PRINCIPAL COMPONENT ANALYSIS WITH DROP RANK COVARIANCE MATRIX

Yitong Guo and Bingo Wing-Kuen Ling*

School of Information Engineering, Guangdong University of Technology
Guangzhou, 510006, China

(Communicated by Bin Li)

Abstract. This paper considers the principal component analysis when the covariance matrix of the input vectors drops rank. This case sometimes happens when the total number of the input vectors is very limited. First, it is found that the eigen decomposition of the covariance matrix is not uniquely defined. This implies that different transform matrices could be obtained for performing the principal component analysis. Hence, the generalized form of the eigen decomposition of the covariance matrix is given. Also, it is found that the matrix with its columns being the eigenvectors of the covariance matrix is not necessary to be unitary. This implies that the transform for performing the principal component analysis may not be energy preserved. To address this issue, the necessary and sufficient condition for the matrix with its columns being the eigenvectors of the covariance matrix to be unitary is derived. Moreover, since the design of the unitary transform matrix for performing the principal component analysis is usually formulated as an optimization problem, the necessary and sufficient condition for the first order derivative of the Lagrange function to be equal to the zero vector is derived. In fact, the unitary matrix with its columns being the eigenvectors of the covariance matrix is only a particular case of the condition. Furthermore, the necessary and sufficient condition for the second order derivative of the Lagrange function to be a positive definite function is derived. It is found that the unitary matrix with its columns being the eigenvectors of the covariance matrix does not satisfy this condition. Computer numerical simulation results are given to valid the results.

1. **Introduction.** Nowadays, datasets with high dimensional vectors are used for conducting research in various fields of science and engineering. In order to efficiently analyze the vectors in these datasets and implement various machine intelligence algorithms on the vectors, it is required to reduce the dimensions of the vectors but preserve most of their information [18], [7], [9], [6].

There are many dimension reduction methods. Among them, the principal component analysis [2], [9], [16] which is also known as the Karhunen Loeve transform, is the most common method used in many science and engineering applications. This is a statistical approach. Its main idea is to map the vectors in the high dimensional space to a low dimensional space via a unitary matrix. Here, the columns of the matrix are the eigenvectors of the covariance matrix of the vectors, where

---

the eigenvectors are sorted according to the descending order of the magnitudes of the corresponding eigenvalues. As the eigenvalues are sorted accordingly, the maximal information of the vectors is preserved if most of the vectors are localized in a certain subspace span by the first several eigenvectors.

Currently, the research on the principal component analysis mainly includes the following aspects. First, apply the principal component analysis on the high dimensional vectors to perform the dimension reductions for the machine learning and the data analysis applications [8], [20]. This includes the pattern recognition applications [15] of the subspace mapping [1] and the subspace learning [14]. For example, it can be applied to the optimal design of the agile supply chain network under an uncertainty [12] and the reduction of the total number of the dimensions of a logistic regression model with the continuous covariates and the avoidance of the multi-collinearity [4]. Second, develop the enhanced versions of the principal component analysis [3] such as the sparse principal component analysis [23], [21] and the robust principal component analysis [19], [10], [5]. The sparse principal component analysis produces the amendatory principal components with the sparse coefficients. On the other hand, the robust principal component analysis considers the problem of recovering a low rank matrix from the sparsely noise corrupted vectors. Finally, study the high dimensional principal component analysis such as the tensor principal component analysis [22], [17] and its applications. The tensor principal component analysis is an extension of the principal component analysis for the three multi-dimensional tensor data. It is widely used to exploit the spatial covariances of the hyperspectral images and the videos.

It is worth noting that the existing principal component analysis algorithms assume that the covariance matrix of the vectors is full rank. Or in other words, the covariance matrix is strictly positive definite. In this case, the eigen decomposition is unique and the matrix with its columns being these eigenvectors is unitary. However, when the total number of the vectors is very limited, the covariance matrix may drop rank. In this case, the covariance matrix becomes semi-positive definite and the eigen decomposition is no longer unique. Also, the matrix with its columns being these eigenvectors does not guarantee to be unitary. Hence, it is difficult to perform the principal component analysis.

This paper studies the properties of the principal component analysis when the covariance matrix of the vectors drops rank. The outline of this paper is as follows. The existing principal component analysis algorithm is reviewed in Section 2. Section 3 presents the properties of the principal component analysis when the covariance matrix drops rank. In Section 4, the computer numerical simulation results are presented. Finally, a conclusion is drawn in Section 5.

2. **Review on existing principal component analysis algorithm.** The detail procedures for performing the conventional principal component analysis are summarized as follows [9]:

Step 1: Acquire the input vectors. Let M and N be the dimensions and the total number of the input vectors, respectively. In this section, it is assumed that $M < N$. This is because this is the situation found in many practical situations. Denote the input vectors as $x_n = \begin{bmatrix} x_{1,n} & \cdots & x_{M,n} \end{bmatrix}^T$ for $n = 1, \cdots, N$ and the matrix with its columns being the input vectors as

$$X = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix}. \tag{1}$$

Step 2: Remove the mean of the input vectors. Let $\mu = \begin{bmatrix} \mu_1 & \cdots & \mu_M \end{bmatrix}^T$ be the mean of the input vectors. That is,

$$\mu_m = \frac{1}{N} \sum_{n=1}^{N} x_{m,n} \quad \text{for} \quad m = 1, \cdots, M. \tag{2}$$

Denote the operator for removal the mean of x_n as $x_n - \mu$ for $n = 1, \cdots, N$ and the vectors after removal the mean of x_n for $n = 1, \cdots, N$. That is, $\bar{x}_n = x_n - \mu$ for $n = 1, \cdots, N$. Let the matrix with its columns being the vectors after removal the mean of x_n be $\bar{X} = \begin{bmatrix} \bar{x}_1 & \cdots & \bar{x}_N \end{bmatrix}$.

Step 3: Calculate the covariance matrix. Let C be the covariance matrix of the input vectors. That is,

$$C = \frac{1}{N} \bar{X} \bar{X}^T. \tag{3}$$

Step 4: Perform the eigen decomposition on C. Here, the eigenvectors and the eigenvalues of C are computed. Denote $\theta_1, \cdots, \theta_M$ as the eigenvalues of C. Let $U_1, \cdots, U_M$ be the corresponding eigenvectors.

Step 5: Select the principal components. Arange $U_1, \cdots, U_M$ according to the descending orders of the magnitudes of $\theta_1, \cdots, \theta_M$. It is worth noting that the eigenvector with the largest eigenvalue is the principal component of the original dataset. Suppose that the first k eigenvectors are selected. Obviously, $k < M$. Let the unitary matrix for performing the principal component analysis be $\bar{U}$. It is formed by setting the columns of $\bar{U}$ as these first k eigenvectors. That is, $\bar{U} = \begin{bmatrix} U_1 & \cdots & U_k \end{bmatrix}$.

Step 6: Reduce the dimensions of the vectors by multiplying the transpose of $\bar{U}$ to $\bar{X}$. Let $\bar{y}_n = \begin{bmatrix} \bar{y}_{1,n} & \cdots & \bar{y}_{k,n} \end{bmatrix}^T$ for $n = 1, \cdots, N$ be the transformed vectors and $\bar{Y} = \begin{bmatrix} \bar{y}_1 & \cdots & \bar{y}_N \end{bmatrix}$. That is,

$$\bar{Y} = \bar{U}^T \bar{X}. \tag{4}$$

Obviously, the columns of $\bar{Y}$ are the transformed vectors with the reduced dimensions.

3. **Properties of principal component analysis when covariance matrix drops rank.** As the principal component analysis involves the eigen decomposition of C and $C = \frac{1}{N} \bar{X} \bar{X}^T$, it is important to study the matrix decomposition of $\bar{X}$. In this section, two types of matrix decompositions of $\bar{X}$ are studied. Also, the relationship between these two types of matrix decompositions is explored.

For the first type of the matrix decomposition of $\bar{X}$, let R be the total number of the independent columns of $\bar{X}$. That is, $rank(\bar{X}) = R$. Then, there exists an $M \times R$ matrix F and there exists an $R \times N$ matrix S such that $\bar{X} = FS$. It is worth noting that the decomposition of $\bar{X}$ is not uniquely defined. Hence, it is required to find one of the realizations of F and S. To achieve this goal, we pick out R columns of $\bar{X}$ where these R columns are linearly independent. Denote $\check{F}$ as the matrix with its columns being these linearly independent columns. Now, F can be expressed as the product of $\check{F}$ and an $R \times R$ random matrix denoted as H. That is, $F = \check{F}H$. Since $\bar{X} = FS$, we have $F^T \bar{X} = F^T FS$ or $S = (F^T F)^{-1} F^T \bar{X}$.

As the principal component analysis involves the eigen decomposition of C, $C = \frac{1}{N} \bar{X} \bar{X}^T$ and $\bar{X} = FS$, it is required to study the decomposition of F in order to study the properties of the principal component analysis when the covariance matrix drops rank, that is, $R < M$. Let U_F, V_F and D_F be the left unitary matrix, the right unitary matrix and the square diagonal matrix obtained by performing the

singular value decomposition on F, respectively. That is, $F = U_F \begin{bmatrix} D_F \\ 0 \end{bmatrix} V_F^T$. Here, it is assumed that the singular values are sorted in the descending order. Denote $\tilde{S} = S^T V_F$ and $\hat{S} = D_F \tilde{S}^T \tilde{S} D_F$. Let U be the matrix used for performing the dimension reduction in the principal component analysis. Let $\tilde{U} = U_F^T U$.

For the completeness, it is required to study the uniqueness of the decomposition of F. It is worth noting that the vectors in a null space are not uniquely defined. Therefore, the decomposition of F is not uniquely defined.

Corollary 1. *Let* $U_F = \begin{bmatrix} U_{F_1} & U_{F_2} \end{bmatrix}$. *The columns of* U_{F_2} *are in the null space of* F^T. *That is,* $F^T U_{F_2} = 0$.

