

經濟學에 있어서의 要因分析의 適用

邊 衡 尹

<目 次>	
I.	序 言
II.	要因分析
III.	統計技術의 適用에 대하여
IV.	經濟發展論에의 適用에 대하여(其1)
V.	經濟發展論에의 適用에 대하여(其2)
VI.	結 言

I. 序 言

C. Spearman 이 1904 年에 發表한 General Intelligence, Objectively Determined and Measured, *American Journal of Psychology* No. XV, 라는 論文이 要因分析(factor analysis)의 始初를 이루는 것이라고 한다. 이에서도 짐작할 수 있는 바와 같이 要因分析은 元來는 心理學의 領域에서 使用되기 시작한 것이다.

그러나 그 後 그것은 社會學, 經濟學⁽¹⁾ 등에까지 그 領域을 擴大했다. 그리하여 오늘 날에 와서는 經濟成長 내지 經濟開發의 理論에서까지 要因分析의 適用例를 찾아 볼 수 있게끔 되었다.

論者의 關心은 주로 이 새로운 適用分野에 있다. 그러나 그들을 살피기에 앞서서 우선 要因分析의 內容을 簡單히 살필 必要가 있다. 그리고 R. Stone의 業績은 비록 統計技術的 適用例라고는 하지만 이 統計技術的 適用의 分野에서 가장 先進된것으로 看做되고 있는 T. Kloek 및 L. B. M. Mennes의 業績의 先驅를 이루고 있고 또 要因分析의 內容을 理解하는데 도움이 된다는 意味에서 큰 意義를 갖고 있다.

따라서 以下에서는 우선 簡單히 要因分析의 內容과 R. Stone의 業績을 살핀 後 經濟成長 내지 經濟開發의 理論에서의 適用例인 M. Megee의 業績과 I. Adelman 및 C. T. Morris

(1) 社會學에서의 適用例로서는 H. F. Goesnell and M. Schmidt, "Factorial and Correlation Analysis of the 1934 Vote in Chicago," *Journal of the American Statistical Association* (1936)을 들 수 있다. 그리고 本論稿에서 다루어지는 것 外의 比較의 새로운 經濟學에서의 適用例로서는 R. J. Meyer and E. Kuh, *The Investment Decision* (1956)을 들 수 있다.

의業績을 各各 살피고 끝으로 結論에 對해서 言及하는 順序를 取하기로 한다.

II. 要因分析

經濟學에서 主로 應用되는 要因分析은 H. Hotelling 이 考案한 主成分分析(principal-component analysis)⁽²⁾이다. 이 分析의 概要는 다음과 같다.

지금 標準化된 變數(standardized variable)⁽³⁾ z_1, z_2, \dots, z_p 를 생각하여 이것을 보다 基本的인 變數 u_1, u_2, \dots, u_p 로 代置할 것을 試圖한다고 하자. u_1, u_2, \dots, u_p 를 다음과 같이 定義한다.

$$\begin{aligned} z_1 &= k_{11} u_1 + k_{12} u_2 + \dots + k_{1p} u_p \\ z_2 &= k_{21} u_1 + k_{22} u_2 + \dots + k_{2p} u_p \\ &\dots\dots\dots \\ z_p &= k_{p1} u_1 + k_{p2} u_2 + \dots + k_{pp} u_p \end{aligned} \tag{2.1}$$

이 u_1, u_2, \dots, u_p 를 主成分(principal component)이라고 한다. 이때 이들 u 는 直交條件 즉

$$\sum_{i=1}^N u_{ji} u_{ji} = 0 \quad (i \neq j) \tag{2.2}$$

을 滿足한다.

k_{ij} 는 常數이며 z_i 와 z_j 의 相關係數 r_{ij} 는

$$r_{ij} = k_{i1} k_{j1} + k_{i2} k_{j2} + \dots + k_{ip} k_{jp} \quad (i, j = 1, 2, \dots, p) \tag{2.3}$$

이다.

$$S_i = k_{i1}^2 + k_{i2}^2 + \dots + k_{ip}^2 \quad (i = 1, 2, \dots, p) \tag{2.4}$$

로 하면 이 S_i 는 i 번째 主成分 z_i 의 모든 標準化된 變數 u_i 의 分散(variance)에 대한 寄與度를 나타내는 것이 된다.

따라서 (2.3)의 制約下에 첫번째 主成分의 寄與度,

$$S_1 = \sum_{i=1}^p k_{i1}^2 \tag{2.5}$$

(2) H. Hotelling, "Analysis of a Complex of Statistical Variables into Principal Components," *Journal of Educational Psychology*(1933)

이 principal-component analysis 外에 principal-factor analysis가 있다.

(3) X_{it} ($i=1, 2, \dots, p; t=1, 2, \dots, N$)를 確率變數의 集合, $\bar{X}_i = \frac{1}{N} \sum_{t=1}^N X_{it}$ ($i=1, 2, \dots, p$), $x_{it} = X_{it} - \bar{X}_i$ ($i=1, 2, \dots, p; t=1, 2, \dots, N$), $S_{ij} = \sum_{t=1}^N x_{it} x_{jt}$ ($i, j=1, 2, \dots, p$), $a_{ij} = \frac{S_{ij}}{N-1}$ ($i, j=1, 2, \dots, p$)로 할때 다음의 Z_{it}

$$Z_{it} = \frac{x_{it}}{\sqrt{a_{ii}}}$$

즉 X_{it} 의 그 平均으로부터의 偏差를 標準偏差로 나눈 값을 標準化된 變數라고 한다.

를 極大로 하기 위해서는

$$F = \sum_{i=1}^p k_{i1}^2 - \sum_{i=1}^p \sum_{j=1}^p \left(\mu_{ij} \sum_{s=1}^p k_{is} k_{js} - r_{ij} \right) \quad (2.6)$$

의 極大值를 求하면 된다. 但 (2.4)의 μ_{ij} 는 Lagrange의 未定乘數이다. F 를 첫번째 主成分의 係數인 $k_{is}(i=1, 2, \dots, p, s=1, 2, \dots, p)$ 에 관해서 偏微分하면 다음 式을 얻는다.

$$\begin{aligned} k_{i1} - \sum_{j=1}^p \mu_{ij} k_{j1} &= 0 \\ - \sum_{j=1}^p \mu_{ij} k_{js} &= 0 \quad (s=2, 3, \dots, p) \end{aligned} \quad (2.7)$$

但 $\mu_{ij} = \mu_{ji}$ 이다. 이 (2.7)의 兩邊에 k_{i1} 을 곱하고 i 에 관해서 合計하면 다음과 같이 된다.

$$\begin{aligned} \sum_{i=1}^p k_{i1} - \sum_{i=1}^p \sum_{j=1}^p \mu_{ij} k_{i1} k_{j1} &= 0 \\ - \sum_{i=1}^p \sum_{j=1}^p \mu_{ij} k_{i1} k_{js} &= 0 \quad (s=2, 3, \dots, p) \end{aligned} \quad (2.8)$$

그러나 (2.7)의 第1式에서 $\sum_{j=1}^p \mu_{ij} k_{j1} = k_{i1}$ 이다. 따라서 $\sum_{i=1}^p k_{i1}^2 = \lambda_1$ 로 놓으면 (2.8)은 다음과 같이 된다.

$$\begin{aligned} \lambda_1 - \sum_{j=1}^p k_{j1}^2 &= 0 \\ - \sum_{j=1}^p k_{j1} k_{js} &= 0 \quad (s=2, 3, \dots, p) \end{aligned} \quad (2.9)$$

이 (2.9)의 各式에 k_{is} 를 곱하고 s 에 관해서 合計하면 다음과 같이 된다.

$$\lambda_1 k_{i1} - \sum_{j=1}^p \sum_{s=1}^p k_{j1} k_{js} k_{is} = 0 \quad (i=1, 2, \dots, p) \quad (2.10)$$

그러나 (2.3)을 考慮하면 이 (2.10)은 다음과 같이 된다.

$$\begin{aligned} k_{11} + r_{12} k_{21} + \dots + r_{1p} k_{p1} &= \lambda_1 k_{11} \\ r_{21} k_{11} + k_{21} + \dots + r_{2p} k_{p1} &= \lambda_1 k_{21} \\ \dots & \\ r_{p1} k_{11} + r_{p2} k_{21} + \dots + k_{p1} &= \lambda_1 k_{p1} \end{aligned} \quad (2.11)$$

이 (2.11)은 k_{i1} 에 관한 聯立同次線形方程式이므로 零이 아닌 根을 갖기 위해서는 다음 行列式이 0이 되지 않으면 안된다.

$$\begin{vmatrix} (1-\lambda_1) & r_{12} & \dots & r_{1p} \\ r_{21} & (1-\lambda_1) & \dots & r_{2p} \\ \dots & \dots & \dots & \dots \\ r_{p1} & r_{p2} & \dots & (1-\lambda_1) \end{vmatrix} = 0 \quad (2.12)$$

이 行列方程式(2.12)의 最大根은 첫번째 主成分과 關聯을 갖고 있다. 이것은 (2.9)의

第1式에서 알수있다. 첫번째 主成分의 係數는 (2.11)에서 計算된다.

