*Currently on leave from the University of New South Wales

ESTIMATION IN ONE DIMENSIONAL BILATERAL PROCESSES

by

C. A. McGILCHRIST*

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 930
June, 1974

Estimation in One Dimensional Bilateral Processes

C. A. McGilchrist
University of North Carolina*

SUMMARY

A method of fitting a bilateral, autoregressive, moving average process is described with emphasis placed on the case in which only short sequences of observations are available. Such a process is fitted to individual plant yields in a monoculture plot.

Keywords: Bilateral, autoregression, moving average, plant competition.

1. INTRODUCTION A sequence of observations $X_1, X_2, \ldots$ forms a one-dimensional bilateral process if the observation at position t depends on observations at either side of t. A model of the form,

$$X_t + \phi_1 X_{t-1} + \ldots + \phi_p X_{t-p} + \phi_{-1} X_{t+1} + \ldots + \phi_{-r} X_{t+r} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} + \theta_{-1}\varepsilon_{t+1} + \ldots + \theta_{-s}\varepsilon_{t+s} ,$$

may be described as a bilateral, autoregressive, moving average process of order p, q, r, s . In the model the ε_t are independent $N(0, \sigma^2)$ variates and ϕ_i, θ_j, σ^2 are unknown parameters. The model is analagous to those of Box and Jenkins (1970) and using a similar notation to these authors we call the model the (p, r, o, q, s) model. Letting B be the backward shift operator $BX_t = X_{t-1}$, the model may be written

---

* on leave from the University of New South Wales

$$(1 + \phi_1 B + \dots + \phi_p B^p + \phi_{-1} B^{-1} + \dots + \phi_{-r} B^{-r}) X_t =$$

$$= (1 + \theta_1 B + \dots + \theta_q B^q + \theta_{-1} B^{-1} + \dots + \theta_{-s} B^{-s}) \varepsilon_t \ ,$$

or more simply $$\phi(B) X_t = \phi(B) \varepsilon_t \ .$$

Usually the above model is not sufficient to describe experimental data because all means are zero. It is enough for our purposes to add to the model a deterministic trend $p_{d-1}(t)$, which is a polynomial in t of order $d - 1$. Thus

$$X_t = p_{d-1}(t) + \phi^{-1}(B) \theta(B) \varepsilon_t$$

and this model is referred to as the (p, r, d, q, s) model. In many applications, including the one discussed in this paper, d equals 1 indicating a process with unknown mean.

The problem which motivates this study is that of fitting a bilateral process to the dry weight of individual plants growing in monoculture plots. In a particular plant competition study, monoculture plots of each species contained 20 plants; two rows each containing 10 plants. For purposes of estimating appropriate variance components (which turn up in other analyses that need not concern us here), it is enough to consider the average ℓn (dry plant weight) for each column. The basic data we handle are such average yields X_t , $t = 1, 2, \dots , 10$, and each plot contributes such a series.

Bilateral models for one and two dimensional processes have previously been considered by Whittle (1954) and Matern(1960). Whittle recommended fitting the one dimensional, bilateral, autoregressive process by fitting a unilateral process with the same spectral density function. The parameters of the unilateral process derived in this way from a stationary bilateral process are such that the unilateral process is nonstationary and attempts at fitting such a model using procedures in McGilchrist (1974), which are specifically suited to small samples, produced nonsense results.

In the following sections we advance a direct method of fitting the bilateral process and, like the method of McGilchrist (1974), the procedure is intended particularly for small samples. Since the general model is as easily handled as any simple model we discuss the general case.

2. REMOVAL OF TREND In a short time series the order of a polynomial trend can be obtained by simply graphing the series. The usual Box-Jenkins procedure for elimation of trend is used, viz. the difference operator $(1 - B)^d$ acting on $p_{d-1}(t)$ removes it so that

$$Y_t = (1 - B)^d X_t = (1 - B)^d p_{d-1}(t) + (1 - B)^d \phi^{-1}(B)\theta(B)\varepsilon_t$$
$$= \phi^{-1}(B)(1 - B)^d \theta(B)\varepsilon_t \ .$$

Thus $\phi(B)Y_t = \theta^*(B)\varepsilon_t$, where $(1 - B)^d\theta(B) = \theta^*(B)$. The coefficients θ_j^* of $\theta^*(B)$ are related to the coefficients of $\theta(B)$.

