

Comité de suivi : présentation de milieu de thèse

Baptiste Plaquevent-Jourdain

Encadrement : Jean-Pierre Dussault, Université de Sherbrooke
et Jean Charles Gilbert, INRIA Paris

30 Août 2023

Plan

1 Présentation

2 Contenu de la thèse

- Complémentarité et optimisation
 - Complémentarité et hyperplans

3 Bibliography

Parcours

- Classes préparatoires, filière MP (1er cycle sur 2 ans) : cours intensif en mathématiques, physique, algorithmique
- ENSTA Paris (2nd cycle sur 3 ans) : formation générale avec spécialisations mathématiques puis optimisation
- Master optimisation en parallèle
- Doctorat, depuis janvier 2022

- M1 - début maîtrise : EDP non-linéaires simples
- césure : modélisation physique en entreprise
- césure : opt. multiobjective, BFGS sur l'hypervolume
- M2 - avec les encadrants du doctorat

Parcours

- Classes préparatoires, filière MP (1er cycle sur 2 ans) : cours intensif en mathématiques, physique, algorithmique
- ENSTA Paris (2nd cycle sur 3 ans) : formation générale avec spécialisations mathématiques puis optimisation
- Master optimisation en parallèle
- Doctorat, depuis janvier 2022

- M1 - début maîtrise : EDP non-linéaires simples
- césure : modélisation physique en entreprise
- césure : opt. multiobjective, BFGS sur l'hypervolume
- M2 - avec les encadrants du doctorat

Organisation générale

- été 2021 : stage de fin d'études
- automne-hiver 2021 : 3 mois de CDD
- janvier 2022 début thèse

Avril-Septembre : Sherbrooke (Québec), sinon Inria Paris

Visios \simeq 1/semaine, échanges par courriels

JPD est venu à Paris en mars-avril 2023.

Bourse de mobilité MITACS pour le séjour canadien à l'été 2023.

Organisation générale

- été 2021 : stage de fin d'études
- automne-hiver 2021 : 3 mois de CDD
- janvier 2022 début thèse

Avril-Septembre : Sherbrooke (Québec), sinon Inria Paris

Visios \simeq 1/semaine, échanges par courriels

JPD est venu à Paris en mars-avril 2023.

Bourse de mobilité MITACS pour le séjour canadien à l'été 2023.

Enseignements

- AO101 (ENSTA Paris) : introduction à l'optimisation niveau L3 (fin bachelor), donné en 2022, 2023, a priori 2024 (bonne connaissance des détails)
- OPT201 (ENSTA Paris), 2022 (, 2024) : optimisation pour la spécialisation mathématiques (*)
- OPT202 (ENSTA Paris), 2022-2023 (, 2024) : optimisation avancée (*)
- ROP631 (UdeS), 2023 : optimisation avancée

(*) aide à l'élaboration des cours (nouveaux enseignants)

Plan

1 Présentation

2 Contenu de la thèse

- Complémentarité et optimisation
- Complémentarité et hyperplans

3 Bibliography

Introduction : problèmes de complémentarité

Pour $F, G : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (régulières), trouver un x vérifiant

$$0 \leq F(x) \perp G(x) \geq 0 \\ \Leftrightarrow \forall i \in [1 : n], F_i(x) \geq 0, G_i(x) \geq 0, F_i(x)G_i(x) = 0 \quad (1)$$

Important pour nous :

$$F(x) \equiv Ax + a, \quad G(x) \equiv Bx + b, \quad A, B \in \mathbb{R}^{n \times n}, \quad a, b \in \mathbb{R}^n$$

$$0 \leq (Ax + a)^\top (Bx + b) \leq 0 \quad (2)$$

Ex : généralisation des conditions d'optimalité, chimie, contact...

L'astuce du minimum

$$[n=1] \quad \begin{aligned} 0 \leq f(x) \perp g(x) \geq 0 &\Leftrightarrow \min(f(x), g(x)) = 0 \\ 0 \leq F(x) \perp G(x) \geq 0 &\Leftrightarrow \min(F(x), G(x)) = 0 \end{aligned} \quad (3)$$

Il existe d'autres fonctions vérifiant (3), moins 'simples'

L'astuce du minimum

$$\begin{aligned} [n=1] \quad 0 \leq f(x) \perp g(x) \geq 0 &\Leftrightarrow \min(f(x), g(x)) = 0 \\ 0 \leq F(x) \perp G(x) \geq 0 &\Leftrightarrow \min(F(x), G(x)) = 0 \end{aligned} \quad (3)$$

Il existe d'autres fonctions vérifiant (3), moins 'simples'

Les problèmes de complémentarité (1),(2) deviennent :

$$\min(F(x), G(x)) = 0, \quad \min(Ax + a, Bx + b) = 0 \quad (4)$$

→ systèmes d'équations non-linéaires *non-lisses*.

