

Een wiskundig model van de coronavirusepidemie in Frankrijk

Nicolas Bacaër

Institut de Recherche pour le Développement
Wiskundige en computermодellering van complexe systemen
Les Cordeliers, Parijs, Frankrijk
nicolas.bacaer@ird.fr

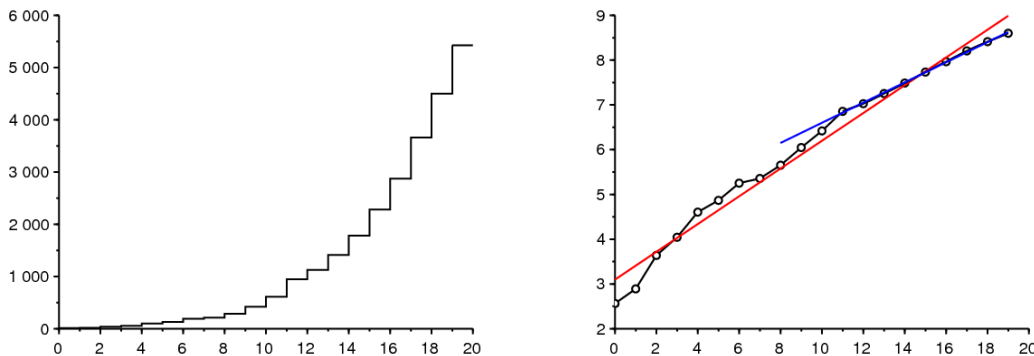
<https://hal.archives-ouvertes.fr/hal-02509142>

Overzicht

We passen een epidemisch model van het SEIR-type aan op de epidemische gegevens van het nieuwe coronavirus in Frankrijk in 2020. Als de contacten vanaf een bepaalde datum dicht bij het begin van de epidemie tot nul worden teruggebracht, is de uiteindelijke omvang van de epidemie dichtbij van die verkregen door het cumulatieve aantal gevallen op deze datum te vermenigvuldigen met de reproduceerbaarheid \mathcal{R}_0 van de epidemie. Met onze schattingen, $\mathcal{R}_0 \simeq 2,3$ in Frankrijk.

Het model

Figuur 1^a toont het cumulatieve aantal gedetecteerde gevallen van coronavirus in Frankrijk tussen 25 februari en 15 maart 2020. Tussen deze twee datums ging het cumulatieve aantal van 13 naar 5427. We zien een groei die exponentieel lijkt. Figuur 1^b laat zien hoe een regressielijn zich aanpast in logaritmische coördinaten. We vinden dat de curve toeneemt als $e^{\lambda t}$ met een groeisnelheid $\lambda \simeq 0,31$ per dag (rechts rood). De verdubbelingstijd is $\log(2)/\lambda \simeq 2,2$ dagen. Als we ons daarentegen beperken tot de laatste 9 dagen, die bijzonder goed zijn uitgelijnd op een logaritmische schaal, verkrijgen we $\lambda \simeq 0,225$ per dag en een verdubbelingstijd van 3,1 dagen (blauw rechts). Aangezien de gegevens aan het begin van de epidemie worden verstoord door een groot deel van de nieuwe gevallen die worden geïmporteerd, is het waarschijnlijk de tweede schatting die betrouwbaarder is; dit is degene die we hieronder zullen gebruiken.



Figuur 1. a) Cumulatief aantal gevallen ontdekt in Frankrijk tussen 25 februari en 15 maart 2020, volgens Santé publique France. b) Natuurlijke logaritme van het cumulatieve aantal gevallen en lineaire regressielijnen.

