

## 参数估计定理

徐大专, 屠伟林, 施超, 周颖

(南京航空航天大学电子信息工程学院, 南京, 211106)

**摘要:** 空间信息论是关于雷达等目标参数估计系统信息获取一般规律的基本理论, 这里的空间信息指被测目标相对于雷达的距离、方向和散射信息。参数估计定理是空间信息论的重要组成部分, 本文提出参数估计定理的证明框架, 该框架由3个原创性概念和1个定理构成, 其中3个概念是空间信息、熵误差和抽样后验概率估计。参数估计定理证明, 熵误差是可达的, 反之, 任何估计器的熵误差不低于熵误差。空间信息论的建立将对雷达探测的系统理论和方法产生巨大的推动作用。

**关键词:** 空间信息; 熵误差; 经验熵误差; 抽样后验概率估计; 参数估计定理

**中图分类号:** TN911      **文献标志码:** A

## Parameter Estimation Theorem

XU Dazhuan, TU Weilin, SHI Chao, ZHOU Ying

(College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, 211106, China)

**Abstract:** Spatial information theory is the fundamental theory of radar and other target detection systems. The spatial information is associated with the range, the direction and scattering signal of targets. Parameter estimation theorem is an important part of the spatial information theory. The proof framework of the parameter estimation theorem is proposed in this paper, which consists of three original concepts of spatial information, entropy error and sampling posterior probability estimation, and one theorem. The parameter estimation theorem shows that the entropy error is reachable, otherwise the empirical entropy error of any estimator is no less than the entropy error. The establishment of spatial information theory will play a significant role in promoting the development of systematic theories and methods of radar.

**Key words:** spatial information; entropy error; empirical entropy error; sampling a posterior estimation; parameter estimation theorem

## 引言

雷达、声纳和医学成像等目标参数估计系统可以从反射信号中获取目标的距离、方向和幅度等空间信息, 正在国防和国民经济部门发挥越来越重要的作用。雷达参数估计的主要任务是目标检测、参数估计和区域成像。除了距离和散射信息之外, 相控阵雷达和合成孔径雷达还可获得目标的方向信息, 或者对观测区域进行成像。干涉合成孔径雷达甚至可以获得观测区域的三维空间信息。目前, 雷

达与通信技术相融合的趋势越来越明显,雷达和通信系统中MIMO多天线技术的应用越来越广泛<sup>[1-6]</sup>,雷达通信一体化技术<sup>[7-14]</sup>的不断发展,从共用天线到共用射频,再到共用波形,融合的程度也在加深。

雷达是一种典型的信息获取系统,而通信是一种信息传输系统,那么,雷达和通信两种信息系统能否在信息理论的基础上进行统一的描述和刻画?信息论与雷达的关系是国际学术界长期关注,而未能解决的基础理论问题。2017年,国际信息论期刊《Entropy》在“雷达与信息理论”特刊的征稿启示<sup>[15]</sup>中充分表达了国际学术界的迫切期望。

从Woodward和Davies研究距离信息<sup>[16-18]</sup>至今已有近70年历史,特别是Bell的研究工作<sup>[19]</sup>发表后,以互信息和相对熵为测度进行雷达系统的波形设计引起了国内外学术界的广泛关注。关于距离互信息的研究可追溯至20世纪50年代,Woodward和Davies<sup>[16-18]</sup>采用逆概率原理研究距离信息问题,得到了单个恒模散射目标的距离互信息与时带积、信噪比(Signal noise ratio, SNR)的近似关系。然而,Woodward和Davies只研究了雷达参数估计系统中目标的距离信息问题,而没有涉及散射信息问题,并且逆概率研究方法的局限性导致他们的工作无法推广到多目标情况。一个颇为不解的现象是,关于距离信息的研究工作在此后近70年中再无人涉足。1988年,Bell首先将互信息测度用于雷达系统的波形设计<sup>[19]</sup>,根据Neyman-Pearson准则,Bell证明了最佳波形设计与信道容量的最优功率注水解相对应<sup>[20]</sup>,正好与通信系统的最优功率分配问题相吻合。在Bell的系统模型中,目标的距离信息隐含于冲激响应中。由于实际环境中目标位置是不断变化的,必须采用自适应的波形设计方法。由于Bell的工作是针对目标检测问题提出来的,其模型并不区分不同的目标,因此,从本质上说,Bell的工作只研究了空间信息中的散射信息问题,而没有涉及雷达参数估计系统中更重要的距离信息问题。

2017年,笔者提出了雷达参数估计的空间信息概念,将空间信息定义为接收信号与目标距离及散射的联合互信息,从而将距离信息和幅度信息纳入的统一定义框架中<sup>[21]</sup>;2019年进一步提出熵误差概念<sup>[22]</sup>,并用熵误差评价参数估计系统的性能。熵误差是均方误差指标的推广,在高SNR条件下退化为均方误差,证明了每获得1 bit位置信息等价于熵偏差缩小一半,或参数估计精度提高一倍。笔者团队已经通过理论推导得到了单目标/多目标距离信息和给定目标位置条件下散射信息的闭合表达式<sup>[23-24]</sup>。

2018年,笔者将空间信息的概念推广到传感器阵列,给出了波达方向(Direction of arrival, DOA)信息和散射的统一定义<sup>[23]</sup>,并针对恒模信源,得到了DOA互信息的闭合表达式。最近,进一步研究相控阵雷达的空间信息问题,把目标的距离信息、方向信息和散射信息纳入到空间信息的统一定义框架<sup>[25]</sup>。用空间信息理论解释雷达分辨率是笔者近期研究工作<sup>[26]</sup>,发现可以通过特征信道的散射信息定义空间分辨率 $\Delta = 2/\rho\beta$ ,这种分辨率不仅与信号带宽有关,还与SNR有关。由于新分辨率在形式上与Cramér-Rao界(CRB)一致,故称Cramér-Rao分辨率。Cramér-Rao分辨率在理论上指出超分辨的可能性,并已得到多维最大似然估计方法的验证。指出传统分辨率是一维最大似然估计方法的分辨能力,而分辨率问题本质上是多维问题,因此,多维最大似然估计方法的超分辨率与传统分辨率并不矛盾。文献<sup>[27]</sup>总结了空间信息论研究的最新进展情况。

