11 Ito's Lemma

11.1 Summary

If a process is given as a stochastic Riemann and/or Ito integral, then one may wish to determine how a function of the process looks. This is achieved by Ito's lemma as an ingredient of stochastic calculus. In particular, stochastic integrals can be determined and stochastic differential equations can be solved with it; we will get to know stochastic variants of familiar rules of differentiation (chain and product rule). For this purpose we approach Ito's lemma step by step by first discussing it for Wiener processes, then by generalizing it for diffusion processes and finally by considering some extensions.

11.2 The Univariate Case

The WP itself is a special case of a diffusion as defined in (10.6). With

$$\mu\,(t, W(t)) = 0 \quad \text{and} \quad \sigma\,(t, W(t)) = 1$$

Eq. (10.6) becomes (with probability one)

$$W(t) = W(0) + \int_0^t dW(s) = \int_0^t dW(s)\,.$$

Thus, we consider this special case first.

U. Hassler, *Stochastic Processes and Calculus*, Springer Texts in Business and Economics, DOI 10.1007/978-3-319-23428-1_11

For Wiener Processes

As a revision, let us recall (10.3), which can be written equivalently as

$$2\int_0^t W(s)\,dW(s) = W^2(t) - t\,.$$

If $g(W) = W^2$ is defined with derivatives $g'(W) = 2W$ and $g''(W) = 2$, then this equation can also be formulated as follows:

$$\begin{aligned}\int_0^t g'\left(W(s)\right)\,dW(s) &= g\left(W(t)\right) - t\\ &= g\left(W(t)\right) - \frac{1}{2}\int_0^t g''\left(W(s)\right)\,ds\,.\end{aligned}$$

Now, this is just the form of Ito's lemma for functions g of a Wiener process. It is a corollary of the more general case (Proposition 11.1) which will be covered in the following. Throughout, we will assume that g has a continuous second derivative ("twice continuously differentiable").

Corollary 11.1 (Ito's Lemma for WP) *Let $g : \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable. Then it holds that*

$$dg\left(W(t)\right) = g'\left(W(t)\right)\,dW(t) + \frac{1}{2}\,g''\left(W(t)\right)\,dt\,.$$

In integral form this corollary to Ito's lemma is to be read as follows:

$$g\left(W(t)\right) = g\left(W(0)\right) + \int_0^t g'\left(W(s)\right)\,dW(s) + \frac{1}{2}\int_0^t g''\left(W(s)\right)\,ds\,.$$

Strictly speaking, this integral equation is the statement of the corollary, which is abbreviated by the differential notation. However, in doing so it must not be forgotten that the WP is not differentiable. Sometimes one also writes even more briefly:

$$dg\left(W\right) = g'(W)\,dW + \frac{1}{2}\,g''(W)\,dt\,.$$

Example 11.1 (Powers of the WP) For $g(W) = \frac{1}{2}W^2$ this special case of Ito's lemma just proves (10.3). In general, one obtains for $m \geq 2$ from Corollary 11.1 with $g(W) = \frac{W^m}{m}$:

$$d\left(\frac{W^m(t)}{m}\right) = W^{m-1}(t)\,dW(t) + \frac{m-1}{2}\,W^{m-2}(t)\,dt\,,$$

or in integral notation

$$W^m(t) = m \int_0^t W^{m-1}(s)\, dW(s) + \frac{m(m-1)}{2} \int_0^t W^{m-2}(s)\, ds\,. \quad \blacksquare$$

Explanation and Proof

Corollary 11.1 can be considered as a stochastic chain rule and can loosely be formulated as follows: the derivative of $g(W(t))$ results as the product of the outer derivative ($g'(W)$) and the inner derivative (dW), plus an Ito-specific extra term consisting of the second derivative of g times $\frac{1}{2}$. Where this term comes from (second order Taylor series expansion) and why no further terms occur (higher order derivatives), we want to clarify now. For this purpose we prove Corollary 11.1 (almost completely) although it is, as mentioned above, a corollary to Proposition 11.1.

With $s_n = t$ and $s_0 = 0$ it holds due to (8.3) that:

$$g(W(t)) = g(W(0)) + \sum_{i=1}^{n} \left(g(W(s_i)) - g(W(s_{i-1}))\right) .$$

Now, on the right-hand side a second order Taylor expansion of $g(W(s_i))$ about $W(s_{i-1})$ yields

$$\begin{aligned} g(W(s_i)) &= g(W(s_{i-1})) + g'(W(s_{i-1}))\ (W(s_i) - W(s_{i-1})) \\ &\quad + \frac{g''(\theta_i)}{2}\ (W(s_i) - W(s_{i-1}))^2\ , \end{aligned}$$

with θ_i between $W(s_{i-1})$ and $W(s_i)$:

$$|\ \theta_i - W(s_{i-1})| \in (0, |\ W(s_i) - W(s_{i-1})|)\ .$$

By substitution of $g(W(s_i)) - g(W(s_{i-1}))$, $g(W(t)) - g(W(0))$ can be expressed by two sums:

$$g(W(t)) - g(W(0)) = \Sigma_1 + \Sigma_2$$

with

$$\Sigma_1 = \sum_{i=1}^{n} g'(W(s_{i-1}))\, (W(s_i) - W(s_{i-1}))\,,$$

$$\Sigma_2 = \frac{1}{2} \sum_{i=1}^{n} g''(\theta_i)\, (W(s_i) - W(s_{i-1}))^2\,.$$

