On Some Properties of ARMA(1,1) Model Fitting to AR(2) Processes

Minoru Tanaka

School of Network and Information, Senshu University,
2-2-2, Higasimita, Tama-ku, Kawasaki,
Kanagawa 214-8580, Japan

Abstract. This paper gives a discussion on a misspecified ARMA(1,1) model fitting to an AR(2) process. The problem concerning a number of globally and locally maximal points of the conditional likelihood function is investigated when the sample size tends to infinity. We shall detect the conditions of AR(2) parameters on which the ARMA(1,1) conditional likelihood function has more than one locally maximal points in the stationary and invertible parameter space.

Keywords: *ARMA(1,1) model fitting, AR(2) process, misspecification, locally minimal points, Catastrophe Theory.*

1. Introduction

This paper is a sequel to "On a moving average time series model fitting" contributed with Mr. Kenji Aoki in 1991 ([9]).

In time series analysis, we usually apply the suitable linear model for a given time series data to predict a future value using the model. When fitting a model to the data, the parameters of a model will be estimated, and then we assume some probability distribution and generally the maximum likelihood method is used. If a true model is fitted to a process, then its unknown parameters can be precisely estimated. But when a model is incorrect, the statistical properties of the estimators will be known very little. The problem concerning maximum likelihood estimation of misspecified models has been investigated by many authors. In particular, the asymptotic properties (consistency) of the estimators of the parameters of a misspecified ARMA model have been discussed ([7]).

Also it is known that when we fit an MA(1) model to some special time series data which is not followed by MA(1) process, the MA(1) parameter does not have an unique Gaussian quasi-maximum likelihood estimator. Tanaka and Huzii [10] have given the conditions of AR(2) parameters on which the MA(1) quasi-likelihood function has more than one local maximal points in the invertible parameter space (-1, 1). Furthermore, Tanaka and Aoki [9] gave the region for the AR(2) parameters on which the MA(1) quasi-likelihood function has more than one local maximal points in the parameter space. In this case, maximizing the likelihood function is equivalent to minimizing the following function S(x; a, b) when the data length is large (see [10]). Here x is an MA(1) parameter and a and b are AR(2) parameters.

$$S(x;\, a,\, b) = \frac{1+b-a\,(1-b)\,x-b\,(1+b)\,x^2}{(1-b)\left(1-a^2+2\,b+b^2\right)\left(1-x^2\right)\left(1+a\,x+b\,x^2\right)}. \qquad (1.1)$$

From Tanaka and Huzii [10], we have two minimal points of the function $S(x;a,b) = S(x)$, say. For example, in the case of an AR(2) process with a = -0.1, b = 0.8, the function $S(x)$ has a graph shown in the following figure.

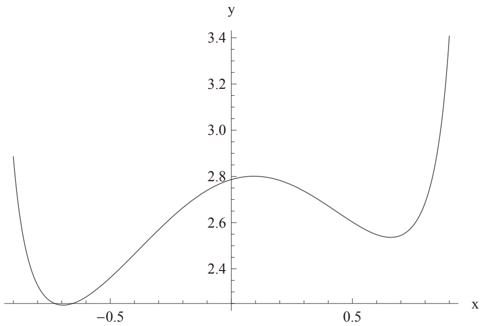


Figure 1. Graph of *S(x;a,b)* with a = -0.1, b = 0.8.

In order to have the conditions on which the function has two local minimal points in the parameter space, we should consider the differentiation *DS(x) = 0*. And we specified the case where the solution of the equation *DS(x) = 0* changed from three to two. That is, the value of the resultant ([5]) was able to formalize the contour line for zero (the bifurcation set). We set the domain D1 with a deep color surrounded with the curve of the shape of a wedge given in the upper part of Fig. 2. Its boundary is the bifurcation set. It will be locally seen a cusp.

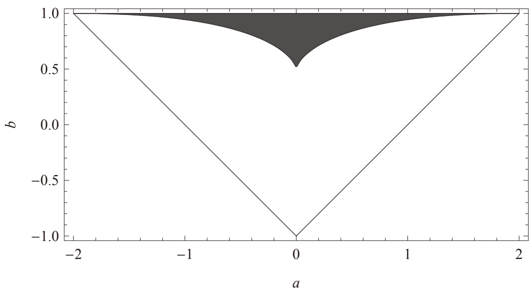


Figure 2. Bifurcation set and the domain D1 for *S(x;a,b)*.

The function S(x) has the two minimum points separated by a maximum with in D1, whereas outside it S(x) has a single minimum, which was given by Prof. Aoki using the concept of the cusp of Catastrophe theory with a potential S(x). It is also seen that the two minimum points are put together and S(x) has only one minimum point at the tip of the cusp (refer to information science research No.12 [9], and also [4] and [8] for details).

In this paper, we shall extend the model to the autoregressive moving average ARMA (1, 1) model and consider a problem similar to the misspecified MA(1) model fitting to AR(2) processes.

This paper is supported by the computer software *Mathematica V9.0* and its application *Time Series Pack for Mathematica* ([6]).

