极值理论检验

如何准确刻画投资中的极值事件，度量金融业所面临的极值风险，一直是投资者以及金融监管者关注的问题。一系列金融危机事件的发生催生了极值理论(Extreme Value Theory，EVT）在投资风险管理中的应用：1995 年 2 月，世纪著名商业银行巴林银行其新加坡分行一名交易员 Nick Leeson 的交易就成了将近九亿美元的损失；几个月后，日本大和银行发现一交易员 Toshihide Iguchi 造成了七亿美元的亏损。这类事件的连续发生促使美联储主席 Alan Greenspan 建议在金融风险领域运用极值理论（EVT)。

事实上，极值理论的不断深化和发展为度量极值风险提供了一个有效的工具。在利用极值理论进行风险管理时，首先必须对极值事件的统计规律进行分析，得出极值分布中各个参数的估计。因此，本节先简要介绍一下极值理论的基本概念和估计方法，然后基于实际案例介绍极值理论在 VaR 中的应用。

一、极值理论简介

一个随机过程的极值指的是在给定的时间区间上的极大极小值。极值方法注重价格或回报分布的尾部而不是整个分布。

若用 r_t 表示某资产的日收益率，则该资产 n 天中的收益率集合可表示为 $\{r_1,\ldots, r_n\}$。该集合中的最小收益率（也即最小次序统计量）为 $r_{(1)} = \min_{1\le j\le n}\{r_j\}$，最大收益率（也即最大次序统计量）为 $r_{(n)} = \max_{1\le j\le n}\{r_j\}$。因此对于持有该头寸的多头而言，其所持头寸的 VaR 与 $r_{(1)}$ 是密切相关的；对于持有该头寸的空头而言，其所持头寸的 VaR 则与 $r_{(n)}$ 密切相关。实际上两者之间的关系只需通过一个简单的符号变换即可得到（$r_{(1)} = -\max_{1\le j\le n}\{r_j\} = -r_{(n)}$），本文主要讨论 $r_{(n)}$ 的性质，并通过以上转换来计算多头头寸的 VaR。

假设收益率 r_t 是独立的时间序列，且 r_t 的变化范围为 $[l,u]$，累积概率密度分布函数为 F(x)，则 $r_{(n)}$ 的累积概率密度分布函数 $F_{n,n}(x)$ 可由式(1)给出：

$$F_{n,n}(x) = \Pr[r_{(n)} \le x] = \Pr[r_1 \le x,\ldots, r_n \le x] = \prod_{j=1}^{n} \Pr(r_j \le x) = [F(x)]^n \qquad (1)$$

极值理论关心的是找到两个序列 $\{\beta_n\}$ 和 $\{\alpha_n\}$，使得 $r_{(n*)} \equiv (r_{(n)} - \beta_n)/\alpha_n$ 当 n 趋于无穷

时收敛于一个非退化分布。其中序列$\{\beta_n\}$被称为位置序列，$\{\alpha_n\}$被称为尺度因子序列。在以上假定下，标准化的最大收益率$r_{(n*)}$的极限分布为[1]：

$$F_*(x)=\begin{cases}\exp[-(1+\xi x)^{-\frac{1}{\xi}}] & if \quad \xi \neq 0 \\ \exp[-\exp(-x)] & if \quad \xi = 0\end{cases} \tag{2}$$

如果$\xi<0$，则上式对$x<-\dfrac{1}{\xi}$成立，如果$\xi>0$，则上式对$x>-\dfrac{1}{\xi}$成立。ξ被称为形状参数，它控制着极限分布的尾部形状，而$\alpha=\dfrac{1}{\xi}$则称为分布的尾指数。

上式是极限分布的一般形式，它包含了以下三种情形：

1. $\xi=0$，Gumbel 族，其累积概率密度分布函数为：

$$F_*(x)=\exp\left[-\exp(-x)\right] \quad ,-\infty<x<\infty \tag{3}$$

2. $\xi>0$，族，Fréchet 其累积概率密度分布函数为：

$$F_*(x)=\begin{cases}\exp[-(1+\xi x)-^{\frac{1}{\xi}}] & if \quad x>-\dfrac{1}{\xi} \\ 0 & 其他情况\end{cases} \tag{4}$$

3. $\xi<0$，Weibull 族，其累积概率密度分布函数为：

$$F_*(x)=\begin{cases}\exp[-(1+\xi x)^{-\frac{1}{\xi}}] & if \quad x<-\dfrac{1}{\xi} \\ 1 & 其他情况\end{cases} \tag{5}$$

在风险管理中，我们主要对 Fréchet 族感兴趣，它包含了平稳分布和学生 t-分布。极值理论有两个重要的特征：

1. r_t的累积概率密度分布函数$F(x)$的尾部行为而不是具体的分布决定了最大收益率的极限分布。基于此，对于收益率r_t的一个较为广泛的分布形态，极值理论都是实际可行的。
2. 尾指数ξ并不依赖与收益率r_t的时间间隔，也即尾指数在时间累积下是不变的，从

[1] 参见 Tsay(2010)的著作“Analysis of Financial Time Series”，3rd.

