Linjär Algebra II

Kompendium till kursen 1MA024

Martin Hershcend

Innehåll

1 Repetition

I det här kapitlet går vi igenom grundläggande begrepp från linjär algebra som du stött på i tidigare kurser. Dessa kommer att vara kritiska för resten av kursen så det är mycket viktigt att du är väl förtrogen med dem. Se till att lösa tillräckligt med övningsproblem, så att begreppen och metoderna sitter väl.

1.1 Linjära ekvationssystem

Väldigt många metoder i linjär algebra leder i slutändan till ett linjärt ekvationssystem. Vi går nu igenom hur dessa löses i allmänhet. Vår strategi är alltid att skriva ner totalmatrisen för ekvationssystemet och göra radoperationer på den för att uppnå en totalmatris (och motsvarande ekvationssystem) av en lättare form. Vi börjar med ett exempel. För tydlighets skull skriver vi ner både ekvationssystemen och deras totalmatriser. I fortsättningen kommer vi bara skriva ner matriserna.

Exempel 1.1.1

Lös ekvationssystemet

$$\begin{cases} 2x + y = 4 \\ x - y = -1 \end{cases}$$

Lösning:

$$\begin{cases} 2x + y = 4 \\ x - y = -1 \end{cases} \qquad \underbrace{\left(\begin{array}{c|c} \overbrace{\begin{matrix} 2 & 1 \\ 1 & -1 \end{matrix}}^{\text{Koefficientmatris}} & \begin{matrix} 4 \\ -1 \end{matrix} \end{array}\right)}_{\text{Totalmatris}}$$

$$\Leftrightarrow \begin{cases} x - y = -1 & \boxed{-2} \\ 2x + y = 4 & \end{cases} \qquad \left(\begin{array}{cc|c} 1 & -1 & -1 \\ 2 & 1 & 4 \end{array}\right) \boxed{-2}$$

$$\Leftrightarrow \begin{cases} x - y = -1 \\ 3y = 6 & \boxed{\tfrac{1}{3}} \end{cases} \qquad \left(\begin{array}{cc|c} 1 & -1 & -1 \\ 0 & 3 & 6 \end{array}\right) \boxed{\tfrac{1}{3}}$$

$$\Leftrightarrow \begin{cases} x - y = -1 \\ y = 2 & 1 \end{cases} \qquad \left(\begin{array}{cc|c} 1 & -1 & -1 \\ 0 & 1 & 2 \end{array}\right) 1$$

$$\Leftrightarrow \begin{cases} x = 1 \\ y = 2 \end{cases} \qquad \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right)$$

Vi ser att det sista systemet har precis en lösning $(x, y) = (1, 2)$. Geometriskt kan ekvationssystemet tolkas som skärningen mellan två linjer. Vi får alltså att linjerna $2x + y = 4$ och $x - y = -1$ skär i punkten $(1, 2)$. Pröva själv att rita linjerna i xy-planet och se om de verkligen skär i punkten $(1, 2)$.

När vi löser ekvationssystem gör vi många räkningar. Det är då lätt hänt att det blir fel någonstans. Därför är det alltid bra att kolla om lösningen stämmer. Sätter vi in $(x, y) = (1, 2)$ i systemet

$$\begin{cases} 2x + y = 4 \\ x - y = -1 \end{cases}$$

så får vi

$$\begin{cases} 2 \cdot 1 + 2 = 4 \\ 1 - 2 = -1 \end{cases}$$

Så det stämmer!

Vi förklarar nu matrisnotationen för allmänna linjära ekvationssystem och vilken form vi försöker uppnå med radoperationer.

- Ett *linjärt ekvationssystem* med n obekanta $x_1, x_2, \ldots, x_{n-1}, x_n$ och m ekvationer har formen

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad \vdots \qquad \ddots \qquad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

där $a_{ij}, b_j \in \mathbb{R}$ är konstanter. Notera att koefficienterna a_{ij} är indexerade så att det första indexet i anger ekvationen, och de andra indexet j anger den obekanta som koefficienten tillhör. Högerleden b_i är indexerade efter vilken ekvation de tillhör.

- En *lösning* till systemet är en n-tippel $(s_1, s_2, \ldots, s_n)$ där varje $s_i \in \mathbb{R}$, sådan att

$$(x_1, \ldots, x_n) = (s_1, \ldots, s_n)$$

  löser *alla* ekvationerna i systemet.

- Att lösa systemet innebär att hitta alla dess lösningar.

- Två ekvationssystem är ekvivalenta om de har precis samma lösningar.

- Till systemet hör följande matriser.

$$\text{Koefficientmatris: } \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \qquad \text{Högerledsmatris: } \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

$$\text{Totalmatris: } \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right)$$

Vi löser ekvationssystem med hjälp av omskrivningar till ekvivalenta ekvationssystem. Dessa omskrivningar motsvarar radoperationer på systemens totalmatriser. Det finns tre elementära radoperationer som vi kommer att använda:

Definition 1.1: Elementära radoperationer

λ Multiplicera en rad med $\lambda \neq 0$.

Byt plats på två rader.

λ Lägg till λ gånger en rad till en annan.

Om vi kan få en matris B från en matris A med en följd radoperationer kallas A och B radekvivalenta. Vi skriver då $A \sim B$.

Vi kan alltid använda radoperationer för att nå en speciell form på vår matris som gör det lätt att bestämma lösningarna till motsvarande ekvationssystem. Denna form kallas radkanonisk.

Definition 1.2

En matris sägas att vara *radkanonisk* om

1. Eventuella nollrader står längst ner.
2. Det första nollskilda elementet i varje rad (när sådant förekommer) är 1. Dessa element kallas ledande ettor.
3. Om en rad har en ledande etta så har raden ovanför (om den finns) ledande etta längre till vänster.
4. Elementen ovanför ledande ettor är 0.

Innan vi förklarar hur man i allmänhet kan nå radkanonisk form går vi igenom några exempel.

Exempel 1.1.2

Lös ekvationssystemet

$$\begin{cases} x_1 + x_2 + 3x_3 = 5 \\ 2x_1 + x_2 + 3x_3 = 1 \\ \phantom{x_1 +{}} x_2 - x_3 = 3 \\ x_1 \phantom{{}+ x_2} + x_3 = 4 \end{cases}$$

Vi utför radoperationer på totalmatrisen för systemet:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 5 \\ 2 & 1 & 3 & 3 \\ 0 & 1 & -1 & 3 \\ 1 & 0 & 1 & 4 \end{array}\right) \boxed{-2}\boxed{-1} \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & 5 \\ 0 & -1 & -1 & -9 \\ 0 & 1 & -1 & 3 \\ 0 & -1 & 1 & -1 \end{array}\right) \boxed{-1} \sim$$

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 5 \\ 0 & 1 & 1 & 9 \\ 0 & 1 & -1 & 3 \\ 0 & -1 & 1 & -1 \end{array}\right) \boxed{-1}\boxed{1} \sim \left(\begin{array}{ccc|c} 1 & 1 & 2 & 5 \\ 0 & 1 & 1 & 9 \\ 0 & 0 & -2 & 3 \\ 0 & 0 & 0 & 8 \end{array}\right)$$

Den sista raden säger $0 = 8$ eller snarare $0x_1 + 0x_2 + 0x_3 = 8$ vilket inte gäller för några värden på x_1, x_2, x_3. Alltså saknar systemet lösningar. Samma sak händer alltid om det uppstår en nollrad i koefficientmatrisen men något nollskilt står i högerledet i samma rad. Då saknar systemet lösningar.

Vi tar ett exempel till

Exempel 1.1.3

Lös ekvationssystemet

$$\begin{cases} 2x_1 - 4x_2 - x_3 + 6x_4 = -2 \\ -x_1 + 2x_2 + 3x_3 - 3x_4 = 11 \\ x_1 - 2x_2 - 2x_3 + 3x_4 = -7 \end{cases}$$

Vi utför radoperationer på totalmatrisen för systemet:

$$\left(\begin{array}{cccc|c} 2 & -4 & -1 & 6 & -2 \\ -1 & 2 & 3 & -3 & 11 \\ 1 & -2 & -2 & 3 & -7 \end{array}\right) \sim$$

$$\left(\begin{array}{cccc|c} 1 & -2 & -2 & 3 & -7 \\ -1 & 2 & 3 & -3 & 11 \\ 2 & -4 & -1 & 6 & -2 \end{array}\right) \begin{array}{c} 1 \quad -2 \\ \\ \\ \end{array} \sim$$

$$\left(\begin{array}{cccc|c} 1 & -2 & -2 & 3 & -7 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 3 & 0 & 12 \end{array}\right) \begin{array}{c} \\ -3 \quad 2 \\ \\ \end{array} \sim$$

$$\left(\begin{array}{cccc|c} 1 & -2 & 0 & 3 & 1 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Den sista matrisen är radkanonisk och motsvarar systemet

$$\begin{cases} x_1 - 2x_2 \qquad + 3x_4 = 1 \\ \qquad\qquad x_3 \qquad = 4 \\ \qquad\qquad\qquad 0 = 0 \end{cases}$$

som vi kan skriva om till

$$\begin{cases} x_1 = 1 + 2x_2 - 3x_4 \\ x_3 = \qquad\qquad 4 \end{cases}$$

Notera att de obekanta som motsvarar kolonner med ledande ettor i den radkanoniska matrisen nu står i vänsterledet, medan de andra står i högerledet. Vi sätter dessa till parametrar: $x_2 = s$ och $x_4 = t$. Ekvationerna ger oss då värden på x_1 och x_3:

$$\begin{cases} x_1 = 1 + 2s - 3t \\ x_3 = 4 \end{cases}$$

Med andra ord får vi precis en lösning för varje värde på s och t. Vi kan alltså skrivna ner alla lösningar som

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 + 2s - 3t \\ s \\ 4 \\ t \end{pmatrix}, \quad s, t \in \mathbb{R}.$$

Om till exempel $s = 5$ och $t = 1$ får vi lösningen

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 8 \\ 5 \\ 4 \\ 1 \end{pmatrix}.$$

Lösningsmetoden som vi beskrivit i exemplen kallas Gauss-Jordan-elimination. I korthet bygger den på följande steg.

- Hitta totalmatrisen för ekvationssystemet.
- Använd radoperationer för att nå en radkanonisk matris.
- Gå tillbaka till motsvarande ekvationsystem och skriv ner lösningarna.

Vi har sett hur metoden fungerar i exempel, men för att bli helt övertygade om att metoden alltid fungerar går vi igenom hur stegen kan genomföras generellt. Låt T vara totalmatrisen till vårt system. I princip kan T vara vilken matris som helst. Hur kan vi vara säkra på att vi alltid kan nå en radkanonisk matris med radoperationer? Nedan visar vi hur det kan göras systematiskt.

Betrakta de ledande elementen in T, det vill säga de bland de nollskilda som står längst till vänster. Vi börjar med att se till att den första raden har ledande element längst till vänster. Om så inte redan är fallet byter vi plats på första raden och någon annan rad som har det. Vi gör nu om det ledande elementet i första raden till 1 genom att multiplicera med dess invers. Vi använder sedan första raden och dess ledande etta för att eliminera alla element under den ledande ettan. Nu har vi nått en matris som har följande form

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}$$

där $*$ betecknar element som vi inte vet något om. Observera att det ofta sker att den ledande ettan i första raden står direkt i första kolonnen, men i allmänhet kan ett godtyckligt antal nollor (inklusive inga alls) dyka upp innan. Vi utför nu samma procedur som vi gjorde på T på de rader som står under den första raden. Observera att de inledande nollorna inte kommer att påverkas av några av radoperationerna så vi får nu en matris på formen

$$\begin{pmatrix} 0 & \cdots & 0 & 1 & \cdots & * & * & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \end{pmatrix}$$

Genom att fortsätta på samma sätt och justera rad för rad, når vi till slut en matris som uppfyller villkor 1, 2 och 3 i definitionen för radkanonisk matris. För att uppnå villkor 4 går vi igenom raderna nedifrån och eliminerar alla element ovanför de ledande ettorna. Genom att börja nedifrån ser vi till att inte förstöra något av den form vi bygger upp som till slut ger oss en radkanonisk matris.

Låt oss till slut förklara hur vi finner lösningarna. Om vi får en nollrad i koeffcientmatrisen, men med ett nollskilt element i högerledet på samma rad så finns inga lösningar. Annars finns antingen en entydig lösning eller flera beroende på ett antal parametrar. Det finns i allmänhet många sätt att parametrisera lösningarna, men ett sätt som alltid fungerar är att göra som i exemplet: sätt de obekanta som motsvarar kolonner *utan* ledande ettor till parametrar. Den radkanoniska formen gör att övriga obekanta (alltså de som motsvarar kolonner *med* ledande ettor) kan lösas ut i termer av parametrarna. För specifika värden på parametrarna får vi därför precis en lösning till ekvationssystemet.

Vi illustrerar med ytterligare exempel.

Exempel 1.1.4

Lös ekvationssystemet

$$\begin{cases} x_1 + 2x_2 - x_3 + x_4 = 1 \\ 2x_1 + 6x_2 + x_3 - x_4 = 3 \\ x_1 + 4x_2 + 2x_3 - 2x_4 = 2 \\ 3x_1 + 8x_2 \phantom{{}+2x_3} + x_4 = 1 \end{cases}$$

Vi utför Gauss-Jordan elimination på totalmatrisen för systemet:

$$\left(\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 1 \\ 2 & 6 & 1 & -1 & 3 \\ 1 & 4 & 2 & -2 & 2 \\ 3 & 8 & 0 & 1 & 1 \end{array}\right) \boxed{-2}\boxed{-1}\boxed{-3} \sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 1 \\ 0 & 2 & 3 & -3 & 1 \\ 0 & 2 & 3 & -3 & 1 \\ 0 & 2 & 3 & -2 & -2 \end{array}\right) \boxed{-1} \sim$$

$$\sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 1 \\ 0 & 2 & 3 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -3 \end{array}\right) \left(\begin{array}{cccc|c} 1 & 2 & -1 & 1 & 1 \\ 0 & 2 & 3 & -3 & 1 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \boxed{-1}\boxed{-3} \sim$$

$$\sim \left(\begin{array}{cccc|c} 1 & 2 & -1 & 0 & 4 \\ 0 & 2 & 3 & 0 & -8 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \boxed{-1} \sim \left(\begin{array}{cccc|c} 1 & 0 & -4 & 0 & 12 \\ 0 & 2 & 3 & 0 & -8 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \boxed{\tfrac{1}{2}} \sim$$

$$\sim \left(\begin{array}{cccc|c} 1 & 0 & -4 & 0 & 12 \\ 0 & 1 & \frac{3}{2} & 0 & -4 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Notera att vi väntade till sista steget med att göra om 2 i andra raden till 1. Det spelar ingen större roll, men ibland är det praktiskt att vänta med att göra de ledande elementen till ettor för att undvika bråk.

Ekvationssystemet som motsvarar den sista matrisen har precis samma lösningar som det ursprungliga och är:

$$\begin{cases} x_1 \phantom{{}+x_2} - 4x_3 \phantom{{}+x_4} = 12 \\ \phantom{x_1+{}} x_2 + \frac{3}{2}x_3 \phantom{{}+x_4} = -4 \\ \phantom{x_1+x_2+\frac{3}{2}x_3+{}} x_4 = -3 \\ \phantom{x_1+x_2+\frac{3}{2}x_3+{}} 0 = 0 \end{cases} \Leftrightarrow \begin{cases} x_1 = 12 + 4x_3 \\ x_2 = -4 - \frac{3}{2}x_3 \\ x_4 = -3 \end{cases}$$

Vi har isolerat de obekanta som motsvarar ledande ettor i vänsterledet. Vi ser att för ett godtyckligt reellt $x_3 = t$ har vi precis en lösning för x_1, x_2 och x_4 så alla lösningar kan skrivas som:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 12 + 4t \\ -4 - \frac{3}{2}t \\ t \\ -3 \end{pmatrix}, \quad t \in \mathbb{R}.$$

Vi kan som alltid kontrollera vi lösningarna genom att sätta in i det ursprungliga systemet:

$$\begin{cases} (12 + 4t) + 2(-4 - \frac{3}{2}t) - t - 3 = 1 \\ 2(12 + 4t) + 6(-4 - \frac{3}{2}t) + t + 3 = 3 \\ (12 + 4t) + 4(-4 - \frac{3}{2}t) + 2t + 2 \cdot 3 = 2 \\ 3(12 + 4t) + 8(-4 - \frac{3}{2}t) - 3 = 1 \end{cases}$$

Det finns många sätt att använda radoperationer för att nå en radkanonisk matris. Men svaret blir alltid samma. Närmare bestämt gäller följande sats.

Sats 1.3

Varje matris är radekvivalent med *precis en* radkanonisk matris.

1.2 Matris- och vektorräkning

I föregående avsnitt använde vi matriser för att lösa linjära ekvationsystem, men matriser behövs i många andra sammanhang. I det här avsnittet påminner vi om grundläggande operationer som vi kan göra med matriser.

En $(m \times n)$-matris A består av en uppsättning tal a_{ij} där $1 \leq i \leq m$ och $1 \leq j \leq n$, som är ordnade rektangulärt på följande vis:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Notera att elementet som står på rad i och i kolonn j betecknas a_{ij}. På liknande sätt kallar vi ofta motsvarande element i en matris B för b_{ij} och i C för c_{ij} och så vidare.

Vi kan addera matriser av samma storlek och multiplicera matriser med skalärer. Detta sker elementvis. Till exempel har vi

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 2 & 14 & 4 \end{pmatrix} + \begin{pmatrix} 3 & 1 & 0 \\ -1 & -4 & 1 \\ 10 & -9 & 5 \end{pmatrix} = \begin{pmatrix} 1+3 & 2+1 & 3+0 \\ 0-1 & -1-4 & -7+1 \\ 2+10 & 14-9 & 4+5 \end{pmatrix} = \begin{pmatrix} 4 & 3 & 3 \\ -1 & -5 & -6 \\ 12 & 5 & 9 \end{pmatrix}.$$

och

$$4\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -7 \\ 2 & 14 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 8 & 12 \\ 0 & -4 & -28 \\ 8 & 56 & 16 \end{pmatrix}$$

I allmänhet gäller

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix}.$$

och

$$\lambda\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix}$$

Matrismultiplikation fungerar lite annorlunda. Om A är en $(m \times k)$-matris och B är en $(k \times n)$-matris kan vi multiplicera A med B och få en $(m \times n)$-matris AB. För att beräkna elementetet på plats ij i matrisen AB matchar vi de k elementen som står på rad i i A med element de k elementen i kolonn j i B, multiplicerar dessa parvis, och adderar resultaten.

Om till exempel

$$A = \begin{pmatrix} 3 & 4 \\ 2 & 0 \\ 1 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix}$$

så är

$$AB = \begin{pmatrix} 3 & 4 \\ 2 & 0 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \end{pmatrix} =$$

$$\begin{pmatrix} 3\cdot 1+4\cdot 0 & 3\cdot 2+4\cdot(-1) & 3\cdot 3+4\cdot(-2) \\ 2\cdot 1+0\cdot 0 & 2\cdot 2+0\cdot(-1) & 2\cdot 3+0\cdot(-2) \\ 1\cdot 1+5\cdot 0 & 1\cdot 2+5\cdot(-1) & 1\cdot 3+5\cdot(-2) \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 4 & 6 \\ 1 & -3 & -7 \end{pmatrix},$$

Mer allmänt gäller att om

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mk} \end{pmatrix} \text{ och } B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots \\ b_{k1} & \cdots & b_{kn} \end{pmatrix}$$

$$\text{så är } AB = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix}$$

$$\begin{aligned} \text{där } c_{ij} &= a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj} \\ &= \sum_{k=1}^{r} a_{ik}b_{kj} \end{aligned}$$

Med hjälp av matrismultiplikation kan vi skriva om linjära ekvationssystem som matrisekvationer. Till exempel är

$$\begin{cases} 2x + y = 4 \\ x - y = -1 \end{cases}$$

ekvivalent med

$$\begin{pmatrix} 2x+y \\ x-y \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 2 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}.$$

Mer allmänt gäller att om

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

så motsvarar $AX = B$ ekvationssystemet

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \quad \vdots \qquad\quad \vdots \qquad \ddots \qquad\quad \vdots \qquad\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases}$$

Med denna tolkning av linjära ekvationssystem är naturligt att skriva lösningarna som en kolonnmatris X på det sätt vi gjort i föregående avsnitt.

Geometriskt kan vi tolka X som en vektor i $\mathbb{R}^n$, vilket går att visualisera för $n = 2$ och $n = 3$ och kan förstås mer abstrakt för $n > 3$ som en vektor i n dimensioner.

Samma regler för addition av matriser och multiplikation av matris med skalär ger upphov till motsvarande operationer för vektorer i $\mathbb{R}^n$.

$$\text{Låt } \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \text{ och } \vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \text{ vara vektorer i } \mathbb{R}^n.$$

$$\text{Då är } \vec{v} + \vec{w} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}.$$

För $\lambda \in \mathbb{R}$ gäller att

$$\lambda\vec{v} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

1.3 Determinanter

I det här avsnittet påminner vi om determinanten $\det A$ av en $(n \times n)$-matris A. Vi börjar med fallet $n = 2$. Då är

$$\det \begin{pmatrix} a & b \\ c & b \end{pmatrix} = ad - bc$$

Observera att första termen ad är produkten av diagonalelementen:

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

den andra termen $-bc$ är minus produkten av de andra två elementen:

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

För att definiera determinanter av större matriser inför vi lite terminologi.

- Ett *mönster* i en $(n \times n)$-matris A är ett val av n element från A, bestående av precis ett element från varje rad och kolonn.
- *Produkten* av ett mönster i A är produkten av motsvarande element i A.
- Vi säger att två element i ett mönster bildar ett *omvänt par* om det ena står ovanför och till höger om det andra.
- Vi säger att ett mönster är *jämnt* om det har ett jämnt antal omvända par och *udda* om det har ett udda antal omvända par.

Definition 1.4

Determinanten $\det A$ av en kvadratisk matris A är summan av alla produkter av jämna mönster i A minus summan av alla produkter av udda mönster i A.

För att beteckna determinanten av en matris A byter vi ut parenteserna mot raka sträck. Till exempel skriver vi

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Notera att det i detta fall $n = 2$ bara finns ett jämnt mönster med produkt ad och ett udda med produkt bc.

Vi ser vad definitionen ger oss för en (3×3)-matris:

$$A = \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}$$

Totalt har vi 6 mönster som vi markerar med omvända par nedan:

$$\begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}$$

$$\begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}$$

$$\begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}$$

Antalet omvända par är som vi ser 0, 1, 1, 2, 2 respektive 3. Alltså är

$$
\begin{aligned}
\det A = \begin{vmatrix} x & y & z \\ a & b & c \\ d & e & f \end{vmatrix} &= xbf - xce - yaf + ycd + zae - zbd \\
&= x(bf - ce) - y(af - cd) + z(ae - bd) \\
&= x\begin{vmatrix} b & c \\ e & f \end{vmatrix} - y\begin{vmatrix} a & c \\ d & f \end{vmatrix} + z\begin{vmatrix} a & b \\ d & e \end{vmatrix}.
\end{aligned}
$$

Sats 1.5

$$
\begin{vmatrix} x & y & z \\ a & b & c \\ d & e & f \end{vmatrix} = x\begin{vmatrix} b & c \\ e & f \end{vmatrix} - y\begin{vmatrix} a & c \\ d & f \end{vmatrix} + z\begin{vmatrix} a & b \\ d & e \end{vmatrix}
$$

Observera att de tre (2×2)-matriserna vars determinanter dyker upp i högerledet fås genom att stryka första raden i

$$
\begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}
$$

och därefter stryka kolonn 1, 2 respektive 3.

Exempel 1.3.1

Beräkna determinanten av

$$
A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 0 & 4 & 1 \end{pmatrix}.
$$

Vi använder Sats 1.5 och finner att

$$
\begin{aligned}
\begin{vmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 0 & 4 & 1 \end{vmatrix} &= 1\begin{vmatrix} -1 & 3 \\ 4 & 1 \end{vmatrix} - 2\begin{vmatrix} 2 & 3 \\ 0 & 1 \end{vmatrix} + 0\begin{vmatrix} 2 & -1 \\ 0 & 4 \end{vmatrix} \\
&= (-1 - 12) - 2(2 - 0) + 0(8 - 0) = -13 - 4 + 0 = -17
\end{aligned}
$$

Ett annat praktiskt sätt att räkna ut determinanten av en (3×3)-matris är att komma ihåg de jämna mönstren:

$$\begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}$$

för sig och de och de udda mönstren:

$$\begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix} \quad \begin{pmatrix} x & y & z \\ a & b & c \\ d & e & f \end{pmatrix}$$

för sig. Då får vi

$$\begin{vmatrix} x & y & z \\ a & b & c \\ d & e & f \end{vmatrix} = xbf + ycd + zae - zbd - xce - yaf.$$

Exempel 1.3.2

Beräkna determinanten av

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 0 & 4 & 1 \end{pmatrix}.$$

Om vi nu räknar jämna och udda mönster var för sig får vi

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 0 \\ 2 & -1 & 3 \\ 0 & 4 & 1 \end{vmatrix} =& 1 \cdot (-1) \cdot 1 + 2 \cdot 3 \cdot 0 + 0 \cdot 2 \cdot 4 \\ & - 0 \cdot (-1) \cdot 0 - 4 \cdot 3 \cdot 1 - 1 \cdot 2 \cdot 2 \\ =& -1 + 0 + 0 - 0 - 12 - 4 = -17. \end{aligned}$$

Vid det här laget börjar det kanske märkas att förekomsten av nollor i en matris förenklar uträkningen av dess determinant. Det beror på att alla mönster som väljer ut en nolla har produkt 0. Till exempel kommer en matris med en nollrad alltid att ha determinant 0 eftersom varje mönster innehåller en av nollorna i den raden. Vi ska nu se på ett annat exempel då determinanten enkelt kan räknas ut.

Exempel 1.3.3

Beräkna determinanten av

$$A = \begin{pmatrix} 1 & 10 & 18 & 23 \\ 0 & 2 & -2 & 7 \\ 0 & 0 & 3 & 11 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Vi reder ut vilka mönster som har nollskild produkt. Från första kolonnen måste vi välja elementet 1 eftersom övriga är 0. Om vi nu fortsätter med andra kolonnen ser vi att vi måste välja 2 eftersom 10 står i den första raden som redan är upptagen och övriga element är 0. På liknande sätt ser vi att vi i tredje kolonnen måste välja 3 och i fjärde 4. Alltså finns bara ett mönster vars produkt inte är noll nämligen:

$$\begin{pmatrix} 1 & 10 & 18 & 23 \\ 0 & 2 & -2 & 7 \\ 0 & 0 & 3 & 11 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Detta diagonalmönster har inga omvända par och är därför jämnt. Alltså är

$$\begin{vmatrix} 1 & 10 & 18 & 23 \\ 0 & 2 & -2 & 7 \\ 0 & 0 & 3 & 11 \\ 0 & 0 & 0 & 4 \end{vmatrix} = 1 \cdot 2 \cdot 3 \cdot 4 = 24.$$

En matris där alla element under diagonalen är 0 (som i exemplet ovan) kallas för en övertriangulär matris. För en sådan kan vi uppenbarligen räkna ut determinanten genom att multiplicera diagonalelementen.

Sats 1.6

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & \cdots & a_{2n} \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & a_{nn} \end{vmatrix} = a_{11}a_{22}\cdots a_{nn}$$

1.3.1 Beräkning av determinanter med radoperationer

Låt A vara en $(n \times n)$-matris och B vara den matris som uppstår när vi utför någon viss radoperation på A. I det här avsnittet ska vi undersöka vad det finns för samband mellan determinanterna $\det A$ och $\det B$. Vi börjar med att studera fallet $n = 2$ och skriver

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

Då är som vi har sett $\det A = ad - bc$. Om vi multiplicerar första raden i A med en konstant k får vi matrisen

$$B = \begin{pmatrix} ka & kb \\ c & d \end{pmatrix}$$

vars determinant är $\det B = kad - kbc = k(ad - bc) = k \det A$. Om B istället matrisen som uppstår då vi multiplicerar andra raden i A med k så gäller på liknande sätt att $\det B = akd - bkc = k(ad - bc) = k \det A$. Vi har alltså

$$\begin{vmatrix} ka & kb \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ kc & kd \end{vmatrix} = k \begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

Motsvarande gäller för determinanter av större matriser: om vi multiplicerar någon viss rad i en matris med en konstant så förändras determinanten genom multiplikation med samma konstant. Det beror på att varje mönster innehåller precis ett element i raden, så motsvarande produkt multipliceras med konstanten precis en gång.

Låt nu istället B vara matrisen som uppstår då vi byter plats på de två raderna i A:

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

Då är $\det B = cb - da = -(ad - bc) = -\det A$. För determinanter av större matriser gäller på liknande sätt att om två rader byter plats så byter determinanten tecken.

Låt till sist B vara matrisen som uppstår då vi lägger till k gånger första raden i A till den andra. Det vill säga

$$B = \begin{pmatrix} a & b \\ c + ka & d + kb \end{pmatrix}.$$

Då är $\det B = a(d + kb) - b(c + ka) = ad + akb - bc - bka = ad - bc = \det A$. För determinanter av större matriser gäller mer allmänt att om vi lägger till en konstant gånger en rad till en annan rad, så påverkas inte determinantens värde.

Vi sammanfattar hur radoperationer påverkar determinantens värde i följande sats

Sats 1.7

Låt A och B vara $(n \times n)$-matriser. Då gäller att

1) om B fås från A genom att multiplicera en rad med en konstant k så är
$$\det B = k \det A,$$

2) om B fås från A genom att byta plats på två rader så är
$$\det B = -\det A,$$

3) om B fås från A genom lägga till en konstant gånger en rad till en annan rad så är
$$\det B = \det A.$$

Vi kan nu beräkna determinanter genom att utföra radoperationer och hålla reda på hur determinantens värde förändras.

Exempel 1.3.4

Beräkna determinanten av
$$A = \begin{pmatrix} 1 & -1 & 0 & 1 \\ -2 & 2 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 5 & 7 \end{pmatrix}$$

Vi använder radoperationer och håller koll på hur determinanten förändras

$$
\begin{aligned}
\begin{vmatrix} 1 & -1 & 0 & 1 \\ -2 & 2 & 2 & 3 \\ 0 & 1 & 2 & 3 \\ 1 & 1 & 6 & 7 \end{vmatrix} \begin{matrix} 2 \ \ -1 \\ \\ \\ \\ \end{matrix}
&= \begin{vmatrix} 1 & -1 & 0 & 1 \\ 0 & 0 & 2 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 2 & 6 & 6 \end{vmatrix}
= - \begin{vmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 5 \\ 0 & 2 & 6 & 6 \end{vmatrix} \begin{matrix} \\ -2 \\ \\ \\ \end{matrix} \\
&= - \begin{vmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 2 & 0 \end{vmatrix}
= \begin{vmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 2 & 5 \end{vmatrix} \begin{matrix} \\ \\ -1 \\ \\ \end{matrix} \\
&= \begin{vmatrix} 1 & -1 & 0 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 2 & 5 \\ 0 & 0 & 0 & 5 \end{vmatrix} = 1 \cdot 1 \cdot 2 \cdot 5 = 10.
\end{aligned}
$$

Den sista likheten gäller eftersom determinanten av en övertriangulär matris är produkten av diagonalelementen. Från beräkningen kan vi se att matrisen A har rang 4. Alltså är A inverterbar och mycket riktigt är $\det A = 10 \neq 0$. Vi har nu fått en ledtråd till beviset av följande sats.

