Chapter 10
An Introduction to Matrix Variate Stochastics

In this chapter, we introduce the reader to matrix variate stochastics. It is intended to set the scene for Wishart processes, which will be covered in the next chapter. We begin by recalling notation and introducing some basic functions used throughout both chapters. This will bring us in a position to discuss matrix valued random variables, matrix valued stochastic processes, and matrix valued stochastic differential equations. To illustrate these concepts, we apply them to the matrix valued version of the Ornstein-Uhlenbeck process and a multidimensional version of the MMM. The main references for this chapter are Gupta and Nagar (2000) and Pfaffel (2008).

10.1 Basic Definitions and Functions

In this section, we fix primarily notation.

Definition 10.1.1 We employ the following notation:

- we denote by $\mathcal{M}_{m,n}(\Re)$ the set of all $m \times n$ matrices with entries in $\Re$. If $m = n$, we write $\mathcal{M}_n(\Re)$ instead;
- we write $GL(p)$ for the group of all invertible matrices of $\mathcal{M}_p(\Re)$;
- let $\mathcal{S}_p$ denote the linear subspace of all symmetric matrices of $\mathcal{M}_p(\Re)$;
- let $\mathcal{S}_p^+(\mathcal{S}_p^-)$ denote the set of all symmetric positive (negative) definite matrices of $\mathcal{M}_p(\Re)$;
- denote by $\overline{\mathcal{S}_p^+}$ the closure of $\mathcal{S}_p^+$ in $\mathcal{M}_p(\Re)$, that is the set of all symmetric positive semidefinite matrices of $\mathcal{M}_p(\Re)$.

The next definition provides a one-to-one relationship between vectors and matrices.

J. Baldeaux, E. Platen, *Functionals of Multidimensional Diffusions with Applications to Finance*, Bocconi & Springer Series 5, DOI 10.1007/978-3-319-00747-2_10,

Definition 10.1.2 Let $\boldsymbol{A} \in \mathcal{M}_{m,n}(\Re)$ with columns $\boldsymbol{a}_i \in \Re^m$, $i = 1, \ldots, n$. Define the function $vec : \mathcal{M}_{m,n}(\Re) \to \Re^{mn}$ via

$$vec(\boldsymbol{A}) = \begin{pmatrix} \boldsymbol{a}_1 \\ \vdots \\ \boldsymbol{a}_n \end{pmatrix}.$$

Note that $vec(\boldsymbol{A})$ is also an element of $\mathcal{M}_{mn,1}(\Re)$. The next lemma is derived in Gupta and Nagar (2000).

Lemma 10.1.3 *The following properties hold*:

- *for* $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{M}_{m,n}(\Re)$ *it holds that* $tr(\boldsymbol{A}^\top \boldsymbol{B}) = vec(\boldsymbol{A})^\top vec(\boldsymbol{B})$;
- *let* $\boldsymbol{A} \in \mathcal{M}_{p,m}(\Re)$, $\boldsymbol{B} \in \mathcal{M}_{m,n}(\Re)$ *and* $\boldsymbol{C} \in \mathcal{M}_{n,q}(\Re)$. *Then we have*

$$vec(\boldsymbol{AXB}) = \left(\boldsymbol{B}^\top \otimes \boldsymbol{A}\right) vec(\boldsymbol{X}).$$

We now recall from Gupta and Nagar (2000) how a symmetric matrix can be mapped to a vector.

Definition 10.1.4 Let $\boldsymbol{S} \in \mathcal{S}_p$. Define the function $vech : \mathcal{S}_0 \to \Re^{\frac{p(p+1)}{2}}$ via

$$vech(\boldsymbol{S}) = \begin{pmatrix} S_{11} \\ S_{12} \\ S_{22} \\ \vdots \\ S_{1p} \\ \vdots \\ S_{pp} \end{pmatrix},$$

such that $vech(\boldsymbol{S})$ is a vector consisting of the elements of $\boldsymbol{S}$ from above and including the diagonal, taken componentwise.

We point out that the vector $vech$ gives access to the $\frac{p(p+1)}{2}$ distinct values of a symmetric $p \times p$ matrix.

10.2 Integrals over Matrix Domains

The aim of this section is to define integrals over matrix domains. These definitions will be employed in the subsequent sections, e.g. when computing characteristic functions and Laplace transforms of matrix valued random variables. Discussing integration, we need a notion of measurability. The following definition is taken from Pfaffel (2008), see also Jacod and Protter (2004).

Definition 10.2.1 Let $(X, \mathcal{T})$ be a topological space. The Borel σ-algebra on X is then given by the smallest σ-algebra that contains $\mathcal{T}$ and is denoted by $\mathcal{B}(X)$.

In this chapter and the following, we focus on $\Re$, $\Re^n$, and $\mathcal{M}_{m,n}(\Re)$, and employ the notation $\mathcal{B}$ for $\mathcal{B}(\Re)$, $\mathcal{B}^n$ for $\mathcal{B}(\Re^n)$ and $\mathcal{B}^{m,n}$ for $\mathcal{B}(\mathcal{M}_{m,n}(\Re))$. We are now in a position to define integrals over matrices, allowing for matrices of size $m \times n$.

Definition 10.2.2 Let $f : \mathcal{M}_{m,n}(\Re) \to \Re$ be a $\mathcal{B}^{m,n} - \mathcal{B}$-measurable function and $M \in \mathcal{B}^{m,n}$ a measurable subset of $\mathcal{M}_{m,n}(\Re)$ and let λ denote the Lebesgue-measure on $(\Re^{mn}, \mathcal{B}^{mn})$. The integral of f over $\boldsymbol{M}$ is then defined by

$$\int_M f(\boldsymbol{X})\, d\boldsymbol{X} := \int_M f(\boldsymbol{X})\, d(\lambda \circ vec)(\boldsymbol{X}) = \int_{vec(M)} f \circ vec^{-1}(\boldsymbol{x})\, d\lambda(\boldsymbol{x}).$$

We call $\lambda \circ vec$ the Lebesgue-measure on $(\mathcal{M}_{m,n}(\Re), \mathcal{B}^{m,n})$.

As pointed out in Pfaffel (2008), $\mathcal{S}_p$ is isomorphic to $\Re^{\frac{p(p+1)}{2}}$, hence for $p \geq 2$, $\mathcal{S}_p$ is a real subspace of $\mathcal{M}_p(\Re)$, and, consequently, of Lebesgue-measure zero. As this means that any integral over subsets of $\mathcal{S}_p$ is zero, we define another Lebesgue measure on the subspace of symmetric matrices $\mathcal{S}_p$:

Definition 10.2.3 Let $f : \mathcal{S}_p \to \Re$ be a $\mathcal{B}(\mathcal{S}_p)$- $\mathcal{B}$-measurable function and $M \in \mathcal{B}(\mathcal{S}_p)$ a Borel-measurable subset of $\mathcal{S}_p$ and let λ denote the Lebesgue-measure on $(\Re^{\frac{p(p+1)}{2}}, \mathcal{B}^{\frac{p(p+1)}{2}})$. The integral of f over M is then defined by

$$\int_M f(\boldsymbol{X})\, d\boldsymbol{X} := \int_M f(\boldsymbol{X})\, d(\lambda \circ vech)(\boldsymbol{X}) = \int_{vech(M)} f \circ vech^{-1}(\boldsymbol{x})\, d\lambda(\boldsymbol{x}).$$

We call $\lambda \circ vech$ the Lebesgue-measure on $(\mathcal{S}_p, \mathcal{B}(\mathcal{S}_p))$.

