Presented Paper to 16th European Meeting of Statisticians, Marburg, 1984

DETERMINATION OF THE ADDITIVE NOISE VARIANCE IN OBSERVED AUTOREGRESSIVE PROCESSES USING VARIANCE COMPONENT ESTIMATION TECHNIQUE

Wolfgang Förstner

Abstract: The paper discusses the determination of the variances σ_e^2 and σ_n^2 in an observed autoregressive process $y_i = x_i + n_i$, with $x_i = \Sigma a_k x_{i-k} + e_i$. It is shown that approximating the estimated Fourier power spectrum $P_y(u) = |H(u)|^2 \sigma_e^2 + \sigma_n^2$ by a weighted least squares fit is identical with the variance component estimation solution in the spatial domain. The statistical and numerical properties of the procedure are analysed showing the versatility of the approach.

1. Introduction

1.1 Autoregressive (AR) models are widely used for describing the statistical behaviour of one- and two-dimensional randomly varying discrete functions. Examples are time series and digital images. This study was motivated by the application of AR-models in the analysis and prediction of heights of topographic surfaces. Preliminary investigations (Jäckle, 1984) suggest that AR-models are suited to describe the behaviour of the slopes or the curvature in terrain profiles. The results of this paper may be used to derive an optimal, possibly adaptive, filter in the presence of observational errors.

1.2 The model for an AR(p)-process (stochastical variables are underscored)

$$\underline{x}_i = \sum_{k=1}^{p} a_k \underline{x}_{i-k} + \underline{e}_i \tag{1}$$

of order p is fully descibed by the p coefficients a_k and the variance σ_e^2 of the driving process $(\underline{e}_i)$ which in most cases is assumed to be white Gaussian noise. Usually only the observed process $(\underline{y}_i)$

$$\underline{y}_i = \underline{x}_i + \underline{n}_i \tag{2}$$

is available, where the $\underline{n}_i$ are observational errors. Then in addition to a_k and σ_e^2 the additive noise variance σ_n^2 is unknown. For a reconstruction of the sequence (x_i) from the observed sequence $(\underline{y}_i)$, e. g. by using a Wiener Filter or equivalenty by least squares techniques, the power spectra of $\underline{x}_i$ and $\underline{n}_i$, i. e. the parameters a_k and both variances have to be known. The standard techniques for estimating the parameters a_k (cf. e.g. Box/Jenkins, 1976) however neglect the effect of the observational noise $\underline{n}_i$ or assume both variances or at least σ_n^2 to be known (cf. e.g. Yum and Park 1983). A joint estimation of all parameters is desirable.

According to an idea of R. L. Kashyap, which became known to the author after finishing the manuskript, the variances can be derived from a nonlinear equation system (Kashyap and Rao 1976, ch.2h) which is based on the representation of an observed AR-process by an autoregressive moving average (ARMA) process. It will be of great interest to compare this approach with the following one which in addition to the estimates of the variances also offers criteria for their evaluation. The estimation, however, needs not necessarily be accomplished in one step but may be achieved in an iterative manner, by alternatively estimating the process coefficients a_k and the variances σ_e^2 and σ_n^2.

1.3 This paper discusses the determination of the variances of the driving and the observation process assuming the process coefficients to be known. The procedure could be part of an iterative algorithm for the joint estimation of all unknowns or used in cases where the process coefficients are known from experience. As the transfer function $H(u)$ of the AR-process only depends on the process parameters, the proposed estimation procedure will immediately yield an estimate for the signal to noise ratio of the observed signal and allows a proper reconstruction of the process $(\underline{x}_i)$.

1.4 We will first derive estimates for the variances σ_e^2 and σ_n^2 based on the power spectrum of the observed process $\underline{y}_i$ and then show that the resulting equation system is identical to that obtained by a set up in the spatial domain using variance component estimation technique and assuming the process $(\underline{y}_i)$ to be periodic. The 3rd section analyses the estimability or determinability and the identifiability or discernability of the variances for a doubly integrated white noise process which is already in use for height interpolation in photogrammetry (Ebner 1979). The 4th section discusses the numerical effort for the variance component estimation and the versatility of the approach.

