The Annals of Statistics
1979, Vol. 7, No. 1, 96–107

ON THE RELATION BETWEEN FITTING AUTOREGRESSION AND PERIODOGRAM WITH APPLICATIONS

BY HIDEAKI SAKAI[1] TAKASHI SOEDA AND HIDEKATSU TOKUMARU

University of Tokushima and Kyoto University

The relation between fitting autoregression and periodogram is herein presented. More specifically, the asymptotic error covariance matrices of the estimates of the autoregressive parameters using Yule-Walker equations are expressed in terms of the periodogram. These expressions permit the immediate calculation of the error covariance matrices for various time series problems including the autoregressive spectral estimation and fitting autoregression to the data with randomly missed observations.

1. Introduction. In time series analysis, roughly speaking, mainly two methods have been used. One of them gives the analysis in the frequency domain. It is well known that the quantity called the periodogram plays the fundamental role in that method. (For example, see Jenkins and Watts [5].) The other gives the analysis in the time domain by which we mean that one postulates some parametric model and the data are fitted to this model by estimating the parameters. Among many models, an autoregressive (AR) process model is preferred because of its simplicity in estimating the parameters. When only the estimation of the autoregressive parameters of a mixed autoregressive-moving average (ARMA) process is of interest, similar simplicity occurs. That is, in the above two cases, one only requires the solution of Yule-Walker equations. (See, for example, Box and Jenkins [3].)

The statistical properties of the periodogram and error covariance matrices of the estimates of AR parameters using Yule-Walker equations were discussed and obtained separately in the literature. (As for the former, see [5]. As for the latter, the classical paper is Mann and Wald [7]. See also the paper of Baggeroer [2] for a different view point.) As far as the authors are aware, there are few papers discussing the relation between them.

In this paper, we present the clear-cut formulae connecting them. To show the usefulness of the formulae, we apply them to some time series problems including autoregressive spectral estimation and fitting autoregression to the data with randomly missed observations. To the best of our knowledge, problems of time series with missed observations have been treated mainly in the frequency domain, or in other words, several papers, for example [6], [8] and [9], were concerned with spectral analysis and calculations of asymptotic variances of the proposed spectral estimators. Thus our result is the first which gives the expression for the error covariance matrix of the estimate of AR process parameters. However, its derivation depends heavily on the work of Scheinok [9].

Received June 1977; revised December 1977.

[1]Now with Department of Applied Mathematics and Physics, Kyoto University.

AMS 1970 *subject classifications*. Primary 62M10; Secondary 62J05.

Key words and phrases. Fitting autoregression, periodogram, missing data.

As another application, we consider the problem of estimation of AR process parameters based on the data contaminated by white noise, already discussed in Walker [10] where the explicit expression for the error covariance matrix was derived only in the case of the first-order AR process. With our approach, the general formula can easily be obtained. Lastly, our approach is applied to rederive the error covariance matrix of the estimate of the AR parameters of an ARMA process which was originally obtained by Gersch [4].

2. Fitting autoregressions and periodogram. Let us assume for the moment that the time series $\{x_t\}$ under consideration is a zero-mean Gaussian mth order AR process given by the equation

$$x_t - a_1 x_{t-1} - \cdots - a_m x_{t-m} = u_t \tag{1}$$

where $\{u_t\}$ is a sequence of white noise with

$$E[u_t] = 0, \qquad E[u_t u_s] = \sigma^2 \delta_{t,s}.$$

To assure the stationarity of (1), it is also assumed that the roots of the following equation

$$1 - a_1 z^{-1} - \cdots - a_m z^{-m} = 0 \tag{2}$$

lie within the unit circle.

We denote the autocovariance function of $\{x_t\}$ as $r_k = E[x_t x_{t+k}]$. Define an $m \times m$ matrix $\mathbf{R}$, $m \times 1$ vectors $\mathbf{r}$ and $\mathbf{a}$, respectively as follows:

$$(i, k)\text{th element of } \mathbf{R} = R_{i,k} = r_{i-k}$$

$$\mathbf{r} = (r_1, r_2, \cdots, r_m)^T, \qquad \mathbf{a} = (a_1, a_2, \cdots, a_m)^T$$

where "T" denotes the transpose operation. As is well known, the Yule-Walker equation holds.

$$\mathbf{Ra} = \mathbf{r}. \tag{3}$$

When a set of data $\{x_1, x_2, \cdots, x_N\}$ is available, one of the most popular estimators for r_k is usually taken as

$$\hat{r}_k = \frac{1}{N} \Sigma_{i=1}^{N-|k|} x_i x_{i+|k|} \qquad k = 0, \pm 1, \pm 2, \cdots, \pm (N-1). \tag{4}$$

