Large–Scale Tikhonov Regularization for Total Least Squares Problems

Heinrich Voss

voss@tu-harburg.de

Joint work with Jörg Lampe

Hamburg University of Technology
Institute of Numerical Simulation

TUHH
Technische Universität Hamburg-Harburg

Outline

Outline

Total Least Squares Problems

The ordinary Least Squares (LS) method assumes that the system matrix A of a linear model is error free, and all errors are confined to the right hand side b.

However, in engineering applications this assumption is often unrealistic. Many problems in data estimation are obtained by linear systems where both, the matrix A and the right-hand side b, are contaminated by noise, for example if A as well is only available by measurements or if A is an idealized approximation of the true operator.

If the true values of the observed variables satisfy linear relations, and if the errors in the observations are independent random variables with zero mean and equal variance, then the total least squares (TLS) approach often gives better estimates than LS.

Given $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $m \ge n$

Find $\Delta A \in \mathbb{R}^{m\times n}$, $\Delta b \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ such that

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \quad \text{subject to } (A + \Delta A)x = b + \Delta b, \tag{1}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Total Least Squares Problems

The ordinary Least Squares (LS) method assumes that the system matrix A of a linear model is error free, and all errors are confined to the right hand side b.

However, in engineering applications this assumption is often unrealistic. Many problems in data estimation are obtained by linear systems where both, the matrix A and the right-hand side b, are contaminated by noise, for example if A as well is only available by measurements or if A is an idealized approximation of the true operator.

If the true values of the observed variables satisfy linear relations, and if the errors in the observations are independent random variables with zero mean and equal variance, then the total least squares (TLS) approach often gives better estimates than LS.

Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $m \geq n$

Find $\Delta A \in \mathbb{R}^{m \times n}$, $\Delta b \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ such that

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \quad \text{subject to } (A + \Delta A)x = b + \Delta b, \qquad (1)$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Total Least Squares Problems

The ordinary Least Squares (LS) method assumes that the system matrix A of a linear model is error free, and all errors are confined to the right hand side b.

However, in engineering applications this assumption is often unrealistic. Many problems in data estimation are obtained by linear systems where both, the matrix A and the right-hand side b, are contaminated by noise, for example if A as well is only available by measurements or if A is an idealized approximation of the true operator.

If the true values of the observed variables satisfy linear relations, and if the errors in the observations are independent random variables with zero mean and equal variance, then the total least squares (TLS) approach often gives better estimates than LS.

Given $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $m \geq n$

Find $\Delta A \in \mathbb{R}^{m\times n}$, $\Delta b \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ such that

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \quad \text{subject to } (A + \Delta A)x = b + \Delta b, \tag{1}$$

where $\|\cdot\|_F$ denotes the Frobenius norm.

Total Least Squares Problems

The ordinary Least Squares (LS) method assumes that the system matrix A of a linear model is error free, and all errors are confined to the right hand side b.

However, in engineering applications this assumption is often unrealistic. Many problems in data estimation are obtained by linear systems where both, the matrix A and the right-hand side b, are contaminated by noise, for example if A as well is only available by measurements or if A is an idealized approximation of the true operator.

If the true values of the observed variables satisfy linear relations, and if the errors in the observations are independent random variables with zero mean and equal variance, then the total least squares (TLS) approach often gives better estimates than LS.

Given $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $m \geq n$

Find $\Delta A \in \mathbb{R}^{m \times n}$, $\Delta b \in \mathbb{R}^m$ and $x \in \mathbb{R}^n$ such that

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \quad \text{subject to } (A + \Delta A)x = b + \Delta b, \tag{1}$$

where $\| \cdot \|_F$ denotes the Frobenius norm.

Total Least Squares Problems cnt.

Although the name “total least squares” was introduced only recently in the literature by Golub and Van Loan (1980), this fitting method is not new and has a long history in the statistical literature, where it is known as orthogonal regression, errors-in-variables, or measurement errors, and in image deblurring blind deconvolution

The univariate problem ($n = 1$) is already discussed by Adcock (1877), and it was rediscovered many times, often independently.

About 30 – 40 years ago, the technique was extended by Sprent (1969) and Gleser (1981) to the multivariate case ($n > 1$).

More recently, the total least squares method also stimulated interest outside statistics. In numerical linear algebra it was first studied by Golub and Van Loan (1980).

Total Least Squares Problems cnt.

Although the name “total least squares” was introduced only recently in the literature by Golub and Van Loan (1980), this fitting method is not new and has a long history in the statistical literature, where it is known as orthogonal regression, errors-in-variables, or measurement errors, and in image deblurring blind deconvolution

The univariate problem ($n = 1$) is already discussed by Adcock (1877), and it was rediscovered many times, often independently.

