信号処理(II)-Random process의 detection 및 estimation Karhunen · Loève의 展開, 한 画像의 SVD

(Signal Processing(II) — Detection and Estimation of Random Process, Karhunen Loève Expansion, SVD of an Image)

安 秀 桔*

(ANN, Souguil)

要 約

信號處理와 analysis를 위한 여러 基礎的인 技術이 招介되었다. 이들은 먼저 不確定性原理의 介入에 依하여 特히 交換不可能한 operator 들이 作用한 結果의 等號는 tolerance가 있을 수 있음과 random process 처리방법과 maximum entropy estimate 的인 思考方式을 通하여 在來式 確定論的 思考方式으로부터의 離脫을 길잡았다.

마지막으로 檢出, 推定 및 函數推定의 여러 技法과 covariance functron의 positive semi-definiteness 그리고 Karhunen-Loève 展開, 한 画像의 SVD 等이 說明됐다.

Abstract

In this paper several basic techniques for signal processing and analysis are surveyed. Firstly by the intervention of the uncertainty principle, an equality sign may have different degree of precision if non commutable operators are applied. Secondly maximum entropy estimate and randam process based viewpoint must be enhanced to get rid of the well established and reigning deterministic image of science.

Thirdly techniques for the analysis of a signal namely detection, estimation and modulation are explained as well as the positive definiteness of a covariance function, Karhunen-Loève expansion and SVD of an image.

1. 序 論

Es irrt der mensch solang er strebt.

Goethe, Faust

Faust 博士가 아니라도 人間은 역시 알수록 實은 안다고 믿는것이 헛된것으로 더욱 더 모르는것을 切感하게 된다.

Newton의 力學은 簡潔하고 確定論的인 數値를 준다. 그러나 이는 參與하는 element의 數가 많지 않고 멀리 있는 것은 無視할 수 있는 天體의 問題等에 있어서 强力할 수 있는 力學이고 n 個의 變數를 計算하려면 n 個의 서로 一次獨立인 關係式을 갖고 있어야 한다.

우리는 그렇치 못한 狀況에서 버리자니 아까운 많은 資料를 갖고 있을때 그나름대로 어떠한 結論을 얻되 우리가 알지 못하는 element 들 때문에 일어나는 오차가 어느 程度 큰 것인가를 把握하고 未知要素의 影響이 그다지 크지않다면 비록 얻어진 結論은(等號로 나타낼수있을 程度)確定된 값이 아닐지라도 이 計算結果를 指針으로 삼을 수 있다. 實例로서 비근한 fourier 變換을 通한 한時間函數의 spectrum에 對해서도 (1-1)式((1-•)은 前回글의 式)으로서는 式을 成立하지만 實際로 우리는 周波數測定을($-\infty, \infty$)동안 時間을 들여서 行할 수가 없다. 有限된 時間에 限定하면

* 正會員, 서울大學校 工科大學 電子科
(Dept. of Electronics Engineering, Seoul National Univ.)
接受日字: 1979年 11月 27日

不確定性의 介入때문에 이미 嚴密한 뜻에 있어서의 等號를 使用할 수가(비록 가장 蓋然性이 있는 값을 주기는 하지만)없고 time domain에서 時點은 더 밝혀지나 freq. domain에서 誤差가 생기게 된다.[1][2]

Random process에 있어서는 우리의 認識과 無關하게 어떠한 숨어 있는 機構에 依해서 定하여진 確率에 따라서 現象이 나타나고 있으니까 그것을 究明하자는 客觀學派가 있고 우리가 알고 있는 知識범위로서는 그리고 새로 얻어진 情報를 加味하였을때 어떻게 우리의 結論을 修正하여야 얻어진 資料範圍에서 가장 非偏見(non committal)的인가를 알고저하는 主觀學派가 그中에서도 두드러지는데 信號處理의 立場으로서는 兩쪽을 다尊重하게 된다(個人的인 見解로서는 後者가 더 묘미가 있어 보인다).

Noise等에 依해서 오염된 信號는 그原形을 찾아보기가 힘드는데 이때 原信號가 어떠한 形態일 것이라고 想定하는 것은 random process에 있어서는 너무 위험한 일이고 주어진 條件으로서 얻을 수 있는 가장 一般的인 結論을 얻어야 하며 따라서 누가 그 立場에 섰더라도 같은 結論이 나와야 하며 그다음 實際에 일어난 일을 通하여 그推定이 틀렸음이 나타날지라도 그때 立場, 그때 條件으로서는 如前히 똑 같은 結論이 가장 健全한 그러한 結論을 얻기 위한 技法을 찾는것이 우리의 目的이 되겠다. 이것은 統計學的인 推定으로서 주어진 情報를 基礎로하여 할 수 있는 最小偏見의 推定이며 傷失해 버린(또는 아직 얻지못한) 情報에 對해서 最大限으로 non committal한 것이다. 순서나 질서 其他에 對해서 아무런 假定을 하지 않기 때문에 maximum entropy의 推定이라고 할 수 있다. 熱力學에 있어서도 모든 速度의 分子가 完全히 뒤섞여서 추운 곳(平均分子速度 小)과 더운 곳(平均分子速度 大)이 따로 없는 狀態가 maximum entropy 狀態로서 熱的見地에서 完全無秩序의 狀態이다. 모든 秩序는 entropy(그뜻에 있어서의)를 낮게 하는 것이다.

2. Detection, Estimation 및 Modulation Theory

우리가 흔히 使用하고 있는 것과는 달리 信號處理에서는 이 三單語를 다음과 같은 뜻으로 使用한다.