Upon substituting these $\hat{r}_k$'s into the r_k's in (3), the estimator of $\mathbf{a}$ is given by the solution of

$$\hat{\mathbf{R}}\hat{\mathbf{a}} = \hat{\mathbf{r}} \tag{5}$$

where $\hat{\mathbf{R}}$, $\hat{\mathbf{a}}$ and $\hat{\mathbf{r}}$ are defined as above. It is well known that as $N \to \infty$, the $\hat{r}_k$'s and $\hat{\mathbf{a}}$ are asymptotically consistent estimators of the r_k's and $\mathbf{a}$, respectively. Define the estimation errors as $\Delta\mathbf{a} = \hat{\mathbf{a}} - \mathbf{a}$, $\Delta\mathbf{r} = \hat{\mathbf{r}} - \mathbf{r}$ and $\Delta\mathbf{R} = \hat{\mathbf{R}} - \mathbf{R}$. If N is sufficiently large, these errors can be assumed to be small. By substituting these into (5), neglecting the second order term concerning the errors and noting the relation (3), we have

$$\mathbf{R}\Delta\mathbf{a} \approx \Delta\mathbf{r} - (\Delta\mathbf{R})\mathbf{a} = \hat{\mathbf{r}} - \hat{\mathbf{R}}\mathbf{a}. \tag{6}$$

Let us turn our attention to the periodogram defined by

$$I_N(s) = \frac{1}{2\pi N}\left|\Sigma_{i=1}^{N} x_i \exp(-jis)\right|^2 \qquad |s| \leqslant \pi. \tag{7}$$

Since (7) can be represented in terms of the $\hat{r}_k$'s as

$$I_N(s) = \frac{1}{2\pi}\Sigma_{k=-N+1}^{N-1} \hat{r}_k \exp(-jks), \tag{8}$$

conversely the r_k's are represented in terms of the periodogram as

$$\hat{r}_k = \textstyle\int_{-\pi}^{\pi} I_N(s) \exp(jks)\, ds. \tag{9}$$

Hence, the kth element of (6) can be expressed by using (9) as follows:

$$(\mathbf{R}\Delta\mathbf{a})_k \approx \hat{r}_k - \Sigma_{i=1}^{m} \hat{r}_{k-i} a_i = \textstyle\int_{-\pi}^{\pi} B(s) I_N(s) \exp(jks)\, ds \tag{10}$$

where we put

$$B(s) = \Sigma_{i=0}^{m}(-a_i)\exp(-jis) \tag{11}$$

with $-a_0 = 1$. On the other hand, it easily follows from (8) that for $k > 0$,

$$\textstyle\int_{-\pi}^{\pi} B(s) E[I_N(s)] \exp(jks)\, ds = \Sigma_{i=0}^{m}(1 - N^{-1}|k - i|)(-a_i) r_{k-i} = \eta_k N^{-1}$$

where we utilize the relation (3) and define $\eta_k = \Sigma_{i=0}^{m}|k - i| a_i r_{k-i}$ for $k \geqslant 0$. Therefore, by multiplying the kth and ith elements of (10) and taking the expectation, the expression of the error covariance matrix is given by

$$\left(\mathbf{R}E[\Delta\mathbf{a}\Delta\mathbf{a}^T]\mathbf{R}^T\right)_{k,i} = \textstyle\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)\,\mathrm{Cov}[I_N(s), I_N(t)] \times \exp[j(ks + it)]\, ds\, dt + \eta_k \eta_i N^{-2}. \tag{12}$$

As will be seen later, the first term in the right hand side of (12), which is denoted by $A_{k,i}$ henceforth, is of order N^{-1} so that the second term is of no importance.

Next we examine the estimation error of the residual σ^2 which satisfies the well-known identity

$$\sigma^2 = r_0 - \Sigma_{i=1}^{m} r_i a_i. \tag{13}$$

Hence, the estimator for σ^2 usually takes the form of

$$\hat{\sigma}^2 = \hat{r}_0 - \Sigma_{i=1}^{m} \hat{r}_i \hat{a}_i \tag{14}$$

where the a_i's are given by (5). By using the same technique to derive (10), the estimation error can be expressed as

$$\Delta\sigma^2 \approx \Delta r_0 - \Sigma_{i=1}^{m}(a_i \Delta r_i + r_i \Delta a_i) = -\Delta\mathbf{a}^T \cdot \mathbf{r} + \textstyle\int_{-\pi}^{\pi} B(s) I_N(s)\, ds - \sigma^2. \tag{15}$$

In deriving (15), we utilize the relation (13) and the fact that $I_N(s) = I_N(-s)$.

By virtue of multiplying (10) with (15) and taking the expectation, we have

(16) $$\mathbf{R}E[\Delta\mathbf{a}\Delta\sigma^2] \approx -\mathbf{R}E[\Delta\mathbf{a}\Delta\mathbf{a}^T]\mathbf{r} + \int_{-\pi}^{\pi} B(s)E[\mathbf{R}\Delta\mathbf{a}\cdot I_N(s)]\, ds - \sigma^2\mathbf{R}E[\Delta\mathbf{a}].$$

Upon noting from (12) that $\mathbf{R}E[\Delta\mathbf{a}\Delta\mathbf{a}^T] = \mathbf{A}\mathbf{R}^{T-1} = \mathbf{A}\mathbf{R}^{-1}$ with $(\mathbf{A})_{k,i} = A_{k,i}$ and $\mathbf{R}^{-1}\mathbf{r} = \mathbf{a}$, the kth element of the first term of (16) becomes $-\sum_{i=1}^m A_{k,i}a_i$. The kth element of the second term is equal to

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)\,\mathrm{Cov}[I_N(s), I_N(t)]\exp(jkt)\, ds\, dt + (\sigma^2 + \eta_0 N^{-1})\eta_k N^{-1}$$

and the kth element of the third one is $-\sigma^2\eta_k N^{-1}$. Hence we obtain

(17) $$(\mathbf{R}E[\Delta\mathbf{a}\Delta\sigma^2])_k = b_k + \eta_0\eta_k N^{-2}$$

with

(18) $$b_k = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(-s)B(t)\,\mathrm{Cov}[I_N(s), I_N(t)]\exp(jkt)\, ds\, dt.$$

Like (12), the second term of (17) is of no concern.

