Errata for "An Expectation Maximization Algorithm for Continuous Markov Decision Processes with Arbitrary Rewards"

Matt Hoffman, Nando de Freitas, Arnaud Doucet,
Jan Peters, Marc Toussaint

September 3, 2009

In an earlier version of this paper we made the claim that (Toussaint and Storkey, 2006) incorrectly uses the inverse dynamics to obtain the backwards transition model $p(x_n|x_{n+1})$. While using the inverse dynamics in this manner is indeed incorrect (see Klaas et al., 2006 for more details), after discussions with the authors we have determined that (Toussaint and Storkey, 2006) do *not* use this inverse dynamics and compute the Unscented Transform correctly. Because of this confusion, and in order to clarify this point we give a brief overview of their method below.

Assuming a transition model of the form

$$p(x_{n+1}|x_n) = \mathcal{N}(x_{n+1}; \phi(x_n), Q), \tag{1}$$

where we have dropped dependence on the actions u, we want to compute the backwards messages

$$\beta(x_n) = \int p(x_{n+1}|x_n)\, \beta(x_{n+1})\, dx_{n+1}. \tag{2}$$

The key here is that we are trying to compute this integral and *are not* trying to approximate $p(x_n|x_{n+1})$. If the messages are normally distributed $\beta(x_{n+1}) = \mathcal{N}(x_{n+1}; b, B)$ and the noise-free dynamics are linear $\phi(x_n) = Ax_n + a$ then we can carry out this integral in closed form to arrive at

$$\beta(x_n) \propto \mathcal{N}\big(x_n;\ A^{-1}(b-a),\ A^{-1}(B+Q)A^{-T}\big).$$

If, however, ϕ is nonlinear we can instead choose $2M+1$ points from $\beta(x_{n+1})$, propagating them backwards under ϕ^{-1} with a covariance of $\widetilde{\phi}^{-1}Q\widetilde{\phi}^{-T}$. Here $\widetilde{\phi}^{-1}$ is the local linearization of the inverse at each selected point. The Unscented Kalman Filter (UKF) update equations can then be used to obtain the Normal distribution approximating $\beta(x_n)$.

References

M. Klaas, M. Briers, N. de Freitas, A. Doucet, S. Maskell, and D. Lang. Fast particle smoothing: If i had a million particles. In *International Conference on Machine Learning (ICML)*, 2006.

M. Toussaint and A. Storkey. Probabilistic inference for solving discrete and continuous state Markov decision processes. In *International Conference on Machine Learning (ICML)*, 2006.