Caratheodory’s extension theorem

DBW

February 7, 2011

These notes are meant as introductory notes on Caratheodory’s extension theorem. The presentation is not completely my own work; the presentation heavily relies on the presentation of Noel Vaillant on http://www.probability.net/WEBcaratheodory.pdf. To make the line of arguments as clear as possible, the starting point is the notion of a ring on a topological space and not the notion of a semi-ring.

1 Elementary definitions and properties

We fix a topological space Ω. The power set of Ω is denoted \( \mathcal{P}(Ω) \) and consists of all subsets of Ω.

**Definition 1.** A ring on \( Ω \) is a subset \( \mathcal{R} \) of \( \mathcal{P}(Ω) \), such that

(i) \( \emptyset \in \mathcal{R} \)

(ii) \( A, B \in \mathcal{R} \Rightarrow A \cup B \in \mathcal{R} \)

(iii) \( A, B \in \mathcal{R} \Rightarrow A \setminus B \in \mathcal{R} \)

**Definition 2.** A \( σ \)-algebra on \( Ω \) is a subset \( \Sigma \) of \( \mathcal{P}(Ω) \) such that

(i) \( \emptyset \in \Sigma \)

(ii) \( (A_n)_{n∈\mathbb{N}} \in \mathcal{P}(Ω) \Rightarrow \bigcup_n A_n \in \Sigma \)

(iii) \( A \in \Sigma \Rightarrow A^c \in \Sigma \)

Since \( A \cap B = A \setminus (A \setminus B) \) it follows that any ring on \( Ω \) is closed under finite intersections; hence any ring is also a semi-ring. Since \( \bigcap_n A_n = (\bigcup_n A_n)^c \) it follows that any \( σ \)-algebra is closed under arbitrary intersections. And from \( A \setminus B = A \cap B^c \) we deduce that any \( σ \)-algebra is also a ring.

If \( (\mathcal{R}_i)_{i∈I} \) is a set of rings on \( Ω \) then it is clear that \( \bigcap_i \mathcal{R}_i \) is also a ring on \( Ω \). Let \( S \) be any subset of \( \mathcal{P}(Ω) \), then we call the intersection of all rings on \( Ω \) containing \( S \) the ring generated by \( S \).

**Definition 3.** Let \( \mathcal{A} \) be a subset of \( \mathcal{P}(Ω) \). A measure on \( \mathcal{A} \) is a map \( μ : \mathcal{A} → [0, +\infty] \) such that

(i) \( μ(\emptyset) = 0 \)

(ii) If \( A_n ∈ \mathcal{A} \) are disjoint and \( A = \bigcup_n A_n ∈ \mathcal{A} \Rightarrow μ(A) = \sum_n μ(A_n) \).

If \( \mathcal{A} \) is a \( σ \)-algebra, we don’t need to assume that in addition \( \bigcup_n A_n ∈ \mathcal{A} \). By taking all but finitely many \( A_n \) to be the empty set one sees that \( μ(\bigcup_n A_n) = μ(A_1) + \ldots + μ(A_n) \). If \( A ⊂ B \) then \( A \cup (B \setminus A) = B \) and hence \( μ(B) = μ(A) + μ(B \setminus A) \geq μ(A) \).
Definition 4. We call an outer measure on $\Omega$ a map $\lambda : \mathcal{P}(\Omega) \to [0, +\infty]$ with

(i) $\lambda(\emptyset) = 0$
(ii) $A \subset B \implies \lambda(A) \leq \lambda(B)$
(iii) $(A_n)_{n \in \mathbb{N}} \in \mathcal{P}(\Omega)$, $\lambda(\cup_n A_n) \leq \sum_n \lambda(A_n)$

By taking all but finitely many $A_n$ to be the empty set one sees that an outer measure is subadditive: $\lambda(A \cup B) \leq \lambda(A) + \lambda(B)$.

