3 Part 2. Surfaces & planes **2021-22**

3.9 Tangent Spaces and Planes

Definition 1 *A **(differentiable) curve in** $\mathbb{R}^n$ is the image of a $\mathcal{C}^1$-function $\boldsymbol{\gamma}: I \to \mathbb{R}^n$, $t \mapsto \left(\gamma^1(t), \gamma^2(t), ..., \gamma^n(t)\right)^T$, from an interval $I \subseteq \mathbb{R}$.*
*If the $\boldsymbol{\gamma}$ is injective we say the curve is **simple**.*

There is a misuse of notation for I use $\boldsymbol{\gamma}$ for both the function and the image of the function.

Definition 2 *A differentiable curve is **regular** if $\boldsymbol{\gamma}'(t) \neq \mathbf{0}$ for all $t \in I$.*
*A **singular point** is one where $\boldsymbol{\gamma}'(t) = \mathbf{0}$.*

In the next definition I have kept the properties of the set S vague (other than it contains a curve), but think of it as a surface.

Definition 3 *Let $S \subseteq \mathbb{R}^n$, $\mathbf{p} \in S$ and let $\boldsymbol{\gamma}: (-1, 1) \to S$ be a regular curve lying within S such that $\boldsymbol{\gamma}(0) = \mathbf{p}$. Then*

- $\boldsymbol{\gamma}'(0)$ *is **a tangent vector to** S **at** $\mathbf{p}$,*
- $\mathbf{p} + \boldsymbol{\gamma}'(0)\, t$, $t \in \mathbb{R}$, *is **a tangent line to** S **at** $\mathbf{p}$.*

Definition 4 *The **Tangent Space of** S **at** $\mathbf{p}$ is the set of **all** tangent vectors to S at $\mathbf{p} \in S$ and is denoted by $T_{\mathbf{p}}(S)$.*

You should think of the vectors $\mathbf{u}$ in the Tangent space as vectors with a point of application $\mathbf{p}$. They might more correctly be written as $\mathbf{u}_{\mathbf{p}}$, though we will not do this.

We call this the Tangent *Space*, but we should prove that it is, in fact, a vector space. We know that at every surface is locally a graph so it suffices to prove the following.

Theorem 5 *Let the surface $S \subseteq \mathbb{R}^n$ be given by a graph, so*

$$S = \left\{ \begin{pmatrix} \mathbf{u} \\ \boldsymbol{\phi}(\mathbf{u}) \end{pmatrix} : \mathbf{u} \in U \right\},$$

for some $\mathcal{C}^1$-function, $\boldsymbol{\phi}: U \subseteq \mathbb{R}^r \to \mathbb{R}^{n-r}$. Let $\mathbf{p} \in S$, in which case

$$\mathbf{p} = \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\phi}(\mathbf{q}) \end{pmatrix}$$

for some $\mathbf{q} \in U$. *Then*

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} \mathbf{t} \\ J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t} \end{pmatrix} : \mathbf{t} \in \mathbb{R}^r \right\}. \tag{1}$$

Proof Recall the notation

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

(in fact we used $\mathbf{F}_{\boldsymbol{\phi}}$ but I now drop the subscript) and the fact that

$$J\mathbf{F}(\mathbf{u}) = \begin{pmatrix} I_r \\ J\boldsymbol{\phi}(\mathbf{u}) \end{pmatrix}. \tag{2}$$

In this notation (1) becomes

$$T_{\mathbf{p}}(S) = \{ J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r \}. \tag{3}$$

The proof of such a set equality requires the proofs of two set inclusions.

($\supseteq$) Let $\mathbf{v} \in \{ J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r \}$, in which case there exists $\mathbf{y} \in \mathbb{R}^r$ such that $\mathbf{v} = J\mathbf{F}(\mathbf{q})\,\mathbf{y}$. To conclude that $\mathbf{v} \in T_{\mathbf{p}}(S)$ we have to find a curve $\boldsymbol{\gamma}$ lying in S such that $\boldsymbol{\gamma}(0) = \mathbf{p}$ and $\boldsymbol{\gamma}'(0) = \mathbf{v}$.

Since $\mathbf{q} \in U$ and U is an open set there exists $\eta > 0$ such that if $|t| < \eta$ then $\mathbf{q} + t\mathbf{y} \in U$. For such t define $\boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^n$, by

$$\boldsymbol{\gamma}(t) = \mathbf{F}(\mathbf{q} + t\mathbf{y}).$$

This is a curve lying within S by the definition of S as the image of $\mathbf{F}$. Also $\boldsymbol{\gamma}(0) = \mathbf{F}(\mathbf{q}) = \mathbf{p}$. For the derivative, we have that $\boldsymbol{\gamma}$ is a composition when the Chain Rule gives

$$\boldsymbol{\gamma}'(t) = J\mathbf{F}(\mathbf{q} + t\mathbf{y})\,\frac{d}{dt}(\mathbf{q} + t\mathbf{y}) = J\mathbf{F}(\mathbf{q} + t\mathbf{y})\,\mathbf{y},$$

for $|t| < \eta$. Choosing $t = 0$ gives

$$\boldsymbol{\gamma}'(0) = J\mathbf{F}(\mathbf{q})\,\mathbf{y} = \mathbf{v},$$

since $\mathbf{v} = J\mathbf{F}(\mathbf{q})\,\mathbf{y}$ from above. Hence $\mathbf{v} \in T_{\mathbf{p}}(S)$. True for all such $\mathbf{v}$ means

$$T_{\mathbf{p}}(S) \supseteq \{ J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r \}.$$

($\subseteq$) Let $\mathbf{v} \in T_{\mathbf{p}}(S)$. So there exists $\boldsymbol{\gamma} : (-1,1) \to S$, a regular curve such that $\boldsymbol{\gamma}(0) = \mathbf{p}$ and $\boldsymbol{\gamma}'(0) = \mathbf{v}$. Write

$$\boldsymbol{\gamma}(t) = \begin{pmatrix} \boldsymbol{\alpha}(t) \\ \boldsymbol{\beta}(t) \end{pmatrix},$$

where, in the components of $\boldsymbol{\gamma}$,

$$\boldsymbol{\alpha}(t) = \begin{pmatrix} \gamma^1(t) \\ \vdots \\ \gamma^r(t) \end{pmatrix} \in \mathbb{R}^r \quad \text{and} \quad \boldsymbol{\beta}(t) = \begin{pmatrix} \gamma^{r+1}(t) \\ \vdots \\ \gamma^n(t) \end{pmatrix} \in \mathbb{R}^{n-r}.$$

But $\boldsymbol{\gamma}(t) \in S$ means that

$$\boldsymbol{\gamma}(t) = \begin{pmatrix} \boldsymbol{\alpha}(t) \\ \boldsymbol{\beta}(t) \end{pmatrix} = \begin{pmatrix} \boldsymbol{\alpha}(t) \\ \boldsymbol{\phi}(\boldsymbol{\alpha}(t)) \end{pmatrix} = \mathbf{F}(\boldsymbol{\alpha}(t)).$$

