A CONSTRAINED OPTIMIZATION PROBLEM UNDER UNCERTAINTY

ERIK QUAEGHEBEUR[†,‡], KEIVAN SHARIATMADAR[†],
and GERT DE COOMAN[†]

[†]*SYSTeMS Research Group, EESA Department, Ghent University*
Technologiepark-Zwijnaarde 914, Zwijnaarde, 9052, Belgium
E-mail: {Erik.Quaeghebeur,Keivan.Shariatmadar,Gert.deCooman}@UGent.be

[‡]*Department of Philosophy, Baker Hall 135, Carnegie Mellon University*
Pittsburgh, PA 15213-3890, United States
E-mail: ErikQ@andrew.cmu.edu

We investigate a constrained optimization problem for which there is uncertainty about a constraint parameter. Our aim is to reformulate it as a (constrained) optimization problem without uncertainty. This is done by recasting the original problem as a decision problem under uncertainty. We give results for a number of different types of uncertainty models—linear and vacuous previsions, and possibility distributions—and for two different optimality criteria for decision problems under uncertainty—maximinity and maximality.

Keywords: constrained optimization, maximinity, maximality, linear prevision, vacuous prevision, possibility distribution.

1. Introduction

Consider the following optimization problem: maximize a bounded real-valued function f—defined on a set $\mathcal{X}$—over all x in $\mathcal{X}$ that satisfy the constraint xRY, where Y is a random variable taking values in a set $\mathcal{Y}$ and R is a relation on $\mathcal{X} \times \mathcal{Y}$. The aim is to reduce this problem to a (constrained) optimization problem from which the uncertainties present in the description of the constraint are eliminated.

This optimization problem is ill-posed: it is underspecified, as there is no unique way of interpreting what is meant by maximizing a function over an uncertain domain; it might also be overspecified, as the constraint may not be satisfiable for some values Y may take. Therefore, in Sec. 2, we introduce some assumptions and reformulate the optimization problem as a well-posed decision problem: optimal solutions correspond to optimal decisions.

In Sec. 3, we investigate what results can be obtained for different types of uncertainty models for the random variable Y—linear[1] and vacuous[2,3] previsions, and possibility distributions[4]— and for two different optimality criteria[5] for decision problems—maximinity and maximality. We present general results for a number of model-criteria pairings. For illustration purposes, we include a running example in which $\mathcal{X} = \mathcal{Y} \coloneqq \mathbb{R}$ and $R \coloneqq \leq$.

We end in Sec. 4 with some conclusions.

Some notational conventions We always let x and z be elements of $\mathcal{X}$, y of $\mathcal{Y}$, $B \subseteq \mathcal{X}$, and $A \subseteq \mathcal{Y}$. We introduce the following sets:

$$\begin{aligned} xR &\coloneqq \{y \in \mathcal{Y} : xRy\}, \quad \overline{BR} \coloneqq \textstyle\bigcup_{x\in B} xR, \quad \underline{BR} \coloneqq \textstyle\bigcap_{x\in B} xR, \\ Ry &\coloneqq \{x \in \mathcal{X} : xRy\}, \quad \overline{RA} \coloneqq \textstyle\bigcup_{y\in A} Ry, \quad \underline{RA} \coloneqq \textstyle\bigcap_{y\in A} Ry. \end{aligned} \tag{1}$$

Also note that $\underline{BR} \subseteq \overline{BR}$, $\underline{RA} \subseteq \overline{RA}$. (And similarly for $\not R$.)

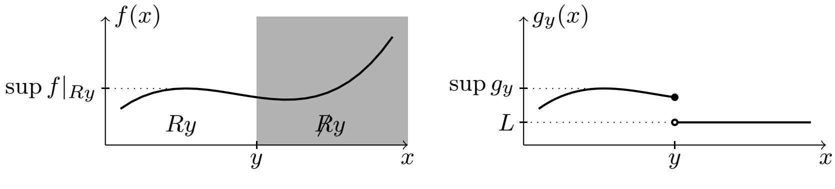
The indicator function of a set C is denoted I_C; it takes the value 1 on C and is 0 elsewhere.

