Single-shot thermo, lecture 4

2.4 Majorization and Lorenz curves

Recall from last time: if $p, q \in \mathbb{R}^m$ are probability vectors, (i.e. $p_i, q_i \geq 0$, $\sum_i p_i = \sum_i q_i = 1$), then

$$p \succ q \Leftrightarrow \sum_{i=1}^{k} p_i^{\downarrow} \geq \sum_{i=1}^{k} q_i^{\downarrow} \quad \text{for all } k = 1, \ldots, m,$$

where $p_i^{\downarrow}$ are the entries of p in non-increasing order.

Furthermore, write $p \xrightarrow{\text{noisy}} q$ if for every $\varepsilon > 0$ there exists q_ε with $\|q - q_\varepsilon\| < \varepsilon$ and a noisy classical operation D_ε with $D_\varepsilon(p) = q_\varepsilon$.

<u>Lemma</u>: Let $p, q \in \mathbb{R}^m$ be probability vectors of the same size. Then $p \xrightarrow{\text{noisy}} q$ if and only if $p \succ q$.

<u>Proof</u>: "$\Rightarrow$": Suppose $p \xrightarrow{\text{noisy}} q$. Consider D_ε, then D_ε maps prob. vectors to prob. vectors, and $D_\varepsilon(1/m, \ldots, 1/m) = (1/m, \ldots, 1/m)$, D_ε linear. Problem 8, Exercise 2 $\Rightarrow$ D_ε is a bistochastic matrix, and thus $p \succ q_\varepsilon$ for every $q_\varepsilon := D_\varepsilon(p)$. But the set of prob. vectors majorized by p is topologically closed $\Rightarrow$ $p \succ q$.

"$\Leftarrow$": Suppose $p \succ q$. Problem 6, Exercises 2 $\Rightarrow$

q can be obtained from p as a mixture of permutations

$$q = \sum_{j=1}^{\ell} \alpha_j \pi_j(p), \quad \alpha_j \geq 0, \ \sum_j \alpha_j = 1, \text{ all } \pi_j \text{ permutations of entries}$$

$(\pi_j(p))_k = p_{\pi_j(k)}$.

Consider only special case $\ell = 2$ ($\ell \geq 3$ analogous).

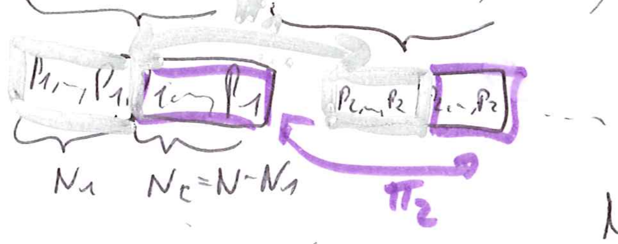
$$\Rightarrow \quad q = \alpha \pi_1(p) + (1-\alpha)\pi_2(p), \qquad 0 \leq \alpha \leq 1.$$

Construct a noisy operation with $p \to q_\varepsilon$:

$p = (p_1, \dots, p_m)$

bring on max. mixed state

$$p \otimes \underbrace{\gamma_N}_{\left(\frac{1}{N}, \dots, \frac{1}{N}\right)} = \frac{1}{N}\Big(\overbrace{p_1, \dots, p_1}^{N}, p_2, \dots, p_2, \dots, p_m, \dots, p_m\Big)$$



global permutation

$$\pi[p \otimes \gamma_N] := \frac{1}{N}\Big(\overbrace{p_{\pi_1(1)}, \dots, p_{\pi_1(1)}}^{N_1}, \overbrace{p_{\pi_2(1)}, \dots, p_{\pi_2(1)}}^{N_2},$$

$$\dots, \underbrace{p_{\pi_1(m)}, \dots, p_{\pi_1(m)}}_{N_1}, \underbrace{p_{\pi_2(m)}, \dots, p_{\pi_2(m)}}_{N_2}\Big)$$

marginal (sum over ancilla) $\to$ sum over N-blocks

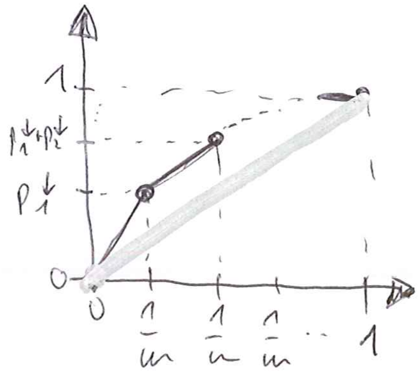
$$(\pi[p \otimes \gamma_N])_{\text{out}} = \frac{1}{N}\Big(N_1 p_{\pi_1(1)} + N_2 p_{\pi_2(1)}, \dots, N_1 p_{\pi_1(m)} + N_2 p_{\pi_2(m)}\Big)$$

