◆ 画像解析等グループ ◆

画像処理における統計数値計算の改良
EM algorithm とその加速化

岡山理科大学総合情報学部

講師　黒 田 正 博

画像処理における統計数値計算の改良
EM algorithm とその加速化

黒田正博（岡山理科大学）

研究協力者
　　榊原道夫（岡山理科大学）
　　Geng Zhi（Peking University）

Overview: the EM algorithm

- The EM algorithm of Dempster, Laird and Rubin (1977) is a very general and popular iterative computational algorithm to find MLEs from incomplete data.
- The EM algorithm is 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.

Overview: 画像処理とEM algorithm

- 画像処理での問題･･･扱う対象が複数同時に存在
  - 画像中に複数の物体が存在している場合の物体の切り出し
- アプローチ･･･
  複数対象のモデル化問題
  ⇒ 確率分布の混合分布の最尤推定問題

- Emission and Transmission Tomography
  - Statistical Modeling: Poisson model
  - Lange, K. and Carson, R. (1984). *Journal of Computer Assisted Tomography*

Overview : the ε-accelerated EM algorithm

- Kuroda and Sakakihara (2006) proposed the ε-accelerated EM algorithm accelerating the convergence of the sequence of EM iterates via the vector ε algorithm of Wynn (1961).
- Wang, Kuroda, Sakakihara and Geng (2006) provided the theoretical results:
  - The ε-accelerated EM algorithm is guaranteed to convergence to the stationary point of the sequence of EM iterates
  - The ε-accelerated EM algorithm accelerates the speed of convergence of the EM algorithm.

■Kuroda, M. and Sakakihara, M. (2006). Accelerating convergence of the EM algorithm using the vector ε algorithm. *Comput. Stat. Data Anal*. (in press).

■Wang, M., Kuroda, M., Sakakihara, M. and Geng, Z. (2006). Acceleration of the EM algorithm using the vector epsilon algorithm. (in review process).

Overview : the ε -accelerated EM algorithm

- The ε -accelerated EM algorithm does not improve the EM algorithm in themselves
but only adding the ε -accelerating process.

- The ε -EM algorithm accelerates the convergence of the EM sequences and then does not depend on the statistical models.

- The ε -accelerated EM algorithm is extended to
the EM algorithm without affecting its stability, flexibility and simplicity.

The ε -accelerated EM algorithm

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

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 : Find the accelerated sequence

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

and check the convergence using

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

where $||x||_\infty = \max_i\{x_i\}$ for a vector $x = \{x_i\}$ and δ is a desired accuracy.

Key results: Convergence and acceleration of convergence

Suppose that $\theta^{(t)}$ converges to a stationary point θ^*.

<u>Theorem 1.</u> The sequence $\dot{\theta}^{(t)}$ generated by the ε-aacelerated EM algorithm converges to the stationary point θ^* of the EM sequence.

<u>Theorem 2.</u> Assume that $\{\theta^{(t)}\}_{t\geq 0}$ is the sequence of EM iterates and the sequence $\{\dot{\theta}^{(t)}\}_{t\geq 0}$ is generated by the equation (2). Then we have

$$\lim_{t\to\infty} \frac{||\dot{\theta}^{(t)} - \theta^*||}{||\theta^{(t+2)} - \theta^*||} = 0.$$

That is, $\{\dot{\theta}^{(t)}\}_{t\geq 0}$ converges to θ^* more quickly than $\{\theta^{(t)}\}_{t\geq 0}$ does.

Sequence of the ε -accelerated EM algorithm

Sequence

$$\textbf{EM} \qquad \theta^{(0)} \quad \theta^{(1)} \quad \theta^{(2)} \quad \theta^{(3)} \quad \ldots \quad \theta^{(S)} \qquad \ldots \quad \underline{\theta^{(T)}(=\theta^*)}$$

$$\varepsilon\textbf{-accelerating} \quad \dot{\theta}^{(0)} \quad \dot{\theta}^{(1)} \quad \dot{\theta}^{(2)} \quad \dot{\theta}^{(3)} \quad \ldots \quad \underline{\dot{\theta}^{(S)}(=\theta^*)}$$

convergence of
the stationary point of
the EM sequence

S(R) program: function of vector epsilon algorithm

```
inv.vec <- function(x)
{
   return(x/(x%*%x));
}

epsilon <- function(p.em)
{
   p.eps <- p.em[2,]+inv.vec(inv.vec(p.em[1,]-p.em[2,])
               +inv.vec(p.em[3,]-p.em[2,]));
   return(p.eps);
}
```

It is easy to implement the vector epsilon algorithm by S (R) languages!!

