Efficient Determination of the Hyperparameter via L-curve in Large Scale Least Squares and Total Least Squares Problems

Jörg Lampe
joerg.lampe@tuhh.de

Hamburg University of Technology
Institute of Numerical Simulation

Joint work with Heinrich Voß

8th GAMM Workshop on Applied and Numerical Linear Algebra

TUHH
Technische Universität Hamburg-Harburg

Outline

Outline

1. Background
   - LS and TLS problems
   - RLS and RTLS problems
2. Solving RLS and RTLS problems
   - Solving RLS problem via QEP
   - Solving RTLS problem via QEPs
3. Determining the hyperparamter
   - L-curve
   - Numerical Examples
4. Conclusions

Least Squares

Consider overdetermined linear system

$$Ax \approx b, \quad A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m, m \geq n$$

with A and b contaminated by noise.

Least Squares (LS) approach:

$$\|Ax - b\|^2 = \text{min!}$$

or equivalently

$$\|\Delta b\|^2 = \text{min!} \quad \text{subject to} \quad Ax = b + \Delta b.$$

Total Least Squares (TLS) approach is suitable:

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

Least Squares

Consider overdetermined linear system

$$Ax \approx b, \quad A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m, m \geq n$$

with A and b contaminated by noise.

Least Squares (LS) approach:

$$\|Ax - b\|^2 = \text{min!}$$

or equivalently

$$\|\Delta b\|^2 = \text{min!} \quad \text{subject to} \quad Ax = b + \Delta b.$$

Total Least Squares (TLS) approach is suitable:

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

Total Least Squares

With singular value decomposition of $[A, b]$

$$[A, b] = U\Sigma V^T, \quad \Sigma = \text{diag}\{\sigma_1, \dots, \sigma_{n+1}\}$$

and $\sigma'_1 \geq \dots \geq \sigma'_n$ singular values of A.

Lemma (Golub, van Loan 1980)

If $\sigma_{n+1}([A, b]) < \sigma'_n(A)$ holds, a unique TLS solution exist.

Closed form solution of TLS problem:

$$x_{TLS} = (A^TA - \sigma_{n+1}^2 I)^{-1}A^Tb.$$

or equivalently

$$\begin{pmatrix} x_{TLS} \\ -1 \end{pmatrix} = -\frac{1}{v^{n+1}(n+1)} v^{n+1}.$$

Total Least Squares

With singular value decomposition of $[A, b]$

$$[A, b] = U\Sigma V^T, \quad \Sigma = \text{diag}\{\sigma_1, \dots, \sigma_{n+1}\}$$

and $\sigma_1' \geq \dots \geq \sigma_n'$ singular values of A.

Lemma (Golub, van Loan 1980)

If $\sigma_{n+1}([A, b]) < \sigma_n'(A)$ holds, a unique TLS solution exist.

Closed form solution of TLS problem:

$$x_{TLS} = (A^T A - \sigma_{n+1}^2 I)^{-1} A^T b.$$

or equivalently

$$\begin{pmatrix} x_{TLS} \\ -1 \end{pmatrix} = -\frac{1}{v^{n+1}(n+1)} v^{n+1}.$$

Outline

Regularized (Total) Least Squares

For ill-conditioned LS and TLS problems regularization is necessary.

Apply Tikhonov regularization with hyperparameter λ or
adding quadratic constraint with hyperparameter δ.

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

with $\delta > 0$ and $L \in \mathbb{R}^{k \times n}$, $k \le n$ defining a seminorm on the solution.

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.

A solution of RLS problem always exists.

Assume active quadratic constraint: $\delta < \|Lx_{LS}\|$ and $\delta < \|Lx_{TLS}\|$ respectively.

Hence $\|Lx\| \le \delta$ replaced by $\|Lx\| = \delta$.

Regularized (Total) Least Squares

For ill-conditioned LS and TLS problems regularization is necessary.

