

TR-A-0165

アニメーションスケジュールの定数倍加速

高畠 一哉

1993. 3.15

ATR 視聴覚機構研究所

〒619-02 京都府相楽郡精華町光台 2-2 ☎07749-5-1411

ATR Auditory and Visual Perception Research Laboratories

2-2, Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02 Japan

Telephone: +81-7749-5-1411

Facsimile: +81-7749-5-1408

アニーリングスケジュールの定数倍加速
Constant Times Acceleration of
Annealing Schedule

高島 一哉
Kazuya Takabatake

(株) ATR視聴覚機構研究所, 京都府
ATR Auditory and Visual Perception Research
Laboratories, Kyoto-fu, 619-02 Japan

あらまし シミュレーテッドアニーリング法で用いられるアニーリングスケジュール $T(t)=c/\ln(t+b)$ (b, c はある定数) によるマルコフ連鎖は強エルゴード性を持つことが知られているがこのアニーリングスケジュールは収束が遅いと言われている。

本論文ではこの時刻を任意の定数倍したアニーリングスケジュール $T(t)=c/\ln(at+b)$ によるマルコフ連鎖も強エルゴード性を持つことを示し収束の速さの理論的評価を示す。さらに一般の時間的・非一様な弱(強)エルゴード性を持つマルコフ連鎖の場合でも比較的緩やかな条件下で任意の定数 $a \in \{1, 2, \dots\}$ に対して次の条件(1), (2)を同時に満たす推移行列の部分列がとれることを示す。

(1) その部分列を推移行列とするマルコフ連鎖も弱(強)エルゴード性を持つ。

(2) 部分列の第 t 項を元の推移行列の時系列の第 $h(t)$ 項とすると $t \rightarrow \infty$ のとき $h(t)/t \rightarrow a$ である。

このことはアニーリングスケジュールをある意味でいくらかでも加速できることを示している。

キーワード シミュレーテッドアニーリング, アニーリングスケジュール, 時間的・非一様なマルコフ連鎖, 弱エルゴード性, 強エルゴード性

1. まえがき

近年最適化問題の解法としてシミュレーテッドアニーリング法^{[3][6][9][10][11]}が提案され様々な問題に応用されている。本論文では有限状態離散時間マルコフ連鎖を用いたシミュレーテッドアニーリング法を扱う。系の取り得る状態の集合を $S := \{1, 2, \dots, N\}$ とする。 S 上の確率分布は点確率測度 $\Pr\{i\}$ を用いて $(\Pr\{1\}, \Pr\{2\}, \dots, \Pr\{N\})$ のように $1 \times N$ ベクトルの形で表すことにする。系が時刻 t で状態 i にいる確率を $p_i(t)$, 時刻 t での確率分布を $p(t) := (p_1(t), p_2(t), \dots, p_N(t))$,

時刻 t での推移行列を $P(t)$ とすると $p(t+1) = p(t)P(t)$ である。

シミュレーテッドアニーリング法はエネルギーと呼ばれる関数 $E: S \rightarrow \mathbb{R}$ があるとき, エネルギーの低い状態 i で高い確率 e^{*}_i を持つような確率分布 $e^* := (e^{*}_1, e^{*}_2, \dots, e^{*}_N)$ に $p(t)$ が収束するようなマルコフ連鎖を考え, このマルコフ連鎖のシミュレーションを行うことにより高い確率でエネルギーの低い状態を得ようとするものである。実際 $p(t)$ が式(1)の e^* に収束するような手法が知られている^{[4][6][7][9]}。

$$\begin{cases} e^* = (e^{*}_1, e^{*}_2, \dots, e^{*}_N) \\ e^{*}_i = \begin{cases} 1/|S^*| & i \in S^* \\ 0 & \text{else} \end{cases} \\ S^* := \{i \mid \forall j \in S [E(i) \leq E(j)]\} \end{cases} \quad (1)$$

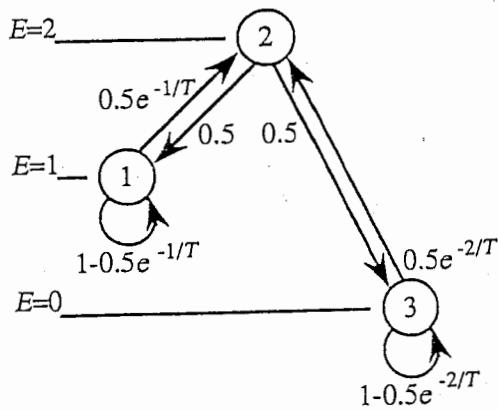
これによれば十分長い時間をかければいくらかでも高い確率で最小エネルギー状態が得られることになる。

式(2)の推移行列 $P = [P_{ij}]$ を持つ時間的一様なマルコフ連鎖が既約で非周期的^[2]であるとき任意の初期分布に対して $p(t)$ は式(3)の分布 π に収束することが知られている^{[9][10][11]}。

$$\begin{cases} P_{ij} = g_{ij} a_{ij} \quad (i \neq j) \\ g_{ij} \geq 0 \\ \sum_j g_{ij} = 1 \\ g_{ij} = g_{ji} \\ g_{ii} = 0 \\ a_{ij} = \min\{1, e^{(E(i) - E(j))/T}\} \\ P_{ii} = 1 - \sum_{j \neq i} P_{ij} \end{cases} \quad (2)$$

$$\begin{cases} \pi = (\pi_1, \pi_2, \dots, \pi_N) \\ \pi_i = \frac{e^{-E(i)/T}}{\sum_j e^{-E(j)/T}} \end{cases} \quad (3)$$

