Iterative Methods for Ill-Posed Problems

Based on joint work with:

James Baglama

Serena Morigi

Fiorella Sgallari

Andriy Shyshkov

Qiang Ye

Fabiana Zama

Bologna, settembre, 2006

Noise-free problem - infinite dimensional

Solve

$$Ax = \hat{b} \qquad (1)$$

with solution $\hat{x}$, where

$A : \mathcal{X} \longrightarrow \mathcal{X}$ bounded linear operator, A^{-1} unbounded,
$\mathcal{X}$ Hilbert space, $\hat{b} \in \mathcal{R}(A)$.

Small perturbations in $\hat{b}$ can result in arbitrarily large perturbations in $\hat{x}$.

The solution of (1) is an ill-posed problem.

Noise-free problem - finite dimensional

Solve

$$Ax = \hat{b} \qquad\qquad (1')$$

with solution $\hat{x}$, where

A $n \times n$ matrix of ill-determined rank, possibly singular, A^{-1} does not exist or has huge entries, $\hat{b} \in \mathcal{R}(A)$.

Small perturbations in $\hat{b}$ can result in large perturbations in $\hat{x}$.

The solution of $(1')$ is an ill-conditioned problem.

Noisy problem - finite or infinite dimensional

Available equation

$$Ax = b, \qquad (2)$$

where

$b = \hat{b} + e$, e=noise,
$\|e\| = \delta$ is assumed known,
eq. (2) might not be consistent.

Determine approximation of $\hat{x}$ from (2).

Computed example: Fredholm integral equation of the first kind

$$\int_0^{\pi} \exp(-st)x(t)dt = 2\frac{\sinh(s)}{s}, \quad 0 \leq s \leq \frac{\pi}{2}$$

Determine solution $x(t) = \sin(t)$.

Discretize integral by Galerkin method using piecewise constant functions. Code baart from Regularization Tools.

This gives a linear system of equations

$$Ax = \hat{b}, \qquad A \in \mathrm{R}^{200\times 200}, \quad \hat{b} \in \mathrm{R}^{200}.$$

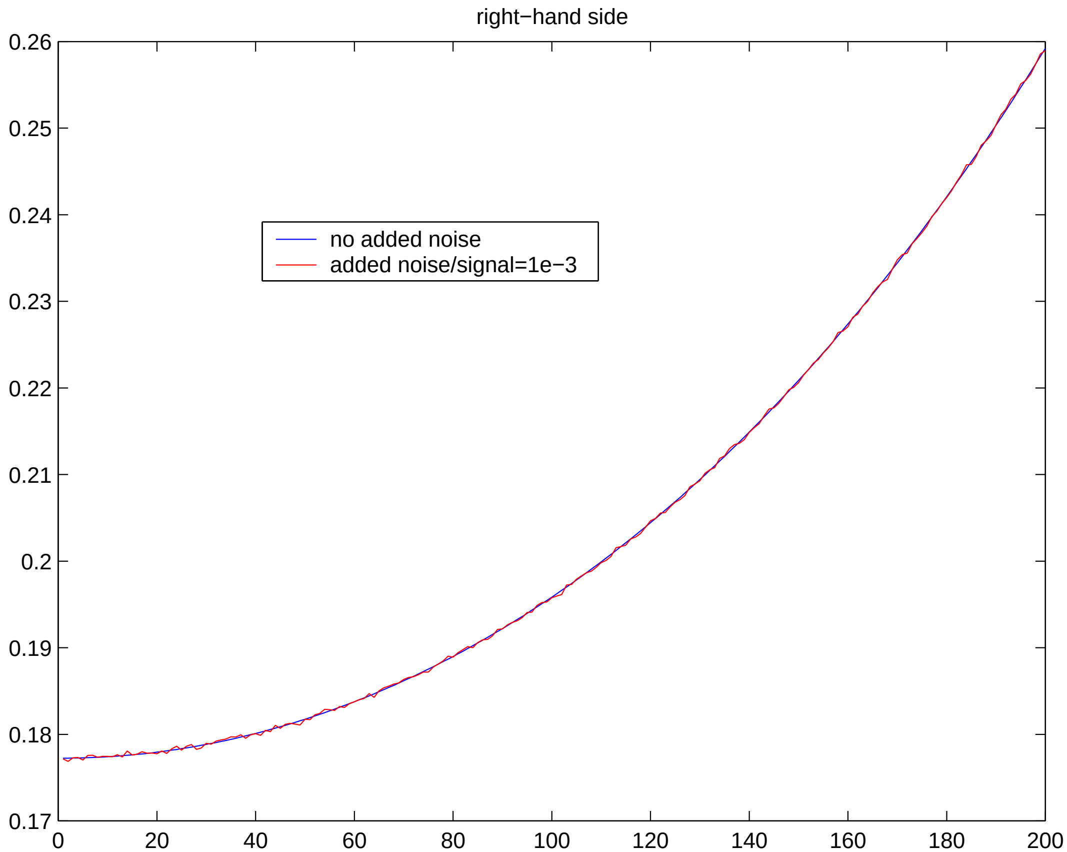
A is numerically singular

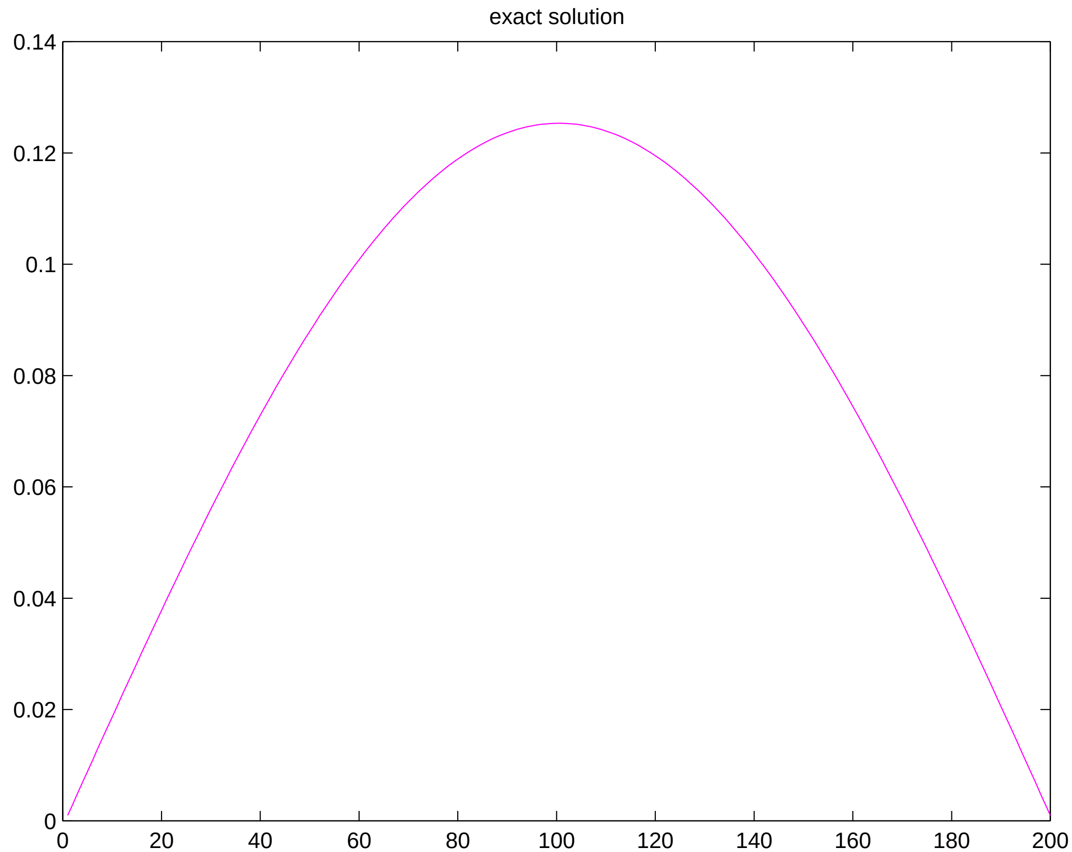
Let the "noise" vector e in b have normally distributed entries with zero mean and

