Advanced Macroeconomics I

Chapter II:
Technical Preliminaries

Prof. Dr. Johannes Pfeifer

University of Cologne

Winter Term 2019

Outline

1. Difference equations
2. Dynamic optimization

Introduction

- In this chapter, we provide an introduction into the main tools and techniques that will be applied throughout this course:
  - (Linear) difference equations
  - Linear state space models
  - (Stochastic) dynamic optimization techniques
  - Markov processes

Outline

First-order difference equations

- Consider the **first-order linear difference equation** (LDE)

$$x_t = ax_{t-1} + e_t,\ t = 1, 2, \ldots \tag{1}$$

  where
  - a is a constant coefficient/parameter
  - x_t and e_t are scalar real-valued functions of time $t = 1, 2, \ldots$:
    - $x_t = x(t)$
    - $e_t = e(t)$
- We assume **perfect foresight**, i.e. e_t is known for all t
- In economic applications, we will refer to x_t as the **endogenous variable** and e_t as the **exogenous forcing variable**
- e_t is assumed to be a bounded scalar real-valued sequence (generalization to vector difference equations straightforward)
- A **solution** to (1) is an expression giving x_t as a function only of past, present, and future values of e_t (and **boundary conditions**)

Boundedness

Definition 1 (Boundedness)

A sequence $\{X_t\}$ is bounded/stable if there exists $M > 0$ such that $||X_t||_\infty < M \; \forall \, t$

- Typically, only bounded solutions are economically relevant as otherwise, there is no scarcity
- Some problems, particularly those involving growth, may require appropriate stationarizing, e.g. transforming into intensive form

Characterizing difference equations: two important distinctions

- The stability of the LDE determines how to actually solve the LDE
  - Is the difference equation stable, i.e. $|a| <= 1$?
    $\rightarrow$ solve the equation backward
  - Or do we have an unstable root, i.e. $|a| > 1$?
    $\rightarrow$ solve the equation forward
- The type of available boundary condition determines how we *should* solve the LDE:
  - Do we have an initial condition that restricts x to a particular value today or in the past?
    $\rightarrow$ solve the equation backward
  - Do we have a terminal condition that restricts x to a particular value in the future?
    $\rightarrow$ solve the equation forward
- Sometimes the type of stability condition does not match the available boundary condition: leads to problems with **uniqueness and existence of bounded solutions**

The lead/lag operator

- LDEs can generally be solved by **recursive substitution**, i.e. iteratively plugging in
- This can be made easier by introducing the lag/lead-operator L defined through

$$L^z a x_t = a x_{t-z} \ , \tag{2}$$

  where a is a constant and z is an integer
- We will first consider the case of a stable difference equation, i.e. $|a| < 1$
- *Exercise 1:* Show that for a scalar $|a| < 1$ it holds that

$$\sum_{i=0}^{\infty} a^i L^i = (1 - aL)^{-1} = 1 + aL + a^2 L^2 + \ldots \tag{3}$$

Solving backward

- To find a solution for x_t, rewrite (1) using the lag operator as

$$x_t(1 - aL) = e_t \tag{4}$$

- *Exercise 2:* Show that the **non-homogenous component**, which expresses x_t as a function of current and past e_t

$$x_t = \sum_{s=-\infty}^{t} a^{t-s} e_s \tag{5}$$

  is a solution
- Adding the **homogenous component** ca^t, with c being an arbitrary constant, delivers the general solution:

$$x_t = \sum_{s=-\infty}^{t} a^{t-s} e_s + ca^t = (1 - aL)^{-1} e_t + ca^t \tag{6}$$

- *Exercise 3:* Show that this is an actual solution to the LDE
- Due to discounting with $|a| < 1$, this backward solution is bounded and the effect of past e_t decays geometrically

Ruling out indeterminacy using a boundary condition

- Problem: (6) will hold for any arbitrary c and there is no way to determine it without further information
- This case is called **indeterminacy**
- But: suppose we have an initial condition stating that x_0 is some given (known) constant.
- From the general solution (6) at time 0 we know that

$$x_0 = \sum_{s=-\infty}^{0} a^{-s} e_s + c a^0 \,, \tag{7}$$

- This equation then determines c uniquely as

$$c = x_0 - \sum_{s=-\infty}^{0} a^{-s} e_s \tag{8}$$

The unique and determinate backward solution

- Inserting expression (8) for c into the solution for x_t, (6), we get

$$x_t = \sum_{s=-\infty}^{t} a^{t-s} e_s + c a^t \overset{(8)}{=} \sum_{s=-\infty}^{t} a^{t-s} e_s + a^t \left(x_0 - \sum_{s=-\infty}^{0} a^{-s} e_s\right) \quad (9)$$

- Using that

$$\sum_{s=-\infty}^{t} a^{t-s} e_s - \sum_{s=-\infty}^{0} a^{t-s} e_s = \sum_{s=1}^{t} a^{t-s} e_s \quad (10)$$

yields

$$x_t = \sum_{s=1}^{t} a^{t-s} e_s + a^t x_0 \quad (11)$$

- Due to discounting with $|a| < 1$, the backward solution with starting value x_0 is unique and bounded

Solving forward

- But what if $|a| > 1$, in which case the backward solution (11) will be unbounded?
- Fortunately, we can then try to solve forward
- Consider the difference equation (1) at time $t+1$

$$x_{t+1} = ax_t + e_{t+1} \tag{12}$$

- *Exercise 4:* Show, using the lead operator, that the non-homogenous component of the forward solution is given by

$$x_t = -\sum_{s=t+1}^{\infty} \left(\frac{1}{a}\right)^{s-t} e_s \tag{13}$$

- Again, there exist more solutions that include the homogenous component ca^t:

$$x_t = -\sum_{s=t+1}^{\infty} \left(\frac{1}{a}\right)^{s-t} e_s + ca^t \tag{14}$$

- Since $|a| > 1$, unless $c = 0, ca^t \to \pm\infty$ for $t \to \infty$

Forward solutions and boundedness

- (14) doesn't look like much progress: we still cannot pin down c and the sequence for x_t grows without bounds for $t \to \infty$
- An initial condition as before would make the problem determinate, but still unbounded
- However: in economic applications sensible solutions typically require economic quantities to be bounded
- This amounts to a **terminal condition** $x_t < \infty$ for $t \to \infty$
- This leaves $c = 0$ as the only possible choice
  $\to$ the initial value x_0 of the sequence has to adjust to ensure that $c = 0$ is satisfied
- Sidenote: justification of terminal conditions can be subject of intense debate (e.g. Cochrane 2011)
- Physical capital stock obviously cannot go to infinity, but what about inflation?

Economic example: stable difference equation

- **Law of motion for capital** accumulation with given initial value k_0

$$k_{t+1} = (1-\delta)k_t + i_{t+1} \tag{15}$$

- With $(1-\delta) \in [0,1)$, one can apply the backward solution (11):

$$\begin{aligned} k_t &= \sum_{s=-\infty}^{t} (1-\delta)^{t-s} i_s + c(1-\delta)^t \\ &= \sum_{s=1}^{t} (1-\delta)^{t-s} i_s + (1-\delta)^t k_0 \end{aligned} \tag{16}$$

- Capital stock at each point in time is the sum of
  1. what depreciation has left from the initial capital stock at $t=0$
  2. the cumulative historical additions through (depreciated) investment
- National accounting often relies on this **perpetual inventory method**

Economic example: unstable difference equation

- No-arbitrage asset pricing relationship between
  - bond with constant coupon/interest rate $r > 0$ and
  - stock with purchase price p_t and known dividend d_t
- Interest rate on bond needs to equal dividend yield plus capital gains yield:

$$1 + r = \frac{d_{t+1} + p_{t+1}}{p_t} \tag{17}$$

- Since $(1 + r) > 1$, the resulting difference equation

$$p_{t+1} = (1 + r)p_t - d_{t+1} \tag{18}$$

can be solved forward using (14):

$$p_t = \sum_{s=t+1}^{\infty} \left(\frac{1}{1 + r} \right)^{-(t-s)} d_s + c\,(1 + r)^t \tag{19}$$

- The stock price is equal to the present discounted value of dividends (**fundamental value**) plus a **bubble component** growing at rate r

Economic example: unstable difference equation II

- Again, without further information, we cannot uniquely pin down c and therefore p_t
- Let's impose a boundary condition of non-explosive/bounded stock prices, i.e. $c = 0$:

$$p_t = \sum_{s=t+1}^{\infty} \left(\frac{1}{1+r}\right)^{-(t-s)} d_s \tag{20}$$

- In this case, the (fundamental) stock price is equal to the discounted future stream of dividends
- But: does this boundary condition make sense?

