主成分分析에 의한 特性値評価에 관한 研究
一身体検査의 例를 中心으로一

(A Study on Evaluation of the Characteristics Value in Principal Component Analysis

崔　震　永*
鄭　官　熹**

Abstract

The methed of principal component analysis is originated by K. Pearson, who considered this as geometrical method. Principal component analysis is the most elementary method, and this means that the information having various type of characteristics which have been correlated among themselves, are summarized by orthogonal transformations of characteristics.

I: Even though we have different result whether this method is applied to homogeneous population or not. In this research we should deal with the case of homogeneous poplation only.

II: On the other hand, we can have different result whether we start from covariance matrix or matrix of correlation-coefficients.

In this research we are studying based on covariance matrix.

1. 序　論

主成分分析(principal component analysis)은 多變量解析에서 가장 基本的인 方法論으로써, 研究對象으로는 工業製品의 品質에 관한 問題로 尺度·強度·電氣抵抗과 같은 物理的 特性이나, 化學的 特性등이 있으며, 人間의 身體에 대해서는 신장·체중·키·몸무게·혈압·안질 등 육체적·정신적의 特性을 分析하며, 또한 企業經營에 있어서는 資本·賣上高·収益性·資産効率과 같은 財務指標 등이 測定되어진다. 이들은 서로 相關이 있는 여러 종류의 特性値가 지닌 情報를 서로 無相關한 몇개의 總合特性値에 要約한다. 이 方法論을 等質集團(homogeneous population)에 適用하는 경우와 異質集團에 適用하는 경우와는 그 結果의 解析方法이 달라지는데 本研究에서는 等質集團에 適用하는 理論과 例를 들기로하고, 異質集團은 言及하지 않겠다. 또 分散, 共分散行列로 出發하느냐, 相關行列로 出發하느냐에 따라 解析이 달라진다. 主成分分析을 最初에 K. Pearson 이 생각했을 때는 幾何學的 解析에 基礎를 두었다는 것이다. 다시말해서 主成分分析은 p個의 特性値 또는 p變量 x_1, x_2, x_3, ……, x_p의 여러 情報를 다음의 2가지 條件을 만족하는 m個($m<p$)의 總合特性値 z_1, z_2, z_3, ……, Z_m 인 것을 제1, 제2, 제3,……, 제 m의 主成分이라하고 이것을 要約하는 方法으로는 (2·1)式이 있으며, 간단히하면 (2·2)式이 된다.[1)]

〔條件 1〕 Z_k 와 $Z_{k'}$($k \neq k'$) ($k = 1, 2, 3, \cdots\cdots$, m, $k' = 1, 2, 3, \cdots\cdots, m$)의 相關은 모두 零이다.

〔條件 2〕 Z_1의 分散은 x_1, x_2, x_3, ……, x_p의 모든 종류가 1次式의 分散이 最大이고, Z_2의 分散은 z_1, z_2, z_3, ……, Z_{m-1} 모두 無相關인 1次式중에서 最大이다. 第2性質은 「主成分 Z_1, Z_2, ……, Z_m 은 모두 特性値 x_1, x_2, ……, x_p 를 가진다」「고로 情報의 損失이 最小가 되게끔 要約하여 結果를 얻는다」 m個를 몇개 취할 것인가에서 情報의 *loss*(損失)을 어느정도 허용하는 것인가에 의해서 決定되나 m은

* 京畿大學 助教授
** 京畿大學 專任講師

1) 奥野忠外 3人, 多變量解析法(東京：日本科技達, 1972), pp.13~15, 159~257.

될 수 있는대로 작은 편이 좋다. 研究目的은 서울 市內 한 私立國民學校 아동 1~6學年 男女學生 1,826 名의 身體發達狀態를 알아보기 위하여 主成分 分析을 하였으며, 研究範圍는 平均・標準偏差・相關係數・固有값(eigenvalue)・固有 vector 등을 計算하여 부록에 실었다. 이들 表에서 x_1(키), x_2(몸무게), x_3(가슴둘레), x_4(앉은키)를 말한다. 이들의 結果値를 보기 위하여, 컴퓨터를 使用하였으며, 특히 IBM의 S. S. P. 책에 있는 Factor-analysis의 Sub-program을 利用하였다. (附錄 3 參照)[2]

2. 主成分의 性質 및 分析法

2-1 提起 및 導出方法

表 2・1 p 個의 特性値와 m 個의 主成分

標本	기 본 특 성 치	추출하여진 주성분
No.	$x_1\ x_2 \cdots x_i \cdots x_p$	$z_1\ z_2 \cdots z_k \cdots z_m$
1	$x_{11}\ x_{12} \cdots x_{1i} \cdots x_{1p}$	$z_{11}\ z_{12} \cdots z_{1k} \cdots z_{1m}$
2	$x_{21}\ x_{22} \cdots x_{2i} \cdots x_{2p}$	$z_{21}\ z_{22} \cdots z_{2k} \cdots z_{2m}$
-	— — … — … —	— — … — … —
-	— — … — … —	— — … — … —
α	$x_{\alpha 1}\ x_{\alpha 2} \cdots \boxed{x_{\alpha i}} \cdots x_{\alpha p}$	$z_{\alpha 1}\ z_{\alpha 2} \cdots \boxed{z_{\alpha k}} \cdots z_{\alpha m}$
-	— — … — … —	— — … — … —
-	— — … — … —	— — … — … —
n	$x_{n1}\ x_{n2} \cdots x_{ni} \cdots x_{np}$	$z_{n1}\ z_{n2} \cdots z_{nk} \cdots z_{nm}$
計	$T_1\ \ T_2 \cdots T_i \cdots T_{np}$	
平 均	$\bar{x}_1\ \ \bar{x}_2 \cdots \bar{x}_i \cdots \bar{x}_p$	$\bar{z}_1\ \bar{z}_2 \cdots \bar{z}_k \cdots \bar{z}_m$
平方合	$S_{11}\ \ S_{22} \cdots S_{ii} \cdots S_{pp}$	
積 合	$S_{ii'}\ (ii' = 1,2,\cdots,p;\ i \neq i')$	

表 2・1에서 $x_{\alpha i}$ $(\alpha = 1,2,3,\cdots\cdots,\ n)$, $(i = 1,2,\cdots\cdots,\ p)$는 n 個의 標本들에 관해서 測定되는 p 種類의 特性値들이다. 여기서 p 의 特性値는 事例 및 考察에서 키(cm), 몸무게(kg), 가슴둘레(cm), 앉은키(cm)가 섞여 있어도 좋다는 것이다. 이렇게 단위가 다를 때는 全體特性値 x를 標準化할 $x_{\alpha i} = (x_{\alpha i} - \bar{x}_i)/S_i$를 使用한다. $x_{\alpha i}$의 分散, 共分散은 實際測定値 $x_{\alpha i}$의 相關行列과 같다.[3]

$$S = (S_{ii'}) = \begin{pmatrix} S_{11} & S_{12} & \cdots\cdots & S_{1p} \\ S_{12} & S_{22} & \cdots\cdots & S_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ S_{1p} & S_{2p} & \cdots\cdots & S_{pp} \end{pmatrix}$$

2) Scientific Subroutine Program IBM

3) John Willy, *An Introduction to Probability theory and its Application*, New York : pp.236~237.
Murray R. Spigel, *Schaum's Outline of theory and Problems of Statistics*, New York : Schaum Publishing Co., 1961.

