

Probability bound of Lemma 2 and 3

Chanwoo Lee

Lemma 2 (Estimation error due to permutation (Expectation version)).

$$\mathbb{E} \|\text{Block}_k(\mathcal{Y} \circ \hat{\sigma}^{-1}) - \text{Block}_k(\mathcal{Y} \circ \sigma^{-1})\|_F^2 \leq d^m \text{Loss}^2(\sigma, \hat{\sigma}).$$

Lemma 2* (Estimation error due to permutation (Probability version)). With probability $1 - k^m \exp(-h^m \epsilon^2)$,

$$\|\text{Block}_k(\mathcal{Y} \circ \hat{\sigma}^{-1}) - \text{Block}_k(\mathcal{Y} \circ \sigma^{-1})\|_F^2 \leq d^m (\epsilon^2 + \text{Loss}^2(\sigma, \hat{\sigma})).$$

Proof. Define $\mathcal{A} := \mathcal{Y} \circ \sigma^{-1}$ and $\hat{\mathcal{A}} := \mathcal{Y} \circ \hat{\sigma}^{-1}$. Notice that

$$\|\text{Block}_k(\mathcal{Y} \circ \hat{\sigma}^{-1}) - \text{Block}_k(\mathcal{Y} \circ \sigma^{-1})\|_F^2 = h^m \sum_{\substack{k_i \in \{0, \dots, k-1\} \\ i=1, \dots, m}} \underbrace{\left(\frac{1}{h^m} \sum_{\substack{h_j \in \{0, \dots, h-1\} \\ j=1, \dots, m}} \hat{\mathcal{A}}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} - \mathcal{A}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} \right)^2}_{(*)}, \quad (1)$$

where $\omega(k_1, \dots, k_m, h_1, \dots, h_m) = (k_1 h + h_1, \dots, k_m h + h_m)$.

Notice that $(*)$ in (1) is divided into three terms by

$$\hat{\mathcal{A}}_\omega - \mathcal{A}_\omega = \underbrace{\hat{\mathcal{A}}_\omega - [\Theta \circ \sigma \circ \hat{\sigma}^{-1}]_\omega}_{(a)} + \underbrace{\mathcal{A}_\omega - \Theta_\omega}_{(b)} + \underbrace{[\Theta \circ \sigma \circ \hat{\sigma}^{-1}]_\omega - \Theta_\omega}_{(c)}$$

Therefore, we bound (a) , (b) , and (c) in $(*)$ respectively. Since $\mathbb{E}(\hat{\mathcal{A}}_\omega) = [\Theta \circ \sigma \circ \hat{\sigma}^{-1}]_\omega$ and $\mathbb{E}(\mathcal{A}_\omega) = \Theta_\omega$, Hoeffding's inequality, we bound (a) and (b) by

$$\begin{aligned} \mathbb{P} \left(\frac{1}{h^m} \left| \sum_{\substack{h_j \in \{0, \dots, h-1\} \\ j=1, \dots, m}} \hat{\mathcal{A}}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} - [\Theta \circ \sigma \circ \hat{\sigma}^{-1}]_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} \right| \geq \epsilon \right) &\leq 2 \exp(-2h^m \epsilon^2), \\ \mathbb{P} \left(\frac{1}{h^m} \left| \sum_{\substack{h_j \in \{0, \dots, h-1\} \\ j=1, \dots, m}} \mathcal{A}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} - \Theta_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} \right| \geq \epsilon \right) &\leq 2 \exp(-2h^m \epsilon^2), \end{aligned} \quad (2)$$

for any $\epsilon > 0$. For (c) term, notice that for any $\omega \in [d]^m$,

$$\begin{aligned} |[\Theta \circ \sigma \circ \hat{\sigma}^{-1}]_\omega - \Theta_\omega| &= |[\Theta \circ \sigma]_{\omega'} - [\Theta \circ \hat{\sigma}]_{\omega'}|, \quad \text{for some } \omega' \in [d]^m \\ &\leq \frac{L|\sigma(\omega') - \hat{\sigma}(\omega')|_1}{d} \\ &\lesssim \text{Loss}(\sigma, \hat{\sigma}). \end{aligned} \quad (3)$$

