SUPPLEMENT (MLSP-2018 PAPER): ONLINE VARIATIONAL MESSAGE PASSING IN THE HIERARCHICAL GAUSSIAN FILTER

Ismail Senoz

Eindhoven University of Technology, Eindhoven, the Netherlands

1. DERIVATION OF MESSAGES

In this supplement we derive the update equations that are given in Table 1 in the MLSP-2018 paper entitled *Online Variational Message Passing in the Hierarchical Gaussian Filter*. We accept mean-field assumptions and use Gaussian distributions for each factor. Before we proceed with the message update rules, we compute the expectations $\mathbf{E}_{q_x}[x^2], \mathbf{E}_{q_u q_\kappa}[\exp(\kappa u)]$ and $\mathbf{E}_{q_\omega}[\exp(\omega)]$ that are needed for computation of messages.

It's straightforward to see that first of these expectations is the uncentralized second moment of a Gaussian distributed random variable, hence $\mathbf{E}_{q_x}\left[x^2\right] = m_x^2 + v_x$.

If we assume independence between κ and u, then following [1] we get

$$u\kappa \overset{approx.}{\sim} \mathcal{N}(m_u m_\kappa, \underbrace{m_u^2 v_\kappa + m_\kappa^2 v_u + v_u v_\kappa}_{\gamma_1}). \quad (1)$$

This means $z \triangleq \exp(u\kappa)$ is log-normally distributed with mean and variance as in Eq. 2a. We approximate the joint expectation of this log-normally distributed product of two independent normal random variables as per Eq. 2b. To arrive at Eq. 2b we use the fact that for a log-normally distributed z, all moments are well defined and analytically given by $\mathbf{E}\left[z^n\right] = \exp\left(n m_z + \frac{n^2 v_z}{2}\right)$. For $n = 1$ we obtain Eq. 2b.

$$\exp(u\kappa) \overset{approx.}{\sim} \log\mathcal{N}(m_u m_\kappa, \gamma_1) \quad (2a)$$

$$\mathbf{E}_{q_u q_\kappa}\left[\exp(\kappa u)\right] \approx \exp\left(m_u m_\kappa + \frac{\gamma_1}{2}\right) \quad (2b)$$

$$\mathbf{E}_{q_u q_\kappa}\left[\frac{1}{\exp(\kappa u)}\right] \approx \exp\left(-m_u m_\kappa + \frac{\gamma_1}{2}\right). \quad (2c)$$

We note that $\mathbf{E}_{q_\omega}\left[\exp(\omega)\right]$ is essentially the same with Eq. 2b, hence it follows that

$$\mathbf{E}_{q_\omega}\left[\exp(\omega)\right] \approx \exp\left(m_\omega + \frac{v_\omega}{2}\right) \quad (3a)$$

$$\mathbf{E}_{q_\omega}\left[\frac{1}{\exp(\omega)}\right] \approx \exp\left(-m_\omega + \frac{v_\omega}{2}\right) \quad (3b)$$

We are now equipped with the needed expectations to obtain expressions of Table 1. We start our derivation with $\overrightarrow{\nu}_x$.

$$\overrightarrow{\nu}_x \propto \exp(\mathbf{E}_{/q_x}\left[\log f(x,u,\kappa,\omega)\right]) \quad (4a)$$

$$\propto \exp\left(\mathbf{E}_{/q_x}\left[\frac{-x^2}{2\exp(\kappa u + \omega)}\right]\right) \quad (4b)$$

Combining Eq. 3b and 2c we see that

$$\overrightarrow{\nu}_x \propto \exp\left(-\frac{x^2}{2\exp\left(m_\kappa m_u + m_\omega - \left(\frac{\gamma_1 + v_\omega}{2}\right)\right)}\right) \quad (5a)$$

$$\propto \mathcal{N}(0, \gamma_2) \quad (5b)$$

where we define

$$\gamma_2 = \exp\left(m_\kappa m_u + m_\omega - \left(\frac{\gamma_1 + v_\omega}{2}\right)\right).$$

Next we derive

$$\overrightarrow{\nu}_u \propto \exp\left(\mathbf{E}_{/q_u}[\log f(x,u,\kappa,\omega)]\right) \quad (6a)$$

