Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Graduate School of Mathematical Sciences, University of Tokyo

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm

Example

Asymptotic Theory

Results

Regularity Condition

The Gibbs sampler

Example

Asymptotic Theory

Results

Regularity Condition

Conclusion

- This talk is concerned with asymptotic properties for the EM and the Gibbs.

Introduction

What is asymptotic?

- This talk is concerned with asymptotic properties for the EM and the Gibbs.
- The meaning of asymptotic is that not only the number of iterations, but also the number of observations goes to infinity.

Introduction
Summary of the talk

- We address some asymptotic properties of the EM and the Gibbs.

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition

The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition

Conclusion

- We address some asymptotic properties of the EM and the Gibbs.
- Using this asymptotic properties, we validate some well known facts.

Introduction
Summary of the talk

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We address some asymptotic properties of the EM and the Gibbs.
- Using this asymptotic properties, we validate some well known facts.
- This type of arguments may be useful for constructing new Monte Carlo algorithms.

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Example: Haplotype Estimation Problem
Description

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We explain the motivation for the EM through an Example.

Example: Haplotype Estimation Problem
Description

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We explain the motivation for the EM through an Example.
- Let $Y_{i,j} : \Omega \to \{0,1\}^2$ $(i = 1, \ldots, n, j = 1, 2)$ be IID rv's such that

$$P_\theta(Y_{i,j} = (0,0)) = \theta_{0,0}, P_\theta(Y_{i,j} = (1,0)) = \theta_{1,0},$$
$$P_\theta(Y_{i,j} = (0,1)) = \theta_{0,1}, P_\theta(Y_{i,j} = (1,1)) = \theta_{1,1}.$$

Example: Haplotype Estimation Problem

Description

- We explain the motivation for the EM through an Example.
- Let $Y_{i,j} : \Omega \to \{0,1\}^2$ $(i = 1, \ldots, n, j = 1, 2)$ be IID rv's such that

$$P_\theta(Y_{i,j} = (0,0)) = \theta_{0,0}, P_\theta(Y_{i,j} = (1,0)) = \theta_{1,0},$$
$$P_\theta(Y_{i,j} = (0,1)) = \theta_{0,1}, P_\theta(Y_{i,j} = (1,1)) = \theta_{1,1}.$$

- Assume that we do not observe $(Y_{i,j})$, but $(X_i = Y_{i,1} + Y_{i,2})$.

Example: Haplotype Estimation Problem
Description

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- Let $Y_i = \{Y_{i,1}, Y_{i,2}\}$. Consider four observations.

$$Y_1 = \{(0,0),(0,0)\}, \qquad Y_2 = \{(1,1),(0,0)\}$$
$$Y_3 = \{(1,1),(1,0)\}, \qquad Y_4 = \{(0,1),(1,0)\}$$

Example: Haplotype Estimation Problem
Description

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- Let $Y_i = \{Y_{i,1}, Y_{i,2}\}$. Consider four observations.

$$Y_1 = \{(0,0),(0,0)\}, \qquad Y_2 = \{(1,1),(0,0)\}$$
$$Y_3 = \{(1,1),(1,0)\}, \qquad Y_4 = \{(0,1),(1,0)\}$$

- Then,

$$X_1 = (0,0), \qquad X_2 = (1,1)$$
$$X_3 = (2,1), \qquad X_4 = (1,1)$$

Example: Haplotype Estimation Problem
Description

- Let $Y_i = \{Y_{i,1}, Y_{i,2}\}$. Consider four observations.

$$Y_1 = \{(0,0),(0,0)\}, \qquad Y_2 = \{(1,1),(0,0)\}$$
$$Y_3 = \{(1,1),(1,0)\}, \qquad Y_4 = \{(0,1),(1,0)\}$$

- Then,

$$X_1 = (0,0), \qquad X_2 = (1,1)$$
$$X_3 = (2,1), \qquad X_4 = (1,1)$$

- Let $T_{a,b}$ denote $\sum_{i,j} 1_{\{Y_{i,j}=(a,b)\}}$ $(a,b = 0,1)$. That is $T_{0,0} = 3$, $T_{0,1} = 1$, $T_{1,0} = 2$, $T_{1,1} = 2$.

