AUGMENTED LAGRANGIAN METHOD FOR RELU DEEP NEURAL NETWORKS

1. ReLU Deep Neural Networks as Complementarity Constraints

Given input output data (x_0^j, y^j), $j = 1, \ldots, N$ we are interested in solving

$$\begin{aligned} &\textbf{minimize} \sum_{j=1}^{N} \|W_N x_N^j - y^i\|^2 \\ &\textbf{subject to } x_{k+1} = \textbf{max}\{W_k x_k^j, 0\}, \qquad k = 0, \ldots, K-1,\ j = 1, \ldots, N \end{aligned}$$

The dynamics can be expressed as

$$\begin{aligned} x_{k+1} = \textbf{max}\{W_k x_k, 0\} &\iff x_{k+1} = -\,\textbf{min}\{-W_k x_k, 0\} \\ &\iff \textbf{min}\{x_{k+1} - W_k x_k, x_{k+1}\} = 0 \\ &\iff x_{k+1}^\top (x_{k+1} - W_k x_k) = 0, x_{k+1} \geq 0, x_{k+1} \geq W_k x_k \end{aligned}$$

Therefore ReLU DNN training can be expressed as the MPEC

$$\begin{aligned} &\textbf{minimize} \sum_{j=1}^{N} \|W_N x_N^j - y^i\|^2 \\ &\textbf{subject to } (x_{k+1}^j - W_k x_k^j)^\top x_{k+1}^j \leq 0, \qquad k = 0, \ldots, K-1,\ j = 1, \ldots, N \\ &\qquad x_{k+1}^j \geq 0, x_{k+1}^j \geq W_k x_k^j, \qquad k = 0, \ldots, K-1,\ j = 1, \ldots, N \end{aligned}$$

Date: November 30, 2018.