

焼きなまし法の温度を下げるスケジュールの高校程度の数学で分かる説明

学籍番号 03-123009 氏名 奥谷 文徳

付録 (論文和訳)

理想的な焼きなまし法の焼きなましスケジュール

BRUCE HAJEK

イリノイ大学アーバナ・シャンペーン校

(提出日時など略)

焼きなまし法と呼ばれるモンテカルロ最適化技法は (最小解ではない) 局所解を脱出するためにランダムな上昇する動きによって修正された降下法である. そのランダム性^{*1}のレベルは T (温度) という制御パラメータによって決定される. その T は決定論的な^{*2}焼きなましスケジュールに従って 0 へと向かっていく. 我々は最小コストの状態へと確率収束する焼きなましスケジュールの簡潔な必要十分条件を与える.

必要十分条件: 焼きなましスケジュールがパラメトリック形式^{*3} $T(t) = \frac{c}{\log(1+t)}$ という特殊な場合 (ただし c は大域的な最小点ではない局所解の深さ以上)

1 序論

有限集合 S に対して最小化すべき関数 V を考える.

S の各要素 x に対して $N(x)$ という状態近傍集合を考える. ($N(x) \subset S$)

一般に $N(x)$ は S の真部分集合である.^{*4}

また, 遷移確率行列 R を定義する. R は $y \in N(x)$ の時 $R(x, y) > 0$ であり, また $y \notin N(x)$ の時 $R(x, y) = 0$ である.

T_1, T_2, \dots はすべて 0 より大きく,

$$T_1 \geq T_2 \geq \dots \tag{1.1}$$

$$\lim_{k \rightarrow \infty} T_k = 0 \tag{1.2}$$

を満たすよう定義する.

次のような X_0, X_1, \dots を構成する順序アルゴリズムを考える. ただし X_0 は与えるものとする.

与えられた $X_k = x$ ならば, 次のあり得る (移りうる) の状態 Y_k は $N(x)$ の中から確率分布関数

$$P[Y_k = y | X_k = x] = R(x, y)$$

で選ばれる.

$$X_{k+1} = \begin{cases} Y_k & \text{with probability } p_k \\ X_k & \text{otherwise} \end{cases}$$
$$p_k = e^{\frac{-[V(Y_k) - V(x)]^+}{T_k}}$$

^{*1} randomization

^{*2} deterministic

^{*3} 媒介変数表示

^{*4} $N(x) \neq S$

によって X_1, X_2, \dots を明記する (定義する)

また, V が最小になるような \mathcal{S} の要素を集めた集合^{*5}を \mathcal{S}^* とする.

私達は

$$\lim_{k \rightarrow \infty} P[X_k \in \mathcal{S}^*] = 1 \quad (1.3)$$

かどうかに決めることに興味がある.^{*6}

アルゴリズムによって生成されるランダムな過程 $X = (X_k : 0 \leq k)$ は離散時間マルコフ連鎖^{*7}である

k ステップ目の遷移確率行列は

$$P_k(x, y) = P[X_{k+1} = y | X_k = x] \\ = \begin{cases} 0 & \text{if } y \notin N(x) \text{ and } y \neq x \\ R(x, y) e^{\frac{-[V(y) - V(x)]^+}{T_k}} & \text{if } y \in N(x) \text{ and } y \neq x \\ 1 - \sum_{z \neq x} P_k(x, z) & \text{if } y = x \end{cases}$$

3つの簡単な仮定のもとでのアルゴリズムを少し考えることによって, アルゴリズムの遷移確率の選択の理由を述べる^{*8}

もし $(x = y)$ もしくは $(0 \leq k < p$ で $x_{k+1} \in N(x_k)$ となるような $p \geq 1$ を用いて $x = x_0, x_1, \dots, x_p = y$ となる順序^{*9}が存在する) ならば y は x から”到達可能”^{*10}という.

一つ目の仮定は (\mathcal{S}, N) は既約であるということ. ここでの既約とは, x と y という二つの状態があったとして, y は x から”到達可能”という意味である.^{*11}

二つ目の仮定はある定数 $T > 0$ において $T_k = T$ であるということ.

三つ目の仮定は R は可逆であるということ, 可逆であるというのは, \mathcal{S} の確率分布 α を考えると平衡分布が R に必要であるということ. つまり任意の $x, y \in \mathcal{S}$ において $\alpha(x)R(x, y) = \alpha(y)R(y, x)$

簡単な例として次のような正当な例があげられる.

$$R(x, y) = \begin{cases} \frac{1}{|N(x)|} & \text{if } y \in N(x) \\ 0 & \text{otherwise} \end{cases}$$

そして状態近傍システムは二つの状態 x, y について $x \in N(y)$ のとき, そしてその時に限り $y \in N(x)$ であるという意味で対称である.

すべての k について $T_k = T$ という過程によって, マルコフ連鎖 X は定常的な1ステップ遷移確率行列 P を持つ. (すべての k において $P = P_k$)

これは R に仮定した可逆性から当然生じる. もし \mathcal{S} に確率分布 π_T を

$$\pi_T(x) = \frac{\alpha(x) e^{-\frac{V(x)}{T}}}{Z_T} \\ Z_T = \sum_x \alpha(x) e^{-\frac{V(x)}{T}}$$

と定義すると P は π_T で改善可能である.

