第五章 水文時間序列資料分析

隨時間變化的氣象水文變數，通常具有可用物理影響因素和數學控制方程式解釋的定率(deterministic)部分，以及無法解釋的隨機分量。當 $X(t)$ 具有隨機分量，且為時間序列變數，則 $X(t)$ 稱為序率過程(stochastic process)。因為前後時間的「隨機分量」往往存在關聯性，若能透過樣本資料分析得到其「統計結構」，則可以發展整合**定率**與**序率**的模式，增強解釋能力（估計準確度）或預報能力。

水文時間序列資料通常是等時間間距的離散樣本，分析的目的包括：建立預報(forecasting)或產生合成資料(data generation)用的數學模式，特別是河川流量和降雨時間序列的分析、合成等應用。

5.1 資料定常性

分析序率過程的目的，是要找出不同時間隨機變數值之間的統計關係，例如聯合機率密度函數。唯時間序列資料往往在每個時間只有一個樣本，必須將不同時間的樣本合併計算，才能估計隨機變數的聯合機率密度函數或動差值。針對這樣的應用目的，必須在序率過程具有時間定常性(Stationarity)條件下[1]，才能將不同時間的樣本合併計算。

5.1.1 嚴謹定常性(Strict-Sense Stationarity, SSS)

一個序率過程(stochastic process) $X(t)$，機率密度函數為 $P\left[X(t)=x\right]$，若且唯若對於任意的 $t_1,t_2,\cdots,t_k$ 和 τ，皆能滿足下列聯合機率密度函數恆等關係式，則 $X(t)$稱作是具有 k 階定常性：

$$\begin{aligned}&P\left[X(t_1)=x_1,\ X(t_2)=x_2,\ \cdots,\ X(t_k)=x_k\right]\\&=P\left[X(t_1+\tau)=x_1,\ X(t_2+\tau)=x_2,\ \cdots,\ X(t_k+\tau)=x_k\right]\end{aligned} \tag{5-1}$$

若對於所有可能的 k 值，$k=1,2,...,n\to\infty$，和所有的 τ 值，上式都能滿足，則稱序率過程 $X(t)$具有嚴謹定常性(SSS)。

[1] 定常性假設是否成立，可以將時間序列分段，設計統計假設檢定試驗測試之。

以上定常性的意思，是指 $X(t)$序率過程的聯合機率密度函數，僅為時間差(time difference)的函數，不是絕對時間參考點(reference)的問題。易言之，將一組序率變數在時間軸上同步移動、關係位置不變，則其聯合機率密度函數不變。圖 5-1 表示隨機變數 $X(t)$具有一階定常性，即在任何不同時間，$X(t)$的機率密度函數分佈均相等。如擬繪圖表示二階定常性，可以用 $\left[X(t_1)=\right]x_1$ 和 $\left[X(t_2)=\right]x_2$ 作為為兩個變數實現值的座標軸，使用機率等值圖方式表現聯合機率密度函數，以二維圖表示三度空間的分部狀況，將第三個軸保留為時間軸；若且唯若二階定常性成立，則沿時間軸的所有聯合機率密度函數機率等值圖都相同。

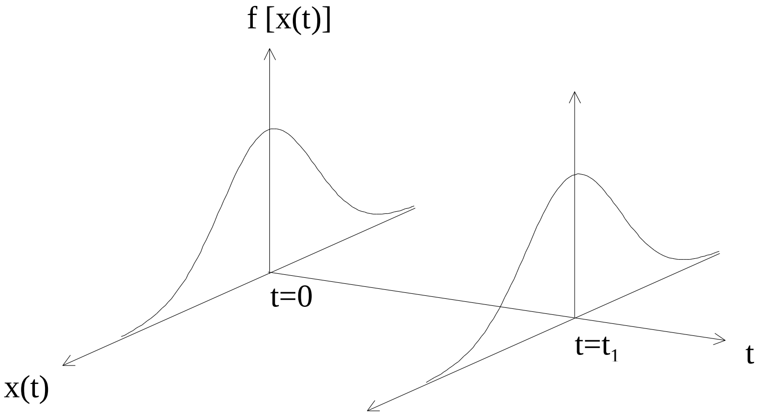


圖5-1 隨機變數一階定常性之示意圖

若序率過程具有高階定常性，則所有較低階的定常性都必然會成立。例如，具備二階定常性 $P\left[X(t_1)=x_1, X(t_2)=x_2\right]=P\left[X(t_1+\tau)=x_1, X(t_2+\tau)=x_2\right]$，可以證明一階定常性必然成立。由於 $P\left[X(t_1)=x_1, X(t_2)=x_2\right]$ 的邊際分佈(marginal distribution) 是 $P\left[X(t_1)=x_1\right]$；$P\left[X(t_1+\tau)=x_1, X(t_2+\tau)=x_2\right]$ 的邊際分佈是 $P\left[X(t_1+\tau)=x_1\right]$ 是因此：

$$\begin{aligned} P\left[X(t_1)=x_1\right] &= \int_{X(t_2)=x_2} P\left[X(t_1)=x_1, X(t_2)=x_2\right]\,dx_2 \\ &= \int_{X(t_2+\tau)=x_2} P\left[X(t_1+\tau)=x_1, X(t_2+\tau)=x_2\right]\,dx_2 \qquad (5\text{-}2) \\ &= P\left[X(t_1+\tau)=x_1\right] \end{aligned}$$

上式推論證實 $P\left[X(t_1)=x_1\right]=P\left[X(t_1+\tau)=x_1\right]$，即一階定常性成立。

例題5-1 $X(t)=A\boldsymbol{cos}(2\pi f_0 t+\theta)$，$Y(t)=A\boldsymbol{sin}(2\pi Ft)$。其中，$A$ 和 f_0 為常數，F 和 θ 都是均勻分佈(uniform distribution)的隨機變數，定義如

下：

$$P_\theta(\theta)=\begin{cases}(2\pi)^{-1} & 0\le\theta\le 2\pi\\ 0 & \text{elsewhere}\end{cases}\quad P_f(F)=\begin{cases}f_o^{-1} & f_o\le F\le 2f_o\\ 0 & \text{elsewhere}\end{cases}$$

闡明$X(t)$具有嚴謹定常性(SSS)；而$Y(t)$則不具有定常性。

解：(1) 因為$2\pi f_o t+\theta$的分布是範圍$\left[2\pi f_o t, 2\pi f_o t+2\pi\right]$的均勻分布，因此對於任意時間$t_1$和$t_2$，$P\left[X(t_1)=x\right]=P\left[X(t_2)=x\right]$，對於所有的$x\in\left(-A,A\right)$都能成立，所以$X(t)$具有一階定常性；同時，$P\left[X(t_1)=x_1, X(t_2)=x_2\right]=P\left[X(t_1+\tau)=x_1, X(t_2+\tau)=x_2\right]$，所以$X(t)$具有二階定常性；依此類推，可知$X(t)$具有任意$k$階的定常性，因此可說$X(t)$具有嚴謹定常性(SSS)。

(2) 隨機變數$Y(t)$的一階動差（期望值）為：

$$E[Y(t)]=\int_{F=f_o}^{2f_o}\frac{A\boldsymbol{sin}(2\pi Ft)}{f_o}\,dF$$

$$=\frac{A}{2\pi f_o t}\left[\boldsymbol{cos}(2\pi f_o t)-\boldsymbol{cos}(4\pi f_o t)\right]\neq\text{constant}$$

對隨機變數F積分，計算隨機變數$Y(t)$的期望值，結果仍是自變數時間t的函數，而不是常數。由於高階定常性（p命題）可以推論低階定常性（q命題）；因此當變數$Y(t)$的一階動差不為常數得到~q的前提命題時，可以得到$\sim p$的推論，意即不具有一階定常性；繼續推論可以得到任意k階皆不具有定常性的結論。

5.1.2 廣義定常性(Wide-Sense Stationarity, WSS)

利用前二階動差，可以定義序率過程的廣義定常性。序率過程$X(t)$若滿足以下兩個條件，便可稱為是具有廣義定常性：

(1) 平均值不變：$E\left[X(t)\right]=\mu_x(t)=\mu_x(t+\tau)=\mu_x$ (constant)

(2) 協變異數不變：

$$\boldsymbol{cov}\left[X(t),X(s)\right]=\boldsymbol{cov}(t,s)=\boldsymbol{cov}(\left|t-s\right|)=\boldsymbol{cov}\left[X(\tau),X(\tau\pm\left|t-s\right|)\right]$$

上式中，t及s分別代表時間軸上兩個不同的時間點，令$\tau=\left|t-s\right|$。上式的等價表示法為：

(1) 平均值：$E\left[X(t)\right]=\mu_x(t)=\mu_x$ (constant)

(2) 協變異數：$\boldsymbol{cov}(t,s)=\boldsymbol{cov}(\tau)$

二階定常性表示「機率密度函數」不隨時間改變，廣義定常性則表示「一階和二階動差」不隨時間改變。若機率密度函數不隨時間改變，則一、二階動差必然不隨時間改變；反之，若一、二階動差不隨時間改變，未必表示機率密度函數不會隨時間改變。因此，若一個序率過程具有二階定常性，則可推知其具有廣義定常性；若已知廣義定常性成立，並不能推論得到二階定常性。

因為常態分布的三階或更高階動差，皆為一、二階動差的函數，所以若序率過程的隨機分量呈常態分布，並且滿足廣義定常性，則該過程即具有嚴謹定常性。意思是，常態分佈的序率變數，只要能得到一、二階動差，即可得到該隨機變數的完整機率分佈。

5.2 統計結構分析

分析水文資料時間序列時，經常將變數 $X(t)$ 分為：(1)長期趨勢(trend)分量 T_t，(2)週期效應(periodicity)分量 P_t，以及(3)序率(stochastic)分量 z_t 等三個分量的和，即 $X_t=T_t+P_t+Z_t$；其中，$E\left[X_t\right]=T_t+P_t$，即前二部份是以**時間**為自變數的定率函數描述的一階分量，有時又將兩者合稱為「趨勢」。序率分量的期望值 $E\left[Z_t\right]=0$，分析序率分量，主要是其二階統計特性，假設其變異數和不同時間差的自相關函數(autocorrelation)皆具有定常性。分析時間序列的程序通常是：(1)將序率過程 X_t 分為趨勢分量、週期效應 T_t+P_t，與序率分量 Z_t；(2)再分析序率分量(stochastic component)的「統計結構」。

分離 X_t 時間序列中定率分量和序率分量時，趨勢分量和週期分量的估計正確性或準確度，會影響序率分量「統計結構」 的估計。估計序率過程 X_t 中定率分量 T_t+P_t 的方法主要有二：(1)若是 X_t 樣本的時間很長，即樣本時間長度遠遠超過「相關距離」或「積分長度」(integral length)，則通常先採用迴歸法估計定率分量 T_t+P_t，再估計 $Z_t=X_t-T_t-P_t$；(2)若 X_t 樣本的時間不是很長（即樣本時間長度並不遠超過「相關距離」），先去除定率項、再估計序率分量統計結構和參數的兩階段處理作法，可能會造成定率分量和序率分量參數估計誤差，則建議採用最大概似法(Maximum Likelihood Estimator)同

時估計定率分量$T_t + P_t$和序率分量Z_t的統計結構（參數）。

分析序率分量Z_t常使用的數學模式，包括移動平均模式(moving average – MA model)、自迴歸模式(autoregressive – AR model)及結合以上二種方法的自迴歸移動平均模式(autoregressive moving average – ARMA model)等。

(一) 轉折點數測試

轉折點數檢測可協助判定測試資料是否為隨機，或是有某種趨勢存在。假設一時間序列為$x_1, x_2, \cdots\cdots x_n$，則任一時間值$x_i$和相鄰兩觀測值的關係共有 6 種：

$$\begin{array}{ll} 1. x_{i-1} < x_i < x_{i+1} & 2. x_{i-1} < x_{i+1} < x_i \\ 3. x_i < x_{i-1} < x_{i+1} & 4. x_i < x_{i+1} < x_{i-1} \\ 5. x_{i+1} < x_{i-1} < x_i & 6. x_{i+1} < x_i < x_{i-1} \end{array} \tag{5-3}$$