Now, Σ_1 just coincides with the Ito sum from (10.5) such that it holds due to Proposition 10.3 that:

$$\Sigma_1 \xrightarrow{2} \int_0^t g'(W(s))\, dW(s)\,.$$

Furthermore, we know from the section on quadratic variation (Proposition 10.8)

$$(dW(s))^2 = ds\,.$$

As the quadratic variation of the WP is not negligible (Proposition 10.7), this suggests the following approximation:

$$\begin{aligned}\Sigma_2 &\approx \frac{1}{2}\int_0^t g''(W(s))\,(dW(s))^2 \\ &= \frac{1}{2}\int_0^t g''(W(s))\, ds\,.\end{aligned}$$

A corresponding convergence in mean square can actually be established, which we will dispense with at this point. Hence, except for this technical detail, Corollary 11.1 is verified.

Additionally, we want to consider why higher order derivatives do not matter for Ito's lemma. For a third order Taylor expansion e.g. it follows

$$\begin{aligned}g(W(s_i)) - g(W(s_{i-1})) = {}& g'(W(s_{i-1}))(W(s_i) - W(s_{i-1})) \\ &+ \frac{g''(W(s_{i-1}))}{2}\,(W(s_i) - W(s_{i-1}))^2 \\ &+ \frac{g'''(\theta_i)}{6}\,(W(s_i) - W(s_{i-1}))^3\,.\end{aligned}$$

Thus, due to the summation, the term

$$\Sigma_3 = \sum_{i=1}^{n} g'''(\theta_i)\ (W(s_i) - W(s_{i-1}))^3$$

occurs. However, it is negligible:

$$\begin{aligned}|\Sigma_3| &\leq \sum_{i=1}^{n} \left|g'''(\theta_i)\right|\ |W(s_i) - W(s_{i-1})|\ (W(s_i) - W(s_{i-1}))^2 \\ &\leq \max_{1\leq i\leq n} \left\{\left|g'''(\theta_i)\right|\ |W(s_i) - W(s_{i-1})|\right\} \cdot Q_n(W,t) \\ &\xrightarrow{2} 0 \cdot t = 0\,,\end{aligned}$$

as the quadratic variation of the WP tends to t and as it furthermore holds that

$$\text{MSE}\left[W(s_i) - W(s_{i-1}),\ 0\right] = \text{Var}\left(W(s_i) - W(s_{i-1})\right) = s_i - s_{i-1} \to 0\,.$$

For Diffusions

Now, we turn to Ito's lemma for diffusions. In this section, we consider the univariate case of only one diffusion that depends on one WP only. The following variant of Ito's lemma is again a kind of stochastic chain rule and the idea for the proof is again based on a second order Taylor expansion.

Proposition 11.1 (Ito's Lemma with One Dependent Variable) *Let $g : \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable and $X(t)$ a diffusion on $[0, T]$ with (10.6), or briefly:*

$$dX(t) = \mu(t)\, dt + \sigma(t)\, dW(t)\,.$$

Then it holds that

$$dg\left(X(t)\right) = g'\left(X(t)\right)\ dX(t) + \frac{1}{2}\, g''\left(X(t)\right)\ \sigma^2(t)\, dt\,.$$

If $X(t) = W(t)$ is a Wiener process, i.e. $\mu(t) = 0$ and $\sigma(t) = 1$, then Corollary 11.1 is obtained as a special case.

The statement in Proposition 11.1 is given somewhat succinctly It can be condensed even more by suppressing the dependence on time:

$$dg\left(X\right) = g'\left(X\right)\ dX + \frac{1}{2}\, g''\left(X\right)\ \sigma^2\, dt\,.$$

However, it needs to be clear that by substituting $dX(t)$ one obtains for the differential $dg\left(X(t)\right)$ the following lengthy expression:

$$\left[g'\left(X(t)\right)\ \mu(t, X(t)) + \frac{1}{2}\, g''\left(X(t)\right)\ \sigma^2(t, X(t))\right] dt + g'\left(X(t)\right)\ \sigma(t, X(t))\, dW(t)\,.$$

The corresponding statement in integral notation naturally looks yet more extensive.

