Markov chains: construction of the path space

Mathav Murugan

February 11, 2020

Definition 0.1. Let \((S, \mathcal{S})\) be a measurable space. We say that \(p : S \times S \to [0, 1]\) is a transition probability on \((S, \mathcal{S})\) if

1. for each \(x \in S\), \(A \mapsto p(x, A)\) is a probability measure on \((S, \mathcal{S})\);
2. for each \(A \in \mathcal{S}\), \(x \mapsto p(x, A)\) is a \((S, \mathcal{S})\)-measurable function.

Let \((\Omega, \mathcal{F}, P)\) be a probability space and let \((\mathcal{F}_n)_{n \in \mathbb{Z}_+}\) be a filtration. A \((\mathcal{F}_n)\)-adapted \(S\)-valued stochastic process \(X = \{X_n : n \in \mathbb{Z}_+\}\) is a time-homogeneous Markov chain with respect to \((\mathcal{F}_n)\) with transition probability \(p\) if it satisfies the following Markov property: for all \(n \in \mathbb{Z}_+\) and for all \(A \in \mathcal{S}\), we have

\[ P(X_{n+1} \in A | \mathcal{F}_n)(\omega) = p(X_n(\omega), A) \quad \text{a.s.} \quad (0.1) \]

The probability distribution \(\mu\) of \(X_0\) is called the initial distribution of \(X\). Note that the above definition implies that \(X_n : (\Omega, \mathcal{F}_n) \to (S, \mathcal{S})\) is measurable for all \(n \in \mathbb{Z}_+\).

Exercise: Show that if \(X\) is a Markov chain with respect to \((\mathcal{F}_n)_{n \in \mathbb{Z}_+}\), then it is also a Markov chain with respect to \((\mathcal{F}_n^X)_{n \in \mathbb{Z}_+}\) (this is the filtration associated to \(X\)).

An equivalent definition of Markov property is the following (use FMCT to show the equivalence): for all \(f \in b\mathcal{S}\)

\[ E(f(X_{n+1}) | \mathcal{F}_n) = \int_{S} f(y) p(X_n, dy), \quad \text{a.s.} \quad (0.2) \]

The first issue is the existence of Markov chains. Given a transition probability \(p\) and an initial distribution \(\mu\), how to construct a probability space \((\Omega, \mathcal{F}, P)\), a filtration \((\mathcal{F}_n)\), a stochastic process \(X = \{X_n : n \in \mathbb{Z}_+\}\) that is a Markov chain with respect to \((\mathcal{F}_n)\)?

The fundamental existence theorem for probability measures is Carathéodory’s extension theorem (See Theorem A.1.3 in the text).

Theorem 0.2 (Carathéodory’s extension theorem). Let \(\mathcal{A}\) be an algebra on \(\Omega\). Let \(\mu : \mathcal{A} \to [0, \infty]\) be \(\sigma\)-finite measure on \(\mathcal{A}\) (that is, a countably additive set function on \(\mathcal{A}\)). Then there exists an unique extension \(\tilde{\mu} : \sigma(\mathcal{A}) \to [0, \infty]\) of \(\mu\) such that \(\tilde{\mu}\) is a measure on \((\Omega, \sigma(\mathcal{A}))\). (Here extension means that \(\tilde{\mu}(A) = \mu(A)\) for all \(A \in \mathcal{A}\).)
We recall the definition of product $\sigma$-field.

**Definition 0.3.** Let $(\Omega_i, \mathcal{F}_i)$ be a family of measurable spaces for $i \in I$. Let $\prod_{i \in I} \Omega_i$ denote the product space (cartesian product of sets). For $i \in I$, let $\pi_i : \prod_{j \in I} \Omega_j \rightarrow \Omega_i$ denote the natural projection map (projection on to the ‘$i$-th component’). A **measurable rectangle** in $\prod_{i \in I} \Omega_i$ is a set of the form

$$\cap_{i \in F} \pi_i^{-1}(A_i),$$

where $F$ is a finite subset of $I$, and $A_i \in \mathcal{F}_i$ for all $i \in F$ (Exercise: The set of measurable rectangles forms a $\pi$-system). The smallest $\sigma$-field containing all the measurable rectangles is called the **product $\sigma$-field** and is denoted by $\prod_{i \in I} \mathcal{F}_i$ (note that this is not the cartesian product of $\mathcal{F}_i$ as sets). It is easy to check that $\prod_{i \in I} \mathcal{F}_i$ is the smallest $\sigma$-field $\mathcal{F}$ on $\prod_{i \in I} \Omega_i$ such that the projection maps $\pi_i : (\prod_{i \in I} \Omega_i, \mathcal{F}) \rightarrow (\Omega_i, \mathcal{F}_i)$ is measurable for all $i \in I$.

The following lemma is often used to prove uniqueness of measures.

