Example

The nonlinear system

$$3x_1 - \cos(x_2x_3) - \frac{1}{2} = 0,$$

$$x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 = 0,$$

$$e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3} = 0$$

was solved in the previous example by the fixed point method. The approximate solution was obtained as $(0.5, 0, -0.52359877)^t$. Apply Newton's method with $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$.

Solution Define

$$\mathbf{F}(x_1, x_2, x_3) = (f_1(x_1, x_2, x_3), f_2(x_1, x_2, x_3), f_3(x_1, x_2, x_3))^t,$$

where

$$f_1(x_1, x_2, x_3) = 3x_1 - \cos(x_2 x_3) - \frac{1}{2},$$

$$f_2(x_1, x_2, x_3) = x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06,$$

and

$$f_3(x_1, x_2, x_3) = e^{-x_1 x_2} + 20x_3 + \frac{10\pi - 3}{3}.$$

The Jacobian matrix $J(\mathbf{x})$ for this system is

$$J(x_1, x_2, x_3) = \begin{bmatrix} 3 & x_3 \sin x_2 x_3 & x_2 \sin x_2 x_3 \\ 2x_1 & -162(x_2 + 0.1) & \cos x_3 \\ -x_2 e^{-x_1 x_2} & -x_1 e^{-x_1 x_2} & 20 \end{bmatrix}$$

Let $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$.

Then $\mathbf{F}(\mathbf{x}^{(0)}) = (-0.199995, -2.269833417, 8.462025346)^t$ and

$$J(\mathbf{x}^{(0)}) = \begin{bmatrix} 3 & 9.999833334 \times 10^{-4} & 9.999833334 \times 10^{-4} \\ 0.2 & -32.4 & 0.9950041653 \\ -0.09900498337 & -0.09900498337 & 20 \end{bmatrix}$$

Solving the linear system, $J(\mathbf{x}^{(0)})\mathbf{y}^{(0)} = -\mathbf{F}(\mathbf{x}^{(0)})$ gives

$$\mathbf{y}^{(0)} = \begin{bmatrix} 0.3998696728 \\ -0.08053315147 \\ -0.4215204718 \end{bmatrix} \quad \text{and} \quad \mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{y}^{(0)} = \begin{bmatrix} 0.4998696782 \\ 0.01946684853 \\ -0.5215204718 \end{bmatrix}$$

Continuing for $k = 2, 3, \ldots,$ we have

$$\begin{bmatrix} x_1^{(k)} \\ x_2^{(k)} \\ x_3^{(k)} \end{bmatrix} = \begin{bmatrix} x_1^{(k-1)} \\ x_2^{(k-1)} \\ x_3^{(k-1)} \end{bmatrix} + \begin{bmatrix} y_1^{(k-1)} \\ y_2^{(k-1)} \\ y_3^{(k-1)} \end{bmatrix}$$

where

$$\begin{bmatrix} y_1^{(k-1)} \\ y_2^{(k-1)} \\ y_3^{(k-1)} \end{bmatrix} = -\left(J\left(x_1^{(k-1)}, x_2^{(k-1)}, x_3^{(k-1)}\right)\right)^{-1} \mathbf{F}\left(x_1^{(k-1)}, x_2^{(k-1)}, x_3^{(k-1)}\right)$$

Thus, at the kth step, the linear system $J\left(\mathbf{x}^{(k-1)}\right)\mathbf{y}^{(k-1)} = -\mathbf{F}\left(\mathbf{x}^{(k-1)}\right)$ must be solved, where

$$J\left(\mathbf{x}^{(k-1)}\right) = \begin{bmatrix} 3 & x_3^{(k-1)} \sin x_2^{(k-1)} x_3^{(k-1)} & x_2^{(k-1)} \sin x_2^{(k-1)} x_3^{(k-1)} \\ 2x_1^{(k-1)} & -162\left(x_2^{(k-1)} + 0.1\right) & \cos x_3^{(k-1)} \\ -x_2^{(k-1)} e^{-x_1^{(k-1)} x_2^{(k-1)}} & -x_1^{(k-1)} e^{-x_1^{(k-1)} x_2^{(k-1)}} & 20 \end{bmatrix}$$

$$\mathbf{y}^{(k-1)} = \begin{bmatrix} y_1^{(k-1)} \\ y_2^{(k-1)} \\ y_3^{(k-1)} \end{bmatrix},$$

