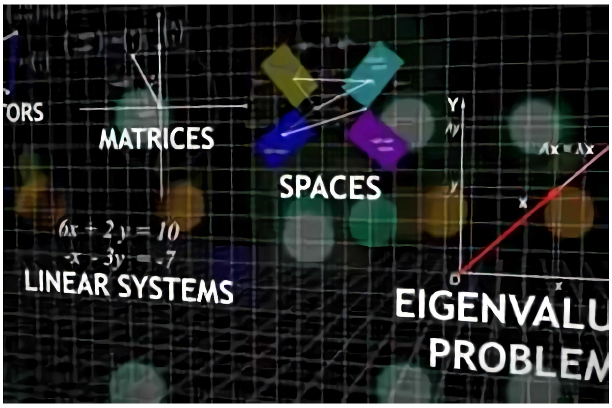
UNIVERSITEIT GENT
Faculteit Wetenschappen
Vakgroep Wiskunde

Lineare Algebra en Meetkunde

Sara Rottey



Academiejaar 2019 – 2020

Inhoudsopgave

Hoofdstuk 3

Vierkante matrices

Hoewel vierkante matrices nog steeds matrices zijn, en er dus eigenlijk niets “speciaals” is aan het feit dat een matrix vierkant is, laten vierkante matrices toch een heleboel zaken toe die bij niet-vierkante onmogelijk zijn. Een notatie als A^2 houdt steek, een inverse kan soms bestaan, en ga zo maar door. In dit hoofdstuk bestuderen we een aantal concepten die enkel voor vierkante matrices zinvol zijn.

3.1 Determinanten

In deze sectie introduceren we een belangrijke afbeelding op $M_n(K)$ met waarden in K, namelijk de *determinant* afbeelding. De determinantafbeelding is een speciale afbeelding

$$\det : M_n(K) \to K$$

die aan elke $n \times n$ matrix over het veld K een getal uit K aan die matrix toekent. Voor we deze kunnen definiëren moeten we eerst enkele nieuwe concepten invoeren.

Definitie 3.1. De (i, j)-*minormatrix* van een matrix is de $(n-1) \times (n-1)$-matrix A_{ij} die we bekomen door in A de i-de rij en de j-de kolom te schrappen.

$$\begin{array}{c|ccc|}
 & & j & \\
\hline
 & & \backslash & \\
i & ///// & \backslash & /// \\
 & & \backslash & \\
\hline
\end{array}$$

Als notatie stellen we de matrix A voor aan de hand van zijn kolomvectoren: $A = (A_1, \ldots, A_n)$, dan kunnen we volgend begrip op de kolommen invoeren.

Definitie 3.2. Een afbeelding $f : M_n(K) \to K$ is *multilineair in de kolommen* als voor elke matrix $A \in M_n(K)$ geldt dat

$$f(A_1, \ldots, A_{k-1}, \sum_{i=1}^{n} c_i A_i, A_{k+1}, \ldots, A_n) = \sum_{i=1}^{n} c_i f(A_1, \ldots, A_{k-1}, A_i, A_{k+1}, \ldots, A_n)$$

voor elke $k \in \{1, \ldots, n\}$ en elke $c_1, \ldots, c_n \in K$.

Een afbeelding is *multilineair op de rijen* als de analoge eigenschap geldt voor de rijen van A. Deze beide eigenschappen zijn equivalent omdat de formule voor de rijen dezelfde is als die voor de kolommen op de getransponeerde matrix, en de eigenschap voor elke $A \in M_n(K)$ moet gelden (en dus ook voor A^t).

Definitie 3.3. De determinant is de unieke afbeelding $\det : M_n(K) \to K$ waarvoor

- $\det I_n = 1$;
- det is multilineair in de kolommen (en dus ook in de rijen);
- als A twee gelijke rijen heeft, of twee gelijke kolommen, of een nulkolom, of een nulrij, dan is $\det(A) = 0$.

We gaan niet bewijzen dat deze uniek is, je mag dit zonder bewijs aannemen. De definiërende formule voor multilineariteit in de kolommen/rijen krijgt door de laatste bullet een vereenvoudigde vorm:

$$\det(A_1, \ldots, A_{k-1}, \sum_{i=1}^{n} c_i A_i, A_{k+1}, \ldots, A_n) = c_k \det(A_1, \ldots, A_{k-1}, A_k, A_{k+1}, \ldots, A_n).$$

Definitie 3.3 is niet erg praktisch om een determinant van een matrix effectief uit te rekenen. We geven daarom volgende eigenschap die de praktische uitrekening van de matrix zal faciliteren.

Stelling 3.4 (ontwikkelen naar een rij/kolom)**.** Zij $A \in M_n(K)$ en $1 \leq i \leq n$. Dan is

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{i,j} = \sum_{j=1}^{n} (-1)^{i+j} a_{ji} \det A_{j,i}.$$

Ook volgende eigenschap mag je zonder bewijs aannemen.

