updated 23 February 2009

Johns Hopkins University
Whiting School of Engineering
Course No. 550.730, Spring 2009

Topics in Statistics:
Roundtable in System Identification and Likelihood Methods

Instructor: Jim Spall, 240-228-4960 (Wash.) or 443-778-4960 (Balt.); james.spall@jhuapl.edu; fax 240-228-8110/443-778-8110.
Office hours: Tuesdays (as available from 10:45AM – 12:30PM; 3:45PM – 7:30PM); Whitehead Hall, Room 303A.
Class time/location: Tuesdays: 2:00PM – 3:20PM for one speaker; 1:30PM – 3:20PM for two speakers, Room 303 (AMS Library), Whitehead Hall.

Summary

The focus of this roundtable-format course will be stochastic modeling as relates to system identification and maximum likelihood. The principles and algorithms being covered in this course have tremendous importance in the world at large. For example, maximum likelihood is arguably the most popular method for parameter estimation in most real-world applications. System identification is the term used in many fields to refer to the process of mathematical model building from experimental data. The system identification process refers to several important aspects of model building, including experimental design, selection of the model form (linear or nonlinear, static or dynamic, etc.), parameter estimation, and model validation. Often, but not necessarily, the models being considered are dynamical models in state-space form. This course will cover topics such as the maximum likelihood formulation and theory, the EM (expectation-maximization) algorithm and its variants, factorial and optimal experimental design, Fisher information, common model structures, online versus offline estimation, the role of feedback in identification (i.e., open-loop versus closed-loop estimation), standard and extended Kalman filtering, and uncertainty characterization (e.g., confidence regions).

Class structure

Typically, each class session will cover an important paper from the literature (or book excerpt) related to the course theme. In each session, the instructor or one or more students will have the responsibility to lead the discussion. Each student will be responsible for leading (or co-leading) several discussions over the duration of the semester. Discussion leaders will be responsible for preparing a well-organized presentation as a vehicle for sparking a dialogue about the subject at hand. (Aside from studying the technical material in the course, it is intended that the roundtable provide a collegial setting for developing presentation skills.) Because course participants will be collectively exploring a relatively advanced topic at most class sessions, the discussion leader need not have a complete grasp of every aspect of the topic being studied; coming to class with pertinent questions and issues for discussion is encouraged as a vehicle for learning.

Prerequisites

Undergraduate-level matrix theory and ordinary differential equations; graduate-level course in probability and statistics (e.g., 550.430; in particular, students should have prior exposure to maximum likelihood and Bayes' rule). Prior experience in data analysis and algorithms will be helpful.

Textbook

None. Journal and other readings will be used throughout the course; students will generally be responsible for obtaining copies of the readings online or at the Eisenhower Library. A copy of Spall (2003) is available in the Eisenhower Library (hard copy and online) for your use as needed; this book includes some background material on matrix theory, convergence, Fisher information, and optimal input design that may be useful for some of the weekly subjects (this background information is generally available in other sources, as well).

Expectations and Grading

The grade in this course will be assigned according to the following proportions:

1. Presentations (oral delivery and hand-in components): 40%
2. Class participation: 10%
3. Homework: 20%
4. Timely submission of acceptable discussion points/questions: 10%
5. Final report: 20%

As this is an upper-level graduate course based on collaborative study, grading is not intended to punitive. I would be happy to give everyone a good grade if possible, but this is not automatic! Students making an honest effort and contributing to the best of their ability should do well in the course. The five components contributing to a student's grade are discussed below:

Presentations. For the presentations component of the grade, each student will be evaluated based on the quality and insight of the presentations they are responsible for giving. It is not necessary to prepare a "fancy" presentation (i.e., fancy graphics or animations), but I will expect students to adhere to basic principles of effective technical presentations as discussed in week 1 of the course. In preparing a presentation, students are free to discuss among themselves (or others) questions or issues relevant to the paper(s) being reviewed. The presentations must include some PowerPoint slides, and may also include some whiteboard presentation. Students must provide a paper copy of the slides to the instructor (one or two slides per page only) and other students (or email the slides to the class at least 90 minutes in advance of the class). Students may (if desired) provide a supplementary one- or two-page summary of the reading(s) being covered; this is not required (the emphasis should be on the presentation).

A tentative schedule is given in the table below. Because of the informal nature of the course, this schedule may be adjusted if it is found that more time will be needed to adequately address a specific topic. There is also the possibility of introducing a topic not shown on the current schedule.

