



Thailand Statistician
October 2021; 19(4): 721-733
<http://statassoc.or.th>
Contributed paper

Equi-covariable Composite – A Connection to Equal-row-sum Covariance Matrix

Saran Ishika Maiti*

Department of Statistics, Visva-Bharati, Santiniketan, India

*Corresponding author; e-mail: saran.ishika@gmail.com

Received: 3 December 2019

Revised: 3 February 2020

Accepted: 26 March 2020

Abstract

The concept of a class of equi-covariable composites (EC class) is introduced when the multivariate data are the outcomes of a repeated-measurement experiment as appears frequently in psychometric, biometric or social studies. Along with interesting properties of such class, the distinction between the generator (generating the EC class) and the best linear unbiased estimator (BLUE), introduced by C.R. Rao, is pointed out. Surprisingly, when the average of the variables belongs to all EC class the covariance matrix (Σ) of the associated variables should maintain equal row (or column) sums (ERS). Possibility of an extended EC class by sequential augmentation of new variables could arise when Σ would have “stair-case” ERS structure which maintains ERS property right from first variable through the last variable.

Keywords: Linear composite, equi-covariability, equal-row-sum matrix, bipolar covariance matrix

1. Introduction

Sometimes the measuring variables in a set refer to the responses collected on the same experimental unit at successive times or under a variety of experimental conditions (experimentations) recorded at the same scale of measurement. This strategy is called repeated-measurement experiment. A popular repeated-measure is the crossover study in which subjects receive a sequence of different treatments (or exposures). Repeated measurement experiments are pretty common in many scientific disciplines, for example psychology, pharmaceutical science, and health-care science. Good sources of theoretical details are Crowder & Hand (1990); Laird et al. (1992); Lindsey (1993); Everitt (1995). In order to analyse repeated measures data, most of the time the bundle of measuring variables are combined or pooled to suitably construct what is termed composite for subsequent uses. Using composite variables controls type I error rate (e.g., when a sample size is not sufficient for testing multiple comparisons), addresses multicollinearity for regression analysis, or organizing multiple highly correlated variables into more digestible interpretation. The composite, constructed as a linear combination of variables, is named as linear composite. The resulting linear composite variable corresponds to the latent dimension in the data that best summarize the overall structure of the

original variables (Song et al. 2013). Carey (1998) defines linear composites as an inevitable outcome of linear transformation. He studies some important properties of linear composites. Further, a more general formulation is constructed on matrix linear component as well. Still by far statistical literature lacks decent amount of discussions in this issue.

The variability of the composite variable depends on the individual variances and the covariances of the original variables. On the basis of the variance pattern of the variables, this linear composite exhibits interesting algebraic results. If the variables are such that they have equal variances and equal covariances creating an intraclass structure in their covariance matrix, the sum (or average) of those variables, as a (linear) composite, would satisfy the special criterion of equal covariability of the composite with all the variables. However, on the converse, there exist situations where an equi-covariable composite may be the sum (or average) for non-intraclass covariance matrix of the variables.

The present paper attributes in searching a class of equi-covariable composites (EC) and the characterization of covariance matrix under equi-covariable concept. Depending upon the covariance matrix, there exists a member belonging to such class, which could generate all other co-members and hence is termed generator of the class. It possesses the largest equal covariability with all other measuring variables. This paper is organized as follows. Section 2 discusses the construction of equi-covariable class followed by some propositions enabling the sum(or average) of the variables as the generator of equi-covariable composite class when the row totals (or column totals) of the covariance matrix are all equal (ERS). An exact statistical test for the tenability of ERS structure of a covariance matrix is formulated in Section 3.

The possibility of pooling two EC classes to create a new EC class (Pooled EC class) is explored in Section 4 while Section 5 delves into the possibility of equi-covariable pooling on ERS covariance structure. Creation of extended EC class through the subsequent inclusion (augmentation) of extra variables (being measured on the same measuring scale) is studied in Section 6. Section 7 documents the evolution of “staircase” ERS covariance structure maintaining certain interesting features followed by an illustrative example in Section 8. Finally some concluding remarks are placed in Section 9.

2. Equi-covariable Composite

Definition 1 For a vector of constants γ , a linear composite $Y = \gamma'X$ is said to have the property of equi-covariability with respect to a random vector $X = (X_1, X_2, \dots, X_p)'$ if $cov(Y, X_1) = cov(Y, X_2) = \dots = cov(Y, X_p)$.

2.1. Construction

Let $\Sigma =$ Covariance matrix of X . $Cov(X_i, Y) = e_i' \Sigma \gamma$ where $e_i = (0, 0, \dots, 1, \dots, 0)'$ with 1 appearing at the i -th position, $i = 1(1)p$. Let g be the common covariance between Y and any component variable [Maiti 2016]. Then $e_i' \Sigma \gamma = g$ for all $i = 1(1)p$. Combining them, a system of non-homogeneous linear equations is obtained as

$$\Sigma \gamma = g \mathbf{J}_p .$$

Under the assumption of positive definiteness of Σ , the solution to the above system is $\gamma = g \Sigma^{-1} \mathbf{J}_p$ where \mathbf{J}_p is the p -component vector of ones. Consequently,

$$Y = g(\Sigma^{-1} \mathbf{J}_p)' \mathbf{X} = g \mathbf{J}'_p \Sigma^{-1} \mathbf{X}.$$

In respect of Σ , a class of equi-covariable composites (EC) of \mathbf{X} may be demarcated with members being identifiable by the scalar $g (> 0)$.

