法（NKLMS）、多核 LMS 算法（MKLMS）、量化核 LMS 算法（QKLMS）等。另一方面，Liu 等人提出了核仿射投影算法（kernel affine projection，KAP），具有更好的效果。Engel 等人则把经典的递归最小二乘（RLS）算法推广到非线性领域，提出了核递归最小二乘算法（kernel recursive least squares，KRLS），并引领了后续的改进和扩展。针对 Alpha 稳定分布噪声环境下的自适应滤波问题，赵知劲和 Gao 等人分别提出了核最小平均 p 范数（kernel LMP，KLMP）自适应滤波算法，成为适用于 Alpha 稳定分布噪声和非线性并存条件的一种有效的自适应滤波方法。

3.6.2　核自适应滤波

1．线性自适应滤波

自适应滤波或自适应滤波器是信号处理中的一个非常重要的分支，自 1959 年 Widrow 提出自适应滤波的概念和方法以来，自适应滤波和自适应滤波器的理论和方法一直受到广泛的重视，并得到不断的发展和完善，经久不衰。

经典的自适应滤波器由线性组合器构成，属于线性滤波器。最常见的滤波器单元，是由

抽头延迟线构成的横向滤波器结构。滤波器的单位冲激响应是一组可调参数的权系数矢量，记为 $\boldsymbol{w}(n)$ 。n 时刻的输入信号矢量记为 $\boldsymbol{x}(n)$ ，滤波器的输出信号 $y(n)$ 与外部输入的期待响应 $d(n)$ 进行比较，产生误差信号 $e(n)$ ，系统利用这个误差信号对权系数矢量进行调整，使得滤波器得到更新。

自适应滤波器常以均方误差（MMSE）最小作为系统的代价函数，即

$$J(n)=E[e^2(n)],\quad n=1,2,\cdots \tag{3.65}$$

经典的最小均方（LMS）算法和递归最小二乘（RLS）算法都是典型的线性自适应滤波算法。

尽管线性自适应滤波器结构简单、易于实现，但是其应用却非常广泛。其强大的信号处理能力使其在噪声抵消、回声对消、信道估计与均衡、阵列自适应波束形成等应用中取得显著成果，并进一步在诸如通信、控制、雷达、声呐、地震信号处理及生物医学工程等多个领域得到广泛的应用。可以说，线性自适应滤波器理论已经发展到一个高度成熟的阶段。

2．非线性自适应滤波

在自然界和工程技术中存在大量的非线性问题，很难依据经典的线性自适应滤波器来解决。这样，非线性自适应滤波理论与技术应运而生。最早的非线性自适应滤波器是通过线性滤波器的串联来实现的，但是这种方法存在三个主要问题，一是其建模能力有限，二是其模型不能独立于问题，三是训练过程中存在局部最优的弊端。这些问题的存在，影响了非线性自适应滤波器的应用和推广。

现有的非线性自适应滤波器中，基于 Volterra 级数展开的非线性自适应滤波器受到广泛重视。理论分析与实践表明，由于这种非线性自适应滤波器综合利用了展开的线性和非线性项，可以很好地逼近非线性过程。然而，Volterra 非线性自适应滤波器的计算复杂度非常高，影响了这种滤波器的进一步推广应用。随后出现的时间抽头多层感知器、径向基函数网络和并发神经网络等，均利用随机梯度法进行实时训练。尽管这些方法也曾获得了一定的成功，但是这些方法本质上均属于非凸优化，限制了其在线应用和进一步发展。

式（3.66）给出了构建一个非线性自适应滤波器所需遵守的序贯学习规则

$$f(n)=f(n-1)+g(n)e(n) \tag{3.66}$$

式中，$f(n)$ 表示时刻 n 由输入信号 $x(n)$ 到输出信号 $y(n)$ 的映射关系，即 $y(n)=f[x(n)]$ ，$e(n)$ 是预测误差函数， $g(n)$ 是与新输入数据有关的增益函数。上式看似简单，但实际上，算法本身可以采用各种不同的目标函数来实现，由此可以产生多种不同的非线性自适应滤波算法。

3．核自适应滤波

核自适应滤波是一类新型的非线性自适应滤波方法，其基本原理是依据 Mercer 定理，采用核方法，通过非线性映射，把输入数据空间的非线性问题映射到高维特征空间，称为再生核希尔伯特空间（RKHS），从而转化为线性问题。然后，再在特征空间使用线性方法进行信号处理。实际上，只要算法可以表示成内积的形式，利用再生核希尔伯特空间的性质和核技巧，就无须直接在高维空间进行计算，使得高维空间的自适应滤波变得非常简单。在这个过程中，RKHS 起到非常关键的作用，它为核自适应滤波提供了非常好的条件，包括线性特性、凸性和通用的逼近能力。有关 RKHS 的详细介绍，可参见本书 1.5 节的内容或参考有关文献。

概括起来，核自适应滤波属于非线性自适应滤波的范畴，但实际上，它又是一种广义的线性自适应滤波。核自适应滤波的核心思想可以概括为两个方面：①使用 Mercer 核将输入数

据空间非线性映射到高维特征空间；②使用线性自适应滤波器进行自适应信号处理。这样，对于诸如自适应均衡、预测、控制、建模和系统辨识等应用中在线、小规模的非线性自适应信号处理，都可以采用最经典的线性自适应滤波，如 LMS 算法或 RLS 算法来完成。

3.6.3　核最小平均 *p* 范数自适应滤波

核最小平均 p 范数（KLMP）自适应滤波是依据核自适应滤波方法对最小平均 p 范数（LMP）自适应滤波进行改造而成的新型自适应滤波方法。其主要特点是适用于输入-输出的非线性映射关系，并能够在 Alpha 稳定分布噪声条件下保持良好的鲁棒性。本节在简要介绍核 LMS（KLMS）算法的基础上，给出对 KLMP 的介绍。

（1）KLMS 算法

LMS 算法是自适应信号处理领域最经典并应用广泛的线性自适应滤波算法。通常，LMS 算法假设其由线性有限脉冲响应（FIR）滤波器构成。若滤波器的期待响应 $d(n)$ 与输入信号 $\boldsymbol{x}(n)$ 之间的映射是高度非线性的，则 LMS 算法的性能会急剧退化。为了改善 LMS 算法在非线性条件下的性能，Principe 团队的 Liu 等人提出了核 LMS 自适应滤波算法，简称为 KLMS 算法。

为了便于对比，现将 LMS 自适应滤波算法的主要公式列于式（3.67）中

$$\begin{cases} \boldsymbol{w}(0)=0 \\ e(n)=d(n)-\boldsymbol{w}^{\mathrm{T}}(n-1)\boldsymbol{x}(n) \\ \boldsymbol{w}(n)=\boldsymbol{w}(n-1)+\mu e(n)\boldsymbol{x}(n) \end{cases} \tag{3.67}$$

式中，$\boldsymbol{w}(n)$ 和 $\boldsymbol{x}(n)$ 分别表示自适应滤波器的权矢量和输入信号矢量，$d(n)$ 为期待响应，$e(n)$ 为误差信号，μ 为自适应滤波器的收敛因子。

