Moments of Truncated Gaussians

Benjamin Marlin, Mohammad Emtiyaz Khan, and Kevin Patrick Murphy
University of British Columbia, Vancouver, Canada

August 22, 2012

Given a Gaussian random variable x with mean μ and variance σ^2, we show how to compute

$$f(\mu, \sigma^2, \boldsymbol{\alpha}) = \int_l^h (ax^2 + bx + c)\mathcal{N}(x|\mu, \sigma^2)dx \tag{1}$$

and its derivatives with respect to μ and σ^2, where $\boldsymbol{\alpha} = [a, b, c, l, h]$ with all elements of the set being real-valued scalars. We introduce the notation $E_l^h[x^m|\mu, \sigma^2]$ to indicate the truncated expectation $\int_l^h x^m \mathcal{N}(x|\mu, \sigma^2)dx$, where m is a non-negative integer. We can then express $f(\mu, \sigma^2, \boldsymbol{\alpha})$ as $(aE_l^h[x^2|\mu, \sigma^2] + bE_l^h[x^1|\mu, \sigma^2] + cE_l^h[x^0|\mu, \sigma^2])$. The computation of $f(\mu, \sigma^2, \boldsymbol{\alpha})$ and its derivatives then follows from the computation of $E_l^h[x^m|\mu, \sigma^2]$ and its derivatives. We use $\phi(x)$ as shorthand for the standard normal probability density function and $\Phi(x)$ as shorthand for the standard normal cumulative distribution function. We define the standardized variables $\tilde{l} = (l - \mu)/\sigma$ and $\tilde{h} = (h - \mu)/\sigma$. It is easy to see that the computational cost of using a piecewise quadratic bound is only marginally higher than using a piecewise linear bound. This is due to the fact that the Gaussian CDF and PDF functions need only be computed twice each per piece for either class of bounds.

The truncated moments of orders zero, one and two are given below. These moments are closely related to the moments of a truncated and re-normalized Gaussian distribution.

$$E_l^h[x^0|\mu, \sigma^2] = \Phi(\tilde{h}) - \Phi(\tilde{l}) \tag{2}$$

$$E_l^h[x^1|\mu, \sigma^2] = \sigma(\phi(\tilde{l}) - \phi(\tilde{h})) + \mu(\Phi(\tilde{h}) - \Phi(\tilde{l})) \tag{3}$$

$$E_l^h[x^2|\mu, \sigma^2] = \sigma^2(\tilde{l}\phi(\tilde{l}) - \tilde{h}\phi(\tilde{h})) + (\sigma^2 + \mu^2)(\Phi(\tilde{h}) - \Phi(\tilde{l})) \tag{4}$$

We now give the derivatives of each truncated moment $E_l^h[x^m|\mu, \sigma^2]$ with respect to μ and σ^2. These are all the derivatives needed to complete the definition of the generalized EM algorithm (Algorithm 1).

$$\frac{\partial E_l^h[x^0|\mu, \sigma^2]}{\partial \mu} = \frac{1}{\sigma^2}\left(\phi(\tilde{l}) - \phi(\tilde{h})\right) \tag{5}$$

$$\frac{\partial E_l^h[x^0|\mu, \sigma^2]}{\partial \sigma^2} = \frac{1}{2\sigma^2}\left(\tilde{l}\phi(\tilde{l}) - \tilde{h}\phi(\tilde{h})\right) \tag{6}$$

$$\frac{\partial E_l^h[x^1|\mu, \sigma^2]}{\partial \mu} = \frac{1}{\sigma}\left(l\phi(\tilde{l}) - h\phi(\tilde{h})\right) + \Phi(\tilde{h}) - \Phi(\tilde{l}) \tag{7}$$

$$\frac{\partial E_l^h[x^1|\mu, \sigma^2]}{\partial \sigma^2} = \frac{l^2 + \sigma^2 - l\mu}{2\sigma^3}\phi(\tilde{l}) - \frac{h^2 + \sigma^2 - h\mu}{2\sigma^3}\phi(\tilde{h}) \tag{8}$$

$$\frac{\partial E_l^h[x^2|\mu, \sigma]}{\partial \mu} = \frac{1}{\sigma}\left((l^2 + 2\sigma^2)\phi(\tilde{l}) - (h^2 + 2\sigma^2)\phi(\tilde{h})\right) + 2\mu\left(\Phi(\tilde{h}) - \Phi(\tilde{l})\right) \tag{9}$$

$$\frac{\partial E_l^h[x^2|\mu, \sigma^2]}{\partial \sigma^2} = (l^3 + 2\sigma^2 l - l^2\mu)\frac{\phi(\tilde{l})}{2\sigma^3} - (h^3 + 2\sigma^2 h - h^2\mu)\frac{\phi(\tilde{h})}{2\sigma^3} + \Phi(\tilde{h}) - \Phi(\tilde{l}) \tag{10}$$