Separable Nonlinear Least Squares Problems in Image Processing

Julianne Chung and James Nagy
Emory University
Atlanta, GA, USA

Collaborators: Eldad Haber (Emory)
Per Christian Hansen (Tech. Univ. of Denmark)
Dianne O’Leary (University of Maryland)
Inverse Problems in Imaging

Imaging problems are often modeled as:

\[ b = Ax + e \]

where

- **A** - large, ill-conditioned matrix
- **b** - known, measured (image) data
- **e** - noise, statistical properties may be known

Goal: Compute approximation of image **x**
A more realistic image formation model is:

\[ b = A(y)x + e \]

where

- \( A(y) \) - large, ill-conditioned matrix
- \( b \) - known, measured (image) data
- \( e \) - noise, statistical properties may be known
- \( y \) - parameters defining \( A \), usually approximated

Goal: Compute approximation of image \( x \) and improve estimate of parameters \( y \)
Application: Image Deblurring

- \( b = A(y) x + e \) = observed image
  where \( y \) describes blurring function

- Given: \( b \) and an estimate of \( y \)

- Standard Image Deblurring:
  Compute approximation of \( x \)

- Better approach:
  Jointly improve estimate of \( y \)
  and compute approximation of \( x \).
Application: Image Deblurring

- \( \mathbf{b} = \mathbf{A}(\mathbf{y}) \mathbf{x} + \mathbf{e} = \) observed image where \( \mathbf{y} \) describes blurring function
- Given: \( \mathbf{b} \) and an estimate of \( \mathbf{y} \)
  - Standard Image Deblurring: Compute approximation of \( \mathbf{x} \)
  - Better approach: Jointly improve estimate of \( \mathbf{y} \) and compute approximation of \( \mathbf{x} \).
The Linear Problem: \( b = Ax + e \)
The Nonlinear Problem: \( b = A(y)x + e \)

Example: Image Deblurring

Concluding Remarks

Application: Image Deblurring

- \( b = A(y)x + e = \) observed image where \( y \) describes blurring function
- Given: \( b \) and an estimate of \( y \)
- Standard Image Deblurring:
  Compute approximation of \( x \)
- Better approach:
  Jointly improve estimate of \( y \)
  and compute approximation of \( x \).

Reconstruction using initial PSF

Julianne Chung and James Nagy
Emory University
Atlanta, GA

Separable Nonlinear Least Squares Problems in Image Processing
Application: Image Deblurring

- \( b = A(y)x + e \) = observed image where \( y \) describes blurring function
- Given: \( b \) and an estimate of \( y \)
- Standard Image Deblurring:
  Compute approximation of \( x \)
- Better approach:
  Jointly improve estimate of \( y \) and compute approximation of \( x \).
Application: Image Data Fusion

- $b_j = A(y_j)x + e_j$
  (collected low resolution images)
Application: Image Data Fusion

- \( b_j = A(y_j)x + e_j \)
  (collected low resolution images)
Application: Image Data Fusion

- $\mathbf{b}_j = \mathbf{A}(\mathbf{y}_j) \mathbf{x} + \mathbf{e}_j$
  (collected low resolution images)
Application: Image Data Fusion

\[ b_j = A(y_j) x + e_j \]
(collected low resolution images)
Application: Image Data Fusion

- \( b_j = A(y_j) x + e_j \)
  (collected low resolution images)
Application: Image Data Fusion

- \( b_j = A(y_j)x + e_j \) (collected low resolution images)
- \[
\begin{bmatrix}
  b_1 \\
  \vdots \\
  b_m \\
\end{bmatrix} = \begin{bmatrix}
  A(y_1) \\
  \vdots \\
  A(y_m) \\
\end{bmatrix} x + \begin{bmatrix}
  e_1 \\
  \vdots \\
  e_m \\
\end{bmatrix}
\]
- \( b = A(y)x + e \)
- \( y = \) registration, blurring, etc., parameters
- Goal: Improve parameters \( y \) and compute \( x \)
Application: Image Data Fusion

- \( b_j = A(y_j)x + e_j \) (collected low resolution images)

\[
\begin{bmatrix}
    b_1 \\
    \vdots \\
    b_m
\end{bmatrix} =
\begin{bmatrix}
    A(y_1) \\
    \vdots \\
    A(y_m)
\end{bmatrix} x +
\begin{bmatrix}
    e_1 \\
    \vdots \\
    e_m
\end{bmatrix}
\]

\[
b = A(y)x + e
\]

- \( y = \) registration, blurring, etc., parameters
- Goal: Improve parameters \( y \) and compute \( x \)
Outline

1. The Linear Problem: $b = Ax + e$
2. The Nonlinear Problem: $b = A(y)x + e$
3. Example: Image Deblurring
4. Concluding Remarks
The Linear Problem

Assume $A = A(y)$ is known exactly.

