Regularized Newton Sketch by Denoising Score Matching for Computed Tomography Reconstruction

Perelli, Alessandro; Andersen, Martin Skovgaard

Published in:
Proceedings of the Workshop on Signal Processing with Adaptive Sparse Structured Representations

Publication date:
2019

Document Version
Peer reviewed version

Citation (APA):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
Regularized Newton Sketch by Denoising Score Matching for Computed Tomography Reconstruction

Alessandro Perelli and Martin S. Andersen
Department of Applied Mathematics and Computer Science, Technical University of Denmark, DK, email: {alper, mskan}@dtu.dk

Abstract—In this work we aim at efficiently solving a model-based maximum-a-posterior (MAP) image reconstruction with application to low-dose transmission X-ray Computed tomography (CT). We propose to solve the regularized optimization problem by a randomized second order method called Newton iterative Hessian sketching for the Poisson likelihood function and to design a regularization term for the MAP problem exploiting the denoising score framework. By approximating the Newton step using a partial Hessian sketch only for the data fit term, it is possible to reduce the complexity by dimensionality reduction while retaining the complex prior structure by a data-driven regularizer. This work shows how to use partial Netwon sketch with denoising score matching and how to efficiently compute the gradient and the Hessian of the likelihood and regularizer. Finally, we show an example for monoenergetic X-ray CT reconstruction.

I. INTRODUCTION
Second-order methods for solving regularized optimization problems with generalized linear models (GLM) have been widely studied but, despite the superior convergence rate, one weakness relies on the computational cumbersome for calculating the Hessian matrix. Additionally, in imaging applications where the input prior is difficult to model, powerful regularization techniques are based on data-driven models or by exploiting denoising functions, like the Plug-and-play [1] approach or the RED [2], [3]. In this work we develop a second order methods which combines the iterative sketching [4], [5] for the log-likelihood GLM function and an explicit regularizer term which can be implemented by a generic denoiser through the score matching formulation. One instance of regression for GLM that we analyze is the X-ray CT application with very low X-ray intensity I0.

II. REGULARIZED NEWTON SKETCH FOR POISSON MODEL

The mono-energetic transmission CT problem is a Poisson process of the form $y_i(x) = \sum_{i=1}^{N} \log p(y_i|x) \frac{a_i e^{-a_i x}}{a_i}$, with $i = 1, \ldots, M$, $N < M$ where $x \in \mathbb{R}^{N}$ is the attenuation coefficient vector, $M$ is the product of the number of detectors and the number of X-ray projections and $a_i$ it the $i$-th row of the Radon matrix $A$. It follows that $p(y_i|x) = \sum_{i=1}^{M} \log p(y_i|x) \frac{a_i e^{-a_i x}}{a_i}$ with $p(y_i|x) = \frac{1}{y^T} (I_0 e^{-A x})$. The negative log-likelihood function is $l(x,y) = -\sum_{i=1}^{M} \log p(y_i|x) \frac{a_i e^{-a_i x}}{a_i} - A x$. The optimization problem aims at minimizing the convex objective function $f(x) = l(x,y) + p(x)$. In this work, we utilize the partially sketched Newton update [4] which means that we perform a sketch of the Hessian $\nabla^2 f(x)$ while using the exact form of the Hessian $\nabla^2 p(x)$ associated with the regularizer. Given a current iterate $x^t \in C$, the update $x^{t+1}$ is obtained by performing a constrained minimization of the second-order Taylor expansion

$$x^{t+1} = \arg \min_{x \in C} \left\{ \frac{1}{2} (x - x^t)^T Q^t (x - x^t) + \langle \nabla f(x^t), x - x^t \rangle \right\}$$

$$Q^t = (S^T \nabla^2 f(x^t,y)) S + \nabla^2 p(x^t)$$

where $S^t$ represent the count sketch matrix, generated at random and at each iteration $t$, of dimensions $D \times N$ with $D << N$. The update for (1) takes the form $x^{t+1} = x^t - Q^{-1} \nabla f(x^t)$. Given the likelihood function in Eq. (II), it results that

$$\nabla f(x,y) = A^T (y - I_0 e^{-A x}), \quad \nabla^2 f(x,y) = \left[ \text{diag} (e^{-A x}) \right] A$$

