Iterative Blind Separation of Gaussian Data of Unknown Order

by

Amy Mueller

Submitted to the Department of Electrical Engineering and Computer Science
on May 9, 2003, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

A method for blind separation of noisy jointly Gaussian multivariate signals $\mathbf{X}$ is presented, where $\mathbf{X} = \mathbf{AP} + \mathbf{N}$, $\mathbf{X}$ is an observed set of vectors, $\mathbf{A}$ is the mixing matrix, $\mathbf{P}$ is the unknown signal matrix, and $\mathbf{N}$ is white noise. The objective is to estimate all matrices on the right-hand side when even their dimensions (the system order) are unknown. The algorithms developed are extensions of the Iterative Order and Noise (ION) estimation algorithm [10]. Improvements made within the iterative structure of ION to better estimate the order and noise yield ION'. The addition of a second-order blind identification algorithm (SOBI, [4]) subsequently yields ONA, which fully characterizes a data set by estimating the (O)rder, (N)oise, and mixing matrix ($\mathbf{A}$). Metrics are developed to evaluate the performance of these algorithms, and their applicability is discussed. Optimum algorithm constants for ION' and ONA are derived, and their range of applicability is outlined.

The algorithms are evaluated through application to three types of data: (1) simulated Gaussian data which spans the problem space, (2) a set of non-Gaussian factory data with 577 variables, and (3) a hyperspectral image with 224 channels. The ONA algorithm is extended to 2D (spatial) hyperspectral problems by exploiting spatial rather than time correlation. ONA produces a full characterization of the data with high signal-to-noise ratios for most unknown parameters in the Gaussian case, though the jointly Gaussian $\mathbf{P}$ is shown to be most difficult to retrieve. In all three cases, ONA reduces the noise in the data, identifies small sets of highly correlated variables, and unmixes latent signals. The spatial ONA identifies surface features in the hyperspectral image and retrieves sources significantly more independent than those retrieved by PCA. Further exploration of the applicability of these algorithms to other types of data and further algorithmic improvement is recommended.

Thesis Supervisor: David H. Staelin
Title: Professor of Electrical Engineering