and

$$\mathbf{F}\left(\mathbf{x}^{(k-1)}\right)=\begin{bmatrix} 3x_1^{(k-1)}-\cos x_2^{(k-1)}x_3^{(k-1)}-\frac{1}{2} \\ \left(x_1^{(k-1)}\right)^2-81\left(x_2^{(k-1)}+0.1\right)^2+\sin x_3^{(k-1)}+1.06 \\ e^{-x_1^{(k-1)}x_2^{(k-1)}}+20x_3^{(k-1)}+\frac{10\pi-3}{3} \end{bmatrix}$$

The results using this iterative procedure are shown in the below table,

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\Vert\mathbf{x}^{(k)}-\mathbf{x}^{(k-1)}\Vert_\infty$
0	0.1000000000	0.1000000000	−0.1000000000	
1	0.4998696728	0.0194668485	−0.5215204718	0.4215204718
2	0.5000142403	0.0015885914	−0.5235569638	1.788×10^{-2}
3	0.5000000113	0.0000124448	−0.5235984500	1.576×10^{-3}
4	0.5000000000	8.516×10^{-10}	−0.5235987755	1.244×10^{-5}
5	0.5000000000	-1.375×10^{-11}	−0.5235987756	8.654×10^{-10}

This example illustrates that using a good initial guess, Newton's method converges faster rather than the fixed point method.

Quasi-Newton Methods

- Newton's method is computationally expensive. Because, Jacobian matrix, at each iteration, must be determined and also the linear system that involves this matrix must be solved.
- In most situations, the exact evaluation of the partial derivatives in the Jacobian matrix is inconvenient. In these cases, finite difference approximations cab be used,

$$\frac{\partial f_j}{\partial x_k}(\mathbf{x}^{(i)}) \approx \frac{f_j(\mathbf{x}^{(i)} + \mathbf{e}_k h) - f_j(\mathbf{x}^{(i)})}{h},$$

where h is small in absolute value and $\mathbf{e}_k$ is the vector whose only nonzero entry is a 1 in the kth coordinate.

- Quasi-Newton methods replace the Jacobian matrix in Newton's method with an approximation matrix that is easily updated at each iteration.
- In quasi-Newton methods number of computational operations rather than Newton's method is decreased. But, the convergence rate is also degraded.
- Broyden's method is a quasi-Newton method that reduces the amount of computations at each step without significantly degrading the speed of convergence.

Broyden's Method

To describe Broyden's method, suppose that an initial approximation $\mathbf{x}^{(0)}$ is given to the solution $\mathbf{p}$ of $\mathbf{F}(\mathbf{x}) = \mathbf{0}$. We calculate the next approximation

$\mathbf{x}^{(1)}$ in the same manner as Newton's method. To compute $\mathbf{x}^{(2)}$, in one dimensional case,

$$x^{(2)} = x^{(1)} - \frac{f(x^{(1)})}{f'(x^{(1)})}$$

We can use the approximation,

$$f'(x_1) \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0}$$

as a replacement for $f'(x_1)$ in the single-variable Newton's method.

For nonlinear systems,

$$x^{(2)} = x^{(1)} - J(x^{(1)})^{-1} F(x^{(1)})$$

In similar fashion, we can replace $J\left(\mathbf{x}^{(1)}\right)$ by a matrix A_1 with the property,

$$A_1\left(\mathbf{x}^{(1)}-\mathbf{x}^{(0)}\right)=\mathbf{F}\left(\mathbf{x}^{(1)}\right)-\mathbf{F}\left(\mathbf{x}^{(0)}\right)$$

Broyden proposed the following matrix,

$$A_1=J\left(\mathbf{x}^{(0)}\right)+\frac{\left[\mathbf{F}\left(\mathbf{x}^{(1)}\right)-\mathbf{F}\left(\mathbf{x}^{(0)}\right)-J\left(\mathbf{x}^{(0)}\right)\left(\mathbf{x}^{(1)}-\mathbf{x}^{(0)}\right)\right]\left(\mathbf{x}^{(1)}-\mathbf{x}^{(0)}\right)^t}{\left\|\mathbf{x}^{(1)}-\mathbf{x}^{(0)}\right\|_2^2}$$

Using this matrix, we have

$$\mathbf{x}^{(2)}=\mathbf{x}^{(1)}-A_1^{-1}\mathbf{F}\left(\mathbf{x}^{(1)}\right)$$

Once $\mathbf{x}^{(2)}$ has been determined, the method is repeated to determine $\mathbf{x}^{(3)}$, using A_1 in place of $A_0\equiv J\left(\mathbf{x}^{(0)}\right)$, and with $\mathbf{x}^{(2)}$ and $\mathbf{x}^{(1)}$ in place of $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(0)}$.

