

# Un bestiaire de contre-exemples en optimisation convexe lisse

EDOUARD PAUWELS (IRIT, TOULOUSE 3)  
en collaboration avec JÉRÔME BOLTE (TSE, TOULOUSE 1)

Séminaire français d'optimisation

Juin 2020



**Avertissement:**

Tout se passe dans  $\mathbb{R}^2$ . Certains résultats se formulent dans  $\mathbb{R}^p$ .

$k$  désigne un indice de régularité  $C^k$ ,  $k \geq 2$ .

Cadre convexe, sous-niveaux compacts,  $C^k$ .

$f$  convexe sur  $\mathbb{R}^p$ , gradient  $L$ -Lipschitz, admet un minimum  $x^*$ ,  $x_0 \in \mathbb{R}^p$ , pour  $k \in \mathbb{N}$ ,

$$x_{k+1} = x_k - \frac{1}{L} \nabla f(x_k).$$

La suite  $(x_k)_{k \in \mathbb{N}}$  converge vers une solution (monotonie de Féjer: pour tout minimum  $x^*$ ,  $\|x_k - x^*\|$  décroît).

$f$  convexe sur  $\mathbb{R}^p$ , gradient  $L$ -Lipschitz, admet un minimum  $x^*$ ,  $x_0 \in \mathbb{R}^p$ , pour  $k \in \mathbb{N}$ ,

$$x_{k+1} = x_k - \frac{1}{L} \nabla f(x_k).$$

La suite  $(x_k)_{k \in \mathbb{N}}$  converge vers une solution (monotonie de Féjer: pour tout minimum  $x^*$ ,  $\|x_k - x^*\|$  décroît).

### Recherche linéaire exacte

$$x_{k+1} \in \arg \min_x f(x), \quad t.q. \quad x = x_k - s \nabla f(x_k), \quad s \in \mathbb{R}_+.$$

$f$  convexe sur  $\mathbb{R}^p$ , gradient  $L$ -Lipschitz, admet un minimum  $x^*$ ,  $x_0 \in \mathbb{R}^p$ , pour  $k \in \mathbb{N}$ ,

$$x_{k+1} = x_k - \frac{1}{L} \nabla f(x_k).$$

La suite  $(x_k)_{k \in \mathbb{N}}$  converge vers une solution (monotonie de Féjer: pour tout minimum  $x^*$ ,  $\|x_k - x^*\|$  décroît).

### Recherche linéaire exacte

$$x_{k+1} \in \arg \min_x f(x), \quad t.q. \quad x = x_k - s \nabla f(x_k), \quad s \in \mathbb{R}_+.$$

**Contribution:** ne converge pas en général,

- dimension 2
- potentiel convexe  $C^k$ , coercif
- minimum unique le long de la trajectoire (sélection dans l'argmin).

## La conjecture du gradient de Thom

$f$  est analytique sur  $\mathbb{R}^p$ , système gradient:  $\dot{x}(t) = -\nabla f(x(t))$ .

## La conjecture du gradient de Thom

$f$  est analytique sur  $\mathbb{R}^p$ , système gradient:  $\dot{x}(t) = -\nabla f(x(t))$ .

**Łojasiewicz (60's):** les solutions bornées convergent.

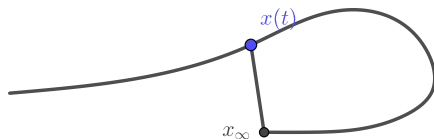
# La conjecture du gradient de Thom

$f$  est analytique sur  $\mathbb{R}^p$ , système gradient:  $\dot{x}(t) = -\nabla f(x(t))$ .

**Łojasiewicz (60's):** les solutions bornées convergent.

**Conjecture du gradient de Thom (70's):** les trajectoires convergentes admettent une tangente à l'infini.

$$\frac{x(t) - x_\infty}{\|x(t) - x_\infty\|} \xrightarrow{t \rightarrow \infty} v \in \mathbb{R}^p.$$





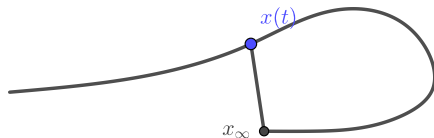
# La conjecture du gradient de Thom

$f$  est analytique sur  $\mathbb{R}^p$ , système gradient:  $\dot{x}(t) = -\nabla f(x(t))$ .

**Łojasiewicz (60's):** les solutions bornées convergent.

**Conjecture du gradient de Thom (70's):** les trajectoires convergentes admettent une tangente à l'infini.

$$\frac{x(t) - x_\infty}{\|x(t) - x_\infty\|} \xrightarrow{t \rightarrow \infty} v \in \mathbb{R}^p.$$



**Kurdyka-Mostowski-Parusinski (2000):** la conjecture du gradient de Thom est vraie.

$f$  convexe  $C^k$  sur  $\mathbb{R}^p$ , admet un minimum  $x^*$ , système gradient,

$$\dot{x}(t) = -\nabla f(x(t)),$$

toutes les solutions convergent vers un minimum (monotonie de Féjer).

$f$  convexe  $C^k$  sur  $\mathbb{R}^p$ , admet un minimum  $x^*$ , système gradient,

$$\dot{x}(t) = -\nabla f(x(t)),$$

toutes les solutions convergent vers un minimum (monotonie de Féjer).

**Convergence des sécantes:**

$$\frac{x(t) - x_\infty}{\|x(t) - x_\infty\|} \xrightarrow{t \rightarrow \infty} v \in \mathbb{R}^p.$$

$f$  convexe  $C^k$  sur  $\mathbb{R}^p$ , admet un minimum  $x^*$ , système gradient,

$$\dot{x}(t) = -\nabla f(x(t)),$$

toutes les solutions convergent vers un minimum (monotonie de Féjer).

