Correlation inequalities

Preliminary: log-concave functions

Definitions: • A function $f : \mathrm{dom}(f) \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is log-concave if $f(x) \geq 0 \ \forall x \in \mathrm{dom}(f)$ and $\log f$ is concave, i.e.:

$$\log f(\theta x + (1-\theta) y) \geq \theta \log(f(x)) + (1-\theta) \log f(y) \quad \left(\begin{array}{c} \forall x, y \in \mathrm{dom}(f) \\ 0 \leq \theta \leq 1 \end{array} \right)$$

equivalently: $$f(\theta x + (1-\theta) y) \geq f(x)^{\theta} \cdot f(y)^{1-\theta}$$

• A function f is quasi-concave if its superlevel sets $\{x \in \mathrm{dom}(f) : f(x) \geq t\}$ are convex $\forall t \geq 0$

Proposition: f is concave $\Rightarrow$ f is log-concave $\Rightarrow$ f is quasi-concave

Likewise, let us define:

- f is <u>log-convex</u> if $f > 0$ on $\mathrm{dom}(f)$ & $\log f$ is convex
- f is <u>quasi-convex</u> if all its sublevel sets are convex

<u>Proposition</u>: log-convex $\Rightarrow$ convex $\Rightarrow$ quasi-convex

<u>Examples of log-concave functions</u>:

- $C \subset \mathbb{R}^n$ convex : $\mathbb{1}_C$ is log-concave
- $A \mapsto \det A$ is log-concave on S^n_{++}
- The joint pdf of a Gaussian vector on $\mathbb{R}^n$ is log-concave

  ($p(x) \doteq \exp(-x^T A x)$ where $A \succeq 0$)

  as many other joint pdfs (unimodal, e.g.)

Operations preserving log-concavity

Multiplication: if f, g are $\begin{cases} \text{log-concave,} \\ \text{log-convex} \end{cases}$ then $f \cdot g$ is $\begin{cases} \text{log-concave} \\ \text{log-convex} \end{cases}$

Addition: if f, g are log-convex, then $f+g$ is log-convex

$(F = \log f,\ G = \log f$: then $\log(f+g) = \log(e^F + e^G)$ is convex$)$

not true for log-concave: take e.g. $f(x) = e^{-\|x\|^2}$ and $g(x) = e^{-\|x-x_0\|^2}$

Integration: if $f : \mathbb{R}^{n \times m} \to \mathbb{R}$ is log-concave,

(pf next page) then $g(x) = \int_{\mathbb{R}^m} f(x,y)\, dy$ is log-concave

Convolution: if f, g are log-concave,

(follows from previous point) then $h(x) = \int_{\mathbb{R}^n} f(y)\, g(x-y)\, dy$ is log-concave

Proof that $g(x) = \int_{\mathbb{R}^m} f(x,y)\,dy$ is log-concave:

This proof uses Prékopa-Leindler's inequality:

To be checked: $g((1-\lambda)x_1 + \lambda x_2) \geq g(x_1)^{1-\lambda} g(x_2)^{\lambda}$

denote $f_x(y) = f(x,y)$; the above inequality may be rewritten as

$$\int_{\mathbb{R}^m} f_{(1-\lambda)x_1+\lambda x_2}(y)\,dy \geq \left(\int_{\mathbb{R}^m} f_{x_1}(y)\,dy\right)^{1-\lambda} \cdot \left(\int f_{x_2}(y)\,dy\right)^{\lambda}$$

To apply PL inequality, we need to show that

$$f_{(1-\lambda)x_1+\lambda x_2}((1-\lambda)y_1 + \lambda y_2) \geq f_{x_1}(y_1)^{1-\lambda} \cdot f_{x_2}(y_2)^{\lambda} \quad \forall y_1, y_2 \in \mathbb{R}^m$$

But this follows directly from the log-concavity of $f: \mathbb{R}^{n\times m} \to \mathbb{R}$. #

More examples

- if X, Y are two indep. r.v.'s with log-concave pdfs then $X+Y$ has a log-concave pdf (follows from convolution)
- if f is a log-concave pdf on $\mathbb{R}$, then its cdf
$$F(x) = \int_{-\infty}^{x} f(y)\,dy = \int_{\mathbb{R}} \underbrace{1_{\{y \leq x\}}}_{\text{log-concave!}} f(y)\,dy \text{ is log-concave}$$
similarly for a multidimensional cdf
- if $C \subset \mathbb{R}^n$ is convex & X is a random vector with log-concave pdf, then $f(x) = \mathbb{P}(X \in C + x)$ is log-concave
- $f(b) = \mathrm{vol}(P_b)$ where $P_b = \{x \in \mathbb{R}^n : Ax \leq b\}$ is log-concave (vol(·) can also be replaced by arbitrary measure with log-concave pdf)

