Game Theory
Dynamic Information

Olivier Gossner, CNRS – Ecole Polytechnique

Master in Economics

Plan

1. Repeated zero-sum games, optimal information revelation
2. Optimal information design
3. Direct communication
4. Learning
5. Reputations

The set-up

Zero-sum repeated games and incomplete information on one side (Aumann-Maschler 68)
$k \in \{k_1, k_2\}$, equally likely. P1, the maximizer, is informed of the k. P2, the minimizer, isn't.
A game takes place in stages between P1 and P2. At the end of each stage, P2 observes P1's action in the previous stage (Not her own payoff, which is kept in a secret bank account). We study optimal information revelation by P1.

The payoff functions to P1 in states k_1 and k_2 are given by:

	L	R
T	1	0
B	0	0

k_1

	L	R
T	0	0
B	0	1

k_2

Analysis 1

1. Assume the game with incomplete information is played only once. What are the strategies for P1? For P2?
2. What is the optimal strategy for P1 in the one-shot Bayesian game? What is the corresponding $\min\max$ payoff?
3. Assume now that P1 is using a completely revealing strategy: a strategy that plays T in state k_1, and B in state k_2. After observing action T, what is the posterior of P2 on the state of nature? Same question after observing B.
4. If P1 uses a completely revealing strategy and P2 plays a best response to it, what is the best payoff that P1 can expect in subsequent repetitions of the game? (in stage 2, stage 3, and so on).
5. Now consider non-revealing strategies of P1, which uses the same mixed strategy in states k_1 and k_2. What is the best non-revealing strategy for P1? What is P1's expected payoff in stage 1 if he uses this non-revealing strategy and P2 plays a best response? What about stage 2 onwards (P1 uses the same non-revealing strategy at every stage)?
6. Is it better for P1 to use a completely revealing strategy, or a non-revealing one?

Set-up and analysis 2

Now the payoffs are given by:

k_1:

	L	R
T	-1	0
B	0	0

k_2:

	L	R
T	0	0
B	0	-1

Same questions as before.

Set-up and analysis 3

Finally consider the payoff functions to P1 in states k_1 and k_2 given by:

	L	M	R
T	4	0	2
B	4	0	-2

k_1

	L	M	R
T	0	4	-2
B	0	4	2

k_2

- Answer the same questions as before.
- Consider the following partially revealing strategies of P1: If the state is k_1, with probability 3/4 play T forever, and with probability 1/4, play B forever. If the state is k_2, with probability 3/4 play B forever, and with probability 1/4, play T forever. Conditional on observing T in the first stage, what is P2's posterior belief on the states of nature? What is P2's best response to P1's strategy from stage 2 on? Same questions when P2 observes P1 playing B in the first stage.
- What expected average payoff per stage does the partially revealing strategy guarantees to P1 in the long-run?
- What is the best strategy for P1, a non revealing strategy, a completely revealing strategy, or a partially revealing strategy?

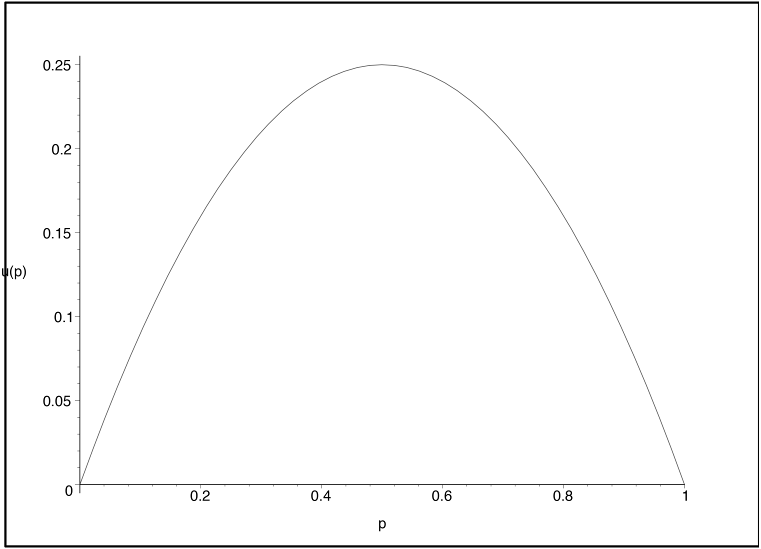
General analysis

For each of the payoff structures:

