

Polytechnique Montréal
Gerad

Analyse de sensibilité pour la réduction de dimension en optimisation sans dérivées

Montréal, 16 juillet 2020

Romain Vanden Bulcke

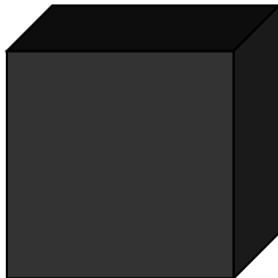
Directeurs : Charles Audet
Sébastien Le Digabel
Superviseur : Miguel Diago Martinez

Jury :
Dominique Orban, président
Stéphane Alarie, membre

Contexte

Contexte du projet

Optimisation sans dérivées et optimisation de **boîtes noires** en grande dimension



Contexte

Contexte du projet

Optimisation sans dérivées et optimisation de **boîtes noires** en grande dimension

Optimisation

- ▶ Recherche d'extrema d'une fonction

Contexte

Contexte du projet

Optimisation sans dérivées et optimisation de **boîtes noires** en grande dimension

Optimisation

- ▶ Recherche d'extrema d'une fonction

Boîtes noires

- ▶ Objet mathématique retournant des valeurs de sortie lorsqu'il reçoit des valeurs en entrée

Contexte

Contexte du projet

Optimisation sans dérivées et optimisation de **boîtes noires** en grande dimension

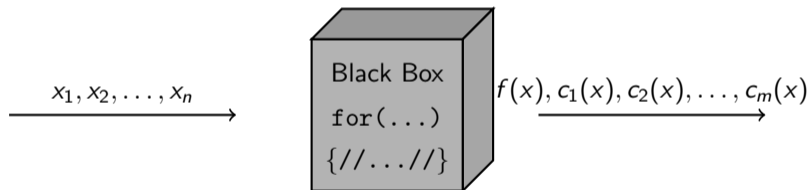
Optimisation

- ▶ Recherche d'extrema d'une fonction

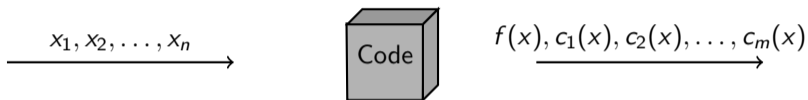
Boîtes noires

- ▶ Objet mathématique retournant des valeurs de sortie lorsqu'il reçoit des valeurs en entrée
- ▶ Généralement sous forme d'une simulation informatique

Caractéristiques des boîtes noires



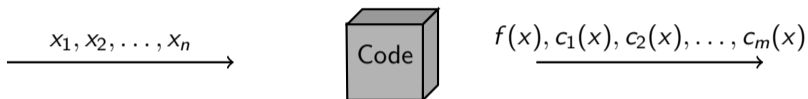
Caractéristiques des boîtes noires



Boîtes noires

- ▶ Dérivées non disponibles

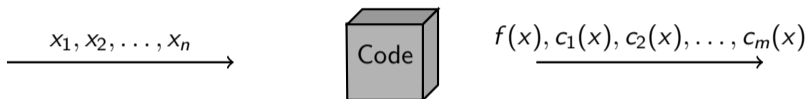
Caractéristiques des boîtes noires



Boîtes noires

- ▶ Dérivées non disponibles
- ▶ Long temps d'exécution

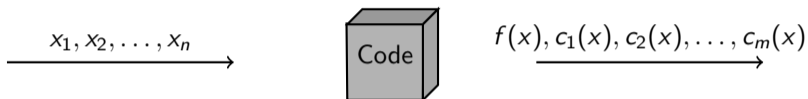
Caractéristiques des boîtes noires



Boîtes noires

- ▶ Dérivées non disponibles
- ▶ Long temps d'exécution
- ▶ Non lisses

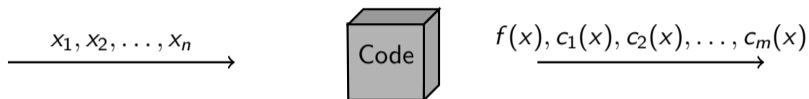
Caractéristiques des boîtes noires



Boîtes noires

- ▶ Dérivées non disponibles
- ▶ Long temps d'exécution
- ▶ Non lisses
- ▶ Parfois bruitées

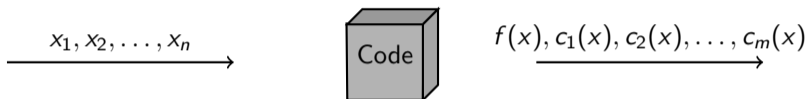
Caractéristiques des boîtes noires



Boîtes noires

- ▷ Dérivées non disponibles
- ▷ Long temps d'exécution
- ▷ Non lisses
- ▷ Parfois bruitées
- ▷ Possibilité d'erreurs

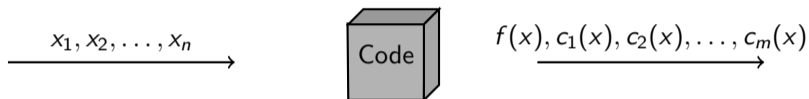
Caractéristiques des boîtes noires



Boîtes noires

- ▶ **Dérivées non disponibles**
- ▶ **Long temps d'exécution**
- ▶ Non lisses
- ▶ Parfois bruitées
- ▶ Possibilité d'erreurs

