Decision-Theoretic Entropy

Pierfrancesco La Mura
Leipzig Graduate School of Management
plamura@hhl.de

Abstract

We introduce an axiomatic approach to the problem of inferring a complete and transitive weak ordering representing the agent's preferences given a set of observed constraints. The axioms characterize a unique inference rule, which amounts to the constrained maximization of a certain formula we derive. The formula can be interpreted as the entropy of the agent's preference ordering, and its unique maximand identifies the simplest rationalization of the observed behavior.

1   Introduction

One of the basic assumptions of decision theory is that an agent's observed behavior can be rationalized in terms of an underlying preference ordering. If the observed behavior is consistent with the ordering, the latter may in principle be inferred from the data at hand. Yet, unless the observations suffice to completely identify the ordering, a variety of different motivations can still be ascribed to the observed behavior. It is natural, in those cases, to look for an inference method that would identify, among all the possible preference orderings which are consistent with the observed behavior, the one which constitutes the most appropriate explanation of the data according to some well-defined and possibly intuitive notion.

---

[0] An earlier version of this work appeared as chapter 6 of the author's Ph.D. dissertation, *Foundations of Multi-Agent Systems*, Stanford GSB (1999). The author would like to thank Matthias Herfert, Hagen Lindstädt, Bob Wilson and Yoav Shoham for useful comments at different stages of preparation.

In this paper we introduce an axiomatic approach to the problem of inferring a complete and transitive weak ordering representing the agent's preferences given a set of observed constraints. The axioms characterize a unique inference rule, which amounts to the constrained maximization of a certain formula we derive. The formula can be interpreted as the entropy of the agent's preference ordering, and its unique maximand identifies the simplest rationalization of the observed behavior. Intuitively, the resulting ordering is simplest in that, whenever possible, it presumes indifference and independent preferences.

The decision-theoretic literature offers a variety of different methods for the estimation of an unknown preference structure from consumer data, among which Conjoint Analysis [Luce and Tukey (1964), Green and Srinivasan (1978)] is the best known, and also the most popular in marketing research. Yet, to the best of our knowledge, the method presented here constitutes the first non-parametric approach to preference estimation.

2   Setup and notation

Let X be a finite set of states, and let A be the Boolean algebra of all the subsets of X.

We assume that the agent's beliefs and preferences admit an expected utility representation (p, u), where p is a probability measure defined on A, which for simplicity we assume to have full support, and u is a function associating to each state $x \in X$ a positive real number $u(x)$.

Next, we extend the utility function from states to general events by defining

$$u(E) = \sum_{x \in X} u(x)p(x|E),$$

and normalize it by requiring that $u(X) = 1$ and $u(\varnothing) = 0$.

Let $v(E) = u(E)p(E)$; we say that v is the *value* of event E. It is easily verified that under the above normalization v is a probability measure. Moreover, since p is strictly positive, for any nonempty event E we can write

$$u(E) = \frac{v(E)}{p(E)}.$$

It follows that, in this perspective, utility is simply the ratio of two probability measures, one representing value and the other belief.

The *entropy* of a probability measure p [Shannon (1948)] is defined as

$$H(p) = -\sum_{x \in X} p(x) \ln p(x).$$

Furthermore, if p and q are two probability measures, the *relative entropy* (or *cross-entropy*) of q with respect to p is defined as

$$D(q||p) = \sum_{x \in X} q(x) \ln \left(\frac{q(x)}{p(x)} \right).$$

This quantity is always nonnegative, and is equal to zero if and only if $q \equiv p$.

The relative entropy of value with respect to probability is given by

$$D(v||p) = \sum v(x) \ln(\frac{v(x)}{p(x)}) = \sum p(x)u(x) \ln u(x) = E_p[u(x) \ln u(x)]. \quad (1)$$

This expectation is always nonnegative, and is equal to zero if and only if $v \equiv p$, in which case $u(x) = 1$ for all x, i.e. there is complete indifference among all states. In analogy to its probabilistic counterpart, we define the *entropy* of a utility function u with respect to p as

$$H_p(u) = -E_p[u(x) \ln u(x)].$$

3   Axioms and main result

Suppose that we are trying to estimate an agent's utility function over a set of possible states X, whose subjective probabilities are known. Further, assume that observations of this agent's behavior correspond to closed and convex constraints on the set of all possible utilities (e.g. monotonicity constraints, constraints on marginal rates of substitution from purchasing behavior, constraints on preferences between lotteries, etc.).

