Behavior Modeling through Inverse Problem Specification

Yoneda, Kiyoshi
Fukuoka University
yoneda@fukuoka-u.ac.jp

Childers, Charles
ForthWorks
crc@forthworks.com

Celaschi, Walter
Faculdades de Campinas
walter@flexsys.com.br

APEA, August 1–2, 2019

Abstract

This is an attempt to model individual behavior through specification of inverse problems with a tool originally devised to endow resource-constrained machines with some ability to make autonomous decisions based on inference. Utility herein is measured by the likelihood that the individual puts the decision into action. Univariate utility functions are specified by number tables rather than as mathematical functions. The number tables are interpolated into continuous functions not in the domain of utilities but in the domain of standardization functions. Having obtained a multiplicatively separable loss function from the univariate utility functions, the method follows the lines taken in maximum likelihood inference. A new performance index is introduced which is a continuous strictly monotone function of standardized variable preserving the information that loss and utility functions embody.

MSC code: 93E24 Least squares and related methods
JEL code: C3 Multiple or Simultaneous Equation Models; Multiple Variables
A preliminary version of this paper was presented at EURO 2019, Dublin.

1 Introduction

A minimal problem example

Here is a minimal example of the kind of problems this paper proposes to solve.

A secretary is going to work on a flexible time basis. Her problem is how long to work and to sleep. Ideally she wishes to work 4 [h], sleep 9 [h], and have 8 [h] free time, where [h] is for hours. The overhead such as dressing and commuting takes 2 [h]. The wage is 1/8 [cu/h] plus 30% extra per hour if over 8 [h] where [cu] is for currency unit; she wishes to earn 2 [cu] per workday.

Causality equations for the minimal example

This is a system of equations describing the problem:

$$
\begin{gathered}
Y\,x \approx \check{y} \\
\begin{bmatrix} Y_{\texttt{work}} \\ Y_{\texttt{sleep}} \\ \hdashline Y_{\texttt{free}} \\ Y_{\texttt{income}} \end{bmatrix}
\begin{bmatrix} x_{\texttt{work}} \\ x_{\texttt{sleep}} \end{bmatrix}
\approx
\begin{bmatrix} \check{y}_{\texttt{work}} \\ \check{y}_{\texttt{sleep}} \\ \hdashline \check{y}_{\texttt{free}} \\ \check{y}_{\texttt{income}} \end{bmatrix} \\
\begin{bmatrix} x_{\texttt{work}} \\ x_{\texttt{sleep}} \\ \hdashline 24 - (2 + x_{\texttt{work}} + x_{\texttt{sleep}}) \\ \begin{cases} (1/8)\, x_{\texttt{work}} & x_{\texttt{work}} \le 8 \\ 1 + (1.3/8)(x_{\texttt{work}} - 8) & 8 \le x_{\texttt{work}} \end{cases} \end{bmatrix}
\approx
\begin{bmatrix} 4 \\ 9 \\ \hdashline 8 \\ 2 \end{bmatrix}
\begin{pmatrix} [\mathrm{h}] \\ [\mathrm{h}] \\ \hdashline [\mathrm{h}] \\ [\mathrm{cu}] \end{pmatrix}
\end{gathered}
\tag{1}
$$

where '[h]' is for hours and '[cu]' is for currency units. The dashed lines separate the outcome into cause and effect parts.

Note that the causality assumed here is deterministic rather than probabilistic. Approximately equal '≈' is used here instead of equality '=' not because the result of an action involves uncertainty but because the system of equations is inconsistent having no solution: for instance, working 4 [h] which is ideal gets only 1/2 [cu] income, way less than the ideal 2 [cu].

In most real-world problems there are many more variables and equations.

Utility tables for the minimal example

Suppose the secretary's value system is as in Table 1. The table indicates, for instance, that the likelihood normalized to $(0, 1]$ that she accepts to work

Table 1: Utility specification

	outcome							
	cause				effect			
	`work`		`sleep`		`free`		`income`	
k	y [h]	u	y [h]	u	y [h]	u	y [cu]	u
1	3	.5	6	.5	3	.5	0.9	.5
2	4	1	9	1	8	1	1.2	.9
3	8	.95	10	.5	10	.5	2	1
4	9	.9					3	.5
5	10	.5						

3 [h] is 1/2, under the condition that she can sleep 9 [h], have 8 [h] free time, and a 2 [cu] income. Ideally she works 4 [h] as has been stated, but not so much against working longer up to about 9 [h]. Likewise with the other items. In most real-world problems there are more of larger tables.

Results for the minimal example

With the above input items the method to be proposed yields a recommendation as in Table 2, with itemwise frustration as shown in Figure 1 in

Table 2: Solution

	`work`	`sleep`	`free`	`income`
Ideal	4	9	8	2
Optimal	**8.9**	**8.0**	**5.1**	**1.2**
Unit	[h]	[h]	[h]	[cu]

which a bar to the left of zero indicates frustration caused by shortage of the item whereas a bar to the right indicates excess. All items fall short except `work` which is too long. Besides the recommendation and its drawbacks, the method is capable of making predictions such as:

- The probability the secretary works between 8 and 9 hours is 23%.

