SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

En analys av Arthur Cayleys arbete "A memoir on the theory of matrices"

av

Oskar Vrethammar

2021 - No K15

En analys av Arthur Cayleys arbete "A memoir on the theory of matrices"

Oskar Vrethammar

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Rikard Bögvad

2021

Innehåll

1 Introduktion

1.1 Vem var Arthur Cayley?

Arthur Cayley (1821-1895) började sin bana som matematiker med att studera matematik på Trinity College i Cambridge, England. Efter att ha avslutat sin utbildning och erhållit titeln senior wrangler (en titel tilldelad studenter med topp-prestationer inom matematik) tog han beslutet att börja jobba som advokat för att försörja sig. Detta hindrade dock inte på något vis hans produktivitet: bara under åren som advokat publicerade Cayley närmare 300 matematiska arbeten. År 1863 valdes han till en professors-position i matematik vid Cambridge. Cayley, som var mest känd som en forskande matematiker, skulle nu bli tvungen att undervisa studenter - som en del av universitetets nya, tysk-inspirerade läroideal. Hans föreläsningar lockade tyvärr inte särskilt många studenter eftersom innehållet ofta var svårförståeligt och irrelevant för studenternas kursmaterial. Icke desto mindre är Arthur Cayley en person som drivit matematiken framåt, med sina närmare 1000 publikationer inom allt från analytisk geometri till matrisalgebra.

1.2 Varför detta arbete om ”A memoir on the theory of matrices?”

År 1858 publicerar Cayley sitt arbete ”A memoir on the theory of matrices”, där han bland annat introducerar abstrakt notation för matriser och etablerar grundläggande matrisalgebra. Drygt 150 år senare ögnar jag igenom detta arbete och blir genast intresserad. Jag slogs av hur mycket av innehållet jag förstod, även med mina begränsade matematikkunskaper. Språkbarriären till trots upplevde jag alltid att huvudbudskapet i varje paragraf sken igenom. Jag fascinerades också av hur Cayley presenterar matematiken i prosa, något som jag nu har lärt mig var typiskt för matematiska arbeten under den här tiden.

Jag har skrivit detta arbete med syftet att försöka förstå Arthur Cayleys tankegångar i arbetet ”A memoir on the theory of matrices”, utifrån mina begränsade kunskaper som student. ”Förståelsen” kan delvis handla om att ge för mig intressanta och relevanta *kommentarer*, som både kan vara spontana intryck jag har fått när jag läst ett stycke i Cayleys arbete - men också mer eftertänkta reflektioner. En annan del av förståelsen handlar för mig om att se detaljerna i Cayleys resonemang och argumentation. De gånger jag har läst något i hans arbete som intresserat mig, men där jag upplever att något kan tilläggas för att förtydliga resonemanget, har jag *kompletterat* med exempelvis ett specifikt exempel eller, om så känns relevant, ett bevis. I så stor utsträckning som möjligt har jag försökt vara Cayley trogen och använt hans notation för matriser och den algebra han introducerar.

1.3 Kort om arbetets struktur

Arbetet är uppbyggt i en "citat-kommentar/kompletterande-material"-struktur. Citat och / eller omformuleringar av Cayleys resonemang i memoaren följs alltså åt av kommenterarer och / eller kompletterande material, där jag anser det lämpligt. Kommentarer är markerade med lådor för att särskilja dem från övriga delar av texten. Det kompletterande materialet tilldelas en rubrik och avslutas med en färgad *Fin.* - notation.

Den observante läsaren kommer för övrigt kanske notera att jag inte har ägnat någon tid åt paragraf 50 och 51 i Cayleys arbete. Det är ett medvetet val: jag kände att dessa delar av arbetet låg lite för långt utanför mitt kompetensområde.

1.4 Tack

Jag vill också passa på att tacka min handledare Rikard Bögvad för all den hjälp han har bidragit med. Diskussionerna med Rikard gav mig en riktning i arbetet som jag inte hade kunnat finna på egen hand.

2 Analys och genomgång av Cayleys memoar

2.1 Införandet av abstrakt notation, fundamentala egenskaper och räkneoperationer på matriser

Låt oss börja med att undersöka Arthur Cayleys införande av abstrakt notation. Matrisen introduceras på första sidan i arbetet (originalet finner läsaren i appendix: A1). Cayley börjar med att beskriva matrisen som objekt:

> *The term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices (...) a set of quantities arranged in the form of a square.*

Med "quantities" syftar Cayley här på tal. Han motiverar införandet av ett sådant matematiskt objekt med att det förenklar lösandet av linjära ekvationssystem, speciellt system med tre okända variabler och tre ekvationer.

När matrisen som objekt väl är definierad börjar Cayley redogöra för de kvadratiska matrisernas fundamentala egenskaper. Han anknyter till de tidigare nämnda systemen med tre ekvationer och tre okända. Paragraf 2 lyder:

The notation

$$\left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'',& b'',& c'' \end{array}\right|\!\!\left)(x,\, y,\, z)\right.$$

represents the set of linear functions

$$\big((a,\, b,\, c)(x,\, y,\, z),\ \ (a',\, b',\, c')(x,\, y,\, z),\ \ (a'',\, b'',\, c'')(x,\, y,\, z)\big)$$

so that calling these (X,Y,Z), we have

$$(\mathrm{X},\,\mathrm{Y},\,\mathrm{Z}) = \left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'',& b'',& c'' \end{array}\right)(x,\,y,\,z)$$

and (...) this formula leads to most of the fundamental notions in the theory.

Kommentar

Cayleys notation för variabelmatrisen $(x,\,y,\,z)$ motsvarar det vi idag skulle kalla för en kolonnmatris med systemets variabler.

Cayley definierar vidare nollmatrisen (i 3x3-fallet) som den kvadratiska matrisen som bara innehåller nollor, samt enhetsmatrisen:

(X,Y,Z) will be identically equal to (x,y,z), if the matrix is

$$\left(\begin{array}{ccc} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{array}\right)$$

and this is said to be the matrix unity (...) The matrix zero may for the most part be represented simply by 0, and the matrix unity by 1.

I de paragrafer som följer (4 - 11) redogör Cayley i detalj för hur räkneoperationerna addition, subtraktion och sammansättning (multiplikation) av matriser fungerar, givet hans definition av den kvadratiska matrisen. Cayley utgår från 3x3-matriser och visar explicit hur koefficienterna hanteras i respektive situation. Han konstaterar att ekvationerna

$$(\mathrm{X},\,\mathrm{Y},\,\mathrm{Z}) = \left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'',& b'',& c'' \end{array}\right)(x,\,y,\,z),\quad (\mathrm{X}',\,\mathrm{Y}',\,\mathrm{Z}') = \left(\begin{array}{ccc} \alpha\ , & \beta\ , & \gamma \\ \alpha'\ , & \beta'\ , & \gamma' \\ \alpha'',& \beta'',& \gamma'' \end{array}\right)(x,\,y,\,z)$$

parvis kan adderas enligt $(\mathrm{X}+\mathrm{X}',\ \mathrm{Y}+\mathrm{Y}',\ \mathrm{Z}+\mathrm{Z}')$ och använder likheten ovan för att dra följande slutsats i paragraf 4:

$$\left(\begin{array}{ccc} a+\alpha\ , & b+\beta\ , & c+\gamma \\ a'+\alpha'\ , & b'+\beta'\ , & c'+\gamma' \\ a''+\alpha'', & b''+\beta'', & c''+\gamma'' \end{array}\right) = \left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a'',& b'',& c'' \end{array}\right) + \left(\begin{array}{ccc} \alpha\ , & \beta\ , & \gamma \\ \alpha'\ , & \beta'\ , & \gamma' \\ \alpha'',& \beta'',& \gamma'' \end{array}\right)$$

[which is] the rule for the addition of matrices; that for their subtraction is of course similar to it.

Kommentar

Att Cayley inte visar att regeln också fungerar för subtraktion beror antaglingen på att han tycker att det är så självklart: Om X, Y och Z alla är ekvationer som består av storheter kan vi ju skriva

$$\mathrm{X}-\mathrm{X}'=\mathrm{X}+(-\mathrm{X}'),\ \ \mathrm{Y}-\mathrm{Y}'=\mathrm{Y}+(-\mathrm{Y}'),\ \ \mathrm{Z}-\mathrm{Z}'=\mathrm{Z}+(-\mathrm{Z}')$$

för att på så vis omvandla subtraktion till addition. Det enda som återstår är hur vi definierar $-\mathrm{X}'$, $-\mathrm{Y}'$ och $-\mathrm{Z}'$. Men detta är ju uppenbart: att bestämma den additiva inversen till en ekvation innebär ju endast att vi byter tecken på samtliga termer i ekvationen. Vi kan således dra slutsatsen att subtraktion av matriser följer samma principiella regel som addition av matriser.

Cayley definierar i paragraf 5 och 6 några av nollmatrisens egenskaper. Han skriver till exempel att två matriser är lika om

(...) the difference of two [such] matrices is the matrix zero.

och ger ett villkor i form av en ekvation för när två matriser kan sägas vara varandras motsatser:

The equation $\mathrm{L}=-\mathrm{M}$*, written in the form* $\mathrm{L}+\mathrm{M}=0$*, expresses that the sum of the matrices* L, M *is equal to the matrix zero, the matrices so related are said to be 'opposite' to each other (...)*

I paragraf 7 påstår Cayley att matrisaddition är kommutativ och associativ, med likheterna

$$\mathrm{L}+\mathrm{M}=\mathrm{M}+\mathrm{L},\ \ (\mathrm{L}+\mathrm{M})+\mathrm{N}=\mathrm{L}+(\mathrm{M}+\mathrm{N})=\mathrm{L}+\mathrm{M}+\mathrm{N}$$

som notation för detta.

Kommentar

Bevisen för dessa lagar är mycket enkla och bygger bara på att vi använder definitionen av matrisaddition, vilket kanske är anledningen till att Cayley själv inte känner det nödvändigt att ge ytterligare exempel. Vad vi kan säga med säkerhet är att Cayley har identifierat dessa lagar som nödvändiga för att denna algebra med matriser ska fungera. Vi känner igen detta från dagens abstrakta algebra: för att kunna använda oss utav objekt som ringar och kroppar i räkningar måste vi först redogöra för hur exempelvis addition och multiplikation fungerar.

I paragraf 10 ser vi tydligt hur Cayley uppfattar matriser som faktiska storheter, när han här hanterar matrisrepresentationen av en skalär. Han skriver:

We have, in particular,

$$m\left(\begin{array}{|ccc|} 1, & 0, & 0 \\ 0, & 1, & 0 \\ 0, & 0, & 1 \end{array}\right)=\left(\begin{array}{|ccc|} m, & 0\ , & 0 \\ 0\ , & m, & 0 \\ 0\ , & 0\ , & m \end{array}\right),$$

or replacing the matrix on the left-hand side by unity, we may write

$$m = \left(\begin{array}{ccc} m, & 0, & 0 \\ 0, & m, & 0 \\ 0, & 0, & m \end{array}\right).$$

The matrix on the right-hand side is said to be the single quantity m considered as involving the matrix unity.

Detta avslutar Cayleys redogörelse för multiplikation mellan matris och skalär. Han kommer slutligen till matrissammansättning och illustrerar hur det fungerar med två matriser i 3x3-fallet (paragraf 11). Resonemanget bygger (implicit) på kompositionen av funktioner och associativitetesegenskapen hos sådana kompositioner. Cayley sammanfattar sitt resultat genom att specifikt beskriva hur den sammansatta matrisens element erhålles:

*It is to be observed, that the operation is not a commutative one; the component matrices may be distinguished as the first or further component matrix, and the second or nearer component matrix, and the rule of composition is as follows, viz. any **line** of the compound matrix is obtained by combining the corresponding **line** of the first or further component matrix successively with the several **columns** of the second or nearer compound matrix.*

Kommentar

Cayley verkar ha insett den icke-kommutativa egenskapen hos matrissammansättning genom att titta på exempel. Han observerade, som nämnt i citatet ovan, att ordningen på matriserna var avgörande för hur rader och kolonner kombinerades i den resulterande matrisen.Vidare ser det ut som att Cayley har definierat matrissammansättning med hjälp av funktioner, en definition som han applicerar på matrisen som objekt. Detta skiljer sig markant från hur matrissammansättning presenteras i moderna läroböcker: definitionen är mer direkt och presenteras i form av en regel. En sådan direkt definition är inte möjlig för Cayleys del, eftersom han inte har några andra auktoriteter att luta sig mot.

