Linear Algebra and its Applications 461 (2014) 1–17

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

On the generalized low rank approximation of the correlation matrices arising in the asset portfolio ☆

Xuefeng Duan *, Jianchao Bai, Maojun Zhang, Xinjun Zhang

College of Mathematics and Computational Science, Guilin University of Electronic Technology, Guilin 541004, PR China

ARTICLE INFO

Article history:
Received 12 October 2013
Accepted 21 July 2014
Available online 13 August 2014
Submitted by R. Brualdi

MSC:
11D07
68W25
65F30

Keywords:
Generalized low rank approximation
Correlation matrix
Asset portfolio
Feasible set
Conjugate gradient algorithm

ABSTRACT

In this paper, we consider the generalized low rank approximation of the correlation matrices problem which arises in the asset portfolio. We first characterize the feasible set by using the Gramian representation together with a special trigonometric function transform, and then transform the generalized low rank approximation of the correlation matrices problem into an unconstrained optimization problem. Finally, we use the conjugate gradient algorithm with the strong Wolfe line search to solve the unconstrained optimization problem. Numerical examples show that our new method is feasible and effective.

Published by Elsevier Inc.

☆ The work was supported by National Natural Science Foundation of China (Nos. 11101100; 11261014; 11301107; 61362021), Natural Science Foundation of Guangxi Province (No. 2012GXNSFBA053006; 2013GXNSFBA019009; 2013GXNSFBB053005; 2013GXNSFDA019030), the Fund for Guangxi Experiment Center of Information Science (20130103), Innovation Project of GUET Graduate Education (GDYCSZ201473), Innovation Project of Guangxi Graduate Education (YCSZ2014137), and Guangxi Key Lab of Wireless Wideband Communication and Signal Processing open grant 2012.

* Corresponding author.
E-mail addresses: duanxuefenghd@aliyun.com (X. Duan), baijianchaok@126.com (J. Bai).

http://dx.doi.org/10.1016/j.laa.2014.07.026
0024-3795/Published by Elsevier Inc.

1. Introduction

Throughout this paper, we use $R^{n\times n}$ and S_n^+ to denote the set of $n\times n$ real matrices and symmetric positive semidefinite matrices, respectively. We use A^T and $\mathrm{tr}(A)$ to represent the transpose and trace of the matrix A, respectively. The symbols $\|A\|_F$ and $\mathrm{rank}(A)$ denote the Frobenius norm and the rank of the matrix A, respectively. The symbol $\mathrm{diag}(Y)$ stands for the vector whose elements lie in the diagonal line of the matrix Y, and the symbol e stands for the vector whose elements are of all ones, i.e., $e=(1,1,\cdots,1)^T$.

In this paper, we consider the following problem named **generalized low rank approximation of the correlation matrices**.

Problem 1.1. Given some correlation matrices $A^{(d)}\in R^{n\times n}$, $d=1,2,\cdots,m$, and a positive integer k, $1\le k<n$, find a correlation matrix $\widehat{Y}$ whose rank is less than and equal to k such that

$$\frac{1}{2}\sum_{d=1}^{m}\big\|A^{(d)}-\widehat{Y}\big\|_F^2=\min_{Y\in S_n^+,\,\mathrm{diag}(Y)=e,\,\mathrm{rank}(Y)\le k}\frac{1}{2}\sum_{d=1}^{m}\big\|A^{(d)}-Y\big\|_F^2. \tag{1.1}$$

Problem (1.1) arises in the asset portfolio (see [10] for more details), which can be stated as follows. Suppose that $R=DCD$ is the covariance matrix of n assets, where C is a correlation matrix and D is a diagonal matrix with positive variances which are specially used to describe the risk of assets. In practice, the covariance matrix is usually estimated by the historical data of the return of each asset, that is, an approximation covariance is obtained by statistics method. Let

$$R^{(d)}=D^{(d)}C^{(d)}D^{(d)}$$

be the approximation covariance with dth sampling some data, where $D^{(d)}$ and $C^{(d)}$ are the dth approximation diagonal matrix and correlation matrix, respectively. Higham [4] proposed a method for finding the nearest low rank approximation of a correlation matrix by only one sampling (i.e., $m=1$). However, it is difficult for the decision maker to choose the best approximation covariance matrix with only one sampling because there is always a noise in the data on the prices of assets. Thus, we develop a repeated sampling method to get a series of approximation covariance matrices, that is, d comes from 1 to m. Obviously, it is very easy to obtain the optimal diagonal matrix $\widehat{D}$ by a series of $D^{(d)}$. The major obstacle to finding the optimal covariance matrix is conducting the optimal correlation matrix $\widehat{C}$ from a series of $C^{(d)}$. The above consideration leads to solving the following problem: given some correlation matrices $A^{(1)},A^{(2)},\cdots,A^{(m)}\in R^{n\times n}$, find a correlation matrix $\widehat{Y}$ such that

$$\frac{1}{2}\sum_{d=1}^{m}\big\|A^{(d)}-\widehat{Y}\big\|_F^2=\min_{Y\in S_n^+,\,\mathrm{diag}(Y)=e}\frac{1}{2}\sum_{d=1}^{m}\big\|A^{(d)}-Y\big\|_F^2. \tag{1.2}$$