

## 第二章 基礎理論

### 2.1 因子克利金

因子克利金主要由半變異元理論、主成份分析與聯合克利金法三部份所組成，以下將個別介紹：

#### 2.1.1 半變異元理論

地質統計學上常以一種空間隨機函數 $Z(X)$ 表示任何與地質有關之參數，稱為區域性變數(Regionalized variables, 以下簡稱Re.V.)，其中 $X=(x,y,z)$ 代表點所在之空間位置。一般而言，Re.V.所呈現的兩大特質可歸納如下：

- (1)隨機性(randomness)，即在所分析區域中任一點值，皆具不確定性。
- (2)結構性(structure)，對任一研究區域而言，Re.V.除具有上述的隨機性，亦同時具有某種統計上的結構性，如Re.V.於空間上可能具有某種趨勢(trend)，又稱為空間傾向值(drift)；以及Re.V.於不同位置的觀測值之間亦可能具有某種程度的相關性(correlation)。在區域性變數理論中，則以半變異元(semivariogram)作為此相關性之量化表示式。

半變異元亦稱為半變異數(semi-variance)，其可以Re.V.沿特定方向但不同位置間之隨機函數或其殘數值(residual)之差的變異程度來表示，其定義式如下：

$$\gamma(h) = \frac{1}{2}[\text{Var}(Z(x) - Z(x+h))] \quad (2.1-1)$$

其中， $h$ 為相異兩點間之分離向量；應用於二維場域且定義域面積為 $S$ 之平面 $A$ 上，其半變異元可以表示成：

$$\gamma(h) = \frac{1}{2S} \int_s [Z(x) - z(x+h)]^2$$

若Re.V.具等向性(isotropic)，則h為任兩點之分離距離。一般傳統所指的等向性是指空間上任一點的參數於任何方向其參數值皆相同，如透水係數(K)在垂直及水平方向的透水能力相同。而在區域性變數理論中等向性(isotropic)是指其半變異元只跟空間上兩點距離有關，與方向無關，即空間上任兩點之相關性僅與兩點距離有關。以下則假設Re.V.為等向性。

就計算上而言，半變異元亦可由下列經驗式求得：

- (1)假設Re.V.符合穩定(stationary)，亦即Re.V.之平均值為常數，則半變異元可由下式估算[Journal，1984]：

$$\gamma(h) = \frac{1}{2n} \sum_{i=1}^n [Z(x) - Z(x+h)]^2 \quad (2.1-2)$$

- (2)假設Re.V.呈現非穩定(nonstationary)狀態，可將隨機函數Z(x)視為一定常數項m(x) — 稱之空間傾向值(drift)，與一隨機項R(x) — 稱之殘數值(residual)之和；亦即

$$E[Z(x)] = m(x)$$

$$Z(x) = m(x) + R(x)$$

則半變異元可以由下式估算：

$$\hat{\gamma}(h) = \frac{1}{2n} \sum_{i=1}^n [R(x) - R(x+h)]^2 \quad (2.1-3)$$

其中，式(2.1-2)與式(2.1-3)中之 $n$ 為落於間距 $h$ 內之任相異兩樣本點的組合配對數(pairs)。

有了以上半變異元的推算公式，可計算出各間距內之半變異元值，而最後半變異元模式之建立則由 $h$ 對 $\gamma(h)$ 之關係圖2.1-1中以迴歸法或觀察方式決定。

繼續並介紹三個與半變異元相關之名詞：

(1)基值(sill)：

在 $h$ 對 $\gamma(h)$ 關係圖中，如圖2.1-1所示，當 $h$ 愈遠則 $\gamma(h)$ 漸增且收斂至一定值 $C$ ，則此 $C$ 值稱之為基值。

(2)影響範圍(range)：

在 $h$ 對 $\gamma(h)$ 關係圖中，基值所對應之分離距離 $h=r$ ，稱之為影響範圍(range)或影響半徑(radius of influence)。

(3)金塊效應(nugget effect)：

就式(2.1-1)而言， $Z(x)$ 應有 $\gamma(0)=0$ ，惟在實際應用上，常有當 $h=0$ 時， $\gamma(0)=C_0 \neq 0$ 的情況發生，此稱為塊金效應。其發生原因可能是 $Z(x)$ 之量測誤差或 $Z(x)$ 在非常小的距離之內即有相當大之變異，而各觀測值所在位置間之距離較大，故無法顯現極小範圍內 $Z(x)$ 之變異狀況。

