ECE 7251: Signal Detection and Estimation

Spring 2002

Prof. Aaron Lanterman
Georgia Institute of Technology

Lecture 1, 1/1/02:
Introduction

What is ECE7251 About?

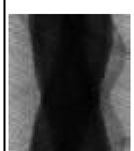
- Goal: Suck useful information out of messy data
- Strategy: Formulate probabilistic model of data y, which depends on underlying parameter(s) $\boldsymbol{q}$
- Terminology depends on parameter space:
  - Detection (simple hypothesis testing):
    - $q \in \{0,1\}$, i.e. 0=target absent, 1=target present
  - Classification (multihypothesis testing):
    - $q \in \{0,1,\ldots,M\}$, i.e. $q \in \{$DC-9, 747, F-15, MiG-31$\}$
  - Estimation
    - $q \in R^n$, C^n, etc. (not too hard)
    - $q \in L^2(R)$, (square-integrable functions), etc. (harder)

A Bit More Termonology

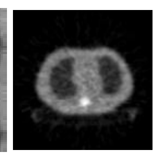
- Suppose $\boldsymbol{q}=(\boldsymbol{q}_1, \boldsymbol{q}_2)$
- If we are only interested in $\boldsymbol{q}_1$, then $\boldsymbol{q}_2$ are called *nuisance parameters*
- If $\boldsymbol{q}_1=\{0,1\}$, and $\boldsymbol{q}_2$ are nuisance parameters, we call it a *composite* hypothesis testing problem

Ex: Positron Emission Tomography

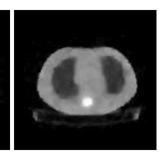
- Simple, traditional linear DSP-based approach
  - Filtered Back Projection (FBP)
- Advanced, estimation-theoretic approach
  - Model *Poisson* "likelihood" of collected data
  - *Markov Random Field* (MRF) "prior" on image
  - Find estimate using *expectation-maximization* algorithm (or similar technique)



Raw Data



FBP



Statistical Estimate

Images from webpage of Prof. Jeffrey Fessler, U. Michigan, Ann Arbor

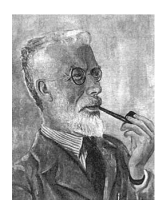
Tasks of Statistical Signal Processing

1. Create statistical model for measured data
2. Find fundamental limitations on our ability to perform inference on the data
   - Cramèr-Rao bounds, Chernov bounds, etc.
3. Develop an optimal (or suboptimal) estimator
4. Asymptotic analysis (i.e., assume we have lots and lots of data) of estimator performance to see if it approaches bounds derived in (2)
5. Hop on the computer: Do simulations and experiments comparing algorithm performance to lower bounds and competing algorithms

(Hero, p. 9)

To be B or not to be B: That is the Bayesian question

Rev. Thomas Bayes



Sir R.A. Fisher

Pics from MacTutor History of Mathematics Archive (http://www-groups.dcs.st-andrews.ac.uk/~history)

To be B

- A Bayesian analysis treats q as a random variable with a "prior" density $p(q)$
- Data generating machinery is specified by a conditional density $p(y|q)$
  - Gives the "likelihood" that the data y resulted from the parameters q
- Inference usually revolves around the posterior density, derived from Bayes' theorem:

$$p(q|y) = \frac{p(y|q)p(q)}{p(y)} = \frac{p(y|q)p(q)}{\int p(y|q)p(q)\, dq}$$

A Quick Side Note

- We use the word "density" in a generalized sense, encapsulating probability mass functions as well. A notation like $p(y|q)$ could indicate discrete or continuous data y, and discrete or continuous parameters q
- What is intended will be clear by context
- Just replace integrals with sums when appropriate; no harm done!
- See p. 4 of Poor for a rigorous discussion

Not to be B

- Trouble with Bayesian analysis: it is not always obvious what a good prior density might be.
  - Often chosen for computational convenience!
- In a non-Bayesian analysis, the likelihood density $p(y;q)$ is a density in y parameterized by a *nonrandom* parameter q.
  - It is not "conditioned" on q in a strict probabilistic sense.
- Inference usually revolves around likelihood $p(y;q)$
- Debates between Bayesians and non-Bayesians have been violent at times!