2.2. Searching of largest equi-covariable composite

As covariance does not numerically exceed the larger variance of the two concerned variables, the present investigation confines to those composites that would satisfy $Cov(\gamma' \mathbf{X}, X_j) \leq V(\gamma' \mathbf{X})$ for all $j = 1, 2, \dots, p$. Let us term the “largest” equi-covariable composite (LEC) when a composite is such that $Cov(\gamma' \mathbf{X}, X_j)$ attains uppermost limit $V(\gamma' \mathbf{X})$ for all $j = 1, 2, \dots, p$. In other words, LEC is the member belonging to EC class when g is chosen as the variance of Y . On fixing $g = Var[(g \Sigma^{-1} \mathbf{J}_p)' \mathbf{X}]$, the “optimum” value of g would be $(\mathbf{J}'_p \Sigma^{-1} \mathbf{J}_p)^{-1}$ providing the expression of LEC as

$$Y_0 = \frac{\mathbf{J}'_p \Sigma^{-1}}{\mathbf{J}'_p \Sigma^{-1} \mathbf{J}_p} \mathbf{X}. \tag{1}$$

Y_0 may be otherwise looked upon as the “generator” generating the class as $Y = g^* Y_0$ where $g^* = g(\mathbf{J}'_p \Sigma^{-1} \mathbf{J}_p), 0 < g^* \leq 1$.

Proposition 1 Based on the covariance matrix Σ of a random vector $\mathbf{X}(p \times 1)$, a class of equi-covariable composites(EC) may be generated by the generator $Y_0 = \gamma'_0 \mathbf{X}$ where $\gamma_0 = \frac{\Sigma^{-1} \mathbf{J}_p}{\mathbf{J}'_p \Sigma^{-1} \mathbf{J}_p}$, the typical member belonging to the family possessing $Y = g_0 \gamma'_0 \mathbf{X}, 0 < g_0 \leq 1$. Any member (Y) of EC class has the variance $V(Y) \leq V(Y_0)$.

Proposition 2 If the equi-covariable loading vector is proportional to certain principal component loading vector, then Σ would have equal row (or column) sums.

Proof: Let $\beta' \mathbf{X}$ be a principal component of \mathbf{X} such that $\Sigma \beta = \lambda \beta (\lambda > 0)$. According to the given condition, for any equivariable composite $Y = \gamma' \mathbf{X}$,

$$\gamma \propto \beta \Rightarrow \Sigma^{-1} \mathbf{J}_p \propto \beta \Rightarrow \mathbf{J}_p \propto \Sigma \beta \Rightarrow \mathbf{J}_p \propto \lambda \beta \Rightarrow \beta \propto \mathbf{J}_p \Rightarrow \Sigma \mathbf{J}_p \propto \mathbf{J}_p$$

which implies that row totals of Σ are all equal.

2.3. Average as a generator

If the entries of Σ are such that they maintain equal-row-sums(ERS), then due to Proposition 2, $\Sigma \mathbf{J}_p \propto \mathbf{J}_p$ as well as $\Sigma^{-1} \mathbf{J}_p \propto \mathbf{J}_p$. Consequently, the generator/LEC Y_0 [vide Equation (1)] simplifies to average of the components of $\mathbf{X}(\bar{X})$. For instance, when Σ possesses an intra-class structure, Y_0 is no other than \bar{X} .

2.4. BLUE and LEC

In the perspective of estimating the same parameter unbiasedly by a number of competent unbiased estimators, the linear combination(composite) having the least variance is termed best linear unbiased estimator (BLUE). BLUE relates to unbiased estimator while LEC to equi-covariability. Eventually, the expression in Equation (1) of LEC would be identical to BLUE when expectations of the component variables are all equal.

Noting that $V(Y_0) = (\mathbf{J}'_p \Sigma^{-1} \mathbf{J}_p)^{-1}$ and $Var(\bar{X}) = (\mathbf{J}'_p \Sigma \mathbf{J}_p)/p^2$, by the extended Cauchy-Schwarz's inequality (Johnson 1996, Chapter 2)

$$Var(\bar{X}) \geq Var(Y_0) . \tag{2}$$

The equality in Equation (2) holds when Σ is ERS. Thus unless Σ satisfies ERS condition, Y_0 would not be reducible to \bar{X} . Conversely, unless the expectations of the component variables are all equal, an LEC \bar{X} would no longer become BLUE.

Proposition 3 (i) Correlation between any two members belonging to a EC class is unity.

(ii) Correlation between any member of a EC and \bar{X} (average) is $\frac{p}{\sqrt{(\mathbf{J}'_p \Sigma \mathbf{J}_p)(\mathbf{J}'_p \Sigma^{-1} \mathbf{J}_p)}}$.

Its maximum value will be unity if and only if Σ is ERS providing \bar{X} as the generator.

Recalling Proposition 2, Proposition 3(ii) has an implication that for a ERS Σ , $\sqrt{p}\bar{X}$ is not only a principal component but also a member of the EC class generated by \bar{X} .

3. A Statistical Test for the Tenability of ERS Σ

A statistical test procedure may be forwarded to test the tenability of ERS structure of Σ under the assumption of multinormality of the random vector $\mathbf{X}(p \times 1)$. Let \mathbf{S} be the data covariance matrix based on the $(p \times n)$ data set comprising the measurement vectors of n individuals.

As \mathbf{J}/\sqrt{p} is necessarily an eigenvector of any ERS matrix, the tenability of ERS structure of Σ may be assured equivalently by the tenability of \mathbf{J}/\sqrt{p} as an eigenvector of Σ . Following the testing procedure as provided by Mallows 1961, an F-statistic is computed using the expression given below.

$$F = \left(\frac{n - p}{p - 1} \right) \left[\frac{(\mathbf{J}'_p \mathbf{S} \mathbf{J}_p)(\mathbf{J}'_p \mathbf{S}^{-1} \mathbf{J}_p)}{p^2} - 1 \right] . \tag{3}$$

If the computed value of F is less than $F_\alpha(p - 1, n - p)$, the tenability of a ERS Σ is asserted at 100 α % level of significance.

