Zeroing in on the Implicit Function Theorem

Dr.A.L.Pathak

Recall

$$y - 3x + \sin(t) = 1$$
$$4y + \cos(t) + 3t = 4$$

$$\Longrightarrow$$

$$x = \frac{1}{3}\sin(t) - \frac{1}{12}\cos(t) - \frac{1}{4}t$$
$$y = 1 - \frac{1}{4}\cos(t) - \frac{3}{4}t$$

$$\Downarrow$$

$$\vec{\mathbf{g}}(t) = \begin{pmatrix} \frac{1}{3}\sin(t) - \frac{1}{12}\cos(t) - \frac{1}{4}t \\ 1 - \frac{1}{4}\cos(t) - \frac{3}{4}t \end{pmatrix}$$

$$\Downarrow$$

$$f_1(x,y,t) = y - 3x + \sin(t) - 1 = 0$$
$$f_2(x,y,t) = 4y + \cos(t) + 3t - 4 = 0$$

$$\Downarrow$$

$$\vec{\mathbf{F}}(x,y,t) = \begin{pmatrix} y - 3x + \sin(t) - 1 \\ 4y + \cos(t) + 3t - 4 \end{pmatrix} = \vec{\mathbf{0}}$$

$$\Longrightarrow$$

$$\vec{\mathbf{F}} : \mathbb{R}^3 \to \mathbb{R}^2$$
$$\vec{\mathbf{g}} : \mathbb{R} \to \mathbb{R}^2$$

In Summary

- Solving a system of m equations in n unknowns is equivalent to finding the "zeros" of a vector valued function from

$$\mathbb{R}^m \rightarrow \mathbb{R}^n.$$

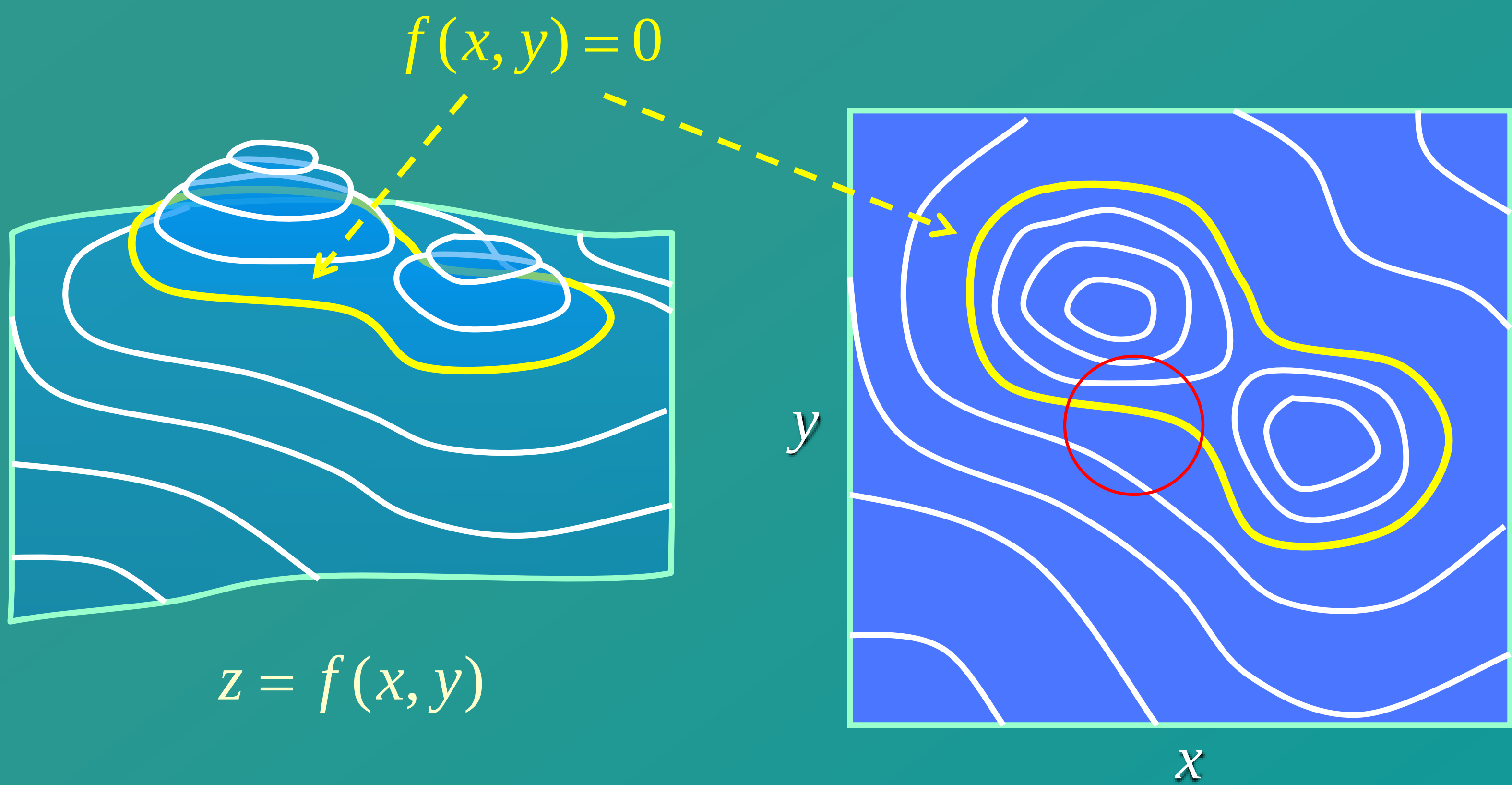
- When $m > n$, such a system will "typically" have infinitely many solutions. In "nice" cases, the solution will be a function from

