Large-Scale Tikhonov Regularization via Reduction by Orthogonal Projection

Heinrich Voss
voss@tuhh.de

Joint work with Jörg Lampe and Lothar Reichel

TU Hamburg–Harburg

Outline

Outline

Problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\| \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$, $m \geq n$, is severely ill-conditioned (not necessarily of full rank).

Problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\| \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$, $m \geq n$, is severely ill-conditioned (not necessarily of full rank).

$b \in \mathbb{R}^m$ represents observations and is assumed to be contaminated by an error $e \in \mathbb{R}^m$, i.e. $b = b_{\text{true}} + e$.

Problem

$$\min_{x\in\mathbb{R}^n} \|Ax - b\| \tag{1}$$

where $A \in \mathbb{R}^{m\times n}$, $m \geq n$, is severely ill-conditioned (not necessarily of full rank).

$b \in \mathbb{R}^m$ represents observations and is assumed to be contaminated by an error $e \in \mathbb{R}^m$, i.e. $b = b_{\text{true}} + e$.

The linear system of equations $Ax = b_{\text{true}}$ associated with the least-squares problem (1) is assumed to be consistent.

Problem

$$\min_{x \in \mathbb{R}^n} \|Ax - b\| \tag{1}$$

where $A \in \mathbb{R}^{m \times n}$, $m \geq n$, is severely ill-conditioned (not necessarily of full rank).

$b \in \mathbb{R}^m$ represents observations and is assumed to be contaminated by an error $e \in \mathbb{R}^m$, i.e. $b = b_{\text{true}} + e$.

The linear system of equations $Ax = b_{\text{true}}$ associated with the least-squares problem (1) is assumed to be consistent.

We would like to determine an approximation of its solution of minimal pseudo-norm by computing a suitable approximate solution of (1).

Problem

Due to the ill-conditioning of A and the error e in b, straightforward solution of (1) often does not yield a meaningful approximation of x_{true}.

Problem

Due to the ill-conditioning of A and the error e in b, straightforward solution of (1) often does not yield a meaningful approximation of x_{true}.

To stabilize the computations apply Tykhonov regularization, i.e. replace (1) by

$$\min_{x\in\mathbb{R}^n} \left\{ \|Ax-b\|^2 + \mu^{-1}\|Lx\|^2 \right\}. \tag{2}$$

Problem

Due to the ill-conditioning of A and the error e in b, straightforward solution of (1) often does not yield a meaningful approximation of x_{true}.

To stabilize the computations apply Tykhonov regularization, i.e. replace (1) by

$$\min_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|^2 + \mu^{-1} \|Lx\|^2 \right\}. \tag{2}$$

The scalar $\mu \in (0, \infty)$ is referred to as the regularization parameter and $L \in \mathbb{R}^{p \times n}$, $p \leq n$, as the regularization matrix.

Problem

Due to the ill-conditioning of A and the error e in b, straightforward solution of (1) often does not yield a meaningful approximation of x_{true}.

To stabilize the computations apply Tykhonov regularization, i.e. replace (1) by

$$\min_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|^2 + \mu^{-1} \|Lx\|^2 \right\}. \tag{2}$$

The scalar $\mu \in (0, \infty)$ is referred to as the regularization parameter and $L \in \mathbb{R}^{p \times n}$, $p \leq n$, as the regularization matrix.

The normal equations associated with (2) are given by

$$(A^T A + \mu^{-1} L^T L)x = A^T b. \tag{3}$$

Problem

Due to the ill-conditioning of A and the error e in b, straightforward solution of (1) often does not yield a meaningful approximation of x_{true}.

To stabilize the computations apply Tykhonov regularization, i.e. replace (1) by

$$\min_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|^2 + \mu^{-1} \|Lx\|^2 \right\}. \tag{2}$$

The scalar $\mu \in (0, \infty)$ is referred to as the regularization parameter and $L \in \mathbb{R}^{p \times n}$, $p \leq n$, as the regularization matrix.

The normal equations associated with (2) are given by

$$(A^T A + \mu^{-1} L^T L)x = A^T b. \tag{3}$$

They have the unique solution

$$x_\mu = (A^T A + \mu^{-1} L^T L)^{-1} A^T b \tag{4}$$

for any $\mu > 0$ when

$$\text{rank} \begin{bmatrix} A \\ L \end{bmatrix} = n. \tag{5}$$

We assume this to be the case.

