The Pseudo-Arclength Method For Turning Points

Stefan Ragnarsson

May 15, 2007

1 Outline

In lecture, we briefly discussed the Newton method for numerical continuation in the absence of singular points (bifurcations, turning points etc.) and mentioned that switching to an arclength parametrization would help continue past turning points. Here we aim to complement the lecture notes by giving a rigorous treatment of the pseudo-arclength method as well as details on actual algorithmic implementations.

2 The Newton method for regular points

The general problem is

$$F(\lambda, u) = 0 \tag{1}$$

where $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. We assume that we have a know solution point (λ_0, u_0) of equation (1). The idea of local continuation is to find a solution at $\lambda_0 + \Delta\lambda$ for a suitably small perturbation $\Delta\lambda$.
We start by using an *Euler prediction*

$$\begin{aligned} (u_0')^0 &= u'(\lambda_0) = -(F_u^0)^{-1} F_\lambda^0 \\ u_1^0 &= u_0 + \Delta\lambda (u_0')^0. \end{aligned}$$

We then applied the Newton method:

$$\begin{aligned} \Delta u_1^k = -(F_u(\lambda_1, u_1^{k-1}))^{-1} F(\lambda_1, u_1^{k-1}), \\ u_1^k = u_1^{k-1} + \Delta u_1^k, \end{aligned}$$

for $k = 1, 2, \ldots$ until convergence. One important advantage of Newton's method is that it can have quadratic convergence. Note that this method relies crucially on F_u being nonsingular.

3 The Pseudo-Arclength Method

Consider the problem

$$F(\lambda, u) = 0,$$

where $F : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$.

Definition: A path

$$\gamma = \{(\lambda(s), u(s)) : u(s) \in \Omega \subseteq \mathbb{R}^n, F(\lambda(s), u(s)) = 0 \text{ for all } s \in I\}$$

is said to be *regular* if the rank of the matrix $[F_\lambda(s), F_u(s)]$ is n for all $s \in I$.

Lemma 1: Rank $[F_\lambda, F_u] = n$ if and only if either

i) F_u is nonsingular,

ii) a) $\dim \mathcal{N}(F_u) = 1$,
  b) $F_\lambda \notin \mathcal{R}(F_u)$.

Definition: A point $(\lambda_0, u_0) = (\lambda(s_0), u(s_0))$ on γ is said to be a *simple limit point* or a *simple fold* if (ii, a, b) hold.

On the path γ we have

$$F(\lambda(s), u(s)) = 0,$$

so that

$$F_\lambda(s)\lambda'(s) + F_u(s)u'(s) = 0. \tag{2}$$

At a fold point $s = s_0$ we have that $\lambda'(s_0) = 0$ since

$$F_\lambda(s_0) \notin \mathcal{R}(F_u(s_0)).$$

Therefore $u'(s_0) \in \mathcal{N}(F_u(s_0))$ at a fold point. Since $\dim \mathcal{N}(F_u^0) = 1$ at a simple fold point (λ_0, u_0), we can take

$$\mathcal{N}(F_u^0) = \text{span}\{\phi\}, \tag{3}$$

and

$$\mathcal{N}((F_u^0)^T) = \text{span}\{\psi\}. \tag{4}$$

etc.

Usual methods of computing paths fail at a fold point. The main idea in pseudo-arclength continuation is to drop the natural parametrization by λ and use some other parametrization.

If (λ_0, u_0) is any point on a regular path and (λ_0', u_0') is the unit tangent to the path, then we ajoin to (1) the scalar normalization

$$N(\lambda, u, \Delta s) := (u_0')^T(u - u_0) + \lambda_0'(\lambda - \lambda_0) - \Delta s = 0. \tag{5}$$

This is the equation of a plane, which is perpendicular to the tangent (λ_0', u_0') at a distance Δs from (λ_0, u_0). This plane will intersect the curve γ if Δs and the curvature of γ are not too large. That is we solve (1) and (5) simultaneously for $(\lambda(s), u(s))$. Using Newton's method on this leads to the linear system

$$\begin{bmatrix} F_\lambda^k & F_u^k \\ \lambda_0' & u_0'^T \end{bmatrix} \begin{bmatrix} \Delta\lambda^k \\ \Delta u^k \end{bmatrix} = -\begin{bmatrix} F^k \\ N^k \end{bmatrix}, \tag{6}$$

where $F_u^k = F_u(\lambda^k, u^k), F_\lambda^k = F_\lambda(\lambda^k, u^k)$, and the iterates are $u^{k+1} = u^k + \Delta u^k$ and $\lambda^{k+1} = \lambda^k + \Delta\lambda^k$.