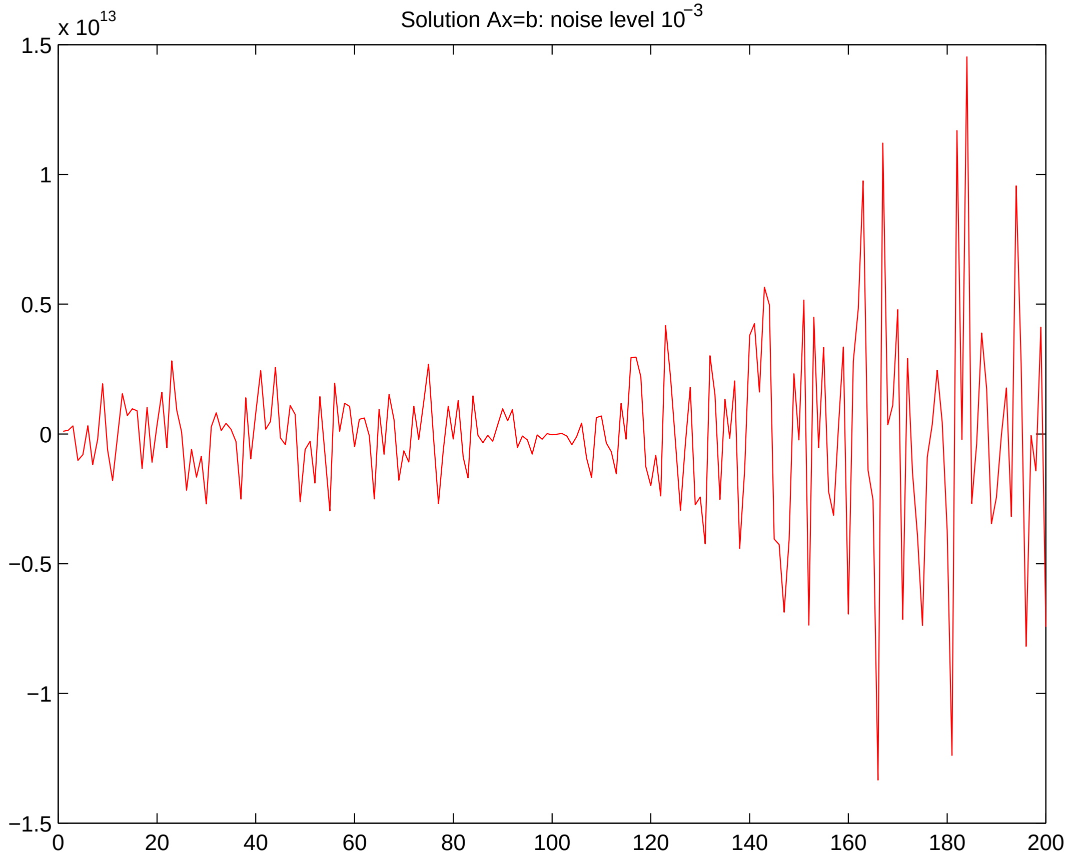
$$\delta = \|e\| = 10^{-3}\|b\|$$

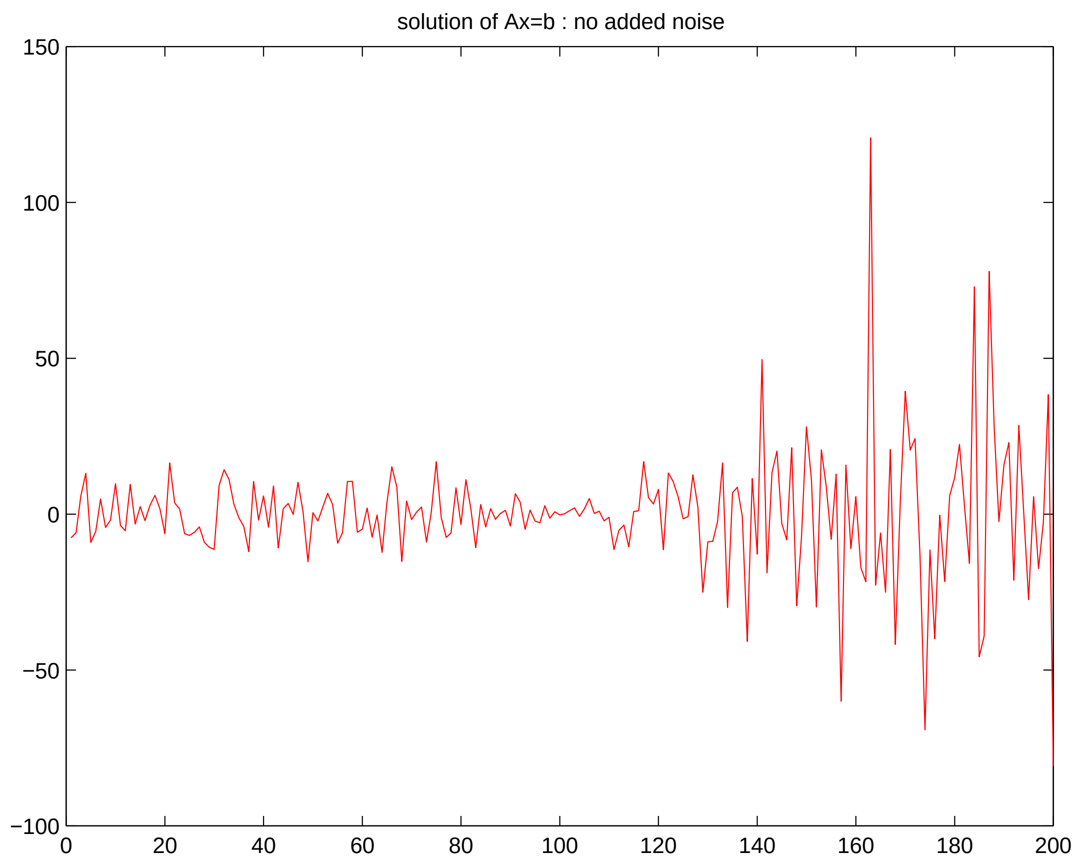
$$b := \hat{b} + e$$

i.e., 0.1% relative noise


right−hand side
0.26
0.25
0.24
0.23
0.22
0.21
0.2
0.19
0.18
0.17
0
20
40
60
80
100
120
140
160
180
200
no added noise
added noise/signal=1e−3


exact solution
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
0 20 40 60 80 100 120 140 160 180 200


Solution Ax=b: noise level 10−3
x 10^13
1.5
1
0.5
0
−0.5
−1
−1.5
0
20
40
60
80
100
120
140
160
180
200


solution of Ax=b : no added noise
150
100
50
0
−50
−100
0
20
40
60
80
100
120
140
160
180
200

Popular solution methods

1. Tikhonov regularization – see Engl, Hanke, Neubauer
2. Iterative regularization methods
   (a) Conjugate gradient method applied to normal equations (CGNR)

$$A^*Ax = A^*b$$

$A^* : \mathcal{X} \longrightarrow \mathcal{X}$ adjoint operator to A,

   (b) GMRES method applied to $Ax = b$.

Outline

- CGNR, GMRES, and the discrepancy principle
- Augmentation and decomposition
- A modified LSQR algorithm
- A multilevel method
- Computation of constrained solutions

CGNR, GMRES, and the Discrepancy Principle

The CGNR method

Define the Krylov subspace

$$\mathcal{K}_k(A^*A, A^*b) = \mathrm{span}\{A^*b, (A^*A)A^*b,$$

$$\ldots, (A^*A)^{k-2}A^*b, (A^*A)^{k-1}A^*b\}.$$

Then $x_k \in \mathcal{K}_k(A^*A, A^*b)$ and

$$\|Ax_k - b\| = \min_{x \in \mathcal{K}_k(A^*A, A^*b)} \|Ax - b\|$$

Therefore discrepancy $d_j = b - Ax_j$ satisfies

$$\|b\| \geq \|d_1\| \geq \ldots \geq \|d_k\|.$$

The GMRES method

The iterate $x_k \in \mathcal{K}_k(A, b)$ satisfies

$$\|Ax_k - b\| = \min_{x \in \mathcal{K}_k(A,b)} \|Ax - b\|$$

where

$$\mathcal{K}_k(A, b) = \operatorname{span}\{b, Ab, \ldots, A^{k-1}b\}.$$

Hence

$$\|b\| \geq \|d_1\| \geq \ldots \geq \|d_k\|.$$

The RRGMRES method

The iterate $x_k \in \mathcal{K}_k(A, Ab)$ satisfies

$$\|Ax_k - b\| = \min_{x \in \mathcal{K}_k(A,Ab)} \|Ax - b\|$$

where

$$\mathcal{K}_k(A, Ab) = \text{span}\{Ab, A^2 b, \ldots, A^k b\}.$$

Hence

$$\|b\| \geq \|d_1\| \geq \ldots \geq \|d_k\|.$$

Stopping Criterion

Discrepancy principle

Let $\alpha > 1$ be fixed, $\|e\| = \|\hat{b} - b\| = \delta$. The iterate x_k satisfies the discrepancy principle if

$$\|Ax_k - b\| \leq \alpha\delta$$

Stopping rule

Terminate the iterations as soon as iterate x_k satisfies

$$\|Ax_k - b\| \leq \alpha\delta$$

$$\|Ax_{k-1} - b\| > \alpha\delta$$

Denote the termination index by k_δ.

An iterative method is a regularization method if

$$\lim_{\delta \searrow 0} \sup_{\|e\| \leq \delta} \|x_{k_\delta} - \hat{x}\| = 0$$

CGNR is a regularization method; see Nemirovskii, Hanke.

GMRES under suitable (stronger) conditions is a regularization method; see Calvetti et al. Numer. Math 2002.

Augmentation and Decomposition

Decomposition

span$\{W\}$ user-supplied subspace:

$$W \in \mathbf{R}^{n\times \ell}, \qquad W^T W = I,$$
$$P_W = WW^T, \qquad P_W^\perp = I - P_W.$$

Decompose the computed approximate solution x_j according to

$$x_j = x_j' + x_j'', \qquad x_j' = P_W x_j, \qquad x_j'' = P_W^\perp x_j.$$

Choose W to allow the representation of pertinent features of $\hat{x}$.

Implementation

Compute QR-factorization

$$AW = QR$$

and define

$$P_Q = QQ^T, \qquad P_Q^\perp = I - P_Q.$$

Then

$$Ax = b$$

can be expressed as

$$\begin{aligned} P_Q A P_W x + P_Q A P_W^\perp x &= P_Q b, \\ P_Q^\perp A P_W^\perp x &= P_Q^\perp b. \end{aligned}$$

The latter equation can be expressed as

$$P_Q^\perp Az = P_Q^\perp b.$$

- Solve this equation with an iterative method.
  Iterates $z_1'', z_2'', \ldots$

- Let $x_j'' = P_W^\perp z_j''$.

- Solve small system for $x_j' \in \mathrm{range}(W)$.

- Let $x_j = x_j' + x_j''$.

Note:

$$\|b - Ax_j\| = \|P_Q^\perp b - P_Q^\perp A z_j''\|.$$

Easy to apply discrepancy principle.

Note: Decomposition with GMRES equivalent to augmented GMRES:

$$\|Ax_j - b\| = \min_{x \in \mathcal{K}_j(A,b) \cup W} \|Ax - b\|, \qquad x_j \in \mathcal{K}_j(A,b) \cup W.$$

Example (deriv2):

$$\int_0^1 k(s,t)x(t)dt = e^s + (1-e)s - 1, \qquad 0 \le s \le 1,$$

with

$$k(s,t) = \begin{cases} s(t-1), & s < t, \\ t(s-1), & s \ge t. \end{cases}$$

Solution $x(t) = e^t$.

Let

$$A \in \mathbf{R}^{400\times 400}, \quad b = \hat{b} + e, \quad \|e\|/\|\hat{b}\| = 10^{-3}.$$

Let

$$\hat{W} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \\ 1 & 400 \end{bmatrix}, \qquad \hat{W} = WR.$$

Terminate iterations by discrepancy principle.

standard GMRES: $\|x_8 - \hat{x}\| = 5.2 \cdot 10^{-1}$

standard RRGMRES: $\|x_{10} - \hat{x}\| = 2.8 \cdot 10^{-1}$

standard LSQR: $\|x_{12} - \hat{x}\| = 2.8 \cdot 10^{-1}$

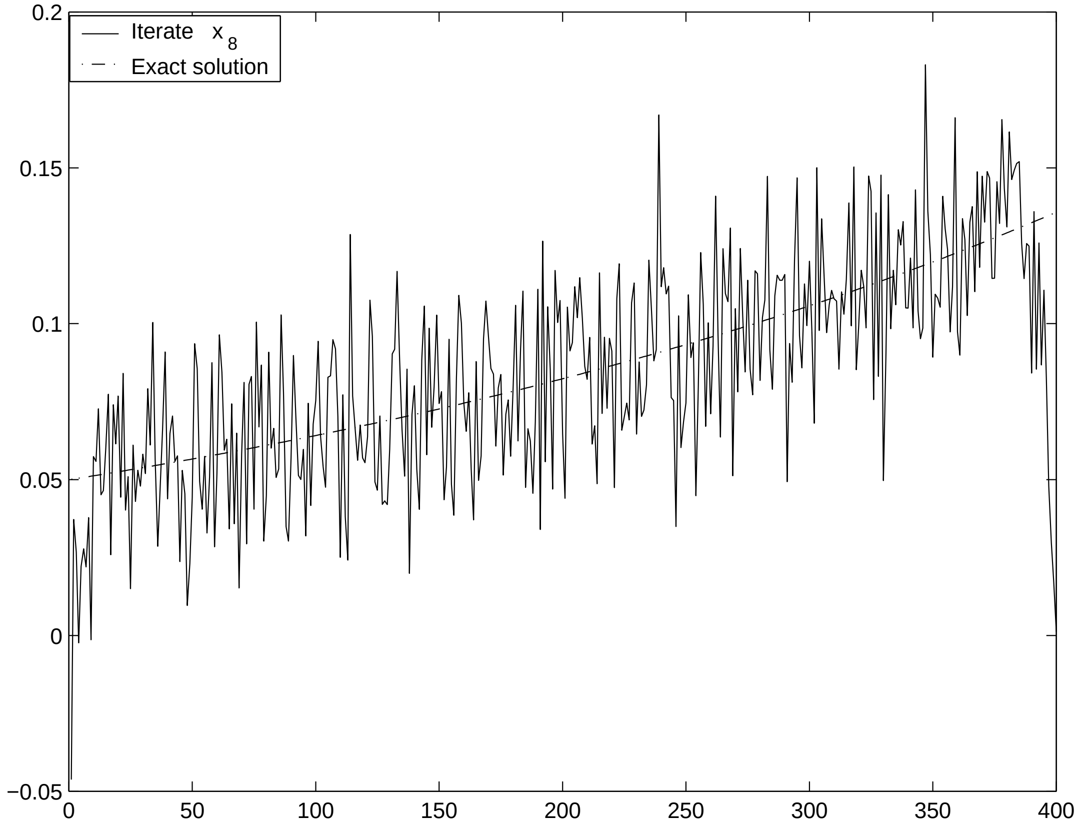
Decomposition with W:

GMRES: $\|x_2 - \hat{x}\| = 6.5 \cdot 10^{-2}$

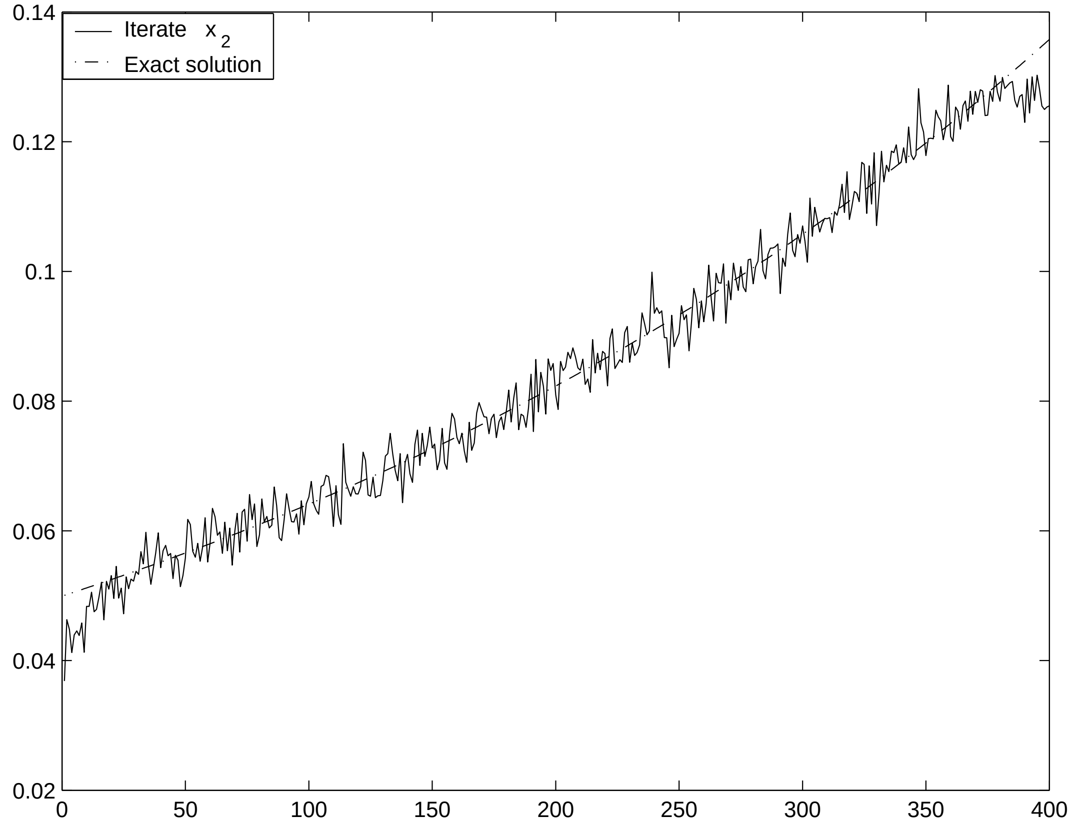
RRGMRES: $\|x_3 - \hat{x}\| = 2.9 \cdot 10^{-2}$

