NORSAR ROYAL NORWEGIAN COUNCIL FOR SCIENTIFIC AND INDUSTRIAL RESEARCH

Scientific Report No. 1-73/74

ESTIMATION OF SIGNALS IN MULTIPLE NOISE

A UNIFIED APPROACH

by
A. Christoffersson
Department of Statistics
Uppsala University
Sweden
and
B. Jansson
Honeywell Bull
Stockholm, Sweden

Kjeller, 10 January 1974

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER F44620-74-C-0001	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Estimation of Signals in Multiple Noise A Unified Approach		5. TYPE OF REPORT & PERIOD COVERED Scientific 1 Jul 73 - 30 Jun 74
		6. PERFORMING ORG. REPORT NUMBER Scientific Rep. 1-73/74
7. AUTHOR(s) Anders Christoffersson and Bo Jansson		8. CONTRACT OR GRANT NUMBER(s) F44620-74-C-0001
9. PERFORMING ORGANIZATION NAME AND ADDRESS NTNF/NORSAR Post Box 51 N-2007 Kjeller, Norway		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NORSAR Phase 3
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research 1400 Wilson Blvd Arlington, Va. 22209 U.S.A.		12. REPORT DATE 1 December 1973
		13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office) European Office of Aerospace Research And Development Keysign House, 429 Oxford Street London W1, England Attn: Maj. Munzlinger		15. SECURITY CLASS. (of this report)
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Array, Beamforming, Signal Estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A unified approach for the estimation of signals in multiple noise is given. In the general form of the method we will allow for several signals to be estimated simultaneously. It is possible to utilize prior knowledge of the signal. Existing methods fall out as special cases as we simplify our model. The large sample statistical properties are studied and it is proved that use of prior knowledge is essential for consistent estimates.

USAF Project Authorization No.	:	VT/4702/B/OSR
Date of Contract	:	30 August 1973
Amount of Contract	:	$ 888,806.00
Contract Termination Date	:	30 June 1974
Project Supervisor	:	Robert Major, NTNF
Project Manager	:	Nils Marås
Title of Contract	:	Norwegian Seismic Array (NORSAR)

ESTIMATION OF SIGNALS IN MULTIPLE NOISE
A UNIFIED APPROACH

SUMMARY

A unified approach for the estimation of signals in multiple noise is given. In the general form of the method we will allow for several signals to be estimated simultaneously. It is possible to utilize prior knowledge of the signal. Existing methods fall out as special cases as we simplify our model. The large sample statistical properties are studied and it is proved that use of prior knowledge is essential for consistent estimates.

INTRODUCTION

Several methods for signal estimation for array stations have been published during the last few years, e.g., Capon et al (1966), Kelly and Levin (1964). Common to these methods is that they operate on the assumption that the signal, or event, in each sensor is identical to that in every other sensor (see (ii) below). In seismological applications it has been found that the signal appears with different amplitude at each sensor for rather close spacing of the sensors. (See Capon et al (1964), page 27.) Hence there will be a degradation in the performance of the methods.

A general model for array data can be written as

$$
\text{(i)} \quad
\begin{aligned}
y_1 &= S_1 + N_1 \\
y_2 &= S_2 + N_2 \\
&\;\;\vdots \\
y_M &= S_M + N_M
\end{aligned}
$$

where y_i is the recording at the i-th sensor, S_i is the signal at the i-th sensor and N_i is the noise at the i-th sensor. The noise is assumed to be uncorrelated with signals and to have expectation zero.

In this model, $y_1, y_2, \ldots, y_M$ are observed and we want to estimate the unknown signals $S_1, S_2, \ldots, S_M$. To estimate the signals we may simply use the observed y_i:s. This will work satisfactorily if the signal-to-noise ratios are high. However, if this is not true, we must try to find some reasonable restrictions to put on the model.

Assuming that the readings are properly shifted, these restrictions may be of the following types:

I. We assume that the signal is the same at every seismometer. The model then becomes

$$
\begin{aligned}
y_1 &= S + N_1 \\
y_2 &= S + N_2 \\
&\;\,\vdots \\
y_M &= S + N_M
\end{aligned}
\qquad \text{(ii)}
$$

To estimate the signal in this model we may simply sum up the traces. However, this model is rather unrealistic. In order to get $S_1 = S_2 = \ldots = S_M$, the sensors have to be spaced rather closely, and this may cause high cross-correlation between the different noise records. This in turn implies that the information about the signal that is contained in one or a few of the sensors is almost the same as the information contained in the whole array. (Note that this is true even if knowledge of the noise coherence is utilized.)

II. To make the model less restrictive and thus more realistic, we may assume that the signal is the same except for an unknown amplitude factor. Then we get

$$
\text{(iii)} \quad
\begin{aligned}
y_1 &= \gamma_1 S + N_1 \\
y_2 &= \gamma_2 S + N_2 \\
&\;\;\vdots \\
y_M &= \gamma_M S + N_M
\end{aligned}
$$

where γ_i:s are the unknown amplitude factors.

The estimation of the unknown signal and the unknown amplitude factors leads to an eigenvalue problem and will be treated in section two of this paper. The stochastic properties of the parameters in the model (iii) are treated in section three. In this model we may also make use of available knowledge about the signal, e.g., that it can be expanded in a finite set of linearly independent functions. (See Broome and Dean (1964).) We then add to the model the following condition

$$\text{(iiia)} \qquad S = \sum_{r}^{p} \phi_r a_r$$

where $a_1, a_2, \ldots, a_p$ is the set of given linearly independent functions and $\phi_1, \phi_2, \ldots, \phi_p$ is a set of unknown parameters.

The square of the signal-to-noise ratio in the estimated signal will, for models (ii) and (iii), increase with the number of seismometers (if the noise coherence is small). For model (iii) with condition (iiia) the square of the signal-to-noise ratio will increase with the number of sensors multiplied with the length of the signal.

The properties of the model with the additional condition (iiia) are treated in sections two and three.

III. The next step to get a more general model is to assume that the space spanned by the signals $S_1, S_2, \ldots, S_M$ in the model (i) have dimension N, where $1 \leq N \leq M$. The model is then

(iv)
$$
\begin{aligned}
y_1 &= \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \ldots + \gamma_{1N}\xi_N + N_1 \\
y_2 &= \gamma_{21}\xi_1 + \gamma_{22}\xi_2 + \ldots + \gamma_{2N}\xi_N + N_2 \\
&\ \ \vdots \\
y_M &= \gamma_{M1}\xi_1 + \gamma_{M2}\xi_2 + \ldots + \gamma_{MN}\xi_N + N_M
\end{aligned}
$$

In this model $\xi_1, \xi_2, \ldots, \xi_N$ form an unknown basis in the space spanned by the signals $S_1, S_2, \ldots, S_M$ and $\gamma_{11}, \gamma_{21}, \ldots, \gamma_{MN}$ are unknown coefficients for the expansion of the signals in terms of this basis, i.e., the signals are

(iva) $$S_i = \sum_j^N \gamma_{ij}\xi_j \qquad \text{for } i = 1,2,\ldots,M$$

The estimate of the space spanned by the signals and the estimates of the signals will be unique. The estimate of the basis on the other hand is unique only up to a non-singular transformation. The assumption that the dimension of the space spanned by the signals equals N can be derived from an assumption that the geological structure under the array consists of N different unknown homogenous regions. It is then possible to transform (rotate) the solution in such a way that each of the basis vectors corresponds to one of the regions. It may also be possible to relate a given sensor to a specific structure. One way to do this will be treated in sections five and six.

