Total Variation Regularization with Barzilai and Borwein Optimization Method

Jenni Helander, Sini Lehtonen and Patrik Lod

INTRODUCTION

In this project work we have a linear inverse problem

$$\mathbf{m} = A\mathbf{f} + \epsilon, \tag{1}$$

where $\mathbf{m} \in \mathbb{R}^k$ is the measurement data, $\mathbf{f} \in \mathbb{R}^n$ is the one we want to recover and A is the measurement matrix. To solve $\mathbf{f}$ we will be using total variation regularization.

The tomographic data will be measured by ourselves and we are going to use a motorcycle toy from a Kinder-surprise as an object of imaging. The resolution of the data will be quite big and we will use only a few angles of the measured data. Therefore we will be using the optimization method of Barzilai and Borwein. This is a matrix-free implementation of total variation regularization.

In total variation regularization we decided to use a resolution-based choice of regularization parameter.

METHODS AND MATERIALS

Total variation regularization

The idea of total variation regularization is to find a vector $\mathbf{f} \in \mathbb{R}^n$, which minimizes the expression

$$\|A\mathbf{f} - \mathbf{m}\|_2^2 + \alpha\|L\mathbf{f}\|_1, \tag{2}$$

where L is a finite difference matrix and $\alpha > 0$ is a regularization parameter. In this project we have a two-dimensional case and therefore $f \in \mathbb{R}^{n \times n}$. To use the formula (2) we treat the matrix f as a vector $\mathbf{f} \in \mathbb{R}^{n^2}$. Now we minimize the expression

$$\|A\mathbf{f} - \mathbf{m}\|_2^2 + \alpha\|L_H\mathbf{f}\|_1 + \alpha\|L_V\mathbf{f}\|_1, \tag{3}$$

where L_H and L_V are horizontal and vertical difference matrices and $\alpha > 0$ is a regularization parameter.

For example in the case $f \in \mathbb{R}^{3\times3}$ implies $\mathbf{f} \in \mathbb{R}^9$. Now the horizontal difference matrix is

$$L_H = \frac{1}{9}\begin{bmatrix} -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

and the vertical difference matrix is

$$L_V = \frac{1}{9}\begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{bmatrix}.$$

To solve the problem (3), we use the gradient descent minimization method of Barzilai and Borwein. This approach is good for large-scale implementations. Let's denote an objective function $G : \mathbb{R}^{n^2} \to \mathbb{R}$

$$\begin{aligned} G(\mathbf{f}) &= \|A\mathbf{f} - \mathbf{m}\|_2^2 + \alpha\|L_H\mathbf{f}\|_1 + \alpha\|L_V\mathbf{f}\|_1 \\ &= \|A\mathbf{f} - \mathbf{m}\|_2^2 + \alpha\sum_{i,j=1}^{n}|f_{i,j} - f_{i,j-1}| + \alpha\sum_{i,j=1}^{n}|f_{i,j} - f_{i-1,j}|, \end{aligned}$$

where indices i and j denote rows and columns, respectively from the matrix $f \in \mathbb{R}^{n\times n}$. We use the periodic boundary conditions. Because G is not continuously differentiable, we replace the absolute value function $|t|$ by

$$|t|_\beta = \sqrt{t^2 + \beta},$$

where $\beta > 0$ is very small. We denote the modified function G by G_β.

Because the method of Barzilai and Borwein is gradient-based, we calculate the gradient of the function G_β:

$$\begin{aligned} \nabla G_\beta(\mathbf{f}) &= \nabla\|A\mathbf{f} - \mathbf{m}\|_2^2 \\ &\quad + \alpha\nabla\left(\sum_{i,j=1}^{n}|f_{i,j} - f_{i,j-1}|_\beta + \sum_{i,j=1}^{n}|f_{i,j} - f_{i-1,j}|_\beta\right). \end{aligned}$$

The gradient of the first term is

$$\nabla\|A\mathbf{f} - \mathbf{m}\|_2^2 = 2A^TA\mathbf{f} - 2A^T\mathbf{m}$$

and the second term consists of the components

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{f}_{k,l}}\left(\sum_{i,j=1}^{n}|f_{i,j} - f_{i,j-1}|_\beta + \sum_{i,j=1}^{n}|f_{i,j} - f_{i-1,j}|_\beta\right) \\ &= \sum_{i,j=1}^{n}\frac{\partial}{\partial \mathbf{f}_{k,l}}\left(\sqrt{(f_{i,j} - f_{i,j-1})^2 + \beta} + \sqrt{(f_{i,j} - f_{i-1,j})^2 + \beta}\right) \\ &= \frac{f_{k,l} - f_{k,l-1}}{\sqrt{(f_{k,l} - f_{k,l-1})^2 + \beta}} - \frac{f_{k,l+1} - f_{k,l}}{\sqrt{(f_{k,l+1} - f_{k,l})^2 + \beta}} \\ &\quad + \frac{f_{k,l} - f_{k-1,l}}{\sqrt{(f_{k,l} - f_{k-1,l})^2 + \beta}} - \frac{f_{k+1,l} - f_{k,l}}{\sqrt{(f_{k+1,l} - f_{k,l})^2 + \beta}}. \end{aligned}$$

