INTELLECTUAL DATA ANALYSIS IN AIRCRAFT DESIGN

V.A. Komarov[1], S.A. Piyavskiy[2]

[1] Samara National Research University, Samara, Russia
[2]Samara State Architectural University, Samara, Russia

Abstract. This article considers the use of confidence judgments method by decision-makers to analyze the information contained in large databases. The comparative analysis of passenger aircrafts shows that it allows flexibly and objectively allocating the most relevant information from the data array.

Keywords: data analysis, aircraft design, large databases, decision-makers, confident judgments.

Citation: Komarov VA, Piyavskiy SA. Intellectual data analysis in aircraft design. CEUR Workshop Proceedings, 2016; 1638: 873-881. DOI: 10.18287/1613-0073-2016-1638-873-881

Introduction

The emergence of large corporate databases opens up new prospects in the field of aircraft design, as well as in other subject fields of project activity. It becomes possible to estimate comprehensively, over a large number of characteristics, both quantitative and qualitative, the efficiency of different variants of design decisions against a background of a huge amount of analogues. It is also important to keep in mind that the project activity has largely heuristic nature, based on a combination of objective quantitative analysis within intuitive the designers' ideas, arising impulses of which may the expand and transfer the attention focus, and even change the design paradigm itself. The mechanism of using large databases for the design of complex, multi-function objects such as aircraft should be oriented towards these features.

In our opinion, it is advisable to use some advanced methods of complex decision theory, such as multi-criteria optimization, during the formation of such mechanism. One of the main advantages of these methods is that they provide an adequate active role of decision-makers along with the use of axiomatic approach to the information analysis. For all the variety of decision-making methods [1,2], a decision-making method under irremovable uncertainly [3,4] and confident judgments method (CJM) are the most efficient methods from this points of view. The article is aimed to demonstrate opportunities offered by the application of these techniques during intellectual data analysis. The considered examples are given with great simplifications.

Each object, denoted by y in the following, is described by a set of data which is useful to divide into two groups. In the first group it is advisable to include the data which determine how an object is arranged and which can be changed by decision-makers. In terms of non-linear mathematical programming, it is usually called design variables. The data, which include the object's behavior characteristics or properties, should be contained in the second group. These data are of interest for products' customers, as a rule, in the form of maximum and minimum values. Further, we will denote them as a particular optimal criterion f i (y). In most cases, the particular efficiency criteria for complex technical objects are contradictory, which generates the well-known problem of multi-disciplinary optimization.
In aviation, for example, two important characteristics are in such conflict: the aircraft weight and aerodynamic efficiency. Increasing the aerodynamic quality is achieved by the wing lengthening, but it increases its weight [7,8]. Introducing new non-dimensional load-carrying coefficient of structural perfection into consideration allows to carry out the optimization of aircraft appearance taking into account both weight and aerodynamic efficiency [9]. However, the design of new aircraft and, particularly, the development of technical specification for its creation, needs analysis and taking into account a variety of parameters, which can allow to predict the success of a new project by the consumer.

Rating Estimation Method

Let us consider a corporate database for the aircrafts as a set Y of objects $y \in Y$, which are characterized by m-dimensional vectors $f(y) = \{f^1(y), f^2(y), ..., f^m(y)\},\ y \in Y$. The components of these vectors are separate efficiency characteristics of the object, which are of interest from the viewpoint of decision-makers. For example, a passenger aircraft has the following characteristics considered from an operational point of view:

- Cruise speed, km/h
- Number of passengers, pers
- Flight range, km
- Service ceiling, m
- Runway length, m
- Minimum price in passenger version, million USD
- Maximum price in passenger version, million USD
- Starting year of manufacturing
- Number of built aircrafts
- Engine power, kgf
- Fuel capacity, l

Traditionally, the simplest way to analyze this data array is to sort by the values of any characteristics $f^j,\ j=1,...,m$. It allows to define the locations according to solution variants for this characteristic among analogues. However, since the solution efficiency, in general, is determined mainly by its characteristics, the analytical value of sorting is not high.