Caractéristiques des boîtes noires



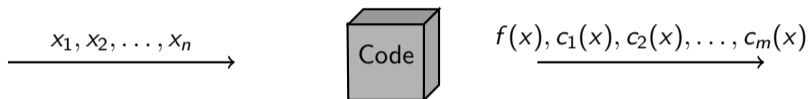
Boîtes noires

- ▶ **Dérivées non disponibles**
- ▶ **Long temps d'exécution**
- ▶ Non lisses
- ▶ Parfois bruitées
- ▶ Possibilité d'erreurs

Contexte du projet

- ▶ Boîtes noires de grande dimension
($n \geq 100$)

Caractéristiques des boîtes noires



Boîtes noires

- ▶ **Dérivées non disponibles**
- ▶ **Long temps d'exécution**
- ▶ Non lisses
- ▶ Parfois bruitées
- ▶ Possibilité d'erreurs

Contexte du projet

- ▶ Boîtes noires de grande dimension ($n \geq 100$)
- ▶ Séquentiel (pas de parallélisme)

Méthodes d'optimisation sans dérivées

- ▶ Heuristiques
- ▶ Méthodes basées sur des modèles
- ▶ Méthodes de recherche directe

Méthodes d'optimisation sans dérivées

- ▷ Heuristiques
 - Algorithme génétique
 - Méthode de Nelder-Mead [6]
 - ...
- ▷ Méthodes basées sur des modèles

- ▷ Méthodes de recherche directe

Méthodes d'optimisation sans dérivées

- ▷ Heuristiques
 - Algorithme génétique
 - Méthode de Nelder-Mead [6]
 - ...
- ▷ Méthodes basées sur des modèles
 - Descente basée sur des modèles
 - Méthodes de région de confiance [7]
 - Optimisation Bayésienne [8]
- ▷ Méthodes de recherche directe

Méthodes d'optimisation sans dérivées

▷ Heuristiques

- Algorithme génétique
- Méthode de Nelder-Mead [6]
- ...

▷ Méthodes basées sur des modèles

- Descente basée sur des modèles
- Méthodes de région de confiance [7]
- Optimisation Bayésienne [8]

▷ Méthodes de recherche directe

- Recherche par coordonnées [3]
- Recherche par motif généralisée [9]
- Recherche directe sur treillis adaptatif (MADS) [1]

Méthodes d'optimisation sans dérivées

- ▷ Heuristiques
 - Algorithme génétique
 - Méthode de Nelder-Mead [6]
 - ...
- ▷ Méthodes basées sur des modèles
 - Descente basée sur des modèles
 - Méthodes de région de confiance [7]
 - Optimisation Bayésienne [8]
- ▷ Méthodes de recherche directe
 - Recherche par coordonnées [3]
 - Recherche par motif généralisée [9]
 - **Recherche directe sur treillis adaptatif (Mads) [1]**

- 1 Algorithme MADS
 - Algorithme STATS-MADS
- 2 Algorithme PCA-MADS
- 3 Tests et résultats
- 4 Conclusion et suite du projet

Plan

- 1 Algorithme MADS
 - Algorithme STATS-MADS
- 2 Algorithme PCA-MADS
- 3 Tests et résultats
- 4 Conclusion et suite du projet

Algorithme MADS

Recherche directe sur treillis adaptatif (MADS)

- ▶ Généralisation de la recherche par coordonnées et la recherche par motifs généralisée

Algorithme MADS

Recherche directe sur treillis adaptatif (MADS)

- ▶ Généralisation de la recherche par coordonnées et la recherche par motifs généralisée
- ▶ Structure itérative en trois étapes

Algorithme MADS

Recherche directe sur treillis adaptatif (MADS)

- ▷ Généralisation de la recherche par coordonnées et la recherche par motifs généralisée
- ▷ Structure itérative en trois étapes
 - ▷ **Etape de recherche** : recherche globale sur le treillis

Algorithme MADS

Recherche directe sur treillis adaptatif (MADS)

- ▷ Généralisation de la recherche par coordonnées et la recherche par motifs généralisée
- ▷ Structure itérative en trois étapes
 - ▷ **Etape de recherche** : recherche globale sur le treillis
 - ▷ **Etape de sonde** : recherche autour de la solution courante

Algorithme MADS

Recherche directe sur treillis adaptatif (MADS)

- ▷ Généralisation de la recherche par coordonnées et la recherche par motifs généralisée
- ▷ Structure itérative en trois étapes
 - ▷ **Etape de recherche** : recherche globale sur le treillis
 - ▷ **Etape de sonde** : recherche autour de la solution courante
 - ▷ Etape de mise à jour et vérification d'un critère d'arrêt

Algorithme MADS

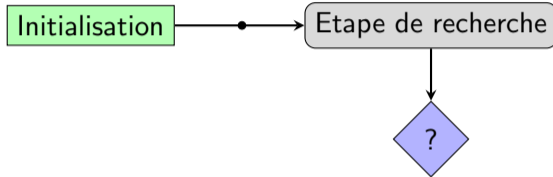
Recherche directe sur treillis adaptatif (MADS)

- ▷ Généralisation de la recherche par coordonnées et la recherche par motifs généralisée
- ▷ Structure itérative en trois étapes
 - ▷ **Etape de recherche** : recherche globale sur le treillis
 - ▷ **Etape de sonde** : recherche autour de la solution courante
 - ▷ Etape de mise à jour et vérification d'un critère d'arrêt
- ▷ Un treillis conceptuel lie les trois étapes

