Discussiones Mathematicae
Probability and Statistics 30 (2010) 103–115

ON THE ESTIMATION OF THE AUTOCORRELATION FUNCTION

Manuel Duarte Ortigueira

UNINOVA, Campus da FCT da UNL,
Quinta da Torre, 2825–114 Monte da Caparica, Portugal

e-mails: mdortigueira@uninova.pt ; mdo@fct.unl.pt

Phone +351 1 2948520, Fax +351 1 2957786

Abstract

The autocorrelation function has a very important role in several application areas involving stochastic processes. In fact, it assumes the theoretical base for Spectral analysis, ARMA (and generalizations) modeling, detection, etc. However and as it is well known, the results obtained with the more current estimates of the autocorrelation function (biased or not) are frequently bad, even when we have access to a large number of points. On the other hand, in some applications, we need to perform fast correlations. The usual estimators do not allow a fast computation, even with the FFT. These facts motivated the search for alternative ways of computing the autocorrelation function. 9 estimators will be presented and a comparison in face to the exact theoretical autocorrelation is done. As we will see, the best is the AR modified Burg estimate.

Keywords: time-series autocorrelation, regression.

2000 Mathematics Subject Classification: 37M10.

1. Introduction

The correlation in general and the autocorrelation function in particular are tools that belong to the daily life in Signal Processing, independently of the application scientific field. In fact everybody knows that Spectral Analysis is based on the commonly named Wiener-Khintchin theorem* which states that the spectrum is the Fourier transform of the autocorrelation function. Other interesting application of the correlation is the ARMA modelling, very important in applications. Other usual uses of the correlation can be found in detection, for example, in Communications and Radar, delay measurement, etc.

There is a slight difference between the engineering definition and the one used in other areas like Statistics or Economy. In engineering the autocorrelation function is defined by [12]:

$$R(\tau) = E\{x(t)x(t+\tau)\} \tag{1}$$

and $R(0)$ is the power of the signal $x(t)$. With η as the mean value of the process $x(t)$, the Autocovariance is always defined by [1, 12]

$$C(\tau) = R(\tau) - \eta^2 \tag{2}$$

leading to the autocorrelation definition used in Statistics or Economy

$$\rho(\tau) = \frac{C(\tau)}{C(0)}. \tag{3}$$

In Engineering this is called Correlogram. As it is a common assumption: η=0

$$R(\tau) = C(\tau) \text{ and } \rho(\tau) = \frac{R(t)}{R(0)}. \tag{4}$$

The difference is merely a normalization useful in comparing different estimates, but without interest for spectral estimation or ARMA modelling.

In almost all the practical applications, the autocorrelation function must be estimated. This may lead to poor results. Here we will study the behavior of several estimators of the autocorrelation function from the point of view of the bias and also by studying the autocorrelation Toeplitz matrix due to its importance in parameter estimation and spectrum analysis.

In the next section we will consider the current estimators and their problems. In Section 3 we present some alternatives that are evaluated in Section 4. At last we will present some conclusions.

*Some authors prefer to rename it as Wiener-Khintchin-Einstein theorem, due to the 1905 Einstein's paper on the Brownian motion.

2. Current estimators and their problems

Let $x_n, n = 0, 1, \ldots, L-1$, be a realization of a stationary stochastic process. Usual estimators are [8,9]:
the unbiased

$$R_u(k) = \frac{1}{L-|k|} \sum_{i=0}^{L-|k|} x_i.x_{i+|k|}, \quad |k| = 0, 1, 2, ... \tag{5}$$

the biased

$$R_b(k) = \frac{1}{L} \sum_{i=0}^{L-|k|} x_i.x_{i+|k|}, \quad |k| = 0, 1, 2, ... \tag{6}$$