다음에 行列方程式(2.12)의 두번째로 큰 根을 計算한다. λ_1 代身에 이 根을 (2.11)에 代入하여 첫번째 主成分이 除去된 後에 標準화된 變數의 分散에 가장 크게 寄與하는 두번째 主成分의 係數를 計算한다. 順次로 이와 同一한 要領을 써서 마지막 主成分까지의 係數를 計算할 수 있다.

以上이 主成分分析의 概要이다. 이것을 좀더 具體的으로 說明하기 위해서 G. Tintner의 業績⁽⁴⁾을 들어둔다.

그는 美國의 耐久財生産指數, 非耐久財生産指數, 鑛業生産指數 및 農業生産指數를 各各 X_1, X_2, X_3 , 및 X_4 로 表示하고 있다. 但 모든 指數는 1935~39年을 基準으로 한 1919~39年間의 것이다.

그리고 그가 計算해서 얻은 相關係數行列은 다음과 같다.

第 2.1 表	相關係數行列			
	X_1	X_2	X_3	X_4
X_1	1.000000	0.495941	0.872836	0.481240
X_2		1.000000	0.768279	0.709807
X_3			1.000000	0.712358
X_4				1.000000

따라서 上記한 (2.11)은 다음과 같이 된다. 즉 첫번째의 그리고 最大의 主成分의 係數를 附與하는 線形方程式體系는 다음과 같이 된다.

$$\begin{aligned}
 &1.000000k_{11} + 0.495941k_{21} + 0.872836k_{31} + 0.481240k_{41} = \lambda_1 k_{11} \\
 &0.495941k_{11} + 1.000000k_{21} + 0.768279k_{31} + 0.709807k_{41} = \lambda_1 k_{21} \\
 &0.872836k_{11} + 0.768279k_{21} + 1.000000k_{31} + 0.712358k_{41} = \lambda_1 k_{31} \\
 &0.481240k_{11} + 0.709807k_{21} + 0.712358k_{31} + 1.000000k_{41} = \lambda_1 k_{41}
 \end{aligned} \tag{2.11}'$$

이 (2.11)'은 上述한 바에서 알 수 있는 바와 같이 그것의 行列式이 0으로 될 때 解를 가질 수 있다. 즉 다음의 (2.12)'이 成立될 때 解를 가질 수 있다.

$$\begin{vmatrix}
 (1.000000 - \lambda_1) & 0.495941 & 0.872836 & 0.481240 \\
 0.495941 & (1.000000 - \lambda_1) & 0.768279 & 0.709807 \\
 0.872836 & 0.768279 & (1.000000 - \lambda_1) & 0.712358 \\
 0.481240 & 0.709807 & 0.712358 & (1.000000 - \lambda_1)
 \end{vmatrix} = 0 \tag{2.12}'$$

이 (2.12)'를 풀면 $\lambda_1 = 3.033424$ 가 된다. 이것은 (2.12)'의 最大根이다. 이 λ_1 을 (2.11)'

(4) G. Tintner, *Econometrics*, pp. 110—112(1953)

에 代入하고, $\sum_{i=1}^4 k_{i1} = \lambda_1 = 3.033424$ 를 考慮하면 첫번째 主成分의 係數가 얻어진다. 그들은 다음과 같다.

$$k_{11} = 0.817391, k_{21} = 0.888102, k_{31} = 0.951934, k_{41} = 0.818076$$

故로 첫번째 主成分의 標準化된 變數의 分散에 대한 寄與度는 다음과 같이 된다.

$$k_{11}^2 = 0.6724, k_{21}^2 = 0.7921, k_{31}^2 = 0.9063, k_{41}^2 = 0.6724$$

따라서 첫번째 主成分은 X_1 의 分散의 約 67%, X_2 의 分散의 約 79%, X_3 의 分散의 約 91%, 그리고 X_4 의 分散의 約 67%를 「說明한다」고 할수 있다.

III. 統計技術的 適用에 대하여

R. Stone⁽⁵⁾은 S. Kuznets 등이 作成한, 1922年에서 1938年까지의 美國의 國民所得및 支出關係의 17個 項目에 관한 資料에 대해서 主成分分析을 適用하여 3個의 主要因 다시 말하면 最大要因을 導出했다. 3個의 最大要因의 各各에 의해서 說明되는 各變數의 分散과

第 3.1 表		各要因에 의해서 說明되는 分散의 比率			
變 數	第 1 要因	第 2 要因	第 3 要因	殘 差	
賃金所得	0.9666	0.0229	0.0089	0.0016	
消費者非耐久財및 生産者耐久財購入	0.9068	0.0009	0.0851	0.0072	
企業純貯蓄+資本再評價	0.5018	0.4918	0.0006	0.0058	
消費者半耐久財및耐久財購入	0.9462	0.0087	0.0296	0.0155	
消費者서비스購入	0.8267	0.1215	0.0431	0.0087	
建築投資	0.7311	0.0051	0.2581	0.0057	
純公共支出	0.4992	0.0019	0.3367	0.1622	
純在庫增加	0.4230	0.1010	0.1093	0.3667	
在庫再評價調整	0.0889	0.7209	0.0252	0.1650	
個人純賃貸料受取	0.5467	0.0332	0.3961	0.0240	
企業家所得	0.8430	0.0558	0.0849	0.0163	
配 當 金	0.7184	0.1382	0.0590	0.0834	
減價償却	0.8270	0.1333	0.0066	0.0331	
利 子	0.0000	0.7486	0.0001	0.2513	
海外로부터의 配當所得등	0.2108	0.0001	0.0537	0.7354	
資本減耗充當	0.7078	0.0885	0.0074	0.1963	
海外收支	0.0447	0.0051	0.3883	0.5619	
合 計	0.8076	0.1059	0.0609	0.0256	

(1) G. Tintner, *Econometrics*, pp. 107-109(1953)

殘差分散 (residual variance)의 比率은 第 3.1 表에 表示한 바와 같다. 이에서 알 수 있는 바와 같이 各主要項目에 대해서는 3 個要因에 의해서 그 分散의 大部分(즉 모든 變數의 總分散의 97% 以上)이 說明되고 있다. 그러나 이에서 더 나가서 注目할만한 일은 이 3 個要因의 各各이 經濟學的으로 說明할 수 있는 것이라는 點이며 R. Stone은 그 中の 第 1 要因을 國民所得(或은 產出高), 第 2 要因을 國民所得의 變化率, 그리고 第 3 要因을 趨勢로 생각 하고 있다. 이와 같이 생각한 根據로서 그가 들고 있는 것은 다음과 같은 3 個의 相關係數이다. (第 3.2 表參照)

第 3.2 表

	第 1 要因	第 2 要因	第 3 要因
國民所得	0.995	-0.041	0.057
國民所得變化率	-0.056	0.948	-0.124
趨 勢	-0.369	-0.282	-0.836

이에서 그의 解釋의 妥當性을 納得할 수 있다. 이와 같이 第 3.1 表의 17 個項目이 實은 3 個要因에 의해서 說明될 수 있음이 明白하게 되었다.

G. Tintner는 이 結果에 대해서 統計學的으로 變數節約의 意味는 認定하면서도 經濟學的 意味의 存在에 대해서는 疑問을 表明하고 있다. 그러나 T. Kloek 및 L.B.M. Mennes의 業績은 이와 같은 R. Stone의 分析을 進一步시켜⁽⁶⁾서 經濟學的으로 하나의 새로운 意味를 附與한 것으로 看做되고 있다.

IV. 經濟發展論에의 適用에 대하여 (其 1)

M. Megee⁽⁷⁾는 두가지 目的에서 그의 論文을 쓰고 있다. 첫째 目的은 數學模型이 經濟成長의 研究에 있어서 가지는 有用性을 論하려는 것이며, 둘째 目的은 要因分析의 어떤 特殊한 適用에서 얻은 諸結果를 簡單히 分析하여 經濟成長分析和 將來의 成長計劃에서의 이方法(要因分析)의 使用이 갖는 長短點과 意義를 究明하려는 것이다. 그러나 이 中에서 本論稿와 關聯을 갖는 것은 둘째 目的이다. 따라서 여기서는 둘째 目的과 關聯된 部分만을 다룬다.

