

# 异常值剔除法在 RCS 测量数据处理中的应用

## 程 度

(解放军 91404 部队 河北 秦皇岛 066001)

**摘要:**对目标的雷达截面(RCS)测量得到大量原始数据,为了对目标的散射特性作进一步分析,必须对这些原始数据进行简化和处理,得到能概括目标散射特征的数据。本文应用格拉布斯检验方法和卡尔曼滤波方法分别对静态和动态测试数据的异常值进行预先判别、剔除处理,然后再对原始数据进行平滑处理和统计分析,提高了处理结果的置信度,更准确地表征目标的 RCS 特性。

**关键词:**雷达截面;数据平滑处理;概率密度函数;累积分布函数

中图分类号:TN95

文献标识码:B

文章编号:1004-373X(2007)19-082-03

## Application of the Method of Outliers Removing in RCS Measurement Data Processing

CHENG Du

(PLA Unit 91404, Qinhuangdao, 066001, China)

**Abstract:** There are a large amount of original data after RCS measuring. In order to analyze the scattering trait of the target, we should simplify and process these data to obtain the data which generalize the scattering trait of the target. This paper applies the methods of Grubbs test and Kalman filter to pre-process the original data to remove the outliers, and then make smoothing process and statistical analysis. In this way, the confidence of the result is promoted and the RCS trait of the target is represented more accurately.

**Keywords:** RCS; data smoothing processing; PDF; CDF

目标的 RCS 测量对了解目标的基本散射现象,取得目标的特征数据等,都具有十分重要的作用。对于典型的测试程序,可以得到不同目标在不同俯仰角、横滚角和不同频率、相位下的大量数据。每一组数据均以数据对(角度、RCS)的形式出现,一般是在部分的或完整的方位区域上每隔  $0.1^\circ$  采样一个数据。原始数据可以真实而形象地反映目标的 RCS 属性和特性,但是对于高频复杂目标,散射图中的各个波瓣相隔可能小于  $0.1^\circ$ ,并且测量的散射图通常是由图纸上的一条黑迹带组成,一个测试程序中九十几百次试验就有几百张图。因此,在有些时候,测试试验的均值、中值、标准差以及测试模型各参量构成的统计学特性更有意义。对测量数据中异常值的预先剔除,将提高分析的准确度。

### 1 数据异常值分析

无论是静态测量还是动态测量,由于测量设备本身、数据传输或者人工操作等原因,都可能使测量数据中包含某些错误的测量,工程上常称为异常值。如果不把测量数据中的这些异常值预先剔除,将使数据处理带来很大的误差。

#### 1.1 静态测量中的异常值判别方法

对静态测量,零假设下观测量  $Y_1, Y_2, \dots, Y_n$  为来自某

分布的随机样本,这里仅考虑正态分布情形,此时的  $Y_i$  ( $1 \leq i \leq n$ ) 分布为  $N(\theta, \sigma^2)$ ,即:

$$Y_i = \theta + V_i \quad (1)$$

式中,  $\theta$  为被测量的真值,  $V_i$  的零分布为  $N(0, \sigma^2)$ 。

把观测量  $Y_1, Y_2, \dots, Y_n$  按取值从小到大排列,记对应的顺序统计量为  $Z_1, Z_2, \dots, Z_n$ ,显然,如果  $Y_1, Y_2, \dots, Y_n$  中存在异常值,必定首先怀疑  $Z_1$  或  $Z_n$ ,其次怀疑接近这两者的那些,因为他们比其他观测量偏离均值  $\theta$  要远。

