CALCUL DIFFERENTIEL ET OPTIMISATION

Soient E et F deux $\mathbb{R}$-espaces vectoriels de dimension finie.

Définition Soit $f : U \to F$ une application de classe $\mathcal{C}^1$ sur U ouvert de E. On dit que f est un $\mathcal{C}^1$-difféomorphisme de U sur $V \subset F$ lorsque
- f est bijective de U sur V,
- f^{-1} est de classe $\mathcal{C}^1$ de V sur U.

Remarque Si f est un $\mathcal{C}^1$-difféomorphisme, alors E et F ont la même dimension, puisque pour un $x \in U$, $df(x)$ est bijective, or c'est une application linéaire de E vers F.
($df(x)$ bijective est obtenu en écrivant la différentielle de la composée $f^{-1} \circ f = \mathrm{Id}_U$.)

Quitte à choisir des bases de E et F respectivement, on peut supposer que $E = F = \mathbb{R}^n$.

Problématique
Soit I un intervalle de $\mathbb{R}$. Soit $f : I \to f(I)$ de classe $\mathcal{C}^1$. On suppose que f' ne s'annule pas sur I. On sait alors que f est un $\mathcal{C}^1$-difféomorphisme.
Peut-on généraliser pour $f : U \to f(U)$ avec $U \subset \mathbb{R}^n$ et $f(U) \subset \mathbb{R}^n$?

La réponse est oui, au moins localement.

Théorème 1 (Inversion locale)
Soit $f \in \mathcal{C}^1(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$. Soit $a \in U$. On suppose que $df(a) \in GL(\mathbb{R}^n)$. Alors il existe un ouvert V de $\mathbb{R}^n$ avec $a \in V \subset U$ et un ouvert W de $\mathbb{R}^n$ tels que $f_{|V}$ réalise un $\mathcal{C}^1$-difféormorphisme de V sur W.

□ Quitte à remplacer f par $x \mapsto df(a)^{-1}(f(a+x) - f(a))$, on peut supposer que $a = 0$, $f(a) = 0$ et $df(a) = \mathrm{Id}_{\mathbb{R}^n}$.
f étant de classe $\mathcal{C}^1$, il existe $r \in \mathbb{R}_+^*$ tel que $B(0,r) \subset U$ et $\forall x \in B(0,r), \quad \||df(x) - \mathrm{Id}_{\mathbb{R}^n}\|| \leq \frac{1}{2}$.
- Soit $x \in B(0,r)$. Montrons que $df(x) \in GL(\mathbb{R}^n)$.
  Soit $h \in \ker(df(x))$. On a $\|0 - h\| \leq \frac{1}{2}\|h\|$ et donc $h = 0$. $df(x)$ est injective et est un endomorphisme de $\mathbb{R}^n$ de dimension finie, donc $df(x)$ est bijective.

Intéressons-nous à $\||df(x)^{-1}\||$.
  $\||\mathrm{Id}_{\mathbb{R}^n} - df(x)^{-1}\|| \leq \||df(x)^{-1}\|| \, \||df(x) - \mathrm{Id}_{\mathbb{R}^n}\|| \leq \frac{1}{2}\||df(x)^{-1}\||$. Alors
  $\||df(x)^{-1}\|| \leq \||\mathrm{Id}_{\mathbb{R}^n}\|| + \frac{1}{2}\||df(x)^{-1}\||$ et finalement, $\||df(x)^{-1}\|| \leq 2$.
- On va montrer que tout élément de $B\left(0, \frac{r}{2}\right)$ possède un unique antécédent par f dans $B(0,r)$.
  Soit $y \in B\left(0, \frac{r}{2}\right)$. Posons $h_y : B'(0,r) \to \mathbb{R}^n$, $x \mapsto y + x - f(x)$. On va montrer que h_y possède un unique point fixe grâce au théorème du point fixe.
  h_y est de classe $\mathcal{C}^1$ sur $B(0,r)$ et pour $x \in B(0,r)$, $dh_y(x) = \mathrm{Id}_{\mathbb{R}^n} - df(x)$ et donc $\||dh_y(x)\|| \leq \frac{1}{2}$.
  Par inégalité des accroissements finis (on peut choisir r pour que $B'(0,r) \subset U$), pour tout $(x, x') \in B'(0,r)^2$, $\|h_y(x) - h_y(x')\| \leq \frac{1}{2}\|x - x'\|$: h_y est $\frac{1}{2}$-lipschitzienne.
  Pour $x \in B'(0,r)$, (avec $x' = 0$) $\|h_y(x)\| \leq \|h_y(0)\| + \frac{1}{2}\|x\| \leq \|y\| + \frac{r}{2} < r$. h_y est donc à valeurs dans $B(0,r)$.
  On peut appliquer le théorème du point fixe : il existe une unique $x \in B'(0,r)$ tel que $h_y(x) = x$.
  Mais $x = h_y(x)$, donc $x \in B(0,r)$. Il existe donc un unique $x \in B(0,r)$ tel que $f(x) = y$.
- Posons $V = f^{-1}\left(B\left(0, \frac{r}{2}\right)\right) \cap B(0,r)$: V est un voisinage ouvert de 0 car $f(0) = 0$ et f est continue en 0. Posons $W = B\left(0, \frac{r}{2}\right)$: W est un ouvert. On vient de montrer que $f_{|V} : V \to W$ est bijective. Notons g sa bijection réciproque. Il reste à voir que g est de classe $\mathcal{C}^1$.