Lets derive practical procedures for solving this system on a regular path. On such a path F_u^0 may be nonsingular or singular but the $(n+1)$ order coefficent matrix should be nonsingular. One proof of the nonsingularity of the coefficient matrices in (6) can be based on the following result:

Lemma 2: Let $\mathcal{B}$ be a banach space and let the linear operator $\tilde{A} : \mathcal{B} \times \mathbb{R}^k \to \mathcal{B} \times \mathbb{R}^k$ have the form

$$\tilde{A} = \begin{bmatrix} A & B \\ C^* & D \end{bmatrix} \tag{7}$$

where $A : \mathcal{B} \to \mathcal{B}$, $B : \mathbb{R}^k \to \mathcal{B}$, $C^* : \mathcal{B} \to \mathbb{R}^k$ and $D : \mathbb{R}^k \to \mathbb{R}^k$.

i) If A is nonsingular the $\tilde{A}$ is nonsingular if and only if

$$D - C^* A^{-1} B$$

is nonsingular.

ii) If A is singular with

$$\dim \mathcal{N}(A) = \text{codim } \mathcal{R}(A) = k,$$

then $\tilde{A}$ is nonsingular if and only if

$$\dim \mathcal{R}(B) = k, \tag{8}$$
$$\mathcal{R}(B) \cap \mathcal{R}(A) = 0, \tag{9}$$
$$\dim \mathcal{R}(C^*) = k, \tag{10}$$
$$\mathcal{N}(A) \cap \mathcal{N}(C^*) = 0. \tag{11}$$

iii) If A is singular with $\dim \mathcal{N}(A) > k$ then $\tilde{A}$ is singular.

In our analysis we will only consider the case $k = 1$ and $\mathcal{B} = \mathbb{R}^n$. Then the conditions in the lemma reduce to

$$B \notin \mathcal{R}(A) \text{ and } C^T \notin \mathcal{R}(A^T), \tag{12}$$

where A^T is the transpose of the matrix A.

Note: Instead of using the normalization (5) it is possible to use other relations. One generalization is

$$N = \theta(u_0')^T(u - u_0) + (1-\theta)\lambda_0'(\lambda - \lambda_0) - \Delta s = 0$$

for $0 < \theta < 1$. Alternatively we can employ a secant, rather than a tangent. If we know two nearby points on γ, say at $s = s_0$ and $s = s_{-1}$ then we can use

$$N(\lambda, u, s) = \theta \left[\frac{u(s_0) - u(s_{-1})}{s_0 - s_{-1}}\right]^T \Delta u + (1-\theta)\left[\frac{\lambda'(s_0) - \lambda'(s_{-1})}{s0 - s_{-1}}\right]\Delta\lambda - \Delta s = 0,$$

where $\Delta u = u(s) - u(s_0), \Delta\lambda = \lambda(s) - \lambda(s_0)$ and $\Delta s = s - s_0$. All of the above are called *pseudo-arclength normalizations.*

