Catalytic majorization in quantum information theory

Guillaume AUBRUN[1]

[1]Université Lyon 1, France
Joint work with avec Ion NECHITA

Samos, 27 June 2007

Majorization

Let $P_d = \{x \in \mathbf{R}^d, x_i \geqslant 0, \sum x_i = 1\}$; $P_d \subset P_{d+1} \subset \dots$
x^* : decreasing rearrangement of x.

Definition

Let $x, y \in P_d$; x is majorized by y ($x \prec y$) if

$$\forall k = 1, \dots, d \qquad \sum_{i=1}^{k} x_i^* \leqslant \sum_{i=1}^{k} y_i^*$$

$\forall x \in P_d, (\frac{1}{d}, \dots, \frac{1}{d}) \prec x \prec (1, 0, \dots, 0)$.

Proposition

$x \prec y \iff$ There is a bistochastic matrix B so that $x = By$.

For $y \in P_d$, let $S_d(y) = \{x \in P_d \text{ s.t. } x \prec y\}$.
Then $S_d(y) = \text{conv}\{(y_{\sigma(1)}, \dots, y_{\sigma(d)}), \sigma \in \mathfrak{S}_d\}$ is a convex polytope.

Majorization

Let $P_d = \{x \in \mathbf{R}^d, x_i \geqslant 0, \sum x_i = 1\}$; $P_d \subset P_{d+1} \subset \dots$
x^* : decreasing rearrangement of x.

Definition

Let $x, y \in P_d$; x is majorized by y ($x \prec y$) if

$$\forall k = 1, \dots, d \qquad \sum_{i=1}^{k} x_i^* \leqslant \sum_{i=1}^{k} y_i^*$$

$\forall x \in P_d, (\frac{1}{d}, \dots, \frac{1}{d}) \prec x \prec (1, 0, \dots, 0)$.

Proposition

$x \prec y \iff$ There is a bistochastic matrix B so that $x = By$.

For $y \in P_d$, let $S_d(y) = \{x \in P_d \text{ s.t. } x \prec y\}$.
Then $S_d(y) = \text{conv}\{(y_{\sigma(1)}, \dots, y_{\sigma(d)}), \sigma \in \mathfrak{S}_d\}$ is a convex polytope.

Majorization

Let $P_d = \{x \in \mathbf{R}^d, x_i \geqslant 0, \sum x_i = 1\}$; $P_d \subset P_{d+1} \subset \dots$
x^* : decreasing rearrangement of x.

Definition

Let $x, y \in P_d$; x is majorized by y ($x \prec y$) if

$$\forall k = 1, \dots, d \qquad \sum_{i=1}^{k} x_i^* \leqslant \sum_{i=1}^{k} y_i^*$$

$\forall x \in P_d, (\frac{1}{d}, \dots, \frac{1}{d}) \prec x \prec (1, 0, \dots, 0)$.

Proposition

$x \prec y \iff$ *There is a bistochastic matrix B so that $x = By$.*

For $y \in P_d$, let $S_d(y) = \{x \in P_d \text{ s.t. } x \prec y\}$.
Then $S_d(y) = \text{conv}\{(y_{\sigma(1)}, \dots, y_{\sigma(d)}), \sigma \in \mathfrak{S}_d\}$ is a convex polytope.

Majorization

Let $P_d = \{x \in \mathbf{R}^d, x_i \geqslant 0, \sum x_i = 1\}$; $P_d \subset P_{d+1} \subset \dots$
x^* : decreasing rearrangement of x.

Definition

Let $x, y \in P_d$; x is majorized by y ($x \prec y$) if

$$\forall k = 1, \dots, d \qquad \sum_{i=1}^{k} x_i^* \leqslant \sum_{i=1}^{k} y_i^*$$

$\forall x \in P_d, (\frac{1}{d}, \dots, \frac{1}{d}) \prec x \prec (1, 0, \dots, 0)$.

Proposition

$x \prec y \iff$ *There is a bistochastic matrix B so that $x = By$.*

For $y \in P_d$, let $S_d(y) = \{x \in P_d \text{ s.t. } x \prec y\}$.
Then $S_d(y) = \text{conv}\{(y_{\sigma(1)}, \dots, y_{\sigma(d)}), \sigma \in \mathfrak{S}_d\}$ is a convex polytope.

Majorization

Let $P_d = \{x \in \mathbf{R}^d, x_i \geqslant 0, \sum x_i = 1\}$; $P_d \subset P_{d+1} \subset \dots$
x^* : decreasing rearrangement of x.

