INTERPOLATION OF GAUSS–MARKOV PROCESSES

PETER MATHÉ AND BERND SCHMIDT

Abstract. We study the problem of simulation of conditioned Gaussian processes. If the number of conditions is large, then such simulation is effective only, if at any specific time, only a few conditions have to be taken into account.

We shall see, that local interpolation is tied to the Markov property. For Gaussian Markov processes we establish some explicit formulae for conditional mean and covariance and discuss some applications.

1. Introduction

Suppose we are given an $\mathbb{R}^d$–valued stochastic process $(X_t)_{t\geq 0}$ with known mean function $m(t) := \mathbf{E}X_t,\ t \geq 0$, and covariance function

$$R_{s,t} = \mathbf{E}\left(X_s - m(s)\right)\left(X_t - m(t)\right)^T.$$

Throughout we shall assume, that the process $(X_t)_{t\geq 0}$ has almost surely continuous trajectories. By the Kolmogorov criterion of continuity, a sufficient condition for this can be given through the continuity of the covariances.

Suppose further, that we have observations $X_{t_j} = x_j,\ j = 1, \dots, n$. Throughout we shall denote $\Delta = \{t_1, \dots, t_n\}$ the design of the given observation sites, and we tacitly assume that the sites are numbered by order, i.e., $t_1 < t_2 < \cdots < t_n$. Our task consists in simulating the process $(X_t)_{t\geq 0}$, conditioned on these observations. Since at the observation sites, the conditional process, which shall be denoted by $\left(X_t^{\Delta}\right)_{t\geq 0}$, coincides with the observations, this is an interpolation problem, which in general may be hard to solve. But, if $(X_t)_{t\geq 0}$ is Gaussian, then all conditional distributions are Gaussian and we can simulate the conditional process knowing the conditional mean and covariance, only.

There are various applications to this task.

Version: October 4, 2000.

2000 *Mathematics Subject Classification.* Primary 65C05; Secondary 60G15, 60J25.

Key words and phrases. Gauss–Markov process, conditional distribution, stochastic differential equation.

(i) The classical problem is the determination of the distribution of X_t, given "previous" observations $X_{t_j},\ j = 1, \dots, n,\ t_j < t$. This is the problem of constructing a process step wise in time, see [4, 2.1.3].

(ii) Our initial interest stems from a paper by S. Prigarin [11]. The author studies boundary value problems for certain solutions $(X_t)_{t \geq 0}$ of stochastic differential equations, given by

$$dX_t = (A(t)X_t + b(t))\, dt + \Sigma(t)\, dB_t,\ t \geq 0, \tag{1}$$

with matrix–functions $A(t),\ \Sigma(t)$, being piece-wise continuous, and known $b(t)$. We are given values $X_\alpha = x_\alpha,\ X_\beta = x_\beta,\ \alpha < \beta$ and we want to simulate $X_t,\ t \in (\alpha, \beta)$, conditioned on the observations.

(iii) More advanced is the optimal design problem (in one dimension): If for given observation sites Δ and values $X_{t_j} = x_j,\ j = 1, \dots, n$, the process $P^\Delta X_t,\ 0 \leq t \leq T$ denotes the orthogonal projection of X_t onto span $\{X_{t_1}, \dots, X_{t_n}\}$ in $L_2(0,T)$, then determine a design Δ which minimizes

$$\int_0^T \mathbf{E} \left| X_t - P^\Delta X_t \right|^2\, dt.$$

In general, results can only be obtained in the asymptotic setting, i.e., when $\mathrm{card}(\Delta) \to \infty$. In this context, the optimal design problem arose from [13, 14]. More recent contributions are [10, 9], see also [8]. In great generality results were obtained for processes of product type, i.e., the covariance splits

$$R_{s,t} = u(\min\{s,t\})v(\max\{s,t\}). \tag{2}$$

(iv) The process X_t^Δ interpolates the original process at the design points. So we may wish to approximate a given process $(X_t)_{t \geq 0}$ by a sequence $X_t^{\Delta_n}$ of conditional processes, based on a nested sequence Δ_n of more and more dense observations. For Brownian motion this results in the Haar function construction (dyadic refinements and piece–wise linear interpolation) [2, Chap. I, §2]. Again, it is required to consecutively determining intermediate values, given previous observations.

Simulation of the conditional process $\left(X_t^\Delta\right)_{t \geq 0}$ can only be done efficiently, if the simulation of X_t given observations from Δ, depends locally on Δ. Best situation is, when only neighboring sites are involved.