ここで $T > 0$ は温度と呼ばれるパラメータである。 T が小さいほど π は式(1)の e^* に近づくので定常状態で系は高い確率で最小エネルギー状態にあることになる反面, 系が定常状態に達するには長い時間がかかる。このことを具体例で見よう。



$$P = \begin{pmatrix} 1-0.5e^{-1/T} & 0.5e^{-1/T} & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 0.5e^{-2/T} & 1-0.5e^{-2/T} \end{pmatrix}$$

図1 推移行列例
Fig.1 Example of transition matrix

$T=0.5$ の場合の方が定常分布 π が最小エネルギー状態で確率1を持つ分布 $(0, 0, 1)$ に近いが $p(t)$ の π への収束は遅い。

そこで T を時刻 t により変化させることが行われる。 T の値は始めは大きく、そして徐々に小さくしてゆく。これはアニーリングスケジュールと呼ばれる。

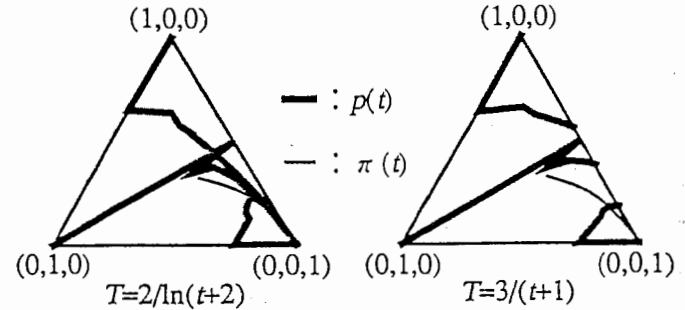


図3 $p(t)$ ($0 \leq t$) の軌跡
Fig.3 Locus of $p(t)$ ($0 \leq t$)

アニーリングスケジュールは初期分布の如何に関わらず $p(t)$ が e^* に収束する性質 (強エルゴード性) を持つように定められていることが望ましい。図3は図1の推移行列を用い $T=2/\ln(t+2)$ とした場合と $T=3/(t+1)$ とした場合について $p(0)=(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ から出発した $\pi(t)$ および $p(t)$ の軌跡の軌跡を示すものである。前者は強エルゴード性を持つアニーリングスケジュールの例、後者は強エルゴード性を持たないアニーリングスケジュールの例である。前者のような $T=c/\ln(t+b)$ の形のアニーリングスケジュールが強エルゴード性を持つことは文献[4][6][7]等により明らかにされている。本論文ではまずこのアニーリングスケジュールの時刻を任意の定数倍つまり $T=c/\ln(at+b)$ としても強エルゴード性が保たれることについて言及しその収束の速さについて理論的な評価を与える。

次に式(2)を離れたより一般的なシミュレーテッドアニーリング法についてもこのような時刻の定数倍による加速が可能かを考察する。ここではシミュレーテッドアニーリング法を単に強エルゴード性あるいはそれより弱い弱エルゴード性を持つ時間的非一様なマルコフ連鎖としてとらえ、推移行列の時系列によるマルコフ連鎖が弱(強)エルゴード性を持つときその推移行列の時系列の部分列によるマルコフ連鎖も弱(強)エルゴード性を保つように部分列をとることを考える。元の推移行列の時系列が $P(0), P(1), \dots, P(t), \dots$ であり得られた部分列が $P(h(0)), P(h(1)), \dots, P(h(t)), \dots$ であるとき $h(t)/t \rightarrow a$ ($t \rightarrow \infty$) ならばアニーリングス

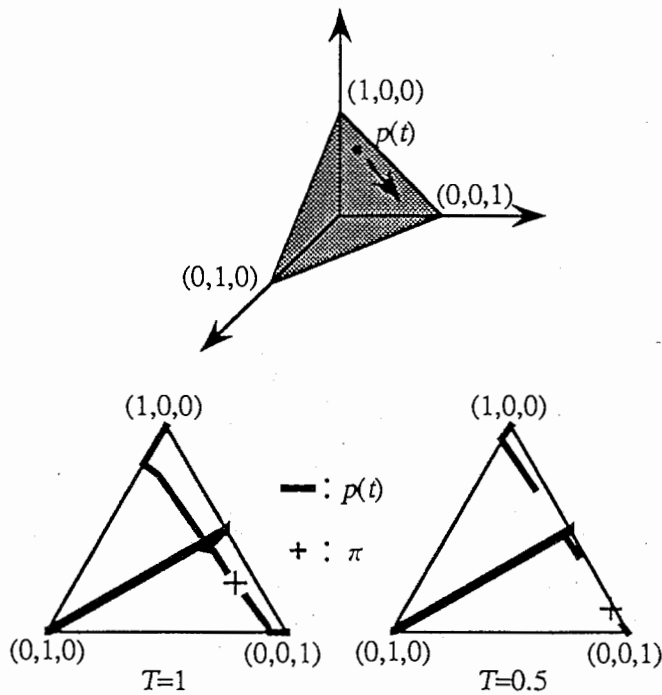


図2 $p(t)$ ($0 \leq t \leq 10$) の軌跡
Fig.2 Locus of $p(t)$ ($0 \leq t \leq 10$)

