AN L-CURVE APPROACH TO OPTIMAL DETERMINATION OF REGULARIZATION PARAMETER IN IMAGE RESTORATION

C. M. Leung and W.-S. Lu

Department of Electrical and Computer Engineering
University of Victoria, Victoria, B.C., V8W 3P6

ABSTRACT

Image restoration refers to the problem of removal or reduction of degradation in noisy blurred images. The image degradation is usually modeled by a linear blur and an additive white noise process, and an image restoration problem can then be considered as an integral equation of the first kind. In many practical image restoration problems, the linear blur involved are always ill-conditioned. This provides a typical example for ill-posed problems for which the solutions are unstable. The method of regularization provides stable solutions to image restoration problems with a tradeoff between accuracy and smoothness of the solutions. The tradeoff is determined by a regularization parameter. In this paper, an L-curve approach to determining this tradeoff is proposed. It is demonstrated that a regularization parameter corresponding to the largest curvature of the L-curve gives a nearly optimal regularized solution of a given image restoration problem.

1. Introduction

Consider a typical image model

$$g = Hf + n \tag{1}$$

where g and f represent the lexicographically ordered recorded and original images, respectively; H is the spatially invariant linear operator that characterizes the image degradation mechanism; and n is the additive white noise. In the image restoration problem, it is usually assumed that the noiseless counterpart of (1)

$$g_I = Hf \tag{2}$$

has a unique solution. Even with this assumption, one finds problem (2) ill-posed as matrix H is usually quite ill-conditioned, meaning that small error in g_I will cause large error in the solution. In restoring the image, it is wished to approximate the solution f to equation (2) using the recorded image g with a known noise level

$$||g - g_I|| = ||g - Hf|| = ||n|| \leq \delta \tag{3}$$

where $||\cdot||$ denotes the 2-norm and δ may be set to be $||n||$ which can be estimated easily in a smooth region of the given image g. The primary interest in the study of image restoration problem (1) is to develop methods to find a smooth and reasonably accurate approximation of f for a given degraded image g. These methods are usually referred to as regularization methods as a regularization parameter λ is always used there to control the degree of solution smoothness. A fixed λ corresponds to a matrix R_λ such that $f_\lambda = R_\lambda g$ is an approximate solution of (1). As λ varies, the family of R_λ is called a regularization scheme of (2). More specifically, a regularization scheme of (2) refers to a family of matrices R_λ, such that

$$\lim_{\lambda \to 0} R_\lambda H f = f \tag{4}$$

pointwisely. One of the well known regularization methods is the Tikhonov regularization scheme where R_λ is given by

$$R_\lambda = (H^T H + \lambda C^T C)^{-1} H^T \tag{5}$$

and $f_\lambda = R_\lambda g$ is the solution of the least squares problem

$$\min_{\hat{f}} \{||H\hat{f} - g||^2 + \lambda ||C\hat{f}||^2\} \tag{6}$$

with C a stabilizing operator [1]-[3]. The term $||Hf_\lambda - g|| = ||HR_\lambda g - g||$ represents the approximation error between R_λ and H^{-1} which can serve as a measure of accuracy of the solution f_λ, and the term $||Cf_\lambda||$ is a measure of the smoothness of the solution f_λ. The operator C is called a stabilizing operator and it should act like a highpass filter. It is a common practice that C is always chosen to be the Laplace operator or similar type in image restoration problems. Each regularization scheme requires a strategy of choosing a λ such that (3) is satisfied and both quantities $||Hf_\lambda - g||$ as well as $||Cf_\lambda||$ are small. To achieve a small approximation error $||Hf_\lambda - g||$, λ has to be small. However to obtain an f_λ with small $||Cf_\lambda||$ requires a large λ.

Concerning the choice of the regularization parameter λ, several methods have been available, see for example reference [4] and the references cited there. The aim of this paper is to propose an L-curve approach for the determination of λ in order to keep both quantities $||Hf-g||^2$ and $||Cf||^2$ as small as possible. The L-curve approach was first proposed in [5] to solve a class of least-squares problems and was further explored recently in [6] to analyse discrete ill-posed problems.

