一様でない遷移確率を用いた焼き鈍し法

宗久・丹沢研究室
T04kf022
鈴木悠也

焼き鈍し法

- 焼きなまし法とは最大値探索問題の解法の1つであり、温度を示すパラメーターを徐々に下げることにより最適な解を求めるものである。

研究目的

今回の研究目的は温度が一定という条件の下でも温度を徐々に下げていく場合と同様に一様でない遷移確率を実現し最適な解を求めることである。

焼き鈍し法のアプローチ

集合Sの要素を変数とする目的関数fの最大値探索問題

$$\downarrow$$

fをエネルギーとみなし、その最大値を与える平衡状態 ${}^t b$ を求める問題と考える

収束状態が最適分布となるようなマルコフ連鎖をつくり、適当な初期状態から状態遷移させる

マルコフ連鎖

次の状態が現在の状態だけに依存するような過程のうち時間のパラメータtが離散的で状態空間Sが有限であるときのもの

過程が時間に依存しないとき

　一様なマルコフ連鎖　であるという

過程が時間に依存するとき

　一様でないマルコフ連鎖　であるという

状態遷移確率P

$i, j \in S$のとき状態iからjに遷移する
状態遷移確率は$P\left(X(t+1)=j \middle| X(t)=i\right)$
と表される

遷移確率行列M

i行j列の成分$M(i,j;t) \equiv P\left(X(t+1)=j \middle| X(t)=i\right)$が

$$M(i,j;t) \geq 0 \quad かつ \quad \sum_{j \in S} M(i,j;t) = 1$$

を満たすとき$M(i,j;t)$をi行j列の成分とする

行列Mをマルコフ連鎖の遷移確率行列という

チャップマン・コルモゴルの方程式

$M(i,j;t,t+m) \equiv P\left(X(t+m)=j \middle| X(t)=i\right)$を

x行y列の要素とする行列は下の式で表される

- 一様でないマルコフ連鎖

$$M(t,t+m)=M(t)\cdot M(t+1)\cdots M(t+(m-1))$$

- 一様なマルコフ連鎖

$$M(t,t+m)=M^m$$

マルコフ連鎖のエルゴード定理

マルコフ連鎖の遷移確率行列Mが既約かつ非周期的であるとき初期状態分布cによらず時間経過とともに平衡分布bに収束する

$$\lim_{m \to \infty} {}^t c M^m = {}^t b$$

- 一様でないマルコフ連鎖でも以下の式が成り立つ

$$\lim_{m\to\infty} {}^t c M(0)\cdot M(1)\cdots M(m) = {}^t b$$

Mは以下の式を満たすものである

$${}^t b M(k) = {}^t b$$

$${}^t\vec{b}: \text{平衡分布}$$

Mを掛けることにより誤差が等しいか小さくなることの証明

$$\left\| {}^t cM - {}^t b \right\| \leq \left\| {}^t c - {}^t b \right\|$$

$$\left\| {}^t c - {}^t b \right\| = \sum_i \left| c_i - b_i \right|$$

$${}^t bM = {}^t b$$

$$\sum_i M_{ji} = 1$$

${}^t c$: 実際の状態分布

$$M_{ji} \geq 0$$

$$\left\| {}^t cM - {}^t b \right\| \le \left\| {}^t c - {}^t b \right\|$$

$$\text{左辺} = \left\| {}^t cM - {}^t bM \right\| = \left\| \left({}^t c - {}^t b \right) M \right\| = \sum_i \left| \sum_j \left(c - b \right)_j M_{ji} \right|$$

$$\sum_i \left| \sum_j \left(c - b \right)_j M_{ji} \right| \le \sum_i \sum_j \left| \left(c - b \right)_j M_{ji} \right|$$

$$\text{右辺} = \sum_j \left| \left(c - b \right)_j \right| \sum_i M_{ji} = \sum_j \left| \left(c - b \right)_j \right| = \left\| {}^t c - {}^t b \right\|$$

