Dimension-free variance bounds for polynomials

Dan Mikulincer

MIT

Joint work with Itay Glazer (Northwestern)

Wishart Tensors

Let $\{X_i\}_{i=1}^k$ be *i.i.d.* copies of an isotropic random vector $X \sim \mu$ in $\mathbb{R}^n$. And consider

$$W := \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left(X_i^{\otimes d} - \mathbb{E}\left[X_i^{\otimes d} \right] \right).$$

Keeping n and d fixed the W converges to a Gaussian vector.

What happens when we allow n (and d) to approach infinity?

Wishart Tensors

Let $\{X_i\}_{i=1}^k$ be *i.i.d.* copies of an isotropic random vector $X \sim \mu$ in $\mathbb{R}^n$. And consider

$$W := \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left(X_i^{\otimes d} - \mathbb{E}\left[X_i^{\otimes d} \right] \right).$$

Keeping n and d fixed the W converges to a Gaussian vector. What happens when we allow n (and d) to approach infinity?

Wishart Tensors

Some motivation to understand the asymptotic normality of W:

1. Empirical moment tensor estimation.
2. Related to random geometric graphs, when $d = 2$.

$$\mathbb{X}\mathbb{X}^T = \sum_{i=1}^{k} X_i \otimes X_i$$

   where $\mathbb{X}$ is a matrix with columns given by X_i.

3. CLT for neural networks, when $d \geq 2$. For fixed $y \in \mathbb{R}^n$,

$$\frac{1}{\sqrt{k}} \sum_{i=1}^{k} \langle X_i, y \rangle^p = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \langle X_i^{\otimes p}, y^{\otimes p} \rangle = \langle \frac{1}{\sqrt{k}} \sum_{\ell=1}^{k} X_i^{\otimes p}, y^{\otimes p} \rangle.$$

Wishart Tensors

Some motivation to understand the asymptotic normality of W:

1. Empirical moment tensor estimation.
2. Related to random geometric graphs, when $d = 2$.

$$\mathbb{X}\mathbb{X}^T = \sum_{i=1}^{k} X_i \otimes X_i$$

where $\mathbb{X}$ is a matrix with columns given by X_i.

3. CLT for neural networks, when $d \geq 2$. For fixed $y \in \mathbb{R}^n$,

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}\langle X_i, y\rangle^p = \frac{1}{\sqrt{k}}\sum_{i=1}^{k}\langle X_i^{\otimes p}, y^{\otimes p}\rangle = \langle\frac{1}{\sqrt{k}}\sum_{\ell=1}^{k} X_i^{\otimes p}, y^{\otimes p}\rangle.$$

Wishart Tensors

Some motivation to understand the asymptotic normality of W:

1. Empirical moment tensor estimation.
2. Related to random geometric graphs, when $d = 2$.

$$\mathbb{X}\mathbb{X}^T = \sum_{i=1}^{k} X_i \otimes X_i$$

   where $\mathbb{X}$ is a matrix with columns given by X_i.

3. CLT for neural networks, when $d \geq 2$. For fixed $y \in \mathbb{R}^n$,

$$\frac{1}{\sqrt{k}} \sum_{i=1}^{k} \langle X_i, y \rangle^p = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \langle X_i^{\otimes p}, y^{\otimes p} \rangle = \langle \frac{1}{\sqrt{k}} \sum_{\ell=1}^{k} X_i^{\otimes p}, y^{\otimes p} \rangle.$$

Known results

When $n^{2d-1} \ll k$, W is asymptotically normal.

- Bubeck, Ding, Eldan, Rácz 15' and Jiang, Li 15' - $d = 2$, standard Gaussian.
- Bubeck, Ganguly 15' - $d = 2$, log-concave product measures.
- Fang, Koike 20' - $d = 2$, product measures.
- Nourdin, Zheng 18'- $d \geq 2$, standard Gaussian.
- M. 20' - $d \geq 2$, unconditional strongly log-concave measures.
- M., Shenfeld 21' - $d \geq 2$, unconditional log-concave measures.

μ is log-concave if $-\log(\frac{d\mu}{dx})$ convex.
μ is unconditional if $\frac{d\mu}{dx}(x_1, \ldots, x_n) = \frac{d\mu}{dx}(\pm x_1, \ldots, \pm x_n)$.

Known results

When $n^{2d-1} \ll k$, W is asymptotically normal.

