Lecture 6: ReLU linearization

Plan for next lectures:

- 1-4 non-algorithmic apx
- 5-8 "algorithmic apx"
- 9-10
- 11 - opt

Last time, this time

* $F \approx F_0$

  $F(x;W_0) + \langle \partial F(x;W_0), W - W_0 \rangle$

$F(x;V) = \sum_j a_j \sigma(v_j^T x)$

* $F_0 \approx$ "infinite-width NTK"

  universal approximator

* kernel perspective

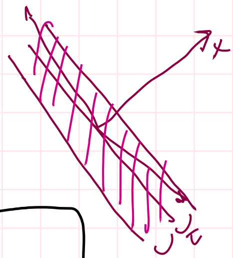
Theorem. Let $B \geq 0$ be given, and $x \in \mathbb{R}^d$. Then, with pr $\geq 1-\delta$, for all $\|W - V_0\| \leq B$, $\|V - V_0\| \leq B$, $\|x\| \leq 1$

$$\left| F(x;W) - \left(F(x;V) + \langle \partial F(x;V), W - V \rangle\right) \right|$$
$$\leq \|a\|_\infty \cdot m^{1/3} \, \cancel{\|x\|} \left(5B^{4/3} + 2B \left(\ln 1/\delta\right)^{1/4}\right).$$

Lemma ("baby anti-concentration").

For any $\tau > 0$, and any $\|x\| > 0$ with probability $\geq 1-\delta$ over $(v_j)_{j=1}^m$, $v_j \sim \mathcal{N}(0, I_d)$,

$$\sum_j \mathbb{1}\left[|v_j^T x| \leq \tau \|x\| \right] \leq m\tau + \sqrt{\frac{m}{2} \ln 1/\delta}.$$



$$\boxed{\mathbb{E}\,|v^T x| \approx 1}$$

$\mathbb{E}\,|v_j^T x| \leq \mathbb{E}\,\|v\| \leq \sqrt{\frac{\mathbb{E}\|v\|^2}{d}}$

$\geq$ (Taylor...)

$\mathbb{E}\,|v^T x| = \mathbb{E}\,|v_1| \cdot \|x\|$

$\approx \|x\|$

Lemma. $P_j := \mathbb{1}\left[|v_j^T x| \leq \tau \|x\| \right]$

$Q_j := \mathbb{1}\left[|g_j| \leq \tau \right]$, $g_j \sim \mathcal{N}(0,1)$.

$v \overset{iid}{=} Mv$

"rotational invariance of Gaussians"

$$\mathbb{1}\left[|(Mv_j)^T x| \leq \tau \|x\| \right] = \mathbb{1}\left[|v_1| \cdot \cancel{\|x\|} \leq \tau \cancel{\|x\|} \right] = \mathbb{1}\left[|v_1| \leq \tau \right]$$

$$M := \begin{bmatrix} | & | & | & | \\ x/\|x\| & & \text{orthogonal} & \\ | & | & | & | \end{bmatrix}$$

$$\mathbb{E}\,Q_j = \Pr\left[|g_j| \leq \tau \right] = \frac{1}{\sqrt{2\pi}} \int_{-\tau}^{\tau} \exp(-g^2/2)\, dg \leq \frac{1}{\sqrt{2\pi}} \int_{-\tau}^{\tau} dg$$

"baby anti-concentration"

$$= \frac{2\tau}{\sqrt{2\pi}} = \tau \sqrt{\frac{2}{\pi}}.$$

By Hoeffding

$$\sum_j \mathbb{1}\left[|v_j^T x| \leq \tau \|x\| \right] \leq m\tau \cdot \sqrt{\frac{2}{\pi}} + \sqrt{\frac{m}{2} \ln 1/\delta}$$

"benign anti-concentration"

Proof.

Let W, V, be given

$$\left|F(x;W) - \left(F(x;V) + \langle \bar\partial F(x;V), W-V\rangle\right)\right|$$

$$= \left|\sum_j a_j \left(\sigma(w_j^T x) - \sigma(v_j^T x) - \sigma'(v_j^T x)\langle x, w_j - v_j\rangle\right)\right|$$

$\sigma'(w_j^T x) w_j^T x$

$$= \left|\sum_j a_j w_j^T x \left(\sigma'(w_j^T x) - \sigma'(v_j^T x)\right)\right|$$

??

$\leq \ldots$

$$= \left|\sum_{j\in S} a_j w_j^T x \left(\sigma'(w_j^T x) - \sigma'(v_j^T x)\right)\right|$$

$$\leq \|a\|_\infty \sum_{j\in S} |w_j^T x| \cdot \left|\sigma'(w_j^T x) - \sigma'(v_j^T x)\right|$$

$$\leq \|a\|_\infty \sum |w_j^T x - v_j^T x| \cdot |\sigma'(w_j^T x) - \sigma'(v_j^T x)|$$

$$\leq \|a\|_\infty \sum_{j\in S} \|w - v\| \cdot 1$$

$$\leq \|a\|_\infty \sqrt{|S|} \sqrt{\textstyle\sum_j \|w_j - v_j\|^2}$$

$$= \|a\|_\infty \sqrt{|S|}\, \|W - V\|.$$

To finish, choose $\tau = B^{2/3}/m^{1/3}$.

