DECISION MAKING UNDER UNCERTAINTY: IS SENSITIVITY ANALYSIS OF ANY USE?

STEIN W. WALLACE

Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Trondheim, Norway
sww@iot.ntnu.no

Sensitivity analysis, combined with parametric optimization, is often presented as a way of checking if the solution of a deterministic linear program is reliable—even if some of the parameters are not fully known but are instead replaced by a best guess, often a sample mean. It is customary to claim that if the region over which a certain basis is optimal is large, one is fairly safe by using the solution of the linear program. If not, the parametric analysis will provide us with alternative solutions that can be tested. This way, sensitivity analysis is used to facilitate decision making under uncertainty by means of a deterministic tool, namely parametric linear programming. We show in this note that this basic idea of stability has little do with optimality of an optimization problem where the parameters are uncertain.

1. WHY SENSITIVITY ANALYSIS

Most, if not all, decisions are made under uncertainty; there is hardly any disagreement about that. Aspects such as price and demand, quality of raw materials, and reliability of machines and operators can hardly ever be viewed as deterministic entities. In addition, the model itself will almost always be an approximation of the real problem and hence represents uncertainty with respect to the quality of the resulting optimal solution.

However, it is important to remember that although all decisions can be viewed as being made under uncertainty, this does not imply that uncertainty is an important aspect of all problems. If, for example, the same decision is the unique optimum for absolutely all possible values of the uncertain parameters, although the objective function value may be very dependent, the true optimal decision can be found simply by solving one single problem, normally the one where all parameters are set at their most likely value. In such a case it is fair to claim that uncertainty is unimportant for making decisions.

An obvious difficulty with this approach is the need to know that it works without actually checking that it does. A common approach is therefore to solve *the expected value problem*—that is, the problem with all uncertain parameters replaced by their expected values—and then perform sensitivity analysis. This approach is taught in most textbooks on mathematical programming and operations research, although it is not common to interpret the linear program that is solved as an expected value problem. If uncertainty is in the objective function, the sensitivity analysis will tell over which area of the parameter set the given (primal) solution is

optimal. If uncertainty rests in the right-hand side, the analysis shows over which area the given basis is primal feasible, and therefore over which area the same primal variables remain nonnegative. In the first case, if the area is large, it is customary to say that the expected value solution (that is, the solution corresponding to the expected value problem) is stable, implying that it is reasonably safe to use it. In the second case, it is the optimal basis that is stable (and also the dual variables), but at least we know which variables (machines, processes, transportation modes) to focus on. Again, in standard texts, the terminology is usually different because the term expected value solution is not used. An advanced user would also consider the reduced costs at the boundaries to make sure the function is not too steep. Correspondingly, if the area of optimality is small, there is a fear that the given solution (or basis) is not very safe to use. In that case, we can use the parametric analysis to look for alternative solutions.

In a newly issued book on sensitivity analysis and parametric programming (Gal and Greenberg 1997), we read in the foreward: "Mathematical programming, especially linear programming and related network and combinatorial methods, usually form the OR/MS deterministic subfield. It is time to recognize that this categorization is restrictive and does not serve the field well. Those of us who work in the area are, in a sense, blessed and lucky. We have in the linear programming mathematical model and in its solution by the simplex method a readily available analysis that answers important data sensitivity questions and, at the same time, yields critical related economic information. Coupling such an analysis with computationally

simple studies provides a rather nondeterministic view of the modeling situation. Thus, those of us who teach and practice mathematical programming have the means of emphasizing and answering concerns about validity, robustness, uncertain data, base case and scenario analysis, and in achieving the truism that modeling is more about gaining insight than in producing numbers. We can and do cut across the dichotomy." We see that the usefulness of sensitivity analysis for understanding uncertainty is shared by the authors in their reference to robustness, uncertain data, and base case and scenario analysis.

As an example from a more typical text book, sensitivity analysis is introduced the following way in Ravindran et al. (1987, §2.11): "In all LP models the coefficients of the objective function and the constraints are supplied as input data or as parameters to the model. The optimal solution obtained by the simplex method is based on the values of these coefficients. In practice the values of these coefficients are seldom known with absolute certainty, because many of them are functions of some uncontrollable parameters. For instance, future demands, the cost of raw materials, or the cost of energy resources cannot be predicted with complete accuracy before the problem is solved. Hence the solution of a practical problem is not complete with the mere determination of the optimal solution. Each variation in the values of the data coefficients changes the LP problem, which may in turn affect the optimal solution found earlier. In order to develop an overall strategy to meet the various contingencies, one has to study how the optimal solution will change with changes in the input (data) coefficients. This is known as *sensitivity analysis* or *post-optimally analysis*."

