Laplace approximation and natural gradient for Gaussian process regression with heteroscedastic Student-t model

Marcelo Hartmann & Jarno Vanhatalo

Supplementary material

The derivatives of each element of $\nabla \log L(\mathbf{y}\,|\,\mathbf{f},\nu)$ and W w.r.t $\mathbf{f}_1$, $\mathbf{f}_2$ and ν are given below. Note that, with respect to W, these are third-order derivatives of the negative of the log-likelihood function and so some derivatives will appear twice since the order of the derivatives can be interchanged. In all the equations presented below we consider that $z_i = \frac{y_i - f_1(\mathbf{x}_i)}{\exp(f_2(\mathbf{x}_i))}$ for $i = 1, \ldots, n$.

$$\frac{\partial \log L(\mathbf{y}\,|\,\mathbf{f},\nu)}{\partial \nu} = \frac{n}{2}\psi\left(\frac{\nu+1}{2}\right) - \frac{n}{2}\psi\left(\frac{\nu}{2}\right) - \frac{n}{2\nu} - \sum_{i=1}^{n} \log\left(1 + \frac{1}{\nu} z_i^2\right) + \frac{z_i^2(\nu+1)}{\nu^2(1+z_i^2/\nu)} \quad (1)$$

The elements of the derivatives of $\nabla_{\mathbf{f}} \log L(\mathbf{y}\,|\,\mathbf{f},\nu)$ w.r.t ν are

$$\begin{aligned} \frac{\partial^2 \log \pi(y_i|f_1(\mathbf{x}_i), f_2(\mathbf{x}_i), \nu)}{\partial \nu \partial f_1(\mathbf{x}_i)} &= \frac{2}{[\exp(f_2(\mathbf{x}_i))]} \frac{z_i^3 - z}{\nu^2(1+z_i^2/\nu)^2} \\ \frac{\partial^2 \log \pi(y_i|f_1(\mathbf{x}_i), f_2(\mathbf{x}_i), \nu)}{\partial \nu \partial f_2(\mathbf{x}_i)} &= \frac{z_i^4 - z_i^2}{\nu^2(1+z_i^2/\nu)^2} \end{aligned} \quad (2)$$

$$\frac{\partial \mathrm{W}}{\partial f_1(\mathbf{x}_i)} = \begin{cases} \frac{1}{[\exp(f_2(\mathbf{x}_i))]^3}\left(1+\frac{1}{\nu}\right)\left[\frac{2z_i}{\nu(1+z_i^2/\nu)^2}\left(\frac{4}{1+z_i^2/\nu} - 1\right)\right] \\ \text{for } i = j = 1, \ldots, n \\ -\frac{2}{[\exp(f_2(\mathbf{x}_i))]^2}\left(1+\frac{1}{\nu}\right)\frac{1}{(1+z_i^2/\nu)^2}\left[1 - \frac{4z_i^2}{\nu(1+z_i^2/\nu)}\right] \\ \text{for } i = 1, \ldots, N \text{ and} \\ j = (i+n)\mathbb{1}_{\{1,\ldots,n\}}(i) + (i-n)\mathbb{1}_{\{n+1,\ldots,N\}}(i) \\ -\frac{1}{\exp(f_2(\mathbf{x}_i))}\left(1+\frac{1}{\nu}\right)\frac{z_i}{(1+z_i^2/\nu)^2}\left[4 - \frac{8z_i^2}{\nu(1+z_i^2/\nu)}\right] \\ \text{for } i = j = n+1, \ldots, N \\ 0, \text{ otherwise.} \end{cases} \quad (3)$$

$$\frac{\partial \mathrm{W}}{\partial f_2(\mathbf{x}_i)} = \begin{cases} -\frac{2}{[\exp(f_2(\mathbf{x}_i))]^2}\left(1+\frac{1}{\nu}\right)\frac{1}{(1+z_i^2/\nu)^2}\left[1-\frac{4z_i^2}{\nu(1+z_i^2/\nu)}\right] \\ \text{for } i=j=1,\ldots,n \\ -\frac{1}{\exp(f_2(\mathbf{x}_i))}\left(1+\frac{1}{\nu}\right)\frac{z_i}{(1+z_i^2/\nu)^2}\left[4-\frac{8z_i^2}{\nu(1+z_i^2/\nu)}\right] \\ \text{for } i=1,\ldots,N \text{ and} \\ j=(i+n)\mathbb{1}_{\{1,\ldots,n\}}(i)+(i-n)\mathbb{1}_{\{n+1,\ldots,N\}}(i) \\ 2-\left(2+\frac{2}{\nu}\right)\left[2z_i^2+\frac{4z_i^2-z_i^4/\nu}{(1+z_i^2/\nu)}-\frac{4z_i^4}{\nu(1+z_i^2/\nu)^2}\right]- \\ \frac{z_i^2-1}{(1+z_i^2/\nu)}+3\Big[-1+\left(1+\frac{1}{\nu}\right)\frac{z_i^2}{(1+z_i^2/\nu)}\Big(1+ \\ \frac{2}{(1+z_i^2/\nu)}\Big)\Big] \\ \text{for } i=j=n+1,\ldots,N \\ 0, \ \text{otherwise.} \end{cases} \tag{4}$$

$$\frac{\partial \mathrm{W}}{\partial \nu} = \begin{cases} \frac{1}{\exp(f_2(\mathbf{x}_i))]^2}\Big\{-\frac{1}{\nu^2}\left[\frac{2}{(1+z_i^2/\nu)^2}-\frac{1}{(1+z_i^2/\nu)}\right] \\ +\left(1+\frac{1}{\nu}\right)\left[\frac{4z_i^2}{(1+z_i^2/\nu)^2}-\frac{z^2}{\nu(1+z_i^2/\nu)}\right]\Big\} \\ \text{for } i=j=1,\ldots,n \\ -\frac{2}{\exp(f_2(\mathbf{x}_i))]}\frac{1}{(1+z_i^2/\nu)^2}\left[\frac{z}{\nu^2}-\left(1+\frac{1}{\nu}\right)\frac{2z_i^3}{\nu^2(1+z_i^2/\nu)^2}\right] \\ \text{for } i=1,\ldots,N \text{ and} \\ j=(i+n)\mathbb{1}_{\{1,\ldots,n\}}(i)+(i-n)\mathbb{1}_{\{n+1,\ldots,N\}}(i) \\ -\frac{2z_i^2}{\nu^2}\left[\frac{2}{(1+z_i^2/\nu)^2}+\frac{1}{(1+z_i^2/\nu)}\right]+ \\ \left(1+\frac{1}{\nu}\right)\frac{z_i^2}{\nu^2}\left[\frac{4}{(1+z_i^2/\nu)^3}+\frac{1}{(1+z_i^2/\nu)^2}\right] \\ +\frac{z_i^4-z_i^2}{\nu^2(1+z_i^2/\nu)^2} \\ \text{for } i=j=n+1,\ldots,N \\ 0, \ \text{otherwise.} \end{cases} \tag{5}$$