As in Gupta and Nagar (2000) and Pfaffel (2008), we use the notation

$$etr(\boldsymbol{A}) := \exp\{tr(\boldsymbol{A})\}. \tag{10.2.1}$$

This notation allows the formulation of the next definition.

Definition 10.2.4 The multivariate gamma function is defined as follows:

$$\Gamma_p(a) := \int_{\mathcal{S}_p^+} etr(-\boldsymbol{A}) det(\boldsymbol{A})^{a-\frac{1}{2}(p+1)}\, d\boldsymbol{A} \quad \forall a > \frac{p-1}{2}.$$

The next result from Gupta and Nagar (2000), shows that for $a > \frac{p-1}{2}$, the matrix variate gamma function can be expressed as a finite product of ordinary gamma functions.

Theorem 10.2.5 *For* $a > \frac{1}{2}(p-1)$,

$$\Gamma_p(a) = \pi^{\frac{1}{4}p(p-1)} \prod_{i=1}^{p} \Gamma\left(a - \frac{1}{2}(i-1)\right).$$

Next, we want to introduce hypergeometric functions of matrix arguments. This requires the definition of zonal polynomials, which in turn requires the definition of *symmetric homogeneous polynomials*. Starting with the latter, a symmetric homogeneous polynomial of degree k in $y_1, \ldots, y_m$ is a polynomial which is unchanged by a permutation of the subscripts and such that every term in the polynomial has degree k. The following example from Muirhead (1982) illustrates this: set $m = 2$, $k = 3$, then

$$y_1^3 + y_2^3 + 10y_1^2 y_2 + 10y_1 y_2^2$$

is a symmetric homogeneous polynomial of degree 3 in y_1 and y_2. Following Gupta and Nagar (2000), we denote by V_k the vector space of symmetric homogeneous polynomials that are of degree k in the $\frac{1}{2}p(p-1)$ distinct elements of $\boldsymbol{S} \in \mathcal{S}_p^+$. As discussed in Gupta and Nagar (2000), the space V_k can be decomposed into a direct sum of irreducible invariant subspaces V_κ, where κ denotes a partition of k, defined as follows: by a partition of k, we mean the p-tuple $\kappa = (k_1, \ldots, k_p)$, where $k_1 \geq \cdots \geq k_p \geq 0$, and furthermore $k_1 + \cdots + k_p = k$. Then the polynomial $tr(\boldsymbol{S})^k \in V_k$ has the unique decomposition into polynomials $C_\kappa(\boldsymbol{S}) \in V_\kappa$ as

$$tr(\boldsymbol{S})^k = \sum_\kappa C_\kappa(\boldsymbol{S}).$$

We now define *zonal polynomials*.

Definition 10.2.6 The *zonal polynomial* $C_\kappa(\boldsymbol{S})$ is the component of $tr(\boldsymbol{S})^k$ in the subspace V_κ.

The next definition from Gupta and Nagar (2000) introduces *hypergeometric functions of matrix arguments*.

Definition 10.2.7 The *hypergeometric function of matrix argument* is defined by

$$_mF_n(a_1, \ldots, a_m; b_1, \ldots, b_n; \boldsymbol{S}) = \sum_{k=0}^{\infty} \sum_\kappa \frac{(a_1)_\kappa \ldots (a_m)_\kappa C_\kappa(\boldsymbol{S})}{(b_1)_\kappa \ldots (b_n)_\kappa k!}, \tag{10.2.2}$$

where $a_i, b_j \in \Re$, $\boldsymbol{S}$ is a symmetric $p \times p$-matrix and $\sum_\kappa$ the summation over all partitions κ of k and $(a)_\kappa = \prod_{j=1}^p (a - \frac{1}{2}(j-1))_{k_j}$ denotes the generalized hypergeometric coefficient, with $(x)_{k_j} = x(x+1)\ldots(x+k_j-1)$.

The following remark provides some properties of hypergeometric functions of matrix arguments.

Remark 10.2.8 Conditions for convergence of the infinite series in Eq. (10.2.2) are of importance, see Gupta and Nagar (2000) for a discussion. The condition $m < n + 1$ is sufficient. We also have the special case

$$_nF_n(a_1, \ldots, a_n; a_1, \ldots, a_n; \boldsymbol{S}) = \sum_{k=0}^{\infty} \frac{(tr(\boldsymbol{S}))^k}{k!} = etr(\boldsymbol{S}).$$

The next lemma will be subsequently employed when computing expectations of functions of non-central Wishart distributed random variables.

Lemma 10.2.9 *Let* $\boldsymbol{Z}, \boldsymbol{T} \in \mathcal{S}_p^+$. *Then*

$$\begin{aligned}&\int_{S_p^+} etr(-\boldsymbol{Z}\boldsymbol{S}) det(\boldsymbol{S})^{a-\frac{p+1}{2}}{}_m F_n(a_1, \ldots, a_m; b_1, \ldots, b_n; \boldsymbol{S}\boldsymbol{T})\, d\boldsymbol{S} \\ &\quad = \Gamma_p(a) det(\boldsymbol{Z})^{-a}{}_{m+1} F_n\big(a_1, \ldots, a_m; b_1, \ldots, b_n; \boldsymbol{Z}^{-1}\boldsymbol{T}\big),\end{aligned}$$

$\forall a > \frac{p-1}{2}$.

Proof The result is a special case of Theorem 1.6.2 in Gupta and Nagar (2000). □

10.3 Matrix Valued Random Variables

In this section, we discuss matrix valued random variables. First, we need to define what we mean by a matrix valued random variable, and consequently associate with it the concepts well-known from the vector and scalar case, such as probability density functions, characteristic functions, and Laplace transforms. We will discuss two examples, first the normal distribution and second the Wishart distribution. The main reference for this section is Gupta and Nagar (2000), see also Pfaffel (2008). As before, we use $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ to denote the filtered probability space.

Definition 10.3.1 An $m \times n$ random matrix $\boldsymbol{X}$ is a measurable function

$$\boldsymbol{X} : (\Omega, \mathcal{F}) \to \big(\mathcal{M}_{m,n}(\Re), \mathcal{B}^{m,n}\big).$$

We now discuss probability density functions of random variables.

Definition 10.3.2 A nonnegative measurable function $f_{\boldsymbol{X}}$ such that

$$P(\boldsymbol{X} \in M) = \int_M f_{\boldsymbol{X}}(\boldsymbol{A})\, d\boldsymbol{A} \quad \forall M \in \mathcal{B}^{m\times n}$$

defines the probability density function of an $m \times n$ random matrix $\boldsymbol{X}$.

We can now introduce expected values.

