Theory of Stochastic Processes
Vol. 22 (38), no. 2, 2017, pp. 1–7

GEORGII A. ALEKSEEV AND EKATERINA V. YUROVA

ON GAUSSIAN CONDITIONAL MEASURES DEPENDING ON A PARAMETER

We prove that if a family of Gaussian measures μ_α on the product of two Souslin locally convex spaces X and Y depends measurably on a parameter α, then it is possible to find conditional measures μ_α^y on X jointly measurable in y and α.

1. Introduction

Suppose we are given two locally convex spaces X and Y and a Radon probability measure μ on $X \times Y$. By a system of conditional measures for μ we shall understand a family of Radon probability measures μ^y on X satisfying the following conditions:

(1) for every Borel subset B of the space $X \times Y$, the function $\mu^y(B^y)$ is Borel measurable, where B^y is the projection onto X of the cross-section of B at the level y, i.e.,
$$B^y = \{x \in X \colon (x, y) \in B\};$$
(2) for every bounded Borel function f on $X \times Y$, the integral of f with respect to μ equals
$$\int_Y \int_X f(x, y)\, \mu^y(dx)\, \nu(dy),$$
where ν is the projection of the measure μ to Y.

It is known that conditional measures exist under broad assumptions, e.g., in case of Souslin spaces.

Suppose now that we are given a family $\{\mu_\alpha\}$ of Radon measures on the product $X \times Y$ depending measurably on a parameter α taking values in a measurable space $(\mathfrak{A}, \mathcal{A})$. The measurability of measures is associated with the weak topology and is understood as the $\mathcal{A}$-measurability in α of the integrals of bounded continuous functions. In case of Souslin spaces, this is also equivalent to the $\mathcal{A}$-measurability in α of the integrals of bounded Borel functions. A naturally arising question is whether it is possible to choose conditional measures so that they depend measurably on (α, y). It has been shown by I.I. Malofeev [8] that if X, Y and $\mathfrak{A}$ are completely regular Souslin spaces (i.e., continuous images of complete separable metric spaces), and $\mathcal{A} = \mathcal{B}(\mathfrak{A})$ is the Borel σ-algebra, then there is a version of conditional measures μ_α^y such that it depends measurably on (α, y) with respect to the σ-algebra $\mathcal{S}(Y \times \mathfrak{A})$ generated by Souslin sets (see Remark 2 below for his precise statement and its corollary). The latter is larger than the Borel σ-algebra, so the corresponding measurability is weaker. It is likely that in the general result of Malofeev the Borel measurability cannot be always guaranteed. In this paper, we show that a Borel measurable choice of conditional measures exists for centered Radon Gaussian measures. In case of Gaussian measures, conditional measures can be constructed explicitly, see [1], [3], [7], and [9].

2010 *Mathematics Subject Classification.* 28C20, 46G12, 60B11.

Key words and phrases. Gaussian measure, conditional measure, measurable dependence on a parameter.

This research has been supported by the Russian Science Foundation Grant N 17-11-01058 (at Lomonosov Moscow State University) .

Let us state the main result of this paper.

Theorem 1.1. *Suppose we are given a family of centered Gaussian measures μ_α on the product $X\times Y$ of two Souslin locally convex spaces that depends measurably on a parameter α taking values in a measurable space $(\mathfrak{A},\mathcal{A})$. Then there are Gaussian conditional measures μ_α^y that depend $\mathcal{B}(Y)\otimes\mathcal{A}$-measurably on (y,α).*

The proof is given in the last section. Note that the projections ν_α of the measures μ_α to Y are obviously measurable in α.

