

# Identification, ARMA Models, and Wold Decomposition

## TA Session 2: Reduced-Form Time Series Analysis

Rafael Lincoln

Econometrics II – NYU PhD Program

Spring 2026

# Today's Roadmap

## Part 1: Theory

### 1. Identification in Time Series

- ▶ Global vs. local identification
- ▶ Observational equivalence and ACGF
- ▶ Causality & invertibility

### 2. Reduced-Form Models: AR, MA, ARMA

- ▶ Definitions, stationarity, and invertibility
- ▶ Observational equivalence in MA processes

### 3. Wold's Decomposition Theorem

## Part 2: Exercises

- ▶ Pset 1 (Problems 1–8): Tim + Medeiros practice problems

# **Part I: Identification in Time Series**

The Challenge of Learning from Data

**Always Remember: Identification is a non-testable assumption!**

# The Identification Problem: Setup

**Setup:** We have a parametric model  $\mathcal{M}(\theta)$  with structural parameters  $\theta \in \Theta$ .

The model generates a family of distributions for observed data:

$$\mathbb{D}(\mathbf{Y}^T \mid \theta), \quad \text{where } \mathbf{Y}^T = \{y_t\}_{t=1}^T$$

## The Fundamental Question:

*Can we uniquely determine  $\theta$  from the distribution of the data?*

### Why This Matters

If two different values of  $\theta$  generate the **same distribution** of data, then we **cannot distinguish** between them using any amount of data!

# Global and Local Identification

## Definition 1 (Global Identification)

$\theta_0 \in \Theta$  is **globally identified** if for all  $\theta \in \Theta$ :

$$\mathbb{D}(\mathbf{Y}^T \mid \theta) = \mathbb{D}(\mathbf{Y}^T \mid \theta_0) \implies \theta = \theta_0$$

## Definition 2 (Local Identification)

$\theta_0 \in \Theta$  is **locally identified** if there exists a neighborhood  $U_{\theta_0}$  such that for all  $\theta \in U_{\theta_0}$ :

$$\mathbb{D}(\mathbf{Y}^T \mid \theta) = \mathbb{D}(\mathbf{Y}^T \mid \theta_0) \implies \theta = \theta_0$$

**Intuition:** Global  $\Rightarrow$  unique in entire space; Local  $\Rightarrow$  unique in a neighborhood.

How to check? → Appendix

# Linear Gaussian Models: Why They Matter for Time Series

For linear Gaussian models:  $\mathbf{Y}^T \sim \mathcal{N}(\mu(\theta), \Sigma(\theta))$

The mean and variance **fully characterize** the distribution.

## Connection to Time Series

**Most time series models we study are linear Gaussian!**

- ▶ AR, MA, ARMA processes with Gaussian innovations
- ▶ VARs and linearized DSGE models
- ▶ Wold decomposition assumes covariance stationarity

⇒ **Second-moment properties** (autocovariances) become the key object for identification!

**Implication:** If two parameter values yield the same autocovariance structure, they are **observationally equivalent**.

# Observational Equivalence

## Definition 3 (Observational Equivalence)

Two models  $\mathcal{M}(\theta_1)$  and  $\mathcal{M}(\theta_2)$  are **observationally equivalent** if:

$$\mathbb{D}(\mathbf{Y}^T \mid \theta_1) = \mathbb{D}(\mathbf{Y}^T \mid \theta_2)$$

**In Time Series:** For covariance-stationary Gaussian processes, observational equivalence reduces to having the **same second-moment properties**.

## The Identification Challenge

Different structural models/parameters can generate **identical autocovariance functions**.

⇒ Data alone cannot distinguish between them!

**Next:** How do we formalize “same second-moment properties”?

# The Autocovariance Generating Function (ACGF)

**Tool for comparing second-moment properties:**

**Definition 4 (ACGF)**

For a covariance-stationary process with autocovariances  $\Gamma(j)$ :

$$\Omega(z) = \sum_{j=-\infty}^{\infty} \Gamma(j)z^j$$

**Key Insight: The Identification Challenge in Macroeconometrics**

**Two time series models are observationally equivalent** if and only if they have the **same ACGF**:

$$\Omega(z; \theta_1) = \Omega(z; \theta_2) \quad \forall z$$

This is the root of why structural interpretation requires **additional identifying assumptions!**

## Preview: Types of Indeterminacy

Observational equivalence can lead to **indeterminacy** — multiple structural interpretations consistent with the same data.

**Three main sources** (covered in detail in the VAR section):

1. **Static Indeterminacy:** Orthogonal rotations of shocks
2. **Dynamic Indeterminacy:** Non-fundamental representations
3. **Size Indeterminacy:** More shocks than observables

Example: MA(1) “Root Flipping”

Consider  $y_t = u_t + \theta u_{t-1}$  with  $|\theta| < 1$ .

The process with  $\theta^* = 1/\theta$  and rescaled variance has **identical autocovariances!**

Both:  $\rho_1 = \frac{\theta}{1+\theta^2}$  (same for  $\theta$  and  $1/\theta$ )

*We'll develop identification strategies when we cover Structural VARs.*

# Causality and Invertibility: General Definitions

Let  $\{Y_t\}$  be a stationary process and  $\{Z_t\}$  an input/shock process.

## Definition 5 (Causality)

$\{Y_t\}$  is **causal with respect to**  $\{Z_t\}$  if  $Y_t$  is  $\mathcal{F}_t^Z$ -measurable for all  $t$ .

## Definition 6 (Invertibility)

$\{Y_t\}$  is **invertible with respect to**  $\{Z_t\}$  if  $Z_t$  is  $\mathcal{F}_t^Y$ -measurable for all  $t$ .

### Interpretations:

- ▶  $Y_t$  depends only on **current and past** values of  $Z$  — no future inputs needed.
- ▶ The shocks  $Z_t$  can be **recovered from current and past observables**  $Y_t$ .

# VARMA: Causality and Invertibility

## Definition 7

A VARMA process  $\{y_t\}$  is **causal** w.r.t.  $\{z_t\}$  if  $y_t \in \text{span}(z_\tau, -\infty < \tau \leq t)$

It is **invertible** w.r.t.  $\{z_t\}$  if  $z_t \in \text{span}(y_\tau, -\infty < \tau \leq t)$

**Sufficient conditions** for  $\Phi(L)y_t = \Theta(L)z_t$ :

- ▶ **Causality:**  $\Phi(L)$  has 1-sided inverse  $\Rightarrow$  VMA( $\infty$ ):  $y_t = \Psi(L)z_t$
- ▶ **Invertibility:**  $\Theta(L)$  has 1-sided inverse  $\Rightarrow$  VAR( $\infty$ ):  $\Pi(L)y_t = z_t$

**Key takeaways:**

- ▶ Structural macro models yield VMA representations  $\Rightarrow$  always causal
- ▶ Invertibility is **far from guaranteed!** (e.g., 5 shocks but 2 observables)

# Why Invertibility Matters

## For estimation and structural analysis:

If the process is invertible:

- ▶ Innovations  $z_t$  can be recovered from data
- ▶ We can estimate VAR and back out the shocks
- ▶ IRFs and variance decompositions become meaningful

If NOT invertible:

- ▶ VAR residuals are **not** the structural shocks
- ▶ Standard SVAR methods may give misleading results
- ▶ Need alternative identification strategies

## Necessary Condition

For invertibility:  $n_{\text{shocks}} \leq n_{\text{observables}}$

But this is **not sufficient!** (Forward-looking behavior can break invertibility)

# Part I: Summary

Concept	Key Point
Global Identification	Unique $\theta$ in entire parameter space
Local Identification	Unique in neighborhood
Linear Gaussian	Second moments fully characterize distribution
Observational Equiv.	Same ACGF $\Rightarrow$ indistinguishable
Causality	$Y_t$ is $\mathcal{F}_t^Z$ -measurable (depends on current/past $Z$ 's)
Invertibility	$Z_t$ is $\mathcal{F}_t^Y$ -measurable (recoverable from current/past $Y$ 's)

**Key Insight:** In linear Gaussian time series, identification hinges on second-moment properties. Observational equivalence  $\Rightarrow$  **indeterminacy**, which requires additional assumptions to resolve (covered in VAR section).

## **Part II: Reduced-Form Models**

AR, MA, and ARMA Processes

# The AR( $p$ ) Process: Definition

## Definition 8 (Autoregressive Process of Order $p$ )

A stochastic process  $\{Y_t\}$  is AR( $p$ ) if:

$$Y_t = \mathbb{E}[Y_t | \mathcal{F}_{t-1}] + u_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + u_t$$

Equivalently:  $\Phi_p(L)y_t = \phi_0 + u_t$ , where:

$$\Phi_p(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$$

and  $u_t$  is white noise:  $\mathbb{E}[u_t | \mathcal{F}_{t-1}] = 0$ ,  $\mathbb{E}[u_t^2 | \mathcal{F}_{t-1}] = \sigma^2$ .

