Decision Making in Economics

A. Postlewaite

March 20, 2007

- How much does investment increase if taxes are reduced?

- How much do aggregate savings increase if we reduce social security payments?

- How much will high school graduation rates go up if we give $5,000 payments to graduates?

Economics addresses questions at the aggregate level but predictions are based on decisions made by individuals

Potential Problem: Economics, game theory and decision theory have at their heart an unrealistic standard of rationality

Is Russia a dangerous country?

Not everyone is aware of the democratic peace phenomenon, yet, no new information is conveyed by it.

Most people have indices of:

Wars by country

Wars by dates

but not Wars by type of regime

Framework: people have a large number facts or observations

From this "knowledge base", they make judgements and decisions about related problem they may face

Standard Assumption: people know all the "regularities" in their knowledge base

Consequence: The only way you can change my behavior is by presenting me with new facts

This seems both implausible and unintuitive

Attempts at modelling rationality more realistically often run into internal inconsistency

Would like a model that recognizes the limits on computational ability in which people are aware of the limitations

In particular, we don't want people to behave subrationally

Fact-Free Learning (American Economic Review, 2005)

Aragones, Gilboa, Postlewaite and Schmeidler

Cases have attributes

Data are given as numbers in $[0,1]$:

x_{ij} – the degree to which case i has attribute j.

Example: Will the US go to war with Iran?

Case	M1	M2	D1	D2	T	W
WWII	.7	1	1	0	0	1
Missile crisis	1	1	1	0	1	0
Gulf War	1	.3	1	0	1	1

Mi – how strong is country i?

Di – is country i a democracy?

T – was it after 1945?

W – did war result?

Formally,

$C = \{1, ..., n\}$ – a set of cases

$A = \{1, ..., m\}$ – a set of attributes

$X : C \times A \to [0, 1]$ – data

$x_{ij} \equiv X(i, j)$ the degree to which case i has attribute j

A naive measure of complexity: number of attributes employed by the rule, $k = |K|$ (more later on this)

Trading off complexity against accuracy, we need a measure of the accuracy of a rule. For linear regressions, $r = R^2$ or $r =$ adjusted R^2

In the k-r space, X has a feasible set of rules $F(X) \subset R_+ \times [0, 1]$.

Given the feasible set $F(X)$, one may find various rules and try to make predictions or decisions based on them.

The decision maker/predictor has to choose a rule optimally trading off accuracy and complexity in the database X.

She does this so as to maximize $v(k, r)$ with $v_k < 0$ and $v_r > 0$.

However, there is a problem:

One does not know the set $F(X)$.

Moreover, one cannot, in general, compute the Pareto frontier of this set.

Linear Regression

Predict a variable Y given the predictors $X_1, ..., X_m$.

Theorem 1: For every $r \in (0, 1]$, the following problem is NP-Complete: Given explanatory variables $X = (X_1, ...X_m)$, a variable Y, and an integer $k \geq 1$, is there a subset K of $\{X_1, ...X_m\}$ such that $|K| \leq k$ and $R^2 \geq r$?

Consider the problem of identifying the determinants of growth. Suppose that you have a data set with 100 variables linked to growth in a particular year. Suppose further that a computer could run 10 million regressions per second.

How long would it take to determine whether there were 13 variables that could account for 75% of the variance in growth?

Answer: about 22 years

Implicit in this is the notion of the "burden" or cost that a complex rule places on the decision maker - the number of non-zero coefficients. One could think of more general costs

$\Phi(b)$ - the "cost" of a rule with coefficients $b = b_1, ..., b_M$

General Problem (GP): $\text{Max}_{b \in R^M} R^2(b)$ s.t. $\Phi(b) \leq C$.

Our rule: $\Phi(b) =$number of non-zero elements of b

Interesting class of rules: $\Phi(b) = \Sigma_j \phi(b_j)$ where $\phi : R \to R$ and $\{b_j\}_{j \in J}$ are the coefficients

Subset of rules used in statistics: $\Sigma_j |b_j|^\gamma$

($\gamma = 2 :$ Ridge regressions; $\gamma = 1 :$ Lasso)

Theorem: (Eilat, 2006)

(i) If ϕ is weakly convex GP is easy

(ii) If ϕ is weakly concave GP is NP-Hard

(iii) If ϕ is non-decreasing and discontinuous at 0, GP is NP-Hard.

Questions:

Are there good approximations that are easy?

What cost functions capture the complexity of finding rules? Experiments might be useful.

R^2 is one measure of accuracy. For what other measures of accuracy is the modified problem easy?

Application: Consumer Choice

Standard approach: Maximize a concave utility function over a budget set, a compact, convex subset of R^L

BUT - This is a cartoon view of real-world problems people face

Real-world problem: Should I take a job at Georgia Tech? Should I buy another car, would my wife work, would I live in the city or the suburbs, how many kids should I have (or at least aim for), what schools should

I send my (existing) kids to, what dentist would I use, should I paint the walls in the kitchen, etc, etc.

Fact: people don't look across all affordable bundles and choose.

Plausible formulations of THIS problem are NP-Hard (Gilboa, Postlewaite and Schmeidler).

People use heuristics: First decide on the job ignoring what color to paint the kitchen walls, and so on, then make a set of “second level” choices, ...

Questions:

When will this sequential approach be optimal?

How hard is it to determine what choices should be in each level?

How bad can the choices be if one makes a “small” mistake?

Application: Incomplete Contracts

Implications

Agents can “agree to disagree”

Suboptimal rules can be “locally” optimal

The rules people employ can be path dependent