Topics in Numerical Analysis II
Computational Inverse Problems

Lecturer: Bangti Jin (btjin@math.cuhk.edu.hk)

Chinese University of Hong Kong

October 12, 2022

Outline

Review

Nonlinear inverse problems

$$F(x) = y,$$

with $F : X \to Y$ being nonlinear operator (possibly compact) between Hilbert spaces X and Y

(i) F is continuous

(ii) F is weakly sequentially closed, i.e., $(x_n) \subset D(F)$, weak convergence of x_n to x^* and weak convergence of $F(x_n)$ to $F(x^*)$ in Y imply $x^* \in X$ and $F(x^*) \in Y$

- electrical impedance tomography
- diffuse optical tomography
- inverse scattering
- ...

Tikhonov regularization

$$J_\alpha(x) = \|F(x) - y\|^2 + \alpha\|x\|^2$$

and approximation x_α

$$x_\alpha \in \arg\min_{x \in D(F)} J_\alpha(x)$$

- existence of a global minimizer, not unique
- stability of the approximation (on subsequence)
- consistency of the approximation as $\delta \to 0$ (subseq.)
- convergence rate

convergence rate analysis Engl-Kunisch-Neubauer Inverse Problems 1989

Let $D(F)$ be **convex**, $y^\delta \in Y$ with $\|y^\delta - y^\dagger\| \leq \delta$, and $x^\dagger$ be an x_0-minimum norm solution. Moreover, the following condition holds

(i) F is Frechet differentiable.

(ii) $\exists\, L > 0$ s.t. $\|F'(x^\dagger) - F'(z)\| \leq L\|x^\dagger - z\|$ for all $z \in D(F)$

(iii) $\exists\, w \in Y$ s.t. $x^\dagger - x_0 = F'(x^\dagger)^* w$ and $L\|w\| < 1$.

Then with $\alpha \sim O(\delta)$, we have

$$\|x_\alpha^\delta - x^\dagger\| = O(\delta^{\frac{1}{2}}).$$

the same result holds for discrepancy principle

Model problem

Most inverse problems for PDEs are nonlinear in nature
model problem:

$$-\Delta u + qu = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega,$$

with given $q \in L^\infty(\Omega)$, $q \geq 0$, and $f \in L^2(\Omega)$

- The forward is well-posed: there exists a unique solution $u \in H_0^1(\Omega)$ (by Lax-Milgram theorem)
- Inverse problem: given $g \approx u^\delta$ in Ω, find q
- The operator $F(q) : q \mapsto u(q)$ is weakly continuous from $L^2(\Omega)$ to $L^2(\Omega)$.
- The operator F is differentiable from $L^2(\Omega)$ to $L^2(\Omega)$
- $q^\dagger = F'(q^\dagger)^* w = -u(q^\dagger)A(q^\dagger)^{-1}w$, with $w \in L^2(\Omega)$, i.e., $\frac{q^\dagger}{u(q^\dagger)} \in H_0^1(\Omega) \cap H^2(\Omega)$

discrete Tikhonov regularization

$$\min_{q_h\in\mathcal{A}_h}\{J_{\alpha,h}(q_h)=\tfrac{1}{2}\|F_h(q_h)-g\|^2_{L^2(\Omega)}+\tfrac{\alpha}{2}\|q_h\|^2_{L^2(\Omega)}\}$$

with $F_h(q_h)$ the FEM solution

- the discrete problem has a solution $q_h^*\in\mathcal{A}_h$
- convergence as $h\to 0^+$ (subsequence)

projected gradient method

Now we need to solve the optimization problem over $q \in \mathcal{A}$

$$J_\alpha(q) = \frac{1}{2}\|F(q) - g\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|q\|^2_{L^2(\Omega)}$$

gradient descent method: given q^0, solve for $k = 1, 2, \ldots$

$$q^{k+1} = P_{\mathcal{A}}(q^k - \tau J'(q^k))$$

- $\tau > 0$: a step size (Amijo's rule etc)
- $J'_\alpha(q^k) \in L^2(\Omega)$ the gradient at q^k: $J'_\alpha(q^k) = -u(q^k)p + \alpha q^k$

$$\int_\Omega \nabla v \cdot \nabla p + \int_\Omega q^k vpdx = \int_\Omega v(F(q^k) - g)dx \quad \forall v \in H^1_0(\Omega)$$

