Modes of Convergence

Definition 1 (convergence almost surely). The matrix-valued sequence of random variables Z_n is said to converge to a random matrix Z *almost surely* (or *with probability one*), written as $Z_n \overset{as}{\to} Z$, if

$$\Pr\left\{\lim_{n\to\infty} Z_n = Z\right\} = 1,$$

i.e. almost every trajectory converges to Z.

Definition 2 (convergence in probability). The matrix-valued sequence of random variables Z_n is said to converge to a random matrix Z *in probability*, written as $Z_n \overset{p}{\to} Z$ or $p\lim Z_n = Z$, if

$$\forall \varepsilon > 0 \quad \lim_{n\to\infty} \Pr\left\{\|Z_n - Z\| > \varepsilon\right\} = 0,$$

i.e. the probability of large deviations converges to 0.

Result. $Z_n \overset{as}{\to} Z \Rightarrow Z_n \overset{p}{\to} Z$.

Definition 3 (convergence in mean square). The matrix-valued sequence of random variables Z_n is said to converge to a random matrix Z *in mean square*, written as $Z_n \overset{ms}{\to} Z$, if

$$\lim_{n\to\infty} \mathbb{E}\left[\|Z_n - Z\|^2\right] = 0,$$

i.e. the mean squared error converges to 0.

Result. $Z_n \overset{ms}{\to} Z \Rightarrow Z_n \overset{p}{\to} Z$.

Definition 4 (convergence in distribution). The vector-valued sequence of random variables Z_n is said to converge to a random vector Z *in distribution*, written as $Z_n \overset{d}{\to} Z$ or $Z_n \overset{d}{\to} \mathcal{D}_Z$, where $\mathcal{D}_Z$ is the distribution of Z, if

$$\lim_{n\to\infty} \Pr\{Z_n \le z\} = \Pr\{Z \le z\}$$

for all continuity points z of $\Pr\{Z \le z\}$.

Result. $Z_n \overset{p}{\to} Z \Rightarrow Z_n \overset{d}{\to} Z$. If Z is constant, $Z_n \overset{p}{\to} Z \Leftrightarrow Z_n \overset{d}{\to} Z$.

Continuous Mapping Theorems

Theorem (Mann–Wald). Suppose that $g(z)$ is a continuous $\mathbb{R}^{k_1\times k_2} \to \mathbb{R}^{\ell_1\times \ell_2}$ function.

- If $Z_n \overset{as}{\to} Z$ as $n\to\infty$, then $g(Z_n) \overset{as}{\to} g(Z)$.
- If $Z_n \overset{p}{\to} Z$ as $n\to\infty$, then $g(Z_n) \overset{p}{\to} g(Z)$.
- If $Z_n \overset{ms}{\to} Z$ as $n\to\infty$ and g is linear, then $g(Z_n) \overset{ms}{\to} g(Z)$.
- If $Z_n \overset{d}{\to} Z$ as $n\to\infty$, then $g(Z_n) \overset{d}{\to} g(Z)$.

If in addition $Z = \text{const}$, then continuity of g only at Z suffices.

Theorem (Slutsky). If $U_n \overset{p}{\to} U = \text{const}$ and $V_n \overset{d}{\to} V$ as $n\to\infty$, then

- $U_n + V_n \overset{d}{\to} U + V$.
- $U_nV_n \overset{d}{\to} UV$, $V_nU_n \overset{d}{\to} VU$.
- $U_n^{-1}V_n \overset{d}{\to} U^{-1}V$, $V_nU_n^{-1} \overset{d}{\to} VU^{-1}$ if $\Pr\{\det(U_n) = 0\} = 0$.

Delta Method

Theorem (Delta Method). Let the sequence of $k\times 1$ random vectors Z_n satisfy

$$\sqrt{n}\,(Z_n - Z) \overset{d}{\to} \mathcal{N}(0,\Sigma),$$

as $n\to\infty$, where $Z = p\lim Z_n$ is constant, and the $\mathbb{R}^k \to \mathbb{R}^\ell$ function $g(z)$ be continuously differentiable at Z. Then

$$\sqrt{n}\,(g(Z_n) - g(Z)) \overset{d}{\to} \mathcal{N}(0, G\Sigma G'),$$

where $G = \left.\dfrac{\partial g(z)}{\partial z'}\right|_{z=Z}$.