We assume $T + d$ observations in each time series so that after differencing there are T observations Y_t which follow an autoregressive, moving average model with zero mean. The coefficients to be estimated are still ϕ_i, θ_j and σ^2 .

3. ESTIMATION The method of maximising the liklihood conditional on fixed border elements is used. The border elements are the first p values of Y_t

and the last r values. Let

$$E_t = \theta^*(B)\mathcal{E}_t \ , \ t = p+1, p+2, \ldots , T-r$$

then the border elements and the E_t determine the Y_t according to the following equation,

$$\begin{bmatrix} 1 & \phi_{-1} & \cdots & \phi_{-r} & & 0 \\ \phi_1 & 1 & & & & \\ \circ & & & & & \\ \phi_p & & & & & \\ & & & & & \phi_{-r} \\ & & & & & \circ \\ & & & & 1 & \phi_{-1} \\ 0 & & \phi_p & \cdots & \phi_1 & 1 \end{bmatrix} \begin{bmatrix} Y_{p+1} \\ Y_{p+2} \\ \circ \\ \circ \\ \circ \\ \circ \\ \circ \\ \circ \\ Y_{T-r} \end{bmatrix} = \begin{bmatrix} E_{p+1} - \phi_1 Y_1 - \cdots - \phi_p Y_p \\ \bullet \\ E_{2p} - \phi_1 Y_{2p-1} \\ E_{2p+1} \\ \circ \\ E_{T-2r} \\ E_{T-2r+1} - \phi_{-1} Y_{T-2r+2} \\ E_{T-r} - \phi_{-1} Y_{T-r+1} \cdots - \phi_{-r} Y_T \end{bmatrix} \quad (3.1)$$

We introduce the following notation. Let $\underset{\sim}{A}_n$ be an $n \times n$ matrix of the type given on the left-hand side of (3.1). Let

$$\underset{\sim}{e}'_n = [E_{p+1}, E_{p+2}, \ldots , E_{p+n}] \ , \ n = 1, 2, \ldots , T-r-p$$

$$\sigma^2 \underset{\sim}{\Sigma}_n = \text{Var } \underset{\sim}{e}_n \ .$$

Now since

$$\text{Cov}(E_t, E_{t+\tau}) = \begin{cases} \sigma^2 \sum_{j=-s}^{q+d-\tau} \theta_j^* \theta_{j+|\tau|}^* \ , & |\tau| \le q+d+s \ , \\ 0 \ , & |\tau| > q+d+s \ , \end{cases}$$

$$= \begin{cases} \sigma_{|\tau|} \ , & |\tau| \le q+d+s \ , \\ 0 \ , & |\tau| > q+d+s \ . \end{cases}$$

then for $s_{i+1} = \sigma_i/\sigma^2$ and $k = q + d + s$

$$\underset{\sim}{\Sigma}_n = \begin{bmatrix} s_1 & s_2 & \cdots & s_k & & 0 \\ s_2 & s_1 & \ddots & & \ddots & \\ \cdot & \ddots & \ddots & \ddots & & s_k \\ s_k & & \ddots & \ddots & \ddots & \cdot \\ & \ddots & & \ddots & s_1 & s_2 \\ 0 & & s_k & \cdots & s_2 & s_1 \end{bmatrix}$$

Let

$$\underset{\sim}{y}'_n = [Y_{p+1}, Y_{p+2}, \ldots , Y_{p+n}]$$

$$Z_t = \phi(B)Y_t$$

$$\underset{\sim}{z}'_n = [Z_{p+1}, Z_{p+2}, \ldots , Z_{p+n}]$$

If we delete the E_t terms from the vector on the right-hand side of (3.1) we obtain a $T - p - r$ dimensional vector in which the first p and last r terms are linear combinations of the border elements and the remaining terms are zero. Let this vector be denoted by $\underset{\sim}{\mu}$. Thus (3.1) becomes

$$\underset{\sim}{A}_{T-p-r}\, \underset{\sim}{y}_{T-p-r} = \underset{\sim}{e}_{T-p-r} + \underset{\sim}{\mu}$$