Definition

Let $x, y \in P_d$; x is majorized by y ($x \prec y$) if

$$\forall k = 1, \dots, d \qquad \sum_{i=1}^{k} x_i^* \leqslant \sum_{i=1}^{k} y_i^*$$

$\forall x \in P_d, (\frac{1}{d}, \dots, \frac{1}{d}) \prec x \prec (1, 0, \dots, 0)$.

Proposition

$x \prec y \iff$ *There is a bistochastic matrix B so that $x = By$.*

For $y \in P_d$, let $S_d(y) = \{x \in P_d \text{ s.t. } x \prec y\}$.
Then $S_d(y) = \text{conv}\{(y_{\sigma(1)}, \dots, y_{\sigma(d)}), \sigma \in \mathfrak{S}_d\}$ is a convex polytope.

In quantum information theory

Alice and Bob share a quantum system: its states are given by unit vectors in a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Theorem (Nielsen)

Let ϕ and ψ be two such states and write their Schmidt decomposition

$$\phi = \sum_{i=1}^{d_1} \sqrt{x_i} e_i^A \otimes e_i^B \quad \text{and} \quad \psi = \sum_{i=1}^{d_2} \sqrt{y_i} f_i^A \otimes f_i^B$$

Then Alice and Bob can transform ϕ into ψ using local quantum operations and classical communication if and only if $x \prec y$.

Example : a state is separable if it is of the form $e^A \otimes e^B$. Separable states can be transformed only into separable states since no vector majorizes $(1, 0, \dots, 0)$.

In quantum information theory

Alice and Bob share a quantum system: its states are given by unit vectors in a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Theorem (Nielsen)

Let ϕ and ψ be two such states and write their Schmidt decomposition

$$\phi = \sum_{i=1}^{d_1} \sqrt{x_i} e_i^A \otimes e_i^B \quad \text{and} \quad \psi = \sum_{i=1}^{d_2} \sqrt{y_i} f_i^A \otimes f_i^B$$

Then Alice and Bob can transform ϕ into ψ using local quantum operations and classical communication if and only if $x \prec y$.

Example : a state is separable if it is of the form $e^A \otimes e^B$. Separable states can be transformed only into separable states since no vector majorizes $(1, 0, \ldots, 0)$.

In quantum information theory

Alice and Bob share a quantum system: its states are given by unit vectors in a Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$.

Theorem (Nielsen)

Let ϕ and ψ be two such states and write their Schmidt decomposition

$$\phi = \sum_{i=1}^{d_1} \sqrt{x_i} e_i^A \otimes e_i^B \quad \text{and} \quad \psi = \sum_{i=1}^{d_2} \sqrt{y_i} f_i^A \otimes f_i^B$$

Then Alice and Bob can transform ϕ into ψ using local quantum operations and classical communication if and only if $x \prec y$.

Example : a state is separable if it is of the form $e^A \otimes e^B$. Separable states can be transformed only into separable states since no vector majorizes $(1, 0, \ldots, 0)$.

Catalysis and multiple-copy transformations

- If $x \in P_d, x' \in P_{d'}$, then $x \otimes x' \in P_{dd'}$.
- Note that if $x_1 \prec y_1$ and $x_2 \prec y_2$, then $x_1 \otimes x_2 \prec y_1 \otimes y_2$ (tensorize bistochastic matrices).
- The converse is false: it can happen that
  1. $x \not\prec y$ but $\exists z$ so that $x \otimes z \prec y \otimes z$ (catalysis).
  2. $x \not\prec y$ but $\exists n$ so that $x^{\otimes n} \prec y^{\otimes n}$ (multiple-copies transformation).
- It makes sense to study these phenomena (you may want to perform massively quantum transformations and/or use surrounding entanglement as a catalyst).

Catalysis and multiple-copy transformations

- If $x \in P_d, x' \in P_{d'}$, then $x \otimes x' \in P_{dd'}$.
- Note that if $x_1 \prec y_1$ and $x_2 \prec y_2$, then $x_1 \otimes x_2 \prec y_1 \otimes y_2$ (tensorize bistochastic matrices).
- The converse is false: it can happen that
  1. $x \not\prec y$ but $\exists z$ so that $x \otimes z \prec y \otimes z$ (catalysis).
  2. $x \not\prec y$ but $\exists n$ so that $x^{\otimes n} \prec y^{\otimes n}$ (multiple-copies transformation).
- It makes sense to study these phenomena (you may want to perform massively quantum transformations and/or use surrounding entanglement as a catalyst).

