ForneyLab Derivations: Gaussian scale model

Bart van Erp

January 8, 2021

Abstract

This document provides the necessary derivations for using Gaussian scale model in ForneyLab.

Contents

1 Overview

1.1 Model specification

The signal model is given as

$$p(X, \xi) = p(X \mid \xi)p(\xi). \tag{1}$$

Here X denotes the complex Fourier coefficients. In this case all calculations are univariate, as presented in the paper. The variable ξ shows equals the deterministic log-power spectrum at its mode and is therefore regarded as the probabilistic log-power spectrum.

The individual factors are represented as

$$p(\xi) = \mathcal{N}(\xi \mid \mu_\xi,\ \sigma_\xi^2) \tag{2}$$

and

$$p(X \mid \xi) = \frac{e^{-\xi}}{\pi} e^{-e^{-\xi}|X|^2} \tag{3}$$

which represents the complex normal distribution, with zero mean and a covariance which increases exponentially as a function of ξ. The complex normal distribution is generally defined as $\mathcal{N}_\mathcal{C}(\boldsymbol{\mu}, \Gamma, C)$, with location parameter $\boldsymbol{\mu}$, covariance matrix Γ and relation matrix C. For an all-zero relation matrix C and a non-singular covariance matrix Γ, the probability density function can be written as

$$\mathcal{N}_\mathcal{C}(\boldsymbol{x} \mid \boldsymbol{\mu},\ \Gamma,\ C = 0) = \frac{1}{\pi^n \det \Gamma} e^{-(\boldsymbol{x}-\boldsymbol{\mu})^H \Gamma^{-1} (\boldsymbol{x}-\boldsymbol{\mu})}. \tag{4}$$

Throughout these derivations the focus is limited to univariate variables, with no relation between the real and imaginary part. From this it can be observed that the conditional factor can be represented as a complex Normal distribution as

$$p(X \mid \xi) = \mathcal{N}_\mathcal{C}(X \mid 0, e^\xi, 0), \tag{5}$$

which shows interesting similarities with the Hierarchical Gaussian Filter. The simplified univariate representation of the complex-Normal distribution can be found as

$$\mathcal{N}_\mathcal{C}(x \mid \mu,\ \sigma^2,\ C = 0) = \frac{1}{\pi\sigma^2} e^{-\frac{1}{\sigma^2}|x-\mu|^2}. \tag{6}$$

1.2 Mean-field approximation

The intractable probabilistic model is being simplified through the mean-field approximation

$$p(X,\ \xi) \approx q(X, \xi) = q(X)q(\xi), \tag{7}$$

where the individual factors are defined as

$$q(X) = \mathcal{N}_\mathcal{C}(X \mid m_X, v_X, 0) \tag{8}$$

and

$$q(\xi) = \mathcal{N}(\xi \mid m_\xi,\ v_\xi). \tag{9}$$

2 Message $\vec{\nu}(X)$

The variational message $\vec{\nu}(X)$ can be calculated as

$$\begin{aligned} \ln \vec{\nu}(X) &= \mathrm{E}_{q(\xi)} \left[\ln p(X \mid \xi) \right] + const \\ &= \mathrm{E}_{q(\xi)} \left[-\ln(\pi) - \xi - e^{-\xi}|X|^2 \right] + const \\ &= -\mathrm{E}_{q(\xi)} \left[e^{-\xi} \right] |X|^2 + const, \end{aligned} \tag{10}$$

where e^{ξ} is log-Normally distributed and where $\mathrm{E}\left[e^{-\xi}\right]$ can be seen as its moment of order -1. This moment is well-defined and therefore

$$\ln \vec{\nu}(X) = -e^{-m_\xi + v_\xi/2}|X|^2 + const \tag{11}$$

holds. From this description the message $\vec{\nu}(X)$ can be determined as

$$\boxed{\vec{\nu}(X) \propto \mathcal{N}_{\mathcal{C}}(0,\ \exp(m_\xi - v_\xi/2),\ 0)} \tag{12}$$

3 Message $\overleftarrow{\nu}(\xi)$

The variational message $\overleftarrow{\nu}(\xi)$ is defined as

$$\begin{aligned}
\ln \overleftarrow{\nu}(\xi) &= \mathrm{E}_{q(X)}\left[\ln p(X \mid \xi)\right] + const \\
&= \mathrm{E}_{q(X)}\left[-\ln(\pi) - \xi - e^{-\xi}|X|^2\right] + const \\
&= -\xi - e^{-\xi}\mathrm{E}\left[|X|^2\right] + const \\
&= -\xi - e^{-\xi}(v_X + m_X m_X^*) + const.
\end{aligned} \tag{13}$$

$$\boxed{\ln \overleftarrow{\nu}(\xi) = -\xi - e^{-\xi}(v_X + m_X m_X^*) + const} \tag{14}$$

As the structure of this log-message is not related to a family of distributions, we need to approximate this message or the resulting marginal. Here we show 3 approaches.

