

**Abstract:** In this talk, we present a brief overview of the theory of self-normalized processes as used in the analysis of least squares over causally dependent data. The self-normalized martingale arises as one of two key terms in the decomposition of the least squares parameter estimation error. We first review the relevant definitions and assumptions required to derive a (almost) time-scale invariant bound on this self-normalized martingale term. We then provide an interpretation of the result, along with a roadmap for how it is applied in the analysis of system identification. We conclude by sketching a proof for the result by employing an approach known as pseudo-maximization.

**Keywords:** System Identification, Least Squares Estimator, Non-asymptotic Statistics, Self-Normalized Processes