##### 1.1.1 单个异常值情形

###### 格拉布斯(Grubbs)检验法

首先考虑单侧情形,如已知异常值出现在右边(左边类似)此时零假设  $H_0: Z_n$  为正常观测量;备选假设  $H_1: Z_n$  为异常值。

假定  $H_0: Y_i \sim N(\theta, \sigma^2)$ ,  $H_1: Y_i \sim N(\mu, \sigma^2)$ ,这里  $\mu > \theta$ 。构造统计量:

$$B = (Z_n - \bar{Z})/S = \max_{1 \leq i \leq n} (Y_i - \bar{Y})/S \quad (2)$$

$$\text{其中: } \bar{Z} = \bar{Y} = \frac{1}{n} \sum_{i=1}^n Y_i, S = \sqrt{\frac{\sum_{i=1}^n (Z_i - \bar{Z})^2}{n-1}}$$

$$\text{记: } B_i = (Y_i - \bar{Y})/S, \quad (1 \leq i \leq n)$$

对于给定的显著性水平  $\alpha$ ,可查得  $B$  的  $1 - \alpha$  分位数  $B_{1-\alpha}$ 。作检验  $B > B_{1-\alpha}$ 。如成立,判断  $Z_n$  为异常值;如不成立,判断  $Z_n$  为正常观测量。

对于双侧情形,假定  $H_0: Y_i \sim N(\theta, \sigma_0^2)$ ,  $H_1: Y_i \sim$

$N(\mu, \sigma_1^2)$ , 其中  $\sigma_1^2 > \sigma_0^2$ , 构造统计量:

$$\begin{aligned} B^* &= \max_{1 \leq i \leq n} |Y_i - \bar{Y}| / S \\ B_{1-\alpha}^* &= B_{1-\alpha/2} \end{aligned} \quad (3)$$

根据给定的显著性水平  $\alpha$ , 可查得  $B_{1-\alpha}^*$  作检验:

$$B^* > B_{1-\alpha}^*$$

如成立, 若  $|Z_1 - \bar{Z}| > |Z_n - \bar{Z}|$ , 判断  $Z_1$  为异常值, 若  $|Z_n - \bar{Z}| > |Z_1 - \bar{Z}|$ , 判断  $Z_n$  为异常值。如上式不成立, 判断无异常值。

### 1.1.2 多个异常值情形

对于异常值个数  $K$  未知的情形, 用检验单个异常值的方法, 首先检验  $Y_1^*$  (使  $|Y_1 - \bar{Y}|$  达极大) 是否异常值, 如是则剔除  $Y_1^*$ ; 第二步对剩余  $n-1$  个观测量, 检验  $Y_{n-1}^*$  是否异常值, 如是则剔除  $Y_{n-1}^*$ ; 如此继续下去, 直至检验到某个  $Y_{n-k}^*$  非异常值为止。

### 1.2 动态测量中异常值的递推判别方法

判别动态测量数据中的异常值与静态情形有很大不同, 后者观测量是独立同分布的, 其中只有接近最大或最小顺序统计量的那些观测量才可能是异常值, 但动态测量中, 观测量  $Y_1, Y_2, \dots, Y_n$  的真值  $x_1, x_2, \dots, x_n$  随时间而变化, 对真值  $x_i$  的变化规律常常用非完全精确的动态模型描述, 在这种情况下, 需要首先用观测量估计  $x_i$ , 然后从观测量  $Y_i$  中分离出测量误差  $V_i$ , 再进行异常值判别。

卡尔曼滤波法:

该方法首先在观测序列的初始找到  $m$  个以上彼此相吻合的观测量, 用他们求得卡尔曼滤波的初值, 然后进行卡尔曼滤波的逐步递推, 在递推的每一步利用预报残差这个统计量对步的观测量进行判别。

记  $k+1$  时刻的观测量为  $Y_{k+1}$ , 是  $m \times 1$  维向量;  $k+1$  时刻状态量为  $X_{k+1}$ , 是  $n \times 1$  维向量, 设量测方程为:

$$Y_{k+1} = B_{k+1} X_{k+1} + V_{k+1} \quad (4)$$

式中:  $B_{k+1}$  表示  $m \times n$  维量测矩阵,  $V_{k+1}$  为量测噪声, 在零假设下,  $\{V_{k+1}\}$  是均值为零, 相互独立, 有方差矩阵  $\{R_{k+1}\}$  的高斯随机量,  $V_{k+1}$  的第  $i$  个分量的方差为  $\sigma_{k+1}^2(i)$ , 对状态  $X_{k+1}$  的预报值为  $\hat{X}_{k+1/k}$ , 预报残差为  $V_{k+1}$ , 有:

$$V_{k+1} = Y_{k+1} - B_{k+1} \hat{X}_{k+1/k} \quad (5)$$

利用预报残差, 可对  $Y_{k+1}$  的每个分量进行判别, 判别式为:

$$|V_{k+1}(i)| \leq c \sqrt{(B_{k+1} P_{k+1/k} B_{k+1}^T)_{ii} + \sigma_{k+1}^2(i)} \quad (6)$$

