

基于重要性抽样的最大似然方位估计方法

李 雄, 黄建国, 张群飞
(西北工业大学航海学院, 陕西西安 710072)

摘要: 最大似然估计是公认的最佳估计器,但是计算量很大.为了解决它的计算量大的问题,本文把蒙特卡罗方法与最大似然方位估计相结合,提出一种基于重要性抽样的最大似然方位估计新方法(Maximum Likelihood DOA Estimator Based on Importance Sampling,简称 ISMLE).研究结果表明,ISMLE方法不但保持了原最大似然方位估计方法的优良性能,而且大大减小了计算量,把原方法的计算复杂度从 $O(L^K)$ 减少到 $O(K \times H)$.

关键词: 方位估计; 最大似然估计; 蒙特卡罗; 重要性抽样; 计算复杂度

中图分类号: TN911.7 **文献标识码:** A **文章编号:** 0372-2112 (2005) 08-1529-04

Maximum Likelihood DOA Estimator Based on Importance Sampling

LI Xiong, HUANG Jian-guo, ZHANG Qun-fei

(College of Marine, Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China)

Abstract: Maximum Likelihood Estimator (MLE) has been shown to be the best performance in parameter estimation. However, the computation burden of MLE is very large. In order to resolve the question of computation burden, Monte Carlo methods are combined with maximum likelihood DOA estimator. A novel Maximum Likelihood DOA Estimator based on Importance Sampling (ISMLE) is proposed. ISMLE not only keeps the excellent performance of the original MLE, but also reduces the computation greatly, from the computational complexity $O(L^K)$ of original method to $O(K \times H)$.

Key words: DOA estimation; maximum likelihood estimator; monte carlo; importance sampling; computational complexity

1 引言

阵列信号处理的高分辨技术一直是国内外十分关注的研究热点,其中目标方位估计的高分辨技术是国内外研究的重心.该项技术的突破对声纳、雷达、地质勘探、生物医学工程等均具有重要意义.

自1979年由Schmidt等提出信号子空间方法以来,以MUSIC方法为典型代表的高分辨定向技术成为研究的热点.这类方法从空间分辨率上突破了波束形成系统的瑞利(Rayleigh)限的制约,这些方法在较理想的应用环境中不同程度取得了较好的估计性能,但是在恶劣条件下,特别是低信噪比下,这些方法性能会迅速降低.实际工程应用迫切需要分辨能力更高、性能更好的方位估计新方法.

最大似然估计器是公认的最佳估计器,不仅在高信噪比下性能逼近克拉美罗下界;而且在低信噪比下也有很好的性能^[1,2].但是最大似然估计需要进行多维网格搜索,计算量随着目标个数的增加呈指数增长,工程上难以实时应用.

为了解决最大似然估计因计算量很大而难以应用的问题,十分需要新的理论和方法,与最大似然估计方法相结合,以期获得新的突破.近年来,蒙特卡罗(Monte Carlo,简称MC)方法在统计信号处理领域的重要性日益突出,国外已经把它

应用于目标参数估计等领域,取得新的进展^[3-6].本文把MC方法重要性抽样(Importance Sampling,简称IS)与最大似然估计方法相结合,提出一种基于重要性抽样的最大似然方位估计新方法(ISMLE).和最大似然方位估计方法相比,新方法不仅十分显著地减少了计算量,而且保持了原最大似然方位估计方法的优越性能.

2 最大似然方位估计方法

K 个窄带信号源,频率分别为 f_1, f_2, \dots, f_K ,中心频率为 f_0 .假设不同快拍、不同阵元上的接收到的噪声都彼此独立,为零均值、方差 σ^2 的复高斯白噪声.阵列为均匀线列阵, M 个阵元,阵元间距为 b ,信号源分别以 $\theta_k(k=1, 2, \dots, K)$ 的入射角到达阵列的各个阵元,如图1所示.

