Calculus for Vector-Valued Functions

Def: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *vector-valued.* Such functions have the form

$$f(x_1, \ldots, x_n) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}$$

Def: Let $a = (a_1, \ldots, a_n)$ and let $b = (b_1, \ldots, b_n)$ be two points in $\mathbb{R}^n$. Then the *distance* between a and b, denoted by $|a - b|$, is

$$|a - b| = \sqrt{(a_1 - b_1)^2 + \cdots + (a_n - b_n)^2}$$

Def: The *length* of a is defined by

$$|a| = \sqrt{a_1^2 + \cdots + a_n^2}$$

Note: The length of a is the same as the distance between a and the origin.

Note: You may have noticed that there is a slight difference between vectors and points. A point is a location in space while a vector can be treated as a set of instructions to get from the origin to its associated point. More generally, remember that vectors in $\mathbb{R}^n$ are a particular case of the mathematical object called a vector. These vectors form a vector space over $\mathbb{R}$ while points do not. To illustrate this, consider the whether it makes any sense to add two points together or to scale a point. All these operations are meaningful in the context of vector spaces, but are meaningless when applied to locations.

Def: The function $f : \mathbb{R}^n \to \mathbb{R}^m$ has the limit

$$L = (L_1, \ldots, L_m) \in \mathbb{R}^m$$

at the point $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ if given any $\epsilon > 0$, there is some $\delta > 0$ such that for all $x \in \mathbb{R}^n$, if

$$0 < |x - a| < \delta$$

we have

$$|f(x) - L| < \epsilon$$

We denote this limit by

$$\lim_{x \to a} f(x) = L$$

or by $f(x) \to L$ as $x \to a$.

Def: The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *continuous* at a point $a \in \mathbb{R}^n$ if

$$\lim_{x \to a} f(x) = f(a)$$

.

Def: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *differentiable* at $a \in \mathbb{R}^n$ if there is an $m \times n$ matrix $A : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{x \to a} \frac{|f(x) - f(a) - A \cdot (x - a)|}{|x - a|} = 0$$

This matrix is called the *Jacobian* and is denoted by $Df(a)$.

Thm. If the Jacobian of a function exists, then it is unique up to change in basis.

Note: This doesn't look like a pure extension of the single variable case, but that's because it extends a different form of the single-variable definition, namely

$$\lim_{x \to a} \frac{|f(x) - f(a) - f'(a)(x - a)|}{|x - a|} = 0$$

This form makes it clear that the Jacobian matrix is really equivalent to the derivative of the multivariate function evaluated at the point a.

Thm. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be given by m differentiable functions such that

$$f(x_1, \ldots, x_n) = \begin{pmatrix} f_1(x_1, \ldots, x_n) \\ \vdots \\ \vdots \\ f_m(x_1, \ldots, x_n) \end{pmatrix}$$

Then f is differentiable and the Jacobian is

$$Df(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Chain Rule: Let $f : \mathbb{R}^n \to \mathbb{R}^m$ and $g : \mathbb{R}^m \to \mathbb{R}^l$ be differentiable functions. Then the composition function

$$g \circ f : \mathbb{R}^n \to \mathbb{R}^l$$

is also differentiable with derivative given by: if $f(a) = b$, then

$$D(g \circ f)(a) = D(g)(b) \cdot D(f)(a)$$

Linear Approximation: The vector $y = f(x)$ can be approximated by

$$y \approx f(a) + Df(a) \cdot (x - a)$$

This is equivalent to finding the tangent line of a curve in the single-variable case.

Inverse Function Theorem: For a vector valued, continuously differentiable function $f : \mathbb{R}^n \to \mathbb{R}^m$, assume that $detDf(a) \neq 0$, at some point $a \in \mathbb{R}^n$. Then there is an open neighborhood U of a in $\mathbb{R}^n$ and an open neighborhood of V of $f(a)$ in $\mathbb{R}^m$ such that $f : U \to V$ is one to one, onto and has a differentiable inverse $g : V \to U$.

Note: In terms of linear algebra, the Jacobian is just a matrix. This means that the derivative is a linear transformation and all the tools from linear algebra can be applied. The Inverse Function Theorem simply notes that when the Jacobian matrix is invertible, the matrix will have an inverse and therefore the function it represents will also have an inverse.

