



de 7480000 kilomètres de la Terre (soit 1/20 de la distance moyenne Terre-Soleil) sont classés dans la catégorie des « Objets potentiellement dangereux » (OPD). C'est par exemple le cas de l'astéroïde Toutatis, en forme de haricot et d'environ 3 kilomètres de diamètre moyen pour une masse de 50 milliards de tonnes.

UNE MENACE BIEN RÉELLE

Les OPD sont étroitement surveillés depuis la fin des années 1990. La prise de conscience de la menace qu'ils représentent est devenue très concrète lors de la spectaculaire collision des fragments de la comète Shoemaker-Lévy 9 avec la planète Jupiter en juillet 1994, et dont les impacts ont été nettement visibles. Cela s'est aussi accompagné de la reconnaissance que des géocroiseurs pouvaient être des causes possibles de grandes extinctions, alors qu'auparavant, les paléontologues misaient plutôt sur les variations du niveau des mers et les volcans.

Actuellement, plusieurs campagnes d'observation dans les domaines radio et infrarouge ayant pour objectif de classer les OPD (à l'aide de diverses échelles de risque) sont menées. Elles visent au recensement exhaustif des OPD et au suivi de leur trajectoire le plus tôt possible. La mission *Near-Earth Object Surveillance* de la Nasa a pour objectif d'envoyer un télescope infrarouge dans l'espace afin d'effectuer ce recensement en une décennie seulement. Cependant, la trajectoire de certains OPD comme Toutatis est particulièrement difficile à prévoir, en raison de leur déviation à chaque fois qu'ils frôlent la Terre ou d'autres planètes.

Et si l'un de ces objets devait percuter la Terre? Des programmes de « défense planétaire » sont à l'étude et financés par plusieurs agences spatiales internationales. C'est le cas de la mission *Dart* (*Double asteroid redirection test*, Test de redirection d'astéroïde binaire), de la Nasa. Son but: dévier légèrement la trajectoire de la lune qui accompagne un astéroïde, une configuration concernant 15% des astéroïdes qui évoluent dans l'environnement de la Terre. On cible le satellite, car ainsi on ne perturbe pas le système complet (la lune représente moins de 1% de la masse totale du système). En l'occurrence, la cible de *Dart* est Didymoon, lune rocheuse de 160 mètres qui accompagne Didymos (mot signifiant d'ailleurs « jumeau » en grec), un astéroïde de 780 mètres de diamètre. Découvert en 1996 grâce au programme *Spacewatch* pour la détection de géocroiseurs de l'université d'Arizona, Didymos a déjà frôlé la Terre le 13 novembre 2003.

La sonde *Dart*, dont le lancement est prévu pour juillet 2021, se propulsera vers l'astéroïde binaire pendant une année grâce à un moteur ionique alimenté par des panneaux solaires. Son impacteur percutera Didymoon à la vitesse

de 6,6 kilomètres par seconde. En théorie, la collision est censée modifier la vitesse de rotation de Didymoon autour de Didymos de 1%, ce qui est suffisant pour tester cette technique d'impact sans trop perturber le mouvement héliocentrique du centre de masse du système. En complément de *Dart*, l'ESA prépare la mission *Hera*, dont Patrick Michel, astrophysicien du CNRS à l'observatoire de la Côte d'Azur, est le coordinateur scientifique. « *Hera* se placera en orbite autour de Didymos et mesurera la déviation de sa lune due à l'impact. La sonde analysera également le cratère produit, les propriétés physiques et la composition de Didymos », résume le chercheur.

Pour protéger la Terre de ces menaces venues du ciel, d'autres solutions existent sur le papier: dévier le géocroiseur d'une façon plus douce, mais continue, en utilisant l'attraction gravitationnelle d'un engin spatial très massif ou en orientant un de ses moteurs vers l'astéroïde, ou encore en modifiant l'albédo du géocroiseur pendant une longue durée – en utilisant alors la force produite par l'écart entre l'absorption solaire et le rayonnement thermique (c'est l'effet Yarkovsky, qui sera étudié par la sonde *OSIRIS-Rex* sur l'astéroïde Benu).

Mais si le temps manque et que la menace est proche, une solution consisterait aussi à faire exploser une charge nucléaire sur le géocroiseur, ou non loin de sa surface. Tester cette solution de la dernière chance est strictement interdit par les traités internationaux, mais c'est pourtant ce que fait Bruce Willis dans *Armageddon* (film de Michael Bay, 1998) pour sauver le monde d'un astéroïde de la taille du Texas.

DES VISITEURS INTERSTELLAIRES

En 2017, le projet de surveillance *Pan-Starrs* (comparable à *Skywatch*) a détecté un nouvel objet baptisé 'Oumuamua. Son originalité: il ne provient pas du Système solaire. Cet objet interstellaire n'a rendu qu'une brève visite au Soleil et sa cohorte de planètes. La forme allongée de 'Oumuamua n'est pas sans rappeler le vaisseau extraterrestre du roman *Rendez-vous avec Rama* d'Arthur C. Clarke, dont on fêtait le centenaire de la naissance en 2017.

L'astronome Abraham Loeb, de l'université Harvard, a alimenté l'idée de l'origine artificielle de 'Oumuamua. En analysant la trajectoire du bolide, il a suggéré que celui-ci était propulsé par une voile solaire. Si des explications plus classiques sont probablement plus raisonnables (des hétérogénéités dans la composition de l'astéroïde, un jet de dégazage, etc.), la science et la fiction semblent se mêler étroitement. En 2019, la comète Borissov a été le deuxième visiteur interstellaire à être détecté. Ceux-ci pourraient donc être plus fréquents qu'on ne le pense et constituent donc une nouvelle classe d'objets à surveiller de près! ■

BIBLIOGRAPHIE

A. Morbidelli et al., **Debiased albedo distribution for Near Earth Objects**, *Icarus*, vol. 340, 113631, 2020.

D. Lauretta et al., **Episodes of particle ejection from the surface of the active asteroid (101955) Bennu**, *Science*, vol. 366, eaay3544, 2019.

