IV

The Indeterminacy Property of Linear Factor Representations

1. *Introduction*

To date, the Central Account has met with very little opposition. Its tenets are re-stated with regularity in countless articles and texts on the theory and application of latent variable modeling. Seldom are these tenets subjected to scrutiny. And yet, there have been dissenters. Very shortly after the birth of factor analysis and classical true-score theory, the American mathematician E.B. Wilson, in a series of reviews (1928a, 1928b, 1929) of the factor theories of Charles Spearman and Truman Kelley, asked pointed questions about certain of what are, herein, identified as tenets of the Central Account. One of Wilson's concerns was that, when a set of tests is described by the linear factor model, the factor scores referred to in the model's equations are not uniquely defined. Spearman's work preceded the extensive use of "unobservability talk" that, nowadays, characterizes latent variable modeling. Wilson quite reasonably read Spearman's stated aim to "objectively determine *g*" as the aim of producing a set of scores, one for each individual under study, that possessed the properties that the two-factor theory claimed that *g*-scores should possess. Thus, in his initial review of Spearman's *Abilities*, he complained that Spearman had not provided an example of a factor analysis worked right down to the scores. Taking matters into his own hands, Wilson used fictitious data to demonstrate the construction of *g*-, or common factor-, scores. In doing so, he also established that there could be constructed *more* than one such set of scores, each set possessing all of the properties of common-factor scores. That is, there did not exist a unique referent of the concept *common factor to* $\underline{\mathbf{X}}$. Wilson quantified the degree of non-uniqueness inherent to his fictitious data by taking the difference between the two most dissimilar sets of common-factor scores. The non-uniqueness property came to be known as the indeterminacy property of the linear factor model, and, over the years, many psychometricians, among them Spearman, Thomson, Camp, Piaggio, Ledermann, Kestelman, Heermann, Guttman, Schonemann, Steiger, Rozeboom, McDonald, and Mulaik, have devoted effort to its clarification. Steiger and Schonemann (1978) provided a comprehensive history of the study of the indeterminacy property.

Whereas the Central Account portrays the referent of *latent variate to* $\underline{\mathbf{X}}$ as a detected, but unobservable, property/attribute (causal source) existing in nature, the indeterminacy property has been interpreted by some as implying that what is signified by this concept are, instead, the members of a set of constructed random variates, ontologically on par with any synthetic or component variate. If this latter interpretation is correct, it would then seem to contradict the most fundamental features of the Central Account, for a constructed random variate certainly cannot be an unobservable cause or measured property of the phenomena represented by a set of manifest variates. Without the Central Account, however, latent variable modeling loses its charm. Not surprisingly, then, there has been no dearth of experts who have felt compelled to step forward to do damage control. And in their responses to the perceived threat, one seldom sees a concern for the correct characterization of latent variable models and modeling, but rather with the protection, at all costs, of the Central Account. In this chapter, the mathematics of indeterminacy are reviewed, and, in the next chapter, the lengthy series of exchanges over the interpretation and implications of the mathematics, which constitute the "indeterminacy debate" of linear factor analysis.

2. *The Mathematics of Indeterminacy*

2a. What the linear factor model says about the factors $\boldsymbol{\theta}$ and $\underline{\boldsymbol{\delta}}$

For particular set of manifest variates, $\mathbf{X}_j$, j=1..p, distributed over a particular population, P_T, and $\underline{\Omega}_T$, the $1\times\frac{1}{2}p(p+1)$ vector containing the non-redundant elements of Σ_T, assume that $\underline{\Omega}_T \subset M_{ulf}$, i.e., $\underline{\mathbf{X}}$ is described by the unidimensional linear factor model. Assume also that the values of the model parameters for P_T are $\underline{\Pi}_T=[\underline{\Lambda}_o,\Psi_o]$. In this case, $\underline{\mathbf{X}}$ will then be said to be *ulcf representable (in P_T)*, and $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$ will be called its ulcf representation. Model (2.8)-(2.9) symbolizes the factors as $\boldsymbol{\theta}$ and $\underline{\boldsymbol{\delta}}$. In contrast to the manifest variates, however, the "meaning" of these symbols is left to be settled by the equations and associated distributional constraints standardly given as "the model", and, of course, the verbal accompaniment that is the Central Account. The equations and distributional specifications alone merely claim that any set of (p+1) random variates, $\{\mathbf{Y},\underline{\mathbf{Z}}\}$, satisfying the requirements

(4.1) $$\underline{\mathbf{X}} = [\underline{\boldsymbol{\Lambda}}_o : \boldsymbol{\Psi}_o^{1/2}]\begin{bmatrix}\mathbf{Y}\\ \underline{\mathbf{Z}}\end{bmatrix}$$

(4.2) $$E\begin{bmatrix}\mathbf{Y}\\ \underline{\mathbf{Z}}\end{bmatrix} = \underline{0},$$

and

(4.3) $$E\begin{bmatrix}\mathbf{Y}\\ \underline{\mathbf{Z}}\end{bmatrix}[\mathbf{Y} : \underline{\mathbf{Z}}'] = \begin{bmatrix}1 & \underline{0}'\\ \underline{0} & I\end{bmatrix}$$

are factors ($\mathbf{Y}$ a common factor, and $\underline{\mathbf{Z}}$, a vector of p specific factors) to $\underline{\mathbf{X}}$. Let *C* be the set which contains all such common factors to $\underline{\mathbf{X}}$. The symbol $\boldsymbol{\theta}$ in the model equations then stands for any of the elements of *C*. Two issues immediately present themselves. First, can there be found a general construction formula for the common factors contained within *C*? Second, what is the cardinality of the set *C*? With regard the first issue, Piaggio (1931, 1933, 1935), working with finite dimensional score arrays, rather than random variates, derived construction formulas for the production of variates that satisfy (4.1)-(4.3). In 1933, he proved the sufficiency of these formulas, i.e., that any scores constructed according to these equations will fit the factor model, and, in 1935, their necessity, i.e., that scores satisfying the factor model are always expressible in terms of these formulas. Kestelman (1952) generalized Piaggio's formulas to the orthogonal, multiple-factor case. Guttman (1955), situating his work within an abstract Euclidean vector space, derived construction formulas applicable to both the sample (data analytic) and population scenarios, and to the case of the oblique multiple factor model, and proved the necessity and sufficiency of these formulas. In regard the second issue, Wilson (1928a; 1928b), once again working with finite dimensional score arrays, originally established that *C* is of infinite cardinality. This property of the linear factor model, i.e., that *C* is of infinite cardinality, is called the indeterminacy property of the linear factor model.

comment: An important consequence of the indeterminacy property, and one that has notational significance, is that any statistical statement involving $\boldsymbol{\theta}$ is shorthand for the set of equivalent statements about each of the elements of *C*. For example, to state that $E(\boldsymbol{\theta}\underline{\mathbf{X}})=\underline{\Lambda}_o$, is to state that "for all $\mathbf{Y}\in C$, $E(\mathbf{Y}\underline{\mathbf{X}})=\underline{\Lambda}_o$".

2b. Constructed factors to $\underline{\mathbf{X}}$

In terms of random variates, the Piaggio-Kestelman-Guttman construction formulas are (for ulf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$)

(4.4) $$\boldsymbol{\theta}_i=\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$$

and

(4.5) $$\underline{\boldsymbol{\delta}}_i=\Psi_o^{1/2}\Sigma^{-1}\underline{\mathbf{X}}-\Psi_o^{-1/2}\underline{\Lambda}_o w^{1/2}\mathbf{s}_i,$$

in which $w=(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)$, and

(4.6) $$C(\mathbf{s}_i,\underline{\mathbf{X}})=\underline{0},\ E(\mathbf{s}_i)=0,\ V(\mathbf{s}_i)=1.$$

Each set of constructed random variates $\{\boldsymbol{\theta},\underline{\boldsymbol{\delta}}\}_i$ is produced by using (4.4)-(4.6) with a particular choice of a random variate $\mathbf{s}_i$ that satisfies (4.6). This is why a given construction is indexed with the subscript i, and suggests that, rather than the $\{\boldsymbol{\theta},\underline{\boldsymbol{\delta}}\}_i$ notation, it might be more suggestive to employ $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}_{(si)}$. Any set of random variates constructed in accord with (4.4)-(4.6) are sufficient for factor-hood, meaning that they do indeed possess all properties required to be called factors to $\underline{\mathbf{X}}$. In particular,

(4.7) $$E(\boldsymbol{\theta}_i)=E(\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)=\underline{\Lambda}_o'\Sigma^{-1}E(\underline{\mathbf{X}})+w^{1/2}E(\mathbf{s}_i)=\underline{\Lambda}_o'\Sigma^{-1}\underline{0}+w^{1/2}*0=0,$$

(4.8) $$E(\underline{\boldsymbol{\delta}}_i)=E(\Psi_o^{1/2}\Sigma^{-1}\underline{\mathbf{X}}-\Psi_o^{-1/2}\underline{\Lambda}_o w^{1/2}\mathbf{s}_i)=\Psi_o^{1/2}\Sigma^{-1}E(\underline{\mathbf{X}})-\Psi_o^{-1/2}\underline{\Lambda}_o w^{1/2}E(\mathbf{s}_i)$$

$$=\Psi_o\Sigma^{-1}\underline{0}-\underline{\Lambda}_o w^{1/2}*0=\underline{0},$$

(4.9) $$V(\boldsymbol{\theta}_i)=V(\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)=\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o+wV(\mathbf{s}_i)+2w^{1/2}\underline{\Lambda}_o'\Sigma^{-1}C(\underline{\mathbf{X}},\mathbf{s}_i)$$

$$=\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o+1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o=1,$$

(4.10) $$C(\underline{\boldsymbol{\delta}}_i)=C(\Psi_o^{1/2}\Sigma^{-1}\underline{\mathbf{X}}-\Psi_o^{-1/2}\underline{\Lambda}_o w^{1/2}\mathbf{s}_i)$$

$$=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}+w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}-w^{1/2}\Psi_o^{1/2}\Sigma^{-1}C(\underline{\mathbf{X}},\mathbf{s}_i)\underline{\Lambda}_o'\Psi_o^{-1/2}$$

$$=-w^{1/2}\Psi_o^{-1/2}\underline{\Lambda}_o C(\mathbf{s}_i,\underline{\mathbf{X}}')\Sigma^{-1}\Psi_o^{1/2}=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}+w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$$

$$=\Psi_o^{-1/2}(\Sigma-\underline{\Lambda}_o\underline{\Lambda}_o')\Sigma^{-1}(\Sigma-\underline{\Lambda}_o\underline{\Lambda}_o')\Psi_o^{-1/2}+(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$$

$$=\Psi_o^{-1/2}(\Sigma-2\underline{\Lambda}_o\underline{\Lambda}_o'+(\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\underline{\Lambda}_o\underline{\Lambda}_o'+(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\underline{\Lambda}_o\underline{\Lambda}_o')\Psi_o^{-1/2}$$

$$=\Psi_o^{-1/2}(\Sigma-\underline{\Lambda}_o\underline{\Lambda}_o')\Psi_o^{-1/2}=I,$$

