Lecture 4:

Diffusion Processes, Stochastic HJB Equations and Kolmogorov Forward Equations

---

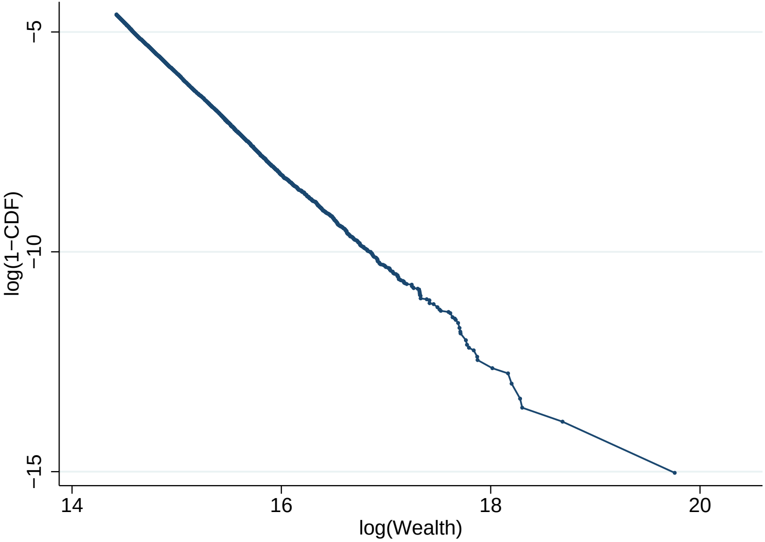
Distributional Macroeconomics

Part II of ECON 2149

Benjamin Moll

Harvard University, Spring 2018

For fun: Pareto Tail of Wealth Distribution in Norway



Outline

1. Diffusion processes
2. Hamilton-Jacobi-Bellman equations in stochastic settings (without derivation)
3. Ito's Lemma
4. Kolmogorov Forward Equations

Diffusion Processes

- A diffusion is simply a continuous-time Markov process (with continuous sample paths, i.e. no jumps)
  - for jumps, use Poisson process: very intuitive, briefly later
- Simplest possible diffusion: standard Brownian motion (sometimes also called "Wiener process")
- Definition: a standard Brownian motion is a stochastic process W which satisfies

$$W(t + \Delta t) - W(t) = \varepsilon_t \sqrt{\Delta t}, \quad \varepsilon_t \sim \mathcal{N}(0, 1), \quad W(0) = 0$$

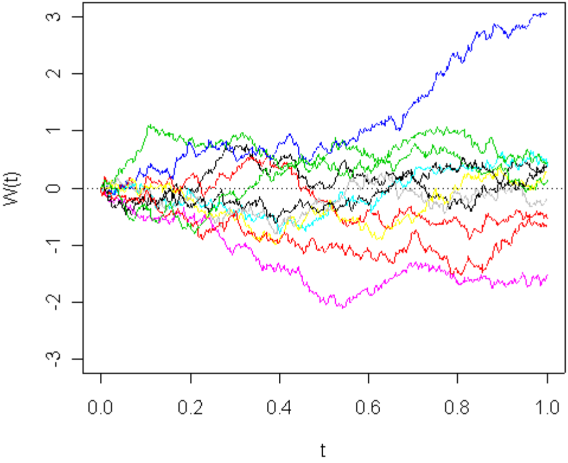
- Not hard to see

$$W(t) \sim \mathcal{N}(0, t)$$

- Continuous time analogue of a discrete time random walk:

$$W_{t+1} = W_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, 1)$$

Standard Brownian Motion



- Note: mean zero, $\mathbb{E}(W(t)) = 0$...
- ... but blows up $\mathrm{Var}(W(t)) = t$

