EEN WISKUNDIG MODEL VAT DE INTERACTIE TUSSEN DE DIVERSE VARIABELEN VAN EEN SYSTEEM SAMEN IN EEN FORMULE EN LAAT ZO TOE OM HET SYSTEEM BETER TE BEGRIJPEN, TE ANALYSEREN, TE ONTWERPEN EN TE VERBETEREN. WISKUNDIGE MODELLEN WORDEN VOORAL INGEZET IN WETENSCHAPPEN EN TECHNIEK, MAAR HEBBEN OOK HEEL WAT PRAKTISCHE TOEPASSINGEN IN ANDERE GEBIEDEN. DE VRAAG IS HOE MEN HET BESTE WISKUNDIGE MODEL VOOR EEN SYSTEEM KAN VINDEN.

Creatieve wiskundige modellen voor realistische toepassingen

Joos Vandewalle



René Magritte (1898-1967), La trahison des images (Ceci n'est pas une pipe), 1929. Los Angeles (CA), Los Angeles County Museum of Art (LACMA). Oil on canvas. Overall: 25 3/8 x 37 in. (64.45 x 93.98 cm). Unframed canvas: 23 11/16 x 31 7/8 inches, 1 1/2 inches deep, 39 5/8 inches diagonal. Purchased with funds provided by the Mr. and Mrs. William Preston Harrison Collection (78.7). © 2010. Digital Image Museum Associates/LACMA/Art Resource NY/Scala, Florence

Een 'wiskundig model' kunnen we op een niet-wiskundige manier als volgt definiëren: 'Een wiskundig model van een systeem beschrijft, niet in woorden maar in formules, hoe de diverse variabelen van het systeem onderling interageren en hoe ze interageren met de buitenwereld.' Wiskundige modellen worden vooral gebruikt in wetenschappen en techniek, zoals in de biologie, de elektrotechniek en de natuurkunde, maar ze worden ook toegepast in vele andere gebieden, zoals de economie, de sociologie en de politicologie. Om de wetenschap van het wiskundig modelleren te kunnen bedrijven is uiteraard ook heel wat kennis van de wiskunde nodig, maar hier gaan we niet dieper op in. Wiskundige modellen vormen vaak een zeer valabel alternatief voor de natuurlijke taal om een systeem of een reëel of fictief object te beschrijven. Een wiskundig model is bovendien noodzakelijk als men computers, apparaten en chips wil gebruiken om bepaalde problemen aan te pakken, of als men computers nuttige opdrachten wil laten vervullen. Door de toename van de rekencapaciteit van computers verwachten we trouwens dat de complexiteit van de systemen die we aankunnen nog zal toenemen.

Men kan zich natuurlijk afvragen waarom men de moeite wil doen om een wiskundig model van een systeem te maken. Een belangrijke motivatie is om het systeem beter te begrijpen. Maar er zijn ook meer praktische redenen om modellen af te leiden en te gebruiken. Zo kan men een wiskundig model gebruiken om een systeem onder diverse omstandigheden te simuleren op een computer, en zo het gedrag nader te bestuderen. Zulke simulaties –

'virtuele prototypes' genoemd – zijn vaak veel economischer dan een fysische uitvoering van het systeem. Deze methode wordt bijvoorbeeld gebruikt bij het ontwerp van nieuwe wagens door autoconstructeurs. Er is software die toelaat om het geluid in de wagen te reduceren, de vermoeidheid van de onderdelen te bestuderen of de trillingen van de wagen te analyseren. Simulatie wordt ook gebruikt voor systemen waarbij fysische experimenten te gevaarlijk zijn – zoals nucleaire reactoren – of waarbij het echte systeem te traag of te groot is – zoals voor berekeningen van ruimteweer. Diverse supercomputercentra, die over massieve parallelle computers beschikken, zijn in belangrijke mate actief in het simuleren van grote systemen.

Een tweede belangrijke klasse van toepassingen zijn allerhande voorspellingen zoals het weer, het klimaat, het elektriciteitsverbruik, of zelfs het dagverbruik van kranten. Een derde klasse van toepassingen betreft de optimalisatietechnieken van systemen, producten en processen. Die vertrekken van een wiskundig model met diverse vrije parameters, en het komt er dan op aan om de waarde te vinden van de parameters waarvoor de doelfunctie (bijvoorbeeld de kostprijs) zo laag mogelijk ligt. Een vierde klasse van gebruik zijn de classificaties van objecten of fenomenen, met belangrijke toepassingen in de automatisering van medische en technische diagnostiek, beeldherkenning en andere. Ten slotte worden modellen gebruikt bij het ontwerp en het operationele gebruik van regelaars voor diverse technische systemen en productiesystemen, zoals in de chemische industrie.

