MINIMISATION D'UNE FONCTION : METHODE DU SIMPLEXE

La méthode de Nelder-Mead est un algorithme d'optimisation non-linéaire qui a été publiée par Nelder et Mead[1] en 1965. C'est une méthode numérique heuristique qui cherche à minimiser une fonction continue dans un espace à plusieurs dimensions.

Appelée également « downhill simplex method[2] », l'algorithme exploite le concept de simplexe qui est un polytope de N+1 sommets dans un espace à N dimensions. Partant initialement d'un tel simplexe, celui-ci subit des transformations simples au cours des itérations : il se déforme, se déplace et se réduit progressivement jusqu'à ce que ses sommets se rapprochent d'un point où la fonction est localement minimale.

A. PRINCIPE

A.1 Généralités

Soit une fonction f (x1, x2,...xn) appartenant à R^n pour laquelle on cherche un minimum. La méthode de **Nelder et Mead**, qui est du type exploratoire, consiste à déplacer un polygone (simplexe) de n+1 points de l'espace à n dimensions, en "réfléchissant" le plus "mauvais" sommet (par un coefficient α) du polygone par rapport à l'hyperplan des autres sommets (Fig. 1). Selon le succès rencontré par cette opération, on peut renouveler cette opération en projetant le même point plus loin dans la même direction ("expansion" par un coefficient γ), ou revenir en deçà de l'hyperplan ("contraction" par un coefficient β). Une procédure de division par deux de la taille du simplexe est prévue si l'on ne peut plus améliorer les valeurs de la fonctions en appliquant les procédures précédentes.

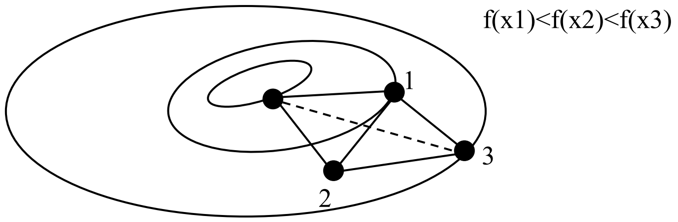


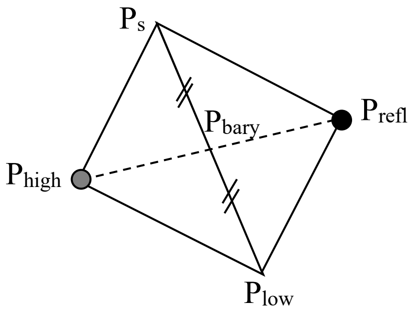
Fig. 1 : transformation d'un simplexe

A.2 Les notations

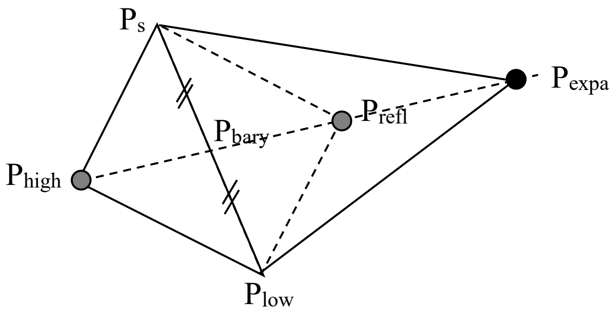
P_{high} le sommet du simplexe tel que :	$f(P_{high}) = \max (f(P_k))$ pour $k=1 \dots n+1$
P_s le sommet du simplexe tel que :	$f(P_s) = \max (f(P_k))$ pour $k \neq high$
P_{low} le sommet du simplexe tel que :	$f(P_{low}) = \min (f(P_k))$ pour $k=1 \dots n+1$
P_{bary} le barycentre du simplexe tel que :	$f(P_{bary}) = f(\sum P_k /n)$ pour $k \neq high$

A.3 Les transformations de base du simplexe

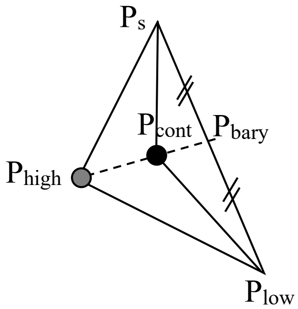
La **réflexion** : on remplace P_{high} par $P_{refl} = (1+\alpha) P_{bary} - \alpha P_{high}$ avec $\alpha > 0$



