《綜 合 報 告》

予 測 理 論 の 発 展

近 藤 次 郎*
松 崎 功 保*

は し が き

規則的でない事象，すなわちランダムな事象にかんする予測の問題は物理学の分野では不確定現象の理論的究明に関連して，ずい分古くから取扱われてきた。

一方，時系列の予測は，統計学では中心課題の一つで古くから研究が行われている。これに対して Wiener[1] は1940年代初期に有名な研究を完成した。これは数学的に完成された優美な理論で，その後多くの研究を刺激した。

また工学の分野では，通信における雑音分離や，自動制御の最適設計の問題等と関連して，最近とくに注目されてきている。

したがって，予測については統計学や数学の理論的研究にかんする論文のほかに近頃は通信工学の分野の文献が多いのもこの方面の技術の重要性によるものである。今後いろいろな分野でこの理論の応用によって複雑な確率的現象が解明されるようになれば，その関連する分野も著しく拡大するものと思われる。

以下には通信工学の論文を参照しつつ，数学，統計学の領域での Wiener 流の予測理論の研究を系統的に紹介して，この理論の今後の発展の動向を明らかにし，研究の方向づけを試みようと考える。

この研究は，機械工業連合会の援助によって実施されたことを付記して感謝の意を表したい。

序 章　予測理論の系統図

第1節　予測理論の発展

通信と制御の理論的，実際的問題のうちで予測理論との関連において重要なものは，信号または入力が統計的性質をもっているものである。

そのような問題としては次の三つ，すなわち

(1) 入力から不規則な信号を検出すること

(2) 不規則な信号と不規則な雑音とを分離すること

(3) 不規則な雑音があるときに，パルスとか，正弦波などの形の分った信号を検出すること

などがある。

雑音や信号の数学的解析の理論的発展の上に重要な貢献をしたもののうちには Rice が，1945

*東京大学　昭和39年7月10日受理　経営科学第8巻1号

年[1]および1948年に発表したランダム雑音の確率論的な考察や，雑音と正弦波が重畳した場合の詳細な数学的解析，1940年に今堀，1946年に Gabor が論じた信号の時間長を制限すると周波数を正確に定めることができず，反対に周波数を正確に定めるためには時間長を長くとらなければならないという，いわゆる時間—周波数の不確定性に関する研究等をあげることができよう。これは，信号の時間長を長くとると定常とは考えられなくなり，そのために推定周波数の誤差が多く入り，逆に時間長を短くとるとその間は信号は定常と考えられるが，データが少なすぎて周波数を正確に定められないためである。

Wiener[1]は，戦時研究として，雑音を除去するためと，信号の未来値を予測するための予測器(predictor)の設計を行った。これは結局雑音を予測してそれを入力から取除くことになるのでこれらは数学的にはほとんど同じものである。この研究はまた独立に Kolmogorov[1]によって行われ，同じような結果が得られている。

Lee[1]は予測器を電子回路を用いて実現し信号対雑音比が −20db という雑音の中に埋もれた信号を検出することを工学的に完成した。〔情報理論〕

Wiener はその開拓的な仕事において先に述べた(1)と(2)の問題は，いわゆる Wiener—Hopf 型積分方程式に導びかれることを示した。そして統計的な性質が厳密な意味で定常で rational なスペクトルをもつという特殊な場合ではあるが実用上重要な問題を Fourier 変換の因子分解の方法で解いた。

Wiener のこの基礎的な業績につづいて，多くの拡張や一般化が行われた。

Zadeh & Ragazzini[1]は影響継続時間が有限の場合を解いた。それは Wiener 問題の物理的解釈を行った Bode & Shannon[1]と同時ではあるが独立に問題の簡単な解法を与えるものであった。これは後に近藤[1]によってさらに周期変動を含む場合に拡張された。

Booton[1]は Wieiner—Hopf 方程式の非定常な場合を扱った。

またこのような線に沿ったものであるが少々変った接近方法によるものとしては Darlington [1]の研究がある

標本抽出した信号 (sampled signal) についての理論の拡張については Franklin & Lee[1]の仕事がある。また日本機械工業連合会の「長期計画の理論的研究委員会」[1]~[3]は近藤[2]を中心として，多重時系列の Wiener 問題を離散的に扱い，電子計算機によって実際に計算を実行している。

その他に Wiener—Hopf 方程式の固有函数系展開の手法を用いているものがあり，Davis[1]をはじめとして Shinbrot,[1]Blum,[1]Pugachev,[1]Solodovnikov & Batkov[1],とつづいており，最近，少し異ってはいるがFarmer[1]の研究が注目される。固有函数系にもとづく方法は，もちろん積分方程式の正統的な解法であるが，さらに従来の方法と違って，これらの研究者は非定常な問題を対象にしていることが特徴である。

一方 Kalman は，時間領域でのベクトル・マトリックスをつかっての手法を用いて基礎的な問

序 章　第1節付表　　第 1 表

年代	線形 定常	線形 非定常	非線形	雑音理論一般
1941	Wiener,[1] Kolmogorov, A[1]			
1944				Whittaker. E. & Robinson. G.
1945				Rice, S. O.[1]
1948				Goldman, S.[1]
1949	Wiener, N[1]			Lawson, J. & Uhlenbeck, G
1950				
〃	Bode, H. W & Shannon. C. E.[1]			
〃		Zadeh, L. A. & Ragazzini, J. R[1]		
〃			Singleton, H. E.	
1951				高橋他
〃			White, W. D.	
1952		Booton, R. C.[1]		
〃		Davis, R. C.[1]		
1953				Woodward, P.
〃				Doob, J. L.
〃				G, llespie, A.
〃				Goldman. S.[2]
〃	河田(竜)委員会			
			Zadeh, L. A[1]	
			Zadeh, L. A[2]	
			ZZadeh, L. A[3]	
1953		近藤次郎[1]		
～1954				Tsien, H[1]
〃				van der Ziel, A
1955				(Chessin, P.)[1]※
〃	Franklin, G.[1]			
1956	Laning. J. H & Battin, R. H[1]			
〃				蓑 妻
〃	Lees, A. B.[1]	Carlton, A. G. & Follin, J. W. Jr.		
〃		Pugachev, V.S.[1]		
〃		Solodovnikov. V.V & Batkov, A.M[1]		
〃		Pugachev[2]		

年代	線形		非線形	雑音理論一般
	定常	非定常		
1957		Hanson, J. E.〔1〕		
〃	Steeg, C. W.〔1〕			
1958				関英男
〃				Freeman, J.
〃		Blum, M〔1〕		
〃		Shinbrot, M〔1〕		
〃	Darlington, S.〔1〕			
〃				Davenport, W. & Root. W
〃				Bendat, J.
〃			Wiener, N〔2〕	
〃			Brilliant, M. B	
1959				Kotelnikov, V. A.
〃	Parzen〔1〕			
〃	Pugachev, V. S〔3〕			
〃		Bucy, R. S.〔1〕		
〃			Chung K. K. & Kazda, L. F.	
〃			George, D. A.	
〃		Darlington, S.〔2〕		
1960				Helsirom, c.
〃				Middleton
〃		Kalman, R. E.〔1〕		
〃		Kalman, R. E.〔2〕		
〃			Fuller, A. T	
〃			Mc Fee, R	
〃	Pugachev, V. S.			
1961				
		Kalman, R. E. & Bucy. R. S.		宮脇一男
		Mayne, D. Q		
1962				
1963			Gibson, J. E.	
〃			Flake, R. H	
〃		Farmer, E. D〔1〕		
〃		Pottle, C		
〃	Parzen, E			

年　代	線　形		非線形	雑音理論一般
	定　常	非定常		
〃	近　藤[2]			大泉，他
〃	P委員会[1]			
1964	P委員会[2]			
〃	P委員会[3]			

註) P委員会：長期計画の理論的研究委員会
※　情報理論の文献目録。その後追加されている。

題と取組んでおり，これは制御理論の観測可能性 (observability)，制御可能性 (controlability) の問題と関連して，今後の理論の発展に大きな役割りを果すものと考えられる。
要するに，予測の理論は，入力が定常なときと非定常なとき，連続な場合と離散的な場合とに分類でき，また入力が単純な場合と多重な場合とに大別できる。

以上は全て線形理論であるが，非線形理論もZadeh[1]~[3]，Wiener[2]等により取扱われて以来，近年多くの論文が発表されているが，工学の問題として，例えば予測器の回路を設計するということとになると未だ未解決の問題も残されており，彷徨函数を用いた Wiener の非線形理論では，理論的には厳密であっても級数の収束性に問題があったりするので，ここでは線形理論のみを取上げることにした。

上に述べたことに少しつけ加えて理論の発展を年代順に系統的に整理したものが第1表である。

第2節　本研究の内容と構成

第1節第1表に予測に関連した文献を全てではないが代表的なものはほとんど挙げてある。これらを全部紹介することは到底不可能である上，余り細部に立入ることによりかえって全体を見喪うおそれもあるので，文献をその内容や手法によって大きくいくつかに分類して，そのうちの代表的なものについて少し詳しく紹介するという方針で研究を進めることにする。

すなわち，序章では Wiener以後の概要を述べ第1表に代表的文献の著者を発表年代にしたがって整理してみた。その著者の文献は巻末にある。共通する記号はまとめて序章第2章の表にしたが，個々の細かな記号については，その都度説明を加えてある。

ここで紹介するのははしがきにも断ったように線形理論である。それを内容から大きく分けるととり扱う時系列の統計的な性質から定常時系列の予測理論と非定常時系列の予測理論の二つになる。

文献の多くは連続な時系列を扱っているが，離散的な扱いをしているものもある。

また手法による分類として周波数領域での手法を用いるものと，時間領域での手法を用いるものがある。定常時系列の解析には周波数領域で考察するのが有効でもあり分りやすいが，非定常時系列の場合，また非線形の予測に際しては直接に時間領域で研究するのが便利である。

ここでは，定常，非定常の二つに分類し，第1章で Wiener と，それにつづく定常時系列の理論を紹介し，第2章でいろいろな非定常時系列の理論について述べる。第3章は，最近の予測理論の紹介として Kalman と Farmer の研究について説明するがこの二つは従来の理論と大きく異っている点があるので別の章にしたのである。

それぞれの章のはじめにそこで紹介される文献の占める位置，意義等につき説明する。

最後に本論文に引用した文献の表を付けてあるが，これは著者名のＡＢＣ順に配列し，序章の年表，本文中の参照文献と対照できるようにした。

はしがきでも述べたが予測にかんする研究は Wiener 以後でもはなはだしい数量に達しているので本論文に引用できなかったものについては，Chessin[1]の文献集等を併わせて参照されたい。

記　号

α　α時間だけ先の予測。一般には正，負，0いずれをもとる。

δ　デルタ函数

λ_i　積分方程式の固有値

ρ　信号の二乗平均値

σ　遅差，積分変数

τ　遅差，積分変数

ϕ　相関函数

Φ　パワー・スペクトル

φ_i　積分方程式の固有函数

ω　角周波数

(*)　最適の意味を示す

$E[\ \ ]$, $\overline{[\quad]}$　期待値，平均値

$e(t)$　誤差

$\overline{[e(t)]^2}$　平均二乗誤差

f_i　入力

f_o　出力

$F_i(\omega)$　f_iのフーリエ変換

$F_o(\omega)$　f_oのフーリエ変換

$f(t)$　時系列（$=s(t)+n(t)$）

$f^*(t+\alpha)$　最適予測値

$F(\omega)$　$f(t)$のフーリエ変換

$g(t)$	時系列の規則成分のうち傾向変動
$h(t)$	抵抗雑音，白色雑音
I	特に予測値に関する平均二乗誤差
$K(t)$	予測子，単位インパルス応答
$m(t)$	時系列の不規則成分
$n(t)$	時系列の不規則成分，雑音
$N(\omega)$	$n(t)$ のフーリエ変換
n	信号に含まれる傾向変動の次数
R, r	相関函数
$s(t)$	時系列のうちの信号
$S(\omega)$	信号のフーリエ変換
t	時間
$x(t)$	時系列，標本函数
X	時系列，離散的時系列

文　献

本　文	系統図，文献表
J. Kondo. [1]	Kondo, J.〔1〕
J. Kondo. [2]	Kondo, J.〔2〕

一著者について一編しか引用されないときも〔1〕をつける。

第１章　定常時系列の予測理論

時系列の予測にかんする理論は最初，A. N. Kolmogorov[1] によって研究された。この研究は離散的な時系列の予測にかんするものである。

N.Wiener(1894—1964)[1] は連続的な時系列について Kolmogorov とは独立に研究をはじめ，その結果は M. I. T. の国防研究委員会 (National Defence Research Commitee) の報告として発表されたが，公表は許可されなかった。第二次大戦後，秘密解除となって John Wiley & Sons 書店より出版された。

これは彼のもう１つの著書 Cybernetics と並んで多くの関心を集めた。とくに通信工学，自動制御等の工学上の応用において画期的な理念を創生し，新しい応用の分野を開拓した。しかしこのような工学方面の応用を離れて時系列の研究という理論的な立場から見ても興味の深い問題が含まれている。

Wiener の原著は時系列のエルゴード性の仮定にもとづき，時間平均をとり，フーリエ変換，一般調和解析を利用して問題を数学的に処理している。

Wiener の理論が公表されたあと続々とそれにかんする論文が発表された。その中で Wiener

の理論の基本的な仮定を明確化し，将来の発展への示唆に富む論文として今や古典的なものは1950年に発表された Bode & Shannon[1] の論文である。

彼らは Wiener と Kolmogorov の最も重要な結果を電気回路の理論を使って，数学的な演算に物理的な意味を与えながら直観的に説明したのが特徴である。Wiener の論文はその黄色い表紙と数学的な難しさの由に“The Yellow Peril”(黄禍) と呼ばれていたが，その物理的な意味づけを行うことは数学の問題を実際に解いて物を作らなければならぬ技術者にとっては非常に大切なことであり事実この論文はその後参照されることが多いものである。そこでここでは以下に第1節で，Wiener 理論の概要を紹介するのに Bode & Shannon に従うことにする。なお他にもわかり易い説明が N. Levinson によって為されており，それは Wiener の書物の附録Cにおさめられている。

第2節では，予測器の物理的な実現性を考慮した Bode & Shannon の手法を雑音のない純予測の場合と雑音のある場合とに分けて説明する。

第3節では，定常時系列の予測理論の Bode & Shannon 以来，今日に至る迄の概要を述べるが，いずれも Wiener 理論の主流をなす問題を扱っているものである。これらの中にあって「長期計画の理論的研究委員会」は従来の研究が，応用上の困難さから，実際の計算例に乏しいのに反して，IBM 7090 を用いていくつかの具体的な予測の問題を数値的に解いていることが注目される。

第4節では，これらの Wiener 理論について，その評価函数のとり方などの基本的な仮定や応用についての一般的な問題点を明らかにしている。

第1節に入るに先だち参考のために以下に簡単に N. Wiener の著書[1] の内容を紹介する。原著は，はしがき，第1章・数学的基本事項のまとめ，第2章・単一時系列に関する線形予測器の理論，第3章・単一時系列の線形沪波器，第4章・多重時系列の線形予測器および沪波器，第5章・本書の技術中に含まれる各種の問題，および附録よりなっていて，第2章，第3章が重要な部分をなしており，その思想ははしがきで，Fourier 解析に関する理論は主として第1章で述べてあり，第4章は，第2章，第3章の場合の拡張となっているが，本質的に新しいものは含まれていない。

第1節　Wiener 理論の物理的考察[註1]

(a)　問題の定式化

いわゆる Wiener 問題の基本的な仮定は3つある。すなわち，

(1)　信号と雑音が厳密に定常的であること (エルゴード性)

