POST OPTIMAL UNCERTAINTY ASSESSMENT

Kalman Žiha
University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture
Ivana Luèiæa 5, 10000 Zagreb, Croatia
kalman.ziha@fsb.hr

Abstract: This paper considers the assessment of uncertainties of objectives and constraints in generally non-linear design analysis and in optimality problems. Some design variables and problem parameters, or even all of them, can be uncertain, either with given tolerance limits or with given statistical properties. The presented procedure can be applied to distinguish solutions with respect to uncertainties, or to indicate a range of solutions in a design or post optimal analysis in which no significant differences can be encountered due to model uncertainties.

Introduction

Due to a large amount of different uncertainties involved in many technical design and optimisation problems, there is often a scepticism about highly sophisticated and very accurate numerical procedures. In practice, there can be a wide range of solutions which cannot be mutualy precisely distinguished due to objective, subjective, numerical, operational and other uncertainties or inaccuracies involved in the design model.

Design analysis and optimisation problems are based on mathematical models defined by (possibly uncertain) free variables and model parameters generally denoted in the text by x. Free variables can be affected by the optimisation procedure, since the parameters are predefined in the global design procedure, and cannot be changed in a single optimisation.

Free variables and parameters are in general represented by their nominal values, denoted in the paper as N_X. In many design and optimisation problems the values of free variables and parameters are considered deterministic. However, the nominal values of variables and parameters must not coincide with their mathematical value. Constants are considered certain and can not be sobject of optimisation. The mathematical model can be uncertain by itself.

The values of variables and parameters x are often, especially in engineering problems, given within a certain tolerance, denoted in the text as t_X. The tolerance represents the bounds of acceptable uncertainties which can be controlled and usually represents the deviations from nominal values or mathematical values in a positive and/or a negative sense, denoted by t_X^{upp} and t_X^{low}. The amount of tolerance can also be expressed as fractions of the considered variable or parameter values.

Some free variables and parameters, or all of them, can be considered as random variables. In some cases, complete statistical information about random variables is available, but there is often the case that only the first two statistical moments, μ_X and σ_X, are known. The correlations between the variables are sometimes also available. The nominal value N_X sometimes coincides with the statistical mean, but in general it can be biased with respect to the mean value. Tolerance can in some problems be expressed in terms of standard deviations.

Stochastic programming can be applied for the solution of optimisation problems when random variables are involved, (e.g. Charnes, Cooper, 1959). The first two statistical moments of random variables can be used to assess the uncertainties of functions of random variables in a post optimal procedure (see e.g. Kapur and Lamberson, 1977).

The idea underlined in the paper is to investigate the effect of uncertainties in the non-linear design analysis and in the optimality problems on the objectives and constraints, with respect to tolerance or statistical properties of variables and parameters in a post-optimal procedure. Presented approach is illustrated by the numerical example.

1. The tolerance limits of non-linear functions

Let us consider a function $Y = f(x_1, x_2, \ldots, x_n)$. The first order Taylor's series expansion in a given linearisation point x^* yields to approximate function value in the adjacent point x, as:

$$Y = f(x_1, x_2, \ldots, x_n) = f(x_1^*, x_2^*, \ldots, x_n^*) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)_{x^*} \left(x_i - x_i^* \right) \qquad (1)$$

The tolerance assessment of linear functions is considered earlier (e.g. Mavriæ, 1990). If there are some uncertainties in the function value, they can be taken into account by additional (possibly subjective) tolerances t_f^{low} and t_f^{upp}. Consider first a non-linear function Y where each of the variables can be separated into "n" single derivable terms, (e.g. $Y=f(x_1,x_2)=x_1^5-x_2^2$). The tolerance limits of such a non-linear function of variables, given with their upper and lower tolerance, can be assessed in a given point X* as follows:

$$Y^{upp} = f(x_1^* + t_{x_1}, x_2^* + t_{x_2}, \ldots, x_n^* + t_{x_n}) + t_f^{upp}, where \begin{cases} t_{x_i} = t_{x_i}^{upp} \; if \; \dfrac{\partial f}{\partial x_i} > 0 \\ t_{x_i} = t_{x_i}^{low} \; if \; \dfrac{\partial f}{\partial x_i} < 0 \end{cases} \qquad (2a)$$