Importantly $\boldsymbol{\gamma}$ is a $\mathcal{C}^1$ function means that $\boldsymbol{\alpha}$ is a $\mathcal{C}^1$ function. Thus we can apply the Chain Rule

$$\boldsymbol{\gamma}'(t) = \frac{d}{dt}\mathbf{F}(\boldsymbol{\alpha}(t)) = J\mathbf{F}(\boldsymbol{\alpha}(t))\,\frac{d}{dt}\boldsymbol{\alpha}(t).$$

Choose $t = 0$. Note that

$$\begin{pmatrix} \mathbf{q} \\ \boldsymbol{\phi}(\mathbf{q}) \end{pmatrix} = \mathbf{p} = \boldsymbol{\gamma}(0) = \begin{pmatrix} \boldsymbol{\alpha}(0) \\ \boldsymbol{\beta}(0) \end{pmatrix}$$

and so $\boldsymbol{\alpha}(0) = \mathbf{q}$. Therefore

$$\mathbf{v} = \boldsymbol{\gamma}'(0) = J\mathbf{F}(\boldsymbol{\alpha}(0))\,\boldsymbol{\alpha}'(0) = J\mathbf{F}(\mathbf{q})\,\boldsymbol{\alpha}'(0).$$

Here $\boldsymbol{\alpha}'(0) \in \mathbb{R}^r$ so $\mathbf{v} = J\mathbf{F}(\mathbf{q})\,\mathbf{t}$ for some $\mathbf{t} \in \mathbb{R}^r$. Hence

$$T_{\mathbf{p}}(S) \subseteq \{J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\}.$$

Finally we must have equality of sets. ■

The conclusion of Theorem 5 may be written in a number of ways.

> If a surface is given as a graph of the function $\boldsymbol{\phi}$ then the Tangent Space is
>
> i. the graph of the linear function $\mathbf{t} \mapsto J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t}$.
>
> ii. the image of the linear map $\mathbf{t} \mapsto J\mathbf{F}(\mathbf{q})\,\mathbf{t}$.

Corollary 6 *Let $S \subseteq \mathbb{R}^n$ be a surface given by a graph of a $\mathcal{C}^1$-function $\boldsymbol{\phi} : U \subseteq \mathbb{R}^r \to \mathbb{R}^{n-r}$. Let $\mathbf{p} \in S$. Then $T_{\mathbf{p}}(S)$ is a vector space of dimension r with a basis of the columns of*

$$\begin{pmatrix} I_r \\ J\boldsymbol{\phi}(\mathbf{q}) \end{pmatrix}.$$

Proof immediate. ■

Now that we know that $T_{\mathbf{p}}(S)$ is a vector space we can make the following definition.

Definition 7 ***The Tangent Plane to** S **at** $\mathbf{p}$ is the set of all tangent lines to S at $\mathbf{p}$. Equivalently, this is*

$$\mathbf{p} + T_{\mathbf{p}}(S) = \{\mathbf{p} + \mathbf{v} : \mathbf{v} \in T_{\mathbf{p}}(S)\}.$$

Corollary 8 *With the notation of Theorem 5 the Tangent Plane to S, the graph of $\boldsymbol{\phi}$, at $\mathbf{p} \in S$ is the graph of the affine function*

$$\mathbf{u} \mapsto \phi(\mathbf{q}) + J\phi(\mathbf{q})(\mathbf{u} - \mathbf{q}).$$

Proof From (1) and the definition of the Tangent plane at $\mathbf{p} \in S$,

$$\begin{aligned} \mathbf{p} + T_{\mathbf{p}}(S) &= \left\{ \begin{pmatrix} \mathbf{q} \\ \phi(\mathbf{q}) \end{pmatrix} + \begin{pmatrix} \mathbf{I}_r \\ J\phi(\mathbf{q}) \end{pmatrix} \mathbf{t} : \mathbf{t} \in \mathbb{R}^r \right\} \\ &= \left\{ \begin{pmatrix} \mathbf{q} + \mathbf{t} \\ \phi(\mathbf{q}) + J\phi(\mathbf{q})\,\mathbf{t} \end{pmatrix} : \mathbf{t} \in \mathbb{R}^r \right\} \\ &= \left\{ \begin{pmatrix} \mathbf{u} \\ \phi(\mathbf{q}) + J\phi(\mathbf{q})(\mathbf{u} - \mathbf{q}) \end{pmatrix} : \mathbf{u} \in \mathbb{R}^r \right\}, \end{aligned}$$

for as $\mathbf{t}$ ranges over $\mathbb{R}^r$ then so does $\mathbf{u} = \mathbf{q} + \mathbf{t}$. ■

Definition 9 *If $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is Fréchet differentiable at $\mathbf{a} \in U$ then the **best affine approximation** to $\mathbf{f}$ at $\mathbf{a}$ is*

$$\mathbf{f}(\mathbf{a}) + J\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}),$$

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

Why is it the **best**? Let $\mathbf{A}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + J\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})$. This is an affine function and **an** approximation to $\mathbf{f}$ in that

$$\lim_{\mathbf{x}\to\mathbf{a}} (\mathbf{f}(\mathbf{x}) - \mathbf{A}(\mathbf{x})) = \mathbf{0},$$

But, since $\mathbf{f}$ is Fréchet differentiable at $\mathbf{a}$, we have

$$\lim_{\mathbf{t}\to\mathbf{0}} \left(\frac{\mathbf{f}(\mathbf{a}+\mathbf{t}) - \mathbf{f}(\mathbf{a}) - J\mathbf{f}(\mathbf{a})\mathbf{t}}{|\mathbf{t}|} \right) = \mathbf{0} \tag{4}$$

Also, by the uniqueness of the Fréchet derivative, $J\mathbf{f}(\mathbf{a})\,\mathbf{t}$ is the **only** linear function for which (4) holds. Thus $\mathbf{A}(\mathbf{x})$ is the **only** affine function for which

$$\lim_{\mathbf{x}\to\mathbf{a}} \left(\frac{\mathbf{f}(\mathbf{x}) - \mathbf{A}(\mathbf{x})}{|\mathbf{x}-\mathbf{a}|} \right) = \mathbf{0}.$$

It is for this reason it is called the **best** affine approximation.

> Corollary 8 says that the Tangent Plane to the graph of $\boldsymbol{\phi}$ at $\mathbf{p}$ is the graph of the best affine approximation to $\boldsymbol{\phi}$ at $\mathbf{p}$.

Example 10 *Let* $\boldsymbol{\phi} : \mathbb{R}^2 \to \mathbb{R}^2$ *be given by*

$$\boldsymbol{\phi}(\mathbf{x}) = \begin{pmatrix} x^4 - y^3 \\ xy \end{pmatrix}.$$

Find the Tangent Space and Tangent Plane to the graph of $\boldsymbol{\phi}$ *at* $\mathbf{p} = (2,2,8,4)^T$, *a point on the graph. Give your answers as graphs. Can you give the answers as level sets?*

Solution *in Problems Class* First note that the point $\mathbf{p}$ can be written as

$$\mathbf{p} = \begin{pmatrix} 2 \\ 2 \\ 6 \\ 4 \end{pmatrix} = \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\phi}(\mathbf{q}) \end{pmatrix},$$

with $\mathbf{q} = (2,2)^T$.

