

Lab0 - Introduction à Python

Automne 2020

Professeur : Benjamin Leener

Chargés de laboratoire : Vicente Enguix, Gabriel Lepetit-Aimon

Installation de l'environnement de développement

Les laboratoires de ce cours seront codés en **Python 3**. Pour chaque laboratoire (y compris celui-ci) un notebook IPython (au format .ipynb) vous sera fourni. Ce document est équivalent au Matlab Live Editor, il contiendra la structure des exercices et le squelette du code organisé en cellule. Le code et les réponses aux questions seront à compléter dans le Notebook que vous remettrez sous Moodle.

Vous aurez besoin d'une version de Python >3.5, des modules **numpy**, **matplotlib**, **pandas**, **scipy** et **cv2** et d'un éditeur de Notebook IPython (jupyter, vscode...).

Voici une des procédures possibles pour installer ces outils :

1. Installer [miniconda](#) un version allégée du gestionnaire d'environnement Anaconda. Miniconda vient avec Python 3.8 et simplifiera l'installation de modules supplémentaires.
2. Lancer la ligne de commande **Anaconda Prompt** sous Windows, ou votre terminal habituel sous linux et macOS et exécuter les lignes de commandes suivantes qui créeront un environnement python **tpGBM** et y installeront les modules nécessaires :

```
conda create -n tpGBM python=3.8 numpy matplotlib pandas scipy ipykernel notebook
conda activate tpGBM
pip install opencv-python
```

3. Installer l'éditeur [vscode](#) puis son extension [ms-python](#).
4. Ouvrir le notebook **tp0.ipynb** dans vscode (**File > Open folder** puis sélectionner le chemin du dossier qui contient le notebook et ouvrir le fichier depuis le panneau **EXPLORER**)

5. Assurer vous que la version de python sélectionnée dans la barre de statut bleu en bas de l'écran est bien **tpGBM** :

Python 3.8.5 64-bit (tpGBM: conda) 0 0

Vérifier aussi que le notebook est indiqué comme **Trusted** en haut à droite de l'éditeur dans les informations de connexions au serveur python. Un popup apparaîtra probablement si ça n'est pas le cas.

6. Tenter d'exécuter la première cellule de code en cliquant sur l'icone ▷ dans le bandeau de la cellule. Le serveur jupyter devrait démarrer automatiquement en arrière plan et l'indicateur d'ordre d'exécution de la cellule devrait indiquer [1]. Il est possible qu'il faille redémarrer vscode pour que les précédents changements soient pris en compte.

Tous ces outils de développement sont libres et ne nécessitent aucune licence, si vous avez des difficultés à les installer, contactez votre chargé de laboratoire.

La suite de ce laboratoire est optionnelle si vous maîtrisez déjà Python et les modules **numpy** et **matplotlib**. Vous trouverez à la fin de ce document et sur moodle une *cheatsheet* qui énumère les fonctions usuelles de ces modules que nous utiliserons durant cette session.

Ce laboratoire présente les deux modules suivant :

- numpy fournit des outils pour manipuler des matrices de dimension arbitraire. Il est importé dans le notebook sous le nom **np**.
- matplotlib.pyplot reproduit les méthodes de matlab pour le tracé de graphs. Il est importé sous le nom de **plt**.



Bien qu'ils soient installés, les modules pythons ne sont pas directement utilisables dans l'interpréteur natif : il est nécessaire de les importer avec la commande **import nom_du_module**. Les fonctions de ce module seront alors disponible dans le reste du document par **nom_du_module.nom_de_fonction()**.



Lors de son import, on peut attribuer un acronyme à un module pour y accéder plus aisément dans le reste du code. Ainsi **import numpy as np** importe le module numpy sous le nom **np**. La variable **pi** du module sera alors accessible via **np.pi**.



Il est aussi possible de n'importer que certains attributs d'un module pour pouvoir y accéder directement dans le code. Ainsi **from numpy import pi** permet d'accéder à la valeur de π directement sous le nom **pi** (plutôt que **np.pi**).

Exercice I : Vecteurs et Courbes

1. Calculer et afficher le résultat de $\sin(0)$ et $\sin(\frac{\pi}{2})$.



Le module **numpy** fourni de nombreuses fonctions mathématiques (notamment `np.sin()`) et de constantes (notamment `np.pi`).

