

Modèles rationnels optimisés de manière exacte pour l'amélioration des procédés chimiques

27 avril 2017

1 Contexte général

1.1 Adéquation à la thématique cible

Cette proposition de thèse s'inscrit dans la thématique transverse énergie du laboratoire CNRS SAMOVAR, en initiant une collaboration avec CentraleSupélec et IFP Energies nouvelles (IFPEN). Ce dernier, établissement public à caractère industriel et commercial créé en 1944, est un acteur majeur de la recherche et de la formation dans les domaines de l'énergie, du transport et de l'environnement. Son action s'articule autour de trois priorités stratégiques : mobilité durable, énergies nouvelles et hydrocarbures responsables. L'apport de solutions aux défis sociétaux de l'énergie et du climat concentre une large partie des efforts d'IFPEN. À ce titre, sa participation apportera des compétences complémentaires à celles du laboratoire SAMOVAR.

1.2 Participants

Ce projet de thèse s'inscrit dans le cadre d'une collaboration entre :

- Marc Castella (**principal encadrant**, non HDR), Maître de Conférences à Télécom SudParis, laboratoire CNRS SAMOVAR,
- Laurent Duval (**encadrant**, non HDR), chef de projet *Data Science* et ingénieur de recherche à IFPEN,
- Jean-Christophe Pesquet (**directeur de thèse**), Professeur à CentraleSupélec.

Des collaborations existent déjà entre les membres sus-mentionnés (publications communes citées dans la description ci-dessous).

2 Description du sujet

2.1 Modèles polynomiaux et rationnels en génie des procédés

La motivation applicative de ce projet réside dans le domaine du génie des procédés physico-chimiques, pour la caractérisation et l'amélioration des performances de catalyseurs. Ce domaine est d'intérêt majeur pour IFPEN, avec de forts impacts industriels via sa filiale Axens. L'une des trois priorités stratégiques d'IFPEN vise à la production d'intermédiaires pétrochimiques et de (bio-)carburants à faible impact environnemental. La catalyse vise à l'accélération ou la réorientation de la direction d'une cinétique de réaction. L'expérimentation haut-débit permet de tester dans des conditions normalisées différentes formulations de catalyseurs, dans l'optique de sélectionner les meilleures pour atteindre les performances souhaitées. Ces compositions sont multifactorielles : supports, prétraitements, métaux rares, dopants, etc. Les tests catalytiques produisent, à partir d'analyses physico-chimiques, quelques dizaines de variables d'intérêt. Ce projet permettra d'améliorer leur efficacité en complétant des modèles physiques mal connus et des observations entachées d'erreurs à l'aide de techniques d'optimisation adaptées et novatrices. Le premier aspect porte sur les mesures physico-chimiques illustrées dans la figure 1.



FIGURE 1 – Chromatogrammes pour la quantification d’espèces chimiques.

Les proportions de différents composés chimiques sont obtenues par analyse quantitative (position, amplitude, aire) des pics positifs correspondant. Ces pics sont de forme similaire h , illustrée en 1D ou 2D (par exemple gaussiennes tronquées ou noyaux gaussiens discrets [13] issus de fonctions de Bessel modifiées, figures 1a et 1b). Ces pics et leurs dérivées sont généralement parcimonieux [16]. Les observations y s’écrivent sous la forme d’un signal inconnu \bar{x} , convolué par h , et perturbé par un bruit de mesure n additif indépendant et identiquement distribué, c’est-à-dire en désignant par k un indice entier multidimensionnel :

$$y_k = \bar{x}_k \star h_k + n_k.$$

La restauration du signal inconnu \bar{x} peut être abordée par minimisation par rapport à x de la fonctionnelle suivante, où p et q sont des polynômes favorisant la parcimonie [4] :

$$\mathcal{J}_1(x) = \|y - Hx\|^2 + \sum \frac{p(x_k)}{q(x_k)}.$$

Cette formulation de type déconvolution [18, 5] permet d’obtenir plus directement les variables d’intérêt pour l’analyse quantitative.

