Taguchi and Robust Optimization

Michael W. Trosset*

March 15, 1997

Abstract

This report is intended to facilitate dialogue between engineers and optimizers about the efficiency of Taguchi methods for robust design, especially in the context of design by computer simulation. Three approaches to robust design are described:

1. Robust optimization, i.e. specifying an objective function f and then minimizing a smoothed (robust) version of f by the methods of numerical optimization.
2. Taguchi's method of specifying the objective function as a certain signal-to-noise ratio, to be optimized by designing, performing and analyzing a single massive experiment.
3. Specifying an expected loss function f and then minimizing a cheap-to-compute surrogate objective function $\hat{f}$, to be obtained by designing and performing a single massive experiment.

Some relations between these approaches are noted and it is emphasized that only the first approach is capable of iteratively progressing toward a solution.

*Adjunct Associate Professor, Department of Computational & Applied Mathematics, Rice University, Houston, TX (e-mail: *trosset@caam.rice.edu*).

1 Introduction

The purpose of design engineering is to design better products. This description is intentionally vague, but it suffices to command the attention of the numerical optimizer. Why not design the *best* products? Why not automatize the design process and replace engineering guesswork with the powerful methods of numerical optimization? In fact, *design optimization* already has become a vitally important subject of industrial and academic research.

The challenges of design optimization are formidable. One of the most perplexing challenges is problem formulation—translating an engineer's various concerns into a concrete mathematical programming problem. This report focuses on one aspect of problem formulation. Specifically, we consider ways to accommodate the oft-encountered desire to find designs that continue to perform well when the design is perturbed, one of many objectives that is encompassed by the phrase *robust design*.

Whatever method the numerical optimizer proposes for *robust design optimization*, the engineer or manager sensitive to current trends is likely to inquire about its relation to the methods of Genichi Taguchi, the "father" of robust design. This report is intended to facilitate dialogue on this subject. In Section 2 we consider some formulations of the purely mathematical problem of *robust optimization*—of finding *neighborhoods* of points with small objective function values rather than singleton point minimizers. These are the formulations of robust design that are likely to seem most natural to a numerical optimizer. In Section 3 we summarize Taguchi's ideas about robust design. In Section 4 we describe an alternative approach to robust design that specializes to one of the formulations of robust optimization in Section 2. In Section 5 we conclude with some thoughts on the importance of numerical optimization in robust design.

2 Robust Optimization

Let $f : \Re^p \to \Re$ be a real-valued function. By "robust optimization" we mean the problem of finding $x_* \in \Re^p$ such that (i) $f(x_*)$ is small, and (ii) x near x_* entails $f(x)$ small. By (ii) we mean more than just the local behavior that any continuous f will necessarily exhibit. Robust optimization is concerned with semi-local behavior and our goal is to avoid simply choosing x_* to be a local minimizer of f if that minimizer lies at the bottom of a very steep and narrow basin.

Robust optimization is inherently multi-objective. One way of operationalizing our objectives is to replace $f(x)$ with some measure of how f behaves near x—say on some $B(x;r)$, the ball of radius r centered at x. (More general approaches immediately suggest themselves, but this one will suffice for the purposes of this report.) A very conservative possibility is then to measure the performance of f on $B(x;r)$ by its supremum, resulting in the robust optimization problem

$$\min_{x \in \Re^p} \max_{\xi \in B(x;r)} f(x+\xi). \tag{1}$$

This formulation protects against worst-case scenarios and is analogous to the minimax principle in statistical decision theory and elsewhere.

Minimax optimization problems have been the subject of considerable study, e.g. by Dem'yanov and Malozemov (1974), but they are not smooth and are likely too conservative for the present application. A more liberal possibility is to measure the performance of f on $B(x;r)$ by some mean value, so that our objective becomes that of minimizing some sort of moving average of f. This can be accomplished by defining a weighting function (presumably a probability density function) $w : B(0;r) \to \Re$, resulting in the robust optimization problem

$$\min_{x \in \Re^p} \int_{B(x;r)} f(x+\xi) w(\xi) d\xi. \tag{2}$$

This formulation addresses typical (as defined by w) scenarios and is analogous to the Bayes principle in statistical decision theory.