2. Estimation of the variances of the driving and the observing process

2.1 The Power Spectrum of Observed AR-Processes

The power spectrum of the AR-process eq. (1) only depends on the parameters a_k and the variance σ_e^2 of the driving process $(\underline{e}_i)$. It is given by (cf. e.g. Lücker (1980), p. 52)

$$P_x(u) = \frac{\sigma_e^2}{\left|1-\sum\limits_{k=1}^{p} a_k\, e^{-j2\pi uk}\right|^2} = T(u)\cdot\sigma_e^2 \quad . \tag{3}$$

With $(\underline{n}_i)$ being white noise and independent of $(\underline{e}_i)$ the power spectrum of $(\underline{y}_i)$ is immediately

$$P_y(u) = P_x(u) + P_n(u) = T(u)\;\sigma_e^2 + \sigma_n^2 \tag{4}$$

$T(u)$ is the squared transfer function of the system yielding $(\underline{x}_i)$ from $(\underline{e}_i)$.

The total <u>variance</u> $\int P_y(u)\,du$ of the process is spread over the frequencies u and consists of two <u>components</u>. As $T(u)$, depending on a_k only, is assumed to be known, each value of the power spectrum is a linear function of the two unknown variance components σ_e^2 and σ_n^2.

<u>Example 1:</u> The non-stationary AR(2) process with $a_1 = +2$ and $a_2 = -1$ can be used to describe the heights in a terrain profile (cf. Ebner, 1979). It is a doubly integrated white noise process. The power spectrum of an observed profile is then given by eq.(4) using

$$T(u) = \frac{1}{16\,\sin^4 \pi u} \tag{5}$$

The Wiener or least squares filter for estimating (x_i) from (y_i) is (cf. e. g. Castleman (1979), cf. also Link (1983)):

$$H(u) = \frac{P_x(u)}{P_x(u)+P_n(u)} = \frac{1}{1+\dfrac{\sigma_n^2}{\sigma_e^2}\cdot 16\,\sin^4 \pi u} \tag{6}$$

and only depends on the ratio σ_n^2/σ_e^2 between the two unknown variances. We will refer to this ratio later in conjunction with the least squares solution in the spatial domain. ■

2.2 Estimating Variance Components from the Power Spectrum $P_y(u)$

The power spectrum $P_y(u)$ may be estimated from $(\underline{y}_i)$ in various ways (cf. the review Kay/Marple, 1981) and leads to an estimate $\hat{\underline{P}}_y(u)$. If the process (y_i) is periodic and Gaussian $\hat{\underline{P}}_y(u)$ reasonably can be estimated using a discrete Fourier transformation. Then the elements of $\hat{\underline{P}}_y(u)$ are independently (cf. Papoulis, 1965, p.368) χ^2_2 distributed with variance $V(\hat{\underline{P}}_y(u)) = P^2_y(u)$, as they are derived from the normally distributed complex amplitude spectrum of $(\underline{y}_i)$. This leads to the following variance component model:

$$E(\hat{\underline{P}}_y(u)) = T(u)\,\sigma^2_e + \sigma^2_n \quad ; \qquad V(\hat{\underline{P}}_y(u)) = P^2_y(u) \tag{7}$$

With approximate values $\sigma^2_e(0)$ and $\sigma^2_n(0)$ one can now derive estimates $\hat{\underline{\phi}}_e$ and $\hat{\underline{\phi}}_n$ for the variance factors using weighted least squares technique. The normal equations are

$$S\,\hat{\underline{\phi}} = \underline{\omega} \tag{8a}$$

$$\begin{bmatrix} \sum\limits_u \dfrac{T^2(u)\cdot\sigma^4_e}{P^2_y(u)} & \sum\limits_u \dfrac{T(u)\,\sigma^2_e\,\sigma^2_n}{P^2_y(u)} \\ \sum\limits_u \dfrac{T(u)\cdot\sigma^2_e\,\sigma^2_n}{P^2_y(u)} & \sum\limits_u \dfrac{\sigma^4_n}{P^2_y(u)} \end{bmatrix} \begin{bmatrix} \hat{\underline{\phi}}_e \\ \hat{\underline{\phi}}_n \end{bmatrix} = \begin{bmatrix} \sum\limits_u \dfrac{T(u)\cdot\sigma^2_e\cdot\hat{\underline{P}}_y(u)}{P^2_y(u)} \\ \sum\limits_u \dfrac{\sigma^2_n\cdot\hat{\underline{P}}_y(u)}{P^2_y(u)} \end{bmatrix} \tag{8b}$$