Consequently, the common row total (T_0) of ERS Σ may be estimated by noticing that the variable $(\mathbf{J}/\sqrt{p})' \mathbf{X}$ is distributed as univariate normal with variance T_0 . Considering the joint distribution of n such variables in respect of n individuals, the maximum likelihood estimate (MLE) of T_0 may be obtained as

$$\hat{T}_0 = \frac{p}{n} \sum_{i=1}^n (\bar{x}_i - \bar{\bar{x}})^2 ,$$

where \bar{x}_i is the average score obtained by i^{th} individual while $\bar{\bar{x}}$ is the overall averaged score ($\frac{1}{n} \sum_{i=1}^n \bar{x}_i$). Noticeably, the least squares estimate of T_0 based on p row totals of \mathbf{S} would be found as $\frac{1}{p} \sum_g \sum_h s_{gh}$ which is algebraically equal to \hat{T}_0 .

The above-discussed test procedure can be devised as a handy tool in testing of circumplex model in psychology. In circumplex model, the variables do not clump orthogonally along the two axes. Rather they group in equal spacings around the circumference of a circle (Guttman 1954; Gurtman et al. 2003). Notably, a circumplex covariance matrix is ERS. Statistical tests for the circumplex covariance structure may be available elsewhere (Nagar et al. 1988). But it might be wise to test on ERS structure before testing on more complex test of circumplex structure.

4. Equi-covariable Pooling of Equi-covariable Classes

Let $Y_1 = g_1(\Sigma_{11}^{-1} \mathbf{J}_p)' \mathbf{X}^{(1)}$, $g_1 > 0$ and $Y_2 = g_2(\Sigma_{22}^{-1} \mathbf{J}_q)' \mathbf{X}^{(2)}$, $g_2 > 0$ be the typical members belonging to two EC classes based on two random vectors $\mathbf{X}^{(1)}$ ($p \times 1$) and $\mathbf{X}^{(2)}$ ($q \times 1$) respectively. Σ_{ii} =Covariance matrix ($\mathbf{X}^{(i)}$), $i = 1, 2$. Our point of interest is to investigate if it is possible to construct a new EC class by pooling these two EC classes. To be more precise, is it possible to get a EC class based on $(p + q)$ -component random vector $\mathbf{X} = (\mathbf{X}^{(1)'} \mathbf{X}^{(2)'})'$, whose typical member Y may be expressed,by definition, as $Y = g(\Sigma^{-1} \mathbf{J}_{p+q})' \mathbf{X}$, $g > 0$? To answer let us consider the partitioned form of Σ along with two forms of its inverse as follows:

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

$$\begin{aligned} \Sigma^{-1} &= \begin{pmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22.1}^{-1} \Sigma_{21} \Sigma_{11}^{-1} & -\Sigma_{11}^{-1} \Sigma_{12} \Sigma_{22.1}^{-1} \\ -\Sigma_{22.1}^{-1} \Sigma_{21} \Sigma_{11}^{-1} & \Sigma_{22.1}^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_{11.2}^{-1} & -\Sigma_{11.2}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \\ -\Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11.2}^{-1} & \Sigma_{22}^{-1} + \Sigma_{22}^{-1} \Sigma_{21} \Sigma_{11.2}^{-1} \Sigma_{12} \Sigma_{22}^{-1} \end{pmatrix}, \end{aligned}$$

where $\Sigma_{11.2} = \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}$ and $\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21} \Sigma_{11}^{-1} \Sigma_{12}$.

Keeping in view the expressions of Y_1, Y_2 and Y , $\Sigma^{-1} \mathbf{J}_{p+q}$ may be reducible to

$$\begin{aligned} &\begin{pmatrix} \Sigma_{11}^{-1} \mathbf{J}_p \\ \Sigma_{22}^{-1} \mathbf{J}_q \end{pmatrix} \text{ under Condition A: } \Sigma_{12} = \mathbf{0}, \\ &\begin{pmatrix} \Sigma_{11}^{-1} \mathbf{J}_p \\ \mathbf{0} \end{pmatrix} \text{ under Condition B: } \Sigma_{21} \Sigma_{11}^{-1} \mathbf{J}_p = \mathbf{J}_q \\ \text{and } &\begin{pmatrix} \mathbf{0} \\ \Sigma_{22}^{-1} \mathbf{J}_q \end{pmatrix} \text{ under Condition C: } \Sigma_{12} \Sigma_{22}^{-1} \mathbf{J}_q = \mathbf{J}_p. \end{aligned}$$

A class of pooled equi-covariable composites (termed **PEC** henceforth) may be obtainable by choosing $g_1 = g_2 (= g, \text{ say})$, i.e., $Y = g(\Sigma^{-1} \mathbf{J}_{p+q})' \mathbf{X} = g(\Sigma_{11}^{-1} \mathbf{J}_p)' \mathbf{X}^{(1)} + g(\Sigma_{22}^{-1} \mathbf{J}_q)' \mathbf{X}^{(2)}$. The generator generating the PEC class is given by

$$Y_0 = \frac{\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{X}^{(1)} + \mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{X}^{(2)}}{\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{J}_p + \mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{J}_q} \tag{4}$$

with variances under conditions **A,B** and **C** respectively as

$$\begin{aligned} Var_{\mathbf{A}}(Y_0) &= (\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{J}_p + \mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{J}_q)^{-1} \\ Var_{\mathbf{B}}(Y_0) &= \frac{\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{J}_p + 3\mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{J}_q}{(\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{J}_p + \mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{J}_q)^2} \\ Var_{\mathbf{C}}(Y_0) &= \frac{3\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{J}_p + \mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{J}_q}{(\mathbf{J}'_p \Sigma_{11}^{-1} \mathbf{J}_p + \mathbf{J}'_q \Sigma_{22}^{-1} \mathbf{J}_q)^2}, \end{aligned}$$

