■1 群（信号・システム）－ 1 編（情報理論）

3 章 情報源モデルとエントロピーレート

（執筆者：西新幹彦）[2011 年 6 月 受領]

■概要■

本章では，情報理論において情報源を定義するのに必要な事項を述べ，情報源のエントロピーレートや漸近等分割性について解説する．また，各種の情報源を紹介し，それらに対するエントロピーレートや漸近等分割性について説明する．本章に登場する情報源は，定常情報源，定常無記憶情報源，エルゴード情報源，AMS 情報源，単純マルコフ情報源，多重マルコフ情報源，隠れマルコフ情報源，ユニフィラー情報源，有限状態情報源，木情報源，FSMX 情報源，単語情報源，一般情報源である．

【本章の構成】

本章では，定常情報源とエントロピーレート（3-1 節）について定義したあと，各種の情報源モデル（3-2 節）を紹介し，それらに関する定理（3-3 節）について述べる．

■1 群 - 1 編 - 3 章

3-1 情報源とエントロピーレート

（執筆者：西新幹彦）［2011 年 6 月 受領］

3-1-1 情報源の数学的モデル

情報を工学的に取り扱うため，情報源は数学的モデルによって表現される．情報源が出力する値，値が選ばれる法則，値が出力される時刻などを数学的に表現することによって情報源が定義される．

(1) 情報源アルファベット

情報源が出力する値の集合を情報源アルファベット（Alphabet）といい，その要素をシンボル（Symbol）または文字（Letter），記号という．情報源アルファベットが可算集合，非可算集合であるような情報源をそれぞれ離散情報源，連続情報源と呼ぶ．また，それぞれの情報源アルファベットを離散アルファベット，連続アルファベットと呼ぶことがある．情報源アルファベットが可算集合のとき，それが有限集合か無限集合かによって一般には異なる性質を持つ．基礎的な考察では情報源アルファベットとして有限集合を仮定する場合が多いが，一般論を論ずるときには可算無限集合や非可算無限集合まで範囲が広げられる．本章では，有限集合を仮定して解説する．

(2) 情報源の確率的性質

通常，情報源は次々にシンボルを出力し，それらのシンボルは確率的に決定されるとみなされる．よって，情報源の出力は確率変数として表される．このことから，多くの場合，情報源は確率過程として表現される．これは情報源の出力が整合性（Consisitency）を持つと仮定することに等しい．すなわち，情報源から出力される複数の連続したシンボルを考えた場合，それらの分布がコルモゴロフの拡張定理を満たすと仮定される．このとき，情報源はシンボルの無限系列全体のうえに定義された一つの確率分布によって表現される．本章では特に断らない限り整合性を持つ情報源を考えるが，14 章にあるように，整合性を仮定しない情報源に対する一般的な理論体系もある．

(3) 情報源と時間

シンボルが一定間隔で出力されるような情報源を離散時間情報源という．情報処理機器は機器に固有のクロックによって処理手順を進めていることから，離散時間情報源を考えることには合理性がある．このとき，シンボルの間隔は便宜的に 1 秒と仮定されたり，あるいは未定義のまま単に単位時間（Unit Time）と呼ばれたりすることが標準的である．一方，電気回路の電圧のようにシンボルが常に出力されている場合や，シンボルが不定期に出力される場合など，シンボルの出力される時刻を実数として表現する連続時間情報源を考えることもある．いずれにせよ，単位時間が実際に何秒であるかや時間の単位が何であるかは理論的な結果とは本質的に無関係である．本章では，単位時間ごとに一つのシンボルを出力する離散時間情報源について述べる．

3-1-2 情報源の定常性

以降，本章では文字列を表すために $x^n \triangleq x_1 x_2 \ldots x_n$ や $x_m^n \triangleq x_m x_{m+1} \ldots x_n$ という記法を用いる．

