Problem Set 3 - Proposed Solutions

Taric Latif Padovani

November 1, 2025

Problem 1.

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \cdots + \phi_p(y_{t-p} - \mu) + \varepsilon_t, \qquad \varepsilon_t \overset{i.i.d.}{\sim} (0, \sigma^2), \qquad \mathcal{I}_t = \{y_t, y_{t-1}, \ldots\}.$$

(a) Companion form. Let

$$Y_t \equiv (\, y_t - \mu,\; y_{t-1} - \mu,\; \ldots,\; y_{t-p+1} - \mu \,)'.$$

Define the $p \times p$ *companion matrix*

$$A \equiv \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix}, \qquad U_t \equiv (\varepsilon_t, 0, \ldots, 0)'.$$

Then the AR(p) can be written as

$$Y_t = AY_{t-1} + U_t, \qquad \forall t,$$

because the first row reproduces the AR(p) equation and the remaining rows simply shift lags down by one position. Moreover,

$$\Omega \;\equiv\; \mathbb{E}[U_t U_t'] \;=\; \operatorname{diag}(\sigma^2, 0, \ldots, 0),$$

i.e., all entries are zero except $(1,1) = \sigma^2$.

(b) Conditional mean. Iterating the state equation,

$$Y_{t+h} = A^h Y_t + \sum_{j=0}^{h-1} A^j U_{t+h-j}.$$

Since U_{t+h-j} is mean-zero and independent of $\mathcal{I}_t$ for $j \geq 0$,

$$\mathbb{E}[Y_{t+h} \mid \mathcal{I}_t] = A^h Y_t.$$

(c) Conditional variance.

From the state-space representation in (b),

$$Y_{t+h} = A^h Y_t + \sum_{j=0}^{h-1} A^j U_{t+h-j}.$$

Taking the conditional variance,

$$\operatorname{Var}(Y_{t+h} \mid \mathcal{I}_t) = \operatorname{Var}\left(\sum_{j=0}^{h-1} A^j U_{t+h-j} \right).$$

Write $Z_j \equiv A^j U_{t+h-j}$. Using U_t white and uncorrelated over time,

$$\operatorname{Cov}(Z_i, Z_j) = A^i \, \mathbb{E}\big[U_{t+h-i} U'_{t+h-j}\big] \, (A')^j = \begin{cases} A^j \, \Omega \, (A')^j, & i = j, \\ 0, & i \neq j. \end{cases}$$

Hence the cross terms vanish and we obtain:

$$\boxed{\Sigma_h \;\equiv\; \operatorname{Var}(Y_{t+h} \mid \mathcal{I}_t) = \sum_{j=0}^{h-1} A^j \, \Omega \, (A')^j} \qquad (\Sigma_0 = 0).$$

(d) Eigenvalues of A. Consider $\det(zI - A)$ (the characteristic polynomial up to a sign):

$$zI - A = \begin{pmatrix} z - \phi_1 & -\phi_2 & -\phi_3 & \cdots & -\phi_p \\ -1 & z & 0 & \cdots & 0 \\ 0 & -1 & z & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -1 & z \end{pmatrix}.$$

Expanding along the first row, the minor of entry $(1,1)$ is lower triangular with diagonal $z, \ldots, z$, hence has determinant z^{p-1}. For $j \geq 2$, the minor of entry $(1, j)$ can be made triangular by column operations (which do not change the determinant); its determinant is $(-1)^{1+j} z^{p-j}$. Therefore,

$$\det(zI - A) = (z - \phi_1) z^{p-1} - \phi_2 z^{p-2} - \cdots - \phi_p = z^p - \phi_1 z^{p-1} - \cdots - \phi_p.$$

Since $\det(A - zI) = (-1)^p \det(zI - A)$, the characteristic polynomial of A is

$$\Phi(z) = (-1)^p \big(z^p - \phi_1 z^{p-1} - \phi_2 z^{p-2} - \cdots - \phi_p \big),$$

so the eigenvalues of A are precisely the roots of the AR polynomial $z^p - \phi_1 z^{p-1} - \cdots - \phi_p = 0$.

