Optimal Information Disclosure

A Linear Programming Approach

Anton Kolotilin

June 14, 2019

Introduction

"Bayesian Persusasion" (Kamenica-Gentzkow, 2011):

- One sender, one receiver (Prosecutor, Judge)
- Sender wants to persuade the receiver to take some action. (Acquit, Convict)
- There is some unknown state of the world, which affects sender's and receiver's utilities.
  - Defendent is either innocent or guilty. Prosecutor wants to maximize probability of guilty verdict, judge wants to make the right decision.
- Sender chooses a "Blackwell experiment", which probabilistically generates a realization depending on the state:
- In other words, sender is choosing a conditional probability distribution $\pi(\gamma|\omega)$, a distribution over signal realizations γ depending on state ω.
- What will be my evidence-gathering procedure? Who will I subpoena? Who will I put on the witness stand?

Introduction

- Receiver observes (i) the experiment and (ii) the realization. Receiver is a rational Bayesian, uses Bayes' rule to infer the state of the world, and then takes an action.
  - Sender has full commitment power.
- **Question:** Can the sender persuade the receiver to change her action? Is there a way for the sender to choose a transparent experiment to maximize utility, even when the receiver is aware of this?
  - In other words, can I do better than $\pi(g|guilty) = 1$ and $\pi(i|innocent) = 1$? Or a completely non-informative signal?
- **Answer:** Yes, depending on the prior distribution, the sender and receiver's utility functions, etc.

Kolotilin (2018)

Similar setting, except now the receiver has private information about their type.

- Utilty functions:
  - Sender: $v(r, s)$ if $a = 1$, otherwise 0. Continuous in r and c.
  - Receiver: $u(r, s)$ if $a = 1$, otherwise 0. Continuously differentiable in r, continuous in c
- Sender's type: $[\underline{s}, \bar{s}]$
- Receiver's type: $[\underline{r}, \bar{r}]$
- Prior distribution of the sender's type: $F(s)$
- Conditional distribution of the types: $G(s|r)$
  - Receiver can glean some information of the sender's type depending on their own type.
  - Also introduces another layer of uncertainty for the sender.
- Sender chooses a mechanism $\mu(m, s)$, a joint distribution between messages $m \in \mathbb{R}$ and sender's type s.
  - Joint distribution instead of conditional distribution.
- Receiver takes some action based on message and their type: $a = \{0, 1\}$

Timing

1. Sender publicly chooses mechanism $\mu(m, s)$.
2. (m, s, r) are drawn according to μ and G.
3. The receiver observes (m, r) and takes some action a.

Toy example

Imagine a university is deciding its grading policy, with the intent of maximizing employment of its students.

- Student type: $\{\text{Good}, \text{Bad}\}$.
  - Some prior distribution $F(\{\text{Good}\})$ and $F(\{\text{Bad}\})$
- Employer conducts some interview, draws their "type" (outcome of interview): $\{\text{Passed}, \text{Failed}\}$
  - The outcome of the interview depended on the student's type: $G(\text{Passed}|\text{Good})$
  - In this case, think of this as the employer's private information.
- University chooses a grading mechanism: $\mu : \{\text{A, B, C}\} \times \{\text{Good}, \text{Bad}\} \rightarrow \mathbb{R}$
  - A joint distribution between student types and grades
  - If I give both the good and bad types an A, this is grade inflation.
  - If I give only the good types an A and only the bad types a C, this is a "full revelation mechanism".
- Employer takes some action: $\{\text{Hire}, \text{Don't Hire}\}$

Building the Problem

- The receiver observes (m, r) and generates a posterior of the sender's type:

$$P(s|m,r) = \frac{P_G(r|s)P_\mu(s|m)}{P_{G,\mu}(r|m)} = \frac{\int_{\underline{s}}^{s} g(r|\hat{s})dP_\mu(\hat{s}|m)}{\int_{\underline{s}}^{\bar{s}} g(r|\hat{s})dP_\mu(\hat{s}|m)}$$

