CHALMERS | UNIVERSITY OF GOTHENBURG

PREPRINT 2009:31

Robust Multi-objective Optimization Based on a User Perspective

CHRISTOFFER CROMVIK
PETER LINDROTH

Department of Mathematical Sciences
Division of Mathematics
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Göteborg Sweden 2009

Preprint 2009:31

Robust multi-objective optimization based on a user perspective

Christoffer Cromvik, Peter Lindroth

Department of Mathematical Sciences
Division of Mathematics
Chalmers University of Technology and University of Gothenburg
SE-412 96 Göteborg, Sweden
Göteborg, August 2009

Preprint 2009:31
ISSN 1652-9715

Matematiska vetenskaper
Göteborg 2009

Robust multi-objective optimization based on a user perspective

Christoffer Cromvik* Peter Lindroth†

Abstract

Solving practical optimization problems that are sensitive to small changes in the variables or model parameters require special attention regarding the robustness of solutions. We present a new definition of robustness for multi-objective optimization problems. The definition is based on an approximation of the underlying utility function for a single decision maker. We further demonstrate an efficient computational procedure to evaluate robustness. This procedure is applied to two numerical examples: one is an analytic test problem while one is a real-world problem in antenna design. The results show that the robustness varies over the Pareto front and that it can be improved if the decision maker is willing to sacrifice some optimality.

Keywords: Multi-objective optimization, Robustness, Multi-criteria decision making

1 Introduction

Many applications of optimization comprises several more or less conflicting objectives, such as cost/quality, expected return/risk etc. These are to be optimized simultaneously and the aim is to find the most appropriate balance between all the objectives. Mathematically, such a problem is denoted a *multi-objective optimization problem* (MOOP) and is formulated as that to

$$\underset{\mathbf{x}\in X}{\text{minimize}} \ (f_1(\mathbf{x}), \dots, f_k(\mathbf{x})) \,. \tag{1}$$

Here, $\mathbf{x} \in \mathbb{R}^n$ denotes a vector of decision variables, $X \subseteq \mathbb{R}^n$ is the feasible decision space, and each $f_i : X \to \mathbb{R}, \ i = 1, \dots, k$, is an objective function to be minimized. Since minimization of a vector in general is not well defined, the notion of optimality for multi-objective problems is somewhat different compared to single-objective problems. Optimality is here based on dominance, and the following definition is used.

Definition 1.1 (Pareto Optimality) *A feasible solution* $\hat{\mathbf{x}} \in X$ *is called* Pareto optimal *if there exists* no *vector* $\mathbf{x} \in X$ *such that* $f_i(\mathbf{x}) \le f_i(\hat{\mathbf{x}}), \ i = 1, \dots, k$, *with at least one inequality holding strictly. The set of all Pareto optimal solutions is denoted* $\mathcal{P} \subseteq X$.

The possibly most intuitive method for solving a MOOP, i.e., to find $\mathcal{P}$ or at least a good approximation of $\mathcal{P}$, is to solve a sequence of standard optimization problems of the following type

$$\underset{\mathbf{x}\in X}{\text{minimize}} \ \sum_{i=1}^{k} w_i f_i(\mathbf{x}), \tag{2}$$

*Corresponding author. Department of Mathematical Sciences, Chalmers University of Technology, and Department of Mathematical Sciences, University of Gothenburg, SE-412 96 Gothenburg, Sweden. Tel.: +46-31-7723515, Fax: +46-31-161973, christoffer.cromvik@chalmers.se

†Volvo 3P, Chassis & Vehicle Dynamics, Chassis Strategies & Vehicle Analysis, SE-405 08 Gothenburg, Sweden. Department of Mathematical Sciences, Chalmers University of Technology, and Department of Mathematical Sciences, University of Gothenburg, SE-412 96 Gothenburg, Sweden. Tel.: +46-31-3231063, peter.lindroth@volvo.com

where the multiple objectives are transformed to different single objective problems by varying the weight vector $\mathbf{w} \in \{\mathbf{v} \in \mathbb{R}^k \mid \sum_{i=1}^k v_i = 1,\ v_i \geq 0\ \forall i\}$. This solution strategy suffers from serious limitations, such that it is only possible to find the subset of $\mathcal{P}$ which is mapped onto the convex part of the Pareto front (Miettinen, 1998), and also that the mapping between $\mathbf{w}$ and the optimal values to (2), i.e., $\mathbb{R}^k \ni \mathbf{w} \mapsto \min_{\mathbf{x}\in X} \sum_{i=1}^k w_i f_i(\mathbf{x}) \in \mathbb{R}^k$, is non-linear and strongly depending on the properties of the actual functions involved (Das and Dennis Jr, 1997). To avoid finding *weakly Pareto optimal solutions* (where the strict inequality requirement in Definition 1.1 is dropped), the weights are required to be strictly positive. Despite of its limitations, the weighting strategy is fundamental, and is used as a basis for the definition of robustness presented in this paper.

1.1 Robustness in single- and multi-objective optimization

An optimal solution which is sensitive to perturbations in the data is often not useful in a practical application. A natural approach to deal with this situation is to incorporate the uncertainty into the model. This approach is used in Stochastic Programming (SP) (cf. Kall and Wallace (1994); Birge and Louveaux (1997)) and Robust Optimization (RO) (cf. Ben-Tal and Nemirovski (2002)). In SP, the objective function is typically the expected value over all uncertain parameters, which implies that an optimal solution is good on average. In RO, feasibility is required for all outcomes of the uncertain parameters, which produces a "conservative" optimal solution. Although most RO theory is restricted to convex problems with an explicit objective function, there are some recent development of RO methods also for non-convex as well as simulation-based problems (cf. Bertsimas et al. (2008, 2009)). Das (2000) views robustness as an objective in itself, and sets the goal to generate solutions that optimize both the unperturbed objective value and the expected objective value in a bi-objective optimization fashion.

There are, however, situations where it is not suitable or even possible to remodel the problem, but where there is an interest in assessing the robustness of an optimal solution in a post-process. This opens up the question of how robustness is evaluated. Considering a single-objective problem, we can use the sensitivity of the objective value at an optimal solution as a measure of robustness, but for multi-objective problems this is less straightforward. For such problems we have to quantify the uncertain responses in the objective space; see Figure 1.

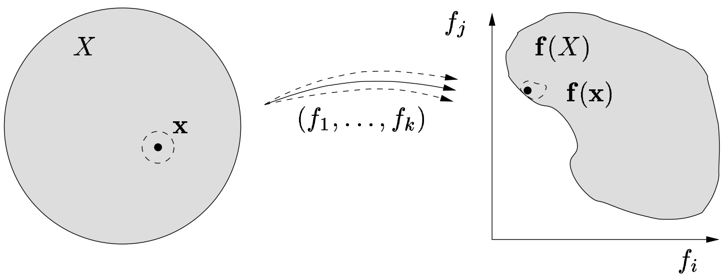


Figure 1: Uncertainties in $\mathbf{x}$ (such as implementation precision) and in $\mathbf{f} = (f_1, \ldots, f_k)$ (e.g. in model parameters) lead to uncertain responses in the objective space.

Among the many papers published on robust optimization, only few concerns multi-objective optimization. One has to distinguish between robust multi-objective optimization for which robustness is one objective and performance is the other (cf. Jin and Sendhoff (2003); Das (2000)), and our interpretation of robust multi-objective optimization where the wish is to find robust solutions to a multi-objective optimization problem. For the latter, Deb and Gupta (2005a,b, 2006) have made a direct extension of SP by using averaged values of the objective functions to define a robust Pareto front.

1.2 Outline

In Section 2, we construct one utility function for each decision maker which measures the objectives. We present a family of utility functions that spans the full range of "hidden objectives", and we also present a few properties of these functions and define two measures of robustness based on them. In Section 3, we discuss the computation of the robustness measures. Depending on the problem in terms of constraints and differentiability, we suggest two approaches to compute approximations of the measures. Section 4 deals with the search for robust solutions. Instead of just assessing the robustness of the Pareto solutions, we state an optimization problem with the goal to find robust, near-optimal solutions. In Section 5 we present two numerical examples. The first uses a known multi-objective test problem and the second considers a real-world problem instance in antenna design. Finally, in Section 6, we summarize the article and suggest some future work.