Summary

- For stable LDEs, we solve backward and need an initial condition for uniqueness
- For unstable LDEs, we solve forward and need a terminal condition for a unique and bounded solution
- This suggests two problems that can arise:
    1. if we have an initial condition paired with an unstable root, no bounded solution exists (called **no solution**)
    2. if we only have a terminal condition paired with a stable root, infinitely many bounded solutions exist (called **indeterminacy**)

LDEs: potential complications

In practical applications, there can be two complications

1. Models may involve difference equations of order higher than 1
2. Economic models usually involve both stable and unstable equations $\rightarrow$ we now know how to handle both of them separately, but what if the equations are simultaneously related?

Higher order linear difference equations

- The second-order linear difference equation

$$x_t = a_1 x_{t-1} + a_2 x_{t-2} + e_t \tag{21}$$

  can be brought into **companion form** and thus handled as a system of two first-order equations

- Defining an auxiliary variable $z_t := x_{t-1}$ we can form a two-dimensional first-order system

$$\begin{bmatrix} x_t \\ z_t \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e_t$$

- Generally, an $n-$th order linear difference equation can be analyzed as an n-dimensional system of first-order LDEs
- But how to deal with these systems?

The n-dimensional case

- Consider the n-dimensional first-order LDE vector system

$$x_t = Ax_{t-1} + e_t \tag{22}$$

  with the square **transition matrix** A containing constant coefficients
- We assume A to be **non-singular**, i.e. invertible
- We will learn how to **decouple** the system into n independent equations that can be solved separately
- One way is to **diagonalize the transition matrix** via a **Jordan eigenvalue decomposition**
- If A has real and distinct eigenvalues, it can be diagonalized using the eigenvalue decomposition as described below
- If these assumptions are not satisfied, other methods like the generalized Schur decomposition (e.g. Klein 2000), are available

Eigenvalues and eigenvectors: definition

- The scalar λ is called an **eigenvalue** (or characteristic root) of A and the vector z a corresponding **eigenvector** if they satisfy

$$\lambda z = Az \tag{23}$$

or equivalently

$$(A - \lambda I)\, z = 0 \tag{24}$$

- A non-zero solution requires the matrix $A - \lambda I$ to be singular $\rightarrow$ determinant must be 0
- Hence, for given matrix A the solution(s) for λ can be derived as the roots of the **characteristic equation**

$$|A - \lambda I| = 0 \tag{25}$$

- The resulting characteristic equation has n solutions and A therefore n (potentially complex) eigenvalues: $\lambda = h \pm iv$.

Stable eigenvalues

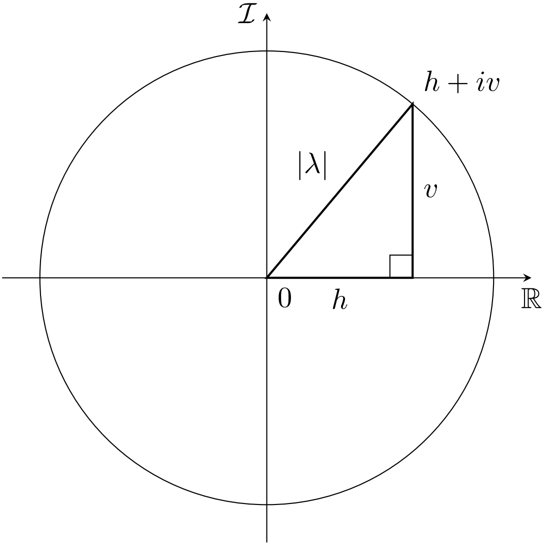
- The **modulus** of an eigenvalue is given by its Euclidean norm $|\lambda| = \sqrt{h^2 + v^2}$

Definition 1

An eigenvalue is stable if its modulus is strictly smaller than one: $|\lambda| < 1$

- For now we will only consider cases with real eigenvalues
  $\rightarrow$ stable eigenvalue is one that is smaller than one in absolute value

Modulus of an eigenvalue



- In the complex plane, a modulus below 1 can be represented as being **inside of the unit circle**
- A thorny issue is the treatment of unit eigenvalues (unit roots)
- In economic problems it often makes sense to lump them together with the stable roots

Jordan Eigenvalue Decomposition

- Define Λ as the diagonal matrix with the eigenvalues of A on its main diagonal:

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} \tag{26}$$

- Let D collect the eigenvectors

$$D = \left[z^1, \ldots, z^n\right] \tag{27}$$

- With distinct real eigenvalues it can be shown that A can be decomposed via the **Jordan Eigenvalue Decomposition** into

$$A = D\Lambda D^{-1} \tag{28}$$

Application to our difference equation

- Premultiplying our system (22) by D^{-1} we get

$$D^{-1}x_t = D^{-1}Ax_{t-1} + D^{-1}e_t \stackrel{(28)}{=} \Lambda D^{-1}x_{t-1} + D^{-1}e_t \tag{29}$$

Defining

$$\zeta_t = D^{-1}x_t \tag{30}$$

$$v_t = D^{-1}e_t \tag{31}$$

we can write this an uncoupled system

$$\zeta_t = \Lambda\zeta_{t-1} + v_t \tag{32}$$

Solving the decoupled system

- The ζ_{it} can be solved separately with the single equation method used above (with the initial/terminal conditions for x transformed into ones for ζ)
- The eigenvalues of A appearing on the diagonal of Λ form the coefficients in the transformed equations
    - Equations associated with an unstable eigenvalue are solved forward
    - Equations associated with a stable eigenvalue are solved backward
- Once the solution ζ_t has been found, the solution for the original variables can be computed from

$$x_t = D\zeta_t \tag{33}$$

The importance of backward-looking stochastic LDEs

- Most of the time we will be concerned with stochastic problems, giving rise to stochastic linear difference equations
- In this case, the perfect foresight assumption on e_t is replaced by e_t being random
- They can be handled very similarly as deterministic ones (as we will see later)
- Typically, solutions to economic problems take the form of backward-looking SLDEs in the **predetermined variables**, i.e. the ones that are restricted by initial conditions
- Because these predetermined variables cannot freely adjust every period, they are also called **state variables**
- In this section, we will characterize stable and unique solutions in the state variables

Stochastic LDEs: Setup

- Consider a process $\{x_t\}$ of n-dimensional vectors, recursively generated by the following first-order linear difference equation:

$$x_{t+1} = A_o x_t + Ce_{t+1}, \text{ for } t = 0, 1, ... \tag{34}$$

  where
  - $\{e_t\}$ for $t = 0, 1, \ldots$ is a sequence of $m \times 1$-dimensional random vectors
  - x_0 is a given initial condition drawn from a distribution with mean μ_0 and covariance $\Sigma_0 = E\left[(x_0 - \mu_0)(x_0 - \mu_0)'\right]$
  - x_t is an $n \times 1$ **state vector** containing variables that are observed by agents at time t
  - A_o is an $n \times n$ transition matrix,
  - C is an $n \times m$ matrix related to the covariance
- The generated $\{x_t\}$ is then also a random sequence (**stochastic process**)
- Note the intimate connection of (34) to VARs from time series econometrics (where $n = m$)