参数估计一直是雷达信号处理所关心的问题,并形成相对成熟的理论体系。但以笔者的观点,参数估计还存在两大问题:第一个问题与参数估计精度的评价指标有关。参数估计的精度通常用均方误差衡量,然而,在中低SNR条件下误差统计量一般不是二阶统计量,这时仍然采用均方误差作为评价指标既不全面,也欠合理。该问题在测量领域也普遍存在。第二个问题与参数估计的最优性有关,具体来说,是否存在某种最优的参数估计方法,如果存在,这种方法在何种意义上最优,且最优估计精度能达到多少?这两个问题正是本文参数估计定理要回答的。

本文的内容安排:第1节给出雷达多目标参数估计的系统模型。第2节给出空间信息的定义。第3节提出熵误差的概念,并用熵误差评价参数估计方法的性能;还给出一种新颖的抽样后验概率估计方

法,该方法是一种随机参数估计方法,不同于传统的确定性参数估计方法。第4节描述并证明参数估计定理。第5节进一步展望空间信息论的发展方向。

### 1 雷达探测系统模型

假设有  $L$  个目标在探测范围中,目标间相互独立,目标的位置和散射信号也相互独立。雷达接收的基带信号可表示为

$$z(t) = s_l \psi(t - \tau_l) + w(t) \tag{1}$$

式中:  $\psi(t)$  表示带宽为  $B/2$  的基带发送信号;  $s_l = \alpha_l e^{j\varphi_l}$  为第  $l$  个目标的复散射系数,其中  $\alpha_l$  为幅度,  $\varphi_l = -2\pi f_c \tau_l + \varphi_0$  为相位;  $\tau_l = 2d_l/v$  为第  $l$  个目标的时延,  $d_l$  为目标位置到雷达接收机位置的距离,  $v$  为信号传播速度;  $w(t)$  表示带宽为  $B/2$ , 均值为零的复高斯白噪声, 它的实部和虚部的功率谱密度均为  $N_0/2$ 。

一般来说, 散射系数的幅度  $\alpha$  是时延的函数, 随距离的增加而减小。为简单起见, 本文将  $\alpha$  看成常数, 隐含地假设观测区间较小, 可以忽略衰减的影响。对于大观测区间, 可将观测区间分成若干小区间, 虽然每个区间的  $\alpha$  不同, 但在一个小区间内可视为常数, 后面的分析方法仍然适用。

为了理论分析方便, 假设观测区间的中心为目标参考点位置, 目标观测区间设为  $[-D/2, D/2]$ , 如图 1(a) 所示。则接收信号对应的时延范围为  $[-T/2, T/2]$ , 如图 1(b) 所示。假设信号模型中发射信号是理想的低通信号, 即基带信号为

$$\psi(t) = \text{sinc}(Bt) = \frac{\sin(\pi Bt)}{\pi Bt} \quad -\frac{T}{2} \leq t \leq \frac{T}{2} \tag{2}$$

式中:  $T$  表示观测时间, 假设  $T \gg 1/B$ , 即  $BT \gg 1$ , 这时信号能量几乎全部位于观测区间之内, 即

$$E_s = \int_{-T/2}^{T/2} \psi^2(t) dt = 1 \tag{3}$$

$\psi(t)$  的频谱为

$$\psi(f) = \begin{cases} \frac{1}{B} & |f| \leq \frac{B}{2} \\ 0 & \text{其他} \end{cases} \tag{4}$$

根据 Shannon-Nyquist 采样定理, 以采样率  $B$  对接收信号  $z(t)$  进行采样, 得到离散形式如图 1(c) 所示。

$$z\left(\frac{n}{B}\right) = \sum_{l=1}^L s_l \psi\left(\frac{n - B\tau_l}{B}\right) + w\left(\frac{n}{B}\right) \quad n = -\frac{N}{2}, \dots, \frac{N}{2} - 1 \tag{5}$$

式中  $N = TB$  为时间带宽积(Time bandwidth product, TBP)。令  $x_l = B\tau_l$ , 表示目标的归一化延迟, 它反映了目标位置与雷达的距离, 那么, 可以得到离散形式系统方程为

$$z(n) = \sum_{l=1}^L s_l \psi(n - x_l) + w(n) \quad n = -\frac{N}{2}, \dots, \frac{N}{2} - 1 \tag{6}$$

带宽为  $B/2$  的高斯噪声的自相关函数为

$$R(\tau) = \frac{N_0 B \sin \pi B \tau}{2 \pi B \tau} \tag{7}$$

以采样率  $B$  得到的离散噪声样值  $w(n)$  间隔为  $1/B$  的整数倍。由式(7)可知, 噪声样值是不相关的, 并且  $w(n)$  是复高斯随机变量, 所以  $w(n)$  互相独立。

