

Modélisation d'une épidémie

1 Description du modèle choisi

Pour modéliser une épidémie, on partage la population entre quatre catégories : les individus sains, infectés, rétablis (ils sont alors immunisés), et décédés. Pour tout $t \in \mathbb{R}$, on note :

- $S(t)$ la proportion d'individus sains à l'instant t ;
- $I(t)$ la proportion d'individus infectés à l'instant t ;
- $R(t)$ la proportion d'individus rétablis à l'instant t ;
- $D(t)$ la proportion d'individus décédés à l'instant t .

On suppose l'existence d'une constante $c \in \mathbb{R}^{+*}$, appelée « taux de contagion », telle que $S' = -cSI$ (le nombre de nouveaux infectés est proportionnel à $S \times I$, lui-même proportionnelle à la probabilité qu'un individu sain entre en contact avec un porteur).

Pour les autres fonctions, on suppose qu'il existe deux constantes d et r telle que

$$\forall t \in \mathbb{R}, \begin{cases} S'(t) = -cS(t)I(t) \\ D'(t) = dI(t) \\ R'(t) = rI(t) \\ I'(t) = cS(t)I(t) - (d+r)I(t) \end{cases} . \quad (1)$$

On notera t_0 l'instant de début de la simulation. On suppose $I(t_0) > 0$, $S(t_0) \geq 0$, $R(t_0) \geq 0$, $D(t_0) \geq 0$ et $(S + I + R + D)(t_0) = 1$.

On propose de programmer cette modélisation dans la partie 3. Dans la partie 4 on verra comment déterminer les paramètres c , d , r ci-dessus à partir des données expérimentales. La partie 2 prouve quelques résultats simples qui seront nécessaires à 4.

Les deux dernières parties sont indépendantes. Dans 5, axée sur la programmation, on joue un peu avec le modèle. Quant à 6, elle propose une analyse mathématique du système pendant laquelle on verra dans quel cas se produit le fameux pic épidémique.

2 Un peu de maths

Ce système n'est pas résoluble analytiquement, ce qui signifie qu'on ne sait pas donner une formule utilisant uniquement les fonctions usuelles. Voyons ce que nous pouvons en dire néanmoins.

Pour simplifier les formules, on prend ici $t_0 = 0$.

1. Vérifier que $S + I + D + R$ est constant, et préciser sa valeur.
2. Au début de l'épidémie, on suppose D , I et R négligeables devant S , donc S équivaut à 1.

Calculer I .

3. Montrer que D est proportionnel à R .

Il est donc inutile de calculer D ; c'est pourquoi (et sans doute aussi un peu par pudeur) il est souvent omis dans la littérature.

3 Euler

Pour tout $t \in \mathbb{R}$, on note $X(t) = \begin{pmatrix} S(t) \\ D(t) \\ R(t) \\ I(t) \end{pmatrix}$. Le vecteur $X(t)$ sera enregistré par un tableau numpy dans Python.

On rappelle que la méthode d'Euler est basée sur un développement limité d'ordre 1 : pour tout $t \in \mathbb{R}^+$, on a :

$$X(t + dt) \underset{dt \rightarrow 0}{=} X(t) + dtX'(t) + O(dt^2)$$

Cette formule permet, une fois choisie une valeur pour dt , de calculer par récurrence des valeurs approchées des $(X(t_0 + idt))_{i \in \mathbb{N}}$.

1. Pourquoi le type `ndarray` des tableaux numpy pour enregistrer X ?
2. Écrire une fonction `Xprime` qui prend $X(t)$, c , r , et d et qui renvoie $X'(t)$.
3. En déduire une fonction `euler` pour simuler l'évolution de S, I, R, D .

Attention : la commande `+=` utilisée sur un tableau numpy fera *muter* ce tableau. Ceci peut avoir des conséquences dérangerantes que je vous laisse découvrir.

Le programme `dessin` du fichier `bib_épidémies` permet de tracer les courbes issues de cette modélisation.

4 Détermination des paramètres

4.1 r et d

Les paramètres r et d ne sont pas connus. En revanche, on connaît la durée moyenne de la maladie, ainsi que son taux de mortalité. Voyons comment déduire les premiers des suivants.

1. Notons m le taux de mortalité (qu'on suppose à peu près connu) de la maladie. C'est le rapport à la fin de l'épidémie du nombre de morts sur le nombre total de malades. Donner une formule liant m , r et d .