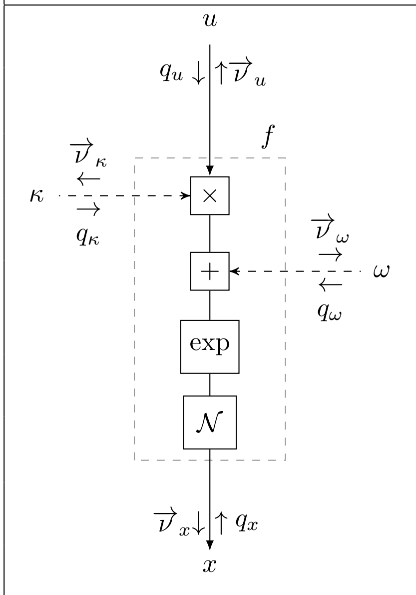
$$\propto \exp\left(\mathbf{E}_{/q_u}\left[-\frac{\kappa u + \omega}{2}\right] + \mathbf{E}_{/q_u}\left[-\frac{x^2}{2\exp(\kappa u + \omega)}\right]\right) \quad (6b)$$

$$\propto \exp\underbrace{\left(-\frac{u m_\kappa + m_\omega}{2} - \frac{m_x^2 + v_x}{\exp(u m_\kappa + m_\omega - \frac{u^2 v_\kappa + v_\omega}{2})}\right)}_{g(u)} \quad (6c)$$

If we assume $v_\kappa \to 0$ then we can expand $g(u)$ around u_0 such that the message becomes Gaussian. We need to find u_0 such that $g'(u_0) = 0$. After a bit of algebraic manipulations we see that $u_0 = \frac{\log(\gamma_3)}{m_\kappa}$, where $\gamma_3 = \left(m_x^2 + v_x\right)\exp\left(-m_\omega + \frac{v_\omega}{2}\right)$. If we expand $g(u)$ around u_0 up to second order, because the first derivative vanishes at u_0 we have,

$$g(u) \approx g(u_0) + \frac{g''(u_0)(u - u_0)^2}{2}. \quad (7)$$

Table 1. Message update rules for the Gaussian-with-Controlled-Variance (GCV) node $f(x, u, \kappa, \omega)$. Assume incoming messages $q_u \sim \mathcal{N}(m_u, v_u)$, $q_x \sim \mathcal{N}(m_x, v_x)$, $q_\omega \sim \mathcal{N}(m_\omega, v_\omega)$ and $q_\kappa \sim \mathcal{N}(m_\kappa, v_\kappa)$. Then, outgoing Gaussian variational messages are as indicated in the table. The middle row specifies the local free energy contribution per Eq. 15 by the GCV node. The final row specifies auxiliary variables $\gamma_\bullet$ that are used in the message update rules.



Node	Message	Update equation
	$\overrightarrow{\nu}(x)$	$\mathcal{N}(0, \gamma_2)$
	$\overrightarrow{\nu}(u)$	$\mathcal{N}\left(\frac{\log(\gamma_3)}{m_\kappa}, \frac{2}{m_\kappa^2}\right)$
	$\overrightarrow{\nu}(\kappa)$	$\mathcal{N}\left(\frac{\log(\gamma_3)}{m_u}, \frac{2}{m_u^2}\right)$
	$\overrightarrow{\nu}(\omega)$	$\mathcal{N}(\log \gamma_4, 1)$
	$F[q]$	$\frac{1}{2}(\log(2\pi) + (m_\kappa m_u + m_\omega)) + \frac{1}{2}\left(m_x^2 + v_x\right)\gamma_2^{-1} - \log\left((2\pi e)^2 \sqrt{v_x v_u v_\kappa v_\omega}\right)$
	γ_1	$m_u^2 v_\kappa + m_\kappa^2 v_u + v_u v_\kappa$
	γ_2	$\exp\left(m_\kappa m_u + m_\omega - \left(\frac{\gamma_1 + v_\omega}{2}\right)\right)$
	γ_3	$\left(m_x^2 + v_x\right)\exp\left(-m_\omega + \frac{v_\omega}{2}\right)$
	γ_4	$\left(m_x^2 + v_x\right)\exp\left(-m_\kappa m_u + \frac{\gamma_1}{2}\right)$

Since $g''(u_0) = 1/m_\kappa^2$ we can approximate

$$\overrightarrow{\nu}_u \propto \mathcal{N}\left(\frac{\log \gamma_3}{m_\kappa}, \frac{2}{m_\kappa^2}\right). \tag{8}$$

