Stabilization of the Gradient Algorithm for Linear Systems

Ibrahima Dione

Independent Researcher

- Gradient Algorithm
- Stabilization of the Gradient Algorithm
- Convergence of the Stabilized Algorithm
- Numerical Examples

Gradient Algorithm

- The gradient method is one of the simplest iterative methods.

- Determine an approximate solution of the linear system of equations

$$\mathbf{A}\boldsymbol{x} = \mathbf{b}, \tag{1}$$

  - $\mathbf{A} \in \mathcal{M}_{n,n}(\mathbb{R})$ is ill-conditioned, non-singular and symmetric matrix,
  - $\mathbf{b}$ is a vector of $\mathbb{R}^n$,
  - $\|\cdot\|_n$ denotes the Euclidean vector norm in $\mathbb{R}^n$.
  - $\langle\cdot,\cdot\rangle_n$ is a scalar product associated with this norm, and
  - $\|\cdot\|$ is a matrix norm induced by the Euclidean vector norm.

- The gradient method is defined by the following iterative process

$$\boldsymbol{x}^{[k+1]} = \boldsymbol{x}^{[k]} - \alpha_k(\mathbf{A}\boldsymbol{x}^{[k]} - \mathbf{b}), \quad k = 0, 1, 2, \cdots, \tag{2}$$

where the step parameter α_k must be chosen.

- The iterative method (2) can be written under the iterative scheme

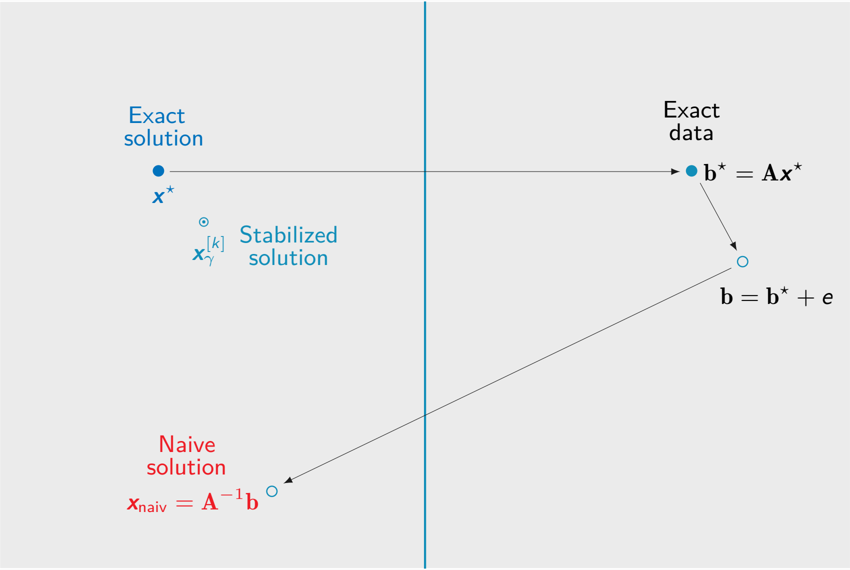
$$\mathbf{M}_{\alpha_k}\boldsymbol{x}^{[k+1]} = \mathbf{N}_{\alpha_k}\boldsymbol{x}^{[k]} + \mathbf{b}. \tag{3}$$

  - where the regular decomposition $\mathbf{A} = \mathbf{M}_\alpha - \mathbf{N}_\alpha$,
  - with the matrices $\mathbf{M}_\alpha = \frac{1}{\alpha}\mathbf{I}$ and $\mathbf{N}_\alpha = \left(\frac{1}{\alpha}\mathbf{I} - \mathbf{A}\right)$.

- This method is based on the idea of choosing the opposite direction of the highest rate of change of the quadratic functional

$$\mathcal{J}(\boldsymbol{x}) = \frac{1}{2}\left\langle \mathbf{A}\boldsymbol{x}, \boldsymbol{x}\right\rangle_n - \left\langle \mathbf{b}, \boldsymbol{x}\right\rangle_n, \tag{4}$$

- namely the direction opposite to the gradient $\nabla\mathcal{J}(\boldsymbol{x}) = \mathbf{A}\boldsymbol{x} - \mathbf{b}$.


Exact solution
$\boldsymbol{x}^\star$
$\boldsymbol{x}_\gamma^{[k]}$
Stabilized solution
Exact data
$\mathbf{b}^\star = \mathbf{A}\boldsymbol{x}^\star$
$\mathbf{b} = \mathbf{b}^\star + e$
Naive solution
$\boldsymbol{x}_{\mathrm{naiv}} = \mathbf{A}^{-1}\mathbf{b}$

Stabilization of the Gradient Algorithm

- Gradient algorithm in (2) can be formulated as the minimization problem:
  - Given $\boldsymbol{x}^{[k]}$,
  - Compute $\boldsymbol{x}^{[k+1]} \in \mathbb{R}^n$ from the minimization problem:

$$\mathcal{F}\left(\boldsymbol{x}^{[k+1]}\right) := \min_{\boldsymbol{x}\in\mathbb{R}^n} \mathcal{F}(\boldsymbol{x}), \tag{5}$$

where the functional $\mathcal{F}(\cdot)$ is defined in the following form

$$\mathcal{F}(\boldsymbol{x}) = \frac{1}{2}\left\|\boldsymbol{x} - \boldsymbol{x}^{[k]}\right\|_n^2 + \alpha_k \left\langle \mathbf{A}\boldsymbol{x}^{[k]} - \mathbf{b},\, \boldsymbol{x} - \boldsymbol{x}^{[k]} \right\rangle_n . \tag{6}$$

- Our approach is introduced from (5) with a focus on **stability**.
- To do so, we minimize the functional (6) under the constraint subset

$$\boldsymbol{K} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \mid \min_{\boldsymbol{x}\in\mathbb{R}^n} \mathcal{J}(\boldsymbol{x}) \right\}. \tag{7}$$

▷ Instead of the problem (5), we therefore consider the following problem:

⋆ Given $\boldsymbol{x}^{[k]}$,

⋆ Compute $\boldsymbol{x}^{[k+1]} \in \boldsymbol{K}$ from

$$\mathcal{F}\left(\boldsymbol{x}^{[k+1]}\right) := \min_{\boldsymbol{x}\in\boldsymbol{K}} \mathcal{F}(\boldsymbol{x}). \tag{8}$$

▷ We relax the constraint in $\boldsymbol{K}$ by regularizing the problem in (8).

