Tikhonov regularization and matrix function evaluation

Paolo Novati
Department of Pure and Applied Mathematics,
University of Padova - Italy

Joint work with
Maria Rosaria Russo (Padova - Italy)
Michela Redivo Zaglia (Padova - Italy)

February 2012

The problem

We consider ill-conditioned linear systems

$$Ax = b$$

We mainly focus the attention on **full-rank** problems in which the singular values of A decay gradually to zero.

- Discretization of compact operators, as in the case of Fredholm integral equations of the first kind.
- Vandermonde type systems arising from interpolation.

Motivations

We want to construct an iterative solver able to overcome some of the typical drawback of the classical iterative solvers:

- **Semi-convergence**: the iterates initially approach the solution but quite rapidly diverge
- **Strong dependence on the parameter-choice strategy**: in order to prevent divergence a reliable stopping criterium has to be used
- **Poor accuracy**: typically holds for Krylov type methods based on the use of $A^T A$ (CGLS)

Outline

- Reformulation of the problem in term of a matrix function evaluation
- Extension to Tikhonov regularization
- Theoretical and numerical error analysis
- Filter factor analysis
- The choice of the regularization parameters
- Numerical experiments

Reformulation of the problem

The basic idea is to solve the system $Ax = b$ in two steps, first solving in some way the regularized system

$$(A + \lambda I)\, x_\lambda = b$$

(Franklin's method) and then recovering the solution x from the system

$$(A + \lambda I)^{-1} Ax = x_\lambda$$

that is equivalent to compute

$$x = f(A)x_\lambda$$

where

$$f(z) = 1 + \lambda z^{-1}$$

The choice of lambda

By the definition of f the attainable accuracy depends on the the conditioning of $(A+\lambda I)^{-1} A$. Theoretically the best situation is attained defining λ such that

$$\kappa(A+\lambda I) = \kappa((A+\lambda I)^{-1} A)$$

In the SPD case taking

$$\lambda = \sqrt{\lambda_1 \lambda_N} \approx 1/\sqrt{\kappa(A)}$$

where λ_1 and λ_N are respectively the smallest and the largest eigenvalue of A, we obtain

$$\kappa(A+\lambda I) = \kappa((A+\lambda I)^{-1} A) = \sqrt{\kappa(A)}$$

The computation of the matrix function - The Arnoldi algorithm

For the computation of $f(A)x_\lambda$ we use the standard Arnoldi method projecting the matrix A onto the Krylov subspaces generated by A and x_λ

$$K_m(A, x_\lambda) = span\{x_\lambda, Ax_\lambda, ..., A^{m-1}x_\lambda\}$$

For the construction of the subspaces $K_m(A, x_\lambda)$ the Arnoldi algorithm generates an orthonormal sequence $\{v_j\}_{j\geq 1}$, with $v_1 = x_\lambda / \|x_\lambda\|$, such that $K_m(A, x_\lambda) = span\{v_1, v_2, ..., v_m\}$. For every m,

$$AV_m = V_m H_m + h_{m+1,m} v_{m+1} e_m^T.$$

where $V_m = [v_1, v_2, ..., v_m]$, H_m is an upper Hessenberg matrix with entries $h_{i,j} = v_i^T A v_j$ and e_j is the j-th vector of the canonical basis of $\mathbb{R}^m$.

The computation of the matrix function - The ASP method

The m-th Arnoldi approximation to $x = f(A)x_\lambda$ is defined as

$$x_m = \|x_\lambda\| \, V_m f(H_m) e_1$$

At each step of the Arnoldi algorithm we have to compute the vector $w_j = Av_j$. The method theoretically converges in a finite number of steps. For the computation of $f(H_m)$ we employ the Schur-Parlett algorithm (Golub and Van Loan 1983).

Extension to Tikhonov regularization - The ATP method

The method can be extended to problems in which the exact right hand side b is affected by noise. We assume to work with a perturbed right-hand side $\overline{b} = b + e_b$. Given λ and H we solve the regularized system

$$(A^T A + \lambda H^T H) x_\lambda = A^T \overline{b}.$$

and then we approximate x by computing

$$\overline{x} = \left(A^T A\right)^{-1} (A^T A + \lambda H^T H) x_\lambda = f(Q) x_\lambda$$

where $Q = \left(H^T H\right)^{-1} \left(A^T A\right)$. As before, for the computation of $f(Q)x_\lambda$ we use the standard Arnoldi method projecting the matrix Q onto the Krylov subspaces generated by Q and x_λ. Now, at each step we have have to compute the vectors $w_j = Qv_j$, $j \geq 1$, with $v_1 = x_\lambda / \|x_\lambda\|$, that is, to solve the systems $\left(H^T H\right) w_j = \left(A^T A\right) v_j$.

