Efficient computation of the L-curve in Large Scale Regularized Total Least Squares

Heinrich Voss
voss@tuhh.de

Hamburg University of Technology
Institute of Numerical Simulation

Joint work with Jörg Lampe

TUHH
Technische Universität Hamburg-Harburg

Outline

1. Total Least Squares
2. Solving RTLS problems
3. Determining the hyperparamter
4. Numerical Examples
5. Conclusions

Least Squares Problem

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.

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

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

$$\|\Delta b\| = \min! \quad \text{subject to } Ax = b + \Delta b.$$

Obviously equivalent to: Find $x \in \mathbb{R}^n$ such that

$$\|Ax - b\| = \min!$$

Least Squares Problem

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.

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

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

$$\|\Delta b\| = \min! \quad \text{subject to } Ax = b + \Delta b.$$

Obviously equivalent to: Find $x \in \mathbb{R}^n$ such that

$$\|Ax - b\| = \min!$$

Least Squares Problem

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.

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

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

$$\|\Delta b\| = \text{min!} \quad \text{subject to } Ax = b + \Delta b.$$

Obviously equivalent to: Find $x \in \mathbb{R}^n$ such that

$$\|Ax - b\| = \text{min!}$$

Total Least Squares

In practical applications it may happen that all data are contaminated by noise, for instance because the matrix A is also obtained from measurements.

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.

In statistics this approach is called errors-in-variables problem or orthogonal regression, in image deblurring blind deconvolution

Total Least Squares

In practical applications it may happen that all data are contaminated by noise, for instance because the matrix A is also obtained from measurements.

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.

In statistics this approach is called errors-in-variables problem or orthogonal regression, in image deblurring blind deconvolution

Total Least Squares

In practical applications it may happen that all data are contaminated by noise, for instance because the matrix A is also obtained from measurements.

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.

In statistics this approach is called errors-in-variables problem or orthogonal regression, in image deblurring blind deconvolution

Total Least Squares

In practical applications it may happen that all data are contaminated by noise, for instance because the matrix A is also obtained from measurements.

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.

In statistics this approach is called errors-in-variables problem or orthogonal regression, in image deblurring blind deconvolution

Regularized Total Least Squares

If A and $[A, b]$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{k \times n}$, $k \leq n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problem reads:

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

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

Lemma (Beck, Ben-Tal 2006)

Let K be an orthonormal basis of $\ker(L)$. If $\sigma_{min}([AK, b]) < \sigma_{min}(AK)$ holds, a solution of the RTLS problem exists.

In general there does not exist a closed form solution x_{RTLS} of the RTLS problem.

Regularized Total Least Squares

If A and $[A, b]$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{k \times n}$, $k \leq n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problem reads:

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

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

Lemma (Beck, Ben-Tal 2006)

Let K be an orthonormal basis of $\ker(L)$. If $\sigma_{min}([AK, b]) < \sigma_{min}(AK)$ holds, a solution of the RTLS problem exists.

In general there does not exist a closed form solution x_{RTLS} of the RTLS problem.

Regularized Total Least Squares

If A and $[A, b]$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{k\times n}$, $k \le n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problem reads:

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

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

Lemma (Beck, Ben-Tal 2006)

Let K be an orthonormal basis of $\ker(L)$. If $\sigma_{min}([AK, b]) < \sigma_{min}(AK)$ holds, a solution of the RTLS problem exists.

In general there does not exist a closed form solution x_{RTLS} of the RTLS problem.

Regularized Total Least Squares

If A and $[A, b]$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{k \times n}$, $k \leq n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problem reads:

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

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

Lemma (Beck, Ben-Tal 2006)

Let K be an orthonormal basis of $\ker(L)$. If $\sigma_{min}([AK, b]) < \sigma_{min}(AK)$ holds, a solution of the RTLS problem exists.

In general there does not exist a closed form solution x_{RTLS} of the RTLS problem.

Regularized Total Least Squares

If A and $[A, b]$ are ill-conditioned, regularization is necessary.

Let $L \in \mathbb{R}^{k \times n}$, $k \leq n$ and $\delta > 0$. Then the quadratically constrained formulation of the Regularized Total Least Squares (RTLS) problem reads:

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

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

Lemma (Beck, Ben-Tal 2006)

Let K be an orthonormal basis of $\ker(L)$. If $\sigma_{min}([AK, b]) < \sigma_{min}(AK)$ holds, a solution of the RTLS problem exists.

