AISM (2006) 58: 311–326
DOI 10.1007/s10463-006-0036-x

Kunihiro Baba · Ritei Shibata

Multiplicative correlations

Received: 13 October 2004 / Revised: 27 July 2005 / Published online: 15 March 2006

Abstract A multivariate distribution is said to have multiplicative correlation if the correlation matrix $R = (r_{ij})$ is written as $r_{ij} = \delta_i\delta_j$ or $r_{ij} = -\delta_i\delta_j (i \neq j)$ for a parameter vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$. We first determine feasible values for $\boldsymbol{\delta}$ and show that variables with such a correlation matrix can always be decomposed into a common "signal" variable plus individual "noise" variables. It is also shown that a special case of this correlation matrix implies a sum constraint among variables and vice versa. Such properties illustrate why many multivariate distributions have such a correlation structure. Furthermore, several invariance properties lead to simple relations among several multivariate distributions.

Keywords Correlation modeling · Factorization of variables · Neural science · Partial correlation · Reduction method

1 Introduction

An n-dimensional multivariate distribution is said to have multiplicative correlation if the correlation matrix $R = (r_{ij})$ can be written as $r_{ij} = \delta_i\delta_j$ or $r_{ij} = -\delta_i\delta_j (i \neq j)$ for a parameter vector $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$. A simple example of multivariate distribution which has a multiplicative correlation of the matrix notation $R = \mathrm{diag}(\mathbf{1} \mp \boldsymbol{\delta}^2) \pm \boldsymbol{\delta}\boldsymbol{\delta}^\mathsf{T}$ is the multivariate Poisson or the multivariate gamma distribution derived by the so called *reduction method* (Mardia, 1970, p. 74.) Such a distribution is derived as the joint distribution of $X_i = Z_0 + Z_i (i = 1, \ldots, n)$ where $Z_0, Z_1, \ldots, Z_n$ are independent Poisson or gamma distributed random variables (see, Johnson, Kotz, & Balakrishnan, 1997, p. 139; Kotz, Balakrishnan, &

K. Baba (✉) · R. Shibata
Department of Mathematics, Keio University, 3-14-1 Hiyoshi, Kohoku,
Yokohama 223-8522, Japan
E-mail: baba@sfc.keio.ac.jp

Johnson, 2000, p. 454). If we denote $E(Z_i) = \mathrm{var}(Z_i) = \theta_i (i = 0, 1, \ldots, n)$, then $\boldsymbol{X} = (X_1, \ldots, X_n)$ has a multiplicative correlations,

$$\mathrm{corr}(X_i, X_j) = \sqrt{\frac{\theta_0}{(\theta_0 + \theta_i)}}\sqrt{\frac{\theta_0}{(\theta_0 + \theta_j)}} \quad (i \neq j = 1, \ldots, n).$$

The correlation matrix of this particular example is not only multiplicative, but also has an *equi-covariance* property, where all pairs of variables share the same covariance. A correlation structure similar to equi-covariance is *equi-correlation*. For example, Abbott and Dayan (1999) have proposed the use of covariance matrix

$$Q_{ij} = \sigma^2\{\delta_{ij} + c(1 - \delta_{ij})\} f_i(x) f_j(x) \tag{1}$$

to describe the firing activities of neurons i and j. Here $f_i(x)$ is the mean firing rate of neuron i for input x, δ_{ij} is Kronecker's delta, and c is a parameter. In this model the same correlation c is shared by all pairs of variables and the correlation matrix is clearly multiplicative. However, multiplicative correlation is not limited to such equi-covariance or equi-correlation. Consider the multinomial distribution. The correlation matrix is

$$\mathrm{corr}(X_i, X_j) = -\sqrt{\frac{p_i p_j}{(1 - p_i)(1 - p_j)}},$$

which is multiplicative but there is no equi-covariance nor equi-correlation structure. In this paper, we will investigate the reasons why such a multiplicative correlation appears so frequently.

In Sect. 2, we first give a necessary and sufficient condition for a multiplicative correlation matrix parameterized by $\boldsymbol{\delta}$ to be a correlation matrix, i.e. non-negative definite matrix with all elements less than or equal to one in absolute value.

In Sect. 3, we investigate several implications of multiplicative correlation for the random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$. One is a decomposition result which shows that each X_i can be factorized into the sum of a common "signal" variable and individual "noise" variables. A special case of multiplicative correlation is also shown to be equivalent to the fact that the sample mean $1/n \sum_{i=1}^n X_i$ is almost surely constant.

In Sect. 4, we show that several interesting relations hold among various families of multivariate distributions with multiplicative correlation using an invariance property of multiplicative correlation structure associated with conditioning. Such families include the homogeneous or Liouville distributions. We also show that multiplicative correlations lead to multiplicative partial correlations with a simple relation holding between the two.

2 Multiplicative correlations and covariances

For later discussion, it is better to distinguish two types of multiplicative correlation matrices parameterized by $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_n)$; *positive multiplicative correlation* matrix $R^+(\boldsymbol{\delta}) = \mathrm{diag}(1 - \boldsymbol{\delta}^2) + \boldsymbol{\delta}\boldsymbol{\delta}^\mathsf{T}$ and *negative multiplicative correlation* matrix

$R^-(\boldsymbol{\delta}) = \mathrm{diag}(1+\boldsymbol{\delta}^2)-\boldsymbol{\delta}\boldsymbol{\delta}^\mathsf{T}$. Here $\mathrm{diag}(1\pm\boldsymbol{\delta}^2)$ is a diagonal matrix with its diagonal elements $1 \pm \delta_1^2, \dots, 1 \pm \delta_n^2$. It is easily seen that the two classes of multiplicative correlation matrices $\{R^+(\boldsymbol{\delta})\}$ and $\{R^-(\boldsymbol{\delta})\}$ have no common element, provided that the $\boldsymbol{\delta}$ has more than three non zero elements. Note that the parameterization in each class is unique except for the sign of $\boldsymbol{\delta}$, provided, once again, that the $\boldsymbol{\delta}$ has more than three non zero elements. Royen (1991) derived a multivariate gamma distribution with multiplicative correlations and described such a correlation structure as *one-factorial*. Positive multiplicative correlation is also called *structure l* in Khatri (1967).

We first derive necessary and sufficient conditions for $R^+(\boldsymbol{\delta})$ or $R^-(\boldsymbol{\delta})$ to be a correlation matrix.

2.1 Positive multiplicative correlation

We first give the following lemma to prove Theorem 2.1, although the Theorem 2.1 itself can be proved by an induction. The lemma is useful in its own sight as will be mentioned in Sect. 5.

