

Notes: ECMA33220 (F25) Introduction to Advanced Macroeconomic Analysis

Lecturer: Harald Uhlig
Notes by: Aden Chen

Tuesday 4th November, 2025

1 Time Series Basics

Definition 1.1. An $\text{AR}(k)$ (autoregressive) process is a time series $\{y_t\}$ satisfying

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_k y_{t-k} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2).$$

Here, y_{t-k} is called a **lag** of order k . A $\text{MA}(k)$ (moving-average) process is a time series $\{y_t\}$ satisfying

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_k \varepsilon_{t-k}, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2).$$

1.1 Forecasting

Consider the AR(1) process $y_t = \rho y_{t-1} + \varepsilon_t$ where $\varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$. Observe that $\mathbb{E}_t[y_{t+k}] = \rho^k y_t$. This can be generalized to the AR(n) case:

Consider the AR(n) process

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_n y_{t-n} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2).$$

We have $x_t = Bx_{t-1} + \varepsilon_t A$, where

$$x_t := \begin{bmatrix} y_t \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-n+1} \end{bmatrix}, \quad B = \begin{bmatrix} \varphi_1 & \varphi_2 & \cdots & \varphi_{n-1} & \varphi_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \quad A = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Note that B is called the **companion matrix**. As above,

$$\mathbb{E}_t[x_{t+k}] = B^k x_t \in \mathbb{R}^{n \times 1}.$$

Corollary 1.2. For an AR(m) process $\{y_t\}$, we have $\mathbb{E}_t[y_{t+k}] = (B^k x_t)_1$.

1.2 Impulse Response Functions

Definition 1.3. The **impulse response** $r(k)$ to a shock one standard deviation in size is the forecast for $y_k - \mathbb{E}[y_k]$, given $\varepsilon_0 = \sigma$ and $y_t = 0, t < 0$.

Remark 1.4. $\mathbb{E}[y_k]$ is the *unconditional expectation* of y_k , i.e., its *deterministic component*.

Proposition 1.5. For an AR(m) process $\{y_t\}$, the impulse response function is given by

$$r(k) = \sigma (B^k e_1)_1,$$

where B is the companion matrix.

Proposition 1.6. For a AR(1) process $y_t = \rho y_{t-1} + \varepsilon_t$,

- The impulse response converges to zero if $|\rho| < 1$,
- The absolutely value of the impulse response converges to a nonzero value if $|\rho| = 1$.
- The absolutely value of the impulse response explodes if $|\rho| > 1$.

If we denote $\theta_k := r(k)$, then if μ_t is the deterministic piece component of y_t , we have

$$y_t - \mu_t = \sum_{k=0}^{\infty} \theta_k \varepsilon_{t-k}$$

is an infinite-order MA representation of the AR process. In fact, we have the

Theorem 1.7 (Wold Representation Theorem). *Any covariance stationary time series can be represented as*

$$y_t = \mu_t + \sum_{k=0}^{\infty} \theta_k \varepsilon_{t-k},$$

where the ε_t 's are the one-step ahead linear forecast errors for the y_t 's, given information on lagged values of y_t , and where μ_t is a deterministic function of time.

Remark 1.8.

- The one-step ahead linear forecast error is defined as $\varepsilon_t = y_t - \mathbb{E}_{t-1}[y_t]$.
- Loosely speaking, a process is **covariance stationary** if it fluctuates around some point, rather than diverges, once deterministic time trends μ_t are removed.



Note that $\mathbb{E}_t[\varepsilon_{t+k}] = 0$ for each $k > 0$. Thus we have the

Corollary 1.9.

$$\mathbb{E}_t[y_{t+k}] = \mu_{t+k} + \sum_{j=k}^{\infty} \theta_j \mathbb{E}_t[\varepsilon_{t+k-j}].$$

1.3 Autocovariances

Definition 1.10. Let $\{y_t\}$ be a time series.

- The k th **autocovariance** at time t is defined as the covariance between y_t and y_{t-k} .
- The k th **autocorrelation** at time t is defined as the correlation between y_t and y_{t-k} .

Moreover, a time series $\{y_t\}$ is called **covariance stationary** if its mean of y_t as well as all autocovariances are finite and do not depend on t .

