ECE 645: Estimation Theory
Spring 2015
Instructor: Prof. Stanley H. Chan

PURDUE UNIVERSITY

Project Guideline

Spring 2015
(Last Update: January 12, 2015)

The purpose of the project is to provide you hands-on experience in writing estimation / detection algorithms, as well as analyzing the performance of the methods. I am in general open to all topics in signal processing, machine learning, and applied statistics. You are also encouraged to propose topics related to your research. However, everything you propose should be within the scope of this course. For example, I do not accept things like MPEG 4, power system dynamics, or data structure, etc. If you have doubts about your proposed project, feel free to send me an email.

Important Dates

1. Jan 30: Deadline to send me an email to describe your initial project idea. (Optional)
2. Feb 27: Proposal deadline. 1-page + $\frac{1}{2}$-page references.
3. May 1: Report deadline. 4-page + 1-page references.
4. Apr 27 - May 1: Oral Presentation. 10 minutes each.

Initial Email (Optional)

If you like to get some quick feedback from me before you start to work on your project, you can send me an email by Jan 30. This email exchange is completely voluntary. The purpose is to help you shape your project focus. Moreover, if you want to chat with me in person, please send me the initial email first. When you send me the email, please include the following items.

1. Title: This is a tentative title of your project. You can change afterwards.
2. Summary: Please use at most 3 sentences to describe what you want to do. Do not write more than 3 sentences, as that usually means either you are not clear of what you want to do, or you are attempting too much. Also, by "sentence" I meant short and simple sentences. Do not write me a lengthy clause.
3. Reference: Tell me which paper / book chapter that your project is based on.

Proposal (20%)

The proposal is extremely important in the sense that it forces you to think before you work. A good proposal can help you work smoothly and achieve a higher goal. When drafting the proposal, please use the LaTeXtemplate available on the course website. I do *not* accept MS Word. Please use font size 10pt (default), and write no more than one page (plus no more than half page of references). Your proposal should include the following elements:

1. Project title: Be specific of what you want to do. I do not accept things like: "Introduction to Estimation Theory".

2. Motivation: State the problem, and tell me why is this problem important (1) to you; (2) to the rest of the world.

3. Uniqueness: What is the unique feature of your project? Why does existing works not sufficient?

4. Tasks: List out the tasks that you need to accomplish. Be realistic.

5. Expected Outcome: What do you expect to obtain?

The percentage of each component is weighted as follows: Title (10%), Motivation (30%), Uniqueness (30%), Tasks (15%), Expected Outcome (15%). As you can tell from the above grade distribution, I put extra emphasis on the motivation and uniqueness of your project. The reason is that I want you to spend the right effort on the right problem. So please think about these carefully when you write your proposal.

Final Report (40%)

Your project report should be self-contained, and should be of an IEEE flagship conference quality (e.g., ICASSP or ICIP). Please use the LaTeXtemplate available on the course website. Maximum number of pages is 4, and you can use the 5th page for reference if needed. Your report should contain the following elements.

1. Project title: This is the finalized project title.

2. Introduction: After you have conducted the study, you should have a better perspective of your problem statement and motivation.

3. Literature Review: Please demonstrate your understanding of your work, and how you would position your work in the literature.

4. Proposed Method / Study: What is your new contribution? New method?

5. Experimental Result: Any experimental data to support your method? Comparison with other methods?

6. Conclusion and Future work.

Your report will be graded based on (1) Clarity of your problem statement and motivation (20%); (2) Creativity of your solution (30%); (3) Completeness of experiment (20%); (4) Writing (30%). Writing is an important part of the report. Please spend enough effort to polish your report.

Oral Presentation (40%)

You are required to give an oral presentation during the last week of the semester. If we cannot accommodate all students during that week, we will use the exam week as well. Date and time will be announced later. The order of the presentation will be decided according to a random draw. Every student will have 10 minutes for the presentation + 2 minutes for Q & A. I aim to have 4 presentations in one class meeting. Your oral presentation will be graded as follows.