In general there does not exist a closed form solution x_{RTLS} of the RTLS problem.

First Order Conditions

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

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

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

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

Two approaches for solving the first order conditions

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

Idea of Sima, Van Huffel, Golub ['04]:

- 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

Idea of Renaut, Guo ['05] is the other way round:

- 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

Two approaches for solving the first order conditions

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

Idea of Sima, Van Huffel, Golub ['04]:

- 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

Idea of Renaut, Guo ['05] is the other way round:

- 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

Two approaches for solving the first order conditions

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

Idea of Sima, Van Huffel, Golub ['04]:

- 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

Idea of Renaut, Guo ['05] is the other way round:

- 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

Solving the Sequence of Eigenproblems

RTLSQEP algorithm contains sequence of quadratic eigenproblems

$$T_k(\lambda)u = (W_k + \lambda I)^2 u - \delta^{-2} h h^T u = 0, \quad \text{for} \quad k = 0, 1 \ldots$$

RTLSEVP algorithm contains sequence of linear eigenproblems

$$T_k(\lambda)u = (B(\theta_k) - \lambda I)u = 0, \quad \text{for} \quad k = 0, 1 \ldots$$

- Sequence of QEPs/EVPs converges $\rightarrow$ Use previously gained information
- A method that can make use of all previous information by performing thick starts is the Nonlinear Arnoldi method

Solving the Sequence of Eigenproblems

RTLSQEP algorithm contains sequence of quadratic eigenproblems

$$T_k(\lambda)u = (W_k + \lambda I)^2 u - \delta^{-2} hh^T u = 0, \quad \text{for} \quad k = 0, 1 \ldots$$

RTLSEVP algorithm contains sequence of linear eigenproblems

$$T_k(\lambda)u = (B(\theta_k) - \lambda I)u = 0, \quad \text{for} \quad k = 0, 1 \ldots$$

- Sequence of QEPs/EVPs converges $\rightarrow$ Use previously gained information
- A method that can make use of all previous information by performing thick starts is the Nonlinear Arnoldi method

Nonlinear Arnoldi

With $\quad T_k(\lambda) = (W_k + \lambda I)^2 - \delta^{-2} h h^T \quad$ or $\quad T_k(\lambda) = B(\theta_k) - \lambda I$

Algorithm 1 Nonlinear Arnoldi [Meerbergen '01, V. '03]

1: Start with initial basis V, $V^T V = I$
2: Determine preconditioner $M \approx T(\sigma)^{-1}$, σ close to wanted eigenvalue
3: Find rightmost/smallest eigenvalue λ of $V^T T_k(\lambda) V y = 0$ and corresponding eigenvector y
4: Set $u = Vy$, $r = T_k(\lambda) u$
5: **while** $\|r\| / \|u\| > \epsilon$ **do**
6: $\quad v = Mr$
7: $\quad v = v - VV^T v$
8: $\quad \tilde{v} = v / \|v\|$, $V = [V, \tilde{v}]$
9: $\quad$ Find rightmost/smallest eigenvalue λ of $V^T T_k(\lambda) V y = 0$ and corresponding eigenvector y
10: $\quad$ Set $u = Vy$, $r = T_k(\lambda) u$
11: **end while**

Comments on Nonlinear Arnoldi

Main advantage: When solving $T_k(\lambda)u = 0$ in step k, start with complete search space V from preceding steps

Projected problems can be updated very cheaply:

RTLSQEP: $V^T T_k(\lambda)V = ((W_k + \lambda I)V)^T((W_k + \lambda I)V) - \delta^{-2}(V^T h)(V^T h)^T y$

with

$$W_k = L^{-T}(A^T A + f(x_k)I)L^{-1} = L^{-T}A^T AL^{-1} + f(x_k)L^{-T}L^{-1}$$

Thus only $L^{-T}A^T AL^{-1}V$, $L^{-T}L^{-1}V$, $V^T h$ and V are needed.