Next, from above the **Tangent Space** is the graph of the linear function $\mathbf{x} \mapsto J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{x}$ while the Tangent Plane is the graph of the best affine approximation to $\boldsymbol{\phi}$ at $\mathbf{q}$.

For both the Space and Plane we need

$$J\boldsymbol{\phi}(\mathbf{q}) = \left.\begin{pmatrix} 4x^3 & -3y^2 \\ y & x \end{pmatrix}\right|_{\mathbf{x}=\mathbf{q}} = \begin{pmatrix} 32 & -12 \\ 2 & 2 \end{pmatrix}.$$

Thus the Tangent space is the graph

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} \mathbf{x} \\ J\phi(\mathbf{q})\,\mathbf{x} \end{pmatrix} : \mathbf{x} \in \mathbb{R}^2 \right\}.$$

Here

$$J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{x} = \begin{pmatrix} 32 & -12 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 32x - 12y \\ 2x + 2y \end{pmatrix}.$$

Hence

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} x \\ y \\ 32x - 12y \\ 2x + 2y \end{pmatrix} : x, y \in \mathbb{R} \right\} = \left\{ x \begin{pmatrix} 1 \\ 0 \\ 32 \\ 2 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ -12 \\ 2 \end{pmatrix} : x, y \in \mathbb{R} \right\}. \tag{5}$$

Thus $(1, 0, 32, 2)^T$, $(0, 1, -12, 2)^T$, the columns of $\left(\begin{smallmatrix} I_2 \\ J\boldsymbol{\phi}(\mathbf{q}) \end{smallmatrix}\right)$, form a basis for $T_{\mathbf{p}}(S)$.

We have the result that the **Tangent plane** is the graph of the best affine approximation to $\boldsymbol{\phi}$ at $\mathbf{q}$. This approximation is

$$\begin{aligned} \boldsymbol{\phi}(\mathbf{q}) + J\boldsymbol{\phi}(\mathbf{q})(\mathbf{x} - \mathbf{q}) &= \begin{pmatrix} 6 \\ 4 \end{pmatrix} + \begin{pmatrix} 32 & -12 \\ 2 & 2 \end{pmatrix} \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \end{pmatrix} \right) \\ &= \begin{pmatrix} 32x - 12y - 34 \\ 2x + 2y - 4 \end{pmatrix} \end{aligned}$$

Hence the Tangent plane, which is the *graph* of the best affine approximation, is

$$\left\{ \begin{pmatrix} x \\ y \\ 32x - 12y - 34 \\ 2x + 2y - 4 \end{pmatrix} : x, y \in \mathbb{R} \right\}. \tag{6}$$

As a **level set** the Tangent *Space* is, from (5), the set of points $(x, y, u, v)^T \in \mathbb{R}^4$ satisfying

$$\begin{aligned} 32x - 12y - u &= 0 \\ 2x + 2y - v &= 0. \end{aligned}$$

As a **level set** the Tangent *Plane* is, from (6), the set of points from $\mathbb{R}^4$ satisfying

$$\begin{gathered} 32x - 12y - u = 34 \\ 2x + 2y - v = 4. \end{gathered}$$

■

Note Recall that the definition of the Tangent *Plane* is $\mathbf{p} + T_{\mathbf{p}}(S)$. In the above example this would lead, from (5) , to

$$\left\{ \begin{pmatrix} x+2 \\ y+2 \\ 32x - 12y + 8 \\ 2x + 2y + 4 \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

Make sure you understand why this is the same set of points as (6). An advantage of (6) is that it is simpler to use it to derive the level set.

In conclusion

> If a surface is given as a **graph** of ϕ then
>
> - the Tangent Space is given by the **graph** of the linear map $d\phi_{\mathbf{q}} : \mathbf{t} \mapsto J\phi(\mathbf{q})\,\mathbf{t}$.
> - the Tangent Plane is given by the **graph** of the best affine approximation to ϕ at $\mathbf{q}$, i.e. $\phi(\mathbf{q}) + J\phi(\mathbf{q})(\mathbf{t} - \mathbf{q})$.

We can now find the Tangent Space and plane for a surface given as a graph. We will use this to find the Space and plane for a surface either given as a level set or parametrically.

3.10 Tangent Space for a Level Set Surface

Theorem 11 *Let the surface S be given as a level set*

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

for some C^1-function $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$. Let $\mathbf{p} \in S$. Then

$$T_{\mathbf{p}}(S) = \{ \mathbf{x} \in \mathbb{R}^n : J\mathbf{f}(\mathbf{p})\, \mathbf{x} = \mathbf{0} \}, \tag{7}$$

and the rows of $J\mathbf{f}(\mathbf{p})$ are a basis for $T_{\mathbf{p}}(S)^{\perp}$.

Proof *not given, see Appendix*.

Idea By, if necessary, relabelling the axes in $\mathbb{R}^n$ we can assume that the final m columns of $J\mathbf{f}(\mathbf{p})$ are linearly independent. Then, by the Implicit Function Theorem there exist open sets $V \subseteq \mathbb{R}^{n-m}$, and $W \subseteq U \subseteq \mathbb{R}^n$ with $\mathbf{p} \in W$, along with $\boldsymbol{\phi} : V \to \mathbb{R}^m$ such that

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

Theorem 5 then implies

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} \mathbf{t} \\ J\boldsymbol{\phi}(\mathbf{q})\, \mathbf{t} \end{pmatrix} : \mathbf{t} \in \mathbb{R}^{n-m} \right\},$$

where $\mathbf{q} \in V$ satisfies $\mathbf{p} = \left(\begin{smallmatrix} \mathbf{q} \\ \boldsymbol{\phi}(\mathbf{q}) \end{smallmatrix} \right)$.

All that remains is to show that this equals the right hand side of (7). For $\mathbf{v} \in V$ we have $\left(\begin{smallmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{smallmatrix} \right) \in S$ and so

$$f\left(\begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} \right) = 0.$$

Apply the Chain Rule; see Appendix.

The fact that the rows of $J\mathbf{f}(\mathbf{p})$ are a basis for $T_{\mathbf{p}}(S)^{\perp}$ follows from earlier results on planes. ■

Corollary 12 *Under the conditions of Theorem 11 the Tangent Plane to S at $\mathbf{p}$ is*

$$\{ \mathbf{p} + \mathbf{v} : \mathbf{v} \in \mathbb{R}^n \text{ and } J\mathbf{f}(\mathbf{p})\, \mathbf{v} = \mathbf{0} \} = \{ \mathbf{x} \in \mathbb{R}^n : J\mathbf{f}(\mathbf{p})(\mathbf{x} - \mathbf{p}) = \mathbf{0} \}.$$

Example 13 *Consider the level set of points in* $\mathbb{R}^4$ *satisfying*

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

Find the Tangent Space and plane at $\mathbf{p} = (1, 0, -1, 2)^T$.

