ELEC 5450: Random Matrix Theory and Applications

Homework Assignment 1

Released on 2nd, March 2019

Department of Electronic and Computer Engineering

The Hong Kong University of Science and Technology

Due Date: 20th, March 2019

Grade: 20%

Submission Instructions: You should submit a hard copy of your solution to the TA during the lecture (you should also make a copy of your solution before submission), and submit to the TA through email the MATLAB code you would have generated to complete the assignment.

Important Note:

1) In order to get full marks, your solution should be very *clearly explained, including all relevant steps*.
2) **Students should conduct their own work, including their own numerical experiments. Any copying of work, including code among students is considered plagiarism and will be dealt with seriously**.

03.02.2019 Released

I. The Jacobi Ensemble

In random matrix theory, three classical ensembles have been extensively studied: the Gaussian, the Laguerre and the Jacobi ensembles. We are here interested in the double Wishart model, which belongs to the family of Jacobi ensembles.

For $m \geq N$, let $\mathbf{Y}$ be an $N \times m$ matrix whose entries are complex Gaussian distributed random variables with zero-mean and whose columns are pairwise independent with covariance matrix $\boldsymbol{\Sigma}$. The matrix $\mathbf{Y}\mathbf{Y}^\dagger$, where $\mathbf{Y}^\dagger$ stands for the conjugate transpose of the complex matrix $\mathbf{Y}$, is then a $N \times N$ complex Wishart matrix with m degrees of freedom and covariance matrix $\boldsymbol{\Sigma}$. We shall denote the complex Wishart distribution as $\mathcal{W}_N(m, \boldsymbol{\Sigma})$. We also use $\mathbb{N}$ to denote the set of natural numbers.

If $\mathbf{A} \sim \mathcal{W}_N(m_1, \boldsymbol{\Sigma})$ and $\mathbf{B} \sim \mathcal{W}_N(m_2, \boldsymbol{\Sigma})$ are independent, it is known that the probability density function of the eigenvalues of the matrix $\mathbf{C} \triangleq \mathbf{A}(\mathbf{A}+\mathbf{B})^{-1}$ is then

$$p(\lambda_1, ..., \lambda_N) \propto \prod_{i=1}^{N} \lambda_i^{m_1-N} (1-\lambda_i)^{m_2-N} \prod_{i<j} (\lambda_i - \lambda_j)^2 \tag{1}$$

where $m_1, m_2 > N$ and $\lambda_1, ..., \lambda_N \in [0,1]$. The matrix $\mathbf{C}$ is known as complex double Wishart matrix, which belongs to the family of Jacobi ensembles.

This ensemble is closely related to the Jacobi polynomials $P_k^{(\alpha,\beta)}(x)$ (here, $\alpha, \beta \in \mathbb{N}$), which are orthogonal on $[0,1]$ with respect to the Jacobi weight: $w_J^{(\alpha,\beta)}(x) \triangleq x^\alpha (1-x)^\beta$, i.e.,

$$\int_0^1 w_J^{(\alpha,\beta)}(x) P_k^{(\alpha,\beta)}(x) P_l^{(\alpha,\beta)}(x)\, dx = \begin{cases} 1 & \text{if } k = l \\ 0 & \text{if } k \neq l \end{cases}. \tag{2}$$

For a given $k \in \mathbb{N}$, the associated Jacobi polynomial $P_k^{(\alpha,\beta)}(x)$ is of degree k, that is, it can be written in the form $P_k^{(\alpha,\beta)}(x) = \sum_i \nu_i x^i$ for a specific set of ν_i, with $\nu_k \neq 0$ and for all $i > k$, $\nu_i = 0$.

1) Prove that $\mathbf{A}+\mathbf{B} \sim \mathcal{W}_N(m_1+m_2, \boldsymbol{\Sigma})$.
2) Prove that it is equivalent to study the eigenvalues of the matrix $\mathbf{A}\mathbf{B}^{-1}$ or the eigenvalues of the matrix $\mathbf{C}$.
3) Justify that the distribution of the eigenvalues of $\mathbf{C}$ is not affected by the covariance matrix $\boldsymbol{\Sigma}$.

Henceforth, without loss of generality, we will assume that $\boldsymbol{\Sigma} = \mathbf{I}_N$ in the following.