(e) Stationarity $\Rightarrow$ eigenvalues inside the unit circle. From item (d) the eigenvalues of the companion matrix A are the roots of

$$\psi(z) \equiv z^p - \phi_1 z^{p-1} - \cdots - \phi_p.$$

Write $\psi(z) = z^p \Phi(1/z)$ with the AR polynomial $\Phi(u) = 1 - \phi_1 u - \cdots - \phi_p u^p$. Covariance stationarity of an AR(p) requires that all roots of $\Phi(u) = 0$ lie *outside* the unit circle, i.e. $|u_i| > 1$. If u_i is such a root, then $\lambda_i = 1/u_i$ is a root of $\psi(z)$, hence an eigenvalue of A, and $|\lambda_i| = |1/u_i| < 1$. Therefore, under stationarity, all eigenvalues of A are strictly smaller than one in modulus.

(f) Long–horizon conditional moments.
From (b)–(c),

$$\mathbb{E}[Y_{t+h} \mid \mathcal{I}_t] = A^h Y_t, \qquad \Sigma_h \equiv \mathrm{Var}(Y_{t+h} \mid \mathcal{I}_t) = \sum_{j=0}^{h-1} A^j \Omega (A')^j.$$

Assume distinct eigenvalues (a similar reasoning would apply with the Jordan Decomposition), A is diagonalizable: $A = P\Lambda P^{-1}$ with $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$. By stationarity (item e)), $|\lambda_i| < 1$ for all i.

Mean. $A^h = P\Lambda^h P^{-1} \to 0$, hence

$$\boxed{\lim_{h\to\infty} \mathbb{E}[Y_{t+h} \mid \mathcal{I}_t] = 0}.$$

Variance. Let $B \equiv P^{-1}\Omega(P^{-1})'$. Then

$$\Sigma_h = P \left(\sum_{j=0}^{h-1} \Lambda^j B (\Lambda')^j \right) P', \quad \left[\Lambda^j B (\Lambda')^j\right]_{ik} = (\lambda_i \lambda_k)^j B_{ik}.$$

Thus, for each (i, k),

$$\sum_{j=0}^{h-1} (\lambda_i \lambda_k)^j B_{ik} = \frac{1 - (\lambda_i \lambda_k)^h}{1 - \lambda_i \lambda_k} B_{ik} \longrightarrow \frac{B_{ik}}{1 - \lambda_i \lambda_k}.$$

Therefore

$$\boxed{\Sigma_\infty = P \left[\frac{B_{ik}}{1 - \lambda_i \lambda_k} \right] P'}.$$

Intuition: as the horizon grows, the influence of today's state Y_t dies out because powers of A shrink (all eigenvalues are < 1 in modulus). The forecast mean therefore reverts to the unconditional mean (zero for Y_t, i.e., μ for y_t), while the forecast variance accumulates the variances of future shocks and converges to the unconditional covariance matrix of Y_t.

Problem 2.

$$y_t = \sum_{j=0}^{\infty} \psi_j \, \varepsilon_{t-j}, \qquad \psi_0 = 1, \quad \psi_j = \frac{1}{j^2} \ (j \geq 1), \qquad \varepsilon_t \overset{i.i.d.}{\sim} (0, \sigma^2).$$

(a) Since $\mathbb{E}[\varepsilon_t] = 0$ and the coefficients are summable,

$$\mathbb{E}[y_t] = \sum_{j=0}^{\infty} \psi_j \, \mathbb{E}[\varepsilon_{t-j}] = 0.$$

Moreover,

$$Var(y_t) = \sigma^2 \sum_{j=0}^{\infty} \psi_j^2 = \sigma^2 \left(1 + \sum_{j=1}^{\infty} \frac{1}{j^4} \right) = \sigma^2(1 + \zeta(4)) = \sigma^2 \left(1 + \frac{\pi^4}{90} \right).$$

(b) For $h \geq 1$,

$$\gamma(h) \equiv Cov(y_t, y_{t-h}) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h} = \sigma^2 \left[\frac{1}{h^2} + \sum_{j=1}^{\infty} \frac{1}{j^2 (j+h)^2} \right].$$

Because $\sum_{j=0}^{\infty} |\psi_j| = 1 + \sum_{j=1}^{\infty} 1/j^2 = 1 + \zeta(2) < \infty$, the linear process is covariance–stationary: the mean is constant, the variance is finite and constant, and $\gamma(h)$ depends only on h.