Implicit Function Theorem: Let $f_1(x,y), \ldots, f_k(x,y)$ be k continuously differentiable functions on $\mathbb{R}^{n+k}$ and suppose that $p = (a,b) \in \mathbb{R}^{n+k}$ is a point for which

$$f_a(a,b) = 0, \ldots, f_k(a,b) = 0$$

Suppose that at the point p the $k \times k$ matrix

$$M = \begin{pmatrix} \frac{\partial f_1}{\partial y_1(p)} & \cdots & \frac{\partial f_1}{\partial y_k(p)} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial y_1(p)} & \cdots & \frac{\partial f_k}{\partial y_k(p)} \end{pmatrix}$$

is invertible. Then in a neighborhood of a in $\mathbb{R}^n$ there are k unique, differentiable functions

$$p_1(x), \ldots, p_k(x)$$

such that

$$f_1(x, p_1(x)) = 0, \ldots, f_k(x, p_k(x)) = 0$$

Exercises:

1a. Compute the Jacobian for the change to polar coordinates.

Answer:
We first write the functions for the coordinate change as a multivariate function.

$$h(x,y) = h(f(r,\theta), g(r,\theta)) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}$$

Now we can take the matrix of partials

$$Dh(x,y) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

1b. At what points is the change of coordinates not well-defined?

Answer:
We can determine this by setting the determinate of the Jacobian to 0

$$(\cos\theta)(r\cos\theta) + (\sin\theta)(r\sin\theta) = 0$$

The change of coordinates function will not be defined at $(0,0)$, or you could also say it is not defined whenever $r = 0$.

1c.   Give a geometric justification for your answer in part b.

Answer:
When $r = 0$, there is a problem going back to Cartesian coordinates. The problem is that when the radius is zero, it gets weird because all the θ values will lead to the same point. (Describing this geometrically is weird when its trivial to show that the inverse of the Jacobian will involve division by zero at $r = 0$)

2.   The roots and the coefficients of degree 2 monic polynomials are related by

$$(x - r_1)(x - r_2) = x^2 + ax + b$$

2a.   Write down the functions giving the change of coordinates from the root space to the coefficient space.

Answer:
Multiply out the LHS

$$(x - r_1)(x - r_2) = x^2 - (r_1 + r_2)x + r_1 r_2 = x^2 + ax + b$$

This gives us the needed equations for a and b

$$a = -(r_1 + r_2), b = r_1 r_2$$

We can therefore write our function as

$$f(r_1, r_2) = \begin{pmatrix} -r_1 - r_2 \\ r_1 r_2 \end{pmatrix}$$

2b.   Compute the Jacobian of the coordinate change.

Answer:
The Jacobian is

$$Df = \begin{pmatrix} -1 & -1 \\ r_2 & r_1 \end{pmatrix}$$

2c.   Find where this coordinate change is not invertible.

Answer:

$$det(Df) = -r_1 + r_2 = 0 \implies r_1 = r_2$$

2d.   Give a geometric interpretation to your answer in part c.

Answer:
This situation is the result of the polynomial having a repeated root. Therefore, the polynomial will only have one root, i.e. one point where the polynomial touches the x-axis. This will be either a max or a min and is therefore the only point when the slope is zero.

3a.   Via the quadratic equation, write down the functions giving the change of coordinates from the coordinate space to the root space.

Answer:
Using the quadratic formula, we get the equations

$$r_1 = \frac{-a + \sqrt{a^2 - 4b}}{2}, r_2 = \frac{-a - \sqrt{a^2 - 4b}}{2}$$

We can write our function as

$$f(a, b) = \begin{pmatrix} \frac{-a+\sqrt{a^2-4b}}{2} \\ \frac{-a-\sqrt{a^2-4b}}{2} \end{pmatrix}$$

3b.   Compute the Jacobian.

Answer:

$$Df(a, b) = \begin{pmatrix} -\frac{1}{2} + \frac{1}{4}(a^2 - 4b)^{-1/2}(2a) & -(a^2 - 4b)^{-1/2} \\ -\frac{1}{2} - \frac{1}{4}(a^2 - 4b)^{-1/2}(2a) & (a^2 - 4b)^{-1/2} \end{pmatrix}$$

3c.   Find where this coordinate change is not invertible.

Answer:

$$det(Df) = \left[-\frac{1}{2} + \frac{1}{4}(a^2 - 4b)^{-1/2}(2a) \right] \left[(a^2 - 4b)^{-1/2} \right] - \left[-\frac{1}{2} - \frac{1}{4}(a^2 - 4b)^{-1/2}(2a) \right] \left[-(a^2 - 4b)^{-1/2} \right]$$

$$= \left[(a^2 - 4b)^{-1/2} \right] [-1] = -\frac{1}{\sqrt{a^2 - 4b}}$$

Therefore, the coordinate change will singular when $a^2 = 4b$.