张志华-统计机器学习笔记
2.概率基础

By Hao ZHAN
Edit 2019.7.28

导言

第二次课主要回顾了概率论的基本知识：
- 样本空间与事件
- σ 域与测度
- 独立事件

Part 01 样本空间与事件 Sample Space and Events

1.样本空间和事件

(1) Sample Space Ω is the set of possible outcome of an experiment.
(2) $\omega \in \Omega$ is called sample outcome or elements.
(3) Subsets of Ω are called Event.

2.补集

给定事件A，定义A的补集:

$$A^c = \{\omega \in \Omega, \omega \notin A\}$$

3.单增与单减

如果一个Sequence of set A1，A2… 是单增 (monotone increasing) 的话，即

$$A_1 \subset A_2 \subset A_3 \dots$$

then, we define $\lim_{n\to\infty} A_n = \bigcup_{i=1}^{\infty} A_i$

如如果是单减(monotone decreasing), 即

$$A_1 \supset A_2 \supset A_3 \dots$$

then, we define $\lim_{n\to\infty} A_n = \bigcap_{i=1}^{\infty} A_i$

Example.01

Let $\Omega = R$, and let $A_i = \left[0, \frac{1}{i}\right)$ for $i = 1,2,3...$

So, $\bigcup_{i=1}^{\infty} A_i = [0,1)$, $\bigcap_{i=1}^{\infty} A_i = \{0\}$

Example. 02

Let $\Omega = R$, and let $A_i = \left(0, \frac{1}{i}\right)$ for $i = 1,2,3...$

So, $\bigcup_{i=1}^{\infty} A_i = (0,1)$, $\bigcap_{i=1}^{\infty} A_i = \emptyset$

Part 02 $\sigma - field$ and Measures

直观解释：测度就是一个测量的衡量标准，是长度、体积的一个推广。

1. σ 域（σ 代数）

Let $\mathscr{A}$ be a collection of subsets of a staple space Ω. $\mathscr{A}$ is called $\sigma - field$ (or $\sigma - algebra$) iff it has the following properties:

(1) The empty set $\emptyset \in \mathscr{A}$.

(2) If $A \in \mathscr{A}$, then $A^C \in \mathscr{A}$.

(3) If $A_i \in \mathscr{A}$, the $\bigcup_{i=1}^{\infty} A_i = \mathscr{A}$

称 $\mathscr{A}$ 为 σ 域。（可以看作是事件）

A pair $(\Omega, \mathscr{A})$ be called a measurable space.
The dements of $\mathscr{A}$ are called measurable set.

Ω 最小的 $\sigma - field$ 是 $\{\emptyset, \Omega\}$
Ω 最大的 $\sigma - field$ 是幂集。

Example. Borel $\sigma - field$
Let $\Omega \in R$, $\mathscr{A}$ is the smallest $\sigma - field$ that contains all the finite open subsets of R is called Borel $\sigma - field$（$B(R)$）.

2. 测度

Let $\{\Omega, \mathscr{A}\}$ be a measurable space a set function v defined in $\mathscr{A}$ is called a measure（测度）iff,

(1) $0 \leq v \leq \infty$ for $A \in \mathscr{A}$.

(2) $v(\emptyset) = 0$

(3) If $A_i \in \mathscr{A}_i$ and A_i are disjoint (i.e. $A_i \bigcap A_j = \emptyset$, for $i \neq j$), the $v\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} v(A_i)$ （可加性）

Triple $(\Omega, \mathscr{A}, v)$ is called a measure space （测度空间）.
当$v(\Omega) = 1$时，称其为概率测度，$(\Omega, \mathscr{A}, P)$ 则是概率测度空间。

若条件（1）更改为$0 \le \upsilon < \infty$，称 υ 为信念（belife）测度。（并且总能够除以 Ω 来将其归一化为概率测度）。

Example.01 计算测度

Let Ω be a sample space, A the collection of all substance and $\upsilon(A)$ the number of elements in A.

Example.02 Lebesgue measure

3.概率测度

$A \subset B \Rightarrow P(A) \le P(B)$

证明：

$$B = A \bigcup (B \backslash A) = A \bigcup \left(A^c \bigcap B\right)$$
$$P(B) = P\left(A \bigcup \left(A^c \bigcap B\right)\right) = P(A) + P\left(A^c \bigcap B\right)$$
$$\because P\left(A^c \bigcap B\right) \ge 0, \therefore P(A) \le P(B)$$

引理，For any events A and B, $P\left(A \bigcup B\right) = P(A) + P(B) - P\left(A \bigcap B\right)$

连续性定理：

If $A_n \to A$, then $P\left(A_n\right) \to P(A)$

Df: A_n monotone increase, then $A_1 \subset A_2 \ldots,\ \lim_{n \to \infty} A_n = \bigcup_{i=1}^{infty} A_i$
证明思路：用测度的可加性来进行证明。但可加性只针对于不想交的情况，因此需要对当前情形进行重新构造：
Let $B_1 = A_1, B_2 = \left\{\omega \in \Omega, w \in A_2, \omega \notin A_1\right\} \ldots$
So, $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$, then $P\left(A_n\right) = P\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} P\left(B_i\right) = P\left(\sum_{i=1}^{\infty} B_i\right) = P(A)$

PART 03 独立事件 Independent Event

1.事件独立

AB事件独立，则$P(AB) = P\left(A \bigcap B\right) = P(A)P(B)$。

多个事件独立：A set of events $\left\{A_i, i \in I\right\}$ is independent, then, $P\left(\bigcap_{i=j} A_i\right) = \prod_{i \in J} P\left(A_i\right)$ for every finite subsets *j* of *I.*

2.条件独立

Assume $P(B) > 0$, A在给定B下的条件概率为，$P\left(A \mid B\right) = \dfrac{P(AB)}{P(B)}$

如果A独立于B，则$P\left(A \mid B\right) = P(A)$

3.Bayes Theorem

（1）The law of total probability

Let $A_1, A_2 \ldots A_k$ be a partition of Ω, then, for any events,

$$P(B) = \sum_{i=1}^{k} P\left(B \mid A_i\right) P\left(A_i\right)$$

证明：构造$C_i = B \bigcap A_i$，then, $\bigcup_{i=1}^{k} C_i = B$

（2）Bayes Theorem

Let $A_1, A_2 \ldots A_k$ be a partition of Ω, such that $P\left(A_i\right) > 0, P(B) > 0$, then,

$$P\left(A_i \mid B\right) = \frac{P\left(B \mid A_i\right) P\left(A_i\right)}{\sum_{j=1}^{k} P\left(B \mid A_i\right) P\left(A_i\right)}$$