**Interpretation:** Current value = conditional expectation given past + innovation.

**Note:** AR processes are **always invertible** (by construction,  
 $u_t = y_t - \mathbb{E}[y_t | \mathcal{F}_{t-1}]$ ).

# The Companion Form

**Idea:** Cast AR( $p$ ) as a VAR(1) by stacking variables.

Definition 9 (Companion Form)

$$\underbrace{\begin{bmatrix} Y_t \\ Y_{t-1} \\ \vdots \\ Y_{t-p+1} \end{bmatrix}}_{\mathbf{X}_t} = \begin{bmatrix} \phi_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \underbrace{\begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}}_F \underbrace{\begin{bmatrix} Y_{t-1} \\ Y_{t-2} \\ \vdots \\ Y_{t-p} \end{bmatrix}}_{\mathbf{X}_{t-1}} + \begin{bmatrix} u_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Compactly:  $\mathbf{X}_t = c + F\mathbf{X}_{t-1} + v_t$

**Why useful?**

- ▶  $\{\mathbf{X}_t\}$  is Markovian even though  $\{Y_t\}$  is not
- ▶ Stationarity analysis reduces to eigenvalue analysis of  $F$

# AR( $p$ ): Stationarity Conditions

## Theorem 10 (Asymptotic Stationarity of AR( $p$ ))

An AR( $p$ ) process is **asymptotically stationary** if the eigenvalues of the companion matrix  $F$  are all **inside the unit circle**.

Sketch.

Iterating backward:  $\mathbf{X}_t = F^t \mathbf{X}_0 + \sum_{i=0}^{t-1} F^i v_{t-i}$ . For this not to explode, need  $F^t \rightarrow 0$ , which requires  $|\lambda_i| < 1$  for all eigenvalues. □

## Corollary 11 (Characteristic Polynomial Version)

AR( $p$ ) is stationary iff roots of  $\Phi_p(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0$  lie **outside the unit circle**.

**Connection:** Eigenvalues of  $F$  are reciprocals of polynomial roots.

# AR( $p$ ): Moments (Yule-Walker Equations)

For a weakly-stationary AR( $p$ ) process:

**Expected value:**

$$\mathbb{E}[y_t] = \frac{\phi_0}{1 - \sum_{i=1}^p \phi_i}$$

**Autocovariance** (Yule-Walker Equations):

$$\gamma_k = \sum_{i=1}^p \phi_i \gamma_{k-i}, \quad k \geq 1$$

**Variance:**

$$\gamma_0 = \sigma^2 + \sum_{i=1}^p \phi_i \gamma_i$$

Key Pattern

ACF decays geometrically — signature of AR processes!

# AR(1) Example: Monetary Policy and Inflation

**Recall from Lecture 1:** The simple monetary transmission model:

$$y_t = -\gamma(i_t - \mathbb{E}_t[\pi_{t+1}]) + u_{1t} \quad (\text{IS curve})$$

$$\pi_t = \lambda y_t + u_{2t} \quad (\text{Phillips curve})$$

$$i_t = \rho \pi_t + i^* \quad (\text{Taylor rule})$$

**Reduced form for inflation:**

$$\pi_t = \phi_0 + \phi_1 \pi_{t-1} + v_t$$

where  $\phi_1 = 1 + \lambda \gamma(\rho - 1)$  (function of structural parameters!)

**Stability requires:**  $|\phi_1| < 1 \Leftrightarrow \text{Taylor Principle: } \rho > 1$

*Key insight:* The reduced-form AR coefficient encodes structural policy parameters!

# AR(1) Inflation Example: Behavior by $\rho$

Cases (with  $\lambda = 0.5$ ,  $\gamma = -0.5$ ):

Taylor Rule $\rho$	Reduced-form $\phi_1$	Behavior
$\rho = 0$	$\phi_1 > 1$	Explosive inflation
$0 < \rho < 1$	$\phi_1 > 1$	Explosive inflation
$\rho = 1$	$\phi_1 = 1$	Random walk
$1 < \rho < 5$	$0 < \phi_1 < 1$	Stationary, persistent
$5 < \rho < 9$	$-1 < \phi_1 < 0$	Stationary, oscillating
$\rho > 9$	$\phi_1 < -1$	Explosive (over-reaction)

**ACF:**  $\rho_k = \phi_1^k$  — geometric decay toward zero.

See figures in Appendix [→ Figures](#)

# The MA( $q$ ) Process: Definition

Definition 12 (Moving Average Process of Order  $q$ )

A process  $\{y_t\}$  is MA( $q$ ) if:

$$y_t = \theta_0 + u_t + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q} = \theta_0 + \theta_q(L)u_t$$

where  $\theta_q(L) = 1 + \theta_1 L + \cdots + \theta_q L^q$  and  $u_t \sim WN(0, \sigma^2)$ .

**Key difference from AR:** Current value is a weighted sum of current and **past innovations** (not past values of  $y$ ).

Proposition 1

*All finite-order MA( $q$ ) processes are **covariance-stationary** and **mean-ergodic**.*

**Intuition:** Finite sum of stationary (white noise) terms is always stationary.

## MA( $q$ ): Moments

**Expected value:**  $\mathbb{E}[y_t] = \theta_0$

**Variance:**

$$\gamma_0 = \sigma^2 \left( 1 + \sum_{j=1}^q \theta_j^2 \right)$$

**Autocovariance:**

$$\gamma_k = \begin{cases} \left( \theta_k + \sum_{j=1}^{q-k} \theta_j \theta_{j+k} \right) \sigma^2 & \text{if } k \leq q \\ 0 & \text{if } k > q \end{cases}$$

### Key Pattern

ACF **cuts off** after lag  $q$  — signature of MA processes!  
(Contrast with AR: geometric decay)

## MA( $q$ ): Invertibility

### Definition 13 (Invertibility)

An MA( $q$ ) process is **invertible** if  $u_t$  lies in the space spanned by current and past observables:  $u_t \in \text{span}(y_\tau, -\infty < \tau \leq t)$ .

Equivalently, MA( $q$ ) can be written as AR( $\infty$ ):

$$u_t = \theta_q(L)^{-1}(y_t - \theta_0) = \sum_{j=0}^{\infty} \psi_j(y_{t-j} - \theta_0)$$

### Theorem 14 (Invertibility Condition)

MA( $q$ ) is invertible iff roots of  $\theta_q(z) = 1 + \theta_1 z + \dots + \theta_q z^q = 0$  lie **outside the unit circle**.

**Practical importance:** Invertibility allows expressing  $u_t$  as function of observables — essential for estimation!

# MA(1): Observational Equivalence Example

**Key fact:** For any MA process, there exists a family of observationally equivalent alternatives with the same mean and autocovariances.

**Example:** MA(1) with  $y_t = u_t + \theta u_{t-1}$ ,  $u_t \sim (0, \sigma^2)$ .

**Autocorrelation at lag 1:**

$$\rho_1 = \frac{\theta}{1 + \theta^2}$$

**Observation:** Both  $\theta$  and  $1/\theta$  give the **same**  $\rho_1$ !

The **invertible** representation ( $|\theta| < 1$ ) is obs. equivalent to a **non-invertible** one:

- ▶ Original:  $\theta, \sigma^2$
- ▶ Alternative:  $\theta^* = 1/\theta, \sigma^{*2} = \theta^2\sigma^2$

## Resolution

We **cannot test** statistically whether a process is invertible!

We impose invertibility as an *a priori* identifying assumption.

