# Numerical Optimization

Lecture Notes #27

Nonlinear Equations — Continuation / Homotopy Methods

Peter Blomgren,
⟨blomgren.peter@gmail.com⟩

Department of Mathematics and Statistics
Dynamical Systems Group
Computational Sciences Research Center
San Diego State University
San Diego, CA 92182-7720

**http://terminus.sdsu.edu/**

Fall 2018





## Outline







## Continuation / Homotopy Methods — The Sales Pitch

**The problem with Newton's method**: Unless $J(\bar{\mathbf{x}})$ is non-singular in the region of interest — something that is very hard to guarantee *a priori* — it may converge to a local minimum of the merit function which does not correspond to a solution of the nonlinear system.

**Continuation methods** go directly for a solution of $\bar{\mathbf{r}}(\bar{\mathbf{x}}) = 0$ and are more likely to converge to such a solution in difficult cases.

**The Idea:** First, solve an "easy" problem where the solution is obvious. Then transform the easy system into the original system $\bar{\mathbf{r}}(\bar{\mathbf{x}}) = 0$, and track the solution as it moves from the easy problem to the full problem.





## The Homotopy Map

We define the **homotopy map**

$$\bar{\mathbf{H}}(\bar{\mathbf{x}}, \lambda) = \lambda \bar{\mathbf{r}}(\bar{\mathbf{x}}) + (1-\lambda)(\bar{\mathbf{x}} - \bar{\mathbf{a}}),$$

and note that

$$\bar{\mathbf{H}}(\bar{\mathbf{x}}, 0) = \bar{\mathbf{x}} - \bar{\mathbf{a}}, \qquad \bar{\mathbf{H}}(\bar{\mathbf{x}}, 1) = \bar{\mathbf{r}}(\bar{\mathbf{x}}).$$

Now, solving $\bar{\mathbf{H}}(\bar{\mathbf{x}}, \lambda) = 0$ is trivial when $\lambda = 0$, the solution is $\bar{\mathbf{x}}_0^* = \bar{\mathbf{a}}$.

**The idea:** If we increase $\lambda$ by "a little," then the root(s) of the equation only move "a little," hence they should be easy to find.

The path from $(\bar{\mathbf{x}}_0^*, \lambda = 0)$ to $(\bar{\mathbf{x}}_1^*, \lambda = 1)$ is known as the **zero path**, it connects the trivial solution to the solution of $\bar{\mathbf{r}}(\bar{\mathbf{x}}) = 0$.

## A Simple Zero Path

A naive implementation of the homotopy method works if there is a **unique** solution $\bar{\mathbf{H}}(\bar{\mathbf{x}}^*_\lambda, \lambda) = 0$, $\forall \lambda \in [0, 1]$, and we get a fairly simple zero path:

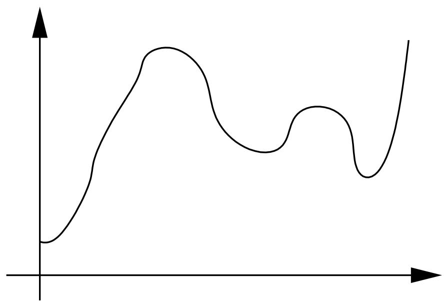

**Figure:** The **zero path** — the trajectory of points $(\bar{\mathbf{x}}, \lambda)$ for which $H(\bar{\mathbf{x}}, \lambda) = 0$. Here, $\lambda$ is increasing from 0 to 1 on the horizontal axis, and the values of $\bar{\mathbf{x}}_k$ are illustrated on the vertical axis.

SAN DIEGO STATE UNIVERSITY





## Problems... ...and Solutions

Unfortunately, we immediately run into trouble when there is more than one root of $H(\bar{\mathbf{x}}, \lambda) = 0$ for some range of $\lambda$.