Theoretical error analysis

The field of values of A is defined as

$$F(A) := \left\{ \frac{y^H A y}{y^H y}, y \in \mathbb{C}^N \setminus \{0\} \right\}$$

Theorem

Assume that $F(A) \subset \mathbb{C}^+$. Then

$$\| f(A)x_\lambda - \|x_\lambda\| \, V_m f(H_m) e_1 \| \leq K \frac{\lambda}{a^{m+1}} \prod_{i=1}^{m} h_{i+1,i} \, \|x_\lambda\| \,,$$

where $a > 0$ is the leftmost point of $F(A)$ and $2 \leq K \leq 11.08$ (Crouzeix 2007; in the symmetric case $K = 1$).

Some general considerations

- The rate of convergence is almost independent of the choice of λ, and this is confirmed by the numerical experiments.
- The error decay is related with the rate of the decay of $\prod_{i=1}^{m} h_{i+1,i}$.

Theorem

(From a result by Nevanlinna 1993) Let σ_j, $j \geq 1$, be the singular values of an operator A. If

$$\sum_{j \geq 1} \sigma_j^p < \infty \text{ for a certain } p > 0$$

then

$$\prod_{i=1}^{m} h_{i+1,i} \leq \left(\frac{\eta e p}{m} \right)^{m/p}$$

where

$$\eta \leq \frac{1+p}{p} \sum_{j \geq 1} \sigma_j^p$$

Some general considerations

- For discrete ill-posed problems the rate of decay of $\prod_{i=1}^{m} h_{i+1,i}$ is superlinear and depends on the p-summability of the singular values of A, i.e., on the degree of ill-posedness of the problem.
- Each Arnoldi-based method (CG, FOM, GMRES) shows the same rate of convergence

Error analysis in computer arithmetics for the ASP method

We need to assume that x_λ, solution of $(A+\lambda I)\, x_\lambda = b$, is approximated by $\overline{x}_\lambda$ with an accuracy depending on the choice of λ and the method used. In this way, the Arnoldi algorithm actually constructs the Krylov subspaces $K_m(A, \overline{x}_\lambda)$. Hence for the error $E_m := x - \|\overline{x}_\lambda\|\, V_m f(H_m) e_1$ we have

$$\begin{aligned} \|E_m\| &= \|f(A)x_\lambda - \|\overline{x}_\lambda\|\, V_m f(H_m) e_1\| \leq \\ &\quad \|f(A)\overline{x}_\lambda - \|\overline{x}_\lambda\|\, V_m f(H_m) e_1\| + \|f(A)\,(x_\lambda - \overline{x}_\lambda)\| \end{aligned}$$

- For small values of λ, $f(A) \approx I$ and we have that $\|E_m\| \approx \|x_\lambda - \overline{x}_\lambda\|$. The method is not able to improve the accuracy provided by the solution of the initial system.
- For large λ we have that $x_\lambda \approx \overline{x}_\lambda$, but even assuming that $\|f(A)\,(x_\lambda - \overline{x}_\lambda)\| \approx 0$, we have another lower bound due the ill-conditioning of $f(A) = A^{-1}\,(A+\lambda I)$.

Error analysis in computer arithmetics for the ATP method

The error is given by

$$E_m := f(Q)x_\lambda - p_{m-1}(Q)\overline{x}_\lambda$$

(p_{m-1} interpolates f at the eigenvalues of Q) where $(A^TA + \lambda H^TH)x_\lambda = A^Tb$ and $(A^TA + \lambda H^TH)\overline{x}_\lambda = A^T\overline{b}$. As before we can write

$$\|E_m\| \leq \|f(Q)\overline{x}_\lambda - p_{m-1}(Q)\overline{x}_\lambda\| + \|f(Q)\left(x_\lambda - \overline{x}_\lambda\right)\|.$$

Theoretically we may choose λ very large, thus oversmoothing, in order to reduce the effect of noise. Unfortunately, the main problem is that, as before, $f(Q)$ may be ill-conditioned for λ large. Even in this case we should find a compromise for the selection of a suitable value of λ, but contrary to the ASP method for noise-free problems it is difficult to design a theoretical strategy. Indeed everything depends on the problem and on the operator H.

