ZJU 2024–2025 学年春夏学期

计算经济学讨论班

Lec 6: 贝叶斯劝说

吴一航　yhwu_is@zju.edu.cn

浙江大学计算机科学与技术学院

CONTENT

目录

检察官和法官的故事

考虑检察官劝说法官判决的例子：假设法官（信号接收者）对于一个被告人，必须做出以下两种决策之一：判决有罪（convict）或无罪释放（acquit）。

- 被告人有两种类型（state of the world）：有罪（guilty）或无罪（innocent）；
- 法官在公正判决下获得的效用为 1：如果有罪被判有罪，无罪被判无罪，否则效用为 0；
- 检察官（信号发送者）为法官提供有关被告的证据（发送信号），如果被告人判有罪，检察官获得效用 1，否则效用为 0；
- 法官和检察官对于被告人的类型有相同的先验概率分布：$\mathbb{P}(\text{guilty}) = 0.3, \mathbb{P}(\text{innocent}) = 0.7$。

检察官和法官的故事

检察官进行调查收集有关被告人的证据，法律要求检察官报告其全部结果。

- 也就是说，检察官收集到的证据必须全部如实提供;
- 因此检察官的策略是选择一个调查收集证据的策略（因为如实披露所以收集的策略就是披露策略），希望改变法官的判决，使得被判有罪的越多越好（检查官效用最大化）。

形式化地说，调查就是一个 $\pi(\cdot \mid \text{guilty})$ 和 $\pi(\cdot \mid \text{innocent})$ 的信号实现，并且这一信号实现应该公开给法官。下面将给出几个信号的例子。

例：精确证据

第一个信号的例子是检察官努力寻找所有精确的证据，例如 DNA 检测、视频监控等，此时的信号为

$$\pi(i \mid \text{innocent}) = 1, \pi(i \mid \text{guilty}) = 0$$
$$\pi(g \mid \text{innocent}) = 0, \pi(g \mid \text{guilty}) = 1$$

因为是精确的证据，所以无罪的被告人的证据一定是 i（无罪证据），有罪的被告人的证据一定是 g（有罪证据）。

由于信号是公开的，法官知道检察官的调查策略，此时法官知道提供证据 i 的被告人一定是无罪，提供证据 g 的被告人一定是有罪，因此法官的最佳决策是根据检察官提供的证据判决。法官的期望效用为 1，检察官的期望效用为 0.3。

例：全有罪证据

自然地，检察官会认为提供精确的证据吃力不讨好，特别是检察官希望被判有罪的越多越好。因此，检察官可能会选择提供全是有罪证据的调查策略，例如寻找证据时只找有罪的证据，此时的信号为

$$\pi(i \mid \text{innocent}) = 0, \pi(i \mid \text{guilty}) = 0$$
$$\pi(g \mid \text{innocent}) = 1, \pi(g \mid \text{guilty}) = 1$$

然而此时法官看见有罪的证据只能保持先验概率，即有罪的概率是 0.3，无罪的概率是 0.7，因此最大化效用的策略是全部判无罪，此时法官的期望效用为 0.7，检察官的期望效用为 0。