Lemma 2.1 *The following inequality holds true for the eigenvalues* $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ *of* $R^+(\boldsymbol{\delta})$.

$$\begin{aligned} 1 - \delta_{k_1}^2 &= \lambda_1 = \cdots = \lambda_{n_1-1} < \lambda_{n_1} < 1 - \delta_{k_2}^2 \\ &= \lambda_{n_1+1} = \cdots = \lambda_{n_1+n_2-1} < \lambda_{n_1+n_2} < \cdots < \lambda_{n-n_m} < 1 - \delta_{k_m}^2 \\ &= \lambda_{n-n_m+1} = \cdots = \lambda_{n-1} < \lambda_n, \end{aligned} \tag{2}$$

where $\delta_{k_1}^2 > \delta_{k_2}^2 > \cdots > \delta_{k_m}^2$ *are m distinct values in* $\delta_1^2, \dots, \delta_n^2$ *and* $n_1, \dots, n_m$ $\left(\sum_{i=1}^m n_i = n\right)$ *are the multiplicities of* $\delta_{k_1}^2, \dots, \delta_{k_m}^2$ *respectively.*

In Lemma 2.1, we have used the convention that a series of equations in Eq. (2) is void when the last suffix of λ is less than the first suffix in the equation. For example, $1 - \delta_{k_1}^2 = \lambda_1 = \cdots = \lambda_{n_1-1} < \lambda_{n_1}$ reduces to a simple inequality $1 - \delta_{k_1}^2 < \lambda_1$ if $n_1 = 1$.

Proof The characteristic equation of $R^+(\boldsymbol{\delta})$ is

$$\prod_{i=1}^{m}(1 - \delta_{k_i}^2 - \lambda)^{n_i-1} \det\left(\mathrm{diag}(1 - \boldsymbol{\delta}_k^2 - \lambda) + \mathrm{diag}(\boldsymbol{n})\boldsymbol{\delta}_k\boldsymbol{\delta}_k^\mathsf{T}\right) = 0, \tag{3}$$

where $\boldsymbol{\delta}_k = (\delta_{k_1}, \dots, \delta_{k_m})$ and $\boldsymbol{n} = (n_1, \dots, n_m)$. As is easily seen, the trivial solutions are $\lambda = 1 - \delta_{k_i}^2$ for which $n_i > 1$. Then it is enough to find m other solutions, since each of the eigenvalues found has multiplicity $n_i - 1$. We seek other solutions, provided that they are different from any of $1 - \delta_{k_i}^2$, $i = 1, \dots, m$. Then the Eq. (3) is equivalent to

$$\sum_{i=1}^{m} \frac{(n_i\delta_{k_i}^2)}{(1 - \delta_{k_i}^2 - \lambda)} = -1.$$

Here we have used the formula $\det(A \pm \boldsymbol{b}\boldsymbol{b}^{\mathsf{T}}) = \det(A)(1 \pm \boldsymbol{b}^{\mathsf{T}} A^{-1}\boldsymbol{b})$ for a non-singular matrix $A = \mathrm{diag}(1 - \boldsymbol{\delta}_k^2 - \lambda)$. Note that the function $f(\lambda) = \sum_{i=1}^{m}(n_i\delta_{k_i}^2)/(1 - \delta_{k_i}^2 - \lambda)$ is a strictly monotone increasing function of λ on each interval $(1 - \delta_{k_i}^2, 1 - \delta_{k_{i+1}}^2)$, and diverges to negative or positive infinity on the boundaries of each interval for $i = 1, \ldots, m-1$. We now have a solution on each interval. Furthermore, we can find one more solution λ_n on the interval $(1 - \delta_{k_m}^2, \infty)$, since $f(\lambda)$ is also a monotone increasing function of λ on this interval and diverges to negative infinity on the left boundary and zero for large enough λ. We have now found the remaining m solutions such that $1 - \delta_{k_1}^2 < \lambda_{n_1} < 1 - \delta_{k_2}^2 < \lambda_{n_1+n_2} < \cdots < \lambda_{n-n_m} < 1 - \delta_{k_m}^2 < \lambda_n$. □

We note that the result Eq. (2) in Lemma 2.1 is not a direct consequence of the well known inequality for eigenvalues of A and B such that $A \le B$ nor of a more sophisticated inequality in the framework of majorization (for example, Theorem 16.F.1 and Theorem 9.G.1.c in Marshall & Olkin, 1979). In fact, the result (2) is much stronger than that obtained from such a general inequality because Lemma 2.1 is specialized for matrices like $R^+(\boldsymbol{\delta}) = \mathrm{diag}(1 - \boldsymbol{\delta}^2) + \boldsymbol{\delta}\boldsymbol{\delta}^{\mathsf{T}}$.

We now have the following theorem.

Theorem 2.1 *Assume that* $|\delta_n| \le \cdots \le |\delta_2| \le |\delta_1|$. $R^+(\boldsymbol{\delta})$ *is a correlation matrix if and only if*

$$|\delta_1| \le 1 \quad \text{or} \quad |\delta_2| < 1 < |\delta_1| \quad \text{and} \quad \sum_{i=1}^{n} \delta_i^2/(1 - \delta_i^2) \le -1.$$

Furthermore, it is a positive definite matrix if and only if

$$1 \ne |\delta_2| \le |\delta_1| \le 1 \quad \text{or} \quad |\delta_2| < 1 < |\delta_1| \quad \text{and} \quad \sum_{i=1}^{n} \delta_i^2/(1 - \delta_i^2) < -1.$$

Proof From Lemma 2.1, non-negativeness of the smallest eigenvalue λ_1 is equivalent to $1 - \delta_1^2 \ge 0$ if $\delta_1 = \delta_2$ and $\lambda_1 \ge 0$ otherwise. For the latter case, since $|\delta_2| < 1 < |\delta_1|$ and $1 - \delta_{k_1}^2 < \lambda_1 < 1 - \delta_{k_2}^2$ from Lemma 2.1, we see that $\lambda_1 \ge 0$ is equivalent to

$$f(0) = \sum_{j=1}^{m} \frac{(n_j\delta_{k_j}^2)}{(1 - \delta_{k_j}^2)} = \sum_{i=1}^{n} \frac{\delta_i^2}{(1 - \delta_i^2)} \le -1.$$

A similar argument establishes the necessary and sufficient condition for the positiveness of λ_1. The condition $|\delta_i\delta_j| \le 1 (i \ne j = 1, \ldots, n)$ is always satisfied when either $|\delta_1| \le 1$ or $|\delta_2| < 1 < |\delta_1|$ and $\sum_{i=1}^{n} \delta_i^2/(1 - \delta_i^2) \le 1$ holds true. For the latter case, it is enough to note that the following inequality holds true:

$$0 \ge 1 + \sum_{i=1}^{n} \frac{\delta_i^2}{(1 - \delta_i^2)} = \frac{(1 - \delta_1^2\delta_2^2)}{(1 - \delta_1^2)(1 - \delta_2^2)} + \sum_{i=3}^{n} \frac{\delta_i^2}{(1 - \delta_i^2)}.$$

□

It is interesting to note that all $|\delta_i|$'s are not necessarily less than or equal to 1. The largest δ_1 can be greater than 1 in absolute value. However, unless all other parameters are very small in absolute value, $|\delta_1|$ cannot be far away from 1 since it must satisfy the side condition $\sum_{i=1}^n \delta_i^2/(1-\delta_i^2) < -1$.

2.2 Negative multiplicative correlation

For negative multiplicative correlation matrix $R^-(\boldsymbol{\delta})$ we need the following lemma, which looks very similar to Lemma 2.1, but is not the same.