3.1 Bordering Algorithm

We write the coefficient matrix of (6) in the form

$$\tilde{A} = \begin{bmatrix} A & b \\ c^T & d \end{bmatrix},$$

where A is an $n \times n$ matrix, $b, c \in \mathbb{R}^n$ and $d \in \mathbb{R}$. Then we consider the linear system

$$\tilde{A}\begin{bmatrix} x \\ \xi \end{bmatrix} = \begin{bmatrix} g \\ \eta \end{bmatrix}, \tag{13}$$

where $x, g \in \mathbb{R}^n$ and $\xi, \eta \in \mathbb{R}$.
Assume that $\tilde{A}$ and A are nonsingular. Then solve the linear systems

$$Ay = b, \tag{14}$$
$$Az = g. \tag{15}$$

The solution to (13) is then

$$\xi = \frac{\eta - c^T z}{d - c^T y},\ x = z - \xi y. \tag{16}$$

Note that $d - c^T y = d - c^T A^{-1} b$ is the *Schur complement* of A in $\tilde{A}$. Hence $d - c^T y \neq 0$ if $\tilde{A}$ is nonsingular by Lemma 2. Thus if $\tilde{A}$ and A are nonsingular then the Bordering

Algorithm is valid.
We could also write the system in (13) as

$$\begin{aligned} Ax + b\xi &= g, && (17) \\ c^T x + d\xi &= \eta. && (18) \end{aligned}$$

To solve this, we can proceed by first eliminating ξ, if $d \neq 0$, to get

$$\xi = \frac{1}{d}(\eta - c^T x). \tag{19}$$

Then for x we have

$$\left(A - \frac{1}{d} b c^T\right) x = g - \frac{\eta}{d} b. \tag{20}$$

Note that $(A - \frac{1}{d} bc^T)$ is a rank-1 modification of A. Hence we can use the Sherman-Morrison formula to easily determine the inverse of $(A - \frac{1}{d} bc^T)$ once we have the inverse of A.

Lemma 3: Let A be an $n \times n$ invertible matrix and $u, v \in \mathbb{R}^n$. Then $A + uv^T$ is nonsingular if and only if $\sigma = 1 + v^T A^{-1} u \neq 0$ and then the inverse of $A + uv^T$ is given by

$$(A + uv^T)^{-1} = A^{-1} - \frac{1}{\sigma} A^{-1} u v^T A^{-1}. \tag{21}$$

Note that this procedure requires that $d \neq 0$, while the bordering algorithm does not. Both require A to be nonsingular, however.

So what can we do when A is singular but $\tilde{A}$ is not? This occurs at simple limit points on the solution paths. Thus we assume

i) $\mathcal{N}(A) = \text{span}\{\phi\}$,

ii) $b \notin \mathcal{R}(A)$,

iii) $c^T \notin \mathcal{R}(A^T)$.

These are precisely the conditions in Lemma 2 for the nonsingularity of $\tilde{A}$ and are equivalent to

$$\psi^T b \neq 0 \text{ and } c^T \phi \neq 0, \tag{22}$$

where ϕ and ψ are nontrivial solutions of

$$A\phi = 0 \text{ and } A^T \psi = 0. \tag{23}$$

We can write the system (13) as

$$\tilde{A} \begin{bmatrix} x_0 \\ \xi_0 \end{bmatrix} = \begin{bmatrix} g \\ \eta \end{bmatrix}, \tag{24}$$

where $x_0, g \in \mathbb{R}^n$ and $\xi_0, \eta \in \mathbb{R}$. That is

$$\begin{aligned} Ax_0 + b\xi_0 &= g, \quad (25) \\ c^T x_0 + d\xi_0 &= \eta. \quad (26) \end{aligned}$$

Multiplying the first equation by ψ^T, we get

$$\xi_0 = \frac{\psi^T g}{\psi^T b}. \quad (27)$$

Hence

$$Ax_0 = g - \frac{\psi^T g}{\psi^T b} b \in \mathcal{R}(A). \quad (28)$$

All solutions x_0 of (28) have the form

$$x_0 = x_p + z\phi, \quad (29)$$

where x_p is any particular solution of (28) and z is obtained by substituting the value of x_0 into (26) to get

$$z = \frac{\eta - d\xi_0 - c^T x_p}{c^T \phi}. \quad (30)$$

Therefore

$$x_0 = \left(x_p - \frac{c^T x_p}{c^T \phi} \phi \right) + \left(\frac{\eta - d\xi_0}{c^T \phi} \right) \phi. \quad (31)$$

Hence the unique solution of (24) is given by (27) and (31). To evaluate this solution we need the vectors ϕ, ψ, x_p and the inner products $\psi^T g, \psi^T b, c^T \phi$ and $c^T x_p$.

