Derivation of Operator-Form Multivariate Itô Lemma with Infinite Brownian Dimensions

Recall 1-D Itô's Lemma:

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu \frac{\partial f}{\partial x} + \frac{1}{2}\sigma^2 \frac{\partial^2 f}{\partial x^2}\right)dt + \sigma \frac{\partial f}{\partial x} dW_t$$

Recall Itô Table:

$$\boxed{dt \cdot dW_t = 0} \qquad \boxed{(dt)^2 = 0} \qquad \boxed{(dW_t)^2 = dt}$$

This way because of Quadratic Variation

$$[W]_t := \lim_{\|\Pi\| \to 0} \sum_{i=1}^{n} (W_t - W_{t-1})^2 = t \implies [dW_t]^2 = dt$$

Multivariate Itô Lemma has a new rule:

$$\boxed{dW_t^{(i)} \cdot dW_t^{(j)} = \rho_{ij}\, dt}$$

if $i = j$,

$$dW_t^{(i)} \cdot dW_t^{(j)} = dt$$

if $i \neq j$,

$$dW_t^{(i)} \cdot dW_t^{(j)} = \rho_{ij}\, dt$$

Thus, ρ_{ij} is a Correlation Coefficient (normalized version of Covariance)

$$\rho_{ij} = \frac{Cov(A,B)}{\sqrt{Var(A)} \cdot \sqrt{Var(B)}} \quad \& \quad \rho_{ij} \in [-1, +1]$$

Alternatively: $\mathbb{E}[dW_t^{(i)} \cdot dW_t^{(j)}] = \rho_{ij}\, dt \rightarrow$ Quadratic Variation

$$\Rightarrow dW_t^{(i)} \cdot dW_t^{(j)} = \rho_{ij}\, dt$$

$$[dW^{(i)}, dW^{(j)}]_t := \lim_{\|\Pi\| \to 0} \sum_{k=1}^{n} dW_{t_k}^{(i)} \cdot dW_{t_k}^{(j)} = \rho_{ij}\, t$$

At the infinitesimal level ρ_{ij} is always multiplied by dt, which shrinks as the # of intervals grows.

$$\sum \underline{\rho_{ij}} \cdot \underline{dt} = \underline{\rho_{ij}} \cdot \underline{t}$$

Changes

Doesn't change because dt properly scales!

$\Rightarrow$ Simple 2D Itô Lemma example:

Assume:
$$dX_t = \mu_x(t)\,dt + \sigma_x(t)\,dW_t^{(1)}$$
$$dY_t = \mu_y(t)\,dt + \sigma_y(t)\,dW_t^{(2)}$$

$\rightarrow$ Apply 2^{nd} order multivariable Taylor expansion

Let $f = f(X_t, Y_t, t)$, then:

$$df = \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dX_t + \frac{\partial f}{\partial y}dY_t + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2} \cdot (dY_t)^2 + \frac{\partial^2 f}{\partial x \partial y}dX_t\, dY_t$$

Alternatively: $df = f_t\,dt + f_x\,dX_t + f_y\,dY_t + \frac{1}{2} f_{xx}(dX_t)^2 + \frac{1}{2} f_{yy}(dY_t)^2 + f_{xy}\,dX_t\,dY_t$

Why go up to the 2nd order?

→ In addition to the 1D Itô Table rules, we have:

$$\boxed{(dX_t)^2 \sim dt} \quad \& \quad \boxed{dX_t \cdot dY_t \sim dt}$$

$$(dX_t)^2 = (\mu_x\,dt + \sigma_x\,dW_t)^2 = \cancel{\mu_x^2 (dt)^2} + \cancel{2\mu_x\,dt\,\sigma_x\,dW_t} + \sigma_x^2 \underbrace{(dW_t)^2}_{\to dt}$$

$$\boxed{(dX_t)^2 = \sigma_x^2\,dt \sim dt}$$

// $dX_t \cdot dY_t = (\mu_x\,dt + \sigma_x\,dW_t^{(1)})(\mu_y\,dt + \sigma_y\,dW_t^{(2)})$

$$= \cancel{\mu_x\mu_y(dt)^2} + \cancel{\mu_x\,dt\,\sigma_y\,dW_t^{(2)}} + \cancel{\sigma_x\,dW_t^{(1)}\mu_y\,dt} + \sigma_x\sigma_y\,dW_t^{(1)} \cdot dW_t^{(2)}$$

