3 Part 1. Surfaces & the Implicit Function Theorem **2021-22**

This is not a course where surfaces are studies as thoroughly as they deserve, for that see MATH41061 Differentiable Manifolds. Here we introduce them as motivation and a source of examples for the theory of functions of several variables. Surfaces are also the natural setting for Extremal Values & Lagrange Multipliers and Stokes' Theorem, subjects of later chapters.

Before our first definition of a surface recall from Linear Algebra that for a matrix the *row rank*, the maximal number of linearly independent rows, equals the *column rank*, the maximal number of linearly independent columns. This common number is called the **rank of the matrix**.

If M is an $m \times n$ matrix then row rank $M \le m$ and column rank $M \le n$. So the common value, rank M, must be $\le$ both m and n. That is, rank $M \le \min(m, n)$.

Definition 1 *A matrix is of* ***full rank*** *if it has the largest possible rank. That is, the $m \times n$ matrix M is of full rank if* $\text{rank}M = \min(m, n)$.

See Appendix for more discussion of full-rank.

Note that a $1 \times n$ matrix is always of full rank unless all entries are 0.

3.1 Surface given as a Level Set; implicit description

Definition 2 *Let $\mathbf{f} : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ for an open set V. Then the* ***level set of*** $\mathbf{f}$ *is*

$$\mathbf{f}^{-1}(\mathbf{0}) = \{\mathbf{x} \in V : \mathbf{f}(\mathbf{x}) = \mathbf{0}.\}$$

Important The notation $\mathbf{f}^{-1}$ does **not** mean that $\mathbf{f}$ is invertible.

Definition 3 *A* ***surface*** *S* ***in*** *$\mathbb{R}^n$ is described as a level set if there exists an open set $V \subseteq \mathbb{R}^n$, and a $\mathcal{C}^1$-function $\mathbf{f} : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$ for some $1 \le m < n$, such that*

$$S = \left\{\mathbf{x} \in V : \mathbf{f}(\mathbf{x}) = \mathbf{0} \textit{ and } J\mathbf{f}(\mathbf{x}) \textit{ is of full-rank}\right\}.$$

Note One consequence of $\mathbf{f}$ being a $\mathcal{C}^1$-function on V is that all the partial derivatives of $\mathbf{f}$ exist and thus $J\mathbf{f}(\mathbf{x})$ is well-defined for all $\mathbf{x} \in V$.

Most examples of a surface in this course will be given as a level set. In any given example we have to check that

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \implies J\mathbf{f}(\mathbf{x}) \text{ is of full-rank.}$$

We most often this by showing the contrapositive;

$$J\mathbf{f}(\mathbf{x}) \text{ not of full-rank} \implies \mathbf{f}(\mathbf{x}) \neq \mathbf{0}.$$

Example 4 *Is the level set of* $\mathbf{x} = (x, y, u, v)^T \in \mathbb{R}^4$ *satisfying* $xy+uv-1 = 0$ *a surface?*

Solution Define $f : \mathbb{R}^4 \to \mathbb{R}$ by $f(\mathbf{x}) = xy+uv-1$. Then $Jf(\mathbf{x}) = (y, x, v, u)$. The Jacobian matrix $Jf(\mathbf{x})$ is **not** of full-rank only if $(y, x, v, u) = \mathbf{0}$, i.e. $\mathbf{x} = \mathbf{0}$. But $f(\mathbf{0}) = -1 \neq 0$, hence we have a surface. ■

The verification that we have a surface can be long, which is why the details of the following example will be given in the Problems Class.

Example 5 *Consider the level set of* $(x, y, u, v)^T \in \mathbb{R}^4$ *satisfying*

$$x^2 + y^2 + u^2 + v^2 = 11 \quad \text{and} \quad xy + uv = 1. \tag{1}$$

First convince yourself this is non-empty and then ask, is it a surface? Is the Jacobian matrix always of full-rank?

Solution *given in Problems Class.* We do not want to waste time on an empty set. Here the set of solutions is not empty; for example, the second equation is satisfied with $x = 0$, $u = v = 1$. and with these choices the first equation reduces to $y^2 = 9$ which has (two) solutions.

Define

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x^2 + y^2 + u^2 + v^2 - 11 \\ xy + uv - 1 \end{pmatrix}.$$

The Jacobian matrix is

$$J\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2x & 2y & 2u & 2v \\ y & x & v & u \end{pmatrix},$$

where $\mathbf{x} = (x, y, u, v)^T$. When is the matrix **not** of full-rank? It is not of full-rank if one row is a multiple of the other, i.e. there exists $\lambda \in \mathbb{R}$:

$$(2x, 2y, 2u, 2v) = \lambda (y, x, v, u).$$

This gives four equations

$$2x = \lambda y,\ 2y = \lambda x,\ 2u = 2v \ \text{ and } \ 2v = \lambda u. \tag{2}$$

Combine the first two equations as $4x = \lambda(2y) = \lambda(\lambda x) = \lambda^2 x$. This means that **either** $\lambda^2 = 4$ **or** $x = 0$ (Students often forget the second case)

If $\lambda^2 = 4$ then either $\lambda = 2$ or $\lambda = -2$.

If $\lambda = 2$ then (2) reduces to $x = y$ and $u = v$. Then

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2x^2 + 2u^2 - 11 \\ x^2 + u^2 - 1 \end{pmatrix}$$

which is never $\mathbf{0}$ (for example, if the second coordinate is 0 then the first will be -9)

If $\lambda = -2$ then (2) reduces to $x = -y$ and $u = -v$. Then

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2x^2 + 2u^2 - 11 \\ -x^2 - u^2 - 1 \end{pmatrix}$$

which is never $\mathbf{0}$ (the second coordinate can never be zero).

If $x = 0$ then $2y = \lambda x$ means $y = 0$. Return to (2) and combine the last two equations to give $4u = \lambda^2 u$. So either $\lambda^2 = 4$ or $u = 0$. We have already dealt with $\lambda^2 = 4$ so only $u = 0$ remains. But $2v = \lambda u$ means $v = 0$, i.e. $x = y = u = v = 0$. That is, $\mathbf{x} = \mathbf{0}$, yet $\mathbf{f}(\mathbf{0}) = (-11, -1)^T \neq \mathbf{0}$.

We have shown that **if** $J\mathbf{f}(\mathbf{x})$ is **not** of full-rank then $\mathbf{x}$ is **not** a solution of (1). The contrapositive of this is our required result; **if** $\mathbf{x}$ **is** a solution of (1) then $J\mathbf{f}(\mathbf{x})$ **is** of full-rank. Hence the level set is a surface. ■

3.2 Surface given as an Image Set; the parametric description

Definition 6 *A **surface** S **in** $\mathbb{R}^n$ is described as an image set if, for some $0 < r < n$, there exists a $\mathcal{C}^1$-function $\mathbf{F} : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$ on the open set U, such that*

$$S = \left\{ \mathbf{F}(\mathbf{u}) : \mathbf{u} \in U \ \textit{and} \ J\mathbf{F}(\mathbf{u}) \ \textit{is of full-rank} \right\}.$$

We also say this is a **parametric** description of the surface, the coordinates of $\mathbf{u}$ being the parameters.

Note One consequence of $\mathbf{F}$ being a $\mathcal{C}^1$-function on U is that all the partial derivatives of $\mathbf{F}$ exist and thus $J\mathbf{F}(\mathbf{u})$ is well-defined for all $\mathbf{u} \in U$.

Example 7 *Is*

$$\left\{ \begin{pmatrix} u^3 \\ v^3 \\ uv \end{pmatrix} : \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^2 \right\} \tag{3}$$

a parametric surface?

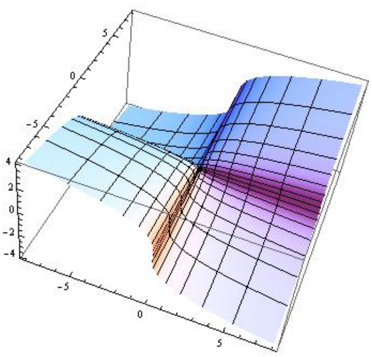
Answer No. Consider the Jacobian matrix at a general point:

$$\begin{pmatrix} 3u^2 & 0 \\ 0 & 3v^2 \\ v & u \end{pmatrix}.$$

If $u \neq 0$ then from the first row we can see that the first column is not a non-zero multiple of the second and so they are linearly independent. Similarly, if $v \neq 0$ then from the second row we can see that the second column is not a non-zero multiple of the first and so they are linearly independent. Thus for $(u, v)^T \neq \mathbf{0}$ the Jacobian matrix is of full rank.

This leaves $(u, v)^T = \mathbf{0}$ when the two columns of the matrix are identical (i.e. $\mathbf{0}$) and so not linearly independent. Hence we have a surface *except at the origin.* ■

Picture of the surface:



For this course $S \subseteq \mathbb{R}^n$ **is a surface** if it can be written in at least one of these two ways; as a level set or parametrically.

3.3 Graphs in $\mathbb{R}^n$.

Definition 8 *The **graph** of a $\mathcal{C}^1$-function $\boldsymbol{\phi} : U \subseteq \mathbb{R}^r \to \mathbb{R}^m$ is the image of the function*

$$\mathbf{F}(\mathbf{u}) = \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\phi}(\mathbf{u}) \end{pmatrix}, \tag{4}$$

for $\mathbf{u} \in U$.

Note that $\mathbf{F} : U \to \mathbb{R}^{m+r}$. If we want $\mathbf{F}$ to map to $\mathbb{R}^n$ we need to choose $m = n - r$.

Example 9 *The graph of $\boldsymbol{\phi}(\mathbf{x}) = x^2 + y^2$, $\mathbf{x} \in \mathbb{R}^2$, is*

$$\left\{ \begin{pmatrix} x \\ y \\ x^2 + y^2 \end{pmatrix} : \mathbf{x} \in \mathbb{R}^2 \right\}. \tag{5}$$

Our definition of graph may appear to be restrictive, since

$$\left\{ \begin{pmatrix} x \\ x^2 + z^2 \\ z \end{pmatrix} : \begin{pmatrix} x \\ z \end{pmatrix} \in \mathbb{R}^2 \right\}$$

is also an example of a graph. Yet, there is a choice in labelling the axes of $\mathbb{R}^3$, and we assume they are labelled (or equivalently, permuted) so the graph looks like (5), or (4) in general.

Graphs as surfaces

Proposition 10 *Graphs are parametric surfaces everywhere.*

Proof Consider the Jacobian matrix of (4) :

$$J\mathbf{F}(\mathbf{u}) = \begin{pmatrix} I_r \\ J\boldsymbol{\phi}(\mathbf{u}) \end{pmatrix}.$$

This is well-defined for all $\mathbf{u} \in U$ since $\boldsymbol{\phi}$ is assumed to be C^1 on U, in particular all it's partial derivatives exist on U. Because of the identity matrix the columns of $J\mathbf{F}(\mathbf{u})$ are linearly independent (For example, the first column has a 1 in the top position. Thus the first column cannot be linear combination of the others, all of which have a 0 in the top position.) Hence the Jacobian matrix is of full rank on U and so the graph (4) is a parametric surface. ■

The converse is **not** true; not all surfaces can be written as a graph.