$$V = (V_{ii'}) = \begin{pmatrix} V_{11} & V_{12} & \cdots\cdots & V_{1p} \\ V_{12} & V_{22} & \cdots\cdots & V_{2p} \\ \vdots & \vdots & & \vdots \\ V_{1p} & V_{2p} & \cdots\cdots & V_{pp} \end{pmatrix}$$

$$R = (r_{ii'}) = \begin{pmatrix} 1 & r_{12} & \cdots\cdots & r_{1p} \\ r_{12} & 1 & \cdots\cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{1p} & r_{2p} & \cdots\cdots & 1 \end{pmatrix}$$

단,

偏差平方合, 積合 $S_{ii'} = \sum_{\alpha=1}^{n} (x_{\alpha i} - \bar{x}_i)(x_{\alpha i'} - \bar{x}_{i'})$

分散, 共分散 $V_{ii'} = S_{ii'}/(n-1)$

標準偏差 $S_i = \sqrt{V_{ii'}}$

相關係數 $r_{ii'} = S_{ii'} / \sqrt{S_{ii} \cdot S_{i'i'}}$
$= V_{ii'} / \sqrt{V_{ii} \cdot V_{i'i'}}$

m 個$(m < p)$인 總合特性値를 다음 式(2・1, 2・2)과 條件(A, B, C)을 만족하도록 定해준다. 이들을 第1, 第2, 第3, ……, 第m의 主成分이라고 한다.

$$\begin{aligned} z_1 &= l_{11} x_1 + l_{12} x_2 + \cdots\cdots + l_{1p} x_p = \sum_{i=1}^{p} l_{1i}\, x_i \\ z_2 &= l_{2i} x_1 + l_{22} x_2 + \cdots\cdots + l_{2p} x_p = \sum_{i=1}^{p} l_{2i}\, x_i \\ &\cdots\cdots \\ z_k &= l_k x_1 + l_{k2} x_2 + \cdots\cdots + l_{kp} x_p = \sum_{i=1}^{p} l_{ki}\, x_i \\ &\cdots\cdots \\ z_m &= l_{m1} x_1 + l_{m2} x_2 + \cdots\cdots + l_{mp} x_p = \sum_{i=1}^{p} l_{mi}\, x_i \end{aligned} \tag{2・1}$$

여기서 $k = 1,2,3,\cdots\cdots,\ m$일 때

$$l_{k1}^2 + l_{k2}^2 + l_{k3}^2 + \cdots\cdots + l_{kp}^2 = \sum_{i=1}^{p} l_{k2}^2 = 1 \tag{2・2}$$

은 考察에서 實例에서 증명되었다.

〔條件 A〕 ① 第1 主成分 Z_1 의 係數 $\{l_{1i}\}$ $(i = 1,2,3,\cdots\cdots,\ P)$는 (2・2)式의 條件下에서 Z_1 의 分散이 最大가 되는 것으로 정한다.

② 第2主成分은 係數 $\{l_{2i}\}$ 로써 ①과 같으며 이럴 때는 分散이 큰 Z_1, Z_2, ……, Z_m 라 한다.

③ 이와 같은 方法으로써 第K의 主成分의 係數 $\{l_{ki}\}$ $(i = 1,2,\cdots\cdots,\ p)$는 Z_k 가 Z_1, Z_2, Z_3, ……, Z_{k-1}과 無相關이 되는 條件下에서 Z_k 의 分散이 最大가 되는 것으로 한다. 이럴 때도 分散이 큰 것부터 차례로 Z_1, Z_2, Z_3, ……, Z_m 라 한다.

〔條件 B〕 變數 x_1, x_2, x_3, ……, x_p에 直交變換을 하여 얻어진 變數를 z_1, z_2, z_3 …… z_p라 한다. 이들 z_1, z_2, z_3, ……, z_p가 서로 無相關이 되도록 (一義的으로 定해진다)하고 그 分散이 큰 것부터 차례로 z_1, z_2, z_3, ……, z_p라 한다.

〔條件 C〕 $\{z_1, z_2, z_3, \cdots\cdots, Z_m\}$를 n個 對象에 주어진 새로운 m個의 指標라고 할 때 實際變數 X_i의 다음의 指標에 관한 重相關係數 R_i의 自乘合 $\sum_{i=1}^{p} R_i^2$이 最大가 되도록 그 $z_1, z_2, z_3, \cdots\cdots, Z_m$을 決定한다. 이 條件들은 다음 式과 같이 쓸 수 있다. 적당하게 뽑은 m個의 指標 $z_1, z_2, z_3, \cdots\cdots, Z_m$의 1次係數는

$$
\begin{array}{l}
X_1 = a_{11} z_1 + a_{12} z_2 + a_{13} z_3 + \cdots\cdots + a_{1m} z_m \\
X_2 = a_{21} z_1 + a_{22} z_2 + a_{23} z_3 + \cdots\cdots + a_{2m} z_m \\
- \quad - \quad - \quad - \quad \cdots\cdots \quad - \\
- \quad - \quad - \quad - \quad \cdots\cdots \quad - \\
- \quad - \quad - \quad - \quad \cdots\cdots \quad - \\
X_p = a_{p1} z_1 + a_{p2} z_2 + a_{p3} z_3 + \cdots\cdots + a_{pm} z_m
\end{array}
$$

라 할 때

$$\sum_{i=1}^{p} \sum_{\alpha=1}^{n} (x_{\alpha i} - x_{\alpha i})^2 / S_i^2$$

이 最小가 되도록 $z_1, z_2, z_3, \cdots\cdots, Z_m$을 決定한다. 따라서, 以上과 같은 提起에 의해서 다음과 같은 導出方法이 이루어진다.

(1) P變量 主成分과 그 導出法

條件 A의 ①, ②에 의하여 z_1, z_2의 分散이 最大가 되는 것으로 定한다. 지금 (2·1)式에 열거한 條件 A에 따라서 主成分을 求한다면 係數 $\{l_{1i}\}$가 決定되고, 第1主成分 z_1의 값이 n個 샘플의 各各에 의해서 計算한 表2·1에서 $\{Z_{\alpha k}\}$와 같이 求해진다. 이때 Z_1의 分散 $V[Z_1]$은

$$
\begin{aligned}
[Z_1] &= \sum_{\alpha=1}^{n} (z_{\alpha 1} - \bar{z}_1)^2 / (n-1) \\
&= \sum_{\alpha=1}^{n} \left\{ \sum_{i=1}^{p} l_{1i}(x_{\alpha i} - \bar{x}_i) \right\}^2 \Big/ (n-1) \\
&= \sum_{i=1}^{p} \sum_{i'=1}^{p} l_{1i}\, l_{1i'} \sum_{\alpha=1}^{n} (x_{\alpha i} - \bar{x}_i)(x_{\alpha i'} - \bar{x}_{i'}) / (n-1) \\
&= \sum_{i} \sum_{i'} l_{1i}\, l_{1i'} V_{ii'}
\end{aligned}
\qquad (2\cdot3)
$$

基準化한 값 x'를 使用하면 (2·3)式에서부터 (2·8)式까지 나타내는 $V_{ii'}$를 $r_{ii'}$로 바꾸면 된다. 條件 A의 ①에 따라서 Z_1의 分散이 最大가 되어야 하므로, langrange의 未定乘數法을 使用하면,[4]

$$Q = \sum_{i=1}^{p} \sum_{i'=1}^{p} l_{1i}\, l_{1i'} V_{ii'} - \lambda \left(\sum_{i=1}^{p} l_{1i}^2 - 1 \right) \ (max)$$

이 Q를 各 l_{1i}로 偏微分하여 零이라 놓으면,

$$\frac{\partial \theta}{\partial l_{1i}} = 0 \text{ 즉, } \sum_{i'} l_{1i'} V_{ii'} - \lambda l_{1i} = 0 \quad (i = 1, 2, \cdots\cdots, p) \qquad (2\cdot4)$$