所謂轉折點即x_i為峰值或谷值，即上述 2、3、4 及 5 共四種情形。若資料為隨機，則轉折點的個數應接近總資料數的$4/6 = 2/3$。但是由於轉折點不會出現在第一點和最後一點，因此轉折點個數 p 的期望值為$\mathrm{E}(p) = \frac{2}{3}(n-2)$，變異數$\mathrm{var}(p) = (16n - 29)/90$。計算時間序列中的轉折點個數$p$，以 2-tailed 常態分佈統計檢定檢測$p$值，可以判斷一個時間序列是否為隨機；轉折點個數太多或太少，均表示可能有某種趨勢或週期存在。

(二) 長期趨勢分析

水文資料具有長期趨勢的原因，可能是由於自然界的影響，或是由於人為環境開發所形成的變遷。長期趨勢函數分析的方式通常是將資料的時間序列繪圖，決定函數型式再加以配套，計算函數係數。趨勢函數可為最簡單的線性迴歸函數：

$$T_t = \alpha_0 + \alpha_1 t \tag{5-4}$$

或是多項式形式：

$$T_t = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \cdots\cdots \alpha_n t^n \tag{5-5}$$

或是指數型式：

$$T_t = \alpha_0 \exp\left(\alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3 + \cdots\cdots \alpha_n t^n\right) \tag{5-6}$$

欲判定一時間序列隨機變數是否具有趨勢，可使用線性迴歸法測試線性趨勢

係數是否顯著、採用 Kendall 排序關係測試(Kendall's rank correlation test)、趨勢或轉折點檢測協助分析。

(三) 利用調和分析辨識週期結構

水文資料週期性的主要原因是地球的公轉和自轉，二者分別造成季節變化和日差(diurnal)變化。除此以外，月亮和地球的關係位置影響到潮汐，形成另一種週期變化。

在去除長期平均趨勢後，時間序列中的週期效應，可表示為一組正弦函數，此分析步驟稱為調和分析(harmonic analysis)。假設時間序列 $x_1, x_2, \cdots\cdots x_n$ 的時間間距為 Δt，已知時間序列含有某種週期性的分量(component)，則完整的時間序列以調和分析表示為：

$$x_t = \mu + \sum_{i=1}^{L} \lambda_i \sin\left(\frac{2i\pi}{T} t + \phi_i\right) + z_t \qquad t = 1, \cdots, n \tag{5-7}$$

其中 $\mu = \mathrm{E}(x)$ 為時間序列的定常期望值，或時間平均值；i/T 和 T/i 分別代表頻率和週期；λ_i 和 ϕ_i 分別為振幅(amplitude)和相位(phase)；z_t 則為去除週期效應之後的序率部份。由於一個週期性函數最少需要三點(或 $2\Delta t$)才能代表一個週期，所以若資料個數 N 為偶數則 $L = N/2$；若 N 為奇數則 $L = (N-1)/2$。

週期 T 通常選擇某種和地球公轉或自轉有關的時間長度，例如分析季節性的變化則 T=1 年；若是分析資料在一天內的日差效應，則取 T=1 天。假設一個時間週期長度 T 之內，共有 p 個 Δt 間距的「季」資料，例如 T=1 年，Δt =1 個月，則 p=12。令 $n = N/p$ 為資料年數。則每季的樣本平均值為：

$$m_\tau = \frac{1}{n} \sum_{i=1}^{n} x_{\tau + p \times (i-1)} \qquad \tau = 1, 2, \cdots, p \tag{5-8}$$

調和分析的目的是將 m_τ 表示為一組正弦和餘弦函數的組合。假設 $\hat{m}_\tau$ 為以調和分析所得到的 m_τ 的擬合值：

$$\begin{aligned} \hat{m}_\tau &= m_\tau - \varepsilon_\tau \\ &= \mu + \sum_{i=1}^{p/2} \alpha_i \sin\left(\frac{2i\pi}{p}\tau\right) + \sum_{i=1}^{p/2} \beta_i \cos\left(\frac{2i\pi}{p}\tau\right) \qquad \tau = 1, 2, \cdots, p \end{aligned} \tag{5-9}$$

上式為週期性季平均資料的表示法，資料數為 p 個，所以最多只能決定 p 個參數。若 p 為偶數，則正弦和餘弦函數最多各取 $p/2$ 項(如此一來，待訂參數似乎變成 1+p 個：μ、α_i 及 β_i，$i = 1, \cdots, p/2$，但是由於 $\alpha_{p/2} \equiv 0$，所以仍是

實際上未知數為 p 個)；若 p 為奇數，則以 $(p-1)/2$ 取代 $p/2$。令估計誤差變異數為：

$$
\begin{aligned}
L &= \sum_{\tau=1}^{p}\left(m_\tau - \hat{m}_\tau\right)^2 \\
&= \sum_{\tau=1}^{p}\left[m_\tau - \mu - \sum_{i=1}^{p/2}\alpha_i \sin\left(\frac{2i\pi}{p}\tau\right) - \sum_{i=1}^{p/2}\beta_i \cos\left(\frac{2i\pi}{p}\tau\right)\right]^2
\end{aligned}
\tag{5-10}
$$

由最小估計誤差原理，分別對未知數 μ 、α_i 及 β_i 取偏微分，並令為零：

$$
\frac{\partial L}{\partial \mu} = -2\sum_{\tau=1}^{p}\left(m_\tau - \mu - \sum_{i=1}^{p/2}\alpha_i \sin\left(\frac{2i\pi}{p}\tau\right) - \sum_{i=1}^{p/2}\beta_i \cos\left(\frac{2i\pi}{p}\tau\right)\right) = 0
$$

$$
\frac{\partial L}{\partial \alpha_i} = 2\sum_{\tau=1}^{p}\left\{\sin\left(\frac{2\pi i}{p}\tau\right)\left[m_\tau - \hat{\mu} - \sum_{i=1}^{p/2}\hat{\alpha}_i \sin\left(\frac{2\pi i}{p}\tau\right) - \sum_{i=1}^{p/2}\hat{\beta}_i \cos\left(\frac{2\pi i}{p}\tau\right)\right]\right\} = 0
$$

$$
\frac{\partial L}{\partial \beta_i} = 2\sum_{\tau=1}^{p}\left\{\cos\left(\frac{2i\pi}{p}\tau\right)\left[m_\tau - \hat{\mu} - \sum_{i=1}^{p/2}\hat{\alpha}_i \sin\left(\frac{2i\pi}{p}\tau\right) - \sum_{i=1}^{p/2}\hat{\beta}_i \cos\left(\frac{2i\pi}{p}\tau\right)\right]\right\} = 0
$$

for $i = 1,2,\cdots,p/2$。若是資料為等間距的資料，則可以使用正弦及餘弦函數的正交特性：

$$
\sum_{\tau=1}^{P}\sin\left(\frac{2i\pi}{P}\tau\right) = \sum_{\tau=1}^{P}\cos\left(\frac{2i\pi}{P}\tau\right) = 0 \qquad \text{for } \forall i = 1,2\cdots\frac{p}{2}
$$

$$
\sum_{\tau=1}^{p}\sin\left(\frac{2i\pi}{p}\tau\right)\cos\left(\frac{2k\pi}{p}\tau\right) = 0 \qquad \text{for } \forall i,k = 1,2,\cdots,\frac{p}{2}
$$

並且當 $i \neq k$ 時：

$$
\sum_{\tau=1}^{p}\sin\left(\frac{2i\pi}{p}\tau\right)\sin\left(\frac{2k\pi}{p}\tau\right) = \sum_{\tau=1}^{p}\cos\left(\frac{2i\pi}{p}\tau\right)\cos\left(\frac{2k\pi}{p}\tau\right) = 0
$$

當 $i = k \neq p/2$ 時：

$$
\sum_{\tau=1}^{p}\sin^2\left(\frac{2i\pi}{p}\tau\right) = \sum_{\tau=1}^{p}\cos^2\left(\frac{2i\pi}{p}\tau\right) = \frac{p}{2} \qquad \text{for } i = 1,2\cdots,\left(\frac{p}{2}-1\right)
$$

當 $i = k = p/2$ 時：

$$
\sum_{\tau=1}^{p}\sin^2\left(\frac{2i\pi}{p}\tau\right) = 0 \qquad \sum_{\tau=1}^{p}\cos^2\left(\frac{2\pi i}{p}\tau\right) = p
$$

將以上結果代入 $\partial L/\partial\mu = 0$ 、$\partial L/\partial\alpha_i = 0$ 及 $\partial L/\partial\beta_i = 0$ 式子中，分別導出未知

數 μ、α_i 和 β_i。

$$\mu = \frac{1}{p}\sum_{\tau=1}^{p} m_\tau$$

$$\hat{\alpha}_i = \frac{2}{p}\sum_{\tau=1}^{P} m_\tau \sin\left(\frac{2i\pi}{p}\tau\right) \quad \text{for } i = 1,2\cdots\cdots,\left(\frac{p}{2}-1\right),\ \hat{\alpha}_{p/2} = 0$$

$$\hat{\beta}_i = \frac{2}{p}\sum_{\tau=1}^{P} m_\tau \cos\left(\frac{2i\pi}{p}\tau\right) \quad \text{for } i = 1,2\cdots\cdots,\left(\frac{p}{2}-1\right),\ \hat{\beta}_{p/2} = \frac{1}{P}\sum_{\tau=1}^{P} m_\tau (-1)^\tau$$

嚴格而言，以上分析只有在自相關係數接近於零(否則 σ_τ^2 的值會彼此互相影響)，並且 $\sigma_\tau^2, \tau = 1,2\cdots p$ 沒有週期現象時才能應用。

以調和分析所得的結果為資料的週期效應部份，將時間序列資料減去長期平均趨勢和週期效應之後的剩餘值是定率部份所無法解釋的「序率」項部份，$z_t = x_t - T_t - P_t$。序率部份可分為統計上可解釋的時間相關性(Sequential Dependence)和統計上也無法解釋的完全獨立的噪音項(noise)。

例 5-1 英國 Teme 河在 Tenbury Wells 地方量測的月流量(轉為降雨深度)，在 1957~1964 年之間的紀錄如下表

	月流量(mm)											
年	1 月	2 月	3 月	4 月	5 月	6 月	7 月	8 月	9 月	10 月	11 月	12 月
1957	37	64	42	15	8	5	4	39	57	27	60	35
1958	56	89	34	17	10	19	16	19	68	85	31	50
1959	84	22	27	39	21	8	5	4	2	3	19	90
1960	122	80	50	36	11	6	4	5	24	102	114	99
1961	61	46	18	38	36	8	6	4	3	20	16	46
1962	84	31	21	48	21	10	7	14	36	19	36	36
1963	23	19	106	49	20	10	13	6	6	5	56	24
1964	15	21	48	24	17	12	8	4	3	4	7	39

試以調和分析求以上資料的年週期平均，假設可忽略因自相關係數造成的 σ_τ^2 偏估和 σ_τ^2 的週期性。

解： 首先求8年各月的平均流量，得到 m_τ=60.25、46.50、43.25、33.25、18.00、9.75、7.88、11.88、24.88、33.13、42.38、52.68。序列的平均值 $\bar{x} = 31.96$。由調和分析求出正弦和餘弦函數的係數for $i = 1,2,\cdots,p/2$ 如下表。

i	$\hat{\alpha}_i$	$\hat{\beta}_i$
1	10.41	21.06
2	-0.04	-0.69
3	1.60	1.35
4	2.27	0.60
5	0.38	-1.10
6	0	-0.18

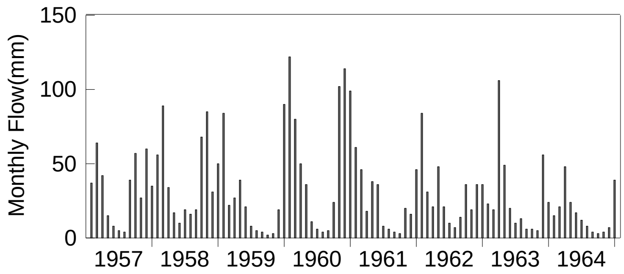
調和分析的目的是分析時間序列中是否有週期現象，虛無假設為第i個調和分量解釋的變異數(variance explained) $\frac{N}{2}\left(\hat{\alpha}_i^2+\hat{\beta}_i^2\right)$為零。

由右表可看出$i=1$的係數α_1、β_1最大，其他的均接近於零。以F分布隨機變數的方式分析個各個調和分量的重要性，由$i=1$開始，逐漸分析「解釋」和「未解釋」變異數的比率。

$$\frac{N}{2}\left(\hat{\alpha}_1^2+\hat{\beta}_1^2\right)=26484 \qquad \text{參數個數(d.o.f)} = 2$$

