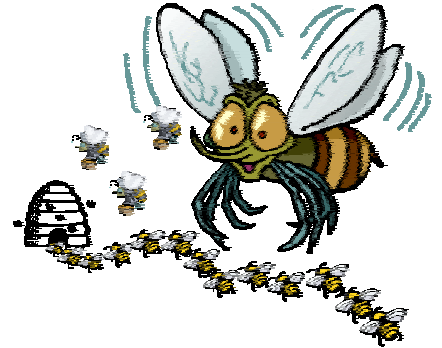




L'essaim de particules

vu comme un système dynamique :

convergence et choix des paramètres

Christian TRELEA

Institut national agronomique Paris-Grignon
UMR Génie et microbiologie des procédés alimentaires
BP 01, 78850 Thiverval-Grignon

Plan

- Motivation
- Algorithme simplifié
- Analyse dynamique
- Expériences d'optimisation
- Conclusion

Motivation

Pourquoi me suis-je intéressé à l'OEP ?

- Problèmes d'optimisation dynamique
- Procédés alimentaires et biologiques
- Espaces de recherche de grande dimension
- Nombreux optima locaux, localisation inconnue *a priori*

Pourquoi analyser l'algorithme ?

- Plusieurs paramètres de réglage,
- à l'effet difficile à caractériser
- mais fort

Algorithme simplifié

Rappel : algorithme OEP standard

Une particule :

$$\vec{v}_{k+1} = \vec{a} \otimes \vec{v}_k + \vec{b}_1 \otimes \vec{r}_1 \otimes (\vec{p}_1 - \vec{x}_k) + \vec{b}_2 \otimes \vec{r}_2 \otimes (\vec{p}_2 - \vec{x}_k)$$

$$\vec{x}_{k+1} = \vec{c} \otimes \vec{x}_k + \vec{d} \otimes \vec{v}_{k+1}$$

k = numéro de l'itération

x = position de la particule (solution candidate)

v = vitesse de la particule

p_1 = meilleure position de la particule en question

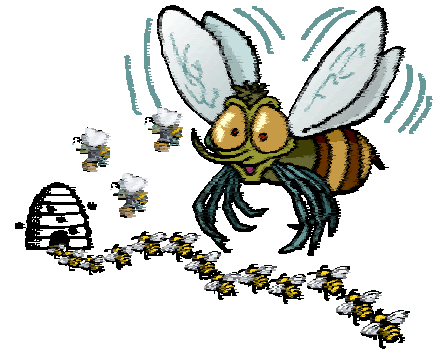
p_2 = meilleure position dans le voisinage (essaim)

a = coefficient d'inertie

b_1, b_2 = coefficients d'attraction

r_1, r_2 = nombres aléatoires de distribution uniforme en $[0 \ 1]$

$\otimes$ = produit vectoriel élément par élément



Algorithme OEP simplifié

- Dimensions traitées indépendamment
- Variables aléatoires remplacées par valeurs moyennes
- Notation $b = (b_1 + b_2) / 2$

$$p = \frac{b_1}{b_1 + b_2} p_1 + \frac{b_2}{b_1 + b_2} p_2$$

- Sans perte de généralité $c = 1, d = 1$

Reste :

$$v_{k+1} = av_k + b(p - x_k)$$

$$x_{k+1} = x_k + v_{k+1}$$

Analyse dynamique

Ecriture matricielle

Notation : $y_k = \begin{bmatrix} x_k \\ v_k \end{bmatrix} \quad A = \begin{bmatrix} 1-b & a \\ -b & a \end{bmatrix} \quad B = \begin{bmatrix} b \\ b \end{bmatrix}$