In general, once $\mathbf{x}^{(i)}$ has been determined, $\mathbf{x}^{(i+1)}$ is computed by

$$A_i = A_{i-1} + \frac{\mathbf{y}_i - A_{i-1}\mathbf{s}_i}{||\mathbf{s}_i||_2^2}\mathbf{s}_i^t \tag{10.13}$$

and

$$\mathbf{x}^{(i+1)} = \mathbf{x}^{(i)} - A_i^{-1}\mathbf{F}\left(\mathbf{x}^{(i)}\right)$$

where

$$\mathbf{y}_i = \mathbf{F}\left(\mathbf{x}^{(i)}\right) - \mathbf{F}\left(\mathbf{x}^{(i-1)}\right)$$

$$\mathbf{s}_i = \mathbf{x}^{(i)} - \mathbf{x}^{(i-1)}$$

Theorem

(Sherman-Morrison Formula)

Suppose that A is a nonsingular matrix and that $\mathbf{x}$ and $\mathbf{y}$ are vectors with $\mathbf{y}^t A^{-1}\mathbf{x} \neq -1$. Then $A + \mathbf{x}\mathbf{y}^t$ is nonsingular and

$$\left(A + \mathbf{x}\mathbf{y}^t\right)^{-1} = A^{-1} - \frac{A^{-1}\mathbf{x}\mathbf{y}^t A^{-1}}{1 + \mathbf{y}^t A^{-1}\mathbf{x}}$$

■

The Sherman-Morrison formula permits A_i^{-1} to be computed directly from A_{i-1}^{-1}, elim inating the need for a matrix inversion with each iteration.

Letting $A = A_{i-1}$, $\mathbf{x} = (\mathbf{y}_i - A_{i-1}\mathbf{s}_i)/||\mathbf{s}_i||_2^2$, and $\mathbf{y} = \mathbf{s}_i$, in Eq. (10.13) gives

$$A_i^{-1} = \left(A_{i-1} + \frac{\mathbf{y}_i - A_{i-1}\mathbf{s}_i}{||\mathbf{s}_i||_2^2}\mathbf{s}_i^t\right)^{-1}$$

$$= A_{i-1}^{-1} - \frac{A_{i-1}^{-1}\left(\frac{\mathbf{y}_i - A_{i-1}\mathbf{s}_i}{||\mathbf{s}_i||_2^2}\mathbf{s}_i^t\right)A_{i-1}^{-1}}{1 + \mathbf{s}_i^t A_{i-1}^{-1}\left(\dfrac{\mathbf{y}_i - A_{i-1}\mathbf{s}_i}{||\mathbf{s}_i||_2^2}\right)}$$

$$= A_{i-1}^{-1} - \frac{\left(A_{i-1}^{-1}y_i - \mathbf{s}_i\right)\mathbf{s}_i^t A_{i-1}^{-1}}{||\mathbf{s}_i||_2^2 + \mathbf{s}_i^t A_{i-1}^{-1}\mathbf{y}_i - ||\mathbf{s}_i||_2^2}$$

so

$$A_i^{-1} = A_{i-1}^{-1} + \frac{\left(\mathbf{s}_i - A_{i-1}^{-1}\mathbf{y}_i\right)\mathbf{s}_i^t A_{i-1}^{-1}}{\mathbf{s}_i^t A_{i-1}^{-1}\mathbf{y}_i}$$

Example

Use Broyden's method with $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$ to approximate the solution to the nonlinear system

$$3x_1 - \cos(x_2x_3) - \frac{1}{2} = 0$$

$$x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06 = 0$$

$$e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3} = 0$$

Solution

$$J(x_1, x_2, x_3) = \begin{bmatrix} 3 & x_3 \sin x_2x_3 & x_2 \sin x_2x_3 \\ 2x_1 & -162(x_2 + 0.1) & \cos x_3 \\ -x_2e^{-x_1x_2} & -x_1e^{-x_1x_2} & 20 \end{bmatrix}$$