- We are given $A$ and $b$, where
  
  $$b = Ax + e$$

- $A$ is an ill-conditioned matrix, and we do not know $e$.
- We want to compute an approximation of $x$.

Bad idea:

- $e$ is small, so ignore it, and
- use $x_{\text{inv}} \approx A^{-1}b$
The Linear Problem

Assume $A = A(y)$ is known exactly.

- We are given $A$ and $b$, where

$$b = Ax + e$$

- $A$ is an ill-conditioned matrix, and we do not know $e$.
- We want to compute an approximation of $x$.

- Bad idea:
  - $e$ is small, so ignore it, and
  - use $x_{\text{inv}} \approx A^{-1}b$
Example: Inverse Heat Equation

Regularization Tools test problem: heat.m
P. C. Hansen, www2.imm.dtu.dk/~pch/Regutools
Example: Inverse Heat Equation

If $A$ and $b$ are known exactly, can get an accurate reconstruction.

Inverse solution $x = A^{-1}b$

Noise free data, $A^*x$
Example: Inverse Heat Equation

But, if $b$ contains a small amount of noise,
Example: Inverse Heat Equation

But, if $b$ contains a small amount of noise, then we get a poor reconstruction!
SVD Analysis

An important linear algebra tool: Singular Value Decomposition

Let \( \mathbf{A} = \mathbf{U} \Sigma \mathbf{V}^T \) where

- \( \Sigma \) = diag\((\sigma_1, \sigma_2, \ldots, \sigma_n)\), \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0 \)
- \( \mathbf{U}^T \mathbf{U} = \mathbf{I} \), \( \mathbf{V}^T \mathbf{V} = \mathbf{I} \)
- \( \mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{bmatrix} \) (left singular vectors)
- \( \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \end{bmatrix} \) (right singular vectors)
The naïve inverse solution can then be represented as:

\[ x = A^{-1}b \]

\[ = V\Sigma^{-1}U^Tb \]

\[ = \sum_{i=1}^{n} \frac{u_i^Tb}{\sigma_i}v_i \]
The naïve inverse solution can then be represented as:

\[ \hat{x} = A^{-1}(b + e) \]

\[ = V \Sigma^{-1} U^T (b + e) \]

\[ = \sum_{i=1}^{n} \frac{u_i^T (b + e)}{\sigma_i} v_i \]
The naïve inverse solution can then be represented as:

\[
\hat{x} = A^{-1}(b + e) \\
= V\Sigma^{-1}U^T(b + e) \\
= \sum_{i=1}^{n} \frac{u_i^T(b + e)}{\sigma_i}v_i \\
= \sum_{i=1}^{n} \frac{u_i^Tb}{\sigma_i}v_i + \sum_{i=1}^{n} \frac{u_i^Te}{\sigma_i}v_i \\
= x + \text{error}
\]
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$. 

Singular values
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$. Large $\sigma_i \leftrightarrow$ smooth (low frequency) $v_i$. 

- **Singular values**
  - $10^{-6}$ to $10^6$
  - Graph shows decreasing trend from left to right.

- **Singular vector, $v_1$**
  - Range from $-0.2$ to $0.2$
  - Graph shows smooth curve from left to right.

---

Julianne Chung and James Nagy
Emory University
Atlanta, GA

Separable Nonlinear Least Squares Problems in Image Processing
The Linear Problem: \( b = Ax + e \)

The Nonlinear Problem: \( b = A(y)x + e \)

Example: Image Deblurring

Example: Inverse Heat Equation

Error term depends on singular values \( \sigma_i \) and singular vectors \( v_i \).

Large \( \sigma_i \) ↔ smooth (low frequency) \( v_i \)
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.

Large $\sigma_i \leftrightarrow$ smooth (low frequency) $v_i$.
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.
Large $\sigma_i \leftrightarrow$ smooth (low frequency) $v_i$
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.

Large $\sigma_i \leftrightarrow$ smooth (low frequency) $v_i$
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$. Small $\sigma_i \leftrightarrow$ oscillating (high frequency) $v_i$.
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.