$$\mathbb{R}^{m-n} \rightarrow \mathbb{R}^n.$$

The Implicit Function theorem will tell us what is meant by the word "typical."

More reminders

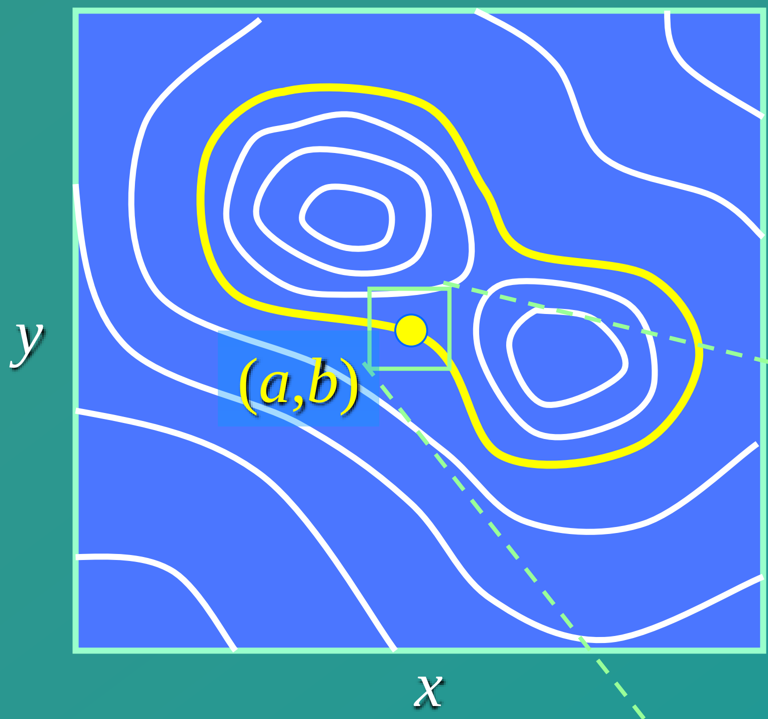
- In general, the 0-level curves are not the graph of a function, but, portions of them may be.
- Though we cannot hope to solve a system “globally,” we can often find a solution function in the neighborhood of a single known solution.
- “Find” is perhaps an overstatement, the implicit function theorem is an existence theorem and we aren’t actually “finding” anything.



Consider the contour line $f(x,y) = 0$ in the xy-plane.

Idea: At least in small regions, this curve might be described by a function $y = g(x)$.

Our goal: find such a function!

