

BIO-109

Introduction à la programmation scientifique

Dominique Gravel

Département de biologie
Université de Sherbrooke
Local D8-3066
819-821-8000 #66589

dominique.gravel@usherbrooke.ca

4 janvier 2017

Objectif général

Les outils informatiques sont utilisés de façon croissante en écologie, que ce soit pour la réalisation d'analyses spatiales, statistiques ou pour la gestion de bases de données. La programmation scientifique intervient à de nombreuses étapes au cours du baccalauréat en biologie, notamment pour l'analyse des données écologiques et pour la réalisation de modèles de simulations. Au terme de ce cours, l'étudiant sera en mesure de conceptualiser un problème qui requiert de la programmation scientifique et de réaliser des tâches courantes de programmation. L'emphasis du cours sera donné sur les bonnes pratiques de programmation scientifique. Le langage de programmation utilisé sera R, mais les techniques de programmation de base qui seront enseignées sont transposables à tout autre langage utilisé couramment en science.

Objectifs spécifiques

Au terme de ce cours, l'étudiant sera en mesure de :

- Charger des données et exporter des résultats d'analyses au moyen du logiciel R ;
- Conceptualiser un problème au moyen de pseudo-code ;
- Manipuler des données ;
- Rédiger des fonctions ;
- Programmer des algorithmes afin de réaliser des tâches complexes, incluant des boucles et des énoncés conditionnels ;
- Réaliser des simulations de Monte Carlo ;

Pré-requis

Ce cours obligatoire est offert aux étudiants en début de programme de baccalauréat en biologie, concentration écologie. Aucun pré-requis n'est exigé pour ce cours. Ce cours sera cependant un pré-requis au cours BIO-500 sur les méthodes computationnelles en écologie, ainsi qu'au cours BIO-300 en biométrie assistée par ordinateur. R est également utilisé dans le cours ECL - 510.

Approche pédagogique

Les connaissances requises pour la programmation scientifique sont minimales, l'apprentissage porte davantage sur l'acquisition de compétences et le développement de capacités à la résolution de problèmes. Les séances seront constituées de courtes leçons magistrales sur des notions de bases de programmation, entrecoupées d'exercices spécifiques destinés à pratiquer les éléments enseignés. Les séances se conclueront sur la réalisation d'un exercice intégrateur à compléter à la maison.

L'ensemble du matériel du cours sera disponible sur un dépôt git à l'adresse :
<https://github.com/EcoNumUdS/BIO109.git>

Contenu

Cours 1 : Introduction à la programmation scientifique

- Présentation du plan de cours
- Historique et motivation au calcul scientifique
- Le pseudo-code
- Bonnes pratiques de programmation
- Installation de R Studio

Cours 2 : Les bases du langage R

- Interagir avec R
- Lire et écrire des fichiers
- Le script R
- Manipulation des objets

Cours 3 : Opérations et fonctions

- Opérations mathématiques
- L'anatomie d'une fonction
- Automatisation d'une série d'opérations

Cours 4 : Algorithmique I

- Boucles
- Opérateurs logiques

Cours 5 : Algorithmique II

- Simulations de Monte Carlo
- Optimisation des scripts

Évaluation

L'évaluation porte sur la participation aux exercices (20%) et sur un travail final (80%). Un exercice simple sera présenté à la fin des séances et chaque étudiant devra remettre la solution de l'exercice sous forme de script avant le début de la séance suivante. Les exercices peuvent être réalisés en groupe, mais chaque étudiant devra remettre sa propre copie, personnalisée. Les points sont attribués pour la participation.

L'évaluation finale portera sur la réalisation d'un projet de programmation en équipe de 4 à remettre deux semaines après la fin du dernier cours, soit au plus tard le **21 février 2017 à 16 :00**. La pénalité sera de 10% par jour de retard. Le rapport final sera évalué à parti de i) le pseudo-code pour le projet de programmation, ii) le respect des bonnes pratiques de programmation et iii) la réussite de l'exercice demandé. Les étudiants devront remettre le script nécessaire à la réalisation du projet.

Références

- Paradis, E. 2005. R pour les débutants.
[ftp ://cran.r-project.org/pub/R/doc/contrib/Paradis-rdebuts_fr.pdf](ftp://cran.r-project.org/pub/R/doc/contrib/Paradis-rdebuts_fr.pdf)
- Venables et al. 2016. An Introduction to R.
[https ://cran.r-project.org/](https://cran.r-project.org/)
- Wickham, H and Golemund, G. 2017. R for Data Science. O'Reilly Media Inc, Sebastopol.
- Centre de la Science de la Biodiversité du Québec. Ateliers R du CSBQ.
[http ://qcbs.ca/wiki/r](http://qcbs.ca/wiki/r)
- Crawley, M.J. 2013. The R Book. John Wiley and Sons, Chichester.
- Adler, J. 2011. R L'essentiel. Pearson Education France, Paris.