Dynamic Programming and Applications

Stochastic Dynamic Programming in Continuous Time

Lectures 5 – 6

Andreas Schaab

Outline

Part 1: Stochastic processes, Brownian motion, and stochastic differential equations

- 1. Stochastic processes in continuous time
- 2. Continuous time Markov chains
- 3. Brownian motion
- 4. Diffusion processes
- 5. Ito's Lemma
- 6. Poisson processes
- 7. The generator of a stochastic process

Outline

Part 2: Optimization with stochastic dynamics

- 1. Stochastic neoclassical growth model
- 2. Stochastic neoclassical growth with diffusion process
- 3. Stochastic neoclassical growth with Poisson process

Outline

Part 3: Applications

- 1. Consumption-savings with stochastic income fluctuations
- 2. Portfolio choice
- 3. Firm profit maximization
- 4. XXX

Part 1: Stochastic Processes

1. Stochastic processes in continuous time

Definition. A **stochastic process** is a time-indexed sequence of random variables.

- A random variable maps an "event" into a scalar, a stochastic process maps an event into a path
- Formally, an event is $\omega \in \Omega$ and a random variable $X : \Omega \to \mathbb{R}$ (learn some basic measure theory!)
- A stochastic process is a sequence $\{X_t\}_{t\geq 0}$ such that each X_t is a random variable. In discrete time $t\in\{0,1,\ldots\}$ and in continuous time $t\in[0,\infty)$
- We can also think of a stochastic process as a random variable, but one that maps into a different space—a function space!

Definition. Let $X = \{X_t\}_{t \geq 0}$ be a sequence of random variables taking values in a finite or countable state space \mathcal{X} . Then X is a *continuous-time Markov chain* if it satisfies the *Markov property*: For any sequence $0 \leq t_1 < t_2 < \ldots < t_n$ of times

$$\mathbb{P}(X_{t_n} = x \mid X_{t_1}, \dots, X_{t_{n-1}}) = \mathbb{P}(X_{t_n} = x \mid X_{t_{n-1}})$$

Definition. Process X is **time-homogeneous** or **stationary** if the conditional probability does not depend on the current time, i.e., for $x, y \in \mathcal{X}$:

$$\mathbb{P}(X_{t+s} = x \mid X_s = y) = \mathbb{P}(X_t = x \mid X_0 = y)$$

• Alternatively: If all possible joint distributions are independent of time *t*.

• The *transition density* of process X is denoted $p(t, x \mid s, y)$ and is defined as

$$\mathbb{P}(X_t \in A \mid Y_s = y) = \int_A p(t, x \mid s, y) dx$$

for any (Borel) set $A \subset \mathcal{X}$. In words: $p(t, x \mid s, y)$ is the probability (density) that process X_t ends up at $X_t = x$ at time t if it started at $X_s = y$ at time s

• Conditional expectation can be written as: $\mathbb{E}[f(X_t) \mid X_0 = y] = \int p(t, x \mid 0, y) f(x) dx$

Definition. A process *X* is a **martingale** if it satisfies

$$\mathbb{E}\left[X_{t+s} \mid \text{ information available at } t\right] = X_t$$

• The most important martingale in macroeconomics: Suppose $\beta R_t = 1$ then

$$\mathbb{E}_t\Big[u'(C_{t+s})\Big]=u'(C_t).$$

- Efficient markets hypothesis in finance $\approx \mathbb{E}_t(P_{t+s}) = P_t$
- A random walk is a martingale (but not every martingale is a random walk)

Example:

- Consider the two-state employment process $z_t \in \{z^L, z^H\}$ with transition rates λ^{LH} (from L to H) and λ^{HL} (from H to L)
- The associated transition matrix (generator) is

$$\mathcal{A}^z = \begin{pmatrix} -\lambda^{LH} & \lambda^{LH} \\ \lambda^{HL} & -\lambda^{HL} \end{pmatrix}$$

- Interpretation: households transition *out of* state i at rate λ^{ij}
- Notice: In discrete time, Markov transition matrix rows sum to 1. Here, rows sum to 0 (mass preservation)

2. Brownian motion

- Brownian motion is the most
- Einstein (1905) uses Brownian motion to model motion of particles
- We will introduce / discuss Brownian motion from 3 perspectives

