Econometric Theory, **19**, 2003, 1123–1127. Printed in the United States of America.
DOI: 10.1017/S0266466603196090

AN EQUIVALENCE RESULT FOR *VC* CLASSES OF SETS

Scott Joslin
Stanford University

Robert P. Sherman
California Institute of Technology

Let $\mathcal{R}$ and Θ be infinite sets and let $A \subseteq \mathcal{R} \times \Theta$. We show that the class of projections of A onto $\mathcal{R}$ is a Vapnik–Chervonenkis (VC) class of sets if and only if the class of projections of A onto Θ is a VC class. We illustrate the result in the context of semiparametric estimation of a transformation model. In this application, the VC property is hard to establish for the projection class of interest but easy to establish for the other projection class.

1. INTRODUCTION

In the course of establishing the uniformity results (uniform laws of large numbers, stochastic equicontinuity) used to establish the limiting behavior of econometric estimators, it is sometimes useful to show that certain classes of sets are Vapnik–Chervonenkis, or *VC*, classes (see, e.g., Pollard, 1984, 1989; Andrews, 1994; and references therein). This is particularly true for estimators that optimize averages or generalized averages of indicator functions of sets involving finite-dimensional parameters. Examples include the maximum score estimators of Manski (1975, 1985) and the rank estimators of Han (1987a, 1987b), Cavanagh and Sherman (1998), Abrevaya (1999, 2002), Khan (2001), Chen (2002), and Asparouhova, Asparouhov, Golanski, Kasprzyk, and Sherman (2002). This paper establishes an equivalence result for *VC* classes of sets that can be used in such settings. This equivalence result had been previously established by van den Dries (1998, Ch. 5, Prop. 2.10) and was discovered by a referee and brought to the attention of the authors after this paper had been accepted for publication.

Let $\mathcal{R}$ and Θ be infinite sets and let A be a subset of the product space $\mathcal{R} \times \Theta$. For each $\theta \in \Theta$, define $\Pi_{\mathcal{R}}(A|\theta) = \{r \in \mathcal{R} : (r,\theta) \in A\}$. We call $\Pi_{\mathcal{R}}(A|\theta)$ the projection of A onto $\mathcal{R}$ given θ. Define the class of projections $\Pi_{\mathcal{R}}(A) = \{\Pi_{\mathcal{R}}(A|\theta) : \theta \in \Theta\}$. Similarly, for each $r \in \mathcal{R}$, define $\Pi_{\Theta}(A|r) = \{\theta \in \Theta : (r,\theta) \in A\}$ and $\Pi_{\Theta}(A) = \{\Pi_{\Theta}(A|r) : r \in \mathcal{R}\}$.

Address correspondence to: Scott Joslin, Stanford Graduate School of Business, 518 Memorial Way, Stanford University, Stanford, CA 94305-5015, USA; e-mail: sherman@hss.caltech.edu.

In a typical econometric application, $\mathcal{R}$ is the range space of the random variables in the econometric model, Θ is a finite-dimensional parameter space, and A has the form $\{(r,\theta) \in \mathcal{R} \times \Theta : h(r,\theta) > 0\}$ where h is a real-valued function on $\mathcal{R} \times \Theta$. Then $\Pi_{\mathcal{R}}(A)$ is the class of positivity sets $\{\{r \in \mathcal{R} : h(r,\theta) > 0\} : \theta \in \Theta\}$. The objective is to show that $\Pi_{\mathcal{R}}(A)$ is a *VC* class of sets. If the function h is linear in the components of θ, then the *VC* property can be immediately deduced from a standard result in the empirical process literature (see, e.g., Pakes and Pollard, 1989, Lemma 2.4). However, if h is nonlinear in the components of θ, then this result does not apply, and one must argue from first principles to show that $\Pi_{\mathcal{R}}(A)$ is a *VC* class.

In the next section, we show that no matter what form the set A has, $\Pi_{\mathcal{R}}(A)$ is a *VC* class of sets if and only if $\Pi_{\Theta}(A)$ is a *VC* class of sets. We illustrate the usefulness of the result in the context of a semiparametric transformation model involving a function h that is nonlinear in parameters. In this example, it is difficult to show that $\Pi_{\mathcal{R}}(A)$ is a *VC* class but easy to show that $\Pi_{\Theta}(A)$ is a *VC* class.

2. EQUIVALENCE RESULT

We begin by recalling the definition of a *VC* class of sets. A class of subsets $\mathcal{D}$ of a set S *shatters* a set S_0 of V points in S if $\{D \cap S_0 : D \in \mathcal{D}\} = \mathcal{P}(S_0)$, the power set of S_0. In other words, $\mathcal{D}$ shatters S_0 if it picks out all 2^V subsets of S_0. The class $\mathcal{D}$ is a *VC* class of sets if there exists a $V < \infty$ such that $\mathcal{D}$ can shatter no V point set $S_0 \subseteq S$ (see, e.g., Pollard, 1984, Ch. 2). The following result was proved by van den Dries in 1998.

