Mathematics

http://www.aimspress.com/journal/Math

AIMS Mathematics, 3(1): 253–262
DOI:10.3934/Math.2018.1.253
Received: 28 June 2017
Accepted: 07 April 2018
Published: 17 April 2018

Research article

A SOR-like AVM for the maximal correlation problem

Jun He*

School of mathematics, Zunyi Normal College, Zunyi, Guizhou, 563006, P. R. China

* **Correspondence:** hejunfan1@163.com.

Abstract: In this paper, a SOR-like alternating variable method for computing the global solution of the maximal correlation problem is presented. The monotone convergence of the SOR-like alternating variable method is proved. Numerical experiments show the efficiency of our method.

Keywords: multivariate statistics; maximal correlation problem; convergence; multivariate eigenvalue problem; global maximizer; power method
Mathematics Subject Classification: 62H20, 65K05, 15A12, 65F10

1. Introduction

Canonical correlation analysis plays an important role for assessing the relationship between sets of variables. In order to find the linear combination of the set of variables, the maximal correlation problem (MCP) is proposed by Hotelling [1, 2]. When we consider the optimal solution to the MCP, which then corresponds to the global maximizer of the following equality constrained optimization problem:

$$\begin{cases} \text{maximize} & r(x) := x^T A x, \\ \text{subject to} & \|x_i\|_2 = 1, \ i = 1, 2, \cdots, m, \end{cases} \tag{1.1}$$

where $x = \left[x_1^T, \cdots, x_m^T\right] \in \mathbb{R}^n$, $x_i \in \mathbb{R}^{n_i}$, $A_{ij} \in \mathbb{R}^{n_i \times n_j}$ and

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{pmatrix} \in \mathbb{R}^{n \times n}$$

is a symmetric and positive definite matrix.

Using the method of Lagrange multipliers, we will get the following equations:

$$\begin{cases} Ax = \Lambda x, \\ \|x_i\|_2 = 1, \ i = 1, 2, \cdots, m. \end{cases} \tag{1.2}$$

where

$$\Lambda := \operatorname{diag}\left\{ \ \lambda_1 I^{[n_1]}, \quad \lambda_2 I^{[n_2]}, \quad \cdots, \quad \lambda_m I^{[n_m]} \ \right\}$$

with $\lambda_1, \ldots, \lambda_m$ as the Lagrange multipliers and $I^{[n_i]} \in \mathbb{R}^{n_i \times n_i}$ denoting the $n_i \times n_i$ identity matrix. The multivariate eigenvalue problem (MEP) is cast exactly as the system of equations (1.2), which serves only as a necessary condition for the global maximum of the MCP [4]. $\lambda_1, \ldots, \lambda_m$ are usually called the multivariate eigenvalues.

The Horst-Jacobi algorithm as a generalization of the so called power method is introduced to solve MEP in [5], its convergence theory was established much later in [3]. The Gauss-Seidel algorithm has also been suggested in [3] and its monotone convergence is recently established in [7]. Based on a core engine in seeking global maximum of the MCP, the authors proposed an alternating variable method (AVM) in [9]. This algorithm is proved to enjoy the global and monotone convergence and it is also shown that for a nonnegative irreducible matrix A, the algorithm converges to the global maximizer from any nonnegative starting point [8, 9].

In this paper, we propose the SOR-like AVM and the monotone convergence of the SOR-like AVM is proved.

2. Main results

If we express the objective function $r_i(x)$ as

$$r_i(x) = \left(x_i^T A_{ii} x_i + 2x_i^T \sum_{j \neq i} A_{ij} x_j \right) + \sum_{k \neq i} \sum_{j \neq i} x_k^T A_{kj} x_j,$$

It is very clear that if $x^* = \left[(x_1^*)^T, \cdots, (x_m^*)^T\right]^T \in \mathbb{R}^n$, is a global maximizer of the MCP (1.1), $x_i^* \in \mathbb{R}^{n_i}$ must be a global maximizer for the following subproblem

$$\max_{\|x_i\|_2 = 1} r_i(x_1^*, \cdots, x_{i-1}^*, x_i, x_{i+1}^*, \cdots, x_m^*). \tag{2.1}$$

Then we can get the following AVM algorithm, which have shown its efficient in [9].

