Computing Nonlinear Active Subspaces for Highly Parameterized Optimization Problems using Autoencoders

Gabriele Boncoraglio, gbonco@stanford.edu

Overview

A novel dimensionality reduction approach related to solving highly parameterized optimization problems is introduced specifically for problems involving running expensive computer simulations such as aircraft design.

Issue: A method used for dimensionality reduction is Active Subspace (AS); it consists in a linear approximation of the subspace and it is not always feasible for nonlinear problems.

Objective: Building an autoencoder using Neural Networks for computing a nonlinear approximation of the subspace and solving the optimization problem in the subspace.

Results: the proposed approach is applied to a series of optimization problems and the results show that the autoencoder can efficiently represent a nonlinear manifold reducing the dimensionality of the problem and solve the optimization problem.

Background & Methodology

Assuming the optimization problem has the following form:

$$\max_{x \in \mathbb{R}^n} \ f(x) \quad \text{subject to} \quad z(x) \leq 0 \quad x^a \leq x \leq x^b$$

The method of active subspace compresses with SVD the gradients of the objective function $f(x)$ to learn a linear low-dimensional representation of the input parameters:

$$\mathbb{R}^n \approx \mathbb{R}^d$$

Using linear Active subspace, the following problem is solved:

$$\max_{x \in \mathbb{R}^d} f(U_k x_k) \quad \text{subject to} \quad U_k x_k \leq U_k x_k \leq U_k x_k$$

The autoencoder learns a nonlinear low-dimensional representation of the input parameters:

$$x \approx g(x)$$

Using nonlinear Active subspace, the following problem is solved:

$$\max_{g(x) \in \mathbb{R}^d} \ f(g(x)) \quad \text{subject to} \quad g(x) \leq g(x) \leq g(x)$$

Data & Features

The dataset for training the autoencoder contains gradients $\nabla f$. The data is generated using computer simulations: For a randomly sampled vector of parameters $x$, the gradient $\nabla f(x)$ is computed.

The dataset is the matrix $M$, (rows=$N_x$ and cols=$m$)

$$M = [\nabla f(x_1), \ldots, \nabla f(x_m)]$$

The size of the dataset, $m$, is chosen using an iterative process

Model & Model Selection

Three models considered. A grid search for selecting the model with the lowest dev set loss is run after training with the mAE-Wing dataset

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Dev set Loss</td>
<td>36.6</td>
<td>21.8</td>
<td>36.7</td>
</tr>
</tbody>
</table>

The optimal architecture has 3 layers: the first layer has 33 filters, the second layer has 23 filters and the third layer has 3 filters (equal to $n_x$)

Experiments & Results

Unconstrained optimization, $N_x=33$, $n_x=2$

Conclusions and Future Work

• Using a nonlinear manifold for representing a subspace has advantages especially when a linear active subspace fails to capture nonlinearities in the problem
• The linear active subspace is a particular case of the nonlinear active subspace obtained with the autoencoder
• In the future, I would like to study how to make the training of the autoencoder cheaper and to study the performances of this methodology in other additional problems.