Lecture 18 — State-Space Workflows, Brownian Motion, and Diffusion Processes

Chapter 7 to Chapter 9: Kalman filtering in practice, stochastic differentials, and the bridge to continuous time

Jiajing Sun

School of Economics and Management, University of Chinese Academy of Sciences

Econometrics and Time Series Methods
Spring 2026

Why Lecture 18 matters

Lecture 17 gave us the state-space framework and the Kalman recursion. Lecture 18 does three new things:

- it turns that theory into an **R workflow** for estimation, filtering, smoothing, and diagnostics;
- it introduces **Brownian motion** and **stochastic differentials**, which are the natural language of continuous-time econometrics;
- it connects discrete-time partial sums to diffusion limits through the **functional central limit theorem**.

Big transition

We are moving from latent-state models in discrete time to the continuous-time objects that appear in diffusion models, high-frequency finance, and asymptotic theory.

Where Lecture 18 fits in the course

- Lecture 16 introduced deterministic filtering.
- Lecture 17 studied probabilistic filtering with latent states.
- **Lecture 18** first shows how to implement those ideas, then begins Chapter 9 by introducing the continuous-time probability tools behind diffusion models.
- Lecture 19 will continue with likelihood inference for diffusion models and high-frequency volatility.

Teaching logic

Hour 1 is still closely tied to Lecture 16. Hours 2 and 3 explain why Brownian motion and diffusion processes are the right continuous-time limits of many econometric objects.

Learning goals

By the end of the lecture, you should be able to:

1. implement a local level model in R and interpret filtered and smoothed states;
2. explain how the prediction-error decomposition produces the likelihood;
3. write down the transformed stochastic-volatility state-space model;
4. define Brownian motion and state its key moment and path properties;
5. explain why Brownian motion is continuous but nowhere differentiable;
6. interpret stochastic differential notation such as $dX_t = \mu\, dt + \sigma\, dB_t$;
7. define diffusion processes and recognize canonical examples such as GBM, OU, and CIR;
8. explain the discrete-to-continuous bridge given by Donsker's theorem.

Three-hour plan

Hour 1

R block: state-space estimation, Kalman filtering, smoothing, and a stochastic-volatility illustration.

Hour 2

Brownian motion, stochastic differentials, quadratic variation, and Brownian bridges.

Hour 3

Diffusion processes, first-passage ideas, Euler approximations, and the functional central limit theorem.

From lecture theory to an R workflow

In practice, the state-space workflow has five steps:

1. specify the observation and state equations;
2. choose unknown hyperparameters such as H and Q;
3. evaluate the prediction-error likelihood recursively;
4. optimize the likelihood numerically;
5. run the filter and smoother using the fitted parameters.

$$y_t = Z_t\alpha_t + \eta_t, \qquad \alpha_t = T_t\alpha_{t-1} + R_t\varepsilon_t.$$

Practical message

The Kalman filter is not only a recursive estimator. It is also the engine that turns a latent-state model into a likelihood-based estimation problem.

Local level model in software

The local level model is the cleanest scalar example:

$$y_t = \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim N(0, H),$$
$$\mu_t = \mu_{t-1} + \eta_t, \qquad \eta_t \sim N(0, Q).$$

Unknown quantities

- latent state path: $\{\mu_t\}_{t=1}^T$;
- variance parameters: H and Q;
- initial state mean and variance: $a_0 = \mathbb{E}(\mu_0)$, $P_0 = \mathrm{Var}(\mu_0)$.

Interpretation

H controls measurement noise. Q controls how quickly the latent level moves over time.

Prediction-error decomposition of the likelihood

For each t, define the one-step-ahead prediction error

$$v_t = y_t - a_t, \qquad a_t = \mathbb{E}(\mu_t \mid \mathcal{F}_{t-1}), \qquad F_t = \mathrm{Var}(v_t \mid \mathcal{F}_{t-1}).$$

In the linear Gaussian case,

$$\ell(\theta) = -\frac{1}{2} \sum_{t=1}^{T} \left[\log(2\pi) + \log F_t + \frac{v_t^2}{F_t} \right].$$

- The filter provides v_t and F_t recursively.
- Therefore likelihood evaluation costs only one forward pass.
- Numerical estimation becomes feasible even when the state is unobserved.