- The receiver's expected utility under this posterior is:

$$E_{s|m,r}[u(r,s)] = \int_{\underline{s}}^{\bar{s}} u(r,s)dP(s|m,r) = \frac{\int_{\underline{s}}^{\bar{s}} u(r,s)g(r|s)dP(s|m)}{\int_{\underline{s}}^{\bar{s}} g(r|s)dP_\mu(s|m)}$$

- If $\int_{\underline{s}}^{\bar{s}} \tilde{u}(r,s)dP(s|m) = \int_{\underline{s}}^{\bar{s}} u(r,s)g(r|s)dP(s|m) > 0$, the receiver chooses $a = 1$

Assumption

Definition

(Single-Crossing Property): For every distribution $P(s|m)$ induced by m, there exists a cutoff $r_m \in R$ such that

$$E_{s|m,r}[u(r,s)] = \int_{\underline{s}}^{\bar{s}} \tilde{u}(r,s)dP(s|m) \geq 0 \text{ for all } r \geq r_m.$$

Moreover, there exists a strictly decreasing function $r^* : S \to R$ such that $u(r^*(s), s) = 0$.

- WLOG only consider mechanisms such that each message m induces a cutoff m.
- The feasible set of messages is defined to be: $M^* := r^*(S)$.
  - In other words, the messages that make the receiver indifferent between action and inaction (when the receiver is indifferent, they act).

Assumption

I consider this to be a statement of the structure of the receiver's private information and the mechanism:

- Receiver's behavior completely characterized by their action cutoff.
- By the revelation principle, it is without loss for the sender to make cutoff recommendation and impose an incentive compatibility constraint on their recommendation.
  - "Cutoff recommendation": The mechanism produces a cutoff depending on the state ("if your type is above this m you should act")
  - "Incentive compatible": The receiver cannot be made worse off if they follow the sender's advice.
- Furthermore, for higher state s, the cutoff $r^*(s)$ will be lower ("the bar is lower")

We can now describe the sender's expected utility:

$$V(m, s) := E_{r|s,m}[v(r, s)] = \int_m^{\bar{r}} v(r, s) g(r|s) dr$$

Alternative Model

It is possible to use the same framework but interpret $R = [\underline{r}, \bar{r}]$ as the receiver's action.

- The single crossing assumption makes the receiver's utility single-peaked in her action.
- A message m induces the receiver to take an action r_m.

Primal

$$
\begin{aligned}
\sup_{\mu \in \mathcal{P}(M^* \times S)} & \int_{M^* \times S} V(m,s) d\mu \\
\text{s.t.} & \int_{M^* \times \hat{S}} d\mu = \int_{\hat{S}} dF, \ \forall \hat{S} \in \Sigma_S \\
& \int_{\hat{M} \times S} \tilde{u}(m,s) d\mu = 0, \ \forall \hat{M} \in \Sigma_{M^*}
\end{aligned}
$$

- Constraint 1: Bayes-plausibility
- Constraint 2: Incentive-compatibility
- (Slight abuse of notation here: Since $M^* = r^*(S) \subset R$, it is valid to define $\tilde{u}(r,s)$ over M^*).

Another way to write the primal

$$\sup_{\mu \in \mathcal{P}(M^* \times S)} \langle V, \mu \rangle$$

$$\text{s.t. } A(\mu) = \begin{pmatrix} F \\ 0 \end{pmatrix}$$

$$A(\mu) := \begin{pmatrix} P_S \mu \\ P_{M^*} \mu_u \end{pmatrix}$$

$$\mu_u(M \times S) := \int_{M \times S} \tilde{u}(m, s) d\mu$$

Monge-Kantorovich

Has a similar flavor to a class of LP problems in "Optimal Transport"