Small $\sigma_i \leftrightarrow$ oscillating (high frequency) $v_i$

![Graph showing singular values and singular vector](image-url)
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.
Small $\sigma_i \leftrightarrow$ oscillating (high frequency) $v_i$
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.
Small $\sigma_i \leftrightarrow$ oscillating (high frequency) $v_i$
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$. Small $\sigma_i \leftrightarrow$ oscillating (high frequency) $v_i$. 

![Singular values](image1.png)  
![Singular vector, $v_{125}$](image2.png)
Example: Inverse Heat Equation

Error term depends on singular values $\sigma_i$ and singular vectors $v_i$.
Small $\sigma_i \leftrightarrow$ oscillating (high frequency) $v_i$.
The naïve inverse solution can then be represented as:

\[ \hat{x} = A^{-1}(b + e) \]

\[ = V\Sigma^{-1}U^T(b + e) \]

\[ = \sum_{i=1}^{n} \frac{u_i^T(b + e)}{\sigma_i} v_i \]

\[ = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i + \sum_{i=1}^{n} \frac{u_i^T e}{\sigma_i} v_i \]

\[ = x + \text{error} \]
Regularization by Filtering

**Basic Idea:** Filter out effects of small singular values.  
(Hansen, SIAM, 1997)

\[ x_{\text{reg}} = A_{\text{reg}}^{-1}b = V \Phi \Sigma^{-1} U^T b = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i , \]

where \( \Phi = \text{diag}(\phi_1, \phi_2, \ldots, \phi_n) \)

The "filter factors" satisfy

\[ \phi_i \approx \begin{cases} 1 & \text{if } \sigma_i \text{ is large} \\ 0 & \text{if } \sigma_i \text{ is small} \end{cases} \]
The Linear Problem: $b = Ax + e$

The Nonlinear Problem: $b = A(y)x + e$

Example: Image Deblurring

Concluding Remarks

An Example: Tikhonov Regularization

$$\min_x \left\{ \|b - Ax\|_2^2 + \lambda^2 \|x\|_2^2 \right\} \iff \min_x \left\| \begin{bmatrix} b \\ 0 \end{bmatrix} - \begin{bmatrix} A \\ \lambda I \end{bmatrix} x \right\|_2^2$$
The Linear Problem: \( b = Ax + e \)

The Nonlinear Problem: \( b = A(y)x + e \)

Example: Image Deblurring

Concluding Remarks

An Example: Tikhonov Regularization

\[
\min_x \left\{ \| b - Ax \|^2_2 + \lambda^2 \| x \|^2_2 \right\} \quad \Leftrightarrow \quad \min_x \left\| \begin{bmatrix} b \\ 0 \end{bmatrix} - \begin{bmatrix} A \\ \lambda I \end{bmatrix} x \right\|^2_2
\]

An equivalent SVD filtering formulation:

\[
x_{tik} = \sum_{i=1}^n \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{u_i^T b}{\sigma_i} v_i
\]
The Linear Problem: \( b = Ax + e \)

The Nonlinear Problem: \( b = A(y) x + e \)

Example: Image Deblurring

Concluding Remarks

An Example: Tikhonov Regularization

\[
\min_x \left\{ \| b - Ax \|^2_2 + \lambda^2 \| x \|^2_2 \right\} \quad \Leftrightarrow \quad \min_x \left\| \begin{bmatrix} b \\ 0 \end{bmatrix} - \begin{bmatrix} A \\ \lambda I \end{bmatrix} x \right\|^2_2
\]

An equivalent SVD filtering formulation:

\[
x_{\text{tik}} = \sum_{i=1}^{n} \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2} \frac{u_i^T b}{\sigma_i} v_i
\]

Julianne Chung and James Nagy  Emory University  Atlanta, GA  Separable Nonlinear Least Squares Problems  in Image Processing
Choosing Regularization Parameters

Lots of choices: Generalized Cross Validation (GCV), L-curve, discrepancy principle, ...
Choosing Regularization Parameters

Lots of choices: Generalized Cross Validation (GCV), L-curve, discrepancy principle, ... 

GCV and Tikhonov: Choose $\lambda$ to minimize

$$GCV(\lambda) = \frac{n \sum_{i=1}^{n} \left( \frac{u_i^T b}{\sigma_i^2 + \lambda^2} \right)^2}{\left( \sum_{i=1}^{n} \frac{1}{\sigma_i^2 + \lambda^2} \right)^2}$$
Example: Inverse Heat Equation

Reconstruction using Tikhonov reg. can be better than $x_{\text{inv}}$. Quality of reconstruction depends on $\lambda$. But $\lambda$ depends on $A$ and $b$. 

