249A Final Project Theoretical Background for Scenario Reachability

Liz Dietrich

1. Tiling with Basis Functions

1.1. This is an excerpt from the paper ”Nonconvex Scenario Optimization for Data-Driven Reachability”

We first consider a method convex in θ but nonconvex in x and construct a sublevel set function $g(x, \theta)$. While convex scenario optimization methods can be used to analyze this approach, they require large sample sizes and are not computationally efficient. We show that by utilizing nonconvex scenario optimization tools, we can significantly improve upon these limitations. To construct $g(x, \theta)$, select a finite set of basis functions $f_1(x), ..., f_m(x) : \mathbb{R}^D \to \mathbb{R}$ not convex in x and take g to be $g(x, \theta) = \sum_{i=1}^{m} \theta_i f_i(x)$. In this section, we will use this approach to construct a *tiling* of the state space and estimate the reachable set of a given problem.

To create this *tiling*, assume that the reachable set, $\mathcal{R}$, is contained in a subset $A \subseteq \mathbb{R}^D$. We then partition A into m cells, creating a collection of sets $A_1, ..., A_m$ such that $\cup_{i=1}^{m} A_i = A$ and $A_i \cap A_j = \emptyset \; \forall i$. This approach can produce arbitrarily fine estimates of the reachable set, depending on how refined the partition we choose is. The accuracy of the partition increases as m increases. Further, we define $\mathbb{1}_{A_i}$ to be the zero-one indicator function for the set A_i, so that $\mathbb{1}_{A_i}(x) = 1$ if $x \in A_i$ and $\mathbb{1}_{A_i}(x) = 0$ otherwise. Therefore, the reachable set estimate is a union of the partitioned cells $\mathcal{R}(\theta) = \{A_i \in \mathcal{R}(\theta) \text{ iff } \theta_i = 0\}$. We write this as a constrained optimization problem:

$$\begin{aligned} \underset{\theta}{\text{minimize}} \quad & -\sum_{i=1}^{m} \theta_i \\ \text{subject to} \quad & \sum_{i=1}^{m} \theta_i \mathbb{1}_{A_i}(x) \leq 0, \quad i = 1, \ldots, N \\ & \theta \in [0, 1]^m, \end{aligned} \tag{1}$$

Which can also be written in the form of:

$$\begin{aligned} \underset{\theta}{\text{minimize}} \quad & -\sum_{i=1}^{m} \theta_i \\ \text{subject to} \quad & \theta \in \bigcap_{i=1,\ldots,m} \{\theta_i : g(\delta^{(i)}, \theta_i) \leq 0\} \end{aligned} \tag{2}$$

To satisfy the scenario constraints, we set $\theta_i = 0 \; \forall i$ such that scenario $\delta_j \in A_i$ for at least one $j \in \{1, ..., N\}$. To minimize the objective while respecting $\theta \in [0, 1]^m$ we set $\theta_i = 1$ for all other A_i. Therefore, $\mathcal{R}(\theta)$ is the solution to 1, the union of cells A_i that contain one or more scenarios δ_j. This is the minimum volume union of cells that contains all scenarios, $\delta_1, ..., \delta_N$.

We define a support scenario as a cell, A_i that contains exactly one scenario, δ_j. If this scenario, $\delta_j \in A_i$ were to be removed from the set of all scenarios, $\delta_1, ..., \delta_N$, the reachable set would no longer contain A_i, therefore changing the solution of 1. This satisfies the conditions needed to obtain the set of s_N^* support scenarios. We calculate ϵ a-posteriori using Equation **??** with s_N^* and the number of optimization variables, d (the number of cells in $\mathcal{R}(\theta)$). If the obtained ϵ does not satisfy the necessary bounds, we iteratively increase the number of samples, N, and repeat the process. This reachability algorithm based on partition $A_1, ..., A_m$ is described in Algorithm 1.

Algorithm 1 : Scenario reachability through tiling

Input: Black-box transition function model $\phi(t_1; t_0, x_0, r)$; Random variables X_0 and R; Partition dimension m; confidence parameter $\delta \in (0, 1)$.
Output: $\theta_1, ..., \theta_m$ corresponding to union of cells A_j such that $A_j \in \mathcal{R}(\theta)$ iff $\theta = 0$; $\Theta_1, ..., \Theta_m$ corresponding to support scenarios A_j such that $A_j \in \mathcal{R}(\theta)$ iff $\Theta = 0$.
Initialize $\theta_j = 1; \Theta_j = 1; j = 1, ..., m; N = 1000.$
forall $i \in \{1, ..., N\}$ **do**

  Take samples $x_{0i} \sim X_0, r_i \sim R$

  Evaluate $\delta_i = \phi(t_1; t_0, x_{0i}, r_i)$

  If $\delta_i \in A_i$, then set $\theta_i = 0$

  If $|A_i| = 1$, then set $\Theta_i = 0$. Else set $\Theta_i = 1$

end
return $\theta_1, \ldots, \theta_m; \Theta_1, \ldots, \Theta_m$.
calculate ϵ. If ϵ is too large, increase N and repeat algorithm.