Constructing Graphical Models for Bayesian Ensemble Learning from Simple Building Blocks

Harri Valpola, Tapani Raiko, and Juha Karhunen

Helsinki University of Technology, Neural Networks Research Centre,
P.O.Box 5400, FIN-02015 HUT, Finland.
Email: `{Harri.Valpola, Tapani.Raiko, Juha.Karhunen}@hut.fi`
URL: `http://www.cis.hut.fi/projects/ica/bayes/`

Bayesian ensemble learning [3, 8] and related variational methods [5, 6] have recently been successfully applied to various extensions of linear Gaussian factor analysis and standard independent component analysis (ICA) [4]. These include for example linear independent factor analysis [1], some other extensions of the basic ICA model [9], nonlinear units [2], as well as MLP networks for modeling nonlinear observation mappings [7, 4] and nonlinear dynamics of the latent variables (source signals) [10]. Ensemble learning has also been applied to large discrete models such as belief networks and hidden Markov models.

In this work, we introduce a small number of basic blocks for building hierarchical models which are then learned using ensemble learning. We have thus far considered unsupervised learning tasks, but models compiled of these blocks can be used in context with supervised learning as well. A wide variety of factor analysis type latent variable models can be constructed by combining the basic blocks suitably. Ensemble learning then provides a cost function which can be used for updating the variables as well as for optimizing the model structure. The blocks are designed so that they fit together and yield efficient update rules. Emphasis is on local computations which result in linear computational complexity.

The building blocks include addition and multiplication, Gaussian variables possibly followed by a nonlinearity, as well as discrete variables and switching units. Various model structures proposed in the literature, for example the independent factor analysis model by Attias [1], can be easily build out of these elements. We also present some new model structures. They utilize Gaussian variables which model the variance of other Gaussian variables, allowing the variance to have a hierarchical or dynamical model. Such variance neurons are often useful in practical applications in which variances of the signals involved can be nonstationary.

In particular, we propose and test a structure with a hierachical nonlinear model for variances and means. It is applied successfully to a hierarchical extension of the well-known bars problem. The structure of the model is correctly found in a completely unsupervised manner. We have also carried out some promising preliminary tests with real-world image data. A more detailed description of the building blocks and some experimental results can be found in [11].

Category: Graphical models. Preference: Oral or poster.

References

[1] H. Attias. Independent factor analysis. *Neural Computation*, 11(4):803–851, 1999.

[2] B. Frey and G. Hinton. Variational learning in nonlinear Gaussian belief networks. *Neural Computation*, 11(4):193–214, 1999.

[3] G. Hinton and D. van Camp. Keeping neural networks simple by minimizing the description length of the weights. In *Proc. of the 6th Ann. ACM Conf. on Computational Learning Theory*, pages 5–13, Santa Cruz, CA, USA, 1993.

[4] A. Hyvärinen, J. Karhunen, and E. Oja. *Independent Component Analysis*. J. Wiley, 2001.

[5] M. Jordan, Z. Ghahramani, T. Jaakkola, and L. Saul. An introduction to variational methods for graphical models. In M. Jordan, editor, *Learning in Graphical Models*, pages 105–161. The MIT Press, Cambridge, MA, USA, 1999.

[6] M. Jordan and T. Sejnowski, editors. *Graphical Models: Foundation of Neural Computation*. The MIT Press, Cambridge, MA, USA, 2001.

[7] H. Lappalainen and A. Honkela. Bayesian nonlinear independent component analysis by multi-layer perceptrons. In M. Girolami, editor, *Advances in Independent Component Analysis*, pages 93–121. Springer-Verlag, 2000.

[8] H. Lappalainen and J. Miskin. Ensemble learning. In M. Girolami, editor, *Advances in Independent Component Analysis*, pages 75–92. Springer, Berlin, 2000.

[9] S. Roberts and R. Everson, editors. *Independent Component Analysis: Principles and Practice*. Cambridge Univ. Press, 2001.

[10] H. Valpola and J. Karhunen. An unsupervised ensemble learning method for nonlinear dynamic state-space models. *Neural Computation*, 2002. Submitted, under revision. Available at http://www.cis.hut.fi/juha/.

[11] H. Valpola, T. Raiko, and J. Karhunen. Building blocks for hierarchical latent variable models. In *Proc. of the 3rd Int. Workshop on Independent Component Analysis and Signal Separation (ICA2001)*, pages 710–715, San Diego, California, December 2001. Available also at http://www.cis.hut.fi/harri/.