Anais IX Simpósio Brasileiro de Sensoriamento Remoto, Santos, Brasil, 11-18 setembro 1998, INPE, p. 1231-1237.

Bias Correction for Covariance Parameters Estimates in Polarimetric SAR Data Models

LUCIANO B. DA SILVA[1]
KLAUS L. P. VASCONCELLOS[1]
ALEJANDRO C. FRERY[2]

[1]UFPE--Universidade Federal de Pernambuco
Depto. Estatística, Cidade Universitária, 50740 - 540 Recife, PE, Brasil
{lbs,klaus}@de.ufpe.br

[2]UFPE--Universidade Federal de Pernambuco
Depto. Informática, CP 7851, 50732 - 970 Recife, PE, Brasil
frery@di.ufpe.br

Abstract. This paper obtains an expression for the second-order bias of the maximum likelihood estimates of the parameters in the covariance matrix of the polarimetric SAR observation vectors, which are assumed to follow a complex normal distribution. The second-order biases are very simple functions of the parameters, and, consequently, the second-order bias corrected estimates are very easy to compute. In particular, we discuss the performance of the corrected estimate of the absolute value of the complex correlation coefficient, which is a crucial parameter in the practical study of polarimetric SAR data.

Keywords: Bias correction, complex correlation coefficient, complex normal distribution, maximum likelihood, polarimetric SAR.

1 Introduction

The synthetic aperture radar – SAR- is an important conquest of the modern remote sensing technology. Through this kind of radar, important characteristics of the target returns can be studied, such as phase difference information, intensity ratio of different polarisations, etc. Applications include geology, oceanography, environmental sciences, planing for the use of the land, among others.

Data captured by this sort of radar are very rich in information and increasingly sophisticated techniques have been employed to extract this information.

As many other sciences working in the field of remote sensing, statistics has been an increasingly powerful tool in the hands of specialists in the area. This area of knowledge has served to model the information acquiring process and, in particular, to help the interpretation and classification of data through specific techniques.

In this work, we depart from a commonly accepted hypothesis about the statistical properties of targets: the normality hypothesis for homogeneous areas. From this hypothesis, we consider the maximum likelihood estimation of a very important parameter, the complex correlation coefficient between two elements of the polarisation vector. Then, we calculate a bias-corrected estimator of the absolute value of this complex correlation coefficient.

The rest of this paper is structured as follows: Section 2 sets up a complex normal model for the polarimetric data we are working with, in Section 3 we obtain the second-order bias of the MLE (maximum likelihood estimate) of the covariance parameters in this model, and, finally, in Section 4, the main conclusions are summarised.

2 A Model for Polarimetric Data

A polarimetric radar is a type of radar which works in all possible polarisations. In other words, it emits and receives signals in parallel polarisation (HH,VV) or crossed polarisation (HV,VH) and data captured by its sensors form the complete polarimetric matrix, defined as follows:

$$\begin{pmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{pmatrix} \qquad (1)$$

where S_{kj} is the signal that is emitted in polarisation k and received in polarisation j, where $k,j \in \{H,V\}$. Each element of this matrix is modelled as a complex random variable, i.e., $S_{kj} = X_{kj} + iY_{kj}$, where $i = \sqrt{-1}$.

In general it is assumed that $\mathrm{Cov}(S_{kk}, S_{kj}) = 0$, and this hypothesis has been used for the calibration of radar sensors.

This work is restricted to monostatic SAR, where the signal is emitted and received with the same antenna configuration. In this case, the matrix presented in eq. (1) is somewhat redundant. Since for this kind of sensors it holds that $S_{HV} = S_{VH}$, it is possible to use a more compact scattering matrix of the form

$$\begin{pmatrix} S_1 \\ S_2 \\ S_3 \end{pmatrix} \qquad (2)$$

where, for the sake of simplicity, the following notation is used $S_{HH} = S_1$, $S_{VV} = S_2$ and $S_{HV} = S_3$.

Following Lee at al., 1994 let us admit that the polarimetric vector presented in eq. (2) can be modelled as a complex normal random vector. A complex random vector U is said to be complex normal if its real and imaginary parts, denoted by X and Y respectively, have bivariate normal joint distribution. In other words, $(X,Y) \sim N_2(m,C)$ where m and C are, respectively, the vector of means and the covariance matrix. Therefore, the density of U is given by (Wooding, 1956)

$$f_U(u) = \frac{1}{p^p\,|C|}\exp\left(-u^* C^{-1} u\right) \qquad (3)$$

where u^* is the transpose complex conjugate of u.

