Total Variation
Regularization and L-curve method for the selection of regularization parameter

Vivek Agarwal
ECE 599
Summer 2003
Abstract

The objective of this project is three fold. First, is to present a concise and self-constrained overview on the L-curve method for the selection of regularization parameter. Second, presents an elaborate overview of total variation (TV) images restoration models and few existing numerical methods for solving. Third, is to present a brief discussion on the mathematical background of the TV method and results of my implementation of fixed-point iterative TV regularization.

Most of the reconstruction algorithms smooth out the edges of the reconstructed images but total variation regularization method preserve the edge information without any prior knowledge about the blurred image geometric details. The regularization parameter is another important criterion that should be carefully chosen for a good regularized output. The approach towards the objective involves a comprehensive review of various ways of implementing total variation regularization that are applicable to Tikhonov regularization and Truncated Singular Value Decomposition (TSVD). The L-curve methodology and its criterion for the selection of regularization parameters in comparison to other methods is also evaluated in order to select the regularization parameters and compares advantages and limitations. The fixed-point iteration method for obtaining total variation regularization is selected and is applied to an image that is corrupted with Gaussian noise and whose Point Spread Function (PSF) is also Gaussian.

The outcome is comprised of a reconstructed image from a noisy, blurred image using the total variation regularization method. A least square method output is also presented.
Table of Contents

Abstract............................................................................................................................... i

1. Introduction......................................................................................................................... 1
   1.1 Motivation......................................................................................................................... 2
   1.2 Background....................................................................................................................... 2

2. Theory ................................................................................................................................. 4
   2.1 Theory on L-curve and its numerical analysis: ................................................................. 4
   2.2 Total Variation (TV) Regularization and its numerical analysis: ................................. 7

3. Implementation .................................................................................................................. 12

4. Conclusions......................................................................................................................... 21

References.............................................................................................................................. 23

Appendix................................................................................................................................. 27
1. Introduction

Regularization and optimization of an inverse problem has been an active area of research for more than a decade. An inverse problem is an ambiguous problem, the best way of describing it is according to a famous quotation by J.B Keller [18], “We call two problems inverse of one another if the formulation of each involves all or part of the solution of the other. Often, for some historical reasons one of the two problems has been studied extensively for some time, while the other has never been studied and is not so well understood. In such cases, the former is called the direct problem and the latter is called the inverse problem.” An example, of such problem might be the formation of an original image, i.e., an image captured using an imaging device (pure data). The obtained image is blurred and corrupted by a noise (direct problem). The process of recovering the original data from the blurred and corrupted image is inverse problem. There is often a trade-off between the regularized output and the original sets of data. In order to obtain a balance or minimize the trade-off, the optimal selection of the regularization parameter ($\lambda$) becomes important. There are various reasons for the proper selection of $\lambda$, which are discussed in the following sections. The most general form of regularization is the Tikhonov regularization. It is given by the expression, $\|Ax-b\|^2 + \lambda \|x\|^2$ where $\|Ax-b\|$ - least Square solution, $\|x\|$ - regularized output, ‘$\lambda$’ the regularization parameter and $\| \|$ represents the norm.

The most common concern of the inverse problem is the solution being ill posed which is opposite to the well-posed solution. J.Hadamard introduced the concept of well-posed problem in a paper published in 1902. A problem is said to be well posed if its solution is unique and exits for the arbitrary and continuous data set. Reverse of well pose is the problem of ill-posed problem i.e. the solution is not unique. The unique solution computed is unacceptable in a physical sense because it’s an approximate solution. It does not reproduce the exact solution but has experimental errors and the original data is not completely reproduced.

Fundamentally the inverse problem is that of solving a Fredholm first kind of integral equation whose kernel is smooth and non-degenerate. This problem is ill posed even for a small perturbation, which produces large oscillation for a small change in the
data. Standard regularization methods like Tikhonov regularization and Truncated Singular Value Decomposition (TSVD) assume the data set to be smooth and continuous. Most of the regularization techniques produce results that are smooth. The total variation regularization method is not the same. This method does not assume data sets to be smooth or continuous but rather only condition or prior knowledge required is that the data (image) has discontinuity.

