数理解析研究所講究録
第504巻 1983年 1-25

Gaussian Channel における
相互情報量とCapacity

愛媛大 理　井原俊輔（Shunsuke Ihara）

§1. はじめに

Gaussian channel は理論的にもまた応用上からも大変興味のある通信路のひとつである．情報理論＝通信の数学的理論の枠組は Shannon [20] によってつくられたが，[20] の後半は Gaussian channel の研究に割かれている．Shannon の仕事は確率論の研究に大きな影響を与えた．その一方，情報理論は確率論や関数解析の興味ある応用の場であり，確率論や関数解析の成果が情報理論に進歩をもたらしている．

ここでは Gaussian channel（以後 GC と略記する）における相互情報量，capacity（容量）について論じる．相互情報量の計算は測度の絶対連続性とその Radon-Nikodym の導関数の計算と密接に関連している．この方面の確率論および関数解析における近年の成果を用いて情報量や capacity を計算する．さらに Gauss過程の Lévy-Hida-Cramér の標準表現を

用いることにより情報量, capacity についてより具体的な結果が得られることも示す.

§2ではGCの数学的記述を与え, 問題を定式化する. §3では相互情報量, §4では capacity についてこれまでに知られていることを述べる. 計算にあたっては確率論的手法と関数解析的手法がある. ここでは前者を中心に述べるが, 両手法の特徴, 違いについても触れる. 定常 Gauss過程, Brown 運動と equivalent (互に絶対連続) なGauss 過程の場合には標準表現の具体的な形が知られている. それを用いて, §5ではGCにおける雑音が定常 Gauss過程の場合に情報量の公式の新しい導き方を与え, §6では雑音が Brown運動と equivalentな Gauss 過程の場合の情報量と capacity について若干の結果を与える.

§2. Gaussian channel

GC (Gaussian channel) は, 入力信号に Gauss型雑音 (つまり雑音が Gauss 過程) がプラスされたものが出力信号として受信される通信路で, 概略次頁の図のように表わされる.

GCの数学的定式化を与えよう. 通信時間は $[0, T]$ ($T < \infty$) とする. θ, $\hat{\theta} = \{\hat{\theta}(t); 0 \leq t \leq T\}$, $X = \{X(t); 0 \leq t \leq T\}$, $Y = \{Y(t); 0 \leq t \leq T\}$, $Z = \{Z(t); 0 \leq t \leq T\}$ で各々

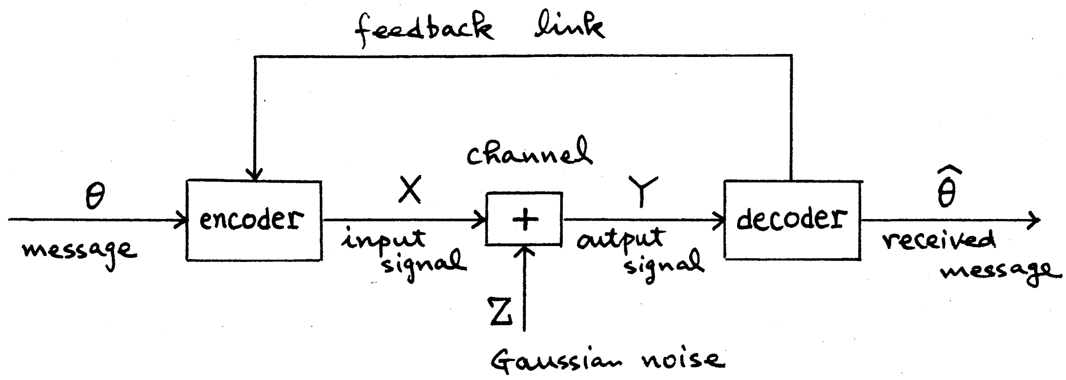

図1. Gaussian channel

message, received message, 入力信号, 出力信号, 雑音を表わす. これらは皆, 基礎確率空間 $(\Omega, \mathfrak{F}, P)$ 上で定義されている実確率過程とする. ただし θ はどんな可測空間の値をとる確率変数でもよく, 有限次元確率変数でも確率過程でもよい. 以下確率過程の平均は常に0と仮定する. こうしても一般性は失われない. GC は次式で与えられる.

$$Y(t) = X(t) + Z(t), \qquad 0 \leq t \leq T. \qquad (1)$$

ただし以下の仮定がみたされているものとする.

(A.1) Z は Gauss 過程.

(A.2) message θ と雑音 Z は独立. ($\theta \perp\!\!\!\perp Z$ と記す)

(A.3) 各 t で, $X(t)$ は $\mathfrak{F}(\theta) \vee \mathfrak{F}_t(Y)$ 可測. ただし $\mathfrak{F}(\theta)$ は θ を, $\mathfrak{F}_t(Y)$ は $Y(s)$, $0 \leq s \leq t$, を可測にする最小の σ-field, $\mathfrak{F}(\theta) \vee \mathfrak{F}_t(Y)$ はこの両者を含む最小の σ-field.

(A.4) 確率方程式(1)は一意な解$Y(\cdot)$をもつ.

仮定(A.3)は直観的に言えば, $X(t)$はθと$Y_0^t \equiv \{Y(s);\ 0 \leq s \leq t\}$の関数であるということである. このことは, この GC は時間の遅れも雑音もない feedback をもつことを意味している.