3.1 Approach 1: VMP with LaPlace approximate messages

This approach approximates the variational message by a Gaussian distribution under the LaPlace approximation.

The mode of the log-message can be found by differentiating with respect to ξ and by setting the equation to 0 as

$$\frac{\partial \ln \overleftarrow{\nu}(\xi)}{\partial \xi} = -1 + e^{-\xi}(v_X + m_X m_X^*) = 0. \tag{15}$$

The mode can be found as

$$\xi_0 = \ln(v_X + m_X m_X^*), \tag{16}$$

which will represent the mean of the Normal approximation.

The true message is expanded around its mode up to a second order, where the first order vanishes.

$$\ln \overleftarrow{\nu}(\xi) \approx \ln \overleftarrow{\nu}(\xi_0) + \frac{1}{2}\frac{\partial^2 \ln \overleftarrow{\nu}(\xi)}{\partial \xi^2}\bigg|_{\xi=\xi_0} (\xi - \xi_0)^2 \tag{17}$$

This second derivative evaluated at the mode can be found as

$$\frac{\partial^2 \ln \overleftarrow{\nu}(\xi)}{\partial \xi^2}\Big|_{\xi=\xi_0} = -e^{-\xi}(v_X + m_X m_X^*)\Big|_{\xi=\xi_0} = -1. \tag{18}$$

The log-message is therefore approximated as

$$\ln \overleftarrow{\nu}(\xi) \approx \ln \overleftarrow{\nu}(\xi_0) - \frac{1}{2}\left(\xi - \xi_0\right)^2. \tag{19}$$

From this the message can be identified as

$$\boxed{\overleftarrow{\nu}(\xi) \propto \mathcal{N}(\ln(v_X + m_X m_X^*),\ 1)} \tag{20}$$

3.2 Approach 2: VMP with functional messages (Gauss-Hermite quadrature)

This approach sends out a functional message and approximates the mean and variance of the resulting marginal (assuming a incoming Gaussian message) using Gauss-Hermite quadrature integration.

Gauss-Hermite quadrature assumes the following approximation

$$\int_{-\infty}^{\infty} e^{-x^2} f(x)\mathrm{d}x \approx \sum_{i=1}^{n} w_i f(x_i), \tag{21}$$

where n is the number of sample points used. x_i are the roots of the roots of the Hermite polynomial $H_n(x)$ and the associated weights are given by

$$w_i = \frac{2^{n-1} n! \sqrt{\pi}}{n^2 (H_{n-1}(x_i))^2} \tag{22}$$

The marginal of ξ is formally defined as

$$q(\xi) = \frac{\overrightarrow{\nu}(\xi)\overleftarrow{\nu}(\xi)}{Z_\xi} = \frac{\overrightarrow{\nu}(\xi)\overleftarrow{\nu}(\xi)}{\int \overrightarrow{\nu}(\xi)\overleftarrow{\nu}(\xi)\mathrm{d}\xi}, \tag{23}$$

where $\overrightarrow{\nu}(\xi) = \mathcal{N}(\xi \mid m_\xi, v_\xi)$ and $\overleftarrow{\nu}(\xi) = c \cdot e^{-\xi - e^{-\xi}(v_X + |m_X|^2)}$. The normalization constant can be calculated using Gauss-Hermite quadrature as

$$\begin{aligned} Z_\xi &= \int_{-\infty}^{\infty} \overrightarrow{\nu}(\xi)\overleftarrow{\nu}(\xi)\mathrm{d}\xi \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi v_\xi}} e^{-\frac{1}{2v_\xi}(\xi - m_\xi)^2} c e^{-\xi - e^{-\xi}(v_X + |m_X|^2)} \mathrm{d}\xi \\ &\overset{\hat{\xi} = \frac{\xi - m_\xi}{\sqrt{2v_\xi}}}{=} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-\hat{\xi}^2} \overleftarrow{\nu}(\hat{\xi}\sqrt{2v_\xi} + m_\xi)\mathrm{d}\hat{\xi} \\ &\approx \frac{1}{\sqrt{\pi}} \sum_{i=1}^{n} w_i \overleftarrow{\nu}(\hat{\xi}_i\sqrt{2v_\xi} + m_\xi) \end{aligned} \tag{24}$$