等号が成立するのはすべての成分に対して

$$\left(c-b\right)_j M_{ji} \geq 0 \quad \text{か} \quad \left(c-b\right)_j M_{ji} \leq 0$$

が成り立つときである

ただし $c=b$ 以外では

$$\sum_i c_i = 1, c_i \geq 0, \sum_i b_i = 1, b_i \geq 0$$

なのですべての成分に対して $\left(c-b\right)_j \leq 0$

$$\left(c-b\right)_j \geq 0$$

が成り立つことはないので

等号が成立するには

$M_{ji} = 0$　が存在する必要がある

しかし

$$\left[M(1)M(2)\cdots M(n)\right]_{ij} \neq 0$$

nは数回程度で上の式は成り立つので状態がすでに平衡状態に達しているとき以外ではほぼ等号は成立はしない

マルコフ連鎖を作るための平衡分布として
ボルツマン分布がある

$$b_i(T) = \frac{1}{\sum_{i \in S} \exp\left(f_i / T \right)} \exp\left(f_i / T \right)$$

order-3だまし問題を使用する

関数fの値は表1の値を使用する

表１：関数ｆの値

000	001	010	011	100	101	110	111
28	26	24	0	14	0	0	30

マルコフ連鎖の設計

受理行列Aと遷移行列Q（t）を用いてi行j列の成分M(i,j;t)が次式で表される行列Mを作る

$$M(i,j;t) \equiv \begin{cases} Q(i,j;t)A(i,j) & for \quad i \neq j \\ 1 - \sum_{j \neq x} Q(i,j;t)A(i,j) & for \quad i = j \end{cases}$$

マルコフ連鎖の設計（2）

行と列の要素数が集合Sの要素数に等しく、
i行j列の成分A(i,j)が次式で表されるよう
な行列Aを考える

$$A(i,j) \equiv W\left(g(j;T) / g(i;T) \right)$$

$$W(s) \equiv sW(1/s) \quad for \quad \forall s \in (0,\infty)$$

$$W(s) = \begin{cases} 1 & s \geq 1 \\ g(j;T)/g(i;T) & s < 1 \end{cases}$$

マルコフ連鎖の設計(3)

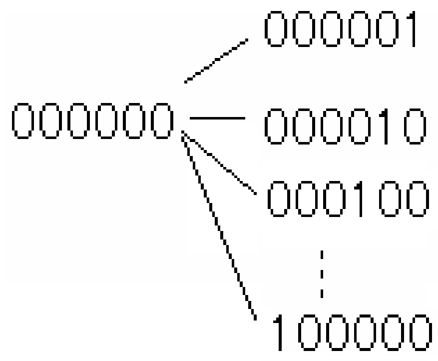
受理行列Aと等しい要素数の行と列を持ち、i行j列の成分Q(i,j;t)が次式を満たすような既約で対称な遷移行列Qを考える

$$Q(i,j;t) = \begin{cases} 0 & for \quad i = j \\ Q(j,i;t)(\geq 0) & for \quad i \neq j \end{cases}$$

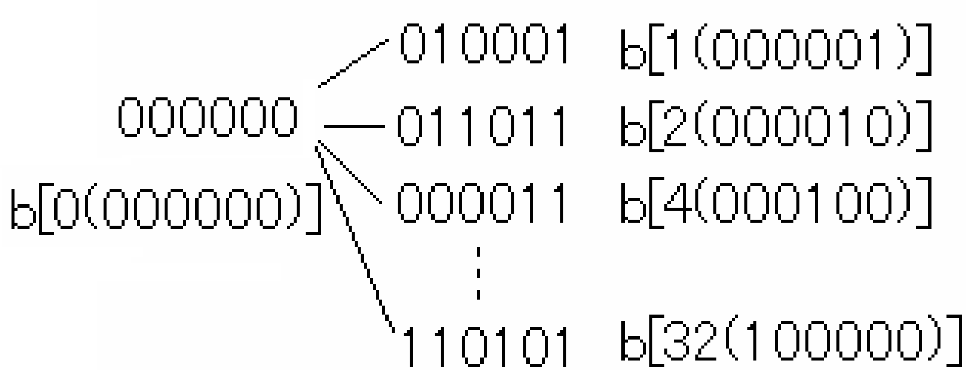
$$\sum_{y \in S} Q(i,j;t) = 1, \exists n : Q^n(i,j;t) \geq 0$$

