

ECOLE D'ÉTÉ DE PEYRESQ 2013

NOUVELLES TECHNIQUES D'OPTIMISATION :
APPLICATIONS EN SIGNAL, IMAGES ET TÉLÉCOMMUNICATIONS

<http://www.gretsi.fr/peyresq13/>

La théorie de l'optimisation et les techniques qui en découlent occupent une place centrale dans les domaines des signaux, des images et des télécommunications. Ces dernières années, elles ont produit de grandes avancées dans des domaines aussi divers que

- La théorie de l'estimation,
- La reconstruction de signaux ou d'images,
- Les représentations parcimonieuses de signaux et d'images,
- Les réseaux de capteurs et l'optimisation dans les systèmes distribués,
- Les environnements non stationnaires,
- Le contrôle optimal,
- L'allocation de ressources pour les réseaux de télécommunication,
- La conception de pré-codeurs et de récepteurs pour les systèmes de transmission,
- ...

La recherche dans les domaines du signal et de l'image a largement bénéficié des progrès en théorie de l'optimisation : dans une compréhension plus profonde de la structure de problèmes d'optimisation rencontrés en traitement du signal, dans la recherche de solutions acceptables à des problèmes complexes, typiquement via des techniques de relaxation, dans le développement d'algorithmes simples et efficaces, ou dans la mise en place d'algorithmes décentralisés.

En guise d'indicateur de l'intérêt que porte notre communauté aux techniques d'optimisation, une recherche sur IEEEXplore montre que le terme « Optimization » apparaît 14 500 fois approximativement dans les « Index Terms » des articles publiés dans les revues de l'IEEE depuis 2005, et 95 800 fois si on y inclut les congrès.

L'école d'été s'articule autour de deux axes : le premier concerne les outils d'optimisation récents et le second les contextes applicatifs typiques dans les domaines de recherche du Gretsi.

Elle consiste en trois cours fondamentaux de cinq heures, trois cours à orientation applicative de deux heures, et des exposés proposés par les participants sur des résultats de recherche récents.

COURS FONDAMENTAUX (3 × 5H)

Opérateurs maximaux monotones et algorithmes de recherche de zéro. Applications en optimisation convexe. Jean-Christophe Pesquet (Université de Marne-la-Vallée) et Nelly Pustelnik (ENS de Lyon).

Plan :

- (1) Opérateurs monotones
 - (a) Propriétés
 - (b) Exemples
- (2) Recherche de zéros d'opérateurs maximaux monotones
 - (a) Algorithmes primaux
 - (b) Algorithmes primaux-duaux
- (3) Exemples d'applications en traitement des signaux et des images

Pénalisation et contraintes convexes en problèmes inverses. Jean-François Giovannelli (IMS, université Bordeaux 1).

Plan :

- (1) Motivation : problèmes inverses en imagerie.
 - (a) Exemples de problèmes inverses, cas de la déconvolution
 - (b) Caractère mal-posé et nécessité de la régularisation
 - (c) Approche bayésienne (contexte stochastique) et énergétique (contexte déterministe)
 - (d) Approches énergétiques
 - (i) Pénalisation
 - (ii) Contrainte
 - (iii) Paramétrisation

La suite du cours ne concerne que les approches énergétiques dans un contexte déterministe qui débouchent sur des problèmes de minimisation de critères en général pénalisés.

- (2) Critère quadratique et régularité des images
 - (a) Propositions et motivations des critères, régularité des images
 - (b) Algorithmes de gradient sous diverses versions
 - (c) Approximation circulante et calculs rapides par FFT, filtrage de Wiener-Hunt.
 - (d) Cas contraints : Augmented Lagrangian and Alternative Directions of Multipliers (ADMM)
- (3) Critère convexe non-quadratique et préservation des discontinuités
 - (a) Propositions et motivations des critères, préservation des discontinuités
 - (b) Approches semi-quadratiques et convexe-conjuguée, transformée de Legendre-Fenchel

- (c) Deux écritures : à la Geman et Yang / à la Geman et Reynolds. Lien avec les mélanges de distributions gaussiennes
- (d) Cas contraints : Augmented Lagrangian and Alternative Directions of Multipliers (ADMM)
- (4) Deux incursions au delà :
 - (a) Cas non-convexe et détection de contours
 - (i) Propositions et motivations des critères, détection des discontinuités
 - (ii) Extension des approches algorithmiques précédentes
 - (b) Apport de l'approche bayésienne
 - (i) Quantification des erreurs, optimalités des approches
 - (ii) Estimation des hyperparamètres (approches non-supervisée)
 - (iii) Estimation de paramètres instruments (myope / semi-aveugle / autocalibration)

Bases mathématiques de la théorie des jeux, applications aux jeux de congestion. Rida Laraki (CNRS, Ecole Polytechnique).