Example 1.11. Consider the mean zero AR(1) process $y_t = \rho y_{t-1} + \varepsilon_t$. The variance is given by

$$\mathbb{E}[y_t y_t] = \rho^2 \mathbb{E}[y_{t-1} y_{t-1}] + \sigma^2.$$

- If $|\rho| \geq 1$, then $\mathbb{V}(y_t) = \infty$.
- If $|\rho| < 1$, we have

$$\mathbb{V}(y_t) = \frac{\sigma^2}{1 - \rho^2}.$$

In this case the autocovariance is given by

$$\begin{aligned}\mathbb{E}[y_t y_{t-k}] &= \mathbb{E}[\varepsilon_t y_{t-k}] + \rho \mathbb{E}[y_{t-1} y_{t-k}] + \dots + \rho^k \mathbb{E}[y_{t-k} y_{t-k}] \\ &= \frac{\sigma^2 \rho^k}{1 - \rho^2},\end{aligned}$$

since $y_t = \rho^k y_{t-k} + \rho^{k-1} \varepsilon_{t-k-1} + \dots + \rho \varepsilon_{t-1} + \varepsilon_t$. Thus the k th autocorrelation is given by ρ^k .



From this discussion we have the

Proposition 1.12. *An AR(1) process with constant and finite mean and iid shocks is covariance stationary if and only if $|\rho| < 1$.*

To derive the autocovariances of a general AR(n) process, we first recall the following:

Let A, B be two matrices. The **Kronecker product** $A \otimes B$ is defined as the block matrix

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1n}B \\ A_{21}B & A_{22}B & \cdots & A_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mn}B \end{bmatrix},$$

where each element of A is multiplied with the entire matrix B . Let $\text{vec}(\cdot)$ denote the column wise vectorization operator. We have the following:

Proposition 1.13. *Let A, B, C be three matrices such that the matrix product ABC is well-defined. Then,*

$$\text{vec}(ABC) = (C' \otimes A) \text{vec}(B).$$

Now consider the mean zero covariance stationary AR(m) process $\{y_t\}$ and write $x_t = Bx_{t-1} + \varepsilon_t A$, $\mathbb{E}[\varepsilon_t^2] = \sigma^2$ as before. Define the autocovariances $\Gamma_k = \mathbb{E}[x_t x'_{t-k}]$. We have

$$\begin{aligned}\Gamma_0 &\coloneqq \mathbb{E}[x_t x'_t] = B \mathbb{E}[x_{t-1} x'_{t-1}] B' + A \sigma^2 A' \\ &= B \Gamma_0 B' + A \sigma^2 A'.\end{aligned}$$

Applying the $\text{vec}(\cdot)$ operator and using the Kronecker product property, we have

$$\text{vec}(\Gamma_0) = (B \otimes B) \text{vec}(\Gamma_0) + (A \otimes A)\sigma^2.$$

It turns out that if all eigenvalues of B lie strictly within the unit circle, then $(I - B \otimes B)$ is invertible and we have

Proposition 1.14. *For a covariance stationary AR(m) process $\{y_t\}$, the variance-covariance matrix of x_t is given by*

$$\text{vec}(\Gamma_0) = (I_{m^2} - B \otimes B)^{-1}((A \otimes A)\sigma^2).$$

Recall here that $(\Gamma_0)_{1,1}$ is the variance of y_t , and $(\Gamma_0)_{1,j}$ is the $(j-1)$ th autocovariance of y_t . Now observe the

Proposition 1.15 (Yule-Walker Equation).

$$\Gamma_k := \mathbb{E}[x_t x'_{t-k}] = B \mathbb{E}[x_{t-1} x'_{t-k}] = B \Gamma_{k-1}.$$

With this, we can iteratively compute the k th autocovariance of the AR(m) process by $(B^k \Gamma_0)_{1,1}$.

1.4 Lag Operator Calculus

Definition 1.16. The **lag operator** L shifts a time series back by one period: $(Ly)_t = y_{t-1}$.

