Pseudo-Transient Continuation

C. T. Kelley
NC State University
tim_kelley@ncsu.edu
Joint with Liqun Qi, Li-Zhi Liao, Moody Chu, Corey Winton

ORNL, November 2008

Outline

Newton's method

Problem: solve $F(u) = 0$
$F : R^N \rightarrow R^N$ is Lipschitz continuously differentiable.
Newton's method

$$u_+ = u_c + s.$$

The step is

$$s = -F'(u_c)^{-1} F(u_c)$$

$F'(u_c)$ is the Jacobian matrix

Implementation

Inexact formulation:

$$\|F'(u_c)s + F(u_c)\| \le \eta_c \|F(u_c)\|.$$

$\eta = 0$ for direct solvers + analytic Jacobians.
If $F(u^*) = 0$, $F'(u^*)$ is nonsingular, and u_c is close to u^*

$$\|u_+ - u^*\| = O(\eta_c \|u_c - u^*\| + \|u_c - u^*\|^2)$$

But what if u_0 is far from u^*?

Armijo Rule: Find the least integer $m \geq 0$ such that

$$\|F(u_c + 2^{-m}s)\| \leq (1 - \alpha 2^{-m})\|F(u_c)\|$$

- $m = 0$ is Newton's method.
- Make it fancy by replacing 2^{-m}.
- $\alpha = 10^{-4}$ is standard.

Theory

If F is smooth and you get s with a direct solve or GMRES then either

- **BAD:** the iteration is unbounded, i. e. $\limsup \|u_n\| = \infty$,
- **BAD:** the derivatives tend to singularity, i. e. $\limsup \|F'(u_n)^{-1}\| = \infty$, or
- **GOOD:** the iteration converges to a solution u^* in the terminal phase, $m = 0$, and

$$\|u_{n+1} - u^*\| = O(\eta_n \|u_n - u^*\| + \|u_n - u^*\|^2).$$

Bottom line: you get an answer or an easy-to-detect failure.

Why worry?

- Stagnation at singularity of F' really happens.
  - steady flow $\longrightarrow$ shocks in CFD
- Non-physical results
  - fires go out
  - negative concentrations
- Nonsmooth nonlinearities
  - are not uncommon: flux limiters, constitutive laws
  - globalization is harder
  - finite diff directional derivatives may be wrong

Ψtc is one way to fix some of these things.

Steady-state Solutioins

Think about a PDE

$$\frac{du}{dt} = -F(u), u(0) = u_0,$$

and its solution $u(t)$.
$F(u)$ contains

- the nonlinearity,
- boundary conditions, and
- spatial derivatives.

We want the steady-state solution: $u^* = \lim_{t \to \infty} u(t)$.

What can go wrong?

If u_0 is separated from u^* by

- complex features like shocks,
- stiff transient behavior, or
- unstable equlibria,

the Newton-Armijo iteration can

- stagnate at a singular Jacobian, or
- find a solution of $F(u) = 0$ that is not the one you want.

A Questionable Idea

One way to guarantee that you get u^* is

- Find a high-quality temporal integration code.
- Set the error tolerances to very small values.
- Integrate the PDE to steady state.
  - Continue in time until $u(t)$ isn't changing much.
- Then apply Newton to make sure you have it right.

Problem: you may not live to see the results.

Ψtc

Integrate

$$\frac{du}{dt} = -F(u)$$

to steady state in a stable way with increasing time steps.
Equation for Ψtc Newton step:

$$\left(\delta_c^{-1} I + F'(u_c)\right) s = -F(u_c),$$

or

$$\| \left(\delta_c^{-1} I + F'(u_c)\right) s + F(u_c)\| \leq \eta_c \|F(u_c)\|.$$

Ψtc as an Integrator

Implicit Euler for $y' = -F(y)$

$$u_{n+1} = u_n + \delta F(u_{n+1})$$

u_{n+1} is the solution of

$$G(u) = u - u_n + \delta F(u) = 0.$$

Since $G'(u) = I + \delta F'(u)$, a single Newton iterate from $u_c = u_n$ is

$$\begin{aligned} u_+ &= u_c - (I + \delta F'(u_c))^{-1}(u_c - u_n + \delta F(u_c)) \\ &= u_c - (\delta^{-1} I + F'(u_c))^{-1} F(u_c), \end{aligned}$$

since $u_c - u_n = 0$.