(2·4)式을 다시 쓰면,

$$
\begin{array}{l}
(V_{11} - \lambda)\, l_{11} + V_{12}\, l_{12} + \cdots\cdots + V_{1i}\, l_{1i} + \cdots\cdots + V_{1p}\, l_{1p} = 0 \\
V_{12}\, l_{11} + (V_{22} - \lambda)\, l_{12} + \cdots\cdots + V_{2i}\, l_{1i} + \cdots\cdots + V_{2p}\, l_{2p} = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
V_{1i}\, l_{11} + V_{2i}\, l_{12} + \cdots\cdots + (V_{2i} - \lambda)\, l_{1i} + \cdots\cdots + V_{ip}\, l_{1p} = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
V_{1p}\, l_{11} + V_{2p}\, l_{12} + \cdots\cdots + V_{ip}\, l_{1i} + \cdots\cdots + (V_{pp} - \lambda)\, l_{1p} = 0
\end{array}
\qquad (2\cdot5)
$$

(2·5)式에서 p個의 未知數 $l_{11}, l_{12}, \cdots\cdots, l_{1p}$이고 常數項이 零이므로 p個의 式이 서로 独立하면, 解는

$$l_{11} = l_{12} = \cdots\cdots = \lambda_{1p} = 0 \qquad (2\cdot6)$$

이때 (2·5)式 係數의 行列은 零으로 되어야 된다. 지금 대각선상 要素는 1, 다른 것은 모두 零인 行列을 I로 나타내면 다음과 같다.

$$
|V - \lambda \mathrm{I}| = \begin{vmatrix}
V_{11} - \lambda & V_{12} & \cdots\cdots & V_{1p} \\
V_{12} & V_{22} - \lambda & \cdots & V_{2p} \\
\vdots & \vdots & \ddots & \vdots \\
V_{1p} & V_{2p} & \cdots\cdots & V_{pp} - \lambda
\end{vmatrix} = 0 \qquad (2\cdot7)
$$

이 行列式을 λ의 多項式으로 전개하면 $\alpha_0 + \alpha_1 \lambda + \alpha_2 \lambda^2 + \alpha_3 \lambda^3 + \cdots\cdots + \alpha_p \lambda^p = 0$와 같이 표시된다. 이것을 만족하는 λ의 값은 p個임을 안다. 또 이 p個의 根 $\lambda_1, \lambda_2, \lambda_3, \cdots\cdots, \lambda_p$는 어느 것이나 實數이고, 正 또는 零인 것으로 알 수 있다. 그러므로 그것을 크기順으로 나열하면,

$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots\cdots \geq \lambda_p \geq 0$으로 된다. 「이 $\{\lambda_k\}$ $(K = 1, 2, 3, \cdots\cdots, p)$를 行列 V의 固有値(eigenvalue)라고 한다.」[5] 여기서 λ_k중 어느것 하나를 (2·4)式에 代入하여보면, 이 p個의 式 적어도 1개는 從屬이 된다. 따라서 独立하지 않은 式은 없애고, 그 代身에 (2·2)式의 條件을 使用해서 係數 $\{l_{ki}\}$ $(K = 1, 2, 3, \cdots\cdots, p)$를 求한다. 「이 解

4) 安恩明, 現代統計學 (서울 : 日新社, 1970), p. 101.

5) 鄭英眞, 近代統計學 理論과 實際(서울 : 普晋齋, 1971), pp. 174～200.

를 行列 V의 固有 vector 라 부른다』 以上에서는 Q를 1回偏微分해서 零이라고 하는條件下에서 p個의 解를 얻었다. Q를 最大로하는 이들 解中에 있다. 이 解들은 어느 것이나 (2·2)式을 만족하기 때문에

$$\sum_{i'} l_{ki}' V_{ii}' = \lambda_k \lambda_{ki} \ (K = 1, 2, 3, \cdots\cdots, p)$$

이 式을 (2·3)式에 代入하면,

$$\boxed{\begin{aligned} V[z_1] &= \sum_i \sum_{i'} l_{ki}\, l_{ki}' V_{ii}' \\ &= \sum_i l_{ki}\ (\lambda_k\, l_{ki}) \\ &= \lambda_k \sum_i l_{ki}^2 = \lambda_k \end{aligned}} \qquad (2\cdot 8)$$

을 얻는다. 그러므로 最大가 될 때 z_1의 分散은 λ_k ($K = 1, 2, 3, \cdots\cdots, p$)의 어느 것과 同等하다고 할 수 있다. 故로 λ_k 중에 最大인 것으로하는 λ_1을 잡고, λ_1에 對應하는 固有 vector 를 $\{l_{1i}\}$ ($i = 1, 2, 3, \cdots\cdots, p$)라고 하면 된다.

第2의 主成分도 거의 위와 같은 方法으로 求해지므로 結果的으로 m個의 主成分을 求하는 係數 $\{l_{ki}\}$ ($K = 1, 2, 3, \cdots\cdots, m$; $i = 1, 2, \cdots\cdots, p$)는 基本變數 x_1, x_2, $x_3 \cdots\cdots$, x_p의 分散, 共分散行列 V(또는 相關行列 R)의 固有値 $\lambda_1 \geqq \lambda_2 \geqq \lambda_3 \geqq \cdots\cdots \geqq \lambda_m$를 取해 對應하는 固有 Vector $[\{l_{1i}\}, \{l_{2i}\}, \{l_{3i}\} \cdots\cdots, \{l_{mi}\}]$를 充當하면 된다.

2-2 性 質

相關行列 R에서 出發하여 求하는 경우를 說明하고, 다음은 分散, 共分散行列 V에서 出發하여 求하는 경우 相違點을 들어보기로 한다.[6)]

表2·2에서 나타난 것과 같이 V에서 出發할 때 $\sqrt{\lambda_k}\, l_{ki}/S_i$ 와 $\sum r^2 (Z_k, x_i) S_i^2$ 으로 주어지며 여기서는 V의 trace (對角要素의 合)을 $t_r(V)$라 쓰면 $\sum_{K=1}^{p} \lambda_K \fallingdotseq t_r(V) = \sum_{i=1}^{p} V_{ii}$가 된다. 主成分의 性質을 나열하면 다음 7가지가 있는데[7)] ① 主成分 Z_K의 分散은 固有値 λ_K로 同等하며, $V[Z_K] = \lambda_K$ ($K = 1, 2, 3, \cdots\cdots, m \leqq p$)이다. 여기서 λ_K는 (2·8)式에 있다. λ_K는 큰 順으로 主成分 Z_K를 잡는다고 하면 된다. ② x_i의 分散은 $V[x_i'] = 1$이므로 그 合은 $\sum_{i=1}^{p} V[x_i'] = p$이다. 主成分 Z_K 分散의 總分散에 대한 比率은 λ_K/p이고, m個의 主成分 $\{Z_1, Z_2, \cdots\cdots, Z_m\}$의 累積寄與率은 $\sum_{i=1}^{p} \lambda_K/p$로 주어진다. 여기서 $m = p$라고 하면, 固有値(eiqenvalue)合計는 $\sum_{K=1}^{p} \lambda_K = p$가 되며 누적기여율은 1로 되기 때문에 이것으로 p개의 基本의 變數로 內包하고 있는 情報의 全部를 아는 것이 된다. $\sum_{i=1}^{p} V[x_i] = \sum_{i=1}^{p} V_{ii}$ 이

表 2·2 x_i 와 z_k의 相關 關係

	基本變數	主 成 分	2乘合 $\sum_k r^2 (z_p, x_i)$
	$x_1\ x_2 \cdots\cdots\cdots x_p$	$z_1\ z_2 \cdots\cdots\cdots z_m$	
x_1 x_2 $\vdots$ x_p	$\begin{matrix} 1 & & \\ & 1 & \\ & & \ddots \\ & & & 1 \end{matrix}$	$r(z_k, x_i)\sqrt{\lambda_k}\ l_{ki}$ (因子負荷量) $\cdots\cdots\cdots\cdots\cdots\cdots$	$\nu_i = R^2 (x_i : z_1 \cdots z_n)$ (x_i에 對応하는 寄與率)
2乘合	$\sum r^2 (z_k, x_i)$	$\lambda_1\ \lambda_2 \cdots\cdots\cdots \lambda_m$	
		$\begin{matrix} 1 & & 0 \\ & 1 \ddots & \\ 0 & & 1 \end{matrix}$	

므로, m個의 主成分 $\{Z_1, Z_2, \cdots\cdots, Z_m\}$의 누적기여율은 $\sum_{K=1}^{m} \lambda_K \Big/ \sum_{i=1}^{p} V_{ii}'$로 주어진다.