図1は式(2)の推移行列の一例、図2は図1の推移行列 P によって $p(0)$ を $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ にとったときの $t=10$ までの $p(t)$ の軌跡である。 $T=1$ の場合に比べ

スケジュールはa倍加速されたものと考えられよう。

2. 準備

以下の議論で用いる記号や語句の定義，既知の定理を示す。

$E: S \rightarrow \mathbb{R}$: エネルギー，推移確率は状態間のエネルギーの差によって決まるので以下の議論ではEの最小値は0と仮定する．最小値が0でないEを扱う場合は $E(i) \rightarrow E(i) - E_{\min}$ と置き換えれば以下の議論はそのまま成り立つ。

$S := \{1, 2, \dots, N\}$: 状態の集合

$p_i(t)$: 時刻tにおいて状態iにある確率

$p(t) := (p_1(t), p_2(t), \dots, p_N(t))$

$P(t)$: 推移行列， $P_{ij}(t)$ を(i, j)成分に持つ $N \times N$ 行列

$P[t_1, t_2]$: 時刻 t_1 から t_2 までの推移をまとめた推移行列， $p(t_2) = p(t_1)P[t_1, t_2]$

$$P[t_1, t_2] := \prod_{t=t_1}^{t_2-1} P(t)$$

$P[t, t] := I$ (単位行列)

$P[t_1, t_2]P[t_2, t_3] = P[t_1, t_3]$

$DV := \{(p_1, p_2, \dots, p_N) \mid p_i \geq 0, \sum_1 p_i = 1\}$

$TM := \{\text{全ての行がDVの元である } N \times N \text{ 行列}\}$

$p \in DV$ のとき $\|p\| := \sum_1 |p_i|$ (L_1 ノルム)

$f(t) \sim g(t) \Leftrightarrow t \rightarrow \infty$ のとき $f(t)/g(t) \rightarrow 1$

$f(t) \leq g(t) \Leftrightarrow t \rightarrow \infty$ のとき $f(t)/g(t) \rightarrow 0$

$\tau: TM \rightarrow [0, 1]$: エルゴード係数^{[1][6]}

$\tau(P) := \sup \{ \|(a-b)P\| / \|a-b\| \mid a, b \in DV, a \neq b \}$

(4)

実際に τ は次式で求められる。

$$\tau(P) = \max \{ \sum_k |P_{1k} - P_{jk}| / 2 \mid i, j \in S \} \\ = 1 - \min \{ \sum_k \min \{ P_{1k}, P_{jk} \} \mid i, j \in S \} \quad (5)$$

エルゴード係数は $\forall P [\tau(P) \leq 1]$ なる性質を持つ。

これは異なる分布ベクトルに対して状態推移を行うと分布ベクトル間の L_1 距離が小さくなる(少なくとも大きくならない)ことを示している。また $\forall P, Q [\tau(PQ) \leq \tau(P)\tau(Q)]$ が成り立つ。

$$\forall m \left[\lim_{n \rightarrow \infty} \sup \{ \|pP[m, n] - qP[m, n]\| \mid p, q \in DV \} = 0 \right] \quad (6)$$

であるとき推移行列の時系列 $\{P(t)\}$ によるマルコフ連鎖は弱エルゴード的^{[1][6]}であるという。弱エルゴード的なマルコフ連鎖の推移行列の時系列の集合をWEとする。

$$\exists e^* \forall m \left[\lim_{n \rightarrow \infty} \sup \{ \|pP[m, n] - e^*\| \mid p \in DV \} = 0 \right] \quad (7)$$

であるとき推移行列の時系列 $\{P(t)\}$ によるマルコフ連鎖は強エルゴード的^{[1][6]}であるという。これは

$$\exists e^* \forall m [P[m, \infty) = u^T e^*] \quad (8)$$

(u は全成分が1である $1 \times N$ ベクトル, T は転置を表す)と同値である。強エルゴード的なマルコフ連鎖の推移行列の時系列の集合をSEとする。SE \subseteq WEである。

[定理1]^{[1][6]} $\{P(t)\} \in WE$ であるための必要十分条件は

$$\sum_{m=0}^{\infty} (1 - \tau(P[k(m), k(m+1)])) = \infty \quad (9)$$

を満たす自然数の狭義単調増加列 $\{k(m)\}$ が存在することである。■

[定理2]^[6] $\{P(t)\} \in WE$ であり，各 $P(t)$ に対する定常分布 $\pi(t)$ (i.e. $\pi(t)P(t) = \pi(t)$)が有界変動，つまり

$$\sum_{t=0}^{\infty} \|\pi(t+1) - \pi(t)\| < \infty \quad (10)$$

であるとき $\{P(t)\} \in SE$ でもある。定常分布は $P(t)$ に対して一意に与えられない場合もあるがこのときは式(10)を満たすものが1つでもあれば $\{P(t)\} \in SE$ となる。

さらに式(10)が満たされるとき $\pi(t) \rightarrow e^*$ ($t \rightarrow \infty$)なる e^* が存在して

$$\forall m \left[\lim_{n \rightarrow \infty} \sup \{ \|pP[m, n] - e^*\| \mid p \in DV \} = 0 \right] \quad (11)$$