2. Results on misspecified ARMA(1,1) model fitting

Let {Z(t)} be a weakly stationary process with $EZ(t) = 0$. {Z(t)} is said to satisfy a autoregressive moving average model of order p and q (ARMA(p, q) model) if {Z(t)} is expressed as

$$(1 - a_1 B - \ldots - a_p B^p)\, Z(t) = (1 + b_1 B + \ldots + b_q B^q)\, e(t), \qquad (2.1)$$

where {e(t)}, t being an integer, consists of independently and identically distributed random variables with $E[e(t)] = 0$, $E\left[e(t)^2\right] = \sigma^2$, the a_p's and b_q's are constants which are independent of t, and B is the usual backshift operator such that $B[e(t)] = e(t\text{-}1)$ and $B^k[e(t)] = B\left[B^{k-1}[e(t)]\right]$ for k =1,2,.. (see, for example, [2], [3]).

Let

$$\phi(B) = 1 - a_1 B - \ldots - a_p B^p = \prod_{k=1}^{p} (1 - \phi_k B), \qquad (2.2)$$

$$\theta(B) = 1 + b_1 B + \ldots + b_q B^q = \prod_{k=1}^{q} (1 - \theta_k B). \qquad (2.3)$$

In our model fitting, it is assumed that $|\phi_h| < 1$, $|\theta_k| \leq 1$ for all h = 1, 2,$\cdots$, p, and k = 1, 2,$\cdots$, q. Let $\Theta = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta q)$ be a (p+q)-dimensional unknown parameter, and let $\{F_k(\Theta)\}$ be a sequence of functions of Θ, which are defined in the following way. For t > 0,

$$\mathrm{e(t)} = \{\prod_{k=1}^{p} (1 - \phi_k B) \prod_{k=1}^{q} (1 - \theta_k B)^{-1}\}\mathrm{Z(t)} = \left\{\sum_{k=1}^{\infty} F_k(\Theta) B^k\right\} Z(t). \quad (2.4)$$

For evaluating the asymptotic properties of the conditional quasi-maximum likelihood estimators of Θ when the sample size tends to infinity, we should attend to a function

$$S_{p,q}(\Theta) = E\left[e(t)^2\right] \qquad (2.5)$$

$$= \int_{-1/2}^{1/2} \frac{\left|\prod_{k=1}^{p}[1-\phi_k \exp(-2\pi i\omega)]\right|^2}{\left|\prod_{j=1}^{q}[1-\theta_j \exp(-2\pi i\omega)]\right|^2} f_Z(\omega)\, d\omega.$$

The value $\hat{\Theta}$ which minimizes $S_{p,q}(\Theta)$ with respect to Θ should be obtained (see Tanaka and Huzii [10] and also Huzii [5]).

The spectrum of an ARMA(p,q) process $f_Z(\omega)$ is given by

$$f_Z(\omega) = \frac{\sigma^2}{2\pi} \frac{|\theta(e^{-i\omega})|^2}{|\phi(e^{-i\omega})|^2}. \qquad (2.6)$$

AR and MA spectra are special cases of this spectrum when $\theta(x) = 1$ and $\phi(x) = 1$, respectively.

Therefore if the process {Z(t)} is an ARMA(p,q) process and is correctly fitted by the ARMA(p,q) model, then we have $S_{p,q}(\Theta) = \frac{\sigma^2}{2\pi}$, which is a spectral density of a white noise process.

Let {X(t)} be a weakly stationary process with mean $E[X(t)] = 0$ and spectral density $f_X(\omega)$. When we consider an ARMA(p,q) model fitting to this process {X(t)}, then $S_{p,q}(\Theta)$ is expressed as

$$S_{p,q}(\Theta) = \int_{-1/2}^{1/2} \frac{\left|\prod_{k=1}^{p}[1-\phi_k \exp(-2\pi i\omega)]\right|^2}{\left|\prod_{j=1}^{q}[1-\theta_j \exp(-2\pi i\omega)]\right|^2} f_X(\omega)\, d\omega. \qquad (2.7)$$

In this paper, consideration is given to the case when an ARMA(1,1) model is fitted incorrectly to an AR(2) process {X(t)}; $(1 - a B - b B^2)$X(t) = e(t). Here we set the ARMA(1,1) model parameters (x, y) in stead of (ϕ, θ). In this case, $S_{p,q}(\Theta)$ can be derived from (2.7), ignoring the constant term $\frac{\sigma^2}{2\pi}$, as

$$S_{11}(x,\, y) \;= S_{1,1}(x,\, y\,;\;\, a,\, b)$$

$$= \frac{1+b-2\,a\,x+(1+b)\,x^2+\left(-a\,(1-b)+2\left(1-b^2\right)x-a\,(1-b)\,x^2\right)y-b\left(1+b-2\,a\,x+(1+b)\,x^2\right)y^2}{(1-b)\left(1-a^2+2\,b+b^2\right)\left(1-y^2\right)\left(1+a\,y+b\,y^2\right)}. \qquad (2.8)$$

If we fit the ARMA(1, 1) model to a special AR(2) process, the function $S_{11}(x,\, y)$ will have two minimal points. For a example, we have the following graph for an AR(2) process whose parameters are $a = 0.4$, $b = 0.9$.