About 30 – 40 years ago, the technique was extended by Sprent (1969) and Gleser (1981) to the multivariate case ($n > 1$).

More recently, the total least squares method also stimulated interest outside statistics. In numerical linear algebra it was first studied by Golub and Van Loan (1980).

Total Least Squares Problems cnt.

Although the name “total least squares” was introduced only recently in the literature by Golub and Van Loan (1980), this fitting method is not new and has a long history in the statistical literature, where it is known as orthogonal regression, errors-in-variables, or measurement errors, and in image deblurring blind deconvolution

The univariate problem ($n = 1$) is already discussed by Adcock (1877), and it was rediscovered many times, often independently.

About 30 – 40 years ago, the technique was extended by Sprent (1969) and Gleser (1981) to the multivariate case ($n > 1$).

More recently, the total least squares method also stimulated interest outside statistics. In numerical linear algebra it was first studied by Golub and Van Loan (1980).

Total Least Squares Problems cnt.

Although the name “total least squares” was introduced only recently in the literature by Golub and Van Loan (1980), this fitting method is not new and has a long history in the statistical literature, where it is known as orthogonal regression, errors-in-variables, or measurement errors, and in image deblurring blind deconvolution

The univariate problem ($n = 1$) is already discussed by Adcock (1877), and it was rediscovered many times, often independently.

About 30 – 40 years ago, the technique was extended by Sprent (1969) and Gleser (1981) to the multivariate case ($n > 1$).

More recently, the total least squares method also stimulated interest outside statistics. In numerical linear algebra it was first studied by Golub and Van Loan (1980).

Total Least Squares Problems cnt.

The TLS problem can be analyzed in terms of the singular value decomposition of the augmented matrix $[A, b] = U\Sigma V^T$.

A TLS solution exists if and only if the right singular subspace $\mathcal{V}_{min}$ corresponding to σ_{n+1} contains at least one vector with a nonzero last component.

It is unique if it holds that $\sigma'_n > \sigma_{n+1}$ where σ'_n denotes the smallest singular value of A, and then it is given by

$$x_{TLS} = -\frac{1}{V(n+1, n+1)} V(1:n, n+1).$$

Total Least Squares Problems cnt.

The TLS problem can be analyzed in terms of the singular value decomposition of the augmented matrix $[A, b] = U\Sigma V^T$.

A TLS solution exists if and only if the right singular subspace $\mathcal{V}_{min}$ corresponding to σ_{n+1} contains at least one vector with a nonzero last component.

It is unique if it holds that $\sigma'_n > \sigma_{n+1}$ where σ'_n denotes the smallest singular value of A, and then it is given by

$$x_{TLS} = -\frac{1}{V(n+1, n+1)} V(1:n, n+1).$$

Total Least Squares Problems cnt.

The TLS problem can be analyzed in terms of the singular value decomposition of the augmented matrix $[A, b] = U\Sigma V^T$.

A TLS solution exists if and only if the right singular subspace $\mathcal{V}_{min}$ corresponding to σ_{n+1} contains at least one vector with a nonzero last component.

It is unique if it holds that $\sigma'_n > \sigma_{n+1}$ where σ'_n denotes the smallest singular value of A, and then it is given by

$$x_{TLS} = -\frac{1}{V(n+1, n+1)} V(1 : n, n+1).$$

Outline

Regularization of TLS Problems

When solving practical problems they are usually ill-conditioned, and regularization is necessary to stabilize the computed solution.

Fierro, Golub, Hansen and O'Leary (1997) suggested to filter its solution by truncating the small singular values of the TLS matrix $[A, b]$, and they proposed an iterative algorithm based on Lanczos bidiagonalization for computing truncated TLS solutions.

Regularization of TLS Problems

When solving practical problems they are usually ill-conditioned, and regularization is necessary to stabilize the computed solution.

Fierro, Golub, Hansen and O'Leary (1997) suggested to filter its solution by truncating the small singular values of the TLS matrix $[A, b]$, and they proposed an iterative algorithm based on Lanczos bidiagonalization for computing truncated TLS solutions.

Regularization Adding a Quadratic Constraint

Sima, van Huffel, and Golub (2004) suggest to regularize the TLS problem adding a quadratic constraint

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \quad \text{subject to } (A + \Delta A)x = b + \Delta b, \ \|Lx\| \le \delta,$$

where $\delta > 0$ and the regularization matrix $L \in \mathbb{R}^{p \times n}$, $p \le n$ defines a (semi-) norm on the solution through which the size of the solution is bounded or a certain degree of smoothness can be imposed.