P. Michel et al., **European component of the AIDA mission to a binary asteroid: Characterization and interpretation of the impact of the DART mission**, *Advances in Space Research*, vol. 62, pp. 2261-2272, 2018.

D. Lauretta, **Un vol de sept ans pour quelques grammes d'astéroïde**, *Pour la Science*, n° 468, octobre 2016.

R

ENDEZ-VOUS

P.58 Logique & calcul
P.64 Idées de physique
P.68 Chroniques de l'évolution
P.72 Science & gastronomie
P.74 À picorer

DESSINER UN ÉLÉPHANT AVEC UN SEUL PARAMÈTRE

Reproduire n'importe quel jeu de données en nombre fini, par exemple des points qui esquissent la silhouette d'un animal: c'est l'exploit que permet une fonction paramétrisée par un unique nombre réel.

L'AUTEUR



JEAN-PAUL DELAHAYE
professeur émérite
à l'université de Lille
et chercheur au
laboratoire Cristal
(Centre de recherche
en informatique, signal
et automatique de Lille)



Jean-Paul Delahaye
a notamment publié :
**Les Mathématiciens
se plient au jeu**,
une sélection de ses
chroniques parues
dans *Pour la Science*
(Belin, 2017).

Supposons que l'on dispose de données et que l'on recherche une explication simple à ces données, ou au moins un modèle qui permette de les reproduire. On partira par exemple de valeurs de mesures qui, pour $x = 0, 1, \dots, n$, donnent $y_0, y_1, \dots, y_n$. Si l'on trouve une fonction f qui vérifie $f(0) = y_0, f(1) = y_1, \dots, f(n) = y_n$, on sera assez satisfait. On utilisera f pour prédire $f(n+1), f(n+2), \dots$ et l'on analysera f pour comprendre les données étudiées et peut-être pour formuler une théorie les expliquant.

Bien évidemment, on recherche des fonctions f les plus simples possible. Les polynômes, c'est-à-dire les fonctions de la forme $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ sont parmi les plus simples. On sera donc assez satisfait si l'on trouve un polynôme qui reproduit $y_0, y_1, \dots, y_n$.

Or le monde mathématique est généreux et permet souvent de faire tout ce que l'on veut. Un joli résultat, dénommé «théorème d'interpolation de Lagrange», indique que quelles que soient les données $y_0, y_1, \dots, y_n$, il existe un polynôme de degré au plus n qui redonne $y_0, y_1, \dots, y_n$ pour $x = 0, 1, \dots, n$ (voir l'encadré 2). Mais ce polynôme est un surajustement (ou *overfitting*) : on se donne trop de paramètres, ici les coefficients du polynôme, et ce que l'on trouve n'a pas vraiment d'intérêt car on était certain d'y arriver !

Aussi le calcul du polynôme d'interpolation de Lagrange ne permettra que très rarement d'anticiper ce qui se passe pour $n+1, n+2, \dots$, et ce polynôme ne sera pas utile sur un plan

théorique. Heureusement, il se peut que le polynôme trouvé soit de faible degré. Si c'est le cas, cela signifie que sa découverte aurait résulté de la prise en compte de moins de données (autant que le degré du polynôme, plus une unité) et donc que le polynôme retrouve les autres données. Dans une telle situation, il ne sera pas absurde de croire que le polynôme calculé permet d'anticiper $f(n+1), f(n+2), \dots$.

Le risque de surajustement est un problème important dans le domaine de l'apprentissage profond en intelligence artificielle, car il conduit à des réseaux de neurones qui sont bien adaptés à un grand ensemble de données... et à rien d'autre.

À propos du choix des paramètres dans un modèle, le mathématicien d'origine hongroise John von Neumann a fait une remarque amusante : «Avec quatre paramètres, je peux faire un éléphant, et avec cinq, je peux faire remuer sa trompe.»

UN SEUL PARAMÈTRE SUFFIT POUR TOUT AJUSTEMENT

Un article récent de Steven Piantadosi, du Département de psychologie de l'université de Californie à Berkeley, a montré que von Neumann aurait pu remplacer le nombre «quatre» de sa phrase par le nombre «un». L'article est bien sûr un jeu, mais ce qu'il montre, en mettant en œuvre des mathématiques assez simples, est remarquable et doit servir de mise en garde : un seul paramètre permet d'ajuster une fonction pour qu'elle redonne mille données ou un milliard si l'on s'y prend bien.

UN SEUL PARAMÈTRE POUR L'ÉLÉPHANT

1

Steven Piantadosi a proposé la fonction : $f_p(x) = \sin^2(2^{kx} \arcsin(\sqrt{p}))$, où le nombre réel p est un paramètre, et a montré qu'elle présente une extraordinaire propriété d'ajustement. La valeur de k est fixée une fois pour toutes, et indique la précision avec laquelle on approchera les données. On considère par exemple des valeurs $y_0, y_1, \dots, y_n$ telles que les points de coordonnées $(0, y_0), (1, y_1), \dots, (n, y_n)$ dessinent un éléphant, en serrant assez les abscisses et en dessinant des points assez gros.

La méthode de Steven Piantadosi (expliquée dans

les encadrés 3 et 4) permet alors de déterminer la valeur du paramètre p pour que : $f_p(0) = y_0, f_p(1) = y_1, \dots, f_p(n) = y_n$, autrement dit pour que le graphe de $f_p(x)$ dessine un éléphant. Un seul paramètre suffit pour avoir un éléphant ! Bien sûr, on peut obtenir tous les dessins qu'on veut en changeant la valeur de p .

Laurent Boué a écrit des programmes disponibles sur internet qui permettent le calcul de p pour les dessins souhaités et a fait les calculs pour un oiseau et une tortue (voir la bibliographie).