Vi avslutar med två viktiga satser om determinanter.

Sats 1.8

En kvadratisk matris A är inverterbar om och endast om $\det A \neq 0$.

Sats 1.9

Låt A och B vara $(n \times n)$-matriser. Då är $\det(AB) = \det A \det B$.

1.3.2 Utveckling efter rader och kolonner

I Sats 1.5 såg vi att determinanten av en (3×3)-matris A kan beräknas utifrån de tre deteminanter som dyker upp om vi stryker första raden och var och en av de tre kolonnerna i A. I detta stycke presenterar vi en generalisering av denna metod.

Låt A vara en godtycklig $(n \times n)$-matris. Låt M_{ij} beteckna determinanten av den matris som vi får om vi stryker rad i och kolonn j i A med. En sådan determinant kallas för en minor till A.

Exempel 1.3.5

Låt

$$A = \begin{pmatrix} 0 & 1 & 0 & -4 \\ 1 & 2 & -1 & 7 \\ 1 & 0 & 0 & -1 \\ 2 & 5 & 6 & 1 \end{pmatrix}$$

Minoren som motsvarar rad 2 och kolonn 3 är då

$$M_{23} = \begin{vmatrix} 0 & 1 & -4 \\ 1 & 0 & -1 \\ 2 & 5 & 1 \end{vmatrix} = 0 - 2 - 20 - 0 - 0 - 1 = -23.$$

Om

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

så är enligt Sats 1.5

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

eller mer kompakt

$$\det A = a_{11}M_{11} - a_{12}M_{12} + a_{13}M_{13}.$$

Detta sätt att räkna ut $\det A$ kallas för utveckling längs rad 1. Faktum är att man på liknade sätt kan utveckla längs någon av de andra raderna eller till och med längs någon kolonn. Det går även att hantera större determinanter på likande sätt. Mer precist gäller följande sats.

Sats 1.10

Låt A vara en $(n \times n)$-matris. Kalla elementet i A på rad i och kolonn j för a_{ij} och motsvarande minor för M_{ij}. Då gäller följande formler:

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}, \quad \text{utveckling efter rad } i.$$

$$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} M_{ij}, \quad \text{utveckling efter kolonn } j.$$

2 Vektorrum

I tidigare kurser har du stött på vektorer i planet och i rummet (det vill säga $\mathbb{R}^2$ och $\mathbb{R}^3$). Dessa kan generaliseras till vektorer i $\mathbb{R}^n$, som vi skriver på följande sätt:

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

Som vi såg i föregående kapitel finns det två grundläggande operationer som vi kan utföra på vektorer i $\mathbb{R}^n$. Addition:

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix}$$

och skalning (som även kallas multiplikation med skalär $\lambda \in \mathbb{R}$):

$$\lambda\vec{v} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Det finns många andra typer av objekt som går att addera och skala om på liknande sätt. Till exempel kan vi ta två polynom $1 + x$ och $3 - x^2$ och addera dem på vanligt vis

$$(1 + x) + (3 - x^2) = 4 + x - x^2$$

eller skala om polynomet $4 - x + x^2$ med en faktor 3:

$$3(4 - x + x^2) = 12 - 3x + 3x^2.$$

Mer generellt kan vi addera och skala om funktioner. Om $f(x) = \sin x$ och $g(x) = x$ så är

$$(f + g)(x) = f(x) + g(x) = \sin x + x$$

och

$$(5f)(x) = 5f(x) = 5\sin x$$

Matriser av någon viss storlek går också att addera och skala. Till exempel är

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} \text{ och } 5\begin{pmatrix} 1 & -1 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 5 & -5 \\ 0 & 10 \end{pmatrix}.$$

Vektorrum är ett gemensamt begrepp för dessa baserat på grundläggande räkneregler.

2.1 Definition och exempel

Definition 2.1

En mängd $\mathcal{V}$ kallas ett *reellt vektorrum* om den är utrustad med två operationer

- *addition*, som till $\vec{v}, \vec{w} \in \mathcal{V}$ tillordnar en vektor $\vec{v} + \vec{w} \in \mathcal{V}$,
- *skalning* eller *multipliation med skalär* som till $\lambda \in \mathbb{R}$, $\vec{v} \in \mathcal{V}$ tillordnar en vektor $\lambda\vec{v} \in \mathcal{V}$.

som uppfyller följande villkor:

1. $\vec{v} + \vec{w} = \vec{w} + \vec{v}$,
2. $\vec{u} + (\vec{v} + \vec{w}) = (\vec{u} + \vec{v}) + \vec{w}$,
3. det finns en *nollvektor* $\vec{0} \in \mathcal{V}$ som uppfyller $\vec{v} + \vec{0} = \vec{v}$,
4. till varje $\vec{v} \in \mathcal{V}$ finns en *additiv invers* $-\vec{v} \in \mathcal{V}$ som uppfyller $\vec{v} + (-\vec{v}) = \vec{0}$,
5. $1\vec{v} = \vec{v}$,
6. $\lambda(\mu\vec{v}) = (\lambda\mu)\vec{v}$,
7. $(\lambda + \mu)\vec{v} = \lambda\vec{v} + \mu\vec{v}$,
8. $\lambda(\vec{u} + \vec{v}) = \lambda\vec{u} + \lambda\vec{v}$,

för alla $\lambda, \mu \in \mathbb{R}$ och $\vec{u}, \vec{v}, \vec{w} \in \mathcal{V}$. De åtta reglerna ovan kallas vektorrumsaxiomen. De kan ses som krav på $\mathcal{V}$ och operationerna som ska gälla för att $\mathcal{V}$ ska förtjäna beteckningen vektorrum.

Vi kallar elementen i $\mathcal{V}$ för vektorer även om $\mathcal{V} \neq \mathbb{R}^n$. Operationerna kan vara definierade hur som helst så länge vektorrumsaxiomen är uppfyllda. Det betyder också att vi inte vet något om nollvektorn $\vec{0}$ än att $\vec{v} + \vec{0} = \vec{v}$ gäller. På samma sätt är det enda vi vet om den additiva inversen $-\vec{v}$ att $\vec{v} + (-\vec{v}) = \vec{0}$ gäller.

Vi går nu igenom några viktiga exempel på vektorrum.

Exempel 2.1.1

Mängden $\mathbb{R}^n$ med $+$ och $\cdot$ definierade som tidigare är ett vektorrum. Vektorrumsaxiomen gäller i detta exempel på grund av att de gäller för reella tal och kan verifieras i $\mathbb{R}^n$ koordinatvis. Till exempel gäller axiom 1 på grund av

$$\vec{v} + \vec{w} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{pmatrix} = \begin{pmatrix} w_1 + v_1 \\ w_2 + v_2 \\ \vdots \\ w_n + v_n \end{pmatrix} = \vec{w} + \vec{v}.$$

Axiom 2 kan visas på liknande sätt. För att verifiera axiom 3 behöver vi påvisa ett nollelement $\vec{0} \in \mathbb{R}^n$. Vi ser att

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

ger

$$\vec{v} + \vec{0} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} v_1 + 0 \\ v_2 + 0 \\ \vdots \\ v_n + 0 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \vec{v}.$$

För att visa axiom 4, behöver vi ange $-\vec{v}$. Sätter vi

$$-\vec{v} = \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix}$$

får vi att

$$\vec{v} + (-\vec{v}) = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} + \begin{pmatrix} -v_1 \\ -v_2 \\ \vdots \\ -v_n \end{pmatrix} = \begin{pmatrix} v_1 - v_1 \\ v_2 - v_2 \\ \vdots \\ v_n - v_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \vec{0}.$$

Vi lämnar verifikationen av axiom 5, 6, 7 och 8 som övning.

Vektorrummet $\mathbb{R}^n$ är i någon mening vårt huvudexempel, men det finns som sagt många andra.

Exempel 2.1.2

Mängden $\mathcal{M}_{m\times n}$ av $(m \times n)$-matriser med $+$ och $\cdot$ definierade enligt vad som gäller för matrisräkning är ett vektorrum. Vektorrumsaxiomen är faktiskt exempel på regler som gäller för matrisräkning. Nollelementet i $\mathcal{M}_{m\times n}$ är matrisen vars alla element är 0. Om till exempel $m = n = 2$ så är nollelementet

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

Som vi nämnde kan vi addera och skala polynom. Det ger upphov till ett antal olika vektorrum som vi går igenom i nästa exempel

Exempel 2.1.3

Låt $\mathcal{P}$ vara mängden av alla reella polynom, det vill säga

$$\mathcal{P} = \{a_0 + a_1 x + \cdots a_N x^N \mid N \geq 0, a_i \in \mathbb{R}\}.$$

Mängden $\mathcal{P}$ med addition och skalning definierade på vanligt sätt är ett vektorrum. Notera att $N \geq 0$ kan vara hur stort som helst. Det finns alltså ingen begränsning på hur många tal vi kan behöva precisera för att ange ett element i $\mathcal{P}$. Det gör att $\mathcal{P}$ är ganska annorlunda jämfört med säg $\mathbb{R}^{10}$ där varje vektor kan anges genom att precisera 10 koordinater.

För att få ett vektorrum som är mer likt $\mathbb{R}^n$ kan vi inskränka oss till polynom av begränsad grad. Låt $n \geq 0$ vara något heltal och $\mathcal{P}_n$ vara mängden av alla polynom av grad högst n, det vill säga

$$\mathcal{P}_n = \{a_0 + a_1 x + \cdots a_n x^n \mid a_i \in \mathbb{R}\}.$$

Då är $\mathcal{P}_n$ ett vektorrum med samma operationer som $\mathcal{P}$. Till exempel är

$$\mathcal{P}_2 = \{a_0 + a_1 x + a_2 x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\},$$

ett vektorrum. Notera att elementen i $\mathcal{P}_2$ kan anges genom att precisera tre tal a_0, a_1, a_2. På så sätt är $\mathcal{P}_2$ ungefär som $\mathbb{R}^3$.

I exemplet ovan tar vi upp polynom som vi kan förstå som funktioner från $\mathbb{R}$ till $\mathbb{R}$. Det finns många likande exempel på vektrrum vars element är funktioner från någon mängd till $\mathbb{R}$.

Exempel 2.1.4

Låt M vara en mängd och $\mathcal{F}(M)$ vara mängden av funktioner från M till $\mathbb{R}$, det vill säga

$$\mathcal{F}(M) = \{f \colon M \to \mathbb{R}\}.$$

Vi kan addera funktioner och skala om funktioner genom att addera och skala varje funktionsvärde. Närmare bestämt definierar vi för $f, g \in \mathcal{F}(M)$ och $\lambda \in \mathbb{R}$ nya funktioner $f+g \in \mathcal{F}(M)$ och $\lambda f \in \mathcal{F}(M)$ genom

$$(f+g)(x) = f(x) + g(x) \quad \text{och } (\lambda f)(x) = \lambda f(x).s$$

Ofta är vi intresserade av fallet då $M \subseteq \mathbb{R}$, till exempel $M = [0, 1]$. Då är det ibland relevant att bara studera kontinuerliga funktioner. Eftersom summan av två kontinuerliga funktioner är kontinuerlig och en skalning av en kontinuerlig funktion är kontinuerlig får vi ett vektorrum:

$$\mathcal{C}(M) = \{f \in \mathcal{F}(M) \mid f: \text{ kontinuerlig}\}.$$

Ett annat intressant fall är då M är en ändlig mängd, till exempel $M = \{1, 2, 3\}$. I så fall kan vi precisera elementen i $\mathcal{F}(M)$ genom helt enkelt ange funktionsvärdena $f(x)$ för alla $x \in M$. Om igen $M = \{1, 2, 3\}$ så bestäms f av tre tal $f(1)$, $f(2)$ och $f(3)$. Med andra ord är $\mathcal{F}(\{1, 2, 3\})$ ungefär som $\mathbb{R}^3$.

2.2 Viktiga egenskaper

Varje vektorrum uppfyller ytterligare räkneregler som följer av axiomen. I det här avsnittet ska vi gå igenom några av de viktigaste. Men det finns många fler. Genom hela kursen kommer vi utveckla teorin för vektorrum och på så sätt få fler och fler verktyg att använda. Det fina med att visa allt från de åtta vektorrumsaxiomen är att vi kan vara säkra på att räknereglerna och andra resultat som vi visar kommer att gälla för alla vektorrum, inklusive de som vi stötte på i föregående avsnitt.

Sats 2.2

Låt $\mathcal{V}$ vara ett vektorrum. Då gäller följande:

a) Nollvektorn är unik. Det finns alltså bara ett element $\vec{0} \in \mathcal{V}$ som uppfyller $\vec{v} + \vec{0} = \vec{v}$ för alla $\vec{v} \in \mathcal{V}$.

b) Om $\vec{v} + \vec{u} = \vec{w} + \vec{u}$ så är $\vec{v} = \vec{w}$.

c) Den additiva inversen till $\vec{v}$ är unik. Det finns alltså bara ett element $-\vec{v} \in \mathcal{V}$ som uppfyller $\vec{v} + (-\vec{v}) = \vec{0}$.

d) $0\vec{v} = \vec{0}$ gäller för alla $\vec{v} \in \mathcal{V}$.

e) $\lambda\vec{0} = \vec{0}$ gäller för alla $\lambda \in \mathcal{V}$.

f) $(-1)\vec{v} = -\vec{v}$.

Bevis: a) Eftersom $\mathcal{V}$ är ett vektorrum finns $\vec{0} \in \mathcal{V}$ så att $\vec{v} + \vec{0} = \vec{v}$. Antag nu $\vec{n} \in \mathcal{V}$ är någon vektor som uppfyller samma sak, $\vec{v} + \vec{n} = \vec{v}$. Vi behöver då visa att $\vec{n} = \vec{0}$. Vi gör det genom att räkna ut $\vec{0} + \vec{n}$ på två sätt. Å ena sidan ger $\vec{v} + \vec{n} = \vec{v}$ att

$$\vec{0} + \vec{n} = \vec{0}$$

Å andra sidan ger $\vec{v} + \vec{0} = \vec{v}$ att

$$\vec{0} + \vec{n} = \vec{n} + \vec{0} = \vec{n}$$

Alltså är $\vec{n} = \vec{0}$.

b) Vi lägger till $-\vec{u}$ i båda led av ekvationen $\vec{v} + \vec{u} = \vec{w} + \vec{u}$:

$$(\vec{v} + \vec{u}) + (-\vec{u}) = (\vec{w} + \vec{u}) + (-\vec{u}).$$

Vänsterledet är

$$(\vec{v}+\vec{u})+(-\vec{u})=\vec{v}+(\vec{u}+(-\vec{u}))=\vec{v}+\vec{0}=\vec{v}.$$

På liknande sätt är högerledet

$$(\vec{w}+\vec{u})+(-\vec{u})=\vec{w}+(\vec{u}+(-\vec{u}))=\vec{w}+\vec{0}=\vec{w}.$$

Alltså är $\vec{v}=\vec{w}$.

c) Denna del liknar a). Eftersom $\mathcal{V}$ är ett vektorrum finns till $\vec{v}$ en vektor $-\vec{v}$ så att $\vec{v}+(-\vec{v})=\vec{0}$. Antag nu $\vec{a}\in\mathcal{V}$ också uppfyller $\vec{v}+\vec{a}=\vec{0}$. Vi behöver visa att $\vec{a}=-\vec{v}$. Vi får att $\vec{v}+(-\vec{v})=\vec{0}=\vec{v}+\vec{a}$ och. Från b) (och axiom 1) följer nu $-\vec{v}=\vec{a}$.

d) Vi skriver

$$0\vec{v}=(0+0)\vec{v}=0\vec{v}+0\vec{v}.$$

Eftersom vänsterledet är $0\vec{v}=\vec{0}+0\vec{v}$ får vi från b) att $0\vec{v}=\vec{0}$.

e) Det här likar d). Vi skriver

$$\lambda\vec{0}=\lambda(\vec{0}+\vec{0})=\lambda\vec{0}+\lambda\vec{0}$$

Som i d) följer nu $\lambda\vec{0}=\vec{0}$.

f) Vi kollar att $(-1)\vec{v}$ uppfyller definitionen av $-\vec{v}$:

$$\vec{v}+(-1)\vec{v}=1\vec{v}+(-1)\vec{v}=(1-1)\vec{v}=0\vec{v}=\vec{0}.$$

Som vi ser i satsen fungerar räkning med vektorer i ett godtyckligt vektorrum mycket likt hur det fungerar i $\mathbb{R}^n$. Vi tar ett exempel för att illustrera detta.

Exempel 2.2.1

Låt $\mathcal{V}$ vara ett vektorrum och $\vec{v}\in\mathcal{V}$. Hitta den vektor $\vec{x}\in\mathcal{V}$ som uppfyller

$$\vec{v}+2\vec{x}=\vec{0}$$

Vi använder räknereglerna stegvis för att lösa ut $\vec{x}$.

$$\begin{aligned}\vec{v} + 2\vec{x} = \vec{0} &\Rightarrow 2\vec{x} + \vec{v} = \vec{0} \\ &\Rightarrow (2\vec{x} + \vec{v}) - \vec{v} = \vec{0} - \vec{v} \\ &\Rightarrow 2\vec{x} + (\vec{v} - \vec{v}) = -\vec{v} \\ &\Rightarrow 2\vec{x} + \vec{0} = -\vec{v} \\ &\Rightarrow 2\vec{x} = -\vec{v} \\ &\Rightarrow \frac{1}{2}(2\vec{x}) = \frac{1}{2}(-\vec{v}) = \frac{1}{2}((-1)\vec{v}) = (-\frac{1}{2})\vec{v}\end{aligned}$$

Å andra sidan är $\frac{1}{2}(2\vec{x}) = \frac{2}{2}\vec{x} = 1\vec{x} = \vec{x}$, så

$$\vec{x} = (-\frac{1}{2})\vec{v}$$

är den enda möjligheten.

En viktig skillnad mellan räkning med reella tal och räkning i ett vektorrum är att vi inte kan dela med vektorer. Men det finns några räkneregler som är nästan lika bra som att dela. Vi går igenom dem i nästa sats.

Sats 2.3

Låt $\mathcal{V}$ vara ett vektorrum och $\vec{v} \in \mathcal{V}$ så att $\vec{v} \neq \vec{0}$. Då gäller följande:

a) Om $\lambda\vec{v} = \vec{0}$ så $\lambda = 0$

b) Om $\lambda\vec{v} = \mu\vec{v}$ så $\lambda = \mu$.

Bevis: a) Antag att $\lambda\vec{v} = \vec{0}$. Om $\lambda \neq 0$ så får vi

$$\vec{v} = \frac{1}{\lambda}(\lambda\vec{v}) = \frac{1}{\lambda}\vec{0} = \vec{0},$$

vilket motsäger vårt antagande $\vec{v} \neq \vec{0}$. Alltså måste $\lambda = 0$.

b) Från a) får vi

$$\begin{aligned}\lambda\vec{v} = \mu\vec{v} &\Rightarrow \lambda\vec{v} - \mu\vec{v} = \vec{0} \\ &\Rightarrow (\lambda - \mu)\vec{v} = \vec{0} \\ &\Rightarrow \lambda - \mu = 0 \\ &\Rightarrow \lambda = \mu\end{aligned}$$

2.3 Delrum

Vi har sett flera exempel på när ett vektorrum är en delmängd av ett annat vektorrum. Till exempel är $\mathcal{P}_2 \subseteq \mathcal{P}_3$:

$$\{a_0+a_1x+a_2x^2 \mid a_0, a_1, a_2 \in \mathbb{R}\} \subseteq \{a_0+a_1x+a_2x^2+a_3x^3 \mid a_0, a_1, a_2, a_3 \in \mathbb{R}\}.$$

Observera att $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3$ uppfyller $p(x) \in \mathcal{P}_2$ om och endast om $a_3 = 0$.

Ett annat exempel är $\mathcal{C}(\mathbb{R}) \subseteq \mathcal{F}(\mathbb{R})$. Här gäller att $f \in \mathcal{C}(\mathbb{R})$ om och endast om f är en kontinuerlig funktion.

Som ett tredje exempel kan vi låta

$$\mathcal{U} = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \middle| \, x, y \in \mathbb{R} \right\} \subseteq \mathbb{R}^3.$$

Med andra ord är $\mathcal{U}$ xy-planet som en delmängd av xyz-rummet. Om vi adderar och skalar vektorer i $\mathcal{U}$ på samma sätt som i $\mathbb{R}^3$ så får vi ett vektorrum. Vi kan förstå $\mathcal{U}$ som $\mathbb{R}^2$ där vi lagt till en tredje koordinat som är 0 för varje vektor. Nyckeln till att detta fungerar är att om vi adderar eller skalar om vektorer i $\mathcal{U}$ så får vi återigen vektorer i $\mathcal{U}$.

De fall vi gick igenom ovan är alla exempel på delrum. Mer allmänt definieras de såhär:

Definition 2.4

En icketom delmängd $\mathcal{U}$ av ett vektorrum $\mathcal{V}$ kallas ett *delrum* om

1. för $\vec{u}, \vec{v} \in \mathcal{U}$ gäller $\vec{u} + \vec{v} \in \mathcal{U}$,
2. för $\lambda \in \mathbb{R}$ och $\vec{u} \in \mathcal{U}$ gäller $\lambda\vec{u} \in \mathcal{U}$.

Från de två villkoren i definitionen följer ytterligare egenskaper. Faktiskt är $\mathcal{U}$ självt ett vektorrum.

Sats 2.5

Om $\mathcal{V}$ är ett vektorrum och $\mathcal{U} \subseteq \mathcal{V}$ är ett delrum så gäller följande:

a) Om $\vec{u} \in \mathcal{U}$ så $-\vec{u} \in \mathcal{U}$.

b) $\vec{0} \in \mathcal{U}$.

c) $\mathcal{U}$ är ett vektorrum med samma operationer som i $\mathcal{V}$.

Bevis: a) Vi har $-\vec{u} = (-1)\vec{u} \in \mathcal{U}$ enligt villkor 2.
b) Låt $\vec{u} \in \mathcal{U}$. Från a) får vi $-\vec{u} \in \mathcal{U}$. Nu följer $\vec{0} = \vec{u} + (-\vec{u}) \in \mathcal{U}$ enligt villkor 1.
c) Axiom 3 och 4 är precis b) och a). Övriga axiom följer direkt för $\mathcal{U}$ på grund av att de gäller för $\mathcal{V}$.

I definitionen av delrum krävs att $\mathcal{U}$ är en icketom mängd, alltså att det finns någon vektor $\vec{u} \in \mathcal{U}$. Från satsen ser vi att vi alltid har att $\vec{0} \in \mathcal{U}$ om $\mathcal{U}$ är ett delrum. Så för att påvisa att det finns åtminstone en vektor i $\mathcal{U}$ kan vi alltid ta nollvektorn. I praktiken är det en bra strategi för att kolla om en delmängd $\mathcal{U}$ är ett underum. Kolla först om $\vec{0} \in \mathcal{U}$. Om svaret är nej kan vi direkt säga att $\mathcal{U}$ inte är ett delrum. Om svaret är ja går vi vidare och kollar de två villkoren.

Exempel 2.3.1

Låt
$$\mathcal{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 + 2x_2 - x_3 = 0 \right\}.$$
Visa att $\mathcal{U}$ är ett delrum av $\mathbb{R}^3$.

Vi börjar med att observera att $\vec{0} \in \mathcal{U}$. Om
$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
får vi i ekvationen
$$0 + 2 \cdot 0 - 0 = 0,$$
vilket stämmer så $\vec{0} \in \mathcal{U}$. Alltså är $\mathcal{U}$ inte tom.

Låt nu $\vec{u}, \vec{v} \in \mathcal{U}$ och $\lambda \in \mathbb{R}$. Vi vill visa $\vec{u} + \vec{v} \in \mathcal{U}$ och $\lambda\vec{u} \in \mathcal{U}$. Vi skriver
$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \text{ och } \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}.$$
Eftersom $\vec{u}, \vec{v} \in \mathcal{U}$ får vi
$$u_1 + 2u_2 - u_3 = 0 \text{ och } v_1 + 2v_2 - v_3 = 0.$$

Vi sätter nu in $\vec{u}+\vec{v}$ i ekvationen som beskriver $\mathcal{U}$. I vänsterledet får vi

$$(u_1+v_1)+2(u_2+v_2)-(u_3+v_3) = u_1+2u_2-u_3+v_1+2v_2-v_3 = 0+0 = 0.$$

Alltså gäller $\vec{u}+\vec{v} \in \mathcal{U}$. På likande sätt får vi för $\lambda\vec{u}$ att

$$\lambda u_1 + 2\lambda u_2 - \lambda u_3 = \lambda(u_1 + 2u_2 - u_3) = \lambda 0 = 0.$$

Alltså gäller även $\lambda\vec{u} \in \mathcal{U}$ och vi får att $\mathcal{U}$ är ett delrum.

Vi undersöker en snarlik delmängd (vad är skillnaden?).

Exempel 2.3.2

Låt

$$\mathcal{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 + 2x_2 - x_3 = 4 \right\}.$$

Visa att $\mathcal{U}$ inte är ett delrum av $\mathbb{R}^3$.

Sätter vi in nollvektorn

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

i ekvationen så får vi $0+2\cdot 0-0 = 0 \neq 4$, så vi ser att $\vec{0} \notin \mathcal{U}$. Alltså kan $\mathcal{U}$ inte vara ett delrum.

Observera att det är två mycket olika saker att visa att en delmängd är ett delrum respektive inte är ett delrum. För att verifiera att en delmängd är ett delrum behöver vi visa att villkoren i definitionen stämmer. För att visa att en delmängd inte är ett delrum behöver vi bara visa att någonting som gäller för delrum går fel.

Exempel 2.3.3

Visa att $\mathcal{C}(\mathbb{R}) \subseteq \mathcal{F}(\mathbb{R})$ är ett delrum.

Nollelementet är funktionen som är konstant lika med 0 och den tillhör $\mathcal{C}(\mathbb{R})$. Om $f, g \in \mathcal{C}(\mathbb{R})$ så är f och g kontinuerliga funktioner. Då är även $(f+g)(x) = f(x) + g(x)$ en kontinuerlig funktion så

$f + g \in \mathcal{C}(\mathbb{R})$. Dessutom är varje skalning λf av f kontinuerlig så $\lambda f \in \mathcal{C}(\mathbb{R})$. Alltså är $\mathcal{C}(\mathbb{R})$ ett delrum.

Delrum kan ofta ges en geometrisk tolkning. Vi går igenom alla möjliga delrum i $\mathbb{R}^2$ och $\mathbb{R}^3$ som exempel.

2.3.1 Delrum i $\mathbb{R}^2$

Låt $\mathcal{U}$ vara ett delrum i $\mathbb{R}^2$, då finns följande alternativ:

- $\mathcal{U} = \{\vec{0}\}$, alltså består $\mathcal{U}$ endast av origo.
- $\mathcal{U}$ är en linje genom origo.
- $\mathcal{U} = \mathbb{R}^2$, alltså är $\mathcal{U}$ hela planet.

2.3.2 Delrum i $\mathbb{R}^3$

Låt $\mathcal{U}$ vara ett delrum i $\mathbb{R}^3$, då finns följande alternativ:

- $\mathcal{U} = \{\vec{0}\}$, alltså består $\mathcal{U}$ endast av origo.
- $\mathcal{U}$ är en linje genom origo.
- $\mathcal{U}$ är ett plan genom origo.
- $\mathcal{U} = \mathbb{R}^3$, alltså är $\mathcal{U}$ hela rummet.

Notera att dessa exempel kan beskrivas som lösningsmängder till linjära ekvationssystem. Att vi alltid har med $\vec{0}$ i lösningsmängden betyder att högerleden i ekvationssystemen måste vara 0. Sådana linjära ekvationssystem kallas homogena och ger i allmänhet upphov till delrum.

Sats 2.6

Låt $\mathcal{L}$ vara lösningsmängden till ett homogent linjärt ekvationssystem:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \quad\vdots \qquad\quad \vdots \qquad\quad \ddots \qquad\quad \vdots \quad\; \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases}$$

Då är $\mathcal{L} \subseteq \mathbb{R}^n$ ett delrum.

Bevis: Om $\vec{x} = \vec{u}$ och $\vec{x} = \vec{v}$ är lösningar till systemet är även $\vec{x} = \vec{u} + \vec{v}$ och $\vec{x} = \lambda\vec{u}$ lösningar. Mycket riktigt får vi för $\vec{x} = \vec{u} + \vec{v}$ i ekvation i att

$$\begin{aligned}a_{i1}(u_1 + v_1) + a_{i2}(u_2 + v_2) + \cdots + a_{in}(u_n + v_n) = \\ a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n + \\ a_{i1}v_1 + a_{i2}v_2 + \cdots + a_{in}v_n = 0 + 0 = 0.\end{aligned}$$

På likande sätt får vi för $x = \lambda\vec{u}$ i ekvation i att

$$\begin{aligned}a_{i1}\lambda u_1 + a_{i2}\lambda u_2 + \cdots + a_{in}\lambda u_n = \\ \lambda(a_{i1}u_1 + a_{i2}u_2 + \cdots + a_{in}u_n) = 0.\end{aligned}$$

På grund av denna sats kan vi vara säkra på att linjer genom origo är delrum av $\mathbb{R}^2$. Vi får även att linjer och plan genom origo är delrum i $\mathbb{R}^3$. Men hur kan vi vara säkra på att det inte finns några andra delrum i $\mathbb{R}^2$ och $\mathbb{R}^3$ (bortsett från origo själv och hela planet/rummet)? Det kommer vi se i nästa avsnitt.

2.4 Linjärkombinationer

Låt oss gå tillbaka till exempelt som vi började med:

$$\mathcal{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 + 2x_2 - x_3 = 0 \right\}$$

Vi har visat att $\mathcal{U}$ är ett delrum av $\mathbb{R}^3$ som kan förstås geometrisk som ett plan genom origo. Vi har infört $\mathcal{U}$ som lösningsmängden till ekvationen $x_1 + 2x_2 - x_3 = 0$. Det visade sig vara en praktisk beskrivning för att se att $\mathcal{U}$ är ett delrum, men för andra syften kan det vara praktiskt att beskriva $\mathcal{U}$ på parameterform. Vi får denna form genom att parametrisera lösningarna till $x_1 + 2x_2 - x_3 = 0$. Sätter vi $x_2 = s$ och $x_3 = t$ så får vi $x_1 = -2s + t$. Med andra ord är lösningarna

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -2s + t \\ s \\ t \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \text{där } s, t \in \mathbb{R}.$$

Vi skriver

$$\vec{v}_1 = \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} \text{ och } \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

och får då att

$$\mathcal{U} = \{s\vec{v}_1 + t\vec{v}_2 \mid s, t \in \mathbb{R}\}.$$

Uttrycket $s\vec{v}_1 + t\vec{v}_2$ beskriver en vektor som fås genom att skala om $\vec{v}_1$ och $\vec{v}_2$ och sedan addera resultaten. Ett sånt sätt att kombinera två vektorer kallas för en linjärkombination. Mer allmänt kan man ta linjärkombinationer av godtyckligt antal vektorer i ett vektorrum. Definitionen är som följer.

