Automatic parameter setting for Arnoldi-Tikhonov methods

Silvia Gazzola
Joint work with Paolo Novati

Department of Mathematics
University of Padova - Italy

3rd Dolomites Workshop on Constructive Approximation and Application
Outline

1. The Arnoldi-Tikhonov method
   - The problem
   - The (standard) Arnoldi-Tikhonov (AT) method
   - The Generalized AT (GAT) method

2. The parameter selection strategy
   - Geometric interpretation

3. Examples
   - Common test problems
   - Image restoration

4. Noise level detection
   - The algorithm
   - Example

5. Final remarks
Outline

1 The Arnoldi-Tikhonov method
   - The problem
   - The (standard) Arnoldi-Tikhonov (AT) method
   - The Generalized AT (GAT) method

2 The parameter selection strategy
   - Geometric interpretation

3 Examples
   - Common test problems
   - Image restoration

4 Noise level detection
   - The algorithm
   - Example

5 Final remarks
Outline

1. The Arnoldi-Tikhonov method
   - The problem
   - The (standard) Arnoldi-Tikhonov (AT) method
   - The Generalized AT (GAT) method

2. The parameter selection strategy
   - Geometric interpretation

3. Examples
   - Common test problems
   - Image restoration

4. Noise level detection
   - The algorithm
   - Example

5. Final remarks
Outline

1. The Arnoldi-Tikhonov method
   - The problem
   - The (standard) Arnoldi-Tikhonov (AT) method
   - The Generalized AT (GAT) method

2. The parameter selection strategy
   - Geometric interpretation

3. Examples
   - Common test problems
   - Image restoration

4. Noise level detection
   - The algorithm
   - Example

5. Final remarks
Outline

1. The Arnoldi-Tikhonov method
   - The problem
   - The (standard) Arnoldi-Tikhonov (AT) method
   - The Generalized AT (GAT) method

2. The parameter selection strategy
   - Geometric interpretation

3. Examples
   - Common test problems
   - Image restoration

4. Noise level detection
   - The algorithm
   - Example

5. Final remarks
The problem

We consider linear systems of equations

\[ Ax = b, \quad A \in \mathbb{R}^{N \times N}, \quad b \in \mathbb{R}^N, \]

in which the matrix \( A \) is assumed to have singular values that rapidly decay and cluster near zero\(^1\).

We assume that

- the available right-hand side vector \( b \) is affected by noise, that is

  \[ b = \bar{b} + e, \]

  where \( \bar{b} \) represents the unknown noise-free right-hand side;

- the quantity \( \varepsilon \approx \|e\| \) is known;

\(^1\)Hansen (1998), *Rank-deficient and Discrete Ill-Posed Problems*
The problem

We consider linear systems of equations

\[ Ax = b, \quad A \in \mathbb{R}^{N \times N}, \quad b \in \mathbb{R}^N, \]

in which the matrix \( A \) is assumed to have singular values that rapidly decay and cluster near zero\(^1\).

We assume that

- the available right-hand side vector \( b \) is affected by noise, that is
  
  \[ b = \bar{b} + e, \]

  where \( \bar{b} \) represents the unknown noise-free right-hand side;

- the quantity \( \varepsilon \approx \| e \| \) is known;

\(^1\text{Hansen(1998), } \textit{Rank-deficient and Discrete Ill-Posed Problems} \)
The problem

We consider linear systems of equations

$$Ax = b, \quad A \in \mathbb{R}^{N \times N}, \quad b \in \mathbb{R}^N,$$

in which the matrix $A$ is assumed to have singular values that rapidly decay and cluster near zero$^1$.

We assume that

- the available right-hand side vector $b$ is affected by noise, that is

$$b = \bar{b} + e,$$

where $\bar{b}$ represents the unknown noise-free right-hand side;

- the quantity $\varepsilon \approx \|e\|$ is known;

---

$^1$Hansen (1998), *Rank-deficient and Discrete Ill-Posed Problems*
The Arnoldi-Tikhonov (AT) method

Given $\lambda$, $L$, $x_0$ consider the Tikhonov regularization

$$
\min_{x \in \mathbb{R}^N} \left\{ \|Ax - b\|^2 + \lambda\|L(x - x_0)\|^2 \right\}.
$$

For the special case of $L = I_N$ and $x_0 = 0$, the Arnoldi-Tikhonov method\(^2\) is based on the reduction to a problem of much smaller dimension, projecting the matrix $A$ onto the Krylov subspaces generated by $A$ and the vector $b$,

$$
K_m(A, b) = \text{span}\{b, Ab, \ldots, A^{m-1}b\}, \quad m \ll N.
$$

For the construction of the Krylov subspaces the AT method uses the Arnoldi algorithm.

\(^2\)Calvetti-Morigi-Reichel-Sgallari (2000), Tikhonov regularization and the L-curve for large discrete ill-posed problems
The Arnoldi-Tikhonov (AT) method

Given \( \lambda, L, x_0 \) consider the Tikhonov regularization

\[
\min_{x \in \mathbb{R}^N} \left\{ \| Ax - b \|^2 + \lambda \| L(x - x_0) \|^2 \right\}.
\]

For the special case of \( L = I_N \) and \( x_0 = 0 \), the Arnoldi-Tikhonov method\(^2\) is based on the reduction to a problem of much smaller dimension, projecting the matrix \( A \) onto the Krylov subspaces generated by \( A \) and the vector \( b \),

\[
\mathcal{K}_m(A, b) = \text{span}\{ b, Ab, \ldots, A^{m-1} b \}, \quad m \ll N.
\]

For the construction of the Krylov subspaces the AT method uses the Arnoldi algorithm.