Solution *in Problems Class* The Jacobian matrix at $\mathbf{p}$ is

$$J\mathbf{f}(\mathbf{p}) = \begin{pmatrix} 6 & 0 & -4 & 4 \\ 2 & 3 & 2 & 0 \end{pmatrix}. \tag{8}$$

This is of full-rank and so $T_{\mathbf{p}}(S)$ is the set of $\mathbf{x} = (x, y, u, v) \in \mathbb{R}^4$ such that $J\mathbf{f}(\mathbf{p})\,\mathbf{x} = \mathbf{0}$, i.e.

$$\begin{aligned} 6x - 4u + 4v = 0 \\ 2x + 3y + 2u = 0. \end{aligned} \tag{9}$$

The Tangent Plane is the set of $\mathbf{x}$ such that $J\mathbf{f}(\mathbf{p})\,(\mathbf{x} - \mathbf{p}) = \mathbf{0}$, i.e.

$$\begin{aligned} 6x - 4u + 4v &= 18 \\ 2x + 3y + 2u &= 0. \end{aligned}$$

■

Aside *Left to Tutorial* As noted above the rows of $J\mathbf{f}(\mathbf{p})$ form a basis of the Normal Space $T_{\mathbf{p}}(S)^{\perp}$. So in this example $(6, 0, -4, 4)^T$ and $(2, 3, 2, 0)^T$ are a basis for $T_{\mathbf{p}}(S)^{\perp}$. How to use a basis for $T_{\mathbf{p}}(S)^{\perp}$ to find a basis for $T_{\mathbf{p}}(S)$?

Method 1: The Implicit Function Theorem tells us the level set is locally the graph of a function ϕ. Then the columns of $\begin{pmatrix} I_{n-m} \\ J\boldsymbol{\phi}(\mathbf{q}) \end{pmatrix}$ give a basis of $T_{\mathbf{p}}(S)$. But the Implicit Function Theorem is an existence result, it tells us that $\boldsymbol{\phi}$ exists but not what it is. Nonetheless we can still calculate $J\boldsymbol{\phi}(\mathbf{q})$.

We do this for the example above. Since the last two columns of (8) are linearly independent the u and v can be given as $\mathcal{C}^1$-functions of x and y (in fact $u = \phi^1$ and $v = \phi^2$).

Differentiate $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ with respect to x.

$$\begin{aligned} 2x - 2\frac{\partial u}{\partial x}v - 2u\frac{\partial v}{\partial x} + 2v + 2x\frac{\partial v}{\partial x} &= 0 \\ 2y - \frac{\partial u}{\partial x}y + \frac{\partial v}{\partial x}x + v + \frac{\partial u}{\partial x}v + u\frac{\partial v}{\partial x} &= 0. \end{aligned}$$

Evaluate at $\mathbf{p} = (1, 0, -1, 2)^T$.

$$\begin{aligned} 2 - 4\frac{\partial u}{\partial x} + 2\frac{\partial v}{\partial x} + 4 + 2\frac{\partial v}{\partial x} &= 0 \\ \frac{\partial v}{\partial x} + 2 + 2\frac{\partial u}{\partial x} - \frac{\partial v}{\partial x} &= 0. \end{aligned}$$

Solve,

$$\frac{\partial u}{\partial x} = -1 \quad \text{and} \quad \frac{\partial v}{\partial x} = -\frac{5}{2}.$$

Repeat but with respect to y to find

$$\frac{\partial u}{\partial y} = -\frac{3}{2} \quad \text{and} \quad \frac{\partial v}{\partial y} = -\frac{3}{2}.$$

Then, in this example

$$\begin{pmatrix} I_{n-m} \\ J\boldsymbol{\phi}(\mathbf{q}) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -3/2 \\ -5/2 & -3/1 \end{pmatrix}.$$

Therefore $(1, 0, -1, -5/2)^T$ and $(0, 1, -3/2, -3/2)^T$ are a basis for $T_{\mathbf{p}}(S)$.

Perhaps a 'better' pair might be $(-2, 0, 2, 5)^T$ and $(0, -2, 3, 3)^T$. (Check that these two vectors are both orthogonal to $(6, 0, -4, 4)^T$ and $(2, 3, 2, 0)^T$).

Method 2: Instead of using the Implicit Function Theorem to say the original level set is locally a graph write the solution set as a graph. For example solve (9) for u and v as functions of x and y, i.e. add twice the second equation to the first to get

$$\begin{aligned} 10x + 6y + 4v &= 0 \\ 2x + 3y + 2u &= 0. \end{aligned}$$

We can then express the Tangent Space as the graph

$$\left\{ \begin{pmatrix} x \\ y \\ -x - 3y/2 \\ -5x/2 - 3y/2 \end{pmatrix} : \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \right\}.$$

But

$$\begin{pmatrix} x \\ y \\ -x-3y/2 \\ -5x/2-3y/2 \end{pmatrix} = x\begin{pmatrix} 1 \\ 0 \\ -1 \\ -5/2 \end{pmatrix} + y\begin{pmatrix} 0 \\ 1 \\ -3/2 \\ -3/2 \end{pmatrix}.$$

This method is formalised in the Appendix.

Why does the second method appear to be quicker? Because it uses Theorem 11; the proof of this requires the Chain Rule which is what we essentially did in Method 1.

End of aside.

For a further example

Example 14 *Let $S \subseteq \mathbb{R}^4$ be given by the system of equations*

$$\begin{aligned} x^4 - y^3 - u &= 0 \\ xy - v &= 0, \end{aligned}$$

for $(x, y, u, v)^T \in \mathbb{R}^4$. Find the Tangent plane at $\mathbf{p} = (2, 2, 8, 4)^T \in S$.

Solution *in Problems Class.* The Jacobian matrix is

$$\begin{pmatrix} 4x^3 & -3y^2 & -1 & 0 \\ y & x & 0 & -1 \end{pmatrix}_{\mathbf{x}=\mathbf{p}} = \begin{pmatrix} 32 & -12 & -1 & 0 \\ 2 & 2 & 0 & -1 \end{pmatrix}$$

at $(2, 2, 8, 4)^T$. This is well-defined and of full-rank. Then the Tangent Plane is the set of points $\mathbf{x} = (x, y, u, v)^T$ satisfying $J\mathbf{f}(\mathbf{p})(\mathbf{x} - \mathbf{p}) = \mathbf{0}$, i.e.

$$\mathbf{0} = \begin{pmatrix} 32 & -12 & -1 & 0 \\ 2 & 2 & 0 & -1 \end{pmatrix}\begin{pmatrix} x-2 \\ y-2 \\ u-8 \\ v-4 \end{pmatrix} = \begin{pmatrix} 32x-12y-u-32 \\ 2x+2y-v-4 \end{pmatrix}.$$

Hence the tangent plane is the solution set to the system

$$32x - 12y - u = 32,$$

$$2x + 2y - v = 4.$$

■

Aside In this example $(32, -12, -1, 0)^T$ and $(2, 2, 0, -1)^T$ are a basis for $T_{\mathbf{p}}(S)^{\perp}$. As in the aside above, we can find $(1, 0, 32, 2)^T$ and $(0, 1, -12, 2)^T$ as a basis for $T_{\mathbf{p}}(S)$.

End of aside.

