MASSACHUSETTS INSTITUTE OF TECHNOLOGY

6.265/15.070J Lecture 25 May 17, SP17

Lecturer: Yury Polyanskiy Scribe notes by Frederic Koehler

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. They are posted to serve class purposes.*

Hypercontractivity and Noise Stability

Content.

1. Proof of KKL via Hypercontractivity
2. Proof of Hypercontractivity
3. Noise Stability

Note: Everything below will be in the $p = 1/2$ setting.

1 Proof of KKL via Hypercontractivity

Last time: Fourier decomposition

$$f(x) = \sum_{S \subset [n]} \hat{f}(S) u_S(x)$$

where

$$u_S(x) = \prod_{i \in S} (-1)^{x_i}$$

and $\|u_S\|_2 = 1$ and the u_S are orthogonal.

Therefore

$$\hat{f}(S) = E[f u_S]$$

Then

$$I_i^{1/2}(f) = 4 \sum_{S \ni i} (\hat{f})^2(S)$$

so

$$I^{1/2}(f) = 4 \sum_S |S| \hat{F}^2(S)$$

Theorem 1 (KKL 88)**.** *There exists a universal $c > 0$ such that for all n and $f : \{0,1\}^n \to \{0,1\}$,*

$$I(f) \geq cVar(f)\log(1/\tau)$$

where

$$\tau = \max_i I_i(f)$$

Theorem 2 (Hypercontractivity)**.** *For any $f : \{0,1\}^n \to \mathbb{R}$, for all $t > 0$*

$$\sum_S e^{-2t|S|}\hat{f}^2(S) = (E|f(x)|^q)^{2/q}$$

where $q = 1 + e^{-2t}$.

Proof (KKL using Hypercontractivity). Poincare tells us that

$$4\sum |S|\hat{f}^2(S) \geq 4\sum_{S\neq\emptyset}\hat{f}^2(S) = 4Var(f)$$

so

$$I \geq 4Var(f)$$

But this is very loose in symmetric case where all coordinates have similar small influence.

Recall

$$\nabla_i f(x) = f(x) - f(x+e_i)$$

If f is boolean then $|\nabla_i f|^2$ is also boolean. Also

$$I_i(f) = E|\nabla_i f|^\alpha$$

for any $\alpha > 0$.

Last time we checked that $\widehat{\nabla_i f}(S)$ is 0 for $S \ni i$ and is $2\hat{f}(S)$ for $S \ni i$. Thus from hypercontractivity

$$4\sum_{S\ni i}\hat{f}^2(S)e^{-2t|S|} \leq (E|\nabla_i f|^q)^{2/q} = I_i^{2/q}$$

and for $q < 2$ we know $2/q > 1$

We only need hypercontractivity for some t, we pick $2t = \log 2$, then $q = 1 + 1/2 = 3/2$ so $2/q = 4/3$. Then we get the following inequality:

$$4\sum_{S\ni i}\hat{f}^2(S)2^{-|S|} \leq I_i^{4/3}$$

Summing over all i we get

$$4\sum_S |S|\hat{f}^2(S)2^{-|S|} \le \sum_i I_i^{4/3} \le \tau^{1/3}\sum_i I_i = \tau^{1/3}I(f)$$

and we can lowerbound for any L by

$$4\cdot 2^{-L}\sum_{|S|\le L}\hat{f}^2(S) \le 4\sum_S |S|\hat{f}^2(S)2^{-|S|}.$$

Thus

$$\sum_{0<|S|\le L}\hat{f}(S)^2 \le \frac{1}{4}2^L\tau^{1/3}I(f).$$

Also

$$\sum_{|S|>L}\hat{f}^2(S) \le \frac{1}{L}\sum_{|S|>L}|S|\hat{f}^2(S) \le \frac{1}{L}\sum_S |S|\hat{f}^2(S) = \frac{I}{4L}$$

Also from Poincare we know

$$\sum_{S\neq\emptyset}\hat{f}^2(S) \le \frac{1}{4}\sum_S |S|\hat{f}^2(S)$$

Observe

$$Var(f) = \sum_{S\neq\emptyset}\hat{f}^2(S) = \sum_{|S|\le L}\hat{f}^2(S) + \sum_{|S|>L}\hat{f}^2(S) \le \frac{1}{4}(2^L\tau^{1/3}+1/L)I$$

What remains is to balance the terms in the sum. We take $L = c_1\log(1/\tau) - c_2\log\log(1/\tau)$ ($c_1 = (\log 2)/3$) and then we get an upper bound of

$$c(\frac{1}{\log(1/\tau)})I(f)$$

which is the desired result. □

2 Proof of Hypercontractivity

Remaining question: how to prove hypercontractivity?

