Statistics > Machine Learning

Closed-form EM for Sparse Coding and its Application to Source Separation

Jörg Lücke, Abdul-Saboor Sheikh

(Submitted on 12 May 2011 (v1), last revised 2 Mar 2012 (this version, v6))

We define and discuss the first sparse coding algorithm based on closed-form EM updates and continuous latent variables. The underlying generative model consists of a standard `spike-and-slab' prior and a Gaussian noise model. Closed-form solutions for E- and M-step equations are derived by generalizing probabilistic PCA. The resulting EM algorithm can take all modes of a potentially multi-modal posterior into account. The computational cost of the algorithm scales exponentially with the number of hidden dimensions. However, with current computational resources, it is still possible to efficiently learn model parameters for medium-scale problems. Thus the model can be applied to the typical range of source separation tasks. In numerical experiments on artificial data we verify likelihood maximization and show that the derived algorithm recovers the sparse directions of standard sparse coding distributions. On source separation benchmarks comprised of realistic data we show that the algorithm is competitive with other recent methods.

Comments:	joint first authorship
Subjects:	**Machine Learning (stat.ML)**
Journal reference:	L\"ucke, J. and Sheikh, A.-S. Proc. LVA/ICA, LNCS pp. 213-221, 2012
Cite as:	**arXiv:1105.2493 [stat.ML]**
	(or **arXiv:1105.2493v6 [stat.ML]** for this version)

Submission history

From: Abdul-Saboor Sheikh [view email]
[v1] Thu, 12 May 2011 14:50:09 GMT (303kb)
[v2] Tue, 28 Jun 2011 09:26:35 GMT (629kb)
[v3] Sun, 16 Oct 2011 17:32:36 GMT (54kb,D)
[v4] Mon, 26 Dec 2011 14:16:12 GMT (54kb,D)
[v5] Wed, 11 Jan 2012 15:40:45 GMT (153kb,D)
[v6] Fri, 2 Mar 2012 10:46:38 GMT (153kb,D)

Which authors of this paper are endorsers?