RTLSEVP: $V^T T_k(\lambda)V = V^T(B(\theta_k) - \lambda I)V = V^T MV + \theta_k V^T NV - \lambda I$

where

$$\begin{aligned} V^T MV &= ([A, b]V)^T([A, b]V) \\ \theta_k V^T NV &= \theta_k([L, 0]V)^T([L, 0]V) - \delta^2\theta_k(e_{n+1}^T V)^T(e_{n+1}^T V) \end{aligned}$$

Thus only have to store $[A, b]V$, $[L, 0]V$ and V.

Comments on Nonlinear Arnoldi

Main advantage: When solving $T_k(\lambda)u = 0$ in step k, start with complete search space V from preceding steps

Projected problems can be updated very cheaply:

RTLSQEP: $V^T T_k(\lambda) V = ((W_k + \lambda I)V)^T ((W_k + \lambda I)V) - \delta^{-2}(V^T h)(V^T h)^T y$

with

$$W_k = L^{-T}(A^T A + f(x_k)I)L^{-1} = L^{-T}A^T A L^{-1} + f(x_k)L^{-T}L^{-1}$$

Thus only $L^{-T}A^T A L^{-1} V$, $L^{-T}L^{-1}V$, $V^T h$ and V are needed.

RTLSEVP: $V^T T_k(\lambda) V = V^T(B(\theta_k) - \lambda I)V = V^T M V + \theta_k V^T N V - \lambda I$

where

$$\begin{aligned} V^T M V &= ([A,b]V)^T([A,b]V) \\ \theta_k V^T N V &= \theta_k([L,0]V)^T([L,0]V) - \delta^2\theta_k(e_{n+1}^T V)^T(e_{n+1}^T V) \end{aligned}$$

Thus only have to store $[A,b]V$, $[L,0]V$ and V.

Comments on Nonlinear Arnoldi

Main advantage: When solving $T_k(\lambda)u = 0$ in step k, start with complete search space V from preceding steps

Projected problems can be updated very cheaply:

RTLSQEP: $V^T T_k(\lambda) V = ((W_k + \lambda I)V)^T ((W_k + \lambda I)V) - \delta^{-2}(V^T h)(V^T h)^T y$

with

$$W_k = L^{-T}(A^T A + f(x_k)I)L^{-1} = L^{-T}A^T A L^{-1} + f(x_k)L^{-T}L^{-1}$$

Thus only $L^{-T}A^T A L^{-1}V$, $L^{-T}L^{-1}V$, $V^T h$ and V are needed.

RTLSEVP: $V^T T_k(\lambda) V = V^T(B(\theta_k) - \lambda I)V = V^T M V + \theta_k V^T N V - \lambda I$

where

$$\begin{aligned} V^T M V &= ([A, b]V)^T([A, b]V) \\ \theta_k V^T N V &= \theta_k([L, 0]V)^T([L, 0]V) - \delta^2 \theta_k (e_{n+1}^T V)^T (e_{n+1}^T V) \end{aligned}$$

Thus only have to store $[A, b]V$, $[L, 0]V$ and V.

Comments on Nonlinear Arnoldi

- Main part of W_k is not changing within the sequence of QEPs
- W_k is not needed explicitely
- Also $M = [A, b]^T [A, b]$ and $N = \begin{pmatrix} L^T L & 0 \\ 0 & -\delta^2 \end{pmatrix}$ not needed explicitely
- All stored small matrices are independent of θ_k, $f(x_k)$ and δ
- Just append one column every inner iteration step of Nonlinear Arnoldi
- No matrix-matrix multiplication is performed
- Complexity is of the order $\mathcal{O}(mn)$, i.e. several matrix-vector multiplications with A

L-curve

How to determine hyperparamter δ in constraint condition $\|Lx\| \leq \delta$?

Several methods are available:
Discrepancy principle, Cross validation, Information Criteria, L-curve

Idea of the L-curve:

- Developed to balance $\|Ax_\lambda - b\|^2$ and $\|Lx_\lambda\|^2$ in Tikhonov approach $\|Ax - b\|^2 + \lambda\|Lx\|^2 = \min_x$
- Works as well for $\|Ax_\delta - b\|^2 = \min_x$ subject to $\|Lx_\delta\| \leq \delta$
- Can be extended to $f(x_\delta) = \dfrac{\|Ax_\delta - b\|^2}{1 + \|x_\delta\|^2} = \min_x$ s.t. $\|Lx_\delta\| \leq \delta$
- Choose set of δ_i, $i = 1, \dots \ell$ and solve one RTLS problem for each δ_i

Note, there is no discrete Picard condition for RTLS problems,
nor there exists a tool like GSVD to obtain L-curve analytically.