We may use the lag operator to rewrite an AR(1) process $y_t = \rho y_{t-1} + \varepsilon_t$ as $(1 - \rho L)y_t = \varepsilon_t$. More generally, an AR(m) process can be written as

$$(1 - \varphi_1 L - \varphi_2 L^2 - \cdots - \varphi_m L^m) y_t = \varepsilon_t.$$

Definition 1.17. For an AR(m) process $(1 - \rho(L))y_t = \varepsilon_t$, the **characteristic polynomial** is defined as

$$\begin{aligned} p(\lambda) &:= \lambda^m (1 - \varphi(\lambda^{-1})) \\ &= \lambda^m - \varphi_1 \lambda^{m-1} - \varphi_2 \lambda^{m-2} - \cdots - \varphi_m. \end{aligned}$$

The complex-valued solutions $\lambda_1, \dots, \lambda_m$ of $p(\lambda) = 0$ are called the **roots** of the characteristic polynomial.

Proposition 1.18. *An AR(m) process is covariance stationary if and only if all roots of its characteristic polynomial lie strictly within the unit circle.*

Remark 1.19. Real-valued roots induce impulse response dynamics as seen in the AR(1) case.

- > 1 : explosive.
- $= 1$: unit root; shock never dies out.
- < 1 : shock eventually dies out.

Complex valued roots always come in conjugate pairs $a \pm bi$ and induce (damped or explosive) oscillatory impulse response dynamics with frequency $\arctan(b/a)$. 

2 Vector Autoregressions

Definition 2.1. A **vector autoregression** of order k for an m -dimensional vector y_t is a vector time series $\{y_t\}$ satisfying

$$y_t = B_1 y_{t-1} + \cdots + B_k y_{t-k} + u_t, \quad u_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \Sigma),$$

where $B_j \in \mathbb{R}^{m \times m}$ and where u_t is called the **one-step ahead linear forecast error**.

As before, we will ignore deterministic components such as time trends.

Definition 2.2. Let a be some m -dimensional vector. The **impulse response** $r_a(k)$ to the vector a is the forecast for y_k given $u_0 = a$ and $y_t = 0, t < 0$.

Example 2.3. For a VAR(1) process, $r_a(k) = B^k a$. 

Again, we can write a VAR(k) process as a VAR(1) process in companion form, by defining

$$\mathbf{x}_t := \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-k+1} \end{bmatrix}, \quad \mathbf{B} := \begin{bmatrix} B_1 & B_2 & \cdots & B_{k-1} & B_k \\ I_m & 0 & \cdots & 0 & 0 \\ 0 & I_m & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_m & 0 \end{bmatrix}, \quad \mathbf{A} := \begin{bmatrix} I_m \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

We thus will next only consider VAR(1) processes.

2.1 Diagonalization and Decoupling

Fix a covariance stationary VAR(1) process $y_t = B y_{t-1} + u_t$, $u_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \Sigma)$. If $B = V D V^{-1}$ is diagonalizable, then $B^k = V D^k V^{-1}$ and

$$r_a(k) = B^k a = V D^k V^{-1} a.$$

Remark 2.4. Not all matrices are diagonalizable, but all matrices admit a Jordan normal form, which can be used similarly. 

Write

$$V = [v_1 \ \cdots \ v_m], \quad D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_m \end{bmatrix}.$$

Note that the impulse response to $a = v_j$ is given by

$$r_{v_j}(k) = \lambda_j^k v_j.$$

Thus, the dynamics of a VAR(1) with a diagonalizable B can be described by the dynamics along each eigenvector v_j . The dynamics has been **decoupled**.

Proposition 2.5. *The roots $\lambda_1, \dots, \lambda_m$ can be calculated as the roots of the characteristic polynomial*

$$p(\lambda) := \det(\lambda I_m - B).$$

Remark 2.6. One can show that this is equivalent to the definition of characteristic polynomial we gave before. 

Proof. Write $B = VDV^{-1}$ and observe that

$$\begin{aligned} p(\lambda) &= \det(\lambda I - B) = \det(V) \det(\lambda I - D) \det(V^{-1}) \\ &= \det(\lambda I - D) = \prod_{j=1}^m (\lambda - \lambda_j). \end{aligned}$$

□

Proposition 2.7. ; A VAR(1) process is covariance stationary if and only if all roots of its characteristic polynomial lie strictly within the unit circle.