The pros and cons

- relatively well developed theory (existence, convergence)
  currently the most popular approach for inverse problems
- the choice of the penalty and regularization parameter is crucial
- requires repeated solution of optimization problem via optimizers
- ...

Regularization by truncating iterative solvers

for the linear system

$$Ax = y,$$

with $\mathbb{R}^{m\times n}$, $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$

- there are a lot of iterative methods for solving linear systems: conjugate gradient, Krylov subspace (GMRES, MINRES, ...), ...
- an iterative solver attempts to solve the problem by finding successive approximations, starting from some x^0, and then

$$x^{n+1} = f_n(x^n, x^{n-1}, ...)$$

- Typically the update involves multiplications by A and its adjoint $A^\top$, but not explicit computation of the inverse.

why iterative methods

- sometimes the only feasible choice if the problem involves a large number of variables, making the direct methods (e.g., Gauss elimination) prohibitive
- especially practical if multiplications by A / A^* are cheap e.g., dedicated implementation on GPUs
- usually not designed for ill-posed equations, but often posses self-regularizing properties: if the iterations are terminated before the solution starts to fit to noise (i.e., early stopping), one often obtains reasonable solutions for inverse problems.
- iteration index plays the role of *regularization parameter*

How to construct an iterative solver for

$$Ax = y?$$

minimize the residual

$$J(x) = \frac{1}{2}\|Ax - y\|^2$$

The global minimizer in $\mathbb{R}^n$ is given by $x^\dagger := A^\dagger y \in \mathbb{R}^n$, i.e., solving normal equation

$$A^\top A x = A^\top y$$

and is orthogonal to $\ker(A)$. (this is not good for inverse problems.)

Instead we proceed iteratively: given x^0, compute

$$\begin{aligned} x^{k+1} &= x^k - \beta \nabla J(x^k) \\ &= x^k - \beta A^\top (Ax^k - y) \end{aligned}$$

Can one use it to construct approximation ?

- $\beta > 0$, step size (learning rate)
- This method is known as the Landweber method Landweber Amer. J. Math. 1951
- Landweber + early stopping is regularizing (later)
- optimal convergence rates (later)
- very slow ...

main tool for well-posed problems: Banach fixed point iteration
Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a vector-valued function.

- A set $S \subset \mathbb{R}^n$ is called an invariant set for T if

$$T(S) \subset S, \quad \text{i.e. } T(x) \in S \quad \forall x \in S.$$

- T is a contraction on an invariant set S if there exists $0 \leq k < 1$ s.t.

$$\|T(x) - T(y)\| \leq \kappa \|x - y\|, \quad \forall x, y \in S.$$

- a vector $x \in \mathbb{R}^n$ is called a fixed point of T if

$$T(x) = x$$

Theorem (Banach fixed point theorem)

Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a contraction on the closed invariant set S. Then there exists a unique fixed point $x \in S$ of T. Further, the fixed point can be found by following fixed point iteration:

$$x = \lim_{k\to\infty} x_k, \quad x_{k+1} = T(x_k),$$

for any $x_0 \in S$.

This result is valid on any metric space.

