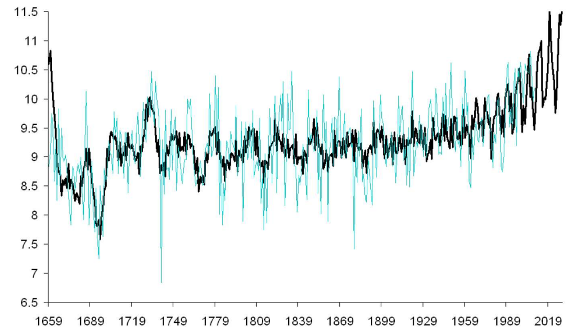
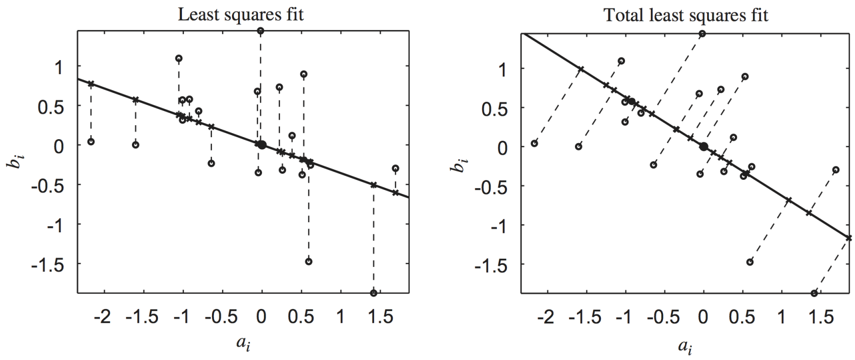
Sujet de stage master 2 recherche
Approximation structurée de rang faible

Mots-clés : approximation structurée de rang faible, décomposition en valeurs singulières, optimisation non convexe, débruitage de données
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



En bleu, température moyenne annuelle en Angleterre, sur la période 1659–2009. En noir, débruitage de la série par approximation de rang 10, et son extrapolation sur la période 2009–2029. Figure tirée de "Cadzow's basic algorithm, alternating projections and singular spectrum analysis" de J. Gillard, 2010.

La régression linéaire (résolution d'un problème des moindres carrés) permet de déterminer les paramètres d'un modèles à partir de données bruitées. Mais considérons l'exemple où les vecteurs a et b contiennent des mesures bruitées de l'intensité et de la tension aux bornes d'une résistance. En l'absence de bruit, on aurait la relation $b_i = Ra_i$, où R est la valeur de résistance que l'on cherche. La régression linéaire détermine R minimisant $\|aR-b\|^2$. Ce résultat serait bon si uniquement les valeurs b_i étaient bruitées, et pas les a_i. Ainsi, la méthode adaptée à ce cas est celle des **moindres carrés totaux** (*total least squares*). La différence entre les deux méthodes a une interprétation géométrique ci-dessous (figure tirée de "Overview of total least-squares methods" de I. Markovsky et S. Van Huffel, 2007).



En fait, il est connu que les moindres carrés totaux reviennent à faire de l'**approximation matricielle de rang faible** : la solution de l'exemple est obtenue en cherchant la matrice **débruitée** $\tilde{M} = [\tilde{a}\ \tilde{b}]$ de rang 1 minimisant la norme (de Frobenius) $\|\tilde{M} - M\|$, où $M = [a\ b]$ est la matrice construite à partir des données. $\tilde{M}$ s'obtient en tronquant la **décomposition en valeurs singulières** de M. Finalement, on a $\tilde{b} = R\tilde{a}$, où R est le paramètre recherché.

Souvent, la matrice formée à partir des données, que l'on veut approcher par une matrice de rang plus faible, a une certaine **structure** affine, par exemple c'est une matrice Toeplitz par blocs. On souhaite que la matrice détruitée de rang faible ait la même structure. Ce problème d'**approximation structurée de rang faible** (*structured low rank approximation*, SLRA), est très difficile (et même NP-difficile en général). Il faut donc se contenter en pratique d'un minimum local de ce problème d'**optimisation non convexe**. Les problèmes de SLRA ont de **nombreuses applications**, en automatique (identification de systèmes), traitement d'image et de la parole, apprentissage, finance (optimisation de portefeuille), analyse spectrale et débruitage, extrapolation de séries temporelles (en météorologie, génomique, chimiométrie, astronomie, économétrie...), etc.

Ce stage vise à appliquer, pour les comparer en termes de qualité et de rapidité, quelques algorithmes de SLRA proposés récemment (dont un par le tuteur de stage), sur quelques problèmes tests. Du code C ou Matlab est déjà disponible pour ces algorithmes, mais il sera nécessaire de l'adapter aux problèmes rencontrés. La finalité de ce stage est une poursuite en thèse, visant à développer de nouveaux algorithmes de SLRA, et à les appliquer à des problèmes pratiques de traitement du signal et des images.