Apply Tikhonov regularization with hyperparameter λ or
adding quadratic constraint with hyperparameter δ.

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

with $\delta > 0$ and $L \in \mathbb{R}^{k \times n}$, $k \le n$ defining a seminorm on the solution.

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.

A solution of RLS problem always exists.

Assume active quadratic constraint: $\delta < \|Lx_{LS}\|$ and $\delta < \|Lx_{TLS}\|$ respectively.

Hence $\|Lx\| \le \delta$ replaced by $\|Lx\| = \delta$.

Outline

Regularized Least Squares

How to solve a quadratically constrained Least Squares Problem?
One possibility is by one QEP (another by LSTRS):

Consider Lagrangian

$$\mathcal{L}(x,\mu) = \|Ax - b\|^2 + \mu(\|Lx\|^2 - \delta^2)$$

with first-order optimality conditions

$$\begin{aligned} 2(A^TA + \mu L^TL)x &= 2A^Tb, \\ \|Lx\|^2 &= \delta^2. \end{aligned}$$

Lemma (Gander 1981)

Choose largest value of μ to obtain the RLS solution.

Assume L is square and nonsingular:
Substitute $z := Lx$, $W := L^{-T}A^TAL^{-1}$ and $h := L^{-T}A^Tb$

$$\begin{aligned} Wz + \mu z &= h, \\ z^Tz &= \delta^2. \end{aligned}$$

Regularized Least Squares

How to solve a quadratically constrained Least Squares Problem?
One possibility is by one QEP (another by LSTRS):

Consider Lagrangian

$$\mathcal{L}(x,\mu) = \|Ax - b\|^2 + \mu(\|Lx\|^2 - \delta^2)$$

with first-order optimality conditions

$$\begin{aligned} 2(A^TA + \mu L^TL)x &= 2A^Tb, \\ \|Lx\|^2 &= \delta^2. \end{aligned}$$

Lemma (Gander 1981)

Choose largest value of μ to obtain the RLS solution.

Assume L is square and nonsingular:
Substitute $z := Lx$, $W := L^{-T}A^TAL^{-1}$ and $h := L^{-T}A^Tb$

$$\begin{aligned} Wz + \mu z &= h, \\ z^Tz &= \delta^2. \end{aligned}$$

Connection to QEP

Remark: If $rank(L) = k < n$, basis for range and kernel is needed.

Denoting $u := (W + \mu I)^{-2} h \quad \Rightarrow \quad h^T u = z^T z = \delta^2 \quad \Rightarrow \quad h = \delta^{-2} h h^T u$

$$(W + \mu I)^2 u - \delta^{-2} h h^T u = 0. \qquad \text{(Gander, Golub, von Matt 1989)}$$

Reconstruct x_{RLS} from rightmost eigenpair $(\hat{\mu}, \hat{u})$:
Scale $\tilde{u} = \delta^2 \frac{\hat{u}}{h^T \hat{u}}$, then it holds

$$x_{RLS} = L^{-T}(W + \hat{\mu} I)\tilde{u}.$$

- The same idea can be carried over to RTLS problems
- In general there exist no closed form solution for x_{RTLS}
- This leads to a converging sequence of QEPs

Connection to QEP

Remark: If $rank(L) = k < n$, basis for range and kernel is needed.

Denoting $u := (W + \mu I)^{-2} h \quad \Rightarrow \quad h^T u = z^T z = \delta^2 \quad \Rightarrow \quad h = \delta^{-2} h h^T u$

$$(W + \mu I)^2 u - \delta^{-2} h h^T u = 0. \qquad \text{(Gander, Golub, von Matt 1989)}$$

Reconstruct x_{RLS} from rightmost eigenpair $(\hat{\mu}, \hat{u})$:
Scale $\tilde{u} = \delta^2 \frac{\hat{u}}{h^T \hat{u}}$, then it holds

$$x_{RLS} = L^{-T} (W + \hat{\mu} I) \tilde{u}.$$

- The same idea can be carried over to RTLS problems
- In general there exist no closed form solution for x_{RTLS}
- This leads to a converging sequence of QEPs