为了描述方便, 将式(6)写成矢量形式

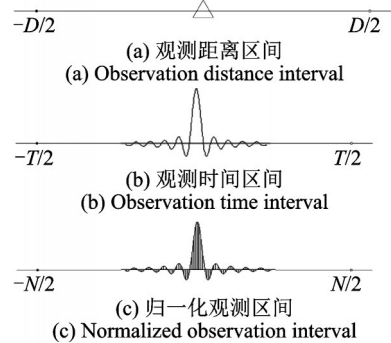


图 1 三种观测区间与信号波形  
Fig.1 Three observation intervals and signal waveforms

$$\mathbf{z} = \mathbf{U}(\mathbf{x})\mathbf{s} + \mathbf{w} \quad (8)$$

式中:  $\mathbf{z} = [z(-N/2), \dots, z(N/2-1)]^T$  表示离散接收信号;  $\mathbf{U}(\mathbf{x}) = [\mathbf{u}(x_1), \dots, \mathbf{u}(x_L)]^T$  表示由发射信号波形和目标时延确定的位置矩阵, 其第  $l$  列矢量  $\mathbf{u}(x_l) = [\sin c(-N/2 - x_l), \dots, \sin c(N/2 - 1 - x_l)]^T$  是经过第  $l$  个目标时延后的回波;  $\mathbf{s} = [s_1, \dots, s_L]^T$  表示目标散射矢量;  $\mathbf{w} = [w(-N/2), \dots, w(N/2-1)]^T$  表示噪声矢量, 其分量是独立同分布的、均值为 0、方差为  $N_0$  的复高斯随机变量。

## 2 空间信息的概念及其定量

### 2.1 目标和信道的统计模型

雷达参数估计系统目标的统计特性对应于通信系统的信源。目标统计特性是距离和散射的联合分布  $p(\mathbf{x}, \mathbf{s})$ , 一般目标的位置和散射是不相关的, 即

$$p(\mathbf{x}, \mathbf{s}) = p(\mathbf{x})p(\mathbf{s}) \quad (9)$$

式中:  $p(\mathbf{x})$  表示归一化距离的先验概率密度函数 (Probability density function, PDF),  $p(\mathbf{s})$  表示散射信号的 PDF。

在没有任何先验信息的情况下, 目标的距离通常假定为在观测区间上均匀分布的随机变量。

雷达目标散射特性十分复杂, 目前已经针对不同散射场景建立散射信号统计模型, 本文只考虑恒模 (Swierling 0) 和复高斯 (Swierling 1) 两种典型的雷达电磁散射信号统计模型。恒模散射系数模型刻画的目标处于准静止状态, 其功率恒定, 相位  $\varphi$  在  $[0, 2\pi]$  上均匀分布。复高斯散射模型刻画的目标由许多随机分布的散射体组成, 且其中没有起主要作用的强散射体。散射系数的实部和虚部为相同的高斯分布且互相独立。换句话说, 复高斯散射系数的幅度呈瑞利分布。两种散射信号的统计特性可表示为

$$p(\mathbf{s}) = p(\alpha, \varphi) = p(\alpha)p(\varphi) = \begin{cases} \frac{1}{2\pi} \delta(\alpha - \alpha_0) & \text{恒模散射模型 (Swierling 0)} \\ \frac{1}{2\pi} \frac{\alpha}{\sigma_a^2} \exp\left(-\frac{\alpha^2}{2\sigma_a^2}\right) & \text{复高斯散射模型 (Swierling 1)} \end{cases} \quad (10)$$

设信道是复高斯加性白噪声信道 (Complex Gaussian additive white noise, CGAWN),  $N$  维复高斯噪声矢量的概率密度函数为

$$p(\mathbf{w}) = \frac{1}{(\pi N_0)^N} \exp\left(-\frac{1}{N_0} \mathbf{w}^T \mathbf{w}\right) \quad (11)$$

那么, 给定  $\mathbf{X}$  和  $\mathbf{S}$  时  $\mathbf{Z}$  的  $N$  维条件概率密度函数为

$$p(\mathbf{z} | \mathbf{x}, \mathbf{s}) = \frac{1}{(\pi N_0)^N} \exp\left(-\frac{1}{N_0} (\mathbf{z} - \mathbf{U}(\mathbf{x})\mathbf{s})^H (\mathbf{z} - \mathbf{U}(\mathbf{x})\mathbf{s})\right) \quad (12)$$

式(12)展开可得

$$p(\mathbf{z} | \mathbf{x}, \mathbf{s}) = \frac{1}{(\pi N_0)^N} \exp\left(-\frac{1}{N_0} \left(z^H z - 2\Re(\mathbf{s}^H \mathbf{U}^H(\mathbf{x})\mathbf{z}) + \mathbf{s}^H \mathbf{U}^H(\mathbf{x})\mathbf{U}(\mathbf{x})\mathbf{s}\right)\right) \quad (13)$$

式中  $\Re(\cdot)$  表示取实部。式(13)定义了雷达参数估计系统的信道特性。

### 2.2 目标空间信息的定义

从统计的观点处理系统方程, 令  $\mathbf{X}$  和  $\mathbf{S}$  分别表示多目标的归一化随机距离矢量和散射矢量,  $\mathbf{Z}$  和  $\mathbf{W}$  为随机接收信号矢量和噪声矢量。现有如下定义:

**定义 1 (空间信息)** 目标的空间信息定义为接收信号与距离和散射信号的联合互信息, 即

$$I(Z; X, S) = E \left[ \log \frac{p(z|x, s)}{p(z)} \right] \quad (14)$$

式中

$$p(z) = \iint p(x, s) p(z|x, s) dx ds \quad (15)$$

表示接收信号的无条件分布。

由互信息的可加性可以证明,空间信息为距离信息  $I(Z; X)$  与已知距离的条件散射信息  $I(Z; S|X)$  之和,即

$$I(Z; X, S) = I(Z; X) + I(Z; S|X) \quad (16)$$

从式(16)可以看出,空间信息的计算过程为:(1)计算目标的距离信息  $I(Z; X)$ ; (2)计算已知距离的条件散射信息  $I(Z; S|X)$ 。

空间信息的定义是笔者在文献[21]中首次给出的,在此之前,距离信息的研究工作很少,而散射信息的研究则很多,但都没有将两者放在一起研究。通常认为距离和散射有不同的单位,很难将两者纳入一个统一概念框架之中。空间信息的定义是一个正式的标志,意味着香农信息论正式成为雷达等信息获取系统的理论基础。从此,同信息传输系统一样,信息获取系统也可以用比特作为单位进行统一量了。