Méthode de résolution : Newton 'semi-lisse'

Fin maîtrise [été 2021, inachevé]

Pour $H(x) = \min(F(x), G(x)) = 0$, minimum global de $\theta(x) = \|H(x)\|^2/2$?

Pas différentiable : adapter la notion de direction de descente.

Difficultés sur les indices i tels que $F_i(x_k) = G_i(x_k)$

Objectif : x stationnaire $\Leftrightarrow ??$

Idée : combinaisons convexes entre F_i et G_i . Un choix judicieux des poids permet d'obtenir une direction de descente à partir d'un "gradient".

Fin maîtrise [été 2021, inachevé]

Pour $H(x) = \min(F(x), G(x)) = 0$, minimum global de $\theta(x) = \|H(x)\|^2/2$?

Pas différentiable : adapter la notion de direction de descente.

Difficultés sur les indices i tels que $F_i(x_k) = G_i(x_k)$

Objectif : x stationnaire \Leftrightarrow ??

Idée : combinaisons convexes entre F_i et G_i . Un choix judicieux des poids permet d'obtenir une direction de descente à partir d'un "gradient".

suite

Ce choix requiert un calcul technique (équation non-linéaire non-lisse simple).

caractérisation : revient à vérifier si un zonotope est inclus dans un autre.

(Zonotope ($= (p+)M \times [0, 1]^m$)).)

Utilise des propriétés techniques (projection cartésienne de polyèdres, inclusion de zonotopes...) ; inachevé

Illustration

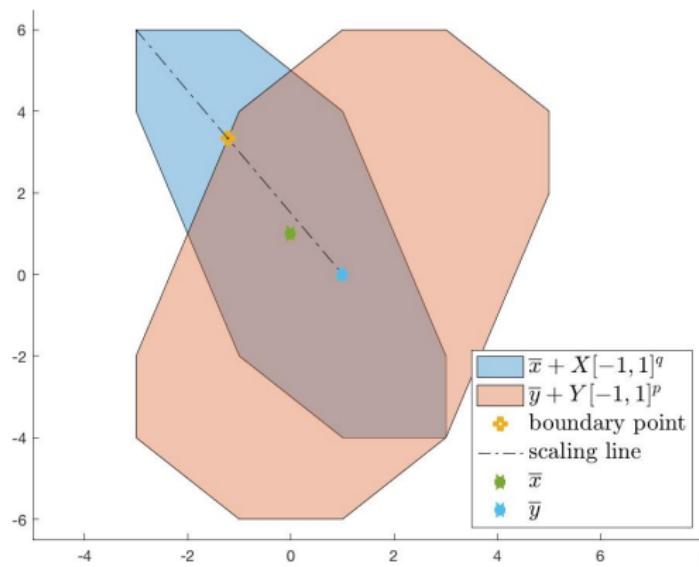


Illustration de deux zonotopes non-inclus dans l'autre.

B-différentiel

Éléments théoriques pour résoudre $H(x) = 0$:

Newton : convergence quadratique locale si $\nabla H(x^*)$ inversible.

Newton semi-lisse : il faut $\forall J \in \partial_B H(x^*), J$ inversible, avec

$$\partial_B H(x) = \{J : \exists (d_k)_k \rightarrow 0, H'(x + d_k) \rightarrow J\} \quad (5)$$

les limites en des points $x_k = x + d_k$ où H est différentiable.

Le B-différentiel sert aussi pour le C(larke)-différentiel

$$\partial_C H(x) = \text{conv}(\partial_B H(x))$$

Peut-on obtenir ce B-différentiel ?