Outline

Small dimension

When A and L are small, solutions x_μ of (2) can be determined easily for many values of $\mu > 0$ by first computing the generalized singular value decomposition (GSVD) of the matrix pair $\{A, L\}$ (Van Loan (1976), Varah (1979)).

Small dimension

When A and L are small, solutions x_μ of (2) can be determined easily for many values of $\mu > 0$ by first computing the generalized singular value decomposition (GSVD) of the matrix pair $\{A, L\}$ (Van Loan (1976), Varah (1979)).

For large-scale problems and a fixed $\mu > 0$, an approximation of x_μ can be determined by applying an iterative method, such as LSQR, to (3). However, generally, a suitable value of the parameter μ is not known a priori and has to be determined during the solution process.

Small dimension

When A and L are small, solutions x_μ of (2) can be determined easily for many values of $\mu > 0$ by first computing the generalized singular value decomposition (GSVD) of the matrix pair $\{A, L\}$ (Van Loan (1976), Varah (1979)).

For large-scale problems and a fixed $\mu > 0$, an approximation of x_μ can be determined by applying an iterative method, such as LSQR, to (3). However, generally, a suitable value of the parameter μ is not known a priori and has to be determined during the solution process.

Many approaches to determining an appropriate value of μ, including the L-curve criterion, the discrepancy principle, generalized cross validation, and information criteria, require the normal equations (3) to be solved repeatedly for many different values of the parameter μ. This can make application of LSQR costly.

Standard case

Approximations of the solution x_μ of problems in standard form ($L = I$) can be computed by partial Lanczos bidiagonalization of A (Björck (1988)).

Standard case

Approximations of the solution x_μ of problems in standard form ($L = I$) can be computed by partial Lanczos bidiagonalization of A (Björck (1988)).

The computed approximation, x_μ^ℓ, lives in the Krylov subspace

$$\mathcal{K}_\ell(A^TA, A^Tb) = \text{span}\{A^Tb, (A^TA)A^Tb, \dots, (A^TA)^{\ell-1}A^Tb\} \tag{6}$$

for some $\ell \geq 1$.

Standard case

Approximations of the solution x_μ of problems in standard form ($L = I$) can be computed by partial Lanczos bidiagonalization of A (Björck (1988)).

The computed approximation, x_μ^ℓ, lives in the Krylov subspace

$$\mathcal{K}_\ell(A^TA, A^Tb) = \text{span}\{A^Tb, (A^TA)A^Tb, \dots, (A^TA)^{\ell-1}A^Tb\} \qquad (6)$$

for some $\ell \geq 1$.

Due to the shift invariance of Krylov subspaces, i.e., the property that

$$\mathcal{K}_\ell(A^TA, A^Tb) = \mathcal{K}_\ell(A^TA + \mu^{-1}I, A^Tb), \qquad \mu > 0,$$

the spaces (6) can be used for all values of $\mu > 0$.

Standard case

Approximations of the solution x_μ of problems in standard form ($L = I$) can be computed by partial Lanczos bidiagonalization of A (Björck (1988)).

The computed approximation, x_μ^ℓ, lives in the Krylov subspace

$$\mathcal{K}_\ell(A^TA, A^Tb) = \text{span}\{A^Tb, (A^TA)A^Tb, \dots, (A^TA)^{\ell-1}A^Tb\} \tag{6}$$

for some $\ell \geq 1$.

Due to the shift invariance of Krylov subspaces, i.e., the property that

$$\mathcal{K}_\ell(A^TA, A^Tb) = \mathcal{K}_\ell(A^TA + \mu^{-1}I, A^Tb), \qquad \mu > 0,$$

the spaces (6) can be used for all values of $\mu > 0$.

This makes it possible to first project the problem (3) onto a Krylov subspace (6) and then regularize the projected problem by Tikhonov's method.

A-weighted pseudoinverse

Solution by partial Lanczos bidiagonalization also can be applied to Tikhonov regularization problems (2) with $L \neq I$, provided that the regularization matrix can be transformed to standard form without too much effort.

