Parameter Selection Tool for Solution to Ill-Posed Problems

Brianna R. Cash[1]   Advisor: Dianne O'Leary[2]

[1]Applied Math Scientific Computing (AMSC)
University of Maryland, College Park

[2]Department of Computer Science
University of Maryland, College Park

May 11, 2012

Outline

The problem:

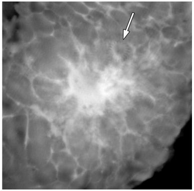

Figure: Tomography image of a mastectomy specimen. Stevens G M et al. Radiology 2003;228:569-575

- In application images can be expensive to produce.
- Used for making decisions
- Images can be distorted and/or noisy.
  - physics of the measurement
  - non-homogeneous material

The discrete model: $\mathbf{Ax} + \epsilon = \mathbf{b}$, $\epsilon \sim \mathbf{N}(\mathbf{0}, \mathbf{S^2})$ where

- $\mathbf{A}$ is a known $m \times n$ matrix where $m \geq n$ (Blurring matrix)
- $\mathbf{x}$ is unknown $n \times 1$ vector (true image) where $n = n_h * n_v$
- ϵ is a $m \times 1$ vector (noise)
- $\mathbf{S^2}$ is known $m \times m$ (variance matrix for ϵ)
- $\mathbf{b}$ is a known $m \times 1$ vector (blurred and noisy image) where $m = m_h * m_v$

Inherent to image deblurring and seismic tomographic problems, $\mathbf{A}$ is ill-conditioned.

Formulation of regularization problem:

$$\min \frac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|_2^2 + \gamma R(x)$$

Where $R(x)$ is a penalty function and γ is the regularization parameter.

Tikhonov: $R(x) = \|\mathbf{x}\|_2^2$.

Total Variation: $R(x) = TV(\mathbf{x})$ where $TV(\mathbf{x}) = \|\nabla \mathbf{x}\|_1$.

Selecting a method and good regularization parameter is problem dependent and subject to bias:

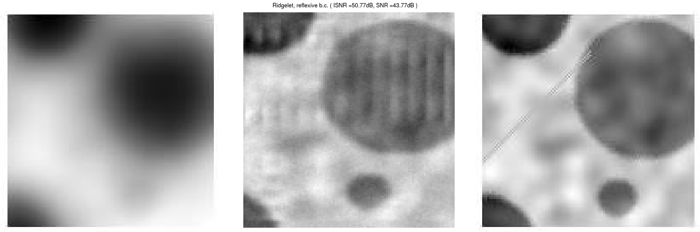

Figure: Images courtesy of Dianne O'Leary

But what is unexpected is often what we are interested in!

Selecting a method and good regularization parameter is problem dependent and subject to bias:

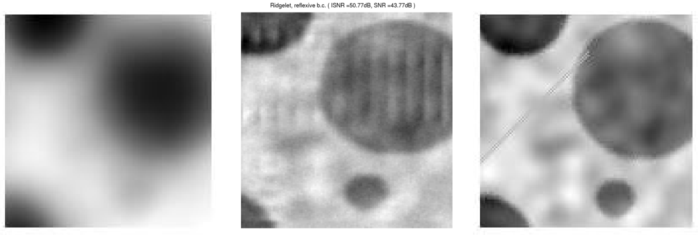

Figure: Images courtesy of Dianne O'Leary

But what is unexpected is often what we are interested in!

Outline

Singular Value Decomposition (SVD) based methods:
SVD of $\mathbf{A}$ is give by

$$\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^{\mathbf{T}} = \sum_{i=1}^{n} \mathbf{u_i}\sigma_i\mathbf{v_i^T}.$$

where $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_n)$ is a $m \times n$ matrix and $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_n)$ is a $n \times n$ matrix both with orthonormal columns, and $\mathbf{S} = diag(\sigma_1, \ldots, \sigma_n)$ is a matrix of the non-negative singular values that appear in decreasing order.

The Tikhonov regularization solution is given by

$$\mathbf{x}_{tik} = \sum_{i=1}^{n} \frac{\sigma_i \mathbf{u_i}\mathbf{v_i^T}\mathbf{b}}{\sigma_i^2 + \gamma}.$$

And Truncated SVD regularization solution is given by

$$\mathbf{x}_{TSVD} = \sum_{i=1}^{n} \phi_i \frac{\mathbf{u_i}\mathbf{v_i^T}\mathbf{b}}{\sigma_i}$$ where $\phi_i = 1$ for $i = 1, \ldots, k$ and $\phi_i = 0$ for $i = k+1, \ldots, n$.