**Lemma 0.4.** Let $\mathcal{P}$ be a $\pi$-system on $\Omega$, and let $P$ and $Q$ be two probability measures on $(\Omega, \sigma(\mathcal{P}))$. If $P(A) = Q(A)$ for all $A \in \mathcal{P}$, then

$$P(A) = Q(A) \quad \text{for all } A \in \sigma(\mathcal{P}).$$

**Proof.** Define

$$\mathcal{L} = \{A \in \sigma(\mathcal{P}) : P(A) = Q(A)\}.$$

It is easy to check that $\mathcal{L}$ is a $\lambda$-system such that $\mathcal{L} \supseteq \mathcal{P}$. The conclusion $\mathcal{L} = \sigma(\mathcal{P})$ follows from $\pi$-$\lambda$ theorem. \qed

The following existence theorem is a basic building block for the construction of Markov chains. Our exposition is based on [Ash, Chapter 2].

**Theorem 0.5.** Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ be a probability space and let $(\Omega_2, \mathcal{F}_2)$ be a measurable space. Let $\mu_2 : \Omega_1 \times \mathcal{F}_2 \rightarrow [0, 1]$ be such that

(a) $A \mapsto \mu_2(\omega_1, A)$ is a probability measure on $(\Omega_2, \mathcal{F}_2)$ for all $\omega_1 \in \Omega_1$.

(b) $\omega_1 \mapsto \mu_2(\omega_1, A)$ is a $(\Omega_1, \mathcal{F}_1)$-measurable function for all $A \in \mathcal{F}_2$.

Then there is a unique probability measure $\mu$ on the product measurable space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ such that

$$\mu(A \times B) = \int_A \mu_2(\omega_1, B) \mu_1(d\omega_1), \quad \text{for all } A \in \mathcal{F}_1, B \in \mathcal{F}_2.$$

This measure $\mu$ is given by

$$\mu(F) = \int_{\Omega_1} \mu_2(\omega_1, F(\omega_1)) \mu_1(d\omega_1), \quad \text{for all } F \in \mathcal{F}_1 \times \mathcal{F}_2,$$

(0.3)

where $F(\omega_1)$ denotes the **section** of $F$ at $\omega_1$:

$$F(\omega_1) = \{\omega_2 \in \Omega_2 : (\omega_1, \omega_2) \in F\}.$$

(Warning: $\mathcal{F}_1 \times \mathcal{F}_2$ is the product $\sigma$-field, and not the cartesian product of $\mathcal{F}_1$ and $\mathcal{F}_2$.)
Proof. The uniqueness of \( \mu \) follows immediately from Lemma 0.4 (by letting \( \mathcal{P} \) to be the \( \pi \)-system of measurable cylinders). The remainder of the proof is devoted to the existence.

We claim that

\[
F(\omega_1) \in \mathcal{F}_2 \text{ for all } \omega_1 \in \Omega_1 \text{ and for all } F \in \mathcal{F}_1 \times \mathcal{F}_2. \tag{0.4}
\]

Define \( \mathcal{L}_1 = \{ F \in \mathcal{F}_1 \times \mathcal{F}_2 : F(\omega_1) \in \mathcal{F}_2 \text{ for all } \omega_1 \in \Omega_1 \} \). Then \( \mathcal{L}_1 \) is \( \lambda \)-system that contains all the measurable rectangles (why?). By \( \pi \)-\( \lambda \) theorem, we have \( \mathcal{L}_1 = \mathcal{F}_1 \times \mathcal{F}_2 \). This completes the proof of (0.4) \( \square \)

Our next claim is that

\[
\omega_1 \mapsto \mu_2(\omega_1, F(\omega_1)) \text{ is a measurable function on } (\Omega_1, \mathcal{F}_1), \text{ for all } F \in \mathcal{F}_1 \times \mathcal{F}_2. \tag{0.5}
\]

Note that \( \mu_2(\omega_1, F(\omega_1)) \in [0, 1] \) for all for all \( F \in \mathcal{F}_1 \times \mathcal{F}_2 \) and for all \( \omega_1 \in \Omega_1 \) by (0.4). As above, we define

\[
\mathcal{L}_2 = \{ F \in \mathcal{F}_1 \times \mathcal{F}_2 : \omega_1 \mapsto \mu_2(\omega_1, F(\omega_1)) \text{ is a measurable function on } (\Omega_1, \mathcal{F}_1) \}.
\]

If \( F = A_1 \times A_2 \) where \( A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2 \), we have \( \mu_2(\omega_1, F(\omega_1)) = \mu_2(\omega_1, A_2)1_{A_1}(\omega_1) \). Hence \( \omega_1 \mapsto \mu_2(\omega_1, F(\omega_1)) \) is \( (\Omega_1, \mathcal{F}_1) \)-measurable, as it is a product of measurable functions \( \omega_1 \mapsto \mu_2(\omega_1, A_2) \). Therefore \( \mathcal{L}_2 \) contains all the measurable rectangles. \( \mathcal{L}_2 \) is a \( \lambda \)-system (Exercise). By \( \pi \)-\( \lambda \) theorem we conclude \( \mathcal{L}_2 = \mathcal{F}_1 \times \mathcal{F}_2 \). Hence we obtain (0.5).