As $P_y(u)$ is unknown but needed for determining the weights

$$w(u) = 1\,/\,V(\hat{\underline{P}}_y(u)) \;, \tag{9}$$

it may be substituted by $\hat{\underline{P}}_y(u)$. Then eq.(8) becomes nonlinear in the unknowns and has to be solved iteratively (cf. Schaffrin, 1983) by setting

$$\hat{\underline{\sigma}}^2_i(\nu+1) = \hat{\underline{\sigma}}^2_i(\nu)\cdot\hat{\underline{\phi}}_i(\nu) \quad ; \quad i = e,\ n \tag{10}$$

2.3 Estimating Variance Components from the Process (y_i)

The derivation in the spatial domain is more extensive than in the spectral domain. We start from the AR-model eqs. (1) and (2) and rewrite it in the form of a linear Gauss-Markov model (GMM):

$$0 = E(\underline{e}_i) = -x_i + \sum_{k=1}^{p} a_k\, x_{i-k} \quad ; \qquad V(\underline{e}) = \sigma^2_e \qquad i = p+1,\ m \tag{11}$$

$$E(\underline{y}_i) = x_i \qquad\qquad ; \qquad V(\underline{y}) = \sigma^2_n \qquad i = 1,\ m \tag{12}$$

The expectation of the $m-p$ prediction errors $\underline{e}_i$ are zero. They can be treated as fictions observations with value 0 and variance σ^2_e. Together with the m observed values $\underline{y}_i$ we have $2m-p$ observations for the unknown x_i of the AR-process, leading to a redundancy of $r = m-p$. In general not all x_i need to be observed, moreover they need not be observed directly, allowing irregular gaps between the observations, or observations at arbitrary points t between two grid points t_i and t_{i+1}.

In this context we restrict the process y_i to be periodic making a comparison with the previous results possible. Then additional p equations in the form of eq.(11) are available extending the range of the parameter i $(i = 1,m)$ and thus increasing the redundancy to $r = m$. The complete GMM then reads as:

$$E(\underline{l}) = Ax \quad ; \quad V(\underline{l}) = C = \sum_{i=1}^{2} \phi_i Q_i \tag{13}$$

with

$$\underline{l} = \begin{pmatrix} \underline{e} \\ \underline{y} \end{pmatrix} ; \; \begin{matrix} \underline{e} = (\underline{e}_i) \\ \underline{y} = (\underline{y}_i) \end{matrix} \quad A = \begin{pmatrix} A_1 \\ I \end{pmatrix} ; \quad Q_1 = \begin{pmatrix} \sigma_e^2 I & 0 \\ 0 & 0 \end{pmatrix} ; \quad Q_2 = \begin{pmatrix} 0 & 0 \\ 0 & \sigma_n^2 I \end{pmatrix}$$

and

$$A_1 = \begin{bmatrix} -1 & a_1 & a_2 & \cdots\cdots & a_k & \cdots & 0 \\ 0 & -1 & a_1 & \cdots\cdots & a_{k-1} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & & \vdots \\ a_1 & a_2 & \cdots & a_k & 0 \cdots & 0 \cdots & -1 \end{bmatrix} = circ \begin{bmatrix} -1 \\ 0 \\ \vdots \\ 0 \\ a_k \\ \vdots \\ a_1 \end{bmatrix}$$

The periodicy of (y_i) is reflected in the circulant matrix A_1 with kernel vector $[-1\ 0..0\ a_k..a_1]^T$. The filtered values $\hat{\underline{x}}_i$ can be estimated from the normal equation system

$$N\,\hat{\underline{x}} = \underline{h} \quad \text{with} \quad N = \sigma_e^{-2} A_1^T A_1 + \sigma_n^{-2} I \quad ; \quad \underline{h} = \sigma_n^{-2} \underline{y} \tag{14}$$

if the variances σ_e^2 and σ_n^2, or at least the ratio σ_n^2/σ_e^2 is known.

The variances however can be estimated using variance component estimation technique (Helmert 1924, Grafarend/d'Hone 1978, Koch 1979, Förstner 1979).