(\mathcal{S}, N) が既約であるという仮定 (そして $P(i, i) > 0$ を満たす i が存在するために $\mathcal{S}^* \neq \mathcal{S}$ ならば P は非周期であ

^{*5} 複数ある可能性を含めて集合としている?

^{*6} 直訳っぽい

^{*7} 参考:<http://ja.wikipedia.org/wiki/マルコフ連鎖モンテカルロ法>

^{*8} motivate

^{*9} sequence

^{*10} 到達可能 (reachable) の説明

^{*11} 恐らくどの状態からでもどの状態へと映ることを示している

るという場合) マルコフエルゴード (的収束)^{*12}定理 [?] は必然的に^{*13}

$$\lim_{k \rightarrow \infty} P[X_k \in \mathcal{S}^*] = \sum_{x \in \mathcal{S}^*} \pi_T(x) \quad (1.4)$$

を示す.

π_T を考察すると $\sum_{x \in \mathcal{S}^*} \pi_T(x)$ は T を小さく取ると十分 1 に近くなることがすぐに分かる. つまり

$$\lim_{T \rightarrow 0} \left[\lim_{k \rightarrow \infty, T_k \equiv T} P[X_k \in \mathcal{S}^*] \right] = 1$$

焼きなまし法アルゴリズムの考えは T_k を k が ∞ へと近づくとともに 0 へと近づけることで

$$\lim_{k \rightarrow \infty} P[X_k \in \mathcal{S}^*] = 1 \quad (1.3 \text{ 再掲})$$

を成立させる試みである.

一度最初の T_k 順列が非増加^{*14}で 0 を下限に持つということに戻って見よう.

R が可逆であるという条件は必要でないだろう, その代わり, 次の定義の助けを借りてずっと弱い仮定が生まれるだろう. もし $(x = y \text{ かつ } V(x) \leq E)$ もしくは $(0 \leq k < p \text{ の範囲で } x_{k+1} \in N(x_k) \text{ かつ } V(x_k) \leq E)$ であるような $x = x_0, x_1, \dots, x_p = x$ である 1 以上の p が存在する) である時, 状態 x から状態 y は高さ E で到達可能であるという.

これから (\mathcal{S}, V, N) は次の性質を持つと仮定する.

弱可逆性^{*15}: 実数 E と二つの状態 x, y において, x が高さ E で y から到達可能であり, かつ y のみが x から高さ E で到達可能である (性質). もし $V(y) < V(x)$ なる状態 y が状態 x から高さ $V(x)$ で到達可能でないなら x は局所的最小点^{*16}と言われる. 私達は局所的最小点 x の”深さ”^{*17}を, x が大局的最小点である時 $+\infty$ になるように定義する. またその一方, x の深さは $V(y) < V(x)$ を満たし, x から高さ $V(x) + E$ で到達可能である状態 y が存在する最小の $E (E > 0)$ とする. 図示すると図 1 である.

(\mathcal{S}, V, N) の”谷”を, ある E について「それぞれの $x \in C$ を満たす x について, $C = \{y : y \text{ は } x \text{ から高さ } E \text{ で到達可能である}\}$ 状態の集合 C であると定義する. 例えば弱可逆性によって, もし $E \geq V(x)$ の時 x から高さ E で到達可能である状態の集合は谷である. 谷 C が与えられれば $\underline{V}(C) = \min\{V(x) : x \in C\}$ と $\bar{V}(C) = \min\{V(y) : y \notin C \text{ かつ } C \text{ の中に } y \in N(x) \text{ を満たす } x \text{ が存在する.}\}$ ^{*22}もし $C = \mathcal{S}$ であれば $\bar{V}(C)$ は空集合であり, $\bar{V}(\mathcal{S}) = +\infty$ とする. $B = \{x \in C : \bar{V}(x) = \bar{V}(C)\}$ で定義される C の部分集合 B を谷の底集合 B と, $d(C) = \bar{V}(C) - \underline{V}(C)$ で定義される $d(C)$ を谷の深さとする. これらを図 2 に示す. 深さ d の局所的最小点は, 深さ d の谷の要素である.

定理 1: 図 1, 2 が持っている様に (\mathcal{S}, V, N) が既約であり弱可逆性を満たしている時,

- (a) 局所的最小点でない任意の状態 x について $\lim_{k \rightarrow \infty} P[X_k = x] = 0$
- (b) 状態の集合 B が深さ d の谷の底であり, B の状態は深さ d の局所的最小点であるとき, $\sum_{k=1}^{\infty} e^{-\frac{d}{T_k}} = +\infty \iff \lim_{k \rightarrow \infty} P[X_k \in B] = 0$
- (c) ((a) と (b) の結果) d^* を大局的最小点ではないが局所的最小点である状態の深さの最大とし, 大局的最小点の集合を \mathcal{S}^* と表す. すると

$$\lim_{k \rightarrow \infty} P[X_k \in \mathcal{S}^*] = 1 \quad (1.5)$$

^{*12} Markov ergodic convergence

^{*13} imply

^{*14} 弱い意味での単調減少

^{*15} 性質 WR(Weak Reversibility)

^{*16} 局所解?