となる。■

[定理3]^{[6][7]} 式(2)において $T = c/\ln(t+b)$ とすると強エルゴード的なマルコフ連鎖が得られる。ただし

$$\left\{ \begin{array}{l} b > 1 \\ c \geq rL \\ r := \min \{ j \mid \forall m [\tau(P[mj, mj+j]) > 0] \} \\ L := \max \{ E(j) - E(i) \mid g_{ij} > 0 \} \end{array} \right\} \quad (12)$$

である。■

3. $T = c/\ln(at+b)$ による収束

定理3は定理4のように拡張できる。

[定理4] 式(2)において $T = c/\ln(at+b)$ とすると強エルゴード的なマルコフ連鎖が得られる。ただし $a > 0$ で b, c は式(12)の条件を満たすものとする。■

証明は付録に示す。ここで注目すべきことは a は $a > 0$ でありさえすれば任意に選べるということである。それでは a を変化させることは如何なる意味を持つであろうか。

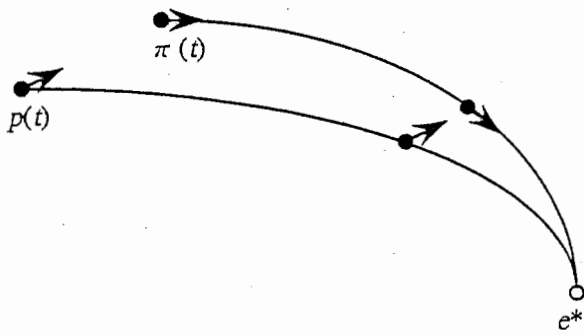


図4 $\pi(t), p(t)$ の軌跡
Fig.4 Locus of $\pi(t), p(t)$

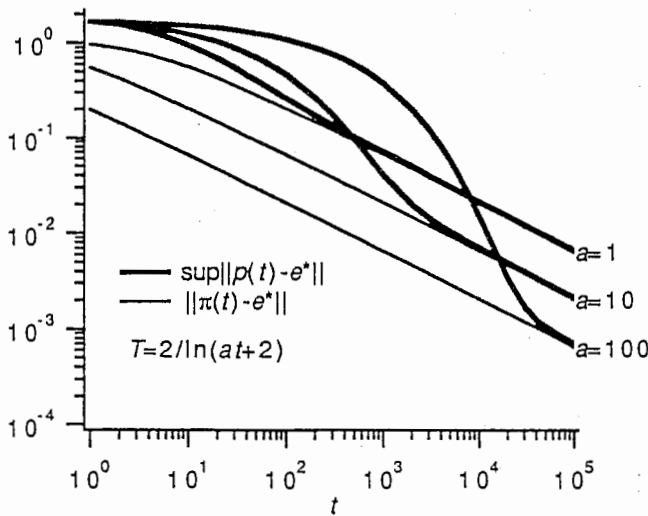


図5 $\|p(t)-e^*\|, \|\pi(t)-e^*\|$
Fig.5 $\|p(t)-e^*\|, \|\pi(t)-e^*\|$

図5は図1において $T=2/\ln(at+2)$ (これは式(12)の条件を満たす)としたときの各 $P(t)$ の持つ定常分布 $\pi(t)$ と e^* 間の L_1 距離, 及び $p(t)$ と e^* 間の L_1 距離の $p(0)$ を任意に取ったときの上限を示したものである. 時変推移行列 $\{P(t)\}$ を用い各々の $P(t)$ の定常確率を $\pi(t)$ とすると $p(t)$ を $\pi(t)$ に引き寄せつつ $\pi(t)$ を徐々に e^* に持ってゆくことにより $p(t)$ を e^* に収束させるのが時間的非一様なマルコフ連鎖を用いたシミュレーテッドアニーリングの原理である(図4)ので $p(t)$ を $\pi(t)$ よりも速く e^* に収束させることはできない. a を大きくとることは $\pi(t)$ の e^* への収束を加速することになるが一方 $p(t)$ の $\pi(t)$ への収束は遅くなる.

$$\|p(t)-e^*\| \leq \|\pi(t)-e^*\| + \|p(t)-\pi(t)\| \quad (13)$$

式(13)の右辺各項の大きさを評価してみる.

$$\begin{aligned} \|\pi(t)-e^*\| &= \sum_{i \in S^*} |\pi_i(t)-e^*| \\ &+ \sum_{i \in S-S^*} \pi_i(t) \\ &= \sum_{i \in S^*} (e^* - \pi_i(t)) \\ &+ \sum_{i \in S-S^*} \pi_i(t) \\ &= 2 \sum_{i \in S-S^*} \pi_i(t) \end{aligned} \quad (14)$$

ここで

$$\pi_i(t) = \frac{(at+b)^{-E(i)/c}}{\sum_j (at+b)^{-E(j)/c}} \quad (15)$$

である. t が十分大きいときのことを考えると

$$\|\pi(t)-e^*\| \sim 2|S^*| (at+b)^{-E/c} \quad (16)$$

が得られる. ここで $E := \min\{E(i) | i \in S-S^*\}$ つまり S 内で2番目に低いエネルギーである. $\|p(t)-\pi(t)\|$ については $c > rL$ のとき

$$\lim_{t \rightarrow \infty} \frac{\|p(t)-\pi(t)\|}{\|\pi(t)-e^*\|} = 0 \quad (17)$$

となることが示される(導出は付録に示す). このことは t が十分大きいとき式(13)右辺において第1項が支配的となることを示している. したがって a の値が大きいほど $\|p(t)-e^*\|$ の収束も速いということになる. ただしこれは t が十分大きなきのことであり t が小さいときには式(13)右辺において第1項が支配的であるとはいえないので a いくらでも大きくすればよいというわけにはいかない.