We shall denote

$$\mathcal{F}_\Delta := \sigma\{X_{t_j}, \quad j = 1, \dots, n\}.$$

Given a process $(X_t)_{t\geq 0}$, we agree to call an interpolation *local*, if for any design Δ and any t the distribution of X_t^Δ depends only on the neighboring sites $\Theta(t)$, i.e.,

$$P(X_t \in A|\mathcal{F}_\Delta) = P(X_t \in A|\mathcal{F}_{\Theta_\Delta(t)}),$$

where

$$\Theta_\Delta(t) := \begin{cases} \{t_{k-1}, t_k\} & \text{if there is } k \leq n, \text{ for which } t_{k-1} \leq t < t_k, \\ \{t_1\} & \text{if } t < t_1 \text{ and finally} \\ \{t_n\} & \text{if } t \geq t_n. \end{cases}$$

In this note we shall establish, that locality of interpolation is equivalent to the Markov property, which means

$$P(X_t \in A|\mathcal{F}_\Delta) = P(X_t \in A|X_{t_n})$$

whenever $t_1 < t_2 < \dots < t_n < t$. For Gaussian Markov processes (Gauss–Markov processes) this allows to provide explicit formulae for conditional mean and covariance. Thus we are able to efficiently answer problems like (i)–(iv) above.

We add some structural results on Gauss–Markov processes, which provide additional insight, how the Markov property implies locality of interpolation for this particular type of processes.

We conclude with an example of linear stochastic differential equations, generalizing some results mentioned in (ii).

2. Locality of interpolation and the Markov property

As discussed in the Introduction we shall consider vector–valued processes $(X_t)_{t\geq 0}$, given on some probability space $(\Omega, \mathcal{F}, P)$ and assume that we have observations $X_{t_j} = x_j$, $j = 1, \dots, n$. Thus we aim at simulating X_t conditioned on $\mathcal{F}_\Delta$. More precisely, given any design Δ, let the regular conditional distributions be

$$P(X_{s_1} \in A_1, \dots, X_{s_l} \in A_l|\mathcal{F}_\Delta)(x) := \mathbf{E}(\prod_{j=1}^{l} \mathbf{1}_{A_j}(X_{s_j})|\mathcal{F}_\Delta)(x),$$

for any $x \in (\mathbb{R}^d)^n$ and $l \geq 1$, $s_1, \dots, s_l$. For fundamentals on conditional expectations and distributions we refer to [1, 6]. Since P is the law of the stochastic process $(X_t)_{t\geq 0}$ with continuous trajectories, we infer, that for almost all $x \in (\mathbb{R}^d)^n$, the family of conditional distributions constitutes a consistent family of probabilities, thus gives rise to (a family of) processes $\left(X_t^\Delta(x)\right)_{t\geq 0}$ with distributions

$$P(X_t^\Delta(x) \in A) = P(X_t \in A|\mathcal{F}_\Delta)(x).$$

Therefore, given design Δ and realization $x = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$, we shall call $\left(X_t^\Delta(x)\right)_{t\geq 0}$ the conditional process of $(X_t)_{t\geq 0}$, given observations at Δ with $X_{t_j} = x_j,\ j = 1, \ldots, n$. As we shall see below, the conditional processes have the Markov property, if the original processes had.

First, we shall derive the following result for any given point $t \geq 0$. We need to know the conditional distributions $P(X_t \in A|\mathcal{F}_\Delta)$. It is immediate from the definition of locality, that it implies the Markov property. Now we shall prove, that the converse is also true: The Markov property implies locality of interpolation.

Theorem 1. *For any Markov process $(X_t)_{t\geq 0}$ with values in $\mathbb{R}^d$ interpolation is local. Precisely, for any finite design Δ and $t > 0$ we have*

$$P(X_t \in A|\mathcal{F}_\Delta) = P(X_t \in A|\mathcal{F}_{\Theta_\Delta(t)}), \quad A \in \mathbb{R}^d \text{ measurable.} \tag{3}$$

The proof will follow from Propositions 1 and 2 below. It is enough to restrict to the case of times t, for which $\Theta_\Delta(t)$ consists of two neighboring points. If t was right from the design, then (3) is just the Markov property. Also, if t was left from the design, this is a consequence of the Markov property, now for the time reversed process. (For equivalent formulations of the Markov property see e.g. [3, p. 2].)

Given $T \subset \mathbb{R}_+$, let $a, b \in T$ with $a < b$ and define σ–algebras $\mathcal{V} := \sigma(X_t; t \in T, t \leq b)$ und $\mathcal{Z} := \sigma(X_t; t \in T, t \geq a)$.

Proposition 1. *The σ–algebras $\mathcal{V}$ and $\mathcal{Z}$ are independent, given $\mathcal{P} := \sigma(X_t; t \in T, a \leq t \leq b)$.*

Proof. Since $\mathcal{P} \subset \mathcal{V}$, it is enough to show

$$\mathbf{E}(\mathbf{1}_Z|\mathcal{V}) = \mathbf{E}(\mathbf{1}_Z|\mathcal{P}) \quad \forall Z \in \mathcal{Z}, \tag{4}$$

see e.g. [1, Thm. 4.5.3] or [6, Lem. 2.26]. If we let $\mathcal{E} := \sigma(X_t; t \in T, t \geq b)$, then $\mathcal{Z} = \sigma(\mathcal{P} \cup \mathcal{E})$. The set of those $Z \in \mathcal{Z}$, which obey (4) is a Dynkin system, such that it suffices to show (4) for sets of the form