但, 모든信號가 noise로서 심히 汚染되어 있어서 推定以外에 原形을 알길이 없는 것으로 한다.

① Detection theory

보내진 信號가 digital 信號로서 許容된 level 中에서 어떤 것이 보내진지를 判斷하는것으로서 이 單語는 判斷理論으로 飜譯할 性質의 것이다.

② Estimation Theory

이것은 보내진 信號가 analog 즉 continuous 한 것으로서 無限히 微細하게 갈라질 수 있는 값일 境遇로서 이러한 continuous 量을 되찾아 내는 것으로서 前者보다 어렵다. 이는 數値의 推定에 該當되는 것이다.

③ Modulation theory(coutinuous waveform estimation)

이單語는 가장 誤解되기 쉬운 立場에 있는데 이는時間的으로 變動하고 있는 어떤 parameter를 繼續 推定하여 그 時間函數的인 값의 變動의 全體를 把握하는 것으로서 가장 어렵다. 이는 continuous waveform estimation이라고 불리워야 할 性質의 것으로 modulation이라는 말에 廣義的으로 包含될 수 있다손 치더라도 狹義的인(우리가 잘알고 있는)뜻과 混同되기 쉬우므로 이것을 "變調理論" 이란 말로는 飜譯하지 않는 것이 좋을듯하다.

이러한 推定은 推定의 優劣이 立場에따라 다르기 때문에 그優劣을 판가름하는 criterion이 주어져야 좋다, 나쁘다 할 수 있고 우리들의 直觀이 全體일수는 없다. 따라서 optimal 한推定이란 얼마라도 있을 수 있고 그中 어느 것이 가장 좋은가는 criterion에 따라 다르기때문에 먼저 criterion을 定하고 그다음에 變分法에 依하여 optimum 한推定을 行한다.

우리들은 source 內에서 일어나고 있는 現實을 把握할 수 없고 우리에게 觀察될때에는 이미 原形에서 離脫하여 있기때문에 原信號는 推定하는 수 밖에 없다는 不可知論的인 立場에 서는 것이다. 즉 우리는 어느 現象에 對한 把握을 觀察된 結果量으로부터 出發하고 이러한 觀察量이 n 個이라면 이들은 n 次元空間(observation space라고 함)內의 點으로서 나타난다. 그리고 이點과 原點을 맺는 vector를 다음과 같이 나타낸다.[3][4]

$$\mathbf{r} \triangleq \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ \vdots \\ r_n \end{bmatrix} \quad \cdots\cdots\cdots (2-1)$$

發生하고 있는信號가 binary일지라도 觀察空間에 對한 mapping結果는 smooth 한 連續으로 나타날때가 많다.

"1" 이 일어났다고 假定하는 것을 Hypothese H_1 으로 나타내고 "0" 이 일어났다고 假定하는 것을 Hypothese H_0 로 나타낸다면 H_0 下에 觀察空間內 $\mathbf{R}$ 點에 그것이 mapping되는 確率(conditional)은 $P_{\mathbf{r}|H_0}(\mathbf{R}|H_0)$로 나타낸다. "0" 이 일어났는데 H_0로 取扱하면 옳고 H_1으로 取扱하면 틀린것인데 이때 罰則

으로 먹이는 값을(cost라고 함) C_{00}, C_{10}로 나타내면 C_{00}는(옳았던 경우이니까) 당연히 C_{10}보다 작게 잡아야 할 것이다.[5][6]

H_1下에 觀察空間內 **R** 點에 mapping되는 確率을 $P_{r|H_1}(\mathbf{R}|H_1)$으로 나타내고 "1"이 일어났는데 H_0로 取扱한다면 判定이 틀린것이고 H_1으로 取扱한다면 맞는 것인데 이때 Cost를 C_{01}, C_{11}으로 나타내면 다음과 같이 하여야 할 것이다.

$$\left.\begin{array}{l} C_{10} > C_{00} \\ C_{01} > C_{11} \end{array}\right\} \quad \cdots\cdots\cdots\cdots (2-2)$$

3. Bayes Criterion

이때 criterion으로서 다음式으로 定義되는 Risk R를 擇하고

$$\begin{aligned} R = & P_0 C_{00} \int_{Z_0} P_{r|H_0}(\mathbf{R}|H_0)\,d\mathbf{R} \\ & + P_0 C_{10} \int_{Z-Z_0} P_{r|H_0}(\mathbf{R}|H_0)\,d\mathbf{R} \\ & + P_1 C_{01} \int_{Z_0} P_{r|H_1}(\mathbf{R}|H_1)\,d\mathbf{R} \\ & + P_1 C_{11} \int_{Z-Z_0} P_{r|H_1}(\mathbf{R}|H_1)\,d\mathbf{R} \quad \cdots\cdots (2-3) \end{aligned}$$

인 R가 最少가 되게끔하면 다음式을 얻는다.
但, 이때 積分은 n重積分이고 Z_0는 觀察空間內에서 source의 "0" 發生에 對應하는 部分이고 $Z-Z_0$는 "1" 發生에 對應하는 部分이며 $Z = Z_0 \cup Z_1$이다.

$$\frac{P_{r|H_1}(\mathbf{R}|H_1)}{P_{r|H_0}(\mathbf{R}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P_0\,(C_{10} - C_{00})}{P_1\,(C_{01} - C_{11})} \Bigg\} \cdots\cdots (2-4)$$

이는 左項이 右項보다 크면 H_1으로 생각하고 그反對이면 H_0로 생각하는 것이 第一 Risk가 적다는 것을 나타내고 이것을 bayes criterion이라고 한다. (2-4)式의 左邊을 likelihood ratio라 부르고 $\Lambda(\mathbf{R})$로서 흔히 나타낸다. 이러한 判定을 LRT(likelihood ratio test)라고 하고 兩邊의 logarithm을 使用하는 수도 있다.[7][8]

이에 反하여 cost를 주지않고 例를 들어 P_F(前回序論參照)를 어떠한 값으로 定하여 놓고 P_D(上同)을 그 constraint 下에 最大로 하는(變分法的)方法을 Neyman-pearson criterion 이라 한다.