Laten we een wiskundig model voorstellen voor deze epidemie. Laten we de Franse bevolking verdelen in vier compartimenten volgens het klassieke SEIR-model (zie bijvoorbeeld [2, p. 61]): waarschijnlijk besmet (S), besmet maar niet besmettelijk (E) besmettelijk zonder bescherming (I) en de infectie waarschijnlijk niet overdraagt (d.w.z. infectieus maar beperkt, genezen of overleden, R). We kunnen dit model tot in het oneindige verfijnen om het realistischer te maken (zie bijvoorbeeld onder vele andere referenties [3,7]), maar we hebben geprobeerd het aantal onbekende parameters zoveel mogelijk te beperken; we streven ook naar een meer theoretisch dan praktisch resultaat.

Opmerking N de totale bevolking, a het effectieve contactpercentage, b de snelheid waarmee besmette mensen besmettelijk worden en c de snelheid waarmee besmettelijke mensen worden geïsoleerd:

$$\frac{dS}{dt} = -a S \frac{I}{N} \quad (1)$$

$$\frac{dE}{dt} = a S \frac{I}{N} - b E \quad (2)$$

$$\frac{dI}{dt} = b E - c I \quad (3)$$

$$\frac{dR}{dt} = c I, \quad (4)$$

met $N = S(t) + E(t) + I(t) + R(t)$. Om de koppeling met de data te maken, kunnen we dat bedenken $R(t)$ komt overeen met het cumulatieve aantal gevallen dat op dit moment bekend is t .

Aan het begin van de epidemie blijft het aantal gevallen erg klein in vergelijking met de totale bevolking, zodat $S(t) \simeq N$, wat leidt tot linearisatie

$$\frac{dE}{dt} \simeq a I - b E, \quad \frac{dI}{dt} \simeq b E - c I.$$

De epidemie heeft daarom de neiging exponentieel te groeien als $e^{\lambda t}$, waar λ is de grootste eigenwaarde van de matrix

$$\begin{pmatrix} -b & a \\ b & -c \end{pmatrix}. \quad (5)$$

De karakteristieke veelterm is

$$\lambda^2 + (b+c)\lambda + b(c-a) = 0. \quad (6)$$

Dus

$$\lambda = \frac{-(b+c) + \sqrt{(b+c)^2 - 4b(c-a)}}{2} = \frac{-(b+c) + \sqrt{(b-c)^2 + 4ab}}{2}. \quad (7)$$

Gemiddelde duur in het compartiment E , wat de moeite waard is 1/ben die we assimileren met de incubatieperiode, zou in de orde van 4 dagen liggen [4]. Neem daarom $b = 0,25$ per dag. Gemiddelde duur in het compartiment I voor isolatie, wat de moeite waard is 1/c is moeilijker in te schatten omdat het niet alleen afhangt van de biologische kenmerken van het virus, maar ook van de snelheid waarmee de gevallen worden geïsoleerd, die van land tot land verschillen. Stel dat het ongeveer 1 dag geleden is dat de epidemie in Frankrijk plaatsvindt wanneer de bewoners al goed op de hoogte zijn van het bestaan van de pandemie; de patiënten zijn niet lang geïsoleerd. Dus $c = 1$ per dag. We leiden dat af uit formule (7)

$$a = \frac{(2\lambda + b + c)^2 - (b - c)^2}{4b} = (\lambda + c) \left(1 + \frac{\lambda}{b}\right). \quad (8)$$

waarmee het effectieve contactpercentage numeriek kan worden berekend a van de waargenomen groeisnelheid λ .

Stel je voor dat maatregelen voor de volksgezondheid het effectieve contactpercentage door een aantal zouden kunnen delen k die groter is dan 1. Hoeveel moet minstens de moeite waard zijn k de epidemie stoppen? Deze waarde van k , traditioneel genoteerd \mathcal{R}_0 Lotka volgend en door hem "reproduceerbaarheid" genoemd [8, p. 102], wordt eenvoudig verkregen door op te merken dat wanneer a wordt vervangen door $a' = a/\mathcal{R}_0$, het nieuwe groeipercentage van de epidemie λ' moet nul zijn, wat volgens vergelijking (6) leidt tot $c - a'/\mathcal{R}_0 = 0$ en naar

$$\mathcal{R}_0 = \frac{a}{c} = \left(1 + \frac{\lambda}{b}\right) \left(1 + \frac{\lambda}{c}\right) \simeq 2,33$$

als we de numerieke waarde gebruiken ($\lambda \simeq 0,225$ per dag) voorgesteld door de epidemische curve van figuur 1. Bekijk de onzekerheden over de parameters b en c kan dit slechts een geschatte waarde zijn.

