18.338J/16.394J: The Mathematics of Infinite Random Matrices Project Ideas

Alan Edelman

Handout #8, Thursday, September 30, 2004

A requirement of this course is that the students experiment with a random matrix problem that is of interest to them. Often these explorations take the shape of a little bit of theory and a little bit of computational experimentation. Some of you might already have some ideas that are relevant to your current research. Regardless, we thought we'd put together some ideas for projects. Feel free to adapt them based on your interests; if you want more information about a particular idea, please feel to contact us.

Mid-term project presentations are tentatively scheduled for October 28th and November 2nd as indicated on the course calendar. We will provide more details about this in a subsequent email to the class. The basic purpose of this mid-term project is to get your feet wet thinking about a problem that is of interest to you using the tools you've learned about so far or to learn new tools for that purpose.

1   Covariance Matrices in Signal Processing

Sample covariance matrices come up in many signal processing applications such as adaptive filtering. Let G be the $n \times N$ "Gaussian" random matrix that we've encountered in class. In signal processing language, the Wishart matrix

$$W = \frac{1}{N}GG^* \tag{1}$$

is a rectangularly windowed estimator. In adaptive filtering applications, the covariance matrix that comes up can be expressed in terms of the Wishart matrix as $A = W^{-1}$. We've already seen how the Marčenko-Pastur theorem can be used to characterize the density of the eigenvalues of the Wishart matrix. The density of A can hence be computed using standard Jacobian tricks.

A more general covariance matrix estimator is referred to as being exponentially windowed. To see where the "windowing" comes in, let us first rewrite the Wishart matrix as

$$W = \frac{1}{N}\sum_{i=1}^{N} \underline{g}_i \underline{g}_i^H \tag{2}$$

where $\underline{g}_i$ is the i-th column of G. The rectangular windowing comes in from the fact that each of the N vectors $\underline{g}_i$ is weighted equally by $1/N$ in forming W. Hence, the exponentially weighted estimator is simply

$$W_\lambda = \frac{1}{N}\sum_{i=1}^{N} \lambda^{N-i} \underline{g}_i \underline{g}_i^H \tag{3}$$

where $\lambda < 1$ is the weighting factor.

In many signal processing publications, the consensus is that, loosely speaking,

$$E[W_\lambda^{-1}] = (1 - \lambda)E[W^{-1}]. \tag{4}$$

The proofs for this involve perturbation theory arguments. We feel that we can get some insight on this problem using random matrix techniques such as the ones you've learned in class.

Project Idea Use random matrix arguments to rederive or verify this proof. Indicate what values of λ this is valid for. Discuss any convergence issues i.e. what values of n and N is this approximately valid.

2 Stochastic Perturbation Theory

Allow us to sincerely flatter G. W. (Pete) Stewart by including his words on this subject verbatim from his survey paper on Stochastic Perturbation theory [1] which has been included in the course reader (Random Matrices - II).

Let A be a matrix and let F be a matrix valued function of A. Two principal problems of matrix perturbation theory are the following. Given a matrix E, presumed small,

1. Approximate $F(A + E)$,

2. Bound $\| F(A + E) - F(A) \|$ in terms of $\| E \|$.

Here $\| . \|$ is some norm of interest.

The first problem is usually, but not always, solved by assuming that F is differentiable at A with derivative $F'(A)$. Then

$$F(A + E) = F(A) + F'_A(E) + o(\| E \|), \tag{5}$$

so that for E sufficiently small $F'_A(E)$ is the required approximation. The problem then reduces to finding tractable expressions for $F'_A(E)$ which in itself is often a nontrivial task. The second problem may be treated a variety of ways; the basic idea being that the size of the perturbation is used as a bound even thought it is likely to be an overestimate.

Stochastic perturbation theory is different from this approach because it is, in some sense, intermediate of the other two. We let E be a stochastic matrix and compute expectations of quantities derived from the perturbation expansion.

Project Idea Pete Stewart's paper was published in 1990. The work of Marčenko and Pastur was rediscovered around that time; hence you will notice that there are no references to their work. An interesting project would be to condense what Pete Stewart has to say and to determine if what we learned in class can strengthen any of the ideas presented in his paper. Thinking of applications of this would be a bonus.

3 Other ideas

These some other ideas based on the survey papers on the website. If you want any clarification feel free to contact us.

- What is the replica method?

- What is the connection between statistical physics, spin glasses and random matrices?

- What can random matrix theory tell us about real world graphs?

- Suppose we were given the moments of the semi-circle (numerically). Could you compute its density using techniques described by what is known as the "Classical Moment Problem"?

- Construct a bijection between the MANOVA matrix and McKay's theorem (ask us for the exact relationship).

- Derive Jonsson's result in a more direct manner using the bipartite graph bijection (possibly).

- What is the connection between Jack Polynomials and Free probability.

- What aspects of random matrix theory can be useful for principal component analysis?

References

[1] G. W. Stewart. Stochastic perturbation theory. *SIAM Rev,* 32(4):579–610, 1990.