II. Application: Distribution of Interference-Limited MIMO

Consider a Multiple-Input Multiple-Output (MIMO) wireless communication channel in which a user with n_t antennas transmits information to a receiver with n_r antennas, as shown in Figure 1. The received signal is also affected by signals being transmitted by K interferers, with each interferer having n_t antennas. The total interference power is P_I, which is split equally among all interferers. In addition, each interferer

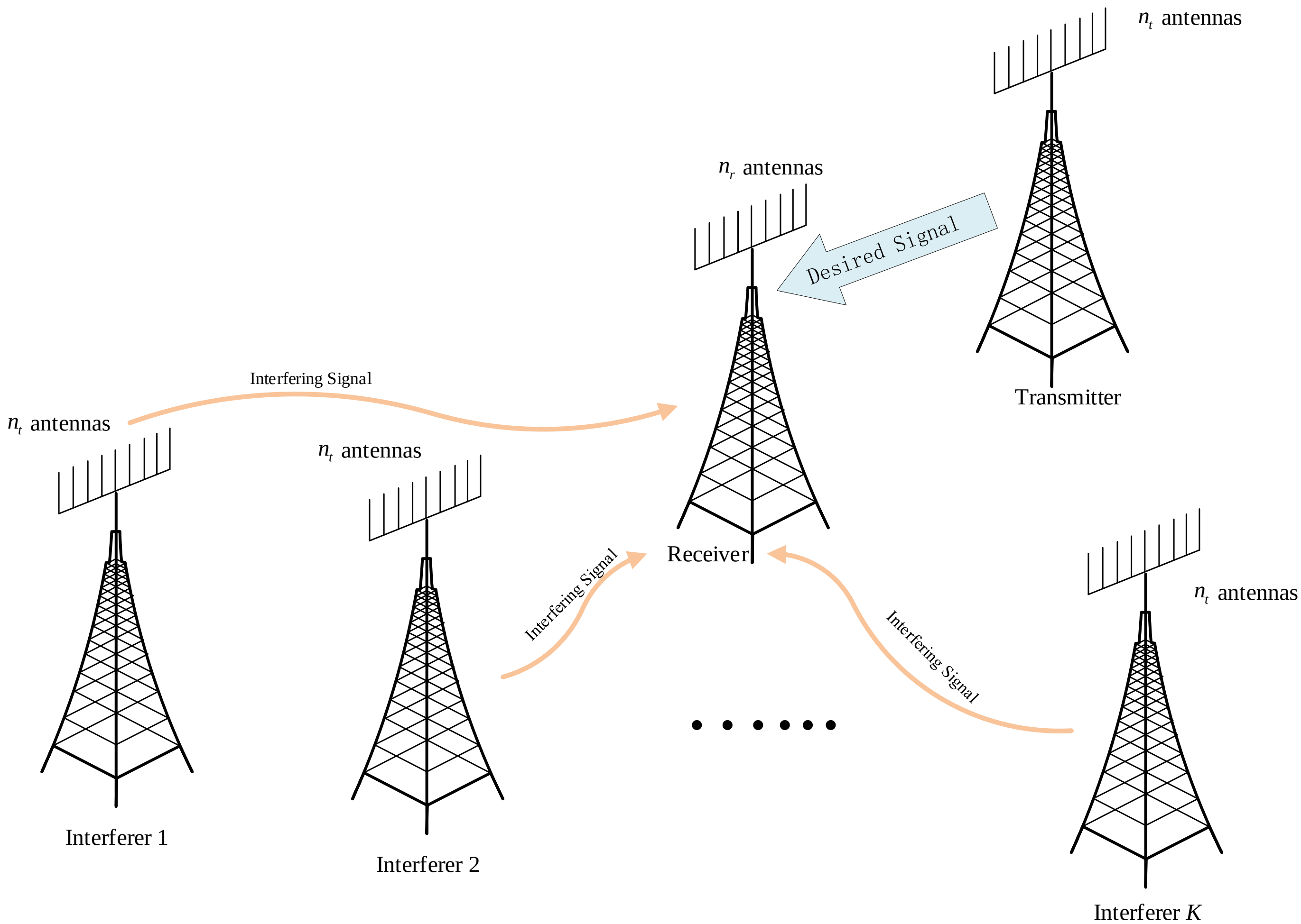

Figure 1. Interference-Limited MIMO system.

allocates its P_I/K power equally among their n_t transmit antennas. Furthermore, we assume that the system is interference-limited, which implies that the noise at the receiver can be neglected compared to the interference.