▷ We thus introduce the following regularized problem:

⋆ Given $\boldsymbol{x}_\gamma^{[k]}$,

⋆ Compute $\boldsymbol{x}_\gamma^{[k+1]} \in \mathbb{R}^n$ from

$$\mathcal{F}_\gamma(\boldsymbol{x}_\gamma^{[k+1]}) := \min_{\boldsymbol{x}\in\mathbb{R}^n} \mathcal{F}_\gamma(\boldsymbol{x}),\ \forall\,\gamma > 0, \tag{9}$$

where the regularized (or **stabilized**) functional $\mathcal{F}_\gamma(\cdot)$ is defined by

$$\mathcal{F}_\gamma(\boldsymbol{x}) := \mathcal{F}(\boldsymbol{x}) + \gamma\mathcal{J}(\boldsymbol{x}),\ \forall\,\gamma > 0. \tag{10}$$

▷ Taking the necessary first-order optimality condition, we obtain:

▷ Given $\boldsymbol{x}_\gamma^{[k]}$, calculate $\boldsymbol{x}_\gamma^{[k+1]} \in \mathbb{R}^n$ from the new iterative scheme:

$$(\mathbf{I} + \gamma \mathbf{A})\, \boldsymbol{x}_\gamma^{[k+1]} = (\mathbf{I} - \alpha_k \mathbf{A})\, \boldsymbol{x}_\gamma^{[k]} + (\alpha_k + \gamma)\, \mathbf{b}, \tag{11}$$

for any positive **stabilization parameter** $\gamma > 0$, and $k = 0, 1, 2, \cdots$.

▷ The scheme (11) may be seen as a preconditioning method in the sense:

$$\mathbf{M}_\gamma \boldsymbol{x}_\gamma^{[k+1]} = \mathbf{N}_k \boldsymbol{x}_\gamma^{[k]} + \boldsymbol{c}_{(\gamma,k)}, \; \forall\, \gamma > 0, \; \forall\, k = 0, 1, 2, \cdots , \tag{12}$$

- ⋆ where $\mathbf{N}_k = (\mathbf{I} - \alpha_k \mathbf{A})$, $\boldsymbol{c}_{(\gamma,k)} = (\alpha_k + \gamma)\, \mathbf{b}$, and
- ⋆ $\mathbf{M}_\gamma = (\mathbf{I} + \gamma \mathbf{A})$ is a non-singular matrix.

Convergence of the Stabilized Algorithm

Lemma

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a symmetric and non-singular matrix with the decomposition $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$, where $\mathbf{V} \in \mathbb{R}^{n\times n}$ is an orthogonal matrix and $\mathbf{D} \in \mathbb{R}^{n\times n}$ is a diagonal matrix $\mathbf{D} = \mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)$ with entries the eigenvalues of $\mathbf{A}$. We obtain the estimate

$$\left\|(\mathbf{I} + \gamma\mathbf{A})^{-1}\right\| \leq \max_{i=1,\cdots,n} \left(\frac{\kappa(\mathbf{V})}{|1+\gamma\lambda_i|}\right), \ \forall\, \gamma > 0, \tag{13}$$

where $\kappa(\mathbf{V}) = \|\mathbf{V}\|\|\mathbf{V}^{-1}\|$ is the conditioning of the orthogonal matrix $\mathbf{V}$.

Proof:

▷ Decomposition $\mathbf{A} = \mathbf{V}\mathbf{D}\mathbf{V}^{-1}$ and submultiplicity of the matrix norm give

$$\|(\mathbf{I}+\gamma\mathbf{A})^{-1}\| = \|\mathbf{V}(\mathbf{I}+\gamma\mathbf{D})^{-1}\mathbf{V}^{-1}\| \leq \|\mathbf{V}\|\|\mathbf{V}^{-1}\|\|(\mathbf{I}+\gamma\mathbf{D})^{-1}\|.$$

▷ The matrix norm being induced from the Euclidean vector norm, we obtain

$$\left\|(\mathbf{I} + \gamma\mathbf{A})^{-1}\right\| \leq \kappa(\mathbf{V}) \max_{i=1,\cdots,n} \left(\frac{1}{|1+\gamma\lambda_i|}\right), \ \forall\, \gamma > 0,$$

where $\kappa(\mathbf{V}) = \|\mathbf{V}\|\|\mathbf{V}^{-1}\|$ is the conditioningconditioning of $\mathbf{V}$. □

Theorem

Assume that $\mathbf{A} \in \mathbb{R}^{n\times n}$ is a non-singular $n \times n$ matrix. For any $\alpha_k \in \mathbb{R}$ and $\gamma > 0$, the sequence $\left\{\boldsymbol{x}_\gamma^{[k]}\right\}_{k\in\mathbb{N}^\star}$ from the stabilized gradient method (11) verifies the error estimate

$$\|\boldsymbol{x}_\gamma^{[k+1]} - \boldsymbol{x}^\star\|_n \leq \left\|(\mathbf{I}+\gamma\mathbf{A})^{-1}\right\| \, \|\mathbf{I}-\alpha_k\mathbf{A}\| \, \|\boldsymbol{x}_\gamma^{[k]} - \boldsymbol{x}^\star\|_n, \; \forall\, \gamma > 0, \; \forall\, k \in \mathbb{N}^\star. \quad (14)$$

Proof:

- ▷ The gradient method (2) is consistent with the linear system (1)

$$\boldsymbol{x}^\star = \boldsymbol{x}^\star - \alpha_k(\mathbf{A}\boldsymbol{x}^\star - \mathbf{b}), \; k = 0, 1, 2, \cdots . \quad (15)$$

