PARAMETER ESTIMATION OF MOVING AVERAGE PROCESSES USING CUMULANTS AND NONLINEAR OPTIMIZATION ALGORITHMS

M. Boulouird, M. M. Hassani
Laboratoire d'Electronique et Instrumentation LEI (FSSM/UCAM)
Department of Physics, Faculty of Sciences Semlalia
Avenue Prince My Abdellah, B.P 2390, 40001 Marrakesh, Morocco

G. Favier
Laboratoire d'Informatique, Signaux et Systèmes de Sophia-Antipolis I3S (CNRS/UNSA)
Les algorithmes - Bâtiment Euclide B, 2000 route des Lucioles
B.P 121, Sophia Antipolis Cedex France

Keywords: MA system identification, Higher-Order Statistics, Estimation parameter, Linear algebra solution, Gradient descent algorithm, Gauss-Newton algorithm, Cumulants.

Abstract: In this paper nonlinear optimization algorithms, namely the Gradient descent and the Gauss-Newton algorithms, are proposed for blind identification of MA models. A relationship between third and fourth order cumulants of the noisy system output and the MA parameters is exploited to build a set of nonlinear equations that is solved by means of the two nonlinear optimization algorithms above cited. Simulation results are presented to compare the performance of the proposed algorithms.

1 INTRODUCTION

Numerous methods have been proposed in the literature for blind identification of MA models using cumulants. The present paper is concerned with the linear algebra solutions approach. It consists in constructing a system of equations obtained from explicit relations that link third and fourth order cumulants of the noisy output with the MA parameters and solving this system by the least-squares method ((Alshebeili, 1993), (Giannakis, 1989), (Martin, 1996), (Na, 1995), (Srinivas, 1995), (Stogioglou, 1996), (Tugnait, 1990), (Tugnait, 1991)). In order to take the redundancy in the unknown parameters vector into account, (Abderrahim, 2001) proposed a constrained optimization based solution.

In this paper, we propose another approach to reduce this redundancy. It consists in exploiting the nonlinearity existing in the unknown parameters estimated vector. In the literature, the parameters of the vector to be estimated are regarded as independent, but actually it isn't the case. Thus the major contribution of this paper lies in the estimates of a non redundant vector of unknown parameters.

The organization of this paper is as follows. The problem formulation is given in Section 2. In Section 3, the resolution with least squares and the nonlinear optimization algorithms, in the event Gradient descent and Gauss-Newton algorithms, well be developed. Computer simulation results are given in Section 4 to show the effectiveness of the proposed techniques. Finally, the paper is concluded in Section 5.

2 PROBLEM FORMULATION

We consider the discrete, causal, linear time-invariant process represented on figure 1, with the following assumptions :

H.1. The input $w(k)$ is a zero mean, independent and identically distributed (i.i.d), stationary non-Gaussian, non measurable real sequence, with unknown distribution, and :

$$C_{m,w}(\tau_1, \tau_2, \ldots, \tau_{m-1}) = \gamma_{m,w}\ \delta(\tau_1, \tau_2, \ldots, \tau_{m-1})$$

where :

- ⋄ $C_{m,w}(\tau_1, \tau_2, \ldots, \tau_{m-1})$ is the m**th**-order cumulant of the input signal of the MA model.
- ⋄ $\gamma_{m,w} \neq 0, \quad \forall \ \ m \geq 2$
- ⋄ $\gamma_{m,w} = C_{m,w}(\underbrace{0, 0, \ldots, 0}_{m-1})$

Boulouird M., M. Hassani M. and Favier G. (2005).
PARAMETER ESTIMATION OF MOVING AVERAGE PROCESSES USING CUMULANTS AND NONLINEAR OPTIMIZATION ALGORITHMS.
In *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics - Signal Processing, Systems Modeling and Control*, pages 11-15
DOI: 10.5220/0001182900110015

- $\diamond$ $\gamma_{2,w} = \sigma_w^2 = E\left\{w(k)^2\right\}$ is the variance of $w(k)$.
- $\diamond$ $\gamma_{3,w} = E\left\{w(k)^3\right\}$ is the skewness of $w(k)$.
- $\diamond$ $\gamma_{4,w} = E\left\{w(k)^4\right\} - 3\left[E\left\{w(k)^2\right\}\right]^2$ is the kurtosis of $w(k)$.