- Montrons que g est continue. Pour ce faire, on va montrer que g est lipschitzienne.
  Soit $(y, y') \in W^2$. On a vu que $h_0 : x \mapsto x - f(x)$ est $\frac{1}{2}$-lipschitzienne. Pour tout $x \in B(0, r)$, $x = f(x) + h_0(x)$. C'est en particulier vrai pour $g(y)$ et $g(y')$.

$$\|g(y) - g(y')\| = \|f(g(y)) - f(g(y')) + h_0(g(y)) - h_0(g(y'))\|$$

$$\|g(y) - g(y')\| \leq \|y - y'\| + \frac{1}{2}\|g(y) - g(y')\|$$

  et finalement,

$$\|g(y) - g(y')\| \leq 2\|y - y'\|.$$

  g est 2-lipschitzienne et donc continue.
- Montrons que g est différentiable.
  Soit $y \in W$. Posons $x \in V$ tel que $g(y) = x$ (et donc $y = f(x)$).
  Soit $\varepsilon \in \mathbb{R}^n$ tel que $y + \varepsilon \in W$. Posons $\Delta(\varepsilon) = g(y + \varepsilon) - g(y) - df(x)^{-1}(\varepsilon)$.
  Cette expression a bien un sens, puisque $x \in B(0, r)$ et on a vu tout au début de la démonstration qu'alors $df(x) \in GL(\mathbb{R}^n)$. Si on note $\tilde{\varepsilon} = g(y + \varepsilon) - g(y)$,

$$\|\tilde{\varepsilon}\| \leq 2\|y + \varepsilon - y\| \leq 2\|\varepsilon\|.$$

  On a, par ailleurs,

$$\Delta(\varepsilon) = df(x)^{-1}\left(df(x)(\tilde{\varepsilon}) - \varepsilon\right)$$

  et donc

$$\|\Delta(\varepsilon)\| \leq |||df(x)^{-1}|||\ \|df(x)(\tilde{\varepsilon}) - \varepsilon\|.$$

  Or $x + \tilde{\varepsilon} = g(y + \varepsilon)$, donc $\varepsilon = y + \varepsilon - y = f(x + \tilde{\varepsilon}) - f(x)$. Il s'en suit

$$\|\Delta(\varepsilon)\| \leq 2\|f(x + \tilde{\varepsilon}) - f(x) - df(x)(\tilde{\varepsilon})\|.$$

  on a donc $\Delta(\varepsilon) = o(\tilde{\varepsilon})$ et finalement, $\Delta(\varepsilon) = o(\varepsilon)$.
- On a donc montré que g est différentiable sur W et que pour $y \in W$, $dg(y) = df(g(y))^{-1}$.

g est de classe $\mathcal{C}^1$ car g est continue, df est continue et $\varphi \mapsto \varphi^{-1}$ est continue. □

Théorème 2 (**Inversion globale**)
Soit $f \in \mathcal{C}^1(U, \mathbb{R}^n)$ avec U ouvert de $\mathbb{R}^n$. On suppose que pour tout $x \in U$, $df(x) \in GL(\mathbb{R}^n)$. Alors $W = f(U)$ est un ouvert de $\mathbb{R}^n$.
Si de plus f est injective, alors f est un $\mathcal{C}^1$-difféomorphisme de U sur W.

□ • Soit $y \in W = f(U)$. Il existe $x \in U$ tel que $y = f(x)$. $df(x)$ est inversible. On peut appliquer le théorème d'inversion locale. Il existe un voisinage ouvert V de x dans U et un ouvert W de $\mathbb{R}^n$ tel que f réalise une bijection de V sur W. $x \in V$, donc $y = f(x) \in W$ et comme $W = f(V) \subset f(U)$, W est un voisinage ouvert de y dans $f(U)$. On a bien montré que $W = f(U)$ est ouvert.
• Supposons de plus f injective. Alors f est bien bijective de U sur $W = f(U)$. Par le théorème d'inversion locale, f^{-1} est localement de classe $\mathcal{C}^1$, donc est de classe $\mathcal{C}^1$. □

On va maintenant étudier une application importante du théorème d'inversion locale : le théorème des fonctions implicites. Il dit, par exemple, que si on a une courbe ou une surface donnée par une équation cartésienne et qu'en un point de cette courbe ou surface, la dérivée partielle par rapport à une des coordonnées est non nulle, alors on peut paramétrer, au voisinage de ce point, la courbe ou la surface par les autres coordonnées.