Kalman recursion objects in the scalar local level model

$$a_t = a_{t-1|t-1}, \qquad P_t = P_{t-1|t-1} + Q.$$

$$v_t = y_t - a_t, \qquad F_t = P_t + H, \qquad K_t = \frac{P_t}{F_t}.$$

$$a_{t|t} = a_t + K_t v_t, \qquad P_{t|t} = P_t - K_t^2 F_t = (1 - K_t) P_t.$$

Signal-to-noise interpretation

If H is large relative to P_t, then K_t is small and the filter trusts the state prediction more than the new observation.

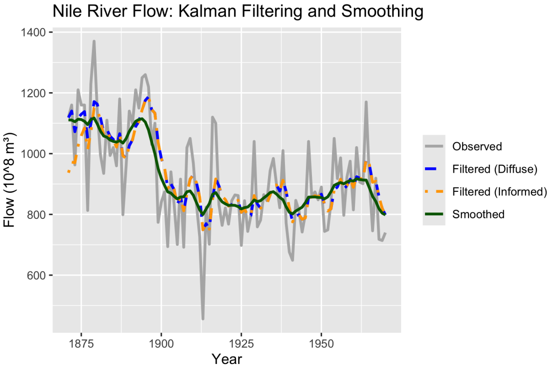
R workflow for a local level model

```
library(dlm)

build_model <- function(theta) {
  dlmModPoly(order = 1,
             dV = exp(theta[1]), # H
             dW = exp(theta[2])) # Q
}

init_theta <- log(c(var(y, na.rm = TRUE),
                   var(y, na.rm = TRUE) / 10))
fit <- dlmMLE(y, parm = init_theta, build = build_model)
mod <- build_model(fit$par)

filt <- dlmFilter(y, mod)
smth <- dlmSmooth(filt)

level_filt <- dropFirst(filt$m)
level_smth <- dropFirst(smth$s)
```

Workflow

Parameterize variances in logs, optimize by MLE, then run `dlmFilter()` and `dlmSmooth()`.

Nile example: filtered and smoothed level



How to read the Nile output

- The **filtered level** uses information only up to date t.
- The **smoothed level** uses the full sample $\{y_1, \ldots, y_T\}$.
- The two are close in the middle of the sample but can differ near turning points or near the sample edges.

Why smoothing helps

Filtering is real-time inference. Smoothing is retrospective inference. Once future data arrive, they help re-estimate past states.

Likelihood connection

The same model that produces these filtered and smoothed paths is the model whose parameters were estimated by the prediction-error likelihood.

Initialization and missing data

The filter is especially useful when data are missing or irregular:

- if y_t is missing, skip the measurement update and propagate the state;
- the predictive variance still grows through the state equation;
- once an observation reappears, the filter absorbs the accumulated uncertainty.

$$\text{If } y_t \text{ missing:} \qquad a_{t|t} = a_t, \qquad P_{t|t} = P_t.$$

Initialization issue

Diffuse priors use a very large P_0. Informed priors place tighter prior mass on the initial state. Early-sample filtering can depend materially on this choice.

A state-space view of stochastic volatility

Stochastic-volatility models introduce a second layer of randomness:

- the return innovation z_t is random;
- the volatility scale itself evolves stochastically.

$$x_t = \sigma_t z_t, \qquad \log(\sigma_t) = \log(\sigma_{t-1}) + \omega_t,$$

where $z_t \sim N(0, 1)$ and $\omega_t \sim N(0, Q)$.

Why this matters

Unlike GARCH, volatility is not a deterministic function of past squared shocks. It is an unobserved latent process.

Basic stochastic-volatility model

Let $s_t = \log(\sigma_t)$. Then the textbook model is

$$\begin{aligned} x_t &= e^{s_t} z_t, \\ s_t &= s_{t-1} + \omega_t, \end{aligned} \qquad z_t \sim N(0,1), \quad \omega_t \sim N(0,Q).$$

- s_t follows a random walk, so volatility can drift over time.
- Large shocks in x_t may come from large z_t, large σ_t, or both.
- The state s_t is hidden, so the inferential problem is naturally state-space.

Log-squared transformation

To obtain a measurement equation, square and log the returns:

$$y_t = \log(x_t^2) = 2s_t + \varepsilon_t, \qquad s_t = s_{t-1} + \omega_t,$$

where

$$\varepsilon_t = \log(z_t^2).$$

Payoff

The model is now linear in the latent state s_t.

But not yet fully Gaussian

$\varepsilon_t \sim \log(\chi_1^2)$, so the observation error is not normal.