Denote $\mathbf{R}^{-1} = (q_{k,i})$. Then,

(19) $$\Delta a_k \approx \sum_{i=1}^m q_{k,i}\int_{-\pi}^{\pi} B(t)I_N(t)\exp(jit)\, dt.$$

On the other hand, we find that

(20) $$E[(\Delta\sigma^2)^2] \approx -\sum_{i=1}^m r_i E[\Delta\sigma^2\Delta a_i] + \int_{-\pi}^{\pi} B(s)E[\Delta\sigma^2 I_N(s)]\, ds - E[\sigma^2\Delta\sigma^2].$$

Using (17) and the identity

(21) $$\sum_{i=1}^m r_i q_{k,i} = a_k,$$

the first term of (20) is equal to $-\sum_{i=1}^m b_i a_i$. The second term can be rewritten by substituting (15) into $\Delta\sigma^2$ as

(22) $$-\sum_{k=1}^m r_k\int_{-\pi}^{\pi} B(s)E[\Delta a_k\cdot I_N(s)]\, ds + \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)E[I_N(s)I_N(t)]\, ds\, dt - \sigma^2\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)E[I_N(s)]\, ds.$$

Upon substituting (19) into Δa_k in the first term of (22) and using (21), this first term is given by

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)\,\mathrm{Cov}[I_N(s), I_N(t)]\sum_{i=1}^m(-a_i)\exp(jit)\, ds\, dt - \sum_{i=1}^m a_i(\sigma^2 + \eta_0 N^{-1})\eta_i N^{-1}.$$

Hence, (22) finally becomes

$$\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)B(-t)\,\mathrm{Cov}[I_N(s), I_N(t)]\, ds\, dt - \sum_{i=1}^m a_i(\sigma^2 + \eta_0 N^{-1})\eta_i N^{-1} + (\sigma^2 + \eta_0 N^{-1})^2 - \sigma^2(\sigma^2 + \eta_0 N^{-1}).$$

Similarly, the third term of (20) becomes $\sigma^2\sum_{i=1}^{m} a_i\eta_i N^{-1} - \sigma^2(\sigma^2 + \eta_0 N^{-1}) + \sigma^4$. Thereby we finally obtain

$$(23) \quad E\left[(\Delta\sigma^2)^2\right] \approx \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(-s)B(t)B(-t)\,\mathrm{Cov}\left[I_N(s), I_N(t)\right] ds\,dt.$$

The formulae (12), (17) and (23) give the clear-cut relations between the error covariances in fitting autoregressions and the covariance of the periodogram.

On the other hand, if in general $\{x_t\}$ is a stationary Gaussian time series with spectral density $f(\omega)$, the covariance of the periodogram is asymptotically expressed as

$$(24) \qquad \mathrm{Cov}\left[I_N(s), I_N(t)\right] \approx f(s)f(t)N^{-2}\{F_N(s+t) + F_N(s-t)\}$$

where $F_N(\cdot)$ is the Fejér kernel and is defined by

$$F_N(x) = \left(\frac{\sin Nx/2}{\sin x/2}\right)^2.$$

(See, for example, [5, page 250].) Also, from the theory of Fejér kernel, it is well known that

$$\int_{-\pi}^{\pi} g(y)F_N(x-y)\,dy \approx 2\pi N g(x).$$

Hence, $F_N(\cdot)$ can be approximated by Dirac's delta function as $2\pi N\,\delta(\cdot)$. Thus (24) becomes $\mathrm{Cov}[I_N(s), I_N(t)] \approx 2\pi N^{-1} f(s)f(t)\{\delta(s+t) + \delta(s-t)\}$. Substitution of this into (12) gives

$$(25) \qquad A_{k,i} = 2\pi N^{-1}\int_{-\pi}^{\pi}\left\{B(s)B(-s)f(s)^2 \exp\left[j(k-i)s\right] + B(s)^2 f(s)^2 \exp\left[j(k+i)s\right]\right\} ds$$

