<講 座>

스펙트럼 解析

金 治 弘*

8. Fourier 級數展開의 뜻

Fourier 級數를 基礎로하는 스펙트럼解析의 手法은 아주 當然한것 처럼 極히 넓게 받아들이고 있다. 그러나 랜담變動을 여러가지 周波數의 正弦 餘弦函數의 荷重合으로서 表示한다는 것은 數學的으로 본다면 相當히 特殊한 方法이라고 할 수 있을것이다. 或은 이렇게 말하면 좋을것인지? Fourier 展開라는 方法은 가장 一般的인 槪念으로 擴張할 수 있고 또 그것과는 달리 別系統의 分野에서 發展해온 다른 考察法과도 關聯을 부칠수 가 있다. 本章에서는 지금까지 詳細히 說明 해온 스펙트럼 解析을 보다 一般的인 立場에서 다시 살피기로 한다.

1 벡타의 分解와 函數의 展開

어떤 函數를 正弦 餘弦函數로 分解해서 表示한다는 것은 하나의 벡타를 어떤 座標系에 關하여 成分으로 分解한다는 幾何學的 考察法과 等價라는 것이 다음과 같이 쉽게 表示 할 수 있다.

8.1.1 函數와 벡타

지금 (0, 1)區間의 函數 $x(t)$를 생각한다. 이區間을 Δt의 間隔으로 區分하면 $n+1$個의 點의 x의 값

$x(0), x(\Delta t), x(2\Delta t), \cdots\cdots, x(n\Delta t)$

로 이 函數를 離散的으로 表現 할 수 있다. $x(i\Delta t) = x_i$라고 쓰면 函數 $x(t)$는 順序가 붙은 $(n+1)$個의 數의 組

$$(x_0, x_1, x_2, \cdots\cdots, x_i, \cdots\cdots, x_n) \cdots\cdots \quad (8.1)$$

로서 近似的으로 表現 할 수 있다. 이 順序가 붙은 $(n+1)$個의 數値의 組는 곧 $(n+1)$次元 벡타이고 x_{i-1}은 벡타의 第 i 成分이다. 따라서 函數≒벡타 라는 것과 같이 2個의 槪念을 結付를 시킬수가 있다.

萬若 $\Delta t \to 0$ 라하면 $n \to \infty$가 되고 數列(8,1)은 函數 그 自體가 된다. 따라서 函數는 無限次元의 벡타로 생각 할 수가 있다.

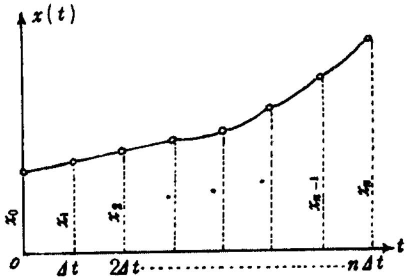


圖 8.1 函水로부터 벡타로

8.1.2 벡타의 直交와 函數의 直交

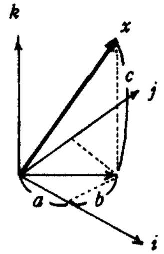
2個의 벡타 **X**, **Y**의 內積이 零

$$\mathbf{X}\cdot\mathbf{Y} = x_1 y_1 + x_2 y_2 + \cdots\cdots + x_n y_n = 0$$

일때 2個의 벡타는 幾何學的으로 直交하고 있다.

2個의 函數 $x(t)$와 $y(t)$를 $(n+1)$點의 離散値로 近似시켜 벡타的 內積을 만든다. 이것에 Δt를 곱해 그 뒤에 $\Delta t \to 0$이라고 하면

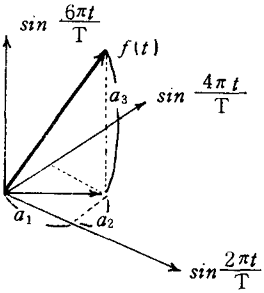
$$[x_0 y_0 + x_1 y_1 + \cdots\cdots + x_n y_n]\Delta t$$
$$= [x(0)y(0) + x(\Delta t)y(\Delta t) + \cdots\cdots + x(n\Delta t) y(n\Delta t)]\Delta t \to \int_0^1 x(t)y(t)\Delta t \quad (\Delta t \to 0) \cdots (8.2)$$



