Newton's method for nonlinear systems of equations

So far we discussed numerical methods to determine the ZEROS of scalar functions in one variable, i.e. $f : [a,b] \to \mathbb{R}$. Next, we consider the problem of determining the ZEROS of vector-valued functions of many variables, i.e. $\vec{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$. In other words, we would like to find the zeros of a nonlinear system of equations in the form:

$$\bar{f}(\bar{x}) = \bar{0} \qquad \bar{x} \in D \subseteq \mathbb{R}^n$$

In an expanded notation the previous system looks like:

$$\begin{cases} f_1(x_1, \cdots, x_n) = 0 \\ \vdots \\ f_n(x_1, \cdots, x_n) = 0 \end{cases}$$

<u>Example</u>: (linear systems)

In the case of linear systems we have that $\bar{f}(\bar{x})$ is linear in $\bar{x}$, i.e., it has the general form

$$\bar{f}(\bar{x}) = A\bar{x} - \bar{b}$$

For example,

$$\begin{cases} x_1 - 3x_2 - 1 = 0 \\ 3x_1 + x_2 + 5 = 0 \end{cases} \Rightarrow \begin{bmatrix} 1 & -3 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

<u>Remark</u> (Geometric meaning of rootfinding for nonlinear systems of equations)

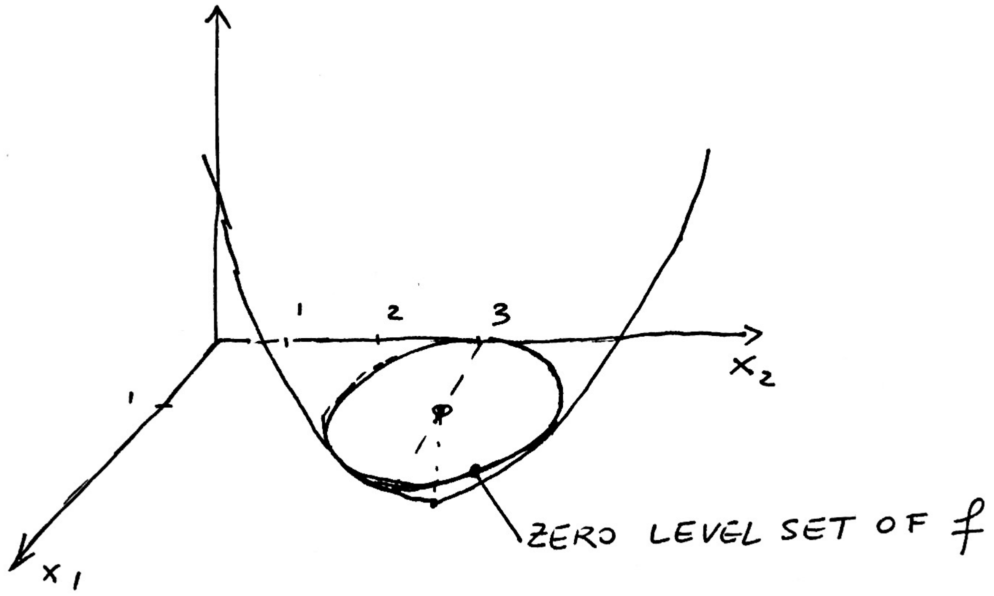
We are basically intersecting the ZERO LEVEL SETS of all surfaces defined by

$y_i = f_i(x_1, \dots, x_n) \quad i = 1, \dots, n$

The zero level set is the intersection of $f_i(x_1, \dots, x_n)$ with the "hyperplane" $y_i = 0$

Example: (zero level set)

$$f(x_1, x_2) = (x_1 - 1)^2 + (x_2 - 3)^2 - 1$$

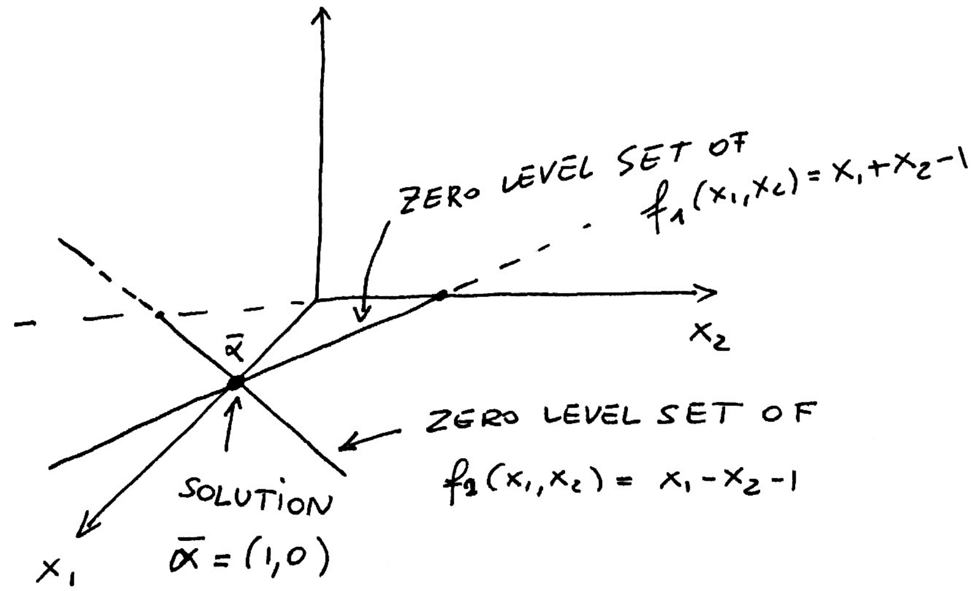


Example (solution to linear systems)

