

Lecture10: 贝叶斯劝说

一、贝叶斯劝说：背景与例子

考虑导师写推荐信将学生推荐至企业的例子：

- 两个参与人：导师（信号发送者）和企业（信号接收者）
 - 导师的任务是向企业为每位学生写推荐信，通过推荐信的好坏向企业发送信号；
 - 企业的任务是对一个学生，在接收到导师的推荐信（信号）后必须做出以下两种决策之一：雇用（hiring）和不雇用（not hiring）；
 - 学生不是博弈参与方，因为学生只被动接受结果，没有自己的策略。
- 学生有两种类型：优秀（excellent）或一般（average）
 - 学生的类型对导师而言是已知的。对企业则是不完全信息；
 - 与不完全信息博弈中的假设一致，可以认为学生的类型是自然按一定的先验概率随机抽取的；
 - 企业对学生类型有先验分布 $\mu_0(\text{average}) = 0.75, \mu_0(\text{excellent}) = 0.25$ ，这一先验概率分布也符合导师已知的实际学生类型分布，因此如果随机抽取一个学生，则企业和导师对学生类型有共同的先验分布。

还需要定义参与人的效用函数：

- 假设导师希望推荐出去的学生越多越好，因此企业只要雇用一個学生，则导师获得效用 1，否则效用为 0；
- 企业则希望招收到优秀的学生，因此在招收到优秀的学生时获得效用为 1，招收普通学生时获得效用为 -0.5。

因此接下来需要形式化定义导师的策略，即形式化“发信号”这一策略。

形象地说，发信号就是通过好或坏的推荐信来向企业表明学生是优秀的或一般的。形式化地说，导师写推荐信的策略就是如下两个条件概率分布 $\pi(\cdot | \text{excellent})$ 和 $\pi(\cdot | \text{average})$ （又称信号机制（signaling scheme））：

$$\pi(e | \text{excellent}), \pi(a | \text{excellent})$$

$$\pi(e | \text{average}) | \pi(a | \text{average})$$

- 其中 e 和 a 分别表示描述学生为优秀类型和一般类型的推荐信；
- $\pi(A|B)$ 表示当学生属于 B 类型时，导师在推荐信中给学生描述的类型为 A 的概率；

一个值得注意的点是，导师和企业之间存在长期关系，因此导师的策略是企业看到推荐信之前就已知的，因为企业在与导师的长期关系中可以验证导师的策略。

1. 诚实推荐

第一个信号的例子是，导师完全诚实地推荐学生，即为优秀的学生写好的推荐信，为一般的学生写一般的推荐信，故此时信号机制为

$$\pi(e | \text{excellent}) = 1, \pi(a | \text{excellent}) = 0$$

$$\pi(e | \text{average}) = 0, \pi(a | \text{average}) = 1$$

企业知道导师的推荐信是诚实的，因此将接收所有推荐信中写优秀的学生，拒绝所有推荐信中写一般的学生；

- 此时导师把优秀的同学推出，故在每个学生上的期望效用为 0.25；
- 企业接收所有优秀同学，故在每个学生上的期望效用也为 0.25。

2. 完全不诚实推荐

自然地，导师会认为提供诚实的推荐信能推荐出的学生太少，因此极端的导师可能会选择为每个学生都写好的推荐信，这就构成了第二个信号的例子：因此此时企业看到的全是好的推荐信，因此只能保持先验概率去判断学生的好坏：

- 由于每个学生是优秀类型的概率只有 0.25，因此如果企业雇用任意一

个学生，其期望效用为 $0.25 \times 1 - 0.75 \times 0.5 = -0.125$ ，因此企业不会雇用任何一个学生；

- 此时导师和企业的期望效用均为 0。

3. 部分诚实推荐

从之前的例子可以看出，完全诚实和完全不诚实的策略都不是最优的，因此需要更精妙的设计。

之后会证明如下信号最优（即导师效用最大化）：

$$\pi(e \mid \text{excellent}) = 1, \pi(a \mid \text{excellent}) = 0$$

$$\pi(e \mid \text{average}) = \frac{2}{3}, \pi(a \mid \text{average}) = \frac{1}{3}$$