Catalysis and multiple-copy transformations

- If $x \in P_d, x' \in P_{d'}$, then $x \otimes x' \in P_{dd'}$.
- Note that if $x_1 \prec y_1$ and $x_2 \prec y_2$, then $x_1 \otimes x_2 \prec y_1 \otimes y_2$ (tensorize bistochastic matrices).
- The converse is false: it can happen that
  1. $x \not\prec y$ but $\exists z$ so that $x \otimes z \prec y \otimes z$ (catalysis).
  2. $x \not\prec y$ but $\exists n$ so that $x^{\otimes n} \prec y^{\otimes n}$ (multiple-copies transformation).
- It makes sense to study these phenomena (you may want to perform massively quantum transformations and/or use surrounding entanglement as a catalyst).

Catalysis and multiple-copy transformations

- If $x \in P_d, x' \in P_{d'}$, then $x \otimes x' \in P_{dd'}$.
- Note that if $x_1 \prec y_1$ and $x_2 \prec y_2$, then $x_1 \otimes x_2 \prec y_1 \otimes y_2$ (tensorize bistochastic matrices).
- The converse is false: it can happen that
  1. $x \not\prec y$ but $\exists z$ so that $x \otimes z \prec y \otimes z$ (catalysis).
  2. $x \not\prec y$ but $\exists n$ so that $x^{\otimes n} \prec y^{\otimes n}$ (multiple-copies transformation).
- It makes sense to study these phenomena (you may want to perform massively quantum transformations and/or use surrounding entanglement as a catalyst).

Catalysis and multiple-copy transformations (2)

Introduce the following sets which generalize $S_d(y)$:

$$C_d(y) = \{x \in P_d \text{ s.t. } \exists k \in \mathbf{N}, \exists z \in P_k, x \otimes z \prec y \otimes z\}$$

$$M_d(y) = \{x \in P_d \text{ s.t. } \exists n \in \mathbf{N}, x^{\otimes n} \prec y^{\otimes n}\}$$

1. $C_d(y)$ is convex
2. $M_d(y) \subset C_d(y)$ — use $z = \frac{1}{n}(x^{\otimes(n-1)} \oplus x^{\otimes(n-2)} \otimes y \oplus \cdots \oplus y^{\otimes(n-1)})$
3. One may need arbitrarily large z and n
4. These sets are typically not closed

Question (#4 in R.Werner's list of QIT open problems)

How to describe the convex body $\overline{C_d(y)}$? Does $\overline{C_d(y)} = \overline{M_d(y)}$?

Catalysis and multiple-copy transformations (2)

Introduce the following sets which generalize $S_d(y)$:

$$C_d(y) = \{x \in P_d \text{ s.t. } \exists k \in \mathbf{N}, \exists z \in P_k, x \otimes z \prec y \otimes z\}$$

$$M_d(y) = \{x \in P_d \text{ s.t. } \exists n \in \mathbf{N}, x^{\otimes n} \prec y^{\otimes n}\}$$

1. $C_d(y)$ is convex
2. $M_d(y) \subset C_d(y)$ — use $z = \frac{1}{n}(x^{\otimes(n-1)} \oplus x^{\otimes(n-2)} \otimes y \oplus \cdots \oplus y^{\otimes(n-1)})$
3. One may need arbitrarily large z and n
4. These sets are typically not closed

Question (#4 in R.Werner's list of QIT open problems)

How to describe the convex body $\overline{C_d(y)}$? Does $\overline{C_d(y)} = \overline{M_d(y)}$?

Catalysis and multiple-copy transformations (2)

Introduce the following sets which generalize $S_d(y)$:

$$C_d(y) = \{x \in P_d \text{ s.t. } \exists k \in \mathbf{N}, \exists z \in P_k, x \otimes z \prec y \otimes z\}$$

$$M_d(y) = \{x \in P_d \text{ s.t. } \exists n \in \mathbf{N}, x^{\otimes n} \prec y^{\otimes n}\}$$

1. $C_d(y)$ is convex
2. $M_d(y) \subset C_d(y)$ — use $z = \frac{1}{n}(x^{\otimes(n-1)} \oplus x^{\otimes(n-2)} \otimes y \oplus \cdots \oplus y^{\otimes(n-1)})$
3. One may need arbitrarily large z and n
4. These sets are typically not closed

Question (#4 in R.Werner's list of QIT open problems)

How to describe the convex body $\overline{C_d(y)}$? Does $\overline{C_d(y)} = \overline{M_d(y)}$?

