Lecture 13

Lecture date: Oct 11 *Scribe: Peter Ralph*

notation: Where the definition of μ is clear from context, we let $|f|_q = \left(\int |f(x)|^q \mu(dx)\right)^{1/q}$ denote the $L^q(\mu)$ norm. Similiarly, let $\langle f, g\rangle = \langle f, g\rangle_\mu = \sum f(x)g(x)\mu(dx)$ denote the inner product of $L^2(\mu)$.

Theorem 1 *For all $p < 2 < q$, any discrete probability measure μ whose smallest atom is of size α has the same $(2, q)$– and $(p, 2)$–hypercontractivity constants as the measure μ_α, that assigns mass α and $1-\alpha$ to 0 and 1, respectively.*

Proof:

The proof follows Wolff (2005). We will prove the result for the $(2, q)$–hypercontractivity contants, and the $(p, 2)$ case will follow by duality.

It is easy to show that if μ is $(2, q, \eta)$–hypercontractive, then so is μ_α. Indeed, suppose otherwise. Then by definition, there exists an $f : \{0, 1\} \to \mathbb{R}$ such that $|T_\eta f|_{L^q(\mu_\alpha)} > |f|_{L^2(\mu_\alpha)}$, where T_η is the Bonami–Bechner operator. Let x be such that $\mu(x) = \alpha$. If we then define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(y) = \begin{cases} f(0) & \text{if } y = x \\ f(1) & \text{o.w. } , \end{cases}$$

then we see that $|g|_{L^p(\mu)} = |f|_{L^p(\mu_\alpha)}$ and $|T_\eta g|_{L^p(\mu)} = |T_\eta f|_{L^p(\mu_\alpha)}$ for all p. However, by $(2, q, \eta)$–hypercontractivity of μ, $|T_\eta g|_{L^2(\mu)} \le |g|_{L^2(\mu)}$, which is a contradiction.

To show the converse, we will prove the following. Suppose that μ is not $(2, q, \eta)$–hypercontractive, and let f_0 maximize $|T_\eta f|_q$ among $\{f : |f|_2 = 1\}$. We will show that f_0 obtains at most (and hence exactly) two values. This will complete the proof of the theorem, since an f_0 taking two values induces a two-point measure μ_β on those values that is not $(2, q, \eta)$–hypercontractive. Since $\beta \ge \alpha$, and the hypercontractivity constants for μ_γ are monotone in γ, we have that μ_α is not $(2, q, \eta)$–hypercontractive either.

Exercise 2 (1 point) *Show that the hypercontractive constants for μ_α is monotone in α.*

The proof of the size of the range of f_0 is as follows. Let $I(f) = |T_\eta f|_q^q$, and $J(f) = |f|_2^2$. We wish to use the method of Lagrange multipliers (see www.wikipedia.org/wiki/Lagrange_Multiplier for background), to which end we will think of I and J as acting on the finite-dimensional real vector space of functions $f : \mathbf{spt}(\mu) \to \mathbb{R}$. Note that any linear function on this space can be represented uniquely as $\langle f, \cdot \rangle_\mu$ for some appropriate f, so we may write the derivate of, say, I, evaluated at a function g, as $DI(g) = \langle f(g), \cdot \rangle_\mu$, for some f depending on g.

By the method of Lagrange multipliers, $DI(f_0) = cDJ(f_0)$, for some constant c. Simple computation reveals that $DJ(f) = 2\langle f, \cdot \rangle_\mu$, and the chain rule allows us to also compute

$$\begin{aligned} DI(f_0) &= (D|\cdot|_q^q) \circ \eta(T_\eta f_0) \\ &= \eta \langle (T_\eta f_0)^{q-1}, T_\eta \cdot \rangle_\mu \\ &= \eta \langle T_\eta (T_\eta f_0)^{q-1}, \cdot \rangle_\mu, \end{aligned}$$

since $(DT_\eta)f = \eta$ (here D is the derivative from $\mathbb{R}^n \to \mathbb{R}^n$), and T_η is self-adjoint with respect to $\langle \cdot, \cdot \rangle_\mu$. Since $DI(f_0) = cDJ(f_0)$, by uniqueness, we have that $f_0 = CT_\eta (T_\eta f_0)^{q-1}$ for some constant C.

Let $g_0 = T_\eta f_0 + (1 - \eta)\mathbf{E}(f_0)$, and note that

$$f_0 = \frac{1}{\eta} g_0 - \frac{1-\eta}{\eta} \mathbf{E}(g_0), \tag{1}$$

and also that

$$f_0 = CT_\eta g_0^{q-1} = C(\eta g_0^{q-1} + (1-\eta)\mathbf{E}(g_0^{q-1})). \tag{2}$$

However, note that for each x, the first equation (1) is linear in $g_0(x)$, while the second equation (2) is strictly convex in $g_0(x)$. A linear function meets a strictly convex function in at most two points, so there are at most two solutions $(g_0(x), f_0(x))$ to (1) = (2), and f_0 takes at most two values.

□

Now we move on to the notion of hypercontractivity of random variables taking values in a separable Banach space (e.g. $\mathbb{R}^n$ with any of the usual norms), and relate it to our previous definition.