Example 11.2 (Differential of the Exponential Function) Let a diffusion $X(t)$ be given,

$$dX(t) = \mu(t)\, dt + \sigma(t)\, dW(t)\,.$$

Then, how does the differential of $e^{X(t)}$ read? This example is particularly easy to calculate as it holds for $g(x) = e^x$ that:

$$g''(x) = g'(x) = g(x) = e^x.$$

Hence Ito's lemma yields:

$$\begin{aligned} de^{X(t)} &= e^{X(t)} dX(t) + \frac{e^{X(t)}}{2}\sigma^2(t)dt \\ &= e^{X(t)}\left(\mu(t) + \frac{\sigma^2(t)}{2}\right) dt + e^{X(t)}\,\sigma(t)dW(t). \end{aligned}$$

If $X(t)$ is deterministic, i.e. $\sigma(t) = 0$, then it results

$$\frac{de^{X(t)}}{dt} = e^{X(t)}\frac{dX(t)}{dt},$$

which just corresponds to the traditional chain rule (outer derivative times inner derivative). ■

On the Proof

Just as for the proof of Corollary 11.1, one obtains with θ_i, where

$$|\theta_i - X(s_{i-1})| \in (0,\ |X(s_i) - X(s_{i-1})|)\ ,$$

from the Taylor expansion:

$$\begin{aligned} g(X(t)) - g(X(0)) &= \Sigma_1 + \Sigma_2\,, \\ \Sigma_1 &= \sum_{i=1}^{n} g'\left(X(s_{i-1})\right)\ \left(X(s_i) - X(s_{i-1})\right) \\ \Sigma_2 &= \frac{1}{2}\sum_{i=1}^{n} g''(\theta_i)\ \left(X(s_i) - X(s_{i-1})\right)^2\ . \end{aligned}$$

The first sum is approximated as desired:

$$\Sigma_1 \approx \int_0^t g'(X(s))\,dX(s)\,.$$

The second sum is approximated by

$$\Sigma_2 \approx \frac{1}{2} \int_0^t g''(X(s))\, (dX(s))^2 \,.$$

By multiplying out the square of the differential of the Ito process,

$$(dX(s))^2 = \mu^2(s)(ds)^2 + 2\,\mu(s)\,\sigma(s)\,dW(s)\,ds + \sigma^2(s)\,(dW(s))^2 \,,$$

one shows due to (cf. Proposition 10.8),

$$(ds)^2 = 0\,, \quad dW(s)\,ds = 0\,, \quad (dW(s))^2 = ds,$$

for the second sum:

$$\Sigma_2 \approx \frac{1}{2} \int_0^t g''(X(s))\,\sigma^2(s)\,ds \,.$$

This verifies Proposition 11.1 at least heuristically.

11.3 Bivariate Diffusions with One WP

A generalization of Proposition 11.1, which is sometimes needed, is presented by the following variant of Ito's lemma. The function g be dependent on two diffusions X_1 and X_2, where both are driven by the very same Wiener process. Occasionally, we will call this case (referring to the literature on interest rate models) the one-factor case as it is the identical factor $W(t)$ driving both diffusions.

One-Factor Case

Let g be a function in two arguments, whose partial derivatives are denoted by

$$\frac{\partial g(X_1, X_2)}{\partial X_i} \quad \text{and} \quad \frac{\partial^2 g(X_1, X_2)}{\partial X_i\, \partial X_j} \,.$$

Then, the following proposition is a special case of Proposition 11.3.

Proposition 11.2 (Ito's Lemma with Two Dependent Variables) *Let g: $\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ be twice continuously differentiable with respect to both arguments, and let $X_i(t)$ be diffusions on $[0, T]$ with the same WP:*

$$dX_i(t) = \mu_i(t)\,dt + \sigma_i(t)\,dW(t)\,, \quad i = 1, 2\,.$$

Then it holds that

$$\begin{aligned} dg\left(X_1(t),\, X_2(t)\right) &= \frac{\partial g\left(X_1(t),\, X_2(t)\right)}{\partial X_1}\, dX_1(t) + \frac{\partial g\left(X_1(t),\, X_2(t)\right)}{\partial X_2}\, dX_2(t) \\ &+ \frac{1}{2}\left[\frac{\partial^2 g\left(X_1(t),\, X_2(t)\right)}{\partial X_1^2}\, \sigma_1^2(t) + \frac{\partial^2 g\left(X_1(t),\, X_2(t)\right)}{\partial X_2^2}\, \sigma_2^2(t)\right] dt \\ &+ \frac{\partial^2 g\left(X_1(t),\, X_2(t)\right)}{\partial X_1\, \partial X_2}\, \sigma_1(t)\, \sigma_2(t)\, dt\,. \end{aligned}$$

Note that substitution of $dX_i(t)$ in Proposition 11.2 leads again to an integral equation for the process $g\left(X_1(t),\, X_2(t)\right)$ including Riemann and Ito integrals.

