<table>
<thead>
<tr>
<th>Title</th>
<th>On the existence of bivariate kernel with given marginal kernels (Information and mathematics of non-additivity and non-extensivity: contacts with nonlinearity and non-commutativity)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Machida, Motoya; Shibakov, Alexander</td>
</tr>
<tr>
<td>Citation</td>
<td>数理解析研究所講究録 (2008), 1585: 47-50</td>
</tr>
<tr>
<td>Issue Date</td>
<td>2008-02</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/81519">http://hdl.handle.net/2433/81519</a></td>
</tr>
<tr>
<td>Type</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>Textversion</td>
<td>publisher</td>
</tr>
</tbody>
</table>

Kyoto University
On the existence of bivariate kernel with given marginal kernels

Motoya Machida and Alexander Shibakov

Tennessee Technological University

Abstract

We raise a question on jointly Markovian sample paths given marginal Markov chains, and prove that such a bivariate Markov chain exists when the state space is a compact Polish space.

1 Introduction

Strassen [4] proved the existence of a probability measure $\lambda$ to realize a pair $(X, X')$ of random variables whose marginals, say $p$ and $p'$, are given. Kamae, Krengel and O'Brien [3] investigated extensively the realization of an ordered pair $X \leq X'$, and associate Strassen's result with stochastic ordering when the underlying space $S$ is equipped with a partial ordering $\leq$. A probability measure $p$ on $S$ is said to be stochastically smaller than $p'$, denoted by $p \preceq p'$, if $\int f(s) p(ds) \leq \int f(s) p'(ds)$ for every real-valued increasing function $f$ on $S$. Then $p \preceq p'$ is a necessary and sufficient condition for the existence of probability measure $\lambda$ whose marginals are $p$ and $p'$, and whose support lies on the set $\Delta = \{(s, s') \in S \times S : s \leq s'\}$ (the Nachbin-Strassen theorem; see Theorem 1 of [3]). This existence theorem was immediately applied to that of Markov chains. A Markov transition kernel $k$ is said to be stochastically cross-monotone to a kernel $k'$ (or, $k$ stochastically dominates $k'$), if $k(r, \cdot) \preceq k'(r', \cdot)$ whenever $r \leq r'$. Assuming the cross-monotonicity between $k$ and $k'$, the respective Markov chain sample paths

$$(1.1) \quad X = (X_0, X_1, \ldots) \text{ and } X' = (X'_0, X'_1, \ldots)$$

can be realized so as to maintain the pairwise order $X'_n \leq X'_n$ for all $n \geq 0$ if the initial distribution $\pi_0$ for $X_0$ is also stochastically smaller than $\pi'_0$ for $X'_0$ (Theorem 2 of [3]).

The paired sample path $(X_n, X'_n)_{n=0,1,\ldots}$ in (1.1) is not necessarily Markovian. But if so, there is a bivariate kernel $K$ on $\Delta$ satisfying the marginal conditions

$$(1.2) \quad k(r, E) = K((r, r'), E \times S) \text{ and } k'(r', E') = K((r, r'), S \times E')$$
2 Measure space and selections

Let $\mathcal{C}$ be the space of real-valued continuous functions on the product space $\mathcal{S} \times \mathcal{S}$. The space $\mathcal{C}$ becomes a Banach space with the norm $\|f\| = \sup|f(S \times S)|$. A Radon measure $\lambda$ is a continuous linear functional on $\mathcal{C}$, and the functional has an integral form $\lambda(f) = \int f(r)\lambda(dr)$. The space $\mathcal{M}$ of Radon measures is a complete lattice, and the positive cone $\mathcal{M}^+$ consists of positive Radon measures (Theorem 11.2 of Choquet [2]). The space $\mathcal{M}$ is equipped with weak* topology, and the cone $\mathcal{M}^+$ is metrizable and separable (Theorem 12.10 of [2]). Let $\mathcal{D}$ be a countable dense subset of $\mathcal{C}$. The family of the semi-norms, $|\lambda(f)|$ for $f \in \mathcal{D}$, introduces the topology on $\mathcal{M}$, and it coincides with the weak* topology on the cone $\mathcal{M}^+$. Then we can form a countable subbase via

$$U_{f,q} := \{\lambda \in \mathcal{M}^+ : \lambda(f) > q\}, \quad f \in \mathcal{D}, q \in \mathbb{Q},$$

where $\mathbb{Q}$ denotes the set of all rational numbers.