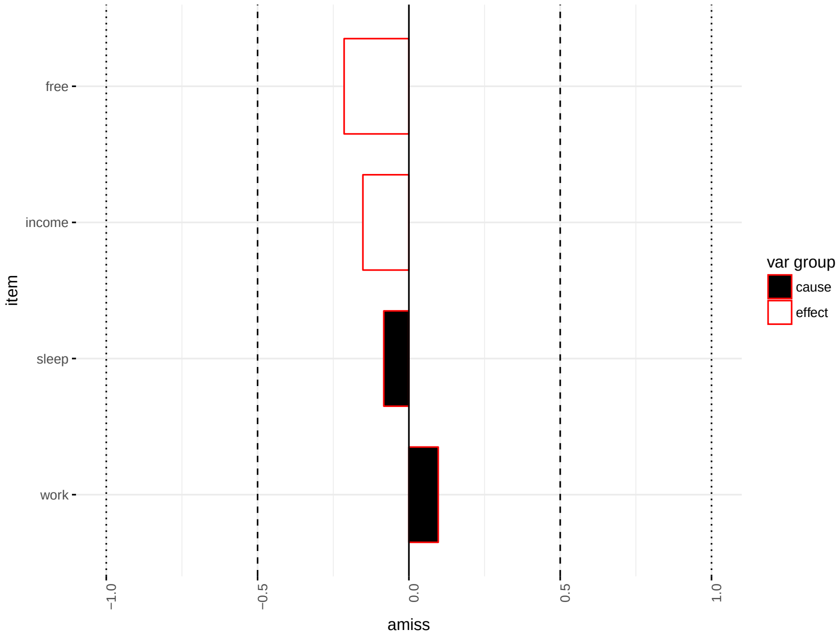

Figure 1: Rejectability chart

Prior research

There is a plethora of literature from all behavior sciences including psychology, economics, decision theory, control theory and so forth, with recent additions such as machine learning, many among which adopting the view that decision making is an inverse problem. This paper takes the classic position that statistical inference is decision making, drawing on the matching law (Herrnstein, *et al.*, 1997) and microeconomic models such as (Wichers, 1996).

Outline

The rest of this paper is organized as follows. The formulation is described in Section 2. The standard itemwise loss function and its equivalent standard itemwise utility function are defined in Section 3. Section 5 uses piecewise linear standardization functions to convert item quantities with measurement units into unit-free values enabling their mutual comparisons. Section 4 introduces a possibly new performance index to be called rejectability which is a continuous strictly monotone increasing function of the standardized variable and hence has all information carried by the standard itemwise utility function or the standard loss function. With these tools prepared, the solution of the minimal example is obtained in 6. Comparison of the proposed method against other approaches is made in Section 7.

2 Formulation

Given a decision vector x, a causality function $Y \coloneqq Y^x : x \mapsto y$ and a target vector $\check{y}$, consider attempting to solve $Y\,x \coloneqq Y(x) = \check{y}$ with respect to x . The decision x constitutes the first part of the outcomes $y \coloneqq [y_1 \;\cdots\; y_n]$,

$$x = [x_1 \;\cdots\; x_m] = [y_1 \;\cdots\; y_m] \qquad\qquad m \leq n \ ,$$

so that in $Y = \left[\cdots\; Y_i \;\cdots\right]$ for $i \leq m$, $Y_i : x \mapsto y_i$ gives $Y_i\,x = x_i = y_i$. Here Y_i is assumed continuous and monotone. For the general case $m < n$ the system of equations has no solution, so a system of relaxed equations $Y\,x \approx \check{y}$ has to be solved instead, which is called the inverse problem.

Approximate solutions are possible if '$\approx$' reflects the decision maker's value system. A typical case is statistical parameter estimation in which x is for unknown parameters, $Y_i : x \mapsto y_i$ predictors, and $\check{y}$ observation data.

The notation for functions of the form $x \mapsto y$ is Y^x in this paper, Y being the upper case of the target variable y . This shorthand notation is hoped to reduce confusion caused by the large number of variables and functions involved. Other conventions are summarized in Appendix together with a list of symbols.

Assuming that the discrepancy ℓ_i between the goal $\check{y}_i$ and prediction y_i for item i can be assessed independently of the other items $j \neq i$, it is customary to take the total amount of adjustments

$$L^y_\bullet : y \mapsto \ell_\bullet \coloneqq \sum \ell_i$$

as a measure of adequacy of the predicted result $y = Y\,x$, which is in turn the adequacy of x . Thus a solution of an inverse problem is usually defined as the minimizer of

$$L^x_\bullet : x \mapsto \ell_\bullet \coloneqq \sum \ell_i = \sum (L^y_i \circ Y_i)\,x$$

where '$\circ$' is for right-to-left function composition:

$$\hat{x} \coloneqq \arg\min_x L^x_\bullet\,x \ . \tag{2}$$

In statistical estimation $L^x_\bullet$ is often a negative log-likelihood function.

It is convenient to split L^y_i into an itemwise standardization function $Z_i : y_i \mapsto z_i$ and a standard loss function $L^z : z_i \mapsto \ell_i$ independent of item

i , as in the lower triangle of the diagram

$$\begin{tikzcd} x \arrow[r, mapsto, "Y_i"] \arrow[d, mapsto, dashed, "L_i^x"'] & y_i \arrow[d, mapsto, "Z_i"] \arrow[dl, mapsto, "L_i^y"'] & \text{measured} \\ \ell_i & z_i \arrow[l, mapsto, "L^z"] & \text{unit-free.} \end{tikzcd} \tag{3}$$

The standardization function performs nondimensionalization; x and y may come with measurement units while z and ℓ are unit-free. A standardization function Z_i is strictly monotone increasing and $Z_i(\check{y}_i) = 0$.

Since inverse problems arise in many applications, it is important that $L_\bullet^x$ be specified easily by those who are familiar with the problem but not necessarily versed in mathematics and programming, in such a way that the resulting inverse problems may be solved by resource-constrained devices.