Definition 2.7

Låt $\mathcal{V}$ vara ett vektorrum, $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in \mathcal{V}$ och $c_1, c_2, \dots, c_n \in \mathbb{R}$. Då kallas vektorn

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n$$

en *linjärkombination* av $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$. Mängden av alla sådana linjärkombinationer kallas *spannet* (ibland det *linjära höljet*) av $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$, vilket betecknas

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n\} = \{c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \mid c_1, c_2, \dots, c_n \in \mathbb{R}\}.$$

Exempel 2.4.1

Från vår undersökning av ekvationen $x_1 + 2x_2 - x_3 = 0$ får vi att

$$\begin{aligned} \mathcal{U} &= \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 + 2x_2 - x_3 = 0 \right\} \\ &= \left\{ s\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \middle| \, s, t \in \mathbb{R} \right\} \\ &= \operatorname{span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\} \end{aligned}$$

Spannet är alltid ett delrum, mer precist har vi följande sats.

Sats 2.8

Låt $\mathcal{V}$ vara ett vektorrum och $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathcal{V}$. Då gäller att:

a) $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \mathcal{V}$ är ett delrum.

b) Om $\mathcal{U} \subseteq \mathcal{V}$ är ett delrum så att $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathcal{U}$, så är

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \mathcal{U}.$$

Bevis: a) Nollvektorn $\vec{0}$ är en linjärkombination av varje uppsättning vektorer så $\vec{0} \in \operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$. Ta två linjärkombinationer

$$\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n, \quad \vec{v} = d_1\vec{v}_1 + d_2\vec{v}_2 + \cdots + d_n\vec{v}_n.$$

Då är deras summa

$$\begin{aligned} \vec{u} + \vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n + \\ d_1\vec{v}_1 + d_2\vec{v}_2 + \cdots + d_n\vec{v}_n = \\ (c_1 + d_1)\vec{v}_1 + (c_2 + d_2)\vec{v}_2 + \cdots + (c_n + d_n)\vec{v}_n \in \mathcal{U}. \end{aligned}$$

På liknande sätt är skalningen av $\vec{u}$ med $\lambda \in \mathbb{R}$ vektorn

$$\begin{aligned} \lambda\vec{u} = \lambda(c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n) = \\ (\lambda c_1)\vec{v}_1 + (\lambda c_2)\vec{v}_2 + \cdots + (\lambda c_n)\vec{v}_n \in \mathcal{U}. \end{aligned}$$

b) Antag nu $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathcal{U}$. Eftersom $\mathcal{U}$ är ett delrum får vi att $c_1\vec{v}_1, c_2\vec{v}_2, \ldots, c_n\vec{v}_n \in \mathcal{U}$ och

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \in \mathcal{U}.$$

Alltså gäller

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \mathcal{U}.$$

Del b) i satsen säger alltså att om $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathcal{U}$ så gäller

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \mathcal{U}.$$

Ifall likhet gäller, det vill säga

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = \mathcal{U}$$

säger vi att vektorerna $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ spänner upp $\mathcal{U}$.

Med hjälp av ovanstående sats kan vi resonera oss fram till att det inte finns några andra delrum i $\mathbb{R}^2$ och $\mathbb{R}^3$ än de vi diskuterat.

Låt $\mathcal{U} \subseteq \mathcal{V}$ vara ett delrum där $\mathcal{V} = \mathbb{R}^2$ eller $\mathcal{V} = \mathbb{R}^3$. Om $\mathcal{U} \neq \{\vec{0}\}$ så finns $\vec{u} \in \mathcal{U}$ så att $\vec{u} \neq \vec{0}$. Då är $\operatorname{span}\{\vec{u}\} \subseteq \mathcal{U}$ en linje genom origo. Antingen är då $\operatorname{span}\{\vec{u}\} = \mathcal{U}$ eller så finns $\vec{v} \in \mathcal{U}$ så att $\vec{v} \notin \operatorname{span}\{\vec{u}\}$. I det andra fallet så är $\operatorname{span}\{\vec{u}, \vec{v}\} \subseteq \mathcal{U}$ ett plan genom origo. Om $\mathcal{V} = \mathbb{R}^2$, så tar möjligheterna slut eftersom $\mathcal{U} = \operatorname{span}\{\vec{u}, \vec{v}\} = \mathbb{R}^2$ då är den enda möjligheten. Om $\mathcal{V} = \mathbb{R}^3$ så finns två möjligheter. Antingen är $\mathcal{U} = \operatorname{span}\{\vec{u}, \vec{v}\}$ eller så finns $\vec{w} \in \mathcal{U}$ så att $\vec{w} \notin \operatorname{span}\{\vec{u}, \vec{v}\}$. Då måste vi ha $\mathbb{R}^3 = \operatorname{span}\{\vec{u}, \vec{v}, \vec{w}\} \subseteq \mathcal{U}$ och $\mathcal{U} = \mathbb{R}^3$.

Vi avslutar med att visa hur man avgör om en vektor tillhör ett spann.

Exempel 2.4.2

Avgör om

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \in \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Vi undersöker om det finns $s, t \in \mathbb{R}$ så att

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = s \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} + t \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} s + 3t \\ -2s + t \\ s + t \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ -2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} s \\ t \end{pmatrix}.$$

Likheten motsvarar ett linjärt ekvationssystem med totalmatris

$$\left(\begin{array}{cc|c} 1 & 3 & 1 \\ -2 & 1 & 2 \\ 1 & 1 & 3 \end{array}\right) \begin{matrix} 2 \quad -1 \\ \\ \\ \end{matrix} \sim \left(\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 7 & 4 \\ 0 & -2 & 2 \end{array}\right) \begin{matrix} \\ \\ -\frac{1}{2} \end{matrix} \sim$$

$$\left(\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & -1 \\ 0 & 7 & 4 \end{array}\right) \begin{matrix} \\ -7 \\ \\ \end{matrix} \sim \left(\begin{array}{cc|c} 1 & 3 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 11 \end{array}\right)$$

Den sista raden motsvarar $0 = 11$. Alltså saknas lösning. Det betyder att

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \notin \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

2.5 Linjärt oberoende

I föregående avsnitt gick vi igenom det spannet av en uppsättning vektorer, det vill säga mängden av alla vektorer som kan fås genom att ta linjärkombinationer av vektorerna i uppsättningen. Ju fler vektorer vi har detso fler linjärkombinationer kan vi bilda, så vi kan förvänta oss att spannet växer om vi lägger till fler vektorer i vår uppsättning. Om så faktiskt sker beror på hur vektorerna förhåller sig till varandra. I det här avsnittet ska vi reda ut hur detta fungerar. Vi börjar med att se på två exempel av spann av tre vektorer i $\mathbb{R}^3$.

Exempel 2.5.1

Skriv $\vec{v} = \begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix}$ som en linjärkombination av

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Beskriv även spannet $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ som delrum av $\mathbb{R}^3$.

Vi söker $c_1, c_2, c_3 \in \mathbb{R}$ så att

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3,$$

vilket är ekvivalent med matrisekvationen

$$\begin{pmatrix} 4 \\ 5 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

som ger

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 4 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & 3 \end{array}\right) \begin{array}{c} \\ \\ -1\;\,-1 \end{array} \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right) \begin{array}{c} \\ -1 \\ \\ \end{array} \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}\right)$$

Vi får en lösning $c_1 = -1$, $c_2 = 2$ och $c_3 = 3$. Det vill säga

$$\vec{v} = -\vec{v}_1 + 2\vec{v}_2 + 3\vec{v}_3.$$

Notera att vi inte har några nollrader i den sista koefficinetmatrisen. Det betyder att om vi ändrar $\vec{v}$ till någon annan vektor

$$\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$$

kommer systemet alltid gå att lösa. För tydlighets skull löser vi motsvarande ekvationssystem igen:

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & x \\ 0 & 1 & 1 & y \\ 0 & 0 & 1 & z \end{array}\right) \begin{array}{l} \\ \\ \boxed{-1}\boxed{-1} \end{array} \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & x-z \\ 0 & 1 & 0 & y-z \\ 0 & 0 & 1 & z \end{array}\right) \begin{array}{l} \\ \boxed{-1} \\ \\ \end{array} \sim$$
$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & x-y \\ 0 & 1 & 0 & y-z \\ 0 & 0 & 1 & z \end{array}\right)$$

Alltså gäller $\vec{v} = (x-y)\vec{v}_1 + (y-z)\vec{v}_2 + z\vec{v}_3$. Sätter vi in $x = 4$, $y = 5$ och $z = 3$ får vi samma lösning som tidigare. I synnerhet tillhör varje vektor $\vec{v} \in \mathbb{R}^3$ spannet $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Med andra ord är

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \mathbb{R}^3.$$

Vi tar ett till exempel på att beskriva spannet.

Exempel 2.5.2

Vilka vektorer $\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ tillhör $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$, där

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}?$$

Vi vill bestämma de vektorer $\vec{v}$, för vilka det finns $c_1, c_2, c_3 \in \mathbb{R}$ så att

$$\vec{v} = c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3.$$

Det leder till ett linjärt ekvationssystem med följande totalmatris

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & x \\ 0 & 2 & 4 & y \\ 1 & 0 & 5 & z \end{array}\right) \boxed{-1} \sim \left(\begin{array}{ccc|c} 1 & -1 & 3 & x \\ 0 & 2 & 4 & y \\ 0 & 1 & 2 & -x+z \end{array}\right) \sim$$

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & x \\ 0 & 1 & 2 & -x+z \\ 0 & 2 & 4 & y \end{array}\right) \boxed{-2} \sim \left(\begin{array}{ccc|c} 1 & -1 & 3 & x \\ 0 & 1 & 2 & -x+z \\ 0 & 0 & 0 & 2x+y-2z \end{array}\right)$$

Från sista raden ser vi att det finns lösning om och endast om $2x + y - 2z = 0$. Med andra ord är

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, 2x + y - 2z = 0 \right\},$$

vilket är ett plan genom origo. Alltså ligger alla tre vektorerna $\vec{v}_1, \vec{v}_2, \vec{v}_3$ i planet som ges av ekvationen $2x + y - 2z = 0$. Men detta plan är också spannet av de två första vektorerna, det vill säga

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2\} = \operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, 2x + y - 2z = 0 \right\}.$$

Särskilt gäller att $\vec{v}_3 \in \operatorname{span}\{\vec{v}_1, \vec{v}_2\}$. Mycket riktigt är

$$5\vec{v}_1 + 2\vec{v}_2 = 5 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} = \vec{v}_3$$

vilket ger en annan förklaring till varför $\operatorname{span}\{\vec{v}_1, \vec{v}_2\} = \operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. Hur kan vi hitta ett sådant förhållande? Ett sätt är faktiskt genom att skriva $\vec{0}$ som en linjärkombination av $\vec{v}_1, \vec{v}_2, \vec{v}_3$, det vill säga

$$c_1\vec{v}_1 + c_2\vec{v}_2 + c_3\vec{v}_3 = \vec{0}.$$

Det ger samma system som ovan fast nu är $x = 0, y = 0, z = 0$. Vi får alltså

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 2 & 4 & 0 \\ 1 & 0 & 5 & 0 \end{array}\right) \boxed{-1} \sim \left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \sim$$

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \boxed{-2} 1 \sim \qquad \left(\begin{array}{ccc|c} 1 & 0 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Sätter vi den sista obekanta till t får vi $c_1 = -5t$, $c_2 = -2t$ och $c_3 = t$, där $t \in \mathbb{R}$. Lösningen då $t = 0$ är inte så spännande:

$$0\vec{v}_1 + 0\vec{v}_2 + 0\vec{v}_3 = \vec{0}.$$

Men väljer vi till exempel $t = 1$ så får vi

$$-5\vec{v}_1 - 2\vec{v}_2 + \vec{v}_3 = \vec{0},$$

som vi kan skriva om till $\vec{v}_3 = 5\vec{v}_1 + 2\vec{v}_2$.

I exemplet ovan såg vi tre vektorer vars spann är det samma som spannet av bara två av dem. Det berodde på att den tredje vektorn var en linjärkombination av de två första. I en sådan situation säger vi att vektorerna är linjärt beroende. Mer allmänt gör vi följande definition.

Definition 2.9

Låt $\mathcal{V}$ vara ett vektorrum. Då kallas en uppsättning vektorer $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathcal{V}$ *linjärt oberoende* om ekvationen

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n = \vec{0}$$

bara har lösningen $x_1 = 0, x_2 = 0, \ldots, x_n = 0$. Annars kallas vektorerna *linjärt beroende*.

Ekvationen i definitionen kallas beroendeekvationen och lösningen $x_1 = 0, x_2 = 0, \ldots, x_n = 0$, som alltid finns kallas för den triviala lösningen. För att testa om en uppsättning vektorer är linjärt beroende kan vi alltså lösa beroendeekvationen. Om vi hittar en lösning som inte är den triviala är de linjärt beroende. Om bara den triviala lösningen finns så är de linjärt oberoende.

Vi återbesöker nu exemplen som vi började med.

Exempel 2.5.3

Vektorerna

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

är linjärt oberoende.

Beroendeekvationen $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{0}$ ger

$$\left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \boxed{-1}\boxed{-1} \sim \left(\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right) \boxed{-1} \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right)$$

Vi har bara en lösning $x_1 = 0$, $x_2 = 0$ och $x_3 = 0$. Alltså är vektorerna linjärt oberoende.

Vektorerna

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$$

är linjärt beroende.

Beroendeekvationen $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{0}$ ger

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 2 & 4 & 0 \\ 1 & 0 & 5 & 0 \end{array}\right) \boxed{-1} \sim \left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 2 & 4 & 0 \\ 0 & 1 & 2 & 0 \end{array}\right) \sim$$

$$\left(\begin{array}{ccc|c} 1 & -1 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 2 & 4 & 0 \end{array}\right) \boxed{-2}\,1 \sim \left(\begin{array}{ccc|c} 1 & 0 & 5 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Vi har oändligt många lösningar $x_1 = -5t$, $x_2 = -2t$ och $x_3 = t$, där $t \in \mathbb{R}$. Alltså är vektorerna linjärt beroende.

Observera vad som gjorde skillnaden i exemplen ovan. Skillnaden kan ses i koefficientmatrisen för den radkanoniska formen vi når i slutet av lösningen. I det fösta exemplet hade varje kolonn en ledande etta. Vi fick därför inga parametrar och således bara den triviala lösningen. I det andra exemplet saknade den sista kolonnen ledande etta och vi fick därför parameterlösningar.

Att en uppsättning vektorer är linjärt beroende innebär att vi kan skriva någon av vektorerna (eller kanske flera) som en linjärkombination av de andra. Mer precist gäller följande sats.

Sats 2.10

Låt $\mathcal{V}$ vara ett vektorrum och $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n \in \mathcal{V}$. Då är vektorerna linjärt beroende om och endast om någon av vektorerna $\vec{v}_i$ är en linärkombination av de andra, det vill säga

$$\vec{v}_i \in \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}.$$

Då gäller att

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} = \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}$$

.

Den sista likheten i satsen säger att vi kan stryka vektorn $\vec{v}_i$ utan att påverka spannet. Notera att satsen inte säger något om vilken av vektorerna som kan strykas. Ofta kan flera vektorer strykas utan att påverka spannet. Det satsen säger att det måste finnas minst en sådan vektor. I beviset får vi en ledtråd om vilka vektorer som kan strykas.

Bevis: Om vektorerna är linjärt beroende så har vi

$$x_1\vec{v}_1 + \cdots + x_i\vec{v}_i + \cdots + x_n\vec{v}_n = \vec{0}$$

där något av talen x_i måste vara skilt från noll. Vi kan då lösa ut $\vec{v}_i$ från ekvationen ovan på följande vis. Vi får

$$x_i\vec{v}_i = -x_1\vec{v}_1 \cdots - x_{i-1}\vec{v}_{i-1} - x_{i+1}\vec{v}_{i+1} \cdots - x_n\vec{v}_n.$$

Eftersom $x_i \neq 0$ kan vi dela med x_i och får då

$$\vec{v}_i = -\frac{x_1}{x_i}\vec{v}_1 \cdots - \frac{x_{i-1}}{x_i}\vec{v}_{i-1} - \frac{x_{i+1}}{x_i}\vec{v}_{i+1} \cdots - \frac{x_n}{x_i}\vec{v}_n,$$

det vill säga

$$\vec{v}_i \in \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}.$$

Eftersom $\operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}$ är ett delrum av $\mathcal{V}$ som innehåller vektorerna $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$, så gäller

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\} \subseteq \operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}.$$

Men den omvända inklusionen gäller också, så inklusionen måste vara en likhet.

På likande sätt kan vi se att villkoret

$$\vec{v}_i \in \text{span}\{\vec{v}_1, \ldots, \vec{v}_{i-1}, \vec{v}_{i+1}, \ldots, \vec{v}_n\}.$$

ger en lösning till ekvationen

$$x_1\vec{v}_1 + \cdots + x_i\vec{v}_i + \cdots + x_n\vec{v}_n = \vec{0}$$

där $x_i \neq 0$ så vektorerna är linjärt beroende i detta fall.

Exempel 2.5.4

Låt

$$\vec{v}_1 = \begin{pmatrix}1\\2\\4\end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix}0\\1\\1\end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix}2\\1\\5\end{pmatrix}.$$

Av gör om $\vec{v}_1, \vec{v}_2, \vec{v}_3$ är linjärt oberoende. Skriv om möjligt var och en av vektorerna som en linjärkombination av de andra.

Beroendeekvationen $x_1\vec{v}_1 + x_2\vec{v}_2 + x_3\vec{v}_3 = \vec{0}$ ger

$$\left(\begin{array}{ccc|c}1&0&2&0\\2&1&1&0\\4&1&5&0\end{array}\right) \boxed{-2}\boxed{-4} \sim \left(\begin{array}{ccc|c}1&0&2&0\\0&1&-3&0\\0&1&-3&0\end{array}\right) \boxed{-1} \sim \left(\begin{array}{ccc|c}1&0&2&0\\0&1&-3&0\\0&0&0&0\end{array}\right)$$

Vi har oändligt många lösningar $x_1 = -2t$, $x_2 = 3t$ och $x_3 = t$, där $t \in \mathbb{R}$. Alltså är vektorerna linjärt beroende. Lösningen då $t = 1$ ger

$$-2\vec{v}_1 + 3\vec{v}_2 + \vec{v}_3 = \vec{0}.$$

Eftersom alla $x_i \neq 0$ i detta fall kan vi lösa ut varje vektor ur ekvationen och får på så sätt:

$$\begin{aligned}\vec{v}_1 &= \frac{3}{2}\vec{v}_2 + \frac{1}{2}\vec{v}_3,\\ \vec{v}_2 &= \frac{2}{3}\vec{v}_1 - \frac{1}{3}\vec{v}_3,\\ \vec{v}_3 &= 2\vec{v}_1 - 3\vec{v}_2.\end{aligned}$$

I det här fallet är alltså

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{span}\{\vec{v}_2, \vec{v}_3\} = \operatorname{span}\{\vec{v}_1, \vec{v}_3\} = \operatorname{span}\{\vec{v}_1, \vec{v}_2\}.$$

I exemplen ovan har vi sett på vektorer i $\mathbb{R}^3$, så vi tar nu några exempel med andra vektorrum.

Exempel 2.5.5

Låt $p_1, p_2, p_3 \in \mathcal{P}_3$ vara polynomen

$$p_1(t) = 1 + t^3, \quad p_2(t) = t + t^3, \quad p_3(t) = 1 + t^2,$$

Avgör om p_1, p_2, p_3 är linjärt oberoende.

Som tidigare undersöker vi ekvationen

$$x_1 p_1(t) + x_2 p_2(t) + x_3 p_3(t) = 0$$

Denna ekvation ska gälla för alla t så vi skriver om vänsterledet och samlar termer av samma grad för sig:

$$x_1(1+t^3) + x_2(t+t^3) + x_3(1+t^2) = x_1 + x_3 + x_2 t + x_3 t^2 + (x_1 + x_2)t^3$$

Ekvationen gäller då för alla t om och endast om

$$\begin{cases} x_1 \qquad\quad + x_3 = 0 \\ \qquad x_2 \qquad\quad = 0 \\ \qquad\qquad + x_3 = 0 \\ x_1 + x_2 \qquad\; = 0 \end{cases}$$

vilket ger ett linjärt ekvationssystem med totalmatris

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \end{array}\right) \begin{array}{l} \boxed{-1} \\ \\ \\ \\ \end{array} \sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \end{array}\right) \begin{array}{l} \\ \boxed{-1} \\ \\ \\ \end{array} \sim$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 \end{array}\right) \begin{array}{l} \\ \\ 1 \; \boxed{-1} \\ \\ \end{array} \sim \left(\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

Vi har entydig lösning $x_1 = 0$, $x_2 = 0$, $x_3 = 0$. Alltså är polynomen linjärt oberoende. Observera att återigen fick vi resultatet linjärt oberoende på grund av att vi har ledande ettor i varje kolonn.

Exempel 2.5.6

Låt $\mathcal{C}(\mathbb{R})$ vara vektorrummet av kontinuerliga funktioner från $\mathbb{R}$ till $\mathbb{R}$.

Vi börjar med att studera funktionerna $f, g \in \mathcal{C}(\mathbb{R})$ givna av $f(t) = e^t$ och $g = \sin t$. Vi undersöker om $f, g \in \mathcal{C}(\mathbb{R})$ är linjärt oberoende. Som tidigare ser vi på ekvationen

$$x_1 f(t) + x_2 g(t) = 0$$

som ska gälla för alla t. Det betyder att vi kan välja värden på t och på så sätt få villkor på x_1 och x_2. Väljer vi $t = 0$ så är $f(0) = e^0 = 1$ och $g(0) = \sin 0 = 0$, vilket ger $x_1 = 0$. Väljer vi $t = \frac{\pi}{2}$ så är $f(\frac{\pi}{2}) = e^{\frac{\pi}{2}}$ och $g(\frac{\pi}{2}) = \sin\frac{\pi}{2} = 1$, vilket ger $x_1 e^{\frac{\pi}{2}} + x_2 = 0$. I och med att vi vet att $x_1 = 0$ får vi då $x_2 = 0$ så f, g är linjärt oberoende.

Låt nu $f, g, h \in \mathcal{C}(\mathbb{R})$ vara tre andra funktioner, nämligen

$$f(t) = \sin^2 t, \quad g(t) = \cos^2 t, \quad h(t) = \cos(2t),$$

Eftersom $h(t) = \cos(2t) = \cos^2 t - \sin^2 t = g(t) - f(t)$ gäller för alla $t \in \mathbb{R}$, ser vi att f, g, h är linjärt beroende.

Vi avslutar med att undersöka vad det betyder att två vektorer $\vec{v}, \vec{w}$ i ett vektorrum $\mathcal{W}$ är linjärt beroende. Beroendeekvationen är

$$x_1\vec{v} + x_2\vec{w} = \vec{0}.$$

Om $\vec{v} = \vec{0}$ så kan vi ta $x_1 = 1$ och $x_2 = 0$. Om $\vec{w} = \vec{0}$ så kan vi ta $x_1 = 0$ och $x_2 = 1$. I dessa fall är alltså vektorerna alltid linjärt beroende.

Om både $\vec{v} \neq 0$ och $\vec{w} \neq 0$, så är $x_1 = 0$ om och endast om $x_2 = 0$. Om vi har en icketrival lösning måste då $\vec{v} = -\frac{x_2}{x_1}\vec{w}$ och $\vec{w} = -\frac{x_1}{x_2}\vec{v}$, med andra ord är $\vec{v}$ och $\vec{w}$ omskalningar av varandra. Vi har då $\operatorname{span}\{\vec{v}, \vec{w}\} = \operatorname{span}\{\vec{w}\} = \operatorname{span}\{\vec{v}\}$.

2.6 Bas och dimension

Vi har sett många exempel på vektorrum där vektorerna kan beskrivas genom att ange ett visst antal reella tal. Ofta kan antalet ses som rummets dimension. Vi ska nu reda ut vad som pågår i dessa exempel och definiera vad exakt vi menar med dimensionen av ett vektorrum. Vi börjar med att ta upp några av exemplen som vi sett.

Exempel 2.6.1

a) Vektorrumet $\mathbb{R}^2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \middle| \, x_1, x_2 \in \mathbb{R} \right\}$ är det tvådimensionella planet. Eftersom

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \end{pmatrix} = x_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

ser vi att

$$\mathbb{R}^2 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

b) Vektorrumet $\mathbb{R}^3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1, x_2, x_3 \in \mathbb{R} \right\}$ är det tredimensionella rummet. På liknande sätt som för $\mathbb{R}^2$ har vi att

$$\mathbb{R}^3 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

c) Vektorrumet $\mathcal{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 + 2x_2 - x_3 = 0 \right\}$ är ett plan och bör ses som tvådimensionellt. Genom att parametrisera lösningarna till $x_1 + 2x_2 - x_3 = 0$ ser vi att

$$\mathcal{U} = \left\{ \begin{pmatrix} -2s + t \\ s \\ t \end{pmatrix} \middle| \, s, t \in \mathbb{R} \right\} = \operatorname{span}\left\{ \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

d) Låt

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}.$$

Vi har tidigare sett att dessa vektorer är linjärt beroende eftersom $\vec{v}_3 = 2\vec{v}_1 - 3\vec{v}_2$. Alltså är $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{span}\{\vec{v}_1, \vec{v}_2\}$. Däremot är inte $\vec{v}_1, \vec{v}_2$ linjärt beroende (de är inte omskalningar av varandra). Delrummet $\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ är ett plan som vi bör se som tvådimensionellt.

Notera att vi i vart och ett av exemplen kan skriva vektorrummet som spannet av några linjärt oberoende vektorer. Dessutom är antalet vektorer det vi uppfattar som rummets dimension. En sådan uppsättning vektorer kallas en bas. I nästa delavsnitt går vi igenom baser och deras egenskaper innan vi definierar dimension.

2.6.1 Baser och koordinater

Motiverade av exemplen ovan gör vi följande definition

Definition 2.11

Låt $\mathcal{V}$ vara ett vektorrum. Då kallas en ordnad uppsättning vektorer $\underline{u} = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ i $\mathcal{V}$ en *bas* om

1. $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\} = \mathcal{V}$,
2. $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ är linjärt oberoende.

Exempel 2.6.2

a) $(\vec{u}_1, \vec{u}_2)$ där $\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ och $\vec{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ är en bas i $\mathbb{R}^2$. Vektorerna är linjärt oberoende och spänner upp $\mathbb{R}^2$.

b) $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ där $\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ och $\vec{u}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ är inte en bas i $\mathbb{R}^2$. Vektorerna är ej linjärt oberoende, men spänner upp $\mathbb{R}^2$.

c) $(\vec{u}_1)$ där $\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ är inte en bas i $\mathbb{R}^2$. Som uppsättning är den enstaka vektron linjärt oberoende, men spänner inte upp hela $\mathbb{R}^2$.

d) $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ där $\vec{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ och $\vec{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ är en bas i $\mathbb{R}^3$. Vektorerna är linjärt oberoende och spänner upp $\mathbb{R}^3$.

e) $(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ där $\vec{u}_1 = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$, $\vec{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ och $\vec{u}_3 = \begin{pmatrix} 2 \\ 1 \\ 5 \end{pmatrix}$ är inte en bas i $\mathbb{R}^3$. Som vi sett i tidigare exempel är vektorerna linjärt beroende och spänner bara upp ett plan i $\mathbb{R}^3$.

I exempel b) spänner vektorerna upp hela vektorrummet, men de är linjärt beroende. I någon mening är de för många för att vara en bas: om vi plockar bort en av vektorerna får vi en bas. I exempel c) har vi istället för få vektorer: vi kan få en bas genom att lägga till en vektor. I exempel e) har vi samma antal vektorer som i d), men de är ändå inte en bas. De uppfyller inget av villkoren i definitionen. Dessa exempel visar på typiska fall för när en uppsättning vektorer inte är en bas.

I vektorrummet $\mathbb{R}^n$ finns en speciell bas som kallas *standardbasen.* Den består av följande vektorer

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Vi skriver $\underline{e} = (\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n)$. Det som gör basen speciell är att

$$\begin{aligned} \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix}, + \dots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, + \dots x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \end{aligned}$$

Med andra ord är $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \dots x_n\vec{e}_n$. Koordinaterna $x_1, x_2, \dots, x_n$ är precis de tal som vi skalar om basvektorerna med för att får $\vec{x}$.

Något likande sker i vektorrummet

$$\mathcal{P}_n = \{a_0 + a_1 t + \cdots a_n t^n \mid a_i \in \mathbb{R}\}.$$

Om vi sätter $p_0(t) = 1, p_1(t) = t, \dots p_n(t) = t^n$ så ser vi att ett godtyckligt polynom

$$p(t) = a_0 + a_1 t + \cdots a_n t^n = a_0 p_0(t) + a_1 p_1(t) + \cdots + a_n p_n(t)$$

För att komma ihåg polynomet $p(t)$ räcker det med att komma ihåg talen $a_0, a_1, \ldots, a_n$ och vi kan förstå dessa tal som ett slags koordinater för $p(t)$. På samma sätt kan man i varje vektorrum med en bas representera vektorerna med koordinater.

Sats 2.12

Låt $\underline{u} = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ vara en bas i ett vektorrum $\mathcal{V}$. För varje $\vec{v} \in \mathcal{V}$ finns entydigt bestämda tal $x_1, x_2, \ldots x_n$ så att

$$\vec{v} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n.$$

Bevis: Att $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\} = \mathcal{V}$ säger precis att det finns tal $x_1, x_2, \ldots x_n$ så att

$$\vec{v} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n.$$

Vi vill nu visa att de är entydiga. Antag därför att

$$\vec{v} = y_1\vec{u}_1 + y_2\vec{u}_2 + \cdots + y_n\vec{u}_n$$

för några tal $y_1, y_2, \ldots, y_n \in \mathbb{R}$. Vi tar differensen av vänsterled respektive högerled och får då

$$\vec{0} = (x_1 - y_1)\vec{u}_1 + (x_2 - y_2)\vec{u}_2 + \cdots + (x_n - y_n)\vec{u}_n.$$

Att $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ är linjärt oberoende, säger nu att:

$$\left\{\begin{aligned} x_1 - y_1 &= 0 \\ x_2 - y_2 &= 0 \\ &\vdots \\ x_n - y_n &= 0 \end{aligned}\right. \Leftrightarrow \left\{\begin{aligned} x_1 &= y_1 \\ x_2 &= y_2 \\ &\vdots \\ x_n &= y_n \end{aligned}\right.$$

Talen $x_1, x_2, \ldots x_n$ kallas *koordinaterna* för $\vec{v}$ i basen $\underline{u}$. Kolonnvektorn som innehåller dessa tal kallas för *koordinatvektorn* av $\vec{v}$ i basen $\underline{u}$ och betecknas $\vec{v}_{\underline{u}}$. Med andra ord har vi

$$\vec{v}_{\underline{u}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Leftrightarrow \vec{v} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n.$$

Att talen $x_1, x_2, \dots x_n$ är entydigt bestämda av $\vec{v}$ gör att vi kan översätta från vektorer i $\mathcal{V}$ till koordinater i $\mathbb{R}^n$ och tillbaka. Vi kan på så sätt föra över frågor om vektorer i $\mathcal{V}$ till förhoppningsvis mer hanterbara frågor om koordinatvektorer i $\mathbb{R}^n$. Särskilt gäller att vi kan beräkna addition och skalning av vektorer i $\mathcal{V}$ genom att utföra motsvarande operationer på deras koordinater.

