

Multi-Objective Optimization

Dante Piotto

fall semester 2023

Contents

1	Multi-Objective Optimization	5
1.1	Domination	5
1.2	Relations derived from domination	6
2	Scalar methods for M-o Optimization	9
2.1	A priori articulation methods	9
2.1.1	weighted-sum-of-objective-functions method	9
2.1.2	Distance-to-a-reference-objective method	9
2.1.3	Goal attainment method	10
2.1.4	Goal programming method	11
2.1.5	Lexicographic method	11
2.1.6	Compromise method	11
2.1.7	Hybrid method	12
2.2	A posteriori articulation of preferences	12
2.2.1	Normal boundary intersection method	12
2.2.2	Min-max method	12

Chapter 1

Multi-Objective Optimization

Optimization problem:

$$\begin{aligned} \min & J(x) \\ \text{subj. to} & g(x) \leq 0 \\ & h(x) = 0 \end{aligned}$$

with $x \in \mathbb{R}^n$, $J \in \mathbb{R}^k$.

1.1 Domination

Definition 1: Pareto optimal points

A point c_0 in the attainable set A is Pareto optimal iff

$$\nexists c \in A \text{ s.t. } c_i \leq c_{0i} \forall i \text{ and } c_i < c_{0i} \text{ for at least one } i$$

The set of Pareto optimal points is the **Pareto set**

A point in the design space is a Pareto point if no feasible point exists that reduces one criterion without increasing the value of some other ones

Definition 2: Domination relation

A vector $J(x_1)$ dominates a vector $J(x_2)$ if:

- $J(x_1)$ is at least as good as $J(x_2)$ for all the objectives
- $J(x_1)$ is strictly better than $J(x_2)$ for at least one objective

Definition 3: Local optimality in the Pareto sense

A vector $J(x)$ is locally optimal in the Pareto sense if there exists $\delta > 0$ such that there is no vector $J(x')$ which dominates $J(x)$, with

$$x' \in \mathbb{R}^n \cap B(x, \delta)$$

where $B(x, \delta)$ is a ball of center x and radius δ

Definition 4: Global optimality in the Pareto sense

A vector $J(\bar{x})$ is globally optimal in the Pareto sense (or optimal in the Pareto sense) if there does not exist any $J(x')$ that dominates $J(\bar{x})$

Definition 5: Negative cone

A negative cone is defined in the following way:

$$C^- = \{x \mid J(x) \in \mathbb{R}^k \text{ and } J(x) \leq 0\}$$

Theorem 1: Contact Theorem

vector x is optimal in the Pareto sense for a multi-objective optimisation problem if

$$(C^- + x) \cap F = x$$

where F corresponds to the feasible subspace

1.2 Relations derived from domination

Definition 6: Optimality in the lexicographic sense

A solution $J^* \in S^k$ is optimal in the lexicographic sense if

$$J^* \leq_{\text{lex}} J, \quad \forall J \in S^k \setminus \{J^*\}$$

where

$$J \leq_{\text{lex}} J', \quad \forall J, J' \in S^k$$

if there exists an index value q^* such that

$$J_q = J'_q, \forall q \mid 1 \leq q \leq q^* - 1 \text{ and } J_{q^*} < J'_{q^*}$$

The definition involves the requirement that the decision maker has **sorted** the objective functions by decreasing preference

Definition 7: Extremal optimality

A solution $J^* \in S^k$ is extreme optimal if, given a weight vector

$$\lambda \in \mathbb{R}^k, \quad \text{such that } \lambda_i \geq 0 \text{ and } \sum_i \lambda_i = 1$$

J^* is an optimal solution of the mono-criterion minimisation problem with the following objective function:

$$\sum_{i=1}^k \lambda_i \cdot J_i$$

which means that

$$\sum_{i=1}^k \lambda_i \cdot J_i^* \leq \sum_{i=1}^k \lambda_i \cdot J_i, \quad \forall J \in S^k \setminus \{J^*\}$$

This relation allows to add preferences between objective functions by using weights

Definition 8: Maximal optimality

A solution $J^* \in S^k$ is max-optimal if the value of the worst objective is as small as possible, i.e.:

$$\max_{q \in \{1, \dots, k\}} J_q^* \leq \max_{q \in \{1, \dots, k\}} J, \quad \forall J \in S^k \setminus \{J^*\}$$

This relation does **not** allow one to add preferences between objective functions.