Brownian motion: perspective #1

• Consider a process X(t). Every Δ time step, the process either goes up or down by h

$$\Delta X \equiv X(t+\Delta) - X(t) = egin{cases} +h & \text{with probability } p \ -h & \text{with probability } q = 1-p \end{cases}$$

Then we have:

$$\mathbb{E}(\Delta X) = ph - qh = (p - q)h$$

$$\mathbb{E}((\Delta X)^2) = ph^2 + qh^2 = h^2$$

$$Var(\Delta X) = \mathbb{E}[\Delta X - \mathbb{E}(\Delta X)]^2 = 4pqh^2$$

• Notice that X(t) - X(0) is a binomial random variable with

$$\mathbb{E}[X(t) - X(0)] = n(p - q)h = t(p - q)\frac{h}{\Delta}$$

$$Var[X(t) - X(0)] = n4pqh^2 = t4pq\frac{h^2}{\Delta}$$

where $n = t/\Delta$ is the number of jumps in interval [0, t]

Next, let

$$h=\sigma\sqrt{\Delta}$$
 and $p=rac{1}{2}\Big[1+rac{\mu}{\sigma}\sqrt{\Delta}\Big]$

This implies

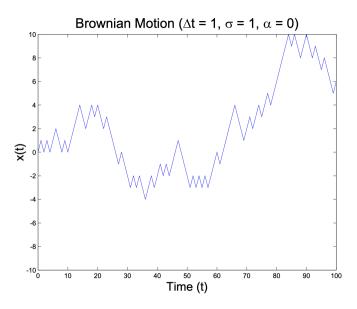
$$(p-q) = \frac{\mu}{\sigma} \sqrt{\Delta}$$

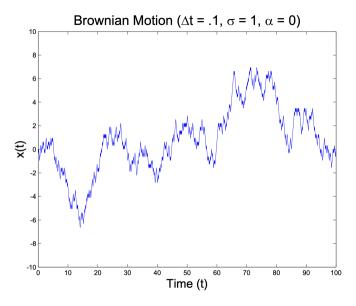
$$\mathbb{E}[X(t) - X(0)] = t \frac{\mu}{\sigma} \sqrt{\Delta} \sigma \frac{\sqrt{\Delta}}{\Delta} = \mu t$$

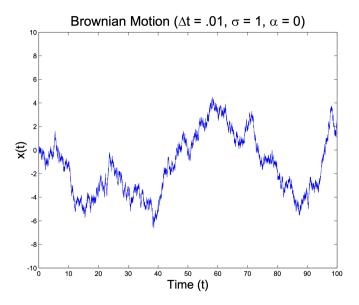
$$Var[X(t) - X(0)] \longrightarrow_{\mathsf{as}\Delta \to 0} \sigma^2 t$$

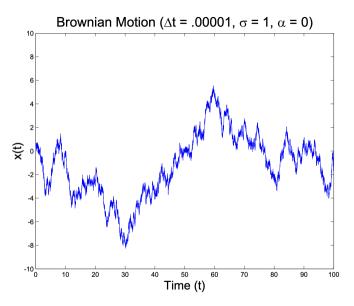
Implications:

- Vertical movements proportional to $\sqrt{\Delta}$ not Δ
- Convergence in distribution: $X(t) X(0) \rightarrow^D \mathcal{N}(\mu t, \sigma t)$ since Binomial \rightarrow^D Normal
- Distance traveled (length of curve) during $t \in [0,1]$ is $= nh = \frac{1}{\Delta}\sigma\sqrt{\Delta} = \sigma\frac{1}{\sqrt{\Delta}} \to \infty$
- Time derivative $\mathbb{E} rac{dX}{dt}$ doesn't exist: $rac{\Delta X}{\Delta} = rac{\pm \sigma \sqrt{\Delta}}{\Delta} = rac{\pm \sigma}{\sqrt{\Delta}}
 ightarrow \infty$
- $\frac{\mathbb{E}\Delta X}{\Delta} = \frac{\mu h^2/\sigma^2}{\Delta} = \mu$ so we can write $\mathbb{E}(dX) = \mu dt$
- $\frac{Var(\Delta X)}{\Delta} \rightarrow \sigma^2$ so we can write $Var(dX) = \sigma^2 dt$









Brownian motion: perspective #2

Consider a random walk process in discrete time:

$$W_{t+1} = W_t + \epsilon_t$$
, where $W_0 = 0$ and $\epsilon_t \sim \mathcal{N}(0, 1)$

