2013

Solving Separable Nonlinear Equations Using LU Factorization

Yun-Qiu Shen  
*Western Washington University, yunqiu.shen@wwu.edu*

Tjalling Ypma  
*Western Washington University, tjalling ypma@wwu.edu*

Follow this and additional works at: https://cedar.wwu.edu/math_facpubs

Part of the Mathematics Commons

**Recommended Citation**  

This Article is brought to you for free and open access by the College of Science and Engineering at Western CEDAR. It has been accepted for inclusion in Mathematics by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.
Research Article

Solving Separable Nonlinear Equations Using LU Factorization

Yun-Qiu Shen and Tjalling J. Ypma

Department of Mathematics, Western Washington University, Bellingham, WA 98225-9063, USA

Correspondence should be addressed to Tjalling J. Ypma; tjalling.ypma@wwu.edu

Received 15 May 2013; Accepted 4 June 2013

Academic Editors: R. Barrio and B. Djafari Rouhani

Copyright © 2013 Y.-Q. Shen and T. J. Ypma. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Separable nonlinear equations have the form

\[ F(y, z) = A(y)z + b(y) = 0, \]

where the matrix \( A(y) \in \mathbb{R}^{m \times N} \) and the vector \( b(y) \in \mathbb{R}^{m} \) are continuously differentiable functions of \( y \in \mathbb{R}^{n} \) and \( z \in \mathbb{R}^{N} \). We assume that \( m \geq N + n \), and \( F^\prime(y, z) \) has full rank. We present a numerical method to compute the solution \((y^\ast, z^\ast)\) for fully determined systems \((m = N + n)\) and compatible overdetermined systems \((m > N + n)\). Our method reduces the original system to a smaller system \( f(y) = 0 \) of \( m - N \geq n \) equations in \( y \) alone. The iterative process to solve the smaller system only requires the LU factorization of one \( m \times m \) matrix per step, and the convergence is quadratic. Once \( y^\ast \) has been obtained, \( z^\ast \) is computed by direct solution of a linear system. Details of the numerical implementation are provided and several examples are presented.

1. Introduction

Many applications [1, 2] lead to a system of separable nonlinear equations:

\[ F(y, z) = A(y)z + b(y) = 0, \quad (1) \]

where the matrix \( A(y) \in \mathbb{R}^{m \times N} \) and the vector \( b(y) \in \mathbb{R}^{m} \) are continuously differentiable functions of \( y \in \mathbb{R}^{n} \) and \( z \in \mathbb{R}^{N} \) with \( m \geq N + n \). Typically \( n \) is very small, and for compatible systems (i.e., those with exact solutions) \( m \) is usually close to \( N \). We assume that \( F^\prime(y, z) \) is Lipschitz continuous and has full rank \( N + n \) in a neighborhood of a solution \((y^\ast, z^\ast)\); thus, we assume that \( A(y) \) has full rank \( N \) in this neighborhood.

Standard projection methods, such as VARPRO [3, 4], transform the problem (1) to the minimization of a function in \( y \) alone:

\[ \min_{y} \| (I - A(y)A^\prime(y))b(y) \|, \quad (2) \]

where we use the Euclidean norm throughout this paper.

In [5, 6], we proposed a different method, using left orthonormal null vectors of \( A(y) \), to reduce (1) to an equation of the form

\[ f(y) = 0 \quad (3) \]

in \( y \) alone. We assumed that the system is fully determined \((m = N + n)\), and \( A(y) \) has full rank \( N \). That algorithm was extended to overdetermined systems \((m > N + n)\) with full rank \( A(y) \) in [7]. One QR factorization of \( A(y) \) is required in each iterative step of the methods used to solve these smaller systems for \( y \). For details of these methods and their relationship to other methods, see [1, 7].

