# Efficient Numerical Methods for Least-Norm Regularization

D.C. Sorensen

- Collaborator: M. Rojas
- Support: AFOSR and NSF

AMS Meeting  Lexington, KY  March 2010

# *Least Norm Regularization*

$$\min_{\mathbf{x}} \|\mathbf{x}\|, \quad s.t. \ \ \|\mathbf{b} - \mathbf{A}\mathbf{x}\| \le \epsilon$$

- **KKT Conditions and the Secular Equation**
- **LNR: Newton's Method for Dense Problems**
- **Re-formulation of KKT for Large Scale Problems**
- **Nonlinear Lanczos: LNR_NLLr & LNR_NLLx**
- **Pre-Conditioning: Newton-Like Iterations**
- **Computational Results**

# *Brief Background*

- Standard Trust Region Problem (TRS):
  $\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$ s.t. $\|\mathbf{x}\| \leq \Delta$.
- Secular Equation:
  Hebden('73), Moré('77), Morozov('84)
- TRS:
  Eldén ('77), Gander('81), Golub & von Mat ('91)
- Large Scale TRS:
  S.('97), Hager('01), Rendl & Wolkowicz ('01),
  Reichel et. al., Rojas & S. ('02), Rojas Santos & S. ('08)
- Nonlinear Iterations:
  Voss ('04): Non-Linear Arnoldi/Lanczos,
  Lampe, Voss, Rojas & S.('09) Improved LSTRS

# The KKT Conditions

Underlying Problem : $\min_{\mathbf{x}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|$

Assumption: All error is measurement error in R.H.S.

**b** is perturbation of exact data,

$$\mathbf{b} = \mathbf{b}_o + \overline{n} \ \text{ with } \ \mathbf{A}\mathbf{x}_o = \mathbf{b}_o, \ \ \epsilon \geq \|\overline{n}\|.$$

Assures solution $\mathbf{x}_o$ is feasible .

Lagrangian :

$$\mathcal{L} := \|\mathbf{x}\|^2 + \lambda(\|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 - \epsilon^2).$$

KKT conditions :

$$\mathbf{x} + \lambda \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{0}, \ \ \lambda(\|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 - \epsilon^2) = 0, \ \ \lambda \geq 0.$$

# *Positive $\lambda$ KKT Conditions*

Some Observations:

- $\|\mathbf{b}\| \le \epsilon \Leftrightarrow \mathbf{x} = \mathbf{0}$ is a solution,
- $\lambda = 0 \Rightarrow \mathbf{x} = \mathbf{0}$,
- $\lambda > 0 \Leftrightarrow \mathbf{x} \neq \mathbf{0}$ and $\|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \epsilon^2$.

KKT conditions with positive $\lambda$ :

$$\mathbf{x} + \lambda \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{0}, \quad \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \epsilon^2, \quad \lambda > 0.$$

**KKT - Necessary and Sufficient**

# Optimality Conditions: SVD version

Let $\mathbf{A} = \mathbf{USV}^T$ (short form SVD)
Let $\mathbf{b} = \mathbf{Ub}_1 + \mathbf{b}_2$ with $\mathbf{U}^T\mathbf{b}_2 = \mathbf{0}$.
Then

$$\|\mathbf{b} - \mathbf{Ax}\|^2 \le \epsilon^2 \iff \|\mathbf{b}_1 - \mathbf{SV}^T\mathbf{x}\|^2 \le \epsilon^2 - \|\mathbf{b}_2\|^2 =: \delta^2,$$

Must assume $\mathbf{b} = \mathbf{Ub}_1 + \mathbf{b}_2$ with $\|\mathbf{b}_2\| \le \epsilon$
($\|\mathbf{b}_2\| > \epsilon \Rightarrow$ no feasible point).

$$\mathbf{x} + \lambda\mathbf{VS}(\mathbf{SV}^T\mathbf{x} - \mathbf{b}_1) = \mathbf{0}, \quad \|\mathbf{b}_1 - \mathbf{SV}^T\mathbf{x}\|^2 = \delta^2, \quad \lambda > 0.$$

Manipulate KKT into more useful form:

$$(\mathbf{I} + \lambda\mathbf{S}^2)\mathbf{z} = \mathbf{b}_1, \quad \|\mathbf{z}\|^2 \le \delta^2, \quad \lambda > 0.$$

$$\mathbf{x} = \lambda\mathbf{VSz} \;\; \textit{with} \;\; \mathbf{z} := \mathbf{b}_1 - \mathbf{SV}^T\mathbf{x}.$$