Our observations will be generally compatible with an infinite number of possible utility functions; among those, we would like to identify a unique

function representing – in some sense – the simplest, or least informative, preference structure compatible with the observed data. We do so by imposing a set of axiomatic conditions on the selection rule, based on invariance considerations. Our conditions mirror those introduced by Shore and Johnson [Shore and Johnson (1980), Johnson and Shore (1983)] for the selection of posterior probabilities, and can be informally stated as follows.

1. **No information**: if there is no information, assume indifference.

2. **Uniqueness**: the result should be unique.

3. **Invariance**: the choice of coordinate system should not matter.

4. **System Independence**: it should not matter whether one accounts for independent information about independent systems together or separately.

5. **Subset Independence**: it should not matter whether one incorporates information on mutually disjoint subsets of system states in terms of separate conditional utilities or in terms of the full system utility.

Let U be the set of normalized expected utility functions on a set X, and let $\widehat{C}$ be the set of all closed and convex subsets of U. Then any element $\widehat{I}$ of $\widehat{C}$ can be expressed in terms of a (possibly infinite) set of inequalities of the type

$$\sum u(x)f(x) \geq 0,$$

and conversely any such set of constraints identifies an element of $\widehat{C}$. We assume that our information on the agent's preferences always takes the form of a closed and convex constraint on the agent's utility.

Without loss of generality, we can state our inference problem as follows: find u such that

$$\widehat{H}(u,p) = \min_{u' \in \widehat{I}} \widehat{H}(u',p),$$

where $\widehat{H}$ is a suitable functional, p is given, and $\widehat{I}$ is a closed and convex constraint.

38

Equivalently, since $u = v/p$, the problem can be restated as $\min_{v \in I} H(v,p)$ for a suitable H, where I is the (closed and convex) constraint on v corresponding to the closed and convex constraint $\widehat{I}$. We shall work with this equivalent formulation, as it is more convenient for our treatment, and denote by C the set of all closed and convex constraints on v. We remark that, if probabilities are known (as we assumed throughout), then the observed constraints – which come in the form of inequalities on the utilities – can be equivalently restated in terms of value inequalities.

We now introduce an inference operator $\circ$, which associates to any constraint I and probability measure p a value function $v = p \circ I$. This operation can be regarded as the outcome of the above minimization for some functional H.

Our assumptions can now be formally stated as follows.

A1
(No information). Let I_0 be the null constraint $\{u \in U\}$. Then $p \circ I_0 = p$.

Assumption 1 is appealing: it says that in the absence of any information about the utilities, the outcome of the inference process should be the trivial utility function representing complete indifference ($v \equiv p$).

A2
(Uniqueness). $p \circ I$ is unique for any prior p and new information $I \in C$.

Assumption 2 is also quite appealing: it says that the outcome of the inference process should be unambiguously determined.

A3
(Invariance). Let Γ be a transformation from $x \in X$ to $y \in X'$, with $(\Gamma v)(y) := v(\Gamma^{-1}(y))$. Let ΓI be the set of value measures on X' corresponding to the original constraint I. Then, for any p and I,

$$(\Gamma p) \circ (\Gamma I) = \Gamma(p \circ I).$$

Assumption 3 says that coordinate transformations should not matter: the inference method should be invariant with respect to relabelings.

Now suppose that the state space is composed of two independent subspaces, i.e. $X = X_i \times X_2$, and that p_i, v_i are probability measures on X_i

$(i = 1, 2)$. Moreover, let I_1 and I_2 two constraints on v_1 and v_2 respectively. We then require that the following holds.

A4
(System Independence). $(p_1 \times p_2) \circ (I_1 \times I_2) = (p_1 \circ I_1) \times (p_2 \circ I_2)$.

The assumption says that it should not matter if the inference rule on the two independent subsystems is applied jointly or separately.