2 Robustness based on a utility function

To quantify the change in the objective space due to uncertainties in the decision space and in the objective itself, we use the notion of a *hidden objective* in a multi-objective problem. The hidden objective is tailored for each decision maker and captures his/her preferences. The robustness of a solution is then measured by this objective. With this approach, the computation of robustness must be considered as a post-process, since the preferences of the decision maker depend on the Pareto front.

The idea is to present a set of candidate solutions that are robust and constitute a reasonable approximation of the Pareto front. This implies that robustness can be treated as an objective itself, which is natural in a multi-objective setting.

2.1 Hidden objectives in multi-objective optimization

As mentioned previously, a multi-objective problem can often be viewed as a hidden single-objective optimization problem, where hidden means that the objective function is not explicitly known. A decision maker seeks *one* final solution which is optimal to him/her in the sense of balancing the different criteria. The reason for using a multi-objective formulation is to push forward the decisions until more knowledge is revealed about the characteristics and the limitations of the problem at hand. This single-objective optimization problem can be formulated as that to

$$\underset{\mathbf{x}\in X}{\text{minimize}} \;\; u\left(f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})\right), \tag{3}$$

where $u : \mathbb{R}^k \to \mathbb{R}$ is the hidden single objective. The observation of this formulation is the core of the ideas developed in this paper. Form here on, we refer to the hidden objective as the utility function, and use a convention that a smaller utility value is better than a larger.

Definition 2.1 (Rationality) *A utility function* $u : \mathbb{R}^k \to \mathbb{R}$ *is* rational *if for* $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$*,* $\mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\mathbf{y})$ *implies that* $(u \circ \mathbf{f})(\mathbf{x}) \leq (u \circ \mathbf{f})(\mathbf{y})$*. A decision maker is* rational *if his/her associated utility function is rational.*

Rationality means that if a point $\mathbf{y}$ is dominated by a point $\mathbf{x}$, then $\mathbf{x}$ must be appreciated as at least as good as $\mathbf{y}$.

With the above definition of rationality, the following proposition shows how rational utility functions can be characterized.

Proposition 2.2 *The utility function u is rational if and only if $u(f_1, \ldots, f_k)$ is monotonically increasing with each f_i, $i = 1, \ldots, k$.*

Proof. If u is monotonically increasing in every argument it holds that $u(\mathbf{f}(\mathbf{x})) \leq u(\mathbf{f}(\mathbf{y}))$ whenever $\mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\mathbf{y})$, i.e., u is rational. Suppose now that u is rational, but not monotonically increasing, i.e., $\exists\, \mathbf{f} \in \mathbb{R}^k$, $j \in \{1, \ldots, k\}$ and $\varepsilon > 0$ such that $u(f_1, \ldots, f_j, \ldots, f_k) > u(f_1, \ldots, f_j + \varepsilon, \ldots, f_k)$. But u is rational and hence since $(f_1, \ldots, f_j, \ldots, f_k) \leq (f_1, \ldots, f_j + \varepsilon, \ldots, f_k)$ it holds that $u(f_1, \ldots, f_j, \ldots, f_k) \leq u(f_1, \ldots, f_j + \varepsilon, \ldots, f_k)$. This is a contradiction, whence u must be monotonically increasing. □

We also make the following assumption on the function values of the Pareto solutions.

Assumption A

The objective values are scaled such that $\mathbf{f}(\mathcal{P}) \subseteq (0, 1]^k$.

If the range of f over X is bounded, it is always possible to scale the objectives such that Assumption A holds true.

2.2 The utility function

We assume that the utility function has the following form:

$$u(\mathbf{f}) := \sum_{i=1}^{k} w_i f_i^{\alpha}, \tag{4}$$

where $\mathbf{w} \in \mathbb{R}_+^k$ are weights and $\alpha \geq 1$ is a parameter related to curvature. We also define a family of utility functions.

Definition 2.3 *A family of* attainable *utility functions $\mathcal{U}$ is defined as*

$$\mathcal{U} = \left\{ \sum_{i=1}^{k} w_i f_i^{\alpha} \;\middle|\; w_i > 0,\ i = 1, \ldots, k;\ \alpha \in [1, \infty) \right\}. \tag{5}$$

We associate a utility function to each candidate vector $\bar{\mathbf{x}} \in X$, i.e., to any solution that a decision maker is interested in. If $\bar{\mathbf{x}} \in \mathcal{P}$, then $\mathbf{w}$ and α are chosen such that α is as small as possible and $\bar{\mathbf{x}} \in \operatorname{argmin} \{u \circ \mathbf{f}(\mathbf{x}) \mid \|\nabla_f (u \circ \mathbf{f})(\mathbf{x})\|_1 = 1\}$.

In the following, we present a few properties of the family of utility functions (5). The main goal is to show that the family is rational, and also complete with respect to certain Pareto optimal points in a sense to be defined below. These are points that can be reached using a utility function in the family $\mathcal{U}$, and we will use the notion of *proper Pareto optimality* to identify them.

We first define completeness for a general family of utility functions.

Definition 2.4 (Completeness) *A family of utility functions $\mathcal{U}$ is* complete with respect to a set *$\hat{\mathcal{P}} \subseteq \mathcal{P}$ if for every $\mathbf{x}^* \in \hat{\mathcal{P}}$ there exists a $u \in \mathcal{U}$ such that*

$$\mathbf{x}^* \in \arg\min_{\mathbf{x} \in X} u\big(f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})\big).$$

That is, in a complete family, for each $\mathbf{x}^* \in \hat{\mathcal{P}} \subseteq \mathcal{P}$ there is at least one utility function that evaluates $\mathbf{x}^*$ as a best one. A good family of utility functions is both rational and complete with respect to a set which is a close approximation to $\mathcal{P}$. We will show that the family (5) is a good one.

Proposition 2.5 *The family of utility functions defined by* (4) *is rational.*

Proof. Since $\mathbf{w} \geq \mathbf{0}^k$ and $\alpha \geq 1$, all $u \in \mathcal{U}$ are monotonically increasing in all their arguments; the result follows then immediately from Prop. 2.2. $\square$

Geoffrion (1968) introduced the notion of *proper* Pareto optimality to exclude some Pareto optimal solutions that are insensible to reasonable decision makers.

Definition 2.6 (Proper Pareto optimality) *A feasible solution $\hat{\mathbf{x}} \in X$ to* (1) *is called* proper Pareto optimal in the sense of Geoffrion *if it is Pareto optimal in* (1) *and if there exists a number $M > 0$ such that for each $i \in \{1, \ldots, k\}$ and each $\mathbf{x} \in X$ satisfying $f_i(\mathbf{x}) < f_i(\hat{\mathbf{x}})$, there exists a $j \in \{1, \ldots, k\} \backslash \{i\}$ such that $f_j(\hat{\mathbf{x}}) < f_j(\mathbf{x})$ and*

$$\frac{f_i(\hat{\mathbf{x}}) - f_i(\mathbf{x})}{f_j(\mathbf{x}) - f_j(\hat{\mathbf{x}})} \leq M. \tag{6}$$

We denote the set of all proper Pareto vectors in the sense of Geoffrion by $\check{\mathcal{P}}$.

A vector $\mathbf{x}$ is properly Pareto optimal in the sense of Geoffrion if it has finite trade-offs between the objectives. We make a somewhat different definition of proper Pareto optimality based on the family of utility functions (5).

Definition 2.7 (Firmly proper Pareto optimality) *A feasible solution to* (1) *is called* firmly proper Pareto optimal *if it is the minimizer of* (3) *for some utility function u in the family $\mathcal{U}$ defined in* (5). *We denote the set of all firmly proper Pareto vectors by $\mathcal{P}'$.*

Figure 2 illustrates some firmly proper, proper and non-proper Pareto optimal solutions. The definition of firmly proper Pareto optimal points implies that the family of utility functions is complete with respect to these points. The question now is which points are firmly proper.

The two following propositions show that firmly proper Pareto optimal solutions are indeed Pareto optimal, and that these solutions are also proper in the sense of Geoffrion.

