

〈論 文〉

多重回歸模型으로 推定된 母數에 의한 最適單位流量圖의 誘導에 관한 研究

A Study on the Derivation of the Unit Hydrograph using Multiple Regression Model

李 種 南* 金 蔡 元** 黃 晶 鉉***
LEE Chong Nam KIM Chai Won HWANG Jung Hyum

Abstract □ A Study on the Derivation of the Unit Hydrograph using Multiple Regression Model.

The purpose of this study is to derive an optimal unit hydrograph using the multiple regression model, particularly when only small amount of data is available. The presence of multicollinearity among the input data can cause serious oscillations in the derivation of the unit hydrograph. In this case, the oscillations in the unit hydrograph ordinate are eliminated by combining the data.

The data used in this study are based upon the collection and arrangement of rainfall—runoff data (1977—1989) at the Soyang—river Dam site.

When the matrix X is the rainfall series, the condition number and the reciprocal of the minimum eigenvalue of XTX are calculated by the Jacobi an method, and are compared with the oscillation in the unit hydrograph. The optimal unit hydrograph is derived by combining the numerous rainfall—runoff data.

The conclusions are as follows:

- 1) The oscillations in the derived unit hydrograph are reduced by combining the data from each flood event.
- 2) The reciprocals of the minimum eigenvalue of XTX , $1/k$ and the condition number CN are increased when the oscillations are active in the derived unit hydrograph.
- 3) The parameter estimates are validated by extending the model to the Soyang river Dam site with elimination of the autocorrelation in the disturbances.

Finally, this paper illustrates the application of the multiple regression model to drive an optimal unit hydrograph dealing with the multicollinearity and the autocorrelation which cause some problems.

要 旨: 本 研究의 目的은 작은 양의 資料만이 사용 가능할 때 多重回歸模型을 利用하여 最適의 單位圖를 誘導하는 데 있다. 入力 資料들 사이에 多重共線性的의 存在는 單位圖 誘導에 있어서 심각한 振動을 일으킨다. 이 경우 資料의 結合을 통하여 振動을 제거함으로써 최적의 단위도를 유도할 수 있다. 多重回歸模型을 利用하여 單位圖를 誘導 하여 行列 XTX 의 $1/k$ 과 CN 를 比較 檢討한 후 獨立된 洪水事像資料의 結合으로 最適의 單位圖를 誘導 하였다.

本 研究에 使用된 資料는 昭陽江댐 支點의 강우—유출 資料(1977~1989)를 收集 整理하여 回歸模型에 適用한 結果는 다음과 같다.

* 수문학회장, 공학박사, 경희대학교 교수

** 정회원, 인덕전문대학, 부교수

*** 정회원, 우보기술단, 수자원부

- 1) 誘導된 單位圖의 振動은 수 개의 獨立된 洪水事像을 結合시킨 결과 減少하였다.
 - 2) 行列 XTX 의 最小 固有值의 逆數인 $1/k$ 과 條件數 CN 은 單位圖에 있어 심한 振動이 있을 때 增加하였다.
 - 3) 模型의 自己相關性을 除去함으로써 推定母數는 信賴性을 갖게 되었다.
- 따라서 問題가 되는 多重共線性과 自己相關性을 처리하여 最適의 單位圖를 誘導할 수 있도록 多重回歸模型의 適用方法을 제시하였다.

1. 序 論

雨量에 대한 流出解析은 새로운 水工構造物의 設計 및 既存의 構造物에 대한 安定性의 評價에 利用된다. 어떤 流域의 特性을 갖는 豪雨事象으로 부터 발생되는 流出을 解析함으로써 豫測不許의 豪雨에 대한 流出을 豫測해 낼 수 있으며, 降雨-流出 解析은 該當 流域에서 最適의 單位圖가 誘導되어진다면 다른 方法보다 쉽게 구하여 질 수 있겠다.

單位圖의 理論은 Sherman(1932)이 豪雨로 부터 河川流量에 대한 研究에서 提示한 이후 Snyder(1955)가 복합호우로부터 最小自乘法을 적용하여 單位圖를 誘導하였으며 Mays와 Coles(1980)는 線型計劃法을 적용하여 單位圖를 유도하는 方法을 제시하였다(Singh, 1988).

Kavvas와 Schulz(1973)는 유도된 單位圖의 2/3는 振動이 존재하며, 특히 單位圖의 後尾部分에 많이 존재한다고 언급하고 low-pass filter를 사용하여 振動을 除去하였다. Delleur와 Rao(1971)는 入力인 雨量과 出力인 直接流出量을 가지고 核 函數에서 發生하는 振動 除去에 관한 연구를 했다.