### 3 参数估计方法及性能评价

考虑复高斯散射目标和CAWGN信道,这时接收信号  $Z$  也是复高斯的,其协方差矩阵  $R$  为

$$R = E_{s, w} [z z^H] \quad (17)$$

将  $z = U(x)s + w$  代入式(17)可得

$$\begin{aligned} R &= E \left[ (U(x)s + w)(U(x)s + w)^H \right] = \\ &U(x) E[ss^H] U(x) + E[ww^H] = \\ &N_0 \left( \sum_{l=1}^L \frac{\rho_l^2}{2} u_l(x) u_l^H(x) + I \right) \end{aligned} \quad (18)$$

式中  $\rho_l^2 = 2E \left[ |s_l|^2 \right] / N_0$  表示第  $l$  个目标的平均 SNR。注意,协方差矩阵是目标距离矢量  $X$  的函数。给定  $X$  时  $Z$  的条件概率密度为

$$p(z|x) = \frac{1}{\pi^N |R|} \exp(-z^H R^{-1} z) \quad (19)$$

条件概率分布刻画信道的统计特性,又称为似然函数。使式(19)达到最大的估计值  $\hat{x}$  称为距离  $x$  的最大似然估计,记为  $\hat{x}_{ML}$ ,则

$$\hat{x}_{ML} = \arg \max_x \frac{1}{|R|} \exp(-z^H R^{-1} z) \quad (20)$$

实际中常用对数似然函数,那么

$$\hat{x}_{ML} = \arg \max_x \left\{ -z^H R^{-1} z - \ln |R| \right\} \quad (21)$$

等价于

$$\hat{x}_{ML} = \arg \min_x \left\{ z^H R^{-1} z + \ln |R| \right\} \quad (22)$$

式(22)是多目标距离参数的最大似然估计表达式,显然,普通的单目标匹配滤波估计器不是最大似然估计器。从式(22)还可以看到,协方差矩阵在多目标参数估计中占有核心地位,如何充分利用统



计信息对多目标参数估计的性能有很大影响。

设目标距离的先验分布为  $p(x)$ , 由贝叶斯公式得后验概率分布为

$$p(x|z) = \frac{\frac{1}{|R|} p(x) \exp(-z^H R^{-1} z)}{\oint \frac{1}{|R|} p(x) \exp(-z^H R^{-1} z) dx} \quad (23)$$

式(23)的分母为归一化常数, 先验分布的作用可看成一个多维权函数对似然函数加权。

使式(23)达到最大的估计值  $\hat{x}$  称为目标距离  $x$  的最大后验概率估计, 记为  $\hat{x}_{\text{MAP}}$ , 则

$$\hat{x}_{\text{MAP}} = \arg \max_x \frac{1}{|R|} p(x) \exp(-z^H R^{-1} z) \quad (24)$$

如果目标距离在观测区间内服从均匀分布, 则已知接收信号时目标归一化时延的后验概率分布为

$$p(x|z) = \frac{\frac{1}{|R|} \exp(-z^H R^{-1} z)}{\oint \frac{1}{|R|} \exp(-z^H R^{-1} z) dx} \quad (25)$$

式(25)的分母是定义在观测区间上的多重积分, 积分结果是一个归一化常数。

显然, 当先验分布与目标位置无关时, 最大后验概率估计等同于最大似然估计。先验分布相当于一个窗函数, 对最大似然估计统计量进行加权。实际中第一次参数估计时, 因没有先验信息故使用最大似然估计。在后续参数估计和目标跟踪时, 利用已有的先验信息作最大后验概率估计, 不断提高参数估计精度。

**定义 2 (抽样后验概率估计)** 对后验概率分布  $p(x|z)$  进行抽样产生的估计值  $\hat{x}$  称为目标距离  $x$  的抽样后验 (Sampling a posterior, SAP) 概率估计, 记为  $\hat{x}_{\text{SAP}}$ , 则

$$\hat{x}_{\text{SAP}} = \arg \text{sam}_x \frac{1}{|R|} p(x) \exp(-z^H R^{-1} z) \quad (26)$$

常见的最大似然估计和最大后验概率估计都是确定性估计, 对给定接收信号序列的估计值是唯一的, 而抽样后验概率估计是一种随机估计器, 对给定接收信号的估值  $\hat{x}_{\text{SAP}}$  是不确定的。本文提出抽样后验概率估计的目的是为了证明后面的参数估计定理, 因为它的性能取决于后验概率分布, 而最大后验概率估计的性能不容易确定。这种思想与香农编码定理采用的随机编码一脉相承。抽样后验概率估计方法还具有重要的实际应用价值, 它避免了确定性估计方法遇到的谱峰搜索问题, 在多维参数估计应用场景具有低复杂度优势。

参数估计的性能通常用均方误差来衡量, 然而, 在中低 SNR 比条件下判决统计量一般不是二阶的, 这时用均方误差作为评价指标既不全面, 也欠合理。那么, 如何评价估计器性能呢? 后验微分熵  $h(x|z)$  表示不确定性,  $h(x|z)$  越小意味着估计越准确。为了从信息论角度评价估计器性能, 有

**定义 3 (熵误差)** 设某一距离估计器的后验概率分布为  $p(x|z)$ , 那么

$$\sigma_{\text{EE}}^2 = \frac{2^{2h(x|z)}}{2\pi e} \quad (27)$$

称为距离估计的熵误差, 其中

$$h(x|z) = - \oint p(x|z) \log p(x|z) dx dz \quad (28)$$

为对应于  $p(x|z)$  的微分熵。

熵误差简记为  $\sigma_{\text{EE}}^2$ , 这里熵误差表示参数矢量的联合估计精度。在此之前, 由于参数间存在相关性,

多参数估计的精度需要一个矩阵去描述,可参考费雪信息矩阵。熵误差用一个标量就可以表示联合估计的总体性能,充分反映信息论方法在描述估计器性能方面的优势。

**定义4 (熵偏差)** 熵误差的平方根  $\sigma_{EE}$  称为熵偏差。

以上定义的熵误差和熵偏差只与目标和信道的统计特性有关,与具体的参数估计方法无关,有时又称理论熵误差,以区别于后面与估计方法相关的经验熵误差。事实上,熵误差就是后验概率分布的熵功率,它是均方误差的推广。在高 SNR 条件下,后验概率分布逼近高斯分布,这时熵误差退化为均方误差,因此,熵误差相比均方误差对各种 SNR 条件具有普适性。

