J. Op. Res. Japan, Vol. 2, No. 2, 1959

PREDICTION THEORY AND DYNAMIC PROGRAMMING

TOSIO ODANKA

The Metropolitan Technical Collge

(Received January 31. 1959)

(Presented at the 4th meeting, November. 16, 1958)

1. INTRODUCTION

In this paper we wish to discuss the applications of the theory of dynamic programming to the study of prediction theory of the type that arise in two fields of time series in statistics and operations research, and of engineering analysis, electrical and mechanical as well.

We shall first formulate the Wiener Roots Mean Square (R. M. S.) error criterion in filter design.

The purpose of the paper is to present a simple methods, requring a minimal mathematical background, which can be used to treat a large class of prediction theory, of nonstationary stochastic processes, of multiple time series.

Finally, we shall discuss the prediction and turn directly to computational solution of some typical prediction problem.

2. The WIENER R. M. S. ERROR CRITERION IN FILTER DESIGN

Here the discussion will be lined to limited filtering devices in fields of communication engineering.

If we denote a signal by the sequence b_k and a message contained in the signal by the sequence a_k, then we can regards a noise as a sequence of differences, $b_k - a_k$. It is our purpose to find the best way to treat the signal, that is the b_k, so as to obtain the information, the a_k.

Let us try to determine the nature of a linear filter which, with input b_k, will have an output as close as possible to a_k. We see that our problem is to determine the number A_n so that the

$$\varepsilon_k = a_k - \sum_{n=1}^{M} A_n b_{n-k}, \tag{1}$$

are as small as possible.

We want to choose A_n so that rms of the ε_k

$$I = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{k=-N}^{N} (a_k - \sum_{n=1}^{M} A_n b_{k-n})^2 \tag{2}$$

should be a minimum. We introduce the auto-correlation

$$R_a(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{l=-N}^{M} a_l a_{l-k},$$

$$R_b(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{l=-N}^{N} b_l b_{l-k}, \tag{3}$$

and the cross correlation function

$$R_{ba}(k) = \lim_{N \to \infty} \frac{1}{2N+1} \sum_{l=-N}^{N} a_l b_{l-k}.$$

We can write Eq. (2) as

$$I_M = R_a(0) - 2\sum_{n=0}^{M} A_n R_{ba}(n) + \sum_{n,m=0}^{M} A_n A_m R_b(m-n). \tag{4}$$

Our probiems is to close the A_n so as to make I_M a minimum.

If we normalize Eq. (1) by dividing by $R_a(0)$;

$$\frac{I_M}{R_a(0)} = 1 - 2\sum_{n=0}^{M} A_n \frac{R_b(n)}{R_a(0)} + \sum_{n,m=0}^{M} A_n A_m \frac{R_b(m-n)}{R_a(0)} \tag{5}$$

If we now call $I_M/R_a(0)$, V_M, and if we set

$$\frac{R_{ba}(n)}{R_a(0)}=\varphi_n,$$

$$\frac{R_b(a)}{R_a(0)}=\psi_n. \tag{6}$$

then we have

$$V_M=1-E_M, \tag{7}$$

where

$$E_M=2\sum_{n=0}^{M}A_n\psi_n-\sum_{n,m=0}^{M}A_nA_m\varphi_{m-n}. \tag{8}$$

we see that

$$0\leqq E_M\leqq 1, \tag{9}$$

and our problem is to determie the maximum of the inhomogeneous form.

3. DYNAMIC PROGRAMMING APPROACH

To detemine the maximum of the inhomogeneous form (8), let us define the auxiliary sequence of function

$$f_M(z)=\max_{(An)}[2\sum_{n=0}^{M-1}A_n\psi_n+2zA^M-\sum_{n,m=0}^{M}A_nA_m\varphi_{n-m}] \tag{10}$$

We wish to determine $f_M(0)$ and the $\{A_n\}$ at which the maximum is attained.

We see that a measure of the effictiveness of the filter output

$$\sum_{k=1}^{M}A_nb_{n-k} \tag{11}$$

in representing the massage a_{k_t} was gisen by $f_M(z)$.