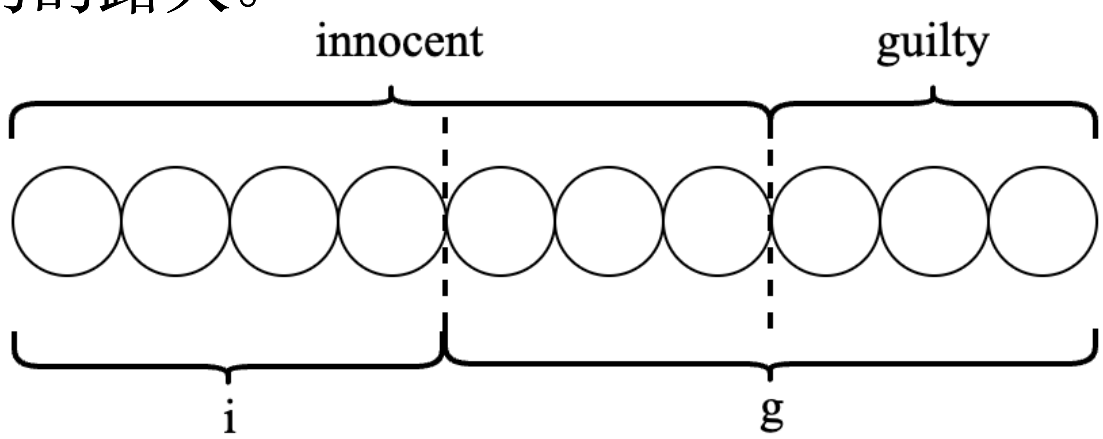
例：模糊证据

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：

$$\pi(i \mid \text{innocent}) = \frac{4}{7}, \pi(i \mid \text{guilty}) = 0$$

$$\pi(g \mid \text{innocent}) = \frac{3}{7}, \pi(g \mid \text{guilty}) = 1$$

即检察官寻找的证据对于有罪的被告而言一定会指向其犯罪事实，而对无罪的被告只有一定的概率指向犯罪。这样的证据是模糊的，例如传唤被告身边的人描述其品行，或者传唤距离案发地点较远的的路人。



例：模糊证据

此时，法官将根据贝叶斯公式计算看到无罪和有罪证据下被告人有罪或无罪的后验概率（为了与原论文记号统一，记 $\mathbb{P}(\text{innocent} \mid i) = \mu_i(\text{innocent})$，其它类似）：

$$\mu_i(\text{innocent}) = 1, \mu_i(\text{guilty}) = 0$$

$$\mu_g(\text{innocent}) = 0.5, \mu_g(\text{guilty}) = 0.5$$

因此法官看到无罪证据时，认为被告人无罪的概率是 1，最优决策是判无罪；看到有罪证据时，认为被告人有罪的概率是 0.5，最优决策是判有罪（实际上是无差异的，但我们假设无差异时信号接收者的行为是最大化发送者效用的行为）。不难得到法官的期望效用为 0.7，检察官的期望效用为 0.6。

直观上来看，如果进一步增加有罪证据的权重，法官看到有罪证据时将倾向于判无罪，因此这一使得法官判决有罪无罪无差异的信号是最优的。

CONTENT

目录

贝叶斯劝说：模型描述

从检察官与法官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（检察官）和信号接收者（法官）；
- 他们对世界的真实状态 $\omega \in \Omega$（一个犯人有罪 / 无罪）有相同的先验分布 $\mu_0 \in \mathrm{int}(\Delta(\Omega))$；
  - 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 $u(a, \omega)$；
  - 检察官的效用为 $v(\text{convict}, \omega) = 1, v(\text{acquit}, \omega) = 0$（与 ω 无关）；
  - 法官的效用为 $u(\text{convict}, \text{guilty}) = 1$ 等。
- 博弈的行动顺序如下：

贝叶斯劝说：模型描述

1. 发送者公开（或承诺（commit））其信号机制 $(S, \pi(s \mid \omega)), \forall s \in S, \omega \in \Omega$;
   - S 称为信号实现空间，例如前面的例子中 $S = \{i, g\}$;
   - 因此信号机制包含信号实现空间 S 以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 $\omega \in \Omega$;
3. 发送者以概率 $\pi(s \mid \omega)$ 发送信号 $s \in S$;
4. 接收者收到信号 s 并选择一个行动 $a \in A$（判有罪 / 无罪）;
   - a 的选择应当最大化接收者的效用，即

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

   如果有多个最大化效用的选择，假设其会选择最大化发送者效用的行动。
5. 发送者获得效用 $v(a, \omega)$，接收者获得效用 $u(a, \omega)$。

与 Cheap Talk 的区别

贝叶斯劝说与（狭义的）信号博弈的区别是显然的，因为狭义的信号博弈发送信号需要成本，因此需要比较的是贝叶斯劝说和扑克牌博弈中的 Cheap Talk 模型有什么区别?

- 显然的区别在于 Cheap Talk 中发送方的策略是不公开的，而在贝叶斯劝说中信号机制是提前公布的；
  - 例如检察官提供证据必须如实，类似的例子还有导师写推荐信、广告宣传等，因为这些情况下提供的信息都是可验证的（verifiable）；
- 因此贝叶斯劝说中的接收方可以直接计算出后验概率，然后根据后验概率选择最优行动；
  - 因此贝叶斯劝说的均衡更像是子博弈完美，可以用逆向归纳的思想分析，无需关注完美贝叶斯均衡；
- 介绍了最优信号机制后会通过计算具体显示二者的区别。

贝叶斯劝说的目标

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？
- 信号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？

