Objective Selection of Hyperparameter for EIT

Brad Graham, Andy Adler
School of Information Technology and Engineering (SITE)
University of Ottawa, Canada

June 7, 2005

Introduction

Electrical impedance tomography (EIT) uses body surface electrodes to make measurements from which an image of the conductivity distribution is calculated. Reconstructions are challenging due to the ill-conditioning of the system matrix and susceptibility of the data to corruption from noise. Regularization methods are used to calculate stable reconstructions by imposing additional conditions, such as image smoothness, on a solution. A difficulty with experimental and clinical EIT reconstruction algorithms is the requirement to select a scalar hyperparameter, λ, to control the amount of regularization used to achieve a good reconstruction.

In the broader field of inverse problems work on automatic hyperparameter selection has produced methods such as Generalized Cross Validation (GCV), L-Curve, and the Discrepancy Principle (Hansen 1992). However within the field of EIT by far the most common method of hyperparameter selection is heuristic selection; researchers inspect reconstructions for a range of hyperparameter values and select one.

The absence of objective hyperparameter selection methods results in several issues: 1) users of clinical EIT systems would be uncomfortable using "fiddle" adjustments to modify images, 2) comparisons of EIT reconstruction algorithms can be subjective due to the necessity of manual tuning of hyperparameter values, and 3) meta-algorithms, such as detection of electrode errors (Asfaw and Adler 2005), require a method to fix these values. In this paper we develop and evaluate some objective hyperparameter selection techniques.

Methods

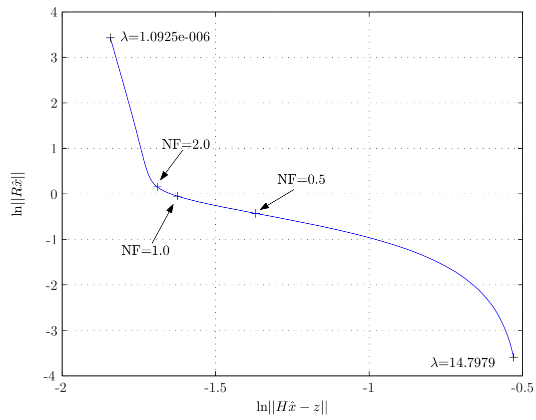
We consider the class of one-step linearized reconstruction algorithms that calculate the proportional change in a finite element conductivity distribution, $\mathbf{x}$, from a proportional change in boundary voltage difference signal, $\mathbf{z}$. For small changes around a background conductivity the relationship between $\mathbf{x}$ and $\mathbf{z}$ may be linearized as

$$\mathbf{z} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

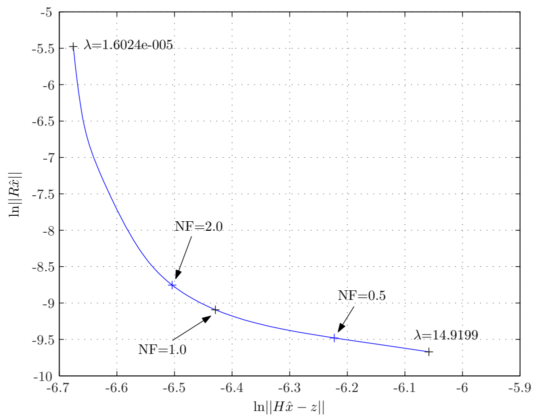
where $\mathbf{H}$ is the Jacobian or sensitivity matrix and $\mathbf{n}$ is the measurement system noise, assumed to be uncorrelated and Gaussian. In order to overcome the ill-conditioning of $\mathbf{H}$ we solve (1) using the following regularized inverse

$$\hat{\mathbf{x}} = (\mathbf{H}^{\mathbf{T}}\mathbf{H} + \lambda\mathbf{R})^{-\mathbf{1}}\mathbf{H}^{\mathbf{T}}\mathbf{z} = \mathbf{B}\mathbf{z} \tag{2}$$