Let $x_n = T(x_{n-1})$. Then

$$\|x_{n+1} - x_n\| \leq \kappa\|x_n - x_{n-1}\| \leq \cdots \leq \kappa^n\|x_1 - x_0\|$$

the sequence (x_n) is Cauchy:

$$\begin{aligned}\|x_m - x_n\| &\leq \|x_m - x_{m-1}\| + \ldots + \|x_{n+1} - x_n\| \\ &\leq \kappa^{m-1}\|x_1 - x_0\| + \ldots \kappa^n\|x_1 - x_0\| \\ &= \kappa^n\|x_1 - x_0\| \sum_{j=0}^{m-1-n} \kappa^j = \frac{\kappa^n(1-\kappa^{m-n})}{1-\kappa}\|x_1 - x_0\| \\ &\leq \frac{\kappa^n}{1-\kappa}\|x_1 - x_0\|\end{aligned}$$

Thus it is Cauchy. There exists a limit x^*, and x^* satisfies

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} T(x_{n-1}) = T(\lim_{n\to\infty} x_{n-1}) = T(x^*)$$

uniqueness follows by contradiction.

define a mapping T by

$$T(x) = x + \beta(A^\top y - A^\top Ax), \quad \beta \in \mathbb{R}$$

any solution to the normal equation is a fixed point of T.
message: if β is small, there is a unique fixed point of T in $\ker(A)^\perp$, i.e. $x^\dagger$, and it can be reached by fixed point iteration if $x_0 = 0$.

Theorem

Let $0 < \beta < 2s_1^{-2}$ be fixed. Then the fixed point iteration

$$x_{k+1} = T(x_k), \quad x_0 = 0$$

converges towards $x^\dagger$ as $k \to \infty$.

Let $S = \ker(A)^\perp = \mathrm{range}(A^\top)$. Then $T(S) \subset S$ since

$$T(x) = x + A^\top(\beta y - \beta Ax) \in \mathrm{range}(A^\top)$$

for all $x \in \mathrm{range}(A^\top)$. Thus, S is invariant under T.
Note also that by singular system $(s_j, u_j, v_j)_{j=1}^r$

$$Ax = \sum_{j=1}^r s_j(x, v_j)u_j, \quad A^\top y = \sum_{j=1}^r s_j(y, u_j)v_j$$

and also

$$x = \sum_{j=1}^r (x, v_j)v_j, \quad \forall x \in S.$$

Let $x, z \in S$ and $x - z \in S$. Then

$$\begin{aligned} T(x) - T(z) &= (x - z) - \beta A^\top A(x - z) \\ &= \sum_{j=1}^{r} (v_j, x - z) v_j - \beta \sum_{j=1}^{r} s_j^2 (v_j, x - z) v_j \\ &= \sum_{j=1}^{r} (1 - \beta s_j^2)(v_j, x - z) v_j \end{aligned}$$

Since s_1 is the largest singular value, by assumption, we have

$$-1 < \beta s_j^2 - 1 \leq \beta s_1^2 - 1 < 1, \quad j = 1, \ldots, r.$$

Hence, we have

$$\kappa := \max_{j=1,\ldots,r} |\beta s_j^2 - 1| < 1.$$

Consequently,

$$\begin{aligned}\|T(x) - T(z)\|^2 &= \sum_{j=1}^{r} (1 - \beta s_j^2)^2 (v_j, x - z)^2 \\ &\leq \kappa^2 \sum_{j=1}^{r} (v_j, x - z)^2 = \kappa^2 \|x - z\|^2.\end{aligned}$$

Thus, T is a contraction on S. Since S is also a closed invariant subset of T, there exists a unique fixed point of T in S.
To complete the proof, recall that $x^\dagger = A^\dagger y$ belongs to $S = \ker(A)^\perp$ and satisfies the normal equation. Further, since $x_0 = 0$ is in S, the fixed point iteration starting from x_0 converges to $x^\dagger$.