③ 各 主成分은 서로 無相關하다. 왜냐하면 基準化 했던 變數 x_{ai}, x_{ai}'의 共分散은 相關係數 r_{ii}' ($i \neq i'$일 때 1)임을 利用하면

$$\begin{aligned} \mathrm{Cov}\,[z_k, z_{k'}] &= \sum_{\alpha=1}^{n} (z_{\alpha k})(z_{\alpha k'})/(n-1) \\ &= \sum_i \sum_{i'} l_{ki}\, l_{k'i'}\, r_{ii'} \\ &= \lambda_{k'} \sum l_{ki}\, l_{k'i} \\ &= 0 \quad (k \neq k') \end{aligned}$$

6) 金俊輔外 2人譯, 數理統計學(서울 : 高麗大學校 出版部, 1971), pp. 182~186.
Hoel, *Introduction to Mathematical Statistics* (New York : John Willy, 1960), pp.117~136.

7) 具野忠外 3人, 前揭書, pp. 13~15, 159~257.
淺野啓三, 行列·行列式(東京 : 共立出版, 1961), pp. 50~56.
岩原信丸郎, 新教育統計法(東京 : 日本文化科學社, 1962), pp. 136~148.

이 되기 때문에 여기서 $K=K'$에 對應하는 固有 vector l_K와 l_K'는 서로 直交한다.

④ 主成分 Z_K와 特性値 x_i와의 相關은 다음과 같다.

$$r(z_k, x_i) = \frac{\mathrm{Cov}(z_k, x_i)}{\sqrt{V(z_k)\cdot V(x_i)}} = \frac{\sum_{i'} l_{ki'}\cdot r_{ii'}}{\sqrt{\lambda_k}} = \frac{\lambda_k \cdot l_{ki}}{\sqrt{\lambda_k}} = \sqrt{\lambda_k}\cdot l_{ki}$$

$(k=1,2,3,\cdots\cdots,\ m : i=1,2,3,\cdots\cdots,\ p)$

⑤ 이 相關係數를 2乘하여 i에 따라서 더해가면 λ_K가 된다.

$$\sum_{i=1}^{p} r^2(z_k, x_i) = \sum_i \lambda_k l_{ki}^2 = \lambda_k$$

다시 Z_k의 x_i에 對應하는 기여율 V_{ii}를 係數로 하는 合은 λ_k와 같으므로 그럴 때는

$$\sum_{i=1}^{p} V_{ii} r^2(z_k, x_i) = \sum_i \lambda_k l_{ki}^2 = \lambda_k$$

로 된다.

⑥ m個의 主成分 $\{Z_1, Z_2, Z_3, \cdots\cdots Z_m\}$의 基本變數 x_i에 對應하는 기여율 $r_i(x_i)$의 $\{Z_1, Z_2, \cdots\cdots Z_m\}$에 對應하는 重相關係數의 2乘合은 다음 式으로 주어진다.

$$\nu_i = \sum_{k=1}^{m} r^2(z_k, x_i) = \sum_{k=1}^{m} \lambda_k l_{ki}^2$$

여기서 Z_k와 Z_k'가 無相關하므로 ③의 性質을 利用했다. 만약 $m=p$까지 잡으면

$$\sum_{k=1}^{p} r^2(z_k, x_i) = \sum_{k=1}^{p} \lambda_k l_{ki}^2 = 1$$

이 됨을 알 수 있다. 故로

$$\nu_i = \sum_{k=1}^{m} r^2(z_k, x_i) = \sum_{k=1}^{m} \lambda_k \cdot l_{ki}^2 / V_{ii}$$

으로 된다.

⑦ 各變數 x_i에 m個의 主成分$(Z_1, Z_2, Z_3, \cdots\cdots Z_m)$에 對應하는 重回歸式은 다음 式(2·9)와 같다. 여기서 x_i는 回歸推定値를 表示한다.

$$\begin{aligned} x_1 &= l_{11} z_1 + l_{21} z_2 + \cdots\cdots + l_{m1} z_m \\ x_2 &= l_{12} z_1 + l_{22} z_2 + \cdots\cdots + l_{m2} z_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_i &= l_{1i} z_1 + l_{2i} z_2 + \cdots\cdots + l_{mi} z_m \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_p &= l_{1p} z_1 + l_{2p} z_2 + \cdots\cdots + l_{mp} z_m \end{aligned} \qquad \cdots\cdots(2\cdot 9)$$

2-3 等質集団에서의 分析法

샘플(sample)을 構成하는 n個의 對象은 等質集團에서 꺼내졌던 경우를 생각한다. 對象集團이 먼저 몇개의 組로 나눌 수 있는 것과 같은 경우에 대해서는 實例에서 보기로 하고, 여기서는 等質集團을 適用하는 데 있어서 P·C·A를 研究하기로 한다.

① 變數의 選擇

入力한 多變數의 全部에 의해 P·C·A를 計算한 것뿐만 아니라 그 一部를 任意로 취해 꺼내서 計算할 수 있도록 놓아야 한다. 變數의 變換은 1次結合에 編入한 것에는 測定値 그대로가 아니고 適當한 變換을 實施하는 것이 좋다.

② 行列의 選擇

P·C·A는 相關行列 R에 分散, 共分數行列 V 혹은 偏差平方合, 積合行列, S에 適用하여도 그 結果는 一致하지 않는다. 또 原點에서 2次積率(moment)의 行列을 使用하는 便이 좋은 경우도 있다. 等質集團에서 m까지의 主成分의 누적기여율은 $\sum_{k=1}^{m} \lambda_k = p$ 또는 $\sum_{k=1}^{m} \lambda_k \big/ \sum_{i=1}^{p} V_{ii}'$式에서 60%로 해야 좋을지 80% 以上으로 해야 좋을지는 解析의 目的이 된다. 큰 情報만을 빼내려고 할 때는 60% 前後가 좋고, 情報損失을 작게하려면 80% 以外가 좋다. 第m번째 主成分의 기여율이 平均以上으로 있을 경우, 그 기준치는 固有値(eigenvalue) λ_m가 다음 式을 만족하는 경우에 해당한다. 相關行列 R부터일 때

$$\lambda_m \geq 1.0$$

分散, 共分散行列 V_{ii}일 때

$$\lambda_m \geq \sum_{i=1}^{p} V_{ii} \big/ p$$

로 되며, 基本變數 x_i가 지닌 정보량은 그 分散 V_{ii}로 있으니까, 그 平均은 $\sum_i V_{ii}/p$이다. R부터 出發할 때는 全部 標準化되여 있고 $V_{ii}=1.0$이 되고 있다. 만약 어느 1개의 變數 x_i가 다른 모든 變數와 無相關에 있었다고 한다면 이때 R에서 出發하면 $\lambda=1$이 지닌 主成分 어딘가에 나타나고, 그것은 x_i의 自身이 될 것이다. m個의 主成分의 各 變數 x_i에 대한 기여율 $\nu_i = \sum_{k=1}^{m} r^2(Z_k, x_i) = \sum_{k=1}^{m} \lambda_K, l_{ki}^2$가 어느 x_i에 대해서도 50% 以上이 되도록 하는 規定은 엄격하다. 이것을 完全히 만족시킨다고 하면 빼내야만 하는 主成分數 m가 있으면, 차라리 x_i를 제거하여 計算하여 고치는 것이 좋다. 그 x_i는 다른

變數와는 協調性이 작은 特殊의 變數로써 처음부터 같이 取扱한 것이다.[8)]

3. 事例 및 分析結果 考察

3-1 身体計測値에 관한 例

이 例는 서울市内 한 私立國民學校의 4가지 特性値를 測定하였다. 그 測定對象은 表3·1과 같다.