Define

\[
\mu(F) = \int_{\Omega_1} \mu_2(\omega_1, F(\omega_1)) \mu_1(d\omega_1), \text{ for all } F \in \mathcal{F}_1 \times \mathcal{F}_2.
\]

By (0.5), the above integral exists and belongs to \([0, 1]\). To prove that \( \mu \) is a measure, consider \( F_1, F_2, \ldots \) denote a sequence of pairwise disjoint sets in \( \mathcal{F}_1 \times \mathcal{F}_2 \). Then

\[
\mu \left( \bigcup_{n=1}^{\infty} F_n \right) = \int_{\Omega_1} \mu_2(\omega_1, \bigcup_{n=1}^{\infty} F_n(\omega_1)) \mu_1(d\omega_1)
\]

\[
= \int_{\Omega_1} \sum_{n=1}^{\infty} \mu_2(\omega_1, F_n(\omega_1)) \mu_1(d\omega_1) \quad \text{(since } (F_n)_{n \geq 1} \text{ are disjoint)}
\]

\[
= \sum_{n=1}^{\infty} \int_{\Omega_1} \mu_2(\omega_1, F_n(\omega_1)) \mu_1(d\omega_1) \quad \text{(by Monotone convergence theorem)}
\]

\[
= \sum_{n=1}^{\infty} \mu(F_n),
\]

proving that \( \mu \) is a measure. If \( A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2 \), we have

\[
\mu(A_1 \times A_2) = \int_{\Omega_1} \mu_2(\omega_1, (A_1 \times A_2)(\omega_1)) \mu_1(d\omega_1)
\]

\[
= \int_{\Omega_1} \mu_2(\omega_1, A_2)1_{A_1}(\omega_1) \mu_1(d\omega_1) = \int_{A_1} \mu_2(\omega_1, A_2) \mu_1(d\omega_1).
\]
Substituting \( A_1 = \Omega_1, A_2 = \Omega_2 \) in the above formula yields that \( \mu \) is a probability measure.  

\[ \square \]

**Notation:** If \((\Omega, \mathcal{F})\) is a measurable space, we denote by \( m\mathcal{F} \) the space of real valued Borel measurable functions on \((\Omega, \mathcal{F})\). By \( b\mathcal{F} \), we denote the bounded functions in \( m\mathcal{F} \).

**Theorem 0.6.** Assume the hypothesis of Theorem 0.5. Let \((\Omega, \mathcal{F}) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)\) denote the product measurable space. Then

(a) For all \( f \in m\mathcal{F} \) and for all \( \omega_1 \in \Omega_1, \omega_2 \mapsto f(\omega_1, \omega_2) \) belongs to \( m\mathcal{F}_2 \).

(b) For all \( f \in b\mathcal{F} \), the function \( \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \) belongs to \( b\mathcal{F}_1 \).

(c) For all \( f \in b\mathcal{F} \), we have

\[
\int_{\Omega} f(\omega) \, d\mu = \int_{\Omega_1} \left( \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \right) \mu_1(d\omega_1).
\]

**Proof.** (a) Fix \( \omega_1 \in \Omega_1 \) and define \( g(\omega_2) = f(\omega_1, \omega_2) \), where \( f \in m\mathcal{F} \). Let \( B \) be a Borel subset of \( \mathbb{R} \). Note that \( g^{-1}(B) = F(\omega_1) \), where \( F = f^{-1}(B) \in \mathcal{F} \). By (0.4), \( g^{-1}(B) \in \mathcal{F}_2 \) for all Borel subsets \( B \). Therefore \( \omega_2 \mapsto f(\omega_1, \omega_2) \) belongs to \( m\mathcal{F} \) for all \( f \in m\mathcal{F} \) and for all \( \omega_1 \in \Omega_1 \).

(b) By (a), \( \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \) exists and is finite for all \( f \in b\mathcal{F}, \omega_1 \in \Omega_1 \). Since \( \mu_2(\omega_1, \cdot) \) is a probability measure on \((\Omega_2, \mathcal{F}_2)\), we have

\[
\left| \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \right| \leq \int_{\Omega_2} \left( \sup_{\Omega} |f| \right) \mu_2(\omega_1, d\omega_2) = \sup_{\Omega} |f|, \quad \text{for all } \omega_1 \in \Omega_1, f \in b\mathcal{F}.
\]

Therefore \( \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \) is bounded. It remains to show measurability. To this end, we define

\[
\mathcal{H}_1 = \left\{ f \in b\mathcal{F} : \omega_1 \mapsto \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \text{ belongs to } m\mathcal{F}_1 \right\}.
\]

Let \( \mathcal{P} \) denote the \( \pi \)-system of measurable cylinders. Let \( A_1 \times A_2 \in \mathcal{P} \), where \( A_1 \in \mathcal{F}_1, A_2 \in \mathcal{F}_2 \). Then \( \omega_1 \mapsto \int_{\Omega_2} 1_{A_1 \times A_2}(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) = 1_{A_1}(\omega_1) \mu_2(\omega_1, A_2) \) belongs to \( m\mathcal{F}_1 \) as it is a product of measurable functions \( \omega_1 \mapsto 1_{A_1}(\omega_1) \) and \( \omega_1 \mapsto \mu_2(\omega_1, A_2) \). By linearity of integrals \( \mathcal{H}_1 \) is a vector space. Let \( f_n \uparrow f, 0 \leq f_n \in \mathcal{H}_1 \) for all \( n \in \mathbb{N} \) and \( f \in b\mathcal{F} \). By MCT,

\[
\int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) = \lim_{n \to \infty} \int_{\Omega_2} f_n(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2), \quad \text{for all } \omega_1 \in \Omega_1.
\]

Since limit of measurable functions is measurable, \( f \in \mathcal{H}_1 \). By FMCT, \( \mathcal{H}_1 = b\mathcal{F} \).