What type of decision problems are we talking about? First, we see from the very wording that uncertainty is an element of the decision problem. Second, it seems clear that some decisions must be made before all the actual parameter values become known, and finally, that in some way it is costly if the decision fits badly to the realized parameter set. The problem therefore has at least two *stages*, that is, at least two points in time where costs (profits) are incurred, and these points are separated by another point in time where uncertainty is revealed. Typical examples would be production to meet an uncertain demand (possibly at uncertain prices), or the construction of a vehicle schedule without knowing the exact travel times and costs resulting from uncertain traffic patterns.

2. PRODUCING ALTERNATIVE SOLUTIONS

As long as the expected value solution is not optimal for all possible parameter sets, it is clear that in *hindsight*, that is after the fact, it may turn out that a different solution would have been better. This in itself is unavoidable and does not imply that the expected value solution is not the right solution.

If we accept the interpretation of the decision structure given in the previous section, it is quite clear that we have some problems with the objective function. What does it mean to minimize costs in a setting where costs occur twice, and the second occurrence is random? In fact, it means nothing unless we specify it more clearly. Very often it seems that the underlying assumption (implicit and hidden as it may be) is to minimize the sum of the immediate costs arising from the decision, and the expected future costs. We will proceed here as if that was the case. A few aspects of this issue will be discussed in the next section.

Many will argue that we have already gone too far, as they are not willing to operate with probability distributions at all. In fact, many approaches claim that one of their benefits is that distributions are not needed. And, of course, probability distributions are not discussed in textbooks on sensitivity analysis and parametric (linear) programming. The arguments made in this paper do not depend on either the willingness or the ability to estimate probability distributions. Furthermore, we do not imply that users are willing or able to solve the resulting complex optimization problem. However, we do assume that users would accept that probability distributions would make it possible to find better solutions *if* the distributions were available and all calculations came for free. In other words, we do not disregard practical arguments against probability distributions, but we do assume that terms like "expected future costs" make sense, even though they possibly cannot be calculated from a practical point of view.

Let us add that any approach that implicitly or explicitly accepts that we do in fact have a problem with at least two stages (simply meaning that some decisions are made before all relevant parameters are known with certainty) must make some kind of assumption about the relationship between the immediate costs and the future costs. There is no free lunch here.

So, again we conclude that although the expected value solution is not optimal (in hindsight) for all parameter sets, it may still minimize expected costs and hence be optimal. But, of course, there may be other solutions that have a better expected performances.

How would one go about finding alternative solutions? Again, if we refer to textbooks (for example, Ravindran 1987) as well as practice, there are a few basic approaches.

• Because it is likely that the solutions close to the expected value solution are good (after all, they are optimal for some parameters), we may use parametric linear programming to seek out candidates. It is sometimes the case that we can list *all* solutions that are optimal for some parameters. In the latter case we know for sure that in our collection we have the solution that will turn out to be optimal in hindsight.

• If the number of possible optimal solutions is too large (or if we are solving for example an integer program), we may sample parameters and solve the deterministic problem corresponding to each sample. This is sometimes referred to as *scenario analysis*. It is hoped that this will produce a representative set of possible optimal solutions, and therefore it will very likely also produce the solution that is optimal in hindsight. During the sampling we may, of course,

use any trick of the trade from sampling theory to decrease variance.

• We may carry out "what-if" sessions. Many software packages are set up for this. Note that there is no principal difference between a "what-if" session and sampling as defined in the previous item.

• Based on the candidate solutions obtained by any of the three previous methods, we may search for common features to capture properties that are in all or most of them. From this we construct some new candidate solutions that represent the essence of what we have found out from the analysis. In the continuous case we may take convex combinations of other solutions.

It is worth noting that we do not list simulation as a way of producing candidate solutions. A simulation model can be used to evaluate solutions, not to find them. In particular, a simulation model can say that strategy x is better than strategy y, but it cannot generally tell if x is good or bad.

3. PROBABILITY DISTRIBUTIONS

As mentioned, many methods make a point of not using probability distributions. The arguments are slightly varying. The most common argument is that in most situations it is very hard or even impossible to obtain such distributions, and then if we could, the resulting model would not be solvable. This is a valid argument, but it is important to realize that it is practical, not a principal, argument. Also, it is not always true, even though the cost of estimation and solution may be high. It would become a principal argument if one claimed that it was *incorrect* to use a probability distribution even if it was available, and the resulting model could be solved. This section will discuss some approaches that do not use probability distributions. Using these methods does not necessarily imply that one has a principal objection against distributions.

