人工智能中的数学讲义

方聪

北京大学

摘要

本讲义收录了人工智能中的数学课程中的主要概念与课程习题。概率与统计讲义内容摘录于陈家鼎、郑忠国《概率与统计》教材与复熹和张原概率与统计课程课件。图论内容摘录于耿素云、屈婉玲、王捍贫《离散数学教程》。本讲义版权归上述作者，不会出版。讲义仅供于上该课程的同学们学习参考，讲义的错误会不断修正。感谢张乙沐、张海涵对讲义整理的帮助。

1.1 随机事件及其运算

1.1.1 随机事件

样本空间和样本点：随机实验 E 中所有可能结果组成的集合称为 E 的**样本空间**，记为 Ω。样本空间中的元素称为样本点，记为 ω

- E_1：抛掷硬币，观察正面 H，反面 T 出现的情况。

$$\Omega_1 = \{H, T\}.$$

- E_2：抛掷一枚硬币 3 次，观察正面 H，反面 T 出现的情况。

$$\Omega_2 = \{HHH, HHT, HTH, THH\ HTT, THT, TTH, TTT\}.$$

- E_3：抛掷一枚硬币 3 次，观察正面出现的次数。

$$\Omega_3 = \{0, 1, 2, 3\}.$$

随机现象的某些样本点组成的集合称为**随机事件**，简称为事件，常用 $A, B, C, \cdots$ 表示

例如, E 为抛掷一枚骰子, 事件 A =“出现奇数点”, 即 $A = \{1, 3, 5\}$, 是样本空间 $\Omega = \{1, 2, 3, 4, 5, 6\}$ 的一个子集

事件的频率：设 μ 是 n 次实验中事件 A 发生的次数，则事件 A 发生的频率 $\frac{\mu}{n}$，随着实验次数 n 增大，频率会在某一数值 p 附近摆动，称为该事件的**概率**，记为 $P(A) = p$

由于频率 $\frac{\mu}{n}$ 总在 $0, 1$ 之间，我们有：

$$0 \leqslant P(A) \leqslant 1$$

例如投一枚硬币 n 次，出现 μ 次正面，则 $\frac{\mu}{n} \overset{n\to\infty}{\to} p$。其中，主观概率 p 为事件的置信度，概率是可能性大小的度量。大概率事情易发生，小概率事情不易发生。

1.1.1.1 事件的交和并

定义 2.1 设有事件 A 和事件 B，如果 A 发生，则 B 必发生，那么称事件 B 包含事件 A（或称事件 A 在 B 中），并记为 $$A \subset B \text{ (或 } B \supset A)$$

定义 2.2 如果事件 A 包含事件 B，同时事件 B 包含事件 A，则事件 A 和事件 B 相等，并记为

$$A = B$$

定义 2.3 设 A 和 B 都是事件，则 “A 或 B” 表示这样的事件 C：C 发生当且仅当 A 或 B 中至少有一个发生，该事件 C 叫做 A 与 B 的并，记为 $A \cup B$。

例 2.1 （对应郑书例 2.1）在桌面上，投掷两枚匀称的硬币，A 表示 “恰好一枚国旗朝上”，B 表示 “两枚国旗朝上”，C 表示 “至少一枚国旗朝上”，则 $C = A \cup B$.

对于并运算，有以下性质，我们恒记必然事件为 U，不可能事件为 V：

$$A \cup B = B \cup A$$
$$A \cup U = U,\ A \cup V = V$$

定义 2.4 设 A 和 B 都是事件，则 “A 且 B” 表示这样的事件 C：C 发生当且仅当 A 和 B 都发生，该事件 C 叫做 A 与 B 的交，记为 $A \cap B$，也简记为 AB。

在例 2.1 中，$A \cap C = A$，$B \cap C = C$，$A \cap B = V$

对于交运算，有以下性质：

$$A \cap B = B \cap A$$
$$A \cap U = A,\ A \cap V = V$$

1.1.1.2 事件的余和差

定义 2.5 设 A 是事件，称 “非 A” 是 A 的对立事件（或称余是事件），其含义为，“非 A” 发生当且仅当 A 不发生，常常用 $\overline{A}$ 表示 “非 A”，也用 A^c 表示 “非 A”。

由定义知 $\overline{(\overline{A})} = A$，$\overline{U} = V$，$\overline{V} = U$

定义 2.6 设 A 和 B 都是事件，则两个事件的差 “A 减去 B” 表示这样的事件 C：C 发生当且仅当 A 发生而 B 不发生，该事件 C 记为 $A - B$（或 $A \backslash B$）