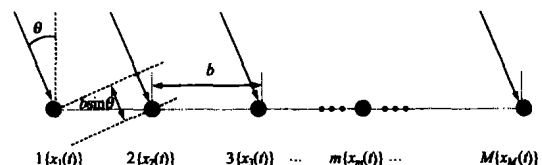


图1 均匀线列阵

设第1个阵元为参考阵元,相对于参考阵元,以 θ_k 为入射方向的第 k 个信号在第 m 个阵元上产生的时间延迟为

收稿日期:2004-09-05;修回日期:2005-01-28

基金项目:国家自然科学基金(No. 60272076);高等学校博士学科点专项科研基金

$m(k) = (m-1) \frac{b \sin \theta_k}{c}$, 其中 c 为信号的传播速度.

令 $X(t) = [x_1(t), x_2(t), \dots, x_M(t)]^T$, $n(t) = [n_1(t), n_2(t), \dots, n_M(t)]^T$, 其中 $x_m(t)$, $n_m(t)$ 分别是第 m 个阵元在 t 时刻接收到的信号和噪声. $S(t) = [\tilde{S}_1(t), \tilde{S}_2(t), \dots, \tilde{S}_K(t)]^T$, 其中 $\tilde{S}_k(t)$ 是第 k 个信号源在 t 时刻的值. 设 $a(\theta_k) = [1, \exp\left\{j2 f_k \frac{b \sin \theta_k}{c}\right\}, \dots, \exp\left\{j2 f_k (M-1) \frac{b \sin \theta_k}{c}\right\}]^T$, 表示未知的信号方位向量, $A(\theta) = [a(\theta_1), a(\theta_2), \dots, a(\theta_K)]$.

则阵列在 t 时刻的输出信号表示为^[7]

$$X(t) = A(\theta) S(t) + n(t) \quad (1)$$

下面研究确定性最大似然估计. 用阵列对空间的 K 个信号在等间隔的 N 个时刻做齐次采样, 由于 N 个采样 X_1, X_2, \dots, X_N 独立同分布, 则由式(1)和噪声的相关假设, 得到采样数据的联合概率密度函数为

$$p(X_1, X_2, \dots, X_N) = \frac{1}{\det(\Sigma)^N} \exp\left\{-\frac{1}{2} \sum_{t=1}^N [X(t) - AS(t)]^H \Sigma^{-1} [X(t) - AS(t)]\right\} \quad (2)$$

使上式最大, 由文献[1, 2, 7]得到参数 θ 的最大似然估计为

$$\hat{\theta} = \arg \min_{\theta} \text{Tr}\left\{(I - A(A^H A)^{-1} A^H) R\right\} = \arg \max_{\theta} \text{Tr}\left\{A(A^H A)^{-1} A^H R\right\} \quad (3)$$

其中, $R = \frac{1}{N} \sum_{t=1}^N X(t) X^H(t)$ 为采样协方差矩阵.

通过对 θ 进行多维网格搜索, 得到 K 个峰值对应的角度就是被估计的 K 个信号源的方位. 研究表明最大似然方位估计方法性能十分优良, 尤其是在低信噪比条件下, 要比子空间类方法性能好的多^[1, 2, 7]. 但是, 由于要进行多维网格搜索, 计算量非常大, 随着目标维数的增加计算量呈指数增长, 难以工程实时应用. 因此迫切需要研究它的快速算法. 采用 Monte Carlo 方法^[8]来解决最大似然估计方法计算量大的问题.

令 $L(\theta) = \text{Tr}\left\{A(A^H A)^{-1} A^H R\right\}$, 把它归一化, $\overline{L(\theta)} = \frac{L(\theta)}{L(\theta) d}$. 把 $\overline{L(\theta)}$ 视作随机变量, 则 $\overline{L(\theta)}$ 符合概率密度函数

(Probability Density Function, 简称 PDF) 的要求, 可以看作一种伪概率密度函数 (pseudo-PDF). 则方位参数 $\theta = (\theta_1, \theta_2, \dots, \theta_K)$ 的估计值为

$$\hat{\theta}_k = \dots \int \overline{L(\theta)} d\theta, k=1, 2, \dots, K \quad (4)$$

这是一个多维积分, 积分函数很复杂, 用数值方法很难得到积分的结果. 但是它可以看作是数学期望, 可以用 Monte Carlo 方法来求解. 采用直接抽样的 MC 方法从 $\overline{L(\theta)}$ 抽取样本, 得到样本向量 $\theta_h, h=1, 2, \dots, H$, 然后把样本均值作为的估计值. 但是由于 $\overline{L(\theta)}$ 很复杂, 是高维的非线性函数, 直接产生样本仍然要求多维积分^[8], 很难实现, 因此不能直接抽取样本. 采用重要性抽样方法可以解决这个问题.

