THEORY

TRS Approach:
$$\begin{array}{rl} \min & \|Gx-d\|_2^2 \\ \text{subject to} & \|x\|_2^2 \le \varepsilon^2 \end{array}$$

Tikhonov Regularization: $(G^TG+\alpha^2 I)x_\alpha = G^Td$

TRS reformulated
$$\mu_\varepsilon := \mu(A,a,\varepsilon) := \begin{array}{rl} \min & q(x) := x^TAx - 2a^Tx \\ \text{subject to} & \|x\|_2^2 \le \varepsilon^2, \end{array}$$

where:

$A = G^TG$ is $n \times n$ symmetric (nonsingular, ill-cond.)
$a = G^Td \in \mathbb{R}^n$
$\varepsilon > 0$
$x \in \mathbb{R}^n$
define λ^* optimal Lagrange multiplier
$x(0) = A^{-1}a = G^{-1}d$ is unconstrained optimum

Both TRS and Tikhonov approaches are shown to be equivalent. The Trust Region radius ε can be viewed as the regularization parameter.

One-to-one correspondence between any two of controlling parameters is established. This equivalence is exploited by the method to compute the step length.

Problem can be parameterized by four parameters:
- t, control parameter in *k(t)* and *D(t)*
- ε, trust region radius, norm of the solution
- α, Tikhonov regularization parameter
- λ, optimal Lagrange multiplier for TRS

$$D(t) = \begin{pmatrix} t & -a^T \\ -a & A \end{pmatrix}$$

$$k(t) = (\varepsilon^2+1)\lambda_{\min}(D(t)) - t$$

$$\mu_\varepsilon = \max_t k(t)$$

Relationships (isotonic):
- $-\infty < \lambda = \lambda_1(D(t)) = -\alpha^2 \le 0$
- $0 < t = \lambda + d^TG(G^TG - \lambda I)^{-1}G^Td \le \|d\|^2$
- $0 < \varepsilon = \|(G^TG-\lambda I)^{-1}G^Td\| \le \|G^{-1}d\|^2$

Upper bound corresponds to the LLSS.

PROBLEM

$Gx = Gx_{\text{true}} + \eta = d = d_{\text{true}} + \eta$
$G, n \times n$ is *singular* and *ill-conditioned*
d is usually contaminated by *noise*

Comes from discretizations of
linear equations $Tx = d$,
where T is a compact linear operator
with an unbounded inverse,
x is not continuous function of d.

Important applications:
- medical imaging – CAT, PET (computer assisted tomography)
- image restoration – space telescopes, satellite pictures
- financial computations – risk assessment, market analysis
- numerical differentiation and integration
- interior-point algorithms
- and many, many, many more...

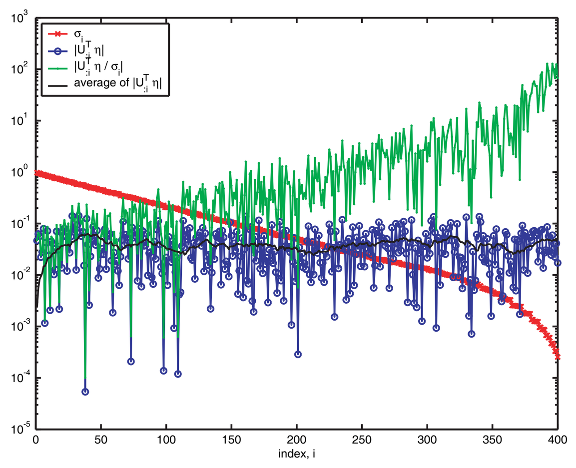
Various methods of *regularization* are often used to obtain meaningful solutions to the mathematical models that are *ill-posed*. The aim is to find algorithms for constructing *generalized solutions* that are stable under small changes in the data *d*. Regularized solutions are parameterized by one or more regularization parameters, which usually control the amount of smoothness. Solving for the best regularized solution is often associated with finding an optimal value of the regularization parameter.

Observed image:
a real telescope photo

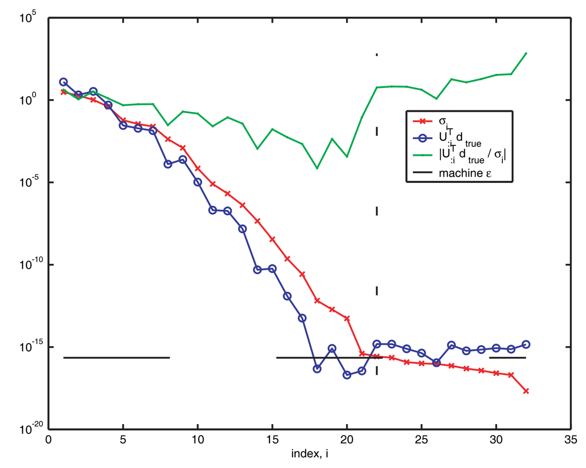
?

