

L11 Sequence bounded in probability

1. Relations of strong convergence and weak convergence

(1) Recall

Let $D_n(\epsilon) = [\|X_n - X\| > \epsilon]$. Then

$$\lim_{N \rightarrow \infty} P(\cup_{n=N}^{\infty} D_n(\epsilon)) = 0 \quad \forall \epsilon > 0 \iff X_n \xrightarrow{a.s.} X$$

$$\lim_{n \rightarrow \infty} P(D_n(\epsilon)) = 0 \quad \forall \epsilon > 0 \iff X_n \xrightarrow{p} X.$$

Then $X_n \xrightarrow{a.s.} X \implies X_n \xrightarrow{p} X$ since

$$X_n \xrightarrow{a.s.} X \iff \lim_{N \rightarrow \infty} P(\cup_{n=N}^{\infty} D_n(\epsilon)) = 0 \implies \lim_{N \rightarrow \infty} P(D_N(\epsilon)) = 0 \iff X_n \xrightarrow{p} X.$$

Borel-Cantelli lemma: $\sum_{n=1}^{\infty} P(D_n(\epsilon)) < \infty \implies X_n \xrightarrow{a.s.} X$ since

$$\begin{aligned} \sum_{n=1}^{\infty} P(D_n(\epsilon)) < \infty &\iff \lim_{N \rightarrow \infty} \sum_{n=N}^{\infty} P(D_n(\epsilon)) = 0 \\ &\implies \lim_{N \rightarrow \infty} P(\cup_{n=N}^{\infty} D_n(\epsilon)) = 0 \iff X_n \xrightarrow{a.s.} X. \end{aligned}$$

(2) $X_n \xrightarrow{p} X \implies X_{n_k} \xrightarrow{a.s.} X$ for some X_{n_k} .

Proof: Note that

$$\begin{aligned} X_n \xrightarrow{p} X &\iff P(D_n(\epsilon)) \rightarrow 0 \quad \forall \epsilon > 0 \iff P(D_n(\frac{1}{k})) \rightarrow 0 \quad \forall k > 0 \\ &\implies \exists n_k \text{ such that } P(D_{n_k}(\frac{1}{k})) < \frac{1}{k^2}. \end{aligned}$$

But there exists K such that $\frac{1}{K} < \epsilon$. Thus $D_n(\epsilon) \subset D_{n_k}(\frac{1}{k})$ for $k \geq K$.

$$\begin{aligned} \sum_{k=1}^{\infty} P(D_{n_k}(\epsilon)) &\leq \sum_{k=1}^{K-1} P(D_{n_k}(\epsilon)) + \sum_{k=K}^{\infty} P(D_{n_k}(\frac{1}{k})) \\ &\leq \sum_{k=1}^{K-1} P(D_{n_k}(\epsilon)) + \sum_{k=K}^{\infty} \frac{1}{k^2} < \infty. \end{aligned}$$

By Borel-Cantelli lemma, $X_{n_k} \xrightarrow{a.s.} X$. $\square$

(3) $X_n \xrightarrow{p} X \iff \forall n_k, \exists n_{k_i} \text{ such that } X_{n_{k_i}} \xrightarrow{a.s.} X$.

Proof. $\implies$: $X_n \xrightarrow{p} X \implies \forall n_k, X_{n_k} \xrightarrow{p} X \implies X_{n_{k_i}} \xrightarrow{a.s.} X$ for some $X_{n_{k_i}}$.

$\Leftarrow$: We show that if $X_n \not\xrightarrow{p} X$, then $\exists n_k, \forall n_{k_i}, X_{n_{k_i}} \not\xrightarrow{a.s.} X$.