Lemma 2.2 *The following inequality holds true for the eigenvalues* $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ *of* $R^-(\boldsymbol{\delta})$.

$$\begin{aligned}\lambda_1 &< 1+\delta_{k_1}^2 = \lambda_2 = \cdots = \lambda_{n_1} < \lambda_{n_1+1} < 1+\delta_{k_2}^2\\ &\qquad = \lambda_{n_1+2} = \cdots = \lambda_{n_1+n_2} < \lambda_{n_1+n_2+1} < \cdots < \lambda_{n-n_m+1} < 1+\delta_{k_m}^2\\ &\qquad = \lambda_{n-n_m+2} = \cdots = \lambda_n,\end{aligned}$$

where $\delta_{k_1}^2 < \delta_{k_2}^2 < \cdots < \delta_{k_m}^2$ *are* m *distinct values in* $\delta_1^2, \ldots, \delta_n^2$ *and* $n_1, \ldots, n_m$ $\left(\sum_{i=1}^m n_i = n\right)$ *are multiplicities of* $\delta_{k_1}^2, \ldots, \delta_{k_m}^2$ *respectively.*

Proof Since the characteristic equation of $R^-(\boldsymbol{\delta})$ is

$$\prod_{i=1}^m (1+\delta_{k_i}^2-\lambda)^{n_i-1} \det\left(\operatorname{diag}(1+\boldsymbol{\delta}_k^2-\lambda) + \operatorname{diag}(\boldsymbol{n})\boldsymbol{\delta}_k\boldsymbol{\delta}_k^\mathsf{T}\right) = 0,$$

a similar argument follows as in the proof of Lemma 2.1. To find m non-trivial solutions, it is enough to note that the equation above is equivalent to $\sum_{i=1}^m (n_i\delta_{k_i}^2)/(1+\delta_{k_i}^2-\lambda) = 1$ provided that λ is equal to none of the $1+\delta_{k_i}^2$'s. We then find a solution on each interval $(1+\delta_{k_i}^2, 1+\delta_{k_i+1}^2)$ for $i = 1, \ldots, m-1$, by noting that $g(\lambda) = \sum_{i=1}^m (n_i\delta_{k_i}^2)/(1+\delta_{k_i}^2-\lambda)$ has the same properties as those of $f(\lambda)$ in Lemma 2.1. The remaining solution on $(-\infty, 1+\delta_{k_1}^2)$ can be found since $g(\lambda)$ is strictly monotone increasing, converges to zero as λ tends to negative infinity and $g(1+\delta_{k_1}^2) = 0$. □

Theorem 2.2 $R^-(\boldsymbol{\delta})$ *is a correlation matrix if and only if* $\sum_{i=1}^n \delta_i^2/(1+\delta_i^2) \le 1$. *It is a positive definite matrix if and only if* $\sum_{i=1}^n \delta_i^2/(1+\delta_i^2) < 1$.

Proof From the proof of Lemma 2.2, non-negativeness of the minimum eigenvalue λ_1 is equivalent to

$$g(0) = \sum_{j=1}^m \frac{(n_j\delta_{k_j}^2)}{(1+\delta_{k_j}^2)} = \sum_{i=1}^n \frac{\delta_i^2}{(1+\delta_i^2)} \le 1.$$

And the condition $|\delta_1\delta_2| \le 1$ follows from

$$1 \ge \sum_{i=1}^n \frac{\delta_i^2}{(1+\delta_i^2)} \ge \sum_{i=1}^2 \frac{\delta_i^2}{(1+\delta_i^2)}.$$

□

Compared with the condition in Theorem 2.1 for positive multiplicative correlation, the condition in Theorem 2.2 for negative multiplicative correlation looks simpler. However, it seems more restrictive, since, although there is no explicit restriction to the individual values of $|\delta_i|$ (such as most of them should be less than 1), the total contribution of the terms $\delta_i^2/(1+\delta_i^2)$ should not exceed 1.

Example 1 (Equi-correlation) From Theorem 2.1 and Theorem 2.2, we can easily see that the choice of parameter c for the equi-correlation model Q_{ij} given in Eq. (1) is limited to the range $-1/(n-1) \le c \le 1$.

The covariance matrix corresponding to a multiplicative correlation is apparently of the form $\mathrm{diag}(\boldsymbol{b}) \pm \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}$ and we can discuss multiplicative structure either through correlation or covariance. A multiplicative covariance can be derived from a multiplicative correlation matrix by giving the variances, $\mathrm{var}(X_i) = \sigma_i^2, i = 1, \ldots, n$. We hereafter write such a covariance matrix as $\Sigma^+(\boldsymbol{a}, \boldsymbol{b}) = \mathrm{diag}(\boldsymbol{b}) + \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}$ in case of $R^+(\boldsymbol{\delta})$ and call it a *positive multiplicative covariance matrix*. Similarly, we define a covariance matrix $\Sigma^-(\boldsymbol{a}, \boldsymbol{b}) = \mathrm{diag}(\boldsymbol{b}) - \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}$ in case of $R^-(\boldsymbol{\delta})$ and call it a *negative multiplicative covariance matrix*. Here $\boldsymbol{a} = (a_1, \ldots, a_n)$ is a vector such that $a_i = \sigma_i \delta_i, i = 1, \ldots, n$, and $\boldsymbol{b} = (b_1, \ldots, b_n)$ is a vector such that $b_i = \sigma_i^2(1-\delta_i^2), i = 1, \ldots, n$, for $R^+(\boldsymbol{\delta})$ and $b_i = \sigma_i^2(1+\delta_i^2)$ for $R^-(\boldsymbol{\delta})$. Note that the converse is not always true. Multiplicative covariance matrix does not necessarily imply multiplicative correlation matrix, since we allow any σ_i^2 to be 0.

Theorem 2.3 *Assume that* $b_1 \le b_2 \le \cdots \le b_n$. *The matrix* $\Sigma^+(\boldsymbol{a}, \boldsymbol{b})$ *is a covariance matrix if and only if*

$$0 \le b_1 \qquad \text{or} \qquad b_1 < 0 < b_2 \quad \text{and} \quad 1 + \sum_{i=1}^{n} \frac{a_i^2}{b_i} \le 0.$$

It is a positive definite matrix if and only if

$$0 \le b_1 \le b_2 \ne 0 \qquad \text{or} \qquad b_1 < 0 < b_2 \quad \text{and} \quad 1 + \sum_{i=1}^{n} \frac{a_i^2}{b_i} < 0.$$

The matrix $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$ *is a covariance matrix if and only if* $0 < b_1$ *and* $\sum_{i=1}^n a_i^2/b_i \le 1$. *It is a positive definite matrix if* $\sum_{i=1}^n a_i^2/b_i < 1$.

Theorem 2.3 is a direct consequence of Theorems 2.1 and 2.2. However, the distribution of the eigenvalues is also of interest. Provided that $\boldsymbol{a}$ has no zero elements, the following inequalities are rather trivial in view of Lemmas 2.1 and 2.2.