Een belangrijke vraag is vervolgens hoe men een wiskundig model voor een systeem kan vinden. Hier onderscheidt men twee essentieel verschillende methodes. Een eerste klasse van methodes steunt op de interne gedragingen of het interne mechanisme van het systeem. In het Engels heeft men het over de 'first principles', zoals de wetten van de fysica, bijvoorbeeld de wetten van Newton voor de mechanica, de wetten van Kirchoff voor de elektriciteit, de wetten van de stromingsleer, of de wetten van de thermodynamica. Voor vele domeinen zijn die principes of basiswetten goed gekend. Men kan dan de wiskundige modellen van het systeem afleiden als de onderdelen en de onderlinge verbindingen gekend zijn. Omdat het systeem volledig inspecteerbaar of meetbaar moet zijn, noemen we die methodes de '*whitebox*methodes'. Die methodiek is uiteraard sterk afhankelijk van het domein en wordt vaak gebruikt in vele domeinen van de techniek.

Daarnaast zijn er ook de '*blackbox*methodes', waarbij alleen metingen van de ingangen en de uitgangen van het systeem vereist zijn. Hierbij wordt een wiskundige vergelijking bepaald die zo goed mogelijk bij die metingen aansluit. Het grote voordeel van die klasse van methodes is dat ze geen interne inspectie van het systeem vereisen, en dus ook kunnen worden toegepast op domeinen waar de interne wetten nog niet goed gekend of te complex zijn. Bovendien zijn die blackboxmodellen universeel toepasselijk en dus vaak interdisciplinair en grensoverschrijdend. Deze methodes noemt men ook wel 'data mining' of gegevensontginning. Soms combineert men de twee methodes, namelijk als de interne gedragingen van een deel van het systeem bekend zijn en van het andere deel onbekend zijn. In dit geval heeft men het over een '*greybox*model'. In het vervolg van deze tekst zullen we meer specifiek ingaan op de blackboxmodellen omdat ze veruit het minst bekend zijn en een groot potentieel hebben. Voorts is het ook nuttig om het onderscheid te maken tussen 'lineaire' en 'niet-lineaire' modellen. Bij lineaire modellen zal met een dubbele ingang ook een dubbele uitgang overeenkomen, zoals we dit gewoon zijn in een regel van drie. Heel wat fenomenen beantwoorden hier echter niet aan en die noemt men dan de niet-lineaire systemen.

Een wiskundig model is een adequate voorstelling van de realiteit binnen een bepaald kader of voor bepaalde doelstellingen

Om te vermijden dat een model verkeerd wordt gebruikt, is het belangrijk om de band van het wiskundige model met de realiteit te bestuderen. Toen René Magritte zijn befaamde schilderij *Ceci n'est pas une pipe* maakte, reflecteerde hij eigenlijk over de band tussen de realiteit van een pijp, de vlakke voorstelling ervan in zijn schilderij en de beschrijving in woorden van het voorwerp 'pijp'. Precies zoals de verwoording in de taal of de figuur in het schilderij geen 'echte' pijp zijn – een pijp die je kunt vasthouden en roken – zo is een wiskundig model ook geen equivalent van de realiteit, maar is het wel een goede of relevante afspiegeling van die realiteit voor een bepaald gebruik. Anders gezegd, het wiskundige model is een adequate voorstelling van de realiteit binnen een bepaald kader of voor bepaalde doelstellingen. Zo zijn de wiskundige vergelijkingen – die zowel met blackboxmethodes als met whiteboxmethodes worden verkregen – alleen geldig binnen een bepaald werkingsgebied waarvoor ze afgeleid zijn. Ze kunnen of mogen daarbuiten dus niet worden gebruikt. De wet van Ohm bijvoorbeeld, die het verband uitdrukt tussen de stroom en de spanning van een weerstand (namelijk dat de spanning gelijk is aan de stroom maal de weerstand), is niet meer geldig als er een zo hoge spanning wordt aangelegd aan de weerstand dat die heet wordt en smelt. De relevantie van whiteboxmodellen en zeker ook van blackboxmodellen vergroot enorm als die modellen worden geverifieerd in experimenten en metingen die nog niet gebruikt werden bij de afleiding van de modellen. Dit heet de 'validatie' van het model.