Proposition 2.8 *Under Assumption* A, *each firmly proper Pareto optimal solution is a Pareto optimal solution, i.e., $\mathcal{P}' \subseteq \mathcal{P}$.*

Proof. Suppose that $\mathbf{x} \in X \backslash \mathcal{P}$. Then, $\exists\, \mathbf{y} \in X$ such that $f_i(\mathbf{y}) \leq f_i(\mathbf{x}),\ i = 1, \ldots, k$, with $f_j(\mathbf{y}) < f_j(\mathbf{x})$ for some index j. It follows that $u(\mathbf{f}(\mathbf{y})) = \sum_{i=1}^k w_i f_i(\mathbf{y})^\alpha < \sum_{i=1}^k w_i f_i(\mathbf{x})^\alpha = u(\mathbf{f}(\mathbf{x}))$ since $w_j > 0$, $f_j(\mathbf{y}) < f_j(\mathbf{x})$, $\alpha \geq 1$, and $\mathbf{f}(\mathbf{y}) > \mathbf{0}^k$. Thus $\mathbf{x} \notin \mathcal{P}'$ and the proposition follows. $\square$

Proposition 2.9 *Under Assumption* A, *each firmly proper Pareto optimal solution to* (1) *is a proper Pareto optimal solution, i.e., $\mathcal{P}' \subseteq \check{\mathcal{P}}$.*

Proof. Let $\mathbf{x}^* \in \mathcal{P}'$. Prop. 2.8 implies that $\mathbf{x}^* \in \mathcal{P}$. Suppose that $\mathbf{x}^*$ does not fulfill (6). Then for every $M > 0$, there exists an i and an $\mathbf{x} \in X$ with $f_i(\mathbf{x}) < f_i(\mathbf{x}^*)$ such that $\frac{f_i(\mathbf{x}^*) - f_i(\mathbf{x})}{f_j(\mathbf{x}) - f_j(\mathbf{x}^*)} > M$ for all $j \in \{1, \ldots, k\} \setminus \{i\}$ with $f_j(\mathbf{x}^*) < f_j(\mathbf{x})$.

Let us first consider a problem with two objectives, f_i and f_j. Let (w_i, w_j) and α be the parameters for a utility function with minimum at $\mathbf{x}^*$, and let $f_j(\mathbf{x}^*) > f_j(\mathbf{x}) - \frac{f_i(\mathbf{x}^*) - f_i(\mathbf{x})}{M}$, for every $M > 0$. Then we have that

$$\begin{aligned} u(f_i(\mathbf{x}^*), f_j(\mathbf{x}^*)) &= w_i f_i(\mathbf{x}^*)^\alpha + w_j f_j(\mathbf{x}^*)^\alpha \\ &\geq \lim_{M \to \infty} w_i f_i(\mathbf{x}^*)^\alpha + w_j \left(f_j(\mathbf{x}) - \tfrac{f_i(\mathbf{x}^*) - f_i(\mathbf{x})}{M} \right)^\alpha \\ &= w_i f_i(\mathbf{x}^*)^\alpha + w_j f_j(\mathbf{x})^\alpha \\ &> w_i f_i(\mathbf{x})^\alpha + w_j f_j(\mathbf{x})^\alpha \\ &= u(f_i(\mathbf{x}), f_j(\mathbf{x})). \end{aligned}$$

Hence $\mathbf{x}^*$ is not optimal in (3), i.e., $\mathbf{x}^* \notin \mathcal{P}'$. This leads to a contradiction.

Let us consider k objectives. For all $\mathbf{x}$, we can partition the objectives into three sets, $I_1(x) = \{i \mid f_i(\mathbf{x}) < f_i(\mathbf{x}^*)\}$, $I_2(x) = \{j \mid f_j(\mathbf{x}) > f_j(\mathbf{x}^*)\}$ and $I_3(x) = \{e \mid f_e(\mathbf{x}) = f_e(\mathbf{x}^*)\}$. In the inequality chain above, only the indices in I_2 are potentially harmful. But each $j \in I_2$ is above shown to result in a non-strict inequality, and also corresponding to an index resulting in a strict inequality. Therefore, altogether we get $u\big(f_1(\mathbf{x}^*), \ldots, f_k(\mathbf{x}^*)\big) > u\big(f_1(\mathbf{x}), \ldots, f_k(\mathbf{x})\big)$, which is the contradiction sought. □

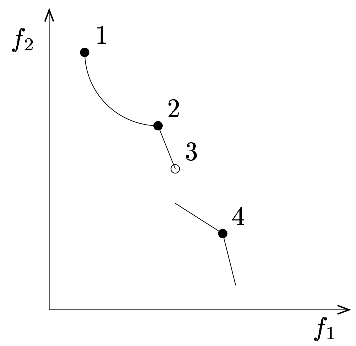


Figure 2: An illustration of the Pareto optimal set for a problem with two objectives. All points except the four marked are proper Pareto optimal points. Points 1 and 2 are not proper, and point 3 is not even Pareto optimal. Point 4 is proper but not firmly proper. Note that point 4 has points arbitrary close on both sides with different values of the trade-offs; this point can therefore be seen as insensible.

We will next identify which points on the Pareto front that are firmly proper. It turns out that for convex multi-objective problems, i.e., with all f_i convex and X convex, it is sufficient with $\alpha = 1$ and $\mathbf{w} \in \mathbb{R}^k_+$ in (4) to make the family (2.3) complete with respect to $\mathcal{P}'$. Since we require that the weights are strictly positive, there may be a few non-proper solutions; however, almost all Pareto optimal points to convex problems are firmly proper.

In the following proposition and corollary we show that also certain non-convex multi-objective problems have Pareto fronts consisting of only firmly proper Pareto points. The proposition is similar to what is shown in (Li, 1996); however, we assume that the objectives are scaled such that $\mathbf{f}(\mathcal{P}) \subseteq (0, 1]^k$. This enables another line of arguments, leading to a significantly shorter proof.

Proposition 2.10 (Convexification) *Consider the problem* (1) *under the Assumption* A. *Let the Pareto front* $\mathbf{f}(\mathcal{P})$ *be parameterized by* $f_k = \phi(f_1, \cdots, f_{k-1})$, *and let* $\mathbf{x}^* \in \mathcal{P}$. *Assume that the local trade-offs on the Pareto front between the pairs of objectives are continuous at* $\mathbf{f}(\mathbf{x}^*)$, *and assume that* ϕ *is twice continuously differentiable. Then, for a sufficiently large* $p \in [0, \infty)$, *the Pareto front of the problem* $\min_{\mathbf{x} \in X}(f_1(\mathbf{x})^p, \ldots, f_k(\mathbf{x})^p)$ *is convex at* $\mathbf{x}^*$.

Proof. Let $\bar{\mathbf{f}} = \{f_1, \ldots, f_{k-1}\}$ and $h(\bar{\mathbf{f}}) = \phi(\bar{\mathbf{f}})^p$. We will show that $\nabla^2_{(\bar{\mathbf{f}}^p)^2} h$ is positive semi-definite at $\mathbf{f}(\mathbf{x}^*)$.

