ECE499/CS499 Statistical Signal Processing (Fall 2024)

4 credits, MW 10:00-11:50am, Merryfield Hall 112

1. **Instructor**: Prof. Jinsub Kim (jinsub.kim@oregonstate.edu)
2. **Target audience**:
   a. *Engineering* students who want to learn fundamentals of statistical inference and machine learning as well as how they are applied to many engineering problems
   b. *Computer Science* students (interested in *machine learning*) who want to learn about classical theory of statistical inference and the statistical perspective of machine learning (note that machine learning problems are essentially statistical inference problems!)
3. **Course description:**

Statistical inference problems exist in almost every engineering discipline. For instance, in digital communications, the receiver needs to *detect* a sequence of bits sent by the transmitter based on the received signal, which is corrupted by additive random noise (see the right figure). In controlling an unmanned aerial vehicle (UAV), a localization technique is needed to *estimate* UAV's location and speed based on noisy sensor measurements. For an autonomous driving system, we need to equip it with techniques to *detect* and *classify* objects in its environment (e.g., pedestrians) based on LiDAR and vision sensor measurements. All these problems essentially boil down to the following question:

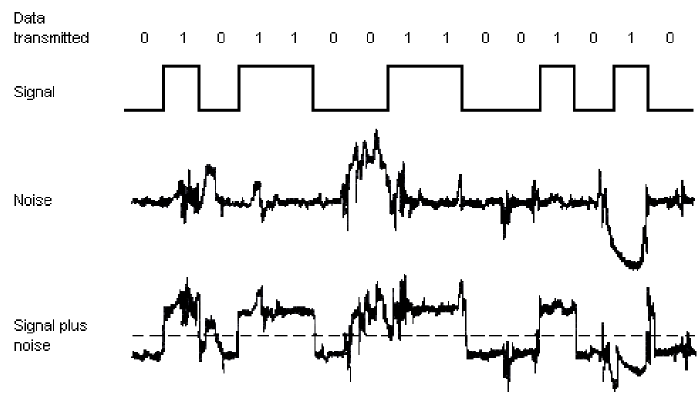

Figure 1. Transmitted signal (top), random channel noise (middle), and random noisy signal at the receiver (bottom) due to the additive channel noise.

"Given an observation of random variable X (e.g., UAV sensor data), how can we estimate the value of an unknown Y (e.g., UAV's location) accurately?"

At the successful completion of this course, students will be able to answer the above question.

4. **Course topics:**

The course consists of two parts. In **Part 1 (Detection and Estimation)**, assuming that the probability distribution of X and Y is known, we will discuss optimal detection and estimation rules that we can use to draw an inference about Y based on observation of X optimally with respect to some popular criteria. The topics in Part 1 include Bayesian detection (MAP detector), Non-Bayesian detection (Likelihood ratio test and Neyman-Pearson Lemma), Bayesian estimation (MMSE estimator and linear MMSE estimator), Non-Bayesian estimation (Maximum likelihood estimator), and various applications in engineering.

In **Part 2 (Machine Learning)**, we will discuss the same question, but in the setting where the probability distribution of X and Y is *unknown*; instead, we are given some samples from the distribution (the so-called *training dataset*), from which we can learn about the statistical dependency between X and Y.

This is a typical supervised learning formulation considered in machine learning and has been popularly employed for many engineering problems due to the difficulty of coming up with an accurate probability model for complex data. In this course, we will put emphasis on the *statistical perspective* of machine learning; the main discussion will be about how the ideas of optimal detection and estimation (learned in Part 1) can be adapted easily for solving machine learning problems. The topics include the bias-variance tradeoff, parametric approaches (linear regression, logistics regression, neural network), non-parametric approaches (k-Nearest Neighbor), and cross-validation.

5. **Prerequisite**:
    a. Probability (ECE 353) **or** Introduction to Statistics (ST 314)
    b. Familiarity with Python **or** MATLAB (students will be asked to implement some algorithms and perform experiments using either Python or MATLAB; each student can choose)

6. **Course logistics:**

   There will be weekly homework (40%), a take-home midterm (30%), and a final project (30%). The final project will be a guided project that will provide you with an opportunity to develop and evaluate statistical inference techniques for real-life applications.

7. **Textbooks**: Detailed lecture slides will be provided. In addition, we will use the following two resources, which are available online for free.

    a. Lecture notes by A. V. Oppenheim and G. C. Verghese, *MIT EECS 6.011 Introduction to Communication, Control, and Signal Processing (Spring 2010) (Available through MIT OpenCourseWare)*

    b. G. James, D Witten, T. Hastie, and R. Tibshirani, "*An Introduction to Statistical Learning with Applications in R*", 2nd Ed., Springer, 2021 (PDF available at https://www.statlearning.com)