Definition 10.3.3 Let $\boldsymbol{X}$ be an $m \times n$-random matrix. For every function $h = (h_{i,j})_{i,j} : \mathcal{M}_{m,n}(\Re) \to \mathcal{M}_{r,s}(\Re)$ with $h_{i,j} : \mathcal{M}_{m,n}(\Re) \to \Re$, $1 \le i \le r$, $1 \le j \le s$, the expected value $E(h(\boldsymbol{X}))$ of $h(\boldsymbol{X})$ is an element of $\mathcal{M}_{r,s}(\Re)$ with elements

$$E\big(h(\boldsymbol{X})\big)_{i,j} = E\big(h_{i,j}(\boldsymbol{X})\big) = \int_{\mathcal{M}_{m,n}(\Re)} h_{i,j}(\boldsymbol{A}) P^{\boldsymbol{X}}(d\boldsymbol{A}).$$

We point out that if $\boldsymbol{X}$ has a probability density function $f_{\boldsymbol{X}}$, then we have

$$E\big(h(\boldsymbol{X})\big)_{i,j} = \int_{\mathcal{M}_{m,n}(\Re)} h_{i,j}(\boldsymbol{A}) f_{\boldsymbol{X}}(\boldsymbol{A})\, d\boldsymbol{A}.$$

The characteristic function or Fourier transform of matrix-valued random variables is now defined.

Definition 10.3.4 The characteristic function of an $m \times n$-random matrix $\boldsymbol{X}$ with probability density function $f_{\boldsymbol{X}}$ is defined as

$$E\big(etr\big(\imath \boldsymbol{X}\boldsymbol{Z}^\top\big)\big) = \int_{\mathcal{M}_{m,n}(\Re)} etr\big(\imath \boldsymbol{A}\boldsymbol{Z}^\top\big) f_{\boldsymbol{X}}(\boldsymbol{A})\, d\boldsymbol{A}, \tag{10.3.3}$$

for every $\boldsymbol{Z} \in \mathcal{M}_p(\Re)$.

Due to the fact that $|\exp(\imath x)| = 1,\ \forall x \in \Re$, the integral in (10.3.3) always exists. Furthermore, (10.3.3) is the Fourier transform of the measure $P^{\boldsymbol{X}}$ at point $\boldsymbol{Z}$.

Definition 10.3.5 The Laplace transform of a $p \times p$-random matrix $\boldsymbol{X} \in \mathcal{S}_p^+$ with probability density function $f_{\boldsymbol{X}}$ is defined as

$$E\big(etr(-\boldsymbol{U}\boldsymbol{X})\big) = \int_{\mathcal{S}_p^+} etr(-\boldsymbol{U}\boldsymbol{A}) f_{\boldsymbol{X}}(\boldsymbol{A})\, d\boldsymbol{A}, \tag{10.3.4}$$

for every $\boldsymbol{U} \in \mathcal{S}_p^+$.

Remark 10.3.6 Recall that the Laplace transform of a positive scalar random variable is always well-defined. For $\boldsymbol{A}, \boldsymbol{U} \in \mathcal{S}_p^+$, we have that $tr(-\boldsymbol{U}\boldsymbol{A}) = -tr(\sqrt{\boldsymbol{U}}\boldsymbol{A}\sqrt{\boldsymbol{U}}) < 0$, since $\sqrt{\boldsymbol{U}}\boldsymbol{A}\sqrt{\boldsymbol{U}}$ is positive definite, hence the integral in Eq. (10.3.4) is well-defined, where $\boldsymbol{X} \in \mathcal{S}_p^+$ is the analogue of a positive random variable.

Next, we introduce covariance matrices for matrix valued random variables.

Definition 10.3.7 Let $\boldsymbol{X}$ be an $m \times n$ random matrix and $\boldsymbol{Y}$ be a $p \times q$ random matrix. Then the $mn \times pq$ covariance matrix is defined as

$$\begin{aligned} cov(\boldsymbol{X}, \boldsymbol{Y}) &= cov\big(vec\big(\boldsymbol{X}^\top\big), vec\big(\boldsymbol{Y}^\top\big)\big) \\ &= E\big(vec\big(\boldsymbol{X}^\top\big)vec\big(\boldsymbol{Y}^\top\big)^\top\big) - E\big(vec\big(\boldsymbol{X}^\top\big)\big)E\big(vec\big(\boldsymbol{Y}^\top\big)\big)^\top, \end{aligned}$$

i.e. $cov(\boldsymbol{X}, \boldsymbol{Y})$ is an $m \times p$ block matrix with blocks $cov(\tilde{\boldsymbol{x}}_i^\top, \tilde{\boldsymbol{y}}_j^\top) \in \mathcal{M}_{n,q}(\Re)$ where $\tilde{\boldsymbol{x}}_i$ (or $\tilde{\boldsymbol{y}}_i$ respectively) denote the rows of $\boldsymbol{X}$ (respectively $\boldsymbol{Y}$).

Having these definitions at hand, we can now discuss some examples. We begin with an example involving the normal distribution.

Definition 10.3.8 A $p \times n$ random matrix is said to have a matrix variate normal distribution with mean $\boldsymbol{M} \in \mathcal{M}_{p,n}(\Re)$ and covariance $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$, where $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$, $\boldsymbol{\Psi} \in \mathcal{S}_n^+$, if $vec(\boldsymbol{X}^\top) \sim \mathcal{N}_{pn}(vec(\boldsymbol{M}^\top), \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$, where $\mathcal{N}_{pn}$ denotes the multivariate normal distribution on $\Re^{pn}$ with mean $vec(\boldsymbol{M}^\top)$ and covariance $\boldsymbol{\Sigma} \otimes \boldsymbol{\Psi}$. We will use the notation $\boldsymbol{X} \sim \mathcal{N}_{p,n}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$.

We now recall a result from Gupta and Nagar (2000).

Theorem 10.3.9 *If $\boldsymbol{X} \sim \mathcal{N}_{p,n}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$, then $\boldsymbol{X}^\top \sim \mathcal{N}_{p,n}(\boldsymbol{M}^\top, \boldsymbol{\Psi} \otimes \boldsymbol{\Sigma})$.*

Proof The proof is given by the one of Theorem 2.3.1 in Gupta and Nagar (2000). □

The next result gives the characteristic function of the normal distribution.

Theorem 10.3.10 *Let $\boldsymbol{X} \sim \mathcal{N}_{p,n}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$. Then the characteristic function of $\boldsymbol{X}$ is given by*

$$E\left(etr\left(\imath \boldsymbol{X}\boldsymbol{Z}^\top\right)\right) = etr\left(\imath \boldsymbol{Z}^\top \boldsymbol{M} - \frac{1}{2}\boldsymbol{Z}^\top \boldsymbol{\Sigma}\boldsymbol{Z}\boldsymbol{\Psi}\right).$$

By employing Theorem 10.3.10 one proves the matrix analogue of the linear transformation property of normal random variables, see Pfaffel (2008).

Theorem 10.3.11 *Let $\boldsymbol{X} \sim \mathcal{N}_{p,n}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{\Psi})$, $\boldsymbol{A} \in \mathcal{M}_{m,q}(\Re)$, $\boldsymbol{B} \in \mathcal{M}_{m,p}(\Re)$ and $\boldsymbol{C} \in \mathcal{M}_{n,q}(\Re)$. Then $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{X}\boldsymbol{C} \sim \mathcal{N}_{m,q}(\boldsymbol{A} + \boldsymbol{B}\boldsymbol{M}\boldsymbol{C}, (\boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}^\top) \otimes (\boldsymbol{C}^\top \boldsymbol{\Psi}\boldsymbol{C}))$.*

Next, we discuss an example involving the Wishart distribution.

