Théorème des extrema liés

Mohamed NASSIRI

Références:

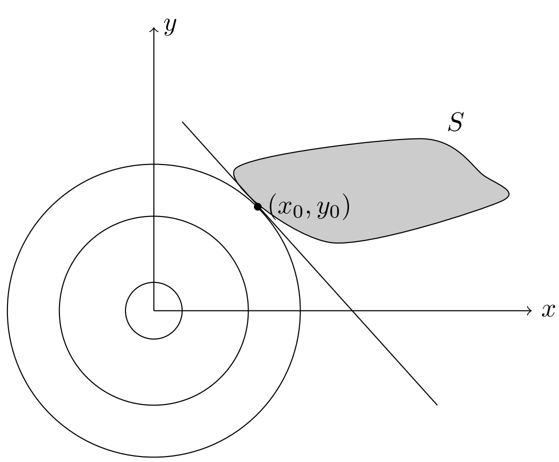
Calcul différentiel topologique élémentaire, Wolfgang Bertram - p.191 $\rightarrow$ 195

Recasage :

●●●●● 214 : Théorème d'inversion locale, théorème des fonctions implicites. Exemples et applications en analyse et en géométrie.
●●●●● 214 : Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.
●●●●● 219 : Extremums : existence, caractérisation, recherche. Exemples et applications.
●●●●○ 159 : Formes linéaires et dualité en dimension finie. Exemples et applications.

Résumé:

Un développement quasi-incontournable pour qui parle d'extrema et d'application différentiables. En guise d'introduction, tentons de visualiser la recherche d'extrema liés, aussi appelés extrema sous contrainte. Prenons la fonction f définie par $f(x, y) = x^2 + y^2$ (qui n'est rien d'autre que la fonction distance au carré) et cherchons son minimum sur la surface S :



Noter que le minimum (libre) de la fonction f est 0 et il est atteint en $(0, 0)$ mais ce point n'est pas dans S.
Pour trouver le minimum sur S de f, on trace les lignes de niveaux de f ($x^2 + y^2 = cte$) et on regarde quand la "première" ligne de niveau touche S. Le point d'intersection sera le minimum recherché.

Mais pourquoi diable parler d'espace tangent ? Comme le montre notre exemple, le minimum recherché se situe sur la tangente commune de la ligne de niveau de f et de la surface S. Le théorème des extrema liés généralise la situation ...

Prérequis:

Applications différentiables - Théorème du rang constant (admis) - Géométrie différentielle (espace tangent) - Formes linéaires

---

Théorème : Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^n$, $S = \{x \in \mathcal{U} \mid g(x) = 0\}$, où $g : \mathcal{U} \to \mathbb{R}^m$ $(m \leq n)$ est de classe C^1. On suppose que, pour tout $x \in S$, $Dg(x)$ est surjective. Soit aussi $f : \mathcal{U} \to \mathbf{R}$ une fonction de classe C^1.
Si $p \in S$ est un point critique de $f_{|S}$, alors il existe une forme linéaire $\lambda : \mathbb{R}^m \to \mathbb{R}$ telle que

$$Df(p) = \lambda \circ Dg(p)$$

Démonstration.
La démonstration se fait en trois étapes. La première étape nous permet de faire le lien entre l'espace tangent de S en p (noté T_pS) et $\mathrm{Ker}(Dg(p))$, ce qui, pour la seconde étape, nous donnera l'inclusion $\mathrm{Ker}(Dg(p)) \subset \mathrm{Ker}(Df(p))$ pour finir par l'existence de la forme linéaire λ.

Etape 1 : Montrons que $T_pS = \mathrm{Ker}(Dg(p))$ *:*
\- $T_pS \subseteq \mathrm{Ker}(Dg(p))$:
Soit $v \in T_pS$. Par définition de l'espace tangent, il existe un intervalle non vide $I =]-\epsilon, \epsilon[$ $(\epsilon > 0)$ et une courbe $\gamma : I \to S$ de classe C^1 telle que $\gamma(0) = p$ et $\gamma'(0) = v$. Donc

$$0 = \left.\frac{d}{dt}\right|_{t=0} g(\gamma(t)) = Dg(\gamma'(0))\gamma'(0) = Dg(p)(v)$$

et donc $v \in \mathrm{Ker}(Dg(p))$.

\- $T_pS \supseteq \mathrm{Ker}(Dg(p))$:
Comme $Dg(p)$ est surjective; d'après le théorème du rang constant, il existe un voisinage U_p de p tel que $Dg(x)$ soit surjective pour tout $x = (x_1, ..., x_n) \in U_p$ et un difféomorphisme local h tel que $h(0) = p$ et

$$g(h(x_1, ..., x_n)) = \Pi_m(x_1, ..., x_n) = (x_1, ..., x_m)$$

Soit $v \in \mathrm{Ker}(Dg(p))$. On pose $w := (Dh(0))^{-1}v$. La courbe $\gamma(t) := h(tw)$, pour t dans un intervalle non vide I, vérifie :

$$(i)\ \gamma(0) = h(0) = p, \quad \gamma'(0) = Dh(0)(Dh(0))^{-1}v = v, \text{ et}$$

$$(iii)\ g(\gamma(t)) = g(h(tw)) = \Pi_m(tw) = tD\Pi_m(0)(w) = tD(g \circ h)(0)(w) = tDg(h(0))Dh(0)(w) = tDg(p)v = 0$$

Donc $\gamma(I) \subset S$ et on a montré que $v \in T_pS$.

