CCIT REPORT #748 October 2009

Anomaly Preserving $\ell_{2,\infty}$-Optimal Dimensionality Reduction over a Grassmann Manifold

Oleg Kuybeda, David Malah, and Meir Barzohar

Abstract

In this paper, we address the problem of redundancy reduction of high-dimensional noisy signals that may contain anomaly (rare) vectors, which we wish to preserve. Since anomaly data vectors contribute weakly to the ℓ_2-norm of the signal as compared to the noise, ℓ_2-based criteria are unsatisfactory for obtaining a good representation of these vectors. As a remedy, a new approach, named Min-Max-SVD (MX-SVD) was recently proposed for signal-subspace estimation by attempting to minimize the *maximum* of data-residual ℓ_2-norms, denoted as $\ell_{2,\infty}$ and designed to represent well both abundant and anomaly measurements. However, the MX-SVD algorithm is greedy and only approximately minimizes the proposed $\ell_{2,\infty}$-norm of the residuals. In this paper we develop an optimal algorithm for the minization of the $\ell_{2,\infty}$-norm of data misrepresentation residuals, which we call *Maximum Orthogonal complements Optimal Subspace Estimation* (MOOSE). The optimization is performed via a natural conjugate gradient learning approach carried out on the set of n dimensional subspaces in $\mathbb{R}^m$, $m > n$, which is a Grassmann manifold. The results of applying MOOSE, MX-SVD, and ℓ_2 - based approaches are demonstrated both on simulated and real hyperspectral data.

Index Terms

Signal-subspace rank, singular value decomposition (SVD), Min-Max-SVD (MX-SVD), Maximum Orthogonal-Complements Analysis (MOCA), Hyperspectral Signal Identification by Minimum Error (HySime), anomaly detection, dimensionality reduction, redundancy reduction, hyperspectral images, Grassmann manifold.