Most of the results discussed in LN10 extend fairly easily to random vectors in \( \mathbb{R}^k \). As before, if \( X \) is a \( k \)-dimensional random vector, then \( E[X] \) is the \( k \)-vector of expected values of the components of \( X \), and \( V[X] \) is the \( k \times k \) symmetric positive definite matrix of variances and covariances.

For vectors \( a = (a_1, a_2, \ldots, a_k)' \) and \( b = (b_1, b_2, \ldots, b_k)' \), let \( a \leq b \) mean that each component of \( a \) is less than or equal to each corresponding component of \( b \):

\[
a_j \leq b_j \quad j = 1, 2, \ldots, k.
\]

Then we can define the CDF of the random vector \( X \) as:

\[
F_X(x) = \Pr(X \leq x), \quad x \in \mathbb{R}^k.
\]

Also, for a vector \( a = (a_1, \ldots, a_k) \), let

\[
||a|| = \left( \sum_{j=1}^{k} a_j^2 \right)^{1/2}
\]

be the usual Euclidean norm. Note that \( ||a|| = \sqrt{a' a} \).

With this notation, we can define convergence in probability, convergence almost surely, and convergence in distribution in a similar way to the scalar case:

**Convergence in Probability**: a sequence of random vectors \( \{X_n\} \) converges in probability to vector \( a \) if for all \( \epsilon > 0 \),

\[
P(||X_n - a|| > \epsilon) \to 0,
\]
as \( n \to \infty \). We usually write this as \( X_n \xrightarrow{p} a \) or \( \text{plim}_{n \to \infty} X_n = a \).

**Convergence Almost Surely**: a sequence of random vectors \( \{X_n\} \) converges almost surely to \( a \) if

\[
P\left( \lim_{n \to \infty} ||X_n - a|| = 0 \right) = 1.
\]

We usually write this as \( X_n \xrightarrow{as} a \).

As with the scalar case, convergence almost surely implies convergence in probability, but the converse does not hold.
**Convergence in Distribution:** a sequence of random variables \( \{X_n\} \) is said to converge in distribution to a random variable \( X \) if

\[
P(X_n \leq x) \longrightarrow P(X \leq x)
\]

at every point \( x \) at which the limit distribution function \( P(X \leq x) \) is continuous. This will be denoted by \( X_n \overset{d}{\longrightarrow} X \).

**Basic Asymptotic Results:**

**Weak Law of Large Numbers (WLLN):** Let \( X_1, X_2, \ldots \) be a sequence of i.i.d. random variables such that \( E(||X_1||) < \infty \). Then

\[
\overline{X}_n \equiv \frac{1}{n} \sum_i X_i \overset{p}{\longrightarrow} E(X_1).
\]

**Strong Law of Large Numbers (SLLN):** Let \( X_1, X_2, \ldots \) be a sequence of i.i.d. random variables such that \( E(||X_1||) < \infty \). Then

\[
\overline{X}_n \equiv \frac{1}{n} \sum_i X_i \overset{as}{\longrightarrow} E(X_1).
\]

**Multivariate Central Limit Theorem (CLT):** Let \( X_1, X_2, \ldots \) be i.i.d. random vectors in \( \mathbb{R}^k \) with \( E(||X_1||^2) < \infty \). Let \( \mu = E[X_1] \) and \( \Sigma = V[X_1] \). Then

\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu) = \sqrt{n}(\overline{X}_n - \mu) \overset{d}{\longrightarrow} N(0, \Sigma).
\]

**Continuous Mapping Theorem (CMT):** Let \( g(\cdot) \) be a function from \( \mathbb{R}^k \) to \( \mathbb{R}^m \), and suppose \( g \) is continuous at every point in a set \( C \) s.t. \( P(X \in C) = 1 \). Then

(i) If \( X_n \overset{p}{\longrightarrow} X \), then \( g(X_n) \overset{p}{\longrightarrow} g(X) \);

(ii) If \( X_n \overset{d}{\longrightarrow} X \), then \( g(X_n) \overset{d}{\longrightarrow} g(X) \).

Note that matrix addition and multiplication are continuous functions.
There are various useful facts about convergence in probability and convergence in distribution:

**Result:**

(i) Convergence in probability implies convergence in distribution:

\[ X_i \xrightarrow{p} X \Rightarrow X_i \xrightarrow{d} X. \]

(ii) Convergence in distribution to a constant, implies convergence in probability:

\[ X_i \xrightarrow{d} a \quad \text{(a constant)} \Rightarrow X_i \xrightarrow{p} a. \]

(iii) If \( X_n \xrightarrow{d} X \) and \( ||X_n - Y_n|| \xrightarrow{p} 0 \), then \( Y_n \xrightarrow{d} X \).

(iv) If \( X_n \xrightarrow{d} X \) and \( Y_n \xrightarrow{p} a \) where \( a \) is a constant, then the vector \( (X_n, Y_n) \xrightarrow{d} (X, a) \).

(v) If \( X_n \xrightarrow{p} X \) and \( Y_n \xrightarrow{p} Y \), then \( (X_n, Y_n) \xrightarrow{p} (X, Y) \). (This is not true for convergence in distribution.)

A useful corollary of the previous result:

**Slutsky’s Lemma:** Let \( X_n, X, \) and \( Y_n \) be random vectors or matrices. If \( X_n \xrightarrow{d} X \) and \( Y_n \xrightarrow{p} c \), where \( c \) is a conformable constant vector or matrix, then

(i) \( X_n + Y_n \xrightarrow{d} X + c; \)

(ii) \( Y_n X_n \xrightarrow{d} cX; \)

(iii) \( Y_n^{-1} X_n \xrightarrow{d} c^{-1} X, \) provided \( c \) is invertible.

**Delta Method:** Let \( X_n \) be a sequence of \( d \)-dimensional random vectors such that

\[ \sqrt{n}(X_n - \mu) \xrightarrow{d} N(0, \Sigma), \]

where \( \Sigma \) is positive definite and finite. Let \( g \) denote a continuously differentiable function from \( \mathbb{R}^d \) into \( \mathbb{R}^k \), and let \( G(x) = \partial g / \partial x \) denote the \( k \times d \) matrix of partial derivatives. Then

\[ \sqrt{n}(g(X_n) - g(\mu)) \xrightarrow{d} N(0, G(\mu) \Sigma G(\mu)^\prime). \]