

통신시스템에서의 信號處理

安 淳 臣*

차 례

1. 序 論	2. 3 Modulation 理論
2. Detection, Estimation, Modulation 理論	2. 3. 1 Nonrandom Waveform 추정
2. 1 Detection 理論	2. 3. 2 Random Waveform 추정
2. 1. 1 관찰공간이 벡터인 경우	2. 4 Recursive Estimation
2. 1. 2 관찰공간이 파형인 경우	3. Data Compression
2. 2 Estimation 理論	4. 結 論
2. 2. 1 관찰공간이 벡터인 경우	參 考 文 獻
2. 2. 2 관찰공간이 파형공간인 경우	

1. 序 論

일반적인 통신 system은 송신측과 수신측 사이에 에러(error)없는 교신을 목적으로 하고 있다. 그러나 실제로는 통신경로(channel)에 雜音(noise) 또는 device 고장(例: 컴퓨터 통신에서의 소프트웨어의 파손, 電原의 고장등)등으로 error 없는 통신 system이라는 것은 거의 불가능하다.

그래서 통신 system의 목적이 error 없는 통신경로(ideal channel)의 제공이라면 실제의 channel이 이상적인 channel에 가까우면 가까울 수록 좋은 channel이다. 가능한 이러한 channel을 제공하기 위하여는 일반적으로 통신신호의 두면을 고려해야만 한다. 한 면은 정보전달의 매체인 신호의 파형에 관련된 면과 다른 하나는 신호의 언어적인 면이다. 신호파형의 전달이라는 면의 연구(例: 관성이론, 추정이론등)는 이미 오래전부터 해왔으나 [2][6][8] 신호의 언어적인 면은 computer communication의 대두와 함께 많은 연구가 행해지게 되었으며 그역사는 비교적 짧다. [3][4][5][26][26]

신호의 언어면은 computer network의 protocol과 연결이 되어있으며 이 protocol은 network의 구조와 데이터의 유동제어(flow control), network 고장 또는 error 복귀와 밀접한 관계가 있다. 신호의 언어적인 면의 설명은 참고문헌으로 미루고 [3]~[5] 여기서는 다음의 순서로 설명하려한다.

2. Detection, Estimation, Modulation 理論

2.1 Detection 理論

Detection 이론에는 parametric 방법과 nonparametric 방법이 있는데 여기서는 parametric 방법만 논하고 nonparametric 방법은 참고문헌으로 미룬다. 13 detection 이론은 보내진 신호가 discrete한 파형으로 되어 있을 때 수신측이 받은 signal를 가지고 (이것을 관찰공간이라함) 송신측 신호가 어떤 signal를 보냈는지 판단하는 것이다. 이 경우에 관찰공간이 벡터공간인 경우와 파형공간인 경우 두가지가 있다. 두 경우로 나누어 설명하도록 한다.

2.1.1 관찰공간이 벡터인 경우

보통 주어진 조건은 channel noise의 확률밀도와 송신측의 source output의 확률분포를 안다는 가정하에 위험부담(risk)를 최소로 하는 decision region

* 亞州大 工大 電子工學科 助敎授·工博

를 구한다. (bayesian test) 일반적인 수식은 송신 측 source output 이 M종류일때 (M-ary hypothesis)

$$R(P_j, Z_i) = \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} P_j C_{ij} \int_{Z_i} P(\mathbf{R}/H_j) d\mathbf{R} \quad (1)$$

이다. 여기서 C_{ij} 는 송신측의 source output 이 H_j 인데 수신측에서는 H_i 라고 판단했을 때의 비용(cost) 이고 Z_i 는 H_j 를 보냈다는 가정하에 H_i 라고 판단할 영역을 말한다. P_j 는 H_j 를 실제로 보낼 확률이다. 여기서 risk는 priori 확률과 decision region의 함수이다. $P(\mathbf{R}/H_j)$ 는 channel의 특성으로 transition 확률이다. decision region Z_i 는 discrete할 수도 있고 continuous할 수도 있으며 discrete한 경우는 적분부분이 합으로 바뀐다. 판단이론(detection theory)는 priori 확률, P_j , 이 주어져 있을때 risk를 최소로 하는 decision region을 결정하는 것이다. 만일 priori 확률, P_j , 가 주어져 있지 않는 경우는 보통 min-max receiver이라 하여 $\min_{Z_i} \max_{P_j} R(P_j, Z_i)$ 를 만족하는 decision region, Z_i 을 구한다. 만일 transition 확률분포에 random parameter가 있는 경우는 composite hypothesis라 하여 (1)식의 transition 확률분포가 $\int P(\mathbf{R}/\theta, H_j) P(\theta/H_j) d\theta$ 로 바뀌며 이 의미는 average transition 확률분포라 할 수 있다. channel 특성이 잡음외에 random성격을 갖고 있을 때가 이 경우에 속한다. (fading channel, radar problem이 이 경우에 속함) 例: M-ary hypothesis에서 $M=2$ 인 경우는 (1)식은 다음과 같이

$$P = C_{00} P_0 \int_{Z_0} P(\mathbf{R}/H_0) d\mathbf{R} + C_{10} P_0 \int_{Z_1} P(\mathbf{R}/H_0) d\mathbf{R} + C_{11} P_1 \int_{Z_1} P(\mathbf{R}/H_1) d\mathbf{R} + C_{01} P_1 \int_{Z_0} P(\mathbf{R}/H_1) d\mathbf{R} \quad (2)$$

로 된다.

$$\int_{Z_0} P(\mathbf{R}/H_0) d\mathbf{R} = \int_{Z_1} P(\mathbf{R}/H_1) d\mathbf{R} = 1 \quad (3)$$

를 이용하여 (2)식을 정리하면

$$R = P_0 C_{00} + P_1 C_{11} + \int_{Z_0} (P_1 (C_{01} - C_{11})) P(\mathbf{R}/H_1) - P_0 (C_{10} - C_{00}) P(\mathbf{R}/H_0) d\mathbf{R} \quad (4)$$

이 된다. 만일 여기서

$$\lambda(\mathbf{R}) = \frac{P(\mathbf{R}/H_1)}{P(\mathbf{R}/H_0)} < \frac{P_0 (C_{10} - C_{00})}{P_1 (C_{01} - C_{11})} = \gamma \quad (5)$$

를 만족하는 지역을 Z_0 라 하면 적분안의 괄호가 음수값을 갖게 되므로 risk R_{01} 최소값을 갖게 된다. 이때의 Z_1 은 $Z - Z_0$ 이다. (즉 받은 signal \mathbf{R} 이 Z_0 영역에 속하면 H_0 라 판단하고 Z_1 영역에 속하면 H_1 이라 판단함을 의미한다.) 식(5)의 좌변을 likelihood ratio $\lambda(\mathbf{R})$ 라고 한다.

만일 $C_{00} = C_{11} = 0$, $C_{10} = C_{01} = 1$.

$$P(\mathbf{R}/H_0) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(R-1)^2}{2\sigma^2} \right\}$$

$$P(\mathbf{R}/H_1) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(R+1)^2}{2\sigma^2} \right\}$$

이면 식(5)는

$$\lambda(\mathbf{R}) = \exp(-4R) \begin{cases} < 1 \\ > 1 \end{cases} \quad (6)$$

즉 이 식은 $R > 0$ 이면 H_0 라 판단하고 $R < 0$ 면 H_1 이라 판단한다.