From $X_n \not\xrightarrow{p} X \iff P(D_n(\epsilon)) \not\rightarrow 0 \quad \forall \epsilon > 0$,

$$X_n \not\xrightarrow{p} X \iff \exists \epsilon_0 > 0 \text{ such that } P(D_n(\epsilon_0)) \not\rightarrow 0$$

But $P(D_n(\epsilon)) \rightarrow 0 \iff \forall \delta > 0, \exists N > 0, \forall n > N, P(D_n(\epsilon)) \leq \delta$. Hence

$P(D_n(\epsilon_0)) \not\rightarrow 0 \iff \exists \delta_0 > 0, \forall N > 0, \exists n_N > N$ such that $P(D_{n_N}(\epsilon_0)) > \delta_0$

With $P(D_{n_k}(\epsilon_0)) > \delta_0$ for all n_k , for all n_{k_i} $P(D_{n_{k_i}}(\epsilon_0)) > \delta_0$.

Hence For all n_{k_i} , $X_{n_{k_i}} \not\xrightarrow{p} X$. Consequently $X_{n_{k_i}} \not\xrightarrow{a.s.} X$. $\square$

2. Random X bounded in probability

(1) Definitions

Random X is bounded in probability $\xLeftrightarrow{def} \forall \epsilon > 0 \exists M_\epsilon > 0$
such that $P(\|X\| \leq M_\epsilon) > 1 - \epsilon$.

(2) Sufficient condition I

If $E(\|X\|^2)$ exists, then X is bounded in probability.

Proof. For $\epsilon > 0$ let $M_\epsilon > \sqrt{\frac{E(\|X\|^2)}{\epsilon}}$. By Chebyshev inequality,

$$P(\|X\| > M_\epsilon) \leq \frac{E(\|X\|^2)}{M_\epsilon^2} < \frac{E(\|X\|^2)}{E(\|X\|^2)/\epsilon} = \epsilon.$$

So $P(\|X\| \leq M_\epsilon) > 1 - \epsilon$. Hence X is bounded in probability. $\square$

(2) Sufficient condition II

If $X \sim (\mu, \Sigma)$, then X is bounded in probability

Proof. $E(\|X\|^2) = E(X'X) = \sum_{i=1}^k E(X_i^2) = \sum_{i=1}^k (\mu_i^2 + \sigma_i^2) = \mu'\mu + \text{tr}(\Sigma)$.

By (1), X is bounded in probability. $\square$

Comment: $E(X'AX) = \mu' A \mu + \text{tr}(A\Sigma)$ (Stat 776).

3. Random sequence bounded in probability

(1) Definition

Random sequence X_n is bounded in probability

$$\stackrel{\text{def}}{\iff} \forall \epsilon > 0 \exists M_\epsilon > 0 \text{ and } N_\epsilon > 0, \forall n > N_\epsilon, P(\|X_n\| \leq M_\epsilon) > 1 - \epsilon.$$

Random X bounded in probability can be treated as a random sequence $X_n \equiv X$ bounded in probability.

(2) Property I

If $X_n \xrightarrow{p} X$, then $X_n - X$ is bounded in probability.

Proof. $X_n \xrightarrow{p} X$. So for $M > 0$, $\lim_n P(\|X_n - X\| < M) = 1$.

Thus $\forall \epsilon > 0 \exists N > 0 \forall n > N P(\|X_n - X\| < M) > 1 - \epsilon$.

Hence $X_n - X$ is bounded in probability. $\square$

(3) Property II

If in $Z_n = X_n + Y_n$ both X_n and Y_n are bounded in probability, so is Z_n .

Proof: $\forall \epsilon > 0, \exists M_{x\epsilon} > 0 \text{ and } N_{x\epsilon} > 0, \forall n > N_{x\epsilon}, P(\|X_n\| \leq M_{x\epsilon}) > 1 - \epsilon/2$.

$\exists M_{y\epsilon} > 0 \text{ and } N_{y\epsilon} > 0, \forall n > N_{y\epsilon}, P(\|Y_n\| \leq M_{y\epsilon}) > 1 - \epsilon/2$.

Let $M_\epsilon = M_{x\epsilon} + M_{y\epsilon}$ and $N = \max[N_{x\epsilon}, N_{y\epsilon}]$.