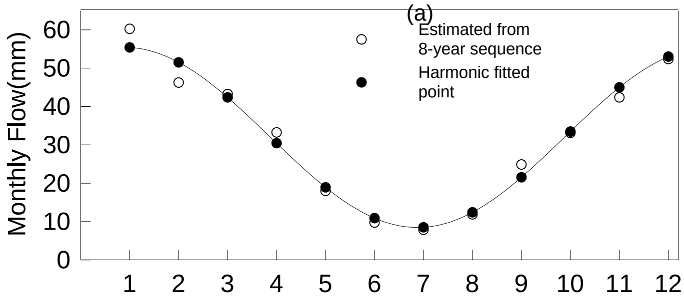
$$\sum_{i=2}^{6}\frac{N}{2}\left(\hat{\alpha}_i^2+\hat{\beta}_i^2\right)=597 \qquad \text{d.o.f} = 9 \ \ (\because \hat{\alpha}_6 \cong 0)$$

8年的月流量繪圖如右之(a)圖。



(a)

8年的月平均流量調和分析的結果，只包含$i=1$項的第一調和分量繪圖如右之(b)圖。



(b)

以F分布隨機變數分析：

$$\text{Total variance} = \sum_{t=1}^{96}\left(x_t-\bar{x}\right)^2=\sum_{t=1}^{96}\left(x_t-31.96\right)^2=78844$$

Total variance的自由度d.o.f. = 96-1 = 95

第一個harmonic的顯著測試：

$$\text{F} = \frac{(26484/2)}{(78844-26484)/(95-2)}=23.52$$

查表$F_{(2,93)}=7.5$ for $\alpha=0.001$，$F > F_{(2,93),0.1\%}$，因此第一調和分量非常顯著$\Rightarrow$選擇H_1 Hypothesis。

其次，將$i = 2, 3, \ldots, 6$各項的「解釋」和「未解釋」變異數比值求出，檢測是否顯著。總變異數= 78844；第一調和分量解釋變異數= 26486，

尚餘 78844-26486 = 52358

d.o.f = 95-2 = 93。

變數分析表

Variance	*Sum of squares*	*Degree of freedom*	*F ratio*
(1) α_i， β_i， i =2, 3, 4, 5, 6	$\sum_{i=2}^{6} \frac{N}{2}\left(\hat{\alpha}_i^2 + \hat{\beta}_i^2\right)$ = 597	9	$(597/9)/(51761/84)$ $= 0.11$
(2) α_1， β_1	$\frac{N}{2}\left(\hat{\alpha}_1^2 + \hat{\beta}_1^2\right)$ = 26486	2	$(26484/2)/(52358/93)$ $= 23.52$
(3) Residuals	51761	84	
(4) Total	$\sum_{t=1}^{96} (x_t - \bar{x})^2$ = 78844	*N*-1 = 95	

在52358之中，第2~6個調和分量又解釋了597的變異數，使用參數值= 9。剩餘的變異數等於52358-579 = 51761，及d.o.f. = 93-9 = 84。

$$\therefore \mathrm{F} = {(597/9)}\Big/{(51761/84)} = 0.11$$

統計不顯著，因此接受虛無假設，並令$\alpha_i = \beta_i = 0$, for $i = 2, 3, \ldots, 6$。

5.3 時間序列自相關分析

假設$X(t)$為某單站的定常時間序列隨機變數，兩個相距時間(稽延時間)為$k = |i - j|$的隨機變數的協變異數定義為：

$$C_{ij} = \boldsymbol{cov}(X_{t_i}, X_{t_j}) = \boldsymbol{E}\left[(X_{t_i} - \mu)(X_{t_j} - \mu)\right] \qquad (5\text{-}11)$$

其中，μ為$X(t)$的期望值，k=0 時$C_{ii} = \sigma^2$。自相關係數定義為：

$$\rho_{ij} = C_{ij} \big/ \sigma^2$$

假設隨機變數的實現值(realization)為一組離散樣本$x_1, x_2, \cdots, x_n$，則樣本的參數估計如下：

(1) 平均值：$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$ (5-12a)

(2) 變異數：$s^2=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\bar{x}\right)^2$ (5-12b)

(3) 協變異數：$\hat{C}_{ij}=\boldsymbol{c\hat{o}v}\left(X_{t_i},X_{t_j}\right)=\frac{\sum_{l=1}^{n-k}\left(x_l-\bar{x}\right)\left(x_{l+k}-\bar{x}\right)}{n-k-1},\ k=\left|i-j\right|$ (5-12c)

(4) 自相關係數：$r_{ij}=\frac{\sum_{l=1}^{n-k}\left(x_l-\bar{x}\right)\left(x_{l+k}-\bar{x}\right)}{\sqrt{\left[\sum_{l=1}^{n-k}\left(x_l-\bar{x}\right)^2\right]\cdot\left[\sum_{m=k+1}^{n}\left(x_m-\bar{x}\right)^2\right]}},\ k=\left|i-j\right|$ (5-12d)

因為時間序列資料間彼此相依，所以，以上公式估計的動差，除了(5-12a)外，其他三個都是偏估的。

(一) 平均值的期望值分析

$$\boldsymbol{E}\left[\bar{X}\right]=\boldsymbol{E}\left[\frac{1}{n}\sum_{i=1}^{n}X_i\right]=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{E}\left[X_i\right]=\frac{1}{n}\cdot\sum_{i=1}^{n}\mu=\mu \tag{5-13a}$$

(二) 平均值的變異數分析

$$\boldsymbol{E}\left[(\bar{X}-\mu)^2\right]=\boldsymbol{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu\right)\right)^2\right]=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\boldsymbol{E}\left[\left(X_i-\mu\right)\left(X_j-\mu\right)\right]$$

$$=\frac{1}{n^2}\begin{bmatrix}\sigma^2+\rho_1\sigma^2+\cdots+\rho_{n-1}\sigma^2\\+\rho_1\sigma^2+\sigma^2+\cdots+\rho_{n-2}\sigma^2\\\cdots\\+\rho_{n-1}\sigma^2+\rho_{n-2}\sigma^2+\cdots+\sigma^2\end{bmatrix}=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}C_{ij} \tag{5-13b}$$

其中C_{ij}是一個只與時間差有關的函數。若資料間為彼此不相關，相關係數$\rho_{ij}\approx 0$，則$\boldsymbol{E}\left[(\bar{X}-\mu)^2\right]=\frac{n\sigma^2}{n^2}=\frac{\sigma^2}{n}$；若資料間為完全相關，相關係數$\rho_{ij}\approx 1$，則 $\boldsymbol{E}\left[(\bar{X}-\mu)^2\right]=\sigma^2$。

(三) 變異數的偏估分析

$$\boldsymbol{E}\left[S^2\right]=\frac{1}{n-1}\sum_{i=1}^{n}\boldsymbol{E}\left[\left(X_i-\bar{X}\right)^2\right]$$

$$=\frac{1}{n-1}\sum_{i=1}^{n}\boldsymbol{E}\left[\left\{\left(X_i-\mu\right)-\frac{1}{n}\sum_{j=1}^{n}\left(X_j-\mu\right)\right\}^2\right]$$

$$
=\frac{1}{n-1}\sum_{i=1}^{n}\left\{\begin{array}{l}\boldsymbol{E}\left[\left(X_i-\mu\right)^2\right]-\frac{2}{n}\sum_{j=1}^{n}\boldsymbol{E}\left[\left(X_i-\mu\right)\left(X_j-\mu\right)\right]\\+\frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n}\boldsymbol{E}\left[\left(X_j-\mu\right)\left(X_k-\mu\right)\right]\end{array}\right\} \tag{5-14}
$$

$$
=\frac{n\sigma^2}{n-1}-\frac{1}{n\left(n-1\right)}\sum_{i=1}^{n}\sum_{j=1}^{n}C_{ij}
$$

若資料間為彼此不相關，則 $\boldsymbol{E}\left[S^2\right]=\frac{n\sigma^2}{n-1}-\frac{n\sigma^2}{n\left(n-1\right)}=\sigma^2$ ，期望的偏估誤差值為 0。若資料為完全相關，則 $\boldsymbol{E}\left[S^2\right]=\frac{n\sigma^2}{n-1}-\frac{n^2\sigma^2}{n\left(n-1\right)}=0$ ，期望值的偏估誤差為 $-\sigma^2$ 。

(四) 協變異數的偏估分析

$$
\boldsymbol{E}\left[\hat{\boldsymbol{cov}}(X_t, X_{t+k})\right]=\boldsymbol{E}\left[\frac{1}{n-k-1}\sum_{i=1}^{n-k}\left(X_i-\bar{X}\right)\left(X_{i+k}-\bar{X}\right)\right]
$$

$$
=\frac{1}{n-k-1}\sum_{i=1}^{n-k}\boldsymbol{E}\left[\left\{\left(X_i-\mu\right)-\frac{1}{n}\left(\sum_{j=1}^{n}\left(X_j-\mu\right)\right)\right\}\left\{\left(X_{i+k}-\mu\right)-\frac{1}{n}\left(\sum_{l=1}^{n}\left(X_l-\mu\right)\right)\right\}\right]
$$

$$
=\frac{1}{n-k-1}\sum_{i=1}^{n-k}\left\{\begin{array}{l}\boldsymbol{E}\left[\left(X_i-\mu\right)\left(X_{i+k}-\mu\right)\right]+\frac{1}{n^2}\sum_{j=1}^{n}\sum_{l=1}^{n}\boldsymbol{E}\left[\left(X_j-\mu\right)\left(X_l-\mu\right)\right]\\-\frac{1}{n}\sum_{l=1}^{n}\boldsymbol{E}\left[\left(X_i-\mu\right)\left(X_l-\mu\right)\right]-\frac{1}{n}\sum_{j=1}^{n}\boldsymbol{E}\left[\left(X_{i+k}-\mu\right)\left(X_j-\mu\right)\right]\end{array}\right\}
$$

$$
=\frac{n-k}{n-k-1}\left\{C_{i,i+k}+\frac{1}{n^2}\sum_{j=1}^{n}\sum_{l=1}^{n}C_{jl}-\frac{1}{n(n-k)}\sum_{i=1}^{n-k}\sum_{j=1}^{n}\left[C_{ij}+C_{i+k,j}\right]\right\} \tag{5-15}
$$

$\hat{\boldsymbol{cov}}\left(X_i\ ,X_{i+k}\right)$的理論值為 $C_{i,i+k}$ ，即以上推導結果的第一項，所以後三項的總和，便是偏估量。利用以下例子可以說明偏估的情形。

例 5-2 假設時間序列為一階馬可夫鏈，並且一個時間差的實際相關係數 (lag-one correlation coefficient) ρ_1=0.8。(1)若樣本資料個數為 10，計算協變異數偏估百分比（以理論值為分母）；(2)若樣本資料個數為 100，計算一個時間差協變異數的偏估百分比。

解： 因為是一階馬可夫鏈，所以 $\rho_n=\rho_1^n$ ，將此事實代入(5-15)式協變異數的偏估量，分項計算偏估量如下：

偏估第一項：

$$\frac{1}{n^2}\sum_{j=1}^{n}\sum_{l=1}^{n}C_{jl}=\frac{\sigma^2}{n^2}\left[n+2(n-1)\rho_1+2(n-2)\rho_1^2+\cdots+2\rho_1^{n-1}\right]$$

令 $\left[(n-1)\rho_1+(n-2)\rho_1^2+\cdots+\rho_1^{n-1}\right]=\mathrm{A}$ ，則

$$\begin{aligned}
&\mathrm{A}-\rho_1\mathrm{A}\\
&=(n-1)\rho_1+(n-2)\rho_1^2+\cdots+\rho_1^{n-1}-\left[(n-1)\rho_1^2+(n-2)\rho_1^3+\cdots+\rho_1^n\right]\\
&=n\rho_1-\left[\rho_1+\rho_1^2+\cdots+\rho_1^n\right]\\
&=n\rho_1-\frac{\rho_1\left(1-\rho_1^n\right)}{1-\rho_1}
\end{aligned}$$

得到 $\left[(n-1)\rho_1+(n-2)\rho_1^2+\cdots+\rho_1^{n-1}\right]=\dfrac{n\rho_1}{1-\rho_1}-\dfrac{\rho_1\left(1-\rho_1^n\right)}{\left(1-\rho_1\right)^2}$

