DIFFERENTIAL CALCULUS ON THE SIMPLEX

C. Barceló-Vidal, J. A. Martín-Fernández
Depart. d’Informàtica i Matemàtica Aplicada
Universitat de Girona - e-mail: carles.barcelo@udg.es

Summary. The purpose of this paper is to introduce the concept of differentiability of a vector-valued function on the simplex. In particular, the concepts of compositional gradient and compositional directional derivative of a real-valued function on the simplex are exposed and discussed.

1. Introduction

In some cases a response variable $y$ is assumed to depend only on the proportions $x_1, \ldots, x_D$ of $D$ ingredients or parts present in a specific mixture and not on the amount of the mixture. These proportions are often expressed by volume, by weight, by mole fraction, etc. In mathematical terms the response variable $y$ can be interpreted as a real or vector-valued function $\varphi$ whose domain is a subset of the simplex space.

In many practical situations the expression $y = \varphi(x_1, \ldots, x_D)$ is unknown and the emphasis is on fitting the simplest model to sample or experimental data. However, in some cases, the function $\varphi$ can be deduced from physical laws. One example is the rubidium-strontium method of dating $Rb$-minerals —based on the law of radioactivity— which uses the following function to determine the “age” $t$ of a mineral:

$$t = \frac{1}{\lambda} \log \left[ \frac{\text{Sr}^{\text{Rb}}}{\text{Sr}^{\text{Sr}}} - \xi_0 \right] + 1,$$

where $\text{Sr}^{\text{Rb}}$, $\text{Sr}^{\text{Sr}}$ and $\text{Rb}$ are, respectively, the total number of atoms of these isotopes in a unit weight of the mineral at the present time; $\xi_0$ is the $\text{Sr}^{\text{Sr}}/\text{Sr}^{\text{Sr}}$ ratio of strontium that was incorporated into the mineral at the time of its formation; $\lambda$ is the decay constant of $\text{Rb}$ in units of reciprocal years ($\lambda = 1.42 \times 10^{-11}$ yr$^{-1}$); and $t$ is the time elapsed in years since the time of formation of mineral, that is, the “age” of the mineral. The number of $\text{Sr}^{\text{Sr}}$ atoms is constant because this isotope is stable and not produced by decay of a naturally occurring isotope of another element. The function (1) is the basis for age determination by the $Rb$-$Sr$ method when the mineral has remained a “closed system” with respect to rubidium and strontium, and when the assumed value $\xi_0$ of the initial $\text{Sr}^{\text{Sr}}/\text{Sr}^{\text{Sr}}$ ratio is appropriate. From a mathematical point of view the expression (1) is a real-valued function defined on a subset in the simplex $S^3$ of 3-parts $\text{Sr}^{\text{Sr}}, \text{Sr}^{\text{Sr}}, \text{Rb}$.

In the next sections we introduce the main topics on differential calculus of the vector-valued functions on the compositional space. We suppose known the terminology associated to the metric vector space structure of the compositional space $C^d$ introduced in Barceló-Vidal et al. (2001) and the notation introduced in Aitchison et al. (2002).
2. Vector-valued functions on $C^d$

2.1 Scale-invariant vector-valued functions on $\mathbb{R}^D_+$

Let $f$ be a vector-valued function whose domain is a subset $A$ in $\mathbb{R}^D_+$ and with range contained in $\mathbb{R}^m$. Thus $f$ assigns each $w$ in $A$ a value $f(w) = (f_1(w), \ldots, f_m(w))'$, an $m$-tuple in $\mathbb{R}^m$.

We say $f$ is scale-invariant —or homogeneous of degree 0— if $f(kw) = f(w)$ for every positive real $k$, and for every $w$ in $A$ for which $kw$ in $A$. Obviously, if a vector-valued function $f(w) = (f_1(w), \ldots, f_m(w))'$, with domain $A \subset \mathbb{R}^D_+$, is scale-invariant, its components $f_1, \ldots, f_m$ are also real-valued scale-invariant functions on $A$. And inversely, if all the components $f_1, \ldots, f_m$ are real-valued scale-invariant functions, the vector-valued function $f(w) = (f_1(w), \ldots, f_m(w))'$ is scale-invariant.