Algorithm 1 (The framework of the alternating variable method (AVM)).
Select $x^{(0)} \in \mathcal{M}$, and set $k := 0$.
while the stoping criterion is not met **do**
for $i = 1, \cdots, m$ **do**
$x_i^{(k+1)} := \arg\max_{\|x_i\|_2 = 1} r_i(x_1^{(k+1)}, \cdots, x_{i-1}^{(k+1)}, x_i^{(k)}, x_{i+1}^{(k)}, \cdots, x_m^{(k)})$
end for
$k := k + 1$
end while

Denote

$$
\begin{aligned}
x^{(k+1)} &: = \left[(x_1^{(k+1)})^T, \cdots, (x_{i-1}^{(k+1)})^T, (x_i^{(k+1)})^T, \cdots, (x_m^{(k+1)})^T\right]^T \in \mathbb{R}^n, \\
g_i^{(k)} &: = \sum_{j=1}^{i-1} A_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{m} A_{ij} x_j^{(k)} \in \mathbb{R}^{n_i}, \\
m^{(k)}(x_i) &: = r_i\left(x_1^{(k+1)}, \cdots x_{i-1}^{(k+1)}, x_i^{(k)}, x_{i+1}^{(k)}, \cdots, x_m^{(k)}\right) \in \mathbb{R}, \\
c_i^{(k)} &: = m^{(k)}(x_i) - (x_i^{(k)})^T A_{ii} x_i^{(k)} - 2(x_i^{(k)})^T g_i^{(k)} \in \mathbb{R}.
\end{aligned} \tag{2.2}
$$

Then, the ith inner-loop iteration of Algorithm 1 at the kth outer-loop iteration is then to solve the following subproblem:

$$x_i^{(k+1)} := \arg \max_{\|x_i\|_2=1} m^{(k)}(x_i) = (x_i^{(k)})^T A_{ii} x_i^{(k)} + 2(x_i^{(k)})^T g_i^{(k)} + c_i^{(k)}.$$

Let $\lambda_i^{(k+1)} = (x_i^{(k+1)})^T A_i x^{(k+1)}$, we can get the necessary and sufficient condition for the global solution $x_i^{(k+1)}$:

$$\left(-A_{ii} + \lambda_i^{(k+1)} I^{[n_i]}\right) x_i^{(k+1)} = g_i^{(k)}, \tag{2.3}$$

$$\left\|x_i^{(k+1)}\right\|_2 = 1, \tag{2.4}$$

$$\left(-A_{ii} + \lambda_i^{(k+1)} I^{[n_i]}\right) \quad \text{is positive semidefinite.} \tag{2.5}$$

If we split A as

$$A = D + U + U^T, \tag{2.6}$$

where U is the strictly block triangular matrix of A, and

$$D = \operatorname{diag}\{A_{11}, \cdots, A_{mm} \in \mathbb{R}^{n \times n}\} \tag{2.7}$$

is the block diagonal matrix of A. Then Algorithm 1 may be written in matrix form:

$$\Lambda^{(k+1)} x^{(k+1)} = (D + U^T) x^{(k+1)} + U x^{(k)}. \tag{2.8}$$

We give the following SOR-like AVM.

Algorithm 2 (SOR-like AVM).
Select $x^{(0)} \in \mathcal{M}$, and set $k := 0$.
while the stoping criterion is not met **do**
for $i = 1, \cdots, m$ **do**
$y_i^{(k+1)} := \arg\max_{\|x_i\|_2=1} r_i(x_1^{(k+1)}, \cdots, x_{i-1}^{(k+1)}, x_i^{(k)}, x_{i+1}^{(k)}, \cdots, x_m^{(k)})$,
$\bar{x}_i^{(k+1)} := \omega_i y_i^{(k+1)} + (1 - \omega_i) x_i^{(k)}$,
$x_i^{(k+1)} := \frac{\bar{x}_i^{(k+1)}}{\left\|\bar{x}_i^{(k+1)}\right\|}$.
end for
$k := k + 1$
end while