Example: Haplotype Estimation Problem

The EM algorithm

E-Step: Calculate
$Q(\theta^{i-1}, \vartheta) = 2\sum_{a,b=0,1} \log \vartheta_{a,b} \mathbf{E}_{\theta^{i-1}}[T_{a,b}|X_1, \ldots, X_n]$
M-Step: Set $\theta^i = \arg\max_\vartheta Q(\theta^{i-1}, \vartheta)$.

The M-step is,

$$\theta^i_{a,b} = \mathbf{E}_{\theta^{i-1}}[\frac{T_{a,b}}{2n}|X_1, \ldots, X_n].$$

Example: Haplotype Estimation Problem

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- This EM algorithm works well when we take a good initial point θ^0.

Example: Haplotype Estimation Problem

- This EM algorithm works well when we take a good initial point θ^0.
- Does the sequence converge to MLE $\hat{\theta}_n$?

Example: Haplotype Estimation Problem

- This EM algorithm works well when we take a good initial point θ^0.
- Does the sequence converge to MLE $\hat{\theta}_n$?
- Computer simulation tells us good information. But, how about theoretical validation?

Example: Haplotype Estimation Problem

- We have,

$$\frac{1}{n}\partial_\theta \log L_n(\theta^{i-1}) = I_{1,2}(\theta^{i-1})(\theta^i - \theta^{i-1}),$$

where $L_n(\theta)$ is the likelihood and $I_{1,2}(\theta)$ is the Fisher information of (X, Y).

Example: Haplotype Estimation Problem

- We have,

$$\frac{1}{n}\partial_\theta \log L_n(\theta^{i-1}) = I_{1,2}(\theta^{i-1})(\theta^i - \theta^{i-1}),$$

where $L_n(\theta)$ is the likelihood and $I_{1,2}(\theta)$ is the Fisher information of (X, Y).

- Under some regularity conditions,

$$\frac{1}{n}(\partial_\theta \log L_n(\theta^{i-1}) - \partial_\theta \log L_n(\hat{\theta}_n)) \sim -I(\theta^{i-1})(\theta^{i-1} - \hat{\theta}_n).$$

Example: Haplotype Estimation Problem

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We have,

$$\frac{1}{n}\partial_\theta \log L_n(\theta^{i-1}) = I_{1,2}(\theta^{i-1})(\theta^i - \theta^{i-1}),$$

where $L_n(\theta)$ is the likelihood and $I_{1,2}(\theta)$ is the Fisher information of (X, Y).

- Under some regularity conditions,

$$\frac{1}{n}(\partial_\theta \log L_n(\theta^{i-1}) - \partial_\theta \log L_n(\hat{\theta}_n)) \sim -I(\theta^{i-1})(\theta^{i-1} - \hat{\theta}_n).$$

- Therefore, by easy calculation,

$$\theta^i - \hat{\theta}_n \sim J(\theta^{i-1})(\theta^{i-1} - \hat{\theta}_n),$$

where $J(\theta) = I - I_{1,2}(\theta)^{-1} I(\theta_0)$.

Example: Haplotype Estimation Problem

- Under some regularity conditions, we have

$$(\theta^i - \hat{\theta}_n) \sim J(\hat{\theta}_n)^i(\theta^0 - \hat{\theta}_n) \sim 0.$$

Example: Haplotype Estimation Problem

- Under some regularity conditions, we have

$$(\theta^i - \hat{\theta}_n) \sim J(\hat{\theta}_n)^i(\theta^0 - \hat{\theta}_n) \sim 0.$$

- A similar expansion as above is well known.

