On the relation between certain stochastic control problems and probabilistic inference

Manfred Opper

Computer Science

- Control problem $\underbrace{\longrightarrow}_{\text{Bellman}}$ inference problem (Bert Kappen & Emanuel Todorov)

- Inference problem $\underbrace{\longrightarrow}_{\text{Var method}}$ control problems

- Things that could be learnt from this and possible extensions

Discrete times: MDPS

- Assume Markov process with transition probabilities $q(x'|x,u)$ tuned by a 'control' variable u.

- Try to minimise total expected costs

$$V_0(x) = \sum_{t=0}^{T} E^u \left[L_t(X_t, u_t) | x_0 = x \right]$$

- Define 'Value' of state x

$$V_t(x) = \sum_{\tau \geq t} E^u \left[L_\tau(X_\tau, u_\tau) | X_t = x \right]$$

- Solution via **Bellman equation**

$$V_t(x) = \min_u \left\{ L_t(x,u) + \sum_{x'} q_t(x'|x,u) V_{t+\Delta t}(x') \right\}$$

Continuous time: SDEs

- (Ito) stochastic differential equation for state $X(t) \in R^d$

$$dX(t) = (\underbrace{u(X_t, t) + f(X(t))}_{\text{Drift}} dt + \underbrace{D^{1/2}(X(t))}_{\text{Diffusion}} dW(t)$$

$W(t)$ vector of independent *Wiener processes*.

- Limit of discrete time process X_k

$$\Delta X_k \equiv X_{k+1} - X_k = (u_t + f(X_k))\Delta t + D^{1/2}(X_k)\sqrt{\Delta t}\, \epsilon_k \ .$$

ϵ_k i.i.d. Gaussian.

Continuous time ctd

- Try to minimise total expected costs

$$V_0(x) = \int_{t=0}^{T} E^u \left[L_t(X(t), u(t)) | X(0) = x \right]$$

- Define 'Value' of state x

$$V_t(x) = \int_{t}^{T} E^u \left[L_s(X(s), u(s)) | X(t) = x \right] \, ds$$

- Solution via **Hamilton - Jacobi - Bellman equation**

$$-\frac{\partial V_t(x)}{\partial t} = \min_u \left\{ L_t(x,u) + (u+f)^\top \nabla V_t(x) + \frac{1}{2} \mathrm{Tr}(D \nabla^\top \nabla) V_t(x) \right\}$$

- Specialise to

$$L_t(x,u) = \frac{1}{2}u(t)^\top R u(t) + U(x(t),t)$$

- Minimisation leads to $u_t = -R^{-1}\nabla V_t$

- and a a nonlinear PDE!

$$-\frac{\partial V_t(x)}{\partial t} = -\frac{1}{2}(\nabla V_t)^\top R^{-1}(\nabla V_t) + f^\top(x)\nabla V_t + \frac{1}{2}\mathrm{Tr}(D\nabla^\top\nabla)V_t + U(x,t)$$

Exact Linearisation (Kappen, 2005)

- Assume that $D = R^{-1}$ and using the transformation $V_t(x) = -\ln Z_t(x)$ we get the equation

$$\left\{ \frac{\partial}{\partial t} + \mathcal{L}^\dagger \right\} Z_t(x) = 0$$

with the **linear operator**

$$\mathcal{L}^\dagger = f^\top \nabla + \frac{1}{2} \mathrm{Tr}(D \nabla^\top \nabla) - U(x,t)$$

- and a **path integral representation**

$$Z_t(x) = E^{u=0} \left[e^{-\int_t^T U_\tau(X(\tau),\tau)\, d\tau} | X(t) = x \right]$$

Now all kinds of inference tricks apply!