针对单目标情况,后验熵的近似表达式为<sup>[28]</sup>

$$h(x|z) = p_s \log \frac{\sqrt{2\pi e}}{\rho\beta} + (1-p_s) \log \frac{N\rho\sqrt{2\pi}}{e^{\frac{1}{2}(\rho^2+1)}} + H(p_s) \quad (29)$$

式中离散熵为

$$H(p_s) = -p_s \log p_s - (1-p_s) \log (1-p_s) \quad (30)$$

而

$$p_s = \frac{\exp\left(\frac{1}{2}\rho^2 + 1\right)}{N\rho^2\beta + \exp\left(\frac{1}{2}\rho^2 + 1\right)} \quad (31)$$

称为检测因子,它刻画了检测到目标的概率,当  $\rho^2 \rightarrow 0$  时,  $p_s \rightarrow 0$ ; 当  $\rho^2 \rightarrow \infty$  时,  $p_s \rightarrow 1$ 。

将式(29)代入式(27)可得熵误差的近似公式为

$$\begin{aligned} \sigma_{EE}^2 &= \frac{1}{2\pi e} 2^{2h(x|z)} = \\ &= \frac{1}{2\pi e} \left(\frac{2\pi e}{\rho^2\beta^2}\right)^{p_s} \left(\frac{N^2\rho^2 2\pi}{e^{(\rho^2+1)}}\right)^{1-p_s} \left(\frac{1}{p_s}\right)^{2p_s} \left(\frac{1}{1-p_s}\right)^{2(1-p_s)} = \\ &= \left(\frac{N\rho^2\beta + \exp\left(\frac{1}{2}\rho^2 + 1\right)}{\rho\beta \exp\left(\frac{1}{2}\rho^2 + 1\right)}\right)^2 = \frac{1}{\rho^2\beta^2 p_s^2} \end{aligned} \quad (32)$$

式(32)表明熵误差是克拉美罗界在各种 SNR 条件下的推广,在 SNR 条件下,  $p_s$  趋近于 1,熵误差退化为克拉美罗界。

图 2 比较了目标距离估计的理论熵误差、均方误差、克拉美罗界和熵误差的近似公式。由图 2 可见,后验概率分布随着 SNR 的提高逐渐逼近高斯分布,这时熵误差渐渐退化为均方误差,并逼近克拉美罗界。

熵误差/熵偏差和距离信息具有密切的联系,设  $\sigma_{EE}(X)$  是先验分布  $p(x)$  定义的熵偏差,即

$$\sigma_{EE}(X) = \frac{2^{h(X)}}{\sqrt{2\pi e}} \quad (33)$$

$\sigma_{EE}(X|Z)$  是后验概率分布  $p(x|z)$  定义的熵偏差,即

$$\sigma_{EE}(X|Z) = \frac{2^{h(X|Z)}}{\sqrt{2\pi e}} \quad (34)$$

那么有

$$\frac{\sigma_{EE}(X|Z)}{\sigma_{EE}(X)} = 2^{h(X) - h(X|Z)} = 2^{-I(Z;X)} \quad (35)$$

式中  $I(Z;X)$  表示距离信息。那么,有

**定理 1** 每获取 1bit 距离信息等价于熵偏差缩小一半,或估计精度提高一倍。

定理 1 表明,距离信息和熵偏差都可以表征参数估计系统性能,距离信息代表获取信息的数量,熵偏差代表参数估计的精度,两者是完全等价的。该定理充分说明,用熵误差/熵偏差评价参数估计系统性能是合理的。

参数估计信息与参数估计精度之间的关系也有助于理解人类的学习与认知过程。人们对事物的认知程度由某种概率分布刻画,尽管通常并不能给出这种概率分布的具体数学形式。学习过程就是信息获取的过程,获取多少信息量取决于认知概率的变化。如果认知概率分布的熵偏差越小,则获得的信息就越多。随着认知的逐渐深入,不确定性会不断降低,对事物的认识程度也不断提高。在人工智能和机器学习领域,还没有一个认知系统模型,学习过程由数据驱动,所以通常很难获得具体的认知概率分布,但可以通过机器学习结果来评估所获得的信息量。

#### 4 参数估计定理

香农信息论的核心内容是编码定理,那么,空间信息论是否存在类似的定理呢?在信号处理和参数估计领域,通常默认最大似然估计和最大后验概率估计是最佳的,并作为估计准则。人们长期忽视的基本理论问题是:(1)何种估计器是最优的;(2)在什么意义上最优;(3)最优的性能是什么?