2. Définir le vecteur t composé de 401 valeurs réparties uniformément entre 0 et 4 avec `np.linspace`, puis afficher la longueur du vecteur.



La longueur de tout objet similaire à une liste (liste, tuple, vecteur, chaîne de caractères...) peut être lu avec `len(t)`.



Dans vscode vous pouvez, depuis la barre d'outils propre au Notebook, afficher la liste des variables définies dans le document.



3. Calculer les signaux $s_0 = \sin(\pi t)$, $s_1 = \frac{\sin(3\pi t)}{3}$ et $s_2 = \frac{\sin(5\pi t)}{5}$ où t est la valeur définie à la question précédente.



Les fonctions mathématiques comme les opérations arithmétiques, lorsqu'elles sont appliquées à une matrice numpy, s'appliquent à chacun des éléments de la matrice.

4. Affichez sur le même graphique : s_0 , s_1 et s_2 et leur somme. Pour rendre le diagramme plus lisible tracer les 3 signaux en pointillés. N'oubliez pas le titre et la légende !



Les méthodes à utiliser et un lien vers leur documentation sont disponibles dans la liste de fonction présente à la fin de ce document.

On peut observer que la somme des signaux reste périodique mais ressemble de moins en moins aux sinusoïdes dont elle est composée. Essayons d'augmenter encore le nombre de signaux sommés.

5. Calculer et afficher le signal : $\sum_{i=0}^{50} \frac{\sin((2i+1)\pi t)}{2i+1}$.



Vous pourriez définir un vecteur nul de même taille que t avec `s = np.zeros_like(t)` puis utiliser une boucle `for i in range(50):` pour ajouter successivement à `s` chacun des signaux de la somme.



En Python, les blocs de codes lors de déclaration des boucles, des conditions, des fonctions, etc, ne sont pas délimité par des `{ }` (comme en C) ou par `end` (comme en Matlab), mais par leur niveau d'indentation (généralement 4 espaces ou une tabulation par niveau). Par exemple :

```
for i in range(10):  
    → if i % 2:  
    →     → print(i, " est impair")  
    → else:  
    →     → print(i, " est pair")  
print("Fin de boucle")
```

6. Réessayer pour i allant de 0 à 500 et afficher le signal. Vous venez d'approximer un signal carré par une somme de sinusoïdales !



En traitement du signal, on nomme *décomposition en série de Fourier* le procédé qui permet de décomposer un signal périodique en une somme infinie de sinusoïdes de fréquences supérieures. Sa généralisation à tout signal complexe continu par morceaux s'appelle *la transformé de Fourier* que vous serez amené à utiliser dans ce cours.

Exercice II : Images et fonctions

Cet exercice présente le stockage et la manipulation d'images sous forme de matrices `numpy`.

1. Charger et afficher l'image `chat.png`. Dans la suite de l'énoncé on considèrera qu'elle est stockée dans la variable `img`.



Vous pourrez utiliser les méthodes `imread` et `imshow` du module `matplotlib.pyplot`.

2. Vérifier que l'image est stockée sous la forme d'une matrice `numpy` : `numpy.ndarray` en affichant son type (accessible par `type(img)`). Afficher le format de stockage des données et les dimensions de la matrice. Identifier quel canal correspond à la hauteur, la largeur et au nombre de canaux de l'image.



Les matrices `numpy` possèdent des attributs qui décrivent les propriétés de la matrice. Ainsi `img.dtype` contient le format des données, `img.ndim` contient le nombre de dimensions de la matrice et `img.shape` sa taille selon chaque dimension.



Le format de données détermine les types de nombres stockable par la matrice : flottant ou entier, signé ou non-signé (valeur uniquement positive), nombre d'octet de stockage... Pour limiter la taille de stockage des images, les pixels sont généralement codés en `uint8` : entier non-signé sur un octet (valeurs comprises entre 0 et 255) mais pour le traitement elles sont parfois converties en `float32` (flottant signé sur 4 octets, valeur comprises entre 0 et 1). Lorsque vous manipulez des images, portez attention au format attendu par les différents modules de traitement.