\(^2\)Calvetti-Morigi-Reichel-Sgallari(2000),

Tikhonov regularization and the L-curve for large discrete ill-posed problems

S.Gazzola (University of Padova)  Arnoldi-Tikhonov methods  DWCAA 2012
The Arnoldi-Tikhonov (AT) method

Given $\lambda$, $L$, $x_0$ consider the Tikhonov regularization

$$\min_{x \in \mathbb{R}^N} \left\{ \|Ax - b\|^2 + \lambda \|L(x - x_0)\|^2 \right\}.$$

For the special case of $L = I_N$ and $x_0 = 0$, the Arnoldi-Tikhonov method\(^2\) is based on the reduction to a problem of much smaller dimension, projecting the matrix $A$ onto the Krylov subspaces generated by $A$ and the vector $b$,

$$K_m(A, b) = \text{span}\{b, Ab, \ldots, A^{m-1}b\}, \quad m \ll N.$$

For the construction of the Krylov subspaces the AT method uses the Arnoldi algorithm.

\(^2\)Calvetti-Morigi-Reichel-Sgallari (2000),

*Tikhonov regularization and the L-curve for large discrete ill-posed problems*
The Arnoldi-Tikhonov method

The Arnoldi algorithm yields the decomposition

$$AV_m = V_{m+1} \overline{H}_m,$$

where

- $V_{m+1} = [v_1, \ldots, v_{m+1}] \in \mathbb{R}^{N \times (m+1)}$ has orthonormal columns which span the Krylov subspace $\mathcal{K}_{m+1}(A, b)$;
  - $v_1$ is defined as $b/\|b\|$.
- $\overline{H}_m \in \mathbb{R}^{(m+1) \times m}$ is an upper Hessenberg matrix.

The AT method searches for approximations belonging to $\mathcal{K}_m(A, b)$. Taking $x = V_m y_m$ ($y_m \in \mathbb{R}^m$) we obtain the reduced minimization problem

$$\min_{y_m \in \mathbb{R}^m} \left\{ \| \overline{H}_m y_m - V_{m+1}^T b \|^2 + \lambda \| y_m \|^2 \right\}.$$

Parameter choice strategy: $L_m$-curve criterium.
The Arnoldi algorithm yields the decomposition

\[ AV_m = V_{m+1} \overline{H}_m, \]

where

- \( V_{m+1} = [v_1, ..., v_{m+1}] \in \mathbb{R}^{N \times (m+1)} \) has orthonormal columns which span the Krylov subspace \( \mathcal{K}_{m+1}(A, b) \);
  - \( v_1 \) is defined as \( b/\| b \| \).
- \( \overline{H}_m \in \mathbb{R}^{(m+1) \times m} \) is an upper Hessenberg matrix.

The AT method searches for approximations belonging to \( \mathcal{K}_m(A, b) \).

Taking \( x = V_m y_m \) (\( y_m \in \mathbb{R}^m \)) we obtain the reduced minimization problem

\[
\min_{y_m \in \mathbb{R}^m} \left\{ \left\| \overline{H}_m y_m - V_{m+1}^T b \right\|^2 + \lambda \| y_m \|^2 \right\}.
\]

Parameter choice strategy: \( L_m \)-curve criterium.
The Arnoldi algorithm yields the decomposition

$$AV_m = V_{m+1} \overline{H}_m,$$

where

- \( V_{m+1} = [v_1, \ldots, v_{m+1}] \in \mathbb{R}^{N \times (m+1)} \) has orthonormal columns which span the Krylov subspace \( \mathcal{K}_{m+1}(A, b) \);
- \( v_1 \) is defined as \( b/\|b\| \).
- \( \overline{H}_m \in \mathbb{R}^{(m+1) \times m} \) is an upper Hessenberg matrix.

The AT method searches for approximations belonging to \( \mathcal{K}_m(A, b) \).

Taking \( x = V_m y_m \ (y_m \in \mathbb{R}^m) \) we obtain the reduced minimization problem

$$\min_{y_m \in \mathbb{R}^m} \left\{ \| \overline{H}_m y_m - V_{m+1}^T b \|^2 + \lambda \|y_m\|^2 \right\}.$$  

Parameter choice strategy: \( L_m \)-curve criterium.
The Arnoldi-Tikhonov method

The Arnoldi algorithm yields the decomposition

\[ AV_m = V_{m+1} \overline{H}_m, \]

where

- \( V_{m+1} = [v_1, \ldots, v_{m+1}] \in \mathbb{R}^{N \times (m+1)} \) has orthonormal columns which span the Krylov subspace \( \mathcal{K}_{m+1}(A, b) \);
- \( v_1 \) is defined as \( b / \| b \| \).
- \( \overline{H}_m \in \mathbb{R}^{(m+1) \times m} \) is an upper Hessenberg matrix.

The AT method searches for approximations belonging to \( \mathcal{K}_m(A, b) \).