Definition 9: Tradeoff surface

The set of globally Pareto optimal points constitutes the **tradeoff surface** or **Pareto front**

Definition 10: Characteristic points

- The point with coordinates obtained by minimising each objective function separately is called **Ideal point**
- The point with coordinates corresponding to the worst values obtained for each objective function when the solution set is restricted to the tradeoff surface is called **Nadir point**

Chapter 2

Scalar methods for M-o Optimization

2.1 A priori articulation methods

2.1.1 weighted-sum-of-objective-functions method

Take the weighted sum of objective functions

$$\begin{aligned} \min J_{eq}(x) &= \sum_{i=1}^k w_i \cdot J_i(x) \\ \text{subj. to } g(x) &\leq 0 \\ h(x) &= 0 \end{aligned}$$

Where $w_i \geq 0$. Let us consider a case with two objective functions:

$$J_{eq}(x) = w_1 J_1(x) + w_2 J_2(x)$$

Which is the expression of a line in the J_1, J_2 plane. Minimising $J_{eq}(x)$ is equivalent to looking for a constant C of the following line equation, which must be as small as possible:

$$J_2(x) = -\frac{w_1}{w_2} J_1(x) + C$$

This method consists of finding lines tangent to the set S . If the process is repeated for several weight values, the solutions found produce a tradeoff surface, **only** when the set S is convex.

2.1.2 Distance-to-a-reference-objective method

The new mono-objective function corresponds to a "distance". Let us introduce the vector F , whose k components correspond to an ideal objective, in a sense defined by the decision maker.

Definition 11: Absolute distance

$$L_r(J(x)) = \left[\sum_{i=1}^k |F_i - J_i(x)|^r \right]^{\frac{1}{r}}$$

with $1 \leq r < \infty$

when $r \rightarrow \infty$, the multi-objective method is called the **min-max method** (or the Tchebichev method), and

$$L_\infty(J(x)) = \max_{i \in \{1, \dots, k\}} |F_i - J_i(x)|$$

On the other hand, the relative distance is defined as

Definition 12: Relative distance

$$L_r^{\text{Rel}}(J(x)) = \left[\sum_{i=1}^k \left| \frac{F_i - J_i(x)}{F_i} \right|^r \right]^{\frac{1}{r}}$$

with $1 \leq r < \infty$

If the L_2 distance is used, the following problem is obtained:

$$\begin{aligned} \min J_{eq}(x) &= \sqrt{\sum_{i=1}^k [F_i - J_i(x)]^2} \\ \text{subj. to } g(x) &\leq 0 \\ h(x) &= 0 \end{aligned}$$

When these methods are used to search for Pareto optimal solutions, it is not necessary to include the square root operator. The choice of a reference point has a lot of influence on the quality of the result. This method allows us to approximate the tradeoff surface, but the surface is "seen" from a reference point.

To plot an iso-value curve, we proceed in the same way as with the weighted-sum-of-objective-functions method. We have to find x such that C in the following expression is minimal:

$$C = w_1[F_1 - J_1(x)]^2 + w_2[F_2 - J_2(x)]^2$$

So, the goal is to find the smallest C such that the ellipse is "tangential" to S

In the min-max method (with weights), the iso-value curve is simpler:

$$\begin{cases} J_1(x) = C & \text{if } w_1[J_1(x) - F_1] > w_2[J_2(x) - F_2] \\ J_2(x) = C & \text{if } w_1[J_1(x) - F_1] < w_2[J_2(x) - F_2] \end{cases}$$

A good reference point must be found, however this method allows to discover solutions hidden inside concavities under some conditions.

2.1.3 Goal attainment method

An equivalent formulation to the weighted min-max method is to minimise a scalar coefficient λ which represents the gap relative to the initial objective F along a search direction w

$$\begin{aligned} &\text{minimise } \lambda \\ &\text{with } J_1(x) - w_1 \cdot \lambda \leq F_1 \\ &\quad \vdots \\ &\quad J_k(x) - w_k \cdot \lambda \leq F_k \\ &\quad g(x) \leq 0 \\ &\quad h(x) = 0 \end{aligned}$$

To obtain the tradeoff surface, choose several search directions. This method is quite sensitive to the choice of the weights, with the weights that do not reflect any preference as with the weighted-sum-of-objective-functions method.