即导师会给所有优秀的学生写好的推荐信，对一般的学生则以 $2/3$ 的概率写好的推荐信。

- 当企业看到好的推荐信时，应当有 $\frac{1}{3}$ 的概率认为学生是真的优秀，有 $\frac{2}{3}$ 的概率认为学生一般；
- 企业看到一般的推荐信时则一定认为学生一般。

即企业看到推荐信后对学生类型的后验概率为（记 $\mu_A(B)$ 是看到 A 类型推荐信后认为学生属于类型 B 的概率，其实就是 $\mu(B \mid A)$ ）：

$$\mu_e(\text{excellent}) = \frac{1}{3}, \mu_e(\text{average}) = \frac{2}{3}$$

$$\mu_a(\text{excellent}) = 0, \mu_a(\text{average}) = 1$$

根据后验概率，看到一般的推荐信代表学生一般，故企业不会雇用；

当看到好的推荐信时，有 $\frac{1}{3}$ 的概率认为是优秀学生， $\frac{2}{3}$ 的概率认为是一般学生，故对于一个好推荐信对应的学生；

- 企业雇用带来的效用为 $\frac{1}{3} \times 1 - \frac{2}{3} \times 0.5 = 0$ ，即是无差异的；
- 假设在信号接收者策略无差异的情况下，信号接收者会选择有利于信号发送者的决策，即这里的企业会选择雇用；
- 故企业期望效用为 0，而根据导师的策略 π ，导师将全部的优秀学生以及 $\frac{2}{3}$ 的一般学生推荐进入企业，因此导师的期望效用为 $0.25 + 0.75 \times \frac{2}{3} = 0.75$

之后的讨论会严谨说明这一信号是最优的，但现在可以理解这一信号是最优的直观：

- 优秀的学生应当全部被雇用，然后应当尽可能地让一般的学生被雇用；
- 在上述信号机制中，如果进一步增大一般学生写好推荐信的比例，企业看到好推荐信时会认为一般学生比例太大，因此将倾向于不雇用，因此使得企业雇用和不雇用无差异的信号是最优的。

4. 贝叶斯公式

再次回顾贝叶斯公式：设 B 和 A_1, \dots, A_n 为一系列事件，则有

$$\mathbb{P}(A_k \mid B) = \frac{\mathbb{P}(A_k, B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(B \mid A_k) \cdot \mathbb{P}(A_k)}{\sum_{A_i} \mathbb{P}(B \mid A_i) \cdot \mathbb{P}(A_i)}$$

在贝叶斯劝说的场景下，贝叶斯公式应当表达为：

$$\mu_s(\omega) = \frac{\pi(s|\omega)\mu_0(\omega)}{\sum_{\omega' \in \Omega} \pi(s|\omega')\mu_0(\omega')}$$

部分诚实推荐策略下给出的后验概率是符合贝叶斯公式的。

二、模型描述与问题转化

1. 贝叶斯劝说：模型描述

从导师写推荐信的例子中可以提炼出一般的贝叶斯劝说（Bayesian persuasion）模型：

- 两个参与人：信号发送者（导师）和信号接收者（企业）；
- 他们对自然的真实状态 $\omega \in \Omega$ （一个学生优秀 / 一般）有相同的先验分布 $\mu_0 \in \text{int}(\Delta(\Omega))$ ，信号发送者知道状态的实现值（即具体每个学生是优秀还是一般的），但信号接收者不知道；
 - $\Delta(\Omega)$ 表示 Ω 上的概率分布；
 - int 含义是内点，即先验分布保证每个状态的概率都是正的；
- 假定双方都是理性的，即追求效用最大化的，并且都是按照贝叶斯公式

更新信念的；

- 发送者的效用为 $v(a, \omega)$ ，接收者的效用为 $u(a, \omega)$ ；
 - 导师的效用为 $v(\text{hiring}, \omega) = 1, v(\text{not hiring}, \omega) = 0$ （与 ω 无关）；
 - 法官的效用为 $u(\text{hiring}, \text{average}) = -0.5$ 等。