1.1 Motivation

Regularization of the inverse problem in imaging has ignited a new perspective for looking at the problems in image processing. It provides us with an opportunity to rediscover the origin of the source (images) by performing a reverse operation, i.e., recovering original information from the information at hand. This phenomenon is not natural but is very useful in the study of astronomy and solar systems. It opens up a new horizon of research in the field of image processing and its application. One of the applications is the reconstruction/restoration of images by preserving its edge information i.e. the total variation regularization technique. While most of the regularization methods use prior knowledge and also tend to smooth the restored images, total variation preserves edges and does not require any prior information about the blurred image. This challenge to solve numerically the total variation regularization was the motivation for the work on this topic.

1.2 Background

A substantial amount of research has been completed in the field of regularization and optimization of inverse ill-posed problems. The least square method was a first step in this direction but it was realized that the method failed and this led to the Tikhonov regularization. Most regularization techniques are based on the assumption of smooth and continuous information from the image to be reconstructed; total variation is independent of these assumptions and it preserves the edge information in the reconstructed image. Another important aspect to be taken into consideration in order to obtain a regularized
output is proper and optimal selection of regularization parameters. There are various methods for the selection of regularization parameters. Some of them are:

- Discrepancy Principle
- Generalized Cross Validation
- L-curve

Various methods were introduced [15] [16] [17] for the selection of parameters in order to obtain regularized output. In recent years, L-curve despite its limitations has gained attention for computing the selection of regularization parameters. It’s a log-log plot of the regularized solution against the squared norm of the regularized residual for a range of values of regularization parameters. The numerical computation and limitation of the L-curve is explained in [1] and [2].

If the image is blocky, then in order to restore its edge information the TV method was introduced [17] by Rudin et al in 1992. The time marching scheme introduced by Rudin et al proposes to denoise image by minimizing the total variation norm of the estimated solution. A constrained minimization algorithm was derived based on a time dependent nonlinear PDE, the constrained was determined from the noise statistics. This method is equivalent to steepest descent method. It uses a projected gradient method to handle the noise but has its own limitation. However the implicit method can be applied to avoid this limitation but it results in a non-linear operation to obtain a linear solution. Later the fixed-point lagged diffusivity iteration method [3] [5] [9] [13] was introduced.

This method employed linearization of a non-linear differential term by lagging the diffusion coefficient $1/\lambda$ behind one iteration. This method is proved by generalized Weiszfeld’s method to be a monotonically convergent [3] [23], in the sense that the objective function evaluated at the each iteration monotonically decreases and the convergence rate is linear. T.Chan and Zhou [8] introduced the Newton type method. This method is similar to [17]. C.R. Vogel in his book [19] and T.F. Chan et al [24] introduced a primal-dual Newton method. The basic idea behind the introduction of the primal-dual method was to avoid some of the singularity caused by the non-differentiability of the quantity in the TV definition. It was also proved experimentationally [24] between the fixed point, time marching and primal-dual methods that fixed point and time marching are linearly convergent whereas primal-dual
is quadratically convergent. Moreover, the domain of convergence of the primal-dual is smaller when compared to the fixed-point iterative method.

The report is organized in the following manner. Section 2 consists of the TV regularization technique and L-curve methodology. The mathematical formulation is also discussed in section 2. Section 3 consists of experimental results, followed by conclusion at the end.

2. Theory

In this section, review on L-curve is presented in subsection 2.1 and comprehensive explanation on total variation regularization is presented in subsection 2.2.

2.1 Theory on L-curve and its numerical analysis:

In practice, most of the regularization of the inverse problems suffers from a trade-off between the “size” of the regularized solution and the quantity of the fit that it provides to the given data. Different regularization techniques differ on the basis on how they minimize this trade off. This trade-off can be controlled by the selection of proper regularization parameter. Various methods have been developed for the optimal selection of these regularization parameters.

- Discrepancy principle
- L-curve
- Generalized cross validation method

The need for optimizing the regularization parameter and effect of over, under and approximate sampling is demonstrated in [1]. Regularization is important for solving inverse problems because the output of least squares is affected by data error and rounding error. By introducing regularization these errors are damped. However [1], showed if the regularization is too much, regularized solution does not fits the given data properly as the residual error $\|Ax-b\|^2$ is too large and if the regularization is too small, the fit will be good but data error will be more.

From the above-mentioned methods for selection of regularization parameters, L-curve method has gained attention is recent years. “L-curve is a log-log plot between the
squared norm of the regularized solution and the squared norm of the regularized residual for a range of values of regularization parameter”. The figure 1 shows the theoretical representation of the L-curve of Tikhonov regularization. The name implies that the shape of the curve should resemble ‘L’ letter closely.