Example: Haplotype Estimation Problem

- Under some regularity conditions, we have

$$(\theta^i - \hat{\theta}_n) \sim J(\hat{\theta}_n)^i(\theta^0 - \hat{\theta}_n) \sim 0.$$

- A similar expansion as above is well known.
- We want to know what is the regularity condition. I tried to find any paper related to the validation of the above expansion.

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

Example: Haplotype Estimation Problem

- Under some regularity conditions, we have

$$(\theta^i - \hat{\theta}_n) \sim J(\hat{\theta}_n)^i(\theta^0 - \hat{\theta}_n) \sim 0.$$

- A similar expansion as above is well known.
- We want to know what is the regularity condition. I tried to find any paper related to the validation of the above expansion.
- Surprisingly, I could not find any validation for the above expansion!

Example: Haplotype Estimation Problem

- Under some regularity conditions, we have

$$(\theta^i - \hat{\theta}_n) \sim J(\hat{\theta}_n)^i(\theta^0 - \hat{\theta}_n) \sim 0.$$

- A similar expansion as above is well known.
- We want to know what is the regularity condition. I tried to find any paper related to the validation of the above expansion.
- Surprisingly, I could not find any validation for the above expansion!
- So we address the validity issue.

Some relative merit for large sample properties

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We can show that the sequence of the EM algorithm converges to the MLE. It is difficult under small sample framework.

Some relative merit for large sample properties

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We can show that the sequence of the EM algorithm converges to the MLE. It is difficult under small sample framework.
- The rate matrix determines the convergence rate. This fact is validated under large sample statistical theory.

Some relative merit for large sample properties

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm

Example

Asymptotic Theory

Results

Regularity Condition

The Gibbs sampler

Example

Asymptotic Theory

Results

Regularity Condition

Conclusion

- We can show that the sequence of the EM algorithm converges to the MLE. It is difficult under small sample framework.
- The rate matrix determines the convergence rate. This fact is validated under large sample statistical theory.
- It may be useful for other complicated algorithms.

Some drawbacks for large sample properties

- We assume a large sample size. Sometimes it is difficult to assume.

Some drawbacks for large sample properties

- We assume a large sample size. Sometimes it is difficult to assume.
- We assume that the initial point θ^0 is in a neighborhood of the true value θ_0. (A kind of moment estimator works well.)

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Asymptotic Theory for IID

- Let θ_0 be a true parameter. Then, under general assumptions,

$$\log L_{h,n} = \langle h, Z_n \rangle - \frac{1}{2}\langle h, I(\theta_0)h \rangle + o_{P^n_{\theta_0}}(1),$$

where $\log L_{h,n} = \sum \log p_{\theta_0 + hn^{-1/2}}(x_i)/p_{\theta_0}(x_i)$ and $Z_n = n^{-1/2} \sum \partial \log p_{\theta_0}(x_i)$.

Asymptotic Theory for IID

- Let θ_0 be a true parameter. Then, under general assumptions,

$$\log L_{h,n} = \langle h, Z_n \rangle - \frac{1}{2}\langle h, I(\theta_0)h \rangle + o_{P^n_{\theta_0}}(1),$$

where $\log L_{h,n} = \sum \log p_{\theta_0 + hn^{-1/2}}(x_i)/p_{\theta_0}(x_i)$ and $Z_n = n^{-1/2} \sum \partial \log p_{\theta_0}(x_i)$.

- The right hand side is maximized at $h = I(\theta_0)^{-1} Z_n$.

Asymptotic Theory for IID

- Let θ_0 be a true parameter. Then, under general assumptions,

$$\log L_{h,n} = \langle h, Z_n \rangle - \frac{1}{2}\langle h, I(\theta_0)h \rangle + o_{P^n_{\theta_0}}(1),$$

where $\log L_{h,n} = \sum \log p_{\theta_0 + hn^{-1/2}}(x_i)/p_{\theta_0}(x_i)$ and $Z_n = n^{-1/2} \sum \partial \log p_{\theta_0}(x_i)$.

