Parallel Universes: Multi-Criteria Optimization

Claus Weihs[1], Heike Trautmann[2]

[1] Dortmund University, Department of Computational Statistics
44221 Dortmund, Germany
weihs@statistik.uni-dortmund.de

[2] Dortmund University, Department of Computational Statistics
44221 Dortmund, Germany
trautmann@statistik.uni-dortmund.de

Abstract. In this paper parallel universes are defined by their relation to multi-criteria optimization combined with an explicit or implicit link for the unambiguous identification of an optimum. As an explicit link function the desirability index is introduced. Desirabilities are also used for restricting the Pareto set to desired parts.

Keywords. Parallel Universes, Multi-Criteria Optimization, Pareto Optimization, Desirability Index, Desirability Functions

1 Parallel Universes

In this paper it is shown how multi-criteria optimization (MCO) fits into the possible notion of parallel universes. We think of one universe as one view of the universe, i.e. as one quality of the universe represented by one response. Analogously we think of parallel universes as many views of the universe, i.e. many qualities or many criteria connected by a link for comparison of outcomes of the individual criteria. This link appears to be an important feature of parallel universes since without it the single universes could well be studied independently. Moreover, for this paper we assume that all criteria depend (at least in principle) on the same set of influential factors, and all criteria are optimized in terms of (the same) influential factors. If the criteria depend on different sets of factors one could just combine these factor sets to one joint set and imagine that all criteria in principle depend on all these factors. The notion of parallel universes appears to be important only when optima corresponding to different criteria are inconsistent, and one is interested in something like a compromise to resolve conflict. This could be done by means of the link criterion, or by assessing the multivariate 'optimum candidates' to identify (a set of) joint (Pareto) optima and afterwards evaluate them by means of external information. This would then relate to an unformalized link. But let us discuss all these in more detail, including a formal definition for parallel universes.

Dagstuhl Seminar Proceedings 07181
Parallel Universes and Local Patterns
http://drops.dagstuhl.de/opus/volltexte/2007/1255

2 Expert knowledge

Expert knowledge is most important not only for the definition of criteria for the parallel universes, but also for the definition of the link between the universes. Prior knowledge may be expressed in an explicit link criterion before optimization. This gives the opportunity to replace the multivariate optimization problem by a univariate one leading (hopefully) to an unambiguous solution. Examples for such explicit link functions are desirability indices to be defined later, and the well-known utility functions.
In multivariate optimization, posterior knowledge may be expressed by criteria for the evaluation of the elements of a set of optimal solutions. One could possibly judge a set of Pareto optima after optimization based on expert knowledge in order to select the most appropriate optimum.
In this paper we even suggest a third possibility combining prior and posterior knowledge. We first assess the criterion values by means of so-called desirabilities, this way restricting the set of Pareto solutions, and then select from this restricted set by means of posterior knowledge.

3 Multi-Criteria Optimization (MCO)

Definition: Standard MCO Problem for the minimization case
The (error free) Standard MCO Problem is defined by:
Minimize $Y = f(X) = (f_1(X), \ldots, f_k(X))$
with $Y = (Y_1, \ldots, Y_k) \in \mathcal{Y}$, quality criteria
$X = (X_1, \ldots, X_n) \in \mathcal{X}$, influential factors
$f_i, \quad i = 1, \ldots, k.$ mathematical models

Note that criteria are modelled here as functions of influential factors without noise. However, model extensions dealing with noise are possible.

Examples are
Y_1 = distance from optimum red color portion, Y_2 = distance from optimum green color portion, X = spectrum, or
Y_1 = error, Y_2 = - interpretability, where X = (observable factors, class), or
Y_i = distance to target in parallel universe $i = 1, 2$.

Possible problems in a MCO are non-comparable optima because of non-comparable scales of criteria, not well-defined order in criteria space $\mathbb{R}^k$, leading to many non-comparable optimum quality criteria vectors.

4 Parallel Universes: Definition

Let us now study parallel universes more formally.

Definition: Parallel universes
A parallel universe is a 4-tuple (X, Y, f, L) where

with	$X = (X_1, \ldots, X_n) \in \mathcal{X}$,	influential factors
	$Y = (Y_1, \ldots, Y_k) \in \mathcal{Y}$,	quality criteria
	$f = (f_1, \ldots, f_k)$,	mathematical models
	$L : \mathbb{R}^k \to \mathbb{R}$.	link function.