Op een meer technische manier (zie bijvoorbeeld [9]) had je dat kunnen merken \mathcal{R}_0 was ook de spectrale straal van de matrix

$$\begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \begin{pmatrix} b & 0 \\ -b & c \end{pmatrix}^{-1}.$$

Laten we teruggaan naar het SEIR-model (1) - (4). Laten we niet vergeten hoe we de uiteindelijke omvang van de epidemie kunnen bepalen zonder volledige interventie; het is een gemakkelijke aanpassing van de methode die wordt gebruikt voor het SIR-model (zie bijvoorbeeld [6, p. 76]). Vergelijking (1) laat dat zien

$$\frac{d}{dt} \log S = -\frac{a}{N} I(t),$$

dus door te integreren $t = 0$ bij $t = +\infty$:

$$\log S(\infty) - \log S(0) = -\frac{a}{N} \int_0^{\infty} I(t) dt. \quad (9)$$

Aan het begin van de epidemie was er nog niemand in het compartiment R , dus $R(0) = 0$. Vergelijking (4) laat dat zien

$$R(\infty) = c \int_0^{\infty} I(t) dt, \quad (10)$$

waar $R(\infty)$ geeft de limiet aan wanneer $t \rightarrow +\infty$ van de functie $R(t)$, die toeneemt en toeneemt met N . Daarnaast hebben we te allen tijde $S(t) + E(t) + I(t) + R(t) = N$. Wanneer $t \rightarrow +\infty$ houdt de epidemie uiteindelijk op, zodat $E(t)$ en $I(t)$ neigen naar 0. Op de limiet, dus alleen de mensen die aan de epidemie ontsnapten en degenen die besmet waren maar die het compartiment binnengingen R :

$$S(\infty) + R(\infty) = N. \quad (11)$$

Door (9), (10) en (11) te combineren, zien we dat

$$N - R(\infty) = S(0) \exp\left(-\frac{a}{c} \frac{R(\infty)}{N}\right).$$

Aan het begin van de epidemie waren er maar een paar besmette mensen in de bevolking, dus $S(0) \simeq N$. De vergelijking voor de uiteindelijke omvang van de epidemie kan worden geschreven als

$$1 - \frac{R(\infty)}{N} \simeq \exp\left(-\mathcal{R}_0 \frac{R(\infty)}{N}\right),$$

die toevallig dezelfde vorm heeft als voor het SIR-model [6]. Met $\mathcal{R}_0 \simeq 2,33$, vinden we digitaal $R(\infty)/N \simeq 87\%$.

Tweede fase met ingrijpende ingreep

Stel je voor dat op een bepaalde datum τ worden er drastische maatregelen genomen om het nieuwe effectieve contactpercentage terug te brengen tot 0 terwijl dat wel het geval is $R(\tau)$ cumulatieve gevallen. Er waren bijvoorbeeld 5427 cumulatieve gevallen vanaf 15 maart, de datum waarop de maatregelen betreffende scholen en openbare plaatsen van kracht werden. [5] Kunnen we dan voorspellen wat de nieuwe uiteindelijke omvang van de epidemie zal zijn?