We can verify easily that $R_b(n) = R_u(n).w(n)$, where $w(n) = 1 - \frac{|n|}{L}; |n| < L$, the so-called Bartlett (triangular) window. Their main properties can be found in [12]. These estimators use a variable summation in the sense that for a given k, we perform $L - |k|$ additions leading to a triangular effect. Alternatively we can use a fixed number of additions (FS)

$$R_f(k) = \frac{1}{L-|N|} \sum_{i=0}^{L-N} x_i.x_{i+|k|}, \quad |k| = 0, 1, 2, ..., N \tag{7}$$

or a half delay definition[†]

$$R(k) = \frac{1}{L-|N|} \sum_{i=0}^{L-N} x_{i-k/2}.x_{i+k/2}, \quad |k| = 0, 1, 2, ..., N, \tag{8}$$

that gives similar results. Normally, these estimators lead to poor estimates forcing us to look for alternatives or algorithms that do not use it [4]. In the AR case there are several better alternatives [2, 3, 6, 7, 10].

To have an idea of the problems we find in parameter estimation, we present in the following tables several results obtained in estimating the AR parameters. We used the well known Yule-Walker method and a modified

[†] See [11] for the fractional delay definition.

Burg method (MBM)[‡] [10]. An AR model with order 6 was randomly generated to create AR signals. We present the exact AR parameters and the average estimates obtained over 100 trials and for two lengths of signals, 500 and 200.

Table 1. AR (6) estimations using 500 points.

Exact	MBM	YW		Exact	MBM	YW
−3.7783	−3.7654	−2.2129		0.9163	0.9087	0.9059
6.4247	6.3821	1.5645		0.1767	0.1660	0.1647
−6.2929	−6.2322	0.0151		−0.0506	−0.0500	−0.0496
3.7838	3.7391	−0.2975		−0.0156	−0.0081	−0.0081
−1.3909	−1.3737	−0.2563		0.0007	0.0041	0.0043
0.2684	0.2653	0.2334		0.0003	0.0045	0.0045

As we can see, the MBM method gives slightly better results. However, if we reduce the number of data points, the results become worst for both methods and the Yule-Walker performs poorly.

Table 2. AR (6) estimations using 200 points.

Exact	MBM	YW	Exact	MBM	YW
−1.8363	−1.8213	−1.6996	−1.9243	−1.9180	−1.5814
1.3869	1.3643	1.1346	2.2417	2.2236	1.5538
−0.4606	−0.4438	−0.2661	−1.2058	−1.1935	−0.4551
0.1111	0.1124	0.0511	0.4000	0.4039	0.0395
−0.0160	−0.0307	−0.0180	0.0418	0.0250	0.1052
0.0014	0.0140	0.0127	0.0052	0.0161	0.0452

[‡]It is better than the original Burg method.

3. NONLINEAR ALTERNATIVE ESTIMATORS

We are going to present several alternative ways of estimating the autocorrelation function.

Method	Intermediate functions	Autocorrelation $\lvert k\rvert=0,1,2,\ldots,N$
Polarity coincidence	$A(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \sin\left[x_i . x_{i+\lvert k\rvert}\right]$ sgn(.) is the signum function	$R(k) = \sin[A(k).\pi/2]$
hybrid sign (HS)	$A(k)=\frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \sin[x_i . x_{i+\lvert k\rvert}].\lvert x_i\rvert$ $B(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \lvert x_i\rvert$	$R(k) = \frac{A(k)}{B(k)}$
modified hybrid sign (MHS)	$A(k)=\frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \sin[x_i . x_{i+\lvert k\rvert}].\lvert x_i\rvert$ $B(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \{\lvert x_i\rvert + \lvert x_{i+k}\rvert\}$	$R(k) = \frac{A(k)}{B(k)}$
absolute difference average (ADA)	$A(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \lvert x_i - x_{i+k}\rvert$ $B(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \{\lvert x_i\rvert + \lvert x_{i+k}\rvert\}$	$R(k) = 1 - 2\left[\frac{A(k)}{B(k)}\right]^2$
reversed ADA (RADA)	$A(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \lvert x_i - x_{i+k}\rvert$ $B(k) = \frac{1}{L-\lvert N\rvert} \sum_{i=0}^{L-N} \{\lvert x_i\rvert + \lvert x_{i+k}\rvert\}$	$R(k) = 2\left[\frac{A(k)}{B(k)}\right]^2 - 1$

