Automated Parameter Selection Tool for Solution to Ill-Posed Problems

An Application to Image Processing

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

AMSC 663, 2011

Outline

Outline

The problem.

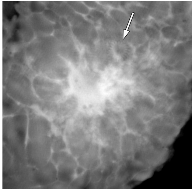

Figure: Stevens G M et al. Radiology 2003;228:569-575

- Images are expensive to produce.
- Images can be distorted and/or noisy.
  - physics of the measurement
  - structure of the material (humans)
- Used for making medical decisions

Deblurring/Denoising medical images are an example of ill-posed inverse problem.

- Ill-posed: Not well-posed
- Inverse: We are seeking the input image given the output image
- Solve by replacing the problem with an approximate well-posed problem by introducing a constraint or regularization parameter
- Even if the blur and noise level are known finding a good solution is difficult.

The challenge: Selecting a good regularization parameter

Deblurring/Denoising medical images are an example of ill-posed inverse problem.

- Ill-posed: Not well-posed
- Inverse: We are seeking the input image given the output image
- Solve by replacing the problem with an approximate well-posed problem by introducing a constraint or regularization parameter
- Even if the blur and noise level are known finding a good solution is difficult.

The challenge: Selecting a good regularization parameter

Selecting a good regularization parameter

- Expensive
- Problem dependent
- Subject to bias

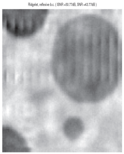

Figure: Images and Data courtesy of Dianne O'Leary

But what is unexpected is often what we are interested in

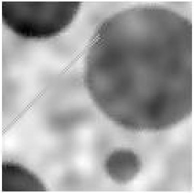

Figure: Images and Data courtesy of Dianne O'Leary

Outline

1. Motivation
   - An application to medical images
2. Project Goal and Implementation
   - Automated tool for parameter selection
   - Regularization method for ill-posed problems
   - Initial parameter selection
   - Test parameters to validate candidate solutions
3. Testing and Validation
   - Data
   - Validation
4. Schedule and Milestones

Building a software package for parameter selection

Frontend

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

Backend

- Regularization method
  - Regularization methods from *RestoreTool*
  - Implement code for Total Variation regularization method
- Initial parameter selection
  - Generalized Cross-Validation (GCV) in *RestoreTool* for regularization methods included.
  - Implement code for GCV for Total Variation
- Test parameters to validate candidate solutions
  - Adapt existing code for statistical diagnostics from Dianne O'Leary

Making a GUI in MATLAB

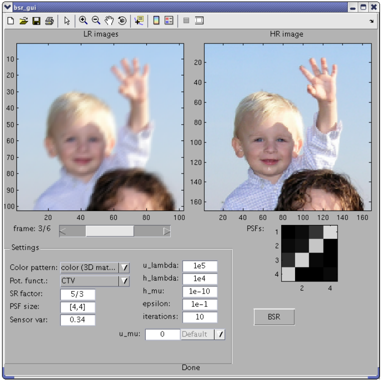

Figure: Filip Šroubek, Academy of Sciences of the Czech Republic

Outline

Defining the discrete ill-posed problem

$$\mathbf{Ax} = \mathbf{b}$$

where

- **A** is a known $m \times n$ matrix where $m \geq n$
  - Point Spread Function (PSF) or blurring functions
  - Ill-conditioned
- **b** is a known $m \times 1$ vector (blurred image)
- **x** is unknown $n \times 1$ vector (true image)

Because **A** is ill-conditioned solving $\mathbf{Ax} = \mathbf{b}$ directly or the equivalent least square problem min $\|\mathbf{Ax} - \mathbf{b}\|_2^2$ is not feasible.

Solving ill-posed problems using regularization parameters

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

- Where $\Omega(x)$ is smoothing function or penalty function and γ is the regularization parameter

Tikhonov: $\Omega(x) = \|\mathbf{Lx}\|_2^2$ and $\gamma = \lambda^2$ where **L** is the identity matrix, approximation of the first derivative operator, a diagonal weighting matrix [Hansen1998], or any other type of operator based on the problem and the desired features.