and the liklihood is the probability density function of $\underset{\sim}{y}_{T-p-r}$ with border elements held fixed. Since

$$E(\underset{\sim}{y}_{T-p-r}) = \underset{\sim}{A}^{-1}_{T-p-r}\, \underset{\sim}{\mu} \;,\; \mathrm{Var}(\underset{\sim}{y}_{T-p-r}) = \sigma^2\, \underset{\sim}{A}^{-1}_{T-p-r}\, \underset{\sim}{\Sigma}_{T-p-r} (\underset{\sim}{A}'_{T-p-r})^{-1}$$

the liklihood function is

$$L = (2\pi\sigma^2)^{-\frac{1}{2}(T-p-r)} |\underset{\sim}{A}^{-1}_{T-p-r}\, \underset{\sim}{\Sigma}_{T-p-r} (\underset{\sim}{A}'_{T-p-r})^{-1}|^{-\frac{1}{2}} \exp(-Q^2_{T-p-r}/2\sigma^2) \;,$$

where

$$Q^2_{T-p-r} = (\underset{\sim}{y}_{T-p-r} - \underset{\sim}{A}^{-1}_{T-p-r}\,\underset{\sim}{\mu})'\,\underset{\sim}{A}^{*}_{T-p-r}\,\underset{\sim}{\Sigma}^{-1}_{T-p-r}\,\underset{\sim}{A}_{T-p-r}(\underset{\sim}{y}_{T-p-r} - \underset{\sim}{A}^{-1}_{T-p-r}\mu)$$

$$= (\underset{\sim}{A}_{T-p-r}\,\underset{\sim}{y}_{T-p-r} - \underset{\sim}{\mu})'\,\underset{\sim}{\Sigma}^{-1}_{T-p-r}(\underset{\sim}{A}_{T-p-r}\,\underset{\sim}{y}_{T-p-r} - \underset{\sim}{\mu})\ .$$

Now $\underset{\sim}{z}_{T-p-r} = \underset{\sim}{A}_{T-p-r}\,\underset{\sim}{y}_{T-p-r} - \underset{\sim}{\mu}$ so that $Q^2_{T-p-r} = \underset{\sim}{z}'_{T-p-r}\,\underset{\sim}{\Sigma}^{-1}_{T-p-r}\underset{\sim}{\dot{z}}_{T-p-r}$

and

$$-2\ell nL = (T-p-r)\ell n(2\pi\sigma^2) + \ell n|\underset{\sim}{\Sigma}_{T-p-r}| - \ell n|\underset{\sim}{A}_{T-p-r}|^2 + Q^2_{T-p-r}/\sigma^2\ .$$

On maximising with respect to σ^2 we find

$$\hat{\sigma}^2 = Q^2_{T-p-r}/(T-p-r)$$

and $-2\ell nL(\hat{\sigma}^2) = \text{constant} + (T-p-r)\ell nQ^2_{T-p-r} + \ell n|\underset{\sim}{\Sigma}_{T-p-r}| - \ell n|\underset{\sim}{A}_{T-p-r}|^2$.

In what follows we choose estimates by minimising

$$R_{T-p-r} = (T-p-r)\ell nQ^2_{T-p-r} + \ell n|\underset{\sim}{\Sigma}_{T-p-r}| - \ell n|\underset{\sim}{A}_{T-p-r}|^2$$

and we do so by finding an algorithm to evaluate R_{T-p-r} for any specific parameter values. The minimum is found by standard iterative procedures.

4. COMPUTATION OF LIKELIHOOD

The method of computing R_{T-p-r} does not involve approximations and is particularly suited to small samples. Some of the computing procedure is the same as that reported in McGilchrist (1974) but results are given here for completeness.

