Appendix:
$AR(1)$ with a deterministic time trend

by Vitor Possebom

In class, there was some confusion on how to derive the results associated with an $AR(p)$ model with a deterministic time trend. To illustrate this model and provide intuition, I focus on the $AR(1)$ model with a deterministic time trend.[1] I like the $AR(1)$ model because I can write the entire matrices without using the "$\cdots$" notation, allowing us to focus on the stuff that actually matters for our understanding.

Let $\{Y_t\}$ be a stochastic process satisfying

$$Y_t = \alpha + \delta \cdot t + \phi \cdot Y_{t-1} + \epsilon_t$$

where ϵ_t is i.i.d with mean zero, variance σ^2 and finite fourth moment, and $|\phi| < 1$.

This problem becomes easier when we use matrix notation:

$$Y_t = X_t' \beta + \epsilon_t,$$

where $X_t = \begin{bmatrix} Y_{t-1} & 1 & t \end{bmatrix}'$ and $\beta = \begin{bmatrix} \phi & \alpha & \delta \end{bmatrix}'$.

Our OLS estimator is given by

$$b_T := \begin{bmatrix} \hat{\phi}_T & \hat{\alpha}_T & \hat{\delta}_T \end{bmatrix}' = \left[\sum_{t=1}^{T} X_t X_t' \right]^{-1} \left[\sum_{t=1}^{T} X_t Y_t \right].$$

[1]Hamilton's Chapter 16 goes through the math for the $AR(p)$ model with a deterministic time trend.

We want to transform our model so that it includes only a constant term, a time trend, and zero-mean weakly stationary random variables. To do so, we define

$$G' := \begin{bmatrix} 1 & 0 & 0 \\ -\alpha+\delta & 1 & 0 \\ -\delta & 0 & 1 \end{bmatrix},$$

$$\begin{aligned} X_t^* := GX_t &= \begin{bmatrix} Y_{t-1}^* \\ 1 \\ t \end{bmatrix} \\ &= \begin{bmatrix} 1 & -\alpha+\delta & -\delta \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} Y_{t-1} \\ 1 \\ t \end{bmatrix} = \begin{bmatrix} Y_{t-1} - \alpha - \delta - \delta \cdot t \\ 1 \\ t \end{bmatrix} \\ &= \begin{bmatrix} Y_{t-1} - \alpha - \delta \cdot (t-1) \\ 1 \\ t \end{bmatrix}, \end{aligned}$$

and

$$\begin{aligned} \beta^* := (G')^{-1} \beta &= \begin{bmatrix} \phi^* \\ \alpha^* \\ \delta^* \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ -\alpha+\delta & 1 & 0 \\ -\delta & 0 & 1 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \phi \\ \alpha \\ \delta \end{bmatrix} \end{aligned}$$

$$
= \begin{bmatrix} 1 & 0 & 0 \\ \alpha - \delta & 1 & 0 \\ \delta & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \phi \\ \alpha \\ \delta \end{bmatrix} = \begin{bmatrix} \phi \\ \alpha \cdot \phi - \delta \cdot \phi + \alpha \\ \delta \cdot \phi + \delta \end{bmatrix}
$$

$$
= \begin{bmatrix} \phi \\ \alpha \cdot (1 + \phi) - \delta \cdot \phi \\ \delta \cdot (1 + \phi) \end{bmatrix}.
$$

Using these definitions, we can rewrite our model as

$$
\begin{aligned}
Y_t &= X_t'\beta + \epsilon_t = X_t' \cdot G' \cdot (G')^{-1} \beta + \epsilon_t \\
&= (G \cdot X_t)' \cdot \beta^* + \epsilon_t \\
&= (X_t^*)' \cdot \beta^* + \epsilon_t.
\end{aligned}
$$

Now, we can define the unfeasible OLS estimator for the transformed model as

$$
b_T^* := \begin{bmatrix} \hat{\phi}_T^* & \hat{\alpha}_T^* & \hat{\delta}_T^* \end{bmatrix}' = \left[\sum_{t=1}^{T} X_t^* \cdot (X_t^*)' \right]^{-1} \left[\sum_{t=1}^{T} X_t^* Y_t \right].
$$