L'énoncé démontré par Steven Piantadosi est le suivant.

- On se fixe une précision ϵ , par exemple $1/1000$, ce qui signifie que l'on recherchera à avoir des égalités avec une erreur inférieure à ϵ .

- Il existe une fonction $f_p(x)$ dépendant d'un paramètre réel p , telle que, quelles que soient les données $y_0, y_1, \dots, y_n$ en nombre fini, il existe une valeur p du paramètre pour laquelle $f_p(0) = y_0, f_p(1) = y_1, \dots, f_p(n) = y_n$, avec une erreur inférieure à ϵ pour chaque égalité.

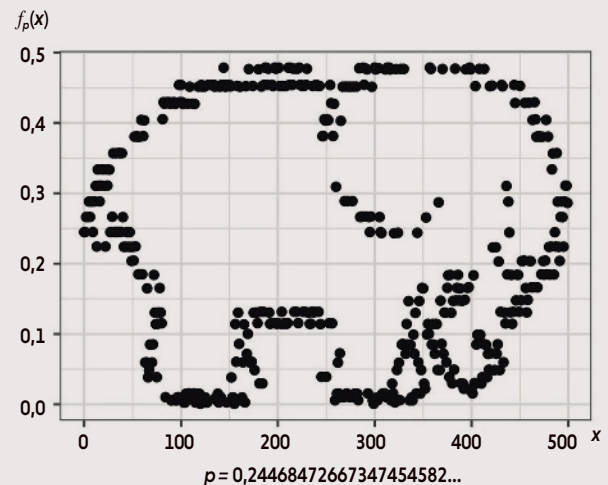
En clair, en faisant varier le paramètre p , la fonction $f_p(x)$ s'ajuste avec une précision ϵ à tout jeu de données fixé à l'avance, aussi grand soit-il.

La fonction miraculeuse de Steven Piantadosi est : $f_p(x) = \sin^2(2^{kx} \arcsin(\sqrt{p}))$. Rappelons que $\arcsin(x)$ est la fonction inverse de $\sin(x)$, ce qui signifie que pour tout x compris entre $-\pi/2$ et $\pi/2, y = \sin(x) \Leftrightarrow x = \arcsin(y)$.

La valeur k qui apparaît dépend de la précision ϵ souhaitée, mais une fois ϵ fixée, k l'est aussi. Bien évidemment, le chercheur a calculé le paramètre p qui donne un éléphant. L'encadré 1 montre le résultat.

D'OÙ VIENT L'EXPLOIT ?

Comprendre en détail comment la fonction $f_p(x)$ réalise cet exploit, et comment s'y prendre pour trouver les bonnes valeurs du paramètre p , est un jeu mathématique qui exige un peu de calcul (voir les encadrés 3 et 4). Cependant, l'idée générale est simple : on va cacher dans le paramètre p qui, parce qu'il est un nombre réel, peut contenir une infinité d'informations, les nombres $y_0, y_1, \dots, y_n$. Ensuite, la



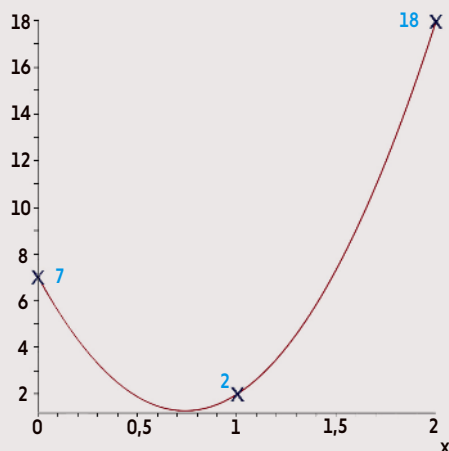
$p = 0,74933466...$



$p = 0,70704013...$

2

L'INTERPOLATION



Imaginons que l'on veuille trouver un polynôme $P(x)$ qui prenne la valeur 7 pour $x = 0$, la valeur 2 pour $x = 1$, et la valeur 18 pour $x = 2$. Voici une méthode bien connue qui permet de le construire et qui se généralise à un nombre quelconque de données.

$$\text{On écrit : } P(x) = 7(x-1)(x-2)/[(0-1)(0-2)] + 2(x-0)(x-2)/[(1-0)(1-2)] + 18(x-0)(x-1)/[(2-0)(2-1)].$$

Le premier terme est nul pour $x = 1$ et $x = 2$, le second pour $x = 0$ et $x = 2$, le troisième pour $x = 0$ et $x = 1$. Donc, seul le premier terme n'est pas nul pour $x = 0$, seul le second n'est pas nul pour $x = 1$, seul le troisième n'est pas nul pour $x = 2$.

De plus, les facteurs 7, 2 ou 18 au début de chaque terme assurent que le terme non nul pour $x = 0$ prenne la valeur 7, pour $x = 1$ prenne la valeur 2, pour $x = 2$ prenne la valeur 18. C'est ce que l'on recherchait. En développant et en simplifiant, on trouve :

$$P(x) = 21x^2/2 - 31x/2 + 7.$$

On vérifie aisément que $P(0) = 7$, $P(1) = 2$, $P(2) = 18$. La méthode se généralise et permet, si $y_0, y_1, \dots, y_n$ sont des valeurs données, de construire sans mal un polynôme qui prendra la valeur y_0 pour $x = 0$, y_1 pour $x = 1$, ..., y_n pour $x = n$. Telle est la « méthode d'interpolation de Lagrange ».

Elle montre qu'autoriser beaucoup de paramètres, ici les coefficients du polynôme, permet sans mal de faire passer un polynôme par les valeurs souhaitées, quelles qu'elles soient.

Cependant, en général, la fonction trouvée oscillera beaucoup, alors qu'on préférerait une fonction aussi régulière et stable que possible : son ajustement aux valeurs voulues est exact, mais insatisfaisant. Il existe de meilleures méthodes permettant d'approcher les valeurs voulues en limitant les oscillations de la fonction résultat.