(A.3)の代りに

(A.3′) 各tで, $X(t)$は$\mathfrak{F}(\theta)$可測.

をみたすとき, feedback のない GC という. このときはθとXは同一視してよい.

(A.1)の代りに

(A.1′) Zは Brown 運動.

をみたす GC を white Gaussian channel (WGC) という.

我々の問題は GC (1) における message θ と output Y との間の相互情報量の計算と channel capacity を求めることである. 相互情報量と capacity の正確な定義を与える.

ξ_i $(i=1,2)$をΩ上で定義され可測空間$(S_i, \mathcal{B}_i)$の値をとる確率変数とする. ξ_iの分布をμ_{ξ_i}とする: $\mu_{\xi_i}(B) = P(\xi_i \in B)$, $B \in \mathcal{B}_i$. ξ_1とξ_2の同時分布をμ_{ξ_1,ξ_2}, μ_{ξ_1}とμ_{ξ_2}の直積測度を$\mu_{\xi_1} \otimes \mu_{\xi_2}$とする. ξ_1とξ_2の間の相互情報量 $I(\xi_1, \xi_2)$ は, $\mu_{\xi_1,\xi_2} \ll \mu_{\xi_1} \otimes \mu_{\xi_2}$ (絶対連続) のときは

$$I(\xi_1, \xi_2) = \int_{S_1 \times S_2} \log \frac{d\mu_{\xi_1,\xi_2}}{d\mu_{\xi_1} \otimes \mu_{\xi_2}}(x_1, x_2)\, d\mu_{\xi_1,\xi_2}(x_1, x_2) \qquad (2)$$

($d\mu_{\xi_1\xi_2}/d\mu_{\xi_1}\otimes\mu_{\xi_2}$ は Radon-Nikodym の導関数)で, $\mu_{\xi_1\xi_2} \not\ll \mu_{\xi_1}\otimes\mu_{\xi_2}$ のときは $I(\xi_1,\xi_2)=\infty$ である.

GC (1) における message θ と時刻 t までの output $Y_0^t \equiv \{Y(s);\ 0\leq s\leq t\}$ の間の相互情報量を $I_t(\theta, Y)=I(\theta, Y_0^t)$ で表わす. 同様に $I_t(X,Y)=I(X_0^t, Y_0^t)$ と記す.

この channel によって伝送し得る最大の相互情報量が capacity である. 普通ある一定の条件, それを $\mathbb{A}$ とする, をみたす signal のみが送信可能である. そこで, feedback のある GC (1) の制限 $\mathbb{A}$ の下での(時刻 t までの) capacity を

$$C_t^f \equiv C_t^f(\mathbb{A}) = \sup I_t(\theta, Y) \qquad (3)$$

と定義する. ここで上限は (A.2)~(A.4) をみたしかつ $X\in\mathbb{A}$ である θ と X の全体についてとる. 上限をとる範囲を (A.2)~(A.4) の代りに (A.2), (A.3′) をみたすものにしたのが feedback のない場合の capacity $C_t^\circ \equiv C_t^\circ(\mathbb{A})$ である. feedback のないときは message θ と input X を同一視してよいから

$$C_t^\circ = C_t^\circ(\mathbb{A}) = \sup\{I_t(X,Y);\ X \perp\!\!\!\perp Z,\ X\in\mathbb{A}\} \qquad (4)$$

になる. もちろん $C_t^\circ(\mathbb{A}) \leq C_t^f(\mathbb{A})$ である.

注意 我々は signal や雑音を確率過程としたが feedback のない場合, 数学的にはより一般に, ある線形空間(たとえば $L^2[0,T]$)の値をとる確率変数として扱うことができる.

§3. Gaussian channel における相互情報量

GCの研究はShannon [20] に始まる．Shannonはsignal，雑音が帯域制限定常過程の場合に，いわゆる標本化定理により問題を離散時間定常過程の場合に帰着することにより，情報量やcapacityを計算した．次の進展はKolmogorovを中心とするソ連の数学者達（[15, 7, 18]）によってなされた．それはGauss過程をKarhunen-Loève式に直交展開することによりなされた．この方法はfeedbackのない場合に有効である．Bakerの研究はこの流れをくんでいる．これについては本節後半で述べる．

これらとは異なり，Wiener測度との絶対連続性およびそのときのRadon-Nikodymの導関数に関する結果を使うことによる情報量の計算がKadota 他 [14]，Duncan [4,5] によってなされた．それを紹介する前に次の基本的な事実に注意しておく．Gauss過程 $Z = \{Z(t)\}$ の共分散関数を再生核にもつ再生核Hilbert空間 $\mathcal{H}(Z)$ を Z のRKHSという．

<u>定理 3.1</u>（[19, 10, 2]） feedbackのないGC (1) において，input X はGauss過程とする．$I(X,Y) \equiv I_T(X,Y) < \infty$ であるための必要かつ十分条件は，確率1でsample path $X(\cdot,\omega)$ が雑音 Z のRKHS $\mathcal{H}(Z)$ に属することである．

Brown運動 $B = \{B(t)\,;\, 0 \leq t \leq T\}$ に対しては

$$\mathcal{H}(B) = \{ G = \{G(t), 0 \leq t \leq T\} ; \exists g \in L^2[0,T] \quad G(t) = \int_0^t g(u)\,du, \ \forall t \} \tag{5}$$