More powerful tool for intellectual analysis is the allocation of the total array of objects which are Pareto efficient. The object is considered as Pareto efficient if there do not exist at least one dominant object on the entire considered set. It means that any object according to the characteristics not worse, and at least one - better. Thus, among the aircrafts, whose characteristics are given in Table 1 (data are taken from [7] and other sources, partly modeled and have purely methodological nature), Pareto efficient are Boeing 737-200, Boeing 737-400, Boeing 737-500 and Boeing 737-200 Advance is not efficient as it is dominated by Boeing 737-500.

Table 1. Some characteristics of Boeing's aircrafts

Aircrafts	Flight range, km	Ceiling, m	Run way length, m	Engine thrust, kgf	Fuel capacity, l
Boeing 737-200 Advance	2960	10670	1830	15780	19535
Boeing 737-300	4670	10200	1940	19940	20105
Boeing 737-400	3870	11300	1920	21340	23825
Boeing 737-500	5550	11300	1530	18160	20105

The rigorous formulation of Pareto efficient object is the following:

$$\neg\exists\, \hat{y} \in Y : (\, f^{\,j}(\hat{y}) \le f^{\,j}(\bar{y}) \;\; j = 1,\dots,m) \wedge (\exists j \in \{1,\dots,m\} : f^{\,j}(\hat{y}) < f^{\,j}(\bar{y})\,))$$

Analysis of Pareto efficiency allows decision-makers to exclude obviously inefficient objects from consideration, but it does not provide information about how much the objects which remain in consideration, are relatively effective. It is necessary to use techniques that allow to proportion the comparative significance of individual objects from the position of a holistic estimation of their efficiency. It means that we need to find an adequate way of comparing the individual characteristics of objects S after which the object's comprehensive efficiency estimation $y \in Y$ is determined by purely mathematical way as $F(y) \equiv F_s(\,f(y)),\; y \in Y$. There are a number of established proportion methods in each subject field. In aviation, "fuel efficiency" and "weight efficiency", as well as several others are used as a complex criterion during aircraft's comparative estimation. The disadvantage of this approach is that the objective criteria are important, but they express only one property and are not universal. Therefore, the conclusions obtained with their use are questionable, since the use of other, to the same extend authoritative criterion, could lead to other conclusions.

The **universal construction methods** of criteria convolution are more reliable. The most famous of these is the linear convolution method, in which various characteristics are assigned numerical weight coefficients of relative importance. It is considered, that they can be obtained by averaging the opinion of many experts, involved by decision-makers for this purpose. Then

$$\bar{F}(f) = \sum_{j=1}^{m} \alpha^j f^j,\ \alpha^j \ge 0,\ j = 1,\ldots,m,\ \sum_{j=1}^{m} \alpha^j = 1,$$

where $\alpha^j,\ j = 1,\ldots,m$ - weight coefficients.

The use of this method cannot be recommended during the design of such important objects as an aircraft for two main reasons.

Confident judgments

Let us notice, that the decision-maker made two judgments by choosing it:

- First of all, exactly this kind of account method for uncertainty in the form of linear convolution is fully adequate for this decision-making task,
- Secondly, exactly the chosen experts, the examination organization and methods of expert opinion processing load to absolutely reliable values of weight coefficients.