Notice that the restriction to $B(x;r)$ is superfluous in Problem (2), as it can be incorporated into the specification of a weighting function on $\Re^p$. Furthermore, it is immaterial whether we define the weighting function on $\xi \in \Re^p$ or on $-\xi \in \Re^p$. Hence, the objective function in Problem (2) can be written as

$$F(x) = \int_{\Re^p} f(x-\xi) w(\xi) d\xi = [f * w](x). \tag{3}$$

The problem of minimizing F arises in some approaches to global optimization in which one endeavors to smooth the objective function in order to avoid becoming trapped prematurely by a local minimizer. In such methods, however, the weighting function is indexed by a continuation parameter h in such a way that $f * w_h \to f$ as $h \to 0$. What distinguishes robust optimization is that we are *not* interested in eventually recovering the original objective function.

Both Problems (1) and (2) were posed by Das (1996), who preferred the latter because of its smoothness. Das reported that this problem has been studied in the linear, but not the nonlinear programming literature. The major technical difficulty is evidently the evaluation of the p-dimensional integrals that appear in the objective function—especially in light of the fact that p is often quite large in engineering design problems and evaluating f may be quite expensive. As described in Davis and Rabinowitz (1984) and Flournoy and Tsutakawa (1991), various quadrature and Monte Carlo methods exist for multidimensional integration. Unfortunately, the curse of dimensionality guarantees that even highly efficient methods require evaluating the integrand at many points. Accordingly, Das proposed an approximate objective function that is derived by replacing $f(x+\xi)$ with its second-order Taylor series expansion at x. Another possibility might be to use Laplace's method, which tends to work well if f is smooth and w is (approximately) normal. This method entails writing the integrand as $\exp(\log(f(x+\xi)w(\xi)))$, then replacing $\log(f(x+\xi)w(\xi))$ with its second-order Taylor series expansion at its maximizer. The efficacy of this approach obviously depends on how easily one can determine the maximizer about which to expand; however, Laplace's method has proven itself a useful device for calculating posterior expectations in Bayesian statistics, where it has been studied by Tierney and Kadane (1986); Kass, Tierney and Kadane (1989); and Wong and Li (1992).

3 Taguchi and Robust Design

We now turn to the ideas of Genichi Taguchi. Taguchi's contributions to quality engineering have been far-ranging and we do not attempt a comprehensive survey in this report. The reader interested in a broader context and a more detailed discussion should turn to the literature on quality control, e.g. Roy (1990), which is the primary source of the material in this section.

Historically, quality engineering relied on product inspection. This remains an important aspect of quality control, but it is evidently a passive approach to developing new and better products. Decades ago, Taguchi proposed that engineers actively design quality into their products, an approach sometimes called *off-line quality control.* Thus, Taguchi was an early proponent of design optimization.

Taguchi envisioned a three-stage process for design optimization. The purpose of the first stage, *systems design*, is to determine the feasible region for the subsequent optimization problem. The purpose of the second stage, *parameter design*, is to optimize the objective function that operationalizes one's notion of quality. This is the stage of particular relevance to the present report. The purpose of the third stage, *tolerance design*, is to fine-tune the (approximately) optimal design obtained in the second stage.

Taguchi had very specific ideas about what objective functions to optimize for parameter design. Perhaps his most fundamental contribution to quality engineering was his observation that one should design a product in such a way as to make its performance insensitive to variation in variables beyond the designer's control. In acknowledgement of the value of this observation, Taguchi is widely regarded as the father of robust design and parameter design is sometimes called robust design.

We now describe some details of Taguchi's strategy for robust design. Taguchi's methods distinguish two types of inputs to a system: "control parameters" (or "control factors") are the inputs that can be easily controlled or manipulated by the designer, hence the inputs that constitute the optimization variables x; "noise variables" (or "noise factors") are the inputs that are difficult or expensive to control, hence the inputs ξ to whose variation product performance is desired to be insensitive. For example, x might specify the design of (a part of) a photocopier and ξ might specify the environment (temperature, humidity, etc.) in which it is to operate.

To obtain Taguchi's objective function(s), let y denote the quality characteristic of interest. This characteristic may be one of three types, for each of which Taguchi defined a measure of mean squared deviation (MSD). Let $\{y_1, \ldots, y_n\}$ denote a sample obtained by varying ξ for a fixed x. If quality is measured by how close y comes to a target value m, then

$$MSD = \frac{1}{n} \sum_{i=1}^{n} (y_i - m)^2;$$

if quality is measured by how small y is, then

$$MSD = \frac{1}{n} \sum_{i=1}^{n} y_i^2;$$

if quality is measured by how large y is, then

$$MSD = \frac{1}{n} \sum_{i=1}^{n} y_i^{-2}.$$

The objective function to be minimized by varying x is then the *signal-to-noise ratio* (SNR), $-10\log_{10}(MSD)$. Strong emphasis of these SNRs as the objective functions that operationalize the concerns of robust design is a hallmark feature of Taguchi's methods.

