MECH 513 – Laboratory 7

"Surrogate Modeling - Part 2"

Thomas H. Bradley, PhD.
(970) 491-3539, Thomas.Bradley@Colostate.edu

Introduction:

This last week, we have been developing the mathematics to be able to propagate uncertainty through models of computationally costly subroutines or sub-simulations.

MATLAB has a series of tools that will allow us to perform multidimensional response surface fitting, multidimensional artificial neural network fitting and visualization.

Learning Objectives:

1) Use response surfaces to create surrogate models to allow for the use of computationally costly models during system sensitivity analysis

Task 1 – Response Surface Modeling (RSM) in MATLAB

I have constructed a complicated and computationally intense problem for you all. Here is its general shape for your reference. This problem contains only two dimensions, making it amenable to visualization; a graph of the response is shown below. You can run the code using the following MATLAB script:

```
x = [1,1]
y = giant_CFD_code(x)
```

The behavior is highly multi-modal and non-linear, making it suitable for testing modeling techniques designed to handle highly non-linear functions. Consider the CFD simulation to be defined over the space:

$: x_1 \in [-2,2] \; x_2 \in [-1,1]$

You will have to develop a design of experiments to populate the design space for your model. The CFD simulation takes about 10 seconds to run, so be prepared to run your simulation for an hour or more to get ~hundreds of data points.

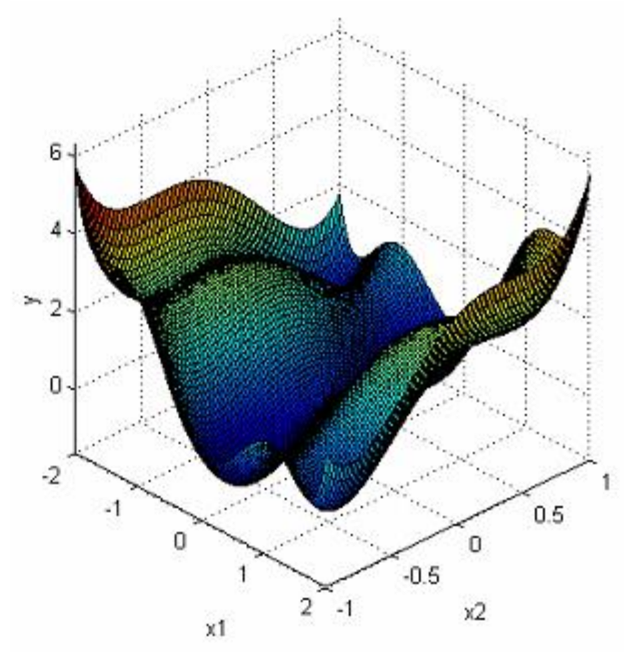

Figure 1 . Visualization of the complicated CFD simulation output (y) as a function of its two inputs (x1, x2)

You will have to generate enough data that you can have some approximately

70% of your data available for training (fitting)
15% of your data available for validation (not necessary for the RSM)
15% of your data available for testing

We will fit a response surface to this data to see if it is less expensive than actual routine. For a full quadratic response surface, we are going to fit the data using the following equation.

$$y = \beta_1 + \beta_2 x_1 + \beta_3 x_2 + \beta_4 x_1 x_2 + \\ \beta_5 x_1^2 + \beta_6 x_2^2$$

To do this in MATLAB simply call up the response surface equation fitting routine:

```
>> rstool
```

Check out the prediction profiler, and then export the data directly to MATLAB. The outputs of this fitting routine are the values of beta as defined in the canonical response surface equation shown above (except that Beta in MATLAB starts at index =1).

Task 2 – Neural Network Surrogate Modeling in MATLAB

Now we can do this same type of surrogate modeling using neural networks as the means for regressing the experimental data from our CFD Simulation. MATLAB has a great neural network toolbox that we can use to do this fitting. To launch the toolbox, call `nnstart`. We will use the `fitting tool` today to do this regression.

Follow the prompts through the GUI and train the artificial neural network. Check the regression plots to see what types of R^2 values your surrogate model has achieved. If you can save your network as a network named net, then you can evaluate your neural network using the function:

```
output_neuralnet(i) = sim(net,inputs)
```

You will have to choose the characteristics of the artificial neural network that you deem appropriate.

Task 3 – Homework

1) Provide a table to compare the characteristics of your two surrogate models on the bases of
    a. Computational time required to evaluate each surrogate model 100 times
    b. Actual by predicted plot for each data set (training, validation and testing), use the Matlab function `plotregression`
    c. Average Absolute Value of Relative Error associated with your surrogate models.

    Which surrogate model to you prefer?

2) Perform a numerical optimization on your ANN ***surrogate-CFD*** model to derive the minimum value of y, as a function of x_1, and x_2. You may initialize your model at any of the corner points of the design space.

    You can run the following MATLAB code to get your optimizer to run on the actual CFD model. It may take multiple hours to get your optimization to run using the CFD code. Compare the results of this optimization to the results of the same optimization on your surrogate model (the ANN) of the CFD code.

    ```
    options = optimset('Display','iter')
    [X,FVAL] = fminsearch('giant_CFD_code',[1,1],options)
    ```

    Report the minimum value of y, and the values of x_1 and x_2 corresponding to the minimum value.

3) Perform a 2^2-DOE-based global sensitivity analysis on your ANN ***surrogate-CFD*** model at the minimum value of y, so as to calculate the sensitivity of the output y to the inputs x1, and x2. **Report the sensitivities $S_{y,x1}$, and $S_{y,x2}$**

Fully document your homework solutions using this handout as a template. All problem statements must be copied to the solutions. All diagrams and plots must be labeled with units and symbols, and must be captioned. Please complete this assignment and turn a pdf into CANVAS. This assignment is due on 4/1/20 4:59pm in CANVAS.