Conclusion : $T_pS = \mathrm{Ker}(Dg(p))$

Etape 2 : Montrons que $\mathrm{Ker}(Dg(p)) \subset \mathrm{Ker}(Df(p))$ *:*

Soit $v \in T_pS = \mathrm{Ker}(Dg(p))$, avec p un point critique de $f|_S$. Alors 0 est un point critique de $f \circ \gamma$, et donc

$$0 = (f(\gamma(0)))' = Df(\gamma(0))(\gamma'(0)) = Df(p)(v)$$

Par conséquent, $v \in \mathrm{Ker}(Df(p))$, et ainsi $T_pS = \mathrm{Ker}(Dg(p)) \subset \mathrm{Ker}(Df(p))$.

Etape 3 : Les multiplicateurs de Lagrange :

Grâce à l'inclusion $\mathrm{Ker}(Dg(p)) \subset \mathrm{Ker}(Df(p))$ et du lemme suivant, on conclut.

Lemme :
Soit $\alpha : \mathbb{R}^n \to \mathbb{R}^m$ et $\beta : \mathbb{R}^n \to \mathbb{R}$ des applications linéaires telles que $\mathrm{Ker}(\alpha) \subset \mathrm{Ker}(\beta)$. Alors il existe une forme linéaire $\lambda : \mathbb{R}^m \to \mathbb{R}$ telle que $\beta = \lambda \circ \alpha$.

Démonstraton du lemme.
Soit B la matrice de β (matrice ligne) et A celle de α (la i^e ligne L_i de A est la matrice de la composante de $\alpha_i : \mathbb{R}^n \to \mathbb{R}$). Soit C la matrice obtenue en rajoutant à A la matrice B comme dernière ligne.

$$C = \begin{pmatrix} \boxed{\quad A \quad} \\ \boxed{\quad B \quad} \end{pmatrix}$$

Par hypothèse, $Ax = 0 \Rightarrow Bx = 0$; et donc $Ax = 0 \Leftrightarrow Cx = 0$. Ainsi, A et C ont le même noyau et donc le même rang, par le théorème du rang.
Or cela signifie que la dernière ligne de C est combinaison linéaire des m autres lignes, et il existe ainsi $\lambda_1, ..., \lambda_m \in \mathbb{R}$ tels que $B = \sum_{i=1}^m \lambda_i L_i$, ce qui veut dire que $\beta(x) = \sum_{i=1}^m \lambda_i \alpha_i(x) = \lambda(\alpha(x))$ pour tout $x \in \mathbb{R}^n$

□

□

Remarques :

- On peut encore écrire le théorème des extrema liés de la façon suivante :
  Théorème (bis) : Pour tout $v \in \mathbb{R}^n$, nous avons $Df(p)(v) = \lambda(Dg(p)(v))$, ou encore, en écrivant $\lambda(x) = \sum_{i=1}^m \lambda_i x_i$: il existe $\lambda_1, ...\lambda_m \in \mathbb{R}$ (dits des multiplicateurs de Lagrange) tels que, pour tout $v \in \mathbb{R}^n$,
  $$Df(p)(v) = \sum_{i=1}^m \lambda_i Dg_i(p)(v)$$
- Grâce au théorème des extrema liés, on peut démontrer de façon amusante le théorème spectral :
  Application : Soit $(E, \langle ., . \rangle)$ un espace euclidien, $u \in \mathcal{L}(E)$ symétrique. Alors il existe une base orthonormée de E formée de vecteurs propres de u.
  Considérons
  $$f : \begin{array}{l} E \to \mathbb{R} \\ x \mapsto \langle u(x), x \rangle \end{array} \qquad g : \begin{array}{l} E \to \mathbb{R} \\ x \mapsto \langle x, x \rangle - 1 \end{array}$$
  Alors $S = \{x \in E \mid g(x) = 0\}$ est la sphère unité qui est compacte et donc f, étant continue, atteint son maximum sur S en e_1.
  D'autre part $Df(x)(h) = 2\langle u(x), h \rangle$ et $Dg(x)(h) = 2\langle x, h \rangle$. D'après le théorème des extrema liés, il existe $\lambda_1 \in \mathbb{R}$ tel que $Df(e_1)(h) = \lambda_1 Dg(e_1)(h)$ pour tout $h \in E$ (*i.e.*) $u(e_1) = \lambda_1 e_1$.
  On a donc trouvé une valeur propre λ_1 de vecteur propre e_1 de norme 1. On raisonne alors par récurrence sur $F = e_1^{\perp}$.
- Mises en garde sur le développement : Le choix de cette preuve "version géométrique" est motivé par deux extraits du rapport du jury 2018 :
  "En ce qui concerne la preuve du théorème des extrema liés, la présentation de la preuve par raisonnement « sous-matriciel » est souvent obscure ; on privilégiera si possible une présentation géométrique s'appuyant sur l'espace tangent."
  "Une preuve géométrique des extrema liés sera fortement valorisée par rapport à une preuve algébrique, formelle et souvent mal maîtrisée."