Let $F \in \mathbb{R}^{n \times k}$ be a matrix whose columns form an orthonormal basis of the nullspace of the regularization matrix L. If it holds that

$$\sigma_{min}([AF, b]) < \sigma_{min}(AF)$$

then the solution x_{RTLS} of the constrained TLS problem is attained (Beck, Ben Tal 2006)

Regularization Adding a Quadratic Constraint

Sima, van Huffel, and Golub (2004) suggest to regularize the TLS problem adding a quadratic constraint

$$\|[\Delta A, \Delta b]\|_F^2 = \min! \quad \text{subject to } (A + \Delta A)x = b + \Delta b,\ \|Lx\| \le \delta,$$

where $\delta > 0$ and the regularization matrix $L \in \mathbb{R}^{p \times n}$, $p \le n$ defines a (semi-) norm on the solution through which the size of the solution is bounded or a certain degree of smoothness can be imposed.

Let $F \in \mathbb{R}^{n \times k}$ be a matrix whose columns form an orthonormal basis of the nullspace of the regularization matrix L. If it holds that

$$\sigma_{min}([AF, b]) < \sigma_{min}(AF)$$

then the solution x_{RTLS} of the constrained TLS problem is attained (Beck, Ben Tal 2006)

First Order Conditions; Golub, Hansen, O'Leary 1999

Assume x_{RTLS} exists and constraint is active, then (RTLS) is equivalent to

$$f(x) := \frac{\|Ax - b\|^2}{1 + \|x\|^2} = \min! \quad \text{subject to} \quad \|Lx\|^2 = \delta^2.$$

First-order optimality conditions are equivalent to

$$\begin{aligned}(A^TA + \lambda_I I + \lambda_L L^T L)x &= A^T b,\\ \mu \geq 0, \qquad \|Lx\|^2 &= \delta^2\end{aligned}$$

with

$$\lambda_I = -\frac{\|Ax - b\|^2}{1 + \|x\|^2}, \quad \lambda_L = \mu(1 + \|x\|^2), \quad \mu = \frac{b^T(b - Ax) + \lambda_I}{\delta^2(1 + \|x\|^2)}.$$

First Order Conditions; Golub, Hansen, O'Leary 1999

Assume x_{RTLS} exists and constraint is active, then (RTLS) is equivalent to

$$f(x) := \frac{\|Ax - b\|^2}{1 + \|x\|^2} = \min! \quad \text{subject to} \quad \|Lx\|^2 = \delta^2.$$

First-order optimality conditions are equivalent to

$$\begin{aligned} (A^T A + \lambda_I I + \lambda_L L^T L)x &= A^T b, \\ \mu \geq 0, \qquad \|Lx\|^2 &= \delta^2 \end{aligned}$$

with

$$\lambda_I = -\frac{\|Ax - b\|^2}{1 + \|x\|^2}, \quad \lambda_L = \mu(1 + \|x\|^2), \quad \mu = \frac{b^T(b - Ax) + \lambda_I}{\delta^2(1 + \|x\|^2)}.$$

Two Iterative Algorithms based on EVPs

Two approaches for solving the first order conditions

$$\left(A^TA + \lambda_I(x)I + \lambda_L(x)L^TL\right)x = A^Tb \qquad (*)$$

1. Quadratic EVPs: Sima, Van Huffel, Golub (2004), Lampe, V. (2007,2008)
   - Iterative algorithm based on updating λ_I
   - With fixed λ_I reformulate $(*)$ into QEP
   - Determine rightmost eigenvalue, i.e. the free parameter λ_L
   - Use corresponding eigenvector to update λ_I

2. Linear EVPs: Renaut, Guo (2005), Lampe, V. (2008)
   - Iterative algorithm based on updating λ_L
   - With fixed λ_L reformulate $(*)$ into linear EVP
   - Determine smallest eigenvalue, i.e. the free parameter λ_I
   - Use corresponding eigenvector to update λ_L

Two Iterative Algorithms based on EVPs

Two approaches for solving the first order conditions

$$\left(A^T A + \lambda_I(x) I + \lambda_L(x) L^T L\right) x = A^T b \qquad (*)$$

1. Quadratic EVPs: Sima, Van Huffel, Golub (2004), Lampe, V. (2007,2008)
   - Iterative algorithm based on updating λ_I
   - With fixed λ_I reformulate $(*)$ into QEP
   - Determine rightmost eigenvalue, i.e. the free parameter λ_L
   - Use corresponding eigenvector to update λ_I
2. Linear EVPs: Renaut, Guo (2005), Lampe, V. (2008)
   - Iterative algorithm based on updating λ_L
   - With fixed λ_L reformulate $(*)$ into linear EVP
   - Determine smallest eigenvalue, i.e. the free parameter λ_I
   - Use corresponding eigenvector to update λ_L