- The right hand side is maximized at $h = I(\theta_0)^{-1} Z_n$.
- Using this expansion, under some regularity condition, we have

$$n^{1/2}(\hat{\theta}_n - \theta_0) \Rightarrow N(0, I(\theta_0)^{-1}).$$

Asymptotic Theory for the EM algorithm

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We take two parametric families $(\mathcal{X}, \mathcal{A}, P_\theta(dx) = p_\theta(x)dx; \theta \in \Theta)$ and $(\mathcal{Y}, \mathcal{B}, P^{2|1}_{x,\theta}(dy) = p^{2|1}_{x,\theta}(y)dy; \theta \in \Theta, x \in \mathcal{X})$.

Asymptotic Theory for the EM algorithm

- We take two parametric families $(\mathcal{X}, \mathcal{A}, P_\theta(dx) = p_\theta(x)dx; \theta \in \Theta)$ and $(\mathcal{Y}, \mathcal{B}, P^{2|1}_{x,\theta}(dy) = p^{2|1}_{x,\theta}(y)dy; \theta \in \Theta, x \in \mathcal{X})$.
- Let $p^{1.2}_s(x, y) = p_s(x)p^{2|1}_{x,s}(y)$.

Asymptotic Theory for the EM algorithm

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We take two parametric families $(\mathcal{X}, \mathcal{A}, P_\theta(dx) = p_\theta(x)dx; \theta \in \Theta)$ and $(\mathcal{Y}, \mathcal{B}, P^{2|1}_{x,\theta}(dy) = p^{2|1}_{x,\theta}(y)dy; \theta \in \Theta, x \in \mathcal{X})$.
- Let $p^{1.2}_s(x, y) = p_s(x)p^{2|1}_{x,s}(y)$.
- Let

$$Q_{s,t}(x) = \int_{\mathcal{Y}} \log \frac{p^{1,2}_t(x, y)}{p^{1,2}_s(x, y)} dP^{2|1}_{x,s}(dy)$$

and

$$Q_{s,t,n}(x^{(n)}) = \sum_{i=1}^{n} Q_{\theta_0+sn^{-1/2},\theta_0+tn^{-1/2}}(x_i).$$

Asymptotic Theory for the EM algorithm

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm

Example

Asymptotic Theory

Results

Regularity Condition

The Gibbs sampler

Example

Asymptotic Theory

Results

Regularity Condition

Conclusion

- Under some regularity conditions, $Q_{s,t,n}(x^{(n)}) = \overline{Q}_{s,t,n}(x^{(n)}) + o_{P^n_{\theta_0}}(1)$, where $\overline{Q}_{s,t,n}$ is

$$-\frac{1}{2}\langle t, I_{1,2}(\theta_0)t\rangle + \langle t, Z_n + (I_{1,2}(\theta_0) - I(\theta_0))s\rangle + c_n(s, \theta_0, x^{(n)})$$

Asymptotic Theory for the EM algorithm

- Under some regularity conditions, $Q_{s,t,n}(x^{(n)}) = \overline{Q}_{s,t,n}(x^{(n)}) + o_{P^n_{\theta_0}}(1)$, where $\overline{Q}_{s,t,n}$ is

$$-\frac{1}{2}\langle t, I_{1,2}(\theta_0)t\rangle + \langle t, Z_n + (I_{1,2}(\theta_0) - I(\theta_0))s\rangle + c_n(s, \theta_0, x^{(n)})$$

- Note that, the maximizer of $\overline{Q}_{s,t,n}$ with respect to t satisfies

$$t - \mu_n = J(\theta_0)(s - \mu_n),$$

where $\mu_n = I(\theta_0)^{-1} Z_n$.

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

The Result

Theorem (Convergence Theorem for the EM)

Under some regularity conditions, for any $m_n \to \infty$,

$$n^{1/2}(\theta^{m_n} - \hat{\theta}_n) = o_{P^n_{\theta_0}}(1),$$

where $\theta^0, \theta^1, \ldots$ is the sequence of the EM algorithm.

