Rosenblatt's Contributions to Deconvolution

By Keh-Shin Lii

A simple model of deconvolution can be described as observing $\{x(t)\}$ which is a convolution of a signal $\{s(t)\}$ with a filter $\{f(j)\}$, $x = s * f$. More specifically, we have

$$x(t) = \sum_{j=-\infty}^{\infty} f(j)s(t-j).$$

The problem of deconvolution is to recover $\{s(t)\}$ based on the output process $\{x(t)\}$. If the filter $\{f(j)\}$ is known then the problem is fairly straightforward. The blind deconvolution, in signal processing terminology, is to recover $\{s(t)\}$ based solely on $\{x(t)\}$ without knowing $\{f(j)\}$. Statisticians may be more interested in the estimation of $\{f(j)\}$ under certain conditions on $\{s(t)\}$ and $\{f(j)\}$. This problem and its many variations have very broad applications in signal processing, image restoration, geo-exploration, seismology, radio astronomy among others [11, 33, 38, 39].

Assume that the signal random variables $\{s(t)\}$ are independent and identically distributed with mean 0 and variance 1. Let the filter $\{f(j)\}$ be a sequence of real constants such that

$$\sum_{j=-\infty}^{\infty} f^2(j) < \infty$$

and $f(z) = \sum_j f(j)z^j$ be the z-transform corresponding to the process $\{x(t)\}$. Then

$$f(e^{-i\lambda}) = \sum_j f(j)e^{-ij\lambda} = |f(e^{-i\lambda})| \exp\{ih(\lambda)\}$$

is the frequency response function or the transfer function where $h(\lambda)$ is the phase function of the transfer function. If we know $f(e^{-i\lambda})$ for all $\lambda \in [0, 2\pi]$, then we can obtain $\{f(j)\}$ for all j. The modulus $f(j)$, $|f(j)|$, of the frequency response function can be obtained from

R.A. Davis et al. (eds.), *Selected Works of Murray Rosenblatt*, Selected Works in Probability and Statistics,
DOI 10.1007/978-1-4419-8339-8_4, © Springer Science+Business Media, LLC 2011

the spectral density of $\{x(t)\}$ which is

$$g(\lambda) = \frac{1}{2\pi}|f(e^{-i\lambda})|^2.$$

It is clear that if the $\{x(t)\}$ process or the random signal process $\{s(t)\}$ is Gaussian then the full probability structure of $\{x(t)\}$ is determined by $g(\lambda)$ or equivalently by $|f(e^{-i\lambda})|$ which is determined by the second order covariance property of the process. The phase information $h(\lambda)$ of $f(e^{-i\lambda})$ is not identifiable in the Gaussian case. Any hope of getting information on the phase function will require the process to be nonGaussian.

Murray Rosenblatt's interest and insight in this problem may have stemmed from his interest in the higher order spectra, especially higher-order cumulant spectra which is fundamental in dealing with nonGaussian processes [34, 4, 5, 19]. The kth-order cumulant spectral density of $\{x(t)\}$ is given by

$$b_k(\lambda_1, \lambda_2, \cdots, \lambda_{k-1}) = \frac{\gamma_k}{(2\pi)^{k-1}} f(e^{-i\lambda_1}) f(e^{-i\lambda_2}) \cdots f(e^{-i\lambda_{k-1}}) f(e^{i(\lambda_1+\lambda_2+\cdots+\lambda_{k-1})}),$$

where γ_k is the kth order cumulant of $\{s(t)\}$. Therefore the phase of the kth order spectrum $b_k(\lambda_1, \lambda_2, \cdots, \lambda_{k-1})$ is related to the phase $h(\lambda)$ of the the transfer function $f(e^{i\lambda})$.

In the paper [35], Murray laid out the basic idea on how the phase function $h(\lambda)$ can be identified up to a linear shift of $c\lambda$ using the phase of the bispectra of the process. It was noted that the same result holds by using any kth-order spectra with $k > 2$. A similar but different idea of using bispectra to obtain phase information of the transfer function is outlined in Brillinger [6]. Of course for different c's we will get different transfer functions and the deconvolution can not be realized.