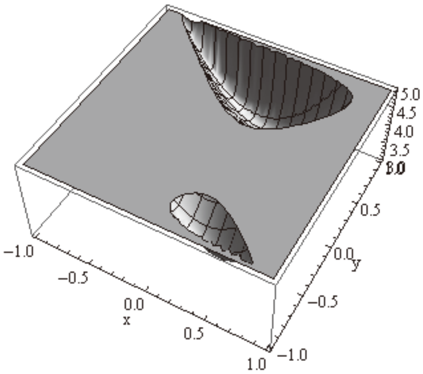


Figure 3. Cross section of $S_{11}(x,\, y)$ with $a = 0.4$, $b = 0.9$.

In order to investigate the minimal point of the function $S_{11}(x,\, y)$, it is first necessary to consider its locally minimal points on the admissible parameter space (Ω) of AR(2) process with parameters a and b, where

$$\Omega = \{(a,b);\; 0 \le (b+a+1)(b-a+1),\; -2 \le a \le 2,\; -1 \le b \le 1\}.$$

The locally minimal and maximal points satisfy the following equations,

$$\frac{\partial\, S_{11}(x,\, y)}{\partial_x} \;=\; 0, \qquad (2.9)$$

$$\frac{\partial\, S_{11}(x,\, y)}{\partial_y} \;=\; 0. \qquad (2.10)$$

We shall solve the equations. The equation (2.9) is equivalent to

$$a - x - b\,x - y + b^2\,y + a\,x\,y - a\,b\,x\,y - a\,b\,y^2 + b\,x\,y^2 + b^2\,x\,y^2 \;=\; 0. \qquad (2.11)$$

Then we have

$$x \;=\; \frac{a+\left(-1+b^2\right)y-a\,b\,y^2}{1+b-a\,y+a\,b\,y-b\,y^2-b^2\,y^2}. \qquad (2.12)$$

Also the equation (2.10) is equivalent to the following equation,

$$a - x - a^2 x + b^2 x + a x^2 - y + b y + 2 b^2 y + 2 a x y - 4 a b x y - x^2 y + b x^2 y + 2 b^2 x^2 y - a y^2 - 2 a b y^2 + a b^2 y^2 - x y^2 + 3 a^2 x y^2 + b x y^2 - a^2 b x y^2 + b^2 x y^2 - b^3 x y^2 - a x^2 y^2 - 2 a b x^2 y^2 + a b^2 x^2 y^2 + a^2 y^3 - 2 b y^3 - a^2 b y^3 - 2 b^2 y^3 - 2 a x y^3 + 4 a b x y^3 + 2 a b^2 x y^3 + a^2 x^2 y^3 - 2 b x^2 y^3 - a^2 b x^2 y^3 - 2 b^2 x^2 y^3 + 2 a b y^4 - a b^2 y^4 - 3 b x y^4 - a^2 b x y^4 + 3 b^3 x y^4 + 2 a b x^2 y^4 - a b^2 x^2 y^4 + b^2 y^5 + b^3 y^5 - 2 a b^2 x y^5 + b^2 x^2 y^5 + b^3 x^2 y^5 = 0. \qquad (2.13)$$

From (2.12) and (2.13),

$$\frac{(1-a+b)\,(1+a+b)\,(a+b\,(1+b)\,y)\left(-1+y^2\right)^2\left(b-a\,y+a\,b\,y-2\,b\,y^2+b^3\,y^4\right)}{\left(-1+a\,y+b^2\,y^2+b\left(-1-a\,y+y^2\right)\right)^2} = 0. \qquad (2.14)$$

Therefore in order to have a real solution (x, y) of the equations (2.9) and (2.10), it is necessary to have a real solution y of the equation (2.14) on the parameter space Ω. Then it is essentially equivalent to

$$(a + b\,y + b^2\,y)\,(b - a\,y + a\,b\,y - 2\,b\,y^2 + b^3\,y^4) = 0. \qquad (2.15)$$

In general, it is very difficult to solve the equation, but to know the number of the real solutions it is sufficient to consider the resultant of the polynomial

$$f(y) = (a + b\,y + b^2\,y)\,(b - a\,y + a\,b\,y - 2\,b\,y^2 + b^3\,y^4). \qquad (2.16)$$

Since the derivative of f is given by

$$\frac{\partial}{\partial y} f(y) = -a^2 + a^2 b + b^2 + b^3 - 6\,a\,b\,y + 2\,a\,b^3\,y - 6\,b^2\,y^2 - 6\,b^3\,y^2 + 4\,a\,b^3\,y^3 + 5\,b^4\,y^4 + 5\,b^5\,y^4, \qquad (2.17)$$

the resultant of the two polynomials (2.16) and (2.17) on y is give as

$$R(a,\,b) = (-1 + a - b)^2\,(-1 + b)^2\,b^{16}\,(1 + b)\,(1 + a + b)^2\,(a - b - b^2)^2 \times (a + b + b^2)^2\,(32\,a^2 - 27\,a^4 + 54\,a^4\,b + 256\,b^2 - 288\,a^2\,b^2 - 27\,a^4\,b^2 + 512\,b^3 + 256\,b^4). \qquad (2.18)$$

From the Catastrophe theory, a number of locally minimum points of $S_{11}(x,\,y)$ on Ω for AR(2) process with parameters (*a, b*) is explained by considering a change for the sign of the resultant *R(a, b)*. If the two polynomials (2.16) and (2.17) have common zeros, the resultant must be vanished. Hence we consider the conditions for *R(a, b)= 0* on Ω *={(a,b); 0≤ (b+a+1)(b-a+1), -2≤a≤2, -1≤b≤1}*.