2. Notation, terminology, and auxiliary results

Let us recall that a probability measure μ on the Borel σ-algebra of a locally convex space X is called Radon if, for every Borel set B and for every $\varepsilon>0$, there exists a compact set K contained in B such that $\mu(B\setminus K)<\varepsilon$. If X is Souslin, then all Borel measures on X are Radon (see [2, Theorem 7.4.3]). The measure μ is called centered Gaussian if every continuous linear functional l on X is a centered Gaussian random variable on $\bigl(X,\mathcal{B}(X),\mu\bigr)$, i.e., the image measure $\mu\circ l^{-1}$ is either concentrated at zero or has a density of the form $(2\pi\sigma)^{-1/2}\exp(-t^2/(2\sigma))$. The density with $\sigma=1$ is called the standard Gaussian density and the corresponding measure is called standard Gaussian on the real line. An important example of a Gaussian measure is the product of countably many copies of the standard Gaussian measure on the real line defined on the space $\mathbb{R}^\infty$ of all real sequences. This measure is called the standard Gaussian measure on $\mathbb{R}^\infty$. For a survey of Radon and Gaussian measures, see [2], [1], and [3].

We recall that the weak topology on the space $\mathcal{M}(X)$ of Borel measures on a completely regular space X is generated by the seminorms

$$m\mapsto\left|\int_X f(x)\,m(dx)\right|,\quad f\in C_b(X),$$

where $C_b(X)$ is the space of bounded continuous functions on X. If X is Souslin, the space $\mathcal{M}(X)$ equipped with the weak topology is Souslin as well (see [2, Chapter 8]). Furthermore, the Borel σ-algebra of $\mathcal{M}(X)$ is generated by the integrals of bounded continuous functions on X, and the integrals of bounded Borel functions are Borel functionals on $\mathcal{M}(X)$. Therefore, in this case the measurability on the space of measures we discuss is precisely the Borel measurability for the Borel σ-algebra generated by the weak topology.

The subset of probability measures will be denoted by $\mathcal{P}(X)$.

According to the well-known Tsirelson theorem (see [1]), every centered Radon Gaussian measure μ is concentrated on a Souslin subspace and is the image of the standard Gaussian measure on $\mathbb{R}^\infty$ under a linear measurable mapping (more precisely, a Borel measurable linear mapping defined on a Borel linear subspace of full measure). Moreover, if μ is not concentrated on a finite-dimensional subspace, this mapping can be taken to be an isomorphism of Borel linear subspaces of measure 1. That is why we shall assume that spaces X and Y are Souslin although for our purposes it does not follow from the Tsirelson theorem since the subspace may depend on the parameter.

So suppose that spaces X and Y are Souslin locally convex spaces and let us consider the Borel structure of $\mathcal{P}(X\times Y)$. A Souslin locally convex space can be mapped into $\mathbb{R}^\infty$ by a continuous injective linear operator. By doing so, in addition to the original topology, one obtains the topology induced from $\mathbb{R}^\infty$. Since the coordinate functions are continuous in both topologies and separate points, the said topologies generate the same Borel structure (see [2, Theorem 6.8.9]). Moreover, the same reasoning holds true for the space of probability measures with the respective weak topologies (in this case one can find a countable family of bounded continuous functions separating measures). This reduces the general case of Souslin spaces X and Y to that of $\mathbb{R}^\infty$. For this reason we further consider the latter case.

Points of the space $\mathbb{R}^\infty \times \mathbb{R}^\infty$ will be denoted by (x, y), their components will be denoted by x_i and y_i, respectively.

It is known (see, e.g., [3, Theorem 5.14] or [1, Section 3.10]) that for a centered Gaussian measure μ on the product $X \times Y$ there exists a centered Gaussian measure σ on X such that conditional measures for μ can be found in the form

$$\mu^y = \sigma(\,\cdot - Ay), \quad A = \mathbb{E}(x \mid y),$$

where $\mathbb{E}(x \mid y)$ is the conditional expectation of the first component with respect to the σ-algebra generated by the second one. In other words,

$$A\colon Y \to X, \quad Ay = (\xi_1(y), \xi_2(y), \ldots),$$

where ξ_k is the conditional expectation of x_k with respect to the σ-algebra $\mathcal{B}_Y$ generated by coordinate functions y_i. This conditional expectation is the projection in $L^2(\mu)$ of the coordinate function x_k onto the closed linear subspace generated by the coordinate functions y_i. If $\{\eta_i\}$ is the orthogonalization of $\{y_i\}$ in $L^2(\mu)$, then

$$\xi_k = \sum_{i=1}^{\infty} (x_k, \eta_i)_{L^2(\mu)} \eta_i,$$

where the sum converges almost everywhere and in $L^2(\mu)$.