In this paper, we use a special set of linearly independent left null vectors of \( A(y) \) to construct a bordered matrix \( M(y) = [A(y) \mid K] \) which inherits the smoothness of \( A(y) \). We use this to construct a system of \( m - N \) equations of the form \( f(y) = 0 \) in the \( n \) unknowns \( y \) alone. The smaller system inherits the Lipschitz continuity and nonsingularity of the Jacobian matrix of the original system (1) so that quadratic convergence of the Newton or Gauss-Newton method for solving \( f(y) = 0 \) is guaranteed. The QR factorization of \( A(y) \) used in previous methods is here replaced by an LU factorization of \( M(y) \), so the cost of each iterative step may be significantly reduced. The method
works for both fully determined systems and compatible overdetermined systems. This paper extends the work using bordered matrices and LU factorization for underdetermined systems in [8] \((m < N + n)\) to fully determined and compatible overdetermined systems.

In the next section, we show how to reduce the original system to a new system in the variables \(y\) only. The corresponding numerical algorithm is presented in Section 3. Examples and computational results are provided in Section 4.

2. Analysis

Assume that \((y^*, z^*)\) is an exact solution of (1), and that \(F'(y, z)\) is Lipschitz continuous and has full rank \(N\) in a neighborhood of \((y^*, z^*)\). For all \(y\) in this neighborhood, we write the singular value decomposition of \(A(y)\) as

\[
A(y) = U(y) \Sigma(y) V^T(y)
\]

\[
\equiv \begin{bmatrix} u_1(y), \ldots, u_m(y) \\
\sigma_1(y) & \cdots & \sigma_N(y) \\
0 & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & 0
\end{bmatrix}
\times \begin{bmatrix} v_1(y), \ldots, v_N(y) \end{bmatrix}^T,
\]

where \(U(y)\) and \(V(y)\) are orthogonal matrices and \(\sigma_1(y) \geq \cdots \geq \sigma_N(y) > 0\). Since the singular values \(\sigma_1(y), \ldots, \sigma_N(y)\) are continuously dependent on \(y\), and \(A(y^*)\) has full rank \(N\) and thus \(\sigma_N(y^*) > 0\), we have

\[
\sigma_N(y) \geq \frac{1}{2} \sigma_N(y^*)
\]

in a small neighborhood of \(y^*\). In this neighborhood, the following result holds.

**Theorem 1.** Let \(A(y)\) be continuously differentiable and full rank in a neighborhood of \(y^*\). Then, there exists an \(m \times (m - N)\) matrix \(P(y)\) of linearly independent columns that satisfy

\[
\begin{bmatrix} P^T(y) & C^T(y) \end{bmatrix} M(y) = I_m.
\]

We use this to reduce the original system (1) to the form (3) for fully determined systems and compatible overdetermined systems with full rank \(F'(y, z)\) as follows.

**Theorem 2.** (a) \(F(y, z) = A(y)z + b(y) = 0\) if and only if

\[
f(y) \equiv C^T(y) b(y) = 0, \quad z = -\xi(y) \equiv -P^T(y)b(y).
\]

(b) \(F'(y^*, z^*)\) has full rank if and only if \(F'(y^*)\) has full rank.

**Proof.** (a) Using (6) and (11), we have

\[
\begin{bmatrix} P^T(y) & C^T(y) \end{bmatrix} A(y) z + b(y) = \begin{bmatrix} z + P^T(y)b(y) \\
C^T(y)b(y) \end{bmatrix},
\]

which implies the result since \(P^T(y)\) is nonsingular.
(b) Differentiating (13) and using $F(y^*, z^*) = A(y^*)z^* + b(y^*) = 0$, we have
\[
\begin{bmatrix}
    P^T(y^*) \\
    C^T(y^*)
\end{bmatrix}
\begin{bmatrix}
    f'(y^*, z^*)
\end{bmatrix}
= \begin{bmatrix}
    \xi'(y^*) \\
    f'(y^*)
\end{bmatrix}
\begin{bmatrix}
    I_N \\
    0_{(m-N)\times N}
\end{bmatrix},
\]
which implies that $F'(y^*, z^*)$ has full rank if and only if $f'(y^*)$ has full rank.

