Variational Message Passing (VMP)

Univariate Gaussian Distribution

Lets define the likelihood for N i.i.d samples each following a uniform distribution:

$$P(\mathbf{x}|\Omega) = \prod_{n=1}^{N} \mathcal{N}(x_n|\mu, \gamma^{-1})$$

where μ corresponds to the mean parameter and γ corresponds to the precision parameter ($\gamma = \sigma^{-2}$)

We rewrite the Normal in exponential family form

$$\begin{aligned}
\ln \mathcal{N}(x_n|\mu, \gamma^{-1}) &= \begin{bmatrix} \gamma\mu & -\gamma/2 \end{bmatrix} \begin{bmatrix} x_n \\ x_n^2 \end{bmatrix} + \frac{1}{2}\left(\ln\gamma - \gamma\mu^2 - \ln 2\pi\right) \\
&= \begin{bmatrix} \gamma x_n & -\gamma/2 \end{bmatrix} \begin{bmatrix} \mu \\ \mu^2 \end{bmatrix} + \frac{1}{2}(\ln\gamma - \gamma x_n^2 - \ln 2\pi) \\
&= \begin{bmatrix} -\frac{1}{2}(x_n - \mu)^2 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \gamma \\ \ln\gamma \end{bmatrix} - \ln 2\pi
\end{aligned}$$

We impose a Normal prior for the mean parameter and a Gamma prior for the precision parameter

$$\begin{aligned}
\ln P(\mu|m, \beta) &= \begin{bmatrix} \beta m & \beta/2 \end{bmatrix} \begin{bmatrix} \mu \\ \mu^2 \end{bmatrix} + \frac{1}{2}\left(\ln\beta - \beta m^2 - \ln 2\pi\right) \\
\ln P(\gamma|a, b) &= \begin{bmatrix} -b & a-1 \end{bmatrix} \begin{bmatrix} \gamma \\ ln\gamma \end{bmatrix} + a\ln b - \ln\Gamma(a)
\end{aligned}$$

Variational Approximation

$$Q(\mu, \gamma) = Q_\mu(\mu)Q_\gamma(\gamma)$$

$$\ln Q_\mu^*(\mu) = \left[\beta m + \mathbb{E}_{q_\gamma}(\gamma)\sum_{n=1}^{N} x_n \quad -\beta/2 - \mathbb{E}_{q_\gamma}(\gamma)\, N/2\right] \begin{bmatrix} \mu \\ \mu^2 \end{bmatrix} + f_\mu(\mu) + const$$

$$\ln Q_\gamma^*(\gamma) = \begin{bmatrix} -b - 1/2\sum_{n=1}^{N}(x_n^2 - 2x_n\mathbb{E}_{q_\mu}(\mu) + \mathbb{E}_{q_\mu}(\mu^2)) & a - 1 - 1/2 \end{bmatrix} \begin{bmatrix} \gamma \\ \ln\gamma \end{bmatrix} + f_\gamma(\gamma) + \text{const}$$

Therefore,

$$
\begin{aligned}
\ln Q_\mu^*(\mu) &= \left[\begin{array}{cc} \beta_\mu m_\mu & -\beta_\mu/2 \end{array}\right] \left[\begin{array}{c} \mu \\ \mu^2 \end{array}\right] + \frac{1}{2}\left(\ln \beta_\mu - \beta_\mu m_\mu^2 - \ln 2\pi\right) \\
\left[\begin{array}{c} \theta_1 \\ \theta_2 \end{array}\right] &= \left[\begin{array}{c} \beta_\mu m_\mu \\ -\beta_\mu/2 \end{array}\right] \\
f_\mu(\mu) &= \frac{1}{2}\ln 2\pi \\
A(m_\mu, \beta_\mu) &= \frac{1}{2}\left(-\ln \beta_\mu + \beta_\mu m_\mu^2\right) \\
A(\theta_1, \theta_2) &= \frac{1}{2}\left(-\ln(-2\theta_2) - 2\theta_2\left(\frac{\theta_1}{-2\theta_2}\right)^2\right) = \frac{1}{2}\left(-\ln(-2\theta_2) - \frac{\theta_1^2}{2\theta_2}\right) \\
\nabla A(\theta_1, \theta_2) &= \left[\begin{array}{c} -\frac{\theta_1}{2\theta_2} \\ -\frac{1}{2\theta_2} + \frac{\theta_1^2}{4\theta_2^2} \end{array}\right] \rightarrow \mathbb{E}_{q_\mu}\left(\mathbf{u}(\mu)\right) = \left[\begin{array}{c} m_\mu \\ \frac{1}{\beta_\mu} + m_\mu^2 \end{array}\right]
\end{aligned}
$$

$$
\begin{aligned}
\ln Q_\gamma^*(\gamma) &= \left[\begin{array}{cc} -b_\gamma & a_\gamma - 1 \end{array}\right] \left[\begin{array}{c} \gamma \\ \ln \gamma \end{array}\right] + a_\gamma \ln b_\gamma - \ln \Gamma(a_\gamma) \\
\left[\begin{array}{c} \theta_1 \\ \theta_2 \end{array}\right] &= \left[\begin{array}{c} -b_\gamma \\ a_\gamma - 1 \end{array}\right] \\
f_\gamma(\gamma) &= 1 \\
A(a_\gamma, b_\gamma) &= -a_\gamma \ln b_\gamma + \ln \Gamma(a_\gamma) \\
A(\theta_1, \theta_2) &= -(\theta_2 + 1)\ln(-\theta_1) + \ln \Gamma(\theta_2 + 1) \\
\nabla A(\theta_1, \theta_2) &= \left[\begin{array}{c} \frac{-(\theta_2+1)}{\theta_1} \\ -\ln(-\theta_1) + \Psi(\theta_2 + 1) \end{array}\right] \rightarrow \mathbb{E}_{q_\gamma}\left(\mathbf{u}(\gamma)\right) = \left[\begin{array}{c} \frac{a_\gamma}{b_\gamma} \\ \Psi(a_\gamma) - \ln(b_\gamma) \end{array}\right]
\end{aligned}
$$

