Linear Inequalities and Mechanism Design

Rakesh Vohra

MEDS Department,
J.L. Kellogg Graduate School of Management
Northwestern University
2001 Sheridan Road
Evanston IL 60208-5428
r-vohra@kellogg.northwestern.edu

A mechanism is a black box that associates with each profile of preferences of a set of agents and an outcome from a given set of ouctomes. The subject of mechanism design is concerned with the existence (or non-existence) of mechanisms that satisfy 'desirable' properties. For example, do agents have the incentive to truthfully report their preferences? Is it non-discriminatory in the sense that the outcome selected is independent of the names of agents?

The literature has two strands. The first, called Social Choice, beginning with Arrow's Impossibility Theorem, makes very few assumptions about preferences. Specifically, agents have an ordering over alternatives and the goal is to identify a mechanism that will produce an ordering that summarizes the individual preference orderings. The major results are of two kinds. The first are impossibility results: non-existence of mechanisms with some desired set of properties. The second are possibility results; restrictions on the families of preference orderings that admit the existence of mechanisms with attractive features.

To make the questions being considered precise some notation will be useful. Let $\mathcal{A}$ denote the set of alternatives (at least three). Let Σ denote the set of all transitive, antisymmetric and total binary relations on $\mathcal{A}$. An element of Σ is a preference ordering. Notice that this set up excludes indifference. The set of admissible preference orderings for members of a society of n-agents (voters) will be a subset of Σ and denoted Ω. Let Ω^n be the set of all n-tuples of preferences from Ω, called **profiles**. An element of Ω^n will typically be denoted as $\mathbf{P} = (\mathbf{p_1}, \mathbf{p_2}, \ldots, \mathbf{p_n})$, where $\mathbf{p_i}$ is interpreted as the preference ordering of agent i.

An n-person Social Welfare Function is a function $f : \Omega^n \rightarrow \Sigma$. Thus for any $\mathbf{P} \in \Omega^n$, $f(\mathbf{P})$ is an ordering of the alternatives. We write $xf(\mathbf{P})y$ if x is ranked above y under $f(\mathbf{P})$. An n-person **Arrovian Social Welfare Function** (ASWF) on Ω is a function $f : \Omega^n \rightarrow \Sigma$ that satisfies the following two conditions:

1. **Unanimity:** If for $\mathbf{P} \in \Omega^n$ and some $x, y \in \mathcal{A}$ we have $x\mathbf{p_i}y$ for all i then $xf(\mathbf{P})y$.

2. **Independence of Irrelevant Alternatives:** For any $x, y \in \mathcal{A}$ suppose $\exists \mathbf{P}, \mathbf{Q} \in \Omega^n$ such that $x\mathbf{p_i}y$ if an only if $x\mathbf{q_i}y$ for $i = 1, \ldots, n$. Then $xf(\mathbf{P})y$ if an only if $xf(\mathbf{Q})y$.

The first axiom stipulates that if all voters prefer alternative x to alternative y, then the social welfare function f must rank x above y. The second axiom states that the ranking of x and y in f is not affected by how the voters rank the other alternatives. An obvious Social Welfare function that satisfies the two conditions is the *dictatorial rule*: rank the alternatives in the order of the preferences of a particular voter (the dictator). Formally, an ASWF is **dictatorial** if there is an i such that $f(\mathbf{P}) = \mathbf{p_i}$ for all $\mathbf{P} \in \Omega^n$. The domain Ω is **Arrovian** if it admits a non-dictatorial ASWF. A fundamental question is this: which domains are Arrovian?

In the first lecture I will show how the problem of deciding if a domain is Arrovian can be formulated as an integer program. For many domains this integer program is in fact a linear program. The program provides for a systematic and simple way to derive most of the important results in social choice as well as some new ones. It is usually presumed that the preferences to be summarized are reported truthfully. This assumption is sometimes untenable, and a common requirement that is imposed is that the mechanism used must give he agents the incentive to report their preferences truthfully. This property is called incentive compatibility and can be formalized in a number of ways. If very few restrictions are placed on preferences then one obtains an impossibility result of the form that no non-trivial mechanism that is incentive compatible exists.

The second strand, usually associated with Auction Theory, restricts preferences to be quasi-linear i.e., each agents utilities can be denominated on a common monetary scale. Restricting preferences in this way allows one to escape the impossibility results alluded to earlier. It is possible in this case to identify non-trivial mechanisms that are incentive compatible. The problem of characterizing the class of mechanisms that are incentive compatible can be formulated as the solution to a a system of linear inequalities.

References

[1] J. Sethuraman, C.P. Teo and R. Vohra. Integer Programming and Arrovian Social Welfare Functions, *Mathematics of Operations Research*, 28, 309-326, 2003.

[2] J. Sethuraman, C.P. Teo and R. Vohra. Anonymous Monotonic Social Welfare Functions *Journal of Economic Theory*, forthcoming.

[3] H. Gui, R. Muller and R. Vohra. Characterizing Dominant Strategy Mechanisms with Multidimensional Types, report 2004.

[4] A. Malakhov and R. Vohra. Single and Multi-Dimensional Optimal Auctions - A Network Approach, report 2004.