2 風險衡量與相關文獻

2.1 風險值

風險值表示在給定的信心水準 α 下, 在特定時間區間 t 之內, 投資因價格波動, 可能產生的最大損失。計算風險值時可分爲以指數衡量的絕對風險值和以報酬率衡量的相對風險值, 絕對風險值可表示爲

$$P(\triangle W_t < VaR) = \alpha$$

相對風險值可表爲

$$P(r_t < VaR) = \alpha$$

風險值在統計上即爲分量的概念, 優點是直覺上容易理解且計算容易, 故是實務上衡量風險最爲普遍的方法。文獻上對於估計風險值並沒有一個最佳的方法, 以下先簡單介紹不同衡量風險值的方法。

2.2 歷史模擬法 (Historical Simulation)

歷史模擬法爲無母數方法, 主要的想法是假設報酬過去的歷史分配未來仍會再次重演。因此風險值可透過實證資料的分量估計

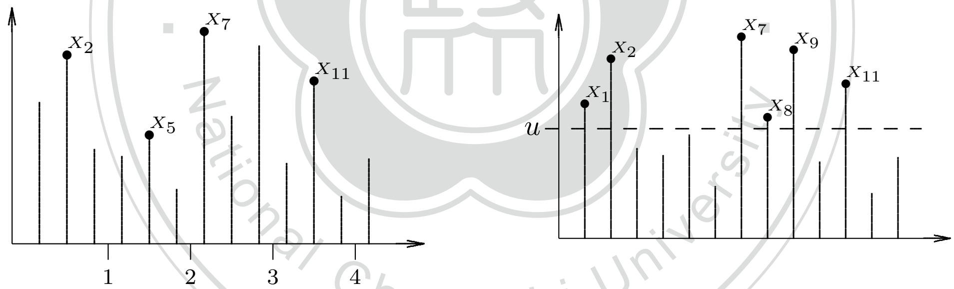
$$VaR_\alpha = Quantile\left\{\{x_i\}_{i=1}^{t}, \alpha\right\} \tag{1}$$

歷史模擬法的優點在於完全不需要分配假設, 也不需要估計任何參數, 操作上相當容易。但是也正因爲如此, 歷史模擬法在樣本外預測時忽略波動的訊息, 模型對環境的適應性較差。

另外, 歷史模擬法受移動窗口的長度影響。由於歷史模擬法給過去各期資料固定的權重, 若選取移動窗口的長度過長, 長久以前的資料仍會影響風險值的估計, 當環境從低波動進入高波動時, 歷史模擬法通常需要一段時間調整; 若移動窗口長度過短, 樣本資料可能不具代表性, 無法準確的估計出風險值。關於歷史模擬法更詳細的說明可參閱 Hendricks (1996) 。

2.3 極值理論 (Extreme Value Theory)

風險值主要探討在尾部極端情況下可能造成的損失, 而極值理論即針對描繪尾部的分配來估計風險值。極值理論可分爲與兩種不同的估計方法, 第一種稱爲 Block Maximum Model (BMM) , 將樣本資料區分爲數個區間, 選取每個區間中最大值爲極端值; 另一種稱爲 Peaks over Threshold method (POT) , 以大於設定門檻值 u 的樣本爲極端值。兩種不同方法選取出的極端值也有所不同, 圖 1 爲以 BMM 法與 POT 法挑選極端值的示意圖例。[2] 以 BMM 法挑選出的極端值爲 X_2 X_5 X_7 X_{11} , 若改以 POT 法, 大於門檻值 u 的極端值爲 X_1 X_2 X_7 X_8 X_9 X_{11} 。比較兩種配適極值分配的方法, 若以 BMM 法挑選極端值, 同一個區間內發生多次極端值時, 會剔除掉部份有用的極端樣本。因此 BMM 法效率通常會低於以門檻值決定極端值的 POT 法。Gilli and Këllezi (2006) 曾以六國股市實證分析, 實證結果顯示以 POT 法估計優於 BMM 法。



(a) Block Maximum Model (b) Peaks over Threshold

圖 1: EVT 挑選極端值示意圖

以下說明如何透過 BMM 法與 POT 法估計風險值。

[2]圖 1 摘錄自 Gilli and Këllezi (2006)

Block Maximum Model (BMM)

