THE BULLETIN OF OKAYAMA UNIVERSITY OF SCIENCE No. 41 A pp. 7-16 (2005)

Improving the EM algorithm using the vector epsilon accelerator

Masahiro KURODA and Michio SAKAKIHARA*

Department of Socio-Information,

* *Department of Information Science,*

Okayama University of Science

1-1 Ridai-cho, Okayama 700-0005, Japan

(Received September 29, 2005; accepted November 7, 2005)

The EM algorithm of Dempster, Laird and Rubin (1977) is a very general and popular iterative computational algorithm to find maximum likelihood estimates from incomplete data and broadly used to statistical analysis with missing data, because of its stability, flexibility and simplicity. However, it is often criticized that the convergence of the EM algorithm is slow. The various algorithms to accelerate the convergence of the EM algorithm have been proposed. In this paper, we propose the EM-εA algorithm accelerating the convergence of the sequence of EM iterates via the vector ε algorithm of Wynn (1961) and show the theoretical property. The EM-εA algorithm is extended to the EM algorithm without affecting its stability, flexibility and simplicity. Numerical experiments illustrate the potential of the EM-εA algorithm.

1 Introduction

The EM algorithm of Dempster, Laird and Rubin (1977) is a very general and popular iterative computational algorithm to find maximum likelihood estimates (MLEs) from incomplete data and broadly applied to statistical analysis with missing data, because of its stability, flexibility and simplicity. However, it is often criticized that the convergence of the EM algorithm is slow, since its rate of convergence is linear. In particular, the convergence of the EM algorithm is quite slow when the proportion of observations that are missing data is high.

In order to speed up the convergence of the EM algorithm, various acceleration algorithms have been proposed. Louis (1982) suggested the EM algorithm incorporating Aitken's acceleration method. Jamshidian and Jennrich (1993) proposed an acceleration algorithm based on conjugate gradients. Lange (1995) used a quasi-Newton algorithm to accelerate the EM algorithm. Applying the algorithms based on the Newton-Raphson algorithm to accelerating the convergence of the EM algorithm, there are potentially including unavoidable problems. Firstly, it requires the computation of the information matrix at each iteration. Then its computation is likely to become rapidly complicated as the number of parameters is increasing. Secondly, the Newton-Raphson algorithm is more sensitive for initial values than the EM algorithm. Because of such possible computational difficulties, their acceleration EM algorithms are lost the attractive features of the EM algorithm, such as its stability, flexibility and simplicity.

As extensions of the EM algorithm, Meng and Rubin (1993) proposed the ECM (Expectation-Conditional Maximization) algorithm that improves the maximization process in the M-step. Liu and Rubin (1994) suggested the ECME (Expectation-Conditional Maximization Either) algorithm that is an extension of the ECM algorithm. McLachlan and Krishnan (1997) provided a comprehensive account of the EM algorithm. Geng, Wan and Tao (2000) presented the partial imputation EM algorithm that imputes a part of missing data such that the observed data and imputed data construct a monotone data pattern.

In this paper, we propose the EM-εA algorithm accelerating the convergence of the sequence of EM iterates using the vector ε algorithm of Wynn (1961). Section 2 describes the vector ε algorithm. In section 3, we show the EM-εA algorithm and give the key result of the EM-εA algorithm that the sequence of EM-εA iterates converges to the stationary point of the sequence of EM iterates and, for scalar sequences, its speed of convergence is faster. Section 4 examines the performance and properties of the EM-εA algorithm by numerical experiments.

2 The vector ε algorithm

The ε algorithm of Wynn (1956) is a nonlinear method for accelerating the convergence of slowly convergent sequences. It is known that its algorithm is powerful for the sequence converging linearly. In this section, we present the vector ε algorithm of Wynn (1961) for the vector sequence.

Let θ be d-dimensional vector and $\theta^{(t)}$ denote values of θ after the t-th iteration. Define the inverse $[x]^{-1}$ of a vector x by

$$[x]^{-1} = \frac{x}{\|x\|^2},$$

where $\|\cdot\|$ is the usual Euclid norm for vectors. For acceleration of convergence of a sequence of iterates $\{\theta^{(t)}\}_{t\geq 0}$, the rule of the vector ε algorithm is

$$\begin{aligned} \varepsilon^{(t,-1)} &= 0, \\ \varepsilon^{(t,0)} &= \theta^{(t)}, \\ \varepsilon^{(t,k+1)} &= \varepsilon^{(t+1,k-1)} + \left[\varepsilon^{(t+1,k)} - \varepsilon^{(t,k)}\right]^{-1}. \end{aligned}$$