Stelling 3.5. Voor elke $A, B \in M_n(K)$ is $\det AB = \det A \cdot \det B$.

Merk op dat in het algemeen $\det A + B \neq \det A + \det B$.

Je zal in de oefeninglessen nog voldoende voorbeelden maken op het uitrekenen van determinanten. Hieronder al enkele theoretische oefeningen, waarvan het resultaat onthouden interessant kan zijn.

Stelling 3.6. Volgende beweringen zijn geldig.

- Voor elke $A \in M_n(K)$ is $\det A = \det A^t$.
- Voor elke $A \in M_n(K)$ en elke $c \in K$ is $\det cA = c^n \det A$.
- Als $A \in M_n(K)$ een diagonaal-, bovendriehoeks- of benedendriehoeksmatrix is, dan is $\det A$ het product van de diagonaalelementen.

Bewijs. Doe zelf als oefening. □

Stelling 3.7. Zij $A \in M_n(K)$, dan zijn de volgende eigenschappen equivalent:

(a) $\det A \neq 0$;

(b) de kolommen van A zijn lineair onafhankelijk (i.e., $\text{rang}(A) = n$);

(c) de rijen van A zijn lineair onafhankelijk.

Bewijs. Zij $A = (A_1, \ldots, A_n)$.

$(a) \Rightarrow (b)$ Als de kolommen lineair afhankelijk zijn, dan is er een eerste kolom, stel A_k, die afhankelijk is van de voorgaande kolommen. Dus $A_k = \lambda_1 A_1 + \cdots + \lambda_{k-1} A_{k-1}$. Vervangen we A_k door deze lineaire combinatie, dan vinden we door de multilineariteit

$$\det A = \sum_{i=1}^{k-1} \lambda_i \det(A_1, \ldots, A_{k-1}, A_i, A_{k+1}, \ldots, A_n) = \sum_{i=1}^{k-1} \lambda_i 0 = 0$$

aangezien de matrices in elke term twee gelijke kolommen hebben.

$(b) \Rightarrow (a)$ Als de kolommen van A lineair onafhankelijk zijn dan volgt uit stelling 2.23 dat de rij-echelon vorm van A gelijk is aan de eenheidsmatrix. De eenheidsmatrix bekomen we dus door de rijen van A te vervangen door lineaire combinaties van die rijen, en de gebruikte scalairen zijn in het algoritme allemaal verschillend van nul, dus zo ook hun product.

$(b) \Leftrightarrow (c)$ Dit volgt nu rechtstreeks uit het feit dat $\det A = \det A^t$.

□

Gevolg 3.8. (1) De determinant van $A \in M_n(K)$ is gelijk aan nul als en slechts als de kolommen van A lineair afhankelijk zijn.

(2) Een matrix $A \in M_n(K)$ is inverteerbaar als en slechts als $\det A \neq 0$.

Bewijs. Punt (1) is een herformulering van stelling 3.7. Punt (2) volgt uit stelling 3.7 en stelling 2.23. □

3.2 Lineaire operatoren, eigenvectoren en eigenwaarden

Definitie 3.9. Zij $T : V \to V$ een lineaire operator K-vectorruimte V. Een vector $v \in V$ waarvoor $T(v) = \lambda v$ voor zekere $\lambda \in K$, noemen we een *eigenvector*. De bijhorende waarde λ noemen we een eigenwaarde. De verzameling $\{v \in V | T(v) = \lambda v\}$ noemen we de *eigenruimte* bij λ.

Voorbeeld 3.10. Zij T een rotatie rond een willekeurige as in de ruimte. Dan is v een eigenvector als en slechts als v op de rotatie-as ligt. De bijhorende eigenwaarde is 1.

In het algemeen is rechtstreeks zoeken naar eigenvectoren enorm moeilijk, omdat we niet weten welke keuzes voor λ mogelijks resultaten zouden afleveren. Als we alle eigenvectoren en bijhorende eigenwaarden willen vinden, dan zoeken we dus eerst de mogelijke eigenwaarden, en zoeken we dan voor elk van hen de mogelijke eigenvectoren. Om de eigenwaarden te vinden, maken we gebruik van het feit dat een eigenvector eigenlijk niets anders dan een element uit $\ker(T - \lambda I)$ is, en dus enkel bestaat als $T - \lambda I$ singulier is.

De gemakkelijkste manier om dit te bepalen, althans in eindigdimensionale vectorruimtes, is als volgt. Zij A een matrixvoorstelling van de operator T, dan is dat singulier zijn equivalent met $\det(A - \lambda I_n) = 0$. Nu is $\det(A - \lambda I_n)$ een veelterm (over K) van graad n in λ, en die kunnen we oplossen naar λ.