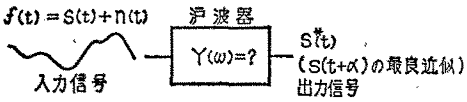
(2)　評価函数は誤差の平均二乗で，それを最小にするものを最適とすること

(3)　予測子またはフィルターの特性が線形であること

である。

Wiener の提起した問題は，信号 $s(t)$ と雑音 $n(t)$ の和として　入力：$f(t)=s(t)+n(t)$ が観

測されたとき，信号 $s(t)$ にできる限り近い値を得るにはどうしたらよいかということである。さらに一般的には $\alpha>0$, $\alpha=0$, $\alpha<0$ のときに $s(t+\alpha)$ の値をできるだけ正確に求める問題となり $\alpha>0$ のときには予測，$\alpha=0$ のときには沪波，$\alpha<0$ のときを平滑ということができる。(第1-1図参照)

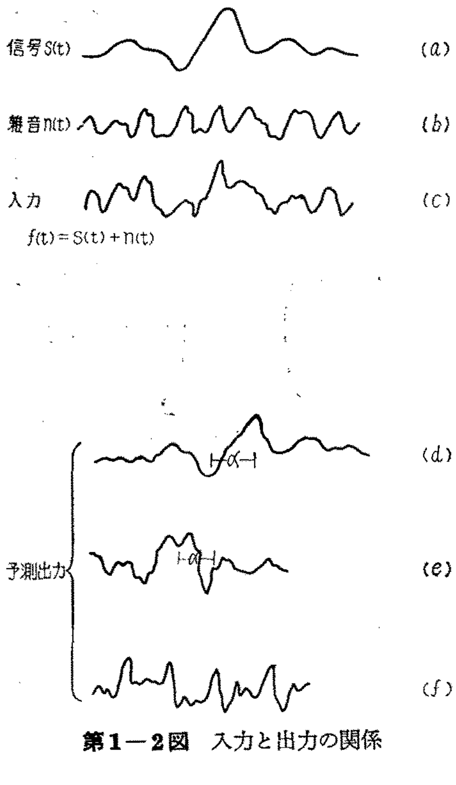


第1－1図 予測，沪波，平滑

またもっと一般的には，希望出力として $s(t)$ の汎函数を求めるということになる。

沪波器の伝達函数を $Y(\omega)$ とすると，$Y(\omega)$ は，集合平均として実際の出力と希望出力との差の2乗平均値をできるだけ小さくするように設計される。いいかえれば沪波器は，$s(t), n(t)$ の集合に対して設計されるものであって，それらの特別な函数にたいして設計されるのではない。ここで取扱う特別の函数 $s(t)$ および $n(t)$ は実際には知ることができないのであるから特別な函数のかわりに函数の集合をとり扱うこの方法は1つの合理的な近似方法である。

研究の目標は 集合平均：$[s^*(t)-s(t+\alpha)]^2$ を最小にすることであるが，時間平均と集合平均とが等しいことはエルゴード性の仮定をおいたことにより保障されている。実際問題としては，$s(t)$, $n(t)$ がそれらの集合の代表的函数であるとき，集合平均が最小であることは，時間平均 $\frac{1}{2T}\int_{-T}^{T}[s^*(t)-s(t+\alpha)]^2\,dt$ についてこの積分を十分長い時間にわたってとったときに，これが最小になることと同じである。



第1－2図 入力と出力の関係

第1－2図には，$s(t), n(t)$ および $s^*(t)$ を図示し，期待しうる概念的な結果を示してある。この図はまた有効な予測は信号間の影響の範囲以外ではできないという結果をも示している。いいかえれば予測は，信号間の影響の知識にもとづく将来への外挿である。第1－2図(d)の出力 $s_1^*(t)$ は α_1 が負ならば，$s(t+\alpha_1)$ の波形にまったく類似し，図(e)では $\alpha_2>0$ でも十分信号間の影響の範囲内にあれば，なお $s(t+\alpha_2)$ の波形に類似していることを示し，図(f)の予測出力 $s_3^*(t)$ は $\alpha_3>0$ で信号間の影響の範囲より大きいとき，細かい波形は $s(t+\alpha_3)$ に無関係であることを示している。

(b) 平均二乗誤差

Tを非常に大きく選べば，平均二乗誤差は

$$\overline{[e(t)]^2}=\frac{1}{2T}\int_{-\infty}^{\infty}[e(t)]^2\,dt$$

$$=\frac{1}{2T}\cdot\frac{1}{2\pi}\int_{-\infty}^{\infty}|E_N(\omega)|^2\,d\omega$$

$$=\int_{-\infty}^{\infty}\Phi_e(\omega)d\omega \tag{1—1}$$

となる。ここに $|E_N(\omega)|^2$ は $e(t)$ のエネルギー・スペクトル，$\Phi_e(\omega)$ は $e(t)$ のパワー・スペクトルを表わす。

ところで，入，出力の関係から

$$F_0(\omega)=Y(\omega)[S(\omega)+N(\omega)] \tag{1—2}$$

$$S_\alpha(\omega)=S(\omega)e^{j\omega\alpha} \tag{1—3}$$

である。$F_0(\omega)$，$S(\omega)$，$N(\omega)$，S_α はそれぞれ $f_o(t)$，$s(t)$，$n(t)$，$s(t+\alpha)$ のフーリエ変換で，$Y(\omega)$ は予測子の伝達函数である。なお $f_o(t)$ は入力 $f_i(t)$ に対する出力を意味する。(1—3)式の関係は T の有限な値に対しては近似的に成り立つだけであるが $T\to\infty$ のときは正しい表式となっている。

したがって，$e(t)$ のフーリエ変換 $E(\omega)$ は

$$E(\omega)=S_\alpha(\omega)-F_0(\omega)$$

$$=S(\omega)e^{j\omega\alpha}-Y(\omega)[S(\omega)+N(\omega)]$$

$$=[e^{j\omega\alpha}-Y(\omega)]S(\omega)-Y(\omega)N(\omega) \tag{1—4}$$

と書ける。

$s(t)$ と $n(t)$ が独立なら，

$$\Phi_e(\omega)=|e^{j\omega\alpha}-Y(\omega)|^2\Phi_s(\omega)+|Y(\omega)|^2N(\omega) \tag{1—5}$$

となるから，これを (1—1) 式に代入すれば，

$$\overline{[e(t)]^2}=\int_{-\infty}^{\infty}[|e^{j\omega\alpha}-Y(\omega)|^2\Phi_s(\omega)+|Y(\omega)|^2\Phi_n(\omega)]\,d\omega \tag{1—6}$$

となる。積分記号の中の第1項は予測子が完全に理想的でないための誤差に対応するものであり，第2項は，信号に重畳された雑音による誤差である。

(1—6) 式が周波数領域での平均二乗誤差の一般表示式である。これより $\overline{[e(t)]^2}$ が $Y(\omega)$，α，Φ_s，Φ_n のみにより，$f(t)$ の形によらないことが分る。

最適予測器は上の誤差を最小にするようにきめられる。

(c) 物理的実現性を考慮しない予測器——Wiener の予測器

(1—6) の右辺の積分を最小にする函数 $Y(\omega)$ を求めよう。それにはまず，$Y(\omega)$ を複素形に

$$Y(\omega)=G(\omega)\,e^{j\beta(\omega)} \tag{1—7}$$

と書こう。ここに $G(\omega)$ は正の実数，$\beta(\omega)$ は実数である。(1—7) を (1—6) に代入すると

$$\overline{[e(t)]^2}=\int_{-\infty}^{\infty}\Big\{[G^2+1-2G\cos(\omega\alpha-\beta)]\Phi_s+G^2\Phi_N\Big\}d\omega \tag{1-8}$$

となる。

$\beta(\omega)$ については，$\cos(\omega\alpha-\beta)$ が最大値をとるときこの積分が最小になる。

したがって直ちに β は求まり

$$\beta=\omega\alpha$$

である。 (1－9)

(1—9) が成り立つとき (1—8) は

$$\overline{[e(t)]^2}=\int_{-\infty}^{\infty}[(C^2+1-2G)\Phi_s+G^2\Phi_n]\,d\omega \tag{1-10}$$

となる。これは次のように変形できる。

$$\overline{[e(t)]^2}=\int_{-\infty}^{\infty}\Big\{[\sqrt{\Phi_s+\Phi_n}\;G-\frac{\Phi_s}{\sqrt{\Phi_s+\Phi_n}}]^2+\frac{\Phi_s\;\;\Phi^n}{\Phi_s+\Phi_n}\Big\}d\omega \tag{1-11}$$

G は実数であるから積分の中の [　　]^2 の項は $\geqq 0$ である。そして G はこの項にのみ入っているから G については

$$\sqrt{\Phi_s+\Phi_n}\,G-\frac{\Phi_s}{\sqrt{\Phi_s+\Phi_n}}=0 \tag{1-12}$$

のとき，すなわち

$$G=\frac{\Phi_s}{\Phi_s+\Phi_n} \tag{1-13}$$

のとき積分は最小となる。

(1—13)，(1－9) を (1—7) に代入すれば，平均二乗誤差を最小にする $Y(\omega)$ は

$$Y\,(\omega)=\frac{\Phi_s}{\Phi_s+\Phi_n}e^{j\omega\alpha} \tag{1-14}$$

で与えられる。

また最小二乗誤差は，

$$\overline{[e(t)]^2}_{min}=\int_{-\infty}^{\infty}\frac{\Phi_s\Phi_n}{\Phi_s+\Phi_n}d\omega \tag{1-15}$$

となる。

(1—14) で表わされた $Y(\omega)$ が物理的に実現可能であるかどうかは，これまでは考えてないが，それを調べるには，$Y(\omega)$ のフーリエ逆変換，すなわち，予測器のインパルス応答 $K(t)$ を求めてみれば明らかである。$K(t)$ は

$$K(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}Y(\omega)e^{j\omega t}\,d\omega$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\Phi_s(\omega)}{\Phi_s(\omega)+\Phi_n(\omega)}\;e^{j\omega(t+\alpha)}\,d\omega \tag{1-16}$$

であるが，$Y(\omega)$ が物理的に実現可能であるためには，$K(t)$が $t<0$ に対して消滅することが必要である。一般には，(1−14) で定められる $Y(\omega)$ はそうでないから，Wiener の予測器は物理的に実現可能な予測器であるとはいえない。

しかしこれは，物理的に実現可能であると否とにかかわらず，すべての複素伝達函数のうちで最小二乗誤差を与えるものである。

(1−14) 式の $Y(\omega)$ が物理的に実現可能でないとき，物理的に実現可能な予測器のうちで，最小二乗誤差を与える予測器は (1−15) 式で示される誤差より大きな誤差を生ずる。

(註)

Wiener のもとの解は Bode & Shannon によって (1—14) のように，周波数領域でではなく，時間領域においてもとめられた。Wiener は変分問題の解として，最小二重誤差に対して，

$$\phi(t+\alpha)-\int_0^{\infty}\phi_s(t-\tau)K(\tau)d\tau=0 \qquad (t>0)$$

を導いた。

また雑音を含む場合は，

$$\phi_s(t+\alpha)-\int_0^{\infty}[\phi_s(t-\tau)+\phi_n(t-\tau)]K(\tau)d\tau=0 \qquad (t>0)$$

を導いた。この積分方程式は，$K(t)$について解かれ，最適予測器のインパルス応答を与える。

ここに $\phi_s(t)$，$\phi_n(t)$ はそれぞれ $s(t)$，$n(t)$ の自己相関函数である。

(註1)

本章は，Bode & Shannon[1] および Goldman[2] (邦訳あり)によるところが大きい。

(註2) 物理的実現可能性

物理的実現可能な回路網は理想的受動回路素子 (L. C. R など) から構成されるものであるが，利得の絶対値をかえる機能のみを持つ理想増幅器を含んでもよい。

この回路網は安定である。

第2節 Bode & Shannon の方法[1]——物理的に実現可能な予測器

(a) 純予測の問題

特殊な場合として，雑音のないときを考える。問題は，$\infty<t\leq0$ の $s(t)$ が分っているとき，$s(t+\alpha)$ を予測することである。解は一般に信号と雑音のパワー・スペクトルのみによるということは分っている。ここでは雑音のない純予測の問題を考えているのであるから，解は信号のパワー・スペクトル：$\Phi_s(\omega)$ のみによることになる。したがって，実際の信号をそれと同じパワー・スペクトルをもつ信号に置き換えることができる。そうすると最適予測の問題は以下に述べるように考えることができる。

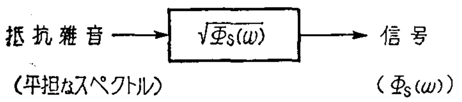
任意のパワー・スペクトルは広帯域の抵抗雑音または白色雑音をゲイン特性が$\sqrt{\Phi_s(\omega)}$ のフィルターを通すことによって得ることができる。広帯域の抵抗雑音のスペクトルは平坦であり，フィルターは求めようとしているパワー・スペクトル $\Phi_s(\omega)$ の平方根の振幅特性を持っていれば$\Phi_s(\omega)$が得られることになる。(第1−3図参照)

このときフィルターの位相特性は物理的に実現可能であるようにとる。ここでは$\sqrt{\Phi_s(\omega)}$ なる

ゲインに対して位相推移が最小になるようにとる。

そうするとフィルターの位相特性は

$$B(\omega_0)=\frac{-\omega_0}{\pi}\int_0^{\infty}\frac{\log\Phi_s(\omega)-\log\Phi_s(\omega_0)}{\omega^2-\omega_0{}^2}\,d\omega \tag{1—17}$$

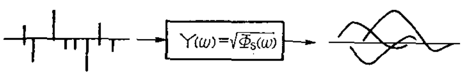


第1－3図

となる。このフィルターの特徴はそれが物理的に実現可能な逆フィルターを持つということである。

$s(t)\ (-\infty<t\leq 0)$ を知っていることは，抵抗雑音 $h(t)\ (-\infty<t\leq 0)$ を知っているのと全く同じことになることが分った。$\sqrt{\Phi_s(\omega)}$ なる特性を持つこのフィルターは，物理的に実現可能な逆フィルターをもつから，$s(t)$ をこの逆フィルターに通することによって $h(t)$ を得ることもできるわけである。

したがって，純予測の問題は，予測値を得るために $h(t)$ にいかなる演算を施したらよいかということになる。



第1－4図　抵抗雑音による応答

抵抗雑音は第1－4図にあるようにインパルスと考えることができ，このインパルスはその振幅の分布が正規分布に従い互いに独立である。

フィルターに入ったインパルスはフィルターのインパルス応答に応じた出力を生ずる。信号はこれらの応答の和である。先にも述べたように知られているのは現在時点迄の $h(t)$ の値であり，現在時点より先の $h(t)$ の値は分らない。

信号の将来値は次の2つの部分よりなっている。すなわち

(1)　現在時点迄に起った各インパルスによる応答の代数和の部分

(2)　現在時点から $t+\alpha$ 迄の間に入る各インパルスによる応答の部分

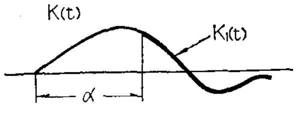
の2つである。

(1)　の部分は，現在迄に起ったインパルスは全て観測されており，インパルス応答は分っているから完全に予測することができる。

しかし(2)の部分は全く予測不能である。それは，現在時点から $t+\alpha$ 迄に起る抵抗雑音は，現在迄に起った抵抗雑音とは独立であるからである。

(1)　の部分の総和，すなわち将来値のうちの予測可能な部分を α 時間先だつ現在時点において得るためには，そのフィルターのインパルス応答を，$Y(\omega)$ のフィルターのインパルス応答の α だけ進めた部分とすればよい。