Why transformed SV is still awkward

The Kalman filter is exact for **linear Gaussian** systems. Here the transformed observation equation is linear, but the noise is non-Gaussian.

$$\varepsilon_t \sim \log(\chi_1^2), \qquad y_t = 2s_t + \varepsilon_t.$$

- We have gained linearity.
- We have not gained Gaussianity.
- Therefore a plain Kalman filter is either approximate or conditionally exact after an additional approximation step.

Core question

How much non-Gaussianity are we willing to ignore, and what approximation is accurate enough for inference?

Route 1: quasi-maximum likelihood

The simplest approach is to pretend the transformed observation noise is Gaussian:

$$y_t = 2s_t + u_t, \qquad u_t \overset{\text{QMLE}}{\sim} N(0, \sigma_u^2).$$

- Then the standard Kalman filter applies directly.
- The resulting estimator is a **quasi-MLE**.
- It is often useful for exploratory work and for obtaining starting values.

Limitation

QMLE is computationally convenient, but it does not fully respect the heavy skewness and kurtosis of $\log(\chi_1^2)$.

Route 2: Gaussian mixture approximation

A more accurate strategy approximates

$$\varepsilon_t \sim \log(\chi_1^2)$$

by a finite mixture of normals:

$$\varepsilon_t \approx \sum_{i=1}^{K} q_i N(m_i, \sigma_i^2).$$

Conditional Gaussianity

Introduce a latent mixture indicator $r_t \in \{1, \ldots, K\}$. Conditional on $r_t = i$,

$$y_t = 2s_t + m_i + \nu_t, \qquad \nu_t \sim N(0, \sigma_i^2).$$

Conditional on the indicators, the state-space system becomes Gaussian again.

Seven-component approximation used in practice

i	q_i	m_i	σ_i^2
1	0.00730	-10.12999	5.79596
2	0.10566	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

Interpretation

These constants are chosen so that the mixture approximates the density of $\log(\chi_1^2)$ closely enough for efficient filtering and smoothing within a simulation-based estimator.

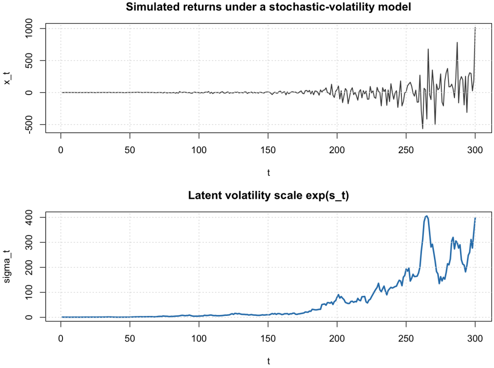
R workflow for stochastic volatility

```
# Step 1: transform returns
y <- log(x^2)

# Step 2: specify latent-state law
s_t = s_{t-1} + omega_t

# Step 3: pick an approximation
# (a) QMLE: treat epsilon_t as Gaussian
# (b) Mixture: epsilon_t | r_t=i ~ N(m_i, sigma_i^2)

# Step 4: conditional on parameters (and on r_t for mixture methods),
# apply Kalman filter and smoother

# Step 5: maximize the approximate likelihood or combine
# filtering with Gibbs / simulation-based updating
```

Practical point

The Kalman filter remains useful even when the original model is not exactly Gaussian, because careful approximations can restore conditional Gaussian structure.

A stochastic-volatility illustration



Hour 1: what you should keep

- State-space models are estimated by recursively evaluating innovation densities.
- Filtering gives real-time state estimates; smoothing uses the full sample.
- Missing observations are easy to handle inside the recursion.
- Stochastic volatility becomes linear after the log-squared transformation, but non-Gaussianity remains.
- QMLE is simple; mixture methods are more faithful to the true observation law.

Bridge to the rest of the lecture

The latent state in Hour 1 lives in discrete time. The latent uncertainty in Hours 2 and 3 will be modeled by continuous-time objects such as Brownian motion and diffusions.

Why continuous-time probability enters econometrics

Brownian motion appears in at least three ways in econometrics:

1. as the building block of continuous-time diffusion models;
2. as the weak limit of normalized partial sums in asymptotic theory;
3. as the noise process behind high-frequency volatility models.