(i) Computation of Q^2_{T-p-r} : Let $\underset{\sim}{p}'_n = [0, \ldots , 0, s_k, \ldots , s_2]$ be an n dimensional vector and $\text{rev}(\underset{\sim}{p}_n)$ stand for the vector obtained by reversing the elements of $\underset{\sim}{p}_n$. First find $Y_t = (1 - B)^d X_t$, $Z_t = \phi(B)Y_t$ and

$\theta^*(B) = (1 - B)^d \theta(B)$. Then for $k = q + d + s$ find directly

$$Q_k^2 = \tilde{z}_k' \tilde{\Sigma}_k^{-1} \tilde{z}_k \ , \ \tilde{g}_k = \tilde{\Sigma}_k^{-1} \tilde{p}_k \ .$$

The following recurrence formulae give successively $\tilde{g}_n$, Q_n^2, $n = k + 1, \ldots ,$ $T - p - r$.

$$c_n^2 = (S_1 - \tilde{p}_n' \tilde{g}_n)^{-1} \ , \ b_n = \text{rev}(\tilde{p}_n')\tilde{g}_n$$

$$\tilde{g}_{n+1} = \begin{bmatrix} 0 \\ \tilde{g}_n \end{bmatrix} + c_n^2 b_n \begin{bmatrix} -1 \\ \text{rev}(\tilde{g}_n) \end{bmatrix}$$

$$Q_{n+1}^2 = Q_n^2 + c_n^2 (Z_{n+1} - \tilde{z}_n' \tilde{g}_n)^2$$

(ii) Computation of $\ell n|\tilde{\Sigma}_{T-p-r}|$: Firstly $|\tilde{\Sigma}_k|$ is calculated by the usual subroutines when inverting $\tilde{\Sigma}_k$. Further

$$\ell n|\tilde{\Sigma}_{n+1}| = \ell n|\tilde{\Sigma}_n| - \ell n(c_n^2)$$

(iii) Computation of $\ell n|\tilde{A}_{T-p-r}|^2$: Let $\tilde{a}_n$, $\tilde{d}_n$ be two n dimensional vectors,

$$\tilde{a}_n = [\phi_{-1}, \phi_{-2}, \ldots , \phi_{-r}, 0, \ldots , 0] \ , \ \tilde{d}_n = [\phi_1, \phi_2, \ldots , \phi_p, 0, \ldots , 0] \ ,$$

let $m = \max(r, p)$. First compute directly $\tilde{A}_m^{-1}$ and

$$\tilde{q}_m = \tilde{A}_m^{-1} \text{rev}(\tilde{a}_m) \ , \ \tilde{h}_m' = \text{rev}(\tilde{d}_m')\tilde{A}_m^{-1} \quad \text{and} \quad |\tilde{A}_m|^2$$

and then successively

$$K_n = [1 - \text{rev}(\tilde{d}_n')\tilde{q}_n]^{-1}$$

$$\underset{\sim}{q}_{n+1} = \begin{bmatrix} 0 \\ \underset{\sim}{q}_n \end{bmatrix} + K_n \underset{\sim}{a}_n' \underset{\sim}{q}_n \begin{bmatrix} -1 \\ \mathrm{rev}(\underset{\sim}{h}_n) \end{bmatrix} , \quad \underset{\sim}{h}_{n+1} = \begin{bmatrix} 0 \\ \underset{\sim}{h}_n \end{bmatrix} + K_n \underset{\sim}{d}_n' \underset{\sim}{h}_n \begin{bmatrix} -1 \\ \mathrm{rev}(\underset{\sim}{q}_n) \end{bmatrix}$$

$$\ell n|A_{n+1}|^2 = \ell n|A_n|^2 - \ell n(K_n^2)$$

The proof of these results for the unsymmetric band matrix $\underset{\sim}{A}_n$ is given in McGilchrist (1974).

5. WHITTLE'S METHOD To indicate what goes wrong with Whittle's method we consider only the simplest case,

$$(1 + \alpha B)(1 + \beta B^{-1})X_t = \varepsilon_t \ , \ t = 2, 3, \ldots , T - 1 \ , \qquad (5.1)$$

where α, β are unknown autoregressive parameters. Given border elements X_1, X_T and $\varepsilon_2, \varepsilon_3, \ldots , \varepsilon_{T-1}$ together with the autoregressive parameters, then X_t may be generated for $t = 2, 3, \ldots , T - 1$. However, the above equation (5.1) may be rearranged as

$$\beta[1 + (\alpha + \beta^{-1})B + (\alpha/\beta)B^2]B^{-1}X_t = \varepsilon_t \ ,$$

or equivalently

$$(1 + \alpha B)(1 + \beta^{-1}B)X_t = \varepsilon_{t-1}/\beta = \varepsilon_t^* \ , \ t = 2, 3, \ldots , T \ . \qquad (5.2)$$

Now equations (5.1) and (5.2) are equivalent and (5.2) describes a unilateral process of order 2 with parameters related to the parameters of the bilateral process. For stationarity of the bilateral process we require $|\alpha| < 1$, $|\beta| < 1$ and these conditions imply that (5.2) is nonstationary.

