1999年度日本オペレーションズ・リサーチ学会
春季研究発表会

Variance-Based Methods for Sensitivity Analysis

01105930 筑波大学 *香田正人 KODA Masato
koda@shako.sk.tsukuba.ac.jp

1. Introduction

Sensitivity analyses are now widely recognized as an essential tool of analysis for model outputs of complex systems (e.g., SAMO'95 and SAMO'98, and references therein). The sensitivity and uncertainty analyses are usually used to understand the behavior of a model, and the coherence between a model and an underlying system. In this study, we will concentrate on the new sensitivity analysis methods, which are referred to as the variance-based method. The method is applicable to model outputs generated from the generic computer codes which are typically used in data mining applications, e.g. decision and regression trees, etc. (Koda (1998)).

A global, nonlinear sensitivity analysis method called the Fourier Amplitude Sensitivity Test (FAST) has been available. The original FAST method was developed in the 70's by Cukier and co-workers (1973 ; 1978), and further improved by Koda, McRae, and Seinfeld (1979). The FAST method explores the multi-dimensional space of input parameters by a single search curve that sweeps throughout the entire parameter space and thus avoids a multi-dimensional (Monte Carlo) integration over the same sample space. It gives a quantitative sensitivity measure defined by a fractional contribution to variance computed from the terms in the Fourier expansion of the model output, and it is further extended recently (e.g., SAMO'98).

The FAST method belongs to the so-called variance-based techniques, whose sensitivity measure can be expressed as

$$\frac{Var_X\{E(Y\mid X)\}}{Var(Y)} \tag{1}$$

where Y denotes the model output and X the input parameter, $E(Y|X)$ denotes the expectation of Y conditional on a fixed value of X, and Var_X denotes the variance taken over all possible values of X. The statistical quantity (1) is well-known in the Design Of Experiment (DOE) study. In simulation, this is known as "what-if" analysis, and DOE uses regression analysis, also called as ANalysis Of VAriance (ANOVA). In ANOVA, the output variance is decomposed into partial variances of increasing order of dimensionality.

The Sobol' sensitivity indices, original extension of DOE by Sobol' (1990), are similar to FAST in the sense that the total variance of model output is assumed to be made up of correlation terms of increasing dimensionality. Hence, Sobol' indices may also provide a global quantitative sensitivity information. Both Sobol' and FAST methods are based on the (ANOVA-like) variance-based sensitivity measure (1), and useful to ascertain if a (small) subset of input variances may account for (most of) the output variance. FAST computes only the first-order terms (main effects) and, hence, is computationally appealing. Different from the FAST method, the Sobol' method evaluates each interaction term by computing a multi-dimensional integral via Monte Carlo techniques.

2. Sobol' Sensitivity Indices

The Sobol' indices (1990) are based on a decomposition of the (output) function *f* into functions of different dimensions. The representation

$$f(x) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{1\le i<j\le n} f_{ij}(x_i, x_j) + \cdots + f_{1,2,\cdots,n}(x_1, x_2, \cdots, x_n), \tag{2}$$

where $x = (x_1, x_2, \cdots, x_n)$ denotes the vector of *n*-dimensional input parameters with each x_i normalized and varies from *0* to *1*, is called a decomposition of summands of different dimensions. The Sobol' sensitivity indices are $S_{i_1,i_2,\cdots,i_s} = D_{i_1,i_2,\cdots,i_s} / D$, where D is the total variance of *f*, and $D_{i_1,i_2,\cdots,i_s} = \int_0^1\int_0^1 \cdots \int_0^1 f^2_{i_1,i_2,\cdots,i_s} dx_{i_1} dx_{i_2} \cdots dx_{i_s}$, for $1 \le i_1 < i_2 < \cdots < i_s \le n$.

We will present a comparative analysis of two variance-based sensitivity techniques, i.e., Sobol' and FAST methods, based on simulation studies.

References

Cukier, R.I., Fortine, C.M., Shuler, K.E., Petshek, A.G., and Schaibly, J.H. (1973) "Study of the sensitivity of coupled reaction systems to uncertainties in rate coefficients. I Theory." *J. Chem Physics*, **59,** 3873-3878.

Cukier, R.I., Levine, H.B., and Shuler, K.E. (1978) "Nonlinear sensitivity analysis of multi-Parameter model systems." *J. Comp. Physics*, **26**, 1-42.

Koda, M. (1998) "Sensitivity analysis in Data Mining." *Discussion Paper Series 788*, Institute of Policy and Planning Sciences, University of Tsukuba.

Koda, M., McRae, G.J., and Seinfeld, J.H. (1979) "Automatic sensitivity analysis of kinetic mechanism." *Int. J. Chemical Kinetics*, **11**, 427-444.

SAMO'95 (Sensitivity Analysis of Model Output) (1995) *Proc.* Int. Symp. on Sensitivity Analysis of Model Output, European Commission, Joint Research Center, Beligirate, Italy.

SAMO'98 (Sensitivity Analysis of Model Output) (1998) *Proc.* Second Int. Symp. on Sensitivity Analysis of Model Output, European Commission, Joint Research Center and University of Venice, Venice, Italy.

Sobol', I.M. (1990) "Sensitivity estimates for nonlinear mathematical models." *Matematicheskoe Modelironanie*, 2, 113-118.