4. Estimation Theory

序論에서 말한 바와 같이 信號源에서 發生하고 있는 어떠한 連續量의 값을 推定하는 것이 estimation으로서 다음과 같이 實例를 通하여 說明하겠다.

信號源에서 a인 連續量이 發生하고 있는데 平均値가 零, variance가 σ_n^2인 Gaussian noise가 重첩되어 있다면 a가 發生하고있는 事象을 A로 나타내어서 觀察空間에는

$$P_{r|a}(R|A) = \frac{1}{\sqrt{2\pi}\,\sigma_n} \exp\left\{-\frac{(R-A)^2}{2\sigma_n^2}\right\} \quad \cdots\cdots\cdots\cdots (2-5)$$

인確率로서 分布된다. 但, A는 noise 全無할때 a가 觀察空間에 mapping 되는 點의 座標이기도 하다. 觀察空間 **R** 點에서 信號가 觀察되었을때 우리는 原信號가 얼마였다고 推定하여야 하는가를 살핀다.

Detection theory의 경우에 使用되었던 Bayes criterion에 準한것이 使用되는데 j事象을 i로 判定($i = j$이면 제대로 判定한 것임)할 경우의 cost를 C_{ij}로 나타내고 i事象이 일어나는 事前確率을 P_i라 하면 signal이 M個 따라서 M hypothese일때 cost는 M×M가지가 되겠다. 이번 estimation의 경우에는 continuous한 量으로서 原信號 a를 그의 觀察空間에의 mapping인 **R**로부터 推定하여 이를 $\hat{a}(\mathbf{R})$라고 하면 cost는 bivariate function이 되고 이것을 $C(a, \hat{a})$라 表記한다. 一般的으로 cost는 a에 無關係로 estimation의 error만의 函數로 定한다. error vector $a_E(\mathbf{R})$를 다음式과 같이 定義하면 cost function은 a_E만의(single variable) 函數가 된다.

$$a_E(\mathbf{R}) \triangleq \hat{a}(\mathbf{R}) - a \quad \cdots\cdots\cdots\cdots (2-6)$$

이 cost function의 設定은 여러 가지 있을 수 있지만 다음과 같은 3가지가 가장 많이 쓰인다.

① Error의 自乘에 比例하는 cost를 課해서 正의 error의 負의 error를 同等하게, 그리고 error 가 클수록 severe하게 課하는 方法. 이 方法에 依한 推定을 MMSE(Minimum Mean Square Error) 推定이라하여 많이 活用된다. 高射砲照準을 爲해 Norbert Wiener가 推定을 硏究하였을 때에도 이것을 criterion으로 使用하였다. 그리고 이 硏究는 stationary한 process의 경우이었다.[9][10][11]

$$C(a_E) = a_E^2 \quad \cdots\cdots\cdots\cdots (2-7)$$

② 다음은

$$C(a_E) = |a_E| \quad \cdots\cdots\cdots\cdots (2-8)$$

로서 cost 를 課하는 方法이 있다. 그리고 또 하나는 ③ 推定이 原값의 $\pm\frac{\Delta}{2}$ 範圍內에 있으면 cost 를 零 그리고 그밖에 있으면 一樣하게 1 을 課하는 方法도 있다.

$$\left.\begin{aligned} C(a_E) &= 0, \ |a_E| \le \frac{\Delta}{2} \\ C(a_E) &= 1, \ |a_E| > \frac{\Delta}{2} \end{aligned}\right\} \quad \cdots\cdots (2-9)$$

Bayes estimation 이란 事前確率과 cost function 을 알면 Risk R이 가장 적게 되게 推定을 하는 것을 말하는데 Risk 는 다음式으로 주어진다.

$$R = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} C(A-\hat{a}(\mathbf{R}))\, P_{a,r}(A,\mathbf{R})\, d\mathbf{R}\right\} dA \quad \cdots\cdots (2-10)$$

但, $P_{a,r}(A,\mathbf{R})$란 A와 $\mathbf{R}$의 joint probability 이다.

MMSE 의 경우는

$$R_{ms} = \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}[A-\hat{a}(\mathbf{R})]^2 P_{a,r}(A,\mathbf{R})\, d\mathbf{R}\right\} dA \quad \cdots\cdots (2-11)$$

처음 積分을 minimize 시킴으로서 얻을수 있다. 처음 積分을 $\hat{a}(\mathbf{R})$에 依해서 微分하여 零과 같게 하여 줌으로서 MMSE $\hat{a}_{ms}$를 다음式과 같이 얻을 수 있다. [12][13]

$$\hat{a}_{ms}(\mathbf{R}) = \int_{-\infty}^{\infty} A P_{a|r}(A|\mathbf{R})\, dA \quad \cdots\cdots (2-12)$$

即, MMSE Criterion 에 依한 a의 推定値란 確率分布 $P_{a|r}(A|\mathbf{R})$에 依한 A의 平均値를 推定値로서 擇하는 것이다. (2-12)式은 P라는 確率分布를 갖는 $r.v.$ A의 ensemble 平均을 求하는 一般式이다. MMSE 推定은 쉽게 首肯이가는 式이지만 이때 確率은 事後確率로서 觀察空間內 $\mathbf{R}$點에 mapping 을 하는 原信號가 a일 確率을 나타내는 即, $P_{a|r}(A|\mathbf{R})$ 인 事實에 조심하여야 한다. 即, MMSE 推定은 事後確率에 依한 A의平均을 $\hat{a}(\mathbf{R})$로서 擇하면 된다.