H.2. The additive noise $v(k)$ is assumed to be an i.i.d Gaussian sequence with unknown variance, zero-mean, and independent of $w(k)$.

H.3. The non measurable output $x(k)$ is assumed to be a nonminimum or minimum phase MA process.

H.4. The order q of the model is assumed to be known.

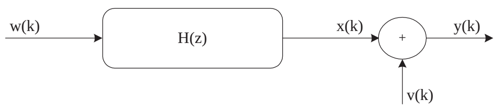

Figure 1: Single-channel system.

The measured noisy MA process $y(k)$ is represented by the following equations :

$$x(k) = \sum_{i=0}^{q} h(i)\, w(k-i); \qquad \{h(0)=1\} \tag{1}$$

$$y(k) = x(k) + v(k) \tag{2}$$

For the MA model described by equation (1) with the assumptions **H.1**, **H.2**, **H.3**, and **H.4**, the m**th** and n**th**-order cumulants of the MA system output (2) are linked by the following relation (Abderrahim, 2001) :

$$\sum_{i=i_{min}}^{i_{max}} h(i)\left[\prod_{k=1}^{m-s-1} h(i+\tau_k)\right] \times C_{n,y}(\beta_1,\beta_2,\ldots,\beta_{n-s-1},i+\alpha_1,i+\alpha_2,\ldots,i+\alpha_s) = \frac{\gamma_{n,w}}{\gamma_{m,w}} \sum_{j=j_{min}}^{j_{max}} h(j)\left[\prod_{k=1}^{n-s-1} h(j+\beta_k)\right] \times C_{m,y}(\tau_1,\tau_2,\ldots,\tau_{m-s-1},j+\alpha_1,j+\alpha_2,\ldots,j+\alpha_s) \tag{3}$$

where $m > 2$, $n > 2$ and s is an arbitrary integer number satisfying : $1 \le s \le min(m,n)-2$,

$$and \quad \left\{\begin{array}{l} i_{min} = max(0,-\tau_1,\cdots,-\tau_{m-s-1}) \\ i_{max} = min(q,q-\tau_1,\cdots,q-\tau_{m-s-1}) \\ j_{min} = max(0,-\beta_1,\cdots,-\beta_{n-s-1}) \\ j_{max} = min(q,q-\beta_1,\cdots,q-\beta_{n-s-1}) \end{array}\right.$$

Setting $n=3$, $m=4$, and $s=1$ in equation (3), yields

$$\sum_{i=i_{min}}^{i_{max}} h(i)h(i+\tau_1)h(i+\tau_2)C_{3,y}(\beta_1,i+\alpha_1) = \frac{\gamma_{3,w}}{\gamma_{4,w}} \sum_{j=j_{min}}^{j_{max}} h(j)h(j+\beta_1)C_{4,y}(\tau_1,\tau_2,j+\alpha_1) \tag{4}$$

$$where \quad \left\{\begin{array}{l} i_{min} = max(0,-\tau_1,-\tau_2) \\ i_{max} = min(q,q-\tau_1,q-\tau_2) \\ j_{min} = max(0,-\beta_1) \\ j_{max} = min(q,q-\beta_1) \end{array}\right.$$

By setting $\tau_1 = \tau_2 = 0$ in (4), we get the relation used in this paper for estimating the parameters $\{h(i)\}_{i=1,2,\ldots,q}$ of the MA model.

$$\sum_{i=0}^{q} h^3(i)C_{3,y}(\beta_1,i+\alpha_1) = \frac{\gamma_{3,w}}{\gamma_{4,w}} \sum_{j=j_{min}}^{j_{max}} h(j)h(j+\beta_1)C_{4,y}(0,0,j+\alpha_1) \tag{5}$$

It is important to determine the range of values of α_1 and β_1 so that the cumulants $\{C_{3,y}(\beta_1,i+\alpha_1)\}_{i=0,\cdots,q}$, $\{C_{4,y}(0,0,j+\alpha_1)\}_{j=j_{min},\cdots,j_{max}}$, and the coefficients $\{h(j+\beta_1)\}$ be not all zero for each equation.