# MA Observational Equivalence: ACGF Perspective

**Tool:** The Autocovariance Generating Function.

**Proposition 2**

For MA( $q$ ):  $y_t - \theta_0 = \theta_q(L)\varepsilon_t$ , the ACGF is:

$$\Gamma_Y(z) = \theta_q(z)\theta_q(z^{-1})\sigma_\varepsilon^2$$

**Two MA processes are observationally equivalent** if they have:

1. Same mean:  $\theta_0^X = \theta_0^Y$
2. Same ACGF:  $\Gamma_X(z) = \Gamma_Y(z)$

**For MA(1):** Matching  $\Gamma(z) = (1 + \theta z)(1 + \theta z^{-1})\sigma^2$  allows the “flip”:

$$\theta \leftrightarrow 1/\theta, \quad \sigma^2 \leftrightarrow \theta^2\sigma^2$$

# The ARMA( $p, q$ ) Process

## Definition 15 (ARMA Process)

A process  $\{Y_t\}$  is ARMA( $p, q$ ) if:

$$Y_t = \phi_0 + \sum_{i=1}^p \phi_i Y_{t-i} + u_t + \sum_{j=1}^q \theta_j u_{t-j}$$

Or:  $\Phi_p(L)Y_t = \phi_0 + \Theta_q(L)u_t$

## Proposition 3 (Stationarity and Invertibility)

- ▶ **Stationarity:** Roots of  $\Phi_p(z) = 0$  outside unit circle (AR part)
- ▶ **Invertibility:** Roots of  $\Theta_q(z) = 0$  outside unit circle (MA part)

**ACGF:**

$$\Gamma(z) = \frac{\Theta_q(z)\Theta_q(z^{-1})}{\Phi_p(z)\Phi_p(z^{-1})}\sigma_\varepsilon^2$$

# ARMA: Common Factor Problem

**Identification issue:** If AR and MA polynomials share a common root, the model is **not identified**.

**Example:** ARMA(1,1) with  $(1 - \phi L)y_t = (1 + \theta L)u_t$

If  $\phi = -\theta$  (common factor):

$$(1 - \phi L)y_t = (1 - \phi L)u_t \implies y_t = u_t$$

The ARMA(1,1) is **observationally equivalent to white noise!**

## Solution

For a **canonical** (identified) ARMA representation, require:  
 $\Phi_p(z)$  and  $\Theta_q(z)$  have **no common roots**.

# ARMA: Equivalent Representations

A stationary and invertible ARMA( $p, q$ ) admits:

## 1. MA( $\infty$ ) Representation:

$$y_t = \mu + \psi(L)u_t = \mu + \sum_{j=0}^{\infty} \psi_j u_{t-j}$$

where  $\psi(L) = \Phi_p(L)^{-1}\Theta_q(L)$

## 2. AR( $\infty$ ) Representation:

$$\pi(L)(y_t - \mu) = u_t$$

where  $\pi(L) = \Theta_q(L)^{-1}\Phi_p(L)$

### Key Insight

ARMA provides a **parsimonious parametric approximation** to MA( $\infty$ )/AR( $\infty$ ) — crucial for Wold decomposition!

## Summary: AR vs MA vs ARMA

Property	AR( $p$ )	MA( $q$ )	ARMA( $p, q$ )
Stationarity	Roots of $\Phi(z)$ outside UC	Always stationary	Roots of $\Phi(z)$ outside UC
Invertibility	Always invertible	Roots of $\Theta(z)$ outside UC	Roots of $\Theta(z)$ outside UC
ACF pattern	Geometric decay	Cuts off after lag $q$	Geometric decay
PACF pattern	Cuts off after lag $p$	Geometric decay	Geometric decay

**UC** = Unit Circle; **PACF** = Partial Autocorrelation Function

## **Part III: Wold's Decomposition Theorem**

The Fundamental Representation of Stationary Processes

# Wold's Representation Theorem

## Theorem 16 (Wold, 1938)

Let  $\{y_t\}$  be an  $n$ -dimensional **covariance-stationary** time series. There exists an  $n \times n$  lag polynomial  $\Psi(L)$  and  $u_t \sim WN(0, \Sigma)$  such that:

$$y_t = \Psi(L)u_t + d_t, \quad \Psi(L) = I + \sum_{\ell=1}^{\infty} \Psi_{\ell} L^{\ell}$$

where:

1.  $\Psi(L)$  is **square-summable**
2.  $u_t = y_t - \mathbb{E}^*[y_t | \{y_{\tau}\}_{-\infty < \tau \leq t-1}]$  (one-step-ahead forecast errors)
3.  $\{y_t\}$  is **invertible** with respect to  $u_t$
4.  $\{d_t\}$  is a **purely deterministic** process

**In words:** Any covariance-stationary time series = VMA( $\infty$ ) + deterministic

# Why Wold Matters: The Key Insight

The Wold decomposition does something **very simple**: it splits a process  $\{y_t\}$  into:

1. **One-step-ahead prediction errors**  $u_t$
2. **A perfectly predictable residual**  $d_t$

## Three Equivalent Representations

For second-order properties, we can **freely map** between:

1. **Autocovariance function**:  $\{\Gamma_k\}_{k=0}^{\infty}$
2. **Spectral density**:  $s_y(\omega)$
3. **Wold decomposition**:  $\{\Psi_j, \Sigma\}$

**Mapping**:  $\Psi_\ell = \text{Cov}(y_t, u_{t-\ell})\Sigma^{-1}$

**Key**: The Wold decomposition is **identifiable from data!**

# Wold: Practical Implications

**Problem:**  $\text{MA}(\infty)$  has infinitely many parameters — cannot estimate!

**Solution:**  $\text{ARMA}(p, q)$  provides a parsimonious approximation:

$$\Phi_p(L)y_t = \Theta_q(L)u_t \quad \Leftrightarrow \quad y_t = \frac{\Theta_q(L)}{\Phi_p(L)}u_t$$

The ratio  $\psi(L) = \Theta_q(L)/\Phi_p(L)$  approximates the Wold coefficients  $\{\Psi_i\}$ .

## Why This Matters

- ▶ We can **estimate** finite-parameter ARMA/VAR models
- ▶ These approximate the infinite-parameter Wold representation
- ▶ From VAR, we can also get  $\text{VAR}(\infty)$  representation:  $A(L)y_t = u_t$

# Wold vs. Structural: A Critical Warning

Wold Innovations  $\neq$  Structural Shocks!

The Wold coefficients  $\{\Psi_j\}$  are **reduced-form** objects.

**Nothing guarantees they are interesting** — they are just coefficients on reduced-form prediction errors!

**Structural VMA** (what we want):

$$y_t = \sum_{\ell=0}^{\infty} \Theta_{\ell} \varepsilon_{t-\ell}, \quad \varepsilon_t = \text{structural shocks}$$

**Wold representation** (what we estimate):

$$y_t = \sum_{j=0}^{\infty} \Psi_j u_{t-j}, \quad u_t = \text{prediction errors}$$

**The gap:** Identifying  $\Theta_{\ell}$  from  $\Psi_j$  requires **additional assumptions**!

⇒ This is the **identification problem** we'll address in Structural VARs.

# Summary

## Identification:

- ▶ Global vs. local: uniqueness in entire space vs. neighborhood
- ▶ Linear Gaussian  $\Rightarrow$  second moments matter
- ▶ Observational equivalence: same ACGF  $\Rightarrow$  indistinguishable
- ▶ Causality:  $y_t$  from current/past  $z$ 's; Invertibility:  $z_t$  from  $y$ 's

## Reduced-Form Models:

- ▶ AR( $p$ ): Stationary if roots outside UC; ACF decays geometrically
- ▶ MA( $q$ ): Always stationary; invertible if roots outside UC; ACF cuts off
- ▶ ARMA: Combines both; common factor  $\Rightarrow$  non-identification

## Wold's Theorem:

- ▶ Any stationary process = VMA( $\infty$ ) + deterministic
- ▶ Wold innovations = prediction errors (**not** structural shocks!)
- ▶ Reduced-form, identifiable; structural interpretation needs more

## **Part IV: Exercises**

# Pset 1 — Problem 3: Finite MA( $q$ ) properties

## Problem

A finite-order MA( $q$ ) process can be expressed as

$$y_t = \sum_{j=0}^q \theta_j \varepsilon_{t-j}, \quad \varepsilon_t \text{ is white noise, } \theta_0 = 1.$$

Show that all finite-order moving-average processes:

- (a) are covariance stationary;
- (b) have autocovariances that truncate at lag  $q$  (i.e.,  $\gamma_k \neq 0$  for  $k \leq q$  and  $\gamma_k = 0$  for  $k > q$ );
- (c) have absolutely summable autocovariances/autocorrelations.

**Idea:** expand products using WN orthogonality ( $\mathbb{E}[\varepsilon_t \varepsilon_s] = 0$  for  $t \neq s$ ).

# Pset 1 — Problem 3: Solution

Let  $\mathbb{E}[\varepsilon_t] = 0$ ,  $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ , and  $\mathbb{E}[\varepsilon_t \varepsilon_s] = 0$  for  $t \neq s$ .

- (a) **Covariance stationarity:**  $\mathbb{E}[y_t] = \sum_{j=0}^q \theta_j \mathbb{E}[\varepsilon_{t-j}] = 0$  (constant).

Variance:

$$\gamma_0 = \text{Var}(y_t) = \sigma_\varepsilon^2 \sum_{j=0}^q \theta_j^2 < \infty.$$

- (b) **Autocovariance truncation:** for  $k \geq 0$ ,

$$\gamma_k = \text{Cov}(y_t, y_{t-k}) = \sigma_\varepsilon^2 \sum_{j=0}^{q-k} \theta_{j+k} \theta_j,$$

because only terms with matching innovation indices survive. If  $k > q$ , the sum is empty  
 $\Rightarrow \gamma_k = 0$ .