Cost function 中에서 ②의 경우는

$$R_{abs} = \int_{-\infty}^{\infty}\left\{P_r(\mathbf{R})\int_{-\infty}^{\infty}[\,|A-\hat{a}(\mathbf{R})|\,]\, P_{a|r}(A|\mathbf{R})\, dA\right\} d\mathbf{R} \quad \cdots\cdots (2-13)$$

으로 나타나고 이를 最少로하는 $\hat{a}(\mathbf{R})$의 값 $\hat{a}_{abs}(\mathbf{R})$는

$$\int_{-\infty}^{\hat{a}_{abs}(\mathbf{R})} P_{a|r}(A|\mathbf{R})\, dA = \int_{\hat{a}_{abs}(\mathbf{R})}^{\infty} P_{a|r}(A|\mathbf{R})\, dA \quad \cdots\cdots (2-14)$$

即, 事後確率 $p.d.f.$의 median(그點을 中心으로 左側의 $p.d.f.$ 曲線下의 面積과 右側의 曲線下의 面積이 같게되는 $r.v.$値)이 된다.

③ 의 경우는 binary 的인 cost로서 맞았는가, 안맞았는가 만으로 cost를 課하는 方法인데 그 結果로 얻는 推定値 $\hat{a}_{map}(\mathbf{R})$는 事後確率 $p.d.f.$에서 勾配가 零인 點 即, 最高確率을 주는 $r.v.$의 값이 된다. [16] 이方法에 依한 判定을 MAP(maximum a posteriori probability) 判定이라고 하고 MMSE의 경우와 함께 가장 愛用된다.

$$\left.\frac{\partial \ln P_{a|r}(A|\mathbf{R})}{\partial A}\right|_{A=\hat{a}_{map}(\mathbf{R})} = 0 \quad \cdots\cdots (2-15)$$

이 세가지 $\hat{a}(\mathbf{R})$는 結局은 같은값을 주게될 때가 많다. 事後確率은 다음 Bayes 의式에 依해서 事前確率과 聯關지어질 수 있다.

$$P_{a|r}(A|\mathbf{R}) = \frac{P_{r|a}(\mathbf{R}|A)\, P_a(A)}{P_r(\mathbf{R})} \quad \cdots\cdots (2-16)$$

$$\ell_n P_{a|r}(A|\mathbf{R}) = \ell_n P_{r|a}(\mathbf{R}|A) + \ell_n P_a(A) - \ell_n P_r(\mathbf{R}) \quad \cdots\cdots (2-17)$$

5. Non random parameter estimation

Parameter 가 random variable 이 아닐때 Bayes test를 行하면 (2-1)式의 경우와 달리 Risk는

$$R(A) \triangleq \int_{-\infty}^{\infty}[\hat{a}(\mathbf{R})-A]^2 P_{r|a}(\mathbf{R}|A)\, d\mathbf{R} \quad (2-28)$$

로서 平均을 하기 위한 積分은 $\mathbf{R}$에 對해서만 行하게 된다.(A는 $r.v.$가 아니고 決定論的 常數이다)

(2-11)에서의 計算때와 같이 最小化는 integrand 인 $[\hat{a}(\mathbf{R})-A]=0$로서 이루어지고 따라서 $\hat{a}_{ms}(\mathbf{R})=A$이지만 A를 알지못 하니까 뜻이 없다. 따라서 推定의 優劣判定을 위한 새로운 measures of quality를 찾아내야 한다. 한편 推定値의 expectation 을 求해보면 다음式으로 주어지는데

$$E[\hat{a}(\mathbf{R})] \triangleq \int_{-\infty}^{\infty} \hat{a}(\mathbf{R})\, P_{r|a}(\mathbf{R}|A)\, d\mathbf{R} \quad \cdots\cdots (2-19)$$

이것이 願하는데로 A와 一致할때도 있지만(이것을

unbiased estimate 라고 함) A와 一定한(A와 無關係한) 差를 가질때도 있고 A에 따라 이 bias가 달라지는 수 까지도 있다(前者를 with known bias 後者를 with unknown bias라고 한다).

• Maximum likelihood estimation

第四節의 實例의 경우를 再活用한다면

$r = A + n$

$$P_{r|a}(R|A) = \frac{1}{\sqrt{2\pi}\,\sigma_n}\exp\left\{-\frac{1}{2\sigma_n^2}(R-A)^2\right\} \quad \cdots\cdots (2-20)$$

parameter a의 값에 따라 觀察空間內 R의 分布의 中心値A가 달라지는 것이지만 R를 觀察하였을때 우선 이R를 A의 값으로 推定하는 것이 옳다(noise의 平均値는 零이니까) 따라서

$$\hat{a}_{ml}(\mathbf{R}) = R \quad \cdots\cdots (2-21)$$

$P_{\mathbf{r}|a}(\mathbf{R}|A)$를 A의 函數로 보았을때 이를 likelihood function, 이의 logarithm를 log likelihood function이라 부르고 이들을 極大로 하는 A의 값을 推定値로 삼는 것이 옳다. 卽,다음式이 成立한다.[18]

$$\left.\frac{\partial \ln P_{\mathbf{r}|a}(\mathbf{R}|A)}{\partial A}\right|_{A=\hat{a}_{mL}(\mathbf{R})} = 0 \quad \cdots\cdots (2-22)$$

卽, 事前確率의 mode(最高點)를 주는 A의 값이 $\hat{a}_{ml}(\mathbf{R})$가 되어야 한다. (22)를 likelihood equation이라한다.