以 BMM 法估計極值分配, 假設有 T 筆樣本觀察值, 將樣本資料區分成 τ 個區間, 每個區間之中包含 N 筆樣本觀察值, 其中 $\tau = T/N$, 原始的樣本經過排序後可表示爲

$$\left\{x_1, ..., x_N \middle| x_{N+1}, ..., x_{2N} \middle| ...... \middle| x_{(\tau-1)N+1}, ..., x_{\tau N}\right\}$$

再挑選出每個區間中的最大值

$$m_i = \max\left(x_{(i-1)N+1}, ..., x_{iN}\right) \ \forall i = 1, ..., \tau$$

根據 Fisher-Tippett Theorem , 若存在實數 $a > 0$, $b \in R$, 則 $y = \dfrac{M-b}{a}$ 會趨近於一非退化分配 H , H 服從下列三種形式

$$Gumbel : \Lambda(y) = \exp(-e^{-y}), \ -\infty < y < \infty$$

$$Fr\grave{e}chet : \Phi_\alpha(y) = \begin{cases} 0 & \text{if } y \leq 0 \\ \exp(-y^{-\alpha}) & \text{if } y > 0 \end{cases}$$

$$Weibull : \Psi_\alpha(y) = \begin{cases} \exp-(-y)^\alpha & \text{if } y \leq 0 \\ 1 & \text{if } y > 0 \end{cases}$$

爲了估計上方便, 令 $\xi = 1/\alpha$, 上述三種分配可以結合成爲單一參數的分配

$$H_\xi(x) = \exp\left\{-\left(1+\xi\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}\right\} \tag{2}$$

$H_\xi(x)$ 稱之爲 generalized extreme value (GEV) distribution 。

以最大概似法估計 GEV 分配參數 $\hat{\xi}$ $\hat{\mu}$ $\hat{\sigma}$, 對數概似函數爲

$$l(\xi,\mu,\sigma) = -\sum_{i=1}^{T}\left\{\ln\sigma + \left(1+\frac{1}{\xi}\right)\ln\left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right] + \left[1+\xi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{\frac{1}{\xi}}\right\}$$

在給定機率水準 α 下, 風險值爲利用第 (2) 式的反函數求取第 100α 分量

$$VaR_\alpha = \hat{u} + \frac{\hat{\sigma}}{\hat{\xi}}\left(-1 + (-n\ln\alpha)^{-\hat{\xi}}\right) \tag{3}$$

以 BMM 法估計極值分配時, 最大的問題在於應該如何決定區間的長度。當區間長度愈長時, 極端值的樣本數會愈少, 參數估計時可能會發生樣本數不足; 當區間長度愈短時, 每個區間內的樣本的數目愈少, 選出的極端值可能會不具代表性。

Peak over Threshold (POT)

使用 EVT 的 POT 法, 首先挑選出資料中超過門檻值 u 的樣本, 配適一般化柏拉圖分配 (generalized Pareto distribution , GPD)。樣本觀察值大於一特定門檻值 u 的分配, 則稱之爲 excess distribution, 可表示爲

$$F_u(y) = Pr\{X \le u + y \mid X > u\} = \frac{F(u+y) - F(u)}{1 - F(u)},\ y > 0 \tag{4}$$

當門檻值 u 夠大時, $F_u(y)$ 會趨近於 GPD , 可表示爲

$$F_u(y) \approx G_{\xi,\sigma}(y) \quad , u \longrightarrow \infty \tag{5}$$

其中

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - \left(1 + \dfrac{\xi y}{\sigma}\right)^{-\frac{1}{\xi}} & \text{if } \xi \neq 0 \\ 1 - e^{-y/\sigma} & \text{if } \xi = 0 \end{cases}$$

其中 $\sigma > 0$, $-\infty < \xi < \infty$

$x = y + u$, 再透過第 (4) 式與第 (5) 式, 可重新改寫 F , 當 $x > u$ 時

$$F(x) = (1 - F(u))G_{\xi,\sigma}(x - u) + F(u) \tag{6}$$

以經驗分配 (empirical distribution) 估計 $F(u)$

$$\hat{F}(u) = \frac{T - N_u}{T} \tag{7}$$

其中 T 為總樣本數, N_u 為樣本中大於門檻值 u 的個數。

透過最大概似法估計 GDP 參數 $\hat{\xi}$ $\hat{\sigma}$, 對數概似函數為

$$l(\xi, \sigma) = -n \ln \sigma - \left(1 + \frac{1}{\xi}\right) \sum_{i=1}^{T} \ln\left(1 + \frac{x_i \xi}{\sigma}\right)$$