其中第三个问题比较简单，我们首先考虑第三个问题。

贝叶斯劝说对接收者的影响

命题

在任意信号机制 $S, \pi(s \mid \omega)$ 下，接收者的效用都不会低于其在没有信号的情况下的效用。

任取信号机制 $S, \pi(s \mid \omega)$，当接收者看到 $s \in S$ 时，其效用为

$$\begin{aligned}\max_{a \in A} \mathbb{E}_{\mu_s}[u(a, w)] &= \max_{a \in A} \sum_{\omega \in \Omega} \mu_s(\omega) u(a, \omega) \\ &= \max_{a \in A} \sum_{\omega \in \Omega} \frac{\pi(s \mid \omega)\mu_0(\omega)}{\sum_{\omega' \in \Omega} \pi(s \mid \omega')\mu_0(\omega')} u(a, \omega)\end{aligned}$$

因此在信号机制 $S, \pi(s \mid \omega)$ 下，其期望效用为

$$
\begin{aligned}
&\sum_{s\in S}\mathbb{P}(s)\cdot\max_{a\in A}\mathbb{E}_{\mu_s}[u(a,w)]\\
&=\sum_{s\in S}\left(\sum_{\omega'\in\Omega}\pi(s\mid\omega')\mu_0(\omega')\right)\cdot\max_{a\in A}\mathbb{E}_{\mu_s}[u(a,w)]\\
&=\sum_{s\in S}\max_{a\in A}\sum_{\omega\in\Omega}\pi(s\mid\omega)\mu_0(\omega)u(a,\omega)\\
&\geqslant\max_{a\in A}\sum_{s\in S}\sum_{\omega\in\Omega}\pi(s\mid\omega)\mu_0(\omega)u(a,\omega)\\
&=\max_{a\in A}\sum_{\omega\in\Omega}\left(\sum_{s\in S}\pi(s\mid\omega)\right)\mu_0(\omega)u(a,\omega)\\
&=\max_{a\in A}\sum_{\omega\in\Omega}\mu_0(\omega)u(a,\omega)=\max_{a\in A}\mathbb{E}_{\mu_0}[u(a,w)]
\end{aligned}
$$

由此可知命题成立。

贝叶斯可行

为了解决剩下的问题，我们首先要定义**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 $(S, \pi(s \mid \omega))$，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 $s \in S, \omega \in \Omega$：

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

由于每个 s 都会导致一个后验概率分布，所以所有的 s 将共同导致一个后验概率分布的分布 $\tau \in \Delta(\Delta(\Omega))$，其中 $\mathrm{Supp}(\tau) = \{\mu_s\}_{s \in S}$，则对任意的 $\mu \in \Delta(\Omega)$：

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

贝叶斯可行

称 μ 由信号导致，如果存在信号机制 $(S, \pi(s \mid \omega))$ 使得 $\tau(\mu) > 0$。称一个后验概率分布的分布 τ 是**贝叶斯可行**的，如果

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

即后验概率的期望等于先验概率。如果 τ 是信号机制导致的，信号机制必然贝叶斯可行：结合上页两式，对任意的 $\omega \in \Omega$ 有

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

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 $(S, \pi(s \mid \omega))$ 使得 τ 是由该信号机制导致的。

"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 $(S, \pi(s \mid \omega))$。根据 $\mathbb{P}(s)\mu_s(\omega) = \pi(s \mid \omega)\mu_0(\omega)$ 自然地可以写出这一构造：定义 $S = \mathrm{Supp}(\tau)$，且对任意的 $s \in S$，定义

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

不难验证这的确构造出了一个合理的信号机制。

问题转化

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。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水平 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 使得接收者的效用达到 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**。

CONTENT

目录

最优信号机制问题

问题转化后，我们需要解决的问题是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。我们首先将问题形式化：记后验概率为 μ 时，接收者的最优行动为 $\hat{a}(\mu)$，则发送者的期望效用为

$$\hat{v}(\mu) = \mathbb{E}_\mu v(\hat{a}(\mu), \omega)$$

此处的期望针对 ω，因此在检察官的例子中是可以省略的。基于此，我们可以定义最优信号机制问题：

$$\max_\tau \mathbb{E}_\tau \hat{v}(\mu)$$

$$\text{s.t.} \sum_{\mathrm{Supp}(\tau)} \mu\tau(\mu) = \mu_0$$

显示原理

然而，问题在转化到最优后验概率设计后，后验概率的设计空间仍然非常大。我们知道，一个后验概率分布对应一个接收者的最优行动，直观来看如果有 $|A|$ 种后验概率（故 $|A|$ 种信号实现）诱导出 n 种行动就足够了。