THEOREM 1. $\Pi_{\mathcal{R}}(A)$ *is a VC class of sets if and only if* $\Pi_{\Theta}(A)$ *is a VC class of sets.*

Proof. By symmetry, it is enough to show that $\Pi_{\Theta}(A)$ not *VC* implies $\Pi_{\mathcal{R}}(A)$ not *VC*. So, suppose that $\Pi_{\Theta}(A)$ is not a *VC* class of sets. We must show that for each $n \geq 1$, there exists an n point set $\mathcal{R}_0 \subseteq \mathcal{R}$ that $\Pi_{\mathcal{R}}(A)$ shatters.

Fix $n \geq 1$. Since $\Pi_{\Theta}(A)$ is not *VC*, it can shatter some V point set for each $V \geq 1$. In particular, there exists a 2^n point set $\Theta_0 = \{\theta_i : 1 \leq i \leq 2^n\} \subseteq \Theta$ that $\Pi_{\Theta}(A)$ shatters. This implies that there exists a 2^{2^n} point set $\mathcal{R}_1 = \{r_j : 1 \leq j \leq 2^{2^n}\} \subseteq \mathcal{R}$ such that $\{\Pi_{\Theta}(A|r_j) \cap \Theta_0 : r_j \in \mathcal{R}_1\} = \mathcal{P}(\Theta_0)$. Define a $2^n \times 2^{2^n}$ matrix M_n such that for $i = 1,\ldots,2^n$ and $j = 1,\ldots,2^{2^n}$,

$$M_n^{ij} = \begin{cases} 1 & \text{if } (r_j,\theta_i) \in A \\ 0 & \text{if } (r_j,\theta_i) \notin A. \end{cases}$$

Thus, the jth column of M_n corresponds to $\Pi_{\Theta}(A|r_j) \cap \Theta_0$, the subset of Θ_0 that $\Pi_{\Theta}(A|r_j)$ picks out. For example, take $n = 2$ and order the r_j's so that

$$M_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

The ninth column of M_2 is $[0,1,1,0]'$, and so $\Pi_\Theta(A|r_9) \cap \Theta_0 = \{\theta_2, \theta_3\}$.

The key insight is that the ith row of M_n corresponds to $\Pi_{\mathcal{R}}(A|\theta_i) \cap \mathcal{R}_1$, the subset of $\mathcal{R}_1$ that $\Pi_{\mathcal{R}}(A|\theta_i)$ picks out. We want to select $\mathcal{R}_0 \subseteq \mathcal{R}_1$ such that $\mathcal{R}_0$ is an n point set and $\{\Pi_{\mathcal{R}}(A|\theta_i) \cap \mathcal{R}_0 : \theta_i \in \Theta_0\} = \mathcal{P}(\mathcal{R}_0)$. This will prove the result.

For $n \geq 1$, define P_n to be the $2^n \times n$ matrix of 0's and 1's such that the rows of P_n correspond to the 2^n subsets of an n element set. For example,

$$P_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Note that the columns of both M_n and P_n have 2^n elements. Moreover, by construction, M_n consists of all possible columns of 2^n elements of 0's and 1's. In particular, the columns of P_n must appear as columns of M_n, say, as columns $j_1, j_2, \ldots, j_n$. Let $\mathcal{R}_0 = \{r_{j_1}, r_{j_2}, \ldots, r_{j_n}\}$. From the previous discussion, $\{\Pi_{\mathcal{R}}(A|\theta_i) \cap \mathcal{R}_0 : \theta_i \in \Theta_0\} = \mathcal{P}(\mathcal{R}_0)$. ∎

Remark. The *VC* dimension of a class of subsets $\mathcal{D}$ of a set S is the largest V for which some set of V points in S is shattered by $\mathcal{D}$. Fix $n \geq 1$. We see from the proof of Theorem 1 that if $\Pi_\Theta(A)$ has *VC* dimension at least 2^n, then $\Pi_{\mathcal{R}}(A)$ has *VC* dimension at least n. See Wenocur and Dudley (1981) and Stengle and Yukich (1989) for complementary results on *VC* classes of sets.[1]

We apply Theorem 1 to an example in Asparouhova et al. (2002). These authors state sufficient conditions for $\sqrt{n}$-consistency and asymptotic normality of rank estimator of Han (1987a) of the parameter vector characterizing the transformation in a semiparametric transformation model. They also state sufficient conditions for $\sqrt{n}$-consistency and asymptotic normality of a new class of rank estimators of this parameter vector.