Remark 1 If both the conditions **B** and **C** hold simultaneously, then $\Sigma_{11.2} \Sigma_{11}^{-1} \mathbf{J}_p = \mathbf{0}$, $\Sigma_{22.1} \Sigma_{22}^{-1} \mathbf{J}_q = \mathbf{0}$ and $\Sigma \text{diag}(\Sigma_{11}^{-1}, -\Sigma_{22}^{-1}) \mathbf{J}_{p+q} = \mathbf{0}$ indicating the singularities of $\Sigma_{11.2}$, $\Sigma_{22.1}$ and Σ .

5. Possibility of Equi-covariable Pooling on ERS Covariance Structures

For ERS Σ_{11} and ERS Σ_{22} ,

$$\Sigma_{11}\mathbf{J}_p = c_1\mathbf{J}_p, \Sigma_{22}\mathbf{J}_q = c_2\mathbf{J}_q, c_1 > 0, c_2 > 0.$$

Consequently, EC classes formed by $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ would have the generators as their component averages $\bar{X}^{(1)}$ and $\bar{X}^{(2)}$. Clearly, using Equation (4) the corresponding PEC class may be generated by the generator as below

$$Y_0 = \frac{\frac{p}{c_1}\bar{X}^{(1)} + \frac{q}{c_2}\bar{X}^{(2)}}{\frac{p}{c_1} + \frac{q}{c_2}}.$$

Subsequently, the variances of Y_0 , under condition **A**, **B** and **C** turn to be the following.

Condition A: Under $\Sigma_{12} = \mathbf{0}$ with corresponding variance of the generator Y_0 ,

$$Var_{\mathbf{A}}(Y_0) = \left(\frac{p}{c_1} + \frac{q}{c_2}\right)^{-1}.$$

Condition B: Under $\Sigma_{21}\mathbf{J}_p = c_1\mathbf{J}_q$ with corresponding variance of the generator Y_0 ,

$$Var_{\mathbf{B}}(Y_0) = \frac{\frac{p}{c_1} + 3\frac{q}{c_2}}{\left(\frac{p}{c_1} + \frac{q}{c_2}\right)^2}.$$

Condition C: Under $\Sigma_{12}\mathbf{J}_q = c_2\mathbf{J}_p$ with corresponding variance of the generator Y_0 ,

$$Var_{\mathbf{C}}(Y_0) = \frac{3\frac{p}{c_1} + \frac{q}{c_2}}{\left(\frac{p}{c_1} + \frac{q}{c_2}\right)^2}.$$

It may be noted that Σ would have ERS property if $c_1 = c_2$ in which case corresponding Y_0 would be equal to $\frac{p\bar{X}^{(1)} + q\bar{X}^{(2)}}{p+q}$.

5.1. A demonstration on bi-polar covariance matrix

A $(p + q)$ - component random vector \mathbf{X} is said to be bipolar if both the 1st p -component subvector $\mathbf{X}^{(1)}$ and 2nd q -component subvector $\mathbf{X}^{(2)}$ have intraclass covariance structures and any component of $\mathbf{X}^{(1)}$ has with any component of $\mathbf{X}^{(2)}$ the same covariance (Roy 1954). The dispersion matrix is thus expressed in the following partitioned form.

$$\Sigma_{(p+q) \times (p+q)} = \left(\begin{array}{c|c} (a-b)\mathbf{I}_p + b\mathbf{J}_p\mathbf{J}_p' & c\mathbf{J}_p\mathbf{J}_q' \\ \hline c\mathbf{J}_q\mathbf{J}_p' & (d-e)\mathbf{I}_q + e\mathbf{J}_q\mathbf{J}_q' \end{array} \right).$$

Its $(p + q)$ eigenvalues are given by,

$$\begin{aligned} \lambda_1, \lambda_2 &= \frac{1}{2}[(a + \overline{p-1}b + d + \overline{q-1}e) \pm \sqrt{(a + \overline{p-1}b - d - \overline{q-1}e)^2 + 4pqc^2}], \\ \lambda_3 = \lambda_4 &= \dots = \lambda_{p+1} = a - b \text{ and } \lambda_{p+2} = \lambda_{p+3} = \dots = \lambda_{p+q} = d - e. \end{aligned}$$

To ensure positive definiteness of Σ_{11} , Σ_{22} and Σ , all the above eigenvalues are to be positive provided that five parameters a, b, c, d and e would maintain the following inequalities.

$$a > 0, d > 0, b < a, e < d, a + \overline{(p-1)}b > 0, d + \overline{(q-1)}e > 0 \text{ and } (a + \overline{p-1}b)(d + \overline{q-1}e) > pqc^2.$$

Assuming positive definiteness of Σ_{11} and Σ_{22} , the constructable EC classes by $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ would have typical members respectively as follows:

$$Y_1 = g_1 \frac{\mathbf{J}'_p \boldsymbol{\Sigma}_{11}^{-1} \mathbf{X}^{(1)}}{\mathbf{J}'_p \boldsymbol{\Sigma}_{11}^{-1} \mathbf{J}_p} = g_1 \bar{X}^{(1)},$$

$$Y_2 = g_2 \frac{\mathbf{J}'_q \boldsymbol{\Sigma}_{22}^{-1} \mathbf{X}^{(2)}}{\mathbf{J}'_q \boldsymbol{\Sigma}_{22}^{-1} \mathbf{J}_q} = g_2 \bar{X}^{(2)},$$

where $g_1 > 0, g_2 > 0$. It is to be pointed out that due to intraclass structures of $\boldsymbol{\Sigma}_{11}$ and $\boldsymbol{\Sigma}_{22}$, both of them have ERS property.

