KUNGL TEKNISKA HÖGSKOLAN

Linjära avbildningar i planet

Matematiska institutionen, 2000

Revison A

1. Introduktion

Det finns hur många böcker som helst om linjär algebra. Detta är ett försök att utan anspråk på fullständighet presentera några av ämnets geometriska idéer till studenter på gymnasiet. Texten är avsedd som självstudiematerial, men är kompakt skriven jämfört med gymnasielitteratur för att man på få sidor ska nå fram till linjära avbildningar och matriser. Detta betyder att läsaren kan behöva stanna upp ofta och tänka efter, rita egna bilder, ta en paus o s v. Lösningarna till övningarna innehåller ibland information som används senare i texten och spelar därför en viktig roll för förståelsen.[1]

Metodik när man ska lösa problem är förstås ett omfattande ämne i sig,[2] låt oss bara i korthet säga att det i allmänhet handlar om att formulera om ett problem till något man kan, och att ”något man kan” även kan vara något man nyss lärt sig. Övning ger faktiskt färdighet, och kom ihåg att om du kör fast så fråga!

Betoningen i den här framställningen ligger på plana problem, d v s två koordinater. Allting vi beskriver här går att göra i tre eller fler dimensioner. Det är nyttigt att då och då fråga sig hur de uttryck man stöter på skulle se ut i högre dimensioner. (Se även Appendix B.)

I texten finns många *definitioner*, d v s införande av nya begrepp, t ex vad en matris är, hur man adderar och multiplicerar matriser, vad som menas med att en operation är kommutativ o s v. Det finns också många *satser*, d v s slutsatser och insikter om olika begrepp, t ex att matrisaddition är kommutativ, men att matrismultiplikation inte är det. I matematisk text är man vanligen tydlig med vad som är definition och vad som är sats, och beviset för satsen utförs ofta efter att satsen formulerats. Vi använder inte det framställningssättet här eftersom det ofta upplevs som mer ”redovisande” och mindre ”upptäckande” för en ovan läsare. Detta med risk för missförstånd om vad som är vad, men med förhoppning om att komma matematiken närmare.

Några ord om figurerna. Att göra snygga och pedagogiska bilder är ett tidsödande arbete. I valet mellan några snygga figurer eller många handritade figurer valde vi det senare. En förhoppning är att bilderna inte avskräcker utan snarare inbjuder läsaren att göra bättre bilder på egen hand. Det är också mycket troligt att det är just genom dessa av studenten egenhändigt ritade bilder som själva förståelsen för ämnet uppkommer.

Den här texten har formats ur diskussioner med Mattias Dahl, Hans Ringstöm, Simon Wigzell och Kristian Bjerklöv.

KTH, januari 2000
Göran Selander

[1] Fotnötterna kan man dock utelämna utan att förlora förståelse.
[2] Se t ex Polya: *How to solve it.*

2. Motivering

Förstoring, förminskning, vridning, spegling och projektion är olika exempel på linjära avbildningar. Vi stöter på dem när vi ändrar storlek på typsnitt i ordbehandlingsprogrammet, speglar en bild i bildbehandlingsprogrammet eller alla dessa snurrande figurer på olika hemsidor (för inte tala om om man spelar TV-spel).

Vi ska här studera linjära avbildningar i planet från en geometrisk utgångspunkt. Under resans gång kommer vi oundvikligen att lära oss om matriser, vilka är rektangulära taltabeller, och hur man räknar med dessa. Vi får då tillgång till ett kraftfullt verktyg som används i många andra sammanhang, t ex när man löser linjära ekvationssystem. Detta enda exempel skulle egentligen vara nog eftersom linjära ekvationssystem är så vanligt förekommande inte minst som approximation till mer komplicerade ekvationer. Men det finns många andra exempel. Heisenberg och Dirac formulerade kvantmekaniken med hjälp av matriser där egenvärden och egenvektorer spelar en viktig roll (kapitel 10). Tensorer i Einsteins relativitetsteori är en generalisering av matriser. Även i den rena matematiken finns många exempel, t ex projektiv geometri, gruppteori, derivata av funktioner av flera variabler, system av differentialekvationer, Markovkedjor. Listan kan göras mycket längre. Bilderna nedan är några exempel.

Grafteori.

Projektion av roterande kub på planet $z = y$. $P = \frac{1}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} \cos \pi t & -\sin \pi t & 0 \\ \sin \pi t & \cos \pi t & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Arnolds kattavbildning. Itereration av $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ modulo 1.[3]

[3]Mannen på fotografiet är John von Neumann, en känd matematiker.

3. Vektorer

Ett sätt att representera punkter i planet är att lägga ut två vinkelräta lika graderade koordinataxlar. Varje punkt i planet svarar på ett entydigt sätt mot en korsning av axelparallella linjer och på axlarna avläses punktens koordinater.

Till varje punkt (a, b) svarar en *vektor* $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$ som vi representerar med en pil från origo till punkten (a, b).[4] Talen a och b kallas för *komponenter* eller element till vektorn $\vec{v}$. En vektor behöver inte starta i origo, genom att parallellförflytta vektorn till origo kan man läsa av dess komponenter.

Man kan också beräkna komponenterna genom att ta koordinaterna för slutpunkten minus koordinaterna för startpunkten:

Exempel 1. Vektorn som startar i $(2, -1)$ och slutar i $(3, 1)$ är $\begin{pmatrix} 3 & - & 2 \\ 1 & - & (-1) \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Övning 3.1. *Beräkna längden av vektorn* $\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}$. *Ledning: Börja med fallet att a och b är positiva. Tänk sedan igenom fallen att a eller b är negativa.*

Längden av vektorn $\vec{v}$ betecknar vi $|\vec{v}\,|$. Avståndet mellan två punkter är detsamma som längden av den vektor som startar i ena punkten och slutar i den andra punkten.

Övning 3.2. *Hur stort är avståndet mellan punkterna (a, b) och (c, d).*

En vektor med längd 1 kallas för en *enhetsvektor*. Det är praktiskt att införa en beteckning för två speciella enhetsvektorer, $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ och $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Låt oss kalla dessa för *basvektorer*.

Övning 3.3. *Rita ut $\vec{e}_1$ och $\vec{e}_2$ i ett koordinatsystem. Låt båda vektorerna starta i origo.*

Vi kan addera vektorer genom att addera respektive komponent.

$$\vec{v}_1 + \vec{v}_2 = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a + c \\ b + d \end{pmatrix} = \vec{v}_3$$

[4] En vanligare notation för vektorn är (a, b), d v s precis samma som för punkten. I stället för $\vec{v}$ skriver man ibland $\bar{v}$ eller $\mathbf{v}$.

vilket är geometriskt detsamma som att parallellförflytta den ena pilen och avsätta den efter den andra och läsa av komponenterna för den resulterande vektorn.

Vid addition av två tal, t ex a och c, gäller förstås att $a + c = c + a$, det spelar ingen roll i vilken ordning två tal adderas. Detta kallas för den *kommutativa lagen* vid addition. När vi nu adderar två vektorer gäller följande ekvation

$$\vec{v}_1 + \vec{v}_2 = \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} a+c \\ b+d \end{pmatrix} = \begin{pmatrix} c+a \\ d+b \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \vec{v}_2 + \vec{v}_1,$$

d v s den kommutativa lagen gäller även addition av vektorer. Detta syns också geometriskt av parallellogrammen.

Vi kan multiplicera vektorer med tal $r\vec{v} = r\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} ra \\ rb \end{pmatrix}$.

Resultatet blir en vektor parallell med $\vec{v}$, pekande åt samma håll om $r > 0$ och åt motsatt håll om $r < 0$, förlängd om $r > 1$ eller $r < -1$ och förkortad om $-1 < r < 1$. Om $r = 0$ så blir $0\vec{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0}$. Denna vektor kallas *nollvektorn*. Att multiplicera en vektor med ett tal kallas för *multiplikation med skalär* eller kort och gott *skalning*.

Övning 3.4. *Antag att* $\vec{v}_1 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ *och* $\vec{v}_2 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$. *Beräkna a)* $\vec{v}_1 + \vec{v}_2$ *b)* $-\vec{v}_1$ *c)* $\vec{v}_1 - 2\vec{v}_2$

Rita ut vektorerna $\vec{v}_1$, $\vec{v}_2$ *samt resultaten i övningarna a) – c) i ett koordinatsystem.*

Övning 3.5. *Bestäm de två enhetsvektorer som är parallella med* $\begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

Med hjälp av addition och multiplikation med skalär kan vi uttrycka alla vektorer med hjälp av basvektorerna $\vec{e}_1$ och $\vec{e}_2$:

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix} = a\vec{e}_1 + b\vec{e}_2$$

Exempel 2.

$$\begin{pmatrix} 2 \\ -3 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} - 3\begin{pmatrix} 0 \\ 1 \end{pmatrix} = 2\vec{e}_1 - 3\vec{e}_2.$$

4. Linjära avbildningar

Vi ska nu införa linjära avbildningar av vektorer. Vi är (förhoppningsvis) vana vid funktioner $y = f(x)$, där man stoppar in ett reellt tal x och får ut ett reellt tal y. En *avbildning* av vektorer är också en funktion, men man stoppar in en vektor och får ut en vektor. Geometriskt är det kanske lättast att förstå om vi ritar upp två koordinatsystem.