Taking \( x = V_m y_m \) (\( y_m \in \mathbb{R}^m \)) we obtain the reduced minimization problem

\[
\min_{y_m \in \mathbb{R}^m} \left\{ \left\| \overline{H}_m y_m - V_{m+1}^T b \right\|^2 + \lambda \left\| y_m \right\|^2 \right\}. 
\]

Parameter choice strategy: \( L_m \)-curve criterium.
The Arnoldi-Tikhonov method yields the decomposition

\[ AV_m = V_{m+1} \overline{H}_m, \]

where

- \( V_{m+1} = [v_1, \ldots, v_{m+1}] \in \mathbb{R}^{N \times (m+1)} \) has orthonormal columns which span the Krylov subspace \( \mathcal{K}_{m+1}(A, b) \);
- \( v_1 \) is defined as \( b/\|b\| \).
- \( \overline{H}_m \in \mathbb{R}^{(m+1) \times m} \) is an upper Hessenberg matrix.

The AT method searches for approximations belonging to \( \mathcal{K}_m(A, b) \).

Taking \( x = V_m y_m \) \( (y_m \in \mathbb{R}^m) \) we obtain the reduced minimization problem

\[ \min_{y_m \in \mathbb{R}^m} \left\{ \| \overline{H}_m y_m - V_{m+1}^T b \|^2 + \lambda \| y_m \|^2 \right\}. \]

Parameter choice strategy: \( L_m \)-curve criterium.
The Generalized Arnoldi-Tikhonov (GAT) method

Extension of the AT method in order to work with a general regularization operator $L \neq I_N$ and with an arbitrary starting vector $x_0$.

Starting from

$$\min_{x \in \mathbb{R}^N} \left\{ \|Ax - b\|^2 + \lambda \|L(x - x_0)\|^2 \right\},$$

we search for approximations of the type

$$x_m = x_0 + V_m y_m,$$

where the columns of $V_m \in \mathbb{R}^{N \times m}$ span the Krylov subspace $K_m(A, r_0)$, where $r_0 = b - Ax_0$.

We obtain the reduced minimization problem

$$\min_{y_m \in \mathbb{R}^m} \left\{ \|AV_m y_m - r_0\|^2 + \lambda \|LV_m y_m\|^2 \right\} = \min_{y_m \in \mathbb{R}^m} \left\{ \|\overline{H}_m y_m - \|r_0\| e_1\|^2 + \lambda \|LV_m y_m\|^2 \right\} = \min_{y_m \in \mathbb{R}^m} \left\| \left( \frac{\overline{H}_m}{\sqrt{\lambda}LV_m} \right) y_m - \left( \begin{array}{c} \|r_0\| e_1 \\ 0 \end{array} \right) \right\|^2.$$
The Generalized Arnoldi-Tikhonov (GAT) method

Extension of the AT method in order to work with a general regularization operator \( L \neq I_N \) and with an arbitrary starting vector \( x_0 \).

Starting from

\[
\min_{x \in \mathbb{R}^N} \left\{ \|Ax - b\|^2 + \lambda \|L(x - x_0)\|^2 \right\},
\]

we search for approximations of the type

\[
x_m = x_0 + V_m y_m,
\]

where the columns of \( V_m \in \mathbb{R}^{N \times m} \) span the Krylov subspace \( \mathcal{K}_m(A, r_0) \), where \( r_0 = b - Ax_0 \).

We obtain the reduced minimization problem

\[
\min_{y_m \in \mathbb{R}^m} \left\{ \|AV_m y_m - r_0\|^2 + \lambda \|LV_m y_m\|^2 \right\}
= \min_{y_m \in \mathbb{R}^m} \left\{ \|\overline{H} y_m - \|r_0\| e_1\|^2 + \lambda \|LV_m y_m\|^2 \right\}
= \min_{y_m \in \mathbb{R}^m} \left\| \begin{pmatrix} \overline{H} \\ \sqrt{\lambda} LV_m \end{pmatrix} y_m - \begin{pmatrix} \|r_0\| e_1 \\ 0 \end{pmatrix} \right\|^2.
\]
The Generalized Arnoldi-Tikhonov (GAT) method

Extension of the AT method in order to work with a general regularization operator $L \neq I_N$ and with an arbitrary starting vector $x_0$. Starting from

$$\min_{x \in \mathbb{R}^N} \left\{ \|Ax - b\|^2 + \lambda \|L(x - x_0)\|^2 \right\},$$

we search for approximations of the type

$$x_m = x_0 + V_m y_m,$$

where the columns of $V_m \in \mathbb{R}^{N \times m}$ span the Krylov subspace $K_m(A, r_0)$, where $r_0 = b - Ax_0$.