Het feit dat een wiskundig model alleen uit metingen van relevante variabelen van een systeem kan worden afgeleid, en dat op basis van dit model voorspellingen kunnen worden gedaan, lijkt op het eerste gezicht voor buitenstaanders misschien wel wat op het principe van een magische kristallen bol. Dit is echter niet juist. Het volstaat om de methodiek van het afleiden van blackboxmodellen conceptueel goed te begrijpen. Hierbij vertrekt men steeds vanuit een verzameling van mogelijke of zinvolle modellen. Lineaire modellen bijvoorbeeld zijn vaak de modellen van eerste orde waarmee gewerkt wordt. De inherente relatie tussen de relevante variabelen van het systeem wordt verkregen door binnen die verzameling van mogelijke wiskundige vergelijkingen dit wiskundige model te selecteren dat zo goed mogelijk aansluit bij de gemeten variabelen. Naargelang de methodiek heeft men het over 'systeemidentificatie', 'machine learning', 'artificiële neurale netwerken' of 'support vector machines'.

Als de modelverzameling deze is van de lineaire systemen met geheugen of dynamica, dan spreken we van 'systeemidentificatie'. Die wordt gebruikt zowel in de econometrie bij Box-Jenkinsmodellen als in de techniek bij procesverbetering met de 'model predictieve controle'. Daarnaast beschouwt men 'lerende systemen' als systemen die uit de modelverzameling van lineaire en niet-lineaire verbanden deze selecteren die zo goed mogelijk bij de metingen aansluiten. Kinderen leren dieren herkennen door herhaaldelijk plaatjes met dierenbeelden zoals een kat, een hond of een vis te benoemen. Bij een juist antwoord worden ze door hun ouders bevestigd, bij een fout antwoord worden ze gecorrigeerd. Op eenzelfde manier zal bij gesuperviseerde lerende systemen de computer tijdens de leerfase steeds maar betere modellen afleiden, tot die modellen geen fouten meer maken. Het is dan cruciaal om aan het wiskundige model nieuwe ongeziene ingangen voor te leggen en daarbij te verifiëren of de juiste uitgang verkregen wordt, net zoals aan de kinderen ook gevraagd wordt om andere beelden van een kat te benoemen. Dit is de validatiestap, die voor alle blackboxmodellen noodzakelijk is om vertrouwen te hebben in de bruikbaarheid van het model.

De noodzaak van validatie betekent meteen ook dat de metingen of data waarover we beschikken, moeten worden opgesplitst in twee delen: een deel dat gebruikt wordt om het beste model mee te berekenen of te vinden (de trainingsdata) en een deel dat gebruikt wordt om de kwaliteit van dit model te verifiëren (de validatiedata). Het is duidelijk dat de opsplitsing van de metingen in trainings- en validatiedatasets evenwichtig moet gebeuren, zodat zowel moeilijke als gemakkelijke gevallen in beide datasets voorkomen. Gewoonlijk wordt hier een vuistregel toegepast waarbij tweederde van de data op willekeurige wijze in de trainingsset komen en de rest in de validatieset. Als nu het model dat na trai-

ning verkregen is, op de validatieset ongeveer evenveel fouten maakt als op de trainingsset, dan beschikken we over een goed getraind model en kunnen we dit model met vertrouwen ook gebruiken op ongeziene maar gelijkaardige data als de validatieset. Als tijdens het trainingsproces van het model de nauwkeurigheid voor de trainingsgegevens nog verbetert, terwijl dit voor de validatieverzameling niet meer verbetert, dan heeft men te maken met het ongewenste fenomeen van 'overtraining'. Hierbij is het wiskundige model bezig met het uit het hoofd leren van de trainingsverzameling, en dit levert geen meerwaarde op buiten de trainingsverzameling. Het wiskundige model wordt juist ontworpen om te functioneren in andere omstandigheden dan die in de trainingsverzameling: dit gewenste fenomeen noemt men 'generalisatie'. Er zijn diverse manieren om het memoriseren tegen te gaan, zoals het stoppen van de training als de fout op de validatieverzameling begint te stijgen en het reduceren van de verzameling modellen waaruit geleerd wordt. Zo zou in ons voorbeeld een kind dat memoriseert en niet generaliseert, een kat die van links naar rechts loopt als een ander dier herkennen dan een kat die van rechts naar links loopt.