Catalysis and multiple-copy transformations (2)

Introduce the following sets which generalize $S_d(y)$:

$$C_d(y) = \{x \in P_d \text{ s.t. } \exists k \in \mathbf{N}, \exists z \in P_k, x \otimes z \prec y \otimes z\}$$

$$M_d(y) = \{x \in P_d \text{ s.t. } \exists n \in \mathbf{N}, x^{\otimes n} \prec y^{\otimes n}\}$$

1. $C_d(y)$ is convex
2. $M_d(y) \subset C_d(y)$ — use $z = \frac{1}{n}(x^{\otimes(n-1)} \oplus x^{\otimes(n-2)} \otimes y \oplus \cdots \oplus y^{\otimes(n-1)})$
3. One may need arbitrarily large z and n
4. These sets are typically not closed

Question (#4 in R.Werner's list of QIT open problems)

How to describe the convex body $\overline{C_d(y)}$? Does $\overline{C_d(y)} = \overline{M_d(y)}$?

Catalysis and multiple-copy transformations (2)

Introduce the following sets which generalize $S_d(y)$:

$$C_d(y) = \{x \in P_d \text{ s.t. } \exists k \in \mathbf{N}, \exists z \in P_k, x \otimes z \prec y \otimes z\}$$

$$M_d(y) = \{x \in P_d \text{ s.t. } \exists n \in \mathbf{N}, x^{\otimes n} \prec y^{\otimes n}\}$$

1. $C_d(y)$ is convex
2. $M_d(y) \subset C_d(y)$ — use $z = \frac{1}{n}(x^{\otimes(n-1)} \oplus x^{\otimes(n-2)} \otimes y \oplus \cdots \oplus y^{\otimes(n-1)})$
3. One may need arbitrarily large z and n
4. These sets are typically not closed

Question (#4 in R.Werner's list of QIT open problems)

How to describe the convex body $\overline{C_d(y)}$? Does $\overline{C_d(y)} = \overline{M_d(y)}$?

Necessary conditions

- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-convex if $x \prec y$ implies $f(x) \leqslant f(y)$.
- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-concave if $x \prec y$ implies $f(x) \geqslant f(y)$.
- For $x \in P_d$ and $p \in \mathbf{R}$, let

$$N_p(x) = \sum_{i=1}^{d} x_i^p.$$

- N_p is Schur-convex for $p \geqslant 1$ or $p \leqslant 0$ and Schur-concave for $0 \leqslant p \leqslant 1$. Also, $N_1(x) = 1$ and $N_0(x) = d$.
- N_p is multiplicative : $N_p(x \otimes z) = N_p(x) N_p(z)$.
- Therefore $x \in C_d(y)$ or $x \in M_d(y)$ imply $N_p(x) \leqslant N_p(y)$ or $N_p(x) \geqslant N_p(y)$ depending on p.
- Are there other conditions ?

Necessary conditions

- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-convex if $x \prec y$ implies $f(x) \leqslant f(y)$.
- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-concave if $x \prec y$ implies $f(x) \geqslant f(y)$.
- For $x \in P_d$ and $p \in \mathbf{R}$, let

$$N_p(x) = \sum_{i=1}^{d} x_i^p.$$

- N_p is Schur-convex for $p \geqslant 1$ or $p \leqslant 0$ and Schur-concave for $0 \leqslant p \leqslant 1$. Also, $N_1(x) = 1$ and $N_0(x) = d$.
- N_p is multiplicative : $N_p(x \otimes z) = N_p(x) N_p(z)$.
- Therefore $x \in C_d(y)$ or $x \in M_d(y)$ imply $N_p(x) \leqslant N_p(y)$ or $N_p(x) \geqslant N_p(y)$ depending on p.
- Are there other conditions ?

Necessary conditions

- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-convex if $x \prec y$ implies $f(x) \leqslant f(y)$.
- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-concave if $x \prec y$ implies $f(x) \geqslant f(y)$.
- For $x \in P_d$ and $p \in \mathbf{R}$, let

$$N_p(x) = \sum_{i=1}^{d} x_i^p.$$

- N_p is Schur-convex for $p \geqslant 1$ or $p \leqslant 0$ and Schur-concave for $0 \leqslant p \leqslant 1$. Also, $N_1(x) = 1$ and $N_0(x) = d$.
- N_p is multiplicative : $N_p(x \otimes z) = N_p(x) N_p(z)$.
- Therefore $x \in C_d(y)$ or $x \in M_d(y)$ imply $N_p(x) \leqslant N_p(y)$ or $N_p(x) \geqslant N_p(y)$ depending on p.
- Are there other conditions ?