- Bubeck, Ding, Eldan, Rácz 15' and Jiang, Li 15' - $d = 2$, standard Gaussian.
- Bubeck, Ganguly 15' - $d = 2$, log-concave product measures.
- Fang, Koike 20' - $d = 2$, product measures.
- Nourdin, Zheng 18'- $d \geq 2$, standard Gaussian.
- M. 20' - $d \geq 2$, unconditional strongly log-concave measures.
- M., Shenfeld 21' - $d \geq 2$, unconditional log-concave measures.

μ is log-concave if $-\log(\frac{d\mu}{dx})$ convex.
μ is unconditional if $\frac{d\mu}{dx}(x_1, \ldots, x_n) = \frac{d\mu}{dx}(\pm x_1, \ldots, \pm x_n)$.

Known results

Important caveat:
Instead of considering the full tensor W, the results apply to its marginal on the subspace of principal (multi-linear) tensors:

$$\mathrm{span}\left\{e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k} | i_1 < i_2 < \cdots < i_k\right\}.$$

Reason: If $X = (X_1, \ldots, X_n)$ is unconditional, the covariance matrix on the principal subspace is diagonal:

$$\mathbb{E}\left[(X_{i_1} \cdot \cdots \cdot X_{i_k})(X_{j_1} \cdot \cdots \cdot X_{j_k})\right] = 0,$$

whenever $(i_1, \ldots, i_k) \neq (j_1, \ldots, j_k)$.

Known results

Important caveat:
Instead of considering the full tensor W, the results apply to its marginal on the subspace of principal (multi-linear) tensors:

$$\mathrm{span}\left\{e_{i_1} \otimes e_{i_2} \otimes \cdots \otimes e_{i_k} | i_1 < i_2 < \cdots < i_k\right\}.$$

Reason: If $X = (X_1, \ldots, X_n)$ is unconditional, the covariance matrix on the principal subspace is diagonal:

$$\mathbb{E}\left[\left(X_{i_1} \cdot \cdots \cdot X_{i_k}\right)\left(X_{j_1} \cdot \cdots \cdot X_{j_k}\right)\right] = 0,$$

whenever $(i_1, \ldots, i_k) \neq (j_1, \ldots, j_k)$.

From tensors to polynomials

Remark

To control convergence rate of the CLT, one needs to understand $\lambda_{\min}\left(\mathrm{Cov}(X^{\otimes d})\right)$ and $\lambda_{\max}\left(\mathrm{Cov}(X^{\otimes d})\right)$

To rephrase, let $f : \mathbb{R}^n \to \mathbb{R}$, be a degree d homogeneous polynomial $f(x) = \sum_I v_I x^I$, where

$$I \in [n]^d \text{ and } x^I = \prod_{i=1}^{d} x_{I_i}.$$

So, $\langle X^{\otimes d}, f\rangle = \sum_{|I|=d} v_I X^I = f(X)$, and

$$\lambda_{\min}\left(\mathrm{Cov}(X^{\otimes d})\right) = \inf_{f : \sum v_I^2 = 1} \mathrm{Var}(f(X)).$$

From tensors to polynomials

Remark

To control convergence rate of the CLT, one needs to understand $\lambda_{\min}\left(\mathrm{Cov}(X^{\otimes d})\right)$ and $\lambda_{\max}\left(\mathrm{Cov}(X^{\otimes d})\right)$

To rephrase, let $f : \mathbb{R}^n \to \mathbb{R}$, be a degree d homogeneous polynomial $f(x) = \sum_I v_I x^I$, where

$$I \in [n]^d \text{ and } x^I = \prod_{i=1}^{d} x_{I_i}.$$

So, $\langle X^{\otimes d}, f\rangle = \sum_{|I|=d} v_I X^I = f(X)$, and

$$\lambda_{\min}\left(\mathrm{Cov}(X^{\otimes d})\right) = \inf_{f:\sum v_I^2=1} \mathrm{Var}(f(X)).$$

From tensors to polynomials

Remark

To control convergence rate of the CLT, one needs to understand $\lambda_{\min}\left(\mathrm{Cov}(X^{\otimes d})\right)$ and $\lambda_{\max}\left(\mathrm{Cov}(X^{\otimes d})\right)$

To rephrase, let $f : \mathbb{R}^n \to \mathbb{R}$, be a degree d homogeneous polynomial $f(x) = \sum_I v_I x^I$, where

$$I \in [n]^d \text{ and } x^I = \prod_{i=1}^{d} x_{I_i}.$$

So, $\langle X^{\otimes d}, f \rangle = \sum_{|I|=d} v_I X^I = f(X)$, and

$$\lambda_{\min}\left(\mathrm{Cov}(X^{\otimes d})\right) = \inf_{f : \sum v_I^2 = 1} \mathrm{Var}(f(X)).$$

A first result - Gaussians

Lemma

Let G be a standard Gaussian in $\mathbb{R}^n$, and let $f(x) = \sum_I v_I x^I$ with $\sum v_I^2 = 1$. Then, $\mathrm{Var}(f(G)) \geq \frac{1}{d!}$.