式中  $P_{k+1/k}$  表示预报误差方差矩阵:

$$P_{k+1/k} = E[(\hat{X}_{k+1/k} - x_{k+1})(\hat{X}_{k+1/k} - x_{k+1})^T] \quad (7)$$

式中下角  $i, i$  表示取矩阵对角线上的第  $i$  个元,  $V_{k+1}(i)$  表示  $V_{k+1}$  的第  $i$  个分量;  $c$  为单位正态随机量的分位数, 根据显著性水平  $\alpha$  确定, 实际应用中可取 3 或 4。

式(6)成立, 判断  $Y_{k+1}(i)$  正常, 否则为异常。

## 2 数据平滑处理

RCS 测量的目的之一是决定一个给定雷达对该目标的可探测性。但是检测是一个统计过程, 多数检测分析对目标起伏采用一个统计模型, 而对不希望的杂散信号, 如: 杂波、多路径和噪声等, 则采用一种统计描述。典型跟踪雷达天线在一个(和波束)或两个(差波束)角度方向上有很高的增益, 相应的波宽度很窄, 典型值  $1^\circ \sim 2^\circ$ , 而雷达接收机则响应于该波束宽度窗口内平均的目标回波。因此, 需要将初始数据在一个窄的角度窗口内进行平滑。为获得平滑数据的均值, 首先需要决定“窗口宽度”, 他限定了用来进行 RCS 平均的角度范围和数据点的多少。典型的窗口宽度范围是  $1^\circ \sim 10^\circ$ , 这取决于散射图的起伏特性。选择窗口宽度原则上至少应包括几个波瓣, 也可以按已有的经验来确定。算术平均值:

$$\bar{\sigma} = \frac{1}{N} \sum_{n=1}^N \sigma_n \quad (8)$$

就是该窗中点的平滑均值,  $\sigma_n$  的单位是  $m^2$ 。

窗口宽度确定后就需要选择“滑动阀”, 他表示窗口中心每次移动的方位角。滑动阀决不应大于窗口宽度, 否则将产生间歇而不能进行平均。常用的滑动阀是  $1^\circ$  或  $2^\circ$ 。更小的滑动阀可产生精细的平滑散射图, 但要求更多的处理时间。

## 3 数据统计分析

每一个简化了的数据可以采取若干种形式: 在一个规定的窗口和滑动阀内按 3 个百分位点的电平(90, 50 和 10)进行平滑数据; 在 3 个规定的角度区域内, 以分贝数(对数空间)表示中值、均值和标准差的区段数据; 以平方米(线性空间)表示均值和标准差数据; 每个区段的概率密度函数(PDF)和累积分布函数(CDF)。

### 3.1 区段均值

区段均值是对给定角扇形范围内的原始数据进行数学平均。对每一对数据(角、RCS)进行检查以确定他是否出现在一个扇形之内。利用计算算术平均值式(8)计算均值, 式中  $N$  是  $\sigma_n$  在规定扇形内出现的总次数。

### 3.2 区段标准差

每个区段的标准差 STD 为:

$$STD = \left[ \frac{\sum_{n=1}^N (\sigma_n - \bar{\sigma})^2}{N-1} \right]^{1/2} \quad (9)$$

式中  $\bar{\sigma}$  为式(8)给出的区段均值。若  $N$  为较大的数, 故  $\frac{1}{N-1} \triangleq \frac{1}{N}$ , 因此式(9)可展开为:

$$\begin{aligned} STD &= \left[ \frac{1}{N} \sum (\sigma_n^2 - 2\bar{\sigma}\bar{\sigma} + \bar{\sigma}^2) \right]^{1/2} \\ &= \left[ \frac{1}{N} \sum \sigma_n^2 - 2\bar{\sigma} \frac{1}{N} \sum \sigma_n + \frac{1}{N} \sum \bar{\sigma}^2 \right]^{1/2} \end{aligned}$$

$$= [\bar{\sigma}^2 - 2\bar{\sigma}\bar{\sigma} + \bar{\sigma}^2]^{1/2} = [\bar{\sigma}^2 - \bar{\sigma}^2]^{1/2} \quad (10)$$