6. General Gaussian 問題

(2 - 5)의 事前確率을 다음과같이 N次元의 경우로 一般化할 수 있다.

$$Pr(\mathbf{R}) = \frac{1}{(2\pi)^{\frac{N}{2}}|\Lambda|^{\frac{1}{2}}}\exp\left[-\frac{1}{2}(\mathbf{R}^T-\mathbf{m}^T)\Lambda^{-1}(\mathbf{R}-\mathbf{m})\right] \quad \cdots\cdots (2-23)$$

但, $E[\mathbf{r}] = \mathbf{m}$ 이다.

(2 - 23)의 指數函數의 指數部인 $(\mathbf{R}^T-\mathbf{m}^T)\Lambda^{-1}(\mathbf{R}-\mathbf{m})$는 二次型式으로 計算結果가 scalar가 되는 것이며 Λ는 (1 - 27b)에서 보는바 covariance matrix 이고

$$\Lambda \triangleq E[(\boldsymbol{V}-\mathbf{m})(\boldsymbol{V}^T-\mathbf{m}^T)] \quad \cdots\cdots (2-24)$$

로 주어진다. $\mathbf{r}$ 는

$$\mathbf{r}^T = [r_1, r_2, \cdots\cdots r_N] \quad \cdots\cdots (2-25)$$

로서 $r.v.$들이고 이 process가 Gaussian이면 이들 사이에 correlation이 없는 座標系를 얻을 수 있다. (eigen vectors) 그때에는

$$\Lambda = \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \cdot & \\ & & & \cdot \\ 0 & & & \sigma_N^2 \end{bmatrix} \quad 또는 \quad \Lambda = \sigma^2 \mathbf{I} \quad \cdots\cdots (2-26)$$

의 型式이 된다. 卽,各各의 $r.v.$ 사이에 correlation이 全혀 없게 된다. (2 - 26)의 두번째 式은 各各의 Gaussion process가 서로 statistically independent 하며 各各의 variance들 마져 서로 같은 경우이다[14]

한 process $\mathbf{r}$이 Gaussian이라면 모든 finite(energy 또는 power) vector $\mathbf{G}$에 對해서 이들間 inner product Z

$$Z = \sum_{i=1}^{N} g_i r_i \triangleq \mathbf{G}^T\mathbf{r} \quad \cdots\cdots (2-27)$$

가 언제나 Gaussian이 된다. 이때 $\mathbf{G}^T$는 vector $\mathbf{r}$에 作用하여 Z라는 scalar를 만들어주기 때문에 co-vecteur라 하고 $\mathbf{r}$ 과 $\mathbf{G}^T$는 dual space를 形成한다.

(2 - 27)은 l_2 space라고 하는 sequence space內에서의 展開이고 sample function $x(t)$를 어떠한 square integrable deterministic function $f(t)$와 곱해서 $[T_\alpha, T_\beta]$ (週期의 整數倍)사이에 積分하면 $r.v.$ x_f를 얻게 되는데 이는(2 - 27)의 L_2 space 에 있어서의 counter part이다.

$$x_f \triangleq \int_{T\alpha}^{T\beta} x(t) f(t)\, dt \quad \cdots\cdots (2-28)$$

그의 Ensemble 平均은

$$E[x_f] = \overline{x_f} \triangleq E\int_{T\alpha}^{T\beta} x(t) f(t)\,dt$$

$$= \int_{T\alpha}^{T\beta} m_x(t) f(t)\,dt \quad \cdots\cdots (2-29)$$

但, $m_x(t)$는 $x(t)$의 ensemble 平均으로 時間의 函數이다. 이때 variance를 求하는 式은

$$v_{ar}(x_f) = E[(x_f - \overline{x_f})^2] = \int_{T\alpha}^{T\beta}[x(t) - m_x(t)]\cdot$$

$$f(t)dt \int_{T\alpha}^{T\beta} [\, x(\mathbf{u}) - m_x(\mathbf{u}) \,]\, f(u)du \cdots\cdots (2-30)$$

(2 - 24)에 準하여 $\mathbf{K}_x(t, u) = E[\{x(t) - m_x(t)\}\{x(u) - m_x(u)\}]$ 로 하면 이는 $r. v.\ x(\cdot)$의 時刻 t 와 時刻 u에 있어서의 값사이의 covariance 이다. 따라서 (2 - 30)은

$$v_{ar}(x_f) = \iint_{T\alpha}^{T\beta} f(t)\, K_x(t, u) f(u)\, dt\, du \cdots\cdots (2-31)$$

Variance 란 두개의 factors 가 同一하여 그의 相乘積이 負가 될 수 없는 것이기 때문에 언제나 non - negative 이다. 따라서 $K_x(t, u)$도 positive semidefinte 하다.[15]

Covariance stationary 한 process 에서는 argument 가 單一한 變數 t - u가 되어서 $K_x(t-u)$로 表記한다. white noise 의 경우에는 다음과 같다.