3. Affichez le minimum, le maximum et la moyenne des valeurs de l'image.



Les matrices `numpy` disposent de nombreuses méthodes qui dérivent des résultats de leurs valeurs. Minimum : `img.min()` ; maximum : `img.max()` ; somme : `img.sum()` ; moyenne : `img.mean()` ; écart-type : `img.std()` ; matrice transposée : `img.transpose()` ...

4. Afficher la moyenne des valeurs pour chaque canal de l'image.



Le comportements des méthodes présentées à la question précédente peut être modifié par des arguments optionnels : en l'occurrence pour `mean` le paramètre **axis** détermine quelles dimensions de la matrice seront moyennées.



En Python, il existe deux types d'arguments : les arguments obligatoire et les arguments optionnels (qui prennent une valeur par défaut lorsqu'il ne sont pas spécifié). Pour spécifier la valeur d'un argument on peut :

- fournir sa valeur : dans ce cas c'est sa position qui déterminera à quel argument elle correspond (la première valeur correspond au premier argument, la seconde au second...). Ici `img.mean((0,1))`.
- spécifier le nom de l'argument. Ici `img.mean(axis=(0,1))`. Cette syntaxe est souvent plus lisible et permet de modifier le 3ème paramètres optionnels sans spécifier la valeur des 2 premiers.

Lorsqu'on spécifie plusieurs arguments, les deux méthodes peuvent être combinées en plaçant d'abord les arguments de position puis les arguments nommés : `plt.plot([0,1,2], [0,1,2], color='red', linestyle='-')`.

5. Afficher uniquement la tête du chat : tronquer l'image pour afficher la fenêtre comprises entre les lignes 93 à 230 et les colonnes 240 à 364.



Pour accéder à des éléments spécifiques d'une matrice numpy (mais ça vaut aussi pour tous les objets similaire à des listes) on utilise la syntaxe : `img[dim1, dim2]` i.e. pour lire la case placée sur la première ligne, dernière colonne : `img[0, -1]`. Il est possible de lire un intervalle de cases consécutives avec la syntaxe :

`img[ymin:ymax, xmin:xmax]`

Si l'index min est omit l'intervalle commencera au premier élément, si l'index max est omit l'intervalle s'étendra jusqu'au dernier élément. Ainsi `img[:, -3:]` sélectionnera les 3 dernières colonnes (de la première à la dernière ligne). L'ensemble des fonctions d'indexages des matrices numpy sont décrites [ici](#).



Attention ! En Python (et contrairement à Matlab) les indexes commencent à 0 et les intervalles vont de l'index min à l'index max *exclu*. Ainsi l'intervalle correspondant aux 3 premiers éléments est `0:3` soit `[0,1,2]`.

6. Affichez le canal bleu de l'image. (Soit toutes les lignes et colonnes du 3^e canal.)

Les smileys doivent être noirs puisqu'ils ne contiennent pas de bleu.



Lorsque `imshow` doit afficher une matrice avec 3 ou 4 canaux, il considère automatiquement que c'est une image RGB (ou RGBA). Cependant lorsqu'il doit afficher une matrice avec 1 canal, il considère que c'est une carte d'intensité et l'affiche tel quel (Intensités min en violet et intensités max en jaune). Dans notre cas, on souhaite afficher le canal bleu en noir et blanc, il faut donc spécifier à `imshow` la *color map* à utiliser avec le paramètre : `cmap='gray'`.



Pour stocker et manipuler les images couleurs, on encode la couleur de chaque pixel dans 3 canaux. Quand cette image est destinée à être affichée sur un écran, on utilise le plus souvent le format **rgb** où chaque canal encode l'intensité du rouge (1^{er} canal), vert (2^e canal), bleu (3^e canal). Parfois un 4^{ème} canal est ajouté pour stocker la transparence du pixel (le canal alpha).

Pour améliorer la lisibilité de notre image on va afficher chaque canal dans sa couleur originale. Pour se faire, l'image affichée doit avoir ses 3 canaux de couleurs dont 2 nuls.

7. Créer une fonction qui retourne une image dont tous les canaux de couleurs ont été mis à 0 sauf un, spécifié en paramètre.

La signature de la fonction est donnée dans le notebook.