2. 1. 2 관찰공간이 파형(Waveform)인 경우

M-ary hypothesis에서 일반적으로 관찰공간이 파형인 경우는 보낸 signal 이 $S_i(t)$ ($i=1, 2, \dots, M$) 일때 수신측의 신호는

$$r(t) = S_i(t) + n(t), \quad 0 \leq t \leq T: H_i \quad (7)$$

로 표시된다. 여기서 $n(t)$ 는 channel noise며 확률밀도는 주어졌다고 가정한다. 식(7)을 vector 형태로 바꾸는 방법은 orthonormal function set를 이용하여 식(7)의 양변을 전개하는 것이다. 즉

$$r(t) = \sum_{j=1}^{\infty} r_j \phi_j(t), \quad S_i(t) = \sum_{j=1}^{\infty} S_{ij} \phi_j(t), \quad n(t) = \sum_{j=1}^{\infty} n_j \phi_j(t) \quad (8)$$

를 이용하면 식(7)은

$$r_j = S_{ij} + n_j, \quad j=1, 2, 3, \dots, i=1, 2, \dots, M \quad (9)$$

가 된다. noise vector n 의 확률밀도는 $n(t)$ 의 확률적 성격으로부터 구해줄 수 있다. 이 noise vector의 확률밀도를 $P_n(n)$ 이라 하면 \mathbf{R} 의 확률밀도는 $P(\mathbf{R}/H_j) = P_n(\mathbf{R} - \mathbf{S}_j)$ 로 된다. 즉 식(1)의 형태로 되었음을 알수 있다. 즉 관찰공간이 파형인 경우도 관찰공간이 vector인 경우로 바뀌며 똑 같이 취급할 수 있음을 보인다.

일반적으로 디지털통신에서의 신크와형전달 문제는 binary hypothesis 문제로 귀착시킬 수 있으며 이 경우의 decision level를 threshold라 한다. 실제로 수신측과 송신측이 먼 경우는 중간에 repeater가 사용되며 repeater에서는 threshold level의 결정문제뿐만 아니라 synchronous통신인 경우에는 clock recovery 문제가 있다. clock recovery를 위하여는 연속의 펄스가 서로 반대의 부호를 갖는 것이 바람직한데 이

를 위하여 HDB 3 (high density bipolar coding) 등이 사용되고 있다.⁽¹⁴⁾ 또한 통신 channel 이 전화망일 때는 펄스의 periodic sequence 는 spectrum 면에서 해석할 때 peak 를 보이는데 이것은 다른 channel 에 영향을 많이 주게 된다. 이 현상을 제거하기 위하여 송신측에 scrambler 와 수신측에 descrambler 을 사용하도록 CCITT 에서 권고하고 있다. 한편 channel 특성이 dispersive 한 성질 (例: attenuation distortion, phase distortion) 을 갖고 있을 때는 수신측에 equalizer 를 부착함으로써 channel 특성을 보상한다.

例: 예로써 $M=2$, $n(t)$ 가 gaussian process 인 경우를 다루도록 하자. ($n(t)$ 가 gaussian 이 아닌 경우는 일반적인 결과를 주기 힘들다.) 이 경우에는

$$\begin{aligned} r(t) &= S_1(t) + n(t) & T_i < t < T_f H_1 \\ r(t) &= S_0(t) + n(t) & T_i < t < T_f H_0 \end{aligned} \quad (11)$$

단 여기서 $E[n(t)n(\tau)] = K_n(t, \tau)$ 이다.

$$\lambda_i \phi_i(t) = \int_{T_i}^{T_f} K_n(t, \tau) \phi_i(\tau) d\tau \quad (12)$$

를 만족하는 orthonormal set $\{\phi_i(t)\}$ $i = 1, 2, \dots$ 로 $n(t)$ 를 전개하여

$$n(t) = \sum_{i=1}^{\infty} n_i \phi_i(t) \quad (13)$$

로 하면

$$\begin{aligned} E(n_i n_j) &= \lambda_i \quad \text{for } i = j \\ &= 0 \quad \text{for } i \neq j \end{aligned} \quad (14)$$

이다. (이것을 karhunen loeve expansion 이라 한다) 이 때에 noise vector \mathbf{n} 의 확률분포는 식 (14) 로 부터

$$P_n(\mathbf{n}) = \prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{n_i^2}{2\lambda_i}} \quad (15)$$

이 된다. 식 (11) 의 양변을 $\{\phi_i(t)\}$ $i = 1, 2, \dots$ 로 전개하면

$$\begin{aligned} r_i &= S_{1i} + n_i & i = 0, 1, 2, \dots; H_1 \\ r_i &= S_{0i} + n_i & i = 0, 1, 2, \dots; H_0 \end{aligned} \quad (16)$$

로 된다. 그러므로 다음 식이 구해진다.

$$\begin{aligned} P(R/H_0) &= P_n(R - S_0) \\ &= \prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{(R_i - S_{0i})^2}{2\lambda_i}} \end{aligned}$$

$$\begin{aligned} P(R/H_1) &= P_n(R - S_1) \\ &= \prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{(R_i - S_{1i})^2}{2\lambda_i}} \end{aligned} \quad (17)$$

likelihood ration 는 2, 1, 1 例의 식 (5) 를 이용하여 구한다.

$$\begin{aligned} \ell_n \lambda(R) &= \sum_{i=1}^{\infty} \frac{1}{2\lambda_i} [2r_i (S_{1i} - S_{0i}) \\ &\quad + S_{0i}^2 - S_{1i}^2] \sum_{H_1}^{H_0} e_n \eta \end{aligned} \quad (18)$$

$$\text{윗식에 } r_i = \int_{T_i}^{T_f} r(t) \phi_i(t) dt, S_{mi} = \int_{T_i}^{T_f} S_m(t)$$

$$\phi_i(t) dt \quad m = 0, 1, n_i = \int_{T_i}^{T_f} n(t) \phi_i(t) dt \text{ 를 식}$$

(18) 에 대입하여 정리하면

$$\begin{aligned} \ell_n \wedge (R) &= \frac{1}{2} \int_{T_i}^{T_f} \int_{T_i}^{T_f} \sum_{i=1}^{\infty} \frac{\phi_i(t) \phi_i(u)}{x} [2r(t) S_1(u) \\ &\quad - 2r(t) S_0(u) + S_0(t) S_0(u) \\ &\quad - S_1(t) S_1(u)] dt du \sum_{H_1}^{H_0} \ell_n \eta \end{aligned} \quad (19)$$

이다. 한편 식 (12) 에서 양변에 $\frac{\phi_i(u)}{\lambda_i}$ 를 곱한 다음 i

에 대하여 합하면

$$\begin{aligned} \sum_{i=1}^{\infty} \phi_i(t) \phi_i(u) \\ &= \int_{T_i}^{T_f} K_n(t, \tau) \left(\sum_{i=1}^{\infty} \frac{\phi_i(\tau) \phi_i(u)}{\lambda_i} \right) d\tau \end{aligned} \quad (20)$$