In the above result, we assume that for $\theta^0 = \theta^0_n$, for any $\epsilon > 0$, there exists $\delta > 0$ such that

$$\limsup_{n\to\infty} P^n_{\theta_0}(n^{1/2}|\theta^0_n - \theta_0| > \delta) \leq \epsilon. \tag{1}$$

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Regularity Condition

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

Assumption (Smoothness Assumption for the EM)

- *There exist some* $\epsilon > 0$*, and some* $M \in L^2(P_{\theta_0})$ *such that for any* $s, t, u, v \in B_\epsilon(\theta_0)$*, we have*

$$|Q_{s,u}(x) - Q_{t,v}(x)| \leq M(x)(|s-t|^2 + |u-v|^2)^{1/2}.$$

- *The Fisher information matrix* $I(\theta)$ *is continuous around* θ_0.

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Example

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We explain the motivation for the Gibbs through the Example, which is used in the previous section.

Example

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We explain the motivation for the Gibbs through the Example, which is used in the previous section.
- The prior distribution is $\mathrm{Dir}(\alpha)$, where $\alpha = (\alpha_{0,0}, \alpha_{0,1}, \alpha_{1,0}, \alpha_{1,1})$.

Example: Haplotype Estimation Problem

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

The Gibbs Sampler

Step i: For $j = 1, \ldots, n$, generate $y_{i,j} = (y_{i,j,1}, y_{i,j,2})$ from $P^{2|1}_{\theta^{i-1}}(dy_{i,j}|x_j)$.
Then calculate $T^i_{a,b,n} = \sum_{j=1}^n T_{a,b}(y_{i,j,1}, y_{i,j,2})$ and set $T^i_n = (T^i_{0,0,n}, T^i_{0,1,n}, T^i_{1,0,n}, T^i_{1,1,n})$.
Generate θ^i from $\mathrm{Dir}(\alpha + T^i_n)$.

Example: Haplotype Estimation Problem

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm

Example

Asymptotic Theory

Results

Regularity Condition

The Gibbs sampler

Example

Asymptotic Theory

Results

Regularity Condition

Conclusion

The above procedure defines a Markov transition kernel $F_{x^{(n)},\theta,n}(d\varphi) = f_{x^{(n)},\theta,n}(\varphi)d\varphi$. We have

$$f_{x^{(n)},\theta,n}(\varphi) \geq \frac{\Gamma(2n + \sum_{a,b=0}^{1} \alpha_{a,b})}{\prod_{a,b=0}^{1} \Gamma(2n + \alpha_{a,b})} \prod_{a,b=0}^{1} \varphi_{a,b}^{2n+\alpha_{a,b}-1},$$

Moreover, we have

$$\|F_{x^{(n)},n} - F^{m}_{x^{(n)},\theta,n}\|_{\mathrm{TV}} \leq \rho^{m}$$

where $F_{x^{(n)},n}$ is the posterior, and $\rho = 1 - \Gamma(2n + \sum_{a,b=0}^{1} \alpha_{a,b})/\Gamma(8n + \sum_{a,b=0}^{1} \alpha_{a,b})$

Some relative merit for large sample properties

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition

The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition

Conclusion

- We approximate the Gibbs sampler by a simple transition kernel, which is defined by simple statistics.

Some relative merit for large sample properties

- We approximate the Gibbs sampler by a simple transition kernel, which is defined by simple statistics.
- Therefore, we can measure the convergence rate of the Gibbs sampler by the simple statistics.

Some relative merit for large sample properties

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We approximate the Gibbs sampler by a simple transition kernel, which is defined by simple statistics.
- Therefore, we can measure the convergence rate of the Gibbs sampler by the simple statistics.
- In some situation, it is reasonable to consider large sample theory.

Some drawbacks for large sample properties

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We assume a large sample. Sometimes it is difficult to assume.

