4 模型估計

我們將資料到$t-1$ 期前的訊息集合表示成爲 F_{t-1}, ψ 表示爲未知的參數, 則給定資料 ψ 的概似函數可表示爲:

$$f(\mathbf{y}|\psi) = \prod_{t=1}^{n} \int f(\mathbf{y}_t|\mathbf{h}_t, \mathbf{B}) f(\mathbf{h}_t, \psi) d\mathbf{h}_t$$

$$= \prod_{t=1}^{n} \int N_p(\mathbf{y}_t|\, 0, \mathbf{\Omega}_t) f(\mathbf{h}_t, \psi) d\mathbf{h}_t$$

上式的積分沒有封閉解, 且維度高難以計算。以我們的實證資料爲例, 三個國家的電子指數及一個因子, 共計有 18 個參數要估計。正因爲沒有封閉解, 所以對隨機波動模型參數的估計, 沒有標準作法。文獻上的估計法有, 準最大概似法, 數值積分法, 以及 MCMC 法。我們選擇以 MCMC 的方法, 對後驗分配進行抽樣, 並以得到的統計量來描述後驗分配的性質。我們在下節簡單的介紹馬可夫鏈蒙地卡羅估計法。

4.1 馬可夫鏈蒙地卡羅估計法

當我們要求取一個函數的積分值時, 最簡單的想法即是利用蒙地卡羅積分,

$$\int_a^b f(x)p(x)dx \cong \frac{1}{n}\sum_{i=1}^{n} f(x_i)$$

其中 $x_1, x_2, \cdots, x_n$ 爲來自 $p(x)$ 的隨機變數。則當 n 趨近於無窮大時, 上式就會趨近於我們想要積分的函數積分。但蒙地卡羅積分法效率不高, 文獻上有許多改良的方法。且當 $p(x)$ 也很複雜, 我們不容易抽取它的隨機變數時, 則我們可以尋求重要抽樣 (important sampling), 即並不一定要從 p 直接抽樣來計算。而是找一個我們已知如何抽樣且與類似的分配, 則

$$\int_b^a f(x)q(x)dx = \int_b^a f(x)(\frac{q(x)}{p(x)})p(x)dx \cong \frac{1}{n}\sum_{i=1}^{n} \frac{f(x_i)q(x_i)}{p(x_i)}$$

但當目標函數爲高維度時蒙地卡羅積分的方法會更複雜收斂變慢, 也變的更難操作, 這時我們就尋求 MCMC 法。M-H 法基本上也不直接從 p 下手, 而是想造一個馬可夫鏈他的轉換機率 (transition

kernel) 的極限分配是我們的目標分配 p , 對轉換機率適當假設之下, 這個馬可夫鏈的極限分配, 會趨近我們要的目標分配, 而從這個馬可夫片鏈所產生的隨機變數, 也會是目標函數所產生的隨機變數。下面簡單介紹馬可夫鏈, 與其收斂所需要的性質。

4.1.1 馬可夫鏈

令 $\{X_t\}$ 爲一隨機過程, 其狀態空間爲 Θ , 爲我們說此隨機過程具有馬可夫性質, 如果其轉換機率的條件分配只跟上一期有關, 也就是滿足

$$P(X_{t+1} = \phi | X_t = \theta, X_{t-1}, \cdots) = P(X_{t+1} = \phi | X_t = \theta), \quad \phi, \theta \in \Theta。$$

