ID de Contribution: **9** Type: **Non spécifié**

Optimisation de réseaux de neurones du point de vue de la théorie de la stabilité

lundi 5 décembre 2022 09:30 (30 minutes)

Les réseaux de neurones ont récemment été utilisés pour remplacer des parties de codes de simulation. Dans ce contexte, les réseaux profonds ou excessivement larges ne sont pas adaptés du point de vue du temps de calcul. Dans cette perspective, ce travail a pour but de construire des réseaux légers et précis. Cependant c'est un véritable défi d'obtenir des réseaux performants à cause du grand nombre d'hyperparamètres à régler. Pour rendre cette étape moins chronophage, ce travail se concentre sur un des plus importants hyperparamètres: l'optimisation des poids. Malgré le grand nombre d'algorithmes d'entraînement disponibles dans la littérature, ils nécessitent un parfait réglage de leurs paramètres pour être fonctionnelles et rapides. Dans le but de s'abstraire de cette étape expérimentale, nous recherchons les propriétés que nous pouvons espérer d'un "bon" optimiseur dans le cas de fonctions non-convexes.
Notre approche se base sur l'analyse numérique des EDPs et met en exergue le rôle essentiel de certaines quantités énergétiques pour l'optimisation. Notre analyse débute en considérant l'optimiseur comme la discrétisation d'un système dynamique. Dès lors il est possible d'interpréter la sensibilité au point initial à partir de la stabilité au sens de Lyapunov. Imposer la décroissance des fonctionnelles de Lyapunov dans le cadre discret amène ainsi à une nouvelle méthodologie pour construire des algorithmes d'entraînement plus robustes tout en réduisant la dimension de l'espace des hyperparamètres. Ces nouveaux optimiseurs atteignent des performances supérieures à l'état de l'art sur différentes tâches classiques avec des modèles largements moins coûteux.

Orateur: BENSAÏD, Bilel (IMB/CEA)