Filter factors

Assume that A is diagonalizable, that is, $A = XDX^{-1}$ where $D = \mathrm{diag}(\lambda_1, ..., \lambda_N)$. For the ASP method we have

$$x_\lambda = \sum_{i=1}^{N} \frac{\lambda_i}{\lambda_i + \lambda} \frac{\left(X^{-1} b\right)_i}{\lambda_i} x_i,$$

where x_i is the eigenvector associated with λ_i. After the first phase, the filter factors are thus $g_i = \lambda_i \left(\lambda_i + \lambda\right)^{-1}$. The Arnoldi algorithm produces approximations of the type $x_m = p_{m-1}(A) x_\lambda$, where p_{m-1} interpolates the function f at the eigenvalues of H_m. Hence

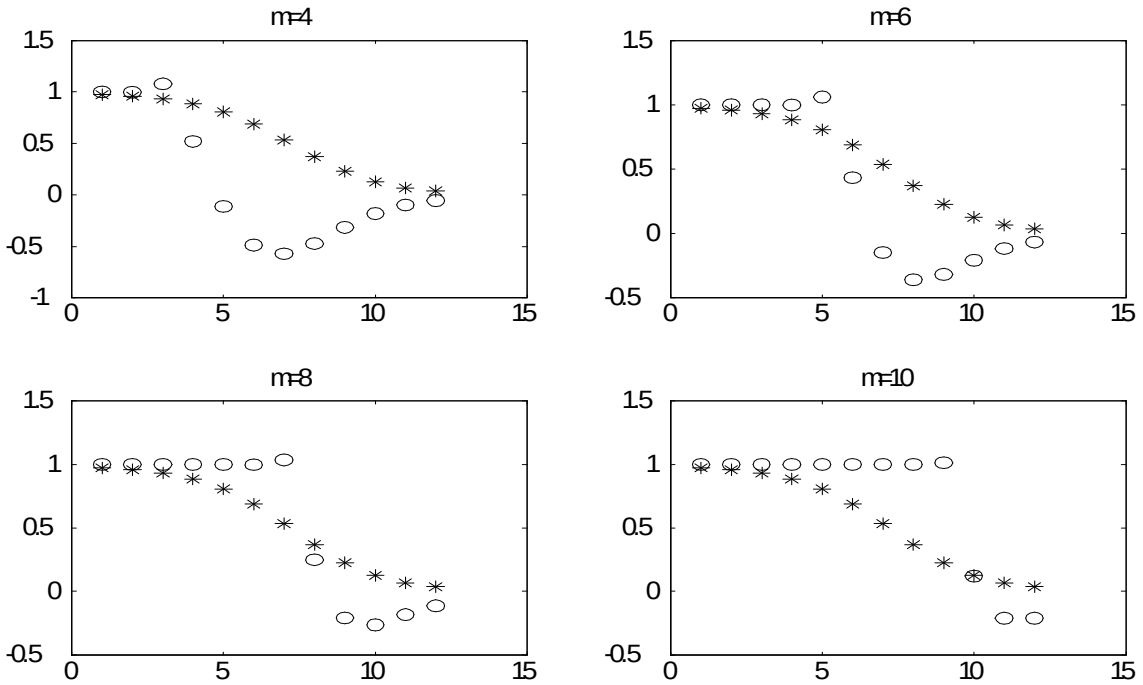
$$x_m = \sum_{i=1}^{N} \frac{\lambda_i p_{m-1}(\lambda_i)}{\lambda_i + \lambda} \frac{\left(X^{-1} b\right)_i}{\lambda_i} x_i.$$

and the filter factors are given by

$$f_i^{(m)} = \frac{\lambda_i p_{m-1}(\lambda_i)}{\lambda_i + \lambda}, \quad i = 1, ..., N.$$

Filter factors - an example

GRAVITY(12) - Filter factors g_i (asterisk) and $f_i^{(m)}$ (circle) with $m = 4, 6, 8, 10$. $\lambda = 1/\sqrt{\kappa(A)}$



The Arnoldi (Lanczos) algorithm initially picks up the largest eigenvalues, hence it automatically corrects the filters corresponding to the low-middle frequencies ($g_i \longrightarrow f_i^{(m)} \approx 1$), keep damping the highest ones.