Since the polynomial;

$$(32\,a^2 - 27\,a^4 + 54\,a^4\,b + 256\,b^2 - 288\,a^2\,b^2 - 27\,a^4\,b^2 + 512\,b^3 + 256\,b^4) \qquad (2.19)$$

in (10) is always non-negative on Ω (see Appendix), it is sufficient to consider the zeros of a polynomial

$$\text{g1}(a,\,b) = (-1 + a - b)\,(-1 + b)\,b\,(1 + b)\,(1 + a + b)\,(a - b - b^2)\,(a + b + b^2). \qquad (2.20)$$

We have the following graph of a contour of g1(a,b) = 0 on Ω.

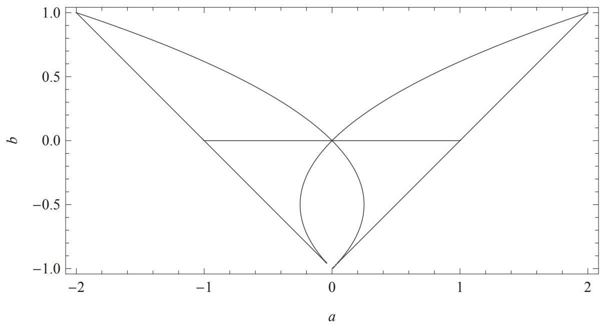


Figure 4. A contour line of g1(a,b) = 0 for AR(2) parameters (a, b).

It turns out that the function $S_{11}(x, y)$ has the two minimum points in a domain (D2) of a portion with a deep color surrounded with the curve in Fig.5, where

$$\mathrm{D2} = \{(a,\ b) \in \Omega;\ (a - b - b^2)\,(a + b + b^2) < 0\,\}. \qquad (2.21)$$

Also we define the (bifurcation) set

$$\mathrm{B2} = \{(a,\ b) \in \Omega;\ (a - b - b^2)\,(a + b + b^2) = 0\,\}. \qquad (2.22)$$

When numerical integration is performed using Mathematica (Ver.9), it turns out that the area of this domain D2 will be 2.0 square, and the rate to the parameter space of a lower triangle will be 50% exactly. It means that the domain D2 where $S_{11}(x, y)$ has 2 minimum points becomes about 3 times larger than the area of the domain D1 shown in Fig.2, since its area is about 0.70 (17.6%).

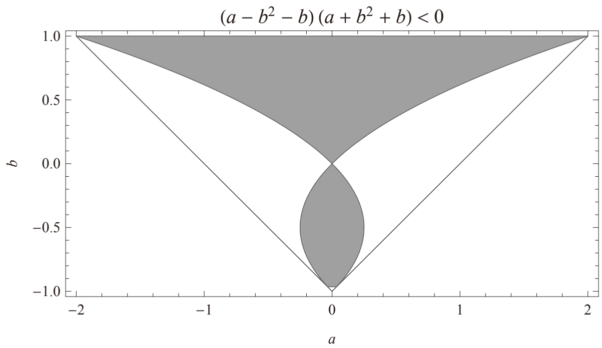


Figure 5. The domain D2 for AR(2) parameters (a, b).

We determine the form of $S_{11}(x, y)$ at every point in D2, by considering only one point within each of the domain in the next section.

3. Illustrations and simulation

3.1 Illustrations

By varying the AR(2) parameters, a and b, continuously and staying inside of D2, for example, going from position P1 to P2 in Fig.6, the system remains in a stable equilibrium that is the function $S_{11}(x, y)$ has two minima. However, if a and b are changed so that the bifurcation set B2 is transversed, something unusual happens. To see this, start in position P2 of Fig.6, where the system is in a stable equilibrium. Moving parallel to the a-axis toward position P3, when position P3 is reached, the system becomes unstablethe and the function $S_{11}(x, y)$ has only one minima. There the system is stable again and remains so while moving onward to position P4. In position P5 inside of D2, it is also seen that the function $S_{11}(x, y)$ has two minima.

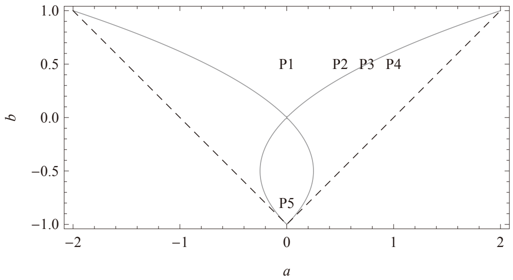


Figure 6. Selected parameters (a, b) of positions P1- P5.

[1] position P1 ; a = 0.0 and b = 0.5. In this case, $S_{11}(x, y)$ has two locally minimum points on the parameter space Ω at {x = -0.5, y = 0.732051} and {x = 0.5, y = -0.732051}, which is shown in Fig.3.1.

[2] position P2 ; a = 0.5 and b = 0.5. In this case, $S_{11}(x, y)$ has two locally minimum points on the parameter space Ω at {x = -0.0417278, y = 0.604608} and{x = 0.867418, y = -0.897478}, which is shown in Fig.3.2.

[3] position P3 ; a = 0.75 and b = 0.5 (lies in B2). In this case, $S_{11}(x, y)$ has only one locally minimum point on the parameter space Ω at {x = 0.208367, y = 0.551929}, which is shown in Fig.3.3.