4. 推移行列の部分列による収束

シミュレーテッドアニーリング法ではその名の示す通り元々熱現象の模倣から式(2)の推移行列によるものが考案されたが, 組み合わせ最適化問題の解法に使うには熱現象の模倣をする必要性はない. どのような推移行列を用いても要は目標とする極限分布に速く法則収束すればよい.

ここではこのような一般のシミュレーテッドアニーリング法についても3.で述べたような定数倍の加速が可能かどうかを考察する. もはや推移行列は温度などというパラメータすら持たない. シミュレーテッドアニーリング法を単に強エルゴード性あるいはそれより弱い弱エルゴード性を持つ時間的非一様なマルコフ連鎖としてとらえ, 推移行列の時系列によるマルコフ連鎖が弱(強)エルゴード性を持つときその推移行列の時系列の部分列によるマルコフ連鎖も弱(強)エル

ゴード性を保つように部分列をとることを考える。元の推移行列の時系列が $P(0), P(1), \dots, P(t), \dots$ であり得られた部分列が $P(h(0)), P(h(1)), \dots, P(h(t)), \dots$ であるとき $h(t)/t \rightarrow a$ ($t \rightarrow \infty$) ならばアニーリングスケジュールは a 倍加速されたものと考えられよう。

4. 1. 弱エルゴード的な部分列

以下の定理が得られた。

[定理5] $\{P(t)\} \in WE$ であり自然数の狭義単調増加列 $\{k(m)\}$ が式(9)及び $k(m+1)/k(m) \rightarrow 1$ ($m \rightarrow \infty$) を満たすとき任意の $a=1, 2, \dots$ に対し次の条件(A), (B)を満たす自然数の狭義単調増加列 $\{h(t)\}$ が存在する。

(A) $\{P \circ h(t)\} \in WE$

(B) $h(t)/t \rightarrow a$ ($t \rightarrow \infty$)

証明は付録に示す。 $k(m+1)/k(m) \rightarrow 1$ ($m \rightarrow \infty$) はつまり $k(m)$ が指数関数的以上には増加しないということであり実際に推移行列を設計する際にはこれは比較的緩い条件であると考えられる。例えば1. で示した例では $k(m) = mr$ (r は式(12)の r) であり $k(m)$ はこの条件を満たす。従って比較的緩い条件下で弱エルゴード性を持つアニーリングスケジュールの弱エルゴード性を保ったままでの定数倍加速が可能であると言えよう。

4. 2. 強エルゴード的な部分列

以下の定理が得られた。

[定理6] 定理5においてさらに $\{P(t)\}$ が式(10)を満たす定常分布の時系列 $\{\pi(t)\}$ を持つならば任意の $a=1, 2, \dots$ に対し次の条件(A), (B), (C)を満たす自然数の狭義単調増加列 $\{h(t)\}$ が存在する。

(A) $\{P \circ h(t)\} \in SE$

(B) $h(t)/t \rightarrow a$ ($t \rightarrow \infty$)

(C) 任意の初期分布に対して $\{P(t)\}$ による $p(t)$ の収束先と $\{P \circ h(t)\}$ による $p(t)$ の収束先は一致する。 ■

証明は付録に示す。 $\{P(t)\} \in WE$ でかつ有界変動である定常分布の時系列を持つならば定理2より $\{P(t)\} \in SE$ であるから比較的緩い条件下で強エルゴード性を持つアニーリングスケジュールの強エルゴード性を保ったままでの定数倍加速が可能であると言えよう。

5. むすび

よく知られた強エルゴード性を持つアニーリングスケジュール $T(t) = c/\ln(t+b)$ の時刻を定数倍したアニーリングスケジュール $T(t) = c/\ln(at+b)$ について考察し以下の結果を得た。

(A) a を任意に大きくしても強エルゴード性を保つ。このことはある意味で $\|\pi(t) - e^x\|$ の収束をいくらでも

加速できることを示している。

(B) c をある値より大きくとった場合十分大きな t に対し $\|\pi(t) - e^x\| \gg \|p(t) - \pi(t)\|$ となることを示した。シミュレーテッドアニーリングでは $\|p(t) - e^x\|$ の値をなるべく小さくしたいわけであるが(A)(B)より t が十分大きな時は a を大きくとることにより $\|p(t) - e^x\|$ の値を小さくできることとなる。

また一般の時間的非一様なマルコフ連鎖の場合上記のような時刻の定数倍ということが出来るかについて考察し以下の結果を得た。

(C) 弱(強)エルゴード性を持つマルコフ連鎖の推移行列の時系列があるとき次の条件(1), (2)を満たす推移行列の部分列をとることを考える。(1)その部分列を推移行列とするマルコフ連鎖も弱(強)エルゴード性を持つ。(2)部分列の第 t 項を元の推移行列の時系列の第 $h(t)$ 項とすると $t \rightarrow \infty$ のとき $h(t)/t \rightarrow a$ である。その結果、比較的緩やかな条件下で任意の定数 $a \in \{1, 2, \dots\}$ に対して(1)(2)を満たす部分列がとれる。このことはある意味で時刻の定数倍によるアニーリングスケジュールの加速がいくらでもできることを示している。