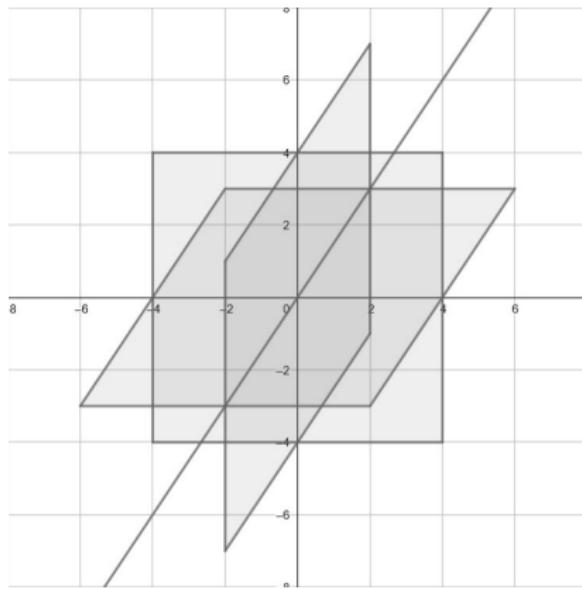
Une recherche surprise

[Les détails sont en annexe]

- le minimum est différentiable sauf si pour une composante i , les deux termes du minimum sont égaux
- $H_i = \min(A_{i,:}(x + d_k) + a_i, B_{i,:}(x + d_k) + b_i)$, égalité dans le min ?
- deux fonctions affines sont égales sur un hyperplan
- tout l'espace sauf des hyperplans :
$$H_i = \{d : (B_{:,i} - A_{:,i})d = 0\} = v_i^\perp$$
 avec $v_i = B_{:,i} - A_{:,i}$
- \Rightarrow il faut déterminer "l'espace privé d'hyperplans"

Complémentarité et hyperplans

$$\mathbb{R}^n = H_i^- \cup H_i \cup H_i^+, \quad H_i^- = \{x \in \mathbb{R}^n : v_i^T x < 0\}$$
$$H_i^+ = \{x \in \mathbb{R}^n : v_i^T x > 0\}$$



Trois hyperplans simples qui séparent l'espace.

Les arrangements d'hyperplans - présentation

On cherche les d tels que $\forall i \in [1 : p], v_i^T d \neq 0$ (donc > 0 ou < 0)

Formalisme abstrait :

- existe-t-il $d : v_i^T d > 0, \forall i$?
- existe-t-il $d : v_i^T d > 0, i \in [1 : p - 1], v_p^T d < 0$?
- ...

Formulation : $v_i^T d > 0$ ou $v_i^T d < 0 \rightarrow s_i v_i^T d > 0$, pour $s_i = -1$ et $s_i = +1$

Nombre combinatoire, 2^p systèmes à vérifier.

Ce sont les "chambres" d'un arrangement d'hyperplans $H_i = v_i^\perp$.
[détail]

Problème assez abstrait, d'autres approches possibles : graphes...

Les arrangements d'hyperplans - état de l'art

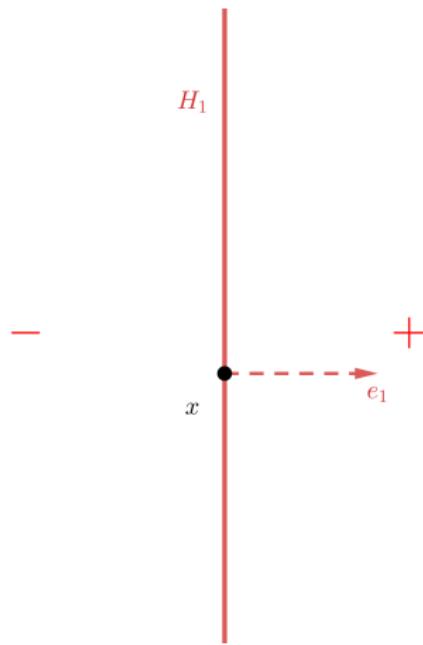
Algèbre-combinatoire : approches / outils / langages très différents
Assez difficile de comprendre (pour nous) ce qui a déjà été fait

