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, arcsin, 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, ..., 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.

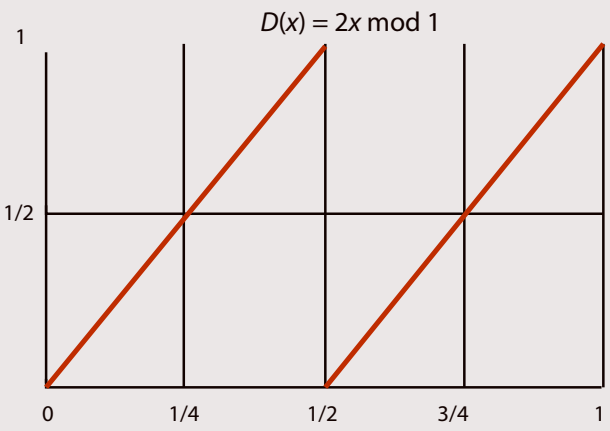
Si l'on dessine (autant qu'on peut le faire) tout le tracé de la fonction $f_p(x)$ entre 0 et n, on ne voit pas du tout l'image que produit le dessin quand on ne retient que les valeurs de $f_p(x)$ pour $x = 0, 1, ..., n$. On voit un zigzag serré qui noircit presque tout le rectangle $[0, n] \times [0, 1]$. Le tracé de la fonction $f_p(x)$ passe bien tout près des points de coordonnées $(0, y_0)$, $(1, y_1)$, ..., (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 mod 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 $x = 0{,}11011110_2$, la fonction $D(x)$ donne un nombre ayant le même développement binaire que x, sauf que le premier chiffre après la virgule a disparu : $D(0{,}11011110_2) = 0{,}1011110_2$. 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 $0{,}11011110_2$ devient $1{,}1011110_2$; d'autre part, $2x$ mod 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 : $D^k(x) = D(D(...(D(x)...))$, 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, $D^4(0{,}11011110_2) = 0{,}1110_2$.

Soient des nombres $y_0, y_1, ..., y_n$ 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 $p = 0{,}1011111001010111_2$ obtenu en mettant les uns derrière les autres les chiffres de y_0, y_1 et y_2. On a :
$p = D^0(p) = 0{,}1011111001010111_2$
$D^1(p) = 0{,}0111110010101_2$
$D^2(p) = 0{,}111110010101_2$
$D^3(p) = 0{,}11110010101_2$
$D^4(p) = 0{,}1110010101_2$
$D^5(p) = 0{,}1100101011_2$, etc.
Ainsi, $D^0(p)$ correspond à y_0 sur les 5 premiers chiffres binaires, $D^5(p)$ 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,..., 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,..., y_n, à chaque fois avec une précision de r chiffres binaires.

En résumé, p contient, cachés dans ses chiffres binaires, les nombres $y_0, y_1, ..., y_n$ 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 $y_0, y_1,..., y_n$ 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 $f_p(x) = \sin^2(2^{kx} \arcsin(\sqrt{p}))$, mais nous y sommes presque. L'encadré 4 explique comment y parvenir.

LA FONCTION DE PIANTADOSI

4

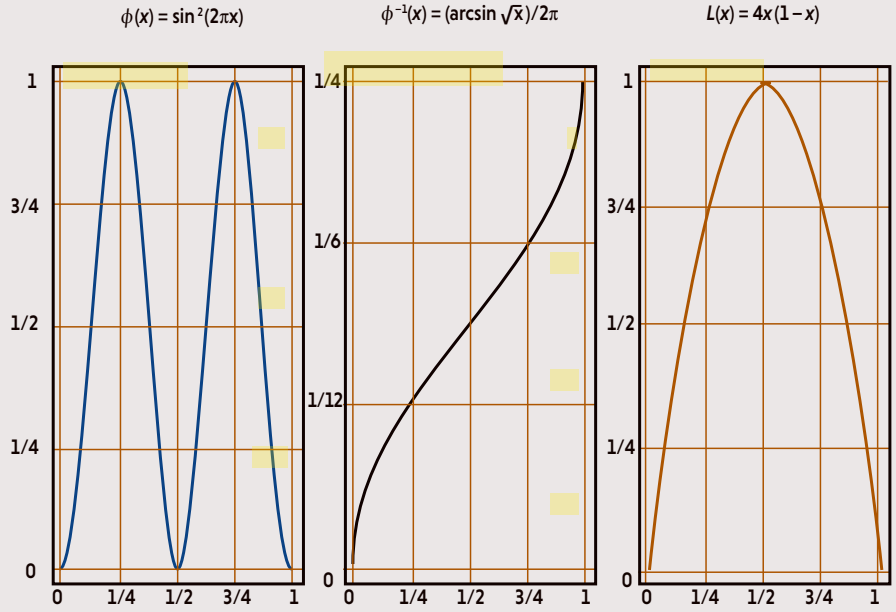
Compte tenu de ce qui a été dit dans l'encadré 3 sur la fonction D, on considère à présent les trois fonctions :
$\phi(x) = \sin^2(2\pi x)$;
$\phi^{-1}(x) = (\arcsin\sqrt{x})/2\pi$;
$L(x) = 4x(1-x)$.
Par définition, $\phi(\phi^{-1}(x)) = x$ pour $0 \leq x \leq 1$ et $\phi^{-1}(\phi(x)) = x$ pour $0 \leq x \leq 1/4$.
Voici des calculs inévitables pour arriver au miracle.
$\phi(D(x)) = \phi(2x) = \sin^2(4\pi x) = 4\sin^2(2\pi x)(1-\sin^2(2\pi x)) = L(\phi(x))$ (on a utilisé $\sin(2a) = 2\sin a \cos a$, $\cos^2 a = 1 - \sin^2 a$ et $\sin(x + 2\pi) = \sin(x)$).
En prenant $x = \phi^{-1}(p)$, soit $p = \phi(x)$, on a :
$\phi(D(\phi^{-1}(p))) = L(p)$,
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 $y_0, y_1, \ldots, y_n$ pour $x = 0, 1, \ldots, n$. 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 $y_0, y_1, \ldots, y_n$ pour $x = 0, 1, \ldots, n$, il suffit donc que $D^{kx}(\phi^{-1}(p))$ donne approximativement $\phi^{-1}(y_0)$, $\phi^{-1}(y_1), \ldots, \phi^{-1}(y_n)$ pour $x = 0, 1, \ldots, n$. 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 $f_p(0)$, $f_p(1)$, ..., $f_p(n)$.

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 derniers les autres :
$q = 0{,}001010000110000110111110001000000000_2 = 0{,}157741434872150421142578$.
On calcule $p = \phi(q) = 0{,}6999652447\ldots$ et on a bien ce qui est attendu :
$f_p(0) = 0{,}6999652447$, soit environ 0,7,
$f_p(1) = 0{,}3996355201$, soit environ 0,4,
$f_p(2) = 0{,}4999999999$, 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 $y_0, y_1, \ldots, y_n$. 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 $1\,000 \times 1\,000$ 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 $y_0, y_1, \ldots, y_n$, 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 :

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

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