Proof.

Gaussian integration by parts:

$$\mathrm{Var}(f(G)) = \sum_{m=1}^{\infty} \frac{\|\mathbb{E}\left[\nabla^m f(G)\right]\|^2}{m!} \geq \frac{\|\mathbb{E}\left[\nabla^d f(G)\right]\|^2}{d!}.$$

But,

$$\frac{d}{dx^I} x^J = I! \delta_{IJ} \implies \frac{d}{dx^I} f = I! v_I.$$

So,

$$\|\mathbb{E}\left[\nabla^d f(G)\right]\|^2 = \sum (I!)^2 v_I^2 \geq 1.$$

A first result - Gaussians

Lemma

Let G be a standard Gaussian in $\mathbb{R}^n$, and let $f(x) = \sum_I v_I x^I$ with $\sum v_I^2 = 1$. Then, $\mathrm{Var}(f(G)) \geq \frac{1}{d!}$.

Proof.

Gaussian integration by parts:

$$\mathrm{Var}(f(G)) = \sum_{m=1}^{\infty} \frac{\|\mathbb{E}\left[\nabla^m f(G)\right]\|^2}{m!} \geq \frac{\|\mathbb{E}\left[\nabla^d f(G)\right]\|^2}{d!}.$$

But,

$$\frac{d}{dx^I} x^J = I!\delta_{IJ} \implies \frac{d}{dx^I} f = I! v_I.$$

So,

$$\|\mathbb{E}\left[\nabla^d f(G)\right]\|^2 = \sum (I!)^2 v_I^2 \geq 1.$$

□

Main question

The previous proof is very Gaussian.

Question

Which isotropic random vectors satisfy,

$$\mathrm{Var}(f(X)) \geq C_d,$$

for any d-homogeneous polynomial with $\sum v_I^2 = 1$?

Specific cases of interest:

1. Product measures.
2. Log-concave measures.

Main question

The previous proof is very Gaussian.

Question

Which isotropic random vectors satisfy,

$$\mathrm{Var}(f(X)) \geq C_d,$$

for any d-homogeneous polynomial with $\sum v_I^2 = 1$?

Specific cases of interest:

1. Product measures.
2. Log-concave measures.

Related results - Carbery-Wright

The celebrated Carbery-Wright inequality connects between log-concave measures and level sets of polynomials

Lemma (Carbery-Wright's inequality)

Let X be a log-concave vector in $\mathbb{R}^m$, then for any polynomial f of degree d, $t \in \mathbb{R}$ and ε.

$$\mathbb{P}(|f(X) - t| < \varepsilon) \lesssim \left(\frac{\varepsilon}{\sqrt{\mathbb{E}\left[f(X)^2\right]}} \right)^{\frac{1}{d}}$$

Problems:

- Need to show that $\mathbb{E}[f^2(X)]$ is comparable to $\sum v_I^2$.
- The proof of Carbery-Wright proceeds by localization, which does not preserve coefficients.

Related results - Carbery-Wright

The celebrated Carbery-Wright inequality connects between log-concave measures and level sets of polynomials

Lemma (Carbery-Wright's inequality)

Let X be a log-concave vector in $\mathbb{R}^m$, then for any polynomial f of degree d, $t \in \mathbb{R}$ and ε.

$$\mathbb{P}(|f(X) - t| < \varepsilon) \lesssim \left(\frac{\varepsilon}{\sqrt{\mathbb{E}\left[f(X)^2\right]}} \right)^{\frac{1}{d}}$$

Problems:

- Need to show that $\mathbb{E}[f^2(X)]$ is comparable to $\sum v_I^2$.
- The proof of Carbery-Wright proceeds by localization, which does not preserve coefficients.

Related results - Carbery-Wright

A more fundamental problem is that Carbery-Wright is too general.