This paper proposes a practical method to specify $L_\bullet^x$, which results in a specification of the standard loss function L^z and the standardization functions $Z = \begin{bmatrix} \cdots & Z_i & \cdots \end{bmatrix}$. The function L_i^y is specified via a utility table which is more intuitive than specifying the function directly in the form of a mathematical expression. Utility tables are preferred over the equivalent loss tables because intuitive meaning can be assigned to utility values while it seems difficult to do the same with loss values. The outline of the procedure for each item i is as follows:

Itemwise loss function specification

1. Let utility $U_i^y : y_i \mapsto u_i$ be the normalized likelihood that the decision maker accepts the outcome y_i given that all other items are fully satisfied.
2. Specify a table of outcomes $\{y_{ik}\}_{k=1 \cdots K}$ of y_i against their utility values $\{u_{ik}\}_{k=1 \cdots K}$.
3. Transform this $[y_{ik}, u_{ik}]$-table to a table of y_{ik} against z_{ik} so that the utility of z_{ik} equals u_{ik} , $U_i^y(z_{ik}) = u_{ik}$.
4. Interpolate and extrapolate the resulting $[y_{ik}, z_{ik}]$-table to produce a continuous standardization function Z_i .
5. Map L^z to L_i^x by (3).

3 Standard loss and utility functions

A typical case of inverse problem formulation is the weighted least squares in which

$$Z_i : y_i \mapsto z_i := (y_i - \mu_i)/\sigma_i \qquad 0 < \sigma_i \qquad \text{linear standardization} \tag{4}$$

with

$$L^z : z_i \mapsto \ell_i := z_i^2 \qquad\qquad \text{square loss} \tag{5}$$

where μ_i and $0 < \sigma_i$ are given constants. Another popular loss function is

$$L^z : z_i \mapsto \ell_i := |z_i| \qquad\qquad \text{absolute value loss} \tag{6}$$

the optimization of which reduces to linear programming if the causality function Y is linear. The loss function used in this paper is

$$L^z : z_i \mapsto \ell_i := \begin{cases} z_i^2 & |z_i| \le \theta \\ |\,2\theta\, z_i\,| - \theta^2 & \theta \le |\,z_i\,| \end{cases} \tag{7}$$

illustrated in Figure 2 which is the square loss near the origin but linear beyond $\pm\theta$ in the same way as the Huber loss function in robust parameter estimation. In addition, (7) has the benefit of requiring a relatively narrow

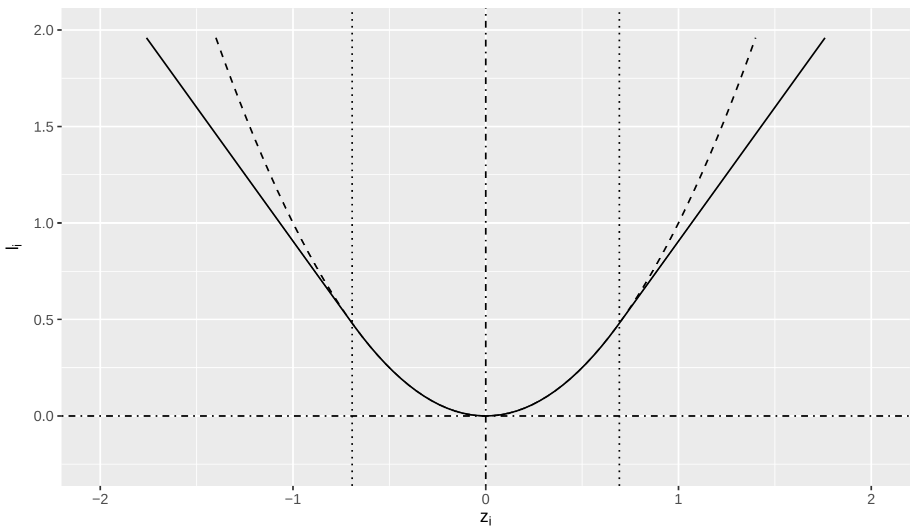

Figure 2: Standard loss function L^z
Dashed line: z_i^2 , $\theta = \log 2 = 0.693$

dynamic range for evaluation. The overall standard loss $L^z_\bullet := \sum L^z(z_i)$ looks as in Figures 3 and 4 if z is bidimensional.

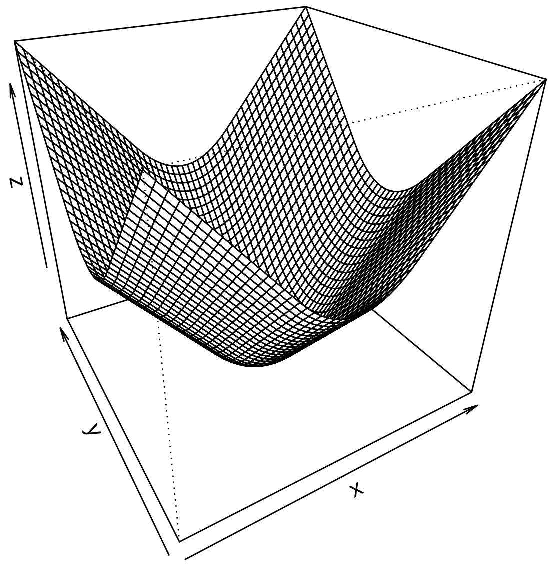

Figure 3: Bivariate standard loss function

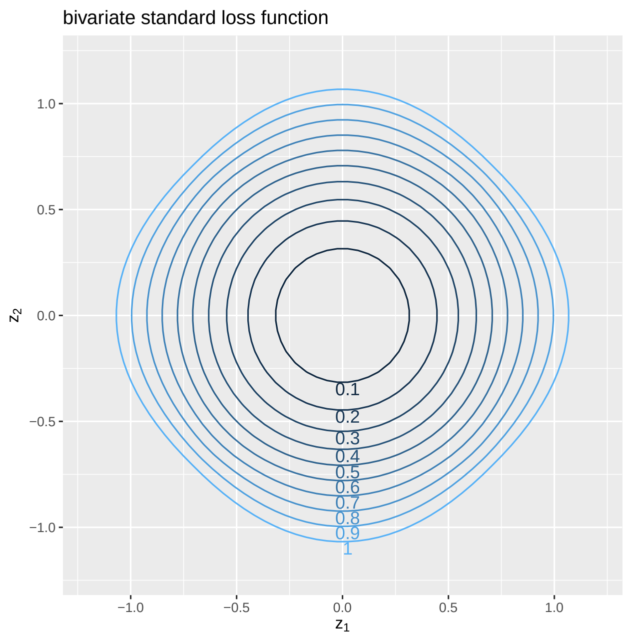

Figure 4: Bivariate standard loss function contours

Let

$$U^z : z_i \mapsto u_i := e^{-\ell_i} \qquad \text{standard utility function}$$

independently of item i, such that u_i fits within 0 and 1, and $\ell_i = 0$ corresponds to $u_i = 1$ as in Figure 5. This is the upside down loss function