1. Attendance (15%). Even if you are not presenting, you should show up to support your fellow classmates.

2. Ask questions (15%). Every student should ask at least one question during the presentation week.

3. Technical content (40%). That includes: (1) Clarity of your problem statement and motivation; (2) Creativity of your solution; (3) Completeness of experiment. (Same for final report).

4. Presentation (30%). Be reminded that you are presenting your work to people who have little or no background knowledge about your topic. You have to make things clear, and understandable.

Suggested Papers

The following is a list of suggested papers for your reference. These papers are biased selections from my Statistical Signal and Information Processing (SSIP) group. You are always welcome to come up with topics outside this list, e.g., estimation and detection in communication systems. However, for students who are interested in joining SSIP, you are advised to choose from the list and talk to me.

Stein Unbiased Risk Estimator

[1] E. Candes, C. Sing-Long, and J. Trzasko, "Unbiased risk estimates for singular value thresholding and spectral estimators," *IEEE Trans. Signal Process.*, vol. 61, pp. 4643–4657, 2013.

[2] Y. Eldar, "Generalized SURE for exponential families: Applications to regularization," available at http://arxiv.org/pdf/0804.3010.pdf, 2008.

[3] S. Ramani, T. Blu, and M. Unser, "Monte-Carlo SURE: A black-box optimization of regularization parameters for general denoising algorithms," *IEEE Trans. Image Process.*, vol. 17, no. 9, pp. 1540–1554, 2008.

[4] D. Van De Ville and M. Kocher, "SURE-based non-local means," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 973–976, Nov. 2009.

Stein's Unbiased Risk Estimator (SURE) is a powerful tool that can be used to estimate regularization parameters in MAP estimations. In a nutshell, SURE is an asymptotically consistent estimate of the ground truth MSE. You can explore various properties of SURE, or apply SURE to solve a problem of your interest.

Signal Processsing on/of Graphs

[1] A. Agaskar and Y. M. Lu, "A spectral graph uncertainty principle," *IEEE Trans. Information Theory*, vol. 59, no. 7, pp. 4338–4356, Jul. 2013.

[2] E. M. Airoldi, T. B. Costa, and S. H. Chan, "Stochastic blockmodel approx- imation of a graphon: Theory and consistent estimation," *Advances in Neural Information Processing Systems (NIPS)*, pp. 692–700, 2013.

[3] S. H. Chan and E. M. Airoldi, "A consistent histogram estimator for exchangeable graph models," *Journal of Machine Learning Research Workshop and Conference Proceedings*, vol. 32, no. 1, pp. 208–216, 2014.

[4] F. Meyer and X. Shen, "Perturbation of the eigenvectors of the graph Laplacian: Application to image denoising," *Applied and Computational Harmonic Analysis*, 2013, In press. Available online at http://arxiv.org/abs/1202.6666.

[5] A. Sandryhaila and J. M. F. Moura, "Big data analysis with signal processing on graphs: Representation and processing of massive data sets with irregular structure," *IEEE Signal Process. Magazine*, vol. 31, no. 5, pp. 80–90, Sep 2014.

[6] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," available at http://arxiv.org/abs/1211.0053.

One important characteristic of modern data, besides large-scale, is the inter-connectivity of various attributes. The field is rapidly evolving, and there are many exciting results developed over the past few years. If you are interested in "big data" type of research topics, I would recommend you to take a look at these papers.

EM Algorithm and Gaussian Mixture Models

[1] O. Cappe and E. Moulines, "Online EM algorithm for latent data models," *Journal of the Royal Statistical Society*, vol. 71, no. 3, pp. 593–613, 2009.

[2] M. Gupta and Y. Chen, "Theory and use of the EM algorithm," *Foundations and Trends in Signal Processing*, vol. 4, no. 3, pp. 223–296, 2010.

[3] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2481–2499, May 2012.

[4] D. Zoran and Y. Weiss, "From learning models of natural image patches to whole image restoration," in *Proc. IEEE International Conference on Computer Vision (ICCV)*, Nov. 2011, pp. 479–486.