(4.11)
$$C(\underline{\boldsymbol{\delta}}_i,\boldsymbol{\theta}_i)=C(\Psi_o^{1/2}\Sigma^{-1}\underline{\mathbf{X}}-\Psi_o^{-1/2}\underline{\Lambda}_o w^{1/2}\mathbf{s}_i,\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)$$

$$=\Psi_o^{1/2}\Sigma^{-1}\underline{\Lambda}_o+w^{1/2}\Psi_o^{1/2}\Sigma^{-1}C(\underline{\mathbf{X}},\mathbf{s}_i)-w^{1/2}\Psi_o^{-1/2}\underline{\Lambda}_o C(\mathbf{s}_i,\underline{\mathbf{X}})\Sigma^{-1}\underline{\Lambda}_o-w\Psi_o^{-1/2}\underline{\Lambda}_o V(\mathbf{s}_i)$$

$$=\Psi_o^{1/2}\Sigma^{-1}\underline{\Lambda}_o-(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\Psi_o^{-1/2}\underline{\Lambda}_o$$

$$=\Psi_o^{-1/2}(\Sigma-\underline{\Lambda}_o\underline{\Lambda}_o')\Sigma^{-1}\underline{\Lambda}_o-(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\Psi_o^{-1/2}\underline{\Lambda}_o$$

$$=\Psi_o^{-1/2}\underline{\Lambda}_o-(\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\Psi_o^{-1/2}\underline{\Lambda}_o-\Psi_o^{-1/2}\underline{\Lambda}_o+(\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\Psi_o^{-1/2}\underline{\Lambda}_o=\underline{0},$$

and

(4.12)
$$\underline{\Lambda}_o\boldsymbol{\theta}_i+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}_i$$

$$=\underline{\Lambda}_o(\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i)+\Psi_o^{1/2}(\Psi_o^{1/2}\Sigma^{-1}\underline{\mathbf{X}}-\Psi_o^{-1/2}\underline{\Lambda}_o w^{1/2}\mathbf{s}_i)$$

$$=\underline{\Lambda}_o\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+\underline{\Lambda}_o w^{1/2}\mathbf{s}_i+(\Sigma-\underline{\Lambda}_o\underline{\Lambda}_o')\Sigma^{-1}\underline{\mathbf{X}}-\underline{\Lambda}_o w^{1/2}\mathbf{s}_i$$

$$=\underline{\Lambda}_o\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+\underline{\Lambda}_o w^{1/2}\mathbf{s}_i+\underline{\mathbf{X}}-\underline{\Lambda}_o\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}-\underline{\Lambda}_o w^{1/2}\mathbf{s}_i=\underline{\mathbf{X}}.$$

As was noted, Piaggio, Kestelman, and Guttman also proved that (4.4)-(4.6) are necessary conditions for factor-hood to $\underline{\mathbf{X}}$. One must, however, be careful in regard the meaning of this. It does not mean that all factors to $\underline{\mathbf{X}}$ must necessarily be constructed random variates (although, it will be argued in Part II that this is precisely the case), but, instead, that any factors to $\underline{\mathbf{X}}$ must be *expressible* as in (4.4) with (4.6). The results of Piaggio, Kestelman, and Guttman, then, give:

Theorem 1 (Definition of factors to $\underline{\mathbf{X}}$; Piaggio, 1933; Kestelman, 1952; Guttman, 1955): Because formulas (4.4)-(4.6) are both a necessary and sufficient for a set of random variates $\{\mathbf{Y},\underline{\mathbf{Z}}\}$ to be factors to $\underline{\mathbf{X}}$ ($\mathbf{Y}$ the common factor, $\underline{\mathbf{Z}}$ the specifics), it follows that the concept *common factor to* $\underline{\mathbf{X}}$ denotes a random variate expressible as in (4.4) with (4.6), and the concept *specific factors to* $\underline{\mathbf{X}}$ denotes a random vector expressible as in (4.5) with (4.6). Set *C* contains just those random variates expressible as in (4.4)-(4.6).

Note, however, that, if $\underline{\mathbf{X}}$ is ulcf representable, one could, in fact, go ahead and choose an $\mathbf{s}_i$ with properties (4.6), and actually construct a random variate as per (4.4), and this constructed random variate would be, by definition, a common factor to $\underline{\mathbf{X}}$. How $\mathbf{s}_i$ would be chosen is not relevant. The symbol $\mathbf{s}_i$ could, for example, stand for the ith random number generator in a large set of such generators. As Wilson (1928a; 1928b), and others, have shown, an endless sequence of such random variates could be produced, and each would be definitionally a common factor to $\underline{\mathbf{X}}$. It would seem, then, fair enough to conclude that:

i.) The set *C* containing the common factors to $\underline{\mathbf{X}}$ contains an infinity of constructed random variates;

ii) If $\underline{\mathbf{X}}$ is ulcf representable, it possesses an infinity of common factors. Hence, there are an infinity of referents of the concept *common factor to* $\underline{\mathbf{X}}$, and this set of referents contains constructed random variates;

iii) Because constructed random variates are contained in *C*, realizations of the common (and, by (4.5) and (4.6), the specific) factors to $\underline{\mathbf{X}}$ can be taken. Hence, the phrase "common factors can only be estimated, but not determined", and variants thereof, are incorrect;

iv) At least some of the factors to $\underline{\mathbf{X}}$, those constructed in accord with (4.4)-(4.6), are not, then, "unobservable", "unknown", or "unmeasureable". Hence, it cannot rightly be said that the "common factor to $\underline{\mathbf{X}}$ is unobservable", at least if, by *unobservable* random variate, one means a random variate on which realizations cannot be taken.[1]

Certainly, a number of prominent psychometricians have drawn these conclusions from a consideration of the indeterminacy property of ulcf representations. However, as will be seen in the next chapter, there has also been a great deal of opposition to these conclusions. This is not surprising, for these conclusions, and what they suggest about linear factor analysis, and, by implication, latent variable analysis in general, are in conflict with the Central Account, and, as was shown in the previous chapter, the discipline of psychometrics is deeply committed to the Central Account. It has been shown that set *C* contains constructed random variates. Whether it can contain anything other than constructed random variates, and, in particular, random variates whose distributions are comprised of measurements with respect a detected unobservable property/attribute (causal source) of the phenomena represented by $\underline{\mathbf{X}}$, these random variates merely *expressible* as in (4.4)-(4.6), will later be investigated. For this, of course, is the issue of the correctness of the Central Account. As already indicated, the aim will be to show that the CA is nonsense.

2c. Indeterminacy property of ulcf representations

As was stated previously, Wilson established that the cardinality of *C* is infinite. In this section, the mathematics that support this conclusion will be reviewed.

Definition (Indeterminate Representation). Let $\underline{\mathbf{X}}$ be ulcf representable, and let $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$ be its ulcf representation. If the cardinality of *C* is greater than unity, the representation is said to be indeterminate. If a representation is not indeterminate, then it is determinate.

It is, of course, possible that a given representation is indeterminate, but that certain of the (p+1) factor variates $\{\boldsymbol{\theta},\underline{\boldsymbol{\delta}}\}$ are individually determinate, meaning that, for each of these variates, the same construction appears in all sets $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}_{(si)}$ that can be constructed. It will, therefore, be necessary to speak, not only in terms of the determinateness of the representation, but also of the individual factors.

---

1 Of course, as was seen in Chapter 3, *unobservability* is taken by latent variable modellers as meaning something quite different from this.

Definition (Indeterminate Factor Variate). Let $\underline{\mathbf{X}}$ be ulcf representable, and let $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$ be its ulf representation. If the same construction $\boldsymbol{\theta}_i$ (or a particular $\boldsymbol{\delta}_{ji}$) appears in all sets $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}_{(\mathbf{si})}$ that can be constructed, then $\boldsymbol{\theta}_i$ (or a particular $\boldsymbol{\delta}_{ji}$) is determinate. Otherwise, it is indeterminate.

Following Wilson (1928), the first terms of the construction formulas (4.4) and (4.5) may be symbolized $\mathbf{D}_{\boldsymbol{\theta}}$ and $\underline{\mathbf{D}}_{\boldsymbol{\delta}}$, respectively, and the second terms as $\mathbf{I}_{\boldsymbol{\theta}(i)}$ and $\underline{\mathbf{I}}_{\boldsymbol{\delta}(i)}$. That is,

(4.13) $$\boldsymbol{\theta}_i=\mathbf{D}_{\boldsymbol{\theta}}+\mathbf{I}_{\boldsymbol{\theta}(i)}$$

and

(4.14) $$\underline{\boldsymbol{\delta}}_i=\underline{\mathbf{D}}_{\boldsymbol{\delta}}+\underline{\mathbf{I}}_{\boldsymbol{\delta}(i)}.$$

Now, $\mathbf{D}_{\boldsymbol{\theta}}$ and $\underline{\mathbf{D}}_{\boldsymbol{\delta}}$ are linear transformations of the manifest variates, while $\mathbf{I}_{\boldsymbol{\theta}(i)}$ and $\underline{\mathbf{I}}_{\boldsymbol{\delta}(i)}$ are not, they each being functions of a random variate that must merely satisfy the weak requirements of (4.6). Clearly, $\mathbf{D}_{\boldsymbol{\theta}}$ and $\underline{\mathbf{D}}_{\boldsymbol{\delta}}$ are constant over the sets of common and specific factors to $\underline{\mathbf{X}}$, respectively, while different "**I**" components enter into each distinct construction. Required distributional properties, e.g., those required by model (2.14)-(2.15), are imparted to $\boldsymbol{\theta}_i$ and $\underline{\boldsymbol{\delta}}_i$ by choosing $\mathbf{s}_i$ to have, in addition to the properties (4.6), particular distributional properties, since the distributions of $\boldsymbol{\theta}_i$ and $\underline{\boldsymbol{\delta}}_i$ are determined jointly by $\underline{\mathbf{X}}$ and $\mathbf{s}_i$. It is clear from construction formulas (4.4) and (4.5) that a given representation is determinate only if $V(\mathbf{I}_{\boldsymbol{\theta}(i)})=0$ and $V(\mathbf{I}_{\boldsymbol{\delta}j(i)})=0\ \forall i,j$, and that individual factor k is determinate only if the variate in the kth row of the (p+1)-vector

(4.15) $$\begin{bmatrix}\mathbf{I}_{\boldsymbol{\theta}(i)}\\ \mathbf{I}_{\boldsymbol{\delta}(i)}\end{bmatrix}$$

has variance equal to zero. Hence, if a representation is determinate, $\underline{\mathbf{X}}$ possesses but one set of factors, $\begin{bmatrix}\boldsymbol{\theta}\\ \underline{\boldsymbol{\delta}}\end{bmatrix}=\begin{bmatrix}\underline{\Lambda}_o\Sigma^{-1}\\ \Psi_o^{1/2}\Sigma^{-1}\end{bmatrix}\underline{\mathbf{X}}$. It must, therefore, be investigated if and when the variances of the variates (4.15) are equal to zero. To do so, it is covenient to define a number of quantities.