From the Algorithm 2, let $\xi_i^{(k)} = \left\|\bar{x}_i^{(k+1)}\right\|$, we have

$$\lambda_i^{(k+1)} y_i^{(k+1)} = \sum_{j=1}^{i} A_{ij} x_j^{(k+1)} + \sum_{j=i+1}^{m} A_{ij} x_j^{(k)},$$

and

$$\xi_i^{(k)} x_i^{(k+1)} = (1 - \omega_i^{(k)}) x_i^{(k)} + \omega_i^{(k)} y_i^{(k+1)}. \tag{2.9}$$

Then, we get

$$\lambda_i^{(k+1)} \xi_i^{(k)} x_i^{(k+1)} = \lambda_i^{(k+1)} (1 - \omega_i^{(k)}) x_i^{(k)} + \omega_i^{(k)} \sum_{j=1}^{i} A_{ij} x_j^{(k+1)} + \omega_i^{(k)} \sum_{j=i+1}^{m} A_{ij} x_j^{(k)}. \tag{2.10}$$

If we define

$$\Xi^{(k)} := \mathrm{diag}\left\{\xi_1^{(k)} I^{[n_1]}, \cdots, \xi_m^{(k)} I^{[n_m]}\right\}, \tag{2.11}$$

and

$$\Omega^{(k)} := \mathrm{diag}\left\{\omega_1^{(k)} I^{[n_1]}, \cdots, \omega_m^{(k)} I^{[n_m]}\right\}, \tag{2.12}$$

then Algorithm 2 may be written in matrix form:

$$\left[\Lambda^{(k+1)} \Xi^{(k)} - \Omega^{(k)} (U^T + D)\right] x^{(k+1)} = \left[(I - \Omega^{(k)}) \Lambda^{(k+1)} + \Omega^{(k)} U\right] x^{(k)}. \tag{2.13}$$

Then, we can get

$$\begin{aligned} r(x^{(k+1)}) &= (x^{(k+1)})^T A x^{(k+1)} \\ &= (x^{(k+1)})^T (D + U + U^T) x^{(k+1)} \\ &= (x^{(k+1)})^T \left[U + (\Omega^{(k)})^{-1} \Lambda^{(k+1)} \Xi^{(k)} - (\Omega^{(k)})^{-1} \left(\Lambda^{(k+1)} \Xi^{(k)} - \Omega^{(k)} (U^T + D)\right)\right] x^{(k+1)} \\ &= (x^{(k+1)})^T \left[U + (\Omega^{(k)})^{-1} \Lambda^{(k+1)} \Xi^{(k)}\right] x^{(k+1)} \\ &\quad - (x^{(k+1)})^T (\Omega^{(k)})^{-1} \left[(I - \Omega^{(k)}) \Lambda^{(k+1)} + \Omega^{(k)} U\right] x^{(k)}, \end{aligned} \tag{2.14}$$

and

$$\begin{aligned} r(x^{(k)}) &= (x^{(k)})^T A x^{(k)} \\ &= (x^{(k)})^T (D + U + U^T) x^{(k)} \\ &= (x^{(k)})^T \left[\left((I - \Omega^{(k)}) \Lambda^{(k+1)} + \Omega^{(k)} U\right) (\Omega^{(k)})^{-1} - (I - \Omega^{(k)}) \Lambda^{(k+1)} (\Omega^{(k)})^{-1} + U^T + D\right] x^{(k)} \\ &= (x^{(k+1)})^T \left[\Lambda^{(k+1)} \Xi^{(k)} - \Omega^{(k)} (U + D)\right] (\Omega^{(k)})^{-1} x^{(k)} \\ &\quad - (x^{(k)})^T \left[-(I - \Omega^{(k)}) \Lambda^{(k+1)} (\Omega^{(k)})^{-1} + U^T + D\right] x^{(k)}. \end{aligned} \tag{2.15}$$

