

## フィードバックをもつガウス型通信路容量

東大・I 公文雅之 ( Masayuki Kumon )

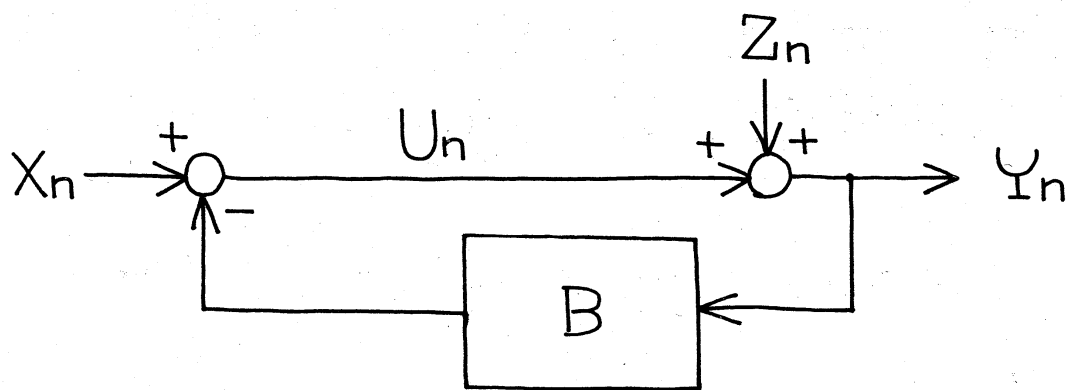


Fig.

### 1. 問題の設定

Figの通信路において、 $X_n \sim \text{input}$ ,  $Y_n \sim \text{output}$ ,  
 $Z_n \sim \text{noise}$ ,  $n = 1, \dots, N$  とし.

$$X := (X_1, \dots, X_N)^t, \quad Y := (Y_1, \dots, Y_N),$$

$$Z := (Z_1, \dots, Z_N)^t, \quad t \sim \text{transpose}$$

とおく。

noiseに関して

$Z \sim N$ :  $n$ -元 Gaussian vector, s.t.  $X \perp Z$ ,

$$E[Z] = 0, \quad R_Z := E[ZZ^t] > 0$$

を仮定する。  $E \sim$  expectation

$I(X; Y) :=$   $X$ と $Y$ の間の相互情報量

$$U = (U_1, \dots, U_N)^t \quad (\text{cf. Fig.})$$

<問題> 電力制限  $J := E[U^t U] \leq P \quad (1)$

のもとで、通信路容量  $C_f := \max_{X, B} I(X; Y)$ ,

及び、 $C_f$ を与える input  $X$ , フィードバック  $B$  を求める。

Noiseに関する仮定より、 $X$ と $B$ に対して

$X \sim$  Gaussian with  $R_X := E[XX^t]$

$B \sim N \times N$  sl (strictly lower triangular)

matrix

$$B = \begin{array}{|c|} \hline \text{斜線部以外} \\ \hline \end{array} \quad \left( \begin{array}{c} \text{斜線部} \\ \hline \end{array} \right) \text{は } 0$$

i.e. causality を満たす linear filter

と制限しても

$$C_f = \max_{R_x, B(s)} I(X; Y)$$

となることが知られている。

2. フィードバック  $B$  による  $J$  の最小化

Fig. より.

$$U = X - BY = (I - K)X - KZ \quad (2)$$

と表わされる。但し

$$K := B(I + B)^{-1} \quad (3)$$

であり、 $B$  と  $K$  は定義 (3) を通じて

$$B \sim s \iff K \sim s$$

なる関係にある。そこで  $J(1)$  を  $K$  と  $R_x$  の functional と見做すと

$$\begin{aligned} J &= J(K, R_x) \\ &= \text{Tr}[K(R_x + R_z)K^t - 2R_x K^t + R_x] \quad (4) \end{aligned}$$

と表わされる。  $\text{Tr} \sim \text{trace}$

$J$  の  $K$  に関する変分を  $\delta J := J(\delta K, R_x)$ .