I figuren ovan verkar avbildningen A. Vi ser hur vektorn $\vec{v}$ ("urbilden") avbildas på vektorn $\vec{w}$ ("bildvektorn"), A vrider och förlänger. Vi skriver: $\vec{w} = A(\vec{v})$ (jämför $y = f(x)$). Det är även vanligt (se ekvation (4.1)) och, som kommer att framgå, väl motiverat att utelämnat parentesen kring $\vec{v}$. Vi kan alltså även skriva $\vec{w} = A\vec{v}$.

Att avbildningen A är *linjär* betyder två saker:

L1. Summan av två vektorer avbildas på bildvektorernas summa: $A(\vec{v}_1 + \vec{v}_2) = A(\vec{v}_1) + A(\vec{v}_2)$.

L2. Om en vektor skalas om så kommer bildvektorn skalas om med samma faktor: $A(r\vec{v}_1) = rA(\vec{v}_1)$.

Övning 4.1. *Antag att A är en linjär avbildning som avbildar $\vec{v}_1$ på $\begin{pmatrix}2\\6\end{pmatrix}$, och $\vec{v}_2$ på $\begin{pmatrix}3\\-2\end{pmatrix}$ (d v s $A\vec{v}_1 = \begin{pmatrix}2\\6\end{pmatrix}$ och $A\vec{v}_2 = \begin{pmatrix}3\\-2\end{pmatrix}$). Hur avbildas $10\vec{v}_1 + 5\vec{v}_2$? (Vi känner inte till komponenterna till $\vec{v}_1$ och $\vec{v}_2$.)*

Övning 4.2. *Antag att A är en linjär avbildning som avbildar basvektorerna $\vec{e}_1$ och $\vec{e}_2$ på $\begin{pmatrix}2\\6\end{pmatrix}$ respektive $\begin{pmatrix}3\\-2\end{pmatrix}$. Hur avbildas $\begin{pmatrix}-4\\5\end{pmatrix}$? Hur avbildas $\begin{pmatrix}x\\y\end{pmatrix}$?*

En finess med en linjär avbildning är att den är entydigt bestämd av hur den avbildar basvektorerna: Om vi vet att

$$A\begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix} \qquad \text{och} \qquad A\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}c\\d\end{pmatrix} \tag{4.1}$$

(vi antar att talen a, b, c och d är kända), så kan vi räkna ut bilden av varje annan okänd vektor, säg $\begin{pmatrix}x\\y\end{pmatrix}$. Vi använder reglerna **L1** och **L2**.

$$A\begin{pmatrix}x\\y\end{pmatrix} = A\left(x\begin{pmatrix}1\\0\end{pmatrix} + y\begin{pmatrix}0\\1\end{pmatrix}\right) = \left[\ \mathbf{L1}\ \right] = A\left(x\begin{pmatrix}1\\0\end{pmatrix}\right) + A\left(y\begin{pmatrix}0\\1\end{pmatrix}\right) = \left[\ \mathbf{L2}\ \right] =$$

$$= xA\begin{pmatrix}1\\0\end{pmatrix} + yA\begin{pmatrix}0\\1\end{pmatrix} = \Big[\text{ ekv. (4.1) }\Big] = x\begin{pmatrix}a\\b\end{pmatrix} + y\begin{pmatrix}c\\d\end{pmatrix} = \begin{pmatrix}ax+cy\\bx+dy\end{pmatrix}$$

Sammanfattningsvis:

$$A\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}ax+cy\\bx+dy\end{pmatrix}. \tag{4.2}$$

Det är nu naturligt att införa begreppet *matris*, vilket betyder en rektangulär tabell med tal innanför en stor parentes. Fast vi har egentligen redan smygstartat:
$\begin{pmatrix}1\\0\end{pmatrix}$ och $\begin{pmatrix}2\\-3\end{pmatrix}$ är exempel på 2×1-matriser, d v s en rektangulär tabell med två *rader* och en *kolonn*. Vi har ju hittills kallat 2×1-matriser för vektorer. Om man vill vara extra tydlig kan man säga att de är *kolonnvektorer*. Det finns även ett annat sätt att skriva en vektor på matrisform nämligen som en 1×2-matris, t ex $(1 \;\; 0)$ eller $(2 \;\; -3)$. Dessa kallas även *radvektorer*. Vi gör ingen geometrisk skillnad på radvektorer och kolonnvektorer, men vi behöver känna till båda när vi ska multiplicera matriser. De matriser vi framför allt kommer använda är 2×2-matriser, som alltså har två rader och två kolonner med element: t ex $\begin{pmatrix}a & c\\b & d\end{pmatrix}$.

Som vi snart ska se är det naturligt att förknippa avbildningen A ovan med just denna matris.

5. Matriser

Precis som man kan addera vektorer komponentvis, kan man addera matriser av samma storlek *elementvis*:

$$\begin{pmatrix}a & c\\b & d\end{pmatrix} + \begin{pmatrix}e & g\\f & h\end{pmatrix} = \begin{pmatrix}a+e & c+g\\b+f & d+h\end{pmatrix} = \begin{pmatrix}e & g\\f & h\end{pmatrix} + \begin{pmatrix}a & c\\b & d\end{pmatrix}$$

Vi ser att additionen är kommutativ. Multiplikation med skalär fungerar som för vektorer, varje element multipliceras med skalfaktorn:

$$\begin{pmatrix}a & c\\b & d\end{pmatrix} = \begin{pmatrix}ra & rc\\rb & rd\end{pmatrix}$$

Övning 5.1. *Beräkna* $\begin{pmatrix}2 & 3\\1 & -1\end{pmatrix} + 6\begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}$.

Nu inför vi multiplikation mellan matriser, först mellan en 1×2-matris och en 2×1-matris.

$$(a \;\; b)\begin{pmatrix}c\\d\end{pmatrix} = ac + bd \tag{5.1}$$

Här är ordningen mycket viktig, därför att produkten är **inte** kommutativ.[5]

Exempel 1.

$$(5 \;\; 4)\begin{pmatrix}2\\-3\end{pmatrix} = 5 \cdot 2 + 4 \cdot (-3) = -2$$

Övning 5.2. *Beräkna a)* $(6 \;\; 4)\begin{pmatrix}2\\-3\end{pmatrix}$ *b)* $(1 \;\; -2)\begin{pmatrix}6\\5\end{pmatrix}$ *c)* $(\frac{3}{5} \;\; \frac{4}{5})\begin{pmatrix}6\\-2\end{pmatrix}$

Denna produkt kan också ses som en produkt mellan vektorerna $\vec{v}_1 = \begin{pmatrix}a\\b\end{pmatrix}$ och $\vec{v}_2 = \begin{pmatrix}c\\d\end{pmatrix}$, om vi gör om den första vektorn till en radvektor:

$$\vec{v}_1 \cdot \vec{v}_2 = (a \;\; b)\begin{pmatrix}c\\d\end{pmatrix} = ac + bd$$

[5] Vi ska strax lära oss multiplicera matriser, låt oss bara som smakprov påpeka att

$$\begin{pmatrix}c\\d\end{pmatrix}(a \;\; b) = \begin{pmatrix}ac & bc\\ad & bd\end{pmatrix}.$$

Detta skall jämföras med (5.1). Resultatet är alltså i det här fallet inte ett tal utan en 2×2-matris.

Produkten kallas för *skalärprodukt* eftersom resultatet blir en skalär, d v s ett tal. Vi betecknar den med en punkt mellan vektorerna.[6]

Övning 5.3. *Visa att skalärprodukten är kommutativ (trots att matrismultiplikation inte är det).*

Skalärprodukten har en geometrisk betydelse. Vi börjar med att införa den *vinkelräta projektionen* av en vektor $\vec{v}$ på en linje L.

Om den ena vektorn är en enhetsvektor parallell med linjen så är skalärprodukten lika med längden av projektionen av den andra vektorn på linjen (med minustecken om vinkeln mellan vektorerna är trubbig).

Övning 5.4. *a) Hur stor är vinkeln mellan vektorerna* $\begin{pmatrix}6\\4\end{pmatrix}$ *och* $\begin{pmatrix}2\\-3\end{pmatrix}$*?*

b)Beräkna skalärprodukten mellan $\vec{v}_1 = \begin{pmatrix}1\\-2\end{pmatrix}$ *och* $\vec{v}_2 = \begin{pmatrix}6\\5\end{pmatrix}$*. Är vinkeln mellan vektorerna spetsig eller trubbig?*

c) Hur lång är den vinkelräta projektionen av vektorn $\begin{pmatrix}6\\-2\end{pmatrix}$ *på en linje som är parallell med vektorn* $\begin{pmatrix}3\\4\end{pmatrix}$*?*

Rita upp vektorerna och se till att du förstår de geometriska frågorna.

Som vi har sett förutsätter skalärprodukten att det är lika många element i båda vektorerna.