其意思爲, 隨機變數 X 上期停留在狀態 θ , 下一期轉換到狀態 ϕ 的機率。以下爲了簡潔起見, 直接表示爲 $p(\phi|\theta)$, 當一個馬可夫鏈具有極限分配如果:

$$\pi(\theta) = \sum_{\theta} \pi(\theta) p(\phi|\theta),$$

如果我們能夠找到一個適當的轉換機率 $p(\phi|\theta)$, 它的極限分配就是我們想要的目標分配 $\pi(\theta)$, 則給定任一起始點, 則在疊代一段時間之後, 停留在狀態 θ 的機率就是 $\pi(\theta)$ 。或著說疊待夠久時, 所得的數值, 相當於從目標分配抽的的隨機變數。底下說明要在什麼條件之下, 極限分配會存在, 令 $r(\theta)$ 表示下一期停留在同一狀態的機率,

$$p(\phi|\theta) = p^*(\phi|\theta) + r(\theta)\delta(\phi|\theta),$$

其中 $p^*(\theta|\theta) = 0$, 且

$$\delta(\theta|\theta) = 1, \quad \text{又當 } \theta \neq \phi \quad , \delta(\phi|\theta) = 0,$$

因爲下一期狀態要不就與這期一樣, 要不就與這期不一樣, 機率加總爲一

$$1 = \sum_{\phi} p(\phi|\theta) = \sum_{\phi} p^*(\phi|\theta) + r(\theta)。$$

又我們說一個轉換機率具可反轉性 (time reversibility), 如果滿足,

$$\pi(\theta) = p^*(\phi|\theta) = \pi(\phi)p^*(\theta|\phi),$$

上式意思是給定 θ 的初始機率 $\pi(\theta)$, 從狀態 θ 轉到狀態 ϕ 的機率, 與一該使停留在 ϕ 最後轉到 θ 的機率是相等的。

$$\begin{aligned}
\sum_\theta \pi(\theta)p(\phi|\theta) &= \sum_\theta \pi(\theta)p^*(\phi|\theta) + \sum_\theta \pi(\theta)r(\theta)\delta(\phi|\theta) \\
&= \sum_\theta \pi(\phi)p^*(\theta|\phi) + \pi(\phi)r(\phi) \\
&= \pi(\phi)\left\{\sum_\theta p^*(\theta|\phi)\right\} + \pi(\phi)r(\phi) \\
&= \pi(\phi)\{1 - r(\phi)\} + \pi(\phi)r(\phi) \\
&= \pi(\phi)
\end{aligned}$$

4.2 Metropolis-Hastings 抽樣

從上面我們知道一個轉換機率是不是具有可反轉性質, 關係到此馬可夫鏈會不會收斂到我們想要的極限分配。令 $q(\phi|\theta)$ 表一候選分配, 是一個已知且容易抽樣的分配, 且滿足, $\sum_\phi q(\phi|\theta) = 1$ 。一般來說, $q(\phi|\theta)$ 並不一定具有可反轉性, 不失一般性我們假設,

$$\pi(\theta)q(\phi|\theta) > \pi(\phi)q(\theta|\phi),$$

這表示從狀態 θ 轉到狀態 ϕ 的機會比從 ϕ 轉到 θ 的機會要大, 爲了要平衡兩邊狀態轉換的機率, 我們對左式乘上一個一個調整項, $\alpha(\phi|\theta)$, 使得

$$\pi(\theta)q(\phi|\theta)\alpha(\phi|\theta) = \pi(\phi)q(\theta|\phi),$$

則

$$\alpha(\phi|\theta) = \frac{\pi(\phi)q(\theta|\theta)}{\pi(\theta)q(\mathbf{p}|\theta)}。$$

$\alpha(\phi|\theta)$ 的作用爲限制從 ϕ 轉到 θ 的次數, 藉以達到可反轉性, 所以 α 又稱爲移動機率 (probability of moving)。因爲 ϕ 轉到 θ 的次數過少, 所以我們不減低其次數, $\alpha(\theta|\phi) = 1$ 。而當前面的不等式符號相反時, 算法一樣, 只不過是將 ϕ 與 θ 相互對調而已。我們可以得到通式,

$$\alpha(\phi|\theta) = \min\left[\frac{\pi(\phi)q(\theta|\theta)}{\pi(\theta)q(\phi|\theta)}, 1\right]。$$