Necessary conditions

- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-convex if $x \prec y$ implies $f(x) \leqslant f(y)$.
- A function $f : \mathbf{R}^d \to \mathbf{R}$ is Schur-concave if $x \prec y$ implies $f(x) \geqslant f(y)$.
- For $x \in P_d$ and $p \in \mathbf{R}$, let

$$N_p(x) = \sum_{i=1}^{d} x_i^p.$$

- N_p is Schur-convex for $p \geqslant 1$ or $p \leqslant 0$ and Schur-concave for $0 \leqslant p \leqslant 1$. Also, $N_1(x) = 1$ and $N_0(x) = d$.
- N_p is multiplicative : $N_p(x \otimes z) = N_p(x) N_p(z)$.
- Therefore $x \in C_d(y)$ or $x \in M_d(y)$ imply $N_p(x) \leqslant N_p(y)$ or $N_p(x) \geqslant N_p(y)$ depending on p.
- Are there other conditions ?

Nielsen's conjecture and our results

Conjecture (Nielsen)

Let $y \in P_d, y_i > 0$. Then $\overline{C_d(y)} = \overline{M_d(y)}$ = the set of $x \in P_d$ so that

(A) $N_p(x) \leqslant N_p(y)$ *for* $p \geqslant 1$.

(B) $N_p(x) \geqslant N_p(y)$ *for* $0 \leqslant p \leqslant 1$.

(C) $N_p(x) \leqslant N_p(y)$ *for* $p \leqslant 0$

We can prove partial results

Theorem (Aubrun+Nechita)

If $x, y \in P_d$ satisfy (A) and (B), then $x \in \overline{M_{d+1}(y)}$; idem with $C_{d+1}(y)$.

Theorem (Aubrun+Nechita)

If $x, y \in P_d$ satisfy (A), then $x \in \overline{\bigcup_{n \geqslant d} M_n(y)}^{\ell_1}$; idem with $\bigcup C_n(y)$.

Nielsen's conjecture and our results

Conjecture (Nielsen)

Let $y \in P_d, y_i > 0$. Then $\overline{C_d(y)} = \overline{M_d(y)}$ = the set of $x \in P_d$ so that

(A) $N_p(x) \leqslant N_p(y)$ *for* $p \geqslant 1$.

(B) $N_p(x) \geqslant N_p(y)$ *for* $0 \leqslant p \leqslant 1$.

(C) $N_p(x) \leqslant N_p(y)$ *for* $p \leqslant 0$

We can prove partial results

Theorem (Aubrun+Nechita)

If $x, y \in P_d$ satisfy (A) and (B), then $x \in \overline{M_{d+1}(y)}$; idem with $C_{d+1}(y)$.

Theorem (Aubrun+Nechita)

If $x, y \in P_d$ satisfy (A), then $x \in \overline{\bigcup_{n \geqslant d} M_n(y)}^{\ell_1}$; idem with $\bigcup C_n(y)$.

Nielsen's conjecture and our results

Conjecture (Nielsen)

Let $y \in P_d, y_i > 0$. Then $\overline{C_d(y)} = \overline{M_d(y)}$ = the set of $x \in P_d$ so that

(A) $N_p(x) \leqslant N_p(y)$ *for* $p \geqslant 1$.

(B) $N_p(x) \geqslant N_p(y)$ *for* $0 \leqslant p \leqslant 1$.

(C) $N_p(x) \leqslant N_p(y)$ *for* $p \leqslant 0$

We can prove partial results

Theorem (Aubrun+Nechita)

If $x, y \in P_d$ satisfy (A) and (B), then $x \in \overline{M_{d+1}(y)}$; idem with $C_{d+1}(y)$.

Theorem (Aubrun+Nechita)

If $x, y \in P_d$ satisfy (A), then $x \in \overline{\bigcup_{n \geqslant d} M_n(y)}^{\ell_1}$ *; idem with* $\bigcup C_n(y)$.

From vectors to random variables

- Idea (cf Kuperberg) : to a vector x associate the random variable

$$V_x \sim \sum_{i=1}^{d} x_i \delta_{\log x_i}.$$

- Then $V_{x \otimes y} \sim V_x \overset{\perp\!\!\perp}{+} V_y$.
- Rather than majorization on vectors, we will use the more flexible stochastic ordering on random variables

$$X \leqslant_{\text{st}} Y \iff \forall t \in \mathbf{R}, \mathbf{P}(X \geqslant t) \leqslant P(Y \geqslant t)$$

Equivalently $\tilde{X} \leqslant \tilde{Y}$ on some probability space (Strassen's theorem).