Tegen het einde van de exponentiële fase waar $t \leq \tau$ en waar het totale aantal gevallen nog steeds een klein deel van de totale bevolking vertegenwoordigt, hebben we

$$E(t) \simeq u e^{\lambda t}, \quad I(t) \simeq v e^{\lambda t}, \quad R(t) \simeq w e^{\lambda t},$$

waar (u, v) is een eigenvector geassocieerd met de grootste eigenwaarde λ van de matrix (5). Dus $-b u + a v = \lambda u$. Met vergelijking (8) vinden we dat

$$u = \frac{a v}{\lambda + b} = \frac{\lambda + c}{b} v.$$

Zoals $dR/dt \simeq \lambda R$ voor $t < \tau$ wanneer t is niet te dicht bij 0, we hebben

$$I(\tau) = \frac{1}{c} \frac{dR}{dt}(\tau) \simeq \frac{\lambda}{c} R(\tau).$$

Maar $I(\tau) \simeq v e^{\lambda\tau}$, dus

$$E(\tau) \simeq u e^{\lambda\tau} = \frac{\lambda + c}{b} v e^{\lambda\tau} \simeq \frac{\lambda + c}{b} I(\tau) \simeq \left(\frac{\lambda^2}{bc} + \frac{\lambda}{b} \right) R(\tau).$$

De contacten moeten worden teruggebracht tot nul, daar hebben we voor $t > \tau$

$$\frac{dS}{dt} = 0 \tag{12}$$

$$\frac{dE}{dt} = -bE \tag{13}$$

$$\frac{dI}{dt} = bE - cI \tag{14}$$

$$\frac{dR}{dt} = cI. \tag{15}$$

Zonder dit systeem op te lossen is het duidelijk dat de uiteindelijke omvang van de epidemie zal zijn $R(\infty) = R(\tau) + E(\tau) + I(\tau)$, aangezien er is $E(\tau) + I(\tau)$ geïnfecteerde personen die zich nog niet in het compartiment bevinden R destijds τ . Dus

$$R(\infty) \simeq R(\tau) \left(1 + \frac{\lambda^2}{bc} + \frac{\lambda}{b} + \frac{\lambda}{c} \right) = R(\tau) \left(1 + \frac{\lambda}{b} \right) \left(1 + \frac{\lambda}{c} \right) = \mathcal{R}_0 R(\tau).$$

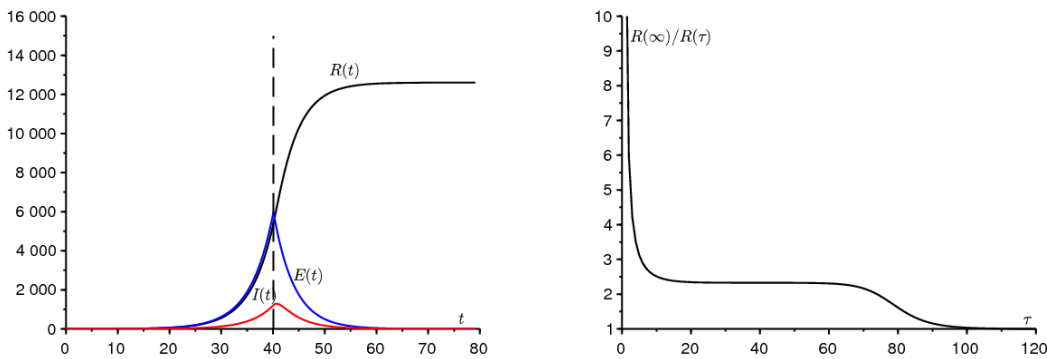
Dus als de contacten vanaf een bepaalde datum dicht bij het begin van de epidemie tot nul worden teruggebracht (dichtbij genoeg om de lineaire benadering nog geldig te maken, maar niet te dichtbij om het gelineariseerde systeem de tijd te hebben gehad om te convergeren) naar de eigenvector geassocieerd met de eerste eigenwaarde, dan ligt de uiteindelijke omvang van de epidemie dicht bij die verkregen door het cumulatieve aantal gevallen op deze datum te vermenigvuldigen met de reproduceerbaarheid \mathcal{R}_0 van de epidemie. Een vergelijkbaar resultaat wordt gemakkelijk op dezelfde manier verkregen voor een SIR-model. Maar de mate van algemeenheid van deze waarneming moet nog worden bepaald.