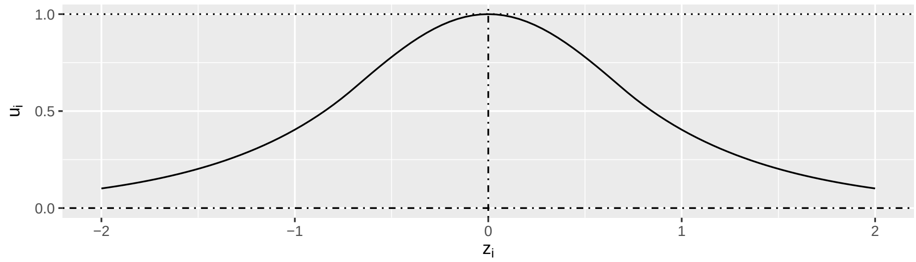

Figure 5: Standard utility function U^z
$\theta = \log 2 = 0.693$

compressed to $[0, 1]$ with higher compression rate towards the bottom. With

$$U^x_\bullet : x \mapsto u_\bullet := \prod u_i$$

minimization of the sum of loss is equivalent to maximization of the product of utility:

$$\hat{x} = \arg\min_x L^x_\bullet \, x = \arg\max_x U^x_\bullet \, x \ . \tag{8}$$

In statistical parameter estimation U^x is proportional to the the likelihood function and the right end of (8) is the maximum likelihood estimation. In the context of inverse problem it would be more appropriate to call U^x the standard acceptability function. For the sake of familiarity U^x is called the standard utility function in the remainder of this paper.

4 Rejectability

To formulate an inverse problem it remains to devise a method to define the standardization functions $Z = \begin{bmatrix} \cdots & Z_i & \cdots \end{bmatrix}$ in (3), which is to be more flexible than the linear standardization in (4). The first requirement for Z_i is that the corresponding values produce the same loss value: $Z : y_i \mapsto z_i \ \Rightarrow L^y(y_i) = L^z(z_i)$. Since $\ell_i \mapsto u_i$ has an inverse, this is equivalent to requiring that $U^y(y_i) = U^z(z_i)$. This condition presents two candidates $\pm z_i$ except when $u_i = 1$, which can be resolved by assuming the utility function to be unimodal.

A way to easily choose one from the two candidates is to define a new performance index which is bijective to z_i while carrying the utility information. Demanding the performance function to be continuous results in a candidate

$$R^z : z_i \mapsto r_i \coloneqq \operatorname{sign}(z_i)\left\{1 - U^z(z_i)\right\} \quad \text{standard rejectability function} \tag{9}$$

$$\operatorname{sign} z_i \coloneqq \begin{cases} -1 & z_i < 0 \\ 0 & 0 = z_i \\ +1 & 0 < z_i \end{cases}$$

illustrated in Figure 6. The function is symmetric with respect to the origin;

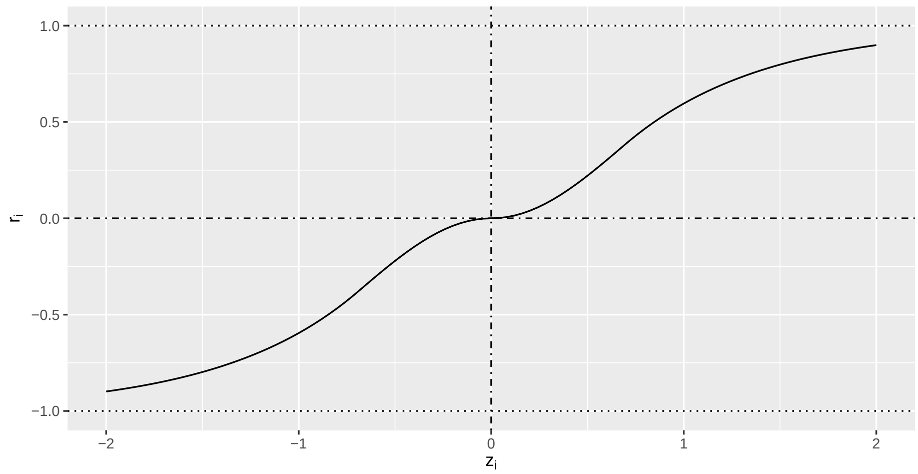

Figure 6: Standard rejectability function R^z
$\theta = \log 2 = 0.693$

the left half is the utility function shifted downwards by one, and the right half is the utility function upside down. Another interpretation is that the rejectability r_i is the standard variable z_i compressed to $[-1, 1]$ in a similar way as the utility u_i is the loss ℓ_i negated and compressed to $[0, 1]$.

The rejectability indicates how off-target the i-th item is in a scale of $[-1, 1]$, with $r_i = 0$ meaning on-target; negative and positive numbers meaning off to the left and to the right, respectively. An interpretation of rejectability is the normalized likelihood of rejecting the value y_i signed to indicate if the rejection is based on too little of y_i or on too much of y_i . An important difference between R^z and U^z is that while the former has a continuous inverse, the latter does not. Thus, given a monotone increasing rejectability function R^y, the standardization function Z_i can be determined uniquely.