Sats 2.13

Låt $\underline{u}$ vara en bas i ett vektorrum $\mathcal{V}$. Då gäller följande för $\vec{v}, \vec{w} \in \mathcal{V}$ och $\lambda \in \mathbb{R}$.

a) $(\vec{v} + \vec{w})_{\underline{u}} = \vec{v}_{\underline{u}} + \vec{w}_{\underline{u}}$,

b) $(\lambda \vec{v})_{\underline{u}} = \lambda \vec{v}_{\underline{u}}$.

Bevis: Kalla basen $\underline{u} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$. Skriv

$$\vec{v}_{\underline{u}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Leftrightarrow \vec{v} = x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_n\vec{u}_n.$$

$$\vec{w}_{\underline{u}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \Leftrightarrow \vec{w} = y_1\vec{u}_1 + y_2\vec{u}_2 + \dots + y_n\vec{u}_n.$$

Då är

$$\vec{v} + \vec{w} = (x_1 + y_1)\vec{u}_1 + (x_2 + y_2)\vec{u}_2 + \dots + (x_n + y_n)\vec{u}_n$$

så

$$(\vec{v} + \vec{w})_{\underline{u}} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \vec{v}_{\underline{u}} + \vec{w}_{\underline{u}}$$

På liknande sätt är

$$\lambda\vec{v} = \lambda(x_1\vec{u}_1 + x_2\vec{u}_2 + \dots + x_n\vec{u}_n) = (\lambda x_1)\vec{u}_1 + (\lambda x_2)\vec{u}_2 + \dots + (\lambda x_n)\vec{u}_n$$

så

$$(\lambda\vec{v})_{\underline{u}} = \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix} = \lambda \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \lambda\vec{v}_{\underline{u}}.$$

Vi går nu igenom ett exempel på hur vi kan översätta mellan vektorer och koordinater i $\mathbb{R}^2$.

Exempel 2.6.3

Vi undersöker en bas och motsvarande koordinater i $\mathbb{R}^2$.

a) Visa att $\underline{u} = (\vec{u}_1, \vec{u}_2)$ där $\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ och $\vec{u}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$ är en bas i $\mathbb{R}^2$.

b) Hitta $\vec{v}_{\underline{u}}$ där $\vec{v} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

c) Vilken vektor $\vec{w}$ uppfyller $\vec{w}_{\underline{u}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$?

a) Vi studerar ekvationen

$$x_1\vec{u}_1 + x_2\vec{u}_2 = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Den ger ett ekvationssystem med totalmatris

$$\left(\begin{array}{cc|c} 1 & -1 & x \\ 1 & 1 & y \end{array}\right) \boxed{-1} \sim \left(\begin{array}{cc|c} 1 & -1 & x \\ 0 & 2 & -x+y \end{array}\right) \boxed{\tfrac{1}{2}} \sim$$

$$\left(\begin{array}{cc|c} 1 & -1 & x \\ 0 & 1 & -\frac{x}{2}+\frac{y}{2} \end{array}\right) 1 \sim \left(\begin{array}{cc|c} 1 & 0 & \frac{x}{2}+\frac{y}{2} \\ 0 & 1 & -\frac{x}{2}+\frac{y}{2} \end{array}\right)$$

Eftersom vi har lösning för varje värde på x och y ser vi att $\operatorname{span}\{\vec{u}_1, \vec{u}_2\} = \mathbb{R}^2$. Dessutom är lösningen entydig, så fallet $x = 0, y = 0$ visar att $\vec{u}_1, \vec{u}_2$ är linjärt oberoende.

b) Vi söker x_1, x_2 i fallet då $x = 1$ och $y = 3$. Detta motsvarar

$$\left(\begin{array}{cc|c} 1 & -1 & 1 \\ 1 & 1 & 3 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 1 & 1 \end{array}\right)$$

Vi får alltså $\vec{v}_{\underline{u}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

d) $\vec{w}_{\underline{u}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ betyder precis att

$$\vec{w} = 1 \cdot u_1 + 1 \cdot u_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}.$$

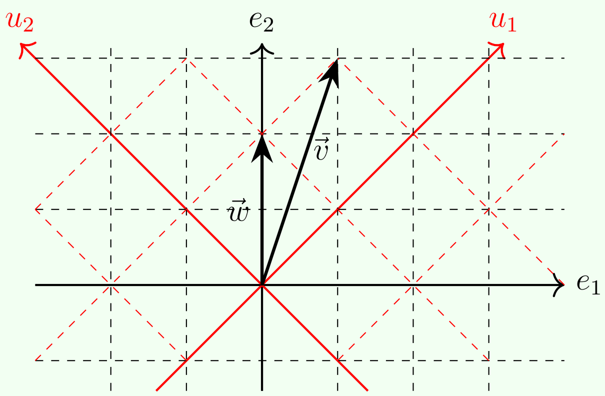
I detta exempel ser vi hur vektorerna $\vec{v}$ och $\vec{w}$ kan beskrivas med sina standardkoordinater

$$\vec{v} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$$

eller med sina koordinater i basen $\underline{u}$.

$$\vec{v}_{\underline{u}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{w}_{\underline{u}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Nedan ritar vi vektorerna $\vec{v}$ och $\vec{w}$ i $\mathbb{R}^2$ med koordinater från basen $\underline{e}$ markerade i svart och koordinater från basen $\underline{u}$ markerade i rött.



Det finns två grundläggande strategier för att hitta en bas i ett vektorrum $\mathcal{V}$. Den första är att först hitta vektorer som spänner upp hela $\mathcal{V}$. Om de är linjärt oberoende får vi en bas. Om inte kan vi ta bort minst en vektor utan att påverka spannet. Vi kan fortsätta att ta bort vektorer tills vi når en linjärt oberoende uppsättning med samma spann vilket då är en bas.

Den andra strategin är att utgå från en linjärt obereonde mängd vektorer. Om de spänner upp hela $\mathcal{V}$ har vi en bas. Om inte kan vi lägga till en vektor i $\mathcal{V}$ utanför spannet. Som nästa sats visar får vi då en linjärt oberoende mängd igen, men som nu har ett större spann. Vi kan upprepa denna process så länge vektorerna inte spänner upp hela $\mathcal{V}$. Om processen tar slut betyder det att vi når en bas i $\mathcal{V}$.

Sats 2.14

Låt $\mathcal{V}$ vara ett vektorrum och $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n \in \mathcal{V}$. Om $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$ är linjärt obereode och $\vec{u}_n \notin \operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}\}$ så är också $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ linjärt oberoende.

Bevis: Vi skriver upp beroendeekvationen

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n = \vec{0}.$$

Vi måste ha $x_n = 0$, för annars skulle vi kunna lösa ut $\vec{u}_n$ vilket skulle ge

$$\vec{u}_n \in \operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}\}.$$

Alltså är

$$x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_{n-1}\vec{u}_{n-1} = \vec{0},$$

vilket ger att alla $x_i = 0$ eftersom $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{n-1}$ är linjärt obereode.

Exempel 2.6.4

Hitta en bas i

$$\mathcal{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 - x_2 - 2x_3 = 0 \right\}.$$

Utvidga den till en bas i hela $\mathbb{R}^3$.

Vi sätter $x_2 = s$ och $x_3 = t$. Då ger $x_1 - x_2 - 2x_3 = 0$ att $x_1 = s + 2t$ och

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} s+2t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

Sätt

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

Då är $(\vec{u}_1, \vec{u}_2)$ en bas i $\mathcal{U}$. Vi väljer nu någon vektor $\vec{u}_3 \notin \mathcal{U}$, till exempel

$$\vec{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Då är $\underline{u} = (\vec{u}_1, \vec{u}_2, \vec{u}_3)$ linjärt oberoende. Dessutom är $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ ett delrum i $\mathbb{R}^3$ som är strikt större än planet $\mathcal{U}$. Den enda möjligheten är $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \mathbb{R}^3$. Alltså är $\underline{u}$ en bas i $\mathbb{R}^3$.

I det sista exemplet är det kanske inte helt klart att $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\} = \mathbb{R}^3$. Hur kan vi vara säkra på att vi inte kan hitta någon ytterligare vektor $\vec{u}_4 \in \mathbb{R}^3$ som är utanför spannet. Då skulle $\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4$ vara linjärt oberoende i $\mathbb{R}^3$, vilket är omöjligt. Mer allmänt gäller följande sats.

Sats 2.15

Låt $\underline{u} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_m)$ vara en bas i ett vektorrum $\mathcal{V}$. Varje uppsättning vektorer $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ där $n > m$ är linjärt beroende.

Bevis: Vi skriver upp beroendeekvationen

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_n\vec{v}_n = \vec{0}$$

och tar koordinater i basen $\underline{u}$ av vänsterlerd och högerled:

$$(x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_n\vec{v}_n)_{\underline{u}} = \vec{0}_{\underline{u}}$$

Vänsterledet är

$$\begin{gathered}(x_1\vec{v}_1 + x_2\vec{v}_2 + \dots + x_n\vec{v}_n)_{\underline{u}} = \\ (x_1\vec{v}_1)_{\underline{u}} + (x_2\vec{v}_2)_{\underline{u}} + \dots + (x_n\vec{v}_n)_{\underline{u}} = \\ x_1(\vec{v}_1)_{\underline{u}} + x_2(\vec{v}_2)_{\underline{u}} + \dots + x_n(\vec{v}_n)_{\underline{u}} = A\vec{x}\end{gathered}$$

där A är matrisen vars kolonner är $(\vec{v}_1)_{\underline{u}}, (\vec{v}_2)_{\underline{u}}, \dots, (\vec{v}_n)_{\underline{u}}$. I synnerhet är A en $(m \times n)$-matris, så A har fler kolonner än rader på grund av $n > m$. Men det betyder att ekvationen $A\vec{x} = \vec{0}$ måste ha parameterlösningar och vi får att $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$ är linjärt beroende.

Exempel 2.6.5

Vektorerna

$$\begin{pmatrix}1\\0\\1\end{pmatrix}, \begin{pmatrix}1\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}$$

i $\mathbb{R}^3$ är linjärt beroende eftersom $\mathbb{R}^3$ har en bas med tre element, men vi har fyra vektorer. Vi kan också se det från beroendeekvationen som ger ett ekvationssystem med totalmatris

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \end{array}\right)$$

Vi har alltså ett homogent system med tre ekvationer och fyra obekanta. Det har garanterat oändligt många lösningar.

Från föregående sats följer nästa viktiga resultat:

Sats 2.16

Låt $\underline{u}$ och $\underline{v}$ vara baser i ett vektorrum $\mathcal{V}$. Då har $\underline{u}$ och $\underline{v}$ samma antal element.

Bevis: Säg att $\underline{u}$ har m element och $\underline{v}$ har n element. Då är $m < n$ omöjligt eftersom $\underline{v}$ består av n linjärt oberoende vektorer, men basen $\underline{u}$ har m element. På samma sätt är $m > n$ omöjlligt. Därmed är enda möjligheten att $m = n$.

Eftersom alla baser har samma antal element kan vi göra följande definition.

Definition 2.17

Om ett vektorrum $\mathcal{V}$ har en bas $\underline{u} = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ med n element så säger vi att $\mathcal{V}$ har dimension n. Vi skriver då $\dim \mathcal{V} = n$. Om ingen sådan ändlig bas finns så säger vi att $\mathcal{V}$ har oändlig dimension och skiver $\dim \mathcal{V} = \infty$. Ett speciellt fall är då $\mathcal{V}$ bara består av nollvektorn. Då skiver vi $\dim \mathcal{V} = 0$.

Hur ska vi tolka vad det betyder att $\mathcal{V}$ inte har någon ändlig bas? Våra strategier för att hitta en bas misslyckas då. Alltså finns inte ens en ändlig uppsättning vektorer som spänner upp $\mathcal{V}$. Från den skulle vi kunna få en bas genom att stryka vektorer tills vi når en linjärt oberoende uppsättning vektorer som spänner upp $\mathcal{V}$. Det betyder även att vår andra strategi kommer att misslyckas. Om vi börjar med $\vec{u}_1 \in \mathcal{V}$, så att $\vec{u}_1 \neq 0$ så är $(\vec{u}_1)$ linjärt oberoende, men inte en bas. Då finns $\vec{u}_2 \in \mathcal{V}$ linjärt oberoende med $\vec{u}_1$. Eftersom inte heller $(\vec{u}_1, \vec{u}_2)$ är en bas i $\mathcal{V}$ finns $\vec{u}_3$ linjärt oberoende med $\vec{u}_1$ och $\vec{u}_2$. Fortsätter vi på samma sätt får vi fler och fler linjärt oberoende vektorer, men når aldrig en bas. Det verkar kanske som en märklig situation, men som vi ska se lite senare finns många exempel på vektorrum av oändlig dimension. De som vi studerar oftast i kursen är dock av ändlig dimension.

Exempel 2.6.6

a) $\dim \mathbb{R}^2 = 2$, planet är tvådimensionellt som förväntat.

b) $\dim \mathbb{R}^3 = 3$, rummet är tredimensionellt som förväntat.

c) $\dim \mathbb{R}^n = n$, följer av att betrakta standardbasen.

d) $\dim \mathcal{P}_n = n + 1$, följer av att betrakta standardbasen.

Det är vettigt att tänka på dimensionen som någon slags storlek. Ur det perspektivet är följande sats rimlig.

Sats 2.18

Låt $\mathcal{V}$ vara ett vektorrum av ändlig dimension och $\mathcal{U} \subseteq \mathcal{V}$ ett delrum. Då gäller $\dim \mathcal{U} \leq \dim \mathcal{V}$. Dessutom gäller $\dim \mathcal{U} = \dim \mathcal{V}$ om och endast om $\mathcal{U} = \mathcal{V}$.

Bevis: Skriv $\dim \mathcal{V} = n$. Vi bygger upp en linjärt oberoende uppsättning vektorer $\vec{u}_1, \vec{u}_2, \ldots \in \mathcal{U}$ på det sätt vi beskrivit tidigare. Processen måste ta slut efter $m \leq n$ steg. Annars skulle vi få fler än n linjärt oberoende vektorer i $\mathcal{V}$, vilket är omöjligt. Det betyder att $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m)$ är en bas i $\mathcal{U}$ och $\dim \mathcal{U} = m \leq n = \dim \mathcal{V}$. Om $\mathcal{U} = \mathcal{V}$ gäller $m = n$ eftersom alla baser har samma antal element. Om $\mathcal{U} \neq \mathcal{V}$ så finns $\vec{u}_{m+1} \in \mathcal{V}$ som inte är i $\mathcal{U}$. Då är $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m, \vec{u}_{m+1}$ linjärt oberoende och $m + 1 \leq n$. Alltså är $\dim \mathcal{U} = m < n = \dim \mathcal{V}$ i detta fall.

Exempel 2.6.7

a) Låt
$$\mathcal{U} = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \middle| \, x_1 + 2x_2 - x_3 = 0 \right\} \subseteq \mathbb{R}^3.$$
Då är $\mathcal{U}$ ett plan och $\dim \mathcal{U} = 2 < 3 = \dim \mathbb{R}^3$.

b) Vi har $\dim \mathcal{P} = \infty$. Om vi hade $\dim \mathcal{P} = n$ för något n så skulle det motsäga $\mathcal{P}_n \subseteq \mathcal{P}$ eftersom $\dim \mathcal{P}_n = n + 1 > n$.

c) Vi kan tolka $\mathcal{P}$ som ett delrum av $\mathcal{C}(\mathbb{R})$ (rummet av kontinuerliga funktioner från $\mathbb{R}$ till $\mathbb{R}$). Alltså ger $\dim \mathcal{P} = \infty$ att $\dim \mathcal{C}(\mathbb{R}) = \infty$.

Om vi söker en bas i ett vektorrum $\mathcal{V}$ sånt att $\dim \mathcal{V} = n$, så vet vi att vi behöver n vektorer. Nästa sats visar att om vi har den situationen behöver vi bara kolla ett av det två villkoren för att vara en bas.

Sats 2.19

Låt $\mathcal{V}$ vara ett vektorrum med $\dim \mathcal{V} = n$ och en uppsättning vektorer $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n \in \mathcal{V}$. Då är följande ekvivalent

a) $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\} = \mathcal{V}$,

b) $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ är linjärt oberoende,

c) $(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n)$ är en bas i $\mathcal{V}$.

Bevis: Vi visar att a) är ekvivalent med b). Då är båda ekvivalenta med c).

Antag a). Om $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ är linjärt beroende kan vi stryka vektorer tills vi når en linjärt oberoende uppsättning vektorer som spänner upp $\mathcal{V}$ och därmed är en bas med färre än n element, vilket är omöjligt, så b) måste gälla.

Antag b). Om $\operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\} \neq \mathcal{V}$ så finns en vektor $\vec{u}_{n+1} \in \mathcal{V}$ som uppfyller $\vec{u}_{n+1} \notin \operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$. Då är $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n, \vec{u}_{n+1}$ linjärt oberoende vilket inte kan stämma då $n + 1 > n = \dim \mathcal{U}$. Alltså gäller a).

2.6.2 Basbyte

I föregående avsnitt har vi sett att om vi har ett vektorrum $\mathcal{V}$ med en bas $\underline{u}$ kan varje vektor $\vec{x} \in \mathcal{V}$ beskrivas av sin koordinatvektor $\vec{x}_{\underline{u}}$. Om vi nu väljer en annan bas $\underline{v}$ kan $\vec{x}$ också beskrivas med $\vec{x}_{\underline{v}}$. Hur översätter vi mellan koordinater i basen $\underline{u}$ och i basen $\underline{v}$, det vill säga hur får vi $\vec{x}_{\underline{u}}$ från $\vec{x}_{\underline{v}}$ och vice versa? Det ska vi gå igenom i detta avsnitt. Vi börjar med ett exempel i $\mathbb{R}^2$ som vi sett tidigare.

Exempel 2.6.8

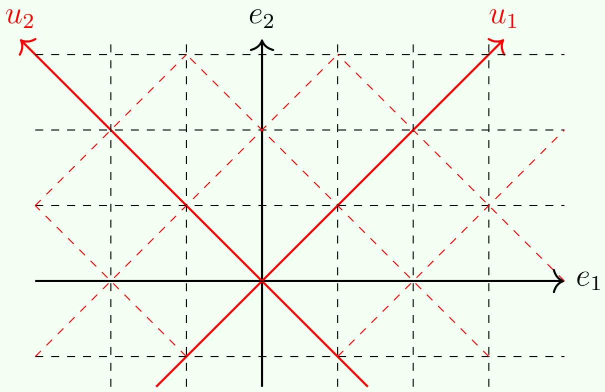
I $\mathbb{R}^2$ låter vi $\underline{e} = (\vec{e}_1, \vec{e}_2)$ vara standardbasen

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

och $\underline{u} = (\vec{u}_1, \vec{u}_2)$ vara basen där

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Vi har då två koordinatsystem



Låt $\vec{v} \in \mathbb{R}^2$ vara en godtycklig vektor, säg med $\underline{u}$-koordinater

$$\vec{v}_{\underline{u}} = \begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow \vec{v} = x\vec{u}_1 + y\vec{u}_2.$$

Om vi tar $\underline{e}$-koordinaterna av vänsterled och högerled i den sista ekvationen så får vi

$$\vec{v}_{\underline{e}} = x(\vec{u}_1)_{\underline{e}} + y(\vec{u}_2)_{\underline{e}} = x\begin{pmatrix} 1 \\ 1 \end{pmatrix} + y\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Vi sätter $A = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ och får

$$\vec{v}_{\underline{e}} = A\vec{v}_{\underline{u}}$$

Notera att kolonnerna i A är precis vektorerna $\vec{u}_1$ och $\vec{u}_2$ (i standardkoordinater).

Om till exempel $\vec{v}_{\underline{u}} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ så får vi

$$\vec{v}_{\underline{e}} = A\vec{v}_{\underline{u}} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Om vi istället vet att $\vec{v}_{\underline{e}} = \begin{pmatrix} 4 \\ 0 \end{pmatrix}$ så får vi att

$$\begin{pmatrix} 4 \\ 0 \end{pmatrix} = A\vec{v}_{\underline{u}} \Leftrightarrow A^{-1} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \vec{v}_{\underline{u}}$$

Inversen till A är $A^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ så vi får

$$\vec{v}_{\underline{u}} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Som i exemplet ovan kan vi alltid byta koordinater med hjälp av en viss matris.

Definition 2.20

Låt $\mathcal{V}$ vara ett vektorrum med två baser $\underline{u} = (\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n)$ och $\underline{v} = (\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n)$. Då är basbytesmatrisen från $\underline{u}$ till $\underline{v}$ den $(n \times n)$-matris $T_{\underline{v}}^{\underline{u}}$ som uppfyller:

$$T_{\underline{v}}^{\underline{u}} \vec{x}_{\underline{u}} = \vec{x}_{\underline{v}}$$

Den fås genom att sätta koordinatvektorerna $(\vec{u}_1)_{\underline{v}}, (\vec{u}_2)_{\underline{v}}, \dots (\vec{u}_n)_{\underline{v}}$ som kolonner.

Vi kollar att formeln stämmer med hur vi definierade kolonnerna i $T_{\underline{v}}^{\underline{u}}$.

Skriv

$$\vec{x}_{\underline{u}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \Leftrightarrow \vec{x} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n.$$

Då är

$$\begin{aligned} \vec{x}_{\underline{v}} &= (x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n)_{\underline{v}} \\ &= (x_1\vec{u}_1)_{\underline{v}} + (x_2\vec{u}_2)_{\underline{v}} + \cdots + (x_n\vec{u}_n)_{\underline{v}} \\ &= x_1(\vec{u}_1)_{\underline{v}} + x_2(\vec{u}_2)_{\underline{v}} + \cdots + x_n(\vec{u}_n)_{\underline{v}} = T_{\underline{v}}^{\underline{u}}\vec{x}_{\underline{v}} \end{aligned}$$

Sats 2.21

Låt $\mathcal{V}$ vara ett vektorrum med baser $\underline{u}, \underline{v}, \underline{w}$. Då gäller följande.

a) Basbytes matrisen $T_{\underline{v}}^{\underline{u}}$ är inverterbar med invers $(T_{\underline{v}}^{\underline{u}})^{-1} = T_{\underline{u}}^{\underline{v}}$.

b) $T_{\underline{w}}^{\underline{v}} T_{\underline{v}}^{\underline{u}} = T_{\underline{w}}^{\underline{u}}$

Bevis: a) vektorerna i basen $\underline{u}$ är linjärt oberoende. Därför är även kolonnerna i $T_{\underline{v}}^{\underline{u}}$ linjärt obereonde. Eftersom $T_{\underline{v}}^{\underline{u}}$ är kvadratisk följer att $T_{\underline{v}}^{\underline{u}}$ är inverterbar. Vidare har vi

$$T_{\underline{v}}^{\underline{u}}\vec{x}_{\underline{u}} = \vec{x}_{\underline{v}} \Leftrightarrow \vec{x}_{\underline{u}} = (T_{\underline{v}}^{\underline{u}})^{-1}\vec{x}_{\underline{v}}.$$

Men det gäller också att $\vec{x}_{\underline{u}} = T_{\underline{u}}^{\underline{v}}\vec{x}_{\underline{v}}$. Eftersom detta gäller för alla vektorer $\vec{x}$ får vi att $(T_{\underline{v}}^{\underline{u}})^{-1} = T_{\underline{u}}^{\underline{v}}$.

b) Vi beräknar $T_{\underline{w}}^{\underline{v}} T_{\underline{v}}^{\underline{u}}\vec{x}_{\underline{u}} = T_{\underline{w}}^{\underline{v}}\vec{x}_{\underline{v}} = \vec{x}_{\underline{w}} = T_{\underline{w}}^{\underline{u}}\vec{x}_{\underline{u}}$. Återigen gäller detta för alla vektorer $\vec{x}$ så vi får att $T_{\underline{w}}^{\underline{v}} T_{\underline{v}}^{\underline{u}} = T_{\underline{w}}^{\underline{u}}$.

Exempel 2.6.9

I $\mathbb{R}^2$ betraktar vi tre baser $\underline{e} = (\vec{e}_1, \vec{e}_2)$, $\underline{u} = (\vec{u}_1, \vec{u}_2)$ och $\underline{v} = (\vec{v}_1, \vec{v}_2)$ där $\underline{e}$ är standardbasen

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

och baserna $\underline{u}$, $\underline{v}$ ges av

$$\vec{u}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Beräkna basbytesmatriserna

a) $T_{\underline{e}}^{\underline{u}}$,

b) $T_{\underline{e}}^{\underline{v}}$,

c) $T_{\underline{v}}^{\underline{u}}$.

a) Kolonnerna i $T_{\underline{e}}^{\underline{u}}$ är $(\vec{u}_1)_{\underline{e}}$ och $(\vec{u}_2)_{\underline{e}}$. Eftersom $\underline{e}$ är standardbasen får vi

$$T_{\underline{e}}^{\underline{u}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

b) På samma sätt som i a) får vi

$$T_{\underline{e}}^{\underline{v}} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

c) Vi kan använda a) och b) för att få fram $T_{\underline{v}}^{\underline{u}} = T_{\underline{v}}^{\underline{e}} T_{\underline{e}}^{\underline{u}}$. Vi beräknar

$$T_{\underline{v}}^{\underline{e}} = (T_{\underline{e}}^{\underline{v}})^{-1} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

så vi får

$$T_{\underline{v}}^{\underline{u}} = T_{\underline{v}}^{\underline{e}} T_{\underline{e}}^{\underline{u}} = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix}.$$

Låt oss dubbelkolla att det stämmer. Kolonnerna i $T_{\underline{v}}^{\underline{u}}$ ska vara $(\vec{u}_1)_{\underline{v}}$ och $(\vec{u}_2)_{\underline{v}}$. Våra beräkningar säger oss att

$$(\vec{u}_1)_{\underline{v}} = \begin{pmatrix} 2 \\ -1 \end{pmatrix} \Leftrightarrow \vec{u}_1 = 2\vec{v}_1 - \vec{v}_2 = 2\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

och

$$(\vec{u}_2)_{\underline{v}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \Leftrightarrow \vec{u}_2 = \vec{v}_1 + 0\vec{v}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

vilket stämmer i båda fall.

2.7 Kolonnrum, radrum och nollrum

I det här avsnittet går vi igenom tre vektorrum som kan associeras till en matris. Vi visar hur en beräkning kan användas för att få fram baser i alla tre vektorrummen.

Definition 2.22

Låt A vara en $(m \times n)$-martis. Då är

1. kolonnrummet av A spannet av kolonnerna i A (alltså ett delrum av $\mathbb{R}^m$),
2. radrummet av A spannet av raderna i A (alltså ett delrum av $\mathbb{R}^n$),
3. nollrummet av A lösningarna till $A\vec{x} = \vec{0}$ (alltså ett delrum av $\mathbb{R}^n$),

Radrummet består av vektorer i $\mathbb{R}^n$ skrivna som rader. Vill vi skriva dem som kolonnvektorer på vanligt sätt så kan vi identifiera radrummet av A med kolonnrummet av $A^\top$.

Vi går nu igenom hur vi kan hitta baser i alla tre rummen med hjälp av samma beräkning. Notera att varje rum har många andra baser. Vår metod är bara ett sätt att hitta några baser. Vi börjar med ett exempel.

Exempel 2.7.1

Låt $A = \begin{pmatrix} 1 & -2 & 1 & 0 \\ 2 & -4 & 3 & 1 \\ -1 & 2 & 5 & 6 \end{pmatrix}$. Vi börjar med att bestämma en bas i nollrummet till A genom att lösa $A\vec{x} = \vec{0}$. Vi får ett ekvationssysstem med totalmatris

$$\left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 0 \\ 2 & -4 & 3 & 1 & 0 \\ -1 & 2 & 5 & 6 & 0 \end{array}\right) \begin{matrix} -2 \;\; 1 \\ \\ \end{matrix} \sim \left(\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 6 & 6 & 0 \end{array}\right) \begin{matrix} \\ -1 \;\; -6 \\ \end{matrix} \sim \left(\begin{array}{cccc|c} 1 & -2 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

Vi har ledande ettor i kolonn 1 och kolonn 3. Alltså sätter vi $x_2 = s$

och $x_4 = t$ vilket ger $x_1 = 2s + t$ och $x_3 = -t$. Alltså är

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 2s+t \\ s \\ -t \\ t \end{pmatrix} = s \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

Vi sätter

$$\vec{u}_1 = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

vilket ger en bas i nollrummet. Notera att parametrisering på detta sätt alltid ger linjärt oberoende vektorer. Sätter vi $s\vec{u}_1 + t\vec{u}_2 = \vec{0}$ får vi genom att se på rad 2 och rad 4 (alltså där parametrarna dyker upp) att $s = 0$ och $t = 0$.

Vi undersöker nu kolonnrummet. Kalla kolonnerna i A för $\vec{v}_1$, $\vec{v}_2$, $\vec{v}_3$ och $\vec{v}_4$. Dessa spänner enligt definition upp kolonnrummet. Men att vi fick parameterlösningar till ekvationen $A\vec{x} = \vec{0}$ säger att de är linjärt beroende. Det betyder att vi kan stryka några av kolonnerna och på så sätt få en bas. Det finns ofta flera sätt att stryka kolonner, men här är ett sätt som alltid fungerar. Vi ser på två speciella lösningar till beroendeekvationen. Först tar vi $s = 1$ och $t = 0$. Då får vi $2\vec{v}_1 + \vec{v}_2 = \vec{0}$. Det betyder att vi kan lösa ut

$$\vec{v}_2 = -2\vec{v}_1.$$

Sedan tar vi $s = 0$ och $t = 1$. Då får vi $\vec{v}_1 - \vec{v}_3 + \vec{v}_4 = \vec{0}$. Det betyder att vi kan lösa ut

$$\vec{v}_4 = -\vec{v}_1 + \vec{v}_3$$

Notera att vi löste ut kolonnerna som motsvarar de obekanta vi satte till parametrar. Eftersom $\vec{v}_2, \vec{v}_4 \in \operatorname{span}\{\vec{v}_1, \vec{v}_3\}$ kan vi stryka både $\vec{v}_2$ och $\vec{v}_4$. Med andra ord är kolonnrummet

$$\operatorname{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4\} = \operatorname{span}\{\vec{v}_1, \vec{v}_3\}$$

Notera att $\vec{v}_1, \vec{v}_3$ är linjärt oberoende eftersom vi tagit bort de kolonner som gav oss parametrar i lösningarna till $A\vec{x} = \vec{0}$. Löser vi samma system igen fast med kolonn 2 och kolonn 4 strukna får vi inga parametrar. Sammanfattningsvis är

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}$$

en bas i kolonnrummet.

Vi undersöker nu radrummet till A. Vi skulle kunna göra en ny beräkning för att tar reda på en bas i kolonnrummet i $A^\top$, men vi kan faktiskt återanvända samma beräkning som vi gjorde ovan. Vi låter B vara den radkanoniska formen av A, det vill säga

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

Vi hävdar att B har samma radrum som A. Om vi gör en radoperation på A och får en matris A' så byter vi en rad i A mot en linjärkombination av rader i A. Dessa kommer alltid att tillhöra radrummet av A, så alla vektorer i radrummet till A' tillhör radrummet i A. Eftersom vi kan ta oss tillbaka till A med en radoperation, så har A och A' precis samma radrum. Radekvivalenta matriser har alltså samma radrum. För matrisen B är det lätt att hitta en bas. De första två första raderna är linjärt oberoende på grund av trappstegsformen. Den sista raden är en nollrad och bidrar inget till spannet. Vi får alltså att

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

Utgör en bas i radrummet. Skrivna som kolonner är dessa vektorer

$$\vec{w}_1 = \begin{pmatrix} 1 \\ -2 \\ 0 \\ -1 \end{pmatrix} \text{ och } \vec{w}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Metoden i exemplet ovan fungerar i allmänhet på följande vis. Låt A vara en $(m \times n)$-martis. Använd Gauss-elimination för att hitta $A \sim B$, där B är radkanonisk.