그는 우선 R 手法(R-technique)을 들고 있다. 이 手法은 數十年間 標準手法으로서 使用되어 오고 있는데 이 手法에서는 「變數에 대한 要因의 寄與度」만이 計算된다. 다음에 W.

(6) T. Kloek and L. B. M. Mennes, "Simultaneous Equations Estimation Based on Principal Components of Predetermined Variables," *Econometrica* (Jan. 1960)

(7) M. Megee, "On Economic Growth and the Factor Analysis Method," *The Southern Economic Journal* (Jan. 1965)

Stephenson 과 G.H. Thomson⁽⁸⁾ 이 1930年代에 導入한 Q手法(Q-technique)을 들고 있다. 이 手法에서는 「變數에 대한 要因의 寄與度」와 「項目 或은 個人에 대한 要因의 寄與度 或은 點數(score)」가 計算된다. 끝으로 M手法(M-technique)을 들고 있다. 이 手法은 時間에 따른 變化를 表示하기 위한 上記 두 手法의 補充物로서의 役割을 行하는데 그것은 異時點의 類似한(或은 相異한) 行列을 同時에 操作하는 것을 內容으로 하고 있다⁽⁹⁾.

그에 의하면 이들 세가지 手法은 經濟發展과 關聯된 세가지 假說의 檢定을 위한 手段으로서 使用된다고 한다. 즉 (1) R은 要因의 寄與度を 使用하여 經濟發展을 惹起시키는 變數에 관한 假說을 檢定할 수 있다. (2) Q는 要因의 點數를 使用하여 經濟發展類型的 地域의 格差에 관한 假說을 檢定할 수 있다. 그리고 (3) M은 時間에 따른 變數 或은 地域의 變化를 內包하는 假說을 檢定할 수 있다고 한다.

그가 要因分析을 통해서 얻은 結論은 다음과 같다. 그는 이 結論을 얻기 위해서 1人當 所得과 第3次産業에 雇傭되는 높은 比率이, 높은 經濟成長段階의 指標라는 것을 假定하고 있다. 즉 그는 第1次産業의 活動에 關聯을 갖고 있는 變數가 특히 美國에서는 經濟成長의 높은 段階를 說明하는 데 그다지 重要하지 않다는 것을 假定하고 있다. 다음의 諸表는 바로 이와 같은 假定下에 美國, 「멕시코」, 蘇聯, 및 世界全體에 대해서 行한 研究에서 얻어진 要因을 表示한다.

이들 表는 다음의 諸點을 示唆하고 있다.

(7) 비록 各 行列에서 相異한 變數들이 使用되어야 하긴 했지만 經濟의 3部門(第1次 第2次 및 第3次産業)은 研究對象이 된 모든 社會의 成長에서 強調되고 있다. 第3次産業 要因은 美國에서만 兩時期에 있어서 두번째로 큰 要因이 됨으로써 相對적으로 重要했다. 第1次産業部門은 美國經濟에서 1940年보다 1960年에 있어서 相對적으로 덜 重要했

美國의 要因들—1940	自乘의 合計, %	美國의 要因들—1960	自乘의 合計, %
1. 에너지 生産	26	1. 製造業과 建設	21
2. 所得과 第3次生産의 活動	11	2. 所得과 第3次産業의 活動	15
3. 厚生과 建設	9	3. 織物과 食糧	7
4. 農業과 敎育	7	4. 서비스, 人口增加, 農業, 化學製品	6
	53%		49%

行列의 크기 : 48×75

(8) W. Stephenson, *The Study of Behavior: Q-technique and Its Methodology* (1953).

G. H. Thomson, *Factor Analysis, Its Hopes and Dangers, Proceedings of the XIIIth International Congress of Psychology* (1951)

(9) 이 手法은 simultaneous programming of matrices 라고 불리우기도 한다.

멕시코의 要因들—1960	自乘의 總合計, %
1. 灌溉와 에너지	13
2. 에너지와 都市化	10
3. 化學製品	8
4. 教育支出과 資本支出	8
	39 %

行列의 크기 : 32×98

蘇聯의 要因들—1960	自乘의 總合計, %
1. 第1次産業品과 金屬製品	12
2. 非耐久消費者財	5
3. 動力과 建設	3
4. 織物과 動力	2
	22 %

行列의 크기 : 15×29

Berry의 世界經濟發展要因들	自乘의 總合計, %
1. 技術規模	84.2
2. 人口規模	4.2
3. 所得과 對外關係에 있어서의 差異	2.5
4. 大國과 小國	1.9
	92.8 %

行列의 크기 : 43×95

다. 그러나 勿論 第2次産業部門은 兩時期에 있어서 重要했다. 「에너지」生産, 農業改革및 灌溉事業과 關聯한 變數는 「멕시코」에서 第1次産業의 活動이 가장 크게 擴張되어야 한다는 것을 表示해주고 있다. 그러나 이들 要因은 生存型農業과 密接히 關係되는 後進國 特有的 農業을 強調하는 것이 아니라 新興都市人口를 相對로하는 經濟作物이나 輸出用 特用作物과 關聯있는 第1次産業部門의 先進된 分野의 開發을 意味하는 變數들로 構成되고 있음을 잊어서는 안된다. 蘇聯에서 導出된 要因은 第1次産業과 第2次産業의 製造業 특히 金屬製品製造業에 重點을 둔 3部門에서의 發展을 反映하고 있다.

(L) 美國에서는 4 要因이 1940년에는 自乘의 總合計(즉 總分散)의 53%를 그리고 1960년에는 49%를 說明하고 있다. 이 兩時期에 있어서의 差異는 1940年以後의 廣範한 技術의 進歩와 그에 따른 經濟의 多樣化에 基因한다.

다음에 「멕시코」에서는 4 要因이 自乘의 總合計의 39%를 그리고 蘇聯에서는 22%를 各各

說明하고 있다. 그런데 이 比率에서의 差異는 行列의 크기의 差異, 計算機型的 差異, 計算에 使用된 資料의 性質의 差異 및 變數의 크기의 差異에 基因한다고 한다.

끝으로 B. J. L. Berry 는 順位화된 資料(ranked data)를 使用해서 「技術規模」라는 單一要因이 自乘의 總合計의 84%나 說明하고 있음을 찾아냈다. 이數値는 行列의 크기를 考慮할 때 너무 큰 感이 드는데 그것은 標準화된 資料나 粗資料를 使用하지 않고 順位화된 資料를 使用한데 基因한다.

(ㄷ) 類似한 變數들이 두 時期以上에 걸쳐서 使用된 때에도 基礎를 이루는 要因들은 美國의 例에서 볼 수 있는 바와 같이 달라지는 것이 普通이다.

(ㄹ) 要因分析은 假定을 檢定하는데 有用하지만 B. J. L. Berry 와 R. B. Cattell 이 指摘한 바와 같이 이들 假定을 必要로 하지 않고 未知의 要因을 찾아내는데 있어서도 理想的인 方法으로 看做되고 있다. 經濟的 社會的 變數의 큰 資料行列을 構成하여 아무런 先驗的 假定없이 要因分析을 行하는 것은 可能的 일이다.

이어서 그는 經濟發展 및 計劃에 관한 研究를 위해서 使用하려고 할 때의 要因分析의 缺陷을 들고 있다. 그것은 다음과 같다.

(ㄱ) 經濟發展行列의 作成을 위해서 必要한 利用可能的 資料의 缺如: 이것은 특히 後進國에 있어서는 實質的으로 問題가 된다. 그러나 이것은 部分的으로는 資料가 缺如되어있는 地域에 대해서 變數値의 平均이나 中位數를 代身使用함으로써 克服할 수 있다. 이것은 또 研究對象 地域으로서 넓은 地域을 使用하지 않고 都市를 使用함으로써 克服할 수 있다. 그러나 地域的으로 廣範한 農村地域을 削除하게 되므로 一國의 經濟發展의 本質을 그릇 判斷하게 할 可能性이 있다. 따라서 두 個의 要因分析(都市에 관한 것과 農村에 관한 것)을 하는 것이 全體的 經濟發展의 本質把握을 위해서 보다 좋을 것이다.