Retenons qu'interpolier exactement est facile, mais que cela ne satisfait pas en général le désir de trouver des lois lisses et régulières.

Cette logique de l'interpolation de Lagrange est poussée jusqu'à l'absurde dans les résultats présentés dans le texte principal : la méthode découverte par Steven Piantadosi, bien plus puissante que l'interpolation de Lagrange, permet, en changeant un seul paramètre p d'une fonction $f_p(x)$, d'obtenir que pour des valeurs quelconques $y_0, y_1, \dots, y_n$, on ait :

$$f_p(0) = y_0, f_p(1) = y_1, \dots, f_p(n) = y_n.$$

> fonction $f_p(x)$ pour $x = 0, 1, \dots, n$ les fera remonter une à une : quand x vaudra 0, y_0 remontera, quand x vaudra 1, y_1 remontera, etc.

L'encadré 3 explique que cette remontée des valeurs approchées de $y_0, y_1, \dots, y_n$ s'opère indirectement grâce à la fonction $2x$. Quand on applique la fonction $2x$ à un nombre réel écrit en base 2, par exemple $x = 0,110011101111_2$, elle fait remonter tous les chiffres vers la gauche : $2x = 1,10011101111_2$ (le 2 en indice indique qu'il s'agit d'une écriture en base 2).

Cette propriété est équivalente, pour la base 2, à ce que l'on connaît bien en base 10 : multiplier par 10 décale toutes les décimales d'un chiffre vers la gauche. Par exemple, $10 \times 3,1415926\dots = 31,415926\dots$

En utilisant la fonction qui enlève ce qui est devant la virgule, la fonction $(x \bmod 1)$, parfois appelée « partie fractionnaire de x », on a alors un moyen de faire remonter les chiffres binaires d'un nombre x donné en effaçant ceux qui sont placés en tête. Posons en effet $D(x) = 2x \bmod 1$, et prenons par exemple $x = 0,1100111011110101_2$. On a alors :

$$D(x) = 0,100111011110101_2,$$

$$D(D(x)) = 0,00111011110101_2,$$

$$D(D(D(x))) = 0,0111011110101_2, \text{ etc.}$$

En exploitant ce moyen de cacher des informations dans un paramètre réel p , ainsi que quelques autres astuces (voir l'encadré 4), on arrive à la fonction qui, avec le bon choix de p , donne y_0 pour $x = 0$, y_1 pour $x = 1$, ..., y_n pour $x = n$ avec une précision fixée à l'avance. Cela montre qu'un seul paramètre p et une seule fonction $f_p(x)$ permettent d'approcher toutes les données que l'on veut, quel que soit leur nombre, pourvu qu'il soit fini !

Ainsi, pas besoin de quatre paramètres pour dessiner un éléphant : avec un paramètre unique, on peut s'arranger, et la même fonction servira toujours, en changeant p , pour dessiner un poisson, une tortue, un oiseau ou tout ce que l'on veut.

NOMBRE DE PARAMÈTRES ET COMPLEXITÉ DES MODÈLES

On mesure parfois la complexité d'un modèle en comptant le nombre de paramètres utilisés par le modèle. C'est une façon de procéder qui ne peut pas se justifier sans précisions supplémentaires sur la nature du modèle. En effet, la fonction $f_p(x)$ de Steven Piantadosi montre que l'on peut toujours se ramener à un seul paramètre. Si l'on veut donner un sens au décompte des paramètres d'un modèle, il faut détailler les contraintes qu'on impose au modèle, par exemple « ne faire intervenir que des polynômes » ce qui interdira d'accepter la fonction $f_p(x)$ ou l'une de ses variantes.

Il faut insister sur l'idée que l'article de Steven Piantadosi ne prétend nullement

proposer une méthode nouvelle pour construire des modèles approchant des données en utilisant une fonction à un seul paramètre (voir l'encadré 5). L'article est seulement un jeu mathématique qui exploite la capacité qu'a un nombre réel de détenir une quantité infinie d'information. Ce qui fait l'intérêt particulier de l'article et le rend étonnant est la simplicité globale de la méthode. Je ne crois pas que l'on ait publié antérieurement au travail de Steven Piantadosi une méthode aussi élémentaire permettant de déterminer et calculer effectivement un paramètre p qui accumule des informations en nombre illimité, et de faire remonter ce qu'on a caché dans p en n'utilisant qu'une seule fonction construite à partir des fonctions élémentaires de l'analyse classique (sin, arccsin, exponentielle, etc.).

Les physiciens utilisent des nombres réels, mais sans jamais prendre en compte plus de 20 ou 30 décimales (voir l'article de cette rubrique de décembre 2019). Aussi ne donneront-ils aucun sens à des modèles comme celui suggéré par la fonction prodigieuse de Steven Piantadosi, qui ne fonctionne que si l'on accepte de manipuler des centaines de chiffres significatifs.

Si l'on examine de près la fonction $f_p(x)$ quand le paramètre p a été fixé, on s'aperçoit qu'elle oscille violemment et que les valeurs qu'elle prend aux nombres entiers $0, 1, \dots, n$ ne sont en rien représentatives de ce qui se passe pour les x compris entre 0 et 1, entre 1 et 2, entre 2 et 3, etc.

Si l'on dessine (autant qu'on peut le faire) tout le tracé de la fonction $f_p(x)$ entre 0 et n , on ne voit pas du tout l'image que produit le dessin quand on ne tient que les valeurs de $f_p(x)$ pour $x = 0, 1, \dots, n$. On voit un zigzag serré qui noircit presque tout le rectangle $[0, n] \times [0, 1]$. Le tracé de la fonction $f_p(x)$ passe bien tout près des points de coordonnées $(0, y_0), (1, y_1), \dots, (n, y_n)$, mais il le fait sans limiter ses variations, alors que c'est ce qu'on attend d'une bonne fonction d'interpolation: on désire cette dernière aussi lisse que possible.