The measure σ coincides with the image of the measure μ under the measurable linear mapping $T\colon\ (x, y) \mapsto x - Ay$. Indeed, let us show that the measure μ coincides with the measure

$$\gamma = \int_Y \mu \circ T^{-1}(\,\cdot\, - Ay)\, \nu(dy).$$

The measure on the right can be written as the convolution of two centered Gaussian measures ($\mu \circ T^{-1}$ and $\nu \circ A^{-1}$). Hence it is centered Gaussian. Therefore, it suffices to verify that μ and γ have equal covariances, i.e., assign equal integrals to every continuous linear functional on $X \times Y$. Such a functional can be written in the form $(x, y) \mapsto f(x) + g(y)$ with two continuous linear functions f and g on X (in our case these are linear combinations of coordinate functions). It is readily seen that $f \circ A$ is the conditional expectation of f with respect to σ_Y. Hence the function $f(x) - f(Ay)$ is orthogonal to all y_j in $L^2(\mu)$ and so is independent with the σ-algebra σ_Y. By the orthogonality of $f(x) - f(Ay)$ and $f(Ay)$ in $L^2(\mu)$ we have

$$\int_{X\times Y} |f(x) + g(y)|^2\, \mu(dx\, dy) = \int_{X\times Y} |f(x) - f(Ay)|^2\, \mu(dx\, dy) + \int_{X\times Y} |f(Ay) + g(y)|^2\, \mu(dx\, dy).$$

On the other hand,

$$\int_{X\times Y} |f(x) + g(y)|^2\, \gamma(dx\, dy) = \int_Y \int_{X+Y} |f(x - Az + Ay) + g(y)|^2\, \mu(dx\, dz)\, \nu(dy)$$
$$= \int_Y \int_{X\times Y} \Bigl(|f(x - Az)|^2 + |f(Ay) + g(y)|^2\Bigr)\, \mu(dx\, dz)\, \nu(dy),$$

which coincides with the previous expression, since ν is the projection of μ on Y.

Therefore, the integral of a bounded Borel function f with respect to σ is calculated by the formula

$$\int_X f(x)\, \sigma(dx) = \int_{X\times Y} f(x - Az)\, \mu(dx\, dz),$$

and the integral with respect to μ^y is given by the formula

$$\int_X f(x)\,\mu^y(dx) = \int_{X\times Y} f(x - Az + Ay)\,\mu(dx\,dz).$$

Thus, for the proof of the theorem it suffices to show that the Gaussian measures $\sigma_\alpha(\,\cdot - A_\alpha y)$ depend measurably on (y,α), which reduces everything to considering the integrals

$$\int_{X\times Y} f(x - A_\alpha z + A_\alpha y)\,\mu_\alpha(dx\,dz).$$

Lemma 2.1. *For every separable metrizable topological vector space X, the mapping*

$$S\colon X\times\mathcal{P}(X)\to\mathcal{P}(X),\quad (h,m)\mapsto m_h,$$

with $m_h(B) = m(B-h)$, is continuous.

