

Seminario de Probabilidades de Chile

Orador: Javier Maass (CMM)

Título: Symmetries in Overparametrized Neural Networks: A Mean-Field View.

Resumen: We develop a Mean-Field (MF) view of the learning dynamics of overparametrized Artificial Neural Networks (NN) under distributional symmetries of the data w.r.t. the action of a general compact group G . We consider for this a class of generalized shallow NNs given by an ensemble of N multi-layer units, jointly trained using stochastic gradient descent (SGD) and possibly symmetry-leveraging (SL) techniques, such as Data Augmentation (DA), Feature Averaging (FA) or Equivariant Architectures (EA). We introduce the notions of weakly and strongly invariant laws (WI and SI) on the parameter space of each single unit, corresponding, respectively, to G -invariant distributions, and to distributions supported on parameters fixed by the group action (which encode EA). This allows us to define symmetric models compatible with taking $N \rightarrow \infty$ and give an interpretation of the asymptotic dynamics of DA, FA and EA in terms of Wasserstein Gradient Flows describing their MF limits. When activations respect the group action, we show that, for symmetric data, DA, FA and freely-trained models obey the exact same MF dynamic, which stays in the space of WI parameter laws and attains therein the population risk's minimizer. We also provide a counterexample to the general attainability of such an optimum over SI laws. Despite this, and quite remarkably, we show that the space of SI laws is also preserved by these MF distributional dynamics even when freely trained. This sharply contrasts the finite- N setting, in which EAs are generally not preserved by unconstrained SGD. We illustrate the validity of our findings as N gets larger, in a teacher-student experimental setting, training a student NN to learn from a WI, SI or arbitrary teacher model through various SL schemes. We lastly deduce a data-driven heuristic to discover the largest subspace of parameters supporting SI distributions for a problem, that could be used for designing EA with minimal generalization error.

Seminario de Probabilidades de Chile

Orador: Arie Worstman (ENS Paris) 16.15

Title: Feature Learning with a structured covariance

Abstract: Recent years have witnessed significant progress in machine learning, largely guided by the success of neural networks in adapting to complex data structures. In this work, we examine how neural networks learn a multi-index model when data is generated from an anisotropic Gaussian distribution with a power-law covariance structure. We focus on weak learning conditions after one gradient descent step, and compare the sample complexity necessary to learn the target function, both at initialization and after training. Our results show that structured data lowers the required sample complexity compared to the isotropic case for both neural networks and random features.

This is joint work with Bruno Loureiro (CNRS - ENS Paris).

El enlace para conectarse al seminario es:

Unirse a la reunión Zoom

<https://reuna.zoom.us/j/84521834914?pwd=OTZ6Y0NWMDpYTGtTbEt3c0luTG96UT09>

ID de reunión: 845 2183 4914

Código de acceso: 997973

Modalidad híbrida en la sala Multimedia CMM, Beauchef 851, Torre Norte Piso 6.
Miércoles 06 de Noviembre a las 15:00 horas.



UNIVERSIDAD
DE CHILE



Universidad
de Concepción



UNIVERSIDAD
ADOLFO IBÁÑEZ



Universidad
Andrés Bello



UCSC



Universidad
de Valparaíso



UNIVERSIDAD
DEL BÍO-BÍO



UNIVERSIDAD TÉCNICA
FEDERICO SANTA MARÍA



FCFM

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE



Agencia
Nacional de
Investigación
y Desarrollo
Ministerio de Ciencia,
Tecnología e Innovación
Gobierno de Chile