Regularized image:
reconstruction of the true picture

Deblurring an image:
Noise from the experimental obervations.



Solving integral equation:
Numerical noise due to finite-precision.



REGULARIZATION

Using a Parameterized Trust Region Subproblem

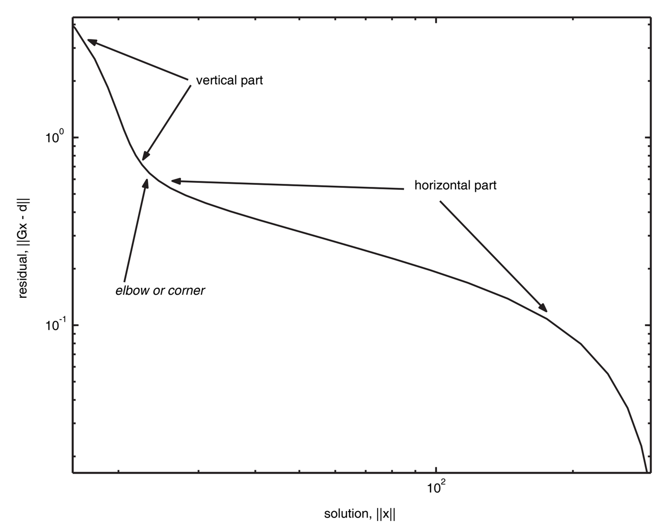
Oleg Grodzevich and Henry Wolkowicz

Department of Management Sciences, Department of Combinatorics and Optimization, University of Waterloo

MITACS

METHOD

Key ideas behind the RPTRS algorithm:
- navigating along the L-curve in a consistent manner (by increasing the regularization parameter)
- exploiting optimality conditions of the TRS to parameterize the L-curve
- stopping criteria based on the L-curve maximum curvature
- estimating the curvature of the L-curve via Gauss Quadratures
- exploiting operator matrix sparsity by using Lanczos-based methods for eigenvalue computations



$$\mathcal{L}(G,d) = \{(\log(\varepsilon), \log\|Gx(\varepsilon)-d\|_2) : \varepsilon > 0,\ x(\varepsilon) \text{ optimal for } TRS\}$$

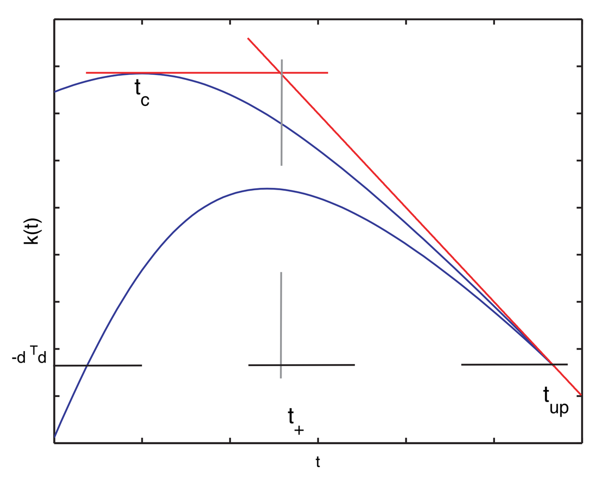
$$\kappa_\varepsilon = \varepsilon^2\mu_\varepsilon\left(2\varepsilon^2\lambda^{*2} - 2\mu_\varepsilon\lambda^* - \varepsilon\mu_\varepsilon\left(\tfrac{\partial\lambda^*}{\partial\varepsilon}\right)\right)\left(\varepsilon^4\lambda^{*2}+\mu_\varepsilon^2\right)^{-3/2}$$ curvature of the L-curve, expensive to compute directly

The L-curve is an essential tool for estimating the regularization parameter. The (nearly) vertical part corresponds to the oversmoothed solutions, while the horizontal part features solutions that are dominated by the perturbation errors to a greater extent. The transition, called the elbow, is given by a point of the maximum curvature and usually corresponds to a desireable value of the regularization parameter.