Ψtc as an Integrator

- Low accuracy PECE integration
  - Trivial predictor
  - Backward Euler corrector + one Newton iteration
  - 1st order Rosenbrock method
    High order possible, Luo, K, Liao, Tam 06
- Begin with small "time step" δ. Resolve transients.
- Grow the "time step" near u^*. Turn into Newton.

Time Step Control

Grow the time step with switched evolution relaxation (SER)

$$\delta_n = \min(\delta_0 \|F(u_0)\| / \|F(u_n)\|, \delta_{max}).$$

If $\delta_{max} = \infty$ then $\delta_n = \delta_{n-1} \|F(u_{n-1})\| / \|F(u_n)\|$.
Alternative with no theory (SER-B):

$$\delta_n = \delta_{n-1} / \|u_n - u_{n-1}\|$$

Temporal Truncation Error (TTE)

Estimate local truncation error by

$$\tau = \frac{\delta_n^2 (u)_i''(t_n)}{2}$$

and approximate $(u)_i''$ by

$$\frac{2}{\delta_{n-1} + \delta_{n-2}} \left[\frac{((u)_i)_n - ((u)_i)_{n-1}}{\delta_{n-1}} - \frac{((u)_i)_{n-1} - ((u)_i)_{n-2}}{\delta_{n-2}} \right]$$

Adjust step so that $\tau = .75$.

Constraints

$$\frac{du}{dt} = -F(u), u(0) = u_0 \in \Omega.$$

$u(t) \in \Omega$, $F(u) \in \mathcal{T}(u)$ (tangent to Ω).
Examples:

- Ω has interior: bound constrained optimization
- Ω smooth manifold: inverse eigen/singular value problems

Problem: Ψtc will drift away from Ω.

Projected Ψtc

$$u_+ = \mathcal{P}(u_c - \left(\delta_c^{-1} I + H(u_c)\right)^{-1} F(u_c))$$

where

- $\mathcal{P}$ is map-to-nearest $R^N \rightarrow \Omega$
  $\|\mathcal{P}'(u)\| = 1$ for $u \in \Omega$.
- $H(u_c)$ makes Newton-like method fast.

General Method

Liao-Qi-K, 2006

F Lipschitz (no smoothness assumptions)

$$u_+ = \mathcal{P}(u_c - (\delta^{-1} I + H(u_c))^{-1} F(u_c)),$$

where H is an approximate Jacobian.

Theory: H bounded, other assumptions imply $u_n \to u^*$ and

$$u_{n+1} = u_{n+1}^N + O(\delta_n^{-1} + \eta_n) \| u_n - u^* \|$$

where

$$u_{n+1}^N = u_n - H(u_n)^{-1} F(u_n)$$

which is as fast as the underlying method.

What are those other assumptions?

- $u(t) \to u^*$
- δ_0 is sufficiently small.
- $\|\mathcal{P}'(u)\| = 1$ or Lip const of $\mathcal{P} = 1$
- u^* is dynamically stable
- $H(u)$ is uniformly well-conditioned near $\{u(t) \,|\, t \geq 0\}$
- $u_+ = u_c - H(u_c)^{-1} F(u_c)$ is rapidly locally convergent near u^*

A word about dynamics

$$\frac{du}{dt} = -F(u), u(0) = u_0$$

implies $u(t) \to u^*$ if $F = \nabla f$ and

- f is real analytic,
- the Lojasiewicz inequality

$$\|\nabla f(u)\| \geq c|f(u) - f(u^*)|$$

  holds, or
- f has bounded level sets and finitely many critical points.