Example 11 *The boundary of the unit disc in* $\mathbb{R}^2$ *is a surface but not a graph.*

Solution The boundary of the unit disc in $\mathbb{R}^2$ can be given parametrically as the image set

$$\left\{ \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} : 0 \le \theta < 2\pi \right\}.$$

Here $J\mathbf{F}(\theta) = (-\sin\theta, \cos\theta)^T \neq \mathbf{0}$ and so is of full-rank for all θ.

Alternatively the boundary can be given as the level set

$$\left\{ (x,y)^T \in \mathbb{R}^2 : x^2 + y^2 = 1 \right\},$$

when $Jf(\mathbf{x}) = (2x, 2y) \neq 0$ for all $\mathbf{x} : x^2 + y^2 = 1$, and so is of full-rank for all $\mathbf{x}$.

But the boundary **cannot** be given as a graph

$$\left\{ \begin{pmatrix} x \\ \phi(x) \end{pmatrix} : x \in I \right\}$$

for some set I since, for every $-1 < x < 1$, there are *two* possible values for $y : x^2 + y^2 = 1$. ■

But the boundary can be *covered*, i.e. is the union of, four graph. Namely

$$\begin{pmatrix} x \\ \sqrt{1-x^2} \end{pmatrix}, \quad \begin{pmatrix} x \\ -\sqrt{1-x^2} \end{pmatrix}, \quad \begin{pmatrix} \sqrt{1-y^2} \\ y \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -\sqrt{1-y^2} \\ y \end{pmatrix},$$

for $-1 < x, y < 1$. (Make sure you understand why we have *four* graphs and not *two*; this is because we require the variables x or y to lie in an *open* set. So the first two graphs above will not include the points $(\pm 1, 0)^T$, these will be covered by the latter two graphs.)

What we take away from this is that though the boundary is not a graph, **every point** on the boundary **is in** the image of some graph.

3.4 Linear Algebra

Vector subspaces of $\mathbb{R}^n$.

The definition that $\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace is that for all $\mathbf{u}, \mathbf{v} \in \mathcal{V}$ and $\lambda \in \mathbb{R}$ we have $\mathbf{u}+\mathbf{v} \in \mathcal{V}$ and $\lambda\mathbf{v} \in \mathcal{V}$. From the first year we know that all vector spaces have a basis. If $\{\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_r\}$ is a basis for $\mathcal{V}$ we say $\mathcal{V}$ has dimension r. Since $\mathcal{V} \subseteq \mathbb{R}^n$ we must have $r \leq n$.

Also from the First year we know that if $M \in M_{n,r}(\mathbb{R})$ then $\{M\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\}$ is a vector subspace of $\mathbb{R}^n$. In fact, for $\mathbf{t} \in \mathbb{R}^r$,

$$M\mathbf{t} = \sum_{i=1}^{r} t^i M\mathbf{e}_i = \sum_{i=1}^{r} t^i \mathbf{c}_i$$

where $\mathbf{c}_i$ is the i-th column of M. Hence

$$\{M\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\} = \text{span}\{\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_r\}. \tag{6}$$

From this we find a connection between full-rank and dimension.

Lemma 12 *$\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace of dimension r iff there exists a full rank matrix $M \in M_{n,r}(\mathbb{R})$ such that $\mathcal{V} = \{M\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\}$.*

Another example in First year linear algebra was that if $N \in M_{m,n}(\mathbb{R})$ then $\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\}$ is a vector subspace of $\mathbb{R}^n$.

Note that matrix multiplication means that $N\mathbf{x} = \mathbf{0}$ iff $\mathbf{x} \cdot \mathbf{r}_j = 0$ for all the rows of N, so $1 \leq j \leq m$. In turn this is equivalent to $\mathbf{x} \cdot \mathbf{u} = 0$ for all $\mathbf{u}$, linear combinations of $\mathbf{r}_j$, $1 \leq j \leq m$, i.e. $\mathbf{u} \in \text{span}\{\mathbf{r}_1, ..., \mathbf{r}_m\}$. To write this in a similar way to (6) we need some new terminology.

Definition 13 *If $\mathcal{S} \subseteq \mathbb{R}^n$ then the* ***orthogonal complement*** *of $\mathcal{S}$ is*

$$\mathcal{S}^\perp = \{\mathbf{x} \in \mathbb{R}^n : \forall \mathbf{s} \in \mathcal{S},\ \mathbf{x} \cdot \mathbf{s} = 0\}.$$

It is easy to check that $\mathcal{S}^\perp$ is a vector space (see problem sheet). The argument before the definition gives

$$\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\} = \text{span}\{\mathbf{r}_1, ..., \mathbf{r}_m\}^\perp.$$

We need then only **assume** the fact that if $\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace then $\dim \mathcal{V}^\perp = n - \dim \mathcal{V}$ (see Appendix) to deduce another connection between dimension and full-rank.

Lemma 14 *$\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace of dimension r iff there exists a full rank matrix $N \in M_{n-r,n}(\mathbb{R})$ such that $\mathcal{V} = \{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\}$.*

For the motivation of two fundamental results of this course we have

Lemma 15 *$\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace of $\mathbb{R}^n$ of dimension r iff it can be written as a graph of a linear function $\mathbb{R}^r \to \mathbb{R}^{n-r}$.*

Proof ($\Longrightarrow$) Assume $\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace of $\mathbb{R}^n$ of dimension r. Let the columns of $M \in M_{n,r}(\mathbb{R})$ be the vectors in a basis of $\mathcal{V}$. The matrix has column rank r and so will have row rank r, i.e. r linearly independent rows. By permuting the axes of $\mathbb{R}^n$ if necessary, assume the first r rows of M are linearly independent. If we write

$$M = \begin{pmatrix} A \\ B \end{pmatrix},$$

with A an $r \times r$ and B an $(n-r) \times r$ matrix then A will be invertible. Hence, starting with Lemma 12.

$$\mathcal{V} = \{M\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\} = \left\{ \begin{pmatrix} A \\ B \end{pmatrix} A^{-1}A\mathbf{t} : \mathbf{t} \in \mathbb{R}^r \right\} = \left\{ \begin{pmatrix} \mathbf{u} \\ BA^{-1}\mathbf{u} \end{pmatrix} : \mathbf{u} \in \mathbb{R}^r \right\},$$

after the change of variable $\mathbf{t} \to \mathbf{u} = A\mathbf{t}$. Since A is invertible $\mathbf{u}$ varies over $\mathbb{R}^r$ as $\mathbf{t}$ varies over $\mathbb{R}^r$. Thus $\mathcal{V}$ is the graph of the linear function $\mathbf{u} \to BA^{-1}\mathbf{u}$.

($\Longleftarrow$) If $\mathcal{V}$ can be written as a graph of a linear function $\mathbb{R}^r \to \mathbb{R}^{n-r}$ then there exists an $(n-r) \times r$ matrix D such that

$$\mathcal{V} = \left\{ \begin{pmatrix} \mathbf{u} \\ D\mathbf{u} \end{pmatrix} : \mathbf{u} \in \mathbb{R}^r \right\} = \left\{ \begin{pmatrix} I_r \\ D \end{pmatrix} \mathbf{u} : \mathbf{u} \in \mathbb{R}^r \right\}.$$

Because of the identity matrix $\left(\begin{smallmatrix} I_r \\ D \end{smallmatrix}\right)$ is of full rank. So, by Lemma 12, $\mathcal{V}$ is a vector subspace of $\mathbb{R}^n$ of dimension r. ■

Planes in $\mathbb{R}^n$.

Definition 16 *A* ***plane*** *P in $\mathbb{R}^n$ of dimension r is the translate of a vector subspace in $\mathbb{R}^n$ of dimension r.*

Some authors would demand a plane be only of dimension 2 and what we are defining for larger dimensions are *affine subspaces*. I will continue, though, with my definition.

The results on vector spaces can be immediately rewritten for planes as

Corollary 17 *$P \subseteq \mathbb{R}^n$ is a plane of dimension r iff*

- *there exists a point $\mathbf{p} \in \mathbb{R}^n$ and a full rank matrix $M \in M_{n,r}(\mathbb{R})$ such that $P = \{\mathbf{p} + M\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\}$,*
- *there exists a point $\mathbf{p} \in \mathbb{R}^n$ and a full rank matrix $N \in M_{n-r,n}(\mathbb{R})$ such that $P = \{\mathbf{x} \in \mathbb{R}^n : N(\mathbf{x} - \mathbf{p}) = \mathbf{0}\}$,*
- *there exists a point $\mathbf{a} \in \mathbb{R}^{n-r}$ and a matrix $D \in M_{n-r,r}(\mathbb{R})$ such that*

$$P = \left\{ \begin{pmatrix} \mathbf{v} \\ \mathbf{a} + D\mathbf{v} \end{pmatrix} : \mathbf{v} \in \mathbb{R}^r \right\}.$$

Proof Perhaps only the last part requires a proof. By Lemma 15

$$P = \mathbf{p} + \left\{ \begin{pmatrix} I_r \\ D \end{pmatrix} \mathbf{u} : \mathbf{u} \in \mathbb{R}^r \right\},$$

for some $\mathbf{p} \in \mathbb{R}^n$ and matrix $D \in M_{n-r,n}(\mathbb{R})$. Write

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \end{pmatrix}$$

with $\mathbf{p}_0 \in \mathbb{R}^r$ and $\mathbf{p}_1 \in \mathbb{R}^{n-r}$. Then

$$P = \left\{ \begin{pmatrix} \mathbf{p}_0 + \mathbf{u} \\ \mathbf{p}_1 + D\mathbf{u} \end{pmatrix} : \mathbf{u} \in \mathbb{R}^r \right\} = \left\{ \begin{pmatrix} \mathbf{v} \\ \mathbf{p}_1 + D(\mathbf{v} - \mathbf{p}_0) \end{pmatrix} : \mathbf{v} \in \mathbb{R}^r \right\}$$

on changing variable from $\mathbf{u}$ to $\mathbf{v} = \mathbf{p}_0 + \mathbf{u}$. Then $\mathbf{a} = \mathbf{p}_1 - D\mathbf{p}_0$. ■

Definition 18 ***An affine function*** $\mathbf{A} : \mathbb{R}^r \to \mathbb{R}^n$ *is the sum of a constant vector* $\mathbf{a} \in \mathbb{R}^n$ *and linear function* $\mathbf{L} : \mathbb{R}^r \to \mathbb{R}^n$*, so* $\mathbf{A} = \mathbf{a} + \mathbf{L}$*.*

Since every linear map is given by multiplication by some matrix we deduce that $\mathbf{A} : \mathbb{R}^r \to \mathbb{R}^n$ is an affine function iff there exists $M \in M_{n,r}(\mathbb{R})$ such that $\mathbf{A}(\mathbf{x}) = \mathbf{a} + M\mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^r$.