(c) Multiply by h^2:

$$\frac{h^2}{\sigma^2} \, \gamma(h) = 1 + h^2 \sum_{j=1}^{\infty} \frac{1}{j^2 (j+h)^2}.$$

Since $(j+h)^2 \geq h^2$ for $j \geq 1$,

$$h^2 \frac{1}{j^2 (j+h)^2} \leq \frac{1}{j^2}, \qquad \Rightarrow \qquad h^2 \sum_{j=1}^{\infty} \frac{1}{j^2 (j+h)^2} \leq \sum_{j=1}^{\infty} \frac{1}{j^2} = \zeta(2) = \frac{\pi^2}{6}.$$

The sum is non–negative, so for all $h \geq 1$,

$$\sigma^2 \ \leq \ h^2 \gamma(h) \ \leq \ \sigma^2 \left(1 + \frac{\pi^2}{6} \right).$$

Hence there exist positive constants $c_1 = \sigma^2$ and $c_2 = \sigma^2(1 + \pi^2/6)$ such that $c_1 \leq h^2 \gamma(h) \leq c_2$, i.e.

$$\boxed{\gamma(h) = O\left(\frac{1}{h^2} \right)}.$$

(d) Assume, by contradiction, that y_t is ARMA(p, q) with finite p, q and stationary AR part. Then the autocovariance of an ARMA process decays *exponentially*: there exist constants $C > 0$ and $\rho \in (0, 1)$ such that

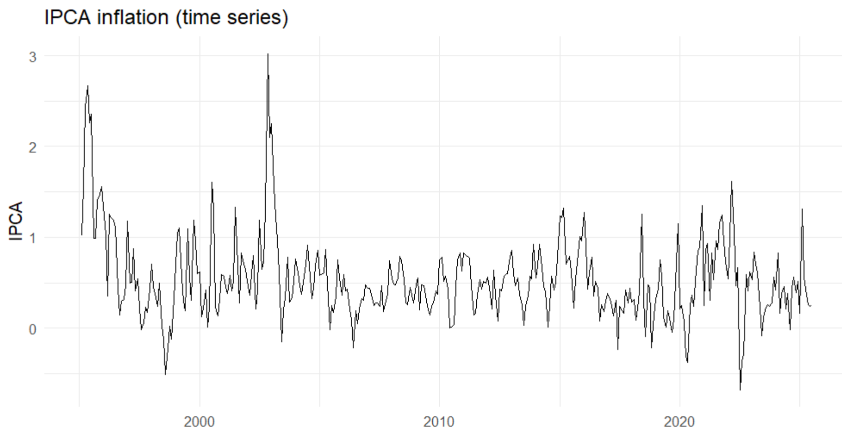
$$|\tilde{\gamma}(h)| \leq C \, \rho^h \qquad (h \to \infty).$$

From item (c), $\gamma(h) = O(1/h^2)$. Therefore

$$\frac{\gamma(h)}{|\tilde{\gamma}(h)|} \;\geq\; \frac{c_1/h^2}{C\rho^h} = \frac{c_1}{C}\,\frac{\rho^{-h}}{h^2} \xrightarrow[h\to\infty]{} \infty,$$

since an exponential dominates any polynomial. The fact this ratio is not constant at 1 implies that y_t can *not* be an ARMA(p,q) process.

Problem 3. (a)



(c)

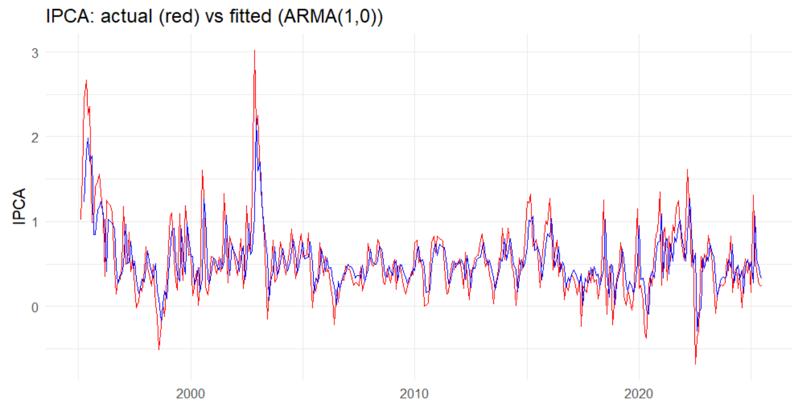
Table 1: ARMA CML estimates (95% Wald intervals)

