Truncated Gaussian Expectation Propagation for motion planning under hard constraints

When truncating a Gaussian distribution (multiplying with a heavyside function) we can approximate the remaining probability mass with another Gaussian, the mean and variance of which can be compute analytically. We can use this for approximate inference of motion trajectories under hard constraints: Collision and joint limit avoidance imply messages of the form of heavyside functions; using EP we can approximate the motion posterior.

Note: Herbrich has already an (unpublished) technical note on EP with truncated Gaussians (http://research.microsoft.com/pubs/74554/EP.pdf) So that's not so novel. But I've never seen it applied.

Problem: Let $x \in \mathbb{R}$ and $g(x) = e^{-x^2/2}$ and $\theta(x) = [x \geq 0]$ (the heavyside function). We want to compute a Gaussian approximation of $g(x)\theta(x)$, that is, the integrals

$$\int_z^\infty e^{-x^2/2}\, x\, dx \quad \text{and} \quad \int_z^\infty e^{-x^2/2}\, x^2\, dx \tag{1}$$

Norm:

$$\int_0^z e^{-t^2}\, dt = \frac{\sqrt{\pi}}{2}\, \text{erf}(z) \tag{2}$$

$$\int_z^\infty e^{-x^2/2}\, dx = \sqrt{2} \int_{z/\sqrt{2}}^\infty e^{-t^2}\, dt = \sqrt{\pi/2}\, [1 - \text{erf}(z/\sqrt{2})] \tag{3}$$

Mean:

$$\int_z^\infty e^{-x^2/2}\, x\, dx \tag{4}$$

$$= -\int_{-\frac{1}{2}z^2}^{-\infty} e^t\, dt = -\Big[e^t\Big]_{-z^2/2}^{-\infty} = -\Big[0 - e^{-z^2/2}\Big] = e^{-z^2/2} \tag{5}$$

Variance:

$$I_n(z,a) = \int_z^\infty e^{-ax^2}\, x^n\, dx \tag{6}$$

$$\frac{\partial}{\partial a} I_n(z,a) = \int_z^\infty e^{-ax^2}\, (-x^2)\, x^n\, dx = -I_{n+2}(z,a) \tag{7}$$

$$\frac{\partial}{\partial a} \text{erf}(\sqrt{a} z) = \frac{a^{-1/2} z}{2}\, \frac{2}{\sqrt{\pi}}\, e^{-az^2} \tag{8}$$

$$I_2(z,a) = -\frac{\partial}{\partial a} N(z,a) \tag{9}$$

Algorithm 1 Truncated Standard Gaussian

Input: z
Output: norm n, mean m, variance v
$n = \sqrt{\pi/2}[1 - \text{erf}(z\sqrt{2})]$
$m = \exp(-z^2/2)/n$
$v = 1 + zm - m^2$

$$= a^{-3/2}\frac{\sqrt{\pi}}{4}\,[1 - \text{erf}(\sqrt{a}z)] + a^{-1/2}\,\frac{\sqrt{\pi}}{2}\,[\frac{a^{-1/2}z}{\sqrt{\pi}}\,e^{-az^2}] \tag{10}$$

$$= a^{-3/2}\frac{\sqrt{\pi}}{4}\,[1 - \text{erf}(\sqrt{a}z)] + \frac{z}{2a}\,e^{-az^2} \tag{11}$$

$$I_2(z, \frac{1}{2}) = \sqrt{\pi/2}\,[1 - \text{erf}(z/\sqrt{2})] + z\,e^{-z^2/2} \tag{12}$$

