Variance Bound

Abstract

Define the basic cost estimator for a solution $\mathcal{S}$

$$E_{\mathcal{S}} := \frac{1}{|\Omega|} \sum_{p \in \Omega} \frac{\text{cost}(\mathcal{A})}{\text{cost}(p, \mathcal{A})} \text{cost}(p, \mathcal{S}).$$

It's expectation is $\text{cost}(\mathcal{S})$. We would like to show that the maximum error is less than $\varepsilon \cdot \text{cost}(\mathcal{S})$ factor by considering the following expectation:

$$\mathbb{E}_{\Omega} \sup_{\mathcal{S}} \left[\frac{|E_{\mathcal{S}} - \text{cost}(\mathcal{S})|}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})} \right]$$

Using the symmetrization argument, we have

$$\mathbb{E}_{\Omega} \sup_{\mathcal{S}} \left[\frac{|E_{\mathcal{S}} - \text{cost}(\mathcal{S})|}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})} \right] \leq O(1) \cdot \mathbb{E}_{\Omega} \mathbb{E}_g \sup_{\mathcal{S}} \left[\frac{\frac{1}{|\Omega|} \sum_{p \in \Omega} \frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})} \text{cost}(p, \mathcal{S}) \cdot g_p}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})} \right]$$

The way we analyse this so far is to first only condition on Ω, with whatever properties it might have, and essentially only use the randomness of g to bound the supremum:

$$\mathbb{E}_{\Omega} \mathbb{E}_g \sup_{\mathcal{S}} \left[\frac{\frac{1}{|\Omega|} \sum_{p \in \Omega} \frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})} \text{cost}(p, \mathcal{S}) \cdot g_p}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})} \right] = \mathbb{E}_{\Omega} \mathbb{E}_g \sup_{\mathcal{S}} \left[\frac{\frac{1}{|\Omega|} \sum_{p \in \Omega} \frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})} \text{cost}(p, \mathcal{S}) \cdot g_p}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})} \;\middle|\; \Omega \right]$$

Whether or not we want to do chaining, we will eventually have to bound the variance of $\frac{\frac{1}{|\Omega|} \sum_{p \in \Omega} \frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})} \text{cost}(p,\mathcal{S}) \cdot g_p}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})}$. Since each is a Gaussian, the sum is a Gaussian with variance

$$\sum_{p \in \Omega} \left(\frac{\text{cost}(\mathcal{A})}{\text{cost}(p, \mathcal{A}) \cdot |\Omega|} \cdot \frac{\text{cost}(p, \mathcal{S})}{\text{cost}(\mathcal{A}) + \text{cost}(\mathcal{S})} \right)^2$$

Here, I am no going to give a variance bound, assuming the following

1. All points in a cluster cost the same, up to constants
2. All clusters cost the same, up to constants
3. Conditioned on Ω, we sample $\sum_{p \in \Omega \cap C} \frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A}) \cdot |\Omega|} = O(|C|)$ for all clusters C of $\mathcal{A}$.
4. All clusters cost in $\mathcal{S}$ roughly 2^i times their cost in $\mathcal{A}$.

Define the clusters that satisfy condition 4 to be L_i and let $|Li| = \alpha \cdot k$ be the number of clusters that satisfy the fourth condition.

$$\begin{aligned}
& \sum_{p\in\Omega}\left(\frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})\cdot|\Omega|}\cdot\frac{\text{cost}(p,\mathcal{S})}{\text{cost}(\mathcal{A})+\text{cost}(\mathcal{S})}\right)^2 \\
\left(\frac{\text{cost}(p,\mathcal{S})}{\text{cost}(p,\mathcal{A})}\approx 2^i\right) \quad \leq \quad & \frac{1}{|\Omega|}\text{cost}(\mathcal{A})\cdot 2^i\cdot\left(\frac{1}{\text{cost}(\mathcal{A})+\text{cost}(\mathcal{S})}\right)^2\sum_{p\in\Omega}\frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})\cdot|\Omega|}\cdot\text{cost}(p,\mathcal{S}) \\
\text{Uniform Cost} \quad \leq \quad & \frac{1}{|\Omega|}\text{cost}(\mathcal{A})\cdot 2^i\cdot\left(\frac{1}{\text{cost}(\mathcal{A})+\text{cost}(\mathcal{S})}\right)^2\sum_{C\in L_i}\frac{\text{cost}(C,\mathcal{S})}{|C|}\sum_{p\in\Omega\cap C}\frac{\text{cost}(\mathcal{A})}{\text{cost}(p,\mathcal{A})\cdot|\Omega|} \\
\text{Condition on } \Omega \quad \leq \quad & \frac{1}{|\Omega|}\text{cost}(\mathcal{A})\cdot 2^i\cdot\left(\frac{1}{\text{cost}(\mathcal{A})+\text{cost}(\mathcal{S})}\right)^2\sum_{C\in L_i}\frac{\text{cost}(C,\mathcal{S})}{|C|}\cdot|C| \\
\leq \quad & \frac{1}{|\Omega|}\text{cost}(\mathcal{A})\cdot 2^i\cdot\left(\frac{1}{\text{cost}(\mathcal{A})+\text{cost}(\mathcal{S})}\right)^2\text{cost}(\mathcal{S})
\end{aligned}$$

This gives us a bound of $\frac{2^i}{|\Omega|} \leq \frac{\varepsilon^{-z}}{|\Omega|}$, if we ignore constants. Note that it doesn't depend on the space, however I doubt that using the space would give us much.

To get the alternative bound of $\frac{k}{|\Omega|}$, we can observe that $\alpha \geq 1/k$, as at least one cluster is in L_i, otherwise we wouldn't be considering it. Then using $\text{cost}(\mathcal{S}) \approx \alpha\cdot 2^i\cdot\mathcal{A}$

$$\begin{aligned}
& \text{cost}(\mathcal{A})\cdot 2^i\cdot\left(\frac{1}{\text{cost}(\mathcal{A})+\text{cost}(\mathcal{S})}\right)^2\text{cost}(\mathcal{S}) \\
\leq \quad & \text{cost}^2(\mathcal{A})\cdot 2^i\cdot\frac{1}{\text{cost}^2(\mathcal{A})(1+\alpha 2^i)^2}\cdot\alpha\cdot 2^i \\
\leq \quad & \frac{\alpha\cdot 2^{2i}}{\alpha^2\cdot 2^{2i}} \leq \frac{1}{\alpha} \leq k
\end{aligned}$$

All of this is without the variance reduction $-q$ trick. But maybe this gives a bit of an idea what one could play with. Unless we modify the estimator (for example using the q vector), this will cost us either an additional k or and additional ε^{-z} factor when doing chaining, assuming we can build the nets. In this case, building the nets is also doable (but also completely disjoint from obtaining the variance bound).