10/23/23

Lecture 13: PAC Learning Cont'd

Contents:

1). Noise stability $\to$ Fourier concentration

2). Stein's lemma (baby version)

Let $f: \{\pm 1\}^n \to \{\pm 1\}$, $\eta \in (0, 1/2)$

Recall noise sensitivity

$$NS_\eta(f) \triangleq \Pr_{x,x'}[f(x) \neq f(x')],$$

where $x \sim \{\pm 1\}^n$ and x' given by flipping each coordinate of x with probability η.

Lemma: $NS_\eta(f) = \frac{1}{2} - \frac{1}{2} \sum_{S \subseteq [n]} (1-2\eta)^{|S|} \hat{f}[S]^2.$

Pf: Note $\mathbb{E}[f(x)f(y)] = \Pr(f(x) = f(y)) - \Pr[f(x) \neq f(y)]$

$$= 1 - 2\Pr[f(x) \neq f(y)]$$

$$= 1 - 2NS_\eta(f),$$

so $$NS_\eta(f) = \frac{1}{2} - \frac{1}{2}\mathbb{E}[f(x)f(y)]$$

$$= \frac{1}{2} - \frac{1}{2}\sum_{S,T} \hat{f}[S]\hat{f}[T]\,\mathbb{E}[x_S y_T]$$

Note: $\mathbb{E}[x_S y_T] = \underbrace{\mathbb{E}[x_{S \setminus T}]\mathbb{E}[y_{T \setminus S}]}_{=0 \text{ if } S \neq T} \cdot \mathbb{E}[x_{S\cap T} y_{S \cap T}]$

and $\mathbb{E}[x_i y_i] = (1-\eta) - \eta = 1-2\eta$,

so $\mathbb{E}[x_S y_T] = \mathbb{1}[S=T] \cdot (1-2\eta)^{|S|}$. $\square$

<u>Lemma:</u>
$$\sum_{|S| \geq 1/\eta} \hat{f}(S)^2 \lesssim NS_\eta(f)$$

<u>Pf:</u>
$$2NS_\eta(f) = 1 - \sum_{S \subseteq [n]} (1-2\eta)^{|S|} \hat{f}[S]^2$$

by previous lemma.

Note $1 = \mathbb{E}[f^2] = \sum_{S \subseteq [n]} \hat{f}[S]^2$, so

$$= \sum_{S \subseteq [n]} \hat{f}(S)^2 \left(1 - (1-2\eta)^{|S|}\right)$$

$$\geq \sum_{|S| \geq 1/\eta} \hat{f}(S)^2 \underbrace{\left(1 - (1-2\eta)^{|S|}\right)}_{\geq 1-e^{-2}},$$

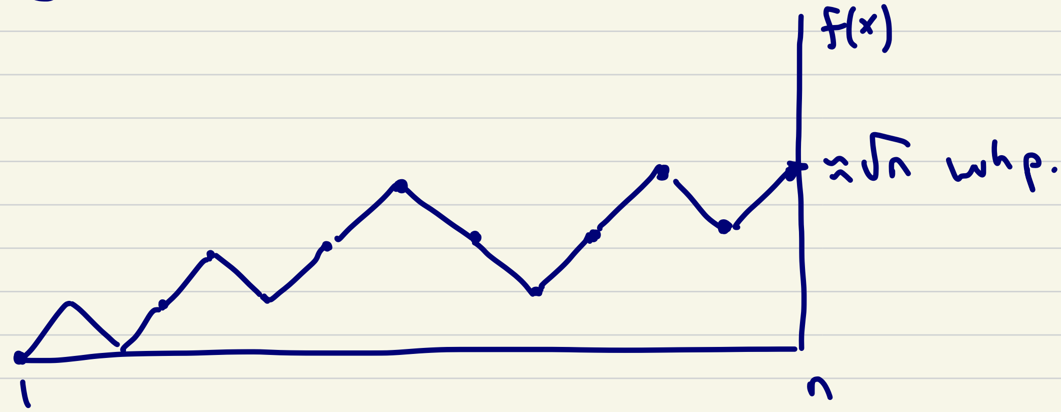
so rearranging gives desired bound. $\square$