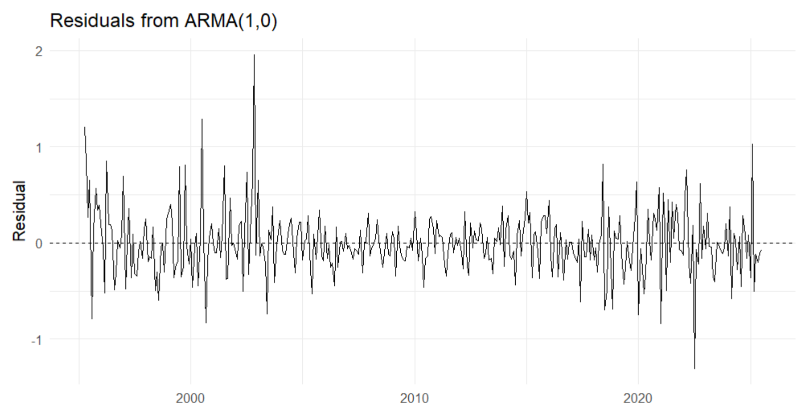
Model	Parameter	Estimate	95% CI	logLik	AIC	BIC
ARMA(1,0)	c	0.169	[0.1165, 0.2223]	-117.02	240.04	251.74
	phi1	0.682	[0.6067, 0.7566]	-117.02	240.04	251.74
ARMA(2,0)	c	0.167	[0.1114, 0.2225]	-115.01	238.03	253.62
	phi1	0.666	[0.5638, 0.7680]	-115.01	238.03	253.62
	phi2	0.017	[-0.0853, 0.1187]	-115.01	238.03	253.62
ARMA(0,1)	c	0.535	[0.4760, 0.5938]	-151.60	309.20	320.90
	theta1	0.565	[0.4961, 0.6343]	-151.60	309.20	320.90
ARMA(1,1)	c	0.168	[0.1001, 0.2369]	-117.02	242.04	257.64
	phi1	0.683	[0.5726, 0.7939]	-117.02	242.04	257.64
	theta1	-0.003	[-0.1549, 0.1488]	-117.02	242.04	257.64
ARMA(2,1)	c	0.152	[-0.0840, 0.3874]	-115.01	240.02	259.50
	phi1	0.754	[-0.5909, 2.0984]	-115.01	240.02	259.50
	phi2	-0.042	[-0.9516, 0.8672]	-115.01	240.02	259.50
	theta1	-0.089	[-1.4342, 1.2569]	-115.01	240.02	259.50
ARMA(0,2)	c	0.531	[0.4626, 0.5999]	-131.13	270.26	285.85
	theta1	0.653	[0.5544, 0.7519]	-131.13	270.26	285.85
	theta2	0.273	[0.1906, 0.3547]	-131.13	270.26	285.85
ARMA(1,2)	c	0.156	[0.0757, 0.2355]	-114.97	239.94	259.43
	phi1	0.704	[0.5666, 0.8413]	-114.97	239.94	259.43
	theta1	-0.036	[-0.2015, 0.1292]	-114.97	239.94	259.43
	theta2	-0.019	[-0.1537, 0.1148]	-114.97	239.94	259.43
ARMA(2,2)	c	0.250	[0.0830, 0.4176]	-114.60	241.19	264.58
	phi1	0.099	[-0.6275, 0.8253]	-114.60	241.19	264.58
	phi2	0.426	[-0.0580, 0.9092]	-114.60	241.19	264.58
	theta1	0.570	[-0.1561, 1.2957]	-114.60	241.19	264.58
	theta2	-0.052	[-0.1947, 0.0904]	-114.60	241.19	264.58

Notes: Gaussian conditional maximum likelihood (CML). Wald 95% CIs. Model-level statistics repeat across parameter rows. Innovation variance rows omitted.