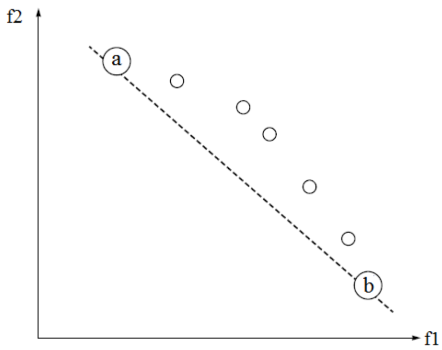

Fig. 1. Example of incorrect linear convolution

Both judgments can be challenged by reasonable positions. First of all, the linear convolution may not see some Pareto-optimal objects for any values of weight coefficients. For instance, on Figure 1 all objects for two minimized objects, images of which lie above the dotted line in a criterion space, will not be recognized as the most rational for any weight coefficient values in linear convolution, although they are Pareto-optimal objects [4]. Thus, this example shows that the use of the linear convolution penalizes a natural requirement for multiple comparison methods of individual object's characteristics *S*: any Pareto-optimal variant from the set of admissible solutions must correspond to at least one function $F_S(f) \in S$, the use of which provide the most rational solution. If this requirement is failure to comply, it reduces the select possibilities of decision-makers by purely mathematical features of the aircraft, which is unacceptable. The subjectivity of weight coefficients' determination by means of expert examination is evident. In addition, the need to attract qualified experts whenever the decision-maker wants to take a new look at the situation, greatly reduces the data analysis capabilities.

Actually, the decision-maker can reasonably make only two types of judgments.

The confident judgment of the first type. *Decision-maker person (DMP) with his confidence may include various particular criteria to different group of importance.* For example, "criteria 1 and 4 are the most important ones, criteria 2 and 6 are merely importance, and criterion 5 has the lowest importance". Let us note, that we do not assume that decision-maker provide a qualitative estimation of the relative importance degree for particular criteria, it refers only to the qualitative comparison which is optional.

The confident judgment of the second type. *If desired, the decision-maker can construct the pairs of Pareto-incomparable vectors of particular criteria, for which he is certain that one of the vectors is better that another.* It is not required that the vectors represent the efficiency of any real objects. If f_1 and f_2 – in which f_1 is surely better than f_2, it implements the following restriction on the set *S*:

$$S \subset \{s)\} : F_s(f_1) \leq F_s(f_2) \ \forall s \in S \ .$$

Based only on these two types of judgments, the method which proportion particular characteristics into a single numerical object's characteristic, represented in the database, was developed in [5,6]. It is called confident judgments method.

Stages of confident judgments' method

Stage 1. The uncertainty profile is constructed for the solving problem. It shows the range of complex efficiency criterion values for this decision within all possible ways to take into account the uncertainty for each design solution. The uncertainty profile is given by a pair of functions, which are defined on the set: minorant $m(y) = \min_{s \in S} F(f(y))$ and majorant$M(y) = \min_{s \in S} F\big(f(y)\big)$.

It should be noted, that obviously irrational decisions $z \in Y$, for which there are better solutions $\bar{z} \in Y$ by complex criterion in all possible ways of uncertainty, can be identified. The identify conditions for such solutions have the following form:

$$\exists \bar{z} \in Y : M(\bar{z}) < m(z) \ .$$

The main purpose of uncertainty profile is to give decision-makers information about the impact of uncertainty during decision-making in the problem. Adding confident judgments, he will be able to estimate how they reduce the uncertainty.

Stage 2. If it is possible, the set of uncertainties narrows by taking into account the decision-maker's confident judgments. The uncertainty accounting methods, which do not correspond to this judgments, are eliminated when we are using confident judgments of the first type. When we have got the confident judgments of the second type, conditions (6) are added to the set description, which eliminates those uncertainty accounting methods that do not carry out these judgments.

At the end of first two stages the initial set of uncertainties can be narrowed. This affect the uncertainty profile of the problem, but it is unlikely to remain only a single element in it, or all variants except one will be eliminated from the plurality of solutions. Thus, the uncertainty retained in the problem. This will be a fatal uncertainty. All uncertainty account methods, which form this fatal uncertainty, are completely equal to the decision-maker as he has already use his ability to make additional content in the problem description using judgments of the first and second types. It is possible that the other types of confident judgments of decision-makers can be found, but they did not fundamentally change the situation: after their usage, fatal uncertainty will remain in the problem.