- (c) **Absolute summability:** since  $\gamma_k = 0$  for  $|k| > q$ ,

$$\sum_{k=-\infty}^{\infty} |\gamma_k| = \sum_{k=-q}^q |\gamma_k| < \infty \quad \Rightarrow \quad \sum_{k=-\infty}^{\infty} |\rho_k| < \infty.$$

## Problem

Consider a second-order moving-average process

$$x_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}, \quad \varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2).$$

- (a) State conditions on  $(\theta_1, \theta_2)$  for the process to be **invertible**.
- (b) Characterize the family of **observationally equivalent**, non-invertible MA(2) representations.

# Pset 1 — Problem 5: Solution

Write  $x_t = \Theta(L)\varepsilon_t$  with  $\Theta(L) = 1 + \theta_1 L + \theta_2 L^2$ .

- (a) **Invertibility**  $\iff$  all roots of  $\Theta(z) = 0$  satisfy  $|z| > 1$ .

Equivalently, factor  $\Theta(L) = (1 - \lambda_1 L)(1 - \lambda_2 L)$  with  $\lambda_1 + \lambda_2 = -\theta_1$ ,  $\lambda_1 \lambda_2 = \theta_2$ . Then invertibility  $\iff |\lambda_1| < 1$  and  $|\lambda_2| < 1$  (roots outside unit circle).

- (b) **Observationally equivalent non-invertible family (root flipping):**

If  $\Theta(L) = \prod_{i=1}^2 (1 - \lambda_i L)$ , pick any subset  $S \subseteq \{1, 2\}$  and define

$$\tilde{\Theta}(L) = \prod_{i \in S} (1 - \lambda_i^{-1} L) \prod_{i \notin S} (1 - \lambda_i L),$$

i.e., replace any factor with  $|\lambda_i| < 1$  by its reciprocal (moving the corresponding root *inside* the unit circle).

With suitably rescaled white noise  $\tilde{\varepsilon}_t$  (variance scaled by  $\prod_{i \in S} \lambda_i^2$ ), the process  $x_t = \tilde{\Theta}(L)\tilde{\varepsilon}_t$  has the **same second moments** as the original MA(2), hence is observationally equivalent.

## Key takeaway

For MA models, **second moments cannot distinguish** between invertible and root-flipped non-invertible representations.

# Pset 1 — Problem 1: Nonstationary bivariate system

## Problem

Consider the non-stationary processes

$$y_{1t} = y_{1,t-1} + \varepsilon_{1t} + \theta \varepsilon_{1,t-1}, \quad y_{2t} = \phi y_{1t} + \varepsilon_{2t},$$

where  $|\theta| < 1$  and  $\phi \neq 0$ .

- Show that the vector of first differences  $[\Delta y_{1t}, \Delta y_{2t}]'$  is a **non-invertible** MA(1) process.
- Are there observationally equivalent, **invertible** representations? Why or why not?

# Pset 1 — Problem 1: Solution

First differences:

$$\Delta y_{1t} = \varepsilon_{1t} + \theta \varepsilon_{1,t-1}, \quad \Delta y_{2t} = \phi \Delta y_{1t} + (\varepsilon_{2t} - \varepsilon_{2,t-1}).$$

Define  $u_t = (\varepsilon_{1t}, \varepsilon_{2t})'$  and write

$$\Delta y_t = \underbrace{\begin{pmatrix} 1 & 0 \\ \phi & 1 \end{pmatrix}}_{A_0} u_t + \underbrace{\begin{pmatrix} \theta & 0 \\ \phi\theta & -1 \end{pmatrix}}_{A_1} u_{t-1} = (A_0 + A_1 L)u_t,$$

so it is a VMA(1).

**Non-invertibility:** check the determinant

$$\det(A_0 + A_1 z) = \det \begin{pmatrix} 1 + \theta z & 0 \\ \phi(1 + \theta z) & 1 - z \end{pmatrix} = (1 + \theta z)(1 - z),$$

which has a root at  $z = 1$  (on the unit circle)  $\Rightarrow$  fails invertibility.

- (b) No observationally equivalent *invertible* MA(1) representation within the usual class: the problematic root is exactly at  $z = 1$  and cannot be “flipped” outside the unit circle.

## Key takeaway

Non-invertibility can be **structural** (unit-circle root), not just a choice of MA parameterization.

# Pset 1 — Problem 2: Measurement error in MA(1)

## Problem

Let  $y_t$ ,  $t = 1, \dots, T$ , be a time series described by an MA(1) process:

$$y_t = \mu + \varepsilon_t + \theta \varepsilon_{t-1}, \quad \varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2)$$

Suppose that  $y_t$  is observed with **additive measurement noise**, i.e.,  $x_t = y_t + u_t$ , where  $u_t \sim \text{IID}(0, \sigma_u^2)$  and  $\mathbb{E}(\varepsilon_t u_s) = 0$  for all  $t, s$ .

## Compute:

- The mean and variance of  $x_t$ .
- The autocovariances of  $x_t$ .
- Which ARMA( $p, q$ ) process best describes  $x_t$ ?

# Pset 1 — Problem 2: Solution (a)

**(a) Mean and Variance of  $x_t$ :**

Since  $x_t = y_t + u_t = \mu + \varepsilon_t + \theta\varepsilon_{t-1} + u_t$ :

**Mean:**

$$\mathbb{E}(x_t) = \mathbb{E}(\mu + \varepsilon_t + \theta\varepsilon_{t-1} + u_t) = \mu$$

**Variance:** Since  $\varepsilon_t$ ,  $\varepsilon_{t-1}$ , and  $u_t$  are mutually uncorrelated:

$$\begin{aligned}\gamma_0 &= \text{Var}(x_t) = \text{Var}(\varepsilon_t) + \theta^2\text{Var}(\varepsilon_{t-1}) + \text{Var}(u_t) \\ &= \sigma_\varepsilon^2 + \theta^2\sigma_\varepsilon^2 + \sigma_u^2 = \boxed{(1 + \theta^2)\sigma_\varepsilon^2 + \sigma_u^2}\end{aligned}$$

# Pset 1 — Problem 2: Solution (b)

## (b) Autocovariances of $x_t$ :

For  $k \geq 1$ :

$$\gamma_k = \text{Cov}(x_t, x_{t-k}) = \text{Cov}(\varepsilon_t + \theta\varepsilon_{t-1} + u_t, \varepsilon_{t-k} + \theta\varepsilon_{t-k-1} + u_{t-k})$$

**For**  $k = 1$ : The only surviving term is  $\text{Cov}(\theta\varepsilon_{t-1}, \varepsilon_{t-1})$ :

$$\gamma_1 = \theta \cdot \sigma_\varepsilon^2 = \boxed{\theta\sigma_\varepsilon^2}$$

**For**  $k \geq 2$ : There are no overlapping terms:

$$\gamma_k = \boxed{0}, \quad k \geq 2$$

## Result

The autocovariance structure is:  $\gamma_0 = (1 + \theta^2)\sigma_\varepsilon^2 + \sigma_u^2$ ,  $\gamma_1 = \theta\sigma_\varepsilon^2$ ,  $\gamma_k = 0$  for  $k \geq 2$ .

# Pset 1 — Problem 2: Solution (c)

(c) Which ARMA( $p, q$ ) describes  $x_t$ ?

The autocovariance “cuts off” after lag 1  $\Rightarrow \text{MA(1)}$  behavior!

Therefore,  $x_t$  can be written as:

$$x_t = \mu + \eta_t + \tilde{\theta}\eta_{t-1}, \quad \eta_t \sim \text{WN}(0, \sigma_\eta^2)$$

**Matching parameters:** We need:

$$\gamma_0^{\text{MA}(1)} = (1 + \tilde{\theta}^2)\sigma_\eta^2 = (1 + \theta^2)\sigma_\varepsilon^2 + \sigma_u^2$$

$$\gamma_1^{\text{MA}(1)} = \tilde{\theta}\sigma_\eta^2 = \theta\sigma_\varepsilon^2$$

## Conclusion

$x_t$  is an **MA(1)** process, but with parameters different from the original MA(1)!  
Measurement noise “hides” the true structural parameters  $(\theta, \sigma_\varepsilon^2)$ .

# Pset 1 — Problem 4: Sum of two AR(1) processes

## Problem

Let  $\{y_t\}$  and  $\{z_t\}$ ,  $t = 1, \dots, T$ , be two stochastic processes:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t, \quad z_t = \phi_2 z_{t-1} + u_t$$

where  $\varepsilon_t \sim \text{IID}(0, \sigma_\varepsilon^2)$ ,  $u_t \sim \text{IID}(0, \sigma_u^2)$ , and  $\mathbb{E}(\varepsilon_t u_s) = 0$  for all  $t, s$ .