本論文では収束の速さをエネルギー最小でない状態では確率が0である分布 e^x への L_1 距離で評価することをおこなったがこれは最小の解以外は何の価値もないという現実的には厳しすぎる条件とも思われる。よりよい評価の方法が今後の課題である

謝辞 日頃より快適な研究環境を与えて下さる(株)ATR視聴覚機構研究所淀川英司所長, 下原勝憲認知機構研究室長ならびに社員の皆様方へ感謝いたします。

文献

- [1] Seneta E. : "Non-negative matrices and Markov chains", Springer-Verlag (1981).
- [2] 羽鳥裕久, 森俊夫 : "有限マルコフ連鎖", 倍風館(1982).
- [3] Kirkpatrick S., Gelatt C.D. and Vecchi M.P. : "Optimization by simulated annealing", Science, 220, pp.671-680 (May 1983).
- [4] Geman S. and Geman D. : "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-6, pp.721-741 (Nov. 1984).
- [5] Isaacson D.L. and Madsen R.W. : "Markov chains, theory and applications", Robert E. Krieger (1985).

- [6] Mitra D., Romeo F. and Sangiovanni-Vincentelli A.L. : "Convergence and finite-time behavior of simulated annealing", Proc. 24th Conf. on Decision and Control, pp.761-767 (Dec. 1985).
- [7] Mitra D., Romeo F. and Sangiovanni-Vincentelli A.L. : "Convergence and finite-time behavior of simulated annealing", Adv. Appl. Prob. 18, pp.747-771 (1986)
- [8] 深尾毅 : "アニーリング法一つの最適化手法-", 信学誌, 70, 12, pp.1247-1250 (1987-12).
- [9] Van Laarhoven P.J.M. and Aarts E.H.L. : "Simulated annealing theory and applications", Kluwer (1987).
- [10] 上坂吉則 : "ニューロダイナミックスの数理", 信学誌, 73, 2, pp.131-136 (1990-02).
- [11] (社) 日本工業技術振興協会ニューロコンピュータ研究部会編 : "ニューロコンピュータの基礎理論", 海文堂 (1990).

付録

(式(17)の導出)

$$\begin{aligned} \tau(P[mr, nr]) &= \prod_{i=m}^{n-1} \tau(P[ir, ir+r]) \\ &= \exp \sum_{i=m}^{n-1} \ln \tau(P[ir, ir+r]) \\ &< \exp \sum_{i=m}^{n-1} (\tau(P[ir, ir+r]) - 1) \\ &< \exp(-\sum_{i=m}^{n-1} g^r (air+b)^{-rL/c}) \\ &< \exp(-\int_m^n g^r (axr+b)^{-rL/c} dx) \end{aligned} \quad (18)$$

c=rLのときは

$$\tau(P[mr, nr]) < \left(\frac{amr+b}{anr+b} \right)^{g^r} \quad (19)$$

c>rLのときはw:=1-rL/c>0とおけば

$$\tau(P[mr, nr]) < \exp \frac{g^r}{w} ((amr+b)^w - (anr+b)^w) \quad (20)$$

となる. d(0):=p(0)-π(0), d(t):=π(t-1)-π(t)とおけば

$$p(t) - \pi(t) = \sum_{u=0}^t d(u) P[u, t] \quad (21)$$

したがって

$$\begin{aligned} &\|p(t) - \pi(t)\| \\ &\leq \sum_{u=0}^t \|d(u)\| \tau(P[u, t]) \\ &= \sum_{u=0}^{s-1} \|d(u)\| \tau(P[u, t]) \\ &\quad + \sum_{u=s}^t \|d(u)\| \tau(P[u, t]) \\ &\leq \tau(P[s, t]) \sum_{u=0}^{s-1} \|d(u)\| \tau(P[u, s-1]) \\ &\quad + \sum_{u=s}^t \|d(u)\| \\ &\leq \tau(P[s, t]) \sum_{u=0}^{\infty} \|d(u)\| \\ &\quad + \sum_{u=s}^t \|d(u)\| \end{aligned} \quad (22)$$

tが十分に大きいときにはπ₁(t)は単調関数となることから^{[5], [7]}sが十分大きければ

$$\sum_{u=s}^t \|d(u)\| = \|\pi(s) - e^*\| - \|\pi(t) - e^*\| \quad (23)$$

ここで

$$s := \max\{s \mid (as+b)^w \leq (at+b)^w - (at+b)^{w/2}\} \quad (24)$$

とおくと式(20)より

$$\tau(P[s, t]) \leq \exp \frac{g^r}{w} (-(at+b)^{w/2}) \quad (25)$$

よって式(16)より

$$\lim_{t \rightarrow \infty} \frac{\tau(P[s, t])}{\|\pi(t) - e^*\|} = 0 \quad (26)$$

また

$$\begin{aligned} \frac{\|\pi(s) - e^*\|}{\|\pi(t) - e^*\|} &\sim \frac{((at+b)^w - (at+b)^{w/2})^{-E/cw}}{(at+b)^{-E/c}} \\ &= (1 - (at+b)^{-w/2})^{-E/cw} \end{aligned} \quad (27)$$