Figure 1: L-curve for Tikhonov Regularization (Theoretical)

For the explanation purpose, we take in to consideration the Tikhonov regularization analysis and compute the Singular value decomposition (SVD) expression. Expression for least squares and regularized output is computed in order to explain the L-curve.

The Tikhonov regularization is given by

\[ \| Ax - b \|^2 + \lambda \| x \|^2 \]

Where

\( \| Ax - b \|^2 \) - Least Square or residual norm

\( \| x \|^2 \) - Regularized norm and

\( \lambda \) – Regularization parameter.

The SVD is applied to the (1) to obtain a Tikhonov solution which is represented as

\[ \eta = \| x \|^2 = \sum_{i=1}^{n} \left( f_i \frac{u_i^T b}{\sigma_i} \right)^2 \]

(2)
The equation (2) and (3) are important for the computation of L-curve. One crucial detail is the precise characterization of the corner of the L-curve. Hansen and O’Leary computed the point of maximum curvature on the L-curve. The point of maximum curvature can be given by the formulation as below.

\[ \hat{\eta} = \log \eta \quad \text{and} \quad \hat{\rho} = \log \rho \]  

(4)

The curvature \( k(\lambda) \) is defined by the expression below

\[
k(\lambda) = \frac{\hat{\rho}'\hat{\eta}'' - \hat{\rho}''\hat{\eta}'}{(\hat{\rho}')^2 + (\hat{\eta}')^2}^{3/2}
\]  

(5)

Where prime denotes the differentiation with respect to \( \lambda \). From the expression in (2) and (3) we can derive the curvature as

\[
k(\lambda) = \frac{\eta\rho [\lambda\eta + \lambda^2 \rho] + [\eta\rho]^2 / \rho'}{(\lambda^2 \eta^2 + \rho^2)^{3/2}}
\]  

(6)

The point of maximum curvature corresponds to the curve portion of the L-curve. A plot is shown in figure 2 between regularization parameter \( \lambda \) and the curvature \( k(\lambda) \).

![Figure 2: A plot of \( \lambda \) versus \( k(\lambda) \)](image-url)
In 1996, Reginska proved that the shape of the log-log L-curve is always concave for the smallest singular value and largest singular values. She also proved that as long as the SVD coefficients \( \left| u_i^T b \right| \) decreases monotonically or increases monotonically with \( i \) or are constant, then there is a good reason to believe that the log-log L-curve is concave. Based on the explanation, the criterion for the selection of the optimal regularization parameter is mainly two fold.

- Taking into consideration both the residual norm and the solution norm
- Some methods use the value of the maximum curvature.

These two are the main criteria utilized in L-curve. As mentioned there are various other methods for the selection of the regularization parameter apart from L-curve. Each of the methods have own its merits and limitations. Some of the merits of the L-curve are that it’s robust and has the ability to treat perturbation caused by correlation noise. Also L-curve is superior to GCV, which produces severe under smoothing effect. As a part of demerits, L-curve also has the limitation of having asymptotic property, i.e. non convergent.

2.2 Total Variation (TV) Regularization and its numerical analysis:

Total variation is a regularization technique that does takes in to consideration the information that the data set is blocky and discontinuous. Most of the regularization methods assume the data sets to be smooth and continuous, but total variation doesn’t assume the same. Its measures the discontinuities in the image data set. The general Tikhonov regularization is given in equation (1). From the above expression, it can be seen that the penalty function in this case is expressed as

\[
x_\lambda = \int x^2 dx
\]  

This is the general form of Tikhonov regularization. Moreover it is of norm \( l_2 \). Since smooth solution is desirable in many applications while others require discontinuity or steep gradient to be computed. One approach is to replace norm \( l_2 \) in Tikhonov regularization with the norm \( l_1 \), i.e., the 1-norm of the first spatial derivation of the
solution. This is called the total variation (TV) regularization. This method will help to obtain the discontinuities or steep gradients in the restored image. The total variation can be expressed as

\[ \| Ax - b \|_2^2 + \lambda TV(x_{\lambda}) \]  

(7)

and

\[ TV(x_{\lambda}) = \int_{\Omega} \sqrt{\| \nabla x_{\lambda} \|^2 + \beta^2} \, dx \, dy \]  

(8)

However, \( TV(x_{\lambda}) \) is not differentiable at zero. So in order to avoid this difficulty a small positive constant value is added to the equation (9)

\[ TV(x_{\lambda}) = \int_{\Omega} \sqrt{\| \nabla x_{\lambda} \|^2 + \beta^2} \, dx \, dy \]  

(9)

Thus the total variation expression can now be expressed as

\[ T(x) = \frac{1}{2} \| Ax - b \|_2^2 + \lambda \int_{\Omega} \sqrt{\| \nabla x_{\lambda} \|^2 + \beta^2} \]  

(10)

The quantity \( \sqrt{\| \nabla x_{\lambda} \|^2 + \beta^2} \) is known as the gradient magnitude. This provides us with the information about the discontinuities in the image.

