# Lecture 2
# Stochastic Processes

Econ 2713A: Computation

University of Pittsburgh, 2018

# Stochastic Optimization

Most stochastic elements in discrete time models have straightforward analogues in continuous time

There are number of useful processes that are useful in a continuous time context because they preserve locality

The cost of introducing these is often the inclusion of higher order derivatives in value function equations

# Wiener Process

Bedrock process is the Wiener process / Brownian motion, characterized by $W_0 = 0$ and

1. Independent, normal increments

   $$W_s - W_t \mid W_{[0,t]} \sim \mathcal{N}(0, s-t)$$

2. Continuous with probability one

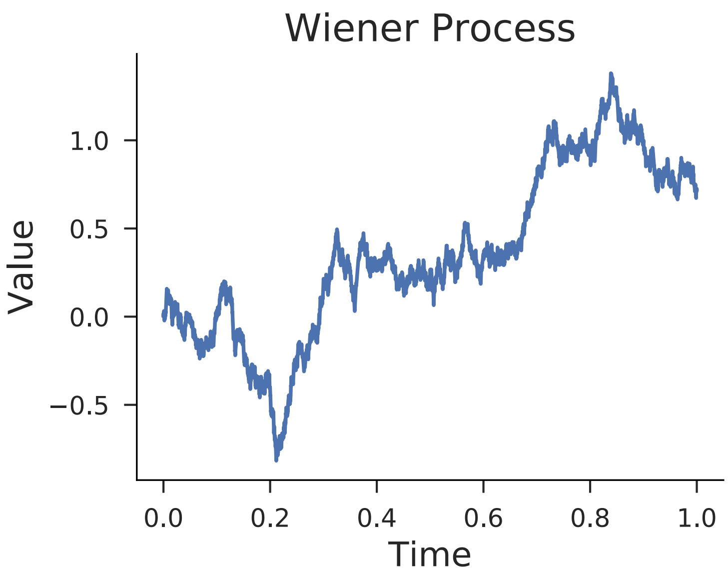
   $$\lim_{h \to 0} \mathbb{P}(|W_{t+h} - W_h| > \varepsilon) = 0$$

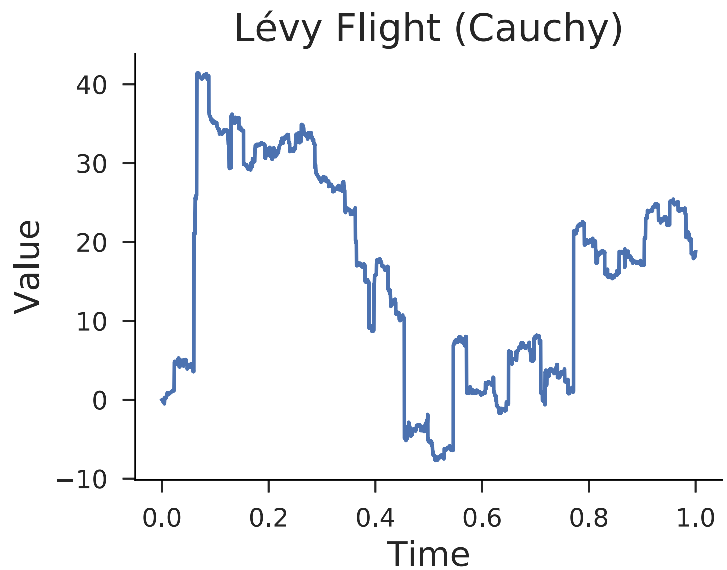
Surprisingly few other options with just (1) as increments must be mean zero and Levy stable

- Cauchy would be another example, but this has undefined mean and variance

# Continuous Processes

Wiener process is the continuous limit of a random walk with increments $\mathcal{N}(0, \Delta)$ as $\Delta \to 0$



# Central Limit Theorem

It turns out that *any* standard random walk with finite mean and variance increments converges to a Wiener process

$$\frac{1}{\sqrt{N}}\sum_{i=0}^{N} z_i \xrightarrow[N]{} \text{Wiener process}$$

Note that Cauchy distribution has infinite variance so Levy Flight does not obey Central Limit Theorem

Random walks are also *recurrent*: they return to the same neighborhood infinitely often in the future. This breaks down for $d \geq 3$ though.

# Characteristic Function

The characteristic function of a distribution is a spectral decomposition closely related to the Fourier transform

$$\varphi_X(t) = \mathbb{E}[\exp(itX)] = \int_{-\infty}^{\infty} \exp(itx) f(x) dx$$

This has the distinctive property that it generates the moments of the distribution

$$\mathbb{E}[X^n] = i^{-n} \left[ \frac{d^n}{dt^n} \varphi_X(t) \right]_{t=0}$$

# Spectral Convolution

We can also express the convultion of two variables in characteristic space. Suppose that $X$ and $Y$ are independent and $Z = X + Y$ is distributed as

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) dx$$

In terms of characteristic functions, convolution simply becomes multiplication as in

$$\varphi_Z(t) = \varphi_X(t) \cdot \varphi_Y(t)$$

# Gaussian Basin

Taking a Taylor expansion, we can see that in general

$$\varphi_X(t) = \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \mathbb{E}[X^n]$$

Suppose that $\mathbb{E}[X] = 0$ and $\mathbb{E}[X^2] = 1$. Consider the distribution of $Z = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i$

$$\varphi_Z(t) = \left[\varphi_X\left(\frac{t}{\sqrt{n}}\right)\right]^n \approx \left(1 - \frac{\frac{1}{2}t^2}{n}\right)^n \longrightarrow \exp\left(-\frac{1}{2}t^2\right)$$

This coincides with the characteristic function of a standard normal, and so $Z \sim \mathcal{N}(0, 1)$

# Drift Diffusion Processes

We can embed this into a more general class known as drift diffusion processes

$$dX_t = \mu \cdot dt + \sigma \cdot dW_t$$

And even generate a continuous time analog of an AR(1) process (the Ornstein–Uhlenbeck process)

$$dX_t = \kappa(\mu - X_t) \cdot dt + \sigma \cdot dW_t$$

# Discrete Events

More general drift diffusion process is represented by class of jump diffusion processes.

We can ditch the continuity assumption but preserve independent increments by bolting on a Poisson jump process.

$$dX_t = \sigma \cdot dW_t + dJ_t$$

Parameter $\lambda$ determines jump rate. Size of jumps can even be random as long as i.i.d.

# Itô's Lemma

Consider the drift process $dx = \mu dt + \sigma dW$. For small $\Delta$, this evolves according to

$$x(t+\Delta) = x(t) + \Delta\mu + \sqrt{\Delta}\sigma z$$

where $z \sim \mathcal{N}(0,1)$

We can then approximate $v$ by taking a Taylor expansion

$$\begin{aligned} v(x(t+\Delta), t+\Delta) - v(x(t),t) \\ \approx \Delta\dot{v}(x(t),t) + \mathbb{E}[(\Delta\mu + \sqrt{\Delta}\sigma z)]v_x(x(t),t) \\ + \frac{1}{2}\mathbb{E}[(\Delta\mu + \sqrt{\Delta}\sigma z)^2]v_{xx}(x(t),t) \\ \approx \Delta\left[\dot{v}(x(t),t) + \mu v_x(x(t),t) + \frac{\sigma^2}{2}v_{xx}(x(t),t)\right] \end{aligned}$$

# Value Function

The general expression for a recursive form value function is

$$
\begin{aligned}
& v(x(t),t) \approx \Delta u(x(t)) + \exp(-\rho\Delta)\mathbb{E}\left[v(x(t+\Delta),t+\Delta)\right] \\
\Rightarrow\ & v(x,t) \approx \Delta u(x) + v(x,t) \\
& + \Delta\left[-\rho v(x,t) + \dot{v}(x,t) + \mu v_x(x,t) + \frac{\sigma^2}{2}v_{xx}(x,t)\right]
\end{aligned}
$$

Eliminating $\Delta$ and dropping explicit dependence yields the value equation

$$
\rho v - \dot{v} = u(x) + \mu v_x + \frac{\sigma^2}{2}v_{xx}
$$

# Distributions

We can apply similar techniques to distribution functions. We will denote the cumulative distribution of a variable $x$ by $F(x)$

Consider the case of a zero-centered O-U process with

$$dx = -\kappa(x - \mu)dt + \sigma dW$$

We can derive conditions for the evolution of the cumulative density $F$ over $x$: these results are sometimes called the **Fokker-Planck** or **Kolmogorov Forward Equations**

# Discrete Limit

Let's approach this again using discrete approximations (note inverted signs)

$$\begin{aligned} F(x,t) &\approx \mathbb{E}\left[F(x+\Delta\kappa(x-\mu)-\sqrt{\Delta}\sigma z, t-\Delta)\right] \\ &\approx F(x,t) - \dot{F}(x,t)\Delta + \mathbb{E}[\Delta\kappa(x-\mu)-\sqrt{\Delta}\sigma z]F_x(x,t) \\ &\quad + \frac{1}{2}\mathbb{E}[(\Delta\kappa(x-\mu)-\sqrt{\Delta}\sigma z)^2]F_{xx}(x,t) \\ &\approx F(x,t) + \Delta\left[-\dot{F}(x,t) + \kappa(x-\mu)\sigma F_x(x,t) + \frac{\sigma^2}{2}F_{xx}(x,t)\right] \end{aligned}$$

Thus we arrive at the desired result

$$\dot{F} = \kappa(x-\mu)F_x + \frac{\sigma^2}{2}F_{xx}$$

# O-U Limiting Distribution

In the limit, we will have $\dot{F} = 0$ so that (here $f = F_x$)

$$\kappa(x - \mu)F_x = -\frac{\sigma^2}{2}F_{xx} \quad \Rightarrow \quad \kappa(x - \mu)f = -\frac{\sigma^2}{2}f_x$$

We can solve the prior equation using the guess $\mathcal{N}(\mu, \gamma)$

$$f(x) = \frac{1}{\sqrt{2\pi\gamma^2}}\exp\left(-\frac{1}{2}\left(\frac{x - \mu}{\gamma}\right)^2\right)$$

Canceling common terms, this yields the following

$$\kappa = \frac{\sigma^2}{2\gamma^2} \quad \Rightarrow \quad \gamma^2 = \frac{\sigma^2}{2\kappa}$$

# Extending NCG

Consider the case of stochastic productivity $z$ as an O-U process

$$f(x,k) = e^x k^\alpha$$

The derived value function is then (same FOC as before)

$$\rho v = u(e^x k^\alpha - i) + (i - \delta k) v_k - \kappa(x-\mu) v_x + \frac{\sigma^2}{2} v_{xx}$$

We can characterize the stationary distribution over $(x,k)$ with

$$(i - \delta k) F_k = \kappa(x-\mu) F_x + \frac{\sigma^2}{2} F_{xx}$$

# Chebyshev Revisited

Such coupled systems are amenable to solution using spectral methods

A convenient construction in such a setting is the Vandermonde matrix

$$\underbrace{V}_{N\times D} = \begin{bmatrix} T_0(x_0) & \cdots & T_{D-1}(x_0) \\ \vdots & \ddots & \vdots \\ T_0(x_{N-1}) & \cdots & T_{D-1}(x_{N-1}) \end{bmatrix}$$

For a vector of coefficients $c$, we then have the Chebyshev approximation at points $(x_0, \ldots, x_{N-1})$ given by $V \cdot c$

# Derivative Matrices

We can calculate similar matrices to deal with arbitrary derivatives of functions as well

To do this, we use the Chebyshev derivative relation

$$\frac{dT_n(x)}{dx} = 2n \sum_{j \text{ odd}}^{n-1} T_j(x)$$

which is itself a linear map over Chebyshev terms and can thus be composed

The above equation can be expressed as a matrix operating on **coefficients** of size $[D \times D]$

# Linear Systems

Note that Chebyshev natively maps from $[-1, 1]$ so there is a transform involved that shows up in derivatives

Letting the derivative constructor be $D$, we can then reexpress differential equations as linear systems, for instance in the deterministic NCG case

$$\rho V = u(e^x k^\alpha - i) + (i - \delta k)VD$$

Thus we can solve a linear system $b = Ac$ where (with implicit broadcasting)

$$\begin{aligned} A &= \rho V - (i - \delta k)VD \\ b &= u(e^x k^\alpha - i) \end{aligned}$$

# Null Space

The case of distributions is slightly more complicated, where we find for the simple O-U case

$$0 = \kappa(x - \mu)VD + \frac{\sigma^2}{2}VD^2$$

Because probability sums to 1, system is singular: can either use a null space solver or impose $F[0] = 0$ and $F[-1] = 1$ directly and use system solver

# Multiple Dimensions

Going to higher dimensions requires some design choices, but the simplest way is to use the cross product of two Chebyshev approximations

$$f(x, y) = \sum_{i=0}^{D_1-1} \sum_{j=1}^{D_2-1} T_i(x) T_j(y)$$

From here we can still construct Vandermonde matrices, but they will be of size $\left[N_1 \times N_2 \times (D_1 \cdot D_2)\right]$

$$V_{ijk\ell} = T_i(x_k) T_j(y_\ell)$$