1. Kolonnerna i A som motsvarar ledande ettor i B ger en bas i kolonnrummet av A.

2. Raderna i B som har ledande ettor ger en bas i radrummet av A.

3. I ekvationen $A\vec{x} = \vec{0}$ kan vi sätta de obekanta som motsvarar kolonner utan ledande ettor i B till parametrar. Från B kan vi lösa ut övriga obekanta och får en bas med en basvektor för varje parameter.

Antalet ledande ettor i B kallas för rangen av A och betecknas $\operatorname{rang}(A)$. Från metoden får vi följande sats.

Sats 2.23

Låt A vara en $(m \times n)$-matris. Beteckna A:s kolonnrum $K(A)$, radrum $R(A)$ och nollrum $N(A)$. Då gäller:

a) $\dim K(A) = \operatorname{rang}(A)$

b) $\dim R(A) = \operatorname{rang}(A)$

c) $\dim N(A) = n - \operatorname{rang}(A)$

Bevis: Metoden som beskrevs ovan ger en bas där antalet element är $\operatorname{rang}(A)$ i fallet av $K(A)$ och $R(A)$. För $N(A)$ är dimensionen antalet parametrar. Eftersom vi får en parameter för varje kolonn utan ledande etta har vi $n - \operatorname{rang}(A)$ parametrar.

En konsekvens av denna sats är att rad- och kolonnoperationer inte påverkar rangen eftersom de inte påverkar radrummet respektive kolonnrummet och därför inte heller deras dimensioner som båda är lika med rangen.

3 Linjära avbildningar

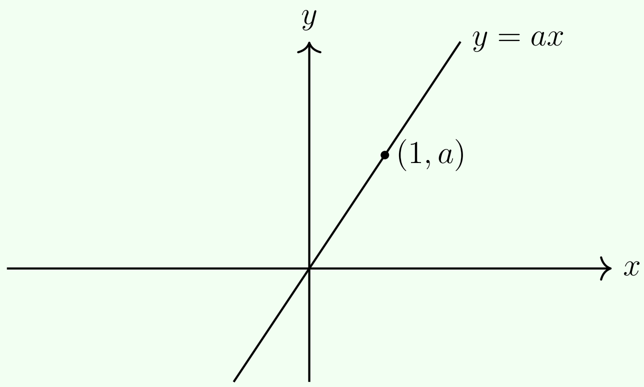
I det här avsnittet ska vi gå igenom linjära avbildningar. Kort sagt är linjära avbildningar funktioner mellan vektorrum som är kompatibla med de två grundläggande vektorrumsoperationerna: addition och skalning. Många viktiga funktioner visar sig vara linjära och allt vi går igenom i det här avsnittet kan användas på dem. Man kan också använda linjära avbildningar för att koppla ihop två olika vektorrum och på så sätt jämföra dem med varandra.

3.1 Definition och egenskaper

Vi börjar med ett exempel.

Exempel 3.1.1

Låt $a \in \mathbb{R}$ och $F : \mathbb{R} \to \mathbb{R}$ vara funktionen som ges av $F(x) = ax$. Grafen $y = F(x)$ beskriver en linje genom origo.



Om vi tar två tal $x_1, x_2 \in \mathbb{R}$ och utvärderar F på deras summa får vi

$$F(x_1 + x_2) = a(x_1 + x_2) = ax_1 + ax_2 = F(x_1) + F(x_2).$$

Om vi skalar om ett tal x med en faktor λ och utvärderar F på resultatet så får vi

$$F(\lambda x) = a\lambda x = \lambda a x = \lambda F(x).$$

Dessa två egenskaper gör det särskilt enkelt att räkna med funktionen F. Funktioner mellan vektorrum som uppfyller liknande räkneregler kallas linjära. Notera att det är viktigt att linjen går genom origo. Om vi istället skulle betrakta $F(x) = ax + b$, där $b \neq 0$ så skulle grafen fortfarande beskriva en linje, men inte uppfylla reglerna. Vi skulle då inte kalla F linjär trots att grafen är en linje.

Definition 3.1

Låt $\mathcal{V}$ och $\mathcal{W}$ vara vektorrum. En funktion $F : \mathcal{V} \to \mathcal{W}$ kallas en linjär avbildning om den uppfyller

1. $F(\vec{v} + \vec{w}) = F(\vec{v}) + F(\vec{w})$ för alla $\vec{v}, \vec{w} \in \mathcal{V}$,
2. $F(\lambda\vec{v}) = \lambda F(\vec{v})$ för alla $\lambda \in \mathbb{R}$ och $\vec{v} \in \mathcal{V}$.

Ibland används benämningarna linjär funktion, linjär transformation, linjär operator för samma sak. Kärt barn har många namn.

I exemplet ovan såg vi att $F : \mathbb{R} \to \mathbb{R}$, där $F(x) = ax$ är en linjär avbildning. Vi kan generalisera exemplet genom att ersätta a med en matris.

Exempel 3.1.2

Låt A vara en $(m \times n)$-matris och låt $F : \mathbb{R}^n \to \mathbb{R}^m$ vara funktionen som ges av $F(\vec{x}) = A\vec{x}$. Notera att A har m rader och n kolonner. Alltså ska $\vec{x}$ tolkas som en kolonn med n element och $A\vec{x}$ är då en kolonn med m element. Reglerna för matrisräkning ger att

$$F(\vec{x}_1 + \vec{x}_2) = A(\vec{x}_1 + \vec{x}_2) = A\vec{x}_1 + A\vec{x}_2 = F(\vec{x}_1) + F(\vec{x}_2)$$

och

$$F(\lambda\vec{x}) = A(\lambda\vec{x}) = \lambda(A\vec{x}) = \lambda F(\vec{x}).$$

Alltså är F en linjär avbildning.

För att göra exemplet konkret ser vi på $F : \mathbb{R}^2 \to \mathbb{R}^3$ där

$$F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ 2x_1 - x_2 \\ -x_1 + 3x_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

I det här fallet är alltså

$$A = \begin{pmatrix} 1 & 1 \\ 2 & -1 \\ -1 & 3 \end{pmatrix}.$$

Vi räknar ut tre värden av funktionen F

$$F\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\1\\2\end{pmatrix}, \quad F\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}1\\-1\\3\end{pmatrix}, \quad F\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}3\\0\\5\end{pmatrix}.$$

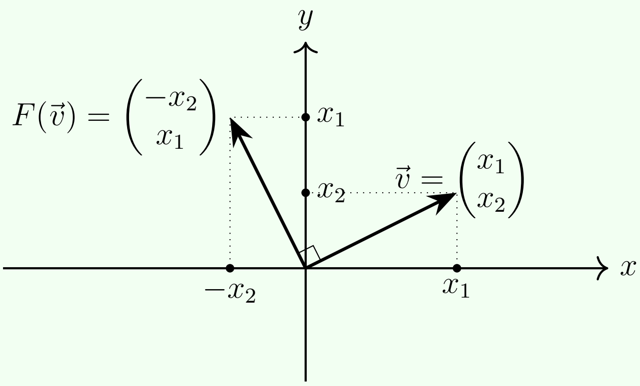
Alltså är

$$\begin{aligned} F\left(\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}0\\1\end{pmatrix}\right) &= F\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}3\\0\\5\end{pmatrix} = \begin{pmatrix}2\\1\\2\end{pmatrix} + \begin{pmatrix}1\\-1\\3\end{pmatrix} \\ &= F\begin{pmatrix}1\\1\end{pmatrix} + F\begin{pmatrix}0\\1\end{pmatrix}. \end{aligned}$$

Vi tar ett till exempel av geometriskt ursprung.

Exempel 3.1.3

Låt $F : \mathbb{R}^2 \to \mathbb{R}^2$ vara funktionen som roterar varje vektor $\vec{v} \in \mathbb{R}^2$ moturs med vinkeln $\frac{\pi}{2}$.



Vi observerar att F ges av en matris

$$F\begin{pmatrix}x_1\\x_2\end{pmatrix} = \begin{pmatrix}-x_2\\x_1\end{pmatrix} = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}.$$

Alltså är F en linjär avbildning. Vi kan också förstå varför geometriskt. Om vi adderar två vektorer och roterar får vi samma sak som om vi först roterar varje vektor och sedan adderar. På liknande sätt får vi samma resultat om vi skalar om en vektor och sedan roterar som om vi först roterar och sedan skalar om.

Som i exemplen visar det sig att vi alltid kan beskriva linjära avbildningar från $\mathbb{R}^n$ till $\mathbb{R}^m$ med hjälp av matriser.

Sats 3.2

Om $F : \mathbb{R}^n \to \mathbb{R}^m$ är en linjär avbildning, så finns en $(m \times n)$-matris A så att $F(\vec{x}) = Ax$. Närmare bestämt har A vektorerna $F(\vec{e}_1), F(\vec{e}_2), \ldots, F(\vec{e}_n)$ som kolonner.

Vi kallar matrisen A i satsen för F:s matris i standardbasen. Vi ska senare se hur vi kan beskriva linjära avbildningar med avseende på andra baser.

Bevis: Skriv

$$\begin{aligned}\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} &= \begin{pmatrix} x_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ x_2 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ x_n \end{pmatrix} \\ &= x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \cdots + x_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \\ &= x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n.\end{aligned}$$

Då är

$$\begin{aligned}F(\vec{x}) &= F(x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n) \\ &= F(x_1\vec{e}_1) + F(x_2\vec{e}_2) + \cdots + F(x_n\vec{e}_n) \\ &= x_1F(\vec{e}_1) + x_2F(\vec{e}_2) + \cdots + x_nF(\vec{e}_n) \\ &= A\vec{x}.\end{aligned}$$

Exempel 3.1.4

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara den linjära avbildningen som ges av

$$F\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_1 - x_2 \end{pmatrix}.$$

Hitta F:s matris i standardbasen.

Vi avbildar standardbasvektorerna i $\mathbb{R}^3$

$$\begin{aligned}
F(\vec{e}_1) &= F\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \\
F(\vec{e}_2) &= F\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \\
F(\vec{e}_3) &= F\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.
\end{aligned}$$

Alltså är F:s matris i standardbasen

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

Vi kan kontrollera att matrisen stämmer

$$A\vec{x} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ x_1 - x_2 \end{pmatrix}.$$

Vi undersöker nu rotation med godtycklig vinkel.

Exempel 3.1.5

Låt $F : \mathbb{R}^2 \to \mathbb{R}^2$ vara funktionen som roterar varje vektor $\vec{v} \in \mathbb{R}^2$ moturs med vinkeln α. Bestäms F:s matris i standardbasen.

Den första basvektorn $\vec{e}_1$ har längd 1 och pekar i x-axelns riktning. Roterar vi $\vec{e}_1$ moturs med vinkeln α får vi en vektor $F(\vec{e}_1)$ som också har längd 1. Med andra ord pekar $F(\vec{e}_1)$ på enhetscirkeln. Dessutom är vinkeln mellan x-axeln och $F(\vec{e}_1)$ precis α. Alltså är

$$F(\vec{e}_1) = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}.$$

Den andra basvektorn $\vec{e}_2$ har längd 1 och pekar i y-axelns riktning. Vinkeln mellan $\vec{e}_2$ och x-axeln är alltså $\frac{\pi}{2}$. Efter att vi roterar $\vec{e}_2$ med vinkeln α får vi då vinkeln $\alpha + \frac{\pi}{2}$. Det ger oss

$$F(\vec{e}_2) = \begin{pmatrix} \cos(\alpha + \frac{\pi}{2}) \\ \sin(\alpha + \frac{\pi}{2}) \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}.$$

Alltså är F:s matris i standardbasen

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Om vi till exempel väljer $\alpha = \frac{\pi}{2}$, så är $\cos\alpha = 0$ och $\sin\alpha = 1$, vilket ger matrisen

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

som vi hittade tidigare.

Det finns många funktioner från $\mathbb{R}^n$ till $\mathbb{R}^m$ som inte är linjära. Dessa kan inte beskrivas med matriser. Vi tar några exempel på sådana funktioner nedan.

Exempel 3.1.6

Följande funktioner är inte linjära. För att se det krävs bara ett motexempel mot de två egenskaperna i definitionen. Efter varje funktion ges ett sådant motexempel.

a) $F : \mathbb{R} \to \mathbb{R}$, $F(x) = x^2$. Då är $F(2 \cdot 1) = 4$, men $2F(1) = 2$.

b) $F : \mathbb{R} \to \mathbb{R}$, $F(x) = 2x + 1$. Då är $F(0 \cdot 1) = 1$, men $0F(1) = 0$.

c) $F : \mathbb{R} \to \mathbb{R}$, $F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 + x_2 \\ x_1 x_2 \end{pmatrix}$. Då är $F\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \end{pmatrix}$, men $2F\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$.

Från definitionen följer många viktiga egenskaper för linjära avbildningar. Vi tar nu upp två grundläggande egenskaper.

Sats 3.3

Om $F : \mathcal{V} \to \mathcal{W}$ är en linjär avbildning så gäller

a) $F(-\vec{v}) = -F(\vec{v})$,

b) $F(\vec{0}) = \vec{0}$.

Bevis: Vi beräknar

$$F(-\vec{v}) = F((-1)\vec{v}) = (-1)F(\vec{v}) = -F(\vec{v})$$

och

$$F(\vec{0}) = F(0 \cdot \vec{0}) = 0F(\vec{0}) = -\vec{0}.$$

Linjära avbildningar dyker upp i många olika sammanhang. Här kommer ett exempel från analys.

Exempel 3.1.7

Låt

$$\mathcal{V} = \mathcal{C}^1(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ deriverbar med } f' \text{ kontinuerlig}\}.$$

och

$$\mathcal{W} = \mathcal{C}(M) = \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ kontinuerlig}\}.$$

Definiera $F : \mathcal{V} \to \mathcal{W}$ genom $F(f) = f'$.

Då är F en linjär avbildning. Om till exempel $f(x) = e^{2x}$ och $g(x) = x^3$ så är $(f+g)(x) = e^{2x} + x^3$ och

$$f'(x) = 2e^{2x}, \qquad g'(x) = 3x^2, \quad (f+g)'(x) = 2e^{2x} + 3x^2.$$

Mer allmänt gäller

$$F(f+g) = (f+g)' = f' + g' = F(f) + F(g)$$

och

$$F(\lambda f) = (\lambda f)' = \lambda f' = \lambda F(f).$$

Kom ihåg att en funktion $f : X \to Y$ kallas inverterbar om det finns en funktion $f^{-1} : Y \to X$ som uppfyller att $f(x) = y$ om och endast om $x = f^{-1}(y)$ för alla $x \in X$ och $y \in Y$. Funktionen f^{-1} kallas inversen till f. Om en linjär avbildning är inverterbar så är även dess invers linjär.

Sats 3.4

Om $F : \mathcal{V} \to \mathcal{W}$ är en inverterbar linjär avbildning så är $F^{-1} : \mathcal{W} \to \mathcal{V}$ också linjär.

Bevis: Att F är linjär ger

$$F(F^{-1}(\vec{v}) + F^{-1}(\vec{w})) = F(F^{-1}(\vec{v})) + F(F^{-1}(\vec{w})) = \vec{v} + \vec{w}.$$

Tar vi F^{-1} av båda led får vi

$$F^{-1}(\vec{v}) + F^{-1}(\vec{w}) = F^{-1}(\vec{v} + \vec{w}).$$

På liknade sätt ger

$$F(\lambda F^{-1}(\vec{v})) = \lambda F(F^{-1}(\vec{v})) = \lambda \vec{v}$$

att

$$\lambda F^{-1}(\vec{v}) = F^{-1}(\lambda \vec{v}).$$

Vi illustrerar satsen genom att åter se på rotationer.

Exempel 3.1.8

Låt $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ vara funktionen som roterar varje vektor $\vec{v} \in \mathbb{R}^2$ moturs med vinkeln α. Vi såg tidigare att matrisen till R_α är

$$\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

Om vi roterar med vinkeln α kan vi återskapa vår ursprungliga vektor genom att rotera med vinkeln $-\alpha$. Alltså är R_α inverterbar med invers $R_{-\alpha}$. Matrisen till $R_{-\alpha}$ är

$$\begin{pmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{pmatrix} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}.$$

Notera att matrisen till $R_{-\alpha}$ är precis inversen av matrisen till R_α:

$$\begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}\begin{pmatrix}\cos\alpha & \sin\alpha\\ -\sin\alpha & \cos\alpha\end{pmatrix} =$$
$$\begin{pmatrix}\cos^2\alpha+\sin^2\alpha & \cos\alpha\sin\alpha-\sin\alpha\cos\alpha\\ \sin\alpha\cos\alpha-\cos\alpha\sin\alpha & \sin^2\alpha+\cos^2\alpha\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

I allmänhet är matrisen av inversen till en avbildning just inversen av avbildnings matris. Vi kommer se det i nästa avsnitt.

Kom ihåg att sammansättningen av en funktion $f : X \to Y$ och en funktion $g : Y \to Z$ är funktionen $g \circ f : X \to Y$ som ges av $(g \circ f)(x) = g(f(x))$. Sammansättningen av två linjära avbildningar är en linjär avbildning

Sats 3.5

Om $F : \mathcal{U} \to \mathcal{V}$ och $G : \mathcal{V} \to \mathcal{W}$ är linjära avbildningar så är $G \circ F : \mathcal{U} \to \mathcal{W}$ också linjär.

Bevis: Vi beräknar

$$G(F(\vec{v}+\vec{w})) = G(F(\vec{v})+F(\vec{w})) = G(F(\vec{v}))+G(F(\vec{w}))$$

och

$$G(F(\lambda\vec{v})) = G(\lambda F(\vec{v})) = \lambda G(F(\vec{v})).$$

Återigen illustrerar vi satsen genom att se på rotationer.

Exempel 3.1.9

Låt $R_\alpha : \mathbb{R}^2 \to \mathbb{R}^2$ vara funktionen som roterar varje vektor $\vec{v} \in \mathbb{R}^2$ moturs med vinkeln α. Om vi sammansätter R_α med R_β så får vi en funktion som först roterar med α och sedan β vilket totalt ger rotation med $\beta+\alpha$. Alltså är $R_\beta \circ R_\alpha = R_{\beta+\alpha}$. Notera att matriserna för dessa avbildningar är

$$\begin{pmatrix}\cos\beta & -\sin\beta\\ \sin\beta & \cos\beta\end{pmatrix}, \begin{pmatrix}\cos\alpha & -\sin\alpha\\ \sin\alpha & \cos\alpha\end{pmatrix}, \begin{pmatrix}\cos(\beta+\alpha) & -\sin(\beta+\alpha)\\ \sin(\beta+\alpha) & \cos(\beta+\alpha)\end{pmatrix}.$$

Multiplicerar vi de två första matriserna så får vi den tredje:

$$\begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} =$$
$$\begin{pmatrix} \cos\beta\cos\alpha - \sin\beta\sin\alpha & -\cos\beta\sin\alpha - \sin\beta\cos\alpha \\ \sin\beta\cos\alpha + \cos\beta\sin\alpha & -\sin\beta\sin\alpha + \cos\beta\cos\alpha \end{pmatrix} =$$
$$\begin{pmatrix} \cos(\beta+\alpha) & -\sin(\beta+\alpha) \\ \sin(\beta+\alpha) & \cos(\beta+\alpha) \end{pmatrix}.$$

I allmänhet är matrisen av sammansättningen av två avbildningar just produkten av avbildningarnas matriser. Precis som med inversen kommer vi bevisa detta i nästa avsnitt.

3.2 Matrisen till en linjär avbildning

I detta avsnitt går vi igenom hur linjära avbildningar kan beskrivas med hjälp av matriser mer allmänt än tidigare.

Vi börjar med att undersöka en speciell sorts linjära avbildningar som kallas koordinatavbildningar. Låt $\mathcal{V}$ vara ett vektorrum och $\vec{u}_1, \vec{u}_1, \dots, \vec{u}_n \in \mathcal{V}$. Då får vi en linjär avbildning

$$F : \mathbb{R}^n \to \mathcal{V}, \quad F\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n.$$

Vi kollar att F är linjär:

$$\begin{aligned} F(\vec{x}+\vec{y}) = F\begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix} &= (x_1+y_1)\vec{u}_1 + (x_2+y_2)\vec{u}_2 + \cdots + (x_n+y_n)\vec{u}_n \\ &= x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n \\ &\quad + y_1\vec{u}_1 + y_2\vec{u}_2 + \cdots + y_n\vec{u}_n \\ &= F(\vec{x}) + F(\vec{y}), \end{aligned}$$

$$
\begin{aligned}
F(\lambda\vec{x}) = F\begin{pmatrix}\lambda x_1\\ \lambda x_2\\ \vdots\\ \lambda x_n\end{pmatrix} &= \lambda x_1\vec{u}_1 + \lambda x_2\vec{u}_2 + \cdots + \lambda x_n\vec{u}_n \\
&= \lambda(x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n) \\
&= \lambda F(\vec{x}).
\end{aligned}
$$

Observera att om $\underline{u} = (\vec{u}_1, \vec{u}_1, \ldots, \vec{u}_n)$ är en bas i $\mathcal{V}$ så är

$$F(\vec{x}) = \vec{v} \Leftrightarrow x_1\vec{u}_1 + x_2\vec{u}_2 + \cdots + x_n\vec{u}_n = \vec{v} \Leftrightarrow \vec{x} = \vec{v}_{\underline{u}}.$$

Alltså är F inverterbar i detta fall och $F^{-1}(\vec{v}) = \vec{v}_{\underline{u}}$. Eftersom F är linjär är även F^{-1} linjär. Det vill säga

$$(\vec{v}+\vec{w})_{\underline{u}} = \vec{v}_{\underline{u}} + \vec{w}_{\underline{u}}, \quad (\lambda\vec{v})_{\underline{u}} = \lambda\vec{v}_{\underline{u}}.$$

Notera att detta ger ett nytt bevis för Sats 5.3.

Avbildningen $F^{-1} : \mathcal{V} \to \mathbb{R}^n$, $\vec{v} \mapsto \vec{v}_{\underline{u}}$ kallas för koordinatavbildningen till basen $\underline{u}$. Att den är inverterbar säger precis att vi kan översätta mellan vektorer i $\mathcal{V}$ och koordinater i $\mathbb{R}^n$ utan att förlora någon information.

Exempel 3.2.1

Låt $\mathcal{V} = \mathcal{P}_3$ och $\underline{u}$ vara basen i $\mathcal{V}$ som består av

$$u_0(t) = 1, \quad u_1(t) = t, \quad u_2(t) = t^2, \quad u_3(t) = t^3.$$

Då ges avbildningen $F : \mathbb{R}^4 \to \mathcal{P}_3$ som vi definierade ovan av

$$F\begin{pmatrix}x_0\\ x_1\\ x_2\\ x_3\end{pmatrix} = x_0 + x_1 t + x_2 t^2 + x_3 t^3.$$

Dess invers är koordinatavbildningen $F^{-1} : \mathcal{P}_3 \to \mathbb{R}^4$ som ges av

$$F(a_0 + a_1 t + a_2 t^2 + a_3 t^3) = \begin{pmatrix}a_0\\ a_1\\ a_2\\ a_3\end{pmatrix}.$$

Funktionerna F och F^{-1} kan alltså användas för att översätta mellan $\mathcal{P}_3$ och $\mathbb{R}^4$.

Vi ska nu använda koordinatavbildningar för att definiera matrisen till en linjär avbildning med avseende på ett val av baser. Låt $\mathcal{V}$ och $\mathcal{W}$ vara vektorrum med baser $\underline{v} = (\vec{v}_1, \vec{v}_1, \dots, \vec{v}_n)$ i $\mathcal{V}$ och $\underline{w} = (\vec{w}_1, \vec{w}_1, \dots, \vec{w}_m)$ i $\mathcal{W}$. Då har vi två koordinatavbildningar $\mathcal{V} \to \mathbb{R}^n$, $\vec{x} \mapsto \vec{x}_{\underline{v}}$ och $\mathcal{W} \to \mathbb{R}^m$, $\vec{y} \mapsto \vec{y}_{\underline{w}}$. Låt nu $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning. Betrakta följande diagram av avbildningar:

$$\begin{tikzcd}
\mathcal{V} \arrow[r, "F"] \arrow[d] & \mathcal{W} \arrow[d] \\
\mathbb{R}^n \arrow[r] & \mathbb{R}^m
\end{tikzcd}$$

där de vertikala pilarna motsvarar koordinatavbildningarna och avbildningen $\mathbb{R}^n \to \mathbb{R}^m$ ges av att först ta inversen till koordinatavbildningen $\mathbb{R}^n \to \mathcal{V}$, sedan tillämpa $F : \mathcal{V} \to \mathcal{W}$ och till sist ta koordinatavbildningen $\mathcal{W} \to \mathbb{R}^m$. Eftersom resultatet är en linjär avbildning $\mathbb{R}^n \to \mathbb{R}^m$ så ges den av en $(m \times n)$-matris. Vi kallar denna för matrisen till F med avseende på baserna $\underline{v}$ i $\mathcal{V}$ och $\underline{w}$ i $\mathcal{W}$ och betecknar den med $(F)^{\underline{v}}_{\underline{w}}$. Matrisen uppfyller $(F)^{\underline{v}}_{\underline{w}} \vec{x}_{\underline{v}} = F(\vec{x})_{\underline{w}}$.

Definition 3.6

Givet

- ett vektorrum $\mathcal{V}$ med en $\underline{v} = (\vec{v}_1, \vec{v}_1, \dots, \vec{v}_n)$,
- ett vektorrum $\mathcal{W}$ med en bas $\underline{w} = (\vec{w}_1, \vec{w}_1, \dots, \vec{w}_m)$,
- en linjär avbildning $F : \mathcal{V} \to \mathcal{W}$,

så är matrisen till en linjär avbildning F med avseende på baserna $\underline{v}$ i $\mathcal{V}$ och $\underline{w}$ i $\mathcal{W}$ är den $(m \times n)$-matris $(F)^{\underline{v}}_{\underline{w}}$ som uppfyller

$$(F)^{\underline{v}}_{\underline{w}} \vec{x}_{\underline{v}} = F(\vec{x})_{\underline{w}}.$$

Kolonnerna i $(F)^{\underline{v}}_{\underline{w}}$ är $F(\vec{v}_1)_{\underline{w}}, F(\vec{v}_2)_{\underline{w}}, \dots, F(\vec{v}_n)_{\underline{w}}$.

Låt oss verifiera att $(F)^{\underline{v}}_{\underline{w}} = A$ där A har kolonnerna $F(\vec{v}_1)_{\underline{w}}, F(\vec{v}_2)_{\underline{w}}, \dots, F(\vec{v}_n)_{\underline{w}}$.

Skriv

$$\vec{x}_{\underline{v}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Vi beräknar högerledet i ekvationen $(F)_{\underline{w}}^{\underline{v}} \vec{x}_{\underline{v}} = F(\vec{x})_{\underline{w}}$:

$$\begin{aligned} F(\vec{x})_{\underline{w}} &= F(x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_n\vec{v}_n)_{\underline{w}} \\ &= (x_1F(\vec{v}_1) + x_2F(\vec{v}_2) + \cdots + x_nF(\vec{v}_n))_{\underline{w}} \\ &= x_1F(\vec{v}_1)_{\underline{w}} + x_2F(\vec{v}_2)_{\underline{w}} + \cdots + x_nF(\vec{v}_n)_{\underline{w}} = A\vec{x}_{\underline{v}}. \end{aligned}$$

Exempel 3.2.2

Låt $F : \mathcal{P}_3 \to \mathcal{P}_2$, där $F(p) = p'$. Välj baser $\underline{v} = (1, t, t^2, t^3)$ i $\mathcal{P}_3$ och $\underline{w} = (1, t, t^2)$ i $\mathcal{P}_2$. Då är

$$F(1) = 0, \quad F(t) = 1, \quad F(t^2) = 2t, \quad F(t^3) = 3t^2$$

och

$$F(1)_{\underline{w}} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \; F(t)_{\underline{w}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; F(t^2)_{\underline{w}} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \; F(t^3)_{\underline{w}} = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}.$$

Alltså är

$$(F)_{\underline{w}}^{\underline{v}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Om vi till exempel väljer $p(t) = 4 + 2t - t^2 + t^3$ så är

$$p_{\underline{v}} = \begin{pmatrix} 4 \\ 2 \\ -1 \\ 1 \end{pmatrix} \text{ och } (F)_{\underline{w}}^{\underline{v}} p_{\underline{v}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ 2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}.$$

Å andra kan vi räkna ut koordinaterna av $F(p) = p'$. Eftersom $p'(t) = 2 - 2t + 3t^2$ får vi att

$$F(p)_{\underline{w}} = \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}.$$

Med andra ord har vi kollat att $(F)_{\underline{w}}^{\underline{v}} p_{\underline{v}} = F(p)_{\underline{w}}$ verkligen stämmer.

Vi visar nu att matrisen av en sammansättning av avbildningar ges av produkten av avbildningarnas matriser.

Sats 3.7

Låt $F : \mathcal{U} \to \mathcal{V}$ och $G : \mathcal{V} \to \mathcal{W}$ vara linjära avbildningar. Givet baser $\underline{u}$ i $\mathcal{U}$, $\underline{v}$ i $\mathcal{V}$ och $\underline{w}$ i $\mathcal{W}$ så är

$$(G \circ F)^{\underline{u}}_{\underline{w}} = (G)^{\underline{v}}_{\underline{w}}(F)^{\underline{u}}_{\underline{v}}.$$

Bevis: Vi beräknar

$$(G)^{\underline{v}}_{\underline{w}}(F)^{\underline{u}}_{\underline{v}}\vec{x}_{\underline{u}} = (G)^{\underline{v}}_{\underline{w}}F(\vec{x})_{\underline{v}} = G(F(\vec{x}))_{\underline{w}} = (G \circ F)^{\underline{u}}_{\underline{w}}\vec{x}_{\underline{u}}.$$

Denna likhet gäller för alla $\vec{x} \in \mathcal{U}$ och därmed för alla kolonnvektorer $\vec{x}_{\underline{u}}$. Det följer att $(G)^{\underline{v}}_{\underline{w}}(F)^{\underline{u}}_{\underline{v}} = (G \circ F)^{\underline{u}}_{\underline{w}}$.

Exempel 3.2.3

Låt $F : \mathcal{P}_3 \to \mathcal{P}_2$, $F(p) = p'$ och $G : \mathcal{P}_2 \to \mathcal{P}_3$, $G(q) = tq(t)$. Välj baser $\underline{v} = (1, t, t^2, t^3)$ i $\mathcal{P}_3$ och $\underline{w} = (1, t, t^2)$ i $\mathcal{P}_2$. Som beräknat tidigare är

$$(F)^{\underline{v}}_{\underline{w}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Vi beräknar nu $(G)^{\underline{w}}_{\underline{v}}$. Eftersom

$$G(1) = t, \quad G(t) = t^2, \quad G(t^2) = t^3,$$

och

$$G(1)_{\underline{w}} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \ G(t)_{\underline{w}} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \ G(t^2)_{\underline{w}} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

så är

$$(G)^{\underline{w}}_{\underline{v}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Enligt satsen är

$$(G\circ F)_{\underline{v}}^{\underline{v}} = (G)_{\underline{v}}^{\underline{w}}(F)_{\underline{w}}^{\underline{v}} = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

Alltså ger kolonnerna i den sista matrisen precis koordinaterna för $G \circ F$ utvärderat på polynomen i basen $\underline{v}$. Vi kollar att det stämmer:

$$G(F(1)) = G(0) = 0, \quad G(F(t)) = G(1) = t,$$

$$G(F(t^2)) = G(2t) = 2t^2, \quad G(F(t^3)) = G(3t^2) = 3t^3.$$

Vi tar ett liknande exempel.