In this work a special case of the complex normal distribution is studied, one that is relevant for polarimetric SAR data modelling. This model, which consists of a restriction in the structure of the covariance matrix (Goodman, 1963; Lee et al., 1994) is defined by the conditions below. Let $S_k = X_k + iY_k$ be any complex component of the vector given in eq. (2), then

i) $\mathrm{E}(X_k) = \mathrm{E}(Y_k) = 0, \forall k$;

ii) $\mathrm{E}(X_k^2) = \mathrm{E}(Y_k^2), \forall k$;

iii) $\mathrm{E}(X_k Y_k) = 0, \forall k \neq j$;
iv) $\mathrm{E}(X_k X_j) = \mathrm{E}(Y_k Y_j), \forall k, j$;
v) $\mathrm{E}(X_k Y_j) = -\mathrm{E}(Y_k X_j), \forall k, j$.

Under the aforementioned restrictions, it is said that the complex random vector defined in (2) obeys a circular complex normal distribution.

For purposes of inference, we remind that data will be always available in one of the following formats: intensity pair, intensity-phase pair, phase difference or intensity ratio. In any of these formats, we can only recover information on two of the three components of the random vector defined in (2). For this reason, we can assume that our observations are two-dimensional complex random vectors, constituting a subset of (2). Our main goal is to obtain, from these two dimensional observations, the MLE of the relevant parameters in the above model, together with their second-order bias correction, as explained in the next section.

3 Second order bias correction

It is well known that MLEs are, in general, biased estimators of the true parameter values. This bias can be ignored in many practical situations, since it is typically of order $O(n^{-1})$, while the asymptotic standard deviation of the estimator has order $O(n^{-1/2})$, where *n* is the sample size. However, for small values of *n*, or, more generally, when there is little information, bias can constitute a problem and correction efforts are worthy.

Let $\hat{q}$ be the MLE of the parameter q. It can be proved that, under very mild conditions, it holds that

$$\mathrm{E}(\hat{q}) = q + B(q) + O(n^{-2})$$

where $B(q)$ is of order $O(n^{-1})$ and it is called "the second order bias" of $\hat{q}$. It can be shown that the bias of $\tilde{q} = \hat{q} - B(\hat{q})$ has order $O(n^{-2})$. Therefore, the estimator $\tilde{q}$ constitutes a bias correction of $\hat{q}$.

We want to derive the second-order bias of the MLEs of the parameters in the covariance structure of the vector presented in (2), where these MLEs have to be obtained from observing a two dimensional sub-vector of (2). In particular, we are interested in obtaining a corrected estimate for the absolute value $|r_c|$ of the complex correlation coefficient. This quantity measures the degree of linear relationship between two complex random variables, i.e., if A and B are two complex random variables with zero mean, then $|r_c| = 1$ iff there is a complex constant z such that $B = zA$.

Following Lee et al. (1994), this complex correlation coefficient $|r_c|$ between any pair of elements S_k and S_j of the polarimetric vector is defined as

$$r_c = \frac{ES_k S_j^*}{\sqrt{E|S_k|^2 \, E|S_j|^2}} = |r_c| e^{iq} \qquad (4)$$

Since $|r_c|$ is directly involved in the densities of the models used for multilook amplitude and intensity SAR data, this is the quantity to be addressed here.

Let S be a two-dimensional sub-vector of U, say $S^T = (S_k, S_j)$ where $S_p = X_p + iY_p$, for $p = k, j$ and $j, k = 1,2,3$. In this case the complex covariance matrix associated to S is defined as

$$\mathrm{E}(SS^*) = \begin{pmatrix} \mathrm{E}S_k S_k^* & \mathrm{E}S_k S_j^* \\ \mathrm{E}S_j S_k^* & \mathrm{E}S_j S_j^* \end{pmatrix} \qquad (5)$$

Since any complex number can be represented as a two-dimensional vector, we can consider a four-dimensional real covariance matrix corresponding to (5).

Let $s_p^2 = \mathrm{E}S_p S_p^* = \mathrm{E}\,|S_p|^2$. Applying restriction (ii) of Section 2, we obtain

$$s_p^2 = \mathrm{E}X_p^2 + \mathrm{E}Y_p^2 = 2\mathrm{E}X_p^2 \Rightarrow \mathrm{E}X_p^2 = \frac{s_p^2}{2} \qquad (6)$$

