Machine Learning Worksheet 12

Variational Inference

1 KL divergence

Problem 1: Compute the KL divergence between two Gaussian distributions $\mathcal{N}(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$ and $\mathcal{N}(\boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2)$ with diagonal covariance matrices.

Hint: If you use the facts you know about normal distribution, you can save yourself a lot of work before taking the straightforward path.

Problem 2: Consider that $p(\boldsymbol{x})$ is some arbitrary fixed distribution that we wish to approximate using an isotropic Gaussian distribution $q(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x} \mid \boldsymbol{\mu}, \boldsymbol{I})$ (covariance matrix is identity matrix).

By writing down the KL divergence $\mathbb{KL}(p\|q)$ and then differentiating w.r.t. $\boldsymbol{\mu}$, show that the optimal setting of the parameter is

$$\boldsymbol{\mu}^* = \arg\min_{\boldsymbol{\mu}} \mathbb{KL}(p\|q) = \mathbb{E}_p[\boldsymbol{x}]$$

2 Mean-field variational inference

Consider a very simple probabilistic model with a 2-D latent variable $\boldsymbol{z} \in \mathbb{R}^2$ and an observed variable $x \in \mathbb{R}$.

The prior over the latent variable is

$$p(\boldsymbol{z}) = \mathcal{N}(\boldsymbol{z} \mid \boldsymbol{0}, \boldsymbol{I}) = \mathcal{N}(z_1 \mid 0, 1) \cdot \mathcal{N}(z_2 \mid 0, 1),$$

and the likelihood is

$$p(x \mid \boldsymbol{z}) = \mathcal{N}(x \mid \boldsymbol{\theta}^T \boldsymbol{z}, 1),$$

where $\boldsymbol{\theta} \in \mathbb{R}^2$ is a known and fixed parameter.

Problem 3: Write down the true posterior distribution $p(\boldsymbol{z} \mid x)$.

Can the posterior be factorized over z_1 and z_2? (i.e. can it be expressed as $p(z_1 \mid x)p(z_2 \mid x)$?)

Problem 4: We approximate the true posterior using a mean-field variational distribution

$$q(\boldsymbol{z}) = q_1(z_1)q_2(z_2) = \mathcal{N}(z_1 \mid m_1, s_1^2) \cdot \mathcal{N}(z_2 \mid m_2, s_2^2)$$

Your task is to derive the optimal updates for q_1 and q_2.

Is $q(\boldsymbol{z})$ able to match the true posterior $p(\boldsymbol{z} \mid x)$?

Upload a single PDF file with your solution to Moodle by 29.01.2018, 9:59am CET. The solution must be typeset (using Latex, Word, etc.). Handwritten solutions are not accepted.