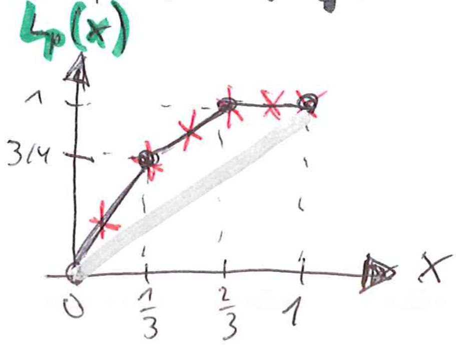
$$= \frac{N_1}{N}\pi_1(p) + \frac{N_2}{N}\pi_2(p).$$

By making $N \to \infty$ and choosing $N_1 = N_1(N)$ appropriately, we can approximate α to arbitrary accuracy. $\square$

Graphical tool: Lorenz curve: linear interpolation of $\left(\sum_{i=1}^{k} p_i^{\downarrow}\right)_{k=0,\dots,m}$

Prob. vector $p \in \mathbb{R}^m$. Ex.: $p=(0,\frac{3}{4},\frac{1}{4})$





Observation: p and $p\otimes\gamma_N = p\otimes(\frac{1}{N},\dots,\frac{1}{N})$ have identical Lorenz curves

e.g. $N=2$: $p\otimes(\frac{1}{2},\frac{1}{2}) = (0,0,\frac{3}{8},\frac{3}{8},\frac{1}{8},\frac{1}{8})$

Lemma: $p \xrightarrow{\text{noisy}} q$ if and only if $L_p(x) \geq L_q(x)$ for all x, even if p and q do not have the same [redacted] dimensionalities (i.e. $p\in\mathbb{R}^m$, $q\in\mathbb{R}^n$).

Proof: If $m=n$ then $L_p \geq L_q \iff p \succ q$ by definition.

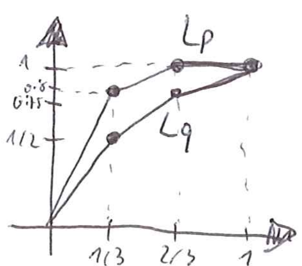
Otherwise: $p \xrightarrow{\text{noisy}} q \iff \underbrace{p\otimes\gamma_n}_{[redacted]} \xrightarrow{\text{noisy}} \underbrace{q\otimes\gamma_m}$

Same size

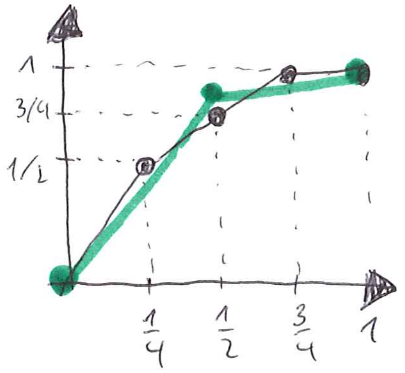
$\iff L_{p\otimes\gamma_n} \succ L_{q\otimes\gamma_m} \iff L_p \geq L_q$.

All Lorenz curves are concave, and lie above the Lorenz curve of $\gamma_n = (\frac{1}{n},\dots,\frac{1}{n})$.

$$\underbrace{\begin{pmatrix}1/5\\0\\4/5\end{pmatrix}}_{p} \succ \underbrace{\begin{pmatrix}1/2\\1/4\\1/4\end{pmatrix}}_{q}$$



p is a more valuable resource than q.



$p = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, 0\right)$, $q = \left(\frac{4}{5}, \frac{1}{5}\right)$

$\Rightarrow p \overset{\text{noisy}}{\not\longrightarrow} q$ and $q \overset{\text{noisy}}{\not\longrightarrow} p$

Neither resource is "more valuable" than the other.

2.5. Nonuniformity monotones

Let $\mathcal{P}_n$ be the set of probability vectors in $\mathbb{R}^n$.

Def.: A function $F: \mathcal{P}_n \to \mathbb{R}$ is Schur-convex (on $\mathcal{P}_n$) if $p \succ q \Rightarrow F(p) \geq F(q)$

(or Schur-concave if $p \succ q \Rightarrow F(p) \leq F(q)$).

Example: $H_\infty(p) := -\log \max_i p_i$

$$p \succ q \Rightarrow p_1^\downarrow \geq q_1^\downarrow \Rightarrow \max_i p_i \geq \max_i q_i$$

$$\Rightarrow H_\infty(p) \leq H_\infty(q)$$

H_∞ is Schur-concave.

We will now prove the intuition that "noisy operations increase the entropy".

Lemma: For every convex (concave) function $f: [0,1] \to \mathbb{R}$, the function $F: \mathcal{P}_n \to \mathbb{R}$ defined by

$$F(p) := \sum_{i=1}^{n} f(p_i)$$

is Schur-convex (Schur-concave).

(f convex $\iff f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$ $\forall \lambda \in [0,1]$.

If f is twice differentiable, then $f'' \geq 0 \implies$ [redacted] f convex.