The variance factors ϕ_e and ϕ_n can be obtained from the equation system

$$S\,\hat{\underline{\phi}} = \underline{\omega} \tag{15}$$

with

$$S = (s_{ij}) = (tr\,(C^{-1} D\, Q_i\, C^{-1} D\, Q_j)) \quad , \quad \hat{\underline{\phi}}^T = (\hat{\underline{\phi}}_e, \hat{\underline{\phi}}_n)$$

$$\underline{\omega} = (\underline{\omega}_i) = (tr\,(\underline{l}^T C^{-1} D\, Q_i\, C^{-1} D\, \underline{l})) \quad , \quad i = e, n$$

$$D = I - A(A^T C^{-1} A)^{-1} A^T C^{-1}$$

The elements s_{ij} in our case are (with $Q_1 = \sigma_e^2 I$ and $Q_2 = \sigma_n^2 I$)

$$s_{11} = tr\,[(I - A_1 N^{-1} A_1^T Q_1^{-1})^2] \tag{16a}$$

$$s_{12} = tr\,[N^{-1} A_1^T Q_1^{-1} A_1 N^{-1} Q_2^{-1}] = s_{21} \tag{16b}$$

$$s_{22} = tr\,[(I - N^{-1} Q_2^{-1})^2] \tag{16c}$$

If we now exploit the special structure of the matrices A_1, N and Q_i, all three being circulant matrices, we can further simplify the expressions.

With the unitarian matrix

$$F = (f_{il}) = \left(\frac{1}{\sqrt{m}}\; e^{-j2\pi i l/m} \right) \tag{17}$$

the circulant matrices A_1 and N can be diagonalized (cf. Zurmühl 1964, Klein 1976):

$$F\,A_1\,F^{-1} = diag(\lambda_u) \quad ; \qquad F\,N\,F^{-1} = diag(\sigma_e^{-2}|\lambda_u|^2 + \sigma_n^{-2})\,;\;\; u = 1, n \tag{18}$$

Pre- and post-multiplying the matrices N, A_1 etc. in eq.(16) with F and $F^{-1} = F^*$ does not change the values s_{ij} but allows to write the traces as sums of the eigenvalues. With $\Lambda(u) = |\lambda_u|^2$ this finally leads to

$$s_{11} = \sum_u \left(1 - \frac{\Lambda(u)\,\sigma_e^{-2}}{\Lambda(u)\,\sigma_e^{-2} + \sigma_n^{-2}}\right)^2 = \sum_u \frac{\sigma_n^{-4}}{(\Lambda(u)\,\sigma_e^{-2} + \sigma_n^{-2})^2}$$

or

$$s_{11} = \sum_u \frac{\sigma_e^4}{(\Lambda(u)\,\sigma_n^2 + \sigma_e^2)^2} = \sum_u \frac{\Lambda^{-1}(u)\,\sigma_e^4}{(\Lambda^{-1}(u)\,\sigma_e^2 + \sigma_n^2)^2} \tag{19a,b}$$

and analogously

$$s_{12} = \sum_u \frac{\Lambda(u)\,\sigma_n^2\,\sigma_e^2}{(\Lambda(u)\,\sigma_n^2 + \sigma_e^2)^2} = \sum_u \frac{\Lambda^{-1}(u)\,\sigma_e^2\,\sigma_n^2}{(\Lambda^{-1}(u)\,\sigma_e^2 + \sigma_n^2)^2} \tag{20a,b}$$

$$s_{22} = \sum_u \frac{\Lambda^2(u)\,\sigma_n^4}{(\Lambda(u)\,\sigma_n^2 + \sigma_e^2)^2} = \sum_u \frac{\sigma_n^4}{(\Lambda^{-1}(u)\,\sigma_e^2 + \sigma_n^2)^2} \tag{21a,b}$$

The right sides of eq.(15) can be treated similarily. If we now substitute $\Lambda^{-1}(u)$ by $T(u)$ in eqs. (19b), (20b) and (21b) we immediately obtain eq. (8b). The equivalence of eq.(8) and (19)-(21) is not surprising as the transformation eq.(18) is identical to the discrete Fourier transformation, which has also been used for the estimation of $P_y(u)$. But we are now able to use both, the spatial and the spectral, version of the estimation procedure to advantage.