這裡我們也可以看出, 因爲 π 出現在分子與分母, 所以當我們計算 α 時, 其實只要知道 π 的同比例分配即可。又當我們的候選分配是對稱分配, 即 $q(\phi|\,\theta) = q(\theta|\phi)$ 時, 接受率變成爲

$$\alpha(\phi|\theta) = \min\left[\frac{\pi(\phi)}{\pi(\theta)}, 1\right]。$$

而這就是 Metropolis 當初的想法。M-H 的演算法如下,

1. 給定初始值 $\theta^{(0)}$,
2. 從 $q(\theta^*|\,\theta^{t-1})$ 抽出 θ^* ,
3. 計算接受率

$$\alpha(\phi|\theta) = \min\left[\frac{\pi(\phi)q(\theta|\theta)}{\pi(\theta)q(\phi|\theta)}, 1\right],$$

4. 從均勻分配抽出一個隨機變數 u , 當 $u \leq \alpha$ 我們令 $\theta^{(t)} = \theta^*$, 否則停留在原值 $\theta^{(t)} = \theta^{(t-1)}$
5. 重複步驟二到步驟四

而經過遞迴而得的數列, $\{\theta^{(1)}, \theta^{(2)}, \cdots, \theta^{(n)}\}$, 要經過一段時間數列才會收斂到我們的目標分配, 所以通常我們會捨棄前面 k 個值, 只保留後面的。要捨棄多少個值沒有定論, 一般來說如果發現產生的數列間的相關性非常大, 則需捨棄較多的數值。在 M-H 法中, 如何選擇候選分配是一門大學問, 只要是定義域相同, 對候選分配其實沒有什麼限制, 然而接受率皆希望能越大越好, 接受率大, 則需要遞迴的次數也就不需要那麼多。不過這端賴經驗, 與試誤法, 一般理論上沒有特定的規則。

4.3 演算法

給定觀察值之下, 參數與隱藏參數的後驗分配爲

$$\pi(\beta, f_t, \boldsymbol{\theta}, \mathbf{h} | \mathbf{y})$$

我們將參數空間切割, 分成幾個區塊, 在固定其他參數之下, 從條件分配中抽樣, 進而反覆疊帶、更新。模型估計的步驟簡述如下:

1. 設定 $(\boldsymbol{\theta}, \mathbf{h})$ 的起始值。
2. 從條件分配 $\mathbf{B}, f_t \mid \mathbf{y}, \mathbf{h}$ 抽取 $\mathbf{B}$ 和 f_t 。
3. 將模型轉換成 $p+k$ 維的單變量模型, 藉用 Chib et al. (2002) 的方法, 從條件分配 $\boldsymbol{\theta}_j, \mathbf{h}_j \mid \mathbf{y}_j, \mathbf{B}, f_t$ 抽取$(\boldsymbol{\theta}_j, \mathbf{h}_j)$, $j = 1, 2, \cdots, p+k$。
4. 回到步驟 2.

步驟 2 到步驟完成一個馬可夫鏈的轉換, 一般蒙地卡羅法需要大量的遞迴次數, 來確保後驗分配的收斂。抽樣的相關係係數越高, 要求的遞迴次數就越多。通常我們也捨棄模擬初始, 明顯未收斂的值 (burn in) , 只留下之後的模擬值。以下詳述各步驟的抽樣方法:

1. 要從條件分配 $\mathbf{B}, f_t \mid \mathbf{y}, \mathbf{h}_t$ 抽取 $\mathbf{B}$ 和 f_t, 我們先給定 f_t 抽取 $\mathbf{B}$, 再固定 $\mathbf{B}$, 抽取 f_t 。其中因爲認定條件的關係, 我們已將 $\mathbf{B}$ 的第一項設定爲 1 , 所以只需對係數未定的 β 做抽樣