Plan :

- (1) Jeux sous forme stratégiques (dominance, équilibre de Nash, valeur)
- (2) Jeux sous forme extensives et liens avec la forme stratégique
- (3) Raffinement : équilibre séquentiel, équilibre propre, liens entre les deux
- (4) Structure des équilibres, dynamique de Nash et stabilité
- (5) Dynamiques de meilleure réponse et du réplicateur, liens entre les deux
- (6) Applications aux jeux de congestion.

Références :

- Hofbauer J., S. Sorin and Y. Viossat (2009), "Time average replicator and best reply dynamics", *Mathematics of Operations Research*, 34, 263-269.
- Laraki R., J. Renault and S. Sorin (2013) "Bases Mathématiques de la Théorie des Jeux", *Editions de l'Ecole Polytechnique*, <https://sites.google.com/site/ridalaraki/>,
- Sandholm (2010), "Population Games and Evolutionary Dynamics", *MIT Press*.

COURS À ORIENTATION APPLICATIVE (3 × 2H)

Jeux non atomiques et routage dans les réseaux. Patrick Maillé (Télécom Bretagne).

Résumé : *We introduce a particular type of games, where players have to select a route in a network with congestion-sensitive link costs, and the effect of an individual player on congestion is negligible. We define the equilibria of those games (the so-called Wardrop equilibria), and show that they correspond to the solution of a convex optimization problem. Using a variational inequality characterization of the equilibria, we establish tight upper bounds on the Price of Anarchy, that quantifies the loss of*

efficiency due to selfishness in the system. We finally show how those methods and results can be applied in the context of wireless networks.

Plan :

- (1) Formal presentation of routing games
 - (a) Traffic routing problem
 - (b) Feasible flow, arc flow
- (2) Equilibria of a routing game
 - (a) Wardrop's principle and Nash equilibria
 - (b) Wardrop equilibrium : existence and uniqueness
 - (c) Variational inequality
- (3) Efficiency considerations
 - (a) Total cost and social optimum
 - (b) The Price of Anarchy
 - (c) How to enforce coordination among users ?
- (4) Conclusion : an example of application

Optimisation distribuée dans les réseaux de capteurs et algorithmes de Gossip. Pascal Bianchi (Télécom ParisTech).

Plan :

- (1) Contexte et objectifs
 - (a) Classes de réseaux multi-agent, réseaux synchrones et asynchrones
 - (b) Revue d'applications : machine learning, estimation statistique, coordination de flottille, communications numériques.
- (2) Méthode ADMM pour les systèmes multi-agent
 - (a) Rappels généraux
 - (b) Algorithme
 - (c) Exemple d'application
- (3) Algorithmes de consensus : calcul distribué de la moyenne sur un graphe
- (4) Méthodes du premier ordre
 - (a) Algorithme du gradient stochastique distribué
 - (b) Outils d'analyse, performances
 - (c) Autres approches

Contrôle optimal : introduction au cas déterministe en dimension finie.

Jean-Baptiste Caillaud (Université de Bourgogne).

Résumé : *Le problème type du contrôle des équations différentielles ordinaires déterministes sera présenté afin qu'un non-spécialiste puisse identifier un tel problème. La condition nécessaire de solution, connue sous le nom de « principe du maximum » (due à Pontryagin, Boltyanskii, Gamkrelidze et Mishchenko), sera abordée et illustrée dans le cas linéaire-quadratique. On commentera également les liens avec le filtrage.*

Plan :

- (1) Formulation
- (2) Principe du maximum de Pontrjagin
- (3) Exemple lineaire-quadratique
- (4) Lien avec le filtrage de Kalman

EXPOSÉS

Itakura-Saito NMF : un modèle probabiliste à facteurs latents pour la transformée de Fourier court-terme. Cédric Févotte (CNRS, laboratoire Lagrange).

Résumé :

Depuis une dizaine d'années, la factorisation en matrices non-négatives (en anglais NMF, pour "nonnegative matrix factorization") connaît un fort essor dans les domaines de l'apprentissage statistique et du traitement du signal. En particulier, de nombreuses recherches sur ce sujet prennent racine dans le domaine de l'audio, où la NMF est utilisée pour des tâches de séparation de sources ou de transcription musicale. Dans ce contexte, la NMF est appliquée au spectrogramme de magnitude ou de puissance du signal à analyser, qui est décomposé en le produit d'une matrice « dictionnaire », qui contient les motifs spectraux se répétant dans le signal, et d'une matrice « activation », qui contient les coefficients de décomposition des données sur ce dictionnaire.

Dans cet exposé nous montrerons que la décomposition NMF du spectrogramme de puissance fondée sur la divergence d'Itakura-Saito sous-tend un modèle probabiliste génératif de la transformée de Fourier court-terme (TFCT, à valeurs complexes). Dans ce modèle, la TFCT s'exprime comme une somme de composantes gaussiennes centrées et structurées par leur variance. Nous montrerons en quoi ce modèle est pertinent pour la décomposition de signaux audio.