が知られている．そこでWGC（white Gaussian channel）の場合，feedbackがありinput Xが必ずしもGauss過程でないときでも，Xは次の形をしているとしても不自然ではない．

$$X(t,\omega) = \int_0^t x(u,\omega)\,du, \quad 0 \leq t \leq T, \quad \omega \in \Omega, \tag{6}$$

ここで$x(t)$は$\mathcal{F}(\theta) \vee \mathcal{F}_t(Y)$可測であり$E[\int_0^T |x(u)|^2 du] < \infty$である．従ってWGCの式は(1)の代りに次のようになる．

$$Y(t) = \int_0^t x(u)\,du + B(t), \quad 0 \leq t \leq T. \tag{7}$$

次の結果はGCの情報量，capacityの計算の基本である．

<u>定理3.2</u>（[14]）feedbackのあるWGC (7)において

$$I_t(\theta, Y) = \frac{1}{2}\int_0^t E|x(u) - \hat{x}(u)|^2 du, \quad 0 \leq t \leq T, \tag{8}$$

ただし$\hat{x}(u) = E[x(u) | \mathcal{F}_u(Y)]$は条件付平均値である．

証明は[16, 17]などにも出ている．

<u>注意</u> 時刻uまでのoutput Y_0^uを観測しそれに基づいて$x(u)$を推定することをfilteringというが，$\hat{x}(u)$はその際の2乗平均誤差を最小にする推定値である．(8)の右辺の$E|x(u) - \hat{x}(u)|^2$はその推定誤差である．このように情報量の計算とfilteringの問題が直接結びついていることがわかる．この方面でいくつかの研究があるが本論文ではとりあげない．

一般のGC (1) を考える. Hitsuda [10] はGauss過程の標準表現を使うことにより話をWGCの場合に帰着させ情報量を計算した. 今Gauss過程 $Z = \{Z(t)\}$ が

$$Z(t) = \sum_{i=1}^{N} \int_0^t F_i(t,u)\, dB_i(u) \qquad (9)$$

とLévy-Hida-Cramérの意味で標準表現されているとする. $F_i(t,u)$ は標準核と呼ばれるVolterra核, dB_i は互に独立なwhite Gaussian noise, $dm_i(u) = E|dB_i(u)|^2$ は continuous measure で $dm_i \gg dm_{i+1}$ となっている. そして重要なことは, $\mathbf{B}(t) \equiv (B_1(t), \cdots, B_N(t))$ とおくとき

$$\mathfrak{F}_t(Z) = \mathfrak{F}_t(\mathbf{B}), \quad 0 \le t \le T, \qquad (10)$$

となっていることである. 一般にはGauss過程の標準表現においては離散スペクトルに対応する成分が存在する. この成分がないと仮定すると標準表現は (9) のようになる. (標準表現については [8] 参照). (9) の Z の RKHS $\mathcal{H}(Z)$ は

$$\mathcal{H}(Z) = \Big\{ G(\cdot)\ ;\ G(t) = \sum_{i=1}^{N} \int_0^t F_i(t,u)\, g_i(u)\, dm_i(u),\ \forall t, \ \sum_{i=1}^{N} \int_0^T |g_i(u)|^2 dm_i(u) < \infty \Big\} \qquad (11)$$

で与えられる. そこで input signal X は

$$X(t,\omega) = \sum_{i=1}^{N} \int_0^t F_i(t,u)\, x_i(u,\omega)\, dm_i(u) \qquad (12)$$

なる形をしているものと仮定する. ただし $x_i(u)$ は $\mathfrak{F}(\theta) \vee \mathfrak{F}_u(Y)$ 可測で $\sum_{i=1}^{N} \int_0^T E|x_i(u)|^2 dm_i(u) < \infty$ である. このとき定理3.2を一般化した次の定理が成り立つ.

<u>定理3.3</u>（[10,11]）feedback のある GC (1) で雑音 Z は (9) で与えられ，input X は (12) の形をしているとき

$$I_t(\theta, Y) = \frac{1}{2}\sum_{i=1}^{N}\int_0^t E|x_i(u) - \hat{x}_i(u)|^2 dm_i(u) \tag{13}$$

である．ここで $\hat{x}_i(u) = E[x_i(u) \mid \mathfrak{F}_u(Y)]$.

<u>注意.</u> Hitsuda [10] は最も一般の場合の Z に対し情報量を計算している．ただし feedback はなく X は Gauss 過程のとき．

以上の議論との比較のため，情報量の関数解析的な計算方法について述べる．詳しいことは [1, 2, 3, 23] を参照してほしい．

feedback はないものとする．input X も Gaussian で，X および雑音 Z の sample space はともに Hilbert 空間 H とする．即ち，X, Z および output Y は皆 H の値をとる確率変数である．よって分布 μ_X, μ_Z は H 上の Gauss 測度である．X と Z が独立なことと，Y = X + Z に注意すると，X と Y の同時分布 μ_{XY}, Y の分布 μ_Y は

$$\mu_{XY}(A) = \mu_X \otimes \mu_Z(\{(x, z);\ (x, x+z) \in A\})$$

$$\mu_Y(B) = \mu_{XY}(H \times B)$$

で与えられる．平均を 0 とすると，H 上の Gauss 測度 μ は covariance operator R により特徴付けられる．R は

$$\langle Ru, v\rangle = \int_H \langle x, u\rangle\langle x, v\rangle\, d\mu(x)$$