At every iteration the algorithm finds an optimal solution of TRS, *x*(ε), for a corresponding TR radius ε. It uses the derivative information to efficiently change the TR radius and move along points of the L-curve. Step length computation involves switching between the parameters ε, *t* and α, which enables the step length to be self-regulated.

$$\boxed{t_{next} = t_{current} - (\varepsilon^2+1)\lambda_{\min}(D(t_{current}))}$$

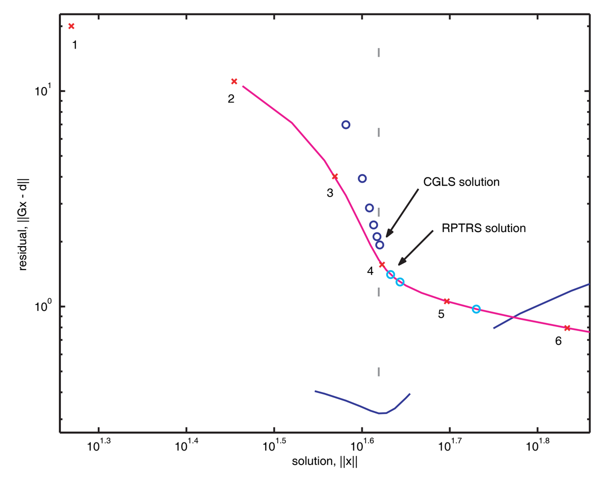
The algorithm starts to the left of the elbow and increases the value of *t* at each iteration. The uncertainty interval for the elbow is shrinked at every step and is usually reduced to a small enough interval in just a few iterations. A good approximation for the elbow is then located by a simple search (e.g. bisection).

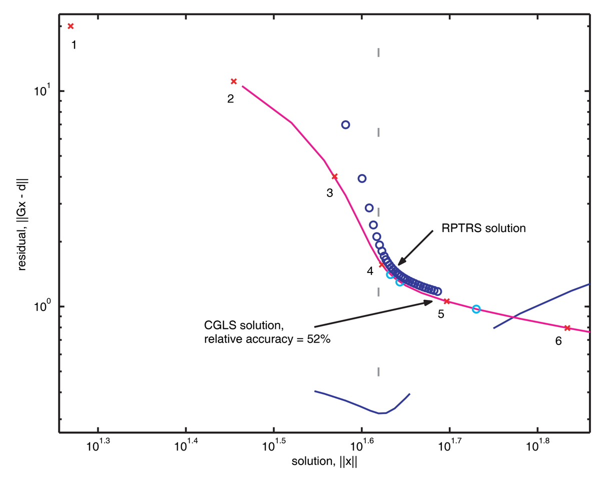


Step length computation is derived via a triangle interpolation technique

Open questions and future development:
- generalization and extension to the Lanczos-based methods, e.g. CGLS, and other types of the regularization
- accelerating eigenvalue computations by exploiting the special structure of the matrix *D(t)*
- analysis of the cases when the maximum curvature point does not correspond to the best regularized solution

RESULTS

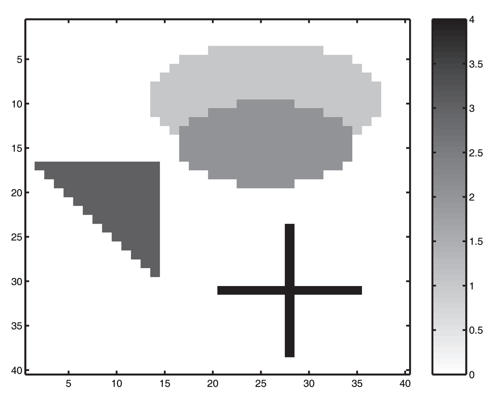


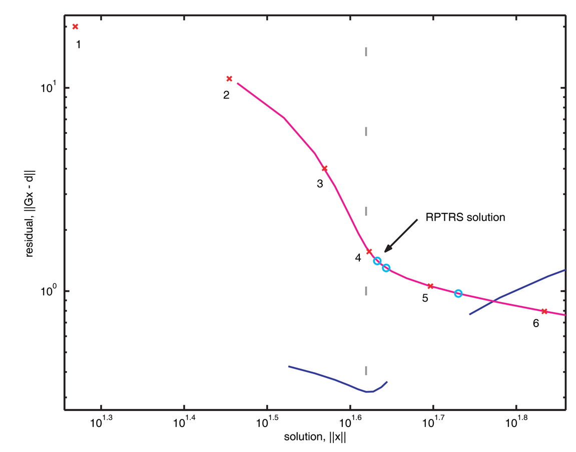


In comparison to the CGLS method the RPTRS produces comparably equivalent results in the case when CGLS is provided with the "true" norm of the error. However, the RPTRS approach does not require an *apriori* value of the norm of the noise. This is an advantage in a sense that CGLS might perform very poorly if supplied with incorrect (slightly smaller or larger) value of the error norm. On the other hand the CGLS is much faster comparing to the RPTRS.