- Suppose we want to move "sand" from a histogram/density in one space to another, and minimize the "effort" involved.
  - We start with the sand in space U distributed according to ν_1 and we want to move the sand to space V so it is distributed according to ν_2, while minimizing the cost $c(u, v)$.
- We can think of this as finding a joint density on $U \times V$, where the marginal in U is ν_1 and the marginal in V is ν_2 that solves the following problem:

$$
\begin{aligned}
\inf_{\rho \in \mathcal{P}(U \times V)} \int_{U \times V} c(u, v) d\rho \\
\text{s.t. } \; P_U \rho &= \nu_1 \\
P_V \rho &= \nu_2 \\
\rho &\geq 0
\end{aligned}
$$

Existence of Primal Solution

1. $\mathcal{P}(M^* \times S)$ is a compact set in the weak topology $\sigma(ca(M^* \times S), C(M^* \times S))$
2. $P_S\mu = \int_{M^* \times S}(1 \circ \varphi_s)d\mu$ where $\varphi_s : (m, s) \to s$, is continuous since the projection map is continuous.
3. $P_{M^*}\mu_u = \int_{M^* \times S}(\tilde{u}(m, s) \circ \varphi_{m^*})d\mu$ is continuous since $u(r, s)g(r|s)$ and the projection map are continuous.
4. But is it non-empty?

Existence of Primal Solution

Definition

(Full Revelation Mechanism) For any measurable set $\hat{M} \in \Sigma_{M^*}$ and $\hat{S} \in \Sigma_S$, the full revelation mechanism is the following:

$$\mu(\hat{M}, \hat{S}) = F(r^{*-1}(\hat{M}) \cap \hat{S})$$

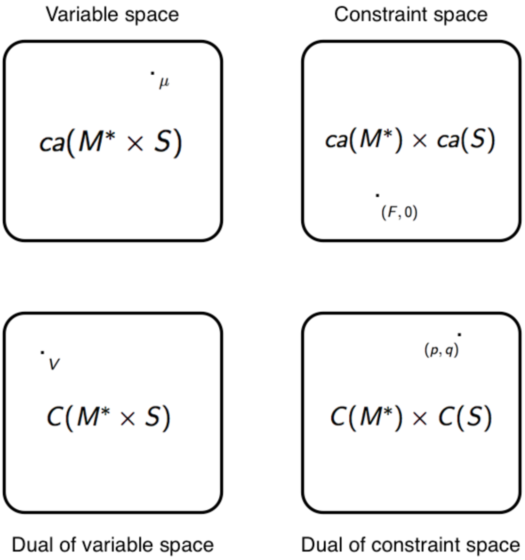
r^* is strictly decreasing, which means that the inverse will exist and so we generate a unique message for each state.

- Constraint 1: $\mu(M^* \times \hat{S}) = F(S \cap \hat{S}) = F(\hat{S})$
- Constraint 2: $\int_{\hat{M} \times S} \tilde{u}(m, s) d\mu = 0$ because the support is $r^*(S)$

Adjoint

- $A : ca(M^* \times S) \to ca(M^*) \times ca(S)$
- $A^* : C(M^*) \times C(S) \to C(M^* \times S)$

Space


Variable space
Constraint space
μ
$ca(M^* \times S)$
$ca(M^*) \times ca(S)$
$(F, 0)$
V
(p, q)
$C(M^* \times S)$
$C(M^*) \times C(S)$
Dual of variable space
Dual of constraint space

Adjoint

$$
\begin{aligned}
\langle A(\mu), (p, q)\rangle &= \langle (P_S(\mu), P_{M^*}(\mu_u)), (p, q)\rangle \\
&= \int_S p(s) dP_S(\mu) + \int_{M^*} q(m) dP_{M^*}(\mu_u) \\
&= \int_S \int_{M^*} p(s) d(P_{M^*}\mu) d(P_S\mu) \\
&+ \int_S \int_{M^*} q(m) d(P_{M^*}\mu_u) d(P_S\mu_u) \\
&= \int_{M^* \times S} (p(s) + q(m)\tilde{u}(m, s))\, d\mu \\
&= \langle p(s) + q(m)\tilde{u}(m, s), \mu\rangle \\
&= \langle \mu, A^*(p, q)\rangle \\
A^*(p, q) &= p(s) + q(m)\tilde{u}(m, s)
\end{aligned}
$$