により与えられる．ここで $\langle\cdot,\cdot\rangle$ は H の内積である．Z の，即ち μ_Z に対する covariance operator を R_Z と記す．

今の場合にも定理 3.1 は成り立つ．Z の RKHS は

$$\mathcal{H}(Z) = \mathrm{range}\,(R_Z^{1/2})$$

であり，内積 $\langle\cdot,\cdot\rangle_{\mathcal{H}(Z)}$ は $u = R_Z^{1/2}x$, $v = R_Z^{1/2}y$ $(x, y \in H)$ とするとき

$$\langle u, v\rangle_{\mathcal{H}(Z)} = \langle x, y\rangle \tag{14}$$

であることがわかる．そして次のことが知られている．

<u>Lemma 3.4</u>　$I(X,Y) < \infty$ であるための必要かつ十分条件は $\mu_X(\mathrm{range}\,(R_Z^{1/2})) = 1$ が成り立つことであり，さらにこのためには次のことが必要かつ十分である．

$$R_X = R_Z^{1/2}\,T\,R_Z^{1/2} \quad かつ \quad \mathrm{range}\,(T) \subset \mathrm{range}\,(R_Z^{1/2}) \tag{15}$$

なる自己共役，非負，trace class 作用素 T が存在する．

そして相互情報量は次によって計算される．

<u>定理 3.5</u>　feedback のない GC で，μ_X も Gauss 測度で (15) が成り立っているとする．T の固有値を $\{\tau_n\}$ とすると

$$I(X,Y) = \frac{1}{2}\sum_{n=1}^{\infty}\log\,(1+\tau_n) \tag{16}$$

である．

<u>注意.</u>　定理 3.2, 3.3 と較べ，Gauss 型雑音 Z の形によらず統一的に取り扱える点が利点である．その一方，この定理は feedback のある場合には適用できない．

§4. Gaussian channel の capacity

再び feedback のある GC (1) を考える. Gauss 型雑音 Z は (9) のように標準表現されているとする. また input signal X については, 確率 1 で $X(\cdot,\omega) \in \mathcal{H}(Z)$ であるとする.

RKHS $\mathcal{H}(Z)$ のノルムを $\|\cdot\|_{\mathcal{H}(Z)}$ とし, input X に対し次のような制限が課せられているものとする.

$$(A_0) \qquad E[\|X\|^2_{\mathcal{H}(Z)}] \leqq P_0 \tag{17}$$

ここで P_0 は定数.

<u>定理 4.1</u> ([11]) 制限 A_0 の下での GC の capacity は feedback の有無で変らず

$$C_T^0(A_0) = C_T^f(A_0) = \frac{1}{2} P_0 \tag{18}$$

である. feedback のある場合には capacity を attain する, i.e.

$$I_T(\theta^*, Y^*) = \frac{1}{2} P_0$$

となる message θ^* と input X^* が存在する (Y^* は対応する output). feedback のない場合には capacity は attain されない.

制限 A_0 についてもう少し詳しく述べる.

雑音 Z が (9) で与えられるとき $G(\cdot) \equiv \sum_{i=1}^{N} \int_0^T F_i(\cdot,u) g_i(u) dm_i(u) \in \mathcal{H}(Z)$ に対し $\|G\|^2_{\mathcal{H}(Z)} = \sum_{i=1}^{N} \int_0^T |g_i(u)|^2 dm_i(u)$ だから, input X を (12) のように表現すれば制限 A_0 は

$$\sum_{i=1}^{N} \int_0^T E|x_i(u)|^2 dm_i(u) \leqq P_0 \tag{19}$$

となる. 特に WGC (7) に対しては, この制限は

$$\int_0^T E|x(u)|^2 du \leq P_0 \qquad (20)$$

となる．制限(20)は平均電力制限と呼ばれているもので，工学的にも自然な条件である．WGCに対しては(18)は昔から知られている(feedbackのある場合は[14])．しかし一般には制限A_0の工学的意味は不明で，現在のところ制限A_0は数学的都合で出てきた条件と言わざるを得ない．

feedbackのない場合，定理4.1は前節後半のようにGCをHilbert空間上の測度を使って表わした場合にも成り立つ．このときの制限A_0の形を調べる．雑音Zのcovariance operator R_Zの0でない固有値をλ_n，対応する固有関数をe_nとする($\langle e_m, e_n\rangle = \delta_{mn}$とする)．これを使うと(14)よりRKHS $\mathcal{H}(Z)$のノルムが

$$\| x \|^2_{\mathcal{H}(Z)} = \sum_n \lambda_n^{-1} |\langle x, e_n\rangle|^2$$

であることが容易に証明できる．よって制限A_0は

$$\sum_n \lambda_n^{-1} \int_H |\langle x, e_n\rangle|^2 d\mu_X(x) \leq P_0 \qquad (21)$$

と書ける．制限(22)の下でのfeedbackのないときのcapacityが(18)で与えられることはBaker[1]が示した．

この他，制限A_0に加え測度μ_Xののっている空間，つまりsupportの次元が制限されている場合もcapacityは具体的に求まる(詳しいことは省略．例えば[1]参照)．

このように，Gauss型雑音Zの形に応じて(19)，(20)，(21)と

色々な形に表わすことも出来るが，要するに制限A_0の下ではGCのcapacityはきれいに求まるのである．しかしこの他の形の制限の下ではcapacityを求めることは難かしくなり，これほどまとまった結果はない．