Example 2: Figure 1 shows the graph of a terrain profile with 150 points derived from aerial photographs of scale 1 : 28 000 using photogrammetric measuring device (Zeiss Planicomp 100). The precision of the operator was determined using the model eq.(13) based on the AR(2)-process from example 1 $(a_1 = 2,\ a_2 = -1)$ (cf. Lindlohr, 1982). The estimated standard error $\hat{\sigma}_n$ was $0.45\,m$. This is ca. 0,1‰ of the flying height of 4300 m over the terrain and in full agreement with photogrammetric experience. ■

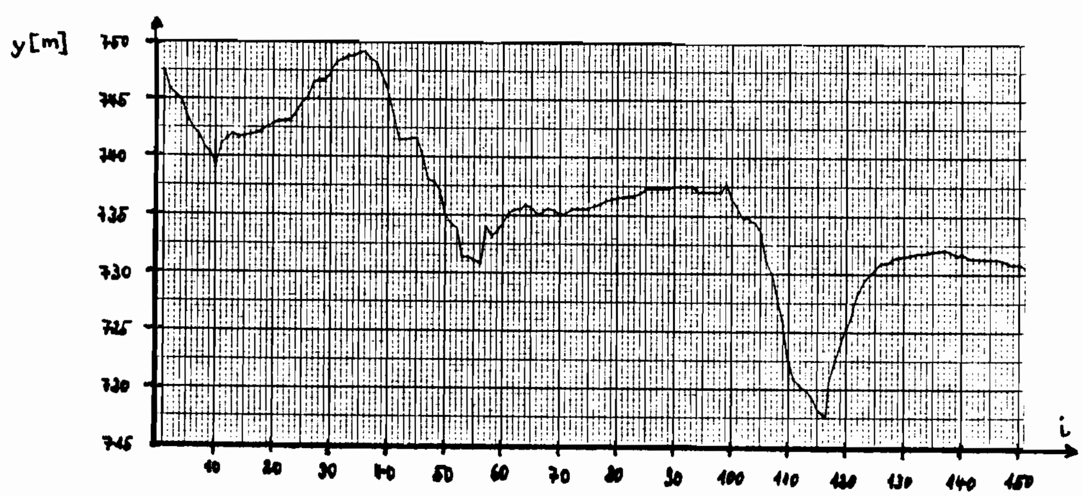

Fig. 1 Terrain profile, point distance 10m

3. Evaluation of the Estimated Variances

3.1 Determinability and Separability

The properties of various estimators for variance components have been analysed by Schaffrin (1983). The estimated variances from eq.(15) are best invariant quadratic estimators with minimum bias (BIQUAMBE). Though their distribution is not known, one can derive their variance under the assumption that $(\underline{y}_i)$ is Gaussian (cf. Koch, 1979, p. 352):

$$V(\hat{\underline{\phi}}_i) = 2 \cdot (S^{-1})_{ii} \qquad (22)$$

They give an indication about the estimability or the determinability of the variances σ_e^2 and σ_n^2. If the standard deviations $\sigma_{\hat{\phi}_i}$ are below 0.2 the variances σ_i^2 can be said to be well determinable, as they are accurate up to 20 %.

The correlation

$$\rho_{12} = \frac{(S^{-1})_{12}}{\sqrt{(S^{-1})_{11}\,(S^{-1})_{22}}} \qquad (23)$$

between the estimates on the other hand is a measure for the discernability or the identifiability. If the correlation coefficient ρ_{12} is less than, say, *75 %* the variances are well discernable. Then with a high probability (of ca. 95 %) one will not identify the observational noise as signal or vice versa. This measure is derived from testing multiple linear hypothesis (Förstner, 1983) and seems to be useful here also.

3.2 Analysis of an AR(2) Process

We will now investigate, under which conditions the signal and the noise in an observed process are estimable or separable. In order to get an idea of the features of the estimation process the already mentioned AR-2 process, with $a_1 = 2$ und $a_2 = -1$, has been analysed in detail.

Fig. 2 shows the variances and the correlation of the estimated variance components σ_e^2 and σ_n^2 in dependency of their ratio σ_n^2/σ_e^2 .