For possible creation of a new EC formable by pooling the members of the above-mentioned EC classes, let us first choose $g_1 = g_2 = g$. Next, let us pay attention to the conditions **A**, **B** and **C**. For condition **A** to hold, $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ should be uncorrelated so as to imply $c = 0$ in which case the desired PEC class would have the typical member $Y = g(\bar{X}^{(1)} + \bar{X}^{(2)})$ with $Var_{\mathbf{A}}(Y) = g^2 \left[\frac{a+p-1b}{p} + \frac{d+q-1e}{q} \right], g > 0$. Or, equivalently, $Y = g^* Y_0, g^* \in (0, 1)$, the generator being

$$Y_0 = \frac{w_1 \bar{X}^{(1)} + w_2 \bar{X}^{(2)}}{w_1 + w_2} \text{ and } Var_{\mathbf{A}}(Y_0) = (w_1 + w_2)^{-1},$$

where $w_1 = \frac{p}{a+p-1b}$ and $w_2 = \frac{q}{d+q-1e}$.

In the case of uncorrelated $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$, $c \neq 0$, condition **B** or **C** is to be checked for the construction of PEC class. Condition **B** requires $c = \frac{d+q-1e}{q}$ while condition **C** requires $c = \frac{a+p-1b}{p}$. $Var_{\mathbf{B}}(Y_0)$ and $Var_{\mathbf{C}}(Y_0)$ may be derived easily.

Remark 2 If both the conditions **B** and **C** hold together, a, b and c would become restricted by the equations

$$\frac{a + \overline{p-1b}}{p} = \frac{d + \overline{q-1e}}{q} = c$$

leading to a singular matrix $\boldsymbol{\Sigma}$ with rank $p + q - 1$ due to $\lambda_2 = 0$. However, PEC class may be obtained by $Y_0 = \frac{\bar{X}^{(1)} + \bar{X}^{(2)}}{2}$ having $Var_{\mathbf{BC}}(Y_0) = c(> 0)$. Under such set-up, canonical correlation between $\mathbf{X}^{(1)}$ and $\mathbf{X}^{(2)}$ would be unity with corresponding canonical composite pairs $(\bar{X}^{(1)}, \bar{X}^{(2)})$.

6. Extension of EC Class by Augmentation

It was noted that the PEC class, while pooling two EC classes, requires the fulfillment of any one of the three conditions **A**, **B** and **C** which are too restrictive. It is rather possible to exercise less restrictive condition when one EC class of $(p - 1)$ variables $(X_1, X_2, \dots, X_{p-1})$ could be extended to a new EC class by appending a new random variable X_p .

Let $\boldsymbol{\Sigma}_{p-1} (> 0)$ and $\boldsymbol{\Sigma}_p (> 0)$ be the covariance matrices of $(X_1, X_2, \dots, X_{p-1})$ and $(X_1, X_2, \dots, X_{p-1}, X_p)$ respectively. If both of them are ERS with row totals g_{p-1} and g_p , then

$$\boldsymbol{\Sigma}_p = \begin{pmatrix} \boldsymbol{\Sigma}_{p-1} & a_p \mathbf{J}_{p-1} \\ a_p \mathbf{J}'_{p-1} & g_p - (p-1)a_p \end{pmatrix}, \quad (5)$$

where $a_p = g_p - g_{p-1}$. To ensure non-singularity (positive definiteness) of $\boldsymbol{\Sigma}_p$, $g_p \neq \frac{p}{p-1} g_{p-1}$. The generators of EC classes corresponding to $\boldsymbol{\Sigma}_{p-1}$ and $\boldsymbol{\Sigma}_p$ are just the averages $(\bar{X}^{(p-1)})$ and $\bar{X}^{(p)}$ respectively of the variables on which they are based.

7. Staircase ERS Structure

An interesting extension of ERS structure can be presented through sequential augmentation of variables and thus obtaining stepwise ERS structure. Let us consider a sequence of random variables by augmentation as $X_1, (X_1, X_2), (X_1, X_2, X_3), \dots, (X_1, X_2, \dots, X_{p-1})$ all maintaining ERS property in their covariance matrices. Then, in effect, the covariance matrix of (X_1, X_2, \dots, X_p) may have a structure formable by a maximum p number of functionally independent parameters, say a_1, a_2, \dots, a_p as follows:

$$\Sigma = \begin{pmatrix} a_1 & a_2 & a_3 & \dots & a_{p-1} & a_p \\ a_2 & g_2 - a_2 & a_3 & \dots & a_{p-1} & a_p \\ a_3 & a_3 & g_3 - 2a_3 & \dots & a_{p-1} & a_p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{p-1} & a_{p-1} & a_{p-1} & \dots & g_{p-1} - (p-2)a_{p-1} & a_p \\ a_p & a_p & a_p & \dots & a_p & g_p - (p-1)a_p \end{pmatrix}, \quad (6)$$

where $g_j = a_1 + a_2 + \dots + a_j, j = 2, \dots, p$ and $g_1 = a_1 (> 0)$. Subscriptized by $p, \Sigma_p (= \Sigma)$ would have the partitioned form as furnished in Equation (5). Notationally, j variables (X_1, X_2, \dots, X_j) would have ERS covariance matrix Σ_j which is identifiable by j basic parameters (a_1, a_2, \dots, a_j) with row totals $g_j = a_1 + \dots + a_j, j = 2(1)p$ and $g_1 = a_1 (> 0)$.

The structure of Σ as shown in Equation (6) is named as **staircase ERS structure**.