The Barzilai and Borwein optimization strategy is an iterative method where $\mathbf{f}^{(1)}$ is the initial guess and the next step is calculated by

$$\mathbf{f}^{(r+1)} = \mathbf{f}^{(r)} - \delta_r\nabla G_\beta(\mathbf{f}^{(r)}), \tag{4}$$

where δ_r is the steplength parameter. Let's denote $y_r = \mathbf{f}^{(r)} - \mathbf{f}^{(r-1)}$ and $g_r = \nabla G_\beta(\mathbf{f}^{(r)}) - \nabla G_\beta(\mathbf{f}^{(r-1)})$ and now the steplength parameter is defined by

$$\delta_r = \frac{y_r^T y_r}{y_r^T g_r}.$$

All of the theory above is based on the book [1].

The choice of the regularization parameter

There are many different methods to choose the regularization parameter α and to this project work we chose the resolution-based choice of total variation regularization parameter. In this method we calculate reconstructions for a few different resolutions and discrete TV norms for them. The optimal α will be the smallest one which gives the norm values very close to each other. Let's define the TV norm by

$$\|\mathbf{f}\|_{\mathrm{TV}} = \|L_H\mathbf{f}\|_1 + \|L_V\mathbf{f}\|_1, \tag{5}$$

where L_H and L_V are horizontal and vertical difference matrices, respectively, as in (3) and we use periodic boundary conditions. This method is introduced in the article [2].

The measurement data

For this project work we chose a motorcycle toy from a Kinder-surprise, shown in Figure 1, to be the object from which we took the X-ray images. To get the measurement data we used the fan-beam geometry with 180 angles. To simplify the problem we consider that the beam was parallel and to have sparse data we chose to use only 30 angles of the measured 180.

Figure 1: A photo of the object to image.

We placed the motorcycle vertically and we took the cross-section of the bike at the rear wheel. In Figure 2 is the first X-ray image and the red line indicates the chosen cross-section.

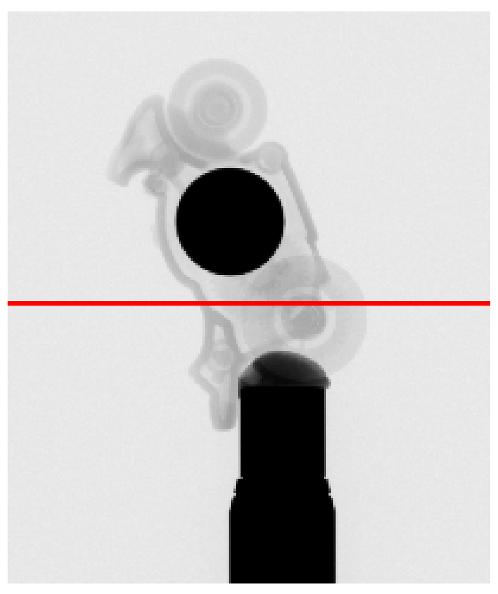

Figure 2: The chosen cross-section for the reconstruction.

RESULTS

We used the resolution-based choice of regularization parameter with two different resolutions: 32×32 and 44×44. For these resolutions we computed reconstructions with Barzilai and Borwein method with 50 different regularization parameters α from the interval $[0.1, 20]$. In the Barzilai and Borwein method we noticed that 2000 iterations is enough so we used it as a number of iterations.

We computed TV norms for the previously mentioned reconstructions with the formula (5). For every α we computed the difference of the two TV norms calculated from the mentioned resolutions. The results can be seen in Figure 3 where the TV norm differences are as a function of α.