Two Iterative Algorithms based on EVPs

Two approaches for solving the first order conditions

$$\left(A^T A + \lambda_I(x) I + \lambda_L(x) L^T L\right) x = A^T b \qquad (*)$$

1. Quadratic EVPs: Sima, Van Huffel, Golub (2004), Lampe, V. (2007,2008)
   - Iterative algorithm based on updating λ_I
   - With fixed λ_I reformulate $(*)$ into QEP
   - Determine rightmost eigenvalue, i.e. the free parameter λ_L
   - Use corresponding eigenvector to update λ_I
2. Linear EVPs: Renaut, Guo (2005), Lampe, V. (2008)
   - Iterative algorithm based on updating λ_L
   - With fixed λ_L reformulate $(*)$ into linear EVP
   - Determine smallest eigenvalue, i.e. the free parameter λ_I
   - Use corresponding eigenvector to update λ_L

Outline

Tikhonov Regularization of TLS problem

$$f(x) + \lambda \|Lx\|^2 = \frac{\|Ax - b\|^2}{1 + \|x\|^2} + \lambda \|Lx\|^2 = \text{min!}.$$

Beck, Ben–Tal (2006) proposed an algorithm where in each iteration step a Cholesky decomposition has to be computed, which is prohibitive for large-scale problems.

We present a method which solves the first order conditions which are equivalent to

$$q(x) := (A^T A + \mu L^T L - f(x) I)x - A^T b = 0, \quad \text{with } \mu := (1 + \|x\|^2)\lambda.$$

via a combination of Newton's method with an iterative projection method.

Tikhonov Regularization of TLS problem

$$f(x) + \lambda\|Lx\|^2 = \frac{\|Ax - b\|^2}{1 + \|x\|^2} + \lambda\|Lx\|^2 = \min!.$$

Beck, Ben–Tal (2006) proposed an algorithm where in each iteration step a Cholesky decomposition has to be computed, which is prohibitive for large-scale problems.

We present a method which solves the first order conditions which are equivalent to

$$q(x) := (A^TA + \mu L^TL - f(x)I)x - A^Tb = 0, \quad \text{with } \mu := (1 + \|x\|^2)\lambda.$$

via a combination of Newton's method with an iterative projection method.

Tikhonov Regularization of TLS problem

$$f(x) + \lambda\|Lx\|^2 = \frac{\|Ax - b\|^2}{1 + \|x\|^2} + \lambda\|Lx\|^2 = \text{min!}.$$

Beck, Ben–Tal (2006) proposed an algorithm where in each iteration step a Cholesky decomposition has to be computed, which is prohibitive for large-scale problems.

We present a method which solves the first order conditions which are equivalent to

$$q(x) := (A^TA + \mu L^TL - f(x)I)x - A^Tb = 0, \quad \text{with } \mu := (1 + \|x\|^2)\lambda.$$

via a combination of Newton's method with an iterative projection method.

Tikhonov Regularization of TLS problem

Newton's method:

$$x^{k+1} = x^k - J(x^k)^{-1} q(x^k)$$

with the Jacobi matrix

$$J(x) = A^T A + \mu L^T L - f(x) I - 2x \frac{x^T A^T A - b^T A - f(x) x^T}{1 + \|x\|^2}.$$

Sherman–Morrison formula yields

$$x^{k+1} = \hat{J}_k^{-1} A^T b - \frac{1}{1 - (v^k)^T \hat{J}_k^{-1} u^k} \hat{J}_k^{-1} u^k (v^k)^T (x^k - \hat{J}_k^{-1} A^T b),$$

with

$$\hat{J}(x) := A^T A + \mu L^T L - f(x) I,$$

$$u^k := 2x^k / (1 + \|x^k\|^2) \quad \text{and} \quad v^k := A^T A x^k - A^T b - f(x^k) x^k.$$

Tikhonov Regularization of TLS problem

Newton's method:

$$x^{k+1} = x^k - J(x^k)^{-1} q(x^k)$$

with the Jacobi matrix

$$J(x) = A^T A + \mu L^T L - f(x) I - 2x \frac{x^T A^T A - b^T A - f(x) x^T}{1 + \|x\|^2}.$$

Sherman–Morrison formula yields

$$x^{k+1} = \hat{J}_k^{-1} A^T b - \frac{1}{1 - (v^k)^T \hat{J}_k^{-1} u^k} \hat{J}_k^{-1} u^k (v^k)^T (x^k - \hat{J}_k^{-1} A^T b),$$

with

$$\hat{J}(x) := A^T A + \mu L^T L - f(x) I,$$

$$u^k := 2x^k / (1 + \|x^k\|^2) \quad \text{and} \quad v^k := A^T A x^k - A^T b - f(x^k) x^k.$$

Tikhonov Regularization of TLS problem

To avoid the solution of the large scale linear systems with varying matrices $\hat{J}_k$ we combine Newton's method with an iterative projection method.