Two-way contingency tables with completely and partially classified data: notation

Let X and Y be dichotomous variables and have a multinomial distribution with

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 and Y as

$$n_X = \{n_X(i)\}_{i=1,2} \quad \textbf{and} \quad n_Y = \{n_Y(j)\}_{j=1,2}.$$

Two-way contingency tables with completely and partially classified data: observed data

proportion of missing data: low → high

convergence speed: slow → extremely slow

change | fix

	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

The number of iterations with each δ

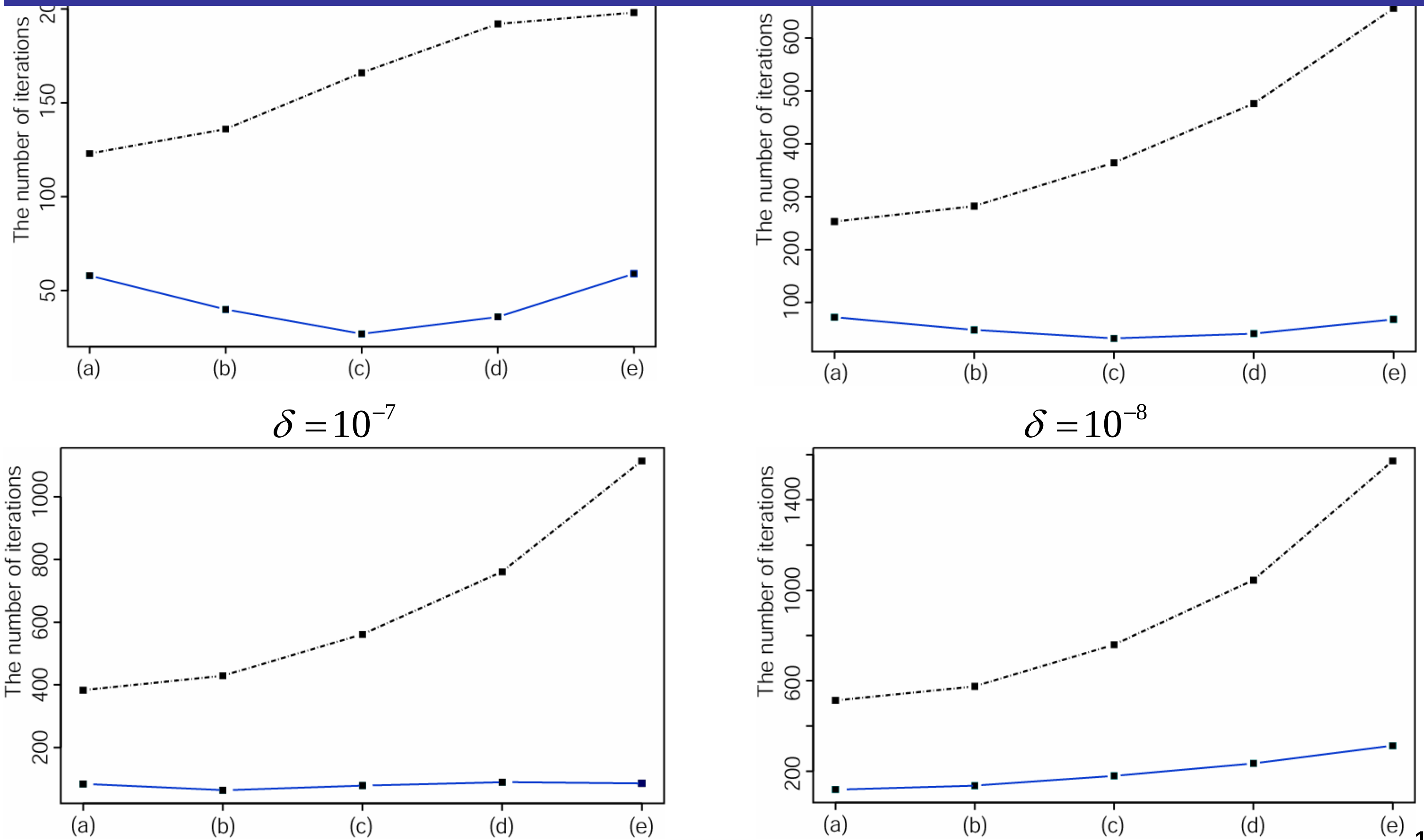
δ		(a)	(b)	(c)	(d)	(e)
10^{-5}	EM	123	136	166	192	198
	ε-accelerated EM	58	40	27	36	59
10^{-6}	EM	253	282	364	476	656
	ε-accelerated EM	72	48	32	41	68
10^{-7}	EM	383	429	561	761	1114
	ε-accelerated EM	84	64	79	90	86
10^{-8}	EM	513	575	759	1045	1572
	ε-accelerated EM	119	136	179	234	313

δ is a desired accuracy

For all data,
the EM algorithm requires the numbers of iterations
more than roughly 3-10 times of the ε -accelerated EM algorithm.

Plots of the numbers of iterations

The EM algorithm increases lineally the numbers of iterations as the data (a) to (e) change, while there are few changes in the numbers of iterations for the ε -accelerated EM algorithm and its convergence is significantly faster.



EM=dash line, ε -accelerated EM =solid line