采用核方法，把原输入空间 $\mathbb{U}$ 中的输入信号 $\boldsymbol{x}(n)$ 通过非线性映射函数 $\boldsymbol{\varphi}$ 映射到高维 RKHS 空间 $\mathbb{H}$，形成 $\boldsymbol{\varphi}(\boldsymbol{x}(n))$，简记为 $\boldsymbol{\varphi}(n)$。由于 $\boldsymbol{x}(n)$ 与 $\boldsymbol{\varphi}(n)$ 的维度差异，$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}(n))$ 是比 $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{x}$ 更强大的模型。进一步地，通过随机梯度下降法寻找权矢量 $\boldsymbol{w}(n)$ 的最优值，是实现非线性滤波的基本方法。在 RKHS 空间 $\mathbb{H}$ 中的新样本序列 $\{\boldsymbol{\varphi}(n), d(n)\}$ 上使用 LMS 算法，可以得到

$$\begin{cases} \boldsymbol{w}(0)=0 \\ e(n)=d(n)-\boldsymbol{w}^{\mathrm{T}}(n-1)\boldsymbol{\varphi}(n) \\ \boldsymbol{w}(n)=\boldsymbol{w}(n-1)+\mu e(n)\boldsymbol{\varphi}(n) \end{cases} \tag{3.68}$$

显然，式（3.68）与式（3.67）具有高度相似性。不过，前者是在 RKHS 空间 $\mathbb{H}$ 中的运算，而后者是在输入信号空间 $\mathbb{U}$ 的运算。

由于 $\boldsymbol{\varphi}$ 的维度很高，且为隐式表达，不方便计算。若对式（3.68）中的权矢量 $\boldsymbol{w}(n)$ 进行反复迭代，则可得

$$\begin{aligned} \boldsymbol{w}(n) &= \boldsymbol{w}(n-1)+\mu e(n)\boldsymbol{\varphi}(n) \\ &= \left[\boldsymbol{w}(n-2)+\mu e(n-1)\boldsymbol{\varphi}(n-1)\right]+\mu e(n)\boldsymbol{\varphi}(n) \\ &= \boldsymbol{w}(n-2)+\mu\left[e(n-1)\boldsymbol{\varphi}(n-1)+e(n)\boldsymbol{\varphi}(n)\right] \\ &\cdots \\ &= \boldsymbol{w}(0)+\mu\sum_{m=1}^{n} e(m)\boldsymbol{\varphi}(m) \end{aligned} \tag{3.69}$$

若 $\boldsymbol{w}(0)=0$，则 $\boldsymbol{w}(n)=\mu\sum_{m=1}^{n}e(m)\boldsymbol{\varphi}(m)$。这表明，经过 n 步训练，权矢量的估计值表示为非线性映射后所有数据的线性组合形式。并且，该自适应系统对于一个新输入 $\boldsymbol{x}'$ 的输出，可以表示成映射后输入信号的内积形式，即

$$\boldsymbol{w}^{\mathrm{T}}(n)\boldsymbol{\varphi}(\boldsymbol{x}')=\left[\mu\sum_{m=1}^{n}e(m)\boldsymbol{\varphi}(\boldsymbol{x}(m))^{\mathrm{T}}\right]\boldsymbol{\varphi}(\boldsymbol{x}')=\mu\sum_{m=1}^{n}e(m)\left[\boldsymbol{\varphi}(\boldsymbol{x}(m))^{\mathrm{T}}\boldsymbol{\varphi}(\boldsymbol{x}')\right]$$

再通过核技巧 $\boldsymbol{\varphi}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{\varphi}(\boldsymbol{x}')=\kappa(\boldsymbol{x},\boldsymbol{x}')$，可以在输入空间 $\mathbb{U}$ 通过核函数非常高效地计算滤波器的输出。

$$\boldsymbol{w}^{\mathrm{T}}(n)\boldsymbol{\varphi}(\boldsymbol{x}')=\mu\sum_{m=1}^{n}e(m)\kappa[\boldsymbol{x}(m),\boldsymbol{x}'] \tag{3.70}$$

显然，上述新算法与 LMS 算法相比，滤波器的输出仅需要内积运算，不涉及权矢量的迭代，计算量显著减小。KLMS 算法完整的序贯学习规则为

$$\begin{cases} f_{n-1}=\mu\sum\limits_{m=1}^{n-1}e(m)\kappa[\boldsymbol{x}(m),\cdot] \\ f_{n-1}(\boldsymbol{x}(n))=\mu\sum\limits_{m=1}^{n-1}e(m)\kappa[\boldsymbol{x}(m),\boldsymbol{x}(n)] \\ e(n)=d(n)-f_{n-1}(\boldsymbol{x}(n)) \\ f_n=f_{n-1}+\mu e(n)\kappa[\boldsymbol{x}(n),\cdot] \end{cases} \tag{3.71}$$

式中，f_n 表示 n 时刻输入-输出非线性映射的估计。KLMS 是 RKHS 空间的 LMS 算法。

（2）KLMP 算法

由于 KLMS 算法不能适应 Alpha 稳定分布条件下的非线性滤波问题，赵知劲和 Gao 等人分别提出了基于最小平均 p 范数的核自适应滤波方法，统称为 KLMP 自适应滤波算法。该算法基于核方法，对 LMP 自适应滤波算法进行改造，可以有效地抑制脉冲噪声的影响，并适用于非线性条件下的自适应滤波。

与 LMS 算法相似，LMP 算法（见本书 3.5 节的介绍）的主要计算式如下：

$$\begin{cases} \boldsymbol{w}(0)=0 \\ e(n)=d(n)-\boldsymbol{w}^{\mathrm{T}}(n-1)\boldsymbol{x}(n) \\ \boldsymbol{w}(n)=\boldsymbol{w}(n-1)+\mu p\,|e(n)|^{p-1}\operatorname{sgn}[e(n)]\boldsymbol{x}(n), \quad 1\leqslant p<\alpha \end{cases} \tag{3.72}$$

式中，$\alpha\in(0,2]$ 为 Alpha 稳定分布的特征指数，p 表示自适应滤波器的阶数，其余参数与 LMS 算法的说明一致。

与 KLMS 算法的推导相似，在非线性条件下，依据核方法，把原输入空间 $\mathbb{U}$ 中的输入信号 $\boldsymbol{x}(n)$ 通过非线性映射函数 $\boldsymbol{\varphi}$ 映射到高维 RKHS 空间 $\mathbb{H}$ 中，形成 $\boldsymbol{\varphi}(\boldsymbol{x}(n))$，同样简记为 $\boldsymbol{\varphi}(n)$。在 RKHS 空间 $\mathbb{H}$ 中的新样本序列 $\{\boldsymbol{\varphi}(n),d(n)\}$ 上使用 LMP 算法，可以得到

$$\begin{cases} \boldsymbol{w}(0)=0 \\ e(n)=d(n)-\boldsymbol{w}^{\mathrm{T}}(n-1)\boldsymbol{\varphi}(n) \\ \boldsymbol{w}(n)=\boldsymbol{w}(n-1)+\mu p\,|e(n)|^{p-1}\operatorname{sgn}[e(n)]\boldsymbol{\varphi}(n), \quad 1\leqslant p<\alpha \end{cases} \tag{3.73}$$

对式（3.73）中的权矢量 $\boldsymbol{w}(n)$ 进行反复迭代，可以得到

$$\boldsymbol{w}(n)=\mu p\sum_{m=1}^{n}|e(m)|^{p-1}\operatorname{sgn}[e(m)]\boldsymbol{\varphi}(m) \tag{3.74}$$

式中假设了 $\boldsymbol{w}(0)=0$。滤波器输出 $y(n)$ 可表示为

$$y(n)=\boldsymbol{w}^{\mathrm{T}}(n-1)\boldsymbol{\varphi}(n)=\mu p\sum_{m=1}^{n-1}|e(m)|^{p-1}\operatorname{sgn}[e(m)]\boldsymbol{\varphi}^{\mathrm{T}}(m)\boldsymbol{\varphi}(n) \tag{3.75}$$