接下来的定理表明这一直观是正确的，这一结论称为贝叶斯劝说的"显示原理"，以表示其与机制设计中的显示原理的关联：

- 因此贝叶斯劝说也称为**信息设计（information design）**，信息设计和机制设计都是使他人行动按照自己设想进行的方式，只是信息设计通过改变他人信念实现，而机制设计通过设计激励实现；
- 这两个显示原理都是缩小信息/机制设计的空间，信息设计中的显示原理是一个信号实现对应接收者的一个行动（直观来看就是为接收者推荐了一个行动），机制设计中的显示原理则表明只需设计直接显示机制即可，此时参与人会诚实显示自己的信息。

显示原理

显示原理

存在一个信号机制使得发送者的效用达到 v^* 当且仅当存在一个**直接（straightforward）信号机制**使得接收者的效用达到 v^*。其中直接信号机制是指满足 $S \subseteq A$ 并且接收者的最优行动等于信号实现的信号。

放在检察官与法官的例子中，直接信号机制指信号实现空间 $S = \{\text{guilty}, \text{innocent}\}$ 且当接收者看到有罪信号时判有罪，看到无罪信号时判无罪地信号。因此显示原理表明，最优信号机制设计所需的信号实现数目（后验概率数目）是不超过接收者行动数目的。

定理的证明是简单的，只需要注意到，如果两个信号实现 s 和 s' 出现时接收者的最优行动相同（记为 i^*），即

$$\sum_{\omega\in\Omega}\mu_s(\omega)u(i^*,\omega)\geqslant\sum_{\omega\in\Omega}\mu_s(\omega)u(i,\omega),\forall i\in A$$

$$\sum_{\omega\in\Omega}\mu_{s'}(\omega)u(i^*,\omega)\geqslant\sum_{\omega\in\Omega}\mu_{s'}(\omega)u(i,\omega),\forall i\in A$$

将第一式乘以 p，第二式乘以 $q=1-p$，可得

$$\sum_{\omega\in\Omega}(p\mu_s(\omega)+q\mu_{s'}(\omega))u(i^*,\omega)\geqslant$$

$$\sum_{\omega\in\Omega}(p\mu_s(\omega)+q\mu_{s'}(\omega))u(i,\omega),\forall i\in A$$

因为 $\mu_{s_{i^*}}=p\mu_s+q\mu_{s'}$ 是合理的后验概率分布，并令 $\tau\left(\mu_{s_{i^*}}\right)=\tau(\mu_s)+\tau(\mu_{s'})$，从而可以计算 s_{i^*} 对应的信号机制替换 s 和 s'。

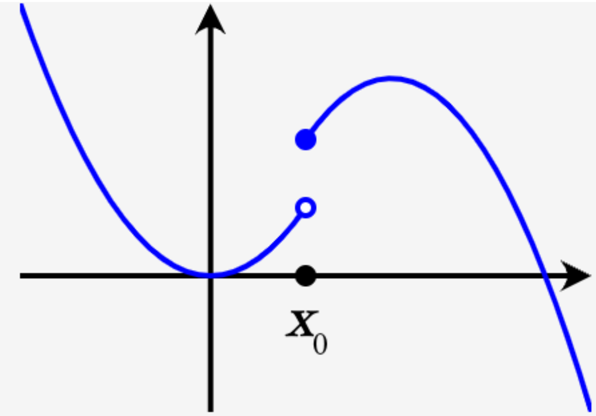
上半连续性

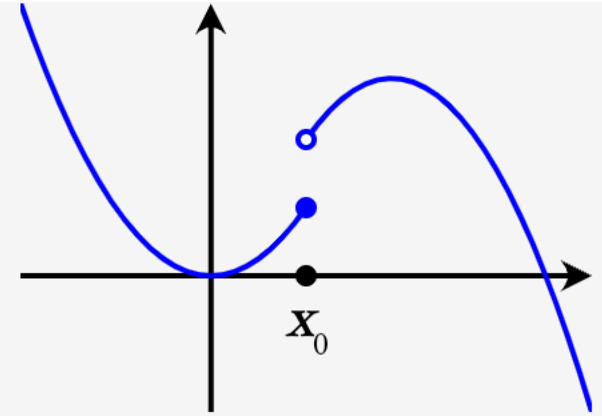
信息设计空间缩小后的问题是，最优信号机制问题是否一定有解？答案是肯定的，这需要引入一个概念：上半连续性。