## Convergence des sécantes:

$$\frac{x(t) - x_\infty}{\|x(t) - x_\infty\|} \xrightarrow{t \rightarrow \infty} v \in \mathbb{R}^p.$$

**Contribution:** faux en général,

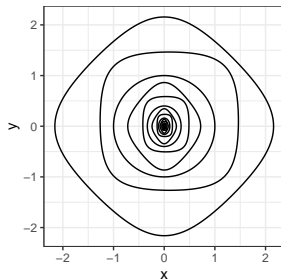
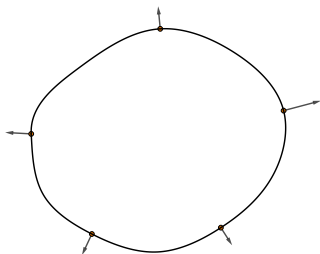
- dimension 2.
- potentiel convexe  $C^k$ , coercif.
- minimum unique.
- hessienne inversible en dehors du minimum.

Montrer l'existence de  $f$ ,  $C^k$ , convexe sur  $\mathbb{R}^2$  dont les gradients oscillent beaucoup autour du minimum.

Montrer l'existence de  $f, C^k$ , convexe sur  $\mathbb{R}^2$  dont les gradients oscillent beaucoup autour du minimum.

## Idée principale:

- Gradient orthogonal aux sous-niveaux.
- Interpolation d'une suite de sous-niveaux.



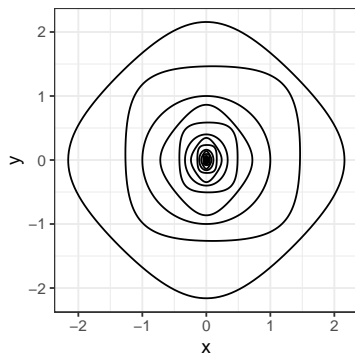
- Interpolation de sous-niveaux convexes à courbure positive.
- Spécification de sous-niveaux convexes à courbure positive.

1. Tour d'horizon pour le problème d'interpolation
2. Rappels sur la courbure et l'approximation de Bernstein
3. Deux résultats d'interpolation dans  $\mathbb{R}^2$
4. Construction des contre exemples



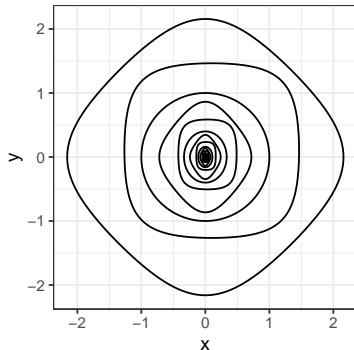
## Description du problème

$(T_i)_{i \in -\mathbb{N}}$  sont convexes compacts tels que  $T_{i-1} \subset \text{int}(T_i) \neq \emptyset$  pour  $i \in -\mathbb{N}$ . Construire  $f$  convexe tel que, pour tout  $i \in -\mathbb{N}$ ,  $T_i$  est un sous-niveau de  $f$ .



## Description du problème

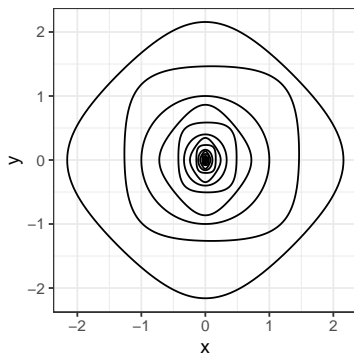
$(T_i)_{i \in -\mathbb{N}}$  sont convexes compacts tels que  $T_{i-1} \subset \text{int}(T_i) \neq \emptyset$  pour  $i \in -\mathbb{N}$ . Construire  $f$  convexe tel que, pour tout  $i \in -\mathbb{N}$ ,  $T_i$  est un sous-niveau de  $f$ .



**Premier questionnements:** de Finetti, Fenchel (50's).

## Description du problème

$(T_i)_{i \in -\mathbb{N}}$  sont convexes compacts tels que  $T_{i-1} \subset \text{int}(T_i) \neq \emptyset$  pour  $i \in -\mathbb{N}$ . Construire  $f$  convexe tel que, pour tout  $i \in -\mathbb{N}$ ,  $T_i$  est un sous-niveau de  $f$ .

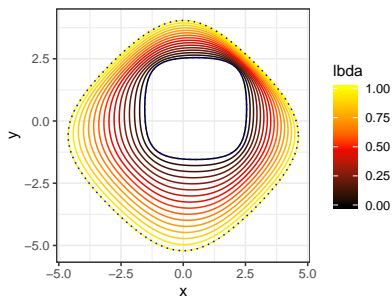


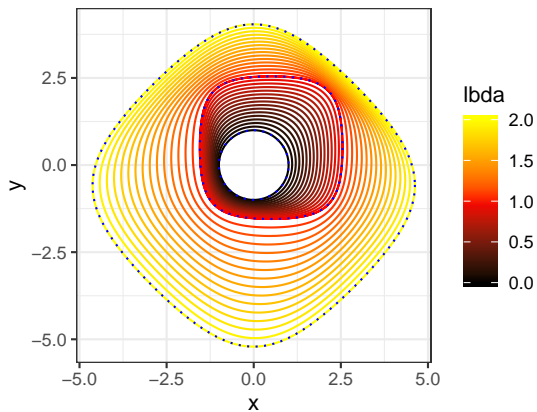
**Premier questionnements:** de Finetti, Fenchel (50's).

**Kannai, Torralba (77, 96):**  $f$  existe.

- $T_0$ , sous-niveau zéro,  $T_1$  sous-niveau 1.
- Pour  $\lambda \in [0, 1]$ ,  $T_\lambda = (1 - \lambda)T_0 + \lambda T_1$  est sous-niveau  $\lambda$ .

- $T_0$ , sous-niveau zéro,  $T_1$  sous-niveau 1.
- Pour  $\lambda \in [0, 1]$ ,  $T_\lambda = (1 - \lambda)T_0 + \lambda T_1$  est sous-niveau  $\lambda$ .





Pour  $S \subset \mathbb{R}^p$ , notons,  $\sigma_S(x) = \sup\{\langle x, z \rangle, z \in S\}$  la fonction support de  $S$ .

