

CD - 07p3 - 17



Mapas, transporte óptimo y análisis de datos

Esteban G. Tabak
Courant Institute of Mathematical Sciences
New York University

1 Descripción

Muchos problemas de análisis de datos pueden formularse en términos de transporte óptimo: encontrar mapas de costo mínimo que transformen una familia de distribuciones de probabilidad en otra, donde algunas de estas distribuciones sólo se conocen a través de muestras. Problemas que pueden plantearse en estos términos incluyen la normalización de datos, estimación de probabilidad, clasificación, separación en clases, simulación de procesos aleatorios, y la identificación y filtrado de factores de variabilidad. Tales planteos permiten desarrollar herramientas robustas y generales para el análisis de datos, al tiempo que sugieren nuevos problemas matemáticos en el área del transporte óptimo.

Este curso introducirá, desarrollará y aplicará estas ideas, con el objetivo de proveer una visión unificada de grandes áreas del análisis de datos, introduciendo herramientas e ideas nuevas, así como nuevos ángulos de interacción con las matemáticas. El curso está orientado a estudiantes de grado y pregrado y a científicos establecidos en matemáticas, computación, estadística u otras áreas con uso intensivo de datos, tales como la biología, las ciencias de la atmósfera y el océano y la economía.

2 Programa

Los temas a tratar incluyen:

- Elementos de transporte óptimo, formulaciones de Monge y Kantorovich.
- Formulación en términos de muestras.
- Estimación de densidad mediante mapas y flujos.
- Clasificación y separación en clases.
- Identificación y filtrado de efectos externos.
- Simulación, *importance sampling* mediante mapas.



3 Metodología

Para facilitar la participación de investigadores y estudiantes del interior del país y del exterior, este curso se ofrecerá en forma intensiva: una semana a ocho horas de clase por día. Estas incluirán cuatro horas de clases teóricas y cuatro horas de clases prácticas frente a computadoras, en colaboración con el doctor Leandro Lombardi, donde se abordarán aplicaciones de interés para los participantes.

4 Evaluación

Cada participante desarrollará a lo largo del curso un proyecto de acuerdo a sus intereses. Dada la corta duración del curso, no se espera que estos proyectos se terminen necesariamente durante el mismo, pero sí que al menos lleguen a un planteo claro. El professor Tabak estará disponible para seguir colaborando con los participantes que deseen continuar desarrollando sus proyectos después de la finalización formal del curso.

Bibliografía

E. G. Tabak y G. Trigila, "Data-Driven Optimal Transport", CPAM 2015.

Esteban G. TABAK