

## Aprendizaje automático en espacios de funciones continuas

Juan Andrés Bazerque

La novedad tecnológica en este trabajo consta en desarrollar teoría y algoritmos de aprendizaje por refuerzo (*reinforcement learning*) en espacios continuos. Se busca así dar un salto significativo en la cantidad de problemas que se podrán abordar con RL, evitando la maldición de la dimensionalidad. Nos apoyaremos en técnicas de análisis funcional basadas en *Reproducing Kernel Hilbert Spaces* (RKHSs) resolviendo un problema de optimización variacional. Más allá de su elegancia teórica estos métodos dan ventajas tecnológicas ya que la solución se interpola hacia zonas en que no se tienen datos. Usaremos asimismo nuevas técnicas de optimización en RKHSs basadas gradientes funcionales. Nuevamente esta incorporación teórica encuentra una razón de ser práctica, permitiéndonos incorporar la experiencia previamente adquirida por los agentes del sistema. La premisa de RL es que los agentes reciben una recompensa cada vez que ejecutan una acción. Estas recompensas son modeladas como muestras de un proceso estocástico cuyas distribuciones de probabilidad se desconocen. Mediante aproximación estocástica se desarrollan algoritmos en que los agentes adquieren datos en forma secuencial y ajustan sus predicciones y políticas de control. De esta forma se aprenden las distribuciones de probabilidad incrementalmente a partir de la información que arroja la experiencia. Los algoritmos estocásticos son también populares pues permiten procesar recursivamente datos masivos, y facilitan la implementación en tiempo real. Se desarrollarán entonces algoritmos estocásticos para aprender la política o acción de control óptima en RL, buscando una solución en un RKHS, y se investigará la convergencia y optimalidad de éstos.

[1] Paternain, Santiago, Juan Andrés Bazerque, Austin Small, and Alejandro Ribeiro. "Stochastic Policy Gradient Ascent in Reproducing Kernel Hilbert Spaces." *arXiv preprint arXiv:1807.11274* (2018).

Temas relacionados:

- a) RKHSs
- b) Reinforcement learning
- c) Stochastic approximation