Statistics and Probability Letters 88 (2014) 1–8

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

On the effect of perturbation of conditional probabilities in total variation

Alessandro Abate [a,b], Frank Redig [c], Ilya Tkachev [b,*]

[a] Department of Computer Science, University of Oxford, United Kingdom
[b] Delft Center for Systems & Control, Delft University of Technology, The Netherlands
[c] Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 18 December 2013
Received in revised form 13 January 2014
Accepted 14 January 2014
Available online 23 January 2014

Keywords:
Conditional probabilities
Total variation
Perturbation
Coupling

ABSTRACT

A celebrated result by A. Ionescu Tulcea provides a construction of a probability measure on a product space given a sequence of regular conditional probabilities. We study how the perturbations of the latter in the total variation metric affect the resulting product probability measure.

1. Introduction

The Ionescu Tulcea extension theorem (Ash, 1972, Section 2.7.2) states that given a sequence of stochastic kernels, there exists a unique probability measure on the product space generated by this sequence, that is a measure whose conditional probabilities equal to these kernels. Such a construction is often used in the theory of general Markov Decision Processes (Bertsekas and Shreve, 1978), and general Markov Chains (Revuz, 1984) in particular. Hence, it is of a certain interest to study how sensitive the resulting product measure is with respect to perturbations of the generating sequence of kernels. A possible direct application of such result concerns numerical methods, where characteristics of the original stochastic process are studied over its simpler approximations, often defined over a finite state space. Such approximations can be further regarded as a perturbation of the original sequence of kernels (Tkachev and Abate, 2013) which connects to the original problem.

Here we specifically focus on the metric between kernels and measures given by the total variation norm. Given the pairwise distances between corresponding transition kernels in this metric, we are interested in bounds on the distance between the resulting product measures. A similar study was given in Roberts and Rosenthal (2013) which used the Borel assumption, that is it assumed that spaces involved are (standard) Borel spaces. However, the bounds obtained in Roberts and Rosenthal (2013) grow linearly with the cardinality of the sequence and hence are not tight: recall that the total variation distance between two probability measures is always bounded from above by 2.

In this paper we elaborate on the result of Roberts and Rosenthal (2013) in the two following directions. First, we generalize linear bounds to the case of arbitrary measurable spaces. Second, we show that under the Borel assumption used in the paper Roberts and Rosenthal (2013) it is possible to derive sharper bounds, that appear to be precise in some special cases—e.g. in case of independent products of measures.

* Corresponding author. Tel.: +31 1527 87171.
E-mail addresses: alessandro.abate@cs.ox.ac.uk (A. Abate), f.h.j.redig@tudelft.nl (F. Redig), i.tkachev@tudelft.nl (I. Tkachev).

0167-7152/$ – see front matter
http://dx.doi.org/10.1016/j.spl.2014.01.009

The rest of the paper is structured as follows. Section 2 gives a problem formulation together with statements of main results. Proofs are given in Section 3, which is followed by the discussion in Section 4 and an enlightening example in Section 5. With regards to the notation, terminology and conventions adopted in this paper, the readers should consult the Appendix.

2. Problem statement

Let us recall the construction of the product measure given the regular conditional probabilities. First of all, we need the following notion of a product of a probability measure and a stochastic kernel which extends a more usual product of two measures.

Proposition 1. *Let $(X, \mathfrak{X})$ and $(Y, \mathfrak{Y})$ be arbitrary measurable spaces. For any probability measure $\mu \in \mathscr{P}(X, \mathfrak{X})$ and any stochastic kernel $K : X \to \mathscr{P}(Y, \mathfrak{Y})$ there exists a unique probability measure $\mathsf{Q} \in \mathscr{P}(X \times Y, \mathfrak{X} \otimes \mathfrak{Y})$, denoted by $\mathsf{Q} := \mu \otimes K$, such that*

$$\mathsf{Q}(A \times B) = \int_A K(x, B)\mu(\mathrm{d}x)$$

for any pair of sets $A \in \mathfrak{X}$ and $B \in \mathfrak{Y}$.

Proof. For a proof, see Ash (1972, Section 2.6.2). □

The construction above immediately extends to any finite sequence of spaces by induction, whereas for the countable products the following result holds true.