(ㄴ) 主觀의 介入: 行列을 構成하는 變數의 選定, 要因을 構成하는 變數의 選擇, 要因의 指定, 經濟的 地域과 社會的 地域의 劃定, 座標軸回轉 或은 轉換法의 選定 및 要因分析法 自體의 選擇은 모두 어느程度의 主觀을 分析에 導入한다. 그러나 要因分析은 보다 널리 使用되고 있는 重(或은 多元) 回歸分析보다도 變數의 獨立性和 從屬性에 관한 先驗的 假定같은 制約을 덜 받고있다.

(ㄷ) 地域間 및 地域內 構造와 「플로우」의 表示不能: 이 點이 「리니어 프로그래밍」, 產業聯關分析 및 其他의 多變量分析과 다르다. 그러나 如斯한 「플로우」는 Q 手法에서 誘導되는 類型的 地域의 分布를 觀察함으로써 推測해 낼 수 있다. 例컨대 製造業部門의 點數가 높은 地域은 工產品供給地域을 나타낼 수 있으며 낮은 地域은 工產品需要地域을 나타낼 수 있다.

더 나아 가서 그는 要因分析의 用途를 밝히고 있다. 그에 의하면 그것은 다음과 같다. 社會科學에 있어서의 要因分析의 가장 重要的 用途는 研究의 初期段階에서 더 研究를 進行하기 위하여 重要的 或은 戰略的인 變數를 分離하며 또 別로 關係가 없는 情報를 除去하는 것이다. 그 然後에 問題에 관한 보다 複雜한 假說은 單 方法으로 檢定되기 위해서 設定될 수 있다. 要因分析模型은 經濟問題를 다루기 위해서 構成되는 다른 通常的인 模型보다 훨씬 많은 變數를 包含할 수 있는 매우 複雜한 模型이다. 그것은 經濟의 어떤 部門의 相對的 重要度를 確認하고 測定하기 위해서 使用될 수 있으며 또 地域格差를 나타내기 위해서 使用될 수 있다.

끝으로 그는 經濟發展과 計劃에 관한 研究를 위해서 要因分析을 使用하려 할 때 앞으로 더 行하여야 할 일을 다음과 같이 들고 있다.

- (ㄱ) 經濟成長分析에서의 變數의 重要度(weight)를 찾아내기 위한 行列의 構成.
- (ㄴ) 經濟成長 및 成長의 先進, 中進 및 後進經濟에서 類似한 要因을 包含하고 있는 程度를 考察하기 위한 比較研究의 構想.
- (ㄷ) 各種의 經濟活動의 地方化를 確認하기 위한 都市에 관한 經濟發展行列의 構成.
- (ㄹ) 보다 많은 政治的 社會的 및 心理的 變數의 經濟成長研究에의 導入.
- (ㄹ) 要因의 性格에 있어서의 있을 수 있는 重要的 變化를 보기 위한 어떤 研究에서 다루어지는 變數의 種類와 數의 變更.

V. 經濟發展論에의 適用에 대하여 (其2)

I. Adelman 및 C.T. Morris,⁽¹⁰⁾는 各種의 社會的 政治的 變化와 經濟開發水準의 相互作用에 관한 半數量的인 洞察力을 얻을 目的에서 그들의 論文을 쓰고 있다. 그들은 이 目的을 위해서 1957~62年間의 74個 後進國에 관한 1人當 所得과 社會的 政治的 構造를 나타내는 많은 指數에 要因分析技術을 適用하고 있다. 그러나 그들은 廣汎한 經濟開發水準과 工業化 및 都市化와 關聯된 社會的 政治的 機構 및 文化價値의 變革間의 相互依存關係의 性質을 分析하기 위해서(1人當 所得 以外의) 純粹經濟變數를 除去하고 있다.

그들은 要因分析을 行하기 위해서 우선 다음의 22個의 社會的 政治的 變數(或은 指標)를 選擇하고 있다.

- ① 傳統的인 農業部門의 크기
- ② 基礎的 社會組織의 性格
- ③ 教育水準
- ④ 「매스콤」의 範圍
- ⑤ 文化의 人種的 同質性의 程度
- ⑥ 土着中間階級の 重要性
- ⑦ 將來展望의 近代化度

(10) I. Adelman and C. T. Morris, "Factor Analysis of Interrelationship between Social and Political Variables and Per Capita Gross National Product", *The Quarterly Journal of Economics*, (Nov. 1965)

⑧ 社會的 移動性的 程度 ⑨ 民主的 機構의 有效性 ⑩ 野黨과 言論의 自由度 ⑪ 政黨의 派黨化度 ⑫ 政黨制度의 基礎 ⑬ 勞動運動의 強力度 ⑭ 軍隊의 政治的 強力度 ⑮ 行政能率의 程度 ⑯ 政治力의 集中度 ⑰ 傳統的인 「엘리트」의 強力度 ⑱ 民族主義와 民族統一意識의 程度 ⑲ 指導者의 經濟開發意欲의 程度 ⑳ 政府의 經濟活動에의 關與度 ㉑ 社會的緊張의 程度 ㉒ 政治制度의 安定度. 여기서 ②의 基礎的 社會組織의 性格은 尙 分化된 親族關係(즉 部族, 種族 및 大家族)이나 核家族(즉 父母子女의 單位)이나를 意味하며, ③의 教育水準은 教育의 近代化度를 意味하며 ⑫의 政黨制度의 基礎는 「이데올로기」的인 것 에 基礎를 두거나 或은 個性的인 것에 基礎를 두는 多元政黨制度나 大衆을 指導하거나 或은 그것에 基礎를 두는 單一政黨制度나를 意味한다. 그런데 높은 1人當 GNP는 ②의 경우에는 尙 分化된 親族關係에서 核家族으로의 分化와, ⑫의 경우에는 單一政黨制度에서 多元政黨制度로의 移動과 關聯되어 있다.

다음에 前記한 22個의 變數를 다음의 세가지로 分類하고 있다.

- (ㄱ) 公表된 統計에 全的으로 依存하는 것.
- (ㄴ) 統計的인 要素와 定性的인 要素를 結合할 必要가 있는 것.
- (ㄷ) 本質에 있어서 純全히 定性的인 것.

그리고서 이 分類에 의거해서 74個國을 分類하고 있다.

셋째로 22個의 變數 하나 하나 別로 74個國의 各各에 A, A-, B, B+ 등과 같은 文字點數(letter score)를 附與하고 있다.

끝으로 이를 文字點數에 다음과 같은 簡單한 線形尺度法을 使用하여 數字點數(numerical score)를 附與하고 있다.

〔分類範圍가 A에서 C까지〕		〔分類範圍가 A에서 D까지〕	
일배의 點數배기는 要領		일배의 點數배기는 要領	
文字點數	數字點數	文字點數	數字點數
A+	100	A+	97
A	90	A	90
A-	80	A-	83
B+	60	B+	67
B	50	B	60
B-	40	B-	54
C+	20	C+	41
C	10	C	35
C-	1	C-	29
		D+	16
		D	10
		D-	4

그들은 以上の 要領을 거쳐 要因分析을 行했지만 實地의 要領은 밝히지 않고 要因分析의 結果만을 提示하고 있다⁽¹¹⁾. 그들의 結果는 要因行列 或은 要因寄與度行列 (factor loadings matrix)이라고 불리우는 第 5.1 表에 要約되어 있다.