where $\hat{\mathbf{x}}$ is an estimate of the true conducitivity distribution, $\mathbf{R}$ is a regularization matrix and λ is a scalar hyperparameter that controls the amount of regularization. With $\mathbf{R} = \mathbf{I}$ (labelled $\mathbf{R_{Tik}}$) equation (2) is the 0^{th} order Tikhonov algorithm. With $\mathbf{R} = \mathbf{diag}(\mathbf{H})$ (labelled $\mathbf{R_{diag}}$) equation (2) is the regularization matrix used in the NOSER algorithm of Cheney et al (1991). With the *maximum a posteriori* (MAP) framework of Adler and Guardo (1996) $\mathbf{R}$ is modeled as a Gaussian spatial high pass filter (labelled $\mathbf{R}_{Gaus}$) with a cut off frequency of 10%. While several other one-step regularized inverse



(a) L-Curve for $\mathbf{R_{Tik}}$



(b) L-Curve that does not work, $\mathbf{R_{diag}}$, 576 element mesh

Figure 1: Example L-Curves

algorithms exist for EIT, in this paper we consider equation (2) with these three regularization matrices as a representative sample with which to compare their effect.

The rest of this paper is organized as follows. Three strategies to objectively select a hyperparameter are developed: 1) L-Curve, 2) Resolution-Curve, 3) Fixed Noise Figure (NF) Method. We then evaluate the effectiveness of each strategy for linearized one-step EIT with respect to heuristic selection.

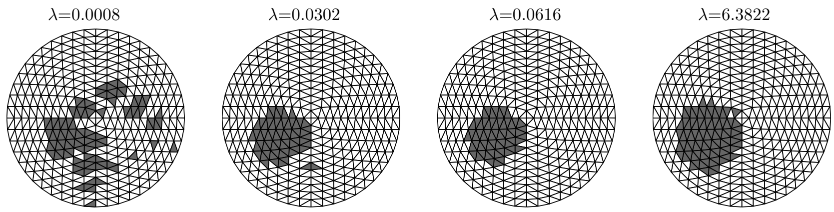
Hyperparameter Selection Methods

A well known method of hyperparameter selection is the L-Curve method (Hansen 1992). This method plots the semi norm of the regularized solution, $\ln \|\mathbf{R}\hat{\mathbf{x}}\|$, versus the norm of the corresponding residual vector, $\ln \|\mathbf{H}\hat{\mathbf{x}} - \mathbf{z}\|$ parametrically over λ. A resulting plot, such as figure 1(a), will often have an "L" shape where the optimal value for λ is located at the point of maximum curvature.

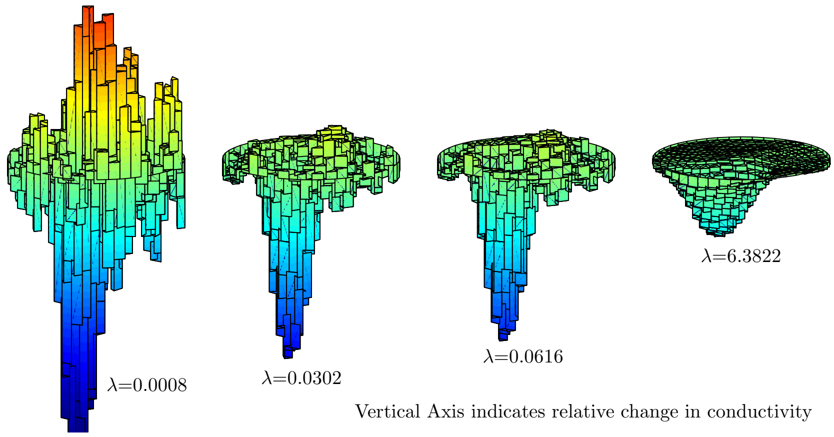
Objective hyperparameter selection can also be made using a resolution maximization rationale. We define Blur Radius (BR) as a measure of the resolution: $BR = r_z/r_0 = \sqrt{A_z/A_0}$ where r_0 and A_0 are the radius and area respectively of the entire 2D medium and r_z and A_z are the radius and area of the reconstructed contrast containing half the magnitude of the reconstructed image (Adler and Guardo 1996). BR calculates the area fraction of the elements that contain 50% of the total image amplitude. We call this the half amplitude (HA) set. Figure 2(a) shows the evolution of the HA set in response to increasing λ for an impulse contrast. With insufficient λ the image is dominated by noise and the HA set is composed of spatially disjoint elements. As λ is increased, noise is filtered through the smoothing action of the prior, image energy starts to concentrate, and the HA set starts to cluster. Excessive regularization blurs the image and expands the HA set. A plot of BR vs λ such as figure 3(a) shows a rapid decrease in BR to a minimum value that occurs after the HA has become contiguous, followed by a slower increase in BR as filtering starts to blur the image. We refer to the point marked OS as the "onset of stability."

