Preface

In many scientific activities, observations are made on moving objects and they are generally corrupted by random errors. The theory of filtering provides means of filtering out the "noise" in the observations and extracting from them the most precise information about the moving object, which itself may or may not be moving in a random fashion. Then the prediction step follows. On the basis of information about the past behavior of the object, one tries to predict its future path. Two simple examples of such moving objects are the motion of a rocket and the changing price of a commodity in the stock market.

The theory of filtering and prediction is highly developed and a huge literature exists that deals with numerous aspects of the subject. Even though the theory at the advanced level requires a basic knowledge of advanced analysis and probability and is highly technical, several texts have been written at the undergraduate level. At that level, the compromises that have been made by other authors seem to be somewhat unnecessary. We feel that a fairly rigorous treatment of the basic ideas can be given in such a way that the interest of a serious undergraduate student can be sustained.

The material in the current text has been offered by us a few times in the past as a one-semester course. Based on the experience, we designed the text basically to consist of two parts: the first three chapters and the five remaining ones.

The first three chapters deal with an introduction to the filtering theory of Markov chains with discrete time and space; interpolation and prediction questions are also introduced. In particular, we treat Hidden Markov Models. Even though one does not need any heavy machinery here, all the definitions are carefully made and the treatment is quite rigorous, yet in this "simple" setting the problems can be quite challenging. We believe there are enough exercises to illustrate the main ideas and also to challenge a strong student. The main purpose, though, is to develop the main ideas. Some of the exercises have an asterisk to indicate that the later text will use the contents of those exercises. Altogether in the book there are about 85 starred exercises and 166 additional ones.

The next five chapters form the second part of the book. This part requires a more mature attitude and more commitment than normally expected from average undergraduates. However, our experience shows that the priming that the student gets by going through the first three chapters does lead to this kind of maturity and commitment. In the Spring semester of 2006 we had ten students finish the course; three of them were graduate students from three different departments and the remaining seven were undergraduates mostly from the department of mathematics. One of the best students turned out to be a junior from our Lower Division.

Chapter 4 introduces the "abstract" notion of conditional expectation. The introduction of this notion in terms of L_2-theory seems most natural in the context of filtering and prediction. We have made a rigorous presentation of the ideas without making it too technical. All the "usual" elementary definitions of conditional probability and expectation are shown to be special cases of this general setting. This chapter lays the foundation for the development of the theory in continuous time and space in the next four chapters.

Chapter 5 deals with discrete time Markov chains whose state space is $\mathbb{R}^d$. The discrete Kalman filter for $d = 1$ and $d \geq 1$, as well as linear filtering, are discussed in this chapter. Special attention is paid to so-called Hidden Markov Models. We also derive general filtering equations for Markov chains whose transition probability functions have a density function (kernel) with respect to Lebesgue measure.

Even though Lebesgue measure or even Riemann integration are not introduced in this text, those unfamiliar with these notions may simply assume the kernel to be jointly continuous. Once integration of continuous functions is understood, the rest is easy to follow and is self-contained.

Chapter 6 deals with continuous time filtering in the important case of the Wiener process. The exercises are designed to help the student with the development of the ideas while minimizing technicalities without sacrificing rigor. Stochastic integral with respect to the Wiener process is defined only for smooth nonrandom summands for obvious reasons. Some of the exercises could be challenging for even the strongest students without helpful hints from the instructor.

Stationary sequences (wide sense) are introduced in Chapter 7. We decided to include an elementary case of the Bochner-Khinchin representation theorem for positive definite functions since it plays an important role in this context. In addition to filtering such sequences we have included some optional sections on the law of large numbers and the spectral representation of stationary sequences. These sections may be skipped if time becomes a factor.

Finally, Chapter 8 contains an introduction to the prediction of stationary sequences. We only deal with autoregressive and autoregressive-moving average sequences. Some ideas from the theory of complex analysis are included rather than referenced in order to make the text self contained as far as possible.

The reader who would like to pursue further study on related topics will certainly benefit from consulting references [1] through [7].

We are thankful to our students for their patience because the text was under development for a number of years, which necessitated changes and revisions. We hope that this is a version that will be of sufficient interest to other colleagues and we hope to get comments from them. We already got quite a few comments and deep observations from J. Baxter, who taught this course in Spring 2007 and whose contribution is greatly appreciated.

B. Fristedt, N. Jain, N. Krylov, May 2007.