Frequently, the time dependence of the processes will be suppressed in order to obtain a more economical formulation of Proposition 11.2:

$$\begin{aligned} dg\left(X_1,\, X_2\right) &= \frac{\partial g\left(X_1,\, X_2\right)}{\partial X_1}\, dX_1 + \frac{\partial g\left(X_1,\, X_2\right)}{\partial X_2}\, dX_2 \\ &+ \frac{1}{2}\left[\frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_1^2}\, \sigma_1^2 + \frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_2^2}\, \sigma_2^2\right] dt \\ &+ \frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_1\, \partial X_2}\, \sigma_1\, \sigma_2\, dt\,. \end{aligned}$$

By this notation one recognizes, that again a second order Taylor expansion hides behind Proposition 11.2, but now of the two-dimensional function g,

$$\begin{aligned} dg\left(X_1,\, X_2\right) &= \frac{\partial g\left(X_1,\, X_2\right)}{\partial X_1}\, dX_1 + \frac{\partial g\left(X_1,\, X_2\right)}{\partial X_2}\, dX_2 \\ &+ \frac{1}{2}\left[\frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_1^2}\, (dX_1)^2 + \frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_2^2}\, (dX_2)^2\right] \\ &+ \frac{1}{2}\left[\frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_1\, \partial X_2}\, dX_1 dX_2 + \frac{\partial^2 g\left(X_1,\, X_2\right)}{\partial X_2\, \partial X_1}\, dX_2 dX_1\right], \end{aligned}$$

because the mixed second derivatives coincide due to the continuity assumed. With (10.7) and (10.8) it can easily be shown that (we again suppress the arguments)

$$(dX_i)^2 = \mu_i^2\, (dt)^2 + 2\, \mu_i\, \sigma_i\, dt dW + \sigma_i^2\, (dW)^2 = 0 + 0 + \sigma_i^2\, dt\,,$$

and for the covariance expression as well

$$dX_1\, dX_2 = \sigma_1\, \sigma_2\, dt\,. \tag{11.1}$$

Example 11.3 (One-Factor Product Rule) Proposition 11.2 provides us with a stochastic product rule for $X_1(t)\,X_2(t)$:

$$d(X_1(t)\,X_2(t)) = X_2(t)\,dX_1(t) + X_1(t)\,dX_2(t) + \sigma_1(t)\,\sigma_2(t)\,dt\,. \tag{11.2}$$

Under $\sigma_1(t) = 0$ or $\sigma_2(t) = 0$ (no stochastics), the well-known product rule is just reproduced. The derivation of (11.2) follows for $g(x_1, x_2) = x_1\,x_2$ with

$$\frac{\partial g}{\partial x_1} = x_2, \quad \frac{\partial^2 g}{\partial x_1^2} = 0,$$
$$\frac{\partial g}{\partial x_2} = x_1, \quad \frac{\partial^2 g}{\partial x_2^2} = 0$$

and

$$\frac{\partial^2 g}{\partial x_1\,\partial x_2} = \frac{\partial^2 g}{\partial x_2\,\partial x_1} = 1.$$

Hence, we obtain an abbreviated form:

$$d\,(X_1\,X_2) = \frac{\partial g\,(X_1,\,X_2)}{\partial X_1}\,dX_1 + \frac{\partial g\,(X_1,\,X_2)}{\partial X_2}\,dX_2 + \sigma_1\,\sigma_2\,dt\,,$$

where the second derivatives were plugged in. If one substitutes the first derivatives, then one obtains the result from (11.2). ■

Time as a Dependent Variable

Frequently it is of interest to consider another special case of Proposition 11.2. Again, g is a function in two arguments; however, the first one is time t, and the second one is a diffusion $X(t)$:

$$\begin{aligned} g: \quad [0,T] \times \mathbb{R} &\to \mathbb{R} \\ (t\,,X) &\mapsto g(t,X)\,. \end{aligned}$$

We consciously suppress the fact that the diffusion is time-dependent as well. Since, when we talk about the derivative of g with respect to time, then we refer strictly formally to the partial derivative with respect to the first argument. Sometimes this is confusing for beginners. For example for

$$g(t, X(t)) = g(t, X) = t\,X(t)$$

the derivative with respect to t refers to:

$$\frac{\partial g\,(t,\,X(t))}{\partial t} = X(t)\,.$$

Hence, for the partial derivatives we purposely do not consider that X itself is a function of t.

With $X_1(t) = t$ and $X(t) = X_2(t)$ from Proposition 11.2 we obtain for $\mu_1(t) = 1$ and $\sigma_1(t) = 0$ the following circumstance.

Corollary 11.2 (Ito's Lemma with Time as a Dependent Variable) *Let* $g : [0, T] \times \mathbb{R} \to \mathbb{R}$ *be twice continuously differentiable with respect to both arguments and let* $X(t)$ *be a diffusion on* $[0, T]$ *with (10.6), or briefly*

$$dX(t) = \mu(t)\,dt + \sigma(t)\,dW(t)\,.$$

Then it holds that

$$dg\,(t,\,X(t)) = \frac{\partial g\,(t,\,X(t))}{\partial t}\,dt + \frac{\partial g\,(t,\,X(t))}{\partial X}\,dX(t) + \frac{1}{2}\,\frac{\partial^2 g\,(t,\,X(t))}{\partial X^2}\,\sigma^2(t)\,dt\,.$$