Outline

Regularized Total Least Squares

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

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

Consider Lagrangian again

$$\mathcal{L} = \frac{\|Ax-b\|^2}{1+\|x\|^2} + \mu(\|Lx\|^2 - \delta^2),$$

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

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

Fixed-point Iteration for RTLS

Look at nonlinear first-order equation again

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

What about the iteration with maximal real Lagrange parameter each step:

$$(A^TA - f(x^k)I)x^{k+1} + \lambda L^TLx^{k+1} = A^Tb, \quad \|Lx^{k+1}\|^2 = \delta^2 \qquad ?$$

Theorem (Lampe, Voß 2007)

Any limit point of the sequence $\{x^k\}$ constructed by the fixed point algorithm above is a global minimizer of

$$f(x) = \frac{\|Ax - b\|^2}{1 + \|x\|^2} \quad s.t. \ \ \|Lx\|^2 = \delta^2.$$

(Under mild condition concerning starting vector x^0)

Fixed-point Iteration for RTLS

Look at nonlinear first-order equation again

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

What about the iteration with maximal real Lagrange parameter each step:

$$(A^T A - f(x^k) I) x^{k+1} + \lambda L^T L x^{k+1} = A^T b, \quad \|L x^{k+1}\|^2 = \delta^2 \qquad ?$$

Theorem (Lampe, Voß 2007)

Any limit point of the sequence $\{x^k\}$ constructed by the fixed point algorithm above is a global minimizer of

$$f(x) = \frac{\|Ax - b\|^2}{1 + \|x\|^2} \quad s.t. \;\; \|Lx\|^2 = \delta^2.$$

(Under mild condition concerning starting vector x^0)

Connection to QEPs

Assume L is square and nonsingular:
Substitute $z := Lx^{k+1}$, $W_k := L^{-T}(A^TA - f(x^k)I)L^{-1}$ and $h := L^{-T}A^Tb$

$$\begin{aligned} W_k z + \lambda z &= h, \\ z^T z &= \delta^2. \end{aligned}$$

Denoting again
$u := (W_k + \lambda I)^{-2} h \quad \Rightarrow \quad h^T u = z^T z = \delta^2 \quad \Rightarrow \quad h = \delta^{-2} h h^T u$

$$(W_k + \lambda I)^2 u - \delta^{-2} h h^T u = 0. \qquad \text{(QEPs)}$$

Reconstruct x^{k+1} from rightmost eigenpair $(\hat{\lambda}, \hat{u})$: With $\tilde{u} = \delta^2 \frac{\hat{u}}{h^T \hat{u}}$ it holds

$$x^{k+1} = L^{-T}(W_k + \hat{\lambda} I)\tilde{u}.$$

Remark: If the matrix $(W_k + \hat{\lambda} I) \geq 0$ is singular, the rightmost eigenpair might not contain the solution of the f.o.c. An additional linear system has to be solved to cope with nonunique solutions. This is not the generic case.

Connection to QEPs

Assume L is square and nonsingular:
Substitute $z := Lx^{k+1}$, $W_k := L^{-T}(A^TA - f(x^k)I)L^{-1}$ and $h := L^{-T}A^Tb$

$$\begin{aligned} W_k z + \lambda z &= h, \\ z^T z &= \delta^2. \end{aligned}$$

Denoting again
$u := (W_k + \lambda I)^{-2} h \quad \Rightarrow \quad h^T u = z^T z = \delta^2 \quad \Rightarrow \quad h = \delta^{-2} h h^T u$

$$(W_k + \lambda I)^2 u - \delta^{-2} h h^T u = 0. \qquad \text{(QEPs)}$$

Reconstruct x^{k+1} from rightmost eigenpair $(\hat{\lambda}, \hat{u})$: With $\tilde{u} = \delta^2 \frac{\hat{u}}{h^T \hat{u}}$ it holds

$$x^{k+1} = L^{-T}(W_k + \hat{\lambda} I)\tilde{u}.$$

Remark: If the matrix $(W_k + \hat{\lambda} I) \geq 0$ is singular, the rightmost eigenpair might not contain the solution of the f.o.c. An additional linear system has to be solved to cope with nonunique solutions. This is not the generic case.