Definition 10.3.12 A $p \times p$-random matrix $\boldsymbol{X}$ in $\mathcal{S}_p^+$ is said to have a non-central Wishart distribution with parameters $p \in \mathcal{N}$, $n \geq p$, $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$ and $\boldsymbol{\Theta} \in \mathcal{M}_p(\Re)$, if its probability density function is of the form

$$f_{\boldsymbol{X}}(\boldsymbol{S}) = \left(2^{\frac{1}{2}np}\Gamma_p\left(\frac{n}{2}\right)det(\boldsymbol{\Sigma})^{\frac{n}{2}}\right)^{-1} etr\left(-\frac{1}{2}\left(\boldsymbol{\Theta} + \boldsymbol{\Sigma}^{-1}\boldsymbol{S}\right)\right)$$
$$\times\, det(\boldsymbol{S})^{\frac{1}{2}(n-p-1)}\,{}_0F_1\left(\frac{n}{2}; \frac{1}{4}\boldsymbol{\Theta}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\right),$$

where $\boldsymbol{S} \in \mathcal{S}_p^+$ and ${}_0F_1$ is the hypergeometric function. We write

$$\boldsymbol{X} \sim \mathcal{W}_p(n, \boldsymbol{\Sigma}, \boldsymbol{\Theta}).$$

We remark that the requirement $n \geq p$ ensures that the matrix variate gamma function is well-defined. If $\boldsymbol{\Theta} = \boldsymbol{0}$, $\boldsymbol{X}$ is said to follow the central Wishart distribution with parameters p, n and $\boldsymbol{\Sigma} \in \mathcal{S}_p^+$, with probability density function

$$\left(2^{\frac{1}{2}np}\Gamma_p\left(\frac{n}{2}\right)det(\boldsymbol{\Sigma})^{\frac{n}{2}}\right)^{-1} etr\left(-\frac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{S}\right)det(\boldsymbol{S})^{\frac{1}{2}(n-p-1)},$$

where $\boldsymbol{S} \in \mathcal{S}_p^+$ and $n \geq p$. Next we provide the Laplace transform of the non-central Wishart distribution, see Pfaffel (2008).

Theorem 10.3.13 *Let $\boldsymbol{S} \sim \mathcal{W}_p(n, \boldsymbol{\Sigma}, \boldsymbol{\Theta})$. Then the Laplace transform of $\boldsymbol{S}$ is given by*

$$E\big(etr(-\boldsymbol{U}\boldsymbol{S})\big) = det(\boldsymbol{I}_p + 2\boldsymbol{\Sigma}\boldsymbol{U})^{-\frac{n}{2}} etr\big(-\boldsymbol{\Theta}(\boldsymbol{I}_p + 2\boldsymbol{\Sigma}\boldsymbol{U})^{-1}\boldsymbol{\Sigma}\boldsymbol{U}\big)$$

with $\boldsymbol{U} \in \mathcal{S}_p^+$.

Now, we list the characteristic function of the non-central Wishart distribution according to Gupta and Nagar (2000).

Theorem 10.3.14 *Let $\boldsymbol{S} \sim \mathcal{W}_p(n, \boldsymbol{\Sigma}, \boldsymbol{\Theta})$. Then the characteristic function of $\boldsymbol{S}$ is given by*

$$E\big(etr(\imath\boldsymbol{Z}\boldsymbol{S})\big) = det(\boldsymbol{I}_p - 2\imath\boldsymbol{\Sigma}\boldsymbol{Z})^{-\frac{n}{2}} etr\big(\imath\boldsymbol{\Theta}(\boldsymbol{I}_p - 2\imath\boldsymbol{\Sigma}\boldsymbol{Z})^{-1}\boldsymbol{\Sigma}\boldsymbol{Z}\big),$$

with $\boldsymbol{Z} \in \mathcal{M}_p(\Re)$.

The next result, which is Theorem 3.5.1 in Gupta and Nagar (2000), shows that the Wishart distribution is the matrix analogue of the non-central χ^2-distribution.

Theorem 10.3.15 *Let $\boldsymbol{X} \sim \mathcal{N}_{p,n}(\boldsymbol{M}, \boldsymbol{\Sigma} \otimes \boldsymbol{I}_n)$, $n \in \{p, p+1, \ldots\}$. Then $\boldsymbol{X}\boldsymbol{X}^\top \sim \mathcal{W}_p(n, \boldsymbol{\Sigma}, \boldsymbol{\Sigma}^{-1}\boldsymbol{M}\boldsymbol{M}^\top)$.*

We remark that $\boldsymbol{\Sigma}$ can be interpreted as a scale parameter and $\boldsymbol{\Theta}$ as a location parameter for $\boldsymbol{S}$. Consequently, a central Wishart distributed matrix is the square of normally distributed matrix random variables with zero mean.

10.4 Matrix Valued Stochastic Processes

This section closely follows Sect. 3.3 in Pfaffel (2008). First, we define matrix valued stochastic processes. Our first example will be matrix valued Brownian motion. Later, we will introduce matrix valued local martingales and semimartingales, which then allow us to formulate an Itô formula for matrix valued semimartingales. The section concludes with an integration-by-parts formula, which is useful when applying the theory presented in this chapter to examples, such as the Ornstein-Uhlenbeck process. We remind the reader that $\Re^+$ refers to the interval of non-negative real numbers $[0, \infty)$.

Definition 10.4.1 A measurable function $\boldsymbol{X} : \Re^+ \times \Omega \to \mathcal{M}_{m,n}(\Re), (t, \omega) \mapsto \boldsymbol{X}(t, \omega) = \boldsymbol{X}_t(\omega)$ is called a matrix valued stochastic process if $\boldsymbol{X}(t, \omega)$ is a random matrix for all $t \in \Re^+$. Moreover, $\boldsymbol{X}$ is called a stochastic process in $\overline{\mathcal{S}_p^+}$ if $\boldsymbol{X} : \Re_+ \times \Omega \to \overline{S_P^+}$.

As noticed in Pfaffel (2008), most definitions applicable to scalar processes can be transferred to matrix valued processes by demanding that they apply to every element of the matrix. The first example is Brownian motion.

Definition 10.4.2 A matrix valued Brownian motion $\boldsymbol{W}$ in $\mathcal{M}_{n,p}(\Re)$ is a matrix consisting of independent one-dimensional Brownian motions, i.e. $\boldsymbol{W} = (W_{i,j})_{i,j}$, where $W_{i,j}$ are independent one-dimensional Brownian motions, $1 \leq i \leq n$, $1 \leq j \leq p$. We write $\boldsymbol{W} \sim \mathcal{BM}_{n,p}$ and $\boldsymbol{W} \sim \mathcal{BM}_n$ if $p = n$.

We now show the obvious distributional properties of Brownian motion.