**de Finetti-Fenchel-Crouzeix:** Soit  $f$  sur  $\mathbb{R}^p$  quasi-convexe,  $T_\lambda$  le sous niveau  $\lambda$ , alors  $f$  est convexe si et seulement si, pour tout  $v$  dans  $\mathbb{R}^p$ ,

$$F_v: \lambda \mapsto \sigma_{T_\lambda}(v) = \sup\{\langle v, x \rangle, f(x) \leq \lambda\}$$

est concave.

Pour  $S \subset \mathbb{R}^p$ , notons,  $\sigma_S(x) = \sup\{\langle x, z \rangle, z \in S\}$  la fonction support de  $S$ .

**de Finetti-Fenchel-Crouzeix:** Soit  $f$  sur  $\mathbb{R}^p$  quasi-convexe,  $T_\lambda$  le sous niveau  $\lambda$ , alors  $f$  est convexe si et seulement si, pour tout  $v$  dans  $\mathbb{R}^p$ ,

$$F_v: \lambda \mapsto \sigma_{T_\lambda}(v) = \sup\{\langle v, x \rangle, f(x) \leq \lambda\}$$

est concave.

**“Linéarité” de la fonction support:** Soient  $A \subset \mathbb{R}^p$  et  $B \subset \mathbb{R}^p$  compacts, convexes. Pour  $\alpha, \beta > 0$ ,

$$\sigma_{\alpha A + \beta B}(\cdot) = \alpha \sigma_A(\cdot) + \beta \sigma_B(\cdot).$$



Pour  $S \subset \mathbb{R}^p$ , notons,  $\sigma_S(x) = \sup\{\langle x, z \rangle, z \in S\}$  la fonction support de  $S$ .

**de Finetti-Fenchel-Crouzeix:** Soit  $f$  sur  $\mathbb{R}^p$  quasi-convexe,  $T_\lambda$  le sous niveau  $\lambda$ , alors  $f$  est convexe si et seulement si, pour tout  $v$  dans  $\mathbb{R}^p$ ,

$$F_v: \lambda \mapsto \sigma_{T_\lambda}(v) = \sup\{\langle v, x \rangle, f(x) \leq \lambda\}$$

est concave.

**“Linéarité” de la fonction support:** Soient  $A \subset \mathbb{R}^p$  et  $B \subset \mathbb{R}^p$  compacts, convexes. Pour  $\alpha, \beta > 0$ ,

$$\sigma_{\alpha A + \beta B}(\cdot) = \alpha \sigma_A(\cdot) + \beta \sigma_B(\cdot).$$

**Kannai-Torralba:** Interpolation “affine par morceaux”:

- $F_v$  devient affine par morceaux  $\rightarrow$  décroissance des pentes.
- Bien affecter les valeurs de  $f$  pour garantir la convexité.
- Condition suffisante faisant intervenir les fonctions support.

$(T_i)_{i \in -\mathbb{N}}$  sont convexes compacts de frontière  $C^k$ , tels que  $T_{i-1} \subset \text{int}(T_i) \neq \emptyset$  pour  $i \in -\mathbb{N}$ . Construire  $f$  convexe  $C^k$ , tel que, pour tout  $i \in -\mathbb{N}$ ,  $T_i$  est un sous-niveau de  $f$ .

$(T_i)_{i \in -\mathbb{N}}$  sont convexes compacts de frontière  $C^k$ , tels que  $T_{i-1} \subset \text{int}(T_i) \neq \emptyset$  pour  $i \in -\mathbb{N}$ . Construire  $f$  convexe  $C^k$ , tel que, pour tout  $i \in -\mathbb{N}$ ,  $T_i$  est un sous-niveau de  $f$ .

**Comment assurer que tous les sous-niveaux sont  $C^k$ ?**

- Dans  $\mathbb{R}^p$ , si  $A$  et  $B$  sont convexes et de frontière  $C^\infty$ ,  $A + B$  n'est pas nécessairement de frontière  $C^2$  (Kiselman 1986)
- Si  $A \subset \mathbb{R}^2$  et  $B \subset \mathbb{R}^2$  sont de frontière  $C^k$ ,  $A + B$  est de frontière  $C^k$ , uniquement pour  $k = 1, 2, 3, 4$  (Kiselman 1987, Boman 1990).

$(T_i)_{i \in -\mathbb{N}}$  sont convexes compacts de frontière  $C^k$ , tels que  $T_{i-1} \subset \text{int}(T_i) \neq \emptyset$  pour  $i \in -\mathbb{N}$ . Construire  $f$  convexe  $C^k$ , tel que, pour tout  $i \in -\mathbb{N}$ ,  $T_i$  est un sous-niveau de  $f$ .

**Comment assurer que tous les sous-niveaux sont  $C^k$ ?**

- Dans  $\mathbb{R}^p$ , si  $A$  et  $B$  sont convexes et de frontière  $C^\infty$ ,  $A + B$  n'est pas nécessairement de frontière  $C^2$  (Kiselman 1986)
- Si  $A \subset \mathbb{R}^2$  et  $B \subset \mathbb{R}^2$  sont de frontière  $C^k$ ,  $A + B$  est de frontière  $C^k$ , uniquement pour  $k = 1, 2, 3, 4$  (Kiselman 1987, Boman 1990).

**Comment assurer un recollement lisse à la frontière?**

1. Tour d'horizon pour le problème d'interpolation
2. Rappels sur la courbure et l'approximation de Bernstein
3. Deux résultats d'interpolation dans  $\mathbb{R}^2$
4. Construction des contre exemples

**Schneider:** Convex Bodies: The Brunn-Minkowski Theory. Section 2.5.

**Schneider:** Convex Bodies: The Brunn-Minkowski Theory. Section 2.5.

**Application de Gauss:** Un convexe compact  $K \subset \mathbb{R}^p$  possédant une frontière  $C^1$ , en chaque  $x$  de  $\partial K$ , il existe un unique vecteur unitaire normal à  $K$ ,  $n(x) \in S_{p-1}$ .

$$n: \partial K \mapsto S_{p-1}$$

est l'application de Gauss. C'est un difféomorphisme si et seulement si,  $\partial K$  est  $C^2$  et partout de courbure non nulle. Cette classe est appelée  $C_+^2$ .