Correlation inequalities

For this entire section, assume X is a centered Gaussian random vector with joint pdf $p(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} \exp(-\frac{1}{2} x^T \Sigma^{-1} x)$ (for some $\Sigma \succ 0$)

Gaussian correlation conjecture [finally proven in 2014 by Royen]

Let $A, B \subset \mathbb{R}^n$ be convex and symmetric (i.e. $A = -A$, $B = -B$)

Then $\mathbb{P}(X \in A \cap B) \geq \mathbb{P}(X \in A) \cdot \mathbb{P}(X \in B)$ (i.e. positive correlation)

<u>Note</u>: When $n=1$, this is true for <u>any</u> pdf p: indeed, $A = [-a, a]$ and $B = [-b, b]$ in this case, so either $A \cap B = A$ or $A \cap B = B$

Say $A \cap B = A$: then $\mathbb{P}(X \in A \cap B) = \mathbb{P}(X \in A) \geq \mathbb{P}(X \in A) \cdot \underbrace{\mathbb{P}(X \in B)}_{\leq 1}$

Consequence (functional formulation of the correlation inequality)

Let $f, g : \mathbb{R}^n \to \mathbb{R}_+$ be quasi-concave & symmetric functions

Then $\mathbb{E}(f(X)g(X)) \geq \mathbb{E}(f(X)) \cdot \mathbb{E}(g(X))$

Proof:

convex, as f is quasi-concave

$$f(x) = \int_0^{f(x)} dt = \int_0^{\infty} \mathbb{1}_{\{f(x) \geq t\}} \, dt \quad , \text{ same for } g$$

So $\mathbb{E}(f(X)g(X)) = \int_{\mathbb{R}^n} f(x) g(x) p(x) \, dx$

$$\overset{\text{Fubini}}{=} \int_0^{\infty} \int_0^{\infty} \Big(\underbrace{\int_{\mathbb{R}^n} \mathbb{1}_{\{f(x) \geq t\}} \mathbb{1}_{\{g(x) \geq s\}} p(x) \, dx}_{= \mathbb{P}(\{f(X) \geq t\} \cap \{g(X) \geq t\}) \underset{\text{"conjecture"}}{\geq} \mathbb{P}(\{f(X) \geq t\}) \cdot \mathbb{P}(\{g(X) \geq t\})} \Big) \, ds \, dt$$

$$\geq \mathbb{E}(f(X)) \cdot \mathbb{E}(g(X))$$

#

Theorem (Khatri, Šidák)

Let A be a symmetric convex set and S be a symmetric slab. Then $\mathbb{P}(X \in A \cap S) \geq \mathbb{P}(X \in A) \cdot \mathbb{P}(X \in S)$

Proof

Without loss of generality, assume $S = \{x \in \mathbb{R}^n : |x_1| \leq 1\}$

Then $\mathbb{P}(X \in A \cap S) = \mathbb{P}((X_1 \ldots X_n) \in A, |X_1| \leq 1)$

$$= \mathbb{E}\left(f(X_1) \cdot 1_{[-1,+1]}(X_1)\right)$$

where $f(x_1) = \mathbb{E}(1_A(x_1, X_2 \ldots X_n))$

$$= \int_{\mathbb{R}^{n-1}} \underbrace{1_A(x_1, x_2, \ldots, x_n)}_{\text{log-concave}} \cdot \underbrace{p(x_2 \ldots x_n)}_{\text{log-concave marginal}} \, dx_2 \ldots dx_n \qquad \text{log-concave}$$

$$\text{So } \mathbb{P}(X \in A \cap S) \underset{\text{1-dim corr. ineq.}}{\geq} \mathbb{E}(f(X_1)) \cdot \mathbb{P}(|X_1| \leq 1) = \mathbb{P}(X_1 \in A) \cdot \mathbb{P}(X_1 \in S)$$

#

Particular case: $\mathbb{P}\left(\max_{1\le i\le n} |X_i| \le t\right) \ge \mathbb{P}(|X_1|\le t)\cdot \mathbb{P}\left(\max_{1<i\le n} |X_i| \le t\right)$

Remark: The Gaussian correlation conjecture is equivalent to

$$\mathbb{P}(|X_1|\le t_1 \dots |X_n|\le t_n) \ge \mathbb{P}(|X_1|\le t_1 \dots |X_k|\le t_k)\cdot \mathbb{P}(|X_{k+1}|\le t_{k+1} \dots |X_n|\le t_n)$$
$\forall 1\le k<n$

(using the fact that a symmetric convex set can be written as a countable intersection of symm. strips)

Anderson's theorem (v1)