Stage 3. Rigid and soft ratings for solution variants are calculated, taking into account unavoidable uncertainty. In order not to introduce the unnecessary for understanding and application complex mathematical apparatus, we shall assume that the set contains a finite number of uncertainty accounting methods S: S= $\{S_k\}_{k=1,\ldots,K}$.

Then the rigid rating $RG(y)$ for $y \in Y$ solution is a fraction of uncertainty accounting method, in which the solution is the best compared to the other solutions:

$$RG(y)=\frac{\sum_{k=1,\ F_k(y)\le F_k(z)\ \forall z\in Y}^{K} 1}{K},\ y\in Y$$

(If any uncertainty accounting method has several best solutions, we should write $\frac{1}{p}$ instead of 1 in the numerator of rigid rating).

Soft Rating *RM(y)* for $y \in Y$ decisions displays the average comprehensive efficiency if this solution compared with solutions, which are the best in different ways of uncertainty concideration:

$$RM = \frac{\sum_{k=1}^{K} \frac{F_{s_k}(f(y))}{\max\limits_{y\in Y} F_{s_k}(f(y))}}{K} .$$

Stage 4. Decision-maker recognize that the possibility of further uncertainty reducing is exhausted due to its confident judgments. Finally, he chooses a solution with the best (lowest) rigid rating as the most efficient solution. If there are several solutions, we will choose the one, which has the best (lowest) soft rating, as the most efficient solution.

Data analysis using confident judgments method

Let us show the use of confident judgments method for data analysis of passenger aircrafts in terms of their operational characteristics. Table 2 shows a fragment of the database.

Table 2. A fragment of database for passenger aircrafts

Aircrafts	Cruise speed, km/h	Passengers, pers	Flight range, km	Service ceiling, m	Take-off weight, t	Weight of empty aircraft, t	Runway length, m	Min. price in passengers variant, mil. USD	Max.price in passengers variant, mil. USD	Year of serial manufacturing	Aircraft Built in Total	Engine power, kgf	Fuel capacity, l
Boeing 737-200	905	120	2960	10670	49,40	27,17	1830	3,00	11,00	1967	3660	13160	10790
Boeing 737-200 Advance	905	120	3700	10670	58,00	31,90	1830	3,00	11,00	1984	865	15780	16250
Boeing 737-300	910	128	4670	10200	62,80	34,54	1940	10,50	44,00	1984	1102	19940	20105
Boeing 737-400	910	168	3870	11300	68,10	37,46	1920	18,50	48,00	1988	456	21340	23825
Boeing 737-500	910	108	5550	11300	60,55	33,30	1530	33,00	39,00	1990	385	18160	20105
Boeing 737-600	925	108	5910	12500	65,09	35,80	1880	32,00	39,00	1998	20	18160	26035
Boeing 737-700	925	128	5920	12500	69,40	38,17	2040	39,00	46,00	1997	15	21830	26035
Boeing r 737-800	925	189	5370	12500	78,24	43,03	2040	48,00	54,00	1998	20	23860	26035
Tu-204	850	210	3700	12600	94,60	52,03	1550	20,00	25,00	1994	15	32280	32000
Tu-204-100	850	210	5200	12600	103,00	56,65	1750	22,00	27,00	1995	20	32280	32000
Tu-204-120	850	210	5200	12100	103,00	56,65	1800	25,00	29,00	1997	26	39000	29900
Tu-204-200	850	210	6200	12100	110,75	60,91	2050	30,00	35,00	1996	23	32280	32000
Tu-204-300	850	210	3400	12600	86,00	47,30	2050	35,00	40,00	1997	22	32280	32000
Airbus Industry A319-110	900	124	4910	11275	68,00	37,40	1750	35,00	35,00	1996	15	21340	23860
Airbus Industry A321-200	900	185	5000	10676	89,00	48,95	2000	46,20	51,00	1996	1000	29960	23700

