Anderson Acceleration: Software, Storage, and a Multi-Physics Example

C. T. Kelley
NC State University
tim_kelley@ncsu.edu

Supported by NSF, DOE, ARO

ICERM, July 26, 2023

Outline

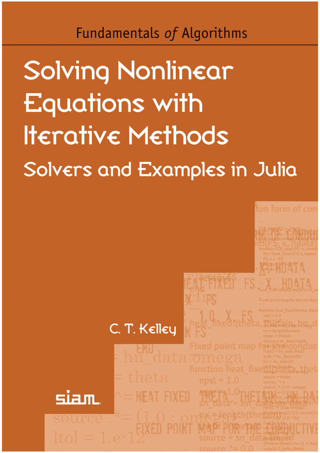
1 New Book

2 Storage Issues in Anderson Acceleration

3 Conductive-Radiative Heat Transport

4 Summary

New Book: Great holiday gift



Coverage

- Sequel to
  C. T. Kelley, Solving Nonlinear Equations with Newton's Method, number 1 in Fundamentals of Algorithms, SIAM, Philadelphia, 2003.
- Differences
  - Completely new code in Julia
  - Deletion: Broyden's method
  - Additions: ΨTC + Anderson Acceleration
  - Software + Jupyter notebook open source on Github

Why Julia?

- Better mixed precision, especially half, support
  C. T. KELLEY, Newton's method in mixed precision, SIAM Review, 64 (2022), pp. 191–211.
- Jupyter notebooks
- Package infrastructure for software
- No licensing pain
  Good for retired people

Three part project

- Print book
- IJulia (aka Jupyter) notebook
  interactive version of print book
  https://github.com/ctkelley/NotebookSIAMFANL
- Julia package with solvers+test problems+examples
  https://github.com/ctkelley/SIAMFANLEquations.jl

Print book is for sale. Notebook and Package are free.

How to do projects like this

- Motivated by R. LeVeque's project:
  - Print book
  - Notebooks
  - notebook $\rightarrow$ print book with Python
- Mappings are different
  - notebook $\rightarrow$ print book
  - print book $\rightarrow$ notebook
- Many things to figure out on your own.
- Randy and Tim were both retired while doing this.

Coverage of Anderson Acceleration

- Nothing fancy (fixed β, no EDIIS, no depth management, ...)
- Theory (Basic stuff, no proofs)
- Discussion of Implementation
- CODE
- Examples

The rest of this talk is about the implementation and one example.

What does AA store?

The simple version of AA fori $x = G(x)$ is $\texttt{anderson}(x_0, G, m)$

$x_1 = G(x_0)$; $F_0 = G(x_0) - x_0$

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

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

$F_k = G(x_k) - x_k$

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

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

end for

Remember the words

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

Solving the Optimization Problem

Solve the linear least squares problem:

$$\min \left\| F_{m_k} - \sum_{j=0}^{m_k-1} \alpha_j^k (F_{k-m_k+j} - F_k) \right\|_2^2,$$

for $\{\alpha_j^k\}_{j=0}^{m_k-1}$ and then

$$\alpha_{m_k}^k = 1 - \sum_{j=0}^{m_k-1} \alpha_j^k.$$

More or less what's in the codes, BUT ...

Storage

- you have to store
  - Coefficient matrix for the optimization problem F
  - Iteration history G
- So a minimum of $2m + O(1)$ vectors.
- Even if you manage $m_k \leq m$, you must store $2m$ vectors.
- Many physics codes use the normal equations and get $2m$. Example RMG
- But the coefficient matrix is very ill-conditioned ...

Walker-Ni Formulation: I

Maintain a QR factorization of the coefficient matrix.
Reformulate the optimization problem:

$$\min_{\theta \in R^{m_k}} \|F(x_k) - \sum_{j=0}^{m_k-1} \theta_j (F(x_{k-m_k+j+1}) - F(x_{k-m_k+j}))\|$$

for $\theta^k \in R^{m_k}$ then

$$x_{k+1} = G(x_k) - \sum_{j=0}^{m_k-1} \theta_j^k (G(x_{k-m_k+j+1}) - G(x_{k-m_k+j})).$$

In terms of the original formulation

$\alpha_0 = \theta_0, \alpha_j = \theta_j - \theta_{j-1}$ for $1 \le j \le m_k - 1$ and $\alpha_{m_k} = 1 - \theta_{m_k-1}$.