A-weighted pseudoinverse

Solution by partial Lanczos bidiagonalization also can be applied to Tikhonov regularization problems (2) with $L \neq I$, provided that the regularization matrix can be transformed to standard form without too much effort.

This transformation is carried out with the aid of the substitutions $y = Lx$ and $x = L_A^\dagger y$, where the matrix

$$L_A^\dagger := (I - (A(I - L^\dagger L))^\dagger A)L^\dagger$$

is referred to as the *A-weighted pseudoinverse of L* (Eldén (1982)).

A-weighted pseudoinverse

Solution by partial Lanczos bidiagonalization also can be applied to Tikhonov regularization problems (2) with $L \neq I$, provided that the regularization matrix can be transformed to standard form without too much effort.

This transformation is carried out with the aid of the substitutions $y = Lx$ and $x = L_A^\dagger y$, where the matrix

$$L_A^\dagger := (I - (A(I - L^\dagger L))^\dagger A)L^\dagger$$

is referred to as the *A-weighted pseudoinverse of L* (Eldén (1982)).

When L is banded with small bandwidth and has a known null space, transformation of (2) to standard form is attractive. Other situations when transformation to standard form is feasible are discussed in Reichel & Ye (2009).

A-weighted pseudoinverse

Solution by partial Lanczos bidiagonalization also can be applied to Tikhonov regularization problems (2) with $L \neq I$, provided that the regularization matrix can be transformed to standard form without too much effort.

This transformation is carried out with the aid of the substitutions $y = Lx$ and $x = L_A^\dagger y$, where the matrix

$$L_A^\dagger := (I - (A(I - L^\dagger L))^\dagger A) L^\dagger$$

is referred to as the *A-weighted pseudoinverse of L* (Eldén (1982)).

When L is banded with small bandwidth and has a known null space, transformation of (2) to standard form is attractive. Other situations when transformation to standard form is feasible are discussed in Reichel & Ye (2009).

However, when (2) is the discretization of an integral equation of the first kind in two or more space-dimensions, the regularization matrix L often is chosen to be a sum of Kronecker products. The expression for $L_A^\dagger$ then is complicated and unattractive to use.

Approximate GSVD

Kilmer, Hansen & Espanol (2007) recently proposed a projection method for large-scale Tikhonov-regularized least-squares problems that are infeasible or expensive to transform to standard form.

Approximate GSVD

Kilmer, Hansen & Espanol (2007) recently proposed a projection method for large-scale Tikhonov-regularized least-squares problems that are infeasible or expensive to transform to standard form.

The method computes an approximation of a partial GSVD of the matrix pair $\{A, L\}$ by an inner-outer iteration scheme.

Approximate GSVD

Kilmer, Hansen & Espanol (2007) recently proposed a projection method for large-scale Tikhonov-regularized least-squares problems that are infeasible or expensive to transform to standard form.

The method computes an approximation of a partial GSVD of the matrix pair $\{A, L\}$ by an inner-outer iteration scheme.

An attractive property of this method is that the computed approximate partial GSVD is independent of μ.

Approximate GSVD

Kilmer, Hansen & Espanol (2007) recently proposed a projection method for large-scale Tikhonov-regularized least-squares problems that are infeasible or expensive to transform to standard form.

The method computes an approximation of a partial GSVD of the matrix pair $\{A, L\}$ by an inner-outer iteration scheme.

An attractive property of this method is that the computed approximate partial GSVD is independent of μ.

However, the inner-outer iteration scheme may be expensive, due to the possibly fairly large number of required matrix-vector product evaluations with A and its transpose A^T.

Outline

Setting

Assume that an estimate of the norm of the error e in the vector b in (1) is available,

$$\varepsilon \approx \|e\|.$$

Setting

Assume that an estimate of the norm of the error e in the vector b in (1) is available,

$$\varepsilon \approx \|e\|.$$

The regularization parameter $\mu = \mu(\varepsilon)$ is determined by the discrepancy principle, i.e. so that the computed approximation $\widetilde{x}_\mu$ of the solution x_μ of (2) satisfies

$$\|A\widetilde{x}_\mu - b\| = \eta\varepsilon =: \delta, \tag{7}$$

where $\eta > 1$ is a user-specified constant, whose size depends on the accuracy of the estimate ε of $\|e\|$.

