The Arnoldi Process for Ill-Posed Problems

Marco Donatelli

Silvia Gazzola

Silvia Noschese

Paolo Novati

Ronny Ramlau

Lothar Reichel

Kalamata, NASCA 2018

Inverse problems

Inverse problems arise when one seeks to determine the cause of an observed effect.

- Computerized tomography.
- Image restoration: Determine the unavailable exact image from an available contaminated version.

Inverse problems often are ill-posed.

Ill-posed problems

A problem is said to be ill-posed if it has at least one of the properties:

- The problem does not have a solution.
- The problem does not have a unique solution.
- The solution does not depend continuously on the data.

Example: Consider the 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 by Hansen.

The code baart gives

- the matrix $A \in \mathbf{R}^{200\times 200}$, which is numerically singular,
- the desired solution $x_{\text{exact}} \in \mathbf{R}^{200}$, and
- the error-free right-hand side $b_{\text{exact}} \in \mathbf{R}^{200}$.

Then

$$A x_{\text{exact}} = b_{\text{exact}}.$$

Assume that b_{exact} is not available. Instead a noise-contaminated vector

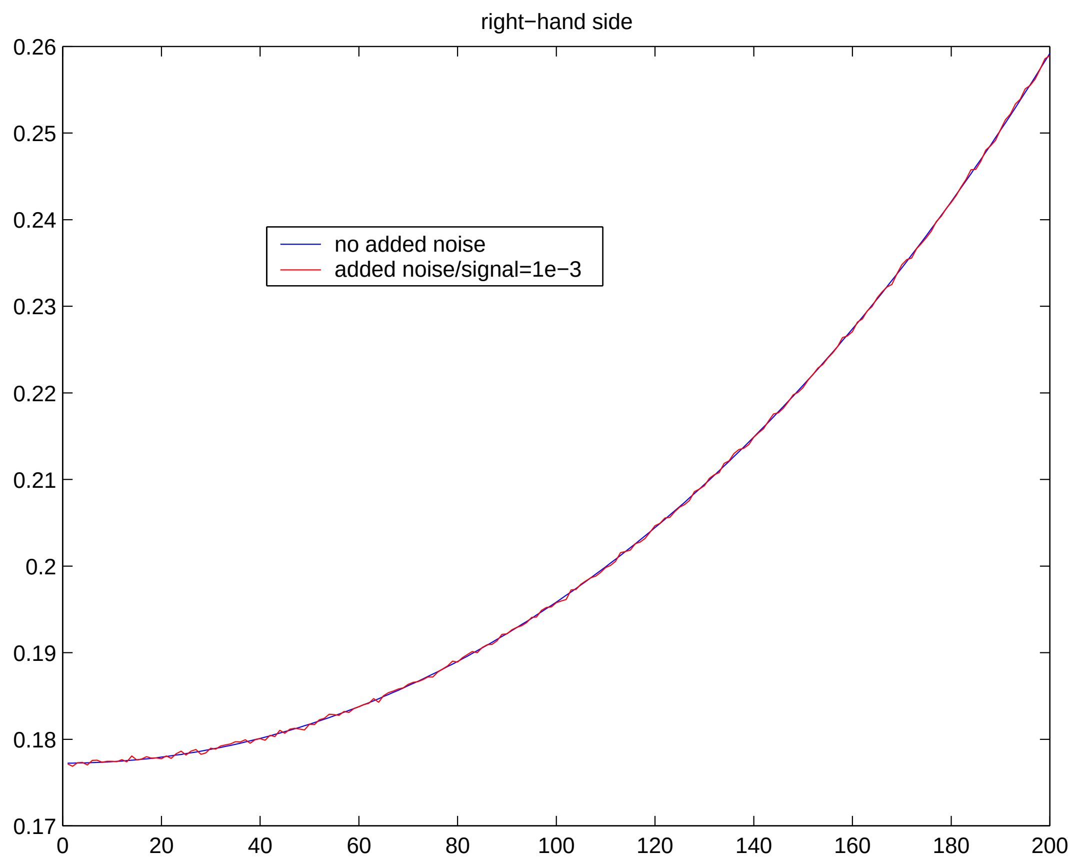
$$b = b_{\text{exact}} + e$$

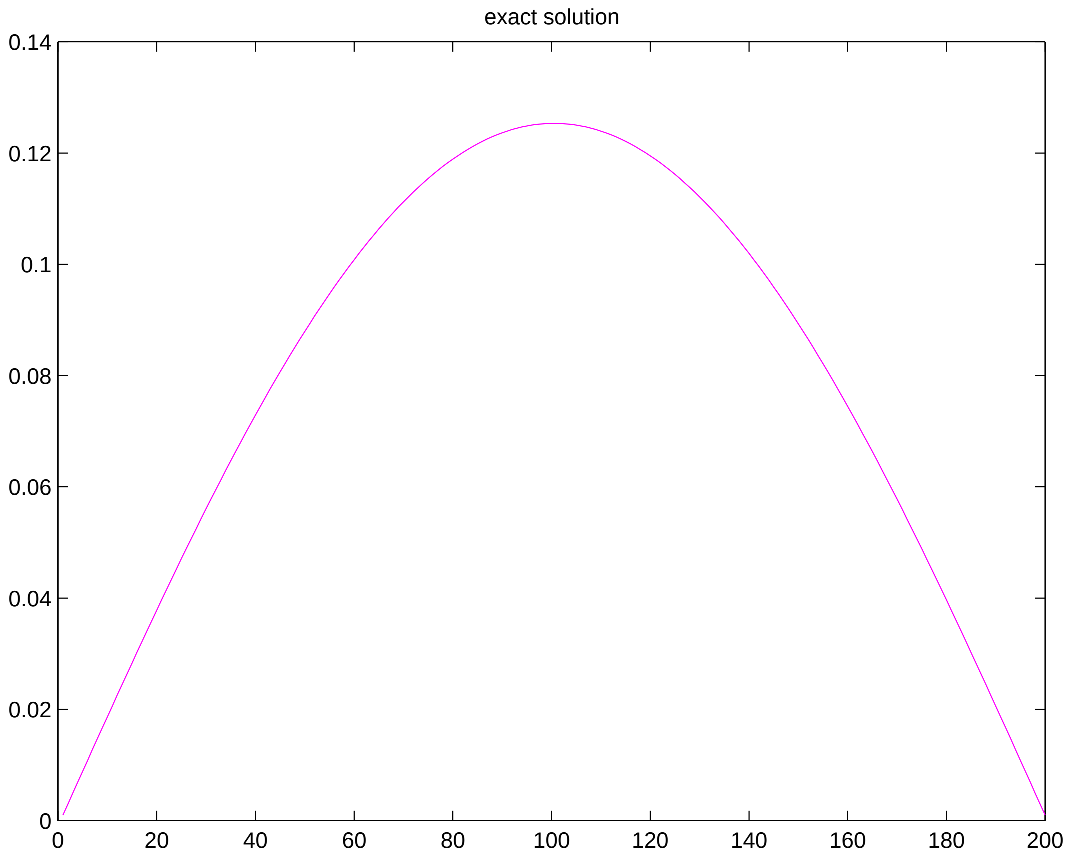
is known. Here e represents white Gaussian noise scaled to correspond to 0.1% relative noise, i.e.,

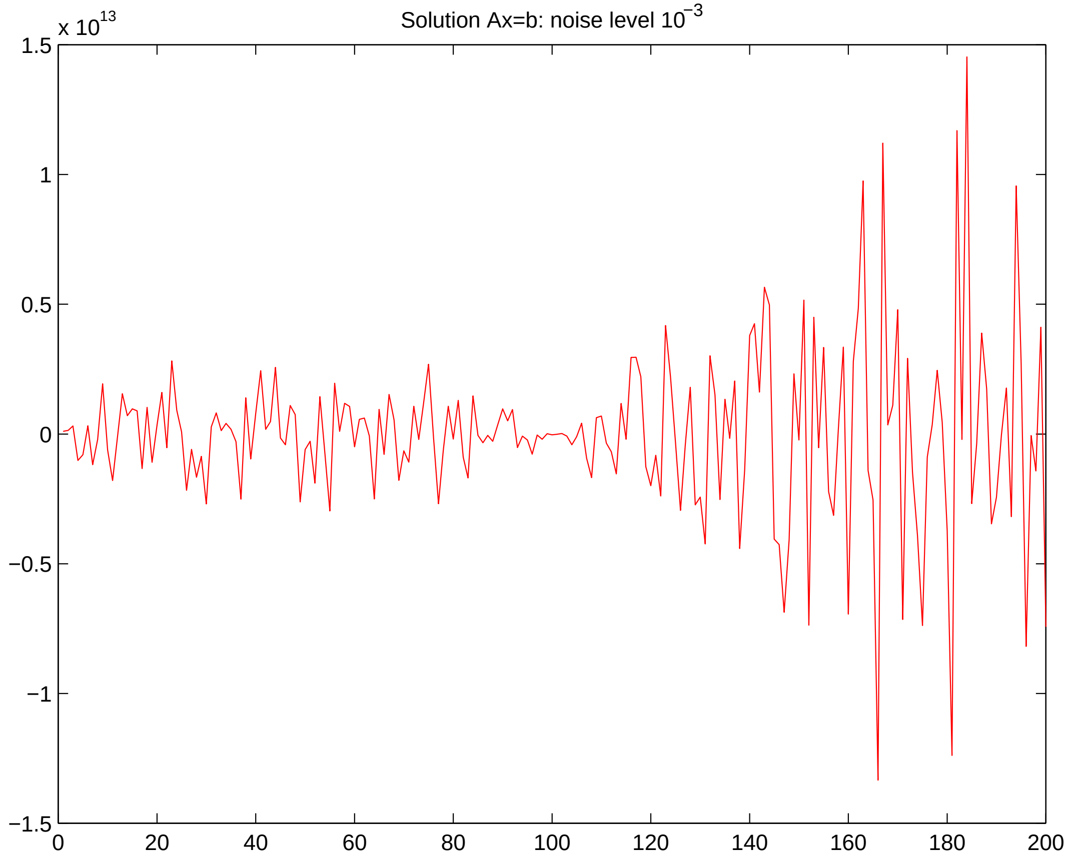
$$\|e\|_2 = 10^{-3}\|b_{\text{exact}}\|_2$$

We would like to determine an approximation of x_{exact} by solving

$$Ax = b.$$


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

Outline

- Theory:
  - Discretization of the operator equation.
  - The Arnoldi process.
  - Tikhonov regularization, taking discretization and data errors into account.
  - Some computed examples.
- Practice:
  - How the Arnoldi process actually works.
  - Preconditioning.
  - More computed examples.

