<table>
<thead>
<tr>
<th>1.0</th>
<th>1.1</th>
<th>1.25</th>
<th>1.4</th>
<th>1.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0</td>
<td>2.2</td>
<td>2.25</td>
<td>2.4</td>
<td>2.5</td>
</tr>
</tbody>
</table>

NATIONAL BUREAU OF STANDARDS

MICROCOPY RESOLUTION TEST CHART
# Methodology for Stochastic Modeling

**Authors:** Herbert E. Cohen

**Performing Organization Name and Address:**

US Army Materiel Systems Analysis Activity
ATTN: AMXSY-MP
Aberdeen Proving Ground, MD 21005-5071

**Report Date:** January 1985

This report supersedes GWD Interim Note G-144.
The requirement to develop stochastic mathematical models arises across the whole range of engineering and applied research where observations are made of a physical process, corrupted by noise, and it is desired to determine the underlying nature, either in time or frequency, of the observed phenomenon. This paper presents some of the current approaches in stochastic modeling, including adaptive autoregressive (AR) models, which have been found to be useful in this area.
ACKNOWLEDGEMENT

The US Army Materiel Systems Analysis Activity (AMSAA) wishes to recognize the following individual for his contribution to this report:

Peer Reviewer: Dr. Keats Pullen - CSD

The author wishes to express his appreciation to John Groff and John Kramar of GWD, AMSAA, for their encouragement on this report. Also, the support and patience of the Reports Processing staff for the many hours they labored over this report.
## CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>3</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>4</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>7</td>
</tr>
<tr>
<td>2. BOX–JENKINS TECHNIQUE</td>
<td>8</td>
</tr>
<tr>
<td>3. OVERDETERMINED RATIONAL MODELS</td>
<td>13</td>
</tr>
<tr>
<td>4. ORDER DETERMINATION</td>
<td>22</td>
</tr>
<tr>
<td>5. ADAPTIVE MODELING</td>
<td>24</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>31</td>
</tr>
<tr>
<td>DISTRIBUTION</td>
<td>33</td>
</tr>
</tbody>
</table>
METHODOLOGY FOR STOCHASTIC MODELING

1. INTRODUCTION

The requirement to develop stochastic mathematical models arises across the whole range of engineering and applied research where observations are made of a physical process, corrupted by noise, and it is desired to determine the underlying nature, either in time or frequency, of the observed phenomenon. This paper presents some of the current approaches in stochastic modeling which have been found to be useful in this area.

Consider a system in which an input \( a_t \) produces an output \( y_t \) as shown in Figure 1.

\[
y_t = \phi_0 + \sum_{k=1}^{\infty} \phi_k y_{t-k} - \sum_{k=0}^{p} \theta_k a_{t-k}
\]

Figure 1

In general we entertain the idea of expressing the output \( y_t \) in terms of its past history \( y_{t-1}, y_{t-2}, \ldots, y_{t-p} \) and the current and past history of the excitation \( a_t, a_{t-1}, a_{t-2} \ldots a_{t-q} \) as shown

\[
y_t = \sum_{k=0}^{\theta_k} \theta_k a_{t-k} - \sum_{k=1}^{\phi_k} \phi_k y_{t-k}
\]

with \( \phi_0 = 1 \)

where the coefficients \( \theta_k \) and \( \phi_k \) are to be estimated as well as the parameters p and q. Equation (1) represents an autoregressive moving average (ARMA) model of order \((p,q)\). It is advantageous to assume that the input or excitation \( a_t \) is a zero mean gaussian white noise. For the case where all \( \phi_k = 0 \) we have what is called a moving average (MA) model.

7
\[ y_t = \sum_{k=0}^{q} \theta_k a_{t-k} \]  

(2)

and for \( \theta_k = 0 \) \((k>1)\) we have an autoregressive (AR) model

\[ y_t = \sum_{k=1}^{p} \phi_k y_{t-k} + a_t \quad \text{(for } \theta_0 = 1) \]  

(3)

2. BOX-JENKINS TECHNIQUE

The Box-Jenkins [1] approach for time series analysis of observed data is widely used and is based on the statistical properties of the data, namely, the autocorrelation function (ACF) and the partial autocorrelation function (PAC) for each of the AR and MA models.

a. Autoregressive (AR) model of order \( p \): An autoregressive model of order \( p \), AR(\( p \)), is characterized as follows:

(1) The autocorrelation function (ACF) is either a decreasing exponential or a damped sine wave.

(2) The partial autocorrelation function (PCF) is non zero for lags less or equal to \( p \) and zero for lags greater than \( p \), i.e., the PCF cutoff after lag \( p \).

When we multiply equation (3) by \( y_{t-k} \) \((k>0)\) and take expectation we obtain the Yule-Walker equations

\[ \rho_k = -\phi_1 \rho_{k-1} - \phi_2 \rho_{k-2} - \cdots - \phi_p \rho_{k-p} \]  

(4)

\((k = 1, 2, \ldots, p)\)

where \( \rho_k \) is the autocorrelation function of the process \( y_t \) of lag \( k \). The Yule-Walker equations can be solved for \( \phi_j \) when we replace the theoretical autocorrelation \( \rho_k \) by the estimated autocorrelation \( r_k \), i.e., where

\[ \hat{\rho}^{-1} = \sum_{p=1}^{p} \Sigma_{p} \]  

(5)
Since the autocorrelation function for an AR(p) is infinite in extent, the partial autocorrelation $\hat{\phi}_{k,k}$, also called "reflection coefficient," can be obtained from the Levinson-Durbin Algorithm [2, 3] namely

$$\hat{\phi}_{k,k} = \begin{cases} r_1 & \quad \text{if } k = 1 \\ r_k - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_{k-j} & \quad \text{if } k = 2, 3, \ldots, L \\ 1 - \sum_{j=1}^{k-1} \hat{\phi}_{k-1,j} r_j & \quad \text{if } k = 2, 3, \ldots, L \end{cases}$$