3 重要性抽样

重要性抽样是经典 Monte Carlo 方法里最常用的抽样方

法. 基本原理如下^[8]

$$I = \int h(x) p(x) dx = \int h(x) \frac{p(x)}{g(x)} g(x) dx \quad (5)$$

其中 $p(x), g(x)$ 是概率密度函数. 则积分 I 估计为

$$\hat{I} = \frac{1}{H} \sum_{h=1}^H h(x_h) \frac{p(x_h)}{g(x_h)} \quad (6)$$

这里, x_h 是根据 $g(x)$ 分布抽取的第 h 个样本. 由文献[8]得到, 估计值 \hat{I} 是无偏的, 其方差取决于 $\frac{p(x)}{g(x)}$ 的选择. $g(x)$ 称为重要函数 (Importance Function, 简称 IF). 重要性抽样的关键是重要函数的选取. 为了减小估计的误差, 重要函数应该近似于原概率分布, 而且为了实现方便, 应该易于从其中抽取样本. 把 IS 方法应用于最大似然方位估计方法, 可以把从难以抽取样本的概率密度 $L(\theta)$ 转换为从重要函数进行抽样.

4 基于重要性抽样的最大似然方位估计方法

把重要性抽样和最大似然方位估计方法相结合, 研究该方法中重要函数的选取. 由前文得知, 参数 θ 的最大似然估计为

$$\hat{\theta} = \arg \max_{\theta} \text{Tr}\left\{A(A^H A)^{-1} A^H R\right\} \quad (7)$$

等价于

$$\hat{\theta} = \arg \max_{\theta} \text{Tr}\left\{MA(A^H A)^{-1} A^H R\right\} \quad (8)$$

$$\begin{aligned} \text{令 } L_1(\theta) &= \exp\left(\text{Tr}\left\{MA(A^H A)^{-1} A^H R\right\}\right) \\ &= \exp\left\{\frac{M}{N} \text{Tr}\left\{A(A^H A)^{-1} A^H \sum_{t=1}^N X(t) X^H(t)\right\}\right\} \end{aligned} \quad (9)$$

设 $Y(t) = [X_1(t), X_2, \dots, X_N(t)]^T$ 则 $R = \frac{1}{N} \sum_{t=1}^N X(t) X^H(t)$

$X^H(t) = \frac{1}{N} Y(t) Y^H(t)$, 故

$$\begin{aligned} L_1(\theta) &= \exp\left\{\frac{M}{N} \text{Tr}\left\{A(A^H A)^{-1} A^H \sum_{t=1}^N X(t) X^H(t)\right\}\right\} \\ &= \exp\left\{\frac{M}{N} \text{Tr}\left\{A(A^H A)^{-1} A^H Y(t) Y^H(t)\right\}\right\} \end{aligned} \quad (10)$$

由式(10)可知, 若取 $(A^H A)^{-1} = I/M$, 其中 I 是 $k \times k$ 维单位矩阵 (实际上, 如果目标的方位角间隔比较大时, 该等式是成立的), 则上式化简为

$$\begin{aligned} L_2(\theta) &= \exp\left\{\frac{M}{N} \text{Tr}\left\{A \left(\frac{I}{M}\right) A^H Y(t) Y^H(t)\right\}\right\} \\ &= \exp\left\{\frac{1}{N} \text{Tr}\left\{AA^H Y(t) Y^H(t)\right\}\right\} \\ &= \exp\left\{\frac{1}{N} \text{Tr}\left\{A^H Y(t) Y^H(t) A\right\}\right\} \\ &= \exp\left\{\frac{1}{N} \left| \sum_{k=1}^K \sum_{n=1}^N \sum_{m=0}^{M-1} Y_m(n) \exp(-j m 2 \frac{1}{2} \sin(\theta_k)) \right|^2\right\} \end{aligned} \quad (11)$$

其中, $Y_m(n)$ 表示 $Y(t)$ 的第 m 行第 n 列元素.

