Supplement (L4DC-2020): Bayesian joint state and parameter tracking in autoregressive models

Ismail Senoz[1] I.SENOZ@TUE.NL
Albert Podusenko[1] A.PODUSENKO@TUE.NL
Wouter M. Kouw[1] W.M.KOUW@TUE.NL
Bert de Vries[1,2] BERT.DE.VRIES@TUE.NL
[1] *Eindhoven University of Technology, Eindhoven, the Netherlands*
[2] *GN Hearing, Eindhoven, the Netherlands*

Derivations for GCV Node

In this section we will derive the messages 8,10,12,13 and 14 that are around the GCV node. We start by computing the message towards y_t. This is equivalent to:

$$\overrightarrow{\nu}(y_t) \propto \int \overrightarrow{\nu}(\boldsymbol{\theta}^\top \mathbf{x}_t) \exp\left(\mathbb{E}_{q(\kappa)q(\omega)q(z_t)}[\log \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, \vartheta_t)]\right) \mathrm{d}\boldsymbol{\theta}\mathrm{d}\mathbf{x}_t. \tag{1}$$

We first analyze the log expectation term.

$$\mathbb{E}_{q(\kappa)q(\omega)q(z_t)}[\log \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, \vartheta_t)] \propto \mathbb{E}\left[-0.5\left(\kappa z_t + \omega + \left(y_t - \boldsymbol{\theta}^\top \mathbf{x}_t\right)^2 \exp(-\kappa z_t - \omega)\right)\right] \tag{2a}$$

$$\propto -0.5\left(\mathbb{E}[\kappa]\mathbb{E}[z_t] + \mathbb{E}[\omega] + \left(y_t - \boldsymbol{\theta}^\top \mathbf{x}_t\right)^2 \mathbb{E}[\exp(-\kappa z_t - \omega)]\right) \tag{2b}$$

If we assume $q(\kappa), q(\omega), q(z_t)$ to be Gaussian we can evaluate the means and the expectation $\mathbb{E}[\exp(-\omega)]$ analytically, i.e.

$$\xi_3 \triangleq \mathbb{E}[\exp(-\omega)] = \exp(-\mathbb{E}[\omega] + \mathbb{V}[\omega]/2). \tag{3}$$

Aroian (1947) approximates the multiplication of two Normally distributed random variable. Using the results of Theorem 2.6 we approximate the expectation

$$\mathbb{E}[\exp(-\kappa z_t)] \approx \xi_2 \triangleq \exp\left(-\mathbb{E}[\kappa]\mathbb{E}[z_t] + 0.5\underbrace{\left(\mathbb{E}[z_t]^2\mathbb{V}[\kappa] + \mathbb{E}[\kappa]^2\mathbb{V}[z_t] + \mathbb{V}[z_t]\mathbb{V}[\kappa]\right)}_{\xi_1}\right). \tag{4}$$

Combining (4) and (3) we can approximate

$$\exp\left(\mathbb{E}_{q(\kappa)q(\omega)q(z_t)}[\log \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, \vartheta_t)]\right) \approx \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, (\xi_2\xi_3)^{-1}). \tag{5}$$

Then the message towards y_t can be computed as the convolution

$$\overrightarrow{\nu}(y_t) \propto \int \overrightarrow{\nu}(\boldsymbol{\theta}^\top \mathbf{x}_t)\mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, (\xi_2\xi_3)^{-1})\mathrm{d}\boldsymbol{\theta}\mathrm{d}\mathbf{x}_t \tag{6a}$$

$$\propto \mathcal{N}(y_t|\mathbb{E}[\boldsymbol{\theta}]^\top \mathbf{x}_t, (\xi_2\xi_3)^{-1}) \tag{6b}$$