   (1) β 的後驗分配爲:

$$\begin{aligned}\pi(\beta \mid \mathbf{y}, \mathbf{h}_t) &\propto p(\beta) \prod_{t=1}^{n} p(\mathbf{y}_t \mid \mathbf{B}, \mathbf{h}_t) \\ &\propto p(\beta) \prod_{t=1}^{n} N_p(\mathbf{y}|0, \mathbf{V}_t + \mathbf{B}\mathbf{D}_t\mathbf{B}')\end{aligned}$$

我們將以 M-H 法藉由多變量 t 分配, $\mathrm{T}(\mathbf{m}, \mathbf{\Sigma}, \nu)$ 來抽取 β。而 $\mathrm{T}(\mathbf{m}, \mathbf{\Sigma}, \nu)$ 中的參數 $(\mathbf{m}, \mathbf{\Sigma})$ 是數值極大化概似函數, $\log\{\prod_{t=1}^{n} N_p(\mathbf{y} \mid 0, \mathbf{\Omega}_t)\}$ 而來。[1]

$$\begin{aligned}\mathbf{m} = \arg\max l(\beta) &= \sum_{t=1}^{n} \log N_p(\mathbf{y} \mid 0, \mathbf{\Omega}) \\ &= const - \frac{1}{2}\sum_{t=1}^{n} \log|\mathbf{\Omega}_t| - \frac{1}{2}\sum_{t=1}^{n} \mathbf{y}_t' \mathbf{\Omega}_t^{-1}\mathbf{y}_t\end{aligned}$$

其中 $\mathbf{m}$ 爲上式的概似估計值, $\mathbf{\Sigma}$ 爲負 Hessian 矩陣的反矩陣。β 的先驗分配假設爲爲常態, 則我們從 $\mathrm{T}(\beta \mid \mathbf{m}, \mathbf{\Sigma}, \nu)$ 抽出 β^* , 接受 β^* 的機率爲:

$$\alpha(\beta, \beta^* \mid \mathbf{y}, \mathbf{h}_t) = \min\left\{1, \frac{p(\beta^*)\prod_{t=1}^{n} N_p(\mathbf{y} \mid 0, \mathbf{V}_t + \mathbf{B}^*\mathbf{D}_t\mathbf{B}^{*\prime})T(\beta \mid \mathbf{m}, \mathbf{\Sigma}, \nu)}{p(\beta)\prod_{t=1}^{n} N_p(\mathbf{y} \mid 0, \mathbf{V}_t + \mathbf{B}\mathbf{D}_t\mathbf{B}')T(\beta^* \mid \mathbf{m}, \mathbf{\Sigma}, \nu)}\right\}$$

如果 β 被拒絕, 則下一輪的 β 保持不變。

(2) 給定了 β 我們可以從 $f_t \mid \mathbf{y}, \mathbf{B}, \mathbf{h}$ 抽出 f 。根據傳統貝氏的計算,

$$\begin{aligned}p(f_t \mid \mathbf{y}, \mathbf{h}_t, \mathbf{B}) &\propto p(f_t) f(\mathbf{y}_t \mid f_t, \mathbf{B}, \mathbf{h}_t) \\ &\propto \exp\left[-\frac{1}{2}(\mathbf{y}_t - \mathbf{B}f_t)'\mathbf{\Omega}_t^{-1}(\mathbf{y}_t - \mathbf{B}f_t)\right] \times \exp\left[-\frac{f_t'\mathbf{D}_t^{-1}f_t}{2}\right] \\ &\propto \exp\left[-\frac{1}{2}(\mathbf{y}_t'\mathbf{\Omega}_t^{-1}\mathbf{y} - 2\mathbf{y}'\mathbf{\Omega}_t^{-1}\mathbf{B}f_t + f_t'\mathbf{B}'\mathbf{\Omega}_t\mathbf{B}f_t + f_t'\mathbf{D}_t^{-1}f_t)\right] \\ &\propto \exp\left\{-\frac{1}{2}\left[f_t'(\mathbf{B}'\mathbf{\Omega}_t^{-1}\mathbf{B} + \mathbf{D}_t^{-1})f_t - 2B'\mathbf{\Omega}_t^{-1}\mathbf{y}\right]\right\} \\ &\propto \exp\left[-\frac{1}{2}(f_t - \hat{f}_t)'\mathbf{F}^{-1}(f_t - \hat{f}_t)\right]\end{aligned}$$