We define image SNR as $SNR = E[\hat{\mathbf{x}}]/\sqrt{\mathrm{var}[\hat{\mathbf{x}}]}$. A plot of image SNR vs λ for an impulse contrast shows that SNR increases rapidly and then stabilizes, after which increases in λ produce smaller increases in SNR. The point of stability occurs just after the OS point of the Resolution-Curve.

The Fixed NF Method is based on a Noise Figure (NF) calculation introduced by Adler and Guardo (1996), where NF is a function of λ and is defined as the ratio of signal-to-noise-ratio in the measurements to signal-to-noise-ratio in the image: $NF = \frac{SNR_{in}}{SNR_{out}} = \left(\frac{mean[\mathbf{z_c}]}{\sqrt{\mathrm{var}[\mathbf{n}]}}\right) \Big/ \left(\frac{mean[\mathbf{Bz}_c]}{\sqrt{\mathrm{var}[\mathbf{Bn}]}}\right)$. The signal used in this definition is $\mathbf{z_c} = \mathbf{Hx_c}$, where $\mathbf{x_c}$ is a small contrast in the centre of the medium. The user selects a NF value and the corresponding λ is found using a bisection search technique. The calculation of NF does not require knowledge of the noise in the measured data, it is strictly a function of the FEM geometry, the regularization matrix, and λ. The Fixed NF Method substitutes the manual selection of



(a) Evolution of HA set (dark triangles) with increasing λ



(b) Evolution of proportional conductivity change image with increasing λ

Figure 2: Sample Reconstructions of impulse phantom, tank data with $\mathbf{R_{Gauss}}$ on 576 element mesh

λ with the manual selection of a NF which the algorithm then maps to λ. The advantage is that the suitable NF range is not algorithm dependent, while λ is. Experience has shown that noise figures in the range 0.5 to 2 consistently lead to good reconstructions.

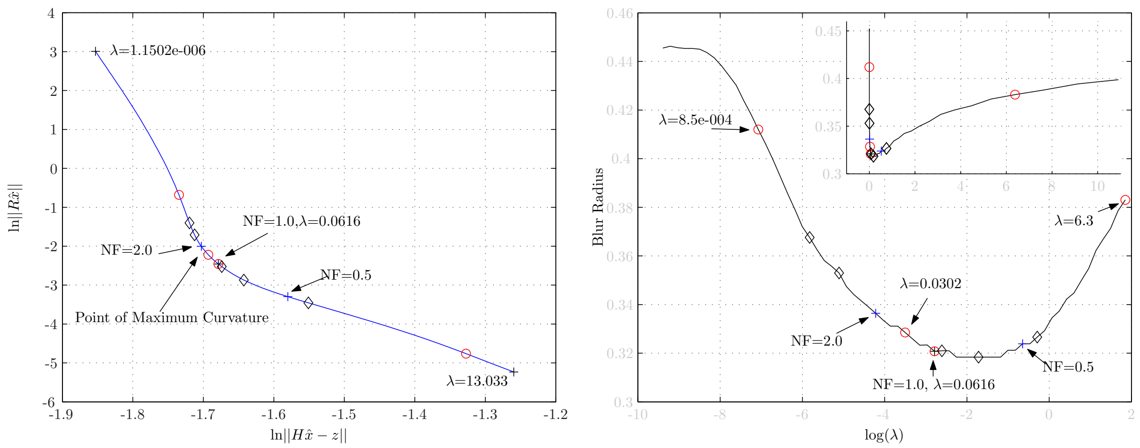
Three sources of test data were used to compare the hyperparameter selection methods: 1) simulated data, generated with a 1968 element mesh with an impulse contrast located halfway between the centre and boundary, 2) simulated data obtained by adding .05% AWGN noise to set #1, and 3) lab data using a 2 cm diameter impulse phantom located halfway along the radius of a 30cm diameter tank containing a saline solution. The data was used to reconstruct images using 18 configurations (6 meshes and 3 regularization matrices). The 6 meshes have 64, 256, 492, 576, 1024, and 1968 elements.