Finally, let $\{X_k\}$ be a partition of X, and let p_{X_k}, v_{X_k} be the corresponding conditional measures. Let I_k be a family of constraints involving only the conditional values v_{X_k}, let $I = \cap_k I_k$, and let M be a constraint of the form $v(X_k) = m_k$, where the m_k are known values. We require that the following holds.

A5
(Subset Independence). Let $v = p \circ (I \cap M)$. Then $v_{X_k} = p_{X_k} \circ I_k$.

This last assumption states that it should not matter if new information on the conditional values of mutually exclusive subsets is incorporated jointly or separately. The following result then holds:

Theorem 1 *There exists an inference rule which satisfies all the above assumptions. Moreover, any such rule is equivalent to the minimization of (1).*

Proof
The result is an immediate corollary of Theorems III and IV in (Shore and Johnson 1980). ∎

Even though our treatment focuses on the case of a finite state space, we remark that axioms $A1 - A5$ also guarantee the existence of an analogous inference rule based on cross-entropy minimization in the continuous case, via the corresponding results in (Shore and Johnson 1980, Johnson and Shore 1983) for probabilistic cross-entropy. In the continuous case, though, unless the utility function is known to be bounded, the requirement that it be also non-negative (necessary for the construction of the value measure) is not without loss of generality.

4   Examples

As a first example, consider an agent who has preferences over combinations Health (H) and Wealth (H). Health and Wealth are independent Boolean events, to which the agent assigns subjective probability equal to 0.8 and 0.3 respectively. All we know about the agent's utility is that $u(H) = 3u(\neg H)$, and $u(W) = 2u(\neg W)$. From this, we would like to infer the agent's utilities on all possible combinations of Health and Wealth. Minimizing $E_p[u \ln u]$ subject to the above constraints returns at once the utility function in the following table (up to a multiplicative normalization constant).

	W	$\neg W$
H	6	3
$\neg H$	2	1

Observe that the resulting utilities are independent, in the following sense: the joint increment in utility relative to a given reference point is the product of the increments along each of the two coordinates. This is always true when the corresponding probabilites are independent.

As a second example, consider a consumer who has preferences over coffee (C) and milk (M). Every week, the consumer is invited to have breakfast at a friend's place. Once there, she is served either coffee, or milk, or both, or neither according to availability, with all four states being equally likely. On the way to her friend's place, the consumer says "I hope there will be coffee today." Then, following her train of thoughts, she adds "In fact I like coffee, but only with milk" (out of politeness, she consumes whatever she is served even if she doesn't like it). In symbols, $u(C) > u(\neg C)$, $u(\neg C \neg M) > u(C \neg M)$.

Compared to the previous example, things are more complicated here since the set of utilities which satisfies the constraints is not closed: for instance, we know that $u(C)$ is greater than $u(\neg C)$ but we don't know by how much. We circumvent this problem by introducing two unknown constants $\alpha, \beta \in (0, 1)$ such that $\alpha u(C) = u(\neg C)$ and $\beta u(\neg C \neg M) = u(C \neg M)$. We then proceed to minimize $E_p[u \ln u]$ subject to the latter constraints. Clearly, the resulting utilities are functions of the unknown parameters α and β. Yet, several inferences about the consumer's preferences can be drawn for all α, β in $(0, 1)$, including:

"If I cannot have coffee and milk, I would rather have neither one."

"I hope there will be milk."

"What I hate most is coffee without milk."

Finally, we present an example with a continuous state space. Specifically, we would like to infer the utility function of a given decision-maker (DM) based on the following information. At the status quo, the DM faces potential returns represented by a non-negative random variable X with density p. The DM's preferences over uncertain monetary prospects admit an expected utility representation in terms of a continuous utility function $u : \mathbb{R} \rightarrow \mathbb{R}$, unique up to positive affine transformations. We also assume that the expected utility of any lottery with bounded losses is finite (note that, by continuity, this implies that the utility function is bounded above). The utility function cannot generally be normalized to be non-negative, but the loss function $w := -au + b$ (where $a > 0$) can. We further normalize the loss function so that $E_p[w] = 1$, where p is the status-quo distribution of returns which (for the sake of this example) we assume to be a negative exponential[1] with parameter λ. Furthermore, we are told that the expected loss of another lottery q, distributed as a $Gamma(0, \frac{1}{\lambda}, 2)$,[2] is $E_q[w] = c$.