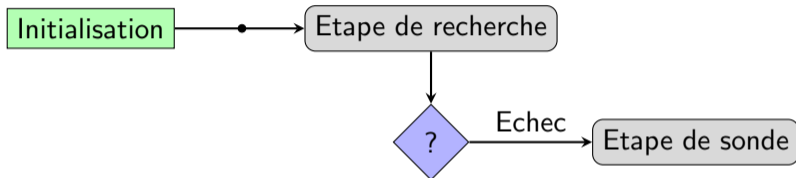
MADS : structure I

Initialisation

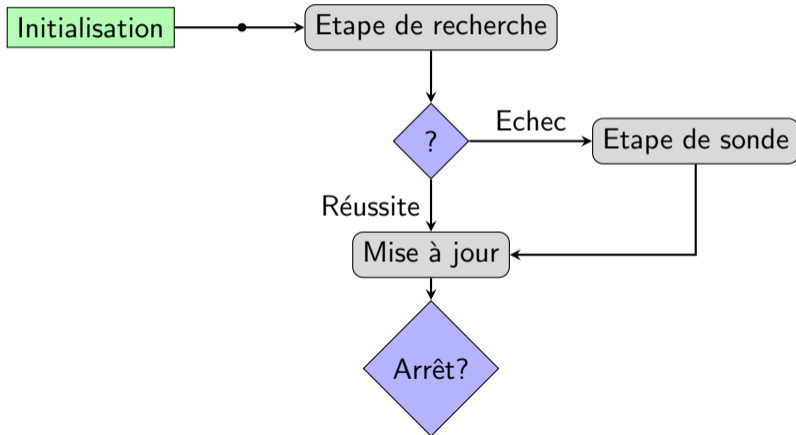
MADS : structure I



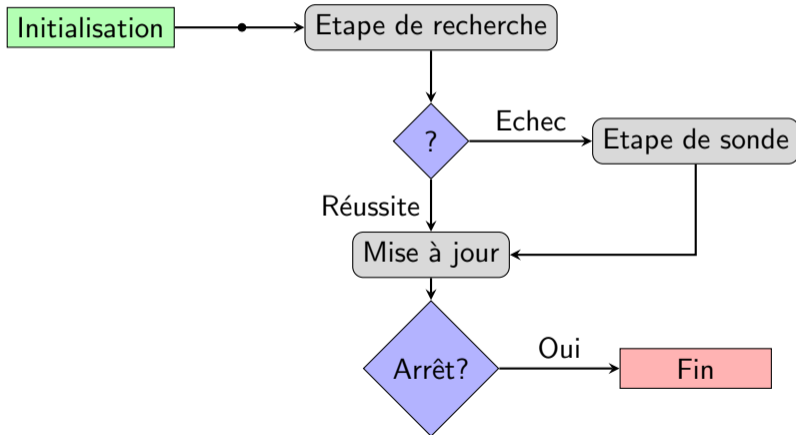
MADS : structure I



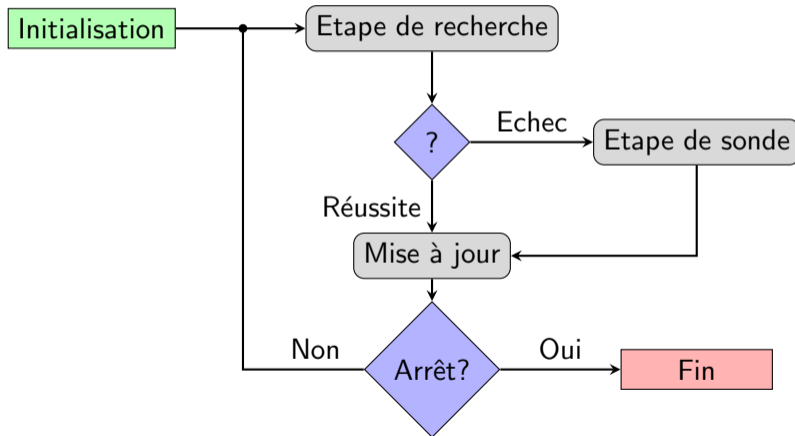
MADS : structure I



MADS : structure I



MADS : structure I



MADS : structure II

Etape de recherche

Etape de recherche

MADS : structure II

Etape de recherche

Etape de recherche

- ▶ Evaluation d'un nombre fini de points sur le treillis
- ▶ Pas nécessaire pour la convergence de l'algorithme

MADS : structure III

Etape de sonde

Etape de sonde

MADS : structure III

Etape de sonde

- ▶ Evaluation de points autour de la solution courante

Etape de sonde

MADS : structure III

Etape de sonde

Etape de sonde

- ▷ Evaluation de points autour de la solution courante
- ▷ Directions de sonde forment une base positive

MADS : structure III

Etape de sonde

Etape de sonde

- ▶ Evaluation de points autour de la solution courante
- ▶ Directions de sonde forment une base positive
- ▶ Nécessaire pour la convergence de l'algorithme

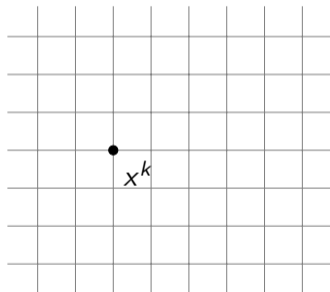
MADS : structure III

Etape de sonde

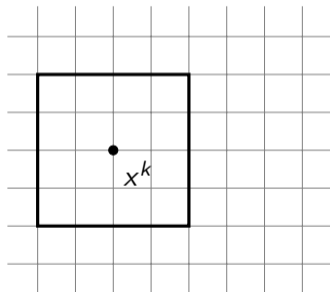
Etape de sonde

- ▷ Evaluation de points autour de la solution courante
- ▷ Directions de sonde forment une base positive
- ▷ Nécessaire pour la convergence de l'algorithme
- ▷ Peut être passée si l'étape de recherche est une réussite