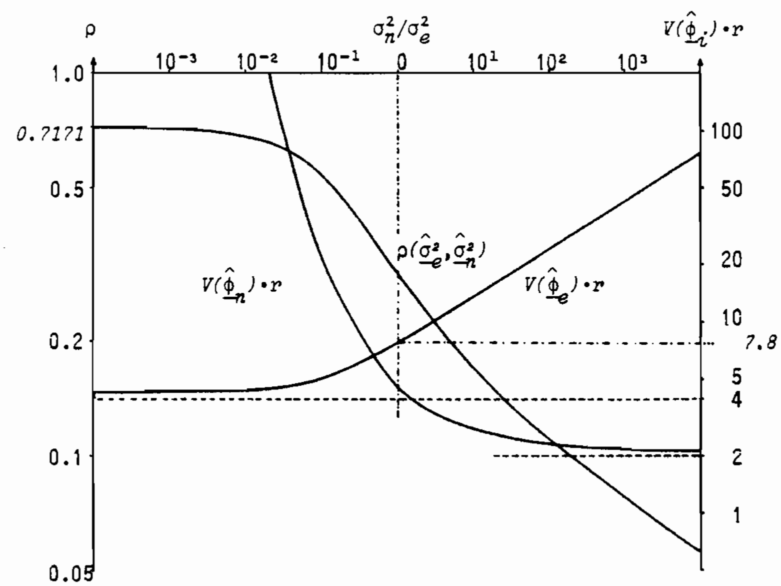

Fig. 2 Estimability and Separability of Variance Components $\hat{\underline{\sigma}}_e^2$ and $\hat{\underline{\sigma}}_n^2$

Instead of $V(\hat{\underline{\sigma}}_i^2)$ the relative accuracy, i. e. the variances $V(\hat{\phi}_i^2) \cdot r$ of the variance factors are given. They also depend on the redundancy r which for periodic sequences equals the number m of the observations. One can derive from fig. 2 that the standard deviation of the additional noise variance $\hat{\underline{\sigma}}_n^2$ is $\hat{\underline{\sigma}}_n^2 \cdot \sqrt{7.8/100} = \hat{\underline{\sigma}}_n^2 \cdot 0.28$, or $28\,\%$, if $\sigma_n^2/\sigma_e^2 = 1$ and $r = m = 100$ observations are made.

The variances and the correlation of the variance components have been calculated from eq.(16) using simulated data. The used processes, with $m = 100$ observations each, were not periodic (Lindlohr,1982). Independently the values were derived from the theoretical power spectra, thus representing periodic processes. The sums in eq. (8) were replaced by integrals assuming a sufficiently large number m of observations. E. g. the element s_{11} then reads as (cf. eq. (19a)):

$$s_{11}\left(\frac{\sigma_n^2}{\sigma_e^2}\right) = m \int_{-1/2}^{+1/2} \frac{T^2(u)\,\sigma_e^4}{P_y^2(u)}\, du = m \int_{-1/2}^{+1/2} \frac{du}{\left(\frac{\sigma_n^2}{\sigma_e^2}\, 16\, \sin^4 \pi u + 1\right)^2} \tag{24}$$

The integrals for s_{11}, s_{12} and s_{22} were solved numerically with a HP 15C computer. The results of both calculations differed not more than 1 %, except for some profiles with very low signal to noise ratio. This demonstrates the low influence of the border effects.

The correlation between the estimated variances never exceeds 75 %. The maximum value is $\sqrt{18/35} = 0.7171$ and is reached for small observational errors $(\sigma_n^2/\sigma_e^2 \to 0)$. Thus signal and noise are always separable, the procedure will not interprete noise as signal or vice versa.

The variances σ_e^2 and σ_n^2 are not always determinable, though. If the variances are of different order, i. e. the ratio σ_n^2/σ_e^2 is very different from 1, then only the larger variance can be estimated with sufficient accuracy. In the extreme cases $(r \to 0, r \to \infty)$ the relative accuracy is $4/r$ and $2/r$ for the variances σ_e^2 and σ_n^2 resp. The last value $2/r$ is identical with the variance of the estimated variance factor σ_o^2 of a least squares estimation. The additional noise variance σ_n^2 obviously can only be determined if it is not much smaller than the variance σ_e^2 of the driving process. On the other hand, even for strongly contaminated signals, σ_e^2 is estimable, though with only moderate accuracy.

These results are representative for observed processes where the spectral properties of signal and noise are different, specifically if the power spectra differ in shape. If, in our case, the noise would have been correlated, e. g. according to an autoregressive scheme of order 1, the separability would have been much less, due to the similarity of the power spectra.

4. Numerical Considerations

4.1 Irregular Observations

Up to now we always have assumed that all signal values x_i have been observed. But the estimation of the variance components is also possible if the sequence of observations is irregular. This is of great practical importance as it increases the flexibility of the procedure.