The Arnoldi process, $n \gg k$:

Input: $A \in \mathbf{R}^{n \times n}$, $b \in \mathbf{R}^n \backslash \{0\}$

$v_1 := b / \|b\|$;

for $j = 1, 2, \ldots, k$ **do**

    $w := A v_j$;

    for $i = 1, 2, \ldots, j$ **do**

        $h_{i,j} := v_i^T w$; $w := w - v_i h_{i,j}$;

    end for

    $h_{j+1,j} := \|w\|$; $v_{j+1} := w / h_{j+1,j}$;

end for

Determines the Arnoldi decomposition:

$$AV_k = V_{k+1} H_{k+1,k}$$

where

$$V_k = [v_1, \dots, v_k], \quad V_{k+1} = [v_1, \dots, v_{k+1}]$$

have orthonormal columns with $v_1 = b/\|b\|$ that span the Krylov subspaces

$$\mathcal{R}(V_j) = \mathbf{K}_j(A, b) = \text{span}\{b, Ab, \dots, A^{j-1}b\}, \quad j = k, k+1$$

and

$$H_{k+1,k} = [h_{i,j}] = \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ & * & * & * & * & * \\ & & * & * & * & * \\ & & & * & * & * \\ & & & & * & * \\ & & & & & * \end{bmatrix} \in \mathbf{R}^{(k+1)\times k}$$

is upper Hessenberg.

Golub–Kahan bidiagonalization (described by Paige and Saunders):

Given A and initial vector b, k steps determines the GK decomposition

$$AV_k = U_{k+1}B_{k+1,k}, \quad A^TU_k = V_kB_{k,k}^T \qquad (1)$$

where columns of $U_{k+1} = [u_1, \ldots, u_{k+1}]$ and $V_k = [v_1, \ldots, v_k]$ orthonormal,

$\mathcal{R}(V_j) = \mathbf{K}_j(A^TA, A^Tb) = \text{span}\{A^Tb, (A^TA)Ab, \ldots, (A^TA)^{j-1}A^Tb\}$,

$u_1 = b/\|b\|$, $B_{k+1,k} \in \mathbf{R}^{(k+1)\times k}$ lower bidiagonal.

The computations of (1) requires k matrix-vector products with A and k with A^T.

Why using the Arnoldi decomposition may be attractive:

- Each step requires only one matrix-vector product evaluation
- Can be applied when the matrix that defines the linear discrete ill-posed problem is not formed, and therefore its transpose is not available:
  - fast multipole methods,
  - nonlinear problems when matrix-vector products with the Jacobian can be evaluated rapidly without computing all partial derivatives.
- The number of matrix-vector products may be fewer than for Golub–Kahan bidiagonalization.

The Arnoldi–Tikhonov method in Hilbert space:

We seek to solve

$$Ax = b$$

where A is a linear operator between Hilbert spaces, A not continuously invertible.

The solution method consists of three steps:

- Discretization of the (infinite-dimensional) operator equation. Requires estimate of the distance between the solutions in Hilbert space and in $\mathbf{R}^n$.

- Definition of a regularized finite-dimensional system. Requires estimate of the distance between the solution in $\mathbf{R}^n$ and its regularized version.
- Compute and approximate solution of the regularized system. Estimate the distance between the solution of regularized finite-dimensional system and its computed approximation.

Discretization of the operator equation:

$$A : X \to Y, \qquad X, Y \quad \text{Hilbert spaces}$$

Define finite-dimensional subspaces

$$X_n \subset X, \qquad \dim(X_n) = n$$

and

$$Y_n = AX_n.$$

Consider least-squares problem

$$\min_{x \in X_n} \|Ax - b\|_Y$$

with unique minimal-norm solution x_n.

Introduce projectors

$$P_n : X \to X_n, \qquad Q_n : Y \to Y_n.$$

Let x_n be a minimal norm solution of the linear system of equations

$$Q_n A P_n x = Q_n b.$$

Define matrix

$$A_n = Q_n A P_n.$$

Let $s_1, \ldots, s_n$ form a convenient basis for X_n. Then

$$x_n = \sum_{j=1}^{n} x_j^{(n)} s_j.$$

Let $\hat{s}_1, \ldots, \hat{s}_n$ be an orthonormal basis for X_n and let

$$[s_1, \ldots, s_n] = M_n[\hat{s}_1, \ldots, \hat{s}_n]$$

with the condition number of the matrix M_n bounded independentlyy of n. Define the vector $\vec{x}_n = [x_1^{(n)}, \ldots, x_n^{(n)}]^T \in \mathbf{R}^n$. Then

$$\|x_n\|_X \sim \|\vec{x}_n\|_2.$$

We can identify X_n with $\mathbf{R}^n$, i.e., $\vec{x}_n$ with x_n.