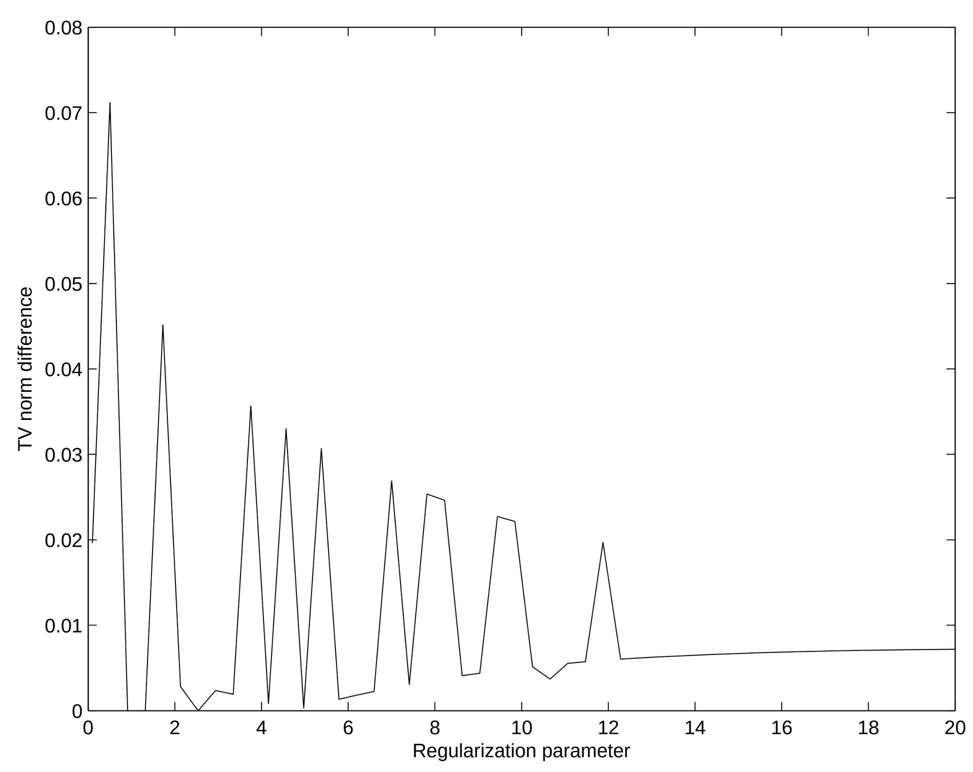

Figure 3: A function to choose α.

From the Figure 3 we can see that the best choice for regularization parameter α is approximately 13. With this parameter choice we computed the reconstruction with original sparse data. The final reconstruction is the size of 494×494 pixels and it can be seen in the Figure 4.

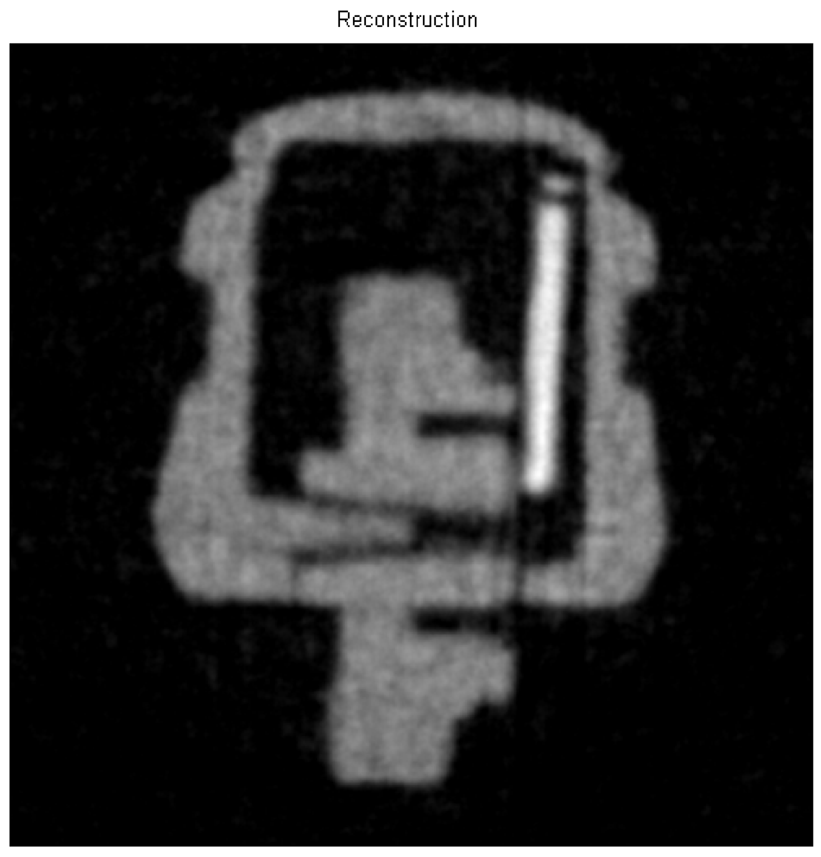

Figure 4: The reconstruction using total variation regularization.

