How to differentiate the general solution

Differential equations for math students, Rafael Ortega

Given an open and connected subset D of $\mathbb{R} \times \mathbb{R}^d$, we consider a continuous vector field

$$X : D \to \mathbb{R}^d, \quad (t, x) \mapsto X(t, x)$$

such that the partial derivative $\frac{\partial X}{\partial x}(t, x)$ exists at each point of D and the matrix-valued function

$$(t, x) \in D \mapsto \frac{\partial X}{\partial x}(t, x) \in \mathbb{R}^{d \times d}$$

is continuous.

Given $(t_0, x_0) \in D$, $x(t; t_0, x_0)$ denotes the unique maximal solution of the initial value problem

$$\dot{x} = X(t, x), \quad x(t_0) = x_0,$$

defined on the maximal interval $]\alpha(t_0, x_0), \omega(t_0, x_0)[$ with

$$-\infty \leq \alpha(t_0, x_0) < t_0 < \omega(t_0, x_0) \leq +\infty.$$

It is known that the set

$$\mathcal{D} = \{(t; t_0, x_0) \in \mathbb{R} \times \mathbb{R}^d : \ (t_0, x_0) \in D, \ \alpha(t_0, x_0) < t < \omega(t_0, x_0)\}$$

is open in $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$. Moreover, the map

$$\mathcal{D} \to \mathbb{R}^d, \quad (t; t_0, x_0) \mapsto x(t; t_0, x_0)$$

is continuous. This map is sometimes called the *general solution*. We want to prove that $x \in C^1(\mathcal{D}, \mathbb{R}^d)$ and the partial derivatives are given by

$$\frac{\partial x}{\partial t}(t; t_0, x_0) = X(t, x(t; t_0, x_0))$$

$$\frac{\partial x}{\partial x_0}(t; t_0, x_0) = Y(t; t_0, x_0)$$

$$\frac{\partial x}{\partial t_0}(t; t_0, x_0) = -Y(t; t_0, x_0)X(t_0, x_0)$$

where $Y(t; t_0, x_0)$ is the matrix solution of

$$\dot{Y} = \frac{\partial X}{\partial x}(t, x(t; t_0, x_0))Y, \quad Y(t_0) = I_{d\times d}.$$

Here $I_{d\times d}$ denotes the identity matrix. The linear equation in Y is sometimes called the *variational equation* or the *linearized equation* along the solution $x(t; t_0, x_0)$.

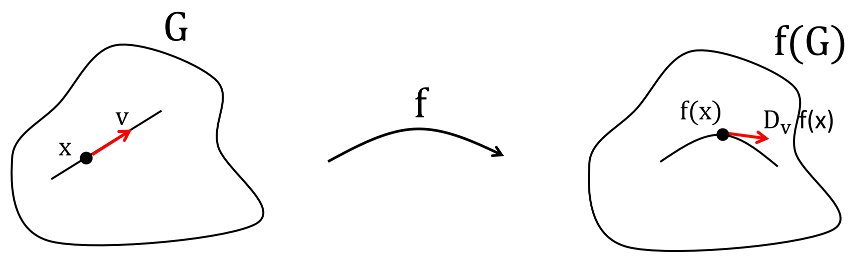
To prepare the proof we need some preliminary remarks on differentiability and continuous dependence.

1. Let G be an open subset of $\mathbb{R}^N$ and $f : G \subset \mathbb{R}^N \to \mathbb{R}^M$. Given $x \in G$ and $v \in \mathbb{R}^N$, the *directional derivative* of f at x in the direction of v is defined as the limit

$$D_v f(x) = \lim_{h\to 0} \frac{1}{h}[f(x + hv) - f(x)]$$

whenever it exists.

The partial derivatives correspond to the cases $v = e_i$ where $\{e_1, \dots, e_N\}$ is the canonical basis of $\mathbb{R}^N$, $\frac{\partial f}{\partial x_i}(x) = D_{e_i} f(x)$, $i = 1, \dots, N$.