一般而言，半變異元模式以以下三種模式為最常見，其分別為：

(1)指數模式(Exponential model)：

$$\gamma(h) = C_0 + Sill[1 - \exp(-3h/Range)] \quad (2.1-4)$$

(2)高斯模式(Gaussian model)：

$$\gamma(h) = C_0 + Sill\{1 - \exp[-(3h/Range)^2]\} \quad (2.1-5)$$

(3)球型模式(Spherical model)：

$$\gamma(h) = C_0 + Sill \left[ \frac{3}{2} (h / Range) - \frac{1}{2} (h / Range)^3 \right] \quad (2.1-6)$$

Re.V除了本身之結構性之外，在不同的Re.V之間，可能也有某種程度上之空間的相關性。在區域性變數理論中，則以複半變異元(Cross-Semivariogram)作為此相關性化之表示。

如同前述之半變異元之定義，複半變異元可定義如下：

$$\gamma_{KK'}(h) = \frac{1}{2} E \left[ (Z_K(x) - Z_K(x+h)) (Z_{K'}(x) - Z_{K'}(x+h)) \right] \quad (2.1-7)$$

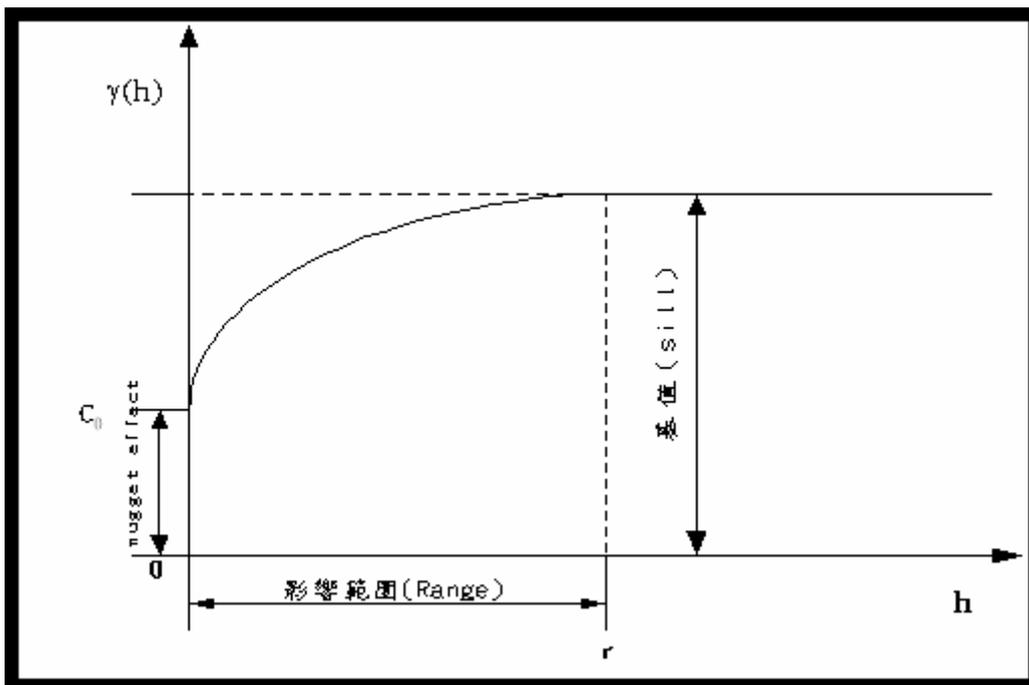


圖 2.1-1 結構分析之示意圖

### 2.1.2 主成份分析

因子分析係指從  $K$  個行為變數萃取出  $J$  個潛伏因子( $K>J$ )，希望能夠在降低變數數目的前提下，找出幾個在概念上有意義，並且彼此之間近於獨立而可以影響原始資料的共同因素，以便我們了解形成某些特定現象的背景原因。

因子分析用來抽取主要因子的因子負荷估計方法有：主成分分析法(Principal component analysis, PCA)、主因子分析法(Principal factor analysis, PFA)、最大概似法(Maximum likelihood method)、映象因子分析法(Image factor analysis, IMAGE)、未加權最小平方法(Unweighted least squares, ULS)與一般化最小平方法(Generalized least squares, GLS)等，一般以主成分分析法使用最廣。其運算元裡與方法如下：

(1)數據標準化處理：

設有  $p$  個原始變量數，有  $n$  個觀測樣本數，則構成一個  $n \times p$  的原始數據矩陣  $X_v$ ：


$$X_v = \begin{bmatrix} xv_{11} & xv_{12} & \cdots & xv_{1p} \\ xv_{21} & xv_{22} & \cdots & xv_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ xv_{n1} & xv_{n2} & \cdots & xv_{np} \end{bmatrix}_{n \times p} \quad (2.1-9)$$