將解出的 A，代回原方程式可以得到：

$$\frac{1}{n^2}\sum_{j=1}^{n}\sum_{l=1}^{n}C_{jl}=\frac{\sigma^2}{n^2}\left[n+\frac{2n\rho_1}{1-\rho_1}-\frac{2\rho_1\left(1-\rho_1^n\right)}{\left(1-\rho_1\right)^2}\right]$$

另外兩個偏估量

$$\begin{aligned}
&\frac{1}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=1}^{n}\left[C_{ij}+C_{i+1,j}\right]=\frac{1}{n(n-1)}\left[2\sum_{i=1}^{n}\sum_{j=1}^{n}C_{ij}-\sum_{j=1}^{n}\left(C_{nj}+C_{1j}\right)\right]\\
&=\frac{2\sigma^2}{n(n-1)}\left[n+\frac{2n\rho_1}{1-\rho_1}-\frac{2\rho_1\left(1-\rho_1^n\right)}{\left(1-\rho_1\right)^2}-\frac{1-\rho_1^n}{1-\rho_1}\right]
\end{aligned}$$

總偏估量 D 為：

$$\begin{aligned}
D&=\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}C_{ij}\ -\frac{1}{n(n-1)}\sum_{i=1}^{n-1}\sum_{j=1}^{n}\left[C_{ij}+C_{i+1,j}\right]\\
&=\left(\frac{1}{n^2}-\frac{2}{n(n-1)}\right)\sum_{j=1}^{n}\sum_{l=1}^{n}C_{ij}+\frac{1}{n(n-1)}\sum_{j=1}^{n}\left(C_{nj}+C_{1j}\right)\\
&=-\left(\frac{(n+1)\sigma^2}{n^2(n-1)}\right)\left[n+\frac{2n\rho_1}{1-\rho_1}-\frac{2\rho_1\left(1-\rho_1^n\right)}{\left(1-\rho_1\right)^2}\right]+\frac{2\sigma^2}{n(n-1)}\frac{1-\rho_1^n}{1-\rho_1}
\end{aligned}$$

$\hat{\mathrm{cov}}\left(X_i, X_{i+1}\right)$的理論值為 $\rho_1\cdot\sigma^2=0.8\sigma^2$ 。

a. 將 n=10， ρ_1=0.8 代入上式，計算結果 D=-0.564σ^2。理論值為 0.8σ^2，估計值的期望值為 0.8σ^2-0.564σ^2=0.236σ^2 ，偏估量的百分比為低估 (0.564/0.8)×100%=70.5%。

b. 將 n=100，ρ_1=0.8 代入上式，計算結果 D=-0.0867σ^2。理論值為 0.8σ^2，

估計值的期望值為 $0.8\sigma^2$-$0.0867\sigma^2$=$0.713\sigma^2$，偏估量的百分比為低估 (0.0867/0.8)×100%=10.8%。

若時間序列資料為完全不相關，則當稽延時間 k=0，相關係數 $\rho_0 \equiv 1$，對所有 k>0，$r_k \approx 0$ 、 $\mathrm{E}[r_k]=0$。在此狀況下 r_k 樣本估計值的變異數為 $\mathrm{var}(\hat{r}_k) \approx \dfrac{1}{n}$，並且 $\hat{r}_k$ 趨近於常態分佈；若時間序列在 $k \le q$ 的相關係數不為零，而在 k>q 時，$r_k = 0$，則當 k>q 時，期望值 $\mathrm{E}[r_k]=0$，變異數 $\mathrm{var}(\hat{r}_k) = \dfrac{1}{n}\left(1+2\sum_{i=1}^{q}\dfrac{r_i^2}{n}\right)$。因此對於一個時間序列檢測在時間上是否相關，可以使用以上關係進行分析。

例 5-3 一組時間序列資料，資料個數為 100，由資料求出自相關係數 $\hat{r}_k$，k=1, 2, …, 10 分別為-0.08，0.02，-0.15，-0.14，0.14，0.10，-0.13，-0.15，0.01， 0.15 和 0.09。以 5%顯著水準檢測此一時間序列是否為彼此不相關。

解：虛無假設 H_0：假設此一時間序列為不相關。常態分佈 5%顯著水準為 z=1.96，若

$$|r_k| > 1.96[\mathrm{var}(r_k)]^{1/2} \Rightarrow \text{不接受虛無假設}$$

$$|r_k| < 1.96[\mathrm{var}(r_k)]^{1/2} \Rightarrow \text{接受虛無假設}$$

若序列彼此為獨立，則 $\mathrm{E}(r_k)=0$，且 $\mathrm{var}(r_k)=1/n$，所以推翻虛無假設的門檻值為 $1.96\times(1/n)^{\frac{1}{2}}=0.196$。由於對 k=1,2,…10，$|\hat{r}_k|$ 均小於 0.196，表示資料較可能是不相關的時間序列，因此接受虛無假設。

計算多站(多變數)的定常性時間序列參數，假設 $X(t)$ 及 $Y(t)$ 分別為站 i 及 j 站隨機變數，其時間序列的實現值分別為 $x_1, x_2, \cdots, x_n$ 及 $y_1, y_2, \cdots, y_n$，則兩站間 lag-k 的互協方差(cross-variance)及相關係數分別估計如下：

(1) 互協方差：
$$C_{xy}^k = \frac{1}{n-k-1}\sum_{i=1}^{n-k}(x_i-\bar{x})(y_{i+k}-\bar{x}) \qquad \text{(5-16a)}$$

(2) Lag-k 互相關係數：
$$r_{xy}^k = \frac{\sum_{l=1}^{n-k}(x_l-\bar{x})(y_{l+k}-\bar{y})}{\sqrt{\left[\sum_{l=1}^{n-k}(x_l-\bar{x})^2\right]\cdot\left[\sum_{m=k+1}^{n}(y_m-\bar{y})^2\right]}} \qquad \text{(5-16b)}$$

5.4 自迴歸模式（Auto-Regressive model or AR model）

分析定常性的時間序列時，常用「移動平均」或「自迴歸」式來表示。自迴歸模式即馬可夫鏈模式，若定常性隨機變數 $X(t)$在時間 t 的實現值 x_t 可以其過去值的線性組合表示為：

$$\begin{aligned} x_t &= \mu + \phi_1(x_{t-1}-\mu) + \phi_2(x_{t-2}-\mu) + \cdots + a_t \\ &= \mu + \sum_{i=1}^{\infty}\phi_i(x_{t-i}-\mu) + a_t \end{aligned} \tag{5-17}$$

則稱其為「自迴歸表現式」(auto-regression representation)，或稱為 AR(∞)模式。所謂「自迴歸」的意思，是表示 x_t 可以由利用「自己」過去的變數值作為影響變數，迴歸估計之。

若令 $z_t = x_t - \mu$，則上式可改寫成：

$$z_t = a_t + \sum_{i=1}^{\infty}\phi_i z_{t-i} \tag{5-18}$$

其中，ϕ_i 為 z_{t-i} 的權重係數，且 $1+\sum_{i=1}^{\infty}|\phi_i| < \infty$，$\sum_{i=1}^{\infty}\phi_i z_{t-i}$ 是自迴歸的部份；a_t 為平均值為 0、變異數為 σ_a^2、與 x_t 及 x_{t-i} 均不相關的常態分布噪音項。Box 和 Jenkins 稱呼滿足 $1+\sum_{i=1}^{\infty}|\phi_i| < \infty$ 條件的自迴歸過程為「可反轉的」(invertible)。自迴歸模式常使用在時間序列的預報上，若自迴歸模式不可反轉，則其在預報上將無意義。

若自迴歸過程中只有前 p 項的自相關性顯著，且當 $k > p$ 時 $\phi_k = 0$，則 x_t 的 p 階自迴歸模式（p 階馬可夫鏈模式），AR(p)，可以寫成：

$$z_t = \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} + a_t \tag{5-19}$$

低階自迴歸模式，如 AR(1)模式（又稱為一階馬可夫鏈），常被用來分析水文時間序列，其形式如下：

$$z_t = \phi_1 z_{t-1} + a_t \tag{5-20}$$

§ 參數估計

由觀測值估計時間序列模式參數的方法有三種：動差法（method of moment）；最小均方差估計（least-square estimation）；以及最大近似法（maximum likelihood method）。最大近似法雖然理論上最正確，但實際應

用因為有相當的困難度故最少使用，此處只介紹前二種參數估計法方法。

(一) 動差法

單變數自迴歸模式 AR(p)，$z_t = \phi_1 z_{t-1} + \cdots + \phi_p z_{t-p} + a_t$，$a_t \sim N(0, \sigma_a^2)$，模式中共有 p+2 個待定參數，包括：$\mu, \phi_1, \cdots, \ \phi_p$ 及 σ_a^2。參數和平均值、噪音項變異數及自相關係數函數（auto-correlation coefficient function, ACF）的關係如下：

$$\mathrm{E}[X_t] = \mu \tag{5-21}$$

將 AR(p)自迴歸模式乘以 z_{t-k} 取期望值，則

$$\boldsymbol{E}\left[z_t z_{t-k}\right] = \sum_{i=1}^{p} \varphi_i \, \boldsymbol{E}\left[z_{t-i} z_{t-k}\right] + \boldsymbol{E}\left[a_t z_{t-k}\right] \tag{5-22}$$

當 k=0 時，

$$C_0 = \phi_1 C_1 + \phi_2 C_2 + \cdots + \phi_p C_p + \sigma_a^2 \tag{5-23a}$$

其中，$C_0 = \sigma_z^2$；$C_m = \mathrm{E}\left[z_t z_{t-m}\right]$。重新整理上式，得到：

$$\sigma_a^2 = \left(1 - \sum_{i=1}^{p} \phi_i \rho_i\right) \sigma_z^2 \tag{5-23b}$$

當 k=1,2,3,...,p 時：

$$C_k = \phi_1 C_{k-1} + \phi_2 C_{k-2} + \cdots + \phi_p C_{k-p} \quad \text{for} \quad k = 1,2,3,\ldots,p \tag{5-24a}$$

將上式除以 σ_z^2，得到：

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \cdots + \phi_p \rho_{k-p} \quad \text{for} \quad k = 1,2,3,\ldots,p \tag{5-24b}$$

上述方程式又稱為 Yule-Walker 方程式。由上述方程式可以得到 AR(1)的變異數及自相關係數的表示式如下：

$$\begin{array}{l} \rho_1 = \phi_1 \rho_0 = \phi_1 \\ \sigma_a^2 = \sigma_z^2 \left(1 - \phi_1^2\right) \end{array} \quad \Leftrightarrow \quad \begin{cases} \phi_1 = \rho_1 \\ \sigma_a^2 = \sigma_z^2 \left(1 - \rho_1^2\right) \end{cases} \tag{5-25}$$

若已知 $x_t, t = 1,2,\cdots n$ 為隨機變數 $X(t)$的觀測樣本，則動差法估計自迴歸模式參數的方式為由觀測樣本估計樣本平均值 $\bar{x}_t$、樣本變異數 s_x^2、以及樣本協變異數 $C_i, i = 1,2,\cdots, p$。計算樣本自相關係數函數 ACF 的估計值 $\hat{\rho}_i, i = 1,2,\cdots, p$，再將 ACF 代入 Yule-Walker 方程式中求得參數的估計值為：

$$
\begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} = \begin{bmatrix} 1 & \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_{p-2} & \hat{\rho}_{p-1} \\ \hat{\rho}_1 & 1 & \hat{\rho}_1 & \cdots & \hat{\rho}_{p-3} & \hat{\rho}_{p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \hat{\rho}_{p-1} & \hat{\rho}_{p-2} & \hat{\rho}_{p-3} & \cdots & \hat{\rho}_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \vdots \\ \hat{\rho}_p \end{bmatrix} \tag{5-26}
$$

以上的參數估計式又稱為 Yule-Walker 估計式。估計噪音項的變異數為：

$$
s_a^2 = s_z^2 \left\{ 1 - \sum_{i=1}^{p} \sum_{j=1}^{p} \hat{\phi}_i \hat{\phi}_j \rho_{|i-j|} \right\} \tag{5-27}
$$

動差法中，由樣本直接計算得到的協變異數 $C_i, i = 1,2,\cdots,p$ 可能為偏估的，因此以動差法求得的參數亦為偏估的，和最大近似法比較，此方法的精確度較低。

