System of non-linear equations:

$$\boxed{\begin{array}{rcl} f_1(x_1, x_2, \ldots, x_n) & = & 0 \\ f_2(x_1, x_2, \ldots, x_n) & = & 0 \\ & \vdots & \\ f_n(x_1, x_2, \ldots, x_n) & = & 0 \end{array}}$$

Vector form:

$$\boxed{\mathbf{f}(\mathbf{x}) = \mathbf{0}\,,} \qquad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad \mathbf{f}(\mathbf{x}) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix}.$$

Let $\mathbf{x}^{(k)}$ be k-th approximation to root $\mathbf{x}^*$ of $\mathbf{f}$ & let $\mathbf{x}^{(k)} + \mathbf{p} = \mathbf{x}^*$:

$$\mathbf{x}^{(k)} = \begin{pmatrix} x_1^{(k)} \\ x_2^{(k)} \\ \vdots \\ x_n^{(k)} \end{pmatrix}, \qquad \mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}, \qquad \mathbf{x}^* = \begin{pmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{pmatrix}.$$

Expand $\mathbf{f}(\mathbf{x}^{(k)} + \mathbf{p})$ as a Taylor series in $\mathbf{p}$ about $\mathbf{x}^{(k)}$, i.e. $\mathbf{p} = \mathbf{0}$,

$$\mathbf{0} = \mathbf{f}(\mathbf{x}^*) = \mathbf{f}(\mathbf{x}^{(k)} + \mathbf{p}) = \mathbf{f}(\mathbf{x}^{(k)}) + \mathbf{J}(\mathbf{x}^{(k)})\mathbf{p} + \ldots \approx \mathbf{f}^{(k)} + \mathbf{J}^{(k)}\mathbf{p}\,,$$

where $\mathbf{f}^{(k)}$ denotes $\mathbf{f}(\mathbf{x}^{(k)})$ & $\mathbf{J}$ = Jacobian matrix of $\mathbf{f}$,

$$\boxed{\mathbf{J}^{(k)} \equiv \mathbf{J}(\mathbf{x}^{(k)}) = \left(\begin{array}{cccc} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} & \cdots & \dfrac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \dfrac{\partial f_n}{\partial x_2} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{array} \right)_{x=x^{(k)}} .}$$

$$\mathbf{p} \approx -\left\{\mathbf{J}^{(k)}\right\}^{-1} \mathbf{f}^{(k)} \qquad \Rightarrow$$

$$\mathbf{x}^* = \mathbf{x}^{(k)} + \mathbf{p} \approx \mathbf{x}^{(k)} - \left\{\mathbf{J}^{(k)}\right\}^{-1} \mathbf{f}^{(k)} .$$

Thus Newton-Raphson method for systems:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \left\{\mathbf{J}^{(k)}\right\}^{-1} \mathbf{f}^{(k)} .$$

Calculation of $\left\{\mathbf{J}^{(k)}\right\}^{-1}$ is inefficient.

Solve for $\mathbf{p}$ using Gaussian elimination & then calculate $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{p}$,

$$\boxed{\mathbf{J}^{(k)}\mathbf{p}^{(k)} = -\mathbf{f}^{(k)} , \qquad \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{p}^{(k)} .}$$

Example 5.1

$$\begin{aligned} f_1(x_1, x_2) &= x_1 x_2 - x_2^3 - 1 \\ f_2(x_1, x_2) &= x_1^2 x_2 + x_2 - 5\,. \end{aligned}$$

Take

$$\mathbf{x}^{(0)} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Solution

$$\boxed{\mathbf{f} = \begin{pmatrix} x_1 x_2 - x_2^3 - 1 \\ x_1^2 x_2 + x_2 - 5 \end{pmatrix}, \qquad \mathbf{J} = \begin{pmatrix} x_2 & x_1 - 3x_2^2 \\ 2x_1 x_2 & x_1^2 + 1 \end{pmatrix}.}$$