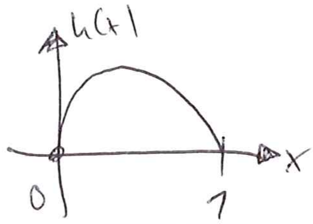
<u>Proof</u>: exercise.

Now consider Shannon entropy

$$H(p) := -\sum_{i=1}^{n} p_i \log p_i \qquad (0 \log 0 := 0).$$

$$= \sum_{i=1}^{n} h(p_i), \qquad h(p_i) = -p_i \log p_i.$$

h concave $\implies H$ Schur-concave.



$h''(x) = -\frac{1}{x} \leq 0$

$p \succ q \implies \underline{H(p) \leq H(q)}$

$\implies$ $\boxed{\text{Noisy operations } D: \mathbb{R}^n \to \mathbb{R}^n \text{ [redacted] never decrease the entropy.}}$

Reminiscent of "Second Law".

For different dimensionalities: $p \in \mathbb{R}^m$, $q \in \mathbb{R}^n$

$$\boxed{p \xrightarrow{\text{noisy}} q} \iff p \otimes \gamma_n \xrightarrow{\text{noisy}} q \otimes \gamma_m \iff p \otimes \gamma_n \succ q \otimes \gamma_m$$

$$\implies H(p \otimes \gamma_m) \leq H(q \otimes \gamma_n)$$

$$\implies H(p) + \log n \leq H(q) + \log m$$

$$\boxed{\implies I(p) \geq I(q) \quad \text{for } I(p) := \log \underbrace{d_p}_{\text{dim. of } p} - H(p) \quad \text{"negentropy"}}$$

This is necessary, but is it <u>sufficient</u>? No.

Recall $p = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, 0\right)$ $\quad q = \left(\frac{4}{5}, \frac{1}{5}\right)$ (for now, $\log \equiv \log_2$)

$\Rightarrow I(p) = \frac{1}{2}$, $\quad I(q) \approx$ [redacted] 0.28

(*) But [redacted] $p \overset{\text{noisy}}{\not\to} q$ (and $q \overset{\text{noisy}}{\not\to} p$). More generally "$\xrightarrow{\text{noisy}}$" is a preorder, but [redacted] I (like every other monotone) introduces an order:

$$p \succeq_I q :\Leftrightarrow I(p) \geq I(q), \text{ all states comparable.}$$

<u>Def</u>: A <u>nonuniformity monotone</u> is a function $F: \bigcup_{n \in \mathbb{N}} \mathcal{P}_n \to \mathbb{R}$ such that

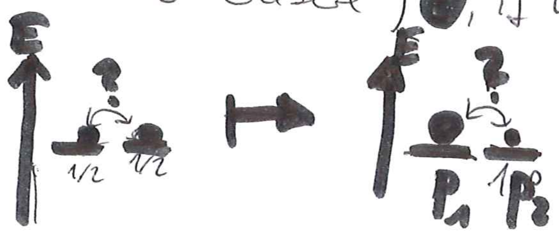
$$p \xrightarrow{\text{noisy}} q \implies F(p) \geq F(q).$$

(Example: negentropy I.)

Because of (*), no single monotone can fully characterize the possible state transitions.

26. Extracting work from states: distillable nonuniformity and nonuniformity of formation

How much work does it cost to create a state (as in Landauer erasure) [redacted], if we want to succeed with probability one?



We will show:

$$\left(\log d_p - H_\infty(p)\right) k_B T \ln 2$$

How much work can we extract from a state (in a Szilard engine)?

We show: $(\log d_p - H_0(p))\, k_B T \ln 2$

where $\underbrace{H_\infty}_{\text{"min entropy"}} \leq \underbrace{H}_{\text{Shannon entropy}} \leq \underbrace{H_0}_{\text{"max entropy"}}$.

- Fundamental irreversibility: have to invest more work than we can get out of [redacted] the state later on.
- But we will show: reversibility is recovered in the thermodynamic limit (work extraction "on average" for many copies).

<u>Def.</u> (Rényi entropies): Let $p \in \mathbb{R}^n$ be any probability vector. For $\alpha > 0$, $\alpha \neq 1$, define $H_\alpha(p) := \frac{1}{1-\alpha} \log \sum_{i=1}^n p_i^\alpha$

By taking the limits, we also get

$H_0(p) := \log \operatorname{rank}(p)$, where $\operatorname{rank}(p) := \#$ of nonzero entries of p,

$H_1(p) = H(p) = -\sum_{i=1}^n p_i \log p_i$ Shannon entropy

$H_\infty(p) = -\log \max_i p_i$.

Analogously for quantum states ρ: $H_\alpha(\rho) = \frac{1}{1-\alpha} \log \operatorname{tr}(\rho^\alpha)$

$H_0(\rho) = \log \operatorname{rank}(\rho)$, $H_1(\rho) = S(\rho)$ von Neumann entropy

$H_\infty(\rho) = -\log \underbrace{\lambda_{max}(\rho)}_{\text{largest eigenvalue}}$