但し  $\delta K \sim s$  とすると

$$\delta J = 2 \text{Tr}[\{(R_x + R_z)K^t - R_x\} \delta K]$$

となる。  $\delta J = 0$  より

$$K(R_x + R_z) - R_x = \Delta^t \quad (5)$$

for some  $\Delta \sim l$  (lower triangular)

が得られる。次に

$$R := R_x + R_z = \Omega \Omega^t \quad (6)$$

$$\Omega \sim l$$

と  $R$  を Choleski 分解する。 (5), (6) より

$$K\Omega - R_x(\Omega^t)^{-1} = \Delta^t(\Omega^t)^{-1} \quad (7)$$

さらに  $R_x(\Omega^t)^{-1}$  を

$$R_x(\Omega^t)^{-1} = \{R_x(\Omega^t)^{-1}\}_{sl} + \{R_x(\Omega^t)^{-1}\}_u \quad (8)$$

と分解する。 (8) より

$$\{ \}_{sl} := \begin{array}{|c|} \hline \text{diagonal and lower triangle shaded} \\ \hline \end{array}, \quad \{ \}_u := \begin{array}{|c|} \hline \text{upper triangle shaded} \\ \hline \end{array}$$

(7), (8) より

$$K\Omega - \{R_x(\Omega^t)^{-1}\}_{sl} = \Delta^t(\Omega^t)^{-1} + \{R_x(\Omega^t)^{-1}\}_u \quad (9)$$

(9) の左辺  $\sim sl$ , (9) の右辺  $\sim u$  であるから.

結局 (9) = 0. 故に

$$\tilde{K} := \{R_x(\Omega^t)^{-1}\}_{sl} \Omega^{-1} \quad (10)$$

とおくと

$$\tilde{J} := \min_{K(sl)} J(K, R_x) = J(\tilde{K}, R_x) \quad (11)$$

となる。  $\tilde{J}$  は  $R_x$  の functional であり.

$$\begin{aligned}\tilde{J} &= \tilde{J}(R_x) \\ &= \text{Tr} [ R_x - \{M\}_{sl} \{M\}_{sl}^t ]\end{aligned}\quad (12)$$

$$\text{但し } M := R_x (\Omega^t)^{-1} \quad (13)$$

と表わされる。

フィードバック  $K(B)$  がない場合の消費電力は、 $\text{Tr}[R_x]$  であるから、(12) はフィードバックを用いることによって  $\text{Tr}[\{M\}_{sl} \{M\}_{sl}^t] \geq 0$  だけ電力が節約できることを示している。

### 3. $\Omega$ の非対角成分による $\tilde{J}$ の最小化

(12) の  $\tilde{J}$  は  $R_x$  の functional であるが、 $R = R_x + R_z$  の functional であるとも見做せる。さらに (6) における Choleski 分解は一意的であるから、 $\Omega$  の functional とも見做せる。

ところで相互情報量  $I(X; Y)$  も  $R_x$  から  $\Omega$  の functional であるが、 $I(X; Y)$  は  $\Omega$  の非対角成分には依存しない。実際、

$$\Omega = \Omega_0 + \Omega_d,$$

$$\Omega_0 \sim sl, \quad \Omega_d \sim d \text{ (diagonal)}$$

と分解すると、

$$\begin{aligned}
I(X; Y) &= \frac{1}{2} \log |I + R_z^{-1} R_x| \\
&= \frac{1}{2} [\log |R| - \log |R_z|] \\
&= \log |\Omega_d| - \frac{1}{2} \log |R_z| \quad (14)
\end{aligned}$$

$$| | \sim \det$$

一方、 $\tilde{J}$ は  $\Omega_0$  にも依るから、電力節約の第2段階として  $\tilde{J}$  を  $\Omega_0$  に関して最小化する必要性が生じる。

まず、

$$\begin{aligned}
\text{Tr}[R_x] &= \text{Tr}[R] - \text{Tr}[R_z] \\
&= \text{Tr}[\Omega_0 \Omega_0^t] + \text{Tr}[\Omega_d^2] - \text{Tr}[R_z] \quad (15)
\end{aligned}$$