Key insight

Continuous time is not just for finance. It is also the language of limit theory for many discrete-time econometric statistics.

$$T^{-1/2} \sum_{t=1}^{\lfloor Ts \rfloor} X_t \;\Rightarrow\; B(s).$$

A short history of Brownian motion

- **Robert Brown** observed erratic pollen motion in water in 1827.
- **Albert Einstein** gave a physical explanation in 1905 and connected it to molecular motion.
- **Norbert Wiener** later provided the rigorous mathematical construction, hence the term **Wiener process**.

Why we care

In our course, Brownian motion is not introduced as a piece of physics. It is introduced because it is the canonical continuous-time noise process and the central limit object for rescaled partial sums.

Start from a discrete random walk

Let $\{\varepsilon_t\}$ be i.i.d. $N(0,1)$, and define

$$z_t - z_0 = \sum_{j=1}^{t} \varepsilon_{t-j}.$$

Then

$$\mathbb{E}(z_t - z_0) = 0, \qquad \mathrm{Var}(z_t - z_0) = t.$$

Discrete-time intuition

The location after t steps is the sum of many independent shocks, so uncertainty accumulates linearly in time and the standard deviation grows like $\sqrt{t}$.

Definition of standard Brownian motion

A process $\{B_t : t \geq 0\}$ is a standard Brownian motion if:

1. $B_0 = 0$ almost surely;
2. it has continuous sample paths;
3. for $0 \leq s < t$, the increment $B_t - B_s \sim N(0, t - s)$;
4. increments over disjoint intervals are independent.

Continuous-time random walk

Brownian motion is the continuous-time analog of a random walk with Gaussian increments whose variance matches elapsed time.

Moments and covariance structure

For $0 \leq s \leq t$,

$$\mathbb{E}(B_t) = 0, \qquad \mathrm{Var}(B_t) = t, \qquad \mathrm{Cov}(B_s, B_t) = s = \min\{s, t\}.$$

Why the covariance looks like $\min\{s, t\}$

Write $B_t = B_s + (B_t - B_s)$. The increment $B_t - B_s$ is independent of B_s, so

$$\mathrm{Cov}(B_s, B_t) = \mathrm{Cov}(B_s, B_s) = \mathrm{Var}(B_s) = s.$$

Martingale and Markov properties

Brownian motion has two structural properties that will matter later:

- **Martingale property:**
$$\mathbb{E}(B_t \mid \mathcal{F}_s) = B_s, \qquad s \leq t.$$
- **Markov property:** the future conditional law depends on the past only through the current state B_s.

Interpretation

The martingale property says the best forecast of future Brownian motion is its current value. The Markov property says the current state summarizes all relevant predictive information.

Independent increments and Gaussian structure

For disjoint intervals,

$$B_{t_2} - B_{t_1}, \quad B_{t_4} - B_{t_3}$$

are independent whenever $[t_1, t_2] \cap [t_3, t_4] = \varnothing$.

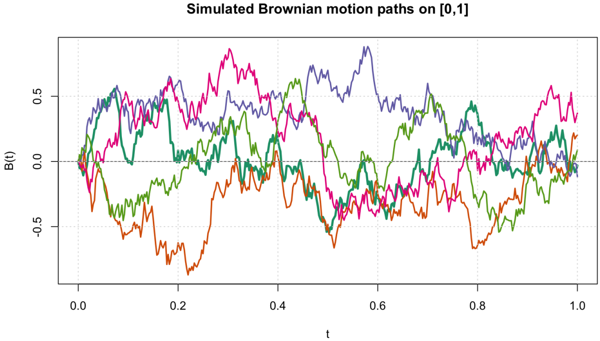
Consequences

- finite-dimensional distributions are jointly Gaussian;
- all inference reduces to means and covariances;
- Brownian motion is the natural noise process in linear diffusion models.

But do not over-interpret independence

The process itself is not a collection of independent points. Only *increments over disjoint intervals* are independent.

What Brownian paths look like



Continuous, but nowhere differentiable

Brownian motion is almost surely continuous, but it is almost surely nowhere differentiable.

Heuristic

Over a small interval of length Δt,

$$B_{t+\Delta t} - B_t = O_p(\sqrt{\Delta t}).$$

If we tried to form a derivative, we would divide by Δt, giving

$$\frac{B_{t+\Delta t} - B_t}{\Delta t} = O_p\left(\frac{1}{\sqrt{\Delta t}}\right),$$

which explodes as $\Delta t \to 0$.

Meaning

Brownian paths are rough at every scale. They are continuous curves, but never smooth curves.