### 3. Computation

Based on (12), we solve the $m-N$ equations in $n$ unknowns $f(y) = 0$ for $y^*$ and then compute $z^* = -\xi(y^*) = -P^T(y^*)b(y^*)$. For the fully determined system, we use Newton’s method
\[
f'(y^{(k)}) (y^{(k+1)} - y^{(k)}) = -f(y^{(k)})
\]
to compute $y^*$, and for compatible overdetermined systems, we use the Gauss-Newton method
\[
\begin{bmatrix}
    (f'(y^{(k)}))^T f'(y^{(k)})
\end{bmatrix} (y^{(k+1)} - y^{(k)}) = -(f'(y^{(k)}))^T f(y^{(k)}).
\]
To evaluate
\[
f(y) = C^T(y) b(y),
\]
we use an LU factorization of $M(y) \equiv [A(y) \mid K]$ to solve for $C^T(y)$ from (11):
\[
C^T(y) M(y) = \begin{bmatrix} 0_{(m-N)\times N} & I_{m-N} \end{bmatrix}.
\]

Now, consider $f'(y)$. Use the following notation: $(B(y))_j^T$ denotes a matrix of the same size as a given matrix $B(y)$ but with each entry being the partial derivative of the corresponding entry of $B(y)$ with respect to $y_j$. Then,
\[
f'(y) = \left[ (C^T(y))^T b(y), \ldots, (C^T(y))^T b(y) \right]_n
\]
\[
+ C^T(y) b'(y).
\]
Differentiating (18), the matrices $(C^T(y))^T_j, j = 1, \ldots, n$ can be computed by solving
\[
(C^T(y))^T_j M(y) = -C^T(y) \left[ (A(y))^T_j \mid 0 \right],
\]
using the previously computed $C^T(y)$ and the LU factorization of $M(y)$.

Notice that Lipschitz continuity of $F'(y, z)$ implies Lipschitz continuity of $(A(y))^T_j, j = 1, \ldots, n$ and $b'(y)$, and also of $M(y)$ and $M^{-1}(y)$. By (18) this implies Lipschitz continuity of $(C^T(y))^T_j, j = 1, \ldots, n$. Using (19), it follows that $f'(y)$ is Lipschitz continuous. Now, Theorem 2(b) guarantees quadratic convergence of the Newton and Gauss-Newton methods for $y^{(0)}$ sufficiently near $y^*$ since the corresponding convergence conditions are satisfied [9–12]. In particular, for the Gauss-Newton method, our assumption that we have a compatible overdetermined system implies that we have a zero residual problem.

An outline of our algorithm follows.

**Algorithm 3.** Given: $F(y, z) = A(y)z + b(y)$ with the corresponding positive integers $m, N, n$, a small positive number $\epsilon$, and a point $y^{(0)}$ near the solution $y^*$. Compute an SVD factorization of $A(y^{(0)})$ and use it to form $K$ as in (7). For $k = 0, 1, 2, \ldots$, do steps (a)–(e).

(a) Form $M(y^{(k)})$ as in (6) and compute its LU factors.

(b) Form $f(y^{(k)})$ as in (17) and (18).

(c) Form $f'(y^{(k)})$ as in (19) and (20).

(d) Find $y^{(k+1)}$ using (15) if $m = N + n$ or (16) if $m > N + n$.

(e) If $\|y^{(k+1)} - y^{(k)}\| < \epsilon$ stop, output $y^* = y^{(k+1)}$ and obtain $z^*$ from (12). Otherwise replace $k$ by $k + 1$ and go to (a).

By (11), in step (e) $z^* = -\xi(y^{(k+1)}) = -P^T(y^{(k+1)})b(y^{(k+1)}) = -[I_N, 0]M^{-1}(y^{(k+1)})b(y^{(k+1)})$. Thus, $z^*$ is the first $N$ component of $-M^{-1}(y^{(k+1)})b(y^{(k+1)})$, which can be computed by using an LU decomposition of $M(y^{(k+1)})$.