As for the model (iii)it is possible to utilize prior knowledge of the signal. Corresponding to (iiia) we then get the condition

(ivb) $$\xi_j = \sum_r^p b_{jr} a_r$$

where $a_1, a_2, \ldots, a_p$ is a set of given linearly independent functions and $b_{11}, b_{12}, \ldots, b_p$ is a set of unknown parameters.

The first two sections of this paper, based on an earlier seminar paper (Christoffersson and Jansson (1966)) which was used in Jansson and Husebye (1968) and Whitcomb (1969), treats the model (iii) and gives the maximum likelihood solution. Section three gives the large sample properties of the estimates. Section four illustrates how the method can be applied to real data. Sections five and six, finally, treat the estimatation of the unknowns in model (iv) and the rotation of the solution so that the basis vectors correspond to signals present in different regions.

The assumption that the noise is normally distributed, which we adopt here, makes the maximum likelihood solution identical to the least-square solution, where no assumption is made about the distribution of the noise.

1. Theoretical Model

The theoretical model is written

$$\eta = \Gamma'(\Phi'\alpha) + \upsilon \qquad (1)$$

where η is a vector variable with M elements (M x 1), Γ is a (N x M) matrix of constants, Φ is a (L x N) matrix of constants in Hilbert space, α is a known non-stochastic variable vector with linearly independent elements, and υ is a stochastic variable vector with $E(\upsilon) = 0$, where the M elements form a joint non-singular normal distribution.

In our application η refers to the measured variables, i.e., the recordings, Γ contains the amplitude factors. In many applications it is known that the signals can be written as linear combinations of some specified variables. Each element in α corresponds to one of the specified variables and the elements of Φ are the unknown coefficients for these linear combinations. If no such knowledge exists about the signal, we can always let the variables in α form a basis in the Hilbert space or more simply we may put $\Xi = \Phi'\alpha$ and rewrite the model as

$$\eta = \Gamma'\Xi + \upsilon \qquad (2)$$

2. The Model when N=1

Assuming that the readings from the M sensors are properly shifted and sampled at K successive timepoints, not necessarily equidistant, we write the model in terms of our observations:

$$Y = A\phi\gamma' + N \qquad (3)$$

where $Y = (y_{ij}) = (\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_M) = \begin{bmatrix} \underline{y}_1^{*\prime} \\ \underline{y}_2^{*\prime} \\ \cdot \\ \cdot \\ \cdot \\ \underline{y}_K^{*\prime} \end{bmatrix}$ is a KxM data matrix

$A = (a_{ij})$ is a K x L, $L \leq K$, matrix where each column contains one of the specified variables, ϕ is a column vector of L constants, γ is a column vector of M constants such that $\gamma'\gamma = 1$, $A\phi$ is the signal, where the columns of A form the basis for the signal. If this basis is unknown, we put A = I.

$N = (n_{ij}) = (\underline{n}_1, \underline{n}_2, \ldots, \underline{n}_M) = \begin{bmatrix} \underline{n}_1^{*\prime} \\ \underline{n}_2^{*\prime} \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_K^{*\prime} \end{bmatrix}$ is a KxM residual matrix, where the elements have joint non-singular normal distribution and are uncorrelated with the signal.

The frequency function of N may be written in two ways.

$$f(N_v) = \text{konst} \cdot e^{-1/2\, N_v' S^{-1} N_v} ;$$

$$f(N_v^*) = \text{konst} \cdot e^{-1/2\, N_v^{*\prime} S^{*-1} N_v^*} \qquad (4)$$

where S and S* denote the moment matrices of N

$$N_v = \begin{bmatrix} \underline{n}_1 \\ \underline{n}_2 \\ \vdots \\ \underline{n}_M \end{bmatrix} \quad \text{and} \quad N_v^* = \begin{bmatrix} \underline{n}_1^* \\ \underline{n}_2^* \\ \vdots \\ \underline{n}_K^* \end{bmatrix}$$

Now let $S^{-1} = Q = [Q_{ij}]$ and $S^{*-1} = Q^* = [Q^*_{ij}]$, where Q_{ij} is a K x K matrix and Q^*_{ij} an M x M matrix.

Then

$$f(N_v) = \text{konst} \cdot e^{-1/2 \sum_i^M \sum_j^M \underline{n}_i' Q_{ij} \underline{n}_j} \tag{5a}$$

and

$$f(N^*_v) = \text{konst} \cdot e^{-1/2 \sum_i^K \sum_j^K \underline{n}^{*\prime}_i Q^*_{ij} \underline{n}^*_j} \tag{5b}$$

Let $X = A\phi$ then

$$\underline{n}_i = \underline{y}_i - \gamma_i A\phi$$

and

$$\underline{n}^*_i = \underline{y}^*_i - x_i\gamma$$

The logarithm of the likelihood function for our observations can be written, except for a constant, as

$$\Omega_O = \sum_i^M \sum_j^M (\underline{y}_i - \gamma_i A\phi)' Q_{ij} (\underline{y}_j - \gamma_j A\phi) =$$

$$= \sum_i^K \sum_j^K (\underline{y}^*_i - x_i\gamma)' Q^*_{ij} (\underline{y}^*_j - x_j\gamma) \tag{6}$$

In order to obtain estimates of ϕ and γ, Ω_O is minimized under the condition that $\gamma'\gamma = 1$, i.e., $\Omega = \Omega_O + \lambda\,(\gamma'\gamma - 1)$ is minimized. Setting the partial derivatives equal to zero we obtain

$$\frac{\partial\Omega}{\partial\phi} = 0 \rightarrow \sum_{i}^{M}\sum_{j}^{M} \gamma_i\gamma_j \; A' \; (Q_{ij} + Q'_{ij}) \; A\phi =$$

$$= \sum_{i}^{M}\sum_{j}^{M} \gamma_i \; A' \; (Q_{ij} + Q'_{ij}) \; \underline{y}_j$$

$$\frac{\partial\Omega}{\partial\gamma} = 0 \rightarrow \left[\sum_{i}^{K}\sum_{j}^{K} x_i x_j \; (Q^*_{ij} + Q^{*\prime}_{ij}) - 2\,\lambda\, I\right] \gamma$$

$$= \sum_{i}^{K}\sum_{j}^{K} x_i \; (Q^*_{ij} + Q^{*\prime}_{ij}) \underline{y}^*_j$$

But since $Q'_{ij} = Q_{ji}$ and $Q^{*\prime}_{ij} = Q^*_{ji}$, we get

$$\sum_{i}^{M}\sum_{j}^{M} \gamma_i\gamma_j \; A'Q_{ij} \; A\phi = \sum_{i}^{M}\sum_{j}^{M} \gamma_i \; A'Q_{ij} \; \underline{y}_j \tag{7a}$$

$$\left[\sum_{i}^{K}\sum_{j}^{K} x_i x_j \; Q^*_{ij} - \lambda I\right] \gamma = \sum_{i}^{K}\sum_{j}^{K} x_i \; Q^*_{ij} \; \underline{y}^*_j \tag{7b}$$

If we replace A in (7a) with I and put all γ_i equal to one, we get a formula for maximum likelihood estimation equivalent to those given in Capon et al 1966 and Kelly and Levin 1964.