(第1－5図参照)



第1－5図　α だけ進めたインパルス応答

この新しいフィルターを Y_1 とすれば，Y_1 に入力として $h(t)$ を与えると，その出力信号は α 時間後の同じ入力に対するもとのフィルターの予測可能な部分を与える。

ところで(2)の部分はどうすればよいであろうか。(2)の部分は，各インパルスが独立であることから予測不能ではあるが， その振幅が正規分布に従うので， インパルスの平均値（期待値）も，インパルス応答の平均値（期待値）も 0 であることが分る。

したがって，予測値として最適なものは(1)の部分のみをとったものであるということができる。

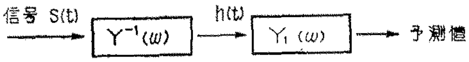
Y_1 をつくる際に抵抗雑音 $h(t)$ を利用できると仮定した。しかし実際には，与えられるものは信号 $s(t)$ である。これから $h(t)$ を得るには，$s(t)$ をその逆フィルター Y^{-1} に通せばよい。

以上の二つのことをまとめれば，与えられた信号に対する演算は

$$Y_1(\omega)Y^{-1}(\omega)$$

とすることが最適であるといえる。

（第1－6図参照）



第1－6図 $s(t)$ に対する演算

純予測の問題の結論を要約すれば次のようになる。

(1) ゲイン特性が $\sqrt{\Phi_s(\omega)}$ であるような位相推移最小のフィルターをつくる。その伝達函数を $Y(\omega)$，インパルス応答を $K(t)$ とする。

(2) インパルス応答が，次のようなフィルターをつくる。

$$K_1(t)=\begin{cases}K(t+\alpha) & t\geqq 0\\ 0 & t<0\end{cases}\qquad(1-18)$$

この伝達函数を $Y_1(\omega)$ とする。

(3) そうすると，最適の予測フィルターは，

$$Y_1(\omega)\cdot Y^{-1}(\omega)$$

なる伝達函数をもったものとなる。

なおこの予測値は， 先の予測不能な部分の平均値（期待値）を 0 にしたものとしたのであるから，評価函数として平均二乗誤差をとるものに適する手法といえる。

この予測の際の誤差は簡単に計算できる。雑音がないので雑音による誤差は考えなくてよく，現在時点から，α 時間先までに起るインパルスによる誤差のみを考えればよいことが分る。これらのインパルスは独立であるから，それぞれのインパルスが引き起す平均二乗誤差は，

$$\begin{aligned}\overline{[e(t)]^2}&=\rho\int_0^{\alpha}K^2(\alpha-t)dt\\ &=\rho\int_0^{\alpha}K^2(t)dt\end{aligned}\qquad(1-19)$$

で与えられる。ここに $\rho=\int\Phi_s(\omega)d\omega$ は信号の平均二乗値である。

同様にして，予測値の平均二乗値は

$$U^2=\rho\int_0^{\infty}K^2(t)dt\qquad(1-20)$$

である。

(b) 雑音のある場合の予測

基本的な考え方は純予測の場合とまったく同じである。(a)では定性的に述べたがここでは少しくわしくみることにする。

平担なスペクトルをもつ出力 $h(t)$ を得るため $f(t)=s(t)+n(t)$ を位相推移が最小であるようなフィルター，すなわち

$$Y^{-1}(\omega)=\frac{1}{[\Phi_s+\Phi_n]^+} \qquad (1-21)$$

に通す。ここで $\Phi_s+\Phi_n=[\Phi_s+\Phi_n]^+[\Phi_s+\Phi_n]^-$ と因子分解できることを仮定する。フィルター Y^{-1} は

$$\int_0^\infty \frac{log\left|\frac{1}{\Phi_s+\Phi_n}\right|}{1+\omega^2}d\omega=\int_0^\infty \frac{-log|\Phi_s+\Phi_n|}{1+\omega^2}d\omega \qquad (1-22)$$

が有限ならば，Paley-Wiener* の判定法によって物理的に実現可能である。

$Y^{-1}(\omega)$ はまた物理的に実現可能なフィルター

$$Y(\omega)=[\Phi_s+\Phi_n]^+$$

をもつから，その出力を知ることと入力を知ることとは，最適予測器を設計するにあたっては等価である。

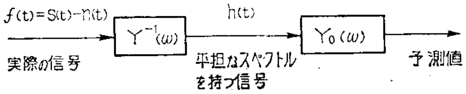
さらに，$s(t+\alpha)$ に近似させるために $Y^{-1}(\omega)$ の出力に施す最適の線形の演算は，入力に対する最適の演算と同じ最適予測値を与える。

物理的な実現可能性を考慮しないならば，$s(t)+n(t)$ を入力として $s(t+\alpha)$ の最適予測値を与える予測器は (1—14) に示したような伝達函数をもつ。

したがってそのような予測値は，予測器が実現可能ではなくても，解析的には $Y^{-1}(\omega)$ の出力を $Y_0(\omega)$ を通すことによって得ることができる。ここに $Y_0(\omega)$ は

$$Y_0(\omega)=\frac{\Phi_s}{([\Phi_s+\Phi_n]^+)^{-1}}e^{j\omega\alpha} \qquad (1-23)$$

によって定義される。(第1—7図参照)



第1—7図 物理的実現性を考慮しない予測器

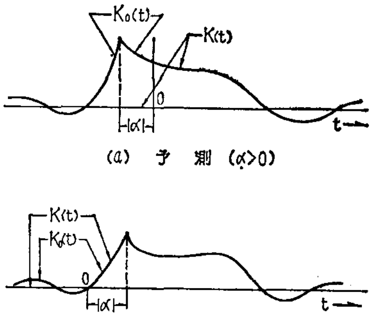
(脚注*) Paley-Wiener の判定法の観点に立った物理的実現可能性のすぐれた議論が H, Wallman の Vol. 18 Radiation Laboratories Series の" *Vacuam Tube Amplifier* の Appendix (A)にのべられている. Paley-Wiener の判定法の導き方は R.E.C.Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, American Math. Soc. Colloquium Publication. Vol. 19, 1934 にある.

また $Y_0(\omega)$ に対するインパルス応答は 第1－8図に示した。$Y_0(\omega)$は物理的に実現可能でないから，正負の無限大にまで尾をひいて拡がるものと考えられる。

$s(t+\alpha)$にたいする最適予測器

$$Y^{-1}(\omega)\,Y_0(\omega)$$

は，同じパワー・スペクトルをもつ他の任意の信号 $s_1(t+\alpha)$に対する最適予測器にもなっている。ただしその場合，$s_1(t)$ に重畳されている雑音は $n(t)$と同じパワー・スペクトルをもつ雑音 $n_1(t)$であるとする。

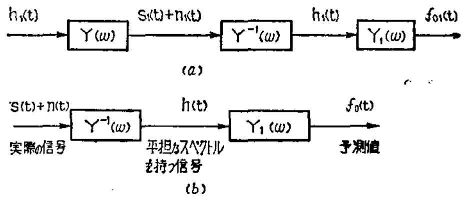


(a) 予 測 ($\alpha>0$)
(b) 平 滑 ($\alpha\leq 0$)

第1－8図 $Y_0(\omega)$のインパルス応答

抵抗雑音からフィルター $Y(\omega)$を通して得られる信号を，$s_1(t)+n_1(t)$とすれば，抵抗雑音$h_1(t)$は $s_1(t)+n_1(t)$をその逆フィルター $Y^{-1}(\omega)$ に通すことによって得ることができる。(第1－9図(a) 参照)

さてつぎに，物理的に実現可能な最適予測器 $Y_1(\omega)$をもとめることを考えよう。ここで最適というのは平均二乗誤差を最小にするものである。



第1－9図

抵抗雑音は，純予測の場合と同じように，互いに独立なパルスと考えてよい。現在時点をt_0とすると，$h_1(t)$が入力であるとき，$Y_0(\omega)$の出力信号 $f_{01}(t)$ は，t_0 以前に生じた各インパルスによるインパルス応答の代数和である。すなわち，インパルス応答を$K_1(\tau)$とすれば，$f_{01}(t_0)$ は

$$f_{01}(t_0)=\int_{-\infty}^{t_0} h_1(t)K_1(t_0-t)dt$$

$$=\int_0^{\infty} h_1(t_0-\tau)K_1(\tau)d\tau \qquad (1\text{—}24)$$

$$(\tau=t_0-t)$$

となる。

ところで，$Y_1(\omega)$ のかわりに，解析的ではあるが，物理的に実現可能でない最適フィルター $Y_0(\omega)$を用いると出力は，

$$\begin{bmatrix}h_1(t)\text{ による }Y_0(\omega)\\ \text{の出力，}(t=t_0)\end{bmatrix}=\int_{-\infty}^{\infty} h_1(t)K_0(t_0-t)d\tau$$

$$=\int_{-\infty}^{\infty} h_1(t_0-\tau)K_0(\tau)d\tau$$

$$=\int_{-\infty}^{0} h_1(t_0-\tau)K_0(\tau)d\tau+\int_0^{\infty} h_1(t_0-\tau)K_0(\tau)d\tau \qquad (1\text{—}25)$$

上の場合は，$Y_0(\omega)$は物理的に実現可能でないから，(1—25) の積分区別は $-\infty$ から $+\infty$ に

亘っている。それは，そのインパルス応答 $K_0(t)$が，過去と未来の両方に拡って尾をひいているからである。

(1—25) は，$s_1(t_0+\alpha)$に対する解析的に可能な，最適値を与える。したがって (1—24) による予測値の誤差より (1—25) の誤差の方が少い。

さて，$Y_0(\omega)$の全出力は次の2つの部分に分けられる。

第1の部分

$$\int_0^\infty h_1(t-\tau)K_0(\tau)d\tau \tag{1—26}$$

は現在時点 t_0 以前に発生した $h_1(t)$ にたいするものであり，

第2の部分

$$\int_{-\infty}^0 h_1(t-\tau)K_0(\tau)d\tau \tag{1—27}$$

は，t_0 以後に発生した $h_1(t)$ に対する応答の和である。

物理的に実現可能なフィルター $Y_1(\omega)$は現在時点以前に発生した $h_1(t)$に対するインパルス応答のみから求めることができる。したがって（1—25）は，物理的に実現可能な最適予測器にかんする重要な情報を与えてくれる。$h_1(t)$ は互いに独立なインパルスよりなると考えてよいし，現在時点以後と以前とについてフィルターのインパルス応答は同じと考えてよいから，物理的に実現可能な場合には（1—25）式は，最適予測器のインパルス応答を与えると考えることができる。そしてこのとき，純予測のときと同じ考え方によって，t_0 以後に発生するインパルスに対する重みは0でなければならない。こうして $K_1(t)$は次のように決定される。

$Y_1(\omega)$は物理的に実現可能であるから，

$$K_1(t)=\begin{cases}K_0(t) & t\geq 0\\ 0 & t<0\end{cases} \tag{1—28}$$

である。$Y_1(\omega)$は $K_1(t)$のフーリエ変換であるから

$$\begin{aligned}Y_1(\omega)&=\int_{-\infty}^{\infty}K_1(t)\,e^{-j\omega t}\,dt\\ &=\int_0^\infty K_0(t)\,e^{-j\omega t}\,dt\end{aligned} \tag{1—29}$$

である。ただし，（1—23）によって

$$\begin{aligned}K_0(t)&=\frac{1}{2\pi}\int_{-\infty}^{\infty}Y_0(\omega)\,e^{j\omega t}\,d\omega\\ &=\frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\Phi_s}{[\Phi_s+\Phi_n]^-}\,e^{j\omega(t+\alpha)}\,d\omega\end{aligned} \tag{1—30}$$

である。

したがって $s_1(t+\alpha)$ を予測するための物理的に実現可能な最適予測器は,

$$Y^{-1}(\omega)\cdot Y_1(\omega)$$

となる。

$Y_1(\omega)$ は（1—29）で $Y^{-1}(\omega)$ は（1—21）で与えられる。

これは $s_1(t+\alpha)$ の最適予測器であるが, $s(t+\alpha)$ の最適予測器でもあることはもちろんである。

予測誤差の平均二乗値は,

$$\overline{[e(t)]^2}_{min}=\frac{1}{2\pi}\int_{-\infty}^{\infty}\Phi_s(\omega)d\omega-\frac{1}{2\pi}\int_{-\infty}^{\infty}(\Phi_s(\omega)+\Phi_n(\omega))|\,Y^{-1}(\omega)\cdot Y_1(\omega)\,|^2\,d\omega$$

で与えられる。（1—31）

第3節　その後の発展——「長期計画の理論的研究委員会」，他

Bode & Shannon の方法は, 後に述べる Zadeh & Ragazzini[1]——近藤[1]の問題や, その他工学的に興味のある一般的な問題に応用することが難しく, もしそれらに応用すると, この手法のもっている有理な点をそこねてしまう。

さらに, Bode & Shannon は Wiener 問題の解法に物理的な解釈を与え予測器設計上の見透しをよくはしたが, 問題を数値的に解くことがこの方法によっては必ずしも容易になったわけではない。

Darlington[1] は Bode & Shannon と同じように電気回路の理論を用いて, 少し異った方法によってより一般的な問題を容易な計算によって扱えるようにした。連続で, 統計的に定常な時系列を対象にしていて, 周波数領域で処理している。数学的な複雑さを避けるために, 統計的な集合のパワー・スペクトルが rational な場合（有理式で表わされる場合）に限定した。実際のスペクトルは rational なスペクトルによって十分近似できるし, 種々の工学上の応用問題に対しても, 十分に一般的といえる。影響継続時間は有限でも半無限でもよく, それに伴う制約条件はあってもなくてもよい。rational なスペクトルを持つという仮定を非定常な時系列に直接あてはめることはできないが, それと非常によく似た仮定で置きかえることができる。そのとき非定常な問題の解法は, 定常な系について開発されたものとまったく似た形になる。それは工学的な目的のためにはあまりにも複雑であるが興味ある問題ではある。(〔1〕の文献では簡単にしか触れてないが Darlington[2] は後に非定常な予測理論を発表している。)一般的な手法をロケットの弾道の予測のような特殊な問題について説明しているが, それは説明の都合と工学的な興味のゆえであり, これは一般の予測問題に応用することが可能である。

Zadeh & Ragazzini[1] も rational なスペクトルの仮定を置いているが, より簡単な周波数領域においての取扱いをせずに時間領域において解析している。

Laning & Battin[1] も同じく rational なスペクトルの仮定の下に時間領域で解析を行っている。

1953年に河田部会[1]は定常的な Wiener 問題の拡張を試みている。

また近藤[2]は，多重時系列の理論を発表した後， 1963年の TIMS (TOKYO) 大会でその後展開された理論を発表した。次にそれを簡単に紹介する。(「長期計画の理論的研究委員会」[1])