이 表의 數值 즉 一般의 表示해서 a_{ij} 는 要因 j 의 社會的 政治的 指標 i 에 대한 影響의 重要性을 나타낸다. 보다 더 具體的으로 말하면 要因寄與度라고 불리우는 이 a_{ij} 는 各要因과 觀察된 變數間의 純相關을 나타낸다. 이 要因寄與度는 이 表의 數值의 自乘을 使用하면 그 解釋이 보다 容易해진다. 各 $(a_{ij})^2$ 은 다른 要因의 寄與度를 考慮한 後의 要

第 5.1 表 回轉된 要因行列 (74 個의 後進國)

政治的 社會的 指標	F ₁	回轉된 F ₂	要因寄與度 F ₃	F ₄	h _i ² (R ²)
1. 1人當 GNP	0.64	0.43	-0.22	0.12	0.661
2. 傳統的인 農業部門의 크기	-0.83	-0.32	0.21	0.03	0.832
3. 基礎的 社會組織의 性格	0.86	0.26	-0.05	-0.04	0.819
4. 教育水準	0.83	0.37	-0.08	0.09	0.840
5. 매스콤의 範圍	0.85	0.34	-0.08	0.04	0.843
6. 文化的人種的 同質性의 程度	0.72	-0.30	0.07	0.17	0.646
7. 土着中間階級의 重要性	0.63	0.31	-0.40	0.05	0.658
8. 將來展望의 近代化度	0.64	0.46	-0.33	0.13	0.753
9. 民主的機構의 有效性	0.36	0.79	-0.17	0.19	0.815
10. 野黨과 言論의 自由度	0.27	0.86	0.04	0.08	0.829
11. 政黨의 派黨化度	0.33	0.78	0.07	-0.30	0.810
12. 政黨制度의 基礎	0.42	0.68	-0.06	-0.11	0.657
13. 勞動運動의 强力度	0.29	0.71	-0.34	0.08	0.715
14. 軍隊의 政治的强力度	0.38	-0.58	0.29	0.36	0.690
15. 行政能率의 程度	0.30	0.53	-0.48	0.20	0.636
16. 政治力의 集中度	-0.03	-0.76	0.20	-0.07	0.617
17. 社會的移動性의 程度	0.42	0.14	-0.55	0.24	0.549
18. 傳統的인 엘리트의 强力度	-0.03	0.15	0.82	-0.04	0.700
19. 民族主義와 民族統一意識의 程度	0.61	-0.05	-0.57	-0.01	0.694
20. 指導者의 經濟開發意欲의 程度	0.11	0.21	-0.75	0.29	0.696
21. 政府의 經濟活動에의 關與度	0.27	-0.41	-0.48	-0.41	0.638
22. 社會的緊張의 程度	-0.23	0.03	0.08	-0.84	0.771
23. 政治制度의 安定度	0.04	0.08	-0.24	0.86	0.808

(11) 그들은 要因分析에 대한 參考文獻으로서 H.H. Harman, *Modern Factor Analysis*(1960)과 L. L. Thurstone, *Multiple Factor Analysis*(1961)을 推薦하고 있다.

(12) $(0.64)^2 \times 100\% = 41.5\%$, $(0.43)^2 \times 100\% = 18.5\%$, $(0.22)^2 \times 100\% = 4.8\%$, $(0.12)^2 \times 100\% = 14\%$

(13) 重相關係數의 自乘을 말한다.

인 j 에 의해서 설명되는 변수 i 의 총단위분산의 비율을 나타낸다. 이 표의 第1行을 보면 1人當의 GNP의 國家間變動의 41.5%⁽³⁾가 要因 I 즉 F_1 에 의해서 그것의 18.5%가 要因 II 즉 F_2 에 의해서, 그것의 4.8%가 要因 III 즉 F_3 에 의해서 그리고 그것의 1.4%가 要因 IV 즉 F_4 에 의해서 설명됨을 알 수 있다.

이 표의 最右側列은 各變數의 要因寄與度의 自乘의 合計 或은 共通度(communality)를 表示하고 있다. 이 共通度는 共通要因 全部에 의해서 설명되는 總單位分散의 비율을 나타내며 따라서 回歸分析의 R^2 ⁽⁴⁾과 類似하다. 例컨대 1人當 GNP의 共通度는

$$(0.64)^2 + (0.43)^2 + (-0.22)^2 + (0.12)^2 = 0.661$$

이다. 즉 1人當 GNP의 國家間變動의 66%가 22個의 社會的 政治的 變數에서 抽出된 4個의 共通要因과 關聯되어 있다.

要因寄與度行列은 觀察된 變數를 說明하는데 있어서의 各 要因의 重要度(weight)를 나타내는 外에 變數를 共通要因으로 「그룹」化하는 基礎를 提供한다. 第 5.1 表는 우선 要因 I에서 가장 큰 寄與度를 갖는 指標를 들고 다음에 要因 II, 要因 III 및 要因 IV에서 가장 큰 寄與度를 갖는 指標를 차례로 들고 있다. 「고직」體數字는 各指標가 配屬되어있는 要因에서의 寄與度를 나타낸다.

一旦 變數를 共通要因에 配屬시키면 이 要因의 「識別」이 必要하게 된다. 이제 各 要因을 識別한 것을 보면 다음과 같다.

要因 I (F_1)

이 要因은 앞에서 본바와 같이 總單位分散의 41.5%를 차지하고 있다. 즉 要因 I은 1人當의 GNP의 國家間變動의 41.5%를 說明하고 있다. ②에서 ⑧까지의 指標를 包括하고 있는 데서 알 수 있는 바와 같이 이 要因은 都市化 및 工業化에 따르는 社會的 文化的 變化를 나타낸다. 좀더 具體的으로 말하면 이 要因은 傳統的인 社會組織의 崩壞와 關聯된 態度 및 機構에 있어서의 變化過程을 나타낸다.

「고직」體數字에서 알 수 있는 바와 같이 (1) ③의 基礎的 社會組織의 性格, ④의 教育水準, ⑤의 「메스콤」의 範圍, ⑥의 言語的 同質性의 程度, ⑦의 土着中間階級의 重要性 및 ⑧의 將來展望의 近代化度는 1人當 GNP와 順相關의 關係에 있다. 그리고 (2) ②의 傳統的인 農業部門의 크기는 1人當 GNP와 逆相關의 關係에 있다. 즉 보다 작은 傳統的인 農業部門의 크기, 보다 큰 核家族으로의 分化度, 보다 높은 教育水準, 보다 큰 「메스콤」의 範圍, 보다 큰 言語的 同質性의 程度, 보다 큰 土着中間階級의 重要性 및 보다 큰 將來展望의 近代化도가 보다 높은 1人當 GNP와 關聯되어 있다.

要컨대 要因 I은 經濟開發과 社會行動, 價値 및 機構의 合理化度間에 強한 相互作用이 있음을 表示한다.

要因 II (F₂)

이 要因은 앞에서 본 바와 같이 總單位分散의 18.5%를 차지하고 있다. 즉 要因 II 는 1 人當 GNP의 國家間變動의 18.5%를 說明하고 있다. ⑨에서 ⑯까지의 指標를 包括하고 있는데서 알 수 있는 바와 같이 이 要因은 政治制度에 있어서의 國家間變動을 나타낸다.

「고직」體數字에서 알 수 있는 바와 같이 (7) ⑨의 民主的 機構의 有效性, ⑩의 野黨과 言論의 自由度, ⑪의 政黨의 派黨化度, ⑫의 政黨制度의 基礎, ⑬의 勞動運動의 強力度, 및 ⑭의 行政能率의 程度는 1 人當 GNP 와 順相關의 關係에 있다. 그리고 (2) ⑭의 軍隊의 政治的 強力度, 및 ⑮의 政治力의 集中度는 1 人當 GNP 와 逆相關의 關係에 있다. 즉 보다 큰 民主的 機構의 有效性, 보다 큰 野黨과 言論의 自由度, 보다 큰 政黨의 派黨化度, 보다 큰 多元政黨制度로의 移動, 보다 큰 勞動運動의 強力度, 보다 큰 行政能率의 程度, 보다 작은 軍隊의 政治的 強力度 및 보다 작은 政治力의 集中度가 보다 높은 1 人當 GNP 와 關聯되어 있다.

要因 III (F₃)

이 要因은 앞에서 본 바와 같이 總單位分散의 4.8%를 차지하고 있다. 즉 要因 III 은 1 人當 GNP 의 國家間變動의 4.8%를 說明하고 있다. ⑰에서 ㉑까지의 指標를 包括하고 있는데 알 수 있는 바와 같이 이 要因은 指導者의 性格과 指導者의 戰略의 性質을 나타낸다.

「고직」體數字에서 알 수 있는 바와 같이 (7) ⑰의 社會的 移動性의 程度, ⑱의 民族主義와 民族의 統一意識의 程度, ㉒의 指導者의 經濟開發意欲의 程度 및 ㉓의 政府의 經濟活動에의 關與度는 1 人當 GNP 와 順相關의 關係에 있다⁽¹⁴⁾. 그리고 (2) ⑱의 傳統的인 「엘리트」의 強力度는 1 人當 GNP 와 逆相關의 關係에 있다. 즉 보다 큰 社會的 移動性의 程度, 보다 큰 民族主義와 民族統一意識의 程度, 보다 큰 指導者의 經濟開發意欲의 程度, 보다 큰 政府의 經濟活動에의 關與度 및 보다 적은 傳統的인 「엘리트」의 強力度가 보다 높은 1 人當 GNP 와 關聯되어 있다.

要因 IV (F₄)

이 要因은 앞에서 본 바와 같이 總單位分散의 1.4%를 차지하고 있다. 즉 要因 IV 는 1 人當 GNP 의 國家的 變動의 1.4%를 說明하고 있다. ㉔와 ㉕의 指標를 包括하고 있는데서 알 수 있는 바와 같이 이 要因은 社會的 政治的 安定性의 性格을 나타낸다.