For the eigenvalues of $\Sigma^+(\boldsymbol{a}, \boldsymbol{b})$

$$\begin{aligned} b_{k_1} &= \lambda_1 = \cdots = \lambda_{n_1-1} < \lambda_{n_1} < b_{k_2} \\ &= \lambda_{n_1+1} = \cdots = \lambda_{n_1+n_2-1} < \lambda_{n_1+n_2} < \cdots < \lambda_{n-n_m} < b_{k_m} \\ &= \lambda_{n-n_m+1} = \cdots = \lambda_{n-1} < \lambda_n, \end{aligned}$$

and for the eigenvalues of $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$

$$\begin{aligned} \lambda_1 < b_{k_1} &= \lambda_2 = \cdots = \lambda_{n_1} < \lambda_{n_1+1} < b_{k_2} \\ &= \lambda_{n_1+2} = \cdots = \lambda_{n_1+n_2} < \lambda_{n_1+n_2+1} < \cdots < \lambda_{n-n_m+1} < b_{k_m} \\ &= \lambda_{n-n_m+2} = \cdots = \lambda_n. \end{aligned}$$

Here $b_{k_1} < b_{k_2} < \cdots < b_{k_m}$ are m distinct elements of $\boldsymbol{b}$ and $n_1, \ldots, n_m$ $\left(\sum_{i=1}^{m} n_i = n\right)$ are the multiplicities of $b_{k_1}, \ldots, b_{k_m}$, as same as in Lemma 2.1 or Lemma 2.2. It is interesting to note that these inequalities depend only on the values of $\boldsymbol{b}$. This can also be seen by noting that the characteristic equation here is $\prod_{i=1}^{m}(b_{k_i} - \lambda)^{n_i - 1} \det\left\{\mathrm{diag}(\boldsymbol{b}_k - \lambda) \pm \boldsymbol{\gamma}\boldsymbol{\gamma}^{\mathsf{T}}\right\} = 0$, where $\boldsymbol{b}_k = (b_{k_1}, \ldots, b_{k_m})$ and $\boldsymbol{\gamma} = (\gamma_{k_1}, \ldots, \gamma_{k_m})$ with $\gamma_{k_i}^2 = \sum_{j:b_j = b_{k_i}} a_j^2$.

If there is a zero element a_i in $\boldsymbol{a}$, the inequalities above should be modified by noting that the corresponding b_i becomes an eigenvalue. Ronning (1982) proved a similar inequality in the cases of the multinomial, Dirichlet or multivariate hypergeometric distributions. Watson (1996) independently noted that such an inequality holds true for the multinomial distribution. However, our result above is not restricted to such special multivariate distributions, but applies, in general, to any multivariate distributions with multiplicative covariances.

Example 2 (Equi-covariance) From Theorem 2.3, we can easily see that the covariance r for the equi-covariance matrix should satisfy

$$0 \le r \le \sigma_1^2, \quad \sigma_1^2 < r < \sigma_2^2 \text{ and } \sum_{i=1}^{n} \frac{r}{(\sigma_i^2 - r)} \le -1, \text{ or } r < 0 \text{ and}$$

$$\sum_{i=1}^{n} \frac{r}{(\sigma_i^2 - r)} \ge -1,$$

where σ_1^2 and σ_2^2 are the minimum and the second minimum of the variances of X_i, respectively.

3 Implications of multiplicative correlations or covariances

In this section, we investigate the implications of multiplicative correlation or covariance matrices. In Sect. 3.1, we first give a factorization theorem. The meaning of multiplicative correlation becomes clearer through the factorization of variables, at least for the case of positive multiplicative correlations. In Sect. 3.2 we give one reason why negative multiplicative correlation matrices arise so frequently, although other explanations are also no doubt possible.

3.1 Characterizations

In view of the reduction method mentioned in Section 1, we might expect that the X_i can be represented as a common variable plus individual variables if the correlation is multiplicative. For the multivariate normal distribution, such a factorization is almost trivial, and Curnow and Dunnett (1962) or Gupta (1963) showed that a simple calculation of the distribution is possible when the correlation is positive multiplicative. Six (1981) extended their results to the case of negative multiplicative correlation. The following theorem gives us a general factorization theorem for positive and negative multiplicative correlations.

Theorem 3.1 *A random vector* $\boldsymbol{X} = (X_1, \dots, X_n)$ *with zero means has a non-singular multiplicative covariance* $\mathrm{var}(\boldsymbol{X}) = \mathrm{diag}(\boldsymbol{b}) \pm \boldsymbol{a}\boldsymbol{a}^\top$ *with* $\boldsymbol{b} > \boldsymbol{0}$ *if and only if each element of* $\boldsymbol{X}$ *is written as*

$$X_i = \gamma a_i Z_0 + \sqrt{b_i} Z_i, \quad i = 1, \dots, n, \tag{4}$$

where $Z_0, Z_1, \dots, Z_n$ *are random variables with zero means and unit variances, in which* $Z_1, \dots, Z_n$ *are uncorrelated each other but correlated with* Z_0 *as*

$$corr(Z_0, Z_i) = \frac{ca_i}{\sqrt{b_i}}, \quad i = 1, \dots, n,$$

where $|c| \leq 1/\kappa$ *for* $\Sigma^+(\boldsymbol{a}, \boldsymbol{b})$ *and* $1 \leq |c| \leq 1/\kappa$ *for* $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$ *with* $\kappa^2 = \sum_{i=1}^n (a_i^2/b_i)$. *The constants* γ *and* c *satisfy the equation,*

$$\gamma^2 + 2\gamma c = \begin{cases} 1 & for \quad \Sigma^+(\boldsymbol{a}, \boldsymbol{b}) \\ -1 & for \quad \Sigma^-(\boldsymbol{a}, \boldsymbol{b}) . \end{cases}$$

Proof If $\boldsymbol{X}$ is represented as in Eq. (4), then a direct calculation yields the multiplicative covariance $\mathrm{var}(\boldsymbol{X}) = \mathrm{diag}(\boldsymbol{b}) \pm \boldsymbol{a}\boldsymbol{a}^\top$. On the other hand, if $\boldsymbol{X}$ has the desired covariance matrix, define a random variable

$$Z_0 = \left\{ \frac{(1 - c^2\kappa^2)}{(1 + \sigma\kappa^2)} \right\}^{1/2} X_0 \pm (c^2 + \sigma)^{1/2} \boldsymbol{a}^\top \Sigma^{-1} \boldsymbol{X}$$

by introducing a new random variable X_0 with mean zero and unit variance but independent of any X_i, $i = 1, \dots, n$. Here σ is 1 for $\Sigma^+(\boldsymbol{a}, \boldsymbol{b})$ or is -1 for $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$, and the sign $\pm$ shows alternative definitions. It can be easily seen that the Z_0 and $Z_i = [X_i - \{\pm(c^2 + \sigma)^{1/2} - c\} a_i Z_0]/\sqrt{b_i}$ satisfy the desired properties from the non-negative definiteness of the covariance of $\boldsymbol{Z}$ and Theorem 2.3. □

The parameter γ in Theorem 3.1 can be written as $\gamma = -c \pm \sqrt{c^2 + 1}$ for $\Sigma^+(\boldsymbol{a}, \boldsymbol{b})$ and $\gamma = -c \pm \sqrt{c^2 - 1}$ for $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$. Hence the factorization is not unique even when c is fixed which determines a global level of correlations of Z_i's with Z_0. It is worthy of noting that Z_i's are always assumed to be uncorrelated with Z_0 in factor analysis.