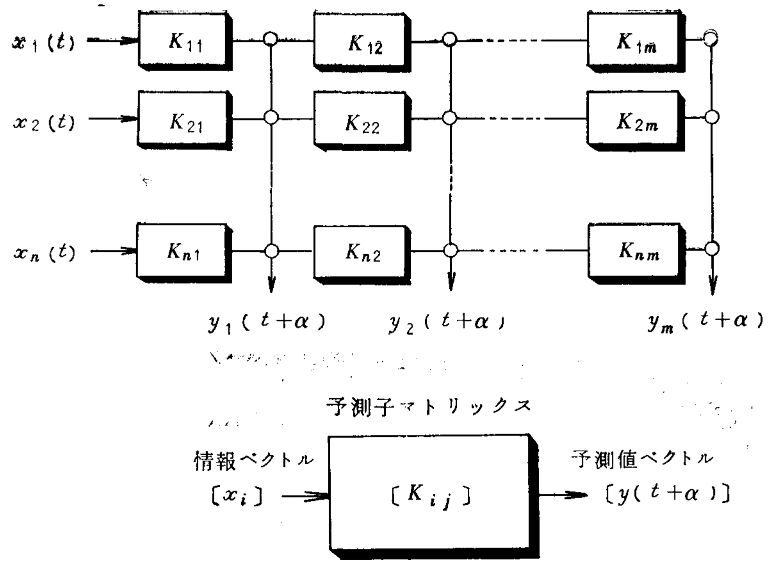
多重時系列の予測子 $K_{ij}(t)$ の満足すべき条件は，平均二乗誤差を最小にするという変分問題の解として，単一時系列と同じように，

$$\phi_{ij}(\alpha+\tau)-\sum_{h=1}^{n}\int_{0}^{\infty}\phi_{ik}(\tau-\sigma)d\,K_{hj}(\sigma)=0 \qquad (1\text{—}32)$$

$$\left(\begin{array}{l} i=1,2,\cdots\cdots,\ n \\ j=1,2,\cdots\cdots,\ m \end{array}\right)$$

を得た。ϕ_{ij}, ϕ_{ih} は相互相関函数である。

(1—32)は n_1の未知函数 $k_{ij}, \cdots\cdots, k_{nj}$ についての連立方程式であり，$j=1,2,\cdots\cdots,\ m$として順に解くと，予測子マトリックス $[K_{ij}]$ が決定できる。この解法は同文献の附録に収められている。(第1—10図参照)



第1—10図

また他の時系列の予測値を用いて，それと相関のある時系列を予測する方法も得ている。

離散的な場合には (1—32) に対応する表式は，次のようになる。

$$\sum_{j=1}^{n}\sum_{r=0}^{s_j}\left(\sum_{r=0}^{m}X_{ji-r}X_{j'i-r'}\right)K_{jr}=\sum_{j=1}^{m}X_{j'\,i-r'}\,Y_{i+\alpha} \qquad (1\text{—}33)$$

$$\left(\begin{array}{ll} i>r & r'=0,\ 1,\ 2,\ \cdots\cdots,\ s_j \\ i'>r' & j'=1,\ 2,\cdots\cdots\cdots,\ n \end{array}\right)$$

のようになるが，これは予測係数にかんする連立方程式であって，これを解けば係数がきまる。

1963年12月には，(1—33) 式に対応する計算プログラムを開発し，株価指数，金利等の予測計算を IBM 7090 を用いて行っている。また影響継続時間決定の問題を扱っている。

Wiener の予測理論の実際上の困難は，時系列の全経歴を考えねばならないことである。

離散時系列

$$\cdots\cdots X_{-2},\ X_{-1},\ X_0,\ X_1,\ X_2\cdots\cdots$$

について，

$$e^2 \equiv \frac{1}{m}\sum_{i=1}^{m}(X_{i+\alpha}-\sum_{r=1}^{m}K_r X_{i-r})^2 \tag{1—34}$$

を最小にするように K_r を決定する。つぎにこの K_r を用い，時系列の将来値と

$$\widehat{X}_{i+\alpha}=\sum_{r=1}^{m}K_r X_{i-r} \tag{1—35}$$

として予測するとき，予測値 $X_{i+\alpha}$ との差を S 個の予測値について調べその二乗和

$$e_\alpha{}^2 = \frac{1}{S}\sum_{\alpha=1}^{S}(X_{i+\alpha}-\sum_{r=1}^{m}K_r X_{i-r})^2 \tag{1—36}$$

を調べる。m を変え $e_\alpha{}^2$ と図示すると，第1—11図のようになれば，$m=m^*$ としたときが予測の誤差を最小にすることがわかる。m^* は条件

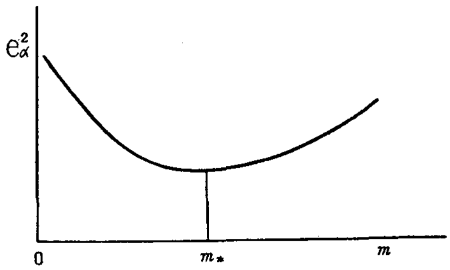
$$\frac{\partial e_\alpha{}^2}{\partial m}=0 \tag{1—37}$$

によって決定できるが，それを解析的な表式として求めることは難しいので例で示している。

同じ報告書で，予測子 $K(t)$ の形を

$$\sum_i C_i t^i\, exp(-\lambda_i t) \tag{1—38}$$

と仮定した場合の母数 C_i, λ_i を最小二乗的に決定する連立方程式も導いている。(「長期計画の理論的研究委員会」[2])



第1—11図　最適継続時間の決定
(予測の誤差と継続時間の関係)

さらに，1964年4月には，上記の手法および，スーツ・モデルを日本の場合に応用して，自動車の需要予測を行っている。(「長期計画の理論的研究委員会」[3])

その後は季節変動等，周期的変動を含む場合の予測の問題に取り組んでいる。(「長期計画の理論的研究委員会」[4])

第4節　予測についての一般的問題点

Wiener 問題の基本的な仮定に対して，Bode & Shannon[1] や Tien[1] が明確でするどい批評を与えているのでその重要な部分を原文から引用する。

Bode & Shannon は次のように云っている。

数学を応用して得た結論は，最初の仮定と同じ程度の信頼性しかもたない。Wiener 理論は，それを適用しようとする現象が，基本的仮定を満足するかどうかを判定することが困難であるために，誤って使われる恐れがある。その仮定は3つあるがその1つ1つについて慎重に考えることが必要である。

信号と雑音が定常であるという仮定は，３つの仮定のうちでおそらく最も無難なものであろう。それはこの仮定が満たされないときは それを知ることが 比較的容易で あるからである。パワー・スペクトル $S(\omega)$ と $N(\omega)$ とから時系列の統計的性質が分るから，それが時間的に変化するかどうかを調べればよい。その変化が他の時間定数に比べて遅ければ，そのような非定常問題は，準定常理論を基礎にして解くことができる。そのとき，線形予測器は伝達函数が各時点で最適になるように設計されよう。

最小二乗の仮定は，その妥当性を正当化することが，さらに困難である。平均二乗誤差を最小にしようとするときには，誤差のうちの大きさものに対して特別な注意を払っていることになる。したがって予測器としては，比較的小さなしかもしばしば生ずる誤差には注意せずにたまにしか起らない大きな誤差をできるだけ小さくするものがえらばれる。しかし多くの実際的な問題では，しばしば起る小さな誤差をできるだけ小さくすることが要求される。また，未来の現象が正規分布をするときには，最も可能な現象は，平均二乗誤差が最小になるようなものであるから，この判定法が最適であるといえる。しかしながら，その分布が，偏奇形 (lopside) であったりするときには，実際に問題が生ずる。

一例として，明日が晴れるかどうかを予測する問題を考えよう。一般には晴れの日の方が多いことおよび雨の日に対する負の雨量の日がないことから考えて，雨量の分布は偏奇形になると考えられる。そのとき平均二乗誤差を最小にする予測で与えられる平均的な予測値はきり雨の日を表わすことになろう。しかしピクニックを計画している人には，その予測値はなんの価値もないであろう。かれは天気が実際に晴れとなる確率に関心があるのであって，ピクニックを雨のためにとりやめねばならないならば，雨量の多少は比較的ささいなことだからである。

第３の仮定，すなわち線形という仮定は予測にもちいられる演算または装置に対する制限である。この仮定の数学的な理由は明らかである。線形の問題はつねに非線形な問題にくらべてはるかに簡単であるからである。応用に際しては，線形の仮定は以下の理由のどれかによって正当とされる。

(1) 正規分布をなす時系列では，線形の最小二乗予測器は，他のいかなる予測器よりも最適である。

(2) 線形予測器は，演算，機構が簡単である。線形予測器を構成することは容易でありかつ広範囲な関連理論があるが，非線形系にはこれに対する理論は少ない。

(3) ほかによい方法がないという理由だけで線形の理論が用いられる。不完全な解でも何もないよりはよいからである。

予測器を線形としたための欠点は何であろうか。予測で非線形理論が重要であることは，明日の天気を予報する問題をふり返ってみれば明らかである。

将来の事象を予測するのに，ある期間にわたっての事象のパターンの方が，個々の現象より重要であることがある。たとえば寒冷前線または温暖前線の通過の際の現象の系列がそのようなも

のの1例である。

さらに，ある現象の効果はその現象の強さに非常に大きく影響されることがある。気圧計が下がることは，悪天候になることを意味するが，同じ時間に2倍の早さで気圧計が下がることは天候が2倍悪化することを意味するよりは，恐らく台風が来るかもしれないことを意味するであろう。

また，線形ということと，最小二乗誤差ということが全ての問題については，両立し得ないことに注意しよう。線形の仮定を無視しても，最小二乗予測値は，どんな場合にも将来の分布の平均値を選ぶ。しかしながらこの分布の平均値は，一般には過去の時系列の非線形な函数であるからである。

次に Tien の云っていることは次のようなことである。同じ相関函数すなわち同じパワー・スペクトルを持った2つのランダムな信号については，その予測器は同一のものとなることを特に注意し，これは設計のための手法としてはある種の粗雑さを示すものであると述べている。もしパワー・スペクトルだけでなく，その他にも信号について知るところがあれば，その情報をつかって，さらによい予測器をつくることができようし，同じパワー・スペクトルを持つ信号でも，もともと違った信号ならばそれを区別することもできるだろう。そのためには，最近発達しつつある情報理論が役に立ってあろうし，そういう手法で問題を扱っている例も挙げている。

第2章　非定常時系列の予測理論

Wiener 理論の非定常な場合は，1950年に Zadeh & Ragazzini[1] によって取扱われて以来，1952年に Booton[1]，1953～4年には近藤[1]によって更に拡張された。

Booton は

$$\phi_{id}(t-\tau_1, t)=\int_0^{\infty}\phi_{ii}(t-\tau_1, t-\tau_2)K(\tau_2, t)d\tau_2$$

$$0<\tau_1<\infty$$

という Wiener-Hopf 方程式と似た形の積分方程式を得た。ここで ϕ_{id} は入力と理想出力との相互相関函数，ϕ_{ii} は入力の自己相関函数である。この積分方程式は一般には解くことができないが，特殊な場合に解法を与えた例はある。

Zadeh & Ragazzini の理論は本質的には Wiener 理論の一般化であるが次の点で異っている。

(1) 与えられた時系列が次の2つの部分よりなっていると考える。

(a) 時間のランダムでない函数で，n 次以下の多項式で表わされる部分。ただし n 次ということは分っていても，その具体的な形は分らない。

(b) 相関函数によって統計的に与えられる定常ではあるが不規則な部分。

(2) 予測器の重み函数，すなわちインパルス応答は，$0\leq t\leq T$ の範囲外で0となる。

(2) 番目は Wiener 理論では T は無限大となっていたが，Zadeh & Ragazzini はそれに反して影響継続時間を有限としたことは注目される。この理論は，離散的時系列にも拡張できるから，Wiener のものより実際的な問題に広く応用できる。

Wiener 理論と同じように，最適予測子の決定の問題は，変形 Wiener-Hopf 型積分方程式に帰する。その方程式の解法が与えられているが，それは特殊な場合として Wiener 理論を含んでいるので，Wiener 問題の解法にも有効であり，かつ計算の点で有利になっている。

近藤[1]は Zadeh & Ragazzini の理論を，週期函数を含む場合に拡張したものである。

時系列 $f(t)$ は規則変動 $g(t)$，$p(t)$ および母集団に固有の偶然変動 $m(t)$，および偶然誤差 $n(t)$ よりなっている，すなわち

$$f(t)=g(t)+p(t)+m(t)+n(t)$$

とする。

規則変動のうち $g(t)$ は傾向変動で，t の多項式で表わされ，その次数は n を越えない。$p(t)$ は週期的変動でその最大週期が T_0 であり，m項の Fourier 正弦級数で表わされる。

その予測子（インパルス応答）$K(t)$ は　$0\leq t\leq T$ の外で 0 になる。このとき T は T_0 の整数倍としておいてよいから，週期変動 $p(t)$ が週期 T の正弦級数に展開されて

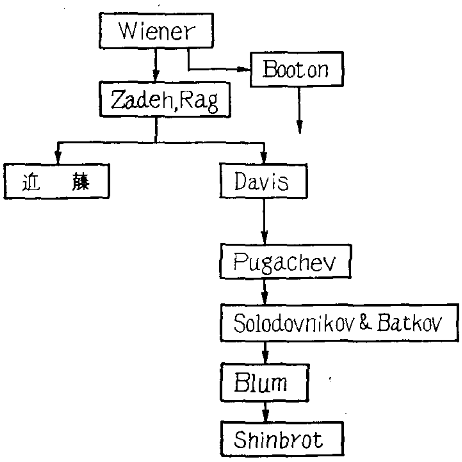
$$p(t)=\sum_{1}^{m}a_i\ sin\frac{i\pi t}{T}$$

と書けるとする。

Wiener の理論は $g(t)\equiv 0$，$p(t)\equiv 0$ の場合である。また Zadeh & Ragazzini は $g(t)$ に対する仮定は上と同じで $p(t)\equiv 0$ の場合となる。注意すべきことは，$g(t)$ や $p(t)$ はその形が確定しているのではなく，$g(t)$ については，単に n 次を越えぬ多項式であること，$p(t)$ については，周期 T の m 項以下の Fourier 正弦級数で表わされることが仮定されている。したがって上の式では，すべての係数 a_i を知る必要はない。

一方 1952 年に Davis[1]が Wiener-Hopf 方程式を固有函数系に展開する方法で，非定常問題を取扱った。それ以後同じような手法で，1956年に Pugachev[1]同じく Solodovnikov & Batkov,[1]，1958年に Blum[1] 同じく Shinbrot[1] 等がそれぞれ理論を展開した。

Davis の扱っている具体的な問題は Zadeh & Ragazzini の近藤とは別の意味での拡張である。Davis は従来のものより さらに 複雑な 場合として，信号，雑音の不規則成分が，非定常な問題を扱っている。その上信号と雑音の間に相関のある



第2－1図　非定常時系列の予測理論

一般的な場合を論じているので，多くの仮定をおいて数学的に処理している。

ここでは，Wiener, Zadeh & Ragazzini をその特殊な場合として含む近藤のものと，固有函数展開による手法を用いたものの代表として Davis のものを取上げることにする。(第2—1図参照)

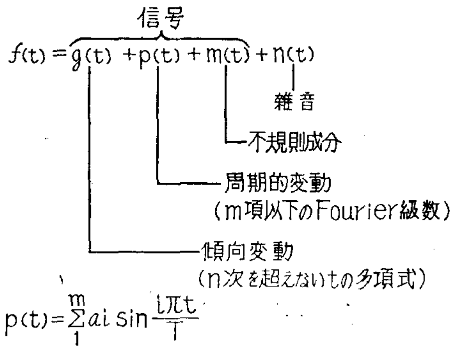
第1節　近藤の理論

時系列

$$f(t)=g(t)+p(t)+m(t)+n(t)$$

において，$g(t)$ は n 次の多項式で表わされる傾向変動，$p(t)$ は m 項以下の Fourier 正弦級数で表わされる周期変動，$m(t)$ は母集団に固有の偶然変動で $n(t)$ は雑音である。