$$Y^{low} = f(x_1^* + t_{x_1}, x_2^* + t_{x_2}, \ldots, x_n^* + t_{x_n}) + t_f^{low}, where \begin{cases} t_{x_i} = t_{x_i}^{low} \; if \; \dfrac{\partial f}{\partial x_i} > 0 \\ t_{x_i} = t_{x_i}^{upp} \; if \; \dfrac{\partial f}{\partial x_i} < 0 \end{cases} \qquad (2b)$$

or approximately, under the same conditions, as follows:

$$Y^{upp} \approx f(x_1^*, x_2^*, \ldots, x_n^*) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)_{X^*} \cdot t_{x_i} \; ; \; Y^{low} \approx f(x_1^*, x_2^*, \ldots, x_n^*) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial x_i} \right)_{X^*} \cdot t_{x_i} \quad (3a\text{-}b)$$

If $\left. \frac{\partial f}{\partial x_i} \right|_{X*} = 0$ and $t_{x_i}^{upp} \neq t_{x_i}^{low}$, discontinuity in function tolerance limits in point X* occur.

The more general case is a non-linear function Y where "n" variables can be separated in more then "n" single derivable terms of a single variable, (e.g. $Y=f(x_1,x_2)=(x_1-x_2^2)(x_1-2x_2^2)$). The problem can be solved by introducing an additional notation for same variables contributing to different terms, (e.g. $Y=f({}_1x_1,{}_2x_1,{}_1x_2,{}_2x_2)=({}_1x_1-{}_1x_2^2)({}_2x_1-2{}_2x_2^2)$). The prefix in the above notation denotes the sequence number of considered terms. Such a function can be considered as a function of more then "n" variables, maximum $m_1+m_2+\ldots m_n$, as presented next:

$$Y = f({}_1x_1, {}_2x_1, \ldots, {}_{m_1}x_1, \ldots {}_1x_2, {}_1x_2, \ldots, {}_{m_2}x_2, \ldots, {}_1x_n, {}_2x_n, \ldots, {}_{m_n}x_n) \;\; where \;\; {}_jx_i = x_i \;\; for \; j = 1, 2, \ldots, m_i .$$

2. The statistical moments of non-linear functions of random variables

The statistical moments of the functions of random variables can be assessed by the **F**irst **O**rder Taylor's series expansion using up to the **S**econd **M**oments of the random variables. Such an approach is denoted earlier as FOSM (e.g. Madsen, Krenk, Lind, 1986.). The mean value and the variance can be assessed in a given linearisation point X^* (e.g. Žiha, 1987) as:

$$\mu_Y = f(X_1^*, X_2^*, \ldots X_n^*,) + \sum_{i=1}^{n} \left(\frac{\partial f}{\partial X_i} \right)_{X^*} \left(\mu_{x_i} - X_i^* \right) \qquad (4)$$

$$\sigma_Y^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \left(\frac{\partial f}{\partial X_i} \right)_{X^*} \left(\frac{\partial f}{\partial X_j} \right)_{X^*} \rho_{X_i X_j} \sigma_{X_i} \sigma_{X_j} \qquad (5)$$

where ρ in the upper equation is the correlation coefficient between the variables X_i and X_j. For statistically uncorrelated variables, the variance can be expressed as:

$$\sigma_Y^2 = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial X_i} \right)_{X^*}^2 \sigma_{X_i}^2 \qquad (6)$$

If the linearisation takes place in the mean value of the random variables X=μ, it yields to the amount of the mean value as follows: $\mu_Y = f(X_1 = \mu_1, X_2 = \mu_2, \dots X_n = \mu_n)$.
The variances in eqns. (5, 6) can be considered as objective (denoted σ_o^2). If there are some subjective uncertainties (denoted σ_s^2) they can be taken into consideration as follows:

$$\sigma_Y^2 = \sigma_o^2 + \sigma_s^2 \qquad (7)$$

3. Testing the tolerance limits and statistical dispersions of two solutions

In a problem considered as deterministic, the mathematical value, the nominal value and the tolerance limits can be assigned to objective functions and to constraint functions in each point. Using tolerance limits of the variables and parameters, the bounds on the resulting objective function can be obtained in two arbitrary points denoted as x^* and x^o, as follows: $f^{low}(x_1^o, x_2^o, \dots, x_n^o), \cdot f^{upp}(x_1^o, x_2^o, \dots, x_n^o), f^{low}(x_1^*, x_2^*, \dots, x_n^*), f^{upp}(x_1^*, x_2^*, \dots, x_n^*)$.
The function values f^* and f^o can be compared in points x^* and x^o, with respect to their tolerance fields in different ways, appropriate to the problem, e.g. overlapping tolerance fields:

$$s^{upp} \cdot f^{o\,upp} - s^{low} \cdot f^{*low} \geq 0 \;\; and \;\; s^{upp} \cdot f^{*\,upp} - s^{low} \cdot f^{o\,low} \geq 0 \qquad (8)$$

Alternatively, the objective function value in the arbitrary point x* can be compared to tolerance limits on the objective function calculated in the reference point x^o, (e.g. in the optimal point or Paretto optimal point), according to eqns. (2 or 3), e.g. as follows:

$$s^{low} \cdot f^{low}(x_1^o, x_2^o, \dots, x_n^o) \leq f(x_1^*, x_2^*, \dots, x_n^*) \leq s^{upp} \cdot f^{upp}(x_1^o, x_2^o, \dots, x_n^o) \qquad (9)$$

In the same way, using eqns. (2 or 3), the bounds on constraints and bounds can be obtained.

$$p^{low} \cdot g^{low}(x_1^o, x_2^o, \dots, x_n^o) \leq g(x_1^*, x_2^*, \dots, x_n^*) \leq p^{upp} \cdot g^{upp}(x_1^o, x_2^o, \dots, x_n^o) \qquad (10)$$

The constraints in eqn. (10) can be defined as either of the inequality type $g(x_1^*, x_2^*, \dots, x_n^*) \leq 0$ or of the equality type $g(x_1^*, x_2^*, \dots, x_n^*) = 0$.
In eqns. (9, 10 and 11), s^{low}, s^{upp} , p^{low} and p^{upp} represent possible correction factors according to specific problem considerations.
Different selection criteria with respect to tolerances in constraint can be applied. For example, the objective function values within tolerance limits in the feasible domain or in the tolerable violated constraint domain can be accepted.
Considering the points x^o and x* as random variables, it is possible to associate each of the selected points with the nominal value, as well as the mean value μ_X and the variance σ^2_X.
If the mean values and variances calculated according to eqns. (4 to 8), can be considered characteristic of the population, the following test can be applied:

$$\mu^* - \mu^o < \frac{\sigma^* + \sigma^o}{r} \qquad (11)$$

where r in eqn. (11) represents a value selected to define the user's acceptance criteria.

4. Example

An illustrative minimisation problem with two variables $0 \leq x_1 \leq 4$, and $0 \leq x_2 \leq 2$, is considered. An objective function is defined as follows:

$$f(x_1,x_2)=A\cdot(x_1-a_1)^{\alpha}(x_2-a_2) \qquad (12)$$

The feasible domain is the interior of a circle given with the constraint function:

$$g(x_1,x_2)=(x_1-b_1)^2+(x_2-b_2)^2-R^2\le 0 \qquad (13)$$

The derivatives of the objective function in eqn. (12) with respect to the parameters and variables are as follows:

$$\frac{\partial f}{\partial A}=f\frac{1}{A};\ \frac{\partial f}{\partial \alpha}=f\cdot \ln(x_1-a_1);\ \frac{\partial f}{\partial a_1}=-f\frac{\alpha}{(x_1-a_1)};\ \frac{\partial f}{\partial a_2}=-f\frac{1}{(x_2-a_2)};\ \frac{\partial f}{\partial x_1}=f\frac{\alpha}{(x_1-a_1)};\ \frac{\partial f}{\partial x_2}=f\frac{1}{(x_2-a_2)} \qquad (14a\text{-}f)$$

The derivatives of the constraint function in eqn. (13) are as shown:

$$\frac{\partial g}{\partial b_1}=-2(x_1-b_1);\quad \frac{\partial g}{\partial b_2}=-2(x_2-b_2);\quad \frac{\partial g}{\partial R}=-2R;\quad \frac{\partial g}{\partial x_1}=2(x_1-b_1);\quad \frac{\partial g}{\partial x_2}=2(x_2-b_2) \qquad (15a\text{-}e)$$

4.1. Optimisation based on the nominal values

Following nominal values of the parameters in the example are given:

N_A=1.25, N_α=0.50, N_{a1}=1.50, N_{a2}=0.00, N_R=1.0, N_{b1}=2.82, N_{b2}=2.20.