According to Brezinski and Zaglia (1991), we have

$$\begin{aligned} &\varepsilon^{(t,2k+2)} \\ &= \varepsilon^{(t+1,2k)} + \left[\left[\varepsilon^{(t,2k)} - \varepsilon^{(t+1,2k)}\right]^{-1} + \left[\varepsilon^{(t+2,2k)} - \varepsilon^{(t+1,2k)}\right]^{-1} - \left[\varepsilon^{(t+2,2k-2)} - \varepsilon^{(t+1,2k)}\right]^{-1}\right]^{-1}. \end{aligned}$$

For the initializations $\varepsilon^{(t,0)} = \theta^{(t)}$ and $\varepsilon^{(t,-2)} = \infty$, we obtain the sequence $\{\dot{\theta}^{(t)}\}_{t\geq 0}$ generated by

$$\dot{\theta}^{(t)} = \varepsilon^{(t,2)} = \theta^{(t+1)} + \left[\left[\theta^{(t)} - \theta^{(t+1)}\right]^{-1} + \left[\theta^{(t+2)} - \theta^{(t+1)}\right]^{-1}\right]^{-1}. \tag{2.1}$$

Note that, at each iteration, the vector ε algorithm is achieved at a cost of $O(d^2)$ while the Newton-Raphson algorithm requires $O(d^3)$ and the computational cost is likely to become more expensive as d becomes large.

3 The EM-εA algorithm

Let y be observed data with a sample space Ω_Y and x be complete data augmented by y with a sample space Ω_X. We assume that there exists some function $h(x) = y$ that relates x to y. Let $f(\cdot|\theta)$ denote a probability density function depending on an unknown parameter vector θ with a parameter space Θ. Define the conditional expectation of the log-likelihood function $\ell(x|\theta) = \log f(x|\theta)$ given y and θ' as

$$Q(\theta|\theta') = \mathrm{E}[\ell(X|\theta)|y, \theta'].$$

The EM algorithm chooses

$$\theta^{(t)} = \arg\max_{\theta\in\Theta} Q(\theta|\theta^{(t-1)}),$$

at each iteration $t = 1, 2, \ldots$.

Given an initial value $\theta^{(0)} \in \Theta$, the EM-εA algorithm adding the ε-accelerating process performs the following steps:

E-step : Calculate

$$Q(\theta|\theta^{(t-1)}) = \mathrm{E}[\ell(X|\theta)|y, \theta^{(t-1)}].$$

M-step : Choose $\theta^{(t)}$ such that

$$Q(\theta^{(t)}|\theta^{(t-1)}) \geq Q(\theta|\theta^{(t-1)})$$

for all $\theta \in \Theta$.

ε-accelerating : Calculate $\dot{\theta}^{(t-2)}$ by equation (2.1) and check the convergence using

$$|\dot{\theta}^{(t-2)} - \dot{\theta}^{(t-3)}| \leq \delta,$$

where δ is a desired accuracy.

As the advantage of the EM algorithm, Dempster, Liard and Rubin (1977) show that the EM algorithm increases $\ell(y|\theta)$ at each iteration, that is,

$$\ell(y|\theta^{(t)}) \geq \ell(y|\theta^{(t-1)}), \tag{3.1}$$

with equality if and only if

$$Q(\theta^{(t+1)}|\theta^{(t)}) = Q(\theta^{(t)}|\theta^{(t)}),$$

and if $\ell(y|\theta)$ is bounded, the sequence of $\ell(y|\theta^{(t)})$ for $t = 0, 1, \ldots$ converges to ℓ^*, the limit of $\ell(y|\theta^{(t)})$. With the ε-accelerating process given by equation (2.1), the following theorem can be immediately shown.