依据 Mercer 核函数简化内积计算规则 $\boldsymbol{\varphi}^{\mathrm{T}}(\boldsymbol{x})\boldsymbol{\varphi}(\boldsymbol{x}')=\kappa(\boldsymbol{x},\boldsymbol{x}')$，可得 KLMP 算法为

$$y(n)=\mu p\sum_{m=1}^{n-1}|e(m)|^{p-1}\operatorname{sgn}[e(m)]\kappa[\boldsymbol{x}(m),\boldsymbol{x}(n)] \tag{3.76}$$

KLMP 算法是 RKHS 空间 $\mathbb{H}$ 的 LMP 算法。由式（3.76）可以看出，KLMP 算法的本质是核函数与误差函数的运算。从输入数据空间 $\mathbb{U}$ 到再生核希尔伯特空间 $\mathbb{H}$ 的非线性映射和在 $\mathbb{H}$ 空间的处理，都隐含在核函数 $\boldsymbol{\kappa}[\boldsymbol{x}(m),\boldsymbol{x}(n)]$ 和误差信号 $e(n)$ 中。当然，KLMP 算法的关键环节和基本原理是所谓的“核技巧”。

例 3.5　试在 Alpha 稳定分布噪声条件下进行 KLMP 算法的仿真，并与 KLMS 算法进行对比。

解　设定仿真条件：输入信号为独立同分布高斯随机信号，均值为 0，方差为 0.25；Alpha 稳定分布噪声的特征指数为 $\alpha=1.5$，分数低阶矩指数取 $p=1.3$；核长参数取 $\sigma=0.35$，收敛因子为 $\mu=0.01$。采用 KLMS 算法与 KLMP 算法进行对比。

图 3.4 给出了两种算法仿真结果的学习曲线。

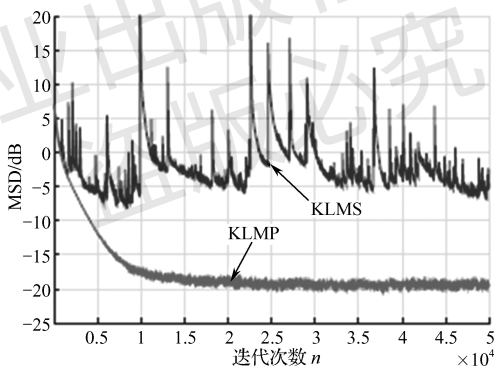


图 3.4　KLMP 算法与 KLMS 算法的仿真结果

图 3.4 中，横轴表示迭代次数，纵轴表示误差的均方差（mean square deviation，MSD）。显然，在 Alpha 稳定分布噪声条件下，KLMP 收敛快速且稳定，而 KLMS 算法则由于 Alpha 稳定分布噪声的影响，产生较大的误差。

3.7　分数低阶统计量信号处理的应用

3.7.1　在时间延迟估计中的应用

1．时间延迟估计的概念与用途

时间延迟（简称为时延），一般指接收器阵列中不同接收器所接收到的同源带噪信号之

间由于信号传输距离不同而引起的时间差。时间延迟估计（time delay estimation，TDE）一般是指利用参数估计和信号处理方法，对上述时间延迟进行估计和测定，并由此进一步确定其他有关参数，如目标信源的距离、方位、运动的方向和速度等。在无线电目标定位等领域，时间延迟估计也常称为到达时差（time difference of arrival，TDOA）估计。

根据目标信源和检测系统的不同，TDE/TDOA 问题大致可分为两种类型，即主动（active）时延估计和被动（passive）时延估计。对于前者，检测系统要发送信号；而对于后者，检测系统不发送信号，仅依据接收信号估计目标信源的参数。主动雷达和主动声呐是主动 TDE 的典型应用，而被动雷达和被动声呐是被动 TDE 的典型应用。在无线电监测的目标定位中，由于监测系统本身不发送信号，仅接收被监测目标发出的信号，因此属于被动 TDE 系统。

在被动 TDE 问题中，通常假定信号在信道中是以无色散球面波传播的。为了简化分析和处理，常假设信号源和接收器在同一平面内，从而把三维空间简化为二维空间。在二维空间中，若进一步地假定信源的距离远大于接收器阵列的几何尺寸，则可以认为信源发出的信号是以二维平面波的方式传播到接收器的。

目标定位是 TDE 或 TDOA 估计的主要应用之一。图 3.5 给出了基于 TDE 的目标信源被动定位示意图。

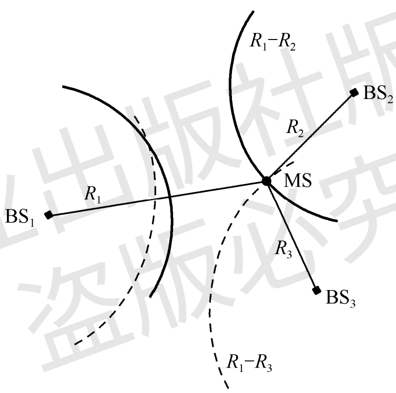


图 3.5　基于 TDE 的目标信源被动定位示意图

如图 3.5 所示，考虑二维平面内的目标定位问题。设待定位目标位于图中 MS 处，三个接收器 BS_1、BS_2 和 BS_3 组成被动 TDE 定位系统。设 MS 信号到达 BS_1 和 BS_2 的时间差为 D_{12}，而到达 BS_1 和 BS_3 的时间差为 D_{13}，通过估计这两个时间差，就可以进一步确定目标信源 MS 的坐标位置。几何定位方程如下：

$$\begin{cases} \sqrt{(x-x_1)^2+(y-y_1)^2}-\sqrt{(x-x_2)^2+(y-y_2)^2}=cD_{12} \\ \sqrt{(x-x_1)^2+(y-y_1)^2}-\sqrt{(x-x_3)^2+(y-y_3)^2}=cD_{13} \end{cases} \tag{3.77}$$

式中，c 表示信号传播速率，(x,y) 表示目标信源 MS 的坐标，(x_1,y_1)、(x_2,y_2) 和 (x_3,y_3) 分别表示已知的三个接收器 BS_1、BS_2 和 BS_3 的坐标。

2．经典的时间延迟估计方法

（1）基于相关分析的 TDE 方法

设两个接收器所接收信号的离散形式 $x_1(n)$ 和 $x_2(n)$ 分别为

$$\begin{cases} x_1(n) = s(n) + v_1(n) \\ x_2(n) = \lambda s(n-D) + v_2(n) \end{cases} \tag{3.78}$$

式中，$s(n)$ 为接收到的目标信源 MS 的纯净信号，$v_1(n)$ 和 $v_2(n)$ 分别表示两个接收器的接收噪声，λ 表示信号的相对衰减，常假定 $\lambda = 1$，D 表示待估计时间延迟。计算 $x_1(n)$ 和 $x_2(n)$ 的互相关函数，并利用信号与噪声的不相关特性，有

$$R_{x_1x_2}(\tau) = E\left[x_1(n)x_2(n+\tau)\right] = R_{ss}(\tau - D) \tag{3.79}$$

根据自相关函数的性质，可知 $R_{ss}(\tau - D)$ 在 $\tau - D = 0$ 处取得最大值，表明此刻两信号 $x_1(n)$ 和 $x_2(n)$ 具有最大相关性。故选择 $R_{ss}(\tau - D)$ 取得最大值处的 τ 值作为时间延迟的估计值，即

$$\hat{D} = \arg\max_{\tau}\left[R_{ss}(\tau - D)\right] \tag{3.80}$$

（2）广义相关 TDE 方法

针对相关分析方法存在的抗噪声和干扰能力不强的局限性，Carter 等人提出了广义相关 TDE 方法，简称为 GCC-TDE 方法。其基本思路是，对两接收信号的互功率谱密度函数进行加权滤波，一方面抑制噪声，另一方面进行白化处理，以获得更好的时间延迟估计。图 3.6 给出了 GCC-TDE 方法的原理框图。

