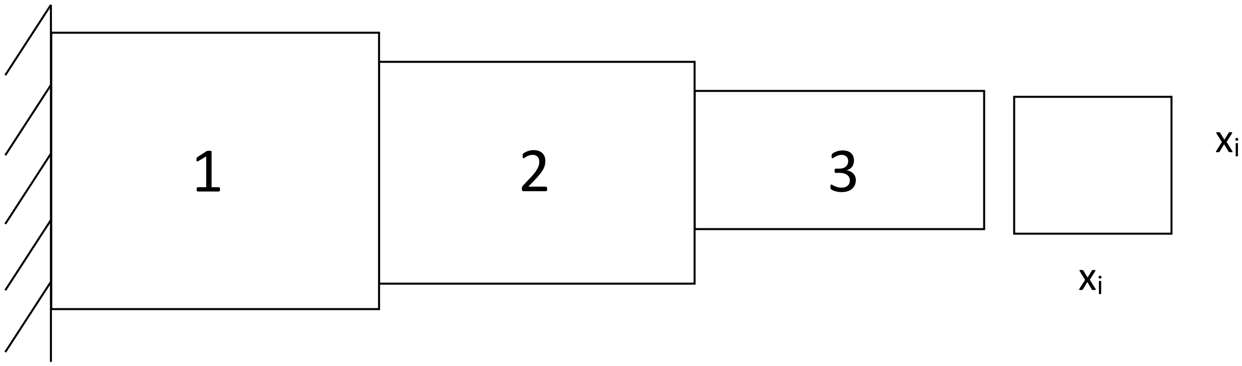
Problem Statement

The stepped cantilever beam described by three sections of equal length has square cross sections



The tip-displacement constraint is given by:

$$g(X) = -\left(\frac{61}{x_1^3} + \frac{37}{x_2^3} + \frac{19}{x_3^3} - 1\right) \geq 0$$

Nine different function approximations are considered for this work, including the following:

- Linear Approximation
- Reciprocal Approximation
- Conservative Approximation
- Quadratic Approximation
- TANA Approximation
- TANA1 Approximation
- TANA2 Approximation
- Response Surface Approximations
  - 2^3 factorial design
  - 4^3 factorial design

For this investigation, the task is to solve the constraint approximation problem using methods developed specifically for designing experiments, and to estimate the number of points necessary to ensure a correlation coefficient of ≥ 0.9.

The first order of business is to define the design space. Two points within the feasible region are chosen as reasonable boundaries for 3 variable factorial designs used for RSM analysis, as they span the feasible design space well.

$$x_1 = [6.5\ 6.5\ 6.5] \qquad x_2 = [9\ 9\ 9]$$

Also, we know…

$$g_L(X) = g(X_0) + \sum_{i=1}^{3} (x_i - x_{oi}) \left.\frac{\partial g}{\partial x_i}\right|_{X_0}$$

$$g_R(X) = g(X_0) + \sum_{i=1}^{3} (x_i - x_{oi}) \frac{x_{0i}}{x_i} \left.\frac{\partial g}{\partial x_i}\right|_{X_0}$$

$$g_C(X) = g(X_0) + \sum_{i=1}^{3} G_i\,(x_i - x_{oi})\,\frac{x_{0i}}{x_i}\,\frac{\partial g}{\partial x_i}\bigg|_{X_0},\quad G_i = \begin{cases} 1 & \text{if } x_{0i}\left(\frac{\partial g}{\partial x_i}\right) \le 0 \\ \frac{x_{0i}}{x_i} & \text{otherwise} \end{cases}$$

$$g_Q(X) = g(X_0) + \sum_{i=1}^{3}(x_i - x_{oi})\frac{\partial g}{\partial x_i}\bigg|_{X_0} + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i - x_{oi})(x_j - x_{oj})\frac{\partial^2 g}{\partial x_i \partial x_j}\bigg|_{X_0}$$