Algorithm for solving RTLS

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

Algorithm 1 RTLSQEP [Sima/van Huffel/Golub 2005]

Let x^0 be an initial guess. Compute $f(x_0) = \frac{\|Ax_0 - b\|^2}{1 + \|x_0\|^2}$.
Set $k = 1$
while Not converged **do**
    Solve

$$T_k(\lambda) u = 0$$

    for eigenpair (u_k, λ_k) corresponding to the rightmost λ_k
    Scale $\tilde{u} = \delta^2 \frac{u_k}{h^T u_k}$
    Compute $x^{k+1} = L^{-1}(W_k + \lambda_k I)\tilde{u}$ and $f(x^{k+1}) = \dfrac{\|Ax^{k+1} - b\|^2}{1 + \|x^{k+1}\|^2}$
    Set $k = k + 1$
end while

Solving the QEPs

RTLSQEP algorithm contains sequence of quadratic eigenproblems

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

- Sequence of QEPs converges $\rightarrow$ Use previously gained information
- Linearization to EVP of double size is not appropriate
- Second Order Krylov Methods use one starting vector each QEP, i.e. use solution vector of previous QEP (SOAR, Li/Ye)
- Can we use more information than only one vector?
- Yes, a method that can make use of all previous information by performing thick starts is the Nonlinear Arnoldi method

Solving the QEPs

RTLSQEP algorithm contains sequence of quadratic eigenproblems

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

- Sequence of QEPs converges $\rightarrow$ Use previously gained information
- Linearization to EVP of double size is not appropriate
- Second Order Krylov Methods use one starting vector each QEP, i.e. use solution vector of previous QEP (SOAR, Li/Ye)
- Can we use more information than only one vector?

- Yes, a method that can make use of all previous information by performing thick starts is the Nonlinear Arnoldi method

