EEL 6935 **_Stochastic Control_** Spring 2020

Control of systems subject to noise and uncertainty

Prof. Sean Meyn, `meyn@ece.ufl.edu` MAE-A 0327, Tues 1:55-2:45, Thur 1:55-3:50

> The first goal is to learn how to formulate models for the purposes of control, in applications ranging from finance to power systems to medicine. Linear and Markov models are chosen to capture essential dynamics and uncertainty. The course will provide several approaches to design control laws based on these models, and methods to approximate the performance of the controlled system. In parallel with these algorithmic objectives, students will be provided with an introduction to basic dynamic programming theory, closely related stability theory for Markovian[1] and linear systems, and simulation and stochastic approximation concepts underlying reinforcement learning.
>
> It is intended for graduate students who have some background in control and stochastic processes. Experience with *Matlab* or *Python* is essential.

Why do we need noisy models? When you introduce the word "stochastic" to control, this just means that you are bringing in a larger range of tools for understanding how to control systems, and evaluate their performance. Name a tool from probability, and you have something useful for control synthesis. In particular, there is the question of *information.* This may mean the data available for control, or information about the system to be controlled. There may be variables of interest that are not directly observed, so we will want to estimate. Tools to be applied include nonlinear filtering and stochastic approximation (a foundation of reinforcement learning).

Office hours Joel Mathias: Weds & Thurs, 4:00-5:00 p.m. in 484 NEB.
Sean Meyn: Tues, 3:00-4:00 p.m. in 455 NEB. We can be reached for questions by electronic mail at joel.mathias@ufl.edu and meyn@ece.ufl.edu (or canvas).

Exams, homework, and grading Homework problems will be assigned on a ∼bi-weekly basis (more frequent at the start), to be uploaded to canvas before lecture on the date due. They will be graded and returned the following week. *Late homework cannot be accepted*.

There will be two evening midterm exams, February 26 and April 22, from 7:20 - 8:50 p.m. You will be allowed *one* sheet of notes ($8\frac{1}{2} \times 11$; both sides) in the first exam, and *two* in the second. Otherwise, the exams are closed-book and closed-notes.

Note that the first week of classes coincides with **Bayes Comp 2020**, an international biennial conference held for the first and only time in Gainesville! It will take place in the Reitz Union at the University of Florida

> The conference and the section both aim to promote original research into computational methods for inference and decision making and to encourage the use of frontier computational tools among practitioners, the development of adapted software, languages, platforms, and dedicated machines, and to translate and disseminate methods developed in other disciplines among statisticians.

The topics are very relevant to this course!

Grading scheme: Homework problems will count 20%[2], the midterm exams 60%, and the final project will count 20% towards the final grade in the course.

[1] A *Markov process* is nothing more than a nonlinear state space model subject to noise.

[2] I encourage collaboration on homework!

References: The following are available free on-line (send your thanks to CUP):

- ⊙ S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability.*
  www.meyn.ece.ufl.edu/archive/spm_pubs.html
- ⊙ S. P. Meyn, *Control Techniques for Complex networks.*
  www.meyn.ece.ufl.edu/archive/spm_pubs.html

The following are valuable background (send your thanks to Profs. Hajek and van Handel):

- ⊙ B. Hajek, *Exploration of Random Processes for Engineers.*
  www.ifp.illinois.edu/~hajek/Papers/randomprocesses.html Review: $(\Omega, \mathcal{F}, \mathsf{P}) \star \mathsf{P}(A) \star \mathsf{E}[X \mid Y]$
- ⊙ R. Van Handel, *Lecture Notes on Hidden Markov Models.*
  web.math.princeton.edu/~rvan/orf557/ (hmm080728.pdf 20-Jun-2018)

The following textbooks are of value, but not needed to follow the course.

- ⊙ D. Bertsekas and J. Tsitsiklis, *Neuro-Dynamic Programming* (see also Sutton's new book on reinforcement learning).
- ⊙ D. Bertsekas and S. Shreve, *Stochastic Optimal Control: The Discrete-Time Case.*
  web.mit.edu/dimitrib/www/soc.html
- ⊙ P. R Kumar and P. Varaiya, *Stochastic systems: Estimation, identification, & adaptive control.*

Course Outline:

I. Nonlinear State Space Models

1) Overview & examples. Review of concepts from optimal control
2) Markov models and more examples
3) Lyapunov theory for stability and performance
4) Numerical techniques and Monte-Carlo for performance estimation

II. Optimal Control

1) Controlled Markov models and MDPs.
2) Approximate dynamic programming.
3) Numerical techniques: Policy and value iteration; LP formulation.
4) Partial information (belief state). Multi-armed bandits (UCB heuristic).

First Midterm Exam

5) Linear quadratic optimal control.

III. Adaptation and Learning

1) Simulation and stochastic approximation: theory & applications
2) Reinforcement learning and approximate dynamic programming, including a superficial look at actor-critic methods.
3) TD Learning.
4) Q Learning

Second Midterm Exam

Review and thoughts for the future