Let $\Lambda$ be a closed set-valued map from $\Delta$ to $\mathcal{M}^+$. If $\Lambda$ is measurable, there exists a selection function $\lambda^{(r,r')} \in \Lambda(r,r')$ such that the map $\lambda^{(r,r')}$ is measurable from $\Delta$ to $\mathcal{M}^+$ (Theorem 8.1.3 of Aubin and Frankowska [1]). In particular, $\lambda^{(r,r')}(f)$ becomes a measurable function on $\Delta$ for each $f \in \mathcal{C}$. To see whether $\Lambda$ is measurable, it suffices to show that

$$\Lambda^{-1}(U_{f,q}) = \{(r, r') \in \Delta : \Lambda(r, r') \cap U_{f,q} \neq \emptyset\}$$

is a Borel measurable subset for every $f \in \mathcal{D}$, and $q \in \mathbb{Q}$ (Definition 8.1.1 of [1]). It is easily observed that $\Lambda(r, r') \cap U_{f,q} \neq \emptyset$ is equivalent to

$$q < H_f(r, r') = \sup_{\lambda \in \Lambda(r,r')} \lambda(f).$$

Thus, the verification of measurability of the set-valued map $\Lambda$ is reduced to that of the function $H_f(r, r')$. 

for $(r, r') \in \Delta$ and measurable sets $E$ and $E'$. Note in (1.2) that we view the measure $K((r, r'), \cdot)$ as if it lies on $S \times S$ and has its support on $\Delta$. Such a bivariate kernel exists via the Nachbin-Strassen theorem when $S$ is discrete (finite or countable). A probability measure $\lambda^{(r,r')}(\cdot)$ on $\Delta$ exists for each pair $(r, r') \in \Delta$ so that it has marginals $k(r, \cdot)$ and $k'(r', \cdot)$. Then $\lambda^{(r,r')}(\cdot)$ can be collectively viewed as a kernel $K((r, r'), \cdot)$. When $S$ is continuous (typically referred to a Polish space), however, the measurability of $\lambda^{(r,r')}$ with respect to $(r, r')$ has to be taken into account. This raises a question on whether $\lambda^{(r,r')}$ can be selected to ensure the measurability, and this expository paper discusses our investigation on a compact Polish space.
3 Validation of measurability

Let $C_S$ be the space of real-valued continuous functions on $S$. We write the direct sum $(f_1 \oplus f_2)(s, s') = f_1(s) + f_2(s')$ for $f_1, f_2 \in C_S$, and the subspace $C_S \oplus C_S = \{ f_1 \oplus f_2 : f_1, f_2 \in C_S \}$ on $C$. A probability measure $p$ is stochastically smaller than $p'$ if and only if $p(f_1) + p(f_2) \leq \sup (f_1 \oplus f_2)(\Delta)$ for any $f_1, f_2 \in C_S$. The Nachbin-Strassen theorem can be similarly stated on a par with this form of stochastic inequality. If $p \preceq p'$ then there exists $\lambda \in \mathcal{M}^+$ satisfying (i) $\lambda(f_1 \oplus f_2) = p(f_1) + p(f_2)$ for any $f_1, f_2 \in C_S$, and (ii) $\lambda(f) \leq \sup f(\Delta)$ for any $f \in C$. The above conditions clearly imply that (i) $\lambda$ has the marginals $p$ and $p'$, and (ii) it has a support on $\Delta$.