Concept algorithmique : rajouter les hyperplans un à un

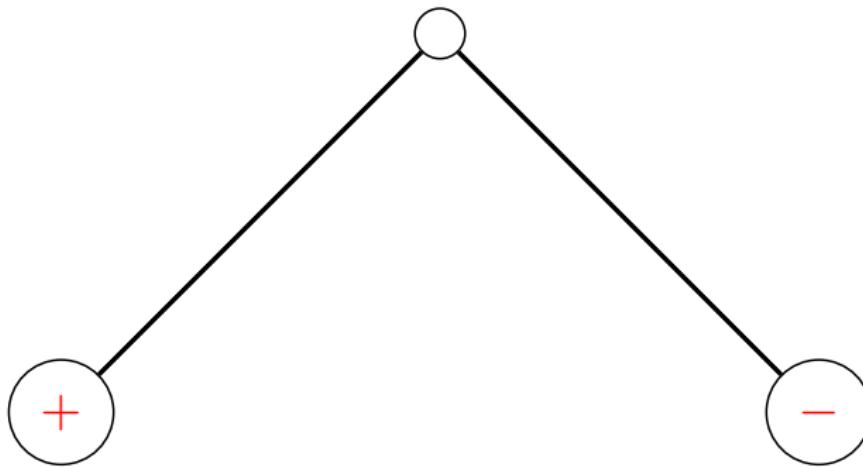
Si $s = (s_1, \dots, s_k)$ a une direction d vérifiant le système, est-ce que $(s^k, +1)$ et $(s^k, -1)$ en ont une aussi ?

Structure d'arbre, chaque niveau (= nouvel hyperplan rajouté) les nœuds ont 1 ou 2 descendants