Thus the above results can be rephrased as, every plane is

- the image of an affine function, $(\mathbf{t} \mapsto \mathbf{p} + M\mathbf{t})$
- the null set of an affine function $(\mathbf{x} \mapsto N(\mathbf{x} - \mathbf{p}))$ and the
- graph of an affine function $\mathbf{v} \mapsto \mathbf{a} + D\mathbf{v}$.

Given a surface $S \subseteq \mathbb{R}^n$, of dimension r, and a point $\mathbf{p} \in S$ then there are many planes of dimension r passing through $\mathbf{p}$. Coming up, we will define the Tangent Plane to S at $\mathbf{p}$ and show that it is the best approximation, in some sense, to S out of all planes containing $\mathbf{p}$. We know from Corollary 17 that this plane can be written as a graph. Since this plane approximates the surface then we might hope that the surface can also be written as a graph, at least 'close to' $\mathbf{p}$. This is, in fact, true and if S is given as a level set then this is the content of the Implicit Function Theorem while if S is given parametrically it is a deduction from the Inverse Function Theorem.

3.5 Level sets are graphs (locally)

We have seen by example that a level set is not necessarily a graph. But it is true 'locally'. The meaning of this word will come from the following, **Fundamental Result**.

Theorem 19 ***Implicit Function Theorem.*** *Suppose that* $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ *is a* $\mathcal{C}^1$*-function on an open set* U *where* $1 \leq m < n$*, and there exists* $\mathbf{p} \in U$ *such that* $\mathbf{f}(\mathbf{p}) = \mathbf{0}$ *and the Jacobian matrix* $J\mathbf{f}(\mathbf{p})$ *has full-rank* m*.*

Suppose that the ***final*** m ***columns*** *of the Jacobian matrix* $J\mathbf{f}(\mathbf{p})$ *are linearly independent. Write*

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_0 \\ \mathbf{p}_1 \end{pmatrix},$$

where $\mathbf{p}_0 \in \mathbb{R}^{n-m}$ *and* $\mathbf{p}_1 \in \mathbb{R}^m$ *.*

Then there exists

- *an open set* $V : \mathbf{p}_0 \in V \subseteq \mathbb{R}^{n-m}$*,*
- *a* $\mathcal{C}^1$*-function* $\boldsymbol{\phi} : V \to \mathbb{R}^m$ *and*
- *an open set* $W \subseteq U \subseteq \mathbb{R}^n$ *containing* $\mathbf{p}$

such that for $\mathbf{w} \in W$*, written as*

$$\mathbf{w} = \begin{pmatrix} \mathbf{v} \\ \mathbf{y} \end{pmatrix}$$

with $\mathbf{v} \in \mathbb{R}^{n-m}$ *and* $\mathbf{y} \in \mathbb{R}^m$*,*

$$\mathbf{f}(\mathbf{w}) = \mathbf{0} \quad \textit{if, and only if,} \quad \mathbf{v} \in V \quad \textit{and} \quad \mathbf{y} = \boldsymbol{\phi}(\mathbf{v})\,.$$

Proof is given in the next Chapter, but it is not examinable. ■

Corollary 20 *With the surface* S *given as a level set, so*

$$S = \{\mathbf{x} : \mathbf{f}(\mathbf{x}) = \mathbf{0},\ J\mathbf{f}(\mathbf{x}) \text{ of full rank}\} \tag{7}$$

and $\mathbf{p} \in S$*, there exists an open set* $W \subseteq U \subseteq \mathbb{R}^n$ *containing* $\mathbf{p}$*, an open set* $V \subseteq \mathbb{R}^{n-m}$ *and a* $\mathcal{C}^1$*-function* $\boldsymbol{\phi} : V \to \mathbb{R}^m$ *such that*

$$S \cap W = \left\{ \begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} : \mathbf{v} \in V \right\}.$$

Proof follows immediately from the Implicit Function Theorem. ■

We interpret Corollary 20 by saying that at every point of the surface (7) it is *locally a graph*.

What if m columns of the Jacobian matrix $J\mathbf{f}(\mathbf{p})$ are linearly independent but **not** the last m columns? The j-th column of $J\mathbf{f}(\mathbf{p})$ is $d_j\mathbf{f}(\mathbf{p})$, the derivative with respect to x^j. Permuting (or relabelling) the variables x^j in $\mathbb{R}^n$ therefore permutes the columns in $J\mathbf{f}(\mathbf{p})$. There is no right way to label the axes in $\mathbb{R}^n$, there is always choice (up to some considerations on orientation which need not concern us). Thus, by permuting the variables in the domain $\mathbb{R}^n$, we can assume that the *last* m columns of $J\mathbf{f}(\mathbf{p})$ are linearly independent. See Appendix for an example.

We return to an earlier example.

Example 21 *Consider the surface of points in* $\mathbb{R}^4$ *satisfying*

$$x^2 + y^2 + u^2 + v^2 = 11 \quad \textit{and} \quad xy + uv = 1.$$

A point on this surface is $\mathbf{p} = (1,1,0,3)^T$. *Show that in some open set of* $\mathbb{R}^2$ *containing* $(x,y)^T = (1,1)^T$ *the solutions of this system can be given by some* C^1*-functions* $u = u(x,y)$, $v = v(x,y)$.

What can we say at the point $\mathbf{p}' = (3,0,1,1)^T$?

Solution The Jacobian matrix at $\mathbf{p}$ is

$$J\mathbf{f}\,(\mathbf{p}) = \begin{pmatrix} 2x & 2y & 2u & 2v \\ y & x & v & u \end{pmatrix}_{\mathbf{x}=\mathbf{p}} = \begin{pmatrix} 2 & 2 & 0 & 6 \\ 1 & 1 & 3 & 0 \end{pmatrix}.$$

The last two columns $(0,3)^T$ and $(6,0)^T$ are linearly independent. So no reordering of columns is necessary to apply the Implicit Function Theorem.

This says that there exists,

- an open set $V : (1,1)^T \in V \subseteq \mathbb{R}^2$,
- a $\mathcal{C}^1$-function $\boldsymbol{\phi} : V \to \mathbb{R}^2$ and
- an open set $W \subseteq U \subseteq \mathbb{R}^4$ containing $\mathbf{p}$

such that $\mathbf{f}(\mathbf{w}) = \mathbf{0}$, $\mathbf{w} \in W$ iff

$$\mathbf{w} = \begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} \quad \text{with} \quad \mathbf{v} \in V.$$

That is, writing $\mathbf{v} = (x, y)^T$,

$$\mathbf{w} = \begin{pmatrix} x \\ y \\ \phi^1(x, y) \\ \phi^2(x, y) \end{pmatrix} \quad \text{with} \quad (x, y)^T \in V.$$

So the solution to the Example is that we choose $u = \phi^1$ and $v = \phi^2$, the component functions of the $\boldsymbol{\phi}$ whose existence is assured by the Implicit Function Theorem.

At $\mathbf{p}'$ the last two columns of

$$J\mathbf{f}(\mathbf{p}') = \begin{pmatrix} 6 & 0 & 2 & 2 \\ 0 & 3 & 1 & 1 \end{pmatrix}$$

are **not** linearly independent so the Implicit Function Theorem does **not** imply that u and v are functions of x and y in an open set containing $(3, 0)^T$. (Note, this does **not** mean that u and v are **not** functions of x and y locally, just that the Implicit Function Theorem doesn't tell us that they are.)

We could note instead that the first and fourth columns are linearly independent and so the Implicit Function Theorem implies that x and v are functions of y and u in an open set of $\mathbb{R}^2$ containing $(0, 1)^T$. ■

3.6 Parametric sets are graphs (locally)

We now quote a result **as fundamental as** the Implicit Function Theorem.

Theorem 22 ***Inverse Function Theorem*** *If* $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *is a* $\mathcal{C}^1$*-function on an open subset* $U \subseteq \mathbb{R}^n$ *such that for some* $\mathbf{a} \in U$ *the Jacobian matrix* $J\mathbf{f}(\mathbf{a})$ *is of full rank then* $\mathbf{f}$ *is locally invertible. This means that there exists an open set* $V : \mathbf{a} \in V \subseteq U$, *such that*

1. $\mathbf{f} : V \to \mathbf{f}(V)$ *is a bijection,*
2. $\mathbf{f}(V)$ *is an open subset of* $\mathbb{R}^n$*,*
3. *the inverse function* $\mathbf{g} = \mathbf{f}^{-1} : \mathbf{f}(V) \to V$ *is* $\mathcal{C}^1$ *and* $d\mathbf{g}_{\mathbf{b}} = d\mathbf{f}_{\mathbf{a}}^{-1}$, *or* $J\mathbf{g}\left(\mathbf{b}\right) = J\mathbf{f}(\mathbf{a})^{-1}$, *where* $\mathbf{b} = \mathbf{f}(\mathbf{a})$.

Proof relegated to a question on a problem sheet, where it is derived from the Implicit Function Theorem. ■

Corollary 23 *Let* S *be a surface given parametrically, so*

$$S = \left\{\mathbf{F}(\mathbf{u}) : \mathbf{u} \in U \text{ and } J\mathbf{F}(\mathbf{u}) \text{ is of full-rank}\right\}, \tag{8}$$

for a $\mathcal{C}^1$*-function* $\mathbf{F} : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$.

Let $\mathbf{p} \in S$ *so there exists* $\mathbf{q} \in U$ *for which* $\mathbf{F}(\mathbf{q}) = \mathbf{p}$. *Assume that the rows* 1 *to* r *of* $J\mathbf{F}(\mathbf{q})$ *are linearly independent.*

Then there exist open sets $V \subseteq U \subseteq \mathbb{R}^r$, *with* $\mathbf{q} \in V$, *and* $T \subseteq \mathbb{R}^r$ *along with a* $\mathcal{C}^1$*-function* $\boldsymbol{\phi} : T \to \mathbb{R}^{n-r}$ *such that*

$$S \cap \mathbf{F}(V) = \left\{ \begin{pmatrix} \mathbf{t} \\ \boldsymbol{\phi}(\mathbf{t}) \end{pmatrix} : \mathbf{t} \in T \right\}.$$

Note The point $\mathbf{p} = \mathbf{F}(\mathbf{q}) \in \mathbf{F}(V)$ and so, on the left, we are looking at points on the surface "local to" $\mathbf{p}$. On the right, the set T contains $\mathbf{h}(\mathbf{q})$ (defined below in (9)) and so the graph is over $\mathbf{t}$ "local to" $\mathbf{h}(\mathbf{q})$.