The Big Picture

$$q \xrightarrow{} \begin{matrix} p(y|q) \\ \text{or} \\ p(y;q) \end{matrix} \xrightarrow{} y \xrightarrow{} \hat{q}(\cdot) \xrightarrow{} \hat{q}(y)$$

Parameter	Random Mapping	Data	Estimator (a function)	Estimate (a specific value)

- In detection and classification, the "estimator" is instead called a "decision rule"
- Note the parameter $\hat{\theta}$ we want to estimate could itself be a function, in which case $\hat{\theta}$ is a functional
- Much of the class will be about ways of designing θ

"Under the Hood" of ECE7251, Spring 2002 offering

- Aaron's quest for a textbook
  - Literature review
- Different ways of teaching/learning the material
  - What to cover?
  - What order to cover it in?

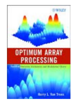
"Detection, Estimation, and Modulation Theory, Part I", Harry L. Van Trees, 1968

- First "engineering" textbook on the topic
- Will be republished periodically for all eternity
- Some homework problems are masters thesis topics!!!
- Every EE should have a copy
- Unfortunately, showing its age
  - All emphasis on continuous time problems, as might be implemented with analog computers
  - Ex: Discrete-time Kalman filtering relegated to a homework problem!!!
- Really wish Van Trees would write a 2nd edition

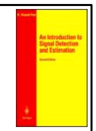
Rest of Van Trees' Masterpiece

- "DEM Theory, Part II: Nonlinear Modulation Theory," 1971
  - Not as popular as the others in the series
- "DEM Theory, Part III: Radar-Sonar Processing and Gaussian Signals in Noise," 1971
  - Much material not readily available elsewhere
  - Finally republished! Hooray!

- "DEM Theory, Part IV: Optimum Array Processing," 2002
  - Van Trees has been working on it for 30 years!!!
  - I'll believe it when I see it…

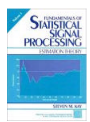


"An Introduction to Signal Detection and Estimation," 2nd Edition, H. Vincent Poor (1988, 1994)



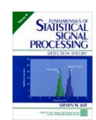
- "If you have some familiarity with the topic you will undoubtedly enjoy this book, but if you are a student tackling with this for the first time it will be demanding reading. You will need considerable fluency in random variable calculus to get the most out of the book, as the author presents many results and derivations as 'straightforward.'" – Amazon Review
- Excellent emphasis on general structure of det. & est. problems
- Most mathematically rigorous of all the engineering texts
- Alas, writing style is often dense and difficult to penetrate

"Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory", Steven Kay, 1993



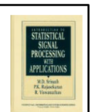
- "...excellent tutorial and research reference book on estimation theory. The theory is illustrated with very concrete examples; the examples give an 'under-the-hood' insight into the solution of some common estimation problems in signal processing. *If you're a statistician, you might not like this book. If you're an engineer, you will like it."* – Amazon review
- "The theory is explaned well and motivated, but what makes the book great are the examples. There are many worked examples and they are chosen to make things very clear." – Amazon review

"Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory," Steven Kay, 1998



- "This is an excellent book and is very easy to follow, unlike Poor's which is too mathematical and hard too read. However, this book does make many references to Vol. 1, the Estimation Theory, so you almost have to get both books to get a full understanding." - Amazon review
- Friendly and easy to read, like J.R.R. Tolkien
- Sprawling and epic, like J.R.R. Tolkien (two book sequence is a bit unwieldy for a single course)
- Seems to view the world as a series individual problems that come up one at a time and are addressed one at a time, like J.R.R. Tolkien (less emphasis on abstract, general structure found in Poor's book)

"Introduction to Statistical Signal Processing," M.D. Srinath, P.K. Rajasek, and R. Viswanthan, 1996



- "The book is well-organized in terms of content and layout, but the mathematics is very poorly presented. Some things just fall out of mid-air. It could use some rigorous presentation." – Amazon Review
- "For a mathematical text this book lacks a lot of details. Things often fall out of nowhere and it is very hard for students to follow how equations are derived." – Amazon Review
- Special topics: Digital Communications, Radar, Target Tracking Pattern Classification, System Identification