Gaussian mixture model and EM algorithm are classical statistical tools for signal processing and machine learning. Despite its long history, people still find them extremely useful in modern applications. If you are interested in working on GMM and EM, you can explore the application of EM in your research. If you are interested in large datasets, you can consider the online EM algorithms.

Photon Limited Imaging

[1] S. H. Chan and Y. M. Lu, "Efficient image reconstruction for giga-pixel quantum image sensors," in *IEEE Global Conf. Signal. Inform. Process.*, 2014.

[2] E. R. Fossum, "The Quantum Image Sensor (QIS): Concepts and challenges," in *Proc. OSA Topical Mtg Computational Optical Sensing and Imaging*, Jul 2011, Paper JTuE1.

[3] A. Kirmani, A. Colaco, F.N.C. Wong, and V.K. Goyal, "CODAC: A compressive depth acquisition camera framework," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP'12)*, Mar. 2012, pp. 5425–5428.

[4] F. Yang, Y. M. Lu, L. Sbaiz, and M. Vetterli, "Bits from photons: Oversampled image acquisition using binary poisson statistics," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1421–1436, Apr. 2012.

As part of Purdue's integrated imaging thrust, SSIP group is working actively on photon limited imaging algorithms. The basic components of photon limited imaging are the maximum likelihood estimation (MLE) and maximum a posterior (MAP) estimation. If you are interested in imaging and estimation theory, you are encouraged to take a look at these papers.

Random Sampling for Matrix Computation

[1] P. Drineas, R. Kannan, and M. Mahoney, "Fast Monte Carlo algorithms for matrices I: Approximating matrix multiplication," *SIAM J. Computing*, vol. 36, pp. 132–157, 2006.

[2] P. Drineas, R. Kannan, and M. Mahoney, "Fast Monte Carlo algorithms for matrices II: Computing low-rank approximations to a matrix,," *SIAM J. Computing*, vol. 36, pp. 158 – 183, 2006.

[3] P. Drineas, R. Kannan, and M. Mahoney, "Fast Monte Carlo algorithms for matrices III: Computing an efficient approximate decomposition of a matrix," *SIAM J. Computing*, vol. 36, pp. 184 – 206, 2006.

[4] P. Drineas and M. W. Mahoney, "A randomized algorithm for a tensor-based generalization of the singular value decomposition," *Linear Algebra and its Applications*, vol. 420, pp. 553–571, 2007.

[5] P. Drineas and M. W. Mahoney, "On the Nystrom method for approximating a Gram matrix for improved kernel-based learning," *J. Machine Learning Research*, vol. 6, pp. 2153–2175, 2005.

Imagine that you are given a huge matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, where m and n are very large, how would you perform matrix computation, e.g., matrix-vector multiplication, or SVD? This topic involves significant amount of probability, estimation, and large deviation analysis.

Image Denoising Theory

[1] S. H. Chan, T. Zickler, and Y. M. Lu, "Monte-Carlo non-local means: Random sampling for large-scale image filtering," *IEEE Trans. Image Process.*, vol. 23, no. 8, pp. 3711–3725, Aug. 2014.

[2] P. Chatterjee and P. Milanfar, "Is denoising dead?," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 895–911, Apr. 2010.

[3] A. Levin and B. Nadler, "Natural image denoising: Optimality and inherent bounds," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Jun. 2011, pp. 2833–2840.

[4] A. Levin, B. Nadler, F. Durand, and W. Freeman, "Patch complexity, finite pixel correlations and optimal denoising," in *Proc. 12th European Conf. Computer Vision (ECCV)*, Oct. 2012, vol. 7576, pp. 73–86.

[5] P. Milanfar, "A tour of modern image filtering," *IEEE Signal Processing Magazine*, vol. 30, pp. 106–128, Jan. 2013.

[6] P. Milanfar, "Symmetrizing smoothing filters," *SIAM J. Imaging Sciences*, vol. 6, no. 1, pp. 263–284, 2013, Submitted to SIAM Journal on Imaging Science. Available at `http://users.soe.ucsc.edu/~milanfar/publications/journal/SIIMSMay-3-12.pdf`.