---

[1]我們亦可提供一階導數, 來增加極大化的速度與準確度,

$$\begin{aligned}\frac{\partial l}{\partial \beta_{ij}} &= \frac{1}{2}\sum_{t=1}^{n}\left[\tilde{y}_t'\Omega_t^{-1}\frac{\partial \Omega_t}{\partial \beta_{ij}}\Omega_t^{-1}\tilde{y}_t - tr\left(\Omega_t^{-1}\frac{\partial \Omega_t}{\partial \beta_{ij}}\right)\right], \\ &= \sum_{t=1}^{n}\left[\tilde{y}_t'\Omega_t^{-1}\frac{\partial B}{\partial \beta_{ij}}D_tB'\Omega_t^{-1}\tilde{y}_t - tr\left(\Omega_t^{-1}BD_t\frac{\partial B'}{\partial \beta_{ij}}\right)\right]。\end{aligned}$$

表 1: 混合常態

s	$\Pr(s=i)$	m_i	v_i^2
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

所以 f_t 的後驗分配爲平均數爲 $\mathbf{F}_t\mathbf{B}'(\mathbf{V}_t^*)^{-1}\mathbf{y}_t$, 變異數爲 $\mathbf{F}_t = \left(\mathbf{B}'(\mathbf{V}_t^*)^{-1}\mathbf{B} + \mathbf{D}_t^{-1}\right)^{-1}$ 的常態分配。

2. 再給定 $\{\mathbf{y}_t, \mathbf{B}, f_t\}$, 且因子與誤差的條件分配爲相互獨立的常態分配, 我們可以將模型分解成 $p+k$ 個常態狀態空間模型。令 $\boldsymbol{\alpha}_t = \mathbf{B} f_t$,

$$y_{jt} - \alpha_{jt} = \exp\left(h_{jt}/2\right)\varepsilon_{jt},$$

$$z_{jt} = \log\left(y_{jt} - \alpha_{jt}\right)^2 = h_j + \log\left(\varepsilon_t\right)^2$$

其中, $\log(\varepsilon_t)^2$ 爲取對數的卡方隨機變數。爲了將上式轉爲常態分配, 我們引用 Kim et al. (1998) 的方法, 以混和的常態分配來模擬取對數的卡方隨機變數。在這裡我們創造出一個指示變數 s_t , 指示變數的作用是指明抽取何種常態分配的隨機變數。理論上以越多階的常態分配來模擬取對數的卡方分配會越準確, 但也越複雜。在可接受的誤差範圍之下, 我們以七個常態分配來模擬 $\log(\varepsilon_t)^2$, 則在給定指示變數 s_t 之下,

$$\log(\varepsilon_t)|s_t \sim \mathbf{N}\left(m_{st}, v_{st}^2\right)^2,$$

$$\Pr(s_t = i) = q_i, \quad i \leq 7,$$

而指示變數的機率分配以及其所代表的常態分配如表 1 。

在給定指示變數 s_{jt} 下,

$$z_{jt}|s_{jt}, h_{jt} \sim N\left(h_{jt} + m_{s_{tj}}, v_{stj}^2\right)$$

$$h_{jt} - \mu_j = \phi_j(h_{jt-1} - \mu_j) + \sigma_j \eta_{jt}, \;\; j \leq p + k$$