Their key assumption, A7, is that a certain class of sets is a *VC* class. In the special case of the Box–Cox (1964) transformation, this class of sets has the form $\Pi_{\mathcal{R}}(A)$, where $\mathcal{R} = \mathbb{R}^4_+$, $\Theta = \mathbb{R}_+$, and $A = \{(r,\theta) \in \mathcal{R} \times \Theta : h(r,\theta) > 0\}$ with $r = (y_1, y_2, y_3, y_4)$ and $h(r,\theta) = y_1^\theta - y_2^\theta - y_3^\theta + y_4^\theta$. The nonlinearity of h as a function of θ makes it difficult to prove directly that

$\Pi_{\mathcal{R}}(A)$ is a *VC* class. Because of this nonlinearity, Lemma 2.4 in Pakes and Pollard (1989) does not apply.

However, showing that $\Pi_{\Theta}(A)$ is a *VC* class is easy. Fix $r \in \mathcal{R}$. Simple calculus shows that $h(r,\cdot)$ has at most three zeros. It follows that $\Pi_{\Theta}(A|r)$ is a union of at most two intervals on Θ. Since the set of all intervals on Θ is a *VC* class, and, for each $k \geq 1$, the set of all unions of k sets from a *VC* class is a *VC* class (see, e.g., Pollard, 1989), it follows that $\Pi_{\Theta}(A)$ is a *VC* class. Deduce from Theorem 1 that $\Pi_{\mathcal{R}}(A)$ is a *VC* class.

Finally, we note that Lemma 1 in Asparouhova et al. (2002) is used to establish that $\Pi_{\mathcal{R}}(A)$ is a *VC* class of sets for the Box–Cox (1964) example. Though the respective proofs are quite different, their Lemma 1 can be viewed as a special case of Theorem 1. Their Lemma 1 requires that (i) the parameter space Θ be a subset of the real line and (ii) the set A have the form $\{(r,\theta) \in \mathcal{R} \times \Theta : h(r,\theta) > 0\}$. Theorem 1, on the other hand, applies to an arbitrary parameter space (e.g., it could be infinite-dimensional) and does not require that the set A have the form specified in (ii); A can be any subset of the product space $\mathcal{R} \times \Theta$.

NOTE

1. We thank an anonymous referee for this remark about *VC* dimension and for these references.

REFERENCES

Abrevaya, J. (1999) Leapfrog estimation of a fixed-effects model with unknown transformation of the dependent variable. *Journal of Econometrics* 93, 203–228.

Abrevaya, J. (2002) Pairwise-Difference Rank Estimation of the Transformation Model. Manuscript, University of Chicago Graduate School of Business.

Andrews, D.W. (1994) Empirical process methods in econometrics. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. 4, pp. 2247–2294. Amsterdam: North-Holland.

Asparouhova, E., T. Asparouhov, R. Golanski, K. Kasprzyk, & R.P. Sherman (2002) Rank estimators for a transformation model. *Econometric Theory* 18, 1099–1120.

Box, G.E.P. & D.R. Cox (1964) An analysis of transformations. *Journal of the Royal Statistical Society, Series B* 34, 187–220.

Cavanagh, C.L. & R.P. Sherman (1998) Rank estimators for monotonic index models. *Journal of Econometrics* 84, 351–381.

Chen, S. (2002) Rank estimation of transformation models. *Econometrica* 70, 1683–1697.

Han, A.K. (1987a) A nonparametric analysis of transformations. *Journal of Econometrics* 35, 191–209.

Han, A.K. (1987b) Non-parametric analysis of a generalized regression model. *Journal of Econometrics* 35, 303–316.

Khan, S. (2001) Two stage rank estimation of quantile index models. *Journal of Econometrics* 100, 319–355.

Manski, C. (1975) Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics* 3, 205–228.

Manski, C. (1985) Semiparametric analysis of discrete response: Asymptotic properties of the maximum score estimator. *Journal of Econometrics* 27, 313–334.

Pakes, A. & D. Pollard (1989) Simulation and the asymptotics of optimization estimators. *Econometrica* 57, 1027–1057.

Pollard, D. (1984) *Convergence of Stochastic Processes*. New York: Springer-Verlag.
Pollard, D. (1989) Asymptotics via empirical processes. *Statistical Science* 4, 341–366.
Stengle, G. & J.E. Yukich (1989) Some new Vapnik–Chervonenkis classes. *Annals of Statistics* 17, 1441–1446.
van den Dries, L. (1998) *Tame Topology and O-minimal Structures*. Cambridge: Cambridge University Press.
Wenocur, R.S. & R.M. Dudley (1981) Some special Vapnik–Chervonenkis classes. *Discrete Mathematics* 33, 313–318.