Results

We compare the techniques of L-Curve, Resolution-Curve, Fixed NF Method (calculated using the algorithm of Adler and Guardo (1996)), and heuristic selection using the plots of figure 3. Each plot indicates λ values for NF of $^1/_2$, 1, and 2 as well as λ values that match the reconstructions of figure 2 (indicated by circular markers). For Tikhonov reconstructions, the L-Curve and Resolution-Curves exhibited more complex behaviour however, the OS and BR minimum point were always evident.

Heuristic selection was performed by 5 graduate students studying image processing. Heuristic selections varied and were not confined to the minimum region of the Resolution-Curve or knee of the L-Curve: no clear preference was shown among images reconstructed using λ from the minimum region of the Resolution-Curve. This suggests that there is no single optimal value of λ, rather there is an optimal region of λ, over which reconstructions are qualitatively indistinguishable. Heuristic selections are indicated by diamond markers on the plots of figure 3.

In all 18 configurations the Resolution-Curve indicated a hyperparameter corresponding to a minimum BR that resulted in a "good" reconstruction. Moreover, the Fixed NF method with NF=1 calculated a hyperparameter that was located in this minimal region of the Resolution-Curve.



(a) L-Curve

(b) Blur Radius vs log(λ).Inner plot is over larger range

Figure 3: (a) L-Curve, (b) Resolution-Curve. Diamonds indicate heuristic selections.

In the six Tikhonov ($\mathbf{R_{Tik}}$) configurations the L-Curve always indicated a clear point of maximum curvature. However there were some $\mathbf{R_{diag}}$ and $\mathbf{R_{Gaus}}$ configurations where the L-Curve did not show a marked and stable "knee"region from which one could select a hyperparameter. Figure 1(b) is such an example. In all cases where the L-Curve was able to indicate a value, the hyperparameter was less than the value selected through inspection of the Resolution-Curve. As a result L-Curve derived images were noisier than Resolution-Curve images. In several instances the point of maximum curvature of the L-Curve occurred prior to the OS indicated by a contiguous HA set.

Dicussion and Conclusion

This paper develops and compares 3 methods to objectively select a hyperparameter for use in one-step image reconstructions and compares them to heuristic selections.

We present four observations: 1) hyperparameters taken from the minimum region of the Resolution-Curve always produce good solutions. 2) with NF=1 the Fixed NF Method calculates a hyperparameter that falls in the minimum region of the Resolution curve. 3) the Fixed NF Method provides a configuration independent method to select λ. 4) the L-Curve is not reliable for all configurations (doesn't indicate a hyperparameter), but when it does it provides a lower hyperparameter value than the other objective methods.

For the class of regularized reconstruction algorithms used in this work the Resolution-Curve provides an objective way to select a minimum value for λ. For both lab and simulated data the Fixed NF Method with NF=1 reliably finds λ on or just after the onset of stability. Thus we recommend the Fixed NF Method with NF=1 as a configuration independent method to efficiently find the minimum region of the Resolution-Curve without having to actually construct the computationally expensive Resolution-Curve.

References

Adler A and Guardo R 1996 Electrical impedance tomography: regularised imaging and contrast detection *IEEE Trans. Med. Imaging* **15** 170-9

Asfaw Y and Adler A 2005 Automatic detection of detached and erroneous electrodes in electrical impedance tomography *Physiol. Meas* **26** S175-S183, 2005.

Cheney M, D. Isaacson, J.C. Newell, S. Simske, and J.C. Goble 1990 NOSER: An algorithm for solving the inverse conductivity problem Internat. *J. Imaging Systems & Technol.* **2**, 66-75

Hansen P C 1992 Analysis of discrete ill-posed problems by means of the L-Curve *SIAM Review* Volume 34 Issue 4 December 1992 561-580