Corollary 3.1 *A random vector* $\boldsymbol{X} = (X_1, \dots, X_n)$ *with zero means has a non-singular multiplicative covariance* $\mathrm{var}(\boldsymbol{X}) = \mathrm{diag}(\boldsymbol{b}) \pm \boldsymbol{a}\boldsymbol{a}^\top$ *with* $\boldsymbol{b} > \boldsymbol{0}$ *if and only if each element of* $\boldsymbol{X}$ *is written as*

$$X_i = a_i Z_0 + \sqrt{b_i} Z_i, \quad i = 1, \dots, n,$$

where $Z_0, Z_1, \dots, Z_n$ *are random variables with zero means and unit variances, in which* $Z_1, \dots, Z_n$ *are uncorrelated each other but it can be correlated with* Z_0 *as*

$$\mathrm{corr}(Z_0, Z_i) = \begin{cases} 0 & for \quad \Sigma^+(\boldsymbol{a}, \boldsymbol{b}) \\ -a_i/\sqrt{b_i} & for \quad \Sigma^-(\boldsymbol{a}, \boldsymbol{b}) \end{cases}, \quad i = 1, \dots, n.$$

The factorization in Corollary 3.1 for positive multiplicative correlation is known as the *fundamental theorem of factor analysis* (see, for example, Steiger, 1979, p. 158). Also, it is clear that the reduction method mentioned in Sect. 1 is based on the factorization with $a_i = a$ for $i = 1, \ldots, n$, and $\mathrm{corr}(Z_0, Z_i) = 0$. However, Corollary 3.1 suggests a more general way of defining new multivariate distributions by choosing different a_i's or allowing the Z_i's to be correlated with Z_0. It is interesting to note that the variables $Z_i (i = 0, 1, \ldots, n)$ should have a specific correlation with Z_0 when $X_i (i = 0, 1, \ldots, n)$ have a negative multiplicative correlation. It is also interesting that the common variable Z_0 is uncorrelated with X_i's, as is easily shown. The following examples suggest how the factorization in Corollary 3.1 can be applied in practice.

Example 3 Variance reduction methods such as *antithetic variates method* (for example, Hammersley and Handscomb, 1964, p. 60) in Monte Carlo simulation typically make use of negatively correlated random numbers. We can generate n negatively correlated random numbers using the formula

$$X_i = \mu + a_i Z_0 + \sqrt{b_i} Z_i, \quad i = 1, \ldots, n,$$

where $Z_1, \ldots, Z_{n+1}$ are independent random variables with zero mean and unit variances and $Z_0 = -\sum_{i=1}^{n} a_i Z_i / \sqrt{b_i} + Z_{n+1}$ for positive a_i's. Then, from Corollary 3.1 it follows that $\boldsymbol{X} = (X_1, \ldots, X_n)$ has a negative multiplicative covariance matrix given by $\mathrm{var}(\boldsymbol{X}) = \mathrm{diag}(\boldsymbol{b}) - \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}$, where all correlation coefficients are negative.

Example 4 Assume that there are two groups of assets; the returns $\boldsymbol{X} = (X_1, \ldots, X_n)$ in a group are uncorrelated with variances $\sigma_1^2, \ldots, \sigma_n^2$, but the returns $\boldsymbol{Y} = (Y_1, \ldots, Y_n)$ in another group have the same variances but a positive multiplicative covariance $\Sigma^+(\boldsymbol{a}, \boldsymbol{\sigma}^2 - \boldsymbol{a}^2)$ where $\boldsymbol{\sigma}^2 - \boldsymbol{a}^2 = (\sigma_1^2 - a_1^2, \ldots, \sigma_n^2 - a_n^2) > \boldsymbol{0}$. It is well known that the minimal risk portfolio $\boldsymbol{w}^{\mathsf{T}}\boldsymbol{Z}$ for the return $\boldsymbol{Z} = (Z_1, \ldots, Z_n)$ with the weights $\boldsymbol{w}$ such that $\boldsymbol{w}^{\mathsf{T}}\boldsymbol{1} = 1$ has the variance $1/\boldsymbol{1}^{\mathsf{T}}\Omega^{-1}\boldsymbol{1}$ where $\mathrm{var}(\boldsymbol{Z}) = \Omega$. Therefore, for example, when $\sum_{i \neq j} a_i a_j / \{(\sigma_i^2 - a_i^2)(\sigma_j^2 - a_j^2)\} < 0$, the risk $1/\boldsymbol{1}^{\mathsf{T}}\left\{\mathrm{diag}(\boldsymbol{\sigma}^2 - \boldsymbol{a}^2) + \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}\right\}^{-1}\boldsymbol{1}$ of the minimal risk portfolio for $\boldsymbol{Y}$ is less than the risk $1/\boldsymbol{1}^{\mathsf{T}}\mathrm{diag}(\boldsymbol{\sigma}^2)^{-1}\boldsymbol{1}$ of the minimal risk portfolio for $\boldsymbol{X}$. Even when $\boldsymbol{Y}$ has a negative multiplicative covariance $\Sigma^-(\boldsymbol{a}, \boldsymbol{\sigma}^2 + \boldsymbol{a}^2)$ for $\boldsymbol{a} < \boldsymbol{1} < \boldsymbol{\sigma}$, the minimal risk $1/\boldsymbol{1}^{\mathsf{T}}\left\{\mathrm{diag}(\boldsymbol{\sigma}^2 + \boldsymbol{a}^2) - \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}\right\}^{-1}\boldsymbol{1}$ for $\boldsymbol{Y}$ is less than the minimal risk for $\boldsymbol{X}$ when $1/2 < \sum_{i=1}^{n} a_i^2/(\sigma_i^2 + a_i^2) < 1$.

The following corollary is a direct consequence of Corollary 3.1 for multiplicative correlations.

Corollary 3.2 *A random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$ with zero means and variances $\sigma_i^2 = \mathrm{var}(X_i), i = 1, \ldots, n$, has a positive definite multiplicative correlation matrix if and only if each element of $\boldsymbol{X}$ is written as*

$$X_i/\sigma_i = \begin{cases} \delta_i Z_0 + (1 - \delta_i^2)^{1/2} Z_i & \textit{for} \quad R^+(\boldsymbol{\delta}) \\ \delta_i Z_0 + (1 + \delta_i^2)^{1/2} Z_i & \textit{for} \quad R^-(\boldsymbol{\delta}) \end{cases}, \quad i = 1, \ldots, n,$$

where $Z_0, Z_1, \ldots, Z_n$ are random variables with zero means and unit variances, in which $Z_1, \ldots, Z_n$ are uncorrelated each other, but can be correlated with Z_0 as

$$\operatorname{corr}(Z_0, Z_i) = \begin{cases} 0 & \textit{for} \quad R^+(\boldsymbol{\delta}) \\ -\delta_i/\left(1+\delta_i^2\right)^{1/2} & \textit{for} \quad R^-(\boldsymbol{\delta}) \end{cases}, \quad i = 1, \ldots, n.$$

3.2 A further characterization of negative multiplicative covariance

The following theorem explains why negative multiplicative covariance matrix arises frequently.