なお，feedbackのある場合のcapacityをattainする符号化については[11〜13]を参照してほしい．

§5．定常Gaussian channelにおける相互情報量

GC (1) においてinput $X=\{X(t)\}$，雑音$Z=\{Z(t)\}$がともに定常Gauss過程でfeedbackがない，即ち$X \perp\!\!\!\perp Z$のとき，GC (1) を<u>stationary Gaussian channel (SGC)</u>という．ただし通信時間は$(-\infty, +\infty)$とする．

定常過程X，Zは純非決定的とする．このときX，Zはスペクトル密度関数$f(\lambda)$，$g(\lambda)$をもつ：

$$E[X(t+s)X(s)] = \int_{-\infty}^{\infty} e^{i\lambda t} f(\lambda)\, d\lambda$$

（平均は0と仮定）．しばしば使うので記号

$$L(f, g) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \log\left(1+\frac{f(\lambda)}{g(\lambda)}\right) d\lambda$$

を導入しておく．

SGCについてはShannon以来よく研究されており，特に情報量レート＝単位時間当りの相互情報量，については

$$\lim_{T\to\infty} T^{-1} I(X_0^T, Y_0^T) = L(f, g) \tag{22}$$

という式が有名である．$f(\lambda)$, $g(\lambda)$ がある条件をみたす場合には (22) はきちんと証明されているが，一般には (22) は数学的に厳密な証明が与えられているとは言い難い．

本節の目的は定常 Gauss 過程 Z の標準表現を使うことにより，次の定理のように (22) と類似な公式を示すことである．

<u>定理 5.1</u>　SGC (1) において，X, Z は純非決定的（この仮定は本質的ではない）でそのスペクトル密度関数 $f(\lambda)$, $g(\lambda)$ は $\int_{-\infty}^{\infty} f(\lambda)/g(\lambda)\, d\lambda < \infty$ をみたしているとする．このとき

$$\lim_{T\to\infty} T^{-1} I(X_0^T, Y_0^T \mid X^0, Y^0) = L(f, g) \tag{23}$$

が成り立つ．ただし $X_0^T = \{X(t);\ 0 \le t \le T\}$, $X^0 = \{X(t);\ t \le 0\}$ で，$I(\cdot, \cdot \mid \cdot)$ は条件付相互情報量を表わす．

<u>注意</u>　Pinsker [18] は我々とは全く違う方法で，$T^{-1} I(X_0^T, Y_{-\infty}^{\infty} \mid X^0)$ の極限が $L(f, g)$ に等しいことを示し，特別の場合には (22) が成り立つことを証明している．

いくつかの補題を使って定理を証明できる．補題の証明は Appendix で与える（省略するのもある）．

<u>定理 5.1 の証明</u>　定常性より Z の標準表現は

$$Z(t) = \int_{-\infty}^{t} F(t-u)\, dB(u), \quad -\infty < t < \infty \tag{24}$$

で与えられる（[8] 参照）．ここで $B = \{B(t)\}$ は Brown 運

動，$F(t-u)$ は標準核である．次のことが成り立つ．

Lemma 5.2. input X は，スペクトル密度関数

$$k(\lambda) = f(\lambda)/2\pi g(\lambda)$$

をもつ定常 Gauss 過程 $x = \{x(t)\}$ を使い

$$X(t) = \int_{-\infty}^{t} F(t-u)x(u)\,du, \quad -\infty < t < \infty \tag{25}$$

と表現できる．$x \perp\!\!\!\perp B$ としてよい．

SGC (1) は (24), (25) より

$$Y(t) = \int_{-\infty}^{t} F(t-u)\,x(u)\,du + \int_{-\infty}^{t} F(t-u)\,dB(u) \tag{26}$$

と書ける．ところで

$$y(t) = \int_{0}^{t} x(u)\,du + B(t), \quad -\infty < t < \infty \tag{27}$$

はひとつの WGC を与えている．Brown 運動の微分，white Gaussian noise は定常（超）過程だから，(27) は SGC（の積分形）と考えてもよい．$\tilde{x} = \{\tilde{x}(t)\}$ を

$$\tilde{x}(t) = x(t) - E[x(t) \mid \mathfrak{F}_0(x)]$$

（$\mathfrak{F}_0(x)$ は $x(u)$, $u \leq 0$, を可測にする最小の σ-field）と定め，(27) を少し変形した WGC

$$\tilde{y}(t) = \int_{0}^{t} \tilde{x}(u)\,du + B(t), \quad t \geq 0 \tag{28}$$

も考える．3つの GC (26) ~ (28) における情報量を比較する．

Lemma 5.3 任意の $T \geq 0$ に対し次式が成り立つ．

$$I(X_0^T, Y_0^T \mid X^0, Y^0) = I(x_0^T, y_0^T \mid x^0, y^0)$$
$$= I(\tilde{x}_0^T, \tilde{y}_0^T). \tag{29}$$

stationary な WGC (27) に対しては次のことが成り立つ.

<u>Lemma 5.4</u>　WGC (27) に対し

$$\lim_{T\to\infty} T^{-1} I(x_0^T, y_0^T) = L(f, g) \tag{30}$$

が成り立つ.

White Gaussian noise $\dot{B}(t)$ は一様なスペクトル密度関数 $g_0(\lambda) \equiv 1/2\pi$ をもつから, $L(f,g) = L(f_1, g_0)$ に注意すると, (30) は定常な WGC (27) に対しては最初に述べた式 (22) が成り立つことを意味している.