Theorem 3.1 *Suppose that $\theta^{(t)}$ converges to a stationary point θ^* with $\ell(y|\theta^{(t)}) = \ell^*$. Then the sequence $\dot{\theta}^{(t)}$ generated by equation (2.1) converges to θ^*, that is,*

$$\theta^* = \lim_{t\to\infty} \theta^{(t)} = \lim_{t\to\infty} \dot{\theta}^{(t)}. \tag{3.2}$$

Proof: Let $\langle \cdot, \cdot \rangle$ denote the inner product and set

$$\eta^{(t)} = \frac{\theta^{(t)} - \theta^{(t+1)}}{||\theta^{(t)} - \theta^{(t+1)}||^2} + \frac{\theta^{(t+2)} - \theta^{(t+1)}}{||\theta^{(t+2)} - \theta^{(t+1)}||^2}. \tag{3.3}$$

From equation (2.1), we obtain

$$\left\langle (\dot{\theta}^{(t)} - \theta^{(t+1)}), \eta^{(t)} \right\rangle = 1. \tag{3.4}$$

According to equation (3.3), $\eta^{(t)}$ diverges to infinity as t tends to infinity, since $\lim_{t\to\infty} \theta^{(t)} = \theta^*$. For any t, equation (3.4) holds, so that $\lim_{t\to\infty}(\dot{\theta}^{(t)} - \theta^{(t+1)}) = \vec{0}$, where $\vec{0}$ is the zero vector. Then we have the results. □

For a scalar sequence $\{\theta^{(t)}\}_{t\geq 0}$, equation (2.1) is formulated by

$$\begin{aligned} \dot{\theta}^{(t)} &= \theta^{(t+1)} + \left(\frac{1}{\theta^{(t)} - \theta^{(t+1)}} + \frac{1}{\theta^{(t+2)} - \theta^{(t+1)}} \right)^{-1} \\ &= \theta^{(t+1)} + \frac{(\theta^{(t)} - \theta^{(t+1)})(\theta^{(t+2)} - \theta^{(t+1)})}{\theta^{(t+2)} - 2\theta^{(t+1)} + \theta^{(t)}} \end{aligned} \tag{3.5}$$

and then is identical to Aitken δ^2 method. Traub (1964) proved that Aitken δ^2 method accelerates the convergence of linear convergent sequences and the rate of convergence is suppler linear. Form the fact that the sequences of EM iterates converge linearly, we can obtain the following theorem.

Theorem 3.2 *Assume that* $\{\theta^{(t)}\}_{t\geq 0}$ *is a scalar sequence of EM iterates. Then the sequence* $\{\dot{\theta}^{(t)}\}_{t\geq 0}$ *generated by equation (3.5) converges to* θ^* *faster than* $\{\theta^{(t)}\}_{t\geq 0}$.

Louis (1982) gave the formula resembles equation (2.1) based on Aitken's acceleration method such as

$$\tilde{\theta}^{(t)} = \theta^{(t-1)} + (I - J^{(t-1)})^{-1}(\theta^{(t)} - \theta^{(t-1)}),$$

where J is the Jacobian of the map $\theta^{(t)} = M(\theta^{(t-1)})$ and I also denotes the $d \times d$ identity matrix. In order to estimate $(I - J^{(t-1)})^{-1}$, Louis suggests making use of the equation

$$I - J^{(t-1)} = \mathrm{E}[\mathcal{I}(\theta^{(t-1)}|X)|y]^{-1}\mathcal{I}(\theta^{(t-1)}|y), \tag{3.6}$$

where

$$\mathcal{I}(\theta^{(t-1)}|x) = \frac{\partial^2}{\partial\theta\partial\theta^T}\ell(x|\theta)\bigg|_{\theta=\theta^{(t-1)}},$$
$$\mathcal{I}(\theta^{(t-1)}|y) = \frac{\partial^2}{\partial\theta\partial\theta^T}\ell(y|\theta)\bigg|_{\theta=\theta^{(t-1)}}.$$

Meilijson (1989) points out that Aitken acceleration method using equation (3.6) is essentially equivalent to the Newton-Raphson algorithm.

We emphasize that the EM-εA algorithm accelerates the convergence of the sequence of EM iterates. Thus the EM-εA algorithm is extended to the EM algorithm without affecting its simplicity and stability. Moreover, the convergence properties of the EM algorithm are preserved, because the EM-εA algorithm does not improve the E- and M-steps in themselves but only adding the ε-accelerating process.

4 Numerical Experiments

In this section, we provide numerical experiments to examine the performance of the EM-εA algorithm in comparison with the EM algorithm.

Example 1: Contingency tables with partially classified observations

Consider a 2×2 contingency table with completely and partially classified observations. Let X and Y be dichotomous variables and $\theta = \{p_{ij}\}_{i,j=1,2}$ be a set of joint probabilities of X and Y. We denote the cross-classified data of X and Y as $n_{XY} = \{n_{XY}(i,j)\}_{i,j=1,2}$, and partially classified data of X as $n_X = \{n_X(i)\}_{i=1,2}$ and Y as $n_Y = \{n_Y(j)\}_{j=1,2}$. Assume that these data have a multinomial distribution with an unknown parameter θ.

