A NEW METHOD FOR EIGENSTRUCTURES EXTRACTION
AND ITS NEURAL NETWORKS IMPLEMENTATION

Yu Shuijun  Liang Diannong
(National University of Defense Technology, Changsha 410073)

Abstract The cost function for eigenstructures extraction is discussed in detail in this paper, one can obtain the largest eigenvector by minimizing the cost function. In order to obtain other eigenvectors, a covariance matrix series is constructed. If one compares the cost function with the energy function of a neural networks, the neural networks can be easily introduced to extract the eigenvectors. Theoretical analysis and computer simulations show that the proposed method is reasonable and feasible.

Key words Eigenstructure; Cost function; Neural networks

I. Introduction

Algorithms based on eigenstructures of a data covariance matrix are being widely used in signal processing. However, due to the complexity and heavy computation burden for extracting the eigenstructures of a covariance matrix, the real applications of the algorithms are limited up to now. In order to decrease the computation burden and reduce its complexity, many researchers have been working for a long time and many methods have been presented. All the proposed methods can be divided into two classes: decomposition methods(DEM) and energy function methods(EMF) or cost function methods. EMF is a new method presented in the latest decades with a prospective future. The existing optimal methods[1,2] and neural networks methods[3,4] belong to the area. At present, the optimal methods focus mainly on theoretical study, the neural networks methods do not construct a proper cost function to describe the problem.

In this paper, a cost function expression is presented for extracting the largest eigenvector of a data covariance matrix $R$, the optimal solution vector of the cost function is corresponding to the largest eigenvector of $R$. For extracting the other eigenvectors, a covariance series matrix is constructed according to the idea that the largest eigenvector of the $j$-th matrix is equal to the $j$-th largest eigenvector of $R$. Having presented the cost function, we introduce a neural networks to solve the presented minimize problem with the energy function of a neural network as a bridge. Theoretical analysis and simulations show that the presented method is valid.

II. Cost Function Expression for Eigenstructures Extraction

Assume $R$ is a $N \times N$ real data covariance matrix, its $N$ eigenvalues satisfy $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ and their corresponding eigenvectors are $q_1, q_2, \cdots, q_N$, $W$ is a $N \times 1$ column
vector.

Cost functions for extracting the eigenvectors of $R$ are constructed as follows in Refs. [5,6]:

(1) computation begins from extracting the smallest eigenvector:

$$J(W_j, \mu) = \sum_{k=1}^{\mu} W_k^T R_k W_j + \mu (W_j^T W_j - 1)^2,$$

where $\mu > \text{trace}(R)/2$, $\beta > \text{trace}(R)$, $L$ is the number of the extracted smallest eigenvectors.

(2) computation begins from extracting the largest eigenvector:

$$J(W_j, \mu) = \sum_{k=1}^{\mu} W_k^T R_k^{-1} W_j + \mu (W_j^T W_j - 1)^2,$$

where $\mu > \text{trace}(R^{-1})/2$, $\beta > \text{trace}(R^{-1})$, $L$ is the number of the extracted largest eigenvectors, $W_k^*$ is the optimal solution vector by minimizing $J(W_k, \mu)$.

From Eqs.(1) and (2), it is obvious that the two unavoidable problems exist if one wants to extract the largest eigenvector of $R$: (a) if the cost function described in Eq.(1) is used, there is a redundant computation. (b) if the cost function described in Eq.(2) is used, it needs to calculate the inverse of $R$. In order to solve these problems, a new cost function is constructed for extracting the largest eigenvector:

$$J(W, \mu) = -2W^T RW + W^T RWW^T W + \mu (W^T W - 1)^2$$

**Theorem 1** The largest eigenvector of $R$ is the global minimizer of Eq.(3), and the other eigenvectors of $R$ are the saddle point of Eq.(3).

**Proof:** Let $W = \sum_{i=1}^{N} a_i q_i$, $\sum_{i=1}^{N} a_i^2 = b^2 > 0$, and substitute these in Eq.(3), there is

$$J(W, \mu) = -2 \sum_{i=1}^{N} a_i^2 \lambda_i + b^2 \sum_{j=1}^{N} a_j^2 \lambda_j + b^4 \mu - 2b^2 \mu + \mu$$

In Eq.(4), $R$ is a positive definite matrix and $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, so there is: (a) while $b^2 - 2 > 0$, if and only if $a_1^2 = a_2^2 = \cdots = a_{N-1}^2 = 0, a_N^2 = b^2$, Eq.(3) can reach its minimum; (b) while $b^2 - 2 < 0$, if and only if $a_2^2 = \cdots = a_{N-1}^2 = a_N^2 = 0, a_1^2 = b^2$, Eq.(3) can reach its minimum. In the case of (a) or (b), one can obtain $q_N$ or $q_1$ respectively by minimizing Eq.(4). Therefore, it is necessary to determine the quantity of $b^2$ before deciding whether the optimal solution vector of Eq.(4) is corresponding to the largest eigenvector of $R$ or not.

Let's assume $b^2 - 2 > 0$, then Eq.(3) can be rewritten as follows:

$$J(W, \mu) = (b^2 - 2)b^2 \lambda_N + b^4 \mu - 2b^2 \mu + \mu$$

Differentiate Eq.(5) with respect to $b$ and let it be equal to zero, that is

$$-4b \lambda_N + 4b^3 \lambda_N + 4b^3 \mu - 4b \mu = 0$$