Regularizing properties of Landweber method

The kth iteration of the Landweber iteration can be written explicitly

$$x_k = \sum_{j=1}^{r} s_j^{-1}(1-(1-\beta s_j^2)^k)(y,u_j)v_j, \quad k = 0, 1, \ldots$$

Since $|1-\beta s_j^2| < 1$ by assumption,

$$(1-\beta s_j^2)^k \to 0 \quad \text{as } k \to \infty.$$

which is expected since

$$x^\dagger = \sum_{j=1}^{r} s_j^{-1}(y,u_j)v_j.$$

However, while $k \in \mathbb{N}$ is finite, the coefficients of $(y, u_j)v_j$ satisfy

$$s_j^{-1}(1-(1-\beta s_j^2)^k) = s_j^{-1}\left(1-\sum_{\ell=0}^{k}\binom{k}{\ell}(-1)^\ell \beta^\ell s_j^{2\ell}\right)$$
$$= \sum_{\ell=1}^{k}\binom{k}{\ell}(-1)^{\ell+1}\beta^\ell s_j^{2\ell-1}$$

which converges to zero as $s_j \to 0$ (for a fixed k)

While, k is small enough, no coefficients of $(y, u_j)v_j$ is so large that the component of the measurement noise in the direction u_j is amplified in an uncontrolled manner.

discrepancy principle

Let y be a noisy version of some underlying exact data $y^\dagger$ and

$$\|y - y^\dagger\| = \delta > 0.$$

The Morozov discrepancy principle works for the Landweber iteration is similar to the truncated SVD and Tikhonov regularization: choose the smallest $k \geq 0$ s.t.

$$\|y^\delta - Ax_k^\delta\| \leq \delta.$$

Such a stopping rule exists if

$$\delta > \|y^\delta - Py^\delta\| = \|y - Ax^{\delta,\dagger}\|.$$

with $x^{\delta,\dagger} = A^\dagger y^\delta$. Indeed since the sequence $(x_k)_k$ converges to $x^{\delta,\dagger} = A^\dagger y^\delta$, for $\delta > \|y^\delta - Ax^{\delta,\dagger}\|$, there exists $k = k_\delta$ s.t.

$$\|x_k^\delta - x^{\delta,\dagger}\| \leq \|A\|^{-1}(\delta - \|y^\delta - Ax^{\delta,\dagger}\|),$$

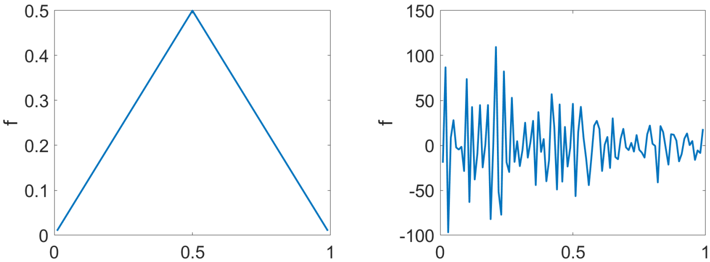
and thus by the reverse triangle inequality

$$\begin{aligned} \|y^\delta - Ax_k^\delta\| - \|y^\delta - Ax^{\delta,\dagger}\| &\leq \|A(x_k^\delta - x^{\delta,\dagger})\| \\ &\leq \|A\|\|x_k^\delta - x^{\delta,\dagger}\| \\ &\leq \delta - \|y^\delta - Ax^{\delta,\dagger}\| \end{aligned}$$

This clearly implies $\|y^\delta - Ax_k^\delta\| \leq \delta$.

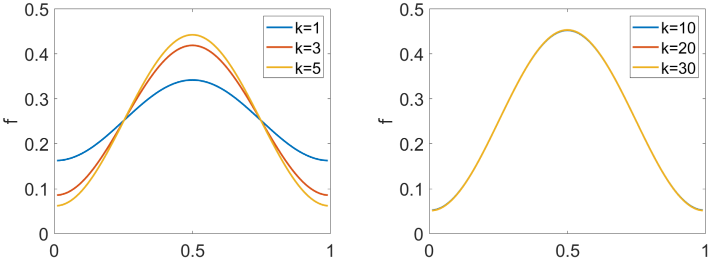
Example: heat conduction

- w: the simulated heat distribution at $T = 0.1$
- $f^\dagger$: wedge function (initial data)
- $A = e^{TB}$ forward operator, $\|A\| = 1$, $\beta = 1$
- a small amount of noise to the measurement, discrepancy principle