# The Dense LNR Scheme

- Compute $\mathbf{A} = \mathbf{U}\mathbf{S}\mathbf{V}^T$;
- Put $\mathbf{b}_1 = \mathbf{U}^T\mathbf{b}$;
- Set $\delta^2 = \epsilon^2 - \|\mathbf{b} - \mathbf{U}\mathbf{b}_1\|^2$;
- Compute $\lambda \geq 0$ and $\mathbf{z}$ s.t.

$$(\mathbf{I} + \lambda\mathbf{S}^2)\mathbf{z} = \mathbf{b}_1, \quad \|\mathbf{z}\|^2 \leq \delta^2;$$

- Put $\mathbf{x} = \lambda\mathbf{V}\mathbf{S}\mathbf{z}$.

**Step 4 requires a solver ...**

# The Secular Equation - Newton's Method

**How to compute $\lambda$:**

We use Newton's Method to solve $\psi(\lambda) = 0$ where

$$\psi(\lambda) := \frac{1}{\|\mathbf{z}_\lambda\|} - \frac{1}{\delta}, \quad \text{where} \quad \mathbf{z}_\lambda := (\mathbf{I} + \lambda \mathbf{S}^2)^{-1}\mathbf{b}_1.$$

Initial Guess -

$$\lambda_1 := \frac{\|\mathbf{b}_1\| - \delta}{\delta\sigma_1^2} < \lambda_o.$$

Note: With $r := \mathrm{rank}(\mathbf{A}) \le n$,

$$\mathbf{z}_\lambda^T \mathbf{z}_\lambda = \sum_{j=1}^{r} \frac{\beta_j^2}{(1 + \lambda\sigma_j^2)^2} + \beta_o^2,$$

$\beta_o^2 := \sum_{j=r+1}^{n} \beta_j^2.$ poles at $-1/\sigma_j^2$ : no problem

# *The Secular Equation*

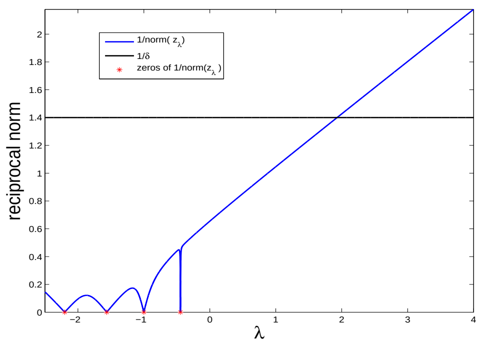



Figure: Graph of Typical Secular Equation

- $\psi(\lambda)$ is concave and monotone increasing for $\lambda \in (0, \infty)$,
- $\psi(\lambda) = 0$ has a unique root at $\lambda = \lambda_o > 0$.
- Newton converges - No safeguarding required

# *Computational Results, Dense LNR*

| Problem | Iter | Time | $\frac{\|\|\mathbf{x}-\mathbf{x}_*\|\|}{\|\|\mathbf{x}_*\|\|}$ |
|---|---|---|---|
| **baart** | 12 | 57.04 | 5.33e-02 |
| **deriv2, ex. 1** | 9 | 57.18 | 6.90e-02 |
| **deriv2, ex. 2** | 9 | 57.93 | 6.59e-02 |
| **foxgood** | 11 | 59.91 | 1.96e-03 |
| **i_laplace, ex. 1** | 12 | 23.04 | 1.67e-01 |
| **i_laplace, ex. 3** | 11 | 22.88 | 1.96e-03 |
| **heat, mild** | 4 | 60.96 | 1.13e-03 |
| **heat, severe** | 9 | 40.27 | 6.95e-03 |
| **phillips** | 9 | 46.97 | 1.32e-03 |
| **shaw** | 11 | 57.25 | 3.14e-02 |

Table: LNR on Regularization Tools problems, $m = n = 1024$.