Walker-Ni Formulation: II

Let D_F^k and D_G^k have columns

$$(D_F^k)_j = F(x_{k-m_k+j+1}) - F(x_{k-m_k+j})$$

and

$$(D_G^k)_j = G(x_{k-m_k+j+1}) - G(x_{k-m_k+j}).$$

Then the optimization problem is

$$\min \| F(x_k) - D_F^k \theta^k \|$$

and

$$x_{k+1} = G(x_k) - D_G^k \theta.$$

Walker-Ni Formulation: III

Walker-Ni update the QR factorization of D_F^k.

- Update D_G by adding a new column and deleteing the leading column (if $k > m_k - 1$) if necessary.
- Update D_F by adding a new column and deleting the old column if necessary.

Updating the QR factorization of D_F is tricky ...

Walker-Ni Formulation: IV

- If $k > m_k - 1$ you have to update QR to be the factorization of D_F with the first column deleted.
- After that update QR with the new column via Gram-Schmidt I like classical GS twice.

The story of the downdate is . . .

Walker-Ni Formulation: IV

- Suppose $A = QR$ is the QR factorization of an $N \times m$

$$A = (a_1, a_2, \ldots, a_m) = QR = Q(r_1, r_2, \ldots, r_m).$$

- I want the QR factorization of $\hat{A} = (a_2, \ldots, a_m)$
- Let $\tilde{R} = (r_2, \ldots, r_m)$, then $\hat{A} = Q\tilde{R}$.
- Factor $\tilde{R} = Q^1 \hat{R}$. Let $\hat{Q} = QQ^1$.
- Then $\hat{A} = \hat{Q}\hat{R}$ is the QR factorization of $\hat{A}$

Storage Problem!

- You must allocate storage for both $\hat{Q}$ and QQ^1,
- then you can overwrite Q with $\hat{Q}$.
- So with AA you are now at $3m$ vectors.
- But there's a hack . . .
  Compute the product QQ^1 in blocks of rows.
  Overwrite the rows of Q as you progress.
  Slower, but the storage is back under control.
- There's always the normal equations.

Conductive-Radiative Heat Transport

(Siewert-Thomas, 91)

- System of two equations
  - Linear Boltzmann transport equation
  - Heat Equation
- Household insulation is the 2D problem
- Coupling via "radiation proportional to $4th$ power of temperature"
- I'm a mathematician, so it's 1-D in space.
- There's Julia code for all of this in the repo.

Geometry and unknowns

- Unknowns:
  - Dimensionless radiation intensity $\psi(x, \mu)$ (angular flux)
  - Dimensionless temperature $\Theta(x)$
- Radiation depends on direction.
  Stand in the sun if you don't believe that.
  μ is cosine of direction angle. $\mu > 0$ means from the left.
  $\mu < 0$ means from the right.

Transport Equation

$$\mu \frac{\partial \psi}{\partial x}(x,\mu) + \psi(x,\mu) = \frac{\omega}{2}\int_{-1}^{1} \psi(x,\mu')\, d\mu' + (1-\omega)\Theta^4(x)$$

for $x \in (0,\tau)$. $0 < \omega \leq 1$ and the boundary conditions are

$$\psi(0,\mu) = \Theta_l^4, \;\; \mu > 0 \text{ and } \psi(\tau,\mu) = \Theta_r^4, \;\; \mu < 0.$$

Some facts about transport with Θ given

Trust me on this. Details in the book.

Given Θ the transport equation is a linear equation for f

- Can formulate as a linear equation for the scalar flux

$$f(x) = \frac{1}{2}\int_{-1}^{1} \psi(x, \mu')\, d\mu'$$

- The operator is a compact perturbation of the identity, so ...
- GMRES works very well as a solver.

Heat equation

$$\frac{\partial^2 \Theta}{\partial x^2} = Q(x), x \in [0, \tau], \ \Theta(0) = \Theta_l, \Theta(\tau) = \Theta_r$$

where

$$Q(x) = \alpha(x)(\Theta^4(x) - f(x)), 0 < x < \tau \ , \ f(x) = \frac{1}{2}\int_{-1}^{1} \psi(x, \mu')\, d\mu'$$

and

$$\alpha(x) = (1 - \omega)/N_c.$$

N_c is the conduction to radiation parameter.

Compact fixed point formulation: $\Theta = \mathcal{G}(\Theta)$

1. Given Θ solve the transport equation with an iterative method to obtain f.
2. Use the solution f of the transport equation from step 1 to compute $Q = \alpha(\Theta^4 - f)$.
3. Compute $\mathcal{G}(\Theta) = T$ as the solution of the heat equation.

$$\frac{\partial^2 T}{\partial x^2} = Q(x), x \in [0, \tau],\ T(0) = \Theta_l,\ T(\tau) = \Theta_r$$

We choose to **expose** Θ. Exposing f is ok.

