ORF524 - Problem Set 5

Bachir EL KHADIR

December 4, 2015

Problem 1

By the triangle inequality $||X|| - ||X_n||_p < ||X - X_n||_p$, and $||X_n||_p - ||X||_p < ||X - X_n||_p$, so $|||X_n|| - ||X||_p| < ||X - X_n||_p \to 0$

The converse is not true. Take $X_n = (-1)^n$, and $X = 1$. $||X_n||_p = 1 = ||X||_p$, but

$$||X_n - X||_p = \begin{cases} 2 & \text{if } n \text{ even} \\ 0 & \text{otherwise} \end{cases}$$

doesn't converge to 0.

Problem 2

1. (a.s) Let's suppose $X_n \to X$ (resp. $Y_n \to Y$ as), and Ω_x (resp. Ω_y) the set where it holds. Then $P(\Omega_x \cap \Omega_y) = 1$ and for $\omega \in \Omega_x \cap \Omega_y$, $X_n(\omega) + Y_n(\omega) \to X(\omega) + Y(\omega)$

   Let's now suppose that $X_n \to X$ and $Y_n \to Y$ in probabolity. Let $\epsilon > 0$, since $|X_n + Y_n - X - Y| \leq |X_n - X| + |Y_n - Y|$, then:

$$\begin{aligned} P(|X_n + Y_n - X - Y| > \epsilon) &\leq P(|X_n - X| + |Y_n - Y| > \epsilon) \\ &\leq P(|X_n - X| > \epsilon) + P(|Y_n - Y| > \epsilon) \\ &\to 0 \end{aligned}$$

   (p) Let's suppose $X_n \to X$ (resp. $Y_n \to Y$ as), and Ω_x (resp. Ω_y) the set where it holds. Then $P(\Omega_x \cap \Omega_y) = 1$, and for $\omega \in \Omega_x \cap \Omega_y$, $X_n(\omega)^T Y_n(\omega) \to X(\omega)^T Y(\omega)$

   Let's now suppose that $X_n \to X$ and $Y_n \to Y$ in probabolity. Let $\epsilon > 0$, since

   (L_p) By the triangular inequality: $||X_n + Y_n - (X + Y)||_p \leq ||X_n - X||_p + ||Y_n - Y||_p \to_n 0$

2. •

$$\begin{aligned} |X_n^T Y_n - X^T Y| &\leq |X_n^T Y_n - X_n^T Y + X_n^T Y - X^T Y| \\ &\leq |X_n^T (Y_n - Y)| + |(X_n - X)^T Y| \\ &\leq |(X_n - X)^T (Y_n - Y)| + |X^T (Y_n - Y)| + |(X_n - X)^T Y| \\ &\leq |X_n - X||Y_n - Y| + |X||Y_n - Y| + |X_n - X||Y| && \text{Cauchy Shwartz} \end{aligned}$$

   , then:

$$P(|X_n^T Y_n - X^T Y| > \epsilon) \leq P(|X_n - X||Y_n - Y| > \epsilon) + P(|X||Y_n - Y| > \epsilon) + P(|X_n - X||Y| > \epsilon)$$

   Let's show that each term in the RHS converges to 0 when n goes to infinity. Indeed:

   – $P(|X_n - X||Y_n - Y| > \epsilon) \leq P(|X_n - X| > \sqrt{\epsilon}) + P(|Y_n - Y| > \sqrt{\epsilon}) \to 0$

- For $A > 0$, $|X| \leq A \wedge |Y_n - Y| \leq \frac{\epsilon}{A} \Rightarrow |X||Y_n - Y| < \epsilon$, so $\{|X||Y_n - Y| > \epsilon\} \subset \{|X| > A\} \cup \{|Y_n - Y|\} > \frac{\epsilon}{A+1}\}$, so (since $|X| < \infty$ a.s.):

$$P(|X||Y_n - Y| > \epsilon) \leq P(|X| > A) + P(|Y_n - Y| > \frac{\epsilon}{A} \rightarrow_n P(|X| > A) \rightarrow_A 0$$

