Surrogates 7020

Chapter 8B: Sensitivity Analysis

Dr. Alex Bledar Konomi

Department of Mathematical Sciences
University of Cincinnati

Sensitivity Analysis

- In any nonparametric regression setting, it's important to understand the role inputs play in predicted outputs.
- When inputs change, how do outputs change?
- In simple linear regression, estimated slope coefficients and their standard errors speak volumes, resulting in t-tests or F-tests to ascertain relevance.
- By contrast, the effect of fitted GP hyperparameters and input settings on predictive surfaces is subtle and sometimes counter-intuitive. In a way, that's what nonparametric means: parameters don't unilaterally dictate what's going on.
- Model fits and predictive equations gain flexibility from data, sometimes with the help of – but equally often in spite of – any estimated tuning or hyperparameters.

Uncertainty distribution

- If we're going to say how sensitive outputs are to changes in inputs, it makes sense to first say what inputs we expect/care about, and how much they themselves may change/vary.
- Underlying the Saltelli method is a reference distribution for $\boldsymbol{x}$, sometimes called an uncertainty distribution $U(\boldsymbol{x})$.
- U can represent uncertainty about future values of $\boldsymbol{x}$, or the relative amount of research interest in various areas of the input space.
- In many applications, the uncertainty distribution is simply uniform over a bounded region.

Uncertainty distribution

- In Bayesian optimization (Chapter 7), U can be used to express prior information from experimentalists or modelers on where to look for solutions.
- For example, when there's a large number of input variables over which an objective function is to be optimized, typically only a small subset will be influential within the confines of their uncertainty distribution.
- Sensitivity analysis can be used to reduce the volume of the search space of such optimizations (Taddy et al. 2009).
- In the case of observational systems such as air-quality or smog levels, $U(\boldsymbol{x})$ may derive from an estimate of the density governing natural occurrence of $\boldsymbol{x}$ factors, e.g., air pressure, temperature, wind and cloud cover.
- In such scenarios, sensitivity analysis attempts to resolve natural variability in responses $Y(\boldsymbol{x})$.

Assume for the moment independent

We treat here the standard and computationally convenient independent specification,

$$U(\boldsymbol{x}) = \prod_{k=1}^{m} u_k(\boldsymbol{x}_k),$$

where u_k, for $k = 1, \ldots, m$, represent densities assigned to the margins of $\boldsymbol{x}$. With U being specified probabilistically, readers may not be surprised to see sampling feature as a principal numerical device for averaging over uncertainties, i.e., over variability in U.

Main Effects

The simplest sensitivity indices are main effects, which deterministically vary one input variable, j, while averaging others over U_{-j}:

$$\begin{aligned} me(\boldsymbol{x}_j) &\equiv E_{U_{-j}}(y|\boldsymbol{x}_j) \\ &= \int \int_{X_{-j}} yp(y|x_1, \ldots, x_m) u_{-j}(x_1, x_{j-1}, x_{j+1}, x_m) dx_{-j} dy. \end{aligned}$$

Above, $u_{-j} = \prod_{k \neq j} u_k(x_k)$ represents density derived from the joint distribution U without coordinate j, i.e. U_{-j} with X_{-j} and x_{-j} defined similarly, and $p(y|x_1, \ldots, x_m) \equiv p(Y|\boldsymbol{x}) \equiv p(Y(\boldsymbol{x}) = y)$ comes from the surrogate, e.g., a GP predictor.

Calculating Main Effects

Algorithmically, calculating main effects proceeds as follows: grid out X-space in each coordinate with values x_{ji}, for $i = 1, \ldots, G$ say; gather $me(x_{ji})$-values holding the jth coordinate fixed at each x_{ji}, in turn, and average over the rest (and y) in an MC fashion; finally plot all $me(x_{ji})$ on a common x-axis. A pseudocode in Algorithm 8.3 formalizes that sequence of steps with some added numerical detail. The algorithm utilizes samples of size N in its MC approximations.

Main Effects or any other quantity

Although Eq. (8.5) and Algorithm 8.3 showcase mean main effects $\mu(\boldsymbol{x}) = \int y p(y|\boldsymbol{x}) dy$, any aspect of $p(y|\boldsymbol{x})$ which may be expressed as an integral can be averaged with respect to U in this manner. Quantiles are popular, for example, since they may be plotted on the same axes as means.

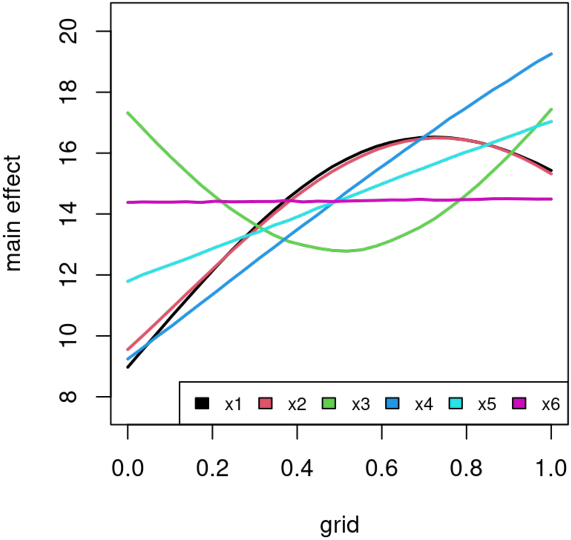

Figure: Mean main effects for Friedman data observed with irrelevant input x_6.