Desired solution, $x$

Inverse solution $x = A^{-1}b$
Example: Inverse Heat Equation

Reconstruction using Tikhonov reg. can be better than $x_{\text{inv}}$. Quality of reconstruction depends on $\lambda$. But $\lambda$ depends on $A$ and $b$. 

![Regularized Solution, $\lambda = 0.0005$](image1.png)

![Inverse solution $x = A^{-1}b$](image2.png)
Example: Inverse Heat Equation

Reconstruction using Tikhonov reg. can be better than $x_{\text{inv}}$.

Quality of reconstruction depends on $\lambda$.

But $\lambda$ depends on $A$ and $b$.

Regularized Solution, $\lambda = 0.05$

Inverse solution $x = A^{-1}b$
Example: Inverse Heat Equation

Reconstruction using Tikhonov reg. can be better than $x_{\text{inv}}$. Quality of reconstruction depends on $\lambda$. But $\lambda$ depends on $A$ and $b$. 

Regularized Solution, $\lambda = 0.005$

Inverse solution $x = A^{-1}b$
Some remarks:

- For large matrices, computing SVD is expensive.

- SVD algorithms do not readily simplify for structured or sparse matrices.

- Alternative for large scale problems: LSQR iteration (Paige and Saunders, ACM TOMS, 1982)
Lanczos Bidiagonalization (LBD)

Given $A$ and $b$, for $k = 1, 2, ..., $, compute

- $W_k = \begin{bmatrix} w_1 & w_2 & \cdots & w_k & w_{k+1} \end{bmatrix}$, $w_1 = b/\|b\|$
- $Z_k = \begin{bmatrix} z_1 & z_2 & \cdots & z_k \end{bmatrix}$

- $B_k = \begin{bmatrix} \alpha_1 & \beta_2 & \alpha_2 & &  \\ & \beta_2 & \alpha_2 & \cdots & \beta_k \\ & & \ddots & \cdots & \alpha_k \\ & & & \beta_k & \alpha_k \\ & & & & \beta_{k+1} \end{bmatrix}$

where $W_k$ and $Z_k$ have orthonormal columns, and

$A^T W_k = Z_k B_k^T + \alpha_{k+1} z_{k+1} e_{k+1}^T$

$A Z_k = W_k B_k$
LBD and LSQR

At $k$th LBD iteration, use $QR$ to solve projected LS problem:

$$\min_{x \in R(Z_k)} \| b - Ax \|_2^2 = \min_f \| W_k^T b - B_k f \|_2^2 = \min_f \| \beta e_1 - B_k f \|_2^2$$

where $x_k = Z_k f$
At $k$th LBD iteration, use $QR$ to solve projected LS problem:

$$
\min_{x \in R(Z_k)} \| b - Ax \|_2^2 = \min_f \| W_k^T b - B_k f \|_2^2 = \min_f \| \beta e_1 - B_k f \|_2^2
$$

where $x_k = Z_k f$

For our ill-posed inverse problems:

- Singular values of $B_k$ converge to $k$ largest sing. values of $A$.
- Thus, $x_k$ is in a subspace that approximates a subspace spanned by the large singular components of $A$.
  - For $k < n$, $x_k$ is a regularized solution.
  - $x_n = x_{\text{inv}} = A^{-1} b$ (bad approximation)
Example: Inverse Heat Equation

Singular values of $B_k$ converge to large singular values of $A$. Thus, for early iterations $k$:

$$f = B_k \backslash W_k b$$

$$x_k = Z_k f$$

is a regularized reconstruction.
Example: Inverse Heat Equation

Singular values of $B_k$ converge to large singular values of $A$. Thus, for early iterations $k$: \[
\begin{align*}
f &= B_k \backslash W_k b \\
x_k &= Z_k f
\end{align*}
\]
is a regularized reconstruction.
Example: Inverse Heat Equation

Singular values of $B_k$ converge to large singular values of $A$. Thus, for later iterations $k$: 

$$f = B_k \backslash W_k b$$

$$x_k = Z_k f$$

is a noisy reconstruction.
Example: Inverse Heat Equation

Singular values of $B_k$ converge to large singular values of $A$. Thus, for later iterations $k$: 

$$ f = B_k \backslash W_k b $$

$$ x_k = Z_k f $$

is a noisy reconstruction.
Lanczos Based Hybrid Methods

To avoid noisy reconstructions, embed regularization in LBD:

- O’Leary and Simmons, SISSC, 1981.
- Björck, Grimme, and Van Dooren, BIT, 1994.
- Chung, N, O’Leary, ETNA 2007 (HyBR Implementation)
Regularize the Projected Least Squares Problem

To stabilize convergence, regularize the projected problem:

$$\min_{f} \left\| \begin{bmatrix} \beta e_1 \\ 0 \end{bmatrix} - \begin{bmatrix} B_k \\ \lambda I \end{bmatrix} f \right\|_2^2$$

Note: $B_k$ is very small compared to $A$, so

- Can use “expensive” methods to choose $\lambda$ (e.g., GCV)
- Very little regularization is needed in early iterations.
- GCV tends to choose too large $\lambda$ for bidiagonal system.
  Our remedy: Use a weighted GCV (Chung, N, O’Leary, 2007)
- Can also use WGCV information to estimate stopping iteration
  (approach similar to Björck, Grimme, and Van Dooren, BIT, 1994).
Example: Inverse Heat Equation

### LSQR (no regularization)

\[
\begin{align*}
\mathbf{f} &= \mathbf{B}_k \backslash \mathbf{W}_k \mathbf{b} \\
\mathbf{x}_k &= \mathbf{Z}_k \mathbf{f}
\end{align*}
\]

### HyBR (Tikhonov regularization)

\[
\begin{align*}
\mathbf{f} &= \left[ \begin{array}{c} \mathbf{B}_k \\ \lambda_k \mathbf{l} \end{array} \right] \backslash \left[ \begin{array}{c} \mathbf{W}_k \mathbf{b} \\ \mathbf{0} \end{array} \right] \\
\mathbf{x}_k &= \mathbf{Z}_k \mathbf{f}
\end{align*}
\]

$\lambda = 0.0115$
Example: Inverse Heat Equation

\[ f = B_k \backslash W_k b \]
\[ x_k = Z_k f \]

**LSQR (no regularization)**

**HyBR (Tikhonov regularization)**

\[ f = \begin{bmatrix} B_k \\ \lambda_k I \end{bmatrix} \backslash \begin{bmatrix} W_k b \\ 0 \end{bmatrix} \]
\[ x_k = Z_k f \]

\[ \lambda = 0.0074 \]

iteration = 15
Example: Inverse Heat Equation

### LSQR (no regularization)

\[ f = B_k \backslash W_k b \]
\[ x_k = Z_k f \]

### HyBR (Tikhonov regularization)

\[ f = \begin{bmatrix} B_k \\ \lambda_k I \end{bmatrix} \backslash \begin{bmatrix} W_k b \\ 0 \end{bmatrix} \]
\[ x_k = Z_k f \]

\[ \lambda = 0.0050 \]
Example: Inverse Heat Equation

LSQR (no regularization)

\[ f = B_k \backslash W_k b \]
\[ x_k = Z_k f \]

HyBR (Tikhonov regularization)

\[ f = \begin{bmatrix} B_k \\ \lambda_k I \end{bmatrix} \backslash \begin{bmatrix} W_k b \\ 0 \end{bmatrix} \]
\[ x_k = Z_k f \]

\[ \lambda = 0.0042 \]

iteration = 35
The Nonlinear Problem

- We want to find \(x\) and \(y\) so that

\[
b = A(y)x + e
\]

- With Tikhonov regularization, solve

\[
\min_{x,y} \left\| \begin{bmatrix} A(y) \\ \lambda I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|^2_2
\]

- As with linear problem, choosing a good regularization parameter \(\lambda\) is important.
- Problem is linear in \(x\), nonlinear in \(y\).
- \(y \in \mathcal{R}^p, \ x \in \mathcal{R}^n\), with \(p \ll n\).
Variable Projection Method:

- Implicitly eliminate linear term.
- Optimize over nonlinear term.