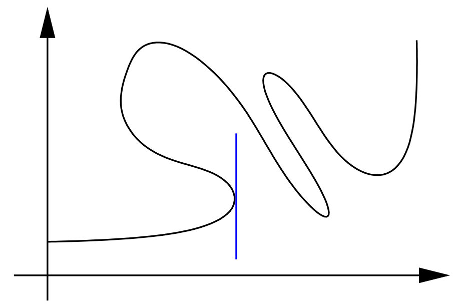

**Figure:** Here the zero path connects $(\bar{\mathbf{x}}_0, 0)$ and $(\bar{\mathbf{x}}_1, 1)$, but if we try to follow the path by monotonically increasing $\lambda$, we will fail at the first **turning point**.

In practical continuation methods, we must allow $\lambda$ to decrease, and sometimes even leave the interval $[0, 1]$...

SAN DIEGO STATE UNIVERSITY





## Practical Continuation Methods 1 of 4

We can model the zero path by letting both $\bar{\mathbf{x}}$ and $\lambda$ be functions of an independent variable $s$, that represents the **arc length** along the path.

| | |
|---|---|
| **Initial point** | $(\bar{\mathbf{x}}(0), \lambda(0)) = (a, 0)$ |
| **Zero path** | $\bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) = 0, \ \forall s \geq 0.$ |
| **Terminal point** | $(\bar{\mathbf{x}}(s^{\text{stop}}), \lambda(s^{\text{stop}})) = (\bar{\mathbf{x}}^*, 1), \quad \bar{\mathbf{r}}(\bar{\mathbf{x}}^*) = 0$ |

We differentiate the zero path condition with respect to $s$, and get

$$\frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \bar{\mathbf{x}}_s + \frac{\partial}{\partial \lambda} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \lambda_s = 0,$$

where $\bar{\mathbf{x}}_s = \frac{d\bar{\mathbf{x}}}{ds}$, and $\lambda_s = \frac{d\lambda}{ds}$.

$\frac{\partial}{\partial \bar{\mathbf{x}}}$ should be interpreted as the gradient operator in the $\bar{\mathbf{x}}$-coordinates, $\nabla_{\bar{\mathbf{x}}}$.

SAN DIEGO STATE UNIVERSITY





## Practical Continuation Methods 2 of 4

The vector $(\bar{\mathbf{x}}_s, \lambda_s)$ is the **tangent vector** to the zero path, and it lies in the **null space** of the $n \times (n+1)$ matrix

$$\left[ \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \quad \frac{\partial}{\partial \lambda} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \right] \qquad (1)$$

since

$$\left[ \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \,\middle|\, \frac{\partial}{\partial \lambda} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \right] \left[ \frac{\bar{\mathbf{x}}_s}{\lambda_s} \right] = 0$$

If this matrix has full rank, its null space has dimension 1. In order to complete the definition of $(\bar{\mathbf{x}}_s, \lambda_s)$, we normalize its length so that

$$\|\bar{\mathbf{x}}_s\|^2 + |\lambda_s|^2 = 1, \quad \forall s$$

this ensures that $s$ is the true arc length along the path.

SAN DIEGO STATE UNIVERSITY

## Practical Continuation Methods



We must now choose the sign (direction) of the tangent vector, so that it points "forward" along the zero path.

Usually, as long as we restrict the change in tangent direction by $\pi/2$ between iterations, we are safe.