Comme souvent en informatique, il y a plusieurs solutions pour coder cette fonction. Une solution consiste à dupliquer l'image puis à faire une série de condition qui annule le 1^{er}, 2^e et 3^e canaux si **channel** est différent de **0**, **1**, **2** respectivement.



Attention en Python tous les objets (à l'exception des types primaires : **int**, **float**, **bool** et **tuple**) sont passés en référence. C'est à dire que la copie d'une variable ne la duplique pas en mémoire : une modification de l'une modifiera l'autre. Par exemple :

```
» a = [0,1,2]
» b = a
» b[0] = 3
» print(a)
[3, 1, 2]
```

Lorsqu'on modifie un argument d'une fonction dans sa définition il faut donc généralement en forcer la copie : **img = img.copy()**. Ainsi la variable passer en argument lors de l'appel de la fonction ne sera pas modifiée en dehors.

8. Affichez côte à côte dans la même fenêtre tous les canaux de l'image.



fig, (ax0, ax1, ax2) = plt.subplots(1, 3) permet de subdiviser l'aire de tracé pour en afficher trois graphiques côte à côte.

Vous devriez obtenir un résultat semblable à :



2. Pour calculer l'expression récursive on va voir besoin de quantifier le nombre de cellule vivante dans le voisinage de chaque pixel. La fonction **N(map)** définie plus bas, renvoie une carte où tous les pixels ont été remplacés par leur nombre de cellule vivante dans le voisinage tel que défini en introduction.

```
def N(carte):
    from scipy.ndimage import convolve
    return convolve(carte*1, np.ones((3,3)), mode='constant')
```

Vérifiez que la fonction **N(carte)** renvoie bien la matrice **n** telle que définie en introduction, en affichant **N(map)**.



convolve implémente d'une opération de traitement d'image appelée *la convolution*. Cette opération somme, pour chaque pixel, son intensité et celles des ses voisins pondérées par un masque (ici un masque unitaire 3x3). Vous la découvrirez plus en détail dans le chapitre 3 et 4.

3. On nomme **e** l'image binaire (où les pixels vivants valent **1** et les pixels morts valent **0**) et **N(e)** la carte contenant pour chaque pixel la valeur **n** telle que définie précédemment (nombre de cellules vivantes dans un carré de 3x3 autour du pixel).

Définir la fonction **compute_next_map(e)** qui prend en argument la carte actuelle **e** et renvoie la carte suivante selon la relation de récursion :

$$e_{i+1} = (N(e_i) == 3) \text{ OU } (e_i == 1 \text{ ET } N(e_i) == 4)$$



En algèbre booléenne où **FAUX** est représenté par **0** et **VRAI** par tout entier positif, on peut remplacer l'opérateur **OU** par **+** et l'opérateur **ET** par *****.



En Python, les boucles **for** sont particulièrement lentes (comparée au C/C++). On n'effectue donc jamais de boucle sur tout les pixels d'une image.



On décrira plutôt l'opération à exécuter élément par élément, et on laissera Numpy l'effectuer pour tous les éléments du tableau.

Ainsi pour une matrice **M** à n dimensions, **M==2** renverra une matrice binaire de même taille que **M** qui vaut 1 pour les cases de **M** égale à 2 et 0 sinon.

De même, **r = (M==1) + (M==4)** stockera dans **r** la somme élément par élément de **M==1** et **M==4**, et est équivalent (en beaucoup plus rapide) à :

```
r = np.zeros_like(M)
for y in range(M.shape[0]):
    → for x in range(M.shape[1]):
    → → r[y,x] = (M[y,x]==1) + (M[y,x]==4)
```

4. Calculer et afficher côte-à-côte les cartes e_0 (la carte initiale égale à la matrice `map`), e_1 , e_2 et e_3 calculées par appels successifs de la méthode `compute_next_map`.
5. Afficher à l'aide d'une boucle `while` l'évolution de la carte binaire soumise à l'algorithme du jeu de la vie, itération après itération.

Au bout d'un certain temps, l'algorithme attendra un état stable : un cycle de 2 images qui se répètent à l'infini. Vous devrez détecter ce cycle pour arrêter la boucle.



Pour mettre à jour la figure affichée sous la cellule de code, vous pouvez utiliser `clear_output(wait=True)` avant `plt.show()` pour remplacer la figure précédente et non en ajouter une nouvelle.