Exempel 3.2.4

Låt $F : \mathcal{P}_2 \to \mathcal{P}_2$, $F(p) = p'$. Välj basen $\underline{v} = (1, t, t^2)$ i $\mathcal{P}_2$. Vi räknar ut matrisen till F i basen $\underline{v}$ som tidigare. Eftersom vi valt att definiera F från ett vektorrum till samma vektorrum kan vi använda samma koordinater för både input och output.

$$(F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

I det här fallet kan vi sammansätta F med sig själv och får

$$(F \circ F)_{\underline{v}}^{\underline{v}} = (F)_{\underline{v}}^{\underline{v}}(F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Notera att $(F\circ F)(p) = F(F(p)) = F(p') = p''$. Alltså är $(F\circ F)(1) = 0$, $(F \circ F)(t) = 0$ och $(F \circ F)(t^2) = 2$, vilket stämmer med matrisen vi beräknade ovan.

För en inverterbar linjär avbildning gäller att matrisen av dess invers är inversen av dess matris, vilket vi nu visar.

Sats 3.8

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning, $\underline{v}$ en bas i $\mathcal{V}$ och $\underline{w}$ en bas i $\mathcal{W}$. Då är F inverterbar om och endast om matrisen $(F)^{\underline{v}}_{\underline{w}}$ är inverterbar. I så fall är $\dim V = \dim W$ och

$$(F^{-1})^{\underline{w}}_{\underline{v}} = ((F)^{\underline{v}}_{\underline{w}})^{-1}.$$

Bevis: Om F är inverterbar så är $(F^{-1} \circ F)(\vec{x}) = \vec{x}$ för alla $\vec{x} \in \mathcal{V}$. Alltså är

$$\vec{x}_{\underline{v}} = (F^{-1} \circ F)(\vec{x})_{\underline{v}} = (F^{-1} \circ F)^{\underline{v}}_{\underline{v}}\vec{x}_{\underline{v}} = (F^{-1})^{\underline{w}}_{\underline{v}}(F)^{\underline{v}}_{\underline{w}}\vec{x}_{\underline{v}}.$$

Eftersom likheten gäller för alla kolonner $\vec{x}_{\underline{v}}$ får vi att $(F^{-1})^{\underline{w}}_{\underline{v}}(F)^{\underline{v}}_{\underline{w}}$ är enhetsmatrisen. På liknande sätt får vi genom att undersöka $(F \circ F^{-1})$ att $(F)^{\underline{v}}_{\underline{w}}(F^{-1})^{\underline{w}}_{\underline{v}}$ också är enhetsmatrisen. Alltså är $(F)^{\underline{v}}_{\underline{w}}$ är inverterbar med invers $(F^{-1})^{\underline{w}}_{\underline{v}}$. Då måste $(F)^{\underline{v}}_{\underline{w}}$ vara en kvadratisk matris och $\dim V = \dim W$.

Antag nu att $(F)^{\underline{v}}_{\underline{w}}$ är inverterbar med invers A. Koordinatavbildningar är inverterbara så det finns en linjär avbildning $G : \mathcal{W} \to \mathcal{V}$ som uppfyller $(G)^{\underline{w}}_{\underline{v}} = A$. Eftersom A är inversen till $(F)^{\underline{v}}_{\underline{w}}$ får vi att $(G \circ F)^{\underline{v}}_{\underline{v}} = (G)^{\underline{w}}_{\underline{v}}(F)^{\underline{v}}_{\underline{w}}$ är enhetsmatrisen. På liknande sätt får vi att $(F \circ G)^{\underline{w}}_{\underline{w}}$ också är enhetsmatrisen. Men då följer $(G \circ F)(\vec{x}) = \vec{x}$ för alla $\vec{x} \in \mathcal{V}$ och $(F \circ G)(\vec{y}) = \vec{y}$ för alla $\vec{y} \in \mathcal{W}$. Alltså är F inverterbar med $F^{-1} = G$.

Exempel 3.2.5

Låt $F : \mathcal{P}_2 \to \mathcal{P}_2$, $F(p) = p(t+1)$. Då är F inverterbar med $F^{-1}(q) = q(t-1)$. Mycket riktigt är $F^{-1}(F(p)) = F^{-1}(p(t+1)) = p(t-1+1) = p(t)$ och $F(F^{-1}(q)) = F(q(t-1)) = q(t+1-1) = q(t)$. Vi beräknar matriserna till F och G med avseende på basen $\underline{v} = (1, t, t^2)$.

$$F(1) = 1, \quad F(t) = t+1, \quad F(t^2) = (t+1)^2 = t^2 + 2t + 1,$$

ger

$$F(1)_{\underline{v}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; F(t)_{\underline{v}} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \; F(t^2)_{\underline{v}} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}.$$

Alltså är

$$(F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}.$$

På likande sätt får vi matrisen till inversen.

$$F^{-1}(1) = 1, \quad F^{-1}(t) = t - 1, \quad F^{-1}(t^2) = (t-1)^2 = t^2 - 2t + 1,$$

ger

$$F^{-1}(1)_{\underline{v}} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \; F^{-1}(t)_{\underline{v}} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \; F^{-1}(t^2)_{\underline{v}} = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.$$

Alltså är

$$(F^{-1})_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Vi kollar att dessa matriser är varandras inverser

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

Det stämmer!

3.3 Basbyte för linjära avbildningar

Som vi såg i föregående avsnitt kan vi beskriva linjära avbildningar med hjälp av matriser genom att välja baser i vektorrummen som är inblandade. Vilken matris vi får beror på vilka baser vi väljer. Vi kan alltså få olika matriser för samma avbildning genom att välja olika baser. Hur jämför vi dessa matriser? De visar sig vara relaterade via basbytesmatriser. För att förstå hur är det bra att komma ihåg hur en matris A ger upphov till en linjär avbildning F. Sambandet är $F(\vec{x}) = A\vec{x}$. Alltså multiplicerar vi A med $\vec{x}$ från höger och går från $\vec{x}$ till $A\vec{x}$. Om vi därefter vill ändra på $A\vec{x}$ kan vi multiplicera med en annan matris B och får då $BA\vec{x}$. Notera att det är lättast att förstå formlerna om vi läser från höger till vänster. Först kommer $\vec{x}$ som blir till $A\vec{x}$ och sedan till $BA\vec{x}$. Med detta i åtanke är följande sats rimlig.

Sats 3.9

Givet

- ett vektorrum $\mathcal{V}$ med baser $\underline{e}$ och $\underline{v}$,
- ett vektorrum $\mathcal{W}$ med baser $\underline{f}$ och $\underline{w}$,
- en linjär avbildning $F : \mathcal{V} \to \mathcal{W}$,

så gäller $(F)^{\underline{e}}_{\underline{f}} = T^{\underline{w}}_{\underline{f}}(F)^{\underline{v}}_{\underline{w}}T^{\underline{e}}_{\underline{v}}$

Om vi läser högerledet från höger till vänster så utförs tre steg:

- $T^{\underline{e}}_{\underline{v}}$: byt från $\underline{e}$-koordinater till $\underline{v}$-koordinater,
- $(F)^{\underline{v}}_{\underline{w}}$: tillämpa F på en vektor given i $\underline{v}$-koordinater och ge resultatet i $\underline{w}$-koordinater,
- $T^{\underline{w}}_{\underline{f}}$: byt från $\underline{w}$-koordinater till $\underline{f}$-koordinater.

Totalt sett motsvarar det vänsterledet $(F)^{\underline{e}}_{\underline{f}}$: tillämpa F på en vektor given i $\underline{e}$-koordinater och ge resultatet i $\underline{f}$-koordinater.

Bevis: Låt $\vec{x} \in \mathcal{V}$. Vi beräknar vänsterled och högerled i ekvationen multiplicerat med $\vec{x}_{\underline{e}}$. Vi får då

$$(F)^{\underline{e}}_{\underline{f}}\vec{x}_{\underline{e}} = F(\vec{x})_{\underline{f}}$$

respektive

$$T^{\underline{w}}_{\underline{f}}(F)^{\underline{v}}_{\underline{w}}T^{\underline{e}}_{\underline{v}}\vec{x}_{\underline{e}} = T^{\underline{w}}_{\underline{f}}(F)^{\underline{v}}_{\underline{w}}\vec{x}_{\underline{v}} = T^{\underline{w}}_{\underline{f}}F(\vec{x})_{\underline{w}} = F(\vec{x})_{\underline{f}}.$$

Eftersom resultat är det samma för varje kolonn $\vec{x}_{\underline{e}}$ så gäller

$$(F)^{\underline{e}}_{\underline{f}} = T^{\underline{w}}_{\underline{f}}(F)^{\underline{v}}_{\underline{w}}T^{\underline{e}}_{\underline{v}}.$$

Ett sätt att använda ovanstående sats är låta $\underline{e}$ och $\underline{f}$ vara baserna som vi vill hitta matrisen till F för. Vi väljer sedan baserna $\underline{v}$ och $\underline{w}$ på ett sätt så att $(F)^{\underline{v}}_{\underline{w}}$ är lätt att räkna ut och använder formeln för att räkna ut $(F)^{\underline{e}}_{\underline{f}}$. Nedan utförs strategin i tre geometriska exempel.

Exempel 3.3.1

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara den linjära avbildningen som roterar varje vektor med vinkeln π runt vektorn $\vec{n} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$. Bestäm matrisen till F i standardbasen.

Precis som rotation i planet ger en linjär avbildning så ger rotation i rummet kring en axel en linjär avbildning. För att bestämma matrisen i standardbasen börjar vi med att bestämma matrisen i en lättare bas. Sedan hittar vi den sökta matrisen genom att byta bas.

Det finns några vektorer som är lättare att rotera än andra. Till att börja med kan vi ta $\vec{v}_1 = \vec{n}$. Eftersom rotationen sker runt $\vec{n}$ får vi att $F(\vec{v}_1) = \vec{v}_1$. Nu behöver vi välja två ytterligare vektorer för att få en bas. Det räcker att se till att vektorerna är linjärt oberoende eftersom vi då har tre linjärt oberoende vektorer i $\mathbb{R}^3$, vilket ger en bas. Vektorerna som är vinkelräta mot $\vec{n}$ bildar ett plan som ges av ekvationen

$$x + 2y + 3z = 0.$$

Om $\vec{v}$ tillhör planet kommer $F(\vec{v}) = -\vec{v}$ eftersom rotationen är med vinkel π. Vi väljer därför $\vec{v}_2$ och $\vec{v}_3$ i planet. Till exempel kan vi ta

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

Vi har då en bas $\underline{v} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$. På grund av

$$F(\vec{v}_1) = \vec{v}_1, \quad F(\vec{v}_2) = -\vec{v}_2, \quad F(\vec{v}_3) = -\vec{v}_3$$

får vi att matrisen till F i basen $\underline{v}$ är

$$(F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Vi betecknar standardbasen i $\mathbb{R}^3$ med $\underline{e}$. Den sökta matrisen är då

$$(F)_{\underline{e}}^{\underline{e}} = T_{\underline{e}}^{\underline{v}} (F)_{\underline{v}}^{\underline{v}} T_{\underline{v}}^{\underline{e}}.$$

Från hur vi valde basen $\underline{v}$ ser vi att

$$T_{\underline{e}}^{\underline{v}} = \begin{pmatrix} 1 & -2 & -3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

Dessutom är

$$T_{\underline{v}}^{\underline{e}} = (T_{\underline{e}}^{\underline{v}})^{-1} = \begin{pmatrix} 1 & -2 & -3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}^{-1} = \frac{1}{14}\begin{pmatrix} 1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix}.$$

Här beräknas inversen med valfri metod.

Vi räknar nu ut $(F)_{\underline{e}}^{\underline{e}} = T_{\underline{e}}^{\underline{v}}(F)_{\underline{v}}^{\underline{v}}T_{\underline{v}}^{\underline{e}}$:

$$\begin{aligned} (F)_{\underline{e}}^{\underline{e}} &= \begin{pmatrix} 1 & -2 & -3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\frac{1}{14}\begin{pmatrix} 1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & -1 & 0 \\ 3 & 0 & -1 \end{pmatrix}\frac{1}{14}\begin{pmatrix} 1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix} = \frac{1}{14}\begin{pmatrix} -12 & 4 & 6 \\ 4 & -6 & 12 \\ 6 & 12 & 4 \end{pmatrix}. \end{aligned}$$

Vi kan testa genom att multiplicera med basvektorerna i $\underline{v}$:

$$(F)_{\underline{e}}^{\underline{e}}(\vec{v}_1)_{\underline{e}} = \frac{1}{14}\begin{pmatrix} -12 & 4 & 6 \\ 4 & -6 & 12 \\ 6 & 12 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 14 \\ 28 \\ 42 \end{pmatrix} = (\vec{v}_1)_{\underline{e}},$$

$$(F)_{\underline{e}}^{\underline{e}}(\vec{v}_2)_{\underline{e}} = \frac{1}{14}\begin{pmatrix} -12 & 4 & 6 \\ 4 & -6 & 12 \\ 6 & 12 & 4 \end{pmatrix}\begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 28 \\ -14 \\ 0 \end{pmatrix} = -(\vec{v}_2)_{\underline{e}},$$

$$(F)_{\underline{e}}^{\underline{e}}(\vec{v}_3)_{\underline{e}} = \frac{1}{14}\begin{pmatrix} -12 & 4 & 6 \\ 4 & -6 & 12 \\ 6 & 12 & 4 \end{pmatrix}\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{14}\begin{pmatrix} 42 \\ 0 \\ -14 \end{pmatrix} = -(\vec{v}_3)_{\underline{e}}.$$

Exempel 3.3.2

Låt $G : \mathbb{R}^3 \to \mathbb{R}^3$ vara den linjära avbildningen som speglar varje vektor i planet

$$x + 2y + 3z = 0.$$

Planet är samma som i förra exemplet och vi väljer basen på samma sätt. Vektorn $\vec{v}_1 = \vec{n}$ är vinkelrät mot planet så den speglas på minus sig själv: $G(\vec{v}_1) = -\vec{v}_1$. Vektorerna i planet speglas på sig själva: $G(\vec{v}_2) = \vec{v}_2$, $G(\vec{v}_3) = \vec{v}_3$. Sammanfattningsvis har vi

$$G(\vec{v}_1) = -\vec{v}_1, \quad G(\vec{v}_2) = \vec{v}_2, \quad G(\vec{v}_3) = \vec{v}_3$$

så matrisen till G i basen $\underline{v}$ är

$$(G)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Basbytesmatriserna är de samma som i förra exemplet. Vi räknar nu ut $(G)_{\underline{e}}^{\underline{e}} = T_{\underline{e}}^{\underline{v}}(G)_{\underline{v}}^{\underline{v}}T_{\underline{v}}^{\underline{e}}$:

$$\begin{aligned}
(G)_{\underline{e}}^{\underline{e}} &= \begin{pmatrix} 1 & -2 & -3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \frac{1}{14} \begin{pmatrix} -1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix} \\
&= \begin{pmatrix} -1 & -2 & -3 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \frac{1}{14} \begin{pmatrix} 1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 12 & -4 & -6 \\ -4 & 6 & -12 \\ -6 & -12 & -4 \end{pmatrix}.
\end{aligned}$$

Observera att $(G)_{\underline{e}}^{\underline{e}} = -(F)_{\underline{e}}^{\underline{e}}$, vilket vi kunde ha sett direkt i basen $\underline{v}$.

$$(G)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = -(F)_{\underline{v}}^{\underline{v}}.$$

En tolkning är att rotation runt en axel följt av spegling i origo är det samma som spegling i planet som är vinkelrätt mot axeln.

Exempel 3.3.3

Låt $H : \mathbb{R}^3 \to \mathbb{R}^3$ vara den linjära avbildningen som projicerar varje vektor ortogonalt på planet

$$x + 2y + 3z = 0.$$

Återigen väljer vi basen på samma sätt. Vektorn $\vec{v}_1 = \vec{n}$ är vinkelrät mot planet så den projiceras på $\vec{0}$. Vektorerna i planet projiceras på sig själva. Alltså får vi

$$H(\vec{v}_1) = \vec{0}, \quad H(\vec{v}_2) = \vec{v}_2, \quad H(\vec{v}_3) = \vec{v}_3$$

och matrisen till F i basen $\underline{v}$ är

$$(H)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Basbytesmatriserna är de samma som i förra exemplet. Vi räknar nu ut $(H)_{\underline{e}}^{\underline{e}} = T_{\underline{e}}^{\underline{v}} (H)_{\underline{v}}^{\underline{v}} T_{\underline{v}}^{\underline{e}}$:

$$\begin{aligned} (H)_{\underline{e}}^{\underline{e}} &= \begin{pmatrix} 1 & -2 & -3 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \frac{1}{14} \begin{pmatrix} -1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2 & -3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \frac{1}{14} \begin{pmatrix} 1 & 2 & 3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix} = \frac{1}{14} \begin{pmatrix} 13 & -2 & -3 \\ -2 & 10 & -6 \\ -3 & -6 & 5 \end{pmatrix}. \end{aligned}$$

Avslutningsvis kommer ett exempel med polynom.

Exempel 3.3.4

Låt $F : \mathcal{P}_2 \to \mathcal{P}_2$ vara den linjära avbildningen som uppfyller

$$F(1+t) = 3t + t^2, \quad F(1-t) = 1 + 2t.$$

Bestäm matrisen för F med avseende på baserna $\underline{e} = (1, t)$ och $\underline{f} = (1, t, t^2)$.

Vi väljer basen $\underline{v} = (1+t, 1-t)$ i $\mathcal{P}_2$. Då är

$$F(1+t)_{\underline{f}} = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix}, \quad F(1-t)_{\underline{f}} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}.$$

Alltså är

$$(F)_{\underline{f}}^{\underline{v}} = \begin{pmatrix} 0 & 1 \\ 3 & 2 \\ 1 & 0 \end{pmatrix}.$$

Eftersom

$$(F)^{\underline{e}}_{\underline{f}} = T^{\underline{f}}_{\underline{f}}(F)^{\underline{v}}_{\underline{f}}T^{\underline{e}}_{\underline{v}} = (F)^{\underline{v}}_{\underline{f}}T^{\underline{e}}_{\underline{v}}$$

söker vi $T^{\underline{e}}_{\underline{v}} = (T^{\underline{v}}_{\underline{e}})^{-1}$. På grund av hur vi valde vår bas $\underline{v} = (1+t, 1-t)$ är

$$T^{\underline{v}}_{\underline{e}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

så

$$T^{\underline{e}}_{\underline{v}} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Till slut får vi

$$(F)^{\underline{e}}_{\underline{f}} = (F)^{\underline{v}}_{\underline{f}}T^{\underline{e}}_{\underline{v}} = \begin{pmatrix} 0 & 1 \\ 3 & 2 \\ 1 & 0 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 5 & 1 \\ 1 & 1 \end{pmatrix}.$$

Som tidigare kan vi testa vårt svar

$$(F)^{\underline{e}}_{\underline{f}}(1+t)_{\underline{e}} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 5 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = (3t+t^2)_{\underline{e}},$$

$$(F)^{\underline{e}}_{\underline{f}}(1-t)_{\underline{e}} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 5 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = (1+2t)_{\underline{e}}.$$

3.4 Kärna och bild

Vi har tidigare nämnt att linjära avbildningar kan användas för att jämföra olika vektorrum. Men hur mycket information som bevaras vid en sådan jämförelse beror på vilken linjär avbildning det rör sig om. Till exempel kan vi definiera en linjär avbildning $F : \mathcal{V} \to \mathcal{W}$ som avbildar varje vektor på nollvektorn, det vill säga $F(\vec{x}) = \vec{0}$. I så fall går all information förlorad när vi passerar från $\vec{x}$ till $F(\vec{x})$ och ingen vidare koppling mellan $\mathcal{V}$ och $\mathcal{W}$ har egentligen gjorts. I det här avsnittet ska vi introducera kärnan och bilden av F. Dessa är vektorrum som säger något om hur stark koppling F ger oss.

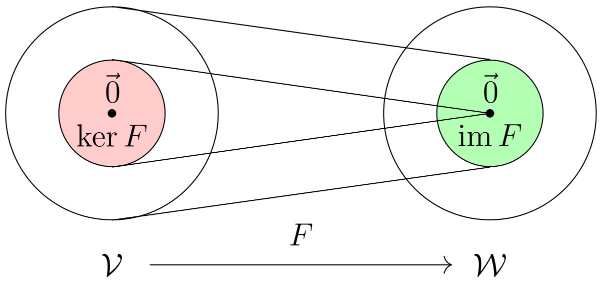
3.4.1 Definition

Definition 3.10

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning.

1. Kärnan av F är $\ker F = \{\vec{x} \in \mathcal{V} \mid F(\vec{x}) = \vec{0}\} \subseteq \mathcal{V}$
2. Bilden av F är $\operatorname{im} F = \{\vec{y} \in \mathcal{W} \mid F(\vec{x}) = y, \text{ för något } \vec{x} \in \mathcal{V}\} \subseteq \mathcal{W}$.

Alltså består $\ker F$ av alla vektorer i $\mathcal{V}$ som avbildas på nollvektorn och $\operatorname{im} F$ av alla vektorer i $\mathcal{W}$ som blir träffade av F. Till exempel visar $F(\vec{0}) = \vec{0}$ att både kärnan och bilden innehåller nollvektorn (i $\mathcal{V}$ respektive $\mathcal{W}$). Vi kan illustrera situationen såhär:



Figur 1: Kärna och bild

Figur 1 ska endast tolkas schematisk, där mängderna $\mathcal{V}$, $\mathcal{W}$, $\ker F$ och $\operatorname{im} F$ beskrivs som punkterna inuti vissa cirklar. I själva verket är $\mathcal{V}$ och $\mathcal{W}$ vektorrum. Det är även $\ker F$ och $\operatorname{im} F$, mer precist har vi följande.

Sats 3.11

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning.

a) Kärnan av F är ett delrum $\ker F \subseteq \mathcal{V}$.

b) Bilden av F är ett delrum $\operatorname{im} F \subseteq \mathcal{W}$.

Bevis:

a) Låt $\vec{u}, \vec{v} \in \ker F$. Då är

$$F(\vec{u} + \vec{v}) = F(\vec{u}) + F(\vec{v}) = \vec{0} + \vec{0} = \vec{0}.$$

Alltså är $\vec{u} + \vec{v} \in \ker F$. På liknande sätt gäller för $\lambda \in \mathbb{R}$ att

$$F(\lambda\vec{u}) = \lambda F(\vec{u}) = \lambda\vec{0} = \vec{0}.$$

Alltså är $\lambda\vec{u} \in \ker F$. Vi har nu visat att $\ker F \subseteq \mathcal{V}$ är ett delrum.

b) Låt $\vec{u}, \vec{v} \in \operatorname{im} F$. Då finns $\vec{x}_1, \vec{x}_2 \in \mathcal{V}$ så att $\vec{u} = F(\vec{x}_1)$ och $\vec{v} = F(\vec{x}_2)$. Vi får

$$\vec{u} + \vec{v} = F(\vec{x}_1) + F(\vec{x}_2) = F(\vec{x}_1 + \vec{x}_2),$$

vilket ger $\vec{u} + \vec{v} \in \operatorname{im} F$. Dessutom är

$$\lambda\vec{u} = \lambda F(\vec{x}_1) = F(\lambda\vec{x}_1),$$

vilket ger $\lambda\vec{u} \in \operatorname{im} F$. Alltså är $\operatorname{im} F \subseteq \mathcal{W}$ ett delrum.

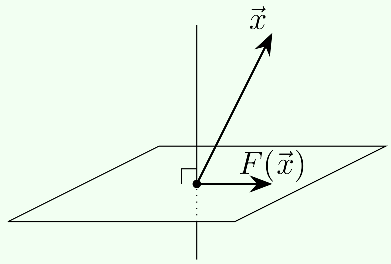
Vi räknar nu ut kärnan och bilden i några exempel. Vi börjar med geometriska exempel.

Exempel 3.4.1

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara ortogonal projektion på planet

$$x + 2y + 3z = 0.$$

Geometrisk kan vi illustrera F så här



En vektor $\vec{x}$ projiceras på nollvektorn om och endast om den är parallell med planets normallinje, det vill säga $\vec{x} = t\vec{n}$, där

$$\vec{n} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Med andra ord är

$$\ker F = \operatorname{span}\{\vec{n}\}.$$

För att bestämma bilden noterar vi att $F(\vec{x})$ tillhör planet för varje $\vec{x}$. Alltså är bilden en delmängd av planet. Faktiskt är bilden precis lika med planet. För att se det kan vi notera att $F(\vec{x}) = \vec{x}$ för alla $\vec{x}$ som tillhör planet. Alltså är

$$\operatorname{im} F = \mathcal{U} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \, x + 2y + 3z = 0 \right\}.$$

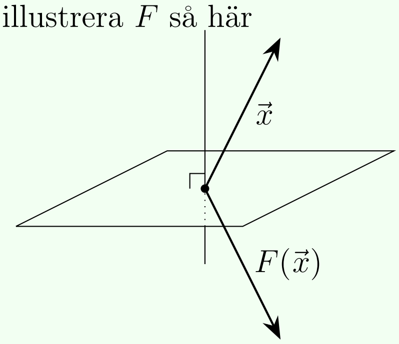
I detta exempel är $\dim \ker F = 1$ och $\dim \operatorname{im} F = 2$.

Exempel 3.4.2

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara spegling i planet

$$x + 2y + 3z = 0.$$

Geometrisk kan vi illustrera F så här



Den enda vektor som speglas på nollvektorn är nollvektorn själv. Alltså är

$$\ker F = \{\vec{0}\}.$$

Om vi speglar en vektor två gånger får vi tillbaka vektorn vi började med. Med andra ord gäller $F(F(\vec{x})) = \vec{x}$ för varje $\vec{x}$. Alltså tillhör varje vektor bilden, det vill säga $\operatorname{im} F = \mathbb{R}^3$.

Notera att i detta exempel är $\dim \ker F = 0$ och $\dim \operatorname{im} F = 3$.

Låt oss nu ta ett exempel som inte är direkt kopplat till geometri.

Exempel 3.4.3

Låt $F : \mathcal{P}_3 \to \mathcal{P}_3$ ges av $F(p) = p'$. Om vi skriver $p(t) = a_0 + a_1t + a_2t^2 + a_3t^3$ så får vi

$$F(a_0 + a_1t + a_2t^2 + a_3t^3) = a_1 + 2a_2t + 3a_3t^2$$

Det ger oss att $F(p) = 0$ om och endast om $p(t) = a_0$. Alltså är $\ker F = \mathcal{P}_0$, mängden av konstanta polynom. Vi ser också att $a_1 + 2a_2t + 3a_3t^2$ kan vara vilket polynom av grad högst två som helst. Alltså är $\operatorname{im} F = \mathcal{P}_2$. I detta exempel har vi att $\dim \ker F = 1$ och $\dim \operatorname{im} F = 3$.

Vi tar nu ett lite mer allmänt exempel.

Exempel 3.4.4

Låt $F : \mathbb{R}^n \to \mathbb{R}^m$ ges av $F(\vec{x}) = A\vec{x}$ där A är någon $(m \times n)$-matris.

Då är $F(\vec{x}) = \vec{0}$ om och endast om $A\vec{x} = \vec{0}$. Alltså är $\ker F$ precis nollrummet till A.

Låt $\vec{y} \in \mathbb{R}^m$. Då gäller $\vec{y} \in \operatorname{im} F$ om och endast om $\vec{y} = A\vec{x}$ för någon vektor $\vec{x}$. Men vektorer på formen $A\vec{x}$ är precis linjärkombinationer av kolonnerna i A, så vi får att $\operatorname{im} F$ är kolonnrummet till A.

Ovanstående exempel är ett specialfall av följande sats.

Sats 3.12

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning, $\underline{v}$ en bas i $\mathcal{V}$ och $\underline{w}$ en bas i $\mathcal{W}$. Sätt $A = (F)^{\underline{v}}_{\underline{w}}$.

a) $\vec{x} \in \ker F$ om och endast om $\vec{x}_{\underline{v}}$ tillhör nollrummet av A.

b) $\vec{y} \in \operatorname{im} F$ om och endast om $\vec{y}_{\underline{w}}$ tillhör kolonnrummet av A.

Bevis:

a) Vi har att $\vec{x} \in \ker F$ om och endast om $F(\vec{x}) = \vec{0}$, vilket är ekvivalent med $F(\vec{x})_{\underline{w}} = \vec{0}$ och $A\vec{x}_{\underline{v}} = \vec{0}$. Det sista villkoret säger precis att $\vec{x}_{\underline{v}}$ tillhör nollrummet av A.

b) Vi har att $\vec{y} \in \operatorname{im} F$ om och endast om $\vec{y} = F(\vec{x})$ för något $\vec{x}$, vilket är ekvivalent med $\vec{y}_{\underline{w}} = F(\vec{x})_{\underline{w}}$ och $\vec{y}_{\underline{w}} = A\vec{x}_{\underline{v}}$. Eftersom $\vec{x}_{\underline{v}}$ kan vara vilken vektor som helst får vi att det sista villkoret är ekvivalent med att $\vec{y}_{\underline{w}}$ tillhör kolonnrummet av A.

3.4.2 Dimensionssatsen

Kärnan motsvarar information som går förlorad då vi tillämpar F (alla vektorer i kärnan avbildas ju på samma sak, nämligen nollvektorn). Bilden motsvarar den information som finns kvar efter att vi tillämpat F. Det är därför rimligt att storlekarna av kärnan och bilden på något sätt balanserar ut varandra. Om kärnan är stor borde bilden vara liten och vice versa. Så är också fallet som vi kan se i följande sats

Sats 3.13: Dimensionssatsen

Låt $\mathcal{V}, \mathcal{W}$ vara ändligt dimensionella vektorrum och $F : \mathcal{V} \to \mathcal{W}$ en linjär avbildning. Då gäller

$$\dim \mathcal{V} = \dim \ker F + \dim \operatorname{im} F$$

Bevis: Eftersom $\mathcal{V}, \mathcal{W}$ är ändligt dimensionella vektorrum kan vi välja baser $\underline{v}$ i $\mathcal{V}$ och $\underline{w}$ i $\mathcal{W}$, så att F beskrivs av en matris $A = (F)^{\underline{v}}_{\underline{w}}$. Enligt Sats 3.12, ger koordinatavbildningen tillhörande $\underline{w}$ en linjär avbildning från $\operatorname{im} F$ till kolonnrummet av A. Koordinatavbildningar är inverterbara så vi får att $\operatorname{im} F$ har samma dimension som kolonnrummet av A, alltså rangen r av A. Vi har därmed

$$\dim \operatorname{im} F = r.$$

På liknande sätt ger oss koordinatavbildningen tillhörande $\underline{v}$ att $\ker F$ har samma dimension som nollrummet av A. Denna dimension är antalet kolonner i A (det vill säga $\dim \mathcal{V}$) minus rangen av A (det vill säga r). Alltså har vi

$$\dim \ker F = \dim \mathcal{V} - r.$$

Nu lägger vi till ekvationen $\dim \operatorname{im} F = r$ till ovanstående och får

$$\dim \ker F + \dim \operatorname{im} F = \dim \mathcal{V} - r + r = \dim \mathcal{V}.$$

Exempel 3.4.5

Vi kan illustrera Dimensionssatsen genom att återbesöka exemplen som vi gick igenom innan. Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$. Då är $\mathcal{V} = \mathbb{R}^3$ och $\dim \mathcal{V} = 3$. Om till exempel F är ortogonal projektion på ett plan så har vi $\dim \ker F = 1$ och $\dim \operatorname{im} F = 2$. Detta stämmer med Dimensionssatsen eftersom $3 = 1+2$. Om i stället F är spegling i ett plan så har vi $\dim \ker F = 0$ och $\dim \operatorname{im} F = 3$, vilket också stämmer då $3 = 0 + 3$.