Setting

Assume that an estimate of the norm of the error e in the vector b in (1) is available,

$$\varepsilon \approx \|e\|.$$

The regularization parameter $\mu = \mu(\varepsilon)$ is determined by the discrepancy principle, i.e. so that the computed approximation $\widetilde{x}_\mu$ of the solution x_μ of (2) satisfies

$$\|A\widetilde{x}_\mu - b\| = \eta\varepsilon =: \delta, \tag{7}$$

where $\eta > 1$ is a user-specified constant, whose size depends on the accuracy of the estimate ε of $\|e\|$.

Let

$$\varphi(\mu) := \|Ax_\mu - b\|^2. \tag{8}$$

We are looking for $\overline{\mu}$ with $\varphi(\overline{\mu}) = \delta^2$.

Setting

Assume that an estimate of the norm of the error e in the vector b in (1) is available,

$$\varepsilon \approx \|e\|.$$

The regularization parameter $\mu = \mu(\varepsilon)$ is determined by the discrepancy principle, i.e. so that the computed approximation $\tilde{x}_\mu$ of the solution x_μ of (2) satisfies

$$\|A\tilde{x}_\mu - b\| = \eta\varepsilon =: \delta, \tag{7}$$

where $\eta > 1$ is a user-specified constant, whose size depends on the accuracy of the estimate ε of $\|e\|$.

Let

$$\varphi(\mu) := \|Ax_\mu - b\|^2. \tag{8}$$

We are looking for $\bar{\mu}$ with $\varphi(\bar{\mu}) = \delta^2$.

The function $\varphi(\mu)$ is convex and monotone such that finding $\bar{\mu}$ is an easy task, but note that for large A the evaluation of $\varphi(\mu)$ is expensive

Setting

Assume that an estimate of the norm of the error e in the vector b in (1) is available,

$$\varepsilon \approx \|e\|.$$

The regularization parameter $\mu = \mu(\varepsilon)$ is determined by the discrepancy principle, i.e. so that the computed approximation $\tilde{x}_\mu$ of the solution x_μ of (2) satisfies

$$\|A\tilde{x}_\mu - b\| = \eta\varepsilon =: \delta, \tag{7}$$

where $\eta > 1$ is a user-specified constant, whose size depends on the accuracy of the estimate ε of $\|e\|$.

Let

$$\varphi(\mu) := \|Ax_\mu - b\|^2. \tag{8}$$

We are looking for $\overline{\mu}$ with $\varphi(\overline{\mu}) = \delta^2$.

The function $\varphi(\mu)$ is convex and monotone such that finding $\overline{\mu}$ is an easy task, but note that for large A the evaluation of $\varphi(\mu)$ is expensive

A numerical method for inexpensively computing upper and lower bounds for $\bar{\mu}$ when $L = I$ is described in Calvetti & Reichel (2003).

Iterative projection method

Assume that L is a regularization matrix such that the computation with $L_A^\dagger$ is unattractive.

Iterative projection method

Assume that L is a regularization matrix such that the computation with $L_A^\dagger$ is unattractive.

Introduce a subspace $\mathcal{V} \subset \mathbb{R}^n$ of small dimension $k \ll n$, and approximate $\phi(\mu)$ by the function

$$\phi(\mu, \mathcal{V}) := \|Ax_\mu^k - b\|^2$$

where x_μ^k is obtained by solving the Tikhonov problem restricted to $\mathcal{V}$.

Iterative projection method

Assume that L is a regularization matrix such that the computation with $L_A^\dagger$ is unattractive.

Introduce a subspace $\mathcal{V} \subset \mathbb{R}^n$ of small dimension $k \ll n$, and approximate $\phi(\mu)$ by the function

$$\phi(\mu, \mathcal{V}) := \|Ax_\mu^k - b\|^2$$

where x_μ^k is obtained by solving the Tikhonov problem restricted to $\mathcal{V}$.

If a given accuracy is not met then expand the search space $\mathcal{V}$.

Iterative projection method

Assume that L is a regularization matrix such that the computation with $L_A^\dagger$ is unattractive.