(二) 最小均方差估計法

最小均方差估計法是以迴歸方式推估參數值，為最常用的參數估計法。若 $x_t, t = 1,2,\cdots n$ 為隨機變數 $X(t)$的觀測樣本，則單變數自迴歸模式 AR(p)的參數估計步驟為令：

$$
\begin{aligned}
z_{p+1} &= \hat{\phi}_1 z_p + \hat{\phi}_2 z_{p-1} + \cdots + \hat{\phi}_p z_1 + a_{p+1} \\
z_{p+2} &= \hat{\phi}_1 z_{p+1} + \hat{\phi}_2 z_p + \cdots + \hat{\phi}_p z_2 + a_{p+2} \\
&\vdots \\
z_n &= \hat{\phi}_1 z_{n-1} + \hat{\phi}_2 z_{n-2} + \cdots + \hat{\phi}_p z_{n-p} + a_n
\end{aligned} \tag{5-28a}
$$

將以上第 p+1 項至第 n 項的線性方程式改為矩陣表示式：

$$
\begin{bmatrix} z_{p+1} \\ z_{p+2} \\ \vdots \\ z_n \end{bmatrix} = \begin{bmatrix} z_p & z_{p-1} & \cdots & z_1 \\ z_{p+1} & z_p & \cdots & z_2 \\ \vdots & \vdots & & \vdots \\ z_{n-1} & z_{n-2} & \cdots & z_{n-p} \end{bmatrix} \begin{bmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \vdots \\ \hat{\phi}_p \end{bmatrix} + \begin{bmatrix} a_{p+1} \\ a_{p+2} \\ \vdots \\ a_n \end{bmatrix} \text{或 } \mathbf{z} = \mathbf{\Omega}\hat{\mathbf{\Phi}} + \mathbf{a} \tag{5-28b}
$$

其中 n-p 必須大於 p。以迴歸方程式思考，則迴歸方程式所不能解釋的估計誤差向量為：$\mathbf{a} = \mathbf{z} - \mathbf{\Omega}\hat{\mathbf{\Phi}}$，估計誤差變異數為：

$$
\begin{aligned}
L(\hat{\mathbf{\Phi}}) &= (\mathbf{z} - \mathbf{\Omega}\hat{\mathbf{\Phi}})^{\mathrm{T}} (\mathbf{z} - \mathbf{\Omega}\hat{\mathbf{\Phi}}) \\
&= \mathbf{z}^{\mathbf{T}}\mathbf{z} - \hat{\mathbf{\Phi}}^{\mathbf{T}}\mathbf{\Omega}^{\mathbf{T}}\mathbf{z} - \mathbf{z}\mathbf{\Omega}\hat{\mathbf{\Phi}} + \hat{\mathbf{\Phi}}^{\mathbf{T}}\mathbf{\Omega}^{\mathbf{T}}\mathbf{\Omega}\hat{\mathbf{\Phi}}
\end{aligned} \tag{5-29}
$$

估計誤差變異數矩陣對於參數微分等於零時估計誤差為最小，得到

$$\mathbf{\Omega}^{\mathrm{T}}\mathbf{\Omega}\hat{\mathbf{\Phi}} = \mathbf{\Omega}^{\mathrm{T}}\mathbf{z} \iff \hat{\mathbf{\Phi}} = (\mathbf{\Omega}^{\mathrm{T}}\mathbf{\Omega})^{-1}\mathbf{\Omega}^{\mathrm{T}}\mathbf{z} \tag{5-30a}$$

上式中，

$$\begin{aligned}
\mathbf{\Omega}^{\mathrm{T}}\mathbf{\Omega} &= \begin{bmatrix} z_p & z_{p-1} & \cdots & z_1 \\ z_{p+1} & z_p & \cdots & z_2 \\ \vdots & \vdots & \ddots & \vdots \\ z_{n-1} & z_{n-2} & \cdots & z_{n-p} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} z_p & z_{p-1} & \cdots & z_1 \\ z_{p+1} & z_p & \cdots & z_2 \\ \vdots & \vdots & \ddots & \vdots \\ z_{n-1} & z_{n-2} & \cdots & z_{n-p} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=p}^{n-1} z_i^2 & \sum_{i=p}^{n-1} z_i z_{i-1} & \cdots & \sum_{i=p}^{n-1} z_i z_{i-p+1} \\ \sum_{i=p-1}^{n-2} z_i z_{i+1} & \sum_{i=p-1}^{n-2} z_i^2 & \cdots & \sum_{i=p-1}^{n-2} z_i z_{i-p+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{n-p} z_i z_{i+p-1} & \sum_{i=1}^{n-p} z_i z_{i+p-2} & \cdots & \sum_{i=1}^{n-p} z_i z_{i+p-1} \end{bmatrix} \\
&\approx (n-p)\sigma_x^2 \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \cdots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \cdots & 1 \end{bmatrix}
\end{aligned} \tag{5-30b}$$

$$\begin{aligned}
\mathbf{\Omega}^{\mathrm{T}}\mathbf{z} &= \begin{bmatrix} z_p & z_{p-1} & \cdots & z_1 \\ z_{p+1} & z_p & \cdots & z_2 \\ \vdots & \vdots & \ddots & \vdots \\ z_{n-1} & z_{n-2} & \cdots & z_{n-p} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} z_{p+1} \\ z_{p+2} \\ \vdots \\ z_n \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=p}^{n-1} z_i z_{i+1} \\ \sum_{i=p}^{n-1} z_{i-1} z_{i+1} \\ \vdots \\ \sum_{i=p}^{n-1} z_{i-p+1} z_{i+1} \end{bmatrix} \approx (n-p)\sigma_x^2 \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix}
\end{aligned} \tag{5-30c}$$

和動差法的結果比較可知，對於 AR(p)模式而言，以最小均方差估計法所求得的參數會和動差法得到的結果相同。

例 5-4 單變數時間序列的二階馬可夫鏈方程式如下：

$$X(t+1) = \phi_1 \cdot X(t) + \phi_2 \cdot X(t-1) + g \cdot W(t+1) + C$$

其中，ϕ_1、ϕ_2、g 和 C 為未知常數係數，$W(t+1)$ 為與 $X(t)$ 無關的

標準常態噪音項，請導出二階馬可夫鏈方程式中的係數。

解：導出係數的條件為廣義定常性的各項條件。

(1) 平均值不變條件 $\boldsymbol{E}\left[\hat{X}(t+1)\right]=\boldsymbol{E}\left[X(t+1)\right]$：

$$\mathrm{E}[X(t+1)]=\mathrm{E}[X(t)]=\mathrm{E}[X(t-1)]=m_x$$

$$\therefore\ m_x=\phi_1 m_x+\phi_2 m_x+C$$

$$\Leftrightarrow\ C=(1-\phi_1-\phi_2)m_x$$

將 C 代入二階馬可夫鏈方程式中得到

$$X(t+1)-m_x=\phi_1[X(t)-m_x]+\phi_2[X(t-1)-m_x]+gW(t+1)$$

令移除平均值的隨機變數 $X(t+1)-m_x=\mathrm{Z}_{t+1}$，其餘類推，則上式可以改寫為：

$$Z_{t+1}=\phi_1 Z_t+\phi_2 Z_{t-1}+gW_{t+1} \qquad \text{(A-1)}$$

(2) 變異數不變條件 $\boldsymbol{var}\left[\hat{Z}_{t+1}\right]=\boldsymbol{var}\left[Z_{t+1}\right]=\sigma_x^2$

$$\mathrm{var}[Z_{t+1}]=\mathrm{var}[Z_t]=\mathrm{var}[Z_{t-1}]=\sigma_x^2$$

$$\begin{aligned}\mathrm{var}[Z_{t+1}]&=\mathrm{E}\left[(\phi_1 Z_t+\phi_2 Z_{t-1}+gW_{t+1})^2\right]\\&=\phi_1^2\sigma_x^2+\phi_2^2\sigma_x^2+2\phi_1\phi_2\rho_1\sigma_x^2+g^2\,\mathrm{var}[W_{t+1}]\\&=\phi_1^2\sigma_x^2+\phi_2^2\sigma_x^2+2\phi_1\phi_2\rho_1\sigma_x^2+g^2\end{aligned}$$

$$\therefore\ g^2=\sigma_x^2\left(1-\phi_1^2-\phi_2^2-2\phi_1\phi_2\rho_1\right) \qquad \text{(A-2)}$$

(3) 將(A-1)式左右兩邊各乘以 Z_t 再取期望值，得到：

$$\rho_1\sigma_x^2=\phi_1\sigma_x^2+\phi_2\rho_1\sigma_x^2$$

$$\therefore\ \rho_1=\phi_1+\phi_2\rho_1 \qquad \text{(A-3)}$$

(4) 將(A-1)式左右兩邊各乘以 Z_{t-1} 再取期望值，得到

$$\rho_2\sigma_x^2=\phi_1\rho_1\sigma_x^2+\phi_2\sigma_x^2$$

$$\therefore\ \rho_2=\phi_1\rho_1+\phi_2 \qquad \text{(A-4)}$$

由(A-3)及(A-4)式聯立求解 ϕ_1 與 ϕ_2，可得到：

$$\phi_1=\frac{\rho_1(1-\rho_2)}{1-\rho_1^2}\text{，}\ \phi_2=\frac{\rho_2-\rho_1^2}{1-\rho_1^2} \qquad \text{(A-5)}$$

將(A-5)式的結果代入(A-1)式，得到：

$$g = \frac{\sigma_x \sqrt{(1-\rho_2)(1+\rho_2-2\rho_1^2)}}{\sqrt{1-\rho_1^2}} \tag{A-6}$$

將(A-5)及(A-6)式代入(A-1)式中即可得到二階馬可夫鏈方程式。另外，以上參數也可以使用 Yule-Walker 方程式求得：

$$\begin{aligned}\begin{bmatrix}\phi_1\\ \phi_2\end{bmatrix} &= \begin{bmatrix}1 & \rho_1\\ \rho_1 & 1\end{bmatrix}^{-1}\begin{bmatrix}\rho_1\\ \rho_2\end{bmatrix} = \frac{1}{1-\rho_1^2}\begin{bmatrix}1 & -\rho_1\\ -\rho_1 & 1\end{bmatrix}\begin{bmatrix}\rho_1\\ \rho_2\end{bmatrix}\\ &= \frac{1}{1-\rho_1^2}\begin{bmatrix}\rho_1(1-\rho_2)\\ \rho_2-\rho_1^2\end{bmatrix}\end{aligned} \tag{A-7}$$

5.5 移動平均模式（Moving Average model or MA model）

分析定常性的時間序列時，除了「自迴歸」模式外，亦常使用「移動平均」模式來表示。若定常性隨機變數 $X(t)$在時間 t 的實現值為 x_t，則 x_t 的幾種可能的移動平均表示式為：

$$y_t = (x_t + x_{t-1} + \cdots + x_{t-k})/(k+1) \tag{5-31a}$$

$$y_t = (x_{t-k} + x_{t-k+1} + \cdots + x_{t+k-1} + x_{t+k})/(2k+1) \quad \text{（centered）} \tag{5-31b}$$

以上的一次移動平均式，若對移動平均的結果再做一次移動平均，稱為二次移動平均（double MA）：

$$z_t = (y_t + y_{t-1} + \cdots + y_{t-k})/(k+1) \tag{5-32a}$$

$$z_t = (y_{t-k} + y_{t-k+1} + \cdots + y_{t+k-1} + y_{t+k})/(2k+1) \tag{5-32b}$$

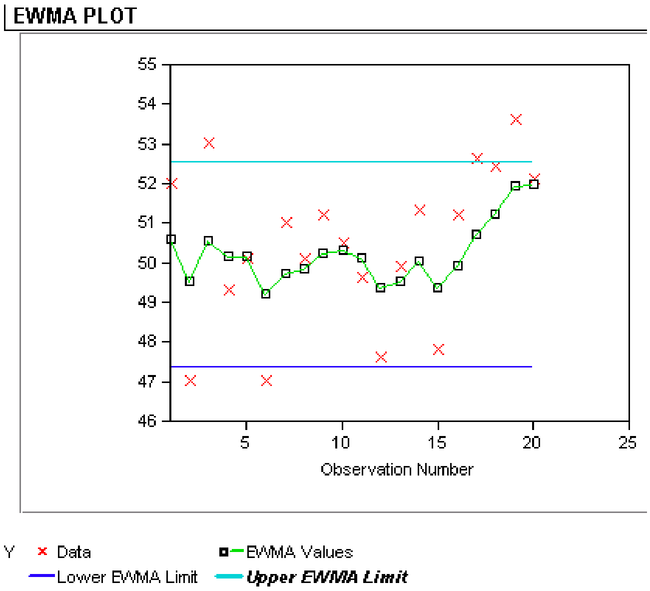
移動平均法經常是用來去除高頻項或噪音項，並且利用一次移動平均或二次移動平均的結果，以外延的方式進行預報。但是當趨勢顯著時，以上模式無論是一次或二次移動平均都無法適當表現時間序列的趨勢。