From the chain rule, we have that

$$\frac{\partial h}{\partial (f_j^p)} = \frac{\partial h}{\partial f_j} \frac{1}{p f_j^{p-1}}$$

and that

$$\frac{\partial^2 h}{\partial (f_j^p)^2} = \frac{\partial^2 h}{\partial f_j^2}\frac{1}{p^2 (f_j^{p-1})^2} - \frac{p-1}{p^2 f^{2p-1}}\frac{\partial h}{\partial f_j},$$
$$\frac{\partial^2 h}{\partial f_j \partial f_i} = \frac{\partial^2 h}{\partial f_j \partial f_i}\frac{1}{p^2 f_j^{p-1} f_i^{p-1}}$$

We denote the exponent of a vector to be component wise, and introduce $D = \operatorname{diag}(\bar{\mathbf{f}}^{p-1})^{-1}$ and $E = D^{p-1}$. With these, we have that

$$\nabla_{\bar{\mathbf{f}}^p} h = \frac{1}{p} D \nabla_{\bar{\mathbf{f}}} h, \tag{7a}$$
$$\nabla_{(\bar{\mathbf{f}}^p)^2} h = \frac{1}{p^2} D \nabla^2_{\bar{\mathbf{f}}^2} h D - \frac{p-1}{p} D \operatorname{diag}(\nabla_{\bar{\mathbf{f}}} h) D E. \tag{7b}$$

Now, since

$$\frac{\partial h}{\partial f_j} = p\phi^{p-1}\frac{\partial \phi}{\partial f_j} \quad \text{and} \quad \frac{\partial^2 h}{\partial f_j^2} = p(p-1)\phi^{p-2}(\frac{\partial \phi}{\partial f_j})^2 + p\phi^{p-1}\frac{\partial^2 \phi}{\partial f_j^2},$$

we get

$$\nabla^2_{\bar{\mathbf{f}}^2} h = p(p-1)\phi^{p-2}\nabla_{\bar{\mathbf{f}}}\phi(\nabla_{\bar{\mathbf{f}}}\phi)^{\mathrm{T}} + p\phi^{p-1}\nabla^2_{\bar{\mathbf{f}}^2}\phi.$$

Finally, by inserting the above expression into (7) we get

$$\nabla_{(\bar{\mathbf{f}}^p)^2} h = \frac{p-1}{p}\phi^{p-2} D \nabla_{\bar{\mathbf{f}}}\phi(\nabla_{\bar{\mathbf{f}}}\phi)^{\mathrm{T}} D + \frac{1}{p}\phi^{p-1} D \nabla^2_{\bar{\mathbf{f}}^2}\phi D$$
$$- \frac{(p-1)^2}{p^2}\phi^{p-1} D \operatorname{diag}(\nabla_{\bar{\mathbf{f}}}\phi) D E.$$

The first term is positive semidefinite, and since $\nabla_{\bar{\mathbf{f}}}\phi < 0$, the last term is also positive semidefinite. When $p \to \infty$, the second term goes to zero faster than the first term, wherefore the result is proved. □

Corollary 2.11 *Let the Pareto front be parameterized by* $f_k = \phi(f_1, \ldots, f_{k-1})$*, and let* $\mathbf{x}^* \in \mathcal{P}$*. Assume that the local trade-offs on the Pareto front between all pairs of objectives are continuous at* $\mathbf{f}(\mathbf{x}^*)$*, and assume that* ϕ *is twice continuously differentiable. Then each proper Pareto optimal point is a firmly proper Pareto optimal point, and therefore* $\check{\mathcal{P}} = \mathcal{P}'$*.*

Proof. It is well known (cf. Ehrgott (2005), Thm. 3.11) that all proper Pareto optimal points to convex multi-objective optimization problems can be found using the standard weighting method with non-negative weights. The result follows from Proposition 2.9. □

The corollary implies that all points on sufficiently smooth Pareto fronts are firmly proper, i.e., for problems with such Pareto fronts, our family of utility functions is complete with respect to the whole of $\mathcal{P}$.

To conclude this section, we have shown that the family $\mathcal{U}$ of utility functions is rational and complete with respect to almost all Pareto solutions arising from convex problems and all Pareto fronts that are smooth enough.

2.3 The robustness index

We present two definitions of robustness for a given decision vector: absolute robustness and relative robustness. Both measures are based on the utility function (4), and for both of them, a smaller value means a more robust point.

Definition 2.12 (Absolute robustness index) *Let $\bar{\mathbf{x}} \in \mathbb{R}^n$ be the point whose robustness is to be measured, and let $\boldsymbol{\eta} \in \Omega \subseteq \mathbb{R}^m$ be a stochastic variable with mean $\boldsymbol{\eta}_0$. Suppose that $u(\cdot)$ is the utility function associated with $\bar{\mathbf{x}}$. The absolute robustness index of $\bar{\mathbf{x}}$ is defined as*

$$R_A(\bar{\mathbf{x}}) = \mathbf{E}[(u \circ \mathbf{f})(\bar{\mathbf{x}}, \boldsymbol{\eta}) - (u \circ \mathbf{f})(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)].$$

Definition 2.13 (Relative robustness index) *Let $\bar{\mathbf{x}} \in \mathbb{R}^n$ be the point whose robustness is to be measured, and let $\boldsymbol{\eta} \in \Omega \subseteq \mathbb{R}^m$ be a stochastic variable with mean $\boldsymbol{\eta}_0$. Suppose that $u(\cdot)$ is the utility function associated with $\bar{\mathbf{x}}$ and that $\mathbf{x}^*(\boldsymbol{\eta}) \in X \cap \arg\min(u \circ \mathbf{f})(\mathbf{x}, \boldsymbol{\eta})$. The relative robustness index of $\bar{\mathbf{x}}$ is defined as*

$$R(\bar{\mathbf{x}}) = \mathbf{E}[(u \circ \mathbf{f})(\bar{\mathbf{x}}, \boldsymbol{\eta}) - (u \circ \mathbf{f})(\mathbf{x}^*(\boldsymbol{\eta}), \boldsymbol{\eta})].$$

Remark 2.14 *Due to Jensen's inequality (cf. Fristedt and Gray (1997), Prop. 12), if $(u \circ \mathbf{f})(\bar{\mathbf{x}}, \cdot)$ is convex, the absolute robustness is non-negative,*

$$R_A(\bar{\mathbf{x}}) \geq (u \circ \mathbf{f})(\bar{\mathbf{x}}, \mathbf{E}[\boldsymbol{\eta}]) - (u \circ \mathbf{f})(\bar{\mathbf{x}}, \boldsymbol{\eta}_0) = 0.$$

In contrast to absolute robustness, relative robustness is not necessarily affected by large changes in the objective space due to different outcomes of $\boldsymbol{\eta}$, since it measures the relative loss to an optimal solution for each $\boldsymbol{\eta}$; see Figure 3.

Which robustness index should be used may be a matter of choice for a decision maker, but practice may motivate the use of one before the other. For example, using relative robustness requires a minimization for each $\boldsymbol{\eta}$ which limits its practical use on some problems. In Section 3, we present procedures for computing approximations of the robustness indices.

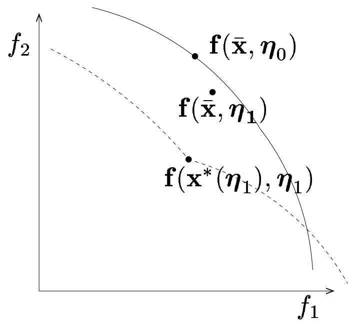


Figure 3: Two Pareto fronts for two realizations of the uncertainty parameter $\boldsymbol{\eta}$. There is a quality loss since the chosen candidate $\bar{\mathbf{x}}$ is not optimal for the outcome $\boldsymbol{\eta}_1$. This quality loss is measured in the relative robustness index.

3 Computation of the utility function and the robustness index

In this section we present practical approaches for computing the utility function and the robustness indices. We start by noting that the computation of robustness is a post-process since it requires a sufficient resolution of the Pareto front. We state this as an assumption:

Assumption B

The Pareto front is computed to a sufficient accuracy and resolution.

Algorithm 1 Calculate robustness index

Input: Candidate $\bar{\mathbf{x}}$, Pareto front $\mathbf{f}(\mathcal{P})$.

1. Approximate the Pareto front by a quadratic implicit curve around $\bar{\mathbf{x}}$.

2. Compute the utility function u for the candidate such that equations (8) and (9) are fulfilled.

3. Compute R or R_A according to the descriptions in subsections 3.1 and 3.2.

The computation of the robustness indices for a specific solution, which we call the candidate, is organized in a series of steps, where the main points are stated in Algorithm 1.