With

$$V^z : z_i \mapsto v_i \coloneqq \operatorname{sign} z_i \cdot L^z(z_i) \qquad \text{signed loss function}$$

an explicit representation of the inverse of R^z is

$$(R^z)^{-1} : z_i \xleftarrow{(V_i^z)^{-1}} v_i \xleftarrow{(R_i^v)^{-1}} r_i$$

$$(R^v)^{-1} : r_i \mapsto v_i = \begin{cases} \operatorname{sign} r_i \cdot \{-\log(1 + r_i)\} & r_i \leq 0 \\ \operatorname{sign} r_i \cdot \{-\log(1 - r_i)\} & 0 \leq r_i \end{cases}$$

$$(V^z)^{-1} : v_i \mapsto z_i = \begin{cases} \operatorname{sign} v_i \cdot |\, v_i \,|^{\frac{1}{2}} & -\frac{1}{2} \leq v_i \leq \frac{1}{2} \\ \{(\log 2)^2 \operatorname{sign} v_i + v_i\}/(2 \log 2) & \text{otherwise.} \end{cases}$$

All functions $F_i^z : z_i \mapsto f_i$ can be mapped to functions $F_i^y : y_i \mapsto f_i$ by prepending the standardization function Z_i as in $F_i^y : y_i \overset{Z_i}{\mapsto} z_i \overset{F_i^z}{\mapsto} f_i$. Similarly, all functions of $G_i^y : y_i \mapsto g_i$ can be mapped to functions of vector x by prepending the causality function Y_i as in $G_i^x : x \overset{Y_i}{\mapsto} y_i \overset{G_i^y}{\mapsto} g_i$.

5 Standardization function specification

The role of itemwise standardization function Z is to convert y_i to z_i. This is needed because the loss function $L_\bullet^y$ has been split into the standardization function Z and the standard loss function $L_\bullet^z$. Similarly, the destandardization function Z^{-1} is needed to convert z_i to y_i because the causality function Y is given in terms of y rather than in z . A solution method for loss function typically involves iterations. Since conversions Z and Z^{-1} are invoked for every iteration, it is necessary that they be simple, stable, and fast.

The standardization function Z is equivalent to rejectability R_i^z, but not U_i^z or L_i^z . However, since rejectability is an unfamiliar concept to potential users, Z will be specified from utility data points. Conversion from utility to rejectability can be made unique by providing $\operatorname{sign} z_i$ as

$$\operatorname{sign} z_i = \begin{cases} -1 & y_i \text{ is on the left of the peak of } U_i^y \\ 0 & y_i \text{ is at the peak of } U_i^y \\ +1 & y_i \text{ is on the right of the peak of } U_i^y \end{cases} \tag{10}$$

This enables determination of Z_i given (y_i, u_i).

To define a standardization functions Z_i , assume that data points $\{[y_{ik}, u_{ik}]\}_{k=1\cdots K}$ are given with: $3 \leq K$; y_{ik} are strictly increasing $y_{ik} < y_{i\,k+1}$; utilities are positive $0 < u_{ik} \leq 1$; the table is unimodal; the peak is at $\check{k}_i$; $2 \leq \check{k}_i \leq K - 1$; and $u_{i\check{k}_i} = 1$.

The $[y_{ik}, u_{ik}]$-table can be converted to $[y_{ik}, r_{ik}]$-table: r_{ik} is negative if $k < \check{k}$, nonnegative otherwise. Converting r_{ik} into z_{ik} by $r_i \mapsto z_i$ produces $[y_{ik}, z_{ik}]$-table.

Interpolating and extrapolating the $[y_{ik}, z_{ik}]$-table linearly results in a piecewise linear standardization function. See Figure 7 for an example. Since Z is piecewise linear and not differentiable, so is $L_\bullet^y$, meaning that the optimization has to be carried out derivative-free.

A utility value u_{ik} is assigned as normalized likelihood that the decision maker accepts the corresponding outcome y_{ik} under the condition that all other outcomes are ideal: $u_{jk} = 1$ for all $j \neq i$.

6 Solution

Itemwise loss functions

This is the core part of the method which converts the utility specification Table 1 into itemwise loss functions L_i^y following the procedure outlined in Section 2.

The parameter in (7) is set to $\theta = 1/2$ since y_i values for normalized acceptance likelihood $u_i < 1/2$ seem to be of insignificant importance.

Interpolating and extrapolating the `work` columns as described in Section 5 results in a piecewise linear standardization function $Z_{\texttt{work}}$ illustrated in Figure 7. The corresponding rejectability function $R^y_{\texttt{work}}$ is $Z_{\texttt{work}}$ compressed vertically as illustrated in Figure 8. The utility function $U^y_{\texttt{work}}$ is $R^y_{\texttt{work}}$ twisted as in Figure 9. Finally, the loss function $L^y_{\texttt{work}}$ is obtained by flipping $U^y_{\texttt{work}}$ upside down and elongating vertically as in Figure 10. Similar functions for the other items are available in the same way.

With the causality the functions Y and the itemwise loss functions L_i^Y at hand, the inverse problem (1) can be solved by (2). The optimization result is summarized in Table 3. Its rejectability chart has been illustrated

Table 3: Solution

i	z_i	y_i	unit	r_i
`work`	0.319	8.946	[h]	0.097
`sleep`	-0.295	7.956	[h]	-0.083
`free`	-0.491	5.098	[h]	-0.214
`income`	-0.405	1.154	[cu]	-0.151

in Figure 1.

Utility as a function of `work` and `sleep` looks as in Figure 11, its contours or indifference curves are in Figure 12.