Hence, by (2.14) and (2.15),

$$\begin{aligned} \Delta r(x^{(k)}) &= r(x^{(k+1)}) - r(x^{(k)}) \\ &= (x^{(k+1)})^T U x^{(k+1)} - (x^{(k)})^T U x^{(k)} + (x^{(k+1)})^T D x^{(k)} - (x^{(k)})^T D x^{(k)} \\ &\quad + \sum_{i=1}^{m} \frac{\lambda_i^{(k+1)}}{\omega_i^{(k)}} (\xi_i^{(k)} + 1 - \omega_i^{(k)}) (1 - (x_i^{(k+1)})^T x_i^{(k)}) \end{aligned} \tag{2.16}$$

$$
\begin{aligned}
&= \frac{1}{2}\left[r(x^{(k+1)}) - r(x^{(k)})\right] - \frac{1}{2}(s^{(k)})^T D s^{(k)} \\
&+ \frac{1}{2}(s^{(k)})^T \left[(\Omega^{(k)})^{-1}\Lambda^{(k+1)}(\Xi^{(k)} + I - \Omega^{(k)})\right] s^{(k)},
\end{aligned} \tag{2.17}
$$

where $s^{(k)} := x^{(k+1)} - x^{(k)}$. Then we have

$$
\Delta r(x^{(k)}) = (s^{(k)})^T \left[-D + (\Omega^{(k)})^{-1}\Lambda^{(k+1)}(\Xi^{(k)} + I - \Omega^{(k)})\right] s^{(k)}. \tag{2.18}
$$

Theorem 2.1. *Let $\left\{x^{(k)}\right\}$ be generated from Algorithm 2, with the corresponding $\left\{\Lambda^{(k)}\right\}$ given in (2.8). If $\left\{-D + (\Omega^{(k)})^{-1}\Lambda^{(k+1)}(\Xi^{(k)} + I - \Omega^{(k)})\right\}$ is uniformly positive definite. Then*

$$
\lim_{k\to+\infty} s^{(k)} = 0. \tag{2.19}
$$

Proof. Under these assumptions, and by (2.18), we know that $\left\{r(x^{(k)})\right\}$ is nondecreasing but bounded. This directly implies (2.19). □

Let $a_i \in \sigma(A_{ii})$, a_i is an eigenvalue of the matrix A_{ii}. Obviously, $a_i > 0$. If

$$
\left\{-D + (\Omega^{(k)})^{-1}\Lambda^{(k+1)}(\Xi^{(k)} + I - \Omega^{(k)})\right\}
$$

is uniformly positive definite, then we have

$$
-a_i + \frac{\lambda_i^{(k+1)}(\xi_i^k - \omega_i^{(k)} + 1)}{\omega_i^{(k)}} > 0.
$$

By direct computation, we have

$$
\omega_i^{(k)} < \frac{\lambda_i^{(k+1)}(\xi_i^k + 1)}{\lambda_i^{(k+1)} + a_i} < \frac{\lambda_i^{(k+1)}(\xi_i^k + 1)}{\lambda_i^{(k+1)}} = \xi_i^k + 1. \tag{2.20}
$$

By the equation (2.9), if $\omega_i^{(k)} \geq 1$, we get

$$
\xi_i^k \left\|x_i^{(k+1)}\right\|_2 = \left\|(1 - \omega_i^{(k)})x_i^{(k)} + \omega_i^{(k)} y_i^{(k+1)}\right\| \geq \omega_i^{(k)} - (\omega_i^{(k)} - 1) = 1.
$$

Then, we have

$$
\omega_i^{(k)} < 2.
$$

And we find that, if $\omega_i^{(k)}$ is very small, then $x_i^{(k+1)} \to x_i^{(k)}$, so we usually let $\omega_i^{(k)} > 0.5$ in our numerical experiments.