The data are shown in Table 1. For these data patterns as Table 1, the convergence of EM algorithm is quite slow, because its convergence is deeply associated with the proportion of missing data. In Table 2, we summarize the numbers of iterations for the EM and EM-εA algorithms with each $\delta = 10^{-5}$ to 10^{-8} and the data (a) to (e), and plot these results in Figure 1. Table 3 gives the estimates of θ by the EM and EM-εA algorithms.

As shown in Table 2 and Figure 1, the EM algorithm increases linearly the numbers of iterations as the data (a) to (e) change, while there are few changes in the numbers of iterations for the EM-εA algorithm and its convergence is significantly faster. For example, with $\delta = 10^{-6}$ and the data (d), the EM-εA algorithm takes only 41 iterations using 43 EM iterations to obtain the final values, while the EM algorithm takes 476 iterations to achieve the same values. For other data, the EM algorithm requires the numbers of iterations more than roughly 3-10 times of the EM-εA algorithm. Table 3 demonstrates that the estimates by the EM algorithm with $\delta = 10^{-5}$ match only three-digits to the MLEs that are the estimates with $\delta = 10^{-7}$ and 10^{-8}. Then the EM-εA algorithm finds the estimates identical to the MLEs.

The plots in Figure 2 illustrate that, for each iteration, the EM-εA algorithm finds the value $\dot{\theta}^{(t)}$ satisfying equations

$$\ell(y|\dot{\theta}^{(t)}) \geq \ell(y|\theta^{(t)}). \tag{4.1}$$

and

$$\ell(y|\dot{\theta}^{(t)}) \geq \ell(y|\dot{\theta}^{(t-1)}).$$

These equations imply that the EM-εA sequence $\{\dot{\theta}^{(t)}\}_{t\geq 0}$ converges to the MLEs faster than the EM sequence $\{\theta^{(t)}\}_{t\geq 0}$ and also the EM-εA algorithm monotonically increases $\ell(y|\theta)$.

Next we investigate the rate of convergence of the EM-εA algorithm. In this numerical experiments, the (i,j)-th componentwise rates of convergence of the EM and EM-εA algorithms are assessed as

$$\begin{aligned} R_{ij}^{EM} &= \lim_{t\to\infty} r_{ij}^{EM(t)} = \lim_{t\to\infty} \frac{\left| p_{ij}^{(t)} - p_{ij}^{MLE} \right|}{\left| p_{ij}^{(t-1)} - p_{ij}^{MLE} \right|}, \\ R_{ij}^{EM-\varepsilon A} &= \lim_{t\to\infty} r_{ij}^{EM-\varepsilon A(t)} = \lim_{t\to\infty} \frac{\left| \dot{p}_{ij}^{(t)} - p_{ij}^{MLE} \right|}{\left| p_{ij}^{(t+2)} - p_{ij}^{MLE} \right|}. \end{aligned}$$

Figure 3 shows the traces of $\{r_{ij}^{EM(t)}\}_{i,j=1,2}$ and $\{r_{ij}^{EM-\varepsilon A(t)}\}_{i,j=1,2}$ of the data (a) to (e) for each iteration with $\delta = 10^{-6}$. It can be seen that the EM-εA algorithm converges faster than the EM algorithm, since $R_{ij}^{EM-\varepsilon A}$ for all i and j converges to zero.

Example 2: Incomplete bivariate normal data

Let (X_1, X_2) be a bivariate normal vector with unknown parameters $\mu = (\mu_1, \mu_2)$ and $\Sigma = (\sigma_{11}, \sigma_{22}, \sigma_{12})$. The observed data are shown in Table 4, where the observed data from No. 1 to 3 are complete, and both the data from No. 4, 5 and No. 6, 7 are incomplete with missing X_2 and X_1, respectively.

The EM algorithm fills in missing data of incompletely observed data in the E-step and estimates μ and Σ in the M-step. With the data (a) and $\delta = 10^{-6}$, after 317 iterations of the EM algorithm, the MLEs of μ and Σ are obtained as follows

$$\mu = (1.3005, 1.4163), \quad \Sigma = (0.2371, 4.9603, -1.0478).$$

The EM-εA algorithm finds same values after 172 iterations using 174 EM iterates. Using both algorithms with the data (b) and $\delta = 10^{-6}$, we obtain the same MLEs of μ and Σ such as

$$\mu = (78.3977, 2247.1084), \quad \Sigma = (70.1051, 79869.7113, 2182.2234).$$

Then the EM-εA algorithm takes 133 iterations while the EM algorithm does 312 iterations.