本 研究에서는 어떤 系(system)의 入力과 出力資料인 有效雨量과 直接流出量을 사용하여 誘導된 單位圖에 대한 振動의 原因이 되는 共線性을 규명하고, 單位圖 誘導에서 發生되는 振動을 回歸模型의 이용으로 除去하여 最適의 單位圖를 誘導하고자 한다.

2. 多重回歸 模型

1. 多重 回歸模型 理論

回歸分析은 變數들의 關係를 연구하는 것으로, 알고 있는 變數 또는 假定된 變數를 사용하여 다른 變數를 推定할 수 있으며 어떤 變數를 推定하기 위해 다른 變數가 사용될 수 있는 變數인지를 識別할 수 있다(김인호, 1985; 박성현, 1981).

多重回歸分析은 獨立變數가 하나인 單純回歸分析

에 대하여 擴張된 것으로 두 개 이상의 獨立變數를 해석하고자 할 때 유용하다(Gaybill, 1980). 본 논문에서는 有效雨量을 獨立變數, 直接流出量을 從屬變數, 單位圖의 縱距를 推定母數로 하여 回歸分析을 시행한다. 回歸分析의 回歸模型은 절편항을 가지고 있으나 有效雨量을 獨立變數로 直接流出量을 從屬變數로 볼 때 獨立變數가 0일 때 반드시 從屬變數도 0이 되어야 한다는 條件을 만족시켜야 하므로 절편이 없는 回歸模型으로 해석하였다. 從屬變數 Y 와 K 개의 獨立變數에서 다음과 같은 多重回歸模型式이 成立된다.

$$Y_i = \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_k X_{ik} + e_i \quad i=1,2,3,\dots,n \quad (2-1)$$

여기서 $\beta_1, \beta_2, \dots, \beta_k$ 는 推定母數이며 e_i 는 發生되는 誤差이고 X_i, Y_i 는 觀測值들이다. 誤差 e_i 項에 대하여 $e_i \sim N(0, \sigma^2)$, $Cov(e_i, e_j) \approx 0, i \neq j$ 로 가정하여 回歸分析理論(Gaybill, 1980; 김인호, 1985; 박성현, 1981; 유석준, 1991; 윤영선, 1990)에 따라 母數를 推定 하였다.

2. 攪亂項의 檢定

單位圖 誘導에 적용된 模型은 回歸假定에서 殘差는 正規的이어야 하고 同分散性을 가져야 하며 自己相關性이 없어야 한다.

母數인 單位圖 從距를 推定함에 있어 最小自乘法을 사용하므로 最小自乘推定量이 正規性的의 假定과 無關하게 不偏的(unbiased)이며 觀測값에 대하여 線形 函數關係에 있고 모든 線形量 중에 가장 적은 分散을 가지고 있다(Gaybill, 1980; Kmenta, 1971; 김인호, 1985; 박성현, 1981). 그리고 Kmenta(1971)는 同分散性이 存在하지 않을 때에도 最小自乘推定量은 不偏的이며 단지 有效性이 缺如된다고 언급하고 있다.

回歸模型은 殘差들이 同分散性과 自己相關性의 가정에 어긋날 때는 模型이 適合하다고 볼 수 없다(김

인호, 1985;박성현, 1981;유석춘, 1991).

同分散性的 가정에서 多重回歸模型에 多數의 說明變數를 사용하게 되므로 둘 이상의 說明變數에 의해 離分散性(heteroscedasticity)이 발생하면 다음과정에 따라 처리한다.

- (1) 最小自乘推定量으로 原回歸模型을 세워 Y_i 에 대한 계산치 \hat{Y}_i 를 얻는다.
- (2) (1)에서 발생한 殘差를 계산치 \hat{Y}_i 에 대해 圖示하여 \hat{Y}_i 축에 대하여 일정하게 分布하고 있는지를 파악한다.
- (3) 原回歸模型에 사용된 全變數의 資料를 回歸分析하여 얻은 \hat{Y}_i 로 나눈다.
- (4) (3)의 變形된 資料를 가지고 回歸分析을 다시 실시한다.

이상의 과정이 加重回歸模型을 적용하는 과정이다 (김인호, 1985).

다음 과정은 自己相關性이 존재할 때 적합한 模型을 위해 資料變換을 통하여 自己相關性을 제거한 模型을 제시하게 된다.

다음의 回歸模型

$$Y_i = \beta_j X_i + e_i \quad j=1, 2, 3, \dots, k \text{ 이고}$$

$$i=1, 2, 3, \dots, n \quad (2-2)$$

에서 殘差의 自己相關은 母數推定에 나쁜 影響을 미치므로 이를 除去하여 母數推定을 실시해야 한다(김인호, 1985;박성현, 1981;소영일, 1989).