Since utility $U^x_\bullet$ is proportional to the likelihood of the position of x , normalizing Figure 12 by

$$N := \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{U}^x_\bullet(x)\, dx_{\texttt{work}}\, dx_{\texttt{sleep}}$$

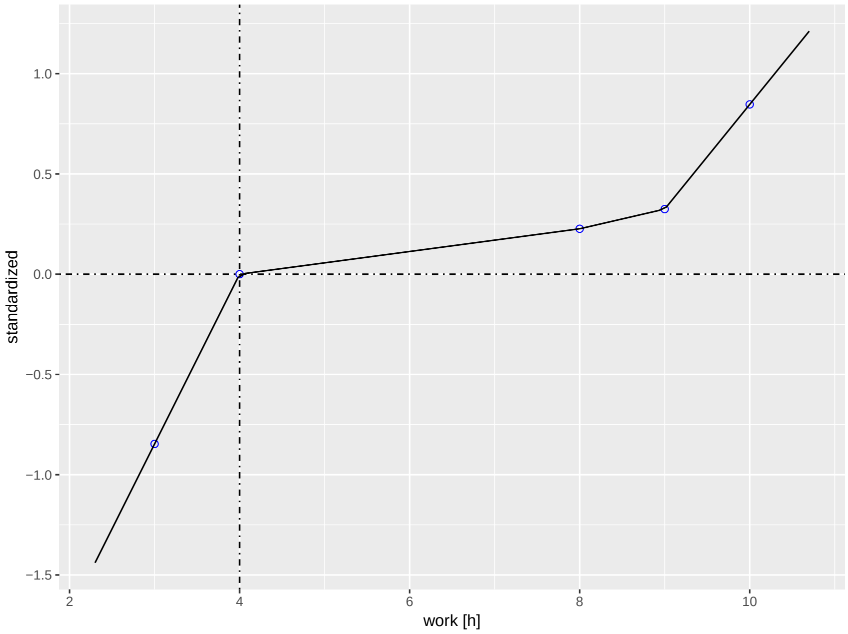

Figure 7: Standardization function $Z_{\texttt{work}}$

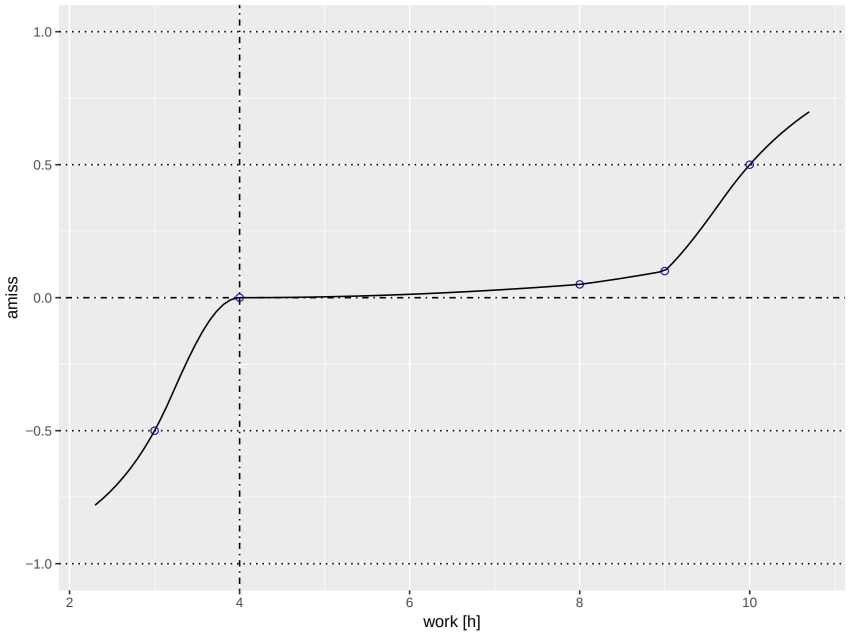

Figure 8: Rejectability function $R^{y}_{\texttt{work}}$

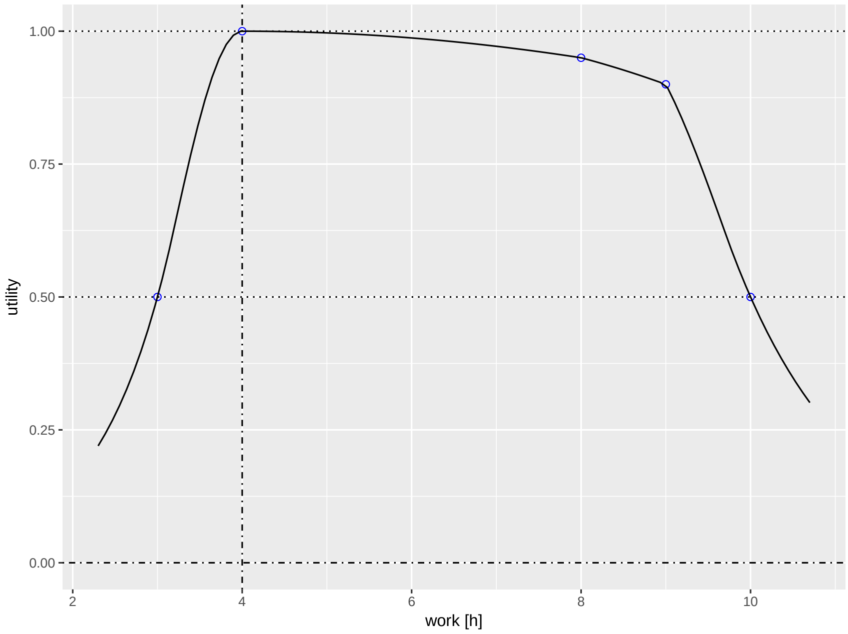

Figure 9: Utility function $U^{y}_{\mathtt{work}}$

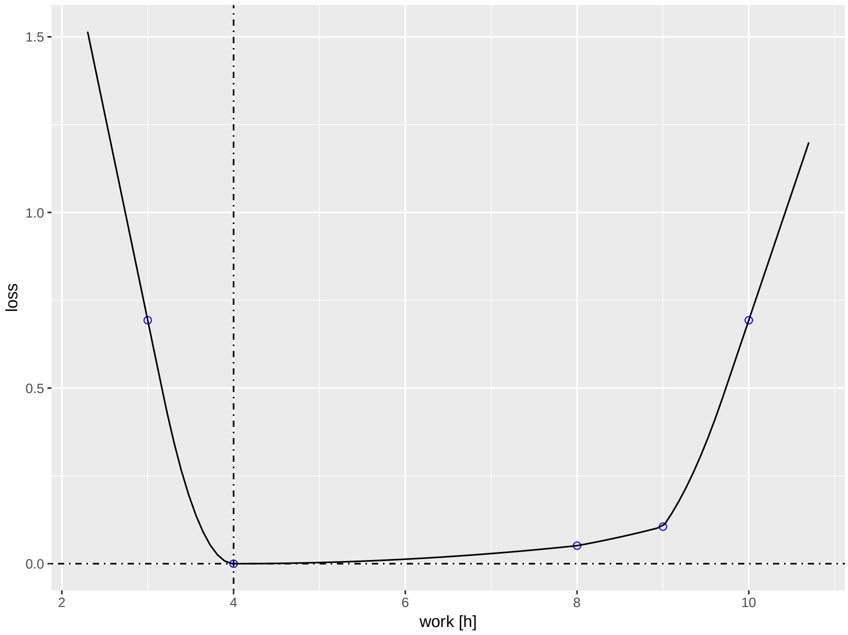

Figure 10: Loss function $L^{y}_{\mathtt{work}}$

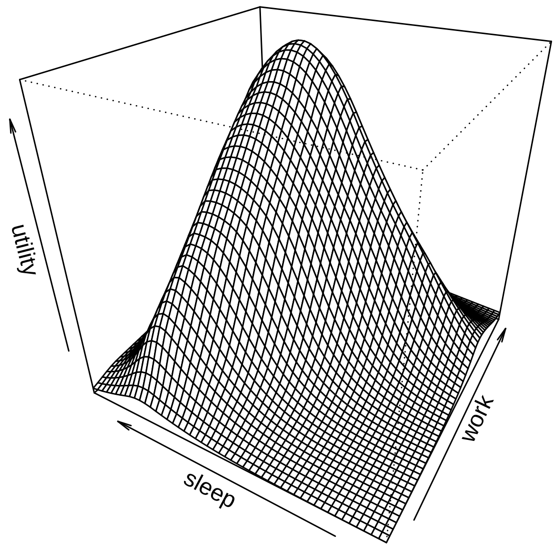

Figure 11: Utility function $U^x_\bullet$ for $x = [x_{\texttt{work}}, x_{\texttt{sleep}}]$

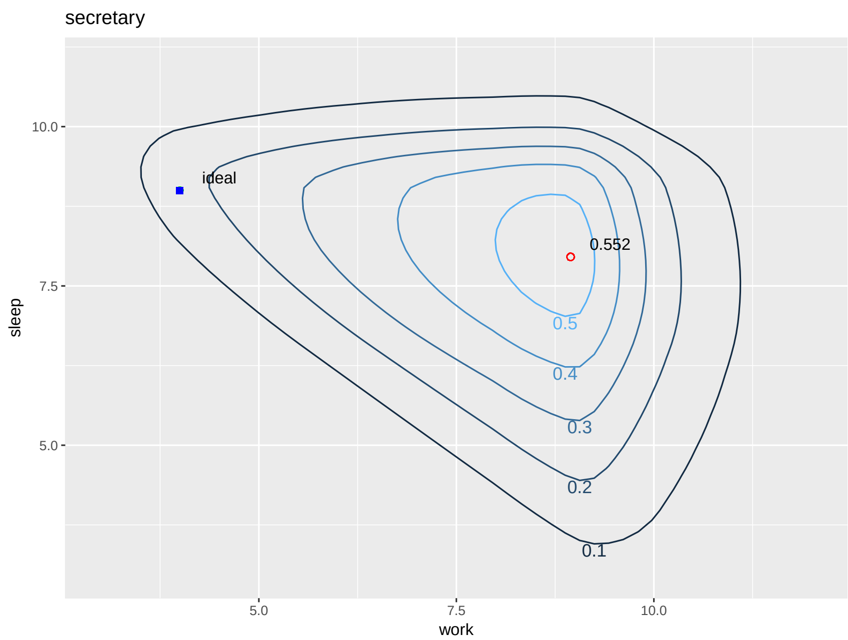

Figure 12: Indifference curves

yields the probability density of x , enabling probabilistic behavior prediction of the decision to be taken. For instance, the probability that the secretary will be working between 8 and 9 [h] is

$$\frac{1}{N}\int_{-\infty}^{\infty}\int_{8}^{9} \mathrm{U}^{x}_{\bullet}(x)\, dx_{\texttt{work}}\, dx_{\texttt{sleep}} = 0.23$$

7 Discussion

The way to fill the utility table as described in Section 5 is the crucial part of the present approach. The utility U^y_i of y_i is optimistically assumed to be measured in terms of acceptability independent of other items. Empirical studies would be awaited to find a better way. Since the overall utility $U^y_\bullet$ is multiplicative, all items i have to be essential, without luxury items.

The method to define a standardization function Z is an application of (Yoneda, Moretti and Poker, 2016). It is possible to adopt sophisticated interpolation methods such as monotone increasing splines enabling optimization methods based on differentiability for fast computation. However, that implies slower destandardization Z^{-1}, lower computational stability, and more complex code.

The major advantage of the present method over the usual least squares is that the piecewise linear standardization functions such as in Figure 7 is more expressive than the linear standardization (4).

While (5) needs a wide dynamic range to store its values since z^2 tends to be a large number, (7) requires a narrower dynamic range, of the same order as (6) so that in principle fixed point numbers rather than floating point numbers should suffice. This is relevant since the goal is to bring inverse problem formulation to resource-constrained devices. Also, the linear interpolation enables fast conversion between y and z . Additional research in this direction is available in (Yoneda, 2018).

