The Covariance Representation on the Sphere

Sergey Bobkov
University of Minnesota

joint work with Devraj Duggal

High Dimensional Probability X
11.06.2023 - 16.06.2023
The Banach Center, Bedlewo, Poland

One dimensional covariance identities

Höffding (1940) Given random variables X and Y with finite second moments, we have

$$\begin{aligned}\mathrm{cov}(X,Y) &= \mathbb{E}XY - \mathbb{E}X\,\mathbb{E}Y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)\,dx\,dy,\end{aligned}$$

where the Höffding kernel is defined by

$$H(x,y) = \mathbb{P}\{X \le x, Y \le y\} - \mathbb{P}\{X \le x\}\,\mathbb{P}\{Y \le y\}.$$

Generalized Höffding formula (Mardia, Sen, Cuadras ...)

$$\mathrm{cov}(u(X), v(Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u'(x)v'(y)\,H(x,y)\,dx\,dy.$$

The case $X = Y$ with distribution μ

Theorem. Given a probability measure μ on the real line,

$$\mathrm{cov}_\mu(u,v) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} u'(x)v'(y)\, d\lambda(x,y)$$

for a unique positive, locally finite measure λ, with density

$$H(x,y) = \frac{d\lambda(x,y)}{dx\, dy} = F(x \wedge y)\,(1 - F(x \vee y)),$$

where

$$F(x) = \mathbb{P}\{X \le x\} = \mu((-\infty, x]),$$

$x \wedge y = \min(x,y)$, $x \vee y = \max(x,y)$.

Definition: λ is called Höffding's measure.

Assumptions:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |u'(x)|\,|u'(y)|\, d\lambda(x,y) < \infty,$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |v'(x)|\,|v'(y)|\, d\lambda(x,y) < \infty.$$

Total mass:

$$\lambda(\mathbb{R}\times\mathbb{R}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(x,y)\, dx\, dy = \mathrm{Var}(X).$$

Marginals of Höffding measures

Let μ have finite second moment with $a = \mathbb{E}X$. The measure λ has equal unimodal marginals: For Borel sets $A \subset \mathbb{R}$,

$$\Lambda(A) = \lambda(A \times \mathbb{R}) = \int_A^\infty \int_{-\infty}^\infty H(x,y)\, dx\, dy$$

with density

$$h(x) = \frac{d\Lambda(x)}{dx} = \int_x^\infty (y-a)\, dF(y).$$

Corollary.

$$\operatorname{cov}(X, u(X)) = \int_{-\infty}^\infty u'(x)\, h(x)\, dx.$$

If μ is supported on an interval Δ and has there a positive density p, this formula may be rewritten as Stein's identity

$$\operatorname{cov}(X, u(X)) = \mathbb{E}\, \tau(X) u'(X),$$

where

$$\tau(x) = \frac{h(x)}{p(x)} = \frac{1}{p(x)} \int_x^\infty (y-a)\, p(y)\, dy, \quad x \in \Delta,$$

is the Stein kernel.

Characterization of Gaussian measures via marginals of Höffding measures

Claim. $\Lambda = \sigma^2 \mu$ for some $\sigma \geq 0$ iff $\mu \sim N(a, \sigma^2)$.

Proof. If $a = 0$, the Fourier-Stieltjes transform of Λ is

$$\hat{\Lambda}(t) = \int_{-\infty}^{\infty} e^{itx}\, h(x)\, dx = -\frac{f'(t)}{t} \quad (t \neq 0),$$

where f is the characteristic function of X. Hence $\Lambda = \sigma^2 \mu$ iff

$$f'(t) = -\sigma^2 t f(t) \quad \text{for all} \;\; t \in \mathbb{R},$$

that is, $\mu \sim N(0, \sigma^2)$.

Standard normal case. If $a = 0$, $\sigma = 1$, we have Stein's identity with $\tau = 1$,

$$\operatorname{cov}(X, u(X)) = \mathbb{E}\, u'(X).$$

Multidimensional extension

Let μ be the standard Gaussian measure on $\mathbb{R}^n$ with density

$$\varphi(x) = \frac{1}{(2\pi)^{n/2}}\, e^{-|x|^2/2}, \quad x \in \mathbb{R}^n.$$

Theorem. For a (unique) probability measure λ on $\mathbb{R}^n \times \mathbb{R}^n$,

$$\mathrm{cov}_\mu(u,v) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle \nabla u(x), \nabla v(y)\rangle \, d\lambda(x,y).$$

Necessarily $\Lambda = \mu$.