[4] position P4 ; a = 1.0 and b = 0.5. In this case, $S_{11}(x, y)$ has only one locally minimum point on the parameter space Ω at {x = 0.467188, y = 0.505418}, which is shown in Fig.3.4.

[5] position P5 ; a = 0.0 and b = -0.8. In this case, $S_{11}(x, y)$ has two locally minimum points on the parameter space Ω at {x = 0.866025, y = -0.732051} and {x = -0.866025, y = 0.732051}, which is shown in Fig.3.5.

The following figures give qualitative graphs of $S_{11}(x, y)$ for the parameters (a, b) of positions P1- P5, respectively.

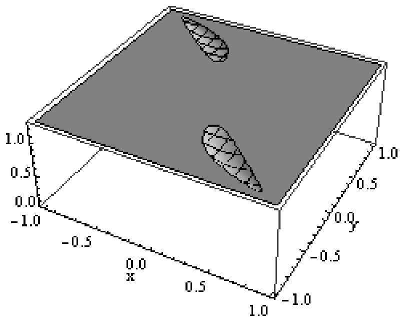


Figure 3.1. $S_{11}(x, y)$ with $a = 0.0$ and $b = 0.5$.

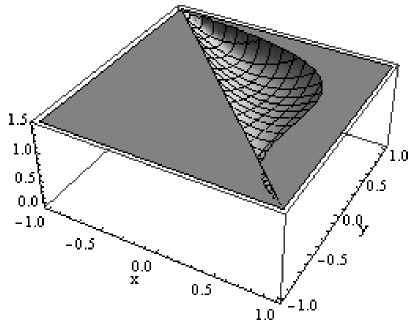


Figure 3.2. $S_{11}(x, y)$ with $a = 0.5$ and $b = 0.5$.

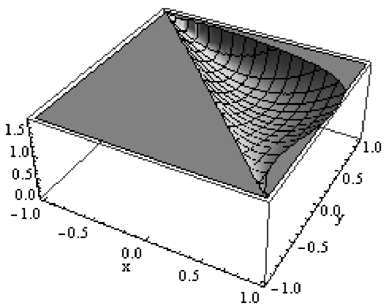


Figure 3.3. $S_{11}(x, y)$ with $a = 0.75$ and $b = 0.5$.

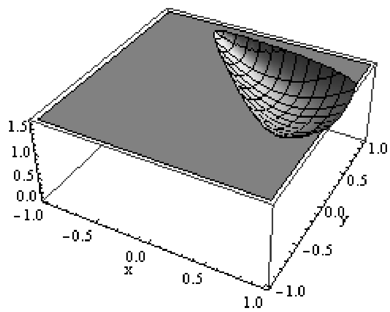


Figure 3.4. $S_{11}(x, y)$ with $a = 1.0$ and $b = 0.5$.

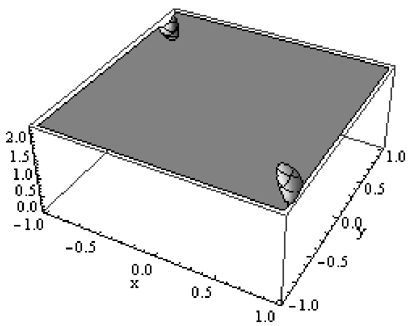
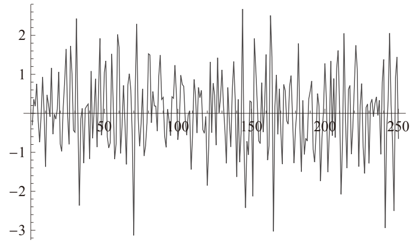


Figure 3.5. $S_{11}(x, y)$ with $a = 0.0$ and $b = -0.8$.

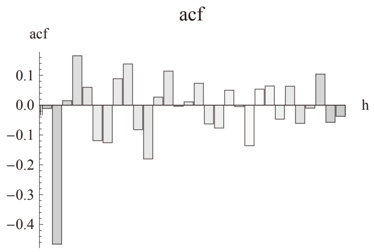
3.2 Simulation

We generate a time series of length n = 250 from the AR(2) models which are discussed above [1], ... , [5], where the noise is generated from the normal distribution with mean 0 and variance 1. Then we fit an ARMA(1,1) model to each of the time series using the conditional maximum likelihood method with initial values of parameters for the arguments (x, *y*) of the model. The calculations below are supported by the computer software *Mathematica* (V.9.0) and an application software ([5]).

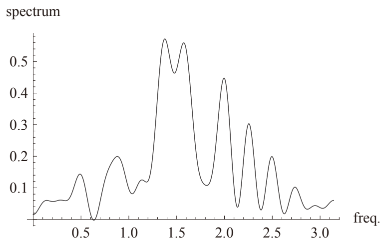
[1] Case when AR(2) process with parameters (*a, b*) = (0.0, 0.5).



Here is a plot of the data.



Here is the plot of the sample correlation function.



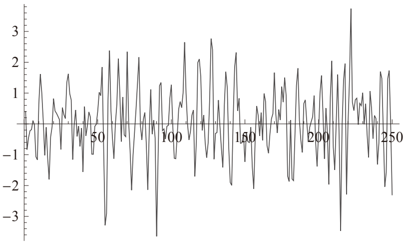
Here is the plot of the sample spectrum.