Stochastic shock structure

- The $\{e_t\}$ process generates a sequence of information sets $\{J_t\}$ with $J_t = \{e_t \ldots, e_1, x_0\}$.
- The sequence $\{e_t\}$ is assumed to satisfy

$$E[e_{t+1}|J_t] = 0, \tag{35}$$

$$E[e_{t+1}e'_{t+1}|J_t] = I_m, \tag{36}$$

  where $E_t \equiv E[\cdot|J_t]$ denotes the conditional expectations based on information set J_t
- Hence, $\{e_t\}$ is a sequence of serially uncorrelated random vectors with an unconditional mean equal to zero (shown below)
- *Exercise 5:* Show that x_t is contained in the information set J_t

Expectations operator: mathematical rules

- $E_t\left[X_{t+1}+Y_{t+1}\right]=E_t\left[X_{t+1}\right]+E_t\left[Y_{t+1}\right].$
- In general: $E_t\left[X_{t+1}\cdot Y_{t+1}\right]\neq E_t\left[X_{t+1}\right]\cdot E_t\left[Y_{t+1}\right]$
- **Law of iterated expectations:**

$$E_t\left[E_{t+1}\left[X_{t+2}\right]\right]=E_t\left[X_{t+2}\right]$$

- Constants and variables contained in the information set:
  - $E_t\left[X_t\right]=X_t.$
  - $E_t\left[X_t\cdot Y_{t+1}\right]=X_t\cdot E_t\left[Y_{t+1}\right].$
  - $E_t\left[b\cdot X_{t+1}\right]=b\cdot E_t\left[X_{t+1}\right].$
- Taking derivatives
  - $\frac{\partial}{\partial Y_{t+1}}E_t\left[f(Y_{t+1})\right]=E_t\left[\frac{\partial}{\partial Y_{t+1}}f(Y_{t+1})\right].$
- *Exercise 6:* Show that

$$E[e_t]=0 \tag{37}$$

Conditional mean and covariance

- *Exercise 7:* Show that, if x_t follows the SLDE in (34), the optimal forecast of x_{t+1} (in the mean squared sense) for a given information set at time t is

$$E[x_{t+1}|J_t] = A_o x_t, \tag{38}$$

- The one-step ahead forecast error is

$$x_{t+1} - E_t x_{t+1} = C e_{t+1} \tag{39}$$

- The **conditional covariance matrix of** x_{t+1} (sometimes called **one-step ahead mean squared prediction error**), can be computed by postmultiplying the prediction error by its transpose and taking conditional expectations:

$$\begin{aligned} & E_t\left[\left(x_{t+1} - E_t x_{t+1}\right)\left(x_{t+1} - E_t x_{t+1}\right)'\right] = E_t\left[C e_{t+1}\left(C e_{t+1}\right)'\right] \\ & = E_t C e_{t+1} e'_{t+1} C' = CC'. \end{aligned} \tag{40}$$

Moving average representation

- The **autoregressive (AR)** first-order linear difference equation (34) can alternatively be written as a **moving average (MA) process**
- *Exercise 8:* Show that

$$x_{t+1} = A_o x_t + Ce_{t+1} = \sum_{i=1}^{t+1} A_o^{t+1-i} Ce_i + A_o^{t+1} x_0 \tag{41}$$

- Leading this by $j-1$-periods

$$x_{t+j} = \sum_{s=t+1}^{t+j} A_o^{t+j-s} Ce_s + A_o^j x_t \tag{42}$$

  and applying conditional expectations yields the j-step ahead prediction

$$E_t x_{t+j} = A_o^j x_t. \tag{43}$$

- The associated j-step ahead prediction error is then given by

$$x_{t+j} - E_t x_{t+j} = \sum_{s=t+1}^{t+j} A_o^{t+j-s} Ce_s = \sum_{i=0}^{j-1} A_o^i Ce_{t+j-i} \tag{44}$$

j-step ahead covariance matrix

- The covariance matrix of the j-step ahead forecast error follows from

$$\begin{aligned}\Xi_t(j) =& E_t\left[\left(x_{t+j}-E_t x_{t+j}\right)\left(x_{t+j}-E_t x_{t+j}\right)'\right]\\ \overset{(44)}{=}& E_t\left[\sum_{i=0}^{j-1} A_o^i C e_{t+j-i}\left(\sum_{i=0}^{j-1} A_o^i C e_{t+j-i}\right)'\right]\\ =& E_t\left[\sum_{i=0}^{j-1} A_o^i C e_{t+j-i}\sum_{i=0}^{j-1} e_{t+j-i}' C'\left(A_o^i\right)'\right]\\ =& \sum_{i=0}^{j-1} A_o^i C C'\left(A_o^i\right)',\end{aligned} \tag{45}$$

  where we used that $E_t(e_{t+i}e_{t+i}') = I$ and $E_t(e_{t+i}e_{t+k}') = 0$ for $i \neq k$ and $i, k > 0$.
- Thus, uncertainty increases in the forecast horizon j

Generic formulation with a constant

- We will consider the concept of **covariance stationarity** using a system where we have separated out constant terms:

$$A_o = \begin{bmatrix} A & \widetilde{A} \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} C_1 \\ 0 \end{bmatrix}, \tag{46}$$

  where
  - A is an $(n-1) \times (n-1)$ matrix with only stable eigenvalues
  - $\widetilde{A}$ is an $(n-1) \times 1$ column vector
- Correspondingly partitioning $x_t' = [x_{1t}' \quad x_{2t}']$, equation (34) can be written as

$$x_{1t+1} = Ax_{1t} + \widetilde{A}x_{2t} + C_1 e_{t+1}, \tag{47}$$

$$x_{2t+1} = x_{2t} \tag{48}$$

- Given that x_{2t} is constant over time (it exhibits a unit root), and thus equal to x_{20}, we define

$$B \equiv \widetilde{A}x_{20} \tag{49}$$

Covariance stationarity

Definition 2

A stochastic process $\{x_t\}$ *is covariance stationary if the unconditional mean is time invariant*

$$Ex_t = Ex_0 \,\forall\, t \tag{50}$$

and the sequence of unconditional autocovariance matrices

$$\Sigma(j) \equiv E(x_{t+j} - Ex_{t+j})(x_t - Ex_t)' \tag{51}$$

is finite and only depends on the separation between dates j*, but not on* t*.*

- Note that conditional heteroscedasticity is allowed, because it refers to conditional moments
- Our sequence $\{x_t\}$ will be covariance stationary, if all eigenvalues of A (i.e. not associated with the constant) have moduli less than one

The goal: a fixed point

- We will be looking for an initial distribution with mean $\mu_0 = Ex_0$ and covariance $\Sigma_0 = E\left[(x_0 - \mu_0)(x_0 - \mu_0)'\right]$ that, if x_0 is drawn from this distribution, makes x_t covariance stationary
- If instead starting from a arbitrary initial value x_0, the process will still converge to this stationary distribution as $t \to \infty$

The unconditional mean

- Taking **unconditional expectations** on both sides of (47), yields

$$Ex_{1t+1} = AEx_{1t} + B. \tag{52}$$

- If μ_0 is the **stationary value** of Ex_{1t}, it needs to satisfy

$$\mu_0 = A\mu_0 + B \Leftrightarrow \mu_0 = (I - A)^{-1}B \tag{53}$$

- As before, using that the eigenvalues are stable, we know that

$$(I - A)^{-1} = I + A + A^2 + A^3 + \ldots \tag{54}$$

Convergence to the unconditional mean

- From the eigenvalue decomposition $A = D\Lambda D^{-1}$, the j^{th} power of A can be written as

$$A^j = \left(D\Lambda D^{-1}\right)\left(D\Lambda D^{-1}\right)\ldots\left(D\Lambda D^{-1}\right) = D\Lambda^j D^{-1} \tag{55}$$

- Thus, the stationary mean μ_0 satisfies

$$\mu_0 = (I - A)^{-1}B \overset{(54)}{=} \sum_{i=0}^{\infty} A^i B \overset{(55)}{=} D\left(\sum_{i=0}^{\infty}\Lambda^i\right)D^{-1}B \tag{56}$$

