

Sixth International Conference on Sensitivity Analysis of Model Output

Applications of sparse grid interpolation: sensitivity analysis and experiment design

Gregery T. Buzzard^{*a}, Daniela Changkuon^a, Maia M. Donahue^b, Ann E. Rundell^c

^aDepartment of Mathematics, Purdue University, West Lafayette, IN 47907

^bDow AgroSciences LLC, 9330 Zionsville Road, Indianapolis, IN 46268

^cWeldon School of Biomedical Engineering, Purdue University, West Lafayette, IN 47907

Abstract

We use adaptive sparse grid interpolation to produce accurate and efficient estimates of the sensitivity coefficients of Sobol' and to guide the design of experiments based on model output. We illustrate this method on two models of the mitogen-activated protein kinase cascade, one with 3 uncertain parameters and one with 18 uncertain parameters.

Keywords: sparse grid; sensitivity analysis; experiment design; MAPK; Sobol'

1. Main text

Sparse grid interpolation provides an efficient method for approximating a computationally expensive function with an interpolating polynomial. By using this polynomial as a surrogate for the original function, we are able to obtain computationally efficient approximate solutions for several problems of interest to modelers. First, we obtain accurate approximations to the variance-based sensitivity coefficients of Sobol'. Second, for a cost function describing the fit of a model to data, we use the polynomial surrogate to guide experiment design in order to maximally reduce the uncertainty in the model dynamics.

In sparse grid interpolation, a C^k smooth function $f(x)$, x in $[-1, 1]^d$ is evaluated at a predetermined set of points, x_1, \dots, x_N . The values of f at these points are used to construct a polynomial, p_N , which agrees with f at these points and which converges to f with error proportional to $N^{-k}(\log N)^{(k+2)(d-1)+1}$ (Barthelmann et al., 2000). Typically, p_N is constructed as a weighted sum of Lagrange interpolating polynomials. A common choice is to use products of 1-dimensional Lagrange interpolating polynomials on the extrema of the Chebyshev polynomials. The applicability of this approach has been extended by means of adaptive selection of additional evaluation points (Klimke, 2006).

We apply adaptive sparse grid interpolation to estimate the sensitivity coefficients of Sobol' (Sobol', 1993) accurately and efficiently. The main effect sensitivity coefficient for a function $f(x_1, \dots, x_d)$ is given by $\text{Var}_{x_j}(E[f|x_j]) / \text{Var}(f)$, where $E[f|x_j]$ is the expected value obtained by fixing a given value of x_j and integrating over the remaining

* Corresponding author. Tel.: +1-765-494-1937; fax: +1-765-494-0548.

E-mail address: buzzard@math.purdue.edu.

variables and Var_{x_j} denotes variance as a function of x_j only (Saltelli et al., 1999). As shown in (Buzzard & Xiu, 2009), the computation of these quantities reduces to the calculation of a few integrals. We evaluate these integrals efficiently and exactly for the interpolating polynomial in place of f , both using the original Lagrange representation and by changing basis to use an expansion in terms of Legendre polynomials. We also compute more general interaction effects with no additional evaluations of the original function.

For experiment design, we note that for many models, particularly in biology, the parameters in a model are not well constrained by existing data. Moreover, additional experiments may be expensive, so it is vital to design experiments that will be nearly optimal among available experiments in terms of constraining the parameters the most. Our sequential approach (Donahue et al., 2010) addresses these issues by using sparse grid interpolation to identify multiple areas of parameter space that are consistent with available data and then clustering these identified parameters based on simulated model response and the limits of experimental measurement. By analyzing the expected experimental variance and the variance due to different model responses, we choose a measurement to provide maximal discrimination among currently acceptable solutions. This experiment design criterion is similar to the Hunter-Reiner criterion (Hunter & Reiner, 1965) since it looks for the largest difference in predicted dynamics, but it also avoids design points with large expected measurement error as recommended by (Buzzi-Ferraris & Forzatti, 1990). This approach further differs from other experiment design methods in that it simultaneously addresses both parameter- and structural- based uncertainty, is applicable to some ill-posed problems where the number of uncertain parameters exceeds the amount of data, places very few requirements on the model type, available data, and *a priori* parameter estimates, and is performed over the global uncertain parameter space. We illustrate this approach on two models of the mitogen-activated protein kinase cascade, one with 3 uncertain parameters and one with 18 uncertain parameters. The results show that system dynamics are highly uncertain with an initial set of limited experimental data. Nevertheless, the algorithm requires only three additional experimental data points to simultaneously discriminate between possible model structures and acceptable parameter values. This sparse grid-based experiment design process provides a systematic and computationally efficient exploration over the entire uncertain parameter space of potential model structures to resolve the uncertainty in nonlinear systems biology model dynamics.

Portions of this work were supported by NSF/NIH grant DMS-0900277.

2. References

- Barthelmann V., Novak E. and Ritter K., 2000: "High dimensional polynomial interpolation on sparse grids," *Adv. Comp. Math.*, 12, 273-288.
- Buzzard G. and Xiu D., 2010: "Variance-based global sensitivity analysis via sparse-grid interpolation and cubature," to appear in *Comm. Comp. Phys.*
- Buzzi-Ferraris G. and Forzatti P., 1990: "An improved version of a sequential design criterion for discriminating among rival multiresponse models," *Chemical Engineering Science*, 45, 477-481.
- Donahue M., Buzzard G., and Rundell, A., 2010: "Experiment design through dynamical characterization of systems biology models utilizing sparse grids," to appear in *IET Systems Biology*.
- Hunter W. and Reiner A., 1965: "Designs for discriminating between two rival models," *Technometrics*, 7, 307-323.
- Klimke A., 2006: "Uncertainty Modeling using Fuzzy Arithmetic and Sparse Grids," PhD Thesis, Universität Stuttgart, Shaker Verlag, Aachen.
- Sobol', I.M., 1993: "Sensitivity estimates for nonlinear mathematical models," *Math. Modeling & Comp. Experiment*, 1, 407-414 [translation of Sobol', I.M., 1990: "Sensitivity estimates for nonlinear mathematical models," *Matematicheskoe Modelirovanie*, 2, 112-118 (in Russian)].
- Saltelli A., Tarantola S., and Chan K. P.-S., 1999: "A quantitative model-independent method for global sensitivity analysis of model output," *Technometrics*, 41, 39-54.