Part B Project Proposal

Ramsay Davis (Exeter)

Proposal

The purpose of this project is to respond to an open problem proposed in a short paper by Marco Mussi, Simone Drago and Alberto Maria Metelli [4]. Given sequential random variables $X_t \subseteq \mathbb{R}^d$, and $Y_t = Ber(\langle \theta_*, X_t \rangle)$ such that θ_* is an unknown parameter vector in $\mathbb{R}^d$, and $X_n = f(X_1, \ldots, X_{n-1}, Y_1, \ldots, Y_{n-1})$, within a Reproducing Kernel Hilbert Space (RKHS), can we construct a tight bound on the confidence set for a penalised least squares estimator of θ_*, hereafter denoted $\hat{\theta}_t$? This can be shown to be equivalent to the open problem, which is phrased in the language of kernelised functions. The result can also be used to improve sequential decision-making algorithms in certain contexts [3]. I will attempt to solve the problem from two different angles. Firstly, to extend the work done in Abbasi-Yadkori's PhD thesis [1], which includes the bound in the subgaussian case:

$$\mathbb{P}\Big(\|\hat{\theta}_t - \theta_*\|_{\overline{V_t}} \leq R\sqrt{\frac{(t-1)L^2}{\lambda} + 2\log\left(\frac{1}{\delta}\right)} + \lambda^{1/2} S \Big) \geq 1 - \delta \tag{1}$$

For $\delta \in (0, 1)$, where $\overline{V_t}$ is the regularised design matrix underlying the covariates. These results can be extended to the Bernoulli setting using results regarding self-concordant functions [2]. Additionally, I will explore using sequential likelihood ratios, which can easily be applied to the finite Bernoulli case [3]. However, extending it to infinity may prove more challenging. I have worked with David Janz in creating this proposal, and he has agreed to supervise this project.

References

[1] Yasin Abbasi-Yadkori. *Online Learning for Linearly Parametrised Control Problems*. PhD thesis, University of Alberta, 2012.

[2] Francis R. Bach. Self-concordant analysis for logistic regression. *CoRR*, abs/0910.4627, 2009.

[3] Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2020.

[4] Marco Mussi, Simone Drago, and Alberto Maria Metelli. Open problem: Tight bounds for kernelized multi-armed bandits with bernoulli rewards, 2024.