^{*17} depth

^{*21} Arc

^{*21} $x_4 (V(x_4) = 5)$ は $V(x) = 10$ まで上がることににより他の局所的最小点へと移動可能であると計算する

^{*22} つまり $\bar{V}(C)$ は C の外の隣接最小値であり, C の内部ではない.

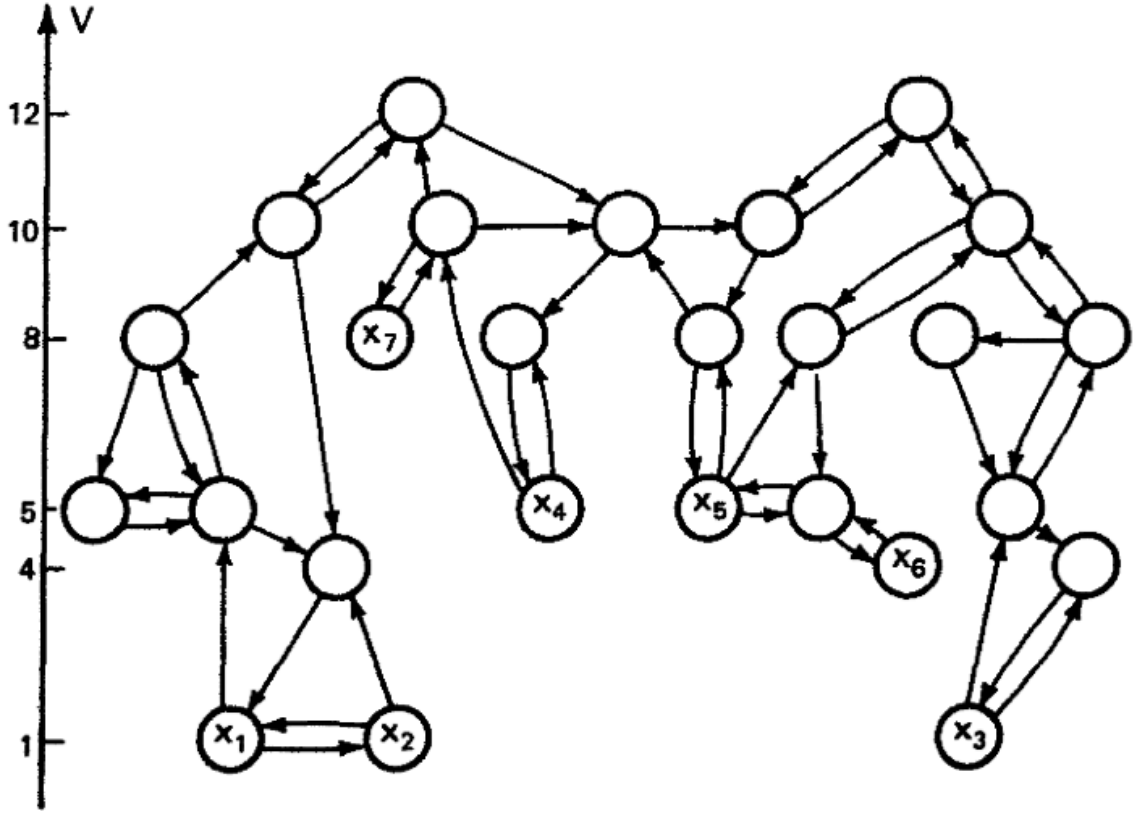


図1 (S, V, N) から生じたものを描写したグラフである. ノードは S の要素に対応している. S の要素 x に対応する $V(x)$ は左の縦軸で示される. 矢印^{*20}は $x \in N(y)$ である状態 (x, y) の順序対を表す. このグラフは弱可逆性を満たしている.

状態 x_1, x_2, x_3 は大局的最小点であり, x_4, x_6, x_7 は深さがそれぞれ 5, 6, 2 である局所的最小点^{*21}であり x_5 は局所的最小点ではない. 状態 x_2 は高さ 1 で x_1 から到達可能であり, x_3 は高さ 12 で x_1 から到達可能である.

またこれは

$$\sum_{k=1}^{\infty} e^{-\frac{d^*}{T_k}} = +\infty \quad (1.6)$$

と同値である.

REMARKS(1)^{*23}

T_k を媒介変数表示

$$T_k = \frac{c}{\log(k+1)} \quad (1.7)$$

で仮定すれば, 式 (1.6) そしてしたがって同値な式 (1.5) を満たす $\iff c \geq d^*$ この結果は Geman と Geman の結果 [?] と矛盾していない. 彼らはここで使われたモデルのほとんどの特殊な場合を考え, 式 (1.5) は T_k が十分大きな定数 c に対して式 (1.7) を満たすことを保証する^{*24}, ことを証明した. 彼らは収束に十分な c の値 ($c > d^*$) を与えた.^{*25}

Gidas[?] も焼きなまし法アルゴリズムの収束性を扱った. 彼が考えたマルコフ連鎖は私達が考えているものより一般的である. 彼は 1 ステップ遷移確率行列 P_k が k が無限へと増えるに連れて収束するという条件を必要とした. 焼

^{*23} 作者の批評, 意見, 注意

^{*24} hold

^{*25} それは実質 d^* より大きい.

