Gaussian approximation with moment matching, aka proj() operator in Expectation Propagation

Daniel Korzekwa (daniel.korzekwa@gmail.com)

11th March 2013

Overview

Moment matching is a technique for approximating a function $p(x)$ with a Gaussian distribution $\tilde{p}(x) \sim \mathcal{N}(m,v)$ by matching expectations $E[x]$ and $E[x^2]$, where

$$E_p[x] = E_{\tilde{p}}[x] \tag{1}$$

$$E_p[x^2] = E_{\tilde{p}}[x^2] \tag{2}$$

Then the mean and the variance of Gaussian approximation $\tilde{p}(x)$ are defined by

$$m_{\tilde{p}} = E_p[x] \tag{3}$$

$$v_{\tilde{p}} = E_p[x^2] - E_p[x]^2 \tag{4}$$

Following Thomas Minka [1] and Kevin P. Murphy [2], in a specific case of a function $p(x) = \frac{f(x)q(x)}{Z(m,v)}$, where $q(x) \sim \mathcal{N}(m,v)$ and $Z(m,v) = \int f(x)q(x)dx$, it can be shown that

$$E_p[x] = m + v\nabla_m logZ \tag{5}$$

$$E_p[x^2] = 2v^2\nabla_v logZ + v + m^2 + 2vm\nabla_m logZ \tag{6}$$

$$m_{\tilde{p}} = E_p[x] = m + v\nabla_m logZ \tag{7}$$

$$v_{\tilde{p}} = E_p[x^2] - E_p[x]^2 = v - v(\nabla_m^2 logZ - 2\nabla_v logZ)v \tag{8}$$

More generally, for any member of exponential family [3] $p(x|\eta) = h(x)g(\eta)exp\{\eta^T u(x)\}$, moments can be computed by differentiating the log partition function $A(\eta) = -lng(\eta)$

Example

Consider a function $p(\theta) = \frac{q(\theta)f(x|\theta)}{\int q(\theta)f(x|\theta)d\theta}$, borrowed from the Clutter Problem [1], where

$$q(\theta) \sim \mathcal{N}(\theta|m,v)$$

$$f(x|\theta) = (1-w)\mathcal{N}(x|\theta,1) + w\mathcal{N}(0,a)$$

In order to compute Gaussian approximation $\tilde{p}(\theta)$ to function $p(\theta)$, evaluated at the value of $x = 3, m = 15, v = 100, w = 0.4, a = 10$, we compute normalisation constant $Z(m, v)$ and derivatives of $log(Z)$ with respect to mean and variance.

$$Z(m, v) = (1 - w)\mathcal{N}(x|m, v + 1) + w\mathcal{N}(x|0, a)$$

$$\nabla_m logZ = (1 - w)\frac{1}{Z}\nabla_m \mathcal{N}(x|m, v + 1)$$

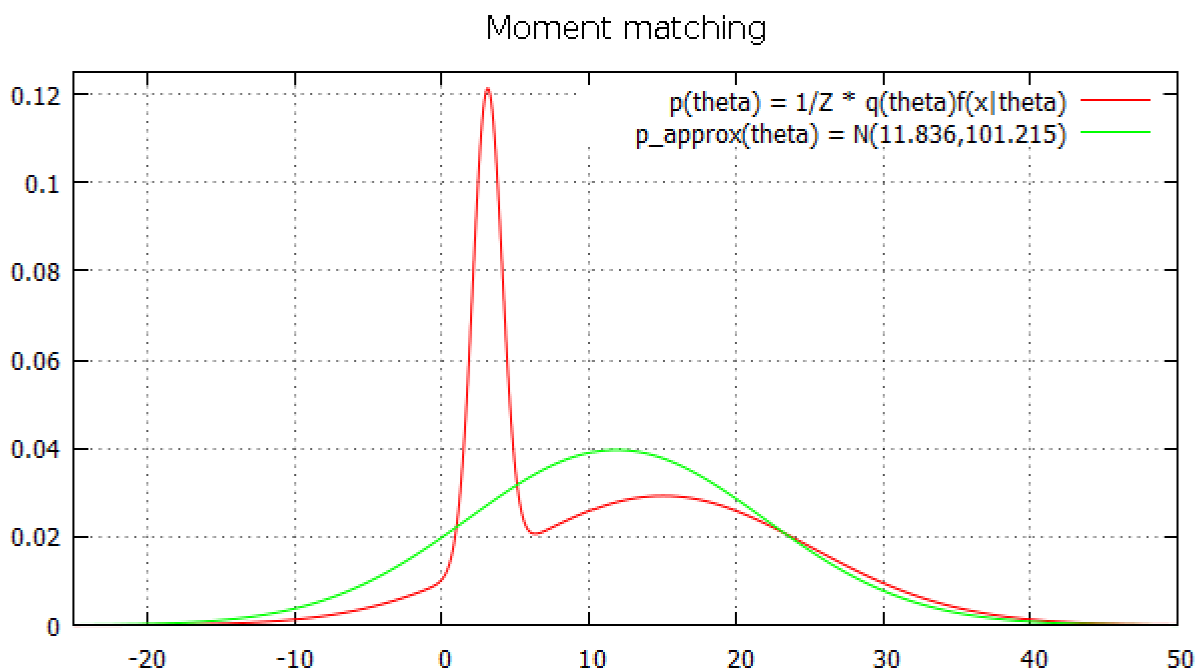
$$\nabla_v logZ = (1 - w)\frac{1}{Z}\nabla_v \mathcal{N}(x|m, v + 1)$$

Then, we can calculate Gaussian approximation $\tilde{p}(\theta)$ with equations 7 and 8

$$m_{\tilde{p}} = 11.8364$$

$$v_{\tilde{p}} = 101.21589$$

The following chart presents both $p(\theta)$ distribution and its Gaussian approximation $\tilde{p}(\theta)$.



Bayes-Scala toolbox [4] provides example implementation of Moment Matching for the Clutter Problem.

References

[1] Thomas P Minka. A family of algorithms for approximate Bayesian inference, 2001

[2] Kevin P. Murphy. From Belief Propagation to Expectation Propagation , 2001

[3] Exponential Family. http://en.wikipedia.org/wiki/Exponential_family

[4] Daniel Korzekwa. Bayes-Scala tool, MomentMatchingTest.scala