
SINA MOHAMMAD-TEHERI, Concordia University

Greedy deep unfolding: New neural networks based on sparse recovery algorithms

In recent years, *algorithm unfolding* (or *unrolling*) has emerged as a promising methodology in various signal-processing applications, notably in biomedical imaging. It aims to combine the strengths of sparse recovery and deep learning by designing deep neural network architectures that implement the iterations of an iterative algorithm as network layers. Despite its success, unfolding *greedy* sparse recovery algorithms like Orthogonal Matching Pursuit (OMP) or Compressive Sampling Matching Pursuit (CoSaMP) has received little attention. The primary challenge is the non-differentiable nature of the *argsort* operator, a key component in greedy algorithms, which hinders gradient backpropagation during training.

To address this, our work introduces a novel approach, termed 'greedy deep unfolding'. We utilize *soft sorting* to approximate the *argsort* operator in a differentiable manner. Additionally, we reinterpret greedy algorithms through a projection-based lens and approximate the permutation matrices from *argsort* with stochastic matrices derived from soft sorting. Our numerical and theoretical analyses show that under certain conditions, the approximation error is minimal, and the performance of the approximated greedy algorithm closely matches the original. We then incorporate this approximate algorithm into a feedforward neural network's layers, integrating learnable *weight* parameters to connect to weighted sparse recovery. Our numerical results demonstrate that this network is trainable and can surpass the performance of the traditional approximate greedy algorithm in certain scenarios.