2. More About Tikhonov Regularization

In some cases, an *a priori* bound for $||Cf||$ might be given

$$||Cf|| \leq E \tag{7}$$

To solve the image restoration problem (1), the following two methods might be considered.

$$\min_f \quad ||Cf|| \tag{8a}$$

$$\text{subject to} \quad ||Hf-g|| \leq ||n|| \tag{8b}$$

and

$$\min_f \quad ||Hf-g|| \tag{9a}$$

$$\text{subject to} \quad ||Cf|| \leq E \tag{9b}$$

In the Tikhonov regularization a "solution" is sought to minimize

$$||Hf-g||^2 + \frac{||n||^2}{E^2}||Cf||^2 \tag{10}$$

Since $||Cf|| \leq E$ and $||Hf-g|| = ||n||$, we obtain

$$||Hf-g||^2 + \frac{||n||^2}{E^2}||Cf||^2 \leq ||n||^2 + ||n||^2 = 2||n||^2$$

This implies that at most a factor of $\sqrt{2}$ is lost in an approximate solution of (1) if the method (10) is used instead of method (8) or method (9). As E is unknown in most cases, a regularized solution of (1) is formulated as the solution of the optimization problem (6), and the solution is

$$f_\lambda = (H^T H + \lambda C^T C)^{-1} H^T g \tag{11}$$

The matrices H and C in its original form are block Toeplitz matrices and can be replaced with block circulant approximations [7]. An important property of block circulant matrices is that all block circulant matrices (with same size) have the same set of eigenvectors which leads to efficient computation of f_λ in (11) using discrete Fourier transform (DFT). To be specific, there exists an unitary matrix U such that any block circulant matrix A can be expressed as

$$A = U\Lambda U^* \tag{12}$$

where Λ is a diagonal matrix determined by the eigenvalues of A and $*$ denotes the operation of complex conjugate transposition. It can be shown that this U acts like a DFT and consequently solution (11) can be obtained using the fast DFT.

3. Determination of the Regularization Parameter via L-Curve Approach

It follows from (13) that

$$H = U\Lambda_H U^*, \qquad C = U\Lambda_C U^* \tag{13}$$

where $\Lambda_H = \text{diag}\{h_i\}$ and $\Lambda_C = \text{diag}\{c_i\}$ are the diagonal matrices with eigenvalues of H and C placed along the diagonal, respectively. Hence solution (11) can be expressed as

$$f_\lambda = U(\Lambda_H^* \Lambda_H + \lambda\Lambda_C^*\Lambda_C)^{-1}\Lambda_H^* U^* g \tag{14}$$

By using (14), it is found that

$$||Cf_\lambda||^2 = \sum_i |\frac{h_i^* c_i}{|h_i|^2 + \lambda|c_i|^2}(u_i^* g)|^2 \tag{15}$$

$$||Hf_\lambda - g||^2 = \sum_i |\frac{\lambda c_i^2}{|h_i|^2 + \lambda|c_i|^2}(u_i^* g)|^2 \tag{16}$$

where u_i is the ith column of U. Eqns. (15) and (16) imply that

$$\frac{d\,||Cf_\lambda||^2}{d\,\lambda} = -2\sum_i \frac{|h_i^*|^2|c_i|^4}{(|h_i|^2 + \lambda|c_i|^2)^3}|u_i^* g|^2 \tag{17}$$

and

$$\frac{d\,||Hf_\lambda - g||^2}{d\,\lambda} = 2\sum_i \frac{\lambda|h_i|^2|c_i|^4}{(|h_i|^2 + \lambda|c_i|^2)^3}|u_i^* g|^2 \tag{18}$$

Hence $||Cf_\lambda||^2$ monotonically decreases with respect to (w. r. t.) λ while $||Hf_\lambda - g||^2$ monotonically increases w. r. t. λ. As a result, if we view λ as a parameter varying from 0 to ∞, the mapping $||Hf_\lambda - g||^2 \longrightarrow ||Cf_\lambda||^2$ defines a curve in the $(||Hf_\lambda - g||^2, ||Cf_\lambda||^2)$ plane. We call this curve the L-curve associated with the optimization problem (6).