現在我們說明如何對 $p + k$ 個常態狀態模型中參數的抽樣。由於模型的因子與誤差的條件分配爲相互獨立, 使得,

$$p(\mathbf{s}, \boldsymbol{\theta}, \mathbf{h} \,|\, \mathbf{z}) = \prod_{j=1}^{k+p} p(\mathbf{s}_j, \boldsymbol{\theta}_j, \mathbf{h}_j|\mathbf{z}_j), \qquad \theta_j = (\mu_j, \phi_j, \sigma_{\eta_j})$$

也就是說我們可以逐一對單變量中的參數做抽樣。我們將 $(\mathbf{s}_j, \boldsymbol{\theta}_j, \mathbf{h}_j|\mathbf{z}_j)$ 分成三個區塊 。

(1) $\mathbf{s}_j$ 的後驗分配爲

$$Pr(\mathbf{s}_j|\mathbf{z}_j, \mathbf{h}_j) \propto Pr(\mathbf{s}_j) N\left(\mathbf{z}_j \,|\, \mathbf{h}_j + \mathbf{m}_{st}, \boldsymbol{v}_{st}^2\right)$$

$$= Pr(\mathbf{s}_j) \prod_{t=1}^{n} \mathbf{N}(z_{jt} \,|\, h_{jt} + m_{s_{jt}}, v_{s_{jt}}^2)$$

$\mathbf{s}_j$ 的先驗分配即如表一。

(2) $\boldsymbol{\theta}_j$ 的後驗分配爲

$$p(\boldsymbol{\theta}_j \,|\, \mathbf{z}_j, \mathbf{s}_j) \propto p(\boldsymbol{\theta}_j) p(\mathbf{z}_j \,|\, \mathbf{s}_j, \boldsymbol{\theta}_j)$$

$$= p(\boldsymbol{\theta}_j) p(z_{j1} \,|\, \mathbf{s}_j, \boldsymbol{\theta}_j) \prod_{t=2}^{n} N(z_{jt}|m_{sjt} + \hat{h}_{t|t-1}, f_{t|t-1}),$$

其中 $\{\hat{h}_{t|t-1}, f_{t|t-1}\}$ 是根據 Kalman filter 遞迴所計算而來, 遞迴公式 (附錄一) 如下:

$$\hat{h}_{t|t-1} = \mu + \phi(\hat{h}_{t|t-1} - \mu), \qquad p_{t|t-1} = \phi^2 p_{t|t-1} + \sigma_\eta^2,$$

$$f_{t|t-1} = p_{t|t-1} + v_{st}^2, \qquad k_t = p_{t|t-1} f_{t|t-1}^{-1},$$

$$\hat{h}_{t|t} = \hat{h}_{t|t-1} + k_t \left(z_j - m_{st} - \hat{h}_{t|t-1}\right) 。$$

對 $\boldsymbol{\theta}_j$ 的抽樣方法, 類似於抽取負荷矩陣 $\mathbf{B}$, 同樣是以 M-H 法從 $\mathrm{T}(\mathbf{m}, \mathbf{V}, \xi)$ 抽出 θ^* , 而接受率爲 α , $(\mathbf{m}, \mathbf{V})$ 求取的方式仍是以數值極大化求解, ξ 任意給定爲 10。在常態假設之下, θ 後驗分配的概似函數可寫成,

$$\log g(\theta) \propto \frac{1}{2}\sum_{t=1}^{N} \log f_{t|t-1} - \frac{1}{2}\sum_{t=1}^{N} \frac{e_t^2}{f_{t|t-1}}$$

$$e_t = z_j - m_{s_t} - \hat{h}_{t|t-1}$$

而

$$\mathbf{m} = arg \max_{\boldsymbol{\theta}} \log(l(\boldsymbol{\theta}))$$

$$\mathbf{V} = \left[\frac{-\partial^2 \log(l(\boldsymbol{\theta}))}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}}\right]^{-1}$$