Théorème 3 (**Cas particulier du théorème des fonctions implicites**)
Soit $f : U \mapsto \mathbb{R}$, U étant un ouvert de $\mathbb{R}^n$. On suppose que f est de classe $\mathcal{C}^1$ sur U. Soit $a \in U$ tel que $\frac{\partial f}{\partial x_n}(a) \neq 0$. On note $a = (a_1, \dots, a_{n-1}, a_n) = (b, c)$, avec $b = (a_1, \dots, a_{n-1})$ et $c = a_n$. Alors il existe un voisinage ouvert W de b, un voisinage ouvert V de a et une application $\varphi \in \mathcal{C}^1(W, \mathbb{R})$ tels que

$$\begin{aligned}&((x_1, \dots, x_{n-1}, x_n) \in V \text{ et } f(x_1, \dots, x_{n-1}, x_n) = f(a))\\ &\qquad\qquad\iff \quad ((x_1, \dots, x_{n-1}) \in W \text{ et } x_n = \varphi(x_1, \dots, x_{n-1}).\end{aligned}$$

□ Posons $F : U \to \mathbb{R}^n$, $(x_1, \dots, x_{n-1}, x_n) \mapsto (x_1, \dots, x_{n-1}, f(x_1, \dots, x_{n-1}, x_n))$.

• F est de classe $\mathcal{C}^1$ sur U et pour $x = (x_1, \dots, x_n) \in U$, et $h = (h_1, \dots, h_n) \in \mathbb{R}^n$,

$$dF(x)(h) = ((h_1, \dots, h_{n-1}), df(x)(h)).$$

En particulier, en a :

$$dF(a)(h) = (h_1, \dots, h_{n-1}, \sum_{k=1}^{n-1} \frac{\partial f}{\partial x_k}(a) h_k + \frac{\partial f}{\partial x_n}(a) h_n).$$

• On va alors exploiter le fait que $\frac{\partial f}{\partial x_n}(a) \neq 0$ pour montrer que $dF(a)$ est inversible.
Soit $H \in \mathbb{R}^n$. On écrit $H = (H_1, \dots, H_n)$ et $K = (H_1, \dots, H_{n-1})$. On pose alors $h = (K, h_n)$ avec

$$h_n = \frac{1}{\frac{\partial f}{\partial x_n}(a)} \left(H_n - \sum_{k=1}^{n-1} \frac{\partial f}{\partial x_k}(a) H_k \right).$$

On a bien $df(a)(h) = H$ et $dF(a)$ est surjective. Comme c'est un endomorphisme de $\mathbb{R}^n$, $df(a)$ est bijective.

• On peut appliquer le théorème d'inversion locale.
Il existe V un voisinage ouvert de a dans U et un ouvert $\widetilde{W}$ de $\mathbb{R}^n$ tels que F soit un $\mathcal{C}^1$-difféomorphisme de V sur $\widetilde{W}$. On note g sa bijection réciproque. Elle s'écrit :
$g : (y_1, \dots, y_{n-1}, y_n) \mapsto (y_1, \dots, y_{n-1}, \tilde{\varphi}(y_1, \dots, y_n))$.
$\widetilde{W}$ est un voisinage ouvert de $F(a) = (b, f(a))$.
Quitte à restreindre, on peut écrire $\widetilde{W} = W \times w$ avec W un voisinage ouvert de b et w un voisinage ouvert de $f(a)$. Comme g est de classe $\mathcal{C}^1$, $\tilde{\varphi}$ est aussi de classe $\mathcal{C}^1$. On a bien

$$\begin{aligned}&(x_1, \dots, x_{n-1}, x_n) \in V \text{ et } f(x_1, \dots, x_{n-1}, x_n) = f(a))\\ &\qquad\qquad\iff \quad (x_1, \dots, x_{n-1}) \in W \text{ et } x_n = \tilde{\varphi}(x_1, \dots, x_{n-1}, f(a)).\end{aligned}$$

Il suffit donc de poser $\varphi : (x_1, \dots, x_{n-1}) \mapsto \tilde{\varphi}(x_1, \dots, x_{n-1}, f(a))$ pour conclure. □

On peut alors prouver l'implication manquante du cours.