<u>Lemma 5.5</u>　WGC (27) と (28) に対し

$$\lim_{T\to\infty} T^{-1} I(x_0^T, y_0^T) = \lim_{T\to\infty} T^{-1} I(\tilde{x}_0^T, \tilde{y}_0^T) \tag{31}$$

が成り立つ.

以上の補題を使えば, (29), (30), (31) より求める式 (23) が得られる. □

我々とは異なる方法により Solev [21] は (22) を証明している. 彼はさらに [22] では (22) における $T\to\infty$ とするときの収束の速さまで評価している (筆者はこの証明を完全に follow していないが).

式 (22) は古くは例えば Fano [6] に "証明" が与えられているが, これは数学的には不完全である.

§6. Brown 運動と equivalent な雑音をもつ Gaussian channel

再び feedback のある GC (1) を考える．本節では Gauss 型雑音 $Z = \{Z(t)\}$ が Brown 運動と equivalent，即ち Z の分布 μ_Z が Wiener 測度と互に絶対連続とする．この場合 Z の標準表現が具体的に知られているので，それを使うことにより相互情報量と capacity について若干の結果が得られる．

標準表現については次のことが知られている．

<u>Lemma 6.1</u> ([9, 8]) Z が Brown 運動と equivalent なための必要十分条件は，Brown 運動 $B = \{B(t)\}$ と Volterra 核 $f(t,s) \in L^2([0,T]^2)$ (Volterra 核 $\Leftrightarrow$ $f(t,s) = 0$ if $t < s$) があって Z が

$$Z(t) = B(t) + \int_0^t \int_0^s f(s,u)\,dB(u)\,ds, \quad 0 \leq t \leq T, \tag{32}$$

と標準表現されることである．

標準表現 (32) は

$$F(t,u) = \begin{cases} 1 + \int_0^t f(s,u)\,ds, & u \leq t \\ 0, & u > t \end{cases}$$

と定めると

$$Z(t) = \int_0^t F(t,u)\,dB(u) \tag{33}$$

となり，これは (9) の特別の場合であることがわかる．従って情報量は定理 3.3 を使って計算できる．定理 3.3 では input

Xに対し(12)を，即ち$X(\cdot,\omega)\in\mathcal{H}(Z)$を仮定したが，今の場合RKHS $\mathcal{H}(Z)$と$\mathcal{H}(B)$は集合として等しいことがわかるから(5)に注意し，Xに対してはWGCのときと同様工学的にもより自然な形の(6)を仮定する．よって我々のGCは

$$Y(t)=\int_0^t \tilde{x}(u)\,du+Z(t),\quad 0\le t\le T \tag{34}$$

となる．Volterra核$f(t,s)$に対しては解核と呼ばれる

$$\begin{cases} f(t,s)+g(t,s)+\int_s^t f(t,u)g(u,s)\,du=0 \\ f(t,s)+g(t,s)+\int_s^t g(t,u)f(u,s)\,du=0 \end{cases} \tag{35}$$

をみたすVolterra核$g(t,s)\in L^2([0,T]^2)$が一意に存在する．核$f(t,s)$，$g(t,s)$に対する$L^2[0,T]$上の積分作用素を各々f，gとすると，(35)は

$$(I+f)(I+g)=(I+g)(I+f)=I\equiv\text{恒等作用素} \tag{36}$$

を意味する．(34)の$\tilde{x}(t)$に対し，$x(t)\equiv x(t,\omega)$を

$$x(t,\omega)=(I+g)\tilde{x}(t,\omega)=\tilde{x}(t,\omega)+\int_0^t g(t,s)\tilde{x}(s,\omega)\,ds \tag{37}$$

と定めると，(36)より逆に

$$\tilde{x}(t,\omega)=(I+f)x(t,\omega)=x(t,\omega)+\int_0^t f(t,s)x(s,\omega)\,ds \tag{38}$$

である．よってinput Xは

$$X(t,\omega)=\int_0^t \tilde{x}(u,\omega)\,du=\int_0^t F(t,u)x(u,\omega)\,du \tag{39}$$

となる．才3項は(12)の右辺に対応する．よって定理3.3の系として次の結果を得る．

<u>定理 6.2</u>　Brown運動とequivalentな雑音をもつGC (34)

において $\int_0^T E|\tilde{x}(t)|^2 dt < \infty$ とすると，相互情報量は

$$I_t(\theta, Y) = \frac{1}{2}\int_0^t E|x(u) - \hat{x}(u)|^2 du \tag{40}$$

で与えられる．ここで $x(u)$ は (37) で与えられ $\hat{x}(u) = E[x(u) | \mathfrak{F}_u(Y)]$ である．

次に話を capacity へ移そう．もち論 input signal に対する制限を定理 4.1 の A_0 とすれば，capacity は feedback の有無によらず $\frac{1}{2}P_0$ である．今の場合制限 A_0 は input X を (39) で表わすとき，(19) より

$$E[\|X\|^2_{\mathcal{H}(Z)}] = \int_0^T E|x(t)|^2 dt \leq P_0 \tag{41}$$

と書けることに注意しておく．

ここでは A_0 ではなく，input signal X に対し工学的意味のはっきりしているところの WGC の場合と同じ制限

(A_1)　X は (39) のように書け

$$\int_0^T E|\tilde{x}(t)|^2 dt \leq P_0, \tag{42}$$

が課されたときの capacity について考える．条件 A_1 は §4 でみたように

$$E[\|X\|^2_{\mathcal{H}(B)}] \leq P_0 \tag{43}$$

と同じことである．制限 A_1 の下での capacity は A_0 の場合のようには簡単ではなく，部分的な結果しか得られていない．それを述べるために次のような積分作用素を考える．式 (32) の Volterra 核 $f(t,s)$ に対し，積分核 $f^*(t,s) = f(s,t)$ をもつ積