For a conclusion recall that if $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ then $\mathbf{x}$ is in the kernel of $\mathbf{f}$. In this language we have

> If the surface is given as the **kernel** of the map $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x})$ then
>
> - the Tangent Space is the **kernel** of the linear map $d\mathbf{f}_{\mathbf{p}} : \mathbf{x} \mapsto J\mathbf{f}(\mathbf{p})\,\mathbf{x}$.
> - the Tangent Plane is the **kernel** of the affine map $\mathbf{x} \mapsto J\mathbf{f}(\mathbf{p})(\mathbf{x} - \mathbf{p})$.

For an alternative conclusion note that if $\mathbf{x}$ satisfies $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ then $\mathbf{x}$ is satisfying a system of equations $f^i(\mathbf{x}) = \mathbf{0}$ for $1 \leq i \leq m$. Recall that the rows of a Jacobian matrix are the transposes of the gradient vector for each component function f^i. Thus if $\mathbf{x} \in T_{\mathbf{p}}(S)$, so $J\mathbf{f}(\mathbf{p})\,\mathbf{x} = \mathbf{0}$, then $\nabla f^i(\mathbf{p})^T\,\mathbf{x} = \mathbf{0}$, i.e. $\nabla f^i(\mathbf{p}) \bullet \mathbf{x} = \mathbf{0}$ for $1 \leq i \leq m$. Each of these is a linear equation in $\mathbf{x}$. Thus

> If a surface is given as a **system of equations**, $f^i(\mathbf{x}) = \mathbf{0}$ for $1 \leq i \leq m$ then
>
> - the Tangent Space can be written as a system of linear equations $\nabla f^i(\mathbf{p}) \bullet \mathbf{x} = \mathbf{0}$ for $1 \leq i \leq m$.
> - the Tangent Plane can be written as a system of linear equations $\nabla f^i(\mathbf{p}) \bullet (\mathbf{x} - \mathbf{p}) = \mathbf{0}$ for $1 \leq i \leq m$.

3.11 Tangent Space for a Parametric Surface.

Theorem 5 can be written as

$$S = \{\mathbf{F}(\mathbf{u}) : \mathbf{u} \in U\} \implies T_{\mathbf{p}}(S) = \{J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\},$$

in the particular case when $\mathbf{F}$ is the image of a graph of a function $\boldsymbol{\phi}$. For a general $\mathbf{F}$ we have to use the fact that a surface given parametrically, i.e. as the image of a function, is locally a graph. We can then apply Theorem 5.

Theorem 15 *Let the surface S be given parametrically as*

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

of a $\mathcal{C}^1$-function $\mathbf{F} : U \subseteq \mathbb{R}^r \to \mathbb{R}^n$. Let $\mathbf{p} \in S$, so $\mathbf{p} = \mathbf{F}(\mathbf{q})$ for some $\mathbf{q} \in U$. Then

$$T_{\mathbf{p}}(S) = \{J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\}, \tag{10}$$

and the columns of $J\mathbf{F}(\mathbf{q})$ are a basis for $T_{\mathbf{p}}(S)$.

Proof By relabelling if necessary the axes in $\mathbb{R}^n$ we can assume that the first r rows of $J\mathbf{F}(\mathbf{q})$ are linearly independent. Then, by the Inverse Function Theorem there exist open sets $V : \mathbf{q} \in V \subseteq U$ and $T \subseteq \mathbb{R}^r$ with a $\mathcal{C}^1$-function $\boldsymbol{\phi} : T \to \mathbb{R}^{n-r}$:

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

Since $\mathbf{q} \in V$ we have that $\mathbf{p} = \mathbf{F}(\mathbf{q})$ lies in this graph, i.e.

$$\mathbf{p} = \begin{pmatrix} \mathbf{r} \\ \boldsymbol{\phi}(\mathbf{r}) \end{pmatrix}$$

for some $\mathbf{r} \in T$. Theorem 5 then implies

$$T_{\mathbf{p}}(S) = \left\{\begin{pmatrix} \mathbf{y} \\ J\boldsymbol{\phi}(\mathbf{r})\,\mathbf{y} \end{pmatrix} : \mathbf{y} \in \mathbb{R}^r\right\}. \tag{11}$$

Recalling the proof of Corollary 23 in the previous notes we write

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

where $\mathbf{h} : U \subseteq \mathbb{R}^r \to \mathbb{R}^r$ and $\mathbf{k} : U \subseteq \mathbb{R}^r \to \mathbb{R}^{n-r}$. If $\mathbf{u} \in V$ then $\mathbf{F}(\mathbf{u})$ is given by the graph of $\boldsymbol{\phi}$ and so $\mathbf{k}(\mathbf{u}) = \boldsymbol{\phi}(\mathbf{h}(\mathbf{u}))$. Apply the Chain Rule to get

$$J\mathbf{k}(\mathbf{u}) = J\boldsymbol{\phi}(\mathbf{h}(\mathbf{u}))\,J\mathbf{h}(\mathbf{u}). \tag{12}$$

Combining some of the above steps we have

$$\begin{pmatrix} \mathbf{h}(\mathbf{q}) \\ \mathbf{k}(\mathbf{q}) \end{pmatrix} = \mathbf{F}(\mathbf{q}) = \mathbf{p} = \begin{pmatrix} \mathbf{r} \\ \boldsymbol{\phi}(\mathbf{r}) \end{pmatrix}.$$

In particular $\mathbf{h}(\mathbf{q}) = \mathbf{r}$. Choosing $\mathbf{u} = \mathbf{q}$ in (12) then gives $J\mathbf{k}(\mathbf{q}) = J\boldsymbol{\phi}(\mathbf{r})\, J\mathbf{h}(\mathbf{q})$.

The assumption that the first r rows of $J\mathbf{F}(\mathbf{q})$ are linearly independent means that $J\mathbf{h}(\mathbf{q})$ is an invertible matrix. Thus, within (11), we have

$$\begin{aligned}
\begin{pmatrix} \mathbf{y} \\ J\boldsymbol{\phi}(\mathbf{r})\,\mathbf{y} \end{pmatrix} &= \begin{pmatrix} I_r \\ J\boldsymbol{\phi}(\mathbf{r}) \end{pmatrix} \mathbf{y} = \begin{pmatrix} I_r \\ J\boldsymbol{\phi}(\mathbf{r}) \end{pmatrix} J\mathbf{h}(\mathbf{q}) J\mathbf{h}(\mathbf{q})^{-1}\mathbf{y} \\
&= \begin{pmatrix} J\mathbf{h}(\mathbf{q}) \\ J\boldsymbol{\phi}(\mathbf{r})\, J\mathbf{h}(\mathbf{q}) \end{pmatrix} J\mathbf{h}(\mathbf{q})^{-1}\mathbf{y} \\
&= \begin{pmatrix} J\mathbf{h}(\mathbf{q}) \\ J\mathbf{k}(\mathbf{q}) \end{pmatrix} J\mathbf{h}(\mathbf{q})^{-1}\mathbf{y} \\
&= J\mathbf{F}(\mathbf{q})\,\mathbf{x},
\end{aligned}$$

having changed variables from $\mathbf{y}$ to $\mathbf{x} = J\mathbf{h}(\mathbf{q})^{-1}\mathbf{y}$. Then the conclusion, (10), follows from (11).