Proof. The weak topology on the space of probability measures on $X=\mathbb{R}^\infty$ is generated by the Kantorovich–Rubinstein norm

$$\|m\|_{\mathrm{KR}} = \sup\biggl\{\int f\,dm\colon\ f\in\mathrm{Lip}_1, \sup_{x\in X}|f(x)|\le 1\biggr\},$$

where Lip_1 is the class of 1-Lipschitz functions with respect to a fixed translation invariant metric ϱ on X generating the topology (such a metric always exists, see, e.g., [5]). Let $x,y\in X$ and $\mu,\nu\in\mathcal{P}(X)$. Let f be a 1-Lipschitz function such that $|f|\le 1$. Its integral with respect to $\mu(\,\cdot - x) - \nu(\,\cdot - y)$ equals

$$\int_X f(z+x)\,\mu(dz) - \int_X f(z+y)\,\nu(dz),$$

which can be bounded as follows:

$$\int_X |f(z+x) - f(z+y)|\,\mu(dz) + \biggl|\int_X f(z+y)\,(\mu-\nu)(dz)\biggr| \le \varrho(x,y) + \|\mu-\nu\|_{\mathrm{KR}}.$$

This shows that the mapping S is Lipschitz. □

It should be noted that this assertion is true for general topological vector spaces, we have required additional assumptions just for the sake of simplicity.

3. Proof of the theorem

Using the facts mentioned in the previous section, we construct a jointly measurable conditional expectation depending on α and then provide an explicit expression for a measurable version of conditional measures.

Proof of Theorem 1. The conditional expectation associated with μ_α will be denoted by $\mathbb{E}_\alpha$. As explained above, we have

$$\mathbb{E}_\alpha(x\mid y) = \Bigl(\mathbb{E}_\alpha(x_1\mid y), \mathbb{E}_\alpha(x_2\mid y),\dots\Bigr),$$

where the conditional expectation $\mathbb{E}_\alpha(x_i\mid y)$ coincides with the projection of the coordinate function x_i onto the closure of the linear span of the coordinate functions y_i in $L^2(\mu_\alpha)$. Certainly, for any fixed α, we have many versions of the conditional expectation. We now show that $\mathbb{E}_\alpha(x_i\mid y)$ can be obtained as a limit of continuous functions of y depending measurably on α.

Let L_n be the linear span of $y_1,\dots,y_n$. Let us show that, for any fixed i and n, for every α one can choose an element $l^n_{i,\alpha}\in L_n$ in such a way that it is the nearest element in L_n to the function x_k with respect to the norm of $L^2(\mu_\alpha)$ and the mapping $\alpha \mapsto l^n_{i,\alpha}$ is measurable in the sense that

$$l^n_{i,\alpha} = \sum_{j=1}^n c_{i,n,j}(\alpha) y_j,$$

where the functions $\alpha \mapsto c_{i,n,j}(\alpha)$ are $\mathcal{A}$-measurable.

At this step, we omit the indication of i and n in the notation for the element l_α and the corresponding coefficients $c_j(\alpha)$. If the matrix

$$G(\alpha) = \Big((y_i, y_j)_{L^2(\mu_\alpha)} \Big)_{i,j \le n}$$

is positive-definite, then the coefficients $c_j(\alpha)$ are uniquely determined from the linear system obtained by taking the inner products of l_α with $y_1, \ldots, y_n$ and are expressed by means of an explicit formula, from which the measurability with respect to α is obvious. The set of values α for which the determinant of $G(\alpha)$ is positive is in $\mathcal{A}$. The remaining values of α belong to finitely many disjoint subsets from $\mathcal{A}$ such that on each subset the rank of $G(\alpha)$ equals some $k < n$ and certain k coordinate functions from $y_1, \ldots, y_n$ are linearly independent in $L^2(\mu_\alpha)$. On the set where $G(\alpha) = 0$ we let $l_\alpha = 0$. On the set where $y_{i_1}, \ldots, y_{i_k}$ are linearly independent in $L^2(\mu_\alpha)$ we repeat the same procedure as above with n independent coordinates.