Throughout, we will denote by $|\cdot|$ the norm coming from the Banach space, and define a family of norms $\|\cdot\|_q$ on random variables in this Banach space by $\|Y\|_q := (\mathbf{E}|Y|^q)^{1/q}$.

Definition 3 *A random variable X taking values in a separable Banach space V is (p, q, σ)–*hypercontractive *for some $0 < p < q$ and $0 < \sigma < 1$ if, for all $v \in V$,*

$$\|v + \sigma X\|_q \le \|v + X\|_p.$$

Theorem 4 *For a finite probability space $(\Omega, \mathcal{F}, \mathbf{P})$, the following are equivalent:*

- *T_η is (p,q)–hypercontractive.*
- *Every mean–zero random variable X taking values in a separable Banach space is (p,q,η)–hypercontractive.*
- *Every mean–zero* **real–valued** *random variable X is (p,q,η)–hypercontractive.*

Proof: Trivially, B $\Rightarrow$ C.

(C $\Rightarrow$ A) Assume C, and let $g = f - \mathbf{E}f$, so that by letting $v = \mathbf{E}f$ and $X = \eta g$, we have $|T_\eta f|_q = |\mathbf{E}f + \eta g|_q \leq |\mathbf{E}f + g|_p = |f|_p$. This shows A.

(A $\Rightarrow$ B) By the triangle inequality, the function $f(x) := |v+x|$ is convex, so using Jensen's inequality twice and that $\mathbf{E}X = 0$,

$$\begin{aligned} T_\eta f(X) &= \eta f(X) + (1-\eta)\mathbf{E}f(X) \\ &\geq \eta f(X) + (1-\eta) f(\mathbf{E}X) \\ &= \eta f(X) + (1-\eta) f(0) \\ &\geq f(\eta X), \end{aligned}$$

and hence

$$\|v + \eta X\|_q = \|f(\eta X)\|_q \leq \|T_\eta f(X)\|_q \leq \|f(X)\|_p = \|v + X\|_p.$$

□

Exercise 5 (1 point) *Prove this lemma:*

Lemma 6 *Let $q > 2$, $\eta > 0$, and let X be a $(2,q,\eta)$–hypercontractive random variable such that $X \neq 0$. Then $\mathbf{E}X = 0$, $\mathbf{E}|X|^q < \infty$, and $\eta < (q-1)^{-1/2}$.*

The following lemma relates (p,q,η)–hypercontractivity of a random variable X to its' moments.

Lemma 7 *Let X be a mean-zero real-valued random variable with $\mathbf{E}|X|^q < \infty$, where $q > 2$. Then X is $(2,q,\eta_q)$–hypercontractive, where*

$$\eta_q = \frac{\|X\|_2}{\sqrt{q-1}\|X\|_q}.$$

Proof:

Let X' be an independent copy of X, and let $Y = X - X'$, so that Y is symmetric. Let ϵ be a further independent random variable taking values $\{+1, -1\}$ with probability $\frac{1}{2}$ each. Note that $\|Y\|_q \leq 2\|X\|_q$ by the triangle inequality, and that $Y =_d \epsilon Y$. We also know that ϵ is $(2, q, 1/\sqrt{q-1})$–hypercontractive.

The idea of this proof (symmetrization of X and using the hypercontractivity of ϵ) is due to Talagrand.

By Jensen's inequality, averaging over the value of X' and using $\mathbf{E}X' = 0$,

$$\|a + \eta_q X\|_q \leq \|a + \eta_q Y\|_q = \|a + \eta_q \epsilon Y\|_q.$$

Using hypercontractivity of ϵ, and where "$\mathbf{E}_Z$" means taking the expectation over Z (and conditioning on everything else),

$$\begin{aligned}
\|a + \eta_q \epsilon Y\|_q &\leq \left(\mathbf{E}_Y \left[\left(\mathbf{E}_\epsilon \left|a + \eta_q \epsilon Y \sqrt{q-1}\right|^2\right)^{q/2}\right]\right)^{1/q} \\
&= \left(\mathbf{E}\left[\left|a^2 + \eta_q^2 Y^2 (q-1)\right|^{q/2}\right]\right)^{1/q} \\
&= \|a^2 + \eta_q^2 Y^2 (q-1)\|_{q/2}^{1/2} \\
&\leq \left(a^2 + (q-1)\eta_q^2 \|Y^2\|_{q/2}\right)^{1/2} \\
&= \left(a^2 + \left(\frac{\|Y\|_q}{2\|X\|_q}\right)^2 \mathbf{E}X^2\right)^{1/2},
\end{aligned}$$

where the second inequality follows from Minkowski's inequality. Using the definition of η_q and that $\|Y\|_q \leq 2\|X\|_q$, we continue the above chain of inequalities to get that

$$\begin{aligned}
\|a + \eta_q X\|_q &\leq \left(a^2 + \mathbf{E}X^2\right)^{1/2} \\
&= \|a + X\|_2,
\end{aligned}$$

where the last equality follows from $\mathbf{E}X = 0$. $\square$