(a) 벡터의 分解

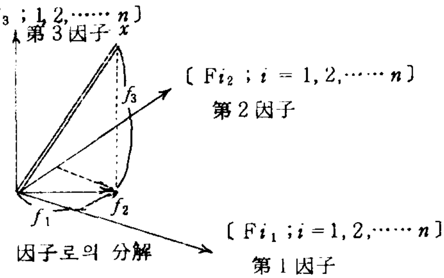
$X = ai + bj + ck$

* 本學會 編輯理事 成均館大學校 工科大學 副敎授



(b) 函數의 Fourier 展開

$$f(f) = a_1\ sin\ \frac{2\pi t}{T} + a_2\ sin \frac{4\pi t}{T} + a_3\ sin \frac{6\pi t}{T}$$



(c) 因子로의 分解

圖 8.2 成分으로의 여러가지 分解法

가 된다. 2個의 函數가 直交한다는것은 따라서 이 2個의 函數의 곱을 函數의 定義域에서 積分했을때 그것이 零이 되는 것이다.

$$\int_0^1 x(t)\ y(t)\ dt \begin{cases} \neq 0 & (x \equiv y) \\ = 0 & (x \neq y) \end{cases} \quad \cdots\cdots\cdots\cdots (8.3)$$

여기서도 벡타와 函數의 槪念이 等價라는 것이 表示되었다. $sin\ 2\pi mt$, $cos\ 2\pi nt$가 위와 같은 性質을 滿足한다는 것은 잘 알려져 있다.

마치 벡타를 成分으로 나누어 表現하는것 처럼 函數는 서로 直交하는 函數列 φ_n 의 荷重合으로 表示 할 수 있다. 但 函數列은 直交性의것 外에 完備(생각하고 있는 函數의 空間을 完全히 cover 한것)여야 한다.

$$x(t) = \sum_{n=0}^{\infty} a_n \varphi_n(t) \quad \cdots\cdots\cdots\cdots (8.4)$$

空間上에 座標系를 任意로 選定 할 수 있으므로 어떤 函數를 Fourier 級數로 表示한다고 定한다는 것은 이쪽에서 마음대로 座標軸을 設定했다는 것에 對應하고 있다. 座標系를 바꾸면 벡타 成分이 變하는것 처럼 函數의 級數展開에는 그外의 여러가지 函數系를 選定 할 수 있다. 但 벡타의 分解의 座標軸이 서로 直交하고 있는 것과 같이 任意의 函數를 展開하는 函數系도 直交하고 있을것 即 이들이 函數的으로 서로 獨立이고 또한 問題로 하고 있는 全函數空間을 表現 할 수 있는 完全系여야 한다. 正弦函數는 이와 같은 性質을 具備한 가장 普遍的인 遜色없는 函數系이다. 여기에 Fourier 級數展開所謂 Fourier 展開에 基底를 둔 스펙트럼解析이 널리 應用되어있는 理由가 있다.

8.2 因子分析(經驗的 直交函數系)展開

萬若 一群의 벡타(x_1, x_2,……, x_r, …)가 한개의 面上에 分布하고 있으면 座標軸中의 2個가 그面에 包含되도록 座標系를 選定하면 벡타의 分解가 簡單히 되는 것 처럼, 랜담變動의 特性에 따라 展開函數系를 選定하는 便이 적은 項數로 랜담變動 $x(t)$를 表示 할 수 있을 것이다(實際Fourie 級數를 써서 計算을 進行하면 級數가 收斂하기 위해서는 20項以上, 要求되는 精度에 따라서는 100項까지 취하지 않으면 안 될때가 있다). 實데아타에 對해서는 이와 같은 分析을 하는것이 主成分分析이든가 因子分析法이다. $0 < t < T$에서의 確率變數 $x(t)$를 離散化하고

$$x(t) = [x\ \ x_2,\ \cdots\cdots,\ x_n]^T \quad \cdots\cdots\cdots\cdots (8.5)$$

로 表示한다. 이때 $x(t)$의 標本마다의 變化는 假想的內在因子 $[F_{j_1}, F_{j_2}, \cdots\cdots, F_{j_m}]$ $(m \leq n,\ j = 1, 2, 3, \cdots\cdots, n)$에 起因한다고 생각하여

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} F_{11} & F_{12} & \cdots & F_{1m} \\ F_{21} & F_{22} & \cdots & F_{2m} \\ \vdots & \vdots & & \vdots \\ F_{n1} & F_{n2} & & F_{nm} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} \quad (8.6)$$

로表示한다. 假想的因子에 걸리는 무게 $[f_1, f_2, \cdots\cdots, f_m]^T$는 x 의 標本마다 變化하는데 行列 $\{\mathbf{F}\}$은 不變이다. 윗式의 左邊의 行列 $\{\mathbf{F}\}$를 因子負荷列(factor loading matrix), 벡타 $\{\mathbf{f}\}$를 因子 評點(factor score)이라고 말한다. 行列$\{\mathbf{F}\}$는 各因子의 基本的인 pattern을 表示하는 것으로서 이것을 因子 pattern 行列이라고 불르고 있다.