If X is uniform on $\sqrt{n}B_2^n$, and $f(x) = \frac{1}{\sqrt{n}}\|x\|^2$, an easy calculation shows,

$$\mathrm{Var}(f(X)) \simeq \frac{1}{n}.$$

More generally, if X is uniform on an isotropic L_p ball, and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\lim_{n\to\infty} \mathrm{Var}(f(X)) = 0.$$

Related results - Carbery-Wright

A more fundamental problem is that Carbery-Wright is too general. If X is uniform on $\sqrt{n}B_2^n$, and $f(x) = \frac{1}{\sqrt{n}}\|x\|^2$, an easy calculation shows,

$$\mathrm{Var}(f(X)) \simeq \frac{1}{n}.$$

More generally, if X is uniform on an isotropic L_p ball, and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\lim_{n\to\infty} \mathrm{Var}(f(X)) = 0.$$

Related results - Carbery-Wright

A more fundamental problem is that Carbery-Wright is too general. If X is uniform on $\sqrt{n}B_2^n$, and $f(x) = \frac{1}{\sqrt{n}}\|x\|^2$, an easy calculation shows,

$$\mathrm{Var}(f(X)) \simeq \frac{1}{n}.$$

More generally, if X is uniform on an isotropic L_p ball, and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\lim_{n\to\infty} \mathrm{Var}(f(X)) = 0.$$

Related results - Fourier analysis

There is also connection between anti-concentration and Fourier transforms that goes back to Esseen:

Lemma (Esseen's inequality)

Let X be a random variable with characteristic function φ, then for any $\varepsilon > 0$ and $t \in \mathbb{R}$,

$$\mathbb{P}(|X - t| < \varepsilon) \leq \varepsilon \int_{-2\pi/\varepsilon}^{2\pi/\varepsilon} |\varphi(\lambda)| d\lambda.$$

In particular, if $|\varphi(\lambda)| \lesssim \frac{1}{|\lambda|^\alpha}$, $\mathbb{P}(|X - t| < \varepsilon) \leq \varepsilon^\alpha$.

Related results - Fourier analysis

There is also connection between anti-concentration and Fourier transforms that goes back to Esseen:

Lemma (Esseen's inequality)

Let X be a random variable with characteristic function φ, then for any $\varepsilon > 0$ and $t \in \mathbb{R}$,

$$\mathbb{P}(|X - t| < \varepsilon) \leq \varepsilon \int_{-2\pi/\varepsilon}^{2\pi/\varepsilon} |\varphi(\lambda)| d\lambda.$$

In particular, if $|\varphi(\lambda)| \lesssim \frac{1}{|\lambda|^\alpha}$, $\mathbb{P}(|X - t| < \varepsilon) \leq \varepsilon^\alpha$.

Related results - Van der Corput lemma

Recall the classical Van der Corput lemma from the 30's. If $h : \mathbb{R} \to \mathbb{R}$, is such that $|h^{(k)}| \geq 1$. Then,

$$\int_{-1}^{1} e^{i\lambda h(x)} dx \lesssim \frac{1}{|\lambda|^{\frac{1}{k}}}.$$

In particular, if X is uniform on $[-1, 1]$ and $f(x) = x^d$,

$$\mathbb{P}(|f(X) - t| < \varepsilon) \lesssim \varepsilon^{\frac{1}{d}}.$$

What happens in higher dimensions?

Related results - Van der Corput lemma

Recall the classical Van der Corput lemma from the 30's. If $h : \mathbb{R} \to \mathbb{R}$, is such that $|h^{(k)}| \geq 1$. Then,

$$\int_{-1}^{1} e^{i\lambda h(x)} dx \lesssim \frac{1}{|\lambda|^{\frac{1}{k}}}.$$

In particular, if X is uniform on $[-1, 1]$ and $f(x) = x^d$,

$$\mathbb{P}(|f(X) - t| < \varepsilon) \lesssim \varepsilon^{\frac{1}{d}}.$$

What happens in higher dimensions?

Related results - Van der Corput lemma

Recall the classical Van der Corput lemma from the 30's. If $h : \mathbb{R} \to \mathbb{R}$, is such that $|h^{(k)}| \geq 1$. Then,

$$\int_{-1}^{1} e^{i\lambda h(x)} dx \lesssim \frac{1}{|\lambda|^{\frac{1}{k}}}.$$

In particular, if X is uniform on $[-1, 1]$ and $f(x) = x^d$,

$$\mathbb{P}(|f(X) - t| < \varepsilon) \lesssim \varepsilon^{\frac{1}{d}}.$$

What happens in higher dimensions?

Related results - Van der Corput lemma

Much work has been done on high-dimensional analogues of the Van der Corput lemma. Carbery, Christ and Wright showed,

$$\int_{[-1,1]^n} e^{i\lambda f(x)} dx \lesssim \frac{\mathrm{poly}(n)}{|\lambda|^{\frac{1}{d}}},$$

if f is a homogeneous degree d polynomial and for some I, $|v_I| \geq 1$.