The following algorithm identifies the tangent direction as described above:

Compute a vector in the null space by QR-factorization with column pivoting

$$Q^T \left[ \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \quad \frac{\partial}{\partial \lambda} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \right] \Pi = \left[ \begin{array}{cc} R & \bar{\mathbf{w}} \end{array} \right]$$

where $Q \in \mathbb{R}^{n \times n}$ orthogonal, $R \in \mathbb{R}^{n \times n}$ upper triangular, $\Pi \in \mathbb{R}^{(n+1)\times(n+1)}$ permutation matrix, and $\bar{\mathbf{w}} \in \mathbb{R}^n$. **Set**

$$\bar{\mathbf{v}} = \Pi \left[ \begin{array}{c} R^{-1}\bar{\mathbf{w}} \\ -1 \end{array} \right], \qquad (\bar{\mathbf{x}}_s, \lambda_s) = \pm \frac{\bar{\mathbf{v}}}{\|\bar{\mathbf{v}}\|} \quad \text{(sign as above).}$$





## Practical Continuation Methods



Now that we have an initial condition

$$(\bar{\mathbf{x}}, \lambda)(0) = (\bar{\mathbf{a}}, 0)$$

and the tangent vector (**think "Ordinary Differential Equation!!!"** [MATH 542])

$$\frac{d}{ds}(\bar{\mathbf{x}}, \lambda)(s) = (\bar{\mathbf{x}}_s, \lambda_s).$$

we can apply any off-the shelf method for solving this ODE (*e.g.* Matlab's `ode23` or `ode45`), and stop the solution when $\lambda(s) = 1$.





## Example #1 — Continuation for $f(x) = \sin(5x) - x$

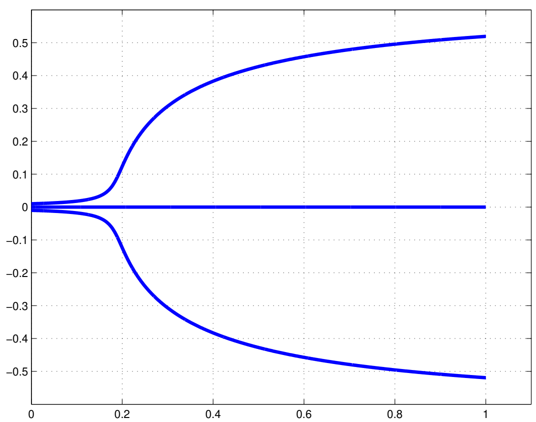


**Figure:** We apply the continuation scheme described to the scalar objective $f(x) = \sin(5x) - x$, *i.e.* we use the **homotopy map** $H^{(a)}(x, \lambda) = \lambda(\sin(5x) - x) + (1 - \lambda)(x - a)$. Depending on the starting value $a$ we get convergence to one of the three roots: $a = 0 \rightsquigarrow x^* = 0$, $a > 0 \rightsquigarrow x^* \approx 0.52$, and $a < 0 \rightsquigarrow x^* \approx -0.52$.