History: Via Ornstein-Uhlenbeck semigroups (Ledoux 1995):

$$\mathrm{cov}_\mu(u,v) = \int_0^\infty \mathbb{E}_\mu \langle \nabla u, \nabla P_t v\rangle \, dt,$$

$$P_t v(x) = \int_{\mathbb{R}^n} v\big(e^{-t}x + \sqrt{1-e^{-2t}}\, y\big)\, d\mu(y).$$

Interpolation (Houdré, Pérez-Abreu 1995, H-PA-Surgailis 1998), B-Houdré-Götze (2001)

Uniqueness of λ: Apply the identity to

$$u(x) = e^{i\langle a,x\rangle}, \quad v(y) = e^{i\langle b,y\rangle} \quad (a,b \in \mathbb{R}^n).$$

Characterization of Gaussian measures via covariance identity

Theorem. Given a probability measure μ on $\mathbb{R}^n$, suppose that

$$\mathrm{cov}_\mu(u,v) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle \nabla u(x), \nabla v(y)\rangle \, d\lambda(x,y)$$

for some finite measure λ on $\mathbb{R}^n \times \mathbb{R}^n$ in the class of all bounded smooth u, v on $\mathbb{R}^n$ with bounded partial derivatives. Then μ is Gaussian with covariance matrix $\sigma^2 I_n$.

Proof is based on the Darmois-Skitovitch theorem.

Gaussian concentration

Let μ be the standard Gaussian measure on $\mathbb{R}^n$.
Let u be a Lipschitz function on $\mathbb{R}^n$ with $\|u\|_{\mathrm{Lip}} \leq 1$ and mean $m = \mathbb{E}u$.

Log-Sobolev: For all $r > 0$,

$$\mu\{|u - m| \geq r\} \leq 2e^{-r^2/2}.$$

Isoperimetry:

$$\mu\{|u - m| \geq r\} \leq 4\,\frac{e^{-r^2/2}}{r}.$$

Corollary (from the covariance identity):

$$\mu\{|u - m| \geq r\} \leq \mathbb{E}\,|u - m|\,\frac{e^{-r^2/2}}{r}.$$

Example: $u(x) = \max(x_1, \ldots, x_n)$.

Proof

Let $m = 0$ and $v = T(u)$ with non-decreasing T. Then

$$\mathrm{cov}_\mu(u, T(u)) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle \nabla u(x), \nabla u(y)\rangle\, T'(u(y))\, d\lambda(x,y)$$

implies

$$\mathbb{E}\, uT(u) \le \mathbb{E}\, T'(u).$$

Choose

$$T(x) = \min\{(x-r)^+, \varepsilon\}, \quad x, \varepsilon > 0.$$

Let $\varepsilon \to 0$, we get

$$V(r) \equiv \int_r^\infty xp(x)\, dx \le p(r),$$

where p is density of u under μ. Equivalently, $V(r)\, e^{r^2/2}$ is non-increasing in $r > 0$, so,

$$V(r)\, e^{r^2/2} \le V(0) = \mathbb{E}\, u^+.$$

But $V(r) \ge r\mu\{u \ge r\}$, and we arrive at

$$\mu\{u \ge r\} \le \mathbb{E}\, u^+\, \frac{e^{-r^2/2}}{r}.$$

Spherical derivatives

Let $\mu = \sigma_{n-1}$ be a uniform distribution on the unit sphere

$$S^{n-1} = \{x \in \mathbb{R}^n : |x| = 1\}, \quad n \geq 2.$$

If f is smooth on the sphere, then $w = \nabla_S f(x)$ at the point $x \in S^{n-1}$ is the shortest vector such that

$$f(y) = f(x) + \langle w, y - x \rangle + o(|y - x|)$$

for $y \to x$, $y \in S^{n-1}$. If f is smooth in a neighborhood of the sphere,

$$\nabla_S f(x) = P_{x^\perp} \nabla f(x) = \nabla f(x) - \langle \nabla f(x), x \rangle\, x.$$