Proof. Let $V_F = \begin{bmatrix} V_{F_1} & V_{F_2} \end{bmatrix}$. Since $F = \begin{bmatrix} U_{F_1} & U_{F_2} \end{bmatrix} \begin{bmatrix} D_F \\ 0 \end{bmatrix} \begin{bmatrix} V_{F_1}^T \\ V_{F_2}^T \end{bmatrix}$, we have $\begin{bmatrix} U_{F_1}^T \\ U_{F_2}^T \end{bmatrix} F = \begin{bmatrix} D_F \\ 0 \end{bmatrix} \begin{bmatrix} V_{F_1}^T \\ V_{F_2}^T \end{bmatrix} = \begin{bmatrix} D_F V_{F_1}^T \\ 0 \end{bmatrix}$. As a result, $U_{F_2}^T F = 0$ or $F^T U_{F_2} = 0$. This completes the proof. $\square$

Since F is not a square matrix, U_F is not uniquely defined. As a result, the decomposition of F is not uniquely defined.

For another type of the matrix decomposition of $\bar{X}$, a generalized singular value decomposition of $\bar{X}$ is considered. The result is summarized in the following corollary:

Corollary 2. $\bar{X}$ *can be decomposed as the product of a left unitary matrix* $U_X = \begin{bmatrix} U_{X_1} & U_{X_2} \end{bmatrix}$, *a right unitary matrix* $V_X = \begin{bmatrix} V_{X_1} & V_{X_2} \end{bmatrix}$ *as well as a block diagonal matrix with one of its sub block matrix being an* $R \times R$ *matrix denoted as* B *and the rest sub block matrix being the zero matrix. That is,* $\bar{X} = U_X \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} V_X^T$. *Also, the columns of* U_{X_2} *are in the null space of* $\bar{X}^T$ *and the columns of* U_{X_1} *are in the null space of* $U_{X_2}^T$. *That is,* $\bar{X}^T U_{X_2} = 0$ *and* $U_{X_2}^T U_{X_1} = 0$. *Similarly, the columns of* V_{X_2} *are in the null space of* $\bar{X}$ *and the columns of* V_{X_1} *are in the null space of* $V_{X_2}^T$. *That is,* $\bar{X} V_{X_2} = 0$ *and* $V_{X_2}^T V_{X_1} = 0$.

Proof. Since $rank(\bar{X}) = R$, there exists an $M \times M$ unitary matrix U_X, an $N \times N$ unitary matrix V_X and an $R \times R$ matrix B such that $\bar{X} = U_X \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} V_X^T$. This implies that $\begin{bmatrix} U_{X_1}^T \\ U_{X_2}^T \end{bmatrix} \bar{X} = \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_{X_1}^T \\ V_{X_2}^T \end{bmatrix}$ or $\begin{bmatrix} U_{X_1}^T \bar{X} \\ U_{X_2}^T \bar{X} \end{bmatrix} = \begin{bmatrix} B V_{X_1}^T \\ 0 \end{bmatrix}$. Hence, we have $U_{X_2}^T \bar{X} = 0$ or $\bar{X}^T U_{X_2} = 0$. As U_X is unitary, we have $U_{X_2}^T U_{X_1} = 0$. Similarly, we have $\bar{X} \begin{bmatrix} V_{X_1} & V_{X_2} \end{bmatrix} = \begin{bmatrix} U_{X_1} & U_{X_2} \end{bmatrix} \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix}$ or $\begin{bmatrix} \bar{X} V_{X_1} & \bar{X} V_{X_2} \end{bmatrix} = \begin{bmatrix} U_{X_1} B & 0 \end{bmatrix}$. As a result, we have $\bar{X} V_{X_2} = 0$. As V_X is unitary, we have $V_{X_2}^T V_{X_1} = 0$. This completes the proof. $\square$

It is worth noting that B is not necessary to a diagonal matrix. If it is a diagonal matrix, then this decomposition becomes the conventional singular value decomposition. Hence, this decomposition can be considered as the generalized singular value decomposition of $\bar{X}$.

Likewise, it is required to study the uniqueness of the decomposition of $\bar{X}$. Similarly, as the vectors in a null space are not uniquely defined, U_X and V_X are not

uniquely defined. As a result, the generalized singular value decomposition of $\bar{X}$ is not uniquely defined.

Since there are two different types of decompositions, it is required to explore the relationship between them. The result is summarized in the following corollary:

Corollary 3. *If $U_X = U_F$, then $\hat{S} = BB^T$.*

Proof. It is worth noting that

$$C = \frac{1}{N}\bar{X}\bar{X}^T = \frac{1}{N}U_X \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} V_X^T V_X \begin{bmatrix} B^T & 0 \\ 0 & 0 \end{bmatrix} U_X^T = \frac{1}{N}U_X \begin{bmatrix} BB^T & 0 \\ 0 & 0 \end{bmatrix} U_X^T \quad (5)$$

and

$$C = \frac{1}{N}FSS^T F^T = \frac{1}{N}U_F \begin{bmatrix} D_F \\ 0 \end{bmatrix} V_F^T SS^T V_F \begin{bmatrix} D_F & 0 \end{bmatrix} U_F^T$$

$$= \frac{1}{N}U_F \begin{bmatrix} D_F \\ 0 \end{bmatrix} \tilde{S}^T \tilde{S} \begin{bmatrix} D_F & 0 \end{bmatrix} U_F^T = \frac{1}{N}U_F \begin{bmatrix} D_F \tilde{S}^T \tilde{S} D_F & 0 \\ 0 & 0 \end{bmatrix} U_F^T \quad (6)$$

$$= \frac{1}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T$$

If $U_X = U_F$, then $\hat{S} = BB^T$. This completes the proof. $\qquad\square$

Here, it can be seen that $\hat{S}$ is strictly positive definite because B is a full rank matrix.

Based on the above results, we have the following results on the eigen decomposition of C:

Theorem 3.1. *The following two sets of statements are equivalent.*
(A) The columns of U are the eigenvectors of C and there exists a diagonal matrix Θ with its diagonal elements being the eigenvalues of C. That is, $CU = U\Theta$, where $U = [\ U_1\ \cdots\ U_M\ ]$ and $\Theta = diag(\ \theta_1,\ \cdots\ ,\theta_M\ )$. Here, it is assumed that the columns of U and the diagonal elements of Θ are sorted according to the descending orders of the magnitudes of the eigenvalues.
(B) There exists an arbitrary $(M-R)\times(M-R)$ matrix $\tilde{U}_B$, there exists a matrix $\tilde{U}_A$ with its columns being the eigenvectors of $\frac{\hat{S}}{N}$ and there exists a diagonal matrix Θ_1 with its diagonal elements being the corresponding eigenvalues of $\frac{\hat{S}}{N}$ such that $\frac{\hat{S}\tilde{U}_A}{N} = \tilde{U}_A\Theta_1$, $\tilde{U} = \begin{bmatrix} \tilde{U}_A & 0 \\ 0 & \tilde{U}_B \end{bmatrix}$ and $\Theta = \begin{bmatrix} \Theta_1 & 0 \\ 0 & 0 \end{bmatrix}$. Here, it is assumed that the diagonal elements of Θ_1 are sorted according to the descending orders of their magnitudes.

Proof. Since $C = \frac{1}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T$, $CU = U\Theta$ implies that $\frac{1}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T U = U\Theta$. In other words, we have $\frac{1}{N} \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T U = U_F^T U\Theta$ or $\frac{1}{N} \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} = \tilde{U}\Theta$. Since Θ is a diagonal matrix, this implies that the columns of $\tilde{U}$ are the eigenvectors of $\frac{1}{N} \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix}$ and the diagonal elements of Θ are the corresponding eigenvalues. Besides, as $rank\left(\frac{1}{N} \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \right) = R$, $\frac{1}{N} \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix}$ only consists of R nonzero eigenvalues. In fact, these nonzero eigenvalues are the eigenvalues of $\frac{\hat{S}}{N}$. That is, Θ_1. As a result,

$\Theta = \begin{bmatrix} \Theta_1 & 0 \\ 0 & 0 \end{bmatrix}$. Let the sub matrices of $\tilde{U}$ be $\tilde{U}_{1,1}$, $\tilde{U}_{1,2}$, $\tilde{U}_{2,1}$ and $\tilde{U}_{2,2}$. That is,
$\tilde{U} = \begin{bmatrix} \tilde{U}_{1,1} & \tilde{U}_{1,2} \\ \tilde{U}_{2,1} & \tilde{U}_{2,2} \end{bmatrix}$. Then, we have $\frac{1}{N} \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{U}_{1,1} & \tilde{U}_{1,2} \\ \tilde{U}_{2,1} & \tilde{U}_{2,2} \end{bmatrix} = \begin{bmatrix} \tilde{U}_{1,1} & \tilde{U}_{1,2} \\ \tilde{U}_{2,1} & \tilde{U}_{2,2} \end{bmatrix} \begin{bmatrix} \Theta_1 & 0 \\ 0 & 0 \end{bmatrix}$.
This implies that $\frac{1}{N} \begin{bmatrix} \hat{S}\tilde{U}_{1,1} & \hat{S}\tilde{U}_{1,2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \tilde{U}_{1,1}\Theta_1 & 0 \\ \tilde{U}_{2,1}\Theta_1 & 0 \end{bmatrix}$. This further implies that
$\frac{\hat{S}\tilde{U}_{1,1}}{N} = \tilde{U}_{1,1}\Theta_1$, $\frac{\hat{S}\tilde{U}_{1,2}}{N} = 0$ and $\tilde{U}_{2,1}\Theta_1 = 0$. As $\hat{S}$ and Θ_1 are full rank matrices, this implies that $\tilde{U}_{1,2} = 0$ and $\tilde{U}_{2,1} = 0$. Also, the columns of $\tilde{U}_{1,1}$ are the eigenvectors of $\frac{\hat{S}}{N}$. This implies that $\tilde{U}_{1,1} = \tilde{U}_A$. As the above eigen relationships are independent of $\tilde{U}_{2,2}$, $\tilde{U}_{2,2}$ can be any arbitrary $(M-R) \times (M-R)$ matrix. Hence, $\tilde{U}_{2,2} = \tilde{U}_B$ and $\tilde{U} = \begin{bmatrix} \tilde{U}_A & 0 \\ 0 & \tilde{U}_B \end{bmatrix}$. As a result, the set of statements (A) implies the set of statements (B).