To solve the system (7a,b) we first show that the Lagrange multipler λ in eq. 7b is equal to zero. To see this, multiply eq. 7a with ϕ'. This gives

$$\phi'A'\left[\sum_{i}^{M}\sum_{j}^{M} \gamma_i\gamma_j Q_{ij}\right] A\phi = \phi'A'\left[\sum_{i}^{M}\sum_{j}^{M} \gamma_i \; Q_{ij} \; \underline{y}_j\right] \tag{8a}$$

but since $\underline{y}_j = \gamma_j A\phi + \underline{n}_j$ we may write

$$(A\phi)' \left[\sum_i^M \sum_j^M \gamma_i\gamma_j Q_{ij}\right] A\phi = (A\phi)' \left[\sum_i^M \sum_j^M \gamma_i\gamma_j Q_{ij}A\phi + \right.$$

$$\left. \sum_i^M \sum_j^M \gamma_i Q_{ij} \underline{n}_j \right] \tag{8b}$$

This implies that $(A\phi)' \left[\sum_i^M \sum_j^M \gamma_i Q_{ij} \underline{n}_j\right] = 0$ (8c)

Multiplying 7b with γ' and using that $\underline{y}_j^* = x_j \gamma + \underline{n}_j^*$ we have

$$\gamma' \left[\sum_i^K \sum_j^K x_i x_j Q_{ij}^* - \lambda I\right] \gamma = \gamma' \left[\sum_i^K \sum_j^K x_i x_j Q_{ij}^* \gamma + \right.$$

$$\left. \sum_i^K \sum_j^K x_i Q_{ij}^* \underline{n}_j^* \right] \tag{8d}$$

But from the definitions of Q_{ij}^*, Q_{ij}, $\underline{n}_j^*$, $\underline{n}_j$ and x_i it follows that

$$\gamma' \left[\sum_i^K \sum_j^K x_i Q_{ij}^* \underline{n}_j^*\right] = (A\phi)' \left[\sum_i^M \sum_j^M \gamma_i Q_{ij} \underline{n}_j\right] = 0 \tag{8e}$$

Thus $\lambda = 0$.

Since there exists no explicit solution to the system (7a,b), we need to apply an iteration technique. The following iteration scheme is proposed: Start with an arbitrary vector γ, normalized so that $\gamma'\gamma = 1$. Solve for ϕ in eq. 7a. Then compute $x = A\phi$ and solve eq. 7b for γ. Normalize γ such that $\gamma'\gamma = 1$ and continue until the desired convergence criteria are satisfied.

To prove that the above iteration scheme always converges, we need an additional result which is proved in the following lemma.

<u>Lemma</u>

Let A be a real symmetric positive definite matrix. Further, let A be partitioned into M^2 matrices of order K,

$$A = \begin{bmatrix} A_{11} A_{12} \cdots A_{1M} \\ \cdot \\ \cdot \\ \cdot \\ A_{M1} \quad \cdots A_{MM} \end{bmatrix} \tag{9}$$

where

$$A_{ij} = (\alpha_{rs}^{ij}) \text{ is a K x K matrix.} \tag{9a}$$

Let $\underline{x}$ be a non-zero vector with M elements. Then the quadratic form

$$B = \sum_{i}^{M} \sum_{j}^{M} x_i x_j A_{ij} \tag{9b}$$

is a symmetric positive definite K x K matrix.

Proof

There exist $n = M \times K$ linearly independent vectors $\underline{y}_j$ with n elements such that $Y'Y = A$ where the $n \times n$ matrix $Y = (\underline{y}_1, \underline{y}_2, \ldots \underline{y}_n)$ has rank n. Then form new vectors $\underline{y}_i^*$ such that

$$\underline{y}_1^* = x_1\underline{y}_1 + x_2\underline{y}_{K+1} + \ldots x_M\underline{y}_{(M-1)K+1}$$

$$\underline{y}_2^* = x_1\underline{y}_2 + x_2\underline{y}_{K+2} + \ldots x_M\underline{y}_{(M-1)K+2}$$

$$\begin{array}{c}\cdot\\ \cdot\\ \cdot\end{array} \tag{10a}$$

$$\underline{y}_K^* = x_1\underline{y}_K + x_2\underline{y}_{K+K} + \ldots x_M\underline{y}_{(M-1)K+K}$$

or in general

$$\underline{y}_j^* = \sum_i^M x_i\underline{y}_{(i-1)\cdot K+j} \tag{10b}$$

Note that these K vectors are linearly independent.

Now let $Y^* = (\underline{y}_1^*, \underline{y}_2^*, \ldots \underline{y}_K^*)$. Since Y^* has rank K, $Y^{*\prime}Y^*$ has also rank K and is a symmetric positive definite matrix. An arbitrary element b_{rs} in B can be written

$$\sum_i^M \sum_j^M x_i x_j \; a_{rs}^{ij}$$

where a_{rs}^{ij} is element rs in A_{ij}. But a_{rs}^{ij} is also an element in A.

Since A = Y'Y it follows that

$$a_{rs}^{ij} = \underline{y}'_{(i-1)K+r} \, \underline{y}_{(j-1)K+s} \tag{11}$$

Hence b_{rs} can be written

$$b_{rs} = \sum_{i}^{M} \sum_{j}^{M} x_i x_j \, \underline{y}'_{(i-1)K+r} \underline{y}_{(j-1)K+s} \tag{12}$$

Now let $Y^{*\prime}Y^{*} = B^{*} = (b_{rs}^{*})$. It then follows that

$$b_{rs}^{*} = \underline{y}_r' \, \underline{y}_s = \left[\sum_{i}^{M} x_i \, \underline{y}'_{(i-1)K+r}\right] \left[\sum_{i}^{M} x_j \, \underline{y}_{(j-1)K+s}\right] =$$

$$= \sum_{i}^{M} \sum_{j}^{M} x_i x_j \, \underline{y}'_{(i-1)K+r} \, \underline{y}_{(j-1)K+s} = b_{rs} \tag{13}$$

Hence B = B* and the lemma is proved.

The lefthand side of eq. 7a can be written

$$A' \left[\sum_{i}^{M} \sum_{j}^{M} \gamma_i \gamma_j \, Q_{ij}\right] A\phi \tag{14}$$

According to the lemma $\sum_{i}^{M} \sum_{j}^{M} \gamma_i \gamma_j \, Q_{ij}$ is a positive definite matrix of order K and thus ϕ in (14) and (7a) is multiplied by a positive definite matrix of order L. This implies that, for given γ, equation 7a uniquely defines ϕ. In the same way for given ϕ, eq. 7b uniquely defines γ. Now the function to be minimized (eq. 6) is a positive definite quadratic form, and since this function is minimized in each step of the iterations convergence is ensured.

If there is no cross-correlation, i.e., $Q_{ij} = 0$ for $i \neq j$, the system reduces to

$$\sum_i^M \gamma_i^2 \, A'Q_{ii} \, A\phi = \sum_i^M \gamma_i \, A'Q_{ii} \, \underline{y}_i \tag{15a}$$

$$\left[\sum_i^K x_i^2 \, Q^*_{ii} - \lambda I\right] \gamma = \sum_i^K x_i \, Q^*_{ii} \, \underline{y}^*_i \tag{15b}$$

and if the residuals have equal variances and are independent we get

$$\sum_i^M \gamma_i^2 \, A'A\phi = \sum_i^M \gamma_i \, A'\underline{y}_i \tag{16a}$$

$$\left[\sum_i^K x_i^2 - \lambda I\right] \gamma = \sum_i^K x_i \, \underline{y}^*_i \tag{16b}$$

Now let the specified variables be orthonormal then

$$\sum_i^M \gamma_i^2 \, \phi = \sum_i^M \gamma_i \, A'\underline{y}_i \tag{17a}$$

$$\left[\sum_i^K x_i^2 - \lambda I\right] \gamma = \sum_i^K x_i \, \underline{y}^*_i \tag{17b}$$

Rao (1964) obtained eq:s (16a,b) when he derived "principal components of instrumental variables" by least squares.

If A is the identity matrix, then this reduces to

$$\sum_{i}^{M} \gamma_i^2 \, \phi = \sum_{i}^{M} \gamma_i \, \underline{y}_i \tag{18a}$$

$$\left[\sum_{i}^{K} x_i^2 - \lambda I\right] \gamma = \sum_{i}^{K} x_i \, \underline{y}_i^* \tag{18b}$$

which is the solution to the ordinary component model.