Let $\mathcal{V}$ be an ansatz space of small dimension k, and let the columns of $V \in \mathbb{R}^{k \times n}$ form an orthonormal basis of $\mathcal{V}$.
Replace $z = \hat{J}_k^{-1} A^T b$ with Vy_1^k where y_1^k solves $V^T \hat{J}_k V y_1^k = A^T b$,
and $w = \hat{J}_k^{-1} u^k$ with Vy_2^k where y_2^k solves $V^T \hat{J}_k V y_2^k = u^k$.

If

$$x_{k+1} = V_k y_1^k - \frac{1}{1 - (v^k)^T V_k y_2^k} V_k y_2^k (v^k)^T (x^k - V_k y_1^k)$$

does not satisfy a prescribed accuracy requirement, then $\mathcal{V}$ is expanded with the residual

$$q(x^{k+1}) = (A^T A + \mu L^T L - f(x^k) I) x^{k+1} - A^T b$$

and the step is repeated until convergence.

Tikhonov Regularization of TLS problem

To avoid the solution of the large scale linear systems with varying matrices $\hat{J}_k$ we combine Newton's method with an iterative projection method.

Let $\mathcal{V}$ be an ansatz space of small dimension k, and let the columns of $V \in \mathbb{R}^{k \times n}$ form an orthonormal basis of $\mathcal{V}$.
Replace $z = \hat{J}_k^{-1} A^T b$ with $V y_1^k$ where y_1^k solves $V^T \hat{J}_k V y_1^k = A^T b$,
and $w = \hat{J}_k^{-1} u^k$ with $V y_2^k$ where y_2^k solves $V^T \hat{J}_k V y_2^k = u^k$.

If

$$x_{k+1} = V_k y_1^k - \frac{1}{1 - (v^k)^T V_k y_2^k} V_k y_2^k (v^k)^T (x^k - V_k y_1^k)$$

does not satisfy a prescribed accuracy requirement, then $\mathcal{V}$ is expanded with the residual

$$q(x^{k+1}) = (A^T A + \mu L^T L - f(x^k) I) x^{k+1} - A^T b$$

and the step is repeated until convergence.

Tikhonov Regularization of TLS problem

To avoid the solution of the large scale linear systems with varying matrices $\hat{J}_k$ we combine Newton's method with an iterative projection method.

Let $\mathcal{V}$ be an ansatz space of small dimension k, and let the columns of $V \in \mathbb{R}^{k\times n}$ form an orthonormal basis of $\mathcal{V}$.
Replace $z = \hat{J}_k^{-1} A^T b$ with Vy_1^k where y_1^k solves $V^T \hat{J}_k V y_1^k = A^T b$,
and $w = \hat{J}_k^{-1} u^k$ with Vy_2^k where y_2^k solves $V^T \hat{J}_k V y_2^k = u^k$.

If

$$x_{k+1} = V_k y_1^k - \frac{1}{1-(v^k)^T V_k y_2^k} V_k y_2^k (v^k)^T (x^k - V_k y_1^k)$$

does not satisfy a prescribed accuracy requirement, then $\mathcal{V}$ is expanded with the residual

$$q(x^{k+1}) = (A^T A + \mu L^T L - f(x^k) I) x^{k+1} - A^T b$$

and the step is repeated until convergence.

Tikhonov Regularization of TLS problem

Initializing the iterative projection method with a Krylov space $\mathcal{V} = \mathcal{K}_\ell(A^T A + \mu L^T L, A^T b)$ the iterates x^k are contained in a Krylov space of $A^T A + \mu L^T L$.

Due to the convergence properties of the Lanczos process the main contributions come from the first singular vectors of $[A; \sqrt{\mu}L]$ which for small μ are close to the first right singular vectors of A.

It is common knowledge that these vectors are not always appropriate basis vectors for a regularized solution, and it may be advantageous to apply the regularization with a general regularization matrix L implicitly.

Tikhonov Regularization of TLS problem

Initializing the iterative projection method with a Krylov space $\mathcal{V} = \mathcal{K}_\ell(A^TA + \mu L^TL, A^Tb)$ the iterates x^k are contained in a Krylov space of $A^TA + \mu L^TL$.