Let b_{exact} be the unknown error-free right-hand side associated with available right-hand side b. Let x_{exact} be the minimal-norm solution of the consistent equation

$$Ax_{\text{exact}} = b_{\text{exact}}.$$

The solution x_n of

$$\min_{x \in X_n} \|Ax - b\|$$

might not be a useful approximation of x_{exact} due to severe propagation of the error $e = b - b_{\text{exact}}$ into x_n. We have to estimate

$$\|x_{\text{exact}} - x_n\|.$$

Let $\Omega \subset \mathbf{R}^N$ and $X = L_2(\Omega)$. Define the Sobolev space $H^\beta = H^\beta(\Omega)$ for $\beta \in \mathbf{R}$. Assume that an "inverse estimate" holds:

$$\|Ax\|_Y \sim \|x\|_{H^{-\beta}} \quad \forall x \in H^{-\beta} \text{ and some } 0 < \beta < \infty, \qquad (*)$$

i.e., $A : H^{-\beta} \to Y$ is continuously invertible.

Example: Consider the mildly ill-posed Volterra integral equation of the first kind

$$[Ax](s) = \int_0^s x(t)dt, \qquad 0 \leq s \leq 1.$$

Then $(*)$ holds with $\beta = 1$. $\square$

Example: Let

$$[Ax](s) = \int_{-6}^{6} k(s-t)x(t)dt, \qquad -6 \le s \le 6,$$

with

$$k(t) = \begin{cases} 1 + \cos\left(\frac{\pi}{3}\right), & -3 \le t \le 3, \\ 0, & \text{otherwise.} \end{cases}$$

This integral operator of the test problem “phillips” in Regularization Tools by Hansen. It is a convolution in 1D. The kernel k has jump discontinuities in the second derivative at $t = \pm 3$. It follows that $\beta = 2$. $\square$

Natterer '77 shows that if an inverse estimate is fulfilled and $x_{\text{exact}} \in H^\eta$, then one obtains

$$\|x_{\text{exact}} - x_n\|_X \leq C \left(h(n)^\eta \|x_{\text{exact}}\|_{H^\eta} + h(n)^{-\beta} \delta \right),$$

where

- C is a constant independent of $h(n)$, x_{exact}, and $\delta = \|b - b_{\text{exact}}\|_Y$,
- $h = h(n) \searrow 0$ as $n \to \infty$ is a discretization parameter that depends on how well x_{exact} can be approximated by an element in X_n.

An optimal dimension of the discretized problem is given by

$$n \sim h^{-1}\left(\left(\frac{\delta}{\|x_{\text{exact}}\|_{H^\eta}}\right)^{1/(\eta+\beta)}\right)$$

and gives

$$\|x_{\text{exact}} - x_n\|_X \leq C' \|x_{\text{exact}}\|_{H^\eta}^{\beta/(\eta+\beta)} \delta^{\eta/(\eta+\beta)}$$

for some constant $C' > 0$; see Natterer '77. Spline and finite element spaces X_n allow for bounds of this kind.

Natterer proposed regularization by discretization with a suitable $h = h(n)$. Instead we will use Tikhonov regularization.

An Arnoldi–Tikhonov method:

The Arnoldi decomposition

$$AV_k = V_{k+1}H_{k+1,k}$$

suggests the approximation

$$A_k := V_{k+1}H_{k+1,k}V_k^T$$

of the matrix A. We require a bound

$$\|A - A_k\| \leq h_k$$

Such a bound can be computed with the Frobenius norm (which is easy to evaluate).

Introduce the Tikhonov functional

$$J_{\alpha,k}(x) := \|A_k x - b\|^2 + \alpha\|x\|^2,$$

where $\alpha > 0$ is a regularization parameter. Let

$$x_{\alpha,k} := \arg\min_{x\in\mathbf{R}^n} \{J_{\alpha,k}(x)\}.$$

Also define the Tikhonov functional obtained by replacing A_k by A,

$$J_\alpha(x) := \|Ax - b\|^2 + \alpha\|x\|^2$$

with solution

$$x_\alpha := \arg\min_{x\in\mathbf{R}^n} \{J_\alpha(x)\}.$$

We would like to choose α and k so that $x_{\alpha,k}$ is an accurate approximation of x_{exact}. This choice is studied by Neubauer '86. Consider the operator equation

$$Tx = b, \qquad T : \widetilde{X} \to \widetilde{Y}$$

for Hilbert spaces $\widetilde{X}$ and $\widetilde{Y}$, discretize, and solve the Tikhonov equation

$$x_{\alpha,k}^{h,\delta} := \left(T_{h,k}^T T_{h,k} + \alpha I\right)^{-1} T_{h,k}^T b.$$

Here $T_{h,k}$ denotes a discretization and modification of T with adjoint $T_{h,k}^T$.

Neubauer '86 requires that

$$\begin{aligned} \|T - T_{h,k}\|_2 &\leq h_k, \\ T_{h,k} &:= R_k T_h, \\ R_k &\to I \text{ point-wise as } k \text{ increases,} \end{aligned}$$

where R_k is an orthogonal projector onto a finite-dimensional subspace $W_k \subset \widetilde{Y}$ and T_h is a discretization of T.