The “norm” can be given by a general expression as

\[ \| x \|_p = \left( \sum_i |x_i|^p \right)^{\frac{1}{p}} \]  

(11)

\( \| x \|_1 = \sum_i |x_i| \) is also known as Manhattan norm because it corresponds to the sum of the distances along the coordinate axes.

\( \| x \|_2 = \sqrt{\sum_i x_i^2} \) is also known as Euclidean norm, or the vector length.
The change to the 1-norm has a dramatic effect on the computation of the solution. Its have been proved by P.C. Hansen that the solution consists of polynomial pieces, and the degree of polynomials is $p-1$.

The minimization of the equation (10) is a penalty approach to the solution of the constrained problem. There are various methods to obtain this minimization

- Time Marching [Rudin et al, 1993]
- Steepest Descent and Newton’s method
- Lagged diffusivity fixed point iterative method
- Primal-Dual Method.

In this report, lagged diffusivity fixed point iterative method of C.R. Vogel and M. Oman is discussed. Since both the lagged diffusivity and primal-dual method requires solution of the non-sparse linear system at each iteration conjugate gradient method (CG) is applied without preconditioning. However CG with preconditioning can be applied to obtain a fast convergence.

In order to minimize the equation (10) the gradient is computed. Therefore in order to compute the gradient the equation (10) is differentiated with respect to ‘x’ and then using integration by parts along with Neumann boundary conditions, the gradient of the total variation function is obtained. The gradient of the total variation can be represented as

$$\nabla TV(x_\lambda) = -\nabla \cdot \left( \frac{\nabla x}{|\nabla x|} \right)$$

(12)

The integration by parts yields a non-linear partial differential equation (PDE), which is represented as

$$\nabla T(x) = -\lambda \nabla \cdot \left( \frac{\nabla x}{|\nabla x|} \right) + (A^*Ax - A^*b) = 0$$

(13)
The nonlinearity of equation (13) poses number of computation challenges [5]

- The operator $\nabla \cdot \left( \frac{\nabla x}{|\nabla x|} \right)$ is highly nonlinear
- The ill conditioning of the operator $\nabla \cdot \left( \frac{\nabla x}{|\nabla x|} \right)$ (when linearized, it is a second order elliptic operator) and $K*K$ (it is convolution operator) can have numerical difficulties.

The equation (13) is solved using the conjugate gradient method without preconditioning information. The method proceeds by generating vector sequences of iterates (i.e., successive approximations to the solution), residuals corresponding to the iterates, and search directions used in updating the iterates and residuals. CG maintains 3 vectors at each step, the approximate solution $x$, its residual $r=Ax-b$, and a search direction, which is also called a conjugate gradient. At each step $x$ is improved by searching for a better solution in the direction, yielding an improved solution. Below is a algorithm to compute the conjugate gradient

1. $v = 0$
2. $f_v = \text{initial guess}$ \hspace{1cm} // Initial guess can be flat image//
3. $g_v = A f_v + b$ \hspace{1cm} // Compute the initial gradient for the solution//
4. $p_v = -g_v$ \hspace{1cm} // Search for the new direction//
5. $\delta_v = ||g_v||^2$

The first 5 steps just computes the initial gradients and search for new direction to proceed. The iterative process after this will converge slowly to the best solution. CG iteration begins

6. $h_v = A p_v$
7. The new line search parameter is computed $\tau_v = \delta_v / (p_v, h_v)$
8. An approximate solution obtained is upgraded $f_{v+1} = f_v + \tau_v p_v$
9. The gradient computed is upgraded $g_{v+1} = g_v + \tau_v h_v$
10. The norm of the new gradient is computed $\delta_{v+1} = ||g_{v+1}||^2$
11. The ratio of new norm to the previous norm is computed $\beta_v = \frac{\delta_{v+1}}{\delta_v}$

12. The search for the line is upgraded $p_{v+1} = -g_v + \beta_v p_v$

**The iteration ends**

The above mentioned algorithm procedure computes the conjugate gradient of the equation (13) and gradually converges in a linear manner. The main difference between the fixed-point iterative method and primal dual method is the rate of convergences and the domain of convergences. The later converges faster and in a quadratic manner. The above steps are used within the loop of fixed-point iterative total variation regularization method. The algorithm for fixed-point iterative method is presented below.