Total variation based regularization method:

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \gamma TV(\mathbf{x}).$$

$TV(\mathbf{x}) = \sum_{i=1}^{n} \|\mathbf{D}_i^T \mathbf{x}\|_2$ where $\mathbf{D}_i^T \mathbf{x} = [\mathbf{x}_{i+n_v} - \mathbf{x}_i, \mathbf{x}_{i+1} - \mathbf{x}_i]^T$.

First order condition:

$$g(\mathbf{x}) = \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) + \gamma \sum_{i=1}^{n} \frac{\mathbf{D}_i \mathbf{D}_i^T \mathbf{x}}{\|\mathbf{D}_i^T \mathbf{x}\|_2} = 0$$

replace $\|\mathbf{D}_i^T \mathbf{x}\|_2$ with $\sqrt{\|\mathbf{D}_i^T \mathbf{x}\|_2^2 + \beta}$ where $\beta > 0$ and small.

Motivation for an improved implementation [Chan1996]:

Consider $g(\mathbf{x})$ (ignoring β for simplicity):

$$g(\mathbf{x}) = \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) + \gamma \sum_{i=1}^{n} \frac{\mathbf{D}_i \mathbf{D}_i^T \mathbf{x}}{\|\mathbf{D}_i^T \mathbf{x}\|_2}$$

Introduce a new variable:

$$\mathbf{y}_i = \frac{\mathbf{D}_i^T \mathbf{x}}{\|\mathbf{D}_i^T \mathbf{x}\|_2} \text{ where } \mathbf{y}_i \text{ is } 2 \times 1.$$

Then the first order condition becomes:

$$g(\mathbf{y}, \mathbf{x}) = \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) + \gamma \sum_{i=1}^{n} \mathbf{D}_i \mathbf{y}_i = 0,$$
$$h(\mathbf{y}, \mathbf{x}) = \|\mathbf{D}_i^T \mathbf{x}\|_2 \mathbf{y}_i - \mathbf{D}_i^T \mathbf{x} = 0 \; \forall i,$$

and $\|\mathbf{y}_i\| \leq 1$.

Primal-Dual Newton's Method [Chan1996]

Motivation for an improved implementation [Chan1996]:

Consider $g(\mathbf{x})$ (ignoring β for simplicity):

$$g(\mathbf{x}) = \mathbf{A}^T(\mathbf{Ax} - \mathbf{b}) + \gamma \sum_{i=1}^{n} \frac{\mathbf{D}_i \mathbf{D}_i^T \mathbf{x}}{\|\mathbf{D}_i^T \mathbf{x}\|_2}$$

Introduce a new variable:

$$\mathbf{y}_i = \frac{\mathbf{D}_i^T \mathbf{x}}{\|\mathbf{D}_i^T \mathbf{x}\|_2} \text{ where } \mathbf{y}_i \text{ is } 2 \times 1.$$

Then the first order condition becomes:

$$g(\mathbf{y}, \mathbf{x}) = \mathbf{A}^T(\mathbf{Ax} - \mathbf{b}) + \gamma \sum_{i=1}^{n} \mathbf{D}_i \mathbf{y}_i = 0,$$
$$h(\mathbf{y}, \mathbf{x}) = \|\mathbf{D}_i^T \mathbf{x}\|_2 \mathbf{y}_i - \mathbf{D}_i^T \mathbf{x} = 0 \; \forall i,$$

and $\|\mathbf{y}_i\| \leq 1$.

Primal-Dual Newton's Method [Chan1996]

Outline

Initial parameter selection (γ):

- Generalized Cross Validation (GCV):
  Minimize $G(\gamma) = \sum_{k=1}^{m} [\mathbf{b}_k - (\mathbf{A}\tilde{\mathbf{x}}_{\gamma}^{(k)})_k]^2$
  where the $\tilde{x}^{(k)}$ minimizes
  $$\frac{1}{2}\|\mathbf{Ax} - \mathbf{b}\|_2^2 + \gamma\mathbf{R}(\mathbf{x})$$
  when the k^{th} measurement of **b** is omitted.
  - Simple implementation for Tikhonov and TSVD regularization
  - Too expensive to use directly for TV method.
- Discrepancy Principle:
  Choose γ such that $\|\mathbf{Ax}_{\gamma} - \mathbf{b}\|_2 = \nu E[\|\epsilon\|_2]$ where $\nu = 2$ is a safety factor.
  - Requires prior knowledge of the distribution of ϵ which we have assumed we know. -> good alternative for TV method

Outline

Motivation for statistical based diagnostics.