(6)

where

$$\hat{\phi}_{k,j} = \hat{\phi}_{k-1,j} - \hat{\phi}_{k-1} \hat{\phi}_{k-2}$$

(j=1,2,..., k-1.)

where $r_j$ is the autocorrelation of the sampled data of lag $j$ with the initial values $\phi_{21}$ and $\phi_{22}$ given by

$$\hat{\phi}_{21} = \frac{r_1 (1-r_2)}{1 - r_1^2}$$

$$\hat{\phi}_{22} = \frac{r_2 - r_1^2}{1 - r_1^2}$$

(7)

The 95 percent confidence interval, or $(+2\pi )$, for $r_k$ and $\phi_{k,k}$ are given by

$$\text{Var} [r_k] \sim \frac{1}{N} \left[ 1 + 2 \sum_{i=1}^{q} r_i^2 \right]^{-\frac{1}{2}}$$

(8)

(Bartell's approximation)
and

\[ \hat{\phi}_{kk} = \frac{1}{\sqrt{N}} ; \quad (k > q) \quad (9) \]

which are useful quantities in determining order of the model.

It is advantageous to plot the ACF \((r_k)\) and PCF \((\phi_{kk})\) for lag 1 up to any specified lag \(k\) with the corresponding 95 percent confidence interval \(\pm 2 \sigma\).

The selection of the form of the model is determined from the behavior of the autocorrelation function and the partial autocorrelation function as follows:

a. If the autocorrelation function decays either exponentially or sinusoidally, and the partial autocorrelation function ("reflection coefficient") cuts off after a lag \(p\), then the model is AR \((p)\) i.e., no moving average terms exist.

b. If the autocorrelation function cuts off after lag \(q\), and the partial autocorrelation function decays either exponentially or sinusoidally, then we have a moving average model MA\((q)\) i.e., no autoregressive terms exist.

c. Should there be no cutoff in either the autocorrelation function or partial autocorrelation function, and they independently decay either exponentially or sinusoidally, then the model is ARMA \((p, q)\) i.e., autoregressive-moving average of order \(p, q\) respectively, so that there exist both autoregressive and moving average terms.

The Box-Jenkins procedures for ARMA modeling are given as follows:

Program #1:

Define expected value and variance as follows:

\[
\text{(Expected value)} \quad x = \frac{1}{N} \sum_{t=1}^{N} x_t
\]

\[
\text{(Variance)} \quad s_x^2 = c_0
\]

Autocovariance function:
\[ c_k = \frac{1}{N-k} \sum_{t=1}^{N} (x_t - \bar{x})(x_{t+k} - \bar{x}) \]

\[ k = 0, 1, 2, \ldots, K \]

Autocorrelation function:

\[ r_k = \frac{c_k}{c_0} \quad k = 0, 1, \ldots, K \]

Partial Autocorrelation function:

\[ \hat{\phi}_{k,l} = \left\{ \begin{array}{ll}
\hat{r}_1 & \text{if } l = 1 \\
\hat{r}_l - \frac{\sum_{j=1}^{l-1} \hat{\phi}_{l-1,j} \hat{r}_{l-j}}{1 - \sum_{j=1}^{l-1} \hat{\phi}_{l-1,j} \hat{r}_j} & \text{if } l = 2, 3, 4 \ldots L
\end{array} \right. \]

where

\[ \hat{\phi}_{l,j} = \hat{\phi}_{l-1,j} - \hat{\phi}_{l-1} \hat{\phi}_{l-1,j} \quad j = 1, 2, \ldots, l-1 \]

Each of the autocorrelation functions and the partial autocorrelation functions are plotted as a function of the lags.

Program #2:

For an ARMA (p,q) model represented by

\[ x_t = \phi_1 x_{t-1} + \ldots + \phi_p x_{t-p} + \theta_0 a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \ldots - \theta_q a_{t-q} \]

we estimate the AR parameters \( \hat{\phi} = (\phi_1, \phi_2, \ldots, \phi_p) \) for \( p > 0 \) by solving the linear equation

\[ A \hat{\phi} = \bar{x} \quad (p > 0) \]

where

\[ A_{ij} = \frac{c_{q+i-j}}{|q+i-j|} \quad \text{for } p+q > 0, K > p+q \]

\[ x_i = c_{q+i} \quad \text{for } c_k, k = 0, 1, \ldots, K \]

\[ \hat{\phi}_0 = (\phi_{10}, \phi_{20}, \ldots, \phi_{p0}) \]
To estimate $\hat{\theta}$ of the moving average parameters we use the autocovariance of $x_t$, namely $c_k$, and calculate

$$c_j = \begin{cases} \sum_{i=0}^{p} \sum_{k=0}^{p} \phi_i \phi_k \sigma_i \sigma_{j+i-k} & p>0 (\phi_{00} = -1) \\ c_j & p=0 \\
\end{cases}$$

(j=0, 1, ..., q)

Box [1] then applies to the Newton-Raphson algorithm

$$\tau^{i+1} = \tau^i - h$$

where

$$T_i h = f_i$$

To calculate the vectors $\tau^{i+1}$ at the $(i+1)$ iteration from its value $\tau^i$ at the $i$th iteration

$$\tau = (\tau_0, \tau_1, ..., \tau_q)$$

$$f_j = \sum_{i=0}^{q-j} \tau_i \tau_{i+j} - c_j$$

$$f^r = (f_0, f_1, ..., f_q)$$

$$T = \begin{bmatrix} \tau_0 & \tau_1 & \cdots & \tau_q \\ \tau_1 & \tau_2 & \cdots & \tau_q \\ \vdots & \vdots & \ddots & \vdots \\ \tau_q & \cdots & \cdots & 0 \end{bmatrix} + \begin{bmatrix} \tau_0 & \tau_1 & \cdots & \tau_{q-1} \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \tau_0 \end{bmatrix}$$

Newton-Raphson algorithm converges when $|f_j| < \epsilon$, $j=0, 1, ..., q$ for some prescribed $\epsilon$ with

$$\hat{\theta}_{j0} = -\tau_j/\tau_0 \quad j = 1, 2, ..., q$$

$$\hat{\theta}_{00} = \begin{cases} \bar{x} (1 - \sum_{i=1}^{p} \hat{\theta}_{i0}) & p>0 \\ \bar{x} & p=0 \end{cases}$$

12
and the estimate $\hat{\sigma}_a^2$ of white noise variance

$$\begin{cases} \tau_i^2 & q > 0 \\ c_0 - \sum_{i=1}^p \phi_i c_1 & q = 0 \end{cases}$$

We therefore have calculated all the parameters in the ARMA $(p, q)$ model.