By taking account of the property of causality of the model and the domain of support for third and fourth order cumulants of an MA(q) process (Mendel, 1991), we obtain :

$$\left\{\begin{array}{l} -q \le \beta_1 \le q \\ -2q \le \alpha_1 \le q \\ -2q+\beta_1 \le \alpha_1 \le q+\beta_1 \end{array}\right. \tag{6}$$

Using the symmetry properties of cumulants (Nikias, 1993), the set of values for α_1 and β_1 is defined by :

$$\left\{\begin{array}{l} -q \le \beta_1 \le 0 \\ -2q \le \alpha_1 \le q+\beta_1 \end{array}\right. \tag{7}$$

3 PARAMETER ESTIMATION

3.1 Least-Squares (LS) Solution

Concatenating (5) for all values of α_1 and β_1 defined by (7), we obtain the following system of equations :

$$M\theta = r \tag{8}$$

where :

$$\theta = [h(1) \cdots h(q) \quad h^2(1) \quad h(1)h(2) \cdots h(1)h(q) \quad h^2(2) \cdots h(2)h(q) \quad h^2(3) \cdots h^2(q) \quad \epsilon_{4,3} \quad \epsilon_{4,3}h^3(1) \cdots \epsilon_{4,3}h^3(q)]^T \quad (9)$$

- ◇ $\epsilon_{4,3} = \gamma_{4,w}/\gamma_{3,w}$.
- ◇ M is a matrix of dimension $\left[\frac{5q^2+7q+2}{2}, \frac{q^2+5q+2}{2}\right]$.
- ◇ θ is a vector of dimension $\left[\frac{q^2+5q+2}{2}, 1\right]$.
- ◇ r is a vector of dimension $\left[\frac{5q^2+7q+2}{2}, 1\right]$.

Assuming that $(M^T M)$ is invertible, the unique **LS** estimate of θ is

$$\hat{\theta} = (M^T M)^{-1} M^T r \quad (10)$$

Solving (8) provides the parameter estimates $\{h(i)\}_{i=1,\cdots,q}$ as the first q components of the estimated parameter vector $\hat{\theta}$ in (10).

3.2 Gradient Descent Algorithm (GDA)

The idea satisfying this paper is to reduce the dimension of the estimated parameter vector (9). In section 3.1, θ is a vector of $\left(\frac{q^2+5q+2}{2}\right)$ elements. The linear algebra solutions regard the elements of the vector θ clarified in relation (9) as independent parameters, but the dependence of these elements is almost obvious. To palliate the problem of redundancy in θ, we propose a new approach based on non-linear optimization algorithm. In this part, the parameters vector θ is a $(q+1)$ length vector. It has this form :

$$\theta_{NL} = [h(1), \cdots, h(q), \epsilon_{4,3}]^T \quad (11)$$

The criterion to be minimized in this case is as follows :

$$J_{LS} = \|r - \phi(\theta_{NL})\|^2$$

The **GDA** solution has the following form :

$$\hat{\theta}^{i+1}_{NL_{gr}} = \hat{\theta}^{i}_{NL_{gr}} + \lambda J^T (r - \phi(\hat{\theta}^{i}_{NL_{gr}})) \quad (12)$$

where :

- ◇ r is defined in section 3.1.
- ◇ ϕ is the system of equations obtained by concatenating (5) for all values of α_1 and β_1 defined by (7).
- ◇ J is the Jacobian matrix of ϕ,

$$J = \left[\frac{\partial \phi_k}{\partial \theta_{NL_l}}\right]_{(k,l)}$$

  where $k = 1, \cdots, \frac{5q^2+7q+2}{2}$, and $l = 1, \cdots, q+1$.
- ◇ λ is the step-size.

The parameter $\epsilon_{4,3}$ must be estimated since we are supposed that we don't know the nature of the distribution of the input signal $w(k)$.

3.3 Gauss-Newton Algorithm (GNA)

This algorithm has this form :

$$\hat{\theta}^{i+1}_{NL_{gn}} = \hat{\theta}^{i}_{NL_{gn}} + \mu (J^T J)^{-1} J^T (r - \phi(\hat{\theta}^{i}_{NL_{gn}})) \quad (13)$$

where :

- ◇ r, ϕ, and J are defined in section 3.2.
- ◇ $\hat{\theta}_{NL_{gn}}$ has the form of (11).
- ◇ μ is the step-size of this algorithm.