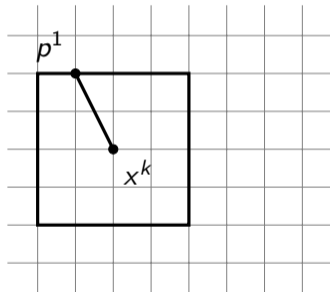
MADS : treillis



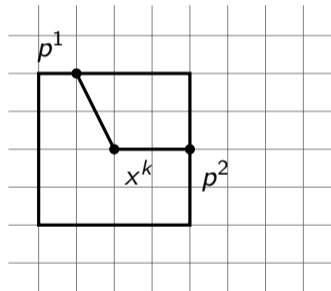
MADS : treillis



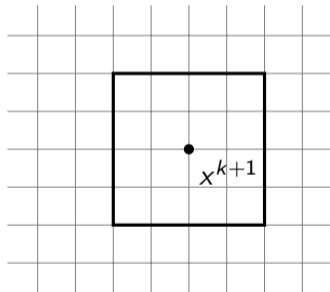
MADS : treillis



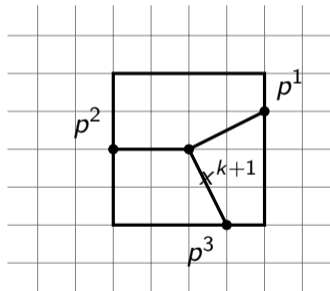
MADS : treillis



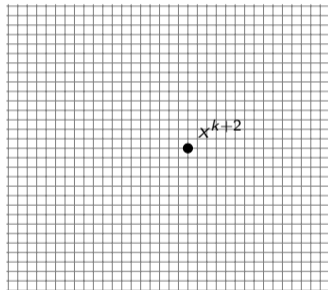
MADS : treillis



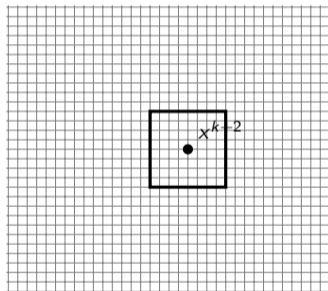
MADS : treillis



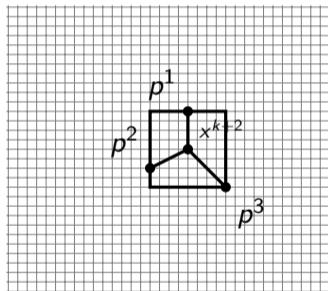
MADS : treillis



MADS : treillis



MADS : treillis



MADS : convergence

Conditions de convergence

- ▷ Repose sur l'étape de sonde
- ▷ Tous les points évalués doivent appartenir au treillis

MADS : convergence

Conditions de convergence

- ▷ Repose sur l'étape de sonde
- ▷ Tous les points évalués doivent appartenir au treillis

Résultats

- ▷ Pour une fonction localement Lipschitz, MADS converge vers un point stationnaire de Clarke
- ▷ Pour une fonction dérivable, MADS converge vers un point stationnaire x^* ,

$$\nabla f(x^*) = 0$$

Algorithme STATS-MADS

Méthode STATS-MADS

- ▶ Adaptation de MADS en grande dimension

Algorithme STATS-MADS

Méthode STATS-MADS

- ▶ Adaptation de MADS en grande dimension
- ▶ Développé par I. Ben Yahia dans son mémoire de maîtrise [2]

Algorithme STATS-MADS

Méthode STATS-MADS

- ▶ Adaptation de MADS en grande dimension
- ▶ Développé par I. Ben Yahia dans son mémoire de maîtrise [2]

Idée

- ▶ Réduction de dimension

Algorithme STATS-MADS

Méthode STATS-MADS

- ▶ Adaptation de MADS en grande dimension
- ▶ Développé par I. Ben Yahia dans son mémoire de maîtrise [2]

Idée

- ▶ Réduction de dimension
- ▶ Analyse de sensibilité à l'aide d'une analyse de variance de la cache

Algorithme STATS-MADS

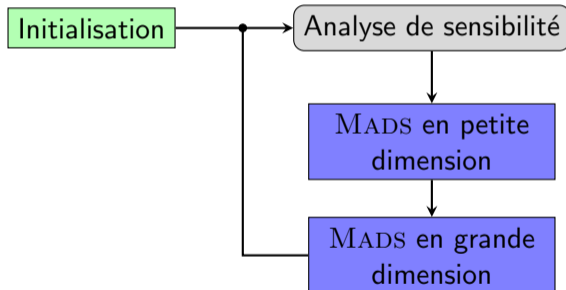
Méthode STATS-MADS

- ▷ Adaptation de MADS en grande dimension
- ▷ Développé par I. Ben Yahia dans son mémoire de maîtrise [2]

Idée

- ▷ Réduction de dimension
- ▷ Analyse de sensibilité à l'aide d'une analyse de variance de la cache
- ▷ Identifier et fixer les variables les moins influentes

Algorithme STATS-MADS : structure I



Algorithme STATS-MADS : structure II

Analyse de sensibilité

Analyse de sensibilité

- ▶ Lire la *cache*

Algorithme STATS-MADS : structure II

Analyse de sensibilité

Analyse de sensibilité

- ▷ Lire la *cache*
- ▷ Appliquer une analyse de variance et fixer les variables les moins influentes

Algorithme STATS-MADS : structure II

Analyse de sensibilité

Analyse de sensibilité

- ▷ Lire la *cache*
- ▷ Appliquer une analyse de variance et fixer les variables les moins influentes