Method	Intermediate functions	Autocorrelation $\|k\|=0,1,2,\ldots,N$
double ADA (DADA)	$A(k)=\frac{1}{L-\|N\|}\sum_{i=0}^{L-N}\|x_i+x_{i+k}\|$ $B(k)=\frac{1}{L-\|N\|}\sum_{i=0}^{L-N}\|x_i-x_{i+k}\|$ $C(k)=\frac{1}{L-\|N\|}\sum_{i=0}^{L-N}\{\|x_i\|+\|x_{i+k}\|\}$	$R(k)=\frac{A^2(k)B^2(k)}{C^2(k)}$
relative magnitude (RM)	$A(k)=\frac{1}{L-\|N\|}\sum_{i=0}^{L-N}\max\{\|x_i\|,\|x_{i+k}\|\}$ $B(k)=\frac{1}{L-\|N\|}\sum_{i=0}^{L-N}\min\{\|x_i\|,\|x_{i+k}\|\}.\operatorname{sgn}(x_i x_{i+k})$	$R(k)=\frac{2A^2(k)B^2(k)}{A^2(k)+B^2(k)}$

3.1. Alternative 8 – multiwindow (MW)

The first multiwindow method for spectral estimation was proposed by Thomson [14, 15]. However, it uses the prolate spheroidal wave functions as windows. These are difficult to generate. Alternatively Riedel and Sidorenko [15] proposed minimum bias multipletaper spectral estimation that uses a set of sinusoidal orthogonal windows, defined by

$$v_n^J=\sqrt{\frac{2}{L+2}}.\sin\frac{\pi jn}{L+1}\quad j=0,...,J-1. \tag{9}$$

They introduced the weights $w_j=1-(\frac{i}{J})^2, j=0,...,J-1$, and a normalization factor $W=\sum_{j=0}^{J-1}w_j$. To compute the autocorrelation estimate, the simplest and efficient procedure is:

Apply the windows to obtain a set of signals $y_n^j = v_n^j . x_n$ and compute the windowed periodograms:

$$Y^j = \left|FT\left[y_n^j\right]\right|^2 . \tag{10}$$

Obtain the multiwindow spectrum:

$$S(e^{j\omega}) = \frac{\sum\limits_{j=0}^{J-1} w_j Y^J}{W} \tag{11}$$

and the autocorrelation function by doing the inverse Fourier transform. Being the inverse of a spectrum, this autocorrelation function leads to a positive definite matrix [5].

4. Comparisons

4.1. Mean square errors for the above alternatives

To do a fair comparison of the above alternatives, we generate three types of signals, MA(10), AR(10), and ARMA(6,4) from 8 different systems. We constrained the poles and zeros of the systems to lie near the unit circle. We generated also the correct autocorrelation, $R(k)$, for $k = 0, 1, \ldots, 10$. For each estimation procedure, we picked 200 points from each realization and used the above estimators to obtain 9 estimates. The mean square errors over 100 realizations were computed and are presented in the following tables.

Table 3. Mean square errors for an MA(10) signal (200 points).

FS	PC	HS	MHS	ADA	RADA	DADA	RM	MW
0.1306	0.2507	0.1824	0.1638	0.1690	0.1585	0.1474	0.1467	0.1015
0.1254	0.2111	0.1505	0.1406	0.1421	0.1609	0.1389	0.1383	0.0934
0.1702	0.2784	0.2258	0.2039	0.2183	0.1924	0.1907	0.1864	0.1090
0.1597	0.2983	0.2088	0.1965	0.1986	0.2068	0.1857	0.1820	0.1162
0.1675	0.2664	0.1964	0.1879	0.1908	0.2065	0.1866	0.1850	0.1210
0.1332	0.2185	0.1717	0.1505	0.1583	0.1670	0.1471	0.1451	0.0739
0.1831	0.2777	0.2080	0.2087	0.2119	0.2029	0.1937	0.1911	0.1510
0.1724	0.2791	0.2067	0.1986	0.2013	0.2144	0.1908	0.1874	0.1353

Table 4. Mean square errors for an AR(10) signal (200 points).