Between the parameters and their nominal values the following relations are assumed:

A=0.8N_A, α=2N_α, a_1=0.66N_{a1}, a_2=N_{a2}, R=1.4142N_R, b_1=1.06N_{b1}, b_2=0.91N_{b2}

The corresponding parameter values are:

A=1.00, α=1.00, a_1=1.00, a_2=0.00, R=1.4142, b_1=3.00, b_2=2.00.

It is assumed that the variables and their nominal values are in following relations: x_1=0.80N_{x1}, x_2=1.25N_{x2}.

The optimal solution of the problem is in the point with mathematical co-ordinates x_1=2.00 and x_2=1.00. The nominal values are obtained by the inverse transformation as follows: N^o_{x1}=1.25x_1=2.50 and N^o_{x2}=0.8x_2=0.80. The corresponding objective function value is f(x_1=2, x_2=1)=1.00, see Fig. 1.

4.2. Post optimal uncertainty analysis

Following tolerance values in parameters and variables are taken:

t_A^{upp}=0.10, t_A^{low}= -0.08
t_α^{upp}=0.05, t_α^{low}= -0.04
t_{a1}^{upp}=0.08, t_{a1}^{low}= -0.03
t_{a2}^{pup}=0.06, t_{a2}^{low}= -0.0
t_R^{pup}=0.06, t_R^{low}= -0.04

t_{b1}^{upp}=0.08, t_{b1}^{low}= -0.05
t_{b2}^{upp}=0.04, t_{b2}^{low}= -0.02
The appropriate tolerances in variables are:
t_{x2}^{pup}=0.05, t_{x2}^{low}= -0.02
t_{x1}^{pup}=0.08, t_{x1}^{low}= -0.04

The mean values taken are the same as the parameter values:

μ_A=1.00, μ_α=1.00, μ_{a1}=1.00, μ_{a2}=0.00, μ_R=1.00, μ_{b1}=3.00, μ_{b2}=2.00.

In addition, following standard deviations are applied:

σ_A=0.06, σ_α=0.04, σ_{a1}=0.10, σ_{a2}=0.20, σ_R=0.15, σ_{b1}=0.45, σ_{b2}=0.30.

For design variables coefficient of variations are given as c.o.$_{x1}$=0.12 and c.o.v.$_{x2}$=0.08.

Substituting values x_1=2.00 and x_2=1.00 obtained in previous section as the optimal solution, to expressions for constraint function derivatives in eqns. (15), following numerical values are obtained: 2, 2, -2.8284, -2, -2, respectively.

The upper values of the constraint function in terms of the values and tolerances of the variables and parameters in optimal point according to eqn. (2) are as shown:

$$g^{upp}(x_1,x_2)=(x_1+t_{x_1}^{upp}-b_1-t_{b_1}^{low})^2+(x_2+t_{x_2}^{upp}-b_2-t_{b_2}^{low})^2-(R+t_R^{upp})^2 \qquad (16)$$

Substituting the tolerance values, the following constraint functions are obtained:

$$g^{upp}(x_1,x_2)=(x_1-2.87)^2+(x_2-1.93)^2-(R+0.06)^2=0 \qquad (17a)$$

The rest of the constraint functions is obtained by appropriate substitution as shown:

$$g^{upp}(x_1,x_2)=(x_1-2.87)^2+(x_2-2.07)^2-(R+0.06)^2=0 \qquad (17b)$$

$$g^{upp}(x_1,x_2) = (x_1 - 3.12)^2 + (x_2 - 1.97)^2 - (R + 0.06)^2 = 0 \qquad (17c)$$

$$g^{upp}(x_1,x_2) = (x_1 - 3.12)^2 + (x_2 - 2.06)^2 - (R + 0.06)^2 = 0 \qquad (17d)$$

