GSMM derivations

Bart van Erp

November 10, 2020

Abstract

This document provides the derivations of variational messages for the Gaussian scale models.

Contents

1 Overview

1.1 Model specification

The signal model as defined in [1], 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 [2]. 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 Preliminaries

2.1 Moments log-normal distribution

A variable Y is called log-Normal distributed if it is defined as $Y = \exp(X)$, where X is Normal distributed as $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$. The moments of the log-normal distribution can be found as

$$\mathrm{E}\left[Y^n\right] = e^{n\mu_X + n^2 \sigma_X^2 / 2}. \tag{10}$$

2.2 Log of Normal probability distribution

The probability distribution of an univariate Normal distribution is defined as

$$\mathcal{N}(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \tag{11}$$

and its natural logarithm can therefore be found as

$$\ln \mathcal{N}(x \mid \mu, \sigma^2) = -\frac{1}{2}\ln 2\pi - \frac{1}{2}\ln \sigma^2 - \frac{1}{2\sigma^2}(x-\mu)^2. \tag{12}$$

2.3 Log of complex-Normal probability distribution

The natural logarithm of (6) can be found as

$$\ln \mathcal{N}_{\mathcal{C}}(x \mid \mu, \sigma^2, C=0) = -\ln(\pi) - \ln(\sigma^2) - \frac{1}{\sigma^2}|x-\mu|^2. \tag{13}$$

2.4 Product of complex-normal distributions

$$p(\boldsymbol{z}) = \frac{1}{\pi^n \sqrt{\det(\Gamma)\det(\Gamma^*)}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} \boldsymbol{z}-\boldsymbol{\mu} \\ \boldsymbol{z}^*-\boldsymbol{\mu}^* \end{pmatrix}^H \begin{pmatrix} \Gamma & 0 \\ 0 & \Gamma^* \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{z}-\boldsymbol{\mu} \\ \boldsymbol{z}^*-\boldsymbol{\mu}^* \end{pmatrix} \right\} \tag{14}$$

$$p(\boldsymbol{z}) = \frac{1}{\pi^n \sqrt{\det(\Gamma\Gamma^*)}} \exp\left\{ -\frac{1}{2} \begin{pmatrix} \boldsymbol{z}-\boldsymbol{\mu} \\ \boldsymbol{z}^*-\boldsymbol{\mu}^* \end{pmatrix}^H \begin{pmatrix} \Gamma^{-1} & 0 \\ 0 & (\Gamma^*)^{-1} \end{pmatrix} \begin{pmatrix} \boldsymbol{z}-\boldsymbol{\mu} \\ \boldsymbol{z}^*-\boldsymbol{\mu}^* \end{pmatrix} \right\} \tag{15}$$

$$p(\boldsymbol{z}) = \frac{1}{\pi^n \sqrt{\det(|\Gamma|^2)}} \exp\left\{ -\frac{1}{2} \left((\boldsymbol{z}-\boldsymbol{\mu})^H \Gamma^{-1} (\boldsymbol{z}-\boldsymbol{\mu}) + (\boldsymbol{z}^*-\boldsymbol{\mu}^*)^H (\Gamma^*)^{-1} (\boldsymbol{z}^*-\boldsymbol{\mu}^*) \right) \right\} \tag{16}$$

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

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

$$\begin{aligned} \ln \vec{\nu}(X) &= \mathrm{E}_{q(\xi)}\left[\ln p(X \mid \xi)\right] + \mathit{const} \\ &= \mathrm{E}_{q(\xi)}\left[-\ln(\pi) - \xi - e^{-\xi}|X|^2\right] + \mathit{const} \\ &= -\mathrm{E}_{q(\xi)}\left[e^{-\xi}\right]|X|^2 + \mathit{const}, \end{aligned} \tag{17}$$

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

$$\ln \vec{\nu}(X) = -e^{-m_\xi + v_\xi/2}|X|^2 + \mathit{const} \tag{18}$$

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

$$\boxed{\vec{\nu}(X) \propto \mathcal{N}_{\mathcal{C}}(0,\ \exp(m_\xi - v_\xi/2),\ 0)} \tag{19}$$

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

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

$$\begin{aligned}
\ln \overleftarrow{\nu}(\xi) &= \mathrm{E}_{q(X)}\left[\ln p(X \mid \xi)\right] + \mathit{const} \\
&= \mathrm{E}_{q(X)}\left[-\ln(\pi) - \xi - e^{-\xi}|X|^2\right] + \mathit{const} \\
&= -\xi - e^{-\xi}\mathrm{E}\left[|X|^2\right] + \mathit{const} \\
&= -\xi - e^{-\xi}(v_X + m_X m_X^*) + \mathit{const}.
\end{aligned} \tag{20}$$

As the structure of this log-message is not related to a family of distributions, we apply the LaPlace approximation in order to approximate the corresponding message with a Normal distribution.

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

$$\frac{\partial \ln \overleftarrow{\nu}(\xi)}{\partial \xi} = -1 + e^{-\xi}(v_X + m_X m_X^*) = 0. \tag{21}$$

The mode can be found as

$$\xi_0 = \ln(v_X + m_X m_X^*), \tag{22}$$

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.

$$\ln \overleftarrow{\nu}(\xi) \approx \ln \overleftarrow{\nu}(\xi_0) + \frac{1}{2}\frac{\partial^2 \ln \overleftarrow{\nu}(\xi)}{\partial \xi^2}\bigg|_{\xi=\xi_0} (\xi - \xi_0)^2 \tag{23}$$

This second derivative evaluated at the mode can be found as

$$\frac{\partial^2 \ln \overleftarrow{\nu}(\xi)}{\partial \xi^2}\bigg|_{\xi=\xi_0} = -e^{-\xi}(v_X + m_X m_X^*)\Big|_{\xi=\xi_0} = -1. \tag{24}$$

The log-message is therefore approximated as

$$\ln \overleftarrow{\nu}(\xi) \approx \ln \overleftarrow{\nu}(\xi_0) - \frac{1}{2}(\xi - \xi_0)^2. \tag{25}$$

From this the message can be identified as

$$\boxed{\overleftarrow{\nu}(\xi) \propto \mathcal{N}(\ln(v_X + m_X m_X^*),\ 1)} \tag{26}$$

References

[1] J. Hao, T.-W. Lee, and T. J. Sejnowski, "Speech Enhancement Using Gaussian Scale Mixture Models," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, pp. 1127–1136, Aug. 2010. Conference Name: IEEE Transactions on Audio, Speech, and Language Processing.

[2] Şenöz and B. de Vries, "Online Variational Message Passing in the Hierarchical Gaussian Filter," in *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, Sept. 2018. ISSN: 1551-2541.