Dual

$$\begin{aligned} \inf_{p,q \in C(S)} & \int_S p(s) dF \\ \text{s.t. } & p(s) + q(m)\tilde{u}(m,s) \geq V(m,s) \\ & \forall m \in M^*, s \in S \end{aligned}$$

Or alternatively...

$$\begin{aligned} \inf_{p \in C(S)} & \langle p, F \rangle \\ \text{s.t. } & A^*(p,q) \geq V \end{aligned}$$

Interpretation of the Dual

- As usual, the dual variables can have a price interpretation.
- $p(s)$ can be interpreted as the sender's shadow price of the Bayes-plausibility constraint (which is a function of the state)
  - "How much would I pay if I could break Bayes' Rule?"
- $q(m)$ can be interpreted as the sender's shadow price of the incentive-compatibility constraint (which is a function of cutoff types)
  - "How much would I pay if I could somehow make the receiver worse off?"
- Hicksian problem: We can think of these constraints as budget constraints. To obtain at least a certain expected utility $V(m, s)$, how can I minimize my "costs"?

Existence of Dual Solution

- First, relax to $L_\infty(S) \times L_\infty(M^*)$, endowed with weak topology $\sigma(L_\infty \times L_\infty, L_1 \times L_1)$ (assume F has an L_1 density f).
- Since this program wants to minimize the objective function (and F is positive), we can bound (p, q), adding an additional constraint, without affecting the value of the dual problem. V is continuous on a compact set so it attains a maximum $\bar{V}$.
- Feasible set:

$$-\bar{V} \leq V(m,s) \leq p(s) + q(m)\tilde{u}(m,s) \leq \bar{V}$$

$$\int_{M^*}\int_S p(s)dsdm + \int_S\int_{M^*} q(m)\tilde{u}(m,s)dmds \leq \int_{M^*}\int_S \bar{V}dsdm$$

$$\left|\left\langle (p,q), (\int_{M^*} dm, \int_S \tilde{u}(m,s)ds) \right\rangle\right| \leq \int_{M^*}\int_S \bar{V}dsdm < \infty$$

- By Banach-Alaoglu, this is a closed subset of a compact set.

Existence of Dual Solution

Now we can construct an optimal solution that is continuous.

- If (p, q) is optimal $(L_\infty(S) \times L_\infty(M^*))$, then (p^*, q) is optimal, where

$$p^*(s) = \sup_{m \in M^*} \{V(m, s) - q(s)\tilde{u}(m, s)\}$$

- $M^* \times S$ is compact so $V(m, s)$ and $u(m, s)g(m|s)$ are uniformly continuous. q is also bounded because $q \in L_\infty$.

Existence of Dual Solution

For any $\epsilon > 0$, there exists a $\delta > 0$ such that $\forall |s - s^{'}| < \delta$ and $\forall m \in M^*$:

$$\left| (V(r,s) - q(s)\tilde{u}(m,s)) - \left(V(r,s^{'}) - q(s^{'})\tilde{u}(m,s^{'}) \right) \right| < \epsilon$$

- Because of the sup definition of p^*, for each s we can always find an m such that

$$\begin{aligned} p^*(s) &< V(m,s) - q(s)\tilde{u}(m,s) + \epsilon \\ p^*(s') &\geq V(m,s^{'}) - q(s^{'})\tilde{u}(m,s^{'}) \\ \implies &> V(m,s) - q(s)\tilde{u}(m,s) - \epsilon \\ &> p^*(s) - 2\epsilon \end{aligned}$$

- Repeat the same procedure for q. (p^*, q^*) is an optimal solution to the problem in $C(S) \times C(M^*)$

Strong Duality

Theorem

(Anderson and Nash) If there exists a point with $A^*(p,q) - V$ in the interior of the dual cone and the dual has a finite value, then there is no duality gap.