However, the two processes are equivalent only if the border elements X_1, X_2 of (5.2) are such as to generate the border elements X_1, X_T of

(5.1) and hence the distribution arising from (5.2) which is equivalent to the distribution arising from (5.1) is a complicated conditional distribution. This approach is not used here.

6. EXAMPLES To test the method a sample of 100 observations were generated for the (1, 1, 1, 0, 0) process with $\phi_{-1} = .7$, $\phi_1 = .2$. Contours of the liklihood surface with a minimum at (.6, .3) are shown in Figure 1. The two contours shown correspond to 90% and 99% confidence regions.

We now come to the problem described in the introduction. The data given in Table 1 are dry weight of vegetative part in grams of three varieties of oats each grown in monoculture. This data is provided by Dr. B. R. Trenbath, Australian National University. Each variety of oats has two monoculture plots and within each plot there are two rows of plants with 10 plants in each row. The same rectangular planting configuration and the same spacing is used in each plot. For convenience we call data from the six plots taken in order, data sets 1, 2, ... , 6 .

As described in the introduction we analyse ℓn (dry weight) and it is sufficient to deal with the average of these values over the columns of each plot resulting in six univariate time series, each of length 10. It is unlikely that we can fit an extensive model to such short series so we fit a (1, 1, 1, 0, 0) process to each of the six sets of data. The likelihood surfaces are evaluated and the minimum and 90% confidence contours are given for each data set in Figures 2(a) and (b). These confidence contours will subsequently be used to estimate within plot variability for the type of diallel analysis reported in McGilchrist and Trenbath (1971).

REFERENCES

Box, G. E. P. and G. M. Jenkins, Time series analysis; forecasting and control , Holden-Day, San Francisco, (1970).

Matern, B., "Spatial variation", Reports of the Forest Research Institute of Sweden, vol. 49, (1960).

McGilchrist, C. A., "Estimation in short time series", submitted to the *Journal of the Royal Statistical Society*, (1974).

McGilchrist, C. A., "Inverse and determinant of a band matrix", submitted to the *International Journal of Computer Mathematics*, (1974).

McGilchrist, C. A. and B. R. Trenbath, "A revised analysis of plant competition experiments", *Biometrics*, vol. 27, pp. 659-71, (1971).

Whittle, P., "On stationary processes in the plane", *Biometrika*, vol. 41, pp. 434-49, (1954).

TABLE 1.
Dry weight (g) of vegetative part of oat plants

Variety	Plot	Dry Weight									
1	1	9.16	4.56	7.23	4.94	3.03	2.26	6.48	7.25	10.70	7.25
		8.40	4.94	4.56	10.32	7.25	5.71	6.09	6.48	9.55	8.78
	2	9.93	10.79	11.92	9.36	6.23	8.51	6.23	6.52	7.08	4.81
		7.08	8.22	7.37	7.65	7.08	7.65	7.08	7.65	5.95	8.22
2	1	2.48	2.16	4.99	5.93	4.04	4.99	6.86	9.37	11.58	15.34
		3.42	11.58	9.99	14.09	6.56	15.62	17.86	5.30	8.12	3.10
	2	21.99	13.99	16.35	9.49	11.87	13.70	8.29	19.03	8.29	19.61
		15.77	7.10	14.84	11.58	13.97	8.29	13.09	11.58	12.19	8.29
3	1	4.55	6.85	12.84	7.75	8.70	7.75	5.01	11.90	19.32	4.55
		10.05	8.21	11.00	8.70	14.69	8.70	4.10	7.30	4.10	3.16
	2	3.93	5.23	18.96	10.54	9.64	9.64	11.41	20.29	14.98	18.09
		7.87	7.43	13.21	10.97	7.43	6.57	14.98	18.52	11.88	8.77

Figure 1

90% and 99% confidence contours of likelihood surface for 100 values generated by (1, 1, 1, 0, 0) process with $\phi_{-1} = .7$, $\phi_1 = .2$. Minimum is marked with a cross.