From these numerical experiments, we can also see that the EM-εA algorithm converges to the MLEs faster than the EM algorithm.

5 Concluding Remarks

In this paper, we proposed the EM-εA algorithm that accelerates the convergence of the EM algorithm incorporating the vector ε algorithm. The vector ε algorithm in itself is very simple computational procedure and also its computational cost is more inexpensive than the one of the Newton-Raphson algorithm. The fundamental difference between the EM-εA algorithm and Newton-type accelerating EM algorithms is that the EM-εA algorithm does not improve the E- and M-steps but accelerates the convergence of the sequence of the EM iterations. Therefore the EM-εA algorithm is an extension algorithm within the framework of the EM algorithm without affecting its simplicity, stability and flexibility.

As the theoretical result, we provide that the EM-εA algorithm is guaranteed to convergence to the stationary point of the sequence of EM iterates and, for scalar sequences of EM iterates, the EM-εA algorithm accelerates the convergence. The numerical experiments demonstrate that the EM-εA algorithm produces the sequence of iterates to converge to MLEs faster than the EM sequence. Then the EM-εA algorithm finds sufficiently accurate estimates using the smaller number of EM iterations.

Our future problem is to evaluate theoretically the rate of convergence of the EM-εA algorithm for vector cases.

Acknowledgement

This research is supported by Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Young Scientists, No 16700264 and Wesco Scientific Promotion Foundation.

References

[1] Brezinski, C. and Zaglia, M.R. (1991). *Extrapolation methods: theory and practice.* Elsevier Science Ltd.

[2] Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B.* **39**, 1 – 22.

[3] Geng, Z., Wan, K. and Tao, F. (2000). Mixed graphical models with missing data and the partial imputation EM algorithm. *Scand. J. Statist.* **27**, pp. 433 – 444.

[4] Jamshidian, M. and Jennrich, R.I. (1993). Conjugate gradient acceleration of the EM algorithm. *J. Amer. Statist. Assoc.* **88**, 221 – 228.

[5] Lange, K. (1995). A gradient algorithm locally equivalent to the EM algorithm. *J. Roy. Statist. Soc. Ser. B* **57**, 425 – 437.

[6] Liu, C. and Rubin, D.B. (1994). The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. *Biometrika* **81**, 633 – 648.

[7] Louis, T.A. (1982). Finding the observed information matrix when using the EM algorithm. *J. R. Stat. Soc. Ser. B.* **44**, 226 – 233.

[8] McLachlan, G.J. and Krishnan, T. (1997). *The EM algorithm and extensions.* New York: Wiley.

[9] Meilijson, I. (1989). A fast improvement to the EM algorithm on its own terms. *J. R. Stat. Soc. Ser. B.* **51**, 127 – 138.

[10] Meng, X.L. and Rubin, D.B. (1993). Maximum likelihood estimation via the ECM algorithm: a general framework. *Biometrika* **80**, 267 – 278.

[11] Wynn, P. (1956). On a procrustean technique for the numerical transformation of slowly convergent sequences and series. *Proc. Cambridge Phil. Sco.* **52**, 663 – 671.

[12] Wynn, P. (1961). The epsilon algorithm and operational formulas of numerical analysis. *Math. Comp.* **15**, 151 – 158.

Table 1: Contingency table with completely and partially classified data

	n_Y		n_X		n_{XY}			
			$i=1$	$i=2$	$i=1$		$i=2$	
	$j=1$	$j=2$			$j=1$	$j=2$	$j=1$	$j=2$
(a)	50	30	300	200	5	4	2	1
(b)	100	60	300	200	5	4	2	1
(c)	250	150	300	200	5	4	2	1
(d)	500	300	300	200	5	4	2	1
(e)	1000	600	300	200	5	4	2	1

Table 2: The numbers of iterations with each δ

δ		(a)	(b)	(c)	(d)	(e)
10^{-5}	EM	123	136	166	192	198
	EM-εA	58	40	27	36	59
10^{-6}	EM	253	282	364	476	656
	EM-εA	72	48	32	41	68
10^{-7}	EM	383	429	561	761	1114
	EM-εA	84	64	79	90	86
10^{-8}	EM	513	575	759	1045	1572
	EM-εA	119	136	179	234	313