The expression (12) together with (25) can also be viewed as the error covariance matrix of the estimate of the optimal linear one-step prediction coefficients of the tapped-delay-line of length m for the general time series. In particular, if $\{x_t\}$ is an mth order AR process, the power spectra is given by

$$(26) \qquad f(s) = \frac{\sigma^2}{2\pi B(s)B(-s)}.$$

Thus the first term of (25) is

$$\sigma^2 N^{-1}\int_{-\pi}^{\pi}\frac{\sigma^2}{2\pi B(s)B(-s)}\exp\left[j(k-i)s\right] ds = \sigma^2 N^{-1} r_{k-i}$$

and the second term is

$$(2\pi N)^{-1}\sigma^4\int_{-\pi}^{\pi}\frac{1}{B(-s)^2}\exp\left[j(k+i)s\right] ds = \frac{\sigma^2 N^{-1}}{2\pi j}\oint_{|z|=1}\frac{z^{k+i-1}\,dz}{B_0(z)^2}$$

where $B_0(z) = 1 - a_1 z - \cdots - a_m z^m$. Since by the assumption concerning (2), roots of $B_0(z) = 0$ lie outside the unit circle and $k + i - 1 > 0$, the integrand of the above complex integral is regular within the unit circle; therefore the integral is zero from Cauchy's theorem. Thus, $\mathbf{A} = \sigma^2 N^{-1}(r_{k-i}) = \sigma^2 N^{-1}\mathbf{R}$. Therefore, the well known result [3, page 281] [7]

(27) $$NE[\Delta\mathbf{a}\Delta\mathbf{a}^T] \approx \sigma^2\mathbf{R}^{-1}$$

follows. In a similar way, one can show that

$$b_k = \frac{\sigma^4 N^{-1}}{\pi j}\oint_{|z|=1}\frac{z^{k-1}}{B_0(z)}\,dz = 0.$$

Hence,

(28) $$NE[\Delta\mathbf{a}\Delta\sigma^2] \approx \mathbf{0}.$$

Also from (23), we have

(29) $$NE[(\Delta\sigma^2)^2] \approx 2\sigma^4.$$

These results are consistent with those obtained in [3, pages 280–281] by evaluating the Fisher information matrix for the maximum likelihood estimate, if one notes that $\Delta\sigma^2 \approx 2\sigma\Delta\sigma$.

3. Autoregressive spectral estimation. In this section, we derive the expression for the covariance of the autoregressive power spectrum estimator originally considered by Akaike [1]. Usually, the estimator takes the form of

(30) $$\hat{f}(s) = \frac{\hat{\sigma}^2}{2\pi\hat{B}(s)\hat{B}(-s)},$$

with

$$\hat{B}(s) = 1 - \Sigma_{i=1}^m \hat{a}_i \exp(-jis).$$

Hence, the estimation error can be approximately represented as

(31) $$\Delta f(s) \approx f(s)\left[\frac{\Delta\sigma^2}{\sigma^2} + \mathbf{H}(s)^T\Delta\mathbf{a}\right]$$

where we define

(32) $$\mathbf{H}(s) = \frac{\mathbf{E}(s)}{B(s)} + \frac{\mathbf{E}(-s)}{B(-s)}$$

with

$$\mathbf{E}(s) = [\exp(-js), \exp(-2js), \cdots, \exp(-jms)]^T.$$

Define the new vector $\mathbf{M}(s)$ by

(33) $$\mathbf{R}^{-1}\mathbf{H}(s) = \mathbf{M}(s) = [M_1(s), M_2(s), \cdots, M_m(s)]^T.$$

Then the covariance between the errors at angular frequencies s and t is given by

(34) $$E[\Delta f(s)\Delta f(t)] \approx f(s)f(t)\{\sigma^{-4}E[(\Delta\sigma^2)^2] + \sigma^{-2}[M(s)^T + M(t)^T] \times \mathbf{R}\,E[\Delta\mathbf{a}\Delta\sigma^2] + \mathbf{M}(s)^T\mathbf{R}\,E[\Delta\mathbf{a}\Delta\mathbf{a}^T]\mathbf{R}^T\mathbf{M}(t)\}.$$

By substituting (12), (17) and (23) into (34) and defining

(35) $$K(s, \mu) = \Sigma_{k=1}^m M_k(s)\exp(jk\mu),$$

we obtain the desired formula

(36) $$E[\Delta f(s)\Delta f(t)] \approx f(s)f(t)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(\mu)B(\nu)\,\mathrm{Cov}[I_N(\mu), I_N(\nu)] \times [K(s,\mu) + \sigma^{-2}B(-\mu)][K(t,\nu) + \sigma^{-2}B(-\nu)]\,d\mu\,d\nu.$$

When $\{x_t\}$ is Gaussian, by the same calculations leading to the results (27), (28) and (29) one can show that

(37) $$NE\left[\frac{\Delta f(s)}{f(s)}\cdot\frac{\Delta f(t)}{f(t)}\right] \approx 2 + \sigma^2\Sigma_{k=1}^{m}\Sigma_{i=1}^{m} M_k(s)M_i(t)r_{k-i}.$$

The second term of (37) is further simplified by using (33) to $\sigma^2\mathbf{H}(s)^T\mathbf{R}^{-1}\mathbf{H}(t)$.

Using the matrix factorization for $\mathbf{R}^{-1}$ described in [1, (3.7)], we obtain the same formula originally given by Akaike [1, (4.12)]. At this point, Akaike's result is more general than ours since the former is free from the Gaussian assumption. However, this drawback will be remedied in the last section.