Lemma 2.2. ([1, Lemma 4.4]). *Let a_k be a bounded sequence of real numbers with the property $|a_k - a_{k+1}| \to 0$ as $k \to 0$. If there are only finitely many limit points for the sequence, then $\{a_k\}$ converges to a unique limit point.*

Theorem 2.3. *Under the assumptions of Theorem 2.1, and assume further that A has n distinct eigenvalues, $\omega_i^{(k)} = \omega$ is a constant. Then the sequence $\left\{\Lambda^{(k)}\right\}$ converges as $k \to +\infty$, the sequence $\left\{x^{(k)}\right\}$ from Algorithm 2 converges monotonically to a solution of the MEP.*

Proof. From the compactness, we know that $\left\{x^{(j_k)}\right\}$ has a convergent subsequence. Without loss of generality, we may assume $\left\{x^{(j_k)}\right\}$ converges to $\bar{x}$ and $\left\{\xi^{(j_k)}\right\}$ converges to $\bar{\xi}$. By (2.9), we get

$$\bar{\xi}\bar{x} = (1-\omega)\bar{x} + \omega \lim_{k\to+\infty} y_i^{(k+1)},$$

and

$$\left\|\lim_{k\to+\infty} y_i^{(k+1)}\right\|_2 = \|\bar{x}\|_2 = 1.$$

Then we get

$$\lim_{k\to+\infty} y_i^{(k+1)} = \bar{x}.$$

Therefore,

$$\begin{aligned}
\lim_{k\to+\infty} \delta x^{(k)} &= \lim_{k\to+\infty} (Ax^{(k)} - \Lambda x^{(k)}) \\
&= \lim_{k\to+\infty} (Ax^{(k)} - \Lambda y^{(k)}) \\
&= \lim_{k\to+\infty} \left[(U + U^T + D)x^{(k)} - (U^T + D)x^{(k)} - Ux^{(k-1)}\right] \\
&= \lim_{k\to+\infty} Us^{(k-1)} = 0.
\end{aligned} \tag{2.21}$$

So, subsequences $\left\{x^{(j_k)}, \Lambda^{(j_k)}\right\}$ converges to $\left\{\bar{x}, \bar{\Lambda}\right\}$, and $A\bar{x} = \bar{\Lambda}\bar{x}$.

From the compactness and the result that the MEP has only finitely many solutions [3], we know that $\left\{x^{(k)}, \Lambda^{(k)}\right\}$ has finitely many limit points. By (2.9), we have

$$\left|\lambda_i^{(k+1)} - \lambda_i^{(k)}\right| \le \sum_{j=1}^{i} \left\|A_{ij}\right\|_2 \left\|x_i^{(k+1)} - x_i^{(k)}\right\|_2 + \sum_{j=i+1}^{m} \left\|A_{ij}\right\|_2 \left\|x_i^{(k)} - x_i^{(k-1)}\right\|_2.$$

By Theorem 2.1, we see that $\left|\lambda_i^{(k+1)} - \lambda_i^{(k)}\right| \to 0$ as $k \to +\infty$. Together with Lemma 2.2, we can get that, the sequence $\left\{\Lambda^{(k)}\right\}$ converges as $k \to +\infty$, and the sequence $\left\{x^{(k)}\right\}$ from Algorithm 2 converges monotonically to a solution of the MEP. □

We shall show that the SOR-like AVM is able to converge globally to the global maximizer of the MCP when A is nonnegative irreducible.

Theorem 2.4. *Suppose that A is a nonnegative irreducible matrix. For any $x^{(0)} \ge_e 0$, let $\left\{x^{(k)}\right\}$ be generated from Algorithm 2, with the corresponding $\left\{\Lambda^{(k)}\right\}$ given in (2.8). If $\left\{-D + (\Omega^{(k)})^{-1}\Lambda^{(k+1)}(\Xi^{(k)} + I - \Omega^{(k)})\right\}$ is uniformly positive definite, then $\left\{x^{(k)}\right\}$ converges to the positive global maximizer of the MCP.*

Proof. Similar to the proof of the Theorem 3.6 in [9]. □

3. Numerical experiments

In this section, we present our numerical experiments of the SOR-like AVM to show its efficiency, and most importantly, the effectiveness in seeking the global maximum of the MCP. All of our tests were conducted in MATLAB 7.0. We run the algorithm starting from 10^3 random initial points, and use $\left\|\delta x^{(k)}\right\|_2 \leq 10^{-6}$ as the stopping criterion.

Example 1. We choose the symmetric and positive definite matrix $BCSSTK04 \in R^{132\times 132}$ in the set BCSSTRUC1 from the Harwell-Boeing collection in Matrix Market as the original matrices A's for the MCP, with $m = 2$, and $\{P = 50, 82\}$, each of which consequently corresponds to a particular MCP problem.