今回の実験では6個の状態が次の遷移候補として選ばれるのでその確率は1/6となる

- 一様なマルコフ連鎖



- 一様でないマルコフ連鎖



ボルツマン分布への収束証明実験

一様なマルコフ連鎖

1ビットの突然変異でMの確率により状態遷移していく

一様でないマルコフ連鎖

乱数により6個の状態を選びMの確率により

状態遷移していく

それぞれの状態が現れる回数を全回数で割ったものがそれぞれの状態が現れる確率となる

初期設定

- 温度　T=10
- 状態遷移数　3000000回
- 初期状態　００１０１０

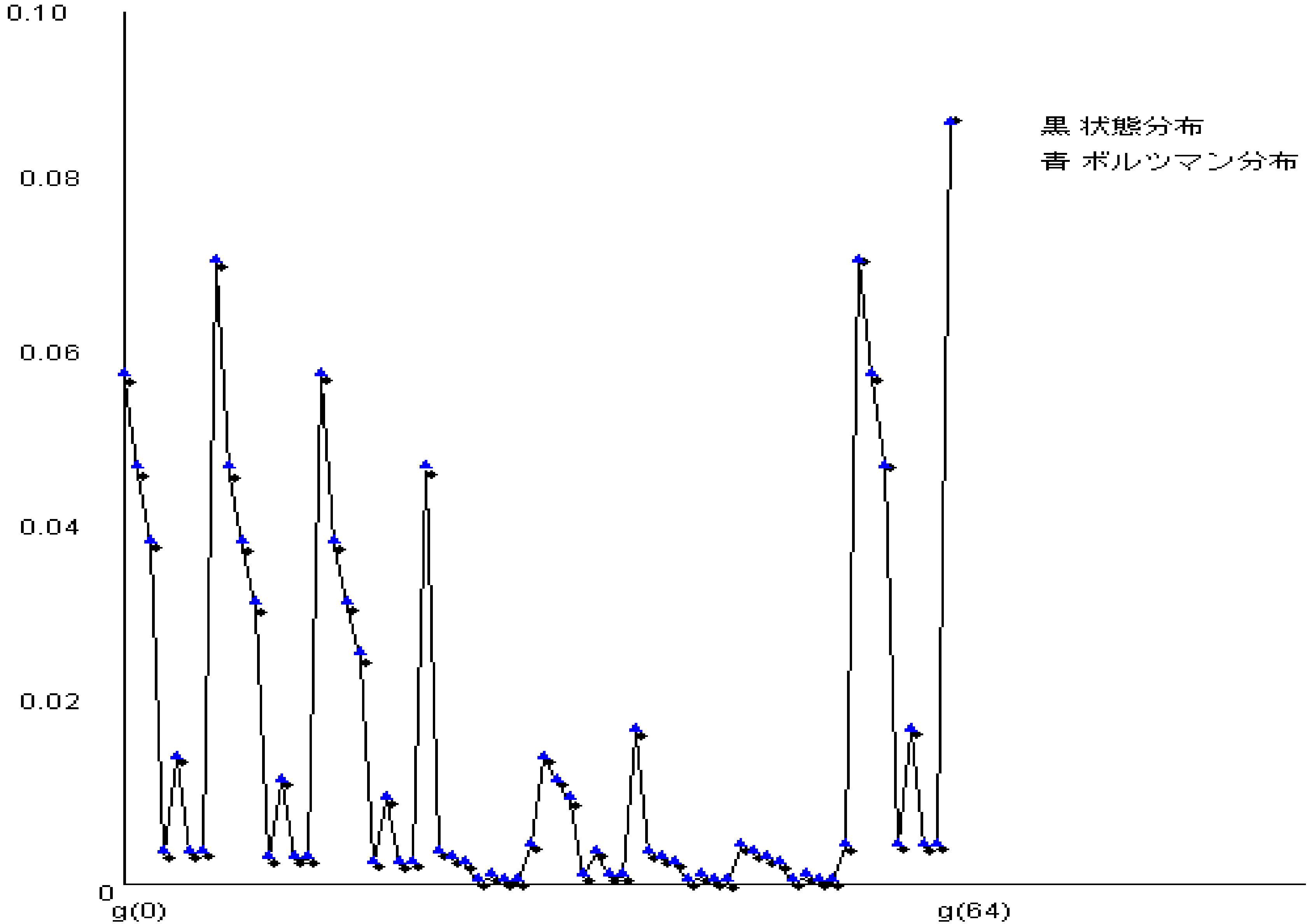


図２：一様なマルコフ連鎖

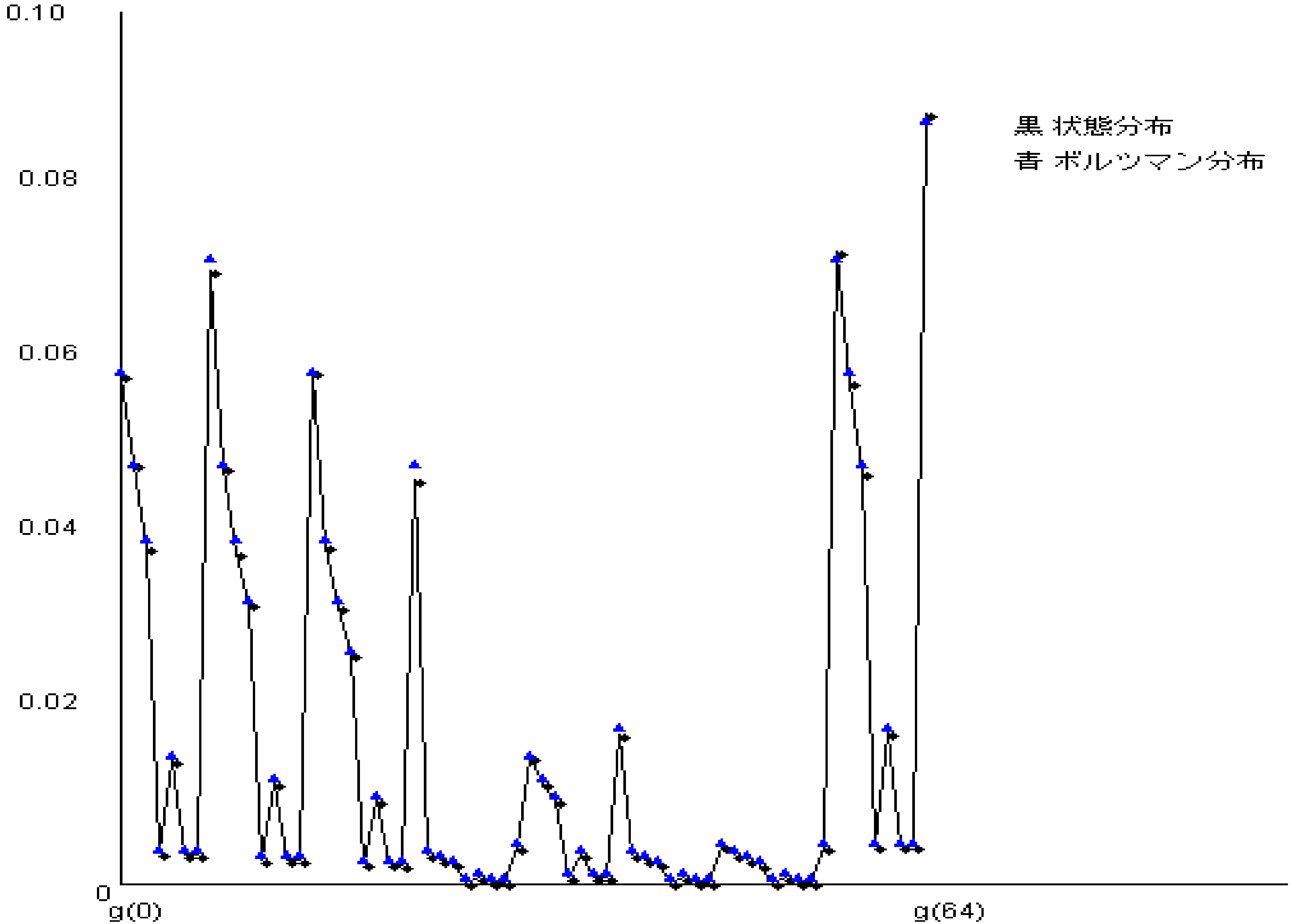


図３：一様でないマルコフ連鎖

ボルツマン分布との誤差の比較実験

- 一様なマルコフ連鎖

$$e(m) = \left\| {}^{t}cM^{m} - {}^{t}b \right\|$$