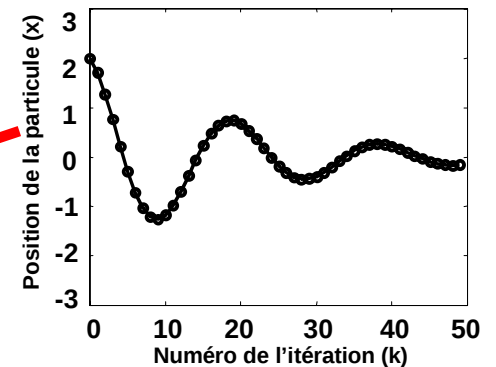
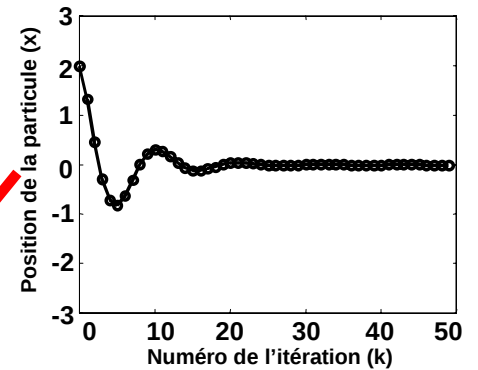
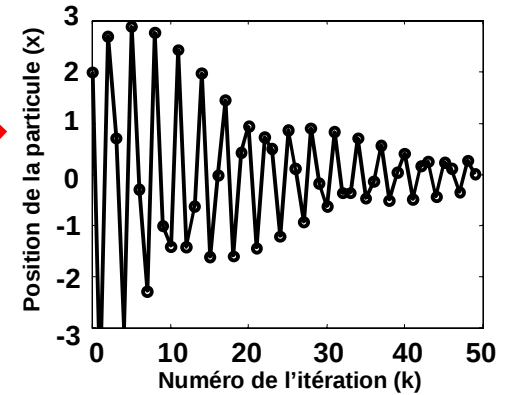
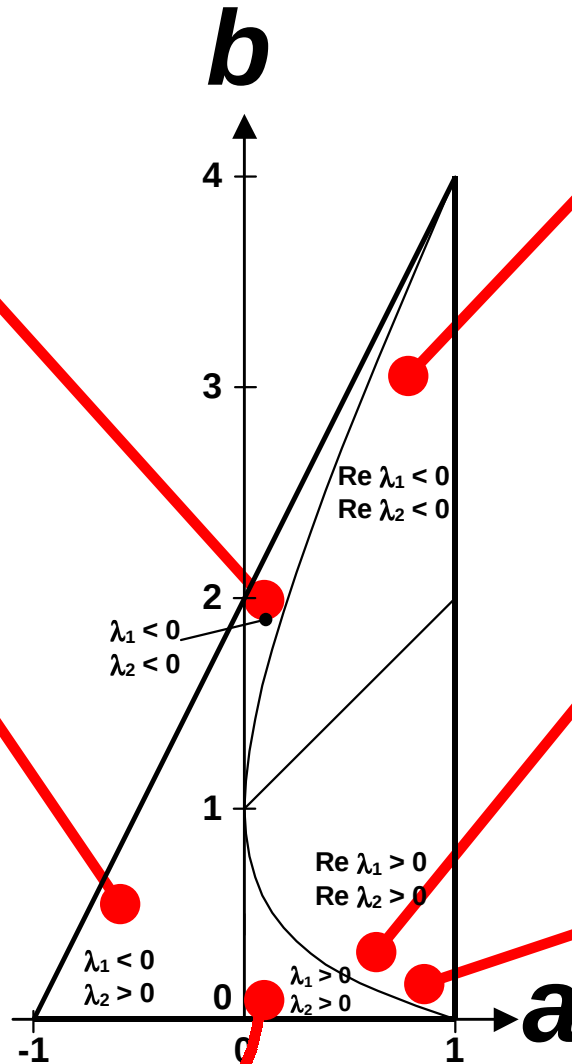
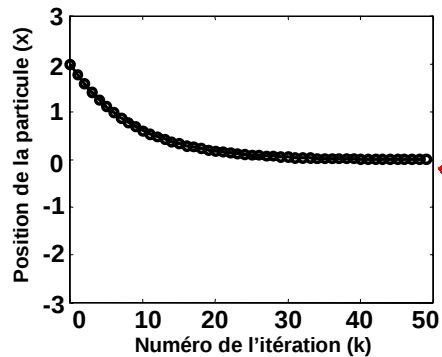
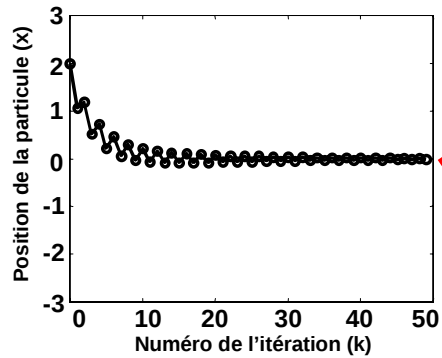
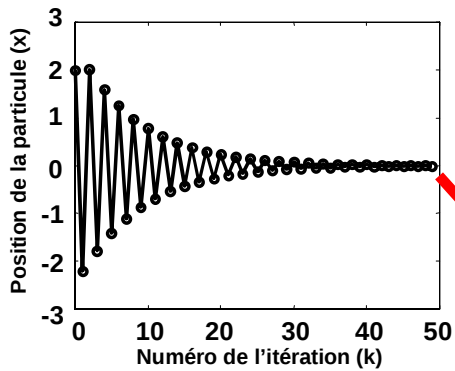
$$y_{k+1} = Ay_k + Bp$$