# KKT For Large Scale Problems

Original Form KKT:

$$\mathbf{x} + \lambda \mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = \mathbf{0}, \;\; \|\mathbf{b} - \mathbf{A}\mathbf{x}\|^2 = \epsilon^2, \;\; \lambda > 0.$$

Solution Space Equations :

$$(\mathbf{I} + \lambda \mathbf{A}^T\mathbf{A})\mathbf{x} = \lambda \mathbf{A}^T\mathbf{b}.$$

Residual Space Equations :
Multiply on left by $\mathbf{A}$ and add $-\mathbf{b}$ to both sides gives:

$$\mathbf{A}\mathbf{x} - \mathbf{b} + \lambda \mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = -\mathbf{b}.$$

Put $\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}$ and adjust $\lambda$ to obtain

$$(\mathbf{I} + \lambda \mathbf{A}\mathbf{A}^T)\mathbf{r} = \mathbf{b}, \quad \|\mathbf{r}_\lambda\| = \epsilon.$$

Set $\mathbf{x}_\lambda = \lambda \mathbf{A}^T\mathbf{r}$.

# *Projected Equations*

Large Scale Framework (J-D Like):

- Build a Search Space $\mathcal{S} = \mathrm{Range}(\mathbf{V})$
- Solve a projected problem restricted to $\mathcal{S}$
- Adjoin new descent direction $\mathbf{v}$ to search space

$$\mathbf{V} \leftarrow [\mathbf{V}, \mathbf{v}]; \ \ \mathcal{S} \leftarrow \mathrm{Range}(\mathbf{V})$$

Solution Space Equations :

Put $\mathbf{x} = \mathbf{V}\hat{\mathbf{x}}$ and multiply on left by $\mathbf{V}^T$

$$(\mathbf{I} + \lambda(\mathbf{AV})^T(\mathbf{AV}))\hat{\mathbf{x}} = \lambda(\mathbf{AV})^T\mathbf{b}.$$

Residual Space Equations :

Put $\mathbf{r} = \mathbf{V}\hat{\mathbf{r}}$ and multiply on left by $\mathbf{V}^T$

$$(\mathbf{I} + \lambda(\mathbf{V}^T\mathbf{A})(\mathbf{V}^T\mathbf{A})^T)\hat{\mathbf{r}} = \mathbf{V}^T\mathbf{b}, \quad \|\mathbf{r}_\lambda\| = \epsilon.$$

Set $\mathbf{x}_\lambda = \lambda\mathbf{A}^T(\mathbf{V}\hat{\mathbf{r}})$.

# *Secular Equation for Projected Equations*

In both **r** and **x** iterations the Secular Equation is

$$\|(\mathbf{I} + \lambda \mathbf{S}^2)^{-1}\mathbf{b}_1\| = \delta$$

Can use Secular Equation Solver from dense LNR
Both Cases: $\mathbf{b}_1 = \mathbf{W}^T\mathbf{V}^T\mathbf{b}$

- **x** - iteration: $\mathbf{WSU}^T = \mathbf{AV}$
- **r** - iteration: $\mathbf{WSU}^T = \mathbf{V}^T\mathbf{A}$

# *Nonlinear Lanczos* **r**-*Iteration*

Repeat until convergence:

- Put $\mathbf{r} = \mathbf{V}\hat{\mathbf{r}}$ and express $\mathbf{b} = \mathbf{V}\hat{\mathbf{b}} + \mathbf{f}$ with $\mathbf{V}^T\mathbf{f} = 0$ .
- Take $\mathbf{WSU}^T = \mathbf{V}^T\mathbf{A}$ (short-form SVD)
- Solve Secular Equation

$$\|(\mathbf{I} + \lambda\mathbf{S}^2)^{-1}\mathbf{b}_1\| = \epsilon \text{ with } \mathbf{b}_1 = \mathbf{W}^T\hat{\mathbf{b}}.$$

- Put $\mathbf{x}_\lambda = \lambda\mathbf{A}^T\mathbf{V}\hat{\mathbf{r}} = \lambda\mathbf{US}(\mathbf{W}^T\hat{\mathbf{r}}) = \mathbf{USz}\lambda,$
  where $\mathbf{z} := \mathbf{W}^T\hat{\mathbf{r}} = (\mathbf{I} + \lambda\mathbf{S}^2)^{-1}\mathbf{b}_1$.
- *Nonlinear Lanczos Step*:
  Get new search direction $\mathbf{v} = (\mathbf{I} - \mathbf{VV}^T)(\mathbf{b} - \mathbf{Ax}_\lambda)$
  Set $\mathbf{v} \leftarrow \mathbf{v}/\|\mathbf{v}\|$
  Update basis $\mathbf{V} \leftarrow [\mathbf{V}, \mathbf{v}]$