Theorem 3.2 *Assume that an n-dimensional random vector $\boldsymbol{X}$ has covariance matrix $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$. Then, $\boldsymbol{b} = (\sum_{i=1}^n a_i)\boldsymbol{a}$ if and only if $\sum_{i=1}^n X_i$ is almost surely constant.*

Proof The fact that $\sum_{i=1}^n X_i$ is almost surely constant is equivalent to

$$\operatorname{var}\left(\sum_i X_i\right) = \mathbf{1}^\mathsf{T}\left\{\operatorname{diag}(\boldsymbol{b}) - \boldsymbol{a}\boldsymbol{a}^\mathsf{T}\right\}\mathbf{1} = \sum_i b_i - \left(\sum_i a_i\right)^2 = 0.$$

From the Schwarz's inequality we have

$$\sum_i b_i = \left(\sum_i a_i\right)^2 = \left(\sum_i \frac{a_i}{\sqrt{b_i}}\sqrt{b_i}\right)^2 \leq \left(\sum_i b_i\right)\left(\sum_i \frac{a_i^2}{b_i}\right)$$

and it is clear that b_i is proportional to a_i since $\sum_{i=1}^n a_i^2/b_i \leq 1$. □

This theorem says that the negative multiplicative covariance matrix takes the form of

$$\operatorname{var}(\boldsymbol{X}) = \left(\sum_{i=1}^n a_i\right)\operatorname{diag}(\boldsymbol{a}) - \boldsymbol{a}\boldsymbol{a}^\mathsf{T}$$

if and only if there is a sum constraint $\sum_{i=1}^n X_i = \text{const}$ a.s. Multivariate Pólya-Eggenberger distribution is an example of family of distributions which have such a special type of negative multiplicative covariance matrix.

Example 5 Multivariate Pólya-Eggenberger distribution is a class of distributions whose joint probability functions take the form of

$$p(x_1, \ldots, x_n) = \binom{t}{x_1, \ldots, x_n}\frac{\left\{\prod_{i=1}^n \alpha_i^{[x_i,c]}\right\}}{\alpha^{[t,c]}},$$

where x_i and $\alpha_i, i = 1, \ldots, n$, are non-negative integers, c is an integer, $t = \sum_{i=1}^n x_i, \alpha = \sum_{i=1}^n \alpha_i$, and $\alpha^{[x,c]} = \alpha(\alpha+c)\cdots\{\alpha+(x-1)c\}$ with $\alpha^{[0,c]} = 1$ (see Johnson et al., 1997, p. 201). Multivariate Pólya-Eggenberger distribution includes multinomial distribution ($c = 0$), multivariate hypergeometric distribution ($c = 1$), or multivariate negative hypergeometric distribution ($c = -1$).

Distributions in the class have negative multiplicative covariance matrices since there is a constraint like $t = \sum_{i=1}^{n} X_i$. In fact, by a direct calculation, we have

$$E(\boldsymbol{X}) = t\frac{\boldsymbol{\alpha}}{\alpha} \quad \text{and} \quad \mathrm{var}(\boldsymbol{X}) = \frac{t(\alpha + tc)}{\alpha^2(\alpha + c)}\left\{\alpha \mathrm{diag}(\boldsymbol{\alpha}) - \boldsymbol{\alpha}\boldsymbol{\alpha}^{\mathsf{T}}\right\},$$

with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$. We can expect that the Dirichlet distribution has a negative multiplicative covariance as in Theorem 3.2, since it is a limit of multivariate Pólya-Eggenberger distribution (see Johnson et al., 1997, p. 211). In fact, the covariance matrix of Dirichlet distribution is written as $\left(\nu \mathrm{diag}(\boldsymbol{\nu}) - \boldsymbol{\nu}\boldsymbol{\nu}^{\mathsf{T}}\right) / \left\{\nu^2(\nu + 1)\right\}$ where $\nu = \sum_{i=1}^{n} \nu_i$.

4 Invariance of multiplicative covariance

We can interpret the result in Theorem 3.2 in the following manner. The conditional distribution of $\boldsymbol{X} = (X_1, \ldots, X_n)$ given $\sum_{i=1}^{n} X_i$ has a special type of negative multiplicative covariance, provided that $\boldsymbol{X}$ has a negative multiplicative covariance. The following Theorem 4.1 is a converse of Theorem 3.2, but it holds true more generally.

Theorem 4.1 *Let $(\boldsymbol{X}, T)$ be an $(n + 1)$-dimensional random vector. Assume that the conditional variance of $\boldsymbol{X}$ given T is*

$$\mathrm{var}(\boldsymbol{X}|T = t) = \sigma(t)\left(\mathrm{diag}(\boldsymbol{b}) \pm \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}\right)$$

for a function $\sigma(t) > 0$. If the conditional expectation takes the form of

$$E(\boldsymbol{X}|T = t) = \mu(t)\boldsymbol{a} + \boldsymbol{c}$$

for an n-dimensional constant vector $\boldsymbol{c}$, then the unconditional covariance is

$$\mathrm{var}(\boldsymbol{X}) = E\left(\sigma(T)\right) \mathrm{diag}(\boldsymbol{b}) + \left\{\mathrm{var}(\mu(T)) \pm E\left(\sigma(T)\right)\right\} \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}},$$

which is again multiplicative.

Proof The result follows by direct calculations;

$$E(X_i|T = t) = a_i\mu(t) + c_i,\ E(X_i^2|T = t) = (b_i \pm a_i^2)\sigma(t) + (a_i\mu(t) + c_i)^2$$

and

$$E(X_iX_j|T = t) = \pm a_ia_j\sigma(t) + (a_i\mu(t) + c_i)\left(a_j\mu(t) + c_j\right).$$

□

It is worth noting that the unconditional covariance matrix can be positive or negative multiplicative irrespective of whether the conditional covariance matrix is positive or negative multiplicative.

Example 6 (Homogeneous Distribution) It is known that $\boldsymbol{X}$ is distributed as a multivariate homogeneous distribution if and only if the conditional distribution of $\boldsymbol{X}$ given the sum $\sum_{i=1}^{n} X_i$ is multinomial (see, Johnson et al., 1997, p. 20). Since the conditional expectation and covariance matrix are written as $E(\boldsymbol{X}|T = t) = t\boldsymbol{p}$ and $\mathrm{var}(\boldsymbol{X}|T = t) = t\left(\mathrm{diag}(\boldsymbol{p}) - \boldsymbol{p}\boldsymbol{p}^{\mathsf{T}}\right)$, the conditions in Theorem 4.1 are satisfied by taking $\sigma(t) = \mu(t) = t$ and $\boldsymbol{c} = \boldsymbol{0}$. Therefore, we see that the homogeneous distribution always has a multiplicative covariance matrix given by

$$\mathrm{var}(\boldsymbol{X}) = E(T)\mathrm{diag}(\boldsymbol{p}) + (\mathrm{var}(T) - E(T))\,\boldsymbol{p}\boldsymbol{p}^{\mathsf{T}}.$$

The sign of $\mathrm{var}(T) - E(T)$ depends on the distribution of T. For example, it is always negative multiplicative if the distribution of T is binomial, and positive multiplicative if the distribution is negative binomial. Although it can be seen from the fact that the resulting distribution of $\boldsymbol{X}$ is multinomial or negative multinomial, respectively, it can also be shown by a direct calculation of $\mathrm{var}(T) - E(T)$. It is trivial, but interesting, to note that $\boldsymbol{X}$ is a vector of orthogonal variables if T is Poisson distributed because $\mathrm{var}(T) = \mathrm{E}(T)$. For the multivariate Pólya-Eggenberger distributions in Example 5, t is a parameter and can be replaced by a non-negative integer valued random variable T. Then it is clear from Theorem 4.1 that the covariance matrix of $\boldsymbol{X}$ is again multiplicative for any choice of T.