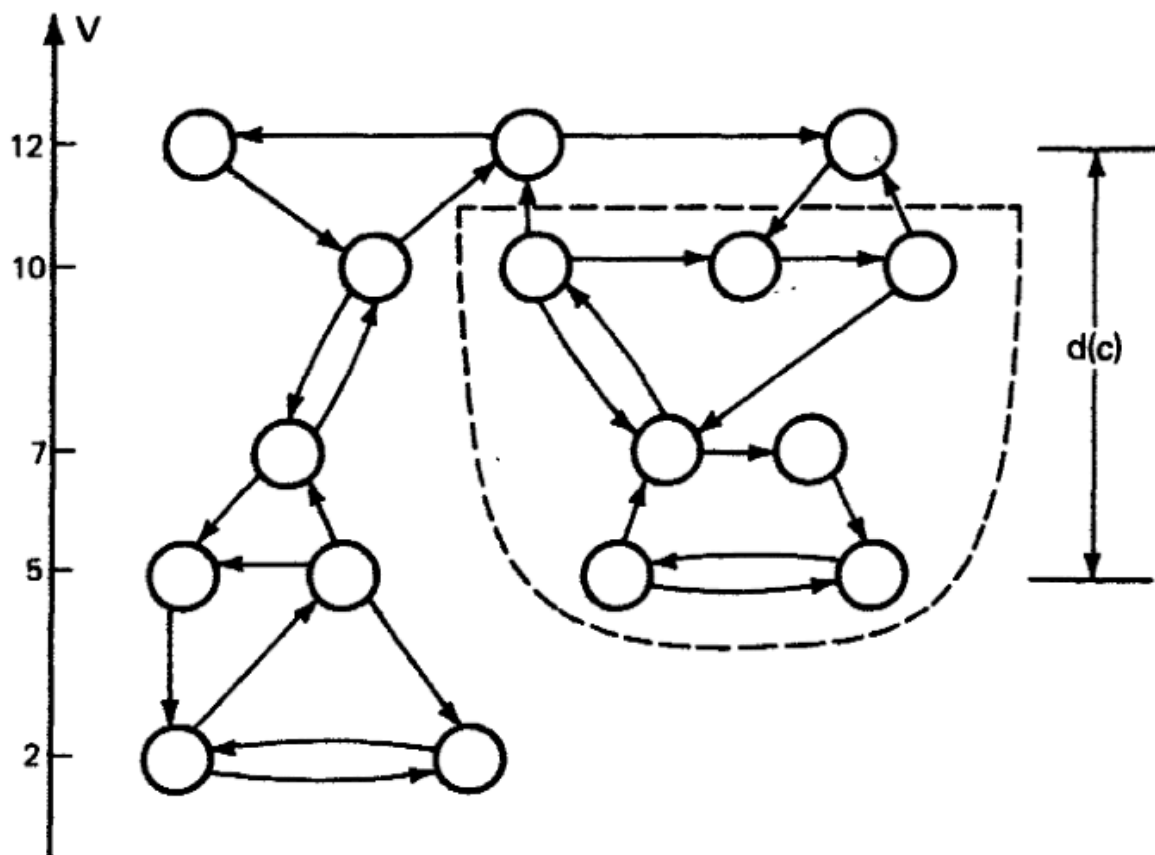


図2 点線で囲まれた谷 C . $\underline{V}(X) = 5$, $\bar{V}(C) = 12$, $d(C) = 7$ であり C の底 B の要素は2つである.

きなまし法の過程の特殊な場合では、彼は c の値 (実際はここで述べている c と Gidas の述べる $\frac{1}{C_0}$ が対応する) を式 (1.7) を満たすような最小値が式 (1.5) を導くと推測した彼の定数はここで述べた定数 d^* とは異なる. Gidas はマルコフ連鎖の興味深い実用的な収束可能性^{*26}の議論をしている.

Geman と Hwang[?] は c が V の最大値と最小値の差以上であれば、式 (1.7) のスケジュールが大局的最小点へと収束する (ことを満たす) 非定常的な分散過程の類似する場合を示している. ここで定義した定数 d^* の明らかな類似によって最小の定数が与えられると推測するが、私達が与えた証明は、この分散の場合を容易に繰り越せない.

REMARKS(2)

式 (1.5) でどの程度早く収束するかについていくつかの情報はそれぞれ定理 3,4 と §3,4, そして [?] の証明から収集できる.

REMARKS(3)

$V(X_k)$ の振る舞いより $\min_{n \leq k} V(X_n)$ を知ることは面白いであろう.

REMARKS(4)

スケジュール (T_k) がランダムでアルゴリズムの状態に従っている時の収束性を考えることは面白いだろう.