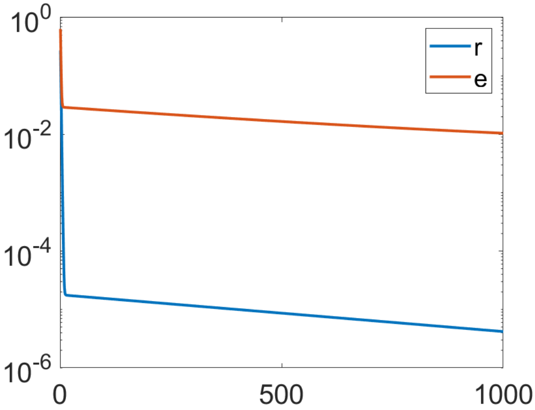


exact solution v.s. least-squares solution

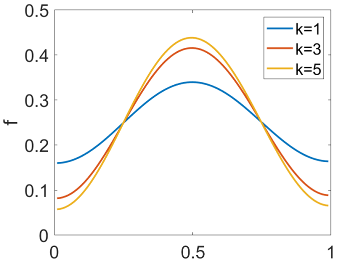
solution for exact data

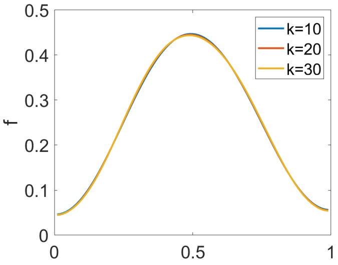


convergence for exact data

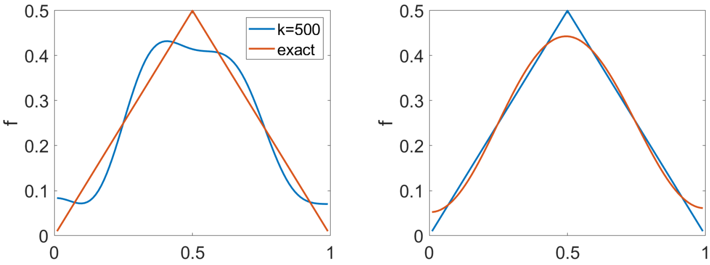


results for noisy data (1% noise)



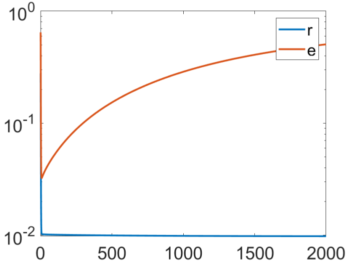


results for noisy data



discrepancy principle stops at $k = 6$

results for noisy data



convergence issue revisited

When the problem is ill-posed, $I - \beta A^* A$ is not a contraction !

problem setting:

$$Ax = y^\delta,$$

with $\|y^\delta - y^\dagger\| \le \delta$, construct an approximation $x^\delta_{k(\delta)}$ by Landweber method, i.e.,

$$x_k^\delta = x_{k-1}^\delta - A^*(Ax_{k-1}^\delta - y^\delta), \quad k = 1, \dots,$$

(with $x_0^\delta = 0$, $\|A\| \le 1$, by rescaling) s.t.

$$\lim_{\delta \to 0^+} x^\delta_{k(\delta)} = x^\dagger?$$

If $y^\dagger \in D(A^\dagger)$, then $x_k \to A^\dagger y^\dagger$ as $k \to \infty$. If $y \notin D(A^\dagger)$, then $\|x_k\| \to \infty$ as $k \to \infty$.