- Same for $P(|Y||X_n - X| > \epsilon)$

- - If $Y_n \rightarrow Y$ on Ω_y of size one, then for $w \in \Omega_y \cap Y \neq 0$ (which is also of size 1) we have that for n large enough $|Y_n - Y|(\omega) < \frac{Y(\omega)}{2}$, and therefore $Y_n(\omega) \neq 0$ and $\frac{1}{Y_n}(\omega) \rightarrow \frac{1}{Y}(\omega)$. eg $\frac{1}{Y_n} \rightarrow \frac{1}{Y}$ a.s. and we can use the last question to prove that

$$\frac{X_n}{Y_n} \rightarrow \frac{X}{Y} \text{ a.s.}$$

  - We suppose Y_n is bounded from below in probability (this is a necessary condition since: $\frac{1}{Y_n}$converges $\Rightarrow \frac{1}{Y_n}$is bounded in probability $\Rightarrow Y_n$ bounded from below in probability
    Let $\alpha > 0$, $\epsilon > 0$, and n large enough so that $P(|Y - Y_n| > \epsilon) < \alpha$.
    Since $Y_n Y$ is bounded from below in probability, $\exists \delta > 0 \forall n P(|Y_n Y| < \delta) < \alpha$
    We have that $|\frac{1}{Y_n} - \frac{1}{Y}| = \frac{|Y - Y_n|}{|Y_n Y|}$, so:

$$\begin{aligned} P(|\frac{1}{Y_n} - \frac{1}{Y}| > \epsilon) &\leq P(|Y - Y_n| > \epsilon |Y_n Y|) \\ &= P(|Y - Y_n| > \epsilon |Y_n Y|, |Y_n Y| \leq \delta) + P(|Y - Y_n| > \epsilon |Y_n Y|, |Y_n Y| > \delta) \\ &\leq \alpha + P(|Y - Y_n| > \epsilon) \leq 2\alpha \end{aligned}$$

    And therefore we have convergence in probability.

- Take $\forall n X_n = Y_n = X = -Y \sim \mathcal{N}(0, 1)$, then:
  - For a), $V(X_n + Y_n) = 2$, but $Var(X + Y) = 0$.
  - For b), $E[X_n Y_n] = E[\mathcal{N}(0, 1)] = 1$, but $E[XY] = -1$
  - For c), $E[\frac{X_n}{Y_n}] = 1, E[\frac{X}{Y}] = -1$

Problem 3

- $\forall \epsilon > 0 \, P(|X_n - X| > \epsilon) = P(|X_n - X|^p > \epsilon^p) \leq \frac{E[|X_n - X|^p]}{\epsilon^p} \rightarrow 0$

- $\Rightarrow$ Assume for every subsequence $X_{n_k} \rightarrow X$. (X_n) is trivially a subsequence of (X_n), so $X_n \rightarrow X$

  $\Leftarrow$ Assume for X_n converges to X, and let $(X_{n_k})_k$ be a subsequence.

  * a.s.: Let Ω the set of measure one in which the convergence holds. Let $\omega \in \Omega$ and $\epsilon > 0$, There exist $N \in \mathbb{N}$ such that for all $n > N$, $|X_n(\omega) - X(\omega)| < \epsilon$, in particular, since $n_k > k$, when $k > N$, $|X_{n_k}(\omega) - X(\omega)| < \epsilon$, and thus $X_{n_k}(\omega) \rightarrow X(\omega)$
  * in probability: $\epsilon > 0$, and $\delta > 0$. There exist $N \in \mathbb{N}$ such that for all $n > N$, $P(|X_n - X| > \epsilon) < \delta$, in particular when $k > N$, $P(|X_{n_k} - X| > \epsilon) < \delta$, and thus $\forall \epsilon P(|X_{n_k} - X| > \epsilon) \rightarrow 0$
  * In distribution:

- if $x \neq c$,

$$F_n(x) \rightarrow_n \begin{cases} 1 & \text{if } x > c \\ 0 & \text{if } x < c \end{cases}$$

$$\forall \epsilon > 0 \, P(|X_n - c| > \epsilon) = E[1_{|X_n - c| > \epsilon}] = E[1_{X_n - c > \epsilon} + 1_{X_n - c < -\epsilon}] \leq (1 - F_n(\epsilon + c)) + F_n(c - \epsilon) \rightarrow 0$$