<u>Thm</u> [Peres '04]: If $f(x) = \mathrm{sgn}(\langle w, x \rangle)$, then

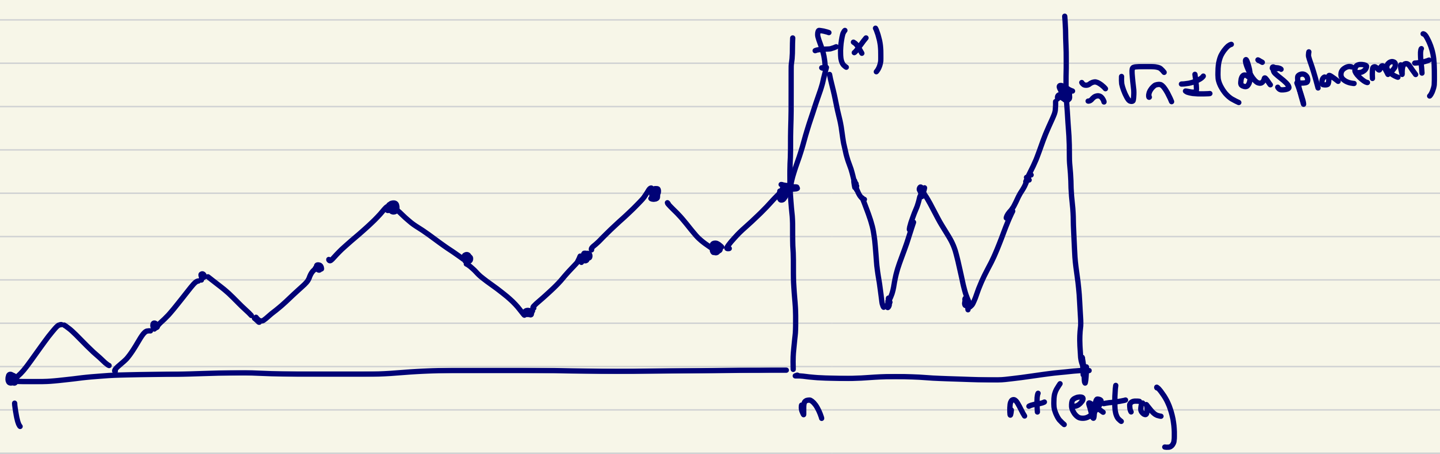
$$NS_\eta(f) \lesssim \sqrt{\eta}.$$

<u>Pf sketch</u> for baby version where $w = (1, \dots, 1)$, i.e. when f is $\mathrm{Majority}(x_1, \dots, x_n)$:

Can view $f(x)$ for $x \sim \{\pm 1\}^n$ as random walk:



Can sort of view $f(y)$ as <u>continuation</u> of this random walk for $\approx \eta n$ steps, where each step we move by 2 instead of 1 (b/c we are flipping coordinates from ± 1 to ∓ 1). So



in order for $f(y) \neq f(x)$, displacement must exceed $\sqrt{n}$. But

$$\Pr\left[\text{displacement} > \sqrt{n}\right] \le \frac{\mathbb{E}(\text{displacement})}{\sqrt{n}}$$

$$\lesssim \sqrt{\eta}. \qquad \square$$

<u>Corollary</u>: Let $h : \{\pm 1\}^k \to \{\pm 1\}$, $w_1, \dots, w_k \in \mathbb{R}^n$. Define $f(x) \triangleq h(\operatorname{sgn}(\langle w_1, x\rangle), \dots, \operatorname{sgn}(\langle w_k, x\rangle))$. Then

$$NS_\eta(f) \lesssim k\sqrt{\eta}$$

<u>Pf</u>:
$$\Pr[f(x) \neq f(y)] \le \sum_{i=1}^{k} \Pr[\operatorname{sgn}(\langle w_i, x\rangle) \neq \operatorname{sgn}(\langle w_i, y\rangle)]$$

$$= k \cdot \max_w NS_\eta(\langle w, x\rangle)$$

$$\lesssim k\sqrt{\eta}. \qquad \square$$

For certain h, can do much better, e.g.

[Kane '14]: If $f = \mathrm{AND}(\mathrm{sgn}(\langle w_1, x\rangle), \ldots, \mathrm{sgn}(\langle w_k, x\rangle))$, then $NS_\eta(f) = O(\sqrt{\lg(k)} \cdot \eta)$, so intersections of halfspaces can be learned in time

$$n^{O(\lg(k)/\varepsilon^2)}.$$

Stein's lemma (baby version):

Let $f: \mathbb{R} \to \mathbb{R}$ be sufficiently "regular". Then for $x \sim N(0,1)$,

① $\mathbb{E}[f(x) \cdot x] = \mathbb{E}[f'(x)]$.

② $\mathbb{E}[f(x) \cdot (x^2 - 1)] = \mathbb{E}[f''(x)]$. $\frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}}$

Pf: Let γ denote Gaussian density. Note:

$$x \cdot \gamma(x) = -\gamma'(x)$$

$$\int_{-\infty}^{\infty} \underline{f(x)} \cdot \underline{x \cdot \gamma(x)}\, dx \underset{\substack{\text{(int. by} \\ \text{parts)}}}{=} \int_{-\infty}^{\infty} f'(x) \cdot \gamma(x)\, dx + \underbrace{f(x)\gamma(x)\Big|_{-\infty}^{\infty}}_{\substack{=0 \\ \text{b/c } \gamma \to 0}}$$

as $|x| \to \infty$

$$= \mathbb{E}[f'(x)]$$

Similarly, note:

$$\gamma''(x) = (-x \cdot \gamma(x))' = (x^2-1)\gamma(x)$$

$$\int_{-\infty}^{\infty} f(x) \cdot (x^2-1)\gamma(x)\,dx = \int_{-\infty}^{\infty} f'(x) \cdot (x \cdot \gamma(x))\,dx + \underbrace{f(x) \cdot x \cdot \gamma(x) \Big|_{-\infty}^{\infty}}_{0}$$

$$= \int_{-\infty}^{\infty} f''(x)\,\gamma(x)\,dx + \underbrace{f'(x) \cdot \gamma(x)\,dx \Big|_{-\infty}^{\infty}}_{0}$$

$$\mathbb{E}[f''(x)].$$

$\square$