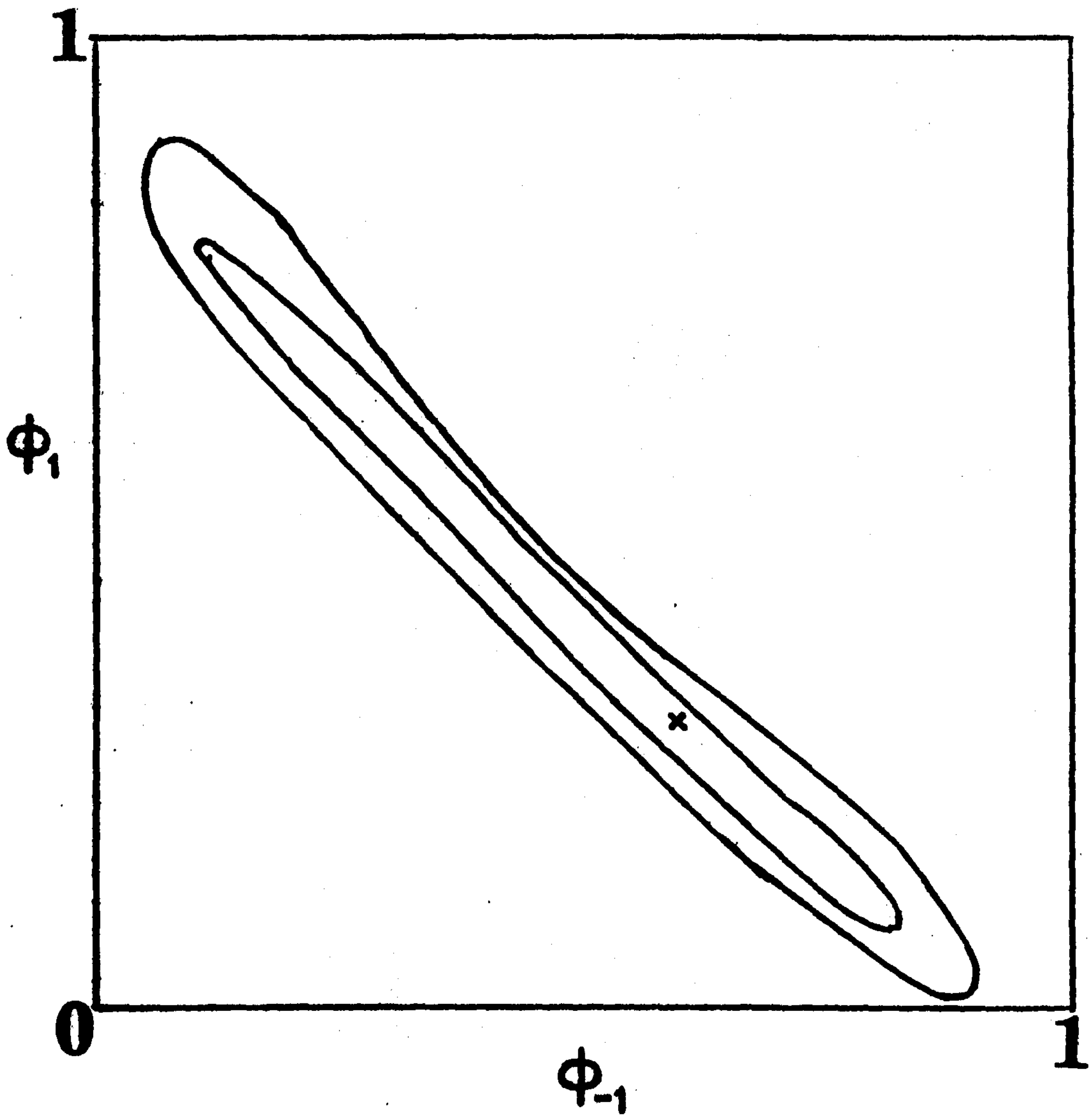

Figure 2(a)

Minima and 90% confidence contours of likelihood surfaces for data sets 1, 2, 3. Minimum is marked with the number of the data set and arrows point to corresponding confidence contours.