Note: the constraint functions in eqns. (17a-d) are shifted circles with enlarged radii, Fig.1.
The lower values of the constraint function in the optimal point can be obtained as

$$g^{low}(x_1,x_2) = (x_1 + t_{x_1}^{low} - b_1 - t_{b_1}^{upp})^2 + (x_2 + t_{x_2}^{low} - b_2 - t_{b_2}^{upp})^2 - (R + t_R^{low})^2 \qquad (18)$$

and after substitution, it reads:

$$g^{low}(x_1,x_2) = (x_1 - 3.12)^2 + (x_2 - 2.06)^2 - (R - 0.04)^2 = 0 \qquad (19)$$

The lower values of the constraint functions are in this example irrelevant since they are anyway in the feasible domain.
Substituting values x_1=2.00 and x_2=1.00 to expressions of objective function derivatives in eqn. (14), obtains the following numerical values respectively: 1, -0, -1, -1, 1, 1.
The upper and lower values of the objective function in terms of the tolerance of the variables, using the derivatives of the objective function are, see eqn. (2):

$$f^{upp}(x_1,x_2) = (A + t_A^{upp})(x_1 + t_{x_1}^{upp} - a_1 - t_{a_1}^{low})^{(\alpha + t_\alpha^{upp})}(x_2 + t_{x_2}^{upp} - a_2 - t_{a_2}^{low}) \qquad (20a)$$

$$f^{low}(x_1,x_2) = (A + t_A^{low})(x_1 + t_{x_1}^{low} - a_1 - t_{a_1}^{upp})^{(\alpha + t_\alpha^{low})}(x_2 + t_{x_2}^{low} - a_2 - t_{a_2}^{upp}) \qquad (20b)$$

A. The upper and lower values of the objective function $f(x_1,x_2)$=1.0 calculated in the optimal point x_1=2.00 and x_2=1.00 according to eqns. (20a-b) are as follows: $f^{upp}(x_1,x_2)$=1.2151 and $f^{low}(x_1,x_2)$=0.7963, see Fig. 1.
The corresponding mean and the standard deviation of the objective function in the optimal point, according to eqns. (4, 6) as well as eqns. (14a-e) are μ_f=1.0 and σ_f=0.2793.
B. Let consider a point x_1=2.80 and x_2=0.90. The objective function value is $f(x_1,x_2)$=1.6200 and the tolerance limits according to eqns. (20a-b), are $f^{upp}(x_1,x_2)$=1.9313, $f^{low}(x_1,x_2)$=0.1.3473. Compared to the optimal value according to relations (9), there is no overlapping in the tolerance field, i.e. there is a gap of -0.1322, and the solution is considered as a different one.
The corresponding mean and the standard deviation of the objective function in the optimal point, according to eqns. (4 and 6) as well as eqns. (14a-e) are μ_f=1.6200 and σ_f=0.3585. According to relation (11) 1.62-1<(.3585-.2793)/2 and the solutions are considered different.
C. Let consider another point x_1=1.80 and x_2=1.50. The objective function value is $f(x_1,x_2)$=1.2000, and the tolerance limits according to eqns. (15a-b) are $f^{upp}(x_1,x_2)$=1.4971 and $f^{low}(x_1,x_2)$=0.9143. Compared to the optimal value according to relations (9), there is an overlapping in the tolerance field of 0.3008, and the solution is considered as indistinguishable.
The corresponding mean and the standard deviation of the objective function in the optimal point, according to eqns. (4 and 6) as well as eqns. (14a-e) are μ_f=1.2000 and σ_f=0.3771. According to relation (11) 1.2-1<(.3771-.2793)/2 and the solutions are indistinguishable.
D. If it is agreed that the acceptance of the solutions within the tolerance limits of the objective functions in the optimal point as the reference one according to relation (10), and if it is expected that the solutions are in the tolerated infeasible domain, the wide range of solutions in variables is available (shadowed in Fig. 1.). The tolerated constraint value violation in the nominal optimal solution x_1=2.00, x_2=1.00, see eqn. (12a), is $g(x_1$=2.00, x_2=1.00)=0.5628.
E. Finally, the optimal solution in the feasible domain, now enlarged for tolerance limits on constraint function, is in the point x_1^o=1.4612 and x_2^o=1.4957 (see Fig. 1.). The corresponding nominal values are N^{ot}_{x1}=1.8265 and N^{ot}_{x2}=1.19656. The corresponding minimal objective function value is $f(x_1$=1.4612, x_2=1.4957)=0.6898 (see Fig. 1.).