이 데이터에 관한 統計處理는 부록 표1과 2에 나타냈으며, 특히 부록 표2에는 主成分分析의 eigen-value 의 값이 나타나 있다.

(1) 身體發達過程分析

測定對象 全員에 대한 男女別, 學年別 및 全體 成長發達過程을 分析해보면 그 結果는 表3·2와 같다.

表 3·1 対象学生人員数 類

	男學生		女學生		全體學生	
	人員數	%	人員數	%	人員數	%
1學年	174	16.4	120	15.7	294	16.1
2 〃	161	15.2	115	15.0	276	15.1
3 〃	193	18.3	132	17.3	325	17.8
4 〃	183	17.2	135	17.6	318	17.5
5 〃	·179	16.9	124	16.1	303	16.5
6 〃	169	16.0	141	18.3	310	17.0
計	1,059	100.0	767	100.0	1,826	100.0

表 3·2 全体 및 男女別·学年別 成長比例表

		키	몸무게	가슴둘레	앉은키
全體學生	1 ~ 2	5.65860	2.36549*	*3.26902	2.61486
	2 ~ 3	5.57891	3.13631	2.59717	2.50740
	計	11.23751	5.50180	5.86619	5.12226
	3 ~ 4	5.09594*	2.86969	2.42315	2.00378*
	4 ~ 5	5.44158	3.40379	2.74319	2.01906
	5 ~ 6	*6.16574	*4.08331	1.76960*	*3.44035
	計	16.70326	10.35679	6.93594	7.46319
	累積表	27.94077	15.85859	12.80213	12.56545
男學生	1 ~ 2	*5.99358	2.41841*	*3.27465	2.46715
	2 ~ 3	4.91926	2.97275	2.35372	2.05312
	計	10.91284	5.39116	5.62837	4.52027
	3 ~ 4	5.47591	2.87240	2.33796	2.35391
	4 ~ 5	4.88800*	3.21759	2.46011	1.73936*
	5 ~ 6	5.49115	3.69269	1.86688*	*2.99404
	計	15.85506	9.78268	6.66495	7.08731
	累積表	26.76790	15.17384	12.29332	6.60758
女學生	1 ~ 2	5.21919	2.31684*	*3.30713	2.84912
	2 ~ 3	6.48606	3.34051	2.89598	3.14679
	計	11.70525	5.65735	6.20311	5.99591
	3 ~ 4	4.61016*	2.90847	2.61109	1.52368*
	4 ~ 5	6.18326	3.63732	3.09910	2.39667
	5 ~ 6	*6.99851	*4.60412	1.11524*	*3.97963
	計	17.79193	11.14991	6.82543	7.89998
	累積表	29.49718	16.80726	13.02854	13.89589

例: * 右側은 群別 最底, 左側은 群別 最高.

표3·2에 나타난 바와 같이 앉은키를 제외하고는 남학생보다는 여학생이 더욱 성숙하게 자라고 있다는 것이 나타나고 있으며, 좀더 세부적으로 서술하며는 다음과 같다.

① 신장은 전체적으로는 5~6학년 때 잘 성장하고, 3~4학년 때 잘 성장하지 않은 편이다. 男•女學生別로 성장비례는 학년과 학년 사이 과정은 表3·2와 같이 나타난다. ② 체중은 신장과 비례하고 男•女學生은 表3·2와 같다. ③ 가슴둘레는 키나 몸무게에 비례하기는커녕 正異例的인 事實이 나타난다. 다시말해서 신장이 성장하거나 몸무게가 늘어난다고 해서 가슴둘레가 늘어난다는 것은 아니다. ④ 앉은키는 전체적으로 신장과 몸무게와 비례한다. ⑤ 以上을 全體的으로 $p=4$에 대하여 고찰하면, 몸무게와 가슴둘레는 반비례하고, 신장과 몸무게는 거의 비례한다.

따라서 서로 同時에 成熟한 部分은 가슴둘레를 제

8) 具滋興外 2人, 統計學概論(서울: 集賢社, 1973), p. 188.
鄭英眞, 近代統計學 理論과 實際(서울: 寶晋斎, 1971), pp. 174~200.

외하고는 거의 比例하고, 低調한 部分에 있어서는 異例的이다.

(2) $p=4$에 관한 相關分析

다음의 分析은 IBM S·S·P을 利用하여 電算處理 하였다.

$p=4$인 特性値의 平均(부록 표1), 標準偏差(부록 표1) 및 相關係數(부록 표2)를 나타냈다.이 부록표에는 全體學生, 男學生, 女學生을 各學年別로 따로 分類되어 있다.

① 標準偏差가 큰 特性値는 키, 가슴둘레, 몸무게, 앉은키순으로 되어 있다.

② 相關係數에서 볼 수 있는 重要한 特徵은 相關의 값이 0.8 以上의 큰 것과 以下의 작은것으로 구분된다. 몸무게 x_2은 身體 各部分의 測定値와 크던 작던 간에 相關을 갖는다. 몸무게 x_2의 例를 들면 가슴둘레 x_3에 대한 相關은 5학년 男學生 係數 0.93745와 같은 높은 相關을 갖고 있다. 이 외에도 몸무게와의 相關이 0.7 ~ 0.9인 것들로 키 등이 있다. 가슴둘레 x_3와 앉은키 x_4는 0.23706으로 낮은 것도 있다.

3·2 結果 考察

지금 $p=4$에 관한 R에 대해서 P·C·A의 결과는 부록 표1·2와 같다. 第1主成分에 대한 λ는 어느 것이나 1보다 크고, 이 累積百分率은 64.445%에서 81.082%까지로 되어 있다. 第2主成分까지의 累積百分率은 85.167%에서 95.379%까지로 되어 있어서 모든 정보가 제1, 제2主成分에 集約되어 있다고 볼 수 있다. 그러므로 第1, 第2主成分만을 使用하기로 한다. 따라서 主成分은 附錄 表2의 固有 Vector 에 의하여 다음과 같은 式이 얻어졌다. 이들 어느 경우에나 係數 l_{Ki} ($i=1, 2, 3, 4$)은 各 K일 때 그 제곱合은 1이 되도록 規準化되어 있다. 또 相關行列 R에서 出發할 때 各主成分의 意義를 解析해보기로 한다. 그러므로 여기서第1, 第2 主成分의 係數는 모두 正·負인 것이 섞여있다.