**Schneider:** Convex Bodies: The Brunn-Minkowski Theory. Section 2.5.

**Application de Gauss:** Un convexe compact  $K \subset \mathbb{R}^p$  possédant une frontière  $C^1$ , en chaque  $x$  de  $\partial K$ , il existe un unique vecteur unitaire normal à  $K$ ,  $n(x) \in S_{p-1}$ .

$$n: \partial K \mapsto S_{p-1}$$

est l'application de Gauss. C'est un difféomorphisme si et seulement si,  $\partial K$  est  $C^2$  et partout de courbure non nulle. Cette classe est appelée  $C_+^2$ .

**Paramétrisation par la normale:** Inverses de l'application de Gauss,  $A \in C_+^2$ ,

$$\begin{aligned}c_A: S_{p-1} &\mapsto \partial A \\ n &\mapsto \operatorname{argmax}_{x \in A} \langle n, x \rangle.\end{aligned}$$



**Schneider:** Convex Bodies: The Brunn-Minkowski Theory. Section 2.5.

**Application de Gauss:** Un convexe compact  $K \subset \mathbb{R}^p$  possédant une frontière  $C^1$ , en chaque  $x$  de  $\partial K$ , il existe un unique vecteur unitaire normal à  $K$ ,  $n(x) \in S_{p-1}$ .

$$n: \partial K \mapsto S_{p-1}$$

est l'application de Gauss. C'est un difféomorphisme si et seulement si,  $\partial K$  est  $C^2$  et partout de courbure non nulle. Cette classe est appelée  $C_+^2$ .

**Paramétrisation par la normale:** Inverses de l'application de Gauss,  $A \in C_+^2$ ,

$$\begin{aligned}c_A: S_{p-1} &\mapsto \partial A \\ n &\mapsto \operatorname{argmax}_{x \in A} \langle n, x \rangle.\end{aligned}$$

Soient  $A, B \in C_+^2$  de frontière  $C^k$ , alors  $A + B \in C_+^2$  et est de frontière  $C^k$ .

$$c_{A+B} = c_A + c_B$$

$T_0 \subset \text{int}(T_1)$ , compacts convexes dans  $C_+^2$ ,  $T_\lambda = (1 - \lambda)T_0 + \lambda T_1$ ,  $\lambda \in [0, 1]$ ,

$$n: \mathbb{R}/2\pi\mathbb{Z} \mapsto \mathbb{R}^2$$

$$\theta \mapsto \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

$T_0 \subset \text{int}(T_1)$ , compacts convexes dans  $C_+^2$ ,  $T_\lambda = (1 - \lambda)T_0 + \lambda T_1$ ,  $\lambda \in [0, 1]$ ,

$$n: \mathbb{R}/2\pi\mathbb{Z} \mapsto \mathbb{R}^2$$

$$\theta \mapsto \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

Pour  $\lambda \in [0, 1]$  et  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ :

$c_\lambda(\theta)$  est l'unique  $x \in \partial T_\lambda$  tel que le vecteur unitaire normal à  $T_\lambda$  en  $x$  soit  $n(\theta)$ .

$c_\lambda$  est  $C^1$  et

$$c_\lambda = (1 - \lambda)c_0 + \lambda c_1.$$

## Paramétrisation par les normales dans $\mathbb{R}^2$

$T_0 \subset \text{int}(T_1)$ , compacts convexes dans  $C_+^2$ ,  $T_\lambda = (1 - \lambda)T_0 + \lambda T_1$ ,  $\lambda \in [0, 1]$ ,

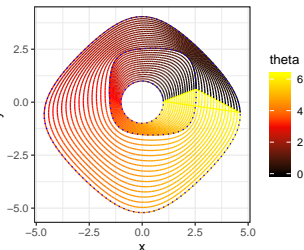
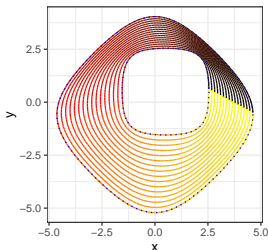
$$n: \mathbb{R}/2\pi\mathbb{Z} \mapsto \mathbb{R}^2$$

$$\theta \mapsto \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$$

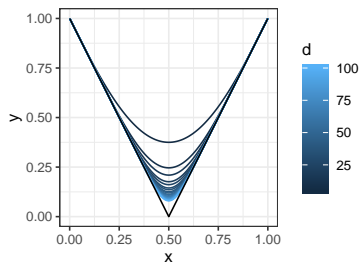
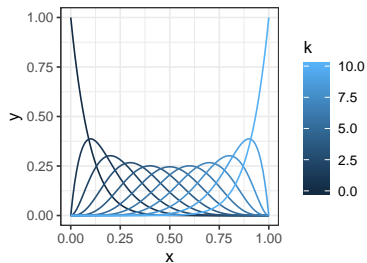
Pour  $\lambda \in [0, 1]$  et  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$ :

$c_\lambda(\theta)$  est l'unique  $x \in \partial T_\lambda$  tel que le vecteur unitaire normal à  $T_\lambda$  en  $x$  soit  $n(\theta)$ .  
 $c_\lambda$  est  $C^1$  et

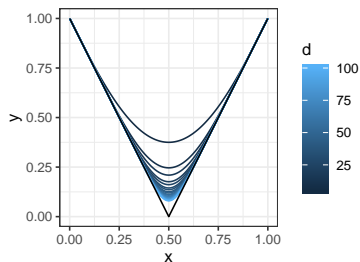
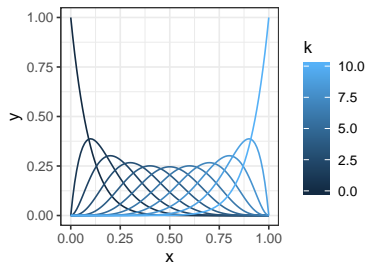
$$c_\lambda = (1 - \lambda)c_0 + \lambda c_1.$$