MADS en petite
dimension

MADS en grande
dimension

Etapas de Mads

Algorithme STATS-MADS : structure II

Analyse de sensibilité

MADS en petite
dimension

MADS en grande
dimension

Analyse de sensibilité

- ▷ Lire la *cache*
- ▷ Appliquer une analyse de variance et fixer les variables les moins influentes

Etapas de Mads

- ▷ Instances de MADS avec certaines variables fixées

Algorithme STATS-MADS : structure II

Analyse de sensibilité

MADS en petite
dimension

MADS en grande
dimension

Analyse de sensibilité

- ▷ Lire la *cache*
- ▷ Appliquer une analyse de variance et fixer les variables les moins influentes

Etapes de Mads

- ▷ Instances de MADS avec certaines variables fixées
- ▷ Utilisation du même treillis

Algorithme STATS-MADS : structure II

Analyse de sensibilité

MADS en petite
dimension

MADS en grande
dimension

Analyse de sensibilité

- ▷ Lire la *cache*
- ▷ Appliquer une analyse de variance et fixer les variables les moins influentes

Etapes de Mads

- ▷ Instances de MADS avec certaines variables fixées
- ▷ Utilisation du même treillis
- ▷ Budget limité pour chaque instance

Algorithme STATS-MADS : convergence et résultats

Convergence de STATS-MADS

- ▶ STATS-MADS est une instance de MADS et hérite des mêmes propriétés de convergence

Algorithme STATS-MADS : convergence et résultats

Convergence de STATS-MADS

- ▶ STATS-MADS est une instance de MADS et hérite des mêmes propriétés de convergence
- ▶ Grâce à la sonde en grande dimension

Algorithme STATS-MADS : convergence et résultats

Convergence de STATS-MADS

- ▷ STATS-MADS est une instance de MADS et hérite des mêmes propriétés de convergence
- ▷ Grâce à la sonde en grande dimension

Performances

- ▷ Bonnes performances sur des problèmes allant jusqu'à 500 variables

Algorithme STATS-MADS : convergence et résultats

Convergence de STATS-MADS

- ▶ STATS-MADS est une instance de MADS et hérite des mêmes propriétés de convergence
- ▶ Grâce à la sonde en grande dimension

Performances

- ▶ Bonnes performances sur des problèmes allant jusqu'à 500 variables
- ▶ Majeure partie de la diminution due à l'instance de MADS en petite dimension

Plan

- 1 Algorithme MADS
 - Algorithme STATS-MADS
- 2 Algorithme PCA-MADS
- 3 Tests et résultats
- 4 Conclusion et suite du projet

Algorithme PCA-MADS I

Contexte

- ▶ Problèmes d'optimisation sans dérivées de grande dimension ($n \geq 100$)

Algorithme PCA-MADS I

Contexte

- ▶ Problèmes d'optimisation sans dérivées de grande dimension ($n \geq 100$)
- ▶ Pas de parallélisme

Algorithme PCA-MADS I

Contexte

- ▶ Problèmes d'optimisation sans dérivées de grande dimension ($n \geq 100$)
- ▶ Pas de parallélisme

Idée

Algorithme PCA-MADS I

Contexte

- ▶ Problèmes d'optimisation sans dérivées de grande dimension ($n \geq 100$)
- ▶ Pas de parallélisme

Idée

- ▶ Algorithme basé sur MADS pour hériter de ses propriétés de convergence

Algorithme PCA-MADS I

Contexte

- ▶ Problèmes d'optimisation sans dérivées de grande dimension ($n \geq 100$)
- ▶ Pas de parallélisme

Idée

- ▶ Algorithme basé sur MADS pour hériter de ses propriétés de convergence
- ▶ Concentrer l'optimisation dans les directions les plus influentes

Algorithme PCA-MADS I

Contexte

- ▶ Problèmes d'optimisation sans dérivées de grande dimension ($n \geq 100$)
- ▶ Pas de parallélisme

Idée

- ▶ Algorithme basé sur MADS pour hériter de ses propriétés de convergence
- ▶ Concentrer l'optimisation dans les directions les plus influentes
- ▶ Réduction de dimension

Algorithme PCA-MADS II

Caractéristiques de STATS-MADS

- ▶ Majeure partie de la diminution de l'objectif lors de l'optimisation en petite dimension

Algorithme PCA-MADS II

Caractéristiques de STATS-MADS

- ▶ Majeure partie de la diminution de l'objectif lors de l'optimisation en petite dimension
- ▶ Optimisation en grande dimension assez coûteuse mais nécessaire à la convergence

Algorithme PCA-MADS II

Caractéristiques de STATS-MADS

- ▶ Majeure partie de la diminution de l'objectif lors de l'optimisation en petite dimension
- ▶ Optimisation en grande dimension assez coûteuse mais nécessaire à la convergence
- ▶ Sous-problème avec un sous-ensemble des variables originales

Algorithme PCA-MADS II

Caractéristiques de STATS-MADS

- ▶ Majeure partie de la diminution de l'objectif lors de l'optimisation en petite dimension
- ▶ Optimisation en grande dimension assez coûteuse mais nécessaire à la convergence
- ▶ Sous-problème avec un sous-ensemble des variables originales

Différences entre STATS-MADS et PCA-MADS

Algorithme PCA-MADS II

Caractéristiques de STATS-MADS

- ▶ Majeure partie de la diminution de l'objectif lors de l'optimisation en petite dimension
- ▶ Optimisation en grande dimension assez coûteuse mais nécessaire à la convergence
- ▶ Sous-problème avec un sous-ensemble des variables originales