定义

设 f 是一个实值函数，若对每个 $\varepsilon > 0$ 都存在 x_0 的开邻域 U 使得 $\forall x \in U, f(x) < f(x_0) + \varepsilon$，则称 f 在 x_0 上半连续。

将定义中大于换成小于即得到下半连续的定义，不难看出上半连续和下半连续是去除了连续函数定义的一半。下图左图给出了一个上半连续但不下半连续的函数，右图给出了一个下半连续但不上半连续的函数。





上半连续性

由于我们假设，当有多个最优行动时，接收者会选择最大化发送者效用的行动，因此最优信号机制问题的目标函数是一个上半连续的函数（发送方效用函数存在间断点，但在间断点上取值在更高的点）。

根据显示原理，τ 的支撑集的元素数目最多为 $|A|$，而每个 τ 中的元素是一个 Ω 上的概率分布，结合一个等式约束，因此现在问题的约束集合显然是一个紧集。下面的结论说明了最优信号机制问题的解的存在性：

定理

设 f 是拓扑空间 X 上的上半连续函数，K 是 X 上的紧集，则 f 在 K 上可以取到最大值。

凹包络

在理解了最优信号机制问题解的存在性后，最后的问题是如何求解最优信号机制。下面我们将介绍一个重要的概念：凹包络。

定义

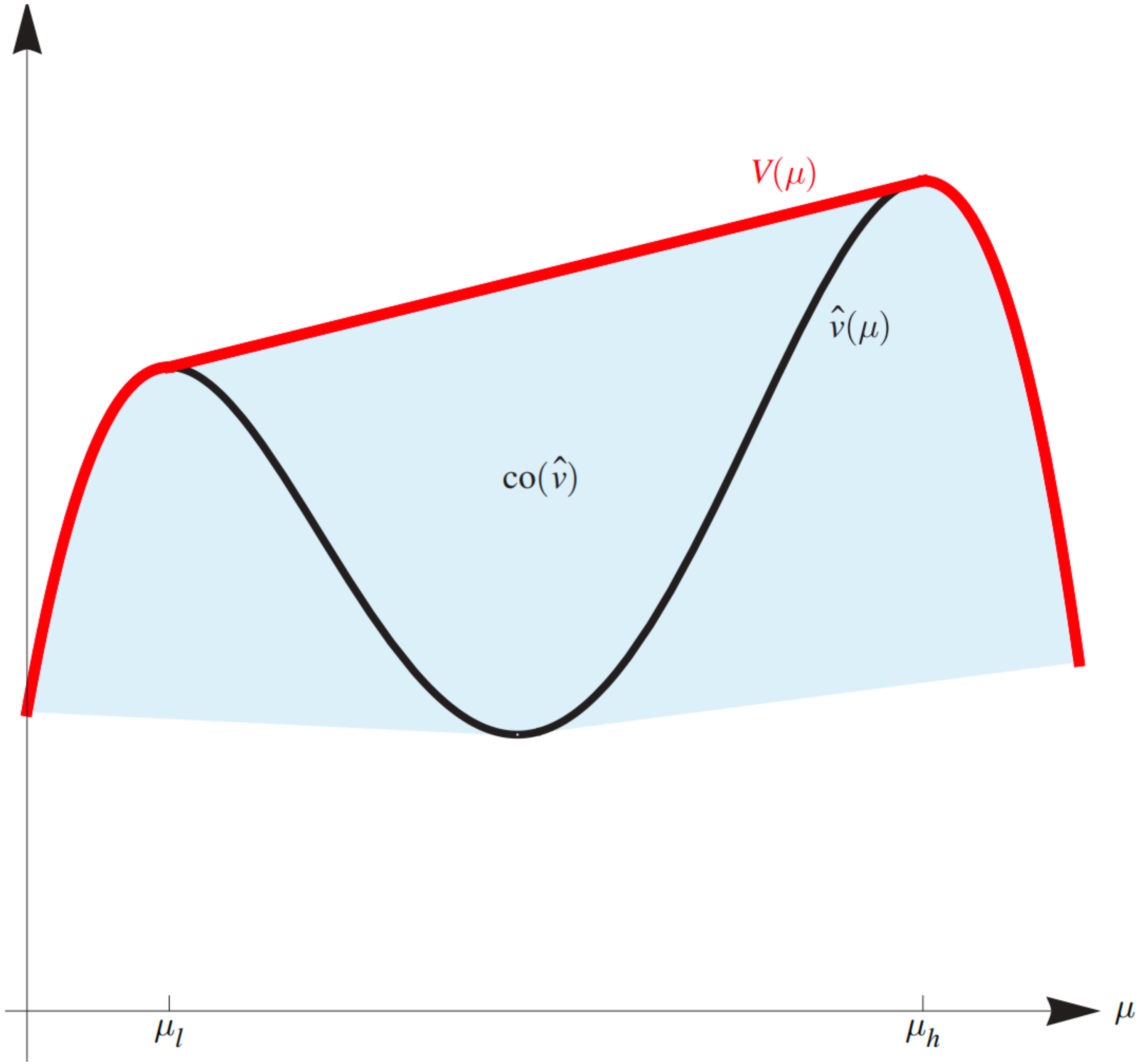
函数 $\hat{v}$ 的**凹包络（concave closure）** V 定义为：