Table 3: Estimates of θ by the EM and EM-εA algorithms with each δ

			$\delta = 10^{-5}$		$\delta = 10^{-6}$		$\delta = 10^{-7}, 10^{-8}$
			EM	EM-εA	EM	EM-εA	EM & EM-εA
(a)	$i=1$	$j=1$	0.3463	0.3457	0.3458	0.3458	0.3458
		$j=2$	0.2572	0.2577	0.2576	0.2577	0.2577
	$i=2$	$j=1$	0.2756	0.2761	0.2761	0.2761	0.2761
		$j=2$	0.1210	0.1205	0.1205	0.1204	0.1204
(b)	$i=1$	$j=1$	0.3471	0.3464	0.3465	0.3465	0.3465
		$j=2$	0.2564	0.2570	0.2569	0.2570	0.2570
	$i=2$	$j=1$	0.2763	0.2769	0.2768	0.2769	0.2769
		$j=2$	0.1203	0.1197	0.1197	0.1197	0.1197
(c)	$i=1$	$j=1$	0.3478	0.3469	0.3470	0.3469	0.3469
		$j=2$	0.2557	0.2565	0.2564	0.2565	0.2565
	$i=2$	$j=1$	0.2765	0.2774	0.2773	0.2774	0.2774
		$j=2$	0.1200	0.1192	0.1193	0.1192	0.1192
(d)	$i=1$	$j=1$	0.3483	0.3471	0.3472	0.3471	0.3471
		$j=2$	0.2551	0.2563	0.2562	0.2563	0.2564
	$i=2$	$j=1$	0.2763	0.2775	0.2774	0.2775	0.2776
		$j=2$	0.1202	0.1190	0.1191	0.1190	0.1190
(e)	$i=1$	$j=1$	0.3491	0.3472	0.3474	0.3472	0.3472
		$j=2$	0.2543	0.2563	0.2561	0.2563	0.2563
	$i=2$	$j=1$	0.2757	0.2776	0.2775	0.2776	0.2776
		$j=2$	0.1209	0.1189	0.1191	0.1190	0.1189

Table 4: Incomplete bivariate normal data

(a)		
No	X_1	X_2
1	1.2	2.3
2	1.7	0.1
3	1.6	-0.7
4	0.2	–
5	1.5	–
6	–	-0.2
7	–	1.6

(b)		
No	X_1	X_2
1	68	2000
2	71	1850
3	72	2100
4	84	–
5	90	–
6	–	2150
7	–	2600

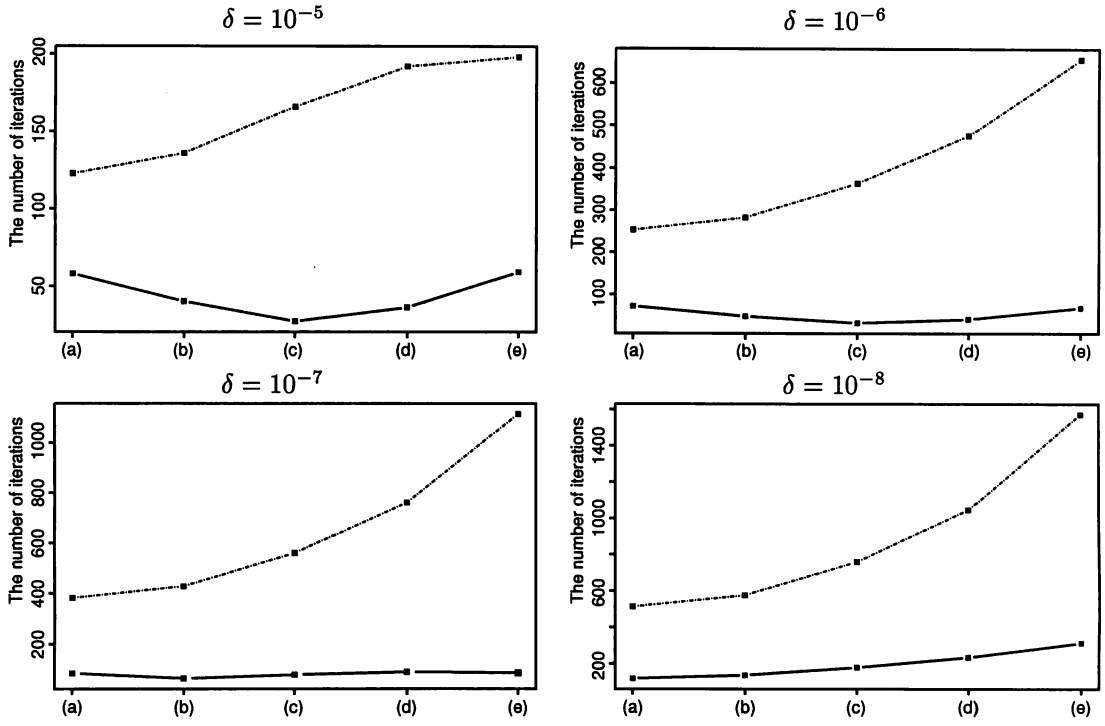

Figure 1: Plots of the numbers of iterations for the EM (dashed line) and the EM-εA (solid line) algorithms with each δ