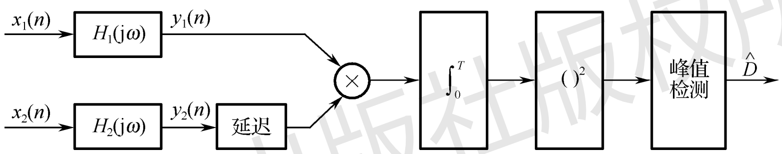


图 3.6　广义相关 TDE 方法原理框图

广义相关函数定义为加权后广义互功率谱函数的傅里叶逆变换

$$R^{\mathrm{G}}_{y_1y_2}(\tau) = F^{-1}[G_{y_1y_2}(\omega)] = F^{-1}[H(\mathrm{j}\omega)G_{x_1x_2}(\omega)] \tag{3.81}$$

式中，$H(\mathrm{j}\omega)$ 为广义相关加权函数。常用的广义相关加权函数有 Roth 处理器 $\dfrac{1}{G_{x_1x_2}(\omega)}$、相位变换 $\dfrac{1}{\left|G_{x_1x_2}(\omega)\right|}$ 和平滑相干变换 $\dfrac{1}{\sqrt{G_{x_1x_1}(\omega)G_{x_2x_2}(\omega)}}$ 等。

3．基于分数低阶统计量的 TDE 方法

上述经典 TDE 方法均是基于信号与噪声的二阶统计量的。若接收噪声服从 Alpha 稳定分布，则这些经典方法性的能显著退化。文献中报道了一类基于分数低阶统计量的 TDE 方法，用以抑制 Alpha 稳定分布噪声的影响，从而在复杂电磁环境中得到更好的 TDE 结果。

（1）基于第一类分数低阶相关的 TDE 方法

两接收信号 $x_1(n)$ 和 $x_2(n)$ 的第一类分数低阶相关为

$$\begin{aligned} R^{\mathrm{C}}(m) &\triangleq E\{x_2(n)[x_1(n+m)]^{\langle p-1\rangle}\} \\ &= E\{x_2(n)\,|\,x_1(n+m)\,|^{p-1}\,\mathrm{sgn}[x_1(n+m)]\},\quad 1 \leqslant p < \alpha \end{aligned} \tag{3.82}$$

在实际计算中，常采用有限时间平均来近似统计平均，则第一类分数低阶相关可以由样本估计得到

$$\hat{R}^{\mathrm{C}}(m) = \frac{1}{L_2 - L_1}\sum_{n=L_1+1}^{L_2} x_2(n)\,|\,x_1(n+m)\,|^{p-1}\,\mathrm{sgn}[x_1(n+m)],\quad 1 \leqslant p < \alpha \tag{3.83}$$

式中，$L_1=\max(0,-m)$，$L_2=\min(N-m,N)$。由此可以得到时间延迟估计为

$$\hat{D}=-\arg\max_m \hat{R}^{\mathrm{C}}(m) \tag{3.84}$$

由 p 的取值范围可见，该算法适用于$1<\alpha\leqslant 2$稳定分布的噪声环境。

（2）基于第二类分数低阶相关的 TDE 方法

两接收信号 $x_1(n)$ 和 $x_2(n)$ 的第二类分数低阶相关为

$$R^{\mathrm{D}}(m)\triangleq E\{[x_2(n)]^{\langle a\rangle}[x_1(n+m)]^{\langle b\rangle}\},\quad 0\leqslant a<\frac{\alpha}{2},\quad 0\leqslant b<\frac{\alpha}{2} \tag{3.85}$$

与第一类分数低阶相关算法相类似，用有限时间平均来近似统计平均，计算公式为

$$\hat{R}^{\mathrm{D}}(m)=\frac{1}{L_2-L_1}\sum_{n=L_1+1}^{L_2}|x_2(n)|^a|x_1(n+m)|^b\operatorname{sgn}(x_2(n)x_1(n+m)) \tag{3.86}$$

由此可以得到时间延迟估计为

$$\hat{D}=-\arg\max_m \hat{R}^{\mathrm{D}}(m) \tag{3.87}$$

该算法的适用范围为$0<\alpha\leqslant 2$。

（3）自适应最小平均 p 范数 TDE 方法

依据最小分散系数（MD）准则，参照经典的 LMS 自适应滤波算法，可以得到自适应最小平均 p 范数（LMP）TDE 方法。定义代价函数

$$J_p(\boldsymbol{w})\triangleq E\left[\left|x_2(n)-\sum_{m=-M}^{M}w(m)x_1(n+m)\right|^p\right],\quad 1<p<\alpha \tag{3.88}$$

式中，$m\in[-M,M]$表示滤波器权矢量的阶数。LMP 算法的自适应迭代公式为

$$\boldsymbol{w}(n+1)=\boldsymbol{w}(n)-\mu p|e(n)|^{p-1}\operatorname{sgn}[e(n)]\boldsymbol{x}_1(n) \tag{3.89}$$

要估计的时间延迟为

$$\hat{D}=\arg[\max_m(\boldsymbol{w})],\quad m=-M,-M+1,\cdots,0,\cdots,M \tag{3.90}$$

在 LMP 算法中，p 值须满足 $p\in[1,\alpha)$，即 LMP 算法只能应用于$1<\alpha\leqslant 2$的条件。

（4）基于分数低阶矩的 ETDE 方法

直接时间延迟估计（explicit time delay estimation，ETDE）算法是 So 等人提出的一种自适应 TDE 方法，其特点是在自适应迭代过程中，直接对时间延迟估计值 $\hat{D}(n)$ 进行递推估计，无须插值即可直接得到非整数时间延迟估计。ETDE 的迭代计算公式为

$$\hat{D}(n+1)=\hat{D}(n)-2\mu e(n)\sum_{k=-M}^{M}x(n-k)f[k-\hat{D}(n)] \tag{3.91}$$

式中，$f(v)=\dfrac{\cos(\pi v)-\operatorname{sinc}(v)}{v}$，$v=k-\hat{D}(n)$。

基于分数低阶矩改进得到的 LETDE 自适应迭代公式为

$$\hat{D}(n+1)=\hat{D}(n)-\mu p|e(n)|^{p-1}\operatorname{sgn}[e(n)]\sum_{k=-M}^{M}x(n-k)f[k-\hat{D}(n)] \tag{3.92}$$

比较式（3.92）与式（3.91），可见 LETDE 是具有时变步长 $\mu(n)$ 的 ETDE 算法，即

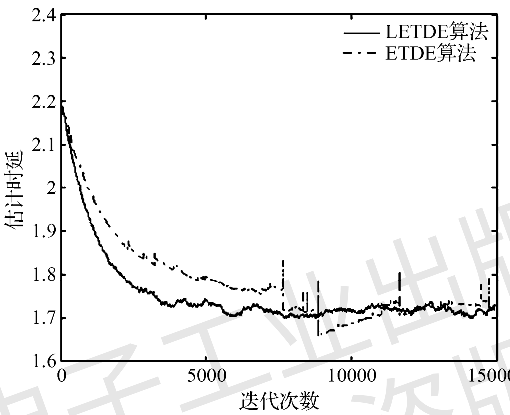
$$\mu(n)=\frac{\mu p|e(n)|^{p-1}}{2} \tag{3.93}$$