- If all eigenvalues of A are stable, μ_0 will have a finite value
- *Exercise 9:* Show that in this case, x_{1t} will converge to the stationary value μ_0 for any initial value of x_{10}

The unconditional covariance

- Subtracting (52) from (47) yields the uncond. period t forecast error

$$(x_{1t} - Ex_{1t}) = A\left(x_{1t-1} - Ex_{1t-1}\right) + C_1 e_t \tag{57}$$

- The uncond. forecast error covariance at t, Σ_t, then satisfies

$$\begin{aligned}
\Sigma_t =& E\left[\left(x_{1t} - Ex_{1t}\right)\left(x_{1t} - Ex_{1t}\right)'\right] \\
\overset{(57)}{=}& E\left[\left(A\left(x_{1t-1} - Ex_{1t-1}\right) + C_1 e_t\right)\left(A\left(x_{1t-1} - Ex_{1t-1}\right) + C_1 e_t\right)'\right] \\
=& E\big[A\left(x_{1t-1} - Ex_{1t-1}\right)\left(x_{1t-1} - Ex_{1t-1}\right)' A' + C_1 e_t e_t' C_1' \\
& \quad + C_1 e_t \left(x_{1t-1} - Ex_{1t-1}\right)' A' + A\left(x_{1t-1} - Ex_{1t-1}\right) e_t' C_1'\big] \\
=& AE\left[\left(x_{1t-1} - Ex_{1t-1}\right)\left(x_{1t-1} - Ex_{1t-1}\right)'\right] A' + E\left[C_1 e_t e_t' C_1'\right] \\
=& A\Sigma_{t-1} A' + C_1 C_1' \ ,
\end{aligned} \tag{58}$$

where the second-to-last line uses that e_t is orthogonal to $x_{1t-1} - Ex_{1t-1}$

The unconditional covariance II

- The result is a **discrete Lyapunov** equation

$$\Sigma_t = A\Sigma_{t-1}A' + C_1C_1' \tag{58}$$

- A stationary value Σ_0 for the covariance matrix of x_{1t} has to satisfy (with $Ex_{1t} = \mu_0$)

$$\Sigma_0 = A\Sigma_0 A' + C_1C_1' \tag{59}$$

- Iterating backwards yields the solution

$$\Sigma_0 = \sum_{k=0}^{\infty} A^k C_1 C_1' A^{k'} \tag{60}$$

- *Exercise 10:* Verify that this solves (59) if A is a stable matrix
- Comparing the j-step ahead forecast error variance, (45), with the unconditional one in (60) reveals that the unconditional covariance is the limit of the former

Inspecting the stationary covariance matrix

- If the eigenvalues of A have moduli less than one, the unconditional covariance matrix Σ_0 takes a finite value:

$$\begin{aligned}\Sigma_0 &\overset{(55)}{=} \sum_{j=0}^{\infty} D\Lambda^j D^{-1} C_1 C_1' \left(D\Lambda^j D^{-1}\right)' \\ &= D \sum_{j=0}^{\infty} \Lambda^j \left[D^{-1}C_1\right]\left[C_1'\left(D^{-1}\right)'\right]\left(\Lambda^j\right)' D' \\ &= D\left(\sum_{j=0}^{\infty} \Lambda^j \widetilde{C}\widetilde{C}'\left(\Lambda^j\right)'\right) D',\end{aligned} \tag{61}$$

  where $\widetilde{C} = D^{-1}C_1$ and $\widetilde{C}' = C_1'(D^{-1})'$

- Thus, if the covariance matrix of x_{10} is Σ_0, it will remain constant over time
- Otherwise, the sequence of covariance matrices will asymptotically converge to Σ_0

The unconditional autocovariances

- This shows that the contemporaneous covariance at lag 0, $\Sigma_0 = \Sigma_0(0)$, only depends on j, but not on t
- But what about other leads and lags?
- The unconditional forecast error at time $t+j$ satisfies

$$\left(x_{1t+j} - Ex_{1t+j}\right) \overset{(44)}{=} A^{j+1}\left(x_{1t-1} - Ex_{1t-1}\right) + C_1 e_{t+j} + \ldots + A^j C_1 e_t \tag{62}$$

- *Exercise 11:* Show that the autocovariance at lead/lag j

$$\Sigma_t\left(j\right) \equiv E\left(x_{1t+j} - Ex_{1t+j}\right)\left(x_{1t} - Ex_{1t}\right)' \tag{63}$$

  is given by

$$\Sigma_t\left(j\right) = A^j \Sigma_t(0) \tag{64}$$

- Thus, if $\Sigma_t(0)$ does not depend on t, neither does $\Sigma_t\left(j\right)$

State space systems

- We have already characterized the **state transition equation**, i.e. the evolution of the predetermined state variables x_t over time

$$x_{t+1} = A_o x_t + C e_{t+1}, \text{ for } t = 0, 1, ... \tag{34}$$

- But often we are also interested in functions of the state variables:

$$y_t = G x_t \ , \tag{65}$$

where y_t is a $n_y \times 1$ vector and G a $n_y \times n_x$ matrix

- Equation (65) linking endogenous non-state variables to state variables is called a **measurement equation/observation equation**
- Equations (34) and (65) together form a **state-space system**, describing the evolution of all variables of interest

Moving average representations (again)

- We already derived the moving average representation for the state variables

$$x_{t+1} = \sum_{i=1}^{t+1} A_o^{t+1-i} C e_i + A_o^{t+1} x_0 \tag{41}$$

- From (65) then follows

$$y_t = G \sum_{i=0}^{t} A_o^{t-i} C e_i + G A_o^t x_0 \tag{66}$$

- *Exercise 12:* Show that

$$E[y_t | x_0] = G A_o^t x_0 \tag{67}$$

and using (45) that

$$E_t \left[\left(y_{t+j} - E_t y_{t+j} \right) \left(y_{t+j} - E_t y_{t+j} \right)' \right] = G \left[\sum_{i=0}^{j-1} A_o^i C C' \left(A_o^i \right)' \right] G' \tag{68}$$

Impulse response functions

- We will often be concerned with the dynamic response of a variable y_{t+h} to a shock ε_t occurring at time t:

$$\frac{\partial y_{t+h}}{\partial \varepsilon_t} \text{ for some } t \text{ and some } h \text{ (typically } \geq 0) \tag{69}$$

- The moving average representations (41) and (66) reveal that the **impulse-response functions** for $h \geq 0$ are given by

$$\frac{\partial y_{t+h}}{\partial \varepsilon_t} = G A_o^h C \tag{70}$$

$$\frac{\partial x_{t+h}}{\partial \varepsilon_t} = A_o^h C \tag{71}$$

- If all eigenvalues are inside of the unit circle, the effect of shocks will die out over time

Outline

Introduction

- We have made great headway in solving and understanding LDEs, but we still do not know where the LDEs come from in economic applications
- They follow from economic optimization problems in discrete time
- We again start with the deterministic case, before turning to the stochastic case

Setup: the sequence problem

1. Choose an infinite sequence of **control variables**, $\{u_t\}_{t=0}^{\infty}$
2. to maximize a discounted time-separable objective

$$\max_{\{u\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t r(x_t, u_t), \tag{72}$$

   where $\beta \in (0,1)$ is a discount factor and r the instantaneous return function,
3. subject to a set of equations constraining the evolution of the **state variables** $\{x_t\}_{t=0}^{\infty}$

$$x_{t+1} \leq g_t(x_t, u_t) \tag{73}$$

4. for a given initial value x_0

Classification of variables

- When discussing LDEs, we already saw the important distinction between predetermined variables and non-predetermined ones
- In economic problems, we classify variables into endogenous and exogenous state variables and control variables
- **State variables**: cannot be changed by the agent in the current period, because they are **predetermined**. They characterize the state of the system at any point in time.
  - **exogenous** state: cannot be changed by the agent at any time, but follow exogenous (stochastic) law of motion
  - **endogenous** state: are predetermined at time t, but can be changed by the agent at time $t + j, j > 0$
- **Control variables**: all variables that the agent can change at time t (can be an endogenous state variable at $t + j, j > 0$)
- We will discuss examples later on

