CS 448I Project Proposal

Samir Agarwala Jared Watrous

February 23, 2024

1 Introduction

In computational imaging, blurry and noisy images are an inherent result of imperfect hardware and measurements. Naturally, the topic of deconvolution arises with the objective of recovering a blur- and noise-free image from a flawed observation. One potentially powerful approach is the Alternating Direction Method of Multipliers (ADMM) [1], which, despite showing promising results, requires hyperparameter tuning and heuristics to optimally deconvolve an image. In this project, we propose an exploration of methods to learn these hyperparameters, and other ways of introducing machine learning to improve the performance of ADMM.

2 Related Work

Modelling optimization approaches using deep learning. Past work has explored using deep learning to model maximum a posteriori (MAP) optimization problems. By using deep learning, methods are able to learn patterns from data and leverage inductive biases encoded in the network architecture to potentially make more intelligence decisions during the optimization process. For instance, [2] proposed a way to leverage information about image formation models into neural networks, and unroll the networks to solve inverse problems. This is done by weight sharing in a neural network and repeatedly applying the same network in the different optimization steps. The paper presents promising results on the deblurring and denoising tasks, showing how using inductive biases or learning priors from data can help us find better solutions to inverse problems.

Adaptations to ADMM. There has been interesting work in ways to adapt ADMM to leverage learned priors. For instance, there has been work [3] focused on proving fixed-point convergence guarantees of ADMM where the z-update steps leverage denoisers that are learned neural networks i.e. there is guaranteed to be convergence to a solution if the initialization lies in a certain region under certain denoisers. Other work [4] has explored modelling the different steps of ADMM using separately learned modules for each of the update steps. However, there is still scope to explore ways in-which we can learn priors to predict hyperparameters for the existing ADMM. For instance, parameters such as ρ and λ in ADMM are often set by hand and by learning priors from data, we can potentially learn task-specific or perhaps even iteration-specific hyperparameters that lead to quicker convergence and better performance in solving inverse problems. Through our work, we hope to contribute towards the efficient selection of good hyperparameters for optimization algorithms such as ADMM.

3 Proposed Method

We choose to explore a straightforward gradient descent optimization of the global hyperparameters λ and ρ. In our proposed method, we fix the number of ADMM iterations N, and unroll the entire algorithm into a single feed-forward computation graph. We then optimize the hyperparameters

λ and ρ using stochastic gradient descent, where the loss is the mean squared error between the deconvolved output image and the ground truth image in pixel space. In this case, the hyperparameters will be optimized to fit the entire training dataset, with no conditioning on the input image. Given computational constraints, we plan on limiting our investigation to relatively smaller datasets such as MNIST [5] and CIFAR [6].

As an extension, if time and compute resources permit, we will introduce a neural network to predict the optimal hyperparameters λ and ρ, conditioned on the input image and possibly the denoising iteration. These approaches would allow the algorithm to adapt the hyperparameters to fit individual images, rather than enforcing globally constant values. Some other possible extensions we are considering time permitting include examining how learning hyperparameters on one dataset transfers to another. For instance, if we learn ρ and λ on CIFAR-10 [6], will it generalise to MNIST [5] and the other way around? Our third potential extension is learning an iteration-specific denoiser for the z-update step in ADMM where we can more intelligently guide the ADMM optimization process.

4 Evaluation

To evaluate our proposed method, we will evaluate both correctness and efficiency. Our baseline for comparison will be the vanilla ADMM algorithm without any learned priors. We will measure correctness by measuring PSNR between denoised images and the GT image we have from our chosen dataset. Additionally, we will compare runtime of the programs to achieve convergence where convergence would be defined as either having a solution close to the GT or reaching a maximum number of iterations which we will define on a holdout validation set.

5 Milestones

The main goals we hope to achieve in our project this quarter are:

- Optimize the ρ and λ parameters in ADMM using stochastic gradient descent on the denosing task using data from MNIST or CIFAR.
- Evaluate the quality of the learned parameters on an unseen test set and compare against the vanilla ADMM algorithm in terms of both accuracy and efficiency

The stretch goals of our project, given time- and compute-related constraints, include:

1. Compare generalisation of learned ADMM hyperparameters for a task across different datasets. For instance, do hyperparameters learned on MNIST transfer to CIFAR and vice-versa?
2. Learn a neural network to predict ADMM hyperparameters on a per-image image or per-iteration basis.
3. Learn an iteration-specific denoiser that can be used in the z-update step in ADMM.

Acknowledgements. We would like to thank Prof. Gordon Wetzstein for helpful discussions and feedback about our project.

References

[1] Stephen Boyd et al. “Distributed optimization and statistical learning via the alternating direction method of multipliers”. In: *Foundations and Trends® in Machine learning* 3.1 (2011), pp. 1–122.

[2] Steven Diamond et al. “Unrolled optimization with deep priors”. In: *arXiv preprint arXiv:1705.08041* (2017).

[3] Stanley H Chan, Xiran Wang, and Omar A Elgendy. “Plug-and-play ADMM for image restoration: Fixed-point convergence and applications”. In: *IEEE Transactions on Computational Imaging* 3.1 (2016), pp. 84–98.

[4] Chao Zhou and Miguel RD Rodrigues. “ADMM-based hyperspectral unmixing networks for abundance and endmember estimation”. In: *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021), pp. 1–18.

[5] Li Deng. “The mnist database of handwritten digit images for machine learning research”. In: *IEEE Signal Processing Magazine* 29.6 (2012), pp. 141–142.

[6] Alex Krizhevsky, Geoffrey Hinton, et al. “Learning multiple layers of features from tiny images”. In: (2009).