次に

$$\begin{aligned}
M &= (R - R_z)(\Omega^t)^{-1} \quad (\text{cf. (13)}) \\
&= \Omega - R_z(\Omega^t)^{-1}
\end{aligned}$$

故に

$$\begin{aligned}
\{M\}_{sl} &= \Omega_0 - \{R_z P\}_{sl} \\
&= \Omega_0 - \{L_z P_0\}_{sl} - L_z P_d \quad (16)
\end{aligned}$$

$$\text{但し, } P := (\Omega^t)^{-1} \sim u \quad (17)$$

$$= P_0 + P_d, \quad P_0 \sim su, \quad P_d = \Omega_d^{-1} \sim d$$

$$L_z := \{R_z\}_{sl} \quad (18)$$

とおいた。

(15), (16) より、 $\tilde{J}$  (12) は  $\Omega = \Omega_0 + \Omega_d$  の functional として次の様に表わされる。

$$\tilde{J} = \tilde{J}_1 + \tilde{J}_2$$

$$\tilde{J}_1 = \tilde{J}_1(\Omega_d)$$

$$= \text{Tr}[\Omega_d^2] - \text{Tr}[L_z^t L_z \Omega_d^{-2}] - \text{Tr}[R_z] \quad (19)$$

$$\tilde{J}_2 = \tilde{J}_2(\Omega)$$

$$= \text{Tr} \left[ \{L_z P_0\}_{sl} \{L_z P_0\}_{sl}^t + 2 \{L_z P_0\}_{sl} P_d L_z^t - 2 \{L_z P_0\}_{sl} \Omega_0^t - 2 L_z P_d \Omega_0^t \right] \quad (20)$$

$\tilde{J}_1$  は  $\Omega_d$  のみに依存するから,  $\tilde{J}_2$  に着目する。特に  $\{L_z P_0\}_{sl}$  の部分を見やすい形で表わすため

$$L_z = [z_{ij}], \quad z_{ij} = 0 \quad \text{for } i \leq j$$

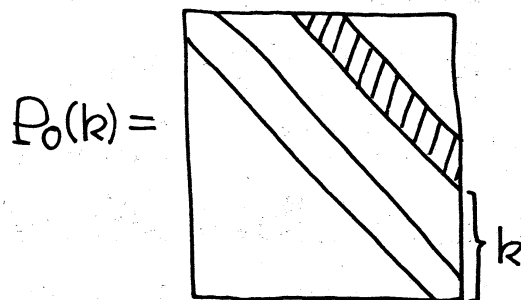
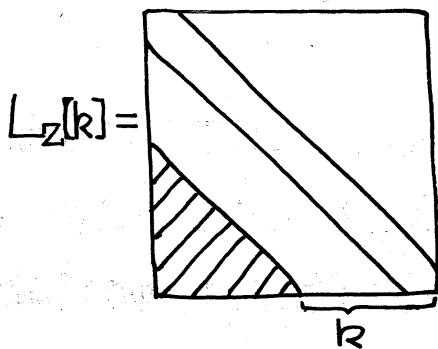
$$P_0 = [\rho_{ij}], \quad \rho_{ij} = 0 \quad \text{for } i \geq j$$

として,

$$L_z[k] := [z_{ij}], \quad z_{ij} = 0 \quad \text{for } i < j+k \quad (21)$$

$$P_0(k) := [\rho_{ij}], \quad \rho_{ij} = 0 \quad \text{for } j \neq i+k \quad (22)$$

と定義する。  $k = 1, \dots, N-1$



(21), (22) によって  $\{L_Z P_0\}_{SL}$  は

$$\{L_Z P_0\}_{SL} = \sum_{k=1}^{N-2} L_Z[k+1] P_0(k) \quad (23)$$

と表わされる。

$L_Z[k]$  は noise の相関の度合いを表わす行列で、例えば  $Z_n$  が  $\ell$  次の moving average 系列  $MA(\ell)$  ならば  $L_Z[k] = 0$ , for  $k \geq \ell + 1$  となる。