**Show that:**

- (a) If  $\phi_1 = \phi_2$ , then  $x_t = y_t + z_t$  is an AR(1) process.
- (b) If  $\phi_1 \neq \phi_2$ , then  $x_t = y_t + z_t$  is an ARMA(2, q) process with  $q \leq 1$ .

# Pset 1 — Problem 4: Solution (a)

(a) Case  $\phi_1 = \phi_2 = \phi$ :

Adding the two equations:

$$\begin{aligned}y_t + z_t &= \phi(y_{t-1} + z_{t-1}) + \varepsilon_t + u_t \\x_t &= \phi x_{t-1} + w_t\end{aligned}$$

where  $w_t = \varepsilon_t + u_t$ .

**Checking that  $w_t$  is white noise:**

- ▶  $\mathbb{E}(w_t) = 0$
- ▶  $\text{Var}(w_t) = \sigma_\varepsilon^2 + \sigma_u^2$
- ▶  $\text{Cov}(w_t, w_{t-k}) = 0$  for  $k \neq 0$  (since  $\varepsilon$  and  $u$  are serially independent and mutually uncorrelated)

Conclusion

$x_t$  is an **AR(1)** process with parameter  $\phi$  and innovation  $w_t \sim \text{WN}(0, \sigma_\varepsilon^2 + \sigma_u^2)$ .

# Pset 1 — Problem 4: Solution (b) Part 1

**(b) Case**  $\phi_1 \neq \phi_2$ :

**Strategy:** Write in lag-operator notation and combine.

The individual processes are:

$$(1 - \phi_1 L)y_t = \varepsilon_t, \quad (1 - \phi_2 L)z_t = u_t$$

To eliminate  $y_t$  and  $z_t$  separately, apply cross-filters:

$$(1 - \phi_2 L)(1 - \phi_1 L)y_t = (1 - \phi_2 L)\varepsilon_t$$

$$(1 - \phi_1 L)(1 - \phi_2 L)z_t = (1 - \phi_1 L)u_t$$

Summing:

$$(1 - \phi_1 L)(1 - \phi_2 L)(y_t + z_t) = (1 - \phi_2 L)\varepsilon_t + (1 - \phi_1 L)u_t$$

# Pset 1 — Problem 4: Solution (b) Part 2

Expanding:

$$[1 - (\phi_1 + \phi_2)L + \phi_1\phi_2L^2]x_t = \varepsilon_t - \phi_2\varepsilon_{t-1} + u_t - \phi_1u_{t-1}$$

**Right-hand side:** Define  $\eta_t = \varepsilon_t + u_t - \phi_2\varepsilon_{t-1} - \phi_1u_{t-1}$ .

Is this an MA(1) process in general? Compute autocovariances of  $\eta_t$ :

- ▶  $\gamma_0^\eta = (1 + \phi_2^2)\sigma_\varepsilon^2 + (1 + \phi_1^2)\sigma_u^2$
- ▶  $\gamma_1^\eta = -\phi_2\sigma_\varepsilon^2 - \phi_1\sigma_u^2$
- ▶  $\gamma_k^\eta = 0$  for  $k \geq 2$

## Conclusion

$x_t$  follows an **ARMA(2,1)** process with:

- ▶ AR part:  $(1 - \phi_1L)(1 - \phi_2L)$  with roots  $1/\phi_1$  and  $1/\phi_2$
- ▶ MA part: order  $q \leq 1$  (it can be 0 if  $\gamma_1^\eta = 0$ )

# Pset 1 — Problem 6: Sum of independent AR(1) processes

## Problem

Consider two independent AR(1) processes  $x_t$  and  $y_t$  defined by:

$$(1 - \rho_1 L)x_t = u_t, |\rho_1| < 1 \quad \text{and} \quad (1 - \rho_2 L)y_t = v_t, |\rho_2| < 1$$

where  $\mathbb{E}(u_t) = \mathbb{E}(v_t) = 0$ ,  $\mathbb{E}(u_t v_\tau) = 0$  for all  $t, \tau$ ,  
 $\mathbb{E}(u_t u_\tau) = \sigma_u^2$  if  $t = \tau$  (0 otherwise),  $\mathbb{E}(v_t v_\tau) = \sigma_v^2$  if  $t = \tau$ .

Let  $z_t = x_t + y_t$ .

## Tasks:

- Compute  $\mathbb{E}(x_{t+1}|\mathcal{F}_t)$ ,  $\mathbb{E}(y_{t+1}|\mathcal{F}_t)$ ,  $\mathbb{E}(z_{t+1}|\mathcal{F}_t)$ , and the forecast error variance  $V_D$ .
- Show that  $z_t$  is ARMA(2,1) and find  $\theta$ .
- Conditions for  $z_t$  to be AR(2)? Show that if  $\rho_1, \rho_2$  have the same sign, then it cannot be AR(2).
- Show that  $\theta \neq \pm 1$

# Pset 1 — Problem 6: Solution (a)

## (a) Conditional forecasts:

Since  $x_t$  and  $y_t$  are AR(1):

$$\mathbb{E}(x_{t+1}|\mathcal{F}_t) = \rho_1 x_t, \quad \mathbb{E}(y_{t+1}|\mathcal{F}_t) = \rho_2 y_t$$

Para  $z_t = x_t + y_t$ :

$$\mathbb{E}(z_{t+1}|\mathcal{F}_t) = \rho_1 x_t + \rho_2 y_t$$

**Forecast error**  $e_{t+1} = z_{t+1} - \mathbb{E}(z_{t+1}|\mathcal{F}_t)$ :

$$e_{t+1} = (x_{t+1} + y_{t+1}) - (\rho_1 x_t + \rho_2 y_t) = u_{t+1} + v_{t+1}$$

**Error variance**  $V_D$ :

$$V_D = \text{Var}(e_{t+1}) = \text{Var}(u_{t+1}) + \text{Var}(v_{t+1}) = \boxed{\sigma_u^2 + \sigma_v^2}$$

# Pset 1 — Problem 6: Solution (b) Part 1

**(b) Showing that  $z_t$  is ARMA(2,1):**

From Exercise 2, we know that when  $\rho_1 \neq \rho_2$ :

$$[1 - (\rho_1 + \rho_2)L + \rho_1\rho_2L^2]z_t = (1 + \theta L)\varepsilon_t$$

where  $\varepsilon_t$  is white noise with variance  $\sigma_\varepsilon^2$ .

**Identification via autocovariances:** For an MA(1),  $\eta_t = (1 + \theta L)\varepsilon_t$ :

$$\gamma_0^\eta = (1 + \theta^2)\sigma_\varepsilon^2, \quad \gamma_1^\eta = \theta\sigma_\varepsilon^2$$

From the previous exercise, the right-hand side has:

$$\gamma_0^\eta = (1 + \rho_2^2)\sigma_u^2 + (1 + \rho_1^2)\sigma_v^2$$

$$\gamma_1^\eta = -\rho_2\sigma_u^2 - \rho_1\sigma_v^2$$

# Pset 1 — Problem 6: Solution (b) Part 2

**Solving for  $\theta$ :**

The lag-1 autocorrelation of an MA(1) is:

$$\rho_1^{\text{MA}} = \frac{\gamma_1}{\gamma_0} = \frac{\theta}{1 + \theta^2}$$

Equating:

$$\frac{\theta}{1 + \theta^2} = \frac{-\rho_2 \sigma_u^2 - \rho_1 \sigma_v^2}{(1 + \rho_2^2) \sigma_u^2 + (1 + \rho_1^2) \sigma_v^2}$$

This is a quadratic equation in  $\theta$ :

$$\theta^2 \cdot (\text{numerador}) - \theta \cdot (\text{denominador}) + (\text{numerador}) = 0$$

Two roots!

There are two solutions,  $\theta$  and  $1/\theta$ , corresponding to observationally equivalent representations (invertible vs. non-invertible).

# Pset 1 — Problem 6: Solution (c)

(c) Conditions for  $z_t$  to be AR(2):

$z_t$  is AR(2) if the MA part disappears, i.e.,  $\theta = 0$ , which requires  $\gamma_1^\eta = 0$ :

$$-\rho_2\sigma_u^2 - \rho_1\sigma_v^2 = 0 \quad \Rightarrow \quad \boxed{\rho_1\sigma_v^2 = -\rho_2\sigma_u^2}$$

If  $\rho_1$  and  $\rho_2$  have the same sign:

The left-hand side  $\rho_1\sigma_v^2$  and the right-hand side  $-\rho_2\sigma_u^2$  have opposite signs (since  $\sigma_u^2, \sigma_v^2 > 0$ ).