$$V(\mu) = \sup\{z \mid (\mu, z) \in \mathrm{co}(\hat{v})\}$$

其中 $\mathrm{co}(\hat{v})$ 表示函数 $\hat{v}$ 的图像的凸包。

下页图给出了一个函数的凹包络的例子。直观而言，一个函数的凹包络就是大于等于这个函数的最小凹函数。

凹包络



最优信号机制的解

函数 $\hat{v}$ 的凹包络是求解最优信号机制问题的关键：

- 注意到如果 $(\mu', z) \in \mathrm{co}(\hat{v})$，则必然存在后验概率分布的分布 τ 使得 $\mathbb{E}_\tau \mu = \mu'$，且 $\mathbb{E}_\tau \hat{v}(\mu) = z$（因为期望也是凸组合）；
- $V(\mu')$ 则是所有这样的 z 中的最大值；

因此 $V(\mu_0)$ 就是最优信号机制问题的解。总而言之我们可以得到下面的推论，从而回答了之前提出的所有问题：

推论

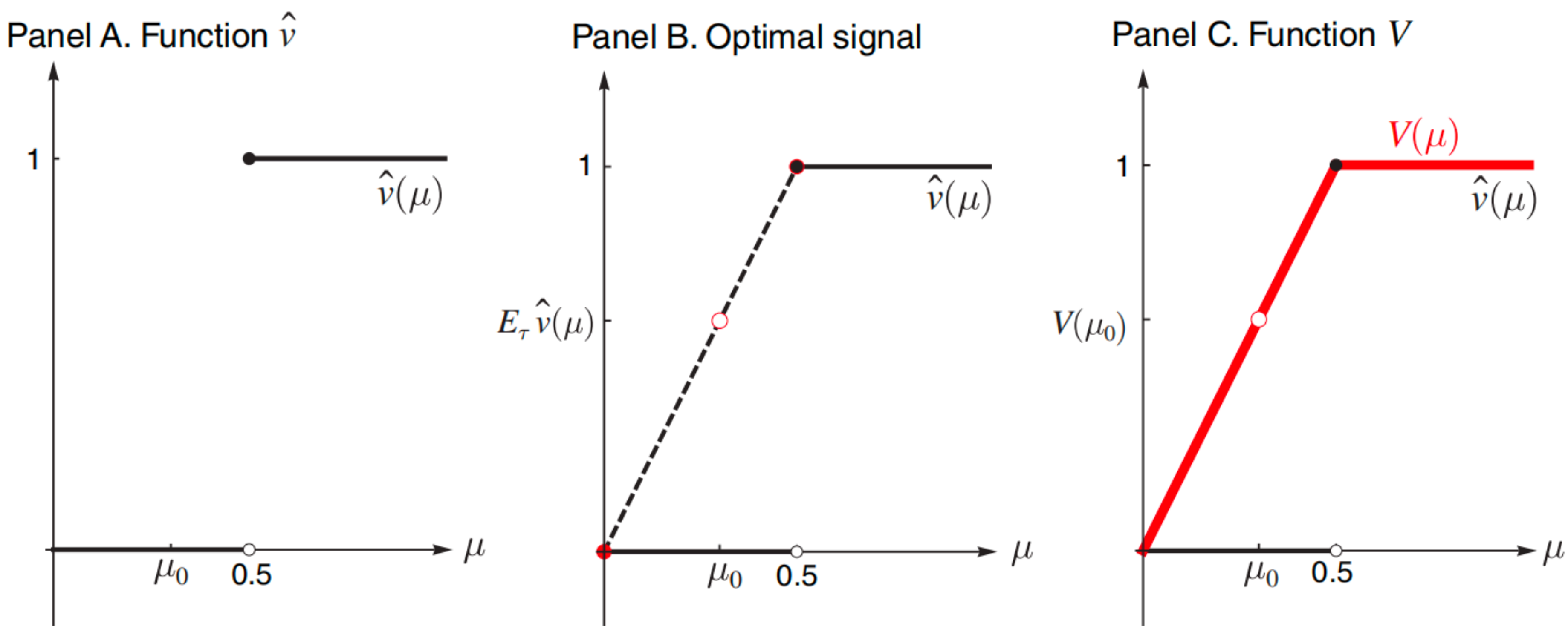
最优信号机制问题的解存在，最大值为 $V(\mu_0)$。进一步地，发送者设计信号能提升自己的效用当且仅当 $V(\mu_0) > \hat{v}(\mu_0)$。