(c) Define

\[
\mathcal{H}_2 = \left\{ f \in b\mathcal{F} : \int_{\Omega} f \, d\mu = \int_{\Omega_1} \left( \int_{\Omega_2} f(\omega_1, \omega_2) \mu_2(\omega_1, d\omega_2) \right) \mu_1(d\omega_1) \right\}.
\]
By (0.3), $1_F \in \mathcal{H}_2$ for all $F \in \mathcal{F}$. By linearity of integrals $\mathcal{H}_2$ is a vector space. If $f_n \uparrow f$ with $0 \leq f_n \in \mathcal{H}_2$ for all $n \in \mathbb{N}$, and $f \in b\mathcal{F}$. By the equation

$$
\lim_{n \to \infty} \int_{\Omega} f_n \, d\mu = \lim_{n \to \infty} \int_{\Omega_1} \left( \int_{\Omega_2} f_n(\omega_1, \omega_2) \, \mu_2(\omega_1, d\omega_2) \right) \, \mu_1(d\omega_1)
$$

and by using MCT (three times; why?), we obtain $f \in \mathcal{H}_2$. By FMCT, $\mathcal{H}_2 = b\mathcal{F}$.

Theorems 0.5 and 0.6 have the following extension to $n$-fold products.

**Theorem 0.7.** Let $(\Omega_j, \mathcal{F}_j)$ be measurable spaces for $j = 1, \ldots, n$. Let $\mu_1$ be a probability measure on $(\Omega_1, \mathcal{F}_1)$, and for each $(\omega_1, \ldots, \omega_j) \in \Omega_1 \times \cdots \times \Omega_j$, let $A \mapsto \mu_{j+1}(\omega_1, \ldots, \omega_j, A)$ be a probability measure on $(\Omega_j, \mathcal{F}_j)$ (for $j = 1, \ldots, n-1$). Assume that for all $j = 1, \ldots, n-1$, $A \in \mathcal{F}_j$, the function $(\omega_1, \ldots, \omega_j) \mapsto \mu_{j+1}(\omega_1, \ldots, \omega_j, A)$ is Borel measurable on $(\Omega_1 \times \cdots \times \Omega_j, \mathcal{F}_1 \times \cdots \times \mathcal{F}_j)$.

Let $(\Omega, \mathcal{F}) = (\Omega_1 \times \cdots \times \Omega_n, \mathcal{F}_1 \times \cdots \times \mathcal{F}_n)$. Then there exists a unique probability measure $\mu$ on $(\Omega, \mathcal{F})$ such that for each measurable rectangle $A_1 \times \cdots \times A_n \in \mathcal{F}$,

$$
\mu(A_1 \times \cdots \times A_n) =
\int_{A_1} \mu_1(d\omega_1) \int_{A_2} \mu_2(d\omega_1, d\omega_2) \cdots \int_{A_{n-1}} \mu_{n-1}(d\omega_1, \ldots, d\omega_{n-1}) \int_{A_n} \mu_n(d\omega_1, \ldots, d\omega_n).
$$

(0.6)

(In the RHS of the above expression, we evaluate the innermost integral over $A_n$ first and move outwards). Furthermore, for all $f \in b\mathcal{F}$,

$$
\int_{\Omega} f \, d\mu = \int_{\Omega_1} \mu_1(d\omega_1) \int_{\Omega_2} \mu_2(d\omega_1, d\omega_2) \cdots \int_{\Omega_{n-1}} \mu_{n-1}(d\omega_1, \ldots, d\omega_{n-1}) \int_{\Omega_n} f(d\omega_1, \ldots, d\omega_n) \mu_n(d\omega_1, \ldots, d\omega_n).
$$

(0.7)

**Proof.** The proof is by induction on $n$. The case $n = 2$ follows from Theorems 0.5 and 0.6.

For the induction step from the case $n-1$ to the case $n$, we apply the $n = 2$ result to the spaces $(\Omega_1 \times \cdots \times \Omega_{n-1}, \mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1})$ and $(\Omega_n, \mathcal{F}_n)$ and using the fact that $(\mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1}) \times \mathcal{F}_n = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$.

**Exercise:** Show that $(\mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1}) \times \mathcal{F}_n = \mathcal{F}_1 \times \cdots \times \mathcal{F}_n$.