原始的複雜數據單位可能不同，首先將原始數據進行標準化處理後，其平均值為 0 而變異數為 1，即

$$zv_{ij} = \frac{xv_{ij} - \overline{Xv_i}}{sv_i} \quad (2.1-10)$$

其中： $X\bar{v}_i = \frac{1}{n} \sum_{j=1}^n xv_{ij}$ ，為第*i*個變數的平均值。

$$sv_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (xv_{ij} - X\bar{v}_i)^2}$$
，為第*i*個變數的標準差。

經標準化後  $zv_{ij}$  的平均值為 0，其變異數為 1，這樣相關矩陣 (correlation matrix) 和共變異矩陣  $V$  (covariance matrix) 會完全相同。且相關矩陣

$$R = ZVZ' \tag{2.1-11}$$

(2) 計算特徵值(eigen values)及特徵向量(eigen vectors)

在  $p$  個變量所構成的相關矩陣 (correlation matrix) 的基礎上，計算其特徵值及特徵向量。從特徵方程式

$$|R - \lambda I| = 0 \tag{2.1-12}$$

式中  $R$  為相關矩陣， $I$  為單位矩陣，求解出特徵值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ ，及其標準特徵向量  $\bar{U}_i$  ( $i=1, 2, \dots, p$ )，特徵向量矩陣為  $U$ ：

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1p} \\ u_{21} & u_{22} & \dots & u_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1} & u_{p2} & \dots & u_{pp} \end{bmatrix} \tag{2.1-13}$$

由主成分分析之原理：設定一新變數為原始的變量乘上權重，可由權重的不同表示原始變量在新變數中的影響大小。

$$Y_v = U'Z_v \tag{2.1-14}$$

(2.1-14)式中 $Y_v$ 是一組由原始變量 $xv_1, xv_2, \dots, xv_p$ 所組成彼此互不相關的新變量矩陣， $Z_v$ 則為經標準化後的數據矩陣，(2.1-14)式可轉變為 $Z_v = UY_v$ ，其中 $Y_v$ 為新變量矩陣，令

$$Y_v = \begin{bmatrix} \sqrt{\lambda_1} & & & 0 \\ & \sqrt{\lambda_2} & & \\ & & \ddots & \\ 0 & & & \sqrt{\lambda_p} \end{bmatrix}$$

所以 $Z_v = UY_v$

$$Z_v = \begin{bmatrix} u_{11}\sqrt{\lambda_1} & u_{12}\sqrt{\lambda_2} & \cdots & u_{1p}\sqrt{\lambda_p} \\ u_{21}\sqrt{\lambda_1} & u_{22}\sqrt{\lambda_2} & \cdots & u_{2p}\sqrt{\lambda_p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1}\sqrt{\lambda_1} & u_{p2}\sqrt{\lambda_2} & \cdots & u_{pp}\sqrt{\lambda_p} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_p \end{bmatrix} \quad (2.1-15)$$

令 $l_{ij} = u_{ij} \cdot \sqrt{\lambda_i}$

因此，因子負荷矩陣 $L$ 可改寫為：

$$L = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1p} \\ l_{21} & l_{22} & \cdots & l_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pp} \end{bmatrix} = \begin{bmatrix} u_{11}\sqrt{\lambda_1} & u_{12}\sqrt{\lambda_2} & \cdots & u_{1p}\sqrt{\lambda_p} \\ u_{21}\sqrt{\lambda_1} & u_{22}\sqrt{\lambda_2} & \cdots & u_{2p}\sqrt{\lambda_p} \\ \vdots & \vdots & \ddots & \vdots \\ u_{p1}\sqrt{\lambda_1} & u_{p2}\sqrt{\lambda_2} & \cdots & u_{pp}\sqrt{\lambda_p} \end{bmatrix}$$

$$\text{而 } F = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_p \end{bmatrix}$$

所以 $Z_v = LF$ ，若取 $m$ 個因子( $m < p$ )，則

$$Z_v = LF + \varepsilon \quad (2.1-16)$$

其中  $LF$  為  $m$  個主因子所能解釋的部份(共同性)， $\varepsilon$  為無法解釋的殘餘部份(獨特性)， $L$  為因子負荷矩陣， $F$  為主要因子矩陣，若忽略  $\varepsilon$  部份之影響而刪除，則(2.1-16)可改寫為

$$Z_v = LF \quad (2.1-17)$$

因此(2.1-17)式的因子模型即改為

$$\begin{cases} z_1 = l_{11}f_1 + l_{12}f_2 + \cdots + l_{1m}f_m \\ z_2 = l_{21}f_1 + l_{22}f_2 + \cdots + l_{2m}f_m \\ \vdots \\ z_p = l_{p1}f_1 + l_{p2}f_2 + \cdots + l_{pm}f_m \end{cases} \quad (2.1-18)$$