Théorème 4 (**Espace tangent à une partie définie par une équation**)
Soit $g : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ sur un ouvert U de E. On pose $X = \{x \in U \mid g(x) = 0\}$.
On suppose $X \neq \emptyset$. Soit $x \in X$ tel que $dg(x) \neq 0$. Alors

$$T_x X = \mathrm{Ker}(dg(x)).$$

$\square$ • On a déjà prouvé $T_x X \subset \ker(dg(x))$: si $v \in T_x X$, il existe $\varepsilon \in \mathbb{R}_+^*$ et $\gamma :]-\varepsilon, \varepsilon[\to X$, dérivable en 0 avec $\gamma(0) = x$ et $\gamma'(0) = v$. Or, pour tout $t \in]-\varepsilon, \varepsilon[$, $g(\gamma(t)) = 0$, donc en dérivant, $dg(\gamma(0))(\gamma'(0)) = 0$, i.e. $dg(x)(v) = 0$: on a bien montré que $v \in \ker(dg(x))$.
• Montrons la réciproque, admise en cours.
Quitte à choisir une base de E on peut supposer que $E = \mathbb{R}^n$. On note aussi $a = x$ pour harmoniser les notations. Comme $dg(a) \neq 0$, on peut supposer, par exemple que $\frac{\partial g}{\partial x_n}(a) \neq 0$.
On applique alors le théorème des fonctions implicites vu précédemment : il existe un voisinage ouvert V de $a = (a_1, \ldots, a_n)$, un voisinage ouvert W de $b = (a_1, \ldots, a_{n-1})$ et une application $\varphi \in \mathcal{C}^1(W, \mathbb{R})$ tels que

$$x = (x_1, \ldots, x_n) \in V \cap X \quad \iff \quad (x_1, \ldots, x_{n-1}) \in W \text{ et } x_n = \varphi(x_1, \ldots, x_{n-1}).$$

Soit $v \in \ker(dg(a))$. Notons $v = (v_1, \ldots, v_n)$. On a

$$\sum_{i=1}^{n} \frac{\partial g}{\partial x_i}(a) v_i = 0 \quad \text{donc} \quad v_n = -\sum_{i=1}^{n-1} \frac{\frac{\partial g}{\partial x_i}(a)}{\frac{\partial g}{\partial x_n}(a)} v_i.$$

Mais on a aussi, pour $(x_1, \ldots, x_{n-1})$ dans W,

$$g(x_1, \ldots, x_{n-1}, \varphi(x_1, \ldots, x_{n-1})) = 0$$

et donc en calculant les dérivées partielles suivant chacune des variables $x_1, \ldots, x_{n-1}$:
pour $i = 1, \ldots, n-1$

$$\frac{\partial g}{\partial x_i}(a) + \frac{\partial g}{\partial x_n}(a) \frac{\partial \varphi}{\partial x_i}(b) = 0.$$

On peut donc écrire

$$v_n = \sum_{i=1}^{n-1} \frac{\partial \varphi}{\partial x_i}(b) v_i.$$

On note $w = (v_1, \ldots, v_{n-1})$. On pose alors $\gamma : t \mapsto (b + tw, \varphi(b + tw))$ qui est bien définie au voisinage de 0 et qui est à valeurs dans X.
On a bien $\gamma(0) = a$. γ est bien dérivable en 0 et

$$\gamma'(0) = \left(w, \sum_{i=1}^{n-1} \frac{\partial \varphi}{\partial x_i}(b) v_i \right) = (w, v_n) = v.$$

On a bien montré que v est un vecteur tangent à X. $\square$

On peut ensuite se demander, ce qui se passerait si on imposait plus d'une contrainte. On va voir ce qui se passe avec deux contraintes : cela donnera une bonne idée de la généralisation à plus de trois contraintes.

On commence par réécrire le théorème des fonctions implicites.

Théorème 5
Soient $f : U \mapsto \mathbb{R}$ et $g : U \mapsto \mathbb{R}$, U étant un ouvert de $\mathbb{R}^n$. On suppose que f et g sont de classe $\mathcal{C}^1$ sur U. Soit $a \in U$ tel que

$$\begin{vmatrix} \frac{\partial f}{\partial x_{n-1}}(a) & \frac{\partial f}{\partial x_n}(a) \\ \frac{\partial g}{\partial x_{n-1}}(a) & \frac{\partial g}{\partial x_n}(a) \end{vmatrix} \neq 0.$$