Introduce a subspace $\mathcal{V} \subset \mathbb{R}^n$ of small dimension $k \ll n$, and approximate $\phi(\mu)$ by the function

$$\phi(\mu, \mathcal{V}) := \|Ax_\mu^k - b\|^2$$

where x_μ^k is obtained by solving the Tikhonov problem restricted to $\mathcal{V}$.

If a given accuracy is not met then expand the search space $\mathcal{V}$.

This yields the following type of iterative projection method.

Iterative projection method

Algorithm 1 Iterative Projection Tikhonov Method

Require: Initial basis V_0 of $\mathcal{V}$, $V_0^T V_0 = I$

1: **for** $i = 0, 1, \dots$ until convergence **do**
2: Find the root μ_i of $f(\mu; V_i) := \phi(\mu, \mathcal{V}) - \delta^2 = 0$
3: Solve $(V_i^T (A^T A + \mu_i^{-1} L^T L) V_i) y_{\mu_i} = V_i^T A^T b$
4: Compute $r_{\mu_i} := (A^T A + \mu_i^{-1} L^T L) V_i y_{\mu_i} - A^T b$
5: Reorthogonalize (optional) $\tilde{r}_{\mu_i} = (I - V_i V_i^T) r_{\mu_i}$
6: Normalize $v_{\text{new}} = \tilde{r}_{\mu_i} / \|\tilde{r}_{\mu_i}\|$
7: Enlarge search space $V_{i+1} = [V_i, v_{\text{new}}]$
8: **end for**
9: Determine approximate Tikhonov solution $x_{\mu_i} = V_i y_{\mu_i}$

Iterative projection method

Algorithm 2 Iterative Projection Tikhonov Method

Require: Initial basis V_0 of $\mathcal{V}$, $V_0^T V_0 = I$
for $i = 0, 1, \dots$ until convergence **do**
    Find the root μ_i of $f(\mu; V_i) := \phi(\mu, \mathcal{V}) - \delta^2 = 0$
    Solve $(V_i^T (A^T A + \mu_i^{-1} L^T L) V_i) y_{\mu_i} = V_i^T A^T b$
    Compute $r_{\mu_i} := (A^T A + \mu_i^{-1} L^T L) V_i y_{\mu_i} - A^T b$
    Reorthogonalize (optional) $\tilde{r}_{\mu_i} = (I - V_i V_i^T) r_{\mu_i}$
    Normalize $v_{\text{new}} = \tilde{r}_{\mu_i} / \|\tilde{r}_{\mu_i}\|$
    Enlarge search space $V_{i+1} = [V_i, v_{\text{new}}]$
end for
Determine approximate Tikhonov solution $x_{\mu_i} = V_i y_{\mu_i}$

In every iteration step have to

- determine μ_i
- expand search space with v_{new}

Solving projected problem

$$y_\mu = (V^T(A^TA + \mu^{-1}L^TL)V)^{-1}V^TA^Tb$$

solves the projected lest squares problem

$$\left\| \begin{bmatrix} AV \\ \mu^{-1/2}LV \end{bmatrix} y_\mu - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|^2 = \text{min!}$$

Solving projected problem

$$y_\mu = (V^T(A^TA + \mu^{-1}L^TL)V)^{-1}V^TA^Tb$$

solves the projected lest squares problem

$$\left\| \begin{bmatrix} AV \\ \mu^{-1/2}LV \end{bmatrix} y_\mu - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|^2 = \min!$$

i.e. with QR-factorizations $AV = Q_A R_A$ and $LV = Q_L R_L$

$$\left\| \begin{bmatrix} R_A \\ \mu^{-1/2}R_L \end{bmatrix} y_\mu - \begin{bmatrix} Q_A^T b \\ 0 \end{bmatrix} \right\|^2 = \min!$$

Solving projected problem

$$y_\mu = (V^T(A^TA + \mu^{-1}L^TL)V)^{-1}V^TA^Tb$$

solves the projected lest squares problem

$$\left\| \begin{bmatrix} AV \\ \mu^{-1/2}LV \end{bmatrix} y_\mu - \begin{bmatrix} b \\ 0 \end{bmatrix} \right\|^2 = \min!$$

i.e. with QR-factorizations $AV = Q_AR_A$ and $LV = Q_LR_L$

$$\left\| \begin{bmatrix} R_A \\ \mu^{-1/2}R_L \end{bmatrix} y_\mu - \begin{bmatrix} Q_A^Tb \\ 0 \end{bmatrix} \right\|^2 = \min!$$

Hence $\phi(\mu; V) = \|AVy_\mu - b\|_2^2$ can be evaluated by solving a least squares problem with a $2k \times k$ system matrix consisting of two stacked triangular matrices.