Hölder continuity and modulus of continuity

Brownian motion is almost surely Hölder continuous of any order $\gamma < 1/2$:

$$|B_t - B_s| \le C|t-s|^\gamma \quad \text{for every } \gamma < \tfrac{1}{2}.$$

But it is not Hölder continuous of order $\gamma > 1/2$, and certainly not differentiable.

$$g(\delta) = \left(2\delta \log \frac{1}{\delta}\right)^{1/2}$$

is the classical modulus-of-continuity scale.

Practical reading

Brownian motion is not wildly discontinuous. Its roughness is very structured: the path regularity stops exactly at the 1/2-barrier.

Stochastic differential notation

When Δt is very small, Brownian increments satisfy

$$B_{t+\Delta t} - B_t \sim N(0, \Delta t).$$

In differential notation we write

$$dB_t.$$

Moment rules

$$\mathbb{E}_t(dB_t) = 0, \qquad \mathrm{Var}_t(dB_t) = dt.$$

Important warning

dB_t is not an ordinary derivative. It is shorthand for a random increment whose scale is $\sqrt{dt}$, not dt.

Generalized Brownian motion

The basic continuous-time signal-plus-noise model is

$$dX_t = \mu \, dt + \sigma \, dB_t.$$

Then

$$\mathbb{E}_t[dX_t] = \mu \, dt, \qquad \mathrm{Var}_t[dX_t] = \sigma^2 dt.$$

Interpretation

$\mu \, dt$ is the deterministic drift over a short interval, while $\sigma \, dB_t$ is the random fluctuation. The drift is first-order in dt; the noise is first-order in $\sqrt{dt}$.

Integral representation

Differential notation becomes more concrete after integration:

$$X_t = X_0 + \int_0^t \mu \, ds + \int_0^t \sigma \, dB_s.$$

- The first integral is an ordinary Riemann integral.
- The second is a stochastic Itô integral.
- If μ and σ are constant, then

$$X_t = X_0 + \mu t + \sigma B_t.$$

Why this matters

Most continuous-time models are really statements about how a process accumulates drift and stochastic shocks over time.

Quadratic variation is the key new object

For a smooth deterministic function, squared increments are negligible. For Brownian motion they are not.

$$\sum_{j=1}^{n} \left(B_{t_j} - B_{t_{j-1}}\right)^2 \;\to\; t$$

in probability as the mesh of the partition goes to zero.

Quadratic variation

Brownian motion has quadratic variation $[B]_t = t$.

Why this matters

Quadratic variation is the reason Itô calculus is different from ordinary calculus. It is also the foundation of realized-volatility estimation in high-frequency finance.

Why $(dB_t)^2 = dt$ is only heuristic notation

In Itô calculus, the useful bookkeeping rules are

$$(dB_t)^2 = dt, \qquad dB_t\,dt = 0, \qquad (dt)^2 = 0.$$

Why these rules make sense

- $dB_t = O_p(\sqrt{dt})$, so $(dB_t)^2 = O_p(dt)$;
- $dB_t\,dt = O_p(dt^{3/2})$, which is negligible;
- $(dt)^2$ is even smaller.

Interpret carefully

This is asymptotic algebra based on orders of magnitude, not literal pointwise multiplication of derivatives.

Brownian bridge as a related continuous-time object

Define the Brownian bridge on $[0, 1]$ by

$$\mathbb{B}(s) = B(s) - sB(1), \qquad 0 \le s \le 1.$$

- $\mathbb{B}(0) = 0$ and $\mathbb{B}(1) = 0$;
- it is Brownian motion with the endpoint component removed;
- it arises naturally after centering, demeaning, residualization, or conditioning on an endpoint.

$$\mathrm{Cov}(\mathbb{B}(s), \mathbb{B}(u)) = \min\{s, u\} - su.$$

Why it belongs here

When we later move from partial-sum processes to weak limits, Brownian motion and Brownian bridge both appear repeatedly.

First-passage questions arise naturally

Many economic questions are threshold questions:

- When does a process first exceed a barrier?
- How likely is it to hit a boundary before time t?
- How long do we wait before an extreme event occurs?

Examples

Barrier options, default thresholds, volatility triggers, and exchange trading halts are all first-passage problems.

Key definition

For Brownian motion, a first-passage time is a stopping time defined by the first boundary crossing of the path.