Dans une seconde partie nous présenterons une variante « dynamique » de ce modèle, permettant de modéliser la structure temporelle des signaux via un système linéaire non-négatif à espace d'état. Des résultats de débruitage de la parole seront présentés.

Références :

- C. Févotte, N. Bertin, and J.-L. Durrieu. Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis. *Neural Computation*, 21(3) :793-830, Mar. 2009.
http://www.unice.fr/cfevotte/publications/journals/neco09_is-nmf.pdf
- C. Févotte, J. Le Roux, and J. R. Hershey. Non-negative dynamical system with application to speech and audio. In *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013.
<http://www.unice.fr/cfevotte/publications/proceedings/icassp13a.pdf>

Transport optimal et ondelettes : applications à l'image. Morgane Henry (Laboratoire Jean Kuntzmann).

Résumé :

Grâce à la découverte de liens avec différents domaines des mathématiques et la récompense par la médaille Fields de Cédric Villani en 2010, le transport optimal,

connu depuis le XVIII^{ème} siècle, a regagné l'intérêt d'un grand nombre de mathématiciens. Le transport optimal détermine une énergie pour optimiser le transport d'une fonction vers une autre fonction. Quand ces fonctions sont des images, plusieurs applications ont été étudiées, par exemple pour l'interpolation ou le recalage d'images. Nous nous plaçons dans le cadre de la formulation de Benamou-Brenier pour le transport, et nous montrons que dans l'espace des contraintes, l'énergie est strictement convexe. Nous étudions ensuite la pertinence de l'usage de bases d'ondelettes à divergence nulle sur $[0, 1]^2$ pour la résolution numérique du problème de transport optimal entre deux fonctions.

Co-factorisation douce en matrices non-négatives. Nicolas Seichepine (Télécom ParisTech).

Résumé :

Il s'agit d'une nouvelle méthode pour une co-factorisation bi-modale en matrices non-négatives. Cette méthode est adaptée aux situations où deux modalités sont liées par une même information sous-jacente. Elle permet une co-factorisation dite douce, qui prend en compte la relation entre les modalités tout en évitant l'hypothèse forte d'un même facteur en commun. Cette méthode n'impose pas que les données associées à chaque modalité aient la même dimension ou soient exprimées dans un même espace. La co-factorisation est obtenue par résolution d'un problème d'optimisation, via une méthode de majoration-minimisation. Son comportement est illustré sur des données synthétiques et un cas réel de regroupement multimodal de locuteurs où l'on exploite la corrélation entre les pistes audio et vidéo dans des débats télévisés.

The association problem in wireless networks : a Policy Gradient Reinforcement Learning approach. Richard Combes, Ilham El Bouloumi, *Stéphane Senecal* and Zwi Altman (Orange Labs).

Résumé :

The purpose of these works is to develop a self-optimized association algorithm based on Policy Gradient Reinforcement Learning (PGRL), which is both scalable, stable and robust. The term robust means that performance degradation in the learning phase should be forbidden or limited to predefined thresholds. The algorithm is model-free (as opposed to Value Iteration) and robust (as opposed to Q-Learning). The association problem is modeled as a Markov Decision Process (MDP). The policy space is parameterized. The parameterized family of policies is then used as expert knowledge for the PGRL. The PGRL converges towards a local optimum and the average cost decreases monotonically during the learning process. The properties of the solution make it a good candidate for practical implementation. Furthermore, the robustness property allows to use the PGRL algorithm in an "always-on" learning mode.

Optimisation rapide basée sur la méthode de sous-espace pour les problèmes Bayésiens Variationnels. Yuling Zheng (Laboratoire des Signaux et Systèmes).

Résumé :

La méthodologie bayésienne variationnelle a été introduite récemment pour résoudre des problèmes inverses en grande dimension. Cette méthode s'appuie sur l'approximation d'une densité de probabilité par des lois aisément calculables. Malheureusement

le problème d'optimisation fonctionnelle introduit dans cette méthodologie ne possède pas de solution explicite et nécessite la mise en oeuvre d'algorithmes itératifs dont la vitesse de convergence est réduite. Pour surmonter cette difficulté, on propose ici des nouvelles approches bayésiennes variationnelles rapides.

En transposant la méthode de descente gradient dans les espaces fonctionnels, le problème d'optimisation introduit par le bayésien variationnel a été résolu d'une manière itérative et plus efficace. Pour obtenir des approches encore plus efficaces, nous proposons ici de transposer l'optimisation par sous-espaces dans notre espace fonctionnel. Nous avons ainsi développé des approches rapides pour les problèmes Bayésiens Variationnels. Ces approches sont appliquées dans un problème de sur-résolution en utilisant un a priori TV. Les comparaisons avec les méthode de l'état de l'art ont montré que nos approches sont plus efficaces que les approches utilisant la méthodologie bayésienne variationnelle classique ainsi que celle basé sur la méthode de descente gradient.