Definition 2.5

Let $p_\theta \in \mathcal{P}$. Define the set of compound sequence matrices of p_θ relative to $\varepsilon \in \mathbf{r}$ as:

$$\begin{aligned} C(\theta,\varepsilon) \ &:= \ \{M;\ m \in N,\ M \in I\!R^{m\times m}, \\ &\quad M_{ij} := p_\theta(s_i \varepsilon t_j)\ \ s_i, t_j \in \mathbf{r}^*\}. \end{aligned}$$

Definition 2.6

Let $p_\theta \in \mathcal{P}$. Define the rank of p_θ relative to $\varepsilon \in \mathbf{r}$ as:

$$r_\theta(\varepsilon) := \sup_{M \in C(\theta,\varepsilon)} \quad (rank\ M)$$

Remark 2.7 If $p_\theta \in \mathcal{P}$ and $\theta = (k, A, B, \pi)$ then:

$$p_\theta(s_i \varepsilon t_j) = \pi M(s_i) M(\varepsilon) M(t_j) e$$

and, from Sylvester's inequality, $r_\theta(\varepsilon) \leq k$.

Definition 2.8

Let $p_\theta \in \mathcal{P}$ and $\theta = (k, A, B, \pi)$.

p_θ is said to be regular if $r_\theta(\varepsilon) = k, \ \ \forall \varepsilon \in \mathbf{r}$.

θ is regular if $\theta \in \Theta'$ and p_θ is regular.

Definitions 2.6 and 2.8 are analogous to notions defined in [5] for deterministic functions of Markov chains.

Theorem 2.9

Regular points of Θ' are identifiable modulo permutations of the states.

Remark 2.10

It is interesting to observe that under the *same* hypotheses theorem 2.9 has been proved in [6] in the case of *fixed order* of the model. For identification it is convenient to have compactness of the parameter space. Define, for $\delta > 0$:

$$\Theta'_\delta := \{\theta \in \Theta';\ a_{ij} \geq \delta,\ b_{j\varepsilon} \geq \delta,\ \forall i, j, \varepsilon\}$$

In what follows we will assume that the observed process $\{Y_t\}$ has a regular pdf $p_Y(\cdot) \in \mathcal{P}'$, i.e. there exists a regular point $\theta_0 = (n, A^0, B^0) \in \Theta'$ such that $p_Y(\cdot) = p_{\theta_0}(\cdot)$.

In practice the choice of δ is reduced to the choice of an upper bound K to the order n of the observed process. If $n \leq K$ and K is large enough taking $\delta := \frac{1}{K}$ we have $\theta_0 \in \Theta'_\delta$

3. Order Determination

We want to analyze the following question: given observations $(y_1, y_2, \ldots y_T)$ with T arbitrarily large, can we determine the order n of $\{Y_t\}$? The answer is yes if $p_{\theta_0} \in \mathcal{P}'_\delta$ and is regular.

It is convenient to decompose Θ'_δ as follows: $\Theta'_\delta = U_{k=1}^K \Theta'_{\delta,k}$ where $\Theta'_{\delta,k}$ is the section of Θ'_δ along k. We identify $\theta = (k, A, B) \in \Theta'_{\delta,k}$ with its projection $\theta = (A, B)$.

Define the random variables on $\mathbf{r}^*$:

$$\begin{aligned} \forall \theta \in \Theta'_\delta \qquad & f_1(\theta, Y(\cdot)) := P_\theta[Y_0] \\ \forall \theta \in \Theta'_\delta \text{ and } T \geq 2 \qquad & f_T(\theta, Y(\cdot)) := P_\theta[Y_0 | Y_{-T-1}^{-1}] \end{aligned}$$

Lemma 3.1[7]

$$f(\theta, Y(\cdot)) := \lim_{T\to\infty} f_T(\theta, Y(\cdot))$$

exists $\forall Y(\cdot) \in \mathbf{r}^*$ and is continuous on $\Theta'_{\delta,k}$ $\forall k \in \mathbf{K}$.