Point d'équilibre : $y^{eq} = \begin{bmatrix} p \\ 0 \end{bmatrix}$

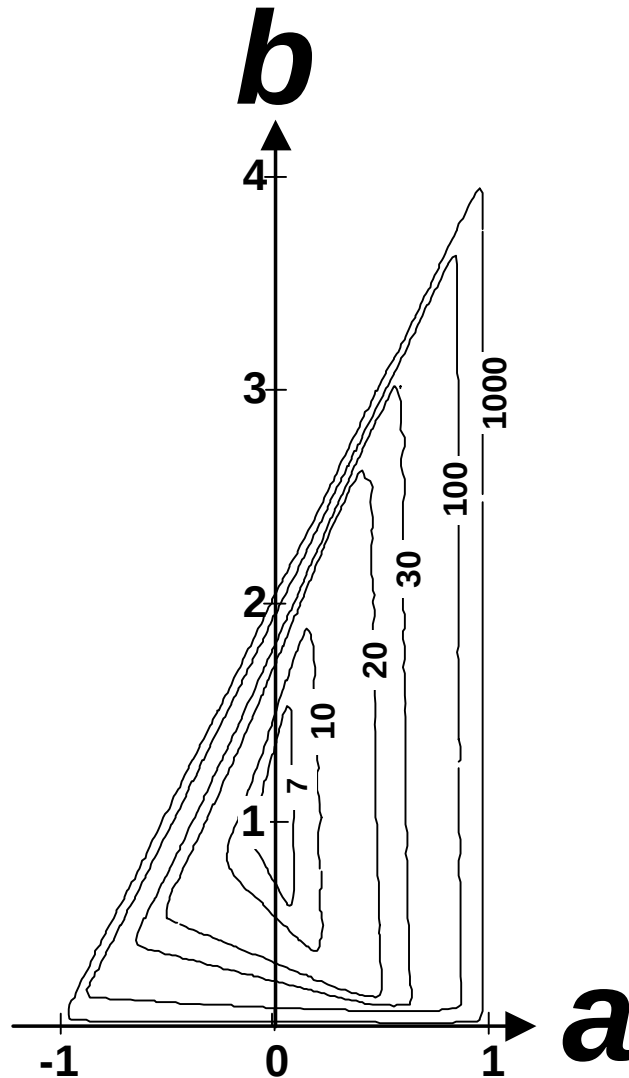
Equation caractéristique :