Definitie 3.11. Zij T een lineaire operator op een eindig dimensionale K-vectorruimte V. De *karakteristieke veelterm* van T is de veelterm $\chi_T(x) := \det(xI_n - A)$, met A een matrixvoorstelling van de operator T.

Opmerking 3.12. Merk op dat deze veelterm enkel afhangt van T, niet van de gekozen matrixvoorstelling A, omdat $\det(A - \lambda I_n) = \det P^{-1}(A - \lambda I_n)P = \det(P^{-1}AP - \lambda I_n)$ voor elke basistransformatie P.

Het eigenwaardeprobleem heeft een oplossing als en slechts als de karakteristieke veelterm van T, $\chi_T(x)$, een wortel heeft in K. Voor elke dergelijke wortel λ (dit zijn de enige kandidaat-eigenwaarden) zijn de eigenvectoren nu precies de niet-nulvectoren in $\ker(T - \lambda I)$. Kiezen we opnieuw een matrixvoorstelling A van T, dan staat daar $(A - \lambda I)v = 0$, en daarvan hebben we al uitgebreid gezien hoe je ze oplost.

Opmerking 3.13. Aangezien een veelterm maar eindig veel wortels heeft, kan een lineaire operator op een eindigdimensionale vectorruimte maar eindig veel eigenwaarden hebben. Uit de hoofdstelling van de algebra (zie verderop in de cursus) volgt dat het eigenwaardeprobleem over het veld $\mathbb{C}$, van de complexe getallen, altijd een oplossing heeft.

Voorbeeld. Als voorbeeld bepalen we de eigenwaarden en eigenvectoren van de matrix

$$A := \begin{pmatrix} 48 & 13 & 51 & 8 \\ 32 & 36 & 32 & 42 \\ 42 & 49 & 39 & 48 \\ 26 & 26 & 5 & 5 \end{pmatrix}.$$

Hier is

$$\chi_A(x) = \det(xI - A) = \begin{vmatrix} x-48 & -13 & -51 & -8 \\ -32 & x-36 & -32 & -42 \\ -42 & -49 & x-39 & -48 \\ -26 & -26 & -5 & x-5 \end{vmatrix}$$

over de veeltermring $\mathbb{C}[x]$. De eigenwaarden van de matrix A zijn de wortels in $\mathbb{C}$ van de veelterm

$$\chi_A(x) = x^4 - 128x^3 - 47x^2 - 6264x - 170628.$$

Maple geeft:

$$128.8221181, 4.309699661 + 11.03239297 * I, -9.441517436, 4.309699661 - 11.03239297 * I$$

We bepalen nog de eigenvectoren bij de eerste eigenwaarde 128.8221181, door het stelsel $A - 128.8221181 \cdot I = 0$ op te lossen. We bekomen bij deze eigenwaarde een 1-dimensionale deelruimte van eigenvectoren:

$$\begin{pmatrix} 1.009283153\, s_2 \\ s_2 \\ 1.270328997\, s_2 \\ 0.473203080\, s_2 \end{pmatrix}, s_2 \in \mathbb{C}.$$

De eigenvectoren bij de andere eigenwaarden vinden we op dezelfde manier. Natuurlijk werden de algoritmes voor het bepalen van eigenwaarden en eigenvectoren in Maple geimplementeerd. De instructie *Eigenvectors*(A) geeft als output een kolomvector en een matrix: in de kolomvector staan de eigenwaarden van A en de kolommen van de matrix zijn de coördinaten van eigenvectoren die horen bij de respectievelijke eigenwaarden. De lezer kan het bovenstaande voorbeeld verifiëren met behulp van deze instructies.

3.3 Diagonaliseren van matrices

Definitie 3.14. Een lineaire operator $T : V \to V$ op een eindigdimensionale vectorruimte V noemen we *diagonaliseerbaar* als er een matrixvoorstelling D van T bestaat met D een diagonaalmatrix.

Als A een matrixvoorstelling is van de operator T ten opzichte van de basis $\mathcal{B}$ voor V, dan is de matrixvoorstelling van T ten opzichte van een basis $\mathcal{B}'$ van de vorm $Q^{-1}AQ$, waarbij Q de transitiematrix is van de basis $\mathcal{B}$ naar de basis $\mathcal{B}'$.

Definitie 3.15. We zeggen dat twee matrices $A, B \in M_n(K)$ *geconjugeerd* of *toegevoegd* zijn als er een *inverteerbare* matrix $P \in M_n(K)$ bestaat zodat $B = P^{-1}AP$. De *conjugatieklasse* van een matrix $A \in M_n(K)$ is de verzameling

$$\{P^{-1}AP | \text{ met } P \text{ een inverteerbare matrix over } K\}.$$

Een matrix $A \in M_n(K)$ heet *diagonaliseerbaar* als de conjugatieklasse van A een diagonaalmatrix bevat.