When we estimate the AR model parameters using the conditional maximum likelihood method, it turns out that AR(2) has a lower AIC value (-0.191707).

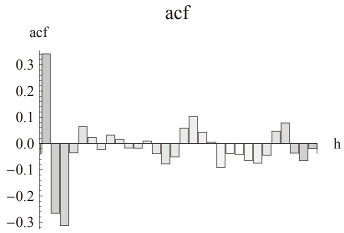
1	ARModel[{-0.0105462}, 1.03342]	0.0408735	1
2	ARModel[{-0.0133694, -0.46747}, 0.812445]	-0.191707	2
3	ARModel[{-0.0113916, -0.467345, 0.00443366}, 0.815708]	-0.179699	3
4	ARModel[{-0.0114683, -0.496194, 0.00494844, -0.0610469}, 0.812564]	-0.175561	4

Next we estimate the ARMA(1,1) model parameters using the conditional maximum likelihood method with some different initial parameter values. The initial parameter values (x = -0.5, y = 0.5) are provided as the arguments of ARMA model (1,1). Then we have ARMA model [{x = -0.553568}, {y =0.788606}, 0.956614] as the conditional maximum likelihood estimate of an ARMA(1,1) model.
On the other hand, different initial values (x = 0.5, y = -0.5) lead to another model, ARMA model [{x = 0.546659}, {y = -0.776022}, 0.963346] .
Therefore we can have two conditional maximum likelihood estimates of an ARMA(1,1) model when we fit the ARMA(1,1) model to the AR(2) process with the parameters (0.0, 0.5), which corresponds to the discussion [5] in 3.1 and also Figure 3.5.

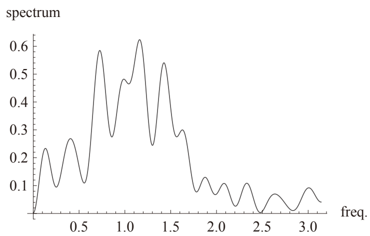
[2] Case when AR(2) process with parameters (*a, b*) = (0.5, 0.5).



Here is a plot of the data.



Here is the plot of the sample correlation function.



Here is the plot of the sample spectrum.

When we estimate the AR model parameters using the conditional maximum likelihood method, it turns out that AR(2) has a lower AIC value (-0.0608918).

1	ARModel[{0.345558}, 1.13956]	0.138642	1
2	ARModel[{0.497681, -0.436744}, 0.92599]	-0.0608918	2
3	ARModel[{0.478082, -0.413745, -0.0474331}, 0.927595]	-0.0511598	3
4	ARModel[{0.479572, -0.403076, -0.0605061, 0.0297809}, 0.930096]	-0.0404674	4

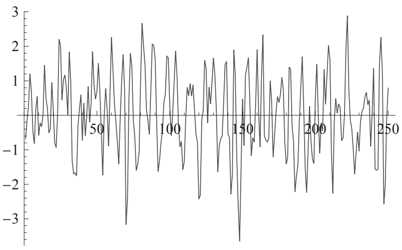
Next we estimate the ARMA(1,1) model parameters using the conditional maximum likelihood method with some different initial parameter values.

The initial parameter values (x = 0.7, y = -0.8) are provided as the arguments of ARMA model (1,1). Then we have ARMA model [{x = 0.877052}, {y = -0.908966}, 1.29477] as the conditional maximum likelihood estimate of an ARMA(1,1) model.

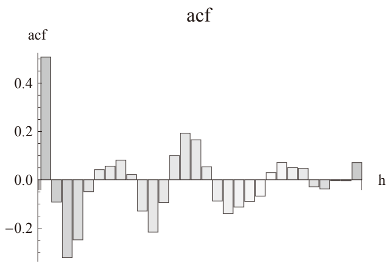
On the other hand, different initial values (x = -0.7, y = 0.8) lead to another model, ARMA model [{x = 0.0119532}, {y = 0.501642}, 1.02417].

Therefore we have two conditional maximum likelihood estimates of an ARMA(1,1) model when we fit the ARMA(1,1) model to the AR(2) process with the parameters (0.5, 0.5), which corresponds to the discussion [2] in 3.1 and also Figure 3.2.

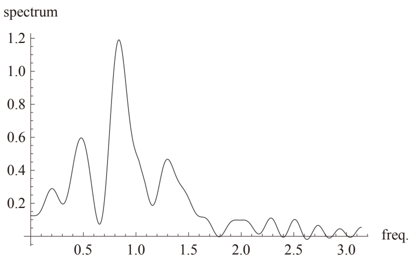
[3] Case when AR(2) process with parameters (*a, b*) = (0.75, 0.5).



Here is a plot of the data.



Here is the plot of the sample correlation function.



Here is the plot of the sample spectrum.

When we estimate the AR model parameters using the conditional maximum likelihood method, it turns out that AR(2) has a lower AIC value (-0.191475).

1	ARModel[{0.508648}, 1.04106]	0.0482427	1
2	ARModel[{0.74941, -0.471886}, 0.812634]	-0.191475	2
3	ARModel[{0.738451, -0.45428, -0.0234112}, 0.815471]	-0.17999	3
4	ARModel[{0.7365, -0.486534, 0.0305719, -0.0714131}, 0.811182]	-0.177263	4

Next we estimate the ARMA(1,1) model parameters using the conditional maximum likelihood method with some different initial parameter values.