- ▷ The stabilized gradient method (11) is also consistent with the system (1)

$$(\mathbf{I}+\gamma\mathbf{A})\,\boldsymbol{x}^\star = (\mathbf{I}-\alpha_k\mathbf{A})\,\boldsymbol{x}^\star + \alpha_k\mathbf{b} + \gamma\mathbf{b}, \; \forall\,\gamma > 0. \quad (16)$$

- ▷ Subtracting (11) and (16), we obtain the equation for $k = 0, 1, 2, 3, \cdots$

$$(\mathbf{I}+\gamma\mathbf{A})\left(\boldsymbol{x}_\gamma^{[k+1]} - \boldsymbol{x}^\star\right) = (\mathbf{I}-\alpha_k\mathbf{A})\left(\boldsymbol{x}_\gamma^{[k]} - \boldsymbol{x}^\star\right), \; \forall\,\gamma > 0. \quad (17)$$

- ▷ We establish (14) using the compatibility of norms and submultiplicity. □

Corollary

Under the assumption of previous lemma and theorem, and for any initial datum $\boldsymbol{x}_\gamma^{[0]}$, the sequence $\left\{\boldsymbol{x}_\gamma^{[k]}\right\}_{k\in\mathbb{N}^\star}$ from the stabilized method (11) verifies:

- If the stepsize α_k is a non-constant parameter, we obtain for $k \in \mathbb{N}^\star$
$$\|\boldsymbol{x}_\gamma^{[k]} - \boldsymbol{x}^\star\|_n \leq \left(\max_{i=1,\cdots,n} \frac{\kappa(\mathbf{V})}{|1+\gamma\lambda_i|}\right)^k \left(\prod_{i=0}^{k-1} \|\mathbf{I} - \alpha_i \mathbf{A}\|\right) \|\boldsymbol{x}_\gamma^{[0]} - \boldsymbol{x}^\star\|_n, \ \forall\, \gamma > 0. \tag{18}$$

- If the stepsize is constant $\alpha = \alpha_k$, $\forall\, k \in \mathbb{N}^\star$, then we obtain
$$\|\boldsymbol{x}_\gamma^{[k]} - \boldsymbol{x}^\star\|_n \leq \left(\max_{i=1,\cdots,n} \frac{\kappa(\mathbf{V})}{|1+\gamma\lambda_i|} \|\mathbf{I} - \alpha \mathbf{A}\|\right)^k \|\boldsymbol{x}_\gamma^{[0]} - \boldsymbol{x}^\star\|_n, \ \forall\, \gamma > 0. \tag{19}$$

Proof:

▷ By multiplying term by term inequalities from (14) for $k = 0, \cdots, k-1$, we establish (18) after simplification.

▷ If the size step is constant, we obtain from (18) the estimate (19). □

Note:

- The convergence speed of the stabilized gradient method (11)
  - is faster as the γ parameter increases;
  - γ has complete control over the convergence of the solution $\boldsymbol{x}_\gamma^{[k]}$, independently of α_k and the sign of the eigenvalues λ_i.
- If $\mathbf{A}$ is positive definite and therefore has real eigenvalues, $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$, then we obtain convergence from (19) if and only if

$$\left(\frac{\kappa\left(\mathbf{V}\right)}{1+\gamma\lambda_n}\right)\|\mathbf{I}-\alpha\mathbf{A}\| < 1$$

$$1-\frac{1+\gamma\lambda_n}{\kappa\left(\mathbf{V}\right)} < \alpha\lambda_i < 1+\frac{1+\gamma\lambda_n}{\kappa\left(\mathbf{V}\right)},\ \forall\,\gamma>0,\ \forall\, i=1,\cdots,n. \tag{20}$$

- From (20), we see that the step α and the eigenvalues λ_i can be of different signs as long as [2]

$$1-\frac{1+\gamma\lambda_n}{\kappa\left(\mathbf{V}\right)} < 0 \iff \gamma > \frac{\kappa\left(\mathbf{V}\right)-1}{\lambda_n} > 0. \tag{21}$$

Numerical Examples

- When gradient method is applied to the standard problem [1]

$$\min_{(x,y)\in\mathbb{R}^2} \left(x^2 + ay^2\right), \tag{22}$$

- This standard problem can be formulated under the quadratic problem

$$\min_{\boldsymbol{x}\in\mathbb{R}^2} \langle \mathbb{A}_a \boldsymbol{x},\, \boldsymbol{x}\rangle_2 + 2\,\langle \mathbf{b},\, \boldsymbol{x}\rangle_2\,, \tag{23}$$

  where $\boldsymbol{x} = (x, y)^t$, $\mathbb{A}_a \in \mathbb{R}^{2\times 2}$ is symmetric and positive definite, and $\mathbf{b} \in \mathbb{R}^2$ is such that:

$$\mathbb{A}_a = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}, \text{ et } \mathbf{b} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{24}$$

- $\boldsymbol{x}^\star = (0,0)^t$ is the unknown exact solution sought.

- An overview on the convergence of the gradient method applied to (23):

  - illustrates the zigzag phenomenon observed with the gradient;

  - therefore the slow convergence with this algorithm [1].

▷ Constant stepsize method: we fix a priori $\alpha_k = \alpha, \ \forall \ k$.

▷ Exact line search method: α_k is a minimizer of $\mathcal{J}$ along $\boldsymbol{x}^{[k]} + \alpha \boldsymbol{d}^{[k]}$

$$\alpha_k = \underset{\alpha \geq 0}{\operatorname{argmin}} \mathcal{J}(\boldsymbol{x}^{[k]} + \alpha \boldsymbol{d}^{[k]}). \tag{25}$$

▷ Backtracking line search method: Requires three parameters: $s > 0$, $\lambda \in (0,1)$, $\beta \in (0,1)$.