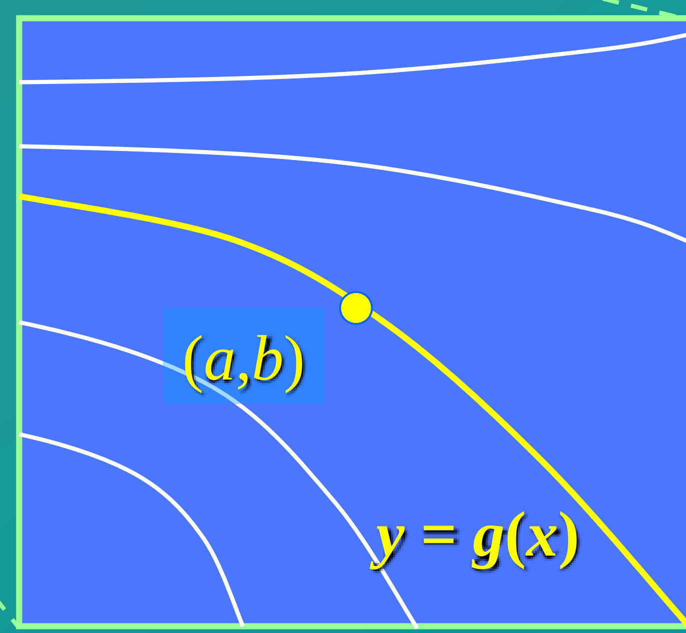


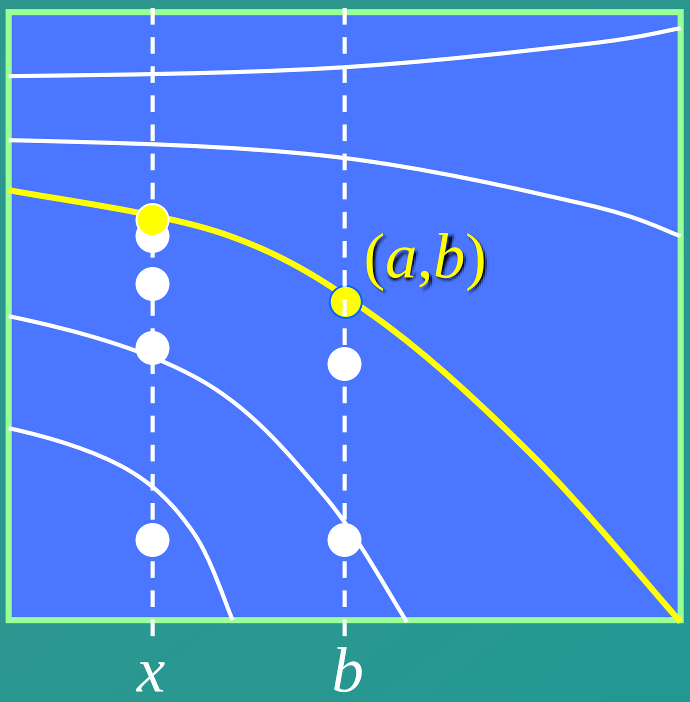
In a small box around (a,b), we can hope to find $g(x)$.

Make sure all of the y-partials in this box are close to D.

Start with a point (a,b) on the contour line, where the contour is not vertical:

$$D = \left.\left(\frac{\partial f}{\partial y}\right)\right|_{(a,b)} \neq 0$$





How to construct $g(x)$

Define

$$\phi_x(y) = y - \frac{f(x,y)}{D}$$

Iterate $\phi_x(y)$ to find its fixed point, where $f(x,y) = 0$. Let fixed point be $g(x)$.

Non-trivial issues:

- Make sure the iterated map converges. (*Quasi-Newton's methods*!)
- Make sure you get the right fixed point. (*Don't leave the box!*)

A bit of notation

- To simplify things, we will write our vector valued function $\mathbf{F} : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^{n}$.
- We will write our “input” variables as concatenations of n-vectors and m-vectors.

  e.g. $(\mathbf{y},\mathbf{x})=(y_1, y_2,. . ., y_n, x_1, x_2,. . ., x_m)$
- So when we solve $\mathbf{F}(\mathbf{y},\mathbf{x})=\mathbf{0}$ we will be solving for the *y*-variables in terms of the **x**-variables.

