Fix to LRZ-2019

Tengyuan Liang

1 Problem on LRZ-2019

We thank Andrea Montanari for pointing out a mistake in our proof. Below Eqn. 18, we use the following claim: If the symmetric PSD matrix satisfies

$$K = K^{\leq \iota} + K^{>\iota} = \Phi \Lambda \Phi^\top + K^{>\iota} ,$$

with $\Phi^\top \Phi = I_{\binom{n+\iota}{\iota}}$ and Λ being a diagonal matrix, and

$$K^{>\iota} \succeq \gamma \cdot I_n, \quad \text{with } \gamma > 0 ,$$

then for $v = \Phi \alpha \in \mathbb{R}^n$ that lies in the span of Φ,

$$v^\top K^{-2} v \leq \big(\lambda_{\min}(\Lambda)\big)^{-2} \|v\|^2 .$$

Unfortunately, this is not true in general. In the current note, we provide a fix to the claim. First, we will show that **(i)** the claim is true up to a multiplicative factor if assumed in addition

$$K^{>\iota} \preceq \kappa \cdot I_n, \quad \text{with } \kappa > 0 .$$

(ii) Second, we will prove why the above assumption is true for our problem.

2 Proof of (i)

For convenience, we define $M := K^{>\iota}$. Recall $K = \Phi \Lambda \Phi^\top + M$, and by assumption (which we will prove later)

$$\gamma \cdot I_n \preceq M \preceq \kappa \cdot I_n .$$

Now we have

$$\begin{aligned} K^{-1} v &= (\Phi \Lambda \Phi^\top + M)^{-1} \Phi \alpha \\ &= M^{-\frac{1}{2}} (M^{-\frac{1}{2}} \Phi \Lambda \Phi^\top M^{-\frac{1}{2}} + I_n)^{-1} M^{-\frac{1}{2}} \Phi \alpha \end{aligned}$$

Therefore

$$\begin{aligned} v^\top K^{-2} v &\leq \big(\lambda_{\min}(M)\big)^{-1} \| (M^{-\frac{1}{2}} \Phi \Lambda \Phi^\top M^{-\frac{1}{2}} + I_n)^{-1} M^{-\frac{1}{2}} \Phi \alpha \|^2 \\ &\leq \gamma^{-1} \cdot \| \underbrace{(M^{-\frac{1}{2}} \Phi \Lambda \Phi^\top M^{-\frac{1}{2}} + I_n)^{-1} M^{-\frac{1}{2}} \Phi \Lambda^{\frac{1}{2}}}_{:=T} \cdot \Lambda^{-\frac{1}{2}} \alpha \|^2 \\ &\leq \gamma^{-1} \lambda_{\max}(T^\top T) \cdot \| \Lambda^{-\frac{1}{2}} \alpha \|^2 . \end{aligned}$$

It is clear that if $\lambda_0 := \lambda_{\min}(M^{-\frac{1}{2}}\Phi\Lambda\Phi^\top M^{-\frac{1}{2}}) > 1$

$$\lambda_{\max}(T^\top T) = \frac{\lambda_0}{(1+\lambda_0)^2} < \lambda_0^{-1} .$$

To lower bound λ_0, we invoke the upper bound on $M \preceq \kappa \cdot I_n$

$$\lambda_0 \geq \kappa^{-1}\lambda_{\min}(\Lambda) \asymp \frac{n}{d^\iota} \gg 1 .$$

Put things together, we have proved that

$$\begin{aligned} v^\top K^{-2} v &\leq \frac{\kappa}{\gamma}(\lambda_{\min}(\Lambda))^{-1}\|\Lambda^{-\frac{1}{2}}\alpha\|^2 \\ &= \frac{\kappa}{\gamma}(\lambda_{\min}(\Lambda))^{-1} \cdot v^\top (K^{\leq \iota})^+ v . \end{aligned}$$

Therefore the problem is fixed with a multiplicative factor $\frac{\kappa}{\gamma}$. In the next section, we will show an upper bound on $\frac{\kappa}{\gamma}$. For the problem in LRZ-2019, by means of the restricted lower isometry, we have $(\lambda_{\min}(\Lambda))^{-1} \precsim \frac{d^\iota}{n}$, and $v^\top (K^{\leq\iota})^+ v = O(1)$.