$$K_x(t-u) = 2D\delta(t-u) \cdots\cdots\cdots\cdots\cdots (2-32)$$

7. Karhunen - Loève 展開

Matrix 의 eigen value 의 槪念에 準해서 函數空間에서도 eigen value λ_j 와 eigen function $\phi_j(u)$이 있고 이들은 다음式을 滿足시킨다.[17][19][3]

$$\int_{T\alpha}^{T\beta} K_x(t, u)\, \phi_j(u)\, du = \lambda_j\, \phi_j(t)$$

$$T_\alpha \leq t \leq T_\beta \cdots\cdots\cdots\cdots\cdots (2-33)$$

$K_x(t, u)$는 x process 의 covariance function 이며 이 積分方程式에서와 같이 한 函數에 作用하여 積分될때 kernel(of an integral equation)이라 불리우고 이경우에 있어서는 covariance 이기 때문에 positive semi -definite이다. eigenvalue가 서로 相異할 경우에는 이들에 對應하는 eigenfunction 들이 서로 直交性을 가진 等 한 matrix 의 eigenvalue 等에 成立하는 特性들이 거의 모두 다 適用된다.

이와 같이 kernel이 한random process의 covariance일때 (2 - 33)式의 eigenfunction 을 座標軸으로하여 (楕圓에서 長軸과 短軸을 座標軸으로 擇하므로서 $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ 로되어 cross term 이 없어지는 것과 analogous 하게) kernel의 固有의 特性을 살려 그 random process 를 數式化할때 이를 Karhunen - Loève의 展開라 하며 process 들이 gaussian 이라면 係數들이 서로 statistically independent하다는 特徵이 있다. 이 座標系는 그 process 自體가 決定하여 주는 것으로서 特히 便利할때가 많다.[22]

T_α, T_β등 區間이 $\pm\infty$가 될때 eigen value들은 discrete하지 않고 continuum을 형성하여 連續的인 energy의 分布를 나타낸다. 卽, spectrum이란 eigen value의 모임이고 이때 eigen function들은 steady 한 週期函數가 된다.

8. Optimum linear filter

Filter의 槪念을 一般化하여 predictor, smoother등을 包含시킨다. 이러한 一般化된 filter의 槪念은 在來의 뜻에 있어서의 filter가 現在 들어 오는 信號로부터 現在의 出力을 얻어내는 것임에 比하여 觀察時點의 前後를 가리지 않고 모든 data 로부터 maximum entropy (序論末尾參照) 推定을 하되 어느 時點까지의 data 로서 같은時點에서의 出力(一般的으로 거의 noise free) 을 얻어 주는 경우라면 또는 周波數에 따라 gain을 달리하여 주는 것이라면 filtering이 되고 그 보다도 늦은 時點에서의 값을 推算하여주는 것이라면 prediction 이 되는 것이며 noise 影響을 除去한 값을 주는 것이 smoothing 이 된다. 이들은 許容誤差가 問題일뿐이지 모두 다 可能하며 Wiener 와 Kalman을 爲始하여 信號處理의 主要手段을 提供하고 있다.[24][25]

우선 filter의 定義로서는 現在, 未來, 過去의 入力 data를 出發點으로 하여 어떠한 目的에 適合하는 出力을 내어주는 裝置라고 말할 수 있다. 未來의 入力을 除外하였을때 causality가 成立하니까 realizable filter 라 말할 수 있고 觀察時點間의 觀察 못하였던 時點에서의 값을 얻는 것이 interpolation, 現在以前의 date로부터 現在以後의 data(가장 妥當한 값)를 얻는것이 extrapolation이 된다. 画面의 處理를 爲해서는 two dimensional sequences와 system을 위한 transform을 定義하여야 한다.

9. Two - dimensional sequences 및 systems

二次元画像을 위해서 x, y discrete 變數를 各各 m 및 n 라 하면 한画面의 밝기를 나타내는 二次元信號 $x(m, n)$의 Fourier 變換은 다음 式으로 定義한다.

$$X(j\omega_1, j\omega_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(m, n)\, e^{-j\omega_1 m} e^{-j\omega_2 n}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots (2-34)$$

이때 逆變換은 다음 式으로 나타낸다.

$$x(m,n) = \frac{1}{4\pi^2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} X(jw_1, jw_2)\cdot$$

$$e^{jw_1m}\, e^{jw_2n}\, dw_1\, dw_2 \cdots\cdots\cdots\cdots (2-35)$$

그리고 한 system을 通하여 한 二次元入力 $x(m,n)$가 그의 出力에서 $y(m,n)$로 바뀔때 다음 式으로 定義되는 $H(jw_1, jw_2)$를 two-dimensional transfer function 이라한다.

$$Y(jw_1, jw_2) = H(jw_1, jw_2)\cdot X(jw_1, jw_2) \cdots\cdots\cdots\cdots (2-36)$$

但, $Y(jw_1, jw_2) = \sum_{n=-\infty}^{\infty} y(m,n)\, e^{-jw_1m}\, e^{-jw_2n}$

어떤 medium에 어떠한 degradation mechanism을 생각할 수 있는데 그때 이것을 나타내는 $H(jw_1, jw_2)$를 생각할 수 있고 이 degradation이 信號의 自由度를 縮小시키는 것이 아니라면 H의 inverse operator를 求할 수 있고 $y(m,n)$로부터 $x(m,n)$를 復元시킬 수도 있다. 이를 deconvolution 이라고 한다. TV와 같이 scanning을 通하여 一次元 信號를 얻어 處理하는 것이 普通이나 digital 信號處理方法에서는 random access memory에 記憶시키고 나서는 水平方向(m 먼저 그리고, n)부터 먼저 處理하거나, 垂直方向(n 먼저 그리고, m)부터 먼저 處理하는 것은 마음대로 이다. Defocussing된 image를 한方向으로 analogue 信號處理方法에 依해서 restore한 case 가 있다. 이는 直交方向으로 되풀이 할 수도 있다.[23]

