

MÉTODOS COMPUTACIONALES AVANZADOS

Jaime E. Forero Romero

NOMBRE DEL CURSO: Métodos Computacionales Avanzados

CÓDIGO DEL CURSO: FISI XXXX

UNIDAD ACADÉMICA: Departamento de Física

PERIODO ACADÉMICO: 201520

HORARIO:

NOMBRE PROFESOR(A) PRINCIPAL: Jaime E. Forero Romero

CORREO ELECTRÓNICO: je.forero@uniandes.edu.co

NOMBRE PROFESOR(A) COMPLEMENTARIO(A): Sebastian Perez Saaibi

CORREO ELECTRÓNICO: spsaaibi@uniandes.edu.co

I Introducción

Los métodos computacionales fundamental el trabajo en todas las áreas técnicas y científicas, ya sean principalmente experimentales o teóricas. Esto se debe en gran parte a que la capacidad de utilizar computadoras de alto rendimiento ha disminuido en costo monetario y en complejidad.

El curso de Métodos Computacionales Avanzados presenta estas posibilidades computacionales a estudiantes de diferentes disciplinas científicas. Para esto se propone profundizar sus conocimientos en tres áreas: resolución de ecuaciones diferenciales, implementación de métodos de machine learning (i.e. algoritmos que aprenden de datos) y utilización de técnicas de cómputo masivamente paralelo.

Se asume que los estudiantes de este curso ya tienen conocimientos básicos en métodos computacionales equivalentes al nivel del curso Métodos Computacionales (FISI-2028).

II Objetivos

Los objetivos principales del curso son:

- Presentar métodos para la resolución de ecuaciones diferenciales parciales y ecuaciones diferenciales estocásticas.
- Demostrar el uso de métodos y librerías de Machine Learning.
- Mostrar diferentes arquitecturas, paradigmas, lenguajes y librerías para cómputo masivamente paralelo.
- Estudiar diferentes aplicaciones prácticas a problemas científicos y del contexto industrial de las metodologías computacionales modernas.

III Competencias a desarrollar

Al finalizar el curso, se espera que el estudiante esté en capacidad de:

- Manejar lenguajes modernos de computación numérica de bajo nivel (i.e. C/C++) y de alto nivel (i.e. Python/R/Julia).
- Tener un esquema para pre-procesar, analizar y generar reportes científicos y técnicos a partir de diversas fuentes de datos utilizando métodos computacionales.
- Desplegar programas en paralelo en un cluster, al igual que en sistemas de cómputo distribuidos tales como Amazon AWS.

- Generar conocimiento a partir del modelamiento teórico y computacional de los conceptos vistos en clase.

IV Contenido por semanas

Semana 1. Unix. Conceptos básicos de simulaciones numéricas. Norma IEEE para aritmética de punto flotante. Repositorios. Makefiles.

Semana 2. Discretización de ecuaciones diferenciales parciales parabólicas (ecuación del calor) e hiperbólicas (ecuación de onda).

Semana 3. Discretización de ecuaciones diferenciales parciales elípticas (ecuación de Laplace).

Semana 4. Leyes de conservación como una expresión hiperbólica. Problema de Riemann. Método de Godunov.

Semana 5. Ecuaciones diferenciales estocásticas.

Semana 6. Métodos Monte Carlo. Cadenas de Markov.

Semana 7. Aprendizaje Supervisado: Árboles de Decisión, Clasificación, Ranking Regresión, Redes Neuronales.

Semana 8. Clustering: k-means, Jerárquico y Maximización del valor esperado (EM).

Semana 9. Reducción de Dimensionalidad y Predicción Estructurada: PCA, MDS, LDA; Modelos Gráficos, Redes Complejas y Análisis de datos Topológico.

Semana 10. Procesamiento Natural de Lenguaje (Naïve Bayes) y Aprendizaje Profundo (Deep Learning).

Semana 11. Fundamentos de programación en paralelo. Taxonomía de arquitecturas para cómputo en paralelo. Máquinas para cómputo en paralelo en Uniandes.

Semana 12. Cómo desplegar máquinas de cómputo en La Nube. Ejemplos de aplicación en Amazon Web Services y Docker.

Semana 13. Fundamentos de programación en paralelo: MPI

Semana 14. Fundamentos de programación en paralelo: OPENMP

Semana 15. Fundamentos de programación en paralelo: CUDA.

V Metodología

Cada clase tendrá una corta presentación teórica (30 minutos aproximadamente) para pasar a practicar todos los conceptos directamente en la computadora/cluster a través de ejercicios de práctica (50 minutos aproximadamente).

VI Criterios de evaluación

En el curso tendrá cinco entregas de trabajos, cada una con un valor del 20 % de la nota definitiva. Los temas de las entregas serán los siguientes:

1. Ecuaciones diferenciales parciales.
2. Ecuaciones diferenciales estocásticas.
3. Machine Learning.
4. Cómputo en paralelo en MPI.
5. Desarrollo de un proyecto propio que utilice cómputo masivamente paralelo y algún otro de los temas vistos en clase.

VII Bibliografía

Bibliografía principal:

- R. L. Burden, J. D. Faires. *Numerical analysis*, 2011. (Biblioteca General - 519.4 B862 2011)
- A. Tveito, H.P. Langtangem, B.F. Nielsen., *Elements of Scientific Computing*, 2010. (Biblioteca General, Recurso Electrónico 510.)
- O. Maimon and L. Rokach, *The Data Mining and Knowledge Discovery Handbook*, 2010. (Biblioteca General, Recurso Electrónico 006.312)

- M. Snir, *MPI : the complete reference*, 1996. (Biblioteca General, 004.35 M637)
- J. Sanders, E. Kandrot. *CUDA by example: an introduction to general-purpose GPU programming*, 2010. (Biblioteca General - 005.275 S152)

Bibliografia complementaria:

- D. Conway and J. M. White. *Machine learning for hackers*, 2012.
- S.Bird. *Natural Language Processing with Python*, 2009.
- Theano Development. *Deep Learning Tutorial* <http://deeplearning.net/tutorial/>