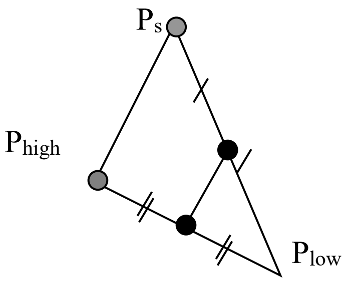
L'expansion : on remplace P_{refl} par $P_{expa} = \gamma P_{refl} + (1- \gamma) P_{bary}$ avec $\gamma > 0$



La **contraction** : on remplace P_{high} par $P_{cont} = (1- \beta) P_{bary} + \beta P_{high}$ avec $\beta > 0$



La **réduction** : on divise tous les côtés par deux : $P_k = (P_k+P_{low})/2$.



A.4 Algorithme de principe

Les opérations à effectuer dans la méthode du simplexe de Nelder et Mead dans $\mathbb{R}^n$ sont celles qui apparaissent dans l'algorithme de principe suivant.

- lire les coordonnées initiales du simplexe → P_i
- calcul des valeurs correspondantes aux différents nœuds, de la fonction à minimiser → y_i
- identifier le nœud du simplexe tel que : $f(P_{high}) = \max (f(P_k))$ pour k=1 ... n+1 → high
- identifier le nœud du simplexe tel que : $f(P_{low}) = \min (f(P_k))$ pour k=1 ... n+1 → low

calculer la position du barycentre $P_{bary} = \sum P_k /n$ pour $k \neq high$

calculer $P_{refl} = (1+\alpha) P_{bary} - \alpha P_{high}$
calculer y_{refl}
si ($y_{refl} > y_{low}$ et $y_{refl} < y_{high}$) alors
    remplacer P_{high} par P_{refl} *transformation par réflexion*
sinon
    si $y_{refl} < y_{low}$ alors
        $P_{expa} = (1+ \gamma) P_{bary} - \gamma P_{high}$
        calculer y_{expa}
        si $y_{expa} < y_{low}$ alors
            remplacer P_{high} par P_{expa} *transformation par expansion*
            mettre à jour y_{high}
        sinon
            remplacer P_{high} par P_{refl}
            mettre à jour y_{high}
        fin si
    sinon *on est dans le cas $y_{refl} > y_{high}$*
        calculer $P_{cont} = (1- \beta) P_{bary} + \beta P_{high}$
        calculer y_{cont}
        si $y_{cont} > y_{high}$ alors *la contraction ne marche pas on fait une réduction*
            remplacer les P_i, tel que $i \neq low$, par $(P_{low}+P_i)/2$ *transformation par réduction*
            mettre à jour les y_i tel que $i \neq low$
        sinon
            remplacer P_{high} par P_{cont} *transformation par contraction*
            mettre à jour y_{high}
        fin si
    fin si
fin si

On répétera ces opérations tant que le critère d'arrêt, définit par $\sqrt{\sum_i (y_i - \bar{y})^2 / n}$ (avec $\bar{y}$: valeur moyenne des y_i) n'est pas atteint. On prendra une précision de 10^{-8}.

B. TRAVAIL A EFFECTUER

B.1 A partir de l'algorithme fourni, développer un programme C.
On prendra $\alpha=1$, $\beta=0.5$ et $\gamma=2$.

B.2 Ecrire le programme. Tester votre programme avec les trois fonctions suivantes :

$$f_1(x, y) = (x-1)^2+(y-1)^2$$

$$f_2(x, y) = 100*(x^2-y)^2+(x-1)^2$$

$$f_3(x, y, z) = (x-1)^2 + (y-0.5)^2 + (y-0.5)*(z-0.75)^2$$

Tester votre programme pour différentes solutions initiales.

B.3 Tester votre programme sur une hypersphère de $\mathbb{R}^n$.

C. REFERENCES

1. J. Nelder, R. Mead, « A simplex method for function minimization », Computer Journal, vol. 7, n° 4, 1965, p. 308-313.

2. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, « Numerical Recipes : The Art of Scientific Computing », Cambridge University Press, Third Edition (2007)