§ Reparameterisation - Full Orthogonalisation

$$T = \bigcup_{E \in \Upsilon} \{ x \in E \mid \|x\|_2 = 1 \} \qquad \Upsilon = \bigcup_{(v,w) \in \tilde{W}} \operatorname{span}[B^{v,w}_{s,t}]_{s,t \in [R]}$$

$$\tilde{W} = \{ v^s \in B_{P_1}, w^t \in B_{P_2} \mid \langle v^s, v^r \rangle = 0 \text{ and } \langle w^s, w^r \rangle = 0 \ \forall r \in [R] \}$$

§ Bounding the Talagrand Functional and covering number

* $\rho_\Upsilon = \dim \operatorname{span}[B^{v,w}_{s,t}]_{s,t \in [R]} \leq n_1 R^2$

* Bounding $\gamma_2(V, \rho_{Fin}) = \inf_{\{\bar{V}_k\}_{k \geq 0}^{\infty}} \sup_{[B^{v,w}_{s,t}] \in \Upsilon} \rho_{Fin}([B^{v,w}_{s,t}], \bar{V}_k)$

The first step is to come up with a good perturbative bound based on Li et al's results. DONE!

First consider constructing an η_k-net

Following ideas from Proof of Theorem 5:

LEMMA [PERTURBATIVE BOUND] : Mode-3 Rank-R

Consider a tensor $\theta = \sum_{r=1}^{R} w^r \circ v^r \circ u^r$ where $w^r \in \mathbb{R}^{n_3}$, $v^r \in \mathbb{R}^{n_2}$, $u^r \in \mathbb{R}^{n_1}$ $\forall r \in [R]$

Consider $A^{vw}_{rq} \in \mathbb{R}^{n \times n_1}$

$$= \sum_{j \in [n_2]} \sum_{k \in [n_3]} A^{jk} v_j^r w_k^q$$

Full Orthogonalisation !

$$\mathcal{A} = \left[A^{vw}_{rq}\right]_{r,q \in [R]} \in \mathbb{R}^{n \times n_1 R^2} \rightsquigarrow \bar{\mathcal{A}} = \left[A^{\bar{v}\bar{w}}_{\bar{r}\bar{q}}\right]_{\bar{r},\bar{q} \in [R]}$$

Let's assume that for $r \in [R]$ $\exists$ $\bar{v}^r \in \mathbb{R}^{n_2}$ and $\bar{w}^r \in \mathbb{R}^{n_3}$ such that $\|\bar{v}^r - v^r\| \leq \eta_0$ and $\|\bar{w}^r - w^r\| \leq \eta_0$

Then, we have the following :

$$\rho_{Fin}\left(\text{span}(\mathcal{A}), \text{span}(\tilde{\mathcal{A}})\right) \leq R^{d-1}\left((1+\eta_0)^{d-1} - 1\right) \cdot \kappa(\mathcal{A})$$

Proof: Following Li et Al, we know that

$$\rho_{Fin}\left(\text{span}(\mathcal{A}), \text{span}(\tilde{\mathcal{A}})\right) \leq \frac{\|\mathcal{A} - \tilde{\mathcal{A}}\|_2 \rightarrow Nr \text{ I}}{\sigma_{min}(\mathcal{A}) \rightarrow Dr \text{ II}}$$

Upper Bounding Numerator I

$$\|\mathcal{A} - \tilde{\mathcal{A}}\|_2 = \left\| \left[\sum_{j,k} A^{jk} v_j^r w_k^q\right]_{r,q} \right\|_2$$

How good is this bound ?

$$\leq \sum_{r,q=1}^{R} \left\| \sum_{j,k} A^{jk} v_j^r w_k^q \right\|_2$$

Comment: $\| [A \,\vdots\, B] \|_2 = \sup_{\|(x,y)^T\|=1} \| [A \,\vdots\, B] (x,y)^T \|_2$ $\quad x \in \mathbb{R}^{n_1}, \; y \in \mathbb{R}^{n_2}$

$A \in \mathbb{R}^{n \times n_1}$, $B \in \mathbb{R}^{n \times n_2}$

$$= \sup_{\|(x,y)^T\|=1} \| Ax^T + By^T \|_2$$

$$\|(x,y)^T\|_2^2 = \|x\|_2^2 + \|y\|_2^2 = 1 \Rightarrow \|x\|_2^2, \|y\|_2^2 \leq 1$$

??

$$\leq \sup_{\|x\| \leq 1} \|Ax^T\| + \sup_{\|y\| \leq 1} \|By^T\|_2$$

$$\leq \|A\| + \|B\|$$

$$\|\mathcal{A} - \tilde{\mathcal{A}}\|_2 \leq \sum_{r,q=1}^{R} \Big\| \sum_{j,k} A^{jk} \left(v_j^r w_k^q - \bar{v}_j^r \bar{w}_k^r \right) \Big\|_2$$

$$\leq \sum_{r,q}^{R} \| A \cdot (v^r \circ w^q \circ \mathbb{1} - \bar{v}^r \circ \bar{w}^q \circ 1) \|_2$$