式(8.6)을

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = f_1 \begin{bmatrix} F_{11} \\ F_{21} \\ \vdots \\ F_{n1} \end{bmatrix} + f_2 \begin{bmatrix} F_{12} \\ F_{22} \\ \vdots \\ F_{n2} \end{bmatrix} + \cdots\cdots + f_n \begin{bmatrix} F_{1n} \\ F_{2n} \\ \vdots \\ F_{nn} \end{bmatrix} \cdots (8.7)$$

라고 다시 써보면 아는것과 같이 $\mathbf{F}$의 等 i 列

$$\mathbf{F}_i = [F_{1i},\ F_{2i}, \cdots\cdots,\ F_{ni}] \cdots\cdots\cdots\cdots (8.8)$$

은 各各 $x(t)$를 展開하는 座標軸의 成分= 直交函數列로 看做할 수 있다. 또 因子評點은 各座標軸으로의 벡타成分, 或은 正規直交函數系의 係數 a_n 이라고 보아진다. 이러한 意味에서 氣象學에서는 因子分析法 乃至는 主成分分析法을 經驗的直交函數系展開(empirical orthonormal functions) 라고 부르고 있다.

그런데 式(8.6),(8.7)을 簡單히 벡타와 行列로 다음과 같이 表示하자.

$$\boldsymbol{x} = \mathbf{F}\boldsymbol{f} \quad \cdots\cdots (8.9)$$

여기서

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{bmatrix} \quad \cdots\cdots (8.10)$$

$$\mathbf{F} = \begin{bmatrix} F_{11} & F_{12} & \cdots\cdots & F_{1m} \\ F_{21} & F_{22} & & F_{2m} \\ \vdots & \vdots & & \vdots \\ F_{n1} & F_{n2} & & F_{nm} \end{bmatrix}$$

벡타 $\boldsymbol{x}$ 와 因子評點벡타 $\boldsymbol{f}$ 는 標本마다 다른 確率벡타이다. 지금 式(8.9)의 轉置形을 만들어

$$\boldsymbol{x}^T = (\mathbf{F}\boldsymbol{f})^T = \boldsymbol{f}^T\mathbf{F}^T \quad \cdots\cdots (8.11)$$

이것을 式(8.9)의 兩邊에 右로부터 곱해 ensemble 平均을 取한다.

$$E[\boldsymbol{x}\,\boldsymbol{x}^T] = \mathbf{F}\cdot E[\boldsymbol{f}\,\boldsymbol{f}^T]\cdot\mathbf{F}^T \quad \cdots\cdots (8.12)$$

$E[\boldsymbol{x}\,\boldsymbol{x}^T]$ 는 相關行列

$$E[\boldsymbol{x}\,\boldsymbol{x}^T] = \{\overline{x_i, x_j}\} = \{R_{ij}\} = \mathbf{C} \quad (8.13)$$

이다. 各因子는 서로 獨立인 條件부터

$$E[f_i f_j] \begin{cases} = \lambda_i & (i = j) \\ = 0 & (i \neq j) \end{cases} \quad \cdots\cdots (8.14)$$

즉

$$E[\boldsymbol{f}\,\boldsymbol{f}^T] = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots\ \lambda_m \end{bmatrix} \quad \cdots\cdots (8.15)$$

따라서 式(8.12)는 다음式이 된다.

$$\mathbf{C}\mathbf{F}_i = \lambda_i\,\mathbf{F}_i \quad \cdots\cdots (8.16)$$

여기서 $\mathbf{F}_i$ 는 式(8.8)로 定義한 벡타이다.

上式(8.16)은 行列C의 固有値 λ_i 와 固有벡타 $\mathbf{F}_i$ 를 決定하는 問題에 지나지 않는다. 벡타 $\mathbf{F}_i$ ($i = 1, 2, \cdots\cdots, n$)는 相關行列 C가 주어지면 一義的으로 決定되나 이것이 因子負荷行列(因子pattern 行列) F 이다.