^{*26} convergence question

REMARKS(5)

もし x が深さ d 局所的最小点でありほかに $V(x) = V(y)$ であり高さ $V(x) + d$ で到達可能な局所的最小点 y が存在しない時、つまり、谷の底に唯一の状態 x であるような深さ d の谷が存在する時、定理 1 の条件下で $\lim_{k \rightarrow \infty} P[X_k = x] = 0 \iff \sum_{k=1}^{\infty} e^{-\frac{d^*}{T_k}} = +\infty$ である。深さ d の任意の局所的最小点においてこの最後の状態は true である推測する。

例 (1)

図 1,2 において, $d^* = 6, d^* = 7^{*27}$

例 (2)

Cerny[?] と Kirkpatrick ら [?] は独立に焼きなまし法を導入した。彼らは (今ではたくさんの研究者が) 焼きなまし法アルゴリズムを困難な組み合わせ問題に応用してきた。この問題が効率的に多項式時間で解けるアルゴリズムが知られているという意味で”簡単”であろうとも最大マッチング問題を簡単に考える。

無向グラフを考える。マッチングしている集合 M は互いに短点を共有しないグラフの枝の集合である。焼きなまし法の状態空間 S をすべての組み合わせになるようにセットアップし, $V(M)$ を M に含まれる枝の数 $\times -1$ とする。最大マッチング問題は S 内部で $V(M)$ が最小になるような M を見つける問題となる。 k 回繰り返す焼きなまし法アルゴリズムの初期状態を M と考える。グラフの枝を一本ランダム (等確率に) に選ぶとする。もしその枝が M 内部でなく M に追加可能であれば、次の状態は M にその枝を追加した状態とする。もし枝が M 内部であれば、次の状態は $e^{-\frac{1}{T_k}}$ の確率で M からその枝が取り除かれた状態とする。そのどちらでもない場合、次の状態は M である (M は状態を維持する)。グラフの枝の数を L とし, M' が M からから枝を一本追加するか取り除くかの動作で得られるものとする、これは $R(M, M') = \frac{1}{L}$ であり, M' がこれに当てはまらず, M でもない場合, $R(M, M') = 0$ である。

もし M が唯一の最大でない場合、一度に二本の枝を追加できる状態になるまで、枝の交換が行えるということは有名である。つまり、グラフの大きさにかかわらず, d^* はせいぜい多くても 1 である (0 であるグラフも存在する)。

例 (3)

(S, V, N) が図 3 を与えるとする。 $d^* = 3$ である。式 (1.1),(1.2)^{*28}を満たす T_k を考える。

$$\sum_{k=1}^{\infty} e^{-\frac{2}{T_k}} = +\infty \quad (1.8)$$

$$\sum_{k=1}^{\infty} e^{-\frac{3}{T_k}} < +\infty \quad (1.9)$$

x_4 が高さ $V(x_1) + 2$ で到達可能であり、式 (1.8) が成立するとすれば、もし「焼きなまし法アルゴリズムが x_1 から開始されたら確率 1 で x_4 へとやがて到達する」と (正しく)^{*29}考える人がいるかもしれない。同様に「やがては必ず x_7 へと到達する」と (誤って) 考える人がいるかもしれない。しかし、定理 1 と式 (1.9) により、 $\lim_{k \rightarrow \infty} P[X_k = x_7] \neq 1$ である。もし k が大きく、 T_k が小さいとき、時間 k において状態 x_4 に過程^{*30}があると考えると、(極めて稀であるが) 状態 x_6 を経由したと考えるより、経由せずに状態 x_1 に過程が存在するとかんがえるほうがずっと確率が高そう

^{*27} 図 1 では x_3 に対して深さ 6 の x_6 が存在している。

^{*28} 式番号あってる? (1.8),(1.9) では?