Corollary 10.4.3 *The following distributional properties regarding a matrix valued Brownian motion* $\boldsymbol{W} = \{\boldsymbol{W}_t,\ t \geq 0\}$ *hold*:

- $\boldsymbol{W}_t \sim \mathcal{N}_{n,p}(0, t\boldsymbol{I}_{np})$;
- $\boldsymbol{W} \sim \mathcal{BM}_{n,}$, $\boldsymbol{A} \in \mathcal{M}_{m,q}(\Re)$, $\boldsymbol{B} \in \mathcal{M}_{m,n}(\Re)$ *and* $\boldsymbol{C} \in \mathcal{M}_{p,q}(\Re)$. *Then* $\boldsymbol{A} + \boldsymbol{B}\boldsymbol{W}_t\boldsymbol{C} \sim \mathcal{N}_{m,q}(\boldsymbol{A}, t(\boldsymbol{B}\boldsymbol{B}^\top) \otimes (\boldsymbol{C}^\top\boldsymbol{C}))$.

Proof For the first part, we need to show that $vec(\boldsymbol{W}_t^\top) \sim \mathcal{N}_{np}(0, t\boldsymbol{I}_{np})$, which is easily verified. The second part follows from Theorem 10.3.11 and the observation that $\boldsymbol{I}_{np} = \boldsymbol{I}_n \otimes \boldsymbol{I}_p$. □

We now define a matrix valued local martingale.

Definition 10.4.4 A matrix valued stochastic process $\boldsymbol{X}$ is called a local martingale, if each component of $\boldsymbol{X}$ is a local martingale, i.e. if there exists a sequence of strictly monotonic increasing stopping times $(T_n)_{n\in\mathcal{N}}$, where $T_n \xrightarrow{a.s.} \infty$, such that $\boldsymbol{X}_{\min(t,T_n),ij}$ forms a martingale for all $i, j, t \geq 0$ and $n \in \{1, 2, \ldots\}$.

The next result is the analogue of Lévy's theorem, which allows us to decide if a given matrix valued continuous local martingale is a Brownian motion. This result appeared in Pfaffel (2008).

Theorem 10.4.5 *Let* $\boldsymbol{B}$ *be a* $p \times p$ *dimensional continuous local martingale such that*

$$[B_{i,j}, B_{k,l}]_t = \begin{cases} t & \text{if } i = k \text{ and } j = l \\ 0 & \text{else} \end{cases}$$

for all $i, j, k \in \{1, \ldots, p\}$. *Then* $\boldsymbol{B}$ *is a* $p \times p$*-dimensional Brownian motion,* $\boldsymbol{B} \sim \mathcal{BM}_p$.

Given the definition of a local martingale, as in the scalar case, we can define semimartingales.

Definition 10.4.6 A matrix valued stochastic process $\boldsymbol{X}$ is called a semimartingale if $\boldsymbol{X}$ can be decomposed into $\boldsymbol{X} = \boldsymbol{X}_0 + \boldsymbol{M} + \boldsymbol{A}$, where $\boldsymbol{M}$ is a local martingale and $\boldsymbol{A}$ an adapted process of finite variation.

We can now consider stochastic integrals. As in the scalar and vector case, we focus on continuous semimartingales, and for an $n \times p$-dimensional Brownian motion $\boldsymbol{W} \sim \mathcal{BM}_{n,p}$, stochastic processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ in $\mathcal{M}_{m,n}(\Re)$ and $\mathcal{M}_{p,q}(\Re)$, respectively, and a stopping time T, the matrix variate stochastic integral on $[0, T]$ is a matrix with entries

$$\left(\int_0^T \boldsymbol{X}_t \, d\boldsymbol{W}_t \, \boldsymbol{Y}_t\right)_{i,j} = \sum_{k=1}^{n} \sum_{l=1}^{p} \int_0^T X_{t,ik} Y_{t,lj} \, dW_{t,kl}, \quad \forall 1 \le i \le m, \ 1 \le j \le q.$$

We are now able to state an Itô formula for matrix-variate semimartingales, see Pfaffel (2008).

Theorem 10.4.7 *Let $U \subseteq \mathcal{M}_{m,n}(\Re)$ be open, $\boldsymbol{X}$ a continuous semimartingale with values in U and let $f : U \to \Re$ be a twice continuously differentiable function. Then $f(\boldsymbol{X})$ is a continuous semimartingale and*

$$f(\boldsymbol{X}_t) = f(\boldsymbol{X}_0) + tr\left(\int_0^t Df(\boldsymbol{X}_s)^\top \, d\boldsymbol{X}_s\right) + \frac{1}{2} \int_0^t \sum_{j,l=1}^{n} \sum_{i,k=1}^{n} \frac{\partial^2}{\partial X_{i,j} \partial X_{k,l}} f(\boldsymbol{X}_s) \, d[X_{i,j}, X_{k,l}]_s \qquad (10.4.5)$$

with $\boldsymbol{D} = (\frac{\partial}{\partial X_{i,j}})_{i,j}$.

The next corollary is given in Pfaffel (2008).

Corollary 10.4.8 *Let $\boldsymbol{X}$ be a continuous semimartingale on a stochastic interval $[0, T]$ with $T = \inf\{t : \boldsymbol{X}_t \notin U\} > 0$ for an open set $U \subseteq \mathcal{M}_{m,n}(\Re)$ and let $f : U \to \Re$ be a twice continuously differentiable function. Then $(f(\boldsymbol{X}_t))_{t \in [0,T]}$ is a continuous semimartingale and (10.4.5) holds for $t \in [0, T)$.*

In order to state a matrix valued integration by parts formula, we need the following definition of covariation for matrix valued stochastic processes.

Definition 10.4.9 For two semimartingales $\boldsymbol{A} \in \mathcal{M}_{d,m}(\Re)$, $\boldsymbol{B} \in \mathcal{M}_{m,n}(\Re)$ the matrix valued quadratic covariation is defined by

$$[\boldsymbol{A}, \boldsymbol{B}]^M_{t,ij} = \sum_{k=1}^{m} [A_{i,k}, B_{k,j}]_t \in \mathcal{M}_{d,n}(\Re).$$

The following integration-by-parts formula will be useful, see Pfaffel (2008).

Theorem 10.4.10 *Let $\boldsymbol{A} \in \mathcal{M}_{d,m}(\Re)$, $\boldsymbol{B} \in \mathcal{M}_{m,n}(\Re)$ be two semimartingales. Then the matrix product $\boldsymbol{A}_t \boldsymbol{B}_t \in \mathcal{M}_{d,n}(\Re)$ is a semimartingale and*

$$\boldsymbol{A}_t \boldsymbol{B}_t = \boldsymbol{A}_0 \boldsymbol{B}_0 + \int_0^t \boldsymbol{A}_s \, d\boldsymbol{B}_s + \int_0^t d\boldsymbol{A}_s \, \boldsymbol{B}_s + [\boldsymbol{A}, \boldsymbol{B}]^M_t.$$

10.5 Matrix Valued Stochastic Differential Equations

In this section, we briefly discuss matrix valued SDEs. The aim is to be able to make sense of the SDEs presented later describing Wishart processes. We follow Pfaffel (2008) and Stelzer (2007), where additional material on matrix valued Markov processes is presented.

As with scalar valued SDEs, we can distinguish between strong and weak solutions. Recall that a strong solution can roughly be thought of as a function of a given Brownian motion. The next definition can be found in Pfaffel (2008).