Since

$$\begin{aligned}
CU &= \frac{1}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T U_F \tilde{U} = \frac{1}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} \\
&= \frac{1}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{U}_A & 0 \\ 0 & \tilde{U}_B \end{bmatrix} = U_F \begin{bmatrix} \frac{\hat{S}\tilde{U}_A}{N} & 0 \\ 0 & 0 \end{bmatrix} \\
&= U_F \begin{bmatrix} \tilde{U}_A\Theta_1 & 0 \\ 0 & 0 \end{bmatrix} = U_F \begin{bmatrix} \tilde{U}_A & 0 \\ 0 & \tilde{U}_B \end{bmatrix} \begin{bmatrix} \Theta_1 & 0 \\ 0 & 0 \end{bmatrix} \\
&= U_F\tilde{U}\Theta = U\Theta
\end{aligned} \tag{7}$$

and Θ is a diagonal matrix, this implies that the columns of U are the eigenvectors of C and diagonal elements of Θ are the corresponding eigenvalues. Hence, the set of statements (B) implies the set of statements (A). As a result, these two sets of statements are equivalent. This completes the proof. $\square$

From here, it can be seen that the eigen decomposition of C depends on whether $\tilde{U}$ being a block diagonal matrix with the block sub matrices being $\tilde{U}_A$ and $\tilde{U}_B$ as well as Θ being a diagonal matrix with the block diagonal sub matrices Θ_1 and the zero matrix such that the columns of $\tilde{U}_A$ are the eigenvectors of $\frac{\hat{S}}{N}$ and the diagonal elements of Θ_1 are their corresponding eigenvalues.

When C is a full rank matrix, the set of statements (B) becomes $\tilde{U} = \tilde{U}_A$, $\Theta = \Theta_1$ and $\frac{\hat{S}\tilde{U}_A}{N} = \tilde{U}_A\Theta_1$. Since $C = \frac{1}{N}U_F\hat{S}U_F^T$, we have

$$\begin{aligned}
CU &= \frac{1}{N}U_F\hat{S}U_F^T U_F\tilde{U} = \frac{1}{N}U_F\hat{S}\tilde{U} = \frac{1}{N}U_F\hat{S}\tilde{U}_A \\
&= U_F\frac{\hat{S}\tilde{U}_A}{N} = U_F\tilde{U}_A\Theta_1 = U_F\tilde{U}\Theta = U\Theta.
\end{aligned} \tag{8}$$

This implies that the columns of U are the eigenvectors of C with the diagonal elements in Θ being the corresponding eigenvalues. This further implies the set of statements (A). On the other hand, $CU = U\Theta$ in the set of statements (A) becomes $\frac{1}{N}U_F\hat{S}U_F^T U_F\tilde{U}_A = U_F\tilde{U}_A\Theta_1$ or $\frac{1}{N}\hat{S}\tilde{U}_A = \tilde{U}_A\Theta_1$. This implies that the columns of $\tilde{U}_A$ are the eigenvectors of $\frac{\hat{S}}{N}$ with the diagonal elements in Θ_1 being the corresponding eigenvalues. This implies the set of statements (B). From here, it

can be seen that the conventional principal component analysis with the full rank matrix C is a particular case of **Theorem 3**.**1**.

From **Theorem 3**.**1**, we find the following two properties.

Corollary 4. *U is not uniquely defined.*

Proof. Since $\tilde{U}_B$ is an arbitrary $(M - R) \times (M - R)$ matrix and $\tilde{U} = U_F^T U$, U is not uniquely defined. This completes the proof. $\square$

Since U is not uniquely defined, different matrices U would be obtained in the **Step 4** of the principal component analysis algorithm.

Corollary 5.
  U is unitary if and only if both $\tilde{U}_A$ and $\tilde{U}_B$ are unitary.

Proof. Since $\tilde{U} = U_F^T U$ and U_F is unitary, U is unitary if and only if $\tilde{U}$ is unitary. As $\tilde{U} = \begin{bmatrix} \tilde{U}_A & 0 \\ 0 & \tilde{U}_B \end{bmatrix}$, $\tilde{U}$ is unitary if and only if both $\tilde{U}_A$ and $\tilde{U}_B$ are unitary. This completes the proof. $\square$

It is worth noting that $\frac{\hat{S}}{N}$ is a strictly positive definite matrix. Therefore, $\tilde{U}_A$ is a unitary matrix. However, as $\tilde{U}_B$ is an arbitrary $(M - R) \times (M - R)$ matrix, it is not necessary to be a unitary matrix. This implies that U is not necessary to be a unitary matrix. As a result, it cannot guarantee that the matrix obtained in the **Step 5** of the principal component analysis algorithm is a unitary matrix.

Now, perform the dimension reduction via the principal component analysis. Let $J(U)$ be the average energy of the difference between the transformed vectors before and after the dimension reduction. That is, $J(U) = \frac{1}{N} \sum_{i=1}^{N} \left\| \left(U^T - \begin{bmatrix} \bar{U}^T \\ 0 \end{bmatrix} \right) \bar{x}_i \right\|^2$. Then, the transform matrix design problem can be formulated as the following optimization problem:
Problem(P)

$$\min_{U} \quad J(U) = \frac{1}{N} \sum_{i=1}^{N} \left\| \left(U^T - \begin{bmatrix} \bar{U}^T \\ 0 \end{bmatrix} \right) \bar{x}_i \right\|^2, \tag{9a}$$

$$\text{subject to} \quad U_i^T U_j = \delta(i - j) \quad \text{for} \quad i = 1, \cdots, M \quad \text{and} \quad \text{for} \quad j = 1, \cdots, M. \tag{9b}$$

Here, $\delta(\cdot)$ denotes the discrete time Dirac delta function. To solve an equality constrained optimization problem, a Lagrange approach is usually employed. Let $\lambda_{i,j}$ for $i = 1, \cdots, M$ and for $j = 1, \cdots, M$ be the optimal Lagrange multipliers. Define $\lambda = \begin{bmatrix} \lambda_{1,1} & \cdots & \lambda_{1,M} \\ \vdots & \ddots & \vdots \\ \lambda_{M,1} & \cdots & \lambda_{M,M} \end{bmatrix}$. Let the Lagrange function be

$$L(U, \lambda) = \frac{1}{N} \sum_{i=1}^{N} \left\| \left(U^T - \begin{bmatrix} \bar{U}^T \\ 0 \end{bmatrix} \right) \bar{x}_i \right\|^2 - \sum_{i=1}^{M} \sum_{j=1}^{M} \lambda_{i,j} (U_r^T U_r - \delta(i - j)). \tag{10}$$

Then, a solution is found in such a way that the first order derivative of the Lagrange function is equal to the zero vector. To find such a solution, we have the following result:

Theorem 3.2. *For an arbitrary $M \times M$ unitary matrix U, an arbitrary $R \times R$ unitary matrix $\hat{V}_{1,1}$ and an arbitrary $(M-R) \times (M-R)$ unitary matrix $\hat{V}_{2,2}$, let $V_\lambda = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} & 0 \\ 0 & \hat{V}_{2,2} \end{bmatrix}$ and $Q_\lambda = \frac{1}{N}\hat{V}_{1,1}^T \hat{S} \hat{V}_{1,1}$. Then, the necessary and sufficient condition for the first order derivative of the Lagrange function to be equal to the zero vector is $\frac{\lambda + \lambda^T}{2} = V_\lambda \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} V_\lambda^T$.*

Proof. First, it is worth noting that

$$
\begin{aligned}
J(U) &= \frac{1}{N}\sum_{i=1}^{N} \left\| \left(U^T - \begin{bmatrix} \bar{U}^T \\ 0 \end{bmatrix} \right) \bar{x}_i \right\|^2 = \frac{1}{N}\sum_{i=1}^{N} \left\| \left(\begin{bmatrix} 0 \\ U_{k+1}^T \\ \vdots \\ U_M^T \end{bmatrix} \right) \bar{x}_i \right\|^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}\sum_{r=k+1}^{M} \left\| U_r^T \bar{x}_i \right\|^2 = \frac{1}{N}\sum_{i=1}^{N}\sum_{r=k+1}^{M} U_r^T \bar{x}_i \bar{x}_i^T U_r \\
&= \frac{1}{N}\sum_{r=k+1}^{M}\sum_{i=1}^{N} U_r^T \bar{x}_i \bar{x}_i^T U_r = \frac{1}{N}\sum_{r=k+1}^{M} U_r^T \bar{X}\bar{X}^T U_r = \sum_{r=k+1}^{M} U_r^T C U_r
\end{aligned}
\tag{11}
$$

Hence, Problem(P) is equivalent to the following optimization problem:

$$
\min_{U} \qquad J(U) = \sum_{r=k+1}^{M} U_r^T C U_r, \tag{12a}
$$

$$
\text{subject to} \qquad U_i^T U_j = \delta(i-j) \quad \text{for} \quad i=1,\cdots,M \quad \text{and} \quad \text{for} \quad j=1,\cdots,M. \tag{12b}
$$

Now, the Lagrange function becomes

$$
L(U,\lambda) = \sum_{r=k+1}^{M} U_r^T C U_r - \sum_{i=1}^{M}\sum_{j=1}^{M} \lambda_{i,j}(U_r^T U_r - \delta(i-j)). \tag{13}
$$