3. Large Sample Properties of the Estimates

We shall here consider the cases where there are no cross-correlation, i.e., we assume that $Q_{ij} = 0$ for every $i \neq j$.

3.1 Unknown basis for the signal and no auto correlation

If there is no auto correlation in the noise and the basis for the signal is unknown, maximum likelihood leads, after suitable normalization of the observations, to eq:s (18a,b). These equations define the estimated amplitudes as the eigenvector corresponding to the largest eigenvalue of covariance matrix for the observations. These estimates are consistent and the large sample variances have been given by Andersson (1963), Lyttkens (1966, see Wold 1966) and others. The variances are

$$D^2(\hat{\gamma}_j) = \frac{(\sigma_s^2 + \sigma^2)\ (1 - \gamma_j^2)\ \sigma^2}{\sigma_s^4\ K} \tag{19}$$

where σ_s^2 is the variance for the signal, σ^2 is the variance for the noise, and ^ denotes the estimate of the corresponding parameters.

The estimated signal is

$$\hat{x} = Y'\ \hat{\gamma} \tag{20}$$

Although the amplitudes are estimated consistently, this <u>will not</u> be the case for the signal. This is easily seen in the following way:

Consider the i:th time point. Then

$$\hat{x}_i = \sum_j^M y_{ij}\ \hat{\gamma}_j = x_i \sum_j^M \hat{\gamma}_j\ \gamma_j + \sum_j^M \hat{\gamma}_j\ n_{ij} \tag{21}$$

As the amplitudes are consistent, we find that the limiting distribution of $\hat{x}_i$ is that of

$$x_i + \sum_j^M \gamma_j n_{ij} = x_i + \bar{n}_i \tag{22}$$

where the variance of $\bar{n}_i$ is equal to σ^2.

Thus, the estimate of the signal is inconsistent with a variance that tends to $\sigma_s^2 + \sigma^2$. This gives the square of the signal-to-noise ratio:

$$(S/N)^2 = \frac{\sigma_s^2}{\sigma^2} \tag{23}$$

It may be of interest to compare this ratio with the corresponding ratio for the "simple summing method", i.e., when the signal is estimated with:

$$\bar{x}_i = \frac{1}{\sqrt{M}} \sum_j^M y_{ij} = \frac{1}{\sqrt{M}} (\sum_j^M \gamma_j) x_i + \frac{1}{\sqrt{M}} \sum_j^M n_{ij} \tag{24}$$

The variance for this estimate is

$$D^2(\bar{x}) = \frac{1}{M} (\sum_j^M \gamma_j)^2 \sigma_s^2 + \frac{1}{M} \sum_j^M \sigma^2 = M \bar{\gamma}^2 \sigma_s^2 + \sigma^2 \tag{25}$$

where $\bar{\gamma}$ is the arithmetic mean of the amplitudes.

This gives

$$(S/N)^2 = \frac{M \bar{\gamma}^2 \sigma_s^2}{\sigma^2} \tag{26}$$

But as

$$\sum_{j}^{M} \gamma_j^2 - M \bar{\gamma}^2 \geq 0 \tag{27}$$

with equality if and only if the amplitudes are equal we find that (23) is always larger than or equal to (26).

3.2 Unknown Basis for the Signal and Autocorrelated Noise

For this situation, the maximum likelihood method leads to eq:s (15a, 7b) with A = I and ϕ = x. The properties of the estimates are to be treated in detail elsewhere, but we mention without proof that the estimates of the amplitudes are consistent and that the covariance matrix for the estimated signal is

$$D^2(x) = cov_x + (\sum_{j}^{M} \gamma_j^2 Q_{jj})^{-1} \tag{28a}$$

where cov_x is the auto-covariance matrix of the signal. That is, as in 3.1 we do not obtain consistent estimates of the signal. And the square of the signal-to-noise ratio may be defined as

$$(S/N)^2 = \frac{\sigma_s^2}{\frac{1}{K} tr(\sum_{j}^{M} \gamma_j^2 Q_{jj})^{-1}} \tag{28b}$$

with $\sigma_s^2 = \frac{1}{K} tr(cov_x)$.

Comment

As an alternative, we might use eq:s (18a,b) to estimate the amplitudes and the signal, although the noise is auto-correlated. The estimated amplitudes will be consistent because the estimated covariance matrix for the observations, under regularity conditions, i.e., stationary noise, tends to the theoretical. For the estimated signal we find

$$D^2(\hat{x}) = \mathrm{cov}_x + \sum_j^M \gamma_j^2 \, Q_{jj}^{-1} \tag{29a}$$

implying that the estimated signal is inconsistent. Analogous to (28b) we get

$$(S/N)^2 = \frac{\sigma_s^2}{\frac{1}{K}\mathrm{tr}(\sum_j^M \gamma_j^2 \, Q_{jj}^{-1})} \tag{29b}$$

If the noise levels are equal, which can be obtained after normalization of the records, (29b) is identical to (23). Further, if the autocovariance matrices for the noise are equal (28b) will be equal to (23) or (29a).

3.3 Known Basis for the Signal and No Auto-correlation

After suitable normalization, maximum likelihood leads to eq:s (16a,b). By choosing an orthogonal basis for the signal, with L<<K and A'A = K·I, (16a,b) changes to (17a,b), which can be written

$$\hat{\gamma}'\hat{\gamma}\,\hat{\phi} = A'Y\,\hat{\gamma} \tag{30a}$$

$$(A'\hat{\phi})'(A\hat{\phi})\,\hat{\gamma} = Y'A\,\hat{\phi} \tag{30b}$$

Using the normalizing condition on the amplitudes and combining these equations we find

$$(A\hat{\phi})'(A\hat{\phi})\,\hat{\gamma} = (A'Y)'(A'Y)\,\hat{\gamma} \tag{31}$$

i.e., $\hat{\gamma}$ is an eigenvector of the symmetric operator (A'Y)'(A'Y) with the corresponding eigenvalue $(A\hat{\phi})'\,A\hat{\phi} = K\,\hat{\phi}'\hat{\phi}$. It is easily seen that the likelihood is maximized when $\hat{\gamma}$ is the eigenvector corresponding to the largest eigenvalue. The coefficient vector is obtained from (30a) and the signal from

$$\hat{x} = A\hat{\phi} \tag{32}$$

The estimates of γ and ϕ are consistent because $\frac{1}{K} A'Y$ tends to $\phi\ \gamma'$ when K increases.

Turning to the variances for the estimated amplitudes and coefficients, we may proceed as follows:

Let $\theta_0 = \begin{bmatrix} \phi_0 \\ \gamma_0 \end{bmatrix}$ and $\hat{\theta} = \begin{bmatrix} \hat{\phi} \\ \hat{\gamma} \end{bmatrix}$ denote the true and estimated parameters. Using Taylor's theorem we get:

$$\left(\frac{\partial\ \Omega}{\partial\ \theta}\right)_{\hat{\theta}} = \left(\frac{\partial\ \Omega}{\partial\ \theta}\right)_{\theta_0} + \left(\frac{\partial^2\ \Omega}{\partial\ \theta^2}\right)_{\theta *} (\theta_0 - \hat{\theta}) \tag{33}$$

where $\theta *$ is between θ_0 and $\hat{\theta}$.

