Design on Non-Convex Regions: Optimal Experiments for Spatial Process Prediction with Applications to Industrial Processes

M. T. Pratola and D. Bingham
Simon Fraser University, Burnaby, BC, V5A 1S6, Canada

Abstract

Modeling a response over a non-convex design region is common in many industrial problems. Our research develops design and analysis methodology for experiments on non-convex regions. The approach uses a Gaussian Process (GP) model with a geodesic distance metric as a regression function. A novel use of Multidimensional Scaling (MDS) enables us to perform design and analysis on such regions.

Introduction

Scientists often make use of regression models in applied research to:

- Model the relationship between experimental factors and response variables
- Maximize response and obtain corresponding factor levels
- Determine robust factor level settings

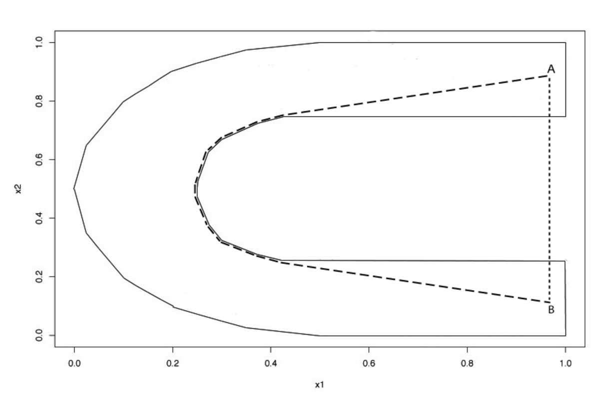

Figure 1: Non-convex design region with dotted line segment (a,b) not contained in region, dashed line segment (a,b) entirely contained in region

Sometimes the experimental variables may not be independent of one another. This can occur because

- it is not physically possible to run an experiment over a rectangular region
- a constraint may be present in the problem
- the response may not be defined over a rectangular region

We call these regions *non-convex* design regions, such as the hypothetical horseshoe region of Figure 1.

Our approach consists of:

- a geodesic distance metric
- transformation of the non-convex region into a new distance-preserving Euclidean space
- GP model using the transformed region
- Optimal design construction for prediction

Examples

Spot Welding: The size of a weld depends on current and time, but too much current or time causes expulsion while too little leaves a weak weld. The acceptable design region is shown in Figure 2.

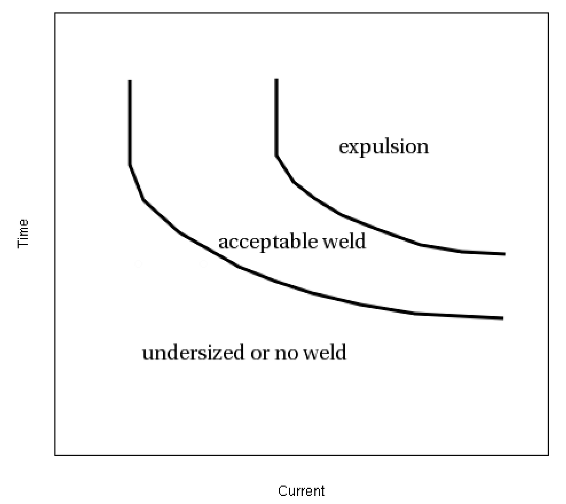

Figure 2: Design region for Spot Welding process

Structural Optimization: The weight of a truss structure is to be minimized subject to a constraint on it's maximum deflection under a dynamic loading. The highly non-convex region is shown in Figure 3 with a heatmap of the response weight.

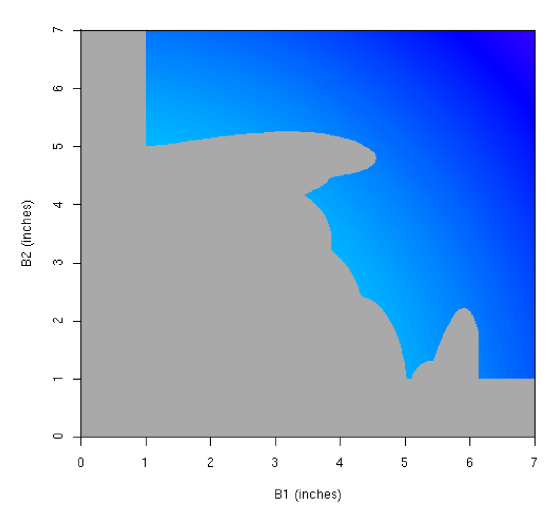

Figure 3: Weight (lb) response over non-convex design region from a constrained structural dynamics optimization problem

Analysis

We develop a GP model for non-convex design regions.

GP Model:

$$Y(x) = \mu + Z(x) + \varepsilon(x),$$
$$Z(x) \sim N(0, \sigma_z^2 R),$$
$$\varepsilon(x) \sim N(0, \sigma_\epsilon^2 I),$$
$$R_{ij} = \prod_{k=1}^{D} \rho_k^{4d_{ijk}^2},$$
$$d_{ijk}^2 = (x_{ik} - x_{jk})^2,$$

and ρ_k is a parameter for each dimension.

To model over a non-convex region, we instead use the *geodesic distance* $d_{g_{ij}}$.

Geodesic Distance: The length of the shortest path between two points such that the path is entirely contained within the region. See dashed line in Figure 1.