Todorov's solvable MDPs (2006)

$$L_t(x,u) = \sum_{x'} \; q(x'|x,u) \ln \frac{q(x'|x,u)}{p(x'|x)} + U_t(x)$$

Bellman equation

$$V_t(x) = \min_u \left\{ U_t(x) + \sum_{x'} q(x'|x,u) \ln \left(\frac{q(x'|x,u)}{p(x'|x)} + V_{t+\Delta t}(x') \right) \right\}$$

The controlled transition probabilities:

$$q(x'|x,u) \propto p(x'|x) \; e^{-V_{t+\Delta t}(x')}$$

with $V_t(x) = -\ln Z_t(x)$ we get the linear equation (Todorov, 2005)

$$Z_t(x) = e^{-U_t(x)} \sum_{x'} p(x'|x) \; Z_{t+\Delta t}(x')$$

Relation to continuous case

- Short time transition probability

$$p\left(x', t+\Delta t | x, t\right) \propto \exp\left[-\frac{1}{2\Delta t}\left\|\Delta x - f(x)\Delta t\right\|_D^2\right]$$

  as $\Delta t \to 0$, with $\|F\|_D^2 = F^\top D^{-1} F$.

- Let p_g and p_f short term transition probabilities for Diffusion processes with drift g and f with **same diffusion** D. Then

$$\int p_g\left(x', t+\Delta t | x, t\right) \ln \frac{p_g\left(x', t+\Delta t | x, t\right)}{p_f\left(x', t+\Delta t | x, t\right)}\, dx' = \simeq \frac{1}{2}\left\|g(x,t) - f(x,t)\right\|_D^2\ \Delta t$$

The KL divergence for Markov processes

Consider probabilities $p(X_{0:T})$, $q(X_{0:T})$ over **entire paths** $X_{0:T}$.

The total KL divergence ....

$$KL\left[q(X_{0:T})\|p(X_{0:T})\right] = \int dx_{0:T}\ q(x_{0:T}) \ln \frac{q(x_{0:T})}{p(x_{0:T})}$$
$$= \sum_{k=1}^{T-1} \int dx_k\ q(x_k) \int dx_{k+1}\ q(x_{k+1}|x_k) \ln \frac{q(x_{k+1}|x_k)}{p(x_{k+1}|x_k)}$$
$$= \sum_{k=1}^{T-1} \int dx_k\ q(x_k)\ KL\left[q(\cdot|X_k)\|p(\cdot|X_k)\right]$$

.... is the expected sum of KLs for transition probabilities.

The global solution

The Kappen / Todorov control problems are of the form:

Minimise the **Variational free energy**

$$V_t(x) = KL\left[q(X_{t:T}) \| p(X_{t:T})\right] + E_q[\sum_{\tau \geq t} U_\tau(X_\tau, \tau)]$$

for fixed $X_t = x$ with respect to q. The **optimal** controlled probability over paths is

$$q_*(X_{t:T}) = \frac{1}{Z_t(x)}\, p(X_{t:T})\, e^{-\sum_{\tau \geq t} U_\tau(X_\tau, \tau)}$$

with the minimal cost (**free energy**)

$$V_t(x) = -\ln Z_t(x) = -\ln E_p\left[\, e^{-\sum_{\tau \geq t} U_\tau(X_\tau, \tau)} | X_t = x\right]$$

This looks like a HMM with '**likelihood**' $e^{-\sum_{\tau \geq t} U_\tau(X_\tau, \tau)}$.

Comments

- For **HMMs**

$$Z_t(x) = E_p\left[\, e^{-\sum_{\tau \geq t} U_\tau(X_\tau,\tau)} | X_t = x \right] \propto P(\text{future data}|X_t = x)$$

  fulfils a **linear backward equation**.

- The **posterior = controlled** process has transition probabilities

$$\frac{q(x_{t+1}|x_t)}{p(x_{t+1}|x_t)} = \frac{Z_{t+1}(x_{t+1})}{Z_t(x_t)}\, e^{-U_t(x_t,t)}$$

- Similar things happen for the continuous case:

$$\left\{\frac{\partial}{\partial t} f^\top \nabla + \frac{1}{2}\text{Tr}(D\nabla^\top\nabla) - U(x,t)\right\} Z_t(x) = 0$$

- Posterior is a diffusion with 'controlled' drift $u_t(x) = D\nabla \ln Z_t(x)$.