Due to the convergence properties of the Lanczos process the main contributions come from the first singular vectors of $[A; \sqrt{\mu}L]$ which for small μ are close to the first right singular vectors of A.

It is common knowledge that these vectors are not always appropriate basis vectors for a regularized solution, and it may be advantageous to apply the regularization with a general regularization matrix L implicitly.

Tikhonov Regularization of TLS problem

Initializing the iterative projection method with a Krylov space $\mathcal{V} = \mathcal{K}_\ell(A^TA + \mu L^TL, A^Tb)$ the iterates x^k are contained in a Krylov space of $A^TA + \mu L^TL$.

Due to the convergence properties of the Lanczos process the main contributions come from the first singular vectors of $[A; \sqrt{\mu}L]$ which for small μ are close to the first right singular vectors of A.

It is common knowledge that these vectors are not always appropriate basis vectors for a regularized solution, and it may be advantageous to apply the regularization with a general regularization matrix L implicitly.

Tikhonov Regularization of TLS problem

Assume that L is nonsingular and use the transformation $x := L^{-1}y$ (for general L we had to use the A-weighted generalized inverse $L_A^\dagger$, cf. Elden 1982) which yields

$$\frac{\|AL^{-1}y - b\|^2}{1 + \|L^{-1}y\|^2} + \lambda\|y\|^2 = \text{min!}.$$

Transforming the first order conditions back and multiplying from the left with L^{-1} one gets

$$(L^T L)^{-1}(A^T Ax + \mu L^T Lx - f(x)x - A^T b) = 0.$$

This equation suggests to precondition the expansion of the search space with $L^T L$ or an approximation $M \approx L^T L$ thereof which yields the following Algorithm.

Tikhonov Regularization of TLS problem

Assume that L is nonsingular and use the transformation $x := L^{-1}y$ (for general L we had to use the A-weighted generalized inverse $L_A^\dagger$, cf. Elden 1982) which yields

$$\frac{\|AL^{-1}y - b\|^2}{1 + \|L^{-1}y\|^2} + \lambda\|y\|^2 = \text{min!}.$$

Transforming the first order conditions back and multiplying from the left with L^{-1} one gets

$$(L^TL)^{-1}(A^TAx + \mu L^TLx - f(x)x - A^Tb) = 0.$$

This equation suggests to precondition the expansion of the search space with L^TL or an approximation $M \approx L^TL$ thereof which yields the following Algorithm.

Tikhonov Regularization of TLS problem

Assume that L is nonsingular and use the transformation $x := L^{-1}y$ (for general L we had to use the A-weighted generalized inverse $L_A^\dagger$, cf. Elden 1982) which yields

$$\frac{\|AL^{-1}y - b\|^2}{1 + \|L^{-1}y\|^2} + \lambda\|y\|^2 = \min!.$$

Transforming the first order conditions back and multiplying from the left with L^{-1} one gets

$$(L^TL)^{-1}(A^TAx + \mu L^TLx - f(x)x - A^Tb) = 0.$$

This equation suggests to precondition the expansion of the search space with L^TL or an approximation $M \approx L^TL$ thereof which yields the following Algorithm.

Tikhonov Regularization of TLS problem

Require: Initial basis V_0 with $V_0^T V_0 = I$, starting vector x^0

for $k = 0, 1, \dots$ until convergence **do**
    Compute $f(x^k) = \|Ax^k - b\|^2/(1 + \|x^k\|^2)$
    Solve $V_k^T \hat{J}_k V_k y_1^k = V_k^T A^T b$ for y_1^k
    Compute $u^k = 2x^k/(1 + \|x^k\|^2)$ and $v^k = A^T A x^k - A^T b - f(x^k) x^k$
    Solve $V_k^T \hat{J}_k V_k y_2^k = V_k^T u^k$ for y_2^k
    Compute $x^{k+1} = V_k y_1^k - \frac{1}{1 - (v^k)^T V_k y_2^k} V_k y_2^k (v^k)^T (x^k - V_k y_1^k)$
    Compute $q^{k+1} = (A^T A + \mu L^T L - f(x^k) I) x^{k+1} - A^T b$
    Compute $\tilde{r} = M^{-1} q^{k+1}$
    Orthogonalize $\hat{r} = (I - V_k V_k^T) \tilde{r}$
    Normalize $v_{\text{new}} = \hat{r}/\|\hat{r}\|$
    Enlarge search space $V_{k+1} = [V_k, v_{\text{new}}]$
end for
Output: Approximate Tikhonov TLS solution x^{k+1}

Tikhonov Regularization of TLS problem

The Tikhonov TLS methods allows for a massive reuse of information from previous iteration steps.