2. Reformulation as a decision problem under uncertainty

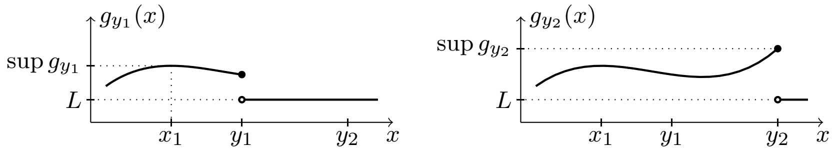
As we came to realize above, we need to decide which elements of $\mathcal{X}$ can be considered as optimal choices for the original optimization problem.

No uncertainty First consider the case without uncertainty about Y, where we know that Y takes some specific value y in $\mathcal{Y}$. We can then define an equivalent *unconstrained* optimization problem by maximizing the real-valued function g_y on $\mathcal{X}$ defined by $g_y \coloneqq fI_{Ry} + LI_{\not Ry} = L + f_L I_{Ry}$, where L is some real number strictly smaller than $\inf f$ and $f_L \coloneqq f - L > 0$. Because then, assuming $Ry \neq \emptyset$, $\sup f|_{Ry} = \sup g_y$, where $f|_{Ry}$ denotes the restriction of f to Ry; if $Ry = \emptyset$, we *also* use the unconstrained problem to replace the overspecified original one. We call L the penalty value, because it penalizes breaking the constraint.

For our running example, we get the following illustrative picture:



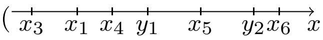
Indeterminacy Now consider the case where Y can be the value y_1 or y_2, and where nothing is specified about the relative likelihood of either value. Which $\mathcal{X}$-values should we now consider as optimal? The objective-function view of our running example does not seem to provide any intuition on how to decide in favor of some x:

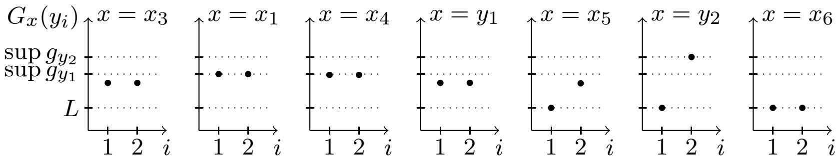


We need to look at each x as a potential optimal solution and compare them on the basis of their consequences. For this, we introduce a so-called utility function G_x on $\mathcal{Y}$ for each x (formally, $G_x(y)$ and $g_y(x)$ are identical):

$$G_x := f(x)I_{xR} + LI_{x\not R} = L + f_L(x)I_{xR}. \tag{2}$$

It returns the utility of choosing x for the different possible values of Y.

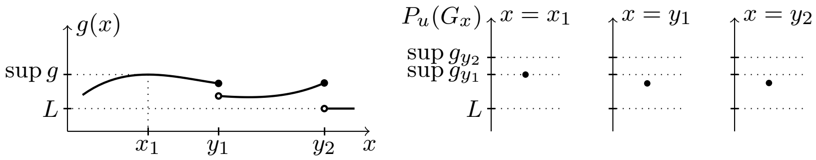
For a selection of $\mathcal{X}$-values (), this gives:



For pessimists, x_1 is a good optimum, because G_{x_1} has the highest minimal value. Optimists could call y_2 optimal, as G_{y_2} has the highest maximum. But we need not take any such extreme stance: pairwise comparisons of the utility functions show that x_1 and y_2 are the only x that could be considered as optimal in the sense that they have undominated utility functions, meaning that $G_z \not> G_x$ (pointwise) for all z.

So we see that there is no uniquely reasonable way of labeling an $\mathcal{X}$-value as optimal. The optimality criteria illustrated in the previous paragraph are respectively called maximinity, maximaxity, and maximality. The second is seldomly used because it can lead to risky decision making. We will be using maximinity and maximality; their formal definitions will follow after we have looked at the impact of probabilistic uncertainty about Y.

