Sujet de stage master 2 recherche
Algorithmes des moindres carrés non négatifs

Mots-clés : optimisation quadratique ; algorithme rapide ;
Lieu : Laboratoire GIPSA-lab (campus St Martin d'Hères)
Encadrant : Laurent Condat, chercheur CNRS à GIPSA-lab.
Contact : Laurent.Condat@gipsa-lab.grenoble-inp.fr

La méthode des moindres carrés est une des méthodes les plus utilisées dans tous les domaines d'ingénierie. Étant donné un vecteur de données y de taille M, contenant généralement du bruit de mesure, et une matrice A de taille $M \times N$ représentant le modèle linéaire fournissant les données à partir des paramètres, on cherche le vecteur de paramètres $\tilde{x} \in \mathbb{R}^N$ tel que

$$\tilde{x} \in \arg\min_{x \in \mathbb{R}^N} \|Ax - y\|^2. \tag{1}$$

Comme chacun sait, ce problème de minimisation quadratique revient à résoudre le système linéaire $A^T A\tilde{x} = A^T y$. La solution existe toujours mais n'est pas nécessairement unique.

Souvent, pour que la solution $\tilde{x}$ soit satisfaisante, il faut tenir compte de contraintes supplémentaires. Par exemple, si les éléments de $\tilde{x}$ sont des quantités physiques dont on sait qu'elles sont positives, comme des températures en Kelvins, on cherchera à résoudre le problème des **moindres carrés non négatifs** (le nom vient de *nonnegative least squares*, NNLS, en anglais, car "positif ou nul" se traduit par *nonnegative*) :

$$\tilde{x}^{\mathrm{NNLS}} \in \arg\min_{x \in \mathbb{R}^N} \|Ax - y\|^2 \quad \text{sous contrainte que } x_n \geq 0,\ \forall n = 1, \ldots, N. \tag{2}$$

À cause des contraintes, la solution de ce problème d'**optimisation quadratique** n'est plus obtenue en résolvant simplement un système linéaire. On peut aussi noter que la solution n'est pas obtenue en résolvant $A^T A\tilde{x} = A^T y$ et en mettant les éléments $\tilde{x}_n$ négatifs à zéro (bien que cette approximation de la solution soit souvent assez bonne). Par contre, si on connaissait l'**ensemble actif** de la solution, c'est-à-dire les éléments $\tilde{x}_n^{\mathrm{NNLS}}$ égaux à zéro, il suffirait de résoudre le système linéaire réduit aux autres variables (dites passives) pour obtenir la solution. Les méthodes d'activation (*active set methods*) reposent sur ce principe : elles effectuent alternativement une mise à jour de l'ensemble actif et une résolution, éventuellement approchée, du système linéaire réduit aux variables passives. La méthode la plus connue et utilisée est l'algorithme proposé par Lawson et Hanson en 1974, entre autres raisons car c'est l'algorithme implémenté dans Matlab, sous la fonction `lsqnonneg`. Il existe par ailleurs dans la littérature des algorithmes rapides pour des problèmes plus généraux d'optimisation contrainte, comme L-BFGS-B.

On s'intéresse ici aux algorithmes d'activation, qui fournissent la solution exacte (à la précision machine près) $\tilde{x}^{\mathrm{NNLS}}$ en un nombre fini d'opérations (au contraire de méthodes d'optimisation plus générales qui convergent vers la solution en un nombre infini d'itérations), dans le cas où $M \geq N$, la matrice A est de rang plein (ce qui rend la solution unique), et M et N ne sont pas trop grands (moins de 10^3). Ce stage vise à passer en revue les stratégies des méthodes de la littérature adaptées au problème NNLS, à les implémenter en C ou en Matlab, et à les comparer, en termes de vitesse, sur des problèmes NNLS tests.

La finalité de ce stage est une poursuite en thèse, visant entre autres à développer de nouveaux algorithmes rapides pour l'optimisation sur le simplexe (c.-à-d. le problème NNLS avec en plus la contrainte $\sum_{n=1}^{N} x_n = 1$), avec de nombreuses applications, par exemple en apprentissage (*machine learning*) et en traitement d'images multispectrales.