Proposition 2 (*Ionescu-Tulcea*)**.** *Let $\{(X_k, \mathfrak{X}_k)\}_{k\in\mathbb{N}_0}$ be a family of arbitrary measurable spaces and let $(\Omega_n, \mathscr{F}_n) = \prod_{k=0}^n (X_k, \mathfrak{X}_k)$ be product spaces for any $n \in \bar{\mathbb{N}}_0$. For any probability measure $P^0 \in \mathscr{P}(X_0, \mathfrak{X}_0)$ and any sequence of stochastic kernels $(P^k)_{k\in\mathbb{N}}$, where $P^k : \Omega_{k-1} \to \mathscr{P}(X_k, \mathfrak{X}_k)$, there exists a unique probability measure $\mathsf{P} \in \mathscr{P}(\Omega_\infty, \mathscr{F}_\infty)$, denoted by $\mathsf{P} := \bigotimes_{k=0}^\infty P^k$, such that the finite-dimensional marginal P^n of P on the measurable space $(\Omega_n, \mathscr{F}_n)$ is given by $\mathsf{P}^n = \bigotimes_{k=0}^n P^k$ for any $n \in \mathbb{N}_0$.*

Proof. For a proof, see (Ash, 1972, Section 2.7.2). □

In the setting of Proposition 2, suppose that we are given another sequence of kernels $(\tilde{P}^k)_{k=0}^\infty$ and let $\tilde{\mathsf{P}} := \bigotimes_{k=0}^\infty \tilde{P}^k$ be the corresponding product measure. Given the assumption that $\|P^k - \tilde{P}^k\| \le c_k$ for any $k \in \mathbb{N}_0$ and some sequence of reals $(c_k)_{k\in\mathbb{N}_0}$, we study how the distance $\|\mathsf{P}^n - \tilde{\mathsf{P}}^n\|$ can be bounded. For the general case of arbitrary measurable spaces, the following result holds true.

Theorem 1. *Let $\{(X_k, \mathfrak{X}_k)\}_{k\in\mathbb{N}_0}$ be any family of measurable spaces and let $\tilde{\mathfrak{X}}_k \subseteq \mathfrak{X}_k$ for any $k \in \mathbb{N}_0$. Denote by $(\Omega_n, \mathscr{F}_n) = \prod_{k=0}^n (X_k, \mathfrak{X}_k)$ and $(\Omega_n, \tilde{\mathscr{F}}_n) = \prod_{k=0}^n (X_k, \tilde{\mathfrak{X}}_k)$ the corresponding product spaces for any $n \in \bar{\mathbb{N}}_0$. Let $P^0 \in \mathscr{P}(X_0, \mathfrak{X}_0)$, $\tilde{P}^0 \in \mathscr{P}(X_0, \tilde{\mathfrak{X}}_0)$ and let kernels $P^k : \Omega_{k-1} \to \mathscr{P}(X_k, \mathfrak{X}_k)$ and $\tilde{P}^k : \Omega_{k-1} \to \mathscr{P}(X_k, \tilde{\mathfrak{X}}_k)$ for $k \in \mathbb{N}$ be $\mathscr{F}_{k-1}$- and $\tilde{\mathscr{F}}_{k-1}$-measurable respectively. If a sequence of reals $(c_k)_{k\in\mathbb{N}_0}$ is such that $\|P^k - \tilde{P}^k\| \le c_k$ for all $k \in \mathbb{N}_0$, then for any $n \in \mathbb{N}_0$ it holds that*

$$\|\mathsf{P}^n - \tilde{\mathsf{P}}^n\| \le \sum_{k=0}^n c_k. \tag{2.1}$$

Remark 1. Through this paper, and in particular in the statement of Theorem 1, we use the following convention. If the domain of one measure is a subset of the domain of another, the total variation distance between them is taken over the smaller domain. For example, in the setting of Theorem 1 we have $\|P^0 - \tilde{P}^0\| = 2 \cdot \sup_{A\in\tilde{\mathfrak{X}}_0} |P^0(A) - \tilde{P}^0(A)|$.

The validity of results of Theorem 1 in some special cases was previously established in Roberts and Rosenthal (2013) and Tkachev and Abate (2013): we discuss these connections in a greater detail in Section 4. As it has been mentioned in Introduction, the bounds (2.1) are not tight. For example, if $c_k = c > 0$ for all $k \in \mathbb{N}_0$ then the right-hand side of (2.1) is $c \cdot n$ and diverges to infinity as $n \to \infty$, whereas the left-hand side stays bounded above by 2. It appears, that under a rather mild assumption that all involved measurable spaces are (standard) Borel, a stronger result can be obtained.

Theorem 2. *Let $\{X_k\}_{k\in\mathbb{N}_0}$ be a family of Borel spaces and let $\Omega_n = \prod_{k=0}^n X_k$ be product spaces for any $n \in \bar{\mathbb{N}}_0$. Let further $P^0, \tilde{P}^0 \in \mathscr{P}(X_0)$ and $P^n, \tilde{P}^k : \Omega_{k-1} \to \mathscr{P}(X_k)$ for $k \in \mathbb{N}$. If a sequence of reals $(c_k)_{n\in\mathbb{N}_0}$ is such that $\|P^k - \tilde{P}^k\| \le c_k$ for all $k \in \mathbb{N}_0$, then for any $n \in \mathbb{N}_0$ it holds that*

$$\|\mathsf{P}^n - \tilde{\mathsf{P}}^n\| \le 2 - 2\prod_{k=0}^n \left(1 - \frac{1}{2}c_k\right). \tag{2.2}$$

The proofs of both theorems are given in the next section.