Define:

$$\begin{aligned} h_k(\cdot) : \Theta'_{\delta,k} I\!R, \to h_k(\theta) &:= E_{\theta_0}[\log f(\theta, Y(\cdot))] \\ h_{k,T}(\cdot) : \Theta'_{\delta,k} \to I\!R, h_{k,T}(\theta) &:= \frac{1}{T} \log P[Y_1^T | \theta] \\ h_k : \mathbf{K} \to I\!R \ , h_k &:= \sup_{\theta \in \Theta_{\delta,k}} h_k(\theta) \\ h_{k,T} : \mathbf{K} \times N \to I\!R \ , h_{k,T} &:= \sup_{\theta \in \Theta_{\delta,k}} h_{k,T}(\theta) \end{aligned}$$

Lemma 3.2[7]

i) $h_k(\cdot)$ is continuous on $\Theta'_{\delta,k}$ $\forall k \in \mathbf{K}$

ii) $h_k(\cdot) = \lim_{T\to\infty} h_{k,T}(\cdot)$ $\forall \theta \in \Theta'_{\delta,k}$ and a.e. P_{θ_0}

Theorem 3.3

i) $h_k(\theta) \leq h_n(\theta_0) \quad \forall k \in \mathbf{K},\ \theta \in \Theta'_{\theta,k}$

ii) $h_k(\theta) = h_n(\theta_0)$ iff $k = n$

This is an important consequence of theorem 2.9. Compactness of Θ'_δ allows us to conclude with:

Corollary 3.4 $h_n > h_k \qquad \forall k \in \mathbf{K}, \qquad k \neq n.$

From the abstract point of view this solves the problem of order determination. To compute the sequence h_k we must start from the data and therefore from the sequence of normalized maximum log-likelihoods $h_{k,T}$. In our situation the asymptotic behavior of $h_{k,T}$ is easy to analyze.

Lemma 3.5: $\quad \delta^T \leq P\left[Y_1^T | \theta\right] \leq (1-\delta)^T \ \ \forall \theta \in \Theta'_\delta$

and therefore:

Lemma 3.6: $\quad h_{k,\infty} := \lim_{T\to\infty} h_{k,T}$ exists $\forall k \in \mathbf{K}$

To be able to apply corollary 3.4 we need $h_{k,\infty} = h_k$ and this is true if:

$$\lim_{T\to\infty} \sup_{\theta \in \Theta'_{\delta,k}} h_{k,T}(\theta) = \sup_{\theta \in \Theta'_{\delta,k}} \lim_{T\to\infty} h_{k,T}(\theta) \tag{3.1}$$

The interchange of limit operations (3.1) is allowed under uniform convergence (for $T \to \infty$) of $h_{k,T}(\theta)$ on $\Theta'_{\delta,k}$. Using the previous results one can verify that actually the sequence $h_{k,T}(\theta)$ satisfies Dini's criterion for uniform convergence.

Observe that (3.1) is verified for every trajectory $Y_t(\omega)$ and from Lemma 3.2 we conclude that $h_{k,T} \longrightarrow h_k$ for almost (P_{θ_o}) every trajectory.

Theorem 3.7

i) *$arg\ max_k\{h_{k,T}\}$ is a strongly consistent estimator of n.*

ii) For (almost) every trajectory there exists T_o such that: $n = arg\ max\ _k\{h_{k,T}\} \qquad (\forall T \geq T_o)$

Remark The practical computation of $h_{k,T}$ (and of the maximizing $\theta = (A, B)$) can be done using the Baum-Eagon algorithm; see [8] for a description and the analysis of its numerical properties. p

References

[1] Heller, A., Math. Sys. Theory, 1, 1967, pp. 197-208.

[2] Picci, G., Springer-Verlag. Lecturer Notes in Economics and Math Sys.Vol. 162 (1978).

[3] Picci, G. van Schuppen, J. H. *On the Weak Finite Stochastic Realization Problem*

[4] Kalman, R. E., Developments in Statistics, Vol. 4, Academic Press, New York, 1983, pp. 97-136.

[5] Gilbert, E. J., Ann. Math. Stat., 30, 1959, pp. 688-697.

[6] Petrie, T., Ann. Math. Stat., 40, 1969, pp. 97-115.

[7] Baum, L., Petrie, T., Ann. Math. Stat., 37, 1966, pp. 1554-1563.

[8] Levinson, S. E., etal, Bell Tech. Jour., 62, 1983, pp. 1035-1074.

[9] Finesso, L., and Baras, J. S., Systems Research Center, TR-87-160, 1987.