On note $a = (a_1, \dots, a_{n-2}, a_{n-1}, a_n) = (b, c)$, avec $b = (a_1, \dots, a_{n-2})$ et $c = (a_{n-1}, a_n)$. Alors il existe un voisinage ouvert W de b, un voisinage ouvert V de a et une application $\varphi \in \mathcal{C}^1(W, \mathbb{R}^2)$ tels que

$$\begin{aligned} &((x_1, \dots, x_{n-1}, x_n) \in V \text{ et } f(x_1, \dots, x_{n-1}, x_n) = f(a) \text{ et } g(x_1, \dots, x_{n-1}, x_n) = g(a)) \\ &\qquad \iff \quad ((x_1, \dots, x_{n-2}) \in W \text{ et } (x_{n-1}, x_n) = \varphi(x_1, \dots, x_{n-2}). \end{aligned}$$

□ Posons

$$\begin{array}{rccl} F & : & U & \rightarrow & \mathbb{R}^n \\ & & (x_1, \dots, x_{n-2}, x_{n-1}, x_n) & \mapsto & (x_1, \dots, x_{n-2}, f(x_1, \dots, x_n), g(x_1, \dots, x_n)). \end{array}$$

- F est de classe $\mathcal{C}^1$ sur U et pour $x = (x_1, \dots, x_n) \in U$, et $h = (h_1, \dots, h_n) \in \mathbb{R}^n$,

$$dF(x)(h) = ((h_1, \dots, h_{n-1}), df(x)(h), dg(x)(h)).$$

En particulier, en a :

$$\begin{aligned} dF(a)(h) \quad = \quad & (h_1, \dots, h_{n-2}, \sum_{k=1}^{n-2} \frac{\partial f}{\partial x_k}(a) h_k + \frac{\partial f}{\partial x_{n-1}}(a) h_{n-1} + \frac{\partial f}{\partial x_n}(a) h_n, \\ & \sum_{k=1}^{n-2} \frac{\partial g}{\partial x_k}(a) h_k + \frac{\partial g}{\partial x_{n-1}}(a) h_{n-1} + \frac{\partial g}{\partial x_n}(a) h_n). \end{aligned}$$

- On va alors exploiter le fait que $\det(M) = \begin{vmatrix} \frac{\partial f}{\partial x_{n-1}}(a) & \frac{\partial f}{\partial x_n}(a) \\ \frac{\partial g}{\partial x_{n-1}}(a) & \frac{\partial g}{\partial x_n}(a) \end{vmatrix} \neq 0$ pour montrer que $dF(a)$ est inversible.

  Soit $H \in \mathbb{R}^n$. On écrit $H = (H_1, \dots, H_n)$ et $K = (H_1, \dots, H_{n-2})$. On pose alors $h = (K, h_{n-1}, h_n)$ avec

$$(h_{n-1}, h_n) = M^{-1} \left(H_{n-1} - \sum_{k=1}^{n-2} \frac{\partial f}{\partial x_k}(a) H_k, H_n - \sum_{k=1}^{n-2} \frac{\partial f}{\partial x_k}(a) H_k \right).$$

  On a bien $df(a)(h) = H$ et $dF(a)$ est surjective. Comme c'est un endomorphisme de $\mathbb{R}^n$, $df(a)$ est bijective.
- On peut appliquer le théorème d'inversion locale.
  Il existe V un voisinage ouvert de a dans U et un ouvert $\widetilde{W}$ de $\mathbb{R}^n$ tels que F soit un $\mathcal{C}^1$-difféomorphisme de V sur $\widetilde{W}$. On note g sa bijection réciproque. Elle s'écrit :
  $g : (y_1, \dots, y_{n-2}, y_{n-1}, y_n) \mapsto (y_1, \dots, y_{n-2}, \tilde{\varphi_1}(y_1, \dots, y_n), \tilde{\varphi_2}(y_1, \dots, y_n))$.
  $\widetilde{W}$ est un voisinage ouvert de $F(a) = (b, f(a), g(a))$.

Quitte à restreindre, on peut écrire $\widetilde{W} = W \times w$ avec W un voisinage ouvert de b et w un voisinage ouvert de $f(a)$. Comme g est de classe $\mathcal{C}^1$, $\tilde{\varphi}$ est aussi de classe $\mathcal{C}^1$. On a bien

$$\begin{aligned}
&(x = (x_1, \ldots, x_n) \in V,\ f(x) = f(a) \text{ et } g(x) = g(a)) \\
\iff \quad &(x_1, \ldots, x_{n-2}) \in W,\ x_{n-1} = \tilde{\varphi}_1(x_1, \ldots, x_{n-2}, f(a), g(a)) \text{ et } x_n = \tilde{\varphi}_2(x_1, \ldots, x_{n-2}, f(a), g(a)).
\end{aligned}$$