In sum, all we know about this DM is the status-quo distribution of potential returns p, and the expected loss of the lottery described above. From this we would like to infer the DM's utility function over monetary returns, u.

We now proceed to maximize the entropy of the loss function, $-E_p[w \log w]$, subject to the constraints $E_p[w] = 1$, $E_q[w] = c$. The variational condition for a maximum is that

$$-(\log w + 1)p + \alpha p + \beta q = 0.$$

Rearranging, one finds that

$$\log w = \alpha - 1 + \beta \frac{q}{p}.$$

Exponentiating, we obtain that

---

[1] The negative exponential of parameter λ has density function $\lambda \exp\{-\lambda y\}$

[2] The distribution $Gamma(0, \gamma, 2)$ has density function $\frac{1}{2\gamma^2} x \exp\left\{-\frac{x}{\gamma}\right\}$.

$$w(x) = e^{\alpha-1+\beta q(x)/p(x)} = e^{\alpha-1}e^{x\beta\lambda/2}.$$

Substituting w in the constraints we can identify α and β. From the first constraint we get

$$e^{\alpha-1} \int_0^{+\infty} \lambda e^{-\lambda x(1-\beta/2)}dx = 1,$$

which implies $e^{\alpha-1} = 1 - \beta/2$. From the second we get

$$e^{\alpha-1} \int_0^{+\infty} e^{x\beta\lambda/2}\frac{\lambda^2}{2} \; x \; e^{-\lambda x}dx = c,$$

or, rearranging,

$$\frac{\lambda^2}{2}e^{\alpha-1} \int_0^{+\infty} x \; e^{-x\lambda(1-\beta/2)}dx = c,$$

which implies

$$\frac{e^{\alpha-1}}{(1-\beta/2)^2} = c, \text{ or } \beta = 2(1 - \frac{1}{c}).$$

It follows that

$$w(x) = \frac{1}{c}e^{-x\lambda(1-c)/c}.$$

Finally, let's recover the general shape of the utility function from the loss function:

$$u = -aw + b = b - \frac{a}{c}e^{-x\lambda(1-c)c}.$$

We conclude that the maximum entropy utility function has constant absolute risk aversion, and is increasing in money if and only if the second lottery is preferred to the status quo (as one would expect, given that p is first-order stochastically dominated by q).

5 Discussion and extensions

We presented a non-parametric method for preference estimation based on a set of axiomatic requirements. The axioms characterize a unique inference rule, which amounts to the maximization of a certain formula that can be interpreted as the entropy of the decision-maker's preference ordering. The axioms mirror those introduced in [Shore and Johnson (1980), Johnson and Shore (1983)] for the case of probabilistic cross-entropy minimization, but are reinterpreted to fit the case of utility entropy maximization given a set of observed constrains and known prior probabilities. Even though our treatment was centered on the finite case, the axioms also guarantee the existence of a corresponding inference rule for the case of continuous state spaces.

One interesting extension would be to provide a unified inference approach suitable when not only the utilities, but also the probabilities are only partially known. In this perspective one would also need to specify which events are under the control of the agent, and introduce rationality constraints to link probabilities and utilities. In principle, such extension would also allow inference in multi-agent systems, provided that the relevant constraints are still closed and convex. For now, we defer that to future work.

6 References

P. E. Green and V. Srinivasan (1978), "Conjoint Analysis in Consumer Research: Issue and Outlook." *Journal of Consumer Research*, Vol. 5, pp. 103-123.

R.W. Johnson and J.E. Shore (1983), "Comments on and corrections to 'Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy'." *IEEE Transactions on Information Theory*, IT-29(6), pp. 942–943.

R. D. Luce and J. W. Tukey (1964), "Simultaneous Conjoint Measurements: A New Type of Fundamental Measurement." *Journal of Mathematical Psychology*, Vol. 1, pp. 1-27.

C.E. Shannon (1948), "A mathematical theory of communication." *The Bell System Technical Journal*, Vol. 27, pp. 379-423,623-656.

J.E. Shore and R. Johnson (1980), "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy." *IEEE Transactions on Information Theory*, IT-26, pp. 26-37.