在(2.1-18)式中， $f_1$ 、 $f_2$ 、...、 $f_m$  是在各個變量中共同出現的因子，稱為主要因子。 $l_{ij}$  是第  $i$  個變量在第  $j$  個主要因子上的負荷量，稱為因子負荷量，也就是第  $i$  個變量與第  $j$  個主要因子的相關係數， $l_{ij}$  愈大則表示主要因子  $f_j$  與變量  $x_i$  的關係愈密切。

### 2.1.3 聯合克利金法

聯合克利金法(Co-kriging method)為克利金法的一種延伸。其主要用途在於利用相異變數之間的相關性來增進不易採樣之變數值之推估精確度。假設今日欲推估於  $X_0$  處之  $Z_j$  值，則聯合克利金法可提供一包含了由許多採樣點所觀測到之觀測值之線性推估方程式。這些觀測值不但包含了欲推估的變數，更包含了其餘易於採樣得到的其他變數觀測值，在此介紹一般聯合克利金法(Ordinary co-kriging method)。

聯合克利金推估式可表示如下：

$$Z_j^*(X_0) = \sum_{i=1}^M \sum_{k=1}^{N_i} \lambda_{ij}^k Z_i(X_k) \quad (2.1-19)$$

其中  $Z_j^*(X_0)$ ：利用聯合克利金法推估之參數

$\lambda_{ij}^k$ ：聯合克利金係數

$Z_i(X_k)$ ：已知之觀測值

$M$ ：變數之總數

$N_i$ ：變數  $Z_i$  之觀測值之數量

式(2.1-19)可寫成以下矩陣型式：

$$\tilde{Z}^*(X_0) = \sum_{k=1}^N \tilde{Z}^*(X_k) \Gamma_k \quad (2.1-20)$$

其中  $\tilde{Z}(X) = [Z_1(X) \cdots Z_M(X)]$  (2.1-21)

$$\Gamma_k = [\lambda^k]_{ij} = \begin{bmatrix} \lambda_{11}^k & \cdots & \lambda_{1M}^k \\ \vdots & \ddots & \vdots \\ \lambda_{M1}^k & \cdots & \lambda_{MM}^k \end{bmatrix} \quad (2.1-22)$$

與克利金法相同，聯合克利金法之聯合克利金係數需滿足不偏性 (unbiasedness) 及最小推估方差 (minimum squared error)。這些條件可寫成：

$$E[Z^*(X_0) - Z(X_0)] = 0 \quad (2.1-23)$$

$$\text{Var}[Z^*(X_0) - Z(X_0)] = \min. \quad (2.1-24)$$

式(2.1-24) 可寫成以下型式：

$$\sum_{k=1}^{N_k} \lambda_{ij}^k = 1$$

$$\text{且 } \sum_{k=1}^{N_k} \lambda_{ij}^k = 0 \quad \text{for } i \neq j \quad (2.1-25)$$

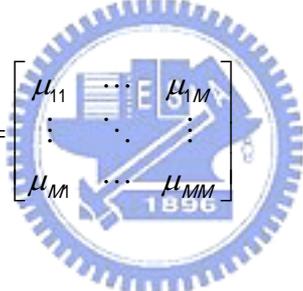
將上述之不偏估條件化作矩陣形式：

$$\sum_{k=1}^N \Gamma_k = I \quad (2.1-26)$$

式(2.1-24)可化寫成以下型式：

$$\bar{\gamma}_{00} - 2 \sum_{i=1}^N \Gamma_i \bar{\gamma}_{x_0} + \sum_{i=1}^N \sum_{j=1}^N \Gamma_i \Gamma_j \bar{\gamma}_{x_i x_j} = \min \quad (2.1-27)$$

式(2.1-27)利用標準拉格蘭茲法，加入一拉格蘭茲矩陣

$$\bar{\mu} = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1M} \\ \vdots & \ddots & \vdots \\ \mu_{M1} & \cdots & \mu_{MM} \end{bmatrix} \quad (2.1-28)$$