Polynomes de Bernstein et approximation de Bernstein sur une valeur absolue:



Polynomes de Bernstein et approximation de Bernstein sur une valeur absolue:



$f: [0, 1] \mapsto \mathbb{R}$ ,  $d \in \mathbb{N}^*$ ,  $m \in \mathbb{N}^*$ ,

$$B_{d,f}: x \mapsto \sum_{k=0}^d f\left(\frac{k}{d}\right) \binom{d}{k} x^k (1-x)^{d-k}$$

$f: [0, 1] \mapsto \mathbb{R}$ ,  $d \in \mathbb{N}^*$ ,  $m \in \mathbb{N}^*$ ,

$$B_{d,f}: x \mapsto \sum_{k=0}^d f\left(\frac{k}{d}\right) \binom{d}{k} x^k (1-x)^{d-k}$$

$f: [0, 1] \mapsto \mathbb{R}$ ,  $d \in \mathbb{N}^*$ ,  $m \in \mathbb{N}^*$ ,

$$B_{d,f}: x \mapsto \sum_{k=0}^d f\left(\frac{k}{d}\right) \binom{d}{k} x^k (1-x)^{d-k}$$

**Différences discrètes:**  $\Delta^1 f(x) = f(x + 1/d) - f(x)$ ,  $\Delta^k f(x) = \Delta^1(\Delta^{k-1} f(x))$ .

Dérivée d'ordre  $m$ :

$$B_{d,f}^{(m)}: x \mapsto d(d-1)\dots(d-m+1) \sum_{k=0}^{d-m} \Delta^m f\left(\frac{k}{d}\right) \binom{d-m}{k} x^k (1-x)^{d-k-m}$$



$f: [0, 1] \mapsto \mathbb{R}$ ,  $d \in \mathbb{N}^*$ ,  $m \in \mathbb{N}^*$ ,

$$B_{d,f}: x \mapsto \sum_{k=0}^d f\left(\frac{k}{d}\right) \binom{d}{k} x^k (1-x)^{d-k}$$

**Différences discrètes:**  $\Delta^1 f(x) = f(x + 1/d) - f(x)$ ,  $\Delta^k f(x) = \Delta^1(\Delta^{k-1} f(x))$ .

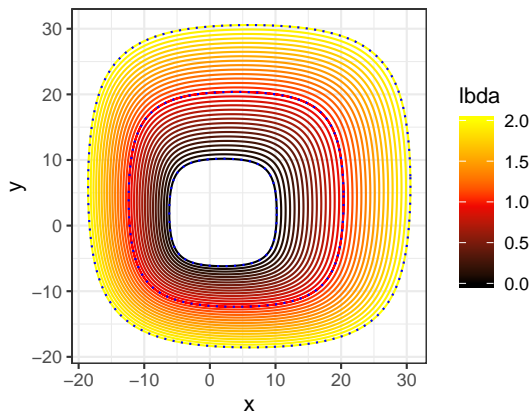
Dérivée d'ordre  $m$ :

$$B_{d,f}^{(m)}: x \mapsto d(d-1)\dots(d-m+1) \sum_{k=0}^{d-m} \Delta^m f\left(\frac{k}{d}\right) \binom{d-m}{k} x^k (1-x)^{d-k-m}$$

- Linéaire en  $f$ .
- Préserve la forme: si  $f$  est monotone, convexe ou concave, alors  $B_{d,f}$  l'est aussi.
- Si  $f$  est affine par morceaux, pour tout  $k \in \mathbb{N}$  et  $d$  suffisamment grand, les dérivées d'ordre  $2 \dots k$  s'annulent aux extrémités.

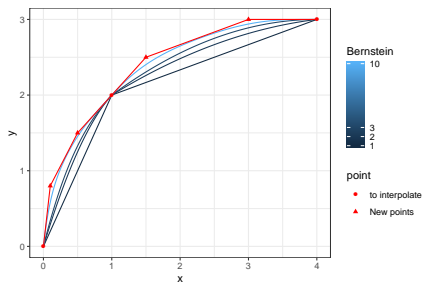
1. Tour d'horizon pour le problème d'interpolation
2. Rappels sur la courbure et l'approximation de Bernstein
3. Deux résultats d'interpolation dans  $\mathbb{R}^2$
4. Construction des contre exemples

Sous niveaux homothétiques.

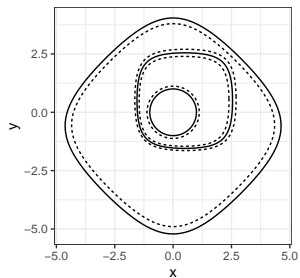
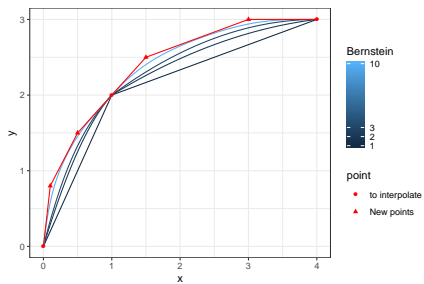


**Idée:** garantir un comportement similaire aux ordres  $0, 1, \dots, k$ .

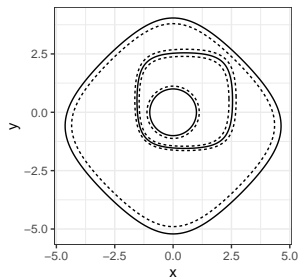
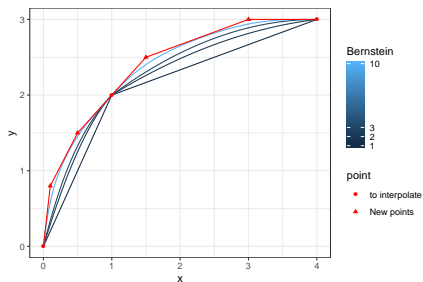
# Interpolation non linéaire, idée des courbes de Bézier