Kompletterande material

För att klargöra Cayleys regel för matrissammansättning kan vi titta på ett specifikt exempel. Låt

$$\mathrm{A} = \left(\begin{array}{cc} 1, & 2 \\ 2, & 2 \end{array}\right)$$

och

$$\mathrm{B} = \left(\begin{array}{cc} 1, & 3 \\ 4, & 1 \end{array}\right)$$

Vi har då att

$$\mathrm{A}\circ\mathrm{B} = \left(\begin{array}{cc} 1, & 2 \\ 2, & 2 \end{array}\right)\!\!\left(\begin{array}{cc} 1, & 3 \\ 4, & 1 \end{array}\right) = \left(\begin{array}{cc} (1,\ 2)(1,\ 4), & (1,\ 2)(3,\ 1) \\ (2,\ 2)(1,\ 4), & (2,\ 2)(3,\ 1) \end{array}\right)$$

Om vi nu ändrar ordningen på sammansättningen, det vill säga istället undersöker B∘A, ser vi att vi omöjligen kan få samma matris. I första steget bildar vi exempelvis storheten (1, 3)(1, 2) men detta är ju inte detsamma som (1, 2)(1, 4). *Fin.*

Cayley generaliserar i paragraf 14 sitt resultat för två matriser till att inkludera ett godtyckligt antal matriser. I samma veva konstaterar han även att matrissammansättning är en associativ operation:

We may in like manner multiply or compound together three or more matrices: the order of arrangement of the factors is of course material (...) any two consecutive factors may be compounded together and replaced by a single matrix, and so on until all the matrices are compounded together, the result being independent of the particular mode in which the composition is effected (...)

Kommentar

Som nämnt i inledningen är detta ett typiskt exempel på hur Cayley troliggör sina påståenden: undersök påståendet i ett specifikt fall (här sammansättning av två 3x3-matriser) och argumentera för att det även måste gälla i det allmänna fallet. Detta sätt att styrka påståenden kallar Thomas Hawkins för ”proof by generic example” och det används flitigt av Cayley genom arbetet.
Vi ser för övrigt hur icke-självklar matrissammansättning är för Cayley. En formulering som i citatet ovan skulle knappast dyka upp i moderna läroböcker, men Cayley rör sig ju bildligt talat genom en djungel av outforskad matematik - med språket som enda verktyg.

Cayley sammanfattar därefter sitt resultat genom att införa punktnotationen

$$\mathrm{L.MN=LM.N=LMN}$$

för associativitet mellan tre matriser.

2.2 Cayleys införande av potenser av matriser

I paragraf 15-20 börjar Cayley hantera potenser av matriser. Han definierar först potensen av en matris med en heltalsexponent och kommenterar även kommuterbarheten mellan sådana potenser av samma matris:

(...) it is to be observed that the different powers of the same matrix are convertible. It is also clear that p and q being positive integers, we have $\mathrm{L}^p.\mathrm{L}^q=\mathrm{L}^{p+q}$ *(...)*

Nu följer ytterligare ett antagande om att påståendet är möjligt att generalisera. Cayley skriver:

The last-mentioned equation, $\mathrm{L}^p.\mathrm{L}^q=\mathrm{L}^{p+q}$*, assumed to be true for all values whatever of the indices p and q, leads to the notion of the powers of the matrix of any form, whatever of the index (...)*

Kommentar

Enligt Cayley är matriser objekt som består av storheter (tal), så det förefaller honom rimligt att potenser med icke-positiva exponenter också existerar för matriser. Om vi accepterar Cayleys resonemang är det inte svårt att förstå varför exempelvis $L^0=1$, eller varför det är möjligt att prata om inversen till en matris L. Dessa är ju självklara egenskaper hos de reella, nollskilda talen - varför skulle de då inte vara lika självklara för matriser?

Determinanten vad vid denna tidpunkt (1800-talets mitt) sedan länge känd. Cayley kopplar i paragraf 17 determinanten för en 3x3-matris till existensen av en invers, givet en kvadratisk matris L. Han börjar med att explicit ge formeln för att hitta koefficienterna till inversen givet en 3x3-determinant (Cramers regel; Katz, s. 668). Eftersom denna algoritm bygger på att determinanten är nollskild drar Cayley inte helt oväntat följande slutsats i paragraf 18:

(...) the notion of the inverse or reciprocal matrix fails altogether when the determinant vanishes: the matrix is in this case said to be indeterminate (...)

Kommentar

Detta påminner om satsen för inverterbarhet hos *nxn*-matriser (Friedberg et. al; s. 223). Här är det förstås inte fråga om något bevis, utan snarare ett konstaterande från Cayleys sida. Om algoritmen, som antas vara sann, inte fungerar i den situation då determinanten är noll - kan vi heller inte använda algoritmen, och då måste följaktligen den givna matrisen sakna invers.

I paragraf 20 ser vi hur Cayley försöker koppla potenser av matriser till algebraiska tal (Cayley refererar specifikt i sitt arbete till "algebraical functions", men det är inte samma matematiska objekt som benämningen idag syftar på). Denna idé om en slags "algebraical function theorem" (ett begrepp myntat av Thomas Hawkins) är något som Cayley är övertygad om är nyckeln till att förstå matrisernas algebra. Han skriver:

(...) to preserve as far as possible the analogy with ordinary algebraical functions, we may restrict the attention to the case where the multiplier is a single quantity (...) We have in this manner a matrix cL^m, and by the addition of any number of such terms we obtain a rational and integral function of the matrix L.

Kommentar

Låt oss för ett ögonblick reflektera över varför Cayley var så övertygad om denna koppling mellan algebraiska funktioner och matriser.
Cayley definierade, som bekant, matriser med hjälp av storheter (läs tal). Precis som hans vän och samtalspartner Sylvester inspirerades av Gauss' och Cauchys hantering av linjära ekvationssystem och determinanter (Katz, s. 740-741); tänkte sig Cayley en matris som ett objekt som ordnar storheter på något förbestämt vis. Om dessa storheter nu är tal, i den vanliga meningen, är det rimligt att anta att matrisen i sig också ärver många av egenskaperna hos dessa tal. Kopplingen mellan matriser och algebraiska tal förefaller nu mer eller mindre naturlig i Cayleys ögon. Det återstår bara för honom att definiera dessa funktioner så att de delar så många av de "vanliga", algebraiska talens egenskaper som möjligt. Kommutativitet är en sådan egenskap.
Vi ser nu att Cayleys avgränsning i citatet ovan syftar till att styrka just denna koppling. Om vi bara betraktar termer av typen $c\mathrm{L}^m$, där c är en skalär, är det uppenbart att $c\mathrm{L}^m = \mathrm{L}^m c$. Termer av typen $\mathrm{L}^m.\mathrm{M}^n$ uppfyller generellt inte kravet på kommutativitet. Vi skulle alltså kunna uttrycka det som att Cayley medvetet begränsar sig till studiet av linjärkombinationer av potenser av en matris, för att stärka kopplingen till tal.

2.3 Karakteristiska ekvationen och Cayley-Hamiltons sats

I paragraf 21-24 introducerar Cayley en av sina största upptäcker i arbetet: det vi känner till som Cayley-Hamiltons sats. Den ordagranna förklaringen av satsens innebörd finner vi i inledningen till Cayleys arbete:

I obtain the remarkable theorem that any matrix whatever satisfies an algebraical equation of its own order, the coefficient of the highest power being unity, and those of the other powers functions of the terms of the matrix, the last coefficient being in fact the determinant (...)

I sitt troliggörande av satsen fokuserar Cayley enbart på 2x2-matriser, men han gör följande kommentar:

(...) I have not thought it necessary to undertake the labour of a formal proof of the theorem in the general case of a matrix of any degree.

För Cayley är det uppenbart att satsen går att generalisera till alla ändliga, kvadratiska matriser - det bara är en fråga om räknebördan.

Kommentar

En intressant kommentar till detta ges av Tony Crilly. Han påpekar att Cayley faktiskt antydde en generaliserad variant av satsen i ett brev till Sylvester år 1857 men att det är fullt möjligt att han antingen glömde bort denna korrespondans, eller bara avstod från att publicera den generella varianten (*Cayley's anticipation of a generalised Cayley-Hamilton theorem*, s. 213-214). Varför han skulle göra ett sådant val går bara att spekulera om. Kanske hade Cayley för avsikt att skapa ett bevis enbart om intresset för satsen växte sig tillräckligt stort eller om den visade sig användbar?

Hur övertygade sig då Cayley om satsens giltighet, i 2x2-fallet? Han börjar med att tänka sig en kvadratisk matris

$$\mathrm{M} = \left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right)$$

Eftersom M kan betraktas som en storhet (ett tal) kan vi uttrycka M på ett alternativt sätt:

$$\mathrm{M} = \left(\begin{array}{cc} \mathrm{M}, & 0 \\ 0, & \mathrm{M} \end{array}\right)$$

Cayley refererar till detta som *the single quantity M considered as involving the matrix unity.* Han bildar nu determinanten

$$\begin{vmatrix} a-\mathrm{M} & b \\ c & d-\mathrm{M} \end{vmatrix}$$

och beräknar den med någon för tiden känd metod. Han får då ett polynom på formen:

$$\mathrm{M}^2-(a+d)\mathrm{M}^1+(ad-bc)\mathrm{M}^0$$

Men storheterna M^2, M^1 och M^0 kan även uttryckas som matriser. Vi ser speciellt att

$$\mathrm{M}^2=\mathrm{M}.\mathrm{M}=\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right)\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right)=\left(\begin{array}{cc} (a, b)(a, c), & (a, b)(b, d) \\ (c, d)(a, c), & (c, d)(b, d) \end{array}\right)=\left(\begin{array}{cc} a^2+bc, & b(a+d) \\ c(a+d), & d^2+bc \end{array}\right),$$

$$\mathrm{M}^1=\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right),\ \mathrm{M}^0=\left(\begin{array}{cc} 1, & 0 \\ 0, & 1 \end{array}\right)$$

Vi kan nu uttrycka polynomet i termer av M^2, M^1 och M^0 och landar då i följande likhet:

$$\begin{vmatrix} a-\mathrm{M} & b \\ c & d-\mathrm{M} \end{vmatrix}=\left(\begin{array}{cc} a^2+bc, & b(a+d) \\ c(a+d), & d^2+bc \end{array}\right)-(a+d)\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right)+(ad-bc)\left(\begin{array}{cc} 1, & 0 \\ 0, & 1 \end{array}\right)$$

vilket efter förenkling ger nollmatrisen. Cayley ersätter nu nollmatrisen med talet 0 och vi ser då att

$$\begin{vmatrix} a-\mathrm{M} & b \\ c & d-\mathrm{M} \end{vmatrix}=0$$

Han förtydligar sitt resultat genom att först poängtera att denna determinants matris är på formen

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix} - \mathrm{M} \begin{pmatrix} 1, & 0 \\ 0, & 1 \end{pmatrix}$$

och sedan med följande kommentar, som återkopplar till hans presentation av upptäckten i början av arbetet:

And this is the general theorem, viz. the determinant, having for its matrix a given matrix less the same matrix considered as a single quantity involving the matrix unity, is equal to zero.

Denna något svårförstådda formulering skulle vi idag uttrycka som att en 2x2-matris, med reella koefficienter, satisfierar sin egen karakteristiska ekvation. Detta är, som tidigare indikerat, ett specialfall av Cayley-Hamiltons sats.

2.3.1 Ett bevis för Cayley-Hamiltons sats

Cayleys bevis är korrekt och är en indikation på att satsen går att generalisera till kvadratiska matriser av högre ordning än två. Det är dock inte trivialt att bevisa detta. För att illustrera det kan vi titta på ett bevis av Ian Kiming som bygger på polynom med matriser som koefficienter. Jag upplever det som ett passande bevis i sammanhanget, givet Cayleys intresse för funktioner.

Låt A vara en $n{\times}n$-matris, och låt

$$g(\lambda){=}b_m\lambda^m{+}b_{m-1}\lambda^{m-1}{+}...{+}b_1\lambda{+}b_0$$

vara ett polynom. Vi definierar nu

$$g(\mathrm{A}){=}b_m{\cdot}\mathrm{A}^m{+}b_{m-1}{\cdot}\mathrm{A}^{m-1}{+}...{+}b_1{\cdot}\mathrm{A}{+}b_0{\cdot}\mathrm{I}. \ (1)$$

Då är $g(\mathrm{A})$ en $n \times n$-matris med reella koefficienter. Givet att B_0, ..., B_m är $n{\times}n$-matriser kan vi betrakta (1) som ett 'polynom med matriskoefficienter':

$$\mathrm{G}(\lambda){=}\mathrm{B}_m{\cdot}\lambda^m{+}\mathrm{B}_{m-1}{\cdot}\lambda^{m-1}{+}...{+}\mathrm{B}_1{\cdot}\lambda{+}\mathrm{B}_0.$$

Vi ersätter λ med den kvadratiska matrisen A:

$$\mathrm{G}(\mathrm{A}){=}\mathrm{B}_m{\cdot}\mathrm{A}^m{+}\mathrm{B}_{m-1}{\cdot}\mathrm{A}^{m-1}{+}...{+}\mathrm{B}_1{\cdot}\mathrm{A}{+}\mathrm{B}_0. \ (2)$$

I det senare beviset kommer vi multiplicera polynom av samma typ som (2) med varandra. Det är därför relevant att titta på en situation då denna multiplikation kanske inte är så intuitiv som vi kanske först tror.