The regularizer $\rho(x)$ is chosen in order to perform a Kernel Density Estimation (KDE)-based MAP estimation using a generic denoiser $D_{\rho, \sigma^2}$ parameterized over $\eta$, to approximate the prior-based optimal MMSE estimator $D_{\text{MMSE,} \sigma^2}$, following the description in [3]. KDE generates a prior for $x$ as $\hat{p}_x(x, \sigma^2) = \frac{1}{N} \sum_{k=1}^{K} \mathcal{N}(x_i, \sigma^2 I)$, given $K$ training images $\{x_i\}_{k=1}^{K}$. By utilizing $p_x(x, \sigma^2)$ as prior for $x$, the regularization term is $\rho(x) = \log \hat{p}_x(x, \sigma^2)$. From Eq. (1), it is needed to evaluate the expression for $\nabla_x \rho(x) = \nabla_x \log \hat{p}_x(x, \sigma^2)$ and $\nabla^2_x \rho(x)$. It was shown in [6] that

$$\nabla_x \log \hat{p}_x(r, \sigma^2) = \frac{1}{\sigma^2} \left( \hat{D}_{\text{MMSE,} \sigma^2}(r) - r \right), \quad r = x + N(0, \sigma^2 I)$$

with $\hat{D}_{\text{MMSE,} \sigma^2}(r) = \mathbb{E}[x|r]$ the MMSE estimator of $x \sim \hat{p}_x$. Calculating $\hat{D}_{\text{MMSE,} \sigma^2}(r)$ is unfeasible in high dimensions, so the idea is to replace the MMSE estimator with a generic denoiser $D_{\sigma^2}(r)$. This is possible by invoking connection between the score by matching and denoising [7]. It was shown in [8] that if we choose a function such that

$$\theta(x, \eta) = D_{\sigma^2}(x) - x, \quad \frac{\partial \theta(x, \eta)}{\partial x} = \frac{1}{\sigma^2} (J[D_{\sigma^2}(x)] - I)$$

then the solution of the least square minimization for the best fit of the score $\hat{\eta} = \arg \min_{\eta} \| \theta(x, \eta) - \nabla_x \log \hat{p}_x(x, \sigma^2) \|^2_2$ and its gradient is equivalent to the solution of the denoising problem $\hat{\eta} = \arg \min_{\eta} \| D_{\sigma^2}(x + N(0, \sigma^2 I)) - x \|^2_2$ which holds also in the case where the parameter of the denoiser $\eta$ is not optimal. $J[D_{\sigma^2}(x)]$ is the Jacobian matrix of the denoising function of dimension $N \times N$ which is computationally prohibitive to calculate and store at each iteration $t$. We exploit the property that $\mathbb{E}_n[\nabla^2 J] = \text{Trace}[J]I$ if $n \sim \mathcal{N}(0,I)$ is independent of $J$. Given a noise realization $n$, we propose to compute the vector $J[D_{\sigma^2}(x)]I$ instead of the matrix $J[D_{\sigma^2}(x)]$. The expectation is calculated by Monte-Carlo simulation and practically [9] one run is sufficient for a good approximation.

III. SIMULATION

The method is tested on a CT simulation with an input image of dimension $N = 128 \times 128$, generated using the TomoPhantom toolbox (model 11) [10], X-ray source photons $I_0 = 10^4$ with an exact Poisson model. The forward radon matrix is implemented using AIRTool [11] which generate an explicit matrix with $N_{\text{angles}} = 1000$ and 256 detector elements. The dimension of the sketch is $D = N/A$. Non Local Means [12] is chosen as denoiser $D_{\sigma^2}$ with a-prior fixed parameters $\eta$. The MSE convergence plot is depicted in Fig. 2.

IV. CONCLUSION AND FUTURE WORKS

We propose a method for a second order optimization which combine dimensionality reduction through matrix sketching and regularization by denoising, applied to low dose X-ray CT reconstruction with Poisson model.

ACKNOWLEDGMENT

The research received funding from EU Horizon 2020 programme under the Marie Skł.-Curie grant no. 713683 (COFUNDfellowsDTU).
Fig. 1. MSE plot over CPU time and the objective error plot over iterations.

Fig. 2. Reconstructed image and residual error.

REFERENCES


