A Consensus-based Decentralized EM for a Mixture of Factor Analyzers

Gene Whispos, Emre Ertin, & Randolph Moses

Abstract
We consider the problem of decentralized learning of a target appearance manifold using a network of sensors. Sensor nodes observe an object from different angles and then, in an unsupervised and distributed manner, learn a joint statistical model for the data manifold. We employ a mixture of factor analyzers (MFA) model, approximating a potentially nonlinear manifold. We derive a consensus-based decentralized expectation maximization (EM) algorithm for learning the parameters of the mixture densities and mixing probabilities. A simulation example demonstrates the efficacy of the algorithm.

Objectives
• Learn object structure from multiple, distributed partial views for detection and classification applications
• Estimate an underlying low-dimensional data manifold of high-dimensional observations from a spatially distributed network of sensors; disseminate learned structure to all nodes
• Develop decentralized algorithms which are scalable and robust to sensor node or communication link failures

Challenges
• Centralized and distributed learning approaches have poor scalability; not robust to sensor node or communication link outages
• Disseminating manifold representation out to nodes requires additional communications and delay

Approach
• Develop robust algorithm that learns a low-dimensional manifold across a network and simultaneously disseminates representation to all nodes
• Centralized models: LLE [2], Isomap [3], MFA [4]
• Model the nonlinear manifold with a MFA
• Generative statistical model of mixture densities
• Each node observes the model with different mixture probabilities
• Derive a consensus-based Expectation-Maximization (EM) algorithm from distributed measurements
• Each node calculates local statistics from its own observations
• Locally estimate global parameters from network average statistics
• Locally configure consensus-based averaging
• Unlike [2,3,4], our model is a soft assignment of linear mappings to each sensor node. Our algorithm is similar to [5,6]; in contrast, our work incorporates a low-dimensional structure - key to accurately modeling high-dimensional data with an intrinsic manifold structure

The MFA Model
\[ x_{m,j} \sim \sum_{j=1}^{J} \alpha_{m,j} N(x_{m,j} | \mu_j, \Psi_j + \Lambda_j \Lambda_j^T), \quad x_{j} = \Lambda_j y_j + \mu_j + w_j \]
\[ x_j \in \mathbb{R}^p, \quad y_j \in \mathbb{R}^r, \quad r < p \]

Consensus-based EM Algorithm

1. E-step:
\[ \theta_{m+1} = \mathbb{E}[\theta | X_{\text{m+1}}] \]

2. E-step:
\[ \theta_{m+1} = \mathbb{E}[\theta | X_{\text{m+1}}] \]

3. Compute local statistics:
\[ b_{m+1}^{(k)} = \sum_{j=1}^{J} \text{log} \sum_{j=1}^{J} \alpha_{m,j} N(x_{m,j} | \mu_j, \Sigma_j) \]

4. Consensus iterations:
\[ b_{m+1} \leftarrow \text{Hastings-based average consensus} \]

5. M-step:
\[ \theta_{m+1} = (1 - \alpha_m) \theta_m + \alpha_m \theta_{m+1} \]

6. Repeat until convergence

Results
• Local computations scale cubically in intrinsic dimension
• Consensus on sample average of MFA local statistics
• Converges very quickly, instead of N² operations shared with "neighbors"
• Consensus iterations disseminates manifold representation
• Metropolis-Hastings-based average consensus
• Requires only mild constraints on the network
• Nodes need only neighbor degrees to setup consensus weights

Path Forward
• Analyze convergence properties of EM of MFA and overall algorithm
• Investigate speedup strategies
• Automate learning of model order and dimensionality

References