We simply try to balance the function value $f(x_\delta)$ and the size of the norm $\|Lx_\delta\|$, which seems to work fine in most cases.

L-curve

How to determine hyperparamter δ in constraint condition $\|Lx\| \leq \delta$?

Several methods are available:
Discrepancy principle, Cross validation, Information Criteria, L-curve

Idea of the L-curve:

- Developed to balance $\|Ax_\lambda - b\|^2$ and $\|Lx_\lambda\|^2$ in Tikhonov approach $\|Ax - b\|^2 + \lambda\|Lx\|^2 = \min_x$
- Works as well for $\|Ax_\delta - b\|^2 = \min_x \quad$ subject to $\quad \|Lx_\delta\| \leq \delta$
- Can be extended to $f(x_\delta) = \dfrac{\|Ax_\delta - b\|^2}{1 + \|x_\delta\|^2} = \min_x \quad$ s.t. $\quad \|Lx_\delta\| \leq \delta$
- Choose set of δ_i, $i = 1, \dots \ell$ and solve one RTLS problem for each δ_i

Note, there is no discrete Picard condition for RTLS problems,
nor there exists a tool like GSVD to obtain L-curve analytically.

We simply try to balance the function value $f(x_\delta)$ and the size of the norm $\|Lx_\delta\|$, which seems to work fine in most cases.

L-curve

How to determine hyperparamter δ in constraint condition $\|Lx\| \leq \delta$?

Several methods are available:
Discrepancy principle, Cross validation, Information Criteria, L-curve

Idea of the L-curve:

- Developed to balance $\|Ax_\lambda - b\|^2$ and $\|Lx_\lambda\|^2$ in Tikhonov approach $\|Ax - b\|^2 + \lambda\|Lx\|^2 = \min_x$
- Works as well for $\|Ax_\delta - b\|^2 = \min_x$ subject to $\|Lx_\delta\| \leq \delta$
- Can be extended to $f(x_\delta) = \dfrac{\|Ax_\delta - b\|^2}{1 + \|x_\delta\|^2} = \min_x$ s.t. $\|Lx_\delta\| \leq \delta$
- Choose set of δ_i, $i = 1, \dots \ell$ and solve one RTLS problem for each δ_i

Note, there is no discrete Picard condition for RTLS problems,
nor there exists a tool like GSVD to obtain L-curve analytically.

We simply try to balance the function value $f(x_\delta)$ and the size of the norm $\|Lx_\delta\|$, which seems to work fine in most cases.

L-curve

- Use Nonlinear Arnoldi to solve sequence of RTLS problems
- This means solving a sequence of a sequence of QEPs or EVPs
- Updating projected problems is trivial, because all stored matrices are independent of hyperparameter δ_i
- Reusing search space V during sequence of sequence of eigenproblems
- If search space grows too large, include restart strategy

Restart strategy:

- If dimension $p \ll n$ of search space span(V) is reached, restart Nonlinear Arnoldi by subspace of $q < p$ eigenvectors corresponding to rightmost/smallest eigenvalues of $T_k(\lambda)$.
- Values can be set in advance (e.g. $p = 45, q = 5$), nice for memory allocation purposes
- Effect of restart: It purges out 'old' information corresponding to values of δ that highlight other parts of the problem

L-curve

- Use Nonlinear Arnoldi to solve sequence of RTLS problems
- This means solving a sequence of a sequence of QEPs or EVPs
- Updating projected problems is trivial, because all stored matrices are independent of hyperparameter δ_i
- Reusing search space V during sequence of sequence of eigenproblems
- If search space grows too large, include restart strategy

Restart strategy:

- If dimension $p \ll n$ of search space span(V) is reached, restart Nonlinear Arnoldi by subspace of $q < p$ eigenvectors corresponding to rightmost/smallest eigenvalues of $T_k(\lambda)$.
- Values can be set in advance (e.g. $p = 45, q = 5$), nice for memory allocation purposes
- Effect of restart: It purges out 'old' information corresponding to values of δ that highlight other parts of the problem

Example 1 – RTLSQEP solves 1 RTLS problem

Problem *deriv2* from Toolbox of Per-Christian Hansen, with a dimension of $m = n = 1000$ and a noise level of 1%.