Le tracé des figure dans vscode ou jupiter prend beaucoup de temps : environ un quart de seconde. Pour atteindre plus vite le cycle stable (après plus de 1700 itérations) vous pouvez ne tracer qu'une itération sur 20.



Pour vérifier si le cycle stable est atteint, il vous faudra en permanence stocker dans une liste les deux dernières itérations de la carte et, à chaque itération, les comparer avec la nouvelle.

Pour comparer si deux matrices Numpy sont égales vous pouvez utiliser :
`(previous_map == current_map).all()`

Raccourcis Clavier

(Notebook)

Exécuter la cellule	Ctrl + Enter
Exécuter et passer à la cellule suivante	Shift + Enter
Auto-complétion (vscode)	Ctrl + Espace
Auto-complétion (jupyter)	TAB

Types Primaires

Chaîne de caractères	str	<code>s = '3.0'</code> ou <code>s = "False"</code>
Nombre Entier	int	<code>i = 3</code>
Nombre Flottant	float	<code>f = 3.0</code>
Boolean	bool	<code>b = True</code> ou <code>b = False</code>
Liste	list	<code>l = ['3.0', False, 3]</code>
Liste immuable	tuple	<code>t = ('3.0', False, 3)</code>
Dictionnaire	dict	<code>d = {'pi': 3.14, 'faux': False}</code>
Ensemble	set	<code>e = {1, 4, 5, 10}</code>
Objet nul		None
Lire le type d'une variable		type(variable)

Liste

Instancier une liste	<code>l = [1, '2', False]</code>
Ajouter un élément	<code>l.append(value)</code>
Insérer un élément	<code>l.insert(i, value)</code>
Supprimer un élément	<code>l.remove(value)</code>
Trier la liste	<code>l.sort()</code>
Concatener deux listes l1 et l2	<code>l1 + l2</code>

Chaîne de caractères

Instancier une chaîne	<code>s = "texte"</code>
Concaténer deux chaînes	<code>s + 'chaîne2'</code>
Test si <i>s</i> commence par 'te'	<code>s.startswith('te')</code>
Test si <i>s</i> fini par 'te'	<code>s.endswith('te')</code>
Remplacer 'tex' par 'chan'	<code>s.replace('tex', 'chan')</code>
Formater	<code>"%i / %s / %.2f" % (i, s, f)</code>

Dictionnaire

Instancier une dictionnaire	<code>d = {'key1': v1, "key1": v2}</code>
Modifier un élément	<code>d['key'] = value</code>
Itérer sur un dictionnaire	<code>for key, value in d.items():</code>

Opérations Générique

Déclarer une fonction	<code>def nom(arg1, arg2 = default):</code> → ... → <code>return valeur</code>
Boucle: For	<code>for i in range(n):</code>
Boucle: While	<code>while condition :</code>
Condition: Si	<code>if condition :</code>
Condition: Sinon si	<code>elif condition :</code>
Condition: Sinon	<code>else:</code>
Afficher dans la console.	<code>print("description:", valeur)</code>
Commentaire	<code># Commentaire</code>

Opérateurs

Égal / Différent	<code>==</code> / <code>!=</code>
Supérieur / Supérieur ou Égal	<code>></code> / <code>>=</code>
Inférieur / Inférieur ou Égal	<code><</code> / <code><=</code>
Inversion booléenne	<code>not</code>
ET booléen / OU booléen	<code>and</code> / <code>or</code>
Test si nul	<code>is None</code>
a^b	<code>a ** b</code>

Commun aux list, tuple, str...

Longueur	<code>len(l)</code>
Supprimer le <i>i</i> ^{ème} élément	<code>del l[i]</code>
Tester si l'objet contient une valeur	<code>value in l</code>
Itérer sur la liste	<code>for element in l:</code>
Appliquer la fonction <i>f</i> à toute la liste	<code>[f(v) for v in l]</code>
Filtrer la liste des objets nuls	<code>[v for v in l if v is not None]</code>

Indexes

Les indexes commencent à 0 !