Solving projected problem

$$\phi'(\mu; V) = 2\left((R_A y_\mu)^T R_A y'_\mu - \mu^{-2}(R_L y_\mu)^T (R_L y_\mu)\right),$$

Solving projected problem

$$\phi'(\mu; V) = 2\left((R_A y_\mu)^T R_A y'_\mu - \mu^{-2}(R_L y_\mu)^T (R_L y_\mu)\right),$$

where y'_μ solves the least squares problem

$$\left\| \begin{bmatrix} R_A \\ \mu^{-1/2} R_L \end{bmatrix} y'_\mu - \begin{bmatrix} 0 \\ \mu^{-3/2} R_L y_\mu \end{bmatrix} \right\| = \text{min!}$$

Solving projected problem

$$\phi'(\mu; V) = 2\left((R_A y_\mu)^T R_A y'_\mu - \mu^{-2}(R_L y_\mu)^T (R_L y_\mu)\right),$$

where y'_μ solves the least squares problem

$$\left\| \begin{bmatrix} R_A \\ \mu^{-1/2} R_L \end{bmatrix} y'_\mu - \begin{bmatrix} 0 \\ \mu^{-3/2} R_L y_\mu \end{bmatrix} \right\| = \text{min!}$$

Notice that the system matrix is the same as in the last problem. Hence, reusing its QR factorization $\phi'(\mu; V)$ can be evaluated very cheaply.

Solving projected problem

$$\phi'(\mu; V) = 2\left((R_A y_\mu)^T R_A y'_\mu - \mu^{-2}(R_L y_\mu)^T (R_L y_\mu)\right),$$

where y'_μ solves the least squares problem

$$\left\| \begin{bmatrix} R_A \\ \mu^{-1/2} R_L \end{bmatrix} y'_\mu - \begin{bmatrix} 0 \\ \mu^{-3/2} R_L y_\mu \end{bmatrix} \right\| = \min!$$

Notice that the system matrix is the same as in the last problem. Hence, reusing its QR factorization $\phi'(\mu; V)$ can be evaluated very cheaply.

Due to the monotonicity and convexity of $f(\cdot; V)$ its root $\bar{\mu}$ can be determined safely with Newton's method with initial approximation $\mu_0 = 0$. However, this will be very time consuming.

Typical behavior of $f(\mu) = \phi(\mu; V) - \delta^2$

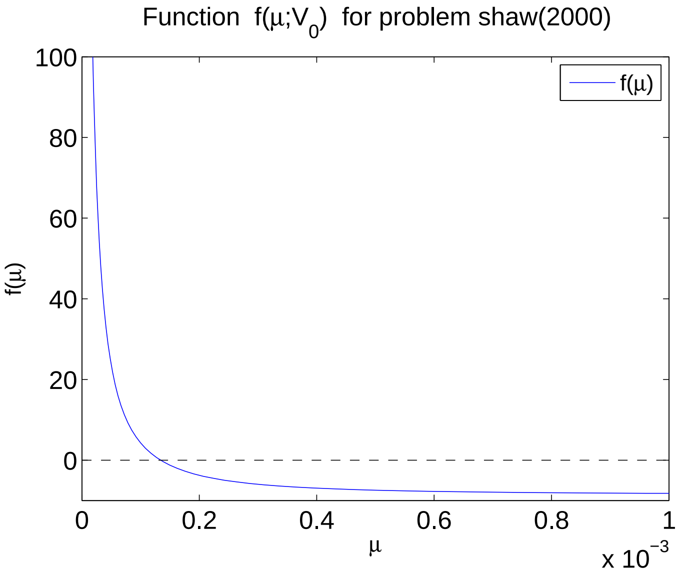


Figure: Plot of a typical function $f(\mu)$

Zero-finder based on rational inverse iteration

Consider a rational model for the inverse of f,

$$f^{-1} \approx h(f) := \frac{p(f)}{f - f_\infty}, \quad p(f) = \sum_{j=0}^{3} a_j f^j,$$

where

$$f_\infty = \lim_{\mu\to\infty} f(\mu; V) = \|b\|^2 - b^T Q_A (R_A R_A^\dagger) Q_A^T b - \delta^2.$$