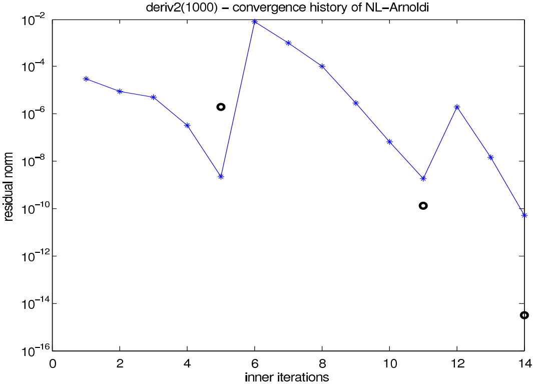

Figure: Convergence history of RTLSQEP

Three quadratic eigenproblems are solved, search space V that is built up during the first two QEPs contains such good information that final QEP is solved within only 2 MatVecs.

Example 2 – RTLSEVP solves 1 RTLS problem

Problem *phillips* from Toolbox of Per-Christian Hansen, with a dimension of $m = n = 2000$ and a noise level of 1%.

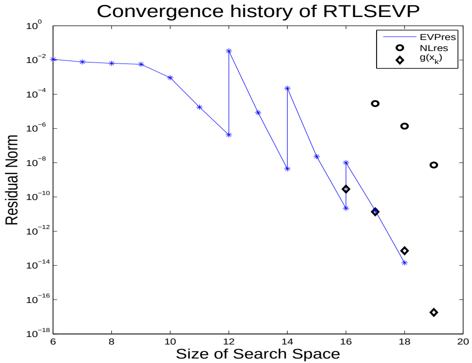

Figure: Convergence history of RTLSEVP

Four eigenproblems are solved, basis V build up during first EVP contains such good information that the last three EVPs solved within 2 MatVecs each.

Example 3 – RLSQEP solves 20 RLS problem

- Problem *phillips*(2000), discretized Fredholm integral equation of first-kind
- L is 1D discrete first order derivative operator
- Noise level 20% of average absolute value of $[A, b]$
- For L-curve: $\delta_i = \delta_{true} \cdot (0,0001 \dots 100)$ with $\delta_{true} = \|Lx_{true}\|$

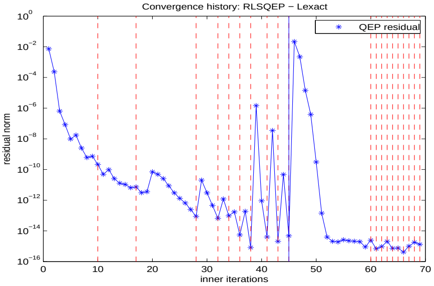

Figure: Convergence history of RLSQEP for different δ_i

Search space build up during first QEP contains such good information that the following QEPs are solved in much less MatVecs.

Example 4 – RTLSQEP solves 20 RTLS problem

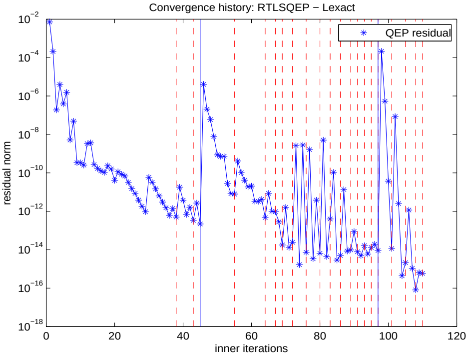

Figure: Convergence history of RTLSQEP for different δ_i

- Restart performed if dimension of search space exeeds $p = 45$
- With subspace corresponding to $q = 5$ rightmost eigenvalues
- Each RTLS problem is solved by very few QEPs
- Search space build up during first RTLS problem contains such good information that following problems are solved within less MatVecs

Example 4 – L-curve of 20 RTLS problem

- $\delta_i = \delta_{true} \cdot (0,0001 \ldots 100)$, $\delta_{true} = \|Lx_{true}\|$, $i = 1, \ldots 20$