It is an important practical question to decide how large to make M. Unless $f_M(z)$ increases appreciably when M is increased, it is not worth while to increase M. In practice, this make desirable a procedure which given us $f_1(z)$, $f_2(z)$, etc., without undue computational difficulty. Our dymamic pragraming approch attained this object.

It is easy to see that

$$f_0(z)=\max_{A_0}[2zA_0-\varphi_0A_0^2]=\frac{z^2}{\varphi_0} \tag{12}$$

and

$$A_0=\frac{z}{\varphi_0} \tag{13}$$

We now wish to derive a recurrence relation connecting f_M with f_{M-1}. If we fix A_M and the minimize over the other A_n, we obtain by relation

$$\begin{aligned} f_M(z)&=\max_{\{A_M\}}[2zA_M-\varphi_0A_M^2+\max_{\{A_M\}}\{2z\sum_{n=0}^{M-2}\psi_nA_n \\ &\quad -2(\psi_{M-1}-\varphi_1A_n)A_{M-1}-\sum_{n,m=1}^{M-1}\varphi_{n-m}A_mA_n\}] \\ &=\max_{\{A_M\}}[2zA_M-\varphi_0A_M^2+f_{M-1}(\psi_{M-1}-\varphi_1A_M)] \end{aligned} \tag{14}$$

4. PREDICTION THEORY AND DYNDMIC PROGRAMMING

In Sec. 2 the problem of separating a message, represented by a sequence a_n, from a signal, represented by a sequence b_k was considered. There the optimum set of numbers A_n was determined in order that a_k, should be represented as closely as possible by

$$\sum_{n=0}^{M} A_m b_{k-n}$$

In Eq. (1) we utlize b_k and eariar values such as b_{k-1}, b_{k-2}, etc., in deriving a_k. There are situations where on the basis of knowing b_k, b_{k-1}, b_{k-2}, etc., we must use Eq. (1) to represent not a_k but a_{k+s}, where s is a positive integer. Here we have a problem involving not only filtering, that is, the separation of message from noise, but also prediction. In other words, even if there were no noise, there would still be the problem of determining a_{k+s} from the knowledge of a_k, a_{k-1}, etc.

Proceeding as in Sec. 2, we now choose the A_M so as minimize the rms of

$$\mathcal{E}_k = a_{k+s} - \sum_{n=0}^{M} A_n b_{k-n} \tag{15}$$

Instead of Eq. (4), we find

$$I_M = R_a(0) - 2\sum_{n=0}^{M} A_n R_{ba}(n+s) + \sum_{n,m=0}^{M} A_n A_m R_b(m-n). \tag{16}$$

In determining the effectiveness of Eq. (19) in representing we get now, instead of Eq. (7) and (8),

$$I_M = 1 - E_M,$$

$$E_M = +2\sum_{n=0}^{M} A_n \varphi_{n+s} - \sum_{n,m=0}^{M} A_n A_m \varphi_{m-n}, \tag{17}$$

where φ_k and ψ_k are defined as Sec. 2.

The iteration formulas given in Sec. 3 can also be generalized to cover the case of predicting together with filtering, and we now turn to this problem.

In place of Eqs. (11), (12) and (13), we have

$$f_0(z) = \max_{A_0}[2zA_0 - \varphi_0 A_0^2] = \frac{z^2}{\varphi_0},$$

$$A_0 = \frac{z}{\varphi_0},$$

$$f_M(z) = \max_{A_M}[2zA_M - \varphi_0 A_M + f_{M-1}(\psi_{M-1+s} - \varphi_1 A_M)]. \qquad (18)$$

We observe that the only difference between these equstions and Eq. (13) is in the index of ψ which is now increased by s.

5. CASE STUDY

Let x_k be the temperature difference from monthly mean value of temperature at Akita and φ_k be its serial correlation coefficients. (Fig. 2)

Our problem is determining x_{k+s} from knowledge of t_k. t_{k-1}, etc, and φ_k, where s is a positive integer. In other words we have a problem of temperature forecasting.