母數推定에서 자료의 變換을 式으로 나타내면

$$Y_i' = Y_i - \rho Y_{i-1} \quad (2-3)$$

$$X_i' = X_i - \rho X_{i-1} \quad (2-4)$$

이 되고 式(2-3)에 式(2-2)를 적용하면

$$Y_i' = (\beta_j X_i + e_i) - \rho (\beta_j X_{i-1} + e_{i-1}) \quad (2-5)$$

로 되어 이를 전개하여 정리하면

$$Y_i' = \beta_j (X_i - \rho X_{i-1}) + (e_i - \rho e_{i-1}) \quad (2-6)$$

이된다. 여기서 $e_i' = (e_i - \rho e_{i-1})$ 으로 놓고 式(2-6)에 式(2-4)를 대입하여

$$Y_i' = \beta_j X_i' + e_i' \quad (2-7)$$

과 같은 變換된 模型이 형성된다.

變換된 模型 式(2-7)을 얻기위한 相關係數 ρ 는 資料로부터 다음과 같은 式을 사용하여 구할 수 있다.

$$\rho = \frac{\sum_{i=1}^n e_i - \rho \sum_{i=1}^n (e_i - \rho)}{\sum_{i=1}^n (e_i - \rho)^2} \quad (2-8)$$

여기서 n 은 표본의 크기이다.

이와 같은 資料變換을 통하여 自己相關性을 제거한다.

3. 多重共線性的 問題

1. 多重共線性

多重共線性이란 獨立變數들중 어떤 變數가 다른 變數와 완전히 相關되거나 다른 變數들과 線形關係에 있을 때 이를 多重共線性이라 한다(김인호, 1985;박성현, 1981;유석춘, 1991). 雨量系列에 있어 어느 한 雨量이 다른 한 雨量값과 완전히 相關되거나 雨量값들이 서로 완전한 線形關係에 있을 때를 多重共線性이라 할 수 있겠다.

完全한 多重共線性關係가 있으면 正規方程式에서 $(X'X)$ 行列의 逆行列 $(X'X)^{-1}$ 이 존재하지 않기 때문에 推定值를 얻을 수 없으며, 또한 雨量變數들 간에 완전한 線形關係는 아닐지라도 약간의 線形關係가 존재하면 逆行列 $(X'X)^{-1}$ 은 존재하지만 單位圖 從距값에 나쁜 影響을 끼친다.

2. 多重共線性的 測定方法

多重共線性(multicollinearity)의 존재를 알아내는 방법은 相關係數行列(correlation matrix), 分散 膨脹 要因(variance inflation factor) 또는 固有值 分析(eigenvalue analysis) 등이 있는데 Belsley(1984)는 條件數(condition number)가 30을 超過할 때 共線性이 크다고 제안하고 있다. 본논문에서는 이들중 固有值 分析을 이용하여 多重共線性的의 존재 여부를 파악하고자 한다.

回歸母數의 信賴性은 最小自乘推定量의 共分散行列 $(X'X)^{-1}$ 에 의하여 나타낼 수 있다. 즉, 推定 量의 分散은 다음과 같이 된다.

$$\text{Var}(b) = \begin{bmatrix} \text{Var}(b_1) & \text{Cov}(b_2, b_1) & \dots & \text{Cov}(b_k, b_1) \\ \text{Cov}(b_2, b_1) & \text{Var}(b_2) & \dots & \text{Cov}(b_k, b_2) \\ \dots & \dots & \dots & \dots \\ \text{Cov}(b_k, b_1) & \text{Cov}(b_k, b_2) & \dots & \text{Vov}(b_k) \end{bmatrix} \quad (3-1)$$

식(3-1)은 다음과 같은 식으로 나타낼 수 있다.

$$\text{Var}(b) = (X^T X)^{-1} \sigma^2 \quad (3-2)$$

式(3-2)에서 $k \times k$ 行列 $X^T X$ 의 固有값을 $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_k$ 이라고 할 때 $(X^T X)^{-1}$ 은 다음과 같이 나타낼 수 있다.

$$\text{tr}(X^T X)^{-1} = \sum_{j=1}^k \lambda_j^{-1} \quad (3-3)$$

式(3-3)을 式(3-2)에 대입하면

$$\sum_{j=1}^k \text{Var}(b_j) = \sigma^2 \sum_{j=1}^k \lambda_j^{-1} \quad (3-4)$$

여기에서 $X^T X$ 行列의 要素들이 완전히 多重共線性을 갖는다면 行列은 非正則性을 갖게되며 k 개의 固有值 중 거의 0에 가까운 값이 존재하게 되므로 式(3-4)에서 推定量의 分散은 增加하게 된다(박성현, 1981; 유석춘, 1991).