$Z = P \cap E$, with $P \in \mathcal{P}$, $E \in \mathcal{E}$. But

$$\begin{aligned} \mathbf{E}(\mathbf{1}_Z|\mathcal{V}) &= \mathbf{E}(\mathbf{1}_P\mathbf{1}_E|\mathcal{V}) = \mathbf{1}_P\mathbf{E}(\mathbf{1}_E|\mathcal{V}) \\ &= \mathbf{1}_P\mathbf{E}(\mathbf{1}_E|X_b) \qquad (5)\\ &= \mathbf{1}_P\mathbf{E}(\mathbf{1}_E|\mathcal{P}) \qquad (6)\\ &= \mathbf{E}(\mathbf{1}_P\mathbf{1}_E|\mathcal{P}) = \mathbf{E}(\mathbf{1}_Z|\mathcal{P}), \end{aligned}$$

using properties if conditional expectations and the Markov property to obtain (5) and (6), respectively. □

Fix time t_0 and assume that $t_0 \in (t_k, t_{k+1})$ for some k. Next we construct a pair $(Y_t)_{t\geq 0}$ of processes, derived from the original one by time reversion. For this purpose define continous $u, v : [0,2] \to [0,\infty)$, u increasing, with $u(0) = t_0$, $u(1) = t_{k+1}$, $u(2) = t_n$, v decreasing and $v(0) = t_0$, $v(1) = t_k$, $v(2) = t_1$. Let $Y_s := (X_{u(s)}, X_{v(s)})$, $s \in [0,2]$.

Proposition 2. *$(Y_s)_{s\in[0,2]}$ is a Markov process with values in $\mathbb{R}^d \times \mathbb{R}^d$.*

Proof. Let $0 \leq s_1 \leq \ldots \leq s_k \leq s \leq 2$ be fixed. It is to show, that

$$P(Y_s \in A|Y_{s_1}, \ldots, Y_{s_k}) = P(Y_s \in A|Y_{s_k})$$

for $A \in \left(\mathbb{R}^d\right)^2$ measurable. Again, the set for which this holds is a Dynkin system and it suffices to consider product sets $A = B \times C$, $B, C \in \mathbb{R}^d$ measurable. We are done once we have shown

$$\begin{aligned} &P(X_{u(s)} \in B, X_{v(s)} \in C|\sigma\{X_r; v(s_k) \leq r \leq u(s_k)\}) \\ &\qquad = P(X_{u(s)} \in B, X_{v(s)} \in C|X_{v(s_k)}, X_{u(s_k)}). \end{aligned}$$

But this follows from Proposition 1 for $a = v(s_k)$, $b = u(s_k)$, since with $\mathcal{P} = \sigma(X_r; v(s_k) \leq r \leq u(s_k))$, the left hand side evaluates to

$$\begin{aligned} P(X_{u(s)} \in B, X_{v(s)} \in C|\mathcal{P}) &= P(X_{u(s)} \in B|\mathcal{P})\ P(X_{v(s)} \in C|\mathcal{P}) \\ &= P(X_{u(s)} \in B|X_{u(s_k)})\ P(X_{v(s)} \in C|X_{v(s_k)}), \end{aligned}$$

by the Markov property of (X_t) and the time reversed process. For the right hand side we use the Markov process $(X_{v(s)}, X_{v(s_k)}, X_{u(s_k)}, X_{u(s)})$, on $T = \{v(s), v(s_k), u(s_k), u(s)\}$, and $\mathcal{P} = \sigma(X_{v(s_k)}, X_{u(s_k)})$ to conclude similarly

$$\begin{aligned} P(X_{u(s)} \in B, X_{v(s)} \in C|\mathcal{P}) &= P(X_{u(s)} \in B|\mathcal{P})\ P(X_{v(s)} \in C|\mathcal{P}) \\ &= P(X_{u(s)} \in B|X_{u(s_k)})\ P(X_{v(s)} \in C|X_{v(s_k)}), \end{aligned}$$

which proves equality as claimed. □

Proof of Theorem 1. Since (Y_s) is Markov, the time reversion is also. This implies $P(Y_0 \in A|\sigma(Y_r; 1 \leq r \leq 2)) = P(Y_0 \in A|Y_1)$ for measurable $A \in \mathbb{R}^d$. Since $Y_0 = (X_{t_0}, X_{t_0})$, $\sigma(Y_r; 1 \leq r \leq 2) \supset \sigma(X_r; r \in$

$[t_1, t_k] \cup [t_{k+1}, t_n]$) and $\sigma(Y_1) = \sigma(X_k, X_{k+1})$, we indeed have for measurable $A \in \mathbb{R}^d$:

$$P(X_{t_0} \in A|\mathcal{F}_\Delta) = P(X_{t_0} \in A|\mathcal{F}_{\Theta_\Delta(t_0)}),$$

which completes the proof of the theorem. $\square$

As an immediate application we have

Corollary 1. *For each design Δ of cardinality say n, the conditional processes $\left(X_t^\Delta(x)\right)_{t\geq 0}$ are Markovian for almost all $x \in (\mathbb{R}^d)^n$.*