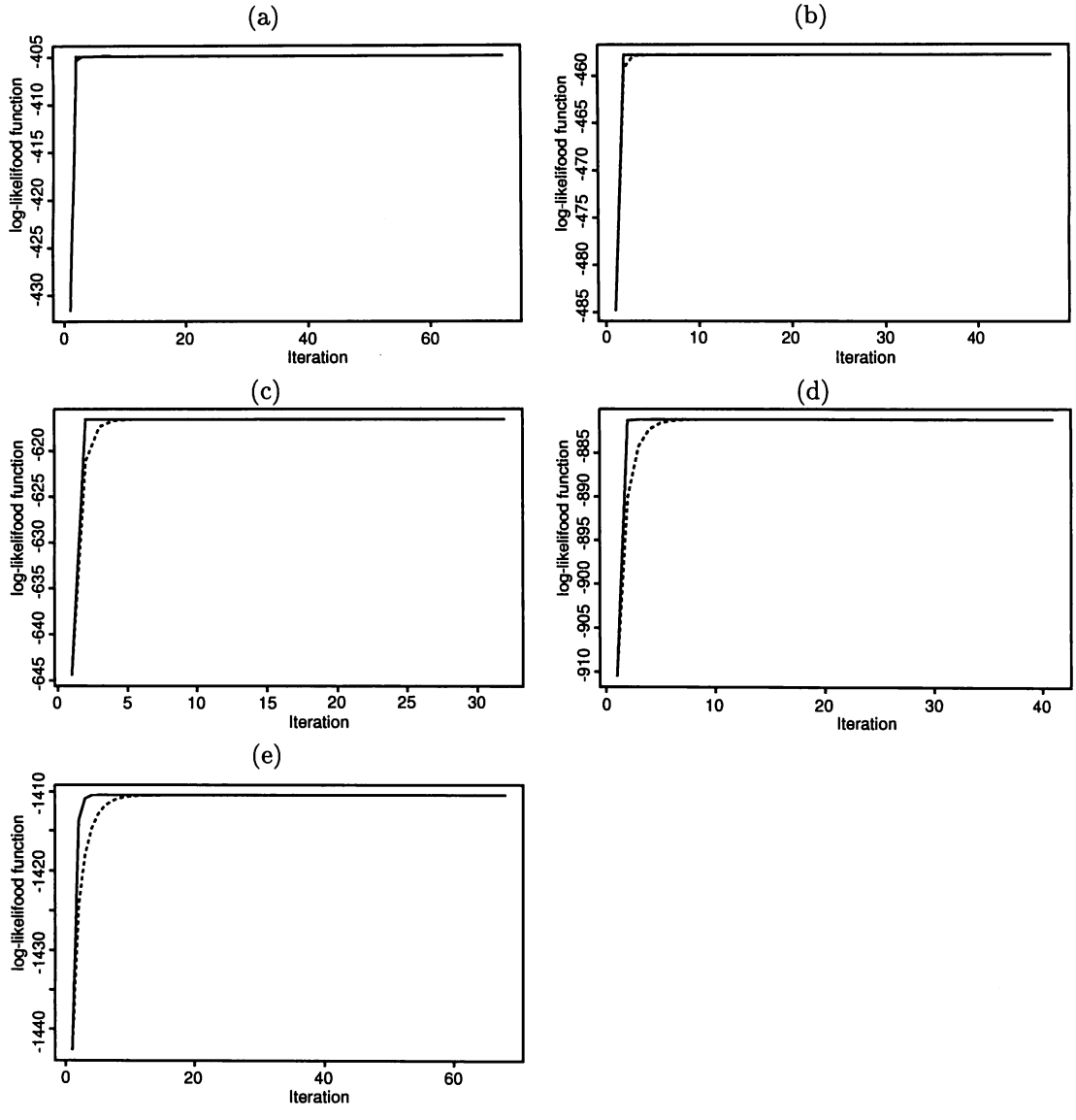

Figure 2: Plots of the log-likelihood functions $\ell(y|\theta)$ (dashed line) and $\ell(y|\dot{\theta})$ (solid line) with $\delta = 10^{-6}$ and the data (a) to (e).