They also asked the question: can the dependence on n be removed form the right hand side?

Related results - Van der Corput lemma

Much work has been done on high-dimensional analogues of the Van der Corput lemma. Carbery, Christ and Wright showed,

$$\int_{[-1,1]^n} e^{i\lambda f(x)} dx \lesssim \frac{\mathrm{poly}(n)}{|\lambda|^{\frac{1}{d}}},$$

if f is a homogeneous degree d polynomial and for some I, $|v_I| \geq 1$.

They also asked the question: can the dependence on n be removed form the right hand side?

Our results

Theorem

Let $X \sim \mu^{\otimes n}$ be a product measure and let $f(x) = \sum_I v_I x^I$ with $\sum v_I^2 = 1$. Then,

$$\mathrm{Var}(f(X)) \geq C_{\mu,d}.$$

Moreover, the constant can be taken to be uniform over all isotropic log-concave product measures.

Corrolary

Let $f(x) = \sum_I v_I x^I$ with $|v_I| \geq 1$ for some I. Then,

$$\int_{[-1,1]^n} e^{i\lambda f(x)} dx \lesssim \frac{1}{|\lambda|^{\frac{1}{d}}},$$

Our results

Theorem

Let $X \sim \mu^{\otimes n}$ be a product measure and let $f(x) = \sum_I v_I x^I$ with $\sum v_I^2 = 1$. Then,

$$\mathrm{Var}(f(X)) \geq C_{\mu,d}.$$

Moreover, the constant can be taken to be uniform over all isotropic log-concave product measures.

Corrolary

Let $f(x) = \sum_I v_I x^I$ with $|v_I| \geq 1$ for some I. Then,

$$\int_{[-1,1]^n} e^{i\lambda f(x)} dx \lesssim \frac{1}{|\lambda|^{\frac{1}{d}}},$$

Variance bound 1d

Let $X \sim \mu$ be random variable with infinite support. Apply the Gram-Schmidt algorithm to $\{1, x, x^2, \dots\}$ in $L^2(\mu)$ and consider the resulting orthogonal polynomials $\{p_k\}_{k=0}^{\infty}$.

Lemma

Let $f(x) = x^d$. Then,

1. $\langle f, p_k \rangle_{L^2(\mu)} = 0$ *for* $k > d$.
2. $\langle f, p_d \rangle_{L^2(\mu)} = \tilde{c}_{\mu,d} \neq 0$.

Proof.

1. p_k is orthogonal to degree d polynomials.
2. $f \notin \mathrm{span}\left(1, x, x^2, \dots, x^{d-1}\right)$.

□

Variance bound 1d

Let $X \sim \mu$ be random variable with infinite support. Apply the Gram-Schmidt algorithm to $\{1, x, x^2, \dots\}$ in $L^2(\mu)$ and consider the resulting orthogonal polynomials $\{p_k\}_{k=0}^{\infty}$.

Lemma

Let $f(x) = x^d$. Then,

1. $\langle f, p_k \rangle_{L^2(\mu)} = 0$ *for* $k > d$.
2. $\langle f, p_d \rangle_{L^2(\mu)} = \tilde{c}_{\mu,d} \neq 0$.

Proof.

1. p_k is orthogonal to degree d polynomials.
2. $f \notin \mathrm{span}\left(1, x, x^2, \dots, x^{d-1}\right)$.

□

Variance bound 1d

Let $X \sim \mu$ be random variable with infinite support. Apply the Gram-Schmidt algorithm to $\{1, x, x^2, \dots\}$ in $L^2(\mu)$ and consider the resulting orthogonal polynomials $\{p_k\}_{k=0}^{\infty}$.

Lemma

Let $f(x) = x^d$. Then,

1. $\langle f, p_k \rangle_{L^2(\mu)} = 0$ *for* $k > d$.
2. $\langle f, p_d \rangle_{L^2(\mu)} = \tilde{c}_{\mu,d} \neq 0$.

Proof.

1. p_k is orthogonal to degree d polynomials.
2. $f \notin \mathrm{span}\left(1, x, x^2, \dots, x^{d-1}\right)$.