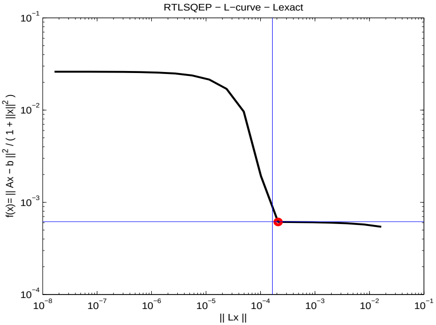

Figure: L-curve of RTLSQEP

- L-curve of RTLS looks similar to L-curve of RLS
- 120 inner iterations for 20 RTLS problems ($\rightarrow$ 500 MatVecs)
- 70 inner iterations for 20 RLS problems ($\rightarrow$ 280 MatVecs)
- Computation time $4,2$sec for RTLS, and $2,4$sec for RLS

Example 5 – RTLSQEP solves 20 RTLS problem

- Problem *deriv2*(2000), discretized Fredholm integral equation of first-kind
- $\delta_i = \delta_{true} \cdot (0,0001 \ldots 100)$, $\delta_{true} = \|Lx_{true}\|$, $i = 1, \ldots 20$

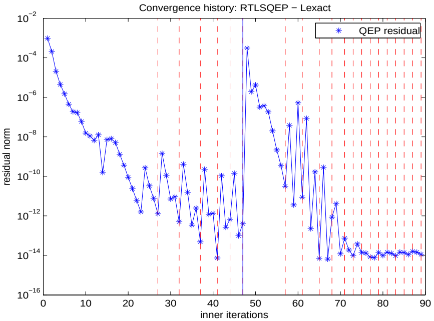

Figure: Convergence history of RTLSQEP for different δ_i

- 90 inner iterations for 20 RTLS problems ($\rightarrow$ 360 MatVecs)
- Computation time 3sec (resp. 2,3sec for 20 RLS problems)

Example 5 – RTLSQEP solves 20 RTLS problem

- Problem *deriv2*(2000), discretized Fredholm integral equation of first-kind
- $\delta_i = \delta_{true} \cdot (0,0001 \ldots 100)$, $\delta_{true} = \|Lx_{true}\|$, $i = 1, \ldots 20$

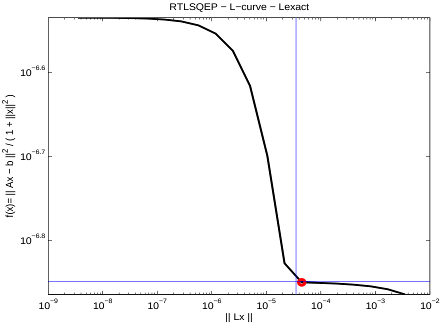

Figure: L-curve of RTLSQEP

- 90 inner iterations for 20 RTLS problems ($\rightarrow$ 360 MatVecs)
- Computation time 3sec (resp. 2, 3sec for 20 RLS problems)

Conclusions

- RTLS problems can be solved efficiently by sequence of EVPs or QEPs
- Nonlinear Arnoldi can reuse all previous information
- Determine hyperparameter δ via L-curve
- Restart strategy necessary
- Computational complexity stays $\mathcal{O}(mn)$, if number of MatVecs smaller n

References

- R. Renaut and H. Guo, Efficient Algorithms for Solution of Regularized Total Least Squares, SIAM J. Matrix Anal. Appl., 2005, vol. 26, p. 457 – 476
- D. Sima, S. Van Huffel and G.H. Golub, Regularized total least squares based on quadratic eigenvalue problem solvers, 2004, BIT Numerical Mathematics, vol. 44, p. 793 – 812
- J. Lampe and H. Voss, On a Quadratic Eigenproblem Occuring in Regularized Total Least Squares, Comput. Stat. Data Anal., 2007, vol. 52/2, p. 1090 – 1102
- J. Lampe and H. Voss, Global convergence of RTLSQEP: a solver of regularized total least squares problems via quadratic eigenproblems, 2008, Math. Modelling Anal., vol. 13, p. 55 – 66
- J. Lampe and H. Voss, A fast algorithm for solving regularized total least squares problems, 2008, Electr. Trans. Numer. Anal., vol. 31, p. 12 – 24
- J. Lampe and H. Voss, Solving Regularized Total Least Squares Problems Based on Eigenproblems, 2008, to appear in Taiwanese Journal of Mathematics
- J. Lampe and H. Voss, Efficient determination of the Hyperparameter in Regularized Total Least Squares Problems, 2009, technical report,