In this case, there were no noise. So, we have the relation $\varphi_k = \psi_k$. By Eq. (18), we have

$$f_M(z) = \max_{A_M}[2zA_M - \varphi_0 A_M^2 + f_{M-1}(\varphi_{M-1+s} - \varphi_1 A_M)],$$

$$(M = 1, 2, 3, \cdots)$$

$$f_0(z) = \frac{z^2}{\varphi_0},$$

$$A_0 = \frac{z}{\varphi_0}, \tag{19}$$

Determining $f_M(t)$ and A_M by means of the methods of successive approximation we obtain Table 1 and Table 2. (Fig. 3. Fig. 4.)

Let us be $k=1941$ January, then we have by references to the appendix,

$$x_{k+1} = \sum_{n=0}^{3} A_n x_{k-n} = 0.66,$$

$$x_{k+2} = \sum_{n=0}^{3} A_n x_{k-n} = 0.181, \tag{20}$$

$$x_{k+3} = \sum_{n=0}^{2} A_n x_{k-n} = 0.03.$$

6. DISCUSSION

A simple methods presented by this paper can be used to treat a large class of prediction theory of nonstationary stochastic processes of multiple time series.

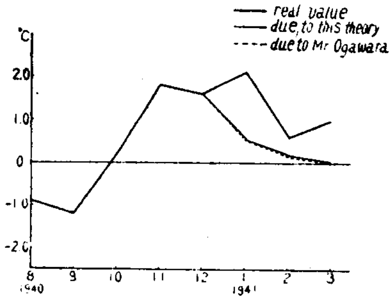
A full acount will occur elsewhere.

BIBLIOGRAPHY

[1] N. Levinson, The Wiener rms (Roots Mean Squars) error criterion in filter design and prediction, *Trans. Math. Phy.* Vol. XXV, No. 4. January (1949).

[2] R. Bellman, *Dynamic Programing*, Princeton University Press, (1957).

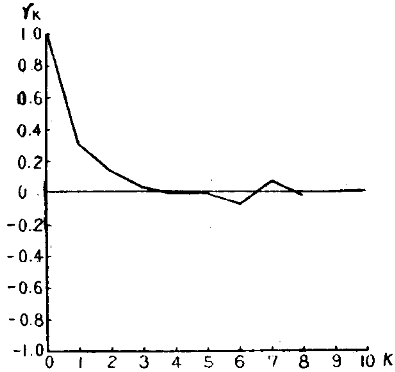
[3] R. Bellman, On some Applications of Dynami Programing to Matrix



k (1941)	1	2	3
due to Mr. Ogawara	0.55	0.16	0
due to this theory	0.66	0.186	0 0.[illegible]

Fig. 1.

An example of extrapolation of monthly mean temperature of Akita



K	0	1	2	3	4	5
γ_K	1.	0.30	0.13	0.02	0.	0.

Fig. 2.

An example of serial correlation coefficients of monthly mean Temperatnre of Akita (due to Mr. Ogawara)

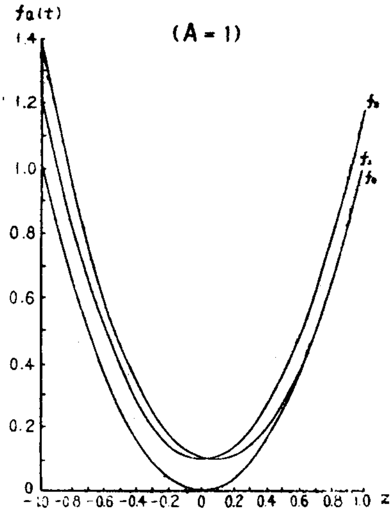


Fig. 3. (a)