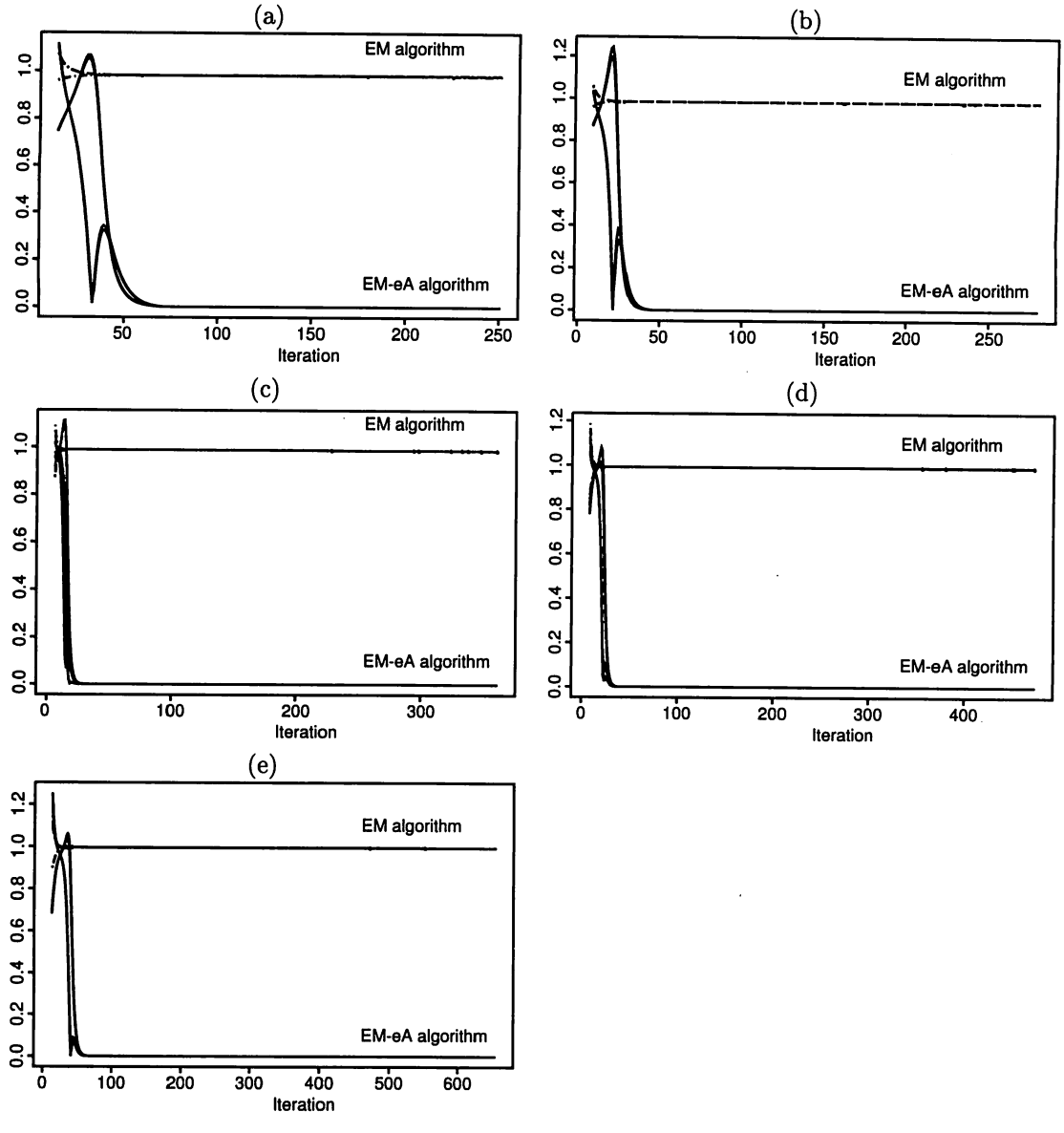

Figure 3: Traces of $\{r_{ij}^{EM(t)}\}_{i,j=1,2}$ and $\{r_{ij}^{EM-\varepsilon A(t)}\}_{i,j=1,2}$ with $\delta = 10^{-6}$ and the data (a) to (e).