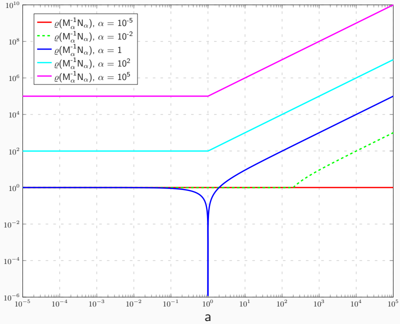
- ⋆ First, α_k is set to be equal to the initial guess s: $\alpha_k \leftarrow s$.
- ⋆ While $\mathcal{J}(\boldsymbol{x}^{[k]}) - \mathcal{J}(\boldsymbol{x}^{[k]} + \alpha_k \boldsymbol{d}^{[k]}) < -\lambda \alpha_k \left(\nabla \mathcal{J}(\boldsymbol{x}^{[k]})\right)^t \boldsymbol{d}^{[k]}$, we set:

$$\alpha_k \leftarrow \beta \alpha_k.$$

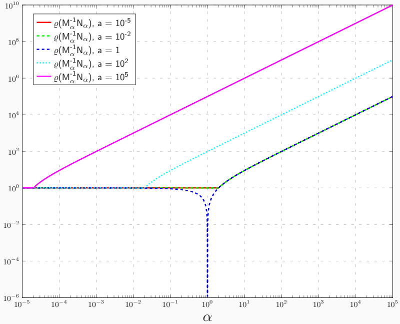
- ⋆ At the end, the stepsize is chosen as fallows:

$$\alpha_k = s\beta^{i_k}, \ i_k \in \mathbb{N}^\star.$$

1 | Spectral radius $\varrho(\mathrm{M}_\alpha^{-1}\mathrm{N}_\alpha)$ of gradient method (3) applied to (23) for $a \in [10^{-5}, 10^5]$

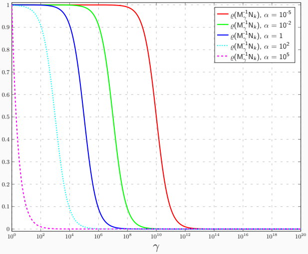


2 | Spectral radius $\varrho(\mathrm{M}_\alpha^{-1}\mathrm{N}_\alpha)$ of gradient method (3) applied to (23) for $\alpha \in [10^{-5}, 10^5]$

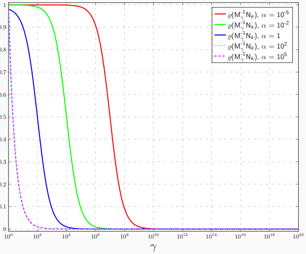


The spectral radius $\varrho(\mathbf{M}_\alpha^{-1}\mathbf{N}_\alpha)$ of the iteration matrix $\mathbf{M}_\alpha^{-1}\mathbf{N}_\alpha$ from the gradient method (3) applied to the quadratic problem (23) and represented in 1 and 2, with respect to the parameter $a \in [10^{-5}, 10^5]$ (for each $\alpha = 10^i$, $i = -5, -2, 0, 2, 5$) and the stepsize $\alpha \in [10^{-5}, 10^5]$ (for each $a = 10^i$, $i = -5, -2, 0, 2, 5$), respectively.

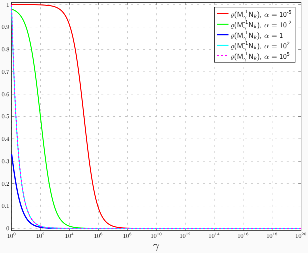
1 | Spectral radius $\varrho(\mathrm{M}_\gamma^{-1}\mathrm{N}_k)$ of scheme (11) with matrix $\mathbb{A}_a$ in (24) where $a = 10^{-5}$



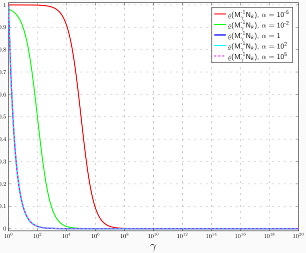
2 | Spectral radius $\varrho(\mathrm{M}_\gamma^{-1}\mathrm{N}_k)$ of scheme (11) with matrix $\mathbb{A}_a$ in (24) where $a = 10^{-2}$



3 | Spectral radius $\varrho(\mathrm{M}_\gamma^{-1}\mathrm{N}_k)$ of scheme (11) with matrix $\mathbb{A}_a$ in (24) where $a = 2$



4 | Spectral radius $\varrho(\mathrm{M}_\gamma^{-1}\mathrm{N}_k)$ of scheme (11) with matrix $\mathbb{A}_a$ in (24) where $a = 10^{5}$



The spectral radius $\varrho(\mathbf{M}_\gamma^{-1}\mathbf{N}_k)$ of the iteration matrix $\mathbf{M}_\gamma^{-1}\mathbf{N}_k$ from the stabilized gradient method (11) applied to the problem (23) and represented in 1, 2, 3 and 4, with respect to the parameter γ (for $\alpha = 10^i$, $i = -5, -2, 0, 2, 5$) and fixed parameter $a = 10^{-5}, 10^{-2}, 2, 10^5$.

Stab. Parameter (γ)	$\|\mathbb{A}_2 \boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_2$	$\frac{\|\mathbb{A}_2 \boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_2}{\|\mathbb{A}_2 \boldsymbol{x}_\gamma^{[0]} + \mathbf{b}\|_2}$	$\|\boldsymbol{x}^\star - \boldsymbol{x}_\gamma^{[k]}\|_2$	# Iters $[k]$	$\boldsymbol{x}_\gamma^{[k]}$
1	$5.556241e^{-06}$	$9.822139e^{-07}$	$2.778121e^{-06}$	10	$1e^{-5}\,(0.277, 0.000)$
10	$1.097558e^{-06}$	$1.940227e^{-07}$	$5.487792e^{-07}$	5	$1e^{-6}\,(0.548, 0.000)$
10^2	$6.834489e^{-07}$	$1.208178e^{-07}$	$3.417245e^{-07}$	3	$1e^{-6}\,(0.341, 0.000)$
10^5	$1.407502e^{-10}$	$2.488135e^{-11}$	$7.037052e^{-11}$	2	$1e^{-10}\,(0.703, 0.004)$
10^7	$2.748737e^{-07}$	$4.859126e^{-08}$	$1.343709e^{-07}$	1	$1e^{-6}\,(0.133, 0.016)$
10^{10}	$2.748737e^{-10}$	$4.859127e^{-11}$	$1.343710e^{-10}$	1	$1e^{-9}\,(0.133, 0.016)$