Opmerking 3.16. Zij $T : V \to V$ een lineaire operator op een eindigdimensionale vectorruimte en zij A een matrixvoorstelling van T. Uit het bovenstaande volgt dat de verzameling van de matrices die geconjugeerd zijn met de matrix A juist bestaat uit alle matrixvoorstellingen van de operator T. Een matrix geassocieerd aan een lineaire operator is dus diagonaliseerbaar als en slechts de operator diagonaliseerbaar is.

Stelling 3.17. Zij T een lineaire operator op een n-dimensionale vectorruimte V over K. Dan is T diagonaliseerbaar als en slechts als er een basis van eigenvectoren bestaat.

Bewijs. Dat de voorwaarde voldoende is volgt eenvoudig door de matrixvoorstelling van T ten opzichte van de basis van eigenvectoren te beschouwen.

Onderstel omgekeerd dat er een basis $\{v_1, \ldots, v_n\}$ bestaat zodat de matrixvoorstelling van T ten opzichte van deze basis een diagonaalmatrix is. Dan is $Tv_i = a_i v_i$, met a_i het i-de element op de diagonaal van de matrixvoorstelling van T. De basis $\{v_1, \ldots, v_n\}$ is dus een basis van eigenvectoren voor T. $\square$

Opmerking 3.18. Vermits de karakteristieke veelterm van een operator T niet afhangt van de matrixvoorstelling van T is het duidelijk dat voor de karakteristieke veelterm van een diagonaliseerbare operator geldt

$$\chi_T(x) = \prod_i (x - \lambda_i)^{n_i},$$

i.e., de veelterm ontbindt volledig in lineaire factoren over het veld K. Het omgekeerde is niet waar, als de karakteristieke veelterm $\chi_T(x)$ volledig in lineaire factoren ontbindt over K, dan is T niet noodzakelijk diagonaliseerbaar.

Definitie 3.19. De *algebraïsche multipliciteit* van een eigenwaarde λ_i is het getal n_i, i.e. de multipliciteit van λ_i als wortel van χ_T. De *meetkundige multipliciteit* van λ is de dimensie van de eigenruimte horend bij λ. Beide zijn ≥ 1 voor elke eigenwaarde λ.

Vermits we geïnteresseerd zijn in diagonaliseerbare operatoren T, onderstellen we voor de rest van deze sectie dat de karakteristieke veelterm van een operator volledig ontbindt in lineaire factoren, i.e. $\chi_T(x) = \prod_i (x - \lambda_i)^{n_i}$ (anders zijn we klaar want de matrix is niet diagonaliseerbaar).

Opmerking 3.20. De ruimte $\mathrm{End}(V)$ van alle lineaire operatoren op een n-dimensionale vectorruimte V over K is een n^2-dimensionale K-vectorruimte. Zij $T \in \mathrm{End}(V)$ dan is het stel

$$\{I, T, T^2, T^3, \ldots\}$$

een lineair afhankelijk stel. Er bestaat dus een d zodat

$$T^d + \sum_i a_i T^i = 0,$$

met $a_i \in K$, (0 staat hier voor de nul-operator). Dit betekent dat de veelterm $x^d + \sum a_i x^i \in K[x]$ nul wordt als we x door T vervangen.

Definitie 3.21. Zij $T \in \mathrm{End}(V)$ en zij $\mu_T(x) \in K[x]$ de veelterm (met hoogstegraadscoëfficiënt gelijk aan 1) van minimale graad zodat $\mu_T(T) = 0$. Dan noemen we $\mu_T(x)$ de *minimaalveelterm* van T.

We geven twee eigenschappen zonder bewijs. De eerste eigenschap is een gevolg van het Euclidisch algoritme voor veeltermen. De tweede eigenschap (de stelling van Cayley-Hamilton) is een basiseigenschap van lineaire operatoren die eenvoudig te bewijzen is voor diagonaliseerbare operatoren.

Lemma 3.22. Zij $T \in \text{End}(V)$ en zij $\phi(x) \in K[x]$ een veelterm zodat $\phi(T) = 0$, dan is de minimaalveelterm $\mu_T(x)$ een deler van $\phi(x)$.

Stelling 3.23 (Cayley-Hamilton)**.** Zij $T \in \text{End}(V)$ en zij $\chi_T(x) \in K[x]$ de karakteristieke veelterm vam T, dan is $\chi_T(T) = 0$. De karakteristieke veelterm van T en de minimaalveelterm van T hebben dezelfde lineaire factoren.

Gevolg 3.24. Zij T een lineaire operator op een n-dimensionale vectorruimte V over K. De operator T, met als eigenwaarden $\lambda_1, \ldots, \lambda_m$, is diagonaliseerbaar als en slechts als het minimaalpolynoom van T gelijk is aan $(x - \lambda_1) \cdots (x - \lambda_m)$.

De volgende stelling (opnieuw zonder bewijs) geeft de nodige en voldoende voorwaarde voor het diagonaliseerbaar zijn van een lineaire operator T op een n-dimensionale vectorruimte V over K.