7.1. Features of Staircase ERS Structure

- i) If the diagonal entries in Σ_p [vide Equation (6)] be equal, then off-diagonal are also equal and thus the resulting Σ_p would have an intraclass covariance structure.
- ii) Σ_p belongs to the spectrally decomposable class in the sense that there exists an orthogonal matrix (\mathbf{P}) such that $\mathbf{P}'\Sigma_p\mathbf{P} = \mathbf{D}_\lambda$, a diagonal matrix with eigenvalues as its diagonal elements where the elements of \mathbf{P} are free from the elements of Σ_p (Mukherjee, 1981). For Equation (6), required \mathbf{P} is the famous Helmert's matrix given by

$$\mathbf{P} = \mathbf{P}_0\mathbf{D}_0 \tag{7}$$

where

$$\mathbf{P}'_0 = \begin{pmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ 1 & 1 & -2 & 0 & \dots & 0 \\ 1 & 1 & 1 & -3 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & 1 & \dots & -(p-1) \\ 1 & 1 & 1 & 1 & \dots & 1 \end{pmatrix}$$

and $\mathbf{D}_0 = \text{diag}(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{12}}, \dots, \frac{1}{\sqrt{p(p-1)}}, \frac{1}{\sqrt{p}})$. The eigenvalues, *i.e.* the diagonal entries of \mathbf{D}_λ are given by

$$\lambda = (\lambda_1, \lambda_2, \dots, \lambda_p) = \mathbf{P}'_0\mathbf{a}.$$

Conversely, $\mathbf{a} = (a_1, a_2, \dots, a_p)' = \mathbf{P}_0\mathbf{D}_0^2\lambda$. Thus there is one-to-one and unique correspondence between λ and \mathbf{a} .

In explicit form, the eigenvalues are given by

$$\begin{aligned} \lambda_j &= j(\bar{a}^{(j)} - a_{j+1}), j = 1, 2, \dots, p - 1 \\ \lambda_p &= p\bar{a}^{(p)}, \end{aligned}$$

where $\bar{a}^{(j)}$ denotes the average of $a_1, a_2, \dots, a_j, j = 2(1)p$ and $\bar{a}^{(1)} = a_1 (> 0)$.

iii) If the eigenvalues of Σ_{p-1} are denoted by $\lambda_1^*, \lambda_2^*, \dots, \lambda_{p-2}^*$ and λ_{p-1}^* , then those of Σ_p would be given by

$$\begin{aligned} \lambda_j &= \lambda_j^*, j = 1, 2, \dots, p - 2 \\ \lambda_{p-1} &= \lambda_{p-1}^* - (p - 1)a_p \\ \lambda_p &= \lambda_{p-1}^* + a_p. \end{aligned}$$

Thus the parameter a_p is involved in the last two eigenvalues only.

iv) For a positive definite Σ_p , λ_j 's are all positive due to which the parameters a_1, a_2, \dots, a_p must maintain the following inequalities.

$$a_2 < \bar{a}^{(1)}, a_3 < \bar{a}^{(2)}, a_4 < \bar{a}^{(3)}, \dots, a_p < \bar{a}^{(p-1)}.$$

And thus $a_1 = \max_{1 \leq j \leq p} a_j$.

7.2. Step-average vector in Staircase ERS Structure

Let us assume that p random variables are subscriptized in accordance with the sequence of p experimentations on the same experimental unit. If the vector $\mathbf{X} = (X_1, X_2, \dots, X_p)'$ maintains the staircase ERS covariance structure Σ_p , the following results may be forwarded.

EC class in respect of a j -component subvector $(X_1, X_2, \dots, X_j)'$ may be generated by j -step average

$$\begin{aligned} \bar{X}^{(j)} &= (X_1 + X_2 + \dots + X_j)/j, j = 1, 2, \dots, p \\ V(\bar{X}^{(j)}) &= (\mathbf{J}'_j \Sigma_j \mathbf{J}_j)/j^2 = \bar{a}^{(j)}, j = 1, 2, \dots, p \\ Cov(\bar{X}^{(j)}, \bar{X}^{(j+t)}) &= \frac{1}{j(j+t)} [\mathbf{J}'_j \Sigma_j \mathbf{J}_j + j(a_{j+1} + a_{j+2} + \dots + a_{j+t})] \\ &= \bar{a}^{(j+t)} \\ &= V(\bar{X}^{(j+t)}). \end{aligned}$$

$$Correlation(\bar{X}^{(j)}, \bar{X}^{(j+t)}) = \sqrt{\frac{\bar{a}^{(j+t)}}{\bar{a}^{(j)}}}, j = 1(1)p, t = 1(1)(p - j),$$

Clearly, $\bar{a}^{(j)}$ is a decreasing function in respect of j . Thus the step-average vector $(\bar{X}^{(1)}, \bar{X}^{(2)}, \dots, \bar{X}^{(p)})'$ would have covariance matrix

$$\mathbf{A} = \begin{pmatrix} \bar{a}^{(1)} & \bar{a}^{(2)} & \bar{a}^{(3)} & \dots & \bar{a}^{(p)} \\ \bar{a}^{(2)} & \bar{a}^{(2)} & \bar{a}^{(3)} & \dots & \bar{a}^{(p)} \\ \bar{a}^{(3)} & \bar{a}^{(3)} & \bar{a}^{(3)} & \dots & \bar{a}^{(p)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \bar{a}^{(p)} & \bar{a}^{(p)} & \bar{a}^{(p)} & \bar{a}^{(p)} & \bar{a}^{(1)} \end{pmatrix}.$$