이 되는데 식 (20) 의 좌변은 orthonormal function 의 성질에 의하여 $\delta(t-u)$ 가 된다. 우변의 괄호 안을 $Q_n(u, \tau)$ 라 하면

$$\delta(t-u) = \int_{T_i}^{T_f} K_n(t, \tau) Q_n(u, \tau) d\tau \quad (21)$$

이 된다. 이 경우 식 (19) 는

$$\begin{aligned} \int_{T_i}^{T_f} \int_{T_i}^{T_f} r(t) Q_n(t, u) [S_1(u) - S_0(u)] \\ dudu \sum_{H_1}^{H_0} \ell_n \eta + \frac{1}{2} \int_{T_i}^{T_f} \int_{T_i}^{T_f} [S_1(t) S_1(u) \\ - S_0(t) S_0(u)] Q_n(t, u) dudt \end{aligned} \quad (22)$$

즉 풀이는 다음과 같이 된다.

i) 식 (21) 부터 $Q_n(t, u)$ 를 구하다. ii) i) 에서 구한 $Q_n(t, u)$ 를 식 (22) 에 대입하여 H_0 를 보냈는지 H_1 를 보냈는지 판단한다.

$$\text{여기서 만일 } K_n(t, u) = \delta(t-u) \quad (23)$$

이라 하면 (이 경우가 white gaussian noise 경우이

다.) $\lambda_i = 1 \quad i = 1, 2, \dots$ 이므로 $Q_n(t, u)$ 의 정의부터 $Q_n(t, u) = \delta(t - u)$ (24)가 된다. 이때 식 (22)은

$$\int_{T_i}^{T_f} r(t) [S_i(t) - S_0(t)] dt \underset{H_1}{\overset{H_0}{\gtrless}} \ell_n \eta + \frac{1}{2} \int_{T_i}^{T_f} [S_i(t)^2 - S_0^2(t)] dt \quad (25)$$

이 된다.

2.2 Estimation 理論

일반적으로 추정이론에는 recursive 형태와 nonrecursive 형태가 있다. 여기서는 nonrecursive 형태를 다루기로 한다. 추정이론에는 parameter estimation과 waveform estimation으로 나누어진다. 후자 경우를 modulation 이론이라고 부르며 이것은 다음 절에서 논하기로 한다. parameter estimation에는 parameter가 random한 경우와 nonrandom한 경우로 갈라질 수 있으며 또한 관찰공간이 벡터공간인 경우와 파형공간인 경우로 나누어 질 수 있다. 수학적인 면에서 볼 때는 파형공간과 벡터공간을 항상 1:1 대응시킬 수 있으므로 먼저 관찰공간이 벡터공간인 경우의 이론을 설명한 다음에 관찰공간이 파형공간인 경우를 다루겠다.

2.2.1 관찰공간이 벡터공간인 경우

이 경우에 추정하고자 하는 parameter가 random한 경우와 nonrandom한 경우를 나누어 설명하겠다.

(a) random parameter estimation

이 경우에 송신측과 수신측의 상황을 modelling하면 그림 1이 된다. 확률분포, $P(\mathbf{R}/\mathbf{A})$ 가 channel의 수학적 model이며 source의 확률분포 $P(\mathbf{A})$ 가 source의 수학적 model이다. $P(\mathbf{A})$ 와 $P(\mathbf{R}/\mathbf{A})$ 가 주어지면 $P(\mathbf{A}, \mathbf{R}) = P(\mathbf{A}) P(\mathbf{R}/\mathbf{A})$ 이다.

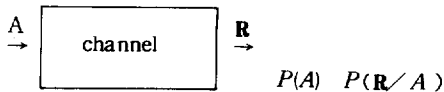


그림 1. system model

일반적으로 estimation한다는 의미는 risk를 최소로 하는 \mathbf{A} 의 추정값을 관찰된 vector \mathbf{R} 로 표시하는 것이다. 이 경우는 bayesian estimation이라 하는데 risk는

$$R \triangleq \int d\mathbf{A} \int C[\mathbf{A}, \hat{\mathbf{a}}(\mathbf{R})] P(\mathbf{A}, \mathbf{R}) d\mathbf{R} \quad (26)$$

로 표시된다. 여기서 $C(\mathbf{A}, \hat{\mathbf{a}}(\mathbf{R}))$ 는 실제의 parameter와 추정된 parameter 값의 function으로 co-

st function을 나타낸다. cost function의 형태를 임의의 형태로 줄 수 있으나 실제로 물리적인 의미와 그리고 수학적으로 다루기 용이한 형태(tractable form)를 갖고 있는 것이 중요하다. 이 두조건을 만족시키는 것으로써

$$C(\mathbf{A}, \hat{\mathbf{a}}(\mathbf{R})) = [\mathbf{A} - \hat{\mathbf{a}}(\mathbf{R})]^T [\mathbf{A} - \hat{\mathbf{a}}(\mathbf{R})] \quad (27)$$

이 있으며 이 quadratic 형태의 criterion은 거의 절대적으로 많이 쓰이고 있다.

이 경우 식 (26)

$$R_{ms} = \int \left\{ \int [\mathbf{A} - \hat{\mathbf{a}}(\mathbf{R})]^T [\mathbf{A} - \hat{\mathbf{a}}(\mathbf{R})] P(\mathbf{A}/\mathbf{R}) d\mathbf{A} \right\} d\mathbf{R} \quad (28)$$

이 된다. 이 식에서 R_{ms} 를 최소로 하는 $\hat{\mathbf{a}}(\mathbf{R})$ 은 다음식이 된다.

$$\hat{\mathbf{a}}_{ms}(\mathbf{R}) = \int \mathbf{A} P(\mathbf{A}/\mathbf{R}) d\mathbf{A} \quad (29)$$

즉 이식은 conditional mean를 의미한다.

한편 random parameter vector \mathbf{A} 에 대한 추정값으로 다른 방법을 취하는 것이 있는데 이것은 $P(\mathbf{A}/\mathbf{R})$ 를 최대로 하는 \mathbf{A} 를 구하는 것이다. (이것을 maximum a posteriori estimate, 간단히 MAP estimate라고 한다.) 즉

$$\frac{\partial \ln P(\mathbf{A}/\mathbf{R})}{\partial \mathbf{A}} \Big|_{\mathbf{A} = \hat{\mathbf{a}}(\mathbf{R})} \text{map} = 0 \quad (30)$$

를 만족하는 \mathbf{A} 를 구하는 것이다. 단, 여기서

$$\frac{\partial}{\partial \mathbf{A}} \text{는} \begin{bmatrix} \frac{\partial}{\partial A_1} \\ \frac{\partial}{\partial A_2} \\ \vdots \end{bmatrix} \text{를 의미한다. 일반적으로 볼때}$$

식 (30)의 해는 식 (26)의 cost 함수 $C(\mathbf{A}, \hat{\mathbf{a}}(\mathbf{R}))$ 이 $\mathbf{A} - \hat{\mathbf{a}}(\mathbf{R})$ 의 함수로 원점을 제외한 uniform cost를 가질 때의 식 (26)의 risk를 최소로 하는 해와 같음을 보일 수 있다. [2] estimation error에 대한 bound가 주어지 있는데 문헌을 참고하기 바란다. [2][15]

$$\text{例: } S_1 = x_1 A_1, \quad S_2 = x_2 A_2$$

$$r_1 = S_1 + n_1, \quad r_2 = S_2 + n_2 \quad (31)$$

라고 하자. 여기서 $x_i \quad i = 1, 2$ 는 알고있는 상수이고 A_1, A_2 는 추정하고자 하는 random parameter로서 서로 독립적이고 $N(m, \sigma_a), N(0, \sigma_a), n_1, n_2$ 는 random noise로 역시 서로 독립적이고 $N(0, \sigma_r)$ 를 만족시킨다고 보자.

