国防科技大学学报
JOURNAL OF NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY
第10卷第1期 1988年8月 Vol.10No.1

雷达测速中的Robust技术

万建伟 皇甫堪

（电子技术系）

摘要 精确测定初速仍然是弹道测量最重要的任务之一。现代多卜勒雷达尽管很实用，但对于处理实际数据还存在许多缺陷。本文运用Robust估计的现代统计法，通过改进原某测速雷达中的数据处理部分，达到即使在最不利的条件下也可以计算出其精度能估计的初速度。使用同一方法还可以确定阻力系数。实验结果证实了这些结论。

关键词 Robust估计，雷达，数据处理，蒙特卡罗

1. 引 言

用于测量初速的现代多卜勒雷达对每发弹初始弹道进行一组测量，然后通过外推方法计算初速。我国某测速雷达，其数据处理的基本过程是：1> 将弹丸的径向速度折算到炮口方向；2> 将弹道上N个点的速度利用最小二乘拟合平滑外推，求得炮口的初速；3> 进行各种参数修正，并计算或然误差。数据处理完毕后，给出结果。其中存在的问题有：

1） 假定了误差U_i服从正态分布，而实际中U_i并不是精确的正态分布；

2） 最小二乘拟合对这种分布的不精确性所引起的偏差很大；

3） 测试中，组与组之间的数据是不关联的，这样就必须打多组炮弹来精确确定炮的性能。

上述问题如果得到解决，不仅在理论上，而且在军事应用价值上有很大意义，而用Robust技术就能较好处理这些问题。

本文第二部分对Robust估计理论进行一些研究，第三部分介绍了Robust技术用于雷达测速的数学原理，第四部分给出实验结果和结论。

2. Robust估计

将模型写为

1987年5月5日收到

$$y_i=\sum_{j=1}^{p}x_{ij}\beta_j+e_i \qquad i=1,\cdots,n \tag{1}$$

这里，$\{x_{ij}\}$是已知系数，$\{y_i\}$表示观察值，$i=1,\cdots,n,j=1,\cdots p$；$\beta_1,\cdots,\beta_p$为未知的回归系数，$e_1,\cdots,e_n$独立同分布、均值为零的误差。最小二乘法是找$\beta_1,\cdots,\beta_p$，使表达式：$\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2$达到最小。如前面指出过的,这个表达式及由此所决定的估计，受少数"离群值"($y_1,\cdots,y_n$中的离群值）的影响过大。为了减少这种离群值的作用，我们可以取一个其增长度比x^2慢的函数$\rho(x)$来代替x^2，即作表达式：

$$D(\beta_1,\cdots,\beta_p)=\sum_{i=1}^{n}\rho\left(y_i-\sum_{j=1}^{p}x_{ij}\beta_j\right) \tag{2}$$

而去找$\beta_1,\cdots,\beta_p$，使之达到最小。这个考虑引导得如下定义：

定义　设函数$\rho(x)$定义于$-\infty<x<+\infty$，满足以下条件：

(1) $\rho(x)$处处连续；

(2) $\rho(0)=0,\rho(x)=\rho(-x)$，且$\rho(x)\geqslant\rho(y)$，当$|x|\geqslant|y|$；

(3) $\lim\limits_{x\to\infty}\rho(x)x^{-2}=0$。

又$\beta_j^*=\beta_j^*(y_1,\cdots,y_n)$，$j=1,\cdots,p$满足条件：

$$D(\beta_1^*,\cdots,\beta_p^*)=\min_{\beta_1,\cdots,\beta_p}D(\beta_1,\cdots,\beta_p) \tag{3}$$

则称$(\beta_1^*,\cdots,\beta_p^*)$为$(\beta_1,\cdots,\beta_p)$的一个$M$估计。

这个概念首先是由Huber在1964年引进，见参考文献[1]。在[1]中，Huber仔细讨论了这种估计的大样本性质。不言而喻，M—估计的性质与所选择的函数ρ有很大的关系。在文献中，并未对M—估计中的函数ρ应具备怎样的最低限度的性质作过要求。然而，任何合理的定义都应当包含我们所给的定义中的三个要求。