Låt oss nu undersöka $F : \mathcal{P}_3 \to \mathcal{P}_3$ där $F(p) = p'$. Då är $\dim \mathcal{P}_3 = 4$. Vi såg att $\dim \ker F = 1$ och $\dim \operatorname{im} F = 3$, vilket stämmer då $4 = 1 + 3$.

3.4.3 Injektiva, surjektiva och bijektiva avbildningar

Vi ska nu koppla begreppen kärna och bild till begreppen injektiv, surjektiv och bijektiv. Vi påminner först om vad dessa ord betyder.

Definition 3.14

Låt A, B vara mängder och $f : A \to B$ vara en funktion. För varje $y \in B$ får vi en ekvation $f(x) = y$. Funktionen f kallas

1. injektiv om $f(x) = y$ har som mest en lösning $x \in A$ för varje $y \in B$,
2. surjektiv om $f(x) = y$ har minst en lösning $x \in A$ för varje $y \in B$,
3. bijektiv om $f(x) = y$ har precis en lösning $x \in A$ för varje $y \in B$.

Att funktionen f inte är injektiv betyder alltså att det finns olika element $x_1, x_2 \in A$ så att $f(x_1) = f(x_2)$. Att f inte är surjektiv betyder att det finns något $y \in B$ som inte är ett funktionsvärde $f(x)$. Det är fullt möjligt att f varken är injektiv eller surjektiv.

Att funktionen f är bijektiv betyder att f är både injektiv och surjektiv. Alltså finns fyra alternativ:

- f bijektiv, till exempel $f : \mathbb{R} \to \mathbb{R},\ f(x) = 2x + 1$,
- f injektiv men inte surjektiv, till exempel $f : \mathbb{R} \to \mathbb{R},\ f(x) = e^x$,

- f surjektiv men inte injektiv, till exempel $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^3 - x$,
- f varken injektiv eller surjektiv, till exempel $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^2$.

Att f är bijektiv innebär att vi kan definiera en inversfunktion $f^{-1} : B \to A$, där $f^{-1}(y) = x$ om och endast om $y = f(x)$. Att en sådan funktion går att definiera på hela B betyder just att $y = f(x)$ har entydig lösning för varje $y \in B$. Alltså är f bijektiv om och endast om f är inverterbar.

Låt oss nu gå tillbaka till våra geometriska exempel.

Exempel 3.4.6

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara ortogonal projektion på ett plan. Då avbildas alla vektorer parallella med normallinjen till planet på samma sak, nämligen nollvektorn, så F är inte injektiv. En vektor utanför planet kan inte träffas av F så F är inte heller surjektiv.

Om vi istället låter $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara spegling i ett plan så är $F(\vec{x}) = y$ om och endast om $\vec{x} = F(\vec{y})$. Vi har alltså precis en lösning $\vec{x}$ för varje $\vec{y}$. Med andra ord är F bijektiv i detta fall. Faktiskt är F sin egen invers.

Vi ska nu se på hur det ligger till med matrisavbildningar.

Sats 3.15

Låt A vara en $(m \times n)$-matris med rang r och $F : \mathbb{R}^n \to \mathbb{R}^m$ avbildningen som ges av $F(\vec{x}) = A\vec{x}$. Då är

a) F injektiv om och endast om $r = n$,

b) F surjektiv om och endast om $r = m$,

c) F bijektiv om och endast om $r = m = n$.

Bevis:

a) F injektiv betyder att $A\vec{x} = \vec{y}$ har som mest en lösning för varje $\vec{y}$. Vi kan förstå $A\vec{x} = \vec{y}$ som ett linjärt ekvationssystem med högerled $\vec{y}$. Det har som mest en lösning för varje högerled om och endast om rangen av A är antalet kolonner, det vill säga $r = n$.

b) F surjektiv betyder på liknande sätt att $A\vec{x} = \vec{y}$ ska ha minst en lösning för varje $\vec{y}$. Det gäller precis då rangen av A är lika med antalet rader, det vill säga $r = m$.

c) F är bijektiv om och endast om F är både injektiv och surjektiv alltså $r = m = n$.

Observera att i satsen ovan gäller $r \leq m$ och $r \leq n$ alltid. Om $m < n$, det vill säga A är en bred matris, så kan F inte vara injektiv eftersom vi då har $r \leq m < n$. Om $m > n$, det vill säga A är en smal matris, så kan F inte vara surjektiv eftersom $r \leq n < m$.

I allmänhet kan vi koppla begreppen injektiv och surjektiv till kärna och bild. Närmare bestämt gäller följande sats.

Sats 3.16

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning. Då är

a) F injektiv om och endast om $\ker F = \{0\}$,

b) F surjektiv om och endast om $\operatorname{im} F = \mathcal{W}$.

Bevis:

a) Om F är injektiv så har $F(\vec{x}) = \vec{0}$ bara en lösning $\vec{x} = \vec{0}$ och $\ker F = \{0\}$. Antag nu $\ker F = \{0\}$. För att visa att F är injektiv tar vi två vektorer $\vec{x}_1, \vec{x}_2 \in \mathcal{V}$ som har samma bild, det vill säga $F(\vec{x}_1) = F(\vec{x}_2)$. Det vi behöver visa är att $\vec{x}_1 = \vec{x}_2$. Vi skriver om vårt antagande

$$F(\vec{x}_1) = F(\vec{x}_2) \Leftrightarrow F(\vec{x}_1) - F(\vec{x}_2) = \vec{0} \Leftrightarrow F(\vec{x}_1 - \vec{x}_2) = \vec{0}.$$

Nu kan vi använda $\ker F = \{0\}$ som ger oss att $\vec{x}_1 - \vec{x}_2 = \vec{0}$ och därmed $\vec{x}_1 = \vec{x}_2$.

b) F är surjektiv om och endast om varje vektor i $\mathcal{W}$ träffas av någon vektor, men det är precis vad $\operatorname{im} F = \mathcal{W}$ säger.

Exempel 3.4.7

a) Låt $F : \mathcal{P}_3 \to \mathcal{P}_3$ ges av $F(p) = p'$. Då är $\ker F = \mathcal{P}_0 \neq \{0\}$ och F är inte injektiv. Till exempel är $F(t) = 1 = F(t + 1)$. Vidare är $\operatorname{im} F = \mathcal{P}_2 \neq \mathcal{P}_3$ så F är inte heller surjektiv.

b) Vi ändrar lite på exemplet och definierar $G : \mathcal{P}_3 \to \mathcal{P}_2$ där $G(p) = p'$. Då är G inte heller injektiv, men surjektiv eftersom vi har $\operatorname{im} G = \mathcal{P}_2$. För att avgöra om en funktion är surjektiv måste målmängden specificeras!

c) Låt $H : \mathcal{P}_2 \to \mathcal{P}_3$ ges av $H(p) = tp(t)$. Då är $H(p) = 0$ om och endast om $p(t) = 0$, så $\ker H = \{0\}$ och H är injektiv. Däremot är H inte surjektiv, eftersom inga konstanta nollskilda polynom tillhör $\operatorname{im} H$.

Vi kan nu koppla ihop ovanstående sats med Dimensionssatsen. Om $\mathcal{V}, \mathcal{W}$ är ändligt dimensionella vektorrum och $F : \mathcal{V} \to \mathcal{W}$ en linjär avbildning, så gäller

$$\dim \mathcal{V} = \dim \ker F + \dim \operatorname{im} F$$

Vi har att $\operatorname{im} F \subseteq \mathcal{W}$ är ett delrum så $\dim \operatorname{im} F \leq \dim \mathcal{W}$. Likhet $\dim \operatorname{im} F = \dim \mathcal{W}$ gäller om och endast om $\operatorname{im} F = \mathcal{W}$, det vill säga F är surjektiv. Vi har nu några olika alternativ

Om $\dim \mathcal{V} < \dim \mathcal{W}$, så är $\dim \operatorname{im} F = \dim \mathcal{V} - \dim \ker F < \dim \dim \mathcal{W}$. Alltså kan F inte vara surjektiv i detta fall.

Om $\dim \mathcal{V} > \dim \mathcal{W}$, så är $\dim \ker F = \dim \mathcal{V} - \dim \operatorname{im} F \geq \dim \mathcal{V} - \dim \mathcal{W} > 0$. Alltså är $\ker F \neq \{0\}$ och F är inte injektiv i detta fall.

Det återstår att undersöka fallet $\dim \mathcal{V} = \dim \mathcal{W}$. Då kan F vara både injektiv och surjektiv, men även här ger Dimensionssatsen en begränsning.

Sats 3.17

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en linjär avbildning antag $\dim \mathcal{V} = \dim \mathcal{W} < \infty$. Då är F injektiv om och endast om F är surjektiv. I detta fall är alltså F bijektiv.

Bevis: F är injektiv om och endast om $\ker F = \{0\}$, det vill säga $\dim \ker F = 0$. Dimensionssatsen ger att detta är ekvivalent med $\dim \operatorname{im} F = \dim \mathcal{V}$. Enligt antagande är $\dim \mathcal{V} = \dim \mathcal{W}$, så vi får att $\ker F = \{0\}$ om och endast om $\dim \operatorname{im} F = \dim \mathcal{W}$, det vill säga då F är surjektiv.

Det ovanstående sats säger är att om F är en linjär avbildning mellan vektorrum av samma (ändliga) dimension så är antingen F bijektiv eller så är F varken injektiv eller surjektiv. Som exempel kan vi åter ta spegling respektive ortogonal projektion på plan i $\mathbb{R}^3$. Speglingen är bijektiv, men den ortogonala projektionen är varken injektiv eller surjektiv.

3.5 Egenvärden och egenvektorer

Egenvärden och egenvektorer är verktyg för att analysera linjära avbildningar $F : \mathcal{V} \to \mathcal{V}$, alltså där F tar in och ger ut vektorer från samma vekorrum $\mathcal{V}$. I en sådan situation kan vi tänka oss att varje vektor i $\mathcal{V}$ på något sätt transformeras av F till någon annan vektor i $\mathcal{V}$. Genom att välja en bas i $\mathcal{V}$ kan vi beskriva F med hjälp av en kvadratisk matris A. På grund av detta samband mellan avbildningar och matriser kan vi också använda egenvärden och egenvektorer för att analysera kvadratiska matriser. Det finns många användningsområden för egenvärden och egenvektorer, till exempel i kvantmekanik och transformteori. I slutet av kursen kommer vi stöta på några specifika tillämpningar. Men innan dess börjar vi med ett geometrisk exempel.

3.5.1 Motiverande exempel

Vi har tidigare sett att det går att bestämma matriser för vissa geometriska avbildningar, som speglingar och projektioner genom att välja en viss bas. Vi ska nu se hur vi kan gå åt andra hållet: givet matrisen för en avbildning, tolka den geometriskt.

Exempel 3.5.1

Den linjära avbildningen $S : \mathbb{R}^3 \to \mathbb{R}^3$ är spegling i ett plan $\mathcal{U}$ genom origo. Matrisen till S i standardbasen är

$$A = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix}.$$

Bestäm planet $\mathcal{U}$.

Eftersom S är spegling i planet $\mathcal{U}$, består $\mathcal{U}$ precis av de vektorer $\vec{x}$ som uppfyller

$$S(\vec{x}) = \vec{x}$$

Vi uttrycker samma likhet med matrisen A och skriver om på följande sätt:

$$A\vec{x} = \vec{x} \Leftrightarrow A\vec{x} - \vec{x} = \vec{0} \Leftrightarrow A\vec{x} - I\vec{x} = \vec{0} \Leftrightarrow (A - I)\vec{x} = \vec{0}.$$

I ovanstående omskrivning ersätter vi $\vec{x}$ med $I\vec{x}$ (vilket vi kan göra då I är enhetsmatrisen) så att vi kan bryta ut $\vec{x}$ enligt regler för matrisräkning. Likheten $(A-I)\vec{x} = \vec{0}$ motsvarar ett homogent linjärt ekvationssystem med koefficientmatris

$$A - I = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} - 1 & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} - 1 & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} - 1 \end{pmatrix}$$
$$= \begin{pmatrix} -\frac{4}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} \end{pmatrix}.$$

Vi löser nu $(A - I)\vec{x} = \vec{0}$.

$$\left(\begin{array}{ccc|c} -\frac{4}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \end{array}\right) -\frac{3}{4} \sim \left(\begin{array}{ccc|c} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \\ \frac{2}{3} & -\frac{1}{3} & -\frac{1}{3} & 0 \end{array}\right) -\frac{2}{3} \; -\frac{2}{3}$$
$$\sim \left(\begin{array}{ccc|c} 1 & -\frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} \frac{1}{2}s + \frac{1}{2}t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \qquad s, t \in R.$$

Alltså är planet

$$\mathcal{U} = \operatorname{span}\left\{ \begin{pmatrix} \frac{1}{2} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} \frac{1}{2} \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Ett alternativt sätt att lösa uppgiften är att istället bestämma normallinjen till planet $\mathcal{U}$. Normallinjen består precis av de vektorer $\vec{x}$ som uppfyller $S(\vec{x}) = -\vec{x}$. Vi skulle då istället få ekvationen $A\vec{x} = -\vec{x}$ som är ekvivalent med $(A + I)\vec{x} = \vec{0}$. Det leder till en linkande beräkning som i första lösningen:

$$A + I = \frac{1}{3}\begin{pmatrix} -1 & 2 & 2 \\ 2 & 2 & -1 \\ 2 & -1 & 2 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{3} + 1 & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} + 1 & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{2}{3} + 1 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{2}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{5}{3} & -\frac{1}{3} \\ \frac{2}{3} & -\frac{1}{3} & \frac{5}{3} \end{pmatrix}.$$

Vi löser $(A+I)\vec{x} = \vec{0}$.

$$\left(\begin{array}{ccc|c} \frac{2}{3} & \frac{2}{3} & \frac{2}{3} & 0 \\ \frac{2}{3} & \frac{5}{3} & -\frac{1}{3} & 0 \\ \frac{2}{3} & -\frac{1}{3} & \frac{5}{3} & 0 \end{array}\right) \begin{matrix} -1 & -1 & -\frac{3}{2} \end{matrix} \sim \left(\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right) \begin{matrix} -1 & 1 \end{matrix}$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} -2t \\ t \\ t \end{pmatrix} = t \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \qquad t \in R.$$

Kom ihåg att vektorerna $\vec{x}$ är ortogonala med planet. Om vi till exempel väljer lösningen då $t = -1$ som normalvektor, så får vi att planet $\mathcal{U}$ ges av ekvationen

$$2x - y - z = 0.$$

3.5.2 Definition och grundläggande metoder

I föregående exempel hittade vi planet $\mathcal{U}$ genom att studera de vektorer som avbildas på sig själva, alltså de som är sin egen bild. Som alternativ lösning hittade vi normalen till planet genom att studera de vektorer som avbildas på minus sig själva. Sådana vektorer, som är sin egen bild upp till omskalning med någon skalär λ kallas för egenvektorer. Närmare bestämt gör vi följande definition.

Definition 3.18

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning.

a) En vektor $\vec{v} \in \mathcal{V}$ kallas en egenvektor till F om $\vec{v} \neq \vec{0}$ och

$$F(\vec{v}) = \lambda \vec{v}$$

för något $\lambda \in \mathbb{R}$.

b) Om det finns en sådan vektor $\vec{v}$ kallas λ för ett egenvärde till F.

c) Egenrummet till F med avseende på λ är

$$\mathcal{E}_\lambda(F) = \{\vec{v} \in \mathcal{V} \mid F(\vec{v}) = \lambda\vec{v}\}.$$

Notera att $\mathcal{E}_\lambda(F)$ består precis av alla egenvektorer $\vec{v}$ med egenvärde λ samt nollvektorn $\vec{v} = \vec{0}$, som enligt definition inte är en egenvektor. Anledningen till att $\vec{0}$ är undantagen är att $F(\vec{0}) = \lambda\vec{0}$ gäller för varje λ, och vi vill att varje egenvektor ska ha ett entydigt egenvärde för att få en smidig terminologi.

Exempel 3.5.2

Låt $S : \mathbb{R}^3 \to \mathbb{R}^3$ vara den linjära avbildningen från föregående exempel. Det vill säga S är spegling i planet $\mathcal{U}$ som ges av $2x-y-z = 0$. Våra tidigare beräkningar ger oss att S har två egenvärden:

$$\lambda_1 = 1 \text{ med } \mathcal{E}_1(S) = \mathcal{U} = \operatorname{span}\left\{\begin{pmatrix}\frac{1}{2}\\1\\0\end{pmatrix}, \begin{pmatrix}\frac{1}{2}\\0\\1\end{pmatrix}\right\}$$

$$\lambda_2 = -1 \text{ med } \mathcal{E}_{-1}(S) = \operatorname{span}\left\{\begin{pmatrix}2\\-1\\-1\end{pmatrix}\right\}.$$

Om vi känner till egenvärdena kan vi räkna ut motsvarande egenvektorer på samma sätt som i exemplet ovan. Men det finns också metoder för att ta fram själva egenvärdena. Ett sätt ges i följande sats.

Sats 3.19

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning, $\underline{v}$ en bas i $\mathcal{V}$ och $A = (F)_{\underline{v}}^{\underline{v}}$. Då gäller:

a) $\mathcal{E}_\lambda(F) = \ker(F - \lambda\,\mathrm{id})$,

b) $\vec{x} \in \mathcal{E}_\lambda(F)$ om och endast om $(A - \lambda I)\vec{x}_{\underline{v}} = \vec{0}$,

c) λ är ett egenvärde till F om och endast om $\det(A - \lambda I) = 0$.

I del a) ovan ska $F - \lambda\,\mathrm{id}$ tolkas på följande sätt. Till att börja med är $\mathrm{id} : \mathcal{V} \to \mathcal{V}$ identitetsavbildningen, alltså den som avbildar varje vektor $\vec{x}$ på $\vec{x}$. Avbildningen $F - \lambda\,\mathrm{id} : \mathcal{V} \to \mathcal{V}$ är då den som uppfyller $(F - \lambda\,\mathrm{id})(\vec{x}) = F(\vec{x}) - \lambda\vec{x}$.

Del b) är mycket lik del a). Det som skiljer är att vi nu ser på koordinatvektorer med avseende på basen $\underline{v}$, för vilka F beskrivs av matrisen A. I en sådan situation talar vi även om egenvärden och egenvektorer till A, vilket enligt satsen är ekvivalent med egenvärden och egenvektorer till F.

Bevis:

a) Vi har att $\vec{x} \in \mathcal{E}_\lambda(F)$ om och endast om $F(\vec{x}) = \lambda\vec{x}$, vilket är ekvivalent med $F(\vec{x}) - \lambda\vec{x} = \vec{0}$ och $(F - \lambda\,\mathrm{id})(\vec{x}) = \vec{0}$. Alltså är $\mathcal{E}_\lambda(F) = \ker(F - \lambda\,\mathrm{id})$.

b) Genom att ta koordinater i basen $\underline{v}$ har vi att $\vec{x} \in \mathcal{E}_\lambda(F)$ om och endast om $A\vec{x} = \lambda\vec{x}_{\underline{v}}$ som är ekvivalent med $A\vec{x} - \lambda\vec{x}_{\underline{v}} = \vec{0}$ och $(A - \lambda I)\vec{x}_{\underline{v}} = \vec{0}$.

c) Från b) får vi att λ är ett egenvärde till F om och endast om $(A - \lambda I)\vec{x}_{\underline{v}} = \vec{0}$ har någon lösning $\vec{x}_{\underline{v}} \neq \vec{0}$. Eftersom $A - \lambda I$ är en kvadratisk matris finns en sådan lösning om och endast om $A - \lambda I$ inte är inverterbar, det vill säga $\det(A - \lambda I) = 0$.

3.5.3 Karaktäristiskt polynom

Uttrycket $\det(A-\lambda I)$ i Sats 3.19 är ett polynom av grad $n = \dim \mathcal{V}$. För att se varför kan vi resonera på fölande sätt. Elementen i $A-\lambda I$ på diagonalen är polynom $a_{ii} - \lambda$ av grad 1, och de utanför diagonalen är konstanter a_{ij}. Enligt determinantens definition är $\det(A-\lambda I)$ en summa av produkter av sådana element (med olika tecken). Detta ger att $\det(A-\lambda I)$ är ett polynom i λ. Termen av högst grad får vi då vi multiplicerar diagonalelementen vilket ger $\prod_{i=1}^n (a_{ii} - \lambda) = (-\lambda)^n + \cdots$, där $\cdots$ är termer av lägre grad.

Eftersom nollställena till polynomet $\det(A - \lambda I)$ är precis egenvärdena till A förtjänar polynomet ett namn.

Definition 3.20

Låt A vara en $(n \times n)$-matris. Det karaktäristiska polynomet till A är

$$\chi_A(\lambda) = \det(A - \lambda I).$$

Sats 3.19 c) säger alltså att λ är ett egenvärde till A om och endast om $\chi_A(\lambda) = 0$. Eftersom $\chi_A(\lambda)$ är ett polynom av grad n har $\chi_A(\lambda) = 0$ som mest n olika lösningar. Det följer alltså att en $(n \times n)$-matris A har som mest n olika egenvärden.

Exempel 3.5.3

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ ges av $F(\vec{x}) = A\vec{x}$, där

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

1. Beräkna alla egenvärden till F.
2. Beräkna alla egenvektorer till F.
3. Visa att vektorer från olika egenrum till F alltid är ortogonala.
4. Beskriv F geometriskt.

För att bestämma egenvärdena till F löser vi $\chi_A(\lambda) = 0$. Eftersom $\chi_A(\lambda) = \det(A - \lambda I)$, börjar vi med att räkna ut

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

Vi räknar nu ut determinanten av ovanstående matris med hjälp av rad- och kolonnoperationer. Vårt mål är att försöka faktorisera så mycket som möjligt då det gör det lättare att hitta nollställena som vi är efter. Om vi ser någon gemensam faktor i en rad eller kolonn kommer vi därför bryta ut den, vilket är möjligt på grund av sambanden mellan determinater och rad- respektive kolonnoperationer. Repetera dessa samband om du känner dig osäker på vad som gäller.

$$\begin{aligned}
\chi_A(\lambda) &= \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & -1 \\ 1 & -1 & 2-\lambda \end{vmatrix} \, (-1) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & -1 \\ 0 & -3+\lambda & 3-\lambda \end{vmatrix} \\
&= (3-\lambda) \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & -1 \\ 0 & -1 & 1 \end{vmatrix} \, (1) = (3-\lambda) \begin{vmatrix} 2-\lambda & 2 & 1 \\ 1 & 1-\lambda & -1 \\ 0 & 0 & 1 \end{vmatrix} \\
&= (3-\lambda) \begin{vmatrix} 2-\lambda & 2 \\ 1 & 1-\lambda \end{vmatrix} = (3-\lambda)((2-\lambda)(1-\lambda) - 2) \\
&= (3-\lambda)(\lambda^2 - 3\lambda + 2 - 2) = (3-\lambda)(\lambda^2 - 3\lambda) \\
&= -(\lambda - 3)(\lambda - 3)\lambda.
\end{aligned}$$

Alltså är $\chi_A(\lambda) = 0$ om och endast om $\lambda = 3$ eller $\lambda = 0$. Det betyder att vi har två egenvärden $\lambda_1 = 3$ och $\lambda_2 = 0$. Vi bestämmer nu egenvektorerna till vart och ett av dessa egenvärden λ, genom att lösa $(A - \lambda I)\vec{x} = \vec{0}$.

Vi börjar med $\lambda_1 = 3$. Egenrummet till 3 består av lösningarna till $(A - 3I)\vec{x} = \vec{0}$:

$$\left(\begin{array}{ccc|c} -1 & 1 & 1 & 0 \\ 1 & -1 & -1 & 0 \\ 1 & -1 & -1 & 0 \end{array}\right) \begin{matrix} 1 \; 1 \; -1 \\ \\ \\ \end{matrix} \sim \left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} s+t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \qquad s, t \in R.$$

Att vi får parameterlösningar bör inte komma som en överraskning, vi bestämde ju λ just så att $A - \lambda I$ inte är invertebar. Entydig lösning skulle därmed betyda att vi räknat fel antingen på egenvärdet eller när vi löste ekvationssystemet. Vi ser nu att:

$$\mathcal{E}_3(F) = \text{span}\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Vi kan lätt kontrollera svaret

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

$$\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} = 3\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

Vi gör på samma sätt med med $\lambda_1 = 0$. Notera att i detta fall är egenvektorerna precis kärnan till F:

$$\left(\begin{array}{ccc|c} 2 & 1 & 1 & 0 \\ 1 & 2 & -1 & 0 \\ 1 & -1 & 2 & 0 \end{array}\right) \sim \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 2 & 1 & 1 & 0 \\ 1 & -1 & 2 & 0 \end{array}\right) \begin{matrix} -2 \;\; -1 \\ \\ \\ \end{matrix}$$

$$\sim \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & -3 & 3 & 0 \\ 0 & -3 & 3 & 0 \end{array}\right) \begin{matrix} \\ -1 \;\; -\tfrac{1}{3} \\ \\ \end{matrix} \sim \left(\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \begin{matrix} \\ -2 \\ \\ \end{matrix}$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} -t \\ t \\ t \end{pmatrix} = t \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \qquad t \in R.$$

Alltså är

$$\mathcal{E}_0(F) = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Återigen kan vi kontrollera svaret

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

Sammanfattningsvis har vi följande egenrum

$$\mathcal{E}_3(F) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{E}_0(F) = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Vi ska nu visa att vektorerna från $\mathcal{E}_3(F)$ är ortogonala med de från $\mathcal{E}_0(F)$. Vi sätter

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Det räcker att visa $\vec{v}_1 \cdot \vec{v}_3 = 0$, $\vec{v}_2 \cdot \vec{v}_3 = 0$. Vi beräknar

$$\vec{v}_1 \cdot \vec{v}_3 = -1 + 1 + 0 = 0, \quad \vec{v}_2 \cdot \vec{v}_3 = -1 + 0 + 1 = 0.$$

Geometriskt kan vi nu tolka situationen på följande sätt. Egenrummet $\mathcal{E}_3(F) = \text{span}\{\vec{v}_1, \vec{v}_2\}$ är ett plan och egenrummet $\mathcal{E}_0(F) = \text{span}\{\vec{v}_3\}$ är dess normallinje. Varje vektor $\vec{x}$ kan skrivas $\vec{x} = \vec{u} + \vec{v}$, där $\vec{u} \in \mathcal{E}_3(F)$ och $\vec{v} \in \mathcal{E}_0(F)$. Vi kan förstå $\vec{u}$ som den ortogonala projektionen av $\vec{x}$ på planet $\text{span}\{\vec{v}_1, \vec{v}_2\}$. Eftersom F är linjär får vi $F(\vec{x}) = F(\vec{u} + \vec{v}) = F(\vec{u}) + F(\vec{v}) = 3\vec{u} + \vec{0} = 3\vec{u}$. Alltså kan vi tolka F som ortogonal projektion på planet $\text{span}\{\vec{v}_1, \vec{v}_2\}$ följt av omskalning med en faktor 3. Notera att $\underline{v} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ är en bas i $\mathbb{R}^3$. Eftersom

$$F(\vec{v}_1) = 3\vec{v}_1, \quad F(\vec{v}_2) = 3\vec{v}_2, \quad F(\vec{v}_3) = \vec{0},$$

får vi

$$(F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 3 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Från denna matrisbeskrivning kan vi göra samma tolkning.

Att matrisen $(F)_{\underline{v}}^{\underline{v}}$ i exemplet ovan har en så enkel form (jämfört med A) beror på att $\underline{v}$ består av egenvektorer till F. I nästa avsnitt ska vi studera sådana baser mer ingående. Men innan dess tar vi några exempel på ett annat vektorrum.

Exempel 3.5.4

Låt $\mathcal{V}$ vara vektorrummet av alla funktioner $f : \mathbb{R} \to \mathbb{R}$ som går att derivera godtyckligt många gånger. Med andra ord är $f \in \mathcal{V}$ om alla dess derivator $f', f'', \ldots$ existerar.

a) Låt $F : \mathcal{V} \to \mathcal{V}$ ges av $F(f) = f'$. Om $f(x) = e^{\lambda x}$ så är $f'(x) = \lambda e^{\lambda x} = \lambda f(x)$. Alltså är f en egenvektor med egenvärde λ. Det betyder att alla $\lambda \in \mathbb{R}$ är egenvärden till F. Att vi har oändligt många egenvärden är möjligt på grund av att $\mathcal{V}$ har oändlig dimension.

b) Låt $G : \mathcal{V} \to \mathcal{V}$ ges av $G(f) = f''$. Om $f(x) = e^{cx}$, så är $f'(x) = ce^{cx}$ och $f''(x) = c^2 e^{cx} = c^2 f(x)$. Alltså är $\lambda = c^2$ ett egenvärde med egenvektor $f(x) = e^{cx}$ (även $f(x) = e^{-cx}$ är en egenvektor med samma egenvärde då $(-c)^2 = c^2$). Alltså är alla $\lambda \geq 0$ egenvärden. Men även $\lambda < 0$ är egenvärden. För att se det kan vi ta $f(x) = \sin cx$. Då är $f'(x) = c \cos cx$

och $f''(x) = -c^2 \sin cx = -c^2 f(x)$, vilket ger oss egenvärdet $\lambda = -c^2$. På liknande sätt är $f(x) = \cos cx$ också en egenvektor med samma egenvärde. Som i föregående fall har vi att alla $\lambda \in \mathbb{R}$ är egenvärden till G.

Dessa exempel på egenvärden och egenvektorer har du nog utnyttjat (kanske utan att veta att de var det) vid lösning av differentialekvationer. Faktum är att egenvärden och egenvektorer dyker upp i massor av olika sammanhang, ibland på oväntade sätt. Håll utkik!