$m(t)$ と $n(t)$ とは時間の定常函数で，有界であり，自己相関函数 $\phi_m(\tau)$，$\phi_n(\tau)$ をもち，それらは連続で Fourier 変換が存在するものと仮定する。さらに $m(t)$ と $n(t)$ とは互いに独立で，その時間平均はともに 0 であるとする。互いに独立の仮定は計算の便利のためであって本質的なものではないが，実際の場合にはこのような仮定が成立することが多い。(第2—2図参照)



第2—2図　(近藤の理論)

予測の問題は，時系列 $f(t)$ の $(t-\tau, t)$ の値を知って将来値 $f(t+\alpha)$ を推定しようということである。

予測子 $K(\tau)$ は，

$$I\,[K]\equiv\lim_{h\to\infty}\frac{1}{2h}\int_{-h}^{h}\{f(t+\alpha)-\int_{0}^{T}f(t-\tau)K(\tau)d\tau\}^2 d\tau \qquad (2-1)$$

が極小値をとり，かつ条件

$$\lim_{h\to\infty}\frac{1}{2h}\int_{-h}^{h}\{f(t+\alpha)-\int_{0}^{T}f(t-\tau)K(\tau)d\tau\}dt=0 \qquad (2-2)$$

を満足するように選ぶのである。

したがってこれは条件つき変分問題となる。この変分問題の解は $K(\tau)$ に関するある積分方程式に帰することを証明しよう。

積　分

$$\int_{0}^{T}g(t-\tau)K(\tau)d\tau \qquad (2-3)$$

において，$g(t-\tau)$ は Taylor 展開により

$$g(t-\tau)=g(t)-\tau g^{(1)}(t)+\frac{\tau^2}{2!}g^{(2)}(t)+\cdots+(-1)\,n\frac{\tau^n}{n!}g^{(n)}(t) \qquad (2-4)$$

と書けるから

$$\int_0^T g(t-\tau)K(\tau)d\tau=\mu_0 g(t)-\mu_1 g^{(1)}(t)+\frac{\mu_2}{2!}g^{(2)}(t)+\cdots\cdots+(-1)n\frac{\mu_n}{n!}g^{(n)}(t) \quad (2-5)$$

となる。ここで $g^{(i)}$ は $g(t)$の第 i 次の導函数で μ_i は次のような第 i 次のモーメント

$$\mu_i=\int_0^T \tau^i K(\tau)d\tau, \quad i=0,\ 1,\ 2, \cdots\cdots,\ n \quad (2-6)$$

である。

また $p(t-\tau)$は

$$p(t-\tau)=\sum a_i \sin\frac{i\pi}{T}(t-\tau)=\sum a_i(\cos\frac{i\pi\tau}{T}\ \sin\frac{i\pi t}{T}-\sin\frac{i\pi t}{T}\cos\frac{i\pi t}{T}) \quad (2-7)$$

となるから，

$$\int_0^T p(t-\tau)K(\tau)d\tau=\sum_{i=1}^m a_i(Kc(i)\sin\frac{i\pi t}{T}-Ks(i)\cos\frac{i\pi t}{T}) \quad (2-8)$$

となる。ここで $Kc(i)$, $Ks(i)$は $K(t)$ の第 i 次で有限 Fourier 余弦変換，および正弦変換で，すなわち

$$\left.\begin{aligned} Kc(i)&=\int_0^T K(\tau)\cos\frac{i\pi\tau}{T}d\tau=\frac{T}{\pi}\int_0^\pi K_1(\xi)\cos i\xi d\xi \\ Ks(i)&=\int_0^T K(\tau)\sin\frac{i\pi\tau}{T}d\tau=\frac{T}{\pi}\int_0^\pi K_1(\xi)\sin i\xi d\xi \end{aligned}\right\} \quad (2-9)$$

である。

したがって

$$\begin{aligned}\int_0^T f(t-\tau)K(\tau)d\tau&=\mu_0 g(t)-\mu_1 g^{(1)}(t)+\frac{\mu^2}{2!}g^{(2)}(t)+\cdots\cdots \\ &+(-1)n\frac{\mu_n}{n!}g^{(n)}(t)+\sum_{i=1}^m a_i(Kc(i)\sin\frac{i\pi t}{T}-Ks(i)\cos\frac{i\pi t}{T}) \\ &+\int_0^T m(t-\tau)K(\tau)d\tau+\int_0^T n(t-\tau)K(\tau)d\tau \end{aligned} \quad (2-10)$$

となる。

さてこの予測値の時間平均が，予測値 $f(t+\alpha)$に一致するためには，$m(t)$, $n(t)$ の時間平均が0であることより，

$$g(t+\alpha)=g(t)+\alpha g^{(1)}(t)+\frac{\alpha^2}{2!}g(t)+\cdots\cdots+\frac{\alpha^n}{n!}g^{(n)}(t), \quad (2-11)$$

$$p(t+\alpha)=\sum a_i\ (\cos\frac{i\pi\alpha}{T}\sin\frac{i\pi t}{T}+\sin\frac{i\pi\alpha}{T}\cos\frac{i\pi t}{T}) \quad (2-12)$$

として条件 (2-2) すなわち

$$\lim_{h\to\infty} \frac{1}{2h}\int_{-h}^{h} f(t+\alpha)\,dt = \lim_{h\to\infty} \frac{1}{2h}\int_{-h}^{h} dt \int_{0}^{T} f(t-\tau)K(\tau)d\tau \qquad (2\text{—}13)$$

の両辺を比較し，

条　件

$$\left.\begin{aligned} &\mu_0 \equiv \int_0^T K(\tau)d\tau = 1, \\ &\mu_1 \equiv \int_0^T \tau K(\tau)d\tau = -\alpha, \\ &\quad\cdots\cdots\cdots\cdots\cdots\cdots \\ &\quad\cdots\cdots\cdots\cdots\cdots\cdots \\ &\mu_n \equiv \int_0^T \tau^n K(\tau)d\tau = (-1)^n \alpha^n \end{aligned}\right\} \qquad (2\text{—}14)$$

および

$$\left.\begin{aligned} &Kc(i) \equiv \int_0^T K(\tau)\cos\frac{i\pi\tau}{T}d\tau = \cos\frac{i\pi\alpha}{T} \\ &Ks\,(i) \equiv \int_0^T K(\tau)\sin\frac{i\pi\tau}{T}d\tau = -\sin\frac{i\pi\alpha}{T} \end{aligned}\right\} \qquad (2\text{—}15)$$

$$i = 1,\ 2, \cdots\cdots,\ m.$$

を得る。

一方，汎函数（２—１）の右辺において，$K(\tau)$ が条件（２—14），（２—15）を満たすときには，

$$I[K] \equiv \lim_{h\to\infty} \frac{1}{2h}\int_{-h}^{h} \{m(t+\alpha)+n(t+\alpha)-\int_0^T m(t-\tau)K(\tau)d\tau$$

$$-\int_0^T n(t-\tau)K(\tau)d\tau\}^2\,dt \qquad (2\text{—}16)$$

となる。したがって右辺を展開し，

$$\left.\begin{aligned} &\phi_m\,(\tau) = \lim_{h\to\infty} \int_{-h}^{h} m(t+\tau)m(t)dt, \\ &\phi_n\,(\tau) = \lim_{h\to\infty} \int_{-h}^{h} n(t+\tau)n(t)dt, \end{aligned}\right\} \qquad (2\text{—}17)$$

および $m(t)$，$n(t)$ の独立性と時間平均が０であるという仮定を用いれば，

$$I[K] = \phi_m\,(0) + \phi_n(0) - 2\int_0^T \phi_m(\alpha+\tau)K(\tau)d\tau - 2\int_0^T \phi_n(\alpha+\tau)K(\tau)d\tau$$

$$+\int_0^T K(\tau)d\tau \int_0^T \phi_m(\tau-\sigma)K(\sigma)\,d\sigma$$

$$+\int_0^T K(\tau)d\tau\int_0^T \phi_n(\tau-\sigma)K(\sigma)d\sigma \tag{2-18}$$

を得る。

したがって条件付き変分問題は，Lagrange の乗数 λ_0，λ_1，……，λ_n，ρ_1，ρ_1'，ρ_2，ρ_2'……，ρ_m，ρ'_m をとって

$$I'[K]\equiv I[K]-2\lambda_0\mu_0-2\lambda_1\mu_1-\cdots\cdots-2\lambda_n\mu_n \tag{2-19}$$

$$-2\sum_1^m{}_i\left(\rho_i W_c(i)+\rho_i' W_c(i)\right)$$

となる。

あるいは，

$$I'[K]=\phi_m(0)+\phi_n(0)-2\int_0^T K(\tau)d\tau[\phi_m(\alpha+\tau)+\phi_n(\alpha+\tau)$$

$$-\int_0^T\phi_m(\tau-\sigma)K(\sigma)\,d\sigma-\int_0^T\phi_n(\tau-\sigma)K(\sigma)d\sigma$$

$$+\lambda_0+\lambda_1\tau+\cdots\cdots+\lambda_n\ \tau_n+\sum_1^n\rho_i\cos\frac{i\pi\tau}{T}+\sum_1^m\rho_i'\sin\frac{i\pi\tau}{T}] \tag{2-20}$$

と書ける。

まえとまったく同様の計算により，$I'[K]$が極小値をとるための必要十分条件として，

$$\int_0^T\phi_m(\tau-\sigma)K(\sigma)d\sigma+\int_0^T\phi_n(\tau-\sigma)K(\sigma)d\sigma$$

$$=\phi_m(\tau+\alpha)+\phi_n(\tau+\alpha)+\lambda_0+\lambda_1\tau+\cdots\cdots+\lambda_n\tau^n$$

$$+\rho_1\cos\frac{\pi\tau}{T}+\rho_2\cos\frac{2\pi\tau}{T}+\cdots\cdots+\rho_n\cos\frac{m\pi\tau}{T}$$

$$+\rho_1\sin\frac{\pi\tau}{T}+\rho_2\sin\frac{2\pi\tau}{T}+\cdots\cdots+\rho_r\sin\frac{m\pi\tau}{T}, \tag{2-21}$$

$$T\geqq\tau\geqq0$$

を得る。

この方程式と（2—14），（2—15）の $2m+n+1$ 個の条件とより，最適の予測子 $K(\tau)$ が決められる。

ここで $n=m=0$，$\phi_m(\tau)+\phi_n(\tau)=\phi(\tau)$，$T\to\infty$ とすれば，

$$\int_0^\infty\phi(\tau-\sigma)K(\sigma)d\sigma=\phi(\tau+\alpha),\qquad\tau\geqq0 \tag{2-22}$$

となり，Wiener[1] の場合に一致する。

また $m(t)\equiv0$，$m=0$ とすれば，

$$\int_0^T \phi_n(\tau-\sigma)K(\sigma)d\sigma=\lambda_0+\lambda_1\tau+\cdots\cdots+\lambda_n\tau^n \tag{2—23}$$

$$T\geqq\tau\geqq 0$$

となって，Phillips および Weiss の得た場合に一致する。

また $m=0$ とすれば，

$$\int_0^T \phi_m(\tau-\sigma)K(\sigma)\,d\sigma+\int_0^T \phi_n(\tau-\sigma)K(\sigma)d\sigma$$

$$=\phi_m(\tau+\alpha)+\phi_n(\tau+\alpha)+\lambda_0+\lambda_1\tau+\cdots\cdots+\lambda_n\tau^n \tag{2—24}$$

となり，これは Zadeh & Ragazzini[1] の場合に一致する。

第2節 Davis の理論[1]

(a) 仮　定

次のような仮定をおく。

(1) 信号の有限な平均値 $Es(t)=m(t)$ が存在する。しかし未知である。

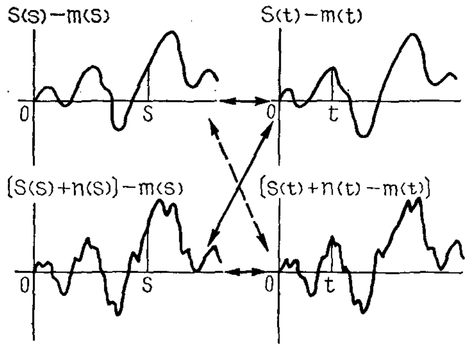
(2) $En(t)\equiv 0$

(3) 次の有限な共分散函数 (covariance function) が存在し，既知で，$0\leq s\leq T$，$0\leq t\leq T$ で連続である。(第2—3図参照)

$$\gamma_s(s,t)\equiv E\{[s(s)-m(s)][s(t)-m(t)]\} \tag{2—25}$$

$$\gamma_{sn}(s,t)\equiv E\{[s(s)-m(s)][s(t)-m(t)+n(t)]\} \tag{2—26}$$

$$R_{sn}(s,e)\equiv E\{[s(s)-m(s)+n(s)][s(t)-m(t)+n(t)]\} \tag{2—27}$$



第2—3図 (Davisの理論)

実際の予測の場合には，$m(t)$ は未知であるが，その性質については何らかの知識をもっているのが普通である。したがってその知識を予測の際にいかにして理論に取入れるかが問題となる。もし十分なデータが得られていれば，$m(t)$をn次の多項式で近似することができる。

(2) の仮定はほとんどの場合成立する。

(3) の仮定は，厳密にいえば，決して満たされることはない。それは共分散函数 (covariance function) を正確に知ることは決してできないからである。しかしそれを現在迄に得られたデータか，または物理的な考察によって求めなければならない。今後，でき得れば，共分散函数についてある程度のあいまいさを残したまま予測の可能な理論をつくり上げることが望まれる。この理論では共分散函数が正確に分ってもその応用には困難が伴う。

また共分散函数が連続であるという仮定がこの理論では非常に大切であるが，大部分の応用上

の問題に際してはこのことは満足される。

論点を明らかにするために簡単な場合だけを考えることにする。

$[0, T]$の間に $s(\tau)+n(t)$ の値を観測して $s(T+\alpha)(\alpha>0)$ の値を予測することを考える。予測の公式を導く前に M. Loève* と K. Karhumen** によって得られた結果を述べる。それがこの手法の基本となっているからである。

まず $[0, T]$ で連続な確率過程 $x(t)$を考える。もし，$h\to 0$ のとき

$$E[x(t+h)-x(t)]^2\to 0 \tag{2-28}$$

が $t\varepsilon[0, T]$ なる全ての t に対して成立するならば $E[m(t)]$ が連続であるということは，共分散函数の連続と等価である，というのである。

もし $x(t)$ が $m(t)$ なる連続な平均値を持ち，その共分散函数を $r(s,t)$とすれば，

$$x(t)=m(t)+\sum_{i=1}^{\infty} E_i\frac{\varphi_i(t)}{(\lambda_i)^{\frac{1}{2}}} \tag{2-29}$$

と書けて，全ての $t\varepsilon[0, T]$ に対して平均値をもつ。ここに $\{\varphi_i(t)\}$ は次の積分方程式の固有函数，$\{\lambda_i\}$ はその固有値である。

$$\varphi(s)=\lambda\int_0^T r(s,t)\varphi(t)dt \tag{2-30}$$

また $\{z_i\}$ は

$$Ez_i=0, \tag{2-31}$$

$$Ez_i\, z_j=\delta_{ij} \tag{2-32}$$

を満足する。

ところで，

$$\int_0^T\int_0^T r(s,t)f(s)f(t)dsdt$$

$$=E\left\{\int_0^T[x(t)-m(t)]\,f(t)dt\right\}^2\geq 0 \tag{2-33}$$

であるから，$r(\tau,t)$は，$0\leq\tau\leq T$，$0\leq t\leq T$ で半無限であることが分る。

$r_s(\tau,t)$，$r_{sn}(\tau,t)$，$R_{sn}(st)$ は仮定(3)によって，分っているから，$s(t)$ と $n(t)$ を独立事象とすると，上の議論によって

$$s(t)+n(t)=m(t)+\sum_{i=1}^{\infty} z_i\frac{\varphi_i(t)}{(\lambda_i)^{\frac{1}{2}}} \tag{2-34}$$

$$t\varepsilon[0, T]$$

---

* M. Loève, compt. rend. 220, 469 (1946)