□

Dimension-free variance bounds

Observe $L^2(\mu^{\otimes n}) = L^2(\mu)^{\otimes n}$. So, an orthonormal basis for $L^2(\mu^{\otimes n})$ is given by $\{p_I\}$, where for $I = (I_1, I_2, \ldots, I_n)$,

$$p_I(x) = \prod_{i=1}^{n} p_{I_i}(x_i).$$

Lemma

Let $f(x) = \sum_I v_I x^I$, be of degree d. Then, for $|J| = d$,

1. $\langle f, p_J \rangle_{L^2(\mu^{\otimes n})} \geq v_J \cdot c_{\mu,d}$

Dimension-free variance bounds

Observe $L^2(\mu^{\otimes n}) = L^2(\mu)^{\otimes n}$. So, an orthonormal basis for $L^2(\mu^{\otimes n})$ is given by $\{p_I\}$, where for $I = (I_1, I_2, \ldots, I_n)$,

$$p_I(x) = \prod_{i=1}^{n} p_{I_i}(x_i).$$

Lemma

Let $f(x) = \sum_I v_I x^I$, be of degree d. Then, for $|J| = d$,

1. $\langle f, p_J \rangle_{L^2(\mu^{\otimes n})} \geq v_J \cdot c_{\mu,d}$

Dimension-free variance bounds

Proof.

$$\langle f, p_J \rangle_{L^2(\mu^{\otimes n})} = \sum v_I \langle x^I, p_J \rangle_{L^2(\mu^{\otimes n})} = v_J \langle x^J, p_J \rangle_{L^2(\mu^{\otimes n})}$$
$$= v_J \prod_{i=1}^{d} \langle x^{J_i}, p_{J_i} \rangle_{L^2(\mu)} = v_J \prod_{i=1}^{d} \tilde{c}_{\mu, J_i}.$$

□

An L^2 decomposition gives

$$\mathrm{Var}(f(X)) = \sum_{I \neq 0} \langle f, p_I \rangle^2_{L^2(\mu^{\otimes n})}.$$

and we are ready to prove the theorem.

Dimension-free variance bounds

Proof.

$$\begin{aligned}\langle f, p_J \rangle_{L^2(\mu^{\otimes n})} &= \sum v_I \langle x^I, p_J \rangle_{L^2(\mu^{\otimes n})} = v_J \langle x^J, p_J \rangle_{L^2(\mu^{\otimes n})} \\ &= v_J \prod_{i=1}^{d} \langle x^{J_i}, p_{J_i} \rangle_{L^2(\mu)} = v_J \prod_{i=1}^{d} \tilde{c}_{\mu, J_i}.\end{aligned}$$

□

An L^2 decomposition gives

$$\mathrm{Var}(f(X)) = \sum_{I \neq 0} \langle f, p_I \rangle^2_{L^2(\mu^{\otimes n})}.$$

and we are ready to prove the theorem.

Dimension-free variance bounds

Proof of Theorem.

$$\mathrm{Var}(X) = \langle f, f \rangle_{L^2(\mu^{\otimes n})} - \langle 1, f \rangle^2_{L^2(\mu^{\otimes n})} = \sum_{I \neq 0} \langle f, p_I \rangle^2_{L^2(\mu^{\otimes n})}$$
$$\geq \sum_{|I|=d} \langle f, p_I \rangle^2_{L^2(\mu^{\otimes n})} \geq c^2_{\mu,d} \sum_{|I|=d} v_I^2$$
$$= c^2_{\mu,d}.$$

□

When μ is log-concave isotropic, by a comparison to an interval, we get $c_{\mu,d} = c^d$.

From variance bounds to sub-level estimates

We can now combine our result with the Carbery-Wright inequality.

Corrolary

Let X be a log-concave with a product law and let $f(x) = \sum_I v_I x^I$, be of degree d. Then, for any $\varepsilon > 0$ and $t \in \mathbb{R}$,

$$\mathbb{P}\left(|f(X) - t| \leq \varepsilon\right) \lesssim \varepsilon^{\frac{1}{d}}.$$

Proof.

$$\mathbb{P}\left(|f(X) - t| \leq \varepsilon\right) \lesssim \left(\frac{\varepsilon}{\sqrt{\mathbb{E}\left[f(X)^2\right]}}\right)^{\frac{1}{d}} \leq \left(\frac{\varepsilon}{c^d}\right)^{\frac{1}{d}}.$$

□

From variance bounds to sub-level estimates

We can now combine our result with the Carbery-Wright inequality.