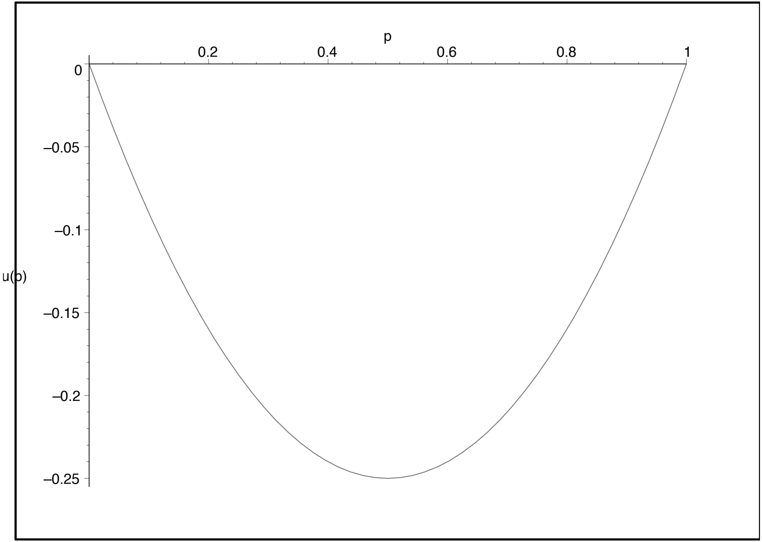
- Assume that the state of nature is k_1 with probability p and k_2 with probability $1-p$, and that P1 plays a non-revealing strategy.
- Graph the value $u(p)$ of the one-shot Bayesian game as a function of p. Draw the function $(\text{cav}\, u)(p)$, the smallest concave function that is larger or equal than u.

Comment on the optimal revelation of informaiton

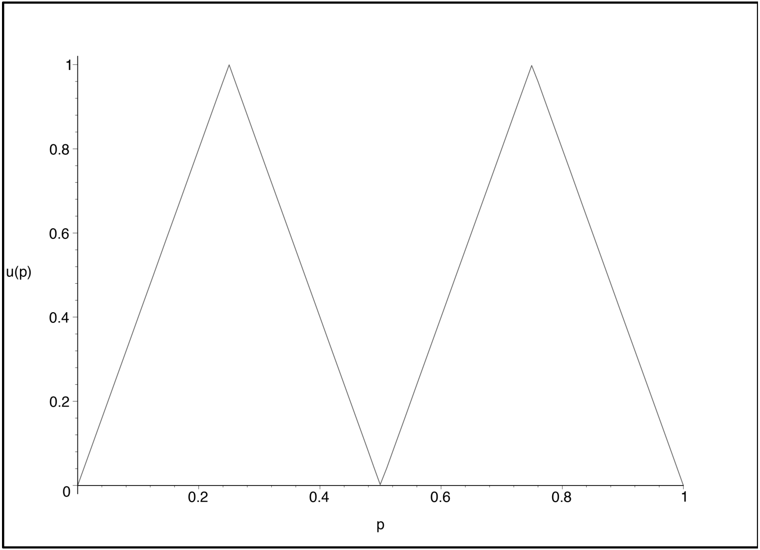
First case



Second case



Third case



Plan

1. Repeated zero-sum games, optimal information revelation
2. Optimal information design
3. Direct communication
4. Learning
5. Reputations

The set-up

Kamenica Gentzkow (2011)

Two players, a sender and a decision maker.

d.m. has belief $p \in \Delta(K)$.

Sender chooses the d.m.'s information structure (full commitment).

After receiving information, receiver chooses $a \in A$.

Payoffs are $g^s(a, k)$, $g^d(a, k)$.

Example

The d.m. is a regulator and the sender is a pharmaceutical company.

The d.m. has initial belief $p = 0.5$ that a certain medicine is efficient.

Approves the medicine iff. the probability that it is efficient > 0.9.

Sender's goal is to maximise the probability that the medicine is approved.