$$\begin{cases} x_1 + x_2 - 1 = 0 \\ x_1 - x_2 - 1 = 0 \end{cases}$$

$x_2 = -x_1 + 1$

$x_2 = x_1 - 1$



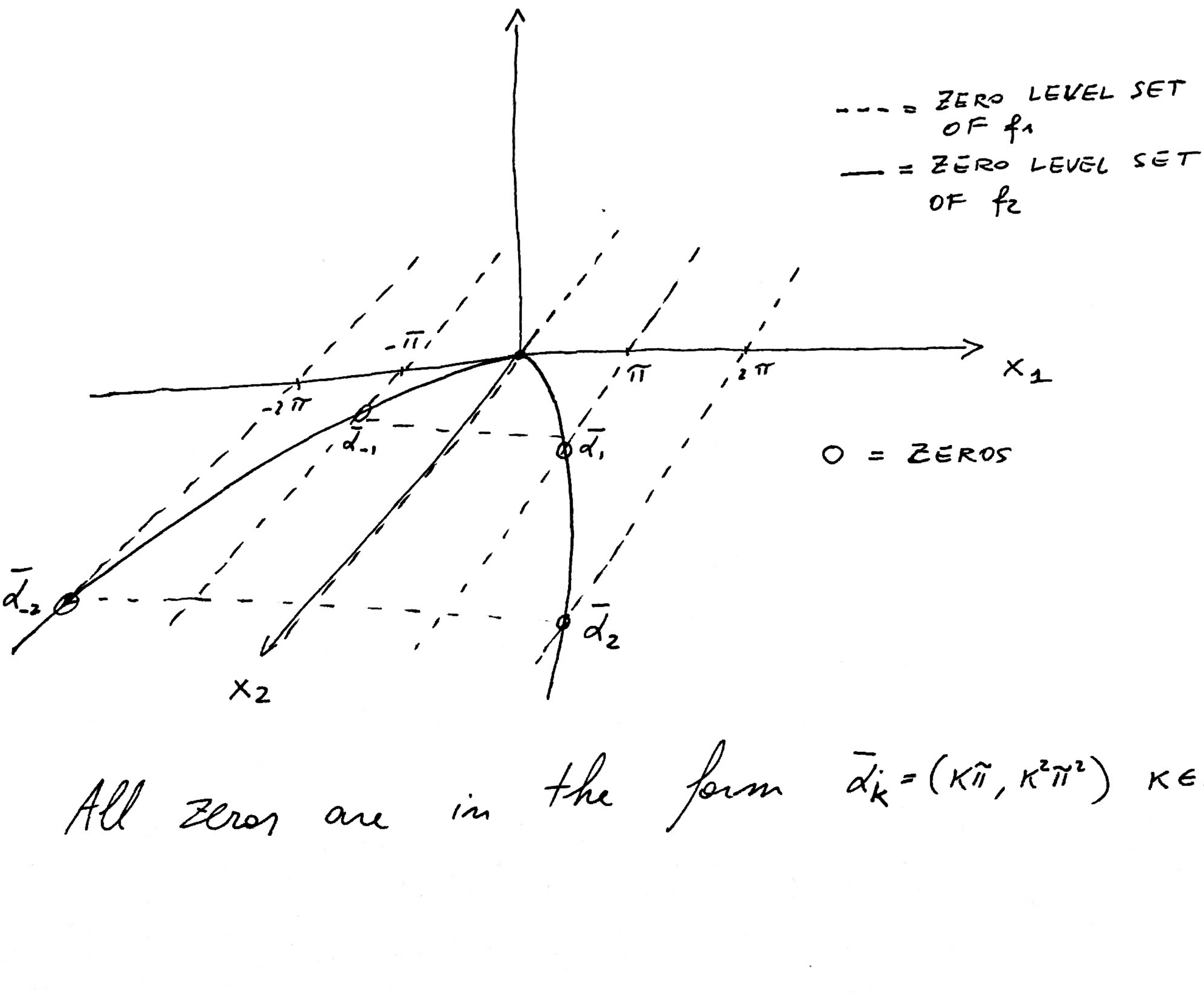
Example (Solution to nonlinear systems)

$$\begin{cases} \sin(x_1) = 0 \\ x_2 - x_1^2 = 0 \end{cases}$$

$\Rightarrow$ the solutions are at the intersections between zero level sets of

$f_1(x_1, x_2) = \sin(x_1)$

$f_2(x_1, x_2) = x_2 - x_1^2$



All zeros are in the form $\bar{\alpha}_k = (k\pi, k^2\pi^2)$ $k \in \mathbb{Z}$

Many methods have been developed to compute the solution to a nonlinear system of equations. Newton's method is one of the simplest. To derive the Newton's method, let us assume that $\bar{f} \in C^{(1)}(D)$ $D \subseteq \mathbb{R}^n$, i.e., that the vector-valued function $\bar{f}(\bar{x})$ is differentiable with continuous derivative in $D \subseteq \mathbb{R}^n$.