例：检察官与法官的解

回到检察官与法官的例子，我们可以利用前面的结论证明之前给出的解是最优的。首先计算

$$\hat{v}(\mu) = \mathbb{E}_{\mu} v(\hat{a}(\mu), \omega) = v(\hat{a}(\mu))$$

又当 $\mu(\text{guilty}) \geqslant 0.5$ 时 $\hat{a}(\mu) = \text{convict}$，反之 $\hat{a}(\mu) = \text{acquit}$，因此可以作出下图，横坐标表示 $\mu(\text{guilty})$：



例：检察官与法官的解

不难看出 $V(\mu_0) = 0.6$，因此检察官的最优效用是 0.6，符合之前的计算。并且此时 τ 的支撑集是 $\{\mu_i, \mu_g\}$，其中

$$\mu_i(\text{innocent}) = 1, \mu_i(\text{guilty}) = 0$$

$$\mu_g(\text{innocent}) = 0.5, \mu_g(\text{guilty}) = 0.5$$

也是符合此前的结果的。进一步可以利用这一后验概率分布的分布计算出对应的信号机制，读者可以自行尝试，也是符合此前的结果的。

当然我们也可以考虑将检察官与法官的博弈放在 Cheap Talk 的视角下进行讨论，计算完美贝叶斯均衡。然而，使用上一讲的技术不难发现不存在和最优信号机制类似的半分离的完美贝叶斯均衡，因为半分离均衡中发送者发送不同信息的收益必须相同（无差异条件），但这样的约束使得不存在半分离均衡。

可以从计算的视角重新看待贝叶斯劝说问题。已知先验分布 μ_0 以及发送者的效用函数 $v(a,\omega)$ 和接收者的效用函数 $u(a,\omega)$。设接收者行动有 n 种，记为 $\{1,2,...,n\}$，则根据显示原理，只需对每个 ω 设计 n 种信号 $\{\pi(s_i \mid \omega)\}_{i=1}^{n}$ 即可，故可以将最优信号机制问题转化为如下线性规划问题：

$$
\begin{aligned}
\max_{\pi(s_i \mid \omega)} \quad & \sum_{\omega\in\Omega}\sum_{i=1}^{n}\pi(s_i \mid \omega)\mu_0(\omega)v(i,\omega) \\
\text{s.t.} \quad & \sum_{\omega\in\Omega}\pi(s_i \mid \omega)\mu_0(\omega)u(i,\omega) \geqslant \pi(s_i \mid \omega)\mu_0(\omega)u(j,\omega), \forall i,j\in[n] \\
& \sum_{i=1}^{n}\pi(s_i \mid \omega) = 1, \forall \omega\in\Omega \\
& \pi(s_i \mid \omega) \geqslant 0, \forall s_i \in S, \omega\in\Omega
\end{aligned}
$$

相信在学习了相关均衡的线性规划表达式后会对上述表达式有更深的理解，事实上约束条件就是贝叶斯相关均衡的条件。