Stelling 3.25. Een operator T op een n-dimensionale vectorruimte V over K is diagonaliseerbaar als en slechts als voor alle eigenwaarden van deze operator T, de algebraïsche multipliciteit gelijk is aan de meetkundige multipliciteit.

Gevolg 3.26. Een lineaire operator T, op een n-dimensionale vectorruimte V over K, met n verschillende eigenwaarden, is altijd diagonaliseerbaar.

Hoewel Stelling 3.25 een werkbaar criterium geeft om de diagonaliseerbaarheid te bepalen, willen we voor praktische toepassingen natuurlijk ook kunnen bepalen welke basisovergang P onze matrix diagonaliseert. Hiervoor voldoet echter eenderwelke basis van eigenvectoren, gezien T op die manier elke basisvector afbeeldt op een veelvoud van zichzelf, wat betekent dat zijn geassocieerde matrixvoorstelling voor die basis diagonaal is. De matrix P die de overgang naar die eigenwaardenbasis regelt bekom je door simpelweg de basisvectoren als kolommen van de matrix te gebruiken. Als de matrix diagonaliseerbaar is, dan zal wegens Stelling 3.25 P vierkant zijn, en zal door wat we in deze paragraaf uitgelegd hebben en $P^{-1}AP$ diagonaal zijn.

Voorbeeld 3.27. De matrix $A = \left(\begin{smallmatrix} 2 & -10 \\ 0 & -3 \end{smallmatrix}\right)$ heeft als eigenwaarden 2 en -3 (het is een bovendriehoeksmatrix). Vermits er twee verschillende eigenwaarden zijn is de matrix diagonaliseerbaar (gevolg 3.26). We zoeken de eigenvectoren, deze vinden we als oplossingen van de stelsels:

$$\begin{pmatrix} 0 & -10 \\ 0 & -5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ en } \begin{pmatrix} 5 & -10 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Het eerste stelsel heeft als oplossingsverzameling $\{(t, 0) | t \in K\}$, de niet-nul vectoren hierin zijn de eigenvectoren met eigenwaarde 2. Het tweede stelsel heeft als oplossingsverzameling $\{(2t, t) | t \in K\}$, de niet-nul vectoren van deze verzameling zijn de eigenvectoren met eigenwaarde -3. We kiezen uit elke verzameling een niet-nul vector, $(1, 0)^t$ (eigenwaarde 2) en $(4, 2)^t$ (eigenwaarde -3). De matrix $P = \left(\begin{smallmatrix} 1 & 4 \\ 0 & 2 \end{smallmatrix}\right)$ is de transitiematrix van de standaardbasis naar de basis gegeven door de eigenvectoren $(1, 0)^t, (4, 2)^t$. De inverse matrix van P is $P^{-1} = \left(\begin{smallmatrix} 1 & -2 \\ 0 & 1/2 \end{smallmatrix}\right)$. We kunnen verifiëren dat $P^{-1}AP = \left(\begin{smallmatrix} 2 & 0 \\ 0 & -3 \end{smallmatrix}\right)$.

Opmerking 3.28. In de praktijk gebruikt men de diagonaalvoorstelling vaak om berekeningen te vereenvoudigen. Immers, als je pakweg A^n wil kennen voor heel hoge n, en A is diagonaliseerbaar, dan is

$$A^n = A \cdot A \cdot A \cdots A = P^{-1}DPP^{-1}DPP^{-1}DP \cdots P^{-1}DP = P^{-1}D^nP,$$

en diagonaalmatrices tot een macht verheffen is heel simpel: gewoon alle elementen op de diagonaal tot die macht verheffen. Dit zal enorm handig blijken in de volgende sectie.

3.4 Lineair iteratieve processen en Markov-processen

Definitie 3.29. Een *lineair iteratief proces* is een proces waarbij elke toestand (op tijdstip) k kan beschreven worden door een vector $p(k) = (x_1, \ldots, x_n)^t$ en waarbij de overgang tussen twee opeenvolgende toestanden $p(k)$ en $p(k + 1)$ gegeven wordt door $p(k + 1) = Ap(k)$, met A een reële $n \times n$-matrix. Hierbij is A onafhankelijk van k.

Voorbeeld 3.30. De Fibonacci rij wordt gedefinieerd door $F_0 = 0$, $F_1 = 1$ en $F_n = F_{n-1} + F_{n-2}$ voor $n \geq 2$. Dit geeft de volgende getallen:

$$0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377, 610, 987, 1597, 2584, 4181, 6765, \ldots$$

Kunnen we een "gesloten formule" vinden, waarmee we F_n kunnen berekenen zonder eerst de vorige Fibonacci-getallen te moeten bepalen?