除了以上等權重的移動平均法外，另一種常見的移動平均法是指數權重移動平均（Exponential or Exponential Weight MA, EWMA），權重係數為指數遞減形式。對於時間序列 x_t，一次指數權重移動平均平滑化（single exponential smoothing）的作法如下：

$s_1 = x_1$ (known value)

$$s_t = \alpha x_t + (1-\alpha)s_{t-1} \quad \text{for} \quad t \geq 2 \tag{5-33}$$

α 稱為平滑常數（smoothing constant），$0 \leq \alpha \leq 1$。一次指數權重移動平均對於資料的處理效果示意如下圖。

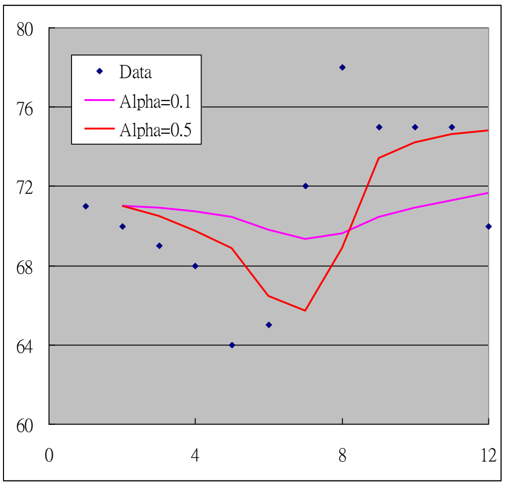


根據以上指數權重移動平均的作法，可以演繹得到，對於 $t \geq 2$：

$$\begin{aligned}
s_t &= \alpha x_t + (1-\alpha)s_{t-1} \\
&= \alpha x_t + (1-\alpha)\alpha x_{t-1} + (1-\alpha)^2 s_{t-2} \\
&= \alpha x_t + (1-\alpha)\alpha x_{t-1} + (1-\alpha)^2 \alpha x_{t-2} + (1-\alpha)^3 s_{t-3} \\
&= \alpha \sum_{i=0}^{t-2} (1-\alpha)^i x_{t-i} + (1-\alpha)^{t-1} s_1 \\
&= \alpha \sum_{i=0}^{t-2} (1-\alpha)^i x_{t-i} + (1-\alpha)^{t-1} x_1 \\
&= \alpha \sum_{i=0}^{t-1} (1-\alpha)^i x_{t-i} + (1-\alpha)^t x_1
\end{aligned}$$

可以看出是以指數遞減的方式權重時間序列 x_t，若 α 值愈趨近於 0，則初始值的影響時間愈長，s_1 型式的選擇也愈重要。

t	x	$\alpha=0.1$	ε^2	$\alpha=0.5$	ε^2
1	71				
2	70	71.00	1.00	71.00	1.00
3	69	70.90	3.61	70.50	2.25
4	68	70.71	7.34	69.75	3.06
5	64	70.44	41.46	68.88	23.77
6	65	69.80	22.99	66.44	2.07
7	72	69.32	7.21	65.72	39.45
8	78	69.58	70.83	68.86	83.55
9	75	70.43	20.92	73.43	2.47
10	75	70.88	16.95	74.21	0.62
11	75	71.29	13.73	74.61	0.15
12	70	71.67	2.77	74.80	23.08
Σ			208.82		181.46



上表中的時間序列 $x(t)$，以 $\alpha=0.1$ 一次指數權重移動平均平滑處理的結果，估計誤差平方和為 208.82；以 $\alpha=0.5$ 處理的結果，估計誤差變異數為 181.46。最佳化的方法，可以採用對切法（bisection），由 $\alpha\in(0,1)$ 逐漸逼近；或是採用統計軟體中非線性最佳化的 Marquardt procedure 求解 α 。

利用一次指數權重移動平均進行一個 Δt 時距預報（forecasting with single exponential smoothing）的基本方程式為：

$$s_{t+1}=\alpha\cdot x_t+(1-\alpha)\cdot s_t=s_t+\alpha(x_t-s_t) \tag{5-34}$$

由於前一次預報為 s_t ，預報的誤差為 $\varepsilon_t=s_t-x_t$ ，因此以上方程式可以改寫為：

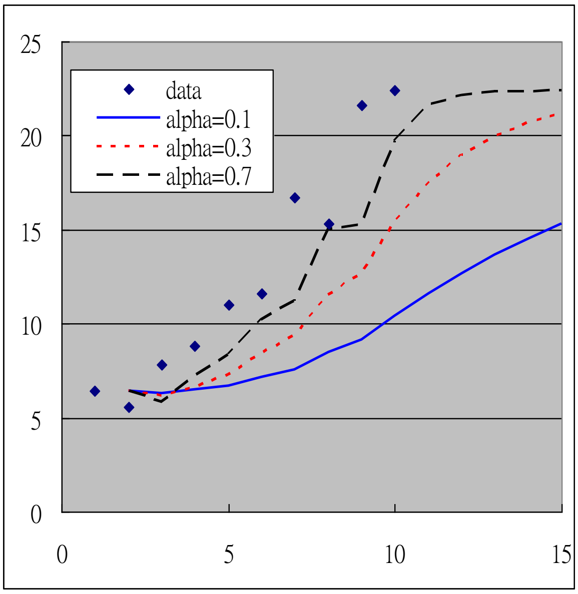
$$s_{t+1}=s_t-\alpha\cdot\varepsilon_t \tag{5-35}$$

如果最後一次觀測時間為 t ，變數的觀測值為 x_t ，欲利用一次指數權重移動平均進行兩個 Δt 時距以上的預報時，預報的自生(bootstrap)方程式為：

$$s_{t+k}=\alpha\cdot x_t+(1-\alpha)\cdot s_{t+k-1}\quad \text{for}\quad k\ge 1 \tag{5-36}$$

利用一次指數權重移動平均法預報，無法正確表現具有趨勢的時間序列，下表數據以 $\alpha=0.1$ 、 $\alpha=0.3$ 和 $\alpha=0.7$ 作平滑和自生預報為例說明。由圖可以了解，一次指數權重移動平均平滑化會低估趨勢的變化率，α 值愈大，偏估情形愈輕微，α 值愈小，則偏估情形愈嚴重。對於自生預報而言，則是 α 值愈大，預報愈接近持續預報；α 值愈小，則前面資料的影響會繼續存在。

t	x	$\alpha=0.1$	$\alpha=0.3$	$\alpha=0.7$
1	6.4			
2	5.6	6.4	6.4	6.4
3	7.8	6.3	6.2	5.8
4	8.8	6.5	6.7	7.2
5	11	6.7	7.3	8.3
6	11.6	7.1	8.4	10.2
7	16.7	7.6	9.4	11.2
8	15.3	8.5	11.6	15.0
9	21.6	9.2	12.7	15.2
10	22.4	10.4	15.4	19.7
11		11.6	17.5	21.6
12		12.7	19.0	22.2
13		13.7	20.0	22.3
14		14.5	20.7	22.4
15		15.3	21.2	22.4



對於一次指數權重移動平均無法正確表現具有趨勢的時間序列的問題，二次指數權重移動平均平滑法（double exponential smoothing）可以克服此問題，作法如下：

$$
\begin{aligned}
s_t &= \alpha x_t + (1-\alpha)\left(s_{t-1}+b_{t-1}\right) \quad \text{for } t \geq 2,\ 0 \leq \alpha \leq 1 \\
b_t &= \gamma\left(s_t - s_{t-1}\right) + (1-\gamma) b_{t-1} \quad \text{for } t \geq 2,\ 0 \leq \gamma \leq 1
\end{aligned}
\tag{5-37}
$$

s 序列的初始條件：$s_1 = x_1$

b 序列的初始條件有幾種選擇：

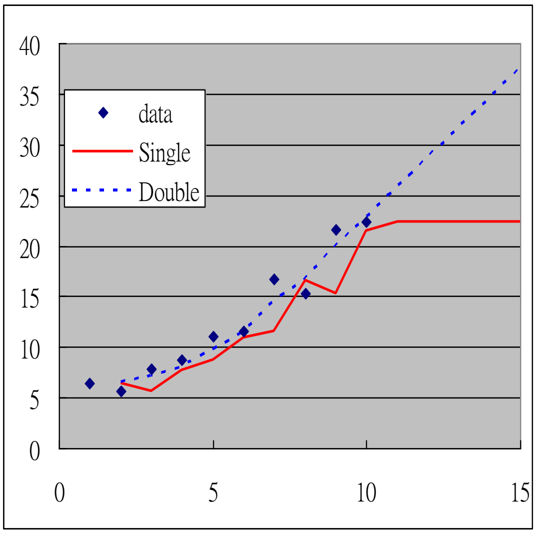
$$b_1 = x_2 - x_1 \text{、} b_1 = \tfrac{1}{2}\left(x_3 - x_1\right) \text{、} b_1 = \tfrac{1}{3}\left(x_4 - x_1\right) \text{、} b_1 = \tfrac{1}{n-1}\left(x_n - x_1\right)$$

以上方程式中的 b 參數為代表趨勢的變數，第一個方程式除了作 single exponential smoothing 之外，並且利用前一時間的趨勢 b_{t-1} 調整 s_t，這樣可以去除延遲（lag）問題；第二個方程式則是以 single exponential smoothing 邏輯建立的趨勢更新方程式。兩個方程式中的平滑化參數 α 和 γ 可以採用統計軟體中非線性最佳化的 Marquardt procedure 求解。

二次指數權重移動平均平滑法（double exponential smoothing）比較有用的是對含有趨勢的時間序列進行預報。預報方程式（forecasting with double exponential smoothing）如下：

$$F_{t+m} = s_t + m b_t \tag{5-38}$$

利用同樣一組資料比較一次指數權重移動平均預報和或二次移動平均預報的差異：



t	x	Single	Double
1	6.4	$\alpha=0.977$	$\alpha=0.3623,\gamma=1$
2	5.6	6.4	6.6
3	7.8	5.6	7.2
4	8.8	7.7	8.1
5	11	8.8	9.8
6	11.6	10.9	11.5
7	16.7	11.6	14.5
8	15.3	16.6	16.7
9	21.6	15.3	19.9
10	22.4	21.5	22.8
11		22.4	25.8
12		22.4	28.7
13		22.4	31.7
14		22.4	34.6
15		22.4	37.6

5.6 ARMA 和 ARIMA 模式

若自迴歸過程中只有前 p 項的自相關性顯著，且當 $k>p$ 時 $\phi_k=0$，則 x_t 的 p 階自迴歸模式（p 階馬可夫鏈模式），AR(p)，可以寫成：

$$\hat{Z}_t=\varphi_1 Z_{t-1}+\cdots+\varphi_p Z_{t-p}\text{，}\quad z_t=\varphi_1 z_{t-1}+\cdots+\varphi_p z_{t-p}+\varepsilon_t$$

相同的，若移動平均過程只有前 q 項的噪音項 ε_t 顯著，則 z_t 的 q 階移動平均模式，MA(q)，可以寫成：

$$z_t=\theta_1\varepsilon_{t-1}+\cdots+\theta_q\varepsilon_{t-q}+\varepsilon_t$$

一個定常性且可反轉的時間序列過程常以移動平均或自迴歸模式的形式來表示。一個高階的時間序列模式，經常會遇到參數太多的問題。因此，一個替代的方法是使用「移動平均自迴歸模式」。「移動平均自迴歸模式」是一個比自迴歸模式更具變化性的模式，可以使用較少的參數。移動平均自迴歸模式可 ARMA(p,q)表示為：

$$z_t=\varphi_1 z_{t-1}+\cdots+\varphi_p z_{t-p}+\theta_1\varepsilon_{t-1}+\cdots+\theta_q\varepsilon_{t-q}+\varepsilon_t \tag{5-39}$$

ARMA(p,q)模式中包含了 p 個自迴歸權重參數 $\varphi_1,\cdots,\varphi_p$ 以及 q 個移動平均權重參數 $\theta_1,\cdots,\theta_q$。若 q=0，則 ARMA(p,0)即為自迴歸 AR(p)模式；若 p=0，則 ARMA(0,q)相當於移動平均 MA(q)模式。以 ARMA(1,1)為例，為確保數列收斂必須要求 $|\phi_1|<1$ 以滿足定常性，$|\theta_1|<1$ 以滿足可反轉性。