Assume that the matrices V_k, $A^T AV_k$, $L^T LV_k$ are stored. Then neglecting multiplications with L and L^T and solves with M the essential cost in every iteration step is only two matrix-vector products with dense matrices A and A^T for extending $A^T AV_k$.

With these matrices $f(x^k)$ in Line 1 can be evaluated as

$$f(x^k) = \frac{1}{1 + \|y^k\|^2} \left((x^k)^T (A^T Ay^k) - 2(y^k)^T V_k^T (A^T b) + \|b\|^2 \right),$$

and q^{k+1} in Line 7 can be determined according to

$$q^{k+1} = (A^T AV_k)y^{k+1} + \mu(L^T LV_k)y^{k+1} - f(x^k)x^{k+1} - A^T b.$$

Since the the number of iteration steps until convergence is usually very small compared to the dimension n, the overall cost of the Algorithm is of the order $\mathcal{O}(mn)$.

Tikhonov Regularization of TLS problem

The Tikhonov TLS methods allows for a massive reuse of information from previous iteration steps.

Assume that the matrices V_k, $A^T AV_k$, $L^T LV_k$ are stored. Then neglecting multiplications with L and L^T and solves with M the essential cost in every iteration step is only two matrix-vector products with dense matrices A and A^T for extending $A^T AV_k$.

With these matrices $f(x^k)$ in Line 1 can be evaluated as

$$f(x^k) = \frac{1}{1 + \|y^k\|^2} \left((x^k)^T (A^T Ay^k) - 2(y^k)^T V_k^T (A^T b) + \|b\|^2 \right),$$

and q^{k+1} in Line 7 can be determined according to

$$q^{k+1} = (A^T AV_k)y^{k+1} + \mu(L^T LV_k)y^{k+1} - f(x^k)x^{k+1} - A^T b.$$

Since the the number of iteration steps until convergence is usually very small compared to the dimension n, the overall cost of the Algorithm is of the order $\mathcal{O}(mn)$.

Tikhonov Regularization of TLS problem

The Tikhonov TLS methods allows for a massive reuse of information from previous iteration steps.

Assume that the matrices V_k, A^TAV_k, L^TLV_k are stored. Then neglecting multiplications with L and L^T and solves with M the essential cost in every iteration step is only two matrix-vector products with dense matrices A and A^T for extending A^TAV_k.

With these matrices $f(x^k)$ in Line 1 can be evaluated as

$$f(x^k) = \frac{1}{1 + \|y^k\|^2} \left((x^k)^T (A^T A y^k) - 2(y^k)^T V_k^T (A^T b) + \|b\|^2 \right),$$

and q^{k+1} in Line 7 can be determined according to

$$q^{k+1} = (A^TAV_k)y^{k+1} + \mu(L^TLV_k)y^{k+1} - f(x^k)x^{k+1} - A^Tb.$$

Since the the number of iteration steps until convergence is usually very small compared to the dimension n, the overall cost of the Algorithm is of the order $\mathcal{O}(mn)$.

Tikhonov Regularization of TLS problem

The Tikhonov TLS methods allows for a massive reuse of information from previous iteration steps.

Assume that the matrices V_k, $A^T AV_k$, $L^T LV_k$ are stored. Then neglecting multiplications with L and L^T and solves with M the essential cost in every iteration step is only two matrix-vector products with dense matrices A and A^T for extending $A^T AV_k$.

With these matrices $f(x^k)$ in Line 1 can be evaluated as

$$f(x^k) = \frac{1}{1 + \|y^k\|^2} \left((x^k)^T (A^T Ay^k) - 2(y^k)^T V_k^T (A^T b) + \|b\|^2 \right),$$

and q^{k+1} in Line 7 can be determined according to

$$q^{k+1} = (A^T AV_k)y^{k+1} + \mu(L^T LV_k)y^{k+1} - f(x^k)x^{k+1} - A^T b.$$

Since the the number of iteration steps until convergence is usually very small compared to the dimension n, the overall cost of the Algorithm is of the order $\mathcal{O}(mn)$.

Outline

Numerical Experiments

Consider several examples from Hansen's Regularization Tools.

The regularization matrix L is chosen to be the nonsingular approximation to the scaled discrete first order derivative operator in one space-dimension.

The numerical tests are carried out on an Intel Core 2 Duo T7200 computer with 2.3 GHz and 2 GB RAM under MATLAB R2009a (actually our numerical examples require less than 0.5 GB RAM).