Example 7 (Random Scaling) It is clear from Theorem 4.1 that the randomly scaled $\boldsymbol{X} = T\boldsymbol{Y}$ has a multiplicative covariance as far as

$$E(\boldsymbol{Y}) = k\boldsymbol{a} \quad \text{for a constant } k, \qquad \mathrm{var}(\boldsymbol{Y}) = \mathrm{diag}(\boldsymbol{b}) \pm \boldsymbol{a}\boldsymbol{a}^{\mathsf{T}}$$

and T is independent of $\boldsymbol{Y}$. Several multivariate continuous distributions can be derived by using such a random scaling. For example, the multivariate Liouville distribution or multivariate beta distribution of the second kind (or inverted Dirichlet) are derived from the Dirichlet distribution by taking respectively Liouville distributed T or second kind beta distributed T (see, Kotz et al., 2000, p. 491, 530; Gupta & Richards, 2001). In view of Theorem 4.1, multiplicative covariance structures are preserved through such a derivation. The random scaling for multivariate continuous distributions plays a similar role to that of the reduction method for multivariate discrete distributions. Unfortunately the converse of Theorem 4.1 is not so simple. It heavily depends on the shape of the distribution and the problem is left for future investigation.

To see another invariance, we need the following proposition, where $\boldsymbol{a}/\boldsymbol{b} = (a_1/b_1, \ldots, a_n/b_n)$.

Proposition 4.1 *Assume that $\Sigma^+(\boldsymbol{a}, \boldsymbol{b})$ and $\Sigma^-(\boldsymbol{a}, \boldsymbol{b})$ are positive definite. Then*

$$\Sigma^+(\boldsymbol{a}, \boldsymbol{b})^{-1} = \Sigma^-\,(c\boldsymbol{a}/\boldsymbol{b}, 1/\boldsymbol{b}) \quad \textit{for} \quad c^2 = 1/(1 + \boldsymbol{a}^{\mathsf{T}}\mathrm{diag}(\boldsymbol{b})^{-1}\boldsymbol{a})$$

and

$$\Sigma^-(\boldsymbol{a}, \boldsymbol{b})^{-1} = \Sigma^+\,(c\boldsymbol{a}/\boldsymbol{b}, 1/\boldsymbol{b}) \quad \textit{for} \quad c^2 = 1/(1 - \boldsymbol{a}^{\mathsf{T}}\mathrm{diag}(\boldsymbol{b})^{-1}\boldsymbol{a}),$$

provided that $\boldsymbol{b} > \boldsymbol{0}$.

We now show another invariance property of multiplicative covariances or correlations.

Theorem 4.2 *Assume that* $\boldsymbol{X} = (X_1, \ldots, X_n)$ *has a multiplicative covariance* $\Sigma^{\pm}(\boldsymbol{a}, \boldsymbol{b})$. *Partition* $\boldsymbol{X}$ *as* $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)$ *with* m*-dimensional vector* $\boldsymbol{X}_1$ *and* $(n-m)$*-dimensional vector* $\boldsymbol{X}_2$, *and its parameter vectors as* $\boldsymbol{a} = (\boldsymbol{a}_1, \boldsymbol{a}_2)$ *and* $\boldsymbol{b} = (\boldsymbol{b}_1, \boldsymbol{b}_2)$, *simultaneously. If* $\mathrm{var}(\boldsymbol{X}_2)$ *is non-singular and all elements of* $\boldsymbol{b}_2$ *are positive, then the partial covariance of* $\boldsymbol{X}_1$ *given* $\boldsymbol{X}_2$ *is also multiplicative and is* $\Sigma^{\pm}(\boldsymbol{a}_1/c, \boldsymbol{b}_1)$ *with* $c^2 = 1 \pm \boldsymbol{a}_2^{\top}\mathrm{diag}(\boldsymbol{b}_2)^{-1}\boldsymbol{a}_2$.

Proof Partition $\Sigma^{\pm}(\boldsymbol{a}, \boldsymbol{b})$ into

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix} = \begin{pmatrix} \mathrm{diag}(\boldsymbol{b}_1) \pm \boldsymbol{a}_1\boldsymbol{a}_1^{\top} & \pm\boldsymbol{a}_1\boldsymbol{a}_2^{\top} \\ \pm\boldsymbol{a}_2\boldsymbol{a}_1^{\top} & \mathrm{diag}(\boldsymbol{b}_2) \pm \boldsymbol{a}_2\boldsymbol{a}_2^{\top} \end{pmatrix}.$$

From Proposition 4.1, the partial covariance matrix of $\boldsymbol{X}_1$ given $\boldsymbol{X}_2$ is written as

$$\begin{aligned} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} &= \mathrm{diag}(\boldsymbol{b}_1) \pm \boldsymbol{a}_1\boldsymbol{a}_1^{\top} - \boldsymbol{a}_1\boldsymbol{a}_2^{\top}\left(\mathrm{diag}(\boldsymbol{b}_2) \pm \boldsymbol{a}_2\boldsymbol{a}_2^{\top}\right)^{-1}\boldsymbol{a}_2\boldsymbol{a}_1^{\top} \\ &= \mathrm{diag}(\boldsymbol{b}_1) \pm \boldsymbol{a}_1\boldsymbol{a}_1^{\top}/\{1 \pm \boldsymbol{a}_2^{\top}\mathrm{diag}(\boldsymbol{b}_2)^{-1}\boldsymbol{a}_2\}. \end{aligned}$$

□

For multiplicative correlations, the following corollary holds true, which is a direct consequence of Theorem 4.2.

Corollary 4.1 *If* $\boldsymbol{X}$ *has a multiplicative correlation* $R^{\pm}(\boldsymbol{\delta})$, *then the partial correlation of* $\boldsymbol{X}_1$ *given* $\boldsymbol{X}_2$ *is also multiplicative and* $R^{\pm}(\tilde{\boldsymbol{\delta}})$ *with*

$$\tilde{\delta}_i = \frac{\delta_i}{\left\{1 \pm c(1 \mp \delta_i^2)\right\}^{1/2}}, \quad i = 1, \ldots, m,$$

where $c = \sum_{j=m+1}^{n} \delta_j^2/(1 \mp \delta_j^2)$.