Lemma

If $V_x \leqslant_{\text{st}} V_y$, then $x \prec y$.

Remark : $V_x \leqslant_{\text{st}} V_y$ can hold only when size$(x) >$ size(y). This is why this approach fails to prove the complete conjecture.

From vectors to random variables

- Idea (cf Kuperberg) : to a vector x associate the random variable

$$V_x \sim \sum_{i=1}^{d} x_i \delta_{\log x_i}.$$

- Then $V_{x \otimes y} \sim V_x \overset{\perp\!\!\!\perp}{+} V_y$.
- Rather than majorization on vectors, we will use the more flexible stochastic ordering on random variables

$$X \leqslant_{\mathrm{st}} Y \iff \forall t \in \mathbf{R}, \mathbf{P}(X \geqslant t) \leqslant P(Y \geqslant t)$$

Equivalently $\tilde{X} \leqslant \tilde{Y}$ on some probability space (Strassen's theorem).

Lemma

If $V_x \leqslant_{\mathrm{st}} V_y$, then $x \prec y$.

Remark : $V_x \leqslant_{\mathrm{st}} V_y$ can hold only when size$(x) >$ size(y). This is why this approach fails to prove the complete conjecture.

From vectors to random variables

- Idea (cf Kuperberg) : to a vector x associate the random variable

$$V_x \sim \sum_{i=1}^{d} x_i \delta_{\log x_i}.$$

- Then $V_{x \otimes y} \sim V_x \overset{\perp\!\!\!\perp}{+} V_y$.
- Rather than majorization on vectors, we will use the more flexible stochastic ordering on random variables

$$X \leqslant_{\text{st}} Y \iff \forall t \in \mathbf{R}, \mathbf{P}(X \geqslant t) \leqslant P(Y \geqslant t)$$

Equivalently $\tilde{X} \leqslant \tilde{Y}$ on some probability space (Strassen's theorem).

Lemma

If $V_x \leqslant_{\text{st}} V_y$, then $x \prec y$.

Remark : $V_x \leqslant_{\text{st}} V_y$ can hold only when $\text{size}(x) > \text{size}(y)$. This is why this approach fails to prove the complete conjecture.

From vectors to random variables

- Idea (cf Kuperberg) : to a vector x associate the random variable

$$V_x \sim \sum_{i=1}^{d} x_i \delta_{\log x_i}.$$

- Then $V_{x \otimes y} \sim V_x \overset{\perp\!\!\!\perp}{+} V_y$.
- Rather than majorization on vectors, we will use the more flexible stochastic ordering on random variables

$$X \leqslant_{\text{st}} Y \iff \forall t \in \mathbf{R}, \mathbf{P}(X \geqslant t) \leqslant P(Y \geqslant t)$$

Equivalently $\tilde{X} \leqslant \tilde{Y}$ on some probability space (Strassen's theorem).

Lemma

If $V_x \leqslant_{\text{st}} V_y$, then $x \prec y$.

Remark : $V_x \leqslant_{\text{st}} V_y$ can hold only when $\text{size}(x) > \text{size}(y)$. This is why this approach fails to prove the complete conjecture.

$N_p(\cdot)$ and Laplace transforms

Recall that $V_x \sim \sum_{i=1}^d x_i \delta_{\log x_i}$. Then $N_{p+1}(x) = \mathbf{E} \exp(pV_x)$.
Inequalities on N_p translate into inequalities on the Laplace transforms.

Theorem (Cramér's large deviations theorem)

Let X be a r.v. and assume $\Lambda(\lambda) := \log \mathbf{E} e^{\lambda X} < +\infty$. Set

$$\Lambda^*(t) = \sup_{\lambda \in \mathbf{R}} \lambda t - \Lambda(\lambda).$$

Then for all $t \in (\min X, \max X)$, we have

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbf{P}(X_1 + \cdots + X_n \geqslant tn) = \begin{cases} 0 \text{ if } t \leqslant \mathbf{E}X \\ -\Lambda^*(t) \text{ if } t \geqslant \mathbf{E}X \end{cases}$$

where (X_i) denote i.i.d. copies of X.

$N_p(\cdot)$ and Laplace transforms

Recall that $V_x \sim \sum_{i=1}^d x_i \delta_{\log x_i}$. Then $N_{p+1}(x) = \mathbf{E} \exp(pV_x)$.
Inequalities on N_p translate into inequalities on the Laplace transforms.