Proof. Let $s_1 < \ldots s_l < t$ and $S := \{s_1, \ldots, s_l\}$. In analogy we let $\mathcal{F}_S^\Delta(x)$ denote the σ–algebra, generated by $X_{s_1}^\Delta(x), \ldots, X_{s_l}^\Delta(x)$. We have to prove, that for a.a. $x \in (\mathbb{R}^d)^n$

$$P(X_t^\Delta(x) \in A|\mathcal{F}_S^\Delta(x))(s) = P(X_t^\Delta(x) \in A|\mathcal{F}_{\{s_l\}}^\Delta(x))(s) \tag{7}$$

holds for a.a. $s \in (\mathbb{R}^d)^l$. It follows from the definition of conditional probabilities, that

$$P(X_t^\Delta(x) \in A|\mathcal{F}_S^\Delta(x))(s) = P(X_t \in A|\mathcal{F}_{S\cup\Delta})(s, x), (s, x) \text{ a.s.}$$

Theorem 1 implies

$$P(X_t^\Delta(x) \in A|\mathcal{F}_S^\Delta(x))(s) = P(X_t \in A|\mathcal{F}_{\Theta_{S\cup\Delta}})(s, x).$$

There are four cases. We indicate these with the respective $\Theta_{S\cup\Delta}(t)$.

(i) $s_l < t < t_1$, thus $\Theta_{S\cup\Delta}(t) = \{s_l, t_1\}$.
(ii) For some k we have $s_l < t_k \leq t < t_{k+1}$, in which case $\Theta_{S\cup\Delta}(t) = \{t_k, t_{k+1}\}$.
(iii) For some k we have $t_k \leq s_l < t < t_{k+1}$, thus $\Theta_{S\cup\Delta}(t) = \{s_l, t_{k+1}\}$, and
(iv) $t_n < s_l < t$, hence $\Theta_{S\cup\Delta}(t) = \{s_l\}$.

In either case we have $\Theta_{S\cup\Delta}(t) = \Theta_{\{s_l\}\cup\Delta}(t)$ and an application of Theorem 1, respectively for design $S \cup \Delta$ and $\{s_l\} \cup \Delta$ completes the proof of the corollary. $\square$

Although Theorem 1 establishes locality of interpolation for Markov processes, this may not help to simulate the conditional distributions unless we have further information. This is the case for Gauss processes and will be studied below.

3. Conditional Gauss–Markov processes

If $(X_t)_{t\geq 0}$ is Gaussian with known mean function $m(t)$, then $X_t^\Delta(x)$ is a family of Gaussian distributions, which is completely determined by the respective (conditional) mean and (conditional) covariance.

Proposition 3. *Let Δ be any given design. For every $t > 0$ and $x = (x_1, \ldots, x_n) \in (\mathbb{R}^d)^n$, the distribution of $X_t^\Delta(x)$ is Gaussian with mean*

$$m^\Delta(t) := m(t) + \mathbf{E}(X_t - m(t)|\mathcal{F}_\Delta)(x_1, \ldots, x_n)$$

and covariance

$$\mathbf{E}\left(X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)\right)\left(X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)\right)^T.$$

It is clear, that we may restrict considerations to the case, that the process $(X_t)_{t\geq 0}$ is centered, i.e., $m(t) = 0,\ t \geq 0$.

Proof. The statement for the conditional mean follows easily from the definition. To derive the representation for the conditional covariance, we use that $X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)$ is independent of $\mathcal{F}_\Delta$, see e.g. [5, Chapt. III,§6]. Thus for fixed Δ and x we conclude

$$\begin{aligned}
&\mathbf{E}\left(X_t^\Delta(x) - \mathbf{E}X_t^\Delta(x)\right)\left(X_t^\Delta(x) - \mathbf{E}X_t^\Delta(x)\right)^T \\
&\quad = \mathbf{E}((X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)(x))\left(X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)(x)\right)^T|\mathcal{F}_\Delta)(x) \\
&\quad = \mathbf{E}((X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta))\left(X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)\right)^T|\mathcal{F}_\Delta)(x) \\
&\quad = \mathbf{E}(X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta))\left(X_t - \mathbf{E}(X_t|\mathcal{F}_\Delta)\right)^T,
\end{aligned}$$

where we used the independence to derive the last equality. □

We note explicitly, that the conditional covariance does not depend on x, but only on the design. Moreover, conditional mean and covariance depend only on the neighboring observation sites for Gauss–Markov processes. Therefore, the explicit description of mean and covariance of the conditional processes X_t^Δ allow to provide explicit formulae, which will be presented next.

4. Some explicit formulae

We start with a centered Gaussian process $(Y_t)_{t\geq 0}$ with independent increments, an important instance of Gauss–Markov processes. Its covariance function evaluates, for $s \leq t$ as

$$\begin{aligned}
R_{s,t}^Y = \mathbf{E}Y_sY_t^T &= \mathbf{E}Y_s(Y_t - Y_s)^T + \mathbf{E}Y_sY_s^T \\
&= \mathbf{E}Y_sY_s^T =: \rho(s),
\end{aligned}$$

which yields $R^Y_{s,t} = \rho(\min\{s,t\})$. In this case, the description of the conditional distributions is particularly simple; also, locality of interpolation can be seen directly in this case. For simplicity of presentation we restrict to the case $\rho(t) - \rho(s) > 0$ [1] if $t > s$.