Probabilistic uncertainty Next consider the case where the possibilities y_1 and y_2 are additionally considered to be equally likely. In such a situation, one usually works with expected utility, *i.e.*, $g := P_u(G_\cdot)$ is used to find the optimal solutions, where P_u is the uniform prevision (linear expectation operator) on $\{y_1, y_2\}$, so $g \coloneqq \frac{1}{2}(g_{y_1} + g_{y_1})$. It turns out that both maximinity and maximality reduce to maximizing expected utility in this case. We get:



Optimality criteria When the uncertainty about Y is expressed using a more general uncertainty model, this reduction does not necessarily happen. In this paper, we consider uncertainty models that can be written as a coherent lower prevision $\underline{P}$, essentially a lower envelope of linear previsions (or expectation operators).[2,3] Its conjugate coherent upper prevision $\overline{P}$ is formally defined by $\overline{P} = -\underline{P}(-\cdot)$.[2,3]

The optimality criteria, exhibiting both the indeterminacy and expected utility aspects, are then expressed as follows:[5]

Maximinity The set of maximin solutions is

$$\operatorname{argsup}_{x\in\mathcal{X}} \underline{P}(G_x) = \operatorname{argsup}_{x\in\mathcal{X}} f_L(x)\underline{P}(xR), \tag{3}$$

where the right-hand side follows from Eq. (2), the irrelevance of additive constants, $f_L(x) > 0$, and the positive homogeneity[2,3] of $\underline{P}$.

Maximality A solution x is maximal if and only if

$$\inf_{z\in\mathcal{X}} \overline{P}(G_x - G_z) = \inf_{z\in\mathcal{X}} \overline{P}\big(f_L(x)I_{xR} - f_L(z)I_{zR}\big) \geq 0. \tag{4}$$

We need to compute the lower probability of events of the type xR and lower previsions of utility function differences $G_x - G_z$. Because of the sublinearity of coherent upper previsions,[2,3] all maximin solutions are maximal.

3. Formulation for specific uncertainty models

We now investigate a number of interesting special cases. For each case, both computing $\underline{P}(xR)$ and $\overline{P}(G_x - G_z)$, and reducing the resulting optimization problems (3) and (4) to a useful form are, in general, nontrivial steps.

Linear previsions When the uncertainty about Y is described by a linear prevision P, both criteria reduce to maximizing expected utility. The set of optimal solutions is $\operatorname{argsup}_{x\in\mathcal{X}} P(G_x) = \operatorname{argsup}_{x\in\mathcal{X}} f_L(x)P(xR)$. Note the influence of L.

For our running example, we see that $xR = [x, +\infty)$. Define the distribution function on $\mathcal{X}$ as $F_P \coloneqq P((-\infty, \cdot])$, then the set of optimal solutions for a linear prevision P with continuous F_P is $\operatorname{argsup} f_L(1 - F_P)$.

Vacuous previsions Vacuous previsions express ignorance. The general case consists of a vacuous prevision relative to an event $A \subseteq \mathcal{Y}$, for which $\underline{P} \coloneqq \inf \cdot|_A$ and $\overline{P} \coloneqq \sup \cdot|_A$.

For our running example, we let $A \coloneqq [a, b] \subset \mathbb{R}$.

For *maximinity*, we combine the vacuous prevision's definition with Eq. (3); the optimal x are those that maximize $\underline{P}(G_x) = L + f_L(x) \inf[I_{xR}|_A]$. So, by evaluating the expression $\inf[I_{xR}|_A] = \inf I_{xR \cap A}$, we discover that $\underline{P}(G_x) = f(x)$ if $A \subseteq xR$ and L otherwise. The expression $A \subseteq xR$ can be expanded to $(\forall y \in A) xRy$ and from this and Eq. (1), we can deduce it is equivalent to $x \in \underline{RA}$. So the set of optimal solutions is $\operatorname{argsup} f|_{\underline{RA}}$. It does not depend on L.

In our running example, $\underline{\leq[a, b]} = \bigcap_{y \in [a,b]} \{x \in \mathcal{X} : x \leq y\} = (-\infty, a]$ by Eq. (1), so the set of solutions is $\operatorname{argsup} f|_{(-\infty, a]}$.

For *maximality*, we combine the vacuous prevision's definition with Eq. (4); those x such that $\overline{P}(G_x - G_z) = \sup\bigl(f_L(x) I_{xR \cap A} - f_L(z) I_{zR \cap A}\bigr) \geq 0$ for all z are optimal.