^{*29} 理由入っていないけど結果を先に言っている

^{*30} process

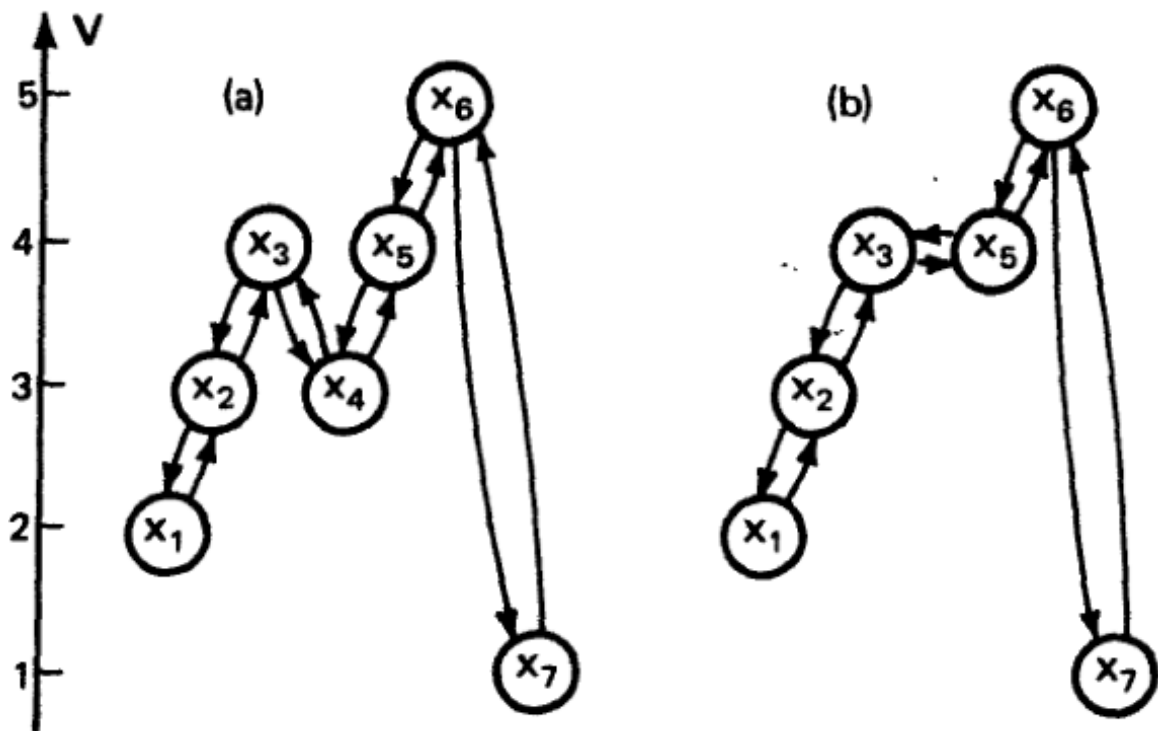


図3 図(a)は (S, V, N) から生成されたグラフ, 図(b)は $\{x_4\}$ という谷を埋めて得られるグラフ

である. 状態 x_4 単独で成る谷が埋まっているとみなす (図3参照) と, x_1 から x_6 まで, 仮定は3段階登らなければいけない. おおまかに言うと, x_4 にある関数 V の小さなくぼみは「3段階を登る動作を一度に2段階上り, 途中で一度休む」という動作へと細工できないということである. もし弱可逆性が成り立たない時は正しいとは限らない. 例えば x_4 から x_3 へと遷移する確率が0である場合などは正しくない.

この論文は次の流れで書かれている. §2に定理1の一般化を述べる. 残りの部分は定理2の証明に費やす. まず, この定理は (S, V, N) が §2で定式化する「継続的な増加」という性質を持っているという仮定で証明されるだろう. §§3,4ではどのように過程が谷から抜け出すかについてを述べた, 定理3,4の証明をする. 「継続的な増加」という仮定の取り除き方を含む証明2は §5で現れる. 必要十分条件 (1.6) と §5の相互作用である.

2 定理1の一般化

証明の前に定理1を一般化することは役立つだろう. 論文の中では (S, V, N) が既約であり, §1で述べた弱可逆性を持つと仮定する. 以降論文では $(\lambda_t : t > 0)$ は $0 < \lambda_t < 1$ を満たすとし, t に対して弱い単調増加であり, $\lim_{t \rightarrow \infty} \lambda_t = 0$ とする. $X_t : t \geq 0$ は, $(U_k, Y_k : k \leq 0)$ が状態空間 $[0, +\infty) \times S$ と, 以下に示す1段階の遷移確率で定義される同次マルコフランダム過程は $U_k \leq t < U_{k+1}$ において $X_t = Y_k$ によって定義される.

$$P[U_{k+1} \geq u, Y_{k+1} = y | (U_i, Y_i) : 0 \leq i \leq k] = \int_u^\infty Q(Y_k, y, t, \lambda_t) \Phi(dt, U_k, Y_k)$$

ただし定数 D, a, c_1, c_2 は正の数であるとする.

A.1 $\Phi(\cdot, s, x)$ は $t \geq 0$, $x \in S$ の任意の点において $\Phi(s-, s, x) = 0$, $\int_0^\infty (t-s) \Phi(dt, s, x) \leq D$, $E[\min\{k : U_k \geq t+a\} | U_0 = t, Y_0 = x] \leq \frac{1}{a}$ である確率分布関数.

A.2 $t \geq 0$, $0 \leq \lambda \leq 1$ において, $Q(\cdot, \cdot, t, \lambda)$ は

$$\begin{aligned} c_1 \lambda^{(V(y)-V(x))^+} &\leq Q(x, y, t, \lambda) \leq c_2 \lambda^{(V(y)-V(x))^+} \text{ if } y \in N(x) \\ Q(x, y, t, \lambda) &= 0 \text{ if } y \notin N(x) \cup \{x\} \end{aligned}$$

である遷移確率行列である。 Y_{k+1} は (U_{k+1}, Y_k) にのみ依存して $(U_0, Y_0, \dots, U_{k-1}, Y_{k-1}, U_k)$ とは独立である。^{*31} 例えば、もし $(T_k : k \geq 0)$ で R が §1 で表されるような形で、 I_A が集合 A の標識変数であるとする、 $\lambda_t = e^{-\frac{1}{T[t]}}$, $\Phi(t, s, x) = I_{\{t \geq s+1\}}$, $Q(x, y, t, \lambda) = R(x, y) \lambda_t^{-(V(y)) - V(x)^+}$ (ただし、 $x \neq y$)。すると、整数時間で標本化した $(X_t : t \geq 0)$ は §1 で $(X_k : k \geq 0)$ が持っているような遷移確率を持つ。 λ, Φ, Q がこの例のように選ばれるとき、定理 2 は定理 1 へと簡約化される。

もし、代わりに $\Phi(t, s, x) = I_{\{t \geq s\}} (1 - e^{-(t-s)})$ とすれば、 X は §1 の過程のアナログ連続時間である。