Since $\frac{\partial}{\partial U_r}L(U,\lambda) = 2CU_r - \sum_{i=1}^{M}(\lambda_{i,r} + \lambda_{r,i})U_r$ for $r=1,\cdots,M$, the condition for the first order derivative of the Lagrange function to be equal to the zero vector is $2CU = U(\lambda + \lambda^T)$ or $\frac{2}{N}U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T U = U(\lambda + \lambda^T)$. This implies that $\frac{2}{N}U^T U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T U = \lambda + \lambda^T$ or $\frac{1}{N}\tilde{U}^T \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} = \frac{\lambda + \lambda^T}{2}$. As $rank\left(\frac{1}{N}\begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \right) = R$, $\frac{\lambda + \lambda^T}{2}$ only consists of R nonzero singular values. Also, as $\frac{\lambda + \lambda^T}{2}$ is a symmetric matrix, there exists a unitary matrix $\bar{V}_\lambda$ and a positive definite matrix $\bar{Q}_\lambda$ such that $\frac{\lambda + \lambda^T}{2} = \bar{V}_\lambda \begin{bmatrix} \bar{Q}_\lambda & 0 \\ 0 & 0 \end{bmatrix} \bar{V}_\lambda^T$. Hence, we have $\frac{1}{N}\tilde{U}^T \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} = \bar{V}_\lambda \begin{bmatrix} \bar{Q}_\lambda & 0 \\ 0 & 0 \end{bmatrix} \bar{V}_\lambda^T$ or $\frac{1}{N}\begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \tilde{U}\bar{V}_\lambda = \tilde{U}\bar{V}_\lambda \begin{bmatrix} \bar{Q}_\lambda & 0 \\ 0 & 0 \end{bmatrix}$. Let $\tilde{U}\bar{V}_\lambda = \check{V} = \begin{bmatrix} \check{V}_{1,1} & \check{V}_{1,2} \\ \check{V}_{2,1} & \check{V}_{2,2} \end{bmatrix}$. Then, we have $\frac{1}{N}\begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} \check{V}_{1,1} & \check{V}_{1,2} \\ \check{V}_{2,1} & \check{V}_{2,2} \end{bmatrix} = \begin{bmatrix} \check{V}_{1,1} & \check{V}_{1,2} \\ \check{V}_{2,1} & \check{V}_{2,2} \end{bmatrix}\begin{bmatrix} \bar{Q}_\lambda & 0 \\ 0 & 0 \end{bmatrix}$. This implies that $\frac{1}{N}\begin{bmatrix} \hat{S}\check{V}_{1,1} & \hat{S}\check{V}_{1,2} \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \check{V}_{1,1}\bar{Q}_\lambda & 0 \\ \check{V}_{2,1}\bar{Q}_\lambda & 0 \end{bmatrix}$. In other words, we have $\frac{\hat{S}\check{V}_{1,1}}{N} = \check{V}_{1,1}\bar{Q}_\lambda$, $\frac{\hat{S}\check{V}_{1,2}}{N} = 0$ and $\check{V}_{2,1}\overline{Q}_\lambda = 0$. As

$\hat{S}$ and $\bar{Q}_\lambda$ are full rank, we have $\check{V}_{1,2} = 0$ and $\check{V}_{2,1} = 0$. Since $\check{V}$ is a unitary matrix as well as $\check{V}_{1,2} = 0$ and $\check{V}_{2,1} = 0$, we have both $\check{V}_{1,1}$ and $\check{V}_{2,2}$ are unitary matrices. As a result, we have $\bar{Q}_\lambda = \frac{1}{N}\check{V}_{1,1}^T \hat{S} \check{V}_{1,1}$. Since U, $\check{V}_{1,1}$ and $\check{V}_{2,2}$ are only constrained to be unitary matrices, they could be any arbitrary $M \times M$ unitary matrix, any arbitrary $R \times R$ unitary matrix and any arbitrary $(M-R) \times (M-R)$ unitary matrix, respectively. Hence, without the loss of generality, let $\hat{V}_{1,1} = \check{V}_{1,1}$ and $\hat{V}_{2,2} = \check{V}_{2,2}$. Then, we have $\bar{V}_\lambda = \tilde{U}^T \begin{bmatrix} \check{V}_{1,1} & \check{V}_{1,2} \\ \check{V}_{2,1} & \check{V}_{2,2} \end{bmatrix} = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} & 0 \\ 0 & \hat{V}_{2,2} \end{bmatrix} = V_\lambda$ and $\bar{Q}_\lambda = \frac{1}{N}\check{V}_{1,1}^T \hat{S}\check{V}_{1,1} = \frac{1}{N}\hat{V}_{1,1}^T \hat{S}\hat{V}_{1,1} = Q_\lambda$. As a result, we have $\frac{\lambda + \lambda^T}{2} = \bar{V}_\lambda \begin{bmatrix} \bar{Q}_\lambda & 0 \\ 0 & 0 \end{bmatrix} \bar{V}_\lambda^T = V_\lambda \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} V_\lambda^T$. This completes the proof of the necessity part.

Now, consider the sufficiency part. Let $\check{V}_{1,1} = \hat{V}_{1,1}$ and $\check{V}_{2,2} = \hat{V}_{2,2}$. Then, we have $V_\lambda = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} & 0 \\ 0 & \hat{V}_{2,2} \end{bmatrix} = \tilde{U}^T \begin{bmatrix} \check{V}_{1,1} & 0 \\ 0 & \check{V}_{2,2} \end{bmatrix}$ and $Q_\lambda = \frac{1}{N}\hat{V}_{1,1}^T \hat{S}\hat{V}_{1,1} = \frac{1}{N}\check{V}_{1,1}^T \hat{S}\check{V}_{1,1}$. Hence,

$$
\begin{aligned}
\frac{\lambda + \lambda^T}{2} &= V_\lambda \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} V_\lambda^T = \tilde{U}^T \begin{bmatrix} \check{V}_{1,1} & 0 \\ 0 & \check{V}_{2,2} \end{bmatrix} \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \check{V}_{1,1}^T & 0 \\ 0 & \check{V}_{2,2}^T \end{bmatrix} \tilde{U} \\
&= \tilde{U}^T \begin{bmatrix} \check{V}_{1,1} Q_\lambda \check{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} = U^T U_F \begin{bmatrix} \check{V}_{1,1} Q_\lambda \check{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} U_F^T U
\end{aligned}
\tag{14}
$$

This implies that

$$
\begin{aligned}
U\left(\frac{\lambda + \lambda^T}{2}\right) U^T &= UU^T U_F \begin{bmatrix} \check{V}_{1,1} Q_\lambda \check{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} U_F^T U U^T \\
&= U_F \begin{bmatrix} \check{V}_{1,1} Q_\lambda \check{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} U_F^T = \frac{1}{N} U_F \begin{bmatrix} \check{V}_{1,1} \check{V}_{1,1}^T \hat{S} \check{V}_{1,1} \check{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} U_F^T \\
&= \frac{1}{N} U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T = C.
\end{aligned}
\tag{15}
$$

This further implies that $2CU = 2U\left(\frac{\lambda + \lambda^T}{2}\right) U^T U = U(\lambda + \lambda^T)$. Since

$$
\frac{\partial}{\partial U_r} L(U, \lambda) = 2CU_r - \sum_{i=1}^{M} (\lambda_{i,r} + \lambda_{r,i}) U_r
\tag{16}
$$

for $r = 1, \cdots, M$, we have $\frac{\partial}{\partial U_r} L(U, \lambda) = 0$ for $r = 1, \cdots, M$. As a result, the condition for the first order derivative of the Lagrange function to be equal to the zero vector is satisfied. This completes the proof of the sufficiency part and it completes the whole proof. □

It is worth noting that

$$
\begin{aligned}
\frac{\lambda + \lambda^T}{2} &= V_\lambda \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} V_\lambda^T = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} & 0 \\ 0 & \hat{V}_{2,2} \end{bmatrix} \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{V}_{1,1}^T & 0 \\ 0 & \hat{V}_{2,2}^T \end{bmatrix} \tilde{U} \\
&= \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} Q_\lambda \hat{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} \frac{1}{N} \hat{V}_{1,1}^T \hat{S} \hat{V}_{1,1} \hat{V}_{1,1}^T & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} \\
&= \frac{1}{N} \tilde{U}^T \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} \tilde{U} = \frac{1}{N} \tilde{U}^T U_F \begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix} U_F^T U = U^T C U.
\end{aligned}
\tag{17}
$$

This implies that $CU = U\left(\frac{\lambda+\lambda^T}{2}\right)$. Hence, if $\frac{\lambda+\lambda^T}{2}$ is a diagonal matrix with its diagonal elements being the eigenvalues of C and the columns of U are the eigenvectors of C, then it reduces to the conventional principal component analysis.

In fact, the second order derivative of the Lagrange function is required to be strictly positive definite in order to guarantee that the obtain solution is a local minimum. To test this condition, we have the following result:

Theorem 3.3. *The necessary and sufficient condition for the second order derivative of the Lagrange function being strictly positive definite is $\frac{\hat{S}}{N} - \lambda_{p,p}I_R$ being strictly positive definite and $\lambda_{p,p} < 0$ for $p = 1, \cdots, M$, as well as $\lambda_{q,p} + \lambda_{p,q} < 0$ for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$.*

Proof. Since

$$\frac{\partial}{\partial U_p}L(U, \lambda) = 2CU_p - \sum_{q=1}^{M}(\lambda_{q,p} + \lambda_{p,q})U_q \tag{18}$$

for $p = 1, \cdots, M$, we have

$$\frac{\partial^2}{\partial U_p \partial U_p}L(U, \lambda) = 2C - 2\lambda_{p,p}I_M \tag{19}$$

for $p = 1, \cdots, M$ and

$$\frac{\partial^2}{\partial U_q \partial U_p}L(U, \lambda) = -(\lambda_{q,p} + \lambda_{p,q})I_M \tag{20}$$

for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$. Here, I_M denotes the $M \times M$ identity matrix. As

$$2C - 2\lambda_{p,p}I_M = 2\frac{1}{N}\begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix}U_F^T - 2\lambda_{p,p}U_FU_F^T$$
$$= 2U_F\left(\begin{bmatrix} \frac{\hat{S}}{N} - \lambda_{p,p}I_R & 0 \\ 0 & -\lambda_{p,p}I_{M-R} \end{bmatrix}\right)U_F^T \tag{21}$$

for $p = 1, \cdots, M$, $\frac{\partial^2}{\partial U_p \partial U_p}L(U, \lambda)$ is strictly positive definite implies that $\frac{\hat{S}}{N} - \lambda_{p,p}I_R$ is strictly positive definite and $\lambda_{p,p} < 0$ for $p = 1, \cdots, M$. Besides, $\frac{\partial^2}{\partial U_q \partial U_p}L(U, \lambda)$ is strictly positive definite implies that $\lambda_{q,p} + \lambda_{p,q} < 0$ for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$. This completes the proof of the necessity part.