Some drawbacks for large sample properties

- We assume a large sample. Sometimes it is difficult to assume.
- We assume that the initial guess θ^0 is in a neighborhood of the true value θ_0. A kind of moment estimator works well.

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Asymptotic Theory for the Gibbs sampler

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline

Introduction

The EM algorithm

Example

Asymptotic Theory

Results

Regularity Condition

The Gibbs sampler

Example

Asymptotic Theory

Results

Regularity Condition

Conclusion

- We showed that the sequence of the Gibbs sampler tends to the following Markov chain:

Asymptotic Theory for the Gibbs sampler

- We showed that the sequence of the Gibbs sampler tends to the following Markov chain:
- The Markov chain $h_1, h_2, \ldots$ satisfies

$$h_i - \mu_n = J(\theta_0)(h_{i-1} - \mu_n) + \epsilon_i$$

where $\epsilon_i = N(0, 2I_{1,2}(\theta_0)^{-1} - I_{1,2}(\theta_0)^{-1} I(\theta_0) I_{1,2}(\theta_0)^{-1})$.

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

The Result

Asymptotic Properties for Monte Carlo Methods
Kengo, KAMATANI
Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

Theorem (Convergence Theorem for the Gibbs)

Under some regularity conditions, we have

$$n^{1/2}|\tilde{\theta}^{I_n} - \tilde{\theta}_n| = o_{P^n_{\theta_0}}(1),$$

Note that $\tilde{\theta}^{I_n}$ is the empirical median of the sequence of the Gibbs sampler such that $\theta^0 = \theta^0_n \sim \nu_{x^{(n)},n}$. We assume that for any $\epsilon > 0$, there exists $\delta > 0$,

$$\limsup_{n\to\infty} P^n_{\theta_0}(n^{1/2}|\theta^0_n - \theta_0| > \delta) \leq \epsilon. \quad (2)$$

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Regularity Condition

Smoothness of the model

Assumption (Smoothness Assumption for the Gibbs)

- *For some* $\epsilon > 0$*, there exists a constant* $C > 0$ *and for any* $s, t \in B_\epsilon(\theta_0)$*, we have*

$$H(P^{2|1}_{x,s}, P^{2|1}_{x,t}) \leq M(x)|s - t|,$$

where $M(x) \in L^2(P_{\theta_0})$*.*

- *The Fisher information matrix* $I(\theta)$ *is continuous around* θ_0*.*

Regularity Condition

Existence of uniformly consistent test

Assumption (Existence of Uniformly Consistent Test)

There exists an integer N and a test $\omega_n = \omega_n(x_1, \ldots, x_n)$ on $(\mathcal{X}^n, \mathcal{A}^n)$, such that there exists a constant $\epsilon_0 \in (0, 1/2)$ and a compact subset K of Θ such that

$$P_{\theta_0}^{(n)}(\omega_n) \leq \epsilon_0, P_{\theta}^{(n)}(1 - \omega_n) \leq \epsilon_0 \ (\forall \theta \in K^c).$$

Outline

1 Introduction

2 The EM algorithm
- Example
- Asymptotic Theory
- Results
- Regularity Condition

3 The Gibbs sampler
- Example
- Asymptotic Theory
- Results
- Regularity Condition

4 Conclusion

Conclusion

Asymptotic Properties for Monte Carlo Methods

Kengo, KAMATANI

Outline
Introduction
The EM algorithm
Example
Asymptotic Theory
Results
Regularity Condition
The Gibbs sampler
Example
Asymptotic Theory
Results
Regularity Condition
Conclusion

- We showed convergence theorems for the EM and the Gibbs.

Conclusion

- We showed convergence theorems for the EM and the Gibbs.
- These two algorithms are approximated by simple algorithms. Using this approximation, we can compare different algorithms.

Conclusion

- We showed convergence theorems for the EM and the Gibbs.
- These two algorithms are approximated by simple algorithms. Using this approximation, we can compare different algorithms.
- It may be useful for constructing a new algorithm, or studying efficiency of existing methods.