A well known result says that f is in $C^1(G, \mathbb{R}^M)$ if the partial derivatives exist at every point of G and the maps

$$x \in G \mapsto \frac{\partial f}{\partial x_i}(x) \in \mathbb{R}^M, \quad i = 1, \ldots, N$$

are continuous.[1] Under these assumptions all directional derivatives can be recovered from the partial derivatives via the formula

$$D_v f(x) = f'(x)v$$

where $f'(x)$ is the Jacobian matrix

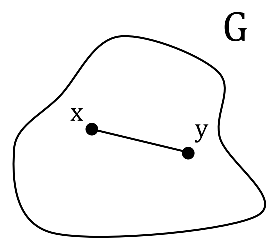
$$f'(x) = (\frac{\partial f}{\partial x_1}(x)| \cdots |\frac{\partial f}{\partial x_N}(x)) \in \mathbb{R}^{M \times N}.$$

2. *A formula for finite increments*
In the previous notations assume that $f \in C^1(G, \mathbb{R}^M)$ and the points $p, q \in G$ are such that the segment joining them is contained in G. Then

$$f(p) - f(q) = [\int_0^1 f'(\lambda p + (1-\lambda)q) d\lambda](p - q).$$

[1] We recall that the definition of C^1 function imposes that f is differentiable and the differential is continous as a map from G to the space of linear maps $\mathcal{L}(\mathbb{R}^N, \mathbb{R}^M)$

Note that the integral in brackets is a matrix in $\mathbb{R}^{M\times N}$ obtained after integrating each entry of the Jacobian matrix.



This formula follows from the application of Barrow's rule to the function

$$\varphi : [0,1] \to \mathbb{R}^M, \quad \varphi(\lambda) = f(\lambda p + (1-\lambda)q).$$

Note that

$$\varphi(1) - \varphi(0) = \int_0^1 \varphi'(\lambda) d\lambda = \int_0^1 [f'(\lambda p + (1-\lambda)q)(p-q)] d\lambda.$$

The last integral belongs to $\mathbb{R}^M$ and can be expressed as above because the vector $p-q$ is constant with respect to λ.

3. *Continous dependence for linear equations: initial conditions and parameters*
Assume that I is an open interval, Λ is an open subset of $\mathbb{R}^m$ and

$$A : I \times \Lambda \to \mathbb{R}^{d\times d}, \quad (t,\lambda) \mapsto A(t,\lambda)$$

is continuous. Given $t_1 \in I$, $y_1 \in \mathbb{R}^d$ and $\lambda \in \Lambda$, the solution of the linear problem

$$\dot{y} = A(t, \lambda)y, \quad y(t_1) = y_1$$

is denoted by $y(t; t_1, y_1, \lambda)$. Then it is known that the map

$$(t; t_1, y_1, \lambda) \in I \times I \times \mathbb{R}^d \times \Lambda \mapsto y(t; t_1, y_1, \lambda) \in \mathbb{R}^d$$

is continuous.

The result has been stated for vector solutions but it is also valid for matrix solutions. Note that each column will be a vector solution.

After this preparation we are ready to prove that the general solution is C^1. We will take a point $(\tau; t_0^*, x_0^*) \in \mathcal{D}$ and we will prove that there exists an open neighborhood $\mathcal{U}$ of this point such that $x \in C^1(\mathcal{U}, \mathbb{R}^M)$. Since the point is arbitrary we can conclude that $x \in C^1(\mathcal{D}, \mathbb{R}^M)$. We proceed in three steps.