- Finite: The dual is bounded above by $\int_S \bar{V} ds$. By weak duality, the dual is bounded below by primal solution (which itself must be bounded below b/c V is continuous).

$$\sup_{\mu \in \mathcal{P}(M^* \times S)} \int_{M^* \times S} V(m,s) d\mu \leq \inf_{p \in C(S)} \int_S p(s) dF$$

- Interior point: $p(s) = 2\bar{V}$ and $q(m) = 0$.

$$\begin{aligned} &A^*(p,q) - V \\ &= 2\bar{V} - V(m,s) > 0 \\ &\forall m \in M^*, s \in S \end{aligned}$$

Complementary Slackness

Lemma

If μ is feasible for the primal and (p, q) is feasible for the dual and

$$\int_{M^* \times S} p(s) + \tilde{u}(m, s)q(m) - V(m, s)d\mu = 0$$

then μ and (p, q) are optimal solutions, and the values of the primal and dual are the same.

Proof

- Follows from strong duality:

$$\int_{M^* \times S} p(s) d\mu = \int_{M^* \times S} V(m,s) d\mu$$

$$\int_{M^* \times S} p(s) - V(m,s) d\mu = 0$$

$$\int_{M^* \times S} p(s) + \tilde{u}(m,s)q(m) - V(m,s) d\mu = 0$$

- $\int_{M^* \times S} \tilde{u}(m,s)q(m)d\mu = 0$ can be derived from the fact that by feasibility of μ (incentive-compatibility constraint), $\int_{M^* \times S} \tilde{u}(m,s)d\mu$ maps to the 0 measure and so $\langle q, 0 \rangle = 0$.

Application

Corollary

The full revelation mechanism is optimal if and only if, for all $s_1, s_2 \in S$ *and* $r \in R$ *such that* $s_2 > s_1$ *and* $r \in (r^*(s_2), r^*(s_1))$*:*

$$\frac{V(r^*(s_2), s_2) - V(r, s_2)}{\tilde{u}(r, s_2)} \geq \frac{V(r^*(s_1), s_1) - V(r, s_1)}{\tilde{u}(r, s_1)}$$

Proof.

By the complementary slackness condition:

$$\int_{M^* \times S} p(s) + \tilde{u}(m,s)q(m) - V(m,s)d\mu_{full} = 0$$

The full revelation mechanism has support only on $(r^*(s), s)$. Furthermore, the integrand is positive, so this is equivalent to:

$$p(s) + \tilde{u}(r^*(s), s)q(r^*(s)) = V(r^*(s), s) \text{ almost everywhere}$$

$\tilde{u}(r^*(s), s) = 0$ by definition, which means $p(s) = V(r^*(s), s)$ (for all s because p is cts).

$$V(r^*(s), s) + \tilde{u}(m,s)q(m) \geq V(m,s)$$

$$\frac{V(r, s_2) - V(r^*(s_2), s_2)}{\tilde{u}(r, s_2)} \geq q(m) \geq \frac{V(r^*(s_1), s_1) - V(r, s_1)}{-\tilde{u}(r, s_1)}$$

□

Application

A more general result than Kamenica-Gentzkow (2011):

- "The full revelation mechanism is optimal if $\hat{V}(Q)$ is convex in Q so that the sender prefers to seperate Q_1 and Q_2 than to pool them at $\alpha Q_1 + (1 - \alpha) Q_2$"
- Kolotilin's criterion allows one to (1) check only classes of discrete distributions (mass on s_1 and s_2) and (2) check only certain types of deviations.