But none of this implies that u^* is dynamically stable.

Fixing TTE and SER-B

If the underlying problem is minimization of f and ...

- you reduce δ until f is reduced,
- δ_0 is sufficiently small, and
- u^* is the unique root of F.

Then either $\delta_n \to 0$ or you converge to u^*.

Example

$$-u_{zz} + \lambda \max(0, u)^p = 0$$

$$z \in (0, 1), u(0) = u(1) = 0,$$

where $p \in (0, 1)$.
Reformulate as a DAE to make the nonlinearity Lipschitz.
Let

$$v = \begin{cases} u^p & \text{if } u \geq 0 \\ u & \text{if } u < 0 \end{cases}$$

Reformulation

Set $x = (u, v)^T$ and solve

$$F(x) = \begin{pmatrix} f(u,v) \\ g(u,v) \end{pmatrix} = \begin{pmatrix} -u_{zz} + \lambda \max(0, v) \\ u - \omega(v) \end{pmatrix} = 0,$$

The nonlinearity is

$$\omega(v) = \begin{cases} v^{1/p} & \text{if } v \geq 0 \\ v & \text{if } v < 0 \end{cases}$$

DAE Dynamics

$$D \begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} u' \\ 0 \end{pmatrix}$$

$$= - \begin{pmatrix} f(u,v) \\ g(u,v) \end{pmatrix} = -F(x), \quad x(0) = x_0,$$

Why not ODE dynamics?

Original time-dependent problem is

$$u_t = u_{zz} - \lambda \max(0, u)^p.$$

Applying Ψtc to

$$v_t = u - \omega(v)$$

rather than using $u - \omega(v) = 0$ as an algebraic constraint

- adds non-physical time dependence,
- changes the problem, and
- doesn't work.

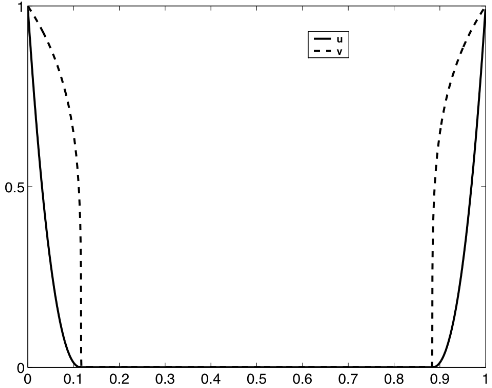
Parameters

- $p = .1$ and $\lambda = 200$. Leads to "dead core".
- $\delta_0 = 1.0$, $\delta_{max} = 10^6$.
- Spatial mesh size $\delta_z = 1/2048$; discrete Laplacian L_{δ_z}
- Terminate nonlinear iteration when either

$$\|F(x_n)\|/\|F(x_0)\| < 10^{-13} \text{ or } \|s_n\| < 10^{-10}.$$

Step is an accurate estimate of error (semismoothness).

Solution



Analytic ∂F

$$F(x) = \begin{pmatrix} f(u,v) \\ g(u,v) \end{pmatrix}$$

$$= \begin{pmatrix} -L_{\delta_z} u \\ u - v - max(0, v^{1/p}) \end{pmatrix} + \begin{pmatrix} \lambda \\ 1 \end{pmatrix} \max(0, v).$$

Since

$$\partial \max(0,v) = \begin{cases} 0, & \text{if } v < 0 \\ [0,1], & \text{if } v = 0 \\ 1, & \text{if } v > 0, \end{cases}$$

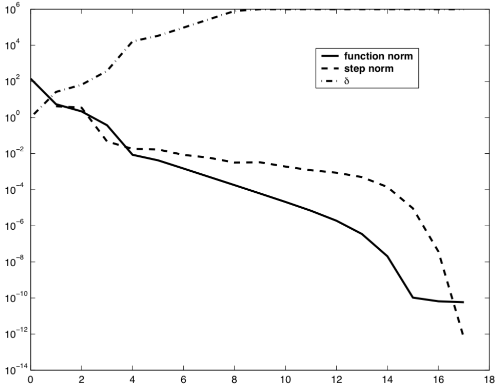
we get ...