而可以用于计算不同时间区间上的 VaR。

以下简要讨论一下极值分布的估计方法。由前文可知，极值分布包含 3 个参数：ξ（形状参数）、β_n（位置参数）和α_n（尺度参数）。对于给定的样本，只有一个单一的最大收益率或最小收益率，而仅用一个极端观测并不能估计出三个参数，因此必须采取其他方法。在具体估计方法的选择上，常用的有两类：

第一类方法称为参数方法。该方法一个基本的想法是将样本分为 g 个互不交互的子样本，且每个子样本有n个观测值，则每个子样本对应于数据区间的一个子区间。当n充分大时，我们希望极值理论对每个子样本都适用，而n的值则由实际情况来决定。这样我们可以得到n个最大值的集合和n个最小值的集合，从而用来估计极值分布的未知参数。通常有两种参数方法：极大似然方法和回归方法。

第二类方法称为非参数方法，直接对变量r_t进行尾估计，并不需要考虑子样，如 Hill 估计。该方法将样本的次序统计量表示为：

$$r_{(1)} \le r_{(2)} \le \ldots \le r_{(T)}$$

令 q 为一个正整数，则 Hill 估计的形状参数ξ定义为：

$$\xi(q) = \frac{1}{q}\sum_{i=1}^{q}[\ln(r_{(T-i+1)}) - \ln(r_{(T-q)})] \qquad (6)$$

其中(q)说明该估计ξ依赖于q。研究表明，Hill 估计仅仅对 Fréchet 族是适用的，此时可以找到一个合适的q^*，使得$q > q^*$当时 Hill 估计$\xi(q)$是稳定的。由于本文采用 R 中的软件包 evir[2]来进行极值理论的参数估计，因此具体的计算过程不再详述。

二、极值理论在 VaR 中的应用

本节中我们使用万科 A（000002）股票从 1991 年 1 月 30 日-2011 年 12 月 26 日的日对数收益率数据进行极值理论分析。图 10-16 给出了子区间长度为 21 个交易日时极端日对数收益率的时间图形。从图可知，1997 年之前万科的最大涨跌幅往往超过 10%，这是因为当时中国股市还没有实施涨跌停制度[3]。因此，在目前的制度环境下，本节以下部分只对 1997 年 1 月 1 日-2011 年 12 月 26 日的日收益率数据进行分析。

[2] R 中用于进行 Hill 估计的命令为 hill，用于进行极大似然估计的命令为 gev。

[3]上海、深圳证券交易所自 1996 年 12 月 16 日起，分别对上市交易的股票（含 A、B 股）、基金类证券的交易实行价格涨跌幅限制。

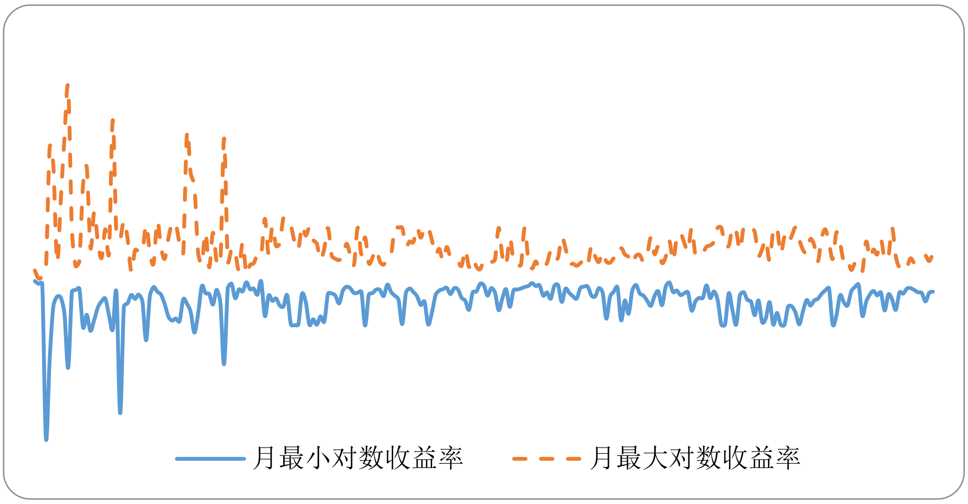

图 1　子区间为 21 个交易日时，万科 A 极端日对数收益率变动

图 2 给出了 Hill 估计 $\xi(q)$ 对 q 的散点图[4]。由图可知，对于最大和最小日对数收益率，$\xi(q)$ 随着 q 的增加而不断上升，这说明 Hill 估计方法对万科 A 的日收益率进行极值分布估计可能是不太合适的。