** K. Karhumen, Ann. Acad. Sci. Feunicae, AI, Helsinki (1947)

とかける。$\{\varphi_i(t)\}$ は次の積分方程式の固有函数，$\{\lambda_i\}$ は固有値である。

$$\varphi(s)=\lambda\int_0^T R_{sn}(s,t)\varphi(t)dt \tag{2-35}$$

(b) 最適線形予測子の決定

予測を $s(t)$ の線形な汎函数に限定する。$s(t)$ を連続とすると，$s(t)$ の線形な汎函数は，

$$L(s)=\int_{-\infty}^{\infty}s(t)dg(t) \tag{2-36}$$

と書くことができる。

問題を簡単ではあるが，重要な場合に限定する。すなわち，［0，T］で $s(t)+n(t)$ が分っているときに，$s(T+\alpha)$，を予測することを考える。予測値は次の汎函数で表わされると考える。

$$\int_0^T[s(t)+n(t)]K(t)dt \tag{2-37}$$

そして，$K(t)$は予測値の時刻 $T+\alpha$ においても有効であるとする。

$K(t)$は $\varphi_i(t)$を完備正規直交系とすると

$$K(t)=\sum_{i=1}^{\infty}a_i\varphi_i(t),\quad 0\leq t\leq T \tag{2-38}$$

と書ける。

ここに

$$x_i=\int_0^T[s(t)+n(t)]\varphi_i(t)dt \tag{2-39}$$

で定義されるフーリエ級数 $\{x_i\}$ を導入すれば，予測値を

$$s^*(T+\alpha)=\sum_{i=1}^{\infty}a_ix_i \tag{2-40}$$

の形に限定したことになる。

この節では $s^*(\tau+\alpha)$ をとくに予測子 (predictor) と呼ぶことにすると，$s(T+\alpha)$ の全ての予測子のうちで平均二乗誤差が最小になる唯一つのバイアスのない予測子が存在することを示そう。

定義によって，

$$E^*(T+\alpha)=m(T+\alpha) \tag{2-41}$$

が成り立つとき不偏 (unbiased) であるという。

任意の予測子による誤差は，

$$s^*(T+\alpha)-s(T+\alpha) \tag{2-42}$$

であるから，最小二乗誤差は，

$$I=E\{s^*(T+\alpha)-s(T+\alpha)\}^2=\gamma_s(T+\alpha,\ T+\alpha)+E[s^*(T+\alpha)-m(T+\alpha)]^2$$

$$+2E\{[s^*(T+\alpha)-m(T+\alpha)][m(m(T+\alpha)-s(t+\alpha)]\}$$

となる。

ところで（2－34），（2—39）によって

$$x_i=(m,\ \varphi_i)+\frac{z^i}{(\lambda_i)^{\frac{1}{2}}} \tag{2—43}$$

である。そして $\{z_i\}$ は互いに独立であるから，（2—40），（2—43）より $s^*(T+\alpha)$が不偏となるための条件は次のようになる。

$$\sum_{i=1}^{\infty} a_i(m,\ \varphi_i)m(T+\alpha) \tag{2—44}$$

以上のことから，

$$s^*(T+\alpha)-m(T+\alpha)=\sum_{i=1}^{\infty}\frac{a_i}{(\lambda_i)^{\frac{1}{2}}}z_i \tag{2—45}$$

であり，

$$E[s^*[T+\alpha]-m(T+\alpha)]^2=\sum_{i=1}^{\infty}\frac{a_i{}^2}{\lambda_i} \tag{2—46}$$

となる。

I の最後の項を考慮すれば，

$$\begin{aligned}&E\{[s^*(T+\alpha)-m(T+\alpha)][m(T+\alpha)-s(T+\alpha)]\}\\&\quad=-E\{[s(T+\alpha)-m(T+\alpha)]\}\\&\qquad\qquad\times\sum_{i=1}^{\infty}a_i[s(t)+n(t)-m(t)\ ;\ \varphi_i(t)]\\&\quad=-\sum_{i=1}^{\infty}a_i(r_{sn}(T+\alpha,\ t),\ \varphi_i(t))\end{aligned} \tag{2—47}$$

である。

ここでもちろん，

$$r_{sN}(T+\alpha,\ t),\varphi_i(t))=\int_0^T r_{sn}(T+\alpha,\ t)\varphi_i(t)dt \tag{2—48}$$

である。

これよりけっきょく

$$I=r_s(T+\alpha,T+\alpha)+\sum_{i=1}^{\infty}\frac{a_i{}^2}{\lambda^i}-2\sum_{i=1}^{\infty}a_i[r_{sn}(T+\alpha,t),\ \varphi_i(t)] \tag{2—49}$$

となる。

$m(t)$ は n 次の多項式と仮定したから，

$$m(t)=\sum_{i=1}^{n}\beta_i\,t^j \tag{2—50}$$

と書ける。

$$(m,\ \varphi_i)=\sum_{j=0}^{n}\beta_j(t^j,\varphi_i) \tag{2—51}$$

であるから（2—41）の条件は，

$$\sum_{i=1}^{\infty}a_i\sum_{j=0}^{n}\beta_j(t^j,\ \varphi_i)=\sum_{j=0}^{n}\beta_j(T+\alpha)^j \tag{2—52}$$

となる。

和の順序を変換して，$\alpha_{ij}=(t^j,\varphi_i)$ とかくと，　（2—53）

$$\sum_{j=0}^{n}\beta_j\left\{\sum_{i=1}^{\infty}\alpha_{ij}\,a_i-(T+\alpha)^j\right\}=0 \tag{2—54}$$

を得る。

この条件は，$m(t)$ の全ての有限な値に対して満足されねばならない。特に全ての線形独立な $\{\beta_j\}$，$j=1,\ 2,\cdots\cdots,\ n$ に対して満足されねばならぬから，次の（2—55）を得る。

$$\sum_{i=1}^{\infty}\alpha_{ij}\,a_i-(T+\alpha)^j=0\,,\qquad j=0,\ 1,\ 2,\cdots\cdots,\ n \tag{2—55}$$

$s(T+\alpha)$ の全ての不偏な最小二乗予測子は（2—55）式の $n+1$ 個の条件の下に，Iを最小にするということから求めることができる。

これらの条件を満足する $\{a_i\}$，$i=1,\ 2,\cdots\cdots,$ が存在することを示そう。

Lagrange の乗数 $2\mu_0$，$2\mu_1$，$\cdots\cdots$，$2\mu_n$ を導入すれば，問題は

$$H=I-2\mu_0\sum_{i=1}^{\infty}\alpha_{i0}\,a_i-2\mu_1\left\{\sum_{i=1}^{\infty}\alpha_{i1}\,a_i-(T+\alpha)\right\}$$

$$\cdots\cdots-2\mu_n\left\{\sum_{i=1}^{\infty}\alpha_{in}\,a_i-(T+\alpha)^n\right\} \tag{2—56}$$

を最小にすることである。

解は，次の条件より得られる。

$$\frac{\partial H}{\partial a_i}=0\,,\qquad i=1,\ 2,\cdots\cdots \tag{2—57}$$

$$\frac{\partial H}{\partial \mu_j}=0\,,\qquad j=0,\ 1,\ 2,\cdots\cdots,\ n \tag{2—58}$$

（2—57）より，

$$a_i=\lambda_i(r_{sn}(T+\alpha,\ t),\ \varphi_i(t))+\mu_0\lambda_i\alpha_{i0}+\cdots\cdots+\mu_n\lambda_i\alpha_{in}\qquad i=1,\ 2,\cdots\cdots \tag{2—59}$$

（2—58）より

$$\sum_{i=1}^{\infty}\lambda_i\alpha_{ij}\{r_{sn}(T+\alpha,\ t),\ \varphi_i(t)+\mu_0\alpha_{i0}+\cdots\cdots+\mu_n\alpha_{in}\}-(T+\alpha)^j=0 \tag{2—60}$$

を得る。

これらを書き直すと,

$$\mu_0\sum_{i=1}^{\infty}\lambda_i\alpha_{i0}\alpha_{ij}+\cdots\cdots+\mu_n\sum_{i=1}^{\infty}\lambda_i\alpha_{in}\alpha_{ij}$$

$$=(T+\alpha)^j-\sum_{i=1}^{\infty}\lambda_i\alpha_{ij}(r_{sn}(T+\alpha,\ t),\ \varphi_i(t)) \qquad (2-61)$$

$$j=0,\ 1,\ 2,\cdots\cdots,\ n$$

となる。

$\alpha_{ij}\alpha_{ik}<T^{jk}M_i^2$ であるから,(2—40)の仮定より,(2—60)の各項は有限であることは明らかである。

(2—59)に(2—60)を代入した(2—61)の $\{\mu_i\}$, $i=0,\ 1,\ 2,\cdots\cdots,\ n$, に対する解は,線形予測子

$$s^*(T+\alpha)=\sum_{i=1}^{\infty}a_ix_i \qquad (2-62)$$

の係数 a_i を与える。

また予測の誤差は,(2—59)で得られた a_i を(2—61)に代入して得られる。

明らかにこの過程が確率的であるための条件は $\alpha>0$ に対して誤差 I が

$$I>0 \qquad (2-63)$$

となることである。

特別な場合として, $s(t)$, $n(t)$ がパワー・スペクトルをもつときには, 非決定的であるための $s(t)$ と $n(t)$ の共分散函数に関する必要十分条件は,M. Loéve によって与えられた。

ここで,$Es(t)\equiv 0$ の場合には,a_i が直接(2—59)から得られる。または(2—49)を直接微分すればよい。a_i は

$$a_j=\lambda_j(r_{sn}(T+\alpha,\ t),\ \varphi_i(t)) \qquad (2-64)$$

である。これを(2—49)に代入して,

$$I_0=r_s(T+\alpha,\ T+\alpha)-\sum_{i=1}^{\infty}\lambda_i(r_{sn}(T+\alpha,\ t),\ \varphi_i(t))^2 \qquad (2-65)$$

となる。そして最適な予測子は,

$$s^*(T+\alpha)=\sum_{i=1}^{\infty}\lambda_ix_i(r_{sn}(T+\alpha,\ t),\ \varphi_i(t)) \qquad (2-66)$$

で与えられる。

また $n(t)\equiv 0$, すなわち, 雑音がない場合には,

$$a_i=\lambda_i\int_0^T r_s(T+\alpha,\ t)\varphi_i(t)dt=\varphi_i^*(T+\alpha) \qquad (2-67)$$

を得る。$\{\varphi_i^*(T+\alpha)\}$ は連続な固有函数である。このとき x_i は

$$x_i=\frac{z_i}{(\lambda_i)^{\frac{1}{2}}} \tag{2-68}$$

となって最適の予測子は

$$s^*(T+\alpha)=\sum_{i=1}^{\infty}\frac{\varphi_i^*(\tau+\alpha)}{(\lambda_i)^{\frac{1}{2}}}z_i \tag{2-69}$$

となる。

第2章　最近の予測理論

最近の予測理論として，Farmer[1] と R. E. Kalman[1] を紹介する。Farmer の論文は1963年の I. F. A. C. (International Federation of Automatic Control) の集りで発表されたもので，実際に電力需要の予測を IBM 7090 使って行い，かなりの精度で予測値を得ている。Kalman は，最近，最適制御理論その他で広く活動しており，その予測理論は基礎的なところで大切な考え方を示している。

まず最初に Farmer の研究を紹介するが，これはある点では Wiener とは根本的に考え方を異にしていると云ってもよい。

Wiener の理論は，厳密な意味で定常な場合を扱った。そして第2章で紹介したようなWiener の拡張された理論においても時系列の平均値はなかば分っているような広義の定常的な場合を扱っている。

その後 Zadeh[1]~[3]等によって，また Wiener[2]自身によって非線形理論も展開されたが，それらはいずれも，そのシステムの満足すべき制約条件を考慮に入れたものではなかった。

一般化された Wiener の理論ではM個の標本函数を $x_m(t)$，$(m=1,\ 2,\cdots\cdots,\ M)$ とすると自己相関函数

$$R(\tau,\tau')=\frac{1}{M}\sum_{m=1}^{M}x_m(\tau)x_m(\tau')$$

が分っていなければならない。理想的には，標本函数 $x_m(t)$ は同一条件の下で，M組の観測を行ってきめなければならない。しかし実際の場合には，同一の条件の下で観測できる標本函数の数は，その過程に含まれるパラメーターの変化をコントロールすることができないので，非常に制限されることになる。

たとえていえば，$x_m(t)$ を毎日の電力需要の曲線とする。そのとき自己相関函数をつくるときに使うことのできるパラメーターの数は，電力負荷に季節変動のあることや経済成長のあることなどによって制限を受ける。

しかしながら Farmer によれば，標本函数は，多くの統計的過程について少数の "Characteristic Modes" によって表わすことができて、それを使えば，自己相関函数を計算するのに，たくさんの標本函数を使うことは必要でないばかりか，かえって望ましくないと述べている。

"Characteristic Modes" はそのプロセスの本質に関する物理的性質を表わすものであり，問題にしている統計的な過程を特徴づけるものは，ただ "modes" のみであって自己相関函数ではないという。

Farmer のつぎには Kalman の研究を紹介する。現在では Wiener-Kolmogorov の沪波と予測の理論は，古典的で，種々の解法手段は十分に確立されており，今後多少の改良や一般化は行われるにしても，本質的な改良は期待できないという見方もある。しかし実際には，その方向の研究活動は，活発に行われている。この Kalman もその1つの例である。

Shinbrot,[1] Steeg,[1] Pugachev,[3] Parzen[1] らは，時間領域における種々の方法を用いて，非定常な，長期にわたる沪波と予測の問題を解いた。

Kalman はこれらの人々と独立に同じ問題を扱い，数々の新しい結果も得ている。時間領域における vector-matrix-method を用いて，従来の Wiener 理論に大きな改良を加え，一般化したのであるが，その結果は何らの修正なしに多重時系列に適用できる。

Kalman の研究の歴史的な背景は次のようなものである。

Follin[1] は，標準的な Wiener 問題の拡張を行っている際に，ある与えられた回路について，時間につれて変動する利得と，誤差変動との間にはある関係のあることをみい出した。後に Hanson[1] は Follin の回路構成が，そこで仮定された統計的性質の下では，最適のものであることを証明し，さらに，最初 Follin が得た誤差変動に関する微分方程式は，厳密に Wiener-Hopf 方程式になることを示した。

Bucy[1] はこれらの結果を一般化し，最適な予測子と，誤差変動との陽表的な関係をみい出した。彼はまた分散方程式と広範囲にわたる，非定常な信号と雑音との統計的性質に関する式を導いた。

これらの研究と独立に，Kalman は従来の予測理論をまったく新しい角度よりとり組んだ。その新しさは，次の2つのよく知られた考え方を結びつけたことにある。

(i) ダイナミックな系を表わすのに「状態」(state) という考え方をとり入れた。

(ii) 線形の予測を ヒルベルト空間における正射影とみなした。

その結果，副産物として，決定論的な制御の問題と，統計的な予測の問題が本質的に同じであることを見い出した。したがって，線形な最適制御の問題において得られた結果はただちに Wiener 問題に応用できることになる。