The problem of indeterminancy of the linear shift of the phase was solved in [20], where it was shown that the phase function can be identified up to an integer shift $\tau\lambda$ with τ an integer. The basic idea is that the transfer function $f(e^{-i\lambda})$ has to be real when $\lambda = \pi$ and hence its phase at π has to be $\tau\pi$ with τ an integer. Since this integer phase shift corresponds to reindexing the noise sequence $\{s(t)\}$, the deconvolution problem is essentially solved in this case. If the linear process $\{x(t)\}$ is a finite parameter ARMA(p,q) process then there are other methods available to identify the phase function. In [20] a few other ideas were laid out to identify the correct ARMA model in terms of the location of the roots of the characteristic functions of the ARMA model. Again, the basic idea here is to match various higher order moments or cumulants of the process with the correct coeffients or roots.

The paper [20] seemed to have generated quite a bit interest in the signal proceesing community. The importance of the phase in signals has been noted in signal and image processing and geoexploration problems [31, 33, 38, 39]. Many papers were published using higher-order statistics and higher-order spectral analysis to analyze signals, images and systems. These activities resulted in the first 'Workshop on Higher-Order Spectral Analysis' held at Vail, Colorado in 1989 with a Proceedings of the Workshop and a 'Special Section on Higher Order Spectral' that appeared in the July 1990 issue of the IEEE Transactions on Acoustics, Speech and Signal Processing. Following the first workshop, four more 'workshops on higher-order statistics' were held in odd years with the location alternating betwwen Europe and North America until the last one in 1997. The last two Proceedings published in the 1995 and 1997 Workshops have more than 450 pages of articles each. The 1995 workshop resulted in IEEE Signal Processing September 1996 issue being the 'special issue on higher-order statistics'. A book [32] was published in 1993 on these topics.

If the output process $\{x(t)\}$ is observed with an independent Gaussian noise $\eta(t)$ it is shown in [21] that we can still estimate the filter $\{f(j)\}$ up to an unknown scale factor c. If the additive Gaussian noise $\eta(t)$ is white then we can estimate the filter consistently up to a time shift. A detailed discussion on the use of fourth-order cumulants to estimate the filter function for deconvolution is given in [22], where a discussion on the sample size relative to various orders of cumulants of $\{s(t)\}$ is given. In [23] it is demonstrated with a well-log data and water-gun signature that the deconvolution can be more effective using the Fourier transform of the tapered sample higher-order cumulants to obtain higher-order spectrum instead of using the smoothed higher-order periodogram. For this reason, theoretical properties of such higher-order spectral estimates are given in [25, 26]. The particular formula using the phase of a higher-order spectrum to obtain the phase of the transfer function in [20] utilized only a subset of the full higher-order spectral phase function. Another method which utilized the whole effective higher-order spectral phase function is given in [28] with better asymptotic convergence properties as a function of sample size. Generlization of the previous results to the deconvolution problem in autoregressive random fields is given in [30, 37].

In the deconvolution problem described above it is generally assumed that the spectral density of the process is strictly positive. In [24] our attention turned to the blind deconvolution problem when the transfer function has zeros, which do occur in geophysical investigations [31]. In [24] it is shown that if the zeros are finitely many and are of finite order then the transfer function can still be consistently estimated without the *minimum phase* assumption when the process is nonGaussian. A procedure is given so that the deconvolution can be effectively carried out. If the transfer function is zero in an interval for $\lambda \in (a, b)$ with $0 < a < b < \pi$, then the smallest mean square error in deconvolution that can be achieved is $(b - a)/2\pi$ even if we assume that the phase can be estimated. It was also shown that the phase is not identifiable if the transfer function is zero in an interval around π with positive length i.e., the process is band-limited. However if all higher order cumulants of $\{x(t)\}$ are available then the phase is identifiable up to a linear phase shit $\delta\lambda$ with δ a real number.