Again, $\Delta = \{t_1, \dots, t_n\}$ is the given design with respective observations $Y_{t_j} = y_j,\ j = 1, \dots, n$. Now, $\mathbf{E}(Y_t|\mathcal{F}_\Delta)$ can be considered as projection $P^\Delta Y_t$. Since $(Y_t)_{t\geq 0}$ has independent increments, the random vectors $Y_{t_1}, Y_{t_2} - Y_{t_1}, \dots, Y_{t_n} - Y_{t_{n-1}}$, are mutually orthogonal and

$$P^\Delta Y_t := \mathbf{E}(Y_t|\mathcal{F}_\Delta) = \sum_{j=1}^n \lambda_j(t)\left(Y_{t_j} - Y_{t_{j-1}}\right),$$

where we put $Y_{t_0} = 0$, for some functions $\lambda_j(t),\ j = 1, \dots, n$. These can be readily computed by noting, that the projection $Y_t - P^\Delta Y_t$ is characterized through orthogonality to each of the increments $Y_{t_j} - Y_{t_{j-1}},\ j = 1, \dots, n$, which yields

$$\begin{aligned}\lambda_j(t) &= \left(R^Y_{t,t_j} - R^Y_{t,t_{j-1}}\right)\left(R^Y_{t_j,t_j} - R^Y_{t_{j-1},t_{j-1}}\right)^{-1} \\ &= \left(\rho(\min\{t,t_j\}) - \rho(\min\{t,t_{j-1}\})\right)\left(\rho(t_j) - \rho(t_{j-1})\right)^{-1}.\end{aligned}$$

More explicitly,

$$\lambda_j(t) = \begin{cases} 0, & \text{if } t < t_{j-1} \\ \left(\rho(t) - \rho(t_{j-1})\right)\left(\rho(t_j) - \rho(t_{j-1})\right)^{-1}, & \text{for } t_{j-1} \leq t < t_j \\ I, & \text{for } t > t_j. \end{cases}$$

We arrive at

$$P^\Delta Y_t = Y_{t_{j-1}} + \left(\rho(t) - \rho(t_{j-1})\right)\left(\rho(t_j) - \rho(t_{j-1})\right)^{-1}\left(Y_{t_j} - Y_{t_{j-1}}\right)$$

if $t \in [t_{j-1}, t_j)$. We summarize our analysis in

Theorem 2. *If $(Y_t)_{t\geq 0}$ is centered Gaussian with independent increments, then interpolation is local.*

For any finite design $\Delta = \{t_1, \dots, t_n\}$, the distribution of Y^Δ_t is Gaussian with mean $m^\Delta(t)$ and covariance $Var(Y^\Delta_t)$, which are given as follows.

(i) *In case $t_1 \leq t \leq t_n$ these are given by*

$$m^\Delta(t) = y_{j-1} + \left(\rho(t) - \rho(t_{j-1})\right)\left(\rho(t_j) - \rho(t_{j-1})\right)^{-1}\left(y_j - y_{j-1}\right),$$

$$Var(Y^\Delta_t) = \left(\rho(t_j) - \rho(t)\right)\left(\rho(t_j) - \rho(t_{j-1})\right)^{-1}\left(\rho(t) - \rho(t_{j-1})\right), \tag{8}$$

provided $t \in [t_{j-1}, t_j)$ for some $j = 1, \dots, n$.

---

[1] For any non–negative definite matrix M we write $M > 0$, if M is positive definite, such that M^{-1} exists.

(ii) *If $t < t_1$, then the respective modifications are*

$$m^{\Delta}(t) = \rho(t)\rho(t_1)^{-1}y_{t_1},$$
$$Var(Y_t^{\Delta}) = (\rho(t_1) - \rho(t))\rho(t_1)^{-1}\rho(t),$$

(which can formally be obtained from (i) *by adding $t_0 = 0$ to the design Δ with respective $\rho(t_0) = 0$).*

(iii) *If $t \geq t_n$, then $m^{\Delta}(t) = y_{t_n}$, $Var(Y_t^{\Delta}) = \rho(t) - \rho(t_n)$.*

Proof. To proof (i), it only remains to establish (8). By definition of Y_t^{Δ} the covariance evaluates as

$$Var(Y_t^{\Delta}) = \rho(t) - \mathbf{E}\left(P^{\Delta}Y_t\right)\left(P^{\Delta}Y_t\right)^T. \tag{9}$$

By orthogonality of the increments we can further conclude, say for $t \in [t_{j-1}, t_j)$, that

$$\mathbf{E}\left(P^{\Delta}Y_t\right)\left(P^{\Delta}Y_t\right)^T =$$
$$\rho(t_{j-1}) + \left(\rho(t) - \rho(t_{j-1})\right)\left(\rho(t_j) - \rho(t_{j-1})\right)^{-1}\left(\rho(t) - \rho(t_{j-1})\right).$$

Together with (9) this finally yields (8).