参数估计定理涉及的内容非常广泛,本节只针对单目标位置参数估计问题给出证明的框架。在证明定理之前先定义需要用到的概念。

**定义 5 (目标时延特性)** 在观测区间上一个目标的归一化时延为随机变量,归一化时延的先验分布称为目标时延特性。

在时延参数估计中目标时延特性相当于信源统计特性,故也简称为目标或信源。

**定义 6 (参数估计信道)** 参数估计信道  $(\mathcal{X}, p(z|x), Z)$  的输入是目标的归一化时延,定义在有限的实观测区间上,信道的输出是由接收复信号序列组成的集合,信道特性由条件概率密度函数  $p(z|x)$  确定。

**定义 7 (估计器)** 估计器是对归一化时延的一个估计函数  $\hat{x} = f(z)$ ,对给定的接收序列输出一个距离的估计值  $\hat{x}$ 。

**定义 8 (联合目标-信道)** 联合目标-信道  $(\mathcal{X}, p(x), p(z|x), Z)$  是指目标时延特性和信道统计特性组成的总体。

联合目标-信道定义了估计器所要面对的参数估计环境,这里假定估计器已知联合目标-信道的全部统计特性。

**定义 9 (参数估计系统)** 参数估计系统  $(\mathcal{X}, p(x), p(z|x), \hat{x} = f(z), Z)$  刻画目标特性、信道特性和估计器组成的总体。

一次参数估计过程由目标、信道和估计器 3 部分组成,简称为一次快拍。多次快拍将产生扩展目标和扩展信道, $M$  次快拍的参数估计过程如图 3 所示。

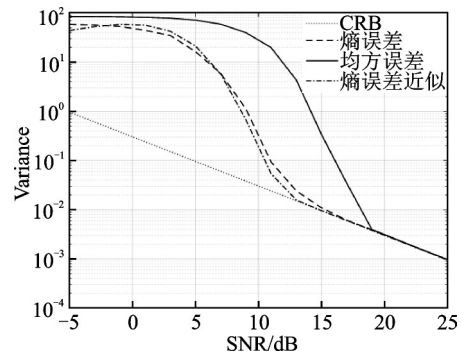


图 2 目标归一化时延的方差

Fig.2 Variance of target's normalized delay



**定义 10** (无记忆扩展目标) 无记忆扩展目标指扩展目标之间相互独立。

**定义 11** (无记忆扩展信道) 无记忆快拍信道(Memoryless snapshot channel, MSC)指多次快拍产生的扩展信道 $(\mathcal{X}^M, p(\mathbf{z}^M|x^M), \mathcal{Z}^M)$ 满足

$$p(\mathbf{z}^M|x^M) = \prod_{m=1}^M p(z_m|x_m) \quad (36)$$

联合目标-信道 $(\mathcal{X}, p(x), p(z|x), \mathcal{Z})$ 确定了后验概率分布 $p(x|z)$ 和后验微分熵 $h(x|z)$ ,进而也确定了理论熵误差为 $\sigma_{EE}^2 = \frac{1}{2\pi e} 2^{2h(x|z)}$ 。本文还需定义另一种与估计器相关的熵误差。

**定义 12** (经验熵误差)  $M$ 次快拍的经验熵定义为 $-\frac{1}{M} \log p(\hat{x}^M|z^M)$ ,经验熵误差为

$$\sigma_{EE}^2(M) = \frac{1}{2\pi e} 2^{-\frac{2}{M} \log p(\hat{x}^M|z^M)} \quad (37)$$

**定义 13** (可达性) 熵误差 $\sigma_{EE}^2$ 称为可达的,如果存在一个估计器,其 $M$ 次快拍的经验熵误差满足

$$\lim_{M \rightarrow \infty} \sigma_{EE}^2(M) = \sigma_{EE}^2 \quad (38)$$

**定义 14** (联合典型序列) 服从联合分布 $p(x, z)$ 的联合典型序列 $\{(x^M, z^M)\}$ 所构成的集合 $A_\epsilon^{(M)}$ 是指它的经验熵与真实熵差值小于 $\epsilon$ 的 $M$ 长序列构成的集合,即

$$A_\epsilon^{(M)} = \left\{ (x^M, z^M) \in \mathcal{X}^M \times \mathcal{Z}^M : \left| -\frac{1}{M} \log p(x^M) - H(X) \right| < \epsilon, \left| -\frac{1}{M} \log p(z^M) - H(Z) \right| < \epsilon, \left| -\frac{1}{M} \log p(x^M, z^M) - H(X, Z) \right| < \epsilon \right\} \quad (39)$$

式中

$$p(x^M, z^M) = p(x_m, z_m)$$

这里定义的联合典型序列与香农信息论完全一致,即扩展信源信道 $(\mathcal{X}^M, p(x^M), p(z^M|x^M), \mathcal{Z}^M)$ 的输入和输出构成联合典型序列。

**引理 1** 对于无记忆快拍信道 $(\mathcal{X}^M, p(z^M|x^M), \mathcal{Z}^M)$ ,如果 $\hat{x}^M$ 是后验概率分布 $p(x|z)$ 的 $M$ 次抽样估计,则 $(\hat{x}^M, z^M)$ 是关于概率分布 $p(\hat{x}^M, z^M)$ 的联合典型序列。

**证明:** 由于 $\hat{x}^M$ 是后验概率分布 $p(x|z)$ 的 $M$ 次抽样估计,则扩展后验概率分布 $p_f(\hat{x}^M|z^M) = p(\hat{x}^M|z^M)$ ,那么

$$p_f(\hat{x}^M, z^M) = p(z^M) p_f(\hat{x}^M|z^M) = p(z^M) p(\hat{x}^M|z^M) = p(\hat{x}^M, z^M) \quad (40)$$

证毕。

抽样后验概率估计的性能完全由后验概率分布 $p(x|z)$ 确定,因此,由抽样后验概率估计得到的扩展序列 $(\hat{x}^M, z^M)$ 是关于概率分布 $p(\hat{x}^M, z^M)$ 联合典型的。

**定理 2** (参数估计定理) 熵误差 $\sigma_{EE}^2$ 是可达的,具体来说,设估计器已知联合信源-信道 $(\mathcal{X}, p(x), p(z|x))$ 统计特性,则,对任意 $\epsilon > 0$ ,必定存在估计器,其经验熵误差满足

$$\sigma_{EE}^2 e^{-4\epsilon} < \sigma_{EE}^2(M) < \sigma_{EE}^2 e^{4\epsilon} \quad (41)$$