2.2 Cointegration

We next analyzes the case where some roots, say $\lambda_1, \dots, \lambda_r$ are stable, while all others are exactly equal to 1: $\lambda_{r+1} = \dots = \lambda_m = 1$. Furthermore, we assume that these **unit roots** affect every entry of y_t . That is, that the last $m - r$ columns of V have at least one nonzero entry in each row.

Definition 2.8. Define the **difference operator** Δ as $\Delta y_t := y_t - y_{t-1} = (1 - L)y_t$. We say a univariate time series y_t is **integrated** if y_t is not stationary but the first difference Δy_t is stationary.

Proposition 2.9. A univariate AR(k) process $\{y_t\}$ is integrated if and only if exactly one of the roots of its characteristic polynomial is equal to 1 and all others are stable.

Definition 2.10. The vector time series y_t is said to be **cointegrated of rank r** if each of the series taken individually is integrated, and r linear combination of the series, $\beta'y_t$, is stationary for some $\beta \in \mathbb{R}^{m \times r}$ of rank $r \geq 1$.

Now, split the matrices V , D , and V^{-1} (where $B = VDV^{-1}$) in the following way:

$$V = [v_r \ v_\perp], \quad D = \begin{bmatrix} D_r & 0 \\ 0 & I_{m-r} \end{bmatrix}, \quad V^{-1} = \begin{bmatrix} \beta' \\ \beta'_\perp \end{bmatrix},$$

where v_r , D_r , and β correspond to the stable eigenvalues, while v_\perp , and β_\perp correspond to the unit roots. (Note that \perp here only differentiates the two parts; it does not imply orthogonality.)

We have the

Proposition 2.11. Let $|\lambda_1|, \dots, |\lambda_r| < 1$ be the stable roots and let $\lambda_{r+1} = \dots = \lambda_m = 1$ be the unit roots. The matrix β as defined above is a cointegrating matrix of rank r for the VAR(1) process $y_t = By_{t-1} + u_t$.

Note also that using the notation above, we have

$$\begin{aligned} B^k a &= \begin{bmatrix} \nu_r & \nu_{\perp} \end{bmatrix} \begin{bmatrix} D_r^k & 0 \\ 0 & I_{m-r} \end{bmatrix} \begin{bmatrix} \beta' \\ \beta'_{\perp} \end{bmatrix} a \\ &= \underbrace{\nu_r D_r^k \beta' a}_{\text{transitory part}} + \underbrace{\nu_{\perp} \beta'_{\perp} a}_{\text{permanent part}}. \end{aligned}$$

In particular,

Proposition 2.12. *The long-run impulse response is given by $\lim_{k \rightarrow \infty} B^k a = \nu_{\perp} \beta'_{\perp} a$.*

2.3 Error Correction

Define $\alpha := \nu(I_r - D_r)$ and note that

$$\Delta y_t - u_t = (B - I)y_{t-1} = V(D - I)V^{-1}y_{t-1} = -\alpha\beta'y_{t-1},$$

where the last equality follows from the fact that $D - I$ is a diagonal matrix with the first r diagonal entries equal to $\lambda_j - 1$ and the last $m - r$ diagonal entries equal to 0. This is called the **error correction representation** of the VAR(1) process.

Remark 2.13.

- Note in particular that $\alpha\beta' = I - B$.
- With the error correction representation, we see that if $\beta'y_{t-1}$ is “large” (i.e., far from its equilibrium), then Δy_t will be “large” in the opposite direction to correct for this deviation.
- The change of y_t is driven by:
 - new shocks u_t ,
 - the convergence of the stationary component $\beta'y_{t-1}$ back to zero, and
 - comovements induced by the cointegrating relationships.



Proposition 2.14 (Engle and Granger, 1987). *A VAR(k) which is cointegrated of rank r can be given an error correction representation of the form*

$$A^*(L)\Delta y_t = -\alpha\beta'y_{t-1} + u_t,$$

Remark 2.15. With cointegration, there are linear combinations $\beta'y_{t-1}$ among the variables that tend to move together and “correct the error” of moving too far from zero.



3 Identification in VARs

Which vectors to choose for impulse responses? A *necessary* requirement is orthogonality. We seek to decompose u_t into mutually orthogonal components $u_t = A\varepsilon_t$, where ε_t are orthogonal shocks such that $\mathbb{E}[\varepsilon_t \varepsilon_t'] = I_m$. This is equivalent to finding A such that $\Sigma = AA'$.