Låt A, B, C vara $n{\times}n$-matriser. Betrakta följande polynom med matriskoefficienter:

$$\mathrm{F}(\lambda){=}\lambda{+}\mathrm{B} \ , \ \mathrm{G}(\lambda){=}\lambda{+}\mathrm{C}.$$

Vi definierar nu

$$\mathrm{H}(\lambda)=\mathrm{F}(\lambda)\cdot\mathrm{G}(\lambda)=(\lambda+\mathrm{B})(\lambda+\mathrm{C})=\lambda^2+(\mathrm{B}+\mathrm{C})\lambda+\mathrm{BC}.$$

Vi ersätter som tidigare skalären λ med matrisen A. Då är

$$\mathrm{F}(\mathrm{A})\cdot\mathrm{H}(\mathrm{A})=(\mathrm{A}+\mathrm{B})(\mathrm{A}+\mathrm{C})=\mathrm{A}^2+\mathrm{AC}+\mathrm{BA}+\mathrm{BC}$$

och

$$\mathrm{H}(\mathrm{A})=\mathrm{A}^2+\mathrm{BA}+\mathrm{CA}+\mathrm{BC}$$

Om likheten

$$\mathrm{F}(\mathrm{A})\mathrm{G}(\mathrm{A})=\mathrm{H}(\mathrm{A})$$

ska gälla måste således

$$\mathrm{AC}=\mathrm{CA}.$$

Det är alltså *inte nödvändigtvis fallet* att F(A)G(A)=H(A). Detta måste vi ta i beaktande i fortsättningen.

Vi har nytta av följande lemma:

Lemma. Låt G(λ) vara ett polynom med $n\times n$-matriskoefficienter. Definiera:

$$\mathrm{H}(\lambda)=\mathrm{G}(\lambda)\cdot(\mathrm{I}\lambda-\mathrm{A}).$$

Då är H(A)=0.

Bevis. Med

$$\mathrm{G}(\lambda)=\mathrm{B}_m\cdot\lambda^m+\mathrm{B}_{m-1}\cdot\lambda^{m-1}+...+\mathrm{B}_1\cdot\lambda+\mathrm{B}_0$$

får vi att

$$\mathrm{H}(\lambda)=(\mathrm{B}_m\cdot\lambda^m+...+\mathrm{B}_0)\cdot(\mathrm{I}\lambda-\mathrm{A})$$
$$=\mathrm{B}_m\lambda^{m+1}+...+\mathrm{B}_1\lambda^2+\mathrm{B}_0\lambda-(\mathrm{B}_m\mathrm{A}\lambda^m+...+\mathrm{B}_1\mathrm{A}\lambda+\mathrm{B}_0\mathrm{A}).$$

Vi ersätter λ med matrisen A och ser då att

$$\mathrm{H}(\mathrm{A})=\mathrm{B}_m\mathrm{A}^{m+1}+...+\mathrm{B}_1\mathrm{A}^2+\mathrm{B}_0\mathrm{A}-(\mathrm{B}_m\mathrm{A}\mathrm{A}^m+...+\mathrm{B}_1\mathrm{AA}+\mathrm{B}_0\mathrm{A})$$

$$=\mathrm{B}_m\mathrm{A}^{m+1}+...+\mathrm{B}_1\mathrm{A}^2+\mathrm{B}_0\mathrm{A}-(\mathrm{B}_m\mathrm{A}^{m+1}+...+\mathrm{B}_1\mathrm{A}^2+\mathrm{B}_0\mathrm{A})=0.$$

$\square$

Slutligen behöver vi definiera minor, kofaktormatris och adjungerad matris. Se appendix (A1) för detta.

Sats. (Cayley-Hamilton) Låt A vara en $n\times n$-matris med karakteristiskt polynom $f(\lambda)$. Då gäller att

$$f(\mathrm{A})=0.$$

Bevis.
Vi utgår från identiteten

$$\text{adj.(A)}\cdot\text{A}=\det(\text{A})\cdot\text{I}$$

och ersätter A med (Iλ−A):

$$\text{adj.}(\text{I}\lambda-\text{A})\cdot(\text{I}\lambda-\text{A})=\det(\text{I}\lambda-\text{A})\cdot\text{I}.$$

Men det(Iλ−A) är ju inget annat än det karakteristiska polynomet $f(\lambda)$. Med

$$f(\lambda)=\lambda^n+\alpha_{n-1}\cdot\lambda^{n-1}+...+\alpha_0$$

ser vi nu att

$$\text{adj.}(\text{I}\lambda-\text{A})\cdot(\text{I}\lambda-\text{A})=\text{I}\lambda^n+\alpha_{n-1}\text{I}\lambda^{n-1}+...+\alpha_0\text{I}:=\text{H}(\lambda)$$

så vi kan tolka H(λ) som ett polynom med $n\times n$-matriskoefficienter.

Matrisen adj.(Iλ−A) måste nu vara en $n\times n$-matris vars element alla är polynom. Vi kan därför tolka adj.(Iλ−A) som ett polynom G(λ) med $n\times n$-matriskoefficienter. Om vi tillämpar lemmat måste H(A)=0, eller

$$\text{A}^n+\alpha_{n-1}\text{A}^{n-1}+...+\alpha_0\text{I}=0$$

□ *Fin.*

2.4 Cayleys hantering av potenser av matriser med rationella exponenter

Paragraf 24-25 utgör i min mening en av de mest intressanta delarna i arbetet. Här försöker Cayley övertyga både sig själv och läsaren om att det finns en koppling mellan funktioner och matriser som är mer allmänn, med Cayley-Hamilton som en första indikation på detta. Thomas Hawkins belyser Cayleys övertygelse i följande citat (The Mathematics of Frobenius in context; s. 216):

”The significance of the Cayley-Hamilton theorem was, in Cayley's view, that it showed that any 'algebraical function' of a matrix M is equal to a polynomial in M of degree less than the order of M”.

Kommentar

Hawkins använder notationen L=p(M), där L är en godtycklig kvadratisk matris och $p(x)$ är en 'algebraisk funktion'. Han visar att påståendet som formulerat ovan stämmer i fallet då p(M) är en polynomfunktion eller då p(M) är en rationell funktion (för detaljer, se Hawkins arbete). Cayley påstår ju dock att det även ska vara möjligt att utvdiga resonemanget till rotfunktioner, exempelvis $f(x)=\sqrt{x}$: han är övertygad om att L=p(M) existerar även i detta fall. Hawkins formulerar detta som Cayleys tro på en slags ”algebraical function theorem”, men påpekar i samma stycke att det *inte* är fråga om en sats och hänvisar till Frobenius' omfattande utläggningar kring detta.

Låt oss undersöka Cayleys argumentation. Cayley skriver i paragraf 26 att han vill illustrera sitt resonemang i ett specifikt fall: målet är att finna L så att $\mathrm{L}=\sqrt{\mathrm{M}}$, där M är en given kvadratisk 2x2-matris. Matrisen eller storheten M satisfierar ekvationen

$$\mathrm{M}^2-(a+d)\mathrm{M}^1+(ad-bc)\mathrm{M}^0$$

Men L är enligt antagandet också en kvadratisk matris. Låt

$$\mathrm{L}=\begin{pmatrix} \alpha, & \beta \\ \gamma, & \delta \end{pmatrix},\ \mathrm{M}=\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}$$

Då måste L satisfiera en ekvation, nämligen

$$\mathrm{L}^2-(\alpha+\beta)\mathrm{L}+\alpha\delta-\beta\gamma=0$$

Cayley använder sig nu av implikationen

$$\mathrm{L}=\sqrt{\mathrm{M}} \Rightarrow \mathrm{L}^2=\mathrm{M}$$

för att uttrycka L som en 'linjär funktion av M'. Med $\mathrm{L}^2=\mathrm{M}$ kan den karakteristiska ekvationen för L skrivas

$$\mathrm{M}-(\alpha+\beta)\mathrm{L}+(\alpha\delta-\beta\gamma)=0$$

Eftersom $(\alpha\delta-\beta\gamma)$ är en storhet, precis som M, är uttrycket ovan inget annat än en vanlig linjär ekvation i Cayleys ögon. Han löser nu ut L:

$$\mathrm{L}=\frac{1}{\alpha+\beta}\left[\mathrm{M}+(\alpha\delta-\beta\gamma)\right]$$

Med substitutionen $\mathrm{X}=\alpha+\delta$, $\mathrm{Y}=\alpha\delta-\beta\gamma$ får han ekvationen

$$\mathrm{L}=\begin{pmatrix} \alpha, & \beta \\ \gamma, & \delta \end{pmatrix}=\begin{pmatrix} \frac{a+\mathrm{Y}}{\mathrm{X}}, & \frac{b}{\mathrm{X}} \\ \frac{c}{\mathrm{X}}, & \frac{d+\mathrm{Y}}{\mathrm{X}} \end{pmatrix}$$

Nu utelämnar Cayley mycket av den algebra som ligger bakom den slutsats som följer. Jag anser det därför lämpligt att här fylla i dessa luckor, så att även en icke-insatt läsare ska förstå hans resonemang. För den icke-intresserade läsaren hänvisar jag till sida 16.

Kompletterande material

Cayley gör troligtvis, i detta läge, en koefficientanalys av matrisernas koefficienter. Vi ser att

$$\alpha=\frac{a+\mathrm{Y}}{\mathrm{X}},\ \beta=\frac{b}{\mathrm{X}},\ \gamma=\frac{c}{\mathrm{X}},\ \delta=\frac{d+\mathrm{Y}}{\mathrm{X}}$$

Då kan vi uttrycka sambandet mellan koefficienterna i respektive matris på följande vis:

$$\mathrm{X}=\frac{(a+d)+2\mathrm{Y}}{\mathrm{X}},\ \mathrm{Y}=\frac{(a+\mathrm{Y})(d+\mathrm{Y})-bc}{\mathrm{X}^2}=\frac{\mathrm{Y}(a+d+\mathrm{Y})+(ad-bc)}{\mathrm{X}^2}$$

Med ytterligare en substitution: $\mathrm{P}{=}a{+}d$, $\mathrm{Q}{=}ad{-}bc$ får vi följande:

$$\mathrm{X}{=}\frac{\mathrm{P}{+}2\mathrm{Y}}{\mathrm{X}}\Rightarrow\mathrm{X}^2{=}\mathrm{P}{+}2\mathrm{Y},$$

$$\mathrm{Y}{=}\frac{\mathrm{Y}(\mathrm{P}{+}\mathrm{Y}){+}\mathrm{Q}}{\mathrm{X}^2}{=}\frac{\mathrm{Y}(\mathrm{P}{+}\mathrm{Y}){+}\mathrm{Q}}{\mathrm{P}{+}2\mathrm{Y}}$$

vilket implicerar att

$$\mathrm{YP}{+}2\mathrm{Y}^2{=}\mathrm{YP}{+}\mathrm{Y}^2{+}\mathrm{Q}$$

eller $\mathrm{Y}{=}\pm\sqrt{\mathrm{Q}}$. Vi ser nu att

$$\mathrm{Y}{=}\sqrt{\mathrm{Q}}\Rightarrow\mathrm{X}^2{=}\mathrm{P}{+}2\sqrt{\mathrm{Q}}\Rightarrow\mathrm{X}{=}\sqrt{\mathrm{P}{+}2\sqrt{\mathrm{Q}}}$$

om vi endast tar hänsyn till den positiva roten till ekvationen $\mathrm{Y}^2{=}\mathrm{Q}$ (det finns ju annars möjligheten att diskriminanten ovan antar ett negativt värde). Eftersom värdena P och Q är kända, är nu även värdena på X och Y kända - och följaktligen även värdena på koefficienterna α, β, γ och δ.