Brownian Motion

- Can be generalized

$$x(t) = x(0) + \mu t + \sigma W(t)$$

- Since $\mathbb{E}(W(t)) = 0$ and $\text{Var}(W(t)) = t$

$$\mathbb{E}[x(t) - x(0)] = \mu t, \quad \text{Var}[x(t) - x(0)] = \sigma^2 t$$

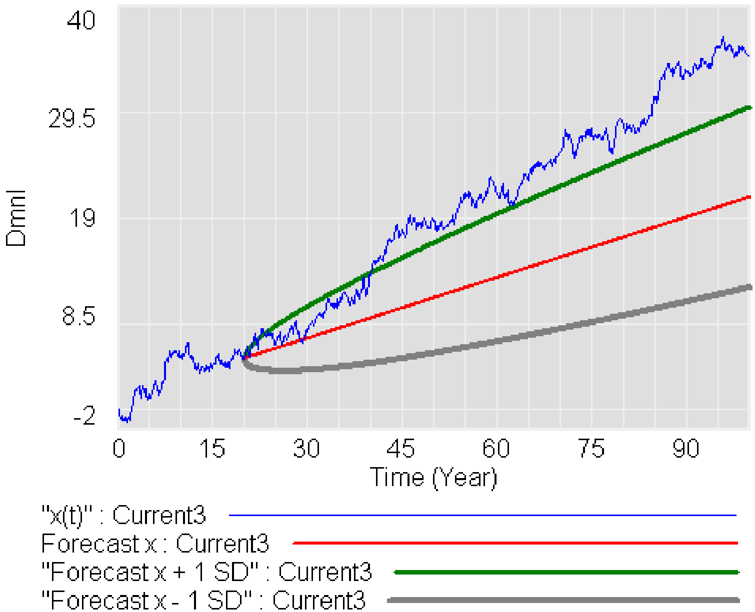
- This is called a Brownian motion with drift μ and variance σ^2
- Often useful to write this in differential form
  - recall $\Delta W(t) := W(t + \Delta t) - W(t) = \varepsilon_t \sqrt{\Delta t}$, $\varepsilon_t \sim \mathcal{N}(0, 1)$
  - use notation $dW(t) := \varepsilon_t \sqrt{dt}$, with $\varepsilon_t \sim \mathcal{N}(0, 1)$ and write

$$dx(t) = \mu dt + \sigma dW(t)$$

- This is called a stochastic differential equation
- Analogue of stochastic difference equation:

$$x_{t+1} = \mu + x_t + \sigma \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, 1)$$

Brownian Motion with Drift



Further Generalizations: Diffusion Processes

- Can be generalized further (suppressing dependence of x, W on t)

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

  where μ and σ are any non-linear etc etc functions.

- This is called a "diffusion process"

- $\mu(\cdot)$ is called the drift and $\sigma(\cdot)$ the diffusion.

- all results can be extended to the case where they depend on t, $\mu(x,t), \sigma(x,t)$ but abstract from this for now.

- The amazing thing about diffusion processes: by choosing functions μ and σ, you can get pretty much any stochastic process you want (except jumps)

Example 1: Ornstein-Uhlenbeck Process

- Brownian motion $dx = \mu dt + \sigma dW$ is not stationary (random walk). But the following process is

$$dx = \theta(\bar{x} - x)dt + \sigma dW$$

- Analogue of AR(1) process, autocorrelation $e^{-\theta} \approx 1 - \theta$

$$x_{t+1} = \theta\bar{x} + (1 - \theta)x_t + \sigma\varepsilon_t$$

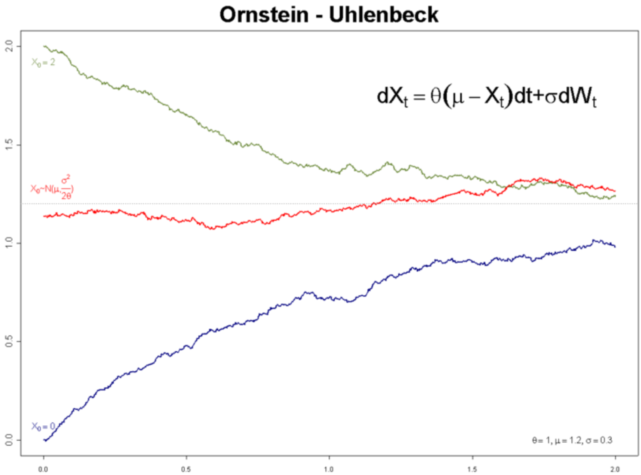
- That is, we just choose

$$\mu(x) = \theta(\bar{x} - x)$$

  and we get a nice stationary process!