## Impossibility

If  $\text{sgn}(\rho_1) = \text{sgn}(\rho_2)$ , then the condition  $\rho_1\sigma_v^2 = -\rho_2\sigma_u^2$  can **never** be satisfied.

Therefore,  $z_t$  **cannot be AR(2)** when  $\rho_1$  and  $\rho_2$  have the same sign.

# Pset 1 — Problem 6: Solution (d)

(d) Showing that  $\theta \neq \pm 1$ :

Suppose  $\theta = 1$ . Then  $\rho_1^{\text{MA}} = \frac{1}{1+1} = \frac{1}{2}$ .

For this to occur, we would need:

$$\frac{-\rho_2\sigma_u^2 - \rho_1\sigma_v^2}{(1 + \rho_2^2)\sigma_u^2 + (1 + \rho_1^2)\sigma_v^2} = \frac{1}{2}$$

Since  $|\rho_1|, |\rho_2| < 1$ , the denominator  $(1 + \rho_2^2)\sigma_u^2 + (1 + \rho_1^2)\sigma_v^2 > 2\sigma_u^2 + 2\sigma_v^2$  (using  $\rho^2 < 1$ ).

The numerator has absolute value  $|\rho_2\sigma_u^2 + \rho_1\sigma_v^2| < \sigma_u^2 + \sigma_v^2$ .

Therefore, the absolute value of the ratio is:

$$|\rho_1^{\text{MA}}| < \frac{\sigma_u^2 + \sigma_v^2}{2(\sigma_u^2 + \sigma_v^2)} = \frac{1}{2}$$

## Conclusion

Since  $|\rho_1^{\text{MA}}| < 1/2$ , and  $\theta = +1$  would imply  $|\rho_1^{\text{MA}}| = 1/2$ , we have  $\boxed{\theta \neq +1}$

# Pset 1 — Problem 6: Solution (e)

(e) Forecasting using the ARMA(2,1):

Do modelo  $[1 - (\rho_1 + \rho_2)L + \rho_1\rho_2L^2]z_t = (1 + \theta L)\varepsilon_t$ :

$$z_t = (\rho_1 + \rho_2)z_{t-1} - \rho_1\rho_2z_{t-2} + \varepsilon_t + \theta\varepsilon_{t-1}$$

**Forecast:**

$$\mathbb{E}(z_{T+1}|\mathcal{F}_T) = (\rho_1 + \rho_2)z_T - \rho_1\rho_2z_{T-1} + \theta\varepsilon_T$$

**Error:**  $e_{T+1} = z_{T+1} - \mathbb{E}(z_{T+1}|\mathcal{F}_T) = \varepsilon_{T+1}$

**Error variance**  $V_A$ :

$$V_A = \text{Var}(\varepsilon_{T+1}) = \sigma_\varepsilon^2$$

**Comparing**  $V_A - V_D$ :  $V_A = \sigma_\varepsilon^2$  and  $V_D = \sigma_u^2 + \sigma_v^2$ .

Since  $(1 + \theta^2)\sigma_\varepsilon^2 = (1 + \rho_2^2)\sigma_u^2 + (1 + \rho_1^2)\sigma_v^2$ , and  $\theta^2 < 1$ ,  $\rho_i^2 < 1$ , we have  $V_A \approx V_D$  (depending on the specific parameter values).

## Partial Adjustment Model

The desired level  $y_t^*$  is related to  $x_t$  by:  $y_t^* = \beta x_t$ .

$x_t$  is covariance-stationary. A fraction  $\gamma$  of the disequilibrium is removed through partial adjustment:

$$y_t - y_{t-1} = \gamma(y_t^* - y_{t-1}) + e_t, \quad e_t \sim N(0, \sigma_e^2), \quad 0 < \gamma < 1$$

## Compare with the adaptive expectations model:

$$y_t = \beta x_{t+1}^e + v_t, \quad v_t \sim N(0, \sigma_v^2)$$

$$x_{t+1}^e - x_t^e = \lambda(x_t - x_t^e), \quad 0 < \lambda < 1$$

**Tasks:** (a) Compare the models. (b) Show that OLS of  $y_t$  on  $y_{t-1}$  and  $x_t$  is consistent for the model (1)-(2), but inconsistent for (3)-(4).

# Pset 1 — Problem 7: Solution (a)

## (a) Comparing the two models:

**Partial Adjustment Model:** Substituting  $y_t^* = \beta x_t$ :

$$y_t - y_{t-1} = \gamma(\beta x_t - y_{t-1}) + e_t$$

$$y_t = \gamma\beta x_t + (1 - \gamma)y_{t-1} + e_t$$

**Adaptive Expectations Model:** Solving for  $x_{t+1}^e$ :

$$x_{t+1}^e = \lambda x_t + (1 - \lambda)x_t^e = \lambda \sum_{j=0}^{\infty} (1 - \lambda)^j x_{t-j}$$

Substituting into  $y_t = \beta x_{t+1}^e + v_t$ :

$$y_t = \beta \lambda \sum_{j=0}^{\infty} (1 - \lambda)^j x_{t-j} + v_t$$

Using the Koyck transformation:  $y_t = \beta \lambda x_t + (1 - \lambda)y_{t-1} + v_t - (1 - \lambda)v_{t-1}$

# Pset 1 — Problem 7: Solution (a) Cont.

Reduced-form representation of both models:

	Partial Adjustment	Adaptive Expectations
Form	$y_t = \alpha x_t + \phi y_{t-1} + e_t$	$y_t = \alpha x_t + \phi y_{t-1} + \eta_t$
$x_t$ coeff.	$\alpha = \gamma\beta$	$\alpha = \beta\lambda$
$y_{t-1}$ coeff.	$\phi = 1 - \gamma$	$\phi = 1 - \lambda$
Error	$e_t$ (WN)	$\eta_t = v_t - (1 - \lambda)v_{t-1}$ (MA(1)!)

## Key Difference

- ▶ In partial adjustment:  $e_t$  is white noise  $\Rightarrow$  exogenous
- ▶ In adaptive expectations:  $\eta_t$  is MA(1)  $\Rightarrow$  correlated with  $y_{t-1}$ !

# Pset 1 — Problem 7: Solution (b)

## (b) OLS consistency:

Consider the regression:  $y_t = \alpha x_t + \phi y_{t-1} + u_t$

**Partial Adjustment Model:**  $u_t = e_t$  is white noise.

- ▶  $\text{Cov}(y_{t-1}, e_t) = 0$  (error is uncorrelated with past variables)
- ▶  $\text{Cov}(x_t, e_t) = 0$  (exogeneidade estrita)
- ▶  $\Rightarrow$  OLS is **consistent** ✓

**Adaptive Expectations Model:**  $u_t = \eta_t = v_t - (1 - \lambda)v_{t-1}$

- ▶  $y_{t-1}$  depends on  $v_{t-1}$  (through the recursive equation)
- ▶  $\text{Cov}(y_{t-1}, \eta_t) = \text{Cov}(y_{t-1}, -(1 - \lambda)v_{t-1}) \neq 0$
- ▶  $\Rightarrow$  OLS is **inconsistent** ✗

## Intuition

The regressor  $y_{t-1}$  is correlated with the error  $\eta_t$  due to the MA(1) component.

# Pset 1 — Problem 7: Solution (c)

## (c) How to check which model is appropriate?

**Strategy:** Test for autocorrelation in the residuals.

1. Estimate by OLS:  $y_t = \hat{\alpha}x_t + \hat{\phi}y_{t-1} + \hat{u}_t$
2. Examine the residuals  $\hat{u}_t$ :
  - ▶ If  $\hat{u}_t$  is white noise  $\Rightarrow$  Partial Adjustment Model
  - ▶ If  $\hat{u}_t$  shows autocorrelation (MA(1))  $\Rightarrow$  Adaptive Expectations
3. Formal tests:
  - ▶ Durbin–Watson test (for AR(1) in the residuals)
  - ▶ Breusch–Godfrey test (for higher-order autocorrelation)
  - ▶ Ljung–Box test on the residual ACF

### Caution

If the correct model is adaptive expectations, use appropriate estimation methods (GMM, NLLS, or MLE) instead of OLS.

# Pset 1 — Problem 8: Vector model (inflation, output gap, policy)

## Macroeconomic Model

$$(1) : \pi_t = \lambda y_t + \pi_t^e + u_{1t}, \quad 0 < \lambda < 1$$

$$(2) : y_t = \gamma(i_{t-1} - \pi_t^e) + u_{2,t}, \quad -1 < \gamma < 0$$

$$(3) : \pi_t^e = \pi_{t-1}$$

$$(4) : i_t = i^* + \rho(\pi_t - \pi^*), \quad \rho \geq 0$$

where  $\pi_t$  = inflation,  $y_t$  = output gap,  $i_t$  = nominal interest rate,  
 $(u_{1t}, u_{2t})' \sim N(0, \Sigma)$ .