例 1. 大氣汚染의 濃度分布에는 風向, 風速에 關聯한 몇個의 pattern 이 있는것은 經驗的으로 느껴지는것이다. 지금 工業市街 地域에 配置된 n點의 monitering station 의 時刻 $k\Delta t$ (Δt : 濃度觀測의 時間間隔, $2\,hr$)에서의 環境濃度를 $x_n(k)$, 그地域의 代表風向風速을 南北 東西成分 U·V로 나누어 이것을 $U(k) = x_{n+1}(k)$, $V(k) = x_{n+2}(k)$ 라고 쓴다. 同一時刻의 이들의 觀測値를 벡타 $\boldsymbol{x}(k)$

$$\boldsymbol{x}(k) = [x_1(k),\ x_2(k),\ \cdots\cdots,\ x_n(k), x_{n+1}(k), x_{n+2}(k)]^T$$

로 表示한다. 지금 벡타 $\boldsymbol{x}$ 가 因子分解되어

$$\boldsymbol{x}(k) = \mathbf{F}\,\boldsymbol{f}(k)$$

로 表示되는 것으로 생각을 해서 因子分析(正準因子分析法에 依함)을 行하여 이것을 圖示하면 濃度分布의 基本 pattern에 關하여 圖 8.3과 같은 結果가 얻어졌다. 圖의 構軸은 各觀測點을 番號順으로 나열하여 最後의 2個는 바람의 南北成分 U와 東西成分 V를 나타낸다. 圖의 縱軸은 因子負荷行列(因子 pattern 行列)의 列 成分 F i, 즉 第i因子의 成分 Fij 의 數値를 記入하고 이들을 基礎로 等濃度線을 記入하면 基本濃度分布를 한층 明白히 表示 할 수 있을 것이다.

第1因子는 거의 南쪽부터의 바람인 경우이고 거의 全域에 걸처 濃度가 높아지는 가장 나타내기 쉬운 基本

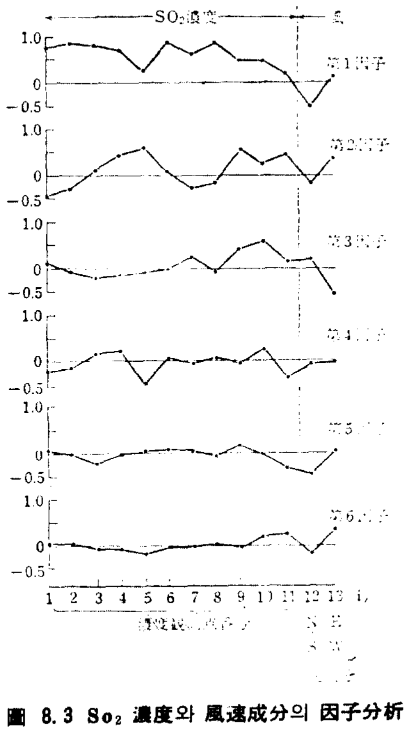


圖 8.3 So₂ 濃度와 風速成分의 因子分析

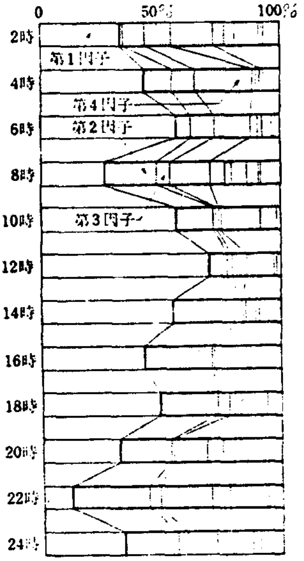


圖 8.4 各因子의 寄與率

形 ; 第2因子는 東쪽 乃至 東南東風이고 一部分의 地點(1, 2, 7, 8)에서는 濃度 가 낮어지는 形 ; 第3因子는 거이 北北西風인 경우의 pattern 에 對應하는것으로 解釋된다. 데이타 x 로부터 逆으로 $f=(F^TF)^{-1}F^Tx$ 에 依해 因子評點을 求하여 各因子의 크기를 計算하여 그 寄與率의 2時間마다의 時間變化로서 表示한 것이 圖 8.4 이다. 第1因子는 各時刻을 通하여 寄與率은 높으나 特히 日中의 寄與率이 높다. 이 工業地帶는 海陸風이 잘 發達된 內海에 南面하고 있고, 그렇다는 것은 南쪽부터의 海風의 影響을 表示한다고 생각 된다.

一般으로 因子分析의 結果에 具體的인 事象解析을 附與하는 것은 困難하다고 말해지고 있지만 本例에서는 벡타 x 의 成分에 汚染濃度外에 바람의 成分을 加하므로서 어느程度 物理的解釋이 可能해지고 있다.