# *Nonlinear Lanczos* **x***-Iteration*

Repeat until convergence:

- Compute $\mathbf{WSU}^T = \mathbf{AV}$ (short form SVD)
  Express $\mathbf{b} = \mathbf{W}\hat{\mathbf{b}} + \mathbf{f}$ with $\mathbf{W}^T\mathbf{f} = 0$
  Set $\delta = \sqrt{\epsilon^2 - \mathbf{f}^T\mathbf{f}}$.
- Solve Secular Equation

$$\|(\mathbf{I} + \lambda \mathbf{S}^2)^{-1}\mathbf{b}_1\| = \delta \quad \text{with} \quad \mathbf{b}_1 = \mathbf{W}^T\hat{\mathbf{b}}.$$

- Put $\mathbf{x}_\lambda = \lambda \mathbf{V}(\mathbf{USz})$ where $\mathbf{z} = (\mathbf{I} + \lambda \mathbf{S}^2)^{-1}\mathbf{b}_1$.
- *Nonlinear Lanczos Step* :
  Compute $\mathbf{r} = \mathbf{b} - \mathbf{Ax}_\lambda$
  Obtain search direction $\mathbf{v} = (\mathbf{I} - \mathbf{VV}^T)(\lambda \mathbf{A}^T\mathbf{r})$
  Normalize $\mathbf{v} \leftarrow \mathbf{v}/\|\mathbf{v}\|$
  Update the basis $\mathbf{V} \leftarrow [\mathbf{V}, \mathbf{v}]$

# *Analysis of Local Minimization Step*

KKT: $(\mathbf{I} + \lambda \mathbf{A}\mathbf{A}^T)\mathbf{r} = \mathbf{b}$ with $\|\mathbf{r}\| = \epsilon$.
Given $\lambda$,

$$\min_{\mathbf{r}} \{\frac{1}{2}\mathbf{r}^T(\mathbf{I} + \lambda \mathbf{A}\mathbf{A}^T)\mathbf{r} - \mathbf{b}^T\mathbf{r}\} \equiv \min_{\mathbf{r}} \varphi(\mathbf{r}, \lambda),$$

Steepest Descent Direction:

$$\mathbf{s} = -\nabla_{\mathbf{r}}\varphi(\mathbf{r}, \lambda) = -[(\mathbf{I} + \lambda \mathbf{A}\mathbf{A}^T)\mathbf{r} - \mathbf{b}] = (\mathbf{b} - \mathbf{A}\mathbf{x} - \mathbf{r}).$$

$$\hat{\mathbf{v}} = (\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{s} = (\mathbf{I} - \mathbf{V}\mathbf{V}^T)(\mathbf{b} - \mathbf{A}\mathbf{x}),$$

Since $\mathbf{r} = \mathbf{V}\hat{\mathbf{r}}$ implies $(\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{r} = \mathbf{0}$.

LNR_NLLr step adjoins full steepest descent direction
Adjoin $\mathbf{v} = \hat{\mathbf{v}}/\|\hat{\mathbf{v}}\|$ to search space $\mathcal{S}_+ = \mathrm{Range}([\mathbf{V}, \mathbf{v}])$
Note: $\mathcal{S}_+$ contains $\min \varphi$ along the steepest descent direction.
Next iterate: Decrease at least as good as steepest descent.

# *Pre-Conditioning: Newton-Like Iterartion*

General Descent Direction:

$$\mathbf{s} = -\mathbf{M}[(\mathbf{I} + \lambda\mathbf{A}\mathbf{A}^T)\mathbf{r} - \mathbf{b}] = \mathbf{M}(\mathbf{b} - \mathbf{A}\mathbf{x} - \mathbf{r}),$$

$\mathbf{M}$ is S.P.D. and $\mathbf{x} = \lambda\mathbf{A}^T\mathbf{r}$.
Orthogonal decomposition (noting $\mathbf{r} = \mathbf{V}\hat{\mathbf{r}}$) will give

$$\mathbf{b} - (\mathbf{I} + \lambda\mathbf{A}\mathbf{A}^T)\mathbf{r} = (\mathbf{I} - \mathbf{V}\mathbf{V}^T)[\mathbf{b} - (\mathbf{I} + \lambda\mathbf{A}\mathbf{A}^T)\mathbf{r}] = (\mathbf{I} - \mathbf{V}\mathbf{V}^T)(\mathbf{b} - \mathbf{A}\mathbf{x}).$$