We obtain the reduced minimization problem

$$\min_{y_m \in \mathbb{R}^m} \left\{ \|AV_m y_m - r_0\|^2 + \lambda \|LV_m y_m\|^2 \right\}$$

$$= \min_{y_m \in \mathbb{R}^m} \left\{ \|\overline{H} y_m - \|r_0\| e_1\|^2 + \lambda \|LV_m y_m\|^2 \right\}$$

$$= \min_{y_m \in \mathbb{R}^m} \left\| \left( \frac{\overline{H}}{\sqrt{\lambda} LV_m} \right) y_m - \left( \frac{\|r_0\| e_1}{0} \right) \right\|^2.$$
The Arnoldi-Tikhonov method

The parameter selection strategy

- Examples
- Noise level detection
- Final remarks

The parameter selection strategy

The discrepancy principle is satisfied as soon as

$$
\phi_m(\lambda) := \| b - Ax_{m,\lambda} \| \leq \eta \varepsilon, \quad \eta \gtrapprox 1.
$$

For the GAT method the approximations are $x_{m,\lambda} = x_0 + V_m y_{m,\lambda}$ and the discrepancy can be rewritten as

$$
\| b - Ax_{m,\lambda} \| = \| r_0 - AV_m y_{m,\lambda} \| = \| c - \overline{H} y_{m,\lambda} \|,
$$

where $c = V_{m+1}^T r_0 = \| r_0 \| e_1 \in \mathbb{R}^{m+1}$. Since $y_{m,\lambda}$ solves the normal equation

$$
(\overline{H}_m^T \overline{H}_m + \lambda V_m^T L^T L V_m) y_{m,\lambda} = \overline{H}_m^T c,
$$

we obtain

$$
\phi_m(\lambda) = \| \overline{H}_m (\overline{H}_m^T \overline{H}_m + \lambda V_m^T L^T L V_m)^{-1} \overline{H}_m^T c - c \|.
$$
The discrepancy principle is satisfied as soon as

\[ \phi_m(\lambda) := \| b - Ax_{m,\lambda} \| \leq \eta \varepsilon, \quad \eta \gtrapprox 1. \]

For the GAT method the approximations are \( x_{m,\lambda} = x_0 + V_m y_{m,\lambda} \) and the discrepancy can be rewritten as

\[ \| b - Ax_{m,\lambda} \| = \| r_0 - AV_{m} y_{m,\lambda} \| = \| c - \overline{H}_m y_{m,\lambda} \|, \]

where \( c = V_{m+1}^T r_0 = \| r_0 \| e_1 \in \mathbb{R}^{m+1} \).

Since \( y_{m,\lambda} \) solves the normal equation

\[ (\overline{H}_m^T \overline{H}_m + \lambda V_m^T L^T L V_m) y_{m,\lambda} = \overline{H}_m^T c, \]

we obtain

\[ \phi_m(\lambda) = \left\| \overline{H}_m (\overline{H}_m^T \overline{H}_m + \lambda V_m^T L^T L V_m)^{-1} \overline{H}_m^T c - c \right\|. \]
The parameter selection strategy

The discrepancy principle is satisfied as soon as

\[ \phi_m(\lambda) := \| b - Ax_{m,\lambda} \| \leq \eta \varepsilon, \quad \eta \gtrsim 1. \]

For the GAT method the approximations are \( x_{m,\lambda} = x_0 + V_m y_{m,\lambda} \) and the discrepancy can be rewritten as

\[ \| b - Ax_{m,\lambda} \| = \| r_0 - AV_{m} y_{m,\lambda} \| = \| c - \overline{H}_m y_{m,\lambda} \|, \]

where \( c = V_{m+1}^T r_0 = \| r_0 \| e_1 \in \mathbb{R}^{m+1} \).

Since \( y_{m,\lambda} \) solves the normal equation

\[ (\overline{H}_m^T \overline{H}_m + \lambda V_m^T L^T L V_m) y_{m,\lambda} = \overline{H}_m^T c, \]

we obtain

\[ \phi_m(\lambda) = \left\| \overline{H}_m (\overline{H}_m^T \overline{H}_m + \lambda V_m^T L^T L V_m)^{-1} \overline{H}_m^T c - c \right\|. \]
The Arnoldi-Tikhonov method

The parameter selection strategy

Examples

Noise level detection

Final remarks

Standard approach\(^3\): solve with respect to \(\lambda\) the nonlinear equation

\[
\phi_m(\lambda) = \eta \varepsilon.
\]

New approach:
Basic hypothesis: the discrepancy can be well approximated by

\[
\phi_m(\lambda) \approx \alpha_m + \lambda \beta_m
\]

in which \(\alpha_m, \beta_m \in \mathbb{R}\) can be easily computed.

Definition of \(\alpha_m\). The Taylor expansion of \(\phi_m(\lambda)\) suggests to chose

\[
\alpha_m = \phi_m(0) = \left\| \overline{H}_m (\overline{H}_m^T \overline{H}_m)^{-1} \overline{H}_m^T c - c \right\|
\]

which is just the residual of the GMRES (computed working in reduced dimension).

\(^3\text{Reichel-Shyshkov}(2008), \text{A new zero-finder for Tikhonov regularization.}\)

\(^4\text{Lewis-Reichel}(2009), \text{Arnoldi-Tikhonov regularization methods.}\)
Standard approach\(^3\):  
solve with respect to \( \lambda \) the nonlinear equation  
\[
\phi_m(\lambda) = \eta \varepsilon.
\]

New approach:  
Basic hypothesis: the discrepancy can be well approximated by  
\[
\phi_m(\lambda) \approx \alpha_m + \lambda \beta_m
\]
  
in which \( \alpha_m, \beta_m \in \mathbb{R} \) can be easily computed.  