「고직」體數字에서 알 수 있는 바와 같이 (7) ㉔의 政治制度의 安定度는 1 人當 GNP 와 順相關의 關係에 있다. 그리고 (2) ㉕의 社會的 緊張의 程度는 1 人當 GNP 와 逆相關의 關係

(14) —0.22 와 對照하여 생각하라.

에 있다. 다시 말하면 社會的 安定的 程度는 1人當 GNP와 順相關의 關係에 있다. 즉 보다 큰 政治制度의 安定度와 社會的인 安定的 程度가 보다 높은 1人當 GNP와 關係되어 있다. 以上은 要因分析을 74個의 後進國에 適用시켜서 얻은 結果이다. 그러나 그들은 이에서 더 나가서 74個國을 「아프리카」, 近東 및 極東, 「라틴아메리카」의 3個地域 「그룹」으로 묶고 이 3個「그룹」에 要因分析을 適用시키고 있다. 그 結果는 다음과 같다.

아프리카(第 5.2 表參照)

「아프리카」에 대한 要因分析의 結果는 74個國의 그것과 매우 類似하다. 이 結果의 重要한 特徵은 經濟開發水準의 國家間變動의 매우 큰 比率이 傳統的인 農業部門의 크기, 教育水準, 「메스콤」의 範圍 및 將來展望의 近代化度를 包括하는 要因I에 의해서 說明되고 있는 點이다. 이 地域에서는 要因I이 1人當 GNP의 國家間 變動의 3/4 以上⁽¹⁵⁾을 說明하고 있다. 이것은 大部分의 「아프리카」國家들이 아직도 傳統的인 社會構造가 成長의 致命

第 5.2 表 回轉된 要因行列
(27個의 아프리카諸國)

政治的 社會的 指標	F ₁	回轉된 要因寄與度			h ² (R ²)
		F ₂	F ₃	F ₄	
1. 1人當 GNP	-0.88	0.04	0.02	0.13	0.800
2. 傳統的인 農業部門의 크기	0.73	-0.20	-0.36	-0.31	0.796
3. 教育水準	- 0.67	0.43	0.21	0.21	0.724
4. 메스콤의 範圍	- 0.93	0.17	0.01	-0.17	0.929
5. 將來展望의 近代化度	- 0.71	0.22	0.30	0.36	0.775
6. 民主的機構의 有效性	-0.27	0.83	-0.14	0.02	0.776
7. 野黨과 言論의 自由度	0.08	0.81	-0.26	-0.12	0.739
8. 政黨의 派黨化度	-0.24	0.85	-0.04	0.33	0.888
9. 政黨制度의 基礎	-0.12	0.64	-0.30	0.15	0.532
10. 勞動運動의 强力度	-0.20	0.75	0.04	0.02	0.596
11. 行政能率의 程度	-0.31	0.56	0.40	0.18	0.598
12. 政治力의 集中度	-0.36	- 0.66	-0.16	0.08	0.593
13. 傳統的인 엘리트의 强力度	-0.39	-0.08	- 0.75	0.15	0.745
14. 社會的移動性의 程度	0.15	0.24	0.47	-0.43	0.485
15. 民族主義와 民族統一意識의 程度	-0.22	-0.26	0.68	-0.01	0.576
16. 指導者의 經濟開發意欲의 程度	-0.22	-0.07	0.76	-0.10	0.639
17. 政府의 經濟活動에의 關與度	-0.25	-0.18	0.51	0.08	0.357
18. 軍隊의 政治的强力度	-0.43	-0.33	-0.01	0.61	0.675
19. 政治制度의 安定度	0.02	-0.15	0.19	- 0.83	0.743
20. 社會的緊張의 程度	-0.28	0.17	-0.04	0.84	0.817
21. 文化的人種의 同質性의 程度	-0.39	-0.16	-0.16	- 0.50	0.456

(15) (0.88)²=0.7744 이다.

의인 障害要因이 되고 있는 「테이크 오프」의 前段階에 놓여 있음을 말해주고 있다.

이에 反하여 要因II와 III이 說明하는 比率는 無視할 수 있을 程度이며 要因IV가 說明하는 比率는 2% 以下이다⁽¹⁶⁾. 그런데 이 II, III, IV의 3個要因이 說明하는 比率가 매우 낮은 것은 「아프리카」國家들이 아직도 最小限의 社會的 機構의 變革를 經驗하지 못했음을 立證하고 있다. 「아프리카」에 대한 要因分析의 結果의 또 다른 重要的 特徵은 要因VI의 性格이다. 이 要因에 있어서 보다 높은 1人當 GNP가 보다 深刻한 社會的 緊張, 보다 작은 政治的 安定, 보다 強한 軍隊의 影響力 및 보다 낮은 文化的 同質性과 關聯되어 있다. 「아프리카」에 있어서의 短期的인 社會的 政治的 安定性과 所得水準間의 逆相關은 傳統的인 社會構造와 規範이 崩壞되며 變革되고 있는 經濟開發의 初期段階에서 볼 수 있는 緊張과 不安에 一部 基因한다.

近東 및 極東(第 5.3 表參照)

近東 및 極東에 대한 要因分析의 結果는 74 個國의 그것과 매우 類似하다. 앞에서와 마찬가지로 要因 I은 經濟開發水準의 差異와 關聯된 가장 重要的 要因이다⁽¹⁷⁾. 이 경우에는 要因 I 속에 社會的 移動性의 程度가 包含되어 있는데 이것은 論理的이라고 할 수 있다. 왜냐하면 要因 I은 工業化에 따르는 社會的 文化的 變化를 나타내기 때문이다. 事實 前近代의인 社會構造의 崩壞는 移動을 自由스럽게 만든다.

要因II는 「아프리카」의 경우와는 對照的으로 總單位分散의 큰 比率를 차지하고 있으며(즉 1人當 GNP의 國家間變動의 큰 比率를 說明하고 있으며), 또 74 個國의 경우 보다도 그 比率가 약간 크다⁽¹⁸⁾. 이것은 이 地域의 經濟成長이 「아프리카」의 경우 보다도 傳統的인 社會構造의 制約을 덜 받고 있는데 基因한다

要因 III은 1人當 GNP의 國家間變動의 2%⁽¹⁹⁾를 說明하고 있다. 但 이 要因에는 中央政府의 行政力을 特徵짓는 3個의 指標外에 社會的 政治的 安定性을 나타내는 2個의 指標가 包含되어있다. 「고직」體數字에서 알 수 있는 바와 같이 (7) ⑩의 將來展望의 近代化度, ⑪의 指導者의 經濟開發意欲의 程度, ⑫의 政治制度의 安定度, ⑬의 行政能率의 程度는 1人當 GNP와 順相關의 關係에 있다. 그리고 (8) ⑭의 社會的 緊張의 程度는 1人當 GNP와 逆相關의 關係에 있다. 즉 보다 큰 將來展望의 近代化度, 보다 큰 指導者의 經濟開發意欲의 程度, 보다 덜 한 社會的 緊張, 보다 큰 政治制度의 安定 및 보다 나은 行政能率이 보다 높은 1人當 GNP와 關聯되어 있다.

要因IV는 1人當 GNP의 國家間變動의 매우 작은 比率(約 0.1%⁽²⁰⁾)만을 說明하고 있

(16) $(0.04)^2 \times 100\% = 0.2\%$, $(0.02)^2 \times 100\% = 0.0\%$, $(0.13)^2 \times 100\% = 1.7$ 이다.

(17) 그 比率는 $(0.72)^2 \times 100\% = 52.0\%$ 이다.

(18) 그 比率는 $(0.47)^2 \times 100\% = 22.0\%$ 이다.

(19) $(0.14)^2 \times 100\% = 2.0\%$ 이다.

(20) $(-0.03)^2 \times 100\% = 0.1\%$ 이다.