Numerical Experiments

Consider several examples from Hansen's Regularization Tools.

The regularization matrix L is chosen to be the nonsingular approximation to the scaled discrete first order derivative operator in one space-dimension.

The numerical tests are carried out on an Intel Core 2 Duo T7200 computer with 2.3 GHz and 2 GB RAM under MATLAB R2009a (actually our numerical examples require less than 0.5 GB RAM).

Numerical Experiments

Consider several examples from Hansen's Regularization Tools.

The regularization matrix L is chosen to be the nonsingular approximation to the scaled discrete first order derivative operator in one space-dimension.

The numerical tests are carried out on an Intel Core 2 Duo T7200 computer with 2.3 GHz and 2 GB RAM under MATLAB R2009a (actually our numerical examples require less than 0.5 GB RAM).

Numerical Experiments

Problem	Method	$\frac{\|q(x^k)\|}{\|A^T b\|}$	Iters	MatVecs	$\frac{\|x - x_{true}\|}{\|x_{true}\|}$
phillips	TTLS	8.5e-16	8.0	25.0	8.9e-2
$\sigma = 1e-3$	RTLSQEP	5.7e-11	3.0	42.0	8.9e-2
	RTLSEVP	7.1e-13	4.0	47.6	8.9e-2
baart	TTLS	2.3e-15	10.1	29.2	1.5e-1
$\sigma = 1e-3$	RTLSQEP	1.0e-07	15.7	182.1	1.4e-1
	RTLSEVP	4.1e-10	7.8	45.6	1.5e-1
shaw	TTLS	9.6e-16	8.3	25.6	7.0e-2
$\sigma = 1e-3$	RTLSQEP	3.7e-09	4.1	76.1	7.0e-2
	RTLSEVP	2.6e-10	3.0	39.0	7.0e-2
deriv2	TTLS	1.2e-15	10.0	29.0	4.9e-2
$\sigma = 1e-3$	RTLSQEP	2.3e-09	3.1	52.3	4.9e-2
	RTLSEVP	2.6e-12	5.0	67.0	4.9e-2
heat(κ=1)	TTLS	8.4e-16	19.9	48.8	1.5e-1
$\sigma = 1e-2$	RTLSQEP	4.1e-08	3.8	89.6	1.5e-1
	RTLSEVP	3.2e-11	4.1	67.2	1.5e-1
heat(κ=5)	TTLS	1.4e-13	25.0	59.0	1.1e-1
$\sigma = 1e-3$	RTLSQEP	6.1e-07	4.6	105.2	1.1e-1
	RTLSEVP	9.8e-11	4.0	65.0	1.1e-1

Outline

Conclusions

We discussed a Tikhonov regularization approach for large total least squares problems.

It is highly advantageous to combine Newton's method with an iterative projection method and to reuse information gathered in previous iteration steps.

Several examples demonstrate that fairly small ansatz spaces are are sufficient to get accurate solutions. Hence, the method is qualified to solve large-scale regularized total least squares problems efficiently

We assumed the regularization parameter λ to be fixed. The same technique of recycling ansatz spaces can be used in an L-curve method to determine a reasonable parameter.

Conclusions

We discussed a Tikhonov regularization approach for large total least squares problems.

It is highly advantageous to combine Newton's method with an iterative projection method and to reuse information gathered in previous iteration steps.

Several examples demonstrate that fairly small ansatz spaces are are sufficient to get accurate solutions. Hence, the method is qualified to solve large-scale regularized total least squares problems efficiently

We assumed the regularization parameter λ to be fixed. The same technique of recycling ansatz spaces can be used in an L-curve method to determine a reasonable parameter.

Conclusions

We discussed a Tikhonov regularization approach for large total least squares problems.

It is highly advantageous to combine Newton's method with an iterative projection method and to reuse information gathered in previous iteration steps.

Several examples demonstrate that fairly small ansatz spaces are are sufficient to get accurate solutions. Hence, the method is qualified to solve large-scale regularized total least squares problems efficiently

We assumed the regularization parameter λ to be fixed. The same technique of recycling ansatz spaces can be used in an L-curve method to determine a reasonable parameter.

Conclusions

We discussed a Tikhonov regularization approach for large total least squares problems.

It is highly advantageous to combine Newton's method with an iterative projection method and to reuse information gathered in previous iteration steps.

Several examples demonstrate that fairly small ansatz spaces are are sufficient to get accurate solutions. Hence, the method is qualified to solve large-scale regularized total least squares problems efficiently

We assumed the regularization parameter λ to be fixed. The same technique of recycling ansatz spaces can be used in an L-curve method to determine a reasonable parameter.