Example 2. The matrix A is given by

$$A = \begin{pmatrix} I & A_{12} & A_{13} \\ A_{12}^T & I & A_{23} \\ A_{13}^T & A_{23}^T & I \end{pmatrix}, A_{12} = \begin{pmatrix} 0.636 & 0.126 & 0.059 \\ -0.021 & 0.633 & 0.049 \\ 0.016 & 0.157 & 0.521 \end{pmatrix},$$

$$A_{13} = \begin{pmatrix} 0.626 & 0.195 & 0.059 \\ 0.035 & 0.459 & 0.129 \\ 0.048 & 0.238 & 0.426 \end{pmatrix}, A_{23} = \begin{pmatrix} 0.709 & 0.050 & -0.002 \\ 0.039 & 0.532 & 0.190 \\ 0.067 & 0.258 & 0.299 \end{pmatrix},$$

with $m = 3$, and $\{P = 2, 3, 4\}$. This example is from [5].

In Table 1, under the columns "Avg. Iter. #" are the average numbers of iterations needed to meet the stopping criterion. Under columns "% to Global" are the sample probabilities, out of the 10^3 random tests, of convergence to a global maximizer. We observe, from the results, that for almost all randomly chosen starting points, the SOR-like AVM is able to reach a global maximizer as the standard AVM, for different choice of the ω, the SOR-like AVM may iterate faster than the standard AVM. This is very attractive and has consequential effect on applications.

Furthermore, to demonstrate more clearly the performance of each algorithm, we plot the history of $r(x^{(k)})$, the residual $\left\|\delta x^{(k)}\right\|_2$, and the multivariate eigenvalues $\lambda_i^{(k)}$ in the following figures. The iteration $\left\{x^{(k)}\right\}$ of Example 1 in Figure 1, Figure 2 and Figure 3 starts from

$$x^{(0)} = \left[\ \frac{1}{\sqrt{50}}, \ldots, \frac{1}{\sqrt{50}}, \frac{1}{\sqrt{82}}, \ldots, \frac{1}{\sqrt{82}}\ \right]^T,$$

while the sequence $\left\{x^{(k)}\right\}$ of Example 2 in Figure 4, Figure 5 and Figure 6 starts from

$$x^{(0)} = \left[\ 1, 0, 1, 0, 0, 1, 0, 0, 0\ \right]^T.$$

Table 1. Performance of the SOR-like AVM on Examples 1 and 2.

Example	ω	Avg. Iter.#	% to Global
Example 1	$\omega = 1(AVM)$	99.44	100.00
	$\omega = 1.4$	33.08	100.00
Example 2	$\omega = 1(AVM)$	22.59	100.00
	$\omega = 1.2$	12.83	100.00

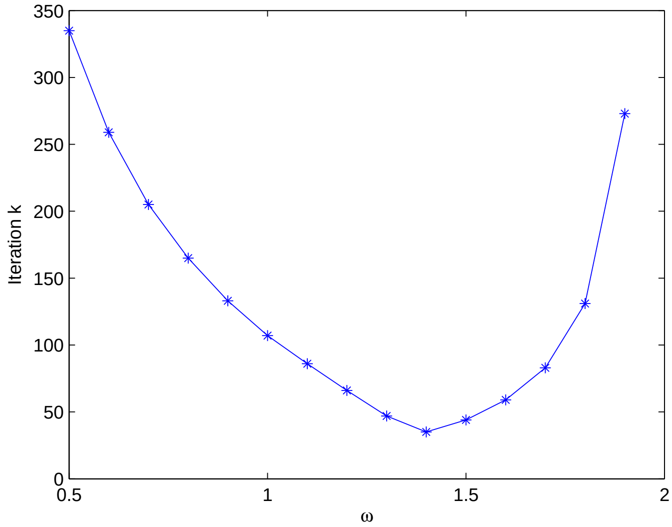

Figure 1. The history of ω for Example 1.