Problem 4

- F is invertible so it is increasing and:

$$P(X^* \leq x) = P(F^{-1}(u) \leq x) = P(u \leq F(x)) = F^(x) = P(X \leq x)$$

-

$$\begin{aligned} X^* \leq x &\Rightarrow F^{-1}(U) \leq x \\ &\Rightarrow \forall \epsilon > 0,\, U \leq F(x+\epsilon) \\ &\Rightarrow U \leq F(x) && \text{(by right continuty)} \end{aligned}$$

So $P(X^* \leq x) \leq P(U \leq F(x)) = F(x)$

By definition of the sup, there exist a sequence $x_n \to F^{-1}(F(x))$ so that $F(x_n) < F(x)$. Since F is increasing $x_n < x$, and therefore by going to the limit $F^{-1}(F(x)) \leq x$ Since F^{-1} is increasing (because it is taking the sup over larger sets):

$$U < F(x) \Rightarrow F^{-1}(U) \leq F^{-1}(F(x)) \leq x$$

So $F(x) = P(U < F(x)) \leq X^* \leq x)$

c/c: F is the cdf of X^*, so X^* has the same distribution of X.

Problem 5

Let $u \in R^d$,

- Lemma: If for all $i \leq d$ $X_n^{(i)}$(the i-th component of X) converges to $X^{(i)}$ in probability, then X_n converges to X in probability.

  Proof: $P(|X_n - X| > \epsilon) \leq P(\sum_i |X_n^{(i)} - X^{(i)}| > \epsilon) \to 0$

$$\begin{aligned} X_n \xrightarrow{a.s.} X &\Rightarrow (\forall u \in R^d)\, u^T X_n \xrightarrow{a.s.} u^T X && \text{By Ex2, b)} \\ &\Rightarrow (\forall u \in R^d)\, u^T X_n \xrightarrow{p} u^T X && \Rightarrow X_n \xrightarrow{p} X \\ &\Rightarrow (\forall u \in R^d)\, u^T X_n \xrightarrow{D} u^T X && \Rightarrow X_n \xrightarrow{D} X \end{aligned}$$

- Let $X_n \to X$ in probability, and X_{n_k} be a subsequence. Let prove by induction on the dimension d of X_n, that there exist a subsequence $X_{n_{k_j}}$ such that $X_{n_{k_j}} \to X$ a.s.

  The case $d = 1$ was proven in class.

  For $d > 1$, let's suppose the induction property true for $d-1$, and let's denote $X_n = \begin{pmatrix} X_n^{(1)} \\ X_n^{(-1)} \end{pmatrix}$, where $X_n^{(1)}$ has dimension 1, and $X_n^{(-1)}$ has dimension $n-1$.

  Since $X_n^{(-1)} \xrightarrow{p} X^{(-1)}$, there exist a subsequence $X_{n_{k_j}}^{(-1)} \xrightarrow{a.s}_j X^{(-1)}$.

  And since $X_{n_{k_j}}^{(1)} \xrightarrow{p}_j X^{(1)}$ we have also that there exist a subsequence $X_{n_{k_{j_r}}}^{(1)} \xrightarrow{a.s}_j X^{(1)}$

  c/c:

$$X_{n_{k_{j_r}}} \xrightarrow{a.s}_j X$$

  And thus the induction proof is complete.

Problem 6

- $P(|X| > a) \to_a 1$, Let Let $\epsilon > 0$ and a large enough so that $P(|X| > a) < \epsilon$.

  $X_n \xrightarrow{D} X$, Let $N \in N$, st for $n > N$, $|E[1_{|X_n|>a}] - E[1_{|X|>a}]| < \epsilon$.

  Then $|E[1_{|X_n|>a}]| \leq \epsilon + |E[1_{|X|>a}]| \leq 2\epsilon$.

  As a result $X_n = O_p(1)$

- $||Y_n^T X_n|| \leq ||Y_n|| ||X_n|| = o_p(1)$ because $||Y_n|| = o_p(1)$ and $||X_n|| = O_p(1)$ and are 1-dimensional.