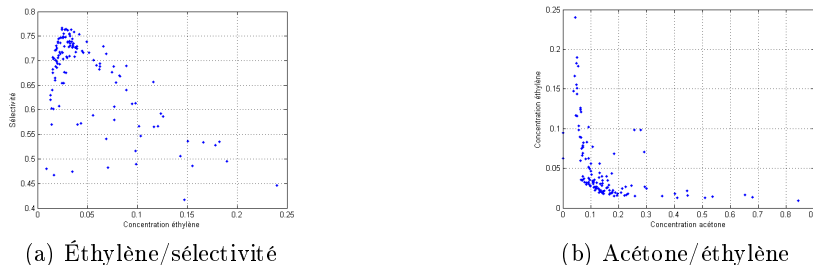


FIGURE 2 – Relations entre variables de concentration et de performance catalytique.

Pour le deuxième aspect, ces variables sont analysées en regard des performances catalytiques (par exemple sélectivité, production, conversion, stabilité), comme illustré dans la figure 2. Un défi majeur consiste à établir des relations latentes de dépendance entre leurs différents n -uplets, dont la combinatoire est importante. Celles-ci peuvent être abordées par des modèles cinétiques, qui sont souvent difficiles à obtenir dans le cas de réactions chimiques complexes. Nous proposons ici une approche plus pragmatique, basée sur des techniques de régression non linéaires, pour identifier des modèles avec un appui faible sur la physique/cinétique. Dans le cas de réactions chimiques comme $aA + bB \rightarrow cC + dD$, en notant $[X]$ la concentration de l’espèce chimique X , on peut obtenir le quotient de réaction sous la forme :

$$Q = \frac{[A]^a [B]^b}{[C]^c [D]^d}.$$

Des fonctions polynomiales ou rationnelles peuvent alors approcher localement des comportements observés entre différentes variables. Or, l’approximation de la courbe en figure 2b par une simple fonction homographe $X \rightarrow a + \frac{1}{bX+c}$ présente la difficulté de paramètres non-linéaires induisant des extrémas locaux. Il est donc courant de modifier les variables par une transformée paramétrique ϕ

ad-hoc basée sur des lois de puissance (Box-Cox [3]) et d’employer des modèles de régression linéaire standard. Néanmoins, la présence de points aberrants ou les contraintes de bornes (concentrations positives, de somme unitaire, performances en pourcentage) ne sont pas prises en compte. Enfin, il peut être intéressant de considérer des fonctions de coûts d’ordre supérieur à 2, comme des puissances paires [19, 2, 8, 7], émulant la norme de Chebyshev (∞ ou max). La régression peut alors prendre une forme plus générique :

$$\mathcal{J}_2(x) = \psi (\|y - \phi(x)\|) .$$

Combinant ces deux problèmes, nous nous intéressons à la minimisation de :

$$\mathcal{J}(x) = \psi (\|y - \phi(Hx)\|) + \sum \frac{p(x_k)}{q(x_k)} ,$$

et aux conditions (sur ψ et ϕ) sous lesquelles le signal \bar{x} dégradé non-linéairement peut être restauré. H pourra représenter selon le cas un opérateur de convolution ou de lissage de points expérimentaux. Dans [5], le cas $\psi(t) = t^2$ et $\phi(t) = \frac{t}{\delta+|t|}$ (modélisant une saturation) est résolu pour H de structure Toeplitz par bande (quand h est de réponse impulsionnelle finie). L’objectif de ce projet est de déterminer les conditions admissibles sur ψ et ϕ , et de concevoir les algorithmes d’optimisation afférents.

2.2 Optimisation polynomiale et rationnelle

Pour les critères et contraintes de type rationnels évoqués ci-dessus, la minimisation du critère est, sauf cas particulier, un problème non convexe. Les techniques usuelles de programmation non linéaire n’offrent aucune garantie de minimisation globale. Les minimas locaux, potentiellement nombreux, sont pourtant non désirés par le praticien soucieux de trouver un minimum global.