**Interpretation:** (1) Phillips curve, (2) IS, (3) adaptive expectations, (4) Taylor rule.

# Pset 1 — Problem 8: Tasks (cont.)

## Tasks:

- (a) Show that  $\pi_t$  follows an ARMA( $p, q$ ). Relate the coefficients to the structural parameters.
- (b) Consider  $\rho = 0, 1, 2$ . What can be said about inflation and monetary policy? Simulate 200 observations in MATLAB.
- (c) Compute  $\hat{\pi}_{t+j|t} = \mathbb{E}[\pi_{t+j} | \mathcal{F}_t]$  and  $\text{Var}(\pi_{t+j} - \hat{\pi}_{t+j|t})$ .
- (d) Let  $\mathbf{x}_t = (\pi_t, y_t)'$ . Find  $\mathbb{D}(\mathbf{x}_t | \mathbf{x}_{t-1})$  and  $\mathbb{E}(\pi_t | y_t, \mathbf{x}_{t-1})$ .
- (e) What would change if  $\pi_t^e = \alpha\pi_{t-1} + (1 - \alpha)\pi_{t-1}^e$ ?

# Pset 1 — Problem 8: Solution (a) Part 1

(a) Deriving the ARMA process for  $\pi_t$ :

**Step 1:** Substitute (3) and (4) into (2):

$$\begin{aligned}y_t &= \gamma(i_{t-1} - \pi_{t-1}) + u_{2t} \\&= \gamma(i^* + \rho(\pi_{t-1} - \pi^*) - \pi_{t-1}) + u_{2t} \\&= \gamma(i^* - \rho\pi^*) + \gamma(\rho - 1)\pi_{t-1} + u_{2t}\end{aligned}$$

**Step 2:** Substitute into (1):

$$\begin{aligned}\pi_t &= \lambda[\gamma(i^* - \rho\pi^*) + \gamma(\rho - 1)\pi_{t-1} + u_{2t}] + \pi_{t-1} + u_{1t} \\&= c + [1 + \lambda\gamma(\rho - 1)]\pi_{t-1} + u_{1t} + \lambda u_{2t}\end{aligned}$$

where  $c = \lambda\gamma(i^* - \rho\pi^*)$  is a constant.

# Pset 1 — Problem 8: Solution (a) Part 2

**Result:**  $\pi_t$  follows an **AR(1)** process:

$$\boxed{\pi_t = c + \phi\pi_{t-1} + \varepsilon_t}$$

where:

- ▶  $\phi = 1 + \lambda\gamma(\rho - 1)$
- ▶  $\varepsilon_t = u_{1t} + \lambda u_{2t}$  (a linear combination of the shocks)
- ▶  $\text{Var}(\varepsilon_t) = \sigma_{11}^2 + 2\lambda\sigma_{12} + \lambda^2\sigma_{22}^2$

## Link to Structural Parameters

- ▶  $\phi < 1$  (stationarity)  $\Leftrightarrow \lambda\gamma(\rho - 1) < 0$
- ▶ Since  $\lambda > 0$  and  $\gamma < 0$ :  $\phi < 1 \Leftrightarrow \rho > 1$
- ▶ This is the **Taylor principle**:  $\rho > 1$  stabilizes inflation!

# Pset 1 — Problem 8: Solution (b)

## (b) Analysis for different values of $\rho$ :

Suppose  $\lambda = 0.5$ ,  $\gamma = -0.5$ . Then  $\phi = 1 + 0.5 \cdot (-0.5) \cdot (\rho - 1) = 1 - 0.25(\rho - 1)$ .

$\rho$	$\phi$	Stationary?	Interpretation
0	1.25	No	Explosive (violates stability/unit-root condition)
1	1.00	No	Random walk
2	0.75	Yes	Stationary inflation

### Economic Interpretation

- ▶  $\rho = 0$ : the central bank does not respond to inflation  $\Rightarrow$  explosive inflation
- ▶  $\rho = 1$ : the central bank responds one-for-one  $\Rightarrow$  constant real rate, inflation is  $I(1)$
- ▶  $\rho > 1$ : the central bank responds more than one-for-one  $\Rightarrow$  stationary inflation

# Pset 1 — Problem 8: Solution (c)

## (c) Forecasts and forecast error variance:

For the AR(1)  $\pi_t = c + \phi\pi_{t-1} + \varepsilon_t$  (assuming  $|\phi| < 1$ ):

**$j$ -step-ahead forecast:**

$$\hat{\pi}_{t+j|t} = \mu(1 - \phi^j) + \phi^j \pi_t$$

where  $\mu = c/(1 - \phi)$  is the unconditional mean.

**Forecast error variance:**

$$\text{Var}(\pi_{t+j} - \hat{\pi}_{t+j|t}) = \sigma_\varepsilon^2 \sum_{i=0}^{j-1} \phi^{2i} = \sigma_\varepsilon^2 \cdot \frac{1 - \phi^{2j}}{1 - \phi^2}$$

**Limit:** As  $j \rightarrow \infty$ :

$$\lim_{j \rightarrow \infty} \text{Var}(\pi_{t+j} - \hat{\pi}_{t+j|t}) = \frac{\sigma_\varepsilon^2}{1 - \phi^2} = \text{Var}(\pi_t)$$

# Pset 1 — Problem 8: Solution (d)

(d) Conditional distribution of  $\mathbf{x}_t = (\pi_t, y_t)'$ :

From the model equations:

$$\begin{aligned}\pi_t &= c_\pi + \phi\pi_{t-1} + \varepsilon_t \\ y_t &= c_y + \gamma(\rho - 1)\pi_{t-1} + u_{2t}\end{aligned}$$

Therefore,  $\mathbf{x}_t | \mathbf{x}_{t-1} \sim N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma})$ , where:

$$\boldsymbol{\mu}_t = \begin{pmatrix} c_\pi + \phi\pi_{t-1} \\ c_y + \gamma(\rho - 1)\pi_{t-1} \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_\varepsilon^2 & \lambda\sigma_{22}^2 + \sigma_{12} \\ \lambda\sigma_{22}^2 + \sigma_{12} & \sigma_{22}^2 \end{pmatrix}$$

**Conditional expectation**  $\mathbb{E}(\pi_t | y_t, \mathbf{x}_{t-1})$ :

Using the conditional normal formula:

$$\mathbb{E}(\pi_t | y_t, \mathbf{x}_{t-1}) = \mathbb{E}(\pi_t | \mathbf{x}_{t-1}) + \frac{\text{Cov}(\pi_t, y_t)}{\text{Var}(y_t)} [y_t - \mathbb{E}(y_t | \mathbf{x}_{t-1})]$$

## Pset 1 — Problem 8: Solution (e)

(e) **Generalized adaptive expectations:**  $\pi_t^e = \alpha\pi_{t-1} + (1 - \alpha)\pi_{t-1}^e$

Solving recursively:

$$\pi_t^e = \alpha \sum_{j=0}^{\infty} (1 - \alpha)^j \pi_{t-1-j}$$

**Effect on the model:** inflation now depends on the entire past history, not only on  $\pi_{t-1}$ .

Using the Koyck transformation:

$$\pi_t^e = \alpha\pi_{t-1} + (1 - \alpha)\pi_{t-1}^e$$

The process for  $\pi_t$  becomes an **ARMA(2,1)** instead of AR(1):

$$\pi_t = c + \phi_1\pi_{t-1} + \phi_2\pi_{t-2} + \varepsilon_t + \theta\varepsilon_{t-1}$$

where the coefficients depend on  $\alpha, \lambda, \gamma, \rho$  in a more complex way.

# Exercise Session: Summary

## Key Skills Practiced:

1. Checking stationarity conditions (roots outside unit circle)
2. Computing autocovariances and ACF for ARMA processes
3. Converting between AR and MA representations
4. Identifying common factors and parameter redundancy
5. Working with vector processes

## For Next Session

Review these exercises and attempt additional problems from the appendix for extra practice.

# **Appendix**

Detailed Material and Extended Examples

# Checking Local Identification: The Information Matrix

**Idea:** Use the Kullback-Leibler divergence between distributions:

$$\Delta_{KL}(\theta \mid \theta_0) = - \int \log \left( \frac{\mathbb{D}(\mathbf{Y}^T \mid \theta)}{\mathbb{D}(\mathbf{Y}^T \mid \theta_0)} \right) \mathbb{D}(\mathbf{Y}^T \mid \theta_0) d\mathbf{Y}^T$$

By Jensen's inequality:  $\Delta_{KL}(\theta \mid \theta_0) \geq 0$ , with equality iff distributions are equal.