Thus

$$\mathbf{f}^{(0)} = \begin{pmatrix} -22 \\ 10 \end{pmatrix}, \qquad \mathbf{J}^{(0)} = \begin{pmatrix} 3 & -25 \\ 12 & 5 \end{pmatrix}.$$

Then

$$\mathbf{J}^{(0)}\mathbf{p}^{(0)} = \begin{pmatrix} 3 & -25 \\ 12 & 5 \end{pmatrix} \begin{pmatrix} p_1^{(0)} \\ p_2^{(0)} \end{pmatrix} = \begin{pmatrix} 22 \\ -10 \end{pmatrix} = -\mathbf{f}^{(0)}\,.$$

Solving for $\mathbf{p}^{(0)}$,

$$\mathbf{p}^{(0)} = \begin{pmatrix} p_1^{(0)} \\ p_2^{(0)} \end{pmatrix} = \begin{pmatrix} -0.4444444 \\ -0.9333333 \end{pmatrix}.$$

Thus

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{p}^{(0)} = \begin{pmatrix} 1.555556 \\ 2.066667 \end{pmatrix}.$$

Iterating gives Table 5.1.

k	$x_1^{(k)}$	$x_2^{(k)}$	$f_1^{(k)}$	$f_2^{(k)}$
0	2.000000	3.000000	-22.00000	10.00000
1	1.555556	2.066667	-6.612147	2.067490
2	1.547205	1.477793	-1.940863	1.5401079×10^{-2}
3	1.780535	1.158865	-0.4929187	-0.1671803
4	1.952843	1.028443	$-7.9391047 \times 10^{-2}$	$-4.9492560 \times 10^{-2}$
5	1.997763	1.001240	$-3.4848466 \times 10^{-3}$	$-2.7525513 \times 10^{-3}$
6	1.999995	1.000003	$-7.3910091 \times 10^{-6}$	$-5.9604918 \times 10^{-6}$
7	2.000000	1.000000	0.0000000	0.0000000

Table 5.1: Results for Example 5.1.

- Newton-Raphson method converges quadratically to simple roots.
- Newton-Raphson method requires n^2 first-order partial derivatives.
- Newton-Raphson method requires solution of $n \times n$ system of linear equations at each iteration.
- Disadvantages can be overcome by using finite-difference approximations to derivatives & only partially updating Jacobian each iteration.
- Newton-Raphson method may not converge & there is no sign f bracketing strategy in $n > 1$ dimensions.

- In $n > 1$ dimensions the Newton-Raphson method may be modified to improve convergence to a solution by only accepting the new iterate $\mathbf{x}^{(k+1)}$ if

$$\boxed{\left\|\mathbf{f}^{(k+1)}\right\|_2 < \left\|\mathbf{f}^{(k)}\right\|_2 \, .}$$

  In the *damped Newton-Raphson method,*

$$\boxed{\mathbf{J}^{(k)}\mathbf{p}^{(k)} = -\mathbf{f}^{(k)} \, , \qquad \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{p}^{(k)}/2^j \, ,}$$

  where $j \geq 0$ is the smallest integer, such that

$$\boxed{\left\|\mathbf{f}(\mathbf{x}^{(k)} + \mathbf{p}^{(k)}/2^{j-1})\right\|_2 \geq \left\|\mathbf{f}(\mathbf{x}^{(k)})\right\|_2 \, , \qquad \left\|\mathbf{f}(\mathbf{x}^{(k)} + \mathbf{p}^{(k)}/2^{j})\right\|_2 < \left\|\mathbf{f}(\mathbf{x}^{(k)})\right\|_2 \, .}$$

- Other modifications to (almost) guarantee convergence include constraining $\mathbf{p}^{(k)}$ to a *trust region,*

$$\left\|\mathbf{p}^{(k)}\right\|_2 \leq \delta \, ,$$

  where δ is chosen by algorithm.