If not all signal values are observed eq.(12) only is valid for the m_n observations y_i. Then, the special structure of A is lost. The normal equation matrix N in eq.(14) is still band limited, with band width p. The prediction of the x_i from the observed values $\underline{y}_i$ needs about $mp^2/6$ operations (cf. Ebner et. al. 1984).

On the other hand, the effort for calculating the elements s_{ij} for the variance estimation is prohibitive, as all elements of the inverse N^{-1} are needed, which requires appr. $m^3/2$ operations. But it is possible to reduce the effort considerably, if one uses a slightly different iteration scheme to solve the nonlinear (cf. the text after eq.(8b)) equation system.

The following equation system (cf. Förstner, 1979)

$$\begin{bmatrix} \bar{s}_{11} & 0 \\ 0 & \bar{s}_{22} \end{bmatrix} \begin{bmatrix} \hat{\underline{\phi}}_e \\ \hat{\underline{\phi}}_n \end{bmatrix} = \begin{bmatrix} s_{11} + s_{12} & 0 \\ 0 & s_{21} + s_{22} \end{bmatrix} \begin{bmatrix} \hat{\underline{\phi}}_e \\ \hat{\underline{\phi}}_n \end{bmatrix} = \begin{bmatrix} \underline{\omega}_e \\ \underline{\omega}_n \end{bmatrix} \tag{25}$$

diagonalizes eq.(15) and after convergence leads to the same result, as $\hat{\underline{\phi}}_e$ and $\hat{\underline{\phi}}_n$ then equal 1 (cf. Schaffrin, 1983). But the sums $\bar{s}_{11}$ and $\bar{s}_{12}$ are much easier to obtain. First observe, that

$$\bar{s}_{11} + \bar{s}_{22} = tr\ D = r = m_n - p \quad , \tag{26}$$

the total redundancy of the system. Now, $\bar{s}_{22}$ can be calculated from

$$\bar{s}_{22} = tr\ [(C - A\ N^{-1} A^T)\ Q_2^{-1}] = m_n - tr\ N^{-1}/\sigma_n^2 \tag{27}$$

where m_n is the number of observed values $\underline{y}_i$, possibly not equal to m. The main effort now is to determine $tr\ N^{-1}$. But as N is band limited and only the elements of N^{-1} within the band are necessary, the number of operations is only $mp^2/2$, which is considerably less than $m^3/2$. Therefore with

$$\bar{s}_{11} = r - \bar{s}_{22} \tag{28}$$

from eq.(26) the solution of eq.(25) can directly be given:

$$\hat{\underline{\phi}}_e = \frac{\underline{\omega}_e}{r - \bar{s}_{22}} = \frac{\hat{\underline{e}}^T \hat{\underline{e}}/\sigma_e^2}{r - \bar{s}_{22}} \quad ; \quad \text{with the prediction errors } \hat{\underline{e}} = A_1\, \hat{\underline{x}} \tag{29}$$

and

$$\hat{\underline{\phi}}_n = \frac{\underline{\omega}_n}{\bar{s}_{22}} = \frac{\hat{\underline{n}}^T \hat{\underline{n}}/\sigma_n^2}{\bar{s}_{22}} \quad ; \quad \text{with the residuals } \hat{\underline{n}}_i = \underline{y}_i - \hat{\underline{x}}_i \tag{30}$$

and $\bar{s}_{22}$ from eq.(27) .

Thus the total effort for estimating the variance components especially the additional noise variance σ_n^2 is appr. 3 times the effort for the prediction of $(\underline{x}_i)$ alone.

The simplification has the disadvantage that the speed of convergence is reduced and the information about the separability is not available. The convergence can be increased by numerical methods, which are discussed by Schaffrin (1983) with special emphasis on variance component estimation. On the other hand, the correlation of the variance components may be approximated by the theoretical values (cf. 3.2).

4.2 Regular Observations

If all values of the process $(\underline{x}_i)$ are observed with no gaps one might distinguish two cases:

a. For rather short sequences ($m < 20$) the direct calculation according to eqs. (15) and (16) seems seems to be feasible, if the sparsity of A_1 and the diagonality of the Q_i are used to advantage.

b. For longer sequences the estimation of the prediction errors $\hat{\underline{e}}_i$ and the residuals $\hat{\underline{n}}_i$ could be achieved from eq. (14). The variance component estimation could reasonably neglect the border effects and the nonperiodicity of the sequence and calculate the elements s_{ij} directly from eqs. (19)-(21), whereas the right sides could be derived from $\hat{\underline{e}}$ and $\hat{\underline{n}}$ (cf. eq. (29)(30)). In this case the additional effort for the variance estimation becomes negligible.