Övning 5.5. *Hur borde man definiera skalärprodukt mellan vektorer med n komponenter, om n är ett heltal $\geqslant 3$? (Hur borde man definiera multiplikation mellan en $1 \times n$-matris och en $n \times 1$-matris?)*

Övning 5.6. *Uttryck längden av en vektor* $\vec{v} = \begin{pmatrix}a\\b\end{pmatrix}$ *med hjälp av en skalärprodukt.*

Nu till litet större matriser. Om vi multiplicerar en 2×2-matris och en 2×1-matris så ser det ut så här:

$$\begin{pmatrix}a & c\\b & d\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}ax+cy\\bx+dy\end{pmatrix} \tag{5.2}$$

Produkten mellan två matriser blir en matris där varje element är en produkt mellan en viss rad i den vänstra matrisen och en viss kolonn i den högra matrisen. En god minnesregel är denna:

$$\begin{array}{cc} & \begin{pmatrix}x\\y\end{pmatrix} \\ \begin{pmatrix}a & c\\b & d\end{pmatrix} & \begin{pmatrix}ax+cy\\bx+dy\end{pmatrix}\end{array}$$

Det första elementet i resultatmatrisen är produkten av den första raden i vänstra matrisen med den första (och enda) kolonnen i den högra matrisen. Det andra elementet i resultatmatrisen är

[6] På engelska kallas skalärprodukt ibland för *dot product.*

produkten av den andra raden i vänstra matrisen med den första (och enda) kolonnen i den högra matrisen.

Exempel 2.

$$\begin{pmatrix} 2 & -1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ -2 \end{pmatrix} = \begin{pmatrix} 8+2 \\ 12+0 \end{pmatrix} = \begin{pmatrix} 10 \\ 12 \end{pmatrix}$$

Övning 5.7. *Beräkna a)* $\begin{pmatrix} 1 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ *b)* $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ *c)* $\begin{pmatrix} -2 & 4 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Nu återvänder vi till de linjära avbildningarna. Jämför ekvationerna (4.2) och (5.2). Vi ser att den linjära avbildningen A är detsamma som multiplikation med matrisen $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Detta är själva poängen med den här texten, stanna upp ett tag och begrunda.

Vi kommer i fortsättningen inte skilja på en linjär avbildning och den matris som vid matrismultiplikation gör samma sak som avbildningen utan skriver $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$.[7]

Resten av texten är ägnad åt att förstå olika egenskaper hos linjära avbildningar, d v s multiplikation av matriser med olika värden på elementen a, b, c och d.

Vi avslutar kapitlet med att diskutera allmän matrismultiplikation. Detta kan man hoppa över vid första genomläsningen om man vill, det kommer inte att användas förrän i kapitel 8.

Den metod vi beskrivit för multiplikation av matriser fungerar även för allmänna matrismultiplikationer. Nu kan man dock inte multiplicera vilka matriser som helst med varandra. Låt oss titta på ett exempel som visar på när det går bra:

$$\begin{pmatrix} 2 & 1 & 6 \\ 3 & -1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 1 & -4 \\ 0 & -2 & 5 & -1 \\ -1 & 0 & 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -2 & 19 & -9 \\ 9 & 2 & -2 & -11 \end{pmatrix}$$

Elementet på första raden och första kolonnen i resultatmatrisen, 0, fick vi fram från produkten av första raden i matrisen till vänster med första kolonnen i matrisen till höger (jämför övning 5.5):

$$(2 \;\; 1 \;\; 6) \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} = 2 \cdot 3 + 1 \cdot 0 + 6 \cdot (-1) = 0$$

De andra elementen i resultatmatrisen är beräknade på samma sätt i enlighet med minnesregeln:

$$\begin{array}{cc} & \begin{pmatrix} 3 \\ 0 \\ -1 \end{pmatrix} \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \\ 2 \end{pmatrix} \begin{pmatrix} -4 \\ -1 \\ 0 \end{pmatrix} \\ \begin{matrix} (2 & 1 & 6) \\ (3 & -1 & 0) \end{matrix} & \begin{pmatrix} 0 & -2 & 19 & -9 \\ 9 & 2 & -2 & -11 \end{pmatrix} \end{array}$$

Som vi ser förutsätter denna multiplikation att antalet kolonner i matrisen till vänster är lika med antalet rader i matrisen till höger (båda = 3), annars blir det ju ingen vettig skalärprodukt. Observera att resultatmatrisen får lika många rader som matrisen till vänster (2 stycken) och lika många kolonner som matrisen till höger (4 stycken).

En minnesregel som sammanfattar båda dessa observationer ges här:

$$(2 \times \cancel{3}) \cdot (\cancel{3} \times 4) = (2 \times 4)$$

Multiplikation av en 2×3-matris och en 3×4-matris går bra eftersom antalet kolonner till vänster = antalet rader till höger (de överstrukna 3:orna) och resultatet blir en 2×4-matris (det som är kvar när vi strukit 3:orna).

En matris har lika många rader som kolonner kallas *kvadratisk*. Kvadratiska matriser av samma storlek (t ex 2×2-matriser) kan alltså multipliceras med varandra.

[7] Det är också därför vi utelämnar parentesen: $A(\vec{v}) = A\vec{v}$

För att återknyta till skalärprodukt vill vi anmärka man inte kan matrismultiplicera en 2×1-matris med en 2×1-matris. Uttrycket $\begin{pmatrix} a \\ b \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix}$ saknar alltså mening som matrismultiplikation. Det var därför vi var tvungna att göra om första vektorn till en radvektor.

Övning 5.8. *Kontrollera att minnesregeln stämmer med de olika matrisprodukter vi stött på tidigare (3 stycken, om man inkluderar fotnötter).*

Övning 5.9. *Avgör om matrisprodukten är möjlig och utför i sådana fall multiplikationen.*

a) $\begin{pmatrix} 5 & 8 & 16 \\ 4 & 1 & 8 \\ -4 & -4 & -11 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$ *b)* $\begin{pmatrix} 2 & -3 & 4 \\ 3 & 2 & 0 \end{pmatrix}\begin{pmatrix} 4 & -1 & 1 \\ 1 & 4 & 2 \end{pmatrix}$ *c)* $\begin{pmatrix} 2 \\ 3 \\ 0 \\ -1 \end{pmatrix}\begin{pmatrix} 4 & 1 \end{pmatrix}$

d) $\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 4 & -1 \\ 1 & 4 \end{pmatrix}$ *e)* $\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

Övning 5.10. *a) Låt* $A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$ *och* $B = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$. *Beräkna* BA *och* AB.

b) Låt C *vara en kvadratisk matris och definiera* $C^2 = CC$. *Varför gäller inte* $(A+B)^2 = A^2 + 2AB + B^2$ *i allmänhet? (*A *och* B *är kvadratiska matriser av samma storlek.)*

6. Geometriska operationer

Vi ska nu försöka förstå mer om linjära avbildningar, vi börjar med ett exempel: $A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$. Observera att $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ och $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, d v s basvektorerna avbildas på de kolonnvektorer som A består av. Kalla dessa $\vec{v}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$ och $\vec{v}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

Vi tittar närmare på enhetskvadraten med hörn i $(0,0)$, $(1,0)$, $(0,1)$ och $(1,1)$. I fortsättningen kommer vi förknippa en punkt (x,y) med vektorn $\begin{pmatrix} x \\ y \end{pmatrix}$ som pekar ut punktens läge om den startar i origo. Vi ser att hörnen i kvadraten $\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ och $\vec{e}_1 + \vec{e}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ avbildas på hörnen i parallellogrammen: $\vec{0}$, $\vec{v}_1$, $\vec{v}_2$ och $\vec{v}_1 + \vec{v}_2$. (Lägg märke till att $A\vec{0} = \vec{0}$ oavsett vilken linjär avbildning A är.[8])

Hur avbildas andra punkter? Om vi väljer en punkt $\begin{pmatrix} x \\ y \end{pmatrix}$ innanför eller utanför kvadraten och vill avbilda den kan vi utnyttja att den går att uttrycka i basvektorerna $\begin{pmatrix} x \\ y \end{pmatrix} = x\vec{e}_1 + y\vec{e}_2$, så

[8]En avbildning som avbildar $\begin{pmatrix} x \\ y \end{pmatrix}$ på $\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$, där $\begin{pmatrix} e \\ f \end{pmatrix}$ är en konstant vektor, kallas för en *translation* och beskriver en förflyttning utan vridning. I det fallet avbildas nollvektorn på $\begin{pmatrix} e \\ f \end{pmatrix}$ så translation är inte en linjär avbildning. En avbildning av typen $A\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}$ (d v s en linjär avbildning plus en translation) kallas för en *affin* avbildning.

$A\begin{pmatrix} x \\ y \end{pmatrix} = x\vec{v}_1 + y\vec{v}_2$. D v s bilden av $\begin{pmatrix} x \\ y \end{pmatrix}$ hittar man om man går x gånger i riktning $\vec{v}_1$ och y gånger i riktning $\vec{v}_2$.