LSQR: $\|x_3 - \hat{x}\| = 3.1 \cdot 10^{-3}$

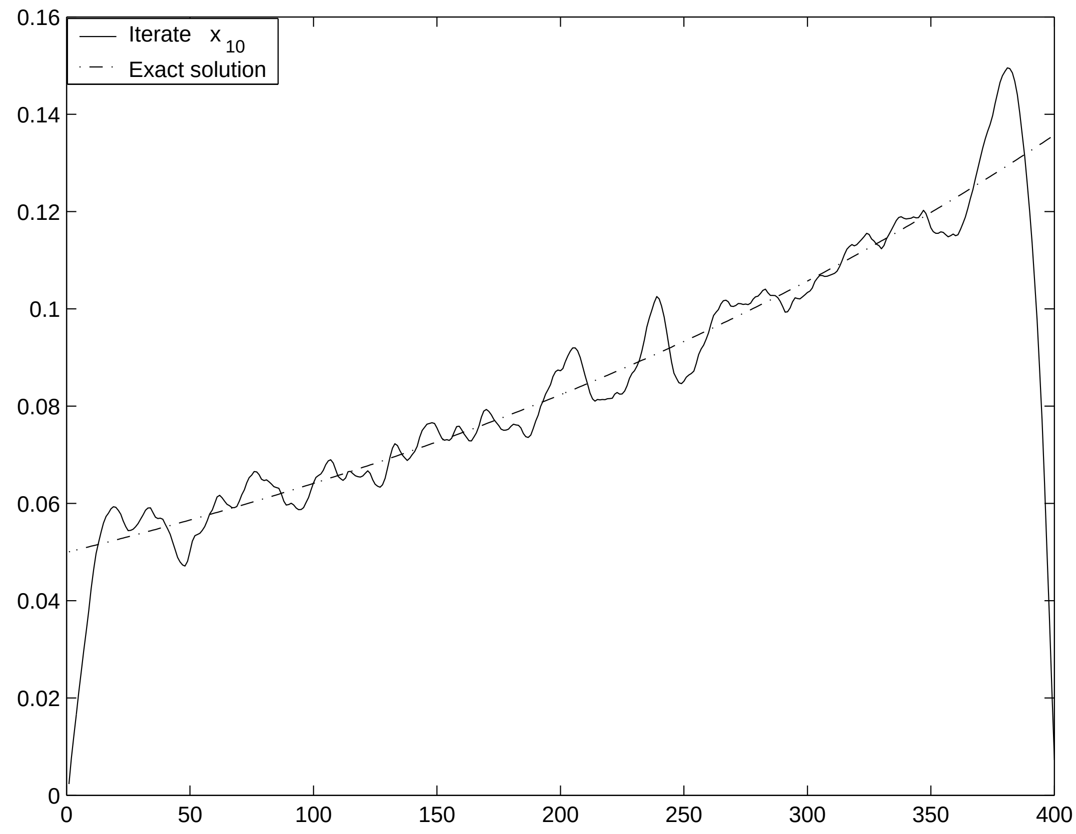
Computed solution by standard GMRES



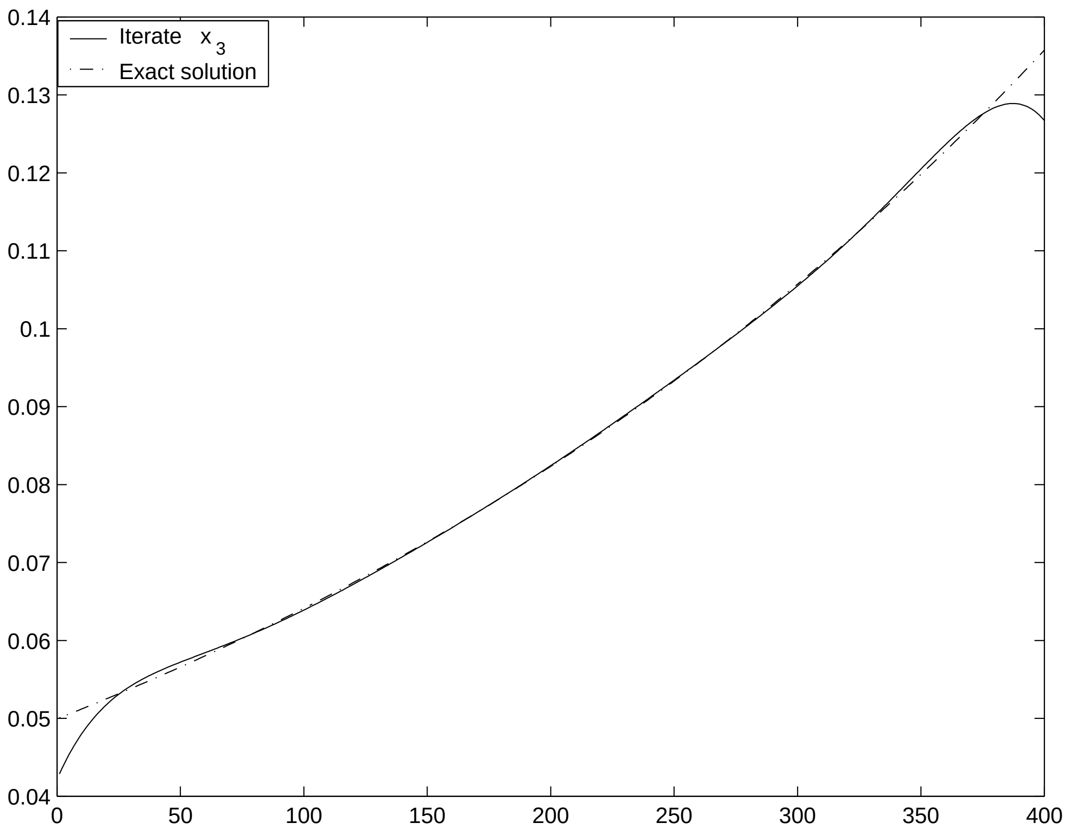
Computed solution by GMRES with W



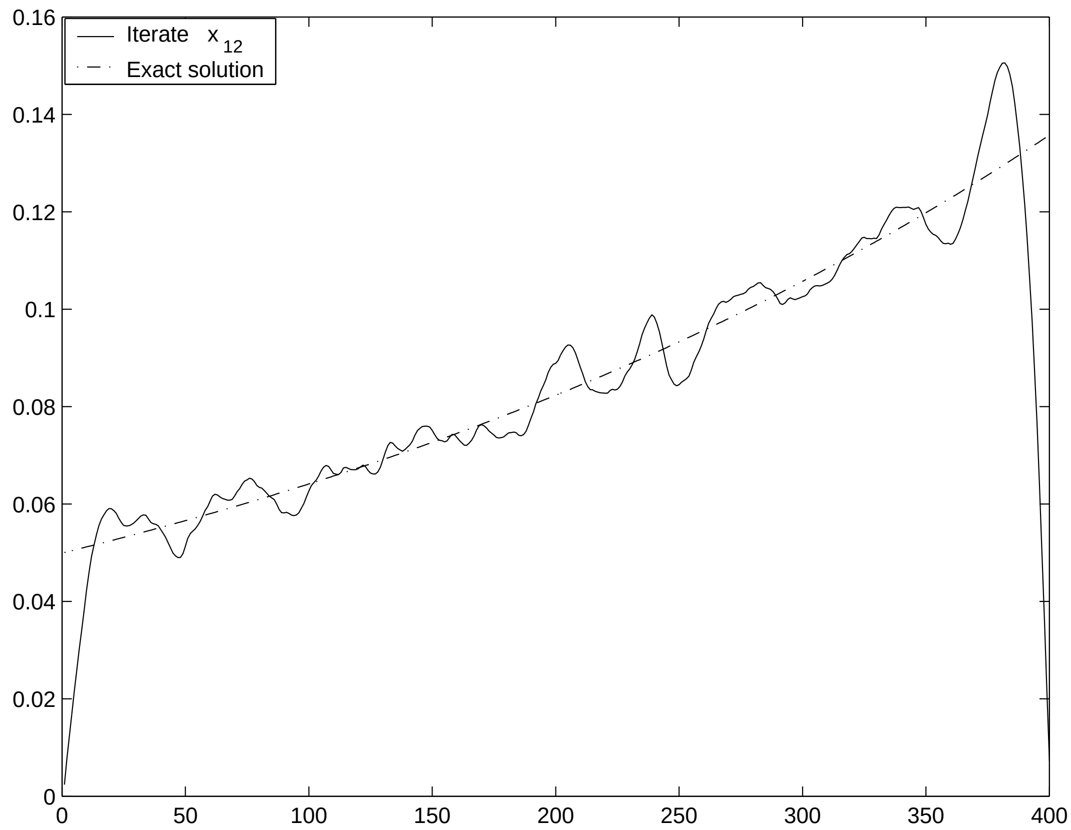
Computed solution by RRGMRES



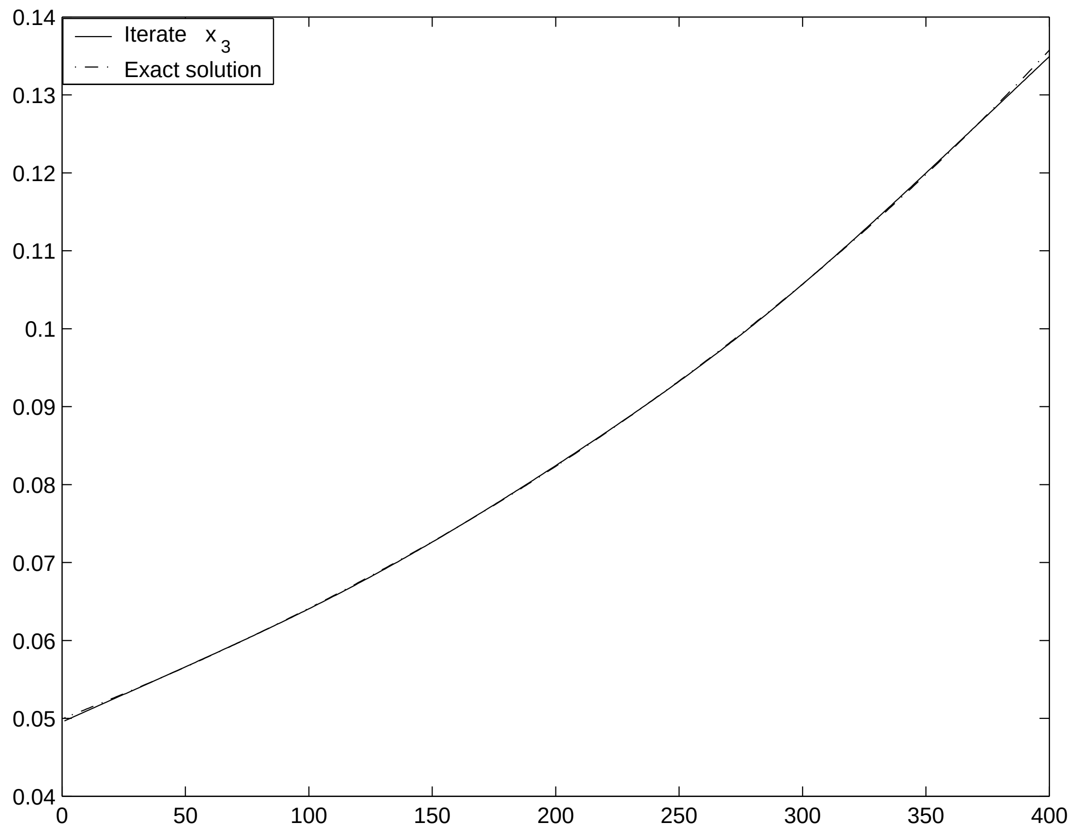
Computed solution by RRGMRES with W



Computed solution by LSQR



Computed solution by LSQR with W



Example (baart+1): Let

$$A \in \mathbf{R}^{200\times 200}, \quad b = \hat{b} + e, \quad \|e\|/\|\hat{b}\| = 10^{-3}.$$

$$W = [1, 1, \ldots, 1]^T/\sqrt{200}.$$

Terminate iterations by discrepancy principle.

standard RRGMRES: $\|x_3 - \hat{x}\| = 6.8 \cdot 10^{-2}$

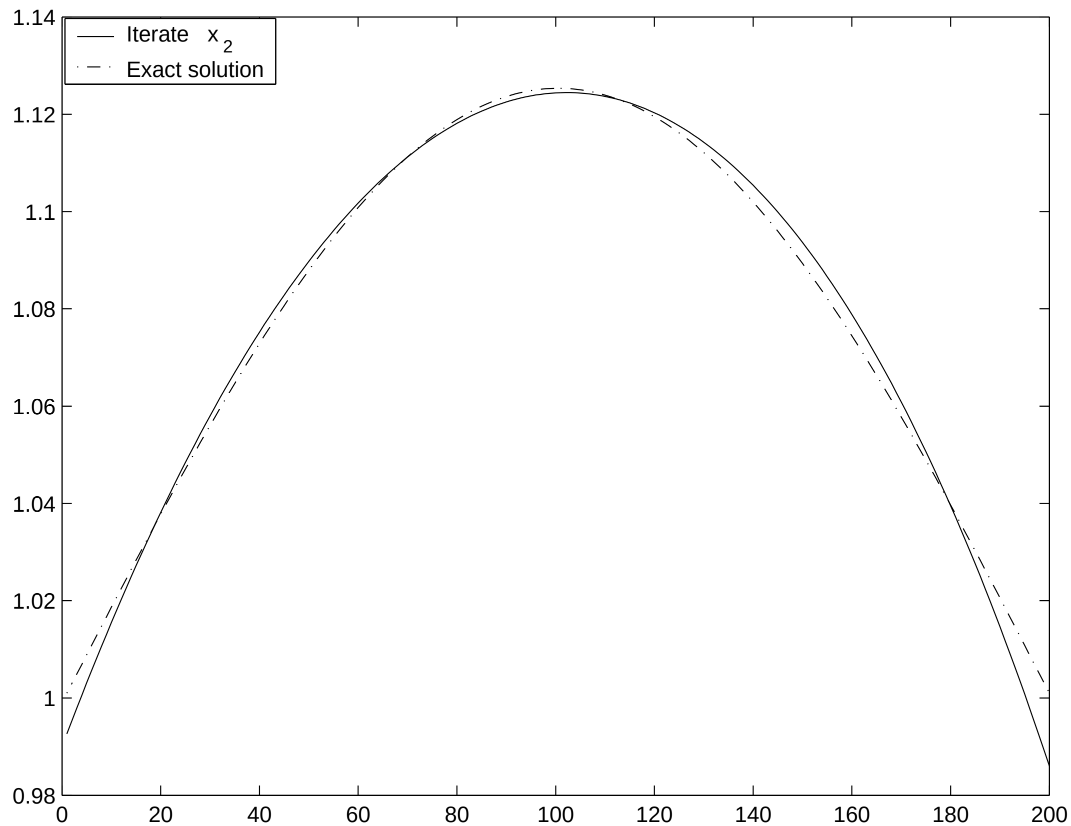
standard LSQR: $\|x_3 - \hat{x}\| = 1.6 \cdot 10^{-1}$

Decomposition with W:

RRGMRES: $\|x_2 - \hat{x}\| = 5.0 \cdot 10^{-2}$

LSQR: $\|x_2 - \hat{x}\| = 1.4 \cdot 10^{-1}$

Computed solution by RRGMRES with W



A Modified LSQR Algorithm

Standard Lanczos bidiagonalization

$$AV_k = U_{k+1}\bar{B}_k, \qquad A^T U_k = V_k B_k^T,$$

with B_k lower bidiagonal, U_{k+1}, V_k orthonormal columns,

$$U_k e_1 = b/\|b\|, \qquad V_k e_1 = A^T b/\|A^T b\|.$$

LSQR uses these decompositions to determine approximate solution x_k in

$$\mathrm{range}(V_k) = \mathcal{K}_k(A^T A, A^T b).$$

Can we change this subspace?

Generalized LSQR

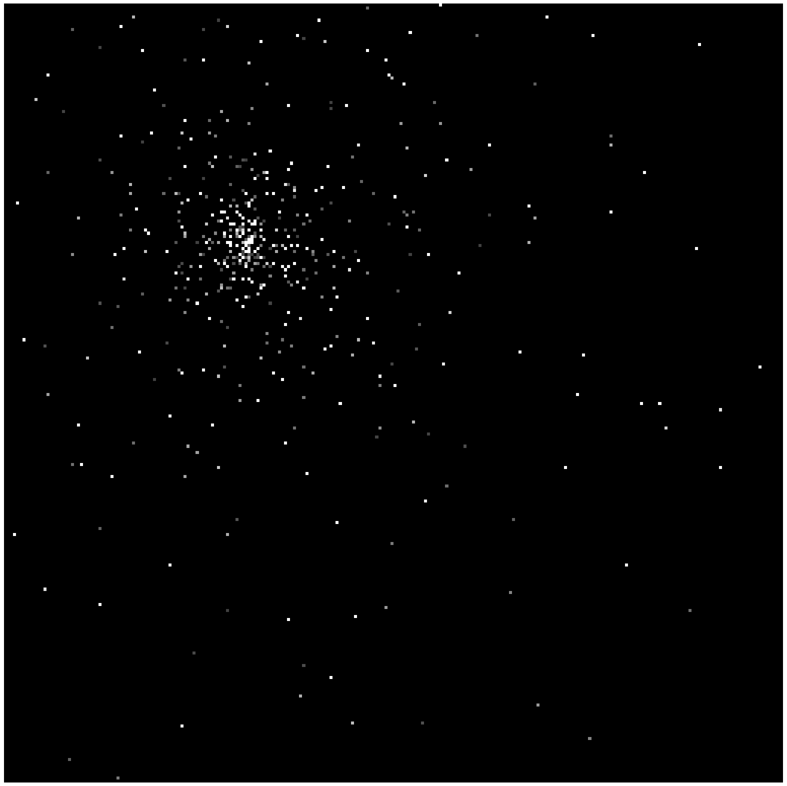
Let $v_1 = V_k e_1$ be arbitrary unit vector. Then

$$AV_k = U_{k+1}\bar{T}_k, \qquad A^T U_k = V_k S_k^T,$$

with $S_k^T = T_k$ tridiagonal.
Breakdowns benign.

