

Review report for Minimax Rates for Estimating the Dimension of a Manifold, Section 3

Jining Qin

February 19, 2016

1 Summary

The estimator is used to decide whether the intrinsic dimensionality of supporting manifold is d_1 or d_2 . It decides the dimension is d_1 if its d_1 -squared length of traveling salesman path is smaller or equal to a threshold. We want to establish an upper bound for the minimax error rate for this estimator.

First we look for the supremum in class P , since it is a natural upper bound of the minimax rate. Lemma 5 shows that if the intrinsic dimension is d_2 the estimator isn't very likely to make mistake, since the TSP length isn't likely to be bounded. Lemma 6 shows that if the intrinsic dimension is d_1 the estimator can always be correct, since under that case the TSP length is shown to be bounded by the threshold. Combining them we know the estimator is always right in d_1 dimension and has an upper bound of error rate under d_2 dimension.

2 General comment

I think the writing is pretty clear (if my summary above isn't obviously wrong somewhere). One suggestion I have is to not mention proposition 7, but use something like 'we want to show this estimator this and that (using less specific language than the exact statement of proposition 7), we achieve this following the steps described below'. And then put a paragraph on the line of 'Putting the above results together, we have blablabla property for the estimator'. The flow of logic might be more fluent if you make this small adjustment. But in general I think you gave the readers a clear big picture. Although I don't understand any of the proof details, I think I get what you are doing in each step and vaguely how the previous lemmas help the final proposition.

3 Specific comments

Paragraph one: should be a space after \hat{dim}_n and should be one fewer 'since' in the end.

Is $d_1 < d_2$ or the other way around? I think there should be a relation between them (and it might be mentioned somewhere else). It doesn't hurt to mention it here, I think.

The sentence 'Choosing an appropriate estimator is critical to get a good bound.' should be move to one paragraph later when you start talking about the estimator. It seems to serve no particular purpose here.

As I mentioned in general comments, I think there should be a summarizing paragraph before or after proposition 7 summarizing briefly how the pieces fit together to prove this result.