Truncated SVD: $\min \|\mathbf{A_k x} - \mathbf{b}\|_2^2$ where $A_k = \sum_{i=1}^{k} u_i \sigma_i v_i^T$ where the regularization parameter is k or the level of truncation.

Total Variation: $\Omega(x) = TV(\mathbf{x})$ and $\gamma = \lambda$ where $TV(\mathbf{x}) = \|\nabla \mathbf{x}\|_1$.

Approach to solving the Total Variation regularization problem.

1. Unconstrained optimization problem:

$$\min \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda TV(\mathbf{x})$$

2. Constrained optimization problem:

$$\min TV(\mathbf{x})$$

subject to

$$\|\mathbf{Ax} - \mathbf{b}\|_2^2 = \sigma^2$$

Select an algorithm to solve one of the above problems

Approach to solving the Total Variation regularization problem.

1. Unconstrained optimization problem:

$$\min \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda TV(\mathbf{x})$$

2. Constrained optimization problem:

$$\min TV(\mathbf{x})$$

subject to

$$\|\mathbf{Ax} - \mathbf{b}\|_2^2 = \sigma^2$$

Select an algorithm to solve one of the above problems

Outline

Generalized Cross-Validation (GCV)
Minimize

$$G(\lambda) = \sum_{k=1}^{m} [b_k - (\mathbf{A}\tilde{\mathbf{x}}_\lambda^{(k)})_k]^2$$

where the $\tilde{x}^{(k)}$ is the estimate when the k^{th} measurement of **b** is omitted.

Find the model that best predicts the missing measurements as a function of the other measurements.

Generalized Cross-Validation (GCV)
Minimize

$$G(\lambda) = \sum_{k=1}^{m} [b_k - (\mathbf{A}\tilde{\mathbf{x}}_\lambda^{(k)})_k]^2$$

where the $\tilde{x}^{(k)}$ is the estimate when the k^{th} measurement of **b** is omitted.

Find the model that best predicts the missing measurements as a function of the other measurements.

Outline

Motivation for Statistical Based Diagnostics.

$$\mathbf{b} = \mathbf{Ax} + \epsilon$$

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

$$\mathbf{r} = \mathbf{b} - \mathbf{Ax}^*$$

Where we expect $\mathbf{r} \sim \mathbf{N}(\mathbf{0}, \mathbf{I_m})$
(If $\epsilon \sim N(\mathbf{0}, \mathbf{S^2})$ the problem can be rescaled by the matrix $\mathbf{S}^{-1}$)

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 $\mathbf{r} \sim \mathbf{N}(\mathbf{0}, \mathbf{I_m})$
(If $\epsilon \sim N(\mathbf{0}, \mathbf{S^2})$ the problem can be rescaled by the matrix $\mathbf{S^{-1}}$)

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* $\|\tilde{\mathbf{r}}\|_2^2 \in [m - 2\sqrt{2m}, m + 2\sqrt{2m}]$ *where* $m = E[\|\epsilon\|_2^2]$.

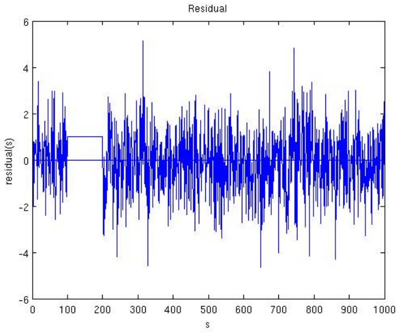

Figure: checkperiod.m by Dianne O'Leary

Diagnostic 2: *Goodness of fit of the normal curve to the histogram of the elements of the residual vector* $\tilde{r}$.

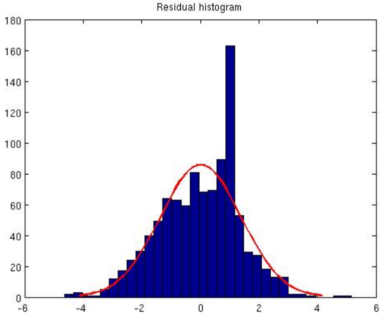

Figure: checkperiod.m by Dianne O'Leary

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 [RustOleary2008].*

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.

