Assignment 4

Convex Optimization SS2022

Martin Zach (martin.zach@icg.tugraz.at)

June 13, 2022

Submission: Upload your report as a single pdf–file ($\leq$ 10MB) to the TeachCenter and include your implementation (main.py) in your submission.

Deadline: June, 27^{th} 2022, 18:00

Computed Tomography Reconstruction

Let $\mathbf{b} \in \mathbb{R}^l$ be the sinogram data acquired by a computed tomography (CT) scan. For a $n = M \times M$ original image $\mathbf{x} \in \mathbb{R}^n$, the sinogram data $\mathbf{b}$ is an image of size $l = M \times L$, where L denotes the number of acquired angles. Each entry of the sinogram data represents a line integral across the original image for a certain offset and angle θ_i, see Figure 1. To account for measurement uncertainty, we assume that the observed sinogram data follows the model

$$\mathbf{b} = \mathcal{A}(\mathbf{x}) + \xi,$$

where $\xi \in \mathbb{R}^l$ is additive Gaussian noise and the linear operator $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^l$ is the Radon transform, which is given by the pylops.signalprocessing.Radon2D function throughout this assignment.

TV-Regularized Reconstruction (25P)

To limit the effect of the noise ξ on the reconstruction, we incorporate the total variation (TV) regularization. The resulting variational reconstruction problem reads as

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ E(\mathbf{x}) \coloneqq \lambda \|D\mathbf{x}\|_{2,1} + \tfrac{1}{2}\|\mathcal{A}(\mathbf{x}) - \mathbf{b}\|_2^2 \right\}, \tag{2}$$

where $D \in \mathbb{R}^{2n\times n}$ is the first-order finite difference operator in x/y-direction as introduced in the previous assignment sheet and the weight parameter λ balances the trade off between regularization and data fidelity. We solve this non-smooth optimization problem by means of the primal dual hybrid gradient (PDHG) algorithm. Therefore, you need to perform the following steps:

1. Dualize the convex non-smooth function to obtain

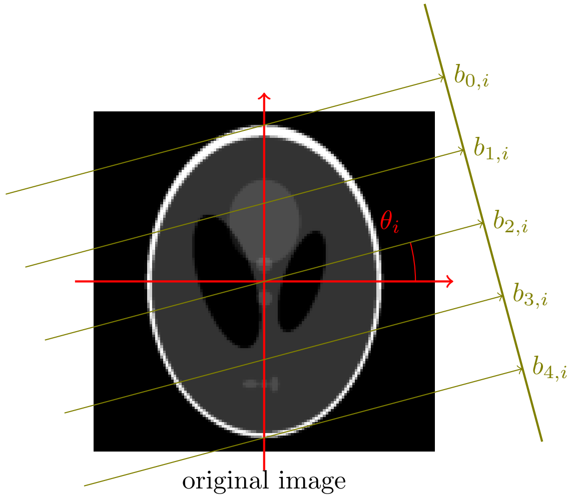
$$\min_{\mathbf{x}\in\mathbb{R}^n} \max_{\mathbf{y}\in\mathbb{R}^{2n}} \langle D\mathbf{x}, \mathbf{y}\rangle - \delta_{B_{\|\cdot\|_{2,\infty}}[0,\lambda]}(\mathbf{y}) + \tfrac{1}{2}\|\mathcal{A}(\mathbf{x}) - \mathbf{b}\|_2^2, \tag{3}$$

   where $\delta_{B_{\|\cdot\|_{2,\infty}}[0,\lambda]}(p)$ is the indicator function of the $\ell^{2,\infty}$-norm ball with radius λ. Explain all steps in your report.

2. By further dualizing the last term in (3), we arrive at the following instantiation of the PDHG algorithm:

$$\begin{aligned}
\mathbf{x}^{k+1} &= \mathbf{x}^k - \tau\left(D^T\mathbf{y}^k + \mathcal{A}^*(\mathbf{z}^k)\right)\\
\mathbf{y}^{k+1} &= \operatorname{proj}_{B_{\|\cdot\|_{2,\infty}}[0,\lambda]}\left(\mathbf{y}^k + \sigma D(2\mathbf{x}^{k+1} - \mathbf{x}^k)\right)\\
\mathbf{z}^{k+1} &= \operatorname{prox}_{\frac{\sigma}{2}\|\mathbf{z}\|_2^2}\left(\mathbf{z}^k + \sigma\left(\mathcal{A}(2\mathbf{x}^{k+1} - \mathbf{x}^k) - \mathbf{b}\right)\right),
\end{aligned}$$

   where $\mathbf{z} \in \mathbb{R}^l$ is the second dual variable due to the dualization of the quadratic ℓ^2-norm of the data fidelity term. State the resulting saddle point formulation of (3) if the quadratic ℓ^2-norm of the data fidelity term is



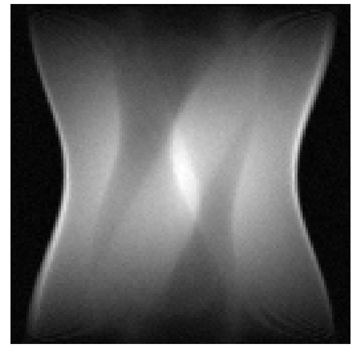


Figure 1: Illustration of the CT reconstruction problem. The measured sinogram $\mathbf{b}$ (right) is obtained by computing line integral across the original image (left) subject to additive Gaussian noise. In detail, each column of the sinogram data $\mathbf{b}$ is the column sum of the original image rotated by the angle θ_i.

dualized. Include all derivations in your report. Derive a condition for the choice of the step sizes τ and σ such that the algorithm converges and include it in your report. *Hint*: Combine all dual variables in one vector and define a corresponding operator $\mathcal{K} : \mathbb{R}^n \to \mathbb{R}^{2n+l}$. Compare the following step size selection methods:

- primal and dual step size such that
$$\tau\sigma\|\mathcal{K}\|^2 \leq 1$$
- 1-preconditioned step sizes such that
$$\begin{pmatrix} T^{-1} & -\mathcal{K}^* \\ -\mathcal{K} & S^{-1} \end{pmatrix} \succeq 0,$$
where $T = \operatorname{diag}(\tau_1, \dots, \tau_n)$ and $S = \operatorname{diag}(\sigma_1, \dots, \sigma_{2n+l})$. A suitable choice is
$$\tau_j = \frac{1}{(\mathcal{K}^*(\mathbf{1}))_j}, \quad \sigma_i = \frac{1}{(\mathcal{K}(\mathbf{1}))_i},$$
where $\mathbf{1}$ is associated vector whose entries are all equal to 1. The step-size multiplication in the algorithm above are then understood element-wise.

3. For $\lambda \in \{0.1, 1.0, 10.0\}$ compare the variants of the PDHG algorithm using $\mathbf{x}^0 = \mathbf{0}$, $\mathbf{y}^0 = \mathbf{0}$, and $\mathbf{z}^0 = \mathbf{0}$ as initial values. Therefore, plot for each choice of λ the primal energy E as defined in (2) for every PDHG variant and all iteration step in a `loglog`-plot.