Zero-finder based on rational inverse iteration

Consider a rational model for the inverse of f,

$$f^{-1} \approx h(f) := \frac{p(f)}{f - f_\infty}, \quad p(f) = \sum_{j=0}^{3} a_j f^j,$$

where

$$f_\infty = \lim_{\mu\to\infty} f(\mu; V) = \|b\|^2 - b^T Q_A (R_A R_A^\dagger) Q_A^T b - \delta^2.$$

For

$$0 < \mu_1 < \mu_2 \quad \text{such that} \quad f(\mu_1) > 0 > f(\mu_2)$$

we determine p such that $h(f(\mu_j)) = \mu_j$ and $h'(f(\mu_j)) = 1/f'(\mu_j)$ for $j = 1, 2$, and $\mu_{new} := h(0)$, and replace the value μ_1 or μ_2 which is on the same side of the root as μ_{new}

Zero-finder based on rational inverse iteration

Consider a rational model for the inverse of f,

$$f^{-1} \approx h(f) := \frac{p(f)}{f - f_\infty}, \quad p(f) = \sum_{j=0}^{3} a_j f^j,$$

where

$$f_\infty = \lim_{\mu\to\infty} f(\mu; V) = \|b\|^2 - b^T Q_A (R_A R_A^\dagger) Q_A^T b - \delta^2.$$

For

$$0 < \mu_1 < \mu_2 \quad \text{such that} \quad f(\mu_1) > 0 > f(\mu_2)$$

we determine p such that $h(f(\mu_j)) = \mu_j$ and $h'(f(\mu_j)) = 1/f'(\mu_j)$ for $j = 1, 2$, and $\mu_{new} := h(0)$, and replace the value μ_1 or μ_2 which is on the same side of the root as μ_{new}

Typically, one requires approximately 5 to 10 iteration steps to find an initial interval, and in the following iterations 2 (or even 1) iterations suffice.

Outline

1. Problem definition
2. Approaches in the literature
3. Iterative projection method
4. **Numerical Example**

Numerical example

Consider the restoration of a grey-scale image which is represented by an array of 256 $\times$ 256 pixels. The pixels are stored column-wise in a vector in $\mathbb{R}^{65536}$.

Numerical example

Consider the restoration of a grey-scale image which is represented by an array of 256×256 pixels. The pixels are stored column-wise in a vector in $\mathbb{R}^{65536}$.

The picture is blurred using the function `blur` with parameters *band* $= 5$ and *sigma* $= 1.0$ and Gaussian noise corresponding to the noise level $\sigma = 10^{-2}$ is added.

Numerical example

Consider the restoration of a grey-scale image which is represented by an array of 256×256 pixels. The pixels are stored column-wise in a vector in $\mathbb{R}^{65536}$.

The picture is blurred using the function `blur` with parameters $band = 5$ and $sigma = 1.0$ and Gaussian noise corresponding to the noise level $\sigma = 10^{-2}$ is added.

To restore the picture we apply the iterative projection method with $\eta = 1.05$ and regularization matrix