Now $\left(\frac{\partial\ \Omega}{\partial\ \theta}\right)_{\hat{\theta}}$ is always zero in the solution, thus

$$\left(\frac{\partial\ \Omega}{\partial\ \theta}\right)_{\theta_0} = -\left(\frac{\partial^2\ \Omega}{\partial\ \theta^2}\right)_{\theta *} (\theta_0 - \hat{\theta}) \tag{34}$$

From the general properties of maximum likelihood estimation $\left(\frac{\partial\ \Omega}{\partial\ \theta}\right)_{\theta_0}$ will tend to a normal distribution with zero mean and covariance matrix equal to

$$E\left(\frac{\partial^2\ \Omega}{\partial\ \theta^2}\right)_{\theta_0} \tag{35}$$

i.e., equal to the expectation of the second order derivative evaluated at $\theta = \theta_0$.

Because of the normalizing conditions on the amplitudes ($\gamma'\gamma = 1$), the rank of (35) will be $(L + M - 1)$. Thus, (35) is singular.

Define

$$B = \begin{pmatrix} I_L & 0 \\ 0 & \Delta^* \end{pmatrix} \quad \text{with} \tag{36a}$$

$$\Delta^* = \begin{pmatrix} 1 & 0 & 0 & . & . & 0 \\ 0 & 1 & 0 & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ \frac{-\gamma_1}{\gamma_M} & \frac{-\gamma_2}{\gamma_M} & . & . & \frac{-\gamma_{M-1}}{\gamma_M} & 1 \end{pmatrix} \quad \text{and } I_L \text{ an L-order identity matrix.} \tag{36b}$$

Then,

$$B' \left(\frac{\partial\ \Omega}{\partial\ \theta}\right)_{\hat{\theta}} = -B'\left(\frac{\partial^2\ \Omega}{\partial\ \theta^2}\right)_{\theta^*} B\ B^{-1}\ (\theta_0 - \hat{\theta}) \tag{37}$$

or

$$Z_0 = -C\Delta_0 \qquad \text{with } \Delta_0 = B^{-1}(\theta_0 - \hat{\theta}) \tag{38}$$

Neglecting terms of second order and higher, we find that the last element of Δ_0 is zero and we may eliminate the last row and column of C. Let C_1 denote the upper left part of C, Z_{01} and Δ_{01} the (M-1) first elements of Z_0 and Δ_0 respectively.

Then

$$Z_{01} = -C_1 \Delta_{01} \tag{39}$$

The limiting distribution of Z_{01} is normal with covariance matrix C_1 (where the derivatives are taken at $\theta^* = \theta_0$).

Further,

$$\Delta_{01} = -C_1^{-1} Z_{01} \tag{40}$$

and we find that the limiting distribution of Δ_{01} is normal with covariance matrix C_1^{-1}.

Define

$$B^* = \begin{pmatrix} I_L & 0 \\ 0 & \Delta^{**} \end{pmatrix} \quad \text{where} \tag{41a}$$

$$\Delta^{**} = \begin{pmatrix} 1 & 0 & 0 & . & 0 \\ 0 & 1 & 0 & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ \frac{-\gamma_1}{\gamma_M} & . & \frac{-\gamma_2}{\gamma_M} & . & \frac{-\gamma_{M-1}}{\gamma_M} \end{pmatrix} = (M \times M-1) \tag{41b}$$

Again neglecting terms of higher order, we obtain

$$B^* \Delta_{01} = (\theta_0 - \hat{\theta}) \tag{42}$$

which implies that the limiting distribution of $\theta_0 - \hat{\theta}$ is normal with covariance matrix

$$\Psi = B^* C_1^{-1} B^{*\prime} \tag{43}$$

To obtain the covariance matrix for the estimated parameters, we thus need the second order derivatives of Ω. These derivatives are:

$$\frac{\partial^2 \Omega}{\partial \phi^2} = K I_L \frac{1}{\sigma^2} \tag{44a}$$

$$\frac{\partial^2 \Omega}{\partial\phi\partial\gamma} = 2K \phi \gamma' \frac{1}{\sigma^2} - A'Y\frac{1}{\sigma^2} \tag{44b}$$

$$\frac{\partial^2 \Omega}{\partial \gamma^2} = (A\phi') A\phi I_M = K \sigma_s^2 I_M \frac{1}{\sigma^2} \tag{44c}$$

We then easily find

$$E\left(\frac{\partial^2 \Omega}{\partial \theta^2}\right)_{\theta_0} = \frac{K}{\sigma^2} \begin{bmatrix} I_L & \phi\gamma' \\ \gamma\phi' & \sigma_s^2 I_M \end{bmatrix} \tag{45}$$

and

$$C = \frac{K}{\sigma^2} B' \begin{bmatrix} I_L & \phi\gamma' \\ \gamma\phi' & \sigma_s^2 I_M \end{bmatrix} B = \frac{K}{\sigma^2} \begin{bmatrix} I_L & C_{12} \\ C_{21} & \sigma_s^2 \Delta^{*\prime}\Delta^* \end{bmatrix} \tag{46a}$$

with

$$C_{12} = \begin{bmatrix} 0 & 0 & \cdot & \cdot & \gamma_M & \cdot & \phi_1 \\ 0 & \cdot & \cdot & \cdot & \gamma_M & \cdot & \phi_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & \cdot & \cdot & \cdot & \gamma_M & \cdot & \phi_L \end{bmatrix} = C_{21}' \tag{46b}$$

After somewhat tedious calculations we obtain

$$C_1^{-1} = \frac{K}{\sigma^2} \begin{bmatrix} I_L & 0 \\ 0 & \frac{1}{\sigma_s^2} C_1^{22} \end{bmatrix} \tag{47a}$$

where

$$C_1^{22} = \begin{bmatrix} 1-\gamma_1^2 & -\gamma_1\gamma_2 & \cdot & -\gamma_1\gamma_{M-1} \\ -\gamma_1\gamma_2 & 1-\gamma_2^2 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ -\gamma_1\gamma_{M-1} & \cdot & \cdot & 1-\gamma_{M-1}^2 \end{bmatrix} \tag{47b}$$

And finally we obtain the covariance matrix for the estimates

$$\Psi = B^* C_1^{-1} B^{*\prime} = \frac{\sigma^2}{K} \begin{bmatrix} I_L & 0 \\ 0 & \frac{1}{\sigma_s^2} \Delta^{**} C_1^{22} \Delta^{**\prime} \end{bmatrix} \tag{48a}$$

$$\Delta^{**}C_1^{22}\Delta^{**'} = \begin{bmatrix} 1-\gamma_1^2 & -\gamma_1\gamma_2 & \cdot & -\gamma_1\gamma_M \\ -\gamma_1\gamma_2 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ -\gamma_1\gamma_M & \cdot & \cdot & 1-\gamma_M^2 \end{bmatrix} \tag{48b}$$

Comments

The variances for the amplitudes are smaller in this situation compared to situation 3.1. If the basis is known, the variance for the amplitudes is

$$D^2(\gamma_j) = \frac{\sigma^2(1 - \gamma_j^2)}{K \cdot \sigma_s^2} \tag{49}$$

which is smaller than the variance under 3.1. This is also reflected by the fact that one cannot rely on the general properties of maximum likelihood estimation when deriving the variance under 3.1.

Turning to the signal-to-noise ratio, we find that the estimated signal can be written

$$\hat{x} = x + (\hat{x} - x) \tag{50}$$

where $(\hat{x} - x)$ is the noise in the estimated signal. The variance for the true signal is

$$\frac{1}{K}\, x'x = \frac{1}{K}\, (A\phi)'(A\phi) = \phi'\phi = \sigma_s^2 \tag{51a}$$

and the variance for $(\hat{x} - x)$ is

$$\frac{1}{K} E\,(\hat{x} - x)'(\hat{x} - x) = \frac{1}{K} E\left[(A(\hat{\phi} - \phi))'A(\hat{\phi} - \phi)\right] =$$

$$\frac{1}{K} E\,(\hat{\phi} - \phi)'A'A(\hat{\phi} - \phi) = E(\hat{\phi} - \phi)'(\hat{\phi} - \phi) = \tag{51b}$$

$$\frac{L\sigma^2}{K}$$

Thus,

$$(S/N)^2 = \frac{K \cdot \sigma_s^2}{L \cdot \sigma^2} \tag{52}$$

and we see that the signal-to-noise ratio tends to infinity with $\sqrt{K}$.