3. OVERDETERMINED RATIONAL MODELS

Cadzow [4] has developed spectral estimates using an overdetermined set of Yule-Walker equations for a rational model. The procedures are summarized below:

1. For spectral models employing exact autocorrelation lag information.
   a. Moving average model (MA), a Blackman-Tukey approach is used for determining the spectrum $S_X(e^{j\omega})$ namely

   $$S_X(e^{j\omega}) = \sum_{n=-q}^{q} w(n) r_X(n) e^{-j\omega n}$$

   where $w(n)$ is a desired data window, a wide variety of which can be located in [5, 6, 7].

   $r_X(n)$ is the exact autocorrelation function of lag $n$ for the observed signal $x(t)$.

   b. Autoregressive model (AR) - Having determined the order $p$ of a purely autoregressive model from Box-Jenkins technique by studying partial autocorrelations, we form a $(p+1)x(p+1)$ AR autocorrelation matrix $R$ whose elements are

   $$R(i,j) = r_X(i-j) \quad 1 \leq i \leq p + 1$$
   $$1 \leq j \leq p + 1$$

   Solve for $R_1 = |b_0|^2 e_1$

   where $a = (1, a_1, a_2, \ldots, a_p)^T$

   $e_1 = (1, 0, \ldots, 0)^T$
and $b_0$ is selected so that the first component of $a$ is one.

Using the value of $a$ just calculated, the spectrum $S_x(e^{j\omega})$ is given by

$$S_x(e^{j\omega}) = \frac{b_0}{\left|1 + a_1 e^{-j\omega} + \ldots + a_p e^{-jp\omega}\right|^2}$$

(13)

A more efficient method for calculating recursively the components of $a$ and at the same time determining the order of the AR model is the Levinson-Durbin algorithm. The procedure for coefficient determination is as follows:

Step #1

$$a_1^{(1)} = -\frac{r_x(1)}{r_x(0)}$$

(14)

$$|b_0^{(1)}|^2 = [1 - |a_1^{(1)}|^2] r_x(0)$$

where we define $a_j^{(k)}$ as the $j$th coefficient associated with the $k$th iteration.

Step #2. For $k = 2, 3, 4 \ldots$

$$a_k^{(k)} = -\left[ r_x(k) + \sum_{m=1}^{k-1} a_m^{(k-1)} r_x(k-m) \right] / |b_0^{(k-1)}|^2$$

(15)

$$a_i^{(k)} = a_i^{(k-1)} + a_k^{(k)} a_{k-1}^{(k-1)} \quad 1 \leq i < (k-1)$$

$$|b_0^{(k)}|^2 = [1 - |a_k^{(k)}|^2] |b_0^{(k-1)}|^2$$

The iteration stops when $b_0^{(k)}$ assumes a constant value or a sufficiently small value after some value of "$k" which thus establishes the order of the model.

c. Autoregressive moving average (ARMA) - In general an ARMA model takes the form

$$x(n) + \sum_{k=1}^{p} a_k x(n-k) = \sum_{k=0}^{q} b_k \epsilon(n-k)$$

(16)
where \( e(n) \) is normalized white noise. For a model to be causal the Yule-Walker equations take on a simple form for \( n > q \)

\[
\sum_{k=0}^{p} a_k r_x(n-k) = 0 \quad n > q + 1
\]

(17)

For an overdetermined model, the extended Yule-Walker equations for \( q+1 \leq n \leq q+t \)
i.e., \( t \) linear equations in \( p \) autoregressive parameter unknowns, is given by

\[
\begin{bmatrix}
    r_x(q+1) & r_x(q) & \ldots & r_x(q-p+1) \\
    r_x(q+2) & r_x(q+1) & \ldots & r_x(q-p+2) \\
    \vdots & \vdots & \ddots & \vdots \\
    -r_x(q+t) & r_x(q+t-1) & \ldots & r_x(q-p+t)
\end{bmatrix}
\begin{bmatrix}
    1 \\
    a_1 \\
    a_2 \\
    \vdots \\
    a_p
\end{bmatrix}
= \begin{bmatrix}
    0 \\
    0 \\
    \vdots \\
    0
\end{bmatrix}
\]

(18)

or simple

\[ R_1 a = 0 \]

where \( R_1 \) is a Toeplitz matrix

\[ R_1(i, j) = r_x(q+1+i-j) \quad 1 < i < t \]

where \( a = (1, a_1, a_2, \ldots, a_p) \)

are the \((p+1)\) autoregressive coefficients and

\[ r_x(n) = E\{x(n+m)\overline{x}(m)\} \]

Using the above developed structure the algorithm for ARMA model is as follows

1. Form the \( t \times (p+1) \) ARMA autocorrelation matrix \( R_1 \). Define \( R_1^* \) as the transpose complex conjugate matrix.