Two-sided first-passage time

For standard Brownian motion, define

$$\tau_a = \inf\{t : |B_t| > a\}.$$

The textbook gives the distribution

$$\mathbb{P}(\tau_a \le t \mid x, a) = 1 - \frac{2}{\pi} \sum_{j=0}^{\infty} \frac{(-1)^j}{j + \frac{1}{2}} \cos\left(\left(j + \frac{1}{2}\right) \frac{\pi x}{a}\right) \exp\left[-\left(j + \frac{1}{2}\right)^2 \frac{\pi^2 t}{2a^2}\right].$$

Message

Even for Brownian motion, boundary-crossing laws are highly nontrivial. Continuous-time probability becomes interesting very quickly.

One-sided crossing probability

Define the one-sided stopping time

$$\tau_a^+ = \inf\{t : B_t > a\}, \qquad a > 0.$$

Starting from $B_0 = 0$, the reflection-principle result is

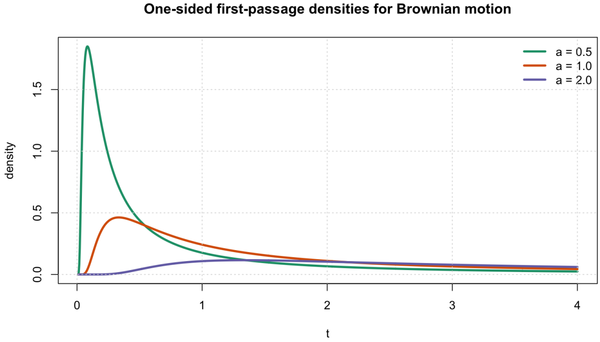
$$\mathbb{P}(\tau_a^+ \le t) = 2\left(1 - \Phi\left(\frac{a}{\sqrt{t}}\right)\right).$$

Its density is

$$f_0(t \mid 0, a) = \frac{a}{\sqrt{2\pi t^3}} \exp\left(-\frac{a^2}{2t}\right).$$

$$\mathbb{E}\left(\frac{1}{\tau_a^+}\right) = \frac{1}{a^2}, \qquad \mathbb{E}(\tau_a^+) = \infty.$$

How the first-passage density changes with the barrier



Stopping times and the strong Markov idea

A random time τ is a stopping time if

$$\{\tau \leq s\} \in \mathcal{F}_s \qquad \text{for all } s.$$

Strong Markov idea

If τ is a stopping time, then after time τ, Brownian motion starts afresh from B_τ in the sense that future increments are independent of the past, conditional on the present state.

- This is why hitting-time problems are analytically tractable.
- It is also why boundary crossing can often be analyzed by restarting the process at the hitting time.

Definition of a diffusion process

A diffusion process is a continuous-time Markov process with continuous sample paths, typically written as the solution of

$$dX_t = \mu(X_t, t)\, dt + \sigma(X_t, t)\, dB_t.$$

Components

- $\mu(X_t, t)$: drift term;
- $\sigma(X_t, t)$: diffusion coefficient;
- B_t: Brownian driver.

Interpretation

Diffusions are continuous-time analogs of nonlinear autoregressions with state-dependent conditional mean and conditional variance.

Integral form of a diffusion SDE

The stochastic differential equation

$$dX_t = \mu(X_t, t)\,dt + \sigma(X_t, t)\,dB_t$$

means

$$X_t = X_0 + \int_0^t \mu(X_s, s)\,ds + \int_0^t \sigma(X_s, s)\,dB_s.$$

- the drift accumulates through an ordinary integral;
- the noise accumulates through an Itô integral;
- future uncertainty depends on the state through $\sigma(X_t, t)$.

Why this is useful

The integral form tells us exactly what must be estimated: a deterministic accumulation term and a stochastic accumulation term.

Existence and uniqueness conditions

The textbook states standard sufficient conditions for a unique strong solution:

1. $\mathbb{E}(X_0^2) < \infty$;
2. $\mu(x, t)$ and $\sigma(x, t)$ are Lipschitz in x;
3. they satisfy a linear-growth bound such as

$$|\mu(x, t)| + |\sigma(x, t)| \leq K(1 + |x|).$$

Meaning

These conditions rule out coefficients that change too wildly as the state moves. They ensure the SDE is well posed.

Geometric Brownian motion

The Black–Scholes diffusion is

$$dX_t = \beta X_t \, dt + \sigma X_t \, dB_t.$$

Properties

- volatility is proportional to the current level;
- positivity is preserved if $X_0 > 0$;
- after logging, the model becomes linear in drift plus Brownian noise.

$$d \log X_t = \left(\beta - \frac{\sigma^2}{2}\right) dt + \sigma \, dB_t.$$

This is the canonical multiplicative-growth diffusion.