Let $\tilde{X}_t$ be a discrete time simple random walk on $\{0,1\}^n$. (not lazy). Define $N_t \sim Pois(n)$ and $X_t = \tilde{X}_{N_t}$ a continuous time markov chain (n-fold speedup of previously introduced continuous time markov chain, which was $Pois(1)$). So

$$X_t = X_0 + Ber(\frac{1-e^{-t}}{2})^{\otimes n}.$$

The kernel of this markov chain is

$$T_t(x,x') = ((1-e^{-t})/2)^{|x-x'|_H}((1+e^{-t})/2)^{n-|x-x'|_H}.$$

The following are basic properties of T_t:

Proposition 3.

1. *$T_tT_s = T_{t+s}$ for $t,s \geq 0$*
2. *$T_tf(x) = Ef(x+Z_t)$ and $Z_t \sim Ber((1-e^{-t})/2)$*
3. *$cov(X_t, X_0) = (1/4)\rho I_n$ where $\rho = e^{-t}$ is the "correlation coefficient" and $X_0 \sim Uniform$.*
4. *$\widehat{T_tf}(S) = e^{-t|S|}\hat{f}(S)$*
5. *$T_t^{(n)} = (T_t^{(1)})^{\otimes n}$, i.e. $T_t((x_1,\ldots,x_n),(x'_1,\ldots,x'_n)) = \prod T_t^{(1)}(x_i,x'_i)$ and*
$$T_t^{(1)} = \begin{bmatrix} (1+e^{-t})/2 & (1-e^{-t})/2 \\ (1=e^{-t})/2 & (1+e^{-t})/2 \end{bmatrix}$$

Only one which is worthwhile to check is 4th one. Observe

$$T_t^{(1)}(-1)^x = (1-\rho)(-1)^X = e^{-t}(-1)^X$$

(column matrix $1\ -1$)

Therefore
$$T_t^{(n)}u_S = \prod_{i\in S}(T_t^{(1)}u_i(x_i)) = e^{-t|S|}u_S$$

Corollary 4.

$$\|T_tf\|_2^2 = \|\sum_S T_t\hat{f}(S)u_S\|_2^2 = \sum_S \hat{f}^2(S)e^{-2t|S|}$$

and observe the rhs is the lhs of hypercontractivity inequality.

Note that T_t is a conditional expectation operator, and so therefore by Jensen

$$T_t\varphi \leq \varphi(T_t)$$

where φ convex. In particular

$$\|T_tf\|_2 \leq \|f\|_2$$

Also we know

$$\|f\|_q \leq \|f\|_p$$

whenever $q < p$.

Hypercontractivity (restated):

$$\|T_t f\|_2 \leq \|f\|_{1+e^{-2t}}$$

note this is stronger than just the 2-norm estimate by Jensen.

Proof: First check $n = 1$.

$$\|\frac{1}{2}\begin{bmatrix}(1+e^{-t})/2 & (1-e^{-t})/2 \\ (1=e^{-t})/2 & (1+e^{-t})/2\end{bmatrix}\begin{bmatrix}a \\ b\end{bmatrix}\|_2 \leq \|\begin{bmatrix}a \\ b\end{bmatrix}(1/2)\|_q$$

where e.g.

$$\|\begin{bmatrix}a \\ b\end{bmatrix}(1/2)\|_q = ((1/2)|a|^q + (1/2)|b|^q)^{1/q}$$

Fix one guy by rescaling, then check in matlab.

Induction Lemma: For matrices A, B suppose: $q < 2$ and

$$\|Af\|_2 \leq \|f\|_q$$

for any f and

$$\|Bg\|_2 \leq \|g\|_q$$

for any g and also $A_{ij}, B_{ij} \geq 0$. Then:

$$\|A \otimes Bh\|_2 \leq \|h\|_q$$

Proof: See Yury's note on Stellar.

Side-note: cool proof of Hypercontractivity is to know that there is equality at time 0 and then just differentiate with respect to t and get logarithmic sobolev inequality to check.

3 Noise Sensitivity

Define for any f boolean

$$N_f(\rho) = Ef(X_t)f(X_0) = (f, T_t f) = (T_{t/2}f, T_{t/2}f)$$

where $\rho = e^{-t}$ as before is the correlation coefficient. Note $t = 0$ corresponds to $\rho = 1$ and $t = \infty$ corresponds to $\rho = 0$. Thus

$$N_f(\rho) = \sum_S \hat{f}^2(S)e^{-t|S|} = \sum_S \hat{f}^2(S)\rho^{|S|}$$

and we see from the last expression that this is convex and increasing. Suppose $E[f] = \delta$ and f is boolean. Then we can visualize the graph of $N_f(\rho)$ as starting at δ^2 and curving upward as we increase the correlation ρ .