The iterate x_k is given by

$$x_{k+1} = \sum_{j=0}^{k} (I - A^*A)^j A^* y^\dagger$$

with $y^\dagger \in D(A^\dagger)$, then $A^*y^\dagger = A^*Ax^\dagger$ for $x^\dagger = A^\dagger y^\dagger$, and

$$x^\dagger - x_{k+1} = x^\dagger - A^*A \sum_{j=0}^{k} (I - A^*A)^j x^\dagger$$

since $A^*A\sum_{j=0}^{k}(I - A^*A)^j x^\dagger = I - (I - A^*A)^{k+1}$, i.e.,

$$x^\dagger - x_{k+1} = (I - A^*A)^{k+1} x^\dagger$$

by the singular system (s_j, u_j, v_j) of A

$$\|x_{k+1} - x^\dagger\|^2 = \sum_j (1 - s_j^2)^{2(k+1)} (x^\dagger, v_j)^2$$

uniformly bounded + Lebesgue dominated convergence theorem convergence rate under the source condition:

$$x^\dagger = A^* w$$

$\Rightarrow$

$$x^\dagger - x_{k+1} = (I - A^*A)^{k+1} x^\dagger = (I - A^*A)^{k+1} A^* w$$

by means of spectral decomposition:

$$\|x_{k+1} - x^\dagger\|^2 = \sum_j s_j^2 (1 - s_j^2)^{2(k+1)} (w, u_j)^2$$

Note that $\sup_{\lambda \in [0,1]} \lambda(1-\lambda)^k \le (k+1)^{-1}$. The error decay

$$\|x_{k+1} - x^\dagger\|^2 \le (2k+3)^{-1} \sum_j (w, u_j)^2 = (2k+3)^{-1} \|w\|^2$$

i.e.,

$$\|x_{k+1} - x^\dagger\| \le (2k+2)^{-1/2} \|w\|$$

Let $y^\dagger, y^\delta$ be a pair of data with $\|y^\delta - y^\dagger\| \le \delta$. Then

$$\|x_k - x_k^\dagger\| \le \sqrt{k}\delta, \quad k \ge 0.$$

$$x_k - x_k^\delta = \sum_{j=0}^{k-1}(I - A^*A)^j A^*(y^\dagger - y^\delta) := R_k(y^\dagger - y^\delta)$$

and

$$\|R_k\|^2 = \|R_k R_k^*\| = \|\sum_{j=0}^{k-1}(I-A^*A)^j(I-(I-A^*A)^k)\| \le \|\sum_{j=0}^{k-1}(I-A^*A)^j\| \le k.$$

error analysis:

$$x^\dagger - x_k^\delta = x^\dagger - x_k + x_k - x_k^\delta$$

- approximation error: $x^\dagger - x_k$ converges to zero as $k \to \infty$
- data error $x_k - x_k^\delta$, of order $k^{\frac{1}{2}}\delta$
- optimal convergence (under $x^\dagger = A^* w$ + a priori choice of $k(\delta)$)

$$\|x_k^\delta - x^\dagger\| \leq c\delta^{\frac{1}{2}}, \quad \text{with } k(\delta) = \delta^{-1}.$$

- semiconvergence: the regularizing property of iterative methods (for ill-posed) problems ultimately depend on reliable stopping discrepancy principle: choose $k(\delta)$ s.t.

$$\|y^\delta - Ax_k^\delta\| \leq \tau\delta,$$

with $\tau > 1$ fixed

motivation

- If $\|y^\delta - Ax_k^\delta\| > 2\delta$, then x_{k+1}^δ approximates $x^\dagger$ better than x_k^δ.
- The sequence $\|y^\delta - Ax_k^\delta\|$ is monotonically decreasing.