〈 1学年 全体 및 男女学生別 主成分〉

$$全體\begin{cases} z_1 = 0.50418\,x_1 + 0.53979\,x_2 + 0.47157\,x_3 + 0.48170\,x_4 \\ z_2 = -0.42655\,x_1 + 0.28727\,x_2 + 0.67178\,x_3 - 0.53311\,x_4 \end{cases}$$

$$男\begin{cases} z_1 = 0.50634\,x_1 + 0.53676\,x_2 + 0.47872\,x_3 + 0.47572\,x_4 \\ z_2 = -0.40702\,x_1 + 0.32208\,x_2 + 0.63640\,x_3 - 0.57059\,x_4 \end{cases}$$

$$女\begin{cases} z_1 = 0.51640\,x_1 + 0.55474\,x_2 + 0.43962\,x_3 + 0.48198\,x_4 \\ z_2 = -0.36906\,x_1 + 0.21183\,x_2 + 0.73893\,x_3 - 0.52238\,x_4 \end{cases}$$

〈 2学年 全体 및 男女学生別 主成分〉

$$全體\begin{cases} z_1 = 0.50011\,x_1 + 0.52710\,x_2 + 0.47471\,x_3 + 0.49664\,x_4 \\ z_2 = -0.47153\,x_1 + 0.32224\,x_2 + 0.65564\,x_3 - 0.49390\,x_4 \end{cases}$$

$$男\begin{cases} z_1 = 0.49831\,x_1 + 0.52976\,x_2 + 0.47768\,x_3 + 0.49279\,x_4 \\ z_2 = -0.46244\,x_1 + 0.31371\,x_2 + 0.65680\,x_3 - 0.50628\,x_4 \end{cases}$$

$$女\begin{cases} z_1 = 0.50145\,x_1 + 0.53071\,x_2 + 0.45992\,x_3 + 0.50533\,x_4 \\ z_2 = -0.49178\,x_1 + 0.31820\,x_2 + 0.67003\,x_3 - 0.45600\,x_4 \end{cases}$$

〈 3学年 全体 및 男女学生別 主成分〉

$$全體\begin{cases} z_1 = 0.57370\,x_1 + 0.58290\,x_2 + 0.54065\,x_3 + 0.32251\,x_4 \\ z_2 = 0.01609\,x_1 - 0.23127\,x_2 - 0.31495\,x_3 + 0.92035\,x_4 \end{cases}$$

$$男\begin{cases} z_1 = 0.50009\,x_1 + 0.53041\,x_2 + 0.47454\,x_3 + 0.49333\,x_4 \\ z_2 = -0.46538\,x_1 + 0.32940\,x_2 + 0.64765\,x_3 - 0.50541\,x_4 \end{cases}$$

$$女\begin{cases} z_1 = 0.51615\,x_1 + 0.59428\,x_2 + 0.55746\,x_3 + 0.26390\,x_4 \\ z_2 = -0.00820\,x_1 - 0.17719\,x_2 - 0.18599\,x_3 + 0.96249\,x_4 \end{cases}$$

〈 4学年 全体 및 男女学生別 主成分〉

$$全體\begin{cases} z_1 = 0.49444\,x_1 + 0.52715\,x_2 + 0.48406\,x_3 + 0.49326\,x_4 \\ z_2 = -0.49510\,x_1 + 0.37721\,x_2 + 0.60316\,x_3 - 0.49875\,x_4 \end{cases}$$

$$男\begin{cases} z_1 = 0.49452\,x_1 + 0.52905\,x_2 + 0.48257\,x_3 + 0.49261\,x_4 \\ z_2 = 0.49190\,x_1 - 0.36412\,x_2 - 0.61386\,x_3 + 0.49860\,x_4 \end{cases}$$

$$女\begin{cases} z_1 = 0.49308\,x_1 + 0.52712\,x_2 + 0.48476\,x_3 + 0.49390\,x_4 \\ z_2 = 0.50209\,x_1 - 0.38887\,x_2 - 0.59266\,x_3 + 0.49540\,x_4 \end{cases}$$

〈 5学年 全体 및 男女学生別 主成分〉

$$全體\begin{cases} z_1 = 0.49784\,x_1 + 0.53275\,x_2 + 0.48122\,x_3 + 0.48656\,x_4 \\ z_2 = -0.46731\,x_1 + 0.36338\,x_2 + 0.61173\,x_3 - 0.52474\,x_4 \end{cases}$$

$$男\begin{cases} z_1 = 0.48884\,x_1 + 0.53567\,x_2 + 0.48132\,x_3 + 0.49234\,x_4 \end{cases}$$

$$女\begin{cases} z_1 = \quad 0.50821\,x_1 + 0.52915\,x_2 + 0.48156\,x_3 + 0.47938\,x_4 \\ z_2 = \quad 0.38576\,x_1 + 0.32862\,x_2 + 0.63087\,x_3 - 0.58752\,x_4 \end{cases}$$

〈 6 学年 全体 및 男女学生別 主成分 〉

$$全體\begin{cases} z_1 = \quad 0.54361\,x_1 + 0.53225\,x_2 + 0.36913\,x_3 + 0.53378\,x_4 \\ z_2 = -0.34881\,x_1 + 0.11471\,x_2 + 0.86015\,x_3 - 0.35399\,x_4 \end{cases}$$

$$男\begin{cases} z_1 = \quad 0.51650\,x_1 + 0.51552\,x_2 + 0.47068\,x_3 + 0.49589\,x_4 \\ z_2 = \quad 0.41910\,x_1 - 0.36044\,x_2 - 0.63436\,x_3 + 0.54034\,x_4 \end{cases}$$

$$女\begin{cases} z_1 = \quad 0.54282\,x_1 + 0.53978\,x_2 + 0.34695\,x_3 + 0.54183\,x_4 \\ z_2 = -0.34080\,x_1 + 0.05470\,x_2 + 0.89404\,x_3 - 0.28554\,x_4 \end{cases}$$

① 第 1 主成分 Z_1 의 係數는 모두 正이고, 大體로 0.3 ~ 0.5 의 範圍內에 들어 있다. 이 性質은 전체 · 男 · 女學生 어느 것에도 共通된다.

② 第 2 主成分 Z_2 의 係數는 모두 正 · 負가 섞여 있으므로 전체 · 남 · 여학생을 正 · 負로 되어 있다. (表 3·3 參照) 이 표에 의하면 正 · 負가 일정하게 일어나는 경우는 1·2·6 學年때이며, 3·4·5 學年일 때는 正 · 負가 섞여 있는 것으로 보아 1·2·6 학년때는 체중성장 속도가 비례적 경향을 나타내고, 3·4·5 학년 때는 거의 비례하지 않는다는 것이 나타난다.

따라서 1·2·6 학년 때는 길이로 나타내는 키와 앉은키의 경우는 負이나 둘레나 무게를 나타내는 가슴둘레와 몸무게는 正인 것으로 나타난다.

3·4·5 학년 때는 前者와는 달리 大部分 길이로 나타내는 키와 앉은키의 경우는 正으로 나타나며, 몸무게와 가슴둘레인 무게나 둘레를 나타내는 것은 負이다.

그러한 고로 키가 크고마른 사람은 第 1 主成分 Z_1 에서 正, 키가 작고 뚱뚱한 사람은 第 2 主成分 Z_2 에서 負의 絶對値가 크다고 할 수 있다.

表 3·3 第 2 主成 (z_2) 正 · 負 分類表

學年	性區分 \ 正負係數 \ 變數名	x_1 身長 正	x_1 身長 負	x_2 體重 正	x_2 體重 負	x_3 胸圍 正	x_3 胸圍 負	x_4 座高 正	x_4 座高 負
1	全 體		×	○		○			×
	男		×	○		○			×
	女		×	○		○			×
2	全 體		×	○		○			×
	男		×	○		○			×
	女		×	○		○			×
3	全 體	○			×		×	○	
	男		×	○		○			×
	女		×		×		×	○	
4	全 體		×	○		○			×
	男	○			×		×	○	
	女	○			×		×	○	
5	全 體		×	○		○			×
	男	○			×		×	○	
	女	○		○		○			×
6	全 體		×	○		○			×
	男	○			×		×	○	
	女		×	○		○			×

따라서 이 제 2 주성분 Z_2 에서는 3·4·5 년 때에 키와 앉은키가 거의 正인 것으로 보아 뚱뚱해지는비율이 크다고 해석을 내릴 수가 있다.