4. Application to some missing data problems. The analysis of time series with missed observations was first treated by Jones [6] where missing instants are assumed to be periodic and the variance of some spectral estimator was obtained. The result of [6] was generalized in Parzen [8]. Also, Scheinok [9] considered the case where missing instants are stochastic and form a Bernoulli sequence.

In this section, we set the same situation as in [9], namely,

$$d_i = 1 \qquad \text{if } x_i \text{ is read}$$
$$d_i = 0 \qquad \text{if } x_i \text{ is not read}$$

where $p = \Pr(d_i = 1)$ is independent of i and known a priori. We also assume that the d_t's are independent not only with each other but also with the time series $\{x_t\}$. With the above assumptions, it is obvious that the estimators for the r_k's

(38) $$\hat{r}_0' = \frac{1}{Np}\Sigma_{i=1}^{N}(d_i x_i)^2$$

$$\hat{r}_k' = \frac{1}{Np^2}\Sigma_{i=1}^{N-|k|} d_i x_i d_{i+|k|} x_{i+|k|} \qquad k \neq 0$$

are consistent as $N \to \infty$. Hence, substitution of (38) into (5) and (14) gives the consistent estimators $\hat{\mathbf{a}}'$, $\hat{\sigma}^{2\prime}$ for $\mathbf{a}$ and σ^2, respectively.

Corresponding to (7), the modified periodogram is now defined as in [9] by

(39)
$$I_N'(s) = (2\pi N)^{-1}\left\{\Sigma_{i=1}^{N}\frac{d_i^2 x_i^2}{p} + \Sigma_{k=1}^{N}\Sigma_{i=1;\, i\neq k}^{N}\frac{d_k d_i}{p^2}x_k x_i \exp[-j(i-k)s]\right\}.$$

Then, corresponding to (9), the following relation between $\hat{r}_k'$ and $I_N'(s)$ holds:

(40) $$\hat{r}_k' = \int_{-\pi}^{\pi} I_N'(s)\exp(jks)\,ds.$$

Also, by the assumption concerning $\{d_t\}$ and (39), it is obvious that

$$E[I_N'(s)] = E[I_N(s)].$$

Hence, all the arguments in the preceding sections are entirely valid by interchanging $I_N(s)$ with $I'_N(s)$. Thus it suffices to know $\mathrm{Cov}[I'_N(s), I'_N(t)]$ but its derivation is just the central theme of [9]. The result is

$$
\begin{aligned}
\mathrm{Cov}[I'_N(s), I'_N(t)] \approx\ & 3(4\pi^2 N)^{-1}\underline{(p^{-1}-1)\alpha} + (\pi N)^{-1}\beta \\
& + (\pi N)^{-1}(3p^{-1}-1)[f(t)+f(s)-\pi^{-1}\alpha] + 2N^{-1}\{f(s)^2+f(t)^2 \\
& -\pi^{-1}\textstyle\int_{-\pi}^{\pi} f(x)[f(x)+f(s)+f(t)]\,dx + \pi^{-2}\alpha^2\} \\
& + [f(s)f(t)]N^{-2}[F_N(s+t)+F_N(s-t)] - 2N^{-1}[f(s)+f(t)]^2 \\
& - (2\pi N^2)^{-1}[F_N(s+t)+F_N(s-t)][f(s)+f(t)]\alpha + 4(\pi N)^{-1} \\
& \times[f(s)+f(t)]\alpha + (\pi N)^{-1}\beta + (4\pi^2N^2)^{-1}[F_N(s+t)+F_N(s-t)-12N]\alpha^2 \\
& + (4\pi^2N^2p^2)^{-1}[F_N(s+t)+F_N(s-t)-2N(3-p^2)]\alpha^2 + (\pi N p^2)^{-1} \\
& \times \textstyle\int_{-\pi}^{\pi} f(x)[f(x+s+t)+f(x+s-t)]\,dx \\
& + (p\pi N)^{-1}\Big\{[f(s)N^{-1}(F_N(s+t)+F_N(s-t)) - 2(3-p)(f(s) \\
& + f(t))]\alpha - (2\pi N)^{-1}\alpha^2[F_N(s+t)+F_N(s-t)-\underline{4N}(3-p)] - \frac{2-p}{2} \\
& \times \textstyle\int_{-\pi}^{\pi} f(x)[f(x+s+t)+f(x-s+t)+f(x+s-t)+f(x-s-t)]\,dx \\
& + 4\pi(2-p)f(s)f(t)\Big\}
\end{aligned}
\tag{41}
$$

where for abbreviation we put

$$\alpha = \textstyle\int_{-\pi}^{\pi} f(x)\,dx, \qquad \beta = \int_{-\pi}^{\pi} f(x)^2\,dx.$$

There are two incorrect terms in (4.3) and (4.10) of the original paper [9] to which the underlined parts of (41) correspond. These errors are corrected in (41).