Om dit te doen, stellen we vast dat

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix}.$$

We stellen $A := \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, dan is

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = A^{n-1} \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} = A^{n-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Om de hoge machten van A te bepalen kunnen we best A diagonaliseren. En dit is mogelijk vermits A twee verschillende eigenwaarden heeft, namelijk de nulpunten van $x^2 - x - 1$: $\lambda_1 = \frac{1+\sqrt{5}}{2}$ en $\lambda_2 = \frac{1-\sqrt{5}}{2}$. We vinden als corresponderende eigenvectoren $v_1 = \begin{pmatrix} \frac{1+\sqrt{5}}{2} \\ 1 \end{pmatrix}$ en $v_2 = \begin{pmatrix} \frac{1-\sqrt{5}}{2} \\ 1 \end{pmatrix}$. Stellen we

$$C := \begin{pmatrix} \frac{1+\sqrt{5}}{2} & \frac{1-\sqrt{5}}{2} \\ 1 & 1 \end{pmatrix},$$

dan is $C^{-1}AC$ een diagonaalmatrix met op de diagonaal de eigenwaarden λ_1, λ_2. We vinden voor de Fibonacci rij

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = A^{n-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = C \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^{n-1} C^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Als we dit product uitwerken zien we dat

$$F_n = \frac{\lambda_1^n - \lambda_2^n}{\sqrt{5}} = \frac{1}{\sqrt{5}} \left((\frac{1+\sqrt{5}}{2})^n - (\frac{1-\sqrt{5}}{2})^n \right).$$

Voorbeeld 3.31. Bij een klinische test van een nieuw medicijn verdeelt men de onderzochte patiënten in drie groepen die we voorstellen met $v = (x_1, x_2, x_3)^t$, met x_1 het percentage patiënten met een beginnend ziektebeeld, x_2 het percentage patiënten met een gevorderd ziektebeeld, en x_3 het percentage herstelde patiënten. Aanvankelijk bestaat de groep uit 20% mensen met een beginnend ziektebeeld en 80% mensen met een gevorderd ziektebeeld. De volgende matrix beschrijft de verandering van de samenstelling van de groep patiënten van week tot week:

$$A = \begin{pmatrix} 0.3 & 0 & 0 \\ 0.2 & 0.8 & 0 \\ 0.5 & 0.2 & 1 \end{pmatrix}.$$

De samenstelling van de groep na 1 week is dan $A \cdot v$, na 2 weken $A^2 \cdot v$, en na k weken $A^k \cdot v$.

Met Maple kunnen we dit voor hoge k vrij snel berekenen en we zien dan dat na 10 weken de toestand van de verschillende groepen gegeven wordt door de vector

$$(0.000118, 9.44888, 90.55100)^t,$$

dus 90,5% van de patiënten is hersteld. We stellen vast dat na 52 weken waarschijnlijk iedereen zal hersteld zijn (99.99919%).

Definitie 3.32. Een vierkante matrix $\mathcal{M}$ over de reële getallen heet een *Markov-matrix* als:

- alle componenten van $\mathcal{M}$ in het gesloten interval $[0, 1]$ liggen,
- de som van de componenten in elke kolom van $\mathcal{M}$ gelijk is aan 1.

Een lineair iteratief proces heet een *Markov-proces* als de matrix A die het proces beschrijft een Markov-matrix is. (Voorbeeld (2) is een voorbeeld van een Markov-proces.)

Het gedrag van een lineair iteratief proces, in het bijzonder van een Markov-proces, kunnen we bestuderen aan de hand van de eigenschappen van de eigenwaarden van het proces.

Zij A de (reële) $n \times n$-matrix van een lineair iteratief proces. We onderstellen dat A diagonaliseerbaar is (over $\mathbb{C}$), dus dat er een basis is van eigenvectoren $v_1, \ldots, v_n$, met eigenwaarden respectievelijk $\lambda_1, \ldots, \lambda_n$, waarbij

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|,$$

met $\lambda_1 \in \mathbb{R}$. Merk op dat dit impliceert dat de multipliciteit van λ_1 gelijk is aan 1, de (veralgemeende) eigenruimte U_1 is dus 1-dimensionaal en we kunnen voor de basis, v_1, een reële eigenvector kiezen. Er geldt dus $Av_i = \lambda_i v_i$, voor alle $i = 1, \ldots, n$.

Zij $p(k)$ de toestandsvector van het proces op het ogenblik k, dus $p(k+1) = Ap(k)$. Zij

$$p(0) = \mu_1 v_1 + \mu_2 v_2 + \cdots + \mu_n v_n,$$

met $\mu_1 \neq 0$. Dan zal

$$\begin{array}{ll} p(k) & = \lambda_1^k \mu_1 v_1 + \lambda_2^k \mu_2 v_2 + \cdots + \lambda_n^k \mu_n v_n \\ & = \lambda_1^k (\mu_1 v_1 + (\frac{\lambda_2}{\lambda_1})^k \mu_2 v_2 + \cdots + (\frac{\lambda_n}{\lambda_1})^k \mu_n v_n) \end{array}.$$

Bijgevolg zal $\lim_{k\to\infty} p(k) = \lim_{k\to\infty} \lambda_1^k \mu_1 v_1$ zodat:

- voor k voldoende groot, zal $p(k) \approx \lambda_1^k \mu_1 v_1$;
- voor $\lambda_1 = 1$, en k voldoende groot, zal $p(k) \approx \mu_1 v_1$;
- voor $\lambda_1 < 1$, en k voldoende groot, zal $p(k) \approx 0$.