Example (star cluster): b represents blurred and noisy 256×256 pixel image of star cluster. Matrix A models atmospheric blur. $\hat{x}$ represents the blur- and noise-free image. $\hat{b}$ represents blurred but noise-free image.

Blur- and noise-free image



Blurred and noisy image. Relative noise 10^{-2}.

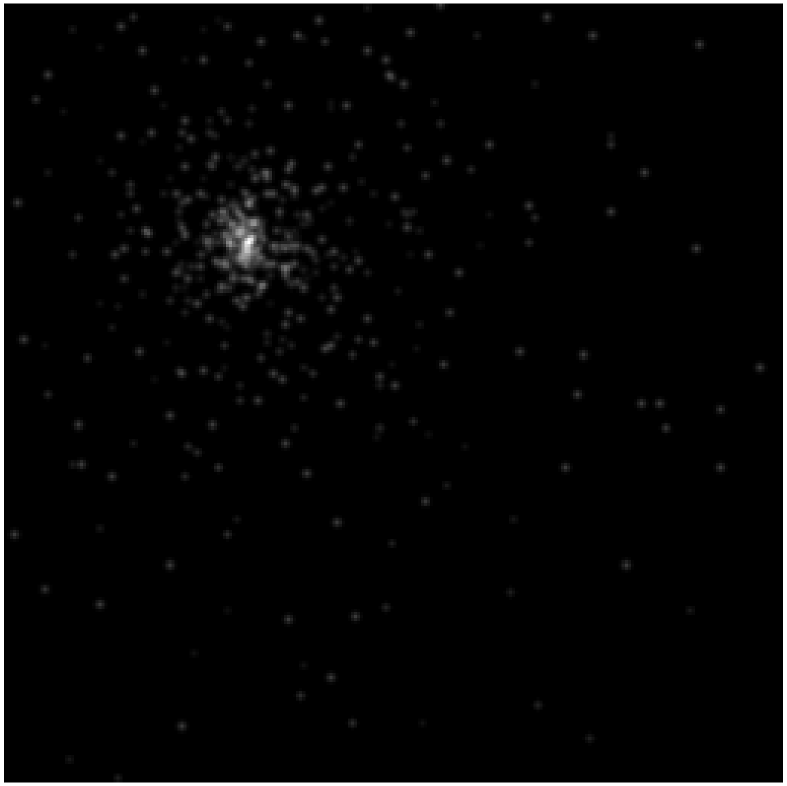


Image restored by 40 iterations with LSQR.

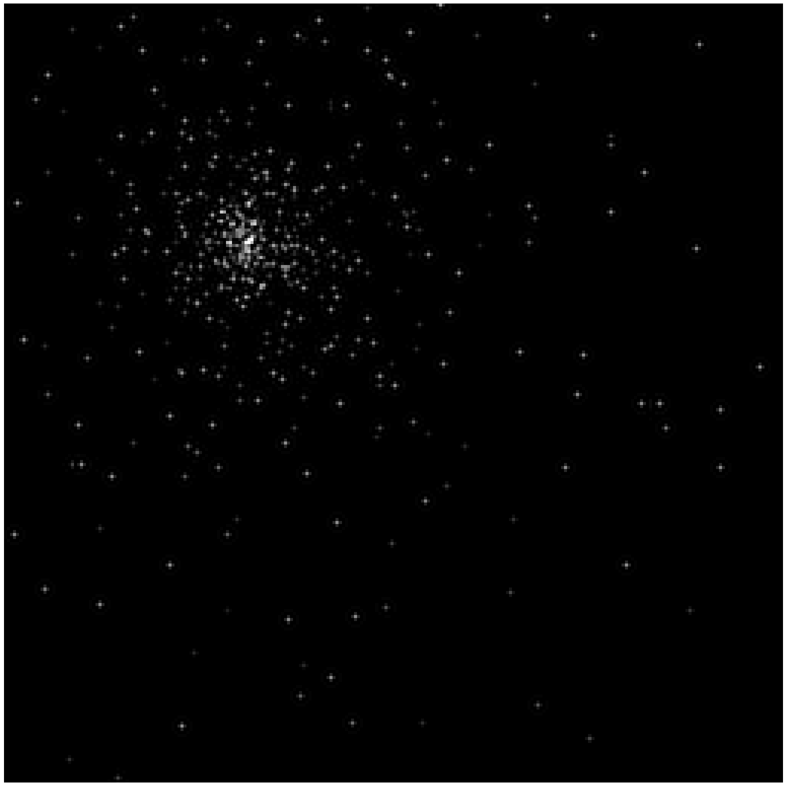
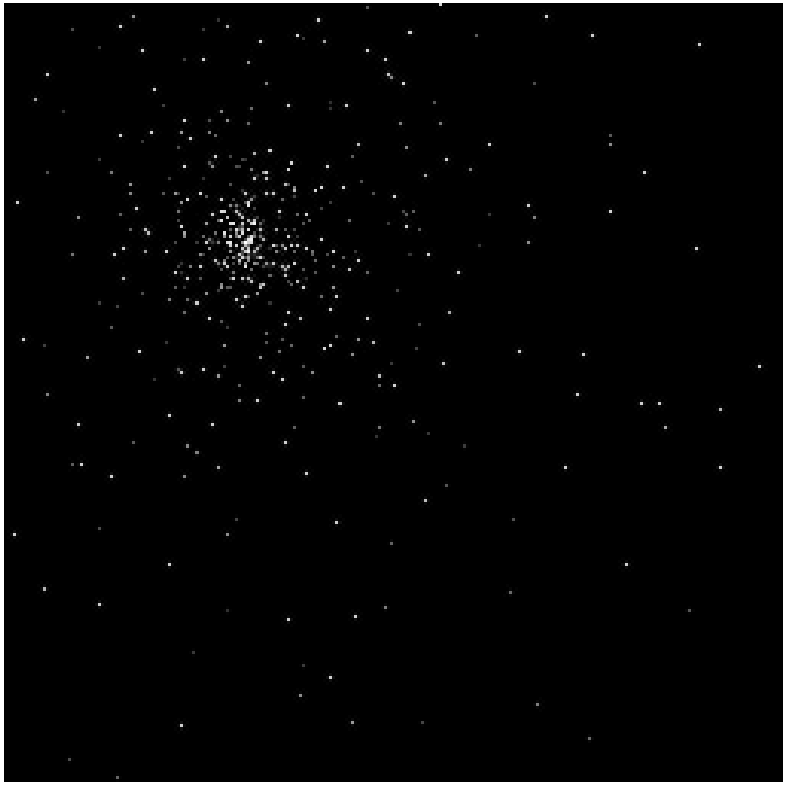
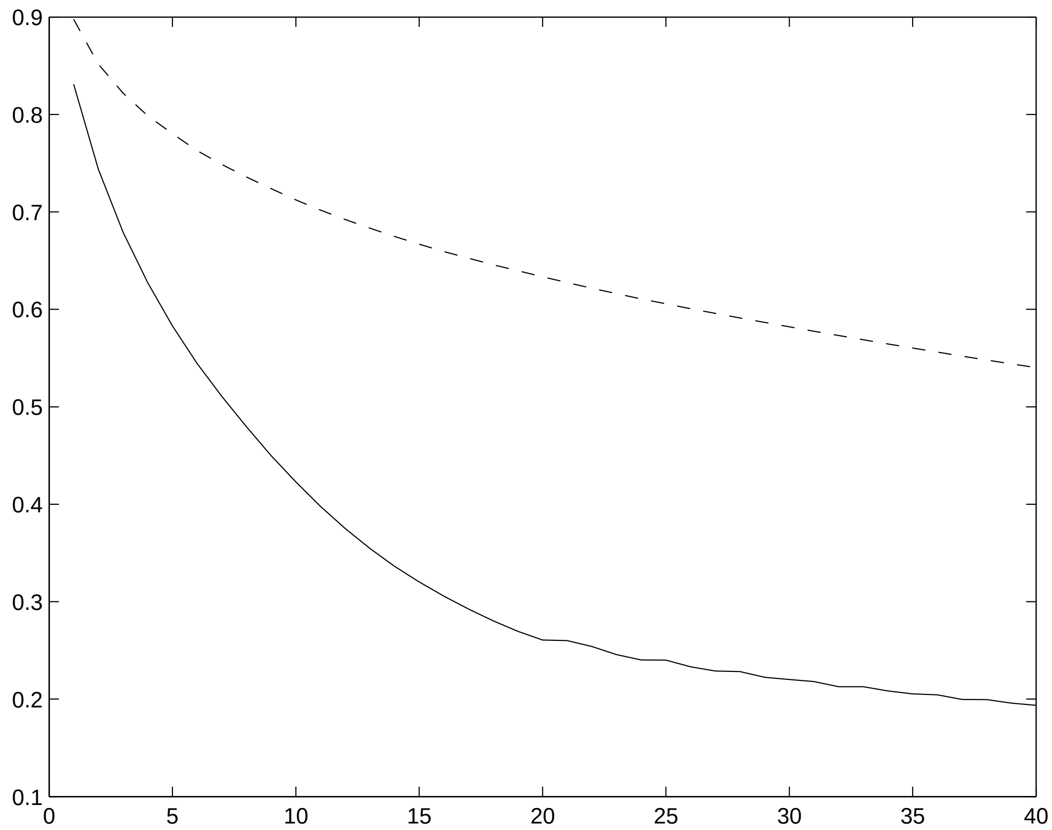


Image restored by 40 iterations with modified LSQR with $V_k e_1 = b/\|b\|$.



Relative error in LSQR iterates (top curve) and modified LSQR iterates (bottom curve).



A Multilevel Method

Consider

$$\int_{\Omega} k(s,t)x(s)ds = b(t), \qquad t \in \Omega$$

Let

$$S_1 \subset S_2 \subset \ldots \subset S_k \subset L_2(\Omega) \qquad \text{nested subspaces}$$

$$R_i : L_2(\Omega) \longrightarrow S_i \qquad \text{restriction operator}$$

$$b_i = R_i b, \qquad \hat{b}_i = R_i \hat{b}, \qquad A_i = R_i A R_i^*$$

$$P_i : S_{i-1} \longrightarrow S_i \qquad \text{prolongation operator}$$

Cascadic multilevel method:

- Solve integral equation in S_1, map solution to S_2,
- Solve integral equation in S_2 for correction, map to S_3,
- Solve integral equation in S_3 for correction, map to S_4 ...

Assume that

$$\|b - \hat{b}\| = \delta$$

and

$$\|b_i - \hat{b}_i\| \leq c\delta, \quad c > 1, \quad i = 1, 2, \ldots$$

Apply CGNR or MR-II in S_1. Yields iterates $x_{1,j}$, $j = 1, 2, \ldots$. Terminate iterations as soon as

$$\|A_1 x_{1,j} - b_1\| \leq c\delta.$$

Proceed similarly on higher levels.