Definition of \( \alpha_m \). The Taylor expansion of \( \phi_m(\lambda) \) suggests to chose  
\[
\alpha_m = \phi_m(0) = \left\| \bar{H}_m(\bar{H}_m^T \bar{H}_m)^{-1} \bar{H}_m^T c - c \right\|
\]
  
which is just the residual of the GMRES (computed working in reduced dimension).

\(^3\)Reichel-Shyshkov(2008), A new zero-finder for Tikhonov regularization.  
\(^4\)Lewis-Reichel(2009), Arnoldi-Tikhonov regularization methods.
**Standard approach**: solve with respect to $\lambda$ the nonlinear equation

$$\phi_m(\lambda) = \eta \epsilon.$$

**New approach:**
Basic hypothesis: the discrepancy can be well approximated by

$$\phi_m(\lambda) \approx \alpha_m + \lambda \beta_m$$

in which $\alpha_m, \beta_m \in \mathbb{R}$ can be easily computed.

**Definition of $\alpha_m$.** The Taylor expansion of $\phi_m(\lambda)$ suggests to chose

$$\alpha_m = \phi_m(0) = \left\| \overline{H}_m (\overline{H}_m^T \overline{H}_m)^{-1} \overline{H}_m^T c - c \right\|,$$

which is just the residual of the GMRES (computed working in reduced dimension).

---


4 Lewis-Reichel (2009), *Arnoldi-Tikhonov regularization methods.*
The Arnoldi-Tikhonov method

The parameter selection strategy

Examples

Noise level detection

Final remarks

---

- **Standard approach**: solve with respect to $\lambda$ the nonlinear equation
  \[ \phi_m(\lambda) = \eta \varepsilon. \]

- **New approach**: Basic hypothesis: the discrepancy can be well approximated by
  \[ \phi_m(\lambda) \approx \alpha_m + \lambda \beta_m \]
  in which $\alpha_m, \beta_m \in \mathbb{R}$ can be easily computed.

  **Definition of $\alpha_m$**: The Taylor expansion of $\phi_m(\lambda)$ suggests to chose
  \[ \alpha_m = \phi_m(0) = \left\| \overline{H}_m (\overline{H}_m^T \overline{H}_m)^{-1} \overline{H}_m^T c - c \right\|, \]
  which is just the residual of the GMRES (computed working in reduced dimension).

---

4. **Lewis-Reichel (2009)**, *Arnoldi-Tikhonov regularization methods*. 

---

S. Gazzola (University of Padova)  
Arnoldi-Tikhonov methods  
DWCAA 2012 8 / 20
Standard approach\(^3\) \(^4\):

solve with respect to \(\lambda\) the nonlinear equation

\[
\phi_m(\lambda) = \eta \varepsilon.
\]

New approach:

Basic hypothesis: the discrepancy can be well approximated by

\[
\phi_m(\lambda) \approx \alpha_m + \lambda \beta_m
\]

in which \(\alpha_m, \beta_m \in \mathbb{R}\) can be easily computed.

Definition of \(\alpha_m\). The Taylor expansion of \(\phi_m(\lambda)\) suggests to chose

\[
\alpha_m = \phi_m(0) = \left\| \overline{H}_m (\overline{H}_m^T \overline{H}_m)^{-1} \overline{H}_m^T c - c \right\|,
\]

which is just the residual of the GMRES (computed working in reduced dimension).

\(^3\)Reichel-Shyshkov (2008), A new zero-finder for Tikhonov regularization.

\(^4\)Lewis-Reichel (2009), Arnoldi-Tikhonov regularization methods.
The Arnoldi-Tikhonov method

The parameter selection strategy

Examples

Noise level detection

Final remarks

- **Standard approach**: 
  solve with respect to \( \lambda \) the nonlinear equation 
  \[ \phi_m(\lambda) = \eta \varepsilon. \]

- **New approach**: 
  Basic hypothesis: the discrepancy can be well approximated by 
  \[ \phi_m(\lambda) \approx \alpha_m + \lambda \beta_m \]
  in which \( \alpha_m, \beta_m \in \mathbb{R} \) can be easily computed.

**Definition of** \( \alpha_m \). 
The Taylor expansion of \( \phi_m(\lambda) \) suggests to chose 
\[ \alpha_m = \phi_m(0) = \left\| \overline{H}_m(\overline{H}_m^T \overline{H}_m)^{-1} \overline{H}_m^T c - c \right\|, \]
which is just the residual of the GMRES (computed working in reduced dimension).

---


4 Lewis-Reichel (2009), *Arnoldi-Tikhonov regularization methods.*
Suppose that, at step $m$, we have used the parameter $\lambda_{m-1}$ (computed at the previous step or, if $m = 1$, given by the user) to compute $y_{m,\lambda_{m-1}}$, by solving the reduced minimization, and the corresponding discrepancy

$$\phi_m(\lambda_{m-1}) = \|c - H_{m}y_{m,\lambda_{m-1}}\|.$$  

Using the linear approximation, we obtain

$$\beta_m \approx \frac{\phi_m(\lambda_{m-1}) - \alpha_m}{\lambda_{m-1}}.$$  

To select $\lambda_m$ for the next step of the Arnoldi algorithm we impose $\phi_m(\lambda_m) = \eta \varepsilon$, and force the approximation