# Interpolation non linéaire, idée des courbes de Bézier



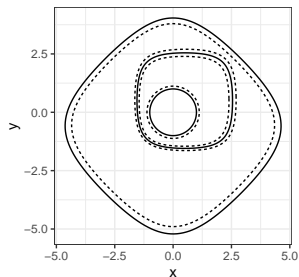
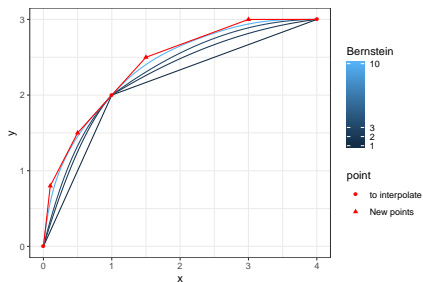
# Interpolation non linéaire, idée des courbes de Bézier



$$T_\lambda = a(\lambda)T_0 + b(\lambda)T_1$$

$a$  et  $b$ : approximation de Bernstein d'une fonction affine par morceaux.

# Interpolation non linéaire, idée des courbes de Bézier



$$T_\lambda = a(\lambda)T_0 + b(\lambda)T_1$$

$a$  et  $b$ : approximation de Bernstein d'une fonction affine par morceaux.

**Calcul explicite:** dérivées d'ordre 1 et 2, repère donné par  $n(\theta)$ .

**Recollement  $C^k$  au bord:**

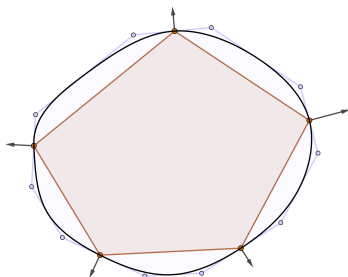
- Direction "normale": contrôle  $C^1$ , les variations d'ordre 2, ...,  $k$  s'annulent.
- Direction "tangente": contrôle  $C^k$ .

**Théorème:**

Soit  $(T_i)_{i \in \mathbb{Z}}$  une suite de sous ensembles de  $\mathbb{R}^2$ , convexes, compacts, à frontière  $C^k$  et à courbure positive, tels que  $T_i \subset \text{int } T_{i+1} \neq \emptyset$  pour tout  $i$  dans  $\mathbb{Z}$ . Il existe une fonction convexe  $C^k$ ,  $f$ , telle que chaque  $T_i$  est un sous-niveau de  $f$ . Sa hessienne est définie positive en dehors de:

$$\text{argmin } f = \bigcap_{i \in \mathbb{Z}} T_i.$$





## Lemma:

Soit  $P = A_1 \dots A_n$  un polygone convexe. Pour tout  $i$ , soit  $V_i$  un vecteur normal en  $A_i$  tel que l'angle entre  $V_i$  et chaque face voisine est dans  $(\frac{\pi}{2}, \pi)$ . Pour tout  $\epsilon > 0$ ,  $m \geq 2$ , il existe un compact convexe  $C \subset \mathbb{R}^2$  tel que  $P \subset C$ , et

- $\partial C$  est  $C^m$  à courbure positive,
- pour tout  $i = 1, \dots, n$ ,  $A_i \in \partial(C)$  et le cône normal à  $C$  en  $A_i$  est donné par  $V_i$ ,
- $\max_{y \in C} \text{dist}(y, P) \leq \epsilon$ .



## Lemma:

Pour tout  $r_-, r_+ > 0$ ,  $t_- > 0$ ,  $t_+ < 0$ ,  $\epsilon > 0$ ,  $m \in \mathbb{N}$ ,  $m \geq 3$ , il existe un polynôme strictement concave  $p: [0, 1] \mapsto [0, \epsilon]$  such that

$$p(0) = 0$$

$$p(1) = 0$$

$$p'(0) = t_-$$

$$p'(1) = t_+$$

$$p''(0) = -r_-$$

$$p''(1) = -r_+.$$

$$p^{(q)}(0) = 0 \quad q \in \{3, \dots, m\}.$$

Approximation de Bernstein, recollement  $C^m$  autour des faces de  $P$ .

1. Tour d'horizon pour le problème d'interpolation
2. Rappels sur la courbure et l'approximation de Bernstein
3. Deux résultats d'interpolation dans  $\mathbb{R}^2$
4. Construction des contre exemples

**Recherche linéaire exacte:**  $f$  convexe  $C^k$ .

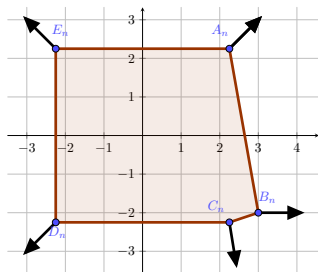
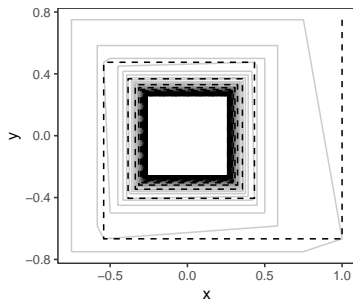
$$x_{k+1} \in \arg \min_x f(x), \quad t.q. \quad x = x_k - s \nabla f(x_k), \quad s \in \mathbb{R}_+.$$

Condition d'optimalité:  $\langle x_{k+1} - x_k, \nabla f(x_{k+1}) \rangle = 0$ .

**Recherche linéaire exacte:**  $f$  convexe  $C^k$ .

$$x_{k+1} \in \arg \min_x f(x), \quad t.q. \quad x = x_k - s \nabla f(x_k), \quad s \in \mathbb{R}_+.$$

Condition d'optimalité:  $\langle x_{k+1} - x_k, \nabla f(x_{k+1}) \rangle = 0$ .