2 The interplay between $\sigma$-algebras and (outer) measures

Let $\lambda$ be an outer measure on $\Omega$. We define $\Sigma_{\lambda}$ to be the set of all subsets $A \subset \Omega$ such that for any $X \subset \Omega$ we have

$$\lambda(X) = \lambda(X \cap A) + \lambda(X \cap A^c).$$

In other words, $\Sigma_{\lambda}$ consists of all subsets $A \subset \Omega$ that cut $\Omega$ in two in a good way. Clearly, $\Omega \in \Sigma_{\lambda}$ and by the very form of the definition of $\Sigma_{\lambda}$, we have $A \in \Sigma_{\lambda} \iff A^c \in \Sigma_{\lambda}$. We can now present the following proposition, whose proof is a bit tedious, but which contains loads of information on the exact interplay.

Proposition 5. Let $\lambda$ be an outer measure on $\Omega$ and let $\Sigma_{\lambda}$ be as defined above. Then $\Sigma_{\lambda}$ is a $\sigma$-algebra on $\Omega$.

Proof. After the preliminary remarks preceding the proposition, it only remains to show that $\Sigma_{\lambda}$ is closed under finite intersections and unions.

Let $A, B \in \Sigma_{\lambda}$ and let $X$ be any subset of $\Omega$. We have $X \cap A^c = X \cap (A \cap B)^c \cap A^c$ since $(A \cap B)^c \supset A^c$. On the other hand we have $(A \cap B)^c = A^c \cup B^c$ and hence $X \cap (A \cap B)^c \cap A = (X \cap A \cap B^c) \cup (X \cap A \cap A^c) = X \cap A \cap B^c$. Therefore we have $\lambda(X \cap (A \cap B)^c) = \lambda(X \cap (A \cap B)^c \cap A) + \lambda(X \cap (A \cap B)^c \cap A^c) = \lambda(X \cap A^c) + \lambda(X \cap A \cap B^c)$. Now adding $\lambda(X \cap A \cap B)$ and using that $\lambda(X \cap A) = \lambda(X \cap A \cap B) + \lambda(X \cap A \cap B^c)$ one obtains $\lambda(X \cap A \cap B) + \lambda(X \cap (A \cap B)^c) = \lambda(X)$. Hence $A \cap B \in \Sigma_{\lambda}$.

Since $A \cup B = (A^c \cap B^c)^c$ and $A \setminus B = A \cap B^c$ we see that $\Sigma_{\lambda}$ is closed under finite unions and the set-theoretic difference. Thus $\Sigma_{\lambda}$ is a ring on $\Omega$.

If $A, B \in \Sigma_{\lambda}$ are disjoint and $X \subset \Omega$ then $\lambda(X \cap (A \cup B)) = \lambda(X) - \lambda(X \cap A^c \cap B^c) = \lambda(X) - \lambda(X \cap A^c) - \lambda(X \cap A^c \cap B) = \lambda(X \cap A) + \lambda(X \cap B)$ as $A^c \cap B = B$. Using induction we obtain $\lambda(X \cap \bigcup_{n=1}^N A_n) = \sum_{n=1}^N \lambda(X \cap A_n)$ whenever $A_n$ are in $\Sigma_{\lambda}$ and pairwise disjoint.

Now we fix a sequence $A_n$ in $\Sigma_{\lambda}$ which are pairwise disjoint and we denote the union $\cup_n A_n$ by $A$. Furthermore, we fix an arbitrary $X \in \Omega$ and an arbitrary large integer $N$.

Since $X \cap A^c \subset X \cap (\bigcup_{n=1}^N A_n)^c$ and $\Sigma_{\lambda}$ is closed under finite unions we have $\lambda(X \cap A^c) + \lambda(X \cap (\bigcup_{n=1}^N A_n)^c) \leq \lambda(X \cap (\bigcup_{n=1}^N A_n)^c) + \sum_n \lambda(X \cap A_n) = \lambda(X)$. But $N$ is arbitrary in this equation and we can safely let it go to $\infty$ and obtain

$$\lambda(X \cap A^c) + \sum_n \lambda(X \cap A_n) \leq \lambda(X). \quad (1)$$
On the other hand we have \( \lambda(X) \leq \lambda(X \cap A^c) + \lambda(X \cap A) \), which again by the definition of an outer measure is less or equal \( \lambda(X \cap A^c) + \sum_n \lambda(X \cap A_n) \). Hence using (1) we obtain
\[
\lambda(X) \leq \lambda(X \cap A^c) + \lambda(X \cap A) \leq \lambda(X \cap A^c) + \sum_n \lambda(X \cap A_n) \leq \lambda(X).
\]

It follows that we must equality. From this we conclude that \( \Sigma_\lambda \) is indeed closed under countable unions and, by taking \( X = A \), that \( \lambda(A) = \sum_n \lambda(A_n) \). Therefore the restriction of \( \lambda \) to \( \Sigma_\lambda \) is a measure on \( \Sigma_\lambda \).