Je kan zien dat dit gedrag van het proces bevestigd wordt in de voorbeelden.

In een *Markov-proces* dat loopt over een reeks van tijdsintervallen is het zo dat de toestand in elke fase afhangt van de vorige toestand en van een *kans*; de volgende voorwaarden moeten hierbij voldaan zijn:

- De verzameling van de mogelijke toestanden is eindig.
- De toestand in een bepaalde fase hangt enkel af van de toestand in de vorige fase en de kans op die toestand.
- De kansen op een bepaalde toestand zijn constant over de tijd.

Een wiskundig model voor een Markov-proces bestaat uit de Markov-matrix A, de *kans matrix*, en de *toestandsvectoren*, v_t met t het tijdstip of fase. De toestand op tijdstip t wordt gegeven door de vergelijking

$$v_t = A v_{t-1}.$$

Wanneer voor t voldoende groot geldt dat $v_t = v_{t+1} = \cdots = v_{t+k} = \cdots$, noemt men de toestand v_t een *stabiele situatie*. Uit de gelijkheid $v_{t+1} = Av_t$ volgt dat een stabiele situatie precies dan bereikt wordt als v_t een *eigenvector* is van A met eigenwaarde 1.

Stelling 3.33. Een Markov-matrix $\mathcal{M}$ over de reële getallen heeft steeds een eigenvector met eigenwaarde 1.

Bewijs. We moeten aantonen dat 1 een eigenwaarde is voor $\mathcal{M}$, dus dat 1 een wortel is van de karakteristieke veelterm van $\mathcal{M}$. Maar de karakteristieke veelterm van $\mathcal{M}$ is gelijk aan de karakteristieke veelterm van de getransponeerde van $\mathcal{M}$ vermits

$$\det(\lambda I_n - \mathcal{M}) = \det((\lambda I_n - \mathcal{M})^t) = \det(\lambda I_n - \mathcal{M}^t).$$

Nu is vanwege de tweede eigenschap van Markov-matrices

$$(1, \ldots, 1) \cdot \mathcal{M} = (1, \ldots, 1)$$

en dus

$$\mathcal{M}^t \cdot \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Vermits $\mathcal{M}^t$ een eigenvector heeft met eigenwaarde 1, is 1 een wortel van de karakteristieke veelterm van $\mathcal{M}^t$ en dus ook van de karakteristieke veelterm van $\mathcal{M}$. $\square$

Opmerking 3.34. Men kan nog meer bewijzen over de eigenwaarden van een Markov-matrix. In het algemeen geldt

$$|\lambda_i| \leq 1, i = 1, \ldots, n,$$

met λ_i de eigenwaarden van $\mathcal{M}$. Als precies één eigenwaarde absolute waarde 1 heeft volgt uit de bovenstaande analyse dat er een *stabiele situatie* is voor het Markov-proces. De vector die deze stabiele situatie voorstelt is een eigenvector v_1 van $\mathcal{M}$ bij de eigenwaarde 1. Dit betekent dat

$$\lim_{k\to\infty} \mathcal{M}^k w = v_1,$$

een eigenvector is van $\mathcal{M}$ met eigenwaarde 1.

Verder is er nog een stelling die zegt dat een Markov-matrix, waarvan alle componenten ongelijk zijn aan nul, juist één eigenwaarde λ_1 gelijk aan 1 heeft en dus dat voor alle andere eigenwaarden geldt

$$|\lambda_i| < 1, \quad i = 2, \ldots, n.$$

Voor dergelijke Markov-matrices is er dus een stabiele situatie voor het Markov-proces.

3.5 Google's algoritme voor de rangorde van webpagina's

We hebben al vernoemd dat het algoritme dat Google gebruikt om webpagina's te ordenen gebaseerd is op een eigenwaardeprobleem. We geven nu een meer gedetailleerde beschrijving van de ideeën achter het algoritme. Het algoritme werd ontworpen[1] door studenten van Stanford, Lawrence Page, Sergey Brin, Motwani Rajeev, en Winograd Terry. Het algoritme is gebaseerd op het model van een Markov-proces.