FS	PC	HS	MHS	ADA	RADA	DADA	RM	MW
0.0674	0.1323	0.0978	0.0827	0.0939	0.0930	0.0785	0.0775	0.0440
0.2123	0.3404	0.2691	0.2588	0.2599	0.2541	0.2461	0.2370	0.2456
0.0652	0.1417	0.0959	0.0826	0.0915	0.0853	0.0758	0.0748	0.0429
0.0815	0.1507	0.1110	0.0967	0.1053	0.1019	0.0906	0.0898	0.0658
0.0753	0.1491	0.1035	0.0906	0.0988	0.0943	0.0845	0.0829	0.0641
0.0743	0.1261	0.1016	0.0866	0.0958	0.0881	0.0845	0.0832	0.3164
0.0744	0.1492	0.1081	0.0926	0.1006	0.0980	0.0860	0.0838	0.0490
0.0619	0.1194	0.0915	0.0725	0.0828	0.0755	0.0665	0.0656	0.0365

It can be seen that the results are generally worst in the MA case than in AR. This could not be expected, since we know that the autocorrelation in the MA case is of finite duration, contrarily to the AR case. The results obtained in the ARMA case are similar to the AR ones.

4.2. On the definiteness of the autocorrelatiom matrix

These results presented above point to consider the multiwindow method to be the best. However to get some insight into deep behavior in a possible use for AR parameter estimate, we constructed, for each estimate, the corresponding autocorrelatiom matrix and computed its eigenvalues. For each simulation run we counted the number of times the autocorrelation matrix was negative definite. In the following tables we show the results corresponding to the situations in Tables 3 and 4.

Table 5. Times the autocorrelation was negative definite for an MA(10) signal (200 points).

FS	PC	HS	MHS	ADA	RADA	DADA	RM	MW
0	65	20	1	7	21	1	5	0
0	86	39	10	30	55	27	26	0
0	54	9	2	8	23	4	4	0
0	84	28	7	21	41	9	11	0
1	95	60	32	63	63	43	52	0
0	40	6	1	3	12	1	1	0
1	96	57	27	55	69	49	51	0
0	82	35	12	30	47	26	17	0

Table 6. Times the autocorrelation was negative definite for an ARMA(6,4) signal (200 points).

FS	PC	HS	MHS	ADA	RADA	DADA	RM	MW
32	87	61	40	90	91	72	66	0
26	87	56	35	72	93	57	65	0
34	82	63	41	85	87	52	52	0
34	92	58	62	93	98	80	78	0
28	95	69	48	86	97	73	87	0
49	96	57	68	88	98	80	94	0
43	97	79	51	88	97	75	86	0
61	99	70	75	90	100	75	95	0

As expected the multiwindow estimator leads to a positive definite autocorrelation matrix.

4.4. Classic vs multiwindow and MBM

The previous tests showed that among the proposed alternatives, the multiwindow is the best. However to try to obtain more definite results, another set of tests were performed using the following estimators: unbiased fix summation length, biased variable summation length (BVS), ujnbiased variable summation length (UVS), multiwindow, and modified Burg method (AR) [10]. Concerning the definiteness of the autocorrelatiom matrix, we know that MW and MBM led to a positive definite autocorrelation matrix. This does not happen with the others as seen in the following table.

Table 7. Times the autocorrelation was negative definite for an MA(10), AR(1), and ARMA(6,4) signal (200 points).

FS	BVS	UVS		FS	BVS	UVS		FS	BVS	UVS
0	0	98		46	46	86		29	29	69
13	17	100		28	26	57		52	42	92
15	15	99		55	51	89		24	25	65
1	1	99		45	46	95		45	46	83
2	2	94		59	50	88		47	52	88
9	6	99		88	85	99		61	63	92
6	4	100		37	43	88		40	40	85
33	31	100		41	35	74		95	93	100

From the above results, it seems clear that, among the classic estimators, the unbiased fixed summation is preferable. To try to find and distictive behaviour between Multiwindow and MBM estimators, we performed a simulation with ARMA(8,6) systems and using only 100 points. We obtained the square errors shown in the following table.

