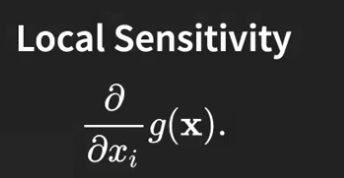
# **Noise** Robustness

Each data point is scaled by a factor, lambda, which is distributed normally (due to the central limit theorem) of mean 1 and std dev .05, creating a new data set. The model is then trained on the artificial data. This process is repeated, a thousand times, and a 95% confidence interval is plotted, between the 2.5 percentile and 97.5 percentile models. The graph of average percentage error against time, with noise application shows us XYZ. (Make comment on residuals).

Hi jakub i’m dying

FAX

Random:



How outputs vary with inputs

Simulate

The output varies locally according to the gradient according to one of the inputs

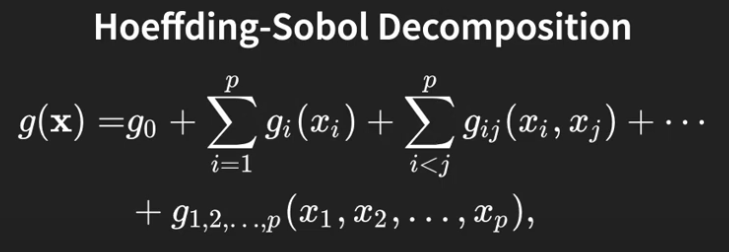
Good for local sensitivity analysis (locally linearise)

# **Parameter Analysis**

## Assumptions

Features are independent of one another. - SEE CROSS TALK IN MULTIVAR REG

## Sobol Indices

The simulated function can be decomposed into a set of unknown functions which is of the form:

Sobol decomposition ^

There are p terms where each term in the first sum is a univariate function for one of the input dimensions. These are called **first order terms**. The second order terms comprise the sum of p bivariate functions according to component parts of the function which is a function of two of the input variables. This decomposition considers all possible combinations of input variables.

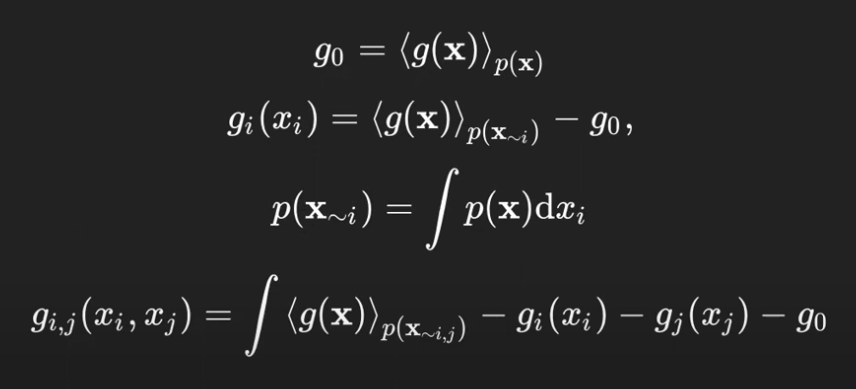
This implies interaction terms, where the terms have been decomposed into 3 types

Independent of any features

Dependent on one feature

Dependent on a combination of n features

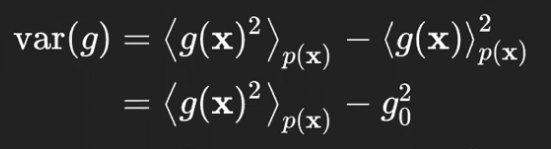
2^p terms, so this is not always expandable.



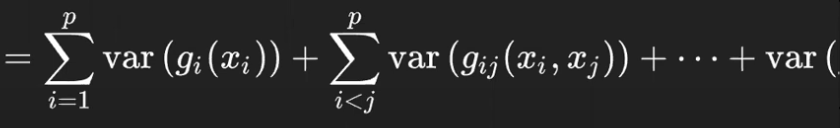
The first term (g\_0) is the expected value of the function in the domain over all the values of the function we expect, denoted p(x).

The other terms can be recursively computed by integrating over all the inputs of g(x) under the distribution p(xNOTi). That is the marginal distribution of p(x) with the relevant component omitted. Since they are independent, this becomes a trivial integral.

Create the other gs and time for the VARIANCE (important bit)



The second line is not a trivial statement - there’s some mathematical trick which is unimportant.



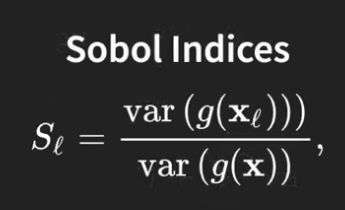
Involves summing variances of all smaller functions (**Total order index**). This is the key idea behind global sensitivity analysis because the sum of these variances (variances are additive). Good thing is that the total variance decomposes into these additive terms.

Decompose it in terms of the source of its variance - what input is responsible for the different components of this output. This kind of stuff is not to be done on a linear underlying function (this is ALL non linear analysis). For linear functions just use the partial derivative (local sensitivity). That’s why linear models are easily interpretable

This guy is spewing a whole lot of yap get to the good bit

We’ve used the sobol decomp. Substituted into formula for variance. Got the decomposition for the variance terms via construction. No one gives a shwat about the proof.

If we care about the variance of this function with respect to one particular variable we are only interested in the first order sobol term.