Some Other Books

- "Elements of Signal Detection and Estimation," Carl Helstrom, 1995
  - Out of print, unfortunately
  - A little bit wordy
  - Covers some interesting (although somewhat obscure) techniques not covered elsewhere: optical communications, saddle point integration, etc.
- "Statistical Signal Processing: Detection, Estimation, and Time Series Analysis," Louis Scharf, 1991
  - Heavy emphasis on matrix analysis; unique viewpoint
  - No continuous time results
  - Not recommended for ECE7251 in particular, but a good reference on the topics it covers; a good book to have close by when doing research

“Statistical Methods for Signal Processing,” Al Hero, 1998-2002

- Book-notes; work in progress
- Prof. Hero noted most recent books downplay statistical theory, and instead emphasize specific engineering applications
- Applications are useful for motivation, but we must be careful to avoid obscuring common underlying structures
- Manuscript strives to more thoroughly cover the theory of mathematical statistics than other books, permitting unification of many diverse areas
- Received permission from Hero to use these notes for ECE7251, Spring 2002; will use in conjunction with Poor’s book and try to combine the best of both
- Feedback welcome!!!

Notation in Different Books

	Param.	Data	Non-B. LL	B. LL	B. Prior	B. Post.
Van Trees	a	r	$p(r\|a)$	$p(r\|a)$	$p(a)$	$p(a\|r)$
Kay	q	x	$p(x;q)$	$p(x\|q)$	$p(q)$	$p(q\|x)$
Srinath	q	z	$p(z\|q)$	$p(z\|q)$	$p(q)$	$p(q\|z)$
Poor	q	y	$p_\theta(y)$	$p_\theta(y)$	$w(q)$	$w(q\|y)$
Hero	q	x	$f(x;q)$	$f(x\|q)$	$f(q)$	$f(q\|x)$
Lanterman	q	y	$p(y;q)$	$p(y\|q)$	$p(q)$	$p(q\|y)$

- Van Trees’ and Srinath’s use of $p(\cdot|\cdot)$ for a non-Bayesian likelihood is potentially misleading
- Poor’s uses of $p_\theta(y)$ for a Bayesian loglikelihood is nonstandard
- When discussion Kalman and Wiener filtering, Poor uses X for parameters and Y for data (common notation in filtering literature)

Approaches of Different Authors

Srinath et. al

Det. with finite/DT data
Det. with CT Data
Est. with finite/DT data
Est. with CT data/params.

Vince Poor/Van Trees

Det. with finite/DT data
Est. with finite/DT data
Det. with CT data
Est. with CT data/params.

Stephen Kay

Est. with finite data
Det. with finite data
(little emphasis on CT)

Al Hero

Est. with finite/DT data
Det. with finite/DT data
Det. with CT data
Est. with CT data/params.
(less emphasis on CT)

Continuous-Time analysis needs more mathematical machinery: *Kahunen-Loéve expansions*

Pedagocial Question: What to do first?

- View 1: Detection first, then Estimation (Van Trees/Poor/Srinath et. al)
  - Detection Theory is easier; introduces concepts used in Estimation Theory in a simple context
  - Detection Theory is more fun
- View 2: Estimation first, then Detection (Kay/Hero)
  - Detection Theory just a special case of estimation theory
  - Detection problems with “unknown parameters” are easier to think about if you’ve already seen estimation theory
  - Estimation Theory is more fun

View 1 seems more common, but we’ll take View 2

Pedagocial Question: When to do Karhunen-Loéve?

- View 1: Don’t bother with at all! (Scharf, Kay)
  - “Most modern implementations done with discrete-time processing anyway”
  - Response from View 2: “continuous-time analysis provides insight into the behavior of discrete-time implementations”
- View 2a: Do after finite-data detection problems, but before estimation (Srinath et. al)
  - Do all detection material at once
- View 2b: Put at end of course (Poor/Van Trees/Hero)
  - Material quite challenging; integral equations are scary!

We’ll take view 2b.