分作用素を f^{*} と記す．積分作用素 F を

$$F = f + f^{*} + f f^{*} \tag{44}$$

で定義する．F は自己共役であり，$I + F = (I+f)(I+f^{*})$ が成り立つ．$(I+F)^{-1}$ が存在し，$\|(I+F)^{-1}\| = \|(I+f)^{-1}\|^{2}$ が容易に示せる．さらに，もし $F \geqq 0$ ならば

$$\|(I+F)^{-1}\| = \|(I+f)^{-1}\| = 1 \tag{45}$$

が成り立つ．実際，F の固有値を $\tau_n \geqq 0$ とすると $(I+F)^{-1}$ の固有値は $(1+\tau_n)^{-1}$ で，$n \to \infty$ のとき $\tau_n \to 0$ だから $1 \geqq (1+\tau_n)^{-1} \to 1$ である．これより (45) が成り立つ．

<u>定理 6.3</u>．(32) の雑音をもつ GC (34) の制限 A_1 の下での capacity $C_T^f(A_1)$, $C_T^o(A_1)$ について以下のことが成り立つ．

(i) $$C_T^f(A_1) \geqq C_T^o(A_1) \geqq \frac{1}{2} P_0, \tag{46}$$

(ii) $$C_T^o(A_1) \leqq C_T^f(A_1) \leqq \frac{1}{2} \|(I+f)^{-1}\|^{2} P_0, \tag{47}$$

(iii) $F \geqq 0$ のときは

$$C_T^f(A_1) = C_T^o(A_1) = \frac{1}{2} P_0. \tag{48}$$

<u>注意</u>．(ii) の上界については柳氏より御教示を受けた．

<u>証明</u>　(i). $C_T^f(A_1) \geqq C_T^o(A_1)$ は自明だから，feedback のない場合に相互情報量が $\frac{1}{2}P_0$ に近い（かより大きい）符号化を実際に与えればよい．N を十分大きな自然数とし区間 $[0,T]$ を N 等分し，n 番目の小区間を Δ_n とする．θ_n, $n=1,\cdots,N$, を互に独立で各々 $N(0, \frac{P_0}{T})$ に従う確率変数とする．message

$\theta = \{\theta(t)\}$ と input $X = \{X(t)\}$, $X(t) = \int_0^t \tilde{x}(u)\,du$ を

$$\tilde{x}(t) \equiv \theta(t) = \theta_n \quad \text{if} \quad t \in \Delta_n$$

と定める．このときの output を Y とすると，詳しい計算は省略するが $N \to \infty$ のとき

$$I_T(\theta, Y) = I_T(X, Y) \geq \frac{1}{2} P_0 + o(1)$$

が証明できる．上の X は (42) をみたしているから (i) が成り立つ．

(ii)　X は制限 A_1 をみたすとする．X を (39) のように表現し，$L^2[0,T]$ のノルムを $\|\cdot\|_2$ と記すと，(37), (36) より

$$E[\|X\|^2_{\mathcal{H}(Z)}] = E\int_0^T |x(t)|^2 dt = E\|x(\cdot,\omega)\|_2^2$$
$$= E\|((I+g)\tilde{x})(\cdot,\omega)\|_2^2 \leq \|I+g\|^2 E\|\tilde{x}(\cdot,\omega)\|_2^2$$
$$= \|(I+f)^{-1}\|^2 E[\|X\|^2_{\mathcal{H}(B)}] \leq \|(I+f)^{-1}\|^2 P_0$$

である．従って定理4.1より (47) が成り立つ．

(iii)　$F \geqq 0$ の場合は，(45) より，(46) と (47) から (48) を得る．　□

雑音 Z の標準表現 (32) の Volterra 核 $f(t,s)$ の言葉で capacity を具体的に求めることが目標であるが，未解決である．定理 4.1，上の定理の (iii) と feedback の有無により capacity が変らない場合だけが出てきたが，(iii) の条件がなければ一般には feedback を使うことにより capacity が増える，i.e., $C_T^f(A_1) \neq C_T^0(A_1)$ が予想される．しかしまだ証明されていない．

Appendix. Lemma 5.2 ～ 5.5 の証明.