Analysis

Solving by backwards induction.
Sender chooses a distribution of posterior beliefs π on $\Delta(\Delta(K))$ s.t.

$$E_\pi q = p$$

For each posterior q, the d.m. chooses

$$a \in A^*(q) = \arg\max_A E_q g^d(a, k)$$

Let

$$v^s(q) = \max_{a \in A^*(q)} E_q g^s(a, k)$$

We can show that the maximal payoff to the sender is

$$(cav\ v^s)(p)$$

Plan

1. Repeated zero-sum games, optimal information revelation
2. Optimal information design
3. Direct communication
4. Learning
5. Reputations

Experts and influence: Nash

Crawford and Sobel 82.

- Set of types of the expert $T = [0, 1]$.
- Set of messages $M = [0, 1]$
- Set of actions of the decision maker $A = [0, 1]$.
- Expert's utility $u_e(a, t) = -(a - (t - b))^2$, $b > 0$.
- Dm's utility $u_d(a, t) = -(a - t)^2$

Solution concept is Nash Equilibrium (no commitment!)

Applications:

- Relationship between a doctor and his patient
- Choice of expenditure on a public project
- Choice of departure time for two friends to take a plane (one having private information about flight time)
- Hierarchical relationships in organizations (e.g., choice = effort level)

Equilibria with 2 messages

Assume $m = 1$ for $t > x$, $m = 0$ for $t \leq x$.
Let σ_d be the dm's strategy.Then,

$$\begin{cases} \sigma_d(0) = x/2 \\ \sigma_d(1) = (x+1)/2 \end{cases}$$

The expert chooses the message inducing the action the closest to $t + b$.

$$\begin{cases} \sigma_e(x) = 0 & \text{if } t + b < x/2 + 1/4 \\ \sigma_e(x) = 1 & \text{if } t + b > x/2 + 1/4 \end{cases}$$

From which we get

$$x = 1/2 - 2b$$

- We must have $b \leq 1/4$
- $x < 1/2$: The second interval is larger than the first

Equilibria with n messages

We look for an equilibrium of the form

$$\sigma_e(x) = m_k \text{ if } t \in [x_{k-1}, x_k]$$

$0 = x_0 < x_1 < x_2 \ldots < x_n = 1$.
x_k indifferent between m_k and m_{k+1}: $x_k + b$ in the middle between $(x_k + x_{k-1})/2$ and $(x_{k+1} + x_k)/2$, so

$$(x_{k+1} - x_k) = (x_k - x_{k-1}) + 4b$$

and

$$x_k = kx_1 + \frac{k(k-1)}{2} 4b$$

$x_n = 1$ gives

$$x_1 = 1/n - 2(n-1)b$$

Such an equilibrium exists if $b < \frac{1}{2n(n-1)}$.

Plan

The basic learning question

- Z finite space of agent's observations
- P law of a process $(z_t)_t$ with values in Z
- Q agent's belief on $(z_t)_t$

Next stage predictions following $z_1 \ldots z_{t-1}$

- $p_t = P(z_t | z_1 \ldots z_{t-1})$ stage t's law conditional on the past
- $q_t = Q(z_t | z_1 \ldots z_{t-1})$ agent's prediction at stage t
- p_t, q_t are r.vs. in $\Delta(Z)$

Does the agent eventually make accurate predictions?

When and in what sense does q_t "converge" to p_t?

Example: P iid. coin tosses, $p \in [0, 1]$. Q puts uniform probability on p

Entropy of a distribution

- X finite set, $p \in \Delta(X)$
- Amount of "surprise" in seeing a realization x

$$\log \frac{1}{p(x)}$$

- Expected amount of surprise, entropy of p

$$H(p) = \sum_x p(x) \log \frac{1}{p(x)}$$

$\log$ is $\log_2$ by convention, $0 \log(\infty) = 0$ by continuity

- $H(p)$ measures the "randomness" of a r.v. with distribution p, or equivalently the amount of information contained in its observation