$$dX_t \cdot dY_t = \sigma_x\sigma_y \underbrace{dW_t^{(1)}\,dW_t^{(2)}} \to dW_t^{(i)}\,dW_t^{(j)} = \rho_{ij}\,dt$$

$$\Rightarrow \boxed{dX_t \cdot dY_t = \sigma_x\sigma_y\rho_{12}\,dt \sim dt}$$

Now plug in dX_t & dY_t into 2nd order multivariable Taylor expansion df

$$
\begin{aligned}
df = {} & f_t\,dt + f_x \mu_x\,dt + f_x \sigma_x\,dW_t^{(1)} + f_y \mu_y\,dt + f_y \sigma_y\,dW_t^{(2)} \\
& + \cancel{\tfrac{1}{2} f_{xx} \mu_x^2 (dt)^2} + \cancel{f_{xx} \mu_x\,dt\,\sigma_x\,dW_t^{(1)}} + \tfrac{1}{2} f_{xx} \sigma_x^2 \underbrace{(dW_t^{(1)})^2}_{\to dt} \\
& + \cancel{\tfrac{1}{2} f_{yy} \mu_y^2 (dt)^2} + \cancel{f_{yy} \mu_y\,dt\,\sigma_y\,dW_t^{(2)}} + \tfrac{1}{2} f_{yy} \sigma_y^2 \underbrace{(dW_t^{(2)})^2}_{\to dt} \\
& + \cancel{f_{xy} \mu_x \mu_y (dt)^2} + \cancel{f_{xy} \mu_x \sigma_y\,dt\,dW_t^{(2)}} + \cancel{f_{xy} \sigma_x \mu_y\,dW_t^{(1)}\,dt} \\
& + f_{xy} \sigma_x \sigma_y \underbrace{(dW_t^{(1)})(dW_t^{(2)})}_{\to \rho_{12}\,dt}
\end{aligned}
$$

$$
\boxed{
\begin{aligned}
df = {} & f_t\,dt + f_x \mu_x\,dt + f_x \sigma_x\,dW_t^{(1)} + f_y \mu_y\,dt + f_y \sigma_y\,dW_t^{(2)} \\
& + \tfrac{1}{2} f_{xx} \sigma_x^2\,dt + \tfrac{1}{2} f_{yy} \sigma_y^2\,dt + f_{xy} \sigma_x \sigma_y \rho_{12}\,dt
\end{aligned}
}
$$

Fully Expanded 2D Itô Lemma w/ SDEs Applied

$$
\boxed{df = \frac{\partial f}{\partial t}\,dt + \sum_i \frac{\partial f}{\partial x_i}\,dX_t^{(i)} + \frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}\,dX_t^{(i)}\,dX_t^{(j)}}
$$

Core General Version of Multivariate Itô Lemma

Generalized Form of the Itô Process

$$
\underbrace{d\vec{X}_t}_{\substack{\text{Vector Valued} \\ \text{Stochastic} \\ \text{Process}}} = \underbrace{\vec{\mu}(t)}_{\substack{\text{Drift} \\ \text{Vector}}}\,dt + \underbrace{\underline{\Sigma}(t)}_{\substack{\text{Volatility} \\ \text{Matrix}}}\,\underbrace{d\vec{W}_t}_{\substack{m\text{-dimensional} \\ \text{Brownian motion} \\ \text{Vector}}}
$$

$$\vec{X}_t = \begin{bmatrix} X_t^{(1)} \\ X_t^{(2)} \\ \vdots \\ X_t^{(n)} \end{bmatrix} \text{ with } \vec{X}_t \in \mathbb{R}^{n \times 1} \qquad \vec{\mu}(t) = \begin{bmatrix} \mu_1(t) \\ \mu_2(t) \\ \vdots \\ \mu_n(t) \end{bmatrix} \text{ with } \vec{\mu}(t) \in \mathbb{R}^{n \times 1}$$

$$\underline{\underline{\sigma}}(t) = \begin{bmatrix} \sigma_{11}(t) & \sigma_{12}(t) & \cdots & \sigma_{1m}(t) \\ \sigma_{21}(t) & \sigma_{22}(t) & \cdots & \sigma_{2m}(t) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1}(t) & \sigma_{n2}(t) & \cdots & \sigma_{nm}(t) \end{bmatrix} \text{ with } \underline{\underline{\sigma}}(t) \in \mathbb{R}^{n \times m}$$

$$\vec{W}_t = \begin{bmatrix} W_t^{(1)} \\ W_t^{(2)} \\ \vdots \\ W_t^{(m)} \end{bmatrix} \text{ with } d\vec{W}_t \in \mathbb{R}^{m \times 1}$$

→ You can have an ∞ # of Brownian terms in a single process — the more Brownian terms, the more accurate.