Filter factors - theoretical analysis

Assume that the Ritz values r_j, $j = 1, \ldots, m$, are distinct. Therefore

$$p_{m-1}(\lambda_i) = \sum_{j=1}^{m} l_j(\lambda_i) f(r_j),$$

where l_j, $j = 1, \ldots, m$ are the Lagrange polynomials. Hence we obtain

$$f_i^{(m)} = \sum_{j=1}^{m} l_j(\lambda_i) \frac{\lambda_i}{r_j} \frac{r_j + \lambda}{\lambda_i + \lambda}, \quad i = 1, \ldots, N.$$

Since the Arnoldi algorithm ensures that $r_j \approx \lambda_j$ for $j = 1, \ldots, m$ we have $f_i^{(m)} \approx 1$ for $i \leq m$. For $i > m$ and when $\lambda_i \approx 0$ we have that

$$f_i^{(m)} \approx p_{m-1}(0) \frac{\lambda_i}{\lambda_i + \lambda},$$

so that the filters are close to the ones of the uncorrected scheme. Therefore, the choice of λ only influences the high frequencies. For this reason, for the ASP method, this choice is more related to the conditioning of the subproblems.

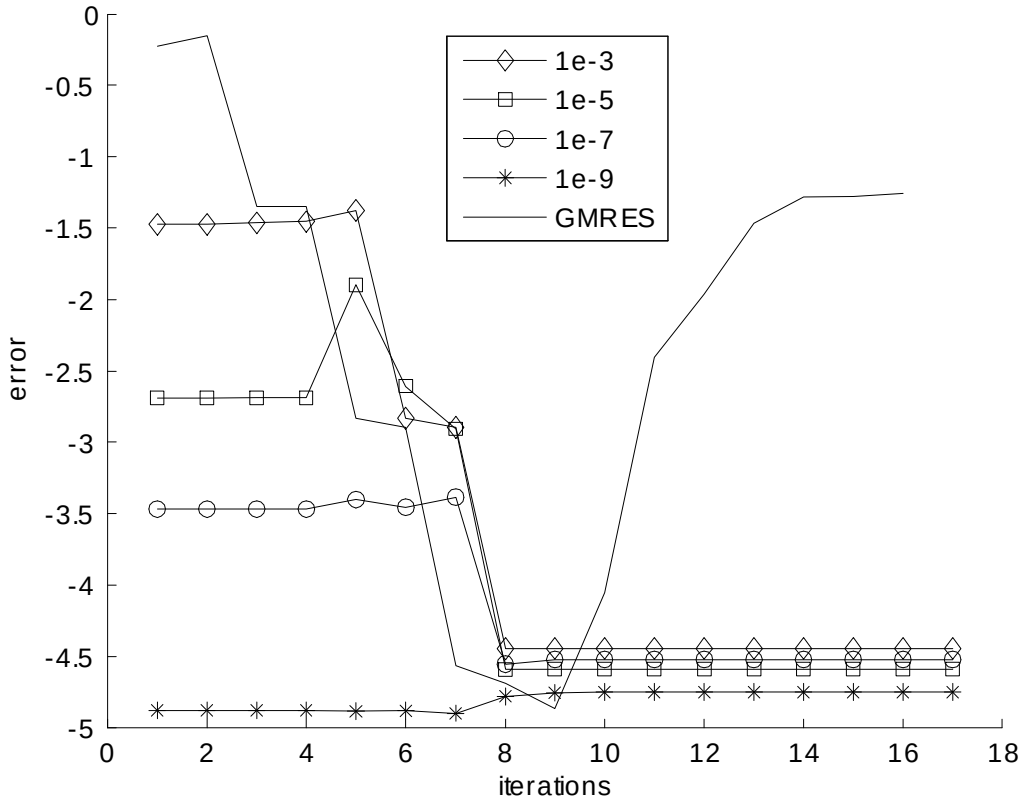
Filter factors - extension to the ATP method

The filter factor analysis remains valid also for the ATP method. Taking $H = I$ and using the SVD decomposition we easily find that the filter factors are now given by