Relative entropy

- $p, q \in \Delta(X)$: p real distribution, q agent's belief

- Expected amount of surprise of the agent with belief q

$$\sum_x p(x) \log \frac{1}{q(x)}$$

$$\sum_x p(x) \log \frac{1}{q(x)} \geq H(p)$$

- The relative entropy is the difference

$$d(p \| q) \quad = \quad \sum_x p(x) \log \frac{p(x)}{q(x)}$$

- It is an information theoretic measure of the agent's prediction error

Fundamental property: Chain Rule

- (x, y) drawn in $X \times Y$ with law P
- Agent's belief on (x, y) is Q

Relative entropy at once

The error in predicting (x, y) is $d(P \| Q)$

Relative entropy in 2 stages: Assume x is observed, then y

- P_X, Q_X marginals of P, Q on X
- The total expected error in predicting x, then y, is

$$d(P_X \| Q_X) + \mathbf{E}_{P_X} d(P(\cdot|x) \| Q(\cdot|x))$$

Chain Rule

$$d(P \| Q) = d(P_X \| Q_X) + \mathbf{E}_{P_X} d(P(\cdot|x) \| Q(\cdot|x))$$

Consequences of the Chain Rule

Relative entropy under grain of truth, $Q = \mu P + (1-\mu)P'$

$$d(P \| Q) \leq -\log \mu$$

Total expected prediction error under grain of truth

Let $Q = \mu P + (1-\mu)P'$ on $Z^{\mathbb{N}}$, for every $T \geq 1$

$$\sum_{t=1}^{T} \mathbf{E}_P d(p_t \| q_t) \leq -\log \mu$$

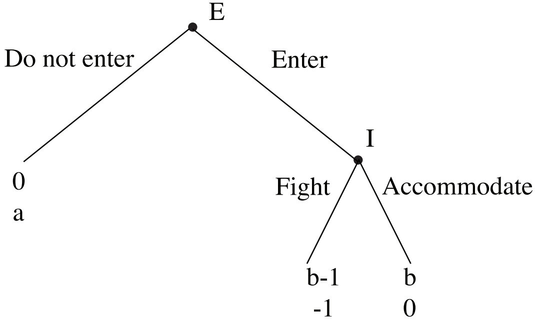
Expected δ-discounted prediction error

$$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_P d(p_t \| q_t) \leq -(1-\delta)\log \mu$$

Plan

The chain-store game

Consider the following entry game, with $a > 1$, $0 < b < 1$:



- What are the NE? The SPNE?
- What happens if a long-run incumbent sequentially faces two short-run entrants?
- What happens with a sequence of 100 entrants?

Possibility of a tough incumbent

With (small) probability α, the long-run player is "tough" and a tough player always fights. The game is thus a game of incomplete information.

For $\alpha > b$ no entrant wishes to enter. Now consider $\alpha < b$.

1. There are no SE in which I always accommodates the first entrant.
2. There are no SE in which I always fights the first entrant.
3. Following "Fight", the second entrant is indifferent between E and N, so his belief that the incumbent is tough is b.
4. The probability of "Fight" f in the first stage satisfies $\alpha = b(\alpha + f)$.
5. The first entrant enters if $\alpha < b^2$, does not enter if $\alpha > b^2$.

Generalization to k entrants, the first entrant does not enter if $\alpha > b^k$.

Lessons from the Chain Store game

The introduction of “crazy types” in the Chain Store game shows that

- A “small” perturbation of a repeated game can lead to equilibrium outcomes that are drastically different from the original game. Questions the robustness of conclusions wrt. modeling assumptions.
- Reputations can be captured by the introduction of “crazy” or “commitment” types.

We study reputation games in which:

- An infinitely lived long-run player faces a series of short-run players
- The long-run player may either be of a normal type, with known stage payoff function and discount factor δ, or of a commitment type who repeatedly plays a given commitment strategy

Reputations (Fudenberg Levine 1989, 1992)

Model

- Long-run P1 facing short-run players 2
- Action spaces A_1, A_2, payoff functions $g_i : A_1 \times A_2 \to \mathbb{R}$
- Long-run of behavioral type with probability μ, or normal type
- Commitment type repeats $\hat{s}_1 \in \Delta(A_1)$
- Normal type uses discount factor δ
- Each P2 knows the past actions of P1

Questions

- Can the long-run player build a "reputation" for playing $\hat{s}_1$?
- Bounds on NE payoffs to P1: asymptotic? explicit?