5. Discussion

The estimation of variances in observed autoregressive processes can be accomplished in a statistically rigorous manner from a single sequence of observations thus not needing more information than the classical identification procedures.

The interpretation of the variance component estimation procedure as weighted least squares solution for the composite power spectrum enables a simple theoretical analysis of the model in the spectral domain and at the same time numerical advantages. The solution in the spatial domain is very flexible allowing irregular gaps, indirect observations such as slopes or curvatures or observations between the grid points. An evaluation can be based on the variance covariance matrix of the estimated variances. Specifically, the estimability and the separability of the variance components can be derived, which for a special process have been discussed in detail, demonstrating the feasibility of the approach.

The method easily can be extended towards more general processes including autoregressive moving-average processes or vector valued processes. The solution in the spectral domain may even be used for processes with arbitrary power spectrum.

REFERENCES

Box G.E.P and Jenkins G.M. (1976): Time Series Analysis, Holden-Day, 1976

Castleman K.R. (1976): Digital Image Processing, Prentice Hall, New Jersey, 1979

Ebner H. (1979): Zwei neue Interpolationsverfahren und Beispiele für ihre Anwendung, Bildmessung und Luftbildwesen, 1979, pp. 15-27

Ebner H., Hössler R. and Joos D. (1984): Random Processes and their Application to linear Prediction Problems - In Memoriam Karl Ramsayer -, Manuscripta Geodaetica, Vol. 9, 1984, pp. 21-43

Förstner, W. (1979): Ein Verfahren zur Schätzung von Varianz- und Kovarianzkomponenten, Allgemeine Vermessungsnachrichten 11-12, 1979, pp.446-453

Förstner W.(1983): Reliability and Discernability of Extended Gauß-Markov-Models, Deutsche Geod. Komm., Reihe A, Heft 98, pp. 79-103

Grafarend E. and d'Hone A. (1978): Gewichtsschätzung in geodätischen Netzen, Deutsche Geodätische Kommission, Reihe A, Nr. 88 München, 1978

Helmert F.R.(1294): Die Ausgleichung nach der Methode der kleinsten Quadrate, 3. Auflage, Teubner, Berlin, 1924

Jäckle R. (1984): Analyse von Geländeprofilen, Selbst. Geod. Arbeit am Inst. f. Photogrammetrie, Stuttgart, 1984

Kashyap R.L. and Rao A.R. (1976): Dynamic Stochastical Models from Empirical Data, Academic Press, 1976

Kay S.M. and Marple S.L. (1981): Spectrum Analysis - A modern Perspective, Proc. IEEE, Vol. 69, 1981, pp. 1380-1418

Klein W. (1976): Finite Systemtheorie, Teubner, Stuttgart, 1976

Koch K.-R.(1979): Parameterschätzung und Hypothesentests in linearen Modellen, Dümmler, 1979

Lindlohr W. (1982): Untersuchungen zur Varianzkomponentenschätzung bei der Interpolation von Geländeprofilen, Diplomarbeit am Inst. f. Photogrammetrie, Stuttgart, 1982

Link M. (1983): Eine Einführung in die Signaltheorie und ihre Anwendung auf die Interpolation mit finiten Elementen, Selbst. Geod. Arbeit am Inst. f. Photogrammetrie, Stuttgart, 1983

*Lücker R. (1980):*Grundlagen digitaler Filter, Springer, 1980

Papoulis A. (1965): Probability, Random Variables and Stochastic Processes, McGraw-Hill, Tokyo, 1965

Sanso F. (1984): The Analysis of Time Series with Applications to Geodetic Control Systems, Internat. School of Geodesy, 3rd Course on Optimization and Design of Geodetic Networks, Erice, 1984

Schaffrin B. (1983): Varianz-Kovarianz-Komponentenschätzung bei der Ausgleichung heterogener Wiederholungsmessungen, Deutsche Geod. Kommission, Reihe C. Nr. 282, München 1983

Yum Y. and Park S.B. (1983): Optimum Recursive Filtering of Noisy Two-Dimensional Data with Sequential Parameter Identification. IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. PAMI-5, 1983, pp.337-344