当输入信号中伴有强脉冲噪声时，采用归一化方法，并补充参数 λ，式（3.93）可写成

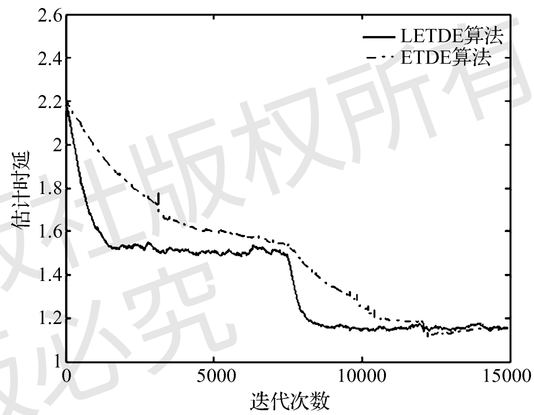
$$\hat{D}(n+1)=\hat{D}(n)-\mu p\left|e(n)\right|^{p-1}\operatorname{sgn}[e(n)]\frac{\sum_{k=-M}^{M}x(n-k)f[k-\hat{D}(n)]}{\left\|\sum_{k=-M}^{M}x(n-k)f[k-\hat{D}(n)]\right\|_p^p+\lambda} \tag{3.94}$$

例 3.6　试利用 MATLAB 编程实现对 LETDE 和 ETDE 两种算法计算机仿真，并分析。

解　设信号 $s(k)$ 为 0 均值高斯白噪声，延迟信号 $s(n-D)$ 由 $s(n)$ 经过单位冲激响应为 $\sum_{k=-M}^{M}\operatorname{sinc}(k-D)z^{-k}$ 的 31 阶 FIR 滤波器产生，噪声 $v_1(n)$ 和 $v_2(n)$ 为 Alpha 稳定分布过程。每次实验的数据长度为 15000 点，每次实验均是 500 次蒙特卡罗仿真的平均。图 3.7 给出了两种算法计算机仿真的结果。由图可见，在 Alpha 稳定分布噪声下，LETDE 算法比 ETDE 算法具有更高的估计精度和更好的跟踪能力。



（a）时延估计收敛曲线
（时延真值为1.7，$\alpha=1.5$，GSNR = 0dB）



（b）对突变时延的跟踪能力
（时延真值从1.5突变为1.2）

图 3.7　两种算法计算机仿真的结果

3.7.2　在波达方向估计中的应用

1．波达方向估计的概念与用途

波达方向（direction of arrival，DOA）估计，又称为波达角（angle of arrival，AOA）估计，其基本概念是依据阵列天线的指向作用，确定接收信号的来波方向，进而通过多组阵列天线的交叉作用，确定信源目标的位置信息。图 3.8 给出了 DOA 估计的示意图。如图所示，无线电接收阵列（或声学接收阵列）通过 DOA 估计算法，可以将其阵列波瓣指向对应目标的来波方向，从而实现目标的方位测定。

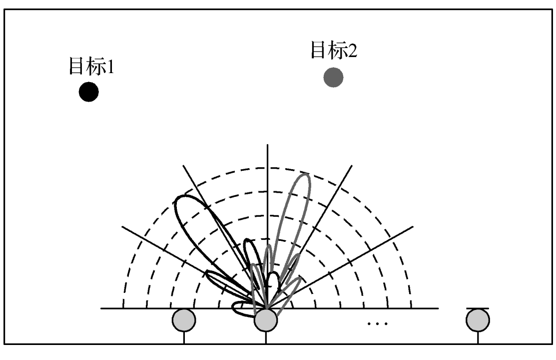


图 3.8　DOA 估计进行目标测向定位示意图

DOA 估计是阵列信号处理中的基本任务之一，是实现目标测向定位的关键技术，广泛

应用于雷达、声呐及无线电监测等多个领域。最典型 DOA 估计方法是以多重信号分类法（multiple signal classification，MUSIC）和旋转不变子空间（estimation of signal parameters via rotational invariance techniques，ESPRIT）为代表的基于特征子空间的超分辨率 DOA 估计方法及其改进和扩展算法。其中，MUSIC 算法的基本思想是对阵列输出数据的协方差矩阵进行子空间分解，利用信号子空间与噪声子空间的正交性构造空间谱，通过谱峰搜索获得 DOA 估计值。ESPRIT 方法同样是基于子空间理论的超分辨算法，该方法避免了 MUSIC 算法中的谱峰搜索，利用信号子空间的旋转不变特性获得入射信号的角度估计值，具有较高的计算速度和较低的计算复杂度。

2．基于 MUSIC 算法的 DOA 估计

（1）均匀线阵与信号模型

均匀线性阵列（uniform linear array，ULA，简称为均匀线阵）是接收阵列中最常见、结构最简单的一种形式，是指所有阵元（或称为传感器）均匀地分布在一直线上的阵列。图 3.9 给出了均匀线阵的示意图。

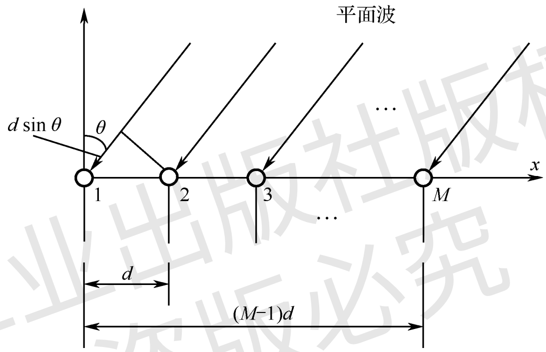


图 3.9　均匀线阵示意图

如图 3.9 所示，在 x 轴方向上等间距地布置 M 个接收阵元，相邻阵元之间的距离为 d，阵列孔径为 $(M-1)d$。通常将第一个阵元在 x 轴的位置定义为坐标原点和参考点。考虑已知中心频率 ω 和入射角度为 $\{\theta_1,\theta_2,\cdots,\theta_K\}$ 的 $K<M$ 个各向同性非相干窄带信号入射到均匀线性阵列。由于跨传感器的传播延迟远小于信号带宽的倒数，因此第 m 个阵元接收到的信号可表示为

$$x_m(t)=\sum_{k=1}^{K}a_m(\theta_k)s_k(t)+v_m(t),\quad m=1,2,\cdots,M \tag{3.95}$$

式中，$a_m(\theta_k)=\exp\left[-\mathrm{j}\dfrac{2\pi}{\lambda}(m-1)d\sin\theta_k\right]$为第 m 个阵元在 θ_k 方向的导向矢量，λ 为载波波长。$s_k(t)$ 是第 m 个阵元接收到的第 k 个信号，$v_m(t)$ 表示第 m 个阵元的接收噪声。

（2）经典 MUSIC 算法

Schmidt 提出的以正交子空间投影为基础的 MUSIC 算法，作为现代谱估计研究的里程碑之一，对 DOA 估计产生了深远的影响。作为一种超分辨率阵列信号处理技术，MUSIC 算法实现了空间谱估计向超分辨率的飞跃。

在均匀线性阵列条件下，基于 MUSIC 算法的 DOA 估计问题，可以归结为从 N 个快拍的数据中估计 K 个信源的波达方向 $\{\theta_1,\theta_2,\cdots,\theta_K\}$ 问题。将式（3.96）所示的信号模型改写为矢量形式，有

$$\boldsymbol{x}(t)=\boldsymbol{A}(\theta)\boldsymbol{s}(t)+\boldsymbol{v}(t) \tag{3.96}$$

式中，$\boldsymbol{x}(t)=[x_1(t), x_2(t), \cdots, x_M(t)]^{\mathrm{T}}$ 和 $\boldsymbol{v}(t)=[v_1(t), v_2(t), \cdots, v_M(t)]^{\mathrm{T}}$ 分别表示阵列输出信号矢量和加性观测噪声矢量，$\boldsymbol{s}(t)=[s_1(t), s_2(t), \cdots, s_K(t)]^{\mathrm{T}}$ 表示纯净信号（即 SOI 信号）矢量，$\boldsymbol{A}(\theta)=[\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_K)]^{\mathrm{T}}$ 表示 $M\times K$ 维导向矩阵。这样，阵列输出信号的协方差阵为

$$\boldsymbol{R}=E[\boldsymbol{x}(t)\boldsymbol{x}^{\mathrm{H}}(t)]=\boldsymbol{A}E[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{H}}(t)]\boldsymbol{A}^{\mathrm{H}}+\sigma^2\boldsymbol{I} \tag{3.97}$$