2. If the rank \((R_1^* R_1) < p+1\) then solve

\[ R_1^* R_1 a = 0 \]

using singular value decomposition (SVD) to determine "p". Discussion on SVD will be presented later in this paper.
3. If the rank \((R_1^* R_1) = p+1\) then we solve
\[
R_1^* R_1 \mathbf{a} = \alpha \mathbf{e}_1 \tag{19}
\]
where \(\mathbf{e}_1 = (1, 0, \ldots, 0)^T\) and \(\alpha\) is a normalizing factor so that the first component of \(\mathbf{a}\) is one or we can find the minimum eigenvalue \(\lambda\) and associated eigenvector \(\mathbf{x}_k\).
\[
\lambda_k < \lambda_{k+1} \quad \text{(eigenvalues)}
\]
so that
\[
\frac{1}{\mathbf{x}_1 (1)} \mathbf{x}_1 \tag{20}
\]
which now gives the AR coefficients.

4. The moving average component to the ARMA model is determined using the AR coefficients.
\[
r_s(n) = \sum_{k=0}^{p} \sum_{m=0}^{p} a_k \overline{a}_m r_x(n+m-k) \tag{21}
\]
with
\[
r_s(n) = 0 \quad n > q
\]
so that the moving average component of the spectrum is given by
\[
S_{MA}(\omega) = \sum_{n=-q}^{q} r_s(n) e^{-j\omega n} \tag{22}
\]
To find the zero's \(z_k\) of \(S_{MA}(\omega)\)
\[
S_{MA}(\omega) = |b_0|^2 \prod_{k=1}^{q} (1-z_k e^{-j\omega}) (1 - \overline{z}_k e^{j\omega}) \tag{23}
\]
where we require \(|z_k| < 1\); all the zero's of \(S_{MA}(\omega)\) will be complex conjugate pairs. With
\[
\sum_{k=0}^{q} b_k e^{-j\omega k} = b_0 \prod_{k=0}^{q} (1 - z_k e^{-j\omega}) \tag{24}
\]
we can compare coefficients of $e^{-j\omega k}$ to determine $b_k$. Thus for the rational ARMA model

$$S_x(e^{j\omega}) = \frac{\sum_{n=-q}^{q} r_s(n) e^{-j\omega n}}{1 + a_1 e^{-j\omega} + \ldots + a_p e^{-jp\omega}}$$  \hspace{1cm} (25)$$

where $S_x(e^{j\omega})$ is the spectrum.

2. For spectral models using only a finite number of observation so that we have only estimates of the autocorrelation function, Cadzow [4] has identified the following procedures.

a. Moving average model - We again use Blackman-Tukey procedure, i.e.,

$$S_x(e^{j\omega}) = \sum_{n=-q}^{q} w(n)\hat{r}_x(n)e^{-j\omega n}$$  \hspace{1cm} (26)$$

where

$$\hat{r}_x(n) = \frac{1}{N} \sum_{k=1}^{N} x(k+n)\overline{x}(k)$$

$$\hat{r}_x(n) = 0 \quad \text{outside} \quad -q < n < q$$

$$w(n) = \begin{cases} 
2n/(N-1) & 0 < n < (N-1)/2 \\
2-2n/(N-1) & (N-1)/2 < n < (N-1) \\
0 & \text{otherwise}
\end{cases}$$  \hspace{1cm} (27)$$

w(n) is known as the Bartlett window.

Any "suitable" window $w(n)$ can be used to weigh $\hat{r}_x(n)$; for the present, we use $w(n)=1$.

b. Autoregressive model (AR) -

1. Compute $\hat{R}(i,j) = r_x(i-j)$ for $1 < i, j < p + 1$

$$\hat{R}(i,j) = \frac{1}{N-|i-j|} \sum_{k=1}^{N} x(k+i-j)\overline{x}(k)$$  \hspace{1cm} (28)$$
2. $\hat{RA} = |b_0|^2 e_1$

where $b_0$ is selected to normalize first component of $a$, i.e., $a(1) = 1$.

3. $S_{AR}(e^{i\omega}) = \frac{b_0}{1 + \sum_{n=1}^{p} a_n e^{-j\omega n}}^2$ (29)

c. Autoregressive moving average (ARMA) model - For an ARMA model the AR components are calculated first, as follows:

1. Compute the autocorrelation matrix estimate $\hat{R}_1$ whose elements $\hat{R}_1(i,j)$ are given by

$$\hat{R}_1(i,j) = r_x(q+1+i-j)$$ (30)

for $1 < i < t$ and $1 < j < p+1$ so that $\hat{R}_1$ is a $t \times (p+1)$ matrix. The rank of $\hat{R}_1$ is equal to the min $(t, p+1)$. If the number of eigenvalues are computed for $R_1^* R_1$ it will be equal to min $(t, p+1)$ so that if $t \gg p+1$ we can be assured that rank $R_1^* R_1 = p+1$. Hence $p$, the order of AR component of our model, is determined. The value of $t$ will be bounded, typically selected, to be $p < t < N - q - 1$.

2. Solve for $a$ in equation $\hat{R}_1^* W \hat{R}_1 = \alpha e_1$ where $\alpha$ is a normalizing constant so that 1st component of $a^0$ is 1 and $W$ is a diagonal weighting matrix which is taken as $I$ (identity) where $\hat{R}_1$ is unbiased.

3. $\hat{R}_1^* W \hat{R}_1(i,j) = \sum_{m=1}^{t} w(m) \hat{r}(q+m+1) \hat{r}(q+m+i-j)$ (31)

for $1 < i, j < p+1$ with $w(m)$ corresponding to diagonal elements of diagonal weighting matrix $W$. 

18
4. Compute \((\hat{R}_1 \hat{W} \hat{R}_1)^{-1}\) and solve for \(\hat{a}_0\)

\[
\hat{a}_0 = \frac{(\hat{R}_1 \hat{W} \hat{R}_1)^{-1} \hat{e}_1}{\hat{a}_0(1)}
\]  

(32)

which normalizes \(a\) so that 1st component of \(\hat{a}_0\) is 1.

\[
|A_p(e^{j\omega})|^2 = |1 + \sum_{k=1}^{p} a_k e^{-j\omega k}|^2
\]

(33)

Moving average model. Several methods exist for computing the moving average coefficients which require spectral factorization of the moving average spectrum.