The goal

- Derive a (time-invariant) **policy function** h that maps the state x_t into the control u_t such that the sequence of controls generated by the two functions

$$u_t = h(x_t) \tag{74}$$
$$x_{t+1} \leq g(x_t, u_t)\,, \tag{73}$$

  starting from the initial condition x_0 at $t = 0$, solves the problem
- Knowing $h : D \to \mathbb{R}$ is equivalent to knowing $\{u_t\}_{t=0}^{\infty}$
- Note that both $\{u_t\}_{t=0}^{\infty}$ and h are infinite-dimensional objects
- The structure of the system (73)-(74) is called **recursive**: starting with x_0, we can generate the sequence for $\{u_t, x_{t+1}\}$ by iteratively plugging in
- We will only encounter functions, but h could be a **correspondence**

Getting the solution

- There are two different solution techniques:
  1. Dynamic programming (covered in Advanced Macro II)
  2. Lagrange-based methods
- Lagrange is typically more intuitive, but requires differentiability
- Dynamic programming works with functions instead of sequences, requiring bigger formal investments, but
  - it has computational advantages, particularly when the problem is not differentiable
  - solution characteristics like existence, uniqueness, boundedness, etc. are often easier to derive

The Lagrangian formulation

- If everything is nicely differentiable, discrete time optimization problems can also be solved by applying the **Lagrange formulation** instead of dynamic programming via Bellman equations
- The Lagrangian sequence problem formulation of the original problem is

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left[r(x_t, u_t) + \lambda_t' \left(g(x_t, u_t) - x_{t+1} \right) \right], \tag{75}$$

  where x is an n-dimensional vector of states, λ_t is an n-dimensional vector of Lagrange multipliers, and x_0 is given
- By convention: discounting applies to the Lagrange multiplier as well
- We seek to maximize the Lagrange function by choosing sequences $\{u_t, x_{t+1}\}_{t=0}^{\infty}$ (and λ_t)
- Note that the Lagrange multiplier part in (75) is equivalent to $-\lambda_t' \left(x_{t+1} - g(x_t, u_t) \right)$

The first order conditions

- The first-order necessary conditions at every point in time t are given by the transversality constraint (TVC) and

$$
\begin{aligned}
0 =& \beta^t \frac{\partial r\left(x_t, u_t\right)}{\partial u_t} + \beta^t \lambda_t \frac{\partial g(x_t, u_t)}{\partial u_t} && (76)\\
0 =& -\beta^t \lambda_t + \beta^{t+1} \frac{\partial r\left(x_{t+1}, u_{t+1}\right)}{\partial x_{t+1}} + \beta^{t+1} \lambda_{t+1} \frac{\partial g(x_{t+1}, u_{t+1})}{\partial x_{t+1}} && (77)\\
0 =& g(x_t, u_t) - x_{t+1}, && (78)
\end{aligned}
$$

  where we used the vector calculus rule that $\frac{\partial \lambda' a}{\partial a} = \lambda$ (see e.g. Greene 2011, p.1009)
- Due to the stationary problem, the FOCs are time-invariant
- Simplifying gives

$$
\frac{\partial r\left(x_t, u_t\right)}{\partial u_t} + \lambda_t \frac{\partial g(x_t, u_t)}{\partial u_t} = 0 \qquad (79)
$$

$$
\beta \frac{\partial r\left(x_{t+1}, u_{t+1}\right)}{\partial x_{t+1}} + \beta \lambda_{t+1} \frac{\partial g(x_{t+1}, u_{t+1})}{\partial x_{t+1}} = \lambda_t \qquad (80)
$$

The transversality condition

- Starting from a finite horizon problem, if we let $T \to \infty$, the terminal condition for capital leads to the **transversality condition**

$$\lim_{T\to\infty} \beta^T \lambda_{1,T} k_{T+1} = 0 \tag{81}$$

- The transversality condition (81) is a necessary condition for the maximization problem
- We will repeatedly face infinite horizon problems leading to this type of condition
- This TVC essentially implies that the discounted value of wealth (i.e., capital) in terms of utility has to be zero for the limiting case where time approaches infinity
- TVCs serve as terminal conditions that have to be taken into account when solving the set of first-order conditions (which will typically take the form of vector difference equations)

The neoclassical growth model

- Consider a Ramsey (1928)-Cass (1965)-Koopmans (1965)-type model with one representative household, supplying labor inelastically: $N = 1$
- Felicity is logarithmic in consumption:

$$u(C_t) = \ln C_t \tag{82}$$

and households have discount factor $\beta \in [0, 1)$

- Production technology is Cobb-Douglas with capital share $\alpha \in (0, 1)$:

$$Y_t = F(K_t, N_t) = A K_t^{\alpha} N_t^{1-\alpha} \, , \tag{83}$$

with TFP $A > 0$ being exogenous

- The law of motion for capital is given by

$$K_{t+1} = (1 - \delta) K_t + I_t \, , \tag{84}$$

where $\delta \in (0, 1]$ is the depreciation rate

- We assume a closed economy, i.e. the resource constraint is

$$Y_t = C_t + I_t \tag{85}$$

Stock notations

- The capital stock is a predetermined variable
- There are two timing conventions in the literature:
  1. the **stock at the end of period notation** uses the capital stock at the end of the period when investment has already been added:

$$K_t = (1-\delta)K_{t-1} + I_t \tag{86}$$

$$Y_t = AK_{t-1}^{\alpha} N_t^{1-\alpha} \tag{87}$$

  2. the **stock at the beginning of period notation** uses the capital stock at the beginning of the current period, before current investment has been added:

$$K_{t+1} = (1-\delta)K_t + I_t \tag{88}$$

$$Y_t = AK_t^{\alpha} N_t^{1-\alpha} \tag{89}$$

- The **stock at the end of period notation** is typically more intuitive, because it endows the capital stock with the timing at which it is decided and makes the predeterminedness explicit

The planner's problem

- Define lower-case letters as per capita variables, i.e. $x_t = X_t / N_t$
- *Exercise 13:* Assuming full deprecation $\delta = 1$, show that

$$k_{t+1} = Ak_t^{\alpha} - c_t \tag{90}$$

- The social planner's program then is:

$$\begin{aligned} &\max_{\{c_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \ln c_t \\ &s.t.\ k_{t+1} = Ak_t^{\alpha} - c_t, \qquad (90) \\ &k_0 > 0 \text{ and TVC given} \end{aligned}$$

- Variable classification
  - State variables: k_t
  - Control variables: c_t, k_{t+1}

The Brock and Mirman (1972)-model

- For the simple neoclassical growth model from above, the Lagrange sequence problem reads

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t \left[\ln c_t + \lambda_t \left(A k_t^{\alpha} - c_t - k_{t+1} \right) \right]. \tag{91}$$

- The first-order conditions (79)-(80) are given by

$$\lambda_t = c_t^{-1}, \tag{92}$$

$$\lambda_t = \beta \lambda_{t+1} \alpha A k_{t+1}^{\alpha-1}, \tag{93}$$

where we used that

$$\frac{\partial r\left(x_{t+1}, u_{t+1}\right)}{\partial x_{t+1}} = \frac{\partial \ln c_{t+1}}{\partial k_{t+1}} = 0 \tag{94}$$

- Eliminating the Lagrange multiplier yields the Euler equation:

$$\frac{c_{t+1}}{c_t} = \beta \alpha A k_{t+1}^{\alpha-1} \tag{95}$$

Reminder: The Keynes-Ramsey rule I

- The Euler, together with the transition law (90), implicitly characterizes the optimal sequence of c_t and k_t for a given initial value k_0 (second order difference equation)
- The Euler equation for consumption per worker yields the **Keynes-Ramsey rule**

$$\frac{\Delta c_{t+1}}{c_t} \approx \log\left(\frac{c_{t+1}}{c_t}\right) = \frac{r_{t+1,t} - \rho}{\sigma} = (r_{t+1,t} - \rho) \times EOIS \tag{96}$$

where we used:

- the approximation $\log(1+x) \approx x$ for small x
- $\beta = 1/(1+\rho)$, where ρ is the rate of time preference
- CRRA preferences of the type $(C_t^{1-\sigma} - 1)/(1-\sigma)$ with risk aversion coefficient σ so that the **intertemporal elasticity of substitution** is $EOIS = 1/\sigma$
- $r_{t+1,t} = f'(k_{t+1}) + (1-\delta) = \alpha A k_{t+1}^{\alpha-1}$

Reminder: The Keynes-Ramsey rule II

- The Euler equation for consumption per worker yields the **Keynes-Ramsey rule**

$$\frac{\Delta c_{t+1}}{c_t} \approx \log\left(\frac{c_{t+1}}{c_t}\right) = \frac{r_{t+1,t} - \rho}{\sigma} = (r_{t+1,t} - \rho) \times EOIS \qquad (96)$$

- If $r > \rho$: Giving up one unit of consumption today and saving it yields r, but the HH only discounts with $\rho \Rightarrow$ HH shifts consumption to the future, i.e. consumption path is increasing
- Marginal utility in the future must be lower and consumption higher than today to make RHS of (95) equal to LHS, given $r_{t+1,t} - \rho > 0$
- In EQ: shifting consumption in any direction at the margin does not increase discounted marginal utility
- The willingness to shift consumption intertemporally depends on the EOIS $1/\sigma$: the higher the EOIS the more consumption responds to a given $r - \rho$

The method of undetermined coefficients

- We are left with a problem: the Euler equation only implicitly characterizes the solution
- In general, it is impossible to immediately obtain the policy functions from these equilibrium conditions
- However: the Brock and Mirman (1972)-model due to its log-linear structure has a simple analytical solution that can by guessed and verified via the **method of undetermined coefficients**
- Idea: guess a functional form for h but leave the coefficients unknown; functional form should only depend on the state of the system and be similar to the model structure
- Plug this guess (repeatedly) into the FOCs and try to solve for the unknown coefficient(s)
- If we can find a value for the coefficients (i.e. a function of only parameters, not variables) that satisfies the FOCs, we have found one (of potentially many) solution

Getting the solution

- *Exercise 14:* Use the guess of a constant savings rate

$$k_{t+1} = \phi A k_t^{\alpha} \tag{97}$$

and show that $\phi = \alpha\beta$ solves the problem

Intro

- This section presents methods of dynamic optimization under **uncertainty**, which will repeatedly be applied in this course

Stochastic structure

- We introduce random events $s_t \in S$ where S is the finite set of possible events $S = \{\sigma_a, \sigma_b, \sigma_c, ..., \sigma_n\}$ that can happen in any period t
- The unconditional probability of an event is given by $\pi(s_t)$ and

$$\sum_{s_t \in S} \pi(s_t) = 1 \tag{98}$$

- The **event history** is given by $s^t = (s_t, s_{t-1}, ..., s_0)$ and summarizes the realizations of all events from period 0 to period t.
- The set of event histories is then $S^t = S \times S \times ... \times S$, with $s^t \in S^t$
- The probability of a particular event history s^t is given by $\pi(s^t)$, where $\pi(s^t) > 0$ for all $s^t \in S^t$ and t
- We could relax the assumption that the same set of events S can realize with some strictly positive probability in each period, but
  - notation would be heavier
  - problem would not be time-invariant

Probability of event history

- The probability of a particular event history $\pi(s^t) = \pi(s_t, s_{t-1}, ...s_0)$ is the joint probability of the events $s_t, \ldots, s_0$, where $\pi(s^0) = 1$ for a given initial state s^0
- Denote the conditional probability of state s_t and thus s^t given the history s^{t-1} at date $t-1$ with $\pi(s^t|s^{t-1})$
- The joint probability can the be factored as

$$\begin{aligned} \pi(s^t) &= \pi(s^t|s^{t-1}) \cdot \pi(s^{t-1}|s^{t-2}) \cdot \ldots \cdot \pi(s^1|s^0) \cdot \pi(s^0) \\ &= \pi(s_t|s^{t-1}) \cdot \pi(s_{t-1}|s^{t-2}) \cdot \ldots \cdot \pi(s_1|s^0) \, , \end{aligned} \tag{99}$$

where we used that $\pi(s^0) = 1$.

Markov processes

- In general the probability of a particular event s_t occuring at time t depends on the complete history of past events
- That makes general processes hard to work with, because past events become states that need to be tracked
- Life becomes easier when working with Markov processes that only require keeping track of some history

Definition 3 (n^{th}-order Markov process)

*s_t is generated by an $n^{\textbf{th}}$**-order Markov process** if the distribution of s_t conditional on n lags is the same as the distribution of s_t conditional on all lags:*

$$\pi\left(s_t | s^{t-1}\right) = \pi\left(s_t | s_{t-1}, ..., s_{t-n}\right) \tag{100}$$

First-order Markov processes

- A first-order Markov process is then characterized by

$$\pi\left(s_t|s_{t-1}\right) = \pi\left(s_t|s^{t-1}\right) \tag{101}$$

- The probability of a particular event only depends on the last event, but not on the entire history of earlier events
- The probability of event history s^t,

$$\pi(s^t) = \pi(s_t, s_{t-1}, ..., s_0) \tag{102}$$

for a given initial state s_0 then simplifies to

$$\begin{aligned} \pi(s^t) &= \pi(s_t|s^{t-1}) \cdot \pi(s_{t-1}|s^{t-2}) \cdot \ldots \cdot \pi(s_1|s^0) \\ &= \pi\left(s_t|s_{t-1}\right) \cdot \pi\left(s_{t-1}|s_{t-2}\right) \cdot \ldots \cdot \pi\left(s_1|s_0\right) \end{aligned} \tag{103}$$

Example

- Consider a 2-state Markov process: each period the event is $s_t \in S = \{\sigma_a, \sigma_b\}$ with $s_0 = \sigma$ being a given fixed event $\in S$
- The history of events up to t is then

$$s^t = (s_t, s_{t-1}, \ldots, s_0) \in S^t = S_t \times S_{t-1} \times \ldots \times S_0 \tag{104}$$

- Thus, for the periods $t = 0, 1, 2$ the sets of event histories are

$$S_0 = \sigma \tag{105}$$

$$S^1 = \{(\sigma_a, \sigma), (\sigma_b, \sigma)\} \tag{106}$$

$$S^2 = \{(\sigma_a, \sigma_a, \sigma), (\sigma_a, \sigma_b, \sigma), (\sigma_b, \sigma_a \, \sigma), (\sigma_b, \sigma_b, \sigma)\} \tag{107}$$

- The unconditional probability for event history $\pi(s^2)$ is given by

$$\begin{aligned} \pi(s^2) &= \pi(s_2|s^1) \cdot \pi(s_1|s^0) \cdot \pi(s^0) \\ &= \pi(s_2|s_1) \cdot \pi(s_1|s_0) \cdot \pi(s_0) \end{aligned} \tag{108}$$

Example continued

- The unconditional probability for the particular history $s^2 = (\sigma_b, \sigma_a\,, \sigma)$ is given by

$$\pi((\sigma_b, \sigma_a\,, \sigma)) = \pi(\sigma_b|\sigma_a) \cdot \pi(\sigma_a|\sigma)\;, \tag{109}$$

where we used that $\pi(s^0) = \pi(\sigma) = 1$.