A Markov transition kernel $k$ on $S$ is a collection of positive Radon measures $k(s, \cdot)$ on $S$ such that $k(s, \cdot)$ is a probability measure for each $s \in S$ and

$$\langle k, f \rangle = \int f(s) k(r, ds)$$

is a measurable function of $r$ for every $f \in C_S$. Suppose that a Markov transition kernel $k$ is cross-monotone to $k'$. Then the cross-monotonicity is equivalently stated as

$$(3.1) \quad (\langle k, f_1 \rangle \oplus \langle k', f_2 \rangle)(r, r') \leq \sup (f_1 \oplus f_2)(\Delta)$$

for any $f_1, f_2 \in C_S$. For each $(r, r') \in \Delta$ we define the subset $\Lambda(r, r')$ consisting of $\lambda^{(r,r')} \in \mathcal{M}^+$ which satisfies the following two conditions.

$$(3.2) \quad \lambda^{(r,r')}(f_1 \oplus f_2) = (\langle k, f_1 \rangle \oplus \langle k', f_2 \rangle)(r, r') \quad \text{for } f_1, f_2 \in C_S;
$$

$$(3.3) \quad \lambda^{(r,r')}(f) \leq \sup f(\Delta) \quad \text{for } f \in C.$$  

It is easily observed that $\Lambda(r, r')$ is closed and that it is nonempty via the Nachbin-Strassen theorem. Let

$$(3.4) \quad H_f(r, r') = \inf_{f_1, f_2 \in C_S} [\sup (f_1 + f_2 + f)(\Delta) - (\langle k, f_1 \rangle \oplus \langle k', f_2 \rangle)(r, r')]$$

for each $(r, r') \in \Delta$ and $f \in C$. Then we have

**Proposition 3.1.** $\sup_{\lambda \in \Lambda(r, r')} \lambda(f) = H_f(r, r').$

**Proof.** Let $(r, r') \in \Delta$ and $f \in C$ be fixed. By replacing $f$ with $f_1 \oplus f_2 + f$ in (3.3), we can immediately observe that

$$(3.5) \quad \lambda(f) \leq H_f(r, r')$$

for every $\lambda \in \Lambda(r, r')$. By (3.2) the equality attains in (3.5) if $f \in C_S \oplus C_S$; thus, it is assumed that $f \not\in C_S \oplus C_S$. Let $\ell(f_1 \oplus f_2) = (\langle k, f_1 \rangle \oplus \langle k', f_2 \rangle)(r, r')$ for
Then $\ell$ is a well-defined linear functional on $C_S \oplus C_S$. By applying (3.1) and (3.2) together, we can observe that

$$-\sup(-f-g)(\Delta) - \ell(g) \leq \sup(f + g')(\Delta) - \ell(g')$$

for any $g, g' \in C_S \oplus C_S$, and therefore, that

$$\kappa_f = \inf_{g \in C_S \oplus C_S} (\sup(f + g)(\Delta) - \ell(g))$$

has a finite value. We can extend the subspace $\tilde{E} = \{g + tf : g \in C_S \oplus C_S, t \in \mathbb{R}\}$ by adding the element $f$, and define

$$\tilde{\ell}(g + tf) := \ell(g) + t\kappa_f$$

for $g \in C_S \oplus C_S$ and $t \in \mathbb{R}$. The map $\tilde{\ell}$ is a well-defined linear functional on $\tilde{E}$ and satisfies (3.2) and (3.3) with $\tilde{\ell}$ and $\tilde{E}$ in place of $\lambda^{(r,r')}$ and $C$. The same argument is essentially recycled to show that $\tilde{\ell}$ on $\tilde{E}$ is extended to $\lambda^{(r,r')} \in \Lambda(r, r')$ via Zorn's lemma. This particular $\lambda^{(r,r')}$ will achieve the equality in (3.5). $\square$

Observe that the space $C_S$ in the infimum of (3.4) can be replaced by a countable dense set, and consequently $H_f(r, r')$ is a measurable function of $(r, r')$ for each $f \in C$. Therefore, $\Lambda$ is a measurable set-valued map from $\Delta$ to $\mathcal{M}^+$, and there exists a measurable selection $\lambda^{(r,r')} \in \Lambda(r, r')$; thus, $K((r, r'), \cdot) = \lambda^{(r,r')}(\cdot)$ becomes a desired bivariate kernel on $\Delta$ satisfying (1.2).

References