The algorithm begins with initialization

1. $v = 0$
2. $f_0 = \text{initial guess}$  // Initial guess includes flat image assumptions/

**The fixed-point iteration begins**

3. $L_v = L(f_v)$. This obtained by discretization of the equation of the Total variation computation
4. The gradient of the non-linear equation is computed.

$$g_v = K^T (K f_v - b) + \lambda L_v f_v$$

5. The approximate Hessian matrix is computed. Hessian matrix is a matrix that consists of a second order partial differential equation

$$H = K^T K + \lambda L_v$$

6. The following method uses quasi-Newton method because of the fact that the method is less sensitive to round-off error than the fixed point form

$$s_{v+1} = -H^{-1} g_v$$
7. The approximate solution obtained at each iteration is updated

\[ f_{v+1} = f_v + s_{v+1} \]

*Fixed-point iteration ends.*

The above fixed point iterative method is convergent and with increase in iteration the method converges linearly. One important issue of concern here is the regularization parameter selection. Though it’s not the primary concern but it is important to choose the appropriate value of \( \lambda \) so that the desired restoration can be obtained. In order, to make a good selection \( \lambda \) a graph can be plotted between the error value of the true image and the restored image for different value of \( \lambda \) for a fixed number of iteration. The expectation is that the curve should descend and minimize at a particular value. The point of minimum descend can be taken as a value of \( \lambda \) for the experimentation.

The implementation of the fixed-point iterative method of total variation regularization is done in MATLAB. The optimization toolbox of MATLAB is utilized.

### 3. Implementation

The equation (1) represents the Tikhonov regularization method. The penalty function in equation (1) is of norm \( l_2 \) but total variation regularization method involves penalty function of norm \( l_1 \). Apart from being 1-norm equation, the equation (10) is also non-linear thereby the minimization of the penalty function is more time consuming and complicated when compared to the Tikhonov regularization. The figure 3 shows the original image used in total variation regularization method. The image is 32 by 32 but has been scaled using the matlab command.
In this regularization technique image vector is used. So original matrix of image is converted into vector using the matlab command `reshapes (...)`. The image in figure 3 is corrupted with Gaussian noise of zero mean and variance of 0.05. The image vector corrupted by noise is convolved with a Point Spread Function (PSF), which is also gaussian in nature. The PSF is computed using the formulation

$$ A_{i,j} = \frac{1}{\sqrt{2\pi}\sigma} \exp \left( -\frac{1}{2} \left( \frac{i-j}{\sigma} \right)^2 \right) $$

(14)

The $\sigma = 7$ was chosen. The PSF ($A$) of size $n$ by $n$ is represented as Kronecker product such that $A \otimes A$ is $n^2 \times n^2$. This can be obtained using the matlab command `kron (A, A)`. The image vector is convolved with the PSF to blur the image. The figure 4 shows the original image corrupted by noise and the blurred image is obtained after convolving with the PSF.
The blurred image in figure 4(b) shows no information about the original image. The objective is to restore the original image from the blurred image i.e., inverse the operation. Least square method was applied to restore the image. It is one of the unbiased approaches with smallest variance values. The expression in equation (15) is minimized using the least square.

\[ \| Af - g \|_2^2 \rightarrow \min \]  

\[ (14) \]

Where

\( A \) - Point Spread Function (PSF); \( f \) – Image to be restored; \( g \) – blurred image

The singular value decomposition of the A is computed to obtain the singular value. The least square solution is given by the expression 15

\[ f_{LS} = (A^T A)^{-1} A^T (g + \varepsilon) \]  

\[ (15) \]

Where \( \varepsilon \) is the noise added to the image. The figure 5 shows the image restoration after applying least squares.
The figure 5 is incorrect restoration. The reason for the least square method failure in restoring the image is because it’s highly dependent on singular values. So regularization techniques are applied to restore the image from the blurred and noise corrupted image. Tikhonov regularization method restores the image, which is smooth, but total variation regularization method restores the images by preserving the edges information. Total variation method involves minimization of the non-linear penalty function as given by the equation (13). Analytical computation of the equation (13) is not possible so numerical analysis method is applied to linearize the equation (13). The numerical analysis of the total variation penalty function is explained in the section 2.2. The fixed-point iteration method of total variation was selected to minimize the non-linear expression

$$f_{i+1} = \left[ A^T A + \lambda L(f_i) \right]^{-1} A^T g$$

$$= f_i - \left[ A^T A + \lambda L(f_i) \right]^{-1} \text{grad } T(f)$$ (16)