Proof Write

$$\mathbf{F}(\mathbf{u}) = \begin{pmatrix} \mathbf{h}(\mathbf{u}) \\ \mathbf{k}(\mathbf{u}) \end{pmatrix}, \tag{9}$$

say, where $\mathbf{h} : U \subseteq \mathbb{R}^r \to \mathbb{R}^r$ and $\mathbf{k} : U \subseteq \mathbb{R}^r \to \mathbb{R}^{n-r}$ are given by

$$\mathbf{h}(\mathbf{u}) = \begin{pmatrix} F^1(\mathbf{u}) \\ \vdots \\ F^r(\mathbf{u}) \end{pmatrix} \quad \text{and} \quad \mathbf{k}(\mathbf{u}) = \begin{pmatrix} F^{r+1}(\mathbf{u}) \\ \vdots \\ F^n(\mathbf{u}) \end{pmatrix}$$

respectively. Then

$$J\mathbf{F}(\mathbf{u}) = \begin{pmatrix} J\mathbf{h}(\mathbf{u}) \\ J\mathbf{k}(\mathbf{u}) \end{pmatrix}.$$

By assumption the first r rows of $J\mathbf{F}(\mathbf{q})$ are linearly independent i.e. **all** the rows of $J\mathbf{h}(\mathbf{q})$ are linearly independent. Hence $J\mathbf{h}(\mathbf{q})$, a square matrix with all its rows linearly independent, is full-rank.

Apply the Inverse Function Theorem to the function $\mathbf{h} : U \subseteq \mathbb{R}^r \to \mathbb{R}^r$ at the point $\mathbf{q} \in U \subseteq \mathbb{R}^r$. The fact that $J\mathbf{h}(\mathbf{q})$ of full-rank thus implies there exists

- an open set $V : \mathbf{q} \in V \subseteq U$,
- $T = \mathbf{h}(V)$ an open set in $\mathbb{R}^r$,
- an inverse $\mathcal{C}^1$-function $\mathbf{h}^{-1} : T \to V$.

Then

$$S \cap \mathbf{F}(V) = \left\{ \begin{pmatrix} \mathbf{h}(\mathbf{u}) \\ \mathbf{k}(\mathbf{u}) \end{pmatrix} : \mathbf{u} \in V \right\} = \left\{ \begin{pmatrix} \mathbf{t} \\ \mathbf{k}\left(\mathbf{h}^{-1}(\mathbf{t})\right) \end{pmatrix} : \mathbf{t} \in T \right\},$$

having changed variable from $\mathbf{u}$ to $\mathbf{t} = \mathbf{h}(\mathbf{u})$. Hence the result follows with $\boldsymbol{\phi} = \mathbf{k} \circ \mathbf{h}^{-1}$. ■

What if r rows of $J\mathbf{F}(\mathbf{q})$ are linearly independent but not the *first* r rows? The image of $\mathbf{q}$ is $\mathbf{F}(\mathbf{q}) \in \mathbb{R}^n$ and so, permuting the coordinates of $\mathbb{R}^n$ will permute the coordinate functions $F^j(\mathbf{q})$ and thus, in turn, the rows of $J\mathbf{F}(\mathbf{q})$. Therefore, by permuting the variables in the domain $\mathbb{R}^n$, we can assume that the *first* r rows of $J\mathbf{f}(\mathbf{p})$ are linearly independent. See Appendix for an example.

The Implicit Function Theorem and Inverse Function Theorem are examples of *existence results*. They assure us that functions with certain properties exist, but give no indication of how to find or construct them.

Example 24 *Show that at each point on*

$$S = \left\{ \begin{pmatrix} u^3 \\ v^3 \\ uv \end{pmatrix} : \mathbf{u} \neq \mathbf{0} \right\},$$

the surface can be given locally as a graph.

Solution (*Problems class*) We know from Example 7 that S is a surface. The Jacobian matrix is

$$J\mathbf{F}(\mathbf{u}) = \begin{pmatrix} 3u^2 & 0 \\ 0 & 3v^2 \\ v & u \end{pmatrix}.$$

If $\mathbf{p}$ is a point on the surface then $\mathbf{p} = J\mathbf{F}(\mathbf{q})$ for some $\mathbf{q} = (k, \ell)^T \in \mathbb{R}^2$. If both $k, \ell \neq 0$ then the two top rows of $J\mathbf{F}(\mathbf{q})$ are linearly independent. We can then apply Corollary 23 directly and points on the surface "local to" $\mathbf{p}$ are in the graph $(x, y, \phi(\mathbf{x}))^T$ for some $\phi : \mathbb{R}^2 \to \mathbb{R}$ when $\mathbf{x} = (x, y)^T$ is "local to" $(k^3, \ell^3)^T$

If $\ell = 0$ then

$$J\mathbf{F}(\mathbf{q}) = \begin{pmatrix} 3k^2 & 0 \\ 0 & 0 \\ 0 & k \end{pmatrix}$$

and the first and third rows are linearly independent. this time, close to $\mathbf{p}$ the points on the surface are given by the graph $(x, \phi(\mathbf{x}), z)^T$ for some $\phi : \mathbb{R}^2 \to \mathbb{R}$ when $\mathbf{x} = (x, z)^T$ is "local to" $(k^3, 0)^T$.

If $k = 0$ then locally the points are given by the graph $(\phi(\mathbf{x}), y, z)^T$ for some $\phi : \mathbb{R}^2 \to \mathbb{R}$ when $\mathbf{x} = (y, z)^T$ is "local to" $(\ell^3, 0)^T$.

3.7 *Manifolds

Not covered in lectures.

We have made vague definitions of a surface as 'it could' be an image set, Definition 6, or 'it could be' a level set, Definition 2. We have then observed that every image or level set is locally a graph. We can, instead, cut out the image and level sets and define a surface as something that is everywhere locally a graph. We don't call it a surface though.

Definition 25 *An r-**dimensional manifold** (or r-**manifold** for short) in $\mathbb{R}^n$ is a subset $M \subseteq \mathbb{R}^n$ such that for all $\mathbf{p} \in M$ there exists*

- *an open set $W \subseteq \mathbb{R}^n$ containing $\mathbf{p}$,*
- *U an open subset of $U \subseteq \mathbb{R}^r$,*
- *a $\mathcal{C}^1$-function $\boldsymbol{\phi} : U \to \mathbb{R}^{n-r}$,*

such that, up to permutations of the coordinates of $\mathbb{R}^n$, $M \cap W$ is the graph of $\boldsymbol{\phi}$, i.e.

$$M \cap W = \left\{ \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\phi}(\mathbf{u}) \end{pmatrix} : \mathbf{u} \in U \right\}.$$

Alternative definitions of a manifold

Definition 25 essentially says that a manifold is a union of graphs. It can similarly be given as a union of parametric sets, Definition 26 below, or a union of level sets, Definition 27.

Definition 26 *An r-**dimensional manifold** (or r-**manifold** for short) in $\mathbb{R}^n$ is a subset $M \subseteq \mathbb{R}^n$ such that for all $\mathbf{p} \in M$ there exists*

- *an open set $V \subseteq \mathbb{R}^r$,*
- *a $\mathcal{C}^1$-function $\mathbf{F} : V \to \mathbb{R}^n$, such that the Jacobian matrix $J\mathbf{F}(\mathbf{v})$ is of full rank for all $\mathbf{v} \in V$,*
- *an open set $W \subseteq \mathbb{R}^n$ with $\mathbf{p} \in W$,*

such that $M \cap W = \mathbf{F}(V)$.

Thus M is the union of the image sets $\mathbf{F}(V)$ for pairs of sets and functions $(V, \mathbf{F})$.

The second alternative is

Definition 27 *An r-**dimensional manifold** (or r-**manifold** for short) in $\mathbb{R}^n$ is a subset $M \subseteq \mathbb{R}^n$ such that for all $\mathbf{p} \in M$ there exists*

- *an open set $U \subseteq \mathbb{R}^n$ with $\mathbf{p} \in U$,*
- *a $\mathcal{C}^1$-function $\mathbf{f} : U \to \mathbb{R}^{n-r}$ with $J\mathbf{f}(\mathbf{x})$ of full-rank for all $\mathbf{x} \in U$,*

such that $M \cap U = \mathbf{f}^{-1}(\mathbf{0})$.

Thus M is the union of the level sets $\mathbf{f}^{-1}(\mathbf{0})$ for functions $\mathbf{f}$.

More might be demanded of the functions $\boldsymbol{\phi}$, $\mathbf{F}$ or $\mathbf{f}$, i.e. that they might be homeomorphisms, diffeomorphisms or have derivatives of higher order.

Also, it may be possible that, for different $\mathbf{p} \in M$, their associated open sets W may overlap. We may well demand that the functions $\boldsymbol{\phi}$, $\mathbf{F}$ or $\mathbf{f}$ associated with each W work in some 'consistent' way on the intersection. All of this is the subject of the studies on Manifolds found in MATH41061 Differentiable Manifolds.

3.8 Conclusion

For a surface $S \subseteq \mathbb{R}^n$ given by the *image set* of a $\mathcal{C}^1$-function $\mathbf{F} : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$,

- it is easy to *find* a point on the surface S since $\mathbf{F}(\mathbf{u})$ lies in S for all $\mathbf{u} \in U$,
- it is hard to *check* if a given $\mathbf{p} \in U$ lies in S when you have to solve $\mathbf{F}(\mathbf{u}) = \mathbf{p}$,
- it is locally a graph.

For a surface $S \subseteq \mathbb{R}^n$ given by the *level set* $\mathbf{f}^{-1}(\mathbf{0})$ for a $\mathcal{C}^1$ function $\mathbf{f} : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$,

- it is hard to *find* a point on S when you have to solve $\mathbf{f}(\mathbf{x}) = \mathbf{0}$,
- it is easy to *check* if a given $\mathbf{p} \in U$ lies in S, you need only check if $\mathbf{f}(\mathbf{p}) = \mathbf{0}$,
- it is locally a graph.

Appendix for Section 3 part 1.

1. Full Rank

Recall from Linear Algebra A set of vectors $\{\mathbf{v}_i\}_{1\leq i\leq r} \subseteq \mathbb{R}^n$ is linearly independent iff given a set of real numbers $\{c_i\}_{1\leq i\leq r} : \sum_{i=1}^{r} c_i\mathbf{v}_i = \mathbf{0}$ we must have $c_i = 0$ for all $1 \leq i \leq r$.

A set of vectors $\{\mathbf{v}_i\}_{1\leq i\leq r} \subseteq \mathbb{R}^n$ is linearly dependent iff there exists a set of real numbers $\{c_i\}_{1\leq i\leq r}$, not all zero, for which $\sum_{i=1}^{r} c_i\mathbf{v}_i = \mathbf{0}$.

If $M \in M_{m,n}(\mathbb{R})$ is a matrix then the **row rank of** M is the number of linearly independent rows and the **column rank of** M is the number of linearly independent columns.

An AMAZING fact about matrices is that row rank equals column rank. This common value is called the **rank of the matrix**.

If M is an $m \times n$ matrix then row rank $M \leq m$ and column rank $M \leq n$. So the common value, rank M, must be $\leq$ both m and n. That is, rank $M \leq \min(m, n)$.

We say M is of **full-rank** if the rank of M equals $\min(m, n)$.