Nonlinear Arnoldi

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

Algorithm 2 Nonlinear Arnoldi [Voß 2003]

```
Start with initial basis V, V^T V = I
Determine preconditioner M ≈ T(σ)^{-1}, σ close to wanted eigenvalue
Find rightmost eigenvalue λ of V^T T_k(λ) V y = 0 and corr. eigenvector y
Set u = V y, r = T_k(λ) u
while ||r||/||u|| > ε do
    v = M r
    v = v − V V^T v
    ṽ = v/||v||, V = [V, ṽ]
    Find rightmost eigenvalue λ of V^T T_k(λ) V y = 0 and corr. eigenvector y
    Set u = V y, r = T_k(λ) u
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

- No preconditioner is needed, i.e. $M = I$
- Projected problems can be updated very cheaply:

$$V^T T_k(\lambda)Vy = \big((W_k + \lambda I)V\big)^T \big((W_k + \lambda I)V\big)y - \delta^{-2}(V^T h)(V^T h)^T y = 0$$

Only V, $W_k V$ and $V^T h$ are needed. A closer look shows

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

- Main part of W_k is not changing within the sequence of QEPs
- Simply store $L^{-T}A^T AL^{-1}V$, $L^{-T}L^{-1}V$, $V^T h$ and V for all QEPs!
- Just append one column every inner iteration step
- No matrix-matrix multiplication is performed

Outline

1 Background
- LS and TLS problems
- RLS and RTLS problems

2 Solving RLS and RTLS problems
- Solving RLS problem via QEP
- Solving RTLS problem via QEPs

3 Determining the hyperparamter
- L-curve
- Numerical Examples

4 Conclusions

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 $\|Lx_\lambda\|$ in Tikhonov $\|Ax - b\| + \lambda\|Lx\| = \min_x!$
- Can be extended to $f(x_\delta) = \dfrac{\|Ax_\delta - b\|^2}{1 + \|x_\delta\|^2}$ and $\delta = \|Lx_\delta\|$
- Choose set of δ_i, $i = 1, \ldots$ and solve one RLS/RTLS problem for each δ_i

Advantage when using Nonlinear Arnoldi method:

- RLS : Search space V reused during sequence of QEPs
- RTLS: Search space V reused during sequence of sequence of QEPs
- If search space grows too large, include restart strategy

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 $\|Lx_\lambda\|$ in Tikhonov
  $\|Ax - b\| + \lambda\|Lx\| = \min_x!$
- Can be extended to $f(x_\delta) = \dfrac{\|Ax_\delta - b\|^2}{1 + \|x_\delta\|^2}$ and $\delta = \|Lx_\delta\|$
- Choose set of δ_i, $i = 1, \ldots$ and solve one RLS/RTLS problem for each δ_i

Advantage when using Nonlinear Arnoldi method:

- RLS : Search space V reused during sequence of QEPs
- RTLS: Search space V reused during sequence of sequence of QEPs
- If search space grows too large, include restart strategy

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 $\|Lx_\lambda\|$ in Tikhonov $\|Ax - b\| + \lambda\|Lx\| = \min_x!$
- Can be extended to $f(x_\delta) = \dfrac{\|Ax_\delta - b\|^2}{1 + \|x_\delta\|^2}$ and $\delta = \|Lx_\delta\|$
- Choose set of δ_i, $i = 1, \ldots$ and solve one RLS/RTLS problem for each δ_i

Advantage when using Nonlinear Arnoldi method:

- RLS : Search space V reused during sequence of QEPs
- RTLS: Search space V reused during sequence of sequence of QEPs
- If search space grows too large, include restart strategy

Outline

1. Background
   - LS and TLS problems
   - RLS and RTLS problems
2. Solving RLS and RTLS problems
   - Solving RLS problem via QEP
   - Solving RTLS problem via QEPs
