Anderson Acceleration for a Class of Nonsmooth Fixed-Point Problems

C. T. Kelley
NC State University
tim_kelley@ncsu.edu
Joint with Xiaojun Chen and Wei Bian.

Supported by NSF, DOE(CASL/ORNL), ARO

SIAM-LA, May 19, 2021

Outline

Anderson Acceleration

$\texttt{anderson}(\mathbf{u}_0, \mathbf{G}, m)$

$\mathbf{u}_1 = \mathbf{G}(\mathbf{u}_0)$; $\mathbf{F}_0 = \mathbf{G}(\mathbf{u}_0) - \mathbf{u}_0$

for $k = 1, \dots$ **do**

$m_k \le \min(m, k)$

$\mathbf{F}_k = \mathbf{G}(\mathbf{u}_k) - \mathbf{u}_k$

Minimize $\| \sum_{j=0}^{m_k} \alpha_j^k \mathbf{F}_{k-m_k+j} \|$ subject to $\sum_{j=0}^{m_k} \alpha_j^k = 1$.

$\mathbf{u}_{k+1} = \sum_{j=0}^{m_k} \alpha_j^k \mathbf{G}(\mathbf{u}_{k-m_k+j})$

end for

Terminology

- m, depth. We refer to Anderson(m).
  Anderson(0) is Picard.
- $\mathbf{F}(\mathbf{u}) = \mathbf{G}(\mathbf{u}) - \mathbf{u}$, residual
- $\{\alpha_j^k\}$, coefficients
  Minimize $\| \sum_{j=0}^{m_k} \alpha_j^k \mathbf{F}_{k-m_k+j} \|$ subject to $\sum_{j=0}^{m_k} \alpha_j^k = 1$.
  is the optimization problem.
- $\| \cdot \|$, ℓ^2, ℓ^1, or ℓ^∞

Typical local convergence theorem (Toth-K 2015, Chen-K 2019)

Assume:

- $\mathbf{G}$ is a C^1 contraction, contractivity constant $c < 1$
- Solution of the optimization is no worse than Picard.
- Either $m = 1$ or there is M_α such that for all $k \geq 0$

$$\sum_{j=1}^{m_k} |\alpha_j| \leq M_\alpha.$$

- You start close to the solution $\mathbf{u}^*$ (warm start, EDIIS?)

Then, pessimistically

$$\limsup_{k \to \infty} \left(\frac{\|\mathbf{F}(\mathbf{u}_k)\|}{\|\mathbf{F}(\mathbf{u}_0)\|} \right)^{1/k} \leq c.$$

Analysis

- Most theory assumes differentiability
- But there are no derivatives in the method
- Exceptions:
  - Global phase of EDIIS (Chen-Kelley, 2019)
  - This talk.

Splittable nonlinearity, (Chen-Yamamoto 89)

Newton: Solve $\mathbf{F}(\mathbf{x}) = 0$

$$\mathbf{F} = \mathbf{F}_S + \mathbf{F}_N$$

- Solution $\mathbf{x}^*$
- $\mathbf{F}_S$ nonsingular near $\mathbf{x}^*$ and $\mathbf{F}_S'$ Lip cont
- $\mathbf{F}_N$ Lipschitz continuous with small Lip const ϵ

Newton: $\mathbf{x}_+ = \mathbf{x}_c - \mathbf{F}_S(\mathbf{x}_c)^{-1}\mathbf{F}(\mathbf{x}_c)$
Then ...

Convergence of $\mathbf{e} = \mathbf{x} - \mathbf{x}^*$ to 0

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \mathbf{F}'_S(\mathbf{x}_n)^{-1}\mathbf{F}(\mathbf{x}_n) \to \mathbf{x}^*$$

and

$$\|\mathbf{e}_{n+1}\| = O(\|\mathbf{e}_n\|^2 + \epsilon\|\mathbf{e}_n\|)$$

Proof:

$$\begin{aligned} \mathbf{e}_{n+1} &= (\mathbf{e}_n - \mathbf{F}'_S(\mathbf{x}_n)^{-1}(\mathbf{F}_S(\mathbf{x}_n) - \mathbf{F}_S(\mathbf{x}^*)) \\ &+\mathbf{F}'_S(\mathbf{x}_n)^{-1}(\mathbf{F}_N(\mathbf{x}^*) - \mathbf{F}_N(\mathbf{x}_n)) \end{aligned}$$

Many appplications but fashion moved on

- Lots of activity in 1990s
  Heinkenschloss, Tran, K 92, Sachs-K 90s
- Semismooth Newton and smoothing Newton took over.
- but . . . CFD! (Coffey, McMullan, McRae, K., 03)

Problem: aside from CFD, you must figure out the splitting.