The last step follows from the convolution property of Gaussians. Nevertheless, we do not add the variance to the term $\xi_2\xi_3$ as $\mathbf{x}_t$ is observed and there is no uncertainty associated with it. Next we proceed to derive the message

$$\overleftarrow{\nu}(\boldsymbol{\theta}^\top \mathbf{x}_t) \propto \int \overleftarrow{\nu}(y_t) \exp\left(\mathbb{E}_{q(\kappa)q(\omega)q(z_t)}[\log \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, \vartheta_t)]\right) \mathrm{d}\boldsymbol{\theta}\mathrm{d}y_t. \tag{7}$$

Using the results from (3), (4) and noting that y_t is given, we again see that

$$\overleftarrow{\nu}(\boldsymbol{\theta}^\top \mathbf{x}_t) \propto \int \overleftarrow{\nu}(y_t)\mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, (\xi_2\xi_3)^{-1})\mathrm{d}\boldsymbol{\theta}\mathrm{d}y_t \tag{8a}$$

$$\propto \mathcal{N}(\boldsymbol{\theta}^\top \mathbf{x}_t|y_t, (\xi_2\xi_3)^{-1}) \tag{8b}$$

Derivation of messages towards z_t and κ is equivalent so we only give it for z_t and just present the result for κ.

$$\overleftarrow{\nu}(z_t) \propto \exp\left(\mathbb{E}_{q(\kappa)q(\omega)q(y_t,\boldsymbol{\theta}^\top \mathbf{x}_t)}[\log \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, \vartheta_t)]\right) \tag{9a}$$

$$\propto \exp\left(-0.5\left(\mathbb{E}[\kappa]z_t + \mathbb{E}[\omega] + \mathbb{E}\left[\left(y_t - \boldsymbol{\theta}^\top \mathbf{x}_t\right)^2\right]\mathbb{E}[\exp(-\kappa z_t - \omega)]\right)\right) \tag{9b}$$

$$\propto \exp\left(-0.5\left(\mathbb{E}[\kappa]z_t + \underbrace{(y_t - \mathbb{E}[\boldsymbol{\theta}^\top]\mathbf{x}_t)^2}_{\xi_4}\mathbb{E}[\exp(-\kappa z_t - \omega)]\right)\right) \tag{9c}$$

$$\propto \exp\left(-0.5\left(\mathbb{E}[\kappa]z_t + \xi_3\xi_4\mathbb{E}[\exp(-\kappa z_t)]\right)\right) \tag{9d}$$

$$\propto \exp\left(-0.5\left(\mathbb{E}[\kappa]z_t + \xi_3\xi_4 \exp\left(-\mathbb{E}[\kappa]z_t + 0.5z_t^2\mathbb{V}[\kappa]\right)\right)\right) \tag{9e}$$

Similarly for κ,

$$\overleftarrow{\nu}(\kappa) \propto \exp\left(-0.5\left(\kappa\mathbb{E}[z_t] + \xi_3\xi_4 \exp\left(-\mathbb{E}[z_t]\kappa + 0.5\kappa^2\mathbb{V}[z_t]\right)\right)\right). \tag{10}$$

The last message we are interested is towards ω edge. It follows;

$$\overleftarrow{\nu}(\omega) \propto \exp\left(\mathbb{E}_{q(\kappa)q(z_t)q(y_t,\boldsymbol{\theta}^\top \mathbf{x}_t)}[\log \mathcal{N}(y_t|\boldsymbol{\theta}^\top \mathbf{x}_t, \vartheta_t)]\right) \tag{11a}$$

$$\propto \exp\left(-0.5\left(\omega + \xi_1\xi_4 \exp(-\omega)\right)\right) \tag{11b}$$

References

Leo A. Aroian. The Probability Function of the Product of Two Normally Distributed Variables. *The Annals of Mathematical Statistics*, 18(2):265–271, June 1947. ISSN 0003-4851, 2168-8990. doi: 10.1214/aoms/1177730442. URL https://projecteuclid.org/euclid.aoms/1177730442.