如果$\rho'(x)=\psi(x)$处处存在，则为使(3)成立，$(\beta_1^*,\cdots,\beta_p^*)$必须满足以下的方程组：

$$\sum_{i=1}^{n}\psi(\delta_i)x_{ij}=0,\ \ j=1,\cdots,p \tag{4}$$

其中

$$\delta_i=y_i-\sum_{j=1}^{p}x_{ij}\beta_j^*,\ \ i=1,\cdots,n \tag{5}$$

为了证明方程组(4)有解，而且这个解满足(3)，必须对函数ρ作进一步限制。下面的定理有重要意义。

定理　设以下的条件成立：

(1) $\rho(x)$为定义在$(-\infty,+\infty)$上的非负凸函数，$\rho'(x)=\psi(x)$处处存在，且：

$$\lim_{x\to\infty}\rho(x)=\infty; \tag{6}$$

(2) 下列矩阵的秩为P：

$$A=\begin{pmatrix}x_{11} & x_{12} & \cdots & x_{1p}\\ x_{21} & x_{22} & \cdots & x_{2p}\\ \cdots & \cdots & \cdots & \cdots\\ x_{n1} & x_{n2} & \cdots & x_{np}\end{pmatrix}$$

则方程组(4)必有解，其解必满足(3)；且任何满足(3)的$(\beta_1^*,\cdots,\beta_p^*)$，必是方程组(4)

的解；若进一步假设 $\rho(x)$的严格凸函数，则方程组(4)的解唯一（定理证明见附录）。

3. 雷达测速中的Robust估计的数学原理

应用文献[2]所表示的ML型统计法，即使在没有误差的情况下什么也不会丢失。另外，这种方法用统计项表示，使方程式符合经典理论的一般方程：

$$\sum_i \psi\left(\frac{X_i - \sum_k C_{ik}\theta_k}{\sigma}\right)C_{ij}=0,\quad j=1,\cdots,p \tag{7}$$

$$\frac{1}{n-p}\sum_i \psi\left(\frac{X_i - \sum_k C_{ik}\theta_k}{\sigma}\right)^2=\beta \tag{8}$$

$$\beta = E_{\phi}\psi(U)^2$$

其中σ表示标准方差，$\theta=(\theta_1,\cdots,\theta_p)$是一个无限集，常数函数 ψ 的作用在于减小误差的影响，去除过大误差。HAMPEL 根据图1(a)提出了函数 ψ，我们稍作修改后如图1(b)所示。修改是为了简化计算。这是由于：根据反复地特性实验，我们取图 1 中的 $c_1=1.2$，$c_2=2.4$，这样既保证了结果的存在和唯一，而且一般经过两至三次迭代即可收敛。令 $C^T=\begin{bmatrix}c_{i_1}\\ c_{i_2}\end{bmatrix}$，$T_i=\sum_k c_{ik}\theta_k^0 \quad i=1,\cdots,n$，对 X_i-T_i 进行泰勒展开，并取线性部分有

$$\sum_i \sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{ij}+\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)[T_i-\sum_k c_{ik}\theta_k]c_{ij}=0 \tag{9}$$

不妨取$k=1,2,j=1,2,p=2$，展开，并整理得：

$$\begin{cases}\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)c_{i_1}c_{i_1}\theta_1+\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)c_{i_2}c_{i_1}\theta_2\\ =\sum_i\sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{i_1}+\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)T_i c_{i_1}\\ \sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)c_{i_1}c_{i_2}\theta_1+\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)c_{i_2}c_{i_2}\theta_2\\ =\sum_i\sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{i_2}+\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)T_i c_{i_2}\end{cases}$$

由于X_i-T_i偏差较小，我们可以利用蒙特卡罗方法计算 $\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)$。

以$m=\frac{1}{n}\sum_i\psi'\left(\frac{X_i-T_i}{\sigma}\right)$取代$\psi'\left(\frac{X_i-T_i}{\sigma}\right)$，则有

$$\begin{cases}m\sum_i c_{i_1}c_{i_1}\theta_1+m\sum_i c_{i_1}c_{i_2}\theta_2=\sum_i\sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{i_1}+m\sum_i T_i c_{i_1}\\ m\sum_i c_{i_1}c_{i_2}\theta_1+m\sum_i c_{i_2}c_{i_2}\theta_2=\sum_i\sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{i_2}+m\sum_i T_i c_{i_2}\end{cases}$$

根据常数部分，分为两部分求解，并令

$$y_j=\sum_i \sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{ij};\quad y=\begin{bmatrix}y_1\\y_2\end{bmatrix}$$

则有

$$m(c^Tc)\begin{bmatrix}\Delta\theta_1\\\Delta\theta_2\end{bmatrix}=\begin{bmatrix}y_1\\y_2\end{bmatrix}\tag{10}$$

因而

$$\Delta\theta=\frac{1}{m}(c^Tc)^{-1}y;\quad m(c^Tc)\begin{bmatrix}\theta_1\\\theta_2\end{bmatrix}=mc^TT\tag{11}$$