Corrolary

Let X be a log-concave with a product law and let $f(x) = \sum_I v_I x^I$, be of degree d. Then, for any $\varepsilon > 0$ and $t \in \mathbb{R}$,

$$\mathbb{P}\left(|f(X) - t| \leq \varepsilon\right) \lesssim \varepsilon^{\frac{1}{d}}.$$

Proof.

$$\mathbb{P}\left(|f(X) - t| \leq \varepsilon\right) \lesssim \left(\frac{\varepsilon}{\sqrt{\mathbb{E}\left[f(X)^2\right]}}\right)^{\frac{1}{d}} \leq \left(\frac{\varepsilon}{c^d}\right)^{\frac{1}{d}}.$$

□

Multivariate Van der Corput

Let $f(x) = \sum_I v_I x^I$ with $|v_I| \geq 1$ for some $I = (I_1, \ldots, I_n)$. We wish to bound,

$$J(\lambda) := \int_{[-1,1]^n} e^{i\lambda f(x)} dx.$$

Define,

$$A := \left\{ x \in [-1,1]^n \middle| \left| \frac{d}{dx_n^{I_n}} f(x) \right| \geq \varepsilon \right\}.$$

So,

$$J(\lambda) \leq \left| \int_A e^{i\lambda f(x)} dx \right| + \left| \int_{\bar{A}} e^{i\lambda f(x)} dx \right|.$$

Multivariate Van der Corput

Let $f(x) = \sum_I v_I x^I$ with $|v_I| \geq 1$ for some $I = (I_1, \ldots, I_n)$. We wish to bound,

$$J(\lambda) := \int_{[-1,1]^n} e^{i\lambda f(x)} dx.$$

Define,

$$A := \left\{ x \in [-1,1]^n \middle| \left| \frac{d}{dx_n^{I_n}} f(x) \right| \geq \varepsilon \right\}.$$

So,

$$J(\lambda) \leq \left| \int_A e^{i\lambda f(x)} dx \right| + \left| \int_{\bar{A}} e^{i\lambda f(x)} dx \right|.$$

Multivariate Van der Corput

Let $f(x) = \sum_I v_I x^I$ with $|v_I| \geq 1$ for some $I = (I_1, \ldots, I_n)$. We wish to bound,

$$J(\lambda) := \int_{[-1,1]^n} e^{i\lambda f(x)} dx.$$

Define,

$$A := \left\{ x \in [-1,1]^n | \left| \frac{d}{dx_n^{I_n}} f(x) \right| \geq \varepsilon \right\}.$$

So,

$$J(\lambda) \leq \left| \int_A e^{i\lambda f(x)} dx \right| + \left| \int_{\bar{A}} e^{i\lambda f(x)} dx \right|.$$

Multivariate Van der Corput

We first bound $\bar{A}$.
The main observation is that $\frac{d}{dx_n^{l_n}} f$ is a polynomial of degree $d - l_n$ with sum of coefficients at least 1,

$$\left| \int_{\bar{A}} e^{i\lambda f(x)} dx \right| \leq \int_{\bar{A}} 1 dx = \mathbb{P}\left(\left| \frac{d}{dx_n^{l_n}} f(X) \right| \leq \varepsilon \right) \lesssim \varepsilon^{\frac{1}{d - l_n}}.$$

Multivariate Van der Corput

We also bound

$$\left| \int_{\bar{A}} e^{i\lambda f(x)} dx \right| \lesssim \frac{1}{(|\lambda|\varepsilon)^{\frac{1}{l_n}}}.$$

High-level idea:

- Decompose $x = (\tilde{x}, x_n)$ and $f_{\tilde{x}}(x_n) = f(x)$.
- On A, for every $\tilde{x}$, $|f_{\tilde{x}}^{(l_n)}| \geq \varepsilon$.
- Use one-dimensional results for $f_{\tilde{x}}$.

Multivariate Van der Corput

We also bound

$$\left| \int_{\bar{A}} e^{i\lambda f(x)} dx \right| \lesssim \frac{1}{(|\lambda|\varepsilon)^{\frac{1}{I_n}}}.$$

High-level idea:

- Decompose $x = (\tilde{x}, x_n)$ and $f_{\tilde{x}}(x_n) = f(x)$.
- On A, for every $\tilde{x}$, $|f_{\tilde{x}}^{(I_n)}| \geq \varepsilon$.
- Use one-dimensional results for $f_{\tilde{x}}$.

Multivariate Van der Corput

$$J(\lambda) \leq \left| \int_A e^{if\lambda(x)} dx \right| + \left| \int_{\bar{A}} e^{if\lambda(x)} dx \right|$$
$$\leq \frac{1}{(|\lambda|\varepsilon)^{\frac{1}{I_n}}} + \varepsilon^{\frac{1}{d - I_n}}.$$

Optimize over ε to get,

$$J(\lambda) \lesssim \frac{1}{|\lambda|^{\frac{1}{d}}}.$$

Question

Is the condition $|v_I| \geq 1$ necessary?