With $A = [\|X_n\| \leq M_{x\epsilon}]$ and $B = [\|Y_n\| \leq M_{y\epsilon}]$, by the triangular inequality $\|Z_n\| \leq \|X_n\| + \|Y_n\|$, $A \cap B \subset [\|Z_n\| \leq M_\epsilon]$.

When $n > N_\epsilon$,

$$\begin{aligned} P(\|Z_n\| \leq M_\epsilon) &\geq P(A \cap B) = P(A) + P(B) - P(A \cup B) \\ &\geq (1 - \frac{\epsilon}{2}) + (1 - \frac{\epsilon}{2}) - P(A \cup B) = 1 - \epsilon + [1 - P(A \cup B)] \\ &\geq 1 - \epsilon. \end{aligned}$$

Hence Z_n is bounded in probability. $\square$.

Ex: $X_n \xrightarrow{p} X$ where $X \sim (\mu, \Sigma)$. Then X_n is bounded in probability.

By (2) in 3, $X_n - X$ is bounded in probability.

By (3) in 2, X is bounded in probability.

By (3) in 3, X_n is bounded in probability.

L12 Properties of sequence bounded in probability

1. Three iff properties for boundedness in probability

(1) X_n is bounded in probability $\iff g(X_n)$ is bounded in probability $\forall$ continuous $g(\cdot)$.

Proof. $\Leftarrow$: With continuous $g(X) = X$, $g(X_n) = X_n$ is bounded in probability.

$\Rightarrow$: X_n is bounded in probability. So

$\forall \epsilon > 0 \exists M_\epsilon > 0$ and $N_\epsilon > 0$, $\forall n > N_\epsilon P(\|X_n\| \leq M_\epsilon) > 1 - \epsilon$.

By the continuity of $Y = g(X)$, $M_{g\epsilon} = \max\{\|Y\| : \|X\| \leq M_\epsilon\}$ exists.

So $\|X_n\| \leq M_\epsilon \implies g(X_n) \leq M_{g\epsilon}$. When $n > N_\epsilon$,

$P(\|g(X_n)\| \leq M_{g\epsilon}) \geq P(\|X_n\| \leq M_\epsilon) > 1 - \epsilon$.

Hence $g(X_n)$ is bounded in probability. $\square$

(2) X_n is bounded in probability $\iff X_{n_k}$ is bounded in probability $\forall n_k$.

Proof. $\Leftarrow$: X_n as its subsequence is bounded in probability.

$\Rightarrow$: X_n is bounded in probability. So

$\forall \epsilon > 0 \exists M_\epsilon > 0$ and $N_\epsilon > 0$, $\forall n > N_\epsilon P(\|X_n\| \leq M_\epsilon) > 1 - \epsilon$.

Thus $\forall n_k \geq N_\epsilon$, $P(\|X_{n_k}\| \leq M_\epsilon) > 1 - \epsilon$.

Hence X_{n_k} is bounded in probability. $\square$

(3) $X_n = \begin{pmatrix} X_{1n} \\ \vdots \\ X_{kn} \end{pmatrix}$ is bounded in probability $\iff$ in $\begin{pmatrix} X_{1n} \\ \vdots \\ X_{kn} \end{pmatrix}$ X_{in} is bounded in probability for all $i = 1, \dots, k$

Proof. $\Rightarrow$: With continuous $g_i(X) = e'_i X$, $X_{in} = g_i(X_n)$ is bounded in probability for all $i = 1, \dots, k$.

$\Leftarrow$: We show $\Leftarrow$ by induction on k .

When $k = 1$, $X_n = X_{1n}$ is bounded in probability. So $\Leftarrow$ holds for $k = 1$.

Assume that $\Leftarrow$ is true for $k = t$. Consider the case of $k = t + 1$.

Write $X_n = \begin{pmatrix} X_{1n} \\ \vdots \\ X_{tn} \\ X_{t+1,n} \end{pmatrix} = \begin{pmatrix} X_{In} \\ X_{t+1,n} \end{pmatrix}$. By given condition and induction assumption both X_{In} and $X_{t+1,n}$ are bounded in probability. So

$\forall \epsilon > 0$, $\exists M_{1\epsilon} > 0$ and $N_{1\epsilon} > 0$, $\forall n > N_{1\epsilon} P(\|X_{In}\| \leq M_{1\epsilon}) > 1 - \epsilon/2$.