As a result, we obtain elements $l^n_{i,\alpha}$ that serve as conditional expectations of the coordinate functions x_i with respect to the measure μ_α and the σ-algebra generated by $y_1, \ldots, y_n$ and have the property that their coefficients $c_{i,j,n}(\alpha)$ are measurable in α. Let us set

$$A^n_\alpha y = \Big(\sum_{j=1}^n c_{1,j,n}(\alpha) y_j, \sum_{j=1}^n c_{2,j,n}(\alpha) y_j, \ldots \Big).$$

It is well-known (see [2, Chapter 10]) that the functions $\sum_{j=1}^n c_{i,j,n}(\alpha) y_j$ converge μ_α-almost everywhere and in $L^2(\mu_\alpha)$ to the functions $\mathbb{E}_\alpha(x_i \mid y)$ as $n \to \infty$. For the desired version of $\mathbb{E}_\alpha(x_i \mid y)$ we take the pointwise limit whenever it exists and take zero otherwise. This version is jointly measurable in (y, α) with respect to the σ-algebra $\mathcal{B}(Y) \otimes \mathcal{A}$.

By Lemma 2.1, in order to ensure the measurability of μ^y_α with respect to (y, α), it is enough to establish the measurability in α of the function

$$\int_{X \times Y} f(x - A_\alpha z)\, \mu_\alpha(dx\, dz)$$

for every bounded continuous function f on X. Moreover, one can easily see from the monotone class theorem (see, e.g., [2, p. 146]) that it suffices to do this for continuous functions in finitely many variables. Moreover, it suffices to consider functions that are linear combinations of exponents $\exp(i(v, x))$. Therefore, we can assume that $f(x) = f(x_1, \ldots, x_N)$ is a function on $\mathbb{R}^N$ of the form $f(x) = \exp(i(v, x))$. By the Lebesgue dominated convergence theorem, the integral we are interested in equals the limit of the integrals

$$\int_{X \times Y} f(x - A^n_\alpha y)\, \mu_\alpha(dx\, dy).$$

Therefore, everything reduces to the finite-dimensional case. Now our assertion becomes obvious, since we deal with the situation where $f(x) = \exp(i(T_\alpha v, x))$ with an operator T_α depending measurably on α and the covariance $Q_\alpha(x, x)$ of μ_α also depends measurably on α, so our integral equals $\exp(-Q_\alpha(T_\alpha v, T_\alpha v))$. It should be noted that the reason for using finite-dimensional approximations is that the measurable linear operator $A_\alpha y$ is not continuous in y, so that we have to take care of a jointly measurable version (in case of a continuous mapping in y, in order to ensure the joint measurability it would be enough to have the $\mathcal{A}$-measurability for every fixed y). □

Remark 3.1. Since the measure σ plays the crucial role in the expression for conditional measures, it is worth noting that its covariance can be expressed as follows:

$$\int_X l(x)^2\,\sigma(dx) = \inf_{c_1,\dots,c_n} \int_{X\times Y} |l(x - c_1 y_1 - \cdots - c_n y_n)|^2\,\mu(dx\,dy),$$

where $l(x) = k_1 x_1 + \cdots + k_n x_n$, $n \in \mathbb{N}$. In the case of a general locally convex space, a similar formula holds true:

$$\int_X l(x)^2\,\sigma(dx) = \inf_{g\in Y^*} \int_{X\times Y} |l(x) - g(y)|^2\,\mu(dx\,dy),$$

where $l \in X^*$. This expression can be used for proving the measurability of σ with respect to a parameter. For measures on $\mathbb{R}^n$ (or on a Hilbert space), this formula can be written in a more explicit form (see, for example, [6]).

Remark 3.2. Let us recall the precise formulation of Malofeev's result from [8]. Suppose that we are given a Borel mapping

$$f\colon\ (x,z) \mapsto f_z(x), \quad X \times Z \to Y,$$

where X, Y, Z are Souslin spaces. Suppose also that for every $z \in Z$ we are given a Borel probability measure μ_z on X such that the mapping

$$z \mapsto \mu_z,\ Z \to \mathcal{P}(X)$$

is Borel measurable provided that the space $\mathcal{P}(X)$ is equipped with the weak topology. Then, there exist conditional Borel probability measures $\{\mu_z^y\}_{y\in Y}$ for all pairs (μ_z, f_z) such that (i) $\mu_z^y(f_z^{-1}(y)) = 1$ for all $y \in f_z(X)$ for all z (such conditional measures are called proper), (ii) for each Borel set B in X, the function