We assume that the Pareto front $\mathbf{f}(\mathcal{P})$ is described by a level set of an implicit function $z(\mathbf{f}(\mathcal{P})) = 0$. By the definition of the utility function u, it is minimized by the candidate $\bar{\mathbf{x}}$. This implies the following two conditions which are also illustrated in Figure 4:

$$\nabla_f u = \gamma \nabla_f z, \tag{8}$$

$$\kappa(u) \geq \kappa(z), \tag{9}$$

where $\gamma \in \mathbb{R}_+$, and $\kappa(\cdot)$ is a measure of curvature. We make a quadratic fit Q of the Pareto front based on the Pareto points within a ball of radius $\tau > 0$. In particular, given a candidate $\bar{\mathbf{x}}$ and the Pareto points in the vicinity, $\mathbf{x}^j$ for $j = 1, \ldots, p$, we solve the following linear least-squares problem

$$\begin{aligned} \underset{c\in\mathbb{R}^m, b\in\mathbb{R}^m}{\text{minimize}} \quad & \sum_{j=1}^{p}\sum_{i=1}^{k} \left(c_i f_i(\mathbf{x}^j)^2 + b_i f_i(\mathbf{x}^j) - 1\right)^2, \\ \text{subject to} \quad & \sum_{i=1} c_i f_i(\bar{\mathbf{x}})^2 + b_i f_i(\bar{\mathbf{x}}) = 1, \end{aligned}$$

and set

$$Q(\mathbf{f}) = \tfrac{1}{2}\mathbf{f}^{\mathrm{T}} \operatorname{diag}(\mathbf{c})\, \mathbf{f} + \mathbf{b}^{\mathrm{T}}\mathbf{f} - 1.$$

This yields an estimate of the normal and curvature of the front. Since $|\partial Q/\partial f_i| > 0$ and Q is twice continuously differentiable, by the implicit function theorem, there exists a twice continuously differentiable ϕ such that $f_k = \phi(f_1, \ldots, f_{k-1})$. This means that all points on Q are firmly proper according to Proposition 2.10. So even though the Pareto front may not be sufficiently smooth, we are always able to reach all points on the approximate front, and there always exists an α such that equations (8) and (9) hold. We use normal curvature in equation (9) and we define it, along a vector d, to be

$$\kappa = \frac{d^{\mathrm{T}} H d}{||d||^2}, \tag{10}$$

where H is the Jacobian of the normal $N \in \mathbb{R}^k$ to the surface,

$$H = \begin{bmatrix} \frac{\partial N_1}{\partial x_1} & \cdots & \frac{\partial N_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial N_k}{\partial x_1} & \cdots & \frac{\partial N_k}{\partial x_k} \end{bmatrix}. \tag{11}$$

We note that in three dimensions, the principal curvatures are the two nonzero eigenvalues to the matrix H (Araujo and Jorge, 2004). Equations (9) and (10) should hold for all directions $d \in \mathbb{R}^k$, although in practice, we only consider a finite set of directions. For the quadratic implicit surface $Q(\mathbf{f})$, equation (11) reduces to (cf. Hughes (2003); Araujo and Jorge (2004))

$$H = \frac{\nabla^2_{ff} Q}{||\nabla_f Q||} - \frac{(\nabla_f Q \nabla_f Q^{\mathrm{T}})\nabla^2_{ff} Q}{||\nabla_f Q||^3}. \tag{12}$$

A utility function which fulfills equations (8) and (9) can be found by iteratively setting the curvature parameter $\alpha \geq 1$ with corresponding weights $\mathbf{w} > \mathbf{0}$ such that equation (8) is satisfied. Equation (9) will then be satisfied for a sufficiently high value of α.

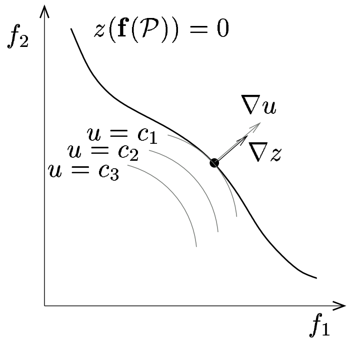


Figure 4: An illustration of the requirements (8) and (9) on the utility function u.

Given a candidate with a corresponding utility function, the next step is to compute R or R_A. This is described in the following subsections. For unconstrained problems with analytic objective functions, we present an approximate closed-form expression for relative robustness. For constrained problems, we show how a Monte-Carlo method can be used. Both methods are used in the numerical experiments in Section 5.

3.1 Unconstrained problem

In addition to assumptions A and B, which we assume holds for $\mathbf{f}(\mathbf{x}, \boldsymbol{\eta}_0)$, we assume the following:

Assumption C

(C1) The functions $f_i(\cdot,\cdot) > 0$, $i = 1, \ldots, k$, are twice continuously differentiable.

(C2) The feasible set is $X = \mathbb{R}^n$.

Under these assumptions, we can formulate a closed-form expression for an approximation of relative robustness. The approximation is based on the second-order Taylor expansion U of the utility function u. With $\hat{u}(\mathbf{x}, \boldsymbol{\eta}) := u(\mathbf{f}(\mathbf{x}, \boldsymbol{\eta}))$, we have

$$\begin{aligned}
U(\mathbf{x}, \boldsymbol{\eta}) &= \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0) + \nabla_x \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)^{\mathrm{T}}(\mathbf{x} - \bar{\mathbf{x}}) + \nabla_\eta \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)^{\mathrm{T}}(\boldsymbol{\eta} - \boldsymbol{\eta}_0) \\
&\quad + (\mathbf{x} - \bar{\mathbf{x}})^{\mathrm{T}} \nabla^2_{x\eta} \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)(\boldsymbol{\eta} - \boldsymbol{\eta}_0) + \tfrac{1}{2}(\mathbf{x} - \bar{\mathbf{x}})^{\mathrm{T}} \nabla^2_{xx} \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)(\mathbf{x} - \bar{\mathbf{x}}) \\
&\quad + \tfrac{1}{2}(\boldsymbol{\eta} - \boldsymbol{\eta}_0)^{\mathrm{T}} \nabla^2_{\eta\eta} \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)(\boldsymbol{\eta} - \boldsymbol{\eta}_0),
\end{aligned}$$

Since the candidate $\bar{\mathbf{x}}$ is defined to minimize the utility function, the Hessian of u is positive semi-definite. If it is positive definite, and thus non-singular, we get an expression for the optimal solution $\mathbf{x}^*(\boldsymbol{\eta}) \in \arg\min_{\mathbf{x} \in X} U(\mathbf{x}, \boldsymbol{\eta})$ as a (linear) function of the uncertainty parameter $\boldsymbol{\eta}$:

$$\mathbf{x}^*(\boldsymbol{\eta}) = \bar{\mathbf{x}} - \nabla^2_{xx} \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)^{-1} \left[\nabla_x \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0) + \nabla^2_{x\eta} \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)^{\mathrm{T}}(\boldsymbol{\eta} - \boldsymbol{\eta}_0) \right].$$

Inserting this into the definition of robustness (2.13) leads to a closed-form expression for the approximate relative robustness index:

$$\begin{aligned}
R^U(\bar{\mathbf{x}}) := \mathbf{E}\left[\hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}) - \hat{u}(\mathbf{x}^*(\boldsymbol{\eta}), \boldsymbol{\eta})\right] = \mathbf{E}\left[\tfrac{1}{2} \nabla_x \hat{u}^{\mathrm{T}} \nabla^2_{xx} \hat{u}^{-1} \nabla_x \hat{u} \right. \\
\left. + \nabla_x \hat{u}^{\mathrm{T}} \nabla^2_{xx} \hat{u}^{-1} \nabla^2_{\eta x} \hat{u}^{\mathrm{T}} (\boldsymbol{\eta} - \boldsymbol{\eta}_0) + \tfrac{1}{2}(\boldsymbol{\eta} - \boldsymbol{\eta}_0)^{\mathrm{T}} \nabla^2_{\eta x} \hat{u} \nabla^2_{xx} \hat{u}^{-1} \nabla^2_{\eta x} \hat{u}^{\mathrm{T}} (\boldsymbol{\eta} - \boldsymbol{\eta}_0) \right].
\end{aligned} \tag{13}$$

Introducing Λ as the covariance matrix of $\boldsymbol{\eta}$, and noting that $\nabla_x \hat{u}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0) = 0$ since $\bar{\mathbf{x}}$ is the minimizer, the expression (13) reduces to

$$R^U(\bar{\mathbf{x}}) = \tfrac{1}{2}\mathrm{tr}\left(\Lambda \nabla^2_{\eta x}\hat{u} \nabla^2_{xx}\hat{u}^{-1} \nabla^2_{\eta x}\hat{u}^{\mathrm{T}}\right). \tag{14}$$

This expression only requires the solution of one linear equation with n unknowns and a few matrix-matrix multiplications, and is thus relatively fast to compute.