We will call \( \Sigma_\lambda \) the \( \sigma \)-algebra related to \( \lambda \).

Now we come to a critical step; we want to associate an outer measure \( \lambda_\mu \) to a given measure \( \lambda \) on some ring \( R \). Of course, we want the restriction of the outer measure \( \lambda_\mu \) to the ring to coincide with the measure \( \mu \).

Let \( R \) be a ring on \( \Omega \) and let \( \mu \) be a measure on \( R \). If \( X \subset \Omega \) is any subset we can cover \( X \) with sets from \( R \) to approximate \( X \) inside \( R \) - we call an \( R \)-cover of \( X \) a countable subset \( (A_n) \) of \( R \) with \( X \subset \cup_n A_n \). This leads to the following definition; for any \( X \subset \Omega \) we define \( \lambda_\mu(X) \) to be the infimum of all \( \sum_n \mu(A_n) \) where \( (A_n) \) is any countable cover of \( X \) with \( A_n \) in \( R \). We need to check that this is an outer measure.

**Proposition 6.** The map \( \lambda_\mu : \mathcal{P}(\Omega) \rightarrow [0, +\infty] \) defined in the above paragraph defines an outer measure on \( \Omega \).

**Proof.** Since \( \emptyset \in R \) we have \( \lambda_\mu(\emptyset) = 0 \). If \( X \subset Y \) are two subsets of \( \Omega \), then any cover of \( Y \) with sets from \( R \) also covers \( X \) and hence \( \lambda_\mu(X) \leq \lambda_\mu(Y) \).

Now let \( X_n \) be any sequence of subsets of \( \Omega \). By the definition of the infimum we can find for each \( \epsilon > 0 \) and for each \( n \) an \( R \)-cover \( (A_{n,m}) \) of \( X_n \) such that \( \sum_{m} \mu(A_{n,m}) < \lambda_\mu(X_n) + \frac{\epsilon}{2^n} \). The sets \( A_{n,m} \) form a countable cover of \( X = \cup_n X_n \); we can for example set \( B_1 = A_{1,1} \), \( B_2 = A_{2,1} \), \( B_3 = A_{1,2} \), \( B_4 = A_{2,1} \) and so on, similar to Cantor’s proof of the countability of \( \mathbb{Q} \). But then \( \lambda_\mu(X_n) = \lambda_\mu(\cup_n X_n) \leq \sum_{m} \mu(A_{n,m}) < \sum_n (\lambda_\mu(X_n) + \frac{\epsilon}{2^n}) = \sum_n \lambda_\mu(X_n) + \epsilon \). But \( \epsilon \) was arbitrary and hence \( \lambda_\mu(\cup_n X_n) \leq \sum_n \lambda_\mu(X_n) \).

**Proposition 7.** The restriction of \( \lambda_\mu \) to \( R \) is \( \mu \).

**Proof.** For any \( A \in R \) the set \( A \) itself forms a cover and hence \( \lambda_\mu(A) \leq \mu(A) \).

On the other hand, let \( (A_n) \) be an \( R \)-cover of \( A \). We define \( B_1 = A_1 \cap A \) and \( B_{n+1} = (A_n \cap A) \cup \cup_{k>n} (A_k \cap A) \) for \( n \geq 1 \). Then clearly \( B_n \in R \), the \( B_n \) are disjoint, \( \cup_n B_n = A \) and \( \mu(B_n) \leq \mu(A_n) \). Since \( \mu \) is a measure on \( R \) we have \( \mu(A) = \sum_n \mu(B_n) \) which is less than or equal to \( \sum_n \mu(A_n) \). Since this holds for any \( R \)-cover of \( A \) we have \( \mu(A) \leq \lambda_\mu(A) \). Therefore equality holds and the proposition is proved.