$$\lambda^2 - (a - b + 1)\lambda + a = 0$$

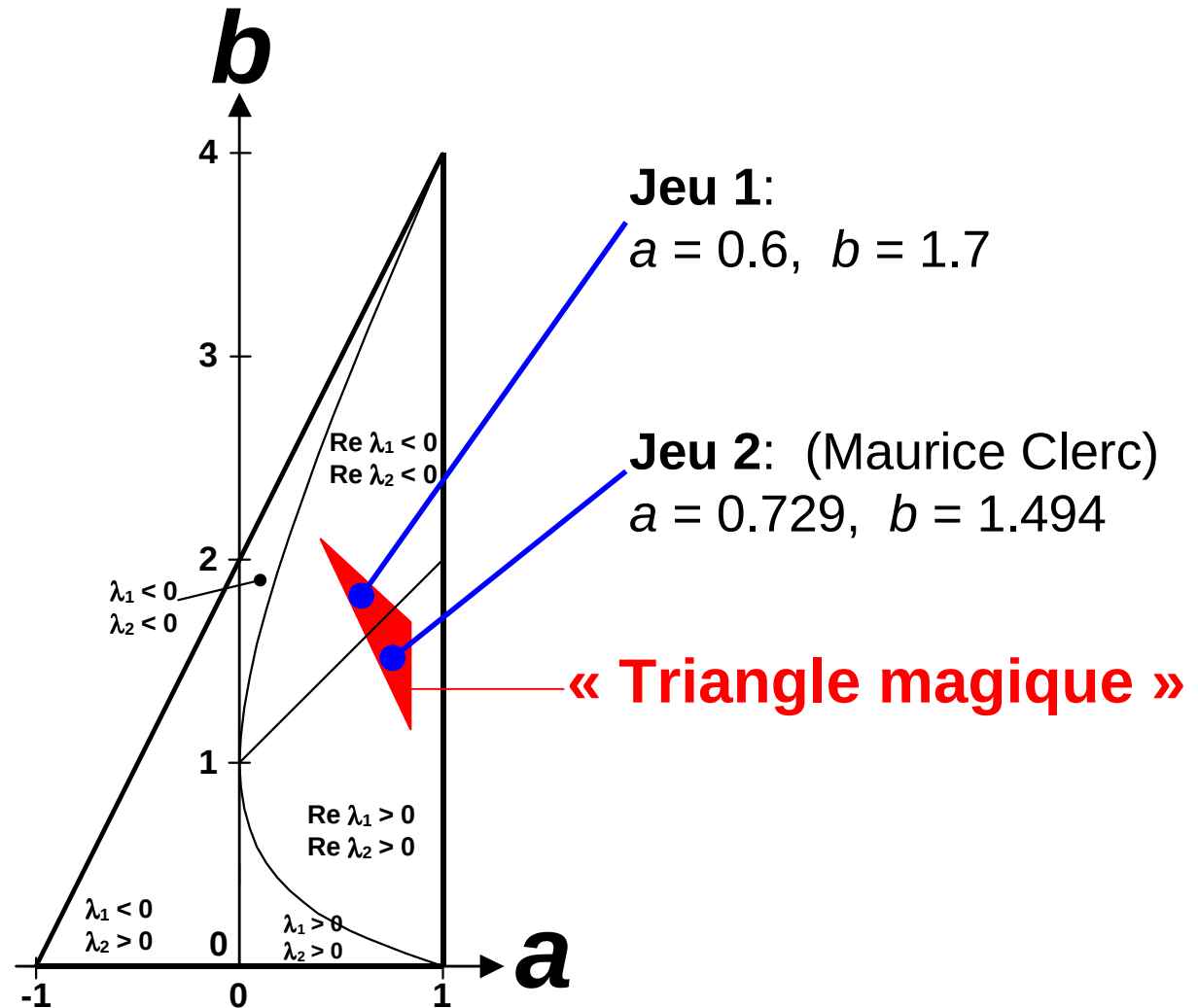
Comportements dynamiques



Vitesse de convergence

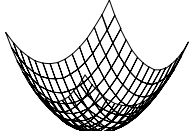
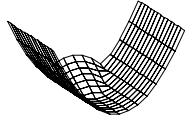
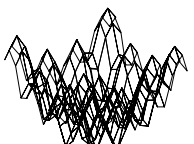
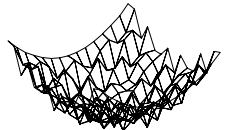
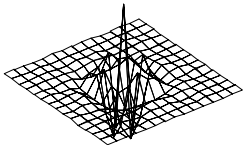


Choix empirique des paramètres



Expériences d'optimisation

Fonctions de test

Nom	Formule	Dim. n	Domaine $[x_{\min}, x_{\max}]$	Optim de f	Objectif pour f	Croquis 2D
Sphère	$f_0(\vec{x}) = \sum_{i=1}^n x_i^2$	30	$[-100, 100]^n$	0	0.01	
Rosenbrock	$f_1(\vec{x}) = \sum_{i=1}^{n-1} (100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	30	$[-30, 30]^n$	0	100	
Rastrigin	$f_2(\vec{x}) = \sum_{i=1}^n (x_i^2 - 10 \cos(2\pi x_i) + 10)$	30	$[-5.12, 5.12]^n$	0	100	
Griewank	$f_3(\vec{x}) = \frac{1}{4000} \sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	$[-600, 600]^n$	0	0.1	
Schaffer f6	$f_6(\vec{x}) = 0.5 - \frac{(\sin \sqrt{x_1^2 + x_2^2})^2 - 0.5}{(1 + 0.001(x_1^2 + x_2^2))^2}$	2	$[-100, 100]^2$	0	10^{-5}	

Conditions de test

- Nombre de particules : 15, 30 et 60
- Topologie: complètement connectée
- Paramètres a et b : Jeu 1 et Jeu 2
(identiques pour toutes les dimensions)
- Algorithme stochastique
- Arrêt : dès que la meilleure particule atteint l'objectif
avec limite à 10 000 itérations
- Répétition de chaque essai : 20 fois