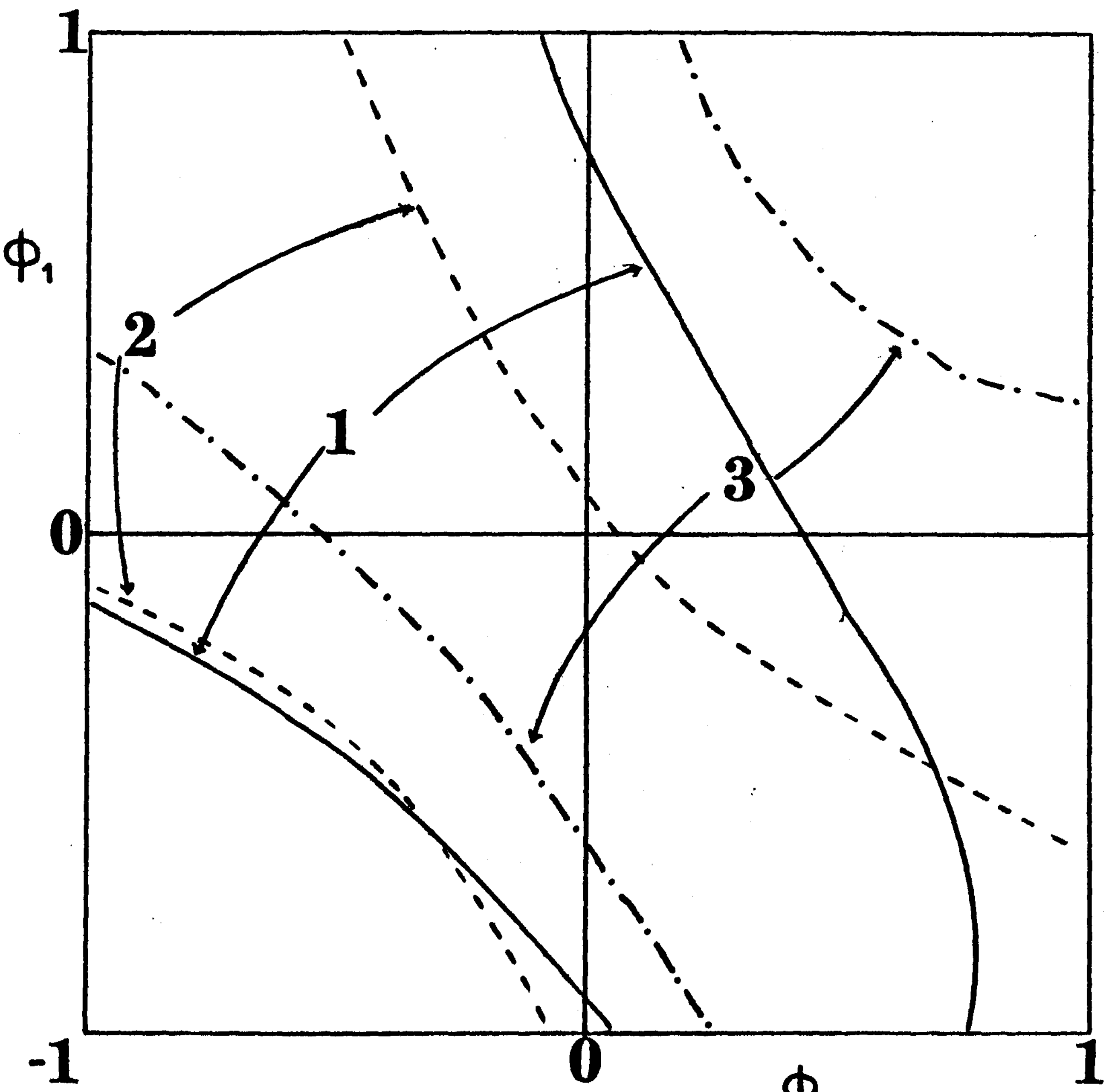

Figure 2(b)

Minima and 90% confidence contours of likelihood surfaces for data sets 4, 5, 6. Minimum is marked with the number of the data set and arrows point to corresponding confidence contours.