Method #1 - The system to be described generates a moving average "residual signal" obtained by passing the observed data \(x(t)\) through a filter whose transfer function is the denominator of the ARMA model, hence an AR\((p)\), thus producing a moving average output whose MA spectrum \(B_q\) corresponds to the original ARMA model. System is described in Figure 2.

![Figure 2.](image)

Define \(s_f(n) = \sum_{k=0}^{p} \hat{a}_k x(n-k)\)  

\(p+1 < n < N\)  

(34)

\(s_b(n) = \sum_{k=0}^{p} a_k x(n+k)\)  

\(1 < n < N-p\)
\[ \hat{r}_s(n) = \frac{1}{N-p-n} \sum_{k=1}^{N-p-n} \left[ s_f(n+p+k)\bar{s}_f(p+k)+s_b(n+k)\bar{s}_b(k) \right] \]  

(35)

for 0<n<q and \( r_s(-n) = r_s(n) \). It should be noted that for a moving average model \( r_s(n)=0 \) for \( n>q+1 \). The spectrum \( S_{MA}(e^{j\omega}) \) is given as follows

\[ S_{MA}(e^{j\omega}) = |B_q(e^{j\omega})|^2 = \sum_{n=-q}^{q} w(n) \hat{r}_s(n)e^{-j\omega n} \]  

(36)

where \( w(n) = \left( \frac{N-p-n}{N-p} \right) \left( \frac{q+1-|n|}{q+1} \right) \)  

(37)

For \( w=1 \)

\[ |B_q(e^{j\omega})|^2 = \sigma_q^2 [1+r_s(1) (z^{-1}+z)+\ldots+r_s(q)(z^{-q}+z^q)] \]  

(38)

where \( z = e^{j\omega} \)

In general we can find the zero's \( z_k \) of \( |B_q(e^{j\omega})|^2 = B_q(e^{j\omega}) \bar{B}_q(e^{j\omega}) \) and carry out a spectral factorization

\[ |B_q(e^{j\omega})|^2 = b_0^2 \prod_{k=1}^{q} (1-z_k e^{-j\omega}) (1-\bar{z}_k e^{j\omega}) \]  

(39)

so that

\[ B_q(e^{j\omega}) = b_0 \prod_{k=1}^{q} (1-z_k e^{-j\omega}) \]  

(40)

Since

\[ B_q(e^{j\omega}) = \sum_{k=0}^{q} b_k e^{-j\omega k} \]  

(41)

we have

\[ \sum_{k=0}^{q} b_k e^{-j\omega k} = b_0 \prod_{k=1}^{q} (1-z_k e^{-j\omega}) \]  

(42)
By equating coefficients we determine the values of $b_k$.

Method #2. Assuming the AR parameters $a_k$ are known, we define the autocorrelation function $r_X(n)$ and its causal image $r_X^+(n)$ as follows

$$r_X(n) = E\left\{ x(n+m) x^*(m) \right\} ; \quad n=0, \pm 1, \pm 2, \ldots$$

$$r_X^+(n) = r_X(n)u(n) - \frac{1}{2} r_X(0) \delta(n)$$

where $u(n) = \text{standard unit step} = \begin{cases} 1 & n > 0 \\ 0 & n < 0 \end{cases}$

$\delta(n) = \text{Kronecker delta sequence}$

$$\begin{align*}
\delta(n) &= \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases} \\
\end{align*}$$

and $r_X(-n) = r_X^*(n)$

Noting that $r_X(n) = r_X^+(n) + r_X^+(-n)^*$

we have

$$S_X(e^{j\omega}) = S_X^+(e^{j\omega}) + S_X^+(e^{-j\omega})^*$$

$$= 2 \text{ Re } [S_X^+(e^{j\omega})]$$

which is the spectrum of the signal

and $c(n) = r_X^+(n) + \sum_{k=1}^{p} a_k r_X^+(n-k) \quad 0 < n < \max(q, p)$

$$= 0 \quad \text{ n outside interval}$$

Define

$$C_s(e^{j\omega}) = \sum_{n=0}^{s} c(n) e^{-j\omega n} \quad \text{where } s = \max(q, p)$$
so that

$$C_s(e^{j\omega}) = [1 + \sum_{n=1}^{p} a_n e^{-j\omega n}] \ S_x^+(e^{j\omega})$$  \hspace{1cm} (46)$$

then

$$S_x(e^{j\omega}) = \frac{C_s(e^{j\omega})}{A_p(e^{j\omega})} + \frac{C_s^*(e^{j\omega})}{A_p^*(e^{j\omega})}$$  \hspace{1cm} (47)$$

$$= \frac{A_p^*(e^{j\omega}) (C_s(e^{j\omega}) + C_s^*(e^{j\omega}) A_p(e^{j\omega})}{A_p(e^{j\omega}) A_p^*(e^{j\omega})}$$

However $B_q(e^{j\omega})B_q^*(e^{j\omega}) = A_p^*(e^{j\omega})C_s(e^{j\omega}) + A_p(e^{j\omega})C_s^*(e^{j\omega})$

so that if we find the zero's of $S_x(e^{j\omega})$ identified as $z_k$ and using spectral factorization we can determine the $b_k$ by equating coefficients in

$$\sum_{k=0}^{q} b_k e^{-j\omega k} = b_0 \prod_{k=1}^{q} (1-z_k z^{-1}) \hspace{1cm} \text{for } |z_k| < 1$$  \hspace{1cm} (48)$$

to maintain minimum phase.