3.2 Constrained problem

If any of the functions f_i are non-differentiable, if the problem includes constraints or if analytic expressions of the functions f_i are not available, the closed-form expression (14) does not apply. The robustness indices can however be computed using a Monte-Carlo method with randomized sampling. The idea is to draw N i.i.d. samples $\boldsymbol{\eta}_i$, $i = 1, \ldots, N$, of $\boldsymbol{\eta}$ and replace the expected value by the sample mean. We only consider the absolute robustness index, since the relative index would require one minimization computation for each sample. The Monte-Carlo estimate is given by

$$\hat{R}_A(\bar{\mathbf{x}}) = \frac{1}{N}\sum_{i=1}^{N}\left[u\big(\mathbf{f}(\bar{\mathbf{x}}, \boldsymbol{\eta}_i)\big) - u\big(\mathbf{f}(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)\big)\right].$$

4 Search for robust solutions

In Section 3, we assume (Assumption B) that the Pareto front is pre-computed, and the practical computational procedures presented refer to candidates on the front. In reality, however, we may forsake optimality of a solution if robustness can be gained. The idea is to move away from a Pareto optimal solution $\bar{\mathbf{x}}$ on the front and search for robust solutions in its neighborhood. Let $\tau > 0$ be the radius of the ball around $\mathbf{f}(\bar{\mathbf{x}})$ used for the quadratic approximation Q of the front, and let $\varepsilon > 0$. We use the utility function u to define the neighborhood. For absolute robustness we formulate the optimization problem to

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & R_A(\mathbf{x}), \\ \text{subject to} \quad & u\left(\mathbf{f}(\mathbf{x})\right) - u\left(\mathbf{f}(\bar{\mathbf{x}})\right) \le \varepsilon, \\ & \|\mathbf{f}(\mathbf{x})) - \mathbf{f}(\bar{\mathbf{x}})\| \le \tau, \\ & \mathbf{x} \in X. \end{aligned} \tag{15}$$

The solution to (15) is the most robust point with at most a decrease of ε in utility compared to $\bar{\mathbf{x}}$ and which is sufficiently close to $\bar{\mathbf{x}}$ in the objective space such that the local approximation of the Pareto front remains valid.

For relative robustness, we have to take into account that inner (non-Pareto) solutions will have a lower utility value for each realization than the optimal solution. Letting $\Delta u(\mathbf{x}) = u(\mathbf{x}, \boldsymbol{\eta}_0) - u(\bar{\mathbf{x}}, \boldsymbol{\eta}_0)$ denote the loss in utility at the unperturbed state, we formulate the optimization problem to

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & R(\mathbf{x}) + \Delta u(\mathbf{x}), \\ \text{subject to} \quad & u\left(\mathbf{f}(\mathbf{x})\right) - u\left(\mathbf{f}(\bar{\mathbf{x}})\right) \le \varepsilon, \\ & \|\mathbf{f}(\mathbf{x})) - \mathbf{f}(\bar{\mathbf{x}})\| \le \tau, \\ & \mathbf{x} \in X. \end{aligned}$$

5 Numerical Experiments

The ideas developed in this article have been applied to both the analytical test functions used in the article by Deb and Gupta (2005a) and on functions derived from real-world numerical data used for antenna optimization (Jakobsson et al., 2008a; Stjernman et al., 2009).

The reader should note that the test functions in the first numerical example are designed to illustrate different principal cases when introducing uncertainty in multi-objective problems, and not designed to imitate practical applications. Our intention with this example is to show how our definition of robustness compares to published results.

The theory developed in this article places no theoretical restriction on the number of objectives. However—for illustrational reasons—in the numerical examples we only consider bi-objective problems.

5.1 Analytical functions

Deb and Gupta (2005a) considers uncertainty in the decision space and formulates a program where each objective function is replaced by its respective average computed over a ball around the intended decision variable, i.e.,

$$f_i^{\text{eff}}(\mathbf{x}) = \frac{1}{|\mathcal{B}(\mathbf{x},\delta)|} \int_{\mathbf{y}\in\mathcal{B}(\mathbf{x},\delta)} f_i(\mathbf{y}) d\mathbf{y}.$$

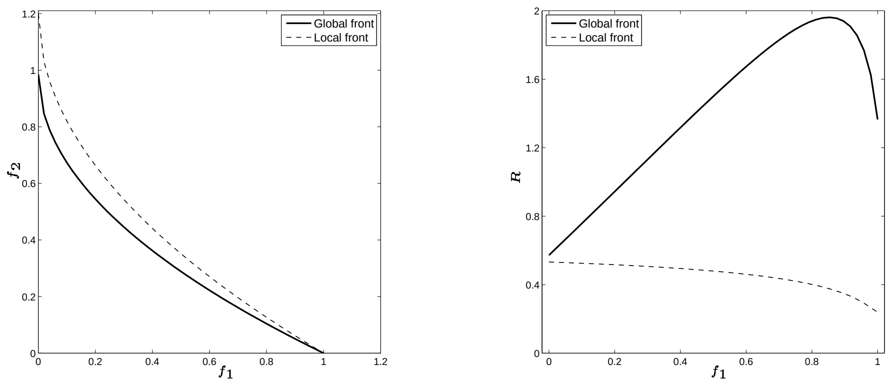
The radius of the ball is given by the parameter $\delta > 0$ which is varied in the numerical tests. A larger value of δ smoothens out the functions and makes sharp global optima less attractive. By using this framework, a "robust" Pareto optimal front is always found, but there is no distinction between the points on this front with respect to robustness. Furthermore, there is no continuous grading of robustness of the points that are not in the robust Pareto set. Deb and Gupta also present an alternative robustness model where they enforce robustness of the resulting solutions using a constraint. The requirement is that the norm of the difference between the (unperturbed) function value and the averaged (or, the worst case) function value must be kept smaller than a certain threshold value. From our point of view, this formulation also suffers from the weakness that it just classifies solutions as robust Pareto optimal or not. It is also possible that large parts of the objective space will not contain any robust solutions if the effect of uncertainty is large. From now on, we will concentrate on Deb and Gupta's first formulation.

Since we derive robustness for the unperturbed front and Deb and Gupta presents a robust front, possibly consisting of completely different solutions, it is difficult to directly compare the respective results. We will, however, show that they are in line in principle.

We present numerical results for one test problem, DEBGUP3, which is one of 4 bi-objective problems from (Deb and Gupta, 2005a). The problem is to

$$\begin{aligned}
\underset{\mathbf{x}}{\text{minimize}} \quad & (f_1(\mathbf{x}), f_2(\mathbf{x})) = \\
& \left(x_1, \left(2 - 0.8e^{-\left(\frac{x_2-0.35}{0.25}\right)^2} - e^{-\left(\frac{x_2-0.85}{0.03}\right)^2}\right)(1-\sqrt{x_1})\sum_{i=3}^{5} 50x_i^2 \right), \\
\text{subject to} \quad & 0 \le x_i \le 1, \quad i = 1,2, \\
& -1 \le x_i \le 1, \quad i = 3,4,5.
\end{aligned} \qquad \text{(DEBGUP3)}$$

The uncertainty appears in the decision space, such that $\mathbf{x}$ is replaced by $\mathbf{x}+\boldsymbol{\eta}$ and $\boldsymbol{\eta}$ is drawn from a uniform distribution, $\boldsymbol{\eta} \in U([-1,1]^5)$. A close study of the functions reveals that the unperturbed problem with $\boldsymbol{\eta} = \mathbf{0}$ has one local and one global Pareto optimal front, where a local Pareto front consists of points that are locally Pareto optimal. The fronts are shown in Figure 5(a). Figure 5(b) presents the relative robustness index (Def. 2.13) for the corresponding points. We have chosen to ignore the bounds when computing relative robustness which enabled the use of the closed-form expression (14). The implication of ignoring the bounds may be that the value of R is overestimated, i.e., the robustness is underestimated.



(a) Function values.

(b) Robustness indices.

Figure 5: The global and the local Pareto fronts for the problem DebGup3. The robustness indices are shown for the corresponding points, parameterized by the values of the first objective.