Example 28 *The matrix*

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 & 0 \end{pmatrix} \tag{10}$$

has row rank 3; the first three rows are linearly independent whilst $r_4 = r_1 + 2r_2 + r_3$.

The matrix

$$\begin{pmatrix} -1 & 1 & 1 & 2 & 1 \\ -1 & 1 & 1 & 2 & 1 \\ 0 & 0 & 3 & 1 & 1 \end{pmatrix} \tag{11}$$

has column rank 2; for example c_2 *and* c_5 *are linearly independent whilst* $c_1 = -c_2, c_3 = 3c_5 - 2c_2$ *and* $c_4 = c_5 + c_2$.

Note how (10) has column rank 3 and (11) has row rank 2 ($r_1 = r_2$).

Note that neither matrix in Example 28 is of full rank.

What if the matrix M has only one row or column? The vectors $\mathbf{v}^i$, $1 \le i \le r$ are linearly *dependent* iff there exist constants c^i, not all zero, such that $\sum_{i=1}^r c^i \mathbf{v}^i = \mathbf{0}$. When you have only one vector write this definition with $r = 1$, so $\mathbf{v}$ is linearly *dependent* iff there exists a non-zero constant such that $c\mathbf{v} = \mathbf{0}$ which can only happen if $\mathbf{v} = \mathbf{0}$. Taking the contrapositive of this, one vector $\mathbf{v}$ is linearly independent iff $\mathbf{v} \neq \mathbf{0}$.

2. The Jacobian matrix of a graph written as an image set.

We build up in steps to a result used, without proof, when looking at the Jacobian matrix of a graph.

a. The Jacobian matrix of the identity map is the identity matrix.

Proof Let $\mathbf{id} : \mathbb{R}^r \to \mathbb{R}^r$ be the identity map, so $\mathbf{id}(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{R}^r$. The i-j-th element of $J\mathbf{id}(\mathbf{x})$ is the j-th partial derivative of the i-th component of $\mathbf{x}$, i.e.

$$\frac{\partial x^i}{\partial x^j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases} \qquad \text{Thus } \left(\frac{\partial x^i}{\partial x^j} \right)_{1 \le i,j \le r} = I_r.$$

Hence $J\mathbf{id}(\mathbf{x}) = I_r$ for all $\mathbf{x} \in \mathbb{R}^r$.

b. Given $\mathbf{g} : \mathbb{R}^r \to \mathbb{R}^p$ and $\mathbf{h} : \mathbb{R}^r \to \mathbb{R}^q$ define $\mathbf{k} : \mathbb{R}^r \to \mathbb{R}^{p+q}$ by

$$\mathbf{k}(\mathbf{u}) = \begin{pmatrix} \mathbf{g}(\mathbf{u}) \\ \mathbf{h}(\mathbf{u}) \end{pmatrix},$$

for $\mathbf{u} \in \mathbb{R}^r$. Then it is easy to see that

$$J\mathbf{k}(\mathbf{u}) = \begin{pmatrix} J\mathbf{g}(\mathbf{u}) \\ J\mathbf{h}(\mathbf{u}) \end{pmatrix}.$$

c. Parts a. and b. are combined in the example of

$$\mathbf{F}(\mathbf{u}) = \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\phi}(\mathbf{u}) \end{pmatrix},$$

for $\mathbf{u} \in \mathbb{R}^r$ and $\boldsymbol{\phi} : \mathbb{R}^r \to \mathbb{R}^{n-r}$. Then the graph of $\boldsymbol{\phi}$ will be the image of $\mathbf{F}$. Yet the image of $\mathbf{F}$ is only a surface at points at which $J\mathbf{F}(\mathbf{u})$ is well defined and of full-rank so we need to know its Jacobian matrix. But we can write $\mathbf{F}(\mathbf{u})$ as

$$\begin{pmatrix} \mathbf{id}(\mathbf{u}) \\ \boldsymbol{\phi}(\mathbf{u}) \end{pmatrix}$$

where $\mathbf{id} : \mathbb{R}^r \to \mathbb{R}^r$ Then, by Parts a. and b. this is

$$J\mathbf{F}(\mathbf{u}) = \begin{pmatrix} J\mathbf{id}(\mathbf{u}) \\ J\boldsymbol{\phi}(\mathbf{u}) \end{pmatrix} = \begin{pmatrix} I_r \\ J\boldsymbol{\phi}(\mathbf{u}) \end{pmatrix}.$$

d. Finally this form of $J\mathbf{F}(\mathbf{u})$ is of full rank. To **not** be of full rank the columns have to be linearly dependent which means that there exists a column that can be written as a linear combination of all the other columns. If this is the i-th column then, because of the identity matrix, the i-th column has 1 in the i-th row while all other columns have 0 and there is no linear combination of the zeros that will give 1, Hence $J\mathbf{F}(\mathbf{u})$ cannot not be full-rank, i.e. it is full rank.

3. Graphs are level sets.

The graph of $\boldsymbol{\phi} : U \subseteq \mathbb{R}^r \to \mathbb{R}^m$ is

$$\begin{pmatrix} \mathbf{u} \\ \boldsymbol{\phi}(\mathbf{u}) \end{pmatrix} = \begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^r \\ \phi^1(\mathbf{u}) \\ \phi^2(\mathbf{u}) \\ \vdots \\ \phi^m(\mathbf{u}) \end{pmatrix} = \begin{pmatrix} u^1 \\ u^2 \\ \vdots \\ u^r \\ u^{r+1} \\ u^{r+2} \\ \vdots \\ u^{r+m} \end{pmatrix}.$$

From this we can see the graph is the level set of points $(u^1, ..., u^{r+m}) \in \mathbb{R}^{r+m}$ satisfying the m equations

$$u^j - \phi^{j-r}\left(u^1, ..., u^r\right) = 0$$

for $r + 1 \leq j \leq r + m$.

4. Jacobian matrix of a graph written as level set.

Repeat the last section but choose $m = n - r$ so the graph is a surface in $\mathbb{R}^n$. Given $\boldsymbol{\phi} : U \subseteq \mathbb{R}^r \to \mathbb{R}^{n-r}$ we define $\mathbf{f}_{\boldsymbol{\phi}} : U \times \mathbb{R}^{n-r} \to \mathbb{R}^{n-r}$ by first writing $\mathbf{x} \in U \times \mathbb{R}^{n-r}$ as

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{y} \end{pmatrix}, \tag{12}$$

with $\mathbf{u} \in U$, $\mathbf{y} \in \mathbb{R}^{n-r}$. Then set

$$\mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x}) = \mathbf{f}_{\boldsymbol{\phi}}\left(\begin{pmatrix} \mathbf{u} \\ \mathbf{y} \end{pmatrix}\right) = \mathbf{y} - \boldsymbol{\phi}(\mathbf{u}).$$

Thus $\mathbf{x} \in G_{\boldsymbol{\phi}}$, the graph of ϕ, if, and only if, $\mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x}) = \mathbf{0}$.

What is next required is the Jacobian matrix $J\mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x})$. In component form, (12) gives

$$x^i = \begin{cases} u^i & \text{if } 1 \leq i \leq r \\ y^j \text{ with } j = i - r & \text{if } r + 1 \leq i \leq n, \end{cases}$$

from (12). Then, if $1 \leq i \leq r$, the i-th column of $J\mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x})$ is

$$d_i \mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x}) = \frac{\partial}{\partial x^i} \mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x}) = \frac{\partial}{\partial x^i} (\mathbf{y} - \boldsymbol{\phi}(\mathbf{u})) = -\frac{\partial}{\partial u^i} \boldsymbol{\phi}(\mathbf{u}) = -d_i \boldsymbol{\phi}(\mathbf{u}),$$

the i-th column of $-J\boldsymbol{\phi}(\mathbf{u})$.

If $r+1 \leq i \leq n$, then

$$d_i \mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x}) = \frac{\partial}{\partial y^j}\left(\mathbf{y} - \boldsymbol{\phi}(\mathbf{x})\right) = \frac{\partial}{\partial y^j}\mathbf{y} = \mathbf{e}_j,$$

the j-th column of the identity matrix I_{n-r}, where $j = i - r$. Make sure you follow all these equalities.

Put these columns together as $J\mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x}) = (-J\boldsymbol{\phi}(\mathbf{u}) \mid I_{n-r})$.

Finally this form of $J\mathbf{f}_{\boldsymbol{\phi}}(\mathbf{x})$ is of full-rank. This is because of the identity matrix for then no row can be written as linear combination of the other rows (as must be the case if the rows are linearly dependent). For example the first row ends in a 1 while all the other rows end in 0. So no linear combination of the other rows will give us the first row and the 1 at the end of it. The same argument works for writing any other row as a linear combination of the remaining rows.

5. $\mathbb{R}^n$ or $\mathbb{R}^r \times \mathbb{R}^{n-r}$?

$\mathbb{R}^n$ is a set of vectors whereas $\mathbb{R}^t \times \mathbb{R}^{n-r}$ is the set of ordered pairs of vectors, so $\mathbb{R}^n$ cannot equal $\mathbb{R}^t \times \mathbb{R}^{n-r}$. But there is a simple bijection. If $\mathbf{x} \in \mathbb{R}^n$ write

$$\mathbf{x} = \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}.$$

Define a map $\mathbb{R}^n \to \mathbb{R}^t \times \mathbb{R}^{n-r}$ by $\mathbf{x} \mapsto (\mathbf{a}, \mathbf{b})$.

Because of this bijection we consider $\mathbb{R}^n$ to be $\mathbb{R}^t \times \mathbb{R}^{n-r}$, and vice-versa, as appropriate. This was used particularly when we looked at level sets as graphs.

6. The surface of the unit ball in $\mathbb{R}^3$.

As an example to show that not every implicitly defined surface or every level set surface can be written as a graph we used the boundary of a disc in $\mathbb{R}^2$. An alternative example is the boundary of a sphere in $\mathbb{R}^3$.

The surface of the unit ball in $\mathbb{R}^3$ is often parameterized as

$$\left\{ \begin{pmatrix} \cos\theta \sin\varphi \\ \sin\theta \sin\varphi \\ \cos\varphi \end{pmatrix} : \begin{matrix} 0 \le \theta < 2\pi \\ 0 \le \varphi \le \pi \end{matrix} \right\}.$$

If φ is restricted to $0 \le \varphi < \pi/2$ then the surface is the top half of the sphere and it is the graph of

$$z = \sqrt{1 - x^2 - y^2} \quad \text{for} \quad (x, y) : x^2 + y^2 < 1.$$

It may have not been obvious but I have defined graphs on *open* sets. In this case $\{(x, y) : x^2 + y^2 < 1\}$. The bottom half of the sphere will be given by the graph of

$$z = -\sqrt{1 - x^2 - y^2} \quad \text{for} \quad (x, y) : x^2 + y^2 < 1.$$

If we take the union of these caps they cover all **but** the equator, $x^2 + y^2 = 1$, of the surface of the sphere. We can then put on other caps

$$\begin{aligned} y &= \sqrt{1 - x^2 - z^2} \quad \text{for} \quad (x, z) : x^2 + z^2 < 1, \\ y &= -\sqrt{1 - x^2 - z^2} \quad \text{for} \quad (x, z) : x^2 + z^2 < 1. \end{aligned}$$

But these four caps will not cover the points $(1, 0, 0)$ and $(-1, 0, 0)$. We will need two more caps for these. Hence we will require 6 graphs to cover the surface of the ball.