3.4 Known Basis for the Signal and Autocorrelated Noise

As under 3.3, the estimates of γ and ϕ from (15a,b) will, under regularity condition, be consistent. The covariance matrix is rather complicated but can be obtained in a way analogous to that used in 3.3.

After some calculations we find the elements of

$$E\left(\frac{\partial^2 \Omega}{\partial \theta}\right)_{\theta_0}$$

to be

$$E\left(\frac{\partial^2 \Omega}{\partial \theta}\right)_{\theta_0} = A'(\sum_i^M \gamma_i^2 Q_{ii})\, A \tag{53a}$$

$$E\left(\frac{\partial^2 \Omega}{\partial \gamma^2}\right)_{\theta_0} = \sum_i^K \sum_j^K x_i x_j Q^*_{ij} \quad \text{with } x = A\phi \tag{53b}$$

$$E\left(\frac{\partial^2 \Omega}{\partial\phi\partial\gamma_i}\right)_{\theta_0} = \gamma_i\, A'Q_{ii}\, A\phi \quad \text{for } i = 1,2...M \tag{53c}$$

The covariance matrix for $\hat{\theta} = \begin{bmatrix} \hat{\phi} \\ \hat{\gamma} \end{bmatrix}$ is then

$$\Psi = B^* C_1^{-1} B^{*\prime} \tag{54}$$

with C_1 and B* defined as in 3.3.

Comments

a) Contrary to 3.3, there is no simple form of the covariance matrix (54). However, it can easily be calculated by computer.

b) Analogous to 3.3, we find for the signal-to-noise ratio (with $A'A = K\cdot I$)

1) The variance for the true signal is $\phi'\phi = \sigma_s^2$,

2) The variance for the noise in the estimated signal

$$\frac{1}{K} E\,(\hat{x} - x)'(\hat{x} - x) = E\,(\hat{\phi} - \phi)'(\hat{\phi} - \phi) =$$

$$\sum_{i=1}^{L} \Psi_{ii}$$

Thus

$$(S/N)^2 = \frac{\sigma_s^2}{\sum_{i=1}^{L} \Psi_{ii}} \tag{55}$$

As under 3.3 the signal-to-noise ratio will tend to inifinity with $\sqrt{K}$.

3) If we use eq:s (30a,b) to estimate ϕ and γ in the autocorrelated case, the estimates will still be consistent because

$$\frac{1}{K} A'Y \quad \text{tends to} \quad \phi\ \gamma'$$

This implies that the signal will be estimated consistently and that the signal-to-noise ratio will tend to infinity with $\sqrt{K}$.

4. Illustration

We shall illustrate the potential usefulness of the methods given above with two examples. The first was an artificial "signal" added to real seismic noise while the other represents two real seismic events recorded at NORSAR.

4.1 Artificial Data

We have used seismic noise from nine Scandinavian stations. The noise was normalized to have equal variances. Since the stations were widely separted the noise showed very little cross-correlation. The signal

$$S_t = \sum_{i=10}^{14} (\cos it + \sin it) \tag{56}$$

was added to each noise trace. The amplitude for this signal was set to 1 for four stations and to 1.5 for the other five. Two tests were made; in the first the signal variance was 0.25 resp. 0.563 of the noise variance. For each test data was analyzed in two ways. First ordinary component analysis was applied (eq:s 18a,b). In the

second try we utilized the knowledge that the signal could be expanded in linear combination of sin 10t,..., sin 14t and cos 10t,..., i.e., that the signal was band-limited. We did not use the knowledge of the noise auto-correlation but used formula (17a,b). The results are shown in Figs. 2 and 3. Fig. 1 shows the signal.

Figure 1. The original signal.

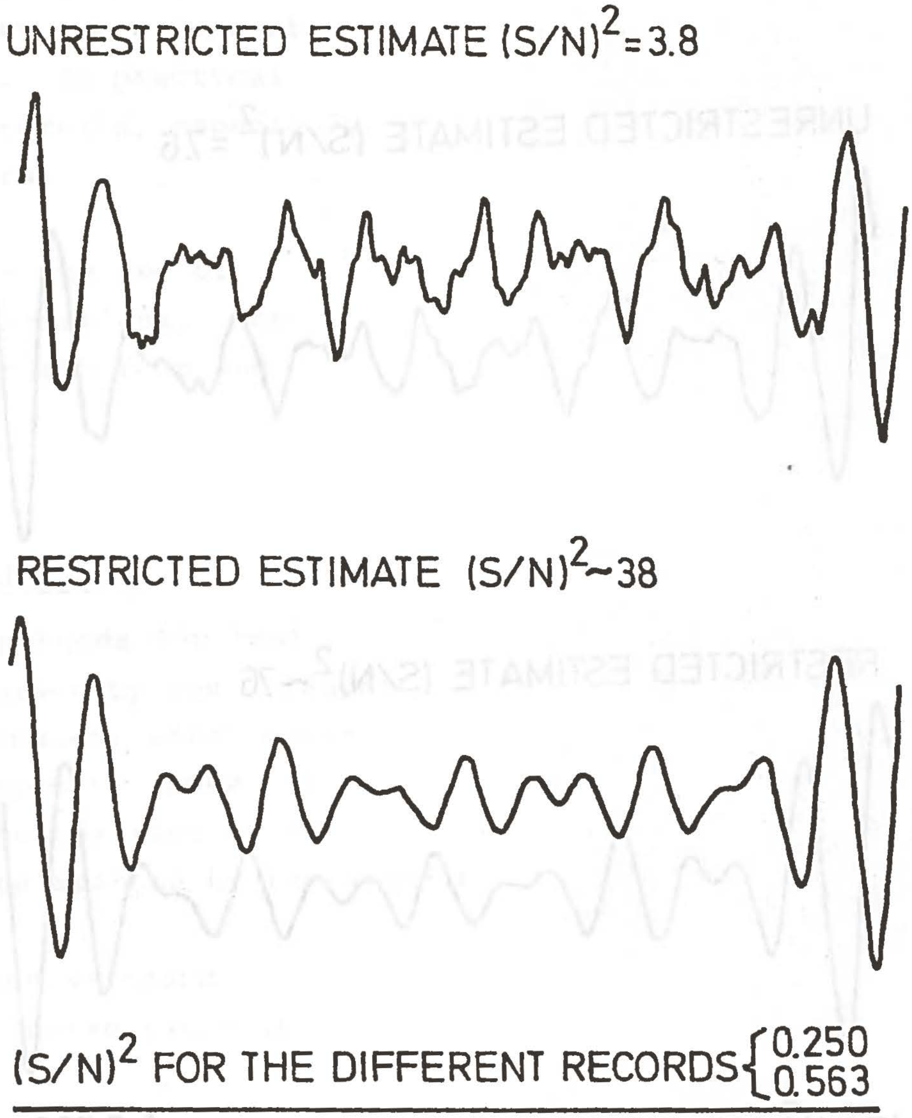

Figure 2. Estimated signals.