Let $A$ be a subset of $\mathbb{R}^D_+$. We define the scale closure of $A$, denoted $A^*$, as the set

$$A^* = \{kw : k > 0, w \in A\}.$$

The set $A$ is said to be scale-closed, if $A^* = A$.

It is clear that any scale-invariant vector-valued function $f$ with domain $A \subset \mathbb{R}^D_+$ can be extended to the scale closure $A^*$ of $A$. It suffices to define $f(kw) = f(w) \quad (k > 0) \ (w \in A)$. Obviously, this extended function with domain $A^*$ is also scale-invariant. Therefore, throughout this paper, we shall implicitly suppose that the domain of any scale-invariant function is always an scale-closed subset.

Let $f : A \rightarrow \mathbb{R}^m$ be an scale-invariant vector-valued function defined on a subset $A$ in $\mathbb{R}^D_+$ with values in $\mathbb{R}^m$. It induces a vector-valued function $\widehat{f} : \mathcal{A} = \text{ccl}A \subset C^d$. Certainly, if $c$ is a composition which belongs to $\mathcal{A}$, there exists at least a vector $w$ in $A$ such that $c\cdot w = c$. Then, we define $f(c)$ as equal to $f(w)$. As $f$ is scale-invariant, it is clear that $f(c)$ is univocally defined. We will denote by $\widehat{f}$ the vector-valued function defined on $\widehat{A} = \text{ccl}A$ in $C^d$ associated to the scalar invariant vector-function $f$ defined on the scalar-closed subset $A$ in $\mathbb{R}^D_+$. Inversely, if $\varphi : A \rightarrow \mathbb{R}^m$ is a vector-valued function defined on a set $\mathcal{A}$ in $C^d$ with values in $\mathbb{R}^m$, it can be proved that there exist an unique scale-invariant function $\varphi$ defined on an scale-closed set $A$ in $\mathbb{R}^D_+$, such that $\varphi = \varphi$. Therefore we can suppose that any vector-valued function defined on a set in $C^d$ is derived from an scalar-invariant vector-valued function defined on an scale-closed set in $\mathbb{R}^D_+$.

2.2 $C$-linear functions on $C^d$

According to the vector space structure of $(C^d, \oplus, \otimes)$, a vector-valued function $\varphi : C^d \rightarrow \mathbb{R}^m$ is $C$-linear if

$$\varphi(w \oplus w^*) = \varphi(w) + \varphi(w^*) \quad (w, w^* \in C^d),$$

and

$$\varphi(\lambda \otimes w) = \lambda \varphi(w) \quad (w \in C^d) \ (\lambda \in \mathbb{R}).$$

It can be proved that, $\varphi$ is $C$-linear if and only if there exists a constant $m \times D$ matrix $A$, constrained by the equality $A1_D = 0_m$, such that $f = \varphi$, where $f : \mathbb{R}^D_+ \rightarrow \mathbb{R}^m$ is
the real-valued function defined by the expression

\[ f(w) = A \log w \quad (w \in \mathbb{R}^D_+). \]