In the previous discussion the process is a general linear process and the approach to the identification of the phase information is the use of higher-order cumulant spectra which in general, is 'nonparametric' in the sense that the probability distribution of the signal process

$\{s(t)\}$ is not used explicitly other than that it is nonGaussian and that a certain kth order cumulant exists and is non-zero. Now if the process $\{x(t)\}$ can be represented by a finite parameter ARMA(p,q) process, then as noted before certain methods based on moments can be used to identify the phase information and in this case it is equivalent to finding the locations of the roots of the charateristic functions associated with the AR and MA parts of the process or eqivalently the coefficients of the ARMA model without the usual causal or invertible conditions under the Gaussian assumption. Maximum likelihood estimation of the parameters of an ARMA model under the Gaussian assumption has been discussed widely in the literature [9, 36]. Of course in this case the phase information is not available and the ARMA model is assumed to be causal and invertible. In [7] it is shown that given a nonGaussian probability density for the independent and identically distributed process $\{s(t)\}$, the stationary AR(p) process

$$x(t) = \phi_1 x(t-1) + \ldots + \phi_p x(t-p) + s(t)$$

is identifiable through a maximum likelihood procedure whether the process is causal or not. The idea is to reparametrize the model by decomposing the autoregressive polynomial into its causal and purely non-causal components and then analyzing the corresponding AR processes that resulted from this decomposition. The likelihood function can be approximated using these processes and the estimates of the parameters of the possibly non-causal AR process are the solutions to the likelihood equations. Similar results for possibly non-invertible MA(q) process are given in [27] and for general ARMA process in [29]. These results give possibly 'efficient' methods for blind deconvolution. If the nonGaussian probability distribution of the $\{s(t)\}$ process is unknown then it is demonstrated in the previous papers that a quasi-likelihood method can be used to estimate the parameters of the ARMA process by assuming that the unknown probability density is Laplace (two-sided exponentail) which leads to a least absolute deviation criterion. A modified least absolute deviation method is shown to be consistent when the input process in the AR(p) model has a stable law distribution with index $\alpha \in (1, 2)$ [13].

Without the causal and invertible conditions for nonGaussian ARMA processes one is naturally led to an interesting subclass of processes called all-pass ARMA models. An ARMA model is all-pass if every root of the AR polynomial $\phi(z)$ is a reciprocal of a root of the MA polynomial $\theta(z)$ matched with multiplicity and vice versa. In such a case the corresponding filter is $f(e^{-i\lambda}) = \phi(e^{-i\lambda})/\theta(e^{-i\lambda})$ and $|f(e^{-i\lambda})|$ is a constant. This means that whatever the spectum of the input process $\{s(t)\}$, it will pass the ARMA filter unchanged except for a multiplicative constant, hence the name all-pass. So the spectral density of the process $\{x(t)\}$

$$g(\lambda) = \frac{1}{2\pi}|\frac{\phi(e^{-i\lambda})}{\theta(e^{-i\lambda})}|^2$$

is a constant. This means that the process $\{x(t)\}$ is uncorrelated or white but not independent if the input independent process is nonGaussian. Processes which are second order uncorrelated or white but with higher order dependence occur often in financial data. The estimation of parameters of such all-pass ARMA models using least absolute deviations is given in [8], using maximum likelihood in [1] and using rank based procedures in [2]. The rank based method can have the same asymtotic efficiency as maximum likelihood estimators and are robust to some distributional assumptions. There are deconvolution problems in signal processsing when the probability distribution of the signal process $\{s(t)\}$ is discrete with finitely many points of support such as in the finite alphabet transmission. These deconvolution problems were treated in [17, 12]. The finite tone image deconvolution or debluriing problems were treated in [18] where the distribution of the pixels is two-tone (or finite-tone) without the stationarity assumption. Methods using the maximization of the standardized cumulant of the deconvolved process to estimate the filter are given in [10].