In this setting, the received signal vector $\mathbf{y} \in \mathbb{C}^{n_r}$ takes the form:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \sum_{i=1}^{K} \mathbf{H}_i \mathbf{x}_i,$$

where $\mathbf{H} \in \mathbb{C}^{n_r \times n_t}$ and $\mathbf{x} \in \mathbb{C}^{n_t}$ are the (fading) channel matrix and transmitted signal vector of the desired transmitter, and where $\mathbf{H}_i \in \mathbb{C}^{n_r \times n_t}$ and $\mathbf{x}_i \in \mathbb{C}^{n_t}$ represents the channel matrix and transmitted signal matrix of the i^{th} interferer, respectively. All channel matrices are assumed to have independent, identically distributed zero-mean and unit-variance complex Gaussian entries. The transmitted signal vectors are modeled as zero-mean complex Gaussian with covariance $E\left[\mathbf{x}\mathbf{x}^\dagger\right] = \frac{P}{n_t}\mathbf{I}_{n_t}$ and $E\left[\mathbf{x}_i\mathbf{x}_i^\dagger\right] = \frac{P_I}{Kn_t}\mathbf{I}_{n_t}$.

The goal of this section is to compute the ergodic capacity of the system, which is expressed as:

$$E\left[C\right]=E\left[\log\det\left(\mathbf{I}_{n_r}+\frac{P}{P_I/K}\mathbf{H}\mathbf{H}^{\dagger}\left(\mathbf{H}_I\mathbf{H}_I^{\dagger}\right)^{-1}\right)\right], \tag{3}$$

where $\mathbf{H}_I \triangleq [\mathbf{H}_1, ..., \mathbf{H}_K] \in \mathbb{C}^{n_r \times K n_t}$ is the aggregate channel matrix of the interferers.

The probability density function of the (unordered) eigenvalues $\mu_1, ..., \mu_{n_r}$ of the matrix $\mathbf{H}\mathbf{H}^{\dagger}\left(\mathbf{H}_I\mathbf{H}_I^{\dagger}\right)^{-1}$ is known to be given by

$$p\left(\mu_1, ..., \mu_{n_r}\right) \propto \prod_{k=1}^{n_r}\frac{\mu_k^{n_t-n_r}}{\left(1+\mu_k\right)^{(K+1)n_t}}\prod_{1\le i<j\le n_r}\left(\mu_i-\mu_j\right)^2 \tag{4}$$

where $\mu_i \in (0, +\infty)$.

1) Show how to rewrite (4) in the form of (1), i.e.,

$$p\left(\lambda_1, ..., \lambda_N\right) \propto \prod_{i=1}^{N}\lambda_i^{m_1-N}\left(1-\lambda_i\right)^{m_2-N}\prod_{i<j}\left(\lambda_i-\lambda_j\right)^2.$$

   In particular, you shall give the explicit change of variables between μ_i and λ_i, and the relation between the sets of parameters (n_t, n_r, K) and (m_1, m_2, N).

2) Consider N polynomials $P_0\left(x\right)$,...,$P_{N-1}\left(x\right)$ written in the form $P_k\left(x\right) = \sum_{i=0}^{k}\nu_i^{(k)}x^i$, such that $\nu_k^{(k)} \neq 0$ (in other words, $P_k\left(x\right)$ is of degree k). Prove that the quantity $\prod_{1\le i<j\le N}\left(\lambda_i-\lambda_j\right)^2$ can be written in terms of a determinant involving the polynomials $P_k\left(x\right)$. Write down the determinant explicitly.

   <u>Remark</u>: It is then from this question clear that $\prod_{1\le i<j\le N}\left(\lambda_i-\lambda_j\right)^2$ can be written in terms of the Jacobi polynomials $P_i^{(\alpha,\beta)}\left(x\right),\ i=1,...,N-1$.