Wiener-Hopf 方程式を直接解かずに，それと等価な，非線形差分方程式をとくことによって最適予測器を設計する為の情報を得ることができる。

第1節 Farmer の研究[1]

m 個の標本函数を $x_m(t)$, $m=1, 2, \cdots\cdots, M$ とする。第1次モードを $\varphi_1(t)$ とすると，誤差 $e_m(t)$ は，

$$e_m(t)=x_m(t)-\lambda^{\frac{1}{2}}\, a_{m1}\varphi_1(t) \qquad (3-1)$$

であり，平均二乗誤差は

$$\overline{[e_{m1}(t)]^2}=\frac{1}{MT}\sum_{m=1}^{M}\int_0^T [e_m(t)]^2dt$$

$$=\frac{1}{MT}\sum_{m=1}^{M}\int_0^T \{x_m(t)-\lambda_1^{\frac{1}{2}}a_{m1}\varphi_1(t)\}^2dt \qquad (3-2)$$

となる。$\overline{[e_{m1}(t)]^2}$ が最小になるように $\varphi_1(t)$ をきめるわけだが，変分問題の解として，次の関係が成り立つことが必要十分であることが得られる。

$$\left.\begin{aligned}&\int_0^T \varphi_1(t)x_m(t)dt=\lambda_1^{\frac{1}{2}}a_{m1}\int_0^T \varphi_1^2(t)dt,\quad m=1,\ 2,\cdots\cdots,\ M.\\&\sum_{m=1}^{M}a_{m1}x_m(t)=\sum_{m=1}^{M}\lambda_1^{\frac{1}{2}}a_{m1}^2\varphi_1(t),\qquad 0<t<T\end{aligned}\right\} \qquad (3-3)$$

ここで $\varphi_1(t)$，a_{m1} を正規化すると，

$$\left.\begin{aligned}&\int_0^T \varphi_1^2(t)dt=1\\&\sum_{m=1}^{M}a_{m1}=M\end{aligned}\right\} \qquad (3-4)$$

となり，（3－3）の条件は，

$$\left.\begin{aligned}&a_{m1}=\lambda_1^{-\frac{1}{2}}\int_0^T \varphi_1(t)x_m(t)dt\\&\lambda_1^{\frac{1}{2}}\varphi_1(t)=\frac{1}{M}\sum_{m=1}^{M}a_{m1}x_m(t)\end{aligned}\right\} \qquad (3-5)$$

となる。（3－5）の２式から a_{m1} を消去すると，

$$\frac{1}{M}\sum_{m=1}^{M}\int_0^T \varphi_1(\tau)x_m(\tau)d\tau\cdot x_m(t)=\lambda_1\varphi_1(t) \qquad (3-6)$$

または，

$$\int_0^T R(t,\tau)\varphi_1(\tau)d\tau=\lambda_1\varphi_1(t) \qquad 0<t<T \qquad (3-7)$$

を得る。

ここに $R(t,\tau)$ は

$$R(t,\tau)=\frac{1}{M}\sum_{m=1}^{M}x_m(t)x_m(\tau)$$

なる相関函数である。

（3－2），（3－7）より，最小二乗誤差は，

$$\overline{[e_{m1}(t)]^2}_{min}=\frac{1}{T}\left\{\int_0^T R(\tau,\tau')\,d\tau-\lambda_1\right\} \quad (3-8)$$

で与えられる。

第2次モード，および誤差を全く同じようにして求められる。

結局，$x_m(t)$は，

$$x_m(t)=\lambda_1^{\frac{1}{2}}a_{m1}\varphi_1(t)+\cdots\cdots+\lambda_2^{\frac{1}{2}}a_{m2}\varphi_2(t)+\cdots\cdots \quad (3-9)$$

となり，$\varphi_k(t)$は，次の積分方程式，

$$\int_0^T R(t,\tau)\varphi_k(\tau)d\tau=\lambda_k\varphi_k(t) \qquad 0<t<T \quad (3-10)$$

$$R=1,\ 2,\ 3,\cdots\cdots$$

の固有函数，λ_kはその固有値として求められる。

ここで，a_{mk} は

$$a_{mk}=\lambda_k^{-\frac{1}{2}}\int_0^T\varphi_k\,x_m(t)dt \quad (3-11)$$

で与えられる。

そして，このときの最小二乗誤差は，

$$\overline{[e_{mk}]^2}_{min}=\frac{1}{T}\left[\int_0^T R(\tau,\tau')\,d\tau'-\sum_{k=1}^{K}\lambda_k\right] \quad (3-12)$$

で与えられる。

これは，次の（3－13），（3－14）式

$$\int_0^T \varphi_k(\tau)\varphi_k'(\tau)d\tau=\delta_{kk'} \quad (3-13)$$

$$R(\tau,\tau')=\sum_{k=1}^{\infty}\lambda_k\varphi_k(\tau)\varphi_k(\tau') \quad (3-14)$$

より出てくる関係式，

$$\int_0^T R(\tau,\tau')d\tau=\sum_{k=1}^{\infty}\lambda_k \quad (3-15)$$

を使えば，

$$\overline{[e_{mk}(t)]^2}_{min}=\frac{1}{T}\sum_{k=K+1}^{\infty}\lambda_k \quad (3-16)$$

と書ける。これより，必要とする精度に応じてKをいくつにとったらよいかが定まる。

次に以上の“characteristic modes”を使っての予測の方法を述べる。

いま観測しつつある時系列を $x(t)$ とする。

これが先にm個の標本函数を使って求めたモードによる線形結合として，

$$x(t)=\sum_{k=1}^{K}c_k\varphi_k(t),\qquad 0<t<T \tag{3—17}$$

という形で表わされるとする。Kは必要とする精度によってきめる。そしてこの表わし方が，$0<t<T_0$（T_0；現在時点）のみならず，予測される時点を含めたところまで成り立つと考えて，係数c_kを決めればよい。

このc_kの決め方は，簡単な方法として，次のようなものがある。

誤差$e(t)$は，

$$e(t)=x(t)-\sum_{k=1}^{K}c_k\varphi_k(t) \tag{3—18}$$

となるから，この誤差の二乗平均値Iを最小になるようにする。

$$\begin{aligned}I&=\int_0^{T_0}[e(t)]^2dt\\&=\int_0^{T_0}\{x(t)-\sum_{k=1}^{K}c_k\varphi_k(t)\}^2dt\end{aligned} \tag{3—19}$$

であり，

$$\frac{\partial I}{\partial c_k}=0\qquad(k=1,\ 2,\ 3,\cdots\cdots,\ K.)$$

より，

$$\sum_{k'=1}^{K}c_{k'}\int_0^{T0}\varphi_k(\tau)\varphi_{k'}(\tau)\,d\tau=\int_0^{T_0}x(\tau)\varphi_k(\tau)d\tau \tag{3—20}$$

$$k=1,\ 2,\cdots\cdots,\ K.$$

となり，これを解いて得られるc_kを使って予測することができる。

なお，Volterra 級数展開により係数を厳密に求める方法も紹介されている。

以上は，連続な$x(t)$についての解法であるが，離散的な形に書き直したものが，附録としてつけられている。

それを用いれば，ディジタル計算機による取扱いが可能である。

Farmer は実際に，イギリスのある地方の電力の需要予測をこの方法で行った。それによれば，“characteristic mode”を決めるには，20週間のデータを用い，3～4時間先の予測を行って，3％の予測誤差で予測値を得ている。その結果も，文献の最後に載せられている。ちなみに，計算時間は，IBM 7090 で約1分とのことである。

第2節 Kalman の研究 — Wiener 問題の解

Kalman は Wiener 問題を状態(state)という見方から解いた。このことは，事実上他の手法では非常に困難になるような一般的な場合を含む問題の解法を導くことになったのである。そし

て Wiener 理論の基本的な事実が速やかに得られ，結果の展望と基礎的な仮定がはっきりした。

ダイナミック・モデルとして次のようなものを考える。（第3－1図参照）

$$x(t+1)=\Phi(t+1,\ t)x(t)+u(t) \tag{3-21}$$

$$y(t)\quad=M(t)x(t) \tag{3-22}$$

ここで

$u(t)$；互いに独立な正規確率過程．一般には n-ベクトルで

$E[u(t)]=0$，　$E[u(t)u(t')]=0$　$(t=t')$ である。

$x(t)$；確率的状態変数，一般には n ベクトル

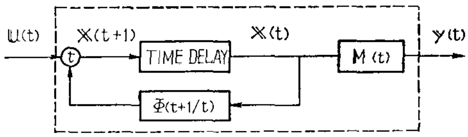
$y(t)$；$x(t)$のうち観測可能なもの，一般には p ベクトル（$p\leqq n$）

$\Phi(t+1,\ t)$；t から $t+1$ への遷移マトリックス（$n\times n$）

$M(t)$；$p\times n$ マトリックス

なお，Φ，Mの要素は時間 t のランダムでない函数である。

このダイナミックモデルは Bode & Shannon の考え方がその基礎となっている。



第3－1図 ダイナミック・モデル

ここで，特に注意すべきことは，$M(t)$ を考えたことである。状態変数 $x(t)$ は，$n_\exists$ 全てが観測されるのではなく，$y(t)$ が観測されるだけである。これは統計学の方で表われる「推定可能」という概念と同じようなものであるが，それを予測の理論にはっきりした形でとり入れたことは重要な意味があり，今後基礎的な理論の展開に際して大きな役割りを果すものと思われる。

このことは，最適制御理論の方でも問題になり，最適ということも，$y(t)$ についての最適ということであって，$x(t)$まで考えたら真に最適かどうかは断定できない。予測の場合も同様で，最適という意味も，$y(t)$ に関するものであって $x(t)$ までならいっていない。

最適の予測値については次の定理が重要である。

<u>定　理</u>

確率過程，$\{x(t)\}$，$\{y(t)\}$ の平均値を 0 と仮定し，さらに次の (a), (b)の仮定をおく。

(a) $\{x(t)\}$，$\{y(t)\}$ は，正規分布をなす。

(b) 評価函数として，誤差の平均二乗値をとり，それを最小にするものを最適とする。また，予測値を観測値の線形な函数とする。

そのとき，$x(t_1)$の最適の予測値 $x^*(t_1/t)$は，観測値 $y(t_0)$，……，$y(t)$が与えられたとき，$Y(t)$ 上への $x(t_1)$ の正射影 $x(t_1/t)$ で与えられるというものである。

ここに $Y(t)$ は観測値 $y(t_0)$，……，$y(t)$ のはる直交空間である。この定理を簡単につぎのように書くことにする。

$$x_1^*(t_1/t)\equiv\bar{x}_1(t_1/t) \tag{3-23}$$

$$=E[x(t_1)/Y(t)]$$

このくわしい説明は，Doob または Pugachev, etc に出ている。

上の定理によって，$x(t_1)$の最適予測値$x^*(t_1/t)$は，$x(t_1)$の $Y(t)$ への正射影で与えられることが分ったが，$x^*(t_1/t)$は，第3－1図に示したダイナミック・モデルの一般的なものによって実現することができる。

いま，$y(t_0), \cdots\cdots, y(t-1)$ が観測されたとする。このことは $Y(t-1)$ が分ったことと同じである。つぎに $y(t)$ が観測される。$y(t)$ の $Y(t-1)$ に直交する成分を $\tilde{y}(t/t-1)$ としよう。そのとき，もし $\tilde{y}(t/t-1)\equiv 0$ なら，$Y(t)=Y(t-1)$となって，$y(t)$を観測したことによって，予測値については何の情報も得られず，予測値を改善することはできないであろう。$\tilde{y}(t/t-1)$ のつくる空間を $Z(t)$ とすると，$Y(t-1)$ と $Z(t)$ とで $Y(t-1)$ を形づくることになる。もちろん $Z(t)$，のすべてのベクトルは $Y(t-1)$ に直交している。

さて，$x^*(t_1-1/t-1)$ がもし分ったとすると$x^*(t_1/t)$ をつぎのようにかくことができる。

$$\begin{aligned} x^*(t_1/t) &= E[x(t_1)/Y(t)] \\ &= E[x(t_1)/Y(t-1)]+E[x(t_1)/Z(t)] \\ &= \Phi(t+1\,;\,t)x^*(t_1-1/t-1)+E[u(t_1-1)/Y(t-1)] \\ &\qquad +E[x(t_1)/Z(t)] \end{aligned} \tag{3—24}$$

（2—24）の等式は（3—21）をつかって得られる。

ここで，簡単のために $t_1=t+1$ と限定しよう。そうすると

（3—24）の最後の項は，確率変数 $\tilde{y}(t/t-1)$ への線形な演算を示しているから

$$E[x(t+1)/Z(t)]=\Delta^*(t)\tilde{y}(t/t-1) \tag{3—25}$$

とかける。$\Delta^*(t)$ は $n\times p$ マトリックスで，* 印は最適の意味である。

$y(t)$ の成分のうち$Y(t-1)$ にあるものを $\bar{y}(t/t-1)$ とすると

$$\bar{y}(t/t-1)=M(t)x^*(t/t-1) \tag{3—26}$$

である。

したがって，

$$\begin{aligned} \tilde{y}(t/t-1) &= y(t)-\bar{y}(t/t-1) \\ &= y(t)-M(t)x^*(t/t-1) \end{aligned} \tag{3—27}$$

となる。（3－24）と（3－27）とから，次の重要な式，

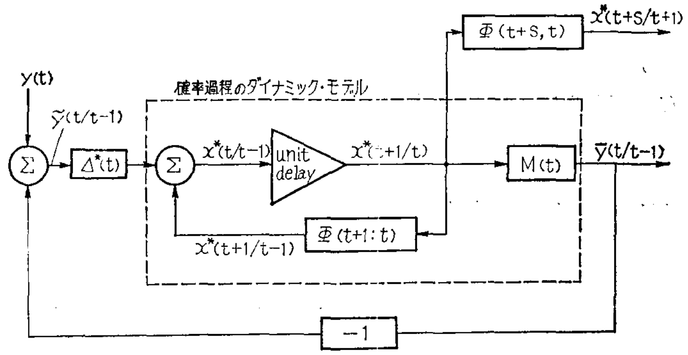
$$x^*(t+1/t)=\Phi^*(t+1\,;\,t)x^*(t/t-1)+\Delta^*(t)y(t) \tag{3—28}$$

を得る。ここで

$$\Phi^*(t+1\,;\,t)\equiv\Phi(t+1\,;t)-\Delta^*(t)M(t) \tag{3—29}$$

である。

（3—28）を（3—21）と比べてみれば，そのダイナミック・システムの形は同じであることが分る。



第3－2図　最適予測器のブロック・ダイアグラム

第3－2図に示すように，その入力は $\boldsymbol{y}(t)$ で，遷移行列 $\boldsymbol{\Phi}^*$ は（3—29）で与えられる。最適予測器を物理的に実現するには，

（i）確率過程のモデル，と

（ii）演算子 $\Delta^*(t)$

とか分ればよいことも分る。

さらに予測誤差もまた，同じような線形のダイナミック・システムによって表わすことができる。実際，予測誤差 $\tilde{\boldsymbol{x}}(t+1/t)$ は，

$$
\begin{aligned}
\tilde{\boldsymbol{x}}(t+1/t) &= \boldsymbol{x}(t+1)-\boldsymbol{x}^*(t+1/t)\\
&= \boldsymbol{\Phi}(t+1\,;\,t)\boldsymbol{x}(t)+\boldsymbol{u}(t)-\boldsymbol{\Phi}^*(t+1\,;\,t)\boldsymbol{x}^*(t/t-1)\\
&\qquad -\boldsymbol{\Delta}^*(t)\boldsymbol{M}(t)\boldsymbol{x}(t)\\
&= \boldsymbol{\Phi}^*(t+1\,;\,t)x(t/t-1)+\boldsymbol{u}(t)
\end{aligned}
\qquad (3—30)
$$