## Example #1b: Several Paths for $f(x) = x^2 - 1$

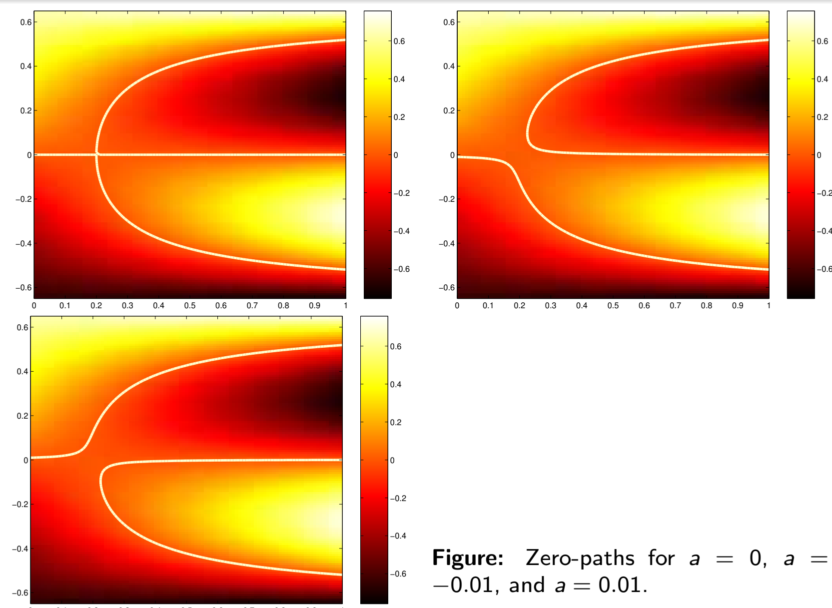

**Figure:** Zero-paths for $a = 0$, $a = -0.01$, and $a = 0.01$.

## Continuation Methods: Robustness



The continuation method we have described relies on the $n \times (n+1)$ matrix (1) on slide#8, having full rank for all $(\bar{\mathbf{x}}, \lambda)$ along the path, so that the tangent vector is well defined.

Theorem

*Suppose that r is twice continuously differentiable. Then for* **almost all** *vectors* $\bar{\mathbf{a}} \in \mathbb{R}^n$*, there is a zero path emanating from* $(\bar{\mathbf{a}}, 0)$ *along which the* $n \times (n+1)$ *matrix*

$$\left[ \begin{array}{cc} \frac{\partial}{\partial \bar{\mathbf{x}}} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) & \frac{\partial}{\partial \lambda} \bar{\mathbf{H}}(\bar{\mathbf{x}}(s), \lambda(s)) \end{array} \right]$$

*has full rank. If this path is bounded for* $\lambda \in [0,1)$*, then it has an accumulation point* $(\bar{\mathbf{x}}^*, 1)$ *such that* $\bar{\mathbf{r}}(\bar{\mathbf{x}}^*) = 0$*. Furthermore, if the Jacobian* $J(\bar{\mathbf{x}}^*)$ *is non-singular, the zero path between* $(\bar{\mathbf{a}}, 0)$ *and* $(\bar{\mathbf{x}}^*, 1)$ *has finite arc length.*





## Continuation Methods: Robustness



The theorem shows that unless we are "very unlucky" in our choice of $\bar{\mathbf{a}}$, our continuation algorithms will be well defined, and will either converge to a point $\bar{\mathbf{x}}^*$ that is a solution $\bar{\mathbf{r}}(\bar{\mathbf{x}}^*) = 0$, or will diverge.

**Example #2: Unbounded Path for $f(x) = x^2 - 1$.**

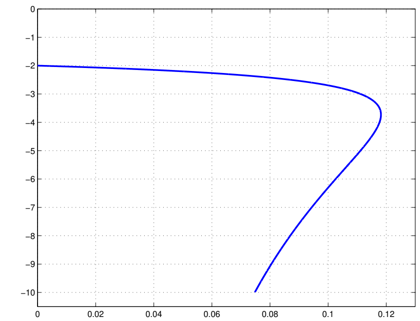


**Figure:** There is no zero path connecting $(-2, 0)$ and either non-degenerate root $(\pm 1, 1)$, hence the continuation method fails (the path continues down to $-\infty$).





## Analysis of Example #2

The homotopy map for $a = -2$ is given by

$$H(x, \lambda) = \lambda(x^2 - 1) + (1-\lambda)(x+2) = \lambda x^2 + (1-\lambda)x + (2 - 3\lambda).$$

For a fixed $\lambda$, the roots of $H(x, \lambda)$ are given by

$$x = \frac{-(1-\lambda) \pm \sqrt{(1-\lambda)^2 - 4\lambda(2-3\lambda)}}{2\lambda}.$$

If (when) the term inside the square root is negative, there are no real roots. This occurs in the range

$$\lambda \in \left( \frac{5 - 2\sqrt{3}}{13}, \frac{5 + 2\sqrt{3}}{13} \right) \approx (0.118,\ 0.651).$$