7.3. A Statistical Test for the Tenability of Staircase ERS Σ

Applying a nonsingular transformation on the step-average vector as

$$\mathbf{Y}' = (Y_1, Y_2, \dots, Y_p) = (\bar{X}^{(1)}, \bar{X}^{(2)}, \dots, \bar{X}^{(p)})\mathbf{U}'_p, \tag{8}$$

where $\mathbf{U}_p = ((u_{ij}))_{p \times p}$ with $u_{ii} = 1$ and $u_{ij} = -1$ for $j = i + 1, i = 1, 2, \dots, p - 1$, corresponding covariance matrix reduces to a diagonal matrix, i.e.

$$\mathbf{U}_p \mathbf{A} \mathbf{U}'_p = \text{diag}[(\bar{a}^{(1)} - \bar{a}^{(2)}), (\bar{a}^{(2)} - \bar{a}^{(3)}), \dots, (\bar{a}^{(p-1)} - \bar{a}^{(p)}), \bar{a}^{(p)}]$$

assuring that Y_1, Y_2, \dots, Y_p are uncorrelated variables.

Under multi-normal set-up on $\mathbf{X}_{p \times 1}$, the testing of null hypothesis of Staircase ERS Σ , presented in Equation (6) is equivalent to testing the stochastic independence of the components of \mathbf{Y} (vide Equation (8)). It may be shown that

$$\mathbf{Y} = \mathbf{D}_0^2 \mathbf{P}'_0 \mathbf{X},$$

where \mathbf{P}_0 and \mathbf{D}_0 are as defined in (7).

The test procedure is based on likelihood ratio criterion on defining

$$\begin{aligned} \Lambda &= \left(\frac{|\mathbf{D}_0^2 \mathbf{P}'_0 \mathbf{S} \mathbf{P}_0 \mathbf{D}_0^2|}{\prod_{j=1}^p (\mathbf{D}_0^2 \mathbf{P}'_0 \mathbf{S} \mathbf{P}_0 \mathbf{D}_0^2)_{jj}} \right)^{\frac{n}{2}} \\ &= \left(\frac{|\mathbf{P}'_0 \mathbf{S} \mathbf{P}_0|}{\prod_{j=1}^p (\mathbf{P}'_0 \mathbf{S} \mathbf{P}_0)_{jj}} \right)^{\frac{n}{2}}, \end{aligned} \tag{9}$$

where \mathbf{S} is the sample covariance matrix based on a random sample of size n and where $(\mathbf{P}'_0 \mathbf{S} \mathbf{P}_0)_{jj}$ denotes the j^{th} diagonal element of $\mathbf{P}'_0 \mathbf{S} \mathbf{P}_0$. $-2 \log_e \Lambda$ is asymptotically distributed as $\chi^2_{p(p-1)/2}$. Instead of $-2 \log_e \Lambda$, Bartlett's modified expression $-2[1 - (2p + 11)/6n] \log_e \Lambda$ may be used for further accuracy. The following remark lights on the number of possibilities of staircase ERS.

Remark 3 If all the variables are allowed to be shuffled, any permutation $(X_{i_1}, X_{i_2}, \dots, X_{i_k}), (i_1, i_2, \dots, i_k) \in (1, 2, \dots, p), k = 2, 3, \dots, p$ may have a candidature for a staircase ERS structure, leaving the situation awesome.

8. An Illustrative Example

A fascinating example on Wechsler Preschool data which is commonly used as a study material in undergrad psychology curriculum is furnished here to check the viability of ERS covariance structure and staircase ERS covariance structure. The data describes the age related trends in the pattern

of mental ability of the children between 4 and $6\frac{1}{2}$ years of age in respect of Wechsler Preschool and Primary Scales (WPPSI), as processed and recorded in the WPPSI Manual (Wechsler 1967). For each age level, 100 boys and 100 girls were administered as many as 11 tests amongst which 8 tests were finally directed to Wechsler Intelligence Scale for Children (WISC). Here, We consider the sample covariance matrix based on the scores of those 8 tests only for 4-year age level (reproduced from Mukherjee (1981, Table 3, p. 221)).

Table 1 Covariances/correlations among eight tests of WPPSI

Test No.	1	2	3	4	5	6	7	8
1. Vocabulary	9.000	0.5100	0.4100	0.3800	0.3800	0.3100	0.4500	0.5700
2. Sentences	4.743	9.610	0.5200	0.3600	0.3400	0.3300	0.3500	0.5300
3. Arithmetic	3.813	4.997	9.610	0.4600	0.4400	0.3900	0.4200	0.5200
4. Block design	3.534	3.460	4.421	9.610	0.3500	0.4300	0.4200	0.3200
5. Geometrical design	3.534	3.267	4.096	3.363	9.610	0.5300	0.4000	0.3800
6. Mazes	2.883	3.171	3.748	4.132	5.093	9.610	0.4700	0.3400
7. Picture completion	3.915	3.146	3.776	3.776	3.596	4.225	8.410	0.4700
8. Block Design	5.310	5.093	4.997	3.075	3.652	3.267	4.225	9.610

(Source: Wechsler Pre-school data, $N = 200$ (100 boys and 100 girls), four-year age level)

It may be observed that row(column) totals of the sample covariance matrix(S) are (36.723, 37.488, 39.458, 35.370, 36.212, 36.130, 35.070, 39.221) indicating a closeness amongst themselves. Such an ERS structure may be hypothesized for the population covariance matrix(Σ). Once we conduct principal component analysis on S , it is noted that the first principal component has the loadings (.352, .361, .381, .334, .344, .341, .332, .379) which are all of same sign (positive) and very close to themselves with the variance 37.038 so as to explain 49.3% of the total variance which is equal to 75.070. More or less equal amount of component loadings indicates the possibility of a ERS structure.