$$f_i^{(m)} = \frac{\sigma_i^2 \, p_{m-1}(\sigma_i^2)}{\sigma_i^2 + \lambda}$$

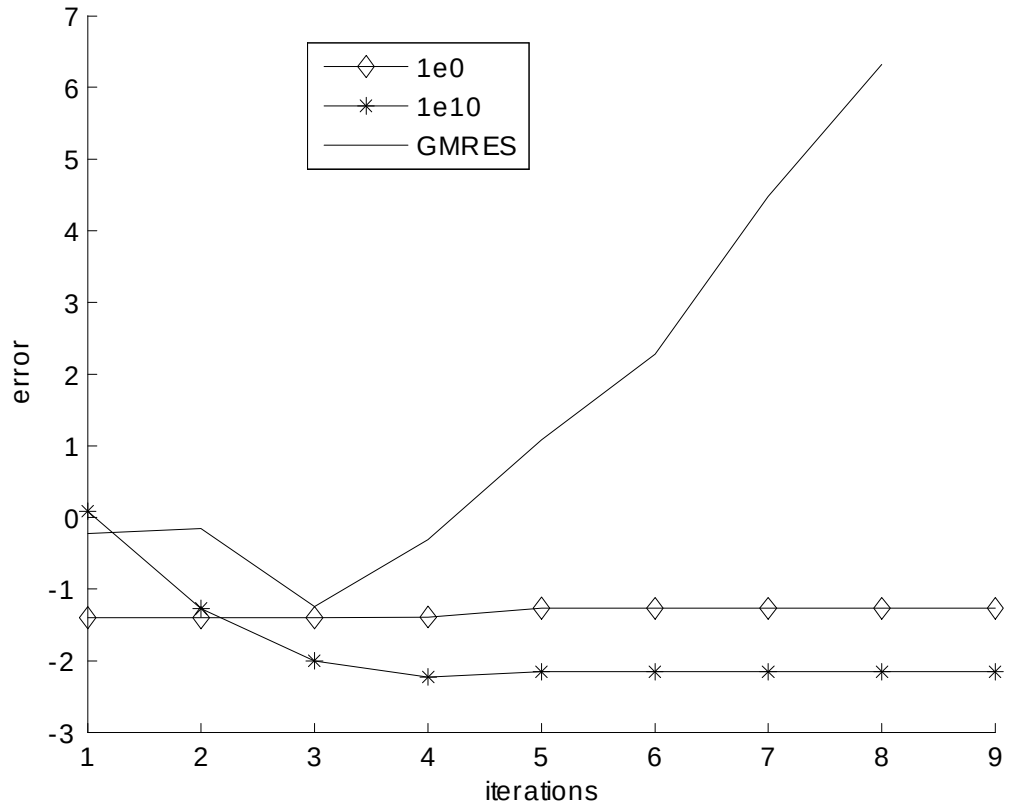
and hence our considerations for the ASP method remains true also for this case. For $H \neq I$ we just need to consider the GSVD. For problems with noise, the choice of λ is of great importance. Anyway the correction phase allows to reproduce the low frequencies independently of this choice. In this sense, in practice we can take λ even very large in order to reduce as much as possible the influence of noise.

Some experiments from Reg. Tools by P.C.Hansen



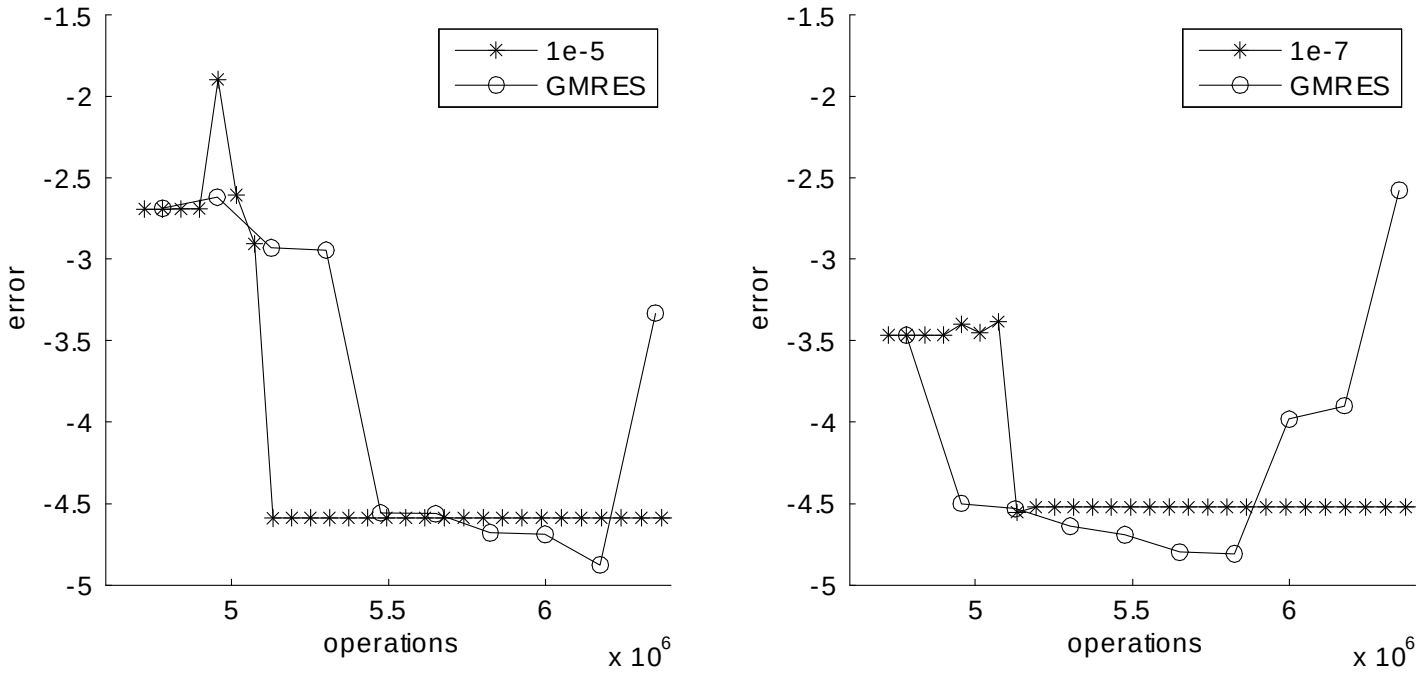
Error behavior of the GMRES and the ASP method with $\lambda = 1e-3$, $1e-5$, $1e-7$, $1e-9$, for BAART(240)

