Part of the difficulty in extending the idea of derivative to more complex spaces is to give an adequate notion of linear approximation. The key idea is to use linear maps. This could be carried out in terms of matrices but it turns out that this neither shortens nor simplifies proofs. In fact, this is often the opposite.

We admit that the more intrinsic definition of the notion of derivative f'_a at a point a of a function $f\colon E \to F$ between two normed (affine) spaces E and F as a linear map requires a greater effort to be grasped, but we feel that the advantages of this definition outweight its degree of abstraction. In particular, it yields a clear notion of the derivative of a function $f\colon \mathrm{M}_m(\mathbb{R}) \to \mathrm{M}_n(\mathbb{R})$ defined from $m \times m$ matrices to $n \times n$ matrices (many definitions make use of partial derivatives with respect to matrices that do make any sense). But more importantly, the definition of the derivative as a linear map makes it clear that whether the space E or the space F is infinite dimensional does not matter. This is important in optimization theory where the natural space of solutions of the problem is often an infinite dimensional function space. Of course, to carry out computations one need to pick finite bases and to use Jacobian matrices, but this is a different matter.

Let us now review the formal definition of the derivative of a real-valued function.

Definition 39.1. Let A be any nonempty open subset of $\mathbb{R}$, and let $a \in A$. For any function $f\colon A \to \mathbb{R}$, the *derivative of f at $a \in A$* is the limit (if it exists)

$$\lim_{h\to 0,\, h\in U} \frac{f(a+h) - f(a)}{h},$$

where $U = \{h \in \mathbb{R} \mid a + h \in A,\ h \neq 0\}$. This limit is denoted by $f'(a)$, or $\mathrm{D}f(a)$, or $\frac{df}{dx}(a)$. If $f'(a)$ exists for every $a \in A$, we say that *f is differentiable on A*. In this case, the map $a \mapsto f'(a)$ is denoted by f', or $\mathrm{D}f$, or $\frac{df}{dx}$.

Note that since A is assumed to be open, $A - \{a\}$ is also open, and since the function $h \mapsto a + h$ is continuous and U is the inverse image of $A - \{a\}$ under this function, U is indeed open and the definition makes sense.

We can also define $f'(a)$ as follows: there is some function ϵ, such that,

$$f(a + h) = f(a) + f'(a) \cdot h + \epsilon(h)h,$$

whenever $a + h \in A$, where $\epsilon(h)$ is defined for all h such that $a + h \in A$, and

$$\lim_{h\to 0,\, h\in U} \epsilon(h) = 0.$$

Remark: We can also define the notion of *derivative of f at a on the left*, and *derivative of f at a on the right*. For example, we say that the *derivative of f at a on the left* is the limit $f'(a_-)$ (if it exists)

$$\lim_{h\to 0,\, h\in U} \frac{f(a+h) - f(a)}{h},$$