結論的으로 實際데이터를 主成分分析을 하여서 主成分이 단 1 개라도 導出되지 않았을 때 統計値는 실효가치가 없을 것이다. 그러나 本研究에 있어서는 以上과 같이 第 1, 第 2 主成分을 式 (2·2) 에 一致되므로 數學的 귀납이 증명되었으며, 上述한 바와 같은 (Z_1, Z_2 의 主成分方程式) 分析結果를 얻을 수 있다.

4. 結 語

本研究는 兒童들의 工産品이 단계적으로 規格化함에 있어서 의류, 자전거, 학생용 책 · 걸상 등의 높이에 참고가 될 것이다. 그 이외에 本研究의 結果가 앞으로 兒童들의 健康狀態를 알아보는데 다소나마 참고가 되기를 바라며, 나아가서 人間工學的인 産業디자인에 기여하기를 바란다.

參 考 文 獻

1) 具滋興外 2 人, 統計學 槪論, 서울 : 集賢社, 1973.

2) 金俊輔外 2 人, 數理統計學, 서울 : 高麗大學校 出版部, 1971.

3) 白雲鵬, 實驗計劃法, 서울 : 博英社, 1974.

4) 安恩明, 現代統計學, 서울 : 日新社, 1970.

5) 鄭英眞, 近代統計學 理論과 實際, 서울 : 寶晋社, 1971.

6) 具野忠外 3 人, 多變量解析法, 東京 : 日科技連, 1972.

7) 淺野啓三, 行列 · 行列式, 東京 : 共立出版, 1961.

8) 岩原信丸郎, 新教育統計法, 東京 : 日本文化科學社, 1962.

9) Hoel, *Iutroduction to Mathematical Statistics*, John Willy (2nd ed.) , New York : 1960.

10) Spigel, Murray R., *Schaum's Outline of theory and problems of statistics*, New York : Schaum Publishing Co., 1961.

11) Willy, John, *A Introduction to Probability theory and its Applications*, New York.

12) Scientific Subroutine Program, IBM.

附錄 表－1

	平均				標準偏差				固有値				固有値의 累積百分率			
	키 x_1	몸무게 x_2	가슴둘레 x_3	앉은키 x_4	키 x_1	몸무게 x_2	가슴둘레 x_3	앉은키 x_4	λ_1	λ_2	λ_3	λ_4	λ_1	λ_2	λ_3	λ_4
1 년 전체	116.17248	20.15076	56.59054	64.94205	4.47598	2.74363	3.10093	2.95407	2.95925	0.65136	0.26600	0.12336	0.73981	0.90265	0.96915	0.99999
2 〃	121.77886	22.51843	58.88143	67.60014	4.51809	3.78073	3.38786	2.43415	3.21695	0.57984	0.12889	0.1	0.80423	0.94920	0.98142	1.0
3 〃	127.48143	26.67506	61.45628	70.20009	5.39944	4.14976	4.13654	6.14924	2.66407	0.82836	0.43381	0.1	0.66601	0.87311	0.99156	1.0
4 〃	132.52456	28.562495	63.98081	72.13889	5.55242	4.97438	4.85399	2.77609	3.13824	0.66356	0.14084	0.1	0.78456	0.95045	0.98566	1.0
5 〃	138.06019	31.99295	66.76041	74.20690	5.68413	6.28901	6.04776	2.85227	3.10493	0.70351	0.14732	0.1	0.77623	0.95211	0.98874	1.0
6 〃	144.30502	36.14133	68.00147	77.69374	6.21522	6.79532	6.66302	3.51539	2.69960	0.80072	0.32558	0.14707	0.67490	0.87508	0.95047	0.99999
1 년 남	116.90455	20.77412	56.71630	65.61506	4.66404	2.94576	3.22555	3.04032	2.99335	0.66248	0.26036	0.12378	0.73833	0.90396	0.96905	0.99999
2 〃	122.89813	23.19253	59.99095	68.08221	4.35628	3.37288	3.52166	2.40468	3.23942	0.55909	0.14386	0.1	0.80985	0.94962	0.98559	1.0
3 〃	127.81730	26.16528	62.34467	70.13533	5.42314	4.27763	4.20178	2.69502	3.16835	0.62617	0.13892	0.1	0.79208	0.94503	0.98336	1.0
4 〃	133.29330	29.03168	64.68263	72.48924	5.52139	5.03026	5.00230	2.71202	3.15428	0.64968	0.14392	0.1	0.78857	0.95099	0.98697	1.0
5 〃	138.18130	32.25527	67.14274	74.22860	5.47016	5.29530	6.07304	2.75150	3.00403	0.80548	0.15076	0.1	0.75100	0.95238	0.99007	1.0
6 〃	143.67245	35.94792	68.50962	77.22264	6.31046	6.68196	5.34506	3.29309	3.09114	0.63303	0.16074	0.11505	0.77272	0.93104	0.97123	0.99999
1 년 남	115.44041	19.52749	54.46479	64.26895	4.74913	2.23410	2.35034	2.63940	2.84916	0.74754	0.27154	0.13174	0.71229	0.89917	0.96706	0.99999
2 〃	120.65960	21.84433	57.77192	57.11807	4.43380	2.71458	2.71717	2.37277	3.27790	0.70258	0.13170	0.1	0.76947	0.94512	0.97804	1.0
3 〃	127.14566	25.18484	60.56790	70.26486	5.36013	3.89829	3.84213	9.10317	2.58142	0.88346	0.45229	0.1	0.64535	0.86622	0.97929	1.0
4 〃	131.75582	28.09331	63.27899	71.78854	5.49144	4.86340	4.53900	2.82085	3.08824	0.70948	0.14194	0.1	0.77206	0.94943	0.98491	1.0
5 〃	137.93908	31.73063	66.37809	74.18521	5.99928	6.29242	6.00683	3.00296	3.24329	0.57189	0.13565	0.1	0.81082	0.95379	0.98771	1.0
6 〃	144.93759	36.33475	67.49333	78.16484	6.04942	6.94675	3.02207	3.71150	2.57822	0.82848	0.35439	0.23888	0.64455	0.85167	0.94027	0.99999