Pour des fonctions rationnelles ou polynomiales, une méthodologie d’optimisation globale a été récemment introduite [10]. Partant du problème de minimisation d’un polynôme p sur un ensemble K , elle consiste à réécrire le problème sous la forme d’un problème généralisé des moments [11], où l’on optimise par rapport à une mesure μ dont le support correspond au(x) point(s) optimal(aux) du problème initial :

$$\inf_{\mu} \int_K p(x) d\mu(x) \quad \text{avec : } \mu \text{ mesure positive sur } K \text{ et } \int_K d\mu(x) = 1 .$$

Le critère ci-dessus est alors simple (linéaire), la difficulté étant transférée sur la variable μ : comme celle-ci n’intervient qu’au travers de ses moments, elle peut être représentée par un vecteur et la méthode se fonde sur une caractérisation des séquences de moments. Une approche liée se base sur la caractérisation des polynômes positifs en tant que sommes de carrés (approche “*sum of squares*” [12]).

Après avoir fixé le degré maximal pour tous les moments ou polynômes qui apparaissent, les caractérisations ci-dessus d’existence d’une mesure ou de positivité peuvent être numériquement certifiées par programmation (convexe) semi-définie (SDP : *semi-definite programming*) et pour laquelle existent des algorithmes efficaces. Le minimum global du problème initial peut alors en théorie être obtenu par une hiérarchie de relaxations SDP d’ordre croissant.

Notons que cette approche offre des garanties d’optimalité *globale* et non pas locale comme dans le cas d’une programmation non linéaire. Cette propriété globale et les garanties associées sont particulièrement intéressantes. Néanmoins, il importe de noter que la certification numérique évoquée passe par la résolution d’un SDP de taille très grande par rapport au nombre de variables du problème initial. Les propriétés globales se payent donc au prix de la résolution d’un problème de bien plus grande dimension.

3 État de l’art et perspectives

La méthode d’optimisation évoquée ci-dessus n’a eu à ce jour qu’un impact très marginal sur les techniques de traitement du signal. Les rares tentatives d’applications de la méthode [1, 9, 20, 15] se sont limitées à des aspects très ponctuels. Au delà de la technicité de l’approche et en plus des difficultés de compréhension entre différentes communautés de chercheurs, une explication possible

tient à la limitation en taille des problèmes qui peuvent être abordés. En effet, l’augmentation de taille du problème lors du passage à la relaxation SDP limite l’approche actuelle à des problèmes de petite ou moyenne taille au vu des solveurs SDP actuellement disponibles [9].

Face aux difficultés identifiées, nous pensons pertinent d’aborder les aspects suivants :

- Comme l’ont montré nos travaux préliminaires [4, 5], face à une taille de problème élevée, des approches qui prennent en compte la structure du problème peuvent offrir une solution. À défaut, un découpage du problème en blocs de petite taille, associé à une exploration aléatoire, peut permettre de dépasser les limites de taille.
- Par ailleurs, nous envisageons d’étudier l’amélioration des solveurs SDP existant, en tirant profit des avancées récentes en optimisation convexe pour les problèmes de grande dimension.
- D’un point de vue numérique et informatique, le calcul dans des bases adaptées de polynômes pourrait se révéler pertinent.
- Dans le cas de modèles rationnels, nous proposons d’étudier des techniques permettant de fixer automatiquement les paramètres du régresseur en fonction des caractéristiques du bruit. On peut notamment songer au développement d’approches de type SURE [6, 17].
- Enfin, la faible diffusion des résultats de [10] et, à notre connaissance, l’absence de toute application pour l’analyse de données en génie des procédés, font du sujet un axe de recherche prometteur et potentiellement à fort impact dans ses applications au domaine de l’énergie (citons une application très récente [14] dans le contexte distinct de regroupement des charges traitées).