従って  $MA(1)$  のように 1 次の相関しかもたない noise の場合、(23) より  $\{L_Z P_0\}_{SL} = 0$ 、故に (20) において

$$\tilde{J}_2 = -2 \operatorname{Tr} [L_Z P_d \Omega_0^\dagger]$$

と、 $\tilde{J}_2$  は  $\Omega_0$  に関して線形な functional となる。従って  $\tilde{J}_2$  の最小値は  $\Omega_0$  の定義域の境界で達成されることになる。

Noise が 2 次以上の相関をもつ場合には、(20) より  $\tilde{J}_2$  は  $\Omega_0$  或いは  $P_0$  に関して非線形となるが、 $P_0$  に関してまづ下に凸な functional であることがわかる。さらに  $\partial \tilde{J}_2 / \partial \Omega_0 = 0$  なる点は必ず  $\Omega_0$  の定義域の外にあることが示される。

以上のことから、 $\Omega_0$  の functional としての  $\tilde{J}$  の最小値は  $\Omega_0$  の定義域の境界で達成されることになり、このことは  $\tilde{J}$  の最小値が  $R_X$  の非対角成分に関してもその定義

域の境界で達成されることを意味している。即ち

$$\tilde{J}^* := \min_{\Omega_0} \tilde{J}(\Omega) \quad (24)$$

とおくと、 $\tilde{J}^*$  は

$$R_x^* := \mathbb{y} \mathbb{y}^t \quad \text{for some } \mathbb{y} = (y_1, \dots, y_N)^t \quad (25)$$

なる  $R_x^*$  によって達成される。

この結果  $\tilde{J}^*$  は  $N$  次元ベクトル  $\mathbb{y}$  の functional となるが、その具体形は次の形で与えられる。

$$\begin{aligned} \tilde{J}^* &= \tilde{J}^*(\mathbb{y}) \\ &= y_1^2 + \frac{|R_z(1)|}{|R^*(1)|} y_2^2 + \frac{|R_z(2)|}{|R^*(2)|} y_3^2 + \dots + \frac{|R_z(N-1)|}{|R^*(N-1)|} y_N^2 \end{aligned} \quad (26)$$

但し、

$$R_x^* = [y_i y_j], \quad R_z = [z_{ij}], \quad 1 \leq i, j \leq N$$

として

$$R_x^*(k) := [y_i y_j], \quad R_z(k) := [z_{ij}], \quad 1 \leq i, j \leq k$$

$$R^*(k) := R_x^*(k) + R_z(k)$$

とおいた。

#### 4. $\tilde{J}^*$ と相互情報量の関係

(25) によって  $\tilde{J}^*$  は  $\mathbb{y} \in \mathbb{R}^N$  の functional となり、従って対応する相互情報量も  $\mathbb{y}$  の functional として表わすことができる。それは以下に定義される量から自然

に導かれる。 かつ

$X(k) := (X_1, \dots, X_k)^t$ ,  $Y(k) := (Y_1, \dots, Y_k)^t$   
 とおき、 $X(k)$  と  $Y(k)$  の間の相互情報量を  $I_k$  とおくと

$$\begin{aligned} I_k &= I[X(k); Y(k)] \\ &= \frac{1}{2} \log [ |R(k)| / |R_Z(k)| ] \end{aligned}$$

さらに

$$I_k^* := \frac{1}{2} \log [ |R^*(k)| / |R_Z(k)| ] \quad (27)$$

とおくと、 $I_k^*$  は  $y_k := (y_1, \dots, y_k)^t$  の functional であり、簡単な計算により

$$\begin{aligned} I_k^* &= I_k^*(y_k) \\ &= \frac{1}{2} \log [ y_k^t R_Z(k)^{-1} y_k + 1 ] \end{aligned} \quad (28)$$

と表わされることがわかる。

ところで (27) より

$$|R_Z(k)| / |R^*(k)| = \exp(-2I_k^*)$$

となるが、左辺は (26) の  $y_{k+1}^2$  の係数に等ならない。

よって (26) は次のようにも表現できる。

$$\tilde{J}^* = y_1^2 + e^{-2I_1^*} y_2^2 + e^{-2I_2^*} y_3^2 + \dots + e^{-2I_{N-1}^*} y_N^2 \quad (29)$$