따라서 雨量系列에서 最小固有值의 逆數인 $1/\lambda_k$ 으로 雨量系列의 多重共線性의 存在를 評價한다.

4. 適用 및 分析

1. 對象流域

本 研究에서 적용하려는 昭陽江댐 流域은 面積이 $2,703\text{km}^2$ 이고 流路延長은 166.2km에 달한다(대한 토목학회, 1962). 本 流域의 降雨觀測所別 支配面積은 表 4.1과 같다(건설부, 1968).

表 4.1 雨量觀測所別 支配面積

관측소	서화	원통	현리	창촌	인제	추암	계
지배면적(km ²)	500.3	469.4	532.6	452.4	383.6	364.7	2,703

表 4.3 水平分離로 구한 3시간 間隔의 直接流出量

事象	直 接 流 出 量 (m ³ /sec)
1	189 199 399 1407 708 510 411 310 300 212 199 80 50 15
2	90 125 255 393 611 832 763 490 385 255 220 180 125 90 80
3	30 50 80 230 990 2310 1965 1795 1625 1525 1400 1265 1140 1265 1140 1005 850 685 505 350 230 150 90 60 40
4	10 15 35 55 155 555 955 1520 1075 855 775 540 470 385 320 260 180 135 115 55 25 15 5
5	19 50 144 512 1130 1884 3656 1409 1257 907 761 590 467 366 290 211 144 95 19
6	155 2455 685 1914 1605 1180 1025 760 655 610 545 435 335 255 81
7	299 722 824 1398 824 662 538 330 321 281 260 231 180 161 81
8	90 170 650 1285 1410 1580 1110 620 410 260 200 100 70 40 10

2. 水文資料

回歸 分析의 入力資料로 對象流域의 有效雨量과 直接流出量을 사용한다. 水文資料는 昭陽江댐의 1976-1989年 洪水水文資料인 降雨-流出量을 圖示하여 最大流出量의 發生時間 및 記錄의 誤差로 인하여 발생할 有效雨量과 流域 平均雨量의 차이를 檢討하여 양호한 雨量과 流出量資料를 택하여 ϕ -index법으로 有效雨量을 산정하였다.

直接流出量과 基底流量의 分離는 水平直線分離法에 의하여 실시하였다.

3시간 간격으로 산정한 有效雨量값과 直接流出量은 表 4.2와 表 4.3과 같다.

表 4.2 ϕ -Index로 算定한 3시간 間隔의 有效雨量

事象	날 짜	有 效 雨 量(cm)
1	1977. 7. 9	1.74 0.26
2	1977. 8. 7	1.57 0.00 0.46
3	1978. 8. 19	0.41 0.09 0.93 2.70 0.95 0.40 0.63 0.34 0.01 0.16 0.71
4	1979. 8. 3	0.29 0.25 0.00 0.00 0.00 0.07 1.61 0.25 0.33 0.32 0.24
5	1981. 7. 11	1.00 1.30 1.50 1.40 0.30 0.0 0.1
6	1985. 10. 12	0.42 2.56 1.23
7	1988. 7. 8	2.58 0.22 0.02
8	1989. 7. 24	0.35 2.70 0.00 0.50

Matrix 解法을 이용하여 얻어진 8개의 單位圖 중 진동이 발생한 4개의 事象은 代表單位圖를 얻는 일반적인 방법에서 제외시켜야 하나 資料가 많을 수록 推定母數를 精確하게 推定할 수 있는 多重回歸模型을 적용하면 귀중하게 얻어진 資料의 손실없이 해당 유역에 대한 最適의 單位圖를 얻을 수 있다.

3. Matrix 解法

Matrix解法으로 單位圖를 구하는 式은 다음과 같다.

$$X^T X U = X^T Y \quad (4-1)$$

따라서

$$U = (X^T X)^{-1} X^T Y \quad (4-2)$$

이고 U는 單位圖 從距, X는 有效雨量, Y는 直接流出量이다.

單位圖의 解를 얻기 위한 聯立方程式을 푸는 수단으로 Gauss-Jordan법을 이용한다(Gerald, 1984; Maron, 1987).