- This is called an "Ornstein-Uhlenbeck process"

Ornstein-Uhlenbeck Process



- Can show: stationary distribution is $\mathcal{N}\left(\bar{x}, \frac{\sigma^2}{2\theta}\right)$

Example 2: "Moll Process"

- Design a process that stays in the interval $[0, 1]$ and mean-reverts around $1/2$

$$\mu(x) = \theta\,(1/2 - x)\,, \quad \sigma(x) = \sigma x(1 - x)$$

- That is

$$dx = \theta\,(1/2 - x)\,dt + \sigma x(1 - x)dW$$

- Note: diffusion goes to zero at boundaries $\sigma(0) = \sigma(1) = 0$ & mean-reverts $\Rightarrow$ always stay in $[0, 1]$

Other Examples

- Geometric Brownian motion:

$$dx = \mu x dt + \sigma x dW$$

$x \in [0, \infty)$, no stationary distribution:

$$\log x(t) \sim \mathcal{N}((\mu - \sigma^2/2)t, \sigma^2 t).$$

- Feller square root process (finance: "Cox-Ingersoll-Ross")

$$dx = \theta(\bar{x} - x)dt + \sigma\sqrt{x}dW$$

$x \in [0, \infty)$, stationary distribution is $Gamma(\gamma, 1/\beta)$, i.e.

$$g_\infty(x) \propto e^{-\beta x} x^{\gamma - 1}, \quad \beta = 2\theta\bar{x}/\sigma^2, \quad \gamma = 2\theta\bar{x}/\sigma^2$$

- Other processes in Wong (1964), "The Construction of a Class of Stationary Markoff Processes"

Stochastic HJB Equations

Stochastic Optimal Control

- Generic problem:

$$v(x_0) = \max_{\{\alpha(t)\}_{t \geq 0}} \mathbb{E}_0 \int_0^\infty e^{-\rho t} r(x(t), \alpha(t))\, dt$$

  subject to the law of motion for the state

$$dx(t) = f(x(t), \alpha(t))\, dt + \sigma(x(t)) dW(t)$$

  and $\alpha(t) \in A$, for $t \geq 0$, $x(0) = x_0$ given

- σ could depend on α as well – easy extension
- Deterministic problem: special case $\sigma(x) \equiv 0$
- In general $x \in \mathbb{R}^N, \alpha \in \mathbb{R}^M$. For now do scalar case.

Stochastic HJB Equation: Scalar Case

- Claim: the HJB equation is

$$\rho v(x) = \max_{\alpha \in A} r(x, \alpha) + v'(x) f(x, \alpha) + \frac{1}{2} v''(x) \sigma^2(x)$$

- Here: on purpose no derivation ("cookbook")

- In case you care, see any textbook, e.g. chapter 2 in Stokey (2008)

Just for Completeness: Multivariate Case

- Let $x \in \mathbb{R}^N, \alpha \in \mathbb{R}^M$
- For fixed x, define the $N \times N$ covariance matrix

$$\sigma^2(x) = \sigma(x)\sigma(x)'$$

(this is a function $\sigma^2 : \mathbb{R}^N \to \mathbb{R}^N \times \mathbb{R}^N$)

- The HJB equation is

$$\rho v(x) = \max_{\alpha \in A} \; r(x,\alpha) + \sum_{i=1}^{N} \frac{\partial v(x)}{\partial x_i} f_i(x,\alpha) + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial^2 v(x)}{\partial x_i \partial x_j} \sigma_{ij}^2(x)$$

- In vector notation

$$\rho v(x) = \max_{\alpha \in A} \; r(x,\alpha) + \nabla_x v(x) \cdot f(x,\alpha) + \frac{1}{2} \mathrm{tr}\left(\Delta_x v(x) \sigma^2(x)\right)$$

- $\nabla_x v(x)$: gradient of v (dimension $N \times 1$)
- $\Delta_x v(x)$: Hessian of v (dimension $N \times N$)