An explicit expression for $\overline{P}(G_x - G_z)$ can be found by considering all possible positions A can be in relative to xR and zR. We find:

$$\begin{aligned} f_L(x) &\text{ if } A \cap xR \cap z\not{R} \neq \emptyset, \\ 0 &\text{ if } A \subseteq x\not{R} \wedge A \cap z\not{R} \neq \emptyset, \\ \max\{0, f(x) - f(z)\} &\text{ if } A \cap xR \cap z\not{R} = \emptyset \wedge A \cap xR \neq \emptyset \wedge A \cap z\not{R} \neq \emptyset, \\ f(x) - f(z) &\text{ if } A \subseteq zR \wedge A \cap xR \neq \emptyset, \\ -f_L(z) &\text{ if } A \subseteq x\not{R} \cap zR. \end{aligned}$$

In the first three cases, $\overline{P}(G_x - G_z)$ is always nonnegative, in the fourth it can be both positive and negative, and in the last it is always negative.

Therefore, only the last two cases are important when checking the condition for an x to be maximal, *i.e.*, to avoid its being nonmaximal. After some predicate logic manipulations, we find:

$$\inf_{z \in \mathcal{X}} \overline{P}(G_x - G_z) \geq 0 \quad \Leftrightarrow \quad \underline{RA} = \emptyset \vee \bigl(x \in \overline{RA} \wedge f(x) \geq \sup f|_{\underline{RA}}\bigr).$$

If $\underline{RA} = \emptyset$, all x in $\mathcal{X}$ are maximal; otherwise, only those x in $\overline{RA}$ such that $f(x) \geq \sup f|_{\underline{RA}}$ are. The set of maximal solutions does not depend on L.

For our running example, as $\underline{\leq[a, b]} = (-\infty, a]$ and $\overline{\leq[a, b]} = (-\infty, b]$ by Eq. (1), we see that those $x \leq b$ such that $f(x) \geq \sup f|_{(-\infty, a]}$ are maximal.

Possibility distributions The general case we consider here consists of a possibility distribution π on $\mathcal{Y}$, so for an event A, $\overline{P}(A) \coloneqq \sup \pi|_A$ and $\underline{P}(A) \coloneqq 1 - \sup \pi|_{A^c}$.[2,4,6]

For our running example, we consider a continuous possibility distribution π with minimal mode $c \in \mathbb{R}$; *i.e.*, $\pi(c) = 1$ and $\pi|_{<c} < 1$.

We only give results for *maximinity*. We combine the possibility distribution's definition with Eq. (3); those x that maximize $\underline{P}(G_x) = L + f_L(x)(1 - \sup \pi|_{x \not\mathcal{R}})$ are optimal. So the set of optimal solutions is $\operatorname{argsup}_{x \in \mathcal{X}} f_L(x)(1 - \sup \pi|_{x \not\mathcal{R}})$. Notice that this set depends on L.

For our running example, as $\sup \pi|_{x \not\leq} = \pi(\min\{x, c\})$, the set of solutions is $\operatorname{argsup} f_L(1 - \pi)|_{<c}$.

4. Conclusions

Using maximinity as an optimality criterion results in less complicated mathematical problems as compared to the maximality criterion. This is visible for the vacuous prevision case, and it is the reason we did not give a maximality solution for the possibility distribution case.

We encountered the same difference in complexity in preliminary investigations of p-boxes[7] and independent products[2] of a vacuous and linear previsions as uncertainty models. However, the maximinity criterion can also result in quite complicated expressions, as we encountered when working with linear-vacuous[2] previsions. Clearly, more work must be done in order to be able to deal with this complexity in practical situations.

Acknowledgments

Erik Quaeghebeur was supported by a Fellowship of the Belgian American Educational Foundation. This research was supported by the IWT SBO project 60043, "Fuzzy Finite Element Method".

Bibliography

1. B. de Finetti, *Theory of Probability* (Wiley, 1974–1975). 2 Volumes.
2. P. Walley, *Statistical Reasoning with Imprecise Probabilities* (Chapman and Hall, London, 1991).
3. E. Miranda, *Int. J. of Approximate Reasoning* **48**, 628 (2008).
4. D. Dubois and H. Prade, *Théorie des possibilités* (Masson, Paris, 1988).
5. M. C. M. Troffaes, *Int. J. of Approximate Reasoning* **45**, 17 (2007).
6. G. de Cooman, *Annals of Mathematics and Artificial Intelligence* **32**, 87 (2001).
7. E. Miranda, M. C. M. Troffaes and S. Destercke, Generalised p-boxes on totally ordered spaces, in *Soft Methods for Handling Variability and Imprecision*, eds. D. Dubois *et al.*, Advances in Soft Computing, Vol. 48 (Springer, 2008).