이 경우

$$P(\mathbf{A}/\mathbf{R}) = \frac{P(\mathbf{R}/\mathbf{A}) P(\mathbf{A})}{P(\mathbf{R})} \quad (32)$$

식 (32)를 이용하여 정리하면

$$P(\mathbf{A}/\mathbf{R}) = \mathbf{K}(\mathbf{R}) \exp \left\{ -\frac{1}{2} \left[\frac{x_1^2}{\sigma_n^2} + \frac{1}{\sigma_a^2} \right] \right. \\ \left. \left(A_1 - \frac{\frac{r_1 x_1}{\sigma_n^2} + \frac{m}{\sigma_a^2}}{\frac{x_1^2}{\sigma_n^2} + \frac{1}{\sigma_a^2}} \right)^2 - \frac{1}{2} \left[\frac{x_2^2}{\sigma_n^2} + \frac{1}{\sigma_a^2} \right] \right. \\ \left. \left(A_2 - \frac{\frac{r_2 x_2}{\sigma_n^2}}{\frac{x_2^2}{\sigma_n^2} + \frac{1}{\sigma_a^2}} \right)^2 \right\} \quad (33)$$

로 된다. 식 33 으로부터

$$\hat{a}_{1ms} = \frac{\frac{r_1 x_1}{\sigma_n^2} + \frac{m}{\sigma_a^2}}{\frac{x_1^2}{\sigma_n^2} + \frac{1}{\sigma_a^2}}, \hat{a}_{2ms} = \frac{\frac{r_2 x_2}{\sigma_n^2}}{\frac{x_2^2}{\sigma_n^2} + \frac{1}{\sigma_a^2}} \quad (34)$$

이다. 식 33 으로부터 \hat{a}_{map} 도 식 34 과 같아짐을 알 수 있다.

(b) nonrandom parameter estimation

일반적으로 nonrandom parameter 에 대하여 추정 하는 방법으로 maximum likelihood 방법이 있다.

이 방법은 $P(\mathbf{R}/\mathbf{A})$ 를 maximum 으로 하는 \mathbf{A} 값 을 \mathbf{R} 의 함수로 표시하는 것이다.

이것을 수식으로 표현하면

$$\frac{\partial \ln P(\mathbf{R}/\mathbf{A})}{\partial \mathbf{A}} = 0 \quad (35)$$

의 해를 구하는 것이다. 이렇게 해서 구한 estimate 의 성능은 estimate error 의 lower bound 하고 비교함으로써 알 수 있는데 이것에 관하여는 참고문헌 으로 미룬다.

例: 주어진 조건이 A_1, A_2 가 nonrandom 이라는 것 을 제외하고는 (a) 의 例하고 같을 때

$$P(\mathbf{R}/\mathbf{A}) = \frac{1}{2\pi\sigma_n^2} e^{-\frac{1}{2\sigma_n^2} [(R_1 - x_1 A_1)^2 + (R_2 - x_2 A_2)^2]} \quad (36)$$

식 35 를 이용하면

$$\hat{a}_{1ml} = \frac{r_1}{x_1}, \hat{a}_{2ml} = \frac{r_2}{x_2} \quad (37)$$

이 된다. (식 34 와 비교하기 바람)

2.2.2 관찰공간이 파형공간인 경우

이 경우는 앞의 detection 경우처럼 파형공간을 vector 공간으로 mapping 시켜서 2.2.1 의 개념을 그대로 적용하여 푼다. 여기서는 개념은 어렵지 않으나 표현상의 문제와 수식의 복잡성 때문에 single para-

meter estimation 를 다루기로 한다.

(a) random parameter estimation

일반적으로 이경우에 modelling된 수식은

$$r(t) = S(t, A) + n(t) \quad T_i \leq t \leq T_f \quad (38)$$

이다. 여기서 $n(t)$ 는 channel 잡음이고 A 는 추정하고자 하는 파라미터이다. 양변을 구간 $T_i \leq t \leq T_f$ 에서 orthonormal function set $\{\phi_i(t)\} \quad i = 1, 2, \dots$ 를 이용하여 전개하면

$$r(t) = \sum_{i=1}^{\infty} r_i \phi_i(t)$$

$$S(t, A) = \sum_{i=1}^{\infty} S_i(A) \phi_i(t) \quad (39)$$

$$n(t) = \sum_{i=1}^{\infty} n_i \phi_i(t)$$

로 전개된다. 식 38 과 식 39 을 이용하 정리하면

$$r_i = S_i(A) + n_i \quad i = 1, 2, \dots \quad (40)$$

로 된다.

$n(t)$ 의 확률분포가 구해지면 잡음 벡터 n 의 확률 분포가 주어지며 A 의 확률분포와 n 의 확률분포가 주어지면 관찰벡터 r 과 파라미터 A 의 joint probability density 를 알 수 있음을 식 40 이 보여준다.

(실제제단의 어려운 정도는 잡음 $n(t)$ 와 A 의 확률 분포에 따라 다르다. 일반적으로 $n(t)$ 가 gaussian 이 아닌 경우는 다루기가 힘들다.)^{[10][16][17]}

noise vector n 의 확률분포를 $P_n(n)$ 이라하면 r 와 A 의 joint probability density 는

$$P(\mathbf{R}, A) = P(A) P_n(\mathbf{R} - \mathbf{S}(A)) \quad (41)$$

이다. 여기서 A 와 n 는 서로 독립적이라고 가정했다. 식 41 를 변수 A 에 대하여 적분함으로써 $P(\mathbf{R})$ 를 구할 수 있으므로 $P(A|\mathbf{R})$ 은

$$P(A|\mathbf{R}) = \frac{P(\mathbf{R}, A)}{P(\mathbf{R})} \quad (42)$$

과 같다. 식 42 부터 a_{ms} 와 a_{map} 를 구할 수 있다.

例: 수신측 파형이

$$r(t) = S(t, A) + n(t) \quad (43)$$

이다. 여기서 $n(t)$ 는 gaussian process 이고 A 는 추정하고자 하는 parameter 로 확률밀도가 $N(0, \sigma_a^2)$ 이다. a_{map} 를 구하기로 하자.

이 문제를 풀기위하여 vector 로 mapping 시키는 orthogonal fcn set $\{\phi_i(t)\} \quad i = 1, 2, \dots$ 을 앞에서처럼

$$\lambda_i \phi_i(t) = \int_{T_i}^{T_f} K_n(t, \tau) \phi_i(\tau) d\tau \quad (44)$$

을 만족시키는 function set 라 하자.