Il suffit donc de poser

$$\varphi\ :\ (x_1, \ldots, x_{n-2}) \mapsto (\tilde{\varphi}_1(x_1, \ldots, x_{n-2}, f(a), g(a)), \tilde{\varphi}_2(x_1, \ldots, x_{n-2}, f(a), g(a)))$$

pour conclure. $\square$

Théorème 6 (**Espace tangent à une partie définie par deux équations**)
Soit $(g_1, g_2) \in \mathcal{C}^1(U, \mathbb{R})^2$ avec U un ouvert de E. On pose $X = \{x \in U \mid g_1(x) = g_2(x) = 0\}$.
On suppose $X \neq \emptyset$. Soit $x \in X$ tel que $(dg_1(x), dg_2(x))$ est une famille libre. Alors

$$T_x X = \operatorname{Ker}(dg_1(x)) \cap \operatorname{Ker}(dg_2(x)).$$

$\square$ • Montrons que $T_x X \subset \operatorname{Ker}(dg_1(x)) \cap \operatorname{Ker}(dg_2(x))$.
Soit $v \in T_x X$, il existe $\varepsilon \in \mathbb{R}_+^*$ et $\gamma :]-\varepsilon, \varepsilon[\to X$, dérivable en 0 avec $\gamma(0) = x$ et $\gamma'(0) = v$. Or, pour tout $t \in]-\varepsilon, \varepsilon[$, $g_1(\gamma(t)) = 0$, donc en dérivant, $dg_1(\gamma(0))(\gamma'(0)) = 0$, i.e. $dg_1(x)(v) = 0$: on a montré que $v \in \ker(dg_1(x))$. On montre de même que $v \in \ker(dg_2(x))$. On a bien montré que $v \in \operatorname{Ker}(dg_1(x)) \cap \operatorname{Ker}(dg_2(x))$.
• Montrons la réciproque.
Quitte à choisir une base de E on peut supposer que $E = \mathbb{R}^n$. On note aussi $a = x$ pour harmoniser les notations. Comme $(dg_1(a), dg_2(a))$ est libre, quitte à échanger les variables, on peut supposer, par exemple, que

$$\begin{vmatrix} \frac{\partial g_1}{\partial x_{n-1}}(a) & \frac{\partial g_1}{\partial x_n}(a) \\ \frac{\partial g_2}{\partial x_{n-1}}(a) & \frac{\partial g_2}{\partial x_n}(a) \end{vmatrix} \neq 0.$$

On applique alors le théorème des fonctions implicites vu précédemment : il existe un voisinage ouvert V de $a = (a_1, \ldots, a_n)$, un voisinage ouvert W de $b = (a_1, \ldots, a_{n-2})$ et une application $\varphi \in \mathcal{C}^1(W, \mathbb{R}^2)$ tels que

$$x = (x_1, \ldots, x_n) \in V \cap X \quad \iff \quad (x_1, \ldots, x_{n-2}) \in W \text{ et } (x_{n-1}, x_n) = \varphi(x_1, \ldots, x_{n-2}).$$

On note $\varphi = (\varphi_1, \varphi_2)$.

Soit $v \in \ker(dg_1(a)) \cap \ker(dg_2(a))$. Notons $v = (v_1, \ldots, v_n)$. On a

$$\sum_{i=1}^{n} \frac{\partial g_1}{\partial x_i}(a) v_i = 0 \quad \text{et} \quad \sum_{i=1}^{n} \frac{\partial g_2}{\partial x_i}(a) v_i = 0$$

donc

$$(v_{n-1}, v_n) = -M^{-1} \left(\sum_{i=1}^{n-2} \frac{\partial g_1}{\partial x_i}(a) v_i, \sum_{i=1}^{n-2} \frac{\partial g_2}{\partial x_i}(a) v_i \right)$$

où

$$M = \begin{pmatrix} \frac{\partial g_1}{\partial x_{n-1}}(a) & \frac{\partial g_1}{\partial x_n}(a) \\ \frac{\partial g_2}{\partial x_{n-1}}(a) & \frac{\partial g_2}{\partial x_n}(a) \end{pmatrix}.$$

Mais on a aussi, pour $(x_1, \dots, x_{n-2})$ dans W,

$$g_1(x_1, \dots, x_{n-2}, \varphi_1(x_1, \dots, x_{n-2}), \varphi_2(x_1, \dots, x_{n-2})) = 0$$

et

$$g_2(x_1, \dots, x_{n-2}, \varphi_1(x_1, \dots, x_{n-2}), \varphi_2(x_1, \dots, x_{n-2})) = 0$$

et donc en calculant les dérivées partielles suivant chacune des variables $x_1, \dots, x_{n-2}$:
pour $i = 1, \dots, n-2$

$$\frac{\partial g_1}{\partial x_i}(a) + \frac{\partial g_1}{\partial x_{n-1}}(a)\frac{\partial \varphi_1}{\partial x_i}(b) + \frac{\partial g_1}{\partial x_n}(a)\frac{\partial \varphi_2}{\partial x_i}(b) = 0$$

et

$$\frac{\partial g_2}{\partial x_i}(a) + \frac{\partial g_2}{\partial x_{n-1}}(a)\frac{\partial \varphi_1}{\partial x_i}(b) + \frac{\partial g_2}{\partial x_n}(a)\frac{\partial \varphi_2}{\partial x_i}(b) = 0.$$