Definition 10.5.1 Let $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ be a filtered probability space satisfying the usual conditions and consider the stochastic differential equation

$$d\boldsymbol{X}_t = b(t, \boldsymbol{X}_t)\, dt + \sigma(t, \boldsymbol{X}_t)\, d\boldsymbol{W}_t, \tag{10.5.1}$$

where $\boldsymbol{X}_0 = \boldsymbol{x}_0$, $b : \Re_+ \times \mathcal{M}_{m,n}(\Re) \to \mathcal{M}_{m,n}(\Re)$ and $\sigma : \Re_+ \times \mathcal{M}_{m,n}(\Re) \to \mathcal{M}_{m,p}(\Re)$ are measurable functions, $\boldsymbol{x}_0 \in \mathcal{M}_{m,n}(\Re)$ and $\boldsymbol{W}$ is a $p \times n$-matrix valued Brownian motion.

(i) A pair $(\boldsymbol{X}, \boldsymbol{W})$ of $\mathcal{A}_t$-adapted continuous processes defined on $(\Omega, \mathcal{A}, \underline{\mathcal{A}}, P)$ is called a solution of the SDE (10.5.1) on $[0, T)$, $T > 0$, if $\boldsymbol{W}$ is an $\underline{\mathcal{A}}$-Brownian motion and

$$\boldsymbol{X}_t = \boldsymbol{x}_0 + \int_0^t b(s, \boldsymbol{X}_s)\, ds + \int_0^t \sigma(s, \boldsymbol{X}_s)\, d\boldsymbol{W}_s \quad \forall t \in [0, T).$$

(ii) Moreover, the pair $(\boldsymbol{X}, \boldsymbol{W})$ is said to be a strong solution of (10.5.1), if $\boldsymbol{X}$ is adapted to the filtration $(\mathcal{A}_t^W)_{t \in \Re_+}$, where $\mathcal{G}_t^W = \sigma_c(W_s, s \le t)$ is the σ-algebra generated by $\boldsymbol{W}_s$, $s \le t$, completed with all P-null sets from $\mathcal{A}$.

(iii) A solution $(\boldsymbol{X}, \boldsymbol{W})$ of the SDE (10.5.1), which is not a strong solution is termed a weak solution of Eq. (10.5.1).

We now discuss the existence of a solution. As in the scalar case, the local Lipschitz condition turns out to suffice, see Stelzer (2007).

Definition 10.5.2 Let $(U, \|\cdot\|_U)$, $(V, \|\cdot\|_V)$ be two normed spaces and $A \subseteq U$ be open. Then a function $f : A \to V$ is called locally Lipschitz, if for every $\boldsymbol{x} \in A$ there exists an open neighborhood $\mathcal{U}(\boldsymbol{x}) \subset A$ and a constant $C(\boldsymbol{x}) \in \Re^+$ such that

$$\left\| f(\boldsymbol{z}) - f(\boldsymbol{y}) \right\|_V \le C(\boldsymbol{x}) \|\boldsymbol{z} - \boldsymbol{y}\|_U \quad \forall \boldsymbol{z}, \boldsymbol{y} \in \mathcal{U}(\boldsymbol{x}).$$

We term $C(\boldsymbol{x})$ the local Lipschitz coefficient. If there is a $K \in \Re^+$ such that $C(\boldsymbol{x}) = K$ can be chosen for all $\boldsymbol{x} \in A$, then f is called globally Lipschitz.

The next theorem states that a local Lipschitz condition is a sufficient condition.

Theorem 10.5.3 *Let U be an open subset of $\mathcal{M}_{d,n}(\Re)$ and $(U_n)_{n \in \mathcal{N}}$ a sequence of convex closed sets such that $U_n \subset U$, $U_n \subseteq U_{n+1}$ $\forall n \in \mathcal{N}$ and $\bigcup_{n \in \mathcal{N}} U_n = U$.*

Assume that $f : U \to \mathcal{M}_{d,m}(\Re)$ is locally Lipschitz and $\boldsymbol{Z}$ in $\mathcal{M}_{m,n}(\Re)$ is a continuous semimartingale. Then for each U-valued $\mathcal{F}_0$-measurable initial value $\boldsymbol{X}_0$ there exists a stopping time T and a unique U-valued strong solution $\boldsymbol{X}$ to the stochastic differential equation

$$d\boldsymbol{X}_t = f(\boldsymbol{X}_t)\, d\boldsymbol{Z}_t \tag{10.5.2}$$

up to the time $T > 0$ a.s., i.e. on the stochastic interval $[0, T)$. At $T < \infty$ we have that either $\boldsymbol{X}$ hits the boundary ∂U of U at T, i.e. $\boldsymbol{X}_T \in \partial U$, or explodes, i.e. $\limsup_{t \to T, t<T} \|\boldsymbol{X}_t\| = \infty$. If f satisfies the linear growth condition

$$\left\| f(\boldsymbol{X}) \right\|^2 \leq K\left(1 + \|\boldsymbol{X}\|^2\right)$$

with some constant $K \in \Re_+$, then no explosion can occur.

We point out that by unique solution we mean that pathwise uniqueness holds for (10.5.2). Two solutions on the same probability space, started from the same initial value and driven by the same semimartingale are then indistinguishable.

We now present a matrix version of the Girsanov theorem for matrix valued stochastic processes. To do so, we recall the notion of stochastic exponentials.

Definition 10.5.4 Let $\boldsymbol{X}$ be a stochastic process. The unique strong solution $\boldsymbol{Z} = \mathcal{E}(\boldsymbol{X})$ of the SDE

$$d\boldsymbol{Z}_t = \boldsymbol{Z}_t\, d\boldsymbol{X}_t, \quad \boldsymbol{Z}_0 = 1 \tag{10.5.3}$$

is called stochastic exponential of $\boldsymbol{X}$.

Theorem 10.5.3 allows us to conclude that the SDE (10.5.3) has a unique strong solution. We now formulate the Girsanov theorem, which will be employed in the next chapter.

Theorem 10.5.5 *Let $T > 0$, $\boldsymbol{W} \sim \mathcal{BM}_p$ and $\boldsymbol{U}$ be an adapted, continuous stochastic process with values in $\mathcal{M}_p(\Re)$ such that*

$$\left(\mathcal{E}\left(tr\left(-\int_0^t \boldsymbol{U}_s^\top d\boldsymbol{W}_s\right)\right)\right)_{t\in[0,T]} \tag{10.5.4}$$

is a martingale, or, which is a sufficient condition for (10.5.4), but not necessary, that the Novikov condition is satisfied

$$E\left(etr\left(\frac{1}{2}\int_0^T \boldsymbol{U}_t^\top \boldsymbol{U}_t\, dt\right)\right) < \infty.$$

Then

$$\hat{Q} = \int \mathcal{E}\left(tr\left(-\int_0^T \boldsymbol{U}_t^\top\, d\boldsymbol{W}_t\right)\right) dP$$

is an equivalent probability measure, and

$$\hat{\boldsymbol{W}}_t = \int_0^t \boldsymbol{U}_s\, ds + \boldsymbol{W}_t$$

is a $\hat{Q}$-Brownian motion on $[0, T)$.

We point out that the Novikov condition presents a general sufficient condition for

$$\left(\mathcal{E}\left(tr\left(-\int_0^t \boldsymbol{U}_s^\top \, d\boldsymbol{W}_s\right)\right)\right)_{t\in[0,T]}$$

to be a martingale. Clearly, for a given matrix valued process $\boldsymbol{U} = \{\boldsymbol{U}_t,\, t \in [0,T]\}$, it can be possible to improve on this sufficient condition, see e.g. Theorem 4.1 and Remark 4.2. in Mayerhofer (2012) for an example involving Wishart processes.