Economic theory tells us to maximize expected utility. In this text we have used expected profit as our objective function. This is equivalent to maximizing expected utility with a linear utility function. The arguments in this paper do not depend on the utility function being linear.

If we object to utility theory, then we need to replace this with another theoretical platform. Some claim that all they want is a "good" solution, but without saying what it is. This paper has little to say to those who just want "good" solutions in a context where optimal is worse than "good"; after all, the starting point is an optimization model with an objective function chosen by the user. In other words, we assume that even though "good" may indeed be good enough, optimal would be better.

In some cases, worst-case analysis is used. This is an approach that avoids probability distributions while accepting the two-stage setting. In itself, this can be valid; but true worst-case models are rare. Normally they are either terribly pessimistic (no goods will be demanded and the price will be zero; all machines will be down all the time) or

rather arbitrary because the worst case is impossible to define (demand will never be less than 23.2 and the price never below 16.6; the machines will never be down more than 25 minutes and 3 seconds per day). We may think of cases like bridge building and other cases where safety is a core property, but these are often not worst-case models. It is impossible to build a bridge that cannot fall down. Rather, one tries to build the bridge as cheaply as possible, given an acceptable probability that it collapses. In other words, these are models that use probability distributions (mostly the tails) in the constraints. In some cases, it is possible to formulate models that are good substitutes of the true worst case. A typical example would be the construction of a telecommunications network under the constraint that the network remains connected if any one arc fails. By *assuming* that only one arc can fail at a time, a meaningful model is achieved.

Another approach often used is to investigate the available candidate solutions as we indicated in item 4 above, and in that way to arrive at a solution. This solution is then accepted without simulation because simulation requires (at least implicitly) a distribution.

Finally, many will argue that because it is impossible to find the correct distribution form, they prefer to solve the deterministic program. This approach has at least two implicit assumptions. The first is that the distribution *form* is so important that if it is wrong, the results are not worth very much. The second is that if all probability mass is put in one point (that is, disregard all information about the parameters except their means, and hence use a distribution with a very extreme form), things are rather good. These two views are hardly consistent. Of course, these statements are not made, but they are implicit in the arguments.

4. WHY SENSITIVITY ANALYSIS/PARAMETRIC OPTIMIZATION MAY NOT DELIVER GOOD CANDIDATE SOLUTIONS

As we have seen, there are different ways of trying to find good overall solutions based on sensitivity analysis. Some use distributions, others do not. However, all but the worst-case analysis have one common aspect: *They build on the assumption that the optimal solution inherits properties from candidate solutions produced by parametric linear programming* (*sensitivity analysis*) *or sampling.* Geometrically, this means that the optimal solution is contained in the space spanned by the deterministic solutions. We shall now demonstrate that this basic assumption is generally false, and hence that all and any of the above approaches, except worst-case analysis, represent a cumbersome way of consciously looking for the wrong solution. The examples are made as simple as possible. For that reason it is easy to see what goes wrong, and one may be tempted to think that these errors would never be made. However, for more complex problems this visual overview is lost, and one is rather likely to make logical mistakes. Of course, even the simple structures below can be embedded in larger structures, making them hard or impossible to detect.

4.1. Example 1

Assume that we can produce three products, in the amounts x, y, and z. The demands for x and y are random between zero and one, and such that their demands add to one. This reflects that the products are substitutes; that is, they satisfy the same needs for the customers. We cannot sell anything that is not demanded. There is an upper limit of 1 on our production capacity. The model assumption here is that the decisions to produce x, y, and z are all made before the demand ξ becomes known. A linear programming model has been formulated and is as follows:

$$\max\ 3x + 2y + z$$

subject to:
$$
\begin{aligned}
x &\leqslant \xi, \\
y &\leqslant 1 - \xi, \\
x + y + z &\leqslant 1, \\
x, y, z &\geqslant 0.
\end{aligned}
$$

In this case it is possible to find the optimal solution for *all* values of ξ. It is the same optimal basis all the time. The optimal solutions are given by $x = t$, $y = 1 - t$, and $z = 0$ for all $\xi = t \in [0, 1]$. If we evaluate these solutions in terms of expected performance, we find that they are all infeasible (or alternatively that they all have an expected value of $-\infty$). This is easily seen by letting $x = \hat{t}$, and $y = 1 - \hat{t}$ for some $0 \leqslant \hat{t} \leqslant 1$. Then either $\hat{t} \leqslant \xi$ or $1 - \hat{t} \leqslant 1 - \xi$ will always be false, except when $\xi = \hat{t}$. Note that we do not need to know the probability distribution to conclude this way, as long as the support is known to be $[0, 1]$.