이 $\{\phi_i(t)\}$ $i = 1, 2, \dots$ 로 식(43)을 전개하여 정리하면

$$P(\mathbf{R} | A) = \prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi\lambda_i}} e^{-\frac{(R_i - S_i(A))^2}{2\lambda_i}}$$

$$P(\mathbf{R} | A) = \left[\prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi\lambda_i}} \right] \exp \left[-\frac{1}{2} \int_{T_i}^{T_f} \int_{T_i}^{T_f} (r(t) - S(t, A))(r(u) - S(u, A)) Q_n(t, u) dt du \right] \quad (45)$$

여기서 $Q_n(t, u)$ 는 앞의 detection 경우처럼

$$\sum_{i=1}^{\infty} \frac{\phi_i(t)\phi_i(u)}{\lambda_i} \text{로써 다음식}$$

$$\delta(t-u) = \int_{T_i}^{T_f} K_n(t, \tau) Q(\tau, u) d\tau \quad (46)$$

를 만족한다. (참고 식(21)) 한편

$$P(A | \mathbf{R}) = \frac{P(\mathbf{R} | A) P(A)}{P(\mathbf{R})} \quad (47)$$

여기서 a_{map} 를 구하기 위하여 식(47)부터

$$\frac{\partial \ln P(A | \mathbf{R})}{\partial A} = \frac{\partial \ln P(\mathbf{R} | A)}{\partial A} + \frac{\partial \ln P(A)}{\partial A}$$

$$= \int_{T_i}^{T_f} \int_{T_i}^{T_f} ((r(t) - S(t, A)) Q_n(t, u) \frac{\partial S(u, A)}{\partial A} du dt - \frac{A}{\sigma_a^2}) = 0$$

그러므로

$$\sigma_a^2 \int_{T_i}^{T_f} \int_{T_i}^{T_f} (r(t) - S(t, A)) Q_n(t, u) \frac{\partial S(u, A)}{\partial A} du dt \Big|_{A=a_{map}} = a_{map} \quad (48)$$

부터 a_{map} 를 구할 수 있다. 이 식은 a_{map} 의 비선형 방정식이다. 여기서 $S(t, A)$ 가 A 의 비선형식일 때는 식(47)부터 a_{ms} 을 구하기가 상당히 어려움을 알 수 있다. 그러나 $S(t, A)$ 가 A 의 선형식일 때는 $a_{map} = a_{ms}$ 임을 보일 수 있다.

(b) nonrandom parameter estimation

위의 random parameter 경우와 똑같은 형태로 수식은,

$$r(t) = S(t, A) + n(t) \quad T_i < t < T_f \quad (49)$$

이다. (a)의 경우와 똑같이 하여 정리하면,

$$P(\mathbf{R} | A) = P_n(\mathbf{R} - S(A)) \text{이 된다.}$$

여기서 a_{me} 을 구하기 위하여

$$\frac{\partial \ln P(\mathbf{R} | A)}{\partial A} = 0 \quad (50)$$

하면 a_{me} 를 구할 수 있다.

例 ; 수신측 입력이

$$r(t) = S(t, A) + n(t) \quad (51)$$

이다. 여기서 $n(t)$ 는 gaussian process라 하고

$E[n(t)n(\tau)] = K_n(t, \tau)$ 라 하자. 앞에서 처럼 식

(44)를 만족하는 function set 을 $\{\phi_i(t)\}_{i=1, 2, \dots}$ 라 하자. 그러면 식(45)에서

$$P(\mathbf{R} | A) = \left[\prod_{i=1}^{\infty} \frac{1}{\sqrt{2\pi\lambda_i}} \right] \exp \left[-\frac{1}{2} \int_{T_i}^{T_f} \int_{T_i}^{T_f} (r(t) - S(t, A))(r(u) - S(u, A)) Q_n(t, u) du dt \right] \quad (52)$$

이다. 그러므로

$$\frac{\partial \ln P(\mathbf{R} | A)}{\partial A} = \int_{T_i}^{T_f} \int_{T_i}^{T_f} (r(t) - S(t, A)) Q_n(t, u) \frac{\partial S(u, A)}{\partial A} du dt \Big|_{A=a_{me}} = 0 \quad (53)$$

를 풀어 $A = a_{me}$ 를 구한다. (식(53)은 $A = a_{me}$ 를 구하기 위한 필요조건이다.)

이 식 역시 앞에서 처럼 a_{me} 의 비선형방정식이다.

2.3 Modulation 理論

일반적으로 waveform estimation를 modulation 이론이라고 한다. 이론적으로 볼 때 parameter 추정 이론의 확장이다. 여기서는 수식상의 복잡성 때문에 잡음 $n(t)$ 가 gaussian process라는 가정하에 이론을 전개하기로 하자.

2.3.1 Nonrandom Waveform 추정

이 경우의 수신측 파형의 적당한 model 은

$$r(t) = S(t, a(t)) + n(t) \quad T_i < t < T_f \quad (54)$$

로 표시된다. 단 여기서 $n(t)$ 는 gaussian process이다. 식(52)처럼 log likelihood fcn를 구하면

$$\ln P(\mathbf{R} | a(t)) = \ln K(\mathbf{R}) + \int_{T_i}^{T_f} \int_{T_i}^{T_f} S(t, a(t)) Q_n(t, u) [r(u) - \frac{1}{2} S(u, a(u))] du dt \quad (55)$$

이다. 여기서 $Q_n(t, u)$ 는 식(21)을 만족한다. 일반적으로 임의의 $S(t, a(t))$ 에 대하여 $\ln P(\mathbf{R} | a(t))$ 를 최대로 하는 $a(t)$ 를 구하는 것은 매우 힘들다. 여기서 $S(t, a(t))$ 가 그리는 공간이 $r(t)$ 가 그리는 공간을 포함할 수 있다고 가정하자. 식(55) $a(t) = a_{me}(t) + a_e(t)$ 라 놓아 calculus of variation를 이용하여 식(55)가 최대되기 위한 필요조건을 구하면 모든 $a_e(t)$ 에 대하여

$$\int_{T_i}^{T_f} a_e(t) dt \left\{ \frac{\partial S(t, \hat{a}_{me}(t))}{\partial \hat{a}_{me}(t)} \int_{T_i}^{T_f} Q_n(t, u) [r(u) - S(u, \hat{a}_{me}(u))] du \right\} = 0 \quad (56)$$

를 만족해야 한다. 즉 이 경우의 해는

$$r(u) = S(u, \hat{a}_{me}(u)), \quad T_i < u < T_f \quad (57)$$

를 만족해야 한다. (위에서 $S(t, a(t))$ 는 $r(t)$ 를 포함한다고 가정했다.) 식(57)부터 $\hat{a}_{me}(u)$ 를 구할 수 있다.

例; 수신측의 입력파형이

$$r(t) = a(t) + n(t) \quad (57)$$

이라 하자. 여기서 $a(t)$ 는 추정하고자 하는 파형이고 $n(t)$ 는 gaussian process 이다.

이 경우 $a(t)$ 의 파형공간이 $r(t)$ 의 파형공간을 나타낼 수 있으므로 식(57)에서 $\hat{a}_{me}(t) = r(t)$ 임을 알 수 있다.

2.3.2 Random Waveform 추정

random waveform estimation를 위하여는 관찰공간이 파형공간이고 추정하고자 하는 parameter 가 multiple 인 경우를 고려해야 한다.