よって

$$\lim_{t \rightarrow \infty} \frac{\|\pi(s) - e^*\| - \|\pi(t) - e^*\|}{\|\pi(t) - e^*\|} = 0 \quad (28)$$

以上より

$$\begin{aligned} & \lim_{t \rightarrow \infty} \frac{\|p(t) - \pi(t)\|}{\|\pi(t) - e^*\|} \\ & \leq \lim_{t \rightarrow \infty} \frac{\tau(P[s, t])}{\|\pi(t) - e^*\|} \\ & + \lim_{t \rightarrow \infty} \frac{\|\pi(s) - e^*\| - \|\pi(t) - e^*\|}{\|\pi(t) - e^*\|} = 0 \end{aligned} \quad (29)$$

左辺は0以上であるので式(17)が得られる。

(導出終わり)

(定理4の証明) これは定理3の証明^{[6], [7]}と同様の証明により示される。

式(5)において $\tau < 1$ であるときは

$$\forall i, j \exists k [\min\{P_{ik}, P_{jk}\} > 0] \quad (30)$$

となるがこのとき

$\min\{P_{ik}, P_{jk}\} \geq \underline{p} := \min\{P_{ij} | P_{ij} > 0\}$ となるので $\tau(P) \geq 1 - \underline{p}$ となる。 $\tau(P[mr, mr+r])$ を評価することを考える。ただし全ての状態が同じエネルギーを持つ場合に r が存在しないことが有り得るがこの場合は元々シミュレーションでアニーリング法を行う必要がないのでこの例外は省いて考える。ある i, j について $P_{ij}[mr, mr+r] > 0$ であるとすると時刻 mr に状態 i から出発して時刻 $mr+r$ に状態 j に至るパス、つまり式(31)を満たす状態のパス $Q = (q(0), q(1), \dots, q(r))$ が存在する。

$$\begin{cases} q(0) = i, q(r) = j \\ \prod_{n=0}^{r-1} P_{q(n)q(n+1)}(mr+n) > 0 \end{cases} \quad (31)$$

$P_{ij}[mr, mr+r]$ はそのパスを通して i から j に到達する確率以上であるので

$$P_{ij}[mr, mr+r] \geq \prod_{t=mr}^{mr+r-1} p(t) \quad (32)$$

である。ここで $\underline{g} := \min\{g_{ij} | g_{ij} > 0\}$ とすると補題1より T が十分小さいとき $\underline{p} \geq \underline{g}e^{-L/T}$ である。 $T \rightarrow 0$ ($t \rightarrow \infty$)なので十分大きな m について

$$\begin{aligned} P[mr, mr+r] & \geq \prod_{t=mr}^{mr+r-1} \underline{g}e^{-L/T(t)} \\ & = \prod_{t=mr}^{mr+r-1} \underline{g}(at+b)^{-rL/c} \\ & > \underline{g}^r (amr+b)^{-rL/c} \end{aligned} \quad (33)$$

である。以上より $b > 1$ ($\because T > 0$) のとき有限個の m を除いて上界

$$\tau(P[mr, mr+r]) \leq 1 - \underline{g}^r (amr+b)^{-rL/c} \quad (34)$$

が得られる。 $c \geq rL$ であれば

$$\sum_{m=0}^{\infty} (1 - \tau(P[mr, mr+r])) \geq \sum_{m=M}^{\infty} \underline{g}^r (amr+b)^{-rL/c} = \infty \quad (35)$$

であるので定理1より $\{P(t)\} \in WE$ であることになる。さらに式(3)により与えられる $\pi(t)$ は有界変動なので定理2より $\{P(t)\} \in SE$ でもあり任意の初期分布に対して $p(t) \rightarrow e^*$ ($t \rightarrow \infty$)となる。(証明終わり)

[補題1] 式(2)で T が十分小さいとき $\underline{p} \geq \underline{g}e^{-L/T}$ である。

(補題1の証明) 全ての i, j に対して $P_{ij} = 0$ かまたは十分小さな T に対して $P_{ij} \geq \underline{g}e^{-L/T}$ であれば良い。以下の場合分けにより証明される。

(I) $i \neq j, g_{ij} = 0$ のとき $P_{ij} = 0$ である。

(II) $g_{ij} > 0$ のとき $P_{ij} \geq \underline{g}e^{-L/T}$ である。

(III) $\forall j [g_{ij} > 0 \Rightarrow E(i) \geq E(j)]$ のとき $P_{ii} = 0$ である。

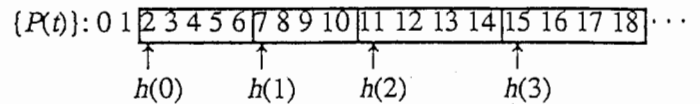
(IV) $\exists j [g_{ij} > 0, E(i) < E(j)]$ のとき $J_- := \{j | E(i) \leq E(j), j \neq i\}$, $J_+ := \{j | E(i) > E(j)\}$ とおくと

$$\begin{aligned} P_{ii} & = 1 - \sum_{j \in J_-} P_{ij} - \sum_{j \in J_+} P_{ij} \\ & = 1 - \sum_{j \in J_-} g_{ij} - \sum_{j \in J_+} g_{ij} a_{ij} \end{aligned} \quad (36)$$