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \epsilon$$

Assumptions: The noise ϵ where $\epsilon \sim N(\mathbf{0}, \mathbf{I_m})$
$\mathbf{x}^*$ is the estimate of $\mathbf{x}$ the the residual vector is

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}^*$$

Where we expect $\mathbf{r} \sim \mathbf{N}(\mathbf{0}, \mathbf{I_m})$

This characteristic of the residual inspired three diagnostics [RustOleary2008]

Motivation for statistical based diagnostics.

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \epsilon$$

Assumptions: The noise ϵ where $\epsilon \sim N(\mathbf{0}, \mathbf{I_m})$
$\mathbf{x}^*$ is the estimate of $\mathbf{x}$ the the residual vector is

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}^*$$

Where we expect $\mathbf{r} \sim \mathbf{N}(\mathbf{0}, \mathbf{I_m})$

This characteristic of the residual inspired three diagnostics [RustOleary2008]

Diagnostic 1: *The residual norm squared should be within two standard deviations of the expected value of* $\|\epsilon\|_2^2$ *of* $\|\mathbf{r}\|_2^2 \in [m - 2\sqrt{2m}, m + 2\sqrt{2m}]$ *where* $m = E[\|\epsilon\|_2^2]$.

Diagnostic 2: *Goodness of fit of the normal curve to the histogram of the elements of the residual vector* **r***.*

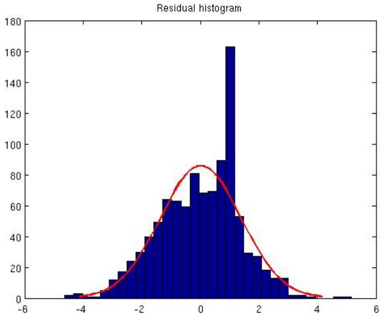

Figure: checkperiod.m by Dianne O'Leary

Diagnostic 3: *Consider the elements* **r** *as time series with index* $j = 1, ..., m$. *Find the cumulative periodogram of the residual time-series and check if it is within* 95% *confidence band of the cumulative periodogram of the time series of white noise.*

The *cumulative periodogram* is the partial sum of the *periodogram* where the *periodogram* is the sum of the square of the real and imaginary parts of the discrete Fourier transform.

Diagnostic 3:

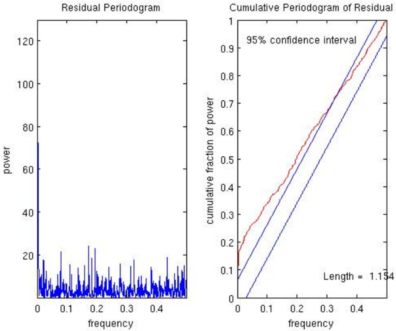

Figure: checkperiod.m by Dianne O'Leary

Outline

Software package

Frontend

- Graphical User Interface (GUI) built using Matlab's GUI toolbox

Backend

- Regularization method
  - Tikhonov and Truncated SVD methods from *RestoreTool* [Nagy2002]
  - Total Variation regularization method [Cash 2012]
- Method for initial parameter selection
  - Generalized Cross-Validation (GCV) in *RestoreTool* for regularization methods included.
  - Discrepancy Principle [Cash 2012]
- Validate candidate solutions using statistical diagnostics
  - Adapt existing code for statistical diagnostics from Dianne O'Leary [Cash 2012]

GUI Demonstration in Matlab

Outline

Effect of SNR on statistical diagnostics

$$SNR = 10\log_{10}(\frac{\|\mathbf{b}\|^2}{\|\epsilon\|^2}).$$

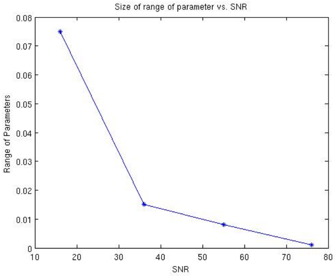

Figure: Length of interval of parameter satisfying Diagnostic 1 for Tikhohnov Method on a 16×16 segment of the image “cell.tif” .

Effects of γ on computation time

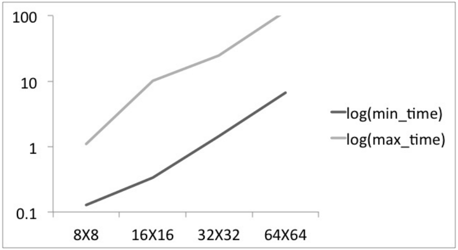

Figure: The difference in log_{10} of the computational time for the TV regularization method for parameters between $\gamma = 1$ and $\gamma = 10^{-9}$

Computation time of the TV regularization method is dependent on the number of CG iterations, preconditioners were explored but included.