Kompletterande material (2)

Vi undersöker Cayleys formel för $\mathrm{L}{=}\sqrt{\mathrm{M}}$ då M är, som Cayley skulle ha uttryckt det ”determinerad” - det vill säga, då $\det(\mathrm{M})\neq 0$. Tag exempelvis matrisen

$$\mathrm{M}{=}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$$

Med X och Y definierade som ovan är

$$\mathrm{L}{=}\begin{pmatrix}\frac{a+\mathrm{Y}}{\mathrm{X}} & \frac{b}{\mathrm{X}}\\ \frac{c}{\mathrm{X}} & \frac{d+\mathrm{Y}}{\mathrm{X}}\end{pmatrix}{=}\begin{pmatrix}\frac{1+\sqrt{2}}{\sqrt{3+2\sqrt{2}}} & 0\\ 0 & \frac{2+\sqrt{2}}{\sqrt{3+2\sqrt{2}}}\end{pmatrix}$$

en lösning till ekvationen. Detta kan vi lätt kontrollera genom att undersöka huruvida $\mathrm{L}^2{=}\mathrm{M}$:

$$\mathrm{L}^2{=}\begin{pmatrix}\frac{1+\sqrt{2}}{\sqrt{3+2\sqrt{2}}} & 0\\ 0 & \frac{2+\sqrt{2}}{\sqrt{3+2\sqrt{2}}}\end{pmatrix}\begin{pmatrix}\frac{1+\sqrt{2}}{\sqrt{3+2\sqrt{2}}} & 0\\ 0 & \frac{2+\sqrt{2}}{\sqrt{3+2\sqrt{2}}}\end{pmatrix}$$

$$={\begin{pmatrix}\frac{3+2\sqrt{2}}{3+2\sqrt{2}} & 0\\ 0 & \frac{2(3+2\sqrt{2})}{3+2\sqrt{2}}\end{pmatrix}}{=}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}{=}\mathrm{M}.$$

Är detta ekvationens enda lösning? Nej! Enligt en känd sats från den linjära algebran kan vi ju bestämma potenser till en diagonalmatris genom att ta potensen av diagonalelementen. Vi kan visa att en fullständig lösning (se appendix: A6) ges av

$$\mathrm{L}{=}\begin{pmatrix}\pm 1 & 0\\ 0 & \pm\sqrt{2}\end{pmatrix}.$$

Det visar sig att Cayleys metod ger oss exakt den matris vars diagonalelement är positiva. Vi observerar nämligen att

$$3+2\sqrt{2}=\left(\sqrt{2}+1\right)^2 \Rightarrow \sqrt{3+2\sqrt{2}}=\sqrt{2}+1$$

vilket låter oss göra följande omskrivning av den med Cayleys formel funna matrisen:

$$\mathrm{L}=\begin{pmatrix} \frac{1+\sqrt{2}}{\sqrt{3+2\sqrt{2}}} & 0 \\ 0 & \frac{2+\sqrt{2}}{\sqrt{3+2\sqrt{2}}} \end{pmatrix}=\begin{pmatrix} \frac{1+\sqrt{2}}{\sqrt{2}+1} & 0 \\ 0 & \frac{2+\sqrt{2}}{\sqrt{2}+1} \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{2} \end{pmatrix} .$$

Med modern, linjär algebra är det för övrigt möjligt att bestämma L även i det fall då M är icke-determinerad (se till exempel Wikipedias artikel: "Square root of a matrix", för en utförlig beskrivning av detta).

Kompletterande material (3)

En alternativ definition av roten ur en matris (som också känns lite mer tidsenlig) kan vi finna om vi betraktar Maclaurinserien för $\sqrt{1+x}$. Med x som en kvadratisk matris A, skulle vi kunna skriva

$$\sqrt{1+\mathrm{A}}=\mathrm{I}+\frac{\mathrm{A}}{2}-\frac{\mathrm{A}^2}{8}+\frac{\mathrm{A}^3}{16}-...$$

I fallet då A är en diagonalmatris fungerar detta relativt bra, eftersom vi enkelt kan undersöka potenser av dessa typer av matriser och därmed avgöra huruvida summan i högerledet konvergerar eller ej. Att Cayley inte använde sig av denna definition beror förstås på den splittring som fortfarande rådde mellan analys och algebra under den här tiden (Katz; s. 710 resp. s. 765): serier av det här slaget hörde till analysen och behandlades enbart inom denna gren av matematiken. *Fin.*

Cayley demonstrerar i ytterligare ett exempel hur hans metod att finna en matris, givet en potens av en annan matris fungerar. Utgångspunkten i paragraf 27 är nu ekvationen $\mathrm{L}^2=\mathrm{M}^2$, med M definierad som ovan. Här är algebran snarlik föregående exempel: för den intresserade läsaren hänvisar jag till appendix: A5. Kort sagt landar Cayley i en falluppdelning, där den första möjligheten är att

$$\mathrm{L}=\pm\mathrm{M},$$

den andra möjligheten att

$$\mathrm{L}=\frac{a+d}{\mathrm{X}}\mathrm{M}-\frac{2(ad-bc)}{\mathrm{X}}$$

med följande kommentar:

It may be remarked that if $\mathrm{M}^2=1$, *the last-mentioned formula fails, for we have X=0; it will be seen presently that the equation* $\mathrm{L}^2=1$ *admits of other solutions besides* $\mathrm{L}=\pm1$.

Kommentar

Eftersom X är kopplat till en variabelsubstitution fungerar ovan och tidigare nämnda lösningsformel inte för alla matriser M. Detta väckte uppenbarligen Cayleys intresse för ekvationer av typen $\mathrm{M}^n=1$, att döma av de paragrafer som följer i hans arbete.
I det här läget är det också hög tid att nämna att koefficienterna $a+d$ och $ad-bc$, som så ofta dyker upp i Cayleys beräkningar, representerar det vi idag kallar för en matris 'spår' respektive 'determinant' (den sistnämnda har vi redan stiftat bekantskap med). Att substitutionerna för P och Q ser ut som de gör är alltså knappast en slump: Cayley har säkerligen stött på dessa koefficienter i samband med skrivandet av sitt arbete om determinanter från 1843, "On a theory of determinants" (se MacTutor).

I paragraf 28 utreder Cayley vad han kallar för en "uppenbar svårighet" förknippad med den karakteristiska ekvationen. Han noterar att matrisen

$$\mathrm{M}=\left(\begin{array}{cc} a, & b \\ c, & d \end{array}\right)$$

som bekant satisfierar ekvationen

$$\mathrm{M}^2-(a+d)\mathrm{M}+ad-bc=0$$

kan antas ha rötterna X_1, X_2. Han drar nu (implicit) en parallell till faktorsatsen och konstaterar följande:

The equation satisfied by the matrix may be written

$$(\mathrm{M}-\mathrm{X}_i)(\mathrm{M}-\mathrm{X}_{ii})=0$$

(...) and it would at first sight seem that we ought to have one of the simple factors equal to zero, which is obviously not the case (...) The explanation is that each of the simple factors is an indeterminate matrix (...) The product of the two factors is thus equal to zero without either of the factors being equal to zero.

Kommentar

Ekvationen ovan är en kompakt omskrivning av matrisekvationen:

$$\left(\begin{array}{cc} a-\mathrm{X}_1, & b \\ c & , \ d-\mathrm{X}_1 \end{array}\right)\left(\begin{array}{cc} a-\mathrm{X}_2, & b \\ c & , \ d-\mathrm{X}_2 \end{array}\right)=\left(\begin{array}{cc} 0, & 0 \\ 0, & 0 \end{array}\right)$$

Vi kan direkt konstatera att (i) ingendera av matriserna ovan behöver vara nollmatrisen för att produkten ska bli nollmatrisen, och (ii) determinanten för respektive matris i vänsterledet måste nödvändigtvis vara noll - precis som Cayley påstår.

Kompletterande material

Vi undersöker Cayleys påstående i ett specifikt exempel. Matrisen

$$\mathrm{M}=\begin{pmatrix} 1, & 0 \\ 0, & 2 \end{pmatrix}$$

uppfyller en ekvation, vars rötter är $\mathrm{X}_1=1$, $\mathrm{X}_2=2$. Vi ser att

$$(\mathrm{M}-\mathrm{X}_1)(\mathrm{M}-\mathrm{X}_2)=$$

$$=\begin{pmatrix} 0, & 0 \\ 0, & 1 \end{pmatrix}\begin{pmatrix} -1, & 0 \\ 0 & , 0 \end{pmatrix}=\begin{pmatrix} 0, & 0 \\ 0, & 0 \end{pmatrix}=0,$$

där vi har använt Cayleys beteckning för nollmatrisen i sista steget. *Fin.*

I paragraf 30 börjar Cayley utforska ekvationer på formen $\mathrm{M}^n=1$, med n som ett heltal. Cayley introducerar läsaren genom att anknyta till den karakteristiska ekvationen. Han inför även begreppet periodicitet:

The equation satisfied by the matrix may be of the form $\mathrm{M}^n=1$*; the matrix is in this case said to be periodic of the n:th order. The preceding considerations apply to the theory of periodic matrices (...)*

Precis som tidigare väljer Cayley nu att utgå ifrån en 2x2-matris, i förhoppningen om att teorin ska kunna utvidgas till större kvadratiska matriser. Antag att vi vill finna en matris M, sådan att M är periodisk av ordning 2. Med M definierad som tidigare vet Cayley att

$$\mathrm{M}^2-(a+d)\mathrm{M}+ad-bc=0$$

Han inför nu även villkoret att M ska vara periodisk av ordning 2:

$$\mathrm{M}^2=1 \Leftrightarrow \mathrm{M}^2-1=0$$

Cayley gör en koefficientanalys och konstaterar att ekvationerna är identiska om

$$a+d=0, \ ad-bc=-1$$

Han formulerar nu sin slutsats, med en antydan om en algoritm för att finna lösningar till ekvationer av typen $\mathrm{M}^n=1$:

(...) And in like manner the matrix M of the order 2 will satisfy the condition $\mathrm{M}^3-1=0$*, or vill be periodic of the third order, if only* M^3-1 *contains as a factor* $\mathrm{M}^2-(a+d)\mathrm{M}+ad-bc$*, and so on.*

Kompletterande material

Låt oss för intressets skull undersöka hur vi med Cayleys metod skulle finna matrisen M, så att $\mathrm{M}^3-1=0$. Antag att M är en 2x2-matris. Matrisen M måste nu uppfylla ekvationerna

$$\mathrm{M}^2-(a+d)\mathrm{M}+ad-bc=0,\ \ \mathrm{M}^3-1=0$$

Vi har att

$$\begin{gathered}\mathrm{M}^3=\mathrm{M}.\mathrm{M}^2=\\ =\mathrm{M}\big((a+d)\mathrm{M}-(ad-bc)\big)=\\ =(a+d)\mathrm{M}^2-(ad-bc)\mathrm{M}\\ =(a+d)\big(a+d)\mathrm{M}-(ad-bc)\big)-\mathrm{M}(ad-bc)\\ =\big[(a+d)^2-(ad-bc)\big]\mathrm{M}-(a+d)(ad-bc)\end{gathered}$$

Vi kan nu identifiera och jämföra koefficienterna. Följande samband måste gälla:

$$(a+d)^2-(ad-bc)=0,\ \ (a+d)(ad-bc)=-1$$

Cayley skulle troligtvis vara nöjd när vi har kommit såhär långt. Vi kan ju nu påstå att matrisen M's koefficienter är kända, även om vi som tidigare får en parameterlösning. Dessutom har vi uttryckt M^3 i termer av en linjär funktion, om vi betraktar M som en variabel, och därmed yrkat på en koppling mellan matriser och funktioner. *Fin.*

Kommentar
Den metod som Cayley beskriver har en allmän idé bakom sig: om vi vill lösa ekvationen $p(\mathrm{M})=0$, där M är en nxn-matris och $p(x)$ är ett polynom får vi villkor på den karakteristiska ekvationen. Dessa villkor kan redas ut med metoder från teorin kring polynomekvationer, något som tyvärr ligger utanför ramen för detta arbete. Detta indikerar dock ännu en gång hur fruktbart Cayleys arbete faktiskt är.

2.5 Cayleys definition av kommuterande matriser och hans försök att finna villkor för att två matriser ska kommutera.

Cayley presenterar i paragraf 33 begreppet "konverterbarhet" (eller kommutativitet, som vi idag skulle uttrycka det). Problemet är följande: Givet en kvadratisk matris M, finn de kvadratiska matriser L som uppfyller att LM=ML. Det vill säga, finn L så att L och M kommuterar. Cayley resonerar på följande vis:

But whatever the form of the function is, it may be reduced to a rational and integral function of an order equal to that of M, less unity, and we have thus the general expression for the matrices convertible with a given matrix, viz.

any such matrix is a rational and integral function (the coefficients being single quantities) of the given matrix, the order being that of the given matrix, less unity. In particular, the general form of the matrix L convertible with a given matrix M of the order 2, is $L = \alpha\mathrm{M}+\beta$, *or what is the same thing, the matrices*

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}, \begin{pmatrix} a', & b' \\ c', & d' \end{pmatrix}$$

will be convertible if $a'-d' : b' : c' = a-d : b : c$.