4 SIMULATION RESULTS

To demonstrate the effectiveness of the proposed techniques, let us examine two examples treated in the literature. In both models the input signal $w(k)$ is a zero-mean exponentially distributed i.i.d noise sequence with $\gamma_{2,w} = \sigma^2_w = 1$ and $\gamma_{3,w} = 2$. We define the Signal-to- Noise Ratio as

$$SNR(dB) = 10\, log_{10} \left(E[x^2(k)]/E[v^2(k)]\right)$$

For each run, we calculate the Normalized Mean Square Error (NMSE) defined as

$$NMSE = \frac{\sum_{i=1}^{q} \left(h(i) - \widehat{h}(i)\right)^2}{\sum_{i=1}^{q} h^2(i)}$$

where $h(i)$ and $\widehat{h}(i)$ are respectively the actual and the estimated impulse responses, respectively. The Error to Signal Ratio (ESR) in decibels is also used as a measure of the estimation error. The ESR is defined as

$$ESR(dB) = 10\, log_{10}(NMSE)$$

Example 1 :

$$y(k) = w(k) - 2.3333w(k-1) + 0.6667w(k-2) + v(k)$$

The zeros of the system transfer function $H(z)$ are located at 2 and 0.3333. This model has also been used in (Abderrahim, 2001), (Giannakis, 1989), and (Srinivas, 1995).

Additive colored noise is generated as the output of the following MA(2) model (Abderrahim, 2001) :

$$v(k) = e(k) + 2.3333e(k-1) - 0.6667e(k-2)$$

where the input sequence $e(k)$ is an i.i.d Gaussian sequence. We carried out Monte Carlo simulations with $K = 100$ different noise sequences, $N = 5120$ data for each run, and three different values of SNR ($20dB$, $10dB$, and $0dB$). The simulation results are summarized in Tables 1, 2, and 3.

Example 2 :

$$y(k) = w(k) + 0.1w(k-1) - 1.87w(k-2) + 3.02w(k-3) - 1.435w(k-4) + 0.49w(k-5) + v(k)$$

Table 1: System identification using signal model MA(2), Mean, Standard deviation, NMSE, and ESR . SNR=20dB.

	True Values		LS	GDA	GNA
				$\lambda = 8.\ 10^{-5}$	$\mu = 0.5$
h(1)	-2.3333	mean	-2.1704	-2.2676	-2.2676
		σ	0.0585	0.0149	0.0149
h(2)	0.6667	mean	0.5101	0.6541	0.6541
		σ	0.0211	0.0060	0.0060
NMSE		mean	0.0093	0.0008	0.0008
		σ	0.0043	0.0003	0.0003
ESR			-20.3041	-30.9500	-30.9500

Table 2: System identification using signal model MA(2), Mean, Standard Deviation, NMSE, and ESR. SNR=10dB.

	True Values		LS	GDA	GNA
				$\lambda = 8.\ 10^{-5}$	$\mu = 0.5$
h(1)	-2.3333	mean	-2.1013	-2.1158	-2.2624
		σ	0.1835	0.4664	0.0511
h(2)	0.6667	mean	0.4987	0.6166	0.6557
		σ	0.0632	0.1275	0.0192
NMSE		mean	0.0203	0.0478	0.0014
		σ	0.0181	0.1490	0.0015
ESR			-16.9250	-13.2057	-28.5387

The zeros of the system transfer function $H(z)$ are located at -2, $0.7 \pm j0.7$ and $0.25 \pm j0.433$. This model has also been used in (Alshebeili, 1993), (Stogioglou, 1996), (Tugnait, 1990), and (Tugnait, 1991).

Additive colored noise is generated as the output of the following MA(3) model (Abderrahim, 2001) and (Na, 1995) :

$$v(k) = e(k) + 0.5e(k-1) - 0.25e(k-2) + 0.5e(k-3)$$

where the input sequence $e(k)$ is an i.i.d Gaussian sequence. In this case $N = 10240$. The simulation results are given in tables 4, 5, and 6.

In these simulations, we initialize the nonlinear optimization algorithms with the LS solution. The advantage of this is to avoid the convergence to local minimum. The good value of the step-size allows also to avoid this problem of the local minima.

Table 1 shows that the parameters estimation via the proposed method is much more powerful than LS solution in term of mean value, standard deviation, or NMSE, and ESR . The Gradient Descent and Gauss-Newton algorithms converge to the same values, but the convergence speed is different. The figures 2 and 3 illustrate this.

Tables 2 and 5 show that the Gradient Descent algorithm is sensitive to the additive noise. However, the Gauss-Newton algorithm is more robust to measurement noise.