Table 1: Numerical results obtained from the stabilized gradient method (11) with the data $\mathbb{A}_a$, $\mathbf{b}$ given in (24), where $a = 2$ ($\textbf{cond}(\mathbb{A}_2) = 2$, $\textbf{rank}(\mathbb{A}_2) = 2$). Computations are made with respect to the different values of the stabilization parameter $\gamma = 10^i$, $i = 0,\ 1,\ 2,\ 5,\ 7,\ 10$, the initial iteration fixed to $\boldsymbol{x}_\gamma^{[0]} = (2, 1)^t$, and the stepsize computed from (25). Stopping criterion $\|\mathbb{A}_a \boldsymbol{x}_\gamma^{[k]} + \mathbf{b}\|_2 \leq \varepsilon$ is used where $\varepsilon = 10^{-5}$.

Stab. Parameter (γ)	$\|\mathbb{A}_2 \boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_2$	$\frac{\|\mathbb{A}_2 \boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_2}{\|\mathbb{A}_2 \boldsymbol{x}_\gamma^{[0]} + \mathbf{b}\|_2}$	$\|\boldsymbol{x}^\star - \boldsymbol{x}_\gamma^{[k]}\|_2$	# Iters $[k]$	$\boldsymbol{x}_\gamma^{[k]}$
1	$5.089472e^{-06}$	$8.997000e^{-07}$	$2.544736e^{-06}$	17	$1e^{-5}\,(0.254, 0.000)$
10	$1.200001e^{-06}$	$2.121321e^{-07}$	$5.999770e^{-07}$	6	$1e^{-6}\,(0.600, 0.003)$
10^2	$2.841455e^{-06}$	$5.023030e^{-07}$	$1.416524e^{-06}$	3	$1e^{-5}\,(0.141, 0.006)$
10^5	$3.302540e^{-10}$	$5.838121e^{-11}$	$1.627850e^{-10}$	2	$1e^{-9}\,(0.162, 0.016)$
10^7	$3.939543e^{-07}$	$6.964193e^{-08}$	$1.843909e^{-07}$	1	$1e^{-6}\,(0.180, 0.040)$
10^{10}	$3.939543e^{-10}$	$6.964194e^{-11}$	$1.843909e^{-10}$	1	$1e^{-9}\,(0.180, 0.040)$

Table 2: Numerical results obtained from the stabilized gradient method (11) with the data $\mathbb{A}_a$, $\mathbf{b}$ given in (24), where $a = 2$ ($\mathbf{cond}(\mathbb{A}_2) = 2$, $\mathbf{rank}(\mathbb{A}_2) = 2$). Computations are made with respect to the different values of the stabilization parameter $\gamma = 10^i$, $i = 0, 1, 2, 5, 7, 10$, the initial iteration $\boldsymbol{x}_\gamma^{[0]} = (2, 1)^t$ and the constant stepsize $\alpha = 0.1$. The stopping criterion $\|\mathbb{A}_a \boldsymbol{x}_\gamma^{[k]} + \mathbf{b}\|_2 \leq \varepsilon$ is used where $\varepsilon = 10^{-5}$.

- We study ill-conditioned linear systems on a small scale.

- The Shaw inverse problem, defined by the integral equation [4]

$$\int_a^b K(s,t)x^\star(t)dt = f(s), \quad c \leq s \leq d, \tag{26}$$

  - where the kernel K and the right-hand side f are given,
  - and $x^\star$ is the unknown solution sought.

- The midpoint quadrature rule applied to approximate (26), provides:

$$\sum_{j=1}^{n} w_j K(s_i, t_j)x^\star(t_j) = f(s_i), \quad i = 1, \cdots, n. \tag{27}$$

- The matrix $\mathbf{A} = (a_{ij})_{1\leq i,j\leq n}$, and the vector $\mathbf{b} = (b_i)_{1\leq i\leq n}$ are given by

$$a_{ij} = w_j K(s_i, t_j), \quad b_i = f(s_i), \quad i = 1, \cdots, n, \quad j = 1, \cdots, n. \tag{28}$$

- The exact solution considered here is defined by the analytical expression

$$\boldsymbol{x}^\star(t) = 2\exp\left(-6\,(t-0.8)^2\right) + \exp\left(-2\,(t+0.5)^2\right), \quad -\frac{\pi}{2} \le t \le \frac{\pi}{2}, \tag{29}$$

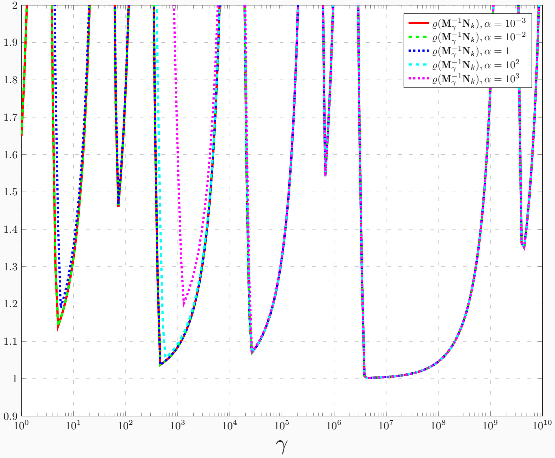
- The exact kernel K is defined analytically by the following expression [4]

$$K(s,t) = (\cos(s)+\cos(t))^2 \left(\frac{\sin(\psi(s,t))}{\psi(s,t)}\right)^2, \quad -\frac{\pi}{2} \le s \le \frac{\pi}{2}, \tag{30}$$

where the function $\psi(s,t) = \pi\,(\sin(s)+\cos(t))$.

- The Matlab code «shaw(n)» in [4] provides the data $\mathbf{A}$, $\mathbf{b}$, and $\boldsymbol{x}^\star$.