式中，$(\cdot)^{\mathrm{H}}$ 表示共轭转置，σ^2 为噪声方差。$\boldsymbol{R}$ 为正定 Hermitain 矩阵，对其进行特征分解，得

$$\boldsymbol{R}=\sum_{k=1}^{K}\rho_k\boldsymbol{u}_k\boldsymbol{u}_k^{\mathrm{H}}+\sigma^2\boldsymbol{I}=\sum_{m=1}^{M}\lambda_m\boldsymbol{u}_m\boldsymbol{u}_m^{\mathrm{H}} \tag{3.98}$$

式中，λ_m 和 $\boldsymbol{u}_m$ 分别表示协方差矩阵的特征值和特征矢量。可以证明，协方差矩阵 $\boldsymbol{R}$ 的前 K 个较大特征值对应的特征矢量（记为 $\boldsymbol{U}$）张成信号子空间，而其余较小特征值对应的特征矢量（记为 $\boldsymbol{G}$）张成噪声子空间。将方向矢量向噪声子空间投影，得到以 DOA 为参数的空间谱函数的 MUSIC 估计方法为

$$P(\theta)=\frac{1}{\boldsymbol{a}^{\mathrm{H}}(\theta)\boldsymbol{G}\boldsymbol{G}^{\mathrm{H}}\boldsymbol{a}(\theta)}=\frac{1}{\boldsymbol{a}^{\mathrm{H}}(\theta)(\boldsymbol{I}-\boldsymbol{U}\boldsymbol{U}^{\mathrm{H}})\boldsymbol{a}(\theta)} \tag{3.99}$$

对式（3.99）取峰值就能得到 L 个信源的 DOA 估计。

（3）基于分数低阶统计量的子空间方法与鲁棒性 MUSIC 波达方向估计

在 Alpha 稳定分布条件下，以 MUSIC 为代表的一类 DOA 估计算法均出现性能退化。针对这个问题，Tsakalides 和 Nikias 提出了基于共变的子空间概念，并提出了基于共变的鲁棒性 MUSIC（ROC-MUSIC）的 DOA 估计算法。

采用与经典 MUSIC 方法相同的均匀线性阵列和信号结构条件，假定信号中的加性噪声服从 $S\alpha S$ 分布。定义阵列输出信号 $\boldsymbol{x}(t)$ 的共变矩阵为 $\boldsymbol{R}^{(\mathrm{C})}$，其第 (i, j) 个元素表示为

$$R_{i,j}^{(\mathrm{C})}=[x_i(t),\ x_j(t)]_{\alpha} \tag{3.100}$$

可以证明，共变矩阵 $\boldsymbol{R}^{(\mathrm{C})}$ 与式（3.98）中的协方差矩阵 $\boldsymbol{R}$ 具有完全相似的结构形式，即

$$\boldsymbol{R}^{(\mathrm{C})}=[\boldsymbol{x}(t), \boldsymbol{x}(t)]_{\alpha}=\boldsymbol{A}(\theta)\boldsymbol{\Lambda}_s\boldsymbol{A}^{\langle\alpha-1\rangle}(\theta)+\gamma_v\boldsymbol{I} \tag{3.101}$$

式中，导向矩阵 $\boldsymbol{A}(\theta)=[\boldsymbol{a}(\theta_1), \boldsymbol{a}(\theta_2), \cdots, \boldsymbol{a}(\theta_K)]^{\mathrm{T}}$，对角阵 $\boldsymbol{\Lambda}_s=\mathrm{diag}[\gamma_{s_1}, \gamma_{s_2}, \cdots, \gamma_{s_K}]$，$\gamma_v$ 表示 $S\alpha S$ 噪声的分散系数，$\gamma_{s_k}=[s_k, s_k]_{\alpha}$，$\boldsymbol{I}$ 为单位阵，$z^{\langle\alpha-1\rangle}=|z|^{\alpha-1}\mathrm{sgn}(z)$。由导向矢量

$$\boldsymbol{a}(\theta_k)=\left[1,\ \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}d\sin\theta_k},\ \cdots,\ \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}(M-1)d\sin\theta_k}\right]^{\mathrm{T}}$$

可以得到 $[\boldsymbol{A}^{\langle\alpha-1\rangle}(\theta)]_{i,j}=[\boldsymbol{A}(\theta)]_{j,i}^{*}$。这样，式（3.101）可改写为

$$\boldsymbol{R}^{(\mathrm{C})}=[\boldsymbol{x}(t), \boldsymbol{x}(t)]_{\alpha}=\boldsymbol{A}(\theta)\boldsymbol{\Lambda}_s\boldsymbol{A}^{\mathrm{H}}(\theta)+\gamma_v\boldsymbol{I} \tag{3.102}$$

仿照经典 MUSIC 算法对共变矩阵 $\boldsymbol{R}^{(\mathrm{C})}$ 进行特征分解，可得基于共变的 MUSIC 算法（ROC-MUSIC）为

$$P_{\text{ROC-MUSIC}}(\theta)=\frac{1}{\sum_{m=K+1}^{M}|\boldsymbol{a}^{\mathrm{H}}(\theta)\boldsymbol{u}_m|^2} \tag{3.103}$$

式中，$\boldsymbol{u}_m$ 为共变矩阵的特征矢量。对式（3.103）进行峰值搜索，可得 K 个信源的 DOA 估计。

（4）其他基于分数低阶统计量的 MUSIC 波达方向估计方法

理论分析和实验验证均表明，共变并不是性能最优的分数低阶统计量，诸如分数低阶矩、

分数低阶协方差和相位分数低阶协方差等，在抑制 Alpha 稳定分布噪声方面，均可能具有更好的性能。

在 ROC-MUSIC 算法的启发下，文献中还报道了多种基于不同分数低阶统计量的改进的 MUSIC 算法以及对应的 DOA 估计方法，如基于分数低阶矩的 MUSIC（FLOM-MUSIC）算法、基于符号协方差矩阵的 MUSIC（SCM-MUSIC）算法和基于相位分数低阶矩的 MUSIC（PFLOM-MUSIC）等。相对于 ROC-MUSIC 算法，这些改进的分数低阶 MUSIC 类算法的性能均取得了一定程度的改善。

定义阵列输出信号 $\boldsymbol{x}(t)$ 的分数低阶矩矩阵为 $\boldsymbol{R}^{(\mathrm{F})}$，其第 (i, j) 个元素表示为

$$R_{i,j}^{(\mathrm{F})} = E[x_i(t)\left|x_j(t)\right|^{p-2} x_j^*(t)],\quad 1 < p < \alpha \tag{3.104}$$

已经证明，$\boldsymbol{R}^{(\mathrm{F})}$ 可以写成式（3.98）的形式，也可以进一步通过子空间分解，得到类似 MUSIC 算法的结果，记为 FLOM-MUSIC。

定义阵列输出信号 $\boldsymbol{x}(t)$ 的相位分数低阶矩矩阵为 $\boldsymbol{R}^{(\mathrm{P})}$，其第 (i, j) 个元素表示为

$$R_{i,j}^{(\mathrm{P})} = E[[x_i(t)]^{\langle p\rangle}[x_j(t)]^{-\langle p\rangle}],\quad 0 < p < \alpha/2 \tag{3.105}$$