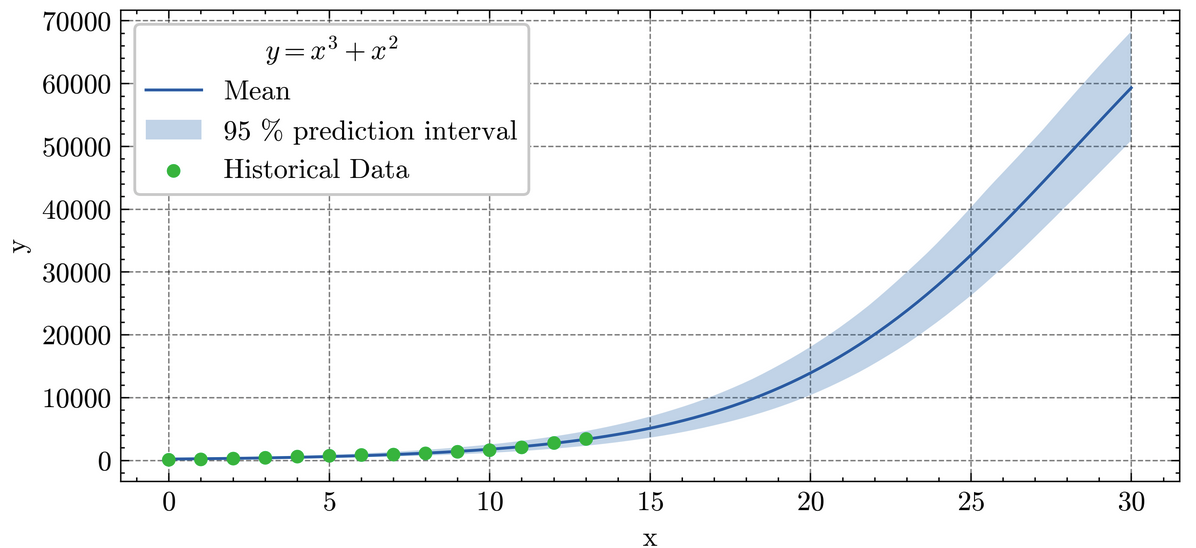
Finally at the cream

The parameters of the model are varied (e.g. L, k and x0 in logistic) by 5%,

Using parameters from model:

Apply range of noise 5% below and above stated values

Using these randomly adjusted models, predictions will vary.



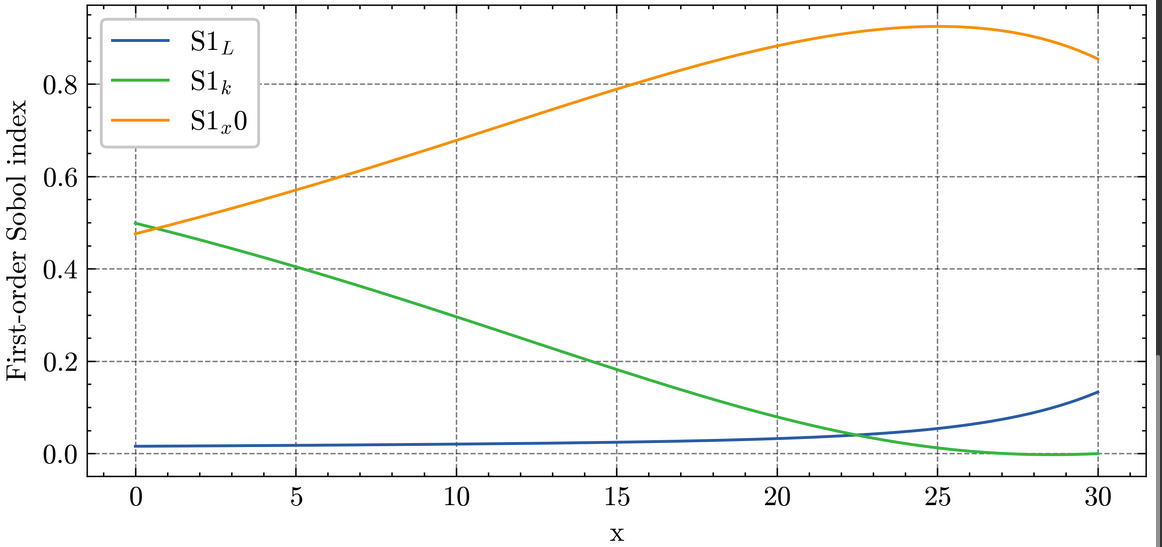
At each discrete time step

| Feature 1 | Feature 2 | Feature 3 | Feature 4 | prediction |
| --- | --- | --- | --- | --- |
| L | k | x | z | x |
| 1.05L | 0,95j | ccc | cc | y |
| 0.97L | .. | .. | .. | z |

Calculate the sobol index of each feature

For each year we do this for each feature

Feature importances over time



This is how sobol works

Order features of data and predicted values in data frame

Variance in predicted values

Do this for a range of time inputs

Describe graph with comment on context (talk about future predictions about relative importances)

References:

<https://en.wikipedia.org/wiki/Interaction_(statistics)>

# **Meyer Notes**

* Noise Robustness
  + Applies 5% noise to data
  + Retrains the model on the noisy data
  + Repeats 1000 times
  + Plots between the 2.5% and 97.5% quartile of results for the 95% confidence interval
  + The average percentage difference from the original prediction after applying noise is also plotted.
* Parameter Analysis
  + Uses SAlib
  + Varies the various parameters of the model (e.g. L, k and x0 in logistic) by 5%
  + This is used to make a prediction into the future
  + The 95% confidence interval is then plotted
  + The second plot is for the “first order indices” (look up SAlib) of the various parameters over time, basically their relative importance over time