Further notice from eqns. (15)-(18) that for small values of λ, both $||Hf_\lambda - g||^2$ and its derivative are small while both $||Cf_\lambda||^2$ and its derivative are relatively large, this implies that

$$\frac{d\,||Cf_\lambda||^2}{d\,||Hf_\lambda - g||^2} = \frac{d\,||Cf_\lambda||^2/d\,\lambda}{d\,||Hf_\lambda - g||^2/d\,\lambda} \tag{19}$$

is fairly negative on an interval $0 \leq ||Hf_\lambda - g||^2 \leq \delta_1$ where δ_1 is a certain small number, and hence the L-curve has a sharp decline there. In other words, the left portion of the L-curve is nearly a vertical line and hence its curvature is small. Obviously, any point on this portion of the L-curve corresponds to a f_λ with good solution accuracy (i.e. small $||Hf_\lambda - g||^2$) but poor smoothness (i.e. large $||Cf_\lambda||^2$). Furthermore, for the large values of λ, eqns. (15)-(18) indicate that both $||Cf_\lambda||^2$ and its derivative are small while $||Hf_\lambda - g||^2$ tends to be independent of λ. Eqn. (19) now implies that the right portion of L-curve corresponding to large λ (and thus large $||Hf_\lambda - g||^2$) is nearly a horizontal line and hence its curvature is also small. Obviously, any point on this portion of the L-curve corresponds to a smooth f_λ with poor solution accuracy (since $||Hf_\lambda - g||^2$ is large). Moreover, from (15)-(18) it can readily be shown that the second-order derivative of $||Cf_\lambda||^2$ w. r. t. $||Hf_\lambda - g||^2$ is always positive meaning that the L-curve is globally convex. Based on the observations made above, we conclude that the shape of the L-curve will look like a letter L as shown in Fig. 1 with the optimal λ corresponding to the corner point c on the curve. Since the maximum curvature will occur at the corner c, we propose an easy-to-implement approach to find the numerical value of this λ as follows. First, a number of points $((||Hf_\lambda - g||^2, ||Cf_\lambda||^2)$ with λ varying over a wide range are calculated. Then a certain interpolation method such as the least squares method is used to obtain a rational function that approximates the L-curve. Next the curvature of this rational function is computed to determine the best value of λ that corresponds to the L-curve's maximum curvature.

To demonstrate the proposed approach, we apply it to two sampled images which are degraded versions of image Lena by a linear motion blur and a defocusing, respectively. Fig. 2 is the original image of Lena, which was linearly blurred with 14 blurring distance units and was contaminated by pseudo-white Gaussian noise with SNR = 30 dB, and the blurred image is shown in Fig. 3(a). The corresponding L-curve obtained using an interpolation with 24 $(||Hf_\lambda - g||^2, ||Cf_\lambda||^2)$ points is depicted in Fig. 1. It is found that $\lambda = 0.01$ corresponds to the maximum value of the curvature. Using this value of λ in (11), the restored Lena is shown in Fig. 3(b) and it has a fairly high improvement signal-to-noise ratio (ISNR) value (ISNR = 10.5). Here the ISNR is defined [3] as

$$\mathrm{ISNR} = 10\log\frac{\sum_k\sum_l(g(k,l) - f(k,l))^2}{\sum_k\sum_l(\hat{f}(k,l) - f(k,l))^2}$$

with f, g and $\hat{f}$ the original, the noisy and blurred, and the restored images, respectively.

Fig. 4(a) shows a defocused Lena with radius of the circle of confusion $r = 7$ which was contaminated by pseudo white Gaussian noise with SNR = 28 dB. The proposed L-curve approach gives $\lambda = 0.005$. Using this value of λ in (11), the restored Lena is shown in Fig. 4(b) and it has a improvement signal-to-noise ratio (ISNR) value (ISNR = 10.95).

4. An Adaptive Restoration Method for Linear Motion Blurred Images

In Tikhonov regularization, the λ is chosen as a compromise between accuracy and smoothness and the same λ is used for the whole input image. The value of λ so chosen tends to be larger than necessary for background regions. As a result, isolated edges lying in a uniform zone of the image may be smoothed out. A possible remedy for this problem is to localize the restoration problem so that λ becomes region-dependent local parameter. For example, if the linear blur is of one-dimensional (1-D), say in the vertical direction, then the 2-D image model (1) can be reduced to

$$g_j = Hf_j + n_j \tag{20}$$

where g_j and f_j represent the jth column of the recorded and the original images, respectively; H is a circulant (not block circulant) matrix corresponding to the 1-D invariant linear blur, assumed to be known; and n_j is the additive white noise. For a given λ_j, the Tikhonov solution of (20), i.e.

$$f_{\lambda_j} = (H^TH + \lambda_jC^TC)^{-1}H^Tg_j \tag{21}$$

can be obtained efficiently using 1-D DFT. For different columns, the optimal λ_j's chosen by the proposed L-curve approach may be different, e.g., for the column with high signal-to-noise ratio (SNR), the corresponding λ_j should be small and for the column with small SNR, the corresponding λ_j should be large. Therefore the restored image can be recovered column by column adaptively with respect to λ_j. It is also noted that the model (20) are computationally more efficient in implementation than its 2-D counterpart (1).