<u>Remark</u> : $\bar{f}(\bar{x}) = (\min(x_1), x_2 - x_1^2)$ is $C^{(\infty)}$ in $\mathbb{R}^2$

Consider the Taylor expansion:

$$\bar{f}(\bar{x}^{(k+1)}) = \bar{f}(\bar{x}^{(k)}) + \underbrace{J_{\bar{f}}(\bar{x}^{(k)})}_{\substack{\text{Jacobian of } \bar{f} \text{ calculated} \\ \text{at } \bar{x}^{(k)}}} (\bar{x}^{(k+1)} - \bar{x}^{(k)}) + \dots$$

Setting $\bar{f}(\bar{x}^{(k+1)}) = \bar{0}$ yields:

$$\boxed{\bar{x}^{(k+1)} = \bar{x}^{(k)} - J_{\bar{f}}^{-1}(\bar{x}^{(k)})\, \bar{f}(\bar{x}^{(k)})}$$

vector-valued discrete dynamical system.

$J_{\bar{f}}^{-1}(\bar{x}^{(k)})$ is the inverse of the jacobian matrix at $\bar{x}^{(k)}$.

Remark (what is a Jacobian matrix?)

Consider n multivariate functions of n variables

$$f_1(x_1, \dots, x_n) \quad , \dots , \quad f_n(x_1, \dots, x_n)$$

The Jacobian of $\bar{f}(\bar{x})$ is a matrix that collects the partial derivatives of f_i with respect to any variable x_J, i.e.,

$$J_{\bar{f}}(\bar{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

Example: $f_1(x_1, x_2) = \sin(x_1) \qquad f_2(x_1, x_2) = x_2 - x_1^2$

$$J_{\bar{f}}(\bar{x}) = \begin{bmatrix} \cos(x_1) & 0 \\ -2x_1 & 1 \end{bmatrix}$$

$\det(J_{\bar{f}}(\bar{x})) = \cos(x_1) \Rightarrow$ the Jacobian is singular at $x_1 = \frac{\pi}{2} + k\pi \quad k \in \mathbb{Z}$ (see previous figure)

① Clearly the Jacobian matrix has to be nonsingular for the method to be applicable (otherwise the inverse does not exist).

Remark: At each iteration we need to solve the linear system:

$$J_{\bar{f}}(\bar{x}^{(k)})\,(\bar{x}^{(k+1)}-\bar{x}^{(k)}) = -\,\bar{f}(\bar{x}^{(k)})$$

(DO NOT COMPUTE J_f^{-1} explicitly)

② In particular, $J_f(\bar{\alpha})$ ($\bar{\alpha}$ is the ZERO WE ARE AFTER) must be nonsingular. The continuity of the derivatives ($\bar{f}$ is $C^{(1)}$) guarantees that if $J_f(\bar{\alpha})$ is nonsingular then there exists a neighborhood of $\bar{\alpha}$ where $J_f(\bar{x})$ is nonsingular.

③ As in the scalar case, we need to select $\bar{x}^{(0)}$ close enough to $\bar{\alpha}$

Theorem (convergence of Newton's method for systems)

Let $\bar{f}: \mathbb{R}^n \to \mathbb{R}^n$ be $C^{(1)}$ in a convex open set $D \subseteq \mathbb{R}^n$, $\bar{\alpha} \in D$. Suppose that there exists $R, L, C \in \mathbb{R}^+$ such that

$$\| J_{\bar{f}}^{-1}(\bar{\alpha}) \| \leq C$$

$$\| J_f(\bar{x}) - J_f(\bar{y}) \| \leq L \, \| \bar{x} - \bar{y} \| \qquad \forall\, \bar{x}, \bar{y} \in \underbrace{B}_{\text{BALL}}(\overset{\text{CENTER}}{\bar{\alpha}}, \underbrace{R}_{\text{RADIUS}})$$

Then there exists $r > 0$ such that for any $\bar{X}^{(0)} \in B(\bar{\alpha}, r)$ the sequence

$$\bar{X}^{(K+1)} = \bar{X}^{(K)} - J_f(x^{(K)})\, \bar{f}(x^{(K)})$$

converges to $\bar{\alpha}$ with order 2, i.e.,

$$\| \bar{X}^{(K+1)} - \bar{\alpha} \| \leq C L \, \| \bar{X}^{(K)} - \bar{\alpha} \|^2$$