Some experiments from Reg. Tools by P.C.Hansen



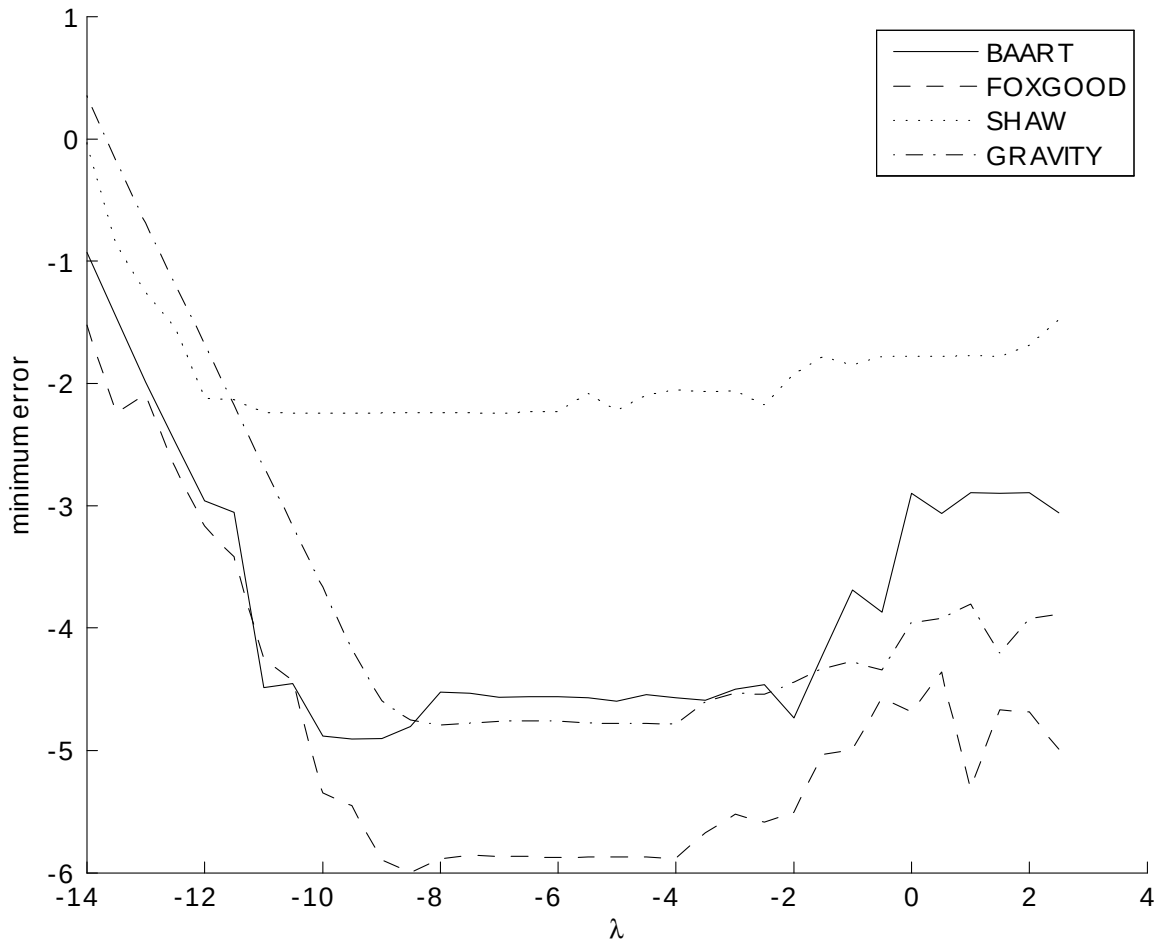
Error behavior of the GMRES and the ATP method with $\lambda = 1$ and $\lambda = 1e10$ for BAART(240) with Gaussian noise (10^{-3})