- Initialisation aléatoire dans $[x_{min} \ x_{max}]$
- Vitesse et position : non limitée

Critère de performance

Nombre attendu d'évaluations de la fonction objectif

= Nombre de particules dans l'essaim

×

nombre moyen d'itérations pour atteindre l'objectif

/

taux de succès

Résultats

Nombre de particules

- $\uparrow$ no. particules: $\downarrow$ no. itérations, $\uparrow$ taux de succès, **mais** $\uparrow$ calcul par itération
- au final: no. de particules intermédiaire meilleur dans la plupart des cas

Jeux de paramètres a et b

- jeu à convergence plus rapide: $\downarrow$ no. itérations, **mais** $\downarrow$ taux de succès
favorise l'exploration au détriment de l'exploitation
- la vitesse de convergence n'est pas le seul facteur important

Forme de la fonction objectif

- optima locaux et absence de courbure de grande échelle
 $\uparrow$ no. itérations et $\downarrow$ taux de succès
- suffisamment de particules et paramètres à convergence suffisamment lente

Restent à étudier

- rôle exact de l'aléatoire
- topologie (topologie fortement connectée ~ paramètres à convergence rapide)
- position de l'optimum dans le domaine (bord, sommet, etc.)
- optima mouvants
- etc.

Conclusion

Méthode d'optimisation relativement jeune

Partiellement comprise sur le plan théorique

Très facile à programmer

Aussi puissante que les algorithmes génétiques

Nombreux perfectionnements astucieux restent à faire