聯立方程式의 解를 推定하는 問題에서 資料에 있어 推定값에 어느 정도 誤差를 확대시키느냐 하는 尺度로 行列 XTX의 條件數를 사용한다. 條件數(condition number) 檢査는 Diskin과 Boneh(1975)가 유클리드 노름을 사용하여 行列 XTX의 最小固有値와 最大固有値의 比로 나타냈다. 따라서 條件數 CN을 구하는 式은

$$CN = \frac{\lambda_{\max}}{\lambda_{\min}} \quad (4-3)$$

와 같으며, 式(4-3)을 계산하는 데 있어 固有値의 계산은 Jacobi법을 이용하여 산정하였다. (Gerald, 1984; Maron, 1987).

Matrix 解法을 이용하여 8개의 洪水事象으로부터 각각에 대한 單位圖를 구하고 條件數(condition number)의 變化가 單位圖의 縱距의 變化에 따라 어느 정도 變하는지를 그림 4.1~그림 4.3에 나타냈으며 單位圖의 振動이 심할수록 條件數 CN값이 增加함을 알 수 있다.

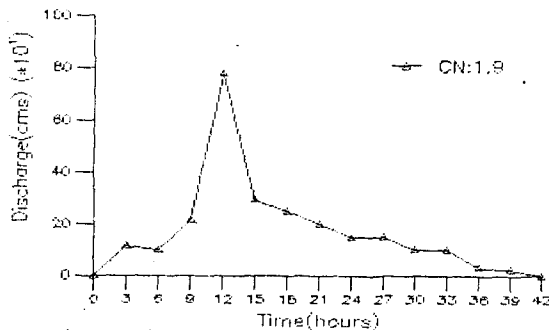


그림 4.1 事象 1에 대한 單位圖

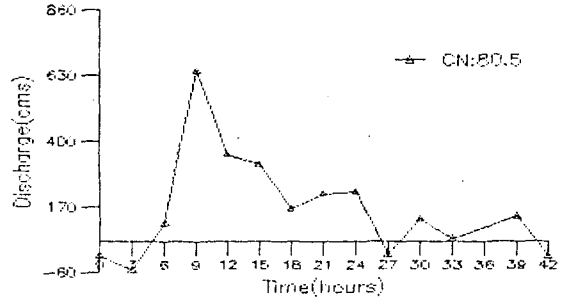


그림 4.2 事象 3에 대한 單位圖

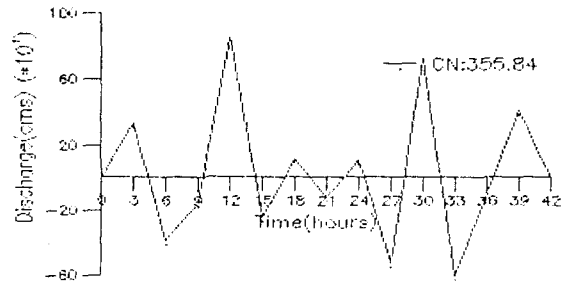


그림 4.3 事象 5에 대한 單位圖

4. 多重回歸模型의 適用

多重回歸模型을 單位圖 誘導를 위한 模型으로 變換하면 다음과 같은 式이 성립된다.

$$Y_i = h_0 X_i + h_1 X_{i-1} + \dots + h_{n-1} X_{i-n+1} + e_i \quad (4-4)$$

여기서 Y는 直接流出量의 縱距이고, X는 有效雨量, h_n는 本 模型에서 구하려는 母數인 單位圖 縱距값이다. 또한 e_i는 說明變數인 有效雨量 X_i로는 說明될 수 없는 項이다.

式(4-4)를 行列形態로 나타내면

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ \vdots \\ Y_m \end{bmatrix} = \begin{bmatrix} X_1 & & & 0 \\ X_2 & X_1 & & \\ X_3 & X_2 & X_1 & \\ \vdots & \vdots & \vdots & X_2 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ X_j & & & 0 \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \\ h_{n-1} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ \vdots \\ e_m \end{bmatrix} \quad (4-5)$$

와 같이 된다. 즉, $\underline{Y} = \underline{X}h + e$ 가 된다.

어떤流域의 單位圖를 誘導할 때 하나의 洪水事象으로부터 얻는 것은 바람직하지 못하며 유도된 單位圖에 振動이 존재하지 않을 때에도 하나의 單位圖만으로 해당 流域의 單位圖로는 적합하지 않다.