The fixed-point iteration is derived by first setting the $\text{grad } T(f_i) = 0$ to obtain the expression $A^T A + \lambda (L(f))f = A^T g$. The regularization operator $L$ is computed [19]. The discretization of total variation penalty function $T(f)$ is performed using the expression (17). The two-dimensional penalty function is discretized as follows
\[ T(f) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \psi \left( \left( D_{x_{i,j}}^x f \right)^2 + \left( D_{y_{i,j}}^y f \right)^2 \right) \]  

(17)

Where

\[ D_{x_{i,j}}^x f = \frac{f_{i,j} - f_{i-1,j}}{\Delta x}, \quad D_{y_{i,j}}^y f = \frac{f_{i,j} - f_{i,j-1}}{\Delta y} \quad \text{and} \quad \psi(t) = 2\sqrt{t + \beta^2} \]

The factors \( \Delta x \) and \( \Delta y \) is removed from the above expressions. The gradient of the equation (17) is computed which is given by the expression

\[
\frac{d}{dt} T(f) = \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{n} \psi'_{i,j} \left( \left( D_{x_{i,j}}^x f \right)^2 + \left( D_{y_{i,j}}^y f \right)^2 \right) 
\]

(18)

Where

\[ \psi'_{i,j} = \psi' \left( \left( D_{x_{i,j}}^x f \right)^2 + \left( D_{y_{i,j}}^y f \right)^2 \right) \]

The resulting gradient is vectorized and corresponding lexicographic ordering of the two dimensional array components is obtained. The resulting \( \text{diag}(\psi'(f)) \) of dimension \( n_x n_y \times n_x n_y \) diagonal matrix was obtained whose diagonal elements are \( \psi'_{i,j} \)’s. The regularization operator is represented using the expression in equation (19).

\[
L(f) = D_x^T \text{diag} \left( \psi'(f) \right) D_x + D_y^T \text{diag} \left( \psi'(f) \right) D_y \\
= \begin{bmatrix} D_x^T D_x & 0 \\ 0 & D_y^T D_y \end{bmatrix} \begin{bmatrix} \text{diag} \left( \psi'(f) \right) \\ \text{diag} \left( \psi'(f) \right) \end{bmatrix} 
\]

(19)

Equation (19) represents the expression of the regularization operator. After the computation of the regularization operator, the gradient of the non-linear expression in (16) is computed. An approximate Hessian matrix is framed and using the quasi-Newton upgrading step is performed and the initial guess of the image is upgraded. These steps are repeated in an iterative process until the fixed-point iteration converges. The convergence of the fixed-point iteration method depends upon the norm of the difference.
between the true image and the restored image. As the number of iteration increases, the norm of the difference between different numbers of iteration approaches zero.

The fixed point iteration method starts with initial guess of the image. In this case, the initial guess was the blank image. The number of iteration and selection of the regulation parameter was arbitrarily to start with. The restoration was initially performed by setting the regularization parameter ($\lambda$) equal to 1e-02 and number of iteration equal to 100. The result of the restoration is shown in the figure 6.

![Figure 6: Restored Image using fixed point iteration method. (No: of iteration=100 and $\lambda=1e-02$)](image)

The selection of optimal value of regularization parameter ($\lambda$) was necessary. For a fixed number of iteration (say, 100) the restoration is performed for a different values of $\lambda$. A norm of the difference between the true image and the restored image at each number of iteration is computed and is plotted for different values of the $\lambda$. A curve is plotted between the number of iteration and norm of the difference. The curve is shown in figure 7(a). From the curve, the optimal value of the regularization parameter ($\lambda$) is selected and it is 1.5e-05. The norm of the difference decreases constantly to a value where it starts to increase again. The point of minimum value is selected as the optimal value of $\lambda$. 
The figure 7(b) is plotted within a scale limit of 1e-06 and 4e-05 in order to show the minimum value clearly.