References

- [1] S. Akaho and F. Jun. Global convergence of independent component analysis based on semidefinite programming relaxation. In *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4264–4267, Prague, Czech Republic, May 22-27, 2011.
- [2] J. O. Berger. Minimax estimation of a multivariate normal mean under polynomial loss. *J. Multivariate Anal.*, 8(2):173–180, 1978.
- [3] G. E. P. Box and D. R. Cox. An analysis of transformations. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 1964.
- [4] M. Castella and J.-C. Pesquet. Optimization of a Geman-McClure like criterion for sparse signal deconvolution. In *Proc. IEEE Int. Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 317–320, Cancun, Mexico, Dec. 13-16, 2015.
- [5] M. Castella and J.-C. Pesquet. Recovery of nonlinearly degraded sparse signals through rational optimization. In *Proc. Signal Processing with Adaptive Sparse Structured Representations (SPARS) workshop (accepted for publication)*, Lisbon, Portugal, Jun. 5-8, 2017.
- [6] C. Chaux, L. Duval, A. Benazza-Benyahia, and J.-C. Pesquet. A nonlinear Stein based estimator for multichannel image denoising. *IEEE Trans. Signal Process.*, 56(8):3855–3870, Aug. 2008.
- [7] D. Fourdrinier, I. Ouassou, and W. E. Strawderman. Estimation of a mean vector under quartic loss. *J. Stat. Plann. Infer.*, 138(12):3841–3857, Dec. 2008.
- [8] K. Helmes and C. Srinivasan. Characterisation of linear mini-max estimators for loss functions of arbitrary power. *Int. Game Theor. Rev.*, 03(2 & 3):203–211, Jun. & Sep. 2001.
- [9] F. W. Kong, P. Parpas, and B. Rustem. Sum of non-concave utilities maximization for MIMO interference systems. *IEEE Trans. Wireless Communications*, 12(4):1744–1751, Apr. 2013.
- [10] J. B. Lasserre. Global optimization with polynomials and the problem of moments. *SIAM J. Optim.*, 11(3):796–817, 2001.
- [11] J. B. Lasserre. *Moments, Positive Polynomials and Their Applications*, volume 1 of *Optimization Series*. Imperial College Press, 2010.
- [12] M. Laurent. Sum of squares, moment matrices and optimization over polynomials. In M. Putinar and S. Sullivan, editors, *Emerging Applications of Algebraic Geometry*, volume 149 of *IMA Volumes in Mathematics and its Applications*, pages 157–270. Springer, 2009.

- [13] T. Lindeberg. Scale-space for discrete signals. *IEEE Trans. Pattern Anal. Mach. Intell.*, 12(3):234–254, Mar. 1990.
- [14] A. Marandi, J. Dahl, and E. de Klerk. A numerical evaluation of the bounded degree sum-of-squares hierarchy of Lasserre, Toh, and Yang on the pooling problem. *Ann. Oper. Res.*, Feb. (online) 2017.
- [15] J. Nie. Sum of squares method for sensor network localization. *Computational Optimization and Applications*, 43(2):151–179, June 2009.
- [16] X. Ning, I. W. Selesnick, and L. Duval. Chromatogram baseline estimation and denoising using sparsity (BEADS). *Chemometr. Intell. Lab. Syst.*, 139:156–167, Dec. 2014.
- [17] J.-C. Pesquet and L. Duval. Generalizations of SURE denoising to polynomial signal dependences and loss functions. *Unpublished manuscript*, 2015.
- [18] A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, and J.-C. Pesquet. Euclid in a taxicab: Sparse blind deconvolution with smoothed ℓ_1/ℓ_2 regularization. *IEEE Signal Process. Lett.*, 22(5):539–543, May 2015.
- [19] A. P. Sage and G. W. Masters. Least-squares curve fitting and discrete optimum fitting. *IEEE Trans. Educ.*, 10(1):29–36, Mar. 1967.
- [20] J. Yan and W.-S. Lu. Global design of perfect-reconstruction orthogonal cosine-modulated filter banks. In *IEEE Conference: 23rd Canadian Conference on Electrical and Computer Engineering (CCECE)*, Calgary, Canada, May 2010.