Thanks to Anton Cheremukhin for these notes

Laws of Large Numbers

Theorem A (Kolmogorov, independent identical observations). Let $\{Z_i\}_{i=1}^{\infty}$ be independent and identically distributed (IID), and let $\mathbb{E}[|Z_i|]$ exist. Then

$$\frac{1}{n}\sum_{i=1}^{n} Z_i \xrightarrow{as} \mathbb{E}[Z_i]$$

as $n \to \infty$.

Theorem B (Kolmogorov, independent heterogeneous observations). Let $\{Z_i\}_{i=1}^{\infty}$ be independent with finite variance σ_i^2. If

$$\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty,$$

then

$$\frac{1}{n}\sum_{i=1}^{n} Z_i - \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} Z_i\right] \xrightarrow{as} 0$$

as $n \to \infty$.

Theorem C (Chebyshev, uncorrelated observations). Let $\{Z_i\}_{i=1}^{\infty}$ be uncorrelated, i.e. $\mathbb{C}\left[Z_i, Z_j\right] = 0$ if $i \neq j$. If

$$\frac{1}{n^2}\sum_{i=1}^{n} \sigma_i^2 \underset{n\to\infty}{\to} 0$$

then

$$\frac{1}{n}\sum_{i=1}^{n} Z_i - \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n} Z_i\right] \xrightarrow{p} 0$$

as $n \to \infty$.

Theorem D (Birkhoff-Khinchin, dependent observations, "Ergodic Theorem"). Let $\{Z_t\}_{t=1}^{\infty}$ be a stationary and ergodic sequence of random variables, and let $\mathbb{E}[|Z_t|] < \infty$. Then

$$\frac{1}{T}\sum_{t=1}^{T} Z_t \xrightarrow{as} \mathbb{E}[Z_t]$$

as $T \to \infty$.

Central Limit Theorems

Theorem E (Lindberg-Levy, independent identical observations). Let $\{Z_i\}_{i=1}^{\infty}$ be independent and identically distributed (IID) with $\mathbb{E}[Z_i] = \mu$ and $\mathbb{V}[Z_i] = \sigma^2$. Then

$$\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} Z_i - \mu\right) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

as $n \to \infty$.

Theorem F (Lyapunov, independent heterogeneous observations). Let $\{Z_i\}_{i=1}^{\infty}$ be independent with $\mathbb{E}[Z_i] = \mu_i$, $\mathbb{V}[Z_i] = \sigma_i^2$ and $\mathbb{E}[|Z_i - \mu_i|^3] = \nu_i$. If

$$\frac{\left(\sum_{i=1}^{n} \nu_i\right)^{\frac{1}{3}}}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{\frac{1}{2}}} \underset{n\to\infty}{\to} 0,$$

then

$$\frac{\sum_{i=1}^{n} (Z_i - \mu_i)}{\left(\sum_{i=1}^{n} \sigma_i^2\right)^{\frac{1}{2}}} \xrightarrow{d} \mathcal{N}(0, 1)$$

as $n \to \infty$.

Theorem G (Billingsley, martingale difference sequences). Let $\{Z_t\}_{t=-\infty}^{+\infty}$ be a stationary and ergodic martingale difference sequence relative to its own past, with $\sigma^2 = \mathbb{E}[Z_t^2] < \infty$. Then

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Z_t \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

as $T \to \infty$.

Theorem H (general dependent observations). Let $\{Z_t\}_{t=-\infty}^{+\infty}$ be a stationary and ergodic sequence of random variables with

$$v_z = \sum_{j=-\infty}^{+\infty} \mathbb{C}[Z_t, Z_{t-j}] < \infty.$$

Then under suitable conditions,

$$\sqrt{T}\left(\frac{1}{T}\sum_{t=1}^{T} Z_t - \mathbb{E}[Z_t]\right) \xrightarrow{d} \mathcal{N}(0, v_z)$$

as $T \to \infty$.

Thanks to Anton Cheremukhin for these notes