6.972: Game Theory April 26, 2005

Lecture 20: Mechanism Design and the Revelation Principle

Lecturer: Asu Ozdaglar

1 Agenda

- Mechanism Design
- Revelation principle (Incentive compatibility and individual rationality)
- "Optimal" Mechanisms (maximizing expected revenue) (Myerson [81])

We will study mechanism design in two contexts:

- Auction theory (indivisible object, risk neutral agents): Krishna Chapter 5.
- Social choice theory (divisible resource, risk averse agents). We will look for "efficient" (Pareto efficient or budget balanced) allocations. (MasColell, Whinston, Greene - Microeconomic theory, Chapter 23)

Today, we will focus on the first topic.

2 Auction theory

An auction is one way of selling an object to potential users with private values. The seller determines the rules/format of the auction. The goal is to determine the best way to allocate the object.

We will achieve this by abstractly defining a selling mechanism involving:

- Strategies of buyers
- Outcomes: Allocations and payments

such that individual actions of self interested buyers maximize our objective.

3 Model

We have a single indivisible object for sale. Define

- N risk neutral buyers, $\mathcal{N}$
- X_i, the private valuation of buyers i. $X_i \sim F_i$ where F is the cumulative distribution function, which is continuously differentiable. $f_i = F_i'$ is the associated probability density function with support $X_i = [0, w_i]$.

Note that this model allows for asymmetry among buyers. We will assume that the X_i are independently distributed. Hence, we have $X = \prod_{j=1}^{N} X_j$ and $X_{-i} = \prod_{j \neq i} X_j$. By independence, we have the joint density of $\underline{x}$, $f(\underline{x}) = f_1(x_1) \dots f_N(x_N)$. Analogously, we define, $f_{-i}(x_{-i})$, the joint density of $x_{-i} = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_N)$.

4 Mechanism

A selling mechanism (B, π, μ) has the following components:

- B_i: A set of possible messages (bids/strategies) for each buyer i. $B = \prod_j B_j$.
- $\pi : B \to \Delta$: Allocation rule where Δ is the probability simplex, i.e. the set of (N-dimensional vector) probability distributions over $\mathcal{N}$, the set of buyers.
- $\mu : B \to \Re^N$: Payment rule.

Given a bid b by buyer i, $\pi_i(b)$ is the probability that i will get the object, and $\mu_i(b)$ is the payment that i must make.

For example, the first and second price auctions are mechanisms with:

$$\begin{aligned} B_i &= X_i \\ \pi_i(b) &= \begin{cases} 1 & \text{if} \quad b_i > max_{j \neq i} b_j \\ 0 & \text{if} \quad b_i < max_{j \neq i} b_j \end{cases} \\ \mu_i^I(b) &= \begin{cases} b_i & \text{if} \quad b_i > max_{j \neq i} b_j \\ 0 & \text{if} \quad b_i < max_{j \neq i} b_j \end{cases} \\ \mu_i^{II}(b) &= \begin{cases} max_{j \neq i} b_j & \text{if} \quad b_i > max_{j \neq i} b_j \\ 0 & \text{if} \quad b_i < max_{j \neq i} b_j \end{cases} \end{aligned}$$

Every mechanism defines a game of incomplete information.

- Strategies: $\beta_i : [0, w_i] \to B_i$.
- Payoffs: Expected payoff for a given strategy profile and selling mechanism.

The mechanism induces a Bayesian Nash Equilibrium (BNE) (Given other people's strategies β_{-i}, $\beta_i(x_i)$ maximizes the expected payoff of player i, $\forall i$ and $\forall x_i$).

Note that as defined, the mechanism design problem could be very complicated because the messages can be arbitrarily complex. However, a surprising result called the revelation principle allows us to narrow our search down.

5 Revelation Principle

The basic idea is that the set of messages can be restricted to sets of values. To formalize this, we need the following definition.

Definition 1 *A* **direct mechanism** (Q, M) *is defined as:*

- $B_i = X_i$.
- $Q : X \to \Delta$: *$Q_i(x)$ is the probability that i will get the object.*
- $M : X \to \Re^N$: *$M_i(x)$ is the payment by i.*

The outcome of the mechanism is $[Q(x), M(x)]$. The mechanism is said to have a **truthful equilibrium** if it is an equilibrium for each buyer to reveal his/her true value.

The idea of the revelation principle, then, is that the outcomes resulting from an equilibrium of any mechanism can be "replicated" by a truthful equilibrium of some direct mechanism.