## Example #2b: Bounded Path for $f(x) = x^2 - 1$

Changing the starting point to $a = -0.1$ yields the following path

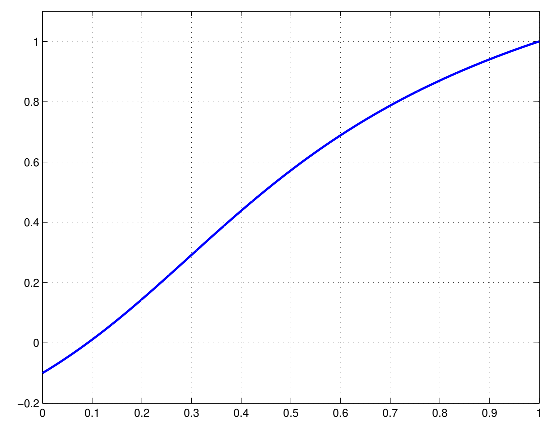


**Figure:** In this case there is a zero path connecting $(-0.1, 0)$ and the non-degenerate root $(1, 1)$.

## Example #2c: Several Paths for $f(x) = x^2 - 1$

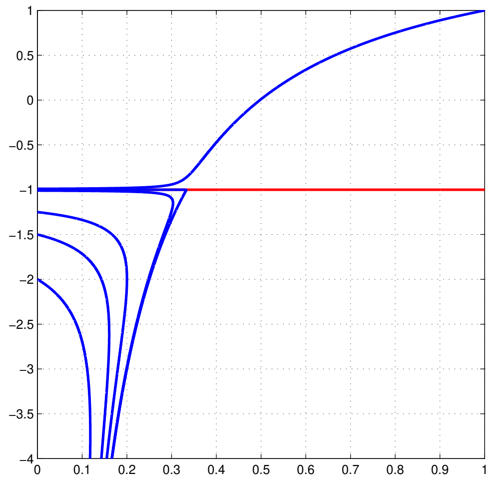


**Figure:** The path $(-1, 0) \rightsquigarrow (-1, 1)$ is the only path to the negative root; for starting points $(a < -1, 0)$ the path becomes unbounded, and $(a > -1, 0) \rightsquigarrow (1, 1)$.





## Example #2d: Several Paths for $f(x) = x^2 - 1$

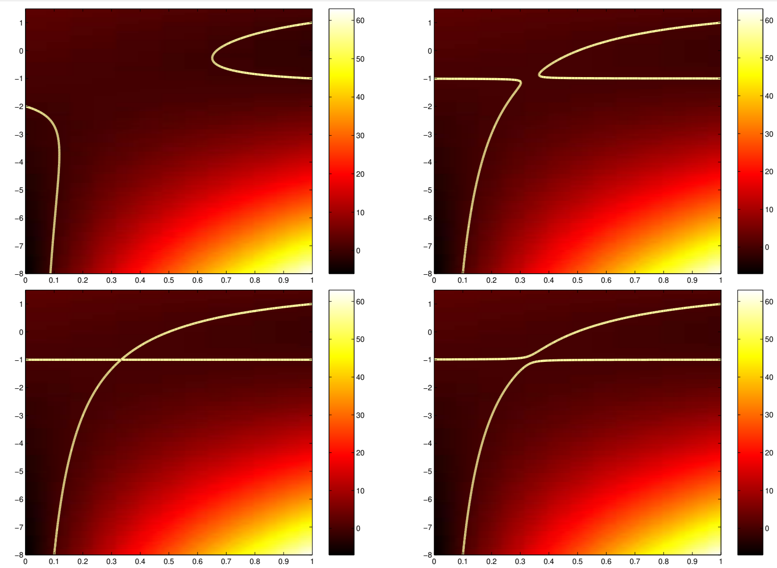

**Figure:** Zero-paths for $a = -2$, $a = -1.01$, $a = -1$, and $a = -0.99$.





## Continuation Methods: Pros and Cons

### The Good

- `[claim]` Generally more reliable than merit-function methods.

### The Bad

- High cost — continuation methods require significantly more function and derivative evaluations, and linear algebra operations than merit-function based methods.





## Index