Again, suppressing time-dependence, this can be condensed to

$$dg\,(t,\,X) = \frac{\partial g\,(t,\,X)}{\partial t}\,dt + \frac{\partial g\,(t,\,X)}{\partial X}\,dX + \frac{1}{2}\,\frac{\partial^2 g\,(t,\,X)}{\partial X^2}\,\sigma^2\,dt\,.$$

Example 11.4 (OUP as a Diffusion) As an application we can just prove that the standard Ornstein-Uhlenbeck process $X_c(t)$ from (9.3) is a diffusion with

$$\mu\,(t, X_c(t)) = c\,X_c(t) \quad \text{and} \quad \sigma\,(t, X_c(t)) = 1\,.$$

For this purpose we define as an auxiliary quantity the process

$$X(t) = \int_0^t e^{-cs} dW(s)\,,$$

or

$$dX(t) = e^{-ct} dW(t)\,.$$

With this variable we define the function g such,

$$g(t, X) = e^{ct} X\,,$$

that it holds for the OUP:

$$X_c(t) = g(t, X(t)) = e^{ct} X(t).$$

With the derivatives

$$\frac{\partial g(t,X)}{\partial t} = c e^{ct} X, \quad \frac{\partial g(t,X)}{\partial X} = e^{ct}, \quad \frac{\partial^2 g(t,X)}{\partial X^2} = 0$$

it follows from Corollary 11.2:

$$\begin{aligned} dX_c(t) &= c e^{ct} X(t) dt + e^{ct} dX(t) + 0 \\ &= c X_c(t) dt + dW(t), \end{aligned}$$

where $dX(t)$ was substituted. ■

Further examples for practicing Corollary 11.2 can be found in the problem section.

K-Variate Diffusions

Concerning the contents, there is no reason why Proposition 11.2 should be written just with two processes. Let us consider as a generalization the case where g depends on K diffusions, all of them given by the identical WP:

$$g\colon \mathbb{R}^K \to \mathbb{R}\,, \quad \text{i.e.} \quad g = g\,(X_1, \ldots, X_K) \in \mathbb{R}\,.$$

Then it holds with $dX_k(t)$, $k = 1, \ldots, K$, due to a second order Taylor expansion that:

$$dg\,(X_1, \ldots, X_K) = \sum_{k=1}^{K} \frac{\partial g}{\partial X_k}\, dX_k + \frac{1}{2} \sum_{k=1}^{K} \sum_{j=1}^{K} \frac{\partial^2 g}{\partial X_k\, \partial X_j}\, dX_k\, dX_j\,.$$

As in the bivariate case, one obtains $dX_k\, dX_j = \sigma_k\, \sigma_j\, dt$, cf. (11.1). Sometimes, as in Corollary 11.2, time as a further variable is allowed for,

$$g\colon [0,T] \times \mathbb{R}^K \to \mathbb{R}\,, \quad \text{i.e.} \quad g = g\,(t, X_1, \ldots, X_K) \in \mathbb{R}\,,$$

and

$$dg\,(t, X_1, \ldots, X_K) = \frac{\partial g}{\partial t}\, dt + \sum_{k=1}^{K} \frac{\partial g}{\partial X_k}\, dX_k + \frac{1}{2} \sum_{k=1}^{K} \sum_{j=1}^{K} \frac{\partial^2 g}{\partial X_k\, \partial X_j}\, dX_k\, dX_j\,.$$

11.4 Generalization for Independent WP

We keep to the multivariate generalization, however, allowing for several, stochastically independent Wiener processes behind the diffusions.

The General Case

Now, $W_1(t), \ldots, W_d(t)$ denote stochastically independent standard Wiener processes. We allow for d factors driving each of the K diffusions. According to this, let $\mathbf{X}(t)$ be a K-dimensional diffusion[1] $\mathbf{X}'(t) = (X_1(t), \ldots, X_K(t))$, defined by d factors $W_j(t), j = 1, \ldots, d$:

$$dX_k(t) = \mu_k(t)dt + \sum_{j=1}^{d} \sigma_{kj}(t)dW_j(t), \quad k = 1, \ldots, K.$$

In order to have a diffusion, it holds for μ_k and σ_{kj} that they may only depend on $\mathbf{X}(t)$ and t:

$$\mu_k(t) = \mu_k(t, \mathbf{X}(t)), \quad k = 1, \ldots K,$$

$$\sigma_{kj}(t) = \sigma_{kj}(t, \mathbf{X}(t)), \quad k = 1, \ldots K, \; j = 1, \ldots d.$$

For a function g, which maps $\mathbf{X}(t)$ to the real numbers, Ito's lemma reads as follows, cf. Øksendal (2003, Theorem 4.2.1).