여러 개의 洪水事象이 유용할 때 單位圖의 最小自乘推定量은 殘差의 自乘을 最小化시키므로 얻어진 8개의 洪水事象으로부터 單一 最適單位圖를 유도한다. 두 개의 洪水事象의 結合形態를 다음과 같이 예시할 수 있다.

$$\underline{Y}_1 = \underline{X}_1 h + e_1 \tag{4-6}$$

$$\underline{Y}_2 = \underline{X}_2 h + e_2 \tag{4-7}$$

式(4-6)과 式(4-7)을 結合시켜 殘差가 最小化하기 위한 自乘은

$$e_1^T e_1 + e_2^T e_2 \tag{4-8}$$

이므로 正規方程式은 다음과 같다.

$$(\underline{X}_1^T \underline{X}_1 + \underline{X}_2^T \underline{X}_2) h = \underline{X}_1^T \underline{Y}_1 + \underline{X}_2^T \underline{Y}_2 \tag{4-9}$$

각 洪水事象의 直接流出量은 m , 單位圖 縱距를 n , 有效雨量을 j 로 할 때

$$n = m - j + 1 \tag{4-10}$$

이 된다. 이 식에서 m 을 조정하여 各事象의 n 이 같도록 지체시킨 雨量系列과 流量系列을 연결하여 길게 배열하면 式(4-9)는 式(4-5)와 같이 各계열을 길게 배열한 것과 같다.

제시된 資料結合에 따라 洪水事象 중 洪水事象 1, 2, 3의 資料를 結合하여 誘導된 單位圖와 개별사상의 單位圖를 그림 4.4에 나타냈다. 그림에서 알 수 있듯이 結合된 洪水事象은 개별 洪水事象의 單位圖보다 어느 정도 振動이 감소한 것을 알 수 있다.

共線性的의 測定으로 雨量系列의 最小固有值의 逆數 $1/\lambda_1$ 를 이용하였으며 資料 하나만의 $1/\lambda_1$ 은 0.45를 보이고 있고 洪水事象 1, 2와 3을 結合한 資料의 雨量系列에 대한 $1/\lambda_1$ 은 0.15로 減少했다. 또한 8개의 모든 事象을 사용했을 때는 0.04까지 減少했다.

條件數 CN를 살펴보면, 條件數 또한 洪水事象 3의 條件數 CN이 큼에도 불구하고 資料의 結合에 의해 洪水事象 1, 2, 3의 資料의 條件數 CN이 작아짐을 알

수 있다. 表 4.4는 各資料에 대한 $1/\lambda_1$ 과 CN의 變化를 나타내고 있다.

表 4.4 雨量系列의 共線性 檢討

洪水事象	1	2	3	3個의 事象	8個의 事象
$1/\lambda_1$	0.45	0.74	1.85	0.15	0.04
CN	1.90	2.97	80.5	7.64	5.34

이상에서 誘導된 單位圖 중에서 振動과 $1/\lambda_1$ 의 樣相을 그림 4.4에 圖示하였다. 그림에서 單位圖 縱距의 振動이 심할수록 $1/\lambda_1$ 값은 크다는 것을 알 수 있다.

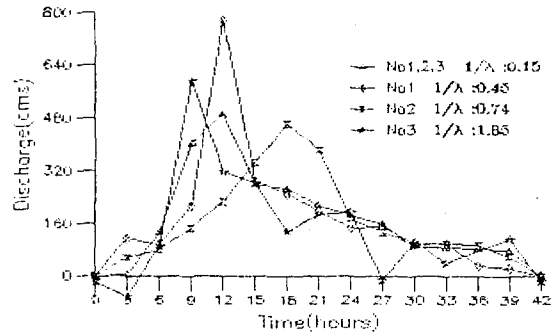


그림 4.4 事象 1, 2, 3과 結合된 3個事象의 單位圖

4.1 適合性 檢討

適合性 檢討에서 分散分析은 절편이 없는 경우로 시행할때 總變動(SST)은 回歸에 의해 說明되는 變動(SSR)과 殘差에 의한變動(SSE)으로 구분되며 總變動은

$$\sum Y_i^2 = \sum (Y_i - \hat{Y}_i)^2 + \sum \hat{Y}_i^2 \tag{4-11}$$

이고 이식을 간단히 정리하면

$$SST = SSE + SSR \tag{4-12}$$

이 된다(박성현, 1981).

8개의 모든 洪水事象 資料分析에서의 修正決定係數(adjusted determinant coefficient) (김인호, 1985) R_a은 0.80의 값을 보이고 있다. 즉, 80%는 잘 說明되고 있다고 볼 수 있다. 有意水準 α 를 1%로 하여 分散分析한 결과는 표4.5와 같다.

表 4.5 變換 資料에 대한 分散分析表

要因	自乗合	自由度	自乗平均	F ₀	F(α)
回歸	3778.7	13	290.7	43.4	2.133
殘差	848.6	126	6.7		

4.2 殘差分析

模型을 적용하여 얻어진 殘差를 分析하여 資料를 變換시킴으로써 보다 適合性을 갖는 模型에 접근할 수 있다.