7 Permuting the coordinate functions in a parametric set.

Assume $\mathbf{F} : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$, $\mathbf{q} \in U$ and that the $n \times r$ Jacobian matrix $J\mathbf{F}(\mathbf{q})$ has full rank. In the lectures it is claimed that, "by permuting the coordinates of $\mathbb{R}^n$ we can assume that the **first** r rows of $J\mathbf{F}(\mathbf{q})$ are linearly independent." We will see this by way of an example.

Example A Define $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^4$ by

$$\mathbf{F}(\mathbf{u}) = \begin{pmatrix} uv^2 \\ u^2 + 2v^2 \\ u^3 + v \\ v^3 + u \end{pmatrix}$$

for $\mathbf{u} = (u, v)^T \in \mathbb{R}^2$. At $\mathbf{q} = (1, -1)^T$ the Jacobian matrix is

$$J\mathbf{F}(\mathbf{q}) = \left. \begin{pmatrix} v^2 & 2uv \\ 2u & 4v \\ 3u^2 & 1 \\ 1 & 3v^2 \end{pmatrix} \right|_{\mathbf{x}=\mathbf{q}} = \begin{pmatrix} 1 & -2 \\ 2 & -4 \\ 3 & 1 \\ 1 & 3 \end{pmatrix}.$$

The columns are linearly independent and so the column rank, and thus the row rank, is two. This means there must be two linearly independent rows but we see that the first two rows are **not** linearly independent. Yet I can permute the coordinate functions of $\mathbf{F}$ to make them so. For example, permuting the 2nd and 3rd coordinate functions of $\mathbf{F}$, i.e. the 2nd and 3rd coordinates of $\mathbb{R}^4$, we get

$$\mathbf{G}(\mathbf{u}) = \begin{pmatrix} uv^2 \\ u^3 + v \\ u^2 + 2v^2 \\ v^3 + u \end{pmatrix},$$

This time

$$J\mathbf{G}(\mathbf{q}) = \begin{pmatrix} 1 & -2 \\ 3 & 1 \\ 2 & -4 \\ 1 & 3 \end{pmatrix}.$$

The first two rows are now linearly independent. ■

To express this mathematically let $\sigma \in S_n$ be a permutation on $\{1, ..., n\}$. Then define the function (given the same name) $\sigma : \mathbb{R}^n \to \mathbb{R}^n$ by $x^i \mapsto x^{\sigma(i)}$ for all $1 \le i \le n$. This can be written as $(\sigma(\mathbf{x}))^i = x^{\sigma(i)}$ for all i. Finally, given $\mathbf{F} : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$, define $\sigma(\mathbf{F}) : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$ by $\sigma(\mathbf{F})(\mathbf{u}) = \sigma(\mathbf{F}(\mathbf{u}))$ for all $\mathbf{u} \in U$.

Next, given $M \in M_{n,r}(\mathbb{R})$ define $\sigma(M)$ by permuting the rows, $\mathbf{r}^i \to \mathbf{r}^{\sigma(i)}$ for $1 \le i \le n$. Then it can easily be seen that the Jacobian matrix of a permuted function equals the permuted Jacobian matrix. That is,

$$J\sigma(\mathbf{F})(\mathbf{u}) = \sigma(J\mathbf{F}(\mathbf{u})).$$

Let now $S = \{\mathbf{F}(\mathbf{u}) : \mathbf{u} \in U,\ J\mathbf{F}(\mathbf{u}) \text{ of full-rank}\}$ be the parametric surface in $\mathbb{R}^n$ given by $\mathbf{F}$. Given $\mathbf{q} \in U$ then $\mathbf{p} = \mathbf{F}(\mathbf{q})$ is a point on the surface. Let $\sigma \in S_n$ be a permutation which ensures that the first r rows of $\sigma(J\mathbf{F}(\mathbf{q}))$ are linearly independent. Let $\mathbf{G} = \sigma(\mathbf{F})$.

Define $S' = \{\mathbf{G}(\mathbf{u}) : \mathbf{u} \in U,\ J\mathbf{G}(\mathbf{u}) \text{ of full-rank}\}$. Note that $J\mathbf{G}(\mathbf{u})$ and $J\mathbf{F}(\mathbf{u})$ have the same rank and so the same $\mathbf{u}$ occur in both S and S'. We can now apply the Inverse Function Theorem in the form of Corollary 23. So, there exists

- an open set $V : \mathbf{q} \in V \subseteq U$,
- T an open set in $\mathbb{R}^r$,
- and a C^1-function $\boldsymbol{\phi} : T \to \mathbb{R}^{n-r}$ such that

$$S' \cap \mathbf{G}(V) = \left\{ \begin{pmatrix} \mathbf{t} \\ \boldsymbol{\phi}(\mathbf{t}) \end{pmatrix} : \mathbf{t} \in T \right\}.$$

By permuting the coordinate functions back we thus get

$$S \cap \mathbf{F}(V) = \left\{ \sigma^{-1} \begin{pmatrix} \mathbf{t} \\ \boldsymbol{\phi}(\mathbf{t}) \end{pmatrix} : \mathbf{t} \in T \right\}.$$

Return to Example A It is easily checked that $J\mathbf{F}(\mathbf{u})$ if of full rank for all $\mathbf{u} \in \mathbb{R}^2$. So $S = \{\mathbf{F}(\mathbf{u}) : \mathbf{u} \in \mathbb{R}^2\}$ is a surface and $\mathbf{p} = \mathbf{F}(\mathbf{q}) = (1,3,2,0)^T$ at $\mathbf{q} = (1,-1)^T$ is a point on S.

To get the Jacobian matrix of the correct form we chose $G^1 = F^1$; $G^2 = F^3$; $G^3 = F^2$ and $G^4 = F^4$. Yet $\mathbf{G} = \sigma(\mathbf{F})$ for some $\sigma \in S_4$, i.e. $G^i = F^{\sigma(i)}$ for $1 \le i \le 4$. So we have chosen $\sigma(1) = 1$, $\sigma(2) = 3$, $\sigma(3) = 2$ and $\sigma(4) = 4$, that is

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 2 & 4 \end{pmatrix}.$$

We apply the Inverse Function Theorem to the 'permuted surface' $S' = \{\mathbf{G}(\mathbf{u}) : \mathbf{u} \in \mathbb{R}^2\}$ at $\mathbf{p}' = \mathbf{G}(\mathbf{q}) = (1,2,3,0)^T$. Thus there exist open sets V and W and C^1 function ϕ such that

$$S' \cap \mathbf{G}(V) = \left\{ \begin{pmatrix} s \\ t \\ \phi^1(s,t) \\ \phi^2(s,t) \end{pmatrix} : \begin{pmatrix} s \\ t \end{pmatrix} \in W \right\}.$$

'Undoing' the permutation gives

$$S \cap \mathbf{F}(V) = \left\{ \begin{pmatrix} s \\ \phi^1(s,t) \\ t \\ \phi^2(s,t) \end{pmatrix} : \begin{pmatrix} s \\ t \end{pmatrix} \in W \right\}.$$

Note that there are many pairs of rows of $J\mathbf{F}(\mathbf{q})$ which are linearly independent, not just the first and third. For example, the last two rows are linearly independent which means there exists $V', W' \subseteq \mathbb{R}^2$ and C^1 function $\boldsymbol{\eta} : W' \to \mathbb{R}^2$ such that

$$S \cap \mathbf{F}(V') = \left\{ \begin{pmatrix} \eta^1(s,t) \\ \eta^2(s,t) \\ s \\ t \end{pmatrix} : \begin{pmatrix} s \\ t \end{pmatrix} \in W' \right\}.$$

8. Permuting the variables in a level set.

Let $S = \mathbf{f}^{-1}(\mathbf{0})$ for some $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and, for $\mathbf{p} \in S$, assume that the $m \times n$ matrix $J\mathbf{f}(\mathbf{p})$ has full rank m. In particular $J\mathbf{f}(\mathbf{p})$ has m linearly independent columns. In the lectures it is claimed that, "by permuting the variables in the domain $\mathbb{R}^n$, we can assume that the *last* m columns of $J\mathbf{f}(\mathbf{p})$ are linearly independent."

To see this consider first an example of a level set,

$$\begin{aligned} x^2 + y^3 + z^4 &= 18, \\ xy + yz + z^3 &= 2, \end{aligned}$$

at $\mathbf{p} = (4, 1, -1)^T$. For $\mathbf{x} \in \mathbb{R}^3$ define

$$\mathbf{f}(\mathbf{x}) = \begin{pmatrix} x^2 + y^3 + z^4 - 18 \\ xy + yz + z^3 - 2 \end{pmatrix}$$

when

$$J\mathbf{f}(\mathbf{p}) = \left.\begin{pmatrix} 2x & 3y^2 & 4z^4 \\ y & x+z & y+3z^2 \end{pmatrix}\right|_{\mathbf{x}=\mathbf{p}} = \begin{pmatrix} 8 & 3 & 4 \\ 1 & 3 & 4 \end{pmatrix}.$$

The Jacobian matrix is of full-rank, the two rows are linearly independent, but the last two columns are **not** linearly independent. Yet if we permute x and y we get the system

$$\begin{aligned} y^2 + x^3 + z^4 &= 18, \\ yx + xz + z^3 &= 2, \end{aligned} \quad \text{i.e.} \quad \begin{aligned} x^3 + y^2 + z^4 &= 18, \\ xy + xz + z^3 &= 2. \end{aligned}$$

For $\mathbf{x} \in \mathbb{R}^3$ define

$$\mathbf{g}(\mathbf{x}) = \begin{pmatrix} x^3 + y^2 + z^4 - 18 \\ xy + xz + z^3 - 2 \end{pmatrix},$$

then

$$\mathbf{g}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{f}\begin{pmatrix} y \\ x \\ z \end{pmatrix},$$

which shows the permutation of variables in $\mathbb{R}^3$. The point $\mathbf{p}$ becomes $\mathbf{p}' = (1, 4, -1)^T$ and the permutation of variables becomes a permutation of the columns in the Jacobian matrix so

$$J\mathbf{g}(\mathbf{p}') = \begin{pmatrix} 3 & 8 & 4 \\ 3 & 1 & 4 \end{pmatrix},$$

and last two columns are now linearly independent.