The link function is applied to the quality criteria which are functions of the influential factors. The link function is typically minimized leading to a standard MCO problem with link function:
Minimize $L(Y) = L(f(X)) = L(f_1(X), \ldots, f_k(X)))$.

An example for the link function L is the negative of a desirability index D defined in the next section.

5 A-priori link: Desirability Index

A desirability index is defined in five steps (cp. Trautmann, Weihs (2006), Trautmann (2004)):

- The influential factors are fixed: $X_1, \ldots, X_n$
- The quality criteria $Y_1, \ldots, Y_k$ and their dependence on the influential factors are identified: $Y_i = f_i(X_1, \ldots, X_n, \varepsilon_i)$
- The Desirability Function mapping the values of the quality criteria to the interval [0,1] is fixed: $d_i(Y_i)(i = 1, \ldots, k)$, $d : \mathbb{R} \to [0, 1]$ bzw. $(0, 1]$
- A summary measure called Desirability Index (DI) is fixed linking the individual desirabilities to one univariate quality measure: $D := f(d_1, \ldots, d_k)$, $D : [0, 1]^k/(0, 1]^k \to [0, 1]/(0, 1)$
- DI is maximized with respect to the influential factors, e.g.: $\min_{X_1,\ldots,X_n} -\hat{D}_g(X_1, \ldots, X_n) = \min_{X_1,\ldots,X_n} - \sqrt[k]{\prod_{i=1}^k d_i(f_i(X_1, ..., X_n, 0))}$

We now give two prominent examples for desirability functions. Harrington (1965) defined the desirability function in two-sided (target) case (cp. Figure 1):

$$d_i(Y_i') = \exp(-|Y_i'|^{n_i}), Y_i' = \frac{2Y_i - (USL_i + LSL_i)}{USL_i - LSL_i}$$

Derringer, Suich (1980) defined in the two-sided case (cp. Figure 2):

$$d_i(Y_i) = \left\{ \begin{array}{ll} 0, & Y_i < LSL_i \\ (\frac{Y_i - LSL_i}{T_i - LSL_i})^{l_i}, & LSL_i \le Y_i \le T_i \\ (\frac{Y_i - USL_i}{T_i - USL_i})^{r_i}, & T_i < Y_i \le USL_i \\ 0, & Y_i > USL_i \end{array} \right\}.$$

The corresponding one-sided case look as follows. The Harrington functions are defined as (cp. Fig. 3):

$$d_i(Y_i') = \exp(-\exp(-|Y_i'|)), \quad Y_i' = b_{0i} + b_{1i}Y_i.$$

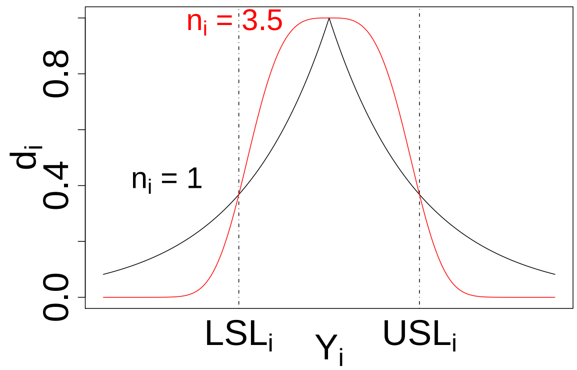

Fig. 1. Harrington desirability: two-sided case

Here, you only have to specify pairs (y_i, d_i), i = 1,2.
The Derringer/Suich functions have the form (cp. Fig. 4):

$$d_i(Y_i) = \begin{cases} 0, & Y_i \leq T_i \\ (\frac{Y_i - USL_i}{T_i - USL_i})^{r_i}, & T_i < Y_i < USL_i \\ 1, & Y_i \geq USL_i \end{cases}$$

Alternative specifications of the Desirability Index are:

$$D_g := (\prod_{i=1}^{k} d_i)^{1/k}, \qquad D_p := \prod_{i=1}^{k} d_i, \qquad D_{min} := \min_{i=1,\dots,k} d_i, \qquad D_m := 1/k \sum_{i=1}^{k} d_i.$$

6 A-posteriori link: Pareto principle

If no explicit link function is specified like a desirability index, in multivariate optimization posterior knowledge may be expressed for the evaluation of the elements of a set of optimal solutions. One could possibly assess a set of Pareto optima based on expert knowledge in order to select the most appropriate optimum.