$$g_T(X) = g(X_2) + \frac{1}{r}\sum_{i=1}^{3} x_{i,2}^{1-r}\,(x_i^r - x_{i,2}^r)\left(\frac{\partial g}{\partial x_i}\right)_{X_2}$$

$$g_{T1}(X) = g(X_1) + \sum_{i=1}^{3}\frac{x_{i,2}^{1-p_i}}{p_i}(x_i^{p_i} - x_{i,2}^{p_i})\left(\frac{\partial g}{\partial x_i}\right)_{X_1} + \varepsilon_1$$

$$g_{T2}(X) = g(X_2) + \sum_{i=1}^{3}\frac{x_{i,2}^{1-p_i}}{p_i}(x_i^{p_i} - x_{i,2}^{p_i})\left(\frac{\partial g}{\partial x_i}\right)_{X_2} + \frac{1}{2}\varepsilon_2\sum_{i=1}^{3}(x_i^{p_i} - x_{i,2}^{p_i})^2$$

Response Surface Methods

The 2^3 and 4^3 factorial designs are considered to choose feasible points with which to develop a response surface model. 2^3 factorial designs are developed using only a low and a high level (-1 and 1) for each of the 3 design variables, resulting in 8 total “black box” points that must be collected. In this case the low level is $x_1$and the high level is x_2. On the other hand, 4^3 factorial designs use 4 factor levels for each design variables (-1, -1/3, 1/3, 1) and require 64 simulations which can in general be considered to have enormous computational cost. For this work, the designs take the following forms using the given data.

2^3 Factorial Design		
x1	x2	x3
6.5	6.5	6.5
6.5	6.5	9
6.5	9	6.5
6.5	9	9
9	6.5	6.5
9	6.5	9
9	9	6.5
9	9	9

4^3 Factorial Design		
x1	x2	x3
6.50	7.17	8.33
6.50	9.00	6.50
7.17	9.00	8.33
8.33	7.17	9.00
9.00	8.33	8.33
6.50	8.33	7.17
7.17	7.17	6.50
7.17	8.33	8.33
9.00	9.00	9.00
...	...	...
6.00	9.00	8.33

Multivariate Linear Least Squares Regression analysis is then used to develop a surrogate model (function approximation) based on the design points. In particular the following problem is solved to estimate least squares β coefficients.

$$Y = X\beta + \varepsilon$$

The orthogonality of $X\beta$and ε is used minimize the least squares error, and the solution for β coefficients is given by:

$$\hat{\beta} + (X^T X)^{-1} X^T Y$$

After the multivariate regression model (i.e. response surface) for both factorial designs is developed, the robustness of each model is tested. Monte Carlo simulations are used to choose pseudo-random sample points from within the design space. The random points within the design space are used to compute values of the actual function g(x) as well as function approximations values based on the two RSM models. The comparison of actual and approximated function values is used to obtain correlation coefficients. For this problem we have

$$R^2 = 1 - \frac{SS_e}{SS_t} = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y_i})^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^{n} (g_i - g_{RSM_i})^2}{\sum_{i=1}^{n} (g_i - \overline{g_{RSM}})^2}$$

An additional analysis for this work was done to consider the accuracy (within the design space) of all function approximations learned thus far, including the Linear, Reciprocal, Conservative, Quadratic, TANA, TANA1, TANA2, and RSM approximations. The function and approximation values are computed about X1 (previous) and X2 (current) points, as well as at surrounding points given by the following expression for all t between -1 and 1.

$$X_i = \left(t - \frac{1}{2}\right) X_2 + (\frac{1}{2} - t) X_1$$

Twenty data points in the design space are then used to compute percent error in approximation of the actual function for each function approximation formulation.