Taguchi's approach to optimizing the SNR was inspired by principles of classical experimental design. The control parameters x are systematically varied according to an orthogonal array, the "control array" or "inner array." At each value of x, the noise variables are systematically varied according to a second orthogonal array, the "noise array" or "outer array." At each value of x, data from the "replications" of the quality characteristic y across the noise array are used to estimate the SNR. One obtains an array of estimated SNRs, which is then analyzed by standard analysis of variance techniques to identify values of x that produce robust performance.

At this point it is crucial to note that, while Taguchi's contributions to the *philosophy* of robust design are almost unanimously considered of fundamental importance, the efficacy of his technical methods for implementing his philosophy is extremely controversial. An excellent survey of these controversies is the panel discussion edited by Nair (1992); in this report we focus on one very specific objection that is particularly damning from the perspective of numerical optimization.

Whatever the details of the SNRs, the control arrays, the noise arrays, the analyses of variance, etc., it is inescapable that Taguchi's methods attempt to optimize an objective function by specifying all of the values of x at which the objective function will be evaluated *prior to observing any function values.* (Data analysis is sometimes supplemented by performing one or more confirmatory experiments, but this is not a fundamental part of the optimization strategy.) Thus, the Taguchi approach violates a fundamental tenet of numerical optimization—that one should avoid doing too much work until one nears a solution. In Taguchi's defense, it should be noted that he was primarily concerned with situations in which sequential experimentation may not be possible, as when an entire manufacturing facility must be dedicated to performing the experiment. In modern engineering design optimization, however, performance is often assessed by computer simulation and the logistic necessity of one-shot experiments disappears.

4 Another Approach to Robust Design

In contrast to Taguchi's emphasis on modeling SNRs, various researchers have advocated directly modeling the response as a function of *both* control and noise factors. The approach that we describe was proposed by Welch, Yu, Kang and Sacks (1990).

Suppose that we measure q quality characteristics, $y_1, \ldots, y_q$. Let $y_i(x, \xi)$ denote the value of the ith quality characteristic when the control and noise factors assume values (x, ξ) and let $l[y_1(x, \xi), \ldots, y_q(x, \xi)]$ denote the loss that accrues from the qualities attained at those values. In the spirit of robust design, we seek a setting of control factors that minimizes the expected loss, computed with respect to the random vector ξ. If the distribution of ξ does not depend on x, then we obtain the objective function

$$L(x) = \int l[y_1(x, \xi), \ldots, y_q(x, \xi)] w(\xi) d\xi, \tag{4}$$

where $w(\xi)$ denotes the probability density function of ξ.

Now consider the case that the loss from design x is $f(x)$ and noise $-\xi$ enters the system by virtue of the impossibility of exactly manufacturing the specified design. Then the design that is actually manufactured is $x - \xi$ and the objective function (4) becomes (3),

$$F(x) = \int_{\Re^p} f(x - \xi) w(\xi) d\xi = [f * w](x).$$

Thus, the robust optimization problem (2) is actually a special case of an approach to robust design that has already received some attention in the statistics literature. That approach, however, is an *alternative* to Taguchi's methods.

Finally, we inquire how statisticians have sought to minimize (4). To a numerical optimizer, the answer may seem more significant for its similarities to Taguchi's methods than for its dissimilarities. First, a design is chosen that specifies the (x_j, ξ_j) at which the y_i are to be evaluated. Instead of inner and outer arrays, this approach results in a single "combined" array. Sacks and Welch (Nair 1992) commented that "the single experimental array for both control and noise factors will usually require far fewer observations than Taguchi's crossed arrays, even when interactions between the control factors are included."

Second, the $y_i(x_j, \xi_j)$ are used to construct cheap-to-compute surrogate models $\hat{y}_i$. These may be regression models, as in Welch, Yu, Kang, and

Sacks (1990), or spatial statistical models for "computer experiments," as in Welch and Sacks (1991).

Third, optimization is carried out using the surrogate objective function

$$\hat{L}(x) = \int l[\hat{y}_1(x,\xi), \ldots, \hat{y}_q(x,\xi)]w(\xi)d\xi.$$

Thus, just as with Taguchi's methods, one relies on a one-shot experiment: all evaluations of the actual objective function are made before any optimization commences.

