High-Dimensional Measures and Geometry

Lecture Notes from Feb 4, 2010

taken by Anando Sen

We apply the martingale method to the boolean cube.

4.1.1 Theorem. *Let* $f : I_n \to \mathbb{R}$ *be* 1*-Lipshitz,* $a_f = \int_{I_n} f d\mu_n$*. Then for all* $t \geq 0$*,*

$$\mu_n(\{x \in I_n : f(x) \geq a_f + t\}) \leq e^{\frac{-2t^2}{n}}$$
$$and \quad \mu_n(\{x \in I_n : f(x) \leq a_f - t\}) \leq e^{\frac{-2t^2}{n}}$$

Proof. Take $\mathcal{F}_0 = \{\emptyset, I_n\}$, $\mathcal{F}_n$ be the maximal σ-algebra and $\mathcal{F}_1 = \{\emptyset, I_n^1, I_n^0, I_n\}$ and proceed by subdividing $I_n^1, I_n^0 \cong I_{n-1}$.

A function f is measurable with respect to $\mathcal{F}_k$ if it depends on the last k coordinates only.

Estimate $|f_n - f_{n-1}|$:
We see that,

$$f_{n-1}(x_1', x_2, \ldots x_n) = \frac{1}{2}\left(f_n(0, x_2, \ldots, x_n) + f_n(1, x_2, \ldots, x_n)\right)$$

By Lipshitz continuity $|f_n(0, x_2, \ldots, x_n) - f_n(1, x_2, \ldots, x_n)| \leq 1$. So we have,

$$|f_{n-1}(x) - f_n(x)| \leq \frac{1}{2} = d_n$$

Proceding iteratively, we get

$$\|f_i - f_{i-1}\|_\infty \leq \frac{1}{2} = d_i$$

Now the theorem gives desired estimate. Rescaling $t = \epsilon\sqrt{n}$ gives,

$$\mu_n(\{x \in I_n : |f(x) - a_f| \geq \epsilon\sqrt{n}\}) \leq 2e^{-2\epsilon^2}$$

□

5 Concentration in Product Spaces

5.1 The martingale method on product spaces

Let $X = X_1 \times X_2 \ldots X_n$ and each X_i be equipped with a probability measure μ_i. Then X can be equipped with a product measure with measurable sets in a σ-algebra generated by, $A = A_1 \times A_2 \ldots A_n$, where each A_i is measurable in X_i and $\mu(A) = \prod_{i=1}^{n} \mu_i(A_i)$.Also let $d(x,y) = |\{i : x_i \neq y_i\}|$ denote the Hemming distance between $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$.

5.1.1 Theorem. *Let $f : X \to \mathbb{R}$ be integrable and $d_1, d_2, \ldots, d_n$ be such that $|f(x) - f(y)| \leq d_i$ if x and y have $x_j = y_j$ for all j except $j = i$. Let $a = \int_X f d\mu_i$, $D = \sum_{j=1}^{n} d_j^2$. Then for $t \geq 0$,*

$$\mu(\{x : f(x) \geq a + t\}) \leq e^{\frac{-t^2}{2D}}$$
$$\mu(\{x : f(x) \leq a - t\}) \leq e^{\frac{-t^2}{2D}}.$$

Proof. Uses Martingale method.
Let $f_0 : X \to \{a\}$, $f_n = f$ and,

$$f_i(x_1, x_2, \ldots, x_i, x'_{i+1}, x'_{i+2} \ldots x'_n) = \int_{X_{i+1} \times \ldots \times X_n} f(x_1, x_2, \ldots x_n) d\mu_{i+1} \ldots d\mu_n.$$

This means for f_i we have averaged over $n - i + 1$ dimensions. This is a conditional expectation with respect to $\mathcal{F}_i$ generated by $A_1 \times A_2 \ldots A_i \times X_{i+1} \times \ldots X_n$ with $A_j \subseteq X_j$ measurable.

If $g(x) = f_i(x) - f_{i-1}(x)$ then, this difference comes from averaging over the ith coordinate and thus,

$$|g_i(x)| = \left| f_i(x'_1, x'_2, \ldots x'_i, \ldots, x'_n) - \int f_i(x'_1, \ldots, x'_{i-1}, x_i, x'_{i+1}, \ldots x'_n) d\mu_i(x_i) \right| \leq d_i$$

for each $1 \leq i \leq n$. Moreover,

$$\begin{aligned} E[g_i|\mathcal{F}_{i-1}] &= E[f_i - f_{i-1}|\mathcal{F}_{i-1}] \\ &= E[f_i|\mathcal{F}_{i-1}] - f_{i-1} \\ &= f_{i-1} - f_{i-1} = 0 \end{aligned}$$

So the Martingale method applies verbatim, as before. □

5.2 Law of large Numbers

Suppose $h : Y \to \mathbb{R}$ is integrable and $a = E[h] = \int_Y h d\nu$. If we "sample" n copies of h independently and average then how far would

$$f(y) = \frac{h(y_1) + \ldots + h(y_n)}{n}$$

typically be from a?

Let $X_i = Y$ and $\mu_i = \nu$ and let $x = (x_1, x_2, \ldots, x_n)$, then consider

$$f(x) = \frac{h(x_1) + \ldots + h(x_n)}{n}$$

Assume $0 \leq h(x_1) \leq d$ for all $x_1 \in X_1$, then changing one coordinate of $(x_1, x_2, \ldots, x_n)$ changes $f(x)$ by at most $\frac{d}{n}$. Thus we can apply the preceding theorem with,

$$D = \sum_{i=1}^{n} \left(\frac{d}{n}\right)^n = \frac{d^2}{n}$$

We conclude that,

$$\mu_n(\{x : f(x) \geq a + t\}) \leq e^{\frac{-nt^2}{2d^2}}$$
$$and \quad \mu_n(\{x : f(x) \leq a - t\}) \leq e^{\frac{-nt^2}{2d^2}}$$

Choosing $t = \frac{\epsilon d}{\sqrt{n}}$ gives,

$$\mu_n\left(\{x : |f(x) - a| \geq \frac{\epsilon d}{\sqrt{n}}\}\right) \leq 2e^{\frac{\epsilon^2}{2}}$$

5.3 Vector Valued Functions

If we have $h_1, h_2, \ldots, h_N$ instead of just one function h, with corresponding averages $a_1, a_2, \ldots, a_N$ and we sample,

$$f_i(x) = \frac{h_i(x_1) + \ldots + h_i(x_n)}{n}$$

then which n should we choose to guarentee that each average f_i is close to a_i on a set of large measure?

Using union bound,

$$\mu_n(\{x : |f_i(x) - a_i| \leq t \;\forall i\}) \geq 1 - 2Ne^{\frac{-nt^2}{2d^2}}$$

Choosing $t = \epsilon d$ gives $1 - 2Ne^{\frac{-n\epsilon^2}{2}}$ on RHS.