1 Proof of Theorem 2

Construction of the projective limit: In the following, we have to explicitly treat the conditional $P_\Theta^\omega(A|\Theta)$ as the function $P_\Theta^\omega(A|\Theta)(\omega)$ for $A \in \mathcal{B}_x^\omega$ and $\omega \in \Omega$. As a function of $\omega$, the conditional is measurable w.r.t. the $\sigma$-algebra $\sigma(\Theta)$. As a regular conditional probability, the function $A \mapsto P_\Theta^\omega(A|\Theta)(\omega)$ is a probability measure for $\mathbb{P}$-almost all $\omega \in \Omega$. The null set of exceptions will be denoted $N^\omega \subset \Omega$. Since the conditional probabilities are conditionally projective, we have $P_{\Theta'}^\omega(A|\Theta \omega)(\omega) = P_{\Theta'}^\omega(\pi_{\Theta'}^\omega(A)|\Theta \omega)(\omega)$ for almost all $\omega$. Again there is a null set of exceptions, which we will denote $N^{\Theta'}$. Denote the union of all exceptions a $N := \bigcup_I N^I$. As a countable union of null sets, $N$ is itself a null set. Now for any fixed $\omega \notin N$, the probability measures $P_\Theta^\omega(\cdot|\Theta \omega)$ form a projective family of measures in the sense of the Kolmogorov theorem. Application of the $\sigma$-algebra $\sigma(\Theta)$ so obtained describes a conditional distribution of $X^\omega$ w.r.t. a $\sigma$-algebra $C^\omega$ if we can show that $\omega \mapsto \nu(A,\omega)$ is $C^\omega$-measurable for every $A \in \mathcal{B}_x^\omega$. This can be shown by means of the $\pi$-$\lambda$ theorem (also called the Dynkin lemma, [2]): First show that $\nu(A,\omega)$ is measurable for all $A$ in a generator of $\mathcal{B}_x^\omega$, and then deduce that this implies measurability for all $A$ by means of the $\pi$-$\lambda$ theorem. As a generator, we choose the “cylinder sets” $Z^\omega := \{ A \in \mathcal{B}_x^\omega | A = \pi_{\Theta'}^\omega(A') \}$, i.e. the set of all sets which are preimages under projection of some finite-dimensional event. Then $\mathcal{B}_x^\omega = \sigma(Z^\omega)$, a fact used for example in the proof of the Kolmogorov theorem (cf [1]). For any $A \in \mathcal{B}_x^\omega$, the function $\nu(A,\omega)$ is measurable: Since $\nu(\pi_{\Theta'}^\omega(A'),\omega) = P_{\Theta'}^\omega(A'|\Theta \omega)(\omega)$, the function $\omega \mapsto \nu(\pi_{\Theta'}^\omega(A'),\omega)$ is $\sigma(\Theta)$-measurable, and therefore $C^\omega$-measurable as $\sigma(\Theta)$ is $C^\omega$. Let $\mathcal{L} \subset \mathcal{B}_x^\omega$ denote the system of all $A$ for which $\nu(A,\omega)$ is $C^\omega$-measurable. In the sense of the $\pi$-$\lambda$ theorem, $\mathcal{L}$ is a $\lambda$-system: For $A = \Omega^\omega_x$, $\nu$ is constant hence measurable. Let $A \in \mathcal{L}$. Then $\nu(\Omega^\omega_x,\omega) = 1 - \nu(A,\omega)$, which is measurable. If $A_n \in \mathcal{L}$ is a pairwise disjoint sequence and $A' = \bigcup_n A_n$, then $\nu(A',\omega) = \lim_n \nu(A_n,\omega)$, which as a limit of measurable functions is measurable. It is well known that the cylinder sets $Z^\omega$ form an algebra [1], so $Z^\omega$ is in particular a $\pi$-$\lambda$ system. Then by the $\pi$-$\lambda$ theorem,

$$B_x^\omega = \sigma(Z^\omega) = \mathcal{L} \subset \mathcal{B}_x^\omega. \quad (1)$$