いま，情報源 $\boldsymbol{X} = X_1X_2\ldots$ を考える．情報源アルファベットを $\mathcal{X}$ と表す．情報源が定常（Stationary）であるとは，確率法則が時刻に依存しないことである．すなわち，任意の 2 つの自然数 n, l と任意の文字列 $x^n \in \mathcal{X}^n$ に対して

$$\Pr\{X^n = x^n\} = \Pr\{X_{l+1}^{l+n} = x^n\} \tag{3・1}$$

が成り立つことをいう．

3-1-3 情報源のエントロピーレート

情報源 $\boldsymbol{X}$ の先頭の n 個の出力 $X^n = X_1X_2\ldots X_n$ のエントロピーを $H(X^n)$ と書く．極限

$$\lim_{n\to\infty}\frac{1}{n}H(X^n) \tag{3・2}$$

が存在する場合，これを情報源 $\boldsymbol{X}$ のエントロピーレートといい，$H(\boldsymbol{X})$ と表す．

エントロピーレートは情報源符号化の理論的限界を表すことが知られており，工学的に大変重要な意味を持つ．詳細は 5 章を参照されたい．

定常情報源はエントロピーレートを持ち，

$$H(\boldsymbol{X}) = \lim_{n\to\infty} H(X_n|X^{n-1}) \tag{3・3}$$

が成り立つ．このとき

$$\frac{1}{n}H(X^n) \geq H(X_n|X^{n-1}) \tag{3・4}$$

が成り立ち，両辺とも n に関して単調減少である．

■1 群 - 1 編 - 3 章

3-2 情報源モデル

（執筆者：西新幹彦）［2011 年 6 月 受領］

情報理論の研究は定常無記憶情報源を対象として始まったと言っても過言ではない．研究が進歩するにつれて対象となる情報源のクラスは次々と広がっていき，複雑かつ一般的なモデルを扱うようになってきた．また，情報源のクラスは，単に現実世界をモデル化するばかりではなく符号化アルゴリズムなどの理論的な性能評価をするために，いろいろなものが考案され導入されてきた．

3-1 節で定常情報源を紹介したが，この節ではその他の各種情報源を紹介する．

3-2-1 定常無記憶情報源

情報源の出力 $X_1, X_2, \ldots$ が互いに独立ですべて同一の分布に従うとき，これを定常無記憶情報源（Stationary Memoryless Source）という．このとき $Q(x) = \Pr\{X_1 = x\}$ とすれば

$$\Pr\{X^n = x^n\} = \prod_{i=1}^{n} Q(x_i) \tag{3･5}$$

が成り立つ．このような情報源は i.i.d.（independently and identically distributed）情報源と呼ばれることも多い．

定常無記憶情報源は，本章に登場するあらゆる情報源の特殊な例である．したがって，あらゆる情報源の特徴をすべて備えていると言ってよい．この意味で，定常無記憶情報源は理論の基礎となる大変重要な情報源クラスである．定常無記憶情報源で成り立つ特徴がどのくらい広いクラスの情報源で成り立つかを見極めることは，情報理論研究における定石の一つである．

3-2-2 エルゴード情報源

情報源アルファベット $\mathcal{X}$ 上の無限系列 $x_1x_2\ldots$ に対して左シフト変換 T を $T(x_1x_2\ldots) \triangleq x_2x_3\ldots$ と定める．情報源を表す確率空間において変換 T がエルゴード的であるとき，その情報源はエルゴード情報源（Ergodic Source）と呼ばれる．ここに，確率空間 $(\Omega, \mathcal{F}, p)$ において変換 T がエルゴード的であるとは，$T^{-1}A = A$ なる任意の $A \in \mathcal{F}$ に対して $p(A) = 0$ または $p(A) = 1$ が成り立つことをいう．

工学では，多くの観測値の算術平均を求めることで，期待値の代用とする場合が多い．このことは「時間平均と集合平均を同一視する」とも言われる．定常エルゴード情報源は，次の定理によってこの考え方が正しいことが保証されている．

定理（G. D. Birkhoff の個別エルゴード定理（The Pointwise Ergodic Theorem）） 定常情報源 $\boldsymbol{X}$ がエルゴード情報源であるための必要十分条件は，任意の自然数 k と $\mathcal{X}^k$ 上に定義された任意の可積分関数 f に対して

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(X_{1+i}^{k+i}) = \mathbb{E}[f(X^k)] \tag{3･6}$$

が確率 1 で成り立つことである．

また，非エルゴードな定常情報源は複数の定常エルゴードな成分に分解して扱えることが知られている．これを定常情報源のエルゴード分解（The Ergodic Decomposition）という．その帰結として次の定理により，観測値の算術平均の極限が存在してその値が定常エルゴードな分布による期待値に等しいことが言える．

定理（定常情報源のエルゴード分解） 情報源 $\boldsymbol{X}$ が定常ならば，任意の自然数 k と $\mathcal{X}^k$ 上に定義された任意の可積分関数 f に対して

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(X_{1+i}^{k+i}) = \mathbb{E}_p[f(X^k)] \tag{3・7}$$

が確率 1 で成り立つ．ここに，p は $X_1, X_2, \ldots$ の値から導かれる定常エルゴードな情報源を表す分布で，$\mathbb{E}_p[\cdot]$ は確率変数が分布 p に従うと仮定した場合の期待値を表す．

3-2-3 AMS 情報源

情報源 $\boldsymbol{X}$ が AMS（Asymptotically Mean Stationary）であるとは，任意の長さ k の系列 $x^k \in \mathcal{X}^k$ に対して，極限

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} \Pr\left\{X_{1+i}^{k+i} = x^k\right\} \tag{3・8}$$

が存在することである．このとき式 (3・8) は定常な確率過程を表す分布となり，$\boldsymbol{X}$ の定常平均と呼ばれる．

AMS 情報源に対してもエルゴード分解ができる．

定理（AMS 情報源のエルゴード分解） 情報源 $\boldsymbol{X}$ が AMS ならば，任意の自然数 k と $\mathcal{X}^k$ 上に定義された任意の可積分関数 f に対して

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1} f(X_{1+i}^{k+i}) = \mathbb{E}_p[f(X^k)] \tag{3・9}$$

が確率 1 で成り立つ．ここに，p は $\boldsymbol{X}$ の定常平均である．

AMS 情報源のクラスはすべての定常情報源とすべてのエルゴード情報源を含んでいる．

3-2-4 単純マルコフ情報源

情報源 $\boldsymbol{X} = X_1X_2\ldots$ がマルコフ連鎖 $X_1 \to X_2 \to \ldots$ をなすとき，単純マルコフ情報源または単にマルコフ情報源（Markov Source）という．マルコフ連鎖では，時刻 $i \geq 2$ の出力 X_i が決まる際に時刻 $i-1$ より前の出力は影響しない．すなわち

$$\Pr\{X_i = x_i | X^{i-1} = x^{i-1}\} = \Pr\{X_i = x_i | X_{i-1} = x_{i-1}\} \tag{3・10}$$

が成り立つ．式 (3・10) の右辺が i によらず

$$\Pr\{X_i = x' | X_{i-1} = x\} = M(x'|x) \tag{3・11}$$

と表されるとき，マルコフ連鎖は斉次（Homogeneous）であるといわれ，行列 $M = (M(x'|x))$ は遷移確率行列（Transition Probability Matrix）と呼ばれる．このときシンボルは情報源の状態（State）と呼ばれることがある．これにより「マルコフ情報源の次の出力の分布は現在の状態によって決まる」と言うことができる．斉次マルコフ情報源の初期分布 $\Pr\{X_1 = x_1\}$ と遷移確率行列 M が与えられると

$$\Pr\{X^n = x^n\} = \Pr\{X_1 = x_1\} \prod_{i=2}^{n} M(x_i|x_{i-1}) \tag{3・12}$$

が成り立つ．$\pi = \pi M$ を満たす分布 π を斉次マルコフ連鎖の定常分布（Stationary Distribution）といい，定常分布を初期分布とすると情報源は定常になる．また，任意の状態から任意の状態に有限の遷移で到達する確率が正である斉次マルコフ連鎖は既約（Irreducible）であるという．既約な斉次マルコフ連鎖の定常分布は唯一であり，この定常分布を初期分布とする情報源は定常エルゴード情報源となる．

3-2-5 多重マルコフ情報源

情報源 $\boldsymbol{X} = X_1X_2\ldots$ が k 次の多重マルコフ連鎖をなすとき，k 次の多重マルコフ情報源という．確率変数の列 $X_1, X_2, \ldots$ が k 次の多重マルコフ連鎖をなすとは，時刻 $i > k$ の出力 X_i が決まる際に時刻 $i - k$ より前の出力は影響しないこと，すなわち

$$\Pr\{X_i = x_i|X^{i-1} = x^{i-1}\} = \Pr\{X_i = x_i|X_{i-k}^{i-1} = x_{i-k}^{i-1}\} \tag{3・13}$$

が成り立つことをいう．式 (3・13) の右辺が i によらず

$$\Pr\{X_i = x'|X_{i-k}^{i-1} = x^k\} = M(x'|x^k) \tag{3・14}$$

と表されるとき，多重マルコフ連鎖は斉次であるといわれる．式 (3・14) の $x^k \in \mathcal{X}^k$ を状態と呼ぶことがある．また，k 次の多重マルコフ情報源は記憶の長さが k であるといわれることがある．$\mathcal{X}^k$ 上の分布 π が k 次の多重マルコフ連鎖の k 次の定常分布であるとは

$$\sum_{x_1 \in \mathcal{X}} \pi(x_1^k) M(x_{k+1}|x_1^k) = \pi(x_2^{k+1}), \quad x_2^{k+1} \in \mathcal{X}^k \tag{3・15}$$

が成り立つことである．k 次の定常分布を初期分布 $\Pr\{X^k = x^k\}$ とすると情報源は定常になる．

連続する k 個の出力 X_i^{i+k-1} をひとまとめにして i 番目の出力 X_i' とみなすと情報源 $\boldsymbol{X}' = X_1'X_2'\ldots$ は単純マルコフ情報源となる．

1 次の多重マルコフ情報源とは，単純マルコフ情報源のことである．

3-2-6 隠れマルコフ情報源

アルファベット $\mathcal{W}$ に値をとる確率変数列がマルコフ連鎖 $W_1 \to W_2 \to \ldots$ をなしているとき，写像 $f : \mathcal{W} \to \mathcal{X}$ を定めることによって確率過程 $(X_1, X_2, \ldots) \triangleq (f(W_1), f(W_2), \ldots)$ が定義される．このとき $\boldsymbol{X} = X_1X_2\ldots$ を出力とする情報源を隠れマルコフ情報源（Hidden Markov Source）と呼ぶ．写像 f が一対一であれば $\boldsymbol{X}$ はマルコフ情報源と等価であるが，一

般にはそうではなく，記憶の長さも有限とは限らない．

隠れマルコフ情報源は，観測対象の状態がマルコフ連鎖で遷移しているが，状態を直接観測することはできず，観測できるのは写像を介した値だけであるという状況をモデル化している．隠れマルコフモデルは音声認識や手書き文字認識の分野で多用されている．

3-2-7 ユニフィラー情報源

隠れマルコフ情報源において，直前の状態 W_{k-1} と現在の観測値 $X_k = f(W_k)$ から現在の状態 W_k が一意に定まるとき，すなわちある写像 g によって $W_k = g(W_{k-1}, f(W_k))$ と表されるとき，これをユニフィラー情報源（Unifilar Source）という．初期状態 W_1 を直接観測できないので，状態の列 $W_1W_2\ldots$ も一般には知ることはできない．しかし，初期状態を知ったもとでは状態の列を知ることができる．

もし f が一対一であれば写像 g の存在は自明であり，情報源はマルコフ情報源と等価になる．

3-2-8 有限状態情報源

有限集合 $\mathcal{S}$ を考え，状態集合と呼ぼう．各時刻 $i = 1, 2, \ldots$ に対して状態 $S_i \in \mathcal{S}$ が存在し，時刻 i の出力 X_i の分布は S_i のみに依存して決まり，次の状態 S_{i+1} の分布は X_i, S_i のみに依存して決まるとしよう．このような情報源を有限状態情報源（Finite State Information Source）と呼ぶ．

この情報源は，特殊な場合として，状態集合を情報源アルファベットと等しいものにし，$X_i = S_i$ とすると単純マルコフ情報源となる．また，S_{i+1} の分布が X_i に依存しないとき隠れマルコフ情報源となる．その一方，任意の有限状態情報源に対して，組 (S_i, X_i) を確率変数 W_i と定めると $W_1 \to W_2 \to \cdots$ はマルコフ連鎖をなし，X_i は W_i の写像として定義されるので，$W_1W_2\ldots$ を状態とする隠れマルコフ情報源とみなすことができる．

3-2-9 木情報源

情報源アルファベット $\mathcal{X}$ 上の系列の集合 $\mathcal{S}$ を考える．$\mathcal{S}$ は次の 2 つの性質を満たすとする．(1) どの系列 $s \in \mathcal{S}$ もほかの系列 $s' \in \mathcal{S}$ の語尾ではない．(2) 任意の左半無限系列 $\cdots x_{n-2}x_{n-1}x_n$ に対して $\mathcal{S}$ の中にその語尾が存在する．このとき，左半無限系列に対して $\mathcal{S}$ の中に見つかる語尾は唯一である．各 $s \in \mathcal{S}$ に対して $\mathcal{X}$ 上の分布 θ_s を対応させる．いま，情報源の出力を十分長く観測したとすると，その観測値 $\cdots x_{n-2}x_{n-1}x_n$ に対して $\mathcal{S}$ の中にその語尾 s が唯一つ見つかり，それによって情報源の出力 X_{n+1} の分布が

$$\Pr\{X_{n+1} = x_{n+1} | \cdots, X_{n-2} = x_{n-2}, X_{n-1} = x_{n-1}, X_n = x_n\} = \theta_s(x_{n+1}) \tag{3・16}$$

のように決まるとする．このような情報源を木情報源（Tree Source）といい，$\mathcal{S}$ をそのモデルという．モデルの要素は文脈（Context）と呼ばれることがある．

木情報源は多重マルコフ情報源の特殊な例であるが，木情報源を多重マルコフ情報源で表した場合，情報源を記述するのに必要なパラメータの数は増える．

3-2-10 FSMX 情報源

情報源アルファベット $\mathcal{X}$ を持つ木情報源において，そのモデルの各要素 $s \in \mathcal{S}$ と任意の $a \in \mathcal{X}$ に対して sa の語尾が $\mathcal{S}$ に属するとき，これを FSMX 情報源（Finite-state Machine X Source）と呼ぶ．FSMX 情報源は有限状態情報源の特別な例であり，このとき $\mathcal{S}$ は状態集合となる．木情報源では，一般に $\mathcal{S}$ を状態集合とみなすことはできない．

3-2-11 単語情報源

有限集合 $\mathcal{W}$ に値をとる確率過程 $\boldsymbol{W} = (W_1, W_2, \ldots)$ を考える．また，$\mathcal{C}$ を有限集合 $\mathcal{X}$ 上の系列の集合とし，その要素を単語（Word）と呼ぶ．写像 $f : \mathcal{W} \to \mathcal{C}$ を用いると単語の列 $f(W_1), f(W_2), \ldots$ を得るが，これらをすべて連接すると $\mathcal{X}$ 上の系列 $\boldsymbol{X}$ ができる．確率過程であるこの $\boldsymbol{X}$ を $\boldsymbol{W}$ から f によって生成される単語情報源（Word-valued Source）と呼ぶ．もし f が一対一で $\mathcal{C}$ が語頭フリーであれば，すなわちどの単語もほかの単語の語頭でないならば，観測された $\boldsymbol{X}$ の値から元の $\boldsymbol{W}$ の値を復元することができる．しかし，一般には $\boldsymbol{X}$ を元の単語の列に一意に復元することは不可能である．また，W_i に対して $f(W_i)$ は長さ 1 以上の単語なので，$\boldsymbol{W}$ と $\boldsymbol{X}$ では時間の進み方が不均一に異なる．特に，もとになる確率過程 $\boldsymbol{W}$ が定常であっても単語情報源 $\boldsymbol{X}$ は定常であるとは限らない．

この情報源は，$\boldsymbol{W}$ を f によって符号化した結果の符号語系列をモデル化していると言える．

3-2-12 一般情報源

整合性を持つ情報源は一つの確率過程として考えることができる．しかし，確率過程としての情報源の様々な性質は，整合性を持たない極めて一般的な情報源を考えることによって，統一的な見地から相互の関係を明確に議論することができる．

情報源アルファベット $\mathcal{X}$ を考え，自然数 n ごとに $\mathcal{X}^n$ に値をとる確率変数 X^n を考える．言い換えると，自然数 n ごとに $\mathcal{X}^n$ 上の分布 P_n を考える．確率変数の列 $\boldsymbol{X} = \{X^n\}_{n=1}^{\infty}$ を一般情報源（General Source）と呼ぶ．一般情報源を考える際には，$n \neq m$ に対して X^n と X^m の同時分布は考えないため，これを確率空間の列と解釈することもできる．

詳細は 14 章を参照されたい．

■1 群 - 1 編 - 3 章

3-3 情報源モデルに関する定理

（執筆者：西新幹彦）［2011 年 6 月 受領］

本節では，前節で述べた情報源モデルに対する主な定理を紹介する．

3-3-1 漸近的等分割性

情報源 $\boldsymbol{X} = X_1X_2\ldots$ がエントロピーレート $H(\boldsymbol{X})$ を持ち

$$\lim_{n\to\infty}\frac{1}{n}\log\frac{1}{P_{X^n}(X^n)} = H(\boldsymbol{X}) \quad \text{（確率収束）} \tag{3・17}$$

が成り立つとき，情報源は漸近的等分割性（AEP：Asymptotically Equipartition Property）を持つという．ここに，$P_{X^n}(\boldsymbol{x}) \triangleq \Pr\{X^n = \boldsymbol{x}\}$ とおいた．

本節では，$\boldsymbol{x} \in \mathcal{X}^n$ が ε-標準系列であるとは，

$$\left|\frac{1}{n}\log\frac{1}{P_{X^n}(\boldsymbol{x})} - H(\boldsymbol{X})\right| \le \varepsilon \tag{3・18}$$

を満たすことをいう．長さ n の ε-標準系列全体を $\mathcal{A}_\varepsilon^n$ と表す．

漸近的等分割性から次の 3 つの性質が導かれる．

1. 任意の長さ n と任意の ε-標準系列 $\boldsymbol{x} \in \mathcal{A}_\varepsilon^n$ に対して

$$\left|\frac{1}{n}\log\frac{1}{P_{X^n}(\boldsymbol{x})} - H(\boldsymbol{X})\right| \le \varepsilon \tag{3・19}$$

が成り立つ（$\mathcal{A}_\varepsilon^n$ の定義そのものである）．

2. 任意の $\lambda > 0$ と十分大きな n に対して

$$\Pr\{X^n \in \mathcal{A}_\varepsilon^n\} \ge 1 - \lambda \tag{3・20}$$

が成り立つ．

3. 任意の $\lambda > 0$ と十分大きな n に対して

$$(1-\lambda)\exp(n(H(\boldsymbol{X})-\varepsilon)) \le |\mathcal{A}_\varepsilon^n| \le \exp(n(H(\boldsymbol{X})+\varepsilon)) \tag{3・21}$$

が成り立つ．

最初の性質より，どの標準系列もほぼ等しく確率 $\exp(-nH(\boldsymbol{X}))$ を持っていると解釈できる．また 2 番目の性質より，何れかの標準系列が出現する確率が漸近的にほぼ 1 に等しい．更に 3 番目の性質は，標準系列の個数はほぼ $\exp(nH(\boldsymbol{X}))$ であると解釈される．

定常エルゴード情報源が漸近的等分割性を持つことは，シャノン-マクミラン-ブライマン（Shannon-MacMillan-Breiman）の定理として知られている．また，式 (3・17) の左辺が確率 1 で存在するという意味で，AMS 情報源が広義の漸近等分割性を持つことも知られている．

3-3-2 定常無記憶情報源のエントロピーレート

定常無記憶情報源 $\boldsymbol{X} = X_1 X_2 \ldots$ はエントロピーレートを持ち，その値は X_1 のエントロピーに等しい．すなわち

$$H(\boldsymbol{X}) = H(X_1) = \sum_{x \in \mathcal{X}} P_{X_1}(x) \log \frac{1}{P_{X_1}(x)} \tag{3・22}$$

である．

3-3-3 定常マルコフ情報源のエントロピーレート

定常分布 π を初期分布とし，遷移確率行列 $M(x'|x)$ を持つ単純斉次マルコフ情報源 $\boldsymbol{X} = X_1 X_2 \ldots$ のエントロピーレートは

$$H(\boldsymbol{X}) = H(X_2|X_1) = \sum_{x \in \mathcal{X}} \pi(x) \sum_{x' \in \mathcal{X}} M(x'|x) \log \frac{1}{M(x'|x)} \tag{3・23}$$

によって求められる．

また，定常かつ斉次な k 次多重マルコフ情報源のエントロピーレートについても同様に

$$H(\boldsymbol{X}) = H(X_{k+1}|X^k) = \sum_{x^k \in \mathcal{X}^k} \pi(x^k) \sum_{x' \in \mathcal{X}} M(x'|x^k) \log \frac{1}{M(x'|x^k)} \tag{3・24}$$

によって求められる．

3-3-4 隠れマルコフ情報源のエントロピーレート

確率過程 $\boldsymbol{W} = (W_1, W_2, \ldots)$ が定常マルコフ連鎖ならば隠れマルコフ情報源 $\boldsymbol{X} = X_1 X_2 \cdots = f(W_1) f(W_2) \ldots$ も定常なので，エントロピーレートを持つ．しかし，定常マルコフ情報源のようなエントロピーレートの計算式は見つかっていない．その代わり，$H(X_n|X^{n-1}W_1)$ という量が単調増加で $H(\boldsymbol{X})$ に収束することが分かっている．$H(X_n|X^{n-1})$ が単調減少で $H(\boldsymbol{X})$ に収束するので，この両者の値を計算しながら任意の精度でエントロピーレートを計算することができる．

3-3-5 単語情報源の漸近等分割性

確率過程 $\boldsymbol{W}$ が AMS ならば $\boldsymbol{W}$ から写像 $f : \mathcal{W} \to \mathcal{C}$ によって生成される単語情報源 $\boldsymbol{X}$ はエントロピーレートを持つ．更に，$\boldsymbol{W}$ が定常エルゴードならば $\boldsymbol{X}$ は漸近等分割性を持つ．また更に，f が一対一で $\mathcal{C}$ が語頭フリーであれば，単語情報源のエントロピーレート $H(\boldsymbol{X})$ と $\boldsymbol{W}$ のエントロピーレート $H(\boldsymbol{W})$ の間には

$$H(\boldsymbol{X}) = \frac{H(\boldsymbol{W})}{\mathbb{E}[|f(W_1)|]} \tag{3・25}$$

の関係が成り立つ．

■参考文献

1) T.M. Cover and J.A. Thomas：“Elements of Information Theory, 2nd ed.,” John Wiley & Sons, 2006.
2) R.M. Gray：“Probability, Random Processes, and Ergodic Properties, 2nd ed.,” Springer-Verlag, 2009.
3) 韓　太舜, 小林欣吾：“情報と符号化の数理,” 培風館, 1999.
4) 韓　太舜：“情報理論における情報スペクトル的方法,” 培風館, 1998.
5) F.M.J. Willems, Y.M. Shtarkov, and T.J. Tjalkens：“The Context-Tree Weighting Method: Basic Properties,” IEEE Trans. Inform. Theory, vol.41, no.3, pp.653–664, 1995.
6) J. Rissanen：“Complexity f Strings in the Class of Markov Sources,” IEEE Trans. Inform. Theory, vol.IT-32, no.4, pp.526–532, 1986.
7) R. Timo, K. Blackmore, and L. Hanlen：“Word-Valued Sources: An Ergodic Theorem, an AEP, and the Conservation of Entropy,” IEEE Trans. Inform. Theory, vol.56, no.7, pp.3139–3148, 2010.