3. Derivatives on \( C^d \)

3.1 Derivative of scale-invariant vector-valued functions on \( \mathbb{R}^D_+ \)

Let \( f = (f_1, \ldots, f_m) : A \rightarrow \mathbb{R}^m \) be a vector-valued function whose domain is an open subset \( A \) in \( \mathbb{R}^D_+ \). If \( f \) is differentiable at \( w^* \in A \), we will denote by

\[
Df(w^*) = \begin{bmatrix}
\frac{\partial f_1}{\partial w_1}(w^*) & \cdots & \frac{\partial f_1}{\partial w_D}(w^*) \\
\vdots & \ddots & \vdots \\
\frac{\partial f_m}{\partial w_1}(w^*) & \cdots & \frac{\partial f_m}{\partial w_D}(w^*)
\end{bmatrix}
\]

the derivative of \( f \) at \( w^* \). Therefore, this derivative is a \( m \times D \) matrix referred to as the matrix of partial derivatives of \( f \) at \( w^* \). In particular, if \( f \) is a real-valued function, its derivative at \( w^* \)

\[
Df(w^*) = \begin{bmatrix}
\frac{\partial f}{\partial w_1}(w^*) & \cdots & \frac{\partial f}{\partial w_D}(w^*)
\end{bmatrix}
\]

is called the gradient of \( f \) at \( w^* \) and denoted by \( \nabla f(w^*) \).

According to the well-known Euler’s theorem for homogeneous functions, if \( f \) is scale-invariant and differentiable at \( w^* \in A \), then

\[
\sum_{j=1}^{D} w_j^* \frac{\partial f_i}{\partial w_j}(w^*) = 0
\]

for each \( i = 1, \ldots, m \). Furthermore, \( f \) is also differentiable at \( kw^* \), for each \( k > 0 \), and

\[
\frac{\partial f_i}{\partial w_j}(kw^*) = \frac{1}{k} \frac{\partial f_i}{\partial w_j}(w^*)
\]

for each \( i = 1, \ldots, m \), and \( j = 1, \ldots, D \).

3.2 Derivative of vector-valued functions on \( C^d \)

Let \( f = (f_1, \ldots, f_m) : A \rightarrow \mathbb{R}^m \) be a scale-invariant vector-valued function whose domain is an open and scale-closed subset \( A \) in \( \mathbb{R}^D_+ \). Let \( f = (f_1, \ldots, f_m) \) be the vector-valued function defined on \( A = \text{ccl} A \) in \( C^d \) associated to function \( f \).

We say that \( f \) is \( C \)-differentiable at \( w^* \) in \( A \) if it exists a \( m \times D \) matrix \( A \), constrained by the equality \( A 1_D = 0_m \), such that

\[
\lim_{u \leq 1_D} \left\| \frac{f(w^* + u) - f(w^*) - A \log u}{\|u\|_c} \right\| = 0,
\]

where \( 1_D = \text{ccl}(1, \ldots, 1)' \) is the neutral element of \( (C^d, \oplus) \).
The element \((i, j)\)-th of matrix \(A\) will be referred to as the \(C\)-partial derivative of \(f_i\) with respect to the \(j\)-th coordinate at \(w^*\), and is denoted by \(\frac{\partial c f_i}{\partial w_j}(w^*)\). It is easy to prove that

\[
\frac{\partial c f_i}{\partial w_j}(w^*) = \lim_{t \to 0} \frac{f_i(w_i^*, \ldots, w_j^* + t, w_{j+1}^*, \ldots, w_D^*) - f_i(w^*)}{t}
\]

From this interpretation of the \(C\)-partial derivatives, it is not difficult to prove that \(f\) is \(C\)-differentiable at \(w^*\) in \(A\) if and only if \(f\) is differentiable at \(w^*\) in \(A\). Furthermore, the \(C\)-partial derivatives of \(f\) at \(w^*\) and the partial derivatives of \(f\) at \(w^*\) are related by the equality

\[
\frac{\partial c f_i}{\partial w_j}(w^*) = w_j^* \frac{\partial f_i}{\partial w_j}(w^*),
\]

for each \(i = 1, \ldots, m\), and \(j = 1, \ldots, D\).