令 $\phi = \frac{1}{2} \sin(\theta)$, $\phi = (\phi_1, \phi_2, \dots, \phi_K)$, $I(\phi) = \frac{1}{N} \sum_{n=1}^N \left| \sum_{m=0}^{M-1} Y_m(n) \exp(-j m 2 \phi) \right|^2$. 可以看出, $I(\phi)$ 是波束形成的空间谱^[9]. 则 $L(\theta) = L(\phi) \Big|_{\phi = \frac{1}{2} \sin(\theta)}$. 由式(11)得到

$$L_1(\theta) = \exp\left\{\frac{1}{N} \left| \sum_{k=1}^K \sum_{n=1}^N \sum_{m=0}^{M-1} Y_m(n) \exp(-j m 2 \phi_k) \right|^2\right\}$$

$$\begin{aligned}
 &= \exp \left\{ \sum_{k=1}^K \left[\frac{1}{N} \sum_{n=1}^N \left| \sum_{m=0}^{M-1} Y_m(n) \exp(-j m 2 \phi_k) \right|^2 \right] \right\} \\
 &= \exp \left(\sum_{k=1}^K I(\phi_k) \right) = \prod_{k=1}^K g(\phi_k) \quad (12)
 \end{aligned}$$

其中, $g(\phi_k) = \exp(I(\phi_k))$. 从上式可以看出, $L_1(\phi)$ 里的每一个 ϕ_k 都被分离出来, 联合概率密度分布可以写成边缘分布的乘积. ϕ_k 的边缘分布都是相同的, 为 $g(\phi) = \exp(I(\phi))$. 选择该分布为重要函数. 因此产生一个样本向量 ϕ , 转换为依据该分布产生 K 个独立的样本 ϕ .

由于似然函数 $L(\cdot)$ 是一个多峰值的函数, 有多个局部极值, 只有一个全局最大值(考虑二维情形, 假设 $\theta_1 < \theta_2$, 则 $\phi_1 < \phi_2$, 因此是单峰的), 如图 2 所示:

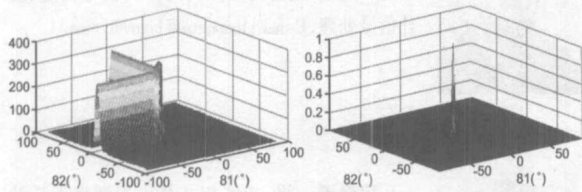


图 2 似然函数 $L(\theta)(\theta_1 < \theta_2)$

图 3 指数归一化后的似然函数 $\rho_1=0.75$

而我们的目的是寻找全局最大值. 但是全局最大值相对局部极值不突出, 为了使样本均值逼近于全局最大值, 全局最大值要相对局部极值更加突出, 曲面要更加尖锐, 同时为了避开数值计算困难, 把它指数归一化如下,

$$\begin{aligned}
 L_2(\phi) &= \exp(L_1(\phi)) / \exp(\max(L_1(\phi))) \\
 &= \exp(L_1(\phi) - \max(L_1(\phi))) \quad (13)
 \end{aligned}$$

比较图 2、图 3, 可以看出, 指数归一化后全局最大值明显突出, 曲面变得非常尖锐.

相应的对重要函数也进行指数化, $g_1(\phi) = \exp(L_2(\phi))$, 得到归一化的重要函数

$$\overline{g_1(\phi)} = \frac{g_1(\phi)}{\int g_1(\phi) d\phi} = \frac{\exp(L_2(\phi))}{\int \exp(L_2(\phi)) d\phi} \quad (14)$$

样本依据该概率密度产生.