On the other hand, if $\frac{\hat{S}}{N} - \lambda_{p,p}I_R$ is strictly positive definite and $\lambda_{p,p} < 0$ for $p = 1, \cdots, M$, then

$$2U_F\left(\begin{bmatrix} \frac{\hat{S}}{N} - \lambda_{p,p}I_R & 0 \\ 0 & -\lambda_{p,p}I_{M-R} \end{bmatrix}\right)U_F^T$$
$$= 2\frac{1}{N}U_F\begin{bmatrix} \hat{S} & 0 \\ 0 & 0 \end{bmatrix}U_F^T - 2\lambda_{p,p}U_FU_F^T = 2C - 2\lambda_{p,p}I_M \tag{22}$$

is strictly positive definite for $p = 1, \cdots, M$. This implies that $\frac{\partial^2}{\partial U_p \partial U_p}L(U, \lambda)$ is strictly positive definite for $p = 1, \cdots, M$. Besides, if $\lambda_{q,p} + \lambda_{p,q} < 0$ for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$, then $\frac{\partial^2}{\partial U_q \partial U_p}L(U, \lambda)$ is strictly positive definite for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$. This completes the proof of the sufficiency part. As a result, this completes the whole proof. $\square$

It is worth noting the conventional principal component analysis employs the columns of U as the eigenvectors of C and $\frac{\lambda+\lambda^T}{2}$ as the diagonal matrix with its diagonal elements being the corresponding eigenvalues. In this case, since $\frac{\hat{S}}{N}$ is strictly positive definite, $\lambda_{p,p} > 0$ for $p = 1, \cdots, M$. As a result, $\frac{\partial^2}{\partial U_p \partial U_p} L(U, \lambda)$ is not strictly positive definite. Also, $\lambda_{q,p} = 0$ for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$, In other words, $\lambda_{q,p} + \lambda_{p,q} = 0$ for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$. As a result, $\frac{\partial^2}{\partial U_q \partial U_p} L(U, \lambda)$ is not strictly positive definite for $p \neq q$, for $p = 1, \cdots, M$ and for $q = 1, \cdots, M$. This implies that the second order derivative condition on the Lagrange function is not satisfied. As a result, the conventional principal component analysis approach can only yield the stationary point of the Lagrange function, but cannot yield the local minima of the Lagrange function.

4. **Computer numerical simulation results.** This Section gives two sets of examples to verify the results derived in Section 3. The input vectors for the first set of examples are generated randomly to illustrate the universality of the developed results. On the other hand, the input vectors for the second set of examples are the wearable non-invasive blood glucose data to illustrate the practicality of the developed results.

4.1. **For the first set of examples.** $M = 4$ and $N = 5$ are chosen in the first case. The set of input vectors is generated randomly. It is found that one of the realizations of $\bar{X}$ is

$$\bar{X} = \begin{bmatrix} -0.0832 & 0.0948 & 0.0012 & 0.2037 & -0.2164 \\ -0.5805 & 0.6595 & 0.0241 & 1.3654 & 1.4685 \\ 0.2193 & -0.2528 & 0.0320 & -0.6574 & 0.6589 \\ -0.2592 & 0.2780 & 0.1980 & -0.0344 & -0.1825 \end{bmatrix}. \tag{23}$$

It can be checked easily that $rank(\bar{X}) = 2$. Hence, $R = 2$. Also, it is found that one of the realizations of U_{X_2} and U_{X_1} are

$$U_{X_2} = \begin{bmatrix} 0.3666 & -0.9208 \\ 0.3471 & 0.2677 \\ 0.8440 & 0.2745 \\ -0.1808 & -0.0715 \end{bmatrix} \tag{24}$$

and

$$U_{X_1} = \begin{bmatrix} 0.1325 & 0.0139 \\ 0.8987 & 0.0130 \\ -0.4646 & -0.2203 \\ 0.1053 & -0.9752 \end{bmatrix}, \tag{25}$$

respectively. It can be checked easily that $\bar{X}^T U_{X_2} = 0$ and $U_{X_2}^T U_{X_1} = 0$. Moreover, it is found that one of the realizations of V_{X_2} and V_{X_2} are

$$V_{X_2} = \begin{bmatrix} 0.4991 & -0.4934 & -0.4443 \\ -0.1240 & -0.7934 & -0.0048 \\ 0.0973 & 0.3175 & -0.8002 \\ -0.4839 & -0.0066 & -0.3621 \\ -0.7013 & -0.1623 & -0.1764 \end{bmatrix} \tag{26}$$

and

$$V_{X_1} = \begin{bmatrix} -0.2663 & 0.4891 \\ 0.3024 & -0.5135 \\ 0.0122 & -0.4993 \\ 0.6225 & 0.4971 \\ -0.6708 & 0.0266 \end{bmatrix}, \qquad (27)$$

respectively. It can be checked easily that $\bar{X}V_{X_2} = 0$ and $V_{X_2}^T V_{X_1} = 0$. Furthermore, it is found that

$$B = \begin{bmatrix} 2.4361 & 0 \\ 0 & 0.4003 \end{bmatrix}. \qquad (28)$$

It can be checked easily that $\bar{X} = U_X \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} V_X^T$. Besides, it is found that one of the realizations of F and S are

$$F = \begin{bmatrix} -0.3221 & 0.0210 \\ -2.1866 & 0.1104 \\ 0.9803 & -0.1354 \\ -0.2750 & -0.3776 \end{bmatrix} \qquad (29)$$

and

$$S = \begin{bmatrix} 0.2895 & -0.3268 & -0.0362 & -0.5979 & 0.6713 \\ 0.4757 & -0.4984 & -0.4981 & 0.5265 & -0.0057 \end{bmatrix}, \qquad (30)$$

respectively. It can be checked easily that $\bar{X} = FS$. In addition, it is found that one of the realizations of U_{F_1} and U_{F_2} are

$$U_{F_2} = \begin{bmatrix} 0.3666 & -0.9208 \\ 0.3471 & 0.2677 \\ 0.8440 & 0.2745 \\ -0.1808 & -0.0715 \end{bmatrix} \qquad (31)$$

and

$$U_{F_1} = \begin{bmatrix} 0.1325 & 0.0139 \\ 0.8987 & 0.0130 \\ -0.4646 & -0.2203 \\ 0.1053 & -0.9752 \end{bmatrix}, \qquad (32)$$

respectively. It can be checked easily that $F^T U_{F_2} = 0$. Also, it is found that

$$D_F = \begin{bmatrix} 2.4361 & 0 \\ 0 & 0.4003 \end{bmatrix} \qquad (33)$$

and

$$V_F = \begin{bmatrix} -0.9988 & 0.0480 \\ 0.0480 & 0.9988 \end{bmatrix}. \qquad (34)$$

It can be checked easily that $F = U_F \begin{bmatrix} D_F \\ 0 \end{bmatrix} V_F^T$. Moreover, it is found that

$$\tilde{S} = \begin{bmatrix} -0.2663 & 0.4891 \\ 0.3024 & -0.5135 \\ 0.0122 & -0.4993 \\ 0.6225 & 0.4971 \\ -0.6708 & 0.0266 \end{bmatrix} \qquad (35)$$

and

$$\hat{S} = \begin{bmatrix} 5.9348 & 0 \\ 0 & 0.1602 \end{bmatrix}. \qquad (36)$$

It can be checked easily that $\tilde{S} = \tilde{S}^T V_F$ and $\hat{S} = D_F \tilde{S}^T \tilde{S} D_F$. Since $U_F = U_X$, $\hat{S} = BB^T$. Furthermore, it is found that

$$C = \begin{bmatrix} 0.0208 & 0.1413 & -0.0637 & 0.0161 \\ 0.1413 & 0.9587 & -0.4317 & 0.1119 \\ -0.0637 & -0.4317 & 0.1959 & -0.0437 \\ 0.0161 & 0.1119 & -0.0437 & 0.0436 \end{bmatrix}. \tag{37}$$

It can be checked easily that $C = \frac{1}{N}\bar{X}\bar{X}^T$. In addition, it is found that

$$\tilde{U}_A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{38}$$

and

$$\Theta_1 = \begin{bmatrix} 1.1870 & 0 \\ 0 & 0.0320 \end{bmatrix}. \tag{39}$$

It can be checked easily that the columns of $\tilde{U}_A$ are the eigenvectors of $\frac{\hat{S}}{N}$ and the diagonal elements of Θ_1 are the corresponding eigenvalues. That is, $\frac{\hat{S}\tilde{U}_A}{N} = \tilde{U}_A \Theta_1$. By arbitrarily generating a random matrix

$$\tilde{U}_B = \begin{bmatrix} -0.9033 & -0.0118 \\ 2.9700 & -1.9995 \end{bmatrix}, \tag{40}$$

we have

$$U = \begin{bmatrix} 0.1325 & 0.0139 & -3.0659 & 1.8368 \\ 0.8987 & 0.0130 & 0.4816 & -0.5394 \\ -0.4046 & -0.2203 & 0.0530 & -0.5589 \\ 0.1053 & -0.9752 & -0.0491 & 0.1452 \end{bmatrix}. \tag{41}$$

It can be checked easily that $\tilde{U} = U_F^T U$ as well as the columns of U are the eigenvectors of C and the diagonal elements of Θ are the corresponding eigenvalues. However, U is not unitary. Besides, one of the realizations of $\hat{V}_{1,1}$ and $\hat{V}_{2,2}$ are chosen as

$$\hat{V}_{1,1} = \begin{bmatrix} -0.9745 & -0.2245 \\ -0.2245 & 0.9745 \end{bmatrix} \tag{42}$$

and

$$\hat{V}_{2,2} = \begin{bmatrix} -0.6353 & 0.7723 \\ 0.7723 & 0.6353 \end{bmatrix}, \tag{43}$$

respectively. Also, it is found that

$$V_\lambda = \begin{bmatrix} -0.9745 & -0.2245 & 0 & 0 \\ -0.2245 & 0.9745 & 0 & 0 \\ 0 & 0 & 2.8675 & 1.1893 \\ 0 & 0 & -1.5367 & -1.2794 \end{bmatrix} \tag{44}$$

and

$$Q_\lambda = \begin{bmatrix} 1.1288 & 0.2526 \\ 0.2526 & 0.0902 \end{bmatrix}. \tag{45}$$