Ornstein–Uhlenbeck and Vasicek diffusion

The mean-reverting Gaussian diffusion is

$$dX_t = \kappa(\alpha - X_t)\,dt + \sigma\,dB_t.$$

- if X_t is below α, the drift is positive;
- if X_t is above α, the drift is negative;
- the process is pulled back toward its long-run level α.

Why econometricians like it

It is analytically tractable, Gaussian, and useful for modeling short rates, spreads, and other approximately mean-reverting series.

CIR or square-root diffusion

The Cox–Ingersoll–Ross diffusion is

$$dX_t = \kappa(\alpha - X_t)\,dt + \sigma\sqrt{X_t}\,dB_t.$$

Why the square root matters

- volatility becomes small near zero;
- the diffusion can stay nonnegative under suitable parameter restrictions;
- it is therefore attractive for short rates and variance processes.

$$\text{Feller condition:} \qquad 2\kappa\alpha \geq \sigma^2$$

is the classical sufficient condition preventing the process from hitting zero.

Euler discretization as a bridge back to discrete time

Over a short interval Δ,

$$X_{t+\Delta} - X_t \approx \mu(X_t, t)\Delta + \sigma(X_t, t)\sqrt{\Delta}\,\varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim N(0, 1).$$

Why this matters

This formula makes a diffusion look like a state-dependent discrete-time model.

- It is the basis of simulation.
- It is the basis of approximate likelihood.
- It is the first step in connecting continuous-time models back to econometric data observed at discrete times.

Functional central limit theorem

Let $\{X_t\}$ be i.i.d. with $\mathbb{E}(X_t) = 0$ and $\mathrm{Var}(X_t) = 1$. Define the partial-sum process

$$S_T(s) = \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Ts \rfloor} X_t, \qquad 0 \le s \le 1.$$

Then

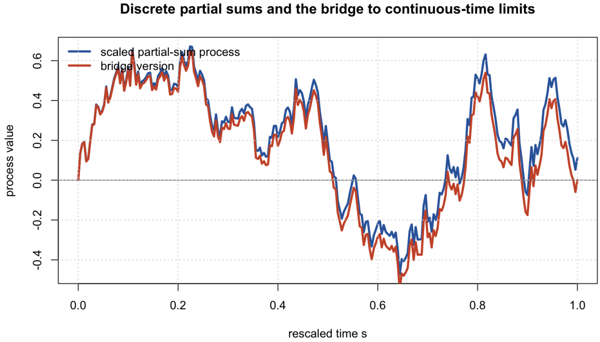
$$S_T(\cdot) \Rightarrow B(\cdot)$$

in $D[0,1]$ as $T \to \infty$.

Interpretation

The usual CLT gives a limit for one number. The functional CLT gives a limit for the entire cumulative-sum path.

A picture of the discrete-to-continuous bridge



Why the bridge matters in econometrics

The FCLT explains why Brownian motion keeps reappearing:

- unit-root asymptotics use partial-sum limits;
- self-normalized and fixed-b statistics use Brownian functionals;
- diffusion models are continuous-time limits of discrete stochastic systems;
- realized-volatility methods rely on quadratic variation, the pathwise counterpart of these limits.

Slogan

Brownian motion is both a model for real continuous-time uncertainty and a limit approximation for many discrete-time econometric objects.

Lecture 18 summary

1. In practice, state-space models are estimated by the prediction-error likelihood.
2. Local level models, missing-data handling, and smoothing are straightforward in R.
3. Stochastic volatility is naturally state-space after the log-squared transformation, but requires a treatment of non-Gaussianity.
4. Brownian motion is the canonical continuous-time noise process.
5. Diffusion models combine drift and Brownian noise in an SDE.
6. First-passage problems and Donsker-type limits show why continuous-time probability matters for econometrics.

Looking ahead to Lecture 19

Lecture 19 continues the continuous-time theme:

- likelihood and approximate likelihood for diffusion models;
- R simulation of Brownian paths and realized volatility;
- high-frequency volatility and covariance estimation under microstructure noise.

What to remember before next time

Keep the three bridges in mind:

1. state-space theory to R implementation;
2. random walks to Brownian motion;
3. discrete-time partial sums to continuous-time diffusion limits.