8.3 Karhunen - Loève

이야기를 다시 Fourier 展開로 돌리자. 確率背數 $x(t)$ 의 Fourier 展開에서는 $x(t)$가 非周期函數이면 係數 a_n 은 서로 無相關이 안 된다.

$$E\{a_n a_m^*\} \neq 0 \quad (n \neq m) \cdots\cdots\cdots\cdots (8.17)$$

그래서 Fourier 展開의 槪念을 一般化하고 確率函數의 展開係數가 直交하겠금(서로 無相關이되는) 한것이 Karhunen — Loève 展開이다. 實은 이것은 前述의 因子分析의 確率函數的表現에 지나지 않음을 以下에 表示하자

任意의 確率函數 $x(t)$ 를

$$x(t) = \sum_{n=1}^{\infty} b_n \varphi_n(t) \cdots\cdots\cdots\cdots (8.18)$$

라고 表示한다. 函數列 $\varphi_n(t)$ 는 區間$(-T/2, T/2)$ 에서 直交하는 完全系이고

$$\int_{-T/2}^{T/2} \varphi_n(t)\, \varphi_m^*(t)\, dt \begin{cases} =1 & (m=n) \\ =0 & (m \neq n) \end{cases} \cdots\cdots\cdots (8.19)$$

또한 係數 bn 이 直交할때(즉 無相關의)

$$E\{bn\ bm^*\} = 0 \quad (m \neq n) \cdots\cdots\cdots (8.20)$$

式 (8.18)을 Karhunen — Loève 展開, 直交函數列 $\varphi_n(t)$ 를 Karhunen — Loève 系라고 부른다. Fourier 展開에서는 $x(t)$ 가 非周期的函數이면 係數는 嚴密히 直交하지 않고 $T\to\infty$ 일때 '近似的으로' 無相關이 된다.

$x(t)$ 의 自己相關函數를 $C(\tau)$ 라하면 函數$\varphi_n(t)$는 다음의 積分方程式의 解로서 주어진다.

$$\int_{-T/2}^{T/2} C(t-\tau)\,\varphi_n(\tau)\,d\tau = \lambda_n \varphi_n(t), \quad |t| < \frac{T}{2} \cdots (8.21)$$

上式은 因子分析의 式(8.16)와 等價이다. 式(8.16) 과 (8.21)과를 比較하면 $\varphi_n(t)$가 $\varphi_i(t)$ 가 因子負荷行列의 第 i 列 F_i 에 對應하고 있는것이 쉽게 理解된다.

$$\varphi_i(t) \longleftrightarrow F_i = [F_{1i}\ F_{2i} \cdots\cdots F_{ni}]^T \qquad (8.22)$$

따라서 K.L 系는 因子負荷行列과 等價이다.

$$\{\varphi_1, \varphi_2, \cdots\cdots, \varphi_k\} \longleftrightarrow F \cdots\cdots\cdots\cdots (8.23)$$

K.L 係의 理論은 pattern 認識의 基礎로서 應用되고 있다.

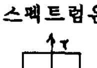
元來 因子分析과 Fourier 展開와는 그 發生과 生長의 過程을 全然 달리 하고 있다. K.L 係와 因子分析과의 關係도 同一하다. 그러나 여기에 表示한 것처럼 이들은 完全히 等價인 內容을 갖는 同一의 槪念임을 理解 했을 것이다.

이와같이 相異한 立場으로 부터 進行해 온 理論이 實은 內容的으로 同一하다는 것은 槪念의 一般化라는 點에서 大端히 興味깊은 일이다.

例 1. 自己相關函數가

$$C(\tau) = \frac{\sin w_0 \tau}{\pi \tau}$$

인 定常確率過程 $x(t)$ 를 생각한다. 添言하면 $x(t)$의 스펙트럼은 $(-w_0, w_0)$ 의 區間에서 一定値가되는



形의 函數이다.

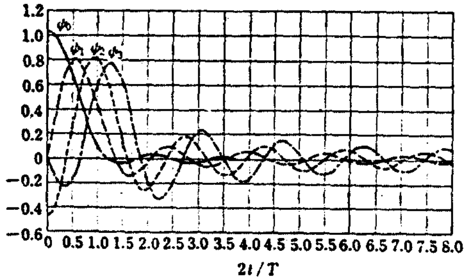


圖 8.5 偏球回轉楕円體波動函數
$\psi_n(t, c)$, $c=4$인 경우

積分方程式, 式(8.21)

$$\int_{-T/2}^{T/2} \frac{\sin w_0(t-\tau)}{\pi(t-\tau)} \varphi(\tau)\, d\tau = \lambda \varphi(t)$$

의 解는 偏球回轉楕円體波動函數(prolate spheroidal wave fanctions)

$$\psi_n(t;\ w_0 T)$$

이다.(圖 8.5)