Analyzing this data, first of all, we will use one on the traditional comprehensive performance criteria – fuel efficiency. In this case, the only Pareto-optimal object is Tu-204-200, which rigid rating is equal to 100% (Column 2, Table 3). Its fuel efficiency is equal to 19,66 44 grams/pass*km, while the nearest Tu-204-120 it is 23.44 grams/pass*km. At the same time, we can use the other criteria – weight efficiency, which is calculated as the aircraft's ration of takeoff weight to the number of passengers. In this instance, the only Pareto-optimal variant with 100% rigid rating is another one object – Boeing Боинг 737-400 (Column 3 of Table 3), the weight efficiency of which is 0,41 t/pass, whereas Tu-204-120 has 0,49.

Table 3. The result of data analysis for passenger aircrafts

Aircraft	Rigid aircraft rating (%)			
	Complex criterion – fuel efficiency	Complex criterion – weight efficiency	CJM (Fuel and weight efficiency)	CJM (six characteristics distributed by three significance groups)
1	2	3	4	5
Boeing 737-200			4	
Boeing 737-200 ADVANCE				
Boeing 737-300				0,1
Boeing 737-400		100	2,9	8
Boeing 737-500				
Boeing 737-600				
Boeing 737-700				
Boeing 737-800			86,2	67,6
Tu-204				
Tu-204-100				
Tu-204-120				
Tu-204-200	100		4,1	
Tu-204-300				
Airbus Industry A319-110				
Airbus Industry A321-200			2,8	24,3
Total	100	100	100	100

Let us use confident judgments method to analyze the data. If we suppose that the decision-maker wants to use both of complex criteria, which were mentioned above, without giving preferences to any of them in order to organize data, we will receive the results shown in column 4 of Table 3. In this case, fire aircrafts are worthy of consideration (Pareto-optimal): Boeing 737-800, Tu-204-200, Boeing 737-200, Boeing 737-400 and Airbus Industry A321-200, while their efficiency was compared in relative scale. According to this, Boeing737-800 leads by a wide margin – its rigid rating is equal to 86,2%. Each of other listed aircrafts has only few percent of rigid rating.

However, applying confident judgments method, there is no need to bring subjectivity in the data analysis, coupled with the use of traditional complex criteria. It is enough to list the primary characteristics which are significant to the maintenance viewpoint. They are:

- Cruise speed,
- Number of passengers,
- Range,
- Minimum price in passenger variant,
- Maximum price in passenger variant,
- Fuel capacity.

Cruise speed and number of passengers are the most significant criteria. Taking into account the variety of routes for various distances, on which aircrafts are operated within its capabilities, the range and fuel capacity are the following on the importance. They also influence on the running costs, as they are transferred to the minimum and maximum ticket price of an aircraft. Thus, the criteria are distributed into three groups of significance. The results are shown in column 5 of Table 3. Boeing 737-800 saves leading positions, and Airbus Industry A321-200 follows it with a considerable margin.

Conclusion

Thus, the article shows that the application of confident judgments method for analysis of large databases opens new flexible opportunities for its users.

References

1. Larichev O. Theory and methods for decision-methods. Logos, 2000; 295.
2. Larichev O. Verbal decision analysis. ISI RAS, Science, 2006; 181.
3. Smirnov O, Padalko S, Piyabskiy S. CAD: the formation and functioning of project modules. Mechanical engineering, 1987; 272.
4. Malyshev V, Piyavskiy B, Piyavsliy S. Decision-making methods taking into account the variety of uncertainty conditions. Izvestiya RAS, Theory and control systems, 2001; 1: 46-61.
5. Piyavskiy S. Two new top-level concepts in the ontology of multi-criteria optimization. Designing ontology, 2013; 1(7): 65-68.
6. Malyshev V, Piyavskiy S. Confident judgments method for making multi-criteria decisions. RAS, Theory and control systems, 2015; 5: 90-101.
7. World passenger aircrafts. Argus, 1997; 336.