Différences entre STATS-MADS et PCA-MADS

- ▶ Possibilité de passer l'optimisation en grande dimension (étape la plus coûteuse)

Algorithme PCA-MADS II

Caractéristiques de STATS-MADS

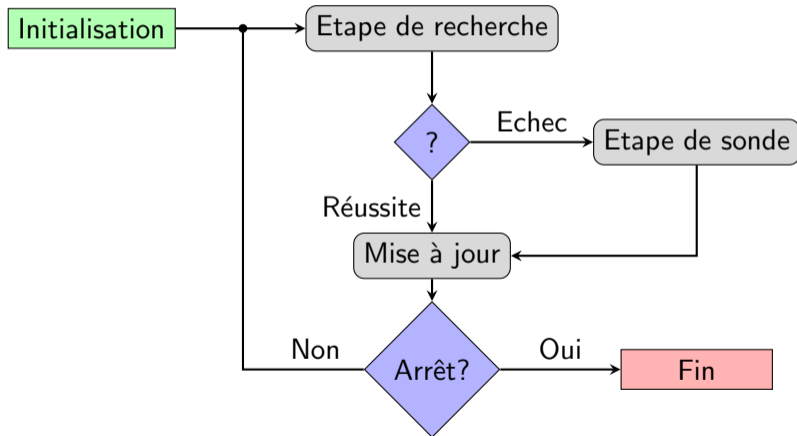
- ▶ Majeure partie de la diminution de l'objectif lors de l'optimisation en petite dimension
- ▶ Optimisation en grande dimension assez coûteuse mais nécessaire à la convergence
- ▶ Sous-problème avec un sous-ensemble des variables originales

Différences entre STATS-MADS et PCA-MADS

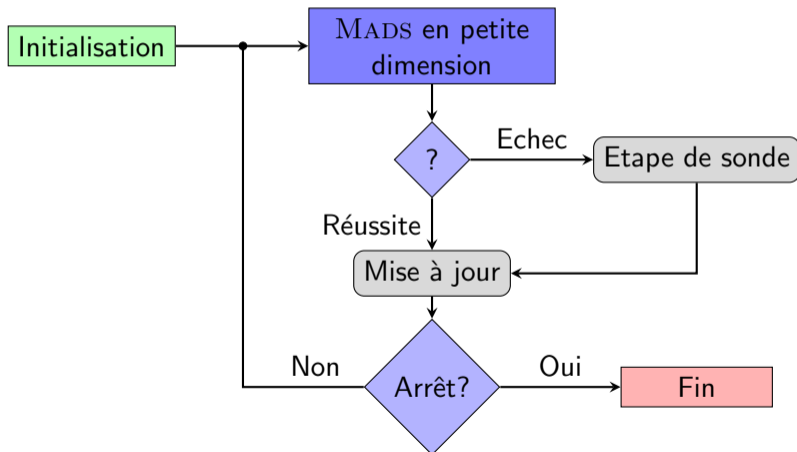
- ▶ Possibilité de passer l'optimisation en grande dimension (étape la plus coûteuse)
- ▶ Faire varier les points selon des combinaisons linéaires des variables originales

PCA-MADS : structure

PCA-MADS : structure



PCA-MADS : structure



PCA-MADS : étape de recherche

Etape de recherche de Pca-Mads

MADS en petite
dimension

PCA-MADS : étape de recherche

Etape de recherche de Pca-Mads

- ▷ Reprendre un ensemble de points de la cache

MADS en petite
dimension

PCA-MADS : étape de recherche

MADS en petite
dimension

Etape de recherche de Pca-Mads

- ▷ Reprendre un ensemble de points de la cache
- ▷ Appliquer une analyse en composante principale (PCA) sur ce nuage de points

PCA-MADS : étape de recherche

MADS en petite
dimension

Etape de recherche de Pca-Mads

- ▷ Reprendre un ensemble de points de la cache
- ▷ Appliquer une analyse en composante principale (PCA) sur ce nuage de points
- ▷ Identifier les combinaisons de variables qui semblent influencer fortement l'objectif