$\exists M_{2\epsilon} > 0$ and $N_{2\epsilon} > 0$, $\forall n > N_{2\epsilon} P(\|X_{t+1,n}\| \leq M_{2\epsilon}) > 1 - \epsilon/2$.

Let $M_\epsilon = \sqrt{M_{1\epsilon}^2 + M_{2\epsilon}^2}$ and $N_\epsilon = \max\{N_{1\epsilon}, N_{2\epsilon}\}$.

With $A = [\|X_{In}\| \leq M_{1\epsilon}]$ and $B = [\|X_{t+1,n}\| \leq M_{2\epsilon}]$,

by $\|X_n\|^2 = \|X_{In}\|^2 + \|X_{t+1,n}\|^2$, $A \cap B \subset [\|X_n\| \leq M_\epsilon]$. When $n > N_\epsilon$,

$$\begin{aligned} P(\|X_n\| \leq M_\epsilon) &\geq P(A \cap B) = P(A) + P(B) - P(A \cup B) \\ &\geq (1 - \epsilon/2) + (1 - \epsilon/2) - P(A \cup B) = 1 - \epsilon + [1 - P(A \cup B)] \\ &\geq 1 - \epsilon. \end{aligned}$$

So $\Leftarrow$ holds for $k = t + 1$. $\Leftarrow$ is established. $\square$

Ex1: When $X_n \xrightarrow{p} X \sim (\mu, \Sigma)$, by (2) of 2 in L11, X is bounded in probability. By

Ex in L11, X_n is bounded in probability. By (3), $\begin{pmatrix} X_n \\ X \end{pmatrix}$ is bounded in probability.

2. Big O notation

(1) $O_p(1)$

$X_n = O_p(1) \stackrel{def}{\iff} X_n \in R^k$ is bounded in probability.

(2) $O_p(Y_n)$

$X_n = O_p(Y_n) \iff X_n \in R^k, 0 \neq Y_n \in R$ and $\frac{X_n}{Y_n} = O_p(1)$.

(3) $O(1)$

$a_n = O(1) \stackrel{def}{\iff}$ Non-random $a_n \in R^k$ is bounded
 $\iff \exists M > 0$ such that $\forall n > 0 \|a_n\| < M$
 $\implies a_n = O_p(1)$.

Ex2: $X_n = O_p(1) \iff g(X_n) = O_p(1)$ for all continuous $g(\cdot)$.

$X_n = O_p(1) \iff X_{n_k} = O_p(1)$ for all subsequence X_{n_k} .

For $X_n = \begin{pmatrix} X_{1n} \\ \vdots \\ X_{kn} \end{pmatrix}$,

$X_n = O_p(1) \iff X_{in} = O_p(1)$ for all $i = 1, \dots, k$.

Ex3: $X_n - Y_n = O_p(a_n) \implies \frac{X_n - Y_n}{a_n} = O_p(1)$.

3. Little o notation

(1) $o_p(1)$

$X_n = o_p(1) \stackrel{def}{\iff} X_n \in R^k$ and $X_n \xrightarrow{p} 0$.

(2) $o_p(Y_n)$

$X_n = o_p(Y_n) \stackrel{def}{\iff} X_n \in R^k, 0 \neq Y_n \in R$ and $\frac{X_n}{Y_n} = o_p(1) \iff \frac{X_n}{Y_n} \longrightarrow 0$

(3) $o(1)$

$a_n = o(1) \stackrel{def}{\iff}$ Non-random $a_n \in R^k$ and $a_n \longrightarrow 0 \implies a_n = o_p(1)$.

Ex4: $X_n \xrightarrow{p} X \in R \iff X_n - X = o_p(1)$

$X_n \xrightarrow{p} X \in R^k \iff \|X_n - X\| = o_p(1)$

$X_n = o_p(1) \implies X_n = O_p(1)$.