In other words, the set of all sets $A$ for which $\omega \mapsto \nu(A,\omega)$ is $C^\omega$-measurable is just $B_x^\omega$. Therefore, $\nu(A,\omega)$ is a regular version of the conditional probability $P_\Theta^\omega(A|C^\omega)(\omega)$. By construction, its marginals are $\pi_{\Theta'}^\omega \nu(A,\omega) = P_{\Theta'}^\omega(A'|\Theta \omega)(\omega)$ almost everywhere.

Interpreting $P_{\Theta'}^\omega(X|\mathcal{C}^\omega)$ as $P_{\Theta'}^\omega(X|\Theta')(\Theta):$ Under the additional assumption $\pi_{\Theta'} \Theta' = \Theta$, define the variable $\Theta'$ as $\Theta' := \bigotimes_{E \in \mathcal{E}} \Theta'(i)$. Then any conditional distribution given $\mathcal{C}^\omega$ can serve as a conditional given $\Theta'$, since the $\sigma$-algebra $\sigma(\Theta')$ generated by $\Theta'$ is just $C^\omega$:

$$\sigma(\Theta') = \Theta'\mathcal{C}(B_x^\omega) = \Theta^{E,\mathcal{C}}(\bigcup_{I \in \mathcal{F}(E)} \sigma(\pi_{\Theta'}^I B_x^\omega)) = \sigma(\bigcup_{I \in \mathcal{F}(E)} \Theta'^{-\mathcal{C}}(\Theta'\mathcal{C}^{-1} \pi_{\Theta'}^I B_x^\omega)) = \sigma(\bigcup_{I \in \mathcal{F}(E)} \Theta'^{-\mathcal{C}} B_x^\omega) = \Theta^{E,\mathcal{C}}. \quad (2)$$

2 Proof of Theorem 3

Proof of (1). We have to construct a candidate for the probability kernel $k^\omega$, and show that $T$ is a posterior index for the projective limit posterior with kernel $k^\omega$. To this end we will show that the conditionals $P_{\Theta'}^\Theta(\Theta'|T^1)$ are conditionally projective and define $k^\omega$ in terms of their projective limit. For each $I \in \mathcal{F}(E)$ and $A' \in \mathcal{B}_\Theta^\omega$, the function $\omega \mapsto k^I(A',\omega) \circ T \circ X^\omega(\omega)$ is $\sigma(T^1 \circ X^\omega)$-measurable, and $k^I(A',\omega) \circ T \circ X^\omega(\omega) = P_{\Theta'}^\Theta(A'|T^1)(\omega)$ a.e. Therefore, $k^I(A',\omega) \circ T \circ X^\omega$ is a version of $P_{\Theta'}^\Theta(A'|T^1)$. The conditional probabilities $P_{\Theta'}^\Theta(\Theta'|T^1)$ are conditionally projective:

$$P_{\Theta'}^\Theta(A'|T^1 = t^1) = P_{\Theta'}^\Theta(A'|X^\omega \in T^{-1}(t^1)) = P_{\Theta'}^\Theta(\pi_{\Theta'}^1 A'|X^\omega \in \pi_{\Theta'}^1 T^{-1}(t^1)) = P_{\Theta'}^\Theta(\pi_{\Theta'}^1 A'|X^\omega \in \pi_{\Theta'}^1 T^{-1}(t^1)) \quad (3)$$

Proofs

Construction of Nonparametric Bayesian Models from Parametric Bayes Equations

Peter Orbanz
Hence \( P_\Theta^E(\pi_{A^1}A^1|T^t) = P_\Theta^E(A^1|T^t) \), which is just the definition of conditional projectiveness. By Theorem 2 there is an a.e.-unique projective limit of the form \( P_\Theta^E(\Theta^E|C^E) \), where \( C^E \) is the \( \sigma \)-algebra \( \sigma := \sigma \left( \bigcup_{I \in \mathcal{F}(E)} \sigma(T^I) \right) \). (4)