5 Discussion

We have emphasized the distinction between a one-shot experiment and an iterative method for numerical optimization. The emphasis in the statistics literature on designed experiments is partly cultural, but in some cases it also derives from concerns about the cost of function evaluation. The following remarks by Sacks and Welch (Nair 1992) are instructive:

> "Taguchi used a deterministic mathematical model to generate data in his Wheatstone-bridge example (Taguchi 1986, chap. 6). There, the mathematical equation is trivial, so it is not clear why one would not simply plug the objective function into a numerical optimizer. Admittedly, this might produce a local optimum, but Taguchi's solution is also suboptimal (Box and Fung 1986). Real computer models can be computationally very expensive, and direct numerical optimization...can require too many function evaluations. Similarly, Taguchi's experimental plan with 1,296 observations is too expensive for real applications. Thus, Taguchi's approach appears (to us anyway) to be overcomplicated for simple deterministic models and too expensive for realistic problems. The strategy we outline has successfully tackled problems much more complex than the Wheatstone-bridge example with far fewer observations."

Mathematically, we can express the cost of function evaluation by imposing a budget—an upper bound V on the total number of function evaluations that will be permitted. If V is sufficiently "large" (one's notion of large or small will depend on various things, e.g. the number of variables, the availability of derivative information, etc.), then presumably one would use traditional iterative methods for numerical optimization. However, if V is sufficiently "small," then designed experiments may make sense. For example, if one is allotted just $V = 5$ function evaluations with which to minimize an objective function in $p = 2$ variables without access to derivative information, then it would hardly be irrational to design an experiment in which the function is evaluated at 4 points, fit a model to the data obtained from this experiment, minimize the model function, and use the final function evaluation to validate the model minimizer.

It is our impression, however, that many engineering applications have an intermediate value of V. For these optimization problems, the challenge is to synthesize ideas from numerical optimization with ideas from designed experiments and surrogate modeling. Indeed, the author is presently participating in a collaboration between The Boeing Company, IBM and Rice University that is dedicated to this challenge. A very general *model management* framework for exploiting surrogate model information for the purpose of optimizing expensive functions has already been proposed by Dennis and Torczon (1996). Recent work by Trosset and Torczon (1997) specializes this framework to models obtained by kriging, the most popular modeling technique in the literature on computer experiments.

Acknowledgements

The subject of this report was brought to my attention by Indraneel Das and John Dennis. This research was supported in part by AFOSR grant F49620-95-1-0210.

References

Box, G. E. P. and Fung, C. A. (1986). Studies in quality improvement: Minimizing transmitted variation by parameter design. Study 8, Center for Quality and Productivity Improvement, University of Wisconsin, Madison, WI.

Das, I. (1996). Multicriteria optimization and robust optimality. Thesis proposal, Department of Computational & Applied Mathematics, Rice University, Houston, TX.

Davis, P. J. and Rabinowitz, P. (1984). *Methods of Numerical Integration.* Academic Press, Orlando, FL, second edition.

Dem'yanov, V. F. and Malozemov, V. N. (1974). *Introduction to Minimax*. Dover Publications, New York.

Dennis, J. E. and Torczon, V. (1996). Managing approximation models in optimization. In Alexandrov, N. and Hussaini, M. Y., editors, *Multidisciplinary Design Optimization: State of the Art*. SIAM, Philadelphia.

Flournoy, N. and Tsutakawa, R. K. (1991). *Statistical Multiple Integration*. American Mathematical Society, Providence, RI.

Kass, R. E., Tierney, L., and Kadane, J. B. (1989). Approximate methods for assessing influence and sensitivity in Bayesian analysis. *Biometrika*, 76:663–674.

Nair, V. N. (1992). Taguchi's parameter design: A panel discussion. *Technometrics*, 34:127–161.

Roy, R. K. (1990). *A Primer on the Taguchi Method*. Competitive Manufacturing. Van Nostrand Reinhold, New York.

Tierney, L. and Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81:82–86.

Trosset, M. W. and Torczon, V. (1997). Numerical optimization using computer experiments. Technical Report 97-02, Department of Computational & Applied Mathematics, Rice University, Houston, TX. Submitted for publication.

Welch, W. J. and Sacks, J. (1991). A system for quality improvement via computer experiments. *Communications in Statistics—Theory and Methods*, 20:477–495.

Welch, W. J., Yu, T. K., Kang, S. M., and Sacks, J. (1990). Computer experiments for quality control by parameter design. *Journal of Quality Technology*, 22:15–22.

Wong, W. H. and Li, B. (1992). Laplace expansion for posterior densities of nonlinear functions of parameters. *Biometrika*, 79:393–398.