J. Math. Study
doi: 10.4208/jms.v48n1.15.05

Vol. **48**, No. 1, pp. 66-78
March 2015

Numerical methods for the Maxnear criterion of multiple-sets canonical analysis

Xin-Guo Liu, Jian-Ping You*

School of Mathematical Sciences, Ocean University of China, Qingdao 266003, Shandong, P.R. China.

Abstract. This paper deals with numerical methods for the Maxnear criterion of multiple-sets canonical analysis. Optimality conditions are derived. Upper and lower bounds of the optimal objective function value are presented. Two iterative methods are proposed. One is an alternating variable method, and the other called Gauss-Seidel method is an inexact version of the alternating variable method. Convergence of these methods are analyzed. A starting point strategy is suggested for both methods. Numerical results are presented to demonstrate the efficiency of these methods and the starting point strategy.

AMS subject classifications: 62H25, 65KO5

Key words: Multiple-sets canonical analysis, Maxnear criterion, alternating variable method, starting point strategy.

1 Introduction

Since Hotelling [3,4] proposed canonical correlation analysis (CCA) as the method for describing the relation between the scores of a set of observation units on two groups of variables, CCA has become an important method in multivariate statistics. It has been widely applied in the econometrics, signal processing, biology, artificial intelligence, and other fields. Several generalizations of canonical correlation analysis for multiple-sets have been proposed by Kettenring [6], Van de Geer [7], Hanafi and Kiers [1] and other scholars. In this paper, we shall concern ourselves with the Maxnear criterion proposed by Van de Geer [7], which can be introduced briefly as follows.

Let $y_i=(y_{i,1},\ldots,y_{i,n_i})^T, i=1,\ldots,m$ be m-sets of random variables. Considering $z_i(t)=t_i^T y_i, t_i\in \boldsymbol{R}^{n_i}$, which is the linear combination of $y_{i,1},\ldots,y_{i,n_i}$, the basic idea of canonical correlation analysis is finding $t_1,\ldots,t_m$ so as to optimize some functions of correlations or

*Corresponding author. *Email addresses:* liuxinguo656@sina.com (X.-G. Liu), you_jian_ping@163.com (J.-P. You)

covariances of $z_1(t),\ldots,z_m(t)$. Therefore, given the covariance matrix A of $y=(y_1,\ldots,y_m)^T$, partitioned as

$$A=(A_{ij})_{m\times m}, A_{ii}\in \boldsymbol{R}^{n_i\times n_i},$$

where A_{ii} is the covariance matrix of y_i, and $A_{ij}(i\neq j)$ is the covariance matrix between y_i and y_j . Suppose A is symmetric and positive definite in the following, and let

$$n=n_1+\ldots+n_m,\ \ D=diag(A_{11},\ldots,A_{mm}).$$

The Maxnear criterion can be described as the following optimization problem:

$$\min\ x^T(mD-A)x,\ \ s.t.\ x\in\Sigma_m, \tag{1.1}$$

$$\Sigma_m=\left\{ x=\begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix}\in \boldsymbol{R}^n : x_i\in\boldsymbol{R}^{n_i}, \| x_i \|_2=1 \right\}.$$

Next, we briefly present a statistical property of Maxnear. Because the matrix A is symmetric and positive definite, it can be factorized as follows:

$$A=P^TP,\ \ P=[P_1,\ldots,P_m],\ \ P_j\in\boldsymbol{R}^{n\times n_j}.$$

Noting that

$$\begin{aligned} Var(y_i^Tx_i-y_j^Tx_j) &= Var(y_i^Tx_i)+Var(y_j^Tx_j)-2cov(y_i^Tx_i,y_j^Tx_j) \\ &= x_i^TA_{ii}x_i+x_j^TA_{jj}x_j-2x_i^TA_{ij}x_j, \end{aligned}$$

adding them up, we have

$$\begin{aligned} &\sum_{i,j=1}^m Var(y_i^Tx_i-y_j^Tx_j) \\ &=2mx^TDx-2x^TAx=2x^T(mD-A)x. \end{aligned}$$

Hence, the Maxnear is equivalent to the following optimization problem:

$$\min\sum_{i,j=1}^m Var(y_i^Tx_i-y_j^Tx_j),\ \ s.t.\ x\in\Sigma_m. \tag{1.2}$$

In this paper, we mainly concentrate on developing efficient algorithm for Maxnear. In fact, lacking of efficient methods is one obstacle of applying Maxnear in practice. All general-purpose optimization algorithms applying to (1.1) are mainly centered around satisfying the first-order necessary condition(see Theorem 2.1 below). Without the global minimizer, the canonical correlation would not be established, making the statistical prediction less reliable. For general $m\geq 2$, several remarks about (1.1) are in order.