Spherical covariance identity

Theorem. On $S^{n-1} \times S^{n-1}$ there exist a probability measure ν with marginals σ_{n-1} and a constant $c_n > 0$ such that

$$\mathrm{cov}_{\sigma_{n-1}}(f,g) = c_n \int_{S^{n-1}} \int_{S^{n-1}} \langle \nabla_S f(x), \nabla_S g(y) \rangle \, d\nu(x,y)$$

for all smooth f, g on S^{n-1}. Moreover, ν has density with respect to $\sigma_{n-1} \otimes \sigma_{n-1}$ of the form $\psi(\langle x, y \rangle)$. We also have

$$\frac{1}{n-1} < c_n < \frac{1}{n-2} \quad (n \geq 3).$$

Notes. 1) Since $\langle x, y \rangle = 1 - \frac{1}{2} |x-y|^2$, the density depends on the distance $|x - y|$.
2) In general no uniqueness. If $S \subset S^{n-1}$ is a circle, and ν' is supported on the set $A = \{(x,y) \in S \times S : \langle x, y \rangle = 0\}$, then

$$\langle \nabla_S f(x), \nabla_S g(y) \rangle = 0 \quad \text{for} \quad (x,y) \in A.$$

Hence, a similar covariance identity is also true for $c_n \nu + \nu'$.

Density in Gaussian covariance representation

The probability measure λ in

$$\mathrm{cov}_\mu(u,v) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \langle \nabla u(x), \nabla v(y)\rangle \, d\lambda(x,y)$$

may be described as

$$\lambda = \int_0^1 \mathcal{L}\big(X,\, tX + sZ\big)\, dt, \quad s = \sqrt{1-t^2}.$$

Here X and Z are independent standard normal random vectors in $\mathbb{R}^n$. It has density on $\mathbb{R}^n \times \mathbb{R}^n$

$$p(x,y) = \frac{1}{(2\pi)^n} \int_0^1 s^{-n} \exp\left[-\frac{|x|^2 + |y|^2 - 2t\,\langle x, y\rangle}{2s^2}\right] dt.$$

Construction of spherical mixing density

With smooth $f, g : S^{n-1} \to \mathbb{R}$, we associate the functions

$$\begin{aligned} u(x) &= f(r^{-1}x) = f(\theta), \qquad r = |x|, \quad \theta = r^{-1}x, \\ v(x) &= g(r^{-1}x) = g(\theta), \end{aligned}$$

on $\mathbb{R}^n \setminus \{0\}$, with their gradients

$$\nabla u(x) = \frac{1}{r}\,\nabla_S f(\theta), \qquad \nabla v(y) = \frac{1}{r'}\,\nabla_S g(\theta')$$

Here r and θ are independent under μ, and $\mathcal{L}_\mu(u) = \mathcal{L}_{\sigma_{n-1}}(f)$. Hence

$$\mathrm{cov}_{\sigma_{n-1}}(f, g) = \mathrm{cov}_\mu(u, v).$$

Applying the Gaussian covariance identity and integrating in polar coordinates, we represent $\mathrm{cov}_{\sigma_{n-1}}(f, g)$ as

$$\int\!\!\int \langle \nabla_S f(\theta), \nabla_S g(\theta') \rangle \, \psi(\langle \theta, \theta' \rangle) \, d\sigma_{n-1}(\theta) \, d\sigma_{n-1}(\theta'),$$

$$\begin{aligned} \psi(\alpha) \ = \ \frac{1}{2^{n-2}\,\Gamma(\frac{n}{2})^2} \int_0^1 s^{n-2} \bigg[\int_0^\infty \int_0^\infty \\ \exp\Big[-\frac{r^2 + r'^2 - 2rr't\alpha}{2} \Big] (rr')^{n-2} \, dr \, dr' \bigg] dt. \end{aligned}$$