Fig. 5(a) shows the image Tiffany with a synthetic horizontal line segment on the top left corner. It is used to examine the performance of the proposed 1-D adaptive restoration algorithm in which the L-curve approach described in Sec. 3 is incoporated. The Tiffany was linearly blurred in the horizontal direction with 9 shifting units and was further contaminated by white noise with SNR = 40 dB. The resulted noisy blurred Tiffany is shown in Fig. 5(b). For each column of g, a λ_j is determined using the L-curve approach. It is found that $\lambda_j = 0.01$ for $j = 1, \cdots, 30$ and the rest

λ_j's vary over the range [0.2 , 0.4]. Four λ_j's values, namely 0.1, 0.25, 0.3 and 0.35 are employed in the 1-D adaptive Tikhonov regularization restoration. The adaptive threshold is set in the following fashion. Let SNR_W be the SNR of the whole image g and let SNR_j be the SNR of the jth column of g. The λ_j in the jth column restoration is chosen as follows

$$\lambda_j = \begin{cases} 0.01 & j = 1, \cdots, 30 \\ 0.25 & \mathrm{SNR}_j < (1-k)\mathrm{SNR}_W \\ 0.3 & (1-k)\mathrm{SNR}_W < \mathrm{SNR}_j < (1+k)\mathrm{SNR}_W \\ 0.35 & \mathrm{SNR}_j > (1+k)\mathrm{SNR}_W \end{cases}$$

where $k = 0.1$ representing 10% standard deviation of SNR_W. The restored image is shown in Fig. 5(c). For comparison purpose, the 1-D conventional Tikhonov regularization with $\lambda = 0.3$ is also applied to the same image and the restored image is shown in Fig. 5(d). It is clear from Fig. 5(c) and (d) the synthetic line segment is better restored by the adaptive restoration method.

Acknowledgement

This work was supported by Micronet and IRIS.

References

[1] A. N. Tikhonov, "Solution of Incorrectly Formulated Problems and the Regularization Method", Doklady Akad. Nauk SSSR, 151, 1963.

[2] A. N. Tikhonov, "Solution of Nonlinear Integral Equations of the First Kind", Doklady Akad. Nauk SSSR, 156, 1 , 1964.

[3] R. L. Lagendijk and J. Biemond, *Iterative Identification and Restoration of Images*, Norwell, MA: Kluwer Academic Publishers, 1991.

[4] A. M. Thompson, J. C. Brown, J W. Kay and D. M. Titterington, "A study of methods of choosing the smoothing parameter in image restoration by regularization," *IEEE Tran. PAMI*, vol. 13, No. 4, pp. 326-339, April 1991.

[5] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*, Englewood Cliffs, NJ: Prentice-Hall, 1974.

[6] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," *SIAM Review*, vol. 34, No. 4, pp. 561-580, Dec. 1992.

[7] B.R. Hunt, "The Application of Constrained Least Squares Estimation to Image Restoration by Digital Computer," *IEEE Trans. Computers*, vol. C-22, pp. 805-812, Sept. 1973.

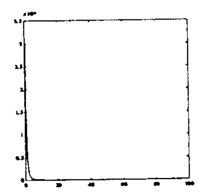


Fig. 1

Fig. 2

Fig. 3(a)

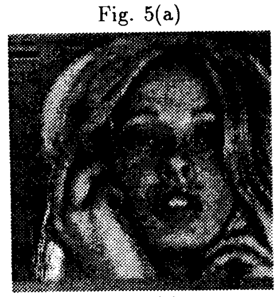
(b)

Fig. 4(a)

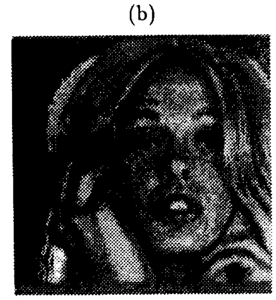
(b)

Fig. 5(a)

(b)



(c)



(d)