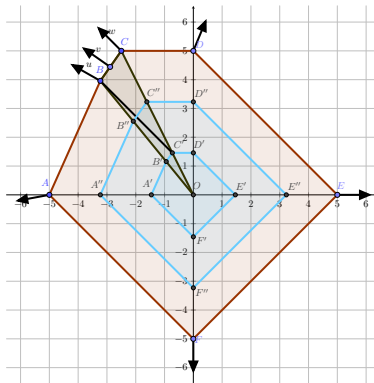
**Convergence des sécantes:**  $f$  convexe  $C^k$ .

$$\dot{x}(t) = -\nabla f(x(t)),$$
$$\frac{x(t) - x_\infty}{\|x(t) - x_\infty\|} \xrightarrow[t \rightarrow \infty]{} v \in \mathbb{R}^2?$$

**Convergence des sécantes:**  $f$  convexe  $C^k$ .

$$\dot{x}(t) = -\nabla f(x(t)),$$

$$\frac{x(t) - x_\infty}{\|x(t) - x_\infty\|} \xrightarrow{t \rightarrow \infty} v \in \mathbb{R}^2?$$



Il existe des fonctions  $f, h$  convexe  $C^k$  sur  $\mathbb{R}^2$  qui génèrent



Il existe des fonctions  $f, h$  convexe  $C^k$  sur  $\mathbb{R}^2$  qui génèrent

- une suite de Gauss-Seidel (minimisation alternée) qui ne converge pas.

$$a_{k+1} \in \arg \min_a f(a, b_k), \quad b_{k+1} \in \arg \min_b f(a_{k+1}, b).$$

- des sous-niveaux qui ne satisfont pas l'inégalité de Kurdyka-Łojasiewicz (Bolte, Daniilidis, Ley, Mazet 2009).
- un chemin de régularisation de Tikhonov de longueur infinie (Torralba 1996).

Il existe des fonctions  $f, h$  convexe  $C^k$  sur  $\mathbb{R}^2$  qui génèrent

- une suite de Gauss-Seidel (minimisation alternée) qui ne converge pas.

$$a_{k+1} \in \arg \min_a f(a, b_k), \quad b_{k+1} \in \arg \min_b f(a_{k+1}, b).$$

- des sous-niveaux qui ne satisfont pas l'inégalité de Kurdyka-Łojasiewicz (Bolte, Daniilidis, Ley, Mazet 2009).
- un chemin de régularisation de Tikhonov de longueur infinie (Torralba 1996).
- une trajectoire de Newton, qui ne converge pas

$$\dot{x} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

Il existe des fonctions  $f, h$  convexe  $C^k$  sur  $\mathbb{R}^2$  qui génèrent

- une suite de Gauss-Seidel (minimisation alternée) qui ne converge pas.

$$a_{k+1} \in \arg \min_a f(a, b_k), \quad b_{k+1} \in \arg \min_b f(a_{k+1}, b).$$

- des sous-niveaux qui ne satisfont pas l'inégalité de Kurdyka-Łojasiewicz (Bolte, Daniilidis, Ley, Mazet 2009).
- un chemin de régularisation de Tikhonov de longueur infinie (Torralba 1996).
- une trajectoire de Newton, qui ne converge pas

$$\dot{x} = -\nabla^2 f(x)^{-1} \nabla f(x)$$

- un chemin central qui ne converge pas

$$x(r) = \operatorname{argmin}\{\langle c, y \rangle + rh(y), y \in [-1, 1]^2\}$$

- une séquence de descente de Bregman (mirror descent) qui ne converge pas.

$$x_{k+1} = \nabla h^*(\nabla h(x_k) - c).$$

- une métrique Hessienne-Riemannienne  $\nabla h^2$  et une trajectoire de gradient qui ne converge pas.

$h$  définie sur  $C$  convexe telle que  $h$  est  $C^1$  sur  $\text{int}(C)$ ,  $\nabla h$  induit une bijection et tends vers l'infini à la frontière de  $C$ .

$h$  définie sur  $C$  convexe telle que  $h$  est  $C^1$  sur  $\text{int}(C)$ ,  $\nabla h$  induit une bijection et tends vers l'infini à la frontière de  $C$ .

Pour tout  $\theta \in \left(-\frac{\pi}{4}, \frac{\pi}{4}\right)$ , il existe une fonction de Legendre  $h: \mathbb{R}^2 \mapsto \mathbb{R}$  dont le domaine est un carré fermé, continue sur son domaine et  $C^k$  sur son intérieur telle que pour tout  $i \in \mathbb{N}$ ,  $\nabla h^*(i, 0)$  est proportionnel à  $(\cos(\theta), (-1)^i \sin(\theta))$ .

$h$  définie sur  $C$  convexe telle que  $h$  est  $C^1$  sur  $\text{int}(C)$ ,  $\nabla h$  induit une bijection et tends vers l'infini à la frontière de  $C$ .

Pour tout  $\theta \in \left(-\frac{\pi}{4}, \frac{\pi}{4}\right)$ , il existe une fonction de Legendre  $h: \mathbb{R}^2 \mapsto \mathbb{R}$  dont le domaine est un carré fermé, continue sur son domaine et  $C^k$  sur son intérieur telle que pour tout  $i \in \mathbb{N}$ ,  $\nabla h^*(i, 0)$  est proportionnel à  $(\cos(\theta), (-1)^i \sin(\theta))$ .

Construction de  $h^*$  en interpolant des carrés homothétiques.

## Résultat principal dans $\mathbb{R}^2$ :

- Interpolation de sous-niveaux convexes à courbure positive.
- Spécification de sous-niveaux convexes à courbure positive.

## Ce que vous n'avez pas vu:

- De nombreux contre-exemples.
- Discussion fine: hessienne, Lipschitz globalement, propriété de Legendre.
- Subtilités calculatoires de la construction.

## Ouverture:

- Plus de contre-exemples.
- Dimension finie arbitraire.
- Interpolation  $C^\infty$ .
- Relachement de l'hypothèse de courbure.









Alvarez, F., Bolte, J. and Brahic, O. (2004).

Hessian Riemannian gradient flows in convex programming.  
SIAM journal on control and optimization, 43(2), 477–501.