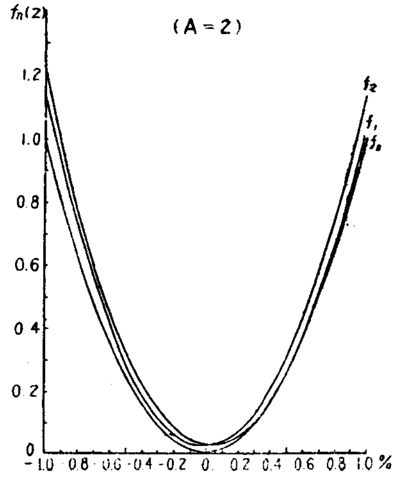


Fig. 3. (b)

(s= 3)

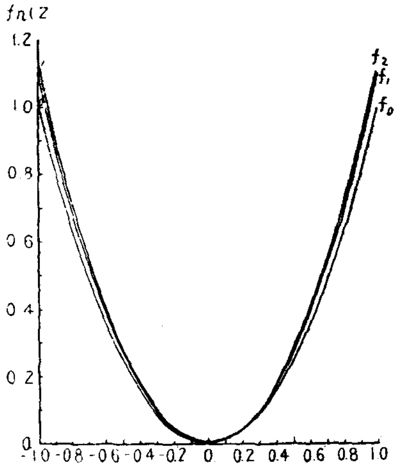


Fig. 3. (c)

(s= 1)

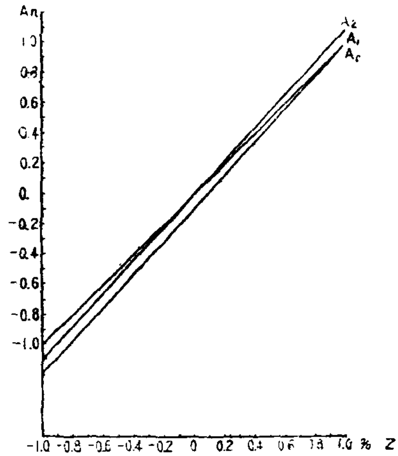


Fig. 4. (a)

(s= 2)

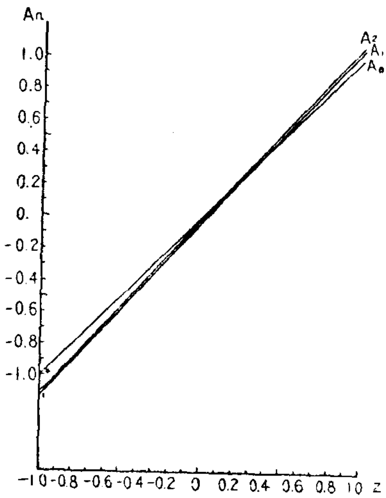


Fig. 4. (b)

(s= 3)

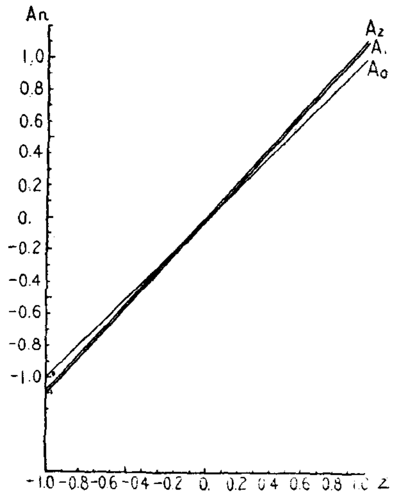


Fig. 4. (c)

Table 1. (a)

s=1

z \ $f_M(z)$	$f_0(z)$	$f_1(z)$	$f_2(z)$
1.	1.	1.	1.1740
0.9	0.81	0.8109	0.9682
0.8	0.64	0.6439	0.7848
0.7	0.49	0.4989	0.6208
0.6	0.36	0.3758	0.4752
0.5	0.25	0.2678	0.3570
0.4	0.16	0.1956	0.2592
0.3	0.09	0.1384	0.1828
0.2	0.04	0.1032	0.1298
0.1	0.01	0.0910	0.1002
0.	0.	0.0987	0.0900
−0.1	0.01	0.1296	0.1021
−0.2	0.04	0.1824	0.1418
−0.3	0.09	0.2571	0.1948
−0.4	0.16	0.3538	0.2786
−0.5	0.25	0.4725	0.3850
−0.6	0.36	0.6131	0.5090
−0.7	0.49	0.7758	0.6514
−0.8	0.64	0.9604	0.8292
−0.9	0.81	1.1670	1.0174
−1	1.	1.3956	1.2300