Each iteration of our method requires one LU factorization of $M(y^{(k)})$ or about $(2/3)m^3$ flops [13]. A QR factorization of the $m \times N$ matrix $A(y^{(k)})$ costs about $2N^2(m - N/3)$ flops [13], so if $m$ is close to $N$ the new method approximately halves this cost. The matrix to be factored in step (d) is only $n \times n$ and typically $n$ is very small, so this cost is negligible. Since $m$ and $N$ are typically large, the computational cost of the other steps is also small relative to the cost of the matrix factorization.

### 4. Examples

The following three examples illustrate the method.

**Example 1.** Consider
\[
A(y)z + b(y) 
= \begin{bmatrix}
    y_1^2 + y_2^2 - y_1 - y_2 - 2 \\
    2y_1 + 2y_2 \\
    0
\end{bmatrix} z
+ \begin{bmatrix}
    0 \\
    0 \\
    -2y_1 + y_2
\end{bmatrix} = 0.
\]
This equation has a solution \((y_1^*, y_2^*) = (2, -1), (z_1^*, z_2^*) = (1, -1)\), at which \(A(y^*)\) has full rank \(N = 2\). We choose \(y = y^{(0)} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}\). An SVD of \(A(y^{(0)})\) produces the matrix block
\[
K = \begin{bmatrix}
0.2935 & -0.6780 \\
0.1668 & 0.7264 \\
-0.1962 & 0.0576 \\
-0.9206 & -0.0968
\end{bmatrix}.
\] (22)

We list the errors in \(y^{(k)}\) in Table 1, which clearly shows quadratic convergence of the Newton iteration. Using \(y^{(3)} \approx y^*\) gives \(z^* \approx -\zeta(y^{(3)})\) with \(\| -\zeta(y^{(3)}) - z^* \| = 4.0868e-12\).

The next example is a discretized version of an interface eigenvalue problem for a differential equation on two intervals. For references on interface problems in differential equations, see [14].

**Example 2.**

Consider
\[
A(y)z + b(y) := \begin{bmatrix}
y_1 E_{2k+1} - I_{2k+1} & 0 \\
0 & y_2 E_{2k+1} - I_{2k+1}
\end{bmatrix} e_{k+1}^T + \begin{bmatrix}
0 \\
0 \\
-1 \\
1
\end{bmatrix} = 0,
\] (23)

where the \((2k + 1) \times (2k + 1)\) tridiagonal matrix
\[
E_{2k+1} = \begin{bmatrix}
2 & -1 & 0 & 0 & \cdots & 0 \\
-1 & 2 & -1 & 0 & \cdots & 0 \\
0 & -1 & 2 & -1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & 0 & \cdots & 2
\end{bmatrix}.
\] (24)

For eigenvalues of such matrices, see [15]. We use \(k = 9\) so that \(n = 2, m = 40,\) and \(N = 38\). This equation has a solution \((y^*)^T = (csc^2(\pi/40)/4, csc^2(3\pi/40)/4)\) and for \(i = 1, \ldots, 19, z^*(i) = \sin(\pi i/20)\) and \(z^*(i + 19) = \sin(3\pi i/20)\), at which \(A(y^*)\) has full rank \(N = 38\). We choose \(y = y^{(0)} = [4^{1/2} \ 5^{1/2}]^T\). An SVD of \(A(y^{(0)})\) produces an appropriate \(40 \times 2\) matrix block \(K\). We list the errors in \(y^{(k)}\) in Table 2, again showing quadratic convergence. Using \(y^{(3)} \approx y^*\) gives \(z^* \approx -\zeta(y^{(3)})\) with \(\| -\zeta(y^{(3)}) - z^* \| = 9.9027e-15\).