우선 同分散性을 검토하기 위해 殘差를 直接流出量 計算值 Y_i에 대하여 그림 4.5에 圖示하였다. 그림에서와 같이 殘差가 일정하게 分布하므로 多重回歸模型의 反復解析은 필요하지 않다.

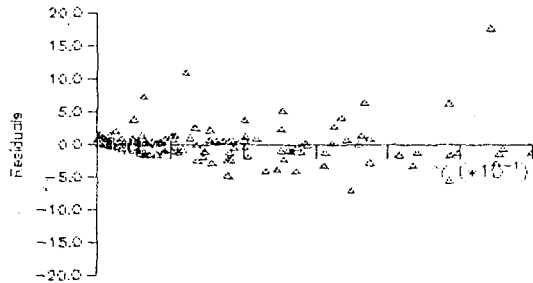


그림 4.5 計算值 Y_i와 殘差 e_i

다음은 殘差의 相關性에 관한 것으로 相關性의 程度를 나타내는 尺度인 統計量은 Durbin-Watson 檢定값 d를 사용한다. Durbin과 Watson은 d값이 1.5보다 작으면 陽의 相關을 갖고 2.5보다 크면 陰의 相關性을 갖는다고 했다. 하지만 일반적으로 資料의 數 m이 비교적 클 때 Durbin-Watson값 d가 0에 가까우면 自己相關性이 크다고 보며 2에 가까우면 自己相關이 存在하지 않는다(Ott, 1988). Durbin-Watson 統計量 d를 구하는 式은

$$d = \frac{\sum_{i=2}^m (e_i - e_{i-1})^2}{\sum_{i=2}^m e_i^2} \quad (4-13)$$

과 같이 된다.

回歸分析된 模型에서의 觀測流出量과 豫測流出量 값의 殘差를 分析하여 얻은 統計量 d는 1.33을 나타냈다. 이것은 統計값 d의 限界值 2에 近似하지 않으므로 相關性이 존재한다고 본다.

II-2절에서 언급한 방법에 따라 相關性을 除去하려면 相關係數 ρ를 우선적으로 구하여야 한다. 결정계수 (R²)를 이용하여 구한 相關係數 ρ는 0.33을 얻었다. 구해진 相關係數를 사용하여 殘差는 다음과 같은 關係를 갖는다고 할 수 있다.

$$e_i = \rho e_{i-1} + u_i \quad (4-14)$$

여기서 u_i는 平均이 0이고 分散이 σ_u²인 確率變數이며, u_i는 e_{i-1}에 獨立이다. 사용된 原模型을 式으로 나타내면 다음과 같다.

$$Y_i = \sum_{j=0}^{n-1} h_j X_{i-j} + e_i \quad (4-15)$$

式(4-15)에서 觀測된 流出量과 雨量資料를 식(2-3)과 식(2-4)를 이용하여 아래와 같은 變換된 式을 誘導 할 수 있다.

$$Y_i - \rho Y_{i-1} = \sum_{j=0}^{n-1} h_j (X_{i-j} - \rho X_{i-j-1}) + u_i \quad (4-16)$$

統計量 d가 2에 近似할 때까지 資料變換을 실시한다.

資料變換을 適用한 두 번째 Durbin-Watson 統計量 d는 1.91를 나타냈으며, 相關係數는 0.04였다. 따라서 統計量 d가 2에 近似하므로 相關性이 除去되었다고 본다. 表 4.6과 表 4.7에서 알 수 있듯이 相關性을 除去하지 않은 最小自乘法과 相關性을 除去한 最小自乘法을 이용하여 推定된 母數인 單位圖 縱距 값은 약간의 차이를 보이고 있다. 여기서 自己相關性을 除去하지 않은 분석에서의 標準偏差는 2.77을 나타냈으며, 相關性을 除去한 분석에서는 2.60을 나타내어 自己相關性을 除去한 분석의 推定量이 보다 더 效率性을 갖는다고 볼 수 있다. 多重回歸模型에 의하여 誘導된 最適單位圖를 1988년 7월 8일 - 7월 10일의 基低流出量 399.4 (m³/sec)를 포함한 실제 發生 雨量에 適用한 결과 그림 4.6과 같다.