The moments of ξ can be found as

$$\begin{aligned}
\mathrm{E}_{q(\xi)}[\xi^k] &= \frac{1}{Z_\xi}\int_{-\infty}^{\infty} \xi^k \vec{\nu}(\xi)\overleftarrow{\nu}(\xi)\mathrm{d}\xi \\
&\overset{\hat{\xi}=\frac{\xi-m_\xi}{\sqrt{2v_\xi}}}{=} \frac{1}{Z_\xi}\int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-\hat{\xi}^2}\left(\sqrt{2v_\xi}\hat{\xi}+m_\xi\right)^k \overleftarrow{\nu}(\sqrt{2v_\xi}\hat{\xi}+m_\xi)\mathrm{d}\hat{\xi} \\
&\approx \frac{1}{Z_\xi\sqrt{\pi}}\sum_{i=1}^{n} w_i \left(\sqrt{2v_\xi}\hat{\xi}_i+m_\xi\right)^k \overleftarrow{\nu}(\sqrt{2v_\xi}\hat{\xi}_i+m_\xi)
\end{aligned} \tag{25}$$

From the approximate first and second moment, the sufficient statics for the approximate Gaussian marginals can be found.

3.3 Approach 3: VMP with functional messages (LaPlace)

The third approach is as follows. A functional message will be send out and will collide with an incoming Gaussian message. The resulting product is then approximated as Gaussian using the LaPlace approximation, which expands the message around its mode as

$$\ln q(\xi) \approx \ln q(\xi_0) + \frac{1}{2}\frac{\partial^2 \ln q(\xi)}{\partial \xi^2}\bigg|_{\xi=\xi_0}(\xi-\xi_0)^2 \tag{26}$$

We can find the resulting log-marginals as

$$\begin{aligned}
\ln q(\xi) &= \ln \vec{\nu}(\xi) + \ln \overleftarrow{\nu}(\xi) + const \\
&= -\frac{1}{2v_\xi}(\xi-m_\xi)^2 - \xi - e^{-\xi}(v_X + |m_X|^2) + const
\end{aligned} \tag{27}$$

For the LaPlace approximation, the mode of the (log-)marginal has to be found. The derivative of the log-marginal with respect to ξ can be found as

$$\frac{\partial}{\partial \xi}\ln q(\xi) = -\frac{1}{v_\xi}(\xi - m_\xi) - 1 + e^{-\xi}(v_X + |m_X|^2), \tag{28}$$

whose root does not have a closed-form solution. Therefore we use Newton's method in order to find a root of this function, through the iterative updating of

$$\xi_i = \xi_{i-1} - \frac{\frac{\partial}{\partial\xi}\ln q(\xi)}{\frac{\partial^2}{\partial\xi^2}\ln q(\xi)}\bigg|_{\xi=\xi_{i-1}} \tag{29}$$

In order to find the curvature around the mode, we calculate the second derivative and finally substitute the found stationary point ξ_0

$$\frac{\partial^2}{\partial\xi^2}\ln q(\xi)\bigg|_{\xi=\xi_0} = -\frac{1}{v_\xi} - e^{-\xi_0}(v_X + |m_X|^2). \tag{30}$$

The resulting log-marginal can be described as

$$\ln q(\xi) \approx \ln q(\xi_0) - \frac{1}{2}\left(\frac{1}{v_\xi} + e^{-\xi_0}(v_X + |m_X|^2)\right)(\xi - \xi_0)^2 \tag{31}$$

From this description, the sufficient statistics are recognized as

$$\mathrm{E}_{q(\xi)}[\xi] = \xi_0 \tag{32}$$

$$\mathrm{E}_{q(\xi)}[(\xi - \xi_0)^2] = \frac{1}{\frac{1}{v_\xi} + e^{-\xi_0}(v_X + |m_X|^2)} \tag{33}$$

3.4 Approach 4 & 5: BP with functional messages (Quadrature & LaPlace)

Another approach follows from using belief propagation or the sum-product algorithm. Here we will derive the outgoing message to ξ from the factor node $p(X \mid \xi)$. We assume that the incoming message over X is represented as $\overleftarrow{\mu}(X) = \mathcal{N}_\mathcal{C}(X \mid m_X, \gamma_X^{-1}, 0)$ and that $\gamma_X \in \mathbb{R}^+$. The outgoing message can be found as