Definition: Pareto optimality
A value of a vector of quality criteria (QC) $Y = (Y_1, \dots, Y_k)'$ is **Pareto optimal** iff there is no other value where process quality is improved for at least one criterion and not diminished for all other criteria. The **Pareto front** gives the points in the target space where the process cannot be improved in one quality without worsening another.
A vector of factor values $X = (X_1, \dots, X_n)'$ is Pareto optimal iff the corresponding vector of quality criteria $Y = f(X_1, \dots, X_n)$ is Pareto optimal.

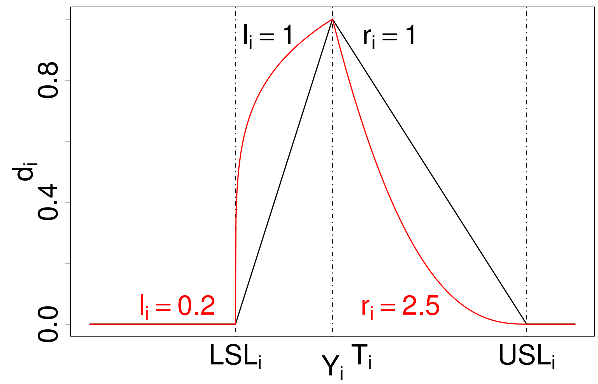

Fig. 2. Derrington/Suich desirability: two-sided case

This set of vectors is called **Pareto set**.

The problem is that in general there is a variety of Pareto optimal locations. Therefore, an a-posteriori link method has to integrate expert knowledge a-posteriori for the selection of (set of possibly) realized optimal factor vector(s).

7 Combination of prior and posterior expert knowledge

Instead of the standard MCO problem (for minimization) we propose to solve the, what we call, Desirability MCO problem (for minimization) (cp. Mehnen, Trautmann (2006)).

Definition: Desirability MCO problem for minimization
The desirability MCO problem is defined as:

Minimize $-d(Y) = -d[f(X)] = -(d_1[f_1(X)], \ldots, d_k[f_k(X)])$

with

$Y = (Y_1, \ldots, Y_k) \in \mathcal{Y}$,		quality criteria
$X = (X_1, \ldots, X_n) \in \mathcal{X}$,		influential factors
$d_i \in [0, 1]$,	$i = 1, \ldots, k$,	desirability functions
f_i,	$i = 1, \ldots, k$.	mathematical models

8 Pareto Optimality

Let uns now consider Pareto optimality for the different ways of linkage between the parallel universes. In the case of prior knowledge linkage by means of desirability indices the situation is optimal since optimal factor values are Pareto

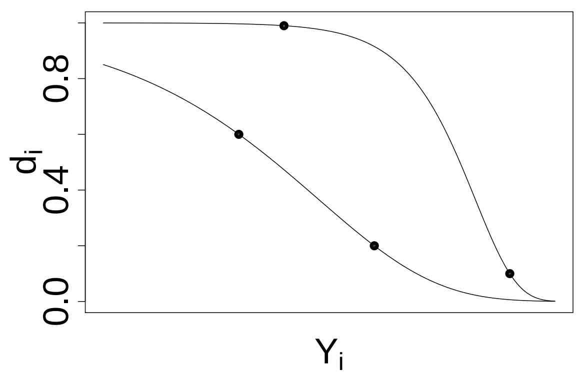

Fig. 3. Harrington desirability: one-sided case

optimal.

Theorem: Optimal factor values found by $\mathbf{D_g}$ are Desirability Pareto optimal.
Proof: Given optimal factor values $X^{opt} = (X_1^{opt}, \ldots, X_n^{opt})'$ derived from $D_g := (\prod_{i=1}^{k} d_i)^{1/k}$ and desirability functions d_i.
Assume X^{opt} is not Desirability Pareto optimal, then

$$\begin{aligned} \exists\, X^* :\ & d_i(Y_i|X^*) > d_i(Y_i|X^{opt}) \ \text{ for } i \in \{1, \ldots, k\} \\ \text{and}\ \ & d_j(Y_j|X^*) \geq d_j(Y_j|X^{opt}) \text{ for } j = 1, \ldots, k; j \neq i. \end{aligned}$$