Thus, orthogonalizing $\mathbf{s}$ against $\mathrm{Range}(\mathbf{V})$ and normalizing gives:

$$\hat{\mathbf{v}} = (\mathbf{I} - \mathbf{V}\mathbf{V}^T)\mathbf{M}(\mathbf{I} - \mathbf{V}\mathbf{V}^T)(\mathbf{b} - \mathbf{A}\mathbf{x}), \quad \mathbf{v} = \hat{\mathbf{v}}/\|\hat{\mathbf{v}}\|,$$

The full pre-conditioned or Newton-like step is adjoined .
Adjoin $\mathbf{v} = \mathbf{s}/\|\mathbf{s}\|$ to search space $\mathcal{S}_+ = \mathrm{Range}([\mathbf{V}, \mathbf{v}])$
Next iterate: Decrease at least as good as Newton-like step.

## *What if* $\hat{\mathbf{v}} = 0$ *?*

Iteration Terminates with Solution

$$0 = (\mathbf{b} - \mathbf{Ax})^T \hat{\mathbf{v}} = (\mathbf{b} - \mathbf{Ax})^T (\mathbf{I} - \mathbf{VV}^T)\mathbf{M}(\mathbf{I} - \mathbf{VV}^T)(\mathbf{b} - \mathbf{Ax}).$$

$\mathbf{M}$ S.P.D. $\Rightarrow$ $\mathbf{0} = (\mathbf{I} - \mathbf{VV}^T)(\mathbf{b} - \mathbf{Ax})$ $\Rightarrow$ $\mathbf{b} - \mathbf{Ax} = \mathbf{Vz}$ for some $\mathbf{z}$

$\mathbf{z} = \mathbf{V}^T(\mathbf{b} - \mathbf{Ax}) = \hat{\mathbf{b}} - \lambda \mathbf{V}^T\mathbf{AA}^T\mathbf{r} = (\mathbf{I} + \lambda \mathbf{V}^T\mathbf{AA}^T\mathbf{V})\hat{\mathbf{r}} - \lambda \mathbf{V}^T\mathbf{AA}^T\mathbf{V}\hat{\mathbf{r}} = \hat{\mathbf{r}}.$

Substitute $\mathbf{z} = \hat{\mathbf{r}}$ to get $\mathbf{b} - \mathbf{Ax} = \mathbf{r}$

$\|\mathbf{r}\| = \epsilon \Rightarrow$ KKT conditions satisfied $\Rightarrow \mathbf{x} = \lambda \mathbf{A}^T\mathbf{r}$ is solution

**x** - iteration LNR_NLLx has analogous properties

# *Computational Results, LNR_NLLr*

| Problem | OuIt | InIt | MVP | Time | Vec | $\frac{\|\|\mathbf{x}-\mathbf{x}_{LNR}\|\|}{\|\|\mathbf{x}_{LNR}\|\|}$ | $\frac{\|\|\mathbf{x}-\mathbf{x}_{*}\|\|}{\|\|\mathbf{x}_{*}\|\|}$ |
|---|---|---|---|---|---|---|---|
| **baart** | 1 | 12.0 | 35 | 0.09 | 12 | 1.5e-11 | 5.3e-02 |
| **deriv2, ex. 1** | 33 | 3.4 | 99 | 0.44 | 44 | 7.8e-03 | 7.0e-02 |
| **deriv2, ex. 2** | 31 | 3.4 | 95 | 0.40 | 42 | 8.2e-03 | 6.6e-02 |
| **foxgood** | 1 | 11.0 | 35 | 0.09 | 12 | 5.3e-13 | 2.0e-03 |
| **i_laplace, ex. 1** | 7 | 4.0 | 47 | 0.16 | 18 | 2.2e-02 | 1.7e-01 |
| **i_laplace, ex. 3** | 4 | 4.0 | 41 | 0.13 | 15 | 2.6e-03 | 3.2e-03 |
| **heat, mild** | 25 | 2.0 | 83 | 0.33 | 36 | 1.2e-03 | 5.5e-04 |
| **heat, severe** | 29 | 3.1 | 91 | 0.37 | 40 | 2.3e-03 | 7.5e-03 |
| **phillips** | 5 | 4.2 | 43 | 0.11 | 16 | 9.4e-04 | 1.4e-03 |
| **shaw** | 1 | 11.0 | 35 | 0.09 | 12 | 2.3e-09 | 3.1e-02 |

Table: LNR_NLLr on Regularization Tools problems, $m = n = 1024$.