Indication :

- (a) Montrer que $\frac{D}{R}$ est une constante, et la préciser.
 - (b) Pour calculer m , se placer lorsque l'épidémie est finie, c'est-à-dire lorsque $I \approx 0$. À ce moment, $m = \frac{D}{D+R}$.
2. Supposons ici que $c = 0$: on étudie une situation où aucun nouvel individu est infecté, et on va calculer la durée moyenne de la maladie.
 - (a) Calculer alors I .
 - (b) La durée moyenne de la maladie, notée τ , est $\frac{1}{I(t_0)} \int_{t_0}^{\infty} -tI'(t)dt$. En effet, penser que $-I'(t)dt$ représente le nombre de personnes pour lesquelles la maladie a duré entre t et $t + dt$. Donc $\frac{-I'(t)dt}{I(0)}$ est la probabilité que la maladie dure entre t et $t + dt$. On reconnaît alors un calcul d'espérance.
L'intégrale sur $[t_0, \infty[$ est définie par $\int_{t_0}^{\infty} -tI'(t)dt = \lim_{t_f \rightarrow \infty} \int_{t_0}^{t_f} -tI'(t)dt$.
Calculer τ en fonction de r et d .
 3. Finalement, exprimer r et d en fonction de m et τ . Définir τ et m en variables globales à partir des données que vous aurez trouvées, dans la presse où sur internet. Définir alors r et d en variables globales également.

4.2 Taux de transmission (c)

Le fichier `time_series_19-covid-Confirmed.csv` contient le nombre de personnes détectée positives au COVID-19 chaque jour depuis le 22 janvier 2020 dans chaque pays du monde. Il y a une entrée par jour.

Dans le fichier `bib_épidémie.py`,

- La fonction `extrait_données` récupère les données de la France. Les paramètres `début` et `fin` permettent de sélectionner une période. Ils sont mesurés en nombre de jours à partir du début des mesures, soit du 22 janvier 2020.
 - La fonction `dessins_données_passé` affiche le graphe issu de ces données, dans un repère classique d'abord puis semilog. En pratique, le deuxième graphe affiché est celui de $\ln(I(t))$ en fonction de t .
1. Au vu de l'allure générale des graphes, est-ce que les calculs de la partie 2 sont confirmés ?
 2. Pour une période située dans la première phase de l'épidémie, le deuxième graphe est supposé être une droite ; donner l'équation théorique de cette droite. On prendra $t_0 = 0$.
 3. La fonction `polyfit` de numpy permet d'effectuer une régression linéaire. Si on lui fournit `tx` le tableau des abscisses et `ty` le tableau des ordonnées d'un ensemble de points, et qu'on lui précise qu'on veut une regression de degré 1 en tapant `np.polyfit(tx,ty,1)`, elle renvoie un couple (a, b) tel que la droite d'équation $y = ax + b$ est la droite passant au plus près de ces points.
Utiliser `np.polyfit` pour trouver $I(0)$ ainsi que la constante λ telle que au début de l'épidémie, $I(t) = I(0)e^{\lambda t}$.
 4. En déduire la valeur de c .

4.3 Conclusion

Tracer l'évolution de l'épidémie avec ces valeurs des paramètres.

5 Variantes

Notons I_c la proportion maximale de population infectée avant que les hopitaux ne soient débordés. Mettons que le taux de mortalité est multiplié par deux lorsque $I(t) > I_c$.

Soit t_{fc} l'instant de fin de confinement. Mettons que c est multiplié par deux après t_{fc} .

1. Écrire une nouvelle variante de votre modélisation pour prendre en compte ces données.
2. Déterminer t_{fc} pour que le nombre de morts au final soit le plus petit possible.
3. Déterminer t_c minimal pour que pour tout $t \geq t_c$, $I(t) \leq I_c$.

6 Plus de maths

1. Retrouver l'énoncé précis du théorème de Cauchy-Lipschitz. J'ai notamment besoin d'une hypothèse plus faible que « lipschitzienne ».
2. Montrer que I ne s'annule pas.
Indication : Cauchy-Lipschitz
3. Montrer que s ne s'annule pas.
4. Montrer l'existence de limites en ∞ pour S , I , R , et D .
5. Montrer que I converge vers 0. *Dans ce modèle, la maladie finit toujours par s'éteindre.*

6. Théorème du seuil épidémique

Notons que l'équation donnant I' peut se réécrire ainsi :

$$\forall t \in \mathbb{R}^+, I'(t) = I(t) (cS(t) - (d + r))$$

De sorte que $I'(t) \geq 0 \Leftrightarrow \frac{cS(t)}{d+r} \geq 1$.

Le nombre $\frac{cS(t)}{d+r}$ est appelé le «taux de contagion à l'instant t ».

- (a) Soit t_1 tel que le taux de contagion à l'instant t_1 soit ≤ 1 . Montrer que I décroît à partir de t_1 .
- (b) Supposons le taux de contagion à l'instant t_0 strictement supérieur à 1. Montrer qu'il existe $t_p \in [t_0, \infty[$ tel que I est croissante sur $[t_0, t_p]$ puis décroissante sur $[t_p, \infty[$.

L'instant t_p est appelé l'instant du pic épidémique.

En particulier, dans une situation où ce modèle est valide (notamment il a bien marché pour la variole), si on connaît les paramètres c , r , et d , on peut calculer la proportion de population à vacciner pour éviter le pic épidémique.