What-If Reasoning with Counterfactual Gaussian Processes

Lorenzo Lagos
April 26, 2017

The objective of Schulam and Saria is to propose a model that can make counterfactual predictions about how time-series respond to sequences of actions taken in continuous time. The authors use Gaussian processes to model counterfactual trajectories within a time interval $[0, \tau]$ where multiple interventions take place. The model's inputs are previous actions and trajectories, which comprise *histories*. The outcomes of interest are future trajectories. In short, the goal is to make counterfactual predictions about the trajectory $\{Y_y : t \in [0, \tau]\}$ under future actions a at time t given the history H_t.

How this works is still a puzzle for me, so here is my interpretation of the model. First, the authors use MPP, which are distributions over sequences of points $\{T_i\}$ where each point is marked by an additional random variable X_i. This MPP is defined by a joint intensity function $\lambda^*(t, x)$ where the star superscript means that it is dependent on the history. This is then factorized into 1) intensity function $\lambda^*(t)$; 2) mark type distribution $p^*(z_y, z_a|t)$; 3) outcome model $p^*(y|t, z_y)$; and 4) action model $p^*(a|y, t, z_a)$. First, (1) describes the intensity at any point t regardless of the mark x [the product of (2)-(4) gives us $p^*(x|t)$]. Second, (2) tells us the distribution of outcomes/actions being active for a given point. Third, (3) depicts the distribution of outcomes for a given point and marker for active outcome. Finally, (4) describes the distribution of actions for a given point, outcome, and marker for active action. The authors then parametrize (3) as a mixture of Gaussian densities, and then assume: A3) the other 3 densities do not depend on the potential trajectories; and A4) something else I am unable to interpret (why don't the authors explain this?). In any case, at least A3 is a very strong assumption since it implies that the action model does not depend on unobserved information that is correlated with future potential outcomes.

In my project, I use value-function approximation on a discretized grid. Thus, I deal with optimal dynamic treatment in discrete time with a single outcome. This paper, instead, focuses on continuous time with trajectories as outcomes. Rather than finding the policy (mapping states to actions) that maximizes some reward, the authors want to predict the effect of actions on future observations. The model is tested with simulations and observational data. First, evaluating counterfactual predictions can only be done with simulations because accuracy cannot be measured otherwise (lacking a randomized experiment). Second, with observational data, model fit on held out data against two baselines can be used to check for any significant improvements. I may explore similar analyses in my project.