I figuren har vi tänkt på $x = 3$ och $y = -1$. Genom att låta x och y vara heltal kan man nu förstå att bilden av det kvadratiska rutnätet som ges av heltalsmultiplar av basvektorerna avbildas på ett rutnät av parallellogram. Vi får alltså ett nytt koordinatsystem där $\vec{v}_1$ och $\vec{v}_2$ fungerar som basvektorer.[9]

Men varken x eller y behöver förstås vara heltal.

Övning 6.1. *Visa, genom att variera x och y mellan 0 och 1 att alla punkter i kvadraten avbildas på alla punkter i parallellogrammen. Vilka värden på x och y ska man välja för att studera hur en viss kant (välj själv) i kvadraten avbildas?*

Man kan också förstå att varje kvadrat med sidorna parallella med koordinataxlarna avbildas på ett parallellogram likformigt med parallellogrammet på höger sida. Om man vill avbilda ett annat område kan man tänka sig att det består av små axelparallella kvadrater, där man väljer dessa kvadraters storlek beroende på hur noggrann man vill vara, och sedan avbildar man kvadraterna.

Vi förstår alltså hela avbildningen genom att studera hur fyra punkter i en kvadrat avbildas. Nu ska vi titta på några linjära avbildningar som svarar mot enkla geometriska operationer.

Exempel 1. Beskriv den linjära avbildningen $A = \begin{pmatrix} 3 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$. Vi ritar ut enhetskvadraten och dess bild.

Vi ser att avbildningen drar ut i x-led (faktor 3) och trycker ihop i y-led (faktor 2). En cirkel med radie r t ex kommer avbildas på en ellips med halvaxlar $3r$ och $r/2$.

[9] Så här fint blir det dock inte alltid, se övning *6.2d* eller slutet på kapitel 7 för exempel på undantag.

Exempel 2. Nu det omvända problemet, låt B vara den linjära avbildningen som rotaterar hela planet 180° kring origo (vi motiverar varför rotation är en linjär avbildning längre fram i detta kapitel). Hur ser matrisen B ut?

Det räcker att titta på basvektorerna. Roterar man $\vec{e}_1$ 180° får man ju $-\vec{e}_1$. Den första kolonnen i B ska alltså vara $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$. På samma sätt gäller $B\vec{e}_2 = -\vec{e}_2$, så den andra kolonnen i B måste vara $\begin{pmatrix} 0 \\ -1 \end{pmatrix}$. Nu kan vi skriva upp matrisen: $B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. Vi kontrollerar genom att avbilda en godtycklig vektor: $B\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x \\ -y \end{pmatrix} = -\begin{pmatrix} x \\ y \end{pmatrix}$.

Övning 6.2. *Beskriv geometriskt de linjära avbildningar som svarar mot följande matriser, genom att avbilda kvadraten med hörn i* $(0,0)$, $(1,0)$, $(0,1)$ *och* $(1,1)$*:*

a) $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ *b)* $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ *c)* $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ *d)* $\begin{pmatrix} 3 & 0 \\ 0 & 0 \end{pmatrix}$ *e)* $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

Övning 6.3. *Skriv upp matrisen för avbildningen "spegling i x-axeln" genom att studera hur basvektorerna avbildas. Varje punkt avbildas alltså på den punkt den ser ut att vara i om den tittar på sig själv i en spegel som ligger längs x-axeln.*

En viktig matris är denna: $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ Den kallas för *identitetsmatrisen.* När man multiplicerar med den får man tillbaka samma vektor, $I\vec{v} = \vec{v}$, oavsett vektorn $\vec{v}$ så allting ser identiskt ut. Den fungerar som talet 1 gör vid multiplikation av tal.

Rotation är förstås en linjär avbildning, för vår övertygelse tänker vi igenom villkoren för linearitet:

Övning 6.4. *Hur ser matrisen R ut som beskriver rotation en vinkel θ? Ledning:*

Kalla bildvektorerna till basvektorerna för $\vec{v}_1$ och $\vec{v}_2$. Observera att $\vec{v}_1$ och $\vec{v}_2$ är också enhetsvektorer. Därför kan man uttrycka koordinaterna till dessa vektorer i bara sinus och cosinus, se figuren. Matrisen ges sedan av $R = (\vec{v}_1 \ \vec{v}_2)$. Kontrollera ditt svar med matrisen B i ett tidigare exempel samt övningarna 6.2 a och c ovan. (*En annan koll du kan göra är att sätta in $\theta = 0$. Då ska du få fram identitetsmatrisen eftersom rotation 0° svarar mot ingen rotation alls.*)

Övning 6.5. *De här övningarna kräver mer arbete. Hitta matriserna till följande linjära avbildningar:*

a) Projektion på linjen $y = kx$.

b) Spegling i linjen $y = kx$.

7. Areaskala och orientering

Vi har sett att $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ avbildar basvektorerna på den parallellogram som ges av kolonnvektorerna i A.

Övning 7.1. *Visa att arean av parallellogrammen som ges av* $\begin{pmatrix} a \\ b \end{pmatrix}$ *och* $\begin{pmatrix} c \\ d \end{pmatrix}$ *i figuren nedan är* $ad - bc$.

Eftersom arean av en enhetskvadrat är 1 så ser vi att arean ändras med en faktor $ad - bc$. Eftersom avbildningen är linjär kommer alla områden som avbildas med denna avbildning att förändra arean med denna faktor. Man säger att *areaskalan* (förstoringsgraden) är $ad - bc$ i det här fallet. Om areaskalan är mindre än 1 är det alltså frågan om en förminskning.

Övning 7.2. *Bestäm areaskalan för* $\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$. *Är avbildningen förstorande eller förminskande?*

Nu kan man förledas att tro att vi bevisat att areaskalan för en matris $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ är $ad - bc$. Det har vi inte gjort, och det är inte riktigt sant heller. Figuren i övning 7.1 är ju inte representativ för alla avbildningar, den bara hänsyn till det fallet att vektorerna ligger på ett visst sätt. Det skulle ju kunna se ut t ex så här istället.

I figuren till vänster kan vi lätt beräkna areaskalan med hjälp av övning 7.1: byt bara namn på vektorerna, d v s byt plats på a och c samt b och d. Resultatet är $bc - ad = -(ad - bc)$, så i det fallet är alltså $ad - bc$ negativ. I stället för att gå igenom alla olika fallen, ger vi i Appendix A ett trigonometriskt bevis för att areaskalan för $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ är $|ad - bc|$, d v s $ad - bc$ om detta är positivt, annars $bc - ad$.

Talet $ad - bc$ kallas för *determinanten* till matrisen $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$ och den betecknas $\det A$. Determinanten är alltså antingen +areaskalan eller −areaskalan. Vad innebär det att determinanten är negativ? Vi jämför bilden av basvektorerna i matriserna $A = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ och $B = \begin{pmatrix} 1 & 3 \\ 2 & 1 \end{pmatrix}$. Beteckningar: $\vec{v}_1 = A\vec{e}_1, \vec{v}_2 = A\vec{e}_2$ samt $\vec{w}_1 = B\vec{e}_1, \vec{w}_2 = B\vec{e}_2$.

I avbildningen A har vi $\det A = 5 > 0$. I detta fall har bildvektorerna ”samma orientering” som basvektorerna: ”Om man tittar i den riktning som vektor 1 har så ligger vektor 2 till vänster”. I avbildningen B gäller $\det B = -5 < 0$ och bildvektorerna har ”motsatt orientering” jämfört med basvektorerna, vektor 2 ligger ju till höger.

Tecknet för determinanten bestämmer alltså om orienteringen är oförändrad eller omkastad.

Övning 7.3. *Studera avbildningarna i kapitel 6. Beräkna determinanten och bestäm areaskala samt avgör om orienteringen är densamma eller omkastad. Notera speciellt determinanten för rotation, spegling och projektion.*

Vi har hoppat över fallet där determinanten är 0, $ad = bc$. I detta fall är kolonnvektorerna i matrisen A parallella eftersom $\frac{a}{b} = \frac{c}{d}$.[10] Hela planet avbildas då på en linje parallell med kolonnvektorerna (eller på origo). Bilden av enhetskvadraten blir inte en parallellogram utan en sträcka (eller origo) som förstås har area 0, varför areaskalan blir 0.

Om $\det A = 0$ så kallas matrisen A för *singulär* (vilket betyder ”speciell” på engelska). Detta är ett viktigt fall att hålla reda på eftersom avbildningen beter sig så annorlunda.