可见,标准差是各数据平方的均值减去均值平方后的平方根。STD 越大,说明数据的起伏或离散性越大。

### 3.3 概率密度函数

概率密度函数(PDF)是一种统计分布表示,他表示了雷达截面为给定电平  $\sigma_0$  的出现概率  $P(\sigma)$ 。描述了  $\sigma$  位于  $\sigma_0$  和  $\sigma_0 + d\sigma$  之间的概率是  $P(\sigma)d\sigma$ ,即:

$$F(\sigma) = \int_{\sigma_0}^{\sigma_0 + d\sigma} P(\sigma)d\sigma = P(\sigma)d\sigma \quad (11)$$

式中,积分号内的  $P(\sigma)$  是概率密度函数,  $F(\sigma)$  是一个  $\leq 1$  的正量。位于  $-\infty \leq \sigma \leq \infty$  内的  $\sigma$  的概率为 1,这就要求:

$$\int_{-\infty}^{\infty} P(\sigma)d\sigma = 1 \quad (12)$$

概率密度函数的单位为面积的倒数(1/m<sup>2</sup>)。PDF 曲线可以表明一组给定的数据与目标模拟分析中使用的某一标准分布函数(例如对数正态分布、瑞利分布、Weibull 分布等)相符合的程度。

### 3.4 累积分布函数

累积分布函数(CDF)为雷达截面低于或等于某个值  $\sigma$  的概率,他是由概率密度函数曲线积分而得。因为位于  $\sigma_1$  和  $\sigma_2$  之间的数据出现的概率由式(13)给出为:

$$\sigma_1 \leq \sigma \leq \sigma_2 \text{ 的概率} = \int_{\sigma_1}^{\sigma_2} P(\sigma)d\sigma \quad (13)$$

由式(13),令  $\sigma_1 = -\infty, \sigma_2 = \sigma$ ,便可得到低于  $\sigma$  的数据出现的概率:

作者简介 程 度 男,1966 年出生,硕士研究生,工程师。主要从事通信工程、电子测试与计量方面的研究。

(上接第 81 页)

(2) 角度测量误差方差增大,定位精度下降较大,说明角度误差对精度产生主要影响,如图 4(a),图 4(b) 所示。

(3) 距离和测量误差方差增大,定位精度变化不明显,说明定位精度对他的变化不敏感,如图 4(a),图 4(c) 所示。

## 5 结语

在电子战中,采用无源定位的方法对运动辐射源进行定位具有重要意义。本文介绍了一种基于外辐射源的固定单站无源定位系统中对目标定位的工程实现方法,利用比幅测角,能够实现瞬时测角。再结合时间延迟  $\Delta t$ ,便可测得目标和雷达站之间的距离。最后通过对定位精度的

作者简介 赵 勇 男,1978 年出生,陕西西安人,长安大学信息工程学院电子信息实验室助工。

隋功浩 男,1979 年出生,山东青岛人,西安电子科技大学电子工程学院硕士研究生。

王 俊 男,1969 年出生,贵州毕节人,西安电子科技大学电子工程学院教授。

赵 婕 男,1979 年出生,陕西西安人,西安邮电学院教务处助工。

$$CDF(\sigma) = \int_{-\infty}^{\sigma} P(\sigma)d\sigma \quad (14)$$

他代表了从  $-\infty$  到  $\sigma$  区间内概率密度函数曲线下面的面积。

### 3.5 中 值

中值为  $CDF(\sigma_{0.5}) = 0.5$  的雷达截面值  $\sigma_{0.5}$ ,即是说比  $\sigma_{0.5}$  高或低的数据出现的概率各占一半。因此在计算 CDF 的同时,也就计算了中值。