$$\phi_m(\lambda_m) \approx \alpha_m + \lambda_m \beta_m.$$  

Hence we define

$$\lambda_m = \frac{\eta \varepsilon - \alpha_m}{\phi_m(\lambda_{m-1}) - \alpha_m} \lambda_{m-1}.$$
Suppose that, at step $m$, we have used the parameter $\lambda_{m-1}$ (computed at the previous step or, if $m = 1$, given by the user) to compute $y_{m,\lambda_{m-1}}$, by solving the reduced minimization, and the corresponding discrepancy

$$\phi_m(\lambda_{m-1}) = \| c - \overline{H}_m y_{m,\lambda_{m-1}} \|.$$  

Using the linear approximation, we obtain

$$\beta_m \approx \frac{\phi_m(\lambda_{m-1}) - \alpha_m}{\lambda_{m-1}}.$$

To select $\lambda_m$ for the next step of the Arnoldi algorithm we impose $\phi_m(\lambda_m) = \eta \varepsilon$, and force the approximation

$$\phi_m(\lambda_m) \approx \alpha_m + \lambda_m \beta_m.$$

Hence we define

$$\lambda_m = \frac{\eta \varepsilon - \alpha_m}{\phi_m(\lambda_{m-1}) - \alpha_m} \lambda_{m-1}.$$
Suppose that, at step \( m \), we have used the parameter \( \lambda_{m-1} \) (computed at the previous step or, if \( m = 1 \), given by the user) to compute \( y_{m,\lambda_{m-1}} \), by solving the reduced minimization, and the corresponding discrepancy

\[
\phi_m(\lambda_{m-1}) = \| c - \bar{H}_{m}y_{m,\lambda_{m-1}} \|.
\]

Using the linear approximation, we obtain

\[
\beta_m \approx \frac{\phi_m(\lambda_{m-1}) - \alpha_m}{\lambda_{m-1}}.
\]

To select \( \lambda_m \) for the next step of the Arnoldi algorithm we impose \( \phi_m(\lambda_m) = \eta \varepsilon \), and force the approximation

\[
\phi_m(\lambda_m) \approx \alpha_m + \lambda_m \beta_m.
\]

Hence we define

\[
\lambda_m = \frac{\eta \varepsilon - \alpha_m}{\phi_m(\lambda_{m-1}) - \alpha_m} \lambda_{m-1}.
\]
Geometric interpretation

We know that $\phi_m(\lambda)$ is a monotonically increasing function such that $\phi_m(0) = \alpha_m$. Hence, the linear function

$$f(\lambda) = \alpha_m + \lambda \left( \frac{\phi_m(\lambda_{m-1}) - \alpha_m}{\lambda_{m-1}} \right),$$

interpolates $\phi_m(\lambda)$ at 0 and $\lambda_{m-1}$, and the new parameter $\lambda_m$ is obtained by solving $f(\lambda) = \eta \varepsilon$.

- The method is just a secant method in which the leftmost point is $(0, \alpha_m)$.
- In the very first steps we may have $\alpha_m > \eta \varepsilon$. In this situation the result of the method may be negative and therefore we use

$$\lambda_m = \frac{\eta \varepsilon - \alpha_m}{\phi_m(\lambda_{m-1}) - \alpha_m} \lambda_{m-1}.$$
Examples

Test problem $\text{shaw}^5$

- $A \in \mathbb{R}^{200 \times 200}$ symmetric; $\text{cond}(A) \simeq 10^{20}$;
- noise level $\tilde{\epsilon} = \|e\|/\|\bar{b}\| = 10^{-3}$;
- $L = I_{200}$;
- $\eta = 1.001$, $\lambda_0 = 1$ $x_0 = 0$ (as always).

- new method
  - $L_m$-curve method
  - $\phi_8(\lambda) = \eta \epsilon$ by Newton’s method

$^5$Hansen (1994), *Regularization Tools*
Comparison of the $L_m$-curve method and of the secant update approach at each iteration. Discrepancy principle satisfied after 8 iterations; we compute 8 extra iterations. The secant approach exhibits a very stable behavior.

○ is the $L_m$-curve method, * is the new method.
Stability with respect to the choice of the initial value $\lambda_0$. We choose $\lambda_0 = 0.1, 0.5, 1, 10, 50$.

Remark: at the beginning we just force $\lambda_m$ to be positive; after a few iterations (when $\alpha_m < \eta \varepsilon$) the new approach is a zero-finder.
Image restoration

- Test image: *peppers.png*, size $256 \times 256$ pixels; $N = 65536$.
- Noise level: $\tilde{\varepsilon} = 10^{-2}$.
- Gaussian blur: $\sigma = 2.5$, $q = 6$.
- Regularization matrix: $D_2 \in \mathbb{R}^{(N-2) \times N}$.

- $L_m$ method
- secant update
(a) original image  
(b) blurred and noisy image  
(c) restored with $L_m$-curve approach  
(d) restored with secant approach
Noise level detection

Assumption: $\varepsilon$ is overestimated by a quantity $\overline{\varepsilon}$.

Therefore, applying the GAT method we can fully satisfy the discrepancy principle (even with $\eta = 1$),

$$\phi_m(\lambda_{m-1}) < \overline{\varepsilon}.$$ 

Applying the secant update method the discrepancy would then stabilize around $\overline{\varepsilon}$.