且

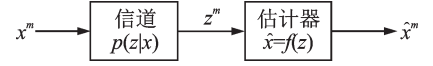


图3  $M$ 次快拍的参数估计系统  
Fig.3 Parameter estimation system of  $M$  snapshots

$$\lim_{M \rightarrow \infty} \sigma_{\text{EE}}^2(M) = \sigma_{\text{EE}}^2 \quad (42)$$

反之,任何估计器的熵误差不可能小于理论熵误差。

定理分为正定理和逆定理两部分,先证明正定理。

正定理的证明:

(1) 根据目标距离特性独立产生  $M$  次扩展目标  $x^M$ ;

(2) 根据扩展目标  $x^M$  和  $M$  次扩展信道特性  $p(z|x)$  产生接收序列  $z^M$ , 经过  $M$  次快拍产生的接收信号  $z^M$  满足

$$p(z^M|x^M) = \prod_{m=1}^M p(z_m|x_m) \quad (43)$$

采用抽样后验概率估计器,令  $\hat{x}^M$  是对于无记忆快拍信道  $p(z^M|x^M)$  的  $M$  次抽样估计,由引理 1 可知,  $(\hat{x}^M, z^M)$  是关于概率分布  $p(\hat{x}^M, z^M)$  的联合典型序列。

根据联合典型序列的定义,对任意  $\epsilon > 0$ , 只要快拍数足够大,有

$$\left| -\frac{1}{M} \log p(\hat{x}^M, z^M) - H(X, Z) \right| < \epsilon \quad (44)$$

$$\left| -\frac{1}{M} \log p(z^M) - H(Z) \right| < \epsilon \quad (45)$$

由于  $p(\hat{x}^M|z^M) = p(\hat{x}^M, z^M) / p(z^M)$ , 那么

$$\left| -\frac{1}{M} \log p(\hat{x}^M|z^M) - H(X|Z) \right| < 2\epsilon \quad (46)$$

也就是

$$H(X|Z) - 2\epsilon < -\frac{1}{M} \log p(\hat{x}^M|z^M) < H(X|Z) + 2\epsilon \quad (47)$$

根据熵误差及经验熵误差的定义,有

$$\sigma_{\text{EE}}^2 e^{-4\epsilon} < \sigma_{\text{EE}}^2(M) < \sigma_{\text{EE}}^2 e^{4\epsilon} \quad (48)$$

根据联合典型序列的定义和切比雪夫不等式,随快拍数  $M \rightarrow \infty$ ,  $\epsilon \rightarrow 0$ , 则

$$\lim_{M \rightarrow \infty} \sigma_{\text{EE}}^2(M) = \sigma_{\text{EE}}^2 \quad (49)$$

这说明采用抽样后验概率估计器的经验熵误差可逼近理论熵误差。

逆定理的证明:

令  $\hat{x}^m = f(z^m)$  是任一估计器,由该估计器获得的互信息记为  $I_f(Z^M, X^M)$ , 那么

$$I_f(Z^M, X^M) = h(X^M) - h_f(X^M|Z^M) \quad (50)$$

式中  $h_f(X^M|Z^M)$  是该估计器的后验微分熵。显然  $(Z^M, \hat{X}^M, X^M)$  组成马尔可夫链,由数据处理定理

$$I_f(Z^M, X^M) \leq I(Z^M, X^M) \quad (51)$$

因此

$$h_f(X^M|Z^M) \geq h(X^M) - I(Z^M, X^M) = h(X^M|Z^M) \quad (52)$$

由熵误差的定义立即有

$$\sigma_{\text{EE}}^2(M)(X^M|Z^M) \geq \sigma_{\text{EE}}^2 \quad (53)$$

证毕。

参数估计定理的证明是构造性的,就是说,抽样后验概率估计是一种实际可使用的估计方法,其性能是渐近最优的。

参数估计定理表明,熵误差是任何参数估计方法的理论极限。在此之前,虽然克拉美罗界也被当作是任何参数估计方法的理论下界,但它只有在高 SNR 条件下是渐近的,而在中低 SNR 条件下明显是不可达的。熵误差下界在各种 SNR 条件下都是可达的,由于实际系统往往更多工作于中低 SNR 条件下,因此,熵误差给出的理论下界具有更重要的实际意义。

## 5 结束语

参数估计定理指出,熵误差是可达的,并且是所有参数估计方法所能达到的理论下界。熵误差之于参数估计就像信道容量之于信道编码,率失真函数之于信源编码的作用一样,为所有参数估计方法指出了理论极限。参数估计定理还表明,本文提出的抽样后验概率估计方法是渐近最优的。随机估计的思想类比于香农信息论的随机编码,对提高参数估计的性能、降低估计方法的复杂性都具有重要意义。参数估计定理之于雷达系统的意义可类比于香农编码定理之于通信系统,可以作为最优参数估计问题的理论基础。

参数估计定理的证明意味着空间信息论的理论框架正在逐步建立。空间信息论统一了雷达和通信的信息论基础,必将对雷达、声纳和医学影像的系统理论和信号处理方法产生巨大的推动作用。空间信息与参数估计精度之间的关系有助于理解人类的学习与认知过程,对人工智能和机器学习相关的研究也有参考价值。