Results

Response Surface Approximations

Linear, Reciprocal, Conservative, Quadratic, TANA, TANA1, and TANA2 approximations are not shown here for brevity. However the Linear RSM models are given by:

- gRSM ***(2^3 Design)*** = 2*(0.02769*x1 + 0.01679*x2 + 0.008624*x3)- 0.1164
- gRSM ***(4^3 Design)*** = 2*(0.02742*x1 + 0.01663*x2 + 0.008541*x3) - 0.09397

Correlation Coefficient Computations

Thirteen MC simulations were run for each model, with the number of design points ranging from 5 to 100, and correlation coefficients were computed between actual and approximation values. Results are summarized below:

Rsquared		
n # points	2 Level	4 Level
5	0.6328	0.925
10	0.8427	0.8897
15	0.8329	0.9188
20	0.8413	0.8938
25	0.7191	0.9195
30	0.7757	0.9281
40	0.8247	0.8987
50	0.7731	0.9217
60	0.8157	0.9117
70	0.753	0.9217
80	0.8098	0.9292
90	0.7975	0.922
100	0.762	0.9282

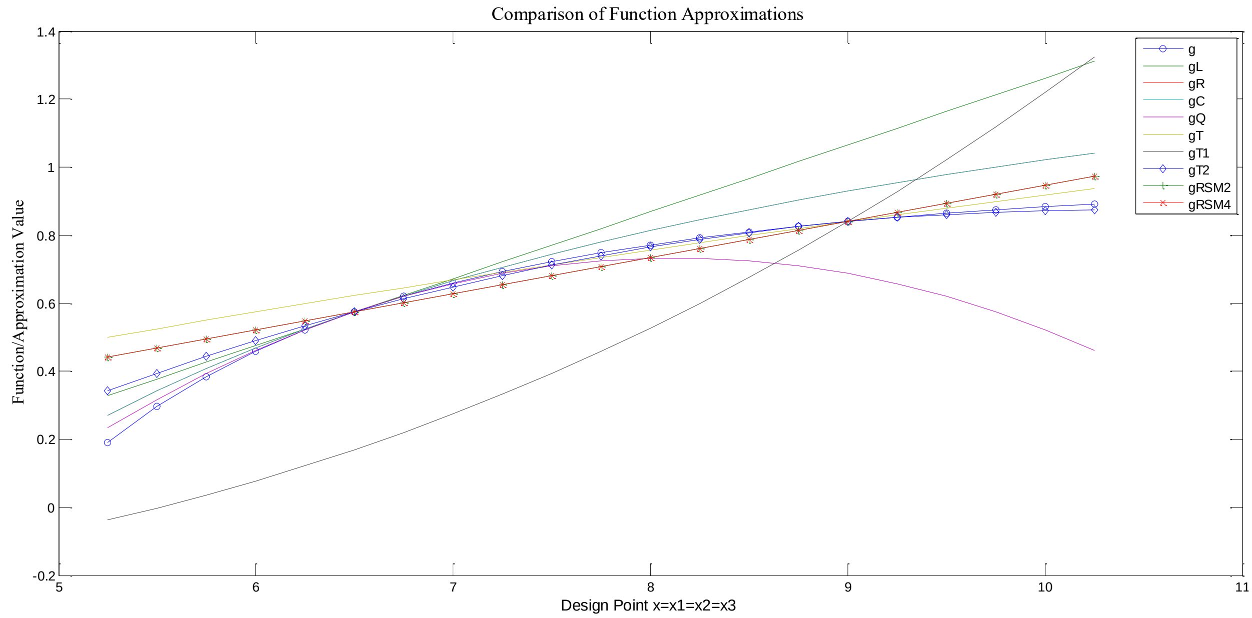
There are a few things that are immediate about these results:

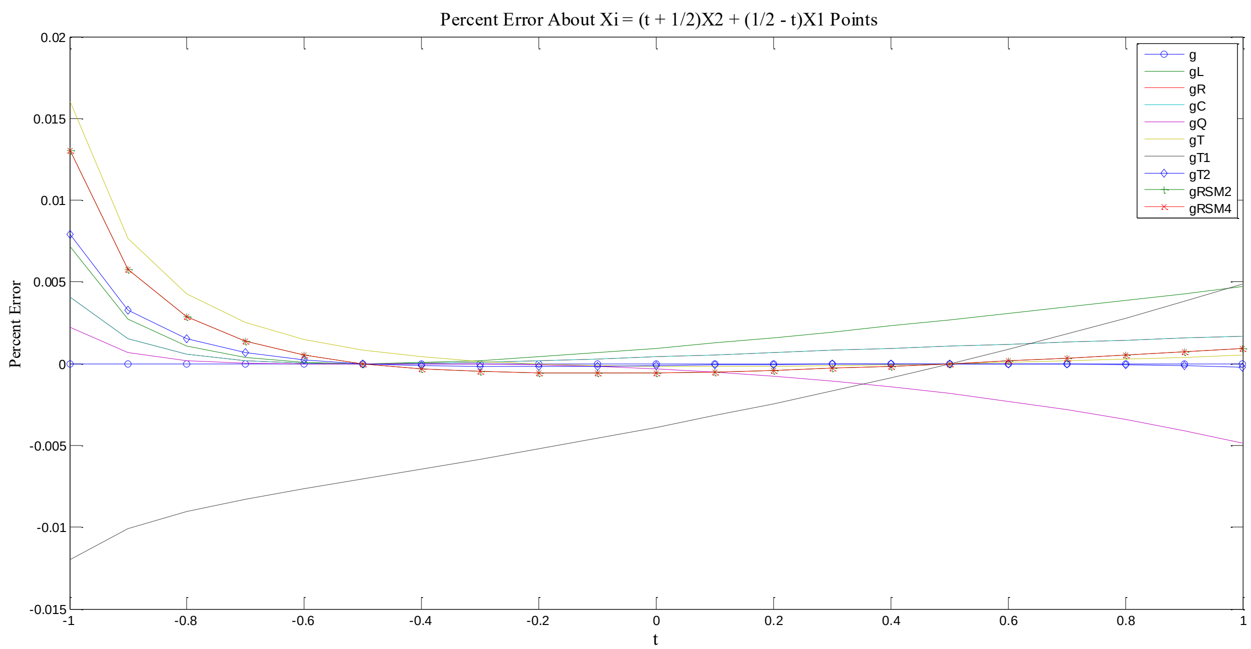
- The 4-level factorial formulation does a much better job ensuring that approximate and actual values are very close to one another throughout the entire design space, and is consistently ≥ 0.9.
- Regardless of how the design space is sampled, the 2-level factorial formulation is never able to ensure the approximation has a correlation coefficient ≥0.9.
- The 4-level is consistent over each design space sampling (i.e. over each n), while the 2-level factorial does poorly with sparse design space sampling (i.e. at n=5).

These results support the idea of using more than 8 design points (2 factor levels per design variable) to create a linear RSM model to represent this problem. While all 64 design points may not have been necessary to get significantly increased accuracy, fractional factorial or other box-type designs could certainly be used. It should be noted also that the cubic nature of the actual function through the design space is captured well by the 4-level factorial experiment, as expected.

Function Approximation Accuracy Analysis

All function approximations are plotted against the true function value in the first plot shown below. In the second plot, the percent error between each function approximation and the true function is computed.





The RSM function approximations are indistinguishable in both figures, despite the 2-level factorial design resulting in much lower R^2 values. This indicates that the 2-level does provide an accurate representation of the actual function within the design space considered.

Further, the TANA2 approximation is by far the best of any approximation considered. Additionally, TANA and TANA2 can be seen to match function values at both X1 and X2 points, while TANA1 can be seen to match the function values at X2 (i.e. the current point) and the gradient of g(x) at X1, as expected. The magnitude of the error in the TANA 1 approximation can be explained by the large size of the design space, the highly nonlinear problem, and the relatively small function approximation and actual function values.