2d. Decompositions

As is well known, under model (2.8)-(2.9)

(4.16) $$E(\boldsymbol{\theta}|\underline{\mathbf{X}}=\underline{x})=\underline{\Lambda}_o'\Sigma^{-1}\underline{x}=\hat{\boldsymbol{\theta}},$$

and, hence,

(4.17) $$V(\hat{\boldsymbol{\theta}})=\underline{\boldsymbol{\Lambda}}_o'\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{\Lambda}}_o.$$

From (4.16) and (4.17), the squared multiple correlation between $\boldsymbol{\theta}$ and $\underline{\mathbf{X}}$ is

(4.18) $$R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}=\frac{C(\hat{\boldsymbol{\theta}},\boldsymbol{\theta})^2}{V(\hat{\boldsymbol{\theta}})V(\boldsymbol{\theta})}=\frac{C(\underline{\boldsymbol{\Lambda}}_o'\boldsymbol{\Sigma}^{-1}\underline{\mathbf{X}},\boldsymbol{\theta})^2}{\underline{\boldsymbol{\Lambda}}_o'\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{\Lambda}}_o}=\frac{(\underline{\boldsymbol{\Lambda}}_o'\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{\Lambda}}_o)^2}{\underline{\boldsymbol{\Lambda}}_o'\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{\Lambda}}_o}=\underline{\boldsymbol{\Lambda}}_o'\boldsymbol{\Sigma}^{-1}\underline{\boldsymbol{\Lambda}}_o,$$

the conditional variance of $\boldsymbol{\theta}$ given $\underline{\mathbf{X}}=\underline{x}$ is

(4.19) $$V(\boldsymbol{\theta}|\underline{\mathbf{X}}=\underline{x})=V(\boldsymbol{\theta})(1-R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}})=(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)=w,$$

and, hence, the expectation (over the range of $\underline{\mathbf{X}}$) of $V(\boldsymbol{\theta}|\underline{\mathbf{X}}=\underline{x})$ is

(4.20) $$E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))=w.$$

Analogously,

(4.21) $$E(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}=\underline{x})=\Psi_o^{1/2}\Sigma^{-1}\underline{x},$$

(4.22) $$C(E(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}))=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2},$$

(4.23) $$C(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}=\underline{x})=I-\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2},$$

and

(4.24) $$E(C(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}))=I-\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}.$$

From these identities are defined the following decompositions:

(4.25) $$\begin{aligned}V(\boldsymbol{\theta})&=V(E(\boldsymbol{\theta}|\underline{\mathbf{X}}))+E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))\\&=\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o+(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)\\&=\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o+w\end{aligned}$$

and

(4.26) $$\begin{aligned}C(\underline{\boldsymbol{\delta}})&=C(E(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}))+E(C(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}))\\&=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}+(I-\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2})\\&=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}+\Psi_o^{1/2}(\Psi_o^{-1}-\Sigma^{-1})\Psi_o^{1/2}\\&=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}+w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}\end{aligned}$$

From (4.13), (4.14), (4.25), and (4.26), it is then seen that $V(\mathbf{D}_{\boldsymbol{\theta}})=\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o$, $V(\mathbf{I}_{\boldsymbol{\theta}(i)})=w$,

$C(\underline{\mathbf{D}}_{\boldsymbol\delta})=\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}$, and $C(\underline{\mathbf{I}}_{\delta(i)})=w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$.

2e. Conditions for Determinacy

From (4.25) and (4.26), the ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol\theta+\Psi_o^{1/2}\underline{\boldsymbol\delta}$ is determinate only if $V(\mathbf{I}_{\theta(i)})=w=0$ and all diagonal elements of $C(\underline{\mathbf{I}}_{\delta(i)})=w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$ are equal to zero, which, because $C(\underline{\mathbf{I}}_{\delta(i)})$ is gramian, implies that $w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$ must be a null matrix. The common factor $\boldsymbol\theta$ is determinate only if w=0, and the jth specific factor, $\boldsymbol\delta_j$, only if either: 1) w=0; 2) $\lambda_j=0$; or 3) w=0 and $\lambda_j=0$. Let

(4.27) $$t=\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o,$$

and

(4.28) $$q=(1+\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o)=1+t.$$

Then (Guttman, 1955, result (60))

(4.29) $$V(\mathbf{I}_{\theta(i)})=E(V(\boldsymbol\theta|\underline{\mathbf{X}}))=w$$

$$=(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)$$

$$=1-\underline{\Lambda}_o'(\Psi_o^{-1}-\Psi_o^{-1}\underline{\Lambda}_o(1+\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o)^{-1}\underline{\Lambda}_o'\Psi_o^{-1})\underline{\Lambda}_o$$

$$=\frac{1}{(1+t)}$$

$$=q^{-1},$$

and

(4.30) $$E(C(\underline{\boldsymbol\delta}|\underline{\mathbf{X}}))$$

$$=w\Psi^{-1/2}\underline{\Lambda\Lambda}'\Psi^{-1/2}$$

$$=q^{-1}\Psi^{-1/2}\underline{\Lambda\Lambda}'\Psi^{-1/2}.$$

Note that (4.29) implies that $R^2_{\boldsymbol\theta.\underline{\mathbf{X}}}$ is equal to $\frac{t}{(1+t)}$. From (4.30), the jth diagonal element of $E(C(\underline{\boldsymbol\delta}|\underline{\mathbf{X}}))$ is equal to

(4.31) $$\frac{\lambda_j^{\,2}}{(1+\underline{\Lambda}'\Psi^{-1}\underline{\Lambda})\sigma^2_{\zeta}}$$

in which σ^2_ζ is the jth diagonal element of Ψ. It follows from (4.29), that:

(4.32) if $t<-1$, then $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))<0$;

(4.33) if $t=-1$, then $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))$ does not exist;

(4.34) if $-1<t<0$, then $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))\in(1,\infty]$;

(4.35) if $t=0$, then $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))=1$.

Now, since $V(\boldsymbol{\theta})=V(E(\boldsymbol{\theta}|\underline{\mathbf{X}}))+E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))=1$, condition (4.34) implies that $V(E(\boldsymbol{\theta}|\underline{\mathbf{X}}))<0$, which cannot be. Note also that, from (4.29),

(4.36) as $t\rightarrow\infty$, $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))\rightarrow 0$.

If, as stipulated by (2.8)-(2.9), Ψ is positive definite, i.e., gramian and non-singular, then $\underline{\alpha}'\Psi\underline{\alpha}>0$ $\forall$ non-null p-vectors $\underline{\alpha}$. Hence, if, for at least one item, $\lambda_j\neq 0$, then $t>0$, and $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))\in[0,1)$. Moreover, for finite p, t must be finite, and, hence,

(4.37) $0<t<\infty$, and $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))\in(0,1)$.

Since $q^{-1}=E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))$, and the matrix $E(C(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}))=q^{-1}\Psi^{-1/2}\underline{\Lambda}\underline{\Lambda}'\Psi^{-1/2}$ is gramian and of rank unity, $E(C(\underline{\boldsymbol{\delta}}|\underline{\mathbf{X}}))$ contains at least one positive diagonal element. Hence, it can be concluded that:

Theorem 2 (Indeterminacy Property of ulcf representations; Guttman, 1955, Theorem 5): If the set of p manifest variates, $\underline{\mathbf{X}}$, is ulcf representable, with ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$, then, for finite p:

i. $0<V(\mathbf{I}_{\boldsymbol{\theta}(i)})<1$, so that the common factor $\boldsymbol{\theta}$ is indeterminate;
ii. $C(\underline{\mathbf{I}}_{\boldsymbol{\delta}(i)})$ possesses at least one positive diagonal element, so that at least one specific factor is indeterminate;
iii. The ulf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$ is indeterminate.

In light of Theorem 2(i), Schonemann and Haagen (1987, p.841) suggest that it is a strange feature of the linear factor model that one begins an analysis with "information" in the form of an N×p matrix whose rows are realizations of $\underline{\mathbf{X}}$, not of $\underline{\mathbf{X}}$ *and* $\boldsymbol{\theta}$, and yet, when one partials from $\boldsymbol{\theta}$ all of the information about $\boldsymbol{\theta}$ that is contained in $\underline{\mathbf{X}}$, one is left with a positive residual variance, $E(V(\boldsymbol{\theta}|\underline{\mathbf{X}}))$! The additional "information" implied by this positive variance comes from $\mathbf{s}_i$, which is wholly arbitrary save for the moment constraints of (4.6).

2f. The Transformation Approach

Schonemann and Wang (1972) named (4.4)-(4.6), and variants thereof, the construction approach to the production of factors to $\underline{\mathbf{X}}$. They also identified another historically important approach that they called the "transformation approach." In the transformation approach, one seeks a (p+1)×(p+1) orthonormal matrix T with the property that

(4.38) $[\underline{\Lambda}:\Psi^{1/2}]\mathrm{T}=[\underline{\Lambda}:\Psi^{1/2}]$.

If such a matrix exists, it follows from (2.8), (2.9), and (4.38) that,

(4.39) $$\underline{\mathbf{X}}=[\underline{\Lambda}:\Psi^{1/2}]\mathrm{T}\begin{bmatrix}\boldsymbol{\theta}\\ \underline{\boldsymbol{\delta}}\end{bmatrix}=[\underline{\Lambda}:\Psi^{1/2}]\begin{bmatrix}\boldsymbol{\theta}^*\\ \underline{\boldsymbol{\delta}}^*\end{bmatrix},$$

and in which both $\begin{bmatrix}\boldsymbol{\theta}\\ \underline{\boldsymbol{\delta}}\end{bmatrix}$ and $\begin{bmatrix}\boldsymbol{\theta}^*\\ \underline{\boldsymbol{\delta}}^*\end{bmatrix}=\mathrm{T}\begin{bmatrix}\boldsymbol{\theta}\\ \underline{\boldsymbol{\delta}}\end{bmatrix}$ contain factors to $\underline{\mathbf{X}}$.

It can be shown that T not only exists, but is not unique. Hence, an infinity of linearly related alternative constructions (a subset of the constructions contained in *C*) are given by this approach. Wilson (1928a) employed this approach to illustrate the indeterminacy property, Thomson (1935) derived an expression for the matrix T for the case of the single-factor model, and Ledermann (1938) extended Thomson's result to the multiple orthogonal case. Heermann (1966) employed Ledermann's result to work out alternative derivations of results due to Guttman (1955), while Schonemann (1971) gave a simplified proof of "Ledermann's lemma."