By the induction hypothesis there is a unique probability measure $\lambda_{n-1}$ on $(\Omega_1 \times \cdots \times \Omega_{n-1}, \mathcal{F}_1 \times \cdots \times \mathcal{F}_{n-1})$ such that for all $A_1 \in \mathcal{F}_1, \ldots, A_{n-1} \in \mathcal{F}_{n-1}$

$$
\lambda_{n-1}(A_1 \times \cdots \times A_{n-1}) = \int_{A_1} \mu_1(d\omega_1) \int_{A_2} \mu_2(d\omega_1, d\omega_2) \cdots \int_{A_{n-1}} \mu_{n-1}(d\omega_1, \ldots, d\omega_{n-1})
$$

\[ \lambda_{n-1}(A_1 \times \cdots \times A_{n-1}) = \int_{A_1} \mu_1(d\omega_1) \int_{A_2} \mu_2(d\omega_1, d\omega_2) \cdots \int_{A_{n-1}} \mu_{n-1}(d\omega_1, \ldots, d\omega_{n-1}) \]
By the $n = 2$ case (Theorem 0.5), there exists a measure $\mu$ such that for all $A \in \mathcal{F}_1 \times \ldots \times \mathcal{F}_{n-1}$, $A_n \in \mathcal{F}_n$,

$$\mu(A \times A_n) = \int_A \mu_n(\omega_1, \ldots, \omega_{n-1}, A_n) \, d\lambda_{n-1}(\omega_1, \ldots, \omega_{n-1})$$

$$= \int_{\Omega_1 \times \ldots \times \Omega_{n-1}} 1_A(\omega_1, \ldots, \omega_{n-1}) \mu_n(\omega_1, \ldots, \omega_{n-1}, A_n) \, d\lambda_{n-1}(\omega_1, \ldots, \omega_{n-1}).$$

If $A = A_1 \times \ldots \times A_{n-1}$, then $1_A(\omega_1, \ldots, \omega_{n-1}) = 1_{A_1}(\omega_1) \ldots 1_{A_{n-1}}(\omega_{n-1})$. Along with induction hypothesis on (0.7) for $n - 1$, we obtain (0.6) for $\mu$.

Let $f \in b\mathcal{F}$. The proof of (0.7) for involves a similar application of the $n = 2$ case (using Theorem 0.6)

$$\int f \, d\mu = \int_{\Omega_1 \times \ldots \times \Omega_{n-1}} \int_{\Omega_n} f(\omega_1, \ldots, \omega_n) \, \mu_n(\omega_1, \ldots, \omega_{n-1}, d\omega_n) \, d\lambda_{n-1}(\omega_1, \ldots, \omega_{n-1}).$$

We use induction hypothesis corresponding to (0.7) for the $(n - 1)$-fold product since $(\omega_1, \ldots, \omega_{n-1}) \mapsto \int_{\Omega_n} f(\omega_1, \ldots, \omega_n) \, \mu_n(\omega_1, \ldots, \omega_{n-1}, d\omega_n)$ is $\mathcal{F}_1 \times \ldots \times \mathcal{F}_{n-1}$-measurable. This concludes the proof of (0.7).

\textit{Definition 0.8.} For each $j = 1, 2, \ldots$, let $(\Omega_j, \mathcal{F}_j)$ be a measurable space. Let $(\Omega, \mathcal{F}) = \left(\prod_{j=0}^{\infty} \Omega_j, \prod_{j=0}^{\infty} \mathcal{F}_j\right)$ denote the product measurable space. An element $\omega \in \Omega$ is identified with a sequence $(\omega_1, \omega_2, \ldots)$, where $\omega_j \in \Omega_j$ for all $j = 1, 2, \ldots$. A set $C \subset \Omega$ is said to be a \textbf{measurable cylinder} if there exists $n \in \mathbb{N}$ and a set $B \subset \Omega_1 \times \ldots \times \Omega_n$ where $B$ belongs to the product $\sigma$-field $\mathcal{F}_1 \times \ldots \mathcal{F}_n$, such that

$$C = \{\omega \in \Omega : (\omega_1, \ldots, \omega_n) \in B\}.$$

The cylinder $C$ above is said to have \textbf{base} $B$, and $n$ is said to be the \textbf{dimension of the base}. The same cylinder can be regarded to have a higher dimensional base. For example, if $B \in \mathcal{F}_1 \times \mathcal{F}_2$, then

$$\{\omega \in \Omega : (\omega_1, \omega_2) \in B\} = \{\omega \in \Omega : (\omega_1, \omega_2, \omega_3) \in B \times \Omega_3\}$$

A set $C \subset \Omega$ is said to be a \textbf{measurable rectangle} if there exists $n \in \mathbb{N}$ and sets $A_j \in \mathcal{F}_j$, $j = 1, \ldots, n$ such that

$$C = \left\{\omega \in \Omega : (\omega_1, \ldots, \omega_n) \in \prod_{j=1}^{n} A_j\right\}.$$

Note that every measurable rectangle is a measurable cylinder. We denote the set of measurable cylinders and measurable rectangles by $\mathcal{C}$ and $\mathcal{C}_\pi$ respectively (\textbf{Exercise}: Show that $\mathcal{C}$ is an algebra on $\Omega$, $\mathcal{C}_\pi$ is a $\pi$-system on $\Omega$ and $\prod_{j=1}^{\infty} \mathcal{F}_j = \sigma(\mathcal{C}) = \sigma(\mathcal{C}_\pi)$).