根据文献[4~6]中的方法, 把计算样本的线性均值转换为计算样本的循环均值, 如下. 因为角度和幅值无关, 所以舍去式中两个很复杂的归一化积分常量. 这样处理不仅简化了计算, 而且减小了估计的误差.

$$\begin{aligned}
 \phi_k &= \frac{1}{2} \frac{1}{H} \sum_{h=1}^H \frac{L_2(\phi_h)}{g_1(\phi_h(k))} \exp(j2 \phi_h(k)) \\
 &= \frac{1}{2} \frac{1}{H} \sum_{h=1}^H \frac{L_2(\phi_h)}{g_1(\phi_h(k))} \exp(j2 \phi_h(k)), \quad k=1, 2, \dots, K \quad (15)
 \end{aligned}$$

其中, $\phi_h, h=1, 2, \dots, H$ 是产生的 H 个样本向量, $\phi_h(k), k=1, 2, \dots, K$ 是 ϕ_h 的第 k 个元素.

现把基于重要性抽样的最大似然方位估计方法的具体实现步骤归纳如下:

(1) 根据文献[4]中的方法, 由 $\overline{g_1(\phi)}$ 产生 H 个样本向量 $\phi_h, h=1, 2, \dots, H$.

(2) 把产生的样本向量 ϕ_h , 代入式(15), 得到 ϕ_k 的估计值 $\hat{\phi}_k, k=1, 2, \dots, K$.

(3) 由 $\hat{\theta}_k = \arcsin(2\hat{\phi}_k), k=1, 2, \dots, K$, 得到方位估计值.

5 仿真性能分析

仿真模型: 均匀线列阵, 阵元数为 12, 快拍数为 100, 两个目标方位角分别为 $\pm 2^\circ$. 样本数为 5000, 取 $\rho_1 = 30, \rho_2 = 0.009$. 做 100 次独立统计实验, 得到不同信噪比下 ISMLE 方法的估计性能, 并和原 MLE、MUSIC、MiniNorm 方法以及 CRB 比较.

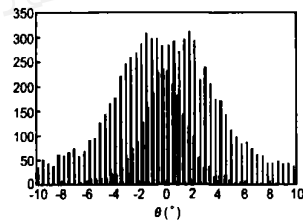


图 4 样本分布图

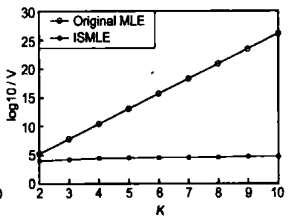


图 5 计算量的比较

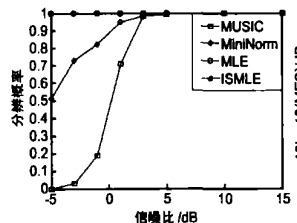


图 6 分辨概率

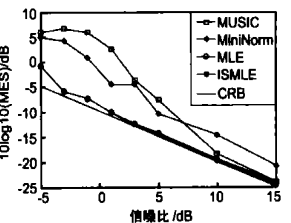


图 7 均方误差

产生样本的频数直方分布如图 4. 从图中可以看出, 样本主要集中在真值 ± 2 附近.

从图 6、7 可以看出, 不管是高信噪比, 还是低信噪比条件下, MLE、ISMLE 方法的估计性能都要比 MUSIC、MiniNorm 方法好. ISMLE 方法保持了原 MLE 方法的优良性能, 分辨概率、均方误差都逼近于原 MLE 方法, 高信噪比下达到 CRB, 尤其是低信噪比条件下比 MUSIC、MiniNorm 方法的性能要好. 而且和原 MLE 方法相比, ISMLE 方法的计算量大大减小. 原 MLE 方法采用 K 维搜索, 计算复杂度为 $O(L^K)$; 而 ISMLE 方法的计算量主要是样本的产生和样本权值的计算. 由于采用黄金分割搜索方法^[10]产生样本, 迭代收敛很快, 计算复杂度近似为 $O(K \times H)$, 其中 H 是样本的个数. H 随目标个数 K 的增加变化不大. 两种方法的计算量随目标源的增加而变化的曲线比较见图 5, 取 $L=400, H=5000$. 从图 5 可以看出, 原 MLE 方法的计算量随着目标个数的增加呈指数增长, 而 ISMLE 方法的计算量随着目标源的个数的增加线性变化, 因此 ISMLE 方法和原 MLE 方法相比, 在保持它的优良性能的同时, 大大减小了计算量.