Murray's interest in higher-order spectra began more than four decades ago [34], which ultimately led to the deconvolution problem [35]. Murray's fundamental contributions in both areas have had a long lasting impact on many aspects of statistical probems and applications as described in this brief article. His influence in the area of blind or noncausal deconvolution is still ongoing [3], and has expanded to many related problems in economics [15], in medicine [14], and in signal processing [16,40], among others.

References

[1] Andrews, B., Davis, R.A., and Breidt, F. J. (2006). "Maximum Likelihood Estimation for All-Pass Time Series Models". J. Multivariate Analysis. 97, 1638-1659

[2] Andrews, B., Davis, R.A., and Breidt, F. J. (2007). "Rank Estimation for All-Pass Time Series Models". Annals of Statistics, 35, 844-869.

[3] Andrews, B., Calder, M. and Davis, R.A. (2009). "Maximum likelihood estimation for α-stable autoregressive processes". Annals of Statistics, 37, 1946-1982.

[4] Brillinger, D. R., Rosenblatt, M. (1967a) "Asymptotic theory of k-th order spectra." In *Spectral Analysis of Time Series*, Ed. B. Harris, 153-188. New York, Wiley.

[5] Brillinger, D. R., Rosenblatt, M. (1967b) "Computation and interpretation of k-th order spectra." In *Spectral Analysis of Time Series*, Ed. B. Harris, 189-232. New York, Wiley.

[6] Brillinger, D. (1977). " The Identification of a Particular Nonlinear Time Series System," Biometrika, 64, 509-515.

[7] Breidt, F. J., Davis, R. A., Lii, K. S., and Rosenblatt, M. (1991). "Maximum Likelihood Estimation for Noncausal Autoregressive Processes", J. Multivariate Analysis, 36, 175-198.

[8] Breidt, F.J., Davis, R.A., and Trindade, A. (2001). "Least Absolute Deviation Estimation for All-Pass Time Series Models". Annals of Statistics 29, 919-946

[9] Brockwell. P. J., and Davis, R. A. (1987). Time Series: Theory and Methods. Springer-Verlag, New York.

[10] Cheng, Q. (1990). "Maximum standardized cumulant deconvolution of non-Gaussian linear processes", Annals of Statistics, 18, 1774-1783.

[11] Donoho, D. (1981). "On minimum entropy deconvolution", In Applied Time Series Analysis, II. Ed. by Findly, D. F. 565-608. Academic Press New York.

[12] Gamboa,F., and Gassiat, E. (1996). "Blind deconvolution of discrete linear systems", Annals of Statistics, 24, 1964-1981.

[13] Jian, H., and Pawitan, Y. (1999). "Consistent estimation for non-Gaussian non-causal autoregressive processes", J. of Time Series Analysis, 20, 417-423.

[14] Kannnan, G,, Milani, A. A., Panahi, I., and Briggs, R. (2008). "Equalizing secondary path effects using the periodicity of fMRI accoustic noise", 30th Annual Intrernational IEEE EMBS Conference, Cancouver, British Columbia, Canada, August 20-24, 2008, 25-28.

[15] Lanne, M., and Saikkonen, P. (2009). "Noncausal vector autoregression", Bank of Finland Research, Discussion papapers 18, 2009.

[16] Larue, A., Mars, J. I., and Jutten, C. (2006). "Frequency-domain blind deconvolution based on mutual information rate", IEEE Transactions on Signal Processing, 54, 1771-1781.

[17] Li, T. H. (1995). "Blind deconvolution of linear systems with multilevel stationary inputs", Annals of Statistics, 23 690-704.

[18] Li, T. H., and Lii, K. S. (2002). "A joint estimation approach for tw-tone image deblurring by blind deconvolution", IEEE Tansactions on Image Processing, 11, 847-858.

[19] Lii, K. S., Rosenblatt, M., and Van Atta, C. (1976). "Bispectra measurements in turbulance", Part 1, J. Fluid Mechanics, 77, 45-62.