## 4 结语

数据处理的要求随用户需要而变,一般所要求的 RCS 数据的形式包括:原始数据、平滑数据、区段数据、统计数据以及中值、均值和标准差的单个数据特征。经过对测量结果的分析表明,在静态及动态测量中预先对异常值进行分析和剔除,有效地排除了杂波及各种干扰因素对测量的影响,得到的数据能更科学合理地反映目标的真实散射特性。

## 参 考 文 献

- [1] [美]Knott E F, Tuely M T, Shaeffer J F. 雷达散射截面 [M]. 阮颖铮, 陈海, 译. 北京: 电子工业出版社, 1988.
- [2] 阮颖铮. 雷达截面与隐身技术 [M]. 北京: 国防工业出版社, 1998.
- [3] 何友. 雷达数据处理及应用 [M]. 北京: 电子工业出版社, 2006.
- [4] 费业泰. 误差理论与数据处理 [M]. 北京: 机械工业出版社, 2005.

分析,揭示了基于外辐射源雷达系统的定位特性。在实际的实验系统中,上述定位方法取得了良好的效果。

## 参 考 文 献

- [1] 孙仲康,周一宇,何黎星. 单多基地有源无源定位技术 [M]. 北京: 国防工业出版社, 1996.
- [2] 刘洪艳, 邵丽鹏, 司锡才. 比幅比相测角系统信号处理技术研究 [J]. 应用科技, 2003, 30(10): 7-9.
- [3] 杨振起, 张永顺, 洛永军. 双(多)基地雷达系统 [M]. 北京: 国防工业出版社, 2001.
- [4] 朱家兵, 陶亮. 一种非合作照射源的无源雷达目标检测方法 [J]. 合肥工业大学学报: 自然科学版, 2006, 29(1): 14-17.
- [5] 王俊, 保铮, 张守宏. 无源探测与跟踪雷达系统技术及其发展 [J]. 雷达科学与技术, 2004, 2(3): 129-135.

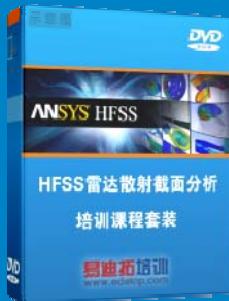
## 雷达散射截面 (RCS) 分析培训课程

易迪拓培训([www.edatop.com](http://www.edatop.com))由数名来自于研发第一线的资深工程师发起成立, 致力和专注于微波、射频、天线设计研发人才的培养, 是国内最大的微波射频和天线设计人才培养基地。客户遍布中兴通讯、研通高频、国人通信等多家国内知名公司, 以及台湾工业技术研究院、永业科技、全一电子等多家台湾地区企业。

雷达散射截面 (Radar Cross Section, 简称 RCS) 是雷达隐身技术中最关键的概念, 也是电磁理论研究的重要课题, 使用 HFSS 软件可以很方便的分析计算各种目标物体的 RCS。

由易迪拓培训推出的《HFSS 雷达散射截面分析培训课程套装》是从零讲起, 系统地向您讲授如何使用 HFSS 软件进行雷达散射截面分析的全过程。该套视频课程由专家讲授, 边操作边讲解, 直观易学。

### HFSS 雷达散射截面分析培训课程套装



套装包含两门视频培训课程, 其中: 《两周学会 HFSS》培训课程是作为 HFSS 的入门培训课程, 帮助您在最短的时间内迅速熟悉、掌握 HFSS 的实际操作和工程应用; 《HFSS 雷达散射截面(RCS)分析》培训课程是专门讲授如何使用 HFSS 来分析计算雷达散射截面, 包括雷达散射截面、单站 RCS、双站 RCS 等的定义, 实例讲解使用 HFSS 分析单站 RCS、双站 RCS 和宽频 RCS 的相关设置和实际操作等。视频课程, 专家讲授, 从零讲起, 直观易学...

课程网址: <http://www.edatop.com/peixun/hfss/130.html>

### ● 更多培训课程:

#### ● HFSS 培训课程

网址: <http://www.edatop.com/peixun/hfss/>

#### ● CST 培训课程

网址: <http://www.edatop.com/peixun/cst/>

#### ● 天线设计培训课程

网址: <http://www.edatop.com/peixun/antenna/>