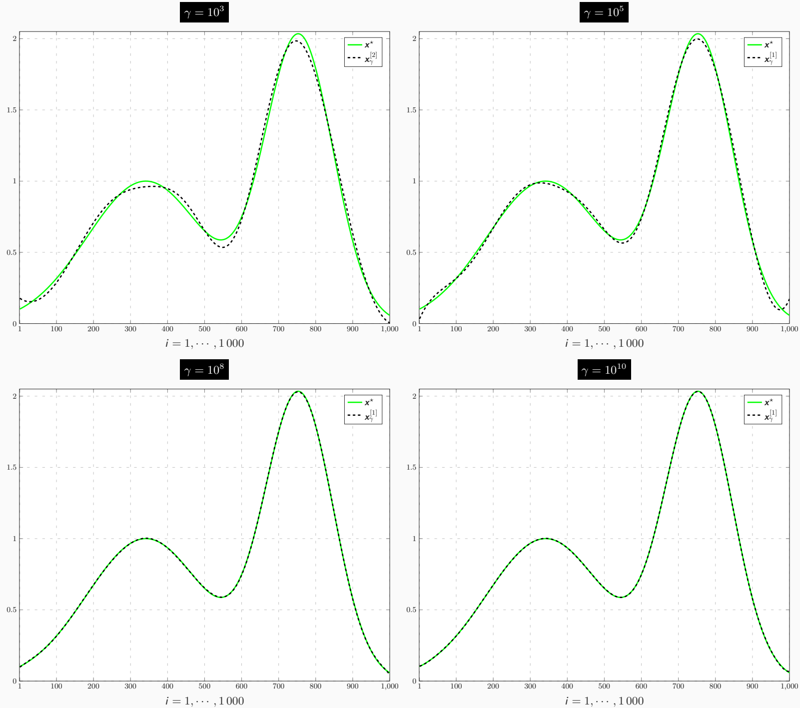
Spectral radius $\varrho(\mathrm{M}_{\gamma}^{-1}\mathrm{N}_k)$ of the scheme (11) from the «shaw(n)» test problem with $n = 1\,000$

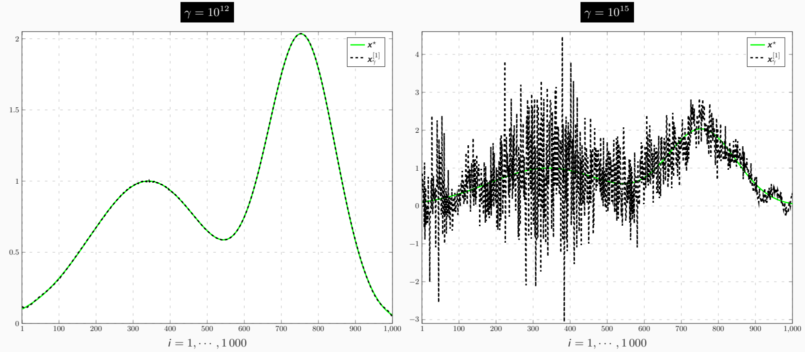


ℹ The spectral radius $\varrho(\mathbf{M}_{\gamma}^{-1}\mathbf{N}_k)$ of the stabilized gradient method (11) applied to the discret linear system (27) for Shaw test problem, and represented with respect to the stabilization parameter γ, for the different values of the stepsize $\alpha = 10^i$, $i = -3$, -2, 0, 2, 3, and under the dimension $n = 1\,000$, where $\textbf{cond}(\mathbf{A}) = 7,697e^{+20}$ and $\textbf{rank}(\mathbf{A}) = 20$.

Stab. Parameter (γ)	$\Vert \mathbf{A}\boldsymbol{x}_{\gamma}^{[k]} - \mathbf{b}\Vert_n$	$\frac{\Vert \mathbf{A}\boldsymbol{x}_{\gamma}^{[k]} - \mathbf{b}\Vert_n}{\Vert \mathbf{A}\boldsymbol{x}_{\gamma}^{[0]} - \mathbf{b}\Vert_n}$	$\Vert \boldsymbol{x}^{\star} - \boldsymbol{x}_{\gamma}^{[k]}\Vert_n$	$\frac{\Vert \boldsymbol{x}^{\star} - \boldsymbol{x}_{\gamma}^{[k]}\Vert_n}{\Vert \boldsymbol{x}^{\star}\Vert_n}$	# Iters $[k]$
10^3	$3.375476e^{-04}$	$4.578985e^{-06}$	$1.147962e^{+00}$	$3.636711e^{-02}$	2
10^5	$5.124204e^{-04}$	$6.951214e^{-06}$	$7.648419e^{-01}$	$2.422998e^{-02}$	1
10^6	$5.124875e^{-05}$	$6.952124e^{-07}$	$5.731437e^{-01}$	$1.815704e^{-02}$	1
10^7	$5.124387e^{-06}$	$6.951463e^{-08}$	$1.229460e^{-01}$	$3.894896e^{-03}$	1
10^8	$5.124346e^{-07}$	$6.951407e^{-09}$	$7.947206e^{-02}$	$2.517653e^{-03}$	1
10^{10}	$5.124286e^{-09}$	$6.951325e^{-11}$	$4.922276e^{-02}$	$1.559364e^{-03}$	1
10^{12}	$5.128325e^{-11}$	$6.956804e^{-13}$	$7.795781e^{-02}$	$2.469682e^{-03}$	1
10^{15}	$1.214243e^{-13}$	$1.647176e^{-15}$	$2.389361e^{+01}$	$7.569430e^{-01}$	1

Table 3: Stabilized method (11) applied to **shaw** problem for different values of the stabilization parameter γ, under the dimension $n = 1\,000$, the stepsize $\alpha = 1$, the maximum iterations $k_{\max} = n$ and the stopping criteria $\frac{\Vert \mathbf{b} - \mathbf{A}\boldsymbol{x}_{\gamma}^{[k]}\Vert_n}{\Vert \mathbf{b} - \mathbf{A}\boldsymbol{x}_{\gamma}^{[0]}\Vert_n} \leq \varepsilon$, $\varepsilon = 10^{-5}$. Data $\mathbf{A}$, $\mathbf{b}$ and $\boldsymbol{x}^{\star}$ computed from the code shaw(n) provided in [4], where $\mathbf{cond}(\mathbf{A}) = 7.697e^{+20}$ and $\mathbf{rank}(\mathbf{A}) = 20$.