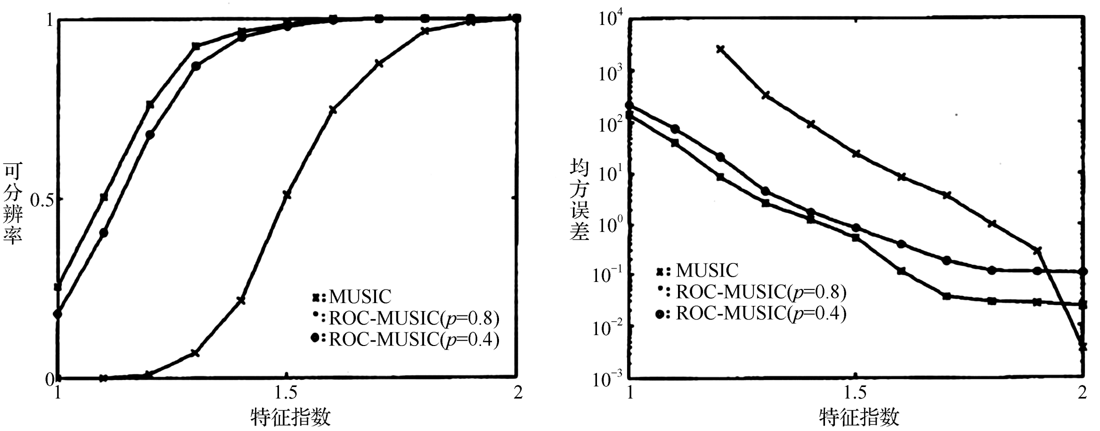
经过同样的处理方法，可以得到基于相位分数低阶矩的 MUSIC 算法，记为 PFLOM-MUSIC。

（5）DOA 估计的计算机仿真

例 3.7　试采用 MATLAB 编程，对经典 MUSIC 算法和 ROC-MUSIC 算法进行 DOA 估计的计算机仿真。并依据 DOA 估计的均方误差 MSE 和可分辨概率来对仿真结果进行评价。

解　定义可分辨概率为正确分辨两入射角次数 N_{res} 与 Monte Carlo 实验总次数 N_{MC} 之比，即 $P_{\mathrm{res}} = N_{\mathrm{res}}/N_{\mathrm{MC}}$。若满足 $P(\theta_m) - \frac{1}{2}\{P(\theta_1) + P(\theta_2)\} > 0$，则称这两个入射角是可分辨的。其中，设 θ_1 和 θ_2 为两个信号的入射角，$\theta_m = (\theta_1 + \theta_2)/2$，而 $P(\theta_m)$ 表示对应角度的 MUSIC 噪声空间谱的倒数。

设 2 个独立的具有相同功率的 QAM 通信信号入射到均匀线性阵列。阵列中包含 5 个阵元，各阵元的间隔为信号的半波长。QAM 信号的加性噪声服从 $S\alpha S$ 分布，其分散系数为 γ。对两种算法分别进行 200 次 Monte Carlo 仿真实验。图 3.10 给出了计算机仿真的结果。显然，ROC-MUSIC 比经典 MUSIC 具有更好的 DOA 估计性能。

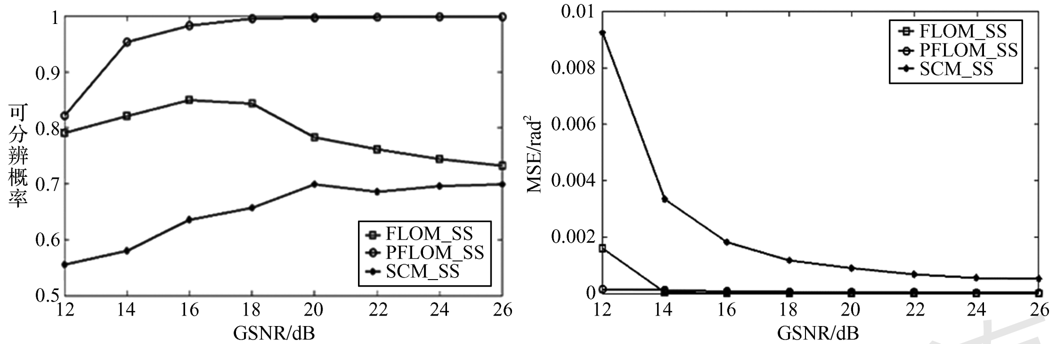


（a）可分辨概率　　（b）均方误差

图 3.10　MUSIC 类算法进行 DOA 估计计算机仿真的结果

例 3.8　试对几种分数低阶 MUSIC 算法进行 DOA 估计的对比。

解　采用 FLOM_SS（分数低阶矩）、PFLOM_SS（相位分数低阶矩）和 SCM_SS（符号协方差矩阵）三种分数低阶 DOA 估计算法，各算法名称的后缀 SS 表示算法进行了空间平滑运算。图 3.11 给出了这三种算法进行 DOA 估计的可分辨概率和均方误差随广义信噪比 GSNR 变化的情况。显然，基于相位分数低阶矩的 MUSIC 算法的 DOA 估计效果更好。



（a）可分辨概率　　（b）均方误差

图 3.11　几种分数低阶 MUSIC 算法进行 DOA 估计的对比

3.7.3　在信道盲均衡中的应用

1．信道均衡的概念

信道均衡（channel equalization）是通信技术中一种重要的技术，是指为了提高衰落信道条件下通信系统的传输性能而采取的一种抗衰落措施。其机理是对信道或整个传输系统特性进行补偿，以消除或减弱宽带通信时多径时延带来的码间串扰（ISI）问题。针对信道恒参或变参特性以及数据速率不同的情况，信道均衡包括多种结构方式，大致可分为线性均衡与非线性均衡两类。

2．经典的多用户恒模算法

多用户恒模算法（multiuser constant modulus algorithm，记为 MU_CMA）是一种受到广泛关注的信道均衡算法，适用于 MIMO 系统。这种算法在常规恒模算法（CAM）的基础上，增加了不同均衡器输出信号间的互相关部分，保证了各均衡器输出信号之间的不相关性，解决了一对多的问题。

MU_CMA 算法代价函数如下：

$$J_{\mathrm{MU\_CMA}}^{l}=E[(\left|y_l(k)\right|^2-R_2)^2]+K\sum_{m=1}^{l-1}\sum_{\delta=-(N+L)}^{N+L}\left|r_{m,\delta}(k)\right|^2 \tag{3.106}$$

式中，$r_{m,\delta}(k)=E[y_l(k)y_m^*(k-\delta)]$ 为第 l 个均衡器输出与第 m，$m=1,2,\cdots,l-1$ 个均衡器输出之间的互相关。考虑到所有可能的信道时延，δ 的取值范围为 $[-(N+L),\ N+L]$。$K\in\mathbf{R}^+$ 为混合参数。以梯度下降法最小化代价函数式(3.106)，可得算法的更新方程，并得到互相关 $r_{m,\delta}(k)$ 的递推表达式。