- 一様でないマルコフ連鎖

$$e(m) = \left\| {}^{t}cM(0)M(1)\cdots M(m) - {}^{t}b \right\|$$

初期設定

- 温度　T=10
- m = 200 回
- 初期状態　００１０１０

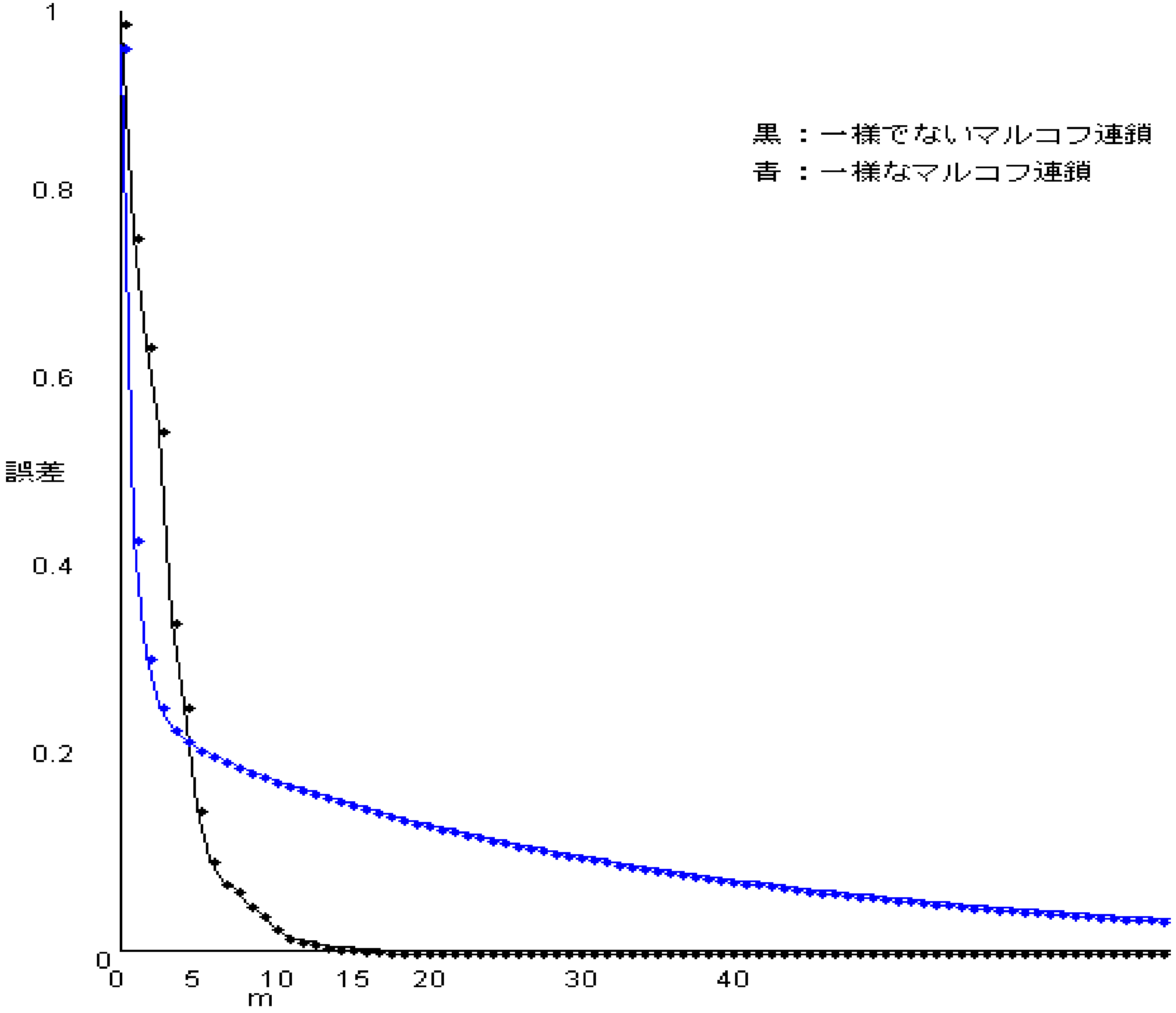


図１：誤差とｍとの関係

Bit数を変えたときの誤差の比較実験

６ビット，１２ビット，１８ビット，２４ビットで初期状態のfの値が最小、最大、平均でそれぞれ実験を行う

表２：Mを50回掛けたときの誤差の値

$6bit$

f	一様なマルコフ連鎖	一様でないマルコフ連鎖
0	0.0051	0.0000022
30	0.11	0.0000059
60	0.18	0.000017

$18bit$

f	一様なマルコフ連鎖	一様でないマルコフ連鎖
0	0.0019	0.0047
90	0.043	0.0051
180	0.20	0.78

$12bit$