Laten we de analogie met het concept van 'groeipotentieel' van bevolkingsgroepen in de demografie opmerken [10, p. 176]. Het is de verhouding tussen de uiteindelijke stationaire populatie en de populatie op een bepaald moment als de vruchtbaarheid op dit moment plotseling wordt gedeeld door de reproduceerbaarheid \mathcal{R}_0 , zodat de bevolking een asymptotisch tempo van nulgroei overhoudt. Net als bij onze berekening, is het door te veronderstellen dat de initiële populatie "stabiel" is in de zin van Lotka (dat wil zeggen gegeven door de eerste eigenvector) dat Keyfitz een relatief eenvoudige formule voor het potentieel van toename, een formule die ook meespeelt \mathcal{R}_0 hoewel op een meer gecompliceerde manier dan voor ons SEIR-model [10, p. 179].

Merk ook op dat de schatting van $R(\tau) + E(\tau) + I(\tau)$ uit de gegevens $R(\tau)$ slechts één is analoog aan het probleem dat zich voordeed aan het begin van de hiv-epidemie om het aantal hiv-positieve mensen te schatten op basis van het aantal gemelde aids-gevallen.

Met $R(\tau) = 5427$ en $\mathcal{R}_0 \simeq 2,33$, geeft het $R(\infty) \simeq 12\,600$. Benadruk nogmaals de onzekerheid rond de parameters b en c , wat tot uiting komt in de waarde van $R(\infty)$.

Figuur 2^a illustreert dit tweefasemodel. We namen $N = 65 \times 10^6$, $S(0) = N - 1$, $E(0) = 1$, $I(0) = 0$ en $R(0) = 0$. De parameter a wordt gegeven door formule (8) met $\lambda = 0,225$ per dag. We namen $\tau = 40,1$ dagen zo $R(\tau) \simeq 5415$ ligt zeer dicht bij de gegevens 5427 van 15 maart. Door de simulatie iets langer voort te zetten dan in de figuur zien we dat numeriek $R(\infty)/R(\tau) \simeq 2,33 \simeq \mathcal{R}_0$.



Figuur 2. a) Voorbeeld van simulatie van het tweefasemodel. b) Het rapport $R(\infty)/R(\tau)$ afhankelijk van τ .

Figuur 2^b laat zien hoe het rapport $R(\infty)/R(\tau)$ varieert afhankelijk van de tijd τ waar het contactpercentage wordt verlaagd tot nul. We kunnen duidelijk een plateau waarnemen waar deze verhouding dichtbij is \mathcal{R}_0 . Wanneer $\tau \rightarrow 0$, men heeft $R(\tau) \rightarrow 0$ en $R(\infty) \rightarrow E(0) + I(0) = 1$, zo het rapport $R(\infty)/R(\tau)$ neigt naar oneindig. Het rapport komt dichterbij \mathcal{R}_0 wanneer τ is in de orde van de inverse van het verschil tussen de twee eigenwaarden van de matrix (5). Integendeel $\tau \rightarrow \infty$, dan komt de interventie te laat; de epidemie is al voorbij en $R(\infty)/R(\tau) \rightarrow 1$. Er wordt verwacht dat de breedte van de plank waar $R(\infty)/R(\tau)$ is dichtbij \mathcal{R}_0 groeit als een constante vermenigvuldigd met $\log(N)$ wanneer $N \rightarrow +\infty$ omdat dit bijvoorbeeld het gedrag is van tijd tot aan de epidemische piek in een SIR-model met constante coëfficiënten.