The off-diagonal terms in (5) can be written as

$$\mathrm{E}S_k S_j^* = \mathrm{E}(X_k + iY_k)(X_j - iY_j) = \mathrm{E}(X_k X_j + Y_k Y_j) + i\mathrm{E}(Y_k X_j - X_k Y_j) \qquad (7)$$

and, from restrictions (iv) and (v) in the last section, we get

$$\mathrm{E}S_k S_j^* = 2\mathrm{E}X_k X_j - i2\mathrm{E}X_k Y_j \qquad (8)$$

Also, it can be readily seen from equation (4) that

$$\mathrm{E}S_k S_j^* = |r_c|(\cos(q) + i\,\mathrm{sen}(q))s_k s_j \qquad (9)$$

and, equating real and imaginary parts in (8) and (9), we get

$$\mathrm{E}X_k X_j = \frac{|r_c|\cos(q)s_k s_j}{2} \quad \text{and} \quad \mathrm{E}X_k Y_j = \frac{-|r_c|\,\mathrm{sen}(q)s_k s_j}{2} \qquad (10)$$

Let’s consider the four-dimensional real random vector Z, containing the real and imaginary parts of the components of S. Hence, we have

$$Z = \begin{pmatrix} X_k \\ Y_k \\ X_j \\ Y_j \end{pmatrix}$$

and the covariance matrix of Z is given by

$$\Sigma = \mathrm{E}ZZ^T = \begin{pmatrix} \mathrm{E}X_k^2 & \mathrm{E}X_k Y_k & \mathrm{E}X_k X_j & \mathrm{E}X_k Y_j \\ \mathrm{E}Y_k X_k & \mathrm{E}Y_k^2 & \mathrm{E}Y_k X_j & \mathrm{E}Y_k Y_j \\ \mathrm{E}X_j X_k & \mathrm{E}X_j Y_k & \mathrm{E}X_j^2 & \mathrm{E}X_j Y_j \\ \mathrm{E}Y_j X_k & \mathrm{E}Y_j Y_k & \mathrm{E}Y_j X_j & \mathrm{E}Y_j^2 \end{pmatrix} \qquad (11)$$

Applying restrictions (i) to (v) of Section 2 and the results in (6) and (10), we arrive at

$$\Sigma = \frac{1}{2}\begin{pmatrix} s_k^2 & 0 & |r_c|\cos(q)s_k s_j & -|r_c|\,\mathrm{sen}(q)s_k s_j \\ 0 & s_k^2 & |r_c|\,\mathrm{sen}(q)s_k s_j & |r_c|\cos(q)s_k s_j \\ |r_c|\cos(q)s_k s_j & |r_c|\,\mathrm{sen}(q)s_k s_j & s_j^2 & 0 \\ -|r_c|\,\mathrm{sen}(q)s_k s_j & |r_c|\cos(q)s_k s_j & 0 & s_j^2 \end{pmatrix} \quad (12)$$

The quantities in (12) are unknown and will, in practice, be estimated. Let $s^T = (s_k, s_j, |r_c|, q)$ be the vector of the parameters in (12). Then, the total log-likelihood for this vector of unknown parameters, given the *n* observable data $s^T = (s_k, s_j, |r_c|, q)$ is

$$s^T = (s_k, s_j, |r_c|, q)$$

and the maximum likelihood estimates are obtained by solving a system of four equations:

$$\sum_i z_i^T \Sigma^{-1} \Sigma_k \Sigma^{-1} z_i = n\, tr(\Sigma^{-1}\Sigma_k); \quad k = 1,\ldots,4 \qquad (13)$$

where Σ_k is a matrix representing the derivative of the covariance matrix with respect to the *k*th element of s and *tr* stands for the trace of the matrix. The equations in (13) are, in general, non-linear, but can be solved by Fisher's scoring method or any iterative re-weighted least squares.

We proceed to the calculation of the second-order bias of the parameters. From Cordeiro and Vasconcellos (1997), the vector $B(s)$ containing the respective second-order biases of the components of the parameter vectors can be calculated as

$$B(s) = (V^T \underset{\sim}{\Sigma}^{-1} V)^{-1} V^T \underset{\sim}{\Sigma}^{-1} x \qquad (14)$$

where

$$x = -\frac{1}{n} W vec[(V^t \underset{\sim}{\Sigma}^{-1} V)^{-1}]$$
$$W = (vec(\Sigma_{11}), vec(\Sigma_{12}), \ldots, vec(\Sigma_{43}), vec(\Sigma_{44}))$$
$$V = (vec(\Sigma_1), \ldots, vec(\Sigma_4))$$
$$\underset{\sim}{\Sigma} = \Sigma \otimes \Sigma$$

with Σ_{kj} being the second derivative of the covariance matrix with respect to the k-th and j-th components of the vector s, for $k, j = 1, \ldots, 4$, $\otimes$ is the Kronecker product and vec is the vec operator, which transforms a matrix into a vector by stacking the columns of the matrix one underneath the other.