第 5.3 表

回轉된 要因行列
(25個의 中東 및 近東諸國)

政治的社會的 指標	F ₁	回轉된 F ₂	要因寄與度 F ₃	F ₄	hi ² (R) ²
1. 1人當 GNP	0.72	0.47	-0.14	-0.03	0.762
2. 傳統的인 農業部門의 크기	-0.77	-0.41	0.17	-0.29	0.881
3. 基礎的社會組織의 性格	0.58	0.35	0.04	0.58	0.793
4. 教育水準	0.64	0.38	-0.15	0.29	0.655
5. 매스콤의 範圍	0.82	0.33	-0.03	0.25	0.850
6. 文化의 人種의 同質性의 程度	0.79	-0.43	-0.11	-0.15	0.831
7. 土着中間階級의 重要性	0.70	0.32	-0.09	0.37	0.747
8. 社會的移動性의 程度	0.77	-0.03	-0.16	0.38	0.765
9. 民主的機構의 有效性	0.27	0.85	-0.26	0.25	0.926
10. 野黨과 言論의 自由度	0.31	0.88	-0.02	0.25	0.935
11. 政黨의 派黨化度	0.28	0.58	0.32	0.54	0.800
12. 勞動運動의 強力度	0.28	0.63	-0.11	0.39	0.639
13. 軍隊의 政治的強力度	-0.04	-0.81	0.14	-0.08	0.688
14. 政治力의 集中度	-0.11	-0.87	-0.07	-0.18	0.810
15. 政府의 政治活動에의 關與度	-0.00	-0.71	0.18	0.34	0.654
16. 將來展望의 近代化度	0.40	0.52	-0.52	0.43	0.882
17. 指導者의 經濟開發意欲의 程度	0.08	0.09	-0.73	0.49	0.781
18. 社會的 緊張의 程度	-0.23	0.01	0.81	0.10	0.714
19. 政治制度의 安定度	-0.01	0.19	-0.92	-0.05	0.883
20. 行政能率의 程度	0.47	0.33	-0.50	0.51	0.851
21. 民族主義와 民族統一意識의 程度	0.48	-0.19	-0.48	0.61	0.871
22. 傳統的인 엘리트의 強力度	-0.51	-0.13	0.09	-0.60	0.650
23. 政黨制度의 基礎	0.18	0.52	0.00	0.60	0.658

다⁽²¹⁾.

라틴 아메리카(第 5.4 表參照)

「라틴 아메리카」에 대한 要因分析의 結果는 3 個地域「그룹」의 가장 典型的인 地域的 類型을 나타낸다.

要因I은 역시 1人當 GNP의 國家間 變動의 큰 比率(約 38%)⁽²²⁾을 說明하고 있다.

要因II는 近東 및 極東地域의 경우와 거의 같은 比率(約 21%)을 說明하고 있다. 그러나 그것이 包括하고 있는 內容에 特色이 있다. 그것은 74 個國의 경우의 要因II에 包含되

(21) 여기서는 省略했지만 要因寄與度에 대한 그들의 解釋에는 잘못이 있는 것 같다. I. Adelman and C. T. Morris, *ibid.* p. 575 參照

(22) $(0.62)^2 \times 100\% = 38.4\%$ 이다. 마찬가지로 해서 $(0.46)^2 \times 100\% = 21.2\%$, $(-0.32)^2 \times 100\% = 10.2\%$, $(-0.13)^2 \times 100\% = 1.7\%$ 이다.

어 있는 5개의 指標 즉 民主的 機構의 有效性, 勞動運動의 強力度, 行政能率의 程度, 軍隊의 政治的 強力度 및 政治力의 集中度 뿐 아니라 指導者의 特徵과 社會的 政治的 安定性을 나타내는 指標를 包含하고 있다.

要因III은 1人當 GNP의 國家間 變動의 約 10%를 說明하고 있다. 「고직」體數字에서 알 수 있는 바와 같이 ⑮의 野黨과 言論의 自由, ⑰의 政黨의 派黨化度 및 ⑳의 政黨制度의 基礎는 1人當 GNP와 順相關의 關係에 있다.

要因IV는 1人當 GNP의 國家間 變動의 約 1.7%를 說明하고 있다. 「고직」體數字에서 알 수 있는 바와 같이 ㉑의 民族主義와 民族統一意識의 程度와 ㉒의 政府의 經濟活動에의 關與度는 1人當 GNP와 順相關의 關係에 있다. 즉 보다 強力한 民族主義와 보다 큰 政府의 經濟活動에의 關與度가 보다 높은 1人當 GNP와 關聯되어 있다.

第 5.4 表

回轉된 要因行列
(21 個의 라틴 아메리카諸國)

政治的 社會的 指標	F ₁	回轉된 F ₂	要因寄與度 F ₃	F ₄	hi ² (R) ²
1. 1人當 GNP	0.62	0.46	-0.32	-0.13	0.706
2. 傳統의인 農業部門의 크기	-0.81	-0.29	0.08	0.08	0.760
3. 教育水準	0.87	0.24	-0.09	0.17	0.856
4. 매스콤의 範圍	0.71	0.23	-0.41	-0.18	0.761
5. 文化的 人種의 同質性의 程度	0.80	-0.26	0.15	-0.14	0.753
6. 土着中間階級의 重要性	0.59	0.32	-0.24	-0.26	0.581
7. 民主的 機構의 有效性	0.52	0.68	-0.27	-0.02	0.814
8. 勞動運動의 強力度	0.23	0.79	-0.08	-0.28	0.771
9. 軍隊의 政治的 強力度	-0.17	-0.70	0.24	0.03	0.577
10. 行政能率의 程度	0.57	0.60	-0.38	-0.04	0.862
11. 政治力의 集中度	-0.12	-0.70	0.31	0.41	0.770
12. 將來展望의 近代化度	0.55	0.70	-0.05	-0.20	0.838
13. 社會的 移動性의 程度	0.35	0.74	0.36	-0.04	0.808
14. 傳統의인 엘리트의 強力度	-0.07	-0.87	-0.22	0.17	0.841
15. 指導者의 經濟開發意慾의 程度	0.05	0.83	-0.27	-0.03	0.770
16. 政治制度의 安定度	0.52	0.51	0.15	0.37	0.688
17. 社會的 緊張의 程度	-0.37	-0.46	-0.26	-0.48	0.647
18. 野黨과 言論의 自由度	0.45	0.27	-0.72	0.07	0.799
19. 政黨의 派黨化度	0.07	-0.07	-0.82	0.05	0.677
20. 政黨制度의 基礎	0.00	0.41	-0.52	-0.48	0.671
21. 民族主義와 民族統一意識의 程度	0.30	0.19	-0.06	-0.82	0.808
22. 政府의 經濟活動에의 關與度	0.02	0.19	0.19	-0.78	0.681

끝으로 그들은 以上の 結果를 要約하고 있는데 끝 部分에 가서 特히 經濟開發水準과 社會的 政治的 構造의 差異間에서 發見된 關係가 결코 因果關係가 아니라는 것을 強調하고 있다.

VI. 結 言

M. Megee가 밝히고 있는 바와 같이 要因分析은 要因行列의 作成에 必要한 利用可能한 資料의 缺如, 主觀의 介入, 地域間 및 地域內 構造와 「플로우」의 表示不能 등의 缺陷을 갖고 있다. 또 要因分析을 經濟發展과 計劃에 관한 研究에 適用시키려면 앞으로 行하여야 할 일이 많다. 즉 經濟成長分析에서의 變數의 重要度を 찾아내기 위한 行列의 構成, 보다 많은 政治的 社會的 및 心理的 變數의 經濟成長研究에의 導入, 變數의 種類와 數의 變更 등을 行하여야 한다. 이 밖에도 要因分析은 計算의 複雜이라는 缺陷을 갖고 있다. 이 點에서 要因分析은 計算機의 制約을 받는다.

그러나 이 要因分析은 亦是 M. Megee가 밝히고 있는 바와 같이 앞으로 더 研究를 進行하기 위해서 重要한 或은 戰略的인 變數를 分離하며 또 別로 關係가 없는 情報를 除去하는 일을 可能케 해준다. 이것이 바로 社會科學에 있어서의 要因分析의 가장 重要한 用途이다. 따라서 要因分析은 앞으로의 研究를 위한 準備作業으로서의 意義도 갖는다고 할 수 있다. 그리고 要因分析模型은 經濟問題를 다루기 위해서 構成되는 다른 通常的인 模型보다 훨씬 많은 變數를 包含할수 있는 매우 複雜한 模型이다.

우리나라에서도 점차로 計量經濟學의 研究가 行해지게 되었다. 이 研究를 위해서는 무엇보다도 優先해서 模型構成(model building)이 行해져야 한다. 要因分析의 用途가 上述한 바와 같은 것이라면 이 模型構成의 過程에서 要因分析은 많은 示唆를 줄 수 있을 것이다. 그리고 要因分析을 行하여 構成된 模型을 使用해서 forecasting을 行하는 편이 보다 더 近似한 豫測值를 賦與하지 않을까? 이런 意味에서 우리나라에서도 이 要因分析에 관한 本格的인 研究를 行할 必要가 있을 것이다.