定理 2

- (a) 局所的最小点ではない任意の状態 x において $\lim_{t \rightarrow \infty} P[X_t = x] = 0$
- (b) B が深さ d の谷 C の底 (集合) であり、 B の内部の状態が深さ d の局所的最小点であるとする。その時、

$$\lim_{t \rightarrow \infty} P[X_t \in B] = 0 \iff \int_0^\infty \lambda_t^d dt = +\infty$$
- (c) (a),(b) より、 d^* を大局的最小点でない局所的最小点のうち最も深いものの深さとし、 S を大局的最小点の集合とすると、

$$\lim_{t \rightarrow \infty} P[X_t \in S^*] = 1 \iff \int_0^\infty \lambda_t^{d^*} dt = +\infty$$

E_0, E_1, \dots, E_q が x が S 上を動くときに $V(x)$ の取りうる値を $E_0 < E_1 < \dots < E_q$ となるように表す。まず、定理 2 は (S, V, N) が以下の様な性質を持つとして証明する。

継続的増加性^{*32}：二つの状態 x, y が与えられ、もし $V(x) = E_i, V(y) = E_j$ で $i \geq i+2$ ならば $y \notin N(x)$ である。^{*33}

図 2 は継続的増加性を有しているが、図 1 の (S, V, N) は継続的増加性を有していない。

3 谷から抜け出す方法

3.1 証明の言明と最初の段階

谷 C を考え、 m, k が谷内部の $V(x)$ の最大値 E_m と最小値 E_k であるような整数であるとする。 $k < q$ だとすると (そうでなければ)、 $C = S) 0 \leq m \leq k < q$. $F = \{y : y \notin C \text{ かつ 谷内部に } y \in N(x) \text{ が存在}\}$ とする。継続的増加性より、 $x \in F$ ならば $V(x) = E_{k+1}$ である。 u と g を $u = E_{k+1} - E_k, g = E_k - E_m$ とする。谷 C の深さ d とは $d = u + g$ の関係がある。 $\Theta = (S, V, N, D, a, c_1, c_2)$ とする。(ただし D, a, c_1, c_2 は A.1, A.2 で現れた定数)

定理 3 0 以上の ϵ が存在し、 $\bar{\lambda} > 0$ が各 $t_0 \geq 0$, 各 $x_o \in C$, 各 $y_o \in F$, 各 $\lambda_{t_0} \leq \bar{\lambda}$ で $\int_{t_0}^\infty \lambda_s^\epsilon ds = +\infty$ である ($\lambda_t : t \geq 0$) であるような Θ と C のみに依存している場合、次の条件を満たす。

$$E \left[\int_{t_0}^{\tau(C)} \lambda_s^d ds | (U_0, Y_0) = (t_o, x_o) \right] \leq \frac{1}{\epsilon} \quad ((a))$$

ただし、 $\tau(C) = U_W$ であり、 $W = \min\{k \geq 0 : Y_k \in F\}$

$$P[Y_W = y_o | (U_0, Y_0) = (t_o, x_o)] \geq \epsilon \quad (1)$$

この定理は次の谷についての仮説を用いて数学的帰納法により証明するだろう。^{*34}

数学的帰納法の仮定：もし C' が谷 C の深さよりも小さい深さしかないような谷であれば 定理 3 は C' についても成立する。

この節の残りでは、この帰納法の仮定を簡単に活用^{*35}していく。2つの節で証明されているような定理 3(a),(b) の二つの性質を満たすよう C を設定する必要がある。状態の集合 C は互いに素な C_1, C_2, \dots, C_p という谷の集合と $A = \{x \in C : V(x) = E_k\} (\max\{V(x) : x \in C_i\} = E_{k-1})$ という集合に分けられる。これを図示すると図??である。

^{*31} ?

^{*32} continuous increase property

^{*33} my memo:つまりレベルが離散的で関数とレベルの交点には状態があるみたいなの？

^{*34} will be proved by induction once we establish it for a cup C with the aid of the folioing hypothesis