ARIMA (Autoregressive Integrated Moving Average)模型為「差分整合移動平均自迴歸模型」，又稱「整合移動平均自迴歸模型」，為時間序列預測

分析方法之一。ARIMA(p, d, q)中，AR 是自迴歸，p 為自迴歸項數；MA 為移動平均，q 為移動平均項數，d 為使之成為平穩序列所做的差分次數（階數）。「差分」一詞雖未出現在 ARIMA 的英文名稱中，卻是關鍵步驟。ARIMA(p, d, q)模型是 ARMA(p, q)模型的擴展；ARIMA(p, d, q)模型可以表示為：

$$\left[1-\sum_{i=1}^{p}\varphi_i L_i\right](1-L)^d Z_t=\left[1+\sum_{j=1}^{q}\theta_j L_j\right]a_t \tag{5-40}$$

其中，L 是滯後算子(Lag operator)，例如 $L_i Z_t = Z_{t-i}$，$L_j \varepsilon_t = \varepsilon_{t-j}$。

辨識 ARMA 和 ARIMA 模型結構和估計模式參數時，可以使用 Matlab 的 Time-Series Model Identification 副程式組。

5.7 多變數時間序列模式

5-7.1 多變數自迴歸模式(VAR model)

水文時間序列資料的分析與模擬在很多時候需要用到多變數模式，例如，分析一個集水區中數個雨量站的降雨序列；記錄一個河川網絡中數個水位站的水位序列；或降雨與水位同時考慮的混合性問題。基本上，多變數時間序列模式的基本觀念與原理與單變數模式者相同，都是由平均值、變異數及協變異數不變的條件出發而推導模式參數；但在此中，數學式需為向量或矩陣形式，故又稱為向量自迴歸 (Vector Auto-Regression, VAR)模型。

考慮一個單一時間向量 $\mathbf{X}_t$，$\mathbf{X}_t=\left[x_t^{(1)} \quad x_t^{(2)} \quad \cdots \quad x_t^{(n)}\right]^{\mathbf{T}}$，向量中的元素 $x_t^{(j)}$ 表示在時間為 t（$t=1,2,\cdots$）時第 j 個觀測點的值(或稱第 j 個變數)，n 為測站或變數個數。Matalas 定義一階多變數自迴歸模式：multivariate AR (1) 為：

$$\mathbf{Z}_t=\mathbf{A}\cdot\mathbf{Z}_{t-1}+\mathbf{B}\varepsilon_t \tag{5-41}$$

其中，$\mathbf{Z}_t=\mathbf{X}_t-\mathbf{M}$，$\mathbf{M}=\left[\mu^{(1)} \quad \mu^{(2)} \quad \cdots \quad \mu^{(n)}\right]^{\mathrm{T}}$，$\mathbf{M}$ 向量中元素 $\mu^{(j)}$ 表示第 j 個測點的平均值或趨勢；參數矩陣 $\mathbf{A}$ 和 $\mathbf{B}$ 均為 $n\times n$ 的方矩陣；$\boldsymbol{\varepsilon}_t=\left[\varepsilon^{(1)} \quad \varepsilon^{(2)} \quad \cdots \quad \varepsilon^{(n)}\right]^{\mathrm{T}}$ 為與 $\mathbf{Z}_{t-1}$ 不相關的常態分佈噪音項，$\boldsymbol{\varepsilon}_t$ 中元素的平均值均為 0，且

$$\mathrm{E}\left[\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_{t\text{-}k}^{\mathrm{T}}\right]=\begin{cases}\mathbf{I} & \text{for } k=0\\ 0 & \text{for } k\neq 0\end{cases} \tag{5-42}$$

其中 $\mathbf{I}$ 為單位矩陣。如前面所推導，$\mathbf{X}_t$ 必須滿足(1)平均值、(2)變異數和(3)協變異數不變之關係，將(5-41)一階多變數自迴歸模式後乘上 $\mathbf{Z}_t^{\mathrm{T}}$ 並取期望值：

$$\begin{aligned}\mathrm{E}\left[\mathbf{Z}_t\mathbf{Z}_t^{\mathrm{T}}\right]&=\mathbf{A}\,\mathrm{E}\left[\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{\mathrm{T}}\right]\mathbf{A}^{\mathrm{T}}+\mathbf{B}\,\mathrm{E}\left[\varepsilon_t\varepsilon_t^{\mathrm{T}}\right]\mathbf{B}^{\mathrm{T}}+\mathbf{A}\,\mathrm{E}\left[\mathbf{Z}_{t-1}\varepsilon_t^{\mathrm{T}}\right]\mathbf{B}^{\mathrm{T}}+\mathbf{B}\,\mathrm{E}\left[\varepsilon_t\mathbf{Z}_{t-1}^{\mathrm{T}}\right]\mathbf{A}^{\mathrm{T}}\\&=\mathbf{A}\,\mathrm{E}\left[\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{\mathrm{T}}\right]\mathbf{A}^{\mathrm{T}}+\mathbf{B}\mathbf{B}^{\mathrm{T}}\end{aligned} \tag{5-43}$$

上式可寫成

$$\mathbf{C}_0=\mathbf{A}\mathbf{C}_0\mathbf{A}^{\mathrm{T}}+\mathbf{B}\mathbf{B}^{\mathrm{T}} \tag{5-44}$$

若(5-41)式乘上 $\mathbf{Z}_{t-1}^{\mathrm{T}}$ 取期望值

$$\begin{aligned}\mathrm{E}\left[\mathbf{Z}_t\mathbf{Z}_{t-1}^{\mathrm{T}}\right]&=\mathbf{A}\,\mathrm{E}\left[\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{\mathrm{T}}\right]+\mathbf{B}\,\mathrm{E}\left[\varepsilon_t\mathbf{Z}_{t-1}^{\mathrm{T}}\right]\\&=\mathbf{A}\,\mathrm{E}\left[\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}^{\mathrm{T}}\right]\end{aligned} \tag{5-45}$$

上式可寫成

$$\mathbf{C}_1=\mathbf{A}\mathbf{C}_0 \tag{5-46}$$

(5-44)及(5-46)式中 $\mathbf{C}_0$ 和 $\mathbf{C}_1$ 分別代表多變數時間序列 $\mathbf{X}_t$ 在同一時間(lag=0)與稽延時間為一(lag=1)的交互協變異矩陣，可由公式(5-11)以歷史資料樣本估計而得。由(5-46)式推導得 $\mathbf{A}$ 為

$$\mathbf{A}=\mathbf{C}_1\mathbf{C}_0^{-1} \tag{5-47}$$

將 A 代入(5-46)式中得到

$$\mathbf{B}\mathbf{B}^{\mathbf{T}}=\mathbf{C}_0-\mathbf{C}_1\mathbf{C}_0^{-1}\mathbf{C}_1^{\mathrm{T}}=\mathbf{C}_0-\mathbf{A}\mathbf{C}_1^{\mathrm{T}} \tag{5-48}$$

令 $\mathbf{B}\mathbf{B}^{\mathrm{T}}=\mathbf{G}$，則矩陣 $\mathbf{B}$ 可用 Cholesky decomposition method 解成一個下三角形矩陣(lower triangular matrix)乘上自己的轉置矩陣，依此得到 $\mathbf{B}$ 的元素 B_{ij} 為

$$B_{11}=\left[G_{11}\right]^{1/2} \tag{5-49a}$$

$$B_{i1}=\frac{G_{i1}}{B_{11}} \qquad \text{for } \ i=1,\cdots,n \tag{5-49b}$$

$$B_{ii} = \left[G_{ii} - \sum_{k=1}^{i-1} (B_{ki})^2 \right]^{\frac{1}{2}} \qquad \text{for} \quad i = 2, \cdots, n \qquad (5\text{-}49c)$$

$$B_{ij} = \frac{1}{B_{ii}} \left(G_{ij} - \sum_{k=1}^{i-1} B_{ki} B_{kj} \right) \qquad \text{for} \quad i = 2, \cdots, n\text{-}1,\ i < j \qquad (5\text{-}49d)$$

$$B_{ij} = 0 \qquad \text{for} \quad i > j \qquad (5\text{-}47e)$$

以上 **B** 矩陣非唯一解，假設 **D** 為一 $n \times n$ 方陣，$\mathbf{DD}^{\mathrm{T}} = \mathbf{I}$，則 **BD** 亦將為以上 **G** 之解。

例 5-5 令 $\mathbf{G} = \begin{bmatrix} 1 & 0.8 \\ 0.8 & 1 \end{bmatrix}$，解 B 矩陣

解：$\mathbf{BB}^{\mathrm{T}} = \begin{bmatrix} B_{11} & 0 \\ B_{12} & B_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix} = \begin{bmatrix} B_{11}^2 & B_{11}B_{12} \\ B_{11}B_{12} & B_{12}^2 + B_{22}^2 \end{bmatrix}$

$$\therefore B_{11} = [G_{11}]^{\frac{1}{2}} = 1$$

$$B_{12} = \frac{G_{12}}{B_{11}} = \frac{0.8}{1} = 0.8$$

$$B_{22} = \left[G_{22} - B_{12}^2 \right]^{\frac{1}{2}} = \sqrt{1 - 0.8^2} = 0.6$$

$$\therefore \mathbf{B} = \begin{bmatrix} 1 & 0 \\ 0.8 & 0.6 \end{bmatrix} \quad \Leftarrow \text{非唯一解}$$

例 5-6 美國紐約市部分的自來水是來自於位於上德拉威河(Upper Delaware River)水區的三個並聯的水庫。年流量的變異數矩陣以及一個稽延時間的交互協變異矩陣是由 50 年的流量紀錄估計而得(m^3/s)

$$\mathbf{C}_0 = \begin{bmatrix} 20.002 & 21.436 & 6.618 \\ 21.436 & 25.141 & 6.978 \\ 6.618 & 6.978 & 2.505 \end{bmatrix}, \quad \mathbf{C}_1 = \begin{bmatrix} 6.487 & 6.818 & 1.638 \\ 7.500 & 7.625 & 1.815 \\ 2.593 & 2.804 & 0.6753 \end{bmatrix}$$

年流量其他的估計值為：

站號	平均流量	標準偏差	r_1
1	20.05	4.472	0.3243
2	23.19	5.014	0.3033

3	7.12	1.583	0.2696

利用以上的資料求 **A** 和 **B** 矩陣。

解：**A** 矩陣為

$$\mathrm{A} = \mathbf{C}_1\mathbf{C}_0^{-1} = \begin{bmatrix} 1.036 & -0.150 & -1.666 \\ 1.441 & -0.308 & -2.226 \\ 0.344 & -0.019 & -0.586 \end{bmatrix}$$

估計 B 矩陣：

$$\mathbf{BB}^{\mathrm{T}} = \mathbf{C}_0 - \mathbf{AC}_0\mathbf{A}^{\mathrm{T}} = \begin{bmatrix} 17.031 & 17.813 & 5.4764 \\ 17.813 & 20.718 & 5.6070 \\ 5.4764 & 5.6070 & 2.06227 \end{bmatrix}$$

注意 $\mathbf{BB}^{\mathrm{T}}$ 必須是對稱的。

若水文時間序列資料具季節週期效應，則一階週期多變數自迴歸模式(periodic multivariate AR model)：multivariate PAR(1)為

$$\mathbf{Z}_{\nu,\tau} = \mathbf{A}_\tau \cdot \mathbf{Z}_{\nu,\tau-1} + \mathbf{B}_\tau \boldsymbol{\varepsilon}_{\nu,\tau} \tag{5-50}$$

ν及τ分別為年與季節，$\mathbf{Z}_{\nu,\tau} = \mathbf{X}_{\nu,\tau} - \boldsymbol{\mu}_\tau$，$\boldsymbol{\mu}_\tau = \left[\mu_\tau^{(1)} \quad \mu_\tau^{(2)} \quad \cdots \quad \mu_\tau^{(n)}\right]^{\mathrm{T}}$，元素 $\mu_\tau^{(j)}$ 表示第 j 個測點在第τ季的趨勢；$\boldsymbol{\varepsilon}_{\nu,\tau}$為不相關的常態分佈噪音項；$\boldsymbol{\mu}_\tau$、$\mathbf{A}_\tau$ 及 $\mathbf{B}_\tau$ 均為週期性參數向量或矩陣，$\mathbf{A}_\tau$ 及 $\mathbf{B}_\tau$ 的關係如下

$$\mathbf{A}_\tau = C_{1,\tau} C_{0,\tau-1}^{-\boldsymbol{1}} \tag{5-51}$$

$$\mathbf{B}_\tau \mathbf{B}_\tau^{\boldsymbol{T}} = C_{0,\tau} - \mathbf{A}_\tau C_{1,\tau}^{\boldsymbol{T}} \tag{5-52}$$

