Aspects prédictifs des modèles mathématiques

Daniel Justens

Des maths pour prévoir l'avenir

La notion de *temps* est impossible à définir. Nous nous bornons à constater sa réalité et à la subir. Il est par contre aisé de *mesurer* des *intervalles de temps* et d'introduire une variable temporelle dans les représentations mentales de fragments de l'univers que l'on nomme « modèles mathématiques », aux côtés d'autres variables plus tangibles. Soit Y un phénomène quantifiable. Si Y dépend de k variables d'état et du temps, on note :

$$Y = f(X_1, \ldots, X_k, t)$$

Lorsque les modèles prennent la forme de fonctions mathématiques usuelles, ces dernières comportent différents paramètres que l'on doit remplacer par des valeurs numériques dans les applications, c'est le calibrage du modèle. Des observations forcément effectuées dans le *passé* permettent alors, sous certaines hypothèses, d'obtenir une forme explicite du modèle. En donnant à t des valeurs correspondant à ce que l'on appelle le futur, on en vient à donner une description du fragment de réalité étudié dans un futur plus ou moins lointain et on obtient un modèle à caractère prédictif qui permet de concrétiser partiellement l'un des plus vieux phantasmes de l'humanité : la maîtrise d'une partie de son avenir.

Dérivabilité et déterminisme

Quelle forme peut prendre la fameuse fonction f ? Supposons que l'état du système étudié soit mesurable et que l'on dispose d'une base de données quantifiant f dans un passé récent et supposons également que cette fonction soit dérivable par rapport au temps. Nous négligeons la référence aux k variables d'état pour simplifier les notations. Affirmer que f est dérivable par rapport au temps à l'instant t_0 (que l'on appellera le *présent*) revient à supposer que la limite pour Δt tendant vers 0 de l'expression

$$\frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}$$

existe quelle que soit la suite de valeurs Δt (tendant vers 0) choisies. On note cette limite $f'(t_0)$. Même pour des fonctions pour lesquelles une forme explicite n'est pas connue, avec des valeurs de Δt suffisamment petites, on peut estimer :

$$f'(t_0) \approx \frac{f(t_0 + \Delta t) - f(t_0)}{\Delta t}$$

En considérant, comme c'est le cas dans notre cadre, uniquement des valeurs négatives pour Δt (observations passées) on en vient à calculer assez correctement $f'(t_0)$, ce qui permet d'écrire au voisinage de l'instant t_0 :

$$f(t_0 + \Delta t) \approx f(t_0) + f'(t_0).\Delta t$$

La valeur $f(t_0)$ décrit l'état présent du phénomène étudié qui est supposé mesurable, celle de $f'(t_0)$ peut être obtenue à partir d'instants antérieurs à t_0. En choisissant à présent des Δt strictement positifs, on en vient à donner une description de l'état *futur* du système étudié à partir du passé proche et du présent. Un phénomène décrit par une fonction dérivable est donc toujours localement prévisible et sa dynamique descriptive prend la forme d'une équation différentielle lorsque sa construction se fait à partir d'hypothèses affectant ses variations instantanées.

On peut aussi voir une fonction dérivable comme localement linéaire (ou affine). Et observer le fait suivant : pour toute fonction f à dérivée continue, n'importe quelle fonction à dérivée continue égale à $f'(t_0)$ *en* t_0 va « prolonger *agréablement* » la fonction f. Mais la seule représentation graphique que l'on peut prolonger de manière unique, localement ou non, sans hypothèses fortes sur son évolution, est la droite.

Remarquons enfin que l'hypothèse de dérivabilité est équivalente à la considération de phénomènes dont la variabilité est proportionnelle, localement, au temps écoulé.

Modèles prédictifs en physique

De nombreux exemples de ce type peuvent se trouver en physique. Certains de ces modèles assurent une description correcte à court terme (chute d'un corps), d'autres à beaucoup plus long terme (planètes du système solaire), jusqu'à ce que des erreurs de mesure des conditions initiales ou les approximations du modèle ne fassent dévier les prévisions des observations qui pourront être réalisées. L'exemple le plus connu est précisément celui des trajectoires des planètes du système solaire dont une composante chaotique a été mise en évidence par l'astronome français Jacques Laskar en 1989, les planètes les plus excentriques étant Mercure, Vénus, Mars et … la Terre. Mais que l'on se rassure : il faudra attendre environ 200 millions d'années pour que le comportement de ces planètes ne devienne imprévisible, étant donnée leur configuration actuelle et le niveau de précision de nos mesures. Les travaux récents de l'astronome évaluent (sur base d'heuristiques) la probabilité associée à des scénarii catastrophes (collisions de planètes ou éjection du système solaire) à moins d'1 % dans les 5 milliards d'années à venir. Voilà de quoi rassurer les âmes sensibles aux discours cataclysmiques des gourous annonciateurs de fin du monde.

On touche là un des éléments majeurs de l'intérêt des modèles prédictifs : la disparition de la peur ou du recours au surnaturel. Imaginons la crainte de nos lointains ancêtres lorsque chaque hiver ils voyaient les jours raccourcir. Sans modèle prédictif validé, seul le recours à la magie pouvait les rassurer et les convaincre qu'ils allaient revoir des jours plus longs, plus chauds, nécessaires à leur survie. Sous cet éclairage, l'apparition de *fêtes de la lumière* lors du solstice d'hiver (fête de Yule en Scandinavie, culte de Mithra et Saturnales dans la Rome pré-chrétienne, …) dans la plupart des cultes dits « païens », prend tout son sens. Selon la tradition catholique, cette fête fut christianisée par le Pape Libère en 354 pour devenir *Noël*, dont une des étymologies possibles se réfère au gaulois « noio » signifiant *nouveau* et « hel », *le soleil*.