A stochastic concave programming problem

- Consider a problem with a time-separable differentiable and concave return function $r(x_t, y_t)$ and discount factor $\beta \in (0, 1)$
- The sequence of control variables y_t is chosen to maximize an intertemporal objective, subject to a set of constraints on the evolution of the endogenous state variables x_t

$$x_{t+1} = g(x_t, y_t) \tag{110}$$

- The set $\{(x_{t+1}, x_t) : x_{t+1} \leq g(x_t, y_t)\}$ is convex and compact for admissible y_t
- We assume events s_t realize at the beginning of each period, i.e. after y_{t-1} has been chosen, but before y_t
- Controls y_t will generally depend on event history s^t, i.e. $y_t(s^t)$, while state variables x_t will depend on history s^{t-1}, i.e. $x_t(s^{t-1})$
- Note that they might depend on more than the random events

Stochastic maximization problem

- Stochastic problems typically involve the maximization of an **expected** sum of discounted returns (e.g. **von Neumann-Morgenstern utility**), subject to set of constraints that has to hold for each state of the world:

$$\max_{\{y_t(s^t), x_{t+1}(s^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{s^t \in S^t} \beta^t \pi(s^t) r\left(x_t(s^{t-1}), y_t(s^t)\right), \tag{111}$$

$$\text{s.t.} \quad x_{t+1}(s^t) = g\left(x_t(s^{t-1}), y_t(s^t), s_t\right),$$

where initial values x_0 and s_0 are given and random events s are generated by a first-order Markov process

- The Lagrangian for this problem is given by

$$\begin{aligned} \mathcal{L}(x_0, s_0) = \sum_{t=0}^{\infty} \sum_{s^t \in S^t} \beta^t \pi(s^t) \Big[& r\left((x_t(s^{t-1}), y_t(s^t)\right) \\ & + \lambda_t'(s^t)\left(g\left(x_t(s^{t-1}), y_t(s^t), s_t\right) - x_{t+1}\left(s^t\right)\right)\Big] \end{aligned} \tag{112}$$

The first order conditions

- The first-order conditions with respect to $y_t(s^t)\ \forall s^t, s^{t+1} \in S^t$ are

$$0 = \beta^t \pi(s^t) \frac{\partial r\left(x_t(s^{t-1}), y_t(s^t)\right)}{\partial y_t(s^t)} + \beta^t \pi(s^t)\lambda_t(s^t) \frac{\partial g\left(x_t(s^{t-1}), y_t(s^t), s_t\right)}{\partial y_t(s^t)}, \tag{113}$$

- The FOCs with respect to $x_{t+1}(s^t)$ are

$$\pi(s^t)\lambda_t(s^t) = \sum_{s^{t+1}|s^t} \beta\pi(s^{t+1}) \frac{\partial r\left((x_{t+1}(s^t), y_{t+1}(s^{t+1})\right)}{\partial x_{t+1}(s^t)} + \sum_{s^{t+1}|s^t} \beta\pi(s^{t+1})\lambda_{t+1}(s^{t+1}) \frac{\partial g\left(x_{t+1}(s^t), y_{t+1}(s^{t+1}), s_{t+1}\right)}{\partial x_{t+1}(s^t)}, \tag{114}$$

  where we sum over all possible histories s^{t+1}, conditional on the realization of a particular history s^t

- In addition, the FOC w.r.t. to λ, i.e. the constraint, has to hold

History-dependence vs. time-invariance

- Solutions for y_t and x_{t+1} that satisfy the two sets of conditions, (113)-(114) and the set of constraints

$$x_{t+1}(s^t) = g\left(x_t(s^{t-1}), y_t(s^t), s_t\right) \tag{115}$$

  are generally **history-dependent**
- That is, objects are in principle time-varying functions of the history of events s^t
- However, we aim at deriving time-invariant functions for y_t and x_{t+1} that do not depend on the entire history, but on a limited set of relevant information.
- At each point in time this information is summarized in the **state**, which includes the random event s_t, i.e. the **exogenous state** and the **endogenous state** x_t

Recursive solutions

- Suppose there exists a solution for the current control y_t that can be written as a time-invariant function of the current endogenous state x_t and the current event s_t rather than of the entire history s^t.

$$y_t = h(x_t, s_t) \tag{116}$$

- In this case, the constraint requires the endogenous state to satisfy

$$x_{t+1} = g\left(x_t, y_t, s_t\right) = g\left(x_t, h(x_t, s_t), s_t\right) \tag{117}$$

- The endogenous state, like the exogenous state, is then generated by a first-order Markov process
  $\rightarrow$ the solutions for y_t and x_{t+1} inherit the Markov property

Rewriting the FOCs

- We can use the Markov property to rewrite the first order conditions (113) as

$$0 = \frac{\partial r\left(x_t, y_t(x_t, s_t)\right)}{\partial y_t(x_t, s_t)} + \lambda_t(x_t, s_t)\frac{\partial g\left(x_t, y_t(x_t, s_t), s_t\right)}{\partial y_t(x_t, s_t)} \tag{113'}$$

  $\rightarrow \lambda_t$ is also a function of x_t and s_t.

- Using the Markov property and $\frac{\pi(s^{t+1})}{\pi(s^t)} = \pi\left(s_{t+1}|s_t\right)$, we can write (114) as

$$\lambda_t(s_t) = \sum_{s_{t+1}|s_t} \beta\pi\left(s_{t+1}|s_t\right)\frac{\partial r\left(x_{t+1}, y_{t+1}(x_{t+1}, s_{t+1})\right)}{\partial x_{t+1}} + \sum_{s_{t+1}|s_t} \beta\pi(s_{t+1}|s_t)\lambda_{t+1}(x_{t+1}, s_{t+1})\frac{\partial g\left(x_{t+1}, y_{t+1}(x_{t+1}, s_{t+1}), s_{t+1}\right)}{\partial x_{t+1}} \tag{114'}$$

Rewriting the FOCs II

- Introducing the expectations operator E_t conditional on information in period t (for any vector z of random variables)

$$E_t z_{t+1} = \sum_{s_{t+1}|s_t} \pi(s_{t+1}|s_t) z_{t+1}(s_{t+1}), \tag{118}$$

allows to rewrite the conditions as

$$0 = \frac{\partial r\left(x_t, y_t\right)}{\partial y_t} + \lambda_t \frac{\partial g\left(x_t, y_t, s_t\right)}{\partial y_t} \tag{113''}$$

$$\lambda_t = \beta E_t \frac{\partial r\left(x_{t+1}, y_{t+1}\right)}{\partial x_{t+1}} + \beta E_t \lambda_{t+1} \frac{\partial g\left(x_{t+1}, y_{t+1}, s_{t+1}\right)}{\partial x_{t+1}} \tag{114''}$$

$$x_{t+1} = g\left(x_t, y_t, s_t\right) \tag{117}$$

Inspecting the solution

- These conditions have to be satisfied by sequences $\{y_t\}_{t=0}^{\infty}$ and $\{x_{t+1}\}_{t=0}^{\infty}$, for a given sequence for s_t and x_0
- If such solutions exist, y_t and x_{t+1} are time-invariant functions of the current states x_t and s_t, i.e. identical functions $\forall\, t \geq 0$
- Note that y_t and x_{t+1} might further depend on the properties of the **distribution** of the random event
- If the system is linear, only the first moment of s_t will affect the solution (**certainty equivalence**)
- Otherwise, higher moments might also affect the solution

A stochastic Brock and Mirman (1972) model

- Consider the planner problem for a Brock and Mirman (1972) model with a random exogenous state variable s_t generated by a first order finite state Markov process

$$\max_{\{c_t(s^t), k_{t+1}(s^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{s^t \in S^t} \beta^t \pi(s^t) u(c_t(s^t)) \tag{119}$$
$$\text{s.t.} \quad k_{t+1}(s^t) + c_t(s^t) = f(k_t(s^{t-1}), s_t),$$

where u denotes a standard concave utility function, c consumption of a representative household, and k capital per capita

- The production function f exhibits neoclassical properties and output (per capita) depends on a random event via stochastic total factor productivity