Theorem (Cramér's large deviations theorem)

Let X be a r.v. and assume $\Lambda(\lambda) := \log \mathbf{E} e^{\lambda X} < +\infty$. Set

$$\Lambda^*(t) = \sup_{\lambda \in \mathbf{R}} \lambda t - \Lambda(\lambda).$$

Then for all $t \in (\min X, \max X)$, we have

$$\lim_{n \to \infty} \frac{1}{n} \log \mathbf{P}(X_1 + \cdots + X_n \geqslant tn) = \begin{cases} 0 \text{ if } t \leqslant \mathbf{E}X \\ -\Lambda^*(t) \text{ if } t \geqslant \mathbf{E}X \end{cases}$$

where (X_i) denote i.i.d. copies of X.

Stochastic ordering of i.i.d. sums

Corollary

Let X, Y be two random variables and assume that

1. $\forall \lambda > 0, \mathbf{E}e^{\lambda X} < \mathbf{E}e^{\lambda Y} < +\infty$,
2. $\forall \lambda < 0, \mathbf{E}e^{\lambda Y} < \mathbf{E}e^{\lambda X} < +\infty$,
3. $\mathbf{E}X < \mathbf{E}Y$
4. $\max X < \max Y$
5. $\min X < \min Y$

Then there exists a $n \geqslant 1$ so that

$$X_1 + \cdots + X_n \leqslant_{\text{st}} Y_1 + \cdots + Y_n$$

where $(X_i), (Y_i)$ are i.i.d. copies of X, Y.

The corollary is false when stated with $\leqslant$ instead of $<$ (we have a counterexample).

Stochastic ordering of i.i.d. sums

Corollary

Let X, Y *be two random variables and assume that*

1. $\forall \lambda > 0, \mathbf{E}e^{\lambda X} < \mathbf{E}e^{\lambda Y} < +\infty,$
2. $\forall \lambda < 0, \mathbf{E}e^{\lambda Y} < \mathbf{E}e^{\lambda X} < +\infty,$
3. $\mathbf{E}X < \mathbf{E}Y$
4. $\max X < \max Y$
5. $\min X < \min Y$

Then there exists a $n \geqslant 1$ *so that*

$$X_1 + \cdots + X_n \leqslant_{\text{st}} Y_1 + \cdots + Y_n$$

where $(X_i), (Y_i)$ *are i.i.d. copies of* X, Y.

The corollary is false when stated with $\leqslant$ instead of $<$ (we have a counterexample).

Sketch of proof

- Define
$$f_n(t) := \mathbf{P}(X_1 + \cdots + X_n \geqslant nt)$$
$$g_n(t) := \mathbf{P}(Y_1 + \cdots + Y_n \geqslant nt)$$
We want to show that $f_n \leqslant g_n$ for n large enough.
- Cramér's theorem gives $\lim \frac{1}{n} \log f_n$ on $[\mathbf{E}X, \max X]$, and also $\lim \frac{1}{n} \log(1 - f_n)$ on $[\min X, \mathbf{E}X]$; and similarly for g_n.
- Because all inequalities were assumed to be strict, we have also a strict inequality between the limits for (f_n) and (g_n). Limit functions are continuous and monotone on a compact set, so the inequality holds uniformly already for some finite n.

Sketch of proof

- Define

$$f_n(t) := \mathbf{P}(X_1 + \cdots + X_n \geqslant nt)$$

$$g_n(t) := \mathbf{P}(Y_1 + \cdots + Y_n \geqslant nt)$$

We want to show that $f_n \leqslant g_n$ for n large enough.

- Cramér's theorem gives $\lim \frac{1}{n} \log f_n$ on $[\mathbf{E}X, \max X]$, and also $\lim \frac{1}{n} \log(1 - f_n)$ on $[\min X, \mathbf{E}X]$; and similarly for g_n.
- Because all inequalities were assumed to be strict, we have also a strict inequality between the limits for (f_n) and (g_n). Limit functions are continuous and monotone on a compact set, so the inequality holds uniformly already for some finite n.

Sketch of proof

- Define

$$f_n(t) := \mathbf{P}(X_1 + \cdots + X_n \geqslant nt)$$

$$g_n(t) := \mathbf{P}(Y_1 + \cdots + Y_n \geqslant nt)$$

We want to show that $f_n \leqslant g_n$ for n large enough.