Modèles prédictifs stochastiques en économie

Nous venons de constater que la différentiabilité se traduisait par la prévisibilité locale. Qu'en est-il mathématiquement de systèmes, fréquents en économie, dont l'évolution est au moins en partie aléatoire ? Un raisonnement élémentaire permet de comprendre le type de relation à mettre en place pour décrire ce type de phénomène.

Considérons une fois encore le présent noté t_0 et envisageons la description d'un phénomène f à l'instant t ($t>t_0$) en divisant l'intervalle $[t_0\ , t[$ en n intervalles temporels de durée identique Δt. Si une partie du phénomène étudié nous est connue et est prévisible, nous pouvons utiliser une fonction dérivable f_1 pour en assurer la description. Pour un système additif, on posera :

$$f = f_1 + f_2.$$

En faisant peser tout le poids de notre ignorance sur la fonction f_2, pour laquelle, sans autre renseignement, il faudra se contenter de représenter les variations sur chaque intervalle Δt par des variables aléatoires X_i forcément équidistribuées et indépendantes étant donné notre absence d'information. On note alors :

$$f(t) = f_1(t) + X_1 + X_2 + X_3 + \dots + X_n$$

Sans information non plus sur la tendance des X_i, cette dernière doit théoriquement être supposée nulle. En effet, toute notre connaissance en tendance du phénomène est par hypothèse contenue dans f_1. Une façon simple d'envisager ce type de modifications aléatoires est la construction d'un arbre binomial pour lequel les changements à chaque instant Δt, sont décrits par les variables aléatoires :

$$\begin{aligned} X_i &= +\,\Delta x \text{ avec probabilité } 0{,}5 \\ &= -\,\Delta x \text{ avec probabilité } 0{,}5 \end{aligned}$$

Sous cette hypothèse et celle d'indépendance, on calcule très aisément espérance et variance de $f(t)$:

$$\mathrm{E}[f(t)] = f_1(t)$$
$$\mathrm{V}[f(t)] = n\,\Delta x^2$$

Dans la dernière égalité, le temps apparaît sous la double forme t et n avec $t = t_0 + n\,\Delta t$. On peut donc écrire :

$$\mathrm{V}[f(t)] = (t-t_0)\,\Delta x^2/\Delta t.$$

La variance mesure une moyenne des carrés des variations. Sa dimension est donc égale au carré de la dimension de la variable étudiée. Pour retrouver un paramètre mesurant l'intensité des variations dans l'unité d'origine, on passe généralement à l'écart-type, racine carrée de la variance, en obtenant ainsi un paramètre difficile à interpréter mais très souvent utilisé.

Apparaît alors un phénomène étrange : l'écart-type de la variation aléatoire observée est proportionnel à la racine carrée de l'intervalle de temps considéré, à savoir $(t\text{-}t_0)$. Une mise sous forme localement linéaire par rapport au temps est tout bonnement impossible. Les phénomènes aléatoires ne peuvent être décrits au moyen de fonctions dérivables. C'est cette considération qui a conduit à la construction des équations différentielles stochastiques.

Remarquons enfin que le choix d'une description additive conduit, lorsque n augmente, à la considération d'une partie aléatoire prenant la forme d'une distribution normale, sous nos hypothèses et étant donné le théorème central limite. Lorsque le phénomène est de type multiplicatif (comme souvent en finance), on passe aux logarithmes pour retrouver le cas précédent. C'est ce qui explique l'usage de distributions lognormales pour la description de produits financiers.

Bref …

Il y a deux types d'approches au problème de la construction de modèles prédictifs. La première prend la forme d'une fonction décrivant l'état futur du système. La deuxième comporte deux composantes, l'une décrivant la *tendance* du système, c'est-à-dire sa partie prévisible, l'autre les variations inattendues relatives à ce *trend*. Ces deux approches conduisent à des considérations mathématiques différentes même en ce qui concerne la partie prévisible des descriptions.

En effet, pour des phénomènes dont on espère une description exacte, on recherche la forme mathématique passant le plus précisément possible par les observations, en considérant les résidus entre valeurs observées et prévues comme dus exclusivement à des erreurs de mesure. Le modèle comporte un volet explicatif, qui va se traduire par la mise en équations (éventuellement différentielles) des variations du phénomène à partir de « lois » et qui nous amène à considérer que l'on a *compris* le fragment de réel modélisé lorsque les résultats obtenus théoriquement ne diffèrent pas trop des observations. C'est le cas pour la représentation de la trajectoire des planètes du système solaire. L'apparition de composantes chaotiques, imprévisibles mêmes dans ce cadre déterministe est relativement marginale. Le modèle proposé permet une description de l'état futur du système sous forme d'une valeur numérique.

Pour des phénomènes partiellement aléatoires, les observations passées utilisées pour le calibrage du modèle comportent deux types d'*erreurs* : celles dues à notre système de mesure forcément imparfait et aussi celles dues à la composante aléatoire du phénomène. On distingue ici deux parties au modèle : l'une prévisible, l'autre pas. Mais l'observation du phénomène dans sa globalité ne permet pas de distinguer clairement les effets de ces deux composantes. Pour déterminer la forme explicite de la fonction décrivant la tendance du système, on doit construire f_1 passant « à peu près » au milieu des observations. Dans ce cas, on aura souvent recours à des méthodes statistiques d'ajustement pour le calibrage, ou à des équations différentielles stochastiques pour la forme à donner aux modèles. La description du futur est toujours possible, mais sous une forme différente : le modèle est composé de la somme (ou du produit) de deux composantes, l'une déterministe donnant la tendance, l'autre prenant la forme d'une variable aléatoire explicite permettant la construction de zones de confiance pour la « trajectoire » des états futurs du phénomène modélisé.