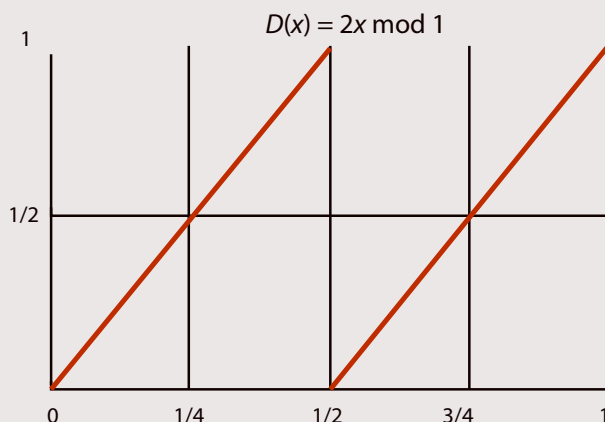
Citons les commentaires de Steven Piantadosi concernant sa fonction miraculeuse:

L'existence d'une fonction aussi simple avec une telle liberté de comportement illustre un problème plus fondamental. Elle montre que la complexité d'un modèle ne peut pas être évaluée en comptant ses paramètres.

Plus généralement, l'utilisation non critique d'une approche par dénombrement de paramètres ignore qu'un seul paramètre à valeur réelle contient potentiellement une quantité illimitée d'informations puisque, pour spécifier entièrement un nombre réel, un nombre infini de bits doit l'être, ce qui d'ailleurs est en soi problématique. Un autre fait notable est que l'ensemble des nombres réels

3

UNE PREMIÈRE FONCTION MIRACULEUSE



Soit D la fonction telle que $D(x) = 2x$ pour $0 \leq x < 1/2$ et $D(x) = 2x - 1$ pour $1/2 \leq x < 1$. Cela peut écrire : $D(x) = 2x \bmod 1$.

La fonction $x \bmod 1$, parfois appelée partie fractionnaire de x , ne garde que ce qui est après la virgule : 2,37 devient 0,37 ; 219,2222 devient 0,2222.

Appliquée à un nombre écrit en binaire, par exemple $x = 0,110111101_2$, la fonction $D(x)$ donne un nombre ayant le même développement binaire que x , sauf que le premier chiffre après la virgule a disparu :

$D(0,110111101_2) = 0,10111101_2$.

L'explication est simple. D'une part, $2x$ décale le développement binaire d'un chiffre vers la gauche (c'est comme en base 10 quand on multiplie un nombre par 10), et donc $0,110111101_2$ devient $1,10111101_2$; d'autre part, $2x \bmod 1$ enlève le 1 devant la virgule s'il y en a un.

Considérons maintenant la fonction D^k obtenue en appliquant k fois de suite la fonction D :

$D^k(x) = D(D(\dots(D(x)\dots)))$, k fois.

La fonction $D^k(x)$ fait glisser les chiffres binaires de x de k places vers la gauche et élimine les k premiers chiffres :

$D^k(x) = 2^k x \bmod 1$. Par exemple, $D^4(0,110111101_2) = 0,11101_2$.

Soient des nombres $y_0, y_1, \dots, y_n$ compris entre 0 et 1, chacun écrit en base 2 avec r chiffres après la virgule. Exemple pour $r = 5$ et $n = 2$:

$y_0 = 0,10111_2$; $y_1 = 0,11001_2$;

$y_2 = 0,01011_2$.

Soit $p = 0,101111100101011_2$ obtenu en mettant les uns derrière les autres les chiffres de y_0, y_1 et y_2 .

On a :

$p = D^0(p) = 0,101111100101011_2$

$D^1(p) = 0,01111100101011_2$

$D^2(p) = 0,1111100101011_2$

$D^3(p) = 0,11100101011_2$

$D^4(p) = 0,1100101011_2$

$D^5(p) = 0,1100101011_2$, etc.

Ainsi, $D^0(p)$ correspond à y_0 sur les 5 premiers chiffres binaires, $D^1(p)$ correspond à y_1 sur les 5 premiers chiffres binaires, $D^{10}(p)$ correspond à y_2 sur les 5 premiers chiffres binaires, etc.

Ce procédé place donc dans les chiffres binaires de p les chiffres binaires des données $y_0, y_1, \dots, y_n$ avec une précision de r chiffres binaires pour chacun, ce qui permet, en appliquant r fois de suite la fonction D , d'avoir successivement $y_0, y_1, \dots, y_n$, à chaque fois avec une précision de r chiffres binaires.

En résumé, p contient, cachés dans ses chiffres binaires, les nombres $y_0, y_1, \dots, y_n$ avec une précision de r chiffres binaires pour chacun, et la fonction D , appliquée de manière répétée à p , retrouve ces nombres cachés.

Soit maintenant la fonction E_p définie par $E_p(x) = D^{rx}(p)$, dont la variable est x et qui dépend du paramètre p . On a :

$E_p(0) = y_0$ avec une précision

de r chiffres binaires,

$E_p(1) = y_1$ avec une précision

de r chiffres binaires,...

$E_p(n) = y_n$ avec une précision

de r chiffres binaires.

En choisissant bien le paramètre p , la fonction $E_p(x)$ est donc capable de redonner les nombres $y_0, y_1, \dots, y_n$ avec une précision de r chiffres binaires. Si les données ne sont pas comprises entre 0 et 1, on s'y ramène en divisant tout par un nombre assez grand.

Ainsi, en ajustant bien son paramètre p , la fonction $E_p(x)$ suffit pour retrouver tout jeu fini de données. Ce n'est pas le résultat annoncé concernant la fonction $f_p(x) = \sin^2(2^{1/x} \arcsin(\sqrt{p}))$, mais nous y sommes presque. L'encadré 4 explique comment y parvenir.