The fact that the columns for $J\mathbf{F}(\mathbf{q})$ are a basis for $T_{\mathbf{p}}(S)$ follows from earlier work on planes. ■

Corollary 16 *Under the conditions of Theorem 15 the Tangent Plane to the image set of* $\mathbf{F}(\mathbf{t}), \mathbf{t} \in U$ *at* $\mathbf{p} = \mathbf{F}(\mathbf{q})$ *is the image set of the Best Affine Approximation to* $\mathbf{F}$ *at* $\mathbf{q}$.

Proof *Problems Class* By Theorem 15,

$$\begin{aligned}
\mathbf{p} + T_{\mathbf{p}}(S) &= \mathbf{F}(\mathbf{q}) + \{J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\} \\
&= \{\mathbf{F}(\mathbf{q}) + J\mathbf{F}(\mathbf{q})\,\mathbf{t} : \mathbf{t} \in \mathbb{R}^r\} \\
&= \{\mathbf{F}(\mathbf{q}) + J\mathbf{F}(\mathbf{q})\,(\mathbf{w} - \mathbf{q}) : \mathbf{w} \in \mathbb{R}^r\}, \qquad (13)
\end{aligned}$$

for as $\mathbf{t}$ ranges over $\mathbb{R}^{n-m}$ then so does $\mathbf{w} = \mathbf{t} + \mathbf{q}$. Here $\mathbf{F}(\mathbf{q}) + J\mathbf{F}(\mathbf{q})\,(\mathbf{w} - \mathbf{q})$ is the best affine approximation to $\mathbf{F}(\mathbf{w})$ at $\mathbf{w} = \mathbf{q}$. ■

Example 17 *Find the Tangent Space & Plane at* $(3,1,4,1)^T$ *to the parameterized surface*

$$\left\{ \begin{pmatrix} x^2+y \\ x-y^2 \\ 5-y+xy \\ 1+x+xy \end{pmatrix} : x^2+y^2<10 \right\}.$$

Solution *in Problems Class.* The point $\mathbf{p}=(3,1,4,1)^T$ on the surface arises from $\mathbf{q}=(2,-1)^T$. The Jacobian matrix at $\mathbf{q}$ is

$$J\mathbf{F}(\mathbf{q}) = \left.\begin{pmatrix} 2x & 1 \\ 1 & -2y \\ y & x-1 \\ 1+y & x \end{pmatrix}\right|_{\mathbf{x}=\mathbf{q}} = \begin{pmatrix} 4 & 1 \\ 1 & 2 \\ -1 & 1 \\ 0 & 2 \end{pmatrix}.$$

The columns are linearly independent so $J\mathbf{F}(\mathbf{q})$ is of full rank and the two columns span the Tangent Space. Hence

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} 4 & 1 \\ 1 & 2 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} : x,y\in\mathbb{R} \right\} = \left\{ \begin{pmatrix} 4x+y \\ x+2y \\ -x+y \\ 2y \end{pmatrix} : x,y\in\mathbb{R} \right\}.$$

The Tangent Plane is given parametrically as the best affine approximation to $\mathbf{F}$ at $\mathbf{q}$, i.e. $\mathbf{F}(\mathbf{q})+J\mathbf{F}(\mathbf{q})(\mathbf{x}-\mathbf{q})$, which is

$$\begin{pmatrix} 3 \\ 1 \\ 4 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 & 1 \\ 1 & 2 \\ -1 & 1 \\ 0 & 2 \end{pmatrix} \left(\begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right) = \begin{pmatrix} 4x+y-4 \\ x+2y+1 \\ -x+y+7 \\ 2y+3 \end{pmatrix},$$

for $(x,y)^T\in\mathbb{R}^2$. ■

In conclusion

> If the surface is given parametrically as the **image** of a map $\mathbf{u}\mapsto\mathbf{F}(\mathbf{u})$ then
>
> - the Tangent Space is the **image** of the linear map $d\mathbf{F}_{\mathbf{q}}:\mathbf{t}\mapsto J\mathbf{F}(\mathbf{q})\,\mathbf{t}$,
> - the Tangent Plane is the **image** of the affine map $\mathbf{t}\mapsto\mathbf{F}(\mathbf{q})+J\mathbf{F}(\mathbf{q})(\mathbf{t}-\mathbf{q})$, the best affine approximation to $\mathbf{F}$.

3.12 Conclusion

For a surface $S \subseteq \mathbb{R}^n$ given as a *graph* of the C^1-function $\boldsymbol{\phi} : U \subseteq \mathbb{R}^r \to \mathbb{R}^{n-r}$,

- the **Tangent Space** is given by the *graph* of the *linear map* $\mathbf{t} \mapsto J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t}$,
- the **Tangent Plane** is given by the *graph* of the *best affine approximation* to $\boldsymbol{\phi}$ at $\mathbf{q}$.

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

- the **Tangent Space** is the *image set* of the *linear map* $\mathbf{t} \mapsto J\mathbf{F}(\mathbf{q})\,\mathbf{t}$ and has a basis of the columns of $J\mathbf{F}(\mathbf{q})$, which are the directional derivatives $d_i\mathbf{F}(\mathbf{q})\,, 1 \le i \le n-m$,
- the **Tangent Plane** is the *image set* of the *best affine approximation* to $\mathbf{F}$ at $\mathbf{q}$.

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

- the **Tangent Space** is the **level set** of the linear function $\mathbf{t} \mapsto J\mathbf{f}(\mathbf{p})\,\mathbf{t}$,
- the **Normal Space** has a basis of the rows of $J\mathbf{f}(\mathbf{p})$, which are the gradient vectors of the component functions, i.e. $\nabla f^i\,(\mathbf{p})$, $1 \le i \le n-m$,

the **Tangent Plane** is the *level set* $J\mathbf{f}(\mathbf{p})\,(\mathbf{x} - \mathbf{p}) = \mathbf{0}$.

Appendix for Section 3 part 2

1. Definition of Tangent Plane

Recall

Definition 7 *The Tangent plane to S at $\mathbf{p}$ is the set of all tangent lines to S at $\mathbf{p}$.*

I claimed that this plane was, in fact, equal to

$$\mathbf{p}+T_{\mathbf{p}}(S)=\{\mathbf{p}+\mathbf{v}:\mathbf{v}\in T_{\mathbf{p}}(S)\}\,.$$

But this requires justification.

If $\mathbf{p}+\mathbf{v}\in\mathbf{p}+T_{\mathbf{p}}(S)$ then, by definition of $T_{\mathbf{p}}(S)$, there exists a curve $\boldsymbol{\alpha}$ in S with $\boldsymbol{\alpha}(0)=\mathbf{p}$ and $\boldsymbol{\alpha}'(0)=\mathbf{v}$. Then $\mathbf{p}+\mathbf{v}=\boldsymbol{\alpha}(0)+\boldsymbol{\alpha}'(0)$ which is a point on the tangent line $\boldsymbol{\alpha}(0)+\boldsymbol{\alpha}'(0)\,t, t\in\mathbb{R}$. Hence $\mathbf{p}+T_{\mathbf{p}}\,(S)$ is contained within the set of all tangent lines.