→ However the law of diminishing returns prevails.

→ The meaningfulness of each additional Brownian term decays exponentially & converges in L^2 sense.

Infinite-Dimensional Itô Process

$$d\vec{X}_t = \vec{\mu}(t)dt + \underbrace{\sum_{k=1}^{\infty} \vec{\varphi}_k(t) \cdot dW_t^{(k)}}$$

The sum of infinite Brownian terms · infinite (& properly scaled) Brownian weights $\varphi_k(t)$

→ The weights $\varphi_k(t)$ decay geometrically

Suppose $\varphi_k(t) = \lambda^k$, if $\lambda = 0.5$, each term is $\frac{1}{2}$ as big as the one before it

$$\vec{\varphi}(t) = \begin{bmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_N(t) \end{bmatrix} = \begin{bmatrix} \lambda^1 \\ \lambda^2 \\ \vdots \\ \lambda^N \end{bmatrix} \text{ with } \lambda \in (0,1)$$

The Brownian term "contribution" is measured via Variance.

$$X_t = X_0 + \int_0^t \mu(s)\,ds + \sum_{k=1}^{\infty} \int_0^t \varphi_k(s)\,dW_s^{(k)}$$

$$Var(X_t) = \underbrace{Var(X_0)}_{\text{Deterministic } \times} + \underbrace{Var\left(\int_0^t \mu(s)\,ds\right)}_{\text{Deterministic } \times} + \underbrace{Var\left(\sum_{k=1}^{\infty} \int_0^t \varphi(s)\,dW_s^{(k)}\right)}_{\text{Random } \checkmark}$$

$$\boxed{Var(X_t) = Var\left(\sum_{k=1}^{\infty} \int_0^t \varphi_k(s)\,dW_s^{(k)}\right)}$$

Variance of an Infinite-Dimensional Brownian Process (the "contribution")

Substitute $dX_t^{(i)} = \mu_i\,dt + \sum_{k=1}^{\infty} \varphi_k^{(i)}(t)\,dW_t^{(k)}$ in df

$$df = \frac{\partial f}{\partial t}dt + \sum_i \frac{\partial f}{\partial x_i}\left(\mu_i\,dt + \sum_{k=1}^{\infty} \varphi_k^{(i)}(t)\,dW_t^{(k)}\right) +$$

$$\frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}\left(\mu_i\,dt + \sum_{k=1}^{\infty} \varphi_k^{(i)}(t)\,dW_t^{(k)}\right)\left(\mu_j\,dt + \sum_{k=1}^{\infty} \varphi_k^{(j)}(t)\,dW_t^{(k)}\right)$$

$$df = \frac{\partial f}{\partial t}dt + \sum_i \frac{\partial f}{\partial x_i}\mu_i\, dt + \sum_i \sum_{k=1}^{\infty} \frac{\partial f}{\partial x_i}\varphi_k^{(i)}(t)\, dW_t^{(k)} +$$

$$\frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}\left(\cancel{\mu_i \mu_j (dt)^2} + \cancel{\mu_i\, dt \sum_{k=1}^{\infty}\varphi_k^{(j)}(t)\, dW_t^{(k)}} + \cancel{\sum_{k=1}^{\infty}\varphi_k^{(i)}(t)\, dW_t^{(k)}\mu_j\, dt} + \sum_{k=1}^{\infty}\varphi_k^{(i)}(t)\varphi_k^{(j)}(t)\underbrace{\left(dW_t^{(k)}\right)^2}_{dt}\right)$$