$$\begin{aligned}
\overleftarrow{\mu}(\xi) &\propto \int p(X \mid \xi)\overleftarrow{\mu}(X)\mathrm{d}X \\
&\propto \int \frac{e^{-\xi}}{\pi} e^{-e^{-\xi}|X|^2} e^{-\gamma_X |X - m_X|^2} \mathrm{d}X \\
&\propto e^{-\xi} \int \exp\left\{-e^{-\xi} X X^* - \gamma_X (X X^* - X m_X^* - X^* m_x + m_X m_X^*)\right\} \mathrm{d}X \\
&\propto e^{-\xi} \int \exp\left\{-(e^{-\xi} + \gamma_X) X X^* + \gamma_X X m_X^* + \gamma_X X^* m_x)\right\} \mathrm{d}X \\
&\propto e^{-\xi} e^{\hat{\gamma}_X \hat{m}_X \hat{m}_X^*} \int \exp\left\{-\hat{\gamma}_X X X^* + \hat{\gamma}_X X \hat{m}_X^* + \hat{\gamma}_X X^* \hat{m}_x - \hat{\gamma}_X \hat{m}_X \hat{m}_X^*\right\} \mathrm{d}X \\
&\propto e^{-\xi} e^{\hat{\gamma}_X \hat{m}_X \hat{m}_X^*} \int \exp\left\{-\hat{\gamma}_X |X - \hat{m}_X|^2\right\} \mathrm{d}X \\
&\propto e^{-\xi} e^{\hat{\gamma}_X \hat{m}_X \hat{m}_X^*} \frac{\pi}{\hat{\gamma}_X} \\
&\propto \frac{e^{-\xi}}{e^{-\xi} + \gamma_X} \cdot \exp\left\{\hat{\gamma}_X \hat{m}_X \hat{m}_X^*\right\} \\
&\propto \frac{e^{-\xi}}{e^{-\xi} + \gamma_X} \cdot \exp\left\{m_X \gamma_X \frac{\gamma_X m_X^*}{e^{-\xi} + \gamma_X}\right\} \\
&\propto \frac{e^{-\xi}}{e^{-\xi} + \gamma_X} \cdot \exp\left\{\frac{\gamma_X^2 |m_X|^2}{e^{-\xi} + \gamma_X}\right\}
\end{aligned} \tag{34}$$

where the intermediate variables $\hat{\gamma}_X(\xi)$ and $\hat{m}_X(\xi)$, which are a function of ξ, are defined as

$$\hat{\gamma}_X(\xi) = e^{-\xi} + \gamma_X \tag{35}$$

and

$$\hat{m}_X(\xi) = m_X \frac{\gamma_X}{\hat{\gamma}_X(\xi)} \tag{36}$$

Concluding, the functional belief propagation message can be represented as

$$\boxed{\overleftarrow{\mu}(\xi) = \frac{e^{-\xi}}{e^{-\xi} + \gamma_X} \cdot \exp\left\{ \frac{\gamma_X^2 |m_X|^2}{e^{-\xi} + \gamma_X} \right\}} \tag{37}$$

The first part of this function can be interpreted as a mirrored sigmoid function, shifted right over $-\ln \gamma_X$. The second part seems to resemble a sigmoid-like function that has not been mirrored. Together they "squeeze" a probability function between both curves.

This message can be propagated similarly as in the previous two section using the Gauss-Hermite quadrature integration or the LaPlace approximation over the resulting marginal using Newtons method. Another approach is to approximate the above functional message by a Gaussian using LaPlace approximation.

3.5 Approach 6: BP with LaPlace approximate messages

This approach approximates the functional message by a Gaussian distribution under the LaPlace approximation.

The log-message can be found as

$$\ln \overleftarrow{\mu}(\xi) = -\xi - \ln(e^{-\xi} + \gamma_X) + \frac{\gamma_X^2 |m_X|^2}{e^{-\xi} + \gamma_X} \tag{38}$$

The mode of the log-message can be found by differentiating with respect to ξ and by setting the equation to 0 as

$$\frac{\partial \ln \overleftarrow{\mu}(\xi)}{\partial \xi} = -1 + \frac{e^{-\xi}}{e^{-\xi} + \gamma_X} \left(1 + \frac{\gamma_X^2 |m_X|^2}{e^{-\xi} + \gamma_X} \right) = 0. \tag{39}$$

The model has a closed form solution given by

$$\xi_0 = -\ln \left(\frac{\gamma_X}{\gamma_X |m_X|^2 - 1} \right) \tag{40}$$

The second derivative can be found as

$$\frac{\partial^2}{\partial \xi^2} \ln \overleftarrow{\mu}(\xi) = \left(\frac{e^{-\xi}}{e^{-\xi} + \gamma_X} \right) \left(\frac{e^{-\xi}}{e^{-\xi} + \gamma_X} - 1 - \frac{\gamma_X^2 |m_X|^2}{e^{-\xi} + \gamma_X} + \frac{2\gamma_X^2 |m_X|^2 e^{-\xi}}{(e^{-\xi} + \gamma_X)^2} \right) \tag{41}$$

4 Average energy

The average energy of the Gaussian scale model can be calculated as

$$\begin{aligned}
\mathrm{U}[p(X \mid \xi)] &= -\mathrm{E}_{q(X)q(\xi)}[\ln p(X \mid \xi)] \\
&= -\mathrm{E}_{q(X)q(\xi)}[-\xi - \ln \pi - e^{-\xi}\|X\|^2] \\
&= m_\xi + \ln \pi + e^{-m_\xi + v_\xi/2}(\|m_X\|^2 + v_X)
\end{aligned} \tag{42}$$