By substituting (41) into (12), (17) and (23) and using delta function approximation to Fejér kernel, one can obtain the asymptotic error covariance of $\hat{\mathbf{a}}'$ and $\hat{\sigma}^{2\prime}$. But the resulting formulae may be rather lengthy and complicated, so we only derive the expression for (12). After some simple calculations, we get

$$
\begin{aligned}
A_{k,i} =\ & N^{-1}\alpha^2 a_k a_i(-9+15p^{-1}-6p^{-2}) \\
& + N^{-1}\sigma^2 r_{k-i} + 2N^{-1}\sigma^2\delta_{k,i}\alpha(p^{-1}-1) \\
& + N^{-1}\alpha^2 e_{k,i}(1+p^{-2}-2p^{-1}) \\
& + N^{-1}\alpha^2 f_{k,i}(1+p^{-2}-2p^{-1}) + 2N^{-1}\alpha g_{k,i} \\
& \times(p^{-1}-1) + (\pi N)^{-1}h_{k,i}(p^{-2}-p^{-1}+0.5) \\
& - (\pi N)^{-1}w_{k,i}(p^{-1}-0.5)
\end{aligned}
\tag{42}
$$

with

$$(43)\qquad e_{k,i} \equiv (2\pi)^{-1}\textstyle\int_{-\pi}^{\pi} B(s)B(-s)\exp[j(k-i)s]\,ds = \sum_{n-n'=k-i}(-a_n)(-a_{n'}) \qquad (n, n' = 0, 1, \cdots, m),$$

$$(44)\qquad f_{k,i} \equiv (2\pi)^{-1}\textstyle\int_{-\pi}^{\pi} B(s)^2 \exp[j(k+i)s]\,ds = \sum_{n+n'=k+i}(-a_n)(-a_{n'}),$$

$$(45)\qquad g_{k,i} \equiv (2\pi)^{-1}\textstyle\int_{-\pi}^{\pi} B(s)^2 f(s)\exp[j(k+i)s]\,ds = \sum_{n=0}^{m}\sum_{n'=0}^{m}(-a_n)(-a_{n'})r_{n+n'-k-i},$$

$$(46)\qquad h_{k,i} \equiv \textstyle\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)\int_{-\pi}^{\pi} f(x)[f(x+s+t)+f(x+s-t)]\,dx \times \exp[j(ks+it)]\,ds\,dt$$

$$(47)\qquad w_{k,i} \equiv \textstyle\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t)\int_{-\pi}^{\pi} f(x)[f(x-s+t)+f(x-s-t)]\,dx \times \exp[j(ks+it)]\,ds\,dt.$$

In (46), by changing x to $-x$ and using $f(x) = f(-x)$, we have $h_{k,i} = w_{k,i}$. Substitution of $f(x+s\pm t) = (2\pi)^{-1}\sum_{i=-\infty}^{\infty} r_i \exp[-ji(x+s\pm t)]$ into (46) yields

$$(48)\qquad h_{k,i} = w_{k,i} = 2\pi\textstyle\sum_{n=-\infty}^{\infty} r_n^2[(-a_{k-n})(-a_{i-n}) + (-a_{k-n})(-a_{i+n})]$$

where we define $-a_n = 0$ if $n > m$ or $n < 0$.

To check the validity of (42), we put $p = 1$, yielding the same result (27) in the previous section. Also to know the explicit value of (42), for example, let $\{x_t\}$ be a first-order AR process with $r_k = a^{|k|}$. In this case, it easily follows that $\alpha = r_0 = 1$, $\sigma^2 = 1 - a^2$, $e_{1,1} = 1 + a^2$, $f_{1,1} = a^2$, $g_{1,1} = 0$ and $h_{1,1} = w_{1,1} = 2\pi \cdot 3a^2$. Therefore, we get

$$(49)\qquad NE[(\Delta a)^2]_{\text{miss}} \approx p^{-2} + a^2(2p^{-2} - 3p^{-1}).$$

Since the missing rate is $1 - p$, the number of net observations can be assumed to be Np with probability one as N tends to infinity. Thus it is reasonable to compare (49) with the error variance from the data of length Np without missed observations. The latter is

$$(50)\qquad NE[(\Delta a)^2]_{\text{cont}} \approx p^{-1} - a^2 p^{-1}.$$

Hence, as long as $p < 1$ holds,

$$(51)\qquad E[(\Delta a)^2]_{\text{miss}} > E[(\Delta a)^2]_{\text{cont}}.$$

This inequality shows the serious effect of missed observations on estimation of parameters. This seriousness is increased as p tends to zero, since $E[(\Delta a)^2]_{\text{miss}}/E[(\Delta a)^2]_{\text{cont}} \approx (1 + 2a^2)p^{-1}/(1 - a^2)$ for small p. Also, it is interesting to note that at $p = \frac{2}{3}$, $NE[(\Delta a)^2]_{\text{miss}} \approx p^{-2}$, independent of the system parameter a.

To conclude this section, the simulation result and the theoretical value (49) are compared. For $a = 0.5$, $N = 500$ and $p = 0.5$, the former, calculated by averaging the squares of the estimation errors over 100 different data sets, is 4.478 while the latter is 4.5. The agreement is fairly good.

The effect of missed observations on the spectral estimator (30) can be calculated by substituting (41) into (36), but the result is rather lengthy and we omit it here.