**Second-order condition:** The Hessian at  $\theta_0$  is the **Fisher Information Matrix**:

$$\mathcal{I}(\theta_0) = -\mathbb{E}_{\theta_0} [\nabla_{\theta^2} \log \mathbb{D}(\mathbf{Y}^T \mid \theta_0)]$$

Theorem 17 (Local Identification Condition)

If  $\mathcal{I}(\theta_0)$  is **positive definite** (full rank), then  $\theta_0$  is locally identified.

# AR(1) Inflation Example: Figures

**Figure 1:** Sensitivity of  $\phi_1$  to Taylor Rule coefficient  $\rho$

*[Figure: param\_combination.png from metrics bible]*

Under the simple monetary policy model with  $\lambda = 0.5$ ,  $\gamma = -0.5$ :

- ▶ Stability requires  $1 < \rho < 9$
- ▶ Taylor principle ( $\rho > 1$ ) ensures stationary inflation

← Back

# AR(1) Inflation: ACF Patterns

**Figure 2:** ACF under different Taylor Rule calibrations

[Figure: *acf\_pi\_1.png, acf\_pi\_2.png* from metrics bible]

**Key observations:**

- ▶  $\rho \in \{2, 3, 4\}$ : Persistent, non-oscillatory decay
- ▶  $\rho \in \{6, 7, 8\}$ : Faster decay, potentially oscillating

## A3: Detailed Proofs

# Proof: MA( $q$ ) Stationarity

**Claim:** All finite-order MA( $q$ ) processes are covariance-stationary.

**Proof:**

**Mean:**

$$\mathbb{E}[y_t] = \mathbb{E} \left[ u_t + \sum_{j=1}^q \theta_j u_{t-j} + \theta_0 \right] = \theta_0$$

**Autocovariance:**

$$\begin{aligned}\Gamma_k &= \mathbb{E}[(y_t - \mu)(y_{t-k} - \mu)] \\ &= \mathbb{E} \left[ \left( \sum_{j=0}^q \theta_j u_{t-j} \right) \left( \sum_{m=0}^q \theta_m u_{t-k-m} \right) \right]\end{aligned}$$

Since  $\mathbb{E}[u_s u_r] = \sigma^2 \mathbf{1}_{s=r}$ :

$$\Gamma_k = \begin{cases} \left( \theta_k + \sum_{j=1}^{q-k} \theta_j \theta_{j+k} \right) \sigma^2 & k \leq q \\ 0 & k > q \end{cases}$$

## Proof: Invertibility Condition for MA( $q$ )

**Claim:** MA( $q$ ) is invertible iff roots of  $\theta_q(z) = 0$  lie outside unit circle.

**Proof:** Start with  $y_t = \theta_0 + \theta_q(L)u_t$ . We need:

$$u_t = \theta_q(L)^{-1}(y_t - \theta_0)$$

Factor the polynomial:  $\theta_q(L) = (1 - \lambda_1 L)(1 - \lambda_2 L) \cdots (1 - \lambda_q L)$   
where  $\lambda_i$  are reciprocals of the roots.

For each factor:  $(1 - \lambda_i L)^{-1} = \sum_{k=0}^{\infty} (\lambda_i L)^k$

This converges iff  $|\lambda_i| < 1$ , i.e., roots  $|1/\lambda_i| > 1$  (outside UC).

If convergent:

$$\theta_q(L)^{-1} = \sum_{j=0}^{\infty} \psi_j L^j \quad \Rightarrow \quad u_t = \sum_{j=0}^{\infty} \psi_j (y_{t-j} - \theta_0) \in \text{span}\{y_t, y_{t-1}, \dots\}$$

□

# MA( $\infty$ ) Representation of AR( $p$ )

**Claim:** A stationary AR( $p$ ) can be written as MA( $\infty$ ).

**Proof** (AR(1) case): Let  $y_t = \phi_1 y_{t-1} + u_t$  with  $|\phi_1| < 1$ .

Iterating backward (assuming process started infinitely far in the past):

$$\begin{aligned}y_t &= \phi_1 y_{t-1} + u_t \\&= \phi_1(\phi_1 y_{t-2} + u_{t-1}) + u_t \\&= \dots \\&= \sum_{i=0}^{\infty} \phi_1^i u_{t-i} + \lim_{n \rightarrow \infty} \phi_1^n y_{t-n}\end{aligned}$$

Since  $|\phi_1| < 1$ :  $\phi_1^n \rightarrow 0$  and  $\sum |\phi_1|^i < \infty$ .

Therefore:  $y_t = \sum_{i=0}^{\infty} \psi_i u_{t-i}$  with  $\psi_i = \phi_1^i$ .

**General case:** Use companion form and iterate.  $\square$

## A4: Extended Examples

## Example: Observational Equivalence in MA(2)

Consider MA(2):  $y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}$

Factor:  $\theta_2(z) = 1 + \theta_1 z + \theta_2 z^2 = (1 - \lambda_1 z)(1 - \lambda_2 z)$

where  $\theta_1 = -(\lambda_1 + \lambda_2)$ ,  $\theta_2 = \lambda_1 \lambda_2$

**Invertible representation:** Roots outside UC  $\Rightarrow |\lambda_i| < 1$

**Non-invertible alternative:** “Flip” each root:

$$\lambda_i^* = 1/\lambda_i \quad \Rightarrow \quad \theta_1^* = \theta_1/\theta_2, \quad \theta_2^* = 1/\theta_2$$

**Match ACGF:** Scale variance by  $\sigma^{*2} = \sigma^2 \theta_2^2$

**Result:** These have **identical** autocovariance structure but different impulse responses!

## Example: NK Model Invertibility (Wolf, 2022)

**Three-equation NK model** with demand, supply, and monetary shocks:

$$\begin{pmatrix} y_t \\ \pi_t \\ i_t \end{pmatrix} = \frac{1}{1 + \phi_\pi \kappa} \begin{pmatrix} \sigma^d & \phi_\pi \sigma^s & -\sigma^m \\ \kappa \sigma^d & -\sigma^s & -\kappa \sigma^m \\ \phi_\pi \kappa \sigma^d & -\phi_\pi \sigma^s & \sigma^m \end{pmatrix} \begin{pmatrix} \varepsilon_t^d \\ \varepsilon_t^s \\ \varepsilon_t^m \end{pmatrix}$$

**With all 3 observables:** Impact matrix  $\Theta$  is invertible for standard parameters.

**With only  $y_t$  and  $\pi_t$ :**

- ▶ Can back out  $\varepsilon_t^s$  (supply shock)
- ▶ **Cannot** disentangle  $\varepsilon_t^d$  and  $\varepsilon_t^m$ !

### Lesson

Even in well-specified models, fewer observables than shocks  $\Rightarrow$  non-invertibility!

## A5: Reference Material

# Key Formulas: Quick Reference

## AR( $p$ ):

- ▶ Process:  $\Phi_p(L)y_t = \phi_0 + u_t$
- ▶ Stationarity: Roots of  $\Phi_p(z) = 0$  outside UC
- ▶ Mean:  $\mu = \phi_0 / (1 - \sum \phi_i)$
- ▶ Yule-Walker:  $\gamma_k = \sum_{i=1}^p \phi_i \gamma_{k-i}$

## MA( $q$ ):

- ▶ Process:  $y_t = \theta_0 + \Theta_q(L)u_t$
- ▶ Always stationary
- ▶ Invertibility: Roots of  $\Theta_q(z) = 0$  outside UC
- ▶  $\gamma_k = 0$  for  $k > q$  (cutoff)

## ARMA( $p, q$ ):

- ▶ Process:  $\Phi_p(L)y_t = \phi_0 + \Theta_q(L)u_t$
- ▶ ACGF:  $\Gamma(z) = \frac{\Theta_q(z)\Theta_q(z^{-1})}{\Phi_p(z)\Phi_p(z^{-1})}\sigma^2$

## References

- ▶ Fernández-Villaverde, Rubio-Ramírez, Schorfheide (2016). “Solution and Estimation Methods for DSGE Models.” *Handbook of Macroeconomics*
- ▶ Hamilton (1994). *Time Series Analysis*. Princeton University Press
- ▶ Wolf (2022). “Linear Time Series” (MIT 14.461 Lecture Notes)
- ▶ Komunjer, Ng (2011). “Dynamic Identification of DSGE Models.” *Econometrica*
- ▶ Lecture Notes in Econometrics (Lincoln, 2025). Chapter 7: Time Series
- ▶ Lecture Notes in Macroeconomics (Lincoln, 2025). Chapter 21: Identification