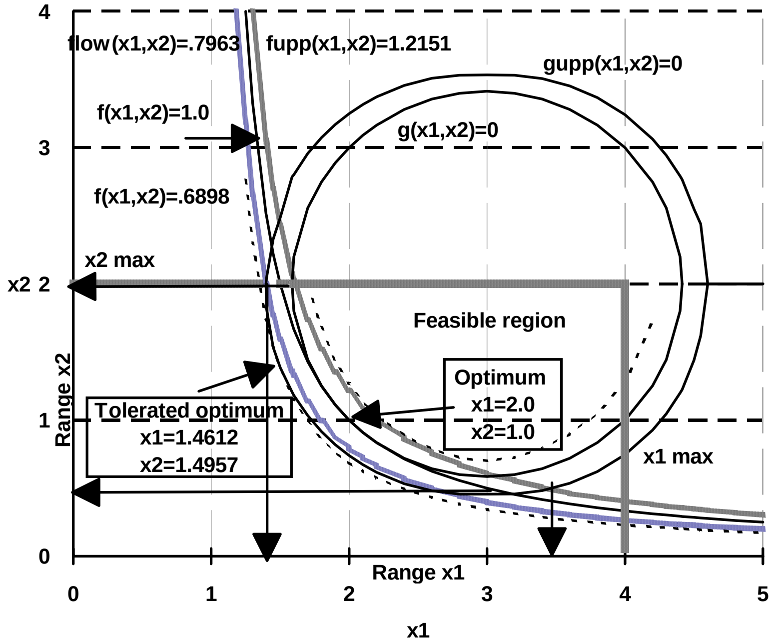

Fig. 1. Optimal solution and tolerance limits as detailed in example

Conclusion

Considering the upper and lower tolerance limits on design model or optimisation variables and parameters, basically defined as deterministic, the constraint and the objective functions can also be assessed by their tolerance limits. For non-linear functions, the tolerance limits can be determined using information from the first order Taylor's series expansion. Solutions to an optimisation problem with the tolerance limits around the optimal solution can be considered as a family of non-distinguishable, i.e. tolerable solutions.

If the optimisation variables and parameters are random variables, with known at least the first two statistical moments, the constraints and the objectives can be considered as functions of random variables. For the non-linear functions, the first two statistical moments of functions of random variables can be assessed using the first order Taylor's series expansion and the first two moments of random variables. The statistical significance of optimality criteria can be defined by a range of values, as established by the user's estimate in his significance assessments.

The analysis of the uncertainty in optimisation problems can indicate a wider range of solutions which can satisfy user requirement. Such an approach can give an adequate explanation to sceptics, that in uncertain conditions the strict, mathematically defined optimum cannot often be clearly distinguished from a number of other sub-optimal solutions. The approach presented in this paper allows for a relatively simple assessment of the dispersion of the solutions of an optimisation problem in a design and post-optimal consideration.

References

1. Charnes, A., Cooper,W.W.(1959),"Chance Constrained Programming", Management Sc., 6.
2. Kapur, K.C., Lamberson, L.R. (1977), Reliability in Engineering Design, Willey, N.Y.
3. Madsen, H, Krenk, S., Lind, N. (1986), Methods of Structural Safety, Prentice-Hall, N.J.
4. Mavric, I. (1990), Contribution to accuracy assessment in ship construction, PhDThesis (in Croatian), Faculty of Mechanical Engineering and Naval Architecture, Zagreb.

5. Žiha, K. (1987), "The Programming Facilities for Computation of Moments of Functions of Random Variables", Proc. 9th Int. Cong. Computer at the University, Cavtat.