또 우리나라에서는 漸次로 GNP 或은 1人當 GNP에 대해서 많은 關心을 갖게 되었다. 그러나 아직 I. Adelman 및 C. T. Morris의 그것과 같은 試圖가 行해지지 않고 있다. 非經濟的 要因 즉 政治的 社會的 要因이 얼마만큼 1人當 GNP의 提高에 寄與하는가를 究明하는 일은 매우 意義있는 일이라고 할 수 있다. 이 일은 더 나가서 計量經濟學의 研究에 있어서 非經濟的 要因을 如何히 다룰 것인가에 대해서 어떤 示唆를 줄 수도 있을 것이다. 이런 意味에서도 要因分析에 대한 本格的인 研究는 우리나라에서 要請된다고 하겠다.

〔筆者 서울大學校商科大學
韓國經濟研究所 研究員
서울大學校商科大學 教授〕

<Summary>

The Application of Factor Analysis in Economics

Hyung-Yoon Byun*

I. Preface

Factor analysis method, which is an applied branch of mathematics, was devised by J. C. Spearman in 1904, and grew up in this century with the field of psychology. It is increasingly being adopted by other fields, such as economics and sociology. In economics it is applied in economic growth or economic development theory.

The present author is especially interested in the latter applications. But before examining the core of the problem, it may well examine factor analysis itself. And R. Stone's 1947 paper, which deals with the application of the factor analysis to statistical technique, and which is the basis of T. Kloeck and L.B.M. Mennes' most advanced paper of the field, is useful for understanding the implication of factor analysis method.

In this paper, therefore, I introduce factor analysis first, examine Stone's paper next, and then M. Megee's paper and I. Adelman and C. T. Morris' paper on the applications of factor analysis to economic development theory.

II. Factor Analysis

Factor analysis method is applied to economics in the form of principal component analysis which was devised by Hotelling to deal with the problem of factor analysis in psychology. By principal component or factor analysis we try to answer the question: Is it possible to analyze a set of variables into a more fundamental set of independent components (or factors) possibly fewer in number? Which portion of the total variance can be accounted for by each component? Here, Hotelling's method of solution is introduced.

Assume that we want to replace a set of standardized variables z_i ($i=1, 2, \dots, p$) by a more fundamental set of variables u_i ($i=1, 2, \dots, p$), such that

$$\mathbf{z} = \mathbf{K}\mathbf{u},$$

where \mathbf{z} is a p -component column vector, \mathbf{K} is a $p \times p$ matrix of coefficients, and \mathbf{u} , a p -component column vector. The components u_i are principal components and are orthogonal to each other.

We want to maximize the contribution of the first principal component u_1 : *i.e.*, maximize S_1 ,

* The author is professor of economics, College of Commerce, Seoul National University. He is also research member of the Institute of Economic Research, Seoul National University.

$$S_1 = \sum_i k_{i1}^2$$

under condition that

$$\mathbf{R} = \mathbf{K}\mathbf{K}'$$

where \mathbf{R} is a $p \times p$ symmetric matrix of correlations between the variables z_i ; and \mathbf{K}' , the transpose of \mathbf{K} . Here it is found that the solutions of k_{i1} are the components of eigenvector of matrix \mathbf{R} corresponding to the largest eigenvalue λ_1 of that matrix multiplied by the square root of λ_1 . Similarly we can find second component (or factor), third component, and so on.

III. On Application of Factor Analysis to Statistical Technique.

R. Stone has applied the method of factor analysis or principal components to national income data of American economy taken from the years 1922-38.

Here seventeen variables are reduced to three factors and these three factors together account for more than 97 per cent of the total variance of all variables. Moreover, the three largest factors can tentatively be identified with income (or output), rate of changes of income, and time trend respectively.

T. Klok and L. B. M. Mennes have, in their 1960 paper, furthered Stone's analysis and give new economic meaning to that method.

IV. On Application of Factor Analysis to Economic Development Theory,

I. M. Megee's paper purposes to discuss the usefulness of mathematical model for analyzing problems of economic growth on the one hand, and to analyze briefly the findings from some particular implementations of factor analysis with the hope of discovering some advantages, disadvantages, and implications for the use of the method in economic growth analysis and planning for future growth on the other. Our interest is in the latter.

She introduces three techniques: R , Q , and M -techniques. These three techniques provide the means of testing three kinds of hypotheses related to economic development: R can test hypotheses about variables causing economic development, using the factor loadings; Q can test hypotheses regarding the area differentiation of patterns of economic development, using the factor scores; and M can test hypotheses involving changes of variables or regions over time.

Although it is quite useful for testing hypotheses, factor analysis is also considered as an ideal way of exploring the unknown without the necessity of assumptions. It is possible to construct a large data matrix of economic and social variables and run factor analysis without any *a priori* assumptions. But it has some problems in the construction of an economic development matrix.

A major problem is lack of availability of data for desired information categories, resulting in their omission from the study. Another major problem is that of subjectivity, the selection of variables, the naming of the factors, the delineation of economic and social regions and the choice of the factor analysis method

itself all introduce a certain amount of subjectivity into the analysis, though it would seem to be less encumbered by restrictions than is multiregression analysis. Unlike other multivariate techniques, factor analysis cannot show inter- and intra-regional structure and flows, though some can be inferred from *Q*-technique.

The most important use of factor analysis for the social scientist is to isolate important or strategic variables for further study and to eliminate extraneous information. And factor analysis model can be used to identify and measure the relative importance of certain sectors of the economy as well as to indicate regional differentiation.

V. On the Application of Factor Analysis to Economic Development Theory, II

In their paper, I. Adelman and C.T. Morris attempt to gain some semiquantitative insights into the interaction of various types of social and political changes with the level of economic development.

For this purpose the techniques of factor analysis are applied to per capita GNP and to twenty-two indexes representing the social and political structure of seventy-four less developed countries in the period 1957-62. And to see the regional differences in the interrelationship between per capita GNP and socio-political influences the countries are divided into three groups: African countries, Near Eastern and Far Eastern countries, and Latin American countries.

In their analysis, the remarkably high percentage of intercountry variations in the level of economic development (66%) are associated with differences in non-economic characteristics. And a strong association was derived between per capita GNP (or the level of economic development) and two aspects of socio-political change: the socio-cultural concomitants of the industrialization-urbanization process (Factor I) and the Westernization of political institutions (Factor II). In contrast, a rather weak relationship appears between per capita GNP and indicators summarizing the character of leadership (Factor III) and the degree of social and political stability in the past decade (Factor IV).

In addition, the results of regional studies indicate that Factor I is overwhelmingly important for low income economies in which the absorptive capacity is sharply limited by the inhibiting nature of the social structure. However, even among countries at higher stages of evolution, the social variables remain the most important element associated with intercountry differences in per capita GNP. As for factor II, it is of negligible importance at the early stages of development. This result probably arises because both the ability to generate sustained economic growth and the evolution of more sophisticated political institutions require fundamental changes in mentality characteristic of Western thought pattern.

And at the last part of the paper they emphasize that the relationships found between levels of economic development and differences in social and political structure are neither caused nor causal.

VI. Conclusion

As M. Megee points out, factor analysis has deficiencies such as lack of availability of data for constitutions of factor matrix, incorporation of subjectivity, and impossibility of showing inter- and intra-regional structure and flows. And in the application of factor analysis to economic development theory, there are many to-be-solved problems: we must construct matrices to find weights for variables on economic development; incorporate political, social and psychological variables into the study of economic development; and vary the numbers and kinds of variables. Further, this analysis is a laborious method because of complexity in calculation and is subject to the type of computer. But this analysis is, as Megee says, useful for isolation of important or strategic variables for further study, and for elimination of extraneous information at an initial stage of research. This is the most important use of factor analysis for the social scientist. And it is also a merit of factor analysis that it is a highly complex model which can incorporate many more variables than the ordinary sorts of models constructed for dealing with economic problems.

In Korea, econometric studies are beginning to be applied to real problems. In these studies efficient model building is the most important process. Here factor analysis method may function usefully. And it may be thought that economic forecasting will be more accurate with the use of factor analysis.

In Korea, more attention is paid to the gross national product or per capita GNP. But there have been no attempts similar to Adelman and Morris'. It may be a valuable attempt to examine how much per capita GNP is influenced by noneconomic or political and social factors. Further, the method can suggest something about treatment of noneconomic factors in econometrics. Therefore, a proper study on factor analysis is much required in Korea.