HJB Equation: Endogenous and Exogenous State

- Lots of problems have the form $x = (x_1, x_2)$
  - x_1: endogenous state
  - x_2: exogenous state

$$\begin{aligned} dx_1 &= \tilde{f}(x_1, x_2, \alpha)dt \\ dx_2 &= \tilde{\mu}(x_2)dt \quad + \quad \tilde{\sigma}(x_2)dW \end{aligned}$$

- Special case with

$$f(x, \alpha) = \begin{bmatrix} \tilde{f}(x_1, x_2, \alpha) \\ \tilde{\mu}(x_2) \end{bmatrix}, \quad \sigma(x) = \begin{bmatrix} 0 \\ \tilde{\sigma}(x_2) \end{bmatrix}$$

- Claim: the HJB equation is

$$\begin{aligned} \rho v(x_1, x_2) = \max_{\alpha \in A} \; & r(x_1, x_2, \alpha) + v_1(x_1, x_2)\tilde{f}(x_1, x_2, \alpha) \\ & + v_2(x_1, x_2)\tilde{\mu}(x_2) + \frac{1}{2} v_{22}(x_1, x_2)\tilde{\sigma}^2(x_2) \end{aligned}$$

Example: Real Business Cycle Model

$$v\left(k_0, z_0\right) = \max_{\{c(t)\}_{t\geq 0}} \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c(t)) dt$$

subject to

$$dk = (zF(k) - \delta k - c)dt$$

$$dz = \tilde{\mu}(z)dt + \tilde{\sigma}(z)dW$$

for $t \geq 0,\ k(0) = k_0,\ z(0) = z_0$ given

Here:

- $x_1 = k, x_2 = z, \alpha = c$
- $r(x, \alpha) = u(\alpha)$
- $f(x, \alpha) = \begin{bmatrix} x_2 F(x_1) - \delta x_1 - \alpha \\ \tilde{\mu}(x_2) \end{bmatrix}, \sigma(x) = \begin{bmatrix} 0 \\ \tilde{\sigma}(x_2) \end{bmatrix}$

Example: Real Business Cycle Model

- HJB equation is

$$\rho v(k,z) = \max_c \; u(c) + v_k(k,z)[zF(k) - \delta k - c]$$

$$+ v_z(k,z)\mu(z) + \frac{1}{2} v_{zz}(k,z)\sigma^2(z)$$

Example: Real Business Cycle Model

- Special Case 1: z is a geometric Brownian motion

$$dz = \mu z dt + \sigma z dW$$

$$\rho v(k,z) = \max_c \; u(c) + v_k(k,z)[zF(k) - \delta k - c] + v_z(k,z)\mu z + \frac{1}{2} v_{zz}(k,z)\sigma^2 z^2$$

See Merton (1975) for an analysis of this case

- Special Case 2: z is a Feller square root process

$$dz = \theta(\bar{z} - z)dt + \sigma\sqrt{z}dW$$

$$\rho v(k,z) = \max_c \; u(c) + v_k(k,z)[zF(k) - \delta k - c] + v_z(k,z)\theta(\bar{z} - z) + \frac{1}{2} v_{zz}(k,z)\sigma^2 z$$

Aside: Poisson Uncertainty

- Simplest way of modeling uncertainty in continuos time: two-state Poisson process
- $z_t \in \{z_1, z_2\}$ Poisson with intensities λ_1, λ_2
- Result: HJB equation is

$$\rho v_i(k) = \max_c \; u(c) + v_i'(k)(z_i F(k) - \delta k - c) + \lambda_i (v_j(k) - v_i(k))$$

for $i = 1, 2, j \neq i$

Special Case: Stochastic AK Model with log Utility

- Preferences: $u(c) = \log c$
- Technology: $zF(k) = zk$ (so maybe "zk model"?)
- Productivity z follows any diffusion

$$\begin{aligned}\rho v(k,z) = \max_c \ & \log c + v_k(k,z)(zk - \delta k - c) \\ & + v_z(k,z)\mu(z) + \frac{1}{2} v_{zz}(k,z)\sigma^2(z)\end{aligned}$$