After some manipulations, we obtain

$$B^T(s) = \left(-\frac{s_k}{8n}, -\frac{s_j}{8n}, \frac{(1 - |r_c|^2)^2}{4n|r_c|}, 0 \right) \qquad (15)$$

where each element of the vector *B* is the second-order bias of the corresponding element of the vector s.

We define the corrected estimate of the absolute value of the complex correlation coefficient as:

$$|\tilde{r}_c| = |\hat{r}_c| - B(|\hat{r}_c|) \Rightarrow$$

$$|\tilde{r}_c| = |\hat{r}_c| - |\hat{r}_c|\left(\frac{(1-|\hat{r}_c|^2)^2}{4n|\hat{r}_c|^2}\right) \Rightarrow$$

$$|\tilde{r}_c| = |\hat{r}_c|\left(1-\left(\frac{(1-|\hat{r}_c|^2)^2}{4n|\hat{r}_c|^2}\right)\right) \qquad (16)$$

On the graphs below, we present the value of the corrected MLE against the uncorrected one, for the sample sizes n = 40, 60, 100, 500. It can be observed that:

i) $|\hat{r}_c| \approx |\tilde{r}_c|$ when n is large (as expected)

ii) the second-order bias correction leads to a reasonable estimator only if the maximum likelihood estimator is greater than $\sqrt{n+1}-\sqrt{n}$ (after this value, the corrected estimator becomes negative) and smaller than 0.5. After this point, the bias correction becomes unnecessary for almost all sample sizes.

Also, we observe that the second-order bias correction can be an important tool to calibrate the image system, when we have little information in our sample or when the sample size is small.

4 Conclusions

We have calculated the second-order bias of the MLE of the complex correlation coefficient, assuming data has a complex normal distribution. From this second-order bias, we obtained a new estimator, corrected up to the second order. This correction is important, since, for small sample sizes, the bias correction can bring our estimates towards the true parameter value. However, the corrected estimator has some deficiencies.

Acknowledgements

This work was partially developed with resources from CNPq (351612/97-1 and 300458/93-2) and FACEPE (APQ-0707-1.03/97).

References

Cordeiro, G.M. and Vasconcellos, K.L.P., (1997), Bias correction for a class of multivariate nonlinear regression models, *Statistics and Probab. Lett.*, 35:155–164.

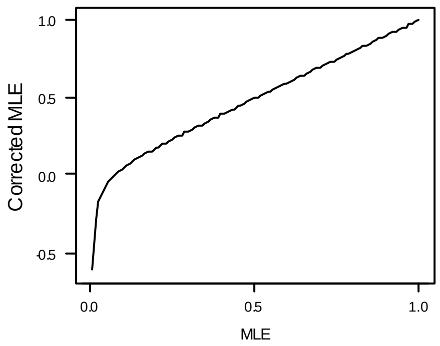
Goodman, N.R., (1963), Statistical analysis based on a certain multivariate complex gaussian distribution (na introduction), *Ann. Of Mathemat. Statist*, 34:152–177.

Lee, J.S., Hoppel, K.W., Mango, S.A. e Miller, A.R., (1994), Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery, *IEEE Trans. Geoci. Remote Sensing*, 32:1017–1028.

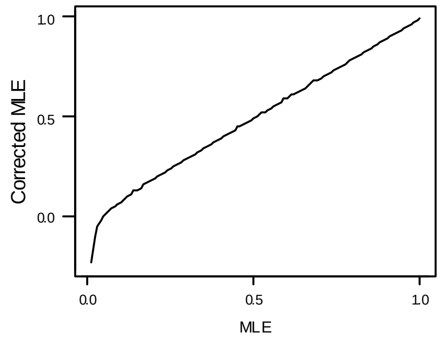
TCI Software Research (New Mexico, USA) Scientific Workplace V. 2.50, 1996, Software set.

Wooding, R.A., (1956), The multivariate distribution of complex normal variables, *Biometrica,* 43:212–215.

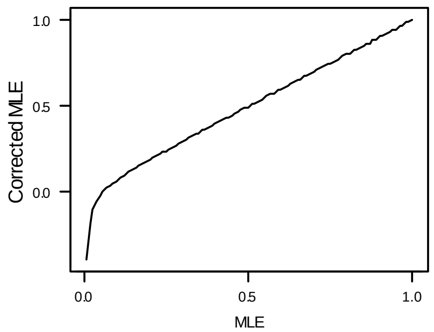
Corrected MLE x MLE of the Abs. Value of Complex Corr. Coef.
n = 40



Corrected MLE x MLE of the Abs. Value of Complex Corr. Coef.
n = 100



Corrected MLE x MLE of the Abs. Value of Complex Corr. Coef.
n = 60



Corrected MLE x MLE of the Abs. Value of Complex Corr. Coef.
n = 500