MAP 방법으로 multiple parameter 를 추정 할 경우의 수식은 잡음이 gaussian 이라는 가정하에 식(48)의 확장으로 다음과 같다.

$$\hat{a}_{i \text{ map}} = \sigma_i^2 \int_{T_i}^{T_f} \int_{T_i}^{T_f} \frac{\partial S(t, \mathbf{A})}{\partial A_i} \Big|_{\mathbf{A} = \hat{\mathbf{A}}_{\text{map}}} Q_n(t, u) [r(u) - S(u, \hat{\mathbf{A}}_{\text{map}})] du dt \quad (58)$$

단, $\sigma_i^2 \triangleq \text{var}(a_i)$, $i = 1, 2, \dots, N$

지금부터 식(58)를 이용하여 random waveform 추정이론을 전개하기로 한다. 일반적으로 수신측의 파형 모델은

$$r(t) = S(t, a(t)) + n(t) \quad (59)$$

와 같다. 여기서 $a(t)$ 는 추정하고자 하는 random 파형이고 $n(t)$ 는 channel 잡음으로 둘 다 gaussian 이다. 한편 $a(t)$ 를

$$\mu_i \psi_i(t) = \int_{T_i}^{T_f} K_a(t, u) \psi_i(u) du, \quad T_i \leq t \leq T_f \quad (60)$$

를 만족하는 orthonormal fcn.set 로 전개하면

$$E(a_i) = 0, \quad E(a_i, a_j) = \mu_i \delta_{ij} \quad (61)$$

을 만족한다. (karhunen-loeve expansion theorem, [2], 또는 mercer's theorem [19])

orthonormal set 의 첫 N 항으로 표시할 수 있는 함수들로 이루어진 fcn.subclass $\{a_N(t)\}$ 를 고려하면 이 경우는

$$a_N(t) = \sum_{i=1}^N a_i \psi_i(t) \quad T_i < t < T_f \quad (62)$$

로 표시된다. 이때에

$$S(t, a(t)) = S(t, \sum_{i=1}^N A_i \psi_i(t)) \quad (63)$$

이 되며 표현을 바꾸면

$$S(t, \mathbf{A}) = S(t, \sum_{i=1}^N A_i \psi_i(t)) \quad (64)$$

이 된다. 식(64)을 식(58)에 대입하면 다음식과 같다.

$$\hat{a}_{i \text{ map}} = a_i^2 \int_{T_i}^{T_f} \int_{T_i}^{T_f} \frac{\partial S(Z, \sum_{i=1}^N A_i \psi_i(Z))}{\partial A_i} \Big|_{\mathbf{A} = \hat{\mathbf{A}}_{\text{map}}} Q(Z, u) [r(u) - S(u, \hat{\mathbf{A}}_{\text{map}})] du dz$$

$$= \mu_i \int_{T_i}^{T_f} \int_{T_i}^{T_f} \frac{\partial S(Z, \hat{a}_N(z))}{\partial \hat{a}_N(Z)} \psi_i(Z) Q_N(Z, u) [r(u) - S(u, \hat{a}_N(u))] du dz \quad (65)$$

여기서 식(61)부터의

$$\sigma_i^2 = \mu_i \quad (66)$$

을 이용했다. 식(65)의 양변에 $\psi_i(t)$ 을 곱하여 i 는 1부터 N 까지 더한 다음에 N 을 ∞ 로 보내면

$$\hat{a}(t) = \int_{T_i}^{T_f} \int_{T_i}^{T_f} \frac{\partial S(Z, \hat{a}(E))}{\partial \hat{a}(Z)} K_a(t, Z) Q_N(E, u) [r(u) - S(u, \hat{a}(u))] du dz \quad (67)$$

을 얻는다. 여기서

$$K_a(t, Z) = \sum_{i=1}^{\infty} \mu_i \psi_i(t) \psi_i(Z) \quad (68)$$

을 이용했다. 식(67)이 $a(t)$ 의 추정파형 $a(t)_{\text{map}}$ 을 주는 식이다. $a(t)$ 의 비선형 적분방정식임을 보여 준다.

例; 위에서 $n(t)$ 가 white gaussian 이고 $S(u, a(t)) = a(t)$ 인 경우를 고려하여 보자. 이 경우는

$$Q_N(Z, u) = \delta(Z - u) \quad (69)$$

식(69)식을 이용하면 식(67)은

$$\hat{a}(t) = \int_{T_i}^{T_f} \int_{T_i}^{T_f} K_a(t, Z) \delta(Z - u) [r(u) - \hat{a}(u)] du dz = \int_{T_i}^{T_f} K_a(t, Z) [r(Z) - \hat{a}(Z)] dz \quad (70)$$

식(70)은 적분방정식이다. 식(60)를 만족하는 orthogonal fcn.set 를 이용하면 쉽게 풀린다. 풀면

$$\hat{a}(t) = \sum_{i=1}^{\infty} \frac{\mu_i r_i}{1 + \mu_i} \psi_i(t) \quad (71)$$

이다. 여기서 $\mu_i \gg 1$ 이라면

$$\hat{a}(t) \approx \sum_{i=1}^{\infty} r_i \psi_i(t) = r(t) \quad (72)$$

이 되는데 식(57)의 해 $\hat{a}_{me}(t)$ 와 같음을 알 수 있다.

또한 μ_i 가 모든 i 에 대하여 같은 값을 갖는다면 즉 $a(t)$ 가 white process 인 경우에는 식(71)에서

$$\hat{a}(t) = \frac{\mu}{1 + \mu} r(t) \quad (73)$$

이 된다.

2.4 Recursive Estimation [1][2][7] - [11][20] - [23]

일반적으로 recursive estimation 은 data 가 시간 에 따라 얻을 수 있는 경우 각 순간마다 추정하되 추

정값을 그전 추정값에 새로 얻은 data 의 정보를 더하여 추정하는 방법을 말한다. 여기에도 선형 추정과 비선형 추정이 있겠으나 일반적인 비선형 추정에 대한 이론은 확립이 안되어 있다. 단지 비선형 추정은 선형추정 방법을 이용하여 근사적인 방법에 의하여 접근하는 방법이 잘 알려져 있다.^{[8][10][11]} 여기서 는 개념적인 설명과 결과만 주겠다. 결과식에 대한 유도는 문헌을 참고하기 바란다.

선형추정 방식은 KALMAN Filter 로 알려져 있으며 근본개념은 source (message)의 modelling 과 관찰과정의 modelling 의 특수성으로 이루어 졌다는 것이다. 일반적으로 source model 은 여러가지가 있는데 (例, AR, MA, ARMA Model, nonlinear model) KALMAN 의 경우는 source model 를

$$X(K+1) = \phi(K+1 | K) X(K) + T(K)W(K) \quad (74)$$

형태로 놓았고 (continuous time 경우도 다음 table 에 주어진 것처럼 비슷하다) observation model 은

$$Y(K+1) = H(K+1) X(K+1) + V(K+1) \quad (75)$$

로 가정했다. 여기서 식(74)는 AR model 아니다. 왜냐하면 $\phi(K+1 | K)$, $T(K)$ 는 일반적으로 상수가 아니라 K 의 함수이기 때문이다. 식(74)와 식(75)에 대하여 table 에 주어진 조건과 가정을 갖고 $X(K)$ 에 대한 recursive estimation 를 구하는 방법은 조건적 확률밀도 $P(X(K) | Y(K))$ 을 갖고 구하는 방법과 Hilbert space 의 orthogonal projection 성질을 이용하여 푸는 방법등이 있다. ^{[1][2][10]}

Table 의 조건과 가정밀에서는 어느 경우도 계산하나 같은 결과를 준다. 일반적으로 Wiener filter 는 Kalman filter 의 특수한 경우이다. 왜냐하면 Wiener filter 의 경우는 source model 과 관찰 modal 이 stationary process model 인 경우이기 때문이다.