We define $\overline{\varepsilon} = \phi_m(\lambda_{m-1})$ as the new approximation of the noise.

We restart the GAT method immediately with the Krylov subspace $\mathcal{K}_\ell(A, b - Ax_m, \lambda_{m-1})$, where $x_m, \lambda_{m-1}$ is the last approximation obtained.

We proceed until the discrepancy is almost constant.
Assumption: $\varepsilon$ is overestimated by a quantity $\overline{\varepsilon}$.
Therefore, applying the GAT method we can fully satisfy the discrepancy principle (even with $\eta = 1$),

$$\phi_m(\lambda_{m-1}) < \overline{\varepsilon}.$$

Applying the secant update method the discrepancy would then stabilize around $\overline{\varepsilon}$.
We define $\overline{\varepsilon} = \phi_m(\lambda_{m-1})$ as the new approximation of the noise.
We restart the GAT method immediately with the Krylov subspace $\mathcal{K}_\ell(A, b - Ax_m, \lambda_{m-1})$, where $x_m, \lambda_{m-1}$ is the last approximation obtained.
We proceed until the discrepancy is almost constant.
Assumption: $\varepsilon$ is overestimated by a quantity $\overline{\varepsilon}$. Therefore, applying the GAT method we can fully satisfy the discrepancy principle (even with $\eta = 1$),

$$\phi_m(\lambda_{m-1}) < \overline{\varepsilon}.$$  

Applying the secant update method the discrepancy would then stabilize around $\overline{\varepsilon}$. We define $\overline{\varepsilon} = \phi_m(\lambda_{m-1})$ as the new approximation of the noise.

We restart the GAT method immediately with the Krylov subspace $\mathcal{K}_\ell(A, b - Ax_m, \lambda_{m-1})$, where $x_m, \lambda_{m-1}$ is the last approximation obtained. We proceed until the discrepancy is almost constant.
Assumption: $\varepsilon$ is overestimated by a quantity $\bar{\varepsilon}$. Therefore, applying the GAT method we can fully satisfy the discrepancy principle (even with $\eta = 1$),

$$\phi_m(\lambda_{m-1}) < \bar{\varepsilon}.$$ 

Applying the secant update method the discrepancy would then stabilize around $\bar{\varepsilon}$. We define $\bar{\varepsilon} = \phi_m(\lambda_{m-1})$ as the new approximation of the noise. We restart the GAT method immediately with the Krylov subspace $\mathcal{K}_\ell(A, b - Ax_m, \lambda_{m-1})$, where $x_m, \lambda_{m-1}$ is the last approximation obtained. We proceed until the discrepancy is almost constant.
Assumption: $\varepsilon$ is overestimated by a quantity $\overline{\varepsilon}$.

Therefore, applying the GAT method we can fully satisfy the discrepancy principle (even with $\eta = 1$),

$$\phi_m(\lambda_{m-1}) < \overline{\varepsilon}.$$ 

Applying the secant update method the discrepancy would then stabilize around $\overline{\varepsilon}$.

We define $\overline{\varepsilon} = \phi_m(\lambda_{m-1})$ as the new approximation of the noise.

We restart the GAT method immediately with the Krylov subspace $\mathcal{K}_\ell(A, b - Ax_m, \lambda_{m-1})$, where $x_m, \lambda_{m-1}$ is the last approximation obtained.

We proceed until the discrepancy is almost constant.
The algorithm

Input: $A$, $b$, $L$, $\lambda^{(0)}$, $\eta$, $\delta$ (threshold parameter), and $\varepsilon_0 = \bar{\varepsilon} > \varepsilon$.

Define $x^{(0)} = 0$.

For $k = 1, 2, \ldots$ until

$$\frac{\|\varepsilon_k - \varepsilon_{k-1}\|}{\|\varepsilon_{k-1}\|} \leq \delta$$

1. Apply GAT method with $x_0 = x^{(k-1)}$, $\varepsilon = \varepsilon_{k-1}$, $\lambda_0 = \lambda^{(k-1)}$. Let $x^{(k)}$ be the last approximation achieved, $\phi^{(k)}$ the corresponding discrepancy norm, and $\lambda^{(k)}$ the last parameter value;

2. Define $\varepsilon_k = \phi^{(k)}$;

3. Define $\lambda^{(k)} = \frac{\phi^{(k)}}{\phi^{(k-1)}} \lambda^{(k)}$. 

S.Gazzola (University of Padova)
The algorithm

Input: $A$, $b$, $L$, $\lambda^{(0)}$, $\eta$, $\delta$ (threshold parameter), and $\varepsilon_0 = \overline{\varepsilon} > \varepsilon$.
Define $x^{(0)} = 0$.
For $k = 1, 2, \ldots$ until

$$\frac{\|\varepsilon_k - \varepsilon_{k-1}\|}{\|\varepsilon_{k-1}\|} \leq \delta$$

1. Apply GAT method with $x_0 = x^{(k-1)}$, $\varepsilon = \varepsilon_{k-1}$, $\lambda_0 = \lambda^{(k-1)}$.
Let $x^{(k)}$ be the last approximation achieved, $\phi^{(k)}$ the corresponding discrepancy norm, and $\lambda^{(k)}$ the last parameter value;
2. Define $\varepsilon_k = \phi^{(k)}$;
3. Define $\lambda^{(k)} = \frac{\phi^{(k)}}{\phi^{(k-1)}} \lambda^{(k)}$. 