Alvarez D. F. and Pérez C., J. M. (1998).

A dynamical system associated with Newton's method for parametric approximations of convex minimization problems.  
Applied Mathematics and Optimization, 38, 193–217.



Auslender, A., Optimisation Méthodes Numériques, Masson, Paris, France, 1976.



Auslender, A. (1999). Penalty and barrier methods: a unified framework. SIAM Journal on Optimization, 10(1), 211–230.



Aubin, J.-P. and Cellina, A. (1984).

Differential inclusions: set-valued maps and viability theory. Springer.

















Bauschke, H. H., Bolte, J. and Teboulle, M. (2016).

A descent lemma beyond Lipschitz gradient continuity: first-order methods revisited and applications.  
Mathematics of Operations Research, 42(2), 330–348.



Bauschke, H. H., and Combettes, P. L. (2011). Convex analysis and monotone operator theory in Hilbert spaces (Vol. 408). New York: Springer.

-  Beck, A. (2017). First-order methods in optimization (Vol. 25). SIAM.
-  Beck, A. and Teboulle, M. (2003).  
Mirror descent and nonlinear projected subgradient methods for convex optimization.  
Operations Research Letters, 31(3), 167–175.
-  Beck, A. and Tetrushvili, L. (2013).  
On the convergence of block coordinate descent type methods.  
SIAM journal on Optimization, 23(4), 2037-2060.
-  Bertsekas, D. P., Scientific, A. (2015). Convex optimization algorithms. Belmont: Athena Scientific.
-  Bolte, J., Daniilidis, A., Ley, O., and Mazet, L. (2010).  
Characterizations of Łojasiewicz inequalities: subgradient flows, talweg, convexity.  
Transactions of the American Mathematical Society, 362(6).
-  Bolte, J., Nguyen, T. P., Peypouquet, J., and Suter, B. W. (2017). From error bounds to the complexity of first-order descent methods for convex functions.  
Mathematical Programming, 165(2), 471–507.
-  Bolte, J., and Teboulle, M. (2003). Barrier operators and associated gradient-like dynamical systems for constrained minimization problems. SIAM journal on control and optimization, 42(4), 1266-1292.

-  Borwein, J. M., Li, G., and Yao, L. (2014). Analysis of the convergence rate for the cyclic projection algorithm applied to basic semialgebraic convex sets. *SIAM Journal on Optimization*, 24(1), 498–527.
-  Boyd, S., and Vandenberghe, L. (2004). *Convex optimization*. Cambridge university press.
-  Chen, C., He, B., Ye, Y., and Yuan, X. (2016). The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. *Mathematical Programming*, 155(1-2), 57–79.
-  Crouzeix, J.-P. (1980). Conditions for convexity of quasiconvex functions. *Mathematics of Operations Research*, 5(1), 120-125.
-  Daniilidis, A., Ley, O., and Sabourau, S. (2010). Asymptotic behaviour of self-contracted planar curves and gradient orbits of convex functions. *Journal de mathématiques pures et appliquées*, 94(2), 183–199.
-  Dragomir, R. A., Taylor, A., d'Aspremont, A., and Bolte, J. (2019). Optimal complexity and certification of Bregman first-order methods. *arXiv preprint arXiv:1911.08510*.
-  Fenchel, W., *Convex Cones, Sets and Functions*, Mimeographed lecture note, Princeton University, 1951.



de Finetti, B. (1949).

Sulle stratificazioni convesse.

Annali di Matematica Pura ed Applicata, 30(1), 173–183.



Gale, D., Klee, V. and Rockafellar, R. T. (1968).

Convex functions on convex polytopes.

Proceedings of the American Mathematical Society, 19(4), 867–873.



Golub, G. H., Hansen, P. C., and O'Leary, D. P. (1999). Tikhonov regularization and total least squares. SIAM Journal on Matrix Analysis and Applications, 21(1), 185–194.



Kannai, Y. (1977). Concavifiability and constructions of concave utility functions. Journal of mathematical Economics, 4(1), 1–56.



Kurdyka, K., Mostowski, T. and Parusinski, A. (2000).

Proof of the gradient conjecture of R. Thom.








Annals of Mathematics, 152(3), 763–792.



Łojasiewicz, S. (1984),

Sur les trajectoires du gradient d'une fonction analytique.

Seminari di Geometria, Bologna (1982/83), Universita' degli Studi di Bologna, Bologna (1984), 115–117.

- 
- Lorentz, G. G. (1954).  
Bernstein polynomials.  
American Mathematical Soc.
- 
- Ma, T. W. (2009). Higher chain formula proved by combinatorics. The electronic journal of combinatorics, 16(1).
- 
- Manselli, P., and Pucci, C. (1991). Maximum length of steepest descent curves for quasi-convex functions. Geometriae Dedicata, 38(2), 211–227.
- 
- A. S. Nemirovsky and D. B. Yudin (1983).  
Problem Complexity and Method Efficiency in Optimization.  
Wiley-Interscience, New York.
- 
- Nesterov, Y. (2003). Lectures on convex optimization (Vol. 137). Springer.
- 
- Nesterov, Y., and Nemirovskii, A. (1994). Interior-point polynomial algorithms in convex programming (Vol. 13). Siam.
- 
- Powell, M. J. (1973).  
On search directions for minimization algorithms.  
Mathematical programming, 4(1), 193–201.



Schneider, R. (1993).

Convex bodies: the Brunn-Minkowski theory (No. 151).  
Cambridge university press.



Torralba, D., Convergence épigraphique et changements d'échelle en analyse variationnelle et optimisation, Ph.D. Thesis (Université Montpellier 2, 1996).



Rockafellar, R. T. (1970).

Convex analysis (Vol. 28). Princeton university press.



Thom, R. (1989).

Problèmes rencontrés dans mon parcours mathématique : un bilan.  
Publications mathématiques de l'IHES, 70, 199–214.



Wright, S. J. (2015).

Coordinate descent algorithms.  
Mathematical Programming, 151(1), 3–34.