Corollary 5. *If $Ef = 1/2$ then*

$$N_f(\rho) \leq (1/4)(1+\rho)$$

(tight iff $f(x) = x_1$ or one of its symmetric partners)

Take f_n a sequence of boolean functions, each with domain $\{0,1\}^n$ s.t. $Ef_n \to \delta$.

Definition 1. *$\{f_n\}$ is noise sensitive if*

$$N_{f_n}(\rho) - N_{f_n}(0) \to_{n\to\infty} 0$$

for all $\rho \in (0,1)$. ("asymptotically flat")

Definition 2. *$\{f_n\}$ is noise-stable if*

$$\sup_n (N_{f_n}(1) - N_{f_n}(\rho)) \to_{\rho\to 1} 0$$

i.e. "uniformly continuous at 1".

These definitions are almost complementary but not strictly so, e.g. a constant function satisfies both.

Idea: noise sensitive should be like, for all t

$$Ef_n(x_t)f_n(x_0) \approx E^2 f_n(x_0)$$

We consider the following example from percolation: what is the probability there exists a L to R crossing on an $n \times n$ square? This is exactly $1/2$ because there is either an left-to-right crossing in the primal lattice or an up-down crossing in the dual lattice. Also if we consider majority, we see that $Ef_{maj}(x) = 1/2$. But these functions are different: majority is noise stable and the left-to-right crossing is noise-sensitive.

Observe

$$N_{f_{max}}(\rho) = Emaj(X_t)maj(X_0)$$

which by CLT is

$$P[Z_1 > 0, \rho Z_1 + \sqrt{1-\rho^2} Z_2 > 0]$$

where Z_1, Z_2 are iid standard normals. And can compute that this is (look at 2d intersection of halfplanes with specified angle)

$$\frac{1}{4} + (arcsin(\rho)/2\pi)$$

Previously we saw dictator is most noise stable but it has very nasty/"unfair" influences. In fact there is a theorem which asserts that among functions without nice influences, majority is asymptotically stablest.

Theorem 6 (MOO '10)**.** *For every $\delta > 0$ there exists $\tau > 0$ such that if $I_{max}(f) < \tau$ then*

$$N_f(\rho) \leq N_{maj}(\rho) + \delta$$

i.e. "majority is stablest".

The next theorem gives a criterion for proving functions are noise sensitive.

Theorem 7 (BKS 99)**.** *If $\sum I_i^2(f_n) \to 0$ as $n \to \infty$ then f_n is noise sensitive. Furthermore this is an iff when f_n are monotone.*

Weaker statement which suffices for most applications: if $\sum I_i^{2-\epsilon} \to 0$ then noise sensitive. Proof: Application of Hypercontractivity, very similar to use in KKL theorem.

$$\sum |S| \hat{f}^2(S) e^{-2t|S|} \leq \sum_i I_i^{2-\epsilon} \to 0$$

which means f is conecentrated on high frequencies and thus noise sensitive.

Recall that our application of KKL to prove the Kesten theorem uses the idea that when something is small (max influence), then something else is very big (derivative equal to total influence). We give an example of this philosophy with percolation and BKS.

Pick a single edge e in percolation on $n \times n$ square lattice. If this edge is influential then there is a crossing through it in both primal and dual lattice. Draw circle of radius r, we call this a "four-arm" event in this circle.

$$I_\rho(f) \leq P[\text{four-arm event in } B(r = d(e, \partial))] = r^{-5/4}$$

where ∂ denotes the boundary and the last equality is well-known to percolation theorists. Then we get $n^2 n^{-10/4} = n^{-1/2} \to 0$ in BKS theorem which proves noise sensitivity of the left-to-right crossing. (Ignoring the boundary)

Let $f_n(x_t)$ be the indicator that there exists L-R crossing in X_t.

Sample X_0 iid bernoulli in $1/2$. Dynamical percolation: suppose we know at time 0 there is a crossing and each edge "blinks" with poisson rate 1

$$E[f_n(X_t) | f_n(X_0) = 1] = 2N_{f_n}(e^{-t})$$

which goes to $2(1/2)^2 = 1/2$ very quickly. So even after a very short time there is basically a $1/2$ chance that there exists a left to right crossing.

Note monotone noise-sensitive functions necessarily have a sharp threshold and we can prove noise sensitivity by BKS, which is one way to see that percolation etc. have sharp thresholds.

Inspirational quote: There are two stages of not knowing, the first is not knowing what exists... hopefully you are at the second stage, not knowing things but knowing that they exist.