Some experiments from Reg. Tools by P.C.Hansen



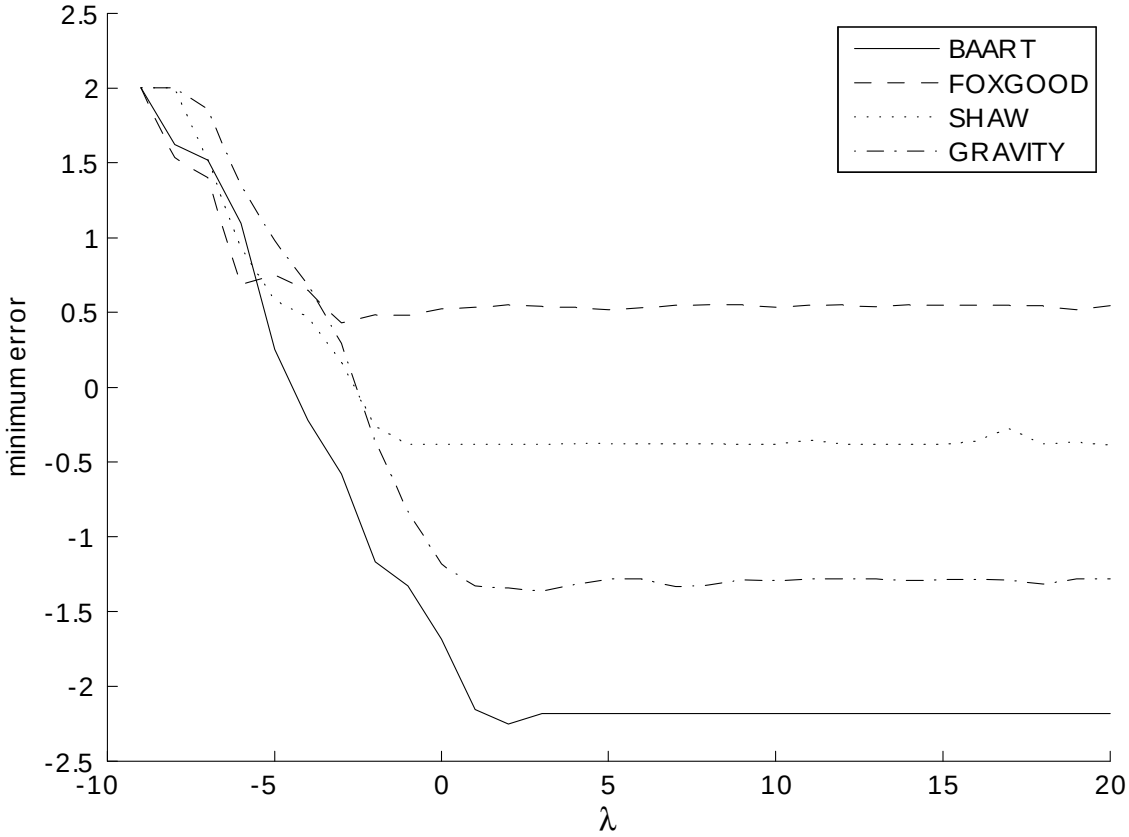
Error behavior of the preconditioned GMRES and the ASP method for BAART(240) with $\lambda = 1e-5$ and $1e-7$

Some experiments from Reg. Tools by P.C.Hansen



Maximum attainable accuracy with respect to the choice of lambda.
$N = 160$.

Some experiments from Reg. Tools by P.C.Hansen



Maximum attainable accuracy with respect to the choice of lambda with right-hand side affected by Gaussian noise (10^{-3}). $N = 160$.