Kompletterande material

Låt oss klargöra Cayleys tankegångar genom att titta på följande tillrättalagda exempel. Antag att vi har en känd matris M och önskar bestämma L, så att L och M kommuterar. Låt

$$\mathrm{M}=\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}, \ \mathrm{L}=\begin{pmatrix} w, & x \\ y, & z \end{pmatrix}$$

Eftersom LM = ML kan vi skriva att

$$\begin{pmatrix} w, & x \\ y, & z \end{pmatrix}\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}=\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}\begin{pmatrix} w, & x \\ y, & z \end{pmatrix}$$

eller ekvivalent

$$\begin{pmatrix} aw+cx, & bw+dx \\ ay+cz, & by+dz \end{pmatrix}=\begin{pmatrix} aw+by, & ax+bz \\ cw+dy, & cx+dz \end{pmatrix}$$

Ovanstående kan efter lite algebraisk manipulation visas implicera följande:

$$\frac{y}{c}=\frac{x}{b}=\alpha,$$

$$z-\frac{dy}{c}=w-\frac{ay}{c}=\beta$$

För dessa värden på α och β kan vi undersöka Cayleys påstående. Vi ser att

$$\alpha\mathrm{M}+\beta=\begin{pmatrix} \alpha a, & \alpha b \\ \alpha c, & \alpha d \end{pmatrix}+\begin{pmatrix} \beta, & 0 \\ 0, & \beta \end{pmatrix}$$

$$=\begin{pmatrix} ya/c+(w-ay/c), & by/c \\ y & , \ yd/c+(z-dy/c) \end{pmatrix}=\begin{pmatrix} w, & x \\ y, & z \end{pmatrix}=\mathrm{L}.$$

Kommentar

Vi känner igen Cayleys sätt att resonera: utgå ifrån en specifik situation och argumentera för att slutsatsen gäller, även i det generella fallet. Liksom i fallet med "algebraical function theorem" tar Cayley tyvärr inte hänsyn till extremfall. Formeln ovan blir exempelvis meningslös då b eller c är 0. Thomas Hawkins påpekar detta och kommenterar samtidigt en annan aspekt som vi har stött på tidigare i arbetet (The Math of Frobenius in context, s. 218):

> "(...) his reasoning is also based on incomplete induction, with general conclusions extrapolated from the (generic) treatment of matrices of small size."

Att poängtera Cayleys ofullständiga resonemang genom att hänvisa till ofullständig induktion är dock knappast rättvist gentemot Cayley: matematisk induktion existerande inte i den form vi känner till idag vid tidpunkten av Cayleys skrivande. Vi bör kanske snarare bektrakta Cayleys hantering av kommuterande matriser som ett första steg mot en mer allmän teori kring dessa typer av matriser (Katz, s. 744).

Kompletterande material (2)

Det räcker att undersöka ett fåtal olika typer av matriser för att inse att problemet att bestämma ett villkor för kommutativitet tyvärr inte är så enkelt som Cayley verkar antyda.

(i) Enhetsmatrisen kommuterar med alla matriser av samma storlek.

(ii) Diagonalmatriser kommuterar med andra diagonalmatriser. Vi har till exempel att

$$\begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 8 & 0 \\ 0 & 5 \end{pmatrix}.$$

(iii) Så kallade *cykliska* matriser kommuterar alltid med varandra. Exempel på sådana är

$$\mathrm{A} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix}, \ \mathrm{B} = \begin{pmatrix} 0 & 1 & 2 \\ 2 & 0 & 1 \\ 1 & 2 & 0 \end{pmatrix}.$$

Vi ser att

$$\mathrm{AB} = \mathrm{BA} = \begin{pmatrix} 7 & 7 & 4 \\ 4 & 7 & 7 \\ 7 & 4 & 7 \end{pmatrix}.$$

(iv) För att ta ett komplext exempel: två *Hermitiska* matriser (se appendix: A2) av ordning 2 kommuterar om de har ett gemensamt egenrum. Speciellt gäller att två sådana Hermitiska matriser med distinkta egenvärden kommuterar om de har samma egenvektorer. Ett exempel på ett par av sådana matriser är

$$\mathrm{M} = \begin{pmatrix} 1 & 1+i \\ 1-i & 1 \end{pmatrix}, \ \mathrm{N} = \begin{pmatrix} -1 & 1+i \\ 1-i & -1 \end{pmatrix}$$

ty vi ser att

$$\text{MN}=\text{NM}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Vi noterar att (i) och (ii), som med moderna ögon skulle uppfattas som tämligen triviala situationer då matriser kommuterar, inte kan karakteriseras med Cayleys villkor.

Om vi endast är ute efter att *finna* en matris som kommuterar med en given, kvadratisk matris kan vi använda oss av en metod som bygger på egenvärden och egenvektorer. För den intresserade läsaren hänvisar jag till appendix: A3. *Fin.*

2.6 Upptakten till Cayleys algoritm för att finna potenser av matriser

Mot slutet av arbetet redogör Cayley även för transponatet till en kvadratisk matris. Specifikt inför han notationen

$$\text{tr.}\left(\begin{array}{cc} a, & b\\ c, & d\end{array}\right)=\left(\begin{array}{cc} a, & c\\ b, & d\end{array}\right)$$

och räknelagar för transponat. Cayley skriver:

It is easy to see that if M be any [square] matrix, then

$$\big(\text{tr. M}\big)^p=\text{tr.}\big(\text{M}^p\big),$$

and in particular,

$$\big(\text{tr. M}\big)^{-1}=\text{tr.}\big(\text{M}^{-1}\big)$$

Regeln för beräknandet av en produkts transponat hanteras på liknande vis.

Kommentar

Vi noterar att detta självfallet inte är en fullständig teori kring matriser och deras transponat. En för Cayley kanske uppenbar egenskap, som han tydligen inte ansåg nödvändig att nämna, är hur vi hanterar transponatet av en summa:

$$\mathrm{tr.}\big(\mathrm{A}+\mathrm{B}\big)=\mathrm{tr.}\big(\mathrm{A}\big)+\mathrm{tr.}\big(\mathrm{B}\big)$$

Givet Cayleys kunskaper om och intresse för determinanter kan det också verka lite märkligt att han inte utvidgar teorin om transponat till att även innefatta just determinanter. Som tidigare anmärkt ville Cayley gärna undvika "räknetunga onödigheter", men att bevisa följande påstående:

$$\mathrm{Det.}\Big(\mathrm{tr.}\big(\mathrm{A}\big)\Big)=\mathrm{Det.}\Big(\mathrm{A}\Big)$$

är knappast räknetungt. Cayley skulle ha kunnat hantera det på samma sätt som övriga påståenden i artikeln: bektrakta ett specifikt exempel, såsom en 2x2-matris med givna koefficienter, och visa att det gäller här - för att sedan "naivt" generalisera till alla ändliga, kvadratiska matriser. En invändning skulle kunna vara att Cayley bedömde att dessa egenskaper inte var nödvändiga för artikeln och att han därför utelämnade dem. En annan möjlighet är att han kanske redan har hanterat detta i sitt tidigare arbete om determinanter (ett arbete som författaren till denna artikel dessvärre inte haft tid att analysera).

I paragraf 40 definierar Cayley en symmetrisk matris som en matris som inte ändras under transposition. Han definierar även skevt symmetriska matriser:

A matrix (...) which by transposition is changed into its opposite, is said to be skew symmetrical.

och ger ett exempel på hur koefficienterna i en sådan matris kan se ut i 3x3-fallet. Syftet med införandet av ovanstående begrepp antyds av Cayley i följande citat:

It is easy to see that any matrix whatever may be expressed as the sum of a symmetrical matrix, and a skew symmetrical matrix (...)

I paragraf 43 illustrerar Cayley en egenskap för sammansättningen av en kvadratisk matris med dess transponat. Han noterar att

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}\begin{pmatrix} a, & c \\ b, & d \end{pmatrix}=\begin{pmatrix} a^2+b^2, & ac+bd \\ ac+bd, & c^2+d^2 \end{pmatrix}$$

(...) which shows that a matrix compounded with the transposed matrix gives rise to a symmetrical matrix.

På sedvanligt vis resonerar han sig sedan fram till att ett jämnt antal sådana kompositioner måste leda till en symmetrisk matris. Cayley skriver:

It is needless to proceed further, since it is clear that

$$\begin{pmatrix} a, & c \\ b, & d \end{pmatrix}\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}\begin{pmatrix} a, & c \\ b, & d \end{pmatrix}\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}=\left(\begin{pmatrix} a, & c \\ b, & d \end{pmatrix}\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}\right)^2$$

och verkar då antyda att denna resulterande matris är symmetrisk. Detta känner vi igen som en av egenskaperna hos symmetriska matriser: givet att matrisen A är symmetrisk är även A^n symmetrisk.

Sista delen av Cayleys arbete (paragraf 46-51) återkopplar till potenser av matriser. Han börjar med att undersöka hur vi utifrån en given kvadratisk matris M kan finna dess potenser. Han utgår som tidigare från den karakteristiska ekvationen för en 2x2-matris och härleder utifrån denna först potenser av M med positiv heltalsexponent, troligtvis genom omskrivningar av typen:

$$\mathrm{M}^3=\mathrm{M}.\mathrm{M}^2,$$

$$\mathrm{M}^4=\mathrm{M}^2.\mathrm{M}^2$$

Jag säger ”troligtvis” eftersom detaljerna har utelämnats av Cayley.

Kompletterande material

Om vi sätter oss in i Cayleys tankebanor för ett ögonblick är det nu inte orimligt att även utveckla resonemanget till att innefatta negativa heltalsexponenter. Låt

$$\mathrm{M}=\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}$$

Nu uppfyller M ekvationen

$$\mathrm{M}^2-(a+d)\mathrm{M}+(ad-bc)=0$$

eller

$$\mathrm{M}^2=(a+d)\mathrm{M}^1-(ad-bc)$$

Multiplikation med M^{-1} från höger i bägge led ger oss att:

$$\mathrm{M}^2.\mathrm{M}^{-1}=(a+d)\mathrm{M}^1.\mathrm{M}^{-1}-(ad-bc)\mathrm{M}^{-1}$$

Cayley har tidigare etablerat räkneregler för potenser av matriser. Applicerar vi dessa på uttrycken i respektive led ovan får vi att

$$\mathrm{M}=(a+d)-(ad-bc)\mathrm{M}^{-1} \Leftrightarrow (ad-bc)\mathrm{M}^{-1}=-\mathrm{M}+(a+d)$$

Om $(ad-bc) \neq 0$ kan vi alltså bestämma den inversa matrisen med denna formel. Cayley konstaterar att det är nu också är möjligt att bestämma övriga negativa potenser:

(...) the other negative powers of M can then be obtained by successive multiplications with M^{-1}.

Kompletterande material (2)

Vi förtydligar Cayleys resonemang med ett specifikt exempel. Låt

$$\mathrm{M}=\begin{pmatrix} 1, & 2 \\ 3, & 4 \end{pmatrix}$$

Vi noterar att

$$\mathrm{Det.(M)}=1\cdot 4-2\cdot 3=-2\neq 0$$

Vi kan nu använda oss av formeln ovan med a=1, b=2, c=3, d=4. Det leder till följande kedja av ekvivalenta ekvationer:

$$(1\cdot 4-2\cdot 3)\mathrm{M}^{-1}=-\begin{pmatrix} 1, & 2 \\ 3, & 4 \end{pmatrix}+(1+4)$$

$$(-2)\mathrm{M}^{-1}=-\begin{pmatrix} 1, & 2 \\ 3, & 4 \end{pmatrix}+\begin{pmatrix} 5, & 0 \\ 0, & 5 \end{pmatrix}$$

$$(-2)\mathrm{M}^{-1}=\begin{pmatrix} 4 & , -2 \\ -3, & 1 \end{pmatrix}$$

Multiplikation med $-1/2$ i bägge led ger oss slutligen att

$$\mathrm{M}^{-1}=\begin{pmatrix} -2\,, & 1 \\ 3/2, & -1/2 \end{pmatrix}$$

Fin.

Kommentar

Det är enkelt att kontrollera att detta verkligen är den inversa matrisen till M. Potenser av typen M^{-2}, M^{-3}, M^{-4}... kan nu beräknas rekursivt: basfallet M^{-1} låter oss beräkna M^{-2}=M^{-1}.M^{-1}. Detta låter oss i sin tur beräkna M^{-3}, och så vidare.