- Same assumptions except **G** splittable
- Same algorithm, same code
- Same theorem except $c \to c + \epsilon$
- Uglier proof

Compact fixed point problems

$$G_l(u)(x) = \int_0^1 g(x,y)\Phi(u(y))\,dy.$$

where $g \in C$ and

$$\Phi(u(x)) = \beta(u(x) + b(x))$$

and $\beta(u)$ is differentiable except for finitely many points.
Example: $\beta(u) = |u|$

The splitting: $\beta(u) = |u|$

Let

$$\Omega_\rho = \{x \,|\, |u^*(x) + b(x)| < 2\rho\} \text{ and } \Omega_\rho^c = \{x \,|\, |u^*(x) + b(x)| \geq 2\rho\}.$$

We define

$$G_N^\rho(u)(x) = \int_{\Omega_\rho} g(x,y)|u(y) + b(y)| \, dy$$

and

$$G_S^\rho(u) = G(u) - G_N^\rho(u) = \int_{\Omega_\rho^c} g(x,y)|u(y) + b(y)| \, dy$$

The splitting's good

- $\|u(x) - u^*(x)\|_\infty < \rho$ implies $u + b$ has the same sign as $u^* + b$ for $u \in \Omega_\rho^c$
  So G^S is smooth
- Lip constant of G_ρ^N is proportional to $m(\Omega_\rho)$
  So if $u^*(x) + b(x) = 0$ only finitely often, $m(\Omega_\rho) \to 0$.

Simple example

Two-point boundary-value problem (Chen-Nashed 2000)

$$-v'' = \lambda \max(v - \alpha, 0), \; v(0) = v_0, v(1) = v_1$$

Convert to fixed point problem by

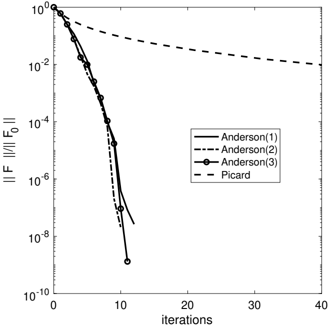
- Set $v = u + \phi$, where $\phi(x) = v_1 x + (1 - x) v_0$
- Let G be the Greens function for $-d^2/dx^2$ with zero bc
- Multiply the equation by G and ...

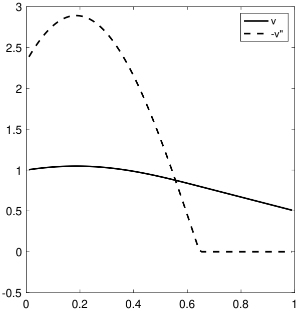
Fixed point formulation

$$u = \lambda G(\max(u + \phi - \alpha, 0))$$

- Solve with Anderson(m) for $m = 0, 1, 2, 3$.
  Increasing m to 2 or 3 makes little difference.
- We plot iteration histories, v, and $-v''$.
  See how constraints on u'' are active.
- $\lambda = 11.65$, $\alpha = .8$
- Central differences with 100 interior grid points.

Residual Histories and solution





Summary

- Nonlinear solver for nonsmooth fixed-point problems
  - Splittable nonlinearities
  - Requires no human intervation, no new algorithm unlike Newton
  - Structure only used in analysis
- Good for 15 minute talk
- W. Bian, X. Chen, and C. T. Kelley, Anderson acceleration for a class of nonsmooth fixed-point problems, SIAM J. Sci. Comp., (2021). doi:10.1137/20M132938X. Published online Jan 20, 2021.