HW2

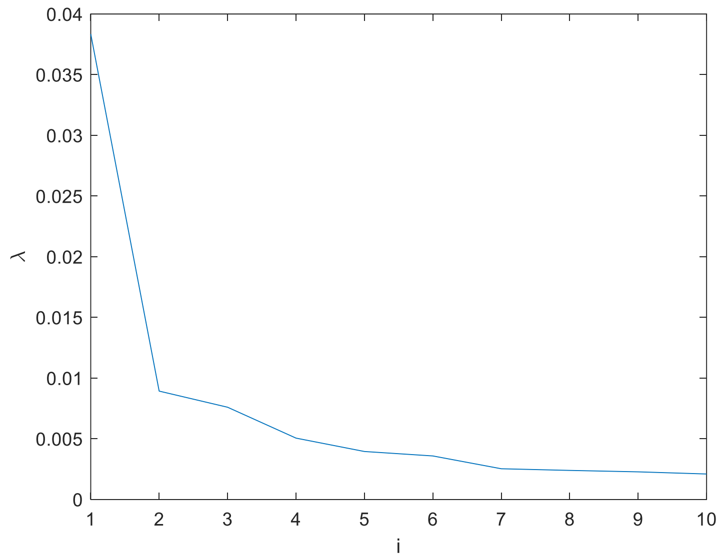
PENG, Han, 14/02/2024

1. Exploring S&P500 Stock Prices:

I put the calculation of problem (a) to (c) into the MATLAB file. Solutions to problem (d) and (e) are shown below.

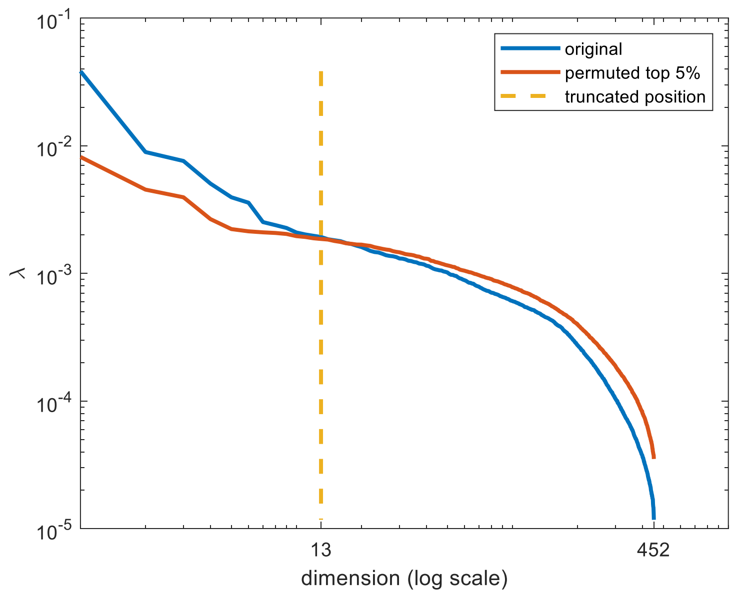
(d)

I plot first 10 eigen values for the covariance matrix in descending order.



(e)

I find first 13 eigen values make $pval_i < 0.05$. The plot is shown below. I observed that large eigen values will not be disturbed by the permutation of the rows of the covariance matrix.



2. Phase transition in PCA "spike" model

(a)

Since the noise variance is not 1, the equation from slides cannot be used directly. We need to replace λ by $\frac{\lambda}{\sigma^2}$ and $SNR = R = \frac{\lambda_0}{\sigma^2}$. Then the primary (largest) eigenvalue of sample covariance matrix is

$$\lambda = (1 + R)(1 + \frac{\gamma}{R})\sigma^2$$

(b)

According to the MP distribution, the upper bound of the eigen value for a Wishart matrix is $b = (1 + \sqrt{\gamma})^2$. And we know the maximum eigenvalue of the sample covariance matrix is within this upper bound if $SNR \leq \sqrt{\gamma}$ and out of this upper bound if $SNR > \sqrt{\gamma}$ as

$$\lambda = (1 + R)\left(1 + \frac{\gamma}{R}\right)\sigma^2 > \left(1 + \sqrt{\gamma}\right)^2 = b$$

Consequently, if we find that the maximum eigenvalue of the sample covariance matrix is less than b, then $SNR < \sqrt{\gamma}$. If maximum eigenvalue is larger or equal to b, this means that $SNR \geq \sqrt{\gamma}$.

(c)

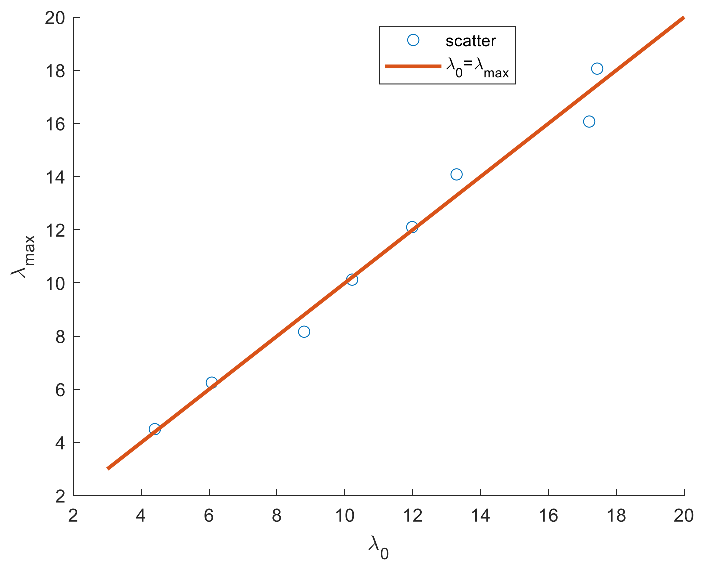
According to the lecture note, the squared correlation between the eigenvector v of the sample covariance matrix (corresponding to the largest eigenvalue λ) and the "true" signal component u converges to