$$
\begin{aligned}
&\|x_{k+1}^\delta - x^\dagger\|^2 = \|x_k^\delta - A^*(Ax_k^\delta - y^\delta) - x^\dagger\|^2 \\
=&\|x^\dagger - x_k^\delta\|^2 - 2(x^\dagger - x_k^\delta, A^*(y^\delta - Ax_k^\delta)) + (y^\delta - Ax_k^\delta, AA^*(y^\delta - Ax_k^\delta)) \\
=&\|x^\dagger - x_k^\delta\|^2 - 2(y^\dagger - y^\delta, y^\delta - Ax_k^\delta) - \|y^\delta - Ax_k^\delta\|^2 \\
&+ (y^\delta - Ax_k^\delta, (AA^* - I)(y^\delta - Ax_k^\delta))
\end{aligned}
$$

since $A^*A - I$ is negative semidefinite,

$$
\|x^\dagger - x_k^\delta\|^2 - \|x^\dagger - x_{k-1}^\delta\|^2 \le \|y^\delta - Ax_k^\delta\|(\|y^\delta - Ax_k^\delta\| - 2\delta) > 0
$$

for $\|y^\delta - Ax_k^\delta\| > 2\delta$

monotonicity of residual

$$
\begin{aligned}
&\|Ax_{k+1}^{\delta} - y^{\delta}\|^2 - \|Ax_k^{\delta} - y^{\delta}\|^2 \\
=&\|Ax_k^{\delta} - y^{\delta} - AA^*(Ax_k^{\delta} - y^{\delta})\|^2 - \|Ax_k^{\delta} - y^{\delta}\|^2 \\
=&-2(Ax_k^{\delta} - y^{\delta}, AA^*(Ax_k^{\delta} - y^{\delta})) + \|AA^*(Ax_k^{\delta} - y^{\delta})\|^2 \\
=&(\|A\|^2 - 2)\|A^*(Ax_k^{\delta} - y^{\delta})\|^2 < 0
\end{aligned}
$$

since $\|A\| \leq 1$ by assumption

The Landweber method with the discrepancy principle is order optimal: for $x^\dagger = A^* w$,

$$\|x_k^\delta - x^\dagger\| \leq c\delta^{\frac{1}{2}},$$

and the required number of iterations is $k(\delta) = O(\delta^{-1})$.

- The Landweber method can be very slow ...
- How to accelerate the computation ...
  - simple: Anderson acceleration
  - simple: Kaczmarz / stochastic gradient descent
  - complex: conjugate gradient, MINRES
  - ...

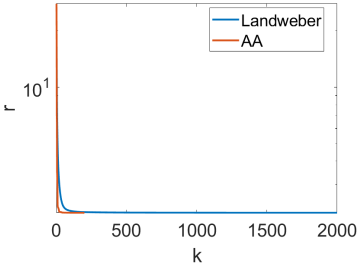
Anderson acceleration for fixed point equation: $x_{n+1} = T(x_n)$

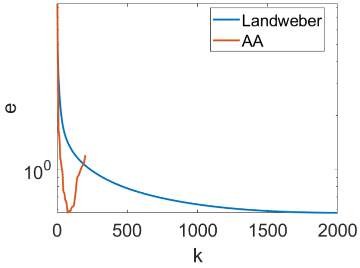
- let $g(x) = T(x) - x$, $g_k = g(x_k)$
- set x_0 and $m \geq 1$ (memory parameter)

D. G. Anderson. Iterative Procedures for Nonlinear Integral Equations. J. the ACM. 1965; 12 (4): 547–560

$x_1 = T(x_0)$
for $k = 1, 2, \ldots$ **do**
  $m_k = \min(m, k)$
  $G_k = [g_{k-m_k} \ \ldots \ g_k]$
  $\alpha_k = \arg\min_{\sum_{i=0}^{m_k} \alpha_i = 1} \|G_k \alpha\|$
  $x_{k+1} = \sum_{i=0}^{m_k} (\alpha_k)_i f(x_{k-m_k+i})$
end for

Landweber iteration v.s. Anderson acceleration for `gravity`





Anderson acceleration

- Landweber: slow convergence vs. slow divergence
- Anderson: fast convergence v.s. fast divergence
- no analysis of the regularizing property of AA ! (open)
-