The case (ii) is proven analogously. The last case (iii) follows immediately, since $(Y_t)_{t\geq 0}$ has independent increments. The theorem is proven. □

Remark 1. Theorem 2 generalizes to Gauss–Markov processes the well known formula for the Brownian bridge B_t^{Δ}, which is the conditional Brownian motion B_t, conditioned at $B_\alpha = y_\alpha$ and $B_\beta = y_\beta$, $\alpha < t < \beta$,

$$m^{\Delta}(t) = \frac{(\beta - t)}{(\beta - \alpha)}y_\alpha + \frac{(t - \alpha)}{(\beta - \alpha)}y_\beta$$
$$Var(B_t^{\Delta}) = \frac{(\beta - t)(t - \alpha)}{(\beta - \alpha)}.$$

In general an explicit representation of the mean and covariance for the conditional distribution can hardly be derived as above. Instead, we may use, that interpolation is local and recourse to the representation of the conditional expectation as orthogonal projection

$$\mathbf{E}(X_t|\mathcal{F}_{t_{j-1},t_j}) = \lambda X_{t_{j-1}} + \mu X_{t_j},\ t \in [t_{j-1}, t_j), \tag{10}$$

which results in the conditions

$$X_t - \lambda X_{t_{j-1}} - \mu X_{t_j} \text{ is orthogonal to } \left\{X_{t_{j-1}}, X_{t_j}\right\},$$

leading to

$$R_{t,t_{j-1}} - \lambda R_{t_{j-1},t_{j-1}} - \mu R_{t_j,t_{j-1}} = 0$$
$$R_{t,t_j} - \lambda R_{t_{j-1},t_j} - \mu R_{t_j,t_j} = 0.$$

From this we may directly compute

$$\lambda = \Big(R_{t,t_{j-1}} R^{-1}_{t_j,t_{j-1}} R_{t_j,t_j} - R_{t,t_j}\Big)\Big(R_{t_{j-1},t_{j-1}} R^{-1}_{t_j,t_{j-1}} R_{t_j,t_j} - R_{t_{j-1},t_j}\Big)^{-1}$$

and

$$\mu = \Big(R_{t,t_j} - R_{t,t_{j-1}} R^{-1}_{t_{j-1},t_{j-1}} R_{t_{j-1},t_j}\Big)\Big(R_{t_j,t_j} - R_{t_j,t_{j-1}} R^{-1}_{t_{j-1},t_{j-1}} R_{t_{j-1},t_j}\Big)^{-1},$$

from which representations of mean and covariance can be computed.

In one space dimension, i.e., when all multiplications are commutative, one obtains as explicit formula

$$Var(X^{\Delta}_t) = \frac{(R_{t_{j-1},t_{j-1}} R_{t_j,t} - R_{t_{j-1},t_j} R_{t_{j-1},t})(R_{t_j,t_j} R_{t_{j-1},t} - R_{t_{j-1},t_j} R_{t_j,t})}{R_{t_{j-1},t_j}(R_{t_{j-1},t_{j-1}} R_{t_j,t_j} - R^2_{t_{j-1},t_j})}.$$

Appendix: Structure of Gauss–Markov processes

Here we shall consider (vector–valued) Gaussian processes $(X_t)_{t\geq 0}$, which are centered, i.e., $\mathbf{E}X_t = 0,\ t \geq 0$. Recall, that $R_{s,t} := \mathbf{E}X_s X_t^T$ denotes the covariance function, which is nonnegative definite for each s, t. For simplicity we shall assume that the covariances $R_{t,t} > 0$ are everywhere invertible. More precisely, we make the following

Basic Assumptions. There is $s_0 \geq 0$ from which on

- $R^{-1}_{s,t}$ exists for all $s, t \geq s_0$.
- The function

$$\rho_{s_0}(t) := R^{-T}_{s_0,t} R_{t,t} R^{-1}_{s_0,t}, \quad t \geq s_0,$$

  is absolutely continuous with respect to the Lebesgue measure.

Remark 2. We introduce s_0, since X_0 may not possess invertible covariance, as this is the case for standard Brownian motion.

Using the orthogonal decomposition (10) we may express the Markov property in terms of the covariance function $R_{s,t}$ as

$$R_{s,u} = R_{s,t} R^{-1}_{t,t} R_{t,u}, \quad s \leq t \leq u, \tag{11}$$

which is the multivariate variant of [12, Chap. III, Ex. 3.13].

As a result of this we conclude, that if the Basic Assumptions are satisfied for some s_0, then they are valid for every $s \geq s_0$. For this reason we shall omit the subscript s_0 henceforth.

We introduce the function $a(t) := R^T_{s_0,t}$.