Lemma 1 (Ledermann, 1938; Schonemann, 1971): Define an "orthogonal right unit" of a p×(p+m) matrix B, to be a (p+m)×(p+m) matrix M, MM'=I, M'M=I, satisfying BM=B. Then B always has an orthogonal right unit, M, and M is a function of an arbitrary m×m orthogonal matrix S.

Proof

Define the singular value decomposition of B to be $B=V\Delta^*W'$, in which V is a p×p orthonormal matrix, Δ^* is the p×(p+m) matrix $[\Delta_{(p\times p)}:0_{(p\times m)}]$, in which $\Delta_{(p\times p)}$ contains the singular values of B, and W is a (p+m)×(p+m) orthonormal matrix. Partition W as $W=[W_{1(p+m)\times p}:W_{2(p+m)\times m}]$, the columns of W_1 a basis for the rows of B, and the columns of W_2 a basis for the m-dimensional subspace of $R^{(p+m)}$ that is orthogonal to the rows of B. It then follows that $B=V\Delta W_1$, $W_1'W_1=I_p$, and $W_1'W_2=0$. Define M to be

(4.40) $$M=W_1W_1'+W_2SW_2',$$

in which $SS'=S'S=I_m$. Then

(4.41) $$BM=V\Delta W_1(W_1W_1'+W_2SW_2')=V\Delta W_1=B,$$

(4.42) $$MM'=(W_1W_1'+W_2SW_2')(W_1W_1'+W_2SW_2')'=$$
$$W_1W_1'W_1W_1'+W_1W_1'W_2S'W_2'+W_2SW_2'W_1W_1'+W_2SW_2'W_2S'W_2'$$
$$=M'M=I_{(p+m),}$$

and M is a function of the arbitrary orthonormal matrix $S_{\lozenge}$

From Lemma 1, the following may be concluded.

Theorem 3 (Transformation Approach): Let $\underline{\mathbf{X}}$ be ulcf representable with ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$, and let $B=[\underline{\Lambda}_o:\Psi_o^{1/2}]$, a $p\times(p+1)$ matrix. Then:

1) B has an orthogonal right unit, M, which is a function of an arbitrary orthonormal matrix S;
2) M is given by (4.40);
3) An infinity of distinct factor constructions (a subset of *C*) are given by the transformation M, each construction tied to a distinct choice of S.

2g. Quantification of the degree of indeterminacy

Theorem 2 established that the ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$ is indeterminate. The cardinality of *C* is infinity. Since *C* contains an infinity of constructions $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}_{(si)}$ each of which is a set of factors to $\underline{\mathbf{X}}$, it is of interest to inquire how "similar" are these distinct sets. Of particular interest are the common factors to $\underline{\mathbf{X}}$. The concept of *similarity* may be defined in many different ways. It may, for example, be defined in terms of the correlation, or functions thereof, between pairs of distinct common factors contained within *C*, or via a comparison of the patterns of relationship between distinct common factors to $\underline{\mathbf{X}}$, and selected external variates. The first sense of similarity is the basis for standard quantifications of indeterminacy, the topic of this section, while the second has been considered under the heading of "external variable theory", later to be reviewed. The two senses are obviously related. The following should be noted:

i. All sets of factors to $\underline{\mathbf{X}}$, $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}_{(si)}$, have, according to (4.7)-(4.11), the same mean vector and covariance matrix, i.e.,

$$E\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}=\underline{0},\ E\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}[\boldsymbol{\theta}_i:\underline{\boldsymbol{\delta}}_i']=\begin{bmatrix}1 & \underline{0}'\\ \underline{0} & I\end{bmatrix}\quad\forall i.$$

ii. For ulcf representations in which $\{\boldsymbol{\theta},\underline{\boldsymbol{\delta}}\}$ must have a particular distribution, e.g., as in model (2.14)-(2.15), all constructions contained within *C* have this distribution.

iii. The fact that each set of factors to $\underline{\mathbf{X}}$ has the same mean vector and covariance matrix, and, when required, the same distribution, does not imply that they are pair-wise, element-wise perfectly correlated. The correlation matrix, R_{ij}, of any two sets of factors, $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}_{(si)}$ and $\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}_{(sj)}$ is

(4.43)

$$\begin{bmatrix} \underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o + w\rho(\mathbf{s}_i,\mathbf{s}_j) & \underline{\Lambda}_o'\Sigma^{-1}\Psi_o^{1/2} - w\underline{\Lambda}_o'\Psi_o^{-1/2}\rho(\mathbf{s}_i,\mathbf{s}_j) \\ \Psi_o^{1/2}\Sigma^{-1}\underline{\Lambda}_o - w\Psi_o^{-1/2}\underline{\Lambda}_o\rho(\mathbf{s}_i,\mathbf{s}_j) & \Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2} - w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}\rho(\mathbf{s}_i,\mathbf{s}_j) \end{bmatrix}$$

in which $\rho(\mathbf{s}_i,\mathbf{s}_j)$ is the correlation between the arbitrary random variates $\mathbf{s}_i$ and $\mathbf{s}_j$. According to (4.43), the degree of similarity, in the sense of mean-square difference, between any pair of alternative common or unique factors to $\underline{\mathbf{X}}$ varies with $\rho(\mathbf{s}_i,\mathbf{s}_j)$.

There have been offered up a number of methods to quantify the indeterminacy inherent to a given ulcf representation. Since, in practice, the common factors to $\underline{\mathbf{X}}$ are of chief interest, these methods have generally centered on the similarity of the elements of set *C*.

Ratio of indeterminate to determinate variance (Spearman, 1927; Holzinger, 1930).

This measure is based on decomposition (4.13). Call $V(\mathbf{D}_\theta)$ the "determinate variance" and $V(\mathbf{I}_{\theta(i)})$ the "indeterminate variance." Because $V(\boldsymbol{\theta}_i)= 1$, $V(\mathbf{D}_\theta)=\underline{\Lambda}'\Sigma^{-1}\underline{\Lambda}=R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}$, and $V(\mathbf{I}_{\theta(i)})=w=1-\underline{\Lambda}'\Sigma^{-1}\underline{\Lambda}=1-R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}$, the ratio of indeterminate to determinate variance is

(4.44) $$RAT_\theta = \frac{V(\mathbf{I}_\theta)}{V(\mathbf{D}_\theta)} = \frac{1-\underline{\mathbf{\Lambda}}_o'\mathbf{\Sigma}^{-1}\underline{\mathbf{\Lambda}}_o}{\underline{\mathbf{\Lambda}}_o'\mathbf{\Sigma}^{-1}\underline{\mathbf{\Lambda}}_o} = \frac{1}{R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}} - 1.$$

Because $0\leq R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}\leq 1$, it follows that $0\leq RAT_\theta\leq\infty$, with large values of RAT_θ indicating a high degree of indeterminacy, and a value of zero indicating a determinate representation (i.e., that Card(*C*)=1).

Ratio of square root of indeterminate variance to square root of determinate variance (Camp, 1932).

This measure is defined as

(4.45) $$RATSD_\theta = \frac{\sqrt{V(\mathbf{I}_\theta)}}{\sqrt{V(\mathbf{D}_\theta)}} = \frac{\sqrt{1-\underline{\mathbf{\Lambda}}_o'\mathbf{\Sigma}^{-1}\underline{\mathbf{\Lambda}}_o}}{\sqrt{\underline{\mathbf{\Lambda}}_o'\mathbf{\Sigma}^{-1}\underline{\mathbf{\Lambda}}_o}} = \frac{\sqrt{1-R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}}}{\sqrt{R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}}}.$$

Since, $0\leq R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}\leq 1$, it follows that $0\leq RATSD_\theta\leq\infty$, with large values of $RATSD_\theta$ indicating a high degree of indeterminacy, and a value of zero indicating a determinate representation.

Minimum correlation between distinct common factors, ρ^* (Guttman,1955).

Consider particular common factor to $\underline{\mathbf{X}}$, $\boldsymbol{\theta}_i=\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$. Guttman (1955) showed that the element within *C* that is mean-square most dissimilar to $\boldsymbol{\theta}_i$, say $\boldsymbol{\theta}_{im}$, is constructed by choosing its arbitrary component to be $-\mathbf{s}_i$:

(4.46) $$\boldsymbol{\theta}_{im}=\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}-w^{1/2}\mathbf{s}_i.$$

The correlation between these most dissimilar common factors to $\underline{\mathbf{X}}$, ρ^*, is Guttman's measure of indeterminacy. Now,

(4.47) $$\rho^*=\rho(\boldsymbol{\theta}_i,\boldsymbol{\theta}_{im})=\rho(\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i,\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}-w^{1/2}\mathbf{s}_i)=\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o-w=$$
$$\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o-(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)=2\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o-1=2R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}-1,$$

so that, as with the previous two measures, ρ^* is a function of $R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}$. Because there is no end to the arbitrary components $\mathbf{s}_i$ that could be produced, there is contained in *C* an infinity of pairs of common factors with precisely the same (minimum) correlation. Hence, ρ^* is a lower bound on the infinity of Pearson correlations defined on the set *C*. Because $0\leq R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}\leq 1$, it follows that $-1\leq\rho^*\leq 1$, with -1 indicating maximum indeterminacy and 1 indicating determinacy. However, it follows from Theorem 1 that, for finite p,
$-1\leq\rho^*<1$. Note that, if $R_{\boldsymbol{\theta}.\underline{\mathbf{X}}}=\frac{1}{\sqrt{2}}=.707$, then $\rho^*=0$, and that if $R_{\boldsymbol{\theta}.\underline{\mathbf{X}}}=0$, then $\rho^*=-1$, and *C* contains common factors to $\underline{\mathbf{X}}$ that are reflections of each other. That is, in the latter case, it would be possible to generate an infinity of pairs of common factors to $\underline{\mathbf{X}}$, the members of each pair in perfect disagreement with each other in regard the ordering of the members of population *P*. In fact, unless $R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}>.707$, *C* will contain pairs of common factors to $\underline{\mathbf{X}}$ whose agreement is no better than that of two uncorrelated variates.

Schonemann and Wang (1972) surveyed a large number of published factor analyses in order to ascertain the degree of indeterminacy that might be expected when the linear factor model is employed in behavioural research. They found that many of these analyses produced factor representations with negative values of ρ^*, prompting them to conclude that it is often the case that a set of "..."factor scores" can be predicted with more success from a set of random numbers than from an equivalent [set] of "factor scores"" (Schonemann and Wang, 1972, p.87). Note that, from (4.29), ρ^* may be expressed as

(4.48) $$\frac{(t-1)}{(t+1)},$$

in which, once again, $t=\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o$.