5. Other applications. Let us now consider the situation where $\{x_t\}$ is not directly observed but the noisy version of it is available. Denote the noisy observation sequence by $\{y_t\}$ with

(52) $$y_t = x_t + v_t$$

where $\{v_t\}$ is a Gaussian white noise sequence uncorrelated with $\{u_t\}$, i.e.,

(53) $$E[v_t] = 0, \qquad E[v_t v_s] = \sigma_v^2 \delta_{t,s}, \qquad E[v_t u_s] = 0.$$

This problem was considered by Walker [10] in which the variances of two different estimators (Method A and Method B) were calculated. In this section, our method is used to derive the error covariance matrix of the Yule-Walker estimate (Method B).

From (1), (52) and (53), it follows that

$$E[y_{t-m-i}(y_t - a_1 y_{t-1} - \cdots - a_m y_{t-m})] = 0 \qquad \text{for } i = 1, 2, \cdots, m$$

or more compactly, the modified Yule-Walker equation $\tilde{\mathbf{R}}\mathbf{a} = \tilde{\mathbf{r}}$ where $(\tilde{\mathbf{R}})_{k,i} = r'_{m+k-i}$, $(\tilde{\mathbf{r}})_i = r'_{m+i}$ and $r'_k = E[y_t y_{t+k}]$. By the same argument developed in Section 2, it follows that the estimation error $\Delta\tilde{\mathbf{a}}$ satisfies

$$(\tilde{\mathbf{R}}\Delta\tilde{\mathbf{a}})_k \approx \int_{-\pi}^{\pi} B(s)\tilde{I}_N(s) \exp[j(k+m)s]\, ds$$

where $\tilde{I}_N(s)$ is the periodogram for the process $\{y_t\}$. Hence, corresponding to (12) we get

(54) $$\left(\tilde{\mathbf{R}}E[\Delta\tilde{\mathbf{a}}\Delta\tilde{\mathbf{a}}^T]\tilde{\mathbf{R}}^T\right)_{k,i} \approx \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} B(s)B(t) \operatorname{Cov}[\tilde{I}_N(s), \tilde{I}_N(t)] \times \exp\{j[(k+m)s + (i+m)t]\}\, ds\, dt = \tilde{A}_{k,i}.$$

By the same approximate calculations as in Section 2 and noting that the spectra of $\{y_t\}$ is given by

$$f_y(s) = (2\pi)^{-1}\left[\sigma_v^2 + \frac{\sigma^2}{B(s)B(-s)}\right],$$

(54) becomes

(55) $$(2\pi N)^{-1}\int_{-\pi}^{\pi} B(s)^2\left[\sigma_v^2 + \frac{\sigma^2}{B(s)B(-s)}\right]^2 \exp[j(k+i+2m)s]\, ds + (2\pi N)^{-1}\int_{-\pi}^{\pi} B(s)B(-s)\left[\sigma_v^2 + \frac{\sigma^2}{B(s)B(-s)}\right]^2 \exp[j(k-i)s]\, ds.$$

If we rewrite the first term of (55) as a complex integral, it can easily be seen that this integral is equal to zero. Hence, using (43), we finally obtain

(56) $$A_{k,i} = N^{-1}\left(e_{k,i}\sigma_v^4 + 2\sigma_v^2\sigma^2\delta_{k,i} + \sigma^2 r_{k-i}\right).$$

As an example, let $\{x_t\}$ be a first-order AR process with $r_k = \sigma^2 a^{|k|}/(1 - a^2)$. Then, (54) together with (56) reduces to

$$NE[(\Delta\tilde{\mathbf{a}})^2] \approx \frac{1 - a^2}{a^2}[1 + 2\lambda(1 - a^2) + \lambda^2(1 - a^4)] \tag{57}$$

with $\lambda = \sigma_v^2/\sigma^2$. This result, of course, agrees with the one obtained by Walker [10]. In a similar way, in principle, one can obtain the error variances and covariances of $\hat{\sigma}^2$ and $\hat{\sigma}_v^2$. However, their derivations will require a large amount of manipulations and are thus omitted here.

As a last application, we rederive the error covariance matrix for the Yule-Walker estimate of AR parameters of a Gaussian ARMA process generated by

$$x_t - a_1 x_{t-1} - \cdots - a_m x_{t-m} = u_t - c_1 u_{t-1} - \cdots - c_n u_{t-n} \tag{58}$$

with $E[u_t] = 0$, $E[u_t u_s] = \sigma^2 \delta_{t,s}$. It is also assumed that the roots of $B_0(z) = 0$ lie outside the unit circle. Obviously, the Yule-Walker equation $\overline{\mathbf{R}}_n \mathbf{a} = \bar{\mathbf{r}}$ holds where $(\overline{\mathbf{R}}_n)_{k,i} = r_{n+k-i}$, $(\bar{r})_i = r_{n+i}$. Hence, as above the covariance matrix is given by

$$\left(\overline{\mathbf{R}}_n E[\Delta\bar{\mathbf{a}}\Delta\bar{\mathbf{a}}^T]\overline{\mathbf{R}}_n^T\right)_{k,i} \approx 2\pi N^{-1} \int_{-\pi}^{\pi} B(s)^2 f(s)^2 \exp[j(k + i + 2n)s]\, ds \tag{59}$$