Thus,

$$D^* = (\prod_{i=1}^{k} d_i(Y_i|X^*))^{1/k} > D^{opt} = (\prod_{i=1}^{k} d_i(Y_i|X^{opt}))^{1/k}. \quad \text{Contradiction!}$$

Therefore, DI can be seen as a method for selection of Pareto optimal factor values.
In the case of the Desirability MCO and one-sided desirability functions of Harrington type (minimization) it is true that Pareto optimal solutions of the desirability MCO problem are Pareto optimal for the standard MCO problem as well. Therefore, desirability functions select relevant local sets of Pareto front and Pareto set! Unfortunately, this is not valid anymore, if at least one two-sided desirability function is involved. However, choosing such a desirability MCO problem this is consciously accepted!
Let us look at some examples.

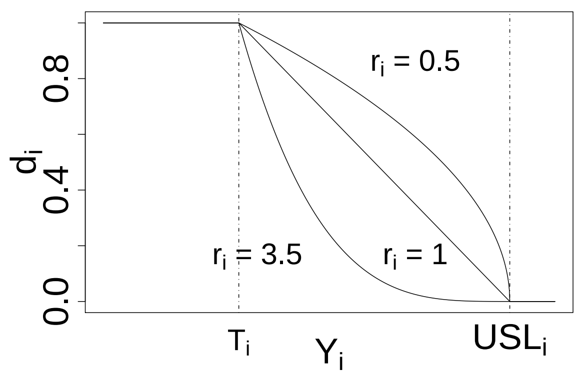

Fig. 4. Derringer/Suich desirability: one-sided case

9 Examples

Let us consider the so-called Binh problem (cp. Binh (1999)):

$$\begin{aligned} \text{Minimize} \;\; & f_1(x_1, x_2) = x_1^2 + x_2^2 \\ \text{Minimize} \;\; & f_2(x_1, x_2) = (x_1 - 5)^2 + (x_2 - 5)^2 \\ \text{with} \;\; & -5 \leq x_1 \leq 10, \quad -5 \leq x_2 \leq 10. \end{aligned}$$

For this example, the Pareto front and the Pareto set are shown in Figure 5.

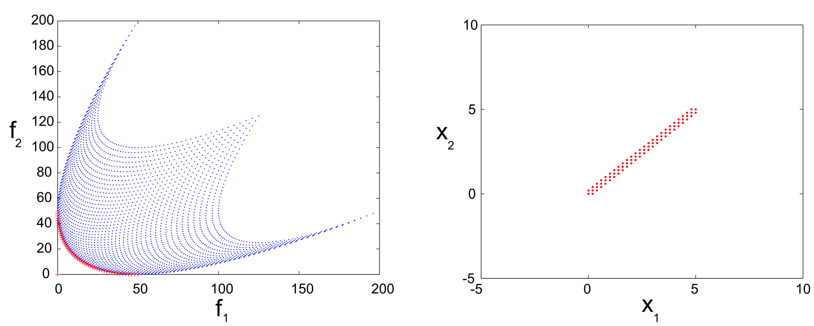

Fig. 5. Binh problem: Pareto front (left) and Pareto set (right)

An analogous Desirability MCO problem could look like:

$$
\begin{aligned}
\text{Minimize} \quad & -d_1(f_1(x_1,x_2)) = -d_1(x_1^2 + x_2^2) \\
\text{Minimize} \quad & -d_2(f_2(x_1,x_2)) = -d_2((x_1-5)^2 + (x_2-5)^2) \\
\text{with} \quad & -5 \le x_1 \le 10, \quad -5 \le x_2 \le 10 \\
\text{and} \quad & y_1^{(1)} = 20, d_1^{(1)} = 0.9, y_1^{(2)} = 50, d_1^{(2)} = 10^{-2}, \\
& y_2^{(1)} = 0.0, d_2^{(1)} = 1 - 10^{-2}, y_2^{(2)} = 10, d_2^{(2)} = 10^{-2}
\end{aligned}
$$

For this problem Pareto front and Pareto set can be found in Figure 6. Obviously, desirabilities lead to the restriction of both Pareto front and Pareto set.