It is immediate from the definition of ρ, that its values are symmetric matrices. Moreover, for a Gauss–Markov process the function ρ is non–decreasing, which means, that for $s \leq t$ the matrix $\rho(t) - \rho(s)$ is nonnegative definite. This can be seen from

$$\mathbf{E}(a(t)^{-1}X_t - a(s)^{-1}X_s)(a(t)^{-1}X_t - a(s)^{-1}X_s)^T \geq 0.$$

Since $\rho(t)$ is moreover assumed to be absolutely continuous, there is a nonnegative function $r(s)$ for which

$$\rho(t) = \rho(s_0) + \int_{s_0}^{t} r(s)\, ds. \tag{12}$$

Finally, the function r splits for some σ as $r(s) = \sigma(s)\sigma(s)^T,\ s \geq s_0$. Representation (12) gives rise to the following Gaussian process, starting at time s_0 with a Gaussian vector Y_{s_0} with covariance $\rho(s_0)$ and following with

$$Y_t := Y_{s_0} + \int_{s_0}^{t} \sigma(s)\, dB_s, \quad t > s_0, \tag{13}$$

which has independent increments and covariance $R^Y_{s,t} = \rho(\min\{s,t\})$ (The process $(B_t)_{t\geq 0}$ denotes standard Brownian motion and the integral in (13) is in the sense of Ito.) .

Our previous analysis results in the following structural result, confer also [12, Chap. III, Ex. 3.13].

Theorem 3. *Let $(X_t)_{t\geq 0}$ be a centered Gaussian process with covariance function $R_{s,t}$, satisfying the Basic Assumptions. Then the following is equivalent*

(i) $(X_t)_{t\geq s_0}$ *is a Gauss–Markov process.*
(ii) *The covariance obeys* (11)
(iii) *The covariance splits*

$$R_{s,t} = a(s)\rho(s)a(t)^T, \text{ for all } s_0 \leq s \leq t,$$

with some invertible function $a(t)$ and increasing $\rho(t)$.
(iv) *The process $\left(\bar{X}_t\right)_{t\geq s_0}$, defined through the process $(Y_t)_{t\geq s_0}$ from* (13) *as*

$$\bar{X}_t := a(t)Y_t, \quad t \geq s_0, \tag{14}$$

has the same distribution as $(X_t)_{t\geq s_0}$, i.e., $R^{\bar{X}}_{s,t} = R_{s,t},\ s,t \geq s_0$.

Sketch of the proof. The equivalence of (i) and(ii) is just equation (11).

That the covariance is of product–type follows from the definition of the involved quantities and from (11).

It is also a routine matter to compute the covariance from the process $\left(\bar{X}_t\right)_{t\geq 0}$, as defined in (14). The existence of the function σ was discussed after Remark 2. □

Remark 3. Item (iii) above may be rephrased, that the covariance function $R_{s,t}$ is of product type, as given in (2), with matrix–functions u, v, such that $v^{-1}(t)u(t),\ t \geq s_0$, is increasing. This was the situation, where results for the optimal design problem are available, see (iii) of the Introduction. One explanation may be, that the solution to the optimal design problem depends only on the covariance structure, so we may assume that the underlying process is Gaussian. Under this additional assumption, property (2) means, that the process is Markovian. As we have seen above, interpolation was local in this case and the decomposition techniques to carry out the asymptotic analysis work.

Example: Linear stochastic differential equations

Here we return to the simulation of conditional processes, arising as solutions of linear stochastic differential equations (1), mentioned in item (ii) of the Introduction as the topic of [11]. To complete the description of the process we let the initial value X_{s_0} be centered Gaussian with covariance $\rho(s_0)$.

We shall briefly indicate, that this situation is covered by Gauss–Markov processes. Of course, this can be seen from the explicit solution, see e.g. [7, 4.4], but it is illuminating to be more explicit.

Let

$$a(t) := \exp\left(\int_{s_0}^t A(s)\,ds\right), \qquad \Phi_s^t = \exp\left(\int_s^t A(u)\,du\right),$$

$$\sigma(t) := a(t)^{-1}\Sigma(t), \qquad m(t) := \int_{s_0}^t \Phi_s^t b(s)\,ds,$$

and finally

$$\rho(t) := \rho(s_0) + \int_{s_0}^t \sigma(s)\sigma(s)^T\,ds.$$

We claim, that the covariance of $(X_t)_{t\geq 0}$ is

$$R_{s,t} = a(s)\rho(s)a(t)^T, \text{ if } s_0 \leq s \leq t. \tag{15}$$

But in the light of Theorem 3 this is easy. Let $(Y_t)_{t\geq 0}$ be Gaussian with independent increments and covariance function ρ and put $\bar{X}_t :=$

$a(t)Y_t + m(t),\ t \ge s_0$. The Ito calculus, see e.g. [12, Chap. IV, §3] readily provides, that $\left(\bar{X}_t\right)_{t\ge s_0}$ then satisfies

$$\begin{aligned} d\bar{X}_t &= (a(t)'Y_t + m'(t))\,dt + a(t)dY_t \\ &= (A(t)\bar{X}_t + b(t))\,dt + a(t)\sigma(t)\,dB_t \\ &= (A(t)\bar{X}_t + b(t))\,dt + \Sigma(t)\,dB_t, \end{aligned}$$

which is equation (1). Thus $\bar{X}_t - m(t),\ t \ge s_0$ is a centered Gauss–Markov process.