10.6 Matrix Valued Ornstein-Uhlenbeck Processes

As an example, we discuss the matrix valued OU-process, see Pfaffel (2008), where we direct the reader for additional results.

Definition 10.6.1 Let $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{M}_p(\Re)$, $\boldsymbol{x}_0 \in \mathcal{M}_{n,p}(\Re)$ a.s. and $\boldsymbol{W} \sim \mathcal{BM}_{n,p}$. A solution $\boldsymbol{X}$ of the SDE

$$d\boldsymbol{X}_t = \boldsymbol{X}_t \boldsymbol{B}\, dt + d\boldsymbol{W}_t\, \boldsymbol{A}, \quad \boldsymbol{X}_0 = \boldsymbol{x}_0, \tag{10.6.5}$$

is called an $n \times p$-dimensional Ornstein-Uhlenbeck process. We write $\boldsymbol{X} \sim \mathcal{OUP}_{n,p}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{x}_0)$ for its probability law.

Since the coefficients $\boldsymbol{X} \mapsto \boldsymbol{X}\boldsymbol{B}$ and $\boldsymbol{X} \mapsto \boldsymbol{A}$ are globally Lipschitz and satisfy the linear growth condition presented in Theorem 10.5.3, we are assured that the SDE (10.6.5) has a unique strong solution on the interval $[0, \infty)$. We can even go further and solve the SDE explicitly.

Theorem 10.6.2 *For a Brownian motion $\boldsymbol{W} \sim \mathcal{BM}_{n,p}$, the unique strong solution of (10.6.5) is given by*

$$\boldsymbol{X}_t = \boldsymbol{x}_0 \exp\{\boldsymbol{B}t\} + \left(\int_0^t d\boldsymbol{W}_s\, \boldsymbol{A} \exp\{-\boldsymbol{B}s\}\right) \exp\{\boldsymbol{B}t\}. \tag{10.6.6}$$

Proof The proof is completed by verifying that (10.6.6) solves (10.6.5). In this regard, the integration-by-parts formula, presented in Theorem 10.4.10, is crucial. We compute

$$\begin{aligned} d\boldsymbol{X}_t &= d\big(\boldsymbol{x}_0 \exp\{\boldsymbol{B}t\}\big) + d\left(\left(\int_0^t d\boldsymbol{W}_s\, \boldsymbol{A} \exp\{-\boldsymbol{B}s\}\right) \exp\{\boldsymbol{B}t\}\right) \\ &= \boldsymbol{x}_0 \exp\{\boldsymbol{B}t\}\boldsymbol{B}\, dt + d\boldsymbol{W}_t\, \boldsymbol{A} \exp\{-\boldsymbol{B}t\} \exp\{\boldsymbol{B}t\} \\ &\quad + \left(\int_0^t d\boldsymbol{W}_s \boldsymbol{A} \exp\{-\boldsymbol{B}s\}\right) \exp\{\boldsymbol{B}t\}\boldsymbol{B}\, dt \\ &\quad + d\left[\int_0^\cdot d\boldsymbol{W}_s\, \boldsymbol{A} \exp\{-\boldsymbol{B}s\}, \exp\{\boldsymbol{B}\cdot\}\right]_t^M \end{aligned}$$

$$= \left(\boldsymbol{x}_0 \exp\{\boldsymbol{B}t\} + \left(\int_0^t d\boldsymbol{W}_s \, \boldsymbol{A} \exp\{-\boldsymbol{B}s\} \right) \exp\{\boldsymbol{B}t\} \right) \boldsymbol{B} \, dt + d\boldsymbol{W}_t \, \boldsymbol{A}$$
$$= \boldsymbol{X}_t \boldsymbol{B} \, dt + d\boldsymbol{W}_t \, \boldsymbol{A}.$$

Finally, we note that (10.6.6) is by construction a strong solution. □

We state the following lemma, which is Lemma 3.48 from Pfaffel (2008).

Lemma 10.6.3 *Let* $\boldsymbol{W} \sim \mathcal{BM}_{n,p}$ *and* $\boldsymbol{X} : \Re_+ \to \mathcal{M}_{p,m}(\Re)$, $t \mapsto \boldsymbol{X}_t$ *be a square integrable, deterministic function. Then*

$$\int_0^t d\boldsymbol{W}_s \, \boldsymbol{X}_s \sim \mathcal{N}_{n,m}\left(\boldsymbol{0}, \boldsymbol{I}_n \otimes \int_0^t \boldsymbol{X}_s^\top \boldsymbol{X}_s \, ds \right).$$

We conclude this section with a result giving the distribution of the matrix valued Ornstein-Uhlenbeck process, see also Theorem 3.49 in Pfaffel (2008) for an alternative presentation.

Theorem 10.6.4 *Let* $\boldsymbol{X} \sim \mathcal{OUP}_{n,p}(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{x}_0)$, *then the distribution of* $\boldsymbol{X}$ *is given by*

$$\boldsymbol{X}_t | \boldsymbol{x}_0 \sim \mathcal{N}_{n,p}\left(\boldsymbol{\mu}, \boldsymbol{\sigma}^2\right),$$

where

$$\boldsymbol{\mu} = \boldsymbol{x}_0 \exp\{\boldsymbol{B}t\},$$
$$\boldsymbol{\sigma}^2 = \boldsymbol{I}_n \otimes \exp\left\{\boldsymbol{B}^\top t\right\} \int_0^t \exp\left\{-\boldsymbol{B}^\top s\right\} \boldsymbol{A}^\top \boldsymbol{A} \exp\{-\boldsymbol{B}s\} \, ds \exp\{\boldsymbol{B}t\}.$$

Proof The proof follows immediately from Lemma 10.6.3, Theorem 10.6.2, and Theorem 10.3.11. □

Finally, we remark that the stationary distribution of the matrix valued Ornstein-Uhlenbeck process can also be computed, see Pfaffel (2008), which is Gaussian.

10.7 A Two-Dimensional Correlated Minimal Market Model

In this section, we discuss how to extend the model for the GOP when denominated in two currencies, as discussed in Sect. 3.3, to allow for a more complex dependence structure. In particular, we introduce a model which allows us to use our knowledge of the Wishart distribution, see Definition 10.3.12. We denote the GOP denominated in the domestic currency by S^a, and the GOP denominated in the foreign currency by S^b. As discussed e.g. in Heath and Platen (2005), an exchange rate at time t can be expressed in terms of the ratio of the two GOPs, see also Sect. 9.7. Assuming the domestic currency is a, then one would pay, at time t, $\frac{S_t^a}{S_t^b}$ units of currency a to

obtain one unit of the foreign currency b. As the domestic currency is a, the price of e.g. a call option on the exchange rate can be expressed as:

$$S_0^a E\left(\frac{(\frac{S_T^a}{S_T^b} - K)^+}{S_T^a}\right). \tag{10.7.7}$$

We now discuss an extension of the model, which is tractable, as we can employ the non-central Wishart distribution to compute (10.7.7). For $k \in \{a, b\}$, we set