Higher order moments... With (7) we can compute any higher order moments

Summary

$$\text{norm}N = \sqrt{\pi/2}\,[1 - \text{erf}(z/\sqrt{2})] \tag{13}$$

$$\text{mean}M = e^{-z^2/2}/N \tag{14}$$

$$\text{variance}V = 1 + zM - M^2 \tag{15}$$

$$\tag{16}$$

1 General case

We now have a n-dim Gaussian and heavyside function

$$f(y) = \mathcal{N}(y|a, A) \propto \exp\{-\frac{1}{2}(y-a)^\top\,A^{\text{-}1}\,(y-a)\} \tag{17}$$

$$\theta(y) = [[c^\top y + d \geq 0]] \tag{18}$$

where $[[\cdot]]$ is the indicator function. We transform this problem such that the Gaussian becomes a standard Gaussian and the constraint is aligned with the x-axis. We need two transformations for this: first a linear transform to standartize the Gaussian, then a rotation to align with the x-axis. Let $A = M^\top M$ be the Cholesky decomposition ($A^{\text{-}1} = M^{\text{-}1}M^{-\top}$) and we define $x = M^{-\top}(y-a)$. We have

$$f(x) = \exp\{-\frac{1}{2}x^\top x\} \tag{19}$$

$$t(x) = [[c^\top(M^\top x + a) + d \geq 0]] = [[v^\top x + z \geq 0]]\;, \tag{20}$$

$$v := Mc/|Mc|\;, \quad z := (c^\top a + d)/|Mc| \tag{21}$$

Note that we defined v to be normalized. (If $|Mc|$ is zero the truncation has no effect or zero likelihood, depending on whether $d > 0$ or $d < 0$, respectively.) We

Algorithm 2 Truncate Gaussian

Input: mean a, covariance A, constraint coeffs c, d
Output: mean b, covariance B
$M^T M = A$ *// Cholesky decomposition*
$z = (c^\top a + d)/|Mc|$
$v = Mc/|Mc|$
$R =$ rotation onto v *// as in equation (22)*
(m, v) = Truncated Standard Gaussian(z)
$b = M^\top R(m, 0, .., 0) + a$
$B = M^\top R \mathrm{diag}(v, 1, .., 1) R^\top M$

define a rotation that rotates the unit vector $\boldsymbol{e} = (1, 0, .., 0)$ onto v. In 4D we choose the rotation

$$R = \begin{pmatrix} v_1 & -v_2 & -v_3 & -v_4 \\ v_2 & v_3 & -v_4 & -v_1 \\ v_3 & v_4 & v_1 & -v_2 \\ v_4 & v_1 & v_2 & v_3 \end{pmatrix} \tag{22}$$

which generalizes to arbitrary dimensionality. We define $x' = R^{\text{-}1}x$. We have $v = R\boldsymbol{e}$ and

$$f(x') = \exp\{-\frac{1}{2}x'^\top x'\} \tag{23}$$

$$\theta(x') = [[v^\top R x' + z \geq 0]] = [[(R^{\text{-}1}v)^\top x' + z \geq 0]] = [[x'_1 + z \geq 0]] \tag{24}$$

That is, $\theta(x')$ truncates along the first axis in the x' coordinate system. Given the mean m and variance v of the z-truncated standard Gaussian, we have

$$f(x')\ \theta(x') \approx \mathcal{N}(x'|b', B') \tag{25}$$

$$b' = (m, 0, .., 0) \tag{26}$$

$$B' = \mathrm{diag}(v, 1, .., 1) \tag{27}$$

We undo the transformation $x' = R^{\text{-}1}M^{-\top}(y - a)$ and get the result

$$f(y)\ \theta(y) \approx \mathcal{N}(y|b, B) \tag{28}$$

$$b = M^\top R b' + a \tag{29}$$

$$B = M^\top R B' R^\top M \tag{30}$$

In summary, the mean and covariance of the truncated Gaussian is

gnuplot

```
a = .37
z = -.5
f(x) = exp(-a*x**2.)
h(x) = (sgn(x-z)+1)/2.

fnorm = sqrt(pi/a)
norm = a**(-1./2.)*sqrt(pi)/2.*(1.-erf(sqrt(a)*z))
mean = exp(-a*(z**2.))/(2.*a)/norm
sumOfSqr  = a**(-3./2.)*sqrt(pi)/4.*(1.-erf(sqrt(a)*z)) + z/(2*a)*exp(-a*z**2)
var = sumOfSqr/norm-mean**2

g(x) = exp(-.5/var*(x-mean)**2)

plot [-3:5] f(x),h(x),norm*g(x)/fnorm

print 'mean=',mean,',  sumOfSqr=', sumOfSqr/norm, ',  var=',var
```