である。 $\underline{g}e^{-L/T(t)}$ は $T \rightarrow 0$ のとき0に収束するのに対し

$$\lim_{T \rightarrow 0} P_{ii} = 1 - \sum_{j \in J_-} g_{ij} > 0 \quad (37)$$

であるので十分小さな T に対して $P_{ii} \geq \underline{g}e^{-L/T}$ である。(証明終わり)



$$\{P \cdot h_1(t)\}: \boxed{2 \ 3 \ 4 \ 5 \ 6} \boxed{11 \ 12 \ 13 \ 14} \dots$$

$$\{P \cdot h_2(t)\}: \boxed{7 \ 8 \ 9 \ 10} \boxed{15 \ 16 \ 17 \ 18} \dots$$

図6 部分列への分配
Fig.6 Distribution to subsequences

(定理5の証明)

図6に示すように $\{P(t)\}$ を区間 $[0, k(0))$ 及び $[k(m), k(m+1))$ に区切って1かたまりとし a 個(図では $a=2$)の部分列に分配して行くことを考える。補題2よりどのような分配の仕方をしてしてもその内少なくとも1つの部分列によるマルコフ連鎖は弱エルゴード的となる。各部分列になるべく均等になるような分配の仕方をしてどの部分列でも $h_i(t)/t \rightarrow a$ ($t \rightarrow \infty$)が成り立つようにできれば所望の部分列が存在する事が保証される。

以下の補題3により $k(m+1)/k(m) \rightarrow 1$ ($m \rightarrow \infty$) であれば 実際このような均等な分配の仕方がある。

SEとなる。 $t \rightarrow \infty$ のとき $\pi(t) \rightarrow e^*$ とすると $\pi \cdot h(t) \rightarrow e^*$ であるので収束先も一致する。 (証明終わり)

(証明終わり)

[補題2] $\{P(t)\} \in WE$ であり自然数の狭義単調増加列 $\{k(m)\}$ が式(9)を満たすとする。自然数の分割 $\{H_i\}$ ($i=1, 2, \dots, a$) が $\forall i [k(m) \in F_i \Rightarrow [k(m), k(m+1)) \subseteq F_i]$ を満たすとする。このとき H_i の元を小さい順に $h_i(0), h_i(1), \dots$ とすると $\exists i [\{P \cdot h_i(t)\} \in WE]$ である。 ■

(補題2の証明)

$$S_i := \sum_{m: [k(m), k(m+1)) \subseteq H_i} (1 - \tau(P[k(m), k(m+1)))) \quad (38)$$

とおく。式(9)より $\sum_i S_i = \infty$ であるので少なくとも1つ $S_i = \infty$ となる H_i が存在するが定理1より $\{P \cdot h_i(t)\} \in WE$ となる。 (証明終わり)

[補題3] 自然数の狭義単調増加列 $\{k(m)\}$ ($m \in \mathbb{N}$) が $k(m+1)/k(m) \rightarrow 1$ ($m \rightarrow \infty$) となるならば次の条件(A), (B)を満たす自然数の分割 $\{H_i\}$ ($i=1, 2, \dots, a$) が存在する。

(A) $\forall i [k(m) \in H_i \Rightarrow [k(m), k(m+1)) \subseteq H_i]$

(B) H_i の元を小さい順に $h_i(0), h_i(1), \dots$ とすると

$$\forall i [h_i(t)/t \rightarrow a \quad (t \rightarrow \infty)] \quad \blacksquare$$

(補題3の証明) 全ての整数区間 $[k(m), k(m+1))$ 及び $[0, k(0))$ を自然数の分割 $\{H_i\}$ ($i=1, 2, \dots, a$) に割り当てていくことを考える。まず $[0, k(0))$ を任意の H_i に割り当てる。以下帰納的に $[0, k(m))$ までの割当てが済んだとき $|H_i \cap [0, k(m))|$ が最も小さい H_i (複数あるときはそのうちのどれでもよい) に $[k(m), k(m+1))$ を割り当てる。こうしてできた H_i について $\sum_i |H_i \cap [0, k(m))| = k(m)$ であるので $\exists i [|H_i \cap [0, k(m))| \leq k(m)/a]$ である。 $k(m) \leq h_i(t) < k(m+1)$ のとき $[k(m), k(m+1))$ は $|H_i \cap [0, k(m))|$ が最も小さい H_i に割り当てられるので

$$\frac{k(m)}{a} \leq t < \frac{k(m)}{a} + k(m+1) - k(m) \quad (39)$$

よって

$$\frac{k(m)}{k(m)/a + k(m+1) - k(m)} < \frac{h_i(t)}{t} < \frac{k(m+1)}{k(m)/a} \quad (40)$$

$t \rightarrow \infty$ のとき $m \rightarrow \infty$ であり $k(m+1)/k(m) \rightarrow 1$ となることから式(40)の両辺は a に収束するので $h_i(t)/t \rightarrow a$ となる。

(証明終わり)

(定理6の証明)

$$\sum_{t=0}^{\infty} \|\pi \cdot h(t+1) - \pi \cdot h(t)\| \leq \sum_{t=0}^{\infty} \|\pi(t+1) - \pi(t)\| \quad (41)$$

より $\{\pi(t)\}$ が有界変動ならば $\{\pi \cdot h(t)\}$ も有界変動である。 $\{P \cdot h(t)\} \in WE$ であるので定理2より $\{P \cdot h(t)\} \in$