Summer Semester 2020
Markus Müller, Felix Binder, Andrew Garner
http://www.mpmueller.net/resource.html

Exercise - Week 1

Problem 1.1: Von Neumann entropy and thermal states (10 points)

Given a quantum system with Hamiltonian H, the thermal state (Gibbs state) $\rho_\beta(H)$ at inverse temperature $\beta = 1/(k_B T)$ is defined as

$$\rho_\beta(H) := \frac{e^{-\beta H}}{Z}, \qquad \text{where } Z := \mathrm{tr}(e^{-\beta H}).$$

The von Neumann entropy S and quantum relative entropy (Kullback-Leibler divergence) D are defined as

$$\begin{aligned} S(\rho) &:= -\mathrm{tr}(\rho \log \rho), \\ D(\rho\|\sigma) &:= \mathrm{tr}(\rho \log \rho - \rho \log \sigma) \end{aligned}$$

where ρ and σ are density matrices. In the following, we use the fact that $D(\rho\|\sigma) \geq 0$, with equality if and only if $\rho = \sigma$. We also define the free energy of a state ρ (at inverse temperature $\beta \neq 0$) as

$$F_\beta(\rho) := \langle H \rangle_\rho - S(\rho)/\beta,$$

where $\langle H \rangle_\rho := \mathrm{tr}(\rho H)$.

(i) Show that if $H = H_A + H_B$ is a non-interacting Hamiltonian on a bipartite quantum system AB, then $\rho_\beta(H) = \rho_\beta(H_A) \otimes \rho_\beta(H_B)$.

(ii) Show that $S\left(\rho_\beta(H)\right) = \beta \langle H \rangle_{\rho_\beta(H)} + \log Z$.

(iii) Show that $\frac{1}{\beta} D(\rho\|\rho_\beta(H)) = F_\beta(\rho) - F_\beta(\rho_\beta(H))$.

(iv) Prove that $\rho_\beta(H)$ is the unique state that minimizes the free energy.

(v) Consider the expression $D(\rho_{AB}\|\rho_A \otimes \rho_B)$, and prove that von Neumann entropy satisfies the subadditivity property $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$, with equality if and only if $\rho_{AB} = \rho_A \otimes \rho_B$. (This is sometimes just written $S(AB) \leq S(A) + S(B)$.)

Summer Semester 2020
Markus Müller, Felix Binder, Andrew Garner
http://www.mpmueller.net/resource.html

Exercises - Week 2

Majorization is concerned with a (pre)order on probability distributions. Writing $x \succ y$ means that y is "closer" to the uniform distribution than x or equivalently y is "more uniform" than x. On this exercise sheet x, y, z will be *probability vectors* of size n – they have non-negative components and their elements sum to one. For each probability vector x we define the vector $x^{\downarrow}$ whose elements are same as in x, but are ordered in decreasing size $x_1^{\downarrow} \geq \ldots \geq x_n^{\downarrow}$. We say x *majorizes* y and write

$$x \succ y \quad \text{iff} \quad \sum_{i=1}^{k} x_i^{\downarrow} \geq \sum_{i=1}^{k} y_i^{\downarrow} \quad \text{for} \quad k = 1, \ldots, n$$

This definition is easy to check if you have two given vectors but it is a bit abstract and not really clear why it captures the notion of "uniformness". The aim of this exercise sheet is to develop a better understanding of majorization, and to present some equivalent statements to the above definition. First convince yourself that $x \succ x$ and if $x \succ y$, then $Px \succ P'y$ for arbitrary permutations P and P'. If $x \succ y$ and $y \succ x$, does this mean that $x \equiv y$?

Problem 2.1: Drawing some Lorenz curves (6 points)

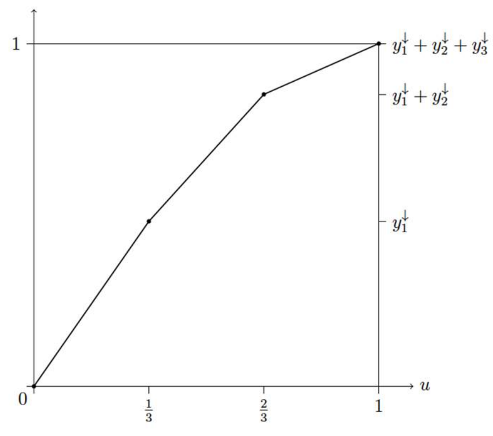
If we have a probability vector x we can define the vector $s(x)$ of partial sums of x as

$$s(x)_i = \sum_{j=1}^{i} x_j.$$

which is an n–component nonnegative vector. If we draw the vector $s(x^{\downarrow})$ as a piecewise linear function we get the *Lorenz curve* (named after the economist Max O. Lorenz).