6 结论

针对最大似然估计方法由于多维搜索而导致的计算量大, 难以工程应用的问题, 利用蒙特卡罗方法, 把重要性抽样与最大似然方位估计方法相结合, 提出了基于重要性抽样的最大似然方位估计新方法 (ISMLE). 给出了完整的理论推导

过程,重点研究了重要函数的选取,并进行了仿真性能研究.结果表明,ISMLE方法不但保持了MLE方法的优良性能,高信噪比下逼近于CRB下界,低信噪比条件下性能明显比MUSIC、MiniNorm方法好;而且把MLE方法的计算复杂度从 $O(L^K)$ 减少到 $O(K \times H)$,大大减少了计算量,为最大似然方位估计方法的工程实现提供了一种新途径.

参考文献:

- [1] Petre Stoica ,Arye Nehorai . MUSIC ,Maximum likelihood ,and Cramer-Rao Bound [J] . IEEE Transactions on Acoustic , Speech , and Signal Processing ,1989 ,37(5) :720 - 740 .
- [2] Petre Stoica ,Arye Nehorai . MUSIC ,Maximum likelihood ,and Cramer-Rao Bound :further results and comparisons [J] . IEEE Transactions on Acoustic , Speech , and Signal Processing ,1990 ,38(12) :2140 - 2150 .
- [3] S M Kay ,Supratim Saha . Design of sparse linear array by Monte Carlo importance sampling [A] . Proc Oceans 2000 MTS/ IEEE [C] . Rhode Island ,USA :IEEE ,2000 . 1501 - 1507 .
- [4] S M Kay ,Supratim Saha . Mean likelihood frequency estimation [J] . IEEE Transactions on Signal Processing ,2000 ,48(7) :1937 - 1946 .
- [5] Supratim Saha . Optimal estimation of nonlinear and linear parameters in general linear models [D] . USA :University of Rhode Island Ph. D. dissertation . 2001 . 43 - 131 .
- [6] S M Kay ,Supratim Saha . Maximum likelihood parameter estimation of superimposed chirps using Monte Carlo Importance sampling [J] . IEEE Transactions on Signal Processing ,2002 ,50(2) :224 - 230 .
- [7] 孙毅 . 基于贝叶斯原理和蒙特卡罗方法的高分辨方位估计新方法研究 [D] . 西安 :西北工业大学硕士学位论文 ,2003 . 39 - 41 .
- [8] 茆诗松 ,王静龙 ,濮晓龙 . 高等数理统计 [M] . 北京 :高等教育出版社 ,2000 . 400 - 428 .
- [9] 陈建峰 . 水下目标定向的高分辨技术研究 [D] . 西安 :西北工业大学硕士学位论文 ,1996 . 12 - 15 .
- [10] 《现代应用数学手册》编委会 . 现代应用数学手册 [M] . 北京 :清华大学出版社 ,1998 . 103 - 105 .

作者简介:



李雄男,1980年9月生于湖北省应城市,2002年毕业于西北工业大学,获学士学位,现为西北工业大学航海学院信号与信息处理专业硕士研究生,主要研究方向为阵列信号处理、统计信号处理. E-mail:lixionga@hotmail.com.



黄建国,男,1945年2月生于湖南省长沙市,西北工业大学航海学院教授、博士生导师,IEEE西安支分会主席,长期从事水下阵列处理、现代信号处理和水声通信等方面的研究工作,发表论文120多篇.

张群飞,男,1968年生于浙江,西北工业大学副教授,从事阵列信号处理和计算机仿真等方面的科研工作,发表学术论文14篇.