4. ORDER DETERMINATION

A fundamental issue in applying the methods that have been presented is that the order of the model needs to be determined. Methods that have been used include:

1. Levinson-Durbin Method for pure AR.

2. Test for autocorrelation function, $r_s(n)=0$ with $n>q+1$ for pure MA.

A new method receiving considerable study is the singular value decomposition (SVD) using the Frobenious norm of a mxn matrix difference, $A-B$,

$$||A - B|| = \left[ \sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij} - b_{ij}|^2 \right]^{1/2}$$  \hspace{1cm} (49)$$
The SVD is carried out as follows:

1. Set \( p_e >> p, q_e >> q \) for \( t >> p \)

2. Form \( R_e \):

\[
R_e = \begin{bmatrix}
  r_x(q_e+1) & r_x(q_e) & \ldots & r_x(q_e-p_e+1) \\
  r_x(q_e+2) & r_x(q_e+1) & \ldots & r_x(q_e-p_e+2) \\
  \vdots & \vdots & \ddots & \vdots \\
  r_x(q_e+t) & r_x(q_e+t-1) & \ldots & r_x(q_e-p_e+t)
\end{bmatrix}
\]

This is a \( t \times (p_e+1) \) matrix which satisfies

\[
R_e \Sigma = \Theta
\]

for which \( R_e = U \Sigma V^* \)

where \( U \) and \( V \) are unitary matrices

3. Determine eigenvalues \( \lambda_{kk} \) of \( R_eR_e^* \)

4. Take \( \sigma_{kk} = \sqrt{\lambda_{kk}} \), called the singular values, which are ordered \( \sigma_1 > \sigma_2 > \ldots > \sigma_h > 0 \) where \( h = \min(t, p_e+1) \)

5. Form the matrix \( R_eR_e^* \) which is non-negative hermitian. Using the Gram-Schmidt method for calculating eigenvectors, determine the columns of \( U \), the matrix, corresponding to ordered orthonormal eigenvectors of \( R_eR_e^* \).

6. The columns of \( V \) are the orthonormal eigenvectors of \( R_eR_e^* \). \( V \) is a \((p_e+1) \times (p_e+1) \) matrix.

7. \( \Sigma \) is a matrix whose elements are zero except possibly along main diagonal, \( \sigma_{kk} \).

Form \( A(k) = U \Sigma_k V^* \)

where \( \Sigma_k \) is obtained by setting to zero all but its "k" largest singular values.
8. Finally, we define

$$v(k) = \frac{||A(k)||}{||A||}$$

so that as $v(k) \rightarrow 1$ as $k \rightarrow p$, the order of our model.

The above approach requires that singular values be generated for carrying out SVD. Improved accuracy for computing singular values has been developed by Dongarra [8].

The work of S. Y. Kung [9] using state space variables with SVD has recently demonstrated superior spectral estimates than estimates obtained in terms of transfer function parameters. The work in this area should be explored for improved estimates for perturbed covariance data.

5. ADAPTIVE MODELING

The goal of this investigation is to develop refined models which adapt to a changing environment. Naturally, it becomes appropriate to develop adaptive models which continuously update the parameters as new time series observations become available ($x(N+1)$, $x(N+2)$,...). To start this investigation it will be appropriate to develop a class of adaptive autocorrelations estimates and adaptive algorithms which from experience have proved most successful. We will define an adaptive class of autocorrelation estimators as follows.

$$\hat{R}(i,j) = \frac{1}{(N + k_2 - k_1) \sum_{k=k_1}^{n+k_2-1} x(k+1-i)x(k+1-j)}$$

where $N$ is the total set of observations
1 < i < p + 1
1 < j < p + 1
1 < k_1, k_2 < p + 1

with \( k_1 \) and \( k_2 \) selected so that the number of lag products \((N + k_2 - k_1)\) is at least equal to \( p + 1 \).

The data matrix \( \hat{R} \) which uses \( \hat{R}(i, j) \) as its \((i, j)\) elements can be written in the form

\[
\hat{R} = \left( \frac{1}{N + k_2 - k_1} \right) X_N^* X_N
\]

(53)

where \( X_N \) is a \((N + k_2 - k_1) \times (p + 1)\) data matrix such that

\( X_N(i, j) = x(k_1 + i - j) \)

and

\[
1 < i < N + k_2 - k_1
\]
\[
1 < j < p + 1
\]

It should be noted that we set \( x = 0 \) whenever \( k_1 + i - j \) falls outside the observation set \( 1 < n < N \).

We have now established the data matrix to be used in adaptive modeling. Experience has shown that \( \hat{R} \) is unbiased and provides consistent statistics when \( k_1 = p + 1 \) and \( k_2 = 1 \). This method is called the covariance method.

**AR Model - Covariance Method:**

As indicated, with \( k_1 = p + 1 \) and \( k_2 = 1 \),

\[
X_N^* X_N \hat{a}_N = (N + k_2 - k_1) |b_0|^2 e_1
\]

(54)

where \( b_0 \) is selected to normalize 1st component of \( \hat{a}_N \) to 1.

Now let

\[
X_{N+1}^* X_{N+1} = X_N^* X_N + X_{N+1}^* X_{N+1} \quad (N > p + 1)
\]

(55)
where \( x_{N+1} = [x(N+1), x(N), \ldots, x(N+1-p)] \)

For \( k_2 = 1 \) and \( N = p+1 \), and the initial value \( x_N^* x_N^* = x_{p+1}^* x_{p+1} \) so that

\[
x_{p+1}^* x_{p+1} = \sum_{k=k_1}^{p+1} x(k+1-i) x(k+1-j)
\]

(56)

for \( 1 \leq i \leq p+1 \) and \( 1 \leq j \leq p+1 \). As can be seen from the expression for \( a_{N+1} \) we need to compute \( [x_{N+1}^* x_{N+1}]^{-1} \). This is given by

\[
[x_{N+1}^* x_{N+1}]^{-1} = [x_N^* x_N]^{-1} - \frac{(y_{N+1}^* y_{N+1})}{(1 + y_{N+1}^* y_{N+1}^{-1})}
\]

(57)

where

\[
y_{N+1} = x_{N+1}^* x_{N+1}^{-1} \text{ for } N > p+k_1
\]

Accordingly, using Gaussian elimination, we can calculate \( [x_N^* x_N]^{-1} \) for \( N = p+k_1 \) so that for all \( N > p+k_1 \), we can calculate \( [x_{N+1}^* x_{N+1}]^{-1} \) using the above expression. This will be used for updating the parameter \( a_{N+1} \).