First step: the construction of $\mathcal{U}_{(\tau; t_0^*, x_0^*)}$
Assume that $\tau \geq t_0^*$. The other case is similar. The definition of $\mathcal{D}$ implies that

$$\alpha(t_0^*, x_0^*) < t_0^* \leq \tau < \omega(t_0^*, x_0^*).$$

We fix $\nu > 0$ such that $x(t; t_0^*, x_0^*)$ is well defined on $[t_0^* - \nu, \tau + \nu]$. The graph of the solution in this interval,

$$\{(t, x(t; t_0^*, x_0^*)) : \; t \in [t_0^* - \nu, \tau + \nu]\}$$

is a compact subset of D. Then we can find $\rho > 0$ such that the tube

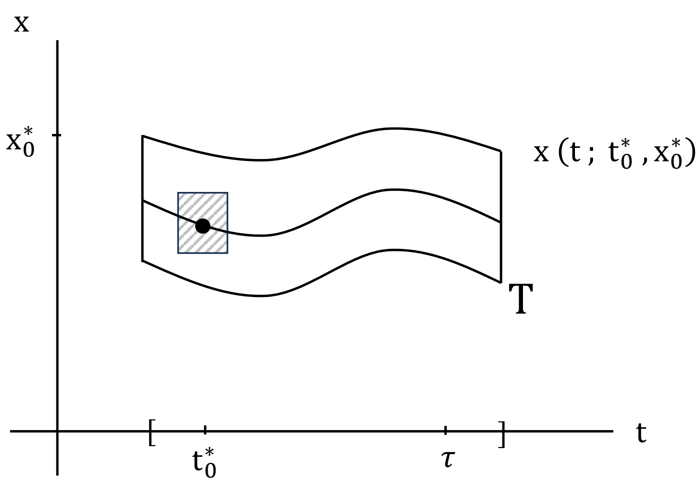
$$T = \{(t, x) \in \mathbb{R} \times \mathbb{R}^d : \; t \in [t_0^* - \nu, \tau + \nu], \; ||x - x(t; t_0^*, x_0^*)|| \leq \rho\}$$

is also contained in D.

Finally we apply the theorem on continuous dependence with respect to initial conditions[2] to find a number $\delta > 0$ such that for any (t_0, x_0) with $|t_0 - t_0^*| \le \delta$ and $||x_0 - x_0^*|| \le \delta$, the solution $x(t; t_0, x_0)$ is well defined on $[t_0^* - \nu, \tau + \nu]$ and

$$(t, x(t; t_0, x_0)) \in T \ \text{ if } t \in [t_0^* - \nu, \tau + \nu].$$

We also assume $\delta < \frac{\nu}{2}$. In this way we know that $t_0^* - \nu < t_0 < \tau + \nu$ if $|t_0 - t_0^*| \le \delta$.



We take $\mathcal{U} = I \times B$ with $I =]t_0^* - \nu, \tau + \nu[$ and

$$B = \{(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^d : \ |t_0 - t_0^*| < \frac{\delta}{2}, \ ||x_0 - x_0^*|| < \frac{\delta}{2}\}.$$

Second step: existence of directional derivatives

[2] $\mathcal{D}$ is open and the general solution is continuous

For each $(t; t_0, x_0) \in \mathcal{U}$ the following directional derivatives exist and satisfy

$$D_{(1;0,0)}x(t; t_0, x_0) = X(t, x(t; t_0, x_0))$$

$$D_{(0;0,v)}x(t; t_0, x_0) = Y(t; t_0, x_0)v, \quad v \in \mathbb{R}^d$$

$$D_{(0;1,0)}x(t; t_0, x_0) = -Y(t; t_0, x_0)X(t_0, x_0).$$

The first identity is a direct consequence of the differential equation because $D_{(1;0,0)}x$ is precisely the time derivative $\dot{x}$.

To prove the second identity we observe that it is not restrictive to assume $||v|| = 1$. With this assumption we can define the finite increment

$$\Delta_h(t) = \frac{1}{h}[x(t; t_0, x_0 + hv) - x(t; t_0, x_0)], \quad t \in I$$

if $0 < |h| < \frac{\delta}{2}$. Note that $\Delta_h(t_0) = v$. The function Δ_h also depends on v but this will not be relevant for the discussion below.