Table 2. (a)

s=1

z \ $f_M(z)$	A_0	A_1	A_2
1.	1.	1.	1.0950
0.9	0.9	0.8901	0.9840
0.8	0.8	0.7802	0.8730
0.7	0.7	0.6703	0.7620
0.6	0.6	0.5604	0.6510
0.5	0.5	0.4505	0.5400
0.4	0.4	0.3406	0.4290
0.3	0.3	0.2307	0.3180
0.2	0.2	0.1208	0.2070
0.1	0.1	0.0109	0.0960
0.	0.	−0.0989	−0.0150
−0.1	−0.1	−0.2087	−0.1256
−0.2	−0.2	−0.3186	−0.2370
−0.3	−0.3	−0.4295	−0.3180
−0.4	−0.4	−0.5384	−0.4370
−0.5	−0.5	−0.6483	−0.5700
−0.6	−0.6	−0.7582	−0.6800
−0.7	−0.7	−0.8681	−0.7910
−0.8	−0.8	−0.9780	−0.9020
−0.9	−0.9	−1.0879	−1.0130
−1.0	−1.0	−1.1978	−1.1240

theory, *Illinois Journal of Mathematics*, 1, No. 2. June (1957).

[4] M. OGAWARA : On Stochastic prediction formula, in Japan Memories of Meteorological Resarch Institute No. 24 (1947)

APPENDIX

On Some Applications of Dynamic Programing to Numerical Soiution of Linear Equations,

The purpose of this appendix is to discuss some applications of the function technique of dynamic programing to some questions of

Table 1. (b)

$s=2$

$f_M(z)$ \ z	$f_0(z)$	$f_1(z)$	$f_2(z)$
1.0	1.	1.0317	1.1419
0.9	0.81	0.8315	0.9279
0.8	0.64	0.6533	0.7387
0.7	0.49	0.4970	0.5693
0.6	0.36	0.3627	0.4247
0.5	0.25	0.2504	0.3047
0.4	0.16	0.1601	0.2000
0.3	0.09	0.0917	0.1198
0.2	0.04	0.0454	0.0634
0.1	0.01	0.0210	0.0529
0.	0.	0.0185	0.0182
−0.1	0.01	0.0381	0.0272
−0.2	0.04	0.0796	0.0560
−0.3	0.09	0.1432	0.1097
−0.4	0.16	0.2287	0.1823
−0.5	0.25	0.3361	0.2845
−0.6	0.36	0.4656	0.4046
−0.7	0.49	0.6170	0.5475
−0.8	0.64	0.7904	0.7151
−0.9	0.81	0.9858	0.9026
−1.0	1.	1.2032	1.1098

Table 2. (b)

$s=2$

A_M \ z	A_0	A_1	A_2
1.0	1.	1.0560	1.0882
0.9	0.9	0.9460	0.9772
0.8	0.8	0.8360	0.8662
0.7	0.7	0.7260	0.7552
0.6	0.6	0.6160	0.6443
0.5	0.5	0.5170	0.5333
0.4	0.4	0.3970	0.4223
0.3	0.3	0.2870	0.3113
0.2	0.2	0.1870	0.2004
0.1	0.1	0.0670	0.0894
0.	0.	−0.0430	−0.0216
−0.1	−0.1	−0.1530	−0.1326
−0.2	−0.2	−0.2630	−0.2423
−0.3	−0.3	−0.3730	−0.3545
−0.4	−0.4	−0.4830	−0.4655
−0.5	−0.5	−0.5920	−0.5765
−0.6	−0.6	−0.7020	−0.6874
−0.7	−0.7	−0.8120	−0.7984
−0.8	−0.8	−0.9220	−0.9094
−0.9	−0.9	−1.0320	−1.0204
−1.0	−1.0	−1.1420	−1.1313

numerical solutions of linear equation.