In our application of the results of Neubauer '86, we let T be the matrix A and $\widetilde{X} = \widetilde{Y} = \mathbf{R}^n$. Moreover,

$$\begin{aligned} T_{h,k} &:= A_k, \qquad \|A - A_k\|_2 \leq h_k, \\ W_k &:= \mathcal{R}(A_k), \\ R_k &:= P_{\mathcal{R}(A_k)}, \end{aligned}$$

where $P_{\mathcal{R}(A_k)}$ denotes the orthogonal projector onto the range of A_k. The operator T_h is not important to us.

Theorem 1: Assume that the Arnoldi process does not break down. Then

$$
\begin{aligned}
W_k &\subset \mathcal{R}(A), \\
\mathcal{R}(R_k A_k) &= \mathcal{W}_k, \\
\|R_k(A - A_k)\| &\leq h_k, \\
\|R_k(b_{\text{exact}} - b)\| &\leq \delta, \\
R_k &\to I \ \text{ point-wise on } \mathcal{R}(A).
\end{aligned}
$$

Theorem 2: Let A be invertible and let the regularization parameter $\alpha > 0$ satisfy

$$\alpha^3 \left\langle \left(A_k A_k^* + \alpha I\right)^{-3} R_k b, R_k b \right\rangle = E\, h_k + C\, \delta^2,$$

with constants $C > 1$ and $E > 3\|x_{\text{exact}}\|$ such that

$$0 \leq E\, h_k + C\, \delta \leq \|R_k b\|.$$

Then the Tikhonov solution satisfies

$$\|x_{\alpha,k}^{h,\delta} - x_{\text{exact}}\| = \mathcal{O}\left((\delta + h_k)^{2/3}\right) + \|(I - R_k)A\|^2\, \|v\|,$$

$\|v\|$ some vector.

Corollary: For k such that

$$\max\{h_k, \|(I - R_k)A_k\|\} \sim \delta,$$

we have

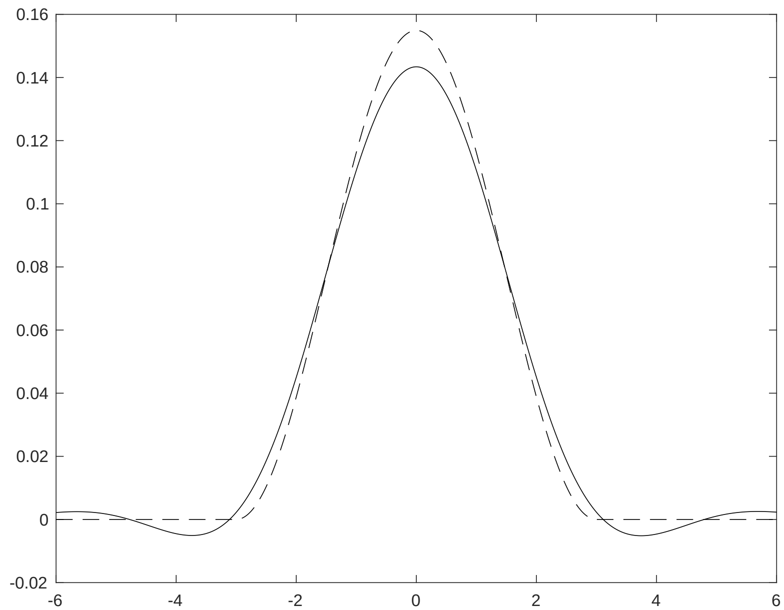
$$\|x_{\alpha,k}^{h,\delta} - x_{\text{exact}}\| = \mathcal{O}\left(\delta^{2/3}\right) \quad \text{as} \quad \delta \searrow 0.$$

Computed examples

Example. Discretization of the integral equation "phillips" from Regularization Tools by Hansen uses a Galerkin method that gives a symmetrix indefinite matrix $A \in \mathbf{R}^{1000\times 1000}$. Relative noise 0.5% gives $\delta = \|b_{\text{exact}} - b\| = 7.65 \cdot 10^{-2}$.

k	h_k	α	$\|x_{\alpha,k}^{h,\delta} - x_{\text{exact}}\|$
20	$2.10 \cdot 10^{-2}$	2.03	$3.98 \cdot 10^{-1}$
30	$5.10 \cdot 10^{-3}$	1.12	$2.78 \cdot 10^{-1}$
40	$2.51 \cdot 10^{-3}$	0.85	$2.36 \cdot 10^{-1}$
50	$1.22 \cdot 10^{-3}$	0.66	$2.04 \cdot 10^{-1}$

x_{exact} (dashed curve) and computed solution $x_{\alpha,k}^{h,\delta}$ (continuous curve) for $n = 2000$, $k = 50$.