Proposition 11.3 (Ito's Lemma (Independent WP)) *Let $g : \mathbb{R}^K \to \mathbb{R}$ be twice continuously differentiable with respect to all the arguments, and let $X_k(t)$ be diffusions on $[0, T]$ depending on d independent Wiener processes:*

$$dX_k(t) = \mu_k(t)dt + \sum_{j=1}^{d} \sigma_{kj}(t)dW_j(t), \quad k = 1, \ldots, K.$$

Then it holds for $\mathbf{X}'(t) = (X_1(t), \ldots, X_K(t))$ that

$$dg(\mathbf{X}(t)) = \sum_{k=1}^{K} \frac{\partial g(\mathbf{X}(t))}{\partial X_k} dX_k(t) + \frac{1}{2} \sum_{i=1}^{K} \sum_{k=1}^{K} \frac{\partial^2 g(\mathbf{X}(t))}{\partial X_i \partial X_k} dX_i(t) dX_k(t)$$

[1] For vectors and matrices, the superscript denotes transposition and not differentiation. Further, the dimension d of the multivariate Wiener process should not be confused with the differential operator denoted by the same symbol.

with

$$dX_i(t)dX_k(t) = \sum_{j=1}^{d} \sigma_{ij}(t)\sigma_{kj}(t)dt. \tag{11.3}$$

Heuristically, (11.3) can be well justified. For this purpose we consider vectors of length d:

$$\boldsymbol{\sigma}_k(t) = \begin{pmatrix} \sigma_{k1}(t) \\ \vdots \\ \sigma_{kd}(t) \end{pmatrix}, \ k = 1, \ldots, K, \quad \text{and } \mathbf{W}(t) = \begin{pmatrix} W_1(t) \\ \vdots \\ W_d(t) \end{pmatrix},$$

such that

$$dX_k(t) = \mu_k(t)dt + \boldsymbol{\sigma}'_k(t)d\mathbf{W}(t).$$

Neglecting the dependence on time, it follows

$$\begin{aligned} dX_i\, dX_k &= \mu_i\, \mu_k (dt)^2 + \mu_i\, \boldsymbol{\sigma}'_k d\mathbf{W}(t)dt + \mu_k\, \boldsymbol{\sigma}'_i d\mathbf{W}(t)dt \\ &\quad + \boldsymbol{\sigma}'_i d\mathbf{W}(t)\boldsymbol{\sigma}'_k d\mathbf{W}(t) \\ &= \boldsymbol{\sigma}'_i d\mathbf{W}(t) d\mathbf{W}'(t)\boldsymbol{\sigma}_k \end{aligned}$$

due to (see Proposition 10.8)

$$(dt)^2 = 0 \quad \text{and} \quad dW_j(t)dt = 0$$

and

$$\boldsymbol{\sigma}'_k d\mathbf{W}(t) = (\boldsymbol{\sigma}'_k d\mathbf{W}(t))' = d\mathbf{W}'(t)\boldsymbol{\sigma}_k.$$

Let us consider the matrix

$$d\mathbf{W}(t)\, d\mathbf{W}'(t) = \begin{pmatrix} (dW_1(t))^2 & dW_1(t)dW_2(t) & \ldots & dW_1(t)dW_d(t) \\ dW_2(t)dW_1(t) & (dW_2(t))^2 & \ldots & dW_2(t)dW_d(t) \\ \vdots & \vdots & \ddots & \vdots \\ dW_d(t)dW_1(t) & dW_d(t)dW_2(t) & \ldots & (dW_d(t))^2 \end{pmatrix}.$$

As is well known, it holds due to (10.7) that:

$$(dW_j(t))^2 = dt.$$

Furthermore, it can be shown for stochastically independent Wiener processes that

$$dW_i(t)dW_k(t) = 0, \quad i \neq k.$$

Overall, we hence obtain

$$d\mathbf{W}(t)d\mathbf{W}'(t) = I_d dt,$$

with the d-dimensional identity matrix I_d. All in all it follows

$$\begin{aligned} dX_i(t)dX_k(t) &= \boldsymbol{\sigma}_i'(t)\, I_d dt\, \boldsymbol{\sigma}_k(t) \\ &= \boldsymbol{\sigma}_i'(t)\boldsymbol{\sigma}_k(t)\, dt\,, \end{aligned}$$

which is given in (11.3).

The 2-Factor Case

Let us consider the case $K = d = 2$. Then, Proposition 11.3 becomes more clearly

$$dg(\mathbf{X}(t)) = \sum_{k=1}^{2} \frac{\partial g(\mathbf{X}(t))}{\partial X_k} dX_k(t) + \frac{1}{2}\sum_{i=1}^{2}\sum_{k=1}^{2} \frac{\partial^2 g(\mathbf{X}(t))}{\partial X_i \partial X_k} dX_i(t)dX_k(t)$$

with

$$dX_1 dX_1 = (\sigma_{11}^2 + \sigma_{12}^2)dt\,,$$

$$dX_2 dX_2 = (\sigma_{21}^2 + \sigma_{22}^2)dt\,,$$

and

$$dX_1 dX_2 = (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})dt.$$

Two interesting special cases result:

1. $\sigma_{12} = \sigma_{22} = 0$ (one-factor model),
2. $\sigma_{12} = \sigma_{21} = 0$ (independent diffusions).

