EE 223: Stochastic Estimation and Control **Spring 2007**

Lecture 27 — May 01

Lecturer: Venkat Anantharam *Scribe: none*

This lecture was not scribed.

We discussed identification for discrete time Markov chains with finite state space $\mathcal{X}$ with finite parameter space Θ, following Chapter 12 of the book of Kumar and Varaiya. We proved that under the assumption that $p_{ij}(\theta) > 0$ for all $i, j \in \mathcal{X}$ and $\theta \in \Theta$ the maximum likelihood parameter estimate converges within a finite random time to the true parameter, almost surely. We then studied identification for controlled finite state Markov chains with finite parameter set and finite set of controls. We illustrated the issue that if the optimal controls do not result in the product space of states and controls being explored then we may fail to identify the true parameter, even if we have $p_{ij}(u, \theta) > 0$ for all i, j, all u, and all θ. We may show this by a simple example: suppose you have two coins, one of which is fair, and the other is biased. You know which coin is fair and which is biased. However, for the biased coin you do not know if the probability of it coming up heads is $\frac{1}{3}$ or $\frac{2}{3}$. You would like to decide which coin to toss at each time, with the goal being to maximize the long term average number of heads you see. A certainty equivalence strategy based on a maximum likelihood parameter estimation rule has a positive probability of tossing the biased coin only a finite number of times, even if it is the better coin. Thus even if the biased coin is better than the fair coin, there is a positive probability that we end up believing the it is worse.

We returned to the ARMAX model and discussed the used of the extended least squares parameter estimator in a certainty equivalence based control loop for minimum variance control (the self-tuning regulator). We discussed the result in the paper of Guo and Chen, "The Åström-Wittenmark self-tuning regulator revisited and ELS-based adaptive trackers". This was a breakthrough paper, proving the consistency of identification and the asymptotic optimality of performance of this adaptive control scheme (with a somewhat modified version of ELS) under a positive real assumption on the polynomial defining the moving average of the noise.