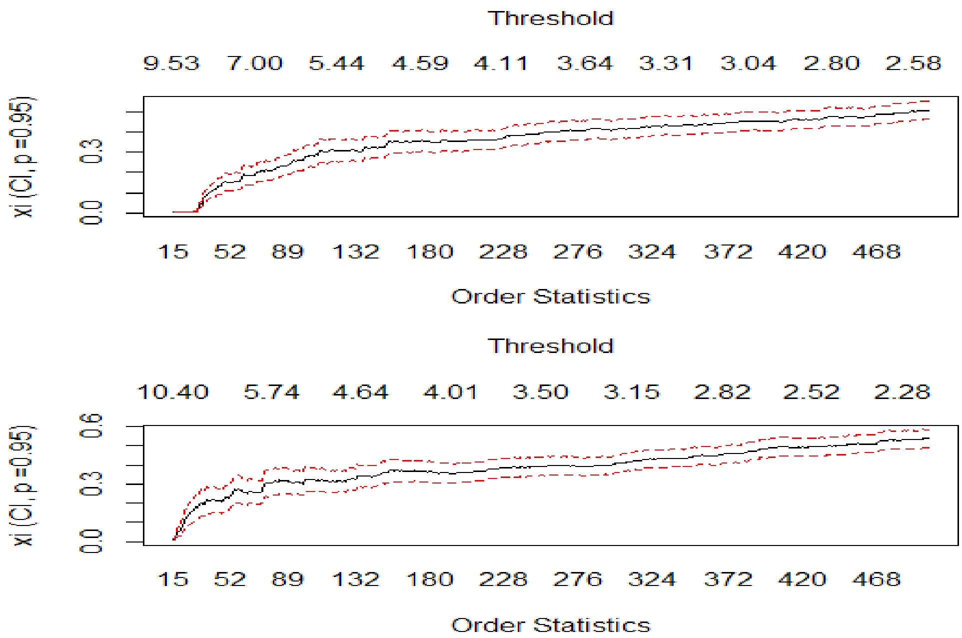

图 2　万科 A 日对数收益率 Hill 估计散点图

（注：上图为最大收益率，下图为最小收益率）

表 1 则给出了对不同长度子区间基于极大似然法估计出的极值分布参数结果[5]。

表 1　万科 A 日对数收益率极值分布的极大似然估计

[4] 估计结果由 R 软件包 evir 中的 hill 命令计算得到。
[5] 估计结果由 R 软件包 evir 中的 gev 命令计算得到。

子区间长度	形状参数 ξ_n	尺度参数 α_n	位置参数 β_n
最小收益率			
1 个月(n=21, g=173)	0.191(0.074)	1.686(0.120)	3.447(0.150)
2 个月(n=42, g=87)	0.135(0.122)	1.999(0.207)	4.343(0.260)
最大收益率			
1 个月(n=21, g=173)	-0.447(0.098)	2.125(0.158)	4.131(0.203)
2 个月(n=42, g=87)	-0.795(0.094)	3.106(0.354)	5.999(0.355)

