Thompson Sampling for Exchangeable Array with Random Function Prior

Dongwoo Kim

Abstract

This is abstract.

1 Exchangeable Array

Definition 1.1 (Exchangeable 2-array)**.** A random 2-array (X_{ij}) is called separately exchangeable if

$$(X_{ij}) \stackrel{\mathrm{d}}{=} (X_{\pi(i)\pi'(j)}) \tag{1}$$

for every pair of permutations π, π' of $\mathbb{N}$.

Theorem 1.1 (Aldous)**.** *A random array (X_{ij}) is separately exchangeable if and only if it can be represented as follows: There is a random function $F : [0,1]^3 \to \mathcal{X}$ such that*

$$(X_{ij}) \stackrel{d}{=} (F(U_i, V_j, U_{ij})) \tag{2}$$

where $(U_i)_{i\in\mathbb{N}}$, $(V_j)_{j\in\mathbb{N}}$, $(U_{ij})_{i,j\in\mathbb{N}}$ are two sequences and an array of i.i.d. Uniform[0,1] random variables, which are independent of F.

2 Model

Based on Theorem 1.1, we define a Bayesian model for exchangeable arrays.

$$(X_{ij}) \stackrel{\mathrm{d}}{=} (F(U_i, U_j, U_{ij})) = H(U_{ij}, \Theta(U_i, V_j)) \tag{3}$$

$\Theta : [0,1]^2 \to W$ and $H : [0,1] \times W \to \mathcal{X}$. For real valued output varialbe X_{ij}, we define H an inverse CDF of gaussian distribution with mean parameter $\Theta(U_i, V_j)$, and variance σ^2.

$$\Theta \sim \mathcal{GP}(0, \kappa) \tag{4}$$

$$U_1, U_2, ... \sim \text{Uniform}[0,1] \tag{5}$$

$$V_1, V_2, ... \sim \text{Uniform}[0,1] \tag{6}$$

$$W_{ij} = \Theta(U_i, V_j) \tag{7}$$

$$X_{ij}|W_{ij} \sim \mathcal{N}(X_{ij}|W_{ij}, \sigma^2_{\mathcal{X}}) \tag{8}$$

3 Posterior Inference

Sampling (U_i) and (V_j) can be done with slice sampling. Hyper parameter of Gaussian process is also optimised given (U_i) and (V_j).

4 Thompson Sampling

Given U and V, the posterior predictive distribution of GP can be analytically computed, of which mean function is a smooth function over the product space.

Posterior predictive distribution of X on unobserved pair U_*, V_* is

$$p(X_*|U_*, V_*, \mathbf{x}, \mathbf{U}, \mathbf{V}) = \mathcal{N}(\mu_*, \sigma_*^2) \tag{9}$$

$$\mu_* = \mathbf{K}_{*N}(\mathbf{K}_N + \sigma^2 I)^{-1}\mathbf{x} \tag{10}$$

$$\sigma_*^2 = K_{**} - \mathbf{K}_{*N}(\mathbf{K}_N + \sigma^2 I)^{-1}\mathbf{K}_{N*} + \sigma^2, \tag{11}$$

where we let $\mathbf{x}$ be the sequence of observations so far, and $\mathbf{K}$ be the corresponding kernel matrix.

To apply Thompson sampling algorithm to this model, at each time t, we have to sample unobserved entry X_{ij} of pair U_i and V_j from this posterior. Again, the number of required posterior sampling increases quadratically as the number of U and V increases. Instead of sampling individual pairs, is there a way to sample a posterior function in an analytic form, so that we do some maximisation on this function? For example, if we place a prior over the hyper-parameters of the kernel function and sample the hyper-parameters [2], and finding a maximum point (U_*, V_*) of the posterior mean function given these hyper-parameters, then can we say that this pair U_* and V_* is a new sample by Thompson sampling? (even this posterior function does not take into account the variance of an individual observation) (or the most close pair U'_* and V'_* from U_* and V_* if corresponding U_* and V_* does not exist) Does this help to reduce the computation complexity of the posterior sampling? Is it better than the UCB-based algorithm? [1]

References

[1] Andreas Krause and Cheng S Ong. Contextual gaussian process bandit optimization. In *Advances in Neural Information Processing Systems*, pages 2447–2455, 2011.

[2] Radford M Neal. Monte carlo implementation of gaussian process models for bayesian regression and classification. *arXiv preprint physics/9701026*, 1997.