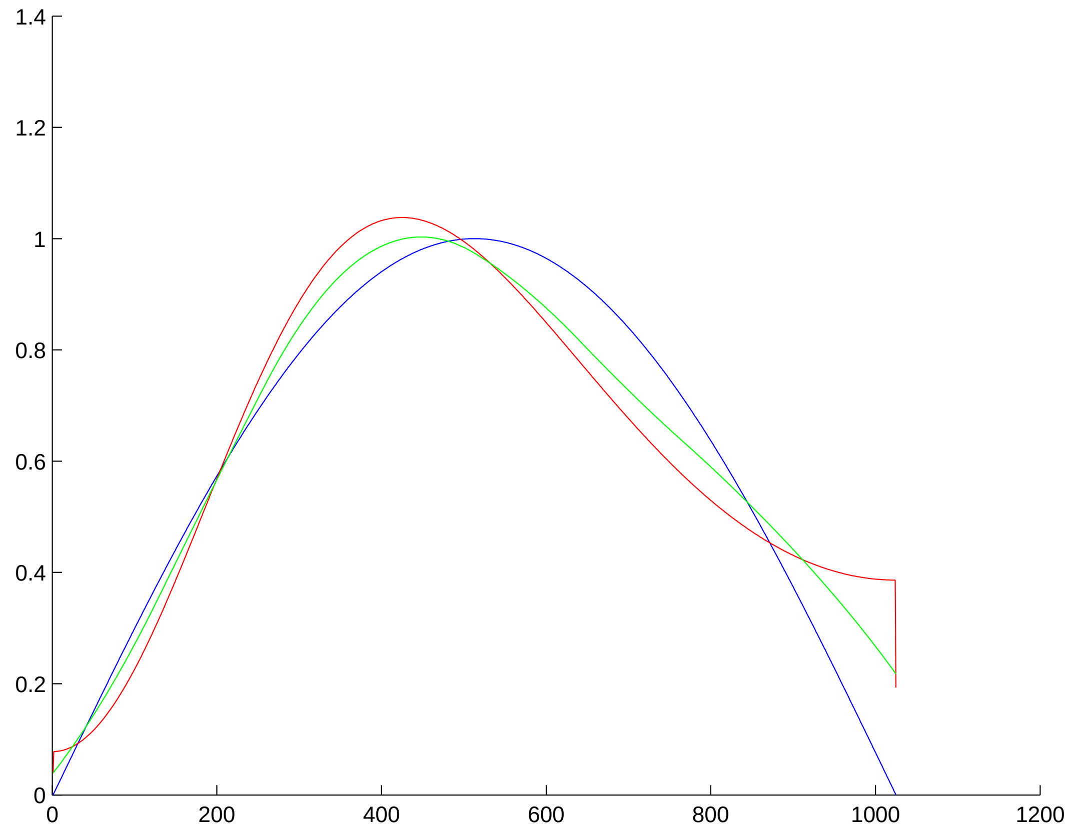
Theorem: The multilevel method outlined is a regularization method.

Example: Baart integral equation discretized by trapeziodal rule on mesh with 1025 point.

	One-Grid CGNR	
$\frac{\delta}{\|b\|}$	$m(\delta)$	$\frac{\|x^{\delta}_{8,m(\delta)}-\hat{x}\|}{\|\hat{x}\|}$
$1 \cdot 10^{-1}$	2	0.3412
$1 \cdot 10^{-2}$	3	0.1662
$1 \cdot 10^{-3}$	3	0.1657
$1 \cdot 10^{-4}$	4	0.1143

	Multilevel CGNR	
$\frac{\delta}{\|b\|}$	$m_i(\delta)$	$\frac{\|x^{\delta}_{8,m_8(\delta)} - \hat{x}\|}{\|\hat{x}\|}$
$1 \cdot 10^{-1}$	2, 1, 1, 1, 1, 1, 1, 1	0.2686
$1 \cdot 10^{-2}$	2, 2, 1, 3, 1, 1, 1, 1	0.1110
$1 \cdot 10^{-3}$	3, 3, 2, 1, 1, 1, 1, 1	0.1065
$1 \cdot 10^{-4}$	4, 3, 3, 3, 2, 1, 1, 1	0.0669

Noise level 10^{-3}: CGNR (red curve), ML-CGNR (green curve), exact solution (blue curve)



Example: Phillips integral equation discretized by trapeziodal rule on mesh with 1025 point.

	One-Grid CGNR	
$\frac{\delta}{\|b\|}$	$m(\delta)$	$\frac{\|x^{\delta}_{8,m(\delta)} - \hat{x}\|}{\|\hat{x}\|}$
$1 \cdot 10^{-1}$	3	0.0934
$1 \cdot 10^{-2}$	4	0.0248
$1 \cdot 10^{-3}$	4	0.0243
$1 \cdot 10^{-4}$	11	0.0064

	Multilevel CGNR	
$\frac{\delta}{\|b\|}$	$m_i(\delta)$	$\frac{\|x^\delta_{8,m_8(\delta)}-\hat{x}\|}{\|\hat{x}\|}$
$1\cdot 10^{-1}$	2, 1, 1, 1, 1, 1, 1, 1	0.0842
$1\cdot 10^{-2}$	5, 5, 3, 2, 1, 1, 1, 1	0.0343
$1\cdot 10^{-3}$	8, 6, 6, 4, 3, 1, 1, 1	0.0243
$1\cdot 10^{-4}$	9, 13, 9, 9, 5, 4, 3, 2	0.0076

Example: Restoration of an image with 409×409 pixels.

Available blurred and noisy image

Restored image, 4 iterations on finest level.

More iterations ...

Even more iterations ...

Computation of
Constrained Solutions

A CGNR-based active set-type method

Determine $x \in \mathcal{S}$, such that

$$\|Ax - b\| \leq \eta\delta,$$

where $\eta \geq 1$ and

$$\mathcal{S} = \{x = [x^{(1)}, x^{(2)}, \ldots, x^{(n)}]^T \in \mathbf{R}^n :$$

$$\ell^{(i)} \leq x^{(i)} \; \forall i \in \mathcal{I}^{(\ell)}, \; x^{(i)} \leq u^{(i)} \; \forall i \in \mathcal{I}^{(u)}\}.$$

Introduce the residual

$$r = [r^{(1)}, r^{(2)}, \ldots, r^{(n)}]^T = A^T A x - A^T b,$$

whose components are Lagrange multipliers.
By the KKT-equations, x solves

$$\min_{x \in \mathcal{S}} \|Ax - b\|$$

if

$$r^{(i)} \geq 0 \quad \forall i \in \mathcal{I}^{(\ell)}, \qquad r^{(i)} \leq 0 \quad \forall i \in \mathcal{I}^{(u)},$$

and

$$r^{(i)} = 0 \quad \forall i \notin \mathcal{I}^{(\ell)} \cup \mathcal{I}^{(u)}.$$

We impose the inequality conditions.

Algorithm:

1. Solve unconstrained problem

   $$\|Ax - b\| \leq \eta\delta \qquad (*)$$

   by CGNR. Gives $\breve{x}$.

2. Project $\breve{x}$ onto $\mathcal{S}$. Gives $\tilde{x}$. If $\tilde{x}$ satisfies (*) then done.

3. Define the active sets

   $$\mathcal{A}^{(\ell)}(\tilde{x}) = \{i \in \mathcal{I}^{(\ell)} : \tilde{x}^{(i)} = \ell^{(i)}\},$$

   $$\mathcal{A}^{(u)}(\tilde{x}) = \{i \in \mathcal{I}^{(u)} : \tilde{x}^{(i)} = u^{(i)}\}.$$

4. Compute residual vector

$$\tilde{r} = [\tilde{r}^{(1)}, \tilde{r}^{(2)}, \ldots, \tilde{r}^{(n)}]^T = A^T A\tilde{x} - A^T b,$$

which yields the Lagrange multipliers.

5. if $i \in \mathcal{A}^{(\ell)}(\tilde{x})$ and $\tilde{r}^{(i)} < 0$ then remove index i from $\mathcal{A}^{(\ell)}(\tilde{x})$.
   if $i \in \mathcal{A}^{(u)}(\tilde{x})$ and $\tilde{r}^{(i)} > 0$ then remove index i from $\mathcal{A}^{(u)}(\tilde{x})$.

6. Introduce $D = \mathrm{diag}[d^{(1)}, d^{(2)}, \ldots, d^{(n)}]$, where

$$d^{(k)} := \begin{cases} 0, & k \in \mathcal{A}^{(\ell)}(\tilde{x}) \cup \mathcal{A}^{(u)}(\tilde{x}), \\ 1, & \text{otherwise.} \end{cases}$$

7. Solve approximately

$$ADz = -\tilde{r},$$

i.e., apply CGNR until an iterate z_j satisfies

$$\|ADz_j + \tilde{r}\| \leq \eta\delta.$$

This gives new approximate solution

$$\breve{x} := \tilde{x} + Dz_j$$

of the original problem. Goto 2.

Computed Examples

Blur, $n = 128^2$, $\delta = \frac{\|e\|}{\|\hat{b}\|} = 10^{-2}$, $\eta = 1$.

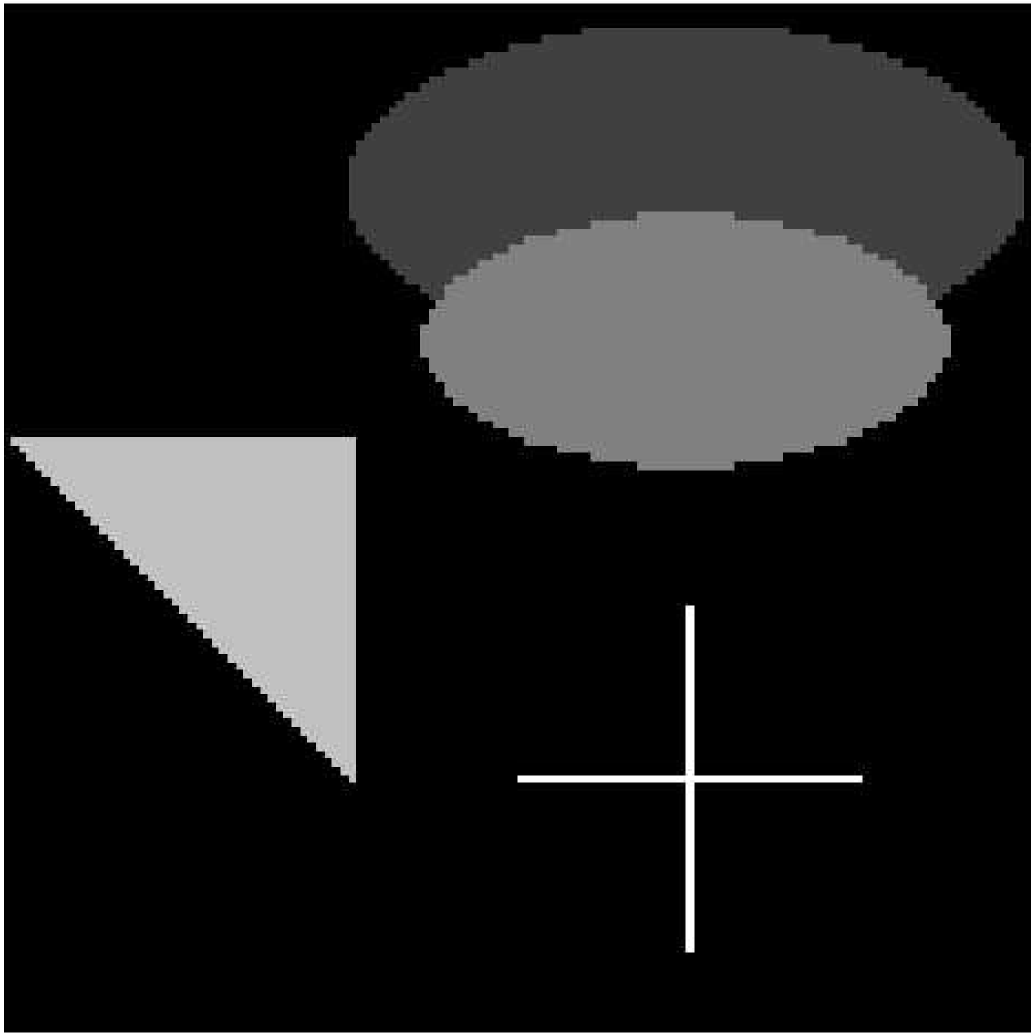
Symmetric blurring matrix models Gaussian blur.