Zij $V = \text{span}(w_1, \ldots, w_m, \ldots, w_n)$, de vectorruimte met als basis de verzameling van alle webpagina's $w_1, \ldots, w_n$ (de dimensie n van deze ruimte is het aantal webpagina's, dit is een getal groter dan $4 \cdot 10^9$). Een vector in V is dus een "lineaire combinatie van webpagina's" $\sum_{i=1}^n \alpha_i w_i$, behalve wanneer $\alpha_j = 1$ en alle andere $\alpha_i = 0$ zijn de vectoren dus geen echte webpagina's. Als we nu bovendien eisen dat $\alpha_i \geq 0$ voor alle i, en $\sum_i \alpha_i = 1$, dan heeft de vector $v = \sum_{i=1}^n \alpha_i w_i$ ook een betekenis: de vector kan namelijk geïnterpreteerd worden als een "kansvector", de coëfficiënt α_i geeft de kans aan dat je (bv. door een link te volgen) op webpagina i terecht komt.

Dit idee, (om de lineaire combinaties $\sum_i \alpha_i w_i$ te interpreteren als kansvectoren), kan gebruikt worden om een rangorde van de webpagina's te geven. Dit is het idee achter het algoritme dat Google gebruikt. We nemen aan dat: *Als men op webpagina i is, men met een zekere kans (bv. 85%) een link op die pagina volgt, anders (in het voorbeeld 15%) gaat men "random" naar een andere webpagina.*

Als er op webpagina w_j bijvoorbeeld 5 links zijn, respectievelijk naar de pagina's $w_{i_1}, \ldots, w_{i_5}$, dan is de kans dat men op pagina w_i, $i \in \{i_1, \ldots, i_5\}$ terecht komt vanaf w_j, gelijk aan $a_{ij} = 0.85 \cdot \frac{1}{5} + 0.15 \cdot \frac{1}{n}$. De kans dat je op een pagina w_l terecht komt vanaf w_j, met $l \notin \{j_1, \ldots, j_5\}$, is dan logischerwijs $a_{lj} = 0.15 \cdot \frac{1}{n}$.

Algemeen noteren we met $k(j)$ het aantal links op pagina w_j, en met $0 \leq p \leq 1$ de kans dat je een link volgt (en dus $1 - p$ de kans dat je naar een willekeurige andere pagina gaat waarvoor er geen link is op pagina w_j). De kans, dat van pagina w_j naar pagina w_i gesurft wordt, is dan $a_{ij} = pm_{ij} + (1-p)\frac{1}{n}$, met

$$m_{ij} := \begin{cases} \frac{1}{k(j)} & \text{als } k(j) \neq 0 \text{ en er een link is van } w_j \text{ naar } w_i \\ 0 & \text{als } k(j) \neq 0 \text{ en er geen link is van } w_j \text{ naar } w_i \\ \frac{1}{n} & \text{als } k(j) = 0 \end{cases}$$

Merk op dat, als $k(j) = 0$, de kans dat je van w_j maar w_i gaat gelijk is aan $p\frac{1}{n} + (1-p)\frac{1}{n} = \frac{1}{n}$, wat overeenkomt met een "random" keuze van een webpagina.

Zetten we dat nu in een matrix $A := (a_{ij})$. Startend van een webpagina, i.e. een vector $w_r = (0, \ldots, 0, 1, 0, \ldots, 0)$ (met dus precies één 1, op plaats r), is de pagina om erna op pagina w_t terecht te komen precies de t-de component van Aw_r, i.e. a_{tr}. We zouden de volgende betekenis kunnen geven aan de werking van de matrix A op de vectorruimte V. De matrix A geeft aan hoe we "door een link te volgen" van de ene webpagina op de andere webpagina komen. Als we op webpagina w zijn dan gaan we door een link te volgen naar "webpagina" Aw. Nu is Aw geen "echte" webpagina maar een kansvector, de componenten geven de kans dat we na de link op een bepaalde webpagina terechtkomen.

De "Google-matrix" A is een Markov-matrix en we hebben er voor gezorgd dat alle componenten van A strikt groter zijn dan 0. De matrix A heeft dus precies één eigenvector v met eigenwaarde 1. Het is een stabiele vector van het Markov-proces. Wat is de betekenis van die $v = (x_1, \ldots, x_n)$? Als je "lang" aan het surfen bent, dan is de kans dat je op dat moment op pagina w_i zit gegeven door x_i. Of ook, zij $x_j = \max_{i=1,\ldots,n}\{x_i\}$, dan is de kans het grootst dat je - als je blijft surfen, en op een willekeurig tijdstip stopt - op dat moment op webpagina w_j zit, waar je ook begint!

We kunnen de componenten van de eigenvector v (de stabiele situatie van het proces) dus ordenen volgens grootte en dit geeft dan een rangorde van de webpagina's. Dit is de rangorde die Google gebruikt. De componenten van de eigenvector geven de *Pagerank* van de webpagina.

[1] Origineel artikel: http://dbpubs.stanford.edu:8090/pub/1999-66 (aanrader!)