Cayley insåg antagligen att detta snabbt leder till tidskrävande beräkningar. Betrakta till exempel potensen M^{-100}, med M definierad som ovan. Först måste vi beräkna M^{-1}, följt av M^{-2}. När M^{-2} är bestämd ser vi att det krävs 50 sådana faktorer för att bilda den potens vi söker. Det är i detta fall möjligt att reducera problemet till 'endast' 25 faktorer på formen M^{-4}, men då kan det ju knappast vara fråga om någon effektiv algoritm.

Cayley föreslår istället en algoritm som bygger på det förhållande han resonerade sig fram till i samband med problemet att finna villkor på kommuterande matriser. Sättet på vilket Cayley kommer fram till algoritmen är relativt svårt att förstå: det involverar ett flertal variabelsubstitutioner och den trigonometriska funktionen cotangens.

Cayley börjar med att associera talen h, b, c med storheterna J, q och talen h', b', c' med storheterna J$'$, q' så att

$$h{:}b{:}c = h'{:}b'{:}c'.$$

(Om vi läser mellan raderna så verkar Cayley påstå att det existerar ett tal t, så att $h'{=}th,\ b'{=}tb,\ c'{=}tc$). Nu definierar han

$$\mathrm{R}{=}-h^2{-}4bc,\ \mathrm{R}'{=}-h'^2{-}4b'c'$$

och konstruerar därefter matrisen

$$(\mathrm{J},\ q) := \begin{pmatrix} \mathrm{J}\left(\cot q - \frac{h}{\sqrt{\mathrm{R}}}\right), & \frac{2b\mathrm{J}}{\sqrt{\mathrm{R}}} \\ \frac{2c\mathrm{J}}{\sqrt{\mathrm{R}}} & ,\ \mathrm{J}\left(\cot q + \frac{h}{\sqrt{\mathrm{R}}}\right) \end{pmatrix}$$

vars signifikans verkar ligga i att produkten $(\mathrm{J},\ q)(\mathrm{J}',\ q')$ ger en matris av samma typ: $(\mathrm{J}'',\ q'')$. Dessutom påstår Cayley att matriserna $(\mathrm{J},\ q)$ och $(\mathrm{J}',\ q')$ kommuterar, vilket är allt annat än självklart.

Kompletterande material

Vi undersöker Cayleys påstående om matriserna $(\mathrm{J},\ q)$ och $(\mathrm{J}',\ q')$'s kommutativitet i ett specifikt fall. Förhållandet

$$2{:}3{:}4{=}1{:}\frac{3}{2}{:}2$$

uppfyller villkoret i premissen. Vi ser nu att

$$(\mathrm{J},\ q){=}\begin{pmatrix} \mathrm{J}\left(\cot q - \frac{1}{\sqrt{13}i}\right), & \frac{3\mathrm{J}}{\sqrt{13}i} \\ \frac{4\mathrm{J}}{\sqrt{13}i} & ,\ \mathrm{J}\left(\cot q + \frac{1}{\sqrt{13}i}\right) \end{pmatrix},$$

$$(\mathrm{J}',\ q'){=}\begin{pmatrix} \mathrm{J}'\left(\cot q' - \frac{2}{2\sqrt{13}i}\right), & \frac{6\mathrm{J}'}{2\sqrt{13}i} \\ \frac{8\mathrm{J}'}{2\sqrt{13}i} & ,\ \mathrm{J}'\left(\cot q' + \frac{2}{2\sqrt{13}i}\right) \end{pmatrix}.$$

Om vi antar att J och J' beter sig som vanliga, reella tal är det möjligt att kontrollera att

$$(\mathrm{J},\ q)(\mathrm{J}',\ q'){=}(\mathrm{J}',\ q')(\mathrm{J},\ q){=}$$

$$={\mathrm{J}\mathrm{J}'k^2}\begin{pmatrix} k\cot(q){-}1, & 3 \\ 4 & ,\ k\cot(q){+}1 \end{pmatrix}\begin{pmatrix} k\cot(q'){-}1, & 3 \\ 4 & ,\ k\cot(q'){+}1 \end{pmatrix}$$

där $k{=}1/\sqrt{13}i$. *Fin.*

I paragraf 49 presenterar så Cayley algoritmen på följande vis:

When the several matrices are each of them equal to

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}^n$$

(...) we find [that]

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}^n{=}\left(\frac{\sin nq}{\sin^n q}\left(\frac{1}{2}\sqrt{\mathrm{R}}\right)^n,\ nq\right)$$

or substituting for the right-hand side, the matrix represented by this notation, and putting for greater simplicity

$$\mathrm{L}=\frac{\sin nq}{\sin q}\left(\frac{1}{2}\sqrt{\mathrm{R}}\right)^{n-1}$$

we find

$$\begin{pmatrix} a, & b \\ c, & d \end{pmatrix}^n=\begin{pmatrix} \tfrac{1}{2}\mathrm{L}\left(\sqrt{\mathrm{R}}\cot nq-(d-a),\right. & \mathrm{L}b \\ \mathrm{L}c & ,\ \tfrac{1}{2}\mathrm{L}\left(\sqrt{\mathrm{R}}\cot nq+(d-a)\right. \end{pmatrix}$$

where it will be remembered that

$$\mathrm{R}=-(d-a)^2-4bc \text{ and } \cot q=\frac{d+a}{\sqrt{\mathrm{R}}}.$$

Kommentar

Cayley verkar först ha skapat en familj av matriser som beter sig ”snyggt” med avseende på multiplikation. Han visar sedan att med lämpligt val av (J, q) kan matrisen reduceras till sin ”vanliga” form, säg

$$\mathrm{M}=\begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Den familj av matriser som Cayley inför verkar vara inspirerad av rotationer, eller rotationsmatriser. En rotation med vinkeln θ kan som bekant skrivas:

$$\mathrm{R}_\theta=\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

En karakteristisk egenskap hos dessa typer av matriser är att $\mathrm{R}_{\theta 1}\cdot\mathrm{R}_{\theta 2}=\mathrm{R}_{\theta_1+\theta_2}$. Cayleys införda matris har exakt samma egenskap, något som kan förklaras med följande omskrivning:

$$(\mathrm{J},\,q)=\begin{pmatrix} \sin q^{-1} & 0 \\ 0 & \sin q^{-1} \end{pmatrix}\begin{pmatrix} \mathrm{J} & 0 \\ 0 & \mathrm{J} \end{pmatrix}\begin{pmatrix} \cos q-\frac{h}{\sqrt{\mathrm{R}}}\sin q & \frac{2b}{\sqrt{\mathrm{R}}}\sin q \\ \frac{2c}{\sqrt{\mathrm{R}}}\sin q & \cos q+\frac{h}{\sqrt{\mathrm{R}}}\sin q \end{pmatrix}.$$

Matrisen längst till höger kan vi benämna med S_q. Ett antal komplicerade och tidskrävande räkningar ger nu att $\mathrm{S}_q\cdot\mathrm{S}_{q'}=\mathrm{S}_{q+q'}$.

Har Cayley alltså funnit en rotationsmatris? Med införandet av komplexa tal och teorin kring egenvärden, egenrum och basbyten visar det sig att vi kan reducera S_q till en vanlig rotationsmatris (se appendix: A4). Oavsett hur Cayley kom fram till sitt resultat var det förmodligen mycket tillfredställande för honom, ty vi observerar att

$$(\mathrm{S}_\theta)^r:=\mathrm{S}_{r\theta}$$

nu är en möjlig definition, vars motsvarighet vi finner i Cayleys arbete (se sida 33 i *A memoir on the theory of matrices*). Cayley har således kommit fram till en metod att bestämma potenser av dessa typer av matriser.

Kompletterande material

Vi visar att Cayleys algoritm faktiskt fungerar i ett specifikt fall. Antag att vi vill beräkna M^2, där

$$\mathrm{M}{=}\begin{pmatrix}0 & -1\\ 2 & 2\end{pmatrix}$$

Då är

$$\mathrm{R}{=}4,\ \ \mathrm{L}{=}2,\ \ q{=}\frac{\pi}{4}.$$

Med Cayleys formel får vi att

$$\mathrm{M}^2{=}\begin{pmatrix}\frac{1}{2}\cdot 2\big(2\ \cot\ \frac{\pi}{2}-(2-0)\big) & -2\cdot 1\\ 2\cdot 2 & \frac{1}{2}\cdot 2\big(2\ \cot\ \frac{\pi}{2}+(2-0)\big)\end{pmatrix}$$

$$={\begin{pmatrix}-2 & -2\\ 4 & 4\end{pmatrix}}$$

vilket mycket riktigt är värdet på M^2. *Fin.*

Kommentar

Det är uppenbart att algoritmen har stora begränsningar. För det första gäller den bara för 2x2-matriser och verkar svår, om inte omöjlig, att generalisera till större matriser (Cayley ger ingen kommentar kring detta i sitt arbete, utan fokuserar enbart på 2x2-matriser). Den kan inte användas om $h{=}a{-}d{=}0$: då fallerar ju relationen som konstruktionen av den ursprungliga matrisen bygger på. Det visar sig också att den för många matriser leder till mer komplicerade beräkningar av potenser, som vi idag skulle bektrakta som enkla. Exempelvis skulle potensen av en diagonalmatris, säg

$$\mathrm{D}{=}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}$$

med Cayleys algoritm behöva beräknas genom att hantera komplexa tal i samband med trigonometriska funktioner. Ytterligare en begränsning finner vi i det faktum att L inte får vara noll, ty då reduceras ju den resulterande matrisen till nollmatrisen - så det måste även gälla att $a{+}d{\neq}0$ (spåret måste vara nollskilt).

Att metoden är elegant går dock knappast att bortse ifrån: den skulle säkerligen imponera på en samtida entusiast som läste Cayleys arbete. Dessutom leder Cayleys algoritm mot mer generella metoder, såsom Lies teorier (se till exempel artikeln om 'Liealgebra' på Wikipedia).

Slutligen påstår Cayley att formeln även fungerar i fallet då exponenten n är ett bråk, eller ett negativt tal:

The formula in fact extends to negative or fractional values of the index n, and when n is a fraction, we must, as usual, in order to exhibit the formula in its proper generality, write $q{+}2m\pi$ instead of q.

Den sista raden verkar indikera att Cayley vill få med alla komplexa rötter, något han tyvärr inte går in på i mer detalj.

Kompletterande material

Vi visar att Cayleys algoritm kan användas för att beräkna $\mathrm{M}^{1/2}$, med M definierad som tidigare. Med n=1/2 får vi att

$$\mathrm{R}{=}4,\ \ \mathrm{L}{=}\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}},\ \ q{=}\frac{\pi}{4}$$

och Cayleys algoritm ger oss att

$$\mathrm{A}{=}\begin{pmatrix} \frac{1}{2}\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}\left(2\ \cot\ \frac{\pi}{8}-2\right) & -\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} \\ 2\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}} & \frac{1}{2}\sqrt{\frac{1}{\sqrt{2}}-\frac{1}{2}}\left(2\ \cot\ \frac{\pi}{8}+2\right) \end{pmatrix}$$

är en rot till ekvationen. En något tidskrävande kontrollräkning visar att

$$\mathrm{A}^2{=}\mathrm{M}.$$

På motsvarande vis kan vi även kontrollera att algoritmen ger oss en korrekt rot B då $n{=}-1/2$ genom att verifiera att

$$\mathrm{B}^2{=}\mathrm{M}^{-1}.$$

Fin.

2.7 Cayleys korta redogörande för rektangulära matriser

I paragraf 52-57 rundar Cayley av sitt arbete med att införa notation och namn på olika typer av rektangulära matriser. Han kallar matriser med fler kolonner än rader för ”breda” matriser och matriser med fler rader än kolonner för ”djupa” matriser. Vidare konstaterar han att

> *The matrix zero subsists in the present theory, but not the matrix unity.*

Kompletterande material

Vi använder ett exempel för att styrka Cayleys påstående om avsaknaden av en rektangulär identitetsmatris. Låt

$$\mathrm{M} = \left(\begin{array}{ccc} a & b & c \\ a' & b' & c' \end{array}\right|$$

Det är nu uppenbart att matrisen

$$\mathrm{N} = \left(\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right|$$

fungerar som nolla (additiv identitet), eftersom

$$\mathrm{M}{+}\mathrm{N} = \left(\begin{matrix} a & b & c \\ a' & b' & c' \end{matrix}\right) + \left(\begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix}\right) = \mathrm{M}.$$

Ett eventuellt identitetselement I kan dock inte existera under multiplikation om vi bektraktar matriser av denna typ: Cayley har ju redan konstaterat att sammansättning mellan två rektangulära matriser per definition kräver att antalet kolonner i den till vänster belägna matrisen överensstämmer med antalet rader i den till höger belägna matrisen. Därmed måste

$$(\mathrm{I}_{a\times a}).(\mathrm{M}_{a\times b}){=}\mathrm{M}_{a\times b}$$

så I är nödvändigtvis en kvadratisk matris. *Fin.*

Cayley ger nu villkor på när addition av två rektangulära matriser är möjlig:

Matrices may be added or subtracted when the number of the lines and the number of the columns of the one matrix are respectively equal to the number of the lines and the number of the columns of the other matrix.