Claim: *There exists a continuous matrix-valued function*

$$(t, h) \in I \times] - \frac{\delta}{2}, \frac{\delta}{2}[\mapsto A(t, h) \in \mathbb{R}^{d \times d}$$

such that, for $h \neq 0$, $\Delta_h(t)$ is the solution of

$$\dot{y} = A(t, h)y, \quad y(t_0) = v.$$

Moreover,

$$A(t, 0) = \frac{\partial X}{\partial x}(t, x(t; t_0, x_0)).$$

To prove this claim we differentiate with respect to time,

$$\dot{\Delta}_h(t) = \frac{1}{h}[X(t, x(t; t_0, x_0 + hv)) - X(t, x(t; t_0, x_0))]$$

and apply the formula for finite increments in **2** to the function $f_t(x) = X(t,x)$ and the points $p = x(t;t_0,x_0+hv)$, $q = x(t;t_0,x_0)$. Due to the geometry of the tube $T \subset D$ we know that f_t is C^1 on the ball centered at $x(t;t_0^*,x_0^*)$ with radius ρ. The points p and q belong to this ball and so the whole segment $[p,q]$ is also contained. In consequence,

$$\dot{\Delta}_h(t) = A(t,h)\Delta_h(t),$$

with

$$A(t,h) = \int_0^1 \frac{\partial X}{\partial x}(t, \lambda x(t;t_0,x_0+hv)+(1-\lambda)x(t;t_0,x_0))d\lambda.$$

The continuity of this matrix is a consequence of the theorem on continuity for integrals depending of parameters. Note that our parameter is (t,h). The assertion on $A(t,0)$ is a direct consequence of the definition of $A(t,h)$.

Once the claim has been proved we apply the continuous dependence for linear equations **3** with $\Lambda =]-\frac{\delta}{2}, \frac{\delta}{2}[$, $t_1 = t_0$, $y_1 = v$. Then, as $h \to 0$, the function $\Delta_h(t)$ converges to the solution of the variational equation $\dot{y} = A(t,0)y$ with initial condition $y(t_0) = v$. This is precisely $Y(t;t_0,x_0)v$, completing the proof of the second identity.

The proof of the third identity is similar but there are some additional details. The increment

$$\Delta_h(t) = \frac{1}{h}[x(t;t_0+h,x_0) - x(t;t_0,x_0)], \quad t \in I$$

is well defined if $0 < |h| < \frac{\delta}{2}$ and satisfies an analogous linear differential equation but now $\Delta_h(t_0)$ is not constant. In fact,

$$\Delta_h(t_0) = \frac{1}{h}[x(t_0;t_0+h,x_0) - x_0]$$

and we must prove that $\Delta_h(t_0) \to -X(t_0, x_0)$ as $h \to 0$. Using Volterra's integral equation we express the increment at t_0 as

$$\Delta_h(t_0) = \frac{1}{h} \int_{t_0+h}^{t_0} X(s, x(s; t_0 + h, x_0))ds.$$

The limit is obtained as an application of the following

Exercise *If $f = f(t, s)$ is continuous in a neighborhood of (T, T) then*

$$\frac{1}{h} \int_{T}^{T+h} f(s, T + h)ds \to f(T, T) \ \ \textit{as } h \to 0.$$

Finally we let h to go to 0 and apply **3** with $\Lambda =]-\frac{\delta}{2}, \frac{\delta}{2}[$, $t_1 = t_0$, $y_1 = \Delta_h(t_0)$.

Third step: conclusion
From the previous step we know that all the partial derivatives of the general solution exist on $\mathcal{U}$ and it remains to prove that they are continuous. This is a consequence of the formulas for the partial derivatives obtained in the second step together with the continuity of $Y(t; t_0, x_0)$ on $\mathcal{U}$. Once again we have applied **3**.