We see that the partial covariance is proportional to the original covariance but the partial correlation is not, although the multiplicative property is retained for both cases. An important implication of Theorem 4.2 or Corollary 4.1 is that zero correlation always implies zero partial correlation. This is due to the multiplicative parameterization of the correlation or covariance matrix, but it is not always true without such a parameterization. A simple example is for the case of $n = 3$. The partial covariance between X_1 and X_2 is $\sigma_{12} - \sigma_{13}\sigma_{23}/\sigma_{33}$ in general, which is not necessarily zero even if the original covariance $\sigma_{12} = 0$. However, σ_{13} or σ_{23} becomes zero if $\sigma_{12} = 0$ under the multiplicative parameterization of the covariance, so that the zero covariance implies zero partial covariance.

5 Concluding remarks

As a concluding remark, we note that many known multivariate distributions have multiplicative correlation or covariance structure. In Johnson et al. (1997), eight families of discrete distributions were introduced and these are given as in Table 1. This table shows whether each of these families has multiplicative correlation structure or not together with reasons. As has been discussed before, multinomial, negative multinomial, Poisson, hypergeometric and Pólya-Eggenberger distribution have multiplicative correlation structure. Although power series distributions or multivariate distributions of order s have no multiplicative correlation in general, subfamilies such as the logarithmic series distributions (Johnson et al., 1997, p. 157), multivariate negative multinomial of order s (p. 255) or multivariate logarithmic series distributions of order s (p. 260) have multiplicative correlation structure. However, at this stage, we do not know the exact reason why such subfamilies have multiplicative correlation structure. Apparently, Ewens distributions have no multiplicative correlation structure.

In terms of continuous distributions, seven families of explicit continuous distributions were introduced in Kotz et al. (2000). Table 2 shows whether each family has multiplicative correlation structure or not in the same way as Table 1. For multivariate normal distributions we define a subfamily where the correlation matrix is multiplicative and call this the multiplicatively correlated normal. Although multivariate exponential, multivariate gamma, multivariate logistic or multivariate Pareto distributions have no multiplicative correlations in general, subfamilies like Moran and Downton's multivariate exponential distributions

Table 1 Discrete multivariate distributions in Johnson et al. (1997)

Family	Subfamily	Positive or negative	Reason
35 Multinomial		Negative	Example 5
36 Negative multinomial		Positive	Example 6
37 Poisson		Positive	Reduction method
38 Power series	Logarithmic series	Positive	?
39 Hypergeometric		Negative	Example 5
40 Pólya-Eggenberger		Negative	Example 5
41 Ewens	–	–	–
42 Distributions of	Negative binomial of order s	Positive	?
order s	Logarithmic distr. of order s	Negative	?

Table 2 Continuous multivariate distributions in Kotz et al. (2000)

Family	Subfamily	Positive or negative	Reason
45 Normal	Multiplicatively correlated normal	Both	
47 Exponential	Moran and Downton's	Positive	Equi-correlation
48 Gamma	Cheriyan and Ramabhadran's	Positive	Reduction Method
49 Dirichlet		Positive	Example 5
49 Inverted Dirichlet		Negative	Example 7
50 Liouvill		Both	Example 7
51 Logistic	Gumbel-Malik-Abraham	Positive	Equi-correlation
	Farlie-Gumbel-Morgenstern	Negative	Equi-correlation
52 Pareto	The first kind	Positive	Equi-correlation

(Kotz et al., 2000, p. 400), Cheriyan and Ramabhadran's multivariate gamma distributions (p. 454), Gumbel-Malik-Abraham's (p. 552) and Farlie-Gumbel-Morgenstern's (p. 561) multivariate logistic distributions, or multivariate Pareto distributions of the first kind (p. 599) have equi-covariance or equi-correlation structure which is multiplicative. Example 5 or Example 7 shows that Dirichlet, inverted Dirichlet or multivariate Liouville distributions have multiplicative correlation structure.

An open problem is the estimation of the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$. For example, a one-factor model such as

$$X_i = E(X_i) + a_i Z_0 + \sqrt{b_i} Z_i, \quad i = 1, \ldots, n,$$

is frequently used in the analysis of asset returns. In our framework, this is equivalent to estimate $\boldsymbol{a}$ or $\boldsymbol{b}$ by assuming a multiplicative model. An advantage of such an approach is that it is distribution free. For example, we can check if multiplicative correlation model fits to data or not by consulting Lemma 2.1 before the estimation of parameters. We note that a multiplicative covariance matrix does not necessarily imply a unique factorization, but allows various choices of orthogonal or non-orthogonal factorizations. We hope to develop an efficient estimation algorithm of the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$. It is also interesting to investigate if measurement exchangeability (Kelderman, 2004) can be extended to non-normal distribution which has a multiplicative correlation structure.

Acknowledgements The authors wish to thank the referees and Dr. Peter Thomson for their constructive suggestions.

References

Abbott, F., Dayan, P. (1999). The effect of correlated variability on the accuracy of a population code. *Neural Computation*, *11*, 91–101.

Baba, K., Shibata, R., Sibuya, M. (2004). Partial correlation and conditional correlation as measures of conditional independence. *Australian & New Zealand Journal of Statistics*, *46*, 657–664.

Curnow, R.N., Dunnett, C.W. (1962). The numerical evaluation of certain multivariate normal integrals. *Annals of Mathematical Statistics*, *33*, 571–579.

Gupta, S.S. (1963). Probability integrals of multivariate normal and multivariate t. *Annals of Mathematical Statistics*, *34*, 792–828.

Gupta, R.D., Richards, D.S.P. (2001). The covariance structure of the multivariate Liouville distributions. In: M.A.G. Viana, D. St. Richards *Algebraic methods in statistics and probability*, pp 125–138, Indiana: American Mathematical Society.

Hammersley, J.M., Handscomb, D.C. (1964). *Monte Carlo Methods*. London: Chapman and Hall.

Johnson, N.L., Kotz, S., Balakrishnan, N. (1997). *Discrete multivariate distributions*. New York: Wiley.

Kelderman, H. (2004). Measurement exchangeability and normal one-factor models. *Biometrika*, *91*, 738–742.

Khatri, C.G. (1967). On certain inequalities for normal distributions and their applications to simultaneous confidence bounds. *Annals of Mathematical Statistics*, *38*(6), 1853–1867.

Kotz, S., Balakrishnan, N., Johnson, N.L. (2000). *Continuous multivariate distributions*. New York: Wiley.

Mardia, K.V. (1970). *Families of bivariate distributions*. London: Griffin

Marshall, A.W., Olkin, I. (1979). *Inequalities: Theory of majorization and Its Applications*. New York: Academic Press.

Ronning, G. (1982). Characteristic values and triangular factorization of the covariance matrix for multinomial, Dirichlet and multivariate hypergeometric distributions and some related results. *Statistische Hefte*, *23*(3), 152–176.

Royen, T. (1991). Multivariate gamma distributions with one-factorial accompanying correlation matrices and applications to the distribution of the multivariate range. *Metrika*, *38*, 299–315.

Six, F.B. (1981). Representations of multivariate normal distributions with special correlation structures. *Communications in statistics: Theory & methods A10*(13), 1285–1295.

Steiger, J.H. (1979). Factor indeterminacy in the 1930's and the 1970's some interesting parallels. *Psychometrika*, *44*(1), 157–167.

Watson, G.S. (1996). Spectral decomposition of the covariance matrix of a multinomial. *Journal of the Royal Statistical Society* B. *58*(1), 289–291.