- the weights do not have any physical significance
- In general, a two-pass procedure has to be elaborated
- The main advantage remains that this method allows to work with a set of solutions S which is not necessarily convex

2.1.4 Goal programming method

The approach is the following:

- Choose an initial vector of objective functions F
- With each objective, associate two new "slack" variables related to the initial vector of objective functions
- Minimise one of the two variables, with the choice of the variable made by using the desired type of overstep (over or under the objective)

$$\begin{aligned}
 & \text{minimise } (d_1^+ \text{ or } d_1^-, \dots, d_k^+ \text{ or } d_k^-) \\
 & \text{with } J_1(x) = F_1 + d_1^+ - d_1^- \qquad \qquad \qquad \vdots \\
 & \qquad \qquad J_k(x) = F_k + d_k^+ - d_k^- \qquad \qquad \qquad g(x) \leq 0 \\
 & \qquad \qquad h(x) = 0
 \end{aligned}$$

where $d_i^+, d_i^- \geq 0$ and $d_i^+ \cdot d_i^- = 0$

This method transforms a Mo-O problem into the minimisation of a vector, and to minimise this vector, we can proceed in various ways:

- Minimise a weighted sum of slack variables. The weights define the preferences relative to the objective functions
- minimise lexicographically the coordinates of the vector. For example, we can proceed by the following steps:
 - Minimise d_1^+
 - Minimise d_2^- while keeping d_1^+ constant
 - Minimise d_3^+ while keeping d_1^+ and d_2^- constant
 - Minimise $d_4^+ + d_4^-$ while keeping d_1^+ , d_2^- and d_3^+ constant

The goal programming method is subject to the same criticisms as the goal attainment method, namely that in some cases it can fail to discover solutions hidden in concavities.

2.1.5 Lexicographic method

This method consists in considering objective functions one after the other and minimising a mono-objective problem while the problem is completed gradually with constraints. This method has a drawback: the choice of the sequence of objective functions to be minimised is somewhat arbitrary. Two distinct lexicographic optimisations with distinct sequences of objective functions do not produce the same solution.

2.1.6 Compromise method

This method allows to transform a Mo-O problem into a mono-objective one with additional constraints as follows:

- Selection of an objective function to be optimised with high priority
- Definition of a vector of initial constraints
- The "main" objective function is conserved, while all other objective functions are transformed into inequality constraints

$$\begin{aligned}
 & \text{minimise } J_1(x) \\
 & \text{with } J_2(x) \leq \varepsilon_2 \\
 & \qquad \qquad \qquad \vdots \\
 & \qquad \qquad \qquad J_k(x) \leq \varepsilon_k \\
 & \qquad \qquad \qquad g(x) \leq 0 \\
 & \qquad \qquad \qquad h(x) = 0
 \end{aligned}$$

2.1.7 Hybrid method

Combination of several methods. The best known is the **Corley method**. It is composed of the weighted-sum-of-objective-functions and compromise methods.

$$\begin{aligned} & \text{minimise } \sum_{i=1}^k w_i \cdot J_i(x) \\ & \text{with } J_j(x) \leq \varepsilon_j \quad j = 1, \dots, k \\ & \quad g(x) \leq 0 \\ & \quad h(x) = 0 \end{aligned}$$

The "new" constraints allow to restrict the search space. This method allows to combine the advantages of two methods, so it is efficient in various kinds of optimisation problems, whether convex or not. Nevertheless, a difficulty arises due to the number of parameters to be tuned.

2.2 A posteriori articulation of preferences

2.2.1 Normal boundary intersection method

It was developed in response to limitations in the weighted-sum approach. This method allows to compute a regular distribution of Pareto optimal points, even with non-convex front of varying curvature.

$$\begin{aligned} & \max \beta \\ & \text{with } \alpha + \beta n = J(x) \\ & \quad g(x) \leq 0 \\ & \quad h(x) = 0 \end{aligned}$$

If α points are chosen uniformly along the **convex hull of individual minima (CHIM)** a set of fairly evenly distributed Pareto points can be reasonably expected

2.2.2 Min-max method

It is formulated as

$$\begin{aligned} & \min \max_{i=1, \dots, k} |J_i(x)| \\ & \text{with } g(x) \leq 0 \\ & \quad h(x) = 0 \end{aligned}$$

Even if the idea is simple, its implementation is not obvious, so it is preferable to rely on this equivalent formulation

$$\begin{aligned} & \min \beta \\ & \text{with } J_i(x) \leq \beta, \quad i = 1, \dots, k \\ & \quad g(x) \leq 0 \\ & \quad h(x) = 0 \end{aligned}$$