LA FONCTION DE PIANTADOSI

4

Compte tenu de ce qui a été dit dans l'encadré 3 sur la fonction D , on considère à présent les trois fonctions :

$$\phi(x) = \sin^2(2\pi x);$$

$$\phi^{-1}(x) = (\arcsin \sqrt{x})/2\pi;$$

$$L(x) = 4x(1-x).$$

Par définition, $\phi(\phi^{-1}(x)) = x$ pour $0 \leq x \leq 1$ et $\phi^{-1}(\phi(x)) = x$ pour $0 \leq x \leq 1/4$.

Voici des calculs inévitables pour arriver au miracle.

$$\phi(D(x)) = \phi(2x) = \sin^2(4\pi x) =$$

$$4 \sin^2(2\pi x) (1 - \sin^2(2\pi x)) =$$

$$L(\phi(x)) \text{ (on a utilisé)}$$

$$\sin(2a) = 2 \sin a \cos a,$$

$$\cos^2 a = 1 - \sin^2 a$$

$$\text{et } \sin(x + 2\pi) = \sin(x).$$

$$\text{En prenant } x = \phi^{-1}(p),$$

$$\text{soit } p = \phi(x), \text{ on a :}$$

$$\phi(D(\phi^{-1}(p))) = L(p),$$

$$\text{d'où l'on déduit :}$$

$$\phi(D^k(\phi^{-1}(p))) = L^k(p).$$

Autrement dit, pour connaître l'effet de L appliqué k fois à p , il suffit d'appliquer ϕ^{-1} à p , puis d'appliquer k fois D , puis d'appliquer ϕ au résultat.

Nous allons maintenant comprendre quel paramètre p choisir pour que $f_p(x)$ donne $y_0, y_1, \dots, y_n$ pour $x = 0, 1, \dots, n$. On note d'abord que $f_p(x) = \sin^2(2^{kx} \arcsin(\sqrt{p})) = \phi(D^{kx}(\phi^{-1}(p))) = L^{kx}(p)$. Si l'on veut que $f_p(x)$ donne approximativement $y_0, y_1, \dots, y_n$ pour $x = 0, 1, \dots, n$, il suffit donc que $D^{kx}(\phi^{-1}(p))$ donne approximativement $\phi^{-1}(y_0),$

$\phi^{-1}(y_1), \dots, \phi^{-1}(y_n)$ pour $x = 0, 1, \dots, n$. D'après l'encadré 3, il suffit donc que $\phi^{-1}(p)$ soit fait de k chiffres binaires de $\phi^{-1}(y_0)$, suivi de k chiffres binaires de $\phi^{-1}(y_1)$, ..., suivi de k chiffres binaires de $\phi^{-1}(y_n)$.

Conclusion : le paramètre p que nous recherchons doit être l'image par ϕ de k chiffres binaires de $\phi^{-1}(y_0)$, suivi de k chiffres binaires de $\phi^{-1}(y_1)$, ..., suivi de k chiffres binaires de $\phi^{-1}(y_n)$. Plus k sera grand, plus les y_i seront donnés avec précision par $f_p(0), f_p(1), \dots, f_p(n)$.

Prenons par exemple $y_0 = 0,7, y_1 = 0,4, y_2 = 0,5$ et $k = 12$. On calcule $\phi^{-1}(0,7)$, $\phi^{-1}(0,4)$ et $\phi^{-1}(0,5)$ avec 12 chiffres binaires significatifs :

$$\phi^{-1}(0,7) = 0,001010000110_2;$$

$$\phi^{-1}(0,4) = 0,000110111110_2;$$

$$\phi^{-1}(0,5) = 0,001000000000_2.$$

On met les chiffres binaires trouvés les uns derrière les autres :

$$q = 0,0010100001100001101111000100000000_2 =$$

$$0,157741434872150421142578.$$

On calcule $p = \phi(q) = 0,6999652447...$ et on a bien ce qui est attendu :

$$f_p(0) = 0,6999652447,$$

$$\text{soit environ } 0,7,$$

$$f_p(1) = 0,3996355201,$$

$$\text{soit environ } 0,4,$$

$$f_p(2) = 0,4999999999,$$

$$\text{soit environ } 0,5.$$

qui peuvent être décrits avec un nombre fini de bits (par exemple en les faisant produire par une machine de Turing) est dénombrable. Cet ensemble de nombres réels a donc une mesure nulle dans l'ensemble de tous les nombres réels. On sait aussi qu'il existe des fonctions injectives (c'est-à-dire ne prenant pas deux fois la même valeur) de $\mathbb{R}^n$ (l'espace de dimension n) vers $\mathbb{R}$. Cela confirme que le nombre de paramètres dans un modèle ne peut pas être une mesure raisonnable de sa complexité.

UNIVERSALITÉ

La fonction $f_p(x) = \sin^2(2^{kx} \arcsin(\sqrt{p}))$ peut être qualifiée d'universelle puisqu'elle s'adapte à toute série de données $y_0, y_1, \dots, y_n$. Ce n'est pas le premier objet universel que l'on rencontre en mathématiques.

Le plus simple est sans doute le nombre π . On conjecture que c'est un nombre-univers et le calcul de plusieurs milliers de milliards de ses décimales le confirme: chaque séquence possible de chiffres décimaux (par exemple 243, 21145 ou 929292), aussi longue qu'on veut, apparaîtrait ainsi dans son développement décimal. Si π est réellement un nombre-univers, on peut le voir comme une fonction qui, quand on fixe un entier n , affiche par exemple 1 million de chiffres décimaux à partir de la n -ième décimale. Avec ce million de chiffres, on peut dessiner une image de 1000×1000 pixels avec 10 niveaux de gris. Si π est un nombre univers, alors pour chaque image de ce format, votre portrait par exemple, il existe au moins un entier n tel que l'image numéro n tirée de π la produise. En ce sens, π serait au moins aussi universel que $f_p(x)$, puisque capable de produire toutes les images du format indiqué.