$$(y,z) \mapsto \mu_z^y(B)$$

on $Y \times Z$ is measurable with respect to the σ-algebra $\mathcal{S}(Y \times Z)$ generated by all Souslin sets in $Y \times Z$, or, equivalently, the mapping

$$(y,z) \mapsto \mu_z^y, \quad Y \times Z \to \mathcal{P}(X)$$

is measurable when $Y \times Z$ is equipped with the σ-algebra $\mathcal{S}(Y \times Z)$ and $\mathcal{P}(X)$ is equipped with the Borel σ-algebra.

A closer look at the proof in [8] shows that in the case where f does not depend on z and is a Borel surjection possessing a Borel right inverse mapping g (which is not always the case), there exists a jointly Borel measurable version of conditional measures μ_z^y. Indeed, according to that proof, if $f_z = f$, then there are only two possible obstacles to obtaining the Borel measurability: either f does not have a Borel right inverse or its image $f(X)$ is not Borel in Y. Therefore, if X is the product of two Souslin spaces X_1 and X_2, $f_z = f$ is the standard projection onto X_2, and $z \mapsto \mu^z$ is Borel measurable, the reasoning in [8] yields existence of jointly Borel measurable conditional measures μ_z^y, since the projection is obviously surjective and, picking an arbitrary element $x_0 \in X_1$, one obtains a Borel right inverse $y \mapsto (x_0, y)$ for it.

In contrast to this setting, the space $\mathfrak{A}$ we consider is a general measurable space and conditional measures are described constructively. In a separate paper, we shall consider the case of a more general linear conditioning defined by a measurable linear mapping also depending on a parameter.

Investigation of dependence of conditional measures on a parameter can be useful for various optimization problems including optimal transportation and stochastic approximation (see, for example, [4], [9], and [10]).

References

1. V. I. Bogachev, *Gaussian measures*, Amer. Math. Soc., Rhode Island, Providence, 1998.
2. V. I. Bogachev, *Measure theory V. 1, 2.*, Springer, New York, 2007.
3. V. I. Bogachev, *Gaussian measures on infinite-dimensional spaces*, Real and Stochastic Analysis. Current Trends (M.M. Rao ed.), World Sci., Singapore, 2014, pp. 1–83.
4. V. I. Bogachev and A. V. Kolesnikov, *The Monge–Kantorovich problem: achievements, connections, and perspectives* Russian Math. Surveys **67** (2012), no. 5, 785–890.
5. V. I. Bogachev O. G. Smolyanov, *Topological vector spaces and their applications*, Springer, New York, 2017.
6. M. Hairer, A. M. Stuart, J. Voss and P. Wiberg, *Analysis of SPDEs arising in path sampling. I. The Gaussian case.*, Commun. Math. Sci. **3** (2005), no. 4, 587–603.
7. T. LaGatta, *Continuous disintegrations of Gaussian processes*, Theory Probab. Appl. **57** (2013), no. 1, 151–162.
8. I. I. Malofeev, *Measurable dependence of conditional measures on a parameter*, Doklady Akad. Nauk **470** (2016), no. 1, 13–17. (in Russian) English transl.: Doklady Math. **94** (2016), no. 2, 493–496.
9. V. Tarieladze and N. Vakhania, *Disintegration of Gaussian measures and average-case optimal algorithms*, J. Complexity **23** (2007), no. 4-6, 851–866.
10. C. Villani, *Optimal transport, old and new*, Springer, New York, 2009.

Department of Mechanics and Mathematics, Moscow State University, 119991 Moscow, Russia

E-mail address: `nlevi@protonmail.com`

Department of Mechanics and Mathematics, Moscow State University, 119991 Moscow, Russia

E-mail address: `yurova-k@rambler.ru`