Math 690 F2017: Topics in Data Analysis and Computation
Homework 2

Xiuyuan Cheng

1. (Wigner's Semi-circle Law) This problem is to observe the semi-circle law numerically,

   (1) The classical case of i.i.d. (symmetric) gaussian matrix:

   Generate a random matrix W of size n-by-n, where $W_{ij} \sim \mathcal{N}(0,1)$ i.i.d. for each $i \leq j$, and $W_{ji} = W_{ij}$, and compute its eigenvalues. Plot the histogram of eigenvalues. Let $n = 10, 100, 1000$. How is the distribution of (real) eigenvalues of W like as n increases? How is the eigenvectors like? (Remark: consider the normalized matrix $\frac{1}{\sqrt{n}}W$, the limiting eigenvalue density converges to a semi-circle, on [-2,2], as $n \to \infty$. The eigenvectors, concatenated into a matrix Ψ, where the j-th column is the j-th eigenvector, converge to the limit where Ψ is a random n-by-n orthogonal matrix.)

   (2) Universality of general Wigner matrix:

   Repeat (1), replacing the distribution of W_{ij} to be Bernoulli instead of $\mathcal{N}(0,1)$, that is $W_{ij} = 1$ with probability $\frac{1}{2}$ and -1 with probability $\frac{1}{2}$. W_{ij} is i.i.d for the upper triangular part same as in (1), and W is symmetric.

   (3) Rank-1 perturbation:

   Let u be a unit-length vector in $\mathbb{R}^n$. For example $u = (1, 0, \cdots, 0)^T$. Let $R > 0$ be a constant, and consider

   $$M = \frac{1}{\sqrt{n}}W + Ruu^T,$$

   where W is the Wigner matrix in (1). Compute the eigenvalues of M and plot the histograms for $R = 0.1, 0.2, \cdots, 3$. What is the largest eigenvalue λ_1 like for different values of R? What is the correlation of the first eigenvector with u, that is $|v_1^T u|$, for different values of R? Repeat the experiment for the Bernoulli case, i.e. W as in (2). (Remark: the result is not sensitive to what unit vector u is. For the Wigner matrix W, when n is large, for any specific n-by-n rotation Q, QWQ^T is statistically "similar" to a Wigner matrix. Thus when u is not as in (1), one can change coordinate by some orthogonal matrix Q so that Qu is as the u in (1), and consider the perturbation of QWQ^T.)

2. (De-noising image patches by PCA) Take a natural image and generate patches of size k-by-k, $k = 7$ or 10 depending on the size and resolution of the image. Add gaussian noise to the clean image and generate patches of the noise-corrupted image. How does projecting to the first d principle components de-noise the patches? You can put back the patches to see how the image is de-noised (if you use overlapping patches then you can take the average on each pixels if it is covered by more than one patch).

   (1) How does the effect of de-noising depend on the choice of d? (Hint: from theory of high-dimensional PCA we know that as k increases, the k-th leading principle components has a larger amount of "noise" than information in it. Meanwhile, suppose that there is no noise, the PCA recovery of the clean image is poorer when d is smaller. Thus the choice of d is a matter of bias-variance trade-off.)

   (2) How does the optimal d change with the noise level, i.e. σ^2, suppose that the pixel-wise gaussian noise is distributed as $\mathcal{N}(0, \sigma^2)$?