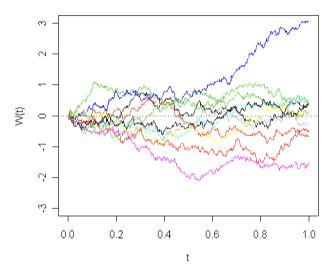
We can extend this to a generalized time step:

$$W_{t+\Delta} = W_t + \sqrt{\Delta} \epsilon_t$$
, where $W_0 = 0$ and $\epsilon_t \sim \mathcal{N}(0,1)$

Why the square root?

$$Var(W_{t+\Delta} - W_t) = Var(\sqrt{\Delta}\epsilon_t) = \Delta Var(\epsilon_t) = \Delta \implies Var(W_t) = t$$

• You get Brownian motion by taking limit $\Delta \to 0$: Brownian motion = continuous time random walk



Brownian motion: perspective #3

Definition. Brownian motion $\{B_t\}_{t\geq 0}$ is a stochastic process with properties:

- (i) $B_0 = 0$
- (ii) (*Independent increments*) For non-overlapping $0 \le t_1 < t_2 < t_3 < t_4$, we have $B_{t_2} B_{t_1}$ independent from $B_{t_4} B_{t_3}$
- (iii) (Normal, stationary increments) $B_t B_s \sim \mathcal{N}(0, t-s)$ for any $0 \le s < t$
- (iv) (*Continuity of paths*) The sample paths of B_t are continuous

- Brownian motion is the only stochastic process with stationary and independent increments that's also continuous
- Brownian motion is a Markov process
- Brownian motion is nowhere differentiable
- Brownian motion is a martingale
- Summary: Brownian motion is a continuous time random walk with zero drift and unit variance

Other properties of Brownian motion

- $B_t \sim B_t B_0 \sim \mathcal{N}(0,t)$
- Important: $dB_t \sim \mathcal{N}(0, dt)$ because

$$dB_t \approx B_{t+\Delta} - B_t \sim \mathcal{N}(0, t + \Delta - t = \mathcal{N}(0, \Delta))$$

and now take $\Delta \to dt$ (continuous-time limit)

• Alternatively: $B_{t+\Delta} - B_t \sim \mathcal{N}(0, \Delta) \sim \epsilon_t \sqrt{\Delta}$ where $\epsilon_t \sim \mathcal{N}(0, 1)$. So as $\Delta \to dt$,

$$\mathbb{E}(dB_t) = \mathbb{E}(\epsilon_t \sqrt{dt}) = 0$$

$$\mathbb{E}[(dB_t)^2] = \mathbb{E}[(\epsilon_t \sqrt{dt})^2] = dt$$

4. Ito's Lemma

- Q: How do functions / transformations of stochastic processes evolve?
- Warm-up: Consider the process x_t characterized by the ODE

$$\dot{x}_t = \frac{dx_t}{dt} = \mu_t \qquad \Longrightarrow \qquad dx_t = \mu_t dt$$

• Let $y_t = f(x_t)$. What do we know about \dot{y}_t or dy_t ? Using calculus, we simply have:

$$\frac{dy_t}{dt} = f'(x_t)\frac{dx_t}{dt} \qquad \Longrightarrow \qquad dy_t = f'(x_t)dx_t = f'(x_t)\mu_t dt$$

• Next, suppose $y_t = f(t, x_t)$. Then using $f_x = \frac{\partial f}{\partial x}$ we have:

$$\frac{dy_t}{dt} = f_x(t, x_t) \frac{dx_t}{dt} + f_t(t, x_t) \qquad \Longrightarrow \qquad dy_t = f_x(t, x_t) \mu_t dt + f_t(t, x_t) dt$$

- Now what about functions of Brownian motion or other diffusion processes?
- What to remember for Ito's lemma: (i) take a 2^{nd} -order Taylor expansion and (ii) remember that

$$(dt)^2 = 0$$
 and $(dt)(dB) = 0$ and $(dB)^2 = 0$

• Suppose $Y_t = f(B_t)$ where B_t is Brownian motion. Show that:

$$dY_t = f'(B_t)dB_t + \frac{1}{2}f''(B_t)dt$$

• Suppose $Y_t = f(t, B_t)$. Show that:

$$dY_t = f_t(t, B_t)dt + f_b(t, B_t)dB_t + \frac{1}{2}f_{bb}(t, B_t)dt$$

• This is Ito's lemma, a core building block of stochastic calculus

5. Diffusion processes

• Consider the class of stochastic processes $\{X_t\}$ characterized by

$$dX_t = \mu_t(X_t)dt + \sigma_t(X_t)dB_t$$

- These are called **diffusion processes** (continuous sample paths)
- For any twice differentiable function $Y_t = f(t, X_t)$, we have

$$dY = f_t dt + f_x dX + \frac{1}{2} f_{xx} (dX)^2$$

= $f_t dt + f_x (\mu(t, X) dt + \sigma(t, X) dB) + \frac{1}{2} f_{xx} \sigma(t, X)^2 dt$
= $\left(f_t + f_x \mu(t, X) + \frac{1}{2} f_{xx} \sigma(t, X)^2 \right) dt + \sigma(t, X) dB$

· Powerful result: Any (twice differentiable) function of a diffusion is also a diffusion

6. Stochastic differential equations

Geometric Brownian motion:

- Stochastic differential equations (SDEs) add noise / uncertainty to ordinary differential equations (ODEs)
- Start with $\dot{X}_t = \mu X_t$ with solution $X_t = X_0 e^{\mu t}$
- Rewrite as $dX_t = \mu X_t dt$ and "add noise" (using Brownian motion):

$$dX_t = \mu X_t dt + \sigma X_t dB_t$$

Solution to geometric Brownian motion (work this out on homework):

$$X_t = X_0 e^{\mu t - \frac{\sigma^2}{2}t + \sigma B_t}$$

Brownian Motion with Drift



"Forecast x - 1 SD": Current3

Ornstein-Uhlenbeck (OU) process:

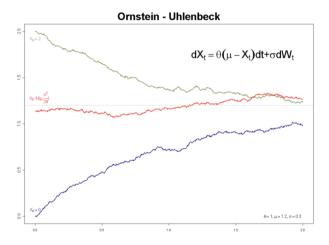
Continuous time analog of the AR(1) process

$$dz_t = \theta(\bar{z} - z_t)dt + \sigma dB_t$$

- Popular model for earnings risk and income fluctuations
- Auto-correlation of $e^{-\theta} \approx 1 \theta$. Compare to:

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

• Stationary distribution is $\mathcal{N}(\bar{z}, \frac{\sigma^2}{2\theta})$



- Diffusion processes allow us to construct many other processes with desired properties by simply choosing $\mu(t,X)$ and $\sigma(t,X)$
- Process that stays in interval [0,1] and mean-reverts around 1/2:

$$dX = \theta(\frac{1}{2} - X)dt + \sigma X(1 - X)dB$$

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

$$dX = \theta(\bar{X} - X)dt + \sigma\sqrt{X}dB$$

with stationary distribution $\sim \Gamma\!\left(\frac{2\theta \bar{X}}{\sigma^2}, \frac{\sigma^2}{2\theta \bar{X}}\right)$

Part 2: Optimization with stochastic dynamics

1. The generator of a stochastic process

Definition. The **generator** of a stochastic process X is defined (for any test function f) as

$$\mathcal{A}f = \lim_{\Delta t \to 0} \mathbb{E}_t \frac{f(t + \Delta t, X(t + \Delta t)) - f(t, X(t))}{\Delta t}$$

- The generator A tells us how the stochastic process is *expected* to evolve
- The generator A is a functional operator

Generator for diffusion processes

We start with the diffusion process

$$dX = \mu(t, X)dt + \sigma(t, X)dB$$

On homework, you will show that:

$$\mathcal{A}f = \partial_t f(t, X) + \mu(t, X)\partial_X f(t, X) + \frac{1}{2}\sigma(t, X)^2\partial_{XX} f(t, X)$$

For the general / multi-dimensional version see Oksendal

Generator for Poisson processes

- Next, we consider the poisson process $\{Y_t\}$ where $Y_t \in \{Y^1, Y^2\}$. This is a two-state Markov chain in continuous time.
- We assume that the Poisson intensity / arrival rate / hazard rate is λ
- · The generator is now given by

$$\mathcal{A}f(Y^j) = \lambda \left[f(Y^{-j}) - f(Y^j) \right]$$

- Intuition: at rate λ you transition, so you lose the value of your current state, $f(Y^j)$, and obtain the value of the new state, $f(Y^{-j})$
- Again see Oksendal for general version of this and more details