The Implicit Function Theorem now says that the two variables represented by the last two columns, y and z, are functions of the remaining variables, x. That is, there exists

- an open set $V : 1 \in V \subseteq \mathbb{R}^{3-2} = \mathbb{R}$,
- a C^1-function $\boldsymbol{\phi} : V \to \mathbb{R}^2$ and
- an open set $W : \mathbf{p}' \in W \subseteq \mathbb{R}^3$

such that for $(x, y, z)^T \in W$

$$\mathbf{g}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0} \quad \text{if, and only if,} \quad x \in V \text{ and } \begin{pmatrix} y \\ z \end{pmatrix} = \boldsymbol{\phi}(x)\,.$$

That is, if $y = \phi^1(x)$ and $z = \phi^2(x)$, the component functions of $\boldsymbol{\phi}(x)$. Hence, $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ and $\mathbf{x} \in W$ iff $\mathbf{x}$ is given as a graph

$$\mathbf{x} = \begin{pmatrix} x \\ \phi^1(x) \\ \phi^2(x) \end{pmatrix}, \quad x \in V.$$

Returning to our original system given by $\mathbf{f}$,

$$\mathbf{f}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0} \quad \text{iff} \quad \mathbf{g}\begin{pmatrix} y \\ x \\ z \end{pmatrix} = \mathbf{0} \quad \text{iff} \quad y \in V,\, x = \phi^1(y) \text{ and } z = \phi^2(y)\,.$$

That is $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ iff $\mathbf{x}$ is given as a point on a graph

$$\mathbf{x} = \begin{pmatrix} \phi^1(y) \\ y \\ \phi^2(y) \end{pmatrix}, \quad y \in V. \tag{13}$$

The **important observation** here is that, even though in

$$J\mathbf{f}(\mathbf{p}) = \begin{pmatrix} 8 & 3 & 4 \\ 1 & 3 & 4 \end{pmatrix},$$

the last two columns are not linearly independent, as required for the Implicit Function as I have given it, the first and third columns are linearly independent. This is sufficient to say that the variables associated with those columns, x and z, can be given as functions of the remaining variables, y, as seen in (13) .

We can recast the above using permutations. By assumption there exists some permutation of the columns of $J\mathbf{f}(\mathbf{p})$ so that the last m are linearly independent. To permute the columns we permute the variables in $\mathbb{R}^n$. Let $\tau \in S_n$ be a permutation such that the last m columns of $J\mathbf{f}(\tau(\mathbf{x}))$ are linearly independent. Let $\mathbf{g}(\mathbf{x}) = \mathbf{f}(\tau(\mathbf{x}))$ and apply the Implicit Function Theorem to $\mathbf{g}$ at $\mathbf{p}' = \tau^{-1}(\mathbf{p})$ to find V, W' with $\mathbf{p}' \in W'$ and $\boldsymbol{\phi} : V \to \mathbb{R}^n$ such that

$$\{\mathbf{x} \in W' : \mathbf{g}(\mathbf{x}) = \mathbf{0}\} = \left\{ \begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} : \mathbf{v} \in V \right\} \tag{14}$$

But, since permutations are bijections,

$$\begin{aligned} \{\mathbf{x} \in W' : \mathbf{g}(\mathbf{x}) = \mathbf{0}\} &= \left\{\tau^{-1}(\mathbf{t}) \in W' : \mathbf{g}\big(\tau^{-1}(\mathbf{t})\big) = \mathbf{0}\right\} \\ &= \tau^{-1}\{\mathbf{t} \in \tau(W') : \mathbf{f}(\mathbf{t}) = \mathbf{0}\}. \end{aligned} \tag{15}$$

Combine (14) with (15) to get

$$\{\mathbf{t} \in W : \mathbf{f}(\mathbf{t}) = \mathbf{0}\} = \left\{ \tau \begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} : \mathbf{v} \in V \right\},$$

where $W = \tau(W')$. Note that $\mathbf{p}' \in W$ means that $\tau(\mathbf{p}') = \tau(\tau^{-1}(\mathbf{p})) = \mathbf{p} \in W$.

9. An example of the use of the Implicit Function Theorem:

Example 29 *Consider the level set of points in* $\mathbb{R}^4$ *satisfying*

$$\begin{aligned} x^2 + y^2 - 2uv + 2xv &= 9 \\ 2xy - uy + vx + uv &= 0. \end{aligned}$$

A solution of the system is $(1, 0, -1, 2)^T$*. Show that in some open set of* $\mathbb{R}^2$ *containing* $(x, y)^T = (1, 0)^T$ *the solutions of this system can be given by some* $\mathcal{C}^1$*-functions* $u = u(x, y)$*,* $v = v(x, y)$*.*

Can we write the solutions as $x = x(u, v)$*,* $y = y(u, v)$ *and if so in the region of what point in* $\mathbb{R}^2$*?*

Solution With $\mathbf{w} = (x, y, u, v)^T \in \mathbb{R}^4$ this level set is $\mathbf{w} : \mathbf{f}(\mathbf{w}) = \mathbf{0}$ where $\mathbf{f} : \mathbb{R}^4 \to \mathbb{R}^2$ is given by

$$\mathbf{f}(\mathbf{w}) = \begin{pmatrix} x^2 + y^2 - 2uv + 2xv - 9 \\ 2xy - uy + vx + uv \end{pmatrix}.$$

At $\mathbf{p} = (1, 0, -1, 2)^T$ the Jacobian matrix is

$$\begin{aligned} J\mathbf{f}(\mathbf{p}) &= \begin{pmatrix} 2x + 2v & 2y & -2v & -2u + 2x \\ 2y + v & 2x - u & -y + v & x + u \end{pmatrix}_{\mathbf{x}=\mathbf{p}} \\ &= \begin{pmatrix} 6 & 0 & -4 & 4 \\ 2 & 3 & 2 & 0 \end{pmatrix}. \end{aligned}$$

The last two columns $(-4, 2)^T$ and $(4, 0)^T$ are linearly independent. So no reordering of columns are necessary to apply the Implicit Function Theorem. We need to write $\mathbf{p} = \left(\mathbf{p}_0^T, \mathbf{p}_1^T\right)^T$ with $\mathbf{p}_0 \in \mathbb{R}^2$ in which case we must have $\mathbf{p}_0 = (1, 0)^T$.

Then the Implicit Function Theorem says there exists, an open set $V : \mathbf{p}_0 \in V \subseteq \mathbb{R}^2$, a $\mathcal{C}^1$-function $\boldsymbol{\phi} : V \to \mathbb{R}^2$ and an open set $W \subseteq U \subseteq \mathbb{R}^4$ containing $\mathbf{p}$ such that $\mathbf{f}(\mathbf{w}) = \mathbf{0}$, $\mathbf{w} \in W$ iff

$$\mathbf{w} = \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\phi}(\mathbf{x}) \end{pmatrix} \quad \text{with} \quad \mathbf{x} \in V.$$

That is,

$$\mathbf{w} = \begin{pmatrix} x \\ y \\ \phi^1(x, y) \\ \phi^2(x, y) \end{pmatrix} \quad \text{with} \quad \begin{pmatrix} x \\ y \end{pmatrix} \in V.$$

So the solution to the Example is that we choose $u = \phi^1$ and $v = \phi^2$, the component functions of the $\boldsymbol{\phi}$ whose existence is assured by the Implicit Function Theorem.

Returning to the Jacobian Matrix $J\mathbf{f}(\mathbf{p})$, we see that the first two columns, $(6, 2)^T$ and $(0, 3)^T$ are linearly independent. We could permute the coordinates in $\mathbb{R}^4$ to ensure these columns were the last two in the Jacobian matrix, and then the conclusion of the Implicit Function Theorem would be that these two variables can be given as functions of the remaining variables, i.e. x and y can be given as $\mathcal{C}^1$ functions of u and v. The values of u and v in $\mathbf{p}$ are -1 and 2, so $x = x(u, v)$ and $y = y(u, v)$ for $(u, v)^T$ in some open set $V : (-1, 2)^T \in V \subseteq \mathbb{R}^2$. ■

10. Background Linear Algebra

In the notes we assumed that if $\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace then $\dim \mathcal{V}^\perp = n - \dim \mathcal{V}$. A way to see this is to start with a basis of $\mathcal{V}$: $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r\}$ say, where $r = \dim V$. You can assume, by applying the Gram-Schmidt method, that this set is orthonormal, i.e. $\mathbf{v}_k \bullet \mathbf{v}_\ell = 0$ if $k \neq \ell$ and 1 if $k = \ell$.

Continuing with the Gram-Schmidt method you can complete this set to a basis of $\mathbb{R}^n$, $\{\mathbf{v}_1, ..., \mathbf{v}_r, \mathbf{w}_1, ..., \mathbf{w}_{n-r}\}$ say, again an orthonormal set.

Then all $\mathbf{x} \in \mathbb{R}^n$ can be written uniquely as

$$\mathbf{x} = \sum_{j=1}^{n-r} \alpha_j \mathbf{w}_j + \sum_{i=1}^{r} \beta_i \mathbf{v}_i,$$

for some $\alpha_j, \beta_i \in \mathbb{R}$. By definition $\mathbf{x} \in \mathcal{V}^\perp$ iff $\mathbf{x} \bullet \mathbf{v} = 0$ for all $\mathbf{v} \in \mathcal{V}$ iff $\mathbf{x} \bullet \mathbf{v}_i = 0$ for all $1 \leq i \leq r$. That is

$$0 = \mathbf{x} \bullet \mathbf{v}_i = \sum_{j=1}^{n-r} \alpha_j \mathbf{w}_j \bullet \mathbf{v}_i + \sum_{i=1}^{r} \beta_i \mathbf{v}_i \bullet \mathbf{v}_i = \beta_i$$

for all $1 \leq i \leq r$. Hence $\mathbf{x} \in \mathcal{V}^\perp$ iff $\mathbf{x} = \sum_{j=1}^{n-r} \alpha_j \mathbf{w}_j$ i.e. $\mathbf{x} \in \text{span}\{\mathbf{w}_1, ..., \mathbf{w}_{n-r}\}$. That is

$$\mathcal{V}^\perp = \text{span}\{\mathbf{w}_1, ..., \mathbf{w}_{n-r}\},$$

and so $\dim \mathcal{V}^\perp = n - r = n - \dim \mathcal{V}$.

If you were to repeat this argument, but starting with $\{\mathbf{w}_1, ..., \mathbf{w}_{n-r}\}$ a basis of $\mathcal{V}^\perp$, and complete it to a basis of $\mathbb{R}^n$ by adding in $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r\}$, you would deduce

$$\left(\mathcal{V}^\perp\right)^\perp = \text{span}(\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_r) = \mathcal{V}.$$

This is used in the proof of

Lemma 14 *$\mathcal{V} \subseteq \mathbf{R}^n$ is a vector subspace of dimension r iff there exists a full rank matrix $N \in M_{n-r,n}(\mathbb{R})$ such that $\mathcal{V} = \{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\}$.*

Proof In the notes it was shown that

$$\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\} = \text{span}\{\mathbf{r}_1, ..., \mathbf{r}_{n-r}\}^\perp,$$

where $r_j, 1 \leq j \leq n - r$, are the rows of N.