$$df = \frac{\partial f}{\partial t}dt + \sum_i \frac{\partial f}{\partial x_i}\mu_i\, dt + \sum_i \sum_{k=1}^{\infty}\frac{\partial f}{\partial x_i}\varphi_k^{(i)}(t)\, dW_t^{(k)} +$$

$$\frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}\sum_{k=1}^{\infty}\varphi_k^{(i)}(t)\varphi_k^{(j)}(t)\, dt$$

$$df = \left(\frac{\partial f}{\partial t} + \sum_i \frac{\partial f}{\partial x_i}\mu_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 f}{\partial x_i \partial x_j}\underline{\sum_{k=1}^{\infty}\varphi_k^{(i)}(t)\varphi_k^{(j)}(t)}\right)dt$$

$$+ \sum_i \sum_{k=1}^{\infty}\frac{\partial f}{\partial x_i}\varphi_k^{(i)}(t)\, dW_t^{(k)}$$

Let $\underline{\underline{\Sigma}}(t) = \sum_{k=1}^{\infty}\vec{\varphi}_k(t)\vec{\varphi}_k(t)^T$ → Instantaneous Covariance Matrix of the noise on the system (covariance kernel)

If using $\{\varphi_k(t)\}_{k=1}^{\infty}$ as weighted volatility vectors instead of $\underline{\underline{\sigma}}(t)$ volatility matrix, we need an infinite-dimensional setting

$$d\vec{X}_t = \vec{\mu}(t)\,dt + \sum_{k=1}^{\infty}\vec{\varphi}_k(t)\, dW_t^{(k)} \quad \longrightarrow \quad \underline{\underline{\Sigma}}(t) = \left[\underline{\underline{\Phi}}(t)\,\underline{\underline{\Phi}}(t)^*\right]$$

Instead of manual summation, we have an infinite matrix multiplied by its adjoint

$\Updownarrow$ Hilbert-Schmidt operator

$$d\vec{X}_t = \vec{\mu}(t)\,dt + \underline{\underline{\Phi}}(t)\, d\vec{W}_t$$

$$\underline{\underline{\Phi}}(t) \sim \begin{bmatrix} | & | & & | \\ \vec{\varphi}_1(t) & \vec{\varphi}_2(t) & \cdots & \vec{\varphi}_k(t) \\ | & | & & | \end{bmatrix} \quad \text{with } \underline{\underline{\Phi}}(t) \in \mathcal{L}_2(\ell^2; \mathbb{R}^n)$$

Thus

$$\frac{1}{2}\underbrace{\sum_{i,j}}\underbrace{\left(\sum_k \varphi_k^{(i)}(t)\,\varphi_k^{(j)}(t)\right)}\underbrace{\frac{\partial^2 f}{\partial x_i \partial x_j}}$$

The trace of their products

$$\left[\underline{\underline{\Phi}}(t)\,\underline{\underline{\Phi}}(t)^*\right] = \underline{\underline{\Sigma}}(t)$$

$$\underline{\underline{\nabla}}^2 f$$

Hessian Matrix

contains all entries just like the double sum.

$$\Rightarrow \boxed{\frac{1}{2}\mathrm{Tr}\left(\underline{\underline{\Sigma}}(t)\,\underline{\underline{\nabla}}^2 f\right)} =$$

2nd order term
Curvature (uncertainty) of function

First order drift term $\displaystyle\sum_i \frac{\partial f}{\partial x_i}\mu_i = \boxed{\vec{\nabla} f^T \vec{\mu}(t)}$

Stochastic Term $\displaystyle\sum_i \sum_{k=1}^{\infty} \frac{\partial f}{\partial x_i}\varphi_k^{(i)}(t)\,dW_t^{(k)} = \boxed{\vec{\nabla} f^T \underline{\underline{\Phi}}(t)\,d\vec{W}_t}$

$\therefore$

$$\boxed{df = \left(\frac{\partial f}{\partial t} + \vec{\nabla} f^T \vec{\mu}(t) + \frac{1}{2}\mathrm{Tr}\left(\underline{\underline{\Sigma}}(t)\,\underline{\underline{\nabla}}^2 f\right)\right)dt + \vec{\nabla} f^T \underline{\underline{\Phi}}(t)\,d\vec{W}_t}$$

Operator-Form Multivariate Itô Lemma w/ infinite Brownian Dimensions 4/22/2025 京都 Joshua Kyan Aalampour