2. Stochastic Neoclassical Growth Model

Preferences. Lifetime utility of representative household

$$\mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt$$

Technology. Final consumption good produced using technology

$$y_t = F(k_t, z_t)$$
 where $F(0, z) = 0$, and $F_k, F_z > 0$, and $F_{kk} < 0$

where z_t is exogenous productivity, and capital accumulation technology is

$$dk_t = (i_t - \delta k_t)dt.$$

Resource constraint for the final good is: $y_t = c_t + i_t$

Endowment. At time t = 0, economy's initial state is (k_0, z_0)

Planning problem. Taking as given (k_0, z_0) , choose an allocation to maximize lifetime utility of representative household subject to technologies and resource constraints:

$$V(k_0, z_0) = \max \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt$$

subject to

$$y_t = F(k_t, z_t)$$

 $y_t = c_t + i_t$
 $dk_t = (i_t - \delta k_t)dt$
 dz_t exogenous

and taking as given (k_0, z_0)

Combining:
$$dk_t = [F(k_t, z_t) - c_t - \delta k_t]dt$$

3. Productivity as a diffusion process

We start with diffusion process:

$$dz_t = -\theta z_t dt + \sigma dB_t,$$

where θ and σ are constants

- This is a continuous-time, mean-reverting AR(1) process called the Ornstein-Uhlenbeck process
- State space is now given by

$$\left\{(k,z)\mid k\in[0,\bar{k}] \text{ and } z\in[\underline{z},\bar{z}]\right\}$$

· In discrete time, we would have

$$V(k_t, z_t) = \max_{c} \left\{ u(c)\Delta t + \frac{1}{1 + \rho \Delta t} \mathbb{E}_t V(k_{t+\Delta t}, z_{t+\Delta t}) \right\}$$

• Difference from previous lecture: \mathbb{E} because there is uncertainty

$$(1 + \rho \Delta t)V(k_t, z_t) = \max_{c} \left\{ (1 + \rho \Delta t)u(c)\Delta t + \mathbb{E}_t V(k_{t+\Delta t}, z_{t+\Delta t}) \right\}$$

$$\rho \Delta t V(k_t, z_t) = \max_{c} \left\{ u(c)\Delta t + \mathbb{E}_t V(k_{t+\Delta t}, z_{t+\Delta t}) - V(k_t, z_t) \right\}$$

$$\rho V(k_t, z_t) = \max_{c} \left\{ u(c) + \mathbb{E}_t \frac{V(k_{t+\Delta t}, z_{t+\Delta t}) - V(k_t, z_t)}{\Delta t} \right\}$$

• Take limit $\Delta t \to 0$ and drop time subscripts:

$$\rho V(k,z) = \max_{c} \left\{ u(c) + \mathbb{E} \frac{dV(k,z)}{dt} \right\}$$

• What remains? Characterizing continuation value $\frac{d}{dt}V(k,z)$ (i.e., characterizing how process dV evolves)

• The generator $\mathcal A$ is exactly the answer to this question! I.e.,

$$\mathbb{E}\frac{dV(k,z)}{dt} = \mathcal{A}V(k,z)$$

$$= \left(F(k,z) - \delta z - c\right)\partial_k V(k,z) - \theta z \partial_z V(k,z) + \frac{\sigma^2}{2}\partial_{zz} V(k,z)$$

Therefore, we arrive at the Hamilton-Jacobi-Bellman equation

$$\rho V(k,z) = \max_{c} \left\{ u(c) + \left(F(k,z) - \delta z - c \right) \partial_{k} V(k,z) - \theta z \partial_{z} V(k,z) + \frac{\sigma^{2}}{2} \partial_{zz} V(k,z) \right\}$$

with first-order condition

$$u'(c(k,z)) = \partial_k V(k,z)$$

4. Productivity as a Poisson process

- Next, consider Poisson process for $\{z_t\}$ with $z_t \in \{z^L, z^H\}$
- · Generator now given by

$$\mathcal{A}V(k,z^{j}) = \left(F(k,z) - \delta z - c\right)\partial_{k}V(k,z) + \lambda\left[V(k,z^{-j}) - V(k,z^{j})\right]$$