所以

$$\theta=(c^Tc)^{-1}c^TT\tag{12}$$

即

$$\theta_{new}=\theta+\Delta\theta$$

又令：$q=\frac{1}{n-p}\sum\psi\left(\frac{X_i-T_i}{\sigma}\right)^2\sigma^2$，则由式(8)有

$$\sigma_{new}=\left(\frac{q}{\beta}\right)^{\frac{1}{2}}\tag{13}$$

在测速雷达中，因 $V(t)=V_0/(1+c_sV_0t)$

线性化为：$V(t)=V_0(1-c_sV_0t)=V_0-V_0^2tc=c_{i_1}\theta_1+c_{i_2}\theta_2$ (14)

即：$c_{i_1}=1$，$\theta_1=V_0$，$c_{i2}=-V_0^2t_i$，$\theta_2=c_s$，其中c_s是阻力衰减系数。

因此，通过Robust估计即可将初速和阻力衰减系数估出。上述过程总结如下：

(1) 取初值 V_0、c、σ；

(2) 计算$(c^Tc)^{-1}$；

(3) 令：$T_i=\sum_k c_{ik}\theta_k$；　$m=\frac{1}{n}\sum\psi'\left(\frac{X_i-T_i}{\sigma}\right)$

$$y_j=\sum_i\sigma\psi\left(\frac{X_i-T_i}{\sigma}\right)c_{ij};\qquad q=\frac{1}{n-p}\sum\psi\left(\frac{X_i-T_i}{\sigma}\right)^2\sigma^2$$

(4) 计算：$\sigma_{new}=\left(\frac{q}{\beta}\right)^{1/2}$；$\Delta\theta=\frac{1}{m}(c^Tc)^{-1}y$；$\theta_{new}=\theta+\Delta\theta$

(5) 结果$\|\Delta\theta\|$比某个预定门限要小，则结束，否则返回2重新开始。

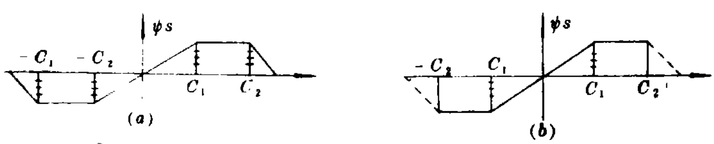

图1 (a) HAMPEL的ψ_s　　(b) 简化的ψ_s

4. 模拟数据仿真和实际数据测试结果

1) 模拟数据仿真

实践表明，对于均匀分布的误差，ML型统计法能满足优化的条件的需要。我们曾经用几组相当差的测量值，即受非均匀误差严重干扰的测试值进行试验。

表 1　Robust估计与最小二乘估计比较（初速750米/秒，衰减因数$c=0.0003$米$^{-1}$）

测量值误差分布($N=30$)	估计方法	V_0	c	σ
$4N(0,15)+5N(0,0.2)+6N(0,2)+7N(0,0.2)+4N(0,2)+4N(0,10)$	P^*	752.0537	/	4.827313
	R^*	750.0848	$3.000817E-04$	0.219928
$5N(0,10)+7N(0,0.2)+3N(0,2)+5N(0,0.2)+6N(0,2)+4N(0.5)$	P	751.8630	/	3.528324
	R	750.0000	$2.999266E-04$	0.2189921
$5N(0,20)+7N(0,0.2)+7N(0,2)+4N(0,0.2)+3N(0,2)+3N(0,15)$	P	754.0724	/	6.983440
	R	749.9338	$2.998595E-04$	0.2205973
$5N(0,8)+7N(0,1)+8N(0,0.2)+4N(0,1)+3N(0,0.2)+3N(0.10)$	P	751.8204	/	3.093524
	R	750.0349	$3.001645E-04$	0.2137124
$5N(0,0.1)+3N(0,8)+7N(0,0.2)+9N(0,0,1)+3N(0,9)+3N(0,0,2)$	P	745.3093	/	2.621149
	R	749.9252	$2.999921E-04$	0.1829986
$6N(0,5)+5N(0,0,2)+4N(0,2)+9N(0,0,2)+3N(0,2)+7N(0,3)$	P	753.5463	/	2.932512
	R	750.047	$3.000013E-04$	0.2134383
$10N(0,0,1)+5N(0,9)+5N(0,0,2)+4N(0,4)+3N(0,2)+3N(0,0,2)$	P	749.4006	/	3.269486
	R	749.9262	$3.000933E-04$	0.1891873
$10N(0,0,1)+14N(0,0)+6N(0,0,1)$	P	750.4172	/	0.1448314
	R	749.932	$2.99968E-04$	0.1267683
$30N(0,0)$	P	749.8784	/	$4.8841E-02$
	R	749.9999	$2.9999E-04$	$2.3565E-03$
$5N(2,2)+22N(0,0)+3N(2,2)$	P	753.6142	/	0.6095116
	R	753.4363	$3.0276E-04$	0.2508543