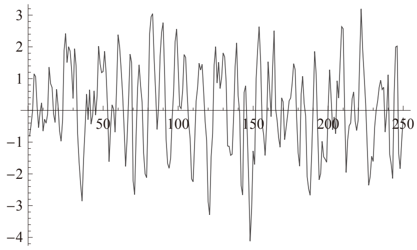
The initial parameter values (x = 0.7, y = -0.8) are provided as the arguments of ARMA model (1,1). Then we have ARMA model[{x = 0.187419}, {y = 0.583655}, 0.86927] as the conditional maximum likelihood estimate of an ARMA(1,1) model.

Different initial values (x = -0.7, y = 0.8) lead to the same model.

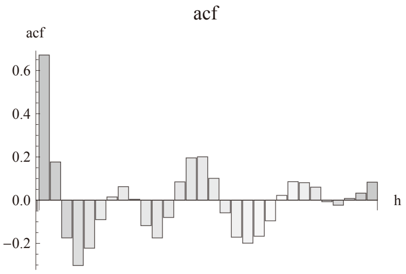
Therefore we have only one conditional maximum likelihood estimate of an ARMA(1,1) model when we fit the ARMA(1,1) model to the AR(2) process with the parameters (0.75, 0.5), which corresponds to the discussion [3] in 3.1

and also Figure 3.3.

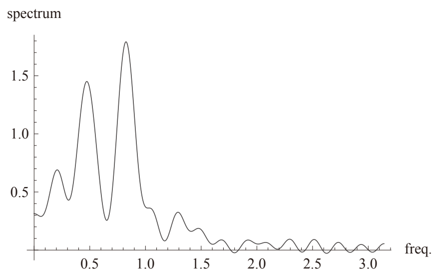
[4] Case when AR(2) process with parameters (a, b) = (1.0, 0.5).



Here is a plot of the data.



Here is the plot of the sample correlation function.



Here is the plot of the sample spectrum.

When we estimate the AR model parameters from the data using the conditional maximum likelihood method, it turns out that AR(2) has a lower AIC value (-0.190295).

1	ARModel[{0.671852}, 1.07934]	0.0843472	1
2	ARModel[{1.00781, -0.499501}, 0.813593]	-0.190295	2
3	ARModel[{0.988389, -0.46024, -0.0390077}, 0.815645]	-0.179776	3
4	ARModel[{0.986122, -0.485076, 0.0157204, -0.053866}, 0.813146]	-0.174845	4

Next we estimate the ARMA(1,1) model parameters using the conditional maximum likelihood method with some different initial parameter values.

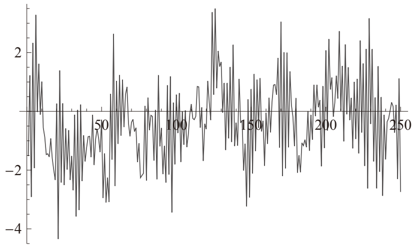
The initial parameter values (x = 0.7, y = -0.8) are provided as the arguments of ARMA model (1,1). Then we have ARMA model[{x = 0.458668}, {y= 0.535122}, 0.874477] as the conditional maximum likelihood estimate of an ARMA(1,1) model.

On the other hand, different initial values (x = 0.5, y = -0.5) lead to another model, ARMA model[{x = 0.458677}, {y

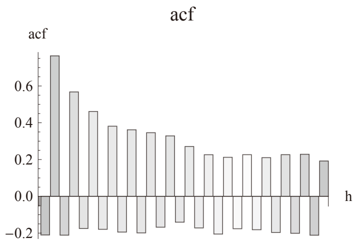
= 0.535107}, 0.874477].

Therefore we have two conditional maximum likelihood estimates of an ARMA(1,1) model when we fit the ARMA(1,1) model to the AR(2) process with the parameters (1.0, 0.5), which corresponds to the discussion [4] in 3.1 and also Figure 3.4.

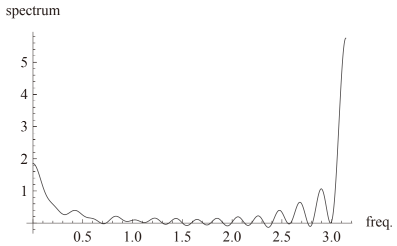
[5] Case when AR(2) process with the parameters (a, b) = (0.0, -0.8).



Here is a plot of the data.



Here is the plot of the sample correlation function.



Here is the plot of the sample spectrum.

When we estimate the AR model parameters from the data using the conditional maximum likelihood method, it turns out that AR(2) has a lower AIC value. If we fit an AR(2) model to the data, the conditional maximum likelihood estimates are given as AR model [{-0.0309124, 0.778491}, 0.907196], which means that a = -0.0309124, b = 0.77849 and σ^2 =0.907196.

1	ARModel[{-0.13487}, 2.26939]	0.82751	1
2	ARModel[{-0.0309124, 0.778491}, 0.907196]	-0.0813971	2
3	ARModel[{0.0045962, 0.7732, -0.0419463}, 0.897542]	-0.0840948	3
4	ARModel[{0.0137434, 0.791136, -0.0465155, -0.0261294}, 0.893215]	-0.0809279	4

Next we estimate the ARMA(1,1) model parameters using the conditional maximum likelihood method with some different initial parameter values.