If we observe that κ and ω are coupled by same functional relationship, $\overrightarrow{\nu}_\kappa$ has a form that is identical to $\overrightarrow{\nu}_u$,

$$\overrightarrow{\nu}_\kappa \propto \mathcal{N}\left(\frac{\log \gamma_3}{m_u}, \frac{2}{m_u^2}\right). \tag{9}$$

The last message we are interested in is $\overrightarrow{\nu}_\omega$. The derivation process is similar as before. We evaluate the expectation and use Laplace approximation again.

$$\overrightarrow{\nu}_\omega \propto \exp\left(\mathbf{E}_{/q_\omega}[\log f(x, u, \kappa, \omega)]\right) \tag{10a}$$

$$\propto \underbrace{\exp\left(-\frac{m_u m_\kappa + \omega}{2} - \frac{m_x^2 + v_x}{\exp\left(m_u m_\kappa + \omega - \frac{\gamma_1}{2}\right)}\right)}_{g(\omega)} \tag{10b}$$

If we choose to expand $g(\omega)$ around $\gamma_4 = \left(m_x^2 + v_x\right)\exp\left(-m_\kappa m_u + \frac{\gamma_1}{2}\right)$, then $g'(\gamma_4) = 0$. This choice of expansion point again allows us to approximate the message as a Gaussian, i.e.,

$$\overrightarrow{\nu}_\omega \propto \mathcal{N}(\log \gamma_4, 1). \tag{11}$$

2. DERIVATION OF FREE ENERGY

Next we want to obtain a free energy expression for the GCV node $f(x, u, \kappa, \omega)$. The free energy functional is defined as

$$F[q] \triangleq U[q] - H[q] \tag{12a}$$

$$U[q] \triangleq -\mathbf{E}_{q_x q_u q_\kappa q_\omega}[\log f] \tag{12b}$$

$$H[q] \triangleq -\mathbf{E}_{q_x q_u q_\kappa q_\omega}[\log q]. \tag{12c}$$

Since each q is a Gaussian distribution, the entropy term $H[q]$ is the sum of all individual entropies, i.e.,

$$H[q] = \log\left((2\pi e)^2 \sqrt{v_x v_u v_\kappa v_\omega}\right). \tag{13}$$

Derivation of the average energy term $U[q]$ is a bit more involved. We can use the expectations that were derived in the previous section to evaluate it.

$$U[q] = \frac{\log 2\pi}{2} + \mathbf{E}_{q_u q_\kappa q_\omega}\left[-\frac{\kappa u + \omega}{2}\right] \tag{14a}$$

$$+ \mathbf{E}_{q_x q_u q_\kappa q_\omega}\left[-\frac{x^2}{2\exp(\kappa u + \omega)}\right] \tag{14b}$$

$$= \frac{1}{2}(\log(2\pi) + (m_\kappa m_u + m_\omega)) + \frac{1}{2}\left(m_x^2 + v_x\right)\gamma_2^{-1} - H[q] \tag{14c}$$

3. LOCAL COMPUTATION OF FREE ENERGY

We will now illustrate why the free energy functional can be computed as a sum of local contributions. Suppose we have a factorization of a global function $g(x) = \prod_{a=1}^{M} g_a(x_a)$, where a is some index set, x_a is a subset of variables that are arguments for g_a and each g_a is a stochastic factor. We will note the variables that are arguments of g_a with $N(a)$. If we accept mean-field assumptions for the recognition distribution $q = \prod_{i=1}^{N} q(x_i)$, then we see that

$$F[q] = \sum_{a=1}^{M} \underbrace{\int - \log g_a(x_a) \prod_{i \in N(a)} q(x_i) \mathrm{d}x_a}_{U_q[g_a]} + \sum_{i=1}^{N} \underbrace{\int q(x_i) \log q(x_i) \mathrm{d}x_i}_{-H[q_i]} . \tag{15}$$

This means each factor g_a contributes an average energy term and each q_i contributes an entropy term. As a result, the total free energy decomposes into a sum of local contributions from nodes and variables in the FFG.

4. REFERENCES

[1] Leo A. Aroian, "The probability function of the product of two normally distributed variables," *The Annals of Mathematical Statistics*, pp. 265–271, 1947.