Let $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$ and

$$\mathbf{F}(x_1, x_2, x_3) = (f_1(x_1, x_2, x_3), f_2(x_1, x_2, x_3), f_3(x_1, x_2, x_3))^t,$$

where

$$f_1(x_1, x_2, x_3) = 3x_1 - \cos(x_2x_3) - \frac{1}{2},$$

$$f_2(x_1, x_2, x_3) = x_1^2 - 81(x_2 + 0.1)^2 + \sin x_3 + 1.06,$$

$$f_3(x_1, x_2, x_3) = e^{-x_1x_2} + 20x_3 + \frac{10\pi - 3}{3}$$

Then

$$\mathbf{F}\left(\mathbf{x}^{(0)}\right) = \begin{bmatrix} -1.199950 \\ -2.269833 \\ 8.462025 \end{bmatrix}.$$

Because

$$\begin{aligned} A_0 &= J\left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\right) \\ &= \begin{bmatrix} 3 & 9.999833 \times 10^{-4} & -9.999833 \times 10^{-4} \\ 0.2 & -32.4 & 0.9950042 \\ -9.900498 \times 10^{-2} & -9.900498 \times 10^{-2} & 20 \end{bmatrix} \end{aligned}$$

we have

$$\begin{aligned} A_0^{-1} &= J\left(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}\right)^{-1} \\ &= \begin{bmatrix} 0.3333332 & 1.023852 \times 10^{-5} & 1.615701 \times 10^{-5} \\ 2.108607 \times 10^{-3} & -3.086883 \times 10^{-2} & 1.535836 \times 10^{-3} \\ 1.660520 \times 10^{-3} & -1.527577 \times 10^{-4} & 5.000768 \times 10^{-2} \end{bmatrix}. \end{aligned}$$

So

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - A_0^{-1}\mathbf{F}\left(\mathbf{x}^{(0)}\right) = \begin{bmatrix} 0.4998697 \\ 1.946685 \times 10^{-2} \\ -0.5215205 \end{bmatrix}$$

$$\mathbf{F}\left(\mathbf{x}^{(1)}\right) = \begin{bmatrix} -3.394465 \times 10^{-4} \\ -0.3443879 \\ 3.188238 \times 10^{-2} \end{bmatrix}$$

$$\mathbf{y}_1 = \mathbf{F}\left(\mathbf{x}^{(1)}\right) - \mathbf{F}\left(\mathbf{x}^{(0)}\right) = \begin{bmatrix} 1.199611 \\ 1.925445 \\ -8.430143 \end{bmatrix}$$

$$\mathbf{s}_1 = \begin{bmatrix} 0.3998697 \\ -8.053315 \times 10^{-2} \\ -0.4215204 \end{bmatrix}$$

$$\mathbf{s}_1^t A_0^{-1} \mathbf{y}_1 = 0.3424604$$

$$A_1^{-1} = A_0^{-1} + (1/0.3424604)\left[\left(\mathbf{s}_1 - A_0^{-1}\mathbf{y}_1\right)\mathbf{s}_1^t A_0^{-1}\right]$$

$$= \begin{bmatrix} 0.3333781 & 1.11050 \times 10^{-5} & 8.967344 \times 10^{-6} \\ -2.021270 \times 10^{-3} & -3.094849 \times 10^{-2} & 2.196906 \times 10^{-3} \\ 1.022214 \times 10^{-3} & -1.650709 \times 10^{-4} & 5.010986 \times 10^{-2} \end{bmatrix}$$

and

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} - A_1^{-1}\mathbf{F}\left(\mathbf{x}^{(1)}\right) = \begin{bmatrix} 0.4999863 \\ 8.737833 \times 10^{-3} \\ -0.5231746 \end{bmatrix}$$

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\Vert\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\Vert_2$
3	0.5000066	8.672157×10^{-4}	−0.5236918	7.88×10^{-3}
4	0.5000003	6.083352×10^{-5}	−0.5235954	8.12×10^{-4}
5	0.5000000	-1.448889×10^{-6}	−0.5235989	6.24×10^{-5}
6	0.5000000	6.059030×10^{-9}	−0.5235988	1.50×10^{-6}