# *Computational Results, LNR_NLLx*

| Problem | OuIt | InIt | MVP | Time | Vec | $\frac{\|\|\mathbf{x}-\mathbf{x}_{LNR}\|\|}{\|\|\mathbf{x}_{LNR}\|\|}$ | $\frac{\|\|\mathbf{x}-\mathbf{x}_{*}\|\|}{\|\|\mathbf{x}_{*}\|\|}$ |
|---|---|---|---|---|---|---|---|
| **baart** | 1 | 7.0 | 67 | 0.16 | 22 | 4.2e-11 | 5.3e-02 |
| **deriv2, ex. 1** | 17 | 3.3 | 115 | 0.32 | 38 | 1.6e-02 | 7.5e-02 |
| **deriv2, ex. 2** | 16 | 3.3 | 112 | 0.30 | 37 | 1.5e-02 | 7.1e-02 |
| **foxgood** | 1 | 6.0 | 67 | 0.16 | 22 | 3.7e-10 | 1.9e-03 |
| **i_laplace, ex. 1** | 1 | 7.0 | 67 | 0.20 | 22 | 1.7e-06 | 1.7e-01 |
| **i_laplace, ex. 3** | 1 | 6.0 | 67 | 0.20 | 22 | 2.7e-08 | 1.9e-03 |
| **heat, mild** | 27 | 2.0 | 145 | 0.48 | 48 | 1.1e-03 | 4.1e-04 |
| **heat, severe** | 15 | 3.1 | 109 | 0.29 | 36 | 3.5e-03 | 9.1e-03 |
| **phillips** | 1 | 5.0 | 67 | 0.16 | 22 | 2.9e-05 | 1.3e-03 |
| **shaw** | 1 | 7.0 | 67 | 0.16 | 22 | 1.7e-12 | 3.1e-02 |

Table: LNR_NLLx on Regularization Tools problems, $m = n = 1024$.

# *Comparison: LNR_NLLr& LNR_NLLx(time and storage)*

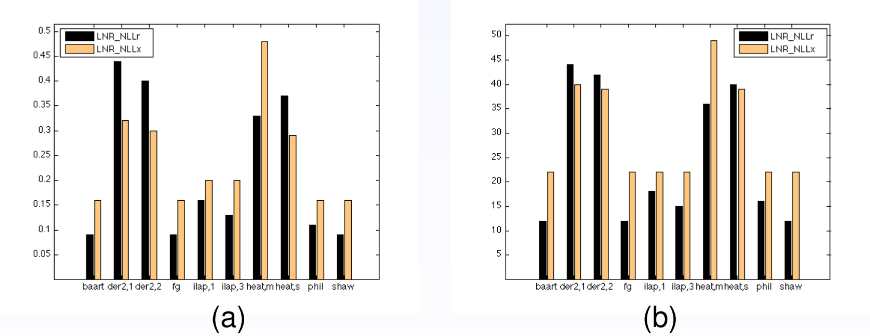



Figure: Time (a) and number of vectors (b) required by LNR_NLLr (dark) and LNR_NLLx (clear), $m = n = 1024$.

# Performance on Rectangular Matrices

Problem **heat, mild**

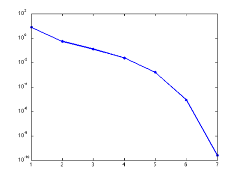
| Method / $m \times n$ | | OutIt | InIt | MVP | Time | Vec | $\frac{\|\|\mathbf{x}-\mathbf{x}_*\|\|}{\|\|\mathbf{x}_*\|\|}$ |
|---|---|---|---|---|---|---|---|
| LNR | $1024 \times 300$ | 7 | – | – | 0.29 | 300 | 5.21e-03 |
| LNR_NLLr | $1024 \times 300$ | 38 | 2.8 | 109 | 0.22 | 49 | 5.22e-03 |
| LNR_NLLx | $1024 \times 300$ | 22 | 3.1 | 130 | 0.21 | 43 | 5.99e-03 |
| LNR | $300 \times 1024$ | 7 | – | – | 0.30 | 1024 | 5.18e-03 |
| LNR_NLLr | $300 \times 1024$ | 37 | 2.8 | 107 | 0.35 | 48 | 5.11e-03 |
| LNR_NLLx | $300 \times 1024$ | 21 | 3.1 | 127 | 0.17 | 42 | 5.76e-03 |