博弈的行动顺序如下（动态博弈需要说明顺序）：

1. 发送者公开（承诺（commit））信号机制 $(S, \pi(s|\omega)), \forall s \in S, \omega \in \Omega$ ；
 - S 称为信号实现空间，例如前面的例子中 $S = \{e, a\}$ ；
 - 故信号机制包含信号实现空间 S 及其在所有现实状态下的条件分布；
 - 于是接收者可以利用贝叶斯公式计算出后验概率 $\mu_s(\omega)$ ；
2. 自然以分布 μ_0 选择 $\omega \in \Omega$ （抽出一个学生是优秀 / 一般的）；
3. 类型为 ω 时发送者以概率 $\pi(s|\omega)$ 发送信号 $s \in S$ ；
4. 接收者收到信号 s 并选择一个行动 $a \in A$ （企业雇用 / 不雇用学生）； a 的选择应当最大化接收者的效用，即

$$a = \arg \max_{a \in A} \mathbb{E}_{\mu_s} [u(a, \omega)]$$

如果有多个最大化效用的选择，假设其选择最大化发送者效用的行动（企业雇用和不雇用无差异时，选择雇用学生）。

5. 发送者获得效用 $v(a, \omega)$ ，接收者获得效用 $u(a, \omega)$ 。

2. 贝叶斯劝说的目标

注意贝叶斯劝说的第一步就是信号发送者公开承诺信号机制：

- 回忆导师写推荐信的例子，这样的情况可以发生在结果可验证的情况；
 - 例如企业可以在雇用后看出学生的能力，或者消费者可以在购买产品后看出产品的实际价值；
 - 因此贝叶斯劝说主要在这样的场景下具有实际价值；
- 此外，贝叶斯劝说模型中，信号发送者优先行动，接收者在看到信号发送者的行动后行动，故最优化问题实际是一个双层优化问题；
 - 此时信号发送者和信号接收者的策略相对于对方的策略都是最优的，并且信号接收者的信念通过贝叶斯公式进行了更新，这一均衡被称为完美贝叶斯均衡（perfect Bayesian equilibrium）。

在理解了贝叶斯劝说的例子、思想以及具体模型后，自然地，我们希望研究有关贝叶斯劝说的如下问题：

- 发送者是否总是可以通过设计信号机制来影响接收者的行为，从而提升自己的效用？如果不是，什么情况下可以？
- 发送者如何设计信号机制以达到最大化自己的效用？最大化效用时信号以及接收者的行为的特点是什么样的？
- 接收者是否愿意接受发送者的信号机制？如果不是，什么情况下可以？

3. 贝叶斯可行

为了解决前两个问题，首先要定义贝叶斯可行（Bayesian plausible）的概念，然后将设计最优信号机制的问题转化为更容易解决的问题。

给定信号机制 $(S, \pi(s | \omega))$ ，任一信号实现 s 都会导致一个后验概率分布 $\mu_s \in \Delta(\Omega)$ ，即对任意的 $s \in S, \omega \in \Omega$ ：

$$\mu_s(\omega) = \frac{\pi(s|\omega)\mu_0(\omega)}{\sum_{\omega' \in \Omega} \pi(s | \omega')\mu_0(\omega')}$$

由于每个 s 都会导致一个后验概率分布，所以所有的 s 将导致 $|S|$ 个后验概率分布，并且所有的后验概率分布本质上都是 Ω 上的分布。根据全概率公式，每个 s 被发出的概率为

$$\mathbb{P}(s) = \sum_{\omega' \in \Omega} \pi(s | \omega')\mu_0(\omega')$$

所以所有 s 将导致一个后验概率分布的分布 $\tau \in \Delta(\Delta(\Omega))$ ，其中概率分布支撑为 $\text{Supp}(\tau) = \{\mu_s\}_{s \in S}$ ，支撑中每一个后验概率 $\mu \in \Delta(\Omega)$ 的概率为：

$$\tau(\mu) = \sum_{s: \mu_s = \mu} \mathbb{P}(s) = \sum_{s: \mu_s = \mu} \sum_{\omega' \in \Omega} \pi(s | \omega')\mu_0(\omega')$$

如果每个后验概率都不同，则支撑中每一个后验概率 $\mu \in \Delta(\Omega)$ 的概率为：

$$\tau(\mu) = \mathbb{P}(s) = \sum_{\omega' \in \Omega} \pi(s | \omega')\mu_0(\omega')$$