### 参考文献:

- [1] LIU F, MASOUIROS C, LI A, et al. MU-MIMO communications with MIMO radar: From co-existence to joint transmission [J]. *IEEE Transactions on Wireless Communications*, 2017, 17(4): 2755-2770.
- [2] LI B, PETROPULU A P. Joint transmit designs for coexistence of MIMO wireless communications and sparse sensing radars in clutter[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2017, 53(6): 2846-2864.
- [3] LIU F, MASOUIROS C, LI A, et al. MIMO radar and cellular coexistence: A power-efficient approach enabled by interference exploitation[J]. *IEEE Transactions on Signal Processing*, 2018, 66(14): 3681-3695.
- [4] CHENG Z, LIAO B, HE Z, et al. Spectrally compatible waveform design for MIMO radar in the presence of multiple targets [J]. *IEEE Transactions on Signal Processing*, 2018, 66(13): 3543-3555.
- [5] QIAN J, HE Z, HUANG N, et al. Transmit designs for spectral coexistence of MIMO radar and MIMO communication systems[J]. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018, 65(12): 2072-2076.
- [6] LIU Y, LIAO G, XU J, et al. Adaptive OFDM integrated radar and communications waveform design based on information theory[J]. *IEEE Communications Letters*, 2017, 21(10): 2174-2177.
- [7] HUANG K W, BIC M, MITRA U, et al. Radar waveform design in spectrum sharing environment: Coexistence and cognition [C]//*Proceedings of IEEE National Radar Conference*. Arlington, VA: IEEE, 2015: 1698-1703.
- [8] BICA M, KOIVUNEN V. Generalized multicarrier radar: Models and performance[J]. *IEEE Transactions on Signal Processing*, 2016, 64(17): 4389-4402.
- [9] LIU F, ZHOU L, MASOUIROS C, et al. Toward dual-functional radar-communication systems: Optimal waveform design[J]. *IEEE Transactions on Signal Processing*, 2018, 66(16): 4264-4279.
- [10] HASSANIEN A, AMIN M G, ZHANG Y D, et al. Signaling strategies for dual-function radar communications: An overview [J]. *IEEE Aerospace & Electronic Systems Magazine*, 2016, 31(10): 36-45.
- [11] HASSANIEN A, AMIN M G, ZHANG Y D, et al. Dual-function radar-communications: Information embedding using sidelobe control and waveform diversity[J]. *IEEE Transactions on Signal Processing*, 2016, 64(8): 2168-2181.
- [12] PAUL B, CHIRIYATH A R, BLISS D W. Survey of RF communications and sensing convergence research[J]. *IEEE Access*, 2017, 5(99): 252-270.
- [13] CHIRIYATH A R, PAUL B, JACYNA G M, et al. Inner bounds on performance of radar and communications co-existence [J]. *IEEE Transactions on Signal Processing*, 2015, 64(2): 464-474.

- [14] CHIRIYATH A R, PAUL B, BLISS D W. Radar-communications convergence: Coexistence, cooperation, and codesign[J]. *IEEE Transactions on Cognitive Communications & Networking*, 2017, 3(1): 1-12.
- [15] NEHORAI A, AKCAKAYA M, SEN S. Radar and information theory special issue information [EB/OL]. (2017-12-20) [2020-05-20]. [https://www.mdpi.com/journal/entropy/special\\_issues/radar\\_and\\_information\\_theory](https://www.mdpi.com/journal/entropy/special_issues/radar_and_information_theory).
- [16] WOODWARD P M, DAVIES I L. A theory of radar information[J]. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1950, 41(321): 1001-1017.
- [17] WOODWARD P. Theory of radar information[J]. *Transactions of the IRE Professional Group on Information Theory*, 1953, 1(1): 108-113.
- [18] WOODWARD P M, DAVIES I L. Information theory and inverse probability in telecommunication[J]. *Proceedings of the IEE-Part III: Radio and Communication Engineering*, 1952, 99(58): 37-44.
- [19] BELL M R. Information theory and radar: Mutual information and the design and analysis of radar waveforms and systems[D]. Pasadena: California Institute of Technology, 1988: 582.
- [20] BELL M R. Information theory and radar waveform design[J]. *Information Theory IEEE Transactions on*, 1993, 39(5): 1578-1597.
- [21] XU S, XU D, LUO H. Information theory of detection in radar systems[C]// *Proceedings of 2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*. Bilbao, Spain: IEEE, 2017: 249-254.
- [22] XU D, YAN X, XU S, et al. Spatial information theory of sensor array and its application in performance evaluation[J]. *IET Communications*, 2019, 13(15): 2304-2312.
- [23] 徐大专, 陈越帅, 陈月, 等. 雷达参数估计系统中目标位置和幅相信息量研究[J]. *数据采集与处理*, 2018, 33(2): 207-214. XU Dazhuan, CHEN Yueshuai, CHEN Yue, et al. Research on location and amplitude-phase information of target in radar detection system[J]. *Journal of Data Acquisition and Processing*, 2018, 33(2): 207-214.
- [24] ZHU S, XU D. Range-scattering information and range resolution of multiple target radar[C]//*Proceedings of International Conference on Communication and Information Systems*. Wuhan, China: [s.n.], 2019.
- [25] SHI C, XU D, ZHOU Y, et al. Range-DOA information in phased-array radar[C]//*Proceedings of IEEE 5th International Conference on Computer and Communications (ICCC)*. Chengdu, China: IEEE, 2019: 747-752.
- [26] ZHOU Y, XU D, TU W, et al. Spatial information and angular resolution of sensor array[J]. *Signal Processing*, 2020: 174.
- [27] 徐大专, 罗浩. 空间信息论的新研究进展[J]. *数据采集与处理*, 2019, 34(6): 941-961. XU Dazhuan, LUO Hao. New research progress in spatial information theory[J]. *Journal of Data Acquisition and Processing*, 2019, 34(6): 941-961.
- [28] LUO H, XU D, TU W, et al. Closed-form asymptotic approximation of target's range information in radar detection systems [J]. *IEEE Access*, 2020(99): 1.

## 作者简介:



徐大专(1963-), 通信作者, 男, 教授, 博士生导师, 研究方向: 空间信息论、信息与编码理论、宽带无线通信, E-mail: xudazhuan@nuaa.edu.cn。



屠伟林(1995-), 男, 硕士研究生, 研究方向: 空间信息论, E-mail: tuweilin@nuaa.edu.cn。



施超(1995-), 男, 硕士研究生, 研究方向: 空间信息论。



周颖(1996-), 女, 硕士研究生, 研究方向: 空间信息论。