A feature of the present formulation as compared to mathematical programming is in the absence of hard constraints, guaranteeing that a solution always exists. The rejectability function R^y plays a role similar to the Lagrangian multiplier or shadow price in constrained programming. The accompanying rejectabiliy chart as in Figure 8 has been found useful in debugging causality equations.

Since all items stand on an equal footing the formulation seeks a balance among them rather than favor an item in detriment of the others. Consider for instance the usual constrained optimization formulation of the diet problem which aims to minimize the cost while minimally satisfying the nutritional requirements. Aiming to hit a balance in various criteria such as nutrition, cost, taste, and other aspects seems at least more humane.

As has been pointed out in Section 5, the optimization (2) should be carried out derivative-free.

Appendix: Notation

A unified notation is set to avoid confusion. The notation for a function of the form $Y : x \mapsto y$ is Y^x, with Y the upper case y for the function value and the superscript x for the argument, which may be omitted when clear from the context. When subscripted Y_i^x, it is a component of a vector of functions $Y^x = \begin{bmatrix} \cdots & Y_i^x & \cdots \end{bmatrix}$, meaning that Y_i^x is short for $(Y_i)^{x_i} : x_i \mapsto y_i$ rather than $(Y_i)^x : x = \begin{bmatrix} \cdots & x_i & \cdots \end{bmatrix} \mapsto y_i$. The suffix i is omitted to Y^x when the function $x_i \mapsto y_i$ is independent of i . A vector-to-scalar function is written as $Y_\bullet^x : x = \begin{bmatrix} \cdots & x_i & \cdots \end{bmatrix} \mapsto y_\bullet$.

Generic

Y^x	Function $x = \begin{bmatrix} \cdots & x_j & \cdots \end{bmatrix} \mapsto y = \begin{bmatrix} \cdots & y_i & \cdots \end{bmatrix}$ or function $x_i \mapsto y_i$ when the function is independent of i .
Y_i^x	Function $(Y_i)^{x_i} : x_i \mapsto y_i$.
$Y_\bullet^x$	Function $x = \begin{bmatrix} \cdots & x_i & \cdots \end{bmatrix} \mapsto y_\bullet \in \mathbb{R}$.

Some specific variables and functions are listed below.

Variables

i	Outcome items, such as `work`, `sleep`, `free`, and `income`.
j	Decision items, such as `work` and `sleep`.
k	Utility specification table row entry number.
$\check{k}_i$	The entry number k such that $u_{ik} = 1$.
x	Vector of decision variables, $= \begin{bmatrix} \cdots & x_j & \cdots \end{bmatrix}$.
$\hat{x}$	Solution of the inverse problem, $= \arg\min_x L_\bullet^x \, x$.
y	Vector of outcomes, $= \begin{bmatrix} \cdots & y_i & \cdots \end{bmatrix} = \begin{bmatrix} \cdots & x_j & \cdots & y_i & \cdots \end{bmatrix}$, the first part of which is identical with x .
$\check{y}$	Vector of target outcomes, $= \begin{bmatrix} \cdots & \check{y}_i & \cdots \end{bmatrix}$.
z	Vector of standardized variables, $= \begin{bmatrix} \cdots & z_i & \cdots \end{bmatrix}$.
ℓ	Loss vector, $= \begin{bmatrix} \cdots & \ell_i & \cdots \end{bmatrix}$, $0 \leq \ell_i$.
$\ell_\bullet$	Loss, $= \sum \ell_i$, $0 \leq \ell_\bullet$.
r	Rejectability vector, $= \begin{bmatrix} \cdots & r_i & \cdots \end{bmatrix}$, $-1 \leq r_i \leq 1$.
u	Utility vector, $= \begin{bmatrix} \cdots & u_i & \cdots \end{bmatrix}$, $0 \leq u_i \leq 1$.

Functions

Y_i	Causality functions $Y_i : x \mapsto y_i$, assumed monotone with respect to each variable x_j .
L_i^y	Itemwise loss functions $y_i \mapsto \ell_i$.
$L_\bullet^y$	Loss function $y \mapsto \ell_\bullet$.
L_i^x	Itemwise Loss functions $x_i \mapsto \ell_i$.
$L_\bullet^x$	Loss function $x \mapsto \ell_\bullet$.
L^z	Standard loss function $z_i \mapsto \ell_i$, independent of i.
R^z	Itemwise standard rejectability function $z_i \mapsto r_i$, independent of i .
U_i^y	Itemwise utility function $y_i \mapsto z_i$.
U^z	Itemwise standard utility function $z_i \mapsto u_i$, independent of i .
Z_i	Standardization function $y_i \mapsto z_i$.

Acknowledgments

The author is grateful to Prof. Maria Helena Hunziker, Universidade de São Paulo, and Prof. João Claudio Todorov, Universidade de Brasília, for discussions from the viewpoint of behavioral psychology.

R (R Core Team, 2018) was used to prepare this paper.

References

(Herrnstein, *et al.*, 1997) Richard J. Herrnstein, Howard Rachlin, David I. Laibson; The Matching Law: Papers in Psychology and Economics; Harvard University Press, 1997.

(R Core Team, 2018) R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. `https://www.R-project.org/`

(Wichers, 1996) Robert Wichers; A Theory of Individual Behavior; Acdemic Press.

(Yoneda, Moretti and Poker, 2016) Kiyoshi Yoneda, Antonio Carlos Moretti, Johan Hendrik Poker, Jr.; Stretching the Least Squares to Embed Loss Functions Tables; Decision Making in Manufacturing and Services, Vol.9, No.2, 2016. `https://journals.agh.edu.pl/dmms/article/view/1207/1435`

(Yoneda, 2018) Kiyoshi Yoneda, Charles Childers; Encoding floating point numbers to shorter integers; submitted for publication.