Reversing the unravelling step

This simplification follows from lemma 2!

$$\leq \underbrace{\sigma_{max}(A)}_{\text{C-F}} \sum_{r,q=1}^{R} \underbrace{\| v^r \circ w^q \circ \mathbb{1} - \bar{v}^r \circ \bar{w}^q \circ 1 \|_2}_{\| (v^r - \bar{v}^r) \circ w^q \circ \mathbb{1} + \bar{v}^r \circ (w^q - \bar{w}^q) \cdot \mathbb{1} \|_2}$$

Replicating Theorem in Paper gives us:

Generalised Result!

$$\| \theta_d \otimes \theta_{d-1} \otimes \cdots \theta_1 - \bar{\theta}_d \otimes \bar{\theta}_{d-1} \otimes \cdots \bar{\theta}_1 \|_2 \leq ?$$

If $\quad \| \theta_i - \bar{\theta}_i \|_2 \leq \eta_0 \quad \forall\, i \in [d]$

Cases : $d = 1$ $\qquad \leq \eta_0$

$d = 2$ $\quad \| \theta_2 \otimes \theta_1 - \bar{\theta}_2 \otimes \bar{\theta}_1 \|$

Does this give bad bounds?

$$\leq \| \theta_2 \otimes \theta_1 - \bar{\theta}_2 \otimes \theta_1 + \bar{\theta}_2 \otimes \theta_1 - \bar{\theta}_2 \otimes \bar{\theta}_1 \|_2$$

$$\leq \| (\theta_2 - \bar{\theta}_2) \otimes \theta_1 \|_2 + \| \bar{\theta}_2 \otimes (\theta_1 - \bar{\theta}_1) \|_2$$

$$\leq \eta + (1+\eta)\eta = \eta(2+\eta)$$

↳ Proof by Induction $\longrightarrow$ Gives better / worse bounds

Or is it the same :/

$$\| \theta_D \circ \theta_{D-1} \circ \cdots \theta_1 - \bar{\theta}_D \circ \bar{\theta}_{D-1} \circ \cdots \bar{\theta}_1 \|_2$$

$$\leq \| (\theta_D - \bar{\theta}_D) \circ \theta_{D-1} \circ \cdots \theta_1 \| + \| \bar{\theta}_D \circ \theta_{D-1} \circ \cdots \theta_1 - \bar{\theta}_D \circ \bar{\theta}_{D-1} \circ \cdots \bar{\theta}_1 \|$$

$$\leq \eta_0 + \| \bar{\theta}_D \| f(d-1) \leq \eta + (1+\eta) f(d-1)$$

SOLVE RECURSION :

$$f(1) = \eta$$

$$f(2) = \eta(2+\eta)$$

$$f(d) = \eta + (1+\eta) f(d-1)$$

$$f(3) = \eta + (1+\eta)\eta(2+\eta)$$

$$= \eta\left(1 + (1+\eta)(2+\eta)\right)$$

$$f(d) = \eta\left(1 + (1+\eta) \cdots (d-1+\eta)\right)$$

It's the same :c

$$= \boxed{(1+\eta)^D - 1}$$

§ UNDERSTANDING LEMMA 3

$$\| \theta_d \circ \theta_{d-1} \circ \cdots \theta_1 - \phi_d \circ \phi_{d-1} \circ \cdots \phi_1 \| = (1+\eta)^D - 1$$

Given $\forall\ i \in [d]$ $\| \theta_d - \phi_d \| \leq \eta$

Proceeding with Numerator Simplification:

$$\sigma_{max}(A) \sum_{r,q=1}^{R} \underbrace{\| v^r \circ w^q \circ \mathbb{1} - \bar{v}^r \circ \bar{w}^q \circ 1 \|_2}_{\| (v^r - \bar{v}^r) \circ w^q \circ \mathbb{1} + \bar{v}^r \circ (w^q - \bar{w}^q) \cdot \mathbb{1} \|_2}$$

C-F

$$\leq \sigma_{max}(A)\, R^2 \left((1+\eta)^{D-1} - 1 \right)$$

$\Rightarrow$ combining numerator and denominator

$$\delta_{Fin}\left(span(\mathcal{A}), span(\hat{\mathcal{A}}) \right) \leq R^{d-1} \left((1+\eta_0)^{D-1} - 1 \right) \cdot K(\mathcal{A})$$

Now, with perturbative bound, moving onto bounding Talagrand

$$\mathcal{A} = [A^{vw}_{rq}]_{r,q=1}^{R}$$

$$\mathcal{T} = \bigcup_{E \in \gamma} \{ x \in E \mid \|x\|_2 = 1 \} \qquad E = \bigcup_{\tilde{W}} [A^{vw}_{rq}]_{r,q=1}^{R}$$

$$\tilde{W} = \{ v^r \in B_{P_2}, w^r \in B_{P_3} \;\; r \in [R] \mid \langle v^r, v^q \rangle = \langle w^r, w^q \rangle = 0 \}$$

$$\gamma_2(\mathcal{V}, \rho_{Fin}) = \inf_{\{\bar{V}_k\}_{k\geq 0}^{\infty}} \sup_{[B_{s,t}^{v,w}] \in \gamma} \rho_{Fin}\left([B_{s,t}^{v,w}], \bar{V}_k\right)$$