On the right we see the Lorenz curve for a three component probability vector. It is a piecewise linear function on the interval $[0, 1]$ which has the value $s(x^{\downarrow})_i$ at the point i/n. The majorization condition can be expressed in terms of Lorenz curves in the following way

$$x \succ y \quad \text{if} \quad s(x^{\downarrow})_i \geq s(y^{\downarrow})_i \quad \forall i$$



Now choose $n = 4$ and write down an arbitrary choice of probability vector y. Find two more probability distributions x and z such that $x \succ y \succ z$ and draw the Lorenz curves for all three in one diagram. (Please don't use $(1, 0, 0, 0)$ or $(1/n, 1/n, 1/n, 1/n)$ for any of x, y or z.)
Furthermore find a probability vector y' such that neither $y \succ y'$ nor $y' \succ y$ and draw the corresponding Lorenz curves. Try to argue graphically that the following holds for arbitrary y:

$$(1, 0, \ldots, 0) \succ y \succ (1/n, \ldots, 1/n). \tag{1}$$

Problem 2.2: The convex set of majorized states (10 points)

In this exercise we will prove that the set of probability vectors majorized by some x is convex. That is:

$$\text{If } x \succ y \text{ and } x \succ z \quad \text{then} \quad x \succ ty + (1-t)z \qquad \text{for } t \in [0,1]. \tag{2}$$

First, show that:

(i) $s(x^{\downarrow})_i \geq s(Px^{\downarrow})_i$ for all i where $Px^{\downarrow}$ is some permutation of the vector $x^{\downarrow}$.

(ii) For real numbers a, b, c with $a \geq b$ and $a \geq c$, it holds that $tb + (1-t)c \leq a$ for $t \in [0,1]$.

Now use these to prove the convexity statement above.
Convince yourself that the following slightly more general statement holds as well:

$$\text{If } x \succ y^i \quad \text{then} \quad x \succ p_1 y^1 + \cdots + p_k y^k,$$

for probability vectors y^i and all probability distribution $p = (p_1, \ldots, p_k)$.
For a given x we always know that the set of its permutations form one such collection of y^i that are majorized by x. Thus, the set of distributions majorized by some x contains all convex mixtures of the permutations of x. It can be further shown that this set in fact contains *all* the distributions majorized by x.

Theorem: The vector $x \succ y$ if and only if there exists a probability distribution p and a set of permutations $\{P_i\}$ such that

$$y = \sum_i p_i \, P_i x. \tag{3}$$

(You do not need to prove this.)
Show that this theorem implies the relations in equation (1).

Problem 2.3: Drawing the set of probability vectors (8 points)

For higher dimensions the set of probability vectors is quite difficult to visualize but for $n = 3$ we can draw some insightful pictures. Particularly, this set is the intersection of the positive octant of $\mathbb{R}^3$ with the plane defined by $x_1 + x_2 + x_3 = 1$. Draw this set and indicate the points $(1,0,0)$, $(0,1,0)$, $(0,0,1)$ and $(1/3, 1/3, 1/3)$.
Since the set is two dimensional it is convenient to forget about its embedding in $\mathbb{R}^3$ and consider only its intersection with the plane. Draw this also, and indicate the above points. Now pick a probability distribution x with $x_1 \neq x_2 \neq x_3$ and indicate it in the set together with all its permutations. Use what you know from the previous exercise to indicate the set of all vectors that are majorized by x. (**Bonus:** *What shape would the set of $n = 4$ probabilities take if we only drew the intersection with $x_1 + x_2 + x_3 + x_4 = 1$?)*

We know that majorization does not really care about the order of the vector components so lets restrict ourselves to the part of the probability distributons which are ordered decreasingly $x_1 \geq x_2 \geq x_3$. Find out which part of the whole set this is and draw it. In this subset now choose again a point x and draw in addition to the set that x is majorizing as well the set of elements that majorize x. (To do work out the latter, you could try to look systematically at some Lorenz curves.) In the end you should end up with three different regions: the points majorizing x, the points majorized by x and the points that are not comparable to x.

Problem 2.4 (Optional): Permutations and other bi-stochastic matrices (10 points)

A stochastic matrix B of size n is a $n \times n$ real matrix whose matrix elements satisfy the first two of the following properties:

$$(i): \quad B_{ij} \geq 0 \qquad (ii): \quad \sum_{i=1}^{n} B_{ij} = 1 \qquad (iii): \quad \sum_{j=1}^{n} B_{ij} = 1.$$

If in addition the matrix elements obeys (iii) we call B a *bi-stochastic matrix* (a.k.a. a *doubly stochastic matrix*).
Prove the following statements:-

(i) The stochastic matrices are exactly the linear maps that map probability vectors into probability vectors.

(ii) Bistochastic matrices are exactly the subset of stochastic matrices that leave the maximally mixed state $(1/n, \ldots, 1/n)$ invariant.

(iii) Permutation matrices are the only bi-stochastic matrices that are *reversible* – that is, they are invertible and their inverse is also a bi-stochastic matrix.

(iv) The set of bi-stochastic matrices is convex. Thus if B_i are bi-stochastic and p is a probaility distribution than $\sum_i p_i B_i$ is bi-stochastic.

The converse of the last statement holds as well:

Theorem (Birkhoff & von Neumann): A matrix B is bi-stochastic if and only if there exist permutations P_i and and a probability distribution p such that

$$B = \sum_i p_i P_i \ .$$

(You do not need to prove this.)
Use this together with the theorem of in Problem 2.2 to argue that:

$$x \succ y \quad \text{if and only if} \quad \exists\ B \text{ bi-stochastic with} \quad Bx = y.$$