The matrix

\[
A = \begin{bmatrix}
\frac{\partial c f_1}{\partial w_1}(w^*) & \cdots & \frac{\partial c f_1}{\partial w_D}(w^*) \\
\cdots & \cdots & \cdots \\
\frac{\partial c f_m}{\partial w_1}(w^*) & \cdots & \frac{\partial c f_m}{\partial w_D}(w^*)
\end{bmatrix}
\]

will be referred to as the matrix of \(C\)-partial derivatives of \(f\) at \(w^*\) and will be denoted by \(D_c f(w^*)\).

In particular, if \(f\) is a real-valued function, the matrix of \(C\) – partial derivatives of \(f\) at \(w^*\)

\[
D_c f(w^*) = \begin{bmatrix}
\frac{\partial f}{\partial w_1}(w^*), & \cdots, & \frac{\partial f}{\partial w_D}(w^*)
\end{bmatrix}
\]

is called the \(C\)-gradient of \(f\) at \(w^*\) and denoted by \(\nabla_c f(w^*)\). The first-order Taylor approximation of \(f\) in a neighbourhood of \(w^*\) is given by

\[
f(w^*; u) = f(w^*) + \sum_{j=1}^{D} \left( \frac{\partial f}{\partial w_j}(w^*) \right) \log u_j + R_1(w^*, u),
\]

where \(u = c(u_1, \ldots, u_D)'\), and \(R_1(w^*, u)/\|u\| \to 0\) in \(\mathbb{R}\) as \(u \to 1_D\) in \(C^d\).

4. Directional derivatives on \(C^d\)

Let \(f : \mathbb{R}_+^D \to \mathbb{R}\) be a scale-invariant real-valued function defined on \(\mathbb{R}_+^D\) and \(f\) be the real-valued function defined on \(C^d\) associated to function \(f\). Let \(w^*, u \in C^d\).

Then, the \(C\)-directional derivative of \(f\) at \(w^*\) along the direction \(u\) is given by

\[
\left( \frac{df(w^* + (t \otimes u))}{dt} \right)_{t=0}
\]

if this exists. From this definition, it is no difficult to prove that the \(C\)-directional derivative can also be defined by the formula

\[
\lim_{t \to 0} \frac{f(w^1 u_1^t, \ldots, w^D u_D^t) - f(w^1, \ldots, w^D)}{t}
\]
If $f$ is differentiable, then all $\mathcal{C}$-directional derivatives of $f$ exist. Furthermore, the $\mathcal{C}$-directional derivative at $\mathbf{w}^* \in \mathcal{C}^d$ along the direction $\mathbf{u} \in \mathcal{C}^d$ is given by

$$
\sum_{j=1}^D w_j^* \frac{\partial f}{\partial w_j}(\mathbf{w}^*) \log u_j = \nabla_{\mathcal{C}} f(\mathbf{w}^*) \log \mathbf{u}.
$$

Therefore, if $\nabla_{\mathcal{C}} f(\mathbf{w}^*) \neq 0$, the composition $\text{ccl} \left\{ \exp \left( \nabla_{\mathcal{C}} f(\mathbf{w}^*) \right) \right\}$ points in the direction along which $f$ is increasing the fastest in a $\mathcal{C}$-neighbourhood of $\mathbf{w}^*$.

We can deduce from (2) that the $\mathcal{C}$-directional derivative of $f$ at $\mathbf{w}^* \in \mathcal{C}^d$ along the direction

$$
\mathbf{u}_j = \text{ccl}(1, \ldots, 1, e, 1, \ldots, 1)^T (j = 1, \ldots, D)
$$

is equal to

$$
\frac{\partial f}{\partial w_j}(\mathbf{w}^*) = w_j^* \frac{\partial f}{\partial w_j}(\mathbf{w}^*). \tag{3}
$$

If the $\mathcal{C}$-directional derivative of $f$ along the direction $\mathbf{u}_j$ is equal to 0 on $\mathcal{C}^d$, we say that $f$ is $\mathcal{C}$-independent of the $j$-th part. It is clear from (3) that $f$ is $\mathcal{C}$-independent of the $j$-th part if and only if $f$ does not depend of the $j$-th coordinate $w_j$ on $\mathbb{R}_+^B$.