Unconstrained problem: $\|\breve{x} - \hat{x}\| / \|\hat{x}\| = 1.9 \cdot 10^{-1}$

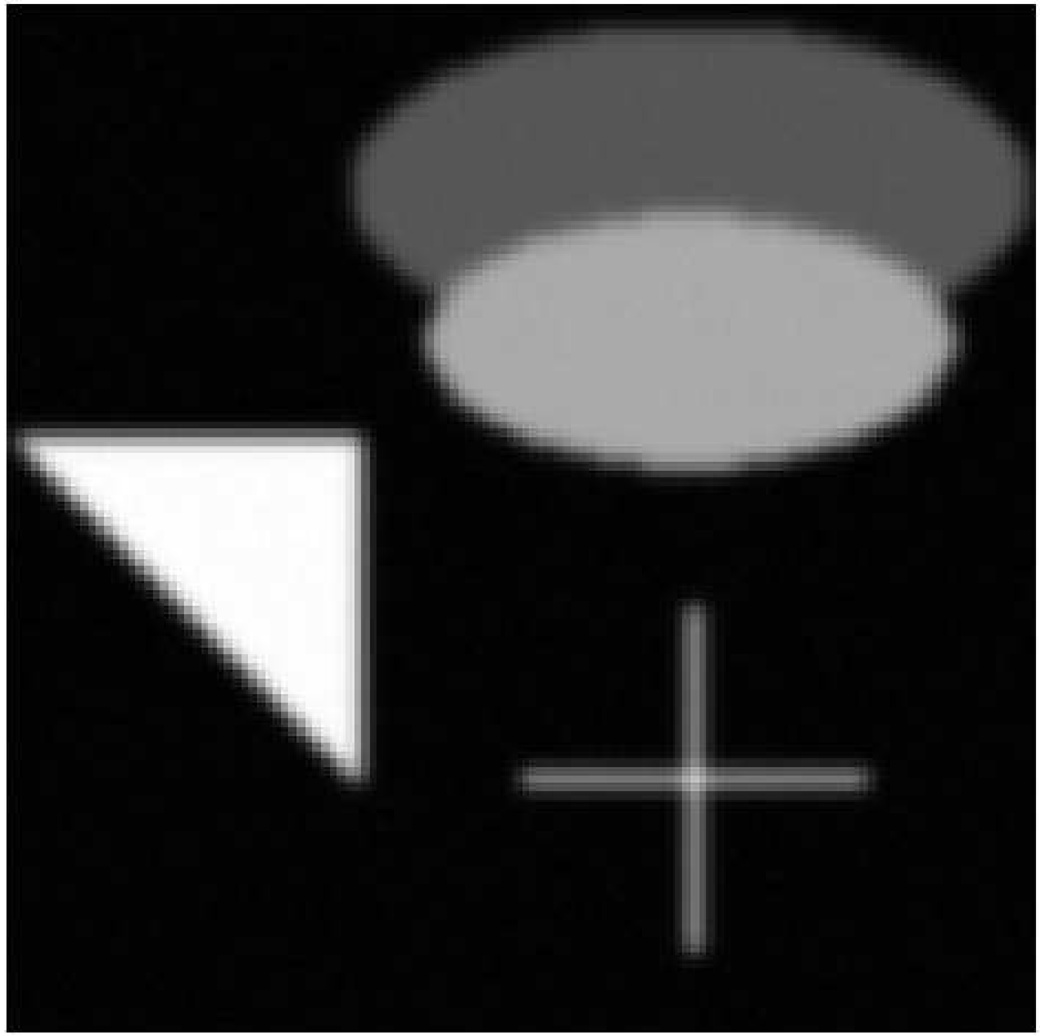
Nonnegatively constrained problem:
$\|\tilde{x} - \hat{x}\| / \|\hat{x}\| = 9.3 \cdot 10^{-2}$

4 outer iterations, 37 inner iterations (CGNR steps), 82 mat-vec prods.

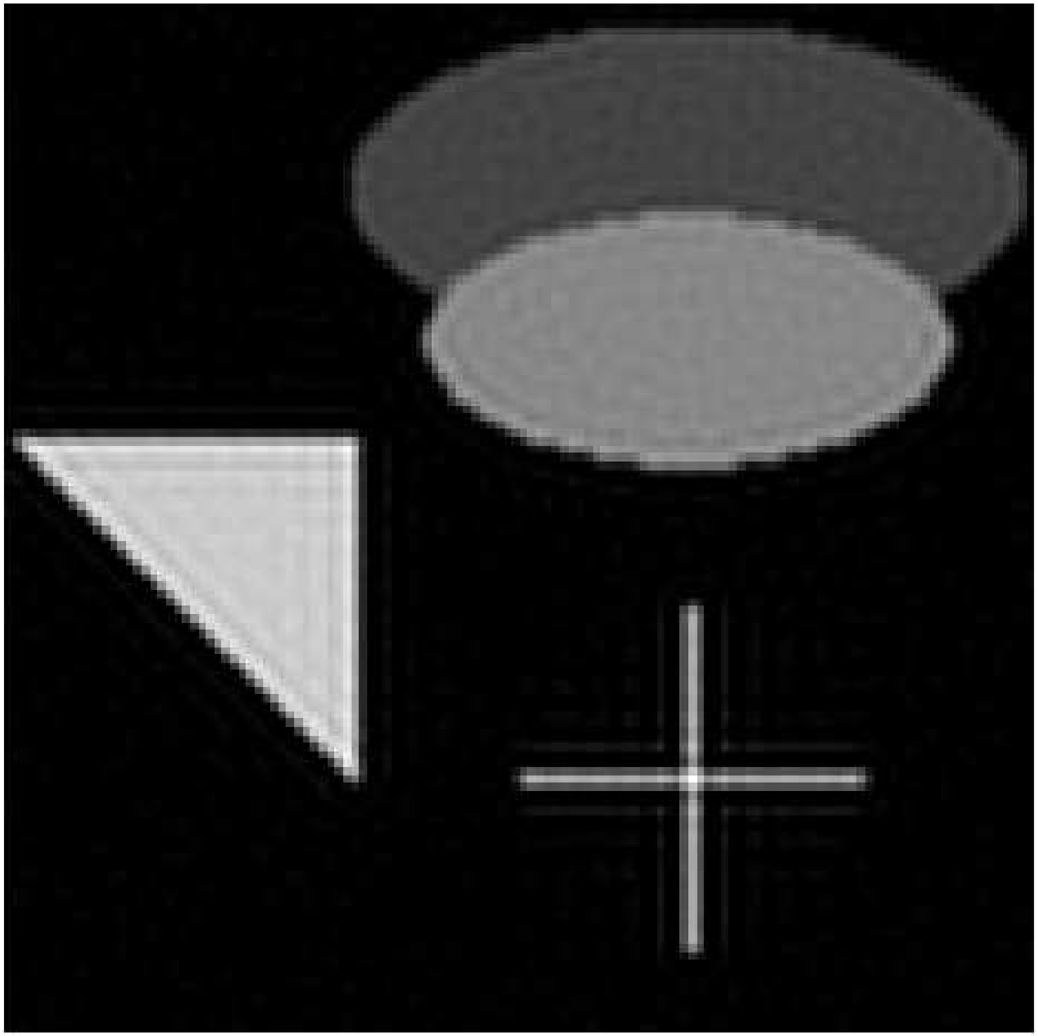
Blur- and noise-free image



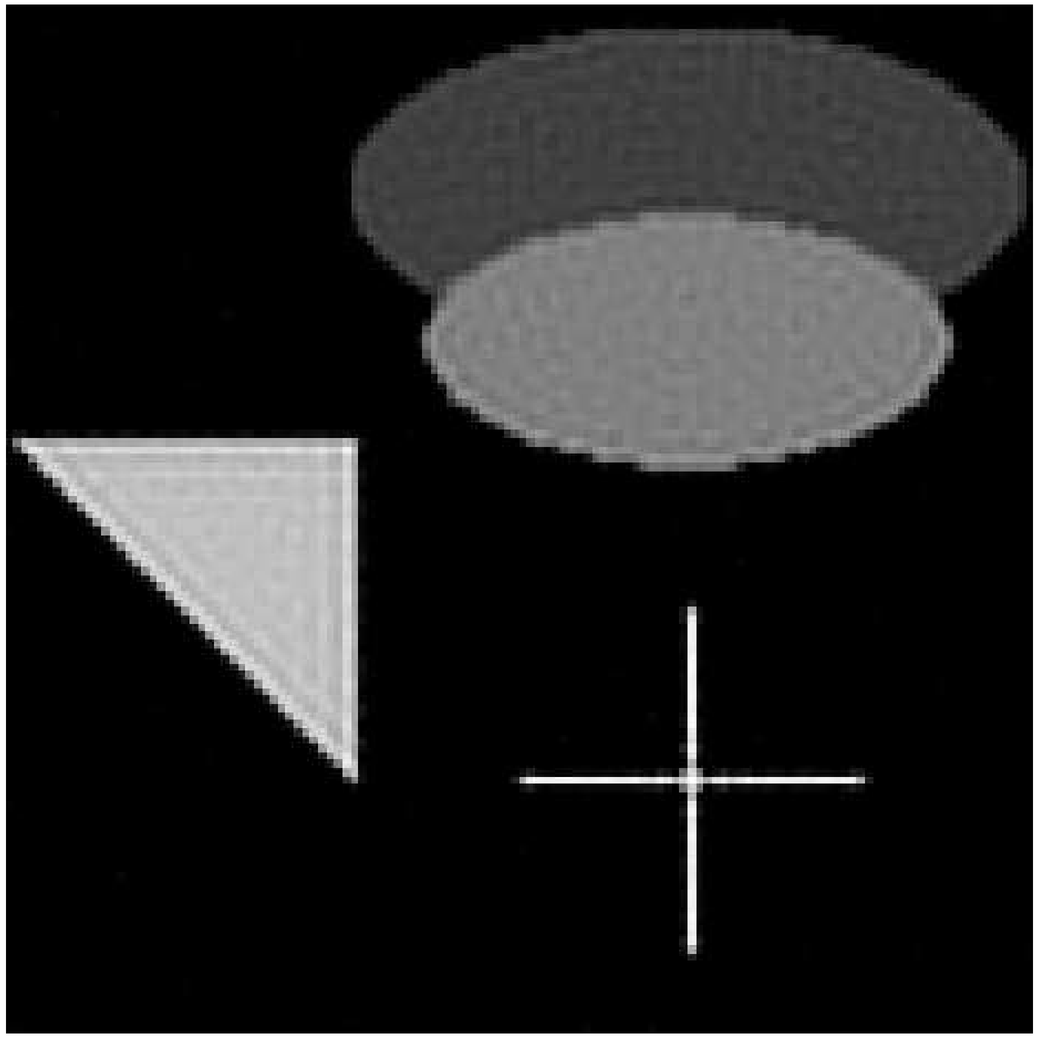
Blurred and noisy image



Computed solution without nonnegativity constraint



Computed solution with nonnegativity constraint



Towers, $\quad n = 256^2, \qquad \delta = \frac{\|e\|}{\|\hat{b}\|} = 10^{-2}$, $\eta = \frac{3}{2}$.