\textbf{Lemma 0.9.} \textit{Let $\mathcal{A}$ be an algebra on a set $\Omega$ and let $\mu$ be a finitely additive set function on $\mathcal{A}$. If $\mu$ is continuous from above at the empty set (that is; $C_n \downarrow \emptyset, C_n \in \mathcal{A}$ for all $n \in \mathbb{N}$ implies $\mu(C_n) \downarrow 0$), then $\mu$ is countably additive on $\mathcal{A}$.}
Proof. Let $A_n, n \in \mathbb{N}$ be pairwise disjoint sets such that $A := \bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$. Let $B_n = \bigcup_{i=1}^{n} A_n$ for all $n$. Then $B_n, A \setminus B_n \in \mathcal{A}$ and $\mu(A) = \mu(B_n) + \mu(A \setminus B_n)$. But $A \setminus B_n \downarrow \emptyset$. Therefore, by hypothesis $\mu(A \setminus B_n) \downarrow 0$, and hence $\mu(B_n) \uparrow \mu(A)$. Therefore

$$\mu(A) = \lim_{n \to \infty} \mu(B_n) = \lim_{n \to \infty} \sum_{i=1}^{n} \mu(A_i) = \sum_{i=1}^{\infty} \mu(A_i),$$

thus proving countable additivity of $\mu$ on $\mathcal{A}$. \hfill \Box

We now prove an infinite dimensional version of Theorem 0.7.

Theorem 0.10. Let $(\Omega_j, \mathcal{F}_j)$ be measurable spaces for $j \in \mathbb{N}$ and let $(\Omega, \mathcal{F}) = (\prod_{j=1}^{\infty} \Omega_j, \prod_{j=1}^{\infty} \mathcal{F}_j)$. Let $\mu_1$ be a probability measure on $(\Omega_1, \mathcal{F}_1)$, and for each $(\omega_1, \ldots, \omega_j) \in \Omega_1 \times \ldots \times \Omega_j$, let $A \mapsto \mu_{j+1}(\omega_1, \ldots, \omega_j, A)$ be a probability measure on $(\Omega_j, \mathcal{F}_j)$ (for $j \in \mathbb{N}$). Assume that for all $j = 1, \ldots, n-1, A \in \mathcal{F}_{j+1}$, the function $(\omega_1, \ldots, \omega_j) \mapsto \mu_{j+1}(\omega_1, \ldots, \omega_j, A)$ is Borel measurable on $(\Omega_1 \times \ldots \times \Omega_j, \mathcal{F}_1 \times \ldots \times \mathcal{F}_j)$.

For $n \in \mathbb{N}$, let $P_n$ be the unique probability measure on $(\prod_{j=1}^{n} \Omega_j, \prod_{j=1}^{n} \mathcal{F}_j)$ such that for all $A \in \prod_{j=1}^{n} \mathcal{F}_j$,

$$P_n(B) = \int_{\Omega_1} \mu_1(d\omega_1) \int_{\Omega_2} \mu_2(\omega_1, d\omega_2) \cdots \int_{\Omega_{n-1}} \mu_{n-1}(\omega_1, \ldots, \omega_{n-2}, d\omega_{n-1}) \int_{\Omega_n} 1_B(\omega_1, \ldots, \omega_n) \mu_n(\omega_1, \ldots, \omega_{n-1}, d\omega_n). \tag{0.8}$$

Such a measure exists by Theorem 0.7.

Then there exists an unique probability measure $P$ on $(\Omega, \mathcal{F})$ such that for all $n \in \mathbb{N}$, $P$ agrees with $P_n$ on $n$-dimensional cylinders; that is

$$P \left( \{ \omega \in \Omega : (\omega_1, \ldots, \omega_n) \in B \} \right) = P_n(B), \quad \text{for all } n \in \mathbb{N}, B \in \prod_{i=1}^{n} \mathcal{F}_i. \tag{0.9}$$

Furthermore, if $f \in b(\mathcal{F}_1 \times \ldots \times \mathcal{F}_n)$, then

$$\int_{\Omega} f(\omega_1, \ldots, \omega_n) \, P(d\omega) = \int_{\Omega_1} \mu_1(d\omega_1) \int_{\Omega_2} \mu_2(\omega_1, d\omega_2) \cdots \int_{\Omega_{n-1}} \mu_{n-1}(\omega_1, \ldots, \omega_{n-2}, d\omega_{n-1}) \int_{\Omega_n} f(\omega_1, \ldots, \omega_n) \mu_n(\omega_1, \ldots, \omega_{n-1}, d\omega_n). \tag{0.10}$$

Proof. Let $\mathcal{C}$ denote the algebra of measurable cylinders. Since the same measurable cylinder can have bases of different dimensions, we need to check that $P$ defined by (0.9) is well defined on measurable cylinders. If a cylinder $C$ has two bases $B^n \in \prod_{j=1}^{n} \mathcal{F}_j, B^m \in \prod_{j=1}^{m} \mathcal{F}_j$ for $m < n$, then $\{ \omega \in \Omega : (\omega_1, \ldots, \omega_n) \in B^n \} = \{ \omega \in \Omega : (\omega_1, \ldots, \omega_m) \in B^m \}$. Therefore $B^n = B^m \times \Omega_{m+1} \times \ldots \times \Omega_n$ and $1_{B^n}(\omega_1, \ldots, \omega_m) = 1_{B^m}(\omega_1, \ldots, \omega_m)$ since $A \mapsto \mu_j(\omega_1, \ldots, \omega_{j-1}, A)$ is a probability measure on $(\Omega_j, \mathcal{F}_j)$, it follows from the definition of $P_n$ that $P_n(B^n) = P_m(B^m)$. 