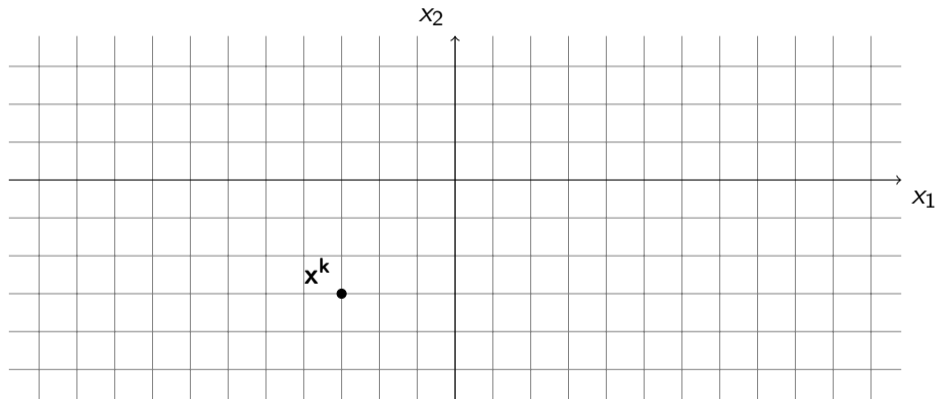
PCA-MADS : étape de recherche

MADS en petite
dimension

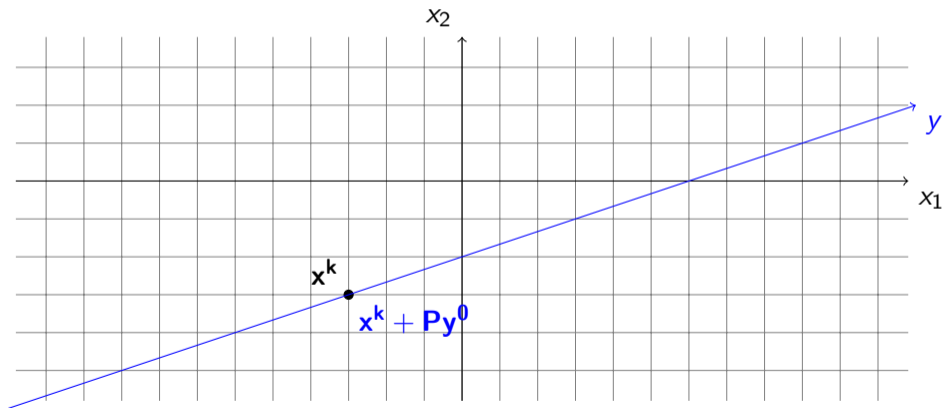
Etape de recherche de Pca-Mads

- ▷ Reprendre un ensemble de points de la cache
- ▷ Appliquer une analyse en composante principale (PCA) sur ce nuage de points
- ▷ Identifier les combinaisons de variables qui semblent influencer fortement l'objectif
- ▷ Lancer une instance de MADS sur le sous-problème contenant ces directions

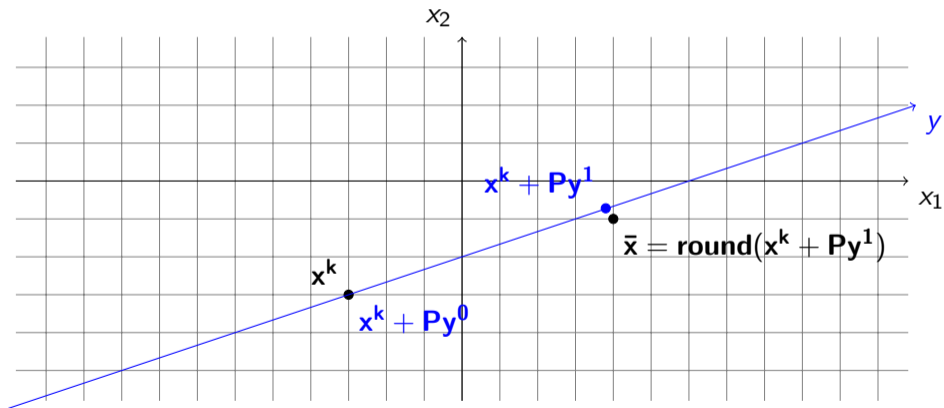
PCA-MADS : construction du sous-problème



PCA-MADS : construction du sous-problème



PCA-MADS : construction du sous-problème



Plan

- 1 Algorithme MADS
 - Algorithme STATS-MADS
- 2 Algorithme PCA-MADS
- 3 Tests et résultats**
- 4 Conclusion et suite du projet

Suite de problèmes

Tests exploratoires

- ▷ Comparaison entre une implémentation de MADS classique et PCA-MADS
- ▷ Fonction de Rosenbrock
- ▷ Petites dimensions ($2 \leq n \leq 20$)
- ▷ Grandes dimensions ($n = 100, 200, \dots, 500$)

Suite de problèmes

Tests exploratoires

- ▷ Comparaison entre une implémentation de MADS classique et PCA-MADS
- ▷ Fonction de Rosenbrock
- ▷ Petites dimensions ($2 \leq n \leq 20$)
- ▷ Grandes dimensions ($n = 100, 200, \dots, 500$)

Comparaison avec d'autres algorithmes

- ▷ Comparaison entre PCA-MADS, STATS-MADS et CMA-ES
- ▷ Suite *BBOB-largescale* de la plateforme *COCO*
- ▷ 56 problèmes de dimension 80, 160, 320 et 640

Résultats : tests exploratoires

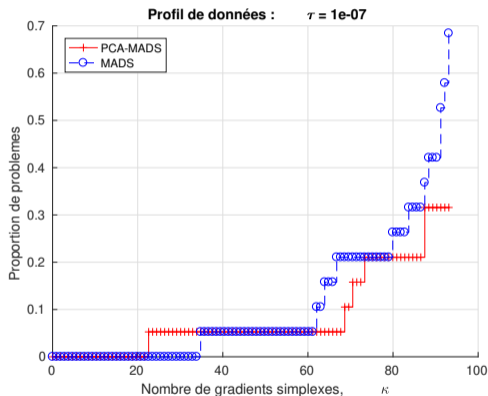


Figure: Ensemble de fonctions de Rosenbrock en petites dimensions, avec un budget de $100n$

Résultats : tests exploratoires

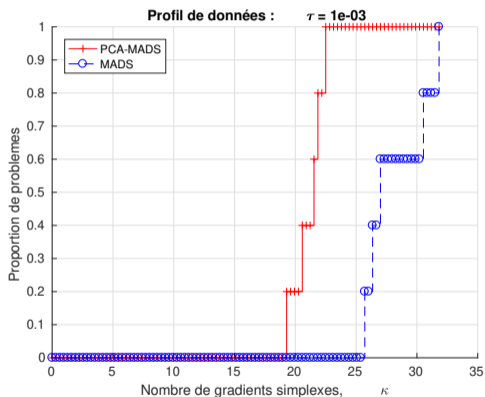


Figure: Profil de données pour l'ensemble de fonctions Rosenbrock en grandes dimensions, avec un budget de $50n$