∂F

$$\partial F = \begin{pmatrix} -L_{\delta_z} & 0 \\ 1 & -1 - (1/p)\max(0, v^{(1-p)/p}) \end{pmatrix}$$

$$+ \begin{pmatrix} 0 & \lambda \\ 0 & 1 \end{pmatrix} \partial \max(0, v).$$

Convergence



Linear Algebra Problem

Chu, 92 ...

Find $c \in R^N$ so that the $M \times N$ matrix

$$B(c) = B_0 + \sum_{k=1}^{N} c_k B_k$$

has prescribed singular values $\{\sigma_i\}_{i=1}^N$.
Data: Frobenius orthogonal $\{B_i\}_{i=1}^N$, $\{\sigma_i\}_{i=1}^N$.

Formulation

Least squares problem

$$\min F(U,V) \equiv \|R(U,V)\|_F^2$$

where

$$R(U,V) = U\Sigma V^T - B_0 - \sum_{k=1}^{N} < U\Sigma V^T, B_k >_F B_k$$

Manifold constraints: U is orthogonal $M \times M$ and V is orthogonal $N \times N$

Dynamic Formulation

$$\Omega = \{ \begin{pmatrix} U \\ V \end{pmatrix} \in R^{M \times M} \oplus R^{N \times N} \mid U \text{ and } V \text{ orthogonal} \}$$

Projected gradinet:

$$g(U,V) = \frac{1}{2} \begin{pmatrix} (R(U,V)V\Sigma^T U^T - U\Sigma V^T R(U,V)^T)U \\ (R(U,V)^T U\Sigma V^T - V\Sigma^T U^T R(U,V))V \end{pmatrix}.$$

ODE:

$$\dot{u} = \begin{pmatrix} \dot{U} \\ \dot{V} \end{pmatrix} = -F(u) \equiv -g(U,V).$$

Projection onto Ω

Higham 86, 04

Projection of square matrix onto orthogonal matrices

$$A \to U_P.$$

where $A = U_P H_P$ is the polar decomposition.
Compute U_P via the SVD $A = U \Sigma V^T$

$$U_P = U V^T.$$

Projection of

$$w = \begin{pmatrix} A \\ B \end{pmatrix}$$

onto Ω is

$$\mathcal{P}(w) = \begin{pmatrix} U_P^A \\ U_P^B \end{pmatrix}.$$

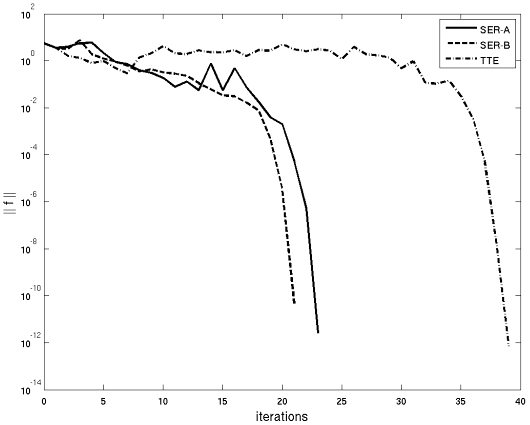
The local method

Given $u \in \Omega$ let $P_T(u) = \mathcal{P}'(u)$ be the projection onto the tangent space to Ω at u. Let

$$H = (I - P_T(u)) + P_T(u)F'(u)P_T(u)$$

Locally (very locally) superlinearly convergent if Ω is OK near u^*.

Inverse Singular Value Problem



Conclusions

- Ψtc computes steady-state solutions.
- Works on some manifolds.
  - Can succeed when traditional methods fail.
  - It is not a general nonlinear solver!
- Theory and practice for many problems
  - ODEs, DAEs
  - Nonsmooth F
  - Inverse eigen/singular value problems.