S. Gazzola (University of Padova)
The algorithm

Input: $A$, $b$, $L$, $\lambda^{(0)}$, $\eta$, $\delta$ (threshold parameter), and $\varepsilon_0 = \bar{\varepsilon} > \varepsilon$.

Define $x^{(0)} = 0$.

For $k = 1, 2, \ldots$ until

$$\frac{\|\varepsilon_k - \varepsilon_{k-1}\|}{\|\varepsilon_{k-1}\|} \leq \delta$$

1. Apply GAT method with $x_0 = x^{(k-1)}$, $\varepsilon = \varepsilon_{k-1}$, $\lambda_0 = \lambda^{(k-1)}$.
Let $x^{(k)}$ be the last approximation achieved, $\phi^{(k)}$ the corresponding discrepancy norm, and $\lambda^{(k)}$ the last parameter value;

2. Define $\varepsilon_k = \phi^{(k)}$;

3. Define $\lambda^{(k)} = \frac{\phi^{(k)}}{\phi^{(k-1)}} \lambda^{(k)}$. 
The algorithm

Input: \( A, b, L, \lambda^{(0)}, \eta, \delta \) (threshold parameter), and \( \varepsilon_0 = \bar{\varepsilon} > \varepsilon \).
Define \( x^{(0)} = 0 \).
For \( k = 1, 2, \ldots \) until

\[
\frac{||\varepsilon_k - \varepsilon_{k-1}||}{||\varepsilon_{k-1}||} \leq \delta
\]

1. Apply GAT method with \( x_0 = x^{(k-1)} \), \( \varepsilon = \varepsilon_{k-1} \), \( \lambda_0 = \lambda^{(k-1)} \). Let \( x^{(k)} \) be the last approximation achieved, \( \phi^{(k)} \) the corresponding discrepancy norm, and \( \lambda^{(k)} \) the last parameter value;
2. Define \( \varepsilon_k = \phi^{(k)} \);
3. Define \( \lambda^{(k)} = \frac{\phi^{(k)}}{\phi^{(k-1)}} \lambda^{(k)} \).
The algorithm

Input: $A$, $b$, $L$, $\lambda^{(0)}$, $\eta$, $\delta$ (threshold parameter), and $\varepsilon_0 = \bar{\varepsilon} > \varepsilon$.
Define $x^{(0)} = 0$.
For $k = 1, 2, \ldots$ until

$$\frac{\|\varepsilon_k - \varepsilon_{k-1}\|}{\|\varepsilon_{k-1}\|} \leq \delta$$

1. Apply GAT method with $x_0 = x^{(k-1)}$, $\varepsilon = \varepsilon_{k-1}$, $\lambda_0 = \lambda^{(k-1)}$.
Let $x^{(k)}$ be the last approximation achieved, $\phi^{(k)}$ the corresponding discrepancy norm, and $\lambda^{(k)}$ the last parameter value;

2. Define $\varepsilon_k = \phi^{(k)}$;

3. Define $\lambda^{(k)} = \frac{\phi^{(k)}}{\phi^{(k-1)}} \lambda^{(k)}$. 
Example

- Test image mri.tif, 128 pixels.
- Gaussian blurring, \( \sigma = 1.5 \) and \( q = 6 \).
- \( \varepsilon / \| b \| = 10^{-3} \)
- \( \overline{\varepsilon} / \| b \| = 10^{-2} \).
- \( \delta = 0.01 \)
- Regularization matrix: \( D_1 \).

- Without step 3;
  56 restarts,
  \( \varepsilon_{24} / \| b \| = 1.03 \cdot 10^{-3} \).
- With step 3;
  24 restarts,
  \( \varepsilon_{56} / \| b \| = 1.05 \cdot 10^{-3} \).
(i) original; (ii) noisy and blurred; (iii) after 4 steps; (iv) after 24 restarts.
Final remarks

- simple and efficient (all the extra computations are performed in reduced dimension);
- simultaneously determine the regularization parameter and the number of iterations;
- can be generalized to the multi-parameter case (S.G., P. Novati, *Multi-parameter Arnoldi-Tikhonov methods*, submitted) and to the Range-Restricted methods.
Final remarks

- simple and efficient (all the extra computations are performed in reduced dimension);
- simultaneously determine the regularization parameter and the number of iterations;
- can be generalized to the multi-parameter case (S.G., P. Novati, *Multi-parameter Arnoldi-Tikhonov methods*, submitted) and to the Range-Restricted methods.
Final remarks

- simple and efficient (all the extra computations are performed in reduced dimension);
- simultaneously determine the regularization parameter and the number of iterations;
- can be generalized to the multi-parameter case (S.G., P. Novati, *Multi-parameter Arnoldi-Tikhonov methods*, submitted) and to the Range-Restricted methods.
Final remarks

- simple and efficient (all the extra computations are performed in reduced dimension);
- simultaneously determine the regularization parameter and the number of iterations;
- can be generalized to the multi-parameter case (S.G., P. Novati, *Multi-parameter Arnoldi-Tikhonov methods*, submitted) and to the Range-Restricted methods.