Example

To make the strengths of the linear relationships easily graspable, we will consider the artificial scenario in which $\underline{\mathbf{X}}$ has population covariance matrix

$$\begin{pmatrix} 1 & .024 & .223 & .197 \\ .024 & 1 & .029 & .025 \\ .223 & .029 & 1 & .233 \\ .197 & .025 & .233 & 1 \end{pmatrix}$$

$\underline{\mathbf{X}}$ then has ulcf representation

$$\underline{\mathbf{X}}=\begin{pmatrix} .434 \\ .056 \\ .513 \\ .455 \end{pmatrix}\boldsymbol{\theta}+\begin{pmatrix} .901 & 0 & 0 & 0 \\ 0 & .998 & 0 & 0 \\ 0 & 0 & .858 & 0 \\ 0 & 0 & 0 & .891 \end{pmatrix}\underline{\boldsymbol{\delta}},$$

$R^2_{\boldsymbol{\theta}.\underline{\mathbf{X}}}$=.46, and ρ^*=-.079. Note that, from (4.29), it follows that $\boldsymbol{\theta}_i=\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i=w\underline{\Lambda}_o'\Psi_o^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i$, in which $w=(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)=\frac{1}{(1+t_o)}$=.54. Hence, the construction formula is

$$\boldsymbol{\theta}_i=\begin{pmatrix} .289 & .954 & .376 & .31 \end{pmatrix}\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \\ \mathbf{X}_4 \end{pmatrix}+.735\mathbf{s}_i$$

in which $C(\mathbf{s}_i,\underline{\mathbf{X}})=\underline{0}$, $E(\mathbf{s}_i)=0$, $V(\mathbf{s}_i)=1$. On the other hand, an $\underline{\mathbf{X}}$ with covariance matrix

$$\begin{pmatrix} 1 & .328 & .223 & .197 \\ .328 & 1 & .388 & .344 \\ .223 & .388 & 1 & .233 \\ .197 & .344 & .233 & 1 \end{pmatrix}$$

has ulcf representation

$$\underline{\mathbf{X}}=\begin{pmatrix} .434 \\ .756 \\ .513 \\ .455 \end{pmatrix}\boldsymbol{\theta}+\begin{pmatrix} .901 & 0 & 0 & 0 \\ 0 & .654 & 0 & 0 \\ 0 & 0 & .858 & 0 \\ 0 & 0 & 0 & .891 \end{pmatrix}\underline{\boldsymbol{\delta}},$$

with $R_{\boldsymbol{\theta}.\underline{\mathbf{X}}}$=.686, ρ^*=.372, w=.314, and construction formula

$$\boldsymbol{\theta}_i=\begin{pmatrix} .168 & .555 & .219 & .18 \end{pmatrix}\begin{pmatrix} \mathbf{X}_1 \\ \mathbf{X}_2 \\ \mathbf{X}_3 \\ \mathbf{X}_4 \end{pmatrix}+.56\mathbf{s}_i$$

.

2h. Minimum average correlation (Schonemann, 1971).

If one considers not just the common factor, but also the specific factors, then the *minimum correlation matrix*, produced by replacing $\rho(\mathbf{s}_i,\mathbf{s}_j)$ with $\rho(\mathbf{s}_i,-\mathbf{s}_i)=-1$ in (4.43), is (Guttman, 1955)

$$(4.49) \qquad R_{min}=\begin{bmatrix} 2\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o-1 & 2\underline{\Lambda}_o'\Sigma^{-1}\Psi_o^{\frac{1}{2}} \\ 2\Psi_o^{\frac{1}{2}}\Sigma^{-1}\underline{\Lambda}_o & 2\Psi_o^{\frac{1}{2}}\Sigma^{-1}\Psi_o^{\frac{1}{2}}-I \end{bmatrix}$$

.

Notice that, since $2\Psi_o^{\frac{1}{2}}\Sigma^{-1}\underline{\Lambda}_o$ will, in general, be non-null, "...this brings us to the unhappy conclusion that common and unique factors are not entirely distinct entities when we consider the set of potential factor axes in total factor space" (Heermann, 1966, p.541). That is, while it is the case that for any particular set, k, of factors to $\underline{\mathbf{X}}$, $E(\boldsymbol{\theta}_k\underline{\boldsymbol{\delta}}_k)=\underline{0}$, it is not necessarily the case that $E(\boldsymbol{\theta}_j\underline{\boldsymbol{\delta}}_k)=\underline{0}$, when $j\neq k$. Hence, some of $\underline{\mathbf{X}}$'s common factors will be correlated with some of its specifics. The *minimum average correlation* is the average of the diagonal elements of (4.49), i.e., the average of the (p+1) minimum correlations:

$$(4.50) \qquad \tau=\frac{tr(R_{min})}{(p+1)}.$$

Using the transformation approach, Schonemann (1971) proved that, for the m factor orthogonal model,

$$(4.51) \qquad \tau=\frac{(p-m)}{(p+m)},$$

which, as he noted, does not depend on Λ_o and Ψ_o, the parameters of the representation. Schonemann (1971) actually derived (4.50) under the mistaken belief that all of the constructions in *C* were related by Ledermann's transformation, and, hence, were pairwise linearly related. He speculated (1971, p.28) that, if this were not the case, $\frac{(p-m)}{(p+m)}$ would be but an upper bound to τ (since, then, there might exist non-linearly related constructions with even lower values of τ). Using the Lawley-Rao basis for the loadings matrix, McDonald (1974) provided an alternative proof of (4.51), and established that: i) (4.51) holds generally, i.e., τ is a minimum over all sets of factors to $\underline{\mathbf{X}}$, and not just linearly related sets; ii) Sets of factors to $\underline{\mathbf{X}}$ are not, in general, pairwise linearly related.

Theorem 3 (Minimum average correlation, unidimensional case): Let $\underline{\mathbf{X}}$ be ulcf representable, and let its ulcf representation be $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$. Then $\tau=\frac{tr(R_{min})}{(p+1)}=\frac{(p-1)}{(p+1)}$.

Proof

It follows from (4.48) that

(4.52) $$\begin{aligned}tr(R_{min})&=2tr(\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)-1+2tr(\Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2})-trI_p=\\ &2tr(\underline{\Lambda}_o'(\Psi_o^{-1}-\Psi_o^{-1}\underline{\Lambda}_o(1+\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o)^{-1}\underline{\Lambda}_o'\Psi_o^{-1})\underline{\Lambda}_o)-1+\\ &2tr(\Psi_o^{1/2}(\Psi_o^{-1}-\Psi_o^{-1}\underline{\Lambda}_o(1+\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o)^{-1}\underline{\Lambda}_o'\Psi_o^{-1})\Psi_o^{1/2})-trI_p\\ &=2\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o-2\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o(1+\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o)^{-1}\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o-1+2trI_p\\ &-2tr\Psi_o^{-1/2}\underline{\Lambda}_o(1+\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o)^{-1}\underline{\Lambda}_o'\Psi_o^{-1/2}-trI_p.\end{aligned}$$

Letting $t=\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o$, (4.52) can then be expressed as

(4.53) $$2t-\frac{2t^2}{(1+t)}-1+2p-\frac{2t}{(1+t)}-p=\frac{2t+2t^2-2t^2-2t}{(1+t)}+(p-1)=(p-1)\ \square$$

2i. The eigenstructure of $\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}$

Let $\underline{\mathbf{X}}$ be ulcf representable, and let its ulcf representation be $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$, so that $\Sigma=\underline{\Lambda}_o\underline{\Lambda}_o'+\Psi_o$. It then follows that

(4.54) $$\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}=\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}+I,$$

and, hence, that

(4.55) $$|\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}-\lambda_*I|=|\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}+I-\lambda_*I|=|\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}-\lambda'I|,$$

in which λ_* is an eigenvalue of $\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}$ and $\lambda'=(\lambda_*-1)$ is an eigenvalue of $\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$. Now, if Σ is of rank p, then $\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}$ is of rank p and possesses p non-zero eigenvalues. Because $\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$ is gramian, all of its eigenvalues λ' are nonnegative. However, the rank of this matrix is unity, so that $\lambda'_1>0$, and $\lambda'_i=0$, i=2...p.[2] It follows then that $\lambda_{*1}>1$, and $\lambda_{*i}=1$, i=2...p. The eigendecomposition of $\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}$ is then

---

2 In the multi-factor case, the number of eigenvalues of $\Psi^{-1/2}\Sigma\Psi^{-1/2}$ that are greater than unity is equal to the number of eigenvalues of $\Psi^{-1/2}\Lambda\Lambda'\Psi^{-1/2}$ that are greater than zero, which, in turn, equals the number of eigenvalues of $(\Sigma-\Psi)$ that are greater than zero, which, finally, is equal to the number of elements of $\underline{\boldsymbol{\theta}}$ (each of whose elements stands for a set of constructed factors of

(4.56) $$Q\Lambda^{*}Q'=\left[\underline{Q}_1 : Q_2\right]\begin{pmatrix}\lambda_{*1} & \underline{0}' \\ \underline{0} & I_{(p-1)}\end{pmatrix}\begin{bmatrix}\underline{Q}_1' \\ Q_2'\end{bmatrix},$$

in which Q_1 is the first eigenvector, and Q_2 the (p-1)×(p-1) matrix whose columns contain eigenvectors 2 through p. From (4.54) and (4.56),

(4.57) $$(\Psi_o^{-1/2}\Sigma\Psi_o^{-1/2}-\Lambda^{*})Q=(\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}+I-\Lambda^{*})Q=$$
$$(\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}-(\Lambda^{*}-I))Q=(\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}-\Lambda')Q,$$

from which it follows that

(4.58) $$\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}=(\lambda_{*1}-1)\underline{Q}_1\underline{Q}_1',$$

since the final (p-1) eigenvalues of $\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}$ are equal to zero. Finally, from (4.58),

(4.59) $$\underline{\Lambda}_o=\Psi_o^{1/2}Q_1(\lambda_{*1}-1)^{1/2}.$$

In the m-factor case, a Λ_o defined as $\Psi_o^{1/2}Q_1(\Pi_{*m}-1)^{1/2}$ satisfies $\Lambda_o'\Sigma^{-1}\Lambda_o$ diagonal, and is said to be defined on the Lawley-Rao basis (see, e.g., McDonald, 1974).