Rather, we would want the solution that maximizes the expected value of $3x + 2y + z$ (where expectation is taken over ξ, and where we remember that decisions must be made *before* ξ is known). We find that the optimal solution is $x = y = 0$ and $z = 1$, with a (deterministic) value of 1. We observe that all the candidate solutions had one thing in common—namely, $z = 0$—and that is *exactly what we do not want.* Another way to put this is to observe that although the same basis is optimal for *all* possible parameters, that is not enough to conclude that we have found the "correct" basis. So not even knowing that the same basis is always optimal in hindsight is enough to make useful conclusions.

It is possible to argue that the model in this example does not function well in the stochastic setting. That is indeed true because it is clearly a deterministic model, not made for stochastic analysis. It does not properly take care of what happens when the production does not match the demand; it simply declares that to be infeasible. However, rather than invalidating the example, that is just another aspect of what may go wrong when sensitivity analysis is applied to a deterministic model to analyze what is understood to be a stochastic decision situation. Models normally will have to be reformulated to make much sense. But note that the model makes full sense as long as it is used deterministically. However, it is not the purpose of this paper to discuss the modeling aspect of passing from deterministic to stochastic formulations.

4.2. Example 2

In the previous example, the uncertainty was in the constraints, and the model was inherently deterministic in its construction. Let us now present an example where the uncertainty is in the objective function, and where the model is explicitly set up for a two-stage analysis. Assume that we can produce a good at the presently unknown price p. To do this we must build a building of capacity $c \leqslant 1$, at a unit cost of 2. If this is to be done, it must be done now. Inside the building we must fit production equipment to facilitate the production. If we build the equipment now, it will come at a unit cost of 2, and its size z is constrained by the building size. We can also wait until after p has become known, but in that case the unit cost will increase to 2.2 because of increased delivery and installation costs. This extra capacity is denoted y, and it is also constrained by c, by the requirement $y + z \leqslant c$. Letting x be the production level, the problem now becomes

$$\max\ px - 2c - 2z - 2.2y$$

subject to:
$$
\begin{aligned}
x &\leqslant y + z, \\
y + z &\leqslant c, \\
c &\leqslant 1, \\
x, y, z, c &\geqslant 0.
\end{aligned}
$$

The two-stage interpretation is that c and z must be determined now. Hence, the goal is to sort out their values. Also here it is easy to find *all possible solutions* by parametric linear programming. If $p \leqslant 4$, the solution is $c = x = y = z = 0$ with a profit of 0. If $p \geqslant 4$, the solution is $c = x = z = 1$ and $y = 0$ with a profit of $p - 4$. The first solution obviously has an expected value of 0, and the second $E\{p\} - 4$. So if $E\{p\} \geqslant 4$, we should build the building and production facility now; if not, we should not nothing.

Only these two solutions are possible from parametric linear programming and sampling. The common feature here is that $y = 0$. This is so because in both cases $z = c$, forcing $y = 0$.

But there is another possible solution, namely $c = 1$ and $z = 0$. The problem facing us *after* p has become known will then be

$$\max\ px - 2.2y - 2$$

subject to:
$$
\begin{aligned}
x &\leqslant y, \\
y &\leqslant 1, \\
x, y &\geqslant 0.
\end{aligned}
$$

In this case, we will clearly do nothing if $p \leqslant 2.2$, but we will build one unit of production equipment and produce one unit of the good if $p > 2.2$.

To compare this solution with the candidate solutions from parametric linear programming above, we must make

some assumption on the distribution. Generally, the new solution is best if

$$\Pr\{p > 2.2\}[\mathrm{E}\{p \mid p > 2.2\} - 2.2] - 2$$

$$> \max\{\mathrm{E}\{p\} - 4, 0\}.$$

If, for example, we assume a uniform distribution over $[0, 9]$, the new solution has an expected value of 0.57, and the best from the parametric linear program an expected value of 0.5. So the new solution is 14% better.

4.3. Conclusion of Examples

What have we obtained? First, it takes only an example to show that the basic idea that the solution that maximizes expected profit can be found or approximated by investigating sampled deterministic solutions is false. In fact, it is more typical that the common features found in the candidate solutions from parametric optimization or sampling is exactly what we do *not* want. Why is this so?

