Solution for Exercise 4 (Wednesday Session)

- Let $\boldsymbol{e_i} \in \mathbb{R}^n$ denotes the vector with a 1 in component i and zeros in the remaining components, and let $\boldsymbol{e} \in \mathbb{R}^n$ denote the all-ones vector.
- Let $W \in \mathbb{R}^n$ denote the random vector that takes value $\boldsymbol{e_i}$ with probability $\mu_i > 0$, and let C denote the covariance matrix of W.

We assume that (after reordering if necessary) $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n$. The following result is obtained by an easy calculation (omitted in this abstract):

Lemma 1. *$\mathbb{E}[W] = \mu$ and the covariance matrix C of W is of the form*

$$C[i,j] = \begin{cases} \mu_i(1-\mu_i) & \text{if } i = j \\ -\mu_i\mu_j & \text{if } i \neq j \end{cases}$$

Lemma 2. *The eigenvalues $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ of C are as follows:*

1. *$\lambda_n = 0$ (with eigenvector $\boldsymbol{e}$).*
2. *Eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{n-1}$ are the zeros of the function*

$$h(\lambda) := \sum_{i=1}^{n} \mu_i \prod_{j:j\neq i} (\lambda - \mu_j) \tag{1}$$

so that

$$\forall k = 1, \ldots, n-1 : \mu_k \geq \lambda_k \geq \mu_{k+1} \ . \tag{2}$$

Proof. Observe that C can be written in the form

$$C = \operatorname{diag}(\mu_1, \ldots, \mu_n) - \mu\mu^\top \ . \tag{3}$$

The sub-additivity of the rank implies that C has rank at least $n-1$. Obviously, $\boldsymbol{e}$ is an eigenvector for eigenvalue $\lambda_n = 0$. Thus, the rank of C is exactly $n-1$. Furthermore, each of the $n-1$ remaining eigenvectors u must be orthogonal to $\boldsymbol{e}$:

$$\boldsymbol{e}^\top u = \sum_{i=1}^{n} u_i = 0 \ . \tag{4}$$

According to (3), equation $Cu = \lambda u$ can be written in the following form:

$$\forall i = 1, \ldots, n : \mu_i u_i - \mu_i \mu^\top u = \lambda u_i \ . \tag{5}$$

We proceed by case analysis.

Case 1: $\mu_1 = \cdots = \mu_n = 1/n$.

Then we can choose as eigenvectors an orthonormal base of the subspace induced by $\boldsymbol{e}^\top u = \mu^\top u = 0$ with eigenvalues

$$\lambda_1 = \cdots = \lambda_{n-1} = \frac{1}{n} \ .$$

For equal probability parameters, (1) collapses to

$$h(\lambda) = \left(\lambda - \frac{1}{n}\right)^{n-1} \ . \tag{6}$$

Note that $1/n$ is a zero of h with multiplicity $n-1$. This shows that $\lambda_1 = \cdots = \lambda_{n-1} = 1/n$ are indeed the zeros of h.

Case 2 $\exists i, j \in \{1, \ldots, n\} : \mu_i \neq \mu_j$.

Inspection of (5) reveals that no vector u satisfying $\mu^\top u = 0$ can be an eigenvector. Thus we can assume w.l.o.g. that an eigenvector u satisfies $\mu^\top u = -1$. Now (5) collapses to

$$\forall i = 1, \ldots, n : \mu_i u_i + \mu_i = \lambda u_i \ , \tag{7}$$

which implies that

$$u_i \prod_{j=1}^{n} (\lambda - \mu_j) = u_i(\lambda - \mu_i) \prod_{j:j\neq i} (\lambda - \mu_j) \overset{(7)}{=} \mu_i \prod_{j:j\neq i} (\lambda - \mu_j) \ . \tag{8}$$

holds for $i = 1, \ldots, n$. Now, we get

$$0 \overset{(4)}{=} \sum_{i=1}^{n} u_i = \sum_{i=1}^{n} u_i \prod_{j=1}^{n} (\lambda - \mu_j) \overset{(8)}{=} \sum_{i=1}^{n} \mu_i \prod_{j:j\neq i} (\lambda - \mu_j) \overset{(1)}{=} h(\lambda) \ .$$

As in case 1, the strictly positive eigenvalues coincide with the zeros of h.

Finally (2) is obtained by observing that

$$\forall i = k, \ldots, n-1 : h(\mu_k) = \mu_k \prod_{j:j\neq k} (\mu_k - \mu_j) \ .$$

This implies that, for $k = 1, \ldots, n-2$,

$$h(\mu_k) = h(\mu_{k+1}) = 0 \text{ or } \operatorname{sign}(h(\mu_k)) \neq \operatorname{sign}(h(\mu_{k+1}))$$

and a simple continuity argument shows that the k-th zero of h is found in the interval $[\mu_{k+1}, \mu_k]$. □