Marginals. Estimation of constants c_n

Repeating integration in polar coordinates, we get

$$\begin{aligned} c_n &= \int\!\!\int \psi(\langle \theta, \theta' \rangle)\, d\sigma_{n-1}(\theta)\, d\sigma_{n-1}(\theta') \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} \frac{p(x,y)}{|x|\,|y|}\, dx\, dy \;=\; \mathbb{E} \int_0^1 \frac{1}{|X|\,|tX + sZ|}\, dt. \end{aligned}$$

By Cauchy's inequality, the latter expectation is smaller than

$$\left(\mathbb{E}\,\frac{1}{|X|^2}\right)^{1/2} \left(\mathbb{E}\,\frac{1}{|tX + sZ|^2}\right)^{1/2} = \mathbb{E}\,\frac{1}{|X|^2} \;=\; \frac{1}{n-2}.$$

Rotational invariance: The density of the marginal of ν is

$$q(\theta) = \int_{S^{n-1}} \psi(\langle \theta, \theta' \rangle)\, d\sigma_{n-1}(\theta').$$

This integral does not depend on θ and is therefore equal to c_n.

Behavior of ψ near the diagonal

Write $\psi = \psi_n$.

Claim: ψ_2 is bounded, while (within n-dependent factors)

$$\psi_3(\langle\theta, \theta'\rangle) \sim \log \frac{1}{|\theta - \theta'|} \quad (n = 3),$$

$$\psi_n(\langle\theta, \theta'\rangle) \sim \frac{1}{|\theta - \theta'|^{n-3}} \quad (n \geq 4).$$

Application to spherical concentration

Let f on S^{n-1} satisfy $\|f\|_{\mathrm{Lip}} \le 1$ and have σ_{n-1}-mean m.

Classical deviation inequality: For all $r > 0$,

$$\sigma_{n-1}\Big\{|f - m| \ge r\Big\} \le 2\, e^{-(n-1)r^2/2}.$$

Corollary. For all $r > 0$,

$$\sigma_{n-1}\Big\{|f - m| \ge r\Big\} \le \frac{1}{r}\, e^{-r^2/2c_n}\ \mathbb{E}_{\sigma_{n-1}}\, |f - m|.$$

Periodic covariance identities

For a probability measure μ on $[0,1)$, consider an identity

$$\mathrm{cov}_\mu(u,v) = \int_0^1 \int_0^1 u'(x)v'(y)\,d\lambda(x,y)$$

in the class of all 1-periodic smooth functions u and v on $\mathbb{R}$, assuming that λ is a signed measure.

Example: The Höffding measure $\lambda = \lambda_\mu$ whose marginal Λ_μ are however not multiples of μ.

Theorem. Subject to the constraint that the marginal distribution of λ is equal to $c\mu$ for a prescribed value $c \in \mathbb{R}$, the mixing measure λ exists, is unique and is given by

$$\begin{aligned} \lambda \ = \ & \lambda_\mu + (\sigma^2 - c)\, m \otimes m \\ & + c\,(\mu \otimes m + m \otimes \mu) - (\Lambda_\mu \otimes m + m \otimes \Lambda_\mu), \end{aligned}$$

where σ^2 is variance of μ and m is the uniform distribution on $(0,1)$.

Uniform distribution

Let μ be uniform on $(0,1)$.

Corollary. Subject to the constraint that the marginal distribution of λ is equal to $c\mu$, $c \in \mathbb{R}$, the measure λ has density

$$\frac{\lambda(x,y)}{dx\,dy} = D(|x-y|) + \Big(c - \frac{1}{24}\Big), \quad x,y \in (0,1),$$

where

$$D(h) = \frac{1}{8}\left[\,1 - 4h(1-h)\right], \quad 0 \le h \le 1.$$

Optimal choice $c = \frac{1}{24}$ leads to

$$\begin{aligned} \operatorname{cov}_\mu(u,v) &= \int_0^1 \int_0^1 u'(x)v'(y)\, D(|x-y|)\, dx\, dy \\ &= \frac{1}{24} \int_0^1 \int_0^1 u'(x)v'(y)\, d\nu(x,y) \end{aligned}$$

with a probability measure $d\nu(x,y) = 24\, D(|x-y|)\, dx\, dy$ on $(0,1) \times (0,1)$. It has μ as a marginal one.