The relative flop count
\[
\frac{2N^2(m - N/3) - 2m^3/3}{2N^2(m - N/3)}
\] (25)

shows that in this example using one LU factorization of \(M(y^{(k)})\) instead of a QR factorization of \(A(y^{(k)})\), as used in alternative methods, reduces the flop count per iteration by approximately 46%. In general, if \(m\) and \(N\) are large and close to each other, as is usually the case in compatible systems, the flop count is approximately halved.

We now give an example of a compatible overdetermined system.

**Example 3.**

Consider
\[
A(y)z + b(y) = \begin{bmatrix} A_1(y_1, y_2) & b_1(y_1, y_2) \\ A_2(y_1, y_2) & b_2(y_1, y_2) \end{bmatrix} z = 0,
\] (26)

where the \((N - 2) \times N\) matrix \(A_1(y_1, y_2)\), the \(5 \times N\) matrix \(A_2(y_1, y_2)\), and \(b_2(y_1, y_2) \in \mathbb{R}^N\) are, respectively,
\[
\begin{bmatrix}
0 & \cdots & 0 & y_1^2 & 1 \\
0 & \cdots & 0 & -y_2^2 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & -y_1^2 & 1
\end{bmatrix},
\]
\[
\begin{bmatrix}
y_1 - y_2 & 1 \\
y_1 & -y_2 & 1 \\
\vdots & \vdots & \ddots & \ddots \\
y_1 & \cdots & \cdots & -y_2 & 1 \\
0 & -y_1 & \cdots & \cdots & -y_2 & 1
\end{bmatrix},
\]
\[
\begin{bmatrix}
-5 \\
-2 & -y_2 \\
-2 \\
-7 \\
-y_1
\end{bmatrix},
\] (27)

and \(b_1(y_1, y_2) = 0 \in \mathbb{R}^{N-2}\). Here, \(n = 2\) and \(m = N + 3\).

We choose \(N = 77\) so that \(m = 80\). This equation has a solution \((y_1^*, y_2^*) = (2, 3)\) and \(z_j^* = 1\) for \(j = 1, \ldots, 77\), at which \(A(y^*)\) has full rank. We choose \(y = y^{(0)} = [\frac{3}{2} \ \frac{1}{2}]^T\). An SVD of \(A(y^{(0)})\) produces an appropriate \(80 \times 3\) matrix.
block $K$. We list the errors in $y^{(k)}$ in Table 3, clearly showing quadratic convergence of the Gauss-Newton method. Using $y^{(4)} \approx y^*$ gives $z^* \approx -\zeta(y^{(4)})$ with $\| -\zeta(y^{(4)}) - z^* \| = 3.4101 \cdot 10^{-13}$. In each iteration of this example, using one LU factorization of $M(y^{(k)})$ instead of one QR factorization of $A(y^{(k)})$ reduces the flop count by approximately 47%, measured by (25).

5. Conclusions

We present an algorithm for solving the separable equation

$$F(y, z) \equiv A(y) z + b(y) = 0,$$

(28)

where the matrix $A(y) \in \mathbb{R}^{m \times N}$ and the vector $b(y) \in \mathbb{R}^{m}$ are continuously differentiable functions of $y \in \mathbb{R}^{n}$ and $z \in \mathbb{R}^{N}$ with $m \geq N + n$. We assume that $F'(y, z)$ is Lipschitz continuous and has full rank $N + n$ in a neighborhood of a solution $(y^*, z^*)$; thus, $A(y)$ has full rank $N$ in this neighborhood. Our technique replaces (28) by $m - N$ equations in $n$ unknowns $f(y) = 0$ which we solve by either the Newton or the Gauss-Newton iterative method, in both cases with quadratic convergence. Our method uses one LU factorization of an $m \times m$ matrix

$$M(y) = [A(y) | K]$$

(29)

instead of a QR factorization of the matrix $A(y)$ per iteration, and may thus be substantially more efficient than competing methods (approximately halving the cost) when $m$ is close to $N$ and $N$ is large, as is usually the case in compatible systems. The method is applicable to fully determined separable systems and to compatible overdetermined separable systems.

References


Submit your manuscripts at http://www.hindawi.com