^{*35} briefly exploit

$m = k$ という特殊な場合では、集合 A にのみ分割される。一般に、 C の二つの状態 x, y と $V(x) < E_k, V(y) < E_k$

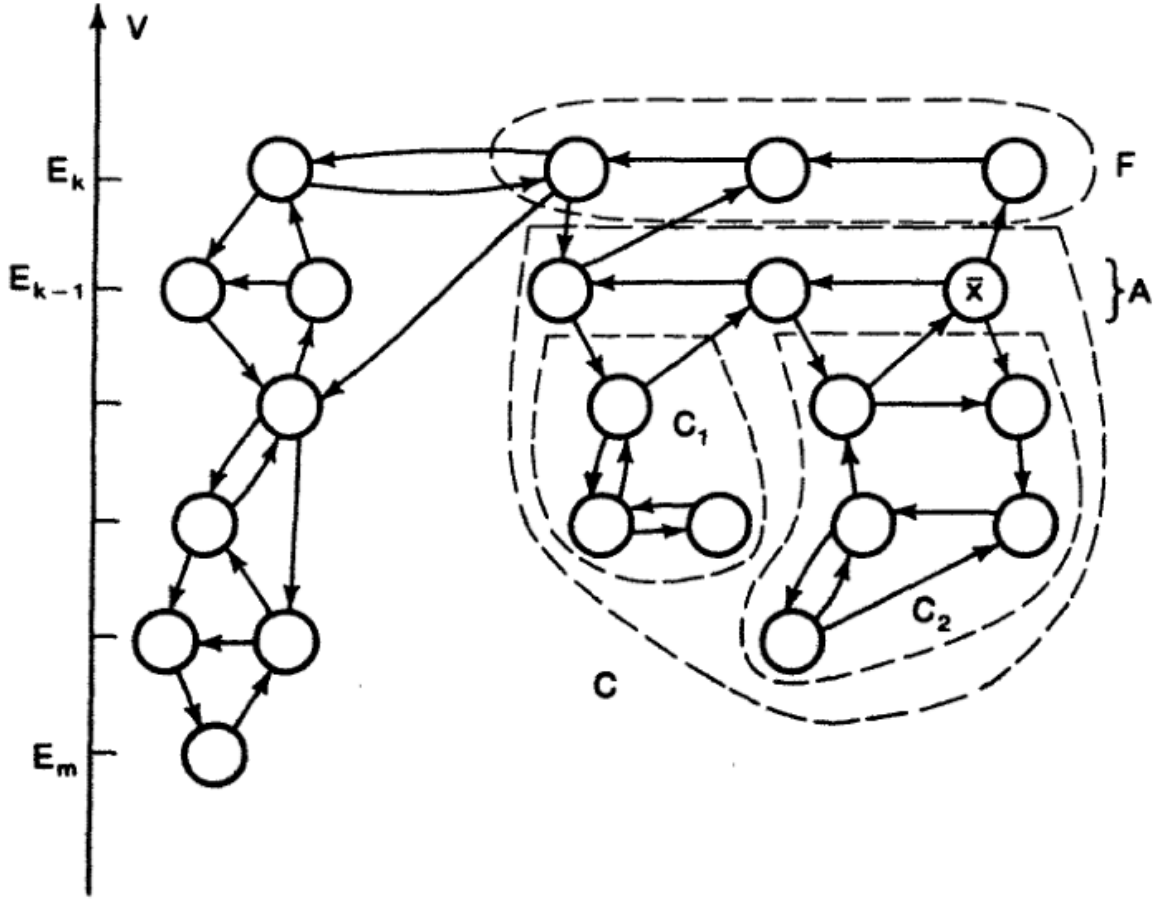


図4 谷 C の、最上層 A と最上層に直接届く谷への分割. A 内の状態 \bar{x} は F から直接届くことを示す

は同じ (分割されたあとの) 谷集合 C_i に含まれるのは x から y へと高さ E_{k-1} で到達可能であるのと同値である。

それぞれの i における C_i の深さはせいぜい g であり、 C の深さ d より小さい。したがって、それぞれの i における C_i に対して帰納法の仮定を適用できる。このようにして各 i に対して ϵ_i と $\bar{\lambda}_i$ を得ることができ、その中で最小の ϵ と $\bar{\lambda}$ を得ることができる。また、それは以下の性質を持つ Θ と C にしか依存していない。もし $\lambda_{t_o} \leq \bar{\lambda}$ で $x_o \in C_i$ となる i が存在するならば、

$$E \left[\int_{t_o}^{\tau(C_i)} \lambda_s^g ds | (U_0, Y_0) = (t_o, x_o) \right] \leq \frac{1}{\epsilon}$$

かつ、ある $x \in C_i$ に対して $y \in A \cap N(x)$ なる y それぞれに

$$P[C_i \text{ から飛び出し } y \text{ へと当たる} | (U_0, Y_0) = (t_o, x_o)] \geq \epsilon$$

が成り立つ。

3.2 定理3の証明 Part.a

この説では帰納法の仮定が成り立つとして定理3のPart aの証明を行う。その結果は§3.1に述べた、 $A^0 = \{x \in A : N(x) \cap F \neq \emptyset\}$

補題 3.2.1 $K = (\min\{k > 0 : Y_k \in A^0\} \wedge W) + 1$ とする。 Θ と各 $x_o \in C$ 、各 $t_o > 0$ であるような C にしか依存

しない D_1 が存在する. $\lambda_{t_o} \leq \bar{\lambda}$ を用いて

$$E \left[\int_{t_o}^{U_K} \lambda_t^d I_{\{t \leq U_W\}} dt | (U_0, Y_0) = (t_o, x_o) \right] \leq D_1 \lambda_{t_o}^u \quad (3.1)$$

証明

$J_0 = 0$ とし, $0 < i$ で $J_{i+1} = \min\{k > J_i : Y_k \in A\} \wedge W$ とし, 離散時間ランダム過程 $\tilde{Y} = (\tilde{Y}_k)$ (ただし, $\tilde{Y}_k = Y_{J_k}$) を考える. §3.1 で述べた帰納法の仮定と Y がいずれかの谷 C_i を抜け出すとき最初に A の状態を訪れるという事実は \tilde{Y} が次のような性質を持つという意味も持っている.

- A のどの状態も他の A の状態からでも到達可能である (1 ステップで到達可能であるという必要はない).
- Θ と C のみに依存している ϵ が存在する. C は \tilde{Y} の直接遷移が可能であるような A の状態 x, y について, 事象 $\{\tilde{Y}_k = x\}$ で, $P[\tilde{Y}_{k+1} = y | \mathcal{F}_k] \geq \epsilon$ である. ただし, \mathcal{F}_k は σ -集合代数で $(U_t : 0 \leq i \leq J_k + 1)$ と $(Y_t : 0 \leq i \leq J_k)$ によって生成されるものである.

...