Definition 3.1. An **impulse vector** a is a column of a matrix $A \in \mathbb{R}^{m \times m}$ such that $\Sigma = AA'$.

Proposition 3.2. Let a be an impulse vector and let $\tilde{A}\tilde{A}' = \Sigma$ be some other decomposition of Σ . If $\det \Sigma \neq 0$, then

$$a = \tilde{A}q,$$

for some vector q of unit length.

Proof. Let $AA' = \Sigma$ be some decomposition of Σ such that a is the j th column of A . Let $Q := \tilde{A}^{-1}A$ and note that

$$QQ' = \tilde{A}^{-1}A(A')(\tilde{A}^{-1})' = I_m.$$

Thus Q is an **orthogonal matrix** (a matrix whose columns are of unit length and mutually orthogonal; Q is orthogonal if and only if Q' is orthogonal). We have $a = \tilde{A}q$, where q is the j th column of Q . \square

Corollary 3.3. The set of decompositions of Σ is given by $\{\tilde{A}Q : QQ' = I_m\}$.

3.1 Cholesky Decomposition

The **Cholesky decomposition** of a positive definite matrix $\Sigma = AA'$ is a decomposition where A is lower triangular with real and positive diagonal entries. In practice, we order “slow-moving” variables first and “fast-moving” variables last, so that the resulting shock from the Cholesky decomposition is consistent with the causal ordering of the variables.

3.2 Blanchard-Quah Decomposition

The **Blanchard-Quah decomposition** decomposes shocks into *permanent* and *transitory* components. To obtain such a decomposition, we seek a vector a which has no permanent effect (using say the characterization of long-run effects from before):

$$0 = \lim_{k \rightarrow \infty} r_a(k) = \lim_{k \rightarrow \infty} B^k a.$$

Typically, this gives a up to a multiplicative constant. One can then find the full decomposition $\Sigma = AA'$ in the following way:

- Find any decomposition $\Sigma = \tilde{A}\tilde{A}'$ (say the Cholesky decomposition).

- We know $a = ca_0$ is a column of A , where a_0 is some particular solution to the permanent effect restriction and c is some constant. Thus, $ca_0 = \tilde{A}q$ for some unit length vector q which satisfies

$$\frac{q}{c} = \tilde{A}^{-1}a_0.$$

Using the unit length restriction on q , we can solve for c , q , and then a .

- The unit vector q can be completed to an orthogonal matrix Q . We then have $A = \tilde{A}Q$.

3.3 Sign Restrictions

Sample impulse responses which satisfy sign restrictions uniformly (Uhlig, 2005).

4 Business Cycles

4.1 The HP Filter

The **Hodrick-Prescott (HP) filter** decomposes a time series y_t into a trend component τ_t and a cyclical component c_t by solving

$$\min_{\tau_t} \sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} ((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2,$$

Remark 4.1. For quarterly data, $\lambda = 1600$ is typically used. Otherwise, for a change in observation frequency by a factor of k , λ should be changed by a factor of k^4 (Ravn and Uhlig, 2002). 

Remark 4.2. Two ways to get pro-cyclical productivity are TFP shocks and fixed costs. 

Remark 4.3. Empirical facts about business cycles:

- Most macroeconomic time series are **procyclical**, i.e., their business cycle component is positively correlated with output.
- Strongly procyclical: consumption, investment, hours worked.
- Mildly procyclical: housing starts, exports, labor productivity, real wages, inflation, short-term interest rates.
- Acyclical: government spending, oil prices, long-term interest rates, M1. 

4.2 A Benchmark Real Business Cycle Model

The social planner solves

$$\max_{c_t, n_t, k_t, y_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u(c_t, n_t) \quad \text{s.t.} \quad c_t + k_t = y_t + (1 - \delta)k_{t-1}, \quad y_t = f\left(A_t \frac{n_t}{k_{t-1}}\right) k_{t-1}$$

and $A_t > 0$ is some exogenous productivity process (E.g., $\log A_t = \rho \log A_{t-1} + \varepsilon_t$). Note that the production function exhibits constant returns to scale.