We can note that for a low SNR ($SNR = 0dB$), the estimate of the parameters is poor what is due to the bad estimate of third and fourth order cumulants. These estimation results could be improved by processing more data, which allows to get better estimates of the cumulants.

Table 3: System identification using signal model MA(2), Mean, Standard Deviation, NMSE, and ESR. SNR=0dB.

	True Values		LS	GDA	GNA
				$\lambda = 8.\ 10^{-5}$	$\mu = 0.5$
h(1)	-2.3333	mean	-1.3781	-0.8631	-1.0952
		σ	0.6033	0.6223	1.0662
h(2)	0.6667	mean	0.3217	0.3006	0.2882
		σ	0.1794	0.2092	0.3499
NMSE		mean	0.2418	0.4623	0.4964
		σ	0.2464	0.2762	0.4890
ESR			-6.1661	-3.3511	-3.0421

Table 4: System identification using signal model MA(5), Mean, Standard Deviation, NMSE, and ESR. SNR=20dB.

	True Values		LS	GDA	GNA
				$\lambda = 6.\ 10^{-9}$	$\mu = 0.5$
h(1)	0.1	mean	0.0474	0.0779	0.0887
		σ	0.0359	0.0384	0.0392
h(2)	-1.87	mean	-1.4212	-1.2822	-1.7798
		σ	0.0845	0.0665	0.1010
h(3)	3.02	mean	2.2489	1.6874	2.8902
		σ	0.1258	0.0667	0.1494
h(4)	-1.435	mean	-1.0779	-1.1768	-1.4206
		σ	0.0671	0.0754	0.0876
h(5)	0.49	mean	0.3640	0.3927	0.4536
		σ	0.0275	0.0301	0.0320
NMSE		mean	0.0651	0.1484	0.0046
		σ	0.0221	0.0184	0.0075
ESR			-11.8660	-8.2867	-23.3527

By increasing the order q of the model, we use the third and fourth order cumulants with large lags that are poorly estimated for a low SNR. The consequence of this is the bad parameters estimated (table 6) much poorer than in the case of a model of order 2 (table 3).

These two algorithms are numerically expensive compared to the Least-Squares algorithm.

5 CONCLUSION

In this paper, a blind identification of the MA models using Higher-Order Statistics (HOS) is exposed. The linear algebra solution is compared with the nonlinear optimization algorithms solution. Computer simulation results prove that it is interesting to use this algorithms in spite of their expensive calculative cost.

REFERENCES

Abderrahim, K., Ben Abdennour, R., Favier, G., Ksouri, M., and Msahli, F. (2001). New Results on FIR Identification using Cumulants. *Journal Européen des Systèmes Automatisés-JESA, Vol. 35, No. 5, pp. 601-622*.

Alshebeili, S. A., Venetsanopoulos, A. N., and Cetin, A. E. (1993, April). Cumulant Based Identification Approaches for Nonminimum Phase FIR Systems. *IEEE Transactions on Signal Processing, Vol. 41, No. 4, pp. 1576-1588*.

Giannakis, G. B., and Mendel, J. M. (1989, March). Identification of Nonminimum Phase Systems Using Higher Order Statistics. *IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, No. 3, pp. 360-377*.

Martin, J. K., and Nandi, A. K. (1996, January). Blind System Identification using Second, Third and Fourth Order Cumulants. *Journal of the Franklin Institute : Enginnering and applied mathematics, Pergamon, Vol. 333(B), No. 1, pp. 1-13*.

Mendel, J. M. (1991, March). Tutorial on Higher-Order Statistics (Spectra) in Signal Processing and System Theory : Theoretical Results and some Applications. *Proceedings of the IEEE, Vol. 79, No. 3, pp. 278-305*.

Na, Y. N., Kim, K. S., Song, I., and Kim, T. (1995, December). Identification of Nonminimum Phase FIR Systems Using the Third- and Fourth-Order Cumulants. *IEEE Transactions on Signal Processing, Vol. 43, No. 8, pp. 2018-2022*.

Nikias, C. L., and Mendel, J. M. (1993, July). Signal Processing With Higher Order Spectra. *IEEE Signal Processing Magazine, pp. 10-37*.

Srinivas, L., and Hari, K. V. S. (1995, December). FIR System Identification Using Higher Order Cumulants - A Generalized Approach. *IEEE Transactions on Signal Processing, Vol. 43, No. 12, pp. 3061-3065*.