3.6 Diagonalisering

Vi har tidigare sett flera exempel på hur vi kan utnyttja egenvärden och egenvektorer för att undersöka linjära avbildningar. I det här avsnittet kommer vi göra samma sak fast mer systematiskt med hjälp av basbyten. Låt oss påminna om hur basbyte fungerar i fallet då $F : \mathcal{V} \to \mathcal{V}$ är en linjär avbildning. Antag att vi har två baser $\underline{e}$ och $\underline{v}$ i $\mathcal{V}$. Det kan vara vilka baser som helst, men det kan underlätta att tänka på $\underline{e}$ som någon given bas vi utgår ifrån då vi vi beskriver vektorerna i $\mathcal{V}$ och $\underline{v}$ som en vald bas anpassad till F. Vi kan då beskriva F med två matriser

$$(F)_{\underline{e}}^{\underline{e}} = A, \quad (F)_{\underline{v}}^{\underline{v}} = B.$$

Notera att A och B är kvadratiska matriser. Mer precis har de storlek $(n \times n)$, där $n = \dim \mathcal{V}$. Skriv nu $T_{\underline{e}}^{\underline{v}} = T$. Då är $T_{\underline{v}}^{\underline{e}} = T^{-1}$. Sambandet $T_{\underline{v}}^{\underline{e}}(F)_{\underline{e}}^{\underline{e}}T_{\underline{e}}^{\underline{v}} = (F)_{\underline{v}}^{\underline{v}}$ ger oss då

$$T^{-1}AT = B \text{ och } B = TAT^{-1}.$$

Vi kan nu använda A eller B för att undersöka egenvärden och egenvektorer till F. Egenvärden är precis nollställena till $\chi_A(\lambda)$ som också är nollställena till $\chi_B(\lambda)$. Faktum är att $\chi_A(\lambda) = \chi_B(\lambda)$, enligt följande sats.

Sats 3.21

Låt A, B och T vara $(n \times n)$-matriser så att T är inverterbar och $T^{-1}AT = B$. Då är $\chi_A(\lambda) = \chi_B(\lambda)$.

Bevis: Enligt definition är $\chi_A(\lambda) = \det(A - \lambda I)$ och $\chi_B(\lambda) = \det(B - \lambda I)$. Observera nu att $T^{-1}(A - \lambda I)T = T^{-1}AT - \lambda T^{-1}IT = B - \lambda I$. Alltså är

$$\begin{aligned} \det(B - \lambda I) &= \det(T^{-1}(A - \lambda I)T) = \det(T^{-1})\det(A - \lambda I)\det(T) \\ &= \det(T)^{-1}\det(A - \lambda I)\det(T) = \det(A - \lambda I), \end{aligned}$$

så $\chi_A(\lambda) = \chi_B(\lambda)$.

På grund av ovanstående ger varje matrisbeskrivning av en linjär avbildning $F : \mathcal{V} \to \mathcal{V}$ upphov till samma karaktäristiska polynom, som vi därför betecknar med $\chi_F(\lambda)$. Alltså är $\chi_F(\lambda) = \chi_A(\lambda)$, där $A = (F)_{\underline{e}}^{\underline{e}}$ för någon bas $\underline{e}$.

3.6.1 Motiverade exempel

Vi ska nu undersöka fallet då $\underline{v}$ är en bas bestående av egenvektorer till F. Vi illustrerar med exemplet från föregående avsnitt.

Exempel 3.6.1

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ ges av $F(\vec{x}) = A\vec{x}$, där

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

Då är $(F)_{\underline{e}}^{\underline{e}} = A$, där $\underline{e}$ är standardbasen. Vi beräknade tidigare att $\underline{v} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ där

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

är en bas av egenvektorer till F. Närmare bestämt såg vi att

$$F(\vec{v}_1) = 3\vec{v}_1, \quad F(\vec{v}_2) = 3\vec{v}_2, \quad F(\vec{v}_3) = \vec{0},$$

vilket ger oss att

$$B = (F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Vi beräknade $\chi_A(\lambda) = -(\lambda-3)(\lambda-3)\lambda$ genom en följd radoperationer. För B är beräkningen mycket enklare:

$$\chi_B(\lambda) = \begin{vmatrix} 3-\lambda & 0 & 0 \\ 0 & 3-\lambda & 0 \\ 0 & 0 & -\lambda \end{vmatrix} = (3-\lambda)(3-\lambda)(-\lambda) = -(\lambda-3)(\lambda-3)\lambda.$$

Det här är ett exempel på hur matrisen B är lättare att räkna med än A.

Matrisen $T = T_{\underline{e}}^{\underline{v}}$ är också lätt att skriva ner eftersom vektorerna i $\underline{v}$ är givna i standardbasen. Vi skriver helt enkelt $\vec{v}_1, \vec{v}_2, \vec{v}_3$ som kolonner

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

Matrisen $T^{-1} = T_{\underline{v}}^{\underline{e}}$, kan vi beräkna om vi behöver den genom att invertera T.

3.6.2 Definition och exempel

Baserat på exemplet i föregående avsnitt gör vi följande definitioner.

Definition 3.22

En diagonalmartis D är en $(n \times n)$-matris vars element utanför diagonalen är 0. Det vill säga

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Definition 3.23

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning.

a) En bas $\underline{v} = (\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n)$ kallas en egenbas till F om $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n$ är egenvektorer till F. Då är $F(\vec{v}_i) = \lambda_i \vec{v}_i$ för något $\lambda_i \in \mathbb{R}$, vilket är det samma som

$$(F)_{\underline{v}}^{\underline{v}} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

det vill säga $(F)_{\underline{v}}^{\underline{v}}$ är en diagonalmatris.

b) Vi säger att F är diagonaliserbar om det finns en sådan egenbas.

Om $F : \mathbb{R}^n \to \mathbb{R}^n$ ges av $F(\vec{x}) = A\vec{x}$ så är F diagonaliserbar om och endast om det finns en inverterbar matris T så att $T^{-1}AT$ är diagonal. I så fall säger vi att A är en diagonaliserbar matris.

Att hitta T inverterbar och D diagonal så att $T^{-1}AT = D$ kallas för att diagonalisera A. Vi ser hur det går till i ett exempel.

Exempel 3.6.2

Låt $F : \mathbb{R}^2 \to \mathbb{R}^2$ ges av $F(\vec{x}) = A\vec{x}$ där

$$A = \begin{pmatrix} -6 & -8 \\ 4 & 6 \end{pmatrix}.$$

a) Bestäm en egenbas till F.

b) Hitta T inverterbar och D diagonal så att $T^{-1}AT = D$.

Vi börjar med att bestämma egenvärden till F. Vi beräknar

$$\begin{aligned}\chi_A(\lambda) &= \begin{vmatrix} -6-\lambda & -8 \\ 4 & 6-\lambda \end{vmatrix} = (-6-\lambda)(6-\lambda) + 32 = \lambda^2 - 36 + 32 \\ &= \lambda^2 - 4 = (\lambda - 2)(\lambda + 2).\end{aligned}$$

Alltså har vi två egenvärden $\lambda_1 = 2$ och $\lambda_2 = -2$.

a) Vi bestämmer nu egenvektorerna till varje egenvärde λ genom att lösa $(A - \lambda I)\vec{x} = \vec{0}$. Vi börjar med $\lambda_1 = 2$:

$$\left(\begin{array}{cc|c} -8 & -8 & 0 \\ 4 & 4 & 0 \end{array}\right) \begin{matrix} -\frac{1}{8} & -4 \end{matrix} \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right) \quad \text{lösningar } \vec{x} = t\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \; t \in \mathbb{R}.$$

Vi fortsätter med $\lambda_2 = -2$:

$$\left(\begin{array}{cc|c} -4 & -8 & 0 \\ 4 & 8 & 0 \end{array}\right) \begin{matrix} -\frac{1}{4} & -4 \end{matrix} \sim \left(\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right) \text{ lösningar } \vec{x} = s\begin{pmatrix} -2 \\ 1 \end{pmatrix}, \; s \in \mathbb{R}.$$

Vi sätter $\underline{v} = (\vec{v}_1, \vec{v}_2)$ där

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

Då är $\underline{v}$ en egenbas. Närmare bestämt är $A\vec{v}_1 = 2\vec{v}_1$ och $A\vec{v}_2 = -2\vec{v}_2$. (Kontrollräkna att det stämmer!).

b) Vi sätter

$$T = T_{\underline{e}}^{\underline{v}} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Eftersom vi satt egenvärdena till $\vec{v}_1$ och $\vec{v}_2$ på diagonalen i D i samma ordning som de kommer i basen $\underline{v}$ så är $T^{-1}AT = D$. Vi kan kontrollräkna genom att först bestämma inversen

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

och sedan multiplicera

$$\begin{aligned} T^{-1}AT &= \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} -6 & -8 \\ 4 & 6 \end{pmatrix} \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} -2 & 4 \\ 2 & -2 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = D. \end{aligned}$$

Strategin för att hitta T och D så att $T^{-1}AT = D$ vi använde i ovanstående exempel är generell.

- T: Sätt kolonnerna till en bas av egenvektorer till A.
- D: Sätt motsvarande egenvärden på diagonalen i samma ordning.

Det enda som kan göra att strategin misslyckas är att om någon sådan egenbas inte finns, det vill säga A är inte diagonaliserbar. När så är fallet ska vi reda ut i nästa avsnitt.

3.6.3 Algebraisk och geometrisk mulitplicitet

I det här avsnittet ska vi ge ett svar på frågan när det finns en egenbas. Vi börjar med ett grundläggande begrepp.

Definition 3.24

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning med ett egenvärde λ. Den geometriska multipliciteten av λ är $\dim \mathcal{E}_\lambda(F)$.

Att λ är ett egenvärde ger att $\mathcal{E}_\lambda(F) \neq \{0\}$ så vi har att den geometriska multipliciteten av λ är minst 1.

Vi har följande strategi för att hitta en egenvas. Välj baser i varje egenrum och slå ihop till en samling vektorer $\underline{u} = (\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m)$ i $\mathcal{V}$. Notera att antalet vektorer m är precis summan av alla geometriska multipliciteter. För att $\underline{u}$ ska ha en chans att vara en bas måste $m = \dim \mathcal{V}$. Detta visar sig vara den enda begränsningen på grund av följande sats som mer eller mindre säger att vektorer från olika egenrum är linjärt oberoende.

Sats 3.25

Vektorer $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ valda som ovan är alltid linjärt oberoende. De bildar därmed en bas om och endast om $m = \dim \mathcal{V}$. Annars är $m < \dim \mathcal{V}$.

Bevis: Vi skriver $F(\vec{u}_i) = \lambda_i \vec{u}_i$ och genomför ett motsägelsebevis. Antag därför att $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_m$ är linjärt beroende. Då finns $k \leq m$ så att $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{k-1}$ är linjärt oberoende, men $\vec{u}_k \in \operatorname{span}\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{k-1}\}$, det vill säga

$$\vec{u}_k = c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_{k-1}\vec{u}_{k-1}. \tag{3.1}$$

Vi tillämpar nu $(F - \lambda_k \operatorname{id})$ på vänsterled och högerled i Ekvation (3.1). Vi observerar först att

$$(F - \lambda_k \operatorname{id})(\vec{u}_j) = F(\vec{u}_j) - \lambda_k\vec{u}_j = \lambda_j\vec{u}_j - \lambda_k\vec{u}_j = (\lambda_j - \lambda_k)\vec{u}_j.$$

I vänsterledet får vi $(F - \lambda_k \operatorname{id})(\vec{u}_k) = \vec{0}$ och i högerledet får vi

$$(F - \lambda_k \operatorname{id})(c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_{k-1}\vec{u}_{k-1}) =$$
$$c_1(\lambda_1 - \lambda_k)\vec{u}_1 + c_2(\lambda_2 - \lambda_k)\vec{u}_2 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\vec{u}_{k-1}.$$

Alltså har vi

$$\vec{0} = c_1(\lambda_1 - \lambda_k)\vec{u}_1 + c_2(\lambda_2 - \lambda_k)\vec{u}_2 + \cdots + c_{k-1}(\lambda_{k-1} - \lambda_k)\vec{u}_{k-1}$$

och eftersom $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_{k-1}$ är linjärt oberoende måste $c_j = 0$ eller $(\lambda_j - \lambda_k) = 0$ för varje $j \leq k-1$. Det första fallet kan vi bortse från att termen $c_j\vec{u}_j$ i Ekvation (3.1). I det andra fallet är $\lambda_j = \lambda_k$ så $\vec{u}_j$ kommer från samma egenrum som $\vec{u}_k$. Men då ger Ekvation (3.1) att $\vec{u}_k$ är en linjärkombination av egenvektorer från samma egenrum, vilket är omöjligt eftersom vi valde en bas i varje egenrum.

Om vi känner till de geometriska multipliciteterna kan vi alltså ta reda på om F är diagonaliserbar eller inte. En svårighet är att dessa multipliciteter kan vara krävande att beräkna. Men det finns ett relaterat begrepp som ofta är lättare att räkna ut.

Definition 3.26

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning med ett egenvärde λ_0. Den algebraiska multipliciteten av λ_0 antalet gånger λ_0 förekommer som rot till ekvationen $\chi_F(\lambda) = 0$. Med andra ord är den algebraiska multipliciteten m om $\chi_F(\lambda) = (\lambda - \lambda_0)^m q(\lambda)$ för något polynom $q(\lambda)$ så att $q(\lambda_0) \neq 0$.

Vi jämför algebraiska och geometriska multipliciteter i några exempel.

Exempel 3.6.3

Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ ges av $F(\vec{x}) = A\vec{x}$. Bestäm algebraiska och geometriska multipliciteter av alla egenvärden då

$$\text{a) } A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \qquad \text{b) } A = \begin{pmatrix} -1 & 2 & 1 \\ -2 & 3 & 2 \\ 0 & 0 & 3 \end{pmatrix}$$

Fallet a) är precis exemplet vi såg i avsnitt 3.6.1. Vi såg att $\chi_A(\lambda) = -(\lambda-3)(\lambda-3)\lambda$, så vi har två egenvärden $\lambda_1 = 3$ och $\lambda_2 = 0$ med algebraisk multiplicitet 2 respektive 1. Dessutom är

$$\mathcal{E}_3(F) = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}, \quad \mathcal{E}_0(F) = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix} \right\}$$

så $\dim \mathcal{E}_3(F) = 2$ och $\dim \mathcal{E}_0(F) = 1$. Alltså är de geometriska multipliciteterna också 2 respektive 1. Sammanfattningsvis har vi följande:

Egenvärde	algebraisk multiplicitet	geometrisk multiplicitet
$\lambda_1 = 3$	2	2
$\lambda_2 = 0$	1	1

Summan av de geometriska multipliciteterna är $1+2 = 3 = \dim \mathbb{R}^3$ och matrisen är diagonaliserbar.

b) Vi undersöker nu den andra matrisen. Vi börjar med att hitta egenvärden.

$$\chi_A(\lambda) = \begin{vmatrix} -1-\lambda & 2 & 1 \\ -2 & 3-\lambda & 2 \\ 0 & 0 & 3-\lambda \end{vmatrix} = (3-\lambda)\begin{vmatrix} -1-\lambda & 2 \\ -2 & 3-\lambda \end{vmatrix}$$
$$= (3-\lambda)((-1-\lambda)(3-\lambda)+4) = (3-\lambda)(\lambda^2-2\lambda-3+4)$$
$$= (3-\lambda)(\lambda^2-2\lambda+1) = (3-\lambda)(\lambda-1)^2.$$

Även i detta fall har vi två egenvärden $\lambda_1 = 1$ med algebraisk multiplicitet 2 och $\lambda_2 = 3$ med algebraisk multiplicitet 1. För att ta reda på geometriska multipliciteter behöver vi undersöka egenrummen.

Vi börjar med $\lambda_1 = 1$ och beräknar

$$\left(\begin{array}{ccc|c} -2 & 2 & 1 & 0 \\ -2 & 2 & 2 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) -\tfrac{1}{2} \sim \left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ -2 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) 2$$
$$\sim \left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 2 & 0 \end{array}\right) 2\ {-1}\ {-1} \sim \left(\begin{array}{ccc|c} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} s \\ s \\ 0 \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \qquad s \in R.$$

Alltså är den geometriska multipliciteten $\dim \mathcal{E}_1(F) = 1$.

Vi fortsätter nu med $\lambda_2 = 3$ och beräknar

$$\left(\begin{array}{ccc|c} -4 & 2 & 1 & 0 \\ -2 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) -\tfrac{1}{2} \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ -4 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) 4$$
$$\sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 2 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) \tfrac{1}{2} \sim \left(\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & -\frac{3}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} 2t \\ 3t \\ 2t \end{pmatrix} = t \begin{pmatrix} 2 \\ 3 \\ 2 \end{pmatrix} \qquad s \in R.$$

Alltså är den geometriska multipliciteten $\dim \mathcal{E}_3(F) = 1$. Sammanfattningsvis har vi följande:

Egenvärde	algebraisk multiplicitet	geometrisk multiplicitet.
$\lambda_1 = 1$	2	1
$\lambda_2 = 3$	1	1

Summan av de geometriska multipliciteterna är $1 + 1 = 2 < 3 = \dim \mathbb{R}^3$ och matrisen är inte diagonaliserbar.

I båda exemplen ovan är de geometrisk multipliciteterna mindre än eller lika med de algebraisk multipliciteterna. Det gäller i allmänhet.

Sats 3.27

För varje egenvärde till en linjär avbildning $F : \mathcal{V} \to \mathcal{V}$ gäller att

$$1 \leq \text{geometrisk multiplicitet} \leq \text{algebraisk multiplicitet}$$

Bevis: Låt m vara den geometriska multipliciteten av ett egenvärde λ_0 till F. Att λ_0 är ett egenvärde ger att $1 \leq m$. Välj en bas $(\vec{v}_1, \ldots, \vec{v}_m)$ i $\mathcal{E}_{\lambda_0}(F)$ och utvidga till en bas $\underline{v} = (\vec{v}_1, \ldots, \vec{v}_n)$ i $\mathcal{V}$. Sätt $B = (F)^{\underline{v}}_{\underline{v}}$. Eftersom $F(\vec{v}_i) = \lambda_0 \vec{v}_i$ för $1 \leq i \leq m$, så är de m första kolonnerna i B samma som i $\lambda_0 I_n$. Därmed kan vi skriva B som en matris uppdelad i fyra block

$$B = \begin{pmatrix} \lambda_0 I_m & X \\ 0 & Y \end{pmatrix}$$

där X är någon $(m \times (n-m))$-matris och Y är någon $((n-m) \times (n-m))$-matris. Nu är

$$\chi_B(\lambda) = \det \begin{pmatrix} \lambda_0 I_m - \lambda I_m & X \\ 0 & Y - \lambda I_{n-m} \end{pmatrix} = (\lambda_0 - \lambda)^m \chi_Y(\lambda).$$

Det följer att den algebraiska multipliciteten är minst m. Notera att den skulle kunna vara större eftersom det kan hända att $\chi_Y(\lambda_0) = 0$.

En viktig konsekvens av ovanstående sats är att om den algebraiska multipliciteten av ett egenvärde λ_0 är 1 så är den geometriska multipliciteten av λ_0 också 1.

Vi är nu redo att karaktärisera när en linjär avbildning är diagonaliserbar

Sats 3.28

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning. Då är följande villkor ekvivalenta

a) F är diagonaliserbar,

b) summan av alla geometriska multipliciteter är $\dim \mathcal{V}$,

c) $\chi_F(\lambda)$ har endast reella nollställen och för varje nollställe (det vill säga egenvärde) λ är den geometriska multipliciteten samma som den algebraiska multipliciteten.

Bevis: Antag c). Vi har tidigare sett att $\chi_F(\lambda)$ är ett polynom av grad $n = \dim \mathcal{V}$. Eftersom $\chi_F(\lambda)$ endast har reella nollställen är summan av antalet reella nollställen räknade med multipliciteter precis n. Alltså är summan av alla algebraiska multipliciteter av egenvärden lika med n. Men eftersom geometriska och algebraiska multipliciteter är lika följer b).

Antag b). Då fungerar vår strategi för att diagonalisera F enligt Sats 3.25, så a) följer.

Antag a). Då finns en egenbas $\underline{v}$ till F, det vill säga $B = (F)_{\underline{v}}^{\underline{v}}$ är diagonal. Eftersom $\chi_F(\lambda) = \chi_B(\lambda)$ får vi att nollställena till $\chi_F(\lambda)$ är precis diagonalelementen i B, vilka är reella. Låt nu λ_0 vara ett nollställe med algebraisk multiplicitet m_0. Det betyder att λ_0 förekommer m_0 gånger på diagonalen i B. Alltså finns det m_0 vektorer i basen $\underline{v}$ som tillhör $\mathcal{E}_{\lambda_0}(F)$. Dessa är linjärt oberoende så $m_0 \le \dim \mathcal{E}_{\lambda_0}(F)$. Men eftersom den geometriska multipliciteten $\dim \mathcal{E}_{\lambda_0}(F)$ också uppfyller $\dim \mathcal{E}_{\lambda_0}(F) \le m_0$ har vi $\dim \mathcal{E}_{\lambda_0}(F) = m_0$. Egenvärdet λ_0 var godtyckligt så c) följer.

Av ovanstående tre fall följer att alla villkoren är ekvivalenta.

Exempel 3.6.4

Avgör vilka av följande matriser som är diagonaliserbara:

a) $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ b) $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ c) $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -5 & 7 & 0 & 0 \\ 10 & -2 & 9 & 0 \\ 1 & 2 & 3 & 4 \end{pmatrix}$

a) Vi beräknar det karaktäristiska polynomet

$$\chi_A(\lambda) = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

Alltså har A inga reella egenvärden och vi kan dra slutsatsen att a inte är diagonaliserbar .

b) Vi beräknar det karaktäristiska polynomet

$$\chi_A(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^2$$

Alltså har vi ett egenvärde $\lambda_1 = 1$ med algebraisk multiplicitet 2. För att beräkna geometrisk mulitplicitet ser vi på

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

Alltså har $(A - I)\vec{x} = \vec{0}$ lösningarna $\vec{x} = t\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $t \in \mathbb{R}$ och den geometriska mulitpliciteten av $\lambda_1 = 1$ är 1. Eftersom det är mindre än den algebraiska mulitpliciteten är A inte diagonaliserbar.

c) Vi beräknar det karaktäristiska polynomet

$$\begin{aligned}\chi_A(\lambda) &= \begin{vmatrix} 1-\lambda & 0 & 0 & 0 \\ -5 & 7-\lambda & 0 & 0 \\ 10 & -2 & 9-\lambda & 0 \\ 1 & 2 & 3 & 4-\lambda \end{vmatrix} \\ &= (1-\lambda)(7-\lambda)(9-\lambda)(4-\lambda)\end{aligned}$$

Alltså har vi fyra egenvärden $\lambda_1 = 1$, $\lambda_2 = 7$, $\lambda_3 = 9$ och $\lambda_4 = 4$. Dessa har alla algebraisk multiplicitet 1 som därför är samma som den geometriska. Vi ser också att summan av de geometriska multipliciteterna är $1 + 1 + 1 + 1 = 4$. Alltså är A diagonaliserbar.

Det sista exemplet går att generalisera till följande sats som följer direkt från Sats 3.28.

Sats 3.29

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning med $\dim \mathcal{V}$ olika egenvärden. Då är F diagonaliserbar.

4 Längder och vinklar

En vektor i rummet har en längd och en riktning. Vi kan därmed studera vinkeln mellan två vektorer eller se på avståndet mellan två vektorer, som är längden av deras differens. Än så länge har vi inte gjort något liknande i godtyckliga vektorrum. Det beror på att vi behöver en ytterligare ingrediens, nämligen skalärprodukter.

4.1 Skalärprodukter

Skalärprodukten av vektorer i rummet (eller planet) är ett praktiskt verktyg för att hantera länger och vinklar. Vi ska nu generalisera detta begrepp till godtyckliga vektorrum, så att vi kan prata om längder och vinklar även för dessa. Till skillnad från den vanliga skalärprodukten som studerats i tidigare kurser kommer vi tillåta olika skalärprodukter (till och med på samma vektorrum) så länge de uppfyller vissa räkneregler. Vilken skalärprodukt som är lämplig beror på sammanhanget. I fortsättningen kommer vi därför alltid specificera vilken skalärprodukt vi syftar på. Både längder och vinklar kommer att bero på vilken skalärprodukt vi använder. Vi börjar med att påminna om de vanliga skalärprodukten i planet och rummet.

4.1.1 Standardskalärprodukt i planet och rummet

Skalärprodukten av två vektorer är

$$\vec{u} \cdot \vec{v} = |\vec{u}||\vec{v}| \cos \alpha$$

där $0 \leq \alpha \leq \pi$ är vinkeln mellan $\vec{u}$ och $\vec{v}$.

Exempel 4.1.1

Eftersom $\cos 0 = 1$, $\cos \pi = -1$, och $\cos \frac{\pi}{2} = 0$ får vi

- $\vec{v} \cdot \vec{v} = |\vec{v}||\vec{v}|$
- $\vec{v} \cdot (-\vec{v}) = -|\vec{v}||\vec{v}|$

- Om $\alpha = \frac{\pi}{2}$, så är $\vec{u} \cdot \vec{v} = 0$.

Längden $|\vec{v}|$ av en vektor $\vec{v}$ och vinkeln α mellan två nollskilda vektorer $\vec{u}$ och $\vec{v}$ kan räknas ut med hjälp av skalärprodukten:

$$|\vec{v}| = \sqrt{\vec{v} \cdot \vec{v}}, \quad \alpha = \cos^{-1}\left(\frac{\vec{u} \cdot \vec{v}}{|\vec{u}||\vec{v}|}\right)$$

För att räkna ut skalärprodukten finns av två vektorer i rummet $\mathbb{R}^3$ finns en enkel formel.

$$\text{Om } \vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \ \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \text{ så } \vec{u} \cdot \vec{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$$

Exempel 4.1.2

Enligt formeln ovan är

$$\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ -1 \\ -2 \end{pmatrix} = 2 + 0 - 6 = -4$$

Eftersom skaläprodukten är negativ måste vinkeln α mellan vektorerna uppfylla $\cos \alpha < 0$. Alltså är det trubbig vinkel mellan vektorerna.

En av anledningarna till att det är lätt att räkna med skalärprodukten är att den uppfyller praktiska räkneregler. Dessa liknar räkneregler för produkt av reella tal, vilket är skälet att vi kallar skalärprodukten för en produkt.

- $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$
- $(\lambda \vec{u}) \cdot \vec{v} = \lambda(\vec{u} \cdot \vec{v})$
- $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$
- Om $\vec{u} \neq \vec{0}$ så $\vec{u} \cdot \vec{u} > 0$.

Reglerna ovan ligger till grund för definitionen av skalärprodukter på allmänna vektorrum, vilka vi nu ska gå igenom.

4.1.2 Skalärprodukter på allmänna vektorrum

Definition 4.1

Låt $\mathcal{V}$ vara ett vektorrum. En *skalärprodukt* på $\mathcal{V}$ är en operation som tillordnar varje par av vektorer $\vec{u}, \vec{v} \in \mathcal{V}$ en skalär $\langle \vec{u} | \vec{v} \rangle \in \mathbb{R}$ så att följande villkor gäller:

a) $\langle \vec{u} + \vec{v} | \vec{w} \rangle = \langle \vec{u} | \vec{w} \rangle + \langle \vec{v} | \vec{w} \rangle, \quad \langle \vec{u} | \vec{v} + \vec{w} \rangle = \langle \vec{u} | \vec{v} \rangle + \langle \vec{u} | \vec{w} \rangle,$

b) $\langle \lambda \vec{u} | \vec{v} \rangle = \lambda \langle \vec{u} | \vec{v} \rangle = \langle \vec{u} | \lambda \vec{v} \rangle,$

c) $\langle \vec{u} | \vec{v} \rangle = \langle \vec{v} | \vec{u} \rangle,$

d) Om $\vec{u} \neq \vec{0}$ så $\langle \vec{u} | \vec{u} \rangle > 0$.

En operation som uppfyller a) och b) sägs vara bilinjär. Uppfyller den c) kallas den symmetrisk. Observera att om c) gäller behöver vi bara kolla en av likheterna i a) respektive b) (den andra följer då automatiskt). Om den sista egenskapen d) gäller kallas operationen för positivt definit. Med andra ord är en skalärprodukt bilinjär, symmetrisk och positivt definit.

Exempel 4.1.3

På $\mathcal{V} = \mathbb{R}^3$ ger den vanliga skalärprodukten en skalärprodukt i den nya bemärkelsen $\langle \vec{u} | \vec{v} \rangle = \vec{u} \cdot \vec{v}$.

Vi har en likande operation på $\mathbb{R}^n$ för varje $n \geq 1$ som kallas standardskalärprodukten.

Definition 4.2

Standardskalärprodukten på $\mathbb{R}^n$ definieras som

$$\langle \vec{u} | \vec{v} \rangle = u_1 v_1 + \cdots + u_n v_n \quad \text{där } \vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, \ \vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}.$$

Observera att vi beräkna standardskalärprodukten genom matrismultiplikation:

$$\langle \vec{u} | \vec{v} \rangle = \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \vec{u}^\top \vec{v}.$$

Egentligen får vi ovan en (1×1)-matris, men vi kommer identifiera sådana med tal i fortsättning då ovanstående formel är mycket praktisk.

Exempel 4.1.4

Vi ska nu gå igenom några exempel på skalärprodukter på funktionsrum.

a) Låt M vara en ändlig mängd och $\mathcal{V} = \{f : M \to \mathbb{R}\}$. Då ges en skalärprodukt av

$$\langle f|g\rangle = \sum_{m \in M} f(m)g(m).$$

b) Låt $\mathcal{V} = \mathcal{C}([a,b])$, det vill säga vektorrummet av alla kontinuerliga funktioner på intervallet $[a,b]$. Då ges en skalärprodukt av

$$\langle f|g\rangle = \int_a^b f(x)g(x)dx.$$

c) Låt $\mathcal{V} = \mathcal{P}$, det vill säga vektorrummet av alla polynom. Då ges en skalärprodukt av

$$\langle f|g\rangle = \int_{-1}^1 f(x)g(x)dx.$$

d) Låt $\mathcal{V} = \mathcal{P}$ igen, men nu med med en annan skalärprodukt

$$\langle f|g\rangle = \int_0^\infty f(x)g(x)e^{-x}dx.$$

I exempel c) och d) har vi introducerat två olika skalärprodukter på samma vektorrum. Vilken ska man välja? Svaret är att det beror på sammanhang. Den första skalärprodukten beror bara på värdena som polynomen antar då $-1 \leq x \leq 1$, om det är det enda vi bryr oss om så är det en rimlig skalärprodukt. Den andra skalärprodukten tar hänsyn till värden då $0 \leq x \leq \infty$, men det finns en viktfunktion e^{-x} som gör att värdena för större och större x har mindre och mindre betydelse. Denna skalärprodukt kan vara användbar om värdena nära

$x = 0$ ska bör ges större betydelse. Observera att detta bara är två möjliga val av skalärprodukt. Varje vektorrum (bortsett från det som bara har nollvektorn) har oändligt många skalärprodukter som vi kan välja bland.

4.1.3 Längder, vinklar och trianglar

I ett vektorrum utrustat med en skalärprodukt kan vi prata om längder och vinklar. Dessa definieras på följande sätt.

Definition 4.3

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\vec{u}, \vec{v} \in \mathcal{V}$.

a) Längden av $\vec{u}$ är talet

$$|\vec{u}| = \sqrt{\langle \vec{u} | \vec{u} \rangle}.$$

b) Om $\vec{u} \neq 0$ och $\vec{v} \neq 0$ definieras vinkeln mellan $\vec{u}$ och $\vec{v}$ som

$$\cos^{-1}\left(\frac{\langle \vec{u} | \vec{v} \rangle}{|\vec{u}||\vec{v}|}\right).$$

c) Vi säger att $\vec{u}$ och $\vec{v}$ är ortