シャノン流情報理論からみた水文現象の不確定構造

京都大学工学部　正員　　髙棹琢馬
京都大学工学部　正員　○池淵周一

1. はじめに

情報理論はシャノン，ウイナーによってその基礎が確立され，それぞれ独立に発展してきた。すでにウイナー流情報理論はフィルター理論として水文現象への適用が試みられ，有用な情報を提供しているが，ブラックボックス的立場はまぬがれない。一方，シャノン流情報理論は確率系の内部構造に立入る積極さをもっており，水文現象の不確定構造と考え合わせれば，本理論の水文学への適用性は十分あるものと確信する。

本研究は，こうした観点からシャノン流情報理論の基本概念を述べるとともに，ウイーナー流情報理論との関係，さらには不確定な水文現象のエントロピー的解釈をいくつか試みたものである。

2. エントロピーとその諸性質

1) 定義式

ある事象 x が起こる確率を p とすれば，x の実現に関する情報量 I(x) は

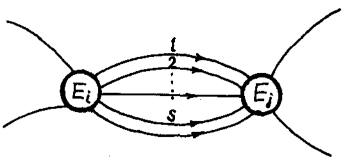
$$I(x) = -\log p \tag{1}$$

で与えられる。この概念を基礎に，情報源の期待情報量としてエントロピーがつぎのように定義される。

i) 離散的な独立型情報源；n 個の異なった値 x_i のみをとり，各値を p_i なる確率で独立に発生する情報源で，エントロピーは次式で与えられる。

$$H = -\sum_{i=1}^{n} p_i \log p_i \tag{2}$$

ii) 離散的なマルコフ型情報源；この場合には，ある確立事象が起こることを図－1のシャノン線図上で一つの状態から一つの状態へ遷移することと考え，その起こり方を状態から状態への遷移確率によって表現する。すなわち，一つの状態 E_i からほかの状態 E_j に第 s 番目の有向線を矢印の方向にたどっていく条件つき確率を $p_i(j, s)$，E_i から E_j への有向線 s に対応して情報源から生起する事象を $L_{ij}^{(s)}$，および状態 E_i にある確率を p_i とすると，この情報源のエントロピーが次式で与えられる。



第 1 図

$$H = -\sum_{i,j,s} p_i \cdot p_i(j, s) \log p_i(j, s) \tag{3}$$

以上は離散的な場合であったが，確率分布が連続的な場合にも，各エントロピーは積分の形で表現することによって，同様に拡張できる。

2) 諸性質

i) 上記の定義式は統計力学でいうエントロピーと同じ形をしており，情報源の中の一つの事象が平均として持っている情報量を表わし，通常，情報量としてのエントロピーと呼んでいる。この量は情報源の中から特定の事象を選ぶときの自由度，あるいはそれを推定するときの不確かさを表わすといってもよい。

ii) 確率的な構造が一様なほど，エントロピーは大きくなる。たとえば，正規分布，指数分布に対してHの値はそれぞれ，$\frac{1}{2}(1+\log 2\pi\sigma^2)$ および $1+\log 1/\alpha$ となり，分散が大きいほどエントロピーは大きくなる。

iii) 時間的特性としてはエントロピーが大きいほど，その過程は定常的である。すなわち，確率的構造が安定し，大数の法則が成立するエルゴード過程に漸近する。

iv) $1 - H/H_{Max}$ を冗長度といい，その大きさは確率過程の再現性に関連しうる。

v) 入，出力情報源とか，シャノン線図の状態と遷移とかのように2つの確率事象 x および y を考えると，5種類の確率分布 $p(x)$，$p(y)$，$p(x, y)$，$p_x(y)$，$p_y(x)$ に相当するエントロピーがつぎのように定義される。

$$
\left.\begin{aligned}
&H(x) = -\sum_{x} p(x)\log p(x), \quad H(y) = -\sum_{y} p(y)\log p(y)\\
&H(x,y) = -\sum_{x,y} p(x,y)\log p(x,y)\\
&H_x(y) = -\sum_{x,y} p(x,y)\log p_x(y), \quad H_y(x) = -\sum_{x,y} p(x,y)\log p_y(x)
\end{aligned}\right\} \tag{4}
$$

ここに，$H(x,y)$ は相互エントロピー，$H_x(y)$ および $H_y(x)$ は条件つきエントロピーあるいは曖昧度とよばれている。なお，これらのエントロピーの間には上記確率分布の相互関係からつぎの関係式がある。

$$H(x) + H(y) \geq H(x,y) = H(x) + H_x(y) = H(y) + H_y(x) \tag{5}$$

不等号の等式は x と y が独立な事象の場合のみ成立する。

3. エントロピーと相関関数，線形フィルタ

シャノン流情報理論とウイーナー流情報理論の関係を論じたものはほとんどない。ここでは両者を関係づける若干の考察を試みる。

1) エントロピーと相関関数

エルゴード過程においては，時間平均と集合平均が一致するから，自己相関関数および相互相関関数の集合平均を与えることができる。

いま，確率過程 $x(t)$, $y(t)$ において $x(t)=x_1$, $x(t+\tau)=x_2$, $y(t+\tau)=y_1$ とすると，相互エントロピー $H(x_1, x_2 : \tau)$ および $H(x_1, y_1 : \tau)$ はそれぞれ次式で定義される。

$$H(x_1, x_2 : \tau) = -\iint p(x_1, x_2 : \tau)\log p(x_1, x_2 : \tau)\, dx_1 dx_2 \tag{6}$$

$$H(x_1, y_1 : \tau) = -\iint p(x_1, y_1 : \tau)\log p(x_1, y_1 : \tau)\, dx_1 dy_1 \tag{7}$$

ここに，$p(x_1, x_2 : \tau)$ および $p(x_1, y_1 : \tau)$ は同時確率分布であるが，いま，もっとも基本的な二変数正規分布を考えると，上式はそれぞれ

$$H(x_1, x_2 : \tau) = 1 + \log 2\pi\sqrt{\sigma_1^4 - \varphi_{11}(\tau)^2} \tag{8}$$

$$H(x_1, y_1 : \tau) = 1 + \log 2\pi\sqrt{\sigma_1^2\sigma_2^2 - \varphi_{12}(\tau)^2} \tag{9}$$

となる。すなわち，各々の相互エントロピーが，自己相関関数 $\varphi_{11}(\tau)$ および相互相関関数 $\varphi_{12}(\tau)$ と結びつけられたわけである。もちろん，(5)式の関係を使えば $H_x(y)$, $H_y(x)$ も相関関数と結びつけられる。なお，σ_1^2, σ_2^2 は x および y の分散である。二変数指数分布についてはベッセル関数を含むので，両者の関係は解析的には求められないが数値的には可能である。

いずれにしても，相関関係が小さいほど相互エントロピーおよび条件つきエントロピーは大きくなり，両者の対応関係が無関係になっていく。

なお，相関関数のフーリエ変換がスペクトルであるので，エントロピーとスペクトルも関係づけられる。

2) 線形フイルタとエントロピー

いま，n 次元の座標 $\mathbf{x}$ から n 次元の他の座標 $\mathbf{y}$ に移る座標変換を行なえば，新しい座標 $\mathbf{y}$ に関する確率密度関数 $g(\mathbf{y})$ は

$$g(\mathbf{y})\, d\mathbf{y} = f(\mathbf{x})\, J(\mathbf{x}/\mathbf{y})\, d\mathbf{y} \tag{10}$$

で与えられる。ここに，$J(\mathbf{x}/\mathbf{y})$ は座標変換のヤコビ行列式である。したがって，新しい座標 $\mathbf{y}$ でのエントロピー $H(\mathbf{Y})$は $d\mathbf{x} = J(\mathbf{x}/\mathbf{y})\, d\mathbf{y}$ を考慮すると，

$$
\begin{aligned}
H(\mathbf{Y}) &= -\int g(\mathbf{y})\log g(\mathbf{y})d\mathbf{y} = -\int f(\mathbf{x})\log f(\mathbf{x})\, d\mathbf{x} - \int f(\mathbf{x})\log J(\mathbf{x}/\mathbf{y})\, d\mathbf{x}\\
&= H(\mathbf{X}) - \int f(\mathbf{x})\log J(\mathbf{x}/\mathbf{y})\, d\mathbf{x}
\end{aligned} \tag{11}
$$

となる。

さて，線形フイルタも本質的には座標変換と同様な役割をもっているので，帯域幅Wの中に制限されて，単位時間当たり $H(\mathbf{X})$ のエントロピーをもっている入力集合が，伝送特性 $K(\omega)$ の線形フイルタ（ω は周波

数）を通ったとき，出力集合の単位時間当たりのエントロピーH($\mathbf{Y}$)も同様に求められる。すなわち，この場合にはヤコビ行列式J($\mathbf{y}/\mathbf{x}$)が

$$J(\mathbf{y}/\mathbf{x}) = \exp\left[\frac{1}{W}\int_w \log|K(\omega)|^2 \, d\omega\right] \tag{12}$$

となり，座標変換によらない定数となるので，(11)式およびJ($\mathbf{x}/\mathbf{y}$) $=1/$J($\mathbf{y}/\mathbf{x}$) を考慮すると，最終的に次式の関係が得られる。

$$H(\mathbf{Y}) = H(\mathbf{X}) + \frac{1}{W}\int_w \log|K(\omega)|^2 \, d\omega \tag{13}$$

右辺第２項はフイルタに加わる雑音の効果を表現している。

この関係は線形フイルタの場合に限って成り立つのであって，フイルタが線形でない場合にはヤコビ行列式の平均値が元のヤコビ行列式と等しくならないので，このように簡単にはならない。

以上の展開はシヤノン流情報理論がウイーナー流情報理論の概念を包含しており，本理論の内容の豊富さをものがたっている。

4. 水文現象のエントロピー的解釈

いままでのエントロピー概念を具体的に水文現象に適用し，その不確定構造を論じる。

1) 短期流出と長期流出

隣りあった２つの直接流出分の終了時の時間間隔を流出の１サイクルとよぶと，短期流出の解析は１サイクルを対象とするものであり（各サイクル内の直接流出は互いに独立），長期流出の解析は多数のサイクルを対象とするので，その不確定構造はおのずと後者の方がより大きい。このことは長期流出のエントロピーが短期流出のエントロピーより大きいこと，サイクル数が多くなると長期流出のエントロピーはさらに大きくなり，定常的性格が強くなることを意味している。

また，(4)式においてxを降水r，yを流量Qと考えれば，H(r,Q)が大きいことはrとQが独立事象に近くなるから，H(r,Q)がH(r)とH(Q)の和に近づき，不確実さが大きくなることを意味しており，時間スケール的には長期流出を，距離スケール的には大流域の流出現象に相当してくる。逆にH(r,Q)が小さいということは，rとQの従属性が強く，短期流出，小流域の流出に相当している。

同様に，流域が大きくなるほど$H_r(Q)$は大きくなり，流域が小さいほど$H_r(Q)$が小さくなって両者の対応がデタミニスティックになる。これらのことは相関関数で表現した場合には，逆の関係になることが3.1)の展開から理解できる。

2) 流況のマクロ的表現としての０交さエントロピー

流況を表現する尺度として，従来から河状係数や流況曲線が用いられているが，異常値を扱っていたり，時間的プロセスを考慮していなかったりして，流況を知る客観的な尺度とはいいがたい。

ここでは，日流量系列のエントロピー的解釈として，流況特性表現の一つの客観的尺度となりうる０交さエントロピーを提示する。

さて，年ないし季節ごとの平均流量からの変動には時間的なものと量的なものがあるが，これらを合わせて流況の平均流量を横切る点を０点とよび，一つの０点時系列あるいは０点状態系列を考え，この系列においてつぎのことを考える。（図－２）。

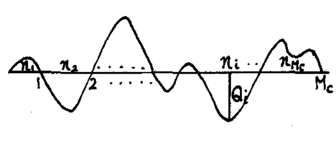
一つの状態の中にはn_iの日数が含まれ，

$$n = n_1 + n_2 + \cdots + n_{MC} \tag{14}$$

である。いま，$n_1 = p_1 \cdot n$，$n_2 = p_2 \cdot n$，…，$n_{MC} = p_{MC} \cdot n$ とおき，nが十分大きいと，この状態系列の日あたりのエントロピーとして，2.1)で定義したように，

$$H_C = -\sum_{i=1}^{MC} p_i \log p_i \tag{15}$$

が得られる。これを０交さエントロピーと名づける。[3)]



図－２

ところで，このH_Cにはエントロピーの性質からもわかるようにつぎのような性質がある。すなわち，i）流域面積が大きくなるとH_Cは大きくなる。これは，日流量に及ぼす種々の因子が流域面積とともに分散していく，つまり曖昧度が大きくなっていくことであり，流域面積が大きくなるにつれて流域の貯留効果が増加して降雨変動が吸収されることも一因していよう。ii）年のエントロピーをH_{cy}，季節のエントロピーをH_{cs}とすると，$H_{cy}/\log 4 + H_{cs}$が1に近づくほど0交さ事象が等確率に近づき，0交さ数M_cが多くなる。このことは同時に，日流量の季節的特性が小さくなることを意味していよう。iii）日流量の標準偏差σと0交さエントロピーH_cとの間には近似的につぎの関係がある。

$$\sigma \propto \exp(-n \cdot H_c/2) \tag{16}$$

これはエントロピーが大きくなると，状態あるいは場合の数が多くなり，大数の法則によって高確率群が確率1でもって1つの所に集まってき，これを除いた低確率群が確率0に近づいていくこと，振幅Q_iと0交さ日数n_iとの関係を調べること，の両者から導かれよう。iv）H_cが大きくなると0交さ回数M_cが多くなり，これはまた日流量系列の自己相関関数のてい減が急になること，あるいはスペクトルが平たくなることにも結びつこう。v）上記の諸性質から，需要水量を年平均流量あるいは季節内平均流量と考えると，H_cは水年あるいは季節内での表流水の資源としての安定性をマクロ的に表現している。

3）流出系の確率的内部構造－状態とその遷移

情報理論でいう雑音のある通信系とは，通信系の状態がβ_tにあるとき，送信信号x_tが送りこまれると受信信号y_tが受信され，通信系の状態がβ_{t+1}に変わる現象が$P_{x_t,\beta_t}(y_t,\beta_{t+1})$で起こるというものである。

われわれが対象としている降水－流出変換系－流量の関係を考えると，雑音のある通信系によく類似している。著者らは，すでに，この類似性に注目し，流出現象をエントロピー的に解釈するとともに，流出系内部の確率的構造解明に状態遷移確率法を提案している。[4] 本研究では，この考えをさらに発展させるべく，図－1のシャノン線図とも対応させながら，長期流出系の確率的内部構造を以下のように考えた。すなわち，流出変換系が有限個の内部状態をもち，その出力に現われる流量は変換系に加わる雑音のために，入力である降水と変換系の内部状態によって確率的に定まる。

内部状態を何で表現するかはきわめて難かしい問題であるが，いま計量可能なものとして流量のm重系列で定義するものとすると，その状態に降雨sが加わると，内部状態は他の内部状態に遷移すると同時に，遅れ時間$t_{ij}^{(s)}$をもつて流量$Q_{ij}^{(s)}$を発生するという状況を想定することができる。そうすると，2.1），ii）で述べたマルコフ型情報源のエントロピー定義式を用いて，単位時間あたりのエントロピーが，次式で与えられる。

$$H' = -\frac{\sum_{i,j,s} p_i\, p_i(j,s) \log p_i(j,s)}{\sum_{i,j,s} p_i\, p_i(j,s)\, t_{ij}^{(s)}} \tag{17}$$

このように考えると，われわれは状態の遷移を遷移確率$p_i(j,s)$で論じようとしているわけで，つぎの段階として上記のエントロピーを用いて遷移確率$p_i(j,s)$を求めなければならない。

ところで，3.1）でも述べたように，長期流出系を対象とする限り，サイクル数がきわめて多く，状態の遷移を多くの系列について調べると，その遷移確率はある定まった値に正確に近づくといえる。このことはとりもなおさず大数の法則の成立を意味し，状態の遷移に関して平衡系を想定しており，上記のエントロピーの極限，すなわち状態遷移に最大エントロピーの仮定をおきうる。

このことを利用すると，降雨に関する情報を満足させながらエントロピーH'を最大にすることによって遷移確率$p_i(j,s)$が次式で求められる。

$$p_i(j,s) = p(s) \cdot e^{-C \cdot t_{ij}^{(s)}} \Big/ \sum_j e^{-C \cdot t_{ij}^{(s)}} \tag{18}$$

ここに，Cは行列式

$$\left| \sum_s \left\{ p(s)\, e^{-C \cdot t_{ij}^{(s)}} \Big/ \sum_j e^{-C \cdot t_{ij}^{(s)}} \right\} - \delta_{ij} \right| = 0 \tag{19}$$

を満たす最大正実根であり，p(s)は降雨sの生起確率，δ_{ij}はクロネッカーのデルタである。

このようにして $p_i(j,s)$ が求められると，シヤノン線図上を連続的に遷移することによって流量系列 $\{Q\}$ が発生することになる。

なお，こうした遷移確率の算出法は Jaynes の最小不偏推定法[5)]にも通じるものがあり，ここにもシヤノン流情報理論の広汎な内容が示されている。

5. あとがき

以上，シヤノン流情報理論の基本概念を述べるとともに，水文現象の不確定構造に関してもこの理論の適用性が十分あることを示唆した。内容の豊富さにもかかわらず，本理論の水文学への適用はまだ緒についたばかりであり，今後は本研究での理論展開を一層きわめるとともに，実流域への適用をはかって具体化していきたい。

〔参考文献〕

情報理論に関しては，たとえば

1) 大泉充郎，本多波雄，野口正一；情報理論，オーム社，昭 37
2) R.M.ファノ著，宇田川銈久訳；情報理論，紀伊国屋書店，昭 40
3) 石原藤次郎，高棹琢馬；表流水源の資源としての安定性，第 21 回土木学会年構，昭 41
4) 高棹琢馬，池淵周一；長期間流出機構の情報理論的研究，京大防災研年報，第 12 号 B. 昭 44
5) Myron Tribus ; Information Theory as the Basis for Thermostatics and Thermodynamics, Trans, ASME, 1961.