由定义知，$A - B = A \cap \bar{B}$
画图法确定关系。

1.1.1.3 事件运算的性质

事件的基本运算还有以下性质：

- $A \cup (B \cup C) = (A \cup B) \cup C$ “并”的结合律
- $A \cap (B \cap C) = (A \cap B) \cap C$ “交”的结合律
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ 分配律
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ 分配律
- $A \cup A = A$，$A \cap A = A$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$ 对偶律
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$ 对偶律

多个事件的交和并：
设 $A_1, A_2, \cdots, A_n$ 是 n 个事件，则“$A_1, A_2, \cdots, A_n$”的并是指这样的事件：它发生当且仅当 $A_1, A_2, \cdots, A_n$ 中至少一个发生，常常用 $\bigcup\limits_{i=1}^{n} A_i$ 表示 $A_1, A_2, \cdots, A_n$ 的并
设 $A_1, A_2, \cdots, A_n$ 是 n 个事件，则“$A_1, A_2, \cdots, A_n$”的交是指这样的事件：它发生当且仅当 $A_1, A_2, \cdots, A_n$ 这 n 个事件都发生，常常用 $\bigcap\limits_{i=1}^{n} A_i$ 表示 $A_1, A_2, \cdots, A_n$ 的交，也用 $A_1 A_2 \cdots A_n$ 表示这个“交”

> 实际应用中，还需定义无穷多事件的并与交
> 设 $A_1, A_2, \cdots, A_i, \cdots$ 是一列事件，则 B 是指这样的事件：B 发生当且仅当这些 $A_i (i = 1, 2, \cdots)$ 中至少一个发生，这个 B 叫做诸 A_i 的并，记为 $\bigcup\limits_{i=1}^{\infty} A_i$，有时也写为 $A_1 \cup A_2 \cup \cdots$
> 设 $A_1, A_2, \cdots, A_i, \cdots$ 是一列事件，则 C 是指这样的事件：C 发生当且仅当这些 $A_i (i = 1, 2, \cdots)$ 都发生，这个 C 叫做诸 A_i 的交，记为 $\bigcap\limits_{i=1}^{\infty} A_i$，有时也写为 $A_1 A_2 \cdots$

例: 取 $X \in \mathbb{R}$, 事件 A_i 为 $X \in [\frac{1}{i+1}, \frac{1}{i}]$, 事件 B_i 为 $X \in [0\ \frac{1}{i}]$。则事件 $\bigcup\limits_{i=1}^{n} A_i$ 发生等价于 $X \in [\frac{1}{n+1}, 1]$, 事件 $\bigcap\limits_{i=1}^{n} B_i$ 发生等价于 $X \in [0, \frac{1}{n}]$。进而当 $n \to \infty$ 时事件 $\bigcup\limits_{i=1}^{\infty} A_i$ 发生等价于 $X \in (0, 1]$, 事件 $\bigcap\limits_{i=1}^{\infty} B_i$ 发生等价于 $X = 0$。

并的更一般定义是，设 $\{A_a, a \in \Gamma\}$ 是一族事件（其中 Γ 是任何非空集，每个 $a \in \Gamma$ 对应一个事件 A_a），这些事件 A_a 的“并”是指这样的事件 B：B 发生当且仅当至少一个 A_a 发生，这个 B 常常记为 $\bigcup\limits_{a \in \Gamma} A_a$，类似可以定义一族事件的交 $\bigcap\limits_{a \in \Gamma} A_a$

例 2.3:（对应郑书例 2.3）一射手向一个目标连续射击，设 $A_1=$“第一次射击，命中”，$A_i=$“前 $i-1$ 次射击都未命中，第 i 次射击命中”（$i=2,3,\cdots$），$B=$“终于命中”，则 $B=\bigcup\limits_{i=1}^{\infty}A_i$

例 2.4:（对应郑书例 2.4）一射手向一个目标连续射击，设 $A_i=$“第 i 次射击，未命中目标”（$i=2,3,\cdots$）则 $\bigcap\limits_{i=1}^{\infty}A_i=$“每次均未命中目标”

不难验证，对可列个事件的并和交有以下规律：

- $A\cup(\bigcap\limits_{i=1}^{\infty}B_i)=\bigcap\limits_{i=1}^{\infty}(A\cup B_i)$ 分配律
- $A\cap(\bigcup\limits_{i=1}^{\infty}B_i)=\bigcup\limits_{i=1}^{\infty}(A\cap B_i)$ 分配律
- $\overline{(\bigcup\limits_{i=1}^{\infty}A_i)}=\bigcap\limits_{i=1}^{\infty}\overline{A_i}$ 对偶律
- $\overline{(\bigcap\limits_{i=1}^{\infty}A_i)}=\bigcup\limits_{i=1}^{\infty}\overline{A_i}$ 对偶律