Let C be a symmetric convex set

Then $\mathbb{P}(X\in C+x) \le \mathbb{P}(X\in C) \quad \forall x\in\mathbb{R}^n$

Proof:

$$\mathbb{P}(X\in C) \underbrace{\ge}_{\text{by log-concavity}} \mathbb{P}(X\in C+x)^{1/2}\cdot \mathbb{P}(X\in C-x)^{1/2} \underbrace{=}_{\text{by symmetry of both } X \,\&\, C} \mathbb{P}(X\in C+x) \quad \#$$

Anderson's theorem (v2)

Let K be a symmetric convex body in $\mathbb{R}^n$

$f: \mathbb{R}^n \to \mathbb{R}_+$ be symmetric, quasi-concave, integrable

Then $\int_K f(x+cy)\,dx \geq \int_K f(x+y)\,dx \quad \forall c \in [0,1]$ & $y \in \mathbb{R}^n$

(Intuition: f = hill centered at 0)

Proof:

- Consider first $f = 1_L$, where L is a symmetric convex body:

1_L is log-concave, so $g(y) = \int_K f(x+y)\,dx = \int_{\mathbb{R}^n} 1_L(x+y)\, 1_K(x)\,dx$ is also log-concave. Since K, L are symmetric, g is symmetric.

Hence: $\cdot\; g(0) \geq g(y)^{1/2} g(-y)^{1/2} = g(y) \quad \forall y \in \mathbb{R}^n$

$\cdot\; g(cy) \geq g(0)^{1-c} g(y)^{c} \geq g(y)$ ✓

- Next, use the fact that $f(x) = \int_0^\infty 1_{\underbrace{\{f(x)\geq t\}}_{=L_t \text{ convex, as } f \text{ is quasi-concave}}} \, dt$:

By Fubini's theorem, we have:

$$\int_K f(x+cy)\,dx = \int_0^\infty \int_K 1_{L_t}(x+cy)\,dx\,dt$$

$$\underset{\text{(by special case)}}{\geq} \int_0^\infty \int_K 1_{L_t}(x+y)\,dx\,dt = \int_K f(x+y)\,dx$$

#

Corollary

Let X be a random vector with pdf f (symmetric & quasi-concave)

Y be another random vector, independent of X, with pdf g

Let K be a symmetric convex body. Then $\mathbb{P}(X+Y\in K) \leq \mathbb{P}(X\in K)$

Proof: $\mathbb{P}(X+Y\in K) = \int_K \left(\int_{\mathbb{R}^n} f(x-y)g(y)\,dy\right)dx \overset{\text{Fubini}}{=} \int_{\mathbb{R}^n}\left(\int_K f(x-y)\,dx\right)g(y)\,dy$

$$\underset{\text{Anderson}}{\leq} \int_{\mathbb{R}^n}\left(\int_K f(x)\,dx\right)g(y)\,dy \underset{\text{Fubini}}{=} \int_K f(x)\,dx \cdot \underbrace{\int_{\mathbb{R}^n} g(y)\,dy}_{=1} = \mathbb{P}(X\in K)$$

#

Appendix: Proof of the Gaussian correlation Conjecture

(main ideas)

Ref: Latała & Matlak 2015, following Royen 2014

$\boxed{\text{To be proven:}}$ $\mathbb{P}(|X_1|\le t_1 \dots |X_n| \le t_n) \ge \mathbb{P}(|X_1|\le t_1 \dots |X_k|\le t_k)\cdot \mathbb{P}(|X_{k+1}|\le t_{k+1} \dots |X_n|\le t_n)$

$\forall t_1 \dots t_n,\ \forall 1\le k<n$

Fix k: Let C be the covariance matrix of X: C may be written as $C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$ where C_{11} is $k\times k$, C_{22} is $(n-k)\times(n-k)$

Let $C(\tau) = \begin{pmatrix} C_{11} & \tau C_{12} \\ \tau C_{21} & C_{22} \end{pmatrix}$ with $0\le \tau \le 1$ and $X(\tau)$ be the centered Gaussian random vector with covariance $C(\tau)$.