Proposition 1 *Given a mechanism (B, π, μ) and an equilibrium β for this mechanism, there exists a direct mechanism in which:*

1. *It is an equilibrium for each buyer to report truthfully.*
2. *The outcomes are the same as in the given equilibrium β of the original mechanism*

Proof Sketch: Define $Q : X \to \Delta$, $M : X \to \Re^N$ as $Q(x) = \pi(\beta(x))$ and $M(x) = \mu(\beta(x))$. Instead of buyers submitting message $b_i = \beta(x_i)$, the mechanism asks the buyer to report their value and makes sure the outcome is the same as if they had submitted $\beta_i(x_i)$. It is easy to see that this mechanism is truthful by considering the case when user i bids z_i instead of x_i. $\square$

6 Incentive Compatibility

We define:

$$\begin{aligned} q_i(z_i) &= \int_{X_{-i}} Q_i(z_i, x_{-i}) f_{-i}(x_{-i}) dx_{-i} \\ m_i(z_i) &= \int_{X_{-i}} M_i(z_i, x_{-i}) f_{-i}(x_{-i}) dx_{-i} \end{aligned}$$

In other words, if player i reports z_i while other players are reporting their true values x_{-i}, $q_i(z_i)$ is the probability that he gets the object, and his expected payment is $m_i(z_i)$. The expected payoff of buyer i in this case is $q_i(z_i)x_i - m_i(z_i)$.

Definition 2 *A direct mechanism (Q, M) is* **incentive compatible** *(IC) iff $\forall i$, $\forall x_i$, we have: $U_i(x_i) = q_i(x_i)x_i - m_i(x_i) \geq q_i(z_i)x_i - m_i(z_i)\ \forall z_i$.*

We will now examine some of the properties of U under IC.

1. $U_i(.)$ is a convex function. This is because $U_i(x_i) = max_{z_i}\{q_i(z_i)x_i - m_i(z_i)\}$ is the maximum of a set of affine (convex) functions, and is therefore itself convex.
2. We can write $\forall x_i$, $\forall z_i$,

$$\begin{aligned} U_i(z_i) &\geq q_i(x_i)z_i - m_i(x_i) \text{ (by IC)} \\ &= q_i(x_i)x_i - m_i(x_i) + q_i(x_i)(z_i - x_i) \\ &= U_i(x_i) + q_i(x_i)(z_i - x_i) \end{aligned}$$

Hence, $IC \Leftrightarrow U_i(z_i) \geq U_i(x_i) + q_i(x_i)(z_i - x_i)\ \forall x_i, z_i, i$.

As can be seen in Figure 1, $q_i(x_i)$ is the subgradient of $u_i(.)$, and is the gradient wherever differentiable.

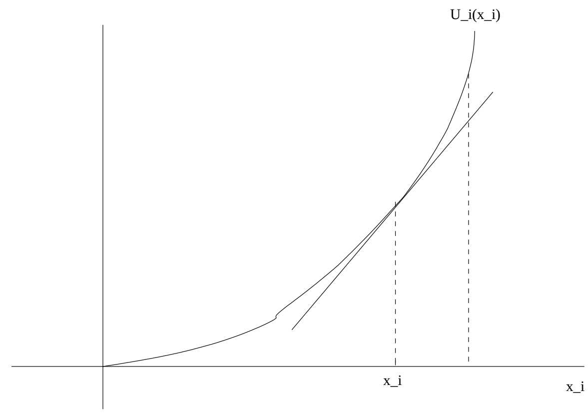

Figure 1: Gradient of U

Hence, $U_i'(x_i) = q_i(x_i)$ at every point x_i where u_i is differentiable. Since U_i is convex, q_i is nondecreasing and hence differentiable almost everywhere. Hence, we have:

$$\boxed{U_i(x_i) = U_i(0) + \int_0^{x_i} q_i(t_i)dt_i} \quad (1)$$

Remark: The expected payoff to a buyer in an IC direct mechanism (Q, M) depends, up to a constant, only on the allocation rule. This payoff equivalence implies revenue equivalence.

Proposition 2 *If the direct mechanism (Q, M) is IC, then $\forall I, x_i$, the expected payment is:*

$$\boxed{m_i(x_i) = m_i(0) + q_i(x_i)x_i - \int_0^{x_i} q_i(t_i)dt_i}$$

Thus, the expected payments in any two IC mechanisms with the same allocation rules are equivalent up to a constant.

Proof Sketch: Use $U_i(x_i) = q_i(x_i)x_i - m_i(x_i)$ and Equation 1. $\square$

*R*emarks:

- Given two BNEs of two different auctions such that, for each:

  1. $\forall(\theta_1, \ldots, \theta_N)$, the probability of getting the object is the same.
  2. the expected payoff at 0 value is the same.

  These equilibria generate the same expected revenue for the seller.

- The above proof generalizes the result from last time. If the buyers are symmetrical and there is an increasing symmetric equilibrium, then the object is allocated to buyer with highest value (in all such auctions, the allocation rule is the same). A technical condition pins down expected payments by assuming $m_i(0) = 0$.

We introduce the notion of **Individual Rationality** (participation constraints), so as to attract all potential buyers.

Definition 3 *A direct mechanism (Q, M) is individually rational (IR) of $\forall I, x_i$, $U_i(x_i) \geq 0$.*

If (Q, M) is IC, then it is IR iff $\forall i, U_i(0) \geq 0$. $U_i(0) = -m_i(0) \Leftrightarrow m_i(0) \leq 0$.