7
Since $P_n$ is a probability measure on $(\prod_{j=1}^n \Omega_j, \prod_{j=1}^n \mathcal{F}_j)$ for all $n \in \mathbb{N}$, $P$ is finitely additive on $\mathcal{C}$ (since, any finite collection of cylinders in $\mathcal{C}$ can be viewed to have bases of the same dimension).

By Carathéodory’s extension theorem (Theorem 0.2), it suffices to show that $P$ is countably additive on $P$. By Lemma 0.9, it suffices to show that for a sequence $B_{n_k} : k \in \mathbb{N}$ of measurable cylinders with dimensions $n_k$ such that $B_n \downarrow \emptyset$, we have $P(B_{n_k}) \downarrow 0$. Without loss of generality, we may assume that $n_k = k$ for all $k \in \mathbb{N}$ (by repeating some cylinders and increasing the dimension by one each time.).

Assume to the contrary that $\lim_{n \to \infty} P(B_n) > 0$. Let $B^n$ denote the basis of $B_n$ for all $n$. Then by (0.8), for all $n > 1$,

$$P(B_n) = \int_{\Omega_n} g_n^{(1)}(\omega_1) \mu_1(d\omega_1),$$

where

$$g_n^{(1)}(\omega_1) = \int_{\Omega_2} \mu_2(\omega_1, d\omega_2) \cdots \int_{\Omega_{n-1}} \mu_{n-1}(\omega_1, \ldots, \omega_{n-2}, d\omega_{n-1}) \int_{\Omega_n} 1_{B^n}(\omega_1, \ldots, \omega_n) \mu_n(\omega_1, \ldots, \omega_{n-1}, d\omega_n).$$

Since $1_{B^{n+1}}(\omega_1, \ldots, \omega_{n+1}) \leq 1_{B^n}(\omega_1, \ldots, \omega_n)$, $g_n^{(1)}(\omega_1)$ decreases as $n$ increases for all $\omega_1 \in \Omega_1$. Therefore $g_n^{(1)} \to h_1$ for some $h_1 \in b\mathcal{F}_1$. Since $g_n^{(1)} \leq 1$ for all $n$, by DCT

$$0 < \lim_{n \to \infty} P(B_n) = \int_{\Omega_1} h_1(\omega_1) \mu_1(d\omega_1).$$

Therefore, there exists $\omega'_1 \in \Omega_1$ such that $\omega'_1 \in B^1$. Otherwise, $1_{B^n}(\omega'_1, \omega_2, \ldots, \omega_n) = 0$ for all $n > 1$, $\omega_j \in \Omega_j$, $j = 2, \ldots, n$ (since $B_n$ is a decreasing sequence of sets).

Similarly,

$$g_n^{(1)}(\omega'_1) = \int_{\Omega_2} g_n^{(2)}(\omega_2) \mu_2(\omega'_1, d\omega_2),$$

where

$$g_n^{(2)}(\omega_2) = \int_{\Omega_2} \mu_3(\omega'_1, \omega_2, d\omega_3) \cdots \int_{\Omega_n} 1_{B^n}(\omega'_1, \ldots, \omega_n) \mu_n(\omega'_1, \ldots, \omega_{n-1}, d\omega_n).$$

As above $g_n^{(2)}(\omega_2) \downarrow h_2(\omega_2)$ for all $\omega_2 \in \Omega_2$; hence by DCT

$$0 < h_1(\omega'_1) = \int_{\Omega_2} h_2(\omega_2) \mu_2(\omega'_1, d\omega_2).$$

Therefore, there exists $\omega'_2 \in \Omega_2$ such that $h_2(\omega'_2) > 0$. Arguing as above, $(\omega'_1, \omega'_2) \in B^2$.

Repeating the process inductively, we obtain a sequence of points $\omega'_1, \omega'_2, \ldots$ such that for each $n$, $(\omega'_1, \omega'_2, \ldots, \omega'_n) \in B^n$. Therefore $(\omega'_1, \omega'_2, \ldots) \in \cap_{n=1}^\infty B_n = \emptyset$, a contradiction. This proves that $P(B_n) \downarrow 0$, and hence $P$ is countably additive on $\mathcal{C}$. The existence and uniqueness of $P$ follows from Theorem 0.2.
Now we turn to the proof of (0.10). To this end, define
\[
\mathcal{H} = \{ f \in b(\mathcal{F}_1 \times \ldots \times \mathcal{F}_n) : f \text{ satisfies } (0.10) \}.
\]
Let \( B^n \in \mathcal{F}_1 \times \ldots \times \mathcal{F}_n \) and let \( f = 1_{B^n}(\omega_1, \ldots, \omega_n) \). Then \( f = 1_B(\omega) \), where \( B \) is the measurable cylinder in \( \Omega \) with base \( B^n \). In this case (0.10) follows from (0.9) and (0.8).