Implicit Function Theorem---in brief

- $\mathbf{F} : \mathbb{R}^{n+m} \longrightarrow \mathbb{R}^{n}$ has continuous partials.
- Suppose $\mathbf{b} \in \mathbb{R}^{n}$ and $\mathbf{a} \in \mathbb{R}^{m}$ with $\mathbf{F}(\mathbf{b},\mathbf{a})=\mathbf{0}$.
- The *n* x *n* matrix that corresponds to the **y** partials of **F** (call it D) is invertible.

Then “near” **a** there exists a unique function **g**(**x**) such that $\mathbf{F}(\mathbf{g}(\mathbf{x}),\mathbf{x})=\mathbf{0}$; moreover **g**(**x**) is continuous.

What on Earth. . .?!

Mysterious hypotheses

Differentiable functions--- (As we know...?)

A vector valued function **F** of several real variables is differentiable at a vector **v** if in some small neighborhood of **v**, the graph of **F** “looks a lot like” an affine function.

That is, there is a linear transformation **A** so that for all **z** “close” to **v**,

$$\mathbf{F}(\mathbf{z}) \approx \mathbf{A}(\mathbf{z}-\mathbf{v})+\mathbf{F}(\mathbf{v})$$

Where **A** is the Jacobian matrix made up of all the partial derivatives of **F**.

Suppose that **F** (**v**) = **0.** When can we solve **F** (**z**) = **0** in some neighborhood of **v**?

Well, our ability to find a solution to a particular system of equations depends on the *geometry* of the associated vector-valued function.

Geometry and Solutions

So suppose we have **F**(**b**,**a**) = **0** and **F** is differentiable at (**b**,**a**), so there exists **A** so that for all (**y**,**x**) "close" to (**b**,**a**)

$$\mathbf{F}(\mathbf{y},\mathbf{x}) \approx \mathbf{A}((\mathbf{y}-\mathbf{b},\mathbf{x}-\mathbf{a})) + \mathbf{F}(\mathbf{b},\mathbf{a})$$

But F(b,a) = 0!

Geometry and Solutions

So suppose we have $\mathbf{F}(\mathbf{b},\mathbf{a}) = \mathbf{0}$ and $\mathbf{F}$ is differentiable at $(\mathbf{b},\mathbf{a})$, so there exists $\mathbf{A}$ so that for all $(\mathbf{y},\mathbf{x})$ "close" to $(\mathbf{b},\mathbf{a})$

$$\mathbf{F}(\mathbf{y},\mathbf{x}) \approx \mathbf{A}((\mathbf{y}-\mathbf{b},\mathbf{x}-\mathbf{a}))$$

Since the geometry of $\mathbf{F}$ near $(\mathbf{b},\mathbf{a})$ is "just like" the geometry of $\mathbf{A}$ near $\mathbf{0}$, we should be able to solve $\mathbf{F}(\mathbf{y},\mathbf{x}) = \mathbf{0}$ for $\mathbf{y}$ in terms of $\mathbf{x}$ "near" the known solution $(\mathbf{b},\mathbf{a})$ provided that we can solve $\mathbf{A}(\mathbf{y},\mathbf{x}) = \mathbf{0}$ for $\mathbf{y}$ in terms of $\mathbf{x}$.

When can we do it?

Since **A** is linear,
A(**x** ,**y**)=0 looks like this.

$$\begin{aligned} 2y_1 + y_2 \qquad\qquad + 4x_1 + 7x_2 &= 0 \\ y_1 + y_2 \qquad\qquad + x_1 + 5x_2 &= 0 \\ -y_1 \qquad + y_3 - 3x_1 \qquad\qquad &= 0 \end{aligned}$$

$$\begin{bmatrix} \boxed{\begin{matrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 3 \end{matrix}} & \begin{matrix} 4 & 7 \\ 1 & 5 \\ -3 & 0 \end{matrix} \end{bmatrix}$$

We know that we will be able to solve for the variables y_1, y_2, and y_3 in terms of x_1 and x_2 if and only if the sub-matrix . . .
is invertible.