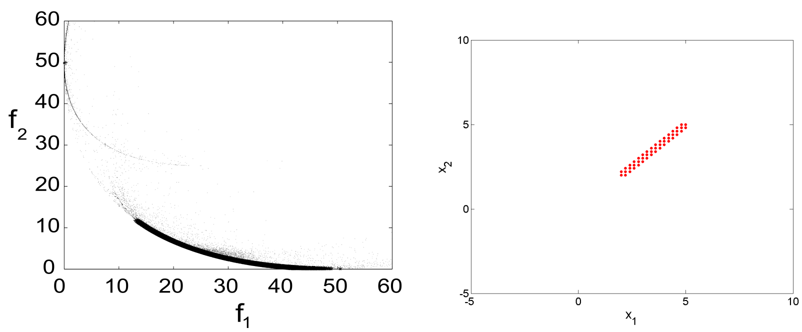

Fig. 6. Binh problem with desirabilities: Pareto front (left) and Pareto set (right)

In order to give you an impression what would happen, if desirability functions are two-sided in Figure 7 examples are given (first row) together with their corresponding Pareto fronts showing that also points are optimal now which were non optimal before transformation by desirabilities.

10 Conclusions

We defined Parallel Universes as Multi-Criteria Optimization in combination with a Link Function. Linkage was allowed to take place a-priori by means of a desirability Index, a-posteriori by means of expert selection from the Pareto set, or by a combination of prior and posterior linkage using desirability functions for each criterion. The latter procedure leads to a restriction of the Pareto front iff desirability functions are one-sided, but to (totally) different Pareto fronts and sets, if there is at least one two-sided desirability function involved. In all three ways, linkage has the aim of identifying an adequate optimum from the Pareto set.

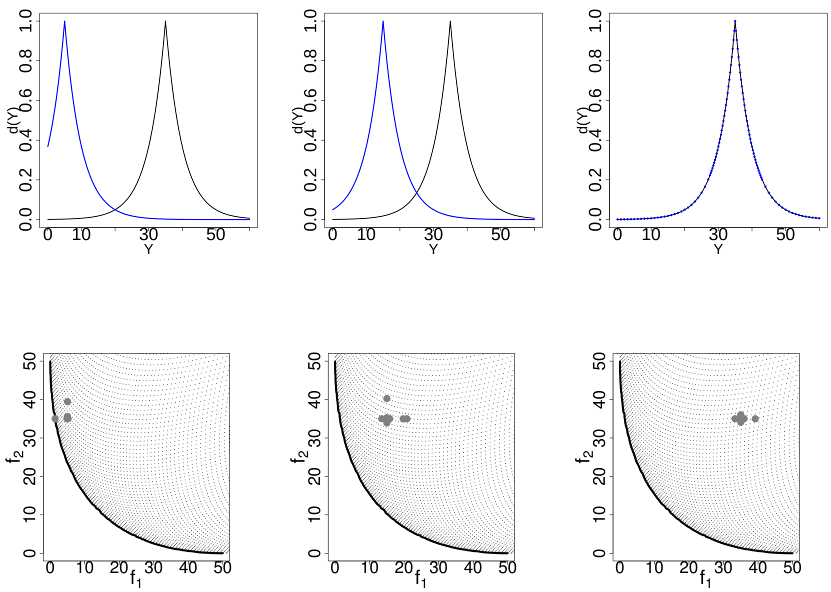

Fig. 7. Two-sided desirability functions (1st row), and corresponding 'Pareto fronts' (2nd row)

References

1. Binh, T.T. (1999): A Multiobjective Evolutionary Algorithm: The Study Cases. Technical report, Institute for Automation and Communication, Berleben, Germany

2. Derringer, G.C., Suich, D. (1980): Simultaeneous optimization of several response variables; Journal of Quality Technology 12, 214-219

3. Harrington, J. (1965): The desirability function; Industrial Quality Control, 21/1, 494-498

4. Mehnen, J., Trautmann, H. (2006): Integration of expert's preferences in pareto optimization by desirability function techniques. In R. Teti (Ed.), Proceedings of the 5th CIRP International Seminar on Intelligent Computation in Manufacturing Engineering (CIRP ICME '06), 293-298, Ischia, Italy

5. Trautmann, H., Weihs, C. (2006): On the distribution of the desirability function index using Harrington's desirability function; Metrika, 62/2, 207-213

6. Trautmann, H. (2004): Qualitätskontrolle in der Industrie anhand von Kontrollkarten für Wünschbarkeitsfunktionen; PhD Thesis, University of Dortmund