Example: Quality choice game

A long-run (cook, firm) may exert effort and produce a produce a high quality good, or produce a low quality good at no cost. Short-run consumers may decide to buy the product, or not.

	b	n
H	$1, 1$	$-2, 0$
L	$3, -1$	$0, 0$

Relating errors and payoffs

- P1 plays $\hat{s}_1$
- P2 plays s_2, BR to his belief s_1
- P2's prediction error on P1's action is

$$d(\hat{s}_1 \| s_1)$$

Min payoff to P1 from $\hat{s}_1$ if P2 makes an error of at most ε

$$v_{\hat{s}_1}(\varepsilon) = \inf g_1(\hat{s}_1, s_2)$$

s_2 BR to some s_1, $d(\hat{s}_1 \| s_1) \leq \varepsilon$

If P1 plays $\hat{s}_1$, P2 BR to his beliefs

Let $p_t = \hat{s}_1$, $q_t = s_1$, $g_{1,t} = g_1(\hat{s}_1, s_{2,t})$

1. $$g_{1,t} \geq v_{\hat{s}_1}(d(p_t \| q_t))$$

2. $$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_P d(p_t \| q_t) \leq -(1-\delta)\log\mu$$

Let $w_{\hat{s}_1}$ be the largest convex mapping below $v_{\hat{s}_1}$:

$$(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} \mathbf{E}_P g_{1,t} \geq w_{\hat{s}_1}(-(1-\delta)\log\mu)$$

Theorem

The worst Nash Equilibrium payoff to P1 is at least

$$w_{\hat{s}_1}(-(1-\delta)\log\mu)$$

Consequences

Assume that there are several behavioral types, $\hat{s}_1$ with probability $\mu(\hat{s}_1)$

Theorem

The worst Nash Equilibrium payoff to P1 is at least

$$\sup_{\hat{s}_1} w_{\hat{s}_1}(-(1-\delta)\log\mu(\hat{s}_1))$$

Let $N(\delta)$ be the worst NE payoff to P1.

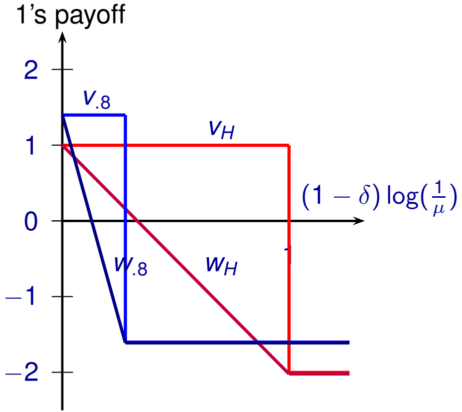
Corollary [Fudenberg Levine 1989, 1992]

If the set of $\hat{s}_1$ such that $\mu(\hat{s}_1) > 0$ is dense in $\Delta(A_1)$, then

$$\liminf_{\delta\to 1} N(\delta) \geq \sup_{\hat{s}_1\in\Delta(A_1)} v_{\hat{s}_1}(0)$$

Example: Quality choice game

	b	n
H	$1, 1$	$-2, 0$
L	$3, -1$	$0, 0$



$\hat{s}_1 = H$

- $d(H \| \frac{1}{2}) = \log 2 = 1$
- $v_H(d) = 1$ if $d < 1$
- $v_H(d) = -2$ if $d \geq 1$

$\hat{s}_1 = \hat{p}H + (1-\hat{p})L$, $\hat{p} > \frac{1}{2}$

- $d(\hat{p} \| \frac{1}{2}) = 1 - H(\hat{p})$
- $v_{\hat{p}}(d) = 3 - 2\hat{p}$ if $d < 1 - H(\hat{p})$
- $v_{\hat{p}}(d) = -2\hat{p}$ if $d \geq 1 - H(\hat{p})$

Section Conclusions

Imperfect monitoring general commitment types,

- Fixing a probability of commitment types with “full support”
- When the long-run becomes arbitrarily patient
- All NE give the long-run player at least a payoff arbitrarily close to the Stackelberg payoff