Cependant, la fonction prodigieuse de Piantadosi a un double avantage sur π :

(a) on est certain qu'elle est universelle, ce n'est pas une conjecture;

(b) on sait calculer la valeur du paramètre p nécessaire pour obtenir des données $y_0, y_1, \dots, y_n$, et ce calcul est faisable, ce qui n'est pas le cas pour le n donnant votre portrait dans π .

UN AUTRE OBJET UNIVERSEL: LA FONCTION ZÊTA DE RIEMANN

Un autre objet universel totalement inattendu a été rencontré en mathématiques en 1975 par le mathématicien russe Sergei Voronin (voir la bibliographie). Il concerne la célèbre fonction zêta de Riemann.

Cette fonction notée $\zeta(s)$ est définie, pour tout nombre réel $s > 1$, par la formule:

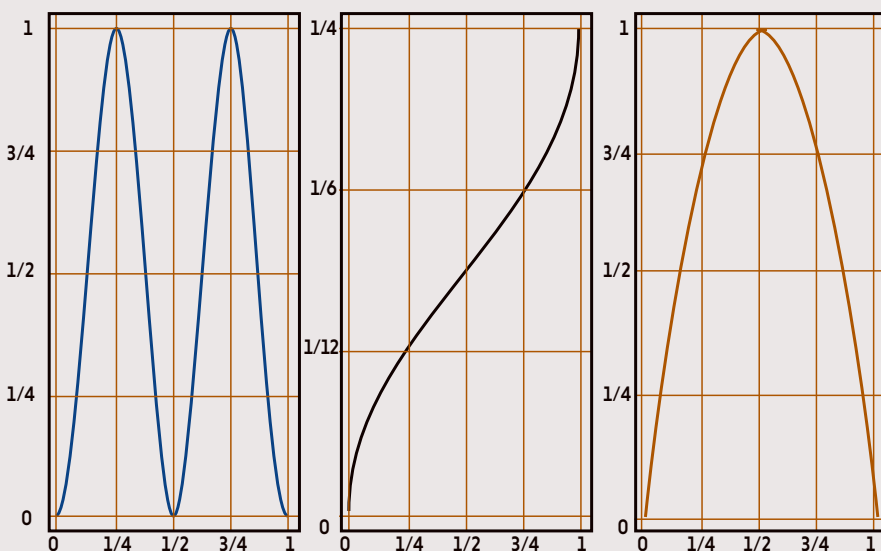
$$\zeta(s) = 1 + 1/2^s + 1/3^s + 1/4^s + \dots$$

On peut lui donner un sens pour toutes les valeurs de s , y compris quand s est un nombre complexe (sauf en $s=1$). On peut la voir

$$\phi(x) = \sin^2(2\pi x)$$

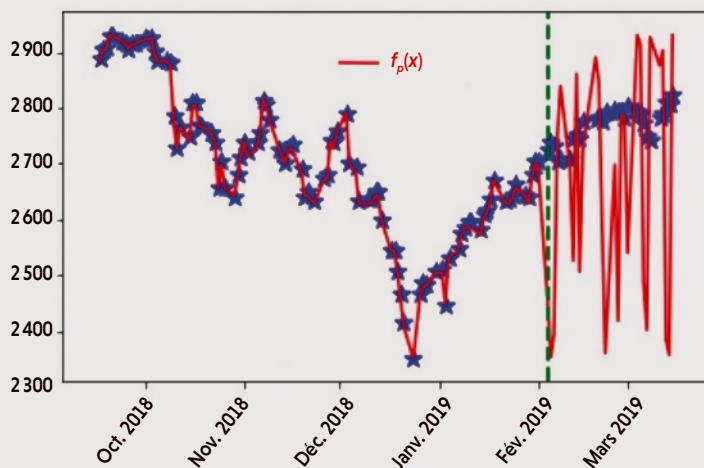
$$\phi^{-1}(x) = (\arcsin \sqrt{x})/2\pi$$

$$L(x) = 4x(1-x)$$



L'ILLUSION DÉÇUE DU SURAJUSTEMENT

5



Laurent Boué a pris 95 valeurs de l'indice boursier américain S&P 500. Il a calculé la valeur de p qui permet à la fonction miraculeuse de Piantadosi, $f_p(x) = \sin^2(2^{kx} \arcsin(\sqrt{p}))$, de redonner ces valeurs. Il a ensuite regardé ce que cela donnait pour les valeurs suivantes de l'indice (au-delà de la barre verticale pointillée verte).

Si la méthode d'approximation par $f_p(x)$ était une bonne façon de découvrir une loi cachée dans les

variations de cet indice boursier, les valeurs suivantes pour le p calculé auraient été proches de celles des données réelles.

Sans surprise, ce n'est pas du tout le cas. Sur le dessin, la courbe rouge est ce que donne $f_p(x)$ et les points bleus sont les valeurs réelles. La fonction $f_p(x)$ permet bien, en changeant p , de s'ajuster à tout ensemble de données, mais ce n'est en rien un moyen pour découvrir des lois cachées. C'est un exemple parfait de surajustement inutile.

comme une fonction qui, à tout point du plan (les nombres complexes sont assimilables à des points du plan), associe un autre point du plan. Cette fonction joue un rôle central en théorie des nombres et, en particulier, la position des valeurs de s qui l'annulent est liée à la répartition des nombres premiers.

Quand on considère des fonctions de ce type, qui produisent des nombres complexes à partir de nombres complexes, les fonctions jouant le rôle des fonctions dérivables se nomment les «fonctions holomorphes». Les fonctions usuelles comme x^n , $\sin(x)$, $\cos(x)$, $\log(x)$ et toutes celles que l'on obtient en les combinant sont bien sûr holomorphes.

