Learned primal dual by Jonas Adler

Zaccharie Ramzi

CEA (Neurospin & Cosmostat)

April, 18 2019

Outline

1. Context
   - MRI
   - General inverse problems formulation
2. Solving inverse problems
   - Model-based approaches
   - Data-driven approaches
   - Mixed approaches

MR image reconstruction

Retrieve the image from its Fourier coefficients (noisy, potentially undersampled, potentially non-uniform).

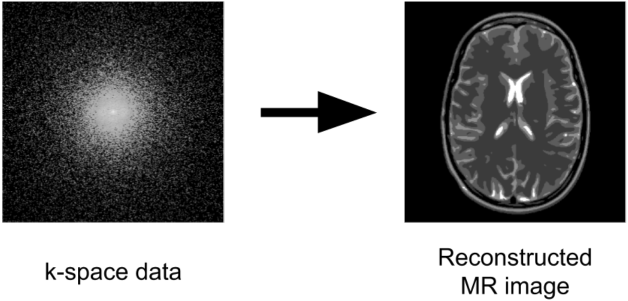

Figure: The goal of MR image reconstruction: going from some points in the Fourier space (k-space) to a reconstructed weighted contrast image

General ill-posed inverse problems formulation

$$\underset{x \in K^n}{\operatorname{argmin}} \frac{1}{2}\|Ax - y\|^2 + \lambda\|\psi x\|_1 \tag{1}$$

- K is generally $\mathbb{R}$ or $\mathbb{C}$
- A is the forward operator
- y is the observed data
- $\|\cdot\|$ is an appropriate norm
- ψ is an operator projecting x in a basis where it has a sparse representation
- the last norm can be either 0 or 1

Model-based approaches

To solve this problem we have 2 main classes of algorithms:

- **Primal-dual approaches**, e.g. Condat-Vu , Primal Dual Hybrid Gradient, ADMM . They can't be easily accelerated.
- **Proximal methods**, e.g. FISTA , POGM . When the regularisation term is not "simple", they need an inner loop to compute its proximity operator.

A similarity between the 2 classes is that they will both make use of the proximity operator of the regularisation term and the gradient of the data fidelity term.

Data-driven approaches

In this case we want to consider the problem as supervised learning problem. Keeping notation consistent with 1:

$$\underset{\theta}{\operatorname{argmin}} \, \mathbb{E}_{X,Y}[\mathcal{L}(f_\theta(Y), X)] \quad \text{“simplified” in} \quad \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{N} \mathcal{L}(f_\theta(Y_i), X_i)$$

- f_θ is our learned reconstruction function. For neural networks, θ are the network's weights.
- $\mathcal{L}$ is our loss function.
- (Y_i, X_i) is our N-size training set, composed of respectively k-space data points and ground truth MR images.

Mixed approaches

One of the main problems with data-driven approaches: going from local to global. In neural networks, this amounts to having dense layers (for a 512×512 image, 6.9×10^{10} parameters).

Jonas Adler et al. address this problem in their work by learning the proximity operators involved in the Primal Dual Hybrid Gradient.

Beck, Amir and Marc Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems". In: *SIAM Journal on Imaging Sciences* 2.1 (2009), pp. 183–202.

Boyd, Stephen et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers". In: *Found. and Trends in Mach. learn.* 3.1 (2011), pp. 1–122.

Chambolle, Antonin and Thomas Pock. "A first-order primal-dual algorithm for convex problems with applications to imaging". In: *Journal of mathematical imaging and vision* 40.1 (2011), pp. 120–145.

Condat, Laurent. "A Primal–Dual Splitting Method for Convex Optimization Involving Lipschitzian, Proximable and Linear Composite Terms". In: *Journal of Optimization Theory and Applications* 158.2 (Aug. 2013), pp. 460–479.

Kim, Donghwan and Jeffrey A Fessler. "Adaptive restart of the optimized gradient method for convex optimization". In: *Journal of Optimization Theory and Applications* 178.1 (2018), pp. 240–263.