Schonemann and Wang (1972) and McDonald (1974) use (4.59) to establish that:

(4.60) $$\underline{\Lambda}_o'\Psi_o^{-1}\underline{\Lambda}_o=(\lambda_{*1}-1),$$

and

(4.61) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o=\frac{(\lambda_{*1}-1)}{\lambda_{*1}}.$$

From (4.56), and the fact that $\Psi^{1/2}\Sigma^{-1}\Psi^{1/2}=(\Psi^{-1/2}\Sigma\Psi^{-1/2})^{-1}$, the eigendecomposition of $\Psi^{1/2}\Sigma^{-1}\Psi^{1/2}$ is

(4.62) $$Q\Lambda^{*-1}Q'=\left[\underline{Q}_1 : Q_2\right]\begin{pmatrix}\lambda_{*1}^{-1} & \underline{0}' \\ \underline{0} & I_{(p-1)}\end{pmatrix}\begin{bmatrix}\underline{Q}_1' \\ Q_{2)}'\end{bmatrix}$$

Hence, from (4.61) and (4.62), R_{min} may be represented as

---

cardinality infinity).

(4.63) $$\begin{pmatrix} 1-\frac{2}{\lambda_{*1}} & \frac{2\sqrt{(\lambda_{*1}-1)}}{\lambda_{*1}}\underline{Q}_1' & \\ \frac{2\sqrt{(\lambda_{*1}-1)}}{\lambda_{*1}}\underline{Q}_1 & \frac{2}{\lambda_{*1}}-1 & \underline{0}' \\ & \underline{0} & I_{(p-1)} \end{pmatrix},$$

from which follows Theorem 3, i.e., that $tr(R_{min})=(1-\frac{2}{\lambda_{*1}})+(\frac{2}{\lambda_{*1}}-1)+tr(I_{(p-1)})=(p-1)$.
McDonald (1974) used this result to show that Schonemann's $tr(R_{min})$ result holds generally, and not just for linearly related factors. Schonemann and Wang (1972, p.69) noted that if λ_{*1} is not greater than two, then, according to (4.63), the minimally related common factor pairs contained in *C* will be negatively correlated, a finding that seems to undermine the heuristic of deciding on the dimensionality of a linear factor representation on the basis of the number of λ_{*i} that are greater than unity.

2j. Linearly and non-linearly related factors to $\underline{\mathbf{X}}$

In the past, there existed uncertainty over the nature of the relationships between the infinity of sets of factors to $\underline{\mathbf{X}}$, and, in particular, whether these factors were pairwise linearly related. McDonald (1974) clarified the issue.

Definition (Linear relatedness; McDonald, 1974). Let $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ each be random t-vectors, with $E(\underline{\mathbf{X}})=E(\underline{\mathbf{Y}})=\underline{0}$, $C(\underline{\mathbf{X}})=C(\underline{\mathbf{Y}})=I$, $E(\underline{\mathbf{X}}\underline{\mathbf{Y}}')=R_{xy}$, and let $C(\underline{\mathbf{X}}|\underline{\mathbf{Y}}=\underline{y})$, the conditional covariance matrix of $\underline{\mathbf{X}}$ given that $\underline{\mathbf{Y}}=\underline{y}$, be homoscedastic over $\underline{y}$. Then $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ are, by definition, linearly related if and only if the diagonal of $C(\underline{\mathbf{X}}|\underline{\mathbf{Y}}=\underline{y})$ contains zeros.

Lemma 2 (Linear relatedness; McDonald, 1974): Let $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ each be random t-vectors, with $E(\underline{\mathbf{X}})=E(\underline{\mathbf{Y}})=\underline{0}$, $C(\underline{\mathbf{X}})=C(\underline{\mathbf{Y}})=I$, $E(\underline{\mathbf{X}}\underline{\mathbf{Y}}')=R_{xy}$, and let $C(\underline{\mathbf{X}}|\underline{\mathbf{Y}}=\underline{y})$, the conditional covariance matrix of $\underline{\mathbf{X}}$ given that $\underline{\mathbf{Y}}=\underline{y}$, be homoscedastic over $\underline{y}$. Then $\underline{\mathbf{X}}$ and $\underline{\mathbf{Y}}$ are linearly related if and only if R_{xy} is orthonormal, in which case $\underline{\mathbf{X}}=R_{xy}\underline{\mathbf{Y}}$.

Proof. $C(\underline{\mathbf{X}}|\underline{\mathbf{Y}}=\underline{y})$ is gramian. Hence, if its diagonal elements are all zero, it must be a null matrix. Since $C(\underline{\mathbf{X}}|\underline{\mathbf{Y}}=\underline{y})=I-R_{xy}R_{xy}'$, this will obtain only if $R_{xy}R_{xy}'=I$, i.e., only if R_{xy} is an orthonormal matrix. In this case, $\underline{\mathbf{X}}=E(\underline{\mathbf{X}}|\underline{\mathbf{Y}})+\underline{\boldsymbol{\varepsilon}}=E(\underline{\mathbf{X}}|\underline{\mathbf{Y}})=R_{xy}\underline{\mathbf{Y}}$ ▫

Theorem 4 (Linear relatedness of factors to $\underline{\mathbf{X}}$; McDonald, 1974): Let $\underline{\mathbf{X}}$ be ulcf representable, let its ulcf representation be $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$, and let $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}$ and $\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$ be any two sets of factors to $\underline{\mathbf{X}}$. Then:

i. The correlation matrix of $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}$ and $\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$, R_{ij}, is as given by (4.43);

ii. R_{ij} is an orthogonal right-unit of $[\underline{\Lambda}_o:\Psi_o^{1/2}]$;

iii. The truth of any of the following implies the truth of the others:

a) $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}and\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$ are related by Ledermann's transformation;

b) $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}and\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$are linearly related;

c) R_{ij} is orthonormal;

d) $\rho^2(\mathbf{s}_i,\mathbf{s}_j)=1$;

e) $\mathbf{s}_i=\rho(\mathbf{s}_i,\mathbf{s}_j)\mathbf{s}_j$.

iv. R_{ij} is not necessarily orthonormal. Hence, $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}and\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$are not necessarily linearly related, and, hence, are not necessarily related by Ledermann's transformation.

Proof

ii) Part (ii) of the theorem is proven by substitution.

iii) (a↔b)

→

Let $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}and\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$ be related by Ledermann's transformation, i.e., let T be a matrix such that $TT'=T'T=I$ and $[\underline{\Lambda}:\Psi^{1/2}]T=[\underline{\Lambda}:\Psi^{1/2}]$, with the consequence that
$\underline{\mathbf{X}}=[\underline{\Lambda}_o:\Psi_o^{1/2}]T\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}=[\underline{\Lambda}_o:\Psi_o^{1/2}]\begin{bmatrix}\boldsymbol{\theta}_j\\ \underline{\boldsymbol{\delta}}_j\end{bmatrix}$. Then $C\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}=C\begin{bmatrix}\boldsymbol{\theta}_j\\ \underline{\boldsymbol{\delta}}_j\end{bmatrix}=I$, and $R_{ij}=E\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}\lfloor\boldsymbol{\theta}_j:\underline{\boldsymbol{\delta}}_j'\rfloor$
$=E\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}\lfloor\boldsymbol{\theta}_j:\underline{\boldsymbol{\delta}}_j'\rfloor T'=T'$. Hence,
$C(\underline{\mathbf{X}}|\underline{\mathbf{Y}}=\underline{y})=I-R_{ij}'R_{ij}=I-T'T=I-I=0$.

←

Let $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}and\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$ be linearly related. Then, by Lemma 2, R_{ij} is orthonormal, and $\begin{bmatrix}\boldsymbol{\theta}_j\\ \underline{\boldsymbol{\delta}}_j\end{bmatrix}=R_{ij}'$
$\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}$. Thus $\underline{\mathbf{X}}=[\underline{\Lambda}_o:\Psi_o^{1/2}]\begin{bmatrix}\boldsymbol{\theta}_j\\ \underline{\boldsymbol{\delta}}_j\end{bmatrix}=[\underline{\Lambda}_o:\Psi_o^{1/2}]R_{ij}'\begin{bmatrix}\boldsymbol{\theta}_i\\ \underline{\boldsymbol{\delta}}_i\end{bmatrix}$, and $[\underline{\Lambda}_o:\Psi_o^{1/2}]=[\underline{\Lambda}_o:\Psi_o^{1/2}]R_{ij}'$. Hence, R_{ij} and R_{ij}' are orthogonal right units to $[\underline{\Lambda}_o:\Psi_o^{1/2}]$, and $\{\boldsymbol{\theta}_i,\underline{\boldsymbol{\delta}}_i\}and\{\boldsymbol{\theta}_j,\underline{\boldsymbol{\delta}}_j\}$are related by Ledermann's transformation▫

(b↔c) From Lemma 2.

(c↔d)

$\rightarrow$

$$R_{ij}R_{ij}'=\begin{bmatrix}\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o+w\rho^2(\mathbf{s}_i,\mathbf{s}_j) & \underline{\Lambda}_o'\Sigma^{-1}\Psi_o^{1/2}-w\underline{\Lambda}_o'\Psi_o^{-1/2}\rho^2(\mathbf{s}_i,\mathbf{s}_j)\\ \Psi_o^{1/2}\Sigma^{-1}\underline{\Lambda}_o-w\Psi_o^{-1/2}\underline{\Lambda}_o\rho^2(\mathbf{s}_i,\mathbf{s}_j) & \Psi_o^{1/2}\Sigma^{-1}\Psi_o^{1/2}-w\Psi_o^{-1/2}\underline{\Lambda}_o\underline{\Lambda}_o'\Psi_o^{-1/2}\rho^2(\mathbf{s}_i,\mathbf{s}_j)\end{bmatrix}$$

If R_{ij} is orthonormal, it must be the case that $\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o+w\rho^2(\mathbf{s}_i,\mathbf{s}_j)=1$. But this is the case only if $\rho^2(\mathbf{s}_i,\mathbf{s}_j)=1$.

$\leftarrow$

On the other hand, if $\rho^2(\mathbf{s}_i,\mathbf{s}_j)=1$ is substituted into $R_{ij}R_{ij}'$, then this matrix product is equal to the identity matrix◊

It may then be concluded that Ledermann's transformation, and Schoneman's restatement of this transformation, does not provide a formula for all constructions, but only the subset of factors to $\underline{\mathbf{X}}$ that are linearly related.

2k. External variate theory.

Let $\underline{\mathbf{X}}$ be ulcf representable, let its ulcf representation be $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$, and consider the set *C* of common factors to $\underline{\mathbf{X}}$. Measures of indeterminacy provide one indication of the similarity of these factors. A second sense of similarity derives from a consideration of the set of correlations of these factors with a variate, **Z**, not belonging to $\underline{\mathbf{X}}$ (i.e., an "external variate"). The indeterminacy inherent to the ulcf representation of $\underline{\mathbf{X}}$ can be quantified by the range of this set of external correlations. Schonemann and Steiger (1978), Steiger (1979), and Schonemann and Haagen (1987) have developed this line of thought.

Let external variate **Z** have E(**Z**)=0 and V(**Z**)=1. To derive the range of $\rho(\boldsymbol{\theta}_i,\mathbf{Z})$ over the set *C* requires the following lemma due to McDonald (1977) and restated in Steiger (1979).