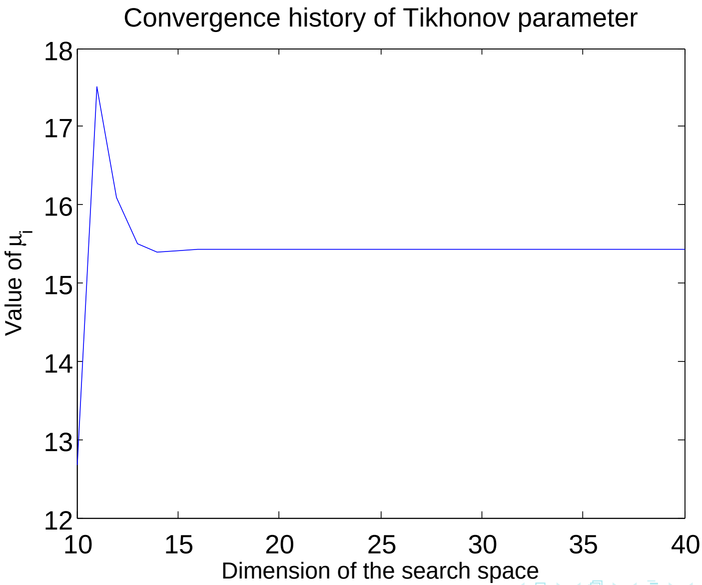
$$L = \begin{bmatrix} L_1 \otimes I_{256} \\ I_{256} \otimes L_1 \end{bmatrix}$$

where $L_1 \in \mathbb{R}^{255 \times 256}$ is the discrete first order derivative in one space dimension.

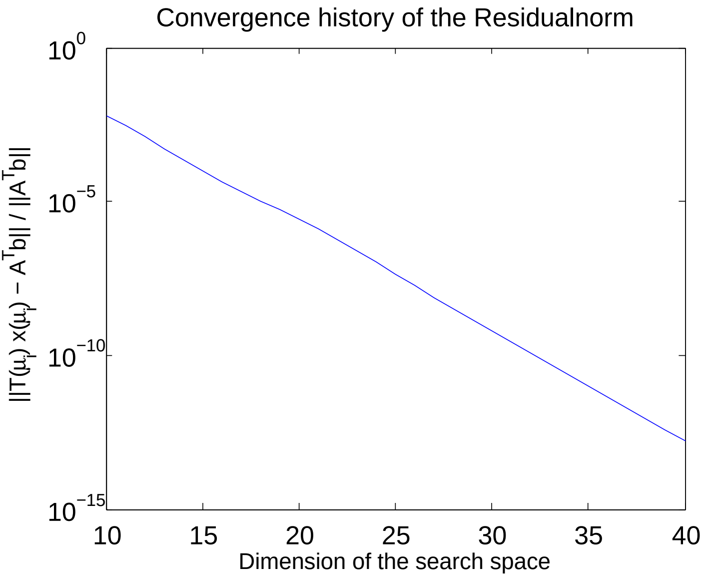
Original image

Blurred image

Convergence history of Tikhonov parameter

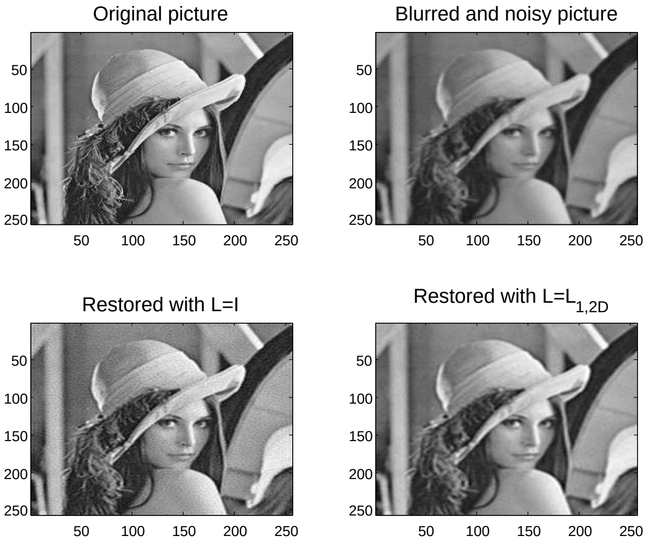


Convergence history of residual norm



restored picture

Four pictures



Conclusions

A new iterative projection method for Tikhonov regularization for large problems with general regularization matrices is presented.

Conclusions

A new iterative projection method for Tikhonov regularization for large problems with general regularization matrices is presented.

An efficient zero-finder based on rational inverse interpolation is suggested.

Conclusions

A new iterative projection method for Tikhonov regularization for large problems with general regularization matrices is presented.

An efficient zero-finder based on rational inverse interpolation is suggested.

Computed examples indicate that search spaces of fairly small dimension suffice.