It can be checked easily that $V_\lambda = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} & 0 \\ 0 & \hat{V}_{2,2} \end{bmatrix}$ and $Q_\lambda = \frac{1}{N}\hat{V}_{1,1}^T \hat{S} \hat{V}_{1,1}$. Moreover, it is found that

$$\frac{\lambda + \lambda^T}{2} = \begin{bmatrix} 1.1870 & 0 & 0 & 0 \\ 0 & 0.0320 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{46}$$

It can be checked easily that $\frac{\lambda + \lambda^T}{2} = V_\lambda \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} V_\lambda^T$ and the condition for the first order derivative of the Lagrange function to be equal to the zero vector is satisfied. That is, $2CU = U(\lambda + \lambda^T)$. Since $\lambda_{i,i} > 0$ for $i = 1, \cdots, R$ and $\lambda_{i,j} = 0$ for $i \neq j$, for $i = 1, \cdots, M$ and $j = 1, \cdots, M$, the condition for the second order derivative of the Lagrange function to be strictly positive definite is not satisfied. To versify this result, a small deviation Δ_U is added to U. Now, a perturbed matrix $U + \Delta_U$ is obtained. It is found that the change of the objective functional value is

$$trace\left((U + \Delta_U)^T C (U + \Delta_U)\right) - trace\left(U^T CU\right) = -0.0014 \tag{47}$$

Since it is negative, the new objective functional value drops. This implies that the obtained solution is not a local minimum. It is just a stationary point of the objective function.

Now, consider another realization of $\tilde{U}_B$. We have

$$U = \begin{bmatrix} 0.1325 & 0.0139 & -0.3280 & -0.9352 \\ 0.8987 & 0.0130 & -0.3580 & 0.2531 \\ -0.4046 & -0.2203 & -0.8547 & 0.2392 \\ 0.1053 & -0.9752 & 0.1836 & -0.0640 \end{bmatrix}. \tag{48}$$

It can be checked easily that $\tilde{U} = U_F^T U$ as well as the columns of U are the eigenvectors of C and the diagonal elements of Θ are the corresponding eigenvalues. In this case, U is a unitary matrix. From here, it can be seen that U is not uniquely defined.

4.2. **Consider the second case.** $M = 4$ and $N = 9$ are chosen. The set of input vectors is also generated randomly. It is found that one of the realizations is

$$\bar{X} = \begin{bmatrix} -0.1133 & 0.0107 & 0.2446 & -0.1513 & -0.0008 & -0.0544 & -0.2621 & 0.2277 & 0.0988 \\ -0.1914 & 0.0067 & 0.1273 & -0.0666 & 0.0512 & -0.0058 & -0.1776 & 0.1013 & 0.1549 \\ -0.6955 & -0.0559 & -0.5896 & 0.3345 & 0.3865 & 0.2791 & 0.4774 & -0.6240 & 0.4874 \\ -0.4422 & 0.0140 & 0.1802 & -0.0697 & 0.1387 & 0.0222 & -0.3160 & 0.1175 & 0.3553 \end{bmatrix}. \tag{49}$$

It can be checked easily that $rank(\bar{X}) = 3$. Hence, $R = 3$. Also, it is found that one of the realizations of U_{X_2} and U_{X_1} are

$$U_{X_2} = \begin{bmatrix} -0.3172 \\ 0.9064 \\ -0.0210 \\ -0.2781 \end{bmatrix} \tag{50}$$

and

$$U_{X_1} = \begin{bmatrix} 0.9064 & -0.0210 & -0.2781 \\ 0.3763 & 0.0144 & 0.1914 \\ 0.0144 & 0.9997 & -0.0044 \\ 0.1914 & -0.0044 & 0.9413 \end{bmatrix}, \tag{51}$$

respectively. It can be checked easily that $\bar{X}^T U_{X_2} = 0$ and $U_{X_2}^T U_{X_1} = 0$. Moreover, it is found that one of the realizations of V_{X_2} and V_{X_1} are

$$V_{X_2} = \begin{bmatrix} 0.2304 & 0.1908 & 0.0512 & -0.3386 & 0.0676 & 0.4575 \\ -0.0188 & 0.3516 & 0.1796 & 0.7517 & -0.3353 & 0.3551 \\ 0.0845 & -0.1051 & 0.1731 & -0.2836 & -0.7513 & -0.1153 \\ 0.5681 & -0.2038 & -0.0434 & 0.0593 & -0.1550 & -0.3354 \\ 0.2597 & 0.8460 & -0.0345 & -0.2818 & -0.0354 & -0.1703 \\ -0.0645 & -0.0389 & 0.9513 & -0.1431 & 0.1553 & -0.0170 \\ 0.2993 & -0.1515 & 0.0616 & 0.1008 & -0.3261 & 0.0301 \\ 0.5886 & -0.0315 & 0.1328 & 0.3023 & 0.4024 & -0.1596 \\ 0.3303 & -0.2152 & -0.0749 & -0.1870 & 0.0291 & 0.6951 \end{bmatrix} \tag{52}$$

and

$$V_{X_1} = \begin{bmatrix} -0.1143 & -0.2553 & -0.7049 \\ -0.1877 & -0.0705 & -0.0047 \\ -0.2257 & 0.4742 & -0.1335 \\ -0.3892 & -0.5693 & 0.1352 \\ 0.1195 & 0.0295 & 0.3014 \\ 0.0096 & -0.1390 & 0.1581 \\ 0.8550 & -0.1725 & -0.0747 \\ -0.0045 & 0.5561 & -0.2165 \\ -0.0626 & 0.1470 & 0.5426 \end{bmatrix}, \tag{53}$$

respectively. It can be checked easily that $\bar{X} V_{X_2} = 0$ and $V_{X_2}^T V_{X_1} = 0$. Furthermore, it is found that

$$B = \begin{bmatrix} -0.2876 & 0.5506 & 0.2284 \\ 0.5256 & -0.6804 & 1.1404 \\ -0.1959 & 0.3206 & 0.5068 \end{bmatrix}. \tag{54}$$

It can be checked easily that $\bar{X} = U_X \begin{bmatrix} B & 0 \\ 0 & 0 \end{bmatrix} V_X^T$. However, B is not a diagonal matrix. Besides, it is found that one of the realizations of F and S are

$$F = \begin{bmatrix} 0.2028 & -0.1534 & 0.1623 \\ 0.1711 & -0.0329 & 0.2901 \\ 0.0368 & 0.7109 & 1.1054 \\ 0.3234 & 0.0141 & 0.6768 \end{bmatrix} \tag{55}$$

and

$$S = \begin{bmatrix} -0.0887 & -0.8045 & -0.0075 & -2.4432 & 0.1858 & -0.3312 & 3.0288 & 1.0287 & -0.5681 \\ -0.0246 & -0.6893 & -1.2905 & -1.0935 & 0.3654 & 0.1164 & 3.6081 & -0.4512 & -0.5408 \\ -0.6104 & 0.4195 & 0.2968 & 1.0871 & 0.1085 & 0.1887 & -1.9893 & -0.3085 & 0.8076 \end{bmatrix}, \tag{56}$$

respectively. It can be checked easily that $\bar{X} = FS$. In addition, it is found that one of the realizations of U_{F_1} and U_{F_2} are

$$U_{F_2} = \begin{bmatrix} 0.3172 \\ -0.9064 \\ 0.0210 \\ 0.2781 \end{bmatrix} \tag{57}$$

and

$$U_{F_1} = \begin{bmatrix} -0.0767 & 0.5086 & -0.7968 \\ 0.1844 & 0.3634 & -0.1111 \\ -0.8716 & -0.4494 & -0.1946 \\ -0.4477 & 0.6382 & 0.5612 \end{bmatrix} \tag{58}$$

respectively. It can be checked easily that $F^T U_{F_2} = 0$. Also, it is found that

$$D_F = \begin{bmatrix} 1.4816 & 0 & 0 \\ 0 & 0.5492 & 0 \\ 0 & 0 & 0.0084 \end{bmatrix} \tag{59}$$

and

$$V_F = \begin{bmatrix} -0.1511 & 0.6466 & -0.7477 \\ -0.4104 & -0.7291 & -0.5476 \\ -0.8993 & 0.2241 & 0.3756 \end{bmatrix}. \tag{60}$$

It can be checked easily that $F = U_F \begin{bmatrix} D_F \\ 0 \end{bmatrix} V_F^T$. Moreover, it is found that

$$\tilde{S} = \begin{bmatrix} 0.5724 & -0.1763 & -0.1495 \\ 0.0272 & 0.0764 & 1.1366 \\ 0.2639 & 1.0026 & 0.8239 \\ -0.1596 & -0.5389 & 2.8339 \\ -0.2756 & -0.1220 & -0.2982 \\ -0.1674 & -0.2567 & 0.2547 \\ -0.1497 & -1.1180 & -4.9876 \\ 0.3072 & 0.9250 & -0.6379 \\ -0.4185 & 0.2079 & 1.0243 \end{bmatrix} \tag{61}$$

and

$$\hat{S} = \begin{bmatrix} 1.7988 & 0.5638 & -0.0016 \\ 0.5638 & 1.0745 & 0.0210 \\ -0.0016 & 0.0210 & 0.0026 \end{bmatrix}. \tag{62}$$

It is worth noting that $\hat{S}$ is not a diagonal matrix. It can be checked easily that $\tilde{S} = S^T V_F$ and $\hat{S} = D_F \tilde{S}^T \tilde{S} D_F$. Since $U_F \neq U_X$, $\hat{S} \neq BB^T$. Furthermore, it is found that

$$C = \begin{bmatrix} 0.0254 & 0.0165 & -0.0390 & 0.0276 \\ 0.0165 & 0.0140 & -0.0021 & 0.0269 \\ -0.0390 & -0.0021 & 0.2254 & 0.0207 \\ 0.0276 & 0.0269 & 0.0207 & 0.0547 \end{bmatrix}. \tag{63}$$