Discretization

- $N = 1001$ point uniform grid in space
- Discrete ordinates Sn method for transport
  - Double (20 pt) Gauss quadrature rule in angle
- Central difference for heat equation.

Examples

Problem parameters: τ, ω, N_c, θ_r, θ_l

Three problems

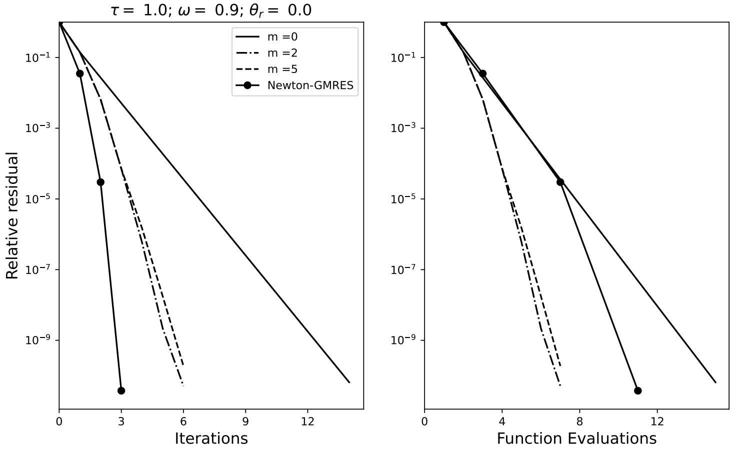
Easy: $N_c = .05, \omega = .9, \tau = 1, \Theta_l = 0, \Theta_r = 0$

Less Easy : $N_c = .05, \omega = .9, \tau = 2, \Theta_l = 0, \Theta_r = 1.8$

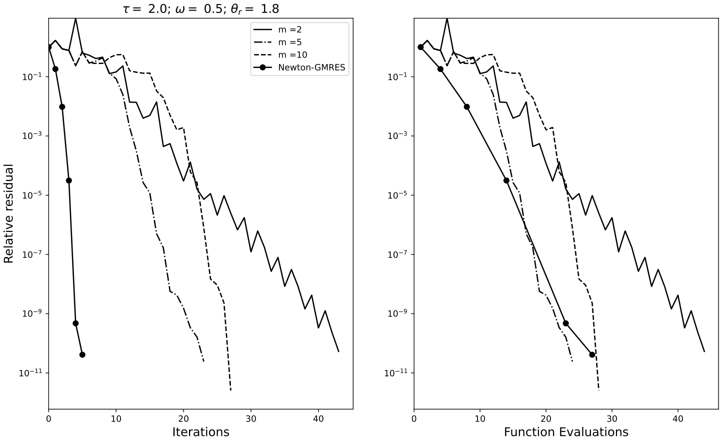
Hard: $N_c = .05, \omega = .9, \tau = 4, \Theta_l = 0, \Theta_r = 2.0$

Compare AA(m) for several m with Newton-GMRES

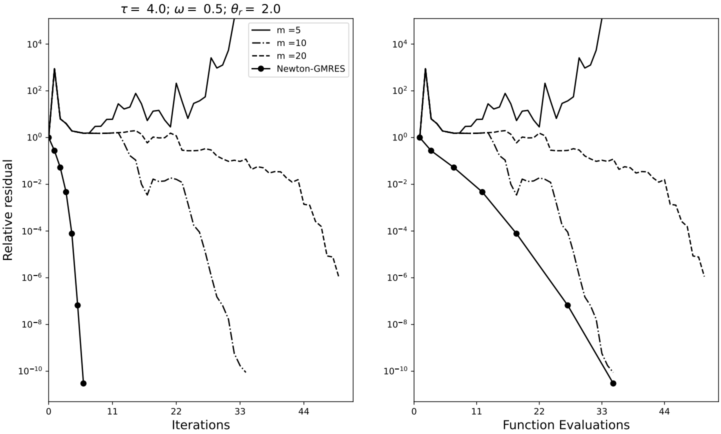
Easy Problem



Less Easy Problem: not contractive



Hard Problem: really not contractive



Summary

- New book: Codes + Examples in Julia
- Storage: Needs thought
- Multi-Physics Example:
  Conductive-Radiative Heat Transport
  AA is not always the thing to do.