In order to apply Mallow’s test we recall Equation (3). Define s_{gh} as (g, h) element of sample covariance matrix S and s^{gh} as (g, h) element of inverse of sample covariance matrix, i.e., S^{-1} . Here $N = 100, p = 8, \mathbf{J}'_p \mathbf{S} \mathbf{J}_p = \sum_g \sum_h s_{gh} = 295.668, \mathbf{J}'_p \mathbf{S}^{-1} \mathbf{J}_p = \sum_g \sum_h s^{gh} = 0.2188$ resulting to $F_{observed} = 0.1419$. As the tabular value of $F(7, 92)$ at 5% level of significance is 2.1107 which is far greater than the $F_{observed}$, we are obliged to accept the tenability of ERS covariance structure. The estimate of the row total (\hat{T}_0) under such ERS structure [vide Equation (3)] is calculated as 36.958.

Keeping the spirit on, we now proceed to testing the tenability of staircase ERS structure. Taking cue to Equation (9), the computed values of $|\mathbf{P}'_0 \mathbf{S} \mathbf{P}_0|$ and $\Pi_{jj=1}^p (\mathbf{P}'_0 \mathbf{S} \mathbf{P}_0)_{jj}$ are computed as 5.9059×10^{15} and 7.7048×10^{15} respectively. Correspondingly, computed value of $-2\log_e \Lambda$ is 53.17 which is greater than $\chi^2_{05}(28) (= 41.337)$. So staircase ERS structure is not tenable at 5% level of significance.

However, adopting the process of discarding the last variables successively could lead to a “reduced” staircase. The Table 2 shows the statistical tenability of “reduced” staircase ERS indicated by (X_1, X_2, \dots, X_5) by successively discarding X_8, X_7 and X_6 . As a consequence, the j -step-average $\bar{X}^{(j)}$ may be treated as the generator of the EC class for the variables $(X_1, X_2, \dots, X_j), j = 1, \dots, 5$.

Table 2 Tenability of a staircase ERS Structure

p	Discarded variables	$\chi_{obs}^2 = -2\log_e\Lambda$	d.f.	P-value	Decision
8	-	53.17	28	0.0028	Reject
7	X_8	41.92	21	0.0043	Reject
6	X_8, X_7	25.60	15	0.0425	Reject
5	X_8, X_7, X_6	12.27	10	0.2670	Accept

9. Concluding Remarks

The past several years have marked a considerable upsurge of interest in the conceptualization and measurement of reliability in the classical test theory, particularly in regard to measures that are formed as linear composites (weighted or unweighted sums) of individual items. The criterion of equi-covariability, as introduced in Section 2, helps create a class of linear composites (EC class) all of whose members could be obtainable by an appropriate generator. When the covariance matrix possesses equal row (column)-sums (ERS), the generator reduces simply to sum (or equivalently to average) of the measuring variables, being a composite involving no variance or covariance parameters in its expression as such. A staircase ERS covariance matrix of a sequence of measurement variables could produce a sequence of EC classes whose generators are step-wise averages of the variables.

In classical test theory, the reliability of a composite draws much attention to the researchers long since. The “scale score” is defined as the total sum of the scores on a number of items in a test (or on a number of tests comprising a battery)[Kano & Azuma 2003, section 1,pp.141]. Under ERS set-up the scale score would be a clear choice for the generator composite to create a EC class.

Acknowledgments

The author would thank the anonymous reviewers and editor for their helpful comments and meticulous checking on the manuscript that led to an improved paper.

References

- Carey G. Linear transformations and linear composites. Available at: psych.colorado.edu/carey/courses/PSYC7291/handouts/transformation.pdf. 1998.
- Crowder MJ, Hand DJ. Analysis of Repeated Measures. London: Chapman and Hall; 1990.
- Everitt BS. The analysis of repeated measures: A practical review with examples. *J Roy Stat Soc D-STA* (1995); 44(1): 113-135.
- Johnson RA, Wichern DW. Applied multivariate statistical analysis, 3rd Ed. New Delhi: Prentice Hall of India Private Limited; 1996.
- Gurtman B, Pincus A. The Circumplex model: methods and research applications. *Handbook of Psychology, Research Methods in Psychology-Vol 2* edited by Irving B. Weiner, New Jersey: John Wiley & Sons, Inc; 2003.
- Guttman L. A new approach to factor analysis: the radex. In P. Lazarsfeld(Ed). *Mathematical thinking in the social sciences*, The Free Press, Glencoe. 1954. 258-348.
- Kano Y, Azuma Y. Use of SEM programs to precisely measure scale reliability. In *new developments in Psychometrics*. Tokyo: Springer-verlag; 2003.

- Laird NM, Donnelly C, Ware JH. Longitudinal studies with continuous responses. *Stat Methods Med Res.* 1992; 1: 225-247.
- Lindsey J. *Models for repeated measurements.* Oxford: Oxford University Press; 1993.
- Maiti SI. Tau proportionate form of congeneric covariance structure. *Int J Stat Prob.* 2016; 5(5): 1-11.
- Mallows CL. Latent vectors of random symmetric matrices. *Biometrika.* 2061; 48: 133-149.
- Mukherjee BN. Analysis of covariance structures and exploratory factor analysis. *Brit J Math Stat Psy.* 1981; 26: 125-154.
- Nagar DK, Jain SK & Gupta A. On testing circular stationarity and related models. *J Stat Comput Sim.* 1988; 29: 225-239.
- Roy J. On some tests of significance in sample from bi-polar normal distribution. *Sankhya.* 1954; 14: 220-239.
- Song M, Lin F, Ward SE, Fine JP. Composite variables when and how. *Nurs Res.* 2013; 62(1): 45-49.
- Wechsler D. *Manual for the Wechsler Preschool and primary scale of intelligence.* New York: Psychological Corporation; 1967.