注：P表示最小二乘多项式回归方法，R表示Robust回归方法。

在初速为750米/秒，衰减因数c为0.0003米$^{-1}$的条件下，我们假设了几组假随机测试结果，其中均方差为0.2米/秒的“好”的测量值若干次，均方差2米/秒的可疑测量

值若干次，还有若干次重大误差。后一种误差被有意识地放在这几组的开始与结束。关于每组测试各种误差测量值所占比例及结果列于表1，并把它与最小二乘多项式回归计算结果相比较。

2） 实际数据测试

我们对某雷达在实际中测试到的数据，利用此方法进行数据处理，其结果列于表2。

3） 结论：

从上面的理论分析和实验结果可以看到，最小二乘估计由于考虑残差的平方和，它

表 2 实际数据的三种估计方法比较

V_0（L 估计）	估计方法	V_0（米/秒）	c（米$^{-1}$）	σ
1471.028/1470.244	P	1475.733	/	3.118711
	R	1469.511	3.181436E-05	0.298558
1473.948/1473.529	P	1471.976	/	0.855578
	R	1472.638	3.114111E-05	0.2858569
1472.977/1472.538	P	1473.026	/	0.8424097
	R	1472.111	3.040602E-05	0.3874447
1469.791/1469.41	P	1474.139	/	4.053387
	R	1469.561	3.136373E-05	0.4526574
1471.046/1469.941	P	1481.362	/	5.831779
	R	1470.016	3.126493E-05	0.4204114
1474.383/1471.597	P	1472.412	/	4.020454
	R	1471.891	3.17312E-05	0.2597771

注：LS—某雷达所用最小二乘估计方法；P—最小二乘多项式回归方法；R—Robust回归方法；"/"以上值为没剔除离群值时的外推值，"/"以下值为已剔除离群值时的外推值。

就使离群值的作用显著增加，而Robust估计则改进了最小二乘估计受离群值影响太大的缺点，并且误差分布不必是正态的。须指出的是：Robust回归估计的目的本不在于缩小残差）实际上，由于最小二乘法估计已使残差平方和达到最小，别的方法在缩小残差上不可能优于它），而在于求得回归系数的较正确的估计。

附录　定理证明

先证以下几个事实：

a)，以β记向量$(\beta_1,\cdots,\beta_p)'$，而$\|\beta\|=(\beta_1^2+\cdots+\beta_p)^{1/2}$，则当$\|\beta\|\to\infty$时，$D(\beta_1,\cdots,\beta_p)\to\infty$，$D$由(2)定义。

事实上，以C记球面$\{\beta:\|\beta\|=1\}$，又令

$$K(\beta)=\|AB\|^2=\sum_{i=1}^{n}\left(\sum_{j=1}^{p}x_{ij}\beta_j\right)^2$$

$K(\beta)$在c上连续，而c是R_p中的有界闭集，故由连续函数的基本性质，存在$\beta_0\in c$，使

$$K(\beta_0)=\min_{\beta\in c}K(\beta)$$

必有$K(\beta_0)>0$。因为若$K(\beta_0)=0$，则将有$A\beta_0=0$，而$\beta_0\neq0$（因$\beta_0\in c$），这显然与矩阵A的秩等于P矛盾。所以当$\|\beta\|\to\infty$时，有：

$$\|A\beta\|=\|\beta\|\cdot\|A(\beta/\|\beta\|)\|\geqslant\|\beta\|\rho\sqrt{K(\beta_0)}\to\infty$$

利用三角形不等式：

$$\sum_{i=1}^{n}\left|y_i-\sum_{j=1}^{p}x_{ij}\beta_j\right|\geqslant\left[\sum_{i=1}^{n}\left(y_i-\sum_{j=1}^{p}x_{ij}\beta_j\right)^2\right]^{1/2}$$
$$\geqslant\|A\beta\|-\left(\sum_{i=1}^{n}y_i^2\right)^{1/2}\to\infty$$