The Arnoldi process in practice

Recall that k steps of the Arnoldi process applied to A with initial vector b yields the decomposition

$$AV_k = V_{k+1}H_{k+1,k}, \quad V_k e_1 = b/\|b\|.$$

Solution methods based on the Arnoldi process

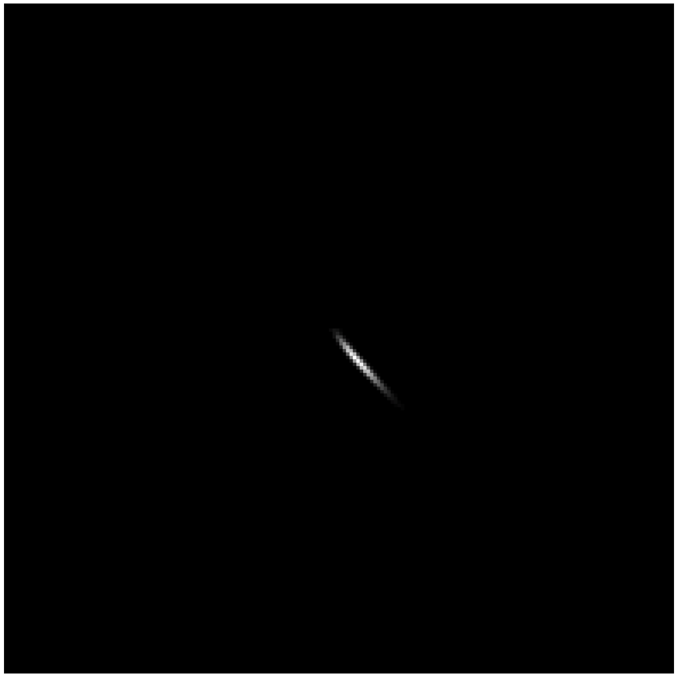
- perform well for many problems, but
- for certain problems low-dimensional solution subspaces generated by the Arnoldi process are poorly suited to represent x_{exact}.

GMRES is a popular iterative method for solving linear systems of equations.

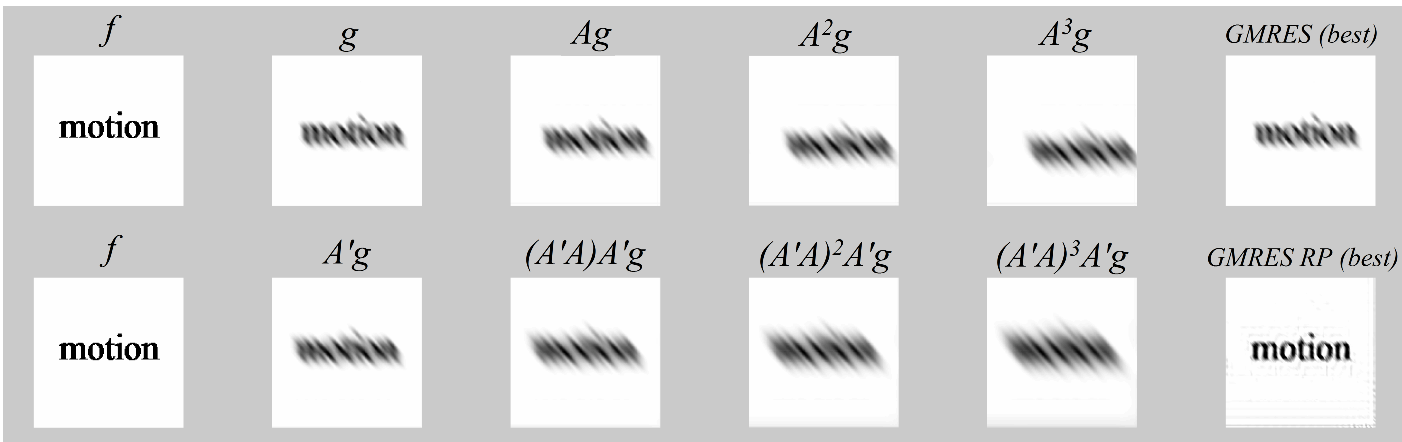
The kth iterate of GMRES applied to A with initial solution $x_0 = 0$ satisfies $x_k \in \mathbf{K}_k(A, b)$ and

$$\|Ax_k - b\| = \min_{x \in \mathbf{K}_k(A,b)} \|Ax - b\| = \min_{y \in \mathbf{R}^k} \|H_{k+1,k}y - e_1\|b\| \|.$$

Example: PSF modeling motion blur.



Sharp image f, blurred images $b = g = Af,\ A^2f,\ A^3f,\ \ldots$



Example: Let $A \in \mathbf{R}^{n\times n}$ be the downshift matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ & 0 & \ddots & \vdots & 0 & 0 \\ & & \ddots & & 0 & 0 \\ & & & 0 & 1 & 0 \end{bmatrix} \qquad b = e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The minimal-norm solution of $Ax = b$ is $x_{\text{exact}} = e_1$.
None of the Krylov subspaces
$\mathbf{K}_k(A, b) = \text{span}\{e_2, e_3, \ldots, e_{k+1}\}$, $1 \le k < n$, contain the solution.

LSQR does not have a problem: It determines approximate solutions in Krylov subspaces $\mathbf{K}_k(A^T A, A^T b)$, and

$$x_{\text{exact}} \in \mathbf{K}_1(A^T A, A^T b) = \text{span}\{A^T b\}.$$

Let $\mathbb{H}$ be the set of Hermitian matrices.

Let $\mathbb{H}_+$ be the set of Hermitian positive semidefinite matrices.

Let $\mathbb{H}_-$ be the set of Hermitian negative semidefinite matrices.