因此樣本尾端估計試可表示為

$$\hat{F}(x) = 1 - \frac{N_u}{T}\left(1 + \hat{\xi}\frac{x-u}{\hat{\sigma}}\right)^{-\frac{1}{\hat{\xi}}} \tag{8}$$

在給定機率水準 α 下, 風險值為利用第 (8) 式的反函數求取第 100α 分量

$$VaR_\alpha = \hat{x}_p = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\frac{N}{N_u}(1-\alpha)^{-\frac{1}{\hat{\xi}}} - 1\right) \tag{9}$$

門檻值的選擇是使用 POT 法估計 GPD 成功的關鍵, 當選擇的門檻值愈大, 可用來估計 GPD 的樣本值會愈少, 估計出的 $\hat{\xi}$ 與 $\hat{\sigma}$ 變異數會過大; 如果選擇的門檻值過小, 估計出的 $\hat{\xi}$ 與 $\hat{\sigma}$ 可能會有偏誤。

門檻值的選擇在文獻上並沒有一致性的準則, McNeil and Frey (2000) 與 Marimoutou , Raggad and Trabelsi (2006) 以 Student t 分配自由度為 4 , 透過模擬 1000 筆樣本資料, 顯示以第 100 到第 150 大的樣本為門檻值, GPD 參數的 MSE 為最小。因此建議可以採用第十分位數做為門檻值。

另外亦可透過 mean excess plot 來判斷。首先定義平均餘額函數 (mean excess function) 為樣本觀察值超過門檻值的條件期望值, 可表示為

$$e(u) = E[x - u | x > u]$$

若超過門檻值 u 的樣本服從 GDP , 當 $\xi < 1$ 時, 平均餘額函數將會是一線性函數

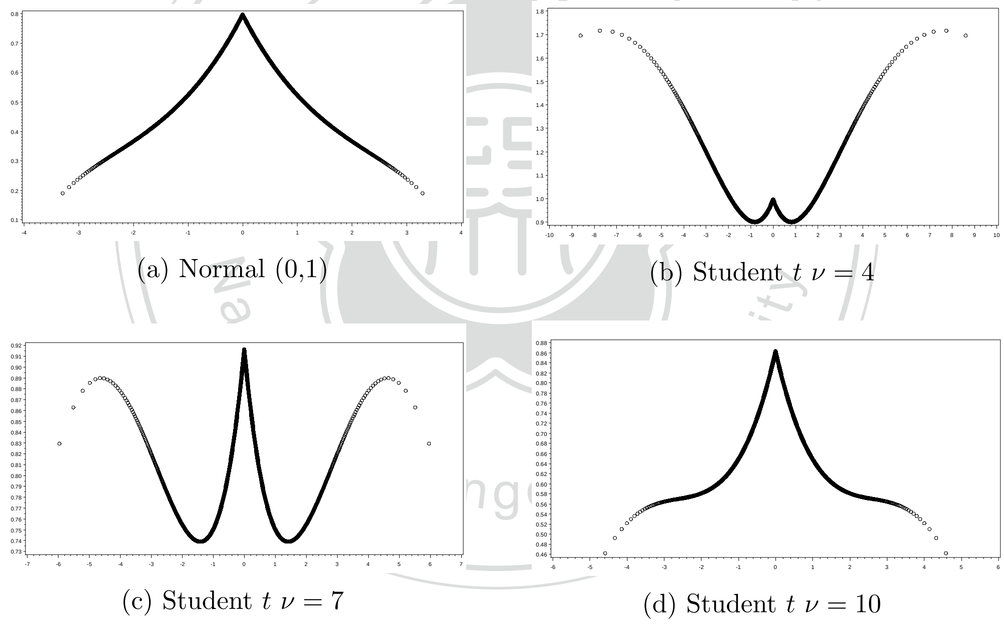
$$e(u) = E[x - u | x > u] = \frac{\sigma + \xi u}{1 - \xi} \tag{10}$$

因此使用 mean excess plot 來判斷門檻值, 首先將樣本觀察值 $\{x_i\}_{i=1}^{t}$ 排序, 得到一組新的樣本 $\{x_{(i)}\}_{i=1}^{t}$, 其中

$$x_{(1)} < x_{(2)} < \ldots\ldots < x_{(t)}$$

計算出在給定不同門檻值下的 $e\left(x_{(i)}\right)$, 以 $x_{(i)}$ 爲橫軸、$e\left(x_{(i)}\right)$ 爲縱軸繪製 mean excess plot 。若圖形在某一門檻值 $x_{(i)}$ 開始近似於直線, 則表示當門檻值 u 爲 $x_{(i)}$ 時, 樣本觀察值大於門檻值 u 的分配會漸進服從 GPD 。