可寫成：

$$\bar{\gamma}_{00} - 2 \sum_{i=1}^N \Gamma_i \bar{\gamma}_{x_0} + \sum_{i=1}^N \sum_{j=1}^N \Gamma_i \Gamma_j \bar{\gamma}_{x_i x_j} + 2 \bar{\mu} \left( \sum_{i=1}^N \Gamma_i - I \right) = \min \quad (2.1-29)$$

式(2.1-29)再對  $\Gamma_i$  偏微分使之等於零，可得下式：

$$\sum_{j=1}^N \bar{\gamma}(x_i - x_j) \Gamma_j + \bar{\mu} = \bar{\gamma}_{x_0} \quad i = 1 \cdots \cdots N \quad (2.1-30)$$

上式表成矩陣型式：

$$\begin{bmatrix} \bar{\gamma}(x_1 - x_1) & \cdots & \bar{\gamma}(x_1 - x_N) & / \\ \vdots & \ddots & \vdots & \vdots \\ \bar{\gamma}(x_N - x_1) & \cdots & \bar{\gamma}(x_N - x_N) & / \\ / & \cdots & / & 0 \end{bmatrix} \begin{bmatrix} \Gamma_1 \\ \vdots \\ \Gamma_N \\ \underline{\mu} \end{bmatrix} = \begin{bmatrix} \bar{\gamma}_{x_1,0} \\ \vdots \\ \bar{\gamma}_{x_1,0} \\ / \end{bmatrix} \quad (2.1-31)$$

其中  $\bar{\gamma}(x_i - x_j) = \begin{bmatrix} \gamma_{11}(x_i - x_j) & \cdots & \gamma_{1M}(x_i - x_j) \\ \vdots & \ddots & \vdots \\ \gamma_{M1}(x_i - x_j) & \cdots & \gamma_{MM}(x_i - x_j) \end{bmatrix}$  (2.1-32)

$$\underline{\mu} = \begin{bmatrix} \mu_{11} & \cdots & \mu_{1M} \\ \vdots & \ddots & \vdots \\ \mu_{M1} & \cdots & \mu_{MM} \end{bmatrix} \quad (2.1-33)$$

$$\Gamma_k = [\lambda^k]_{ij} = \begin{bmatrix} \lambda_{11}^k & \cdots & \lambda_{1M}^k \\ \vdots & \ddots & \vdots \\ \lambda_{M1}^k & \cdots & \lambda_{MM}^k \end{bmatrix} \quad (2.1-34)$$

$\bar{\gamma}_{ij}(x_i - x_j)$  代表了  $Z_i$  與  $Z_j$  於  $x_i$  與  $x_j$  位置之半變異元，其為已知值，而  $\underline{\mu}$  與  $\Gamma_k$  為式(2.1-31)欲求之值。

因此，聯合克利金變異數可進一步寫成：

$$\sigma^2 = Tr\left(\sum_{j=1}^N \bar{\gamma}_{x_j,0} \Gamma_j\right) + Tr(\underline{\mu} - \bar{\gamma}_{00}) \quad (2.1-35)$$

其中  $Tr$  表示矩陣的跡(Trace)

共區域化分析(Coregionalization Analysis)為因子克利金一個重要且基本的假設，以下對共區域化分析做一說明：

● 共區域化分析

一組二階定常(Second-order stationary)隨機函數  $\{Z_i(x); i=1, \dots, N\}$  可分解成不同空間尺度的集合  $\{Z_i^u(x); i=1, \dots, N; u=1, \dots, S\}$ ，而這些不同

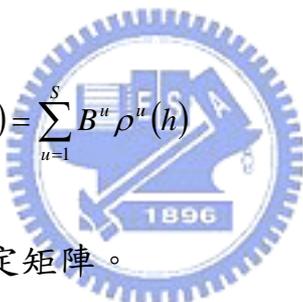
尺度的變數之間不具有空間相關性。其中， $i$ 代表欲探討的變數種類，而 $u$ 代表空間尺度種類。

$$Z_i(x) = \sum_{u=1}^S Z_i^u(x) + m_i \quad (2.1-36)$$

具有多重空間變異尺度的共變異函數 $C_{ij}(h)$ 是由個別空間變異尺度下的基本共變異函數 $\rho^u(h)$ 乘上其所對應的共變異數 $b_{ij}^u$ 所組成，即

$$C_{ij}(h) = \sum_{u=1}^S C_{ij}^u(h) = \sum_{u=1}^S b_{ij}^u \rho^u(h) \quad (2.1-37)$$

