

TP2: Algorithme Backpropagation

Gaétan Marceau Caron

4 décembre 2015

1 Introduction

Au cours du TP1, nous avons étudié le modèle *Softmax*¹ pour traiter le problème de classification probabiliste. Le but était de présenter deux étapes importantes de l'entraînement : la *forward propagation* et la mise à jour des paramètres. Le TP2 reprend le modèle Softmax dans un cadre plus général, celui des réseaux de neurones avec couches cachées.

Dans ce cadre, on peut considérer le modèle Softmax comme un "module" qui prend en entrée des "features", e.g. les pixels d'une image, et qui donne en sortie une loi de probabilité sur les étiquettes. D'un point de vue computationnel, un réseau de neurones est composé de plusieurs modules, transformant simplement les features d'un espace à un autre en fonction des valeurs courantes des paramètres. Ainsi, le but de l'entraînement est d'apprendre les transformations pertinentes, i.e., en modifiant les paramètres, qui permettront de réaliser la tâche associée au module de sortie. En augmentant le nombre de modules (mais aussi de fonctions non-linéaires), on augmente ainsi la complexité du modèle. La thèse du *Deep Learning* nous dit que les modules près des données d'entrée doivent apprendre des features de bas niveau, e.g., filtre de Gabor pour l'image, alors que les modules près de la sortie apprennent des features de haut niveau, e.g., la probabilité qu'il y ait un chat dans l'image. A priori, cette hiérarchie des features n'est pas imposée par le programmeur, mais apparaît naturellement lors de l'entraînement avec l'algorithme de backpropagation (RUMELHART, HINTON et WILLIAMS, 1988).

Le but du TP2 est de programmer les trois étapes essentielles à l'entraînement d'un réseau de neurones : la forward propagation, la backpropagation et la mise à jour des paramètres. Ensuite, il faudra créer un test essentiel à la vérification de votre implémentation : le test des différences finies. Finalement, vous pourrez comparer les performances de votre réseau de neurones avec celles de votre modèle Softmax de la semaine dernière.

1. aussi connu sous le nom de MaxEnt

2 Description

1. Récupérer les sources du projet à ...
2. Implémenter la fonction forward (`nn_ops.py:72`)
3. Implémenter la fonction sigmoid et sa dérivée (`nn_ops.py:149`)
4. Implémenter la fonction backward (`nn_ops.py:93`)
5. Implémenter la fonction update (`nn_ops.py:123`)
6. Implémenter le test des différences finies (`fd_test.py`)

La différence finie est une approximation de la dérivée partielle :

$$\frac{\partial l(w)}{\partial w_i} \approx \frac{l(w + \epsilon e_i) - l(w - \epsilon e_i)}{2\epsilon} \quad (1)$$

où l est une fonction à plusieurs variables avec ses dérivées partielles définies, w est un vecteur, w_i est sa i ème composante, ϵ est la longueur du pas et e_i est le i ème vecteur de la base canonique de l'espace euclidien.

3 Livrable

Date du livrable : avant le 18 décembre 2015

Format du livrable : un fichier compressé nommé *DL_tp2-prénom-nom.zip* contenant le code et le résumé

Dépôt : à l'adresse `gaetan.marceau-caron@inria.fr` avec comme objet du message *DL_tp2-prénom-nom*.

Description :

Le livrable associé au TP2 doit contenir le code de MiniNN complété et accompagné d'un résumé de une à deux pages. Le code doit s'exécuter avec la commande `python miniNN.py` et afficher l'évolution de l'apprentissage (sortie par défaut du programme). Le résumé doit être succinct et se focaliser uniquement sur les points essentiels reliés à l'entraînement des réseaux de neurones. Ce document doit décrire les difficultés que vous avez rencontrées et, dans le cas échéant, les solutions utilisées pour les résoudre. Vous pouvez aussi y décrire vos questions ouvertes et proposer une expérience sur MNIST afin d'y répondre.

Références

RUMELHART, David E., Geoffrey E. HINTON et Ronald J. WILLIAMS (1988). "Neurocomputing: Foundations of Research". In : sous la dir. de James A. ANDERSON et Edward ROSENFELD. Cambridge, MA, USA : MIT Press. Chap. Learning Representations by Back-propagating Errors, p. 696–699. ISBN : 0-262-01097-6.