Table: Performance of LNR, LNR_NLLr, and LNR_NLLx

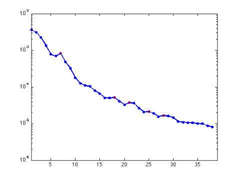
# *Convergence History on Problem* **heat, mild**

Iteration no. vs $|\psi(\lambda_k)|$ (dense) and $\|\mathbf{b} - \mathbf{A}\mathbf{x}_k\|$ (sparse)

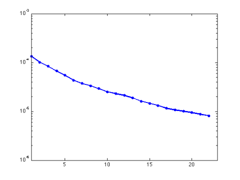
$m = 1024$, $n = 300$




(LNR )

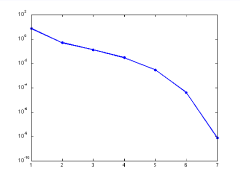



(LNR_NLLr)

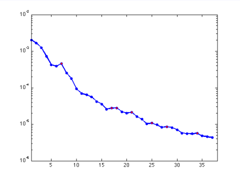



( LNR_NLLx)

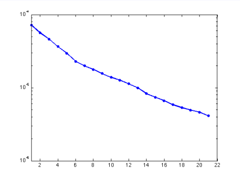
$m = 300$, $n = 1024$




(LNR )




(LNR_NLLr)




( LNR_NLLx)

# *Image Restoration*

- Recover an Image from Blurred and Noisy Data
- Digital Photo was Blurred using **blur** from Hansen
- Data Vector **b**: Blurred and Noisy Image, a one-D array
- Noise Level in $\mathbf{b} = \mathbf{b}_o + \overline{n}$ was $\|\overline{n}\|/\|\mathbf{b}_o\| = 10^{-2}$
- Original photograph: $256 \times 256$ pixels, $n = 65536$.
- **A** = Blurring Operator returned by **blur**.

| Method / noise level | OutIt | InIt | MVP | Time | Vec | $\frac{\|\|\mathbf{x}-\mathbf{x}_*\|\|}{\|\|\mathbf{x}_*\|\|}$ |
|---|---|---|---|---|---|---|
| LNR_NLLr / $10^{-2}$ | 1 | 4.0 | 35 | 0.79 | 12 | 1.08e-01 |
| LNR_NLLx / $10^{-2}$ | 1 | 4.0 | 67 | 2.10 | 22 | 1.08e-01 |
| LNR_NLLr / $10^{-3}$ | 41 | 3.0 | 115 | 46.67 | 52 | 7.13e-02 |
| LNR_NLLx / $10^{-3}$ | 6 | 3.0 | 82 | 3.87 | 27 | 7.86e-02 |

Table: Performance LNR_NLLr , LNR_NLLx on Image Restoration.

# *Image Restoration: Paris Art, $n = 65536$*

True image

Blurred and noisy image

LNR_NLLr restoration

LNR_NLLx restoration

# *Summary*

- **CAAM TR10-08** Efficient Numerical Methods for Least-Norm Regularization, D.C. Sorensen and M. Rojas
- **TR09-26** Accelerating the LSTRS Algorithm, J. Lampe, M. Rojas, D.C. Sorensen, and H. Voss
- http://www.caam.rice.edu/ sorensen/

**Least Norm Regularization:** $\min_{\mathbf{x}} \|\mathbf{x}\|, \;\; s.t. \;\; \|\mathbf{b} - \mathbf{A}\mathbf{x}\| \leq \epsilon$

- KKT Conditions and the Secular Equation
- Newton's Method for Dense Problems
- Re-formulation of KKT for Large Scale Problems
- Nonlinear Lanczos: LNR_NLLr & LNR_NLLx
- Pre-Conditioning: Newton-Like Iterations
- Computational Results