The fixed point restoration is performed for 512 number of iteration for an optimal value of the $\lambda$ selected from the figure 7. The optimization of the number of iteration is also important. To evaluate the convergence of the fixed point iteration method, graph is plotted between the norm of the difference between the true image and the restored image and the regularization parameter $\lambda$. The curve shows the selection of the optimal value of regularization parameter ($\lambda$).
image, and number of iteration for an optimal value of $\lambda$. The result of the convergence is shown in the figure 8. From the figure 8, it can be concluded that the iteration process has converged because the difference between the norm of the difference between the true image and the restored image for different number iterations is approximately zero. After optimal selection of $\lambda$ and number of iterations, the restoration of an image from an initial guess of blank image is performed for optimal values. The figure 9 shows the restored image after 512 numbers of iterations and $\lambda$ value of $1e-05$.

**Figure 8:** Convergence of the fixed point iteration method

**Figure 9:** Restored image (No: of iteration =512, $\lambda = 1e-05$)
The same is repeated with different images for a same 512 number of iteration and the $\lambda$ value of 1.5e-05. The figure 10 and 12 show the different images. The figure 11(a) shows the restored images obtained after 512 numbers of iterations and with the $\lambda$ value equal to 1.5e-05. The figure 11(b) shows the convergences of the fixed point iteration.

![Original Image](image1)

**Figure 10: Original Image**

The figure 12(a) shows the restored image obtained for same numbers of iteration and $\lambda$ value. The figure 12(b) shows the convergence of the fixed point iteration method.

![Restored Image: Total Variation process](image2) ![Difference between the True image and the restored image](image3)

**Figure 11: (a) Restored Image (b) Rate of convergence**
Figure 12: Original Image

The results shown above are obtained using a total variation regularization technique, in particular fixed point iteration method.

4. Conclusions

The results suggested that total variation regularization technique is an edge preserving method of regularization. The restoration does not produce any smoothing effect. Inverse engineering problems are so ill-posed that regularization techniques are important. Moreover in the case of non quadratic functionality such as total variation the iteration method is the best method for restoration. If the image to be restored is piece wise continues then total variation just measures the sum of the magnitudes of the jump.
in the image. Therefore total variation produces good results when the function restored is “blocky”. The rate of convergence of the iterative method can be tested from the numerical optimization. The numerical analysis indicates that the iterative method converges rapidly provided the regularization parameter is small and also the non-smoothing $\beta$ value is also not too small. This is shown in figure 13. The figure 13(a) shows the convergence rate for the $\lambda=1e-02$ and $\beta=0.1$ and figure 13(b) shows the rate of convergence for the $\lambda=1.5e-05$ and $\beta=0.1$. The former takes more number of iteration to converge when compared to the latter. The fixed point iteration method converges globally in a non-quadratic fashion.

Figure 13: (a) Rate of convergence ($\lambda=1e-02$) (b) Rate of convergence ($\lambda=1.5e-05$)
References


[12] Peter Blomgren and Tony F. Chan, "Color TV: Total Variation methods for restoration of vector valued images".


[29] David M. Strong, Peter Blomgren and Tony F. Chan, “Spatially Adaptive Local Feature Driven-Total Variation Minimizing image restoration,” *Department of Mathematics*, University of California, LA


Appendix

Source Code for Total Variation Regularization:

clear all
clc
%--------------------------------------------------------------------------
%                               Total Variation Regularization Method
%--------------------------------------------------------------------------
I = imread('image 1res.pgm');
figure,imagesc(I);
title('Original Image');
xlabel('No:of pixels');
ylabel('No:of pixels');

N = 32;                         % Image Size
I_res = reshape(I,N^2,1);      % Image Vector

% Computing the PSF
b = 10;                        % Band
s = 7;                         % Sigma value

Z = [exp(-((0:b-1).^2)/(2*s^2)),zeros(1,N-b)];
A = toeplitz(Z);
A = (1/(2*pi*s^2))*kron(A,A);  % Point Spread Function

% Addition of Gaussian Noise

noise = imnoise(I,'gaussian',0,0.005);
figure,imagesc(noise);
title('Noise added to the image');
xlabel('No:of pixels');
ylabel('No:of pixels');
% Creating Noise Image.
x = zeros(N,N);
noisex = imnoise(x,'gaussian',0,0.005);
figure,imagesc(x);
figure,imagesc(noisex);
title('Noise');

% Lexicographic Arrangement of Noisy image

Nr = reshape(noise,32^2,1);
lg_Nr = double(Nr);

% g is the blurred image
% I_res is the resized image to 32*1
% I_org is the image restored back to original size of 32*32

$g = A \times lg_Nr$;
I_blur = reshape(g,32,32);
figure,imagesc(I_blur);
title('Blurred and Noisy Image');
xlabel('No:of pixels');
ylabel('No:of pixels');

