Note 3.3 - Chain Rules and Implicit Differentiations

1 Introduction

We have learned the complete relations between differentiation and elementary algebraic operations. Next we develop the formula for differentiation of composite functions.

2 Linear Approximation

The tangent line to the graph $y = f(x)$ at x_0 has equation

Certainly $p_1(x_0) - f(x_0) = 0$. By continuity, p_1 is quite close to f near x_0. But note that the constant function $p_0(x) = f(x_0)$ is always "quite close" to f near x_0:

But our intuition tells us that p_1 is "closer" to f than p_0. This is indeed true, in the sense that

or equivalently

3 The Chain Rule

Let $f(x)$ and $g(x)$ be differentiable functions and

$$h(x) = g \circ f(x) := g(f(x)).$$

We want to compute $h'(x)$ in terms of f, g, and their derivatives. Let's consider the easy (but important) case, where

$$f(x) = ax + b, \quad g(x) = cx + d.$$

The general case follows from the principle of linear approximation above:

We have arrived at the conclusion.

Theorem 3.1. *Given differentiable functions* $f(x) : A \to B$ *and* $g(u) : B \to C$*, let* $h(x) = g(f(x))$*, we have*

$$h'(x) = g'(f(x))f'(x).$$

In Leibniz notations,

$$\frac{dh}{dx} = \left.\frac{dg}{du}\right|_{u=f(x)} \frac{df}{dx}.$$

Let's discuss some examples.

For the composition of multiple functions

$$h(x) = f_n \circ f_{n-1} \circ \cdots \circ f_1(x),$$

we apply the formula repeatedly to obtain

Basically, we differentiate the function one layer at the time, starting from the outermost. Each derivative is evaluated at the composition of functions from f_1 to the function before that layer.

Here are some examples.

4 The Implicit Differentiations

There are some curves in $\mathbb{R}^2$ defined by an equation of x and y (called the *level curve*), but is not a graph of any function. The typical example is the unit circle $F(x, y) = x^2 + y^2 - 1 = 0$:

However, these curves ar nice and smooth with very reasonable tangent line at every point. This is because derivatives are *local* quantities. Recall, on a graph of a function $y = f(x)$, we only need to know the curve *near* $P = (x_0, f(x_0))$ to compute $\frac{dy}{dx}|_P$:

The same quantity makes perfect sense a point P on any curve if the curve *near* P is a graph of some function.

We use the same derivative notation $\frac{dy}{dx}|_P$ (or $\frac{dx}{dy}|_P$) with the understanding that some part of the curve around P is given by the graph $y = f(x)$ (or $x = g(y)$).

Note that points P on the level curve $F(x, y) = 0$ can have neighborhood on which the curve is the graph of $y = f(x)$ or $x = g(y)$ (or both):

There are also points P near which the level curve is not graph of any function:

Computations of $\frac{dy}{dx}$ or $\frac{dx}{dy}$, however, generally do not requiring solving one variable in terms of the others:

5 Derivatives of Inverse Functions

For a differentiable function $f : A \to B$ with inverse, implicit differentiation can be used to compute the derivative of f^{-1}:

Lets derive the derivatives of *inverse trigonometric functions*. But let's first define them appropriately:

Then, we have