Remark 4.4. In the decentralized economy, the representative firm is given w_t, r_t and solves

$$\max_{k_{t-1}, n_t} f\left(A_t \frac{n_t}{k_{t-1}}\right) k_{t-1} - w_t n_t - r_t k_{t-1}.$$

The FOC dictates that

$$w_t = A_t f'\left(A_t \frac{n_t}{k_{t-1}}\right),$$

$$r_t = f\left(A_t \frac{n_t}{k_{t-1}}\right) - f'\left(A_t \frac{n_t}{k_{t-1}}\right) A_t \frac{n_t}{k_{t-1}}.$$

In a competitive equilibrium, profits are zero by Euler's theorem ($F(K, N) = F_K(K, N)K + F_N(K, N)N$). Thus we obtain

$$w_t n_t = f' \left(A_t \frac{n_t}{k_{t-1}} \right) A_t n_t.$$



Remark 4.5. Some specifications of utility functions:

- Log:

$$u(c_t, n_t) = \log c_t - \chi n_t^{1+\frac{1}{\varphi}}.$$

Here, φ is the Frisch elasticity of labor supply.

- Cobb-Douglas:

$$u(c_t, n_t) = c_t^\alpha (1 - n_t)^{1-\alpha}.$$

- Constant relative risk aversion (CRRA):

$$u(c_t, n_t) = \frac{c_t^{1-\eta} - 1}{1 - \eta} - \chi n_t^{1+\frac{1}{\varphi}}.$$

As $\eta \rightarrow 1$, this converges to the log utility specification.

- Constant Frisch Elasticity (CFE):

$$u(c_t, n_t) = \frac{c_t^{1-\eta} \left(1 - (1 - \eta) \chi n_t^{1+\frac{1}{\varphi}} \right)^\eta - 1}{1 - \eta}.$$

Here, φ is the Frisch elasticity of labor supply.

- Greenwood-Hercowitz-Huffman (GHH):

$$u(c_t, n_t) = \frac{\left(c_t - \chi n_t^{1+\frac{1}{\varphi}} \right)^{1-\eta} - 1}{1 - \eta}.$$



Definition 4.6. The **Frisch elasticity of labor supply** is defined as the elasticity of labor supply with respect to the wage, holding the marginal utility λ of wealth constant.

Remark 4.7. From $n = \text{const} \cdot w^\varphi$ we can easily see that the Frisch elasticity is given by φ :

$$\frac{\partial n}{\partial w} = \varphi \cdot \text{const} \cdot w^{\varphi-1} \implies \frac{\partial n}{\partial w} \frac{w}{n} = \varphi.$$



4.3 The Bellman Equation

$$\mathcal{L} = u(c, n) + \beta \mathbb{E}[V(k', A')] + \lambda \left(f\left(A \frac{n}{k}\right) k + (1 - \delta)k - c - k' \right).$$

4.3.1 The Envelope Condition

At the optimum,

$$V(k, A) = \mathcal{L} = u(c, n) + \beta \mathbb{E}[V(k', A')] + \lambda \underbrace{\left(f\left(A \frac{n}{k}\right) k + (1 - \delta)k - c - k' \right)}_{X(k, A)}.$$

Thus the envelope theorem gives

$$V_k(k, A) = \lambda \underbrace{\left(f\left(A \frac{n}{k}\right) - f'\left(A \frac{n}{k}\right) A \frac{n}{k} + (1 - \delta) \right)}_{k(k, A), \text{ rental rate of capital}} =: \lambda R(k, A).$$

Interpretation:

- With the addition of Δk of k , the budget constraint is relaxed by $\Delta X := R(k, A)\Delta k$, each unit of which is valued at λ .
- From $\lambda = u_c$ and $\lambda = \beta \mathbb{E}[V_k(k', A')]$, we have

$$V_k(k, A) = u_c \frac{\partial X}{\partial k}, \quad V_k(k, A) = \beta \mathbb{E}[V_k(k', A')].$$

At the optimum and marginally, consuming and investing are equally valuable.