10. 畵面信號處理

画面信號處理는 二次元일때와 그 以上의 경우가 있는데 代表的인 前者의 例로서는 color TV에서의 edge enhancing 技法等이고 後者의 例는 medical tomography의 경우等이다.[26] 어느 경우에 있어서나 二次元信號處理가 꼭 必要한데 (2-26)式과 같이 出力信號로 變하게 하는 system(또는 媒質)의 特性을 [H]와 같이 나타내면 noise의 影響을 감안하여 다음 式이 成立한다.

$$\mathbf{Y} = [H]\,\mathbf{X} + \mathbf{n} \cdots\cdots\cdots\cdots (2-37)$$

image restore 過程이란 이러한 $\mathbf{Y}$로부터 다음 式을 通하여 $\hat{\mathbf{X}}$를 求하는 것을 ($\mathbf{X}$ 原形은 入手下能) 말한다.[27][29]

$$\hat{\mathbf{X}} = [H]^{+}\,\mathbf{Y} \cdots\cdots\cdots\cdots (2-38)$$

이때 使用된 $[H]^{+}$라는 operator는 [H]의 Moore-penrose pseudo inverse라 한다. 이러한 inverse는 다음과 같이 求하여진다. unitary 變換이란 座標軸의 回轉等에 隨伴하는 vector의 表現式 變化等을 計算하는데 使用되는 operator로서

$$[U][U]^t = [I],\ [V][V]^t = [I] \cdots (2-39)$$

인 性質이 있는 [U]나 [V] matrix로서 나타낼 수 있다. 이들은 unique하지 않는데 그中에서 [H]의 左右에서 $[U]^t[H][V]$와 같이 作用시키면 結果가 Jordan의 形式이 (簡單히 말해 diagonal matrix가)되게 하는 그러한 [U]와 [V]를 選擇하면

$$[\alpha] = [U]^t[H][V] \cdots\cdots\cdots\cdots (2-40)$$

$$[H] = [U][\alpha][V]^t$$

$$= [\mathbf{U}_1, \mathbf{U}_2, \mathbf{U}_3, \cdots \mathbf{U}_N][\alpha]\cdot \begin{bmatrix} \mathbf{V}_1^t \\ \mathbf{V}_2^t \\ \vdots \\ \mathbf{V}_N^t \end{bmatrix}$$

$$= \sum_{i=1}^{N} \alpha_{ii}\, \mathbf{U}_i\, \mathbf{V}_i^t \cdots\cdots\cdots\cdots (2-41)$$

但, $\mathbf{U}_i$ 들은 column vector, $\mathbf{V}_i^t$ 들은 row vector, 그리고 $[\alpha]$는 Jordan 形式 그리고 처음 image $\mathbf{Y}$에 따라서는 實際項數 R가 $R<N$일 수도 있다.

이 展開式을 그 畵像의 SVD(singular value decomposition)이라 부른다. $R<N$일때에는 α의 diagonal element 中 零인 것이 있음을 뜻하고 (2-41)式은 $\alpha_1 \mathbf{U}_1\mathbf{Y}_1^t + \alpha_2\mathbf{U}_2\mathbf{Y}_2 + \cdots + \alpha_N\mathbf{U}_N\mathbf{Y}_N^t$이 되고 各項의 vector의 外積 $\mathbf{U}_i\,\mathbf{V}_i^t$는 matrix를 形成하지만 rank를 1에 不過하고 따라서 (2-41)式의 各項은 自己만이 支配하는 獨立性을 갖고 있어서 이들의 weighted sum이 되며 各項은 一次獨立 卽, 直交의 關係에 있다. 이러한 條件에 맞는 unitary transform이 아니였더라면 (2-41)式은 $\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_{ij}\,\mathbf{U}_i\mathbf{V}_j^t$이였어야 할 것이다. SVD matrix의 eigen vector decomposition과 같은 內容으로서 [H]의 eigen value 들中 N-R個는 零이고 나머지 R個의 eigen value를 λ_i, $i=1, 2, \cdots\cdots R$라 하면 $\alpha_{ii} = \lambda_i^{\frac{1}{2}}$이다. 따라서 (2-41) 式은

$$[H] = [U][\lambda^{\frac{1}{2}}][V]^t$$

한편 $[H][H]^t = [U][\lambda^{\frac{1}{2}}][V]^t[V][\lambda^{\frac{1}{2}}][U]^t$

$$= [U][\lambda][U]^t$$

$[H]^t[H] = [V][\lambda][V]^t$ 이다. $\cdots\cdots$ (2-42)

이때 pseudo inverse$[H]^+$는 다음 式으로 주어진다.

$$[H]^+=[U][\lambda^{-\frac{1}{2}}][V]^t=\sum_{i=1}^{R}\lambda_i^{-\frac{1}{2}}\mathbf{V}_i\mathbf{U}_i^t \quad \cdots\cdots (2-43)$$