% Computing the Least square method
% using in-built command
$lsinbuilt = lsqr(A,g)$;
imagesc(reshape(lsinbuilt,32,32));
title('Least Square Output using the inbuilt Matlab command');
xlabel('No:of pixels');
ylabel('No:of pixels');
% Least suqares using formulation
inverse = inv(transpose(A)*A)*transpose(A)*g;
lsrest = reshape(inverse,32,32);
figure,imagesc(lsrest);
title('Least Square Output using the formulation');
xlabel('No:of pixels');
ylabel('No:of pixels');

fixed_iter = 512; % fixed point iteration
beta = 0.1; % Smoothing factor
lamda = 1.5e-05; % Regularization parameter

% Non-Linear Equation
% Computation of Regularization operator:

n = 32;
nsq = n^2;
Delta_x = 1 / n;
Delta_y = Delta_x;
D = spdiags([-ones(n,1) ones(n,1)], [0 1], n,n) / Delta_x;
I_trunc1 = spdiags(ones(n,1), 0, n,n);
Dx1 = kron(D,I_trunc1); % Forward (upwind) differencing in x
Dy1 = kron(I_trunc1,D); % Forward (upwind) differencing in y
f_fp = zeros(n,n); % Initial Guess of Blank Image
fvec = f_fp(:); % Image Vector
i=1;

for f_iter = 1:fixed_iter
    tic
    t=((Dx1*fvec).^2 + (Dy1*fvec).^2);
psi_prime1 = 1./sqrt(t + (0.1)^2);
Dpsi_prime1 = spdiags(psi_prime1, 0, (n)^2,(n)^2);
L1 = Dx1' * Dpsi_prime1 * Dx1 + Dy1' * Dpsi_prime1 * Dy1;
L = L1 * Delta_x * Delta_y;
pen_fun = lamda * L * fvec; % Computing alpha*operator*initial image

% computing the gradient of the non-Linear equation

Gv = A' * (A * fvec - g) + pen_fun;
Gv_nm = norm(Gv); % Compute the norm of the fixed point gradient
H_residual = lamda * L; % Computation alpha * operator
H = (A' * A) + H_residual; % Computes the Hessian Matrix
S = -(inv(H)) * Gv; % Compute the quasi-newton
fvec = fvec + S; % Upgrading the initial image
df = double(I_res) - double(fvec);
cng(f_iter) = norm(df);
toc
time(f_iter) = toc;
end

semilogy([1:fixed_iter], Gv_nm,'r-');
title('Gradient plot of the fixed point iteration-Convergence');
xlabel('No:of iteration');
ylabel('||gradient||');
grid on

semilogy([1:fixed_iter],cng,'r-');
title('Difference between the True image and the restored image(Lamda=1.5e-05)');
xlabel('No:of iteration');
ylabel('|| True Image - Restored Image ||');
grid on

\[ y = [2.5882, 2.5837, 2.5825, 2.5816, 2.5808, 2.5801, 2.5795, 2.5789, 2.5785, \\
    2.5780, 2.5778, 2.5775, 2.5772, 2.5771, 2.5769, 2.5768, 2.5767, \\
    2.5766, 2.5766, 2.5762, 2.5773, 2.5794, 2.5824, 2.5857, 2.5923, 2.5991, \\
    2.6054, 2.6116, 2.6180, 2.6242]; \]

    5*10^{-6}, 5.5*10^{-6}, 6*10^{-6}, 6.5*10^{-6}, 7*10^{-6}, 7.5*10^{-6}, 8*10^{-6}, \\
    8.5*10^{-6}, 9*10^{-6}, 9.5*10^{-6}, 1*10^{-5}, 1.5*10^{-5}, 2*10^{-5}, 2.5*10^{-5}, \\
    3.0*10^{-5}, 3.5*10^{-5}, 4.5*10^{-5}, 5.5*10^{-5}, 6.5*10^{-5}, 7.5*10^{-5}, \\
    8.5*10^{-5}, 9.5*10^{-5}]; \]

figure
plot(x,y)
xlim([1e-06 4e-05])
xlabel('Regularization Parameter (Lamda)');
ylabel('|| True Image - Restored Image ||');
title('Selection of Optimal Lamda');
grid on

fvec_res = reshape (fvec, 32, 32);
figure, imagesc(fvec_res);
title ('Restored Image: Total Variation process');
xlabel ('No: of pixels');
ylabel ('No: of pixels');