3．基于分数低阶统计量的多用户恒模算法

当信道中存在 Alpha 稳定分布噪声时，式（3.106）不能收敛，从而使算法的应用性能受

到显著影响。若使用分数低阶统计量替代式（3.106）中的二阶统计量，则分数低阶多用户恒模算法（简称为 FMU_CAM）的代价函数表示为

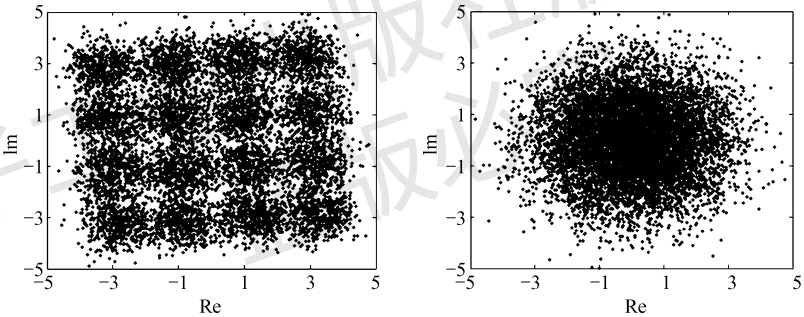
$$J_{\text{FMU\_CMA}}^{l}=E[(|y_l(k)|^p-R_p)^2]+K\sum_{m=1}^{l-1}\sum_{\delta=-(N+L)}^{N+L}\left|r_{m,\delta}^{p}\right|^2,\quad p<\alpha/2 \tag{3.107}$$

式中，p 表示分数低阶矩，$r_{m,\delta}^{p}=E[y_l(k)y_m^{\langle p-1\rangle}(k-\delta)]$ 为第 l 个均衡器输出与第 m（$m=1,2,\cdots,l-1$）个均衡器输出之间的分数低阶互相关。其余参数与式（3.106）相同。采用与 MU_CAM 算法相同的操作，可以得到分数低阶互相关 $r_{m,\delta}^{p}(k)$ 的递推表达式。

例 3.9 试利用 MATLAB 编程，分别仿真实现采用 MU_CAM 算法和 FMU_CAM 算法对 MIMO 系统信道的均衡。

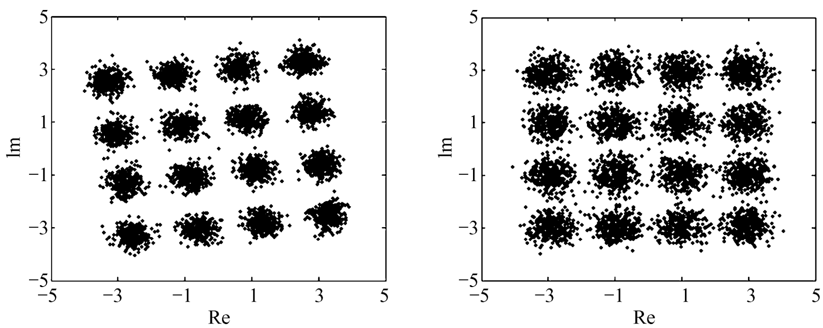
解 仿真中采用的 MIMO 系统为 2 输入 3 输出的 FIR 系统。源信号采用独立同分布的零均值 16-QAM 信号。Alpha 稳定分布噪声的特征指数为 $\alpha=1.85$，广义信噪比设定为 $\text{GSNR}=25\text{dB}$，分数低阶矩指数设定为 $p=0.875$。

图 3.12 和图 3.13 分别给出了采用 MU_CAM 和 FMU_CAM 算法进行信道均衡的结果。显然，MU_CAM 算法第一个均衡器的输出信号勉强可以分离开，而第二个均衡器的输出信号完全混叠在一起。FMU_CAM 算法的两个均衡器的输出信号很好地聚集在各自的星座点上。进一步得到的组合信道单位脉冲响应，表明 FMU_CAM 信道均衡器可以较好地分离源信号并很好地进行恢复，消除了码间干扰的影响。



（a）均衡器1的输出　　（b）均衡器2的输出

图 3.12　MU_CAM 算法的均衡器输出



（a）均衡器1的输出　　（b）均衡器2的输出

图 3.13　FMU_CAM 算法的均衡器输出

3.8　本章小结

针对传统的基于二阶统计量信号处理理论方法在 Alpha 稳定分布噪声下的性能退化问题，本章系统地介绍了分数低阶统计量的概念与基于分数低阶统计量的信号处理基本理论框架，主要包括分数低阶统计量的基本概念、性质和基本原理，Alpha 稳定分布的线性空间与线性理论，基于分数低阶统计量的线性参数模型与信号参数估计等。此外，结合分数低阶统计量的应用，介绍了多种 TDOA、DOA 和信道均衡的典型应用。为了配合后续章节关于相关熵和循环相关熵内容的介绍，本章还特别介绍了 Alpha 稳定分布条件下核方法与核自适应滤波的概念。总之，本章内容在理论和实际应用上为后续章节起到承上启下的作用。

思考题与习题

1．总结归纳随机信号统计量的分布情况。

2．简要说明什么是一阶、二阶统计量，什么是高阶统计量，什么是分数低阶统计量。

3．试解释分数低阶矩的概念，解释为什么分数低阶矩能抑制 Alpha 稳定分布噪声的影响。

4．说明零阶矩和负阶矩的概念。

5．解释几何功率的概念。为什么几何功率更适合评价脉冲噪声的功率？

6．解释共变的概念。如何计算共变与共变系数？

7．学习并掌握共变的主要性质。

8．什么是分数低阶相关？解释第一类和第二类分数低阶相关。

9．说明相位分数低阶矩算子的概念与作用。

10．试解释 Alpha 稳定分布的线性空间理论。

11．试解释 Alpha 稳定分布下最小距离测度、最小误差α范数、最小误差 p 阶矩和最小分散系数准则及其相互之间的联系。

12．解释基于共变的 Yule-Walker 方程的物理意义。

13．说明参数模型的最小 p 范数估计方法。

14．解释基于分数低阶统计量的自适应滤波技术。

15．什么是核方法？试说明再生核希尔伯特空间的概念。

16．说明核自适应滤波的概念与方法。

17．说明基于分数低阶统计量的核自适应滤波的概念与方法。

18．解释基于分数低阶统计量的时间延迟估计的概念与方法。

19．解释基于分数低阶统计量的波达方向估计的概念与方法。

20．试说明基于分数低阶统计量的信道均衡的概念与方法。

21．试分别由高斯分布和柯西分布的特征函数推导各自的概率密度函数。

22．试根据共变系数的定义来分析 p 值对共变系数的影响。

23．设 U_1、U_2 是独立同分布的标准 $S\alpha S$ 分布（$\alpha=1.5$）的随机变量。X 和 Y 是 U_1、U_2 的线性组合：$X=a_1U_1+a_2U_2$，$Y=b_1U_1+b_2U_2$。若 $(a_1,a_2,b_1,b_2)=(-0.75,0.25,0.18,0.78)$，试求共变系数 λ_{XY}。

24．试利用 MATLAB 编程，根据计算机产生标准 $S\alpha S$ 分布（$\alpha=1.5$）随机变量 U_1、U_2 的数据，按照上题构成随机变量 X 和 Y，依据 FLOM 方法估计共变系数 λ_{XY}。其中分别取 $p=0.3, 0.5, 0.7, 1.0, 1.3$。

25．试利用 MATLAB 编程，实现 LMP 自适应滤波算法的计算机仿真，并于 LMS 自适应算法进行对比。

26．试利用 MATLAB 编程，实现核自适应滤波器（KLMP 和 KLMS 算法）的计算机仿真，比较二者的性能。

27．试利用 MATLAB 编程，在 Alpha 稳定分布下（如 $\alpha=1.5$，GSNR = 5dB）实现基于分数低阶统计量的时间延迟估计算法。进一步实现目标信源定位。并与基于二阶统计量的算法对比。

28．改变 α 取值，再做习题 27。

29．改变 GSNR，再做习题 27。

30．试利用 MATLAB 编程，在 Alpha 稳定分布下（如 $\alpha=1.5$，GSNR = 5dB）实现基于分数低阶统计量的波达方向估计算法。并进一步实现目标信源定位。

31．改变 α 取值，再做习题 30。

32．改变 GSNR，再做习题 30。