- Note: derivation of HJB exactly as before $up\ to$ characterizing $\mathbb{E}[dV]$
- · With Poisson process, HJB becomes

$$\rho V(k, z^{j}) = \max_{c} \left\{ u(c) + \left(F(k, z) - \delta z - c \right) \partial_{k} V(k, z) + \lambda \left[V(k, z^{-j}) - V(k, z^{j}) \right] \right\}$$

with first-order condition

$$u'(c(k,z^j)) = \partial_k V(k,z^j)$$

Part 3: Applications

1. Consumption-savings with income fluctuations

- · Economy is populated by representative household that faces income risk
- Household accumulates wealth according to

$$\dot{a}_t = ra_t + e^{z_t} - c_t$$

subject to borrowing constraint $a_t \geq 0$

- Preferences again: $V_0 = \max \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt$
- Income follows diffusion process: $dy_t = -\theta y_t dt + \sigma dB_t$
- · Away from borrowing constraint, HJB given by

$$\rho V = \max_{c} \left\{ u(c) + (ra + e^{z} - c)V_{a} - \theta z V_{z} + \frac{\sigma^{2}}{2}V_{zz} \right\}$$

with $V_a = \partial_a V(a, z)$ (you'll see this often)

2. Firm profit maximization

- Firm maximizes NPV of profit: $V_0 = \max \mathbb{E}_0 \int_0^\infty e^{-rt} \pi_t dt$
- For now, profit given by: $\pi_t = A_t n_t^{\alpha} w_t n_t$ where firm chooses labor n_t Assume $\alpha < 1$, so this is a decreasing-returns production function
- Firm is small and takes wage $\{w_t\}$ as given (wages determined in general equilibrium)
- Productivity follows two-state high-low process, with $A_t \in \{A^{\text{rec}}, A^{\text{boom}}\}$
- Recursive representation: A is only state variable, $w_t = w(A_t)$

$$rV(A^{\mathsf{boom}}) = \max_{n} \left\{ A^{\mathsf{boom}} n^{\alpha} - w(A^{\mathsf{boom}}) n + \lambda \Big[V(A^{\mathsf{rec}}) - V(A^{\mathsf{boom}}) \Big] \right\}$$

with first-order condition

$$n = \left(\frac{\alpha A^j}{w(A^j)}\right)^{\frac{1}{1-\alpha}}$$

3. Capital investment with adjustment cost

- Firm again maximizes NPV of profit: $V_0 = \max \mathbb{E}_0 \int_0^\infty e^{-rt} \pi_t dt$
- Now: let $\psi(\cdot)$ denote an adjustment cost

$$\pi_t = e^{A_t} k_t^{\alpha} - Q_t \iota_t - \psi(\iota_t, k_t)$$
$$dk_t = (\iota_t - \delta k_t) dt$$
$$dA_t = -\theta A_t dt + \sigma dB_t$$

- Firm is small and takes capital price as given
- Recursive representation in terms of (k, A), i.e., $Q_t = Q(k_t, A_t)$

$$rV(k,A) = \max_{\iota} \left\{ e^{A_t} k_t^{\alpha} - Q(A) \iota_t - \psi(\iota_t, k_t) + (\iota - \delta k) \partial_k V(k, A) - \theta A \partial_A V(k, A) + \frac{\sigma^2}{2} \partial_{AA} V(k, A) \right\}$$

with first-order condition: $Q(k, A) + \partial_{\iota} \psi(\iota(k, A), k) = \partial_{k} V(k, A)$

4. Investing in stocks

- Suppose you optimize lifetime utility $V_0 = \mathbb{E}_0 \int_0^\infty u(c_t) dt$
- You can trade two assets: riskfree bond (return rdt), and risky stock

$$dR = (r + \pi)dt + \sigma dB$$
, where π is the equity premium

• You have wealth a_t and invest a share θ_t in stocks, thus,

$$da_t = \theta_t a_t dR_t + (1 - \theta_t) a_t r_t dt + y - c_t$$

or, rearranging, and dropping t subscripts

$$da = ra + \theta a \pi dt + y - c + \theta a \sigma dB$$

HJB becomes:

$$\rho V(a) = \max_{c,\theta} \left\{ u(c) + (ra + \theta a \pi dt + y - c)V'(a) + \frac{1}{2}(\sigma \theta a)^2 V''(a) \right\}$$

with FOCs: (i)
$$u'(c) = V'(a)$$
 and (ii) $\theta = -\frac{\pi}{\sigma^2} \frac{V'(a)}{aV''(a)}$

5. Real Business Cycles