The operational count to actually obtain these vectors is only one inner product more than the count required by the Bordering Algorithm. Thus the solution (x_0, ξ_0) requires only two inner products more.

Next, we will show how to compute ϕ and ψ efficiently.

3.2 Left and right null vectors of A

Assume that A_1 is an $n \times n$ matrix satisfying $(i) - (iii)$ p.5 so that $\text{rank}(A_1) = n - 1$. Then there exist permutation matrices P and Q such that the matrix

$$A = PA_1Q \quad (32)$$

has an LU factorization

$$A = LU = \begin{bmatrix} \hat{L} & 0 \\ \hat{l}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{U} & \hat{u} \\ 0 & 0 \end{bmatrix}. \quad (33)$$

Here $\hat{L}$ and $\hat{U}$ are lower and upper triangular matrices, respectively, of order $(n-1) \times (n-1)$ with

$$\hat{L} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \times & 1 & 0 & \cdots & 0 \\ \times & \times & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \times & \times & \cdots & \times & 1 \end{bmatrix}$$

and

$$\hat{U} = \begin{bmatrix} u_{11} & \times & \times & \cdots & \times \\ 0 & u_{22} & \times & \cdots & \times \\ 0 & 0 & u_{33} & \cdots & \times \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & u_{n-1,n-1} \end{bmatrix}$$

and $\hat{u}, \hat{l} \in \mathbb{R}^{n-1}$.

Since L is nonsingular, $A\phi = 0$ if and only if $U\phi = 0$. So with $\hat{\phi} \in \mathbb{R}^{n-1}$ and $\alpha \in \mathbb{R}$, we seek ϕ in the form

$$\phi = \alpha \begin{bmatrix} \hat{\phi} \\ -1 \end{bmatrix}, \alpha \neq 0. \tag{34}$$

It follows (since $\hat{U}$ is nonsingular) that $\hat{\phi}$ is uniquely determined by

$$\hat{U}\hat{\phi} = \hat{u}. \tag{35}$$

In other words,

$$\hat{\phi} = \hat{U}^{-1}\hat{u}. \tag{36}$$

Since U is triangular, we obtain $\hat{\phi}$ and hence ϕ for only half the cost of an actual back solve operation (solving a general $n-1$ system). Similarly, the nonsingularity of $\hat{U}$ implies that $A^T\psi = 0$ if and only if

$$L^T\psi = \beta \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \beta \in \mathbb{R}. \tag{37}$$

Thus we find that all nontrivial left null vectors are given by

$$\psi = \beta \begin{bmatrix} \hat{\psi} \\ -1 \end{bmatrix}, \beta \neq 0, \tag{38}$$

and $\hat{\psi} \in \mathbb{R}^{n-1}$ is uniquely determined by

$$\hat{L}^T\hat{\psi} = \hat{l}, \tag{39}$$

$$\hat{\psi} = (\hat{L}^T)^{-1}\hat{l}. \tag{40}$$

Again $\hat{\psi}$ and hence ψ are obtained with half of a back solve operation.

3.3 Almost singular A

In actual calculations we do not get exact zeros in the final diagonal position of U, but rather we get an inexact factorization

$$A = A_\epsilon = L_\epsilon U_\epsilon = \begin{bmatrix} \hat{L} & 0 \\ \hat{l}^T & 1 \end{bmatrix} \begin{bmatrix} \hat{U} & \hat{u} \\ 0 & \epsilon \end{bmatrix}, \tag{41}$$

where $\epsilon \ll 1$. If we use full pivoting to determine P and Q then under appropriate conditions on A_1 we can bound ϵ by $C \cdot 10^{-t}$ for t digit arithmetic where C is a constant.