The initial parameter values (x = -0.5, y = 0.5) are provided as the arguments of ARMA model (1,1). Then we have ARMA model [{x = -0.966694}, {y =0.820735}, 1.66297] as the conditional maximum likelihood estimate of an ARMA(1,1) model.

On the other hand, different initial values (x = 0.5, y = -0.5) lead to another model, ARMA model[{x = 0.958326}, {y = -0.845974}, 2.02728].

Therefore, depending on the initial parameter values, we have two conditional maximum likelihood estimates of an ARMA(1,1) model when we fit the ARMA(1,1) model to the AR(2) process with the parameters (0.0, -0.8), which corresponds to the case [5] in 3.1 and also Figure 3.5.

4. Conclusion

In this paper, we have considered the misspecified ARMA(1,1) model fitting to AR(2) processes. The conditions for AR(2) parameters on which ARMA(1,1) quasi-likelihood function has more than one local maximum points in the stationary and invertible parameter space were given as the domain D2 for AR(2) parameters (a, b), and it was shown in Fig.5. It related to critical point theory and the behaviour of degenerate critical points of the function of two variables in Catastrophe theory, considering the ARMA(1,1) quasi-likelihood function as a potential function with two external parameters a and b. On the misspecified MA(1) model fitting to AR(2) processes, it is already seen that the domain for AR(2) parameters on which the MA(1) quasi-likelihood function has more than one local maximum points is related to a cusp catastrophe. Our result presented in this paper will be also explained completely by using Catastrophe Theory.

Applying a stationary ARMA model to time series data in actual data analysis, there is a possibility that two or more candidates for the model parameters exist, and then we cannot estimate the parameters of the model well. We also know that the ARMA (1, 1) model seems to be more sensitive than MA (1) model about incorrect discernment. Therefore, if such a phenomenon appears in the parameter estimation for an ARMA model fitting, the applied model must be different from a true (or proper) model, and then we should change the model immediately.

Appendix

We should check the local maximal or minimal values of

$$\mathrm{g2}(a,b) = (32a^2 - 27a^4 + 54a^4 b + 256b^2 - 288a^2b^2 - 27a^4b^2 + 512b^3 + 256b^4), \text{ say.} \quad \text{(A1)}$$

Then we have

$$\frac{\partial\, \mathrm{g2}(a,b)}{\partial_a} = 64a - 108a^3 + 216a^3 b - 576ab^2 - 108a^3b^2 = 0, \quad \text{(A2)}$$

$$\frac{\partial\, \mathrm{g2}(a,b)}{\partial_b} = 54a^4 + 512b - 576a^2 b - 54a^4 b + 1536b^2 + 1024b^3 = 0\,. \quad \text{(A3)}$$

The solutions of real number for the equations (10) and (11) in Ω *={(a,b); 0≤(b+a+1)(b-a+1), -2≤a≤2, -1≤b≤1}* are

$$\{a=0,\ b=-1\},\ \left\{a=0,\ b=-\frac{1}{2}\right\},\ \{a=0,\ b=0\},$$

$$\left\{a=-\frac{4}{3}\sqrt{-10+\sqrt{105}}\ ,\ b=\frac{1}{12}\left(9-\sqrt{105}\right)\right\} \text{ and } \left\{a=\frac{4}{3}\sqrt{-10+\sqrt{105}}\ ,\ b=\frac{1}{12}\left(9-\sqrt{105}\right)\right\}.$$

The local maximal or minimal values of g2(a,b) at those points are {0, 16, 0, 8.561, 8.561}, respectively. Also *g2(a,b) ≥ 0* on the boundary of Ω. Thus the minimum of the function g2(a,b) is 0 at (a, b) = (0, -1) and (0, 0). □

Acknowledgements

The author is very grateful to the late Prof. Kenji Aoki for many useful comments and suggestions.

References

[1] Åström, K J. and Söderström, T., 1974, "Uniqueness of the maximum likelihood estimates of the parameters of an ARMA model", IEEE Trans. Automat. Contr., 19, 769-773.
[2] Box, G. E. P. and Jenkins, G. M., 1970, *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.
[3] Brockwell, P. J. and Davis, R. A., 1991, *Time Series : Theory and Methods,* Springer, New York.
[4] Castrigiano, D.P.L. and Hayes, S.A., 2004, *Catastrophe theory*, Westview Press.
[5] Huzii , M., 1988, "Some properties of conditional quasi-likelihood functions for time series model fitting", Journal of Time Series Analysis, 9, 345-352.
[6] He,Y., 1995, *Time Series Pack for Mathematica*, Wolfram Research.
[7] Kabaila, P., 1983, "Parameter values of ARMA models minimizing the one-step-ahead prediction error when the true system is not in the model set", J. Appl . Prob., 20, 405-408.
[8] Poston,T. and Stewart, I.N., 1978, *Catastrophe theory and its applications,* Pitman Publishing Limited.
[9] Tanaka, M. and Aoki, K., 1991, "On a moving average time series model fitting" (in Japanese), Bulletin of the Institute of Information Science 12, 42 - 54.
[10] Tanaka, M. and Huzii,M., 1992, "Some properties of moving - average model fittings", J. Japan Statist. Soc., Vol.22, No .1, 33 - 44.