また対応する相互情報量は (28) で  $k=N$  の場合であり。

$$I_N^* = \frac{1}{2} \log [ y^t R_Z^{-1} y + 1 ] \quad (30)$$

以上の結果 〈問題〉は、 $y \in \mathbb{R}^N$  に関する条件付変分問題

$$I_N^*(y) \rightarrow \max \quad \text{under } \tilde{J}_N^*(y) \leq P \quad (31)$$

に帰着された。但し

$$\tilde{J}_k^* := y_1^2 + e^{-2I_1^*} y_2^2 + \cdots + e^{-2I_{k-1}^*} y_k^2 \quad (32)$$

$$k = 2, \dots, N$$

とおいた。

(31)における制約式は実際には等式と見做してよく、その結果

$$y_N^2 = (P - \tilde{J}_{N-1}^*) e^{2I_{N-1}^*}$$

となり、これを  $I_N^*$  (30) に代入することにより結局、 $N-1$ 個の変数  $y_1, \dots, y_{N-1}$  の functional  $I_N^*(y_{N-1})$  の最大値を求めればよいことになる。しかしその具体的な形を与えることは、 $N=2$ の場合でも既に容易なことではない。そこで次節ではこれまでの結果にもとづき、通信路容量  $C_f = \max I_N^*$  の上限を求めることを考えることにする。

## 5. 通信路容量 $C_f$ の上限

$\tilde{J}_N^*$  (29) をフィードバックがない場合の消費電力

$$\text{Tr}[R_x] = y_1^2 + y_2^2 + \cdots + y_N^2$$

と比較してみると、各  $y_k^2$  に関して係数  $0 < e^{-2I_k^*} < 1$  分だけ電力が節約されていることがわかる。そこで次のような  $y \in \mathbb{R}^N$  に関する条件付き変分問題を考える。

$$J_0 := \sum_{k=1}^N y_k^2 \rightarrow \max \quad \text{under } \tilde{J}_N^* = P \quad (33)$$

この問題の解を  $P^* := \max J_0$  とおくと明らかに  $P^* > P$ 。また電力制限  $P$  のもとでの、フィードバックのある場合とない場合の通信路容量をそれぞれ  $C_f(P)$ ,  $C_0(P)$  とおくと

$$C_0(P) < C_f(P) \leq C_0(P^*)$$

が成り立つ。 $C_0(P)$  の具体的な形はよく知られていることであるから、 $C_f(P)$  の上限  $C_0(P^*)$  を求めることは (33) の解  $P^*$  を求めることに帰着される。

前節と同じく (33) の制約式より

$$y_N^2 = (P - \tilde{J}_{N-1}^*) e^{2I_{N-1}^*}$$

となって、これを  $J_0$  に代入することにより  $y_{N-1}$  の functional  $J_0(y_{N-1})$  の最大値を求めることになる。

詳細は省くが  $\max J_0(y_{N-1})$  は  $y_{N-2} = 0$  において達成されることかわかる。このとき

$$\tilde{J}_N^* = y_{N-1}^2 + \frac{1}{\frac{|R_2(N-2)|}{|R_2(N-1)|} y_{N-1}^2 + 1} y_N^2 = P$$

となり、よって

$$y_N^2 = (P - y_{N-1}^2) \left[ \frac{|R_z(N-2)|}{|R_z(N-1)|} y_{N-1}^2 + 1 \right]$$

を  $J_0 = y_{N-1}^2 + y_N^2$  に代入すると、 $J_0$  は

$$J_0 = -\frac{|R_z(N-2)|}{|R_z(N-1)|} \left( y_{N-1}^2 - \frac{P}{2} \right)^2 + P + \frac{|R_z(N-2)|}{|R_z(N-1)|} \frac{P^2}{4}$$

となる。故に

$$P^* = \max J_0 = P + \frac{|R_z(N-2)|}{|R_z(N-1)|} \frac{P^2}{4} \quad (34)$$

が得られる。