**Notes on : A bayesian foundation for learning under uncertainty**

 : sensory input

Generative model to predict next inout

Variables

binary, continuous, continuous

Distinction between states (variables that change quickly) and parameters (variables that change slowly or not at all)

If the brain uses a Bayesian scheme, it is likely that it relies on approximate bayesian inference.

Mean-field approximation within a variational scheme furnishes an efficient solution with biological plausibility and interpretability.

**Generative model description**

**Aim : Inversion**

**Variational inversion**

We make minimal assumptions about the form of the approximate posteriors by following the maximum entropy principle : given knowledge of, or assumptions about, constraints on a distribution, the least arbitrary choice of distribution is the one that maximizes entropy.

To keep the description of the posteriors simple and biologically plausible, we take them to be characterized only by their first 2 moments

First level, Bernoulli. Second level, Gaussian.

**Online inference**

We avoid the circular problem of approximating subsets consecutively until convergence by using the hierarchical structure : because information about the last event cannot have reached higher level, we use the sufficient statistics of the previous time and proceed upward (propagating information from lower to higher levels)

For continuous states, the problem of finding a Gaussian approximation can be reformulated as finding a quadratic approximation to the variational energies.

Solution chosen : taylor expand at the previous mode (in time).

Solution motivated by : (i) bears resemblance to RL update rules (ii) computational efficacy

Closed form update rule

**The update equations**

*Remark on notations*: Hat notation referring to prediction

Prediction of sufficient statistics from previous time higher order before seeing the input.

Those predictions are not classical predictions in the sense that they don’t follow the top-down generation from the top. They constitute predictions from the posterior the time just before.

Quadratic approximation to variational energies : ok.