Brief papers

Reply to "Reply to 'Determining structural identifiability of parameter learning machines'"

CrossMark

Zhi-Yong Ran [a,*], Bao-Gang Hu [b]

[a] *Chongqing Key Laboratory of Computational Intelligence, School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China*
[b] *NLPR&LIAMA, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China*

In the paper "Reply to 'Determining structural identifiability of parameter learning machines'" [1], the author presented a critical comment on the results on global identifiability and parameter redundancy in our previous publication [2]. Based on this comment [1], we summarize the main points addressed by Cole, and give the answer to each point below.

Point 1: Global identifiability result in Theorem 4 in [2].

Answer: Many thanks for pointing out the incorrect part in the theorem. We agree with Cole that Theorem 4 is for locally identifiable rather than for globally identifiable. After a careful check of the entire theoretical development concerning Theorem 4 in [2], we see that, the quantity $\mathbf{H}_{ab}(\theta)$, in Eq. (52) of [2], should be $\mathbf{H}_{ab}(\theta_0)$, so that this result is merely a criterion for checking local identifiability with respect to θ_0.

Point 2: Theorem 6 in [2] is identical to Theorem 2a in [3] but published 4 years earlier.

Answer: The difference of Theorem 6 in [2] between Theorem 2a in [3] can be recognized from the following aspects:

- Following the naming convention in statistics literature [4], Definition 4 in [3] means that

$$s(\theta_1, data) = s(\theta_2, data) \Rightarrow M(\theta_1) = M(\theta_2); \tag{1}$$

this leads to a *sufficient partition* in parameter space. Nevertheless, Definition 3 in [2] means that

$$s(\theta_1) = s(\theta_2) \Leftrightarrow M(\theta_1) = M(\theta_2); \tag{2}$$

this leads to an *identifying partition*. Hence, the **formal** definitions of "exhaustive summary" in [2] and that in [3] are

fundamentally different concepts. For more details, one can see [4] and the references therein.

- From Definition 3 in [2], the exhaustive summary is *solely* the function of unknown parameter θ. However, the exhaustive summary used in [3] and the counterexample 1 in [1] consist of parameter θ and *data*. Therefore, this also indicates the difference.

- As a matter of fact, if exhaustive summary is defined as in [3], one cannot derive a *sufficient and necessary* result; this is just the reason why the complete proof is not provided in [3]. Nevertheless, we presented a mathematically rigorous proof in [2].

Point 3: Global identifiability result in Theorem 1 [2].

Answer: It should be noticed that, the authors placed the additional assumption that "[θ] is a smooth manifold of $\mathbb{R}^{kn}$" in Theorem 1 in [2], and then derived the result. This assumption means that the input-output mapping is unvaried along a smooth curve in parameter space (see [6] for more neuromanifold examples). Note that this assumption also appeared in [5] (see page. 3383). Empirically, this condition occurs if the model contains coupled parameters (see [2,7,8]). However, the counterexample 1 given in [1] does not meet this requirement, since the observationally equivalent parameters are two isolated points, namely (a,b) and $(a, -b)$; these two points cannot form a smooth curve.

Acknowledgement

References

[1] D.J. Cole, Reply to determining structural identifiability of parameter learning machines, Neurocomputing 173 (2016) 2039–2040.

[2] Z.-Y. Ran, B.-G. Hu, Determining structural identifiability of parameter learning machines, Neurocomputing 127 (2014) 88–97.

[3] D.J. Cole, B.J.T. Morgan, D.M. Titterington, Determining the parametric structure of models, Math. Biosci. 228 (2010) 16–30.

[4] C.D.M. Paulino, C.A.D.B. Pereira, On identifiability of parametric statistical models, J. Ital. Stat. Soc. 3 (1994) 125–151.

[5] A. Dasgupta, S.G. Self, S.D. Gupta, Nonidentifiable parametric probability models and reparameterization, J. Stat. Plan. Inference 137 (2007) 3380–3393.

[6] S.I. Amari, H. Park, T. Ozeki, Singularities affect dynamics of learning in neuromanifolds, Neural Comput. 18 (5) (2006) 1007–1065.

[7] S.-H. Yang, B.-G. Hu, P.H. Cournde, Structural identifiability of generalized-constraint neural network models for nonlinear regression, Neurocomputing 72 (2008) 392–400.

[8] B.-G. Hu, H.B. Qu, S.H. Yang, A generalized-constraint neural network model: associating partially known relationships for nonlinear regressions, Inf. Sci. 179 (2009) 1929–1943.

Bao-Gang Hu (M'94-SM'99) received the M.Sc. degree from the University of Science and Technology, Beijing, China, and the pH.D. degree from McMaster University, Hamilton, ON, Canada, both in mechanical engineering, in 1983 and 1993, respectively. He was a Research Engineer and Senior Research Engineer at C-CORE, Memorial University of Newfoundland, St. Johns, NF, Canada, from 1994 to 1997. From 2000 to 2005, he was the Chinese Director of computer science, control, and applied mathematics with the Chinese-French Joint Laboratory, National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences, Beijing, China. He is currently a Professor at NLPR. His current research interests include pattern recognition and plate growth modeling.

Zhi-Yong Ran received the M.Sc. degree in applied mathematics from the Beijing University of Technology, Beijing, China, and the pH.D. degree from the National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2014. Currently he is a lecturer with the School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China. His research interests include system identification, machine learning and pattern recognition.