$(\implies)$ Assume $\mathcal{V} \subseteq \mathbb{R}^n$ is a vector subspace of dimension r. Then $\dim \mathcal{V}^\perp = n - r$. Let $\{\mathbf{r}_1, ..., \mathbf{r}_{n-r}\}$ be a basis of $\mathcal{V}^\perp$, so $\operatorname{span}\{\mathbf{r}_1, ..., \mathbf{r}_{n-r}\} = \mathcal{V}^\perp$. Construct N by choosing the i-th row of N to be $\mathbf{r}_i$ for $1 \le i \le n-r$. Since the rows are linearly independent then N is of full rank. And

$$\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\} = (\operatorname{span}\{\mathbf{r}_1, ..., \mathbf{r}_{n-r}\})^\perp = \left(\mathcal{V}^\perp\right)^\perp = \mathcal{V}.$$

$(\impliedby)$ Assume $N \in M_{n-r,n}(\mathbb{R})$ is of full rank. Then $\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\}$ is a vector space with

$$\begin{aligned} \dim\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\} &= \dim(\operatorname{span}\{\mathbf{r}_1, ..., \mathbf{r}_{n-r}\})^\perp \\ &= n - (n-r) = r. \end{aligned}$$

■

Calculation of basis. If given a vector space as $\{M\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\}$ we can read off a basis from the columns of M. What if we are given the vector space as $\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\}$?

The matrix N does not need to be of full rank for this to be a vector space, but we can remove those equations $\mathbf{r}_i \cdot \mathbf{x} = 0$ which can be written as linear combinations of others and assume N is of full rank.

If N is of size $m \times n$ it must have m linearly independent rows and thus m linearly independent columns. By permuting the axes of $\mathbb{R}^n$ if necessary, assume the *last* m columns of N are linearly independent.

Write $N = (A \,|\, B)$ with A and B matrices of size $m \times (n-m)$ and $m \times m$ respectively. Here B is a square matrix and, by assumption, it's columns are linearly independent which means it is invertible. Then

$$N\mathbf{x} = \mathbf{0} \quad \text{iff} \quad B^{-1}(A \,|\, B)\,\mathbf{x} = \mathbf{0} \quad \text{iff} \quad \left(B^{-1}A \,|\, I_m\right)\mathbf{x} = \mathbf{0}.$$

Write

$$\mathbf{x} = \begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$$

with $\mathbf{u} \in \mathbb{R}^{n-m}$ and $\mathbf{v} \in \mathbb{R}^m$. Then

$$\left(B^{-1}A \,|\, I_m\right)\mathbf{x} = \mathbf{0} \quad \text{iff} \quad B^{-1}A\mathbf{u} + I_m\mathbf{v} = \mathbf{0},$$

i.e. $\mathbf{v} = -B^{-1}A\mathbf{u}$. Hence

$$\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\} = \left\{ \begin{pmatrix} I_{n-m} \\ -B^{-1}A \end{pmatrix} \mathbf{u} : \mathbf{u} \in \mathbb{R}^{n-m} \right\}. \tag{16}$$

Thus the columns of $\binom{I_{n-m}}{-B^{-1}A}$ give a basis for $\{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\}$. Note this also shows that $\dim \{\mathbf{x} \in \mathbb{R}^n : N\mathbf{x} = \mathbf{0}\} = n - m$ which agrees with Lemma 14.

11. Surfaces as level sets $\mathbf{f}^{-1}(\mathbf{0})$. Which $\mathbf{f}$?

Let

$$S = \left\{\mathbf{x} \in V : \mathbf{f}(\mathbf{x}) = \mathbf{0} \text{ and } J\mathbf{f}(\mathbf{x}) \text{ is of full-rank}\right\},$$

for some $\mathbf{f} : V \subseteq \mathbb{R}^n \to \mathbb{R}^m$. This function is not unique, many choices will lead to the same S, so is there a better $\mathbf{f}$ to choose out of all possibilities? Unanswerable since there is no definition of 'better'. But by the methods of the last section we can ensure that $J\mathbf{f}(\mathbf{x})$ has a particular form.

Take a point $\mathbf{p} \in S$ for which the $m \times n$ matrix $J\mathbf{f}(\mathbf{p})$ has full rank m. In particular $J\mathbf{f}(\mathbf{p})$ has m linearly independent columns and by permuting the variables in the domain $\mathbb{R}^n$ if necessary, we can assume that the *last* m columns of $J\mathbf{f}(\mathbf{p})$ are linearly independent.

Write $J\mathbf{f}(\mathbf{p}) = (E \mid F)$ where F is an $m \times m$ matrix. Since the m columns of F are linearly independent it is invertible with inverse H. Then

$$HJ\mathbf{f}(\mathbf{p}) = H\,(E \mid F) = (HE \mid I_m) = (G \mid I_m)\,,$$

say, where G is a $m \times (n-m)$ matrix.

The matrix H is associated with the linear map $\mathbf{L}_H : \mathbb{R}^m \to \mathbb{R}^m, \mathbf{t} \to H\mathbf{t}$. Recall, from the last Chapter, that the Jacobian matrix of a linear function equals the matrix associated with the function, i.e. $J\mathbf{L}_H(\mathbf{a}) = H$ for all $\mathbf{a} \in \mathbb{R}^m$. Consider the composite

$$\widetilde{\mathbf{f}} = \mathbf{L}_H \circ \mathbf{f} : V \subseteq \mathbb{R}^n \to \mathbb{R}^m.$$

Since $\mathbf{L}_H$ is a linear map $\mathbf{L}_H(\mathbf{x}) = \mathbf{0}$ iff $\mathbf{x} = \mathbf{0}$. Thus

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \quad \text{iff} \quad \mathbf{L}_H(\mathbf{f}(\mathbf{x})) = \mathbf{0} \quad \text{i.e.} \quad \mathbf{L}_H \circ \mathbf{f}(\mathbf{x}) = \mathbf{0}.$$

That is, $(\mathbf{L}_H \circ \mathbf{f})^{-1}(\mathbf{0}) = \mathbf{f}^{-1}(\mathbf{0})$, i.e. the level sets are equal.

By the Chain Rule, *for matrices,* we find that

$$J\widetilde{\mathbf{f}}(\mathbf{p}) = J(\mathbf{L}_H \circ \mathbf{f})(\mathbf{p}) = J\mathbf{L}_H(\mathbf{b})\,J\mathbf{f}(\mathbf{p}) = HJ\mathbf{f}(\mathbf{p}) = (G \mid I_m)\,.$$

So it is possible, when necessary in a proof, to assume that $J\mathbf{f}(\mathbf{p})$ not only has full rank but also has the form $(G \mid I_m)$ for some $m \times (n-m)$ matrix G.

12. Proof of the Inverse Function Theorem

Inverse Function Theorem *Suppose that* $\mathbf{f} : U \to \mathbb{R}^n$ *is a* $\mathcal{C}^1$*-function on an open subset* $U \subseteq \mathbb{R}^n$ *such that for some* $\mathbf{a} \in U$ *the Jacobian matrix* $J\mathbf{f}(\mathbf{a})$ *is of full rank then* $\mathbf{f}$ *is locally invertible. This means that there exists an open set* $V \subseteq U$ *with* $\mathbf{a} \in V$*, such that*

1. $\mathbf{f} : V \to \mathbf{f}(V)$ *is a bijection,*
2. $\mathbf{f}(V)$ *is an open subset of* $\mathbb{R}^n$*,*
3. *the inverse function* $\mathbf{g} = \mathbf{f}^{-1} : \mathbf{f}(V) \to V$ *is* C^1,
4. $d\mathbf{g}_{\mathbf{b}} = d\mathbf{f}_{\mathbf{a}}^{-1}$ *(or* $J\mathbf{g}(\mathbf{b}) = J\mathbf{f}(\mathbf{a})^{-1}$*) where* $\mathbf{b} = \mathbf{f}(\mathbf{a})$*.*

Proof Given $\mathbf{f} : U \to \mathbb{R}^n$, a $\mathcal{C}^1$-function on an open subset $U \subseteq \mathbb{R}^n$, define $\mathbf{h} : \mathbb{R}^n \times U \subseteq \mathbb{R}^{2n} \to \mathbb{R}^n$ by first writing $\mathbf{w} \in \mathbb{R}^n \times U$ as

$$\mathbf{w} = \begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix},$$

with $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in U$. Then set

$$\mathbf{h}(\mathbf{w}) = \mathbf{h}\left(\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}\right) = \mathbf{x} - \mathbf{f}(\mathbf{y}).$$

The Jacobian matrix of $\mathbf{h}$ is $J\mathbf{h}(\mathbf{w}) = (I_n \,|\, -J\mathbf{f}(\mathbf{y}))$. Here $J\mathbf{f}(\mathbf{y})$ is an $n \times n$ matrix.

We are given $\mathbf{a} \in U$, a point at which $J\mathbf{f}(\mathbf{a})$ is non-singular, so let

$$\mathbf{p} = \begin{pmatrix} \mathbf{f}(\mathbf{a}) \\ \mathbf{a} \end{pmatrix} \in \mathbb{R}^n \times U.$$

Then $\mathbf{h}(\mathbf{p}) = \mathbf{f}(\mathbf{a}) - \mathbf{f}(\mathbf{a}) = 0$ and $J\mathbf{h}(\mathbf{p}) = (I_n | - J\mathbf{f}(\mathbf{a}))$ has full-rank because of the occurrence of the identity matrix. But further, because of occurrence of the $-J\mathbf{f}(\mathbf{a})$, a non-singular matrix, the last n columns of $J\mathbf{h}(\mathbf{p})$, i.e. **all** the columns, are linearly independent. So we can apply the Implicit Function Theorem and no permutation of coordinates is required. Thus there exists

- an open set $A \subseteq \mathbb{R}^n$ containing $\mathbf{f}(\mathbf{a})$,
- a $\mathcal{C}^1$-function $\mathbf{g} : A \to \mathbb{R}^n$,
- an open set $B \subseteq \mathbb{R}^n \times U$ containing $\mathbf{p}$,

such for $\left(\mathbf{x}^T, \mathbf{y}^T\right)^T \in B$, ($\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^n$),

$$\mathbf{h}\left(\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}\right) = \mathbf{0} \text{ if, and only if, } \mathbf{x} \in A \text{ and } \mathbf{y} = \mathbf{g}(\mathbf{x})\,.$$

Note that the set B can be written as $A \times V$ for some open set $V \subseteq U \subseteq \mathbb{R}^n$. Thus we have $\mathbf{x} - \mathbf{f}(\mathbf{y}) = \mathbf{0}$, i.e. $\mathbf{x} = \mathbf{f}(\mathbf{y})$ with $\mathbf{y} \in V$ if, and only if, $\mathbf{y} = \mathbf{g}(\mathbf{x})$ with $\mathbf{x} \in A$. This gives the existence of a $\mathcal{C}^1$-inverse.

That $d\mathbf{g}_\mathbf{b} = d\mathbf{f}_\mathbf{a}^{-1}$ follows from the Chain Rule. ■