8. Sammansättning av avbildningar

Vi ska nu utföra flera linjära avbildningar efter varandra. Låt oss börja med två avbildningar, A och B. Om vi utför först A och sedan B och följer en vektor $\vec{v}$ genom avbildningarna kan det se ut så här:

Vi antar att $A\vec{v} = \vec{w}$ och $B\vec{w} = \vec{u}$. När vi *sätter samman* dessa två avbildningar får vi en avbildning direkt från första till tredje koordinatsystemet, låt oss kalla denna avbildning C. Nu ska vi se hur man beräknar C utgående från A och B. Från sambandet mellan vektorerna ovan ser vi att

$$\vec{u} = B\vec{w} = BA\vec{v}$$

Vi påminner om att A och B är 2×2-matriser och att $\vec{v}$ är en 2×1-matris. Man kan tycka att vi i ekvationen ovan borde skriva $\vec{u} = B(A\vec{v})$ för att betona att man först ska räkna ut $A\vec{v}$ (d v s $\vec{w}$) och sedan multiplicera med B. Paranteserna är dock inte nödvändiga. På samma sätt som när vi multiplicerar tre tal xyz kan beräkna detta antingen som $(xy)z$ eller $x(yz)$ vilket ger samma resultat, gäller denna regel för matriser också. Detta kallas för den *associativa lagen*

[10] Här antar vi förstås att varken b eller d är 0. Om $b = 0$ så är antingen $a = 0$ eller $d = 0$. I fallet $a = 0$ så är första kolonnvektorn $\binom{0}{0}$ nollvektorn, vilken är parallell med alla vektorer. I fallet $d = 0$ så är kolonnvektorerna $\binom{a}{0}$ och $\binom{c}{0}$, vilka är parallella. På samma sätt får man samma resultat om man startar med $d = 0$.

vid multiplikation.[11] Vi kan t ex skriva $\vec{u} = (BA)\vec{v}$. Men vad är BA för något? Minnesregeln för matrismultiplikation från kapitel 5 ger (gå tillbaka om du hoppade över detta förut): $(2\times\cancel{2})\cdot(\cancel{2}\times 2) = (2\times 2)$ d v s att matrisprodukten går bra och att resultatet är en 2×2-matris.

BA är alltså en 2×2-matris som avbildar $\vec{v}$ på $\vec{u}$. Eftersom det här resonemanget fungerar oavsett vilken vektor $\vec{v}$ vi väljer, måste BA vara den sökta avbildningen C. Alltså är C en linjär avbildning och $C = BA$. Observera ordningen på matriserna A och B.

Övning 8.1. *Låt A vara avbildningen "rotation $90°$ kring origo" och B vara avbildningen "spegling i linjen $y = x$".*

a) Skriv upp matriserna A och B.

b) Låt C vara den sammansatta avbildningen där man först vrider $90°$ och sedan speglar. Beräkna matrisen C. Vilken enkel geometrisk operation utför C?

c) Låt D vara den sammansatta avbildningen där man först speglar och sedan vrider $90°$. Beräkna matrisen D. Vilken enkel geometrisk operation utför D?

Detta är alltså ännu ett exempel som visar att den kommutativa lagen inte gäller matrismultiplikation eftersom $BA \neq AB$

Övning 8.2. *Beräkna sammansättningen av rotation en vinkel α och rotation en vinkel β på två sätt:*

a) Kalla matrisen för rotation en vinkel θ för R_θ. Skriv upp matriserna R_α och R_β (se övning 6.4). Beräkna den sammansatta avbildningen genom att multiplicera matriserna: $R_\beta R_\alpha$

b) Tänk litet och kom på vilken enkel geometrisk operation som den sammansatta avbildningen beskriver. Efter denna insikt kan du direkt skriva upp matrisen för den sammansatta avbildningen (med hjälp av övning 6.4).

Eftersom dessa två matriser beskriver samma linjära avbildning är matriserna lika, d v s varje element i den ena matrisen är lika med motsvarande elementet i den andra matrisen. Du har nu bevisat två trigonometriska formler som du antagligen stött på tidigare.

Kommentar: Efter man satt matrisen i *a)* lika med matrisen i *b)* är det lätt att bevisa att $R_\beta R_\alpha = R_\alpha R_\beta$. Ibland kan man alltså byta ordning i en matrisprodukt. Om den kommutativa lagen gällde så skulle man alltid kunna byta ordning, men bara för att man kan byta ordning ibland betyder det inte att den kommutativa lagen gäller, som vi sett exempel på tidigare.

Övning 8.3. *Beräkna sammansättningen av en projektionen med sig själv. Ta t ex $P = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$ och beräkna P^2. Geometrisk tolkning av resultatet. Beräkna även P^n där n är ett positivt heltal.*

Övning 8.4. *Antag att A och B är 2×2-matriser. Visa att determinanten av produkten är produkten av determinanterna,* $\det AB = (\det A)(\det B)$, *genom att studera areaskalan och orienteringen i den sammansatta avbildningen.*

Om man sätter samman flera linjära avbildningar än två blir det förstås också en linjär avbildning, det är bara att ta två i taget och utnyttja den associativa lagen.

[11] Låt A, B och C vara tre linjära avbildningar som man kan sätta samman i tur och ordning. Den associativa lagen vid multiplikation säger att $(CB)A = C(BA)$. Varför gäller den? Vi följer en godtycklig vektor $\vec{v}_1$ genom avbildningarna som i figuren.

Hur avbildas $\vec{v}_1$ av avbildningarna $(CB)A$ respektive $C(BA)$? Vi ser i figuren att $(CB)A\vec{v}_1 = (CB)\vec{v}_2 = \vec{v}_4$ och på samma sätt får vi $C(BA)\vec{v}_1 = C\vec{v}_3 = \vec{v}_4$. D v s $(CB)A$ är samma avbildning som $C(BA)$.

9. Invers avbildning

Det är lätt att skriva upp matrisen till en linjära avbildning som avbildar enhetskvadraten på ett parallellogram som ges av vektorerna $\begin{pmatrix} a \\ b \end{pmatrix}$ och $\begin{pmatrix} c \\ d \end{pmatrix}$: $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$. Hur ser matrisen ut till den omvända avbildningen, den som avbildar parallellogrammen på enhetskvadraten?

Denna avbildning (om den finns) kallas för den *inversa avbildningen* och betecknas med A^{-1}.[12] Vi ska hitta matrisen till den inversa avbildningen genom att dela upp den i enklare avbildningar och sedan sätta samman dessa. Låt oss först titta på en avbildning där det är lätt att skriva upp den inversa matrisen. Den här sortens avbildning kallas för *skjuvning*:[13]

$$M = \begin{pmatrix} 1 & h \\ 0 & 1 \end{pmatrix} \text{ och } M^{-1} = \begin{pmatrix} 1 & -h \\ 0 & 1 \end{pmatrix}$$

Vi kontrollerar bilden. Det räcker förstås att avbilda vektorerna som är sidorna i parallellogrammen. Klart M avbildar basvektorerna på $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ och $\begin{pmatrix} h \\ 1 \end{pmatrix}$. Nu kollar vi den inversa avbildningen: $M^{-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ och $M^{-1}\begin{pmatrix} h \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Vi får tillbaka basvektorerna, som vi avsåg.

Nu ska vi hitta A^{-1}. Vi utför vissa avbildningar utan att i detalj förklara hur vi kom fram till dessa. Så får man lov att göra ibland för att inte bli för omständig. Det viktiga är att du kontrollerar räkningarna och ser att de gör vad vi påstår att de gör. Först skjuvar vi tillbaka och skalar om i vertikal led:

$$M_1 = \begin{pmatrix} 1 & 0 \\ -b & a \end{pmatrix}$$

[12] Om man vill vara noggrann så definierar man A^{-1} som en matris med egenskapen att $AA^{-1} = I$ och $A^{-1}A = I$. Invers matris motsvarar division mellan matriser. Jämför med två tal a, b där $a \neq 0$: $a^{-1}b = ba^{-1} = \frac{b}{a}$. Däremot kan man inte skriva $\frac{B}{A}$ med två matriser eftersom $A^{-1}B$ och BA^{-1} inte behöver vara lika.

[13] Observera att areaskalan är 1, rombens area är lika med kvadratens.

Sedan skjuvar vi tillbaka i horisontell led.

$$M_2 = \begin{pmatrix} 1 & \frac{-c}{ad-bc} \\ 0 & 1 \end{pmatrix}$$

Slutligen skalar vi om med en faktor $1/a$ i x-led och $1/(ad-bc)$ i y-led.

$$M_3 = \begin{pmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{ad-bc} \end{pmatrix}$$

Nu sätter vi samman avbildningarna. För att påminna oss om ordningen ser vi att

$$M_3 M_2 M_1 \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

och

$$M_3 M_2 M_1 \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

så den sökta inversen är $A^{-1} = M_3 M_2 M_1$.

Övning 9.1. *Visa att* $A^{-1} = M_3 M_2 M_1 = \begin{pmatrix} \frac{d}{ad-bc} & \frac{-c}{ad-bc} \\ \frac{-b}{ad-bc} & \frac{a}{ad-bc} \end{pmatrix} = \frac{1}{ad-bc} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$.

I härledningen antog vi att avbildningen såg ut på ett speciellt sätt. Gäller formeln för inversen i andra fall?

Övning 9.2. *Antag att* $ad - bc \neq 0$. *Visa, genom att använda formeln för* A^{-1} *i övning 9.1, att* $AA^{-1} = I$ *och* $A^{-1}A = I$. *Slutsatsen är att denna formel gäller oavsett matris (så länge* $ad - bc \neq 0$*).*

Om $ad - bc \neq 0$ så kallas matrisen A *inverterbar* (d v s samma sak som icke singulär). Om A är inverterbar så har alltså A en invers och vi kan beräkna den med formeln i övning 9.1.