3) Write down the probability density function of the eigenvalues $\lambda_1, ..., \lambda_N$ in terms of the Jacobi weight $w_J^{(\alpha,\beta)}\left(x\right)$ and the Jacobi polynomials. You don't have to determine the scaling (normalization) factor. Then, prove that it is possible to write $p\left(\lambda_1, ..., \lambda_N\right)$ as

$$p\left(\lambda_1, ..., \lambda_N\right) \propto \det\left(K\left(\lambda_j, \lambda_k\right)\right)_{j,k=1}^{N},$$

   where $K\left(\lambda, \mu\right)$ should be given in terms of the Jacobi weight and the Jacobi polynomials. Prove that $K\left(\lambda, \mu\right)$ has the three properties:

$$K\left(\lambda,\mu\right)=K\left(\mu,\lambda\right) \quad \text{(i)}$$

$$\int_0^1 K\left(\lambda,\lambda\right)d\lambda = N \quad \text{(ii)}$$

$$\int_0^1 K\left(\lambda,\mu\right)K\left(\mu,\nu\right)d\mu = K\left(\lambda,\nu\right) \quad \text{(iii)}$$

4) For the case $N = 2$, derive an exact expression for the marginal eigenvalue distribution $p(\lambda)$ of one arbitrary eigenvalue λ (including the explicit value of the leading normalization constant). Show in general that $p(\lambda)$ can be written as

$$p(\lambda) = \frac{1}{N} \sum_{i=0}^{N-1} \left(P_i^{(\alpha,\beta)}(\lambda) \right)^2 w_J^{(\alpha,\beta)}(\lambda).$$

5) Express the ergodic capacity $C_{\text{erg}} \triangleq E[C]$ in terms of $p(\lambda)$.

III. Numerical Simulations

The simulations in this section should be done in MATLAB.

1) For the case $N = 2$, generate 20000 realizations of the matrix $\mathbf{A} \sim \mathcal{W}_2(2, \mathbf{I}_2)$ and $\mathbf{B} \sim \mathcal{W}_2(2, \mathbf{I}_2)$. Compute the corresponding realizations of matrix $\mathbf{C} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}$. Plot the histogram of the set of eigenvalues of $\mathbf{C}$ produced by the 20000 realizations.

   Plot also the function $p(\lambda)$ (obtained in Section II, Question 4) superimposed on the same graph and verify that the simulated and theoretical curves agree. Repeat the same procedure for $m_2 = 4$, 6 and 8.

   Note: You may refer to the sample code provided on the course website to plot histograms, and to compute the coefficients of the Jacobi polynomials[1].

2) Generate 20000 realizations of the matrix $\mathbf{A} \sim \mathcal{W}_4(5, \mathbf{I}_4)$ and $\mathbf{B} \sim \mathcal{W}_4(6, \mathbf{I}_4)$. Compute the corresponding realizations of matrix $\mathbf{C} = \mathbf{A}(\mathbf{A} + \mathbf{B})^{-1}$. Plot the histogram of the set of eigenvalues of the 20000 matrices $\mathbf{C}$. Plot the function $p(\lambda)$ superimposed on the same graph. Repeat the same procedure for $(N, m_1, m_2) = (20, 25, 30)$, and $(N, m_1, m_2) = (40, 50, 60)$. What do you observe?

3) Generate 20000 realizations of the matrix $\mathbf{A} \sim \mathcal{W}_4(5, \mathbf{I}_4)$ and $\mathbf{B} \sim \mathcal{W}_4(6, \mathbf{I}_4)$. Compute the corresponding realizations of matrix $\tilde{\mathbf{C}} = \frac{1}{m_1}\mathbf{A}\left(\frac{1}{m_1}\mathbf{A} + \frac{1}{m_2}\mathbf{B}\right)^{-1}$. Plot the histogram of the set of eigenvalues of the 20000 matrices $\tilde{\mathbf{C}}$. Superimposed on the same graph, plot for $\lambda \in [0, 1]$ the function $p_{\text{Jac}}(\lambda)$ defined as

$$p_{\text{Jac}}(\lambda) = \frac{\sqrt{4b_0 b_2 - b_1^2}}{2\pi b_2},$$

   where

$$b_0 = (c_1 - c_2)\lambda - c_1 + 2$$
$$b_1 = (2c_1 - 2c_2)\lambda^2 + (2 - 3c_1 + c_2)\lambda + c_1 - 1$$
$$b_2 = (c_1 - c_2)\lambda^3 + (-2c_1 + c_2)\lambda^2 + c_1\lambda,$$

[1]Please refer to the README.txt file associated with MATLAB sample code for details.

with $c_1 = N/m_1$ and $c_2 = N/m_2$.