Table 8. Mean square errors for an ARMA(8,6) signal (100 points).

FS	BVS	UVS	MW	MBM
0.0826	0.0798	0.1406	2.4564	0.0744
0.0715	0.0712	0.1180	2.3559	0.0695
0.2741	0.2677	0.3182	1.4217	0.2459
0.2058	0.2025	0.2600	1.6791	0.1891
0.0107	0.0105	0.0674	2.5880	0.2494
0.2451	0.2355	0.2759	1.6565	0.2200
0.1290	0.1257	0.1846	2.2624	0.1192
0.0212	0.0200	0.0748	2.5798	0.0210

These results show that MBM is clearly better.

5. Conclusions

We made a study of the autocorrelation estimators: the classic, several based on nonlinear transformations, the multiwindow and the modified Burg method. The simulation results showed that the nonlinear transformations based autocorrelations are worst than the classic. These are frequently better in terms of mean square error than the others, but are frequently worst in terms of positiveness of the autocorrelation matrix than the multiwindow or MBM autocorrelations.

In general the behaviour becomes worst with the reduction in the number of available points and the motion of the poles or zeros to near the unit circle. From the results, we can conclude that the best estimator is the MBM estimator.

References

[1] G.E. Box, G.H. Jenkins and G.C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3ª edition, Prentice-Hall 1994.

[2] J.P. Burg, *Maximum Entropy Spectral Analysis*, Proceedings of the 3th Meeting of the Society of Exploration Geophysicists, 1967, in Modern Spectrum Analysis, Ed. Donald Childers, IEEE Press 1978.

[3] J.P. Burg, *A New Analysis Technique for Time Series Data*, Nato Advanced Study Institute on Signal Processing with Emphasis on Underwaters Acoustics, Enschede, Holanda, l968 in Modern Spectrum Analysis, Ed. Donald Childers, IEEE Press.

[4] M.J.L. De Hoon, T.H.J.J. van der Hagen, H. Schoonewelle and H. van Dam, *Why Yule-Walker should not be used for Autoregressive Modelling*, Annals of Nuclear Energy **23** (1996), 1219–1228.

[5] T.L. McWhorter and L.L. Scharf, *Multiwindow Estimators of Correlation*, IEEE Transactions on Signal Processing **46** (2) February 1998.

[6] S.L. Marple, *A New Autoregressive Spectrum Analysis Algorithm*, IEEE Trans. on ASSP Vol. 28, Agosto 1980.

[7] S.L. Marple, *Digital Spectral Analysis with Applications*, Prentice-Hall 1987.

[8] B.J.F. Murteira, D.A. Müller e K.F. Turkman, *Análise de Sucessões Cronológicas*, McGraw-Hill 1993.

[9] A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Clifs 1975.

[10] M.D. Ortigueira and J.M. Tribolet, *Global versus local minimization in least-squares AR Spectral Estimation*, Signal Processing **7** (3) (1984), 267–281.

[11] M.D. Ortigueira, *Introduction to Fractional Linear Systems. Part 2: Discrete-time Systems*, IEE Proceedings Vision, Image and Signal Processing **147** (1) (2000), 71–78.

[12] A. Papoulis and S.U. Pilai, *Probability, Random Variables, and Stachastic Processes*, McGraw-Hill Book Company 2002.

[13] M.C. Sullivan, *Efficient Autocorrelation Estimation Using Relative Magnitudes*, IEEE Transactions On Acoustics, Speech. And Signal Processing **37** (3) (1989).

[14] D.J. Thomson, *Using Local Eigen Expansion to Estimate Spectrum in Terms of Solution of Integral Equations*, Proceedings of IEEE, Setembro 1982.

[15] D.J. Thomson, *Quadratic-inverse spectrum estimates: applications to paleoclimatology*, Philos. Trans. R. Soc. Lond. **332** (1627) (1990), p. 539.

[16] K.S. Riedel and A. Sidorenko, *Minimum bias multipletaper spectral estimation*, IEEE Trans. Signal Processing **43** (1995), 188–195.

Received 14 May 2010