Övning 9.3. *Beräkna inversen till avbildningen "rotation en vinkel* α *kring origo". Beräkna inversen till en spegling. Tänk efter i båda fallen vad svaren borde bli och kolla att det stämmer. Kontrollera slutligen svaret genom att multiplicera matris med invers matris och se att du får identitetsmatrisen.*

Exempel 1. Ekvationslösning.

Vi ska visa hur matrisinvertering kan användas för att lösa ekvationssystem. Betrakta följande ekvationssystem:

$$\begin{aligned} -2x + 3y &= 4 \\ x - y &= 5 \end{aligned}$$

Vi ska hitta tal x och y som löser båda ekvationerna. Vi skriver om det som en matrisekvation:

$$\begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix} \tag{9.1}$$

Definiera $A = \begin{pmatrix} -2 & 3 \\ 1 & -1 \end{pmatrix}$ och beräkna $A^{-1} = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}$.[14] Multiplicera båda led i ekvation (10.2) med A^{-1} **från vänster**:

$$A^{-1}A\begin{pmatrix} x \\ y \end{pmatrix} = A^{-1}\begin{pmatrix} 4 \\ 5 \end{pmatrix}$$

Vänster sida kan vi förenkla på följande sätt:

$$A^{-1}A\begin{pmatrix} x \\ y \end{pmatrix} = I\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

där I förstås är identitetsmatrisen. Utför multiplikationen på höger sida:

$$A^{-1}\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix}$$

Slutsatsen är således

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 8 \\ 12 \end{pmatrix}$$

och vi har löst ekvationssystemet.

Finns det någon invers om determinanten är 0, d v s $ad = bc$ (t ex projektioner)? Vi såg i slutet av kapitel 7 att kolonnvektorerna i detta fall är parallella. Finns det någon möjlighet att avbilda kolonnvektorerna på basvektorerna?

Om vi avbildar $\begin{pmatrix} a \\ b \end{pmatrix}$ på $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, så kommer p g a att A är linjär $\begin{pmatrix} c \\ d \end{pmatrix}$ avbildas på en vektor parallell med $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Bilden av kolonnvektorerna kommer alltså alltid att vara parallella och vi kan inte återskapa basvektorerna. Det finns ingen invers. Projektioner saknar därmed invers.

Övning 9.4. *Antag att A är inverterbar. Visa att* $\det A^{-1} = 1/(\det A)$, *genom att använda resultatet i övning 8.4.*

Exempel 2. Antag att A är en linjär avbildning och

$$A\vec{v} = \vec{0}$$

Vi ska bevisa att om $\vec{v} \neq \vec{0}$ så måste A vara singulär, d v s om $\vec{v}$ inte är nollvektorn, så saknar A invers (detta kommer vi ha användning för i kapitel 10).

Vi börjar beviset på ett sätt som kan tyckas litet bakochfram. Vi antar att A **har** en invers. Vad innebär det? Då kan man multiplicera ekvationen ovan med A^{-1} från vänster:

$$A^{-1}A\vec{v} = A^{-1}\vec{0}$$

Vi förenklar båda sidor (jämför föregående exempel):

$$\vec{v} = \vec{0}$$

Därmed kan vi dra den önskade slutsatsen att om $\vec{v} \neq \vec{0}$ så saknar A invers. Varför då? Jo, om $\vec{v} \neq \vec{0}$ och A har invers så vet vi med resonemanget ovan att $\vec{v} = \vec{0}$. Men det går ju inte att $\vec{v} \neq \vec{0}$ och $\vec{v} = \vec{0}$ samtidigt, alltså kan inte A ha en invers. Stanna upp ett tag och tänk igenom detta.[15]

[14] Vi ser att det finns en invers eftersom $ad - bc = -1 \neq 0$.

[15] Denna bevismetod kallas för *indirekt bevis*.

10. Egenvärden och egenvektorer

Ibland avbildas vissa vektorer på en parallell vektor.

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

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

Vi säger att $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ är *egenvektor* till $\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ med *egenvärde* $r_1 = 2$. På samma sätt är $\vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ egenvektor till $\begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ med egenvärde $r_1 = -1$.

Övning 10.1. *Titta tillbaks på uppgift 5.7 och 5.9a). Kan vi säga något om egenvärden eller egenvektorer till de kvadratiska matriserna?*

Övning 10.2. *Utnyttja den geometriska karaktäriseringen av matriserna i övning 6.2 b)–e) och hitta egenvärden och egenvektorer till matriserna. (Finns det en vektor som avbildas på en parallell vektor? Finns det flera?) Förstå vilka egenvärden en spegling har respektive vilka egenvärden en projektion har.*

I allmänhet defineras egenvärden och egenvektorer på följande vis:

Definition. Antag att vi har en matris A, ett tal r och en vektor $\vec{v} \neq \vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Om

$$A\vec{v} = r\vec{v} \tag{10.1}$$

så kallas $\vec{v}$ för *egenvektor* till A med *egenvärde* r.

Två kommentarer:

1. Talet 0 kan vara ett egenvärde (se övning 6.2 *d*), men $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ räknas inte som egenvektor.

2. Antag att $\vec{v}$ är en egenvektor till A med egenvärde r. Låt $\vec{u}$ vara en parallell vektor: $\vec{u} = t\vec{v}$. Avbilda $\vec{u}$ med A:

$$A\vec{u} = At\vec{v} = tA\vec{v} = tr\vec{v} = rt\vec{v} = r\vec{u}$$

Vi använde här att A är linjär och vårt antagande om $\vec{v}$. Om vi tittar på första och sista ledet, ser vi att $\vec{u}$ också är en egenvektor till A med egenvärde r. Slutsatsen är att egenvektorer inte är entydiga utan varje parallell vektor till en egenvektor (utom nollvektorn) är en egenvektor med samma egenvärde.

Nu till det allmänna problemet, givet en matris $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, hur hittar vi egenvärden och egenvektorer? D v s hur hittar vi r och $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ som uppfyller

$$A\vec{v} - r\vec{v} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{10.2}$$

(vi har bara flyttat om litet i ekvation (10.1)). Vi använder identitetsmatrisen och skriver om $r\vec{v}$:

$$r\vec{v} = rI\vec{v} = \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \vec{v}$$

För in detta i ekvation (10.2), med $\vec{v} = \begin{pmatrix} x \\ y \end{pmatrix}$

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} - \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

För matriser gäller den *distributiva lagarna*, t ex $(A - B)C = AC - BC$. Nu använder vi denna regel baklänges och bryter ut $\begin{pmatrix} x \\ y \end{pmatrix}$ åt höger:

$$\left(\begin{pmatrix} a & c \\ b & d \end{pmatrix} - \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

vilket vi kan skriva som

$$\begin{pmatrix} a - r & c \\ b & d - r \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{10.3}$$

Om $\begin{pmatrix} x \\ y \end{pmatrix}$ skall vara en egenvektor så kan inte både x och y vara 0, enligt definitionen. Därför måste matrisen $\begin{pmatrix} a - r & c \\ b & d - r \end{pmatrix}$ vara singulär, se Exempel 2 i kapitel 9. Detta innebär att determinanten $(a-r)(d-r)-bc$ är 0. Denna ekvation kallas den *karaktärisktiska ekvationen*, vi skriver den som:[16]

$$r^2 - (a + d)r + ad - bc = 0$$

Lösningarna till denna andragradsekvation, är således egenvärdena till matrisen A, kalla lösningarna r_1 och r_2.

Om vi nu sätter $r = r_1$ i ekvation (10.3) och löser ekvationssystemet för x och y så har vi hittat en egenvektor med egenvärde r_1. Kalla egenvektorn $\vec{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, den uppfyller alltså

$$\begin{pmatrix} a - r_1 & c \\ b & d - r_1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

d v s

$$(a - r_1)x_1 + cy_1 = 0 \tag{10.4}$$
$$bx_1 + (d - r_1)y_1 = 0 \tag{10.5}$$

I praktiken är detta ekvationssystem lätt att lösa. Eftersom $\begin{pmatrix} a - r_1 \\ b \end{pmatrix}$ och $\begin{pmatrix} c \\ d - r_1 \end{pmatrix}$ är parallella (de är kolonnvektorerna i en singulär matris) så är varje lösning till ekvation (10.4) också en lösning till ekvation (10.5). Därför kan man välja vilken lösning $\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ som helst till någon av ekvationerna, och det finns många (jämför kommentar 2 tidigare i kapitlet).

Därefter gör man på motsvarande sätt med det andra egenvärdet, r_2, och hittar dess egenvektor som vi kan kalla $\vec{v}_2$.

Exempel 1. Vi ska ta fram egenvärden och egenvektorer till övning 5.7 c). Egenvärdesekvationen blir $r^2 + r - 6 = 0$ med rötter $r_1 = 2$ och $r_2 = -3$. För egenvärdet $r_1 = 2$ måste motsvarande egenvektor $\vec{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ uppfylla t ex ekvation (10.4):

$$-4x_1 + 4y_1 = 0$$

[16] Lägg märke till att determinanten $ad - bc$ finns med i ekvationen. Den andra koefficienten, $a + d$. summan av diagonalelementen i matrisen A, har också ett namn; den kallas för *spåret*.