The first case naturally corresponds to the one from the previous section: Both diffusions only depend on the same WP. The second case is the opposite extreme where both diffusions depend only on one or other of the two stochastically independent processes:

$$dX_k(t) = \mu_k(t)dt + \sigma_{kk}(t)dW_k(t), \quad k = 1, 2.$$

We want to discuss both borderline cases based on two examples.

Example 11.5 (2-Factor Product Rule) Proposition 11.3 with $K = d = 2$ yields with the derivatives from Example 11.3 as product rule:

$$d(X_1 X_2) = X_2\, dX_1 + X_1\, dX_2 + dX_1\, dX_2\,. \tag{11.4}$$

In the borderline case of only one factor, naturally the result from Eq. (11.2) is reproduced. In the second borderline case of stochastically independent diffusions, however, it holds, as in the deterministic case, that

$$d(X_1 X_2) = X_2\, dX_1 + X_1\, dX_2\,.$$

Without restrictions (11.4) reads as follows:

$$d(X_1 X_2) = X_2\, dX_1 + X_1\, dX_2 + (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})dt\,.$$

The example illustrates that a proper application of Ito's lemma needs to account for the number of factors underlying the diffusions. ■

Example 11.6 (2-Factor Quotient Rule) For $X_2(t) \neq 0$ and

$$g(X_1, X_2) = \frac{X_1}{X_2}$$

we obtain:

$$\frac{\partial g}{\partial X_1} = X_2^{-1}, \quad \frac{\partial g}{\partial X_2} = -X_1 X_2^{-2},$$

$$\frac{\partial^2 g}{\partial X_1^2} = 0, \quad \frac{\partial^2 g}{\partial X_2^2} = 2X_1 X_2^{-3}, \quad \frac{\partial^2 g}{\partial X_1 \partial X_2} = -X_2^{-2}.$$

Hence Proposition 11.3 yields with $K = d = 2$ suppressing the arguments:

$$d\left(\frac{X_1}{X_2}\right) = \frac{X_2\, dX_1 - X_1\, dX_2}{X_2^2} + \frac{X_1 X_2^{-1}(\sigma_{21}^2 + \sigma_{22}^2) - (\sigma_{11}\sigma_{21} + \sigma_{12}\sigma_{22})}{X_2^2}\, dt\,. \tag{11.5}$$

If X_2 is a deterministic function ($\sigma_{21} = \sigma_{22} = 0$), then the conventional quotient rule is reproduced. ■

11.5 Problems and Solutions

Problems

11.1 Prove part (a) from Example 9.1.
Hint: Choose $g(t, W) = t\,W$ in Corollary 11.2 or in (11.2).

11.2 Prove part (b) from Example 9.1.
Hint: Choose $g(t, W) = (1-t)\,W$ in Corollary 11.2 or in (11.2).

11.3 Prove part (b) from Proposition 9.1 (integration by parts).
Hint: Choose $g(t, W) = f(t)\,W$ in Corollary 11.2 or in (11.2).

11.4 Prove statement (a) from Proposition 9.4 with Ito's lemma.
Hint: Choose $g(t, W) = e^{-ct}W$.

11.5 Prove for the OUP from Proposition 9.4:

$$\int_0^t X_c(s)\,dW(s) = \frac{1}{2}\left(X_c^2(t) - t\right) - c\int_0^t X_c^2(s)\,ds$$

Note that for $c = 0$ (WP) this reproduces (10.3).
Hint: Choose $g(X_c(t)) = X_c^2(t)$ in Ito's lemma.

11.6 Determine the differential of $W(t)/e^{W(t)}$ according to the one-factor product rule (11.2).

11.7 Determine the differential of $W(t)/e^{W(t)}$ directly from Corollary 11.1.

Solutions

11.1 For the proof we use Corollary 11.2 with

$$g(t, W) = tW.$$

The derivatives needed read:

$$\frac{\partial g(t, W)}{\partial t} = W, \quad \frac{\partial g(t, W)}{\partial W} = t, \quad \frac{\partial^2 g(t, W)}{\partial W^2} = 0.$$

Hence, one determines with Corollary 11.2:

$$d(tW(t)) = W(t)dt + tdW(t) + \frac{0}{2}.$$

Due to $g(0, W(0)) = 0$, we obtain as an integral equation

$$tW(t) = \int_0^t W(s)ds + \int_0^t s dW(s),$$

which was to be shown.

11.2 As in Problem 11.1 we consider

$$g(t, W) = (1 - t)W$$

with

$$\frac{\partial g(t, W)}{\partial t} = -W, \quad \frac{\partial g(t, W)}{\partial W} = (1 - t), \quad \frac{\partial^2 g(t, W)}{\partial W^2} = 0.$$

Therefore, substitution into Corollary 11.2 yields

$$d((1 - t)W(t)) = -W(t)dt + (1 - t)dW(t) + \frac{0}{2}.$$

As $W(0) = 0$ with probability one, it follows that

$$(1 - t)W(t) = -\int_0^t W(s)ds + \int_0^t (1 - s)dW(s),$$

which was to be shown.