Lecture / Écriture du <i>i</i> ^{ème} élément	<code>v = l[i] / l[i] = v</code>
Accéder au <i>i</i> ^{ème} dernier élément	<code>l[-i]</code>
Accéder du 2 ^{ème} au 4 ^{ème} éléments	<code>l[1:4]</code>
Accéder aux 3 premiers éléments	<code>l[:3]</code>
Accéder aux 3 derniers éléments	<code>l[-3:]</code>
Accéder à un élément sur deux	<code>l[::2]</code>
Inverser l'ordre des éléments	<code>l[::-1]</code>

Tableau (Matrices)

`import numpy as np`

Création de tableaux

Convertir une liste en tableau	<code>arr = np.array([0,1,2])</code>
Créer un tableau NxM de 0	<code>arr = np.zeros((N,M))</code>
Créer un tableau NxM de 1	<code>arr = np.ones((N,M))</code>
Créer une matrice identité NxN	<code>arr = np.eye(N)</code>
Vecteur: <code>[j, j+i, ..., k]</code>	<code>arr = np.arange(j, k, i)</code>
Intervalle de j à k avec N points	<code>arr = np.linspace(j, k, N)</code>
Charger depuis un fichier	<code>arr = np.load('path')</code>
Sauvegarder vers un fichier	<code>np.save('path', arr)</code>

Propriétés des tableaux

Type des données	<code>arr.dtype</code>
Nombre de dimensions	<code>arr.ndim</code>
Taille sur chaque dimension	<code>arr.shape</code>

Calculs sur les tableaux

Valeur minimale / maximale	<code>arr.min()</code> / <code>arr.max()</code>
Index de la valeur minimale	<code>arr.argmin()</code>
Index de la valeur maximale	<code>arr.argmax()</code>
Valeur moyenne	<code>arr.mean()</code>
Écart-type	<code>arr.std()</code>
Somme de tous les éléments	<code>arr.sum()</code>
Produit de tous les éléments	<code>arr.prod()</code>
Test si tout le tableau est vrai	<code>arr.all()</code>

Conversion de tableaux

Modifier le type de données	<code>arr.astype(np.dtype)</code>
Modifier les dimensions	<code>arr.reshape((d1, d2, ...))</code>
Permuter les axes	<code>arr.transpose(a1, a2, ...)</code>
Transformer en vecteur	<code>arr.flatten()</code>
Dupliquer le tableau	<code>arr.copy()</code>

Graphique

`import matplotlib.pyplot as plt`

Créer la zone du tracé

Créer un graph	<code>fig, ax = plt.subplots()</code>
Créer plusieurs sous-graphs	<code>fig, axs = plt.subplots(ny, nx)</code>
Afficher le graph	<code>fig.show()</code>
Choisir la taille du graph	<code>fig.set_size_inches(width, height)</code>

Tracé le graphique

Tracer une courbe	<code>ax.plot(x, y, label="nom", color='red', ...)</code>
Tracer un nuage de point	<code>ax.scatter(x, y, label="nom", color='r', ...)</code>
Tracer un histogramme	<code>ax.bar(x, height, label="nom", color='r', ...)</code>
Trace un polygone	<code>ax.fill(x, y, color='r', ...)</code>
Tracer une ligne horizontale	<code>ax.hlines(y, xmin, xmax)</code>
Tracer une ligne verticale	<code>ax.vlines(x, ymin, ymax)</code>
Afficher une annotation	<code>ax.annotate("texte", (xpos, ypos))</code>
Affiche une image	<code>ax.imshow(array, cmap="gray")</code>

Configurer l'apparence du graphique

Spécifier un titre au graph	<code>ax.set_title("Titre")</code>
Spécifier un titre global	<code>fig.suptitle("Titre de la figure")</code>
Spécifier un titre à l'axe X	<code>ax.set_xlabel("Axe X")</code>
Spécifier un titre à l'axe Y	<code>ax.set_ylabel("Axe Y")</code>
Désactiver les axes	<code>ax.axis("off")</code>
Afficher une grille	<code>ax.grid(color='r', linewidth=2)</code>
Affiche la légende	<code>ax.legend()</code>

Manipulation d'images

Charger une image	<code>img = plt.imread("fichier.jpg")</code>
Accès à une région	<code>img[y_min : y_max, x_min : x_max]</code>
Accès à un canal	<code>img[:, :, canal]</code>