となる。このように $\boldsymbol{\Phi}^*(t+1\,;t)$ は予測誤差についてのダイナミック・システムの遷移行列となっている。

（3—30）より，予測誤差 $\boldsymbol{x}$ の分散行列 $\boldsymbol{p}^*$ について次のような漸化式を得ることができる。(′)は転置行列を意味する。

$$
\begin{aligned}
\boldsymbol{p}^*(t+1) &= \boldsymbol{E}\boldsymbol{x}(t+1/t)\tilde{\boldsymbol{x}}'(t+1/t)\\
&= \boldsymbol{\Phi}^*(t+1\,;t)\boldsymbol{E}\tilde{\boldsymbol{x}}(t/t-1)\tilde{\boldsymbol{x}}'(t/t-1)\boldsymbol{\Phi}^{*\prime}(t+1\,;t)+\boldsymbol{Q}(t)\\
&= \boldsymbol{\Phi}^*(t+1\,;t)\boldsymbol{p}^*(t)\boldsymbol{\Phi}'(t+1\,;t)+\boldsymbol{Q}(t)
\end{aligned}
\qquad (3—31)
$$

ここに $\boldsymbol{Q}(t)=\boldsymbol{E}\boldsymbol{u}(t)\boldsymbol{u}'(t)$ である。

残る問題は Δ^*，したがって $\boldsymbol{\Phi}^*$ を求めることである。予測誤差 $\tilde{\boldsymbol{x}}(t+1/t)$ が，$\tilde{\boldsymbol{y}}(t/t-1)$ に直交することを使い，さらに，行列の演算について若干の仮定をおいて，けっきょく，

$$
\boldsymbol{\Delta}^*(t)=\boldsymbol{\Phi}(t+1\,;t)\boldsymbol{p}^*(t)\boldsymbol{M}'(t)[\boldsymbol{M}(t)\boldsymbol{p}^*(t)\boldsymbol{M}'(t)]^{-1} \qquad (3—32)
$$

を得る。

観測は t_0 より開始されるので，$\tilde{\boldsymbol{x}}(t_0/t_0-1)=\boldsymbol{x}(t_0)$であり，（3—31）式によって計算をはじめるにあたって $\boldsymbol{p}^*(t_0)=\boldsymbol{E}\boldsymbol{x}(t_0)\boldsymbol{x}'(t_0)$ をきめる。つぎに（3—32）によって $\boldsymbol{\Delta}^*(t_0)$，（3—29）によって $\boldsymbol{\Phi}^*(t_0+1,t_0)$，ふたたび（3—31）にもどって $\boldsymbol{p}^*(t_0+1)$ というふうに順次計算が実

行できる。

予測値は

$$x^*(t+1/t)=\Phi^*(t+1;t)x^*(t/t-1)+\Delta^*(t)y(t) \tag{3―33}$$

で与えられ，最小二乗誤差は，

$$\sum_{t=1}^{n}E\tilde{x}_i^2(t/t-1)=trace\,p^*(t) \tag{3―34}$$

で与えられる。

なお，(3―29)，(3―31)，(3―32) より，Δ^*, Φ^* を消去すると，p^* についての非線形差分方程式

$$p^*(t+1)=\Phi(t+1;t)\{p^*(t)-p^*(t)M'(t)[M(t)p^*(t)M'(t)]^{-1}$$
$$\times p^*(t)M(t)\}\Phi'(t+1;t)+Q(t) \qquad (t\geqq t_0) \tag{3―35}$$

が得られこれが，従来の Wiener 理論における Wiener-Hopf 方程式に対応するものである。

結　論

予測にかんする Wiener の理論が公にされてから20年以上経過した。その間に夥しい研究が積み重ねられてこの理論が発達してきている。

本論文で，著者達は，まず従来の研究を系統的に分類して，文献の相互の関係を明らかにすることに努力した。

定常確率過程の線形予測にかんしては，Wiener-Kolmogorov の今や古典的な理論が中心で，多くの研究があっても，本質的な改良はあまり顕著ではない。しかし何といってもこの理論は，予測理論の主流をなすものであって，多くの研究もこれに集中している。工学的な応用の面から見ると，Bode & Shannon の研究は重要な意義をもち，これによれば，Wiener の理論の本質もさらに鮮明に解明されている。今後のこの線に沿う研究としては，予測路回路の設計，製作に一層直接後に立つような形式に書き改めることおよび経済的時系列を含めて，統計の時系列の予測に便利なような計算式を作成することがあげられよう。

つぎに，非定常過程の線形予測については，Wiener の流れを汲むもののほかに本質的に新しい要素を含んだ Kalman および Farmer の研究が最近あらわれた。前者は Wiener の予測理論の特徴を分析することから出発して Wiener の理論を特別な場合として包含するような理論を建設した。今後はこのような研究がさらに伸びるものと思われる。

しかるに実際家にとって一番重要なことは，現象の非定常性と非線形性とを考慮に入れた予測手法の確立であることに間違いはない。この第3の分類に属する研究は，その数学的な困難にも拘らず今後ますます伸展するものと期待される。

数学的手法としては，Wiener の用いたいわゆる Wiener-Hopf 型積分方程式の時間預域での解法のほかに周波数預域での解法が工夫され，また積分方程式の直交函数系展開による手法が用いられている。

文　献

Bendat, J. S. "Principles and Applications of Random Noise Theory." John Wiley, 1958.

Blackman, J. "The Representation of Nonlinear Networks." Syracuse University Research Institute, Report No. 81560. for Air Force Cambridge Research Center.

Blum, M. "Recursion Formulas for Growing Memory Digital Filters." Trans. IRE Prof. Group on Information Theory, IT-4, 1958, pp. 24-30.

Bode, H. W. and Shannon, C. E. (1) "A Simplified Derivation of Linear Least Square Smoothing and Prediction Theory." Proceedings IRE, vol. 38, 1950, pp. 417-425.

Booton, R. C. (1) "An Optimization Toeory for Time-Varying Linear Systems with Nonstationary Statistical Inputs." Proceedings IRE, vol. 40, 1952. 977-981.

Brilliant, M. B. "Theory of the Analysis of Nonlinear Systems." Technical Report 345, Research Laboratory of Electronics, MIT 1958.

Bucy, R. S. (1) "Optimimum Finite-Time Filters for a Special Non-stationary Class of Inputs." JHU/APL Internal Memorandom BBD-600 1959.

Carlton, A. G. and Follin, J. W. Jr. (1) "Recent Development in Fixed and Adaptive Filtering." Proceedings of The Second AGARD Guided Missiles Sominar (Guidance and Control) AGARDograph 21, September, 1956.

Chessin, P. (1) "A Bibiography on Noise." IRE Transactions on Information Theory, Sept., 1955.

Chuang, K. K. and Kazda, L. F. "Non-linear Systems with Random Inputs." A.I.E.E. Paper 59-147, Applications and Industry, May, 1959.

長期計画の理論的研究委員会（主査 近藤次郎）(1) "多重時系列の予測" 日本機械工業連合会市場調査研究会資料 38-201 1963.

長期計画の理論的研究委員会（主査 近藤次郎）(2) "時系列の予測 (1)" 日本機械工業連合会市場調査研究会資料 38-201 1963.

長期計画の理論的研究委員会（主査 近藤次郎）(3) "時系列の予測 (2)——乗用自動車の需要予測" 日本機械工業連合会市場調査研究会資料 1964.

Darlington, S. (1) "Linear Least-Square Smoothing and Prediction, with Applications." Bell System Tech. Journal, vol. 37, 1958, pp. 1221-1294.

Darlington, S. (2) "Nonstationary Smoothing and Prediction Using Network Theory Concepts." IRE IT-5 1959 pp. 1221—1294.

Davenport, W. and Root, W. "An Introduction to the Theory of Random Signals and Noise." McGraw-Hill 1958.

Davis, R. C. (1) "On the Theory of Prediction of Nonstationary Stochastic Processes." Journal of Applied Physics, vol. 23, 1952, pp. 1047-1053.

Doob, J. L. "Stochastic Processes." Wiley, New York, 1953.

Farmer, E. D. (1) "A Methos of Prediction for Nonstationary Processed and its Application to the Problem of Load Estimation." Trans. I.F.A.C. 1963.

Flake, R. H. "Volterra Series Representation of Nonlinear Systems." Trans. Amer. Inst. elect. Engrs. Paper No. 62 1189. (1962)

Flake, R. H. "Volterra Series Reoresentation of Time-Varying Nonlinear Systems."

Flanklin, G. (1) "The Optimum Synthesis of Sampled-Data Systems." Doctoral dissertation, Dept. of Elect. Engr., Columbia University, 1955.

Follin (1) → Carlton, A. G. and Follin, J. W. Jr. (1)

Freeman, J. "Principles of Noise." John Wiley, 1958.

Fuller, A. T. "Optimization of Nonlinear control systems with transient inputs." J. Electronics and Control. 8, p. 415, 1960.

George, D. A. "Continous nonlinear systems." Technical Report 355, Research Laboratory of Electronics. MIT 1959.

Gibbon, J. E. "Nonlinear Automatic Control." McGraw-Hill Book Company, Inc. N.Y. 1963.

Gillespie, A. "Signal, Noise and Resolution in Nuclear Counter Amplifier." Pergamon Press, 1953.

Goldman, S. "Frequency Analisis, Modulation and Noise." McGraw-Hill, 1948.

Goldman, S. "Information Theory." Prentice-Hall, 1953.

Hanson, J. E. (1) "Some Notes on the Application of the Calculus of Variations to Smoothing for Finite Time, etc." JHU/APL Internal Memorandom. BBD-346, 1957.

Helstom, C. "Statistical Theory of Signal Detection." Pergamon Press, 1960.

Kalman, R. E. "On the General Theory of Control Systems." Proc. I.F.A.C. Congress Moscow, 1960.

Kalman, R. E. (1) "A New Approach to Linear Filtering and Prediction Problem." Trans. ASME, Series D, Journal of Basic Engineering, vol. 82, 1960, pp. 35-45.

Kalman, R. E. and Bucy, R. S. "New Results in Linear Filtering and Prediction Theory."

Kolmogorov, A. (1) "Interpolation und Extrapolation von Stationären Zuffälligen Folgen." Ball. Acad. Sci. (U.S.S.R.) Ser. Math., vol. 5, 1941, pp. 3-14.

Kotelnikov, V. A. "The Theory of Optimum Noice Immunity." Mc Graw-Hill, 1959.

近藤 次郎 (1) "Wiener 理論にかんする 2, 3 の覚え書き" 統計局研究彙報 第4, 5, 6号 1953~1954.

近藤 次郎 (1) "経営と管理のための数学入門" 日本科学技術連盟ライブラリー ③ 1959. pp. 197-207.

Kondo, J. (2) "A Theory of Multiple Prediction." TIMS. (TOKYO). 1963.

河田 部会 (K 委員会) "時系列および情報の理論とその応用" 日本科学技術連盟, K 委員会, 河田部会報告書 1953,

Laning J. H. and Battin R. H. (1) "Random Processes in Automatic Controll." McGraw-Hill Book Company, Inc., New York, N.Y., 1956.

Lawson, J. and Uhlenbeck, G. "Threshold Signals." McGraw-Hill, 1950.

Lees, A. B. (1) "Interpolation and Extrapolation of Sampled Data." Trans. IRE Prof. Group on Ioformation Theory, It-2, 1956, pp. 173—175.

Levinson, N. "A Heuristic Exposotion of Wiener's Mathematical Theory of Prediction and Filtering." (Wiener's Appendix C.)

Mayne, D. Q. "Optimum Non-Stationary Filters." An Expositions of ADAPTIVE CONTROL. Proceedings of a Symposium held at the Imperial College of Science and Technology, London, 1961. PERGAMON PRESS.

Middleton, D. "An Introduction to Statistical Cmmunication Theory." McGraw-Hill. 1960.

Parzen, E. "Statistical Inference on Time Series by Hilbert Space Methods." I. Tech. Rep. No. 23, Applied Methematics and Statistics Laqoratory, Stanford Univ., 1959.

Parzen, E. "Mathematical Optimization Techniques." edited by R. Bellman. University of California Press. 1963.

Pottle, C. "The Digital Adaptive Control of Linear Process Modulated by Random Noise." IEEE Transactions on Automatic Control. Vol. AC-8. JULY, 1963. No. 3. pp. 228-234.

Pugachev, V. S. (1) "The Use pf Canonical Expansions of Random Functions in Determining an Optimum Linear Systems." Automatics and Remote Control (U.S.S.R.), vol. 17, 1956, pp. 489-499; translation pp. 545-556.

Pugachev, V. S. (2) "General Condition for the Minimum Mean Square Error in a Dynamic System." Autmat. Telemech., Moscow, 17. p. 289, 1956.

Pugachev, V. S. (3) "A Method for Solving the Basic Integral Equations of Statistical Theory of Optimum Systems in Finite Form." Prikl. Math. Mekh., vol. 23, 1959, pp. 3-14. (English Translation: pp. 1-16)

Pugachev, V. S. (4) "Teoriya Sluchainikh Funktsii Ee Primenenie k Zadaham Automaticheskogo Upravlenipa. (Theory of Random Functions and its Application to Automatic Control Problems) (in Russian) second edition, Geotecknizdat, Moscow, 1960.

Rice, S. O. (1) "Mathematical Analysis of Random Noise." Bell System Technical Journal, July, 1944, Jan., 1945.

Sakrison, D. J. "Iterative Design of Optimimum Filters for Non-Mean-Square-Error Performance Criteria." IEEE Transactions on Information Theory. Vol. IT.-9 JULY, 1963, No. 3 pp. 161-167.

Shinbrot, M. (1) "Optimization of Time-Varying Linear Systems with Nonstationary Inputs." Trans. ASME., vol. 80. 1958, pp. 457-462.

Singleton, H. E. "Theory of Nonlimear Transducers." Technical Report No. 160, M.I.T. Research Laboratory of Electronics.

Steeg, C. W. "A Time Domain Synthesis for Optimum Extrapolators." Trans. IRE. Prof. Group on Automatic Control. Nov., 1957, pp. 32-41.

Solodovnikov, V. V. and Batkov, A. M. (1) "On the Theory of Self-Optimizing Systems." (in German and Russian) Proc. Heidelberg Conference on Automatic Control, 1956, pp. 308-323.

Tsien, H. (1) "Engineering Cybernetics." McGraw-Hill. 1954.

Westcott, J. H. "Design of Maltivariable Optimum Filters." Trans. A.S.M.E., 80, p. 463, 1958.

Whittaker, E. and Robinson, G. "The Caluclus of Observations. Blackie and Son, 1944.

Wiener, N. (1) "Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications." John Wiley S. Sons. New York. N.Y. 1949.

Wiener, N. "Non-linear Problems in Random Theory." New York, Technology Press and Wiley. 1958.

Woodward, P. "Probability and Information Theory, with Application to Radar." Pergmon Press 1953.

Zadeh, L. A. and Ragazzini, J. R. (1) "An Extension of Wiener's Theory of Prediction." Journal of Applied Physics, vol. 21. 1950, pp. 645-655.

Zadeh, L. A. (1) "Optimum Non-Linear Filters." J. appl. Phys. 24 (1953).

Zadeh, L. A. (3) "Nonlinear Multipoles." Proc. Nat. Acad. Sei. Wash. 39 (1953). 274.

Zadeh, L. A. (2) "A Contribution to the Theory of Nonlinear Systems." J. Franklin Inst. 255 (1953).

van der Ziel, A. "Noise." Prentice Hall 1954.