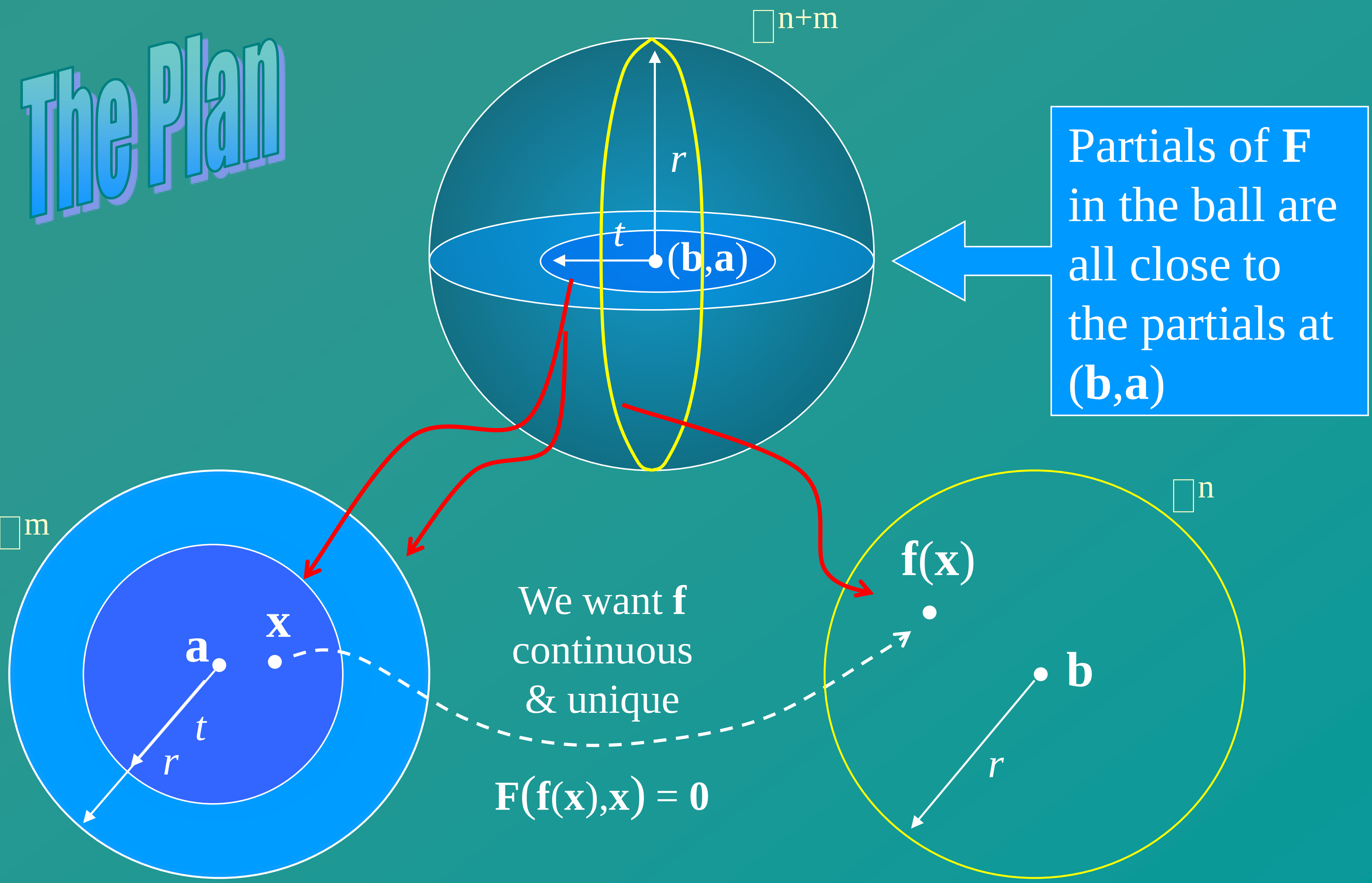
Hypothesis no longer mysterious

When can we solve **F**(**y** , **x**) = **0** for **y** in terms of **x** “near” (**b**,**a**)?

When the domain of the function has more dimensions than the range, and when the “correct” square sub-matrix of **A**=**F’**(**y**,**x**) is invertible.

That is, the square matrix made up of the partial derivatives of the **y**-variables. This sub-matrix is **D** in implicit function theorem.

The Plan



n+m
r
t
(b,a)
Partials of F in the ball are all close to the partials at (b,a)
m
a
x
t
r
We want f continuous & unique
F(f(x),x) = 0
n
f(x)
b
r

The Plan