[7] I. Ram, M. Elad, and I. Cohen, "Image processing using smooth ordering of its patches," *IEEE Trans. Image Process.*, vol. 22, no. 7, pp. 2764–2774, Jul. 2013.

Image denoising is the classical test bed for estimation algorithms. Recently, there are numerous efforts in seeking for the fundamental limits of image denoising. Some uses Cramer-Rao lower bound, some uses Bayesian MMSE estimation, and some uses Gaussian mixture models, etc. The powerfulness of these results is that they are the intrinsic properties of the problem, not of a particular denoising algorithm. If you are interested to learn how to analyze the performance of an estimation problem, denoising is a very good starting point.

Optimization Algorithms

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[2] S.S. Chen, D.L. Donoho, and M.A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.

[3] J. Eckstein and D.P. Bertsekas, "On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators," *Math. Program.*, vol. 55, no. 3, pp. 293–318, Jun. 1992.

Most, if not all, of the estimation methods are iterative. In these situations, the estimation boils down to an optimization problem. If you are interested in learning how to solve a particular estimation problem using numerical optimization tools, you can check these recent papers.

Tensor

[1] C.F. Caiafa and A. Cichocki, "Generalizing the column-row matrix decomposition to multi-way arrays," *Linear Algebra and its Applications*, vol. 433, pp. 557–573, 2010.

[2] J. Chen and Y. Saad, "On the tensor SVD and the optimal low rank orthogonal approximation of tensors," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 4, pp. 1709–1734, Jan. 2009.

[3] M. Ishteva, P.A. Absil, S. Huffel, and L. De Lathauwer, "On the best low multilinear rank approximation of higher-order tensors," in *Recent Advances in Optimization and its Applications in Engineering*, Moritz Diehl, Francois Glineur, Elias Jarlebring, and Wim Michiels, Eds., pp. 145–164. Springer Berlin Heidelberg, 2010.

[4] R. Kannan, "Spectral methods for matrices and tensors," in *Proceedings of the 42nd ACM symposium on Theory of computing*, New York, NY, USA, 2010, STOC '10, pp. 1–12, ACM.

[5] T.G. Kolda and B.W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, Aug. 2009.

[6] T. Kolda and B. Bader, "Tensor decompositions and applications," *SIAM Review*, vol. 51, no. 3, pp. 455–500, 2009.

[7] L. De Lathauwer and D. Nion, "Decompositions of a higher-order tensor in block terms, part III: Alternating least squares algorithms," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1067–1083, Sept. 2008.

Tensor is a natural extension of matrices and vectors. However, efficient computational methods for tensors are yet to be developed. Tensors can be used for network analysis, hyperspectral imaging, and data compression etc. If you are interested in tensors, you can study the applications of tensors; You can also study various tensor computation methods, e.g., SVD.

Sampling and Estimation of Sparse Signals

[1] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: Design of dictionaries for sparse representation," *Proceedings of SPARS*, vol. 5, pp. 9–12, 2005.

[2] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, "A simple proof of the restricted isometry property for random matrices," *Const. Approx.*, vol. 28, no. 3, pp. 253–263, Dec. 2008.

[3] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Rev.*, vol. 51, no. 1, pp. 34–81, Feb. 2009.

[4] E. J. Candès and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Trans. Information Theory*, vol. 57, no. 11, pp. 7235–7254, Nov. 2011.

[5] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Magazine*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

[6] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[7] D.L. Donoho and X. Huo, "Uncertainty principles and ideal atomic decomposition," *IEEE Trans. Info. Theory*, vol. 47, no. 7, pp. 2845–2862, Nov. 2001.

[8] M. Elad, "Sparse and redundant representation modeling – what next?," *IEEE Signal Process. Letters*, vol. 19, no. 12, pp. 922–928, Dec. 2012.

Sparse representation, or dictionary learning, is a young branch of signal processing. (Some people like to call it compressed sensing.) Sparse signal processing involves two important components: sampling and estimation. There are numerous applications of sparse signal processing, spanning from communication theory, image processing, matrix completion, etc. You can work on a new application of sparse signal processing, or you can conduct a comprehensive study of some existing methods.