Let $\mathbb{A}$ be the set of skew-Hermitian matrices.

Let $\mathbb{N}$ be the set of normal matrices.

It is known that Arnoldi-based solution methods (such as GMRES) may converge slowly when the matrix A is far from normal.

Consider the relative distances of the downshift matrix $A \in \mathbf{C}^{n \times n}$ in the Frobenius norm:

$$\frac{\mathrm{dist}_F(A, \mathbb{H})}{\|A\|_F} = 1/\sqrt{2}$$

$$\frac{\mathrm{dist}_F(A, \mathbb{A})}{\|A\|_F} = 1/\sqrt{2}$$

$$
\begin{aligned}
\frac{\operatorname{dist}_F(A, \mathbb{H}_+)}{\|A\|_F} &= \sqrt{3}/2 \\
\frac{\operatorname{dist}_F(A, \mathbb{H}_-)}{\|A\|_F} &= \sqrt{3}/2 \\
\frac{\operatorname{dist}_F(A, \mathbb{N})}{\|A\|_F} &\leq 1/\sqrt{n}
\end{aligned}
$$

Small distance to normality is not enough to secure fast convergence.

Example: Let $A \in \mathbf{R}^{n\times n}$ be the circulant downshift matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ & 0 & \ddots & \vdots & 0 & 0 \\ & & \ddots & & 0 & 0 \\ & & & 0 & 1 & 0 \end{bmatrix} \qquad b = e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

The minimal-norm solution of $Ax = b$ is $x_{\text{exact}} = e_1$. None of the Krylov subspaces $\mathbf{K}_k(A, b) = \text{span}\{e_2, e_3, \ldots, e_{k+1}\}$, $1 \leq k < n-1$, contain the solution.

GMRES typically converges quickly when the matrix A is close to HPD and its spectrum is clustered on a small set in the complex plane away from the origin and is close to the positive real axis.

The performance of GMRES is invariant under "rotation" of the matrix, i.e., under multiplication of the matrix by $\exp(i\theta)$, $i = \sqrt{-1}$.

We define the set $\mathbb{G}$ of generalized Hermitian matrices obtained by multiplying a Hermitian matrix by $\exp(i\theta)$ and of the set $\mathbb{G}_+$ of generalized Hermitian positive semidefinite matrices obtained by multiplying a Hermitian positive semidefinite matrix by $\exp(i\theta)$.

Proposition: The matrix $A \in \mathbf{C}^{n\times n}$ is generalized Hermitian (positive semidefinite) if and only if there is $\theta \in (-\pi, \pi]$ and $\alpha \in \mathbf{C}$ such that

$$A = \exp(i\theta)B + \alpha I,$$

where B is a Hermitian (positive semidefinite) matrix.

Preconditioning techniques

Right preconditioning with the aim of making the matrix in the preconditioned system close to $\mathbb{G}_+$:

$$AMy = b, \qquad x = My.$$

The Arnoldi decomposition

$$AV_m = V_{m+1}H_{m+1,m}$$

suggests that $A \approx A_m := V_{m+1}H_{m+1,m}V_m^*$. Use preconditioners

$$M = A_m^* \qquad \text{or} \qquad M = A_m^* + (I - V_mV_m^*).$$

The downshift matrix suggests that M be chosen a periodic shift matrix (circulant). Consider

$$\min_{C \text{ circulant}} \|A - C\|_F.$$

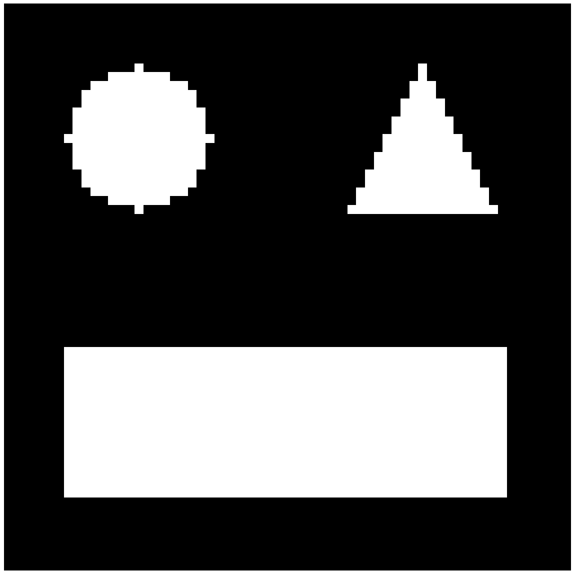
Solution $\widehat{C}$ can be computed with the FFT. Let $M := \widehat{C}^{-1}$

Alternatively, solve

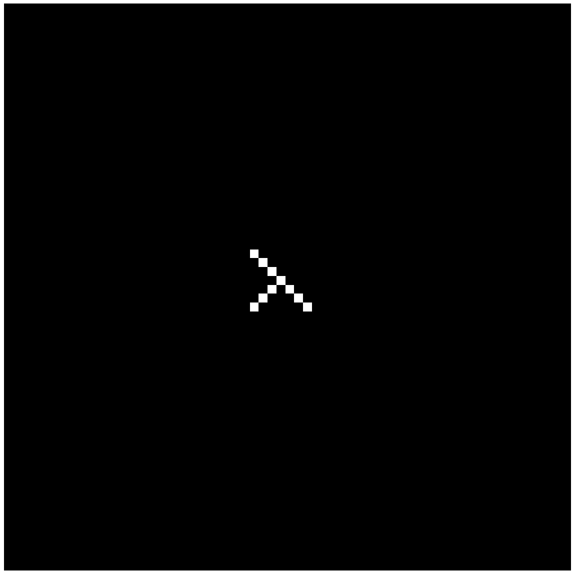
$$\min_{C \text{ circulant}} \|I - C^{-1}A\|_F.$$

Solution $\widehat{C}$ can be computed with the FFT. Let $M := \widehat{C}^{-1}$.

