
Techniques avancées d'apprentissage automatique basées sur DCA et applications à la maintenance prédictive

THÈSE

présentée et soutenue publiquement le 21 octobre 2022

pour l'obtention du

Doctorat de l'Université de Lorraine

(mention informatique)

par

LUU Hoang Phuc Hau

Composition du jury

<i>Rapporteurs :</i>	Emilio CARRIZOSA	Professeur, Université de Séville
	Jalal FADILI	Professeur, ENSI de Caen
<i>Examineurs :</i>	Anne BOYER	Professeur, Université de Lorraine
	Stéphane CANU	Professeur, INSA Rouen Normandie
	Xuan Vinh DOAN	Reader, Université de Warwick
	Yann GUERMEUR	Directeur de Recherche, LORIA
<i>Invité :</i>	Tao PHAM DINH	Professeur émérite, INSA Rouen Normandie
<i>Directrice de thèse :</i>	Hoai An LE THI	Professeur, Université de Lorraine
<i>Co-directeur de thèse :</i>	Hoai Minh LE	Professeur, Université de Lorraine

Résumé

L'optimisation stochastique revêt une importance majeure à l'ère du big data et de l'intelligence artificielle. Ceci est attribué à la prévalence de l'aléatoire/de l'incertitude ainsi qu'à la disponibilité toujours croissante des données, deux facteurs qui rendent l'approche déterministe infaisable. Cette thèse étudie l'optimisation stochastique non convexe et vise à résoudre les défis du monde réel, notamment l'extensibilité, variance élevée, l'incertitude endogène et le bruit corrélé. Le thème principal de la thèse est de concevoir et d'analyser de nouveaux algorithmes stochastiques basés sur la programmation DC (différence de fonctions convexes) et DCA (algorithme DC) pour répondre aux nouvelles problématiques émergeant dans l'apprentissage automatique, en particulier l'apprentissage profond. Comme application industrielle, nous appliquons les méthodes proposées à la maintenance prédictive où le problème central est essentiellement un problème de prévision de séries temporelles.

La thèse se compose de six chapitres. Les préliminaires sur la programmation DC et le DCA sont présentés dans le chapitre 1. Le chapitre 2 étudie une classe de programmes DC dont les fonctions objectives contiennent une structure de somme importante. Nous proposons deux nouveaux schémas DCA stochastiques, DCA-SVRG et DCA-SAGA, qui combinent des techniques de réduction de la variance et étudient deux stratégies d'échantillonnage (avec et sans remplacement). La convergence presque sûre des algorithmes proposés vers les points critiques DC est établie, et la complexité des méthodes est examinée. Le chapitre 3 étudie les programmes DC stochastiques généraux (la distribution de la variable aléatoire associée est arbitraire) où un flux d'échantillons i.i.d. (indépendants et identiquement distribués) de la distribution intéressée est disponible. Nous concevons des schémas DCA stochastiques dans le cadre en ligne pour résoudre directement ce problème d'apprentissage théorique. Le chapitre 4 considère une classe de programmes DC stochastiques où l'incertitude endogène est en jeu et où les échantillons i.i.d. ne sont pas disponibles. Au lieu de cela, nous supposons que seules les chaînes de Markov qui sont ergodiques assez rapidement vers les distributions cibles peuvent être accédées. Nous concevons ensuite un algorithme stochastique appelé DCA stochastique à chaînes de Markov (MCS-DCA) et fournissons une analyse de convergence dans les sens asymptotique et non asymptotique. La méthode proposée est ensuite appliquée à l'apprentissage profond via la régularisation des EDP (équations différentielles partielles), ce qui donne deux réalisations de MCSDCA, MCSDCA-odLD et MCSDCA-udLD, respectivement, basées sur la dynamique de Langevin suramortie et sous-amortie. Les applications de maintenance prédictive sont abordées au chapitre 5. La prédiction de la durée de vie utile restante (RUL) et l'estimation de la capacité sont deux problèmes centraux étudiés, qui peuvent tous deux être formulés comme des problèmes de prédiction de séries temporelles utilisant l'approche guidée par les données. Les modèles MCSDCA-odLD et MCSDCA-udLD établis au chapitre 4 sont utilisés pour former ces modèles à l'aide de réseaux neuronaux profonds appropriés. En comparaison avec divers optimiseurs de base en apprentissage profond, les études numériques montrent que les deux techniques sont supérieures, et les résultats de prédiction correspondent presque aux vraies valeurs de RUL/capacité. Enfin, le chapitre 6 met un terme à la thèse.

Mots-clés: Maintenance Prédictive, Apprentissage profond, Programmation DC et DCA

Abstract

Stochastic optimization is of major importance in the age of big data and artificial intelligence. This is attributed to the prevalence of randomness/uncertainty as well as the ever-growing availability of data, both of which render the deterministic approach infeasible. This thesis studies nonconvex stochastic optimization and aims at resolving real-world challenges, including scalability, high variance, endogenous uncertainty, and correlated noise. The main theme of the thesis is to design and analyze novel stochastic algorithms based on DC (difference-of-convex functions) programming and DCA (DC algorithm) to meet new issues emerging in machine learning, particularly deep learning. As an industrial application, we apply the proposed methods to predictive maintenance where the core problem is essentially a time series forecasting problem.

The thesis consists of six chapters. Preliminaries on DC programming and DCA are presented in Chapter 1. Chapter 2 studies a class of DC programs whose objective functions contain a large-sum structure. We propose two new stochastic DCA schemes, DCA-SVRG and DCA-SAGA, that combine variance reduction techniques and investigate two sampling strategies (with and without replacement). The proposed algorithms' almost sure convergence to DC critical points is established, and the methods' complexity is examined. Chapter 3 studies general stochastic DC programs (the distribution of the associated random variable is arbitrary) where a stream of i.i.d. (independent and identically distributed) samples from the interested distribution is available. We design stochastic DCA schemes in the online setting to directly solve this theoretical learning problem. Chapter 4 considers a class of stochastic DC programs where endogenous uncertainty is in play and i.i.d. samples are *unavailable*. Instead, we assume that only Markov chains that are ergodic fast enough to the target distributions can be accessed. We then design a stochastic algorithm termed Markov chain stochastic DCA (MCSDCA) and provide the convergence analysis in both asymptotic and nonasymptotic senses. The proposed method is then applied to deep learning via PDEs (partial differential equations) regularization, yielding two MCSDCA realizations, MCSDCA-odLD and MCSDCA-udLD, respectively, based on overdamped and underdamped Langevin dynamics. Predictive maintenance applications are discussed in Chapter 5. The remaining useful life (RUL) prediction and capacity estimation are two central problems being investigated, both of which may be framed as time series prediction problems using the data-driven approach. The MCSDCA-odLD and MCSDCA-udLD established in Chapter 4 are used to train these models using appropriate deep neural networks. In comparison to various baseline optimizers in deep learning, numerical studies show that the two techniques are superior, and the prediction results nearly match the true RUL/capacity values. Finally, Chapter 6 brings the thesis to a close.

Keywords: Predictive Maintenance, Deep Learning, DC (Difference of Convex functions) programming and DCA (DC Algorithms)