- Claim: Optimal consumption is $c = \rho k$ and hence capital follows

$$dk = (z - \rho - \delta)k dt$$

$$dz = \mu(z) dt + \sigma(z) dW$$

- Solution properties? Simply simulate two SDEs forward in time

Special Case: Stochastic AK Model with log Utility

- Proof: Guess and verify

$$v(k,z) = \nu(z) + \kappa \log k$$

- FOC:

$$u'(c) = v_k(k,z) \quad \Leftrightarrow \quad \frac{1}{c} = \frac{\kappa}{k} \Leftrightarrow \quad c = \frac{k}{\kappa}$$

- Substitute into HJB equation

$$\rho[\nu(z) + \kappa \log k] = \log k - \log \kappa + \frac{\kappa}{k}[zk - \delta k - k/\kappa] \\ + \nu'(z)\mu(z) + \frac{1}{2}\nu''(z)\sigma^2(z)$$

- Collect terms involving $\log k \Rightarrow \kappa = 1/\rho \Rightarrow c = \rho k$ $\square$

- Remark: log-utility $\Rightarrow$ offsetting income and substitution effects of future $z \Rightarrow$ constant savings rate ρ

General Case: Numerical Solution with FD Method

- Want to solve:

$$\rho v(k,z) = \max_c \; u(c) + v_k(k,z)[zF(k) - \delta k - c]$$

$$+ v_z(k,z)\mu(z) + \frac{1}{2} v_{zz}(k,z)\sigma^2(z)$$

General Case: Numerical Solution with FD Method

It doesn't matter whether you solve ODEs or PDEs

$\Rightarrow$ everything generalizes

`http://www.princeton.edu/~moll/HACTproject/HJB_diffusion_implicit_RBC.m`

General Case: Numerical Solution with FD Method

- Solve on bounded grids $k_i, i = 1, ..., I$ and $z_j, j = 1, ..., J$
- Use short-hand notation $v_{i,j} = v(k_i, z_j)$. Approximate

$$v_k(k_i, z_j) \approx \frac{v_{i+1,j} - v_{i,j}}{\Delta k} \quad \text{or} \quad \frac{v_{i,j} - v_{i-1,j}}{\Delta k}$$

$$v_z(k_i, z_j) \approx \frac{v_{i,j+1} - v_{i,j}}{\Delta z} \quad \text{or} \quad \frac{v_{i,j} - v_{i,j-1}}{\Delta z}$$

$$v_{zz}(k_i, z_j) \approx \frac{v_{i,j+1} - 2v_{i,j} + v_{i,j-1}}{(\Delta z)^2}$$

- Discretized HJB

$$\begin{aligned}
\rho v_{i,j} =& u(c_{i,j}) + v_k(k_i, z_j)(z_j F(k_i) - \delta k_i - c_{i,j}) \\
&+ v_z(k_i, z_j)\mu(z_j) + \frac{1}{2} v_{zz}(k_i, z_j)\sigma^2(z_j) \\
c_{i,j} =& (u')^{-1}[v_k(k_i, z_j)]
\end{aligned}$$

Numerical Solution: Boundary Conditions?

- Upwind method in k-dimension $\Rightarrow$ no boundary conditions needed

- Do need boundary conditions in z-dimension

$$\begin{aligned} v_z(k, z_1) = 0 \quad \text{all } k \quad &\Rightarrow \quad v_{i,0} = v_{i,1} \\ v_z(k, z_J) = 0 \quad \text{all } k \quad &\Rightarrow \quad v_{i,J+1} = v_{i,J} \end{aligned}$$

- These correspond to "reflecting barriers" at lower and upper bounds for productivity, z_1 and z_J (Dixit, 1993)

General Case: Numerical Solution with FD Method

- Stack value function $v_{i,j}$ into vector $\mathbf{v}$ of length $I \times J$
  - I usually stack it as "endogenous state variable first"