Note that the robustness varies both along each single front, and also between the two fronts. The local front is more robust than the global one as is expected from the results in (Deb and Gupta, 2005a). Here, we can distinguish a difference between using the robustness index and using averaged objectives. Depending on the size of the radius δ, the robust front will equal either the local front, the global front, or a combination of these. It is possible to construct problems where the global front equals the robust Pareto front, but having a local Pareto front arbitrarily close and which has much better robustness indices according to our definition. The size of δ highly determines which solutions are presented to a decision maker, whereas the idea in this paper is to partly push forward the decision of how much robustness is desired to the decision maker, and therefore present solutions of different robustness values. The robustness index may also show that robustness may vary along the Pareto front. With more complex objective functions found in real-world applications, we anticipate that there may be more dramatic changes in robustness between close solutions on the Pareto front. In such cases, the decision maker may prefer a solution slightly off his/her ideal (optimal) solution if the robustness properties are better. This situation is presented in the following subsection.

5.2 A real-world example

Designing antennas typically involves a number of conflicting requirements. These may be based on spatial size, so called S-parameters related to electromagnetic properties, functions of the directivity of the antenna, band width, input impedance, or other characteristics of the antenna. In a joint project between the Fraunhofer-Chalmers Centre and the Antenna Research Centre at Ericsson AB a multi-objective optimization approach is taken on the antenna design problem, as described in (Jakobsson et al., 2008a; Stjernman et al., 2009). We have chosen to study this problem using a subset of the proposed objectives. The decision variables are the positions and geometrical dimensions of the antenna, and the objectives chosen are the maximum return loss ($|S_{11}|$) over the frequency band $[750, 850]$ MHz and the area of the hull of the antenna. An approximate Pareto front is shown in Figure 6, where it is clearly shown that the two objectives are conflicting. The objective functions are expensive to evaluate since they are outcomes of long-running computer simulations. For this reason, a surrogate modeling technique (Jin et al., 2001) is used, where approximate functions are constructed using the function values computed at a number of sample points. Jakobsson et al. (2008a,b) have developed a new technique based on interpolation with rational radial basis functions to handle the sharp function behaviors around the resonance levels.

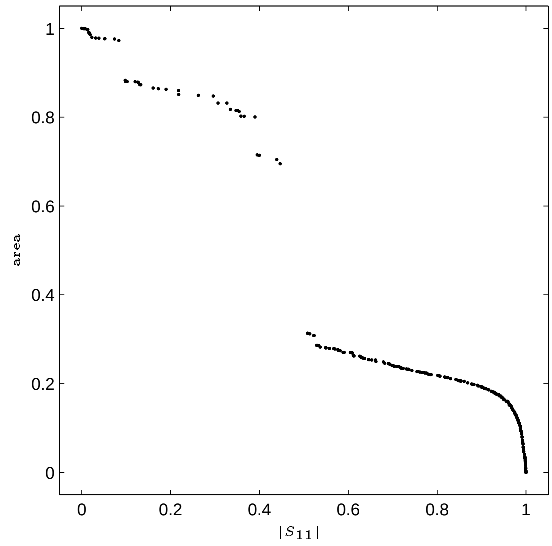


Figure 6: An approximate Pareto front (with the objectives scaled) to the problem found using the NSGA-II algorithm (Deb et al., 2000) with 200 generations and a population size of 48.

The two objectives are interesting for a robustness study. Near resonance, small variations of the decision variables yield large differences in the function values. This is the case for many practical problems where resonance phenomena are part of the problem characteristics. We have noticed that the surrogate models are quite sensitive to the choice of sample points (and this choice is not obvious) and have constructed our numerical study based on this fact.

Originally, the decision space has been sampled at 2000 distinct point chosen using an ad-hoc design–of–experiments strategy, and the surrogate models (or, response surfaces) have been constructed using the rational RBF technique on the function evaluations at these points. In our experiments, we have randomly selected 500 out of the 2000 points and constructed new response surfaces using only these. The uncertainty characteristics depend on which of the 2000 points that are chosen, reflecting the fact that it is not clear from the start which sample points to choose. Obviously, a *robust* solution is a solution for which the randomness does not have a large effect according to our definition of robustness. In Figure 7, the (absolute) robustness index is shown for the points on the Pareto front to the original problem, where the objective functions are the response surfaces constructed using all 2000 data points. The index varies substantially along the front, and for some Pareto points, there are other points on the front that are close in the objective space but with a very different robustness index. This opens up the possibility for a decision maker to choose a point which lies close to his or her ideal point with respect to the function values, but which are much more robust. Doing so will, on average, improve the utility. But since the front is only valid for the unperturbed problem, a decision maker could also be searching for a non-Pareto optimal solution since such a point can be even more robust. In Figure 8, we illustrate such a search. For each (unperturbed) Pareto optimal point, we search for optimal points according to

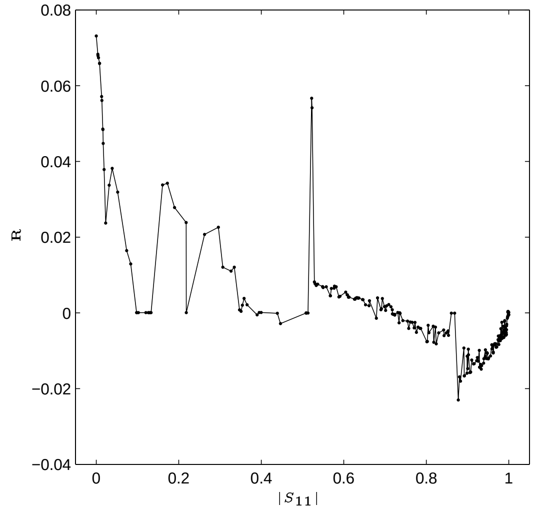


Figure 7: Absolute robustness for the points on the Pareto front, parameterized by the first objective.

the model (15) with the parameter values $\varepsilon = 0.01$ and $\tau = 0.1$. We use the global optimization algorithm `DIRECT` (Jones, 2009), implemented in `TOMLAB` (T). We have also implemented a simple local search strategy to complement the algorithm. In the left figure, the points obtained are shown together with the original (approximate) Pareto front. The right figure shows a histogram of the size of the improvements in robustness when—for each unperturbed point—picking the corresponding robust alternative. One obvious conclusion is that for most Pareto optimal points, there are robust solutions that are close with respect to the value of the utility but with a significantly better robustness index. This fact can be used by a decision maker, who gets an option to choose between robustness and "optimality" for the unperturbed problem.

To further illustrate the framework, we consider the following scenario: Suppose we have presented a Pareto front corresponding to the unperturbed problem to a decision maker, and that he/she has located a candidate solution. Since the problem contains uncertain parameters, the decision maker is also interested in the robustness of this solution. We now solve problem (15) for varying values on the parameter ε. This will produce solutions that are more robust, but with lower utility values. These candidates are then presented to the decision maker, who gets the option to consider how much he or she values robustness considering how much utility is lost. In the spirit of multi-objective optimization, the decision of robustness versus optimality is thus left to the decision maker. Figure 9 shows the results for a specific candidate. In a), the robustness index is shown as a function of the utility for the alternative solutions. In b), the unperturbed Pareto front is shown along with level curves of the utility function for the specific candidate. The results show that the decision maker can substantially improve the robustness if he or she is willing to sacrifice some utility.

6 Summary and conclusions

We have presented a new definition of robustness for multi-objective problems based on the idea that each decision maker has a hidden single objective function defining which of the Pareto optimal points he/she desires. This hidden objective is characterized with a family of utility functions; we present two robustness indices measuring the relative and absolute expected loss in the utility function due to uncertainties. We have shown that the family of utility functions has certain nice properties such as rationality and completeness. We also presented procedures for computing the robustness indices and applied them to two numerical examples: an analytic test problem from (Deb and Gupta, 2005a), and an in antenna optimization from (Jakobsson et al., 2008a; Stjernman et al., 2009).

The formulation of robustness by Deb and Gupta (2005a) for multi-objective optimization, which consists of replacing the objectives by their respective expected values, is very natural and direct, and is suitable for many applications. In line with the main idea of multi-objective optimization, our approach has the advantage that the decisions are pushed further into the future when more information about the problem has been revealed. Also, our method gives a continuous measure of robustness and it does that to all points; it does not only say if a point is a robust Pareto optimal point or not.