Let τ be arbitrary (chosen later), & define

$$S_1 := \{ j : |v_{0,j}^T x| \leq \tau \|x\| \}$$

$$S_2 := \{ j : \|w_j - v_{0,j}\| \geq \tau \}$$

$$S_3 := \{ j : \|v_j - v_{0,j}\| \geq \tau \}.$$

$$S = S_1 \cup S_2 \cup S_3.$$

w/ pr $\geq 1-\delta$, $\quad |S_1| \leq m\tau + \sqrt{\frac{m}{2}\ln 1/\delta}$

$$B^2 \geq \|W - V_0\|^2 = \sum_j \|w_j - v_{0,j}\|^2 \geq \tau^2 |S_2|$$

$$\Rightarrow |S_2| \leq \frac{B^2}{\tau^2}$$

$$\Rightarrow |S| \leq m\tau + \sqrt{\frac{m}{2}\ln 1/\delta} + \frac{2B^2}{\tau^2}$$

Suppose $j \in S^c \Rightarrow |v_{0,j}^T x| > \tau\|x\|$

consider case $v_{0,j}^T x > \tau\|x\|$.

(other 3 cases "similar")

$$w_j^T x = (v_{0,j} + w_j - v_{0,j})^T x \geq \tau\|x\| - \|w_j - v_{0,j}\| \cdot \|x\| > 0.$$

$\Rightarrow j \in S^c$, then $\mathbb{1}[v_j^T x \geq 0] = \mathbb{1}[v_{0,j}^T x \geq 0] = \mathbb{1}[w_j^T x \geq 0]$.

$B^{2/3}/m^{1/3}$

Recall: given x, B, w/ pr $\geq 1-\delta$, $\forall \|W - V_0\| \leq B$, $\|V - V_0\| \leq B$,

$$|F - (t_x, l_r)| \leq \|a\|_\infty m^{1/3}\left(5B^{4/3} + 2B\left(\ln\tfrac{1}{\delta}\right)^{1/4}\right),$$

$\|W - V_0 + V_0 - V\|$

$$\sum_j |w_j^T x| \cdot \left|\sigma'(w_j^T x) - \sigma'(v_j^T x)\right|$$

- either 0
- or $\sigma'(w_j^T x) = 1 \neq 0 = \sigma'(v_j^T x) \Rightarrow |w_j^T x| = w_j^T x \leq w_j^T x - v_j^T x \leq |w_j^T x - v_j^T x|$
- or $\sigma'(w_j^T x) = 0 \neq 1 = \sigma'(v_j^T x) \Rightarrow$ "similar"

Next

Office hours

(Hoeffding in lemma)

$$\sum_j X_j \Rightarrow \text{w/ pr} \geq 1-\delta \quad \sum_j X_j \leq \mathbb{E}\sum_j X_j - \sqrt{\frac{m(b-a)^2}{2}\ln\frac{1}{\delta}}$$

iid $\quad x \in [a,b]$

$$\sum_j R_j = \sum_j Q_j \geq \underbrace{\sum_j \mathbb{E}\, Q_j}_{m \cdot \tau \sqrt{\frac{2}{\pi}}} - \sqrt{\frac{m}{2}\ln\frac{1}{\delta}}$$

("other 3 cases")

$j \in S \Rightarrow |v_{0,j}^T x| > \tau\|x\|$ (because $j \in S_1$)

- $v_{0,j}^T x > 0$
  - $w_j^T x = (v_j + w_j - v_j)^T x \geq \tau\|x\| - \|w_j - v_{0,j}\|\|x\|$ $\Rightarrow \sigma'(w_j^T x) = \sigma'(v_{0,j}^T x)$
    $\geq \tau\|x\| - \tau \cdot \|x\| = 0.$
  - $v_j^T x > 0 \Rightarrow \sigma'(v_j^T x) = \sigma'(v_{0,j}^T x)$
- $v_{0,j}^T x \leq 0 \Rightarrow v_{0,j}^T x < -\tau\|x\|$
  - $w_j^T x \ldots < 0 \Rightarrow$ same signs
  - $v_j^T x \ldots < 0 \Rightarrow$ same signs

e.g., $w_j^T x = \langle v_{0,j} + w_j - v_{0,j}, x\rangle$

$\leq \langle v_{0,j}, x\rangle + \|w_j - v_{0,j}\| \cdot \|x\|$

$< -\tau\|x\| + \tau\|x\|$

$= 0.$