We shall first consider the solution of a system of liner equations

$$\sum_{j=1}^{N} a_{ij}x_j = C_i, \quad (i=1,\ 2,\ \cdots,\ N). \tag{1}$$

where $A=(a_{ij})$ be a positive definite symmetric matrix.

Then we obtain that the problem of solving (1) is equivalent to determining the absolute minimum of the form

Table 1. (c)

$s=3$

z \ $f_M(z)$	$f_0(z)$	$f_1(z)$	$f_2(z)$
1.0	1	1.0864	1.1065
0.9	0.9	0.8784	0.9033
0.8	0.8	0.6924	0.7027
0.7	0.7	0.5304	0.5460
0.6	0.6	0.3884	0.3972
0.5	0.5	0.2680	0.2782
0.4	0.4	0.1704	0.1789
0.3	0.3	0.0954	0.0995
0.2	0.2	0.0424	0.0448
0.1	0.1	0.0104	0.0100
0.	0.	0.0004	0.0000
−0.1	0.1	0.0124	0.0097
−0.2	0.2	0.0464	0.0443
−0.3	0.3	0.1034	0.0987
−0.4	0.4	0.1814	0.1729
−0.5	0.5	0.2814	0.2719
−0.6	0.6	0.4034	0.3906
−0.7	0.7	0.5474	0.5342
−0.8	0.8	0.7144	0.7026
−0.9	0.9	0.9024	0.8958
−1.0	1.0	1.1124	1.1040

Table 2. (c)

$s=3$

z \ A_M	A_0	A_1	A_2
1.0	1.0	1.0920	1.1076
0.9	0.9	0.9820	0.9966
0.8	0.8	0.8730	0.8856
0.7	0.7	0.7630	0.7746
0.6	0.6	0.6530	0.6637
0.5	0.5	0.5430	0.5527
0.4	0.4	0.4330	0.4417
0.3	0.3	0.3230	0.3348
0.2	0.2	0.2130	0.2198
0.1	0.1	0.1030	0.1088
0.	0.	−0.0070	0.0022
−0.1	−0.1	−0.1160	−0.1088
−0.2	−0.2	−0.2260	−0.2198
−0.3	−0.3	−0.3360	−0.3348
−0.4	−0.4	−0.4460	−0.4471
−0.5	−0.5	−0.5560	−0.5527
−0.6	−0.6	−0.6660	−0.6637
−0.7	−0.7	−0.7760	−0.7746
−0.8	−0.8	−0.8860	−0.8856
−0.9	−0.9	−0.9960	−0.9966
−1.0	−1.0	−1.1050	−1.1076

$$Q_N(x)=\sum_{ij=1}^{N} a_{ij}x_ix_j-2\sum_{i=1}^{N} C_ix_i \tag{2}$$

Define this minimum to be

$$f_N(x)=\min Q_N(x). \tag{3}$$

and obtain a recurrence relation connecting f_N and f_{N-1},

$$f_k(x)=\min_{x_k}\,[a_{kk}x_k^2-2zx_k+f_{k-1}(C_{k-1}-a_{k-1,k}x_k)]. \qquad (4)$$

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

To distinguish between the various values A_m assumes as M changes, we introduce the more specific notation, A_n^M.

Let us

$$A_M^{(M)}=A_M,$$

and determining $A_M^{(M)}$, $A_{M-1}^{(M)}$, $\cdots$, $A_{M-t+1}^{(M)}$, then we have

$$f_{M-1}(z)=\max_{A_{M-t}}\,[2zA_{M-t}^{(M)}-\varphi_0A_{M-1}^{(M)2}-f_{M-t+1}(\psi_{M-t-1}-\sum_{n=0}\varphi_nA_{M-t+s})]$$

where

$$A_{M-t}^{(M)}=A_{M-t}.$$