但, $\mathbf{U}_i\ \mathbf{V}_i$는 (2 - 42) 第二, 第三式으로 보아 各各 $[H][H]^t$와 $[H]^t[H]$의 eigenvector 들이다. 이것을 利用하면 (2 - 38)式은

$$\hat{\mathbf{X}}=[H]^+[H]\mathbf{X}+[H]^+\mathbf{n}=\sum_{i=1}^{R}\lambda_i^{-\frac{1}{2}}\mathbf{V}_i\mathbf{U}_i^t$$

$$\{[H]\mathbf{X}\}+\sum_{i=1}^{R}\lambda_i^{-\frac{1}{2}}\mathbf{V}_i\mathbf{U}_i^t\{\mathbf{n}\} \quad \cdots\cdots (2-44)$$

實際 restoration은 다음 式으로서 行한다.

$$\hat{\mathbf{X}}=\sum_{i=1}^{R}\lambda_i^{-\frac{1}{2}}\mathbf{V}_i\mathbf{U}_i^t\mathbf{Y} \cdots\cdots (2-45)$$

이것은 R 만큼의 weighted rank 1 matrice 의 重첩으로 볼 수 있으며 noise에 따라서는 重첩個數를 R 보다 적게 잡아서 도리어 明確하게 나타날때도 있다.

Restoration은 역시 推定에 屬하므로 後述의 Wiener Hopf의 technigue 가 必要하게 될 경우가 많다.

이러한 二次元信號를 手扱하기 爲해서 二次元 filter 가 必要한데 그의 特性은 (2 - 36)式의 H($j\omega_1$, $j\omega_2$) 等으로 나타나고 그에 關한 研究가 活潑하며 學會誌 이번 號에도 여러편이 發表되고 있다.

參 考 文 獻

1. A. Papoulis, The Fourier Integral and its Application, Mc Graw Hill, New york, 1962.
2. D. Slepian and H.O. Pollack, " Prolate spherical wave functions-1," Bell System Tech, J., 40, 1961.
3. W. B. Davenport and W. L. Root, Random Signals and Noise, McGraw Hill, New York, 1958.
4. A. Sommerfeld, An Introduction to the Geometry of N Dimensions, Dutton, New York, 1929.
5. W. Feller, An Introduction to Probability theory and it's Application, Vol 1, Wiley, New York, 1957.
6. M. G. Kendall and A. Stuart, The Advanced Theory of Statistics, Vol 2, Hafner, New York, 1961.
7. H. Cramer, Mathematical Methods of Statistics, Princeton Univ, Press, New Jersey, 1946.
8. R. A. Fisher, "Two New Properties in Mathematical Likelihood," Proc. Royal Society, Rondon, 144, 1934.
9. N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series, with Engineering Applications, Tech. Press and Wiley, New York, 1949.
10. H. W. Bode and C. E. Shannon, "A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory," Proc. I RE, Vol 38, 1950.
11. L. A. Zadeh and J. R. Ragazzini, "Extension of Wiener 'a Theory of Prediction," J. Appl., Phys. Vol 21, 1950.
12. S. Darlington, "Linear Least Squares Smoothing and Prediction, with Applications," Bell system Tech, J., Vol 37, 1958.
13. H. Van Trees, Detection, and Estimation and Modulation Theory, part 1 and 2, Wiley, New York, 1971.
14. R. A. Nash and F. B. Tuteur, "The Effect of Uncertainties in the Noise Covariance Matrices on the Maximum Likehood Estimate of a Vector," IEEE Trans, Automatic Cotrol, Vol. 13, 1968.
15. A. Papoulis, Probability, Random Variables, and Stochastic Processes, McGraw Hill New York, 1965.
16. A. P. Sage, "Maximum a posteriori Filtering and Smoothing Algorithms," Intern. J. Cotrol, Vol 11, 1970.
17. F. Hilderbrand, Methods of Applied Math., Prentice Hall, Inc. New Jersey, 1952.
18. T. Kailath, "A General Likelihood Ratio Formula for Random Signals in Gaussian Noise" IEEE Trans. IT Vol 15, 1969.
19. W. V Lovitt, Linear Integral Equations, McGrawhill, New York, 1924.
20. C. Dolph and H. Woodbury, "On the Relacion between Green's Functions and Covariances of Certain Stochastic Process and It's Application to Unbiased Linear Prediction," Trans Amer. Math. Soc., Vol 72, 1952.
21. S Miklin and K Smolitsky, Approximate Methods for Solution of Differential and Integral Equation, American Elsevier, New York, 1967.

22. D. Slepian and T Kadota, "Four Integral Equacion of Detection Theory," SIAM J. Appl. Math., Vol 17, 1969.
23. Ann Souguil et al., "One Procedure perttining to 2D Image Enhancement by Simple Analog Filters, Convention record communication society, KIEE summer, 1979.
24. R.E. Kalman and B. S. Bucy, "New Results in Linear Filtering and Prediction Theory," Trans. ASME, J. Basic Eng. Vol. 83, 1961.
25. R.E. Kalman, "Mathematical Description of Linear Dynamical Systems," SIAM J. Control. Vol 1, 1963.
26. Cho Zang Hee, "Recent Developments in Imaging Systems and Processings 3 D Computerized Tomography," J. of KIEE Dec. 1978.
27. H. Andrews and C. Patterson, "Singular Value Decompositions and Digital Image Prosessings," IEEE ASSP-24, 1976.
28. W. Pratt and F. Davarian, "Fast Computational Tech. for Pseudoinverse and Wiener Image Restoration," IEEE Trans. Computer C-26, 1977.
29. W. Pratt, "Generalized Wiener Filter Computation Tech.," IEEE Trans. Computer C-21, 1972.