Lösningen kan väljas till $\vec{v}_1 = \begin{pmatrix}1\\1\end{pmatrix}$ eller en multipel därav. Denna egenvektor kände vi redan till från övningen. Det andra egenvärdet $r_2 = -3$ ger på samma sätt följande ekvation för egenvektorn $\vec{v}_2 = \begin{pmatrix}x_2\\y_2\end{pmatrix}$:

$$x_2 + 4y_2 = 0$$

En lösning i detta fall är $\vec{v}_2 = \begin{pmatrix}4\\-1\end{pmatrix}$. Vi kontrollerar att vi räknat rätt:

$$A\vec{v}_2 = \begin{pmatrix}-2 & 4\\1 & 1\end{pmatrix}\begin{pmatrix}4\\-1\end{pmatrix} = \begin{pmatrix}-12\\3\end{pmatrix} = -3\begin{pmatrix}4\\-1\end{pmatrix} = r_2\vec{v}_2$$

Övning 10.3. *Hitta egenvärden och egenvektorer till övning 5.7 b).*

Även om bilden av basvektorerna bestämmer avbildningen, är det inte alltid så tydligt vad avbildningen gör. Vi får mer kunskap om avbildningen med tillgång till egenvärden och egenvektorer. I den riktning som ges av en egenvektorer är ju avbildningen mycket enkel, en förstoring eller förminskning och möjligen en omkastning av riktning beroende på om motsvarande egenvärde r är negativt.

Övning 10.4. *Titta på den första figuren i detta kapitel. Determinanten för avbildningen är -2 och detta är också produkten av egenvärdena: $2 \cdot (-1)$. Bevisa att determinanten alltid är lika med produkten av egenvärdena.*

Övning 10.5. *Hur ser egenvärdena ut till rotationer, t ex övning 6.2 a)? (Vilka sorts lösningar har egenvärdesekvationen? Se även Appendix C.)*

En matris $A = \begin{pmatrix}a & b\\b & d\end{pmatrix}$ kallas *symmetrisk* eftersom den är symmetrisk med avseende på spegling i diagonalen a–d.[17]

Övning 10.6. *Denna övning är svår. Visa att symmetriska matriser har (reella) egenvärden. (Ledning: Skriv upp den allmänna formeln för lösningen till den karaktäristiska ekvationen för en symmetrisk matris. Använd den första och andra kvadreringsregeln under roten. Hur vet man att man alltid har (reella) lösningar?)*

Övning 10.7. *Denna övning är också svår. Visa att symmetriska matriser med olika egenvärden $r_1 \neq r_2$ har vinkelräta egenvektorer $\vec{v}_1 \cdot \vec{v}_2 = 0$. (Ledning: För symmetiska matriser gäller $\vec{u}_1 \cdot A\vec{u}_2 = \vec{u}_2 \cdot A\vec{u}_1$ för alla vektorer $\vec{u}_1$ och $\vec{u}_2$. Visa detta! Tillämpa sedan det på egenvektorerna $\vec{v}_1$ och $\vec{v}_2$.)*

[17] Inom kvantmekaniken är ett slags symmetriska matriser det centrala objektet. T ex energinivåerna i en atom kan inte anta vilka värden som helst. De möjliga energierna är egenvärdena till en matris.

11. Appendix A – Trigonometrisk härledning av areaskalan

Vi ska visa hur stor arean är i ett allmänt parallellogram som ges av $\begin{pmatrix} a \\ b \end{pmatrix}$ och $\begin{pmatrix} c \\ d \end{pmatrix}$.

Antag att den minsta vinkeln mellan vektorerna är α. Parallellogrammen består av två kongruenta trianglar. Låt A, B och C beteckna längden på sidornas i triangeln som i figuren ovan.

$$A^2 = a^2 + b^2$$

$$B^2 = (a-c)^2 + (b-d)^2$$

$$C^2 = c^2 + d^2$$

Enligt cosinussatsen gäller följande samband

$$B^2 = A^2 + C^2 - 2AC\cos\alpha \tag{11.1}$$

Enligt sinussatsen får vi parallellogrammens area P (som är dubbla trianglens area) på följande sätt.

$$P = AC\sin\alpha \tag{11.2}$$

Enligt "trigonometriska ettan" så gäller att $\cos^2\alpha + \sin^2\alpha = 1$ oavsett vinkel α. Vi ska nu använda detta samband och bli av med vinkeln α, som inte intresserar oss. Skriv om ekvation (11.1) som

$$\frac{A^2 + C^2 - B^2}{2} = AC\cos\alpha \tag{11.3}$$

Tag kvadraten på båda sidor av denna ekvation och addera kvadraten på båda sidor av ekvation (11.2):

$$\frac{(A^2 + C^2 - B^2)^2}{4} + P^2 = A^2C^2\cos^2\alpha + A^2C^2\sin^2\alpha =$$

$$= A^2C^2\left(\cos^2\alpha + \sin^2\alpha\right) = A^2C^2$$

där vi brutit ut A^2C^2 och använt den trigonometriska ettan. Vi flyttar över första termen från vänsterledet så vi får P^2 ensamt där, det är ju P vi letar efter.

$$P^2 = A^2C^2 - \frac{(A^2 + C^2 - B^2)^2}{4}$$

Nu förenklar vi: Använd de första formlerna som relaterar stora bokstäver till små. Första termen:

$$A^2C^2 = (a^2 + b^2)(c^2 + d^2) = a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2$$

Det väsentliga i andra termen:

$$A^2 + C^2 - B^2 = a^2 + b^2 + c^2 + d^2 - \Big((a-c)^2 + (b-d)^2\Big) =$$

$$= a^2 + b^2 + c^2 + d^2 - \Big(a^2 - 2ac + c^2 + b^2 - 2bd + d^2\Big) =$$

$$= a^2 + b^2 + c^2 + d^2 - a^2 + 2ac - c^2 - b^2 + 2bd - d^2 =$$

$$= 2ac + 2bd = 2(ac + bd)$$

Alltså blir den andra termen:

$$\frac{(A^2 + B^2 + C^2)^2}{4} = (ac + bd)^2 = a^2c^2 + 2adbc + b^2d^2$$

Subtrahera första termen från andra

$$P^2 = a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 - \left(a^2c^2 + 2adbc + b^2d^2\right) =$$

$$= a^2c^2 + a^2d^2 + b^2c^2 + b^2d^2 - a^2c^2 - 2adbc - b^2d^2 =$$

$$= a^2d^2 + b^2c^2 - 2adbc = (ad - bc)^2$$

där vi i sista ledet använde andra kvadreringsregeln baklänges. Slutligen drar vi roten ur båda led. Positiva kvadratroten måste det vara eftersom P är en area.

$$P = \sqrt{(ad - bc)^2}$$

d v s arean är $ad-bc$ om detta är positivt, annars $bc-ad$. Detta kan man också skriva som $|ad-bc|$.

12. Appendix B – Linjära avbildningar i rummet

I detta avsnitt tittar vi på linjära avbildningar i (det tre-dimensionella) rummet. Resultaten är analoga med motsvarande resultat i planet, vi skriver ner dem utan motivering.

En linjär avbildning i rummet definieras som tidigare, men vektorerna har tre komponenter istället för två. En linjär avbildning i rummet är detsamma som multiplikation med en 3×3-matris. En godtycklig linjär avbildning i rummet ser alltså ut så här.

$$A = \begin{pmatrix} a & d & g \\ b & e & h \\ c & f & i \end{pmatrix}$$

Determinanten har en mer komplicerad formel

$$\det A = aei - ahf + bfg - bdi + cdh - ceg$$

(Det är ingen mening att lära sig denna utantill, när man ska beräkna dessa determinanter finns det andra metoder.) Absolutbeloppet av determinanten är *volymen* av det parallellepiped som spänns upp av kolonnvektorerna i matrisen. Om alla kolonnvektorerna ligger i ett och samma plan är alltså determinanten 0.

Man kan tala om *orientering* av vektorer i rummet, två vektorer definierar ett plan och den tredje kan ligga på den ena eller andra sidan av planet.[18] Tecknet för determinanten avgör om orienteringen bevaras eller omkastas. Byter man plats på två kolonnvektorer eller två radvektorer ändrar determinanten tecken.

Om determinanten är skild från 0, så finns det en invers matris A^{-1} som uppfyller $AA^{-1} = A^{-1}A = I$, där I är identitetsmatrisen

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Det finns ingen enkel formel för inversen till en 3×3-matris A, man beräknar i allmänhet inversen genom att lösa ett ekvationssystem (nämligen $AA^{-1} = I$).