注：括号内的数据位标准误差。

为对以上估计结果进行检验，图 3 给出了子区间长度为 21 个交易日时最小日收益率极值分布拟合的残差结果。由下面的 Q-Q 图可以看出，该拟合是较为合理的。

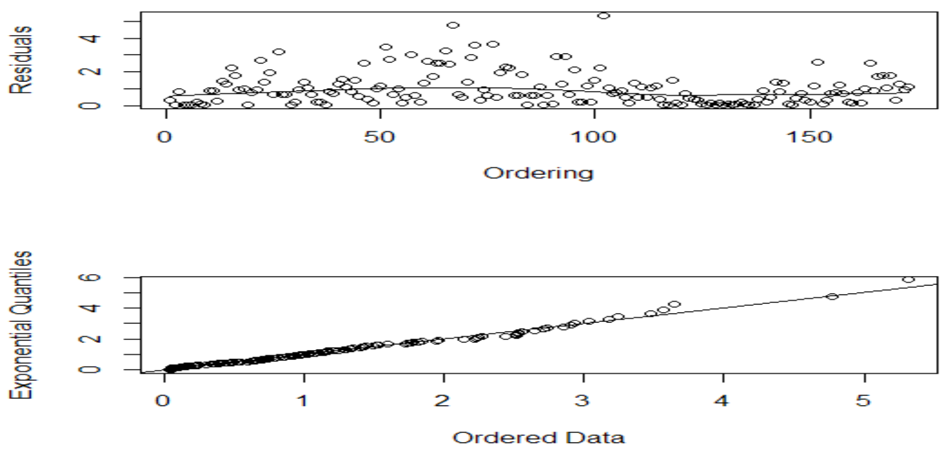

图 3　万科 A 最小日对数收益率极值分布拟合的残差图

令 $r_{(n^*)} \equiv (r_{(n)} - \beta_n)/\alpha_n$，将表 1 中估计出的参数值代入式(2)，则可以得到极值分布在给定概率下的分位数。

具体计算过程如下：

若令 P^* 为该分布的上尾分布概率，则 $r_{(n^*)}$（意味着潜在损失，即前文定义的 VaR）对应着极限分布下某子区间极值的 1-P^* 分位数。此时有

$$1 - P^* = \begin{cases} \exp\{-[1 + \dfrac{\xi_n (r_{(n^*)} - \beta_n)}{\alpha_n}]^{-\frac{1}{\xi_n}}\} & if \quad \xi_n \neq 0 \\ \exp[-\exp(-\dfrac{r_{(n^*)} - \beta_n}{\alpha_n})] & if \quad \xi_n = 0 \end{cases} \qquad (8)$$

整理后可得相应的分位数为：

$$r_{(n^*)}=\begin{cases}\beta_n-\dfrac{\alpha_n}{\xi_n}\{1-[-\ln(1-P^*)]^{-\xi_n}\} & if \quad \xi_n\neq 0\\ \beta_n-\alpha_n\ln[-\ln(1-P^*)] & if \quad \xi_n=0\end{cases} \tag{9}$$

由于$1-P^*=P(r_{n,i}\leq r_{(n^*)})=[P(r_t\leq r_{(n*)})]^n$，则式(10-33)变为:

$$VaR=\begin{cases}\beta_n-\dfrac{\alpha_n}{\xi_n}\{1-[-n\ln(1-P)]^{-\xi_n}\} & if \quad \xi_n\neq 0\\ \beta_n-\alpha_n\ln[-n\ln(1-P)] & if \quad \xi_n=0\end{cases} \tag{10}$$

其中n是子区间长度。以上即是基于极值理论的 VaR 计算公式，其计算过程可概括如下：

1. 选择子区间的长度n，并得到子区间的最小值$\{r_{n,i}\}$，i=1,⋯，g，这里$g=[T/n]$；
2. 得到ξ_n、α_n和β_n的极大似然估计并检查极值模型的充分性和有效性；
3. 应用式(10-34)计算 VaR。

以万科 A 股票为例，对其 1997 年 1 月 1 日至 2011 年 12 月 26 日的日对数收益率数进行极值分布的参数估计。

取n=21，从表-10 可知ξ_n=0.191、α_n=1.686、β_n=3.447。对应的置信度为 0.01 的 VaR 值为：

$$VaR=3.447-\frac{1.686}{0.191}\{1-[-21*\ln(1-0.01)]^{-0.191}\}=6.50$$

也就是说，万科 A 股票在 1%置信水平下日对数收益率不会低于-6.50%。

若取n=42，则有ξ_n=0.135、α_n=1.999、β_n=4.343。对应的置信度为 1%的 VaR 值为：

$$VaR=4.343-\frac{1.999}{0.135}\{1-[-42*\ln(1-0.01)]^{-0.135}\}=6.17$$

此时，万科 A 股票在 1%置信水平下日对数收益率不会低于-6.17%。

注意到以上案例中，不同的参数n对应着不同的 VaR 值，而n=21 对应着更高的 VaR 值，这是由估计极值分布尾行为时所存在的不确定性决定的。由于没有真实的 VaR 来衡量估计精度，因此建议在运用极值分布进行 VaR 估计时综合考虑基于不同n得到的结果来确定 VaR 的估计范围。

对于更长时间的持有期，如何通过极值理论来计算 VaR 呢？式(11)给出了相应的转换公式：

$$VaR(k)=k^{\xi_n}VaR(1) \tag{11}$$

其中VaR(1)是由式(11)计算出的持有期为一天的VaR值,VaR(k)为持有期为k天的VaR值。若取n=21，则持有期为20个交易日的万科A股票1%置信水平下的VaR值为：

$$VaR(20)=20^{0.191}*6.5=11.52$$