例如，回忆导师写推荐信的例子，在最优机制下，信号机制导致的两个后验概率分布分别为

$$\mu_e(\text{excellent}) = \frac{1}{3}, \mu_e(\text{average}) = \frac{2}{3}$$

和

$$\mu_a(\text{excellent}) = 0, \mu_a(\text{average}) = 1$$

这两个后验概率分布不相同，因此 $\text{Supp}(\tau) = \{\mu_e, \mu_a\}$ ，二者概率为

$$\mathbb{P}(e) = \pi(e | \text{excellent})\mu_0(\text{excellent}) + \pi(e | \text{average})\mu_0(\text{average}) = 0.75$$

$$\mathbb{P}(a) = \pi(a | \text{excellent})\mu_0(\text{excellent}) + \pi(a | \text{average})\mu_0(\text{average}) = 0.25$$

基于上述记号，可以给出贝叶斯可行的定义：

定义：称 τ 由信号导致，如果存在信号机制 $(S, \pi(s | \omega))$ 对应的后验概率分布的分布为 τ 。称一个后验概率分布的分布 τ 是贝叶斯可行的，如果

$$\sum_{\text{Supp}(\tau)} \mu \tau(\mu) = \mu_0$$

即后验概率的期望等于先验概率。

再次强调，这里不同的后验概率分布可以求和的原因在于，本质上不同的后验概率分布都是 Ω 上的概率分布。例如可以检查导师写推荐信的例子是否满足贝叶斯可行性：

$$\tau(\mu_e) \cdot \mu_e(\text{excellent}) + \tau(\mu_a) \cdot \mu_a(\text{excellent}) = 0.25 = \mu_0(\text{excellent})$$

$$\tau(\mu_e) \cdot \mu_e(\text{average}) + \tau(\mu_a) \cdot \mu_a(\text{average}) = 0.75 = \mu_0(\text{average})$$

事实上，导师写推荐信的例子满足贝叶斯可行性并非偶然：只要 τ 是信号机制导致的，信号机制必然贝叶斯可行：对任意的 $\omega \in \Omega$ 有

$$\begin{aligned} \sum_{\text{Supp}(\tau)} \mu(\omega) \tau(\mu) &= \sum_{s \in S} \mu_s(\omega) \mathbb{P}(s) \\ &= \sum_{s \in S} \frac{\pi(s | \omega) \mu_0(\omega)}{\sum_{\omega' \in \Omega} \pi(s | \omega') \mu_0(\omega')} \mathbb{P}(s) \\ &= \sum_{s \in S} \pi(s | \omega) \mu_0(\omega) = \mu_0(\omega) \end{aligned}$$

4. 问题转化

由此可知，一个信号机制可以导致一个贝叶斯可行的后验概率分布的分布。反之，一个贝叶斯可行的后验概率分布的分布 τ 是否一定对应一个信号机制呢？答案是肯定的：

定理：一个后验概率分布的分布 $\tau \in \Delta(\Delta(\Omega))$ 是贝叶斯可行的当且仅当存在一个信号机制 $(S, \pi(s | \omega))$ 使得 τ 是由该信号机制导致的。

证明：“当”的部分上页已经说明，“仅当”部分的证明需要从 τ 构造出信号机制 $(S, \pi(s | \omega))$ 。根据 $\mathbb{P}(s) \mu_s(\omega) = \pi(s | \omega) \mu_0(\omega)$ 自然地可以写出这一构造：定义 $S = \text{Supp}(\tau)$ ，且对任意的 $s \in S$ ，定义

$$\pi(s | \omega) = \frac{\tau(\mu_s) \mu_s(\omega)}{\mu_0(\omega)}$$

不难验证这的确构造出了一个合理的信号机制。

- 因此，一个信号机制等价于一个贝叶斯可行的后验概率分布的分布；
 - 进而可以导致接收者行动的分布，因为一个后验概率分布就对应接收者的一个最优行动；
 - 显然，只要接收者行动分布一定，那么发送者的效用也是确定的；
- 因此是否存在一个信号机制使得发送者达到效用 v^* ，只需要考虑是否存在一个贝叶斯可行的后验概率分布的分布 τ 使得发送者效用达到 v^* ；
- 因此设计最优信号机制的问题可以转化为设计一个贝叶斯可行的后验概

率分布的分布 τ 使得接收者的效用最大化。



三、最优信号机制