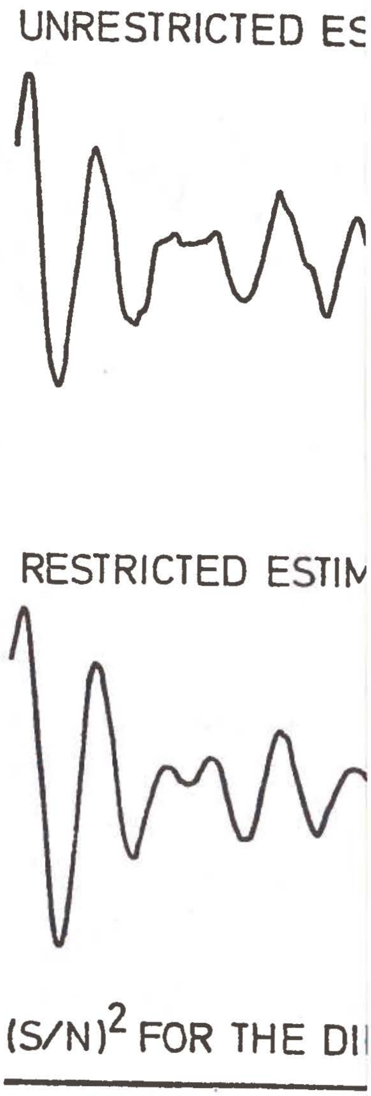

Figure 3. Estimated signals.

In the analysis we have compared fou

1) Sum and delay (Model
2) Sum and delay corrected (Model for variation in noise levels
3) Component model (Model
4) Component model corrected for variation in noise levels. (Model

The correction for variation in nois first estimating root mean square at each subarray by using 20 sec of signal. For Model 2B the correction each subarray beam with the square o RMS while for Model 3B the correctio the subarray beams with the correspo

The first event is a North Atlantic relatively strong and coherent array The gain for models 3A-B is as expec signals) rather small, 1.34 and 1.52 In the figure the left column gives codes, plot scaling factors and rela dsec. The right column gives model for the subarrays, SNR for the stand 2A), and gain in dB for model 2B, 3A the standard beam.

Figure 5 shows a very weak Western tively, the signals may represent a of a local explosion in the Baltic event, model 3A and 3B result in la 7.88 and 8.59 dB respectively.

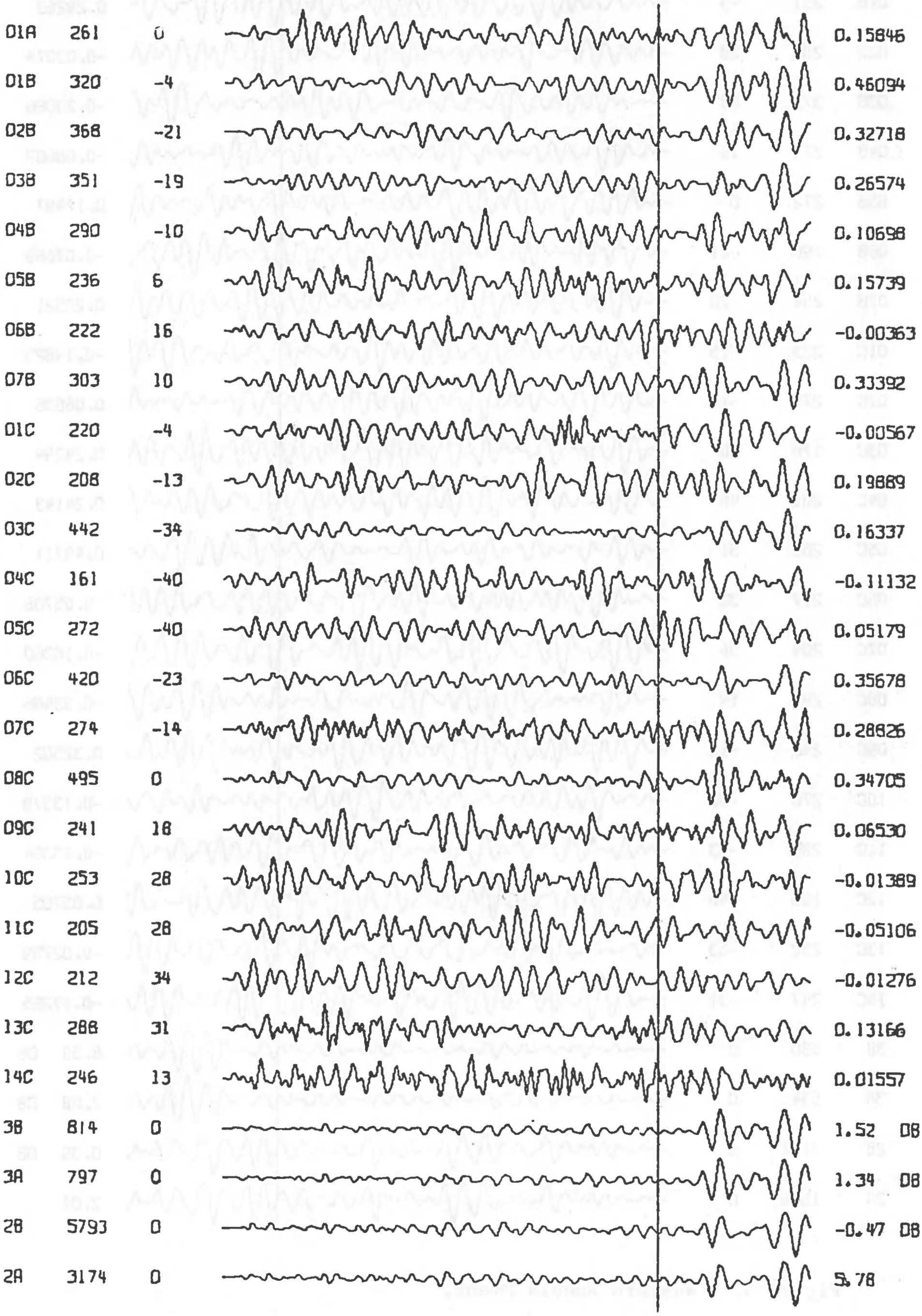

Figure 4. North Atlantic Earthquake.