다음 table 은 continuous kalman filter 와 discrete kalman filter 의 식을 보여준다.

Table 1. Continuous Kalman filter

source model	$\dot{\mathbf{X}}(t) = \mathbf{A}(t)\mathbf{X}(t) + \mathbf{B}(t)\mathbf{W}(t)$
observation model	$\mathbf{Z}(t) = \mathbf{C}(t)\mathbf{X}(t) + \mathbf{V}(t)$
initial condition	$E\{\mathbf{X}(0)\} = \hat{\mathbf{X}}(0), \text{Var}\{\mathbf{X}(0)\} = \mathbf{P}(0)$
other assumption	$E\{\mathbf{W}(t)\mathbf{V}(\tau)^T\} = 0,$ $E\{\mathbf{V}(t)\mathbf{V}^T(\tau)\} = \mathbf{R}(t)\delta(t-\tau)$ $E\{\mathbf{W}(t)\mathbf{W}^T(\tau)\} = \mathbf{Q}(t)\delta(t-\tau)$

$$E\{\mathbf{W}(\tau)\} = E\{\mathbf{V}(\tau)\} = 0,$$

$$\text{Cov}\{\mathbf{V}(\tau), \mathbf{X}(0)\} = \text{Cov}\{\mathbf{W}(\tau), \hat{\mathbf{X}}(0)\} = 0$$

estsmate eq. $\hat{\mathbf{X}}(t) = \mathbf{A}(t)\hat{\mathbf{X}}(t) + \mathbf{K}(t)[\mathbf{Z}(t) - \mathbf{C}(t)\hat{\mathbf{X}}(t)]$

error covariance eq. $\dot{\mathbf{P}}(t) = \mathbf{A}(t)\mathbf{P}(t) + \mathbf{P}(t)\mathbf{A}^T(t) - \mathbf{P}(t)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)\mathbf{C}(t)\mathbf{P}(t) + \mathbf{B}(t)\mathbf{Q}(t)\mathbf{B}^T(t)$

gain eq. $\mathbf{K}(t) = \mathbf{P}(t)\mathbf{C}^T(t)\mathbf{R}^{-1}(t)$

Table 2. Discrete Kalman filter

source model	$\mathbf{X}(K+1) = \phi(K+1 K)\mathbf{X}(K) + \mathbf{T}(K)\mathbf{W}(K)$
observation model	$\mathbf{Z}(K+1) = H(K+1)\mathbf{X}(K+1) + \mathbf{V}(K+1)$
initial condition	$E\{\mathbf{X}(0)\} = \hat{\mathbf{X}}(0), \text{Var}\{\mathbf{X}(0)\} = \mathbf{P}(0)$
other assumptions	$E\{\mathbf{W}(j)\mathbf{V}(K)\} = 0, E\{\mathbf{W}(K)\mathbf{W}(j)^T\} = \mathbf{Q}(K)\delta(K=j)$ $E\{\mathbf{V}(K)\mathbf{V}(j)^T\} = \mathbf{R}(K)\delta(K-j)$ $E\{\mathbf{W}(K)\} = E\{\mathbf{V}(K)\} = 0$ $\text{Cov}\{\mathbf{V}(K), \mathbf{X}(0)\} = \text{Cov}\{\mathbf{W}(K), \mathbf{X}(0)\} = 0$
estimate eq.	$\hat{\mathbf{X}}(K+1) = \phi(K+1 K)\hat{\mathbf{X}}(K) + \mathbf{K}(K+1)[\mathbf{Z}(K+1) - H(K+1)\phi(K+1 K)\hat{\mathbf{X}}(K)]$
gain matrix	$\mathbf{K}(K+1) = \mathbf{P}(K+1 K)\mathbf{H}^T(K+1)[\mathbf{H}(K+1)\mathbf{P}(K+1 K)\mathbf{H}^T(K+1) + \mathbf{R}(K+1)]^{-1}$
error matrix	$\mathbf{P}(K+1 K) = \phi(K+1 K)\mathbf{P}(K K)\phi^T(K+1 K) + \mathbf{T}(K)\mathbf{Q}(K)\mathbf{T}^T(K)$ $\mathbf{P}(K+1 K+1) = [\mathbf{I} - \mathbf{K}(K+1)\mathbf{H}(K+1)]\mathbf{P}(K+1 K)$

3. Data Compression

[21] ~ [24][27][28]

Data compression 은 source coding 의 한 방법으로 주어진 distortion 밀에 transmission rate 를 최소로 줄이는 방법을 찾는 것을 목적으로 한다. 이것은 information source 에 따라 data compression 하는 방법이 다름을 의미한다. discrete source 에 대하여 coding 하는 방법은 각 source element 가 독립적이고 source element 의 확률이 주어져 있다면 Huffman coding 방법이 optimal 이다. 여기서 는 Huff-

ete time continuous source 에 대하여 생각하기로 한다. 일반적으로 data compression 은 source modelling 하고 밀접한 관계가 있다. 왜냐하면 source 의 일반적인 성질을 알기 위하여 source 에 접근하는 방법은 source modelling 을 통해서만이 할 수 있기 때문이다. information source 의 선형(線形) model 로 제일 general 한 형태는

$$X_n = \sum_{l=0}^q b_e n_{n-e} - \sum_{k=1}^p a_k X_{n-k} \quad (76)$$

로 ARMA (Autoregressive Moving Average) model 이다.

여기서 $b_e = 0, e = 1, 2, \dots, q$ 이면 AR (Autoregressive) model 이라고 하고 $a_k = 0, k = 1, 2, \dots, p$ 이며 MA (Moving Average) model 이라 한다.

이 ARMA model 은 각 분야에 많이 이용되나 통신분야에서는 data compression 의 한 방법으로 speech signal modelling 에 많이 이용된다. 특히 AR model 은 all pole model 이라 하여 parameter $a_k, k = 1, 2, \dots, p$ 등을 구하는데 계산의 용이성 때문에 많이 이용되고 있다. 이 경우에 식(76)은

$$X_n = n_n - \sum_{k=1}^p a_k X_{n-k} \quad (77)$$

이 된다.