以矩陣形式表示

$$C(h) = \sum_{u=1}^S B^u \rho^u(h) \quad (2.1-38)$$


式中， $B^u$ 為一半正定矩陣。

而每一特定空間尺度下之區域化變數組成成份(spatial component)  $Z_i^u(x)$ 可以一組互相獨立的因子 $Y_p^u(x)$ 以及相對應的轉換係數(transformation coefficients)  $a_i^{pu}$ 之乘積的和所取代：

$$Z_i^u(x) = \sum_{p=1}^N a_i^{pu} Y_p^u(x) \quad (2.1-39)$$

式中， $Y_p^u(x)$ 稱為區域化因子(regionalized factor)。

對於每一個特定的空間尺度，每個區域化因子具有相同的半變異元。此外，因每個區域化因子為互相獨立，故

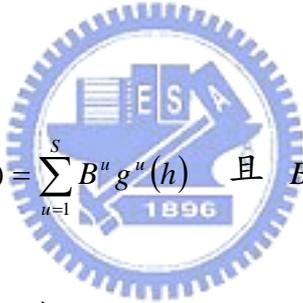
$$\frac{1}{2} E[\{Y_v^u(x) - Y_v^u(x+h)\}\{Y_{v'}^{u'}(x) - Y_{v'}^{u'}(x+h)\}] = g^u(h) \quad (2.1-40)$$

式中，當  $u \neq u'$  且  $v \neq v'$  時等於 0。

如欲以  $g^u(h)$  與  $b_{ij}^u$  表示相同或不同變數間之半變異元，則為

$$\begin{aligned} \gamma_{ij}(h) &= \frac{1}{2} E[\{Z_i(x) - Z_i(x+h)\}\{Z_j(x) - Z_j(x+h)\}] \\ &= \sum_{u=1}^S \sum_{u'=1}^S \sum_{p=1}^K \sum_{p'=1}^K a_{ip}^u a_{jp'}^{u'} \frac{1}{2} E[\{Y_p^u(x) - Y_p^u(x+h)\}\{Y_{p'}^{u'}(x) - Y_{p'}^{u'}(x+h)\}] \\ &= \sum_{u=1}^S \sum_{p=1}^K a_{ip}^u a_{jp}^u g^u(h) \\ &= \sum_{u=1}^S b_{ij}^u g^u(h) \end{aligned} \quad (2.1-41)$$

以矩陣形式表示

$$\Gamma(h) = \sum_{u=1}^S B^u g^u(h) \quad \text{且} \quad B^u = A^u A^{uT} \quad (2.1-42)$$


式中， $B^u$  為半正定矩陣

當自然現象之空間分佈具有多種空間結構特性時，應找出其各別空間結構下之共區域化矩陣  $B^u$ ，在針對各別之共區域化矩陣作多變量分析。

由(2.1-42)，當  $h \rightarrow \infty$  時

$$\begin{aligned} \Gamma(h) &= \sum_{u=1}^S B^u g^u(h) = \sum_{u=1}^S B^u [\rho^u(0) - \rho^u(h)] \\ &= \sum_{u=1}^S B^u [\rho^u(0) - 0] = \sum_{u=1}^S B^u \cdot \rho^u(0) \\ &= \sum_{u=1}^S B^u = V \end{aligned} \quad (2.1-43)$$

可知共變異矩陣  $V$  是由各空間變異尺度下的共區域化矩陣  $B^u$  所線性疊加而成。

由式(2.1-36)與式(2.1-39)可得下式

$$Z_i(x) = \sum_{u=1}^S \sum_{p=1}^K a_{ip}^u Y_p^u(x) + m_i \quad (2.1-44)$$

式中， $Y_p^u(x)$  未知。要求得  $Y_p^u(x)$  必須利用聯合克利金法，故  $Y_p^u(x)$  之聯合克利金推估式為

$$Y_p^{u*}(x) = \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{i\alpha}^u Z_i(x_\alpha) \quad (2.1-45)$$

而聯合克利金系統(Co-kriging system)可表示成

$$\begin{cases} \sum_{j=1}^N \sum_{\beta=1}^{n_j} \lambda_{j\beta}^u C_{ij}^u(x_\alpha - x_\beta) + \mu_i^u = a_{ip}^u \rho^u(x_\alpha - x_0) & i = 1, \dots, N \\ \sum_{\beta=1}^{n_i} \lambda_{i\beta}^u = 0 & \alpha = 1, \dots, n_i \end{cases} \quad (2.1-46)$$

解上述系統方程式，則可得克利金權重 ( $\lambda_{i\beta}^u$ ) 與拉格朗日參數 ( $\mu_i^u$ )，帶回(2.1-45)式即可得推估點上的區域化因子值。