We will call \( \lambda_\mu \) the outer measure associated to \( \mu \).

So we now have two constructions; given a ring and a measure on it we can construct an outer measure. Given an outer measure we can construct a \( \sigma \)-algebra such that the
restriction of the outer measure to the \( \sigma \)-algebra is a measure on the \( \sigma \)-algebra. So it seems feasible that we can construct a measure on a \( \sigma \)-algebra starting from a ring with a measure on it. That this really works and that all things work out nicely is the content of Caratheodory’s theorem.

3 Caratheodory’s theorem: Statement and Proof

**Lemma 8.** Let \( R \) be a ring on \( \Omega \) and let \( \mu \) be a measure on \( R \). Let \( \lambda \) be the outer measure associated to \( \mu \). Let \( \Sigma \) be the \( \sigma \)-algebra related to \( \lambda \). Then \( R \in \Sigma \).

**Proof.** Let \( A \) be an element of \( R \) and let \( X \) be any subset of \( \Omega \). Since \( \lambda \) is an outer measure on \( \Omega \) we have \( \lambda(X) = \lambda((X \cap A) \cup (X \cap A^c)) \leq \lambda(X \cap A) + \lambda(X \cap A^c) \).

Now let \( (A_n)_{n \in \mathbb{N}} \) be any \( R \)-cover of \( X \). Then the \( A_n \cap A \) form an \( R \)-cover of \( X \cap A \) and the \( A_n \cap A^c \) form an \( R \)-cover of \( X \cap A^c \). Hence we have that \( \lambda(X \cap A) + \lambda(X \cap A^c) \leq \sum_n \mu(A_n \cap A) + \sum_n (A_n \cap A^c) = \sum_n \mu(A_n) \), where the last step follows from the fact that \( \mu \) is a measure and hence \( \mu(C \cup D) = \mu(C) + \mu(D) \). Since the inequality holds for any \( R \)-cover of \( X \) we need \( \lambda(X \cap A) + \lambda(X \cap A^c) \leq \lambda(X) \). We thus need equality; for any \( X \subset \Omega \) we have \( \lambda(X) = \lambda(X \cap A) + \lambda(X \cap A^c) \), or in other words \( A \in \Sigma \) and since \( A \) was an arbitrary element of \( R \) the lemma is proved.

**Remark 9.** Any ring generates a \( \sigma \)-algebra; one simply enlarges the ring with countable unions. Or, the \( \sigma \)-algebra generated by the ring \( R \) is the intersection of all \( \sigma \)-algebras that contain \( R \). Therefore the above lemma shows that the \( \sigma \)-algebra generated by \( R \) is contained in \( \Sigma \).

Now we come to Caratheodory’s theorem:

**Theorem 10.** Let \( R \) be a ring on \( \Omega \) and let \( \mu \) be a measure on \( R \). Then there exists a measure \( \mu' \) on the \( \sigma \)-algebra generated by \( R \) such that the restriction of \( \mu' \) to \( R \) coincides with \( \mu \).

**Proof.** Let \( \lambda \) be the outer measure on \( \Omega \) associated to \( \mu \). Let \( \Sigma \) be the \( \sigma \)-algebra associated to \( \lambda \). Then by lemma 8 the \( \sigma \)-algebra generated by \( R \) is contained in \( \Sigma \). Hence \( \lambda \) restricts to a measure on the \( \sigma \)-algebra generated by \( R \). By proposition 7 this restriction of \( \lambda \) to \( R \) coincides with \( \mu \).

**Example 11.** Let \( \Omega \) be the real line. Then the open intervals generate a \( \sigma \)-algebra \( \Sigma \). For any open interval \((a, b)\) with \( a < b \) we can put \( \mu((a, b)) = b - a \). Then there exists a measure \( \mu' \) on \( \Sigma \) such that \( \mu'((a, b)) = b - a \). Indeed, for countable unions of disjoint intervals we can define \( \mu(\cup_{n=1}^{\infty} (a_n, b_n)) = \sum_{n=1}^{\infty} (b_n - a_n) \). Hence \( \mu \) does give rise to a measure on the ring generated by all intervals.