Artificiële neurale netwerken nu vormen een speciale klasse van lerende systemen, die zoals biologische neurale netwerken opgebouwd zijn uit een grote collectie van eenvoudige bouwblokken, neuronen die in een groot netwerk geschakeld zitten. Elk neuron berekent eerst een lineaire combinatie, waarbij de ingangen van het neuron telkens worden vermenigvuldigd met gewichten of parameters en opgeteld, en voert dan op het resultaat een niet-lineaire functie uit om de uitgang van het neuron te bepalen. Typische neurale netwerken bestaan uit twee tot drie lagen van neuronen waarvan de uitgangen van de vorige laag de ingangen van de volgende laag aansturen zoals in de biologische systemen. Het trainen en het gebruiken van zo'n neuraal netwerk gebeurt met softwarepakketten zoals Matlab (toolbox Neural Networks). Alle gewichten van alle neuronen vormen dan samen de vrije parameters die tijdens het leren worden geoptimaliseerd. Hoewel het ontwerp van zo'n neuraal netwerk nogal heuristisch gebeurt – dit wil zeggen met vallen en opstaan – kan toch theoretisch worden gegarandeerd dat neurale netwerken met drie lagen en voldoende neuronen in de lagen elk praktisch bruikbaar continu niet-lineair verband tussen ingangen en uitgangen kunnen weergeven. In 1900 werd dit als open probleem nummer 13 geformuleerd door de Duitse wiskundige Hilbert. In de jaren 1950 werd het probleem opgelost door de Russische wiskundigen Kolmogorov en Arnold. Sindsdien is dit als bewijs gebruikt dat artificiële neurale netwerken met twee tot drie lagen in staat zijn om alle realistische problemen aan te pakken en werd er een hele reeks toepassingen voor gevonden, zoals het voorspellen van het elektriciteitsgebruik en van de dagelijkse afzet van kranten. Hoewel de gevonden modellen enigszins aansluiten bij de biologische systemen, hebben de neurale netwerken wel als nadeel dat ze geen echte verklaring bieden voor de interne mechanismen in het echte systeem.

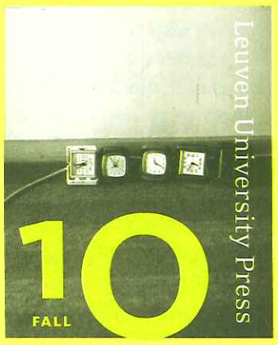
Chaotische systemen vormen een zeer belangrijke en uitdagende klasse van systemen die erg gevoelig zijn aan de beginvoorwaarden

In een recente bijdrage in *Science* wordt nu een alternatief ontwikkeld waarbij een wiskundig model wordt afgeleid dat beter 'fysisch' onderbouwd is dan bijvoorbeeld in een artificieel neuraal netwerk. Dit is zo omdat een veel grotere verzameling van basisfuncties kan worden gebruikt, die bovendien vaak ook betere fysische betekenis hebben zoals kinetische energie of geometrische informatie over afstanden en hoeken. Zo slagen de auteurs erin om een chaotisch systeem als een dubbele slinger te modelleren vanuit geregistreerde metingen. Chaotische systemen vormen een zeer belangrijke en uitdagende klasse van systemen, die erg gevoelig zijn aan de beginvoorwaarden. Zo zal voor een lichtjes verschillende begintoestand van een dubbele slinger het gedrag na korte tijd erg verschillend zijn. Dit levert steeds het verrassende effect op dat toeschouwers zeer moeilijk kunnen inschatten hoe die dubbele slinger verder zal evolueren. Er wordt beweerd dat het weer op aarde zich gedraagt als een chaotisch systeem, waarbij een kleine wijziging ten gevolge van het fladderen van de vleugels van een vlinder in Australië een groot gevolg kan hebben, zoals een tropische storm in Florida.

Wat ook de specifieke methode van de blackbox-modellering is, men moet steeds een afweging maken tussen het generaliserende karakter van het wiskundige model en de zuinigheid – met daaraan gekoppeld de complexiteit van de computerberekeningen. Een eenvoudig model met een beperkt aantal termen kan slechts in beperkte mate generaliseren, maar dit is wel eenvoudiger in gebruik en in berekeningen voor de computer. Anderzijds kan een wiskundig model met meer termen in het model beter generaliseren, maar men zal langer moeten rekenen vooraleer men een optimaal model gevonden heeft. Het heeft dan een beter generaliserend vermogen.

Tot slot zou men zich ook kunnen afvragen of het geautomatiseerde afleiden van wiskundige modellen de rol van de wetenschappers of de ingenieurs overbodig zal maken. Het omgekeerde is eerder het geval, want modellen zullen toelaten om veel effectiever de systemen te analyseren en verbeteren. Wat die trend naar breder gebruik van wiskundige modellen wel met zich meebrengt, is de noodzaak van een betere opleiding van de gebruikers van die methodes en ook de beschikbaarheid van hoogwaardige software.

Michael Schmidt en Hod Lipson, 'Distilling Free-Form Natural Laws from Experimental Data', in: *Science*, 3 april 2009, vol.324, 5923, 81-85.



Nieuw bij Universitaire Pers Leuven

De najaarscatalogus 2010 is er!

Vraag een exemplaar aan via marketing@upers.kuleuven.be of surf naar www.upl.be en ontdek onze nieuwste publicaties.

UNIVERSITAIRE PERS LEUVEN

Minderbroederstraat 4 bus 5602, 3000 Leuven, tel +32 (0)16 32 53 45 - www.upl.be

1425