All the problems implicitly solved by parametric optimization have one thing in common: They are deterministic. Therefore, the chance is that if the solutions have common features, these features reflect the deterministic property. This is not a general result, of course, but it is a very important observation. In example 1, we never produced any z because all the deterministic problems allowed $x + y = 1$. In example 2, the price is either high enough to allow production—in which case we build the building and the *cheap* production facility—or not high enough—in which case we do nothing. The common feature here is that the price is *known*. In such a case there is never a reason to pay 2.2 for something we can get for 2.

Geometrically, what we are observing is that the space spanned by all the possible deterministic solutions does *not* contain the solution that maximizes the expected value of the objective function. This is despite the fact that the solution which is optimal in hindsight is indeed in the space. So, clearly, the idea of looking for common features is generally not a valid approach. But are the examples above extreme and unusual, or are they typical? In other words, does it *normally* work to use parametric analysis and similar methods, even though there are cases where it is not right? This question cannot be answered in any precise way. However, we will end this section by listing two arguments for why the difficulty is more common than not. In addition, we must remember that even if there were a large class of problems where parametric optimization would produce the right candidate, we would need a way to tell if a given problem was in that class or not. Such a method does not currently exist.

• Decision problems usually contain some decisions that produce options; that is, they open up possibilities in the future that may or may not be used. Normally there is a cheaper version available if we buy immediately. In a deterministic optimization problem we will always either not do anything or will buy the cheap version. An example would be to build an oil platform with extra space, in case the field turns out to be larger than expected and we need more production equipment. It costs more to install this extra capacity later (often offshore) than to do it when the platform is under construction. A deterministic model will never suggest the expensive version but will either suggest a smaller platform without the extra production capacity or the larger platform with the capacity installed. It is very often these decisions that are the crucial decisions for this type of investments.

• Decisions that produce options as described above are often much less spectacular than in the example. It is simply that out of many possible ways of solving a problem, some are better for later periods than others. Deterministic methods will consistently pick the wrong ones unless flexibility comes for free. For example, if the goal is to construct a fleet of vehicles from a very large number of possible models, and the goods that are to be transported are unknown or the amounts are unknown, vehicles that are flexible with respect to type of goods will normally not show up in any analysis based on deterministically known quantities. It is not easy to realize that if one analysis shows we should buy three station wagons and the other that we need a medium-sized truck, we should in fact buy two pick-ups.

5. WHAT DOES SENSITIVITY/PARAMETRIC ANALYSIS DO?

After having seen that all approaches based on parametric analysis are principally false for analyzing decision making under uncertainty, we are still left with one question: What does sensitivity analysis do? Clearly, there is such a thing as stability of a solution (or basis) with respect to a parameter set.

In this article, we have discussed solely the issue of decision making under uncertainty. Assume that the parameter under investigation is not random. Rather, it is known, but we are considering changing it. For example, Norwegain authorities have decided that certain rivers must have a minimal flow so that some waterfalls can be enjoyed by tourists. Typically this water is lost for hydropower production. As the owner of the hydropower plant, one may ask: "What if the minimal required amount of water is reduced by 30%?" If the model is otherwise correctly defined with respect to uncertainty (the amount of available water is very uncertain), then sensitivity analysis is appropriate. This is a general rule: namely, that sensitivity analysis is appropriate for variations in *deterministic* parameters.

There is one situation related to uncertainty where sensitivity analysis is an appropriate tool. Assume that next year we will be making an important decision, and when it is being made all parameters will be known with certainty. Now, however, that the parameters are unknown, but even so we need numbers for our budget for the next year. If we now solve the expected value problem, and based on sensitivity analysis find it is very stable, we can be quite confident that the numbers we put into the budget are good.

Note the setting here. We are not making any decisions in face of uncertainty; we are simply predicting what will happen next year when we make a decision under *certainty*. This has nothing to do with decision making under uncertainty—and it is maybe not so surprising that a deterministic tool like sensitivity analysis does in fact analyze only deterministic decision problems?

ACKNOWLEDGMENTS

REFERENCES

Gal, T., H. J. Greenberg (eds.). 1997. *Advances in Sensitivity Analysis and Parametric Programming*. Kluwer Academic, Boston, MA.

Ravindran, A., D. T. Phillips, J.J. Solberg. 1987. *Operations Research—Principles and Practice*. John Wiley & Sons, New York.