A General Framework for Learning in the Complex Domain and its Applications*

Tülay Adalı and Hualiang Li
University of Maryland Baltimore County, Baltimore, Maryland 21250
{adali,lihua1}@umbc.edu

Abstract

We present a framework that greatly simplifies the evaluations and analyses for optimization in the complex plane through the use of a generalized definition of analyticity. We show how the gradient, relative (natural) gradient, Newton, and Newton variation updates can be easily derived within the framework. We demonstrate application of the framework for training of a multi-layer perceptron network and for performing complex-valued independent component analysis.

Summary

Complex-valued signals arise frequently in applications as diverse as communications, radar, and biomedicine, as most practical modulation formats are of complex type and applications such as radar and magnetic resonance imaging lead to data that are inherently complex valued. When the processing has to be done in a transform domain such as Fourier or complex wavelet, again the data are complex valued. The complex domain not only provides a convenient representation for these signals but also a natural way to preserve the physical characteristics of the signals and the transformations they go though, such as the phase and magnitude distortions a communications signal experiences. In this case, the processing also needs to be carried out in the complex domain in such a way that the complete information—represented by the interrelationship of the real and imaginary parts or the magnitude and phase of the signal—can be fully utilized. In this context, both the approximation (choice of filter structure) and the optimization (algorithm derivation) problems have to be carefully addressed. The complex domain, however, presents unique challenges for each, and as a result, most of the algorithms derived for the complex domain have taken shortcuts limiting their usefulness.

In terms of the filter structures utilized, a strictly linear filter has been the dominant filter type—rather than a widely-linear filter [1]—and for the nonlinear case, boundedness has been emphasized and “split”-type functions that process the real and imaginary parts (or the magnitude and phase) of the signal separately have been most commonly used, see *e.g.* [2], [3], [4]. Not surprisingly, both approaches fail to fully utilize the information in the real and imaginary channels and lead to suboptimal solutions, especially when the signals are noncircular [5], [6], [7], *i.e.*, have probability density functions (pdfs) that are *not* rotation-invariant. However, a great number of signals of interest, such as communications signals and magnetic resonance imaging data are noncircular [8], [9]. Another difficulty arises as the derivation of algorithms involves optimization of a cost function, which is typically real valued and hence non-analytic in the complex domain. The problem is more pronounced when these cost functions involve split type nonlinear functions, which themselves are also non-analytic. In the derivations, either optimization is carried out using gradients defined by combining the

*This work was supported by the NSF grants NSF-CCF 0635129 and NSF-IIS 0612076.

gradients of real and imaginary parts, or by defining appropriate transformations $\mathbb{C}^n \mapsto \mathbb{R}^{2n}$ or $\mathbb{C}^n \mapsto \mathbb{C}^{2n}$ and working with gradient vectors or Hessian matrices in this new space [10]. Though these approaches facilitate working in the complex domain, they might lead to redundant representations and usually involve matrices with special block structures. Thus a number of simplifying assumptions are usually done most common of which is circularity to render the analysis and derivations tractable (see *e.g.* [2], [3]).

We show how an elegant result due to Brandwood [11] can be used to introduce a framework for complex-valued learning such that all computations can be carried out in the complex domain. We derive the gradient, relative (natural) gradient, Newton, and the Newton variation updates within the framework. We show how the framework easily accommodates the use of fully complex functions rather than the more commonly utilized bounded but non-analytic functions with an application of training of the multilayer perceptron network. These fully complex functions have recently been shown to provide universal approximation of arbitrary functions in the complex domain with domains of convergence dependent on the type of their singularities [12]. Also, these functions provide attractive alternatives for performing independent component analysis (ICA) by efficiently generating higher-order statistical information. We also present examples of the application of the framework for deriving efficient algorithms to perform ICA in the complex domain, which bypass the the need for simplifying assumptions such as circularity of sources in the derivation and the analysis of the algorithms.

References

[1] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Trans. Signal Processing*, vol. 43, pp. 2030–2033, Aug. 1995.

[2] J. Anemüller, T. J. Sejnowski, and S. Makeig, "Complex independent component analysis of frequency-domain electroencephalographic data," *Neural Networks*, vol. 16, pp. 1311–1323, 2003.

[3] E. Bingham and A. Hyvarinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *Int. J. Neural Systems*, vol. 10, pp. 1–8, 2000.

[4] P. Smaragdis, "Blind separation of convolved mixtures in the frequency domain," *Neurocomputing*, vol. 22, pp. 21–34, 1998.

[5] T. Adalı, T. Kim, and V. Calhoun, "Independent component analysis by complex nonlinearities," in *Proc. ICASSP*, Montreal, Canada, May 2004, vol. 5, pp. 525–528.

[6] M. Novey and T. Adalı, "ICA by maximization of nongaussianity using complex functions," in *Proc. MLSP*, Mystic, CT, 2005.

[7] J.-F. Cardoso and T. Adalı, "The maximum likelihood approach to complex ICA," in *Proc. ICASSP*, Toulouse, France, May 2006.

[8] V. D. Calhoun, T. Adalı, and G. D. Pearlson, "Independent component analysis of fmri data in the complex domain," *Magnetic Resonance in Medicine*, vol. 48, no. 1, pp. 180–192, July 2002.

[9] P. Schreier and L. Scharf, "Detection and estimation of improper complex random signals," *IEEE Trans. on Info. Theory*, vol. 51, no. 1, pp. 306–312, Jan. 2005.

[10] A. van den Bos, "Complex gradient and Hessian," *IEE Proc. Image Signal Processing*, vol. 141, pp. 380–382, Dec. 1994.

[11] D. H. Brandwood, "A complex gradient operator and its application in adaptive array theory," *Proc. Inst. Elect. Eng.*, vol. 130, no. 1, pp. 11–16, Feb. 1983.

[12] T. Kim and T. Adalı, "Approximation by fully complex multilayer perceptrons'," *Neural Comput.*, vol. 15, pp. 1641–1666, 2003.

Topic: Learning theory and algorithms
Preference: Oral/Poster