In werkelijkheid zal het effectieve contactpercentage zeker niet helemaal nul zijn $t > \tau$. De waarde verkregen voor $R(\infty)$ kan niettemin worden beschouwd als een ondergrens van de reële waarde, aangezien het zeker is dat de uiteindelijke omvang van de epidemie groter zal zijn bij niet-nul contacten dan bij nul contacten voor $t > \tau$. In dit verband moet eraan worden herinnerd dat epidemische modellen van het SIR- of SEIR-type met een variabel contactpercentage niet "eentoniig" zijn, in die zin dat een vermindering van het contactpercentage soms kan leiden tot een grotere uiteindelijke omvang van de epidemie. [1].

Als we willen weten wanneer de piek van de aankomst van nieuwe gevallen zal plaatsvinden, die van belang kan zijn voor congestieproblemen in ziekenhuizen, moeten we het systeem oplossen (12) - (15). We vinden $E(t) = E(\tau) e^{-b(t-\tau)}$ en

$$\begin{aligned}
I(t) &= I(\tau)e^{-c(t-\tau)} + b \int_{\tau}^t e^{-c(t-s)} E(s) ds \\
&= I(\tau)e^{-c(t-\tau)} + bE(\tau) \frac{e^{-b(t-\tau)} - e^{-c(t-\tau)}}{c-b} .
\end{aligned}$$

Met de numerieke waarden van figuur 2^a, men heeft $E(\tau) \simeq 5\,970$, $I(\tau) \simeq 1\,218$ en de afgeleide van $I(t)$ annuleert 0,6 dagen nadat het contactpercentage is verlaagd naar 0. Voor een contacttarief dat niet nul is, zou de vertraging uiteraard langer zijn.

Concluderend hebben we het meest gunstige scenario verkend, waarbij het contactpercentage vanaf een bepaalde datum tot nul wordt teruggebracht. We hebben een eenvoudige benadering gevonden voor de uiteindelijke omvang van de epidemie op basis van het aantal gevallen dat werd ontdekt op het moment van de reductie. Het blijft echter om dit resultaat rigoureuzer te verklaren en aan te tonen, ongetwijfeld door de totale bevolking te vergroten N richting $+\infty$.

Dankwoord

We danken Hisashi Inaba en Ali Moussaoui voor hun opmerkingen over het manuscript.

Referenties

1. Bacaër N., Gomes M. G. M. (2009) On the final size of epidemics with seasonality. *Bull. Math. Biol.* 71 : 1954 – 1966.
2. Corlosquet – Habart M., Janssen J., Manca R. (2012) Modélisation stochastique du risque de pandémie : stratégies de couverture et d'assurance. Lavoisier, Cachan.
3. Danchin A., Ng P. T. W., Turinici G. (2020) A new transmission route for the propagation of the SARS – CoV – 2 coronavirus. <https://doi.org/10.1101/2020.02.14.20022939>
4. Guan W. J. et coll. (2020) Clinical characteristics of 2019 novel coronavirus infection in China. *New England Journal of Medicine*, doi : 10.1056/NEJMoa2002032
5. Hecketsweiler C., Pietralunga C. (2020) Coronavirus : les simulations alarmantes des épidémiologistes pour la France. *Le Monde*, publié en ligne le 15/03/2020.
6. A. Hillion (1986) *Les Théories mathématiques des populations*. Presses Universitaires de France, Paris.
7. Lin Q. Y., Zhao S., Gao D. Z., et coll. (2020) A conceptual model for the coronavirus disease 2019 (COVID – 19) outbreak in Wuhan, China with individual reaction and governmental action. *Int. J. Infect. Dis.* 93 : 211 – 216.
8. Lotka A. J. (1939) *Théorie analytique des associations biologiques*, 2e partie. Hermann, Paris.
9. Nkague Nkamba L. (2012) Robustesse des seuils en épidémiologie et stabilité asymptotique d'un modèle à infectivité et susceptibilité différentielle. Thèse, Université de Lorraine et Université Gaston Berger.
10. Pressat R. (1995) *Éléments de démographie mathématique*. AIDELF, Paris.