Lemma 3 (Partial correlation inequality; McDonald, 1977): Let $\underline{\mathbf{Y}}=\begin{bmatrix}\underline{\mathbf{X}}\\ \mathbf{W}_1\\ \mathbf{W}_2\end{bmatrix}$ be a (p+2)×1 random vector, in which $E(\underline{\mathbf{Y}})=\underline{0}$, and

$\Sigma_{\underline{\mathbf{Y}}}=\begin{pmatrix}\Sigma_{\underline{\mathbf{X}}} & \underline{\sigma}_{\underline{\mathbf{X}},W_1} & \underline{\sigma}_{\underline{\mathbf{X}},W_2}\\ \underline{\sigma}'_{\underline{\mathbf{X}},W_1} & 1 & \sigma_{W_1,W_2}\\ \underline{\sigma}'_{\underline{\mathbf{X}},W_2} & \sigma_{W_1,W_2} & 1\end{pmatrix}$. Then the correlation $\rho_{w1,w2}$ satisfies the inequality

(4.64)

$$\underline{\sigma}'_{w1,\underline{x}}\Sigma_{\underline{x}}^{-1}\underline{\sigma}_{\underline{x},w2}-(1-R^2_{w1.\underline{x}})^{1/2}(1-R^2_{w2.\underline{x}})^{1/2}\leq\rho_{w1,w2}\leq\underline{\sigma}'_{w1,\underline{x}}\Sigma_{\underline{x}}^{-1}\underline{\sigma}_{\underline{x},w2}+(1-R^2_{w1.\underline{x}})^{1/2}(1-R^2_{w2.\underline{x}})^{1/2}.$$

Proof

The conditional correlation of $\mathbf{W}_1$ and $\mathbf{W}_2$, having partialed $\underline{\mathbf{X}}$ from each, is

(4.65) $$\rho_{w_1,w_2|\underline{x}} = \frac{\rho_{w_1,w_2} - \underline{\sigma}'_{w_1,\underline{x}}\Sigma^{-1}\underline{\sigma}_{\underline{x},w_2}}{\sqrt{(1-\underline{\sigma}'_{w_1,\underline{x}}\Sigma^{-1}\underline{\sigma}_{\underline{x},w_1})}\sqrt{(1-\underline{\sigma}'_{w_2,\underline{x}}\Sigma^{-1}\underline{\sigma}_{\underline{x},w_2})}}.$$

Inequality (4.64) then follows from the fact that $-1\leq\rho_{w1,w2|\underline{x}}\leq 1$, and by noting the identities $(1-\underline{\sigma}'_{w1,\underline{x}}\Sigma_{\underline{x}}^{-1}\underline{\sigma}'_{w1,\underline{x}})^{1/2}=(1-R^2_{w1.\underline{x}})^{1/2}$ and $(1-\underline{\sigma}'_{w2,\underline{x}}\Sigma_{\underline{x}}^{-1}\underline{\sigma}'_{w2,\underline{x}})^{1/2}=(1-R^2_{w2.\underline{x}})^{1/2}$.

Theorem 5 (Bounds on external correlations, $\rho(\theta_i,Z)$; Steiger, 1979): Let $\underline{\mathbf{X}}$ be ulcf representable with ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$. The set of correlations between the $\boldsymbol{\theta}_i$ contained in *C* and any external variate $\mathbf{Z}$:

i) Has lower bound (lb) equal to

(4.66) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{Z}}-(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})^{1/2}w^{1/2}$$

and upper bound (ub) equal to

(4.67) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{Z}}+(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})^{1/2}w^{1/2}.$$

ii) The $\boldsymbol{\theta}_i$ that yields the lower bound is constructed as

(4.68) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}-(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})^{-1/2}w^{1/2}(\mathbf{Z}-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma^{-1}\underline{\mathbf{X}})$$

and that which yields the upper bound,

(4.69) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})^{-1/2}w^{1/2}(\mathbf{Z}-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma^{-1}\underline{\mathbf{X}}).$$

iii) The range of the set of external correlations is equal to

(4.70) $$\sqrt{2}\,(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})^{1/2}(1-\rho^*)^{1/2}.$$

Proof

i) Given that $\underline{\mathbf{X}}$ is ulcf representable, let $\underline{\mathbf{Y}}=\begin{bmatrix}\underline{\mathbf{X}}\\ \boldsymbol{\theta}_i\\ \mathbf{Z}\end{bmatrix}$ and V($\mathbf{Z}$)=1, so that E($\underline{\mathbf{Y}}$)=$\underline{0}$, and

$\Sigma_{\underline{\mathbf{Y}}}=\begin{pmatrix}\boldsymbol{\Sigma} & \underline{\Lambda}_o & \underline{\sigma}_{\underline{\mathbf{X}},\mathbf{Z}}\\ \underline{\Lambda}_o' & 1 & \rho_{\boldsymbol{\theta}_i,\mathbf{Z}}\\ \underline{\sigma}'_{\underline{\mathbf{X}},\mathbf{Z}} & \rho_{\boldsymbol{\theta}_i,\mathbf{Z}} & 1\end{pmatrix}$. With $\mathbf{W}_1$ set to $\boldsymbol{\theta}_i$ and $\mathbf{W}_2$ set to $\mathbf{Z}$, apply Lemma 3 to get

(4.71) $$\rho_{\boldsymbol{\theta}_i,\mathbf{Z}|\underline{\mathbf{X}}}=\frac{\rho_{\boldsymbol{\theta}_i,\mathbf{Z}}-\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{Z}}}{\sqrt{(1-\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o)}\sqrt{(1-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{Z}})}}$$

from which, upon noting that $(1-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma_{\underline{\mathbf{X}}}^{-1}\underline{\sigma}'_{\underline{\mathbf{X}},\mathbf{Z}})=(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})$ and that $(1-\underline{\Lambda}_o'\Sigma_{\underline{\mathbf{X}}}^{-1}\underline{\Lambda}_o)=w$, (4.66) and (4.67) follow immediately.

ii) Note that

(4.72) $$\rho(\boldsymbol{\theta}_i,\mathbf{Z})=C(\boldsymbol{\theta}_i,\mathbf{Z})=C(\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}+w^{1/2}\mathbf{s}_i,\mathbf{Z})=\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{x}},\mathbf{z}}+w^{1/2}\sigma_{\mathbf{si},\mathbf{z}}.$$

Now, $\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{Z}}$ is constant over all $\boldsymbol{\theta}_i$ in *C*. Hence, the magnitude of $\rho(\boldsymbol{\theta}_i,\mathbf{Z})$ varies only through $\sigma_{\mathbf{si},\mathbf{Z}}$. For the upper bound, $\sigma_{\mathbf{si},\mathbf{z}}$ must be maximized over all admissible random variates $\mathbf{s}_i$. The only restrictions on the choice of $\mathbf{s}_i$ are that E($\mathbf{s}_i$)=0, V($\mathbf{s}_i$)=1, and $\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{si}}=\underline{0}$. Hence, to maximize $\sigma_{\mathbf{si},\mathbf{Z}}$, one chooses $\mathbf{s}_i$ to be the unit variance random variate contained within the space of random variates uncorrelated with $\underline{\mathbf{X}}$, and closest, in a mean-square distance sense, to $\mathbf{Z}$. That is, $\mathbf{s}_i$ is the unit variance counterpart of the residual of $\mathbf{Z}$ after projection into the space of $\underline{\mathbf{X}}$:

(4.73) $$\mathbf{s}_{max}=(\mathbf{Z}-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma^{-1}\underline{\mathbf{X}})(1-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma^{-1}\underline{\sigma}_{\mathbf{X},\mathbf{Z}})^{-1/2}.$$

Because $(1-\underline{\sigma}'_{\mathbf{Z},\underline{\mathbf{X}}}\Sigma^{-1}\underline{\sigma}_{\mathbf{X},\mathbf{Z}})^{-1/2}=(1-R^2_{\mathbf{Z}.\underline{\mathbf{X}}})^{-1/2}$, (4.69) follows immediately by substitution of (4.73) into (4.4). On the other hand,

(4.74) $$\mathbf{s}_{min}=-\mathbf{s}_{max},$$

and (4.68) follows by substitution of $-\mathbf{s}_{max}$ into (4.4).

iii) Because, from (4.47), $\rho^*=2\underline{\Lambda}_o'\Sigma^{-1}\underline{\Lambda}_o-1$, it follows that (4.66) and (4.67) can be re-expressed as

(4.75) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}-\frac{1}{\sqrt{2}}(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2}(1-\rho^*)^{1/2}$$

and

(4.76) $$\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}+\frac{1}{\sqrt{2}}(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2}(1-\rho^*)^{1/2},$$

respectively. The range of the set of correlations is then the difference between (4.76) and (4.75), or

(4.77) $$\sqrt{2}(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2}(1-\rho^*)^{1/2} \lozenge$$

Result (4.77) makes clear that the range of the external correlations is a function of both: i) the degree of linear dependency of **Z** on **X̲**; and ii) the indeterminacy of the ulcf representation of **X̲**. In particular, because, for finite p, $\rho^*<1$, if $R^2_{\mathbf{z}.\underline{\mathbf{X}}}<1$, it follows that (4.77) will be positive. That is, the correlations of the common factors to **X̲** with **Z** will not all be equal.

Note that:

i) If the ulf representation of **X̲** is maximally indeterminate, i.e., $\rho^*=-1$, then, for fixed $R^2_{\mathbf{z}.\underline{\mathbf{X}}}<1$, the set of external correlations has bounds

(4.78) $$[\underline{\Lambda}'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}-(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2},\underline{\Lambda}'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}+(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2}],$$

and the range is equal to $2(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2}$.

ii) If $\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}=\underline{0}$, i.e., the manifest variates are each uncorrelated with **Z**, then the set of external correlations has bounds

(4.79) $$[-\frac{1}{\sqrt{2}}(1-\rho^*)^{1/2},\frac{1}{\sqrt{2}}(1-\rho^*)^{1/2}],$$

and the range is equal to $\sqrt{2}(1-\rho^*)^{1/2}$.

<u>Example</u>

Consider once again the first correlation matrix from the previous example, this for a set of four manifest variates with ulcf representation

$$\underline{\mathbf{X}}=\begin{pmatrix} .434 \\ .056 \\ .513 \\ .455 \end{pmatrix}\boldsymbol{\theta}+\begin{pmatrix} .901 & 0 & 0 & 0 \\ 0 & .998 & 0 & 0 \\ 0 & 0 & .858 & 0 \\ 0 & 0 & 0 & .891 \end{pmatrix}\underline{\boldsymbol{\delta}},$$

$R_{\boldsymbol{\theta}.\underline{\mathbf{X}}}$=.46, and ρ^*=-.079. Imagine that for some external variate **Z**, $\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}$ is equal to

$$\begin{pmatrix} .301 \\ .404 \\ -.121 \\ .523 \end{pmatrix}.$$

Then, lb=-.275, ub=.706, and the range of the correlations of **Z** with the elements of *C*, i.e., over the common factors to **X**, is .981. That is, **X** possesses common factors that have a correlation with **Z** as small as -.275, and common factors that have a correlation with **Z** as large as .706.