$\gamma = 10^3$
$x^\star$
$x_\gamma^{[2]}$
$i = 1, \cdots, 1\,000$
$\gamma = 10^5$
$x^\star$
$x_\gamma^{[1]}$
$i = 1, \cdots, 1\,000$
$\gamma = 10^8$
$x^\star$
$x_\gamma^{[1]}$
$i = 1, \cdots, 1\,000$
$\gamma = 10^{10}$
$x^\star$
$x_\gamma^{[1]}$
$i = 1, \cdots, 1\,000$



The exact solution $\boldsymbol{x}^\star$ (in green solid line) computed from (29), and the stabilized solution $\boldsymbol{x}_\gamma^{[k]}$ (in black dashed line) obtained from (11), presented with respect to the index $i = 1, \cdots, 1\,000$ for different $\gamma = 10^i$, $i = 3, 5, 8, 10, 12, 15$. Data $\mathbf{A}$ and $\mathbf{b}$ are computed from the Matlab code shaw(n) in [4], the dimension $n = 1\,000$, the stepsize $\alpha = 1$, the maximum iterations $k_{\max} = n$ and the tolerance $\varepsilon = 10^{-5}$, with $\mathbf{cond}(\mathbf{A}) = 7.697e^{+20}$, $\mathbf{rank}(\mathbf{A}) = 20$.

- An image deblurring problem is used here to analyze the performance of the stabilized gradient method (11).

- It is modeled by the Fredholm integral equation of the first kind

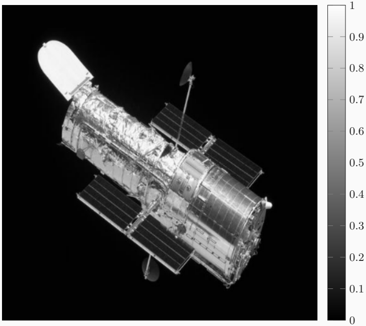
$$\int_a^b K(s,t)\boldsymbol{x}^\star(t)dt = \boldsymbol{f}(s), \;\; c \le s \le d, \tag{31}$$

  - where the kernel $K(s,t)$ specifying how the image is distorted at point $(s,t) \in \mathbb{R}^2$, defines the matrix $\mathbf{A}$.
  - The data $\boldsymbol{f}$ is the blurred image, and defines the right-hand side $\mathbf{b}$.
  - The source $\boldsymbol{x}^\star$ is the sharp image to be reconstructed.

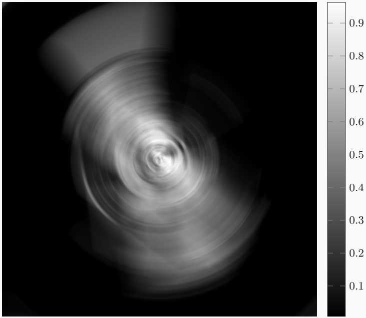
Note: The type of deblurring that interests us is the one where the blurred image is due to a rotational movement around the center of the image [3].

Inverse $2D$ image deblurring problem

1 | True image $x^\star$



2 | Blurred image b

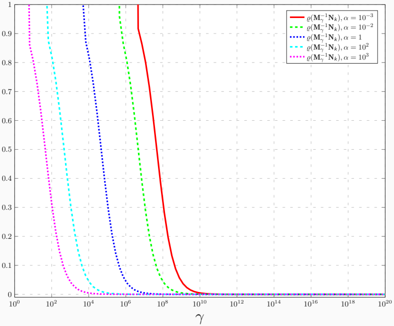


ℹ Real (sharp) image 1 and blurred image 2 obtained under $n = N^2$, $N = 400$, from the Matlab code «PRblurrotation(N)» provided in [3].

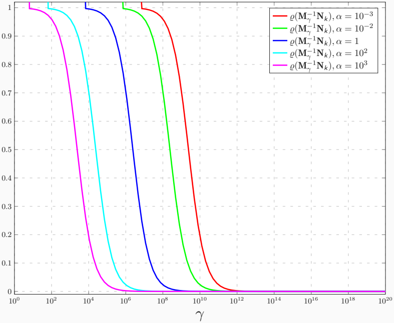
▷ The Matlab code «PRblurrotation(N)» returns:

- ⋆ The vectors $\boldsymbol{x}^\star$ and $\mathbf{b}$ representing the sharp and blurred images,
- ⋆ $\mathbf{A}$, an ill-conditioned matrix, models the blurring operation.

1 | Spectral radius $\varrho(\mathrm{M}_\gamma^{-1}\mathrm{N}_k)$ for the $2D$ image deblurring problem under $N = 100$



2 | Spectral radius $\varrho(\mathrm{M}_\gamma^{-1}\mathrm{N}_k)$ for the $2D$ image deblurring problem under $N = 150$



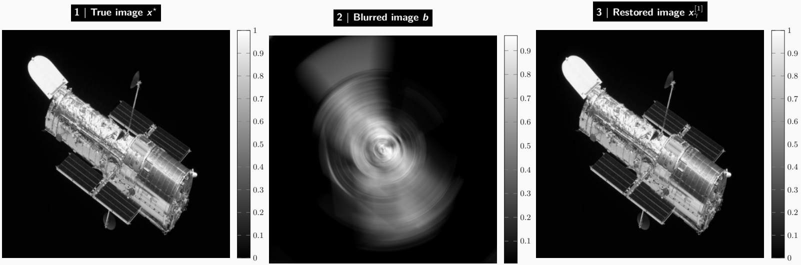
The spectral radius $\varrho(\mathbf{M}_\gamma^{-1}\mathbf{N}_k)$ of the stabilized gradient method (11) applied to the $2D$ image deblurring problem under the dimension $n = N^2$ with $N = 100$ in 1 and $N = 150$ in 2, and represented with respect to the stabilization parameter $\gamma \in [10^0, 10^{20}]$, for the different values of the stepsize $\alpha = 10^i$, $i = -3, -2, 0, 2, 3$.