Dans la partie du plan correspondant à une bande verticale B des points de coordonnées (x, y) avec $1/2 < x < 1$, Sergei Voronin a démontré une incroyable propriété d'universalité de $\zeta(s)$. Pour toute partie P de cette bande sans trou et compacte (c'est-à-dire de taille finie, et qui contient la limite de toutes ses suites convergentes), et pour toute fonction holomorphe f qui ne s'annule pas dans P , il existe un nombre réel r tel que $\zeta(s+r)$ prend les mêmes valeurs dans P que f , à ϵ près (ϵ , comme précédemment, est une précision que l'on fixe une fois pour toutes).

Dit autrement : quelle que soit la fonction holomorphe f définie sur une zone P sans trou de la bande B , la fonction ζ prend presque les mêmes valeurs que f quand on examine ses valeurs dans une version décalée vers le haut de P . C'est un peu comme si $\zeta(s)$, à elle seule, pouvait imiter toutes les fonctions holomorphes ! Dans π , il y a sans doute toutes les suites finies de décimales, mais dans $\zeta(s)$, on est certain qu'il y a toutes les fonctions holomorphes ! Précisons que le résultat de Sergei Voronin a été étendu à d'autres fonctions que la fonction $\zeta(s)$.

Un autre objet universel fascinant a déjà été évoqué dans cette rubrique, le graphe universel de Rado (voir *Pour la Science* de novembre 2018). Ce graphe infini contient en effet comme sous-graphes tous les graphes finis, et tous les graphes infinis ; de plus, il a la propriété que quand on en enlève une partie finie, ce qui reste est identique à ce dont on est parti.

Le monde infini des mathématiques recèle des merveilles et procure des étonnements bien au-delà, du moins c'est mon point de vue de mathématicien, de tout ce qu'on peut trouver dans le monde physique qui, en comparaison, apparaît parfois terne et ennuyeux ! ■

BIBLIOGRAPHIE

Interpolation lagrangienne, Wikipédia, 2020.

Zeta function universality, Wikipedia, 2020.

L. Boué, *Real numbers, data science and chaos : How to fit any dataset with a single parameter*, prépublication arXiv:1904.12320, 2019.

L. Boué, *Programmes pour «Real numbers, data science and chaos: How to fit any dataset with a single parameter»*, 2019 : <https://github.com/Ranlot/single-parameter-fit>

C. Zhang et al., *A study on overfitting in deep reinforcement learning*, prépublication arXiv:1804.06893, 2018.

S. T. Piantadosi, *One parameter is always enough*, *AIP Advances*, vol. 8(9), 095118, 2018 (<https://aip.scitation.org/doi/pdf/10.1063/1.5031956>).

J. Mayer et al., *Drawing an elephant with four complex parameters*, *American Journal of Physics*, vol. 78, pp. 648-649, 2010.

S. Voronin, *Theorem on the "universality" of the Riemann zeta-Function*, *Math. USSR Izv.*, vol. 9, pp. 443-445, 1975.

LES AUTEURS



JEAN-MICHEL COURTY et ÉDOUARD KIERLIK
professeurs de physique à Sorbonne Université, à Paris

LES MASQUES DÉMASQUÉS

En pleine pandémie de Covid-19, l'intérêt pour les masques respiratoires a bondi, alors même qu'ils sont en usage depuis longtemps. Mais comment fonctionnent-ils ?

Norme FFP2, masque chirurgical ou encore masque fabriqué maison... Ces mots ont envahi notre quotidien avec la pandémie de Covid-19. En tant que physiciens, il ne nous appartient pas de donner des recommandations sur l'usage de ces protections pour minimiser les risques de contamination. En revanche, nous pouvons nous intéresser aux mécanismes physiques mis en œuvre dans leur fonctionnement. On pourrait croire que les masques agissent, pour l'essentiel, comme des tamis. Or ce n'est pas du tout le cas !

FILTRER LES AÉROSOLS

En toussant, en éternuant, en parlant ou tout simplement en respirant, nous produisons des aérosols : autrement dit, nous émettons dans l'air expiré des particules de tailles diverses. Ce sont typiquement des gouttelettes d'eau de diamètre compris entre 1 micromètre (un millième de millimètre) et 100 micromètres, qui s'évaporent rapidement et libèrent dans l'air bactéries (0,5 à 5 micromètres) et virus (0,02 à 0,3 micromètre,

0,1 micromètre pour le virus SARS-CoV-2, responsable du Covid-19).

Les particules les plus grosses tombent rapidement au sol. Les plus légères, en revanche, restent en suspension. Dans de l'air calme, la durée de sédimentation sur une hauteur de 3 mètres est en effet de 4 minutes pour des particules de 20 micromètres de diamètre, et cette durée est multipliée par 4 à chaque fois que l'on divise la taille par 2 : elle dépasse déjà 1 heure pour des particules de 5 micromètres.

En situation épidémique, on peut filtrer l'air à l'aide d'un masque pour protéger notre système respiratoire de ces aérosols chargés de pathogènes. Le premier effet filtrant auquel on songe est l'effet de tamis : comme dans une passoire, ne pourraient traverser le masque que les particules dont la taille est inférieure à celle des trous. Il y a cependant une difficulté majeure : plus les trous sont petits, mieux on filtre mais plus il est difficile de respirer à travers. Si l'on veut que le masque soit utilisable sans assistance mécanique, un filtre à trous submicrométriques, qui seraient nécessaires pour tamiser bactéries et virus, est exclu.

Heureusement, d'autres mécanismes sont à l'œuvre et permettent de piéger des particules de toutes tailles dans le masque. Ce dernier comporte en général une couche, un peu épaisse, composée de fibres non tissées, mais plutôt entrelacées (voir l'encadré page ci-contre). Lorsqu'une particule, entraînée par l'air qui traverse le filtre, vient percuter une fibre, elle s'y colle durablement grâce aux forces intermoléculaires dites « de Van der Waals ». Il faut donc comprendre d'où résultent les collisions particule-fibre.

Aux échelles considérées, on montre que l'écoulement de l'air est dominé par les effets de viscosité et qu'il est