On the other hand, if a Gauss–Markov process satisfies the Basic Assumptions with a function $a(t)$, which is differentiable, then it can be modelled by a linear stochastic differential equation (1) with $A(t) := a(t)'a(t)^{-1}$ and $\Sigma(t) := a(t)\sigma(t)$.

Thus our previous interpolation formulae generalize Lemma 1 of [11] to the situation of finitely many observations. In this situation, the following representations are useful. Using the matrix–function Φ_s^t from above , a little elaboration, in particular making use of (11), provides, for $t \in [t_{j-1}, t_j)$ the expressions

$$\begin{aligned} &m^{\Delta}(t) = \\ &\Big((\Phi_t^{t_j})^{-1}R_{t_j,t_j} - R_{t,t}(\Phi_t^{t_j})^T\Big)\Big((\Phi_{t_{j-1}}^{t_j})^{-1}R_{t_j,t_j} - R_{t_{j-1},t_{j-1}}(\Phi_{t_{j-1}}^{t_j})^T\Big)^{-1} x_{t_{j-1}} \\ &\qquad + \Big((\Phi_{t_{j-1}}^{t})^{-1}R_{t,t} - R_{t_{j-1},t_{j-1}}(\Phi_{t_{j-1}}^{t})^T\Big)^T \times \\ &\qquad\qquad \times \Big((\Phi_{t_{j-1}}^{t_j})^{-1}R_{t_j,t_j} - R_{t_{j-1},t_{j-1}}(\Phi_{t_{j-1}}^{t_j})^T\Big)^{-T} x_{t_j}, \end{aligned}$$

and

$$\begin{aligned} &Var(X_t^{\Delta}) = \\ &\Big((\Phi_t^{t_j})^{-1}R_{t_j,t_j} - R_{t,t}(\Phi_t^{t_j})^T\Big)\Big((\Phi_{t_{j-1}}^{t_j})^{-1}R_{t_j,t_j} - R_{t_{j-1},t_{j-1}}(\Phi_{t_{j-1}}^{t_j})^T\Big)^{-1} \\ &\qquad\qquad \times \Big((\Phi_{t_{j-1}}^{t})^{-1}R_{t,t} - R_{t_{j-1},t_{j-1}}(\Phi_{t_{j-1}}^{t})^T\Big). \end{aligned}$$

as a convenient form.

References

[1] Robert B. Ash and Melvin F. Gardner. *Topics in stochastic processes.* Academic Press [Harcourt Brace Jovanovich Publishers], New York, 1975. Probability and Mathematical Statistics, Vol. 27.

[2] Richard F. Bass. *Probabilistic techniques in analysis.* Springer-Verlag, New York, 1995.

[3] Kai Lai Chung. *Lectures from Markov processes to Brownian motion.* Springer-Verlag, New York, 1982.

[4] S. M. Ermakov and G. A. Mikhaĭlov. Статистическое моделирование. "Nauka", Moscow, second edition, 1982.

[5] William Feller. *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons Inc., New York, second edition, 1971.

[6] Wolfgang Hackenbroch and Anton Thalmaier. *Stochastische Analysis.* B. G. Teubner, Stuttgart, 1994. Eine Einführung in die Theorie der stetigen Semimartingale. [An introduction to the theory of continuous semimartingales].

[7] Peter E. Kloeden and Eckhard Platen. *Numerical solution of stochastic differential equations.* Springer-Verlag, Berlin, 1992.

[8] Peter Mathé. Optimal reconstruction of stochastic evolutions. In *The mathematics of numerical analysis (Park City, UT, 1995)*, pages 523–531. Amer. Math. Soc., Providence, RI, 1996.

[9] T. Müller-Gronbach. Optimal designs for approximating a stochastic process with respect to a minimax criterion. *Statistics*, 27:279–296, 1996.

[10] T. Müller-Gronbach. Optimal designs for approximating the path of a stochastic process. *J. Statist. Plann. Inference*, 49:371–385, 1996.

[11] S. M. Prigarin. Numerical solution of boundary value problems for linear systems of stochastic differential equations. *Zh. Vychisl. Mat. Mat. Fiz.*, 38(12):1983–1988, 1998.

[12] Daniel Revuz and Marc Yor. *Continuous martingales and Brownian motion.* Springer-Verlag, Berlin, third edition, 1999.

[13] J. Sacks and D. Ylvisaker. Design for regression problems with correlated errors. *Ann. Math. Statist.*, 37:66 – 89, 1966.

[14] J. Sacks and D. Ylvisaker. Statistical designs and integral approximation. In R. Pyke, editor, *Proc. 12th. Biennial Seminar of the Canad. Math. Congress*, pages 115–136, Montreal, 1970. Canad. Math. Society.

Weierstrass–Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, D– 10117 Berlin, Germany
E-mail address: `mathe@wias-berlin.de`

I. Mathematical Institute, Free University of Berlin, Arnimallee 2-6, D–14195 Berlin, Germany
E-mail address: `bschmidt@math.fu-berlin.de`