Variational Bayesian linear regression

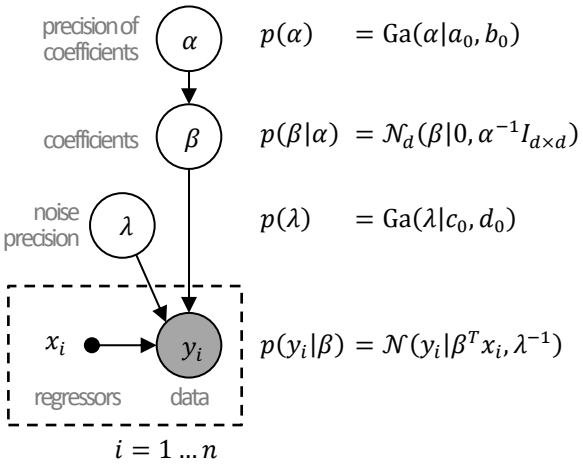
Release v1.01 (r19158) for MATLAB, March 2013

Overview

The conceptual and practical limitations of classical multiple linear regression models of the form $y = X\beta + \epsilon$ can be resolved naturally in a Bayesian framework. Unless based on an overly simplistic parameterization, however, exact inference in Bayesian regression models is analytically intractable. This problem can be overcome using methods for approximate inference. Here, we provide a simple implementation of variational Bayesian inference for a fully Bayesian multiple linear regression model. The code is as easy to use as classical regression implementations, such as regress(), and requires no prerequisites other than MATLAB and the MATLAB Statistics Toolbox.

Model

We consider a multiple linear regression model with a shrinkage prior on the regression coefficients. This model is a generalization of the model illustrated in: Bishop, C.M., *Pattern Recognition and Machine Learning* (2005), Springer, pp. 486–490. We wish to infer on the coefficients β, their precision α, and the noise precision λ. There is no analytical posterior $p(\beta, \alpha, \lambda|y)$. We therefore seek a variational approximation: $q(\beta, \alpha, \lambda) = q_\beta(\beta)\ q_\alpha(\alpha)\ q_\lambda(\lambda)$.



Function vblm() for model estimation and inference

```
% Variational Bayesian multiple linear regression.
%
% Basic usage:
%     q = vblm(y, X)
%
% Full usage:
%     [q, stats, q_trace] = vblm(y, X, a_0, b_0, c_0, d_0)
%
% Args:
%     y:   <n x 1> vector of observations (response variable)
%     X:   <n x d> design matrix (regressors)
%     a_0: shape parameter of the prior precision of coefficients
%     b_0: rate  parameter of the prior precision of coefficients
%     c_0: shape parameter of the prior noise precision
%     d_0: rate  parameter of the prior noise precision
%
% Returns:
%     q: moments of the variational posterior:
%       .mu_n:     posterior expectation of coefficients
%       .Lambda_n: posterior precision matrix of coefficients
%       .a_n:      shape parameter of the posterior precision of coefficients
%       .b_n:      rate  parameter of the posterior precision of coefficients
%       .c_n:      shape parameter of the posterior noise precision
%       .d_n:      rate  parameter of the posterior noise precision
%       .F:        free energy of the model given the data
%       .prior:    struct with a copy of the prior moments a_0, b_0, c_0, d_0
%     stats: additional statistics:
%       .logBF_beta: log Bayes factor between the full model and a reduced model
%                    in which one regressor at a time has been omitted
%       .logBF_null: log Bayes factor between the full model and a null model
%                    whose only regressor is a constant
%     q_trace: all intermediate results
%
% The regression model is
%     p(alpha)            = Ga(alpha | a_0, b_0)
%     p(beta | alpha)     = N(beta | 0, alpha^-1 I)
%     p(lambda)           = Ga(lambda | c_0, d_0)
%     p(y | beta, lambda) = Prod_i N(y_i | beta^T*x_i, lambda^-1)
% where
%     y:      data (response variable)
%     X:      design matrix (regressors)
%     beta:   coefficients
%     alpha:  precision of coefficients
%     lambda: precision of observation noise
%
% The model is inverted by optimizing a variational approximate posterior
% q(alpha, beta, lambda) ~ p(alpha, beta, lambda | y, X) with respect to the
% (negative) free energy, which itself is a lower bound to the log model
% evidence. The variational posterior is derived under the mean-field
% assumption, i.e., q(alpha, beta, lambda) = q(alpha)*q(beta)*q(lambda).
%
% See also:
%     vblm_demo
%     vblm_predict
```

Function vblm_predict() for predicting new values

```
% Posterior predictive density of a variational Bayesian multiple linear
% regression model.
%
% Usage:
%     [m_new, y_new] = vblm_predict(X_new, q)
%
% Args:
%     X_new: <n x d> matrix of new regressors
%     q:     variational posterior, as returned by vblm()
%
% Returns:
%     m_new: <n x 1> vector of posterior predictive response mean
%     t_new: <n x 1> vector of posterior predictive response precision
```

Software note

Kay H. Brodersen
TNU
ETH Zurich
Switzerland
brodersen@biomed.ee.ethz.ch