由(2.1-46)式，化簡後可得因子克利金變異數為：

$$\sigma_{FK}^2 = \sum_{i=1}^N \sum_{\alpha=1}^{n_i} \lambda_{i\alpha}^u C_{j\beta}^u(x_\alpha - x_0) + \mu_i^u - C_{i_0 i_0}^u(x_0 - x_0) \quad (2.1-47)$$

## 2.2 遺傳演算法

### 2.2.1 遺傳演算法之概述

遺傳演算法是模仿自然界演化(Natural Evolution)過程的一種演算法，其觀念源自達爾文演化論中之“物競天擇、適者生存”的道理。而所謂自然界的演化或進化，通常是發生在生物體的染色體(Chromosome)上，一條染色體上包含了數量龐大、組成生物體各部份元件的遺傳訊息而形成一獨特的個體。由生物體的特性組成染色體的過程可視為一編碼(Encoding)的過程，而由染色體去推求所形成的生物體可視為一解碼(Decoding)的過程。

自然界的演化程序有以下幾個特性(Davis, [1991]; 陳莉, [1995]):

1. 演化作用是針對編碼後的參數集合加以搜尋，而非針對參數本身。利用編碼後的參數集合，其最大優點，在於不受函數型態的限制。
2. 自然選取(Natural Selection)的過程取決於生物體的表現(Performance)或適合度(Fitness)，此過程使得表現較好、適合度較高之生物體的染色體具有較大的被複製(Reproduce)的機會，亦即符合了達爾文演化論中「物競天擇、適者生存」的理論。
3. 染色體基因交換(Crossover)使得子代(Offspring)繼承了兩個親代(Parents)中較優的特質，而突變(Mutation)則使得子代有機會形成更異於親代的個體。

由於上述的幾項特性，使得遺傳演算法的發明者 John Holland 認為利用自然界的演化特性可以發展出一種可用以求解複雜問題的搜尋技巧(Search Technique)，且經證明可以解決大部份傳統的解析(Alytic)與數值(Numerical)的最佳化技術所難以求解的非連續(Discontinuous)或不可微分(Non-Differentiable)、以及非凸

(Non-Convex)、多峰(Multiple Peaks)的函數最佳化問題。對於多峰函數的最佳化問題，遺傳演算法一次搜尋多點的群集，同時平行攀登許多個峰值，使其得到整體最佳值(Global Optimum)的可能性增加，而不被區域最佳值(Local Optimum)給誤導(陳莉，[1995])。

### 2.2.2 遺傳演算法的架構

遺傳演算法將欲求解的問題變數或參數以一種類似染色體的資料結構(Chromosome-Like data structure)來編碼，並應用一些遺傳運算元(Operators)如交換(Crossover)、突變(Mutation)對大量的染色體作運算，運算後產生的子代除了保存親代中具優勢的特質外，也有可能因為基因的交換與突變而比親代的表現更佳。基本的遺傳演算法包含下列幾個步驟：

#### 1.將問題的變數編碼:

例如可以二進位字串(Binary string)的形式來表示變數，其間的轉換為二進位與十進位的對應，如將二進位字串 1001 解碼，則可對應於十進位的變數值 9，而 1100 對應於 12，1001 與 1100 可看作是兩條染色體。

#### 2.產生初始群集(Initial population)

以隨機的方式產生多條染色體作為初始解。

#### 3.計算目標函數值(Evaluation)

將初始群集大量的染色體解碼後對應的變數值一一代入問題模式中，計算函數或目標函數值。

#### 4.計算適合度

適合度愈高表示該染色體具有較優的特質，將來被複製(Reproduction)的機會也較大，以搜尋最大化目標值的問題來說，

適合度可以目標函數來表示，若是應用於最小化目標函數之問題時，適合度函數則需由目標函數經適當的轉換而產生。

#### 5.複製(Reproduction)或選取(Selection)

由目前之群集中選取染色體作為繁衍的後代，本研究採用比較選取法，即模仿自然界的生物彼此競爭，當某一個體的適合度愈高，經由比較選取後，存活下來被複製的機會愈高，此選取法有一好處，即是染色體被複製下來的機率與染色體間適合度的相對值大小無關而是取決於相對大小，因此較適合個體間適合度值相對變化很大的問題。