- Cramér's theorem gives $\lim \frac{1}{n} \log f_n$ on $[\mathbf{E}X, \max X]$, and also $\lim \frac{1}{n} \log(1 - f_n)$ on $[\min X, \mathbf{E}X]$; and similarly for g_n.
- Because all inequalities were assumed to be strict, we have also a strict inequality between the limits for (f_n) and (g_n). Limit functions are continuous and monotone on a compact set, so the inequality holds uniformly already for some finite n.

Sketch of proof

- Define

$$f_n(t) := \mathbf{P}(X_1 + \cdots + X_n \geqslant nt)$$
$$g_n(t) := \mathbf{P}(Y_1 + \cdots + Y_n \geqslant nt)$$

  We want to show that $f_n \leqslant g_n$ for n large enough.
- Cramér's theorem gives $\lim \frac{1}{n} \log f_n$ on $[\mathbf{E}X, \max X]$, and also $\lim \frac{1}{n} \log(1 - f_n)$ on $[\min X, \mathbf{E}X]$; and similarly for g_n.
- Because all inequalities were assumed to be strict, we have also a strict inequality between the limits for (f_n) and (g_n). Limit functions are continuous and monotone on a compact set, so the inequality holds uniformly already for some finite n.

Reformulation using majorization

We can reformulate everything in the language of majorization

Corollary

Let $x \in P_{d_x}$ and $y \in P_{d_y}$ with nonzero coordinates. Assume that

1. $N_p(x) < N_p(y)$ *for* $1 < p < +\infty$.
2. $N_p(x) > N_p(y)$ *for* $-\infty < p < 1$.
3. $H(x) < H(y)$ *(where* $H(x) = \sum x_i \log x_i$*).*
4. $x_{\max} < y_{\max}$.
5. $x_{\min} < y_{\min}$.

Then there exists an integer n such that $x^{\otimes n} \prec y^{\otimes n}$ *(i.e.* $x \in M_{d_x}(y)$*).*

From this corollary one can prove our two theorems : replace x by, respectively

$$(x_1 - \frac{\varepsilon}{d}, \ldots, x_d - \frac{\varepsilon}{d}, \varepsilon) \quad \text{or} \quad (x_1 - \frac{\varepsilon}{d}, \ldots, x_d - \frac{\varepsilon}{d}, \frac{\varepsilon}{k}, \ldots, \frac{\varepsilon}{k}),$$

so that condition 2 get satisfied.

Reformulation using majorization

We can reformulate everything in the language of majorization

Corollary

Let $x \in P_{d_x}$ and $y \in P_{d_y}$ with nonzero coordinates. Assume that

1. $N_p(x) < N_p(y)$ *for* $1 < p < +\infty$.
2. $N_p(x) > N_p(y)$ *for* $-\infty < p < 1$.
3. $H(x) < H(y)$ *(where* $H(x) = \sum x_i \log x_i$*).*
4. $x_{\max} < y_{\max}$.
5. $x_{\min} < y_{\min}$.

Then there exists an integer n such that $x^{\otimes n} \prec y^{\otimes n}$ *(i.e.* $x \in M_{d_x}(y)$*).*

From this corollary one can prove our two theorems : replace x by, respectively

$$(x_1 - \frac{\varepsilon}{d}, \dots, x_d - \frac{\varepsilon}{d}, \varepsilon) \quad \text{or} \quad (x_1 - \frac{\varepsilon}{d}, \dots, x_d - \frac{\varepsilon}{d}, \frac{\varepsilon}{k}, \dots, \frac{\varepsilon}{k}),$$

so that condition 2 get satisfied.

Recall the main conjecture :

Conjecture (Nielsen)

Let $y \in P_d, y_i > 0$. Then $\overline{C_d(y)} = \overline{M_d(y)}$ = the set of $x \in P_d$ so that

(A) $N_p(x) \leqslant N_p(y)$ *for* $p \geqslant 1$.

(B) $N_p(x) \geqslant N_p(y)$ *for* $0 \leqslant p \leqslant 1$.

(C) $N_p(x) \leqslant N_p(y)$ *for* $p \leqslant 0$

$$C_d(y) = \{x \in P_d \text{ s.t. } \exists k \in \mathbf{N}, \exists z \in P_k, x \otimes z \prec y \otimes z\}$$

$$M_d(y) = \{x \in P_d \text{ s.t. } \exists n \in \mathbf{N}, x^{\otimes n} \prec y^{\otimes n}\}$$

$$N_p(x) = \sum_{i=1}^{d} x_i^p.$$

Update: this conjecture has been proved by S. Turgut for $C_d(y)$ few days after I gave this talk. Cf arxiv:0707.0444