圖 2 爲標準常態分配與 student t 分配在自由度爲 4 、7 、10 時的 mean excess plot 。[3]



(a) Normal (0,1)

(b) Student t $\nu = 4$

(c) Student t $\nu = 7$

(d) Student t $\nu = 10$

圖 2: mean excess plot

[3]當 $x < 0$ 時, 採用 $e(u) = E[u - x | x < u]$

2.4 GARCH Model

財務時間序列資料通常存在波動叢集的現象, 即低波動伴著隨低波動、高波動伴隨著高波動。因此一般認爲, 時間序列資料的變異數會隨時間而改變。Engle (1982) 提出了自我迴歸異質條件變異數模型 (ARCH) , 改變傳統上對變異數爲固定的假設, 允許資料的條件變異數會隨著時間而改。Bollerslev (1986) 進一步將 ARCH 模型擴充爲一般自我迴歸條件異質變異數模型 (GARCH), 使估計的參數變得精簡而模型也更具有彈性。GARCH 模型假設條件變異數爲過去報酬與變異數的函數, 一般化的 GARCH(m,s) 模型可表示爲

$$\begin{aligned} r_t &= \mu + \varepsilon_t \\ h_t &= \alpha_0 + \sum_{i=1}^{m} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i h_{t-i} \\ \varepsilon_t &\sim WN(0, h_t) \end{aligned}$$

其中 WN 爲白噪音, 要求期望值爲 0 , 變異數爲 1 , 且不存在序列相關。因此以 GARCH 模型估計風險值, 風險值可表示爲

$$VaR_\alpha = \hat{\mu} + \hat{h}_t WN^{-1}(\alpha) \tag{11}$$

若白噪音假設爲常態分配, 風險值爲

$$VaR_\alpha = \hat{\mu} + \hat{h}_t \Phi^{-1}(\alpha)$$

與 EVT 不同的, GARCH 模型著重的是模型整體的配適, 而非風險值著重的分配尾端。因此使用 GARCH 模型, 若分配尾端的行爲與常態分配不同, 估計風險值可能會產生問題。

2.5 動態歷史模擬法 (Filtered Historical Simulation)

動態歷史模擬法結合 GARCH 模型與歷史模擬法, 以 GARCH(1,1) 爲例

$$
\begin{aligned}
&r_t = \mu + \varepsilon_t \\
&h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_i h_{t-1} \\
&\varepsilon_t \sim WN(0, h_t)
\end{aligned}
$$

首先以 GARCH 模型估計 $\hat{\mu}$ 與 $\hat{\sigma}_t$ 後, 求出標準化殘差 $z_t = \dfrac{\hat{\mu}}{\hat{\sigma}_t}$, 以第 (1) 式估計 z_t 的分量, 風險值爲

$$VaR_\alpha = \hat{\mu} + \hat{h}_t \, Quantile \left\{ \{z_i\}_{i=1}^{t}, \alpha \right\} \tag{12}$$

2.6 動態極值理論 (Conditional Extreme Value Theory)

動態極值理論結合 GARCH 模型與極值理論, 以 GARCH(1,1) 爲例

$$
\begin{aligned}
&r_t = \mu + \varepsilon_t \\
&h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_i h_{t-1} \\
&\varepsilon_t \sim WN(0, h_t)
\end{aligned}
$$

首先以 GARCH 模型估計 $\hat{\mu}$ 與 $\hat{\sigma}_t$ 後, 求出標準化殘差 $z_t = \dfrac{\hat{\mu}}{\hat{\sigma}_t}$, 再以第 (3) 式或第 (9) 式估計標準化殘差 z_t 極值分配的風險值, 整體風險值爲

$$VaR_\alpha = \hat{\mu} + \hat{h}_t \, VaR_\alpha(z) \tag{13}$$

Marimoutou , Raggad and Trabelsi (2006) 以石油價格、Ghorbel and Trabsi (2007) 以法國 CAC 40 股票指數與突尼西亞 BVMT 証券交易指數、Angelidis et al.(2006) 以道瓊歐盟指數實証分析結果, 皆顯示以動態歷史模擬法與動態極值理論估計的風險值較傳統方法爲佳。