3 Proof of (ii)

In LRZ-2019, we have already proved

$$K^{>\iota} = K^{(i,2i+1]} + K^{>2\iota+1} \succeq K^{>2\iota+1}$$

and $K^{>2\iota+1}$ is a diagonally dominate matrix that satisfies

$$\gamma \cdot I_n \preceq K^{>2\iota+1} \preceq 2\gamma \cdot I_n .$$

with a constant $\gamma > 0$.

To establish an upper bound on $\|K^{>\iota}\|_{\mathrm{op}}$, we only need to control

$$\|K^{(\iota,2\iota+1]}\|_{\mathrm{op}} .$$

Recall the feature map for the inner product kernel, $\phi_{(\iota,2\iota+1)}(x_j) \in \mathbb{R}^{\binom{d+2\iota+1}{2\iota+1}-\binom{d+\iota}{\iota}}$

$$K^{(i,2i+1]} = [\langle \phi_{(\iota,2\iota+1)}(x_j), \phi_{(\iota,2\iota+1)}(x_k)\rangle]_{1\leq j,k\leq n} .$$

Then bounding the operator norm is the same as bounding the following operator norm

$$\left\| \sum_{j=1}^n \phi_{(\iota,2\iota+1)}(x_j)\phi_{(\iota,2\iota+1)}(x_j)^\top \right\|_{\mathrm{op}} .$$

By the matrix Bernstein's inequality, we have with high at least $1 - d^{-C}$,

$$\left\| \sum_{j=1}^n \phi_{(\iota,2\iota+1)}(x_j)\phi_{(\iota,2\iota+1)}(x_j)^\top - n\,\mathbb{E}[\phi_{(\iota,2\iota+1)}(\mathbf{x})\phi_{(\iota,2\iota+1)}(\mathbf{x})^\top] \right\|_{\mathrm{op}} \precsim \sqrt{n \cdot \mathbf{V}\log(d)} \vee \mathbf{B}\log(d)$$

where

$$\begin{aligned}
\mathbf{V} &= \| \mathbb{E}[\phi_{(\iota,2\iota+1)}(\mathbf{x})\phi_{(\iota,2\iota+1)}(\mathbf{x})^\top \phi_{(\iota,2\iota+1)}(\mathbf{x})\phi_{(\iota,2\iota+1)}(\mathbf{x})^\top] \|_{\mathrm{op}} \le \mathbf{B} \cdot d^{-\iota-1} \ , \\
\mathbf{B} &= \sup_x \|\phi_{(\iota,2\iota+1)}(x)\phi_{(\iota,2\iota+1)}(x)^\top\|_{\mathrm{op}} \ .
\end{aligned}$$

Under the assumption $\sup_x K(x,x) \le C$, $\mathbf{B} \le C$, we have

$$\left\| \sum_{j=1}^n \phi_{(\iota,2\iota+1)}(x_j)\phi_{(\iota,2\iota+1)}(x_j)^\top - n\,\mathbb{E}[\phi_{(\iota,2\iota+1)}(\mathbf{x})\phi_{(\iota,2\iota+1)}(\mathbf{x})^\top] \right\|_{\mathrm{op}} \precsim \sqrt{\frac{n}{d^{\iota+1}}\log(d)} + \log(d)$$

and thus

$$\|K^{(\iota,2\iota+1]}\|_{\mathrm{op}} = \left\| \sum_{j=1}^n \phi_{(\iota,2\iota+1)}(x_j)\phi_{(\iota,2\iota+1)}(x_j)^\top \right\|_{\mathrm{op}} \precsim \frac{n}{d^{\iota+1}} + \sqrt{\frac{n}{d^{\iota+1}}\log(d)} + \log(d) \asymp \log(d) \ ,$$

where the last step uses $d^\iota \ll n \ll d^{\iota+1}$.

So far, we have proved the

$$\kappa \precsim \log(d) \ .$$

Put things together, the variance bound in LRZ-2019 holds true with the following expression

$$\log(d) \cdot \frac{d^\iota}{n} + \frac{n}{d^{\iota+1}} \ .$$