In future work, our methodology should be applied to more numerical examples, and also to problem instances with more than two objectives. The inclusion of constraints for relative robustness should be developed further. It would also be interesting to apply other types of robustness measures based on utility functions.

Acknowledgments

The first author is partially financially supported by the Gothenburg Mathematical Modeling Centre (GMMC) at Mathematical Sciences (Chalmers University of Technology and the University of Gothenburg), whose main sponsor is the Swedish Foundation for Strategic Research (SSF). The second author is financially supported by Volvo 3P.

The authors wish to thank Michael Patriksson and Ann-Brith Strömberg for their insightful comments and helpful suggestions.

References

B. R. Araujo and J. A. P. Jorge. Curvature dependent polygonization of implicit surfaces. In *Proceedings of the 17th Brazilian Symposium on Computer Graphics and Image Processing*, pages 266–273, Washington DC, USA, 2004. IEEE Computer Society.

A. Ben-Tal and A. Nemirovski. Robust optimization – methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.

D. Bertsimas, O. Nohadani, and K. M. Teo. Robust nonconvex optimization for simulation-based problems. *Operations Research (to appear)*, 2008. URL http://www.optimization-online.org.

D. Bertsimas, O. Nohadani, and K. M. Teo. Nonconvex robust optimization for problems with constraints. *Informs Journal on Computing (to appear)*, 2009.

J. R. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer Series in Operations Research. Springer-Verlag, New York, NY, USA, 1997.

I. Das. Robustness optimization for constrained nonlinear programming problems. *Engineering Optimization*, 32(5):585–618, 2000.

I. Das and J. E. Dennis Jr. A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural Optimization*, 14: 63–69, 1997.

K. Deb and H. Gupta. Searching for robust Pareto-optimal solutions in multi-objective optimization. In *Evolutionary Multi-Criterion Optimization*, pages 150–164, 2005a.

K. Deb and H. Gupta. Handling constraints in robust multi-objective optimization. In *Congress on Evolutionary Computation*, volume 1, pages 25–32, 2005b.

K. Deb and H. Gupta. Introducing robustness in multi-objective optimization. *Evolutionary Computation*, 14(4):463–494, 2006.

K. Deb, A. Pratap, Sameer A., and T. Meyarivan. A fast elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6:182–197, 2000.

M. Ehrgott. *Multicriteria Optimization.* Springer-Verlag, Heidelberg, Germany, 2005.

B. Fristedt and L. Gray. *A modern approach to probability theory.* Probability and its Applications. Birkhäuser Boston Inc., Boston, MA, USA, 1997.

A. Geoffrion. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications*, 22:618–630, 1968.

J. Hughes. Differential geometry of implicit surfaces – a primer. Technical report, Brown University, Computer Science, Providence, USA, 2003.

S. Jakobsson, B. Andersson, and F. Edelvik. Multiobjective optimization applied to design of PIFA antennas. In *Proceedings of the 7th International Conference on Scientific Computing in Electrical Engineering (SCEE 2008)*, Helsinki, Finland, October 2008a.

S. Jakobsson, B. Andersson, and F. Edelvik. Rational radial basis function interpolation with applications to antenna design. Technical report, Fraunhofer Chalmers Research Centre, Göteborg, Sweden, 2008b.

R. Jin, W. Chen, and T. W. Simpson. Comparative studies of metamodeling techniques under multiple modeling criteria. *Journal of Structural Optimization*, 23(1):1–13, 2001.

Y. Jin and B. Sendhoff. Trade-off between optimality and robustness: An evolutionary multi-objective approach, in evolutionary multi-criterion optimization. In *Proceeding of the Second International Conference on Evolutionary Multi-criterion Optimization*, pages 237–251, Algarve, Portugal, 2003. Springer-Verlag.

D. R. Jones. Direct global optimization algorithm. In C.A. Floudas and P.M. Padalos, editors, *Encyclopedia of Optimization*, pages 725–735. Kluwer Academic, 2009.

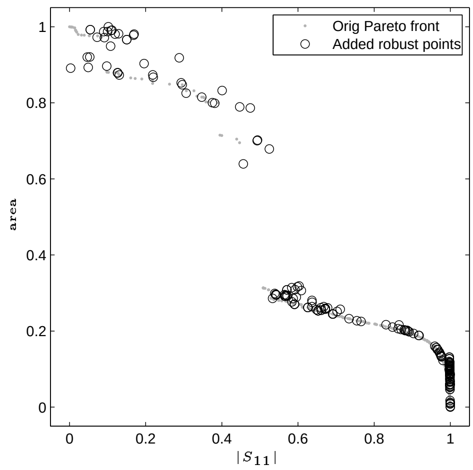
P. Kall and S. W. Wallace. *Stochastic Programming.* Wiley John & Sons, Chichester, England, 1994.

D. Li. Convexification of a noninferior frontier. *Journal of Optimization Theory and Applications*, 88(1):177–196, 1996.

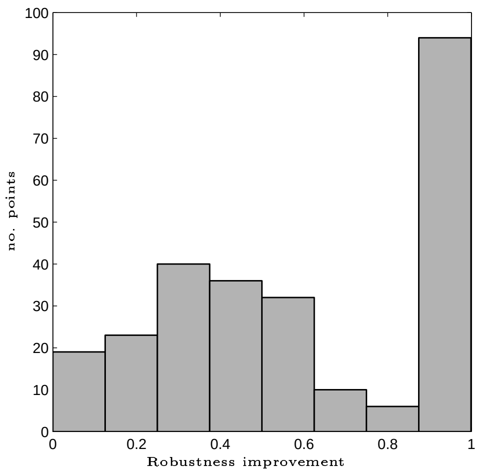
K. Miettinen. *Nonlinear Multiobjective Optimization.* Kluwer Academic Publishers, Boston, MA, USA, 1998.

A. Stjernman, A. Derneryd, S. Jakobsson, B. Andersson, and F. Edelvik. Multi-objective optimization of MIMO antenna systems. In *Proceedings of the 3rd European Conference on Antennas and Propagation*, Berlin, Germany, March 2009.

T. The TOMLAB optimization environment, http://tomopt.com/tomlab/.

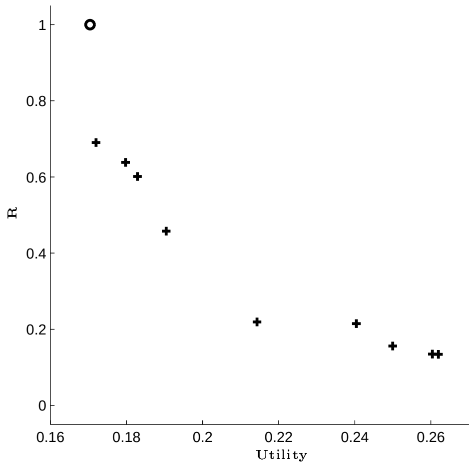


(a)

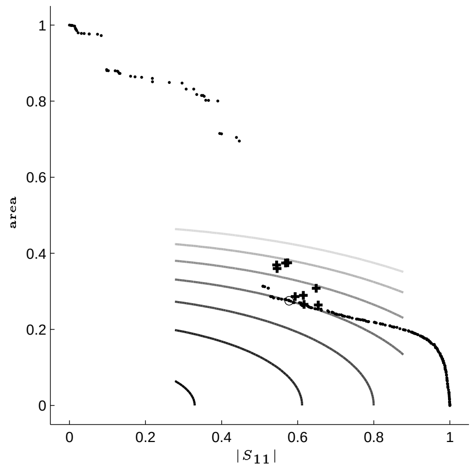


(b)

Figure 8: In a), robust points are added to the (approximate) Pareto front. In b), the relative improvements in the robustness index for the points found are shown.



(a)



(b)

Figure 9: Figure a) shows the utility function values and robustness index values for the alternative solutions. The robustness index is normalized by the original candidate. Figure b) shows the unperturbed function values for the candidates and level curves of the utility function. In both figures, the ring (o) corresponds to the unperturbed Pareto point originally chosen by the decision maker and the plus signs (+) correspond to the alternative points.