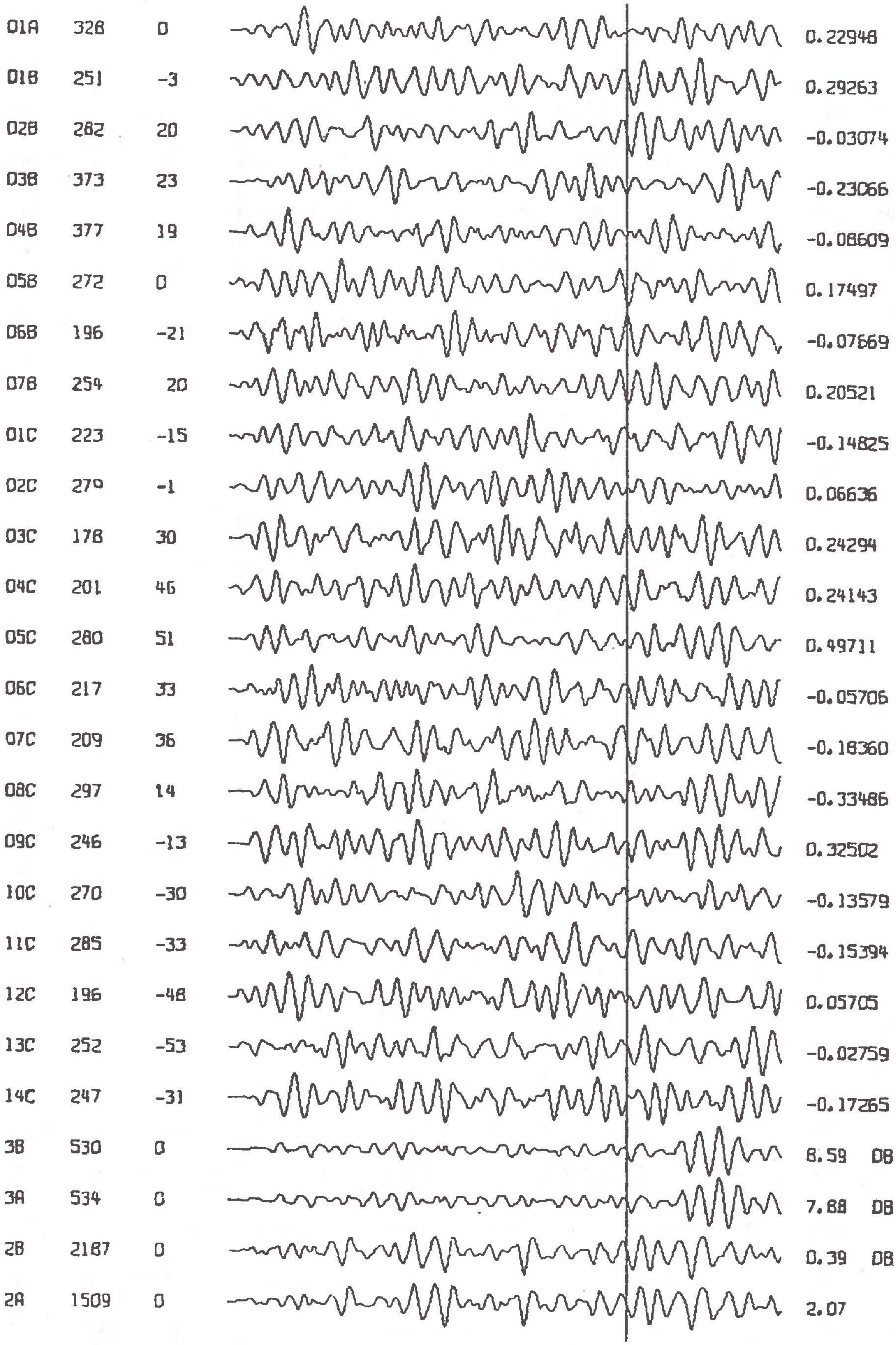

Figure 5. Western Russia Event.

5. The Model when 1<N<M

For this case the model, in terms of our observations, takes the form

$$Y = A\,\Phi\,\Gamma + N \tag{58}$$

where Y, A, N are defined in section 2, and

$$\Phi = \begin{pmatrix} \phi_{11}, \ldots, \phi_{1N} \\ \cdot \\ \phi_{L1}, \ldots, \phi_{LN} \end{pmatrix}$$

The columns of Φ contain these coefficients for the expansion of the basic vectors in terms of the given functions.

$$\Gamma = \begin{pmatrix} \alpha_{11}, \ldots, \alpha_{1M} \\ \cdot \\ \alpha_{N1}, \ldots, \alpha_{NM} \end{pmatrix}$$

The columns of Γ contain the amplitude factors for the basic vectors.

If the basis A is unknown, the model becomes

$$Y = X\Gamma + N \tag{59}$$

where

$$X = \begin{pmatrix} x_{11}, \ldots, x_{1N} \\ \cdot \\ x_{K1}, \ldots, x_{KN} \end{pmatrix}$$

The columns of X are the basic vectors in the space spanned by the signals. The signals are then $A\Phi\Gamma$ (or $X\,\Gamma$).

The estimation of the unknown in the model is carried out in a way similar to that used for the model with N=1. A way of extracting the solution that works also in the general case is to estimate the first basic vectors and the corresponding amplitude factors described in Section2. We then subtract the basic vector multiplied by the appropriate amplitude factors from each of the traces. The second basic vector and corresponding amplitude factor can now be obtained from the residual by repeating the whole process until we have extracted the described number of basic vectors. If the auto and cross correlation are disregarded, the rows of Γ are obtained as the eigenvectors corresponding to the N largest eigenvalues of (31). The elements of Φ are given by (30a).

6. Rotation of the Basis

We have so far just obtained a set of vectors that span our solution space. In most applications we would like to find a set of new vectors each of which is closely related to a signal form observed at one or more sensors. Depending on which estimation formula that has been used, our components are orthogonal or nearly orthogonal. It is not reasonable to assume, however, that the different signal forms should be orthogonal. What one wants to assume is that the output of one sensor should be identical to one vector only, except for a constant. Fortunately, psychologists have been concerned with very similar problems for half a century. For a detailed description, the reader is referred to (Harman 1967); here we will propose a method given by (Jennrich and Sampson 1966).

Define

$$F(B) = \sum_{p}^{N}\sum_{q} \left(\sum_{j}^{M} b_{jp}^2 b_{jq}^2 - \frac{\delta}{n}\sum_{j}^{M} b_{jp}^2 \sum_{j}^{M} b_{jq}^2\right) \qquad (60)$$

where $BT = \Gamma$ and $\delta<0.8$ is a given constant. Our problem is now to find a transformation T that will minimize F(B) under the constraint that diag (T'T) = I. The rational for setting up this formula is that we want the cross-correlation of the squared elements of a pair of amplitude factors to be small. Our constraint is introduced to make the contribution to the variance from each instrument the same as it was before the rotation.

<u>An example</u>: Assume that we have got the regression coefficients presented in the two leftmost columns of Table 1. The two basic vectors seem to be present of every instrument.

Table 1. Hypotetical data (Partly from Harman 1967)

Amplitude factors for eight instruments on two components			
Original Solution		Rotated Solution	
.830	-.396	.883	.065
.818	-.469	.956	-.029
.777	-.470	.926	-.045
.798	-.401	.882	.035
.786	.500	.005	.940
.672	.458	-.006	.803
.594	.444	-.065	.793
.647	.333	.104	.646

If we apply the Jennrich-Sampson method with $\delta=0$, we obtain the two rightmost columns. It is now evident that the signal form represented by the first basic vector is present only at instruments 1-4, while the second basic vector is present at instruments 5-8. This would normally be caused by two different structures. We should also observe that the correlation for the basic vectors has raised to 0.471 from almost zero in the unrotated case.

REFERENCES

Anderson, T.W.: Asymptotic Theory for Principal Component Analysis, Ann. Math. Stat. 34, 1963.

Broome, P., and W.C. Dean: Seismological Applications of Orthogonal Function Expansions, Final Report, Contract AF 49 (638)-1117 UEV, Alexandria, Virginia, 1964.

Capon, J., R.J. Greenfield, and R.T. Lacoss: Off-line Processing Results for the Large Aperture Seismic Array, M.I.T. Lincoln Laboratory Technical Note 1966-37, 1966.

Christoffersson, A., and B. Jansson: Maximum-Likelihood Estimation of an Unknown Signal in Multiple Gaussian Noise with Known Covariance Matrix, Seminar report, University Institute of Statistics, Uppsala, 1966.

Harman, H.H.: Modern factor analysis, University of Chicago Press, 1967.

Jansson, B., and E.S. Husebye: Application of Array Data Processing Techniques to a Network of Ordinary Seismograph Stations, PAGEOP, Vol 69, 1968.

Jennich, R.I., and P.F. Sampson: Rotation for simple loadings, Psychometrika, 31, 1966.

Kelly, E.J., and M.J. Levin: Signal parameter estimation for seismometer arrays, Technical Report No. 339, M.I.T. Lincoln Laboratory, 1964.

Rao, C.R.: The use and interpretation of principal component analysis in applied research, Sankyā 26, 1969.

Whitcomb, J.H.: Array data processing techniques applied to long-period shear waves at Fennoscandian seismograph stations, 1969.

Wold, H.: Nonlinear estimation by iterative least squares procedures, Festschrift Jerzy Neyman, John Wiley & Sons, 1966 (a).

Wold, H.: On estimation of principal components and related models by iterative least-squares. Paper in multivariate analysis., Ed. Krishnaiah, Academic Press, 1966 (b).