이 때 parameter a_k 는 signal X_k 의 correlation 으로 표시될 수 있는데 수식은 다음과 같다.

$$R_{xx}(K) = - \sum_{l=1}^p a_l R_{xx}(k-l) \text{ for } k > 0 \quad (78)$$

$$= - \sum_{l=1}^p a_l R_{xx}(-l) + \sigma^2 \text{ for } k = 0$$

단, 여기서

$$E[n_k n_e] = \begin{cases} \sigma^2 & \text{for } k = l \\ 0 & \text{for } k \neq l \end{cases} \quad (79)$$

이라 가정했다. 식(78)에서 $a_k, k = 1, 2, \dots, p$ 를 구할 수 있다. 구하는 방식중에 Levinson - Durbin algorithm^[21] 이라하여 다음식을 유도할 수 있다.

$$a_{11} = -R_{xx}(1) / R_{xx}(0) \quad (80)$$

$$\sigma_1^2 = (1 - a_{11}^2) R_{xx}(0) \quad (81)$$

$$a_{kk} = -[R_{xx}^*(K) + \sum_{l=1}^{k-1} a_{x-1, l} R_{xx}(k-l)] / \sigma_{x-1}^2 \quad (82)$$

$$a_{ki} = a_{k-1, i} + a_{kk} a_{k-1, k-i}^* \quad (83)$$

$$\sigma_k^2 = (1 - a_{kk}^2) \sigma_{k-1}^2 \quad (84)$$

식(78)과 위에 주어진 식과의 차이는 식(78)에서는 parameter $a_e, e = 1, 2, \dots, p$ 를 구하기 위하여 matrix 를 inverse 를 하여야 하나 식(80)~(84)는 param-

eter 를 구하기 위한 recursive structure 를 제공한다.

여기서 $\{a_{k1}, a_{k2}, \dots, a_{kk}, \sigma_k^2\}$ 은 $k = p$ 일 때 $\{a_1, a_2, \dots, a_p, \sigma^2\}$ 이 된다.

음성신호에서 (78)식에 의해 구한 계수를 이용하여 식(77)처럼 AR model 를 할 때를 LPC (Linear Predictor Coding) 이라 한다. 일반적으로 음성신호를 직접보내는 것보다 한 frame (보통 20 msec ~ 30 msec) 을 잡아 $\sigma^2, \{a_k\} k = 1, 2, \dots, p$, parameter, voiced / unvoiced parameter, pitch parameter 등을 보내는 것이 transmission rate 가 훨씬 적다. 현재 data compression rate 를 더 높이기 위하여 이러한 parameter 의 quantization 에 많은 연구가 행해지고 있다.^{[27][28]}

4. 結 論

Modern communication system 은 error 가 없는 통신 channel 의 제공과 동시에 channel 의 높은 효율성의 제공에 있다고 하겠다. 주어진 channel capacity^[29] 에 대하여 error 가 없는 통신 channel 의 제공이라는 면에서 볼 때 통신 system 의 신호처리 및 data 처리는 user 에게 ideal channel 를 제공하려는 노력의 일면이라 볼 수 있다. 그러나 실제로는 ideal channel 이 존재할 수 없으므로 user 에게 ideal channel 처럼 보이게끔 하는 것 (이것은 virtual channel 이라 말할 수 있다.) 이 필요한데 통신신호처리가 뒷받침을 하고 있다. 다른 말로 표현하면 통신신호처리는 user 가 원하는 virtual channel 과 physical channel 의 interface 역할을 한다. 즉 이것은 hierarchical structure 가 있다는 것을 의미한다. user 가 보는 virtual channel 이 high level 에 속하고 physical channel 이 low level 에 속한다면 통신 신호처리는 high level channel 에 low level channel 에 대한 transparency 를 제공하는 것이다. 이 관점에서 digital 통신을 바라볼 때 detection theory 는 bit transparency 를 제공하는 이론이라 할 수 있다. information theory 의 hamming code 는 byte transparency 를 제공하는 관점에서 볼 수 있다. 이 면에서 estimation theory 를 관찰한다면 신호의 파형 (또는 parameter) 의 transparency 를 제공하려고 하는 학문이라 할 수 있다. 참고로 data communication 의 X25 interface 는 packet level 의 transparency 의 제공을 목적으로 하고 있다.

參 考 文 獻

- [1] A. Papoulis; Probability, Random Variables, and Stochastic Processes, McGraw Hill, New York, 1965.
- [2] H. Van Trees; Detection and Estimation and Modulation Theory, Part 1 and 2 and 3, Wiley New York, 1968, 1971.
- [3] A. TANENBAUM; Network Protocols, Computing Surveys, vol. 13, No. 4, Dec. 1981.
- [4] A. TANENBAUM; Computer Networks Prentice Hall, 1981.
- [5] Davies, D.W. et al; Computer Networks and Their Protocols, John Wiley, 1979.
- [6] J.L. Melsa and D.L. COHN; "Decision and estimation theory," McGraw-Hill, 1978.
- [7] A.P. Safe et al; "Optimum systems control," Prentice Hall, 1977.
- [8] JAZWINSKI; "Stochastic processes and filtering theory," Academic Press, 1969.
- [9] R.E. Kalman and B.S. Bucy; "New results in linear filtering and prediction theory," Trans. ASME, J. Basic Eng vol. 83, 1961.
- [10] Mc GARTY; "Stochastic systems & state estimation" Wiley, 1974.
- [11] Gelb; "Applied optimal estimation.. M.I.T. Press, 1974.
- [12] D.A.S. Fraser; "Nonparametric methods in statistics" Wiley, New York, 1957.
- [13] R.H. RANGLES and D.A. WOLFE; "Introduction to the theory of nonparametric statistics" Wiley, 1979.
- [14] Davies and Barber; "Communication networks for computers" Wiley, 1973.
- [15] M.G. Kendall and A. Stuart; "The advanced theory of statistics" vol. 2, Inference and Relationship, Hafner, New York, 1961.
- [16] V.R. Algazi and R.M. Lerner; "Optimum binary detection in non-gaussian noise" IEEE Trans. Theory IT-12, No. 2, p. 269, 1965.
- [17] H. Sherman and B. Reiffen; "An optimum demodulator for poisson process: photo source detectors," Proc. IEEE 51, pp. 1316-1320 Oct. 1963
- [18] A.J. Viterb; "Principles of coherent communication" McGraw-Hill, New York, 1966.
- [19] F. RIESE and B. SZ-NAZY; "Functional analysis" Frederick Ungar. New York, 1972.
- [20] Box Jenkins; "Time series analysis, forecasting and control", Holden day, 1970.
- [21] Norbert Wiener; "Extrapolation, interpolation and smoothing of stationary time series" MIT Press, 1970.
- [22] A. PAPOULIS; "Maximum entropy and spectral estimation: a review" IEEE Trans. ASSP-29, No. 6, Dec. 1981.
- [23] S. M. KAY and S.L. MAPPLE; "Spectrum analysis - a modern perspective" IEEE Proc. Nov., 1981.
- [24] L.R. Rabiner/R.W. Schafter; "Digital processing of speech signals" Prentice Hall, 1978.
- [25] Proc. IEEE; "Packet communication" 1978 Nov. vol. 66 Packet Communication Special Volume.
- [26] Data Communication Special Issue IEEE Trans. COM 28 No. 4, April 1980.
- [27] A. BUZO, A.H. GRAY et al.; "Speech coding based upon vector quantization" IEEE Trans. ASSP-28 No. 5, Oct., 1980.
- [28] Y. Linde, A. BUZO and R.M. Gray; "An algorithm for vector quantizer design" IEEE Trans. Com. vol. COM-28 pp. 84-95 Jan. 1980.
- [29] Robert G. Gallager; "Information theory and reliable communication" 1968 John, Wiley.