Consider the following ε_k-net

For $v \in B_{n_2}$ and $B_{n_3} \ni w$ consider η_k nets such that

$\| v^{n} - \bar{v} \|_2 \leq \eta_k$ and $\| w^{n} - \bar{w} \|_2 \leq \eta_k$

$\Rightarrow$ Then consider the $\bar{\gamma}_k = \{ \underbrace{\qquad\qquad} \}$ collection.

$$\rho_{Fin}\left(A^{v,w}, A^{\bar{v},\bar{w}}\right) \leq \min\{1, K(A) R^2 ((1+\eta_k)^2 - 1)\}$$

Similar to sketch proof provided, consider decreasing η_k

Find smallest k' such that $K(A) R^2 ((1+\eta_k)^2 - 1) \leq 1$

$$\gamma_2(\gamma, \rho_{Fin}) = \sum_{k=0}^{k'} 2^{k/2} + R^2 2^{k'/2} \sum_{k=1}^{\infty} 2^{k/2} K(A) ((1+\eta_k)^2 - 1)$$

Bounding $\displaystyle\sum_{k=0}^{k'} 2^{k/2} \leq \frac{2^{k'/2}}{\sqrt{2}-1}$

$\underbrace{[v_1 \cdots\cdots v_r]}_{Rn} \quad [w_1 \cdots\cdots w_r]$

From results on ε_s: $\quad |\bar{\gamma}_k| \leq \boxed{\left(\frac{3}{\eta_k}\right)^{Rn_1 + Rn_2}} = \left(\frac{3}{\eta_k}\right)^{R(n_1+n_2)} \gtrsim 2^{2^k}$

$$\Rightarrow \quad R(n_1+n_2) \log\left(\frac{3}{\eta_k}\right) \leq 2^k$$

$$2^{k'/2} \leq \sqrt{R(n_1+n_2) \log\left(\frac{3}{\eta_k}\right)}$$

$$\sum_{k=0}^{k'} 2^{k/2} \lesssim \sqrt{R \log\left(\frac{1}{\eta_{k'}}\right)(n_1+n_2)}$$

$$R^2 d^{k'/2} \sum_{k=1}^{\infty} 2^{k/2} \frac{K(A)}{((1+\eta_k)^2-1)} \lesssim \ ?$$

For $k \geq k'$ How do you choose η_k ?

Binomial Expansion: $R^2 K(A)((1+\eta_k)^2-1) \lesssim 2R^2 K(A)\eta_k$

Take η_k and decrease it very slowly. choose k' such that

if we choose $\eta_{k'+1} = \frac{1}{4R^2 K(A)}$ s.t $\left(\frac{3}{\eta_{k'+1}}\right)^{R(n_1+n_2)} \leq 2^{2^{k'+1}}$

choose $\eta_{k+1} = \eta_k^2$

$$R^2 d^{k'/2} \sum_{k=1}^{\infty} 2^{k/2} \frac{K(A)}{((1+\eta_k)^2-1)}$$

$$\leq 2^{k'/2} \sum_{k=1}^{\infty} 2^{k/2} \underbrace{(2R^2 K(A)\eta_k)}_{1-Cacu}$$

$$\Sigma \lesssim \sqrt{R(n_1+n_2)\log\left(\frac{1}{\eta_{k'}}\right)}$$

$$\varepsilon_0 \leq 2R^2 K(A)\eta_{k'}$$

$$\frac{1}{\eta_{k'}} \leq \frac{2R^2 K(A)}{\varepsilon_0}$$

$$\boxed{\Sigma \lesssim \sqrt{R(n_1+n_2)\log\left(\frac{R^2 K(A)}{\varepsilon_0}\right)}}$$

$$\sum \lesssim \sqrt{R(n_1+n_2\cdots n_d)\,\log\left(\frac{(D-1)\,R^{D-1}\,\kappa(A)}{\varepsilon_0}\right)}$$

$$\gamma_2^{\,2}(\gamma_1 P_{Fin}) \lesssim R(n_1+n_2\cdots+n_d)\,\log\left(\frac{(D-1)\,R^{D-1}\,\kappa(A)}{\varepsilon_0}\right)$$

$$\mathcal{N}(\gamma, P_{Fin}, t) \leq \left(\frac{3}{\varepsilon_0}\right)^{\boxed{}}$$

connecting n_k and ε_k

$$\|\quad\| \leq n_k \quad \Rightarrow \quad \varepsilon_k = \left(2R^2\kappa(A)\left((1+n_k)^D-1\right)\right)$$

$$\left(\frac{3}{n_k}\right)^{Rn_1+Rn_2}$$

$$\sqrt[D]{\frac{\varepsilon_k}{2R^2\kappa(A)}+1} \; - 1 = n_k$$

$$\tilde{\alpha} =$$