Repeat the same procedure for $(N, m_1, m_2) = (20, 25, 30)$, and $(N, m_1, m_2) = (40, 50, 60)$. What do you observe? Is $p_{\text{Jac}}(\lambda)$ a good approximation for the histogram of the set of eigenvalues of the 20000 matrices $\tilde{\mathbf{C}}$ for $(N, m_1, m_2) = (40, 50, 60)$? Is $p_{\text{Jac}}(\lambda)$ a good approximation for the histogram of the set of eigenvalues of the 20000 matrices $\tilde{\mathbf{C}}$ for $(N, m_1, m_2) = (4, 5, 6)$? Justify.

4) Numerically compute and plot the ergodic capacity of the system versus transmit power P (in dB[2]) using equation (3) for $n_t = n_r = 4$, $K = 4$, $P_I = 5$dB and P varying from 0 to 40dB. The capacity is obtained numerically by averaging the $\log\det$ quantity in (3) over many realizations of the random matrix $\mathbf{H}\mathbf{H}^\dagger \left(\mathbf{H}_I\mathbf{H}_I^\dagger\right)^{-1}$. Superimposed on the same graph, plot the analytical ergodic capacity (you derived in Section II, Problem 5) versus P (in dB) for the same parameter configuration. Verify that the simulated and theoretical curves agree. Also comment on the effect of increasing the power of desired transmitter P on the ergodic capacity. As P increases, how does the ergodic capacity scale with P?

   Note: if you need to compute an integral in MATLAB, you may use the fact that for a vector of points $x = 0 : step : 1$, the quantity $step \times \sum_{k=0}^{length(x)} f(x_k)$ is a good approximation of $\int_0^1 f(y)\, dy$ when step is small. In cases where symbolic computation is challenging, using this computational trick may be helpful.

   Recall that P denotes the total power of the desired transmitter (split equally among the antennas such that $E\left[\mathbf{x}\mathbf{x}^\dagger\right] = \frac{P}{n_t}\mathbf{I}_{n_t}$) while P_I denotes the total power across all the interferers (split equally among the interferers and equally across their antennas such that $E\left[\mathbf{x}_i\mathbf{x}_i^\dagger\right] = \frac{P_I}{Kn_t}\mathbf{I}_{n_t}$).

5) Fix $n_t = n_r = 4$. Plot the ergodic capacity of the system introduced in the previous section versus P (in dB, varying from 0 to 40dB) when $P_I = 5$dB and $K = 5$. Repeat this experiment for $K =$ 10, 15, ..., 35 and plot the curves on the same graph. Do the curves converge as K grows? If so, can you explain theoretically what they converge to?

   Plot the ergodic capacity $E[C_{\text{su}}]$ for the single user MIMO system (we encountered in the lecture) versus P (in dB), i.e.,

$$E[C_{\text{su}}] = E\left[\log\det\left(\mathbf{I} + \frac{P_{\text{eqv}}}{n}\mathbf{H}\mathbf{H}^\dagger\right)\right]$$

   for $n = n_r = n_t = 4$, and assume that $P_{\text{eqv}} = \frac{P}{P_I}$, with $P_I = 5$dB and P varying from 0 to 40dB. What do you observe? Explain your observations.

6) Fix $K = 1$, i.e., there is one interferer.

[2]dB is a logrithmic mearsure of a ratio. Mathmatically, $x(\text{in dB}) = 10\log_{10} x$

a) Plot the ergodic capacity $E\left[C\right]$ of the interference-limited MIMO system in (3) versus P (in dB) for $n_r = 2$, $n_t = 2$, $P_I = $ 5dB and P varying from 0 to 40dB. Superimposed on the same graph, plot the same curve for different antenna numbers $(n_r, n_t) = (4,4), (6,6)$ and $(8,8)$. Comment on the effect of increasing the number of antennas on the ergodic capacity.

b) Repeat the same experiment but plot the capacity **per receive antenna** (divide the capacity by n_r) versus P for different antenna size. Do the curves converge? Superimposed on the same graph, plot the capacity **per receive antenna** versus P using the approximation $p_{\mathrm{Jac}}\left(\lambda\right)$ introduced in Section III Question 3. What do you observe?