#### 6.基因交換(Crossover)或重組(Recombination)

由上一項所選取的染色體中，隨機地將染色體兩兩配對，再經由彼此間所進行之基因交換產生子代。藉此，子代可同時具有父母雙方部份之優良基因，合組成更具有適應能力之染色體。在演算處理過程主要有兩步驟：(1) 進行族群之隨機配對；(2) 隨機產生配對點，再依配對點之位置，進行基因互換。一般常採用的方式有三，即單點交配、雙點交配與均勻交配。其中單點交配為隨機產生一配對點，親代保留配對點前之基因，互換其後之基因而形成下一代。圖 2.2-1 為單點交配示意圖，以二進位編碼為例，某兩組染色體為 1011010 與 1111000，先以亂數選取交換點 (Crossover Point)，如選取的交換點在第 4 位元與第 5 位元之間，則交換後之子代為 1011000 與 1111010。

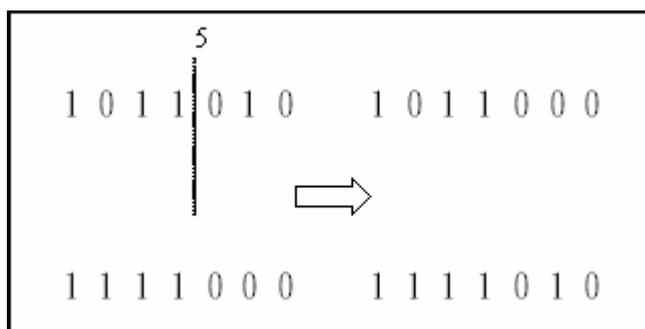


圖 2.2-1 單點交配示意圖

雙點交配係產生兩配對點，親代交換兩點間之基因，而保留其餘部份，如圖 2.2-2 所示。藉由雙點交配理念可擴充成多點交配。

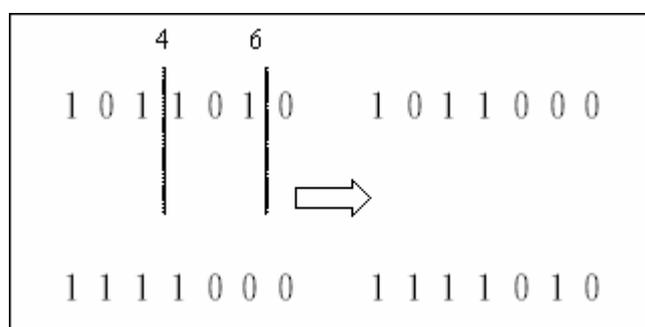


圖 2.2-2 雙點交配示意圖

## 7. 突變

突變即從染色體中隨機任取基因做修正，以防止染色體於複製及交配過程中，遺漏重要訊息或落入局部最佳解。一般處理方式有兩種：(1)基因突變，當產生之亂數低於設定之突變率時，隨機任取基因更動基因值。(2)移轉突變，即變動染色體某一段之基因位置，而不改變其值(曾國雄，1997)。突變率之設定直接影響尋優結果，該值過小，將無法發揮突變功能，可能發生過早收斂現象，反之將破壞子代繼承親代之優良基因。突變率大小隨問題之不同有不同之值，圖 2.2-3 為兩種突變方式之示意圖。

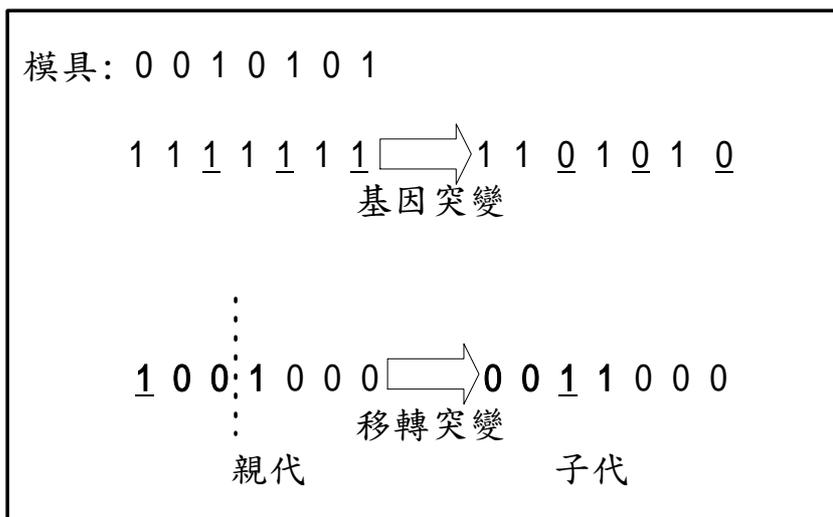


圖 2.2-3 突變示意圖

至此，原來的群集已經演化為第二代，重複上述步驟 3~7，經多代的演算後可望得到較佳的變數解，甚至能夠收斂逼近於整體最佳解。