We computed also the reconstruction for the data with full 180 angles using Matlab's routine `iradon.m`. This reconstruction is shown in Figure 5. We computed this reconstruction to have something to compare the results of total variation regularization. The reconstruction images were normalized to the same scale and then the relative norm difference for these two reconstructions got the value $17,0\%$.

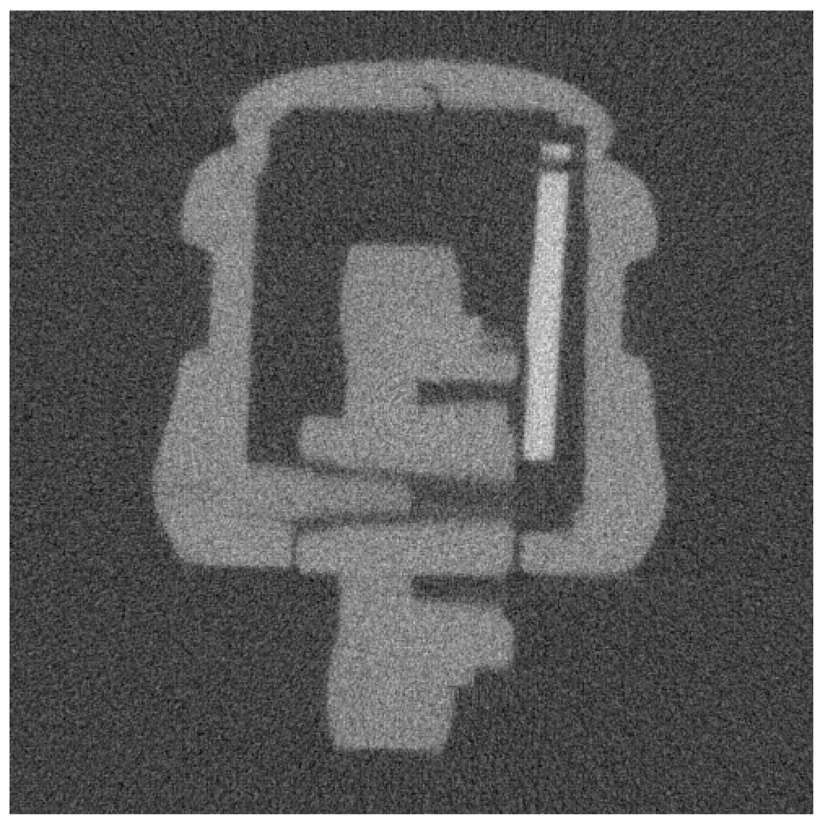

Figure 5: The full-angle reconstruction.

DISCUSSION

From the Figure 4 we see that the result is quite good. The structure of reconstruction is clear and the shape is very recognizable. The edges look quite sharp and we can distinguish the bottom part of the figure is separate from the other structure. This part of the figure is the rear wheel of the motorcycle and in this cross-section it is not even supposed to be attached to the other parts.

In the reconstruction there is a visible amount of noise in the background and also a little in the motorcycle. The background should be black but instead there are some grey spots in it because of the noise. Not every little detail can be seen in the reconstruction. For example in the top of the image there should be a visible joint which can be seen in Figure 5. The lack of details results from having data only from 30 angles.

From the reconstruction we can see that there is on the right side a shape made of different material than the other parts of the motorcycle. This part is made of material that attenuates X-rays more than the other parts and therefore it can be seen lighter in the Figure 4.

In Figure 5 is shown the full-angle reconstruction using Matlab's routine `iradon.m`. This reconstruction is better than the sparse one because we are using all 180 angles, but it is not perfect either. Also in this image there is some amount of noise and because of it the background for example is not as smooth as it could be.

These two reconstructions, the full-angle reconstruction and the sparse reconstruction, are after all quite similar having only 17.0% norm difference. As a conclusion from the results of this project work we can say that the total variation regularization with Barzilai and Borwein optimization method using the resolution-based regularization parameter choice gives good and usable results even with only few angles.

References

[1] Mueller J L and Siltanen S, *Linear and Nonlinear Inverse Problems with Practical Applications*, SIAM 2012.

[2] Kati Niinimäki, Matti Lassas, Keijo Hämäläinen, Aki Kallonen, Ville Kolehmainen, Esa Niemi, and Samuli Siltanen, *Multi-resolution Parameter Choice Method for Total Variation Regularized Tomography*, July 2014.