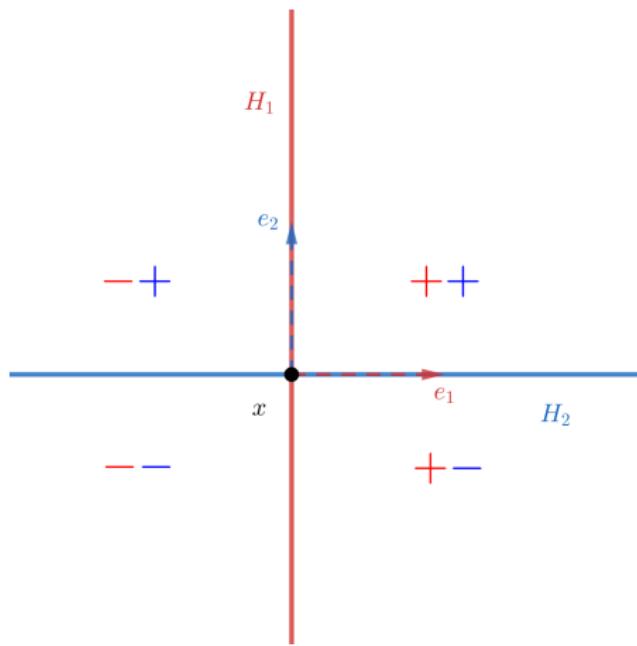
Example avec les chambres et l'arbre



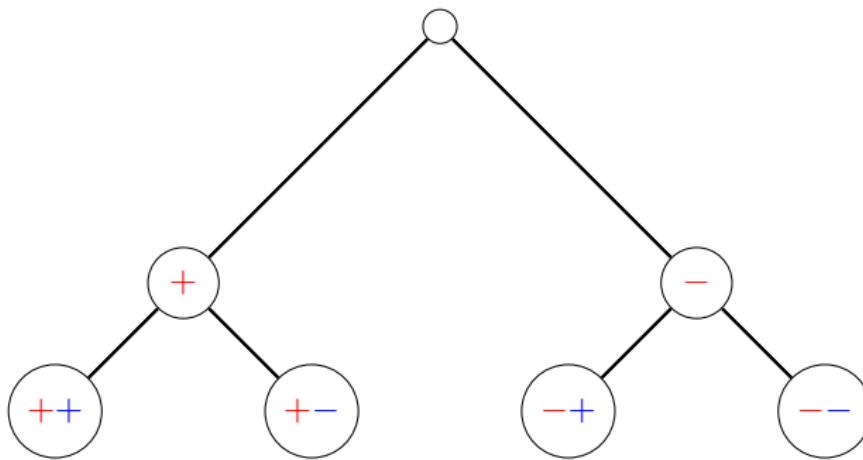
Example avec les chambres et l'arbre



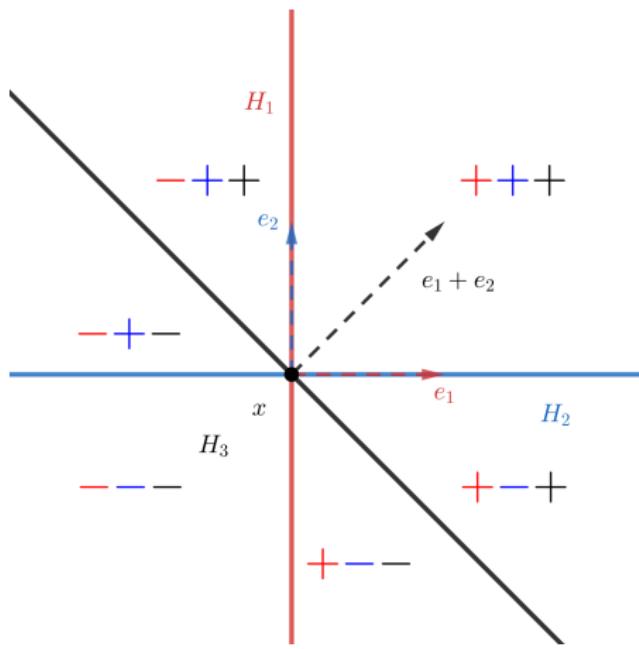
Example avec les chambres et l'arbre



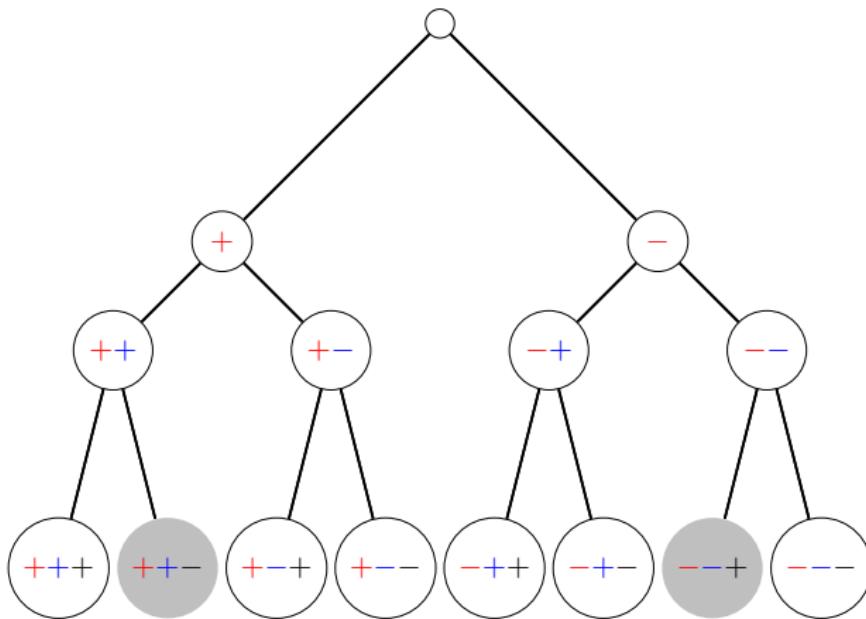
Example avec les chambres et l'arbre



Example avec les chambres et l'arbre



Example avec les chambres et l'arbre



Les arrangements d'hyperplans - contribution

améliorations structurelles

- structure générale de l'arbre : résoudre moins de POL
- ordre heuristique sur "quel hyperplan rajouter à l'étape $k + 1$ "

approche duale et "matroïdes"

Matroïdes : structure générale, qui, ici, regarde les "dépendances linéaires" entre les vecteurs v_i (cf image).

Calculer cette structure pour savoir quand élaguer une branche de l'arbre : un $(s^k, \pm 1)$ est assuré de ne pas avoir de solution.

Repose sur la dualité (Gordan), lien fort avec l'algèbre linéaire

Améliorations très efficaces ! D'autres raffinements (plus laborieux) sont possibles.

Les arrangements d'hyperplans - l'étape suivante

Les hyperplans n'ont pas nécessairement un point commun.

Cas centré [B-différentiel...] / cas général.

Initialement : l'algorithme est pour le cas général, tant mieux !

La question

Peut-on adapter les modifications au cas quelconque ? Si oui, sont-elles aussi efficaces / simples ? Si non, que peut-on faire ?

Les adapter oui, efficacité à voir.

Mais il existe des liens très forts entre le cas quelconque et le cas centré, en changeant légèrement les données adéquatement : une autre option que juste adapter les modifications.

Conclusion

Tout s'axe autour de $\min(F(x), G(x)) = 0$. Propriété structurelle du min d'avoir "soit l'un, soit l'autre".

