# CONSISTENCY OF OLS, PROPERTIES OF CONVERGENCE

Though this result was referred to often in class, and perhaps even proved at some point, a student has pointed out that it does not appear in the notes. OLS is consistent under much weaker conditions that are required for unbiasedness or asymptotic normality. Not even predeterminedness is required.

As usual we assume

$$y_t = X_t\beta + \varepsilon_t\,,\ t = 1,\dots,T\,. \tag{1}$$

**Theorem.** *Suppose*

(i) $X_t, \varepsilon_t$ *are jointly ergodic;*
(ii) $E[X_t'\varepsilon_t] = 0$*;*
(iii) $E[X_t'X_t] = \Sigma_X$ *and* $|\Sigma_X| \neq 0$*.*

*Then the OLS estimator of* $\beta$ *is consistent.*

*Proof.* The OLS estimator is

$$\hat\beta_T = (X'X)^{-1}X'y = (\sum_{t=1}^{T} X_t'X_t)^{-1}\sum_{t=1}^{T} X_t'y_t$$

$$\left(\frac{1}{T}\sum_{t=1}^{T} X_t'X_t\right)^{-1}\frac{1}{T}\sum_{t=1}^{T}(X_t'X_t\beta + X_t'\varepsilon_t)$$

$$= \beta + \underbrace{\left(\frac{1}{T}\sum_{t=1}^{T} X_t'X_t\right)}_{1}^{-1} \underbrace{\frac{1}{T}\sum_{t=1}^{T} X_t'\varepsilon_t}_{2}\,.$$

The two expressions with underbraces are both time averages of functions of an ergodic process, by assumption, so each converges a.s. to its expectation. But the first term converges to a nonsingular limit, and the mapping from a matrix to its inverse is continuous at any nonsingular argument. Therefore the inverse of the $(1/T)X'X$ term also converges a.s., to $\Sigma_X^{-1}$. The second underbracketed expectation converges to zero a.s. by assumption (ii) and ergodicity. Since multiplication is also a continuous function, the product of the limits is the limit of the products and we can conclude that $\hat\beta_T \xrightarrow{a.s.} \beta$. □

Below are some properties of convergence in probability, a.s., etc. that were referred to several times in lectures and precepts, but are not written out in the notes elsewhere.

**Proposition.** (i) $X_t \xrightarrow[T\to\infty]{a.s.} Z \Rightarrow X_t \xrightarrow[T\to\infty]{P} Z$;

(ii) *For any* $p > 0$, $E[|X_t - Z|^p] \to 0 \Rightarrow X_t \xrightarrow{P} Z$;

(iii) *If* $|X_t|$ *is bounded above by* $B$, *uniformly in* $t$, *then for any* $p > 0$, $X_t \xrightarrow{P} Z \Rightarrow E[|X_t - Z|^p] \to 0$;

(iv) $X_t \xrightarrow{P} Z \Rightarrow X_t \xrightarrow{\mathcal{D}} Z$;

(v) *If* $f$ *is continuous at a set of* $Z$ *values with probability one, then* $X_t \xrightarrow{P} Z \Rightarrow f(X_t) \xrightarrow{P} f(Z)$.

It is important to remember which of these relationships are one-way, or depend on strong assumptions. a.s. convergence is stronger than convergence in probability: (i) is one-way. Convergence in $p$'th moment is stronger than convergence in probability: (ii) is one-way, unless we add the boundedness assumption of (iii). Convergence in probability is stronger than convergence in distribution: (iv) is one-way.

Also, as was emphasized in lecture, these convergence notions make assertions about different types of objects. Convergence a.s. makes an assertion about the distribution of *entire random sequences* of $X_t$'s. One has to think of all the $X_t$'s and $Z$ being "drawn" simultaneously. Convergence a.s. says that each such draw will show ordinary calculus convergence of the $x_t$'s to $z$. Convergence in probability makes an assertion about the sequence of pairwise joint distributions of $X_t$ with $Z$. It asserts that these joint distributions make $X_t$ and $Z$ increasingly tightly related, but it makes no claims about entire sequences of $X_t$ draws. Convergence in distribution makes an assertion about the distribution of $X_t$ and the distribution of $Z$. It says these two distributions become more similar, without making any assertion about whether $X_t$ and $Z$ draws are likely to be close to each other.

Check your understanding. In which senses does $X_t$ converge to $Z$ as $t \to \infty$?

(1) $X_t$ i.i.d. across $t$, $Z = X_1$.

(2) $X_t = 1/t$ with probability $1 - 1/t$, $X_t = t^2$ with probability $1/t$. $X_t$ independent across $t$. Note that $E[X_t] \to \infty$. $Z = 0$ with probability one. (Determining whether this converges a.s. is hard. Checking convergence in probability, in distribution, and in $p$'th moment should be easy.)

(3) $Z = 1$ with probability .5, $Z = -1$ with probability .5. $X_t = (e^{Zt} - 1)/(e^{Zt} + 1)$.