The above inequality may be rewritten as:

$$\mathbb{P}(|X_1(1)|\le t_1, \dots, |X_n(1)| \le t_n) \ge \mathbb{P}(|X_1(0)|\le t_1, \dots, |X_n(0)|\le t_n)$$

Defining $Z_i(\tau) = \frac{1}{2} X_i(\tau)^2$ and $s_i = \frac{1}{2} t_i^2$, the above inequality becomes:

$$\mathbb{P}(Z_1(1) \leq s_1, \dots, Z_n(1) \leq s_n) \geq \mathbb{P}(Z_1(0) \leq s_1, \dots, Z_n(0) \leq s_n)$$

This will be proven if one can show that

$$\tau \longmapsto \mathbb{P}(Z_1(\tau) \leq s_1, \dots, Z_n(\tau) \leq s_n) \quad \text{is non-decreasing on } [0,1]$$

Let $f(x,\tau)$ denote the joint pdf of $Z(\tau)$:

$$\frac{\partial}{\partial \tau} \mathbb{P}(Z_1(\tau) \leq s_1, \dots, Z_n(\tau) \leq s_n) = \frac{\partial}{\partial \tau} \int_K f(x,\tau)\, dx \geq 0 \; ?$$

$$\left(\text{where } K = \{x \in \mathbb{R}_+^n : 0 \leq x_i \leq s_i \;\; \forall i\}\right)$$

To solve this question, the idea is to compute the Laplace transform of $f(x,\tau)$.

<u>Laplace transform</u>: fix $\lambda \in \mathbb{R}_+^n$:

$$\int_{\mathbb{R}_+^n} \exp(-\lambda^T x) f(x,\tau)\, dx = \mathbb{E}\Big(-\frac{1}{2}\sum_{j=1}^n \lambda_j X_j(\tau)^2\Big)$$
$$= \det(I+\Lambda C(\tau))^{-\frac{1}{2}} \quad \text{where} \quad \Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

<u>Cauchy-Binet formula</u>: $\det(I+A) = 1 + \sum_{\emptyset \neq J \subset \{1..n\}} \det(A_J)$

so $\det(I+\Lambda C(\tau)) = 1 + \sum_{\emptyset \neq J \subset \{1..n\}} \Big(\prod_{j \in J} \lambda_j\Big) \det(C(\tau)_J)$

For every $\emptyset \neq J \subset \{1..n\}$, $J = J_1 \cup J_2$ where $J_1 \subset \{1..k\}$, $J_2 \subset \{k+1..n\}$

and $C(\tau)_J = \begin{pmatrix} C_{J_1} & \tau C_{J_1 J_2} \\ \tau C_{J_2 J_1} & C_{J_2} \end{pmatrix}$. If $J_1 = \emptyset$ or $J_2 = \emptyset$, then

$\det(C(\tau)_J) = \det(C_J)$ is indep. of τ. If not, then (next page)

$$\det\left(C(\tau)_J\right) = \det\left(C_{J_1}\right)\det\left(C_{J_2}\right)\det\Big(I - \tau^2 \underbrace{C_{J_1}^{-1} C_{J_1 J_2} C_{J_2}^{-1} C_{J_2 J_1}}\Big)$$

<u>fact</u>: the eigenvalues of this matrix $\in [0,1]$

(because C is positive definite)

So $a_J(\tau) = -\frac{\partial}{\partial\tau}\det\left(C(\tau)_J\right) \geq 0$

Therefore:

$$\frac{\partial}{\partial\tau}\det\left(I + \Lambda C(\tau)\right)^{-1/2} = -\frac{1}{2}\det\left(I+\Lambda C(\tau)\right)^{-3/2} \sum_{\phi \neq J \subset \{1..n\}} \prod_{j\in J} \lambda_j \underbrace{\det\left(C(\tau)_J\right)}_{= -a_J(\tau)}$$

and

$$\frac{\partial}{\partial\tau}\int_{\mathbb{R}_n^+} \exp(-\lambda^T x) f(x,\tau)\,dx = \frac{\partial}{\partial\tau}\det\left(I+\Lambda C(\tau)\right)^{-1/2}$$

$$= \frac{1}{2}\det\left(I+\Lambda C(\tau)\right)^{-3/2} \sum_{\phi\neq J\subset\{1..n\}} \Big(\prod_{j\in J}\lambda_j\Big)\, a_J(\tau)$$

Back to $f(x,\tau)$:

The last steps are more technical (and will be mostly skipped):

For a given $\emptyset \subset J \subset \{1..n\}$, compute $g_J(x)$ s.t.

$$\det(I + \Lambda C(\tau))^{-3/2} \prod_{j \in J} \lambda_j = \int_{\mathbb{R}} \exp(-\lambda^T x)\, g_J(x)\, dx$$

This allows to make the identification (provided some exchanges of integrals and derivatives, to be justified):

$$\frac{\partial}{\partial \tau} f(x,\tau) = \sum_{\emptyset \neq J \subset \{1..n\}} \frac{1}{2} a_J(\tau) \cdot g_J(x)$$

and one can finally argue that $\int_K g_J(x)\, dx \geq 0$,

implying that $\frac{\partial}{\partial t} \int_K f(x,\tau)\, dx \geq 0$ (exchange $\frac{\partial}{\partial \tau}$ & $\int$ again) "#"