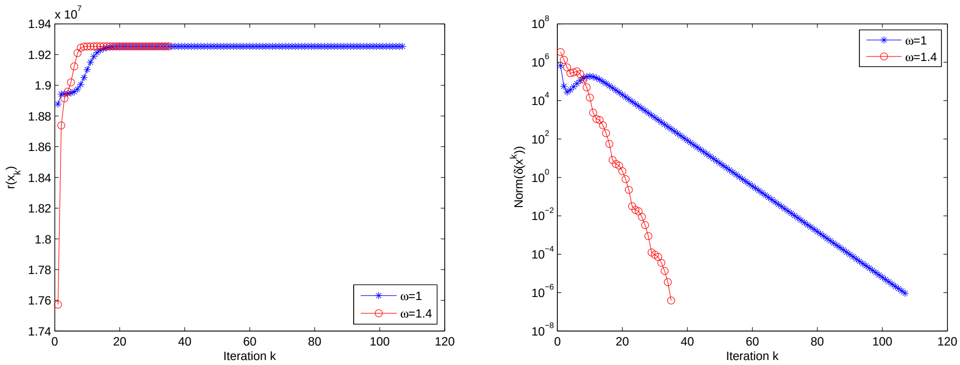

Figure 2. The history of $r(x_k)$ and $Norm(\delta(x^k))$ for Example 1.

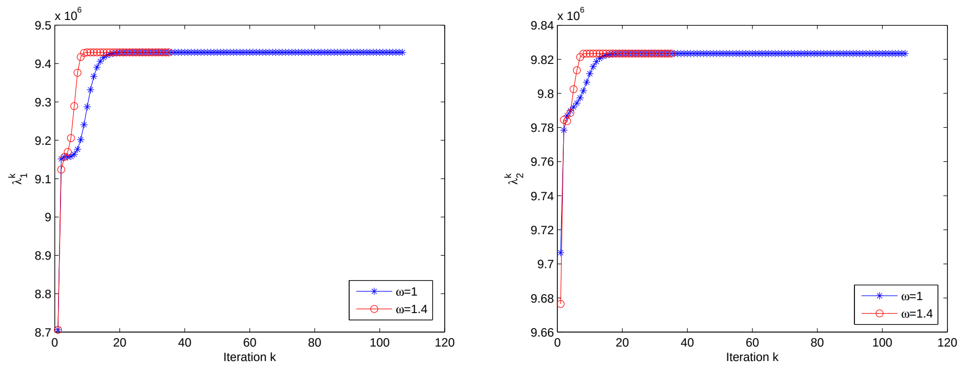

Figure 3. The history of λ_i^k for Example 1.

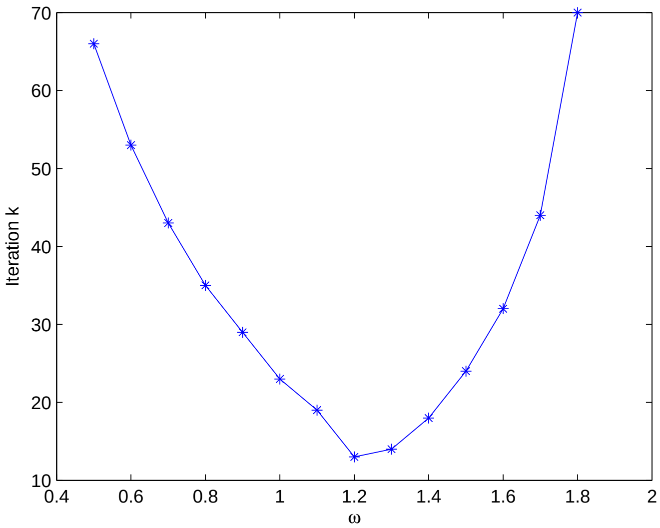

Figure 4. The history of ω for Example 2.