Multivariate Van der Corput

$$J(\lambda) \leq \left| \int_A e^{if\lambda(x)} dx \right| + \left| \int_{\bar{A}} e^{if\lambda(x)} dx \right|$$
$$\leq \frac{1}{(|\lambda|\varepsilon)^{\frac{1}{l_n}}} + \varepsilon^{\frac{1}{d-l_n}}.$$

Optimize over ε to get,

$$J(\lambda) \lesssim \frac{1}{|\lambda|^{\frac{1}{d}}}.$$

Question

Is the condition $|v_I| \geq 1$ necessary?

Beyond products

Recall that if $X \sim \mathrm{Uniform}(B_p^n)$ and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\mathrm{Var}(f(X)) = o(1).$$

However, in these cases we have,

$$\mathbb{E}\left[f(X)^2\right] = \omega(1).$$

Can we get dimension-free estimates on $\mathbb{E}\left[\left(X^{\otimes d}\right)\left(X^{\otimes d}\right)^T\right]$,

instead of $\mathrm{Cov}\left(X^{\otimes d}\right)$?

Is $\mathbb{E}\left[f(X)^2\right]$ large, when $\sum v_I^2 = 1$?

Beyond products

Recall that if $X \sim \mathrm{Uniform}(B_p^n)$ and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\mathrm{Var}(f(X)) = o(1).$$

However, in these cases we have,

$$\mathbb{E}\left[f(X)^2\right] = \omega(1).$$

Can we get dimension-free estimates on $\mathbb{E}\left[\left(X^{\otimes d}\right)\left(X^{\otimes d}\right)^T\right]$,

instead of $\mathrm{Cov}\left(X^{\otimes d}\right)$?

Is $\mathbb{E}\left[f(X)^2\right]$ large, when $\sum v_I^2 = 1$?

Beyond products

Recall that if $X \sim \mathrm{Uniform}(B_p^n)$ and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\mathrm{Var}(f(X)) = o(1).$$

However, in these cases we have,

$$\mathbb{E}\left[f(X)^2\right] = \omega(1).$$

Can we get dimension-free estimates on $\mathbb{E}\left[\left(X^{\otimes d}\right)\left(X^{\otimes d}\right)^T\right]$,
instead of $\mathrm{Cov}\left(X^{\otimes d}\right)$?

Is $\mathbb{E}\left[f(X)^2\right]$ large, when $\sum v_I^2 = 1$?

Beyond products

Recall that if $X \sim \mathrm{Uniform}(B_p^n)$ and $f(x) = \frac{1}{\sqrt{n}}\|x\|_p^p$,

$$\mathrm{Var}(f(X)) = o(1).$$

However, in these cases we have,

$$\mathbb{E}\left[f(X)^2\right] = \omega(1).$$

Can we get dimension-free estimates on $\mathbb{E}\left[\left(X^{\otimes d}\right)\left(X^{\otimes d}\right)^T\right]$,
instead of $\mathrm{Cov}\left(X^{\otimes d}\right)$?
Is $\mathbb{E}\left[f(X)^2\right]$ large, when $\sum v_I^2 = 1$?

Isotropic L_p balls

Let $Y \sim \frac{1}{z} e^{-\|x\|_p^p} dx$ and $U \sim \mathrm{Uniform}([0,1])$. Then

$$X = n^{\frac{1}{p}} U \frac{Z}{\|Z\|_p}.$$

Since Z is a product measure, for any homogeneous function,

$$\mathbb{E}\left[f(X)^2\right] \simeq n^{\frac{2d}{p}} \mathbb{E}\left[\frac{f(Z)^2}{\|Z\|_p^{2d}}\right] \gtrsim 1.$$

Euclidean balls

1. If X is uniform on the isotropic Euclidean ball, we identify all eigenvalues of $\mathrm{Cov}\left(X^{\otimes d}\right)$.
2. Eigenvectors are given by $\|x\|_2^{2k} H_{d-2k}$, where H_{d-2k} are degree $d-2k$ spherical harmonics.
3. If $f(x) = \|x\|_2^2$, $\mathrm{Var}(f(X)) \simeq \frac{1}{n}$.
4. If f is orthogonal to $\|x\|_2^2$, $\mathrm{Var}(f(X)) = \Omega(1)$.

Thank You