$$S_t^k = S_t^{0,k} \bar{S}_t^k,$$

where $S_t^{0,k} = \exp\{r_k t\}$, $S_0^{0,k} = 1$. So $S^{0,k}$ denotes the savings account in currency k, which for simplicity is assumed to be a deterministic function of time. As for the stylized MMM, we model $\bar{S}^k$ as a time-changed squared Bessel process of dimension four. We introduce the 2×4 matrix process $\boldsymbol{X} = \{\boldsymbol{X}_t,\, t \geq 0\}$ via

$$\boldsymbol{X}_t = \begin{bmatrix} (W^{1,1}_{\varphi^1(t)} + w^{1,1}) & (W^{2,1}_{\varphi^1(t)} + w^{2,1}) & (W^{3,1}_{\varphi^1(t)} + w^{3,1}) & (W^{4,1}_{\varphi^1(t)} + w^{4,1}) \\ (W^{1,2}_{\varphi^2(t)} + w^{1,2}) & (W^{2,2}_{\varphi^2(t)} + w^{2,2}) & (W^{3,2}_{\varphi^2(t)} + w^{3,2}) & (W^{4,2}_{\varphi^2(t)} + w^{4,2}) \end{bmatrix}.$$

The processes $W^{i,1}_{\varphi^1}$, $i = 1, \ldots, 4$, denote independent Brownian motions, subjected to a deterministic time-change

$$\varphi^1(t) = \frac{\alpha_0^1}{4\eta^1}\big(\exp\{\eta^1 t\} - 1\big) = \frac{1}{4}\int_0^t \alpha_s^1\, ds,$$

cf. Sect. 3.3, and $W^{i,2}_{\varphi^2}$, $i = 1, \ldots, 4$, denote independent Brownian motions, subjected to the deterministic time change

$$\varphi^2(t) = \frac{\alpha_0^2}{4\eta^2}\big(\exp\{\eta^2 t\} - 1\big) = \frac{1}{4}\int_0^t \alpha_s^2\, ds.$$

Now consider the process $\boldsymbol{Y} = \{\boldsymbol{Y}_t,\, t \geq 0\}$, which assumes values in S_2^+, and is given by

$$\boldsymbol{Y}_t := \boldsymbol{X}_t \boldsymbol{X}_t^\top, \quad t \geq 0,$$

which yields

$$\boldsymbol{Y}_t = \begin{bmatrix} \sum_{i=1}^4 (W^{i,1}_{\varphi^1(t)} + w^{i,1})^2 & \sum_{i=1}^4 \sum_{j=1}^2 (W^{i,j}_{\varphi^j(t)} + w^{i,j}) \\ \sum_{i=1}^4 \sum_{j=1}^2 (W^{i,j}_{\varphi^j(t)} + w^{i,j}) & \sum_{i=1}^4 (W^{i,2}_{\varphi^2(t)} + w^{i,2})^2 \end{bmatrix}.$$

We set

$$\bar{S}_t^a = Y_t^{1,1},$$

and

$$\bar{S}_t^b = Y_t^{2,2}.$$

We use the diagonal elements of $\boldsymbol{Y}_t$ to model the GOP in different currency denominations. Next, we introduce the following dependence structure: the Brownian motions $W^{i,1}$ and $W^{i,2}$, $i = 1, \ldots, 4$, covary as follows,

$$\left[W^{i,1}_{\varphi^1(\cdot)}, W^{i,2}_{\varphi^2(\cdot)}\right]_t = \frac{\varrho}{4} \int_0^t \sqrt{\alpha^1_s \alpha^2_0}\, ds, \quad i = 1, \ldots, 4, \tag{10.7.8}$$

where $-1 < \varrho < 1$. The specification (10.7.8) allows us to employ the non-central Wishart distribution. We work through this example in detail, as it illustrates how to extend the stylized MMM to allow for a non-trivial dependence structure, but still exploit the tractability of the Wishart distribution. As discussed in Sect. 10.3, matrix valued normal random variables are studied by interpreting the matrix as a vector, cf. Definition 10.3.12. We recall that $vec(\boldsymbol{X}_T^\top)$ stacks the two columns of $\boldsymbol{X}_T^\top$, hence

$$vec\left(\boldsymbol{X}_T^\top\right) = \begin{bmatrix} (W^{1,1}_{\varphi^1(T)} + w^{1,1}) \\ \vdots \\ (W^{4,1}_{\varphi^1(T)} + w^{4,1}) \\ (W^{1,2}_{\varphi^2(T)} + w^{1,2}) \\ \vdots \\ (W^{4,2}_{\varphi^2(T)} + w^{4,2}) \end{bmatrix}.$$

It is easily seen that the mean matrix $\boldsymbol{M}$ satisfies

$$vec\left(\boldsymbol{M}^\top\right) = \begin{bmatrix} w^{1,1} \\ \vdots \\ w^{4,1} \\ w^{1,2} \\ \vdots \\ w^{4,2} \end{bmatrix} \tag{10.7.9}$$

and the covariance matrix of $vec(\boldsymbol{X}_T^\top)$ is given by

$$\boldsymbol{\Sigma} \otimes \boldsymbol{I}_4 = \begin{bmatrix} \Sigma^{1,1}\boldsymbol{I}_4 & \Sigma^{1,2}\boldsymbol{I}_4 \\ \Sigma^{2,1}\boldsymbol{I}_4 & \Sigma^{2,2}\boldsymbol{I}_4 \end{bmatrix}, \tag{10.7.10}$$

where $\boldsymbol{\Sigma}$ is a 2×2 matrix with $\Sigma^{1,1} = \varphi^1(T)$, $\Sigma^{2,2} = \varphi^2(T)$, and $\Sigma^{1,2} = \Sigma^{2,1} = \frac{\varrho}{4} \int_0^t \sqrt{\alpha^1_s \alpha^2_s}\, ds$. We remark that assuming $-1 < \varrho < 1$ results in $\boldsymbol{\Sigma}$ being positive definite. It now immediately follows from Theorem 10.3.15 that

$$\boldsymbol{X}_T \boldsymbol{X}_T^\top \sim W_2\left(4, \boldsymbol{\Sigma}, \boldsymbol{\Sigma}^{-1}\boldsymbol{M}\boldsymbol{M}^\top\right),$$

where $\boldsymbol{M}$ and $\boldsymbol{\Sigma}$ are given in Eqs. (10.7.9) and (10.7.10) respectively.

Recall that we set

$$\begin{aligned}\boldsymbol{Y}_t &= \boldsymbol{X}_t \boldsymbol{X}_t^\top,\\ \bar{S}_t^a &= Y_t^{1,1},\\ \bar{S}_t^b &= Y_t^{2,2}.\end{aligned}$$

Hence we can compute (10.7.7) using

$$E\big(f(\boldsymbol{Y}_T)\big),$$

where $f : S_2^+ \to \Re$ is given by

$$f(\boldsymbol{M}) = \frac{\left(\frac{\exp\{r_1 T\} M^{1,1}}{\exp\{r_2 T\} M^{2,2}} - K\right)^+}{\exp\{r_1 T\} M^{1,1}},$$

for $\boldsymbol{M} \in S_2^+$, and $M^{i,i}$, $i = 1, 2$, are the diagonal elements of $\boldsymbol{M}$. The probability density function of $\boldsymbol{Y}_T$ is given in Definition 10.3.12.