Some general references:

- Golub and Pereyra, SINUM 1973 (also IP 2003)
- Kaufman, BIT 1975
- Osborne, SINUM 1975 (also ETNA 2007)
- Ruhe and Wedin, SIREV, 1980

How to apply to inverse problems?
Variable Projection Method

Instead of optimizing over both $x$ and $y$:

$$\min_{x,y} \phi(x, y) = \min_{x,y} \left\| \begin{bmatrix} A(y) \\ \lambda I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2^2$$

Let $x(y)$ be solution of

$$\min_x \phi(x, y) = \min_x \left\| \begin{bmatrix} A(y) \\ \lambda I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2^2$$

and then minimize the reduced cost functional:

$$\min_y \psi(y), \quad \psi(y) = \phi(x(y), y)$$
The Linear Problem: \( b = Ax + e \)

The Nonlinear Problem: \( b = A(y) x + e \)

Example: Image Deblurring

Concluding Remarks

Gauss-Newton Algorithm

Choose initial \( y_0 \)

For \( k = 0, 1, 2, \ldots \)

\[
\begin{align*}
  x_k &= \arg \min_x \left\| \begin{bmatrix} A(y_k) & \lambda_k I \end{bmatrix} x - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|_2 \\
  r_k &= b - A(y_k) x_k \\
  d_k &= \arg \min_d \| J_{\psi} d - r_k \|_2 \\
  y_{k+1} &= y_k + d_k 
\end{align*}
\]

End
Gauss-Newton Algorithm with HyBR

And we use HyBR to solve the linear subproblem:

choose initial $y_0$
for $k = 0, 1, 2, \ldots$

$$x_k = \text{HyBR}(A(y_k), b)$$

$$r_k = b - A(y_k) x_k$$

$$d_k = \arg \min_d \| J_\psi d - r_k \|_2$$

$$y_{k+1} = y_k + d_k$$
end
Example: Image Deblurring

Matrix $A(y)$ is defined by a PSF, which is in turn defined by parameters. Specifically:

$$A(y) = A(P(y))$$

where

- $A$ is $65536 \times 65536$, with entries given by $P$.
- $P$ is $256 \times 256$, with entries:

$$p_{ij} = \exp \left( \frac{(i - k)^2 s_2^2 - (j - l)^2 s_1^2 + 2(i - k)(j - l)\rho^2}{2s_1^2 s_2^2 - 2\rho^4} \right)$$

- $(k, l)$ is the PSF center (location of point source)
- $y$ vector of unknown parameters:

$$y = \begin{bmatrix} s_1 \\ s_2 \\ \rho \end{bmatrix}$$
Can get analytical formula for Jacobian:

\[
J_\psi = \frac{\partial}{\partial y} \left\{ A(P(y)) x \right\} = \frac{\partial}{\partial P} \left\{ A(P(y)) x \right\} \cdot \frac{\partial}{\partial y} \left\{ P(y) \right\} = A(X) \cdot \frac{\partial}{\partial y} \left\{ P(y) \right\}
\]

where \( x = \text{vec}(X) \).

Though in this example, finite difference approximation of \( J_\psi \) works very well.
Example: Image Deblurring

Gauss-Newton Iteration History

<table>
<thead>
<tr>
<th>G-N Iteration</th>
<th>$\Delta y$</th>
<th>$\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.5716</td>
<td>0.1685</td>
</tr>
<tr>
<td>1</td>
<td>0.3345</td>
<td>0.1223</td>
</tr>
<tr>
<td>2</td>
<td>0.2192</td>
<td>0.0985</td>
</tr>
<tr>
<td>3</td>
<td>0.1473</td>
<td>0.0804</td>
</tr>
<tr>
<td>4</td>
<td>0.1006</td>
<td>0.0715</td>
</tr>
<tr>
<td>5</td>
<td>0.0648</td>
<td>0.0676</td>
</tr>
<tr>
<td>6</td>
<td>0.0355</td>
<td>0.0657</td>
</tr>
<tr>
<td>7</td>
<td>0.0144</td>
<td>0.0650</td>
</tr>
</tbody>
</table>
Example: Image Deblurring

Observed Image

Reconstruction using initial PSF

Reconstruction after 8 GN iterations
Imaging applications require solving challenging inverse problems.

Separable nonlinear least squares models exploit high level structure.

Hybrid methods are efficient solvers for large scale linear inverse problems.
  - Automatic estimation of regularization parameter.
  - Automatic estimation of stopping iteration.

Hybrid methods can be effective linear solvers for nonlinear problems.
Questions?

- Other methods to choose regularization parameters?
- Other regularization methods (e.g., total variation)?
- Sparse (in some basis) reconstructions?
- MATLAB Codes and Data?

  www.mathcs.emory.edu/~nagy/WGCV
  www.mathcs.emory.edu/~nagy/RestoreTools
  www2.imm.dtu.dk/~pch/HNO
  www2.imm.dtu.dk/~pch/Regutools