表 4.6 自己相關性이 存在하는 單位圖縱距

시간	3	6	9	12	15	18	21	24	27	30	33	36	39
종 거	35	180	339	461	319	317	251	135	130	96	86	66	60

表 4.7 自己相關을 除去한 單位圖縱距

시간	3	6	9	12	15	18	21	24	27	30	33	36	39
종 거	20	180	340	463	317	315	251	135	127	93	86	65	46

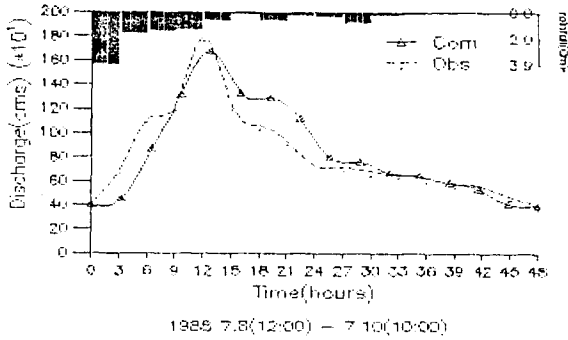


그림 4.6 單位圖를 用한 線 流出水文曲數

5. 結 論

本 研究에서는 一般線形模型인 多重回歸分析을 실시하여 最適의 單位圖를 誘導하는 過程에 있어서 다음과 같은 결과를 얻을 수 있었다.

- (1) 多重共線性的 존재는 單位圖 縱距인 母數推定에 나쁜 影響을 미쳐 該當流域에서의 單位圖는 심한 振動을 야기시킨다는 것을 알았으며 單位圖의 振動은 資料를 結合함으로써 振動이 除去된다는 것을 확인 할 수 있었다.
- (2) 多重共線性的 存在 尺度인 $1/\lambda_k$ 와 Matrix 解法에서 檢討한 不良條件의 尺度인 條件數 CN은 誘導된 單位圖의 振動이 심할 때에는 $1/\lambda_k$ 와 CN의 값이 增加함을 알았다.
- (3) 8개의 모든 洪水事象을 結合시켰을 때, $1/\lambda_k$ 와 CN은 減少하였으며 多重共線性(multicollinearity)도 減少되어 單一의 最適 單位圖를 얻었다.
- (4) 單位圖 誘導에서 問題가 되는 多重共線性和 自己相關性的 處理방법을 예시하므로 最適의 單位圖를 誘導할 수 있도록 한 本例示는 다른 流域에도 적용이 가능하리라 생각된다.

參考文獻

1. 건설부(1968), 소양강 다목적댐 기본설계 보고서.
2. 김인호(1985), 회귀분석론, 비봉출판사, pp.143-168
3. 대한토목학회(1962), 소양강유역 수문조사서, p4.
4. 박성현(1981), 회귀분석, 대영사, pp.171-299.
5. 수자원공사(1977-1989), 소양강댐 홍수수문 기록자료.
6. 소영일(1989), 회귀분석, 창지사, pp.347-352.
7. 유석춘(1991), 예를 중심으로 한 회귀분석, 대광문화사, pp.187-227.
8. 윤영선, 박정식(1990), 현대통계학, 다산출판사, pp.371-457.
9. Belsley, D.A.(1984)Collinearityand Forecasting, Journal of Forecasting, Vol.3, pp.183-196.
10. Delleur, J.W., and A.R.Rao(1971), Linear systems analysis in hydrology, the transform approach, the kernel oscillations and the effect of noise. InSystems Approach to Hydrology, Water Resources Publ. FortCollins, Colo., pp.116-142
11. Diskin, M.H., and A.Boneh(1975), Determine of an optimal IUH forlinear, time invaria systems from multistorm records, Journal of Hydrology, Vol.24, pp.57-76.
12. Gaybill, F.A.(1980), Introduction to the Theory of Statistics, McGrawHil Book Company, New York,pp.117-126.
13. Gerald, C.F., and P.O.Wheatley(1984), Applied Numerical Analysis, Addison Wesley Publishing Company, pp.91-99.
14. Kavvas, M.L., and E.F.Schulz(1973), Removal of unit hydrograph oscillations by filter ing.Floods and Drought, Water Resour. Publ., Fort Collins, Colo. pp. 167-174.
15. Kmenta, J.(1971), Elements of Econometrics, Macmillan, NewYork, pp.430-442.
16. Maron, M.J.(1987), Numerical Analysis, Macmillan Publishing Company, pp.156-159.
17. Mays, L.W., and L.Coles(1980), Optimization of unit hydrograph determination, Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, 106(HY1), pp.85-97.
18. Ott, L.(1988), An Introduction to Statistical Methods and Data Analysis, Pws-Kent Publishing Company, pp.594-598.
19. Sherman, L.K.(1932), Streamflow from rainfall by unit graph method, Engineering News-record108, pp.501-505.
20. Singh, v.p.(1988), Hydrologic System, Rainfall-Runoff Modeling vol 1, Prentice Hall, pp. 55-360.
21. Snyder, W.M.(1955), Hydrograph analysis by the method of least square, Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers, Vol.81, No.793, p.25.