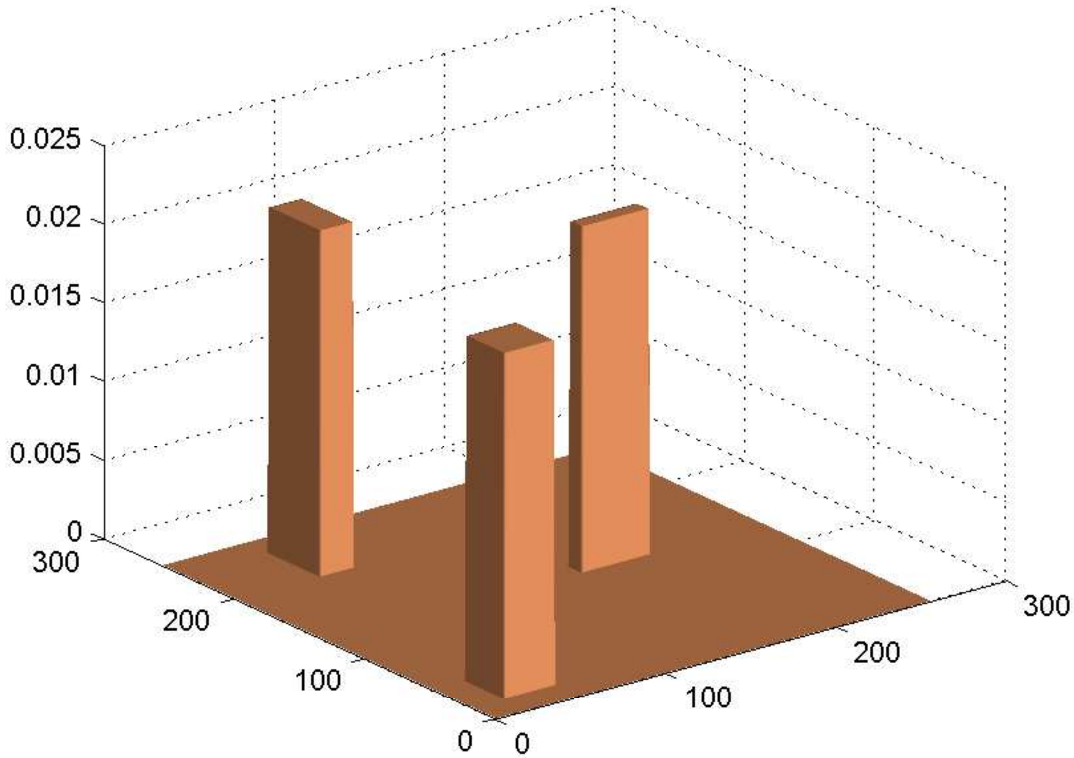
Nonsymmetric BTTB blurring matrix.

Unconstrained problem: $\|\breve{x} - \hat{x}\|/\|\hat{x}\| = 3.3 \cdot 10^{-1}$

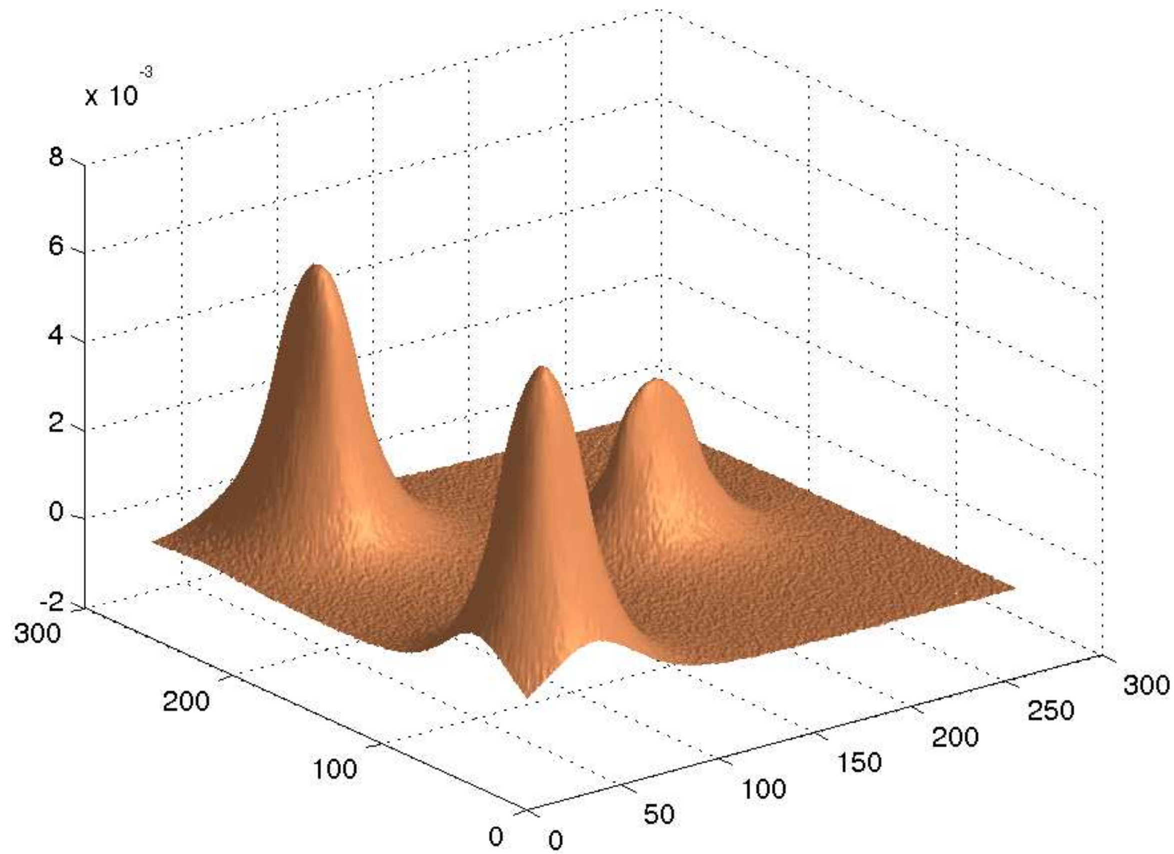
Constrained problem (upper and lower bounds imposed): $\|\tilde{x} - \hat{x}\|/\|\hat{x}\| = 3.0 \cdot 10^{-1}$

4 outer iterations, 49 inner iterations (CGNR steps), 109 mat-vec prods.

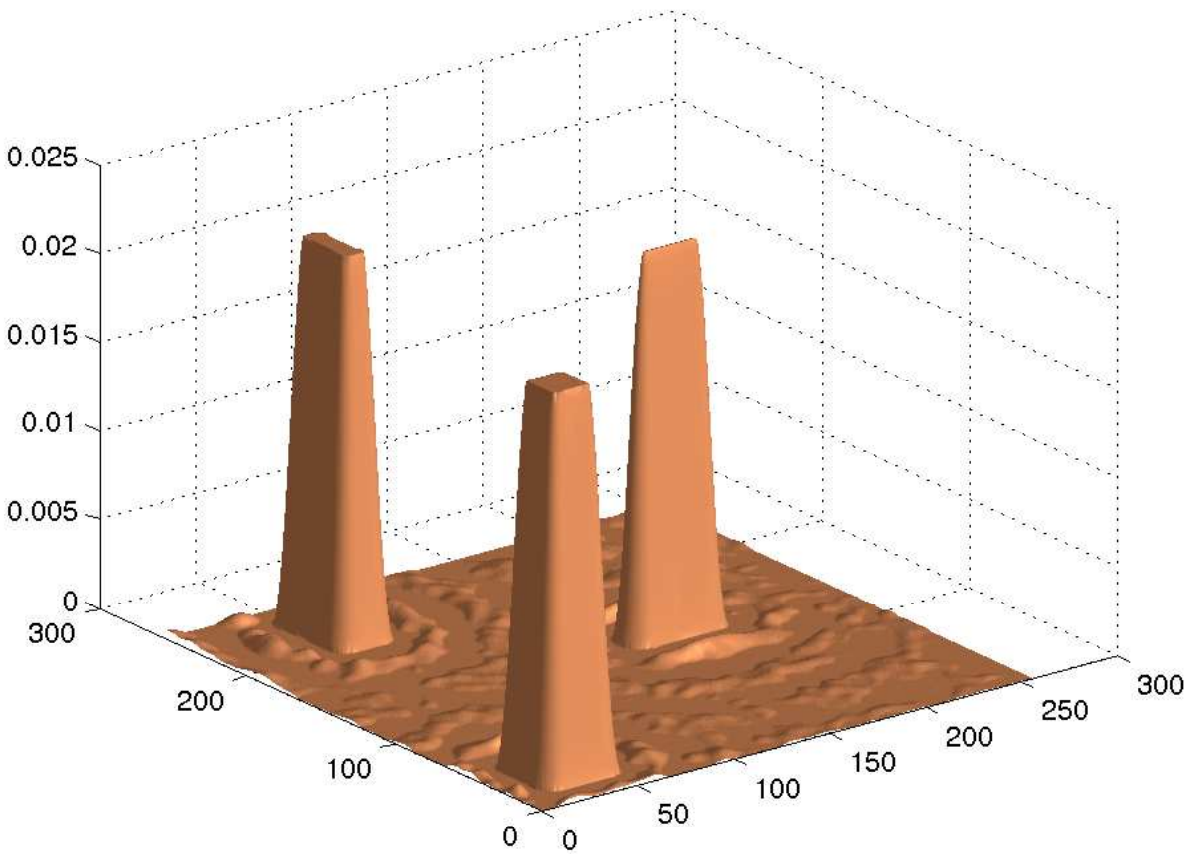
Blur- and noise-free image



Blurred and noisy image



Computed solution without constraints



Computed solution with constraints