5. Example

The function (1) giving the “age” $t$ of a mineral from the $^{87}\text{Sr}/^{86}\text{Sr}$ and $^{87}\text{Rb}/^{86}\text{Sr}$ ratios can be interpreted as a scale-invariant real-valued function $f$

$$
\left( ^{86}\text{Sr}, ^{87}\text{Sr}, ^{87}\text{Rb} \right) \rightarrow \frac{1}{\lambda} \log \left[ \frac{^{87}\text{Sr}}{^{86}\text{Sr}} - \xi_0 \right] + 1
$$

defined on

$$
A = \left\{ \left( ^{86}\text{Sr}, ^{87}\text{Sr}, ^{87}\text{Rb} \right) \in \mathbb{R}_+^3 : \frac{^{87}\text{Sr}}{^{86}\text{Sr}} - \xi_0 \geq 0 \right\}.
$$

Therefore, the function $f$ induces a real-valued function $\bar{f}$ defined on $A = \text{ccl} A \subset \mathcal{C}^2$. This function $\bar{f}$ can also be interpreted as a function on the simplex $S^3$. Figure 1 shows some level contours of this function for $\lambda = 1.42 \times 10^{-11}$ and $\xi_0 = 0.7071$. The projections of these level contours on the simplex are straight segments from a common point on the $^{86}\text{Sr} - ^{87}\text{Sr}$ edge.

The $\mathcal{C}$-gradient of $\bar{f}$ at $\mathbf{w} = \text{ccl} \left( ^{86}\text{Sr}, ^{87}\text{Sr}, ^{87}\text{Rb} \right)^T$ is equal to

$$
\nabla_{\mathcal{C}} f(\mathbf{w}) = \frac{1}{\lambda} \left[ \frac{^{87}\text{Sr}}{^{86}\text{Sr}} - \xi_0 \right] - \xi_0 \left[ \frac{^{87}\text{Sr}}{^{86}\text{Sr}} - \xi_0 \right] - \xi_0 \left[ \frac{^{87}\text{Sr}}{^{86}\text{Sr}} - \xi_0 \right] \tag{4}
$$

From (4) we deduce that $\frac{\partial f}{\partial ^{87}\text{Sr}}(\mathbf{w})$ is always positive in each $\mathbf{w} \in A$. It means that the “age” of a mineral increases when $^{87}\text{Sr}$ increases and the $^{87}\text{Rb}/^{86}\text{Sr}$ ratio remains constant. Similarly, the $\mathcal{C}$-partial derivative $\frac{\partial f}{\partial ^{87}\text{Rb}}(\mathbf{w})$ is negative in each $\mathbf{w}$ in the $\mathcal{C}$-interior of $A$. Therefore, the “age” of a mineral decreases when $^{87}\text{Rb}$ increases and the $^{87}\text{Sr}/^{86}\text{Sr}$ ratio remains constant.
Any of the three dotted paths in Figure (1) represents in each point the direction in which the values of function $f$ change most rapidly, showing the evolution of the ($^{86}\text{Sr}$-$^{87}\text{Sr}$-$^{87}\text{Rb}$)-composition of a mineral from the time of its formation.

Figure 1: Level contours ($t = 0; 10 \times 10^6; 50 \times 10^6; 100 \times 10^6; 150 \times 10^6; 200 \times 10^6$ years) of the function $t = \frac{1}{1.42 \times 10^{-6}} \log \left[ \frac{87\text{Rb}}{87\text{Sr}} - 0.7071 \right] + 1$ on the simplex of 3-parts $^{86}\text{Sr}$-$^{87}\text{Sr}$-$^{87}\text{Rb}$. The dotted lines show the evolution of the ($^{86}\text{Sr}$-$^{87}\text{Sr}$-$^{87}\text{Rb}$)-composition of three minerals from the time of their formation.

References