The Lagrangian sequence problem

- The Lagrangian sequence problem is then given by

$$\mathcal{L} = \sum_{t=0}^{\infty} \sum_{s^t \in S^t} \beta^t \pi(s^t) u(c_t(s^t)) + \sum_{t=0}^{\infty} \sum_{s^t \in S^t} \beta^t \pi(s^t) \lambda_t(s^t) \left(f(k_t(s^{t-1}), s_t) - k_{t+1}(s^t) - c_t(s^t) \right) \tag{120}$$

- The FOCs for c_t and k_{t+1} for given histories s^t and s^{t+1} are

$$\pi(s^t) u_c(c_t(s^t)) - \pi(s^t)\lambda_t(s^t) = 0 \tag{121}$$

$$\pi(s^t)\lambda_t(s^t) - \sum_{s^{t+1}|s^t} \beta \pi(s^{t+1}) \lambda_{t+1}(s^{t+1}) f'(k_{t+1}(s^t), s_{t+1}) = 0 \tag{122}$$

Deriving the Euler equation

- The FOCs (121)-(122) can be simplified to

$$u_c(c_t(s^t)) = \lambda_t(s^t) \tag{123}$$

$$\lambda_t(s^t) = \sum_{s^{t+1}|s^t} \beta \frac{\pi(s^{t+1})}{\pi(s^t)} \lambda_{t+1}(s^{t+1}) f'(k_{t+1}(s^t), s_{t+1}) \tag{124}$$

- Now use that s_t follows a Markov process $\frac{\pi(s^{t+1})}{\pi(s^t)} = \pi\left(s_{t+1}|s_t\right)$ and suppose there exists a solution that can be written as functions of the current states, but does not depend on the previous history of events
- Then we end up with the Euler equation

$$u_c(c_t(s_t)) = \sum_{s_{t+1}|s_t} \beta\pi\left(s_{t+1}|s_t\right) u_c(c_{t+1}(s_{t+1})) f'(g(k_t, s_t), s_{t+1}) \tag{125}$$

or more compactly

$$u_c(c_t(s_t)) = E_t \beta \left[u_c(c_{t+1}(s_{t+1})) f'(g(k_t, s_t), s_{t+1})\right] \tag{126}$$

The equilibrium

- The Euler equation implies that $c_t(s_t)$ depends on current states (s_t, k_t)
- Consumption will also in general depend on the first and higher moments of the random variable s_t
- Given that consumption does not depend on the history of states, we can write the constraint as

$$k_{t+1}(k_t, s_t) = f(k_t, s_t) - c_t(s_t, k_t), \tag{127}$$

  showing that capital is a function of the current states k_t and s_t
- We have demonstrated that a recursive solution might exist
- But we have not shown that it actually exists or whether it is unique
- In principle, there might even be non-recursive solutions to the problem, leading to a non-Markov evolution of the state
- We will disregard these types of solutions throughout the course

Certainty Equivalence I

- *Exercise 15:* Consider the stochastic Brock and Mirman (1972) model with log utility and stochastic TFP A_t:

$$y_t = A_t k_t^{\alpha} \tag{128}$$

Show that the solution is given by

$$k_{t+1} = \alpha\beta A_t k_t^{\alpha} \tag{129}$$

- Due to log utility, full depreciation, and Cobb-Douglas: distribution of shock process does not affect solution
- Regarding the variance, a mean-preserving spread to A_t would not alter the agents' choice
- **Certainty equivalence**: people make decisions under uncertainty by acting as if future stochastic variables were sure to turn out equal to their expected values (Obstfeld and Rogoff 1996, p. 81)
- This also implies the absence of **precautionary behavior**, because people act as if the expected value would occur with 100% certainty

Certainty Equivalence II

- As in this special case, certainty equivalence is a **property of the model**
- But: certainty equivalence can also be an **artifact of a particular solution technique**, although the underlying model does not feature it
- For example, Jensen's inequality effects cannot arise with linear solutions (more on this later)

Equilibria with several agents

- In the Brock and Mirman (1972)-model we had a central planner directly choosing the state
- But: in most macroeconomic models, the law of motion for the relevant endogenous state variables will not only depend on the decisions of one agent
- Agents will e.g. interact via markets where they are price-takers
- We will thus require an equilibrium concept where the plans of different agents are combined in a consistent way

Setup

- Let x_t be the vector of state variables under the control of the agent and X_t the vector of the same variables resulting from "markets"
- Denote with Z_t the vector of exogenous states
- The agents problem then takes the form

$$V(x_t, X_t, Z_t) = \max_{u_t} \{r(x_t, X_t, Z_t, u_t) + \beta V(x_{t+1}, X_{t+1}, Z_{t+1})\}$$

$$s.t.$$

$$x_{t+1} = g(x_t, X_t, Z_t, u_t) \tag{130}$$

$$X_{t+1} = G(X_t, Z_t) \tag{131}$$

$$Z_{t+1} = F(Z_t) \tag{132}$$

- The function g describes the effect of the agents control variable on his own state
- The functions G and F are the **perceived law of motion** and describe the agent's beliefs about the evolution of aggregate states

The solution

- The solution of the problem will take the form

$$u_t = h(x_t, X_t, Z_t) \tag{133}$$

- In a representative agent framework, we will have $x_t = X_t$ in the solution so that

$$X_{t+1} \equiv G_A(X_t, Z_t) = g(X_t, X_t, Z_t, h(X_t, X_t, Z_t)) \tag{134}$$

- Note: symmetry/representativeness is always imposed **after** computing the individual policy functions

Recursive competitive rational expectations equilibrium

Definition 2 (Recursive competitive rational expectations equilibrium)

A recursive competitive rational expectations equilibrium consists of a policy function h, an actual aggregate law of motion G_A, and a perceived aggregate law G such that

i) given the perceived law G, the policy function h solves the agent's optimization problem
ii) the policy function h implies that $G_A = G$, i.e. the perceived and actual law of motion are consistent

- Although terminologically imprecise, shorthand names found in the literature are **recursive competitive equilibrium** or **rational expectations equilibrium**

Bibliography I

Acemoglu, Daron (2008). *Introduction to modern economic growth*. Princeton University Press.

Brock, William and Leonard Mirman (1972). “Optimal economic growth and uncertainty: the discounted case”. *Journal of Economic Theory 4* (3), 479–513.

Cass, David (1965). “Optimum growth in an aggregative model of capital accumulation”. *Review of Economic Studies 32* (3), 233–240.

Chow, Gregory C. (1997). *Dynamic economics: optimization by the Lagrange method*. Oxford: Oxford University Press.

Cochrane, John H. (2011). “Determinacy and identification with Taylor rules”. *Journal of Political Economy 119* (3), 565–615.

Greene, William H. (2011). *Econometric analysis*. 7th ed. Upper Saddle River: Prentice Hall.

Hamilton, James D. (1994). *Time series analysis*. Princeton, NJ: Princeton University Press.

Bibliography II

Heer, Burkhard and Alfred Maußner (2009). *Dynamic general equilibrium modeling*. 2nd ed. Berlin: Springer.

Klein, Paul (2000). "Using the generalized Schur form to solve a multivariate linear rational expectations model". *Journal of Economic Dynamics and Control 24* (10), 1405–1423.

Koopmans, Tjalling C. (1965). "On the concept of optimal economic growth". *The Econometric Approach to Development Planning*. North-Holland. Chap. 4, 225–287.

Ljungqvist, Lars and Thomas J. Sargent (2012). *Recursive macroeconomic theory*. 3rd ed. Cambridge, MA: MIT Press.

Miao, Jianjun (2014). *Economic dynamics in discrete time*. Cambrigde, MA: MIT Press.

Obstfeld, Maurice and Kenneth Rogoff (1996). *Foundations of international macroeconomics*. MIT Press.

Ramsey, Frank P. (1928). "A mathematical theory of saving". *Economic Journal 38* (152), 543–559.

Bibliography III

Romer, David H. (2012). *Advanced macroeconomics*. 4th ed. McGraw-Hill.