Résultats : comparaison

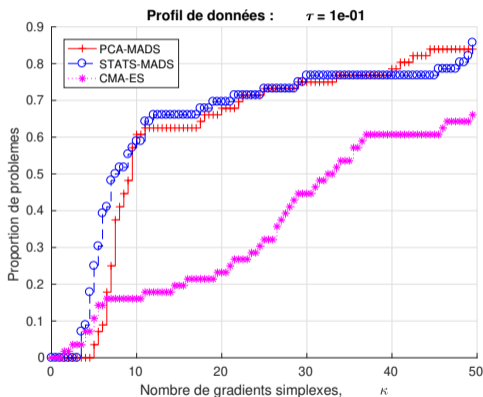


Figure: Comparaison de PCA-MADS, STATS-MADS et CMA-ES

Résultats : comparaison

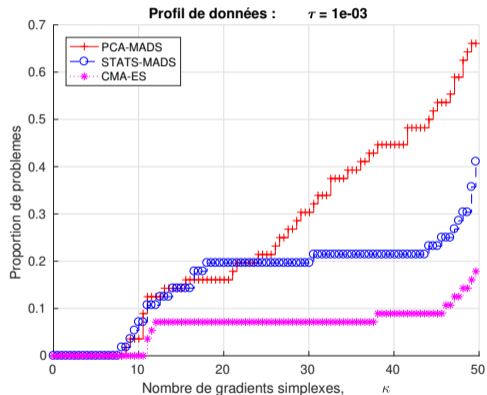


Figure: Comparaison de PCA-MADS, STATS-MADS et CMA-ES

Tests de comparaison entre STATS-MADS et PCA-MADS I

Problème	Dimension	Pca-Mads	Stats-Mads
BROWNAL [5, Problème 27]	$n = 600$	-216305309.391029	-216197339.471389
	$n = 800$	-512459636.278400	-512348752.593470
	$n = 1000$	-1000756562.998527	-1000537007.708609
	$n = 1500$	-3376668083.684607	-3376197820.744400
L1HILB [4]	$n = 600$	0.509424	1.896469
	$n = 800$	0.616835	2.392375
	$n = 1000$	4.490846	2.515649
	$n = 1500$	2.433164	3.902208

Tests de comparaison entre STATS-MADS et PCA-MADS II

Problème	Dimension	Pca-Mads	Stats-Mads
PENALTY1 [5, Problème 23]	$n = 600$	-383.202462	- 603.041571
	$n = 800$	-499.452491	- 805.309644
	$n = 1000$	-599.723565	- 1003.565068
	$n = 1500$	-891.330350	- 1509.736833
VARDIM [5, Problème 25]	$n = 600$	8932322761.775856	11317779579.835079
	$n = 800$	33427194649.599163	38882984612.787933
	$n = 1000$	77736754861.641235	98219498344.552536
	$n = 1500$	463537077338.545654	519702720318.423035

Plan

- 1 Algorithme MADS
 - Algorithme STATS-MADS
- 2 Algorithme PCA-MADS
- 3 Tests et résultats
- 4 Conclusion et suite du projet

Conclusion et travaux futurs

Conclusion

- ▷ Pas efficace en petite dimension
- ▷ L'approche de l'algorithme PCA-MADS intéressante en grande dimension
- ▷ Meilleures performances que STATS-MADS

Conclusion et travaux futurs




Conclusion

- ▷ Pas efficace en petite dimension
- ▷ L'approche de l'algorithme PCA-MADS intéressante en grande dimension
- ▷ Meilleures performances que STATS-MADS




Suite

- ▷ Comparaison sur d'autres problèmes issus de la littérature ou de l'industrie
- ▷ Comparer avec d'autres algorithmes de DFO en grande dimension
- ▷ Optimiser les paramètres de PCA-MADS
- ▷ Intégration à PSD-MADS, une version parallèle de MADS




Références I

-  C. Audet and J.E. Dennis, Jr.
Mesh Adaptive Direct Search Algorithms for Constrained Optimization.
SIAM Journal on Optimization, 17(1):188–217, 2006.
-  I. Ben Yahia.
Identification statistique de variables importantes pour l'optimisation de boîtes noires.
Master's thesis, École Polytechnique de Montréal, 2012.
-  E. Fermi and N. Metropolis.
Numerical solution of a minimum problem.
Los Alamos Unclassified Report LA-1492, Los Alamos National Laboratory, Los Alamos, USA, 1952.

Références II

-  L. Lukšan and J. Vlcek.
Test problems for nonsmooth unconstrained and linearly constrained optimization.
Technical Report 798, Institute of Computer Science, Academy of Sciences of the
Czech Republic, 2000.
-  J.J. Moré, B.S. Garbow, and Kenneth E. Hillstom.
Testing unconstrained optimization software.
ACM Transactions on Mathematical Software, 7(1):17–41, 1981.
-  J.A. Nelder and R. Mead.
A simplex method for function minimization.
The Computer Journal, 7(4):308–313, 1965.

Références III

-  L.M. Rios and N.V. Sahinidis.
Derivative-free optimization: a review of algorithms and comparison of software implementations.
Journal of Global Optimization, 56(3):1247–1293, 2013.
-  B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, and N. De Freitas.
Taking the human out of the loop: A review of Bayesian optimization.
Proceedings of the IEEE, 104(1):148–175, 2015.
-  V. Torczon.
On the convergence of pattern search algorithms.
SIAM Journal on Optimization, 7(1):1–25, 1997.