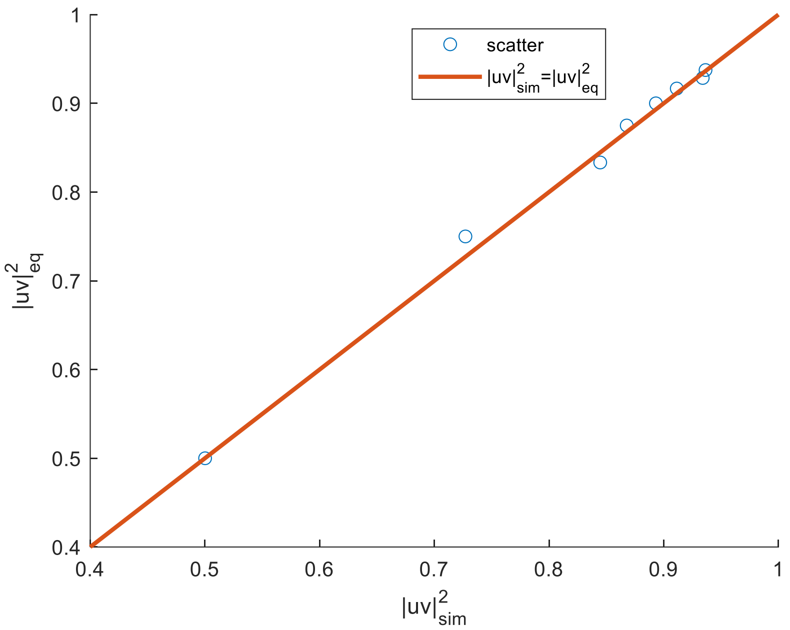
$$|\langle u, v_{max} \rangle|^2 = \begin{cases} 0 & \frac{\lambda_0}{\sigma^2} \leq \sqrt{\gamma} \\ \dfrac{1 - \frac{\gamma\sigma^4}{\lambda_0^2}}{1 + \frac{\gamma\sigma^2}{\lambda_0}} & \frac{\lambda_0}{\sigma^2} > \sqrt{\gamma} \end{cases}$$

(d)

I compared the simulation results λ_{sim}, $|uv|_{sim}^2$ with equation predications λ_{eq}, $|uv|_{eq}^2$. I take $n = 800, p = 400, \lambda_0 = 2,4, \ldots, 16$. Each element in vector u is random generated using normal distribution and then u is normalized.

We can see that the equation predications well match the simulation results.

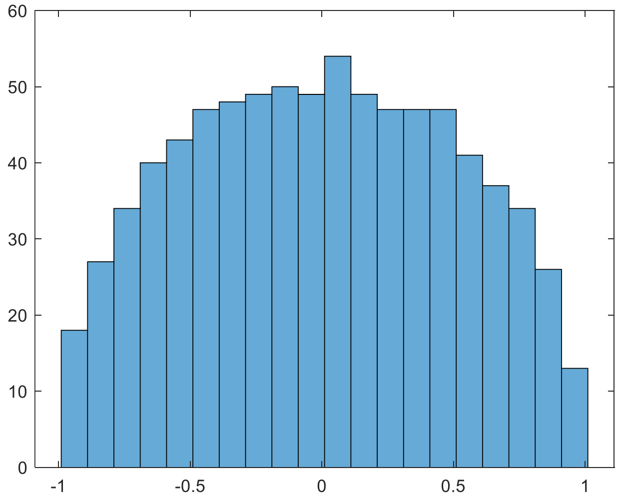




3. Finite rank perturbations of random symmetric matrices

(a)

I used $n = 800$ to generate the random symmetric matrix. The histogram is shown below and we can see that it satisfies Wigner's semi-circle law.



(b)

The general conclusion can be found in [1] and details of the prove can be found in [2]. Due to difficultly to understand the mathematical derivation in [2]. I choose to present the conclusion from [1].

Farzan [1] states that if $\lambda_0 > \sigma^2$, $\lambda_1 \to (R + \frac{1}{R})\sigma^2$, where $R = \frac{\lambda_0}{\sigma^2}$, and if $\lambda_0 < \sigma^2$ every eigenvalue is in the bulk. Consequently, the critical value of λ_0 for phase transfer is σ^2.

The squared correlation between the top eigenvector of $W + \lambda_0 uu^T$ and the vector u as a function of λ_0 is

$$|\langle u, v_{max} \rangle|^2 = \begin{cases} 0 & \frac{\lambda_0}{\sigma^2} \leq 1 \\ 1 - \frac{\sigma^2}{\lambda_0} & \frac{\lambda_0}{\sigma^2} > 1 \end{cases}$$

[1] *Haddadi, Farzan, and Arash Amini. "Eigenvectors of deformed Wigner random matrices."* IEEE Transactions on Information Theory *67.2 (2020): 1069-1079.*

[2] *Benaych-Georges, Florent, and Raj Rao Nadakuditi. "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices."* Advances in Mathematics *227.1 (2011): 494-521.*