f	一様なマルコフ連鎖	一様でないマルコフ連鎖
0	0.0071	0.0029
60	0.0084	0.0034
120	0.21	0.24

$24bit$

f	一様なマルコフ連鎖	一様でないマルコフ連鎖
0	0.00039	0.0024
120	0.020	0.0051
240	0.18	0.95

考察

- 一様なマルコフ連鎖では次の状態遷移候補のfの値は初期状態のfの値との差が30以内におさまるのでbitが大きくなるにつれ初期状態のfの値が大きい場合には一様でないマルコフ連鎖に比べ適していると推測できる

まとめ

- 一様でないマルコフ連鎖の場合でもボルツマン分布に収束する

参考文献

- 長尾智晴：最適化アルゴリズム，P209，昭晃堂（2000）
- 平早哲郎：数学的保証をもつ遺伝的アルゴリズムの改良，山梨大学修士論文（2006）

$\vec{{}^{t}c}$: 実際の状態分布

$\vec{{}^{t}b}$: 最適分布

${}^{t}c$: 実際の状態分布

${}^{t}b$: 平衡分布

^{t}b : 平衡分布

$24bit$

^{t}c : 実際の状態分布

平衡分布

一様なマルコフ連鎖Mのもとでb M＝bが成り立つような状態分布bが存在するなら

↓

その状態分布は以後時間的に変化しない

このような状態分布bは平衡分布と呼ばれる

結果

bitが小さいときは一様でないマルコフ連鎖のほうが初期状態によらず誤差が小さくなるがbitが大きくなるにつれ初期状態のfの値が大きいと平衡分布に収束に時間がかかる

誤差が0.01以下となるmの最小値

表２：mの最小値

初期状態	一様なマルコフ連鎖	一様でないマルコフ連鎖
000000	136	16
001010	134	17
010100	128	17
011110	50	16
101000	122	14
110010	153	14
110111	119	14
111100	154	14
111111	168	25

今後の展開

- 一様でないマルコフ連鎖の理論をさらに研究し、このプログラムを拡張していく