Parameter (γ)	$\|\mathbf{A}\boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_n$	$\frac{\|\mathbf{A}\boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_n}{\|\mathbf{A}\boldsymbol{x}_\gamma^{[0]} - \mathbf{b}\|_n}$	$\|\boldsymbol{x}^\star - \boldsymbol{x}_\gamma^{[k]}\|_n$	$\frac{\|\boldsymbol{x}^\star - \boldsymbol{x}_\gamma^{[k]}\|_n}{\|\boldsymbol{x}^\star\|_n}$	# Iters $[k]$
10^4	$3.107685e^{+02}$	$6.324906e^{+00}$	$3.915976e^{+06}$	$6.598306e^{+04}$	10
10^5	$3.999870e^{-04}$	$8.140723e^{-06}$	$1.217203e^{+00}$	$2.050952e^{-02}$	1
10^8	$3.998163e^{-07}$	$8.137249e^{-09}$	$2.789201e^{-02}$	$4.699723e^{-04}$	1
10^{10}	$3.998161e^{-09}$	$8.137245e^{-11}$	$2.850475e^{-04}$	$4.802968e^{-06}$	1
10^{15}	$2.377694e^{-13}$	$4.839194e^{-15}$	$3.076279e^{-08}$	$5.183441e^{-10}$	1
10^{20}	$2.742299e^{-13}$	$5.581255e^{-15}$	$1.281315e^{-08}$	$2.158979e^{-10}$	1

Table 4: Results from the stabilized method (11) applied to the $2D$ image deblurring problem problem for different values of the stabilization parameter γ, under the dimension $n = N^2$ with $N = 200$, the stepsize $\alpha = 1$, the maximum iterations $k_{\max} = n$ and the criteria $\frac{\|\mathbf{b} - \mathbf{A}\boldsymbol{x}_\gamma^{[k]}\|_n}{\|\mathbf{b} - \mathbf{A}\boldsymbol{x}_\gamma^{[0]}\|_n} \leq \varepsilon$ where the tolerance $\varepsilon = 10^{-5}$. The data $\mathbf{A}$, $\mathbf{b}$ and the exact solution $\boldsymbol{x}^\star$ are computed from the Matlab code PRblurrotation(N) provided in [3] with the conditioning $\mathbf{cond}(\mathbf{A}) = 3.8925e^{+08}$.

Dimension $(n = N^2)$	$\|\mathbf{A}\boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_n$	$\frac{\|\mathbf{A}\boldsymbol{x}_\gamma^{[k]} - \mathbf{b}\|_n}{\|\mathbf{A}\boldsymbol{x}_\gamma^{[0]} - \mathbf{b}\|_n}$	$\|\boldsymbol{x}^\star - \boldsymbol{x}_\gamma^{[k]}\|_n$	$\frac{\|\boldsymbol{x}^\star - \boldsymbol{x}_\gamma^{[k]}\|_n}{\|\boldsymbol{x}^\star\|_n}$	# Itrs $[k]$	Cond$(\mathbf{A})$
$n = 50^2$	$1.793223e^{-14}$	$1.467459e^{-15}$	$1.479070e^{-12}$	$9.937144e^{-14}$	1	$8.4982e^{+04}$
$n = 100^2$	$9.114684e^{-14}$	$3.722874e^{-15}$	$7.886166e^{-11}$	$2.659608e^{-12}$	1	$4.7433e^{+06}$
$n = 200^2$	$2.742299e^{-13}$	$5.581255e^{-15}$	$1.281315e^{-08}$	$2.158979e^{-10}$	1	$3.8925e^{+08}$
$n = 300^2$	$2.532943e^{-13}$	$3.442336e^{-15}$	$2.001594e^{-08}$	$2.253546e^{-10}$	1	$3.9430e^{+08}$
$n = 400^2$	$3.542053e^{-13}$	$3.612476e^{-15}$	$1.202291e^{-08}$	$1.015846e^{-10}$	1	$3.9745e^{+08}$

Table 5: Results from the stabilized method (11) applied to the $2D$ image deblurring problem problem for different values of the dimension $n = N^2$ with $N = 50$, 100, 200, 300, 400, the stepsize $\alpha = 1$, the maximum iterations $k_{\max} = 100$, the criteria $\frac{\|\mathbf{b} - \mathbf{A}\boldsymbol{x}_\gamma^{[k]}\|_n}{\|\mathbf{b} - \mathbf{A}\boldsymbol{x}_\gamma^{[0]}\|_n} \leq \varepsilon$ where the tolerance $\varepsilon = 10^{-5}$, and the fixed stabilization parameter $\gamma = 10^{20}$. The data $\mathbf{A}$, $\mathbf{b}$ and the exact solution $\boldsymbol{x}^\star$ are computed from the code PRblurrotation(N) provided in [3].



ℹ Image deblurring test data computed from PRblurrotation(N) provided in [3] with $n = N^2, N = 400$: 1 true image, 2 blurred noise free image and 3 restored image by the stabilized gradient method (11). The restored image is computed under the parameter $\gamma = 10^{20}$, the stepsize $\alpha = 1$, and the criteria $\frac{\|\mathbf{b} - \mathbf{A}\boldsymbol{x}_\gamma^{[k]}\|_n}{\|\mathbf{b} - \mathbf{A}\boldsymbol{x}_\gamma^{[0]}\|_n} \leq \varepsilon$, where the tolerance $\varepsilon = 10^{-5}$.

Questions or comments?

References:

[1] A. Beck.
Introduction to nonlinear optimization : theory, algorithms, and applications with MATLAB.
Society for Industrial and Applied Mathematics and the Mathematical Optimization Society, 2014.

[2] I. Dione.
Stabilization of the gradient method for solving linear algebraic systems.
Submitted to Numerical Algorithms, 2025.

[3] S. Gazzola, P. C. Hansen, and J. G. Nagy.
Ir tools: a matlab package of iterative regularization methods and large-scale test problems.
Numerical Algorithms, 81(773-811), 2019.

[4] P. C. Hansen.
Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems.
Numerical Algorithms, 46:189–194, 2007.