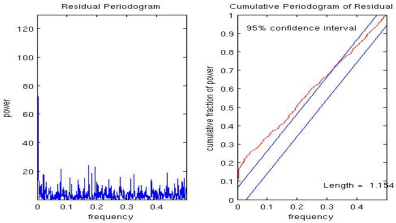

Figure: checkperiod.m by Dianne O'Leary

Outline

1. Motivation
   - An application to medical images
2. Project Goal and Implementation
   - Automated tool for parameter selection
   - Regularization method for ill-posed problems
   - Initial parameter selection
   - Test parameters to validate candidate solutions
3. Testing and Validation
   - Data
   - Validation
4. Schedule and Milestones

Data (images and PSF) for development and validation

Figure: Examples of generated image for development of various sizes

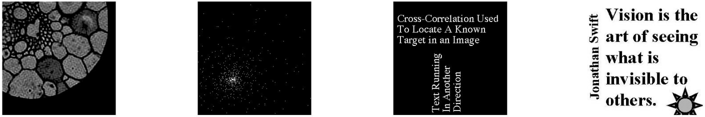

Figure: Test images from RestoreTool, 256x256

Outline

Testing and validation of the software package

- Assumption: Tools from RestoreTool have been thoroughly tested and validated.
- Total Variation (TV) regularization method
  - Use binary test images where without noise the method should return close to the original
  - Compare the expected convergence rate of the method chosen to my implementation of TV regularization
  - Use test problems from RestoreTool where the true image can be compared to result (one way to compare is measuring the root mean square average magnitude of the error).
- Generalized cross-validation (GCV) and Diagnostics test
  - Use test problems where the true image can be compared to the result for the parameter selected by GCV or range of parameters that meet the diagnostics.

Semester 1

- Learn about and become comfortable with RestoreTool (object oriented programming) and Matlab GUI.
- **Oct 15 - Milestone**: A basic frontend for RestoreTool regularization method.
- **Nov 30 - Milestone**: Validated statistics based diagnostics in RestoreTool framework.
- Read literature on Total Variation regularization and determine iterative method.
- **Dec 1 - Milestone**: Develop a GUI for parameter selection using the diagnostic test and RestoreTool.
- **Dec 15 - Milestone**: Outline of Total Variation regularization method and basic implementation Matlab.
- **Dec 15 - Milestone**:Deliver mid-year report.

Semester 2

- Read literature about High-Performance computing and Matlab's parallel toolbox.
- **Feb 1 - Milestone**: Total Variation regularization in RestoreTool framework.
- **Feb 15 - Milestone**Validation Total Variation regularization tool.
- **Feb 30 - Milestone**: Generalized Cross Validation for Total Variation tool.
- **Mar 15 - Milestone**Validation of Generalized Cross Validation for Total Variation tool.
- **Mar 30 - Milestone**: Develop GUI for parameter selection to include regularization method from RestoreTool along with the Total Variation tool.
- **April 15 - Milestone**: Optimize software package (using parallel toolbox in Matlab if available).
- **April 30 - Milestone**: Deliver final software package.
- **May 15 - Milestone**:Deliver Final presentation.

References I

- Tony F. Chan and Jianhong Shen.
  Image Processing and Analysis.
  SIAM, Philadelphia, PA, 2005.
- Per Christian Hansen.
  Rank-Deficient and Discrete Ill-Posed Problems:Numerical Aspects of Linear Inversion.
  SIAM, Philadelphia, PA, 1998.
- Per Christian Hansen.
  Regularization Toolbox.
  http://www2.imm.dtu.dk/ pch/Regutools/regutools.html.
- James G. Nagy.
  RestoreTool.
  http://www.mathcs.emory.edu/ nagy/RestoreTools/.

References II

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.

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 papaer

References III

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