While this class is directed at students who are (or will be) carrying out research, the students are not expected to carry out their own research on the ideas being reviewed. For each presentation, it is sufficient to do a thorough review of relevant reference material. In particular, the following questions/issues may be used as a framework for constructing each presentation. While it is recognized that not every item on this list can be covered in every presentation, items 1 – 3 should be considered mandatory in all presentations:

1. (mandatory) Basic description of the problem(s) being addressed, keeping in mind the prerequisites of the course. It is very important that this be done for non-specialists (which may require some thought on the student's part, as the reference(s) may be written for specialists).
2. (mandatory) Summary of the main contribution of the author(s) of the paper (i.e., the "breakthrough" if relevant). While the presentation may include general background information on whatever topic is being considered, it must be focused on the reading(s) for that week.
3. (mandatory) Positives and negatives of the basic algorithms, methods, etc.
4. If relevant, connection of the topic to work the student may be aware of in the AMS or other Department, including by the student him/herself.
5. Summary of numerical results of the author(s) and perhaps others (including any personal experience the student may have, if relevant).
6. The presentation may include published insights and conclusions of others, as published after the paper(s) being reviewed were published (science citation indices may be useful for this; e.g., http://apps.isiknowledge.com). Note: All standards of proper attribution and scholarship apply if the student uses such insights (i.e., provide specific citation information as relevant, including to results appearing only on the Internet). Also, full citation information must be provided for figures/diagrams that appear elsewhere and that are used in the presentation. Failure to adhere to these principles of scholarship is an ethical violation and potential plagiarism.
7. As the course proceeds, make connections (as relevant) to topics that were discussed in earlier weeks.
8. Any other relevant points relating the material to the student's personal, academic, or work experience.

Class participation. As a roundtable course, it is expected that students other than the discussion leader will review the assigned paper prior to class and to be prepared to pose questions or raise issues (see also "Discussion points/questions" below). The roundtable is intended to be an informal forum for discussion, and will work best if people other than the discussion leader contribute. The discussion leader is also encouraged to raise pertinent questions and issues for discussion, including questions or points of confusion that he/she may have. Do not assume the instructor knows much about many of the topics being studied—he doesn't and he has much to learn about certain topics!

Homework. Homework will be assigned in some classes as a means of demonstrating key points. The homework will tie directly to the theme of the class session in which it is assigned. Unless notified otherwise, students are expected to work independently on the homework. Some homework problems may be assigned that allow for collaborative study; these problems will be clearly indicated when assigned.

Discussion points/questions. To encourage preparation and background reading for all participants, all students other than the presenter are required to submit a comment or question about that week's subject no later than 6:30PM on the Monday before class. These comments/questions should represent some depth of understanding only possible after a serious reading of the assigned material (i.e., obvious and/or trivial comments/questions will be given zero credit). However, it is not necessary that the comments/questions be more than one or two substantive sentences (overly long submissions are discouraged). These comments/questions should be submitted via email to james.spall@jhuapl.edu. All such comments/questions will be forwarded to all in the class Monday evening; the presenter for the week may wish to address some or all of these comments/questions in the presentation. The full 10% towards the final grade will be given for timely submission of "good" (insightful) comments/questions in all weeks (i.e. approximately one percentage point per week will be deducted for late/ no submissions and/or for comments that may be considered trivial); a submission is late if it arrives after 6:30PM and does not get included in the all-class mailing on Monday evening.

Final report. All students will have to write a final report, summarizing aspects of the weekly discussions and possibly answering several questions that will be posed by the instructor (these questions will be based on the weekly discussions) or carrying out an independent research project directly related to at least one of the course themes. There will also be an oral presentation component of the final report (see week 14 in the class schedule below). More information will be provided near the end of the course.

Tentative Class Schedule (May be adapted as circumstances warrant)*

Week	Topics	Discussion leader
1	(1) Giving technical presentations. (2) Introduction to system identification (Ljung, 1999); the likelihood function and some ramifications (Mardia et al., 1979, pp. 96–108; Reid, 2000).	Jim Spall
2	Fisher information matrix and illustration of Monte Carlo sampling (Spall, 2003, Sect. 13.3 [Fisher information matrix] and Spall, 2005 [Monte Carlo sampling]).	Xumeng Cao
3	Maximum likelihood: properties and standard search algorithms. Primary: Efron, 1982; secondary: brief comments on standard numerical methods [steepest descent, Newton-Raphson, scoring], e.g., http://en.wikibooks.org/wiki/Statistics:Numerical_Methods/Optimization#Why_Gradient_Methods.3F or other reference.	Graham Beck
4	Expectation–Maximization (EM). Dempster et al. (1977).	Nash Borges
5	EM, continued. Convergence properties and practical experience. Primary: Wu (1983); secondary: Vaida (2005).	Valentina Staneva