On peut donc écrire

$$(v_{n-1}, v_n) = \left(\sum_{i=1}^{n-2} \frac{\partial \varphi_1}{\partial x_i}(b)v_i, \sum_{i=1}^{n-2} \frac{\partial \varphi_2}{\partial x_i}(b)v_i\right).$$

On note $w = (v_1, \dots, v_{n-2})$. On pose alors $\gamma : t \mapsto (b + tw, \varphi_1(b + tw), \varphi_2(b + tw))$ qui est bien définie au voisinage de 0 et qui est à valeurs dans X.
On a bien $\gamma(0) = a$. γ est bien dérivable en 0 et

$$\gamma'(0) = \left(w, \sum_{i=1}^{n-2} \frac{\partial \varphi_1}{\partial x_i}(b)v_i, \sum_{i=1}^{n-2} \frac{\partial \varphi_2}{\partial x_i}(b)v_i\right) = (w, v_{n-1}, v_n) = v.$$

On a bien montré que v est un vecteur tangent à X. $\square$

On peut alors établir le résultat concernant l'optimisation sous deux contraintes.

Théorème 7 (**Optimisation sous deux contraintes**)
Soient f, g_1 et g_2 dans $\mathcal{C}^1(U, \mathbb{R})$, U ouvert de E. On note $X = \{x \in U \mid g_1(x) = g_2(x) = 0\}$.
Soit $x \in X$ tel que $(dg_1(x), dg_2(x))$ soit libre. On suppose aussi que $f_{|X}$ admet un extremum local en x. Alors il existe $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ tel que $df(x) = \lambda_1 dg_1(x) + \lambda_2 dg_2(x)$.

$\square$ On sait déjà que $T_xX \subset \mathrm{Ker}(df(x))$. Or le résultat précédent donne

$$T_xX = \mathrm{Ker}(dg_1(x)) \cap \mathrm{Ker}(dg_2(x)).$$

Donc

$$\mathrm{Ker}(dg_1(x)) \cap \mathrm{Ker}(dg_2(x)) \subset \mathrm{Ker}(df(x)).$$

On note ψ la forme linéaire $df(x)$, φ_1 la forme linéaire $dg_1(x)$ et φ_2 la forme linéaire $dg_2(x)$.
On note $F = \mathrm{Ker}(\varphi_1) \cap \mathrm{Ker}(\varphi_2)$: $F \subset \mathrm{Ker}(\psi)$.
La famille (φ_1, φ_2) étant libre, on peut trouver deux vecteurs non nuls a et b de E tels que
$E = F \oplus \mathrm{Vect}(a) \oplus \mathrm{Vect}(b)$, $\mathrm{Ker}(\varphi_1) = F \oplus \mathrm{Vect}(a)$ et $\mathrm{Ker}(\varphi_2) = F \oplus \mathrm{Vect}(b)$.
Soit $h \in E$. On écrit $h = y + \alpha a + \beta b$, avec $y \in F$ et $(\alpha, \beta) \in \mathbb{R}^2$.
On a $\psi(y) = \alpha\psi(a) + \beta\psi(b)$, $\varphi_1(h) = \beta\varphi_1(b)$ et $\varphi_2(h) = \alpha\varphi_2(a)$. Or $\varphi_1(b) \neq 0$, sinon, $\varphi_1 = 0$ et de même $\varphi_2(a) \neq 0$. Alors, en posant $\lambda_1 = \frac{\psi(b)}{\varphi_1(b)}$ et $\lambda_2 = \frac{\psi(a)}{\varphi_2(a)}$ on a

$$\psi(h) = \lambda_1\varphi_1(h) + \lambda_2\varphi_2(h).$$

On a bien montré l'existence de deux réels λ_1 et λ_2 tels que $\psi = \lambda_1\varphi_1 + \lambda_2\varphi_2$. $\square$

Pour aller plus loin : voici les énoncés du théorème des fonctions implicites, dans le cas général, et le théorème des extrema liés qui généralise les résultats vus pour l'optimisation avec une ou deux contraintes.