Variational Message Passing

$$\mathbf{m}_{\gamma\to x_n} = \mathbb{E}_{q_\gamma(\gamma|a_\gamma,b_\gamma)}\left(\mathbf{u}_\gamma(\gamma)\right) = \left[\begin{array}{c} \mathbb{E}_{q_\gamma}(\gamma) \\ \mathbb{E}_{q_\gamma}(\ln\gamma) \end{array}\right] = \left[\begin{array}{c} -\frac{\delta g(a_\gamma b_\gamma)}{\delta a_\gamma} \\ -\frac{\delta g(a_\gamma b_\gamma)}{\delta b_\gamma} \end{array}\right] = \left[\begin{array}{c} \frac{a_\gamma}{b_\gamma} \\ \Psi(a_\gamma) - \ln(b_\gamma) \end{array}\right]^T$$

$$\mathbf{m}_{x_n\to\mu} = \left[\begin{array}{c} \mathbb{E}_{q_\gamma}(\gamma)x_n \\ -\mathbb{E}_{q_\gamma}(\gamma)/2 \end{array}\right] = \left[\begin{array}{c} \frac{a_\gamma}{b_\gamma}x_n \\ -\frac{a_\gamma}{2b_\gamma} \end{array}\right]$$

$$\phi_\mu^* = \left[\begin{array}{c} \beta m \\ -\beta/2 \end{array}\right] + \sum_{n=1}^{N}\mathbf{m}_{x_n\to\mu}$$

$$\mathbf{m}_{\mu\to x_n} = \mathbb{E}_{q_\mu(\mu|m_\mu,\beta_\mu)}\left(\mathbf{u}_\mu(\mu)\right) = \left[\begin{array}{c} \mathbb{E}_{q_\mu}(\mu) \\ \mathbb{E}_{q_\mu}(\mu^2) \end{array}\right] = \left[\begin{array}{c} -\frac{\delta g(m_\mu,\beta_\mu)}{\delta m_\mu} \\ -\frac{\delta g(m_\mu\beta_\mu)}{\delta\beta_\mu} \end{array}\right] = \left[\begin{array}{c} m_\mu \\ \frac{1}{\beta_\mu} + m_\mu^2 \end{array}\right]$$

$$\mathbf{m}_{x_n\to\gamma} = \left[\begin{array}{c} -\frac{1}{2}\left(x_n^2 - 2x_n\mathbb{E}_{q_\mu}(\mu) + \mathbb{E}_{q_\mu}(\mu^2)\right) \\ \frac{1}{2} \end{array}\right] = \left[\begin{array}{c} -\frac{1}{2}\left(x_n^2 - 2x_n m_\mu + \frac{1}{\beta_\mu} + m_\mu^2\right) \\ \frac{1}{2} \end{array}\right]$$

$$\phi_\gamma^* = \left[\begin{array}{c} -b \\ a-1 \end{array}\right] + \sum_{n=1}^{N}\mathbf{m}_{x_n\to\gamma}$$

LowerBound

$$\mathcal{L}(Q) = \sum_i \mathcal{L}_i$$

$$\mathcal{L}_{x_n} = \left[\begin{array}{c} \mathbb{E}_{q_\gamma}(\gamma)\mathbb{E}_{q_\mu}(\mu) \\ -\mathbb{E}_{q_\gamma}(\gamma)/2 \end{array}\right]^T \left[\begin{array}{c} x_n \\ x_n^2 \end{array}\right] + \frac{1}{2}\left(\mathbb{E}_{q_\gamma}(\ln\gamma) - \mathbb{E}_{q_\gamma}(\gamma)\mathbb{E}_{q_\mu}(\mu^2) - \ln 2\pi\right)$$

$$\mathcal{L}_\mu = \left[\begin{array}{c} \beta m - \beta_\mu m_\mu \\ -\beta/2 + \beta_\mu/2 \end{array}\right]^T \left[\begin{array}{c} \mathbb{E}_{q_\mu}(\mu) \\ \mathbb{E}_{q_\mu}(\mu^2) \end{array}\right] + \frac{1}{2}\left(\ln\beta - \beta m^2 - \ln\beta_\mu + \beta_\mu m_\mu^2\right)$$

$$\mathcal{L}_\gamma = \left[\begin{array}{c} -b + b_\gamma \\ a - a_\gamma \end{array}\right]^T \left[\begin{array}{c} \mathbb{E}_{q_\gamma}(\gamma) \\ \mathbb{E}_{q_\gamma}(\ln\gamma) \end{array}\right] + a\log b - \log\Gamma(a) - a_\gamma\log b_\gamma + \log\Gamma(a_\gamma)$$