It is straightforward to check that \( \sigma(T) = C^E \), because \( T \) satisfies Eq. (4). Therefore, the projective limit \( P_\Theta^E(\Theta^E|C^E) \) can serve as the conditional distribution \( P_\Theta^E(\Theta^E|T) \). Now define a candidate for the kernel \( k^E \) as

\[
k^E(A, t) := P_\Theta^E(A|T = t) \quad \text{for all } A \in B_\Theta^E, \ t \in \Omega_\Theta^E.
\]  

(5)

What remains to be shown is that \( k^E(A, T(x)) = P_\Theta^E(A|X^E = x) \) a.e. for all \( A \in \mathcal{F}(E) \). If this identity can be shown to hold for \( A \in Z^E \), then it holds for all \( A \). Since \( \sigma(Z^E) = B_\Theta^E \), and since \( Z^E \) is an algebra, the Carathéodory extension theorem is applicable to extend measures from \( Z^E \) to \( B_\Theta^E \). Since the conditional probability \( k^E \) is a Markov kernel, the Carathéodory theorem can be applied pointwise in \( x \). (For a conditional that is not a Markov kernel, the subset of exceptional points \( x \in \Omega_\Theta^E \) on which the conditional is not unique depends on \( A \). Over all \( A \), these could then aggregate into a non-null set.) To show that the identity holds on \( Z^E \), consider any \( A \in Z^E \), i.e. there is some \( I \in \mathcal{F}(E) \) such that \( A = \pi_{A^I}A^I \). Then

\[
k^E(\pi_{A^I}A^I, t \in \pi_{A^I}t^I) = P_\Theta^E(\pi_{A^I}A^I|T \in \pi_{A^I}t^I) = P_\Theta^E(A^I|T^I = t^I) = P_\Theta^E(A^I|X^I \in T^I t^I)
\]

(6)

such that \( k^E(\pi_{A^I}A^I, T(x)) = P_\Theta^E(\pi_{A^I}A^I|X^E = x) \). By the Carathéodory theorem, this implies that \( k^E(\pi_{A^I}A^I, T(x)) = P_\Theta^E(\pi_{A^I}A^I|X^E = x) \), and hence \( T \) is a posterior index for the projective limit posterior \( P_\Theta^E(\Theta^E|X^E) \), and \( k^E \) is the probability kernel corresponding to \( T \).

**Proof of (2).** To prove part (2), we have to show that the posterior index \( T^I \) and corresponding probability kernel \( k^I \) as specified in the theorem make each of the marginal Bayesian systems on the finite-dimensional subspaces \( \Omega_\Theta^B \) conjugate. That is, we have to verify \( k^I(A^I, T^I(x^I)) = P_\Theta^E(A^I|X^I = x^I) \). To this end, write

\[
k^I(A^I, t^I) = k(\pi_{A^I}A^I, t \in \pi_{A^I}t^I) = P_\Theta^E(\pi_{A^I}A^I|X^I \in T^I \pi_{A^I}t^I) = P_\Theta^E(A^I|X^I \in \pi_{A^I}T^I \pi_{A^I}t^I)
\]

(7)

Since, for each \( x^I \), there is some \( t^I \) such that \( T^I(x^I) = t^I \), this means:

\[
k^I(A^I, t^I) = P_\Theta^E(A^I|X^I = x^I) \iff x^I \in T^I(t^I),
\]

(8)

and thus \( P_\Theta^E(A^I|X^I = x^I) = k^I(A^I, T^I(x^I)) \) as we had to show. In other words, the posterior \( P_\Theta^E(A^I|X^I) \) is conjugate with posterior index \( T^I \) and probability kernel \( k^I \).