# Factor Analysis as a Tool for Signal Processing

Stefan Wijnholds[†], Ahmad Mouri Sardarabadi[*] and Alle-Jan van der Veen[*]

[†] ASTRON, Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
[*] TU Delft, Fac. EEMCS, Mekelweg 4, 2628 CD Delft, The Netherlands

***Abstract*—In array signal processing, eigenvalue decompositions are commonly used to analyze covariance matrices, e.g. for subspace estimation. An implicit assumption is that the noise power at each antenna element is the same. If the array is not calibrated and the noise powers are different, a Factor Analysis is a more appropriate alternative. It is not a well-known tool in signal processing.**

**Factor Analysis proves to be very useful for interference mitigation in uncalibrated radio astronomy arrays. It also plays a role in image formation and self-calibration. We will study some of these applications.**

## I. Introduction

Factor analysis considers covariance data models where the noise is uncorrelated but has unknown powers at each sensor, i.e., the noise covariance matrix is an arbitrary diagonal with positive real entries. In these cases the familiar eigenvalue decomposition (EVD) has to be replaced by a more general “Factor Analysis” decomposition (FAD), which then reveals all relevant information. It is a very relevant model for the early stages of data processing in radio astronomy, because at that point the instrument is not yet calibrated and the noise powers on the various antennas may be quite different.

As it turns out, this problem has been studied in the psychometrics, biometrics and statistics literature since the 1930s (but usually for real-valued matrices) [1], [2]. The problem has received much less attention in the signal processing literature. In this presentation, we briefly describe the FAD and some algorithms for computing it, as well as some applications.

## II. Problem formulation

Assume that we have a set of $Q$ narrow-band Gaussian signals impinging on an array of $J$ sensors. The received signal can be described in complex envelope form by

$$\mathbf{x}(n) \;=\; \sum_{q=1}^{Q} \mathbf{a}_q s_q(n) + \mathbf{n}(n) \;=\; \mathbf{A}\mathbf{s}(n) + \mathbf{n}(k) \qquad (1)$$

where $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_Q]$ contains the array response vectors. In this model, $\mathbf{A}$ is unknown, and the array response vectors are unstructured. The source vector $\mathbf{s}(n)$ and noise vector $\mathbf{n}(n)$ are considered i.i.d. Gaussian, i.e., the corresponding covariance matrices are diagonal. Without loss of generality, we can scale the source signals such that the source covariance matrix is identity.

The data covariance matrix thus has the form

$$\mathbf{R} = \mathbf{A}\mathbf{A}^H + \mathbf{D} \qquad (2)$$

where we assume $Q < J$ so that $\mathbf{A}\mathbf{A}^H$ is rank deficient.

Many signal processing algorithms are based on estimates of the signal subspace, i.e. the range of $\mathbf{A}$. If the noise is white ($\mathbf{D} = \sigma^2\mathbf{I}$), this information is provided by the eigenvalue decomposition of $\mathbf{R}$. This does not work if the noise is not uniform. The objective of factor analysis is, for given $\mathbf{R}$, to identify $\mathbf{A}$ and $\mathbf{D}$, as well as the factor dimension $Q$.

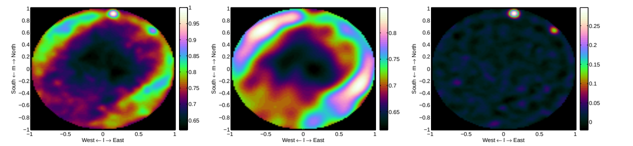

Fig. 1. $(a)$ Calibrated all-sky map for a LOFAR prototype station; $(b)$ image of $\mathbf{R}_n$, corresponding to baselines shorter than 4 wavelengths; $(c)$ image of the residual.

## III. Research issues

Issues to be discussed in the presentation are:

1) *Identifiability:* What constraints provide unique results; what is the maximal factor rank;
2) *Detection:* how can the factor rank be determined;
3) *Estimation:* how can the factors be estimated.

Some answers are obtained by viewing the problem as a form of covariance matching (cf. [3]). An extension of the data model is

$$\mathbf{R} = \mathbf{A}\mathbf{A}^H + \mathbf{R}_n$$

where $\mathbf{R}_n$ is a band matrix. This can be further generalized to more general (sparse) $\mathbf{R}_n$ with known locations of the nonzero entries.

## IV. Applications to radio astronomy

In the context of radio astronomy, factor analysis shows up in a number of applications, see [4] for an overview. Interference cancellation is demonstrated in the accompanying poster presentation. A rank-1 factor analysis problem occurs in the calibration of an array of telescopes pointing at a single calibrator source [5]. As application of the extended factor analysis problem, consider a field with point sources and an extended emission [6]. The extended emission has mostly effect on the short baselines (a band matrix of sorts) whereas the point sources give a low rank contribution. After extended factor analysis, the two components can be imaged separately. Figure 1 shows LOFAR station data, and the resulting image components.

## References

[1] D. N. Lawley and A. E. Maxwell, *Factor Analysis as a Statistical Method*. Butterworth & Co, London, 1971.

[2] K. V. Mardia, J. T. Kent, and J. M. Bibby, *Multivariate Analysis*. Academic Press, 1979.

[3] B. Ottersten, P. Stoica, and R. Roy, “Covariance matching estimation techniques for array signal proce ssing applications,” *Digital Signal Processing, A Review Journal*, vol. 8, pp. 185–210, Jul. 1998.

[4] A. J. van der Veen and S. J. Wijnholds, “Signal processing tools for radio astronomy,” in *Handbook of Signal Processing Systems, 2nd ed.*, S. Bhattacharryya, Ed. Springer, Feb. 2013, p. (40 pp.).

[5] A. J. Boonstra and A. J. van der Veen, “Gain calibration methods for radio telescope arrays,” *IEEE Trans. Signal Processing*, vol. 51, no. 1, pp. 25–38, Jan. 2003.

[6] S. J. Wijnholds and A. J. van der Veen, “Self-calibration of radio astronomical arrays with non-diagonal noise covariance matrix,” in *Proc. 17th European Signal Processing Conference (EUSIPCO 2009)*. Glasgow (UK): Eurasip, Aug. 2009, pp. 1146–1150.