Proposition 4.8. Define $w := f'(An/k)A$.

$$-u_n(c, n) = \lambda w, \quad -\frac{u_n(c, n)}{u_c(c, n)} = w.$$

4.4 The Lucas Asset Pricing Equations

Recall that

$$\begin{aligned} \lambda &= \beta \mathbb{E}[V_k(k', A')|k, A], \\ V_k(k, A) &= \lambda R(k, A). \end{aligned}$$

From these we have

$$1 = \mathbb{E} \left[\underbrace{\beta \frac{\lambda_{t+1}}{\lambda_t}}_{M_{t+1}, \text{ stochastic discount factor}} R_{t+1} | k, A \right].$$

The SDF is the price of consumption in $t + 1$ in units of goods at time t . Thus,

$$\mathbb{E}_t[R_{t+1}] = \frac{1 - \mathbb{C}(M_{t+1}, R_{t+1})}{\mathbb{E}_t[M_{t+1}]}.$$

Interpretation:

- The return R_{t+1} is high if it has a negative covariance with the SDF M_{t+1} .
- Higher expected returns compensate for holding market risk. The market price of risk is encoded in the stochastic discount factor.

4.5 A Lagrangian Approach

$$\mathcal{L} := \max \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[u(c_t, n_t) + \lambda_t \left(f \left(A_t \frac{n_t}{k_{t-1}} \right) k_{t-1} + (1 - \delta) k_{t-1} - c_t - k_t \right) \right].$$

FOC:

$$\begin{aligned} c_t : & u_c(c_t, n_t) = \lambda_t, \\ n_t : & -u_n(c_t, n_t) = -\lambda_t w_t, \\ \lambda_t : & c_t + k_t = y_t + (1 - \delta) k_{t-1}, \\ k_t : & \lambda_t = \beta \mathbb{E}_t [\lambda_{t+1} R_{t+1}], \end{aligned}$$

where the expectation in the last equation is taken conditional on information at time t , since λ_t is chosen at time t .

4.6 A Competitive Equilibrium

Given k_{t-1} and the processes for wages w_t and capital rental rates r_t , the representative household solves

$$V_0 := \max_{c_t, n_t, k_t} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t, n_t^{(s)}) \right] \quad \text{s.t.} \quad c_t + k_t = w_t n_t^{(s)} + r_t k_{t-1} + (1 - \delta) k_{t-1}.$$

Given wages w_t and capital rental rates r_t , the representative firm maximizes profits:

$$\max_{k_{t-1}, n_t^{(d)}} f \left(A_t \frac{n_t^{(d)}}{k_{t-1}} \right) k_{t-1} - w_t n_t^{(d)} - r_t k_{t-1}.$$

We impose market clearing:

$$c_t + k_t = f \left(A_t \frac{n_t^{(d)}}{k_{t-1}} \right) k_{t-1} + (1 - \delta) k_{t-1}, \quad n_t^{(s)} = n_t^{(d)}.$$

4.7 A Balanced Growth Path

A balanced growth path with growth factor ζ is a sequence $\{c_t, y_t, k_t, w_t, n_t, r_t, R_t\}_{t=0}^{\infty}$ such that

$$\begin{aligned} c_t &= \zeta^t c, \\ y_t &= \zeta^t y, \\ k_t &= \zeta^t k, \\ w_t &= \zeta^t w, \\ n_t &= n, \\ r_t &= r, \\ R_t &= R. \end{aligned}$$

Consider the non-stochastic case where $A_t = \gamma^t A_0$ for all t . When are the first order conditions consistent with balanced growth? Recall that we have

$$\begin{aligned} -u_n(\gamma^t c_0, n) &= \lambda_t \gamma^t w_0, \\ u_c(\gamma^t c_0, n) &= \lambda_t, \\ \lambda_t &= \beta \lambda_{t+1} R. \end{aligned}$$

Theorem 4.9 (King-Plosser-Rebelo, 2001). *Under mild, additional assumptions, these equations imply that for some $\eta > 0$ and some function $\Phi(\cdot)$, up to scaling and constants,*

$$u(c, n) = \begin{cases} \frac{(c\Phi(n))^{1-\eta}}{1-\eta}, & \eta \neq 1, \\ \log c + \Phi(n), & \eta = 1. \end{cases}$$

This is satisfied by the log utility, Cobb-Douglas, CFR, but not CRRA or GHH.

