COWLES COMMISSION DISCUSSION PAPER: ECONOMICS NO. 2062

The Problem of Finding Optimal Decisions [1]

Leo Törnqvist

December 15, 1952

We now assume that we have been able to construct an indicator of goodness $g(a)$ to every plan of action considered $a \in A$. The problem is then to invent an optimal method for finding such plans $a \in \hat{A}$ that

$$g(a) = g(\hat{a}) = \operatorname*{Max}_{a \in A} g(a).$$

The problem is thus to find a solution to a conditional optimum problem.

This problem can be mathematically attacked in many different ways:

I. The method of transforming the conditional optimum problem to an unconditional optimum problem for a function $G_\varepsilon(a)$ which, when $\varepsilon \to 0$, converges to $g(a)$ for $a \in A$ and to $-\infty$ for $a \notin A$.

$$\hat{a} = \lim \hat{a}_\varepsilon$$

$$G_\varepsilon(\hat{a}_\varepsilon) = \operatorname*{Max}_{a} G_\varepsilon(a)$$

Such functions $G_\varepsilon(a)$ are for instance

$$G_\varepsilon(a) = g(a) + {}^{2}\log \Phi(a \in A \mid \varepsilon) \text{ where}$$

$$0 < \Phi(a \in A \mid \varepsilon) < |\varepsilon| \text{ for a point } a \text{ such that } a,\ a - \varepsilon$$

and $a + \varepsilon$ are all outside A.

---

1. Research undertaken by the Cowles Commission for Research in Economics under contract Nonr-358(01), NR 047-006 with the Office of Naval Research.

$1 > \Phi(a \in A|\varepsilon) > 1 - |\varepsilon|$ for a point a such that a, $a - \varepsilon$ and $a + \varepsilon$ are all inside A.

II. The method of Lagrange multipliers

$$G\lambda(a) = g(a) + G(\lambda, a)$$

where λ is a table containing so many numbers n_A that there always exists a finite number of tables (a, λ) such that (a, λ) is a solution to the system of equations

$$\begin{cases} G(\lambda, a) = 0 \text{ for every } a \in \hat{a} \subset A. \\ \dfrac{\partial G\lambda(a)}{\partial a_z} = \dfrac{\partial g(a)}{\partial a_z} + \dfrac{\partial G(\lambda, a)}{\partial a_z} = 0 \end{cases}$$

where a_z is one number in the position z of the set of numbers defining a_o. If n_A is reduced to $n_A = 1$ no solutions to II exist. The optimal solutions $\hat{a} \subset a$ are all contained in some of the solutions to the system of equations II.

III. The method of <u>moving</u> from $a_v \in A$ to $a_{v+1} \in A$ whenever $g(a_{v+1}) > g(a_v)$. The numbers a_{v+1} are sought from a neighbourhood $a_v + e_v \circ \Delta_v$ where Δ_v is a table containing only one number $\Delta_v \gtrless 0$ and e_v is a table of numbers containing only the numbers ± 1 and 0 in different positions z of the tables A. An upper limit $\bar{\Delta}_v$ for Δ_v is then defined for every ε_v by the requirement that $a_v + e_v \circ \Delta_v \subset A$. The points $a_v + e_v \Delta_v$ are studied in the order of the size of the sum $N_v = \sum_z |e_{vz}|$. When a direction e_v with $\bar{\Delta}_v > 0$ is found such that $\max_{0 < \Delta_v < \bar{\Delta}_v} g(a_{v+1} = a_v + e_v \Delta_v) > g(a_v)$, we proceed to find a new point a_{v+2} better than a_{v+1} and so on until the process stops, at an $a_n = a_n(a_o)$. If N_v is restricted to a number $N_v \lesseqgtr \bar{N}_v$, ($\bar{N}_v$ for instance = 6), at a table $a_n(a_o)$ we need not necessarily have $g(a_n) = g(\hat{a})$ but the difference $g(\hat{a}) - g(a_n)$ is usually relatively small. By varying the "opening" a_o of the game we can find a set of tables $a_n(a_o)$, which have the property that some kind of "probability" that $\underset{a_o}{\mathrm{Max}}\ (g(\bar{a}) - g(a_n(a_o)) > \varepsilon\ )$ goes to zero very rapidly with an increasing number of trials for openings a_o and/or increasing numbers of N_v and/or ε. For functions g(a) which have no uniquely defined partial derivates $\frac{\partial g(a)}{a_z}$ the method III to construct an increasing sequence $g(a_v) \rightarrow g(\hat{a})$ seems to be the only promising method available. This method is also the only one for finding the over all optimal solutions in the cases when the method II picks out a denumerable set of extremal points (relative maxima or minima). In all complex cases we thus have to rely on a method of successive improvements $g(a_v) \rightarrow g(\hat{a})$. Fortunately this method seems to be also the most promising method for finding optimal decisions in non trivial decision problems. In some cases it is also possible to calculate a measure

$$M(A \cap (g(a) > g)) = M(A,g)$$

for the set of tables a, belonging to the intersection of the sets A, and the set of a, fulfilling $(g(a) > g)$ as a function of g. This measure is zero for $g \geq g(\hat{a})$.

If we then determine the largest value of g for which $M(A,g) \neq 0$, we have the possibility of determining $g(\hat{a})$. The limit sets $(a,g) \subset (A,g)_\varepsilon$ such that

$$\underset{\varepsilon}{\overset{0}{\uparrow}} \begin{cases} A \cap (g(a) > g + \varepsilon\ ) = 0 \\ A \cap (g(a) > g - \varepsilon\ ) \neq 0. \end{cases}$$

gives the solution to the optimal problem considered, but it is often too complicated to use this straightforward way of finding an $a \subset \hat{a}$.

The problem of finding an optimal way for finding optimal decisions is solvable only if we can construct an indicator of goodness for different methods of seeking ways to a whole set of approximative solutions in ε - neighbourhoods of the optimal solutions. A non optimal solution to the decision problem to maximize $g(a)$; $a \subset A$ can namely be a better solution to the problem of maximizing the indicator of goodness

$$g(a_{n\lambda}, F_\lambda, \lambda) \; ; \; F_\lambda(\varepsilon_\lambda) = P(g(a) - g(a_{n\lambda}) \leq \varepsilon_\lambda)$$

for the whole activity to find a method λ to get good decisions $a_n \subset A$. It seems to be clear that the variance of ε_λ is usually > 0 for the best method $\lambda = \hat{\lambda}$.

For a full understanding of the principles expressed, they must be clarified by examples of applications to special optimum problems. This paper is only an introduction to this field of study.