[20] Lii, K. S., and Rosenblatt, M. (1982)."Deconvolution and estimation of the transfer function phase and coefficients for non-Gaussian linear processes," Ann. Statist., 10, 1195-1208.

[21] Lii, K. S., and Rosenblatt, M. (1984). "Remarks on nonGaussian linear processes with additive Gaussian noise". Lecture Notes in Statistics, V. 26, 185-197. In "Robust and Nonlinear Time Series Analysis" Ed. by Franke, J., Hardle, W., and Martin, D.

[22] Lii, K. S., and Rosenblatt, M. (1985). "A fourth order deconvolution technique for nonGaussian linear processes". In Krishnaiah, P. R. (eds), Multivariate Analysis VI, 395-410.

[23] Lii, K. S., and Rosenblatt, M. (1988). "Nonminimum phase non-Gaussian deconvolution", J. Multivariate Analysis, 27, 359-374.

[24] Lii, K. S., and Rosenblatt, M. (1988)."Estimation and Deconvolution When the Transfer Function Has Zeros," J. Theoretical Probability, 93-113.

[25] Lii, K. S., and Rosenblatt, M. (1990)."Cumulant spectral estimates: Bias and covariance," Proc. Third Hungarian Colloquium on Limit Theorems in Probability and Statistics (held at PECS), 365-405.

[26] Lii, K. S., and Rosenblatt, M. (1990)."Asymptotic normality of cumulant spectral estimates," J. Theoretical Probability, 3, 367-385.

[27] Lii, K. S., and Rosenblatt, M. (1992). "An Approximate Maximum Likelihood Estimation for Non-Gaussian Non-Minimum Phase Moving Average processes", J. Multivariate Analysis, 43, 272-299.

[28] Lii, K. S., and Rosenblatt, M. (1993). "Bispectra and phase of nonGaussian linear processes", J. Theoretcal Probability, 6, 579-593.

[29] Lii, K. S., and Rosenblatt, M. (1996). "Maximum Likelihood Estimation for NoGaussian NonMinimum Phase ARMA Sequences", Statistica Sinica, 6, 1-22.

[30] Lii, K. S., and Rosenblatt, M. (1996)."NonGaussian autoregressive sequences and random fields". In Adler, R. J., Muller, P., and Rozovskii, B. (Eds). Stochastic Modeling in Physical Oceanography, 295-309.

[31] Matsuoka, T., and Ulrych, T. J. (1984). "Phase Estimation Using Bispectrum", Proc. IEEE, 72, 1403-1411.

[32] Nikias, C. L., and Petropulu, A. P. (1993) Higher-order Spectra Analysis, a Nonlinear Signal Process Framework, Prentice Hall, Englewood Cliffs, NJ.

[33] Oppenheim, A. V., and Lim, J. S. (1981). " The importance of phase in Signals," Proc. IEEE, 69, 529-541.

[34] Rosenblatt, M., and Van Ness, J. W. (1965). "Estimation of bispectra", Annals of Mathematical Statistics, 36, 1120-1136.

[35] Rosenblatt, M. (1980). "Linear processes and Bispectra", J. Appl. Probab., 17, 265-270.

[36] Rosenblatt, M. (1985). Stationary Sequences and Random Fields. Birkhauser, Boston.

[37] Rosenblatt, M. (2000). Gaussian and Non-Gausian Linear Time Series and Random Fields. Springer, New York.

[38] Scargle, J. D. (1981). "Phase-sensitive deconvolution to model random processes with special reference to astronomical data". In Applied Time Series Analysis, II. Ed. by Findly, D. F. 549-564. Academic Press New York.

[39] Wiggins, R. A. (1978). "Minimum entropy deconvolution", Geoexploration 16, 21-35.

[40] Xia, B., and Zhang, L. (2007). " Blind deconvolution in nonminimum phase systems using cascade structure", EURASIP J. on Advances in Signal Processing, V. 2007, Article ID 48432, 10 pages.