4.8 Steady State

Consider the steady state with $A \equiv 1$:

$$\begin{aligned} 1 &= \beta R \\ R &= r + 1 - \delta \\ r \frac{k}{y} &= 1 - \frac{wn}{y} \\ \frac{c}{y} + \delta \frac{k}{y} &= 1 \\ \theta &:= \frac{wn}{y} = \frac{f'(x)x}{x}, \quad x := \frac{n}{k} \\ \kappa &:= \frac{-u_n n}{u_c c} = \frac{wn}{y} \frac{y}{c}. \end{aligned}$$

Calibration:

- Time t counts quarters.
- Data: $R = 1.01$.¹ Thus $\beta \approx 0.99$.
- Data: $\delta = 0.02$. Thus $r = 0.03$.
- Data: labor share $\theta := wn/y = 0.64$. Thus $k/y = 12$. Thus $c/y = 0.76$. Thus $\kappa = \theta/0.76 = 0.84$.

4.9 A Few Parameters

- We define ω_D by

$$\frac{1}{\omega_D} := \frac{-f''(x)x}{f'(x)}, \quad x := \frac{n}{k}.$$

- The process

$$\log A_t = \rho \log A_{t-1} + \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2),$$

has a estimated $\rho \approx 1$ and $\sigma \approx 0.007$. We use $\rho = 0.95$ to avoid unit root complications.

- The negative of the coefficient of relative risk aversion or the negative inverse of the intertemporal elasticity of substitution

$$\eta := -\frac{cu_{cc}}{u_c}.$$

Many believe $1 \leq \eta \leq 3$; we will pick $\eta = 2$.

- The cross-elasticity between consumption and labor

$$\psi := \frac{u_{cn}c}{u_n}.$$

- With $\kappa = -u_n n / u_c c$,

$$\frac{u_{cn}n}{u_c} = -\kappa\psi.$$

- With the Frisch elasticity φ at the steady state,

$$\frac{u_{nn}n}{u_n} = \frac{1}{\varphi} + \frac{u_{cn}c}{u_n} \frac{u_c}{u_{cc}c} \frac{u_{cn}n}{u_c}.$$

The last four are evaluated at the steady state.

It turns out that with growth consistent utility functions (King-Plosser-Rebelo preferences), then

- the η defined above is consistent with that in the utility function,

¹Back of the envelope: 1% bonds, 7% equity; equity premium 6%. Deducting inflation, 4%. Quarterly 1%.

- $\psi := \frac{u_{cn}c}{u_n} = 1 - \eta.$
- $\kappa := \frac{-u_n n}{u_c c} = -\frac{\Phi'(n)n}{\Phi(n)}.$

Thus with growth-consistent preferences and a value for φ , it remains to pick η .

4.10 Loglinearization

For x_t , let $\hat{x}_t := \log(x_t/x)$ be the log-deviation from the steady state. Then $100\hat{x}_t$ is (approximately) the percentage deviation from the steady state \bar{x} :

$$x_t = \bar{x}e^{\hat{x}_t} \approx \bar{x}(1 + \hat{x}_t).$$

Proposition 4.10 (Log-linear Calculus).

(i)

$$x_t y_t = z_t \implies \hat{x}_t + \hat{y}_t = \hat{z}_t.$$

(ii)

$$x_t + y_t = z_t \implies \bar{x}\hat{x}_t + \bar{y}\hat{y}_t = \bar{z}\hat{z}_t.$$

(iii)

$$y_t = h(x_t) \implies \hat{y}_t = \frac{h'(\bar{x})\bar{x}}{h(\bar{x})}\hat{x}_t.$$

(iv)

$$z_t = h(x_t, y_t) \implies \hat{z}_t = \frac{h_x \bar{x}}{h}\hat{x}_t + \frac{h_y \bar{y}}{h}\hat{y}_t.$$

Proof. (i)

$$\log(x_t y_t = z_t) - \log(\bar{x}\bar{y} = \bar{z}) \implies \hat{x}_t + \hat{y}_t = \hat{z}_t.$$

Alternatively, use

$$\bar{x}(1 + \hat{x}_t) \cdot \bar{y}(1 + \hat{y}_t) = \bar{z}(1 + \hat{z}_t)$$

and ignore second order terms.

(ii)

$$\bar{x}(1 + \hat{x}_t) + \bar{y}(1 + \hat{y}_t) = \bar{z}(1 + \hat{z}_t)$$

and ignore second order terms.

□