Lemma 5.2 の証明. 表現(24)の標準核 $F(t-u)$ に対し

$$2\pi g(\lambda) = \left| \int_{-\infty}^{\infty} F(t) e^{-it\lambda} dt \right|^2$$

が成り立つ. これより $\tilde{X}(t) \equiv \int_{-\infty}^{t} F(t-u)\, x(u)\, du$ はスペクトル密度 $f(\lambda)$ をもつ定常過程であることが示せる. よって $X(t)$ と $\tilde{X}(t)$ を同一視してよい. $\|$

Lemma 5.3 の証明. 条件付相互情報量の性質より, $I(X_0^T, Y_0^T \mid X^0, Y^0) = I(X^T, Y^T \mid X^0, Y^0)$, $I(x_0^T, y_0^T \mid x^0, y^0) = I(x^T, y^T \mid x^0, y^0)$. 標準表現の性質より $\mathfrak{F}_t(X) = \mathfrak{F}_t(x)$, $\mathfrak{F}_t(Y) = \mathfrak{F}_t(y)$ だから, $I(X^T, Y^T \mid X^0, Y^0) = I(x^T, y^T \mid x^0, y^0)$. 以上より (29) の第1の等式が成立. 次に $B = \{B(t)\}$ は独立増分をもつことと, $\tilde{x}(t)$, $t \geq 0$, は x^0 と独立なことから, WGC (28) と x^0 と B^0 が与えられたときの WGC (27) は同じ channel となり

$$I(\tilde{x}_0^T, \tilde{y}_0^T) = I(x_0^T, y_0^T \mid x^0, B^0) = I(x_0^T, y_0^T \mid x^0, y^0)$$

が成り立つ. $\|$

Lemma 5.4 の証明. この場合は Fano [6, § 5.10] に述べられている方針に従い, 厳密に証明できる. 本質的には定理3.5において R_Z が恒等作用素, T が $\{x(t)\}$ の共分散関数 $R_x(t-u) = E[x(t)x(u)]$ を核にもつ $L^2[0,T]$ 上の積分作用素 R_x の場合に対応する. 従って R_x の固有値 $\tau_n \equiv \tau_n(T)$ の T

$\to \infty$ のときの挙動を調べることにより (30) が示せるのである．Solev [22] の考えに沿っても証明できる． $\square$

<u>Lemma 5.5 の証明</u>. WGC (27) と (28) に対し定理 3.2 を使えば，結局 (31) を示すためには

$$\lim_{T\to\infty}\{E|x(T)-\hat{x}(T)|^2 - E|\tilde{x}(T)-\hat{\tilde{x}}(T)|^2\} = 0,$$

を示せばよい（$\hat{x}(T) = E[x(T)\mid \mathfrak{F}_{0,T}(y)]$, $\hat{\tilde{x}}(T) = E[\tilde{x}(T)\mid \mathfrak{F}_{0,T}(\tilde{y})]$, $\mathfrak{F}_{0,T}(y)$ は $y(t)$, $0 \le t \le T$, を可測にする最小の σ-field）．上式の証明は難かしくはないが，煩雑な計算が必要であり，紙数の関係上省略する． $\square$

REFERENCES

[1] Baker,C.R.: Capacity of the Gaussian channel without feedback. Inform. Control, 37 (1978), 70-89.

[2] __________ : Mutual information, strong equivalence, and signal sample path properties for Gaussian processes. Inform. Control, 41 (1979), 156-164.

[3] __________ : Calculation of the Shannon information. J. Math. Anal. Appl., 69 (1979), 115-123.

[4] Duncan,T.E.: On the calculation of mutual information. SIAM J. Appl. Math., 19 (1970), 215-220.

[5] ___________ : Mutual information for stochastic differential equations. Inform. Control, 19 (1971), 265-271.

[6] Fano,R.M.: Transmission of information. MIT Press and Wiley, 1961.

[7] Gel'fand,I.M., Yaglom,A.M.: Calculation of the amount of information about a random function contained in another such function. A.M.S. Transl. Ser. 2, 12 (1959), 199-246.

[8] 飛田, 櫃田 : ガウス過程 紀伊国屋, 1976.

[9] Hitsuda,M.: Representation of Gaussian processes equivalent to Wiener process. Osaka Math. J., 5 (1968), 299-312.

[10] __________ : Mutual information in Gaussian channels. J. Multivariate Anal., 4 (1974), 66-73.

[11] Hitsuda,M., Ihara,S.: Gaussian channels and the optimal coding. J. Multivariate Anal., 5 (1975), 106-118.

[12] Ihara,S.: Coding theory in white Gaussian channel with feedback. J. Multivariate Anal., 4 (1974), 74-87.

[13] ________: On the capacity of the continuous time Gaussian channel with feedback. J. Multivariate Anal., 10 (1980), 319-331.

[14] Kadota,T.T., Zakai,M., Ziv,J.: Mutual information of white Gaussian channel with and without feedback. IEEE Trans. Inform. Theory, IT-17 (1971), 368-371.

[15] Kolmogorov,A.N.: Theory of transmission of information. A.M.S. Transl. Ser. 2, 33 (1963), 291-321.

[16] 国田寛：確率過程の推定. 産業図書, 1976.

[17] Liptser,R.S., Shiryayev,A.N.: Statistics of random processes. Springer, 1977.

[18] Pinsker,M.S.: Information and information stability of random variables and processes. Holden-Day, 1964.

[19] Pitcher,T.S.: On the sample functions of processes which can be added to a Gaussian process. Ann. Math. Statist., 34 (1963), 329-333.

[20] Shannon,C.E.: A mathmatical theory of communication. Bell System Tech. J., 27 (1948), 379-423, 623-656.

[21] Solev,V.N.: Unit-time average of the quantity of information contained in one stationary Gaussian process about another. J. Soviet Math., (1975), 725-732.

[22] __________ : Information in a scheme with additive noise. J. Soviet Math., 16 (1981), 996-1004.

[23] Yanagi,K.: On some properties of Gaussian channels. J. Math. Anal. Appl., 88 (1982), 364-377.