By linearity of integrals \( \mathcal{H} \) is a vector space. By using MCT, we verify that \( \mathcal{H} \) is closed under increasing bounded limits. Therefore by FMCT, \( \mathcal{H} = b(\mathcal{F}_1 \times \ldots \times \mathcal{F}_n) \). \[\Box\]

As a special case of the above theorem, we obtain a construction of Markov chain, given a transition probability \( p \) and an initial distribution \( \mu \).

**Theorem 0.11** (Existence/Construction of Markov chain). Let \((S, \mathcal{S})\) be a measurable space and let \( p \) be a transition probability on \((S, \mathcal{S})\). Let \( \mu_0 \) be a probability measure on \((S, \mathcal{S})\). Let \((\Omega, \mathcal{F}) = (\prod_{j=0}^{\infty} S, \prod_{j=0}^{\infty} \mathcal{S})\) denote the product measurable space. Let \( X_n : (\Omega, \mathcal{F}) \to (S, \mathcal{S}) \) denote the projection on to the \( n \)-th component, \( n \in \mathbb{Z}_+ \) (that is, \( X_n(\omega_0, \omega_1, \ldots) = \omega_n \)). Let \( \mathcal{F}_n = \sigma(X_0, X_1, \ldots, X_n) \) be the natural filtration associated to the process \( X = \{X_n\}_{n \in \mathbb{Z}_+} \). Then there exists a probability measure \( P \) on \((\Omega, \mathcal{F})\) such that \( X \) is \((\mathcal{F}_n)\)-Markov chain with transition probability \( p \) and initial distribution \( \mu \).

**Proof.** We will use Theorem 0.10 with indices \( j \in \mathbb{Z}_+ \) (instead of \( j \in \mathbb{N} \)). We let \((\Omega_j, \mu_j) = (S, \mathcal{S})\) for all \( j \in \mathbb{Z}_+ \), \( \mu_0 = \mu \), \( \mu_n(\omega_0, \ldots, \omega_{n-1}, A) = p(\omega_{n-1}, A) \) for all \( A \in \mathcal{S} \) and \( \omega_0, \ldots, \omega_{n-1} \in \mathcal{S} \) for all \( n \in \mathbb{N} \) in Theorem 0.10. Let \( P \) be the probability measure on \((\Omega, \mathcal{F})\) in Theorem 0.10.

Clearly, \( X_n \) is \( \mathcal{F}_n \)-adapted and \((S, \mathcal{S})\)-valued process. It remains to verify the Markov property (0.1). Let \( A \in \mathcal{S} \). Note that \( \omega \mapsto p(X_n(\omega), dy) \) belongs to \( m\mathcal{F}_n \). Therefore, to verify (0.2), it suffices to check
\[
\int_{\Omega} 1_A(X_{n+1}(\omega))1_B(\omega) P(d\omega) = \int p(X_n(\omega), A)1_B(\omega) P(d\omega) \quad \text{for all } B \in \mathcal{F}_n. \quad (0.11)
\]
Let \( \mathcal{L} = \{ B \in \mathcal{F}_n : B \text{ satisfies } (0.11) \} \). Then, \( \mathcal{L} \) is a \( \lambda \)-system (check!). Since \( \mathcal{F}_n \) is generated by the \( \pi \)-system of measurable cylinders in \( \mathcal{F}_n \), by \( \pi \)-\( \lambda \) theorem, it suffices to verify (0.11) for measurable cylinders \( B \) in \( \mathcal{F}_n \).

To this end, let \( B \) denote a measurable cylinder in \( \mathcal{F}_n \); that is \( B = A_0 \times \ldots \times A_n \times S \times S \times \ldots \), where \( A_0, \ldots, A_n \in \mathcal{S} \). Set \( A_{n+1} = A \). Then (0.11), reduces to the claim
\[
P(A_0 \times \ldots \times A_n \times A_{n+1} \times S \times S \times \ldots) = \int_{\Omega} p(X_n(\omega), A_{n+1}) \prod_{j=0}^{n} 1_{A_j}(X_j(\omega)) dP(\omega). \quad (0.12)
\]
By (0.8) and (0.9), we obtain

\[
P(A_0 \times \ldots \times A_n \times A_{n+1} \times S \times S \times \ldots) = \int_S \mu(d\omega_0) \int_S p(\omega_0, d\omega_1) \cdots \int_S p(\omega_{n-1}, d\omega_n) \int_{\prod_{j=0}^{n+1} 1_{A_j}(\omega_j)p(\omega_n, d\omega_{n+1})}.
\]

\[
= \int_S \mu(d\omega_1) \int_S p(\omega_1, d\omega_2) \cdots \int_S p(\omega_n, A_{n+1}) \prod_{j=0}^n 1_{A_j}(\omega_j) p(\omega_{n-1}, d\omega_n) = \int_{\Omega} p(\omega_n, A_{n+1}) \prod_{j=0}^n 1_{A_j}(\omega_j) dP(\omega) \quad \text{(by (0.10))}.
\]

This proves (0.12) and hence (0.11).

Hence the process \((X_n)_{n \in \mathbb{Z}_+}\) satisfies the Markov property with respect to the filtration \((\mathcal{F}_n)_{n \in \mathbb{Z}_+}\) on the probability space \((\Omega, \mathcal{F}, P)\). \qed

References