Surjectivité de la direction du gradient dans la sphère unité et zéros bornés non critiques

On se place dans l'espace euclidien $(E, \langle \cdot | \cdot \rangle)$ et on note $\|\cdot\|$ la norme euclidienne ainsi que $\mathcal{S} := \mathrm{S}_{\|\cdot\|_2}(0_E, 1)$ la sphère unité.

Soit une fonction $f : E \longrightarrow \mathbb{R}$ différentiable, et notons $\mathcal{Z}_f := \left\{ z \in E \,:\, f(z) = 0 \right\}$ ses zéros, que l'on suppose non critiques :

$$\forall z \in \mathcal{Z}_f, \quad \nabla f(z) \neq 0$$

On suppose de plus que $\mathcal{Z}_f$ est **borné**. Montrons alors que la direction du gradient

$$\begin{array}{rcl} \mathcal{Z}_f & \longrightarrow & \mathcal{S} \\ z & \longmapsto & \dfrac{\nabla f(z)}{\|\nabla f(z)\|} \end{array}$$

est surjective, c'est à dire que le gradient parcourt toutes les directions. Illustration :

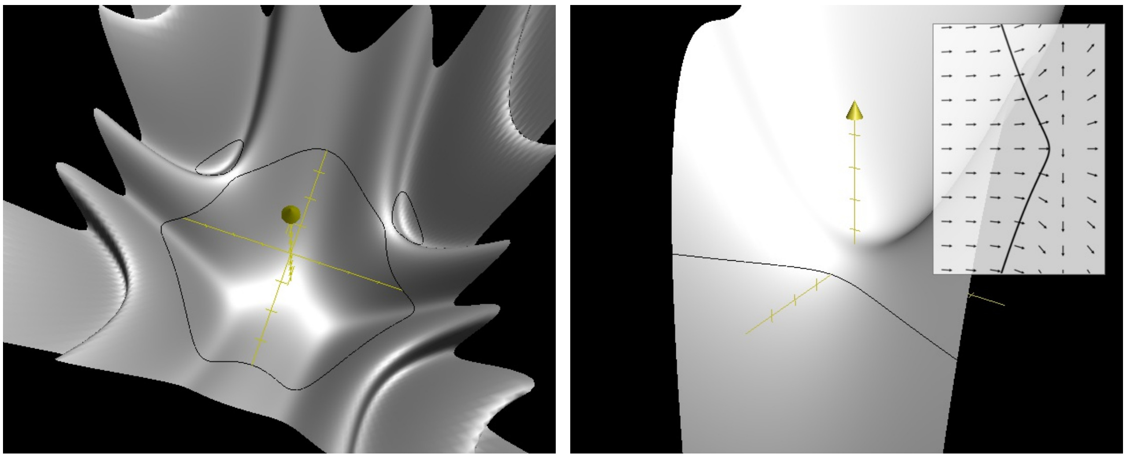


Figure. Zéros de deux fonctions (ligne de niveau noire) :

- à gauche, une fonction pour laquelle $\mathcal{Z}$ est borné : on remarque alors que le gradient (non nul) pointe vers toutes les directions le long de la ligne des zéros
- à droite, une fonction pour laquelle $\mathcal{Z}$ est **non** borné : alors le gradient en $\mathcal{Z}$ ne parcourt visiblement pas toutes les directions

On ne peut avoir un zéro isolé, car ce serait un point critique, donc un point où le gradient est nul.

Remarquons d'abord que $\mathcal{Z}_f$ est **compact** car fermé et borné (dimension finie). En effet, c'est l'image réciproque par contraposée, si $\mathcal{Z}_f$ n'était pas fermé

Soit $v \in \mathcal{S}$. Puisque l'application $z \mapsto \langle v | z \rangle$ est continue sur le compact $\mathcal{Z}_f$ non vide, elle y est bornée et atteint ses bornes. Posons alors

$$\boxed{z \in \mathcal{Z}_f \quad \text{tel que} \quad \langle v | z \rangle = \max_{x \in \mathcal{Z}_f} \langle v | x \rangle}$$

et montrons que $\nabla f(z) \in \mathbb{R}\, v$.

Dans un premier temps, $\mathcal{Z}_f$ étant bornée, $\exists R > 0 \,:\, \mathcal{Z}_f \subset B^{\mathsf{f}}(0, R)$. Posons alors $C := {}^{\complement}B^{\mathsf{f}}(0, R)$, qui est une partie connexe par segments donc **connexe**. Alors, par continuité de f, l'image $f(C) \subset \mathbb{R}$ est connexe (théorème des valeurs intermédiaires), c'est donc un **intervalle**. Par contruction, $C \cap \mathcal{Z}_f = \varnothing$, donc $0 \notin f(C)$, donc $f(C) \subset \mathbb{R}_+^*$ ou $f(C) \subset \mathbb{R}_-^*$. Dans le deuxième cas, on remplace par la suite f par $-f$, sans perte de généralité.