1.1.1.4 互斥事件

互不相容的事件

如果事件 A 和事件 B 不能都发生，即 $A\cap B=V$，则称 A 和 B 是互不相容的事件（也称互斥的事件）

称事件 $A_1,\cdots A_n$ 互不相容，若对任何 $i\neq j(i,j=1,\cdots n)$，A_i 与 A_j 互不相容

例如，抛掷两枚硬币，事件“恰好一枚国徽朝上”和事件“两枚都是国徽朝上”是互不相容的。不难看出，对任何事件 A，A 和 $\overline{A}$ 是互不相容的

- 加法公式：$A_1,A_2,\ldots$ 互不相容，则：

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

- $P(A\bigcup B)=P(A)+P(B)-P(AB)$

1.2 概率的公理化定义

概率空间子类: 设 Ω 为样本空间，$\mathscr{F}$ 为 Ω 的一些子集构成的集类。若 $\mathscr{F}$ 满足以下三个条件：(1) $\Omega \in \mathscr{F}$，(2) $A \in \mathscr{F} \Rightarrow A^c \in \mathscr{F}$，(3) $\{A_n\}_{n\in\mathbb{N}} \subsetneqq \mathscr{F} \Rightarrow \bigcup\limits_{n=1}^{\infty} A_n \in \mathscr{F}$，则称 $\mathscr{F}$ 为概率空间子类

例:

- $\mathscr{F} = \{\emptyset, \Omega\}$ 平凡概率空间子类
- $\mathscr{F} = \{\emptyset, \Omega, A, \overline{A}\}$ 包含 A 的最小概率空间子类
- $\mathscr{F} = \{A|A \subset \Omega\}$ Ω 上的最大概率空间子类
- $\Omega = \{\omega_1, \cdots, \omega_n\}$，则 Ω 所有子集构成的概率空间子类共有 2^n 个元素

定义: 设 $\mathscr{F}$ 是满足上述条件的概率空间子集类。概率 $P = P(\cdot)$ 是 $\mathscr{F}$ 上面定义的实值函数，满足:

- 非负性：$P(A) \geqslant 0$ 对于一切 $A \in \mathscr{F}$
- 规范性：$P(\Omega) = 1$
- 可列可加性：若 $A_n \in \mathscr{F} (n = 1, 2, \cdots)$ 两两不相交, 则

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

$(\Omega, \mathscr{F}, P)$ 为概率空间

例 1: 假定 $\Omega = \{\omega_1, \cdots, \omega_n\}$，$\mathscr{F}$ 为全体子集构成的概率空间子类。
设 $p_1, \cdots, p_n$ 为 n 个非负实数，且满足 $\sum\limits_{i=1}^{n} p_i = 1$。令

$$\mathbb{P}(\emptyset) = 0, \quad \mathbb{P}(A) = \sum_{j=1}^{k} p_{i_j}, \quad A = \{\omega_{i_1}, \cdots, \omega_{i_k}\}, k = 1, \cdots, n$$

则 $\mathbb{P}$ 为 $(\Omega, \mathscr{F})$ 上概率。
概率 P 有以下性质:
(1) $P(\emptyset) = 0$;
(2) 若 $A \in \mathscr{F}$，则 $P(A^c) = 1 - P(A)$;
(3) 若 $A_1, \cdots, A_n$ 都属于 $\mathscr{F}$ 且两两不相交，则

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) \tag{1.2.1}$$

(4) 若 $A \subset B$，$A \in \mathscr{F}$，$B \in \mathscr{F}$，则 $P(A) \leqslant P(B)$，且

$$P(B-A)=P(B)-P(A) \tag{1.2.2}$$

(5) 若 $A_n \subset A_{n+1}$，$A_n \in \mathscr{F}(n=1,2,\cdots)$，则

$$P(\bigcup_{n=1}^{\infty} A_n)=\lim_{n\to\infty} P(A_n) \tag{1.2.3}$$

(6) 若 $A_n \supset A_{n+1}$，$A_n \in \mathscr{F}(n=1,2,\cdots)$，则

$$P(\bigcap_{n=1}^{\infty} A_n)=\lim_{n\to\infty} P(A_n) \tag{1.2.4}$$

(7) 若 $A_n \in \mathscr{F}(n=1,2,\cdots)$，则

$$P(\bigcup_{n=1}^{\infty} A_n) \leqslant \sum_{n=1}^{\infty} P(A_n) \tag{1.2.5}$$