Stogioglou, A. G., and McLaughlin, S. (1996, July). MA Parameter Estimation and Cumulant Enhancement. *IEEE Transactions on Signal Processing, Vol. 44, No. 7, pp. 1704-1718*.

Tugnait, J. K. (1990, July). Approaches to FIR System Identification With Noisy Data Using Higher Order Statistics. *IEEE Transactions on Signal Processing, Vol. 38, No. 7, pp. 1307-1317*.

Tugnait, J. K. (1991, October). New Results on FIR System Identification Using Higher-Order Statistics. *IEEE Transactions on Signal Processing, Vol. 39, No. 10, pp. 2216-2221*.

Table 5: System identification using signal model MA(5), Mean, Standard Deviation, NMSE, and ESR. SNR=10dB.

True Values			LS	GDA	GNA
				$\lambda = 6.\,10^{-9}$	$\mu = 0.5$
h(1)	0.1	mean	-0.0929	-0.0841	-0.0604
		σ	0.1145	0.1193	0.1161
h(2)	-1.87	mean	-1.0274	-0.9433	-1.2443
		σ	0.2466	0.2177	0.3461
h(3)	3.02	mean	1.6917	1.4036	2.1096
		σ	0.3516	0.2240	0.5103
h(4)	-1.435	mean	-0.8187	-0.8680	-0.9787
		σ	0.1748	0.1971	0.2740
h(5)	0.49	mean	0.2760	0.2902	0.3151
		σ	0.0653	0.0692	0.0887
NMSE		mean	0.2122	0.2694	0.1311
		σ	0.1105	0.0959	0.1345
ESR			-6.7328	-5.6965	-8.8246

Table 6: System identification using signal model MA(5), Mean, Standard Deviation, NMSE, ESR. SNR=0dB.

True Values			LS	GDA	GNA
				$\lambda = 6.\,10^{-9}$	$\mu = 0.5$
h(1)	0.1	mean	-0.2698	-0.2888	-0.3315
		σ	0.1083	0.1154	0.1085
h(2)	-1.87	mean	-0.1297	-0.1187	-0.0863
		σ	0.1577	0.1459	0.0811
h(3)	3.02	mean	0.3457	0.3302	0.2627
		σ	0.2129	0.2021	0.1582
h(4)	-1.435	mean	-0.1644	-0.1630	-0.0949
		σ	0.1068	0.1061	0.0845
h(5)	0.49	mean	0.0420	0.0427	0.0237
		σ	0.0366	0.0359	0.0321
NMSE		mean	0.8191	0.8280	0.8732
		σ	0.1353	0.1300	0.0936
ESR			-0.8667	-0.8199	-0.5887

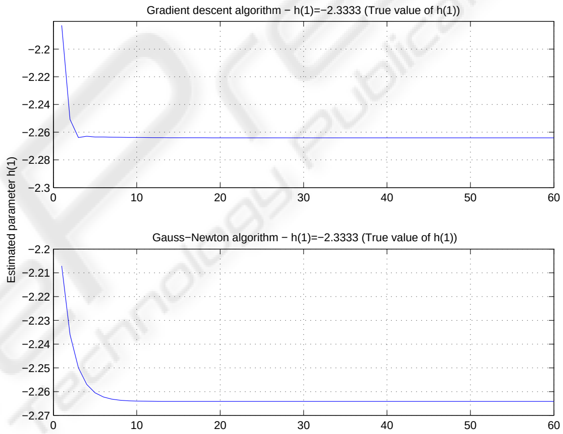

Figure 2: The convergence of h(1).

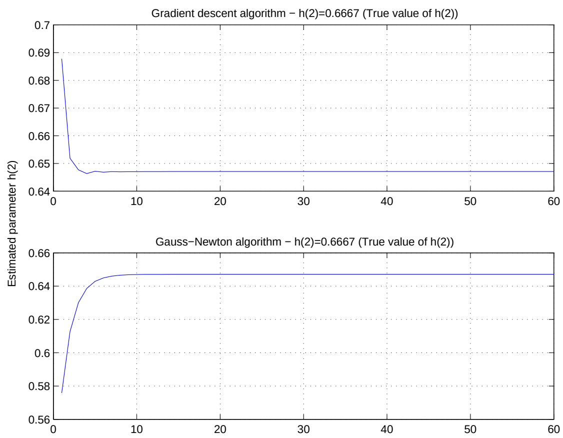

Figure 3: The convergence of h(2).