Théorème 8 (**des fonctions implicites**)
Soient $(p,q) \in (\mathbb{N}^*)^2$. Soit A un ouvert de $\mathbb{R}^p \times \mathbb{R}^q$. Soit $f = (f_1, \dots, f_q) : A \to \mathbb{R}^q$,
$(x,y) \mapsto (f_1(x,y), \dots, f_q(x,y))$ une application de classe $\mathcal{C}^1$. Soit $(a,b) \in A$.
Si $\det\left(\frac{\partial f_i}{\partial y_j}(a,b)\right)_{\substack{1\leq i\leq q\\ 1\leq j\leq q}} \neq 0$, alors il existe un voisinage ouvert W de a, un voisinage ouvert V de (a,b) et une application $\varphi \in \mathcal{C}^1(W, \mathbb{R}^q)$ tels que

$$((x,y) \in V \text{ et } f(x,y) = f(a,b)) \quad \Longleftrightarrow \quad (x \in W \text{ et } y = \varphi(x)).$$

Théorème 9 (**des extrema liés**)
Soient $(n,p) \in (\mathbb{N}^*)^2$, U un ouvert de $\mathbb{R}^n$ et $(f, g_1, \dots, g_p) \in \mathcal{C}^1(U, \mathbb{R})^{p+1}$.
Soit $A = \{x \in U \mid g_1(x) = \dots = g_p(x) = 0\}$.
Si la restriction de f à A admet un extremum relatif en $a \in A$ et si les formes linéaires $dg_1(a), \dots, dg_p(a)$ sont linéairement indépendantes, alors il existe des réels $\lambda_1, \dots, \lambda_p$ tels que

$$df(a) = \sum_{k=1}^{p} \lambda_k dg_k(a).$$

□ On pose $q = n - p$ et on écrit $\mathbb{R}^n = \mathbb{R}^q \times \mathbb{R}^p$. On écrit $a = (b,c)$ avec $b \in \mathbb{R}^q$ et $c \in \mathbb{R}^p$.
$(dg_1(a), \dots, dg_p(a))$ est libre dans $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ de dimension n, donc $p \leq n$.
Si $p = n$, $(dg_1(a), \dots, dg_p(a))$ est une base de $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$ et donc le résultat est immédiat.
On suppose donc $p < n$, i.e. $q \in \mathbb{N}^*$. La jacobienne de $(g_1, \dots, g_p)$ a un rang égal à p : on peut en extraire une sous-matrice carrée de taille p inversible. Quitte à échanger les variables, on peut supposer que $\left(\frac{\partial g_i}{\partial y_j}(a)\right)_{\substack{1\leq i\leq q\\ 1\leq j\leq p}}$ est inversible. On peut alors appliquer le théorème des fonctions implicites : il existe W un voisinage ouvert de b, V un voisinage ouvert de a et $\varphi \in \mathcal{C}^1(W, \mathbb{R}^p)$ tels que

$$((x,y) \in V \text{ et } g(x,y) = g(a)) \quad \Longleftrightarrow \quad (x \in W \text{ et } y = \varphi(x))$$

avec $g = (g_1, \dots, g_p)$. Notons $\varphi = (\varphi_1, \dots, \varphi_p)$ et posons $h : x \mapsto f(x, \varphi(x))$. h admet un extremum local en b et donc pour $i = 1 \cdots q$

$$\frac{\partial f}{\partial x_i}(a) + \sum_{j=1}^{p} \frac{\partial f}{\partial y_j}(a) \frac{\partial \varphi_j}{\partial x_i}(b) = 0.$$

Mais on a aussi pour $x \in W$ et pour $k = 1 \cdots q$, $g_k(x, \varphi(x)) = 0$ et donc $\frac{\partial g_k}{\partial x_i}(a) + \sum\limits_{j=1}^{p} \frac{\partial g_k}{\partial y_j}(a) \frac{\partial \varphi_j}{\partial x_i}(b) = 0$.
Les q premières colonnes de la matrice

$$\begin{pmatrix} \frac{\partial f}{\partial x_1}(a) & \cdots & \frac{\partial f}{\partial x_q}(a) & \frac{\partial f}{\partial y_1}(a) & \cdots & \frac{\partial f}{\partial y_p}(a) \\ \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_q}(a) & \frac{\partial g_1}{\partial y_1}(a) & \cdots & \frac{\partial g_1}{\partial y_p}(a) \\ & \cdots & & & \cdots & \\ \frac{\partial g_p}{\partial x_1}(a) & \cdots & \frac{\partial g_p}{\partial x_q}(a) & \frac{\partial g_p}{\partial y_1}(a) & \cdots & \frac{\partial g_p}{\partial y_p}(a) \end{pmatrix}$$

sont donc des combinaisons linéaires des p dernières : son rang est au plus p, les $p+1$ lignes de cette matrice sont liées. Les p dernières lignes étant libres puisque la matrice carrée extraite avec les p dernières colonnes et les p dernières lignes est inversible, la première ligne est combinaison linéaire des p dernières et c'est le résultat voulu. □