Egenvärden och egenvektorer defineras precis som tidigare. (Vi har ju redan sett ett exempel, övning 5.9 *a)*.) När man ska räkna ut egenvärden och egenvektorer kommer man med samma resonemang som för ekvation (10.3) i kapitel 10 fram till att

$$\det \begin{pmatrix} a-r & d & g \\ b & e-r & h \\ c & f & i-r \end{pmatrix} = 0.$$

Detta leder till en tredjegradsekvation i r och det finns formler för att lösa dessa.[19]

[18] Detta är något som man kan känna igen från fysiken, t ex induktionskraftens riktning på en elektriskt ström i ett magnetfält.

[19] Det finns formler för lösning av allmänna fjärdegradsekvationer också. Dessa formler för tredje- och fjärdegradsekvationer innehåller (precis som formeln för andragradsekvationen) de fyra räknesätten och rotutdragning. Man kan visa att för ekvationer av femte och högre grad finns det däremot inga formler! Med detta menas inte bara det faktum att ingen hittat någon formel, utan det mycket starkare påståendet att det inte går att skriva upp den allmänna lösningen på detta sätt. Detta resultat är en del av ämnet *Galoisteori* och slutsatsen baseras på vissa symmetrier i ekvationer, se Appendix D.

13. Appendix C – Komplexa tal

I övning 10.5 letade vi lösning till ekvationen $x^2 + 1 = 0$. Den saknar reella lösningar, men har lösning om man tillåter lösningen att vara ett *komplext tal*. I detta avsnitt berättar vi litet om komplexa tal.[20]

Det kan tyckas konstigt för oss idag, men det fanns en tid när man funderade mycket över om ekvationen $x^2 - 2 = 0$ hade lösning. Lösningarna $\pm\sqrt{2}$ är självklara för oss. Beteckningen $\sqrt{2}$ är en symbol för ett tal vars kvadrat är lika med 2. Genom att räkna med symbolen och titta på en tallinje och tänka sig en punkt någonstans nära 1.4142 har vi vant oss vid symbolen.

Pythagoréerna aldrig inte ha nöjt sig med detta. För dessa var ett *tal* samma sak som ett *rationellt tal*, d v s ett tal som kan skrivas som $\frac{p}{q}$ där p och q är heltal och $q \neq 0$. Man kan bevisa att $\sqrt{2}$ inte kan skrivas på detta sätt, och $\sqrt{2}$ är därför ett *irrationellt* tal.[21] Så för Pythagoréerna var längden av diagonalen i en enhetskvadrat inte ett tal.

Genom att lägga till de irrationella talen till de rationella talen så får vi de reella talen och tallinjen blir utan hål. Men de reella talen duger alltså inte till att lösa $x^2 + 1 = 0$. För att vänja oss vid lösningar till denna ekvation, inför vi en symbol, i, som uppfyller $i^2 + 1 = 0$ (d v s motsvarar $\sqrt{-1}$). I övrigt behandlar vi den som vilket tal som helst, vi använder de fyra räknesätten (precis som vi gör med $\sqrt{2}$ utan att tänka). Det visar sig att alla tal som man kan bilda på detta sätt kan förenklas (genom att använda $i^2 = -1$ vid strategiska tillfällen) till $a + b \cdot i$, där a och b är reella tal. Dessa tal kallas för *komplexa tal*.

En synonym till ordet "komplex" är "sammansatt" och detta tal är sammansatt av två delar: a kallas *realdel* och b kallas *imaginärdel*. De reella talen är alltså det specialfall av de komplexa talen där imaginärdelen är 0. Bokstaven i kommer förstås från ordet imaginär.

Vad blev det av lösningen till $x^2 + 1 = 0$? Jo, i är förstås en rot och $-i$ är den andra: $(-i)^2 = (-1)^2 i^2 = 1(-1) = -1$. Lösningarna är alltså de komplexa talen med $a = 0$ och $b = 1$ respektive $a = 0$ och $b = -1$. Men nu kommer det fantastiska: Alla lösningar till alla polynomekvationer (inte bara andragradsekvationer) är komplexa tal och kan alltså skrivas $a + bi$.[22] Detta kallas för algebrans fundamentalsats och är ett tillräckligt argument för att man borde känna till komplexa tal. Sedan har de dessutom väldigt tilltalande geometriskt innehåll. Rekommenderas varmt.

[20] Endast reella egenvärden har någon mening i den geometriska bild vi presenterat hittills, så detta avsnitt skänker ingen utökad förståelse för egenvärden. Däremot är komplexa tal i sig en enormt viktig del av matematiken och detta är anledning nog att introducera dem när nu frågan uppkommit.

[21] Det finns ett klassiskt och enkelt bevis för att $\sqrt{2}$ är irrationell som använder sig av en bevisteknik som kallas *motsägelsebevis*.

[22] Däremot finns det ingen allmän formel för lösningen till ekvationer av grad fem och högre, se fotnot 16.

14. Appendix D – Gruppteori

Symmetrier har stor betydelse i matematik och fysik. Om en ekvation har en viss symmetri så säger det något om lösningarna till ekvationen. I detta avsnitt ska vi införa en abstrakt algebraisk konstruktion som kallas för en *grupp*. Vi ska använda symmetrier hos en en enkel geometrisk figur som motivering.

Betrakta följande figur:

Vi ska studera linjära avbildningar som *bevarar* figuren. Dessa avbildningar kännetecknas av att om någon utför en sådan medan vi blundar, så märker vi ingen skillnad när vi öppnar ögonen igen. För att kunna diskutera dessa operationer så märker vi temporärt rektangelns hörn.

Figuren är bevarad under rotation 120°, $R = \frac{1}{2}\begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$.

Figuren är bevarad under rotation 240°, $S = \frac{1}{2}\begin{pmatrix} -1 & \sqrt{3} \\ -\sqrt{3} & -1 \end{pmatrix}$.

Figuren är bevarad under rotation 360°, men det är ju samma sak som ingen rotation alls, d v s identitetsavbildningen I. Vi har hittat alla symmetrioperationer som bevarar figuren. Kalla denna uppsättning linjära avbildningar för $G = \{I, R, S\}$.

Observera att om vi sätter samman två avbildningar i G så blir resultatet alltid en av avbildningarna i G. D v s om vi multiplicerar två element i G så blir produkten ett element i G. Detta kallar man att G är *sluten* med avseende på multiplikation.[23] Vi kontrollerar detta påstående med en multiplikationstabell.

[23] Två andra exempel: Mängden positiva heltal är sluten med avseende på addition. Varför? Om man tar två positiva heltal och adderar dessa så får man ett positivt heltal. Mängden positiva heltal är **inte** sluten med avseende på subtraktion. Varför inte?

	I	R	S
I	I	R	S
R	R	S	I
S	S	I	R

Vi ser också att varje element i G har sitt inversa element i G: I är ju sin egen invers och eftersom $SR = RS = I$ så gäller att $S^{-1} = R$ och $R^{-1} = S$. Produkten är associativ, den egenskapen gäller ju för matriser. Slutligen finns det ett identitetselement i G, när man multiplicerar med I så händer ingenting.

Dessa egenskaper: slutenhet, associativitet, förekomst av en identitet och att alla element har invers sammanfattas i begreppet *grupp*. Vi säger att $G = \{I, R, S\}$ är en *grupp* under multiplikation och att figuren har *symmetrigrupp* G.

Låt oss en liten stund diskutera något helt annat, nämligen moduloräkning (eller ”klockräkning”). En timma efter klockan 23.00 är klockan 0.00. Vi kan skriva det som $23 + 1 = 0$. Detta kallas att räkna modulo 24. Tänk dig nu att dygnet vore tre timmar långt. En timma efter klockan två vore det då nytt dygn vilket vi skriver som $2 + 1 = 0$. Nu ska vi addera andra heltal modulo 3, men det finns bara tre tal att välja på: 0, 1 och 2.[24] Kalla talmängden för $H = \{0, 1, 2\}$. Låt oss ställa upp en additionstabell för H:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

Vi ser att H är sluten under addition modulo 3. Det är klart att denna slags addition är associativ eftersom vanlig addition är det. Det finns ett identitetselement i H: 0. Addition med 0 ändrar ju ingenting. Varje element i H har en invers: Eftersom

$$2 + 1 = 1 + 2 = 0$$

så är 2 inversen till 1, vilket man lämpligen skriver som $-1 = 2$. På samma sätt är 1 inversen till 2, $-2 = 1$, och 0 är sin egen invers $-0 = 0$. H är alltså en grupp under addition modulo 3.

Vid det här laget har du säkert lagt märke till likheten mellan G och H. Trots att G består av vissa matriser som man multiplicerar och H tal som man adderar på ett visst sätt, så är det ingen skillnad på dem som grupper betraktade. Tabellen innehåller all information och vi kan bortse från den underliggande modellen. Istället för att studera dessa två system separat kan man renodla deras egenskaper och studera en abstrakt mängs G en operation * som uppfyller de fyra egenskaperna vi nämnt ovan. De slutsatser man kan dra från dessa förutsättningar gäller då för våra två system ovan och alla andra som har samma gruppstruktur. Detta är idén i ämnet *gruppteori*.

[24]Man kan tala om andra tal t ex talet 4 men i det fallet är det egentligen 1 man menar eftersom $3 = 0$ och 4 är talet efter.