AR adaptive algorithms then require the following steps:

Step 1: Input data: \( x(N+1), [x_N^* x_N]^{-1} \)

Step 2: Compute: \( [x_{N+1}^* x_{N+1}]^{-1} \)

Step 3: Let \( c = [x_{N+1}^* x_{N+1}]^{-1} e_1 \)

Step 4: \( a_{N+1} = c(1)^{-1} c \) where \( c(1) \) is the 1st component of \( c \)

The problem of AR order determination still remains as in the non-adaptive case.

The approach recommended in using raw time series data where \( \hat{R}(i,j) = r_X(i-j) \) for \( 1 \leq i, j \leq p+1 \) is to find the order "\( p_1 \)" for which \( R \) has \( (p-p_1) \) of its eigenvalues sufficiently close to zero for all \( p > p_1 \). This could be carried out using SVD.

ARMA Adaptive Modeling:

As in the AR adaptive model we need to define \( \hat{R}_1(i,j) \) which we do by using

\[
\hat{R}_1(i,j) = \frac{1}{(N+k_2-k_1-q-1)} \sum_{k=k_1}^{N-q+2+k_2} x(k+1-i) x(k+q+2-j)
\]

(58)
with \(1 < i < t\) and \(1 < j < p+1\).

\[ \hat{R}_1(i,j) = r_x(q+1+i-j) \]

The number of lag products, \((N+k_2-k_1-q-1)\), is selected such that \((N+k_2-k_1-q-1)>p+1\) and \(1 < k_1 < t\) \(1 < k_2 < p+1\).

As in the AR adaptive model the covariance method had preferred properties, in the ARMA adaptive model the covariance method with \(k_1=t\) and \(k_2=1\) has similarly been demonstrated as being unbiased and statistically consistent.

**ARMA Covariance Method:**

As indicated we constrain the parameters \(k_1, k_2\) as shown: \(k_1 = t, k_2 = 1\), then we write \(\hat{R}_1\) as follows

\[ \hat{R}_1 = \left( \frac{1}{N+k_2-k_1-q-1} \right) \sum_{i=1}^{N+k_2-k_1-q-1} X_N(x) \]

where

\[ X_N(i,j) = x(k_1+i-q-j) \]

for

\[ 1 < i < N+k_2 - k_1 - q - 1 \]

\[ 1 < j < p+1 \]

with

\[ Y_N(i,j) = x(k_1+i-j) \]

and

\[ 1 < i < N+k_2 - k_1 - q - 1 \]

\[ 1 < j < t \]

with

\[ x(n) = 0 \] for \(n>N\) or \(n<1\)

To determine the AR parameters we need to solve

\[ X_N^* Y_N Y_N X_N A_N = \alpha e_1 \]

where \(\alpha\) is selected for the first component of \(A_N\) equal to 1.

When the updated \(A_{N+1}\) parameter is calculated we will require \(Y_{N+1}^* X_{N+1}\).
This can be calculated using

\[ Y_{N+1}^* X_{N+1} = Y_N^* X_N + Y_N^* X_N \quad (N > t) \]

where \( X_N = [x(N+1), x(N), \ldots, x(N+1-p)] \)

\( Y_N = [x(N-q), x(N-q-1), \ldots, x(N+1-q-t)] \)

so that for \( N > t \)

\[ X_{N+1}^* Y_{N+1}^* X_{N+1} = X_N^* Y_N^* X_N + Z_N^* Z_N + Z_N^* Z_N + (Y_N^* Y_N) X_N^* X_N \quad (61) \]

and

\[ Z_N = Y_N^* X_N \]

The best overall ARMA adaptive model performance is the covariance method for \( k_2 = 1 \) and \( k_1 = t \) (\( k_1 \) can range in the interval \( 1 < k_1 < t \)); the ARMA adaptive model algorithm for \( k_2 = 1 \) and \( 1 < k_1 < t \) is given as follows:

**Step 0:** The input to commence the algorithm at \( N = q + p + k_1 + 1 \) is \( Y_N^* X_N \) and \([X_N^* Y_N^* X_N]^{-1}\)

which can be calculated by Gaussian elimination.

**Step 1:** \( N = q + p + k_1 + 1 \)

**Step 2:** Compute \( Y_{N+1}^* X_{N+1} \) from

\[ Y_{N+1}^* X_{N+1} = Y_N^* X_N + Y_N^* X_N \]

\( X_N = [x(N+1), x(N), \ldots, x(N+1-p)] \)

\( Y_N = [x(N-q), x(N-q-1), \ldots, x(N+1-q-t)] \)

**Step 3:** \( Z_N = Y_N^* X_N \)

**Step 4:** \( u_1 = Z_N \)

\[ y_1 = X_N \]

\[ A_1^{-1} = (X_N^* Y_N^* X_N)^{-1} \]

Compute \([A_1^{-1} + u_1^* v_1]^{-1}\) from
\[
[A + u^* v]^{-1} = A^{-1} - \frac{[A^{-1} u^*] [v A^{-1}]}{(1 + v A^{-1} u^*)}
\]

Step 5: \( u_2 = x_{N+1}^* \), \( v_2 = z_N \)
\( A_2^{-1} = [A_1 + u_1 v_1]^{-1} \)
Compute \([A_2 + u_2^* v_2]^{-1}\) from step 4.

Step 6: \( u_3 = (x_N x_N^*) x_N \)
\( v_3 = x_N \)
\( A_3^{-1} = [A_2 + u_2^* v_2]^{-1} \)
Compute \([A_3 + u_3^* v_3]^{-1} = [x_{N+1}^* y_{N+1}^* y_{N+1} x_{N+1}]^{-1}\) from step 4.