Outline

Results for larger images

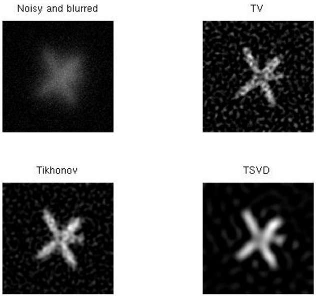

Figure: 256 $\times$ 256 image of Satellite with PSF provided in *RestoreTool* with zero boundary conditions and SNR=9

Results for larger images

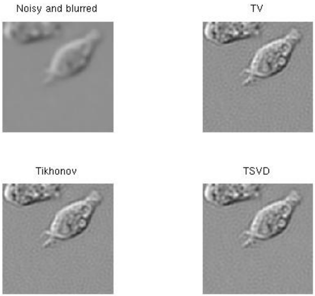

Figure: 129 × 129 image of “cell.tif” with Gaussian blur and zero boundary conditions with SNR of 60

Outline

Validation of software

- Modularly developed
- Modules validated independently with small examples that could be confirmed by hand or by comparing to existing code or functions
    - CG code was both validated by solving a known test problem as well as the code and results could be compared to matlabs *pcg.m*.
    - search direction on dual variable was first verified that solution met the constraints as well results were compared to the Vogel's code implementation independently.
- Results found using a implementations of the Primal-Dual Newton's method implemented by Curtis Vogel

Primal-Dual Newton's method

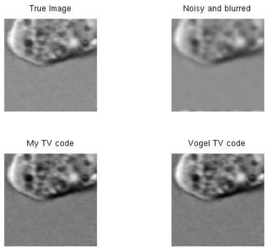

Figure: Results of the TV regularization method for my implementation and code by Curtis Vogel.

Relative error of my implementation was 0.9% compared to 1.02% for the given example.

Outline

Validation of software usefulness

- Presented as part of AMSC 663/664
- Presented and distributed to undergraduate students in Deblurring Digital Images (CMSC/AMSC 498D) as an educational tool
- Presented to the AMSC student seminar
- Proof of concept for tool for picking regularization method and parameter

Project Deliverables:
Parameter Selection Tool for Solution to Ill-Posed Problems

- Graphical user interface that could be used by a researcher/student
- All the necessary code for computing the regularization solution, selecting an initial parameter, validating solutions with the diagnostics.

References I

- Tony F. Chan and Jianhong Shen.
  Image Processing and Analysis.
  SIAM, Philadelphia, PA, 2005.
- Wayne A. Fuller.
  Introduction to Statistical Time Series.
  SIAM, Philadelphia, PA, 1998.
- Per Christian Hansen.
  Rank-Deficient and Discrete Ill-Posed Problems:Numerical Aspects of Linear Inversion.
  Wiley-Interscience, New York, NY, 1996.
- C.T. Kelly.
  Iterative Methods for Linear and Nonlinear Equations.
  SIAM, Philadelphia, PA, 1995.

References II

- Stephen G. Nash.
  Linear and Nonlinear Programming.
  McGraw-Hill, 1996.
- James G. Nagy.
  RestoreTool.
  http://www.mathcs.emory.edu/ nagy/RestoreTools/.
- Dianne P. 0'Leary.
  Scientific Computing with Case Studies.
  SIAM, Philadelphia, PA, 2009.
- Curtis R. Vogel.
  Computational Methods for Inverse Problems.
  SIAM, Philadelphia, PA, 2002.

References III

R. Acar, C.R. Vogel.
Analysis of Bounded Variation Penalty Methods for Ill-Posed Problems
Inverse Problems, Vol.10, pp.1217-1229, 1994.

Tony F. Chan, Gene H. Golub and Pep Mulet.
A nonlinear primal-dual method for total variation-based image restoration.
Lecture Notes in Control and Information Sciences, Vol.219, pp.241-251, 1996.

Katrina P. Lee, James G. Nagy and Lisa Perron.
Iterative Methods for Image Restoration: A Matlab Object Oriented Approach
Numerical Algorithms, Vol.36, pp.73-93, 2004.

References IV

- Peng Liu, Dingsheng Liu.
  Selection of regularization parameter based on generalized cross-validation in total variation remote sensing image restoration.
  Image and Signal Processing for Remote Sensing, Vol.7830, 2010.

- James G. Nagy, K. Palmer and L. Perrone.
  Iterative Methods for Image Deblurring: A Matlab Object Oriented Approach.
  Numerical Algorithms. 36: 73-93, 2004.

- Bert W. Rust and Dianne P. O'Leary.
  Residual programs for choosing regularization parameters for ill-posed problems.
  Inverse Problems. 24:034005 (30 pages), 2008. (invited paper)

References V

Bert W. Rust.
Parameter selection for constrained solutions to ill-posed problems.
SIAM Journal on Scientific Computing, Vol.17, pp.227-238, 1996.