其中 $C_{0,\tau}$ 及 $C_{1,\tau}$ 為同一時間(time lag=0)與稽延時間為一(time lag=1)各站間季對季的交互協變異矩陣。

5-6.2 多變數移動平均自迴歸模式(multivariate ARMA model)

單一時間序列一階多變數移動平均自迴歸模式(multivariate ARMA model)：multivariate ARMA(1,1)可寫為

$$\mathbf{Z}_t = \mathbf{A} \cdot \mathbf{Z}_{t-1} + \mathbf{B}\varepsilon_t - \mathbf{C}\varepsilon_{t-1} \tag{5-53}$$

其中，參數 **C** 為一代表移動平均項的 $n\times n$ 方陣，其餘各項變數及參數的定義均與多變數 AR(1)中者相同。為說明以上的概念，將一個多變數 ARMA(p,q)寫成：

$$\mathbf{Z}_t=\sum_{i=1}^{p}\mathbf{A}_i\cdot\mathbf{Z}_{t-i}+\varepsilon_t-\sum_{i=1}^{q}\mathbf{C}_i\varepsilon_{t-i} \tag{5-54}$$

假設矩陣 $\mathbf{A}_\tau$ 及 $\mathbf{C}_\tau$ 為對角矩陣，則可將上式去耦合(decouple)而將各站的分量表示為：

$$Z_t^{(j)}=\sum_{i=1}^{p}A_i^{(j)}\cdot Z_{t-i}+\varepsilon_t^{(j)}-\sum_{i=1}^{q}C_i^{(j)}\varepsilon_{t\text{-}i}^{(j)} \tag{5-55}$$

上標(j)表示站的編號，$j=1,\cdots,n$，各站的參數 $A_i^{(j)}$ 及 $C_i^{(j)}$ 可分別以單變數 ARMA(p,q)模式各自求得。將 $\boldsymbol{\varepsilon}_t$ 表示為

$$\boldsymbol{\varepsilon}_t=\mathbf{B}\xi_t \tag{5-56}$$

ξ_t 為常態分佈噪音項，$\mathrm{E}[\xi_t\xi_t^{\mathrm{T}}]=\mathbf{I}$，$\mathrm{E}[\xi_t\xi_{t-k}^{\mathrm{T}}]=0$ 且 $k>0$。則對於 ARMA(1,1)，令 $\mathbf{BB}^{\mathrm{T}}=\mathbf{G}$，依與(5-43)至(5-48)式相似的推導過程可得到 **G** 矩陣元素為

$$G^{ij}=\frac{M_0^{ij}(1-A^{(i)}A^{(j)})}{1-A^{(i)}C^{(j)}-A^{(j)}C^{(i)}+C^{(i)}C^{(j)}} \tag{5-57}$$

其中 M_0^{ij} 為 $\mathbf{M}_0$ 的元素。最後，由(5-49a)至(5-49e)式可求得 **B** 矩陣的值。

若水文時間序列資料具季節週期效應，則一階多變數週期移動平均自迴歸模式(periodic multivariate ARMA model)：multivariate PARMA(1,1)為：

$$\mathbf{Z}_{\nu,\tau}=\mathbf{A}_\tau\cdot\mathbf{Z}_{\nu,\tau-1}+\mathbf{B}_\tau\varepsilon_{\nu,\tau}-\mathbf{C}_\tau\varepsilon_{\nu,\tau} \tag{5-58}$$

參數矩陣的推求法與單一時間多變數 ARMA(1,1)類似，亦是將模式簡化為時間性多變數移動平均自迴歸模式：CPARMA，以單變數 PARMA(p,q)模式各自求得以各站的週期性參數 $A_\tau^{(j)}$ 及 $C_\tau^{(j)}$。令 $\mathbf{B}_\tau\mathbf{B}_\tau^{\mathbf{T}}=\mathbf{G}_\tau$，則 $\mathbf{G}_\tau$ 矩陣元素為

$$G_\tau^{(ij)}=e_\tau^{(ij)}+f_\tau^{(ij)}G_{\tau-1}^{(ij)} \tag{-59a}$$

$$e_\tau^{(ij)}=M_{0,\tau}^{(ij)}-A_\tau^{(i)}M_{0,\tau-1}^{(ij)}A_\tau^{(j)} \tag{5-59b}$$

$$f_\tau^{(ij)}=A_\tau^{(i)}C_\tau^{(j)}+A_\tau^{(i)}C_\tau^{(i)}-C_\tau^{(i)}C_\tau^{(j)} \tag{5-59c}$$

其中，上標(ij)表示元素的行列位置，$M_{0,\tau}^{(ij)}$ 表示 $\mathbf{M}_{0,\tau}$ 的元素。

5.8 多變數補遺模式(multivariate infilling model)

以上所介紹的水文時間序列數學模式主要用於合成資料(或預報)之用。對於已發生的事件，若有資料缺漏的情形則必須使用補遺模式。補遺模式與合成資料模式最大的不同在於補遺模式中完全使用歷史水文資料做最佳估計，而不另外加上任何誤差項。

多變數補遺模式(Bras and Rodriguez-Iturbe，1985)主要的方法，是根據歷史水文資料，求出測站與測站之間的協變異數，假設補遺值為其他測站觀測值的線性平均組合，將所有補遺值組成補遺向量，以矩陣法求權重係數矩陣，再將求得之權重係數矩陣代入估計式計算即可得到補遺值向量。

補遺模式可為單一時間或多時間，單一時間多變數補遺模式指的是補遺模式僅使用與遺失資料同一時間的其他測站資料做補遺；而二時間多變數補遺模式則是使用遺失同一時間與前一時間兩個相鄰時間所有測站資料做補遺的工作。

5-7.1 單一時間與二時間基本模式

假設資料需補遺的時間為 t，$\mathbf{Y}_t = \left[y_t^{(1)} \quad y_t^{(2)} \quad \cdots \quad y_t^{(n)} \right]$為需補遺的 n 站之正規化估計值(即以估計值扣除平均值後的殘餘估計值，各站的平均值不必相等)，$\mathbf{X}_t = \left[x_t^{(1)} \quad x_t^{(2)} \quad \cdots \quad x_t^{(p)} \right]$為在時間為 t 時有記錄的 p 站之正規化觀測值(即以觀測值扣除平均值後的殘餘值)，$\mathbf{X}_{t-1} = \left[x_{t-1}^{(1)} \quad x_{t-1}^{(2)} \quad \cdots \quad x_{t-1}^{(q)} \right]$為在時間為 t-1 時有記錄的 q 站之正規化記錄值，則多變數補遺模式的估計式可表示為：

$$\mathbf{Y}_t = \mathbf{A} \begin{bmatrix} \mathbf{X}_{t-1} \\ \mathbf{X}_t \end{bmatrix} \tag{5-60}$$

其中 A 為 $n \times (p+q)$的權重係數未知常數矩陣。將(5-60)式等號兩側均後乘(post-multiply)上$\left[\mathbf{X}_{t-1} \quad \mathbf{X}_t \right]$向量，並取期望值，可得：

$$\left[\mathbf{S}_{Y_t X_{t-1}} \quad \mathbf{S}_{Y_t X_t} \right] = \mathbf{A} \begin{bmatrix} \mathbf{S}_{X_{t-1} X_{t-1}} & \mathbf{S}_{X_{t-1} X_t} \\ \mathbf{S}_{X_t X_{t-1}} & \mathbf{S}_{X_t X_t} \end{bmatrix}$$

$$\mathbf{A} = \left[\mathbf{S}_{Y_t X_{t-1}} \quad \mathbf{S}_{Y_t X_t} \right] \begin{bmatrix} \mathbf{S}_{X_{t-1} X_{t-1}} & \mathbf{S}_{X_{t-1} X_t} \\ \mathbf{S}_{X_t X_{t-1}} & \mathbf{S}_{X_t X_t} \end{bmatrix}^{-1} \tag{5-61}$$

其中$\mathbf{S}_{U_{t-k} V_{t-l}}$代表兩向量$\mathbf{U}_{t-k}$和$\mathbf{V}_{t-l}$間的樣本協變異數，$k$ 和 j 為稽延時間。將(5-61)式求得之 **A** 代入(5-60)式得到補遺模式為：

$$\mathbf{Y}_t = \begin{bmatrix}\mathbf{S}_{Y_tX_{t-1}} & \mathbf{S}_{Y_tX_t}\end{bmatrix}\begin{bmatrix}\mathbf{S}_{X_{t-1}X_{t-1}} & \mathbf{S}_{X_{t-1}X_t} \\ \mathbf{S}_{X_tX_{t-1}} & \mathbf{S}_{X_tX_t}\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{X}_{t\text{-}1} \\ \mathbf{X}_t\end{bmatrix} \tag{5-62}$$

若不考慮前一時間對估計雨量的影響，則單一時間的補遺模式可簡化為：

$$\mathbf{Y}_t = \mathbf{S}_{Y_tX_t}\mathbf{S}_{X_tX_t}^{-1}\mathbf{X}_t \tag{5-63}$$

將 $\mathbf{Y}_t$ 再加回平均值便可計算補遺值。

5-7.2 修正變異數模式

「修正」多變數補遺模式所指的「修正」，是修正估計值變異數使其與歷史觀測資料的變異數相等當補遺站與其他觀測站間資料的相關性較低時，根據統計理論，多變數補遺模式所得之估計會接近其時間平均值，估計值的變異數會偏低，修正的方法為將單一測站多變數補遺模式按照(5-60)式的形式改寫為：

$$y_t^{(j)} = \mathbf{A}^{(j)}\begin{bmatrix}\mathbf{X}_{t-1} \\ \mathbf{X}_t\end{bmatrix} \quad \text{for } j = 1,\cdots n \tag{5-64}$$

其中 $\mathbf{A}^{(j)}$ 為第 j 站的權重向量，上式的變異數為

$$\sigma_{\hat{y}^{(j)}}^2 = \mathrm{E}[\left(y_t^{(j)}\right)^2] = \mathbf{A}^{(j)}\begin{bmatrix}\mathbf{S}_{X_{t-1}X_{t-1}} & \mathbf{S}_{X_{t-1}X_t} \\ \mathbf{S}_{X_tX_{t-1}} & \mathbf{S}_{X_tX_t}\end{bmatrix}\mathbf{A}^{(j)\mathrm{T}}$$

$$\sigma_{\hat{y}^{(j)}} = \left\{\mathbf{A}^{(j)}\begin{bmatrix}\mathbf{S}_{X_{t-1}X_{t-1}} & \mathbf{S}_{X_{t-1}X_t} \\ \mathbf{S}_{X_tX_{t-1}} & \mathbf{S}_{X_tX_t}\end{bmatrix}\mathbf{A}^{(j)\mathrm{T}}\right\}^{1/2} \tag{5-65}$$

以(5-65)式求得的補遺值的標準偏差 $\sigma_{\hat{y}^{(j)}}$ 和歷史資料的標準偏差 $\sigma_{y^{(j)}}$ 比值，修正(5-64)的估計式，則修正變異數的多變數補遺模式為：

$$y_t^{(j)} = \left(\frac{\sigma_{y^{(j)}}}{\sigma_{\hat{y}^{(j)}}}\right)\mathbf{A}^{(j)}\begin{bmatrix}\mathbf{X}_{t-1} \\ \mathbf{X}_t\end{bmatrix} \quad \text{for } j = 1,\cdots n \tag{5-66}$$

若僅考慮單一時間的影響，則以上修正多變數補遺模式可簡化為

$$\hat{y}_t^{(j)} = \frac{\sigma_{y^{(j)}}}{\sqrt{\mathbf{A}^{(j)}\mathbf{S}_{X_tX_t}\mathbf{A}^{(j)\mathrm{T}}}}\mathbf{A}^{(j)}\begin{bmatrix}\mathbf{X}_{t-1} \\ \mathbf{X}_t\end{bmatrix} \quad \text{for } j = 1,\cdots n \tag{5-67}$$