It can be checked easily that $C = \frac{1}{N} \bar{X} \bar{X}^T$. In addition, it is found that

$$\tilde{U}_A = \begin{bmatrix} 0.8776 & 0.4793 & 0.0084 \\ 0.4794 & -0.8773 & -0.0240 \\ 0.0041 & -0.0251 & 0.9997 \end{bmatrix} \tag{64}$$

and

$$\Theta_1 = \begin{bmatrix} 0.2341 & 0 & 0 \\ 0 & 0.0852 & 0 \\ 0 & 0 & 0.0002 \end{bmatrix}. \tag{65}$$

It is worth noting that $\tilde{U}_A$ is not a diagonal matrix. It can be checked easily that the columns of $\tilde{U}_A$ are the eigenvectors of $\frac{\hat{S}}{N}$ and the diagonal elements of Θ_1 are the corresponding eigenvalues. That is, $\frac{\hat{S}\tilde{U}_A}{N} = \tilde{U}_A \Theta_1$. By arbitrarily generating a random matrix

$$\tilde{U}_B = -1.9648, \tag{66}$$

we have

$$U = \begin{bmatrix} 0.1732 & -0.4629 & -0.8094 & -0.6233 \\ 0.0120 & -0.4044 & -0.1214 & 1.7809 \\ -0.9812 & -0.0186 & -0.1911 & -0.0412 \\ -0.0846 & -0.7886 & 0.5419 & -0.5465 \end{bmatrix}. \tag{67}$$

It can be checked easily that $\tilde{U} = U_F^T U$ as well as the columns of U are the eigenvectors of C and the diagonal elements of Θ are the corresponding eigenvalues. However, U is not a unitary matrix. Besides, one of the realizations of $\hat{V}_{1,1}$ and $\hat{V}_{2,2}$ are

$$\hat{V}_{1,1} = \begin{bmatrix} -0.9819 & -0.1765 & -0.0681 \\ -0.1826 & 0.9784 & 0.0967 \\ 0.0496 & 0.1074 & -0.9930 \end{bmatrix} \tag{68}$$

and

$$\hat{V}_{2,2} = 1, \tag{69}$$

respectively. Also, it is found that

$$V_\lambda = \begin{bmatrix} -0.9491 & 0.3146 & -0.0175 & 0 \\ -0.3118 & -0.9456 & -0.0926 & 0 \\ 0.0457 & 0.0824 & -0.9956 & 0 \\ 0 & 0 & 0 & -1.9648 \end{bmatrix} \tag{70}$$

and

$$Q_\lambda = \begin{bmatrix} 0.2191 & -0.0448 & 0.0063 \\ -0.0448 & 0.0994 & 0.0062 \\ 0.0063 & 0.0062 & 0.0010 \end{bmatrix}. \tag{71}$$

It can be checked easily that $V_\lambda = \tilde{U}^T \begin{bmatrix} \hat{V}_{1,1} & 0 \\ 0 & \hat{V}_{2,2} \end{bmatrix}$ and $Q_\lambda = \frac{1}{N} \hat{V}_{1,1}^T \hat{S} \hat{V}_{1,1}$. Moreover, it is found that

$$\frac{\lambda + \lambda^T}{2} = \begin{bmatrix} 0.2341 & 0 & 0 & 0 \\ 0 & 0.0852 & 0 & 0 \\ 0 & 0 & 0.0002 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{72}$$

It can be checked easily that $\frac{\lambda + \lambda^T}{2} = V_\lambda \begin{bmatrix} Q_\lambda & 0 \\ 0 & 0 \end{bmatrix} V_\lambda^T$ and the condition for the first order derivative of the Lagrange function to be equal to the zero vector is satisfied. That is, $2CU = U(\lambda + \lambda^T)$. Since $\lambda_{i,i} > 0$ for $i = 1, \cdots, R$ and $\lambda_{i,j} = 0$ for $i \neq j$, for $i = 1, \cdots, M$ and for $j = 1, \cdots, M$, the condition for the second order derivative of the Lagrange function to be strictly positive definite is not satisfied.

4.3. **Consider the other set of examples.** Here, different signals operated at different near infrared frequencies are generated. The waveforms are incident to the blood vessels and the reflected waveforms are acquired using the sensors fixed in a wrist band. As the measured waveforms contain some useful information for performing the blood glucose estimation [11], the training features are extracted from the measured waveforms. In particular, the features are selected as the frequencies, the energies and the heart rate variabilities of the reflected waveforms. Here, there are totally 84 features and 20 samples. Hence, $M = 84$ and $N = 20$. Each feature of these 20 samples is normalized. That is, each row of $\bar{X}$ is normalized to a unit energy vector. The feature values of these samples are shown in Figure 1. The

training outputs are the corresponding blood glucose levels. A machine learning based regression method [13] is employed for performing the estimation.

To improve the training efficiency and the model accuracy, a principal component analysis based dimension reduction algorithm is applied to the feature vectors. It can be checked easily that $rank(\bar{X}) = 20$. That is, $R = 20$. Also, as $C = \frac{1}{N}\bar{X}\bar{X}^T$, $rank(C) = 20$ even though the dimension of C is 84×84. In this case, C drops rank. Here, there exists an $M \times R$ matrix F and there exists an $R \times N$ matrix S such that $\bar{X} = FS$. As described in **Theorem 3.1**, by defining $\tilde{U} = U_F^T U$, we have $\tilde{U} = \begin{bmatrix} \tilde{U}_A & 0 \\ 0 & \tilde{U}_B \end{bmatrix}$ and $\Theta = \begin{bmatrix} \Theta_1 & 0 \\ 0 & 0 \end{bmatrix}$. It can be checked easily that the columns of $\tilde{U}_A$ are the eigenvectors of $\frac{\hat{S}}{N}$ and the diagonal elements of Θ_1 are the corresponding eigenvalues. That is, $\frac{\hat{S}\tilde{U}_A}{N} = \tilde{U}_A\Theta_1$. Also, it is found that U is not uniquely defined. Moreover, U is not a unitary matrix. This is because $\tilde{U}_B$ is not a unitary matrix. Besides, it can be checked easily that the condition for the first order derivative of the Lagrange function to be equal to the zero vector is satisfied. That is, $2CU = U(\lambda + \lambda^T)$. Since $\lambda_{i,i} > 0$ for $i = 1, \cdots, R$ and $\lambda_{i,j} = 0$ for $i \neq j$, for $i = 1, \cdots, M$ and for $j = 1, \cdots, M$, the condition for the second order derivative of the Lagrange function to be strictly positive definite is not satisfied.
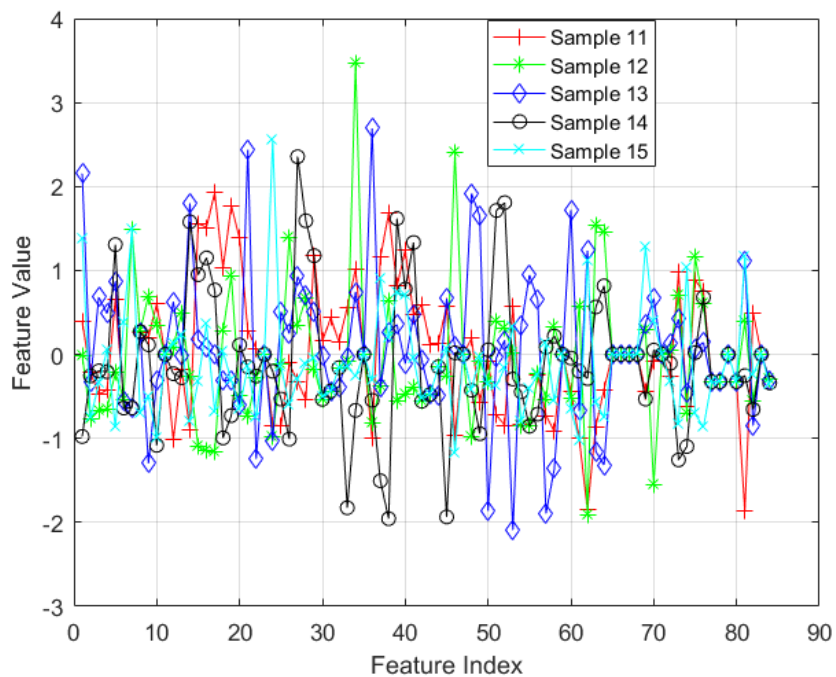
5. **Conclusions.** When the covariance matrix of the input vectors drops rank, it is found that the eigen decomposition of the covariance matrix is not uniquely defined. Also, the matrix with its columns being the eigenvectors of the covariance matrix
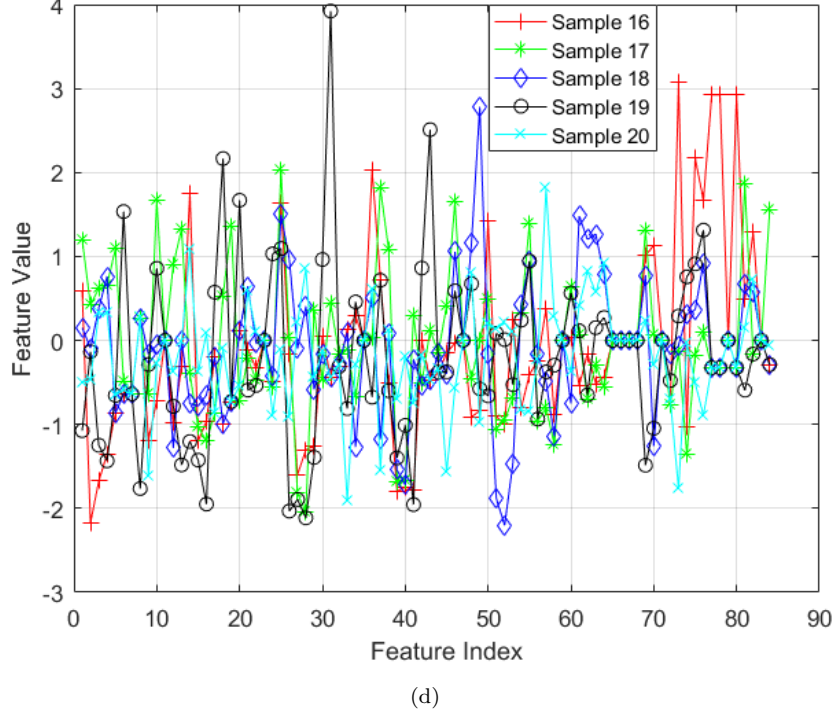

