18.338 Eigenvalues of Random Matrices

Problem Set 1

Due Date: Wed Feb. 18, 2015

Homework

Do at least four out of the following problems (Computational/Mathematical problems are denoted as C/M. Exercises with numbers and pages are from the class notes.)

Read Chapter 1 and 5 from the notes (available from nb.mit.edu). Please comment about what you have read as if you were writing a substantial referee report. We want to know if there are any errors of any kind of course. Also and perhaps more important, we want to hear comments of style. Furthermore, usually there is one place where writing gets harder to follow – I'd like to know where that is for you.

Concentration of Measure for Gaussian Ensembles

It is remarkable how well the semicircle describes the histogram for Gaussian ensembles and other Wigner-type matrices. These mathematical and computational problems investigate the semicircle, how good it is, and how far off we can get.

Section 1.1 and section 5 of reference `http://www-math.mit.edu/~{}Eedelman/homepage/papers/flucts.pdf` are related to this question.

Take as given that the tridiagonal matrix T_n when normalized by $\sqrt{\beta n}$ (i.e. $H_n = T_n/\sqrt{\beta n}$) on Page 93 (Equation (6.3)) of the notes has the same eigenvalues as a Gaussian ensemble, where $\beta = 1$ is the GOE, $\beta = 2$ is the GUE, and any $\beta > 0$ is allowed.

The computational problems allow for investigation. Do as much or as little as interests you. *The main thing is to do something.* Ask us for help.

This following MATLAB code would estimate the kth moments:

```
n=20; beta=1; k=2;
t=4000;v=zeros(t,1);

for i=1:t
   d=sqrt(2)*randn(1,n);
   s=sqrt(chi2rnd(beta*[n-1:-1:1]));
   e=trideig(d,s)/sqrt(n); %install from http://persson.berkeley.edu/mltrid/index.html
   v(i)=mean(e.^k);
end
```

In Julia (*http://julialang.org*), one can use code such as (those students who want to do huge experiments and get into parallelism should contact me (edelman@mit.edu) or Andreas Noack (noack@csail.mit.edu)

```
n=20
beta=1
k=2
t=4000;
v=zeros(t);
for i=1:t
   d=randn(n)*sqrt(2)
    s=float64([sqrt(randchi2(beta*(n-i)) for i=1:(n-1)]))
    e=eigvals(SymTridiagonal(d,s))
    v[i]=mean(e.^k)
```

1. (M) or (C). The first moment (and all odd moments) of the eigenvalues of the Gaussian ensembles has expected value 0. (This is a way of saying that $\mathbb{E}[\mathbf{Tr}(T_n)] = 0$). Mathematically or with a Monte Carlo

simulation or both, conclude that $\mathbf{Tr}(T_n)$ is a scalar Gaussian. If you wish to access to Section 2.3.3 of Anderson, Guionnet, Zeitouni `http://www.math.umn.edu/zeitouni/technion/cupbook.pdf` (book page 42, pdf page 56) you might compare 2.3.10. How close are they?

2. (M) or (C) The second moment is a factor of $n^2/2$ times a χ^2 random variable with $n(n-1)\beta/2+n$ degrees of freedom. Prove this by using simple properties of chi-square. (The degrees of freedom add.)

   For the computationally minded you can compare the following.

```
[a,b]=hist(v,50);
hold off
plot(b, a/sum(a)/(b(2)-b(1)));
hold on
xx=(0:.01:1)*max(b);
j=n*(n-1)*beta/2+n;
x=xx*(n^2/2);
%for n>20 this formula must be approximated
plot(xx, (n^2/2)*(x).^(j/2-1).*exp(-x/2)/2^(j/2)/gamma(j/2),'r')
```

   One might use approximations such as if X has the distribution of χ^2_k then $\sqrt{2X}$ is roughly normal with mean $\sqrt{2k-1}$ (or just $\sqrt{2k}$ with unit variance). Potentially compare the concentration of measure again.

3. (M) What would happen in Problem 1 and 2 if the matrices are Wigner matrices (i.e., diagonal has variance 1 and the off-diagonal has variance 2) as $n \to \infty$? (Hint: use the Central Limit Theorem.)

4. (C) Investigate how other odd moments deviate from 0 or how even moment deviate from the Catalan numbers.

5. (C) Try to investigate how the histograms themselves deviate from the semicircle. One can draw lots of pictures to see the semicircle.

```
[a,b]=hist(e,50); % e are eigenvalues from the previous code
hold off
plot(b, a/sum(a)/(b(2)-b(1)));
hold on
x=[-2:.01:2];
plot(x,sqrt(4-x.^2)/(2*pi),'r')
```

   but what is interesting is to take averages and watch the fluctuations. See if you can estimate the fluctuations to the semicircle over various intervals using normals. One might start by taking the mean and seeing how far off finite n is from infinite n, or one can consider the variance.

6. (C) Perform Monte Carlo experiments on non-Gaussians carefully enough to predict the deviation.

7. (C) Plot the histogram of the square singular values of (for different z's on the complex plane)

$$\texttt{(randn(n,n) + sqrt(-1)*randn(n,n)) / sqrt(2*n)} - z \cdot \mathbb{I},$$

   and compare $|z| < 1$ with $|z| > 1$. One could start with the following Julia code:

```
n = 1000
M = ((randn(n, n) + im*randn(n,n))/sqrt(2*n))
# realPart = -2:0.1:2, imaginaryPart = -2:0.1:2
vals = svdvals(M - complex(realPart, imaginaryPart)*I)
plt.hist(vals, normed=True, bins=50)
```