3. Determining the hyperparamter
   - L-curve
   - Numerical Examples
4. Conclusions

Example *phillips* (2000), RLS

- P.C. Hansen, 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 $\delta_{true} = \|Lx_{true}\|$

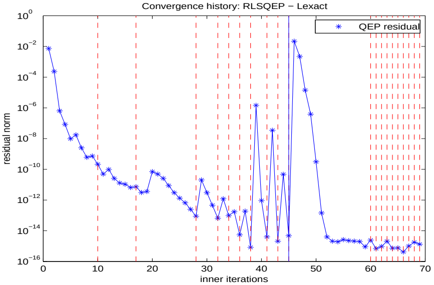

Figure: Convergence history of RLS-QEP for different δ_i

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

Example *phillips* (2000), RTLS

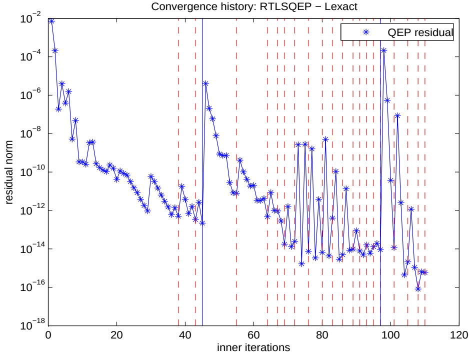

Figure: Convergence history of RTLS-QEP for different δ_i

- Restart performed if dimension of search space > 45
- 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 *phillips* (2000), RTLS

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

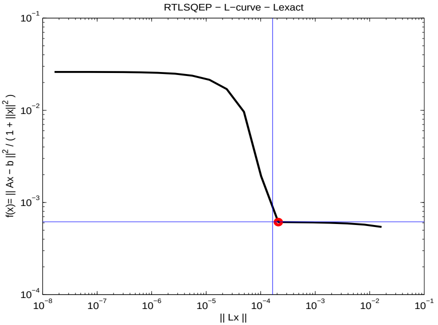

Figure: L-curve of RTLS-QEP

- 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 *deriv2* (2000), RTLS

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

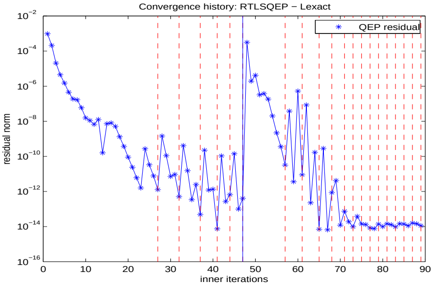

Figure: Convergence history of RTLS-QEP for different δ_i

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

Example *deriv2* (2000), RTLS

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

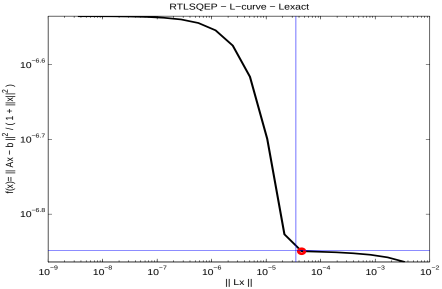

Figure: L-curve of RTLSQEP

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

Example *deriv2* (2000), RLS/RTLS

$$L = \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \quad \text{or} \quad \tilde{L} = \begin{pmatrix} \varepsilon & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix}$$

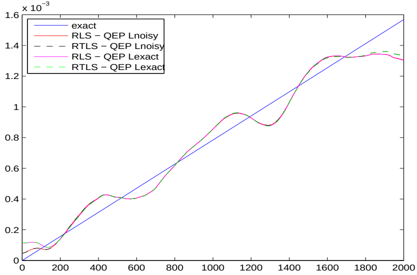

Figure: Solution curves for RLS and RTLS

- Not much difference between RLS and RTLS approach
- Little difference between approaches with different regularization matrices L and $\tilde{L}$

Conclusions

- RLS/RTLS problems can be solved efficiently by one/sequence of QEPs
- Determine δ via L-curve

Improvements of RTLSQEP method:

- Global convergence to RTLS solution was proven
- Nonlinear Arnoldi uses all previous information
- Computational complexity $\mathcal{O}(n^2)$, number of MatVecs smaller n

Remark:
For RTLS problems there is also an approach where instead of QEPs linear EVPs are solved [Renaut/Guo 2005]. When solving a sequence of EVPs via Nonlinear Arnoldi for several δ_i ($\rightarrow$ L-curve), similar advantages due to reusing the search space V are observable.
For RLS problems this can be recognized as well, when solving one problem by sequence of EVPs ($\rightarrow$ LSTRS), but solving by Nonlinear Arnoldi again.

Conclusions

- RLS/RTLS problems can be solved efficiently by one/sequence of QEPs
- Determine δ via L-curve

Improvements of RTLSQEP method:

- Global convergence to RTLS solution was proven
- Nonlinear Arnoldi uses all previous information
- Computational complexity $\mathcal{O}(n^2)$, number of MatVecs smaller n

Remark:
For RTLS problems there is also an approach where instead of QEPs linear EVPs are solved [Renaut/Guo 2005]. When solving a sequence of EVPs via Nonlinear Arnoldi for several δ_i ($\rightarrow$ L-curve), similar advantages due to reusing the search space V are observable.
For RLS problems this can be recognized as well, when solving one problem by sequence of EVPs ($\rightarrow$ LSTRS), but solving by Nonlinear Arnoldi again.