Problem 7

Since convergence a.s and in probability imply convergence in distribution, we only need to show that:

$$X_n \xrightarrow{D} X \Rightarrow E[g(X_n)] \to E[g(X)]$$

By **** representation theorem, there exist $Y_n \overset{D}{\sim} X_n$, $Y \overset{D}{\sim} X$ such that $Y_n \to Y$ a.s. Let Ω be the set where g is continuous, we know that $P_Y(\Omega) = P_X(\Omega) = 1$. Let $\omega \in \Omega$, $t_o = Y(\omega)$ and $\epsilon > 0$.

- g is continuous on t_0, so there exist a $\delta > 0$, so that for $|t - t_0| < \delta$, $|g(t) - g(t_0)| < \epsilon$.
- Let $N \in \mathbb{N}$ so that for $n > N$ $|Y_n(\omega) - Y(\omega)| < \delta$, and therefore $g(Y_n(\omega)) - g(Y(\omega))| < \epsilon$.

As a conclusion $g(Y_n) \to g(Y)$ a.s. Since g is bounded, by dominated convergence theorem we have that: $E[g(Y_n)] \to E[g(Y)]$, and since $E[g(Y_n)] = E[g(X_n)]$ and $E[g(Y)] = E[g(X)]$, we have the result.

Problem 8

- Let

  $$y_i = \begin{pmatrix} x_i \\ x_i^2 \end{pmatrix}$$

  $(y_i)_i$ are iid L_2 integrable r.v, $E[y_i] = (E[X], E[X^2])$

  $$Cov(y_i) = \begin{pmatrix} Var(X) & Cov(X, X^2) \\ Cov(X, X^2) & Var(X^2) \end{pmatrix}$$

  by TCL

  $$\sqrt{n}(\bar{y}_n - E[y]) \xrightarrow{D} \mathcal{N}(0, Cov(Y))$$

  - $S_n^2 = \frac{1}{n} \sum x_i^2 - \bar{x}^2 = g(\bar{y})$, where $g : (u, v) \to v - u^2$, $\nabla g = \left(\begin{smallmatrix} -2u \\ 1 \end{smallmatrix} \right)$
  - $g(E[y]) = \sigma^2$
  - By the continuous mapping theorem: $\sqrt{n}(S_n^2 - \sigma^2) \to \mathcal{N}(0, \nabla g_{E[y]}^T Cov(Y) \nabla g_{E[y]})$
  - Since $\frac{n}{n-1} \to 1$, by Slutsky:
    $\sqrt{n}(S_{n-1}^2 - \frac{n}{n-1}\sigma^2) \to \mathcal{N}(0, \nabla g_{E[y]}^T Cov(Y) \nabla g_{E[y]})$
  - Since $\sqrt{n}(\sigma - \frac{n}{n-1}\sigma) \to 0$, by Slutsky:
    $\sqrt{n}(S_{n-1}^2 - \sigma^2) = \sqrt{n}(S_{n-1}^2 - \frac{n}{n-1}\sigma^2) + \sqrt{n}(\sigma - \frac{n}{n-1}\sigma) \to \mathcal{N}(0, \nabla g_{E[y]}^T Cov(Y) \nabla g_{E[y]})$

- $z_i = \begin{pmatrix} x_i \\ y_i \\ x_i y_i \end{pmatrix}$ iid L_2, so $\sqrt{n}(\bar{z}_i - \begin{pmatrix} E[X] \\ E[Y] \\ E[XY] \end{pmatrix} \to \mathcal{N}(0, Cov(Z))$

  $\hat{C}_n = f(\bar{z})$ where $f(u, v, t) = t - uv$ is differentiable and: $\nabla f = \begin{pmatrix} -v \\ -u \\ 1 \end{pmatrix}$ is continuous

  $\sqrt{n}(\hat{C}_n - Cov(x_1, y_1)) \to \mathcal{N}(0, \nabla f^T cov(Z) \nabla f)$

$$cov(Z) = \begin{pmatrix} Var(X) & Cov(X,Y) & Cov(X,XY) \\ Cov(X,Y) & Var(Y) & Cov(Y,XY \\ Cov(X,XY) & Cov(Y,XY) & Var(XY) \end{pmatrix}$$