A 'real' likelihood for continuous time paths

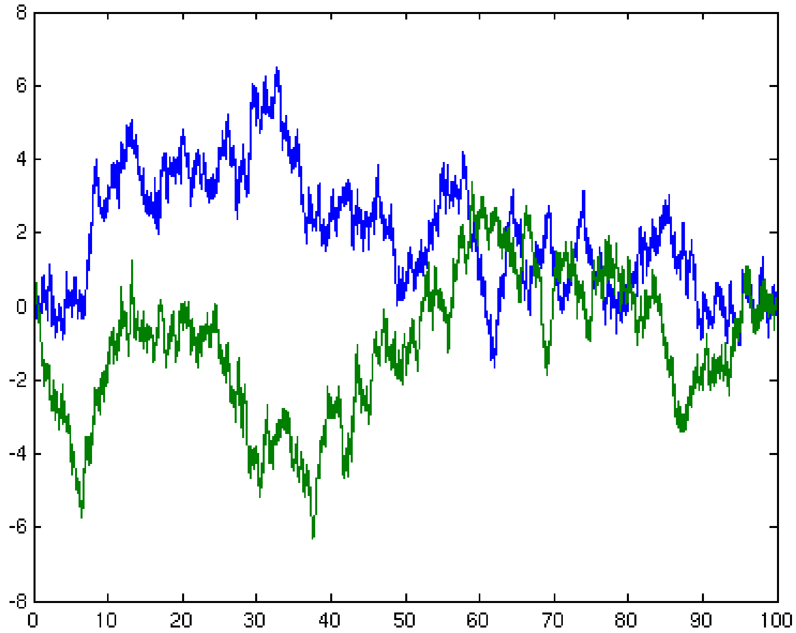
Consider an inhomogeneous Poisson process with rate function $U(X_t)$.

Then

$$\Pr\{\text{No event } \in [0\ T]\} = e^{-\int_t^T U_s(X_s, s)\, ds}$$

Application: Simulate diffusions with constraints

Wiener process with fixed endpoints $x(t = T) = 0$



Solution

$$u_t(x) = \frac{\partial \ln Z_t(x)}{\partial x}$$

$$\frac{\partial Z_t(x)}{\partial t} + \frac{1}{2}\frac{\partial^2 Z_t(x)}{\partial x^2} = 0$$

$$Z_t(x) = \delta(x)$$

is solved by

$$Z_t(x) \propto e^{-\frac{x^2}{2(T-t)}}$$

and leads to

$$u_t(x) = -\frac{x}{T-t}$$

for $0 < t < T$.

Diffusions with constraints on domain: $X(t) \in \Omega$.

1. **Method I:** Kill trajectory if $X(t) \in \partial\Omega$ for some t.

2. **Method II:** Simulate SDEs with drift $u_t(x) = \nabla \ln Z_t(x)$ where

$$Z_t = E\left[e^{-\sum_{\tau \geq t} U_\tau(X_\tau, \tau)} | X_t = x \right]$$

with $U = \infty$ if $x \notin \Omega$ and $U = 0$ else. Hence

$$\frac{\partial Z_t(x)}{\partial t} + \frac{1}{2}\frac{\partial^2 Z_t(x)}{\partial x^2} = 0$$

with $Z_t(x) = 0$ for $X(t) \in \partial\Omega$.

Possible approximations if we haven't got KL losses ?

- Approximate solution to control problem

$$V_0(x) = \int_{t=0}^{T} E^u \left[\frac{1}{2} u(t)^\top R u(t) + U(x(t), t) | X(0) = x \right]$$

  for general matrix R.

- Gaussian measure over paths $X_{0:T}$ induced by linear (approximate) posterior SDE (Archambeau, Cornford, Opper & Shawe - Taylor, 2007)

$$dX(t) = \{-A(t)X + b(t)\}\, dt + D^{1/2} dW$$

  as an approximation to

$$dX(t) = \{u(X, T) + f(X)\}\, dt + D^{1/2} dW$$

  Replace $u(X,T) \approx -A(t)X + b(t) - f(X)$

  $\rightarrow$ **nonlinear ODEs** for moments instead of linear PDEs !

Possible extensions to other losses

Simple non KL losses

$$KL(q\|p) \rightarrow \alpha KL(q\|p_1) - \beta KL(q\|p_2)$$

with $\alpha, \beta > 0$.

Use iterative method (CCCP style) upper bounding

$$-KL(q\|p_2) \leq -E_q[\ln \frac{q_n}{p_2}]$$

where q_n is the present optimiser.