Conversely, given a point on a Tangent Line, i.e. $\boldsymbol{\alpha}(0)+\boldsymbol{\alpha}'(0)\,y$ for some curve α and $y\in\mathbb{R}$, define a new curve $\boldsymbol{\beta}(t)=\boldsymbol{\alpha}(yt)$. Then $\boldsymbol{\beta}(0)=\boldsymbol{\alpha}(0)=\mathbf{p}$ and $\boldsymbol{\beta}'(t)=y\boldsymbol{\alpha}'(yt)$ so $y\boldsymbol{\alpha}'(0)=\boldsymbol{\beta}'(0)\in T_{\mathbf{p}}(S)$, since the Tangent Space contains, by definition, all tangent vectors such as $\boldsymbol{\beta}'(0)$. Thus $\boldsymbol{\alpha}(0)+\boldsymbol{\alpha}'(0)\,y\in\mathbf{p}+T_{\mathbf{p}}(S)$. Hence the set of all tangent lines lies in $\mathbf{p}+T_{\mathbf{p}}(S)$.

2. Tangent Space for a Level Set

We now give the proof of Theorem 11. Recall from earlier appendices that, in a level set $\mathbf{f}^{-1}(\mathbf{0})$ with $\mathbf{f}:U\subseteq\mathbb{R}^n\to\mathbb{R}^m$ and $\mathbf{p}\in\mathbf{f}^{-1}(\mathbf{0})$ such that $J\mathbf{f}(\mathbf{p})$ is of full-rank, we can permute the coordinates of $\mathbb{R}^n$ and choose $\mathbf{f}$ so that

$$J\mathbf{f}(\mathbf{p})=(A\mid I_m)\,, \tag{14}$$

for some $m\times(n-m)$ matrix A.

Theorem 11 *Let*

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

for some $\mathcal{C}^1$-function $\mathbf{f}:U\subseteq\mathbb{R}^n\to\mathbb{R}^m$. Let $\mathbf{p}\in S$ and assume (14) *holds. Then*

$$T_{\mathbf{p}}\,(S)=\{\mathbf{x}\in\mathbb{R}^n:J\mathbf{f}(\mathbf{p})\,\mathbf{x}=\mathbf{0}\}\,. \tag{15}$$

Proof Because of (14) the **final** m columns of $J\mathbf{f}(\mathbf{p})$ are linearly independent and we can apply the Implicit Function Theorem. Hence there exists

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

such that

$$S \cap W = \left\{ \begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} : \mathbf{v} \in V \right\}. \tag{16}$$

Since $\mathbf{p} \in S \cap W$ we have

$$\mathbf{p} = \begin{pmatrix} \mathbf{q} \\ \boldsymbol{\phi}(\mathbf{q}) \end{pmatrix}$$

for some $\mathbf{q} \in V$. Then by Theorem 5

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} \mathbf{t} \\ J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t} \end{pmatrix} : \mathbf{t} \in \mathbb{R}^{n-m} \right\}. \tag{17}$$

We now wish to write this in the form (15) .

From (16) we see that for $\mathbf{v} \in V$ we have

$$\begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} \in S,$$

which, by the definition of S, holds if and only if

$$\mathbf{f}\begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} = 0.$$

The Chain Rule gives

$$J\mathbf{f}\begin{pmatrix} \mathbf{v} \\ \boldsymbol{\phi}(\mathbf{v}) \end{pmatrix}\begin{pmatrix} I_{n-m} \\ J\boldsymbol{\phi}(\mathbf{v}) \end{pmatrix} = 0.$$

Choose $\mathbf{v} = \mathbf{q}$ to get

$$J\mathbf{f}(\mathbf{p})\begin{pmatrix} I_{n-m} \\ J\boldsymbol{\phi}(\mathbf{q}) \end{pmatrix} = 0. \tag{18}$$

Yet, by (14),

$$J\mathbf{f}(\mathbf{p}) = (A \mid I_m)\,,$$

for some $m\times(n-m)$ matrix A. Then (18) gives

$$0 = (A \,|\, I_m)\begin{pmatrix} I_{n-m} \\ J\boldsymbol{\phi}(\mathbf{q}) \end{pmatrix}.$$

Multiplied out this gives $A + J\boldsymbol{\phi}(\mathbf{q}) = 0$ which, in (14) gives

$$J\mathbf{f}(\mathbf{p}) = \big(-J\boldsymbol{\phi}(\mathbf{q}) \,|\, I_m \big).$$

Write $\mathbf{x} \in \mathbb{R}^n$ as

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

with $\mathbf{t} \in \mathbb{R}^{n-m}$, $\mathbf{y} \in \mathbb{R}^m$. Then $J\mathbf{f}(\mathbf{p})\,\mathbf{x} = \mathbf{0}$ if, and only if,

$$\mathbf{0} = J\mathbf{f}(\mathbf{p})\,\mathbf{x} = \big(-J\boldsymbol{\phi}(\mathbf{q}) \,|\, I_m \big)\begin{pmatrix} \mathbf{t} \\ \mathbf{y} \end{pmatrix}.$$

That is $\mathbf{0} = -J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t} + I_m\mathbf{y}$, i.e. $\mathbf{y} = J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t}$, so

$$\mathbf{x} = \begin{pmatrix} \mathbf{t} \\ J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t} \end{pmatrix}.$$

Hence, returning to (17),

$$T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} \mathbf{t} \\ J\boldsymbol{\phi}(\mathbf{q})\,\mathbf{t} \end{pmatrix} : \mathbf{t} \in \mathbb{R}^{n-m} \right\} = \left\{ \mathbf{x} \in \mathbb{R}^n : J\mathbf{f}(\mathbf{p})\,\mathbf{x} = \mathbf{0} \right\}.$$

■

3 More Examples

An example of the use of the Implicit Function Theorem:

Example 18 *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 $\mathbf{f}(\mathbf{w}) = \mathbf{0}$ 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 $V : (-1, 2)^T \in V \subseteq \mathbb{R}^2$. ■

And another example:

Example 19 *Find the Tangent Space and Plane to the surface given by the level set*

$$\left\{\begin{array}{r} x^3 - xyu - uv^2 + v^3 = 2 \\ xu - yv = 3 \end{array}\right.$$

at the point $(x, y, u, v)^T = (1, -1, 2, 1)^T$. *Find also a basis for the Tangent Space.*

Solution Let $\mathbf{x} = (x, y, u, v)^T$ and $\mathbf{p} = (1, -1, 2, 1)^T$. Then

$$\begin{aligned} J\mathbf{f}(\mathbf{p}) &= \left.\begin{pmatrix} 3x^2 - yu & -xu & -xy - v^2 & -2uv + 3v^2 \\ u & -v & x & -y \end{pmatrix}\right|_{\mathbf{x}=\mathbf{p}} \\ &= \begin{pmatrix} 5 & -2 & 0 & -1 \\ 2 & -1 & 1 & 1 \end{pmatrix}. \end{aligned}$$