Combining (2), (3) and triangular inequality, for probability at least $1 - \exp(-h^m \epsilon^2)$,

$$(*) \lesssim \epsilon^2 + \text{Loss}^2(\sigma, \hat{\sigma}).$$

Going back to (1), we show that

$$h^m \sum_{\substack{k_i \in \{0, \dots, k-1\} \\ i=1, \dots, m}} \left(\frac{1}{h^m} \sum_{\substack{h_j \in \{0, \dots, h-1\} \\ j=1, \dots, m}} \hat{\mathcal{A}}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} - \mathcal{A}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} \right)^2 \lesssim d^m (\epsilon^2 + \text{Loss}^2(\sigma, \hat{\sigma})),$$

with probability at least, $1 - k^m \exp(-h^m \epsilon^2)$. \square

Lemma 3 (Denoising error (Expectation version)).

$$\mathbb{E} \|\text{Block}_k(\mathcal{Y} \circ \sigma^{-1}) - \text{Block}_k(\Theta)\|_F^2 \leq k^m.$$

Lemma 3* (Denoising error (Probability version)). With probability $1 - k^m \exp(-h^m \epsilon^2)$,

$$\|\text{Block}_k(\mathcal{Y} \circ \sigma^{-1}) - \text{Block}_k(\Theta)\|_F^2 \leq d^m \epsilon^2.$$

Proof. Remember that we define $\mathcal{A} = \mathcal{Y} \circ \sigma^{-1}$. By definition, we have

$$\begin{aligned} \|\text{Block}_k(\mathcal{Y} \circ \sigma^{-1}) - \text{Block}_k(\Theta)\|_F^2 = & \\ h^m \sum_{\substack{k_i \in \{0, \dots, k-1\} \\ i=1, \dots, m}} \left(\frac{1}{h^m} \sum_{\substack{h_j \in \{0, \dots, h-1\} \\ j=1, \dots, m}} \mathcal{A}_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} - \Theta_{\omega(k_1, \dots, k_m, h_1, \dots, h_m)} \right)^2, \end{aligned} \quad (4)$$

where $\omega(k_1, \dots, k_m, h_1, \dots, h_m) = (k_1 h + h_1, \dots, k_m h + h_m)$. Combining (2) and (4) yields,

$$\|\text{Block}_k(\mathcal{Y} \circ \sigma^{-1}) - \text{Block}_k(\Theta)\|_F^2 \lesssim d^m \epsilon^2,$$

with probability at least $1 - k^m \exp(-2h^m \epsilon^2)$. \square

Based on the changed lemmas, our main theorem becomes

Theorem 0.1 (Sorting-and-blocking under β -monotonicity of degree). With probability at least $1 - k^m \exp(-h^m \epsilon^2)$,

$$\begin{aligned} \|\hat{\Theta} - \Theta\|_F^2 &\lesssim d^m (\epsilon^2 + \text{Loss}^2(\sigma, \hat{\sigma})) + d^m \epsilon^2 + \frac{d^m}{k^2} \\ &\lesssim d^{-\beta(m-1)+m} + d^m \epsilon^2 + \frac{d^m}{k^2}. \end{aligned}$$

Furthermore, setting $\epsilon^2 = \frac{d}{h^m}$ yields,

$$\|\hat{\Theta} - \Theta\|_F^2 \lesssim d^{-\beta(m-1)+m} + dk^m + \frac{d^m}{k^2}$$

$$\lesssim d^{-\beta(m-1)+m} + d^{\frac{-2m+2}{m+2}+m},$$

with probability at least $1 - k^m \exp(-d)$. The last line comes from balancing the block size as $k = \mathcal{O}(d^{\frac{m-1}{m+2}})$.