Consider, additionally, the following two scenarios from Steiger (1979).

i) Imagine that each member of the set of ulcf representable manifest variates $\mathbf{X}_j$ described above is uncorrelated with **Z**, i.e., $\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}=\underline{0}$. Obviously, then, the traditional regression estimator $\underline{\Lambda}_o'\Sigma^{-1}\underline{\mathbf{X}}$ is also uncorrelated with $\underline{\mathbf{Z}}$. Yet, because ρ^*=-.079,

$$\text{lb}=\underline{\Lambda}_o'\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{X}},\mathbf{z}}-\frac{1}{\sqrt{2}}(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{1/2}(1-\rho^*)^{1/2}=-\frac{1}{\sqrt{2}}(1-\rho^*)^{1/2}=-\frac{1}{\sqrt{2}}(1-(-.079))^{1/2}=-.735,\ \text{ub}=.735,$$

and the range of the correlations of **Z** with the common factors to **X** is 1.47. That is, despite the fact that the manifest variates contained in **X** are uncorrelated with **Z**, certain of the common factors to **X** are, nevertheless, strongly positively correlated with **Z**, while others are strongly negatively correlated with **Z**.

ii) Within population *P*, let a set of p variates, **X**, be ulcf representable with ulcf representation $\underline{\mathbf{X}}=\underline{\Lambda}_x\boldsymbol{\theta}_x+\Psi_x^{1/2}\underline{\boldsymbol{\delta}}_x$, and a second set of q variates, **Y**, be ulcf representable with ulcf representation $\underline{\mathbf{Y}}=\underline{\Lambda}_y\boldsymbol{\theta}_y+\Psi_y^{1/2}\underline{\boldsymbol{\delta}}_y$. Consider the two sets C_X and C_Y. The minimum correlation over the set C_X of common factors to **X** is, of course, equal to $2\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\Lambda}_x-1$, and the minimum correlation over the set C_y, $2\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\Lambda}_y-1$. What can be said of the correlations between the $\boldsymbol{\theta}_{ix}$ contained in C_x and the $\boldsymbol{\theta}_{iy}$ contained in C_y? The correlation between a $\boldsymbol{\theta}_{ix}$ contained in C_x and a $\boldsymbol{\theta}_{iy}$ contained in C_y is

$$\rho(\boldsymbol{\theta}_{ix},\boldsymbol{\theta}_{iy})=\rho(\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\mathbf{X}}-w_x^{1/2}\mathbf{s}_{ix},\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\mathbf{Y}}-w_y^{1/2}\mathbf{s}_{iy})=$$

$$\underline{\Lambda}_x'\Sigma_x^{-1}\Sigma_{xy}\Sigma_y^{-1}\underline{\Lambda}_y+w_x^{1/2}\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\sigma}_{ysix}+w_y^{1/2}\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\sigma}_{xsiy}+w_x^{1/2}w_y^{1/2}\sigma_{sixsiy}$$

Consider the special case in which: i) Σ_{xy} is a null matrix, i.e., the manifest variates **X** are pairwise uncorrelated with the manifest variates **Y**; ii) $\rho^*_x=0=\rho^*_y$. Choose $\mathbf{s}_{ix}$ to be $\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\mathbf{Y}}(\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\Lambda}_y)^{-1/2}$ and $\mathbf{s}_{iy}$ to be $\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\mathbf{X}}(\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\Lambda}_x)^{-1/2}$. These choices of $\mathbf{s}_{ix}$ and $\mathbf{s}_{iy}$ are admissible choices because Σ_{xy} is a null matrix and, as a result:

$E(\underline{\mathbf{X}}(\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\mathbf{Y}}(\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\Lambda}_y)^{-1/2}))=\underline{0}$, $E(\underline{\mathbf{Y}}(\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\mathbf{X}}(\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\Lambda}_x)^{-1/2}))=\underline{0}$, $V(\mathbf{s}_{ix})=1$, $V(\mathbf{s}_{iy})=1$, $E(\mathbf{s}_{ix})=0$, and $E(\mathbf{s}_{iy})=0$. It follows, then, that

$$\rho(\boldsymbol{\theta}_{ix},\boldsymbol{\theta}_{iy})=w_x^{1/2}\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\sigma}_{ysix}+w_y^{1/2}\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\sigma}_{xsiy}+w_x^{1/2}w_y^{1/2}\sigma_{sixsiy} \quad \text{(from (i))}$$

$$=w_x^{1/2}(\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\Lambda}_y)^{1/2}+w_y^{1/2}(\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\Lambda}_x)^{1/2} \quad \text{(from (ii))}$$

$$=1,$$

the final step, a consequence of (iii), and the following two identities: 1) $w^{1/2}=\left(\frac{1-\rho^*}{2}\right)^{\frac{1}{2}}$; 2) $\underline{\Lambda}'\Sigma^{-1}\underline{\Lambda}=\left(\frac{1+\rho^*}{2}\right)^{\frac{1}{2}}$.

If, instead, $\mathbf{s}_{ix}$ is chosen to be $-\underline{\Lambda}_y'\Sigma_y^{-1}(\underline{\Lambda}_y'\Sigma_y^{-1}\underline{\Lambda}_y)^{-1/2}\underline{\mathbf{Y}}$ and $\mathbf{s}_{iy}$ is chosen to be $-\underline{\Lambda}_x'\Sigma_x^{-1}(\underline{\Lambda}_x'\Sigma_x^{-1}\underline{\Lambda}_x)^{-1/2}\underline{\mathbf{X}}$, then $\rho(\boldsymbol{\theta}_{ix},\boldsymbol{\theta}_{iy})=-1$. One can then conclude that, even though $E(\mathbf{X}_t\mathbf{Y}_s)=0$ $\forall$ t,s, so that $R_{\mathbf{Y}j.\underline{\mathbf{X}}}=0$, j=1..p, when $\rho^*_x=0=\rho^*_y$, set C_X nevertheless contains common factors that range from being perfectly negatively linearly related to perfectly positively linearly with the common factors contained in set C_Y.

Theorem 6 (Perfect prediction of Z by {$\theta_i,\underline{\delta}_i$}; Schonemann and Steiger, 1978): Let $\underline{\mathbf{X}}$ be ulcf representable, and let its ulcf representation be $\underline{\mathbf{X}}=\Lambda_o\boldsymbol{\theta}+\Psi_o^{1/2}\underline{\boldsymbol{\delta}}$. For any external variate $\mathbf{Z}$, there always exists a set of factors to $\underline{\mathbf{X}}$, say, $\{\boldsymbol{\theta}_z,\underline{\boldsymbol{\delta}}_z\}$ such that $R^2_{\mathbf{Z}.\{\boldsymbol{\theta}z,\underline{\boldsymbol{\delta}}z\}}$, the squared multiple correlation between $\mathbf{Z}$ and $\{\boldsymbol{\theta}_z,\underline{\boldsymbol{\delta}}_z\}$, is equal to unity.

Proof

With no loss of generality, let $E(\mathbf{Z})=0$ and $V(\mathbf{Z})=1$. External variate $\mathbf{Z}$ can be expressed as

$$(4.80)\quad \mathbf{Z}=\underline{\sigma}'_{\underline{\mathbf{X}},\mathbf{z}}\Sigma^{-1}\underline{\mathbf{X}}+(\mathbf{Z}-\underline{\sigma}'_{\underline{\mathbf{X}},\mathbf{z}}\Sigma^{-1}\underline{\mathbf{X}}),$$

in which the first term is the linear predictor of $\mathbf{Z}$ by $\underline{\mathbf{X}}$, and the second term is the corresponding residual variate. Define $\mathbf{s}_z$ to be $(\mathbf{Z}-\underline{\sigma}'_{\underline{\mathbf{X}},\mathbf{z}}\Sigma^{-1}\underline{\mathbf{X}})(1-R^2_{\mathbf{z}.\underline{\mathbf{X}}})^{-1/2}$ so that

$$(4.81)\quad \mathbf{Z}=[\underline{\mathbf{X}}'\ \mathbf{s}_z]\begin{bmatrix}\Sigma^{-1}\underline{\sigma}_{\mathbf{x},\mathbf{z}}\\ (1-R^2_{\mathbf{z}.\underline{\mathbf{x}}})^{\frac{1}{2}}\end{bmatrix}$$

and note that $E(\mathbf{s}_z)=0$, $V(\mathbf{s}_z)=1$, and $C(\underline{\mathbf{X}},\mathbf{s}_z)=\underline{0}$. It then follows that the set of random variates

(4.82) $[\boldsymbol{\theta}_z\ \underline{\boldsymbol{\delta}}_z]=[\underline{\mathbf{X}}'\ \mathbf{s}_i]B$,

in which $B=\begin{bmatrix}\Sigma^{-1}\underline{\Lambda}_o & \Sigma^{-1}\Psi_o^{\frac{1}{2}} \\ w^{\frac{1}{2}} & -\underline{\Lambda}_o\Psi_o^{-\frac{1}{2}}w^{\frac{1}{2}}\end{bmatrix}$ are factors to $\underline{\mathbf{X}}$. Finally, from (4.81) and (4.82), it follows that

(4.83) $\mathbf{Z}=[\boldsymbol{\theta}_z\ \underline{\boldsymbol{\delta}}_z]B^{-1}\begin{bmatrix}\Sigma^{-1}\underline{\sigma}_{\underline{\mathbf{x}},\mathbf{z}} \\ (1-R^2_{\mathbf{z}.\underline{\mathbf{x}}})^{\frac{1}{2}}\end{bmatrix}$,

i.e., that $\mathbf{Z}$ is a linear transformation of $\{\boldsymbol{\theta}_z\ \underline{\boldsymbol{\delta}}_z\}$. Hence, $R^2_{\mathbf{Z}.\{\boldsymbol{\theta}z,\underline{\boldsymbol{\delta}}z\}}$ is equal to unity◊

Working with finite dimensional arrays, Schonemann and Haagen (1987) used this result to illustrate the absurd fact that, if $\underline{\mathbf{X}}$ is ulcf representable, then, for *any* criterion measure whatsoever, e.g., the dates of Easter Sunday from 1960 to 1979, or the shoe sizes in inches of the members of *P*, there is a set of factors to $\underline{\mathbf{X}}$ (one common and p specific) that predicts this criterion with multiple correlation equal to unity.