An example of image restoration

We want to solve $Ax = b$ where $A \in \mathbb{R}^{N\times N}$ is the matrix representing the blurring operator and $b = Ax$ is the blurred image. We generate a noisy image $\overline{b} = b + e_b$, where e_b is a Gaussian noise (10^{-3}). As filters we consider

$$H_{1,2} = \begin{pmatrix} I \otimes H_1 \\ H_1 \otimes I \end{pmatrix}, \text{ where } H_1 = \begin{pmatrix} 1 & -1 & & \\ & \ddots & \ddots & \\ & & 1 & -1 \\ & & & 1 \end{pmatrix} \in \mathbb{R}^{n\times n},$$

$$H_{2,2} = \begin{pmatrix} 4 & -1 & & -1 & \\ -1 & 4 & -1 & & -1 \\ & \ddots & \ddots & \ddots & \\ -1 & & -1 & 4 & -1 \\ & -1 & & -1 & 4 \end{pmatrix} \in \mathbb{R}^{N\times N}.$$

An example of image restoration

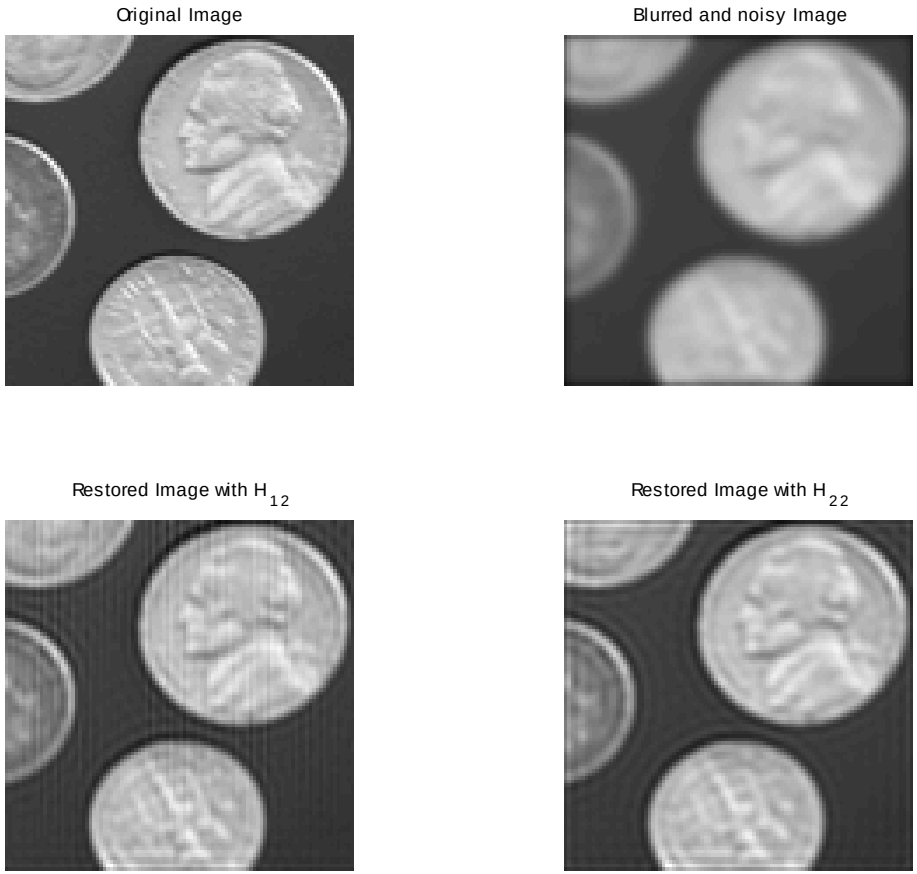


Image restoration with the ATP method using $H_{1,2}$, $H_{2,2}$ and $\lambda = 100$.

Results of image restoration

	1	10^2	10^4	10^6
$H_{1,2}$	0.060	0.060	0.062	0.059
$H_{2,2}$	0.061	0.064	0.069	0.075

Minimum attainable error with respect to the choice of λ

Final remarks

- Both methods are stable w.r.t. the choice of the number of iterations, i.e. they do not require a reliable stopping rule
- They are as fast and accurate as the most effective iterative solvers
- W.r.t. classical preconditioned iterative solvers, only one linear system with the preconditioner has to be solved
- They generally do not require an accurate parameter-choice strategy for λ