Moreover, note that

$$
\begin{aligned}
b_T^* &= \left[\sum_{t=1}^{T} X_t^* \cdot (X_t^*)' \right]^{-1} \left[\sum_{t=1}^{T} X_t^* Y_t \right] \\
&= \left[\sum_{t=1}^{T} G \cdot X_t \cdot (G \cdot X_t^*)' \right]^{-1} \left[\sum_{t=1}^{T} G \cdot X_t^* Y_t \right] \\
&= \left[\sum_{t=1}^{T} X_t \cdot X_t' G' \right]^{-1} \cdot G^{-1} \cdot G \cdot \left[\sum_{t=1}^{T} X_t^* Y_t \right] \\
&= (G')^{-1} \cdot \left[\sum_{t=1}^{T} X_t \cdot X_t' \right]^{-1} \cdot \left[\sum_{t=1}^{T} X_t^* Y_t \right] \\
&= (G')^{-1} \cdot b_T,
\end{aligned}
$$

implying that

$$b_T = G' \cdot b_T^*.$$

Consequently, we have that

$$\begin{bmatrix} \hat{\phi}_T \\ \hat{\alpha}_T \\ \hat{\delta}_T \end{bmatrix} = b_T = G' \cdot b_T^* = \begin{bmatrix} 1 & 0 & 0 \\ -\alpha + \delta & 1 & 0 \\ -\delta & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{\phi}_T^* \\ \hat{\alpha}_T^* \\ \hat{\delta}_T^* \end{bmatrix}. \tag{1}$$

Looking at the second row of Equation (1), we have that

$$\hat{\alpha}_T = \begin{bmatrix} -\alpha + \delta & 1 & 0 \end{bmatrix} \cdot b_T^* = g_\alpha' \cdot b_T^*,$$

where $g_\alpha' := \begin{bmatrix} -\alpha + \delta & 1 & 0 \end{bmatrix}$. Hence, we can derive the asymptotic properties of $\hat{\alpha}_T$ by using the Continuous Mapping Theorem and the asymptotic distribution of b_T^*.

Looking at the third row of Equation (1), we have that

$$\hat{\delta}_T = \begin{bmatrix} -\delta & 0 & 1 \end{bmatrix} \cdot b_T^* = \begin{bmatrix} -\delta & 0 & 0 \end{bmatrix} \cdot b_T^* + \hat{\delta}_T^* = g_\delta' \cdot b_T^* + \hat{\delta}_T^*,$$

where $g_\delta' := \begin{bmatrix} -\delta & 0 & 0 \end{bmatrix}$. We broke this term into two parts because the first part converges at rate $\sqrt{T}$ and the second part converges at rate $T^{3/2}$.

Combining Lemma 2 from the lecture notes with the Continuous Mapping Theorem, we have

$$\sqrt{T} \cdot g_\delta' \cdot (b_T^* - \beta^*) \xrightarrow{d} N\left(0, \sigma^2 g_\delta' \left(Q^*\right)^{-1} g_\delta\right).$$

We also know that $\sqrt{T}\left(\hat{\delta}_T^* - \delta^*\right) \xrightarrow{p} 0$ because $\hat{\delta}_T^*$ converges at rate $T^{3/2}$ and multiplying it by $\sqrt{T}$ is not enough to avoid the collapse of our estimator to a degenerate distribution around the true parameter.[2]

Combining the last three results, we have that

$$
\begin{aligned}
\sqrt{T} \cdot \left(\hat{\delta}_T - \delta\right) &= \sqrt{T} \cdot \left(g_\delta' \cdot b_T^* + \hat{\delta}_T^* - g_\delta' \cdot \beta^* + \delta_T^*\right) \\
&= \sqrt{T} \cdot g_\delta' \cdot \left(b_T^* - \beta^*\right) + \sqrt{T}\left(\hat{\delta}_T^* - \delta^*\right) \\
&\xrightarrow{d} N\left(0, \sigma^2 g_\delta' \left(Q^*\right)^{-1} g_\delta\right).
\end{aligned}
$$

[2]I illustrate this peculiar result using a MC simulation.