$$+ 2\pi N^{-1} \int_{-\pi}^{\pi} B(s)B(-s)f(s)^2 \exp[j(k - i)s]\, ds$$

where $f(s)$ is the spectra of (58) and is given by

$$f(s) = \frac{C(s)C(-s)}{2\pi B(s)B(-s)}\sigma^2$$

with $C(s) \triangleq \sum_{i=0}^{n}(-c_i)\exp(-jis)$ $(-c_0 = 1)$. By rewriting the first term of (59) as a complex integral, one can easily show from the assumption concerning the a_i's that this integral is equal to zero. The second term is

$$\sigma^2 N^{-1} \sum_{q=0}^{n} \sum_{q'=0}^{n} (-c_q)(-c_{q'}) r_{k-i-q+q'}.$$

By defining

$$\gamma_k = \sum_{i=0}^{n-k}(-c_i)(-c_{i+k}), \qquad k = 0, 1, \cdots, n - 1,$$

the right hand side of (59) is written in a matrix form as

$$\overline{\mathbf{A}} = \sigma^2 N^{-1}\left[\gamma_0 \overline{\mathbf{R}}_0 + \sum_{i=1}^{n} \gamma_i (\overline{\mathbf{R}}_i + \overline{\mathbf{R}}_i^T)\right] \tag{60}$$

which coincides with the result of Gersch [4]. In the same way as in Section 4, one can straightforwardly calculate the effect of randomly missed observations on the above Yule-Walker estimate, so that we need not reproduce it here.

6. Discussion and conclusion. The key idea of this paper lies in the derivation of the simple formula (10) which connects the error vector of the Yule-Walker estimate and the periodogram. Thus the examinations of various statistical problems concerning the Yule-Walker estimates are converted into those of the periodogram which have been investigated since the beginning of time series analysis. This relation is particularly powerful when applied to the problem with randomly missed observations for which usual methods such as the maximum likelihood method with the evaluation of Fisher information matrix, may not be applicable.

It should also be noted that the formulae (12), (17), (23), (36) and (54) are all valid without *the Gaussian assumption*. But if one drops this Gaussian assumption, one must add a term of N^{-1} order to the right hand side of (24). This term can be obtained by the following rather rough argument. Let the time series $\{x_t\}$ be generated by passing the innovation sequence $\{u_t\}$ whose variance and fourth cumulant are σ^2 and κ_4, respectively, through the linear filter with the transfer function $G(s)$. Then, by the formula (6.3.15) in [5, page 238] and the argument developed in [5, page 250], the extra term can be expressed as $N^{-1}\kappa_4 |G(s)|^2\sigma^2|G(t)|^2\sigma^2/\sigma^4 = N^{-1}\kappa_4 f(s)f(t)/\sigma^4$ where $f(s)$ is the spectra of $\{x_t\}$.

It can be easily shown that in the presence of this extra term the results (27), (28) and (59) are unchanged while in the right hand sides of (29) and (37) one must add κ_4 and κ_4/σ^4, respectively. Thus we can obtain the results of [7], [4] and [1] *without the Gaussian assumption*.

REFERENCES

[1] AKAIKE, H. (1969). Power spectrum estimation through autoregressive model fitting. *Ann. Inst. Statist. Math.* **21** 407–419.

[2] BAGGEROER, A. B. (1976). Confidence intervals for regression (MEM) spectral estimates. *IEEE Trans. Information Theory* **22** 534–545.

[3] BOX, G. E. P. and JENKINS, G. M. (1970). *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.

[4] GERSCH, W. (1970). Estimation of the autoregressive parameters of a mixed-autoregressive-moving average time series. *IEEE Trans. Automatic Control* **15** 583–588.

[5] JENKINS, G. M. and WATTS, D. G. (1968). *Spectral Analysis and Its Applications.* Holden-Day, San Francisco.

[6] JONES, R. H. (1962). Spectral analysis with regularly missed observations. *Ann. Math. Statist.* **33** 455–461.

[7] MANN, H. B. and WALD, A. (1943). On the statistical treatment of linear stochastic difference equations. *Econometrica* **11** 173–220.

[8] PARZEN, E. (1963). On spectral analysis with missing observations and amplitude modulation. *Sankhya, Ser. A* **25** 383–392.

[9] SCHEINOK, P. A. (1965). Spectral analysis with randomly missed observations: the binomial case. *Ann. Math. Statist.* **36** 971–977.

[10] WALKER, A. M. (1960). Some consequences of superimposed error in time series analysis. *Biometrika* **47** 33–43.

DEPARTMENT OF INFORMATION SCIENCE AND SYSTEMS ENGINEERING
FACULTY OF ENGINEERING
UNIVERSITY OF TOKUSHIMA
TOKUSHIMA, 770, JAPAN

DEPARTMENT OF APPLIED MATHEMATICS AND PHYSICS
FACULTY OF ENGINEERING
KYOTO UNIVERSITY
KYOTO, 606, JAPAN