Han kommenterar även kring när det är möjligt att multiplicera två rektangulära matriser med varandra:

The notion of composition applies to rectangular matrices, but it is neccessary that the number of lines in the second or nearer component matrix should be equal to the number of columns in the first or further component matrix (...)

Cayley ger i paragraf 55 två specifika exempel på sammansättningar mellan två rektangulära matriser. Han hanterar det precis som i det kvadratiska fallet: identifiera respektive matris' element och utför multiplikationen enligt regeln för sammansättning, såsom tidigare definierad.

I paragraf 56 konstaterar han att två rektangulära matriser, sammansatta i olika ordning, ger upphov till två kvadratiska matriser av olika storlek. Han utgår från matriserna

$$\mathrm{M}{=}\left(\begin{matrix} a & b & c \\ d & e & f \end{matrix}\right),\ \mathrm{N}{=}\left(\begin{matrix} a' & d' \\ b' & e' \\ c' & f' \end{matrix}\right)$$

och visar att

$$\mathrm{M.N}{=}\left(\begin{matrix} (a,\, b,\, c \between a',\, b',\, c'), & (a,\, b,\, c \between d',\, e',\, f') \\ (d,\, e,\, f \between a',\, b',\, c'), & (d,\, e,\, f \between d',\, e',\, f') \end{matrix}\right)$$

samt att

$$\mathrm{N.M}{=}\left(\begin{matrix} (a',\, d' \between a,\, d), & (a',\, d' \between b,\, e), & (a',\, d' \between c,\, f) \\ (b',\, e' \between a,\, d), & (b',\, e' \between b,\, e), & (b',\, e \between c,\, f) \\ (c',\, f' \between a,\, d), & (c',\, f' \between b,\, e), & (c',\, f' \between c,\, f) \end{matrix}\right)$$

Cayley drar nu följande slutsats:

The two matrices in the case last considered admit of composition in the two different orders of arrangement, but as the resulting square matrices are not of the same order, the notion of the convertibility of two matrices does not apply even to the case in question.

Kommentar

Vi påminner oss om att Cayley med 'convertibility' avser kommutativitet mellan två matriser. Cayley konstaterar alltså ovan att två rektangulära matriser, sammansatta med varandra i olika ordning, omöjligen kan kommutera. För att två matriser ska kommutera måste ju den resulterande matrisens storlek efter komposition vara oberoende av ordningen på matriserna, vilket uppenbarligen inte är fallet här.

I paragraf 57 kommenterar Cayley kort varför det inte går att tala om inversen eller potensen av en rektangulär matris:

Since a rectangular matrix cannot be compounded with itself, the notions of the inverse or reciprocal matrix and of the powers of the matrix and the whole resulting theory of the functions of the matrix, do not apply to rectangular matrices.

Kommentar

Den kanske största anledningen till att Cayley lägger ner relativt lite tid på teorin kring rektangulära matriser i arbetet antyds ovan: de passar inte in i hans idé om det så kallade 'algebraical function theorem': determinanten är endast definierad för kvadratiska matriser så vi kan inte tala om ett karakteristiskt polynom i det rektangulära fallet. Dessutom är Cayleys motivation till införandet av matriser att det ska underlätta lösandet av linjära ekvationssystem med samma antal ekvationer som okända, men dessa representeras (som Cayley skriver i inledningen till sitt arbete) av just kvadratiska matriser.

Avslutningsvis kopplar Cayley transponatet såsom tidigare definierat till rektangulära matriser. Han konstaterar att transponatet, när det verkar på en "bred" matris, omvandlar den till en "djup" matris och vice versa. Han gör sedan den intressanta iakttagelsen att trippeln (a, b, c) rimligen kan ersättas med en matris på formen

$$\begin{array}{|ccc|} a & b & c \\ & & \end{array}$$

och att detta, tillsammans med transponatet, eventuellt skulle kunna ersätta notationen för

$$(a, b, c)(a', b', c')$$

genom att uttrycka (a', b', c') som en "djup" matris. Cayley poängterar dock att

(...) it is however more convenient to retain the symbol

$$(a,\, b,\, c\between a',\, b',\, c').$$

Cayley illustrerar avslutningsvis i två exempel hur transponatet verkar i sammansättningen mellan två matriser och beskriver utförligt hur elementen i den resulterande matrisen erhålles. Det andra exemplet följs av en regel som vi känner igen sedan tidigare:

$$\left(\begin{array}{cc} a, & b \\ d, & e \end{array}\right)\mathrm{tr.}\left(\begin{array}{cc} a', & b' \\ d', & e' \end{array}\right)=\left(\begin{array}{cc} (a,\, b\between a',\, b'), & (a,\, b\between d',\, e') \\ (d,\, e\between a',\, b'), & (d,\, e\between d',\, e') \end{array}\right)$$

(...) so that in the composition of matrices (square or rectangular), when the second or nearer component matrix is expressed as a matrix preceded by the symbol tr., any 'line' of the compound matrix is obtained by compounding the corresponding 'line' of the first or further component matrix successively with the several 'lines' of the matrix which preceded by tr. gives the second or nearer component matrix.

Kommentar

Man kan fråga sig varför Cayley inte bara ersätter den transponerade matrisen ovan med dess faktiska transponat. Då skulle ju regeln kunna formuleras på precis samma sätt som tidigare, i samband med introduktionen av sammansättning mellan matriser (se sida 7 i mitt arbete) - med det extra steget att först bestämma transponatet. Cayley skulle kanske ha gjort invändningen att ovanstående förklaring är pedagogisk, eftersom den mer direkt kan sägas demonstrera en egenskap hos transponatet under sammansättning - nämligen hur *närvaron* av ett transponat verkar på den resulterande matrisens element.

3 Avslutande tankar och reflektioner

Arthur Cayleys "A memoir on the theory of matrices" har varit ett intressant arbete att jobba med. Som student tillhör jag definitivt inte Cayleys tänkta målgrupp: som forskande matematiker var han mest intresserad av matematiken för matematikens skull och den tänkta publiken bestod av likasinnade entusiaster. Som Tony Crilly skriver i ett av sina många arbeten om Cayley var han tidigt känd som en mycket produktiv matematiker: positionen som professor vid Cambridge fick han mycket tack vare detta (*Arthur Cayley as Sadleirian Professor: A glimpse of Mathematics Teaching at 19th Century Cambridge*; s. 127). Att jag personligen har upplevt Cayleys arbete som aningen svårförståeligt i vissa delar kanske därför inte är så konstigt.

Cayley var uppenbarligen exalterad över att ha funnit och styrkt existensen av Cayley-Hamiltons sats. I sin iver verkar han ha hoppats på att idén om det så kallade "algebraical function theorem" skulle gå att generalisera till alla kvadratiska matriser, med nollskilda determinanter. Det ledde honom till studiet av kommuterande matriser och algoritmen för potenser av matriser. Även om Cayleys resultat så som presenterade i arbetet med moderna, lärobokstillvanda ögon kan uppfattas som bristfälliga, måste det ändå sägas att Cayleys arbete har varit inflytelserikt. Hans abstrakta notation för matriser lever kvar än idag (om än i förenklad form) och de delar av arbetet som handlar om Cayley-Hamiltons sats är fortfarande läsvärda och korrekta ur matematisk synpunkt. Att Cayley överhuvudtaget kom på idén att se matriser som algebraiska objekt är fascinerande i sig.

Jag lämnar läsaren med ett sista citat från Cayley, denna gång orelaterat till hans arbete. Det publicerades i *The World of Mathematics* (New York 1956):

As for everything else, so for a mathematical theory: beauty can be perceived but not explained.

4 Appendix

4.1 A1: Första sidan i Cayleys arbete

II. *A Memoir on the Theory of Matrices.* *By* ARTHUR CAYLEY, *Esq., F.R.S.*

Received December 10, 1857,—Read January 14, 1858.

THE term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices, and the term matrix used without qualification is to be understood as meaning a square matrix; in this restricted sense, a set of quantities arranged in the form of a square, *e. g.*

$$\left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a''\ , & b''\ , & c'' \end{array}\right)$$

is said to be a matrix. The notion of such a matrix arises naturally from an abbreviated notation for a set of linear equations, viz. the equations

$$\begin{aligned} X &= ax + by + cz, \\ Y &= a'x + b'y + c'z, \\ Z &= a''x + b''y + c''z, \end{aligned}$$

may be more simply represented by

$$(X, Y, Z) = \left(\begin{array}{ccc} a\ , & b\ , & c \\ a'\ , & b'\ , & c' \\ a''\ , & b''\ , & c'' \end{array}\right)(x, y, z),$$

and the consideration of such a system of equations leads to most of the fundamental notions in the theory of matrices. It will be seen that matrices (attending only to those of the same order) comport themselves as single quantities; they may be added, multiplied or compounded together, &c.: the law of the addition of matrices is precisely similar to that for the addition of ordinary algebraical quantities; as regards their multiplication (or composition), there is the peculiarity that matrices are not in general convertible; it is nevertheless possible to form the powers (positive or negative, integral or fractional) of a matrix, and thence to arrive at the notion of a rational and integral function, or generally of any algebraical function, of a matrix. I obtain the remarkable theorem that any matrix whatever satisfies an algebraical equation of its own order, the coefficient of the highest power being unity, and those of the other powers functions of the terms of the matrix, the last coefficient being in fact the determinant; the rule for the formation of this equation may be stated in the following condensed form, which will be intelligible after a perusal of the memoir, viz. the determi-

4.2 A2: Några definitioner.

Definition. Låt $\mathrm{A}\in\mathrm{M}_{n\times n}(\mathrm{F})$. Polynomet $f(\lambda)=\det(\mathrm{A}-\lambda\mathrm{I}_n)$ kallas för det *karakteristiska polynomet* till A.

Exempel. Matrisen

$$\mathrm{A}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}\in\mathrm{M}_{2\times 2}(\mathrm{R})$$

har det karakteristiska polynomet

$$f(\lambda)=\det(\mathrm{A}-\lambda\mathrm{I}_2)=\begin{vmatrix}1-\lambda & 2\\ 3 & 4-\lambda\end{vmatrix}=(1-\lambda)(4-\lambda)-6.$$

Definition. En minor $\mathrm{M}_{i,j}$ till en $n\times n$-matris A är en reducerad determinant i en determinant-utveckling som bildas genom att avlägsna den i:te raden och j:te kolumnen i A.

Definition. Kofaktormatrisen till en $n\times n$-matris A är den matris som bildas utifrån kofaktorerna

$$\mathrm{C}_{i,j}:=(-1)^{i+j}\mathrm{M}_{i,j}.$$

Definition. Givet en $n\times n$-matris A definieras dess adjungerande matris som

$$\mathrm{adj}(\mathrm{A})=\mathrm{C}^{\mathrm{T}}$$

där C är kofaktormatrisen till A.

Exempel. Låt

$$\mathrm{A}=\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$$

Minorerna har för denna matris A följande utseende:

$$\mathrm{M}_{1,1}=|4|,\ \mathrm{M}_{1,2}=|3|,\ \mathrm{M}_{2,1}=|2|,\ \mathrm{M}_{2,2}=|1|.$$

Kofaktormatrisen till A blir således

$$\mathrm{C}=\begin{pmatrix}4 & -3\\ -2 & 1\end{pmatrix}$$

och

$$\mathrm{C}^{\mathrm{T}}=\begin{pmatrix}4 & -2\\ -3 & 1\end{pmatrix}.$$

Vi ser nu speciellt att

$$\mathrm{adj.}(\mathrm{A})\cdot\mathrm{A}=\mathrm{C}^{\mathrm{T}}\cdot\mathrm{A}=\begin{pmatrix}4 & -2\\ -3 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 3 & 4\end{pmatrix}$$

$$=\begin{pmatrix}-2 & 0\\ 0 & -2\end{pmatrix}=\det(\mathrm{A})\cdot\mathrm{I}.$$

Definition. En matris $\mathrm{A}\in\mathrm{M}_{n\times n}$ är självadjungerad eller *Hermitisk* om den är sitt eget konjugattransponat, det vill säga, om $\mathrm{A}=\mathrm{A}^*$.