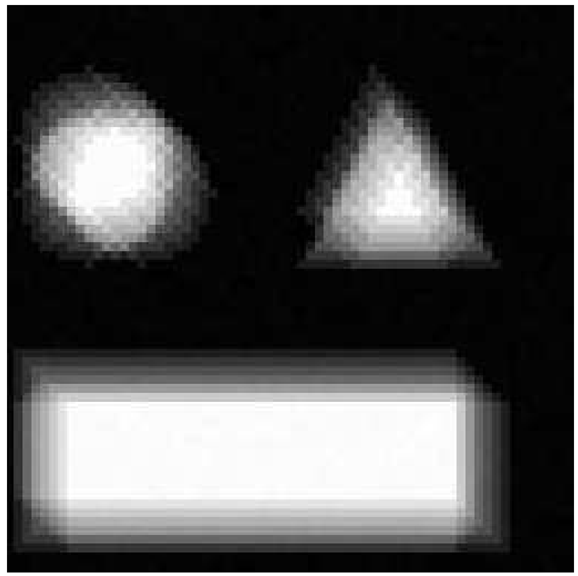
Example: Image restoration. The exact image is 64×64. It is contaminated by motion blur and 2% Gaussian noise. Original image



PSF modeling motion blur in two directions

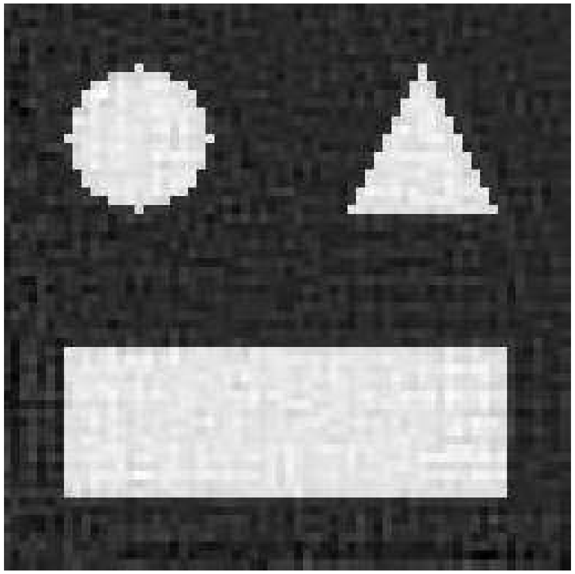


Available blur- and noise-contaminated image



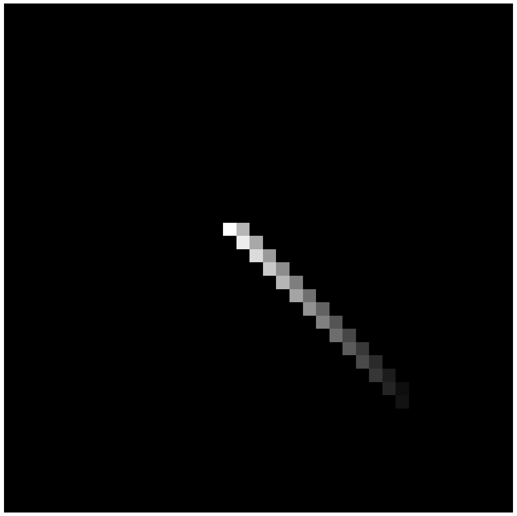
Best possible restoration determined by Arnoli-Tikhonov or GMRES

Best possible restoration determined by Arnoldi-Tikhonov or GMRES with circulant preconditioner



Example: Image restoration. The exact image is 256×256. It is contaminated by motion blur and 0.5% Gaussian noise. Original image

17×17 pixel PSF modeling motion blur



Available blur- and noise-contaminated image

Restoration determined by Arnoldi–Tikhonov with $k = 26$ steps of Arnoldi

Restoration determined by Arnoldi–Tikhonov with preconditioner

$$M = A_p^T, \quad p = 50.$$

Preconditioner reduces error by 5%.

Blow up of last restoration

Conclusions

- Solution methods for ill-posed problems can be based on the Arnoldi process.
- For many problems a solution subspace of fairly small dimension suffices.
- For certain problems, such as the restoration of images that have been contaiminated by motion blur
  - the use of a preconditioner may be beneficial,
  - regularization can be achieved by (preconditioned) Tikhonov regularization or the (preconditioned) truncated SVD method.

$\epsilon\upsilon\chi\alpha\rho\iota\sigma\tau\omega$