其中 $\mathbf{V}$ 爲負的 Hessian 的反矩陣, 則我們的 M-H 爲從 $f_T(\mathbf{m}, \mathbf{V}, \xi)$ 抽出 θ^* 接受機率爲

$$\alpha(\theta, \theta^* \mid z_j, s) = \min\left[1, \frac{g(\theta^*)}{g(\theta)} \frac{f_T(\theta \mid \mathbf{m}, \mathbf{V}, \xi)}{f_T(\theta' \mid \mathbf{m}, \mathbf{V}, \xi)}\right]$$

(3) 狀態變數 h_t 的抽樣

對狀態變數 h_t 的抽樣, 我們運用 de Jong 與 Shephard (1995) 的模擬平滑法 (signal simulatin smoother) , 有別其他方法, 這個方法最大的不同, 是它不是對狀態變數抽樣, 而是給定所有觀察值之下, 藉由對模型中誤差項的事後分配抽樣, 然後形成對狀態變數的模擬。在 kalman recursion 的計算之中, 我們保留 e_t, f_t, k_t , 以向後遞迴的方式逐一抽出 h_t 。首先定義 $n_t = f_{t|t-1}^{-1} + k_t^2 u_t$, $d_t = f_{t|t-1}^{-1} e_t - r_t k_t$, 令 $r_n = u_n = 0$, 則其遞迴公式 (附錄二) 如下,

$$c_t = v_{s_t}^2 - v_{s_t}^4 n_t, \qquad \zeta_t \sim N(0, c_t), \qquad b_t = v_{s_t}^2 (n_t - \phi k_t u_t),$$

$$r_{t-1} = f_{t|t-1}^{-1} e_t + (\phi - k_t) r_{t-1} - b_t \frac{\zeta}{c_t},$$

$$u_{t-1} = f_{t|t-1}^{-1} + (\phi - k_t)^2 u_t + \frac{b_t^2}{c_t}$$

則狀態變數 $h_t = z_j - m_{s_t} - v_{s_t}^2 d_t - \zeta_t,\ t = n, \cdots 1$ 。

4.4 模擬驗證

爲了驗證上述演算法的正確性, 我們依據選定的參數, 模擬數據資料。下表列出參數的眞實值與模型的估算值, 其中樣本數 2000 , 模擬次數 12000 次, 捨棄前面 2000 個值。表列爲重複執行 30 次的平均值。眞實值的選取爲先對實際資料估計, 以得到的估計值再整數化作爲我們模擬眞實設定的依據。估算的經驗中歸納幾點發現:

表 2: 模擬參數估計值

	眞實值				估計值			
	y_1	y_2	y_3	f	y_1	y_2	y_3	f
μ	1.500	1.000	0.000	1.500	1.5312	0.9777	-0.0397	1.4781
					(0.237)	(0.097)	(0.081)	(0.054)
ϕ	0.985	0.950	0.850	0.850	0.9786	0.9214	0.8633	0.8507
					(0.007)	(0.029)	(0.012)	(0.019)
σ_η	0.150	0.200	1.000	0.250	0.1704	0.2575	0.1538	0.2679
					(0.022)	(0.055)	(0.029)	(0.031)
β	1.000	1.200	0.800		1.0000	1.1998	0.8001	
						(0.035)	(0.027)	

表列爲樣本數 2000 , 遞迴次數 12000 次, 捨棄前面 2000 個值, 重複執行 30 次的平均值。括弧內的數值爲 RMSE。

- 程式以 Ox 語言編寫, 在 CPU 1.8 GHz, 記憶體 1GB 的筆記型電腦上執行一次, 大約耗時 10 小時。
- 大抵上估計值都算準確, 不過通常 ϕ 值皆會產生低估, 但低估的程度在可以接受的範圍。另外,

σ_η 值皆會高估, 且高估的程度與原來參數的眞實大小值有關, 眞實的 σ 越大, 高估的比例越大。

- μ 與 β 的值相對比較容易收斂, 其變異程度也較小。
- 改善偏誤的方法, 我們發現以增加樣本數量的方式會比增加遞迴次數來的有效。