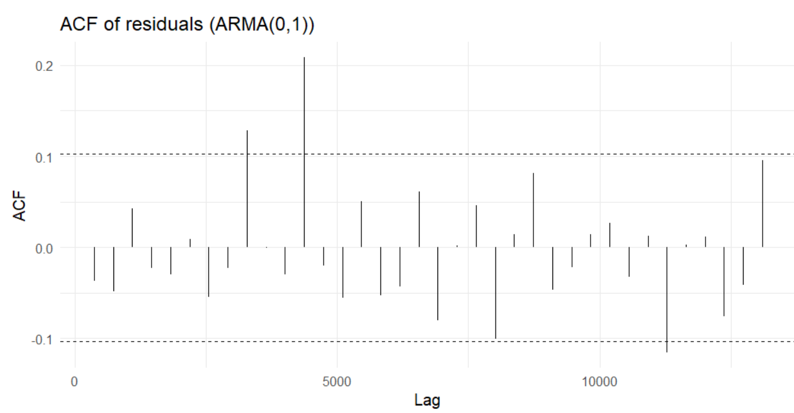
(d) AIC criteria selects ARMA(2,0) and BIC criteria selects ARMA(1,0). I will choose ARMA(1,0) for parsimony.

(e)

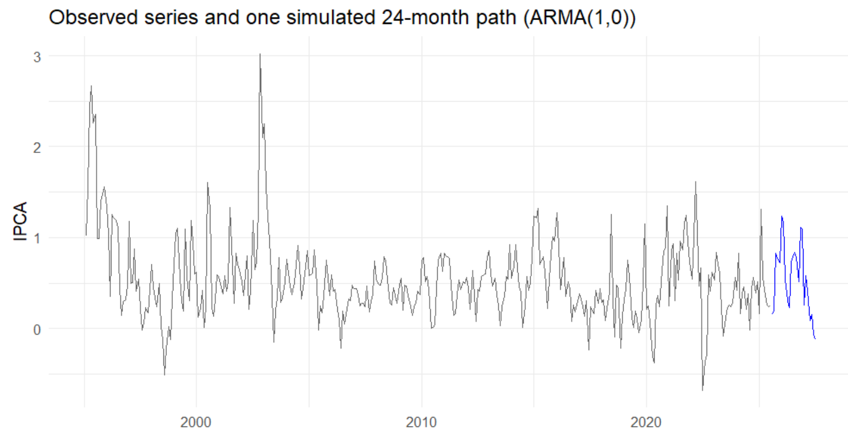

IPCA: actual (red) vs fitted (ARMA(1,0))
IPCA
3
2
1
0
2000
2010
2020


Residuals from ARMA(1,0)
Residual
2
1
0
-1
2000
2010
2020

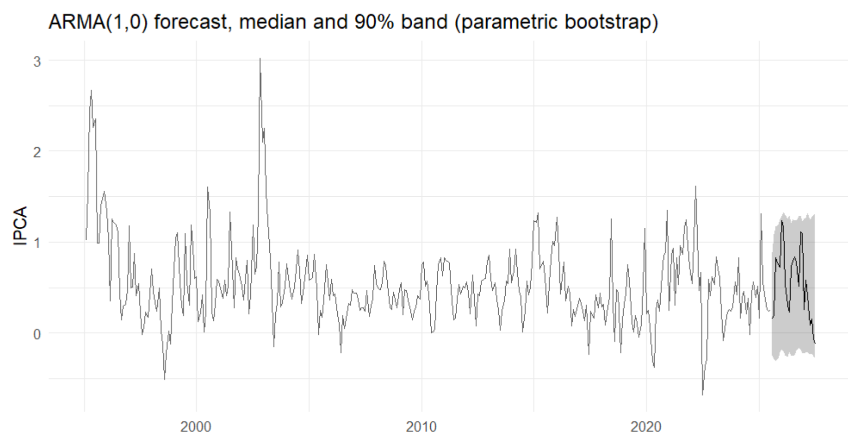
(f)


ACF of residuals (ARMA(0,1))
ACF
0.2
0.1
0.0
-0.1
0
5000
10000
Lag

(g)



(h)



(i) This procedure is a *parametric* bootstrap: uncertainty comes from simulating shocks from the assumed distribution $\mathcal{N}(0, \widehat{\sigma}^2)$ and propagating them through the estimated ARMA(1,0) dynamics. A nonparametric bootstrap would instead resample (possibly block-resample) *empirical* residuals without imposing normality. Parametric bootstrap is more efficient and yields smoother, tighter bands when the model is correctly specified; the nonparametric version is more robust to distributional misspecification (e.g., fat tails or skewness) but must handle serial dependence carefully (e.g., via residual or block bootstrap).