Supplementary Material for 'Gaussian Process Regression with Censored Data Using Expectation Propagation'

Perry Groot, Peter Lucas
Radboud University Nijmegen, the Netherlands
{perry,peterl}cs.ru.nl

Abstract

We provide implementation details and code for censored regression with Gaussian processes as published in (Groot and Lucas, 2012).

1 Implementation details

1.1 Reparametrization

Since the σ parameter is restricted to $(0, \infty)$, the likelihood is reparameterized in order to obtain an unconstrained optimization problem. This can, for example be done by reparameterizing the likelihood in terms of w where

$$w = \log(\sigma^2), \qquad \sigma^2 = \exp(w) \tag{1}$$

The transformed parameter w is unrestricted. The probability density of w given a probability density for σ^2 is given by

$$p_w(w) = |J| p_{\sigma^2}(\exp(w)) = \sigma^2 p_{\sigma^2}(\sigma^2) \tag{2}$$

where $J = \frac{\partial \exp(w)}{\partial w}$ is the Jacobian of the transformation between parameters (Gelman et al., 2003).

1.2 Likelihood derivatives

$$\begin{aligned}
\frac{\partial \log L}{\partial f_i} =& \frac{1}{\sigma}\left[\frac{-\phi(z_l)}{1-\Phi(z_l)}\mathbf{1}_{[y_i=l]} + \frac{\phi(z_u)}{\Phi(z_u)}\mathbf{1}_{[y_i=u]} \right.\\
&\left. + \frac{(y_i - f_i)}{\sigma}\mathbf{1}_{[l<y_i<u]}\right]\\
\frac{\partial \log L}{\partial w} =& \frac{1}{2}\left[\sum_{y_i=l}\frac{z_l\phi(z_l)}{1-\Phi(z_l)} - \sum_{y_i=u}\frac{z_u\phi(z_u)}{\Phi(z_u)} \right.\\
&\left. + \sum_{l<y_i<u}\left(\frac{(y_i-f_i)^2}{\sigma^2} - 1\right)\right]
\end{aligned}$$

where $z_l = (f_i - l)/\sigma$, $z_u = (f_i - u)/\sigma$.

1.3 Site derivatives

The following term is needed when evaluating the gradients of the marginal likelihood estimate Z_{EP} with respect to the likelihood parameters (Seeger, 2005)

$$E_q\left[\frac{\partial \log t_i}{\partial w}\right] = \int \left[\frac{\partial \log t_i}{\partial w}\right] \hat{Z}_i^{-1} q_{\backslash i} t_i \, \mathrm{d}f_i \tag{3}$$

which depends on whether y_i is censored from below or above or non-censored. If $y_i = l$, let $\mathcal{N}(f|l, \sigma^2)q_{\backslash i} = Z_l \mathcal{N}(f|c, C)$. Then

$$E_q\left[\frac{\partial \log t_i}{\partial w}\right] = \frac{1}{2}\hat{Z}_i^{-1} Z_l(c - l) \tag{4}$$

If $l < y_i < u$

$$E_q\left[\frac{\partial \log t_i}{\partial w}\right] = \frac{y_n^2}{2\sigma^2} - \frac{1}{2} - \frac{y_n\hat{\mu}_i}{\sigma^2} + \frac{\hat{\sigma}^2 + \hat{\mu}_i^2}{2\sigma^2} \tag{5}$$

If $y_i = u$, let $\mathcal{N}(f|u, \sigma^2)q_{\backslash i} = Z_u \mathcal{N}(f|c, C)$. Then

$$E_q\left[\frac{\partial \log t_i}{\partial w}\right] = \frac{1}{2}\hat{Z}_i^{-1} Z_u(u - c) \tag{6}$$

where $\hat{Z}_i$, $\hat{\sigma}^2$, $\hat{\mu}_i$ refer to the corresponding moments in (Groot and Lucas, 2012), Section 2.3.2.

1.4 Moments

The moment equations are given in (Groot and Lucas, 2012), Section 2.3.2. Some care has to be taken with the implementation of $\frac{\phi(z)}{\Phi(z)}$ since

$$\lim_{z \to -\infty} \phi(z) = \lim_{z \to -\infty} \Phi(z) = 0 \tag{7}$$

For small values of z one can use the tight upper bound (Rasmussen and Nickisch, 2010)

$$\frac{\phi(z)}{\Phi(z)} \sim \sqrt{\frac{1}{4}z^2 + 1} - \frac{1}{2}z \tag{8}$$

and furthermore make use of

$$\frac{\phi(z)}{1 - \Phi(z)} = \frac{\phi(-z)}{\Phi(-z)} \tag{9}$$

References

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. 2003. *Bayesian Data Analysis, Second Edition (Chapman & Hall/CRC Texts in Statistical Science)*. Chapman and Hall/CRC, 2 edition, July.

P. Groot and P. Lucas. 2012. Gaussian process regression with censored data using Expectation propagation. In *The 6th European Workshop on Probabilistic Graphical Models (PGM)*.

C. E. Rasmussen and H. Nickisch. 2010. Gaussian Processes for Machine Learning (GPML) Toolbox. *Journal of Machine Learning Research*, 11:3011–3015, Nov.

M. Seeger. 2005. Expectation propagation for exponential families. Technical report, Max Planck Institute, Tuebingen, November.