Montrons alors que

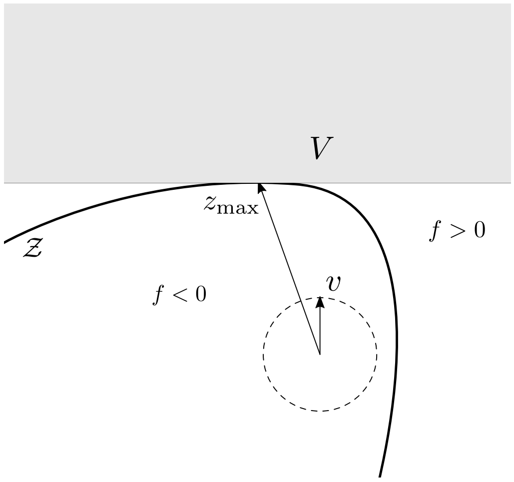
$$\boxed{\begin{array}{c} f(V)\subset\mathbb{R}_+^* \quad \text{avec} \\ V:=\left\{x\in E\,:\,\langle v\,|\,x-z\rangle>0\right\}\end{array}}$$

Soit $x\in V$. Alors $x\in X:=x+\mathbb{R}_+\,v$. On voit que X est une partie connexe par segments intersectant C (en $x+(R+\|x\|)\,v$ par exemple). Ainsi, $X\cup C$ est connexe par segments, donc **connexe**. En appliquant de nouveau le théorème des valeurs intermédiaires, on obtient que $f(X\cup C)$ est un **intervalle**. Or, par contruction de z,

$$\langle v\,|\,x\rangle>\langle v\,|\,z\rangle=\max_{x\in\mathcal{Z}}\langle v\,|\,x\rangle \quad \text{donc} \quad x\notin\mathcal{Z}_f$$

Donc $(X\cup C)\cap\mathcal{Z}_f=\varnothing$, donc $0\notin f(X\cup C)$, donc $f(X\cup C)\subset\mathbb{R}_+^*$. En particulier, $f(x)>0$.

Enfin, on a $\left\{x\in E\,:\,\langle v\,|\,x-z\rangle\geqslant 0\right\}=\bar{V}$ et puisque f est continue, $f(\bar{V})\subset\overline{f(V)}\subset\overline{\mathbb{R}_+^*}=\mathbb{R}_+$.



Situation à v fixé, avec $z_{\max}$ maximisant $\langle v\,|\,z\rangle$. Il est assez peu intuitif que $\nabla f(z_{\max})$ est dans la direction de v. On peut voir les vecteurs u et la frontière de V ($\{v\}^\perp$) comme tengents à $\mathcal{Z}_f$ en $z_{\max}$.

Soit $u\in\{v\}^\perp$, et posons l'application partielle $\begin{array}{rcl} g:\mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & f(z+t\,u)\end{array}$ et montrons qu'elle admet un minimum local en 0. Puisque $g(0)=f(z)=0$, il suffit de montrer que $\forall t\in\mathbb{R},\ g(t)\geqslant 0$:

$$\begin{array}{rl} \forall t\in\mathbb{R}, & \langle v\,|\,(z+t\,u)-z\rangle=\langle v\,|\,u\rangle=0 \quad \text{car} \quad u\perp v \\ & \text{donc} \quad z+t\,u\in\bar{V} \\ & \text{donc} \quad g(t)=f(z+t\,u)\in f(\bar{V})\subset\mathbb{R}_+ \end{array}$$

De plus, g est dérivable en 0 car f différentiable en z et $\partial g(0)=\partial_u f(z)$ par définition. Donc 0 étant un extremum de g, $\partial g(0)=\partial_u f(z)=\mathrm{d}f(z)(u)=\left\langle\nabla f(z)\,|\,u\right\rangle=0$ donc $\nabla f(z)\perp u$. C'est vrai pour tout $u\perp v$, donc $\nabla f(z)\in\{v\}^{\perp\perp}$ donc (dimension finie)

$$\boxed{\nabla f(z)\ \in\ \mathbb{R}\,v}$$

Ce n'est pas encore tout à fait suffisant pour que l'on ait la surjectivité, car il se pourrait que la direction du gradient ne parcourt qu'une hémisphère. Mais il est clair que

$$\forall t\in\mathbb{R}_+^*, \quad z+t\,v\in V \quad \text{donc} \quad f(z+t\,v)\in f(V)\subset\mathbb{R}_+^* \quad \text{donc} \quad \frac{f(z+t\,v)-f(z)}{t}>0$$

donc par passage à la limite,

$$\big\langle\underbrace{\nabla f(z)}_{\in\,\mathbb{R}v}\,|\,v\big\rangle \ = \ \partial_v f(z) \ = \ \lim_{t\to 0^+}\frac{f(z+t\,v)-f(z)}{t} \ \geqslant \ 0 \quad \text{donc} \quad \boxed{\nabla f(z)\ \in\ \mathbb{R}_+\,v}$$

(f étant différentiable, les dérivées partielles à gauche et à droite de f en direction v coïncident). On a donc que $\nabla f(z)=\|\nabla f(z)\|\,v$ puisque v est unitaire, avec $\|\nabla f(z)\|\neq 0$. Concluons :

$$\boxed{\forall v\in\mathcal{S}, \quad \frac{\nabla f(z_v)}{\|\nabla f(z_v)\|}=v \quad \text{avec} \quad z_v=\operatorname*{arg\,max}_{x\in\mathcal{Z}_f}\langle v\,|\,x\rangle\in\mathcal{Z}_f}$$