(a)

(b)

(c)

(d)
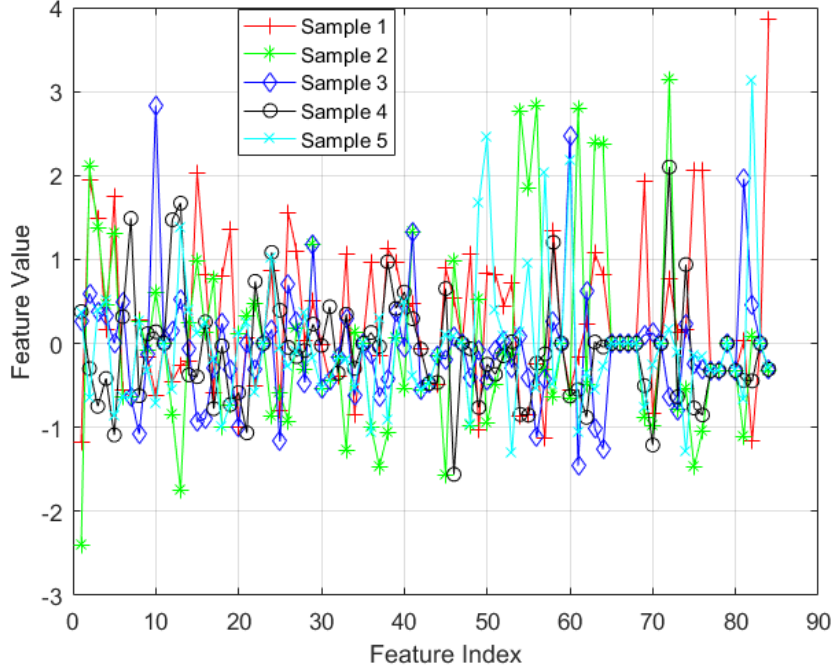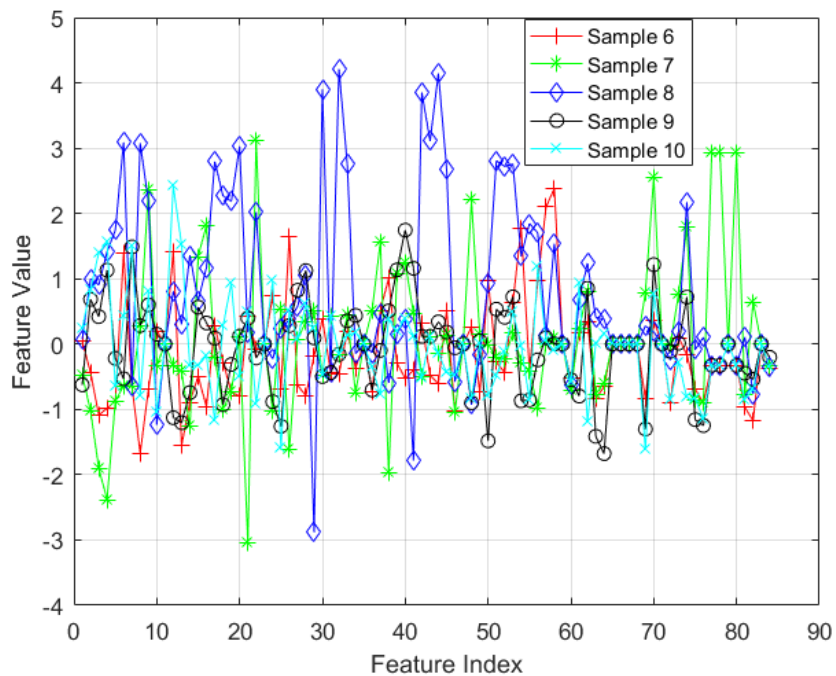
FIGURE 1. Plots of the feature values. (a) Sample I.D.: 1-5. (b) Sample I.D.: 6-10. (c) Sample I.D.: 11-15. (d) Sample I.D.: 16-20.

is not necessary to be a unitary matrix. Moreover, the unitary matrix with its columns being the eigenvectors of the covariance matrix is just a particular solution for the first order derivative of the Lagrange function to be equal to the zero vector. Furthermore, the unitary matrix with its columns being the eigenvectors of the covariance matrix does not satisfy the condition for the second order derivative of the Lagrange function to be strictly positive definite. Computer numerical simulation results are given to verify the validity of the results.

REFERENCES

[1] J. Dehaene, M. Moonen and J. Vandewalle, An improved stochastic gradient algorithm for principal component analysis and subspace tracking, *IEEE Transactions on Signal Processing*, **45** (1997), 2582–2586.

[2] C. L. Fancourt and J. C. Principe, Competitive principal component analysis for locally stationary time series, *IEEE Transactions on Signal Processing*, **46** (1998), 3068–3081.

[3] J. B. O. S. Filho and P. S. R. Diniz, A fixed-point online kernel principal component extraction algorithm, *IEEE Transactions on Signal Processing*, **65** (2017), 6244–6259.

[4] I. A. Guimarães and A. C. Neto, Estimation in polytomous logistic model: Comparison of methods, Journal of Industrial and Management Optimization, *Journal of Industrial and Management Optimization*, **5** (2009), 239–252.

[5] R. He, B. G. Hu, W. S. Zheng and X. W. Kong, Robust principal component analysis based on maximum correntropy criterion, *IEEE Transactions on Image Processing*, **20** (2011), 1485–1494.

[6] S. M. Huang and J. F. Yang, Improved principal component regression for face recognition under illumination variations, *IEEE Signal Processing Letters*, **19** (2012), 179–182.

[7] J. Kang, X. Lin and G. Yang, Research of multi-scale PCA algorithm for face recognition, *International Conference on Information and Communications Technologies,ICT*, (2015), 1–5.

[8] M. S. Kang, J. H. Bae, B. S. Kang and K. T. Kim, ISAR cross-range scaling using iterative processing via principal component analysis and bisection algorithm, *IEEE Transactions on Signal Processing*, **64** (2016), 3909–3918.

[9] S. Karamizadeh, S. M. Abdullah, A. A. Manaf and M. Zamani, An overview of principal component analysis, *Journal of Signal and Information Processing*, **4** (2013), 173–175.

[10] H. C. Liu, S. T. Li and L. Y. Fang, Robust object tracking based on principal component analysis and local sparse representation, *IEEE Transactions on Instrumentation and Measurement*, **64** (2015), 2863–2875.

[11] Z. Li, Q. L. Ye, Y. T. Guo, Z. K. Tian, B. W. K. Ling and R. W. K. Lam, Wearable non-invasive blood glucose estimation via empirical mode decomposition based hierarchical multiresolution analysis and random forest, *23rd IEEE International Conference on Digital Signal Processing,DSP*, (2018), 1–5.

[12] A. Moradi, J. Razmi, R. Babazadeh and A. Sabbaghnia, An integrated Principal Component Analysis and multi-objective mathematical programming approach to agile supply chain network design under uncertainty, *Journal of Industrial and Management Optimization*, **15** (2019), 855–879.

[13] E. M. Moreno, Non-invasive estimate of blood glucose and blood pressure from a photoplethysmography by means of machine learning techniques, *Artificial Intelligence in Medicine*, **53** (2011), 127–138.

[14] P. P. Markopulos, G. N. Karystinos and D. A. Pados, Optimal algorithms for L1-subpace signal processing, *IEEE Transactions on Signal Processing*, **62** (2014), 5046–5058.

[15] O. Ozgonenel and T. Yalcin, Principal component analysis (PCA) based neural network for motor protection, *10th IET International Conference on Developments in Power System Protection,DPSP*, (2010), 1–5.

[16] S. Ouyang and Z. Bao, Fast principal component extraction by a weighted information criterion, *IEEE Transactions on Signal Processing*, **50** (2002), 1994–2002.

[17] Y. M. Ren, L. Liao and S. J. Maybank, Hyperspectral image spectral-spatial feature extraction via tensor principal component analysis, *IEEE Geoscience and Remote Sensing Letters*, **14** (2017), 1431–1435.

[18] C. E. Thomaz and G. A. Giraldi, A new ranking method for principal components analysis and its application to face image analysis, *Image and Vision Computing*, **28** (2009), 902–913.

[19] G. Tang and A. Nehorai, Constrained Cramér-Rao bound on robust principal component analysis, *IEEE Transactions on Signal Processing*, **59** (2011), 5070–5076.

[20] X.C Xiu, Y. Yang, W.Q. Liu, L.C. Kong and M.J. Shang, An improved total variation regularized RPCA for moving object detection with dynamic background, *Journal of Industrial and Management Optimization*, **13** (2017), 1–14.

[21] S. Y. Yi, Z. H. Lai, Z. Y. He, Y. Cheung and Y. Liu, Joint sparse principal component analysis, *Pattern Recognition*, **61** (2016), 524–536.

[22] A. Zare, A. Ozdemir, M. A. Iwen and S. Aviyente, Extension of PCA to higher order data structures: An introduction to tensors, tensor decompositions, and tensor PCA, *Proceedings of the IEEE*, **106** (2018), 1341–1358.

[23] H. Zou and L. Xue, A selective overview of sparse principal component analysis, *Proceedings of the IEEE*, **106** (2018), 1311–1320.

E-mail address: 614041771@qq.com,yitongguo007@gmail.com
E-mail address: yongquanling@gdut.edu.cn