4.3 A3: Att finna en kommuterande matris till en godtycklig, kvadratisk matris A.

Sats. Givet en kvadratisk matris $\mathrm{A}\in\mathrm{M}_{n\times n}(\mathrm{F})$ är följande en stegvis metod för att alltid finna en kommuterande matris till A:

Steg 1. Finn en egenvektor u, tillhörande något egenvärde λ_u till A.

Steg 2. Finn en egenvektor v till A^{T}, tillhörande samma egenvärde λ_u. (Detta är alltid möjligt, eftersom B och B^{T} har samma egenvärden.)

Steg 3. Definiera $\mathrm{B}{=}uv^{\mathrm{T}}$. Då är AB=BA, det vill säga, B är en matris som kommuterar med A.

Bevis. Per definition gäller att $\mathrm{A}u{=}\lambda_u u$. Multiplikation med y^{T} från höger ger

$$\mathrm{A}uv^{\mathrm{T}}{=}\lambda uv^{\mathrm{T}} \quad (1)$$

Men $\mathrm{A}^{\mathrm{T}}v{=}\lambda_u v$, vilket kan uttryckas som att

$$\left(\mathrm{A}^{\mathrm{T}}v\right)^{\mathrm{T}}{=}\left(\lambda_u v\right)^{\mathrm{T}}.$$

Uttrycken i respektive led kan med kända egenskaper hos transponatet istället skrivas

$$v^{\mathrm{T}}\mathrm{A}{=}\lambda_u v^{\mathrm{T}}.$$

Vi ser nu att

$$v^{\mathrm{T}}\mathrm{A}{=}\lambda_u v^{\mathrm{T}}{\Rightarrow}uv^{\mathrm{T}}\mathrm{A}{=}u\lambda_u v^{\mathrm{T}}{=}\lambda_u uv^{\mathrm{T}} \quad (2)$$

där vi i sista likheten har untyttjat kommutativiteten mellan skalär och matris. Ekvation (1) och (2) ger oss slutligen att

$$\mathrm{A}uv^{\mathrm{T}}{=}uv^{\mathrm{T}}\mathrm{A}{\Rightarrow}\mathrm{AB}{=}\mathrm{BA},$$

om vi definierar $\mathrm{B}{=}uv^{\mathrm{T}}$. $\square$

4.4 A4: Detaljer till avsnittet om Cayleys algoritm.

Låt

$$\mathrm{S}_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Den karakteristiska ekvationen till denna matris är

$$\lambda^2 - 2\cos\theta \cdot \lambda + 1 = 0$$

vilket ger oss egenvärdena

$$\lambda_{1,2} = \cos\theta \pm i\sin\theta = e^{\pm i\theta}$$

Eftersom egenvärdena är distinkta vet vi att det existerar en bas av egenvektorer till S_θ i $\mathbb{C}^2$. Kalla egenvektorerna för v_1 och v_2. I denna bas gäller per definition att

$$\mathrm{S}_\theta v_1 = e^{i\theta} v_1, \quad \mathrm{S}_\theta v_2 = e^{-i\theta} v_2.$$

Vi försöker nu införa ett basbyte så att S_θ reduceras till en vanlig rotationsmatris. Låt vektorerna

$$w_1 = v_1 + iv_2, \quad w_2 = iv_1 + v_2$$

representera denna nya bas. Vi löser ut v_1 och v_2 och får då att

$$v_1 = \frac{iw_1 + w_2}{2i}, \quad v_2 = \frac{w_1 + iw_2}{2i}.$$

Vi kan nu bestämma basvektorerna w_1 och w_2 som en linjärkombination av v_1 och v_2. För den andra basvektorn har vi till exempel att

$$\mathrm{S}_\theta w_2 = ie^{i\theta} v_1 + e^{-i\theta} v_2$$

$$= ie^{i\theta}\left(\frac{iw_1 + w_2}{2i}\right) + e^{-i\theta}\left(\frac{w_1 + iw_2}{2i}\right)$$

$$= \frac{e^{i\theta}}{2i}\left(iw_2 - w_1\right) + \frac{e^{-i\theta}}{2i}\left(iw_2 + w_1\right)$$

$$= \left(\frac{e^{-i\theta} - e^{i\theta}}{2i}\right) w_1 + \left(\frac{e^{i\theta} + e^{i\theta}}{2}\right) w_2$$

vilket enligt Eulers (trigonometriska) formler kan skrivas om till

$$-\sin\theta \cdot w_1 + \cos\theta \cdot w_2 = \mathrm{S}_\theta w_2.$$

På samma sätt kan vi resonera oss fram till att

$$\cos\theta \cdot w_1 + \sin\theta \cdot w_2 = \mathrm{S}_\theta w_1.$$

I basen (w_1, w_2) ser vi alltså att S_θ är en vanlig rotationsmatris.

4.5 A5: Ett förslag på hur Cayley kan ha hanterat ekvationen $\mathrm{L}^2{=}\mathrm{M}^2$.

Antag att

$$\mathrm{M}{=}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\ ,\ \ \mathrm{L}{=}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}\ .$$

Då är

$$\mathrm{M}^2{=}\begin{pmatrix} a^2{+}bc & b(a{+}d) \\ c(a{+}d) & d^2{+}bc \end{pmatrix}\ .$$

Vi vet att matrisen L uppfyller ekvationen

$$\mathrm{L}^2{-}(\alpha{+}\delta)\mathrm{L}{+}\alpha\delta{-}\beta\gamma{=}0$$

så med substitutionen $\mathrm{L}^2{=}\mathrm{M}^2$ ser vi att L kan uttryckas enligt

$$\mathrm{L}{=}\frac{1}{\alpha{+}\delta}\left[\mathrm{M}^2{+}\alpha\delta{-}\beta\gamma\right].$$

Med samma typ av substitution som tidigare,

$$\mathrm{X}{=}\alpha{+}\delta,\ \ \mathrm{Y}{=}\alpha\delta{-}\beta\gamma$$

och en koefficientanalys - får vi följande ekvationssystem:

$$\alpha{=}\frac{a^2{+}bc{+}\mathrm{Y}}{\mathrm{X}},\ \ \beta{=}\frac{b(a{+}d)}{\mathrm{X}},\ \ \gamma{=}\frac{c(a{+}d)}{\mathrm{X}},\ \ \delta{=}\frac{d^2{+}bc{+}\mathrm{Y}}{\mathrm{X}}.$$

Om vi utvecklar summan $\alpha{+}\delta$ får vi, efter lite jobb, följande likhet:

$$\alpha{+}\delta{=}\mathrm{X}{=}\frac{\Big((a{+}d)^2{-}2(ad{-}bc)\Big){+}2\mathrm{Y}}{\mathrm{X}}.$$

Vi substituerar uttrycket inom parentes med bokstaven P och får då (återigen) att

$$\mathrm{X}^2{=}\mathrm{P}{+}2\mathrm{Y}.$$

Vi kan nu också utveckla differensen $\alpha\delta{-}\beta\gamma$. I ett första steg ser vi att

$$\alpha\delta{-}\beta\gamma{=}\mathrm{Y}{=}\frac{\big(a^2{+}bc{+}\mathrm{Y}\big)\big(d^2{+}bc{+}\mathrm{Y}\big){-}bc(a{+}d)^2}{\mathrm{X}^2}$$

vilket vi, efter en hel del algebra, kan reducera till att

$$\mathrm{Y}{=}\frac{(ad{-}bc)^2{+}\big((a{+}d)^2{-}2(ad{-}bc)\big)\mathrm{Y}{+}\mathrm{Y}^2}{\mathrm{X}^2}.$$

Med en sista substitution,

$$\mathrm{Q}{=}(ad{-}bc)^2,$$

landar vi i likheten

$$\mathrm{Y}=\frac{\mathrm{Q}+\mathrm{PY}+\mathrm{Y}^2}{\mathrm{X}^2}.$$

Sammanfattningsvis har vi alltså kommit fram till att

$$\mathrm{X}^2=\mathrm{P}+2\mathrm{Y},\quad \mathrm{Y}=\frac{\mathrm{Q}+\mathrm{PY}+\mathrm{Y}^2}{\mathrm{X}^2}$$

vilket är helt identiskt med resultatet som beskrivs i huvuddelen av arbetet. Ovan kan nämligen (återigen) visas motsvara ekvationerna

$$\mathrm{Y}^2=\mathrm{Q},\quad \mathrm{X}=\pm\sqrt{\mathrm{P}+2\sqrt{\mathrm{Q}}}.$$

Om vi uttrycker den första ekvationen i termer av den kända matrisen M's koefficienter, ser vi att

$$(ad-bc)^2=\mathrm{Y}^2 \Rightarrow \mathrm{Y}=\pm(ad-bc).$$

Detta kan ha lett Cayley till den falluppdelning han redovisar i sitt arbete: om vi väljer $\mathrm{Y}=ad-bc$ är det lätt att visa att $\mathrm{X}=\pm(a+d)$. Men detta är ju ekvivalent med att påstå att

$$\mathrm{L}=\pm\mathrm{M}$$

är en lösning till ekvationen $\mathrm{L}^2=\mathrm{M}^2$. Om vi istället väljer $\mathrm{Y}=-ad+bc$, ser vi att

$$\mathrm{X}=\pm\sqrt{(a+d)^2+4bc}.$$

Explicit kan vi nu, med hjälp av ovan, uttrycka matrisen L enligt

$$\mathrm{L}=\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}=\begin{pmatrix}\frac{a^2-ad+2bc}{\mathrm{X}} & \frac{b(a+d)}{\mathrm{X}}\\ \frac{c(a+d)}{\mathrm{X}} & \frac{d^2-ad+2bc}{\mathrm{X}}\end{pmatrix}$$

vilket (som Cayley också visar) kan skrivas

$$\mathrm{L}=\frac{a+d}{\mathrm{X}}\begin{pmatrix}a & b\\ c & d\end{pmatrix}-\frac{2(ad-bc)}{\mathrm{X}}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

eller, om vi ersätter enhetsmatrisen med talet 1 och den kända matrisen med storheten M:

$$\mathrm{L}=\frac{a+d}{\mathrm{X}}\mathrm{M}-\frac{2(ad-bc)}{\mathrm{X}}.$$

4.6 A6: Kompletterande material till diskussionen om matrisekvationen L^2=M.

Låt

$$M=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}.$$

Antag att

$$L=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

är en lösning till matrisekvationen L^2=M. Vi kan då skriva att

$$\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}a & b\\ c & d\end{pmatrix}=\begin{pmatrix}a^2+bc & ab+bd\\ ac+cd & bc+d^2\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}.$$

Detta implicerar att

$$b=0 \lor a+d=0,$$

$$c=0 \lor a+d=0.$$

I fallet då b eller c är 0 ser vi att matrisens element blir:

$$a=\pm 1,\ b=0,\ c=0,\ d=\pm\sqrt{2}.$$

Om istället $a+d=0$, måste det gälla att

$$a^2+bc=bc+d^2$$

men detta måste då återigen implicera att b eller c är 0 - ty

$$a^2+bc=1,\ bc+d^2=2.$$

Vi kan nu dra slutsatsen att matrisekvationen L^2=M har exakt två lösningar, som båda är diagonalmatriser:

$$L=\begin{pmatrix}\pm 1 & 0\\ 0 & \pm\sqrt{2}\end{pmatrix}.$$

5 Referenser

Arthur Cayley (1858), *A memoir on the theory of matrices* (originaltryck); tillgänglig via royalsocietypublishing.org.

Victor J. Katz (2014), *A history of mathematics: an introduction*, 4:e upplagan, s. 740-746.

Tony Crilly (1978), *Cayley's anticipation of a generalised Cayley-Hamilton theorem*, s. 211-219 i *Historia Mathematica*, 5:e volymen.

Tony Crilly (1999), *Arthur Cayley as Sadleirian professor: a glimpse of mathematics teaching at 19th-century Cambridge*, s. 125-160 i *Historia Mathematica*, 26:e volymen.

Thomas Hawkins (2013), *The mathematics of Frobenius in context*, s. 215-218.

Friedberg et. al. (2013), *Linear algebra: Pearson New International Edition*

Robertson och O'Connor, *MacTutor* (digitalt arkiv); Arthur Cayleys biografi och historien om matriser och determinanter.

Stackexchange (2019), "Cayley's matrix notation"; https://tex.stackexchange.com/questions/487643/cayleys-matrix-notation

Wikipedia (2021), "Square root of a matrix"; https://en.wikipedia.org/wiki/Square_root_of_a_matrix