The rows are linearly independent so $J\mathbf{f}(\mathbf{p})$ is of full-rank. Then the Tangent Space is the level set of the linear function $\mathbf{x} \mapsto J\mathbf{f}(\mathbf{p})\,\mathbf{x}$;

$$T_{\mathbf{p}}(S) = \left\{\mathbf{x} \in \mathbb{R}^4 : J\mathbf{f}(\mathbf{p})\,\mathbf{x} = \mathbf{0}\right\} = \left\{(x, y, u, v)^T : \begin{array}{r} 5x - 2y - v = 0 \\ 2x - y + u + v = 0. \end{array}\right\}.$$

The Tangent plane is the level set of the best affine approximation to $\mathbf{f}$ at $\mathbf{p}$, which is $\mathbf{f}(\mathbf{p}) + J\mathbf{f}(\mathbf{p})\,(\mathbf{x} - \mathbf{p}) = J\mathbf{f}(\mathbf{p})\,(\mathbf{x} - \mathbf{p})$ since $\mathbf{f}(\mathbf{p}) = 0$, so

$$\left\{\mathbf{x} \in \mathbb{R}^4 : J\mathbf{f}(\mathbf{p})\,(\mathbf{x} - \mathbf{p}) = \mathbf{0}\right\} = \left\{(x, y, u, v)^T : \begin{array}{r} 5x - 2y - v = 6 \\ 2x - y + u + v = 6. \end{array}\right\}.$$

Look back in the Appendix for part 1 of Section 3 to see how to find the basis of a plane given as a level set. To start, write the Jacobian matrix $J\mathbf{f}(\mathbf{p})$ in the form $(G \,|\, I_2)$. The last two columns of $J\mathbf{f}(\mathbf{p})$ are linearly independent so we can multiply by the inverse of $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$. But this is equivalent to constructing the identity matrix within $J\mathbf{f}(\mathbf{p})$ by row operations:

$$\begin{pmatrix} 5 & -2 & 0 & -1 \\ 2 & -1 & 1 & 1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_2} \begin{pmatrix} 2 & -1 & 1 & 1 \\ 5 & -2 & 0 & -1 \end{pmatrix}$$

$$\xrightarrow{r_1 \to r_1 + r_2} \begin{pmatrix} 7 & -3 & 1 & 0 \\ 5 & -2 & 0 & -1 \end{pmatrix} \xrightarrow{r_2 \to -r_2} \begin{pmatrix} 7 & -3 & 1 & 0 \\ -5 & 2 & 0 & 1 \end{pmatrix}.$$

This is of the form $(G \,|\, I_2)$ in which case, by the theory in the previous appendix,

$$T_{\mathbf{p}}(S) = \left\{\begin{pmatrix} I_2 \\ -G \end{pmatrix} \mathbf{t} : \mathbf{t} \in \mathbb{R}^2\right\} = \left\{\begin{pmatrix} 2 \\ t \\ -7s + 3t \\ 5s - 2t \end{pmatrix} : \begin{pmatrix} s \\ t \end{pmatrix} \in \mathbb{R}^2\right\}.$$

so the basis for $T_{\mathbf{p}}(S)$ consists of the columns of

$$\begin{pmatrix} I_2 \\ -G \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -7 & 3 \\ 5 & -2 \end{pmatrix},$$

i.e. $\mathbf{v}_1 = (1, 0, -7, 5)^T$ and $\mathbf{v}_2 = (0, 1, 3, -2)^T$. (As a check confirm that $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal to $(5, -2, 0, -1)^T$ and $(2, -1, 1, 1)^T$, the rows of the Jacobian matrix $J\mathbf{f}(\mathbf{p})$ and thus a basis for $T_{\mathbf{p}}(S)^{\perp}$.) ■

Second to last observation on this example Go back to the system for the Tangent plane

$$\begin{aligned} 5x - 2y - v - 6 &= 0, \\ 2x - y + u + v - 6 &= 0, \end{aligned}$$

and solve for u and v (perhaps add the two equations to remove v from the second) so

$$\begin{aligned} v &= 5x - 2y - 6, \\ u &= -7x + 3y + 12. \end{aligned}$$

Then the plane can be written as a graph

$$\begin{pmatrix} x \\ y \\ -7x + 3y + 12 \\ 5x - 2y - 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 12 \\ 6 \end{pmatrix} + x \begin{pmatrix} 1 \\ 0 \\ -7 \\ 5 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 3 \\ -2 \end{pmatrix}.$$

We see the same basis vectors $\mathbf{v}_1 = (1, 0, -7, 5)^T$ and $\mathbf{v}_2 = (0, 1, 3, -2)^T$.

Last observation on this example *An ad hoc method to find a basis of $T_{\mathbf{p}}(S)$ given a basis of $T_{\mathbf{p}}(S)^{\perp}$:*

To find a basis for the tangent space we first note that because $J\mathbf{f}(\mathbf{p})$ is of full-rank, the Normal Space is of dimension 2 within $\mathbb{R}^4$, which has dimension 4. Hence the Tangent Space will be of dimension $4 - 2 = 2$. Thus we need to find two linearly independent vectors orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

Recall that if $\mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ then the vector product $\mathbf{v} \times \mathbf{w}$ is orthogonal to both $\mathbf{v}$ and $\mathbf{w}$. Given $\mathbf{a}, \mathbf{b} \in \mathbb{R}^4$ write $\mathbf{a} = (\mathbf{v}, c), c \in \mathbb{R}$ and $\mathbf{b} = (\mathbf{w}, d), d \in \mathbb{R}$. Then $(\mathbf{v} \times \mathbf{w}, 0)$ is orthogonal to both $\mathbf{a}$ and $\mathbf{b}$.

In our case $\mathbf{v}_1 = (5, -2, 0, 1)^T$ and $\mathbf{v}_2 = (2, -1, 1, 1)^T$ are a basis for $T_{\mathbf{p}}(S)^{\perp}$. Then $(5, -2, 0) \times (2, -1, 1) = (-2, -5, -1)$ and thus $(-2, -5, -1, 0)^T$ will be orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

Repeating this but taking, say, the 2$^{\text{nd}}$, 3$^{\text{rd}}$ and 4$^{\text{th}}$ coordinates. Then $(-2, 0, 1) \times (-1, 1, 1) = (1, 3, -2)$ and so $(0, 1, 3, -2)^T$ will be orthogonal to both $\mathbf{v}_1$ and $\mathbf{v}_2$.

Finally, $(2, 5, 1, 0)^T$ and $(0, 1, 3, -2)^T$ are linearly independent, thus they constitute a basis for Tangent Space at $\mathbf{p}$. $\blacksquare$

4. Dimension

I have defined surfaces parametrically as the image set of functions $\mathbf{F} : \mathbb{R}^r \to \mathbb{R}^n$. I have also defined surfaces as level sets, the kernel of functions $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. But I have not defined the dimension of a surface. Yet at every point of a surface we have a Tangent Space which does have a dimension. Further, because of the assumption in the definition of surface that the Jacobian matrix is of full-rank, the dimension of the Tangent Space is the same at all points on the surface; for parametric surfaces it is r and for level sets it is $n - m$. It is not unreasonable that the dimension of the Tangent planes is taken as the dimension of the Surface. This is why, if you have not already noticed, you should think of $n - m$ and r as equal and so $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^{n-r}$.