最后由(6)式得到要证的结果。

b)　$D(\beta)=D(\beta_1,\cdots,\beta_p)$为$\beta$的凸函数，又若$\rho$为严凸，则$D$也为严凸。

记$y=(y_1,\cdots,y_n)'$，且设$c_1\geqslant0$，$c_2\geqslant0$，$c_1+c_2=1$，$\beta^{(1)}$和$\beta^{(2)}$为两个P维向量，$\beta=c_1\beta^{(1)}+c_2\beta^{(2)}$，又对任一向量$\xi$，以$(\xi)_i$记其第$i$分量。由$\rho$的凸性，有：

$$D(\beta)=\sum_{i=1}^{n}\rho[(Y-A\beta)_i]=\sum_{i=1}^{n}\rho[(c_1Y_1+c_2Y_2-c_1A\beta^{(1)}-c_2A\beta^{(2)})_i]$$
$$=\sum_{i=1}^{n}\rho[c_1(Y-A\beta^{(1)})_i+c_2(Y-A\beta^{(2)})_i]$$
$$\geqslant\sum_{i=1}^{n}\{c_1\rho[Y-A\beta^{(1)})_i]+c_2\rho[Y-A\beta^{(2)})_i]\}=c_1D(\beta^{(1)})+c_2D(\beta^{(2)})$$

这证明了D的凸性。若ρ为严凸，$\beta^{(1)}\neq\beta^{(2)}$，而$c_1$、$c_2$皆不为0，则因至少有一个$i$使$(Y-A\beta^{(1)})\neq(Y-A\beta^{(2)})_i$，（否则将有$A\beta^{(1)}=A\beta^{(2)}$，即$A(\beta^{(1)}-\beta^{(2)})=0$，由$A$的秩为$P$将得$\beta^{(1)}-\beta^{(2)}=0$，与$\beta^{(1)}\neq\beta^{(2)}$矛盾），知(7)中的"$\geqslant$"号要改为"$>$"号，这证明了$D$为严凸的。

现在转到定理的证明。由$D(\beta)$的连续性（由ρ的连续性得出）及预备事实a，知$D(\beta)$的最小值必在且只能在某些有限点β^*处达到。又由$D(\beta)$对$\beta_1,\cdots,\beta_p$的偏导数处处存在（由$\rho'(x)$处处存在）知这样的β^*必须满足方程组(1.4)，（这部分证明没有用到ρ的凸性）。

现在证明：方程组(4)的任一解$\beta^*(\beta_1^*,\cdots,\beta_p^*)$必满足(3)。设若不然，则存在$\tilde{\beta}=(\tilde{\beta}_1,\cdots,\tilde{\beta}_p)$，使$D(\tilde{\beta})<D(\beta^*)$。考虑函数$h(t)=D(\tilde{\beta}+t(\beta^*-\tilde{\beta}))$，则由$D(\beta)$的凸性

知，$h(t)$为凸函数，而$h(1)>h(0)$，因此将得$h'(1)>0$。但因β^*为方程组(4)之解，有

$$\left.\frac{\partial D(\beta)}{\partial \beta_j}\right|_{\beta=\beta^*}=0,\quad j=1,\cdots,p$$

故
$$h'(1)=\sum_{j=1}^{p}\left.\frac{\partial D(\beta)}{\partial \beta_j}\right|_{\beta=\beta^*}(\beta_j^*-\tilde{\beta}_j)=0$$

与$h'(1)>0$矛盾。这证明了β^*必满足(3)。

最后，若ρ为严凸，则$D(\beta)$为严凸，因而$D(\beta)$只能有唯一的最小值点（存在性前已证明），故方程组(4)也只能有唯一解。

（证毕）

参考文献

[1] P.J.Huber: Robust estimation of a location parameter, Ann.Math. Staist, 35, 1964,73-101

[2] P.J.Huber: Robust Regression: Asymptotics, conjectares, and Monte Carlo, Ann stat 1(5), 1973,pp799—821

[3] 热尔阿尔·洛伯：用现代统计方法精确测定初速，国外靶场技术

[4] 陈希孺、王松桂编著：近代实用回归分析，广西人民出版社，1984年2月

Robust Technique in Speed Measuring with Radar

Wan Jianwei　Huang Pukan

Abstract

Exactly measuring the raw speed is still one of the most important tasks in trajectory measuring. Modern Dopplar radar is of practical use, but there are many weak points in processing practical datum. This paper uses modern statistical method, called "Robust estimation", to calculate the first speed, the precision of which can be estimated, even in the worst condition by improving the data processing in the raw speed measuring radar. With the same method, the coefficient of resistance can also be determined. The experiments have heen proved to be of these correct conclusions.

Key Words: Robust Estimation, Radar, Data Processing, Monte Carlo