We now apply the Bordering Algorithm to solve

$$\tilde{A} \begin{bmatrix} x \\ \xi \end{bmatrix} = \begin{bmatrix} g \\ \eta \end{bmatrix}, \tag{42}$$

for this inexact factorization. Thus we consider

$$A_\epsilon y_\epsilon = b, \; A_\epsilon z_\epsilon = g, \tag{43}$$

with

$$b = \begin{bmatrix} \hat{b} \\ b_n \end{bmatrix}, g = \begin{bmatrix} \hat{g} \\ g_n \end{bmatrix}. \tag{44}$$

Using our previous formulas we can easily see that

$$y_\epsilon = \begin{bmatrix} (\hat{L}\hat{U})^{-1}\hat{b} \\ 0 \end{bmatrix} + \frac{\psi^T b}{\epsilon}\phi, \tag{45}$$

$$z_\epsilon = \begin{bmatrix} (\hat{L}\hat{U})^{-1}\hat{g} \\ 0 \end{bmatrix} + \frac{\psi^T g}{\epsilon}\phi. \tag{46}$$

Using (16) we get

$$\xi_\epsilon = \frac{\eta - \hat{c}^T(\hat{L}\hat{U})^{-1}\hat{g} - \frac{1}{\epsilon}(\psi^T g)(c^T\phi)}{d - \hat{c}^T(\hat{L}\hat{U})^{-1}\hat{b} - \frac{1}{\epsilon}(\psi^T b)(c^T\phi)}, \tag{47}$$

$$x_\epsilon = \begin{bmatrix} (\hat{L}\hat{U})^{-1}(\hat{g} - \xi_\epsilon \hat{b}) \\ 0 \end{bmatrix} + \frac{1}{\epsilon}[(\psi^T g) - \xi_\epsilon(\psi^T b)]\phi. \tag{48}$$

If we expand this about $\epsilon = 0$ we obtain

$$\xi_\epsilon = \xi_0 + O(\epsilon), \tag{49}$$

$$x_\epsilon = x_0 + O(\epsilon). \tag{50}$$

In other words, our computed solution $(x_\epsilon^T, \xi_\epsilon)^T$ will be a close approximation to the solution $(x_0^T, \xi_0)^T$ (when $\epsilon = 0$).

3.4 The tangent vectors

Lets (briefly) consider how to compute the tangent vectors (λ_0', u_0'). They must satisfy

$$F_\lambda^0 \lambda_0' + F_u^0 u_0' = 0, \tag{51}$$
$$\|u_0'\|^2 + |\lambda_0'|^2 = 1. \tag{52}$$

First consider regular points, where F_u^0 is nonsingular. We find ϕ_0 from

$$F_u^0 \phi_0 = -F_\lambda^0. \tag{53}$$

Then set

$$u_0' = a\phi_0 \text{ and } \lambda_0' = a \tag{54}$$

where a is determined from (52) as

$$a = \frac{\pm 1}{\sqrt{1 + \|\phi_0\|^2}}. \tag{55}$$

The sign of a is determined so that the orientation of the path is preserved. More precisely, if (λ_{-1}', u_{-1}') is the preceding tangent vector then we require

$$(u_{-1}')^T u_0' + \lambda_{-1}' \lambda_0' > 0. \tag{56}$$

Choosing the sign of a is very important in numerical calculations. If we do not choose the sign properly, either we will get trapped in the iteration at some point or it will reverse the direction and hence it will compute the same path already computed!

At a simple fold F_u^0 is singular. Our analysis of the case of almost singular A shows that we get in this case for the solution of (53) by setting

$$\phi_0 = \begin{bmatrix} \hat{\phi}_0 \\ \omega_0 \end{bmatrix}, -F_\lambda^0 = \begin{bmatrix} \hat{g} \\ \eta \end{bmatrix}. \tag{57}$$

The result is

$$\hat{\phi}_0 = -(\hat{L}\hat{U})^{-1}\hat{g} + \frac{\eta - \hat{\psi}^T \hat{g}}{\epsilon}\hat{\phi}, \tag{58}$$
$$\omega_0 = -\frac{\eta - \hat{\psi}^T \hat{g}}{\epsilon}. \tag{59}$$

Using these results in (54) and (55) we find that we indeed get the tangent vector within $O(\epsilon)$.