- B-différentiel : F'_i ou G'_i
- hyperplans : du côté + ou du côté - de l'hyperplan
- zonotopes : aspect dual des hyperplans

Beaucoup de liens entre les différents thèmes (ça part de la même chose), des recherches plus prenantes que prévu (puis après lien avec les hyperplans donc des ouvertures).

Publications

Article envoyé

[DGP23c] - B-différentiel et hyperplans centrés

En préparation - projets

- [DGP23d] - cas général, package Julia
- peut-être quelque chose de court sur d'autres propriétés reliées

Présentations orales : B-différentiel et hyperplans, avancements successifs

- JOPT Montréal, 2022[DGP22b]
- Séminaire d'un groupe du CNRS, [DGP22a]
- JOPT Montréal, 2023[DGP23b]
- Colloque ISM, 2023[DGP23a]

Plan

1 Présentation

2 Contenu de la thèse

- Complémentarité et optimisation
- Complémentarité et hyperplans

3 Bibliography

Bibliographic elements I

- [DGP22a] J.-P. Dussault, J.Ch. Gilbert, and B. Plaquevent-Jourdain. “On the B-differential of the componentwise minimum of two vectorial functions”. In: url. 2022.
- [DGP22b] J.-P. Dussault, J.Ch. Gilbert, and B. Plaquevent-Jourdain. “On the computation of the B-differential of the Min C-function for the balanced linear complementarity problem”. In: url. 2022.
- [DGP23a] J.-P. Dussault, J.Ch. Gilbert, and B. Plaquevent-Jourdain. “Arrangements d’hyperplans : apports de la dualité et de l’optimisation”. In: 2023.

Bibliographic elements II

- [DGP23b] J.-P. Dussault, J.Ch. Gilbert, and B. Plaquevent-Jourdain. “Hyperplane Arrangements and Matroids in Complementarity Problems: the B-differential of the componentwise Minimum”. In: 2023.
- [DGP23c] Jean-Pierre Dussault, J.Ch. Gilbert, and B. Plaquevent-Jourdain. “On the B-differential of the componentwise minimum of two affine vector functions”. In: *Mathematical Programming Computation* (2023).
- [DGP23d] Jean-Pierre Dussault, J.Ch. Gilbert, and B. Plaquevent-Jourdain. *Arrangement of hyperplanes in the space by a dual approach*. 2023.

Bibliographic elements III

- [DGP23e] Jean-Pierre Dussault, J.Ch. Gilbert, and
B. Plaquevent-Jourdain. *Partial description of the
B-differential of the componentwise minimum of two
vector functions by linearization.* 2023.

Détails sur l'obtention des hyperplans

- $(Ax + a)_i < (Bx + b)_i$, vrai en $x + d_k \rightarrow x : H'_i(x + d_k) = A_{:,i}$
- $(Ax + a)_i > (Bx + b)_i$, vrai en $x + d_k \rightarrow x : H'_i(x + d_k) = B_{:,i}$
- $(Ax + a)_i = (Bx + b)_i$: les $\{i \in [1 : n] : (Ax + a)_i = (Bx + b)_i\} = \mathcal{E}(x)$ posent problème

$(Ax + a + Ad_k)_i = (Bx + b + Bd_k)_i$ se simplifie : il reste

$$(Ad_k)_i = A_{:,i}d_k = B_{:,i}d_k = (Bd_k)_i.$$

Si $i \in \mathcal{E}(x)$ et $v_i^T = B_{:,i} - A_{:,i} = 0$, les fonctions affines i sont identiques : le minimum ne fait rien (et donc c'est différentiable).
 $\rightarrow i \in \mathcal{E}^{\neq}(x) = \{i \in \mathcal{E}(x) : v_i \neq 0\}$, donc les hyperplans sont bien définis.

Techniquement on regarde un voisinage de x , mais ensuite on ne considère que ça donc on regarde tout l'espace. [retour]

Détails sur les vecteurs de signes

Il y a 2^P possibilités à vérifier, selon le signe de $v_i^T d$.

Plutôt que d'avoir des $< 0, > 0$, pour conventionner on prend uniquement > 0 :

$$v_i^T d < 0 \Leftrightarrow -v_i^T d > 0 \Leftrightarrow -(v_i^T d) > 0$$

d'où la présence de signes s_i . En optimisation on veut toutes les 'contraintes' avec le même signe.

Donc au lieu de systèmes avec $> 0, < 0$, on rajoute simplement des signes et c'est sur eux que l'on travaille. [retour]