An example of regularization algorithm (RPTRS) applied to the deblurring an image problem.

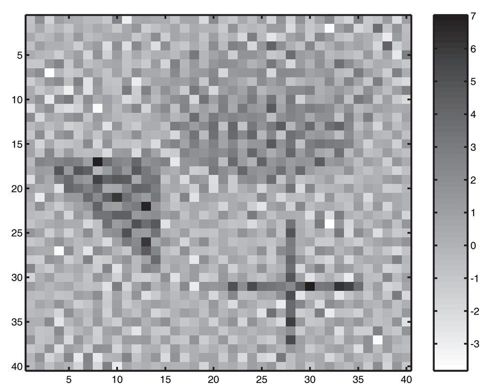
Original picture: true solution



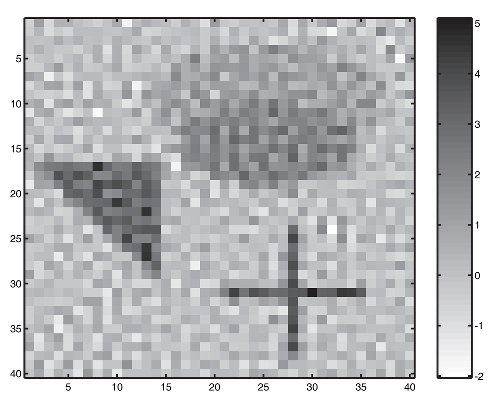


The RPTRS algorithm visits 6 points on the L-curve detecting the change in curvature and backtracks to the point of the maximum (negative) curvature. The solution is indeed very close to the best possible obtainable solution in a sense of Tikhonov regularization.

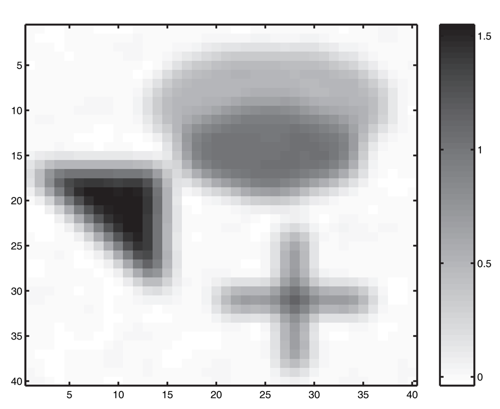
very noisy



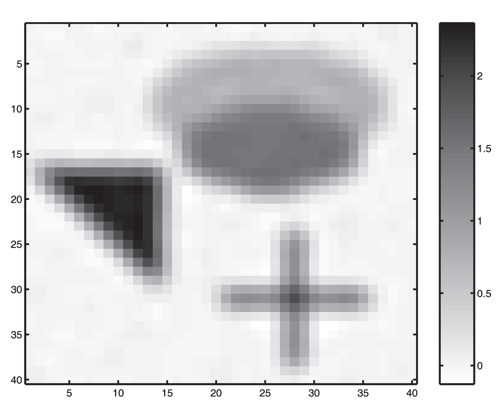
undersmooth

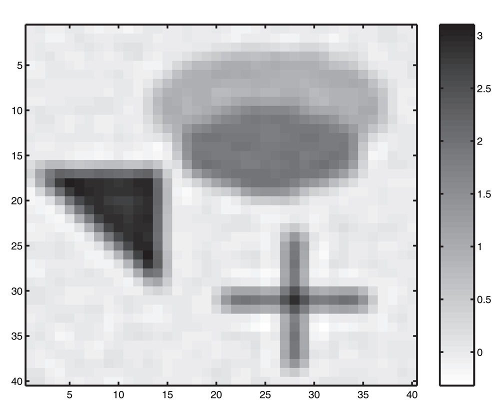


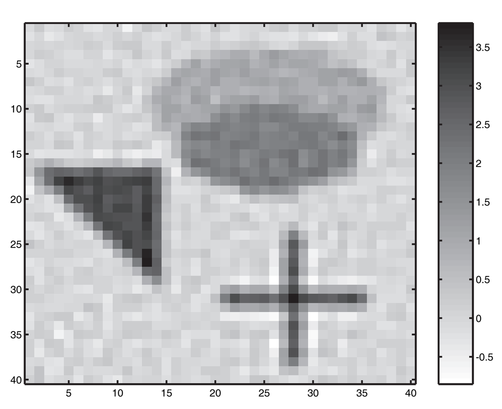
Observed picture: blurred and distorted with Gaussian-type noise



oversmooth







Subsequent solutions as the algorithm navigates along the L-curve, i.e. corresponding to different (increasing) values of the regularization parameter. One may see how noise evolves as picture becomes sharper.