Step 7: \( c = [x_{N+1}^* y_{N+1}^* y_{N+1} x_{N+1}]^{-1} e_1 \)
where \( a_{N+1} = [c(1)]^{-1} c \)
\( c(1) \) is the first component of \( c \).

Step 8: Let \( N = N + 1 \), go to step 2.

We have presented a sufficient number of useful methods that should be explored in stochastic modeling for any application you may have in mind. The application of SVD provides an important method for determining order of the model. The recent results by S. Kung using state space and SVD appear to provide higher resolution models and should be evaluated for Army applications.
REFERENCES


## DISTRIBUTION LIST

<table>
<thead>
<tr>
<th>No. of Copies</th>
<th>Organization</th>
<th>No. of Copies</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Commander</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Defense Technical Information Center</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: DTIC-DDAC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cameron Station, Bldg 5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alexandria, VA 22304-6145</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Commander</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>US Army Materiel Command</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: AMXPA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AMCDMA-M</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5001 Eisenhower Avenue</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alexandria, VA 22333-0001</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Director</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Combat Data Information Center</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: AFWAL/FIES/CDIC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wright-Patterson AFB</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OH 45433-5000</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Pentagon Library</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>ATTN: AN-AL-RS (Army Studies)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Pentagon, Room 1A518</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wash, DC 20310</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Commander</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>US Army Concepts Analysis Agency</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: CSCA-MSI-L</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8120 Woodmont Avenue</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bethesda, MD 20814-2797</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>General Dynamics</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Land Systems Division</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATTN: Mr. Paver/Mr. J. O'Rourke</td>
<td></td>
</tr>
<tr>
<td></td>
<td>P.O. Box 1800</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Warren, MI 48090</td>
<td></td>
</tr>
</tbody>
</table>

**Commander**
US Army Development and Employment Agency
ATTN: MODE-TED-SAB (2 cys)
Ft. Lewis, WA 98433-5000

**Director**
US Army TRADOC Systems
Analysis Activity
ATTN: ATOR-TSL
ATOR-T
ATOR-TE
ATOR-TF

White Sands Missile Range
NM 88002-5502

**Commandant**
US Army Infantry School
ATTN: ATSH-CD-CS-OR
Fort Benning, GA 31905-5400

**Commander**
US Army Tank-Automotive Command
ATTN: AMSTA-Z (Tech Lib)
AMSTA-ZE
AMSTA-ZSA
Warren, MI 48090

HQDA (DACS-CV)
Washington, DC 20310

**Project Manager**
MIEI Tank System
ATTN: AMCPM-MIE1
Warren, MI 48090
<table>
<thead>
<tr>
<th>No. of Copies</th>
<th>Organization</th>
<th>No. of Copies</th>
<th>Organization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Dir, BRL</td>
<td>7</td>
<td>Director, AMSAA</td>
</tr>
<tr>
<td></td>
<td>ATTN: AMXBR-OD-ST</td>
<td></td>
<td>ATTN: AMXSY-G</td>
</tr>
<tr>
<td></td>
<td>Bldg 305</td>
<td></td>
<td>AMXSY-GA (W. Brooks/</td>
</tr>
<tr>
<td></td>
<td>Aberdeen Proving Ground, MD</td>
<td></td>
<td>E. Christman/B. Siegel/</td>
</tr>
<tr>
<td></td>
<td>21005-5066</td>
<td></td>
<td>R. Conroy)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AMXSY-GB (B. Blankenbiller)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AMXSY-C (H. Burke)</td>
</tr>
<tr>
<td>2</td>
<td>Dir, BRL, Bldg 328</td>
<td></td>
<td>APG, MD 21005-5071</td>
</tr>
<tr>
<td></td>
<td>ATTN: AMXBR-OD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AMXBR-SECAD (Dr. Johnson/</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Dr. Wolfe)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Aberdeen Proving Ground, MD</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>21005-5066</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
THE PRINCIPAL FINDINGS and recommendations of the work reported herein are as follows:

N/A

THE MAIN ASSUMPTIONS on which the work reported herein rests are as follows:

N/A

THE PRINCIPAL LIMITATIONS of this work which may affect the findings are as follows:

N/A
THE SCOPE OF THE STUDY
Review methodology in stochastic modeling.

THE STUDY OBJECTIVE
Identify current approaches in stochastic modeling including adaptive autoregressive (AR) models.

THE BASIC APPROACH
Review of professional journals in area of signal processing.

THE REASONS FOR PERFORMING THE STUDY
Requirement to develop stochastic mathematical models arises across the whole range of engineering and applied research where observations are made of a physical process, corrupted by noise, and it is desired to determine the underlying nature either in time or frequency or both of the observed phenomenon.

THE STUDY SPONSOR
Ground Warfare Division and Ballistic Research Laboratory in support of gun dynamics efforts.

THE STUDY EFFORT

COMMENTS AND QUESTIONS
THE PRINCIPAL FINDINGS and recommendations of the work reported herein are as follows:

N/A

THE MAIN ASSUMPTIONS on which the work reported herein rests are as follows:

N/A

THE PRINCIPAL LIMITATIONS of this work which may affect the findings are as follows:

N/A
THE SCOPE OF THE STUDY

Review methodology in stochastic modeling.

THE STUDY OBJECTIVE

Identify current approaches in stochastic modeling including adaptive autoregressive (AR) model.

THE BASIC APPROACH

Review of professional journals in area of signal processing.

THE REASONS FOR PERFORMING THE STUDY

Requirement to develop stochastic mathematical models arises across the whole range of engineering and applied research where observations are made of a physical process, corrupted by noise, and it is desired to determine the underlying nature either in time or frequency or both of the observed phenomenon.

THE STUDY SPONSOR

Ground Warfare Division and Ballistic Research Laboratory in support of gun dynamics efforts.

THE STUDY EFFORT

COMMENTS AND QUESTIONS