6	Dynamical (state-space) models and the Kalman filter. Primary: Rauch et al. (1965) (filter and smoother); secondary: Spall (1988) (historical context) and Duncan and Horn (1972) (first paper in statistics literature, including regression interpretation).	G.B.
7	Parameter estimation for dynamical models. Primary: Gibson and Ninness (2005); secondary: Gupta and Mehra (1974) (continuous time state, discrete time measurement) and Goodrich and Caines (1979) (note especially Table 1, p. 409, for standard Kalman filter-based gradient calculations when using scalar measurements).	X.C.
8	Recursive stochastic algorithms (Ljung, 1977).	V.S.
9	Selecting input values in experimentation: (1) Factorial designs and open-loop design of experiments (Spall, 2009).	N.B.
10	Selecting input values in experimentation: (2) Optimal design in open- and closed-loop. Primary: Pronzato, 2008 (review from identification perspective); secondary: Spall, 2003, Chap. 17 (general review).	V.S.
11	Model structure determination (Ref. TBD).	X.C.
12	Mixture distributions and EM (Redner and Walker, 1984).	N.B.
13	General issues in open-loop versus closed-loop indentification (tentative references: selected parts of Hjalmarsson, 2005, and Agüero and Goodwin, 2007).	G.B.
14	Final class summary presentations.	All

*Possible additional/alternative topics: Extended Kalman filter, case study in system identification, model validation, stochastic EM algorithms, other.

References

Agüero, J. C. and Goodwin, G. C. (2007), "Choosing Between Open and Closed Loop Experiments in Linear System Identification," IEEE Transactions on Automatic Control, vol. 52(8), pp. 1475–1480, 2007.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum Likelihood from Incomplete Data via the EM Algorithm," Journal of the Royal Statistical Society, Series B, vol. 39, pp. 1–38.

Duncan, D. B. and Horn, S. D. (1972), "Linear Dynamic Recursive Estimation from the Viewpoint of Regression Analysis," Journal of the American Statistical Association, vol. 67, pp. 815–821.

Efron, B. (1982), "Maximum Likelihood and Decision Theory," Annals of Statistics, vol. 10, pp. 340–356.

Gibson, S. and Ninness, B. (2005), "Robust Maximum-Likelihood Estimation of Multivariable Dynamic Systems," Automatica, vol. 41(1), pp. 1667–1682.

Goodrich, R. L. and Caines, P. E. (1979), "Linear System Identification from Nonstationary Cross-Sectional Data," IEEE Transactions on Automatic Control, vol. 24, pp. 403–411.

Gupta, N. and Mehra, R. (1974), "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations," IEEE Transactions on Automatic Control, vol. AC-19(6), pp. 774–783.

Hjalmarsson, H. (2005), "From Experiment Design to Closed-Loop Control," Automatica, vol. 41(3), pp. 393–438.

Ljung, L. (1977), "Analysis of Recursive Stochastic Algorithms," IEEE Transactions on Automatic Control, vol. AC-22, pp. 551–575.

Ljung, L. (1999), System Identification—Theory for the User (2nd ed.), Prentice Hall PTR, Upper Saddle River, NJ.

Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979), Multivariate Analysis, Academic Press, New York.

Pronzato, L. (2008), "Optimal Experimental Design and Some Related Control Problems," Automatica, vol. 44, pp. 303–325.

Rauch, H. E., Tung, F., and Striebel, C. T. (1965), "Maximum Likelihood Estimates of Linear Dynamic Systems," AIAA Journal, vol. 3(8), pp. 1445–1450.

Redner, R. A. and Walker, H. F. (1984), "Mixture Densities, Maximum Likelihood, and the EM Algorithm," SIAM Review, vol. 26(2), pp. 195–239.

Reid, N. (2000), "Likelihood," Journal of the American Statistical Association, vol. 95, pp. 1335–1340.

Spall, J. C. (1988), "An Overview of Key Developments in Dynamic Modeling and Estimation," in Bayesian Analysis of Time Series and Dynamic Models (J. C. Spall, ed.), Marcel Dekker, New York, pp. xv–xxvii.

Spall, J. C. (2003), Introduction to Stochastic Search and Optimization: Estimation, Simulation, and Control, Wiley, Hoboken, NJ.

Spall, J. C. (2005), "Monte Carlo Computation of the Fisher Information Matrix in Nonstandard Settings," Journal of Computational and Graphical Statistics (American Statistical Assoc.), vol. 14, pp. 889–909.

Spall, J. C. (2009), "Factorial Design for Choosing Input Values in Experimentation: Generating Informative Data for System Identification," submitted to IEEE Control Systems Magazine.

Vaida, F. (2005), "Parameter Convergence for EM and MM Algorithms," Statistica Sinica, vol. 15, pp. 831–840.

Wu, C. F. J. (1983), "On the Convergence Properties of the EM Algorithm," Annals of Statistics, vol. 11, pp. 95–103.

Statement on Academic Integrity

The strength of the university depends on academic and personal integrity. Work submitted in the class is required to be the student's own unless appropriate acknowledgment and/or citation is given. Students are reminded that it is not acceptable to download a presentation from the Web and call it your own (even with changes); this is plagiarism and is a serious breach of academic standards. Violations of recognized scholarly principles may result in an "F" in the class and possible reporting to the appropriate JHU academic authorities. Unless notified otherwise, all homework is to be done on an individual basis (i.e., no collaboration).