附錄 表 - 2

	상관계수				固有 Vector				成分
	키 x_1	몸무게 x_2	가슴둘레 x_3	앉은키 x_4	λ_1	λ_2	λ_3	λ_4	
1 년전체	1.00000	0.72714	0.50152	0.74486	0.50418	0.53979	0.47157	0.48170	
2 〃	1.00000	0.75275	0.58840	0.87109	0.50011	0.52710	0.47471	0.49604	
3 〃	1.00000	0.73533	0.56444	0.37499	0.57370	0.58290	0.54065	0.32251	
4 〃	1.00000	0.69292	0.55360	0.85957	0.49444	0.52715	0.48406	0.49326	
5 〃	1.00000	0.70046	0.54099	0.85170	0.49784	0.53275	0.48122	0.48656	
6 〃	1.00000	0.69717	0.32657	0.91986	0.54361	0.53225	0.36913	0.53378	
1 년 남	1.00000	0.71609	0.52207	0.74574	0.50634	0.53676	0.47872	0.47572	第 1
2 〃	1.00000	0.16546	0.60443	0.85573	0.49831	0.52976	0.47769	0.49279	
3 〃	1.00000	0.74849	0.56477	0.80081	0.50009	0.53041	0.47454	0.49332	
4 〃	1.00000	0.70857	0.55626	0.85671	0.49452	0.52905	0.48257	0.49261	
5 〃	1.00000	0.63107	0.46436	0.84919	0.48884	0.53567	0.48132	0.49234	
6 〃	1.00000	0.71383	0.58610	0.87967	0.51650	0.51552	0.47068	0.49589	
1 년 여	1.00000	0.74646	0.42883	0.73093	0.51640	0.55474	0.43962	0.48198	主成分
2 〃	1.00000	0.70905	0.48046	0.88000	0.50145	0.53071	0.45992	0.50533	
3 〃	1.00000	0.73023	0.56541	0.25425	0.51615	0.59428	0.55746	0.26390	
4 〃	1.00000	0.66290	0.52361	0.85831	0.49308	0.52712	0.48476	0.49390	
5 〃	1.00000	0.79350	0.64455	0.85486	0.50821	0.52915	0.48156	0.47938	
6 〃	1.00000	0.65841	0.26615	0.75475	0.54282	0.53978	0.34695	0.54183	
1 년전체	0.72714	1.00000	0.81659	0.63505	−0.42655	0.28727	0.67178	−0.53311	
2 〃	0.75275	1.00000	0.89317	0.74112	−0.47153	0.32224	0.65564	−0.49390	
3 〃	0.73553	1.00000	0.89459	0.33640	0.01609	−0.23127	−0.31495	0.92035	
4 〃	0.69292	1.00000	0.92390	0.68390	−0.49510	0.37721	0.00316	−0.49875	
5 〃	0.70046	1.00000	0.93119	0.66707	−0.46731	0.36388	0.01173	−0.52474	
6 〃	0.69717	1.00000	0.51497	0.63876	−0.34881	0.11471	0.86015	−0.35399	
1 년 남	0.71609	1.00000	0.83055	0.60604	−0.40702	0.32208	0.63640	−0.57059	第 2
2 〃	0.76548	1.00000	0.90854	0.75255	−0.46244	0.31371	0.65680	−0.50628	
3 〃	0.73848	1.00000	0.90004	0.71099	−0.46538	0.32940	0.64765	−0.50541	
4 〃	0.73857	1.00000	0.92527	0.69594	0.49190	−0.36412	−0.61386	0.49860	
5 〃	0.63107	1.00000	0.93745	0.63733	0.50788	−0.37835	−0.59764	0.49165	
6 〃	0.71383	1.00000	0.81921	0.55289	0.41910	−0.36048	−0.63436	0.54034	
1 년 여	0.74646	1.00000	0.75367	0.64110	−0.36906	0.21183	0.73893	−0.52238	主成分
2 〃	0.70905	1.00000	0.85556	0.69999	−0.49176	0.31820	0.67003	−0.45600	
3 〃	0.73023	1.00000	0.88711	0.25904	−0.08720	−0.17719	−0.18599	0.96249	
4 〃	0.66290	1.00000	0.92321	0.66055	0.50209	−0.38887	−0.59206	0.49540	
5 〃	0.79350	1.00000	0.92215	0.70864	0.38576	0.32862	0.63087	−0.58752	
6 〃	0.65941	1.00000	0.44165	0.62983	−0.34080	0.05470	0.89404	−0.28554	

	상 관 계 수				固 有				成 分
	키 x_1	몸무게 x_2	가슴둘레 x_3	앉은키 x_4	λ_1	λ_2	λ_3	λ_4	
1 년전체	0.50152	0.81659	1.00000	0.47806	−0.64685	−0.21726	0.23224	0.69313	
2 〃	0.58840	0.89317	1.00000	0.57873	−0.70657	−0.02899	0.03794	0.70602	
3 〃	0.56444	0.89459	1.00000	0.27366	−0.82061	0.12632	0.51153	0.22116	
4 〃	0.55360	0.92390	1.00000	0.55607	−0.69930	−0.05309	0.04733	0.71127	
5 〃	0.54099	0.93119	1.00000	0.50870	−0.71840	−0.00855	0.05131	0.69367	
6 〃	0.32657	0.51497	1.00000	0.32928	0.14853	−0.82136	0.34608	0.42839	
1 년 남	0.52270	0.83055	1.00000	0.47284	−0.67103	−0.19266	0.26475	0.66519	第 3
2 〃	0.60443	0.90854	1.00000	0.58285	−0.72258	0.02803	0.01021	0.69063	
3 〃	0.56477	0.90004	1.00000	0.54560	−0.71024	−0.02912	0.05183	0.70143	
4 〃	0.55628	0.92527	1.00000	0.55822	−0.69737	−0.06218	0.05621	0.71179	
5 〃	0.46436	0.93745	1.00000	0.47817	−0.69972	−0.02555	0.00884	0.71390	
6 〃	0.58610	0.81921	1.00000	0.52273	0.01059	−0.76400	0.01282	0.20155	
1 년 여	0.42883	0.75367	1.00000	0.35975	−0.65352	−0.17656	0.22061	0.70218	主成分
2 〃	0.48046	0.85556	1.00000	0.51586	−0.51486	−0.42050	0.29028	0.68834	
3 〃	0.56541	0.88711	1.00000	0.23706	−0.81501	0.16127	0.55300	0.06270	
4 〃	0.52861	0.92321	1.00000	0.53650	−0.69868	−0.04654	0.03466	0.71306	
5 〃	0.64455	0.92215	1.00000	0.55228	−0.74534	−0.01212	0.15442	0.64843	
6 〃	0.26615	0.44165	1.00000	0.30395	0.76936	−0.82136	0.26696	0.47656	
1 년전체	0.74486	0.63505	0.47806	1.00000	0.38135	−0.76084	0.52189	−0.05746	
2 〃	0.87108	0.74112	0.57873	1.00000	0.00000	−0.00000	0.00000	0.00000	
3 〃	0.37498	0.33640	0.27366	1.00000	0.00000	0.00000	0.00000	0.00000	
4 〃	0.85957	0.69390	0.55607	1.00000	0.00000	0.00000	0.00000	0.00000	
5 〃	0.85170	0.66707	0.50870	1.00000	0.00000	0.00000	0.00000	0.00000	
6 〃	0.81986	0.63876	0.32928	1.00000	−0.74883	0.17002	0.06403	0.63736	
1 년 남	0.74574	0.60504	0.47284	1.00000	0.35728	−0.75565	0.54380	−0.07488	第 4
2 〃	0.85573	0.75268	0.58285	1.00000	0.00000	0.00000	0.00000	0.00000	
3 〃	0.86081	0.71699	0.54500	1.00000	0.00000	0.00000	0.00000	0.00000	
4 〃	0.85671	0.69594	0.55822	1.00000	0.00000	0.00000	0.00000	0.00000	
5 〃	0.84919	0.63733	0.47817	1.00000	0.00000	0.00000	0.00000	0.00000	
6 〃	0.87967	0.65289	0.52273	1.00000	0.74663	0.14344	−0.02178	0.64922	
1 년 여	0.73093	0.64110	0.35975	1.00000	−0.41283	0.78499	−0.46046	0.04170	主成分
2 〃	0.88000	0.69999	0.51586	1.00000	0.00000	0.00000	0.00000	0.00000	
3 〃	0.25425	0.25904	0.23706	1.00000	0.00000	0.00000	0.00000	0.00000	
4 〃	0.85831	0.66065	0.53650	1.00000	0.00000	0.00000	0.00000	0.00000	
5 〃	0.85486	0.70864	0.55228	1.00000	0.00000	0.00000	0.00000	0.00000	
6 〃	0.75475	0.67983	0.30395	1.00000	−0.74866	0.18090	0.09502	0.63066	

附錄 3. 컴퓨터에 의한 데이터의 **處理過程**

IBM System／360 Scientific Subroutine Package