ISOMAP Approximation: A Euclidean space is constructed to approximate geodesic distance with ISOMAP [1].

This estimates d_g using a Euclidean embedding space found via Multidimensional Scaling (MDS). The approximate metric is then d_e.

Embedding Space: The embedding space dimension chosen minimizes $(d_g - d_e)^2$ instead of the usual proportion of variability explained metric.

The resulting embedding space is often of higher dimension than the space of the original problem.

Design

We can construct designs and model

- in the original space, using the distance approximation, or,
- in the higher dimensional space found by MDS.

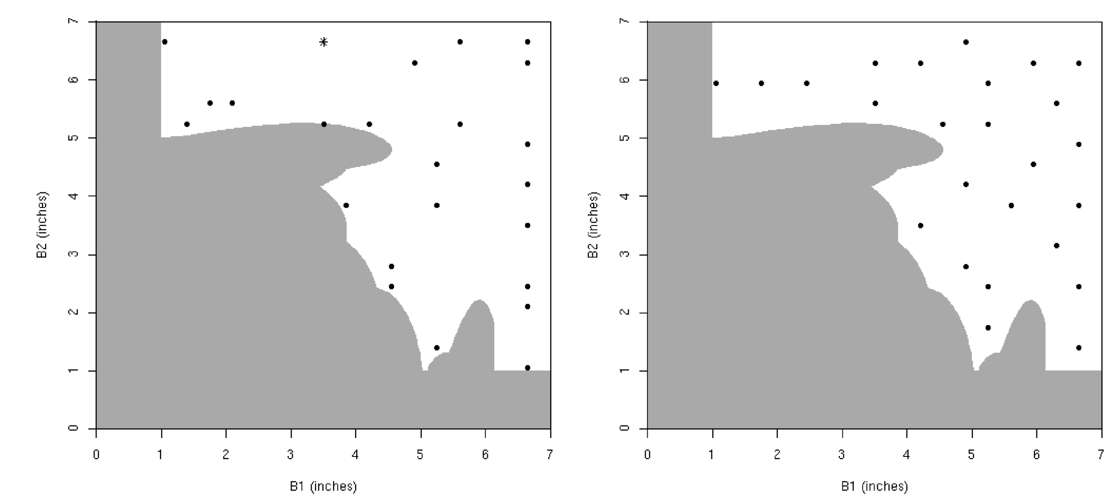

Figure 4: 25-point IMSE optimal design for the constrained optimization problem (left) and 25-point space-filling design for the constrained optimization problem (right)

Design Criterion: A method of constructing *Integrated Mean Squared Error (IMSE) Optimal* designs for non-convex regions was derived. IMSE optimal designs are appropriate when interest is in prediction.

A design $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is found by solving

$$\underset{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N}{\arg\min} \; IMSE\left(\hat{Y}(\mathbf{x}) \mid \rho, \sigma_z^2, \sigma_\epsilon^2\right).$$

Our method constructs such designs directly in the non-convex region or in the approximating embedding space.

An IMSE-optimal design for the constrained structural optimization problem is shown in Figure 4.

Results and Conclusions

Simulation Study: A simulation study was performed over the region in Figure 1 for various assumed responses:

- f_g: GP response in original region (distance approximation)
- f_{e^2}: GP response in 2D embedding space
- f_{e^3}: GP response in 3D embedding space
- $f_{e^3,LM}$: 2^{nd}-order linear model in 3D embedding space

IMSE-optimal designs were constructed assuming a correlation parameter $\rho = 0.2$. We compare our optimal designs to geodesically space-filling designs, and report the relative efficiency:

Function (Truth)	Relative Efficiency (Space Filling to IMSE-Optimal)
$f_{g,\rho=0.2}$	0.88
$f_{g,\rho=0.4}$	0.89
$f_{g,\rho=0.6}$	0.91
$f_{e^2,\rho=0.2}$	0.86
$f_{e^2,\rho=0.4}$	0.96
$f_{e^2,\rho=0.6}$	0.96
$f_{e^3,\rho=0.2}$	0.86
$f_{e^3,\rho=0.4}$	0.92
$f_{e^3,\rho=0.6}$	0.96
$f_{e^3,LM}$	0.75

Relative efficiencies span from 0.75 to 0.96, reflecting an improvement using the IMSE-Optimal design of 4% to 25%.

Structural Optimization Result: An IMSE-optimal and space-filling design for the optimization example is shown in Figure 4. An average improvement of 10% was realized in this real-world example.

Recommendations: A practitioner can realize tangible benefits when modeling over a non-convex region. Using the embedding space as a distance approximation is easiest and resulted in similar prediction ability in most cases.

Conclusion: Our method achieves noticeable improvements, particularly as the response becomes more complex. A complete discussion is available [2]. Contact mpratola@sfu.ca.

References

1. Tenenbaum, J.B. and de Silva, V. and Langford, J.C.: *A global geometric framework for nonlinear dimensionality reduction, Science, pp.2319-2323, vol.290, 2000*
2. *Pratola, M.T.:* Design on Non-Convex Regions: Optimal Experiments for Spatial Process Prediction, M.Sc. Thesis, Dept. of Statistics and Actuarial Science, Simon Fraser University, 2006