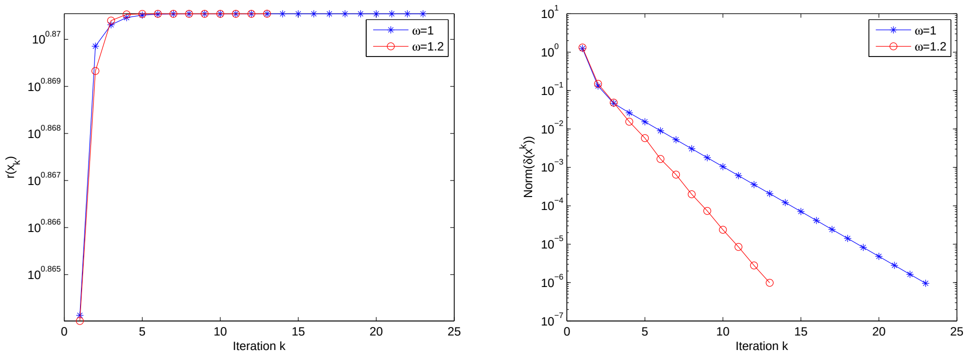

Figure 5. The history of $r(x_k)$ and $Norm(\delta(x^k))$ for Example 2.

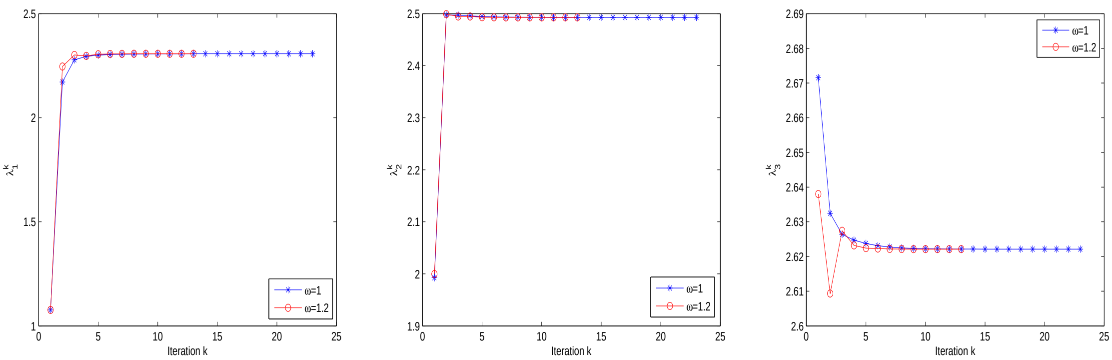

Figure 6. The history of λ_i^k for Example 2.

4. Conclusion

In this paper, we prove the monotone convergence of the SOR-like alternating variable method. The SOR-like AVM shows its computational advantage, for different choice of the ω, the SOR-like AVM may have a better performance than the standard AVM.

Acknowledgments

This research is supported by Science and technology Foundation of Guizhou Province (Qian Ke He Ji Chu [2016]1161); Guizhou Province natural science foundation in China (Qian Jiao He KY [2016]255); The doctoral scientific research foundation of Zunyi Normal College(BS[2015]09);High-level innovative talents of Guizhou Province(Zun Ke He Ren Cai[2017]8).

Conflict of interest

The author declares no conflicts of interest in this paper.

References

1. H. Hotelling, *The most predictable criterion*, J. Educ. Pyschol., **26** (1935), 139–142.
2. H. Hotelling, *Relations between two sets of variates*, Biometrika, **28** (1936), 321–377.
3. M. T. Chu, J. L. Watterson, *On a multivariate eigenvalue problem, part I: Algebraic theory and a power method*, SIAM J. Sci. Comput., **14** (1993), 1089–1106.
4. M. Hanafi, J. M. F. Ten Berge, *Global optimality of the successive Maxbet algorithm*, Psychometrika, **68** (2003), 97–103.
5. P. Horst, *Relations among m sets of measures*, Psychometrika, **26** (1961), 129–149.
6. J. Nocedal, S. J. Wright, *Numerical Optimization*, 2nd edn. Springer, New York, 2006.
7. L. H. Zhang, M. T. Chu, *Computing absolute maximum correlation*, IMA J. Numer. Anal., **32** (2012), 163–184.
8. L. H. Zhang, L. Z. Liao, L. M. Sun, *Towards the global solution of the maximal correlation problem*, J. Global Optim., **49** (2011), 91–107.
9. L. H. Zhang, L. Z. Liao, (2012) *An alternating variable method for the maximal correlation problem*, J. Global Optim., **54** (2012), 199–218.