11.3 As an adequate function g we choose

$$g(t, W) = f(t)W,$$

where $f(t)$ is deterministic. Then, Corollary 11.2 is used with

$$\frac{\partial g(t, W)}{\partial t} = f'(t)W, \quad \frac{\partial g(t, W)}{\partial W} = f(t), \quad \frac{\partial^2 g(t, W)}{\partial W^2} = 0.$$

This yields for the differential:

$$dg(t, W(t)) = f'(t)W(t)dt + f(t)dW(t) + \frac{0}{2}.$$

In integral notation this reads

$$g(t, W(t)) = g(0, W(0)) + \int_0^t f'(s)W(s)ds + \int_0^t f(s)dW(s).$$

As $W(0) = 0$ with probability one, we hence obtain the desired result:

$$f(t)W(t) = \int_0^t f'(s)W(s)ds + \int_0^t f(s)dW(s).$$

11.4 If one chooses $g(t, W) = e^{-ct}W$ with

$$\frac{\partial g(t,W)}{\partial t} = -ce^{-ct}W, \quad \frac{\partial g(t,W)}{\partial W} = e^{-ct}, \quad \frac{\partial^2 g(t,W)}{\partial W^2} = 0,$$

then Corollary 11.2 allows for the following calculation:

$$d(e^{-ct}W(t)) = -ce^{-ct}W(t)dt + e^{-ct}dW(t),$$

i.e.

$$e^{-ct}W(t) = W(0) - c\int_0^t e^{-cs}W(s)ds + \int_0^t e^{-cs}dW(s)$$

or

$$\begin{aligned} W(t) &= -ce^{ct}\int_0^t e^{-cs}W(s)ds + e^{ct}\int_0^t e^{-cs}dW(s) \\ &= -ce^{ct}\int_0^t e^{-cs}W(s)ds + X_c(t). \end{aligned}$$

Rearranging terms completes the proof.

11.5 With the function $g(X_c) = X_c^2$ and its derivatives,

$$g'(X_c) = 2X_c, \quad g''(X_c) = 2,$$

Proposition 11.1 can be applied. We know that $X_c(t)$ is a diffusion with (see Example 11.4):

$$dX_c(t) = cX_c(t)dt + dW(t).$$

Plugging in into Proposition 11.1 shows:

$$\begin{aligned} dX_c^2(t) &= 2X_c(t)dX_c(t) + \frac{2}{2}dt \\ &= (2cX_c^2(t) + 1)dt + 2X_c(t)dW(t). \end{aligned}$$

With starting value $X_c(0) = 0$ this translates into the following integral equation:

$$X_c^2(t) = \int_0^t (2cX_c^2(s) + 1)ds + 2\int_0^t X_c(s)dW(s).$$

This is equivalent to

$$\int_0^t X_c(s)dW(s) = \frac{1}{2}\left(X_c^2(t) - \int_0^t ds\right) - c\int_0^t X_c^2(s)ds,$$

which amounts to the claim.

11.6 We define

$$X_1(t) = W(t), \quad X_2(t) = e^{-W(t)},$$

and we are interested in the differential of the product. In order to apply the one-factor product rule, we need the differentials of the factors. For X_1 it obviously holds that: $dX_1 = dW$. For $e^{-W(t)}$ Example 11.2 yields

$$dX_2 = de^{-W} = -e^{-W}dW + \frac{e^{-W}}{2}dt\,.$$

Hence, we have

$$\sigma_1(t) = 1, \quad \sigma_2(t) = -e^{-W(t)}.$$

Plugging in into the product rule (11.2) yields:

$$\begin{aligned} d\left(We^{-W}\right) &= e^{-W}dX_1 + WdX_2 - e^{-W}dt \\ &= e^{-W}dW + W\left(-e^{-W}dW + \frac{e^{-W}}{2}dt\right) - e^{-W}dt \\ &= e^{-W}\left(\frac{W}{2} - 1\right)dt + e^{-W}(1 - W)dW. \end{aligned}$$

11.7 As $g(W) = \frac{W}{e^W}$ is a simple function of W, Corollary 11.1 yields a direct approach to the differential. For this purpose, we only need the derivatives (quotient rule):

$$g'(W) = \frac{e^W - W\,e^W}{e^{2W}} = \frac{1 - W}{e^W},$$

$$g''(W) = \frac{-e^W - (1 - W)\,e^W}{e^{2W}} = \frac{W - 2}{e^W}.$$

Thus, it follows from Ito's lemma that

$$\begin{aligned} d\left(\frac{W}{e^W}\right) &= g'(W)\,dW + \frac{1}{2}\,g''(W)\,dt \\ &= \frac{1-W}{e^W}\,dW + \frac{W-2}{2e^W}\,dt \\ &= e^{-W}\left(\frac{W}{2}-1\right)dt + e^{-W}(1-W)\,dW. \end{aligned}$$

Of course, this result coincides with the one from the previous problem.

Reference

Øksendal, B. (2003). *Stochastic differential equations: An introduction with applications* (6th ed.). Berlin/New York: Springer.