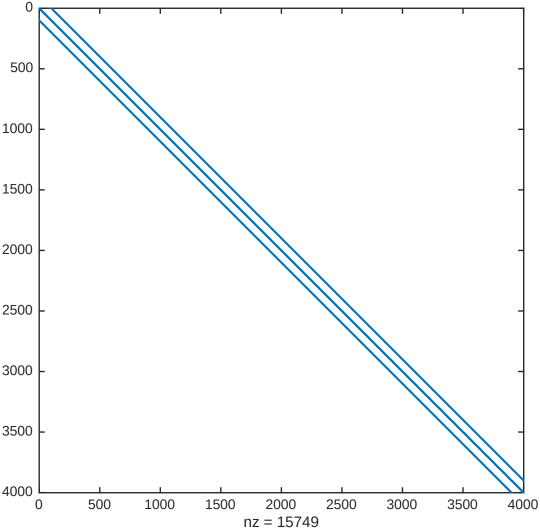
$$\mathbf{v} = (v_{1,1}, v_{2,1}, \ldots, v_{I,1}, v_{1,2}, \ldots, v_{I,2}, v_{1,3}, \ldots, v_{I,J})'$$

  - here: doesn't really matter
- End up with system of $I \times J$ non-linear equations

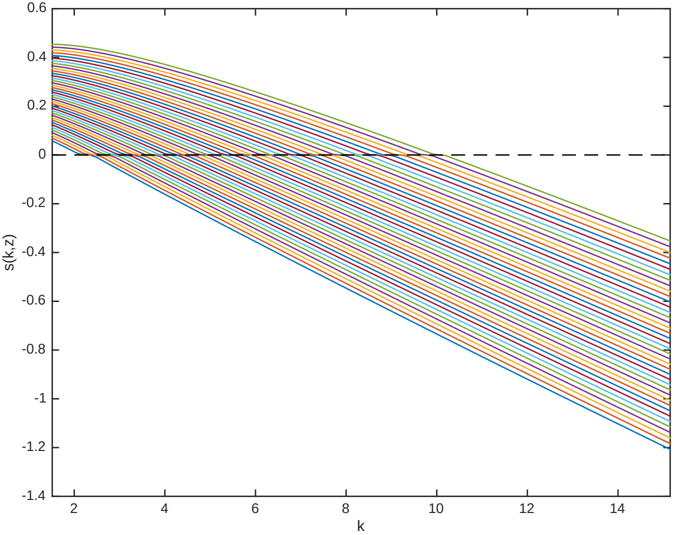
$$\rho \mathbf{v} = \mathbf{u}(\mathbf{v}) + \mathbf{A}(\mathbf{v})\mathbf{v}$$

- Solve exactly as before
  - upwind scheme
  - implicit method preferred to explicit method

Visualization of **A** (output of `spy(A)` in Matlab)



Saving Policy Function



Ito's Lemma

Ito's Lemma

- Let x be a scalar diffusion

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

- We are interested in the evolution of $y(t) = f(x(t))$ where f is any twice differentiable function

- Lemma: $y(t) = f(x(t))$ follows

$$df(x) = \left(\mu(x)f'(x) + \frac{1}{2}\sigma^2(x)f''(x)\right)dt + \sigma(x)f'(x)dW$$

- Extremely powerful because it says that any (twice differentiable) function of a diffusion is also a diffusion

- Can also be extended to vectors

- FYI: this is also where the $v'(x)\mu(x) + \frac{1}{2}v''(x)\sigma^2(x)$ term in the HJB equation comes from (it's $\frac{\mathbb{E}[dv(x)]}{dt}$)

Application: Brownian vs. Geometric Brownian Motion

- Let x be a geometric Brownian motion

$$dx = \mu x dt + \sigma x dW$$

- Claim: $y = \log x$ is a Brownian motion with drift $\mu - \sigma^2/2$ and variance σ^2

- Derivation: $f(x) = \log x$, $f'(x) = 1/x$, $f''(x) = -1/x^2$.
  By Ito's Lemma

$$\begin{aligned} dy = df(x) &= \left(\mu x(1/x) + \frac{1}{2}\sigma^2 x^2(-1/x^2)\right) dt + \sigma x(1/x) dW \\ &= \left(\mu - \sigma^2/2\right) dt + \sigma dW \end{aligned}$$

- Note: naive derivation would have used $dy = dx/x$ and hence

$$dy = \mu dt + \sigma dW \quad \text{wrong unless } \sigma = 0!$$

Just for Completeness: Multivariate Case

- Let $x \in \mathbb{R}^N$. For fixed x, define the $N \times N$ covariance matrix

$$\sigma^2(x) = \sigma(x)\sigma(x)'$$

- Ito's Lemma:

$$df(x) = \left(\sum_{i=1}^{N} \mu_i(x) \frac{\partial f(x)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sigma_{ij}^2(x) \frac{\partial^2 f(x)}{\partial x_i x_j} \right) dt + \sum_{i=1}^{N} \sigma_i(x) \frac{\partial f(x)}{\partial x_i} dW_i$$

- In vector notation

$$df(x) = \left(\nabla_x f(x) \cdot \mu(x) + \frac{1}{2} \operatorname{tr}\left(\Delta_x f(x) \sigma^2(x) \right) \right) dt + \nabla_x f(x) \cdot \sigma(x) dW$$

- $\nabla_x f(x)$: gradient of f (dimension $m \times 1$)
- $\Delta_x f(x)$: Hessian of f (dimension $m \times m$)

Kolmogorov Forward Equations

Kolmogorov Forward Equations

- Let x be a scalar diffusion

$$dx = \mu(x)dt + \sigma(x)dW, \quad x(0) = x_0$$

- Suppose we're interested in the evolution of the distribution of x, $g(x, t)$, and in particular in the stationary distribution $g(x)$
- Natural thing to care about especially in heterogenous agent models
- Example 1: x = wealth
  - $\mu(x)$ determined by savings behavior and return to investments
  - $\sigma(x)$ by return risk
  - microfound later
- Example 2: x = city size, will cover later

Kolmogorov Forward Equations

- Fact: Given an initial distribution $g(x, 0) = g_0(x)$, $g(x, t)$ satisfies the PDE

$$\frac{\partial g(x,t)}{\partial t} = -\frac{\partial}{\partial x}[\mu(x)g(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[\sigma^2(x)g(x,t)]$$

- This PDE is called the "Kolmogorov Forward Equation"
- Note: in math this often called "Fokker-Planck Equation"
- Corollary: if a stationary distribution $g(x)$ exists, it satisfies the ODE

$$0 = -\frac{d}{dx}[\mu(x)g(x)] + \frac{1}{2}\frac{d^2}{dx^2}[\sigma^2(x)g(x)]$$

- Remark: as usual, stationary distribution defined as "if you start there, you stay there"
  - $g(x)$ s.t. if $g(x, t) = g(x)$, then $g(x, \tau) = g(x)$ for all $\tau \geq t$

Just for Completeness: Multivariate Case

- Let $x \in \mathbb{R}^N$
- As before, define the $N \times N$ covariance matrix

$$\sigma^2(x) = \sigma(x)\sigma(x)'$$

- The Kolmogorov Forward Equation is

$$\frac{\partial g(x,t)}{\partial t} = -\sum_{i=1}^{N} \frac{\partial}{\partial x_i}[\mu_i(x)g(x,t)] + \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \frac{\partial^2}{\partial x^2}[\sigma_{ij}^2(x)g(x,t)]$$

Application: Stationary Distribution of RBC Model

- Recall RBC Model

$$\rho v(k,z) = \max_c \; u(c) + v_k(k,z)[zF(k) - \delta k - c] + v_z(k,z)\mu(z) + \frac{1}{2}v_{zz}(k,z)\sigma^2(z)$$

- Denote the optimal policy function by

$$s(k,z) = zF(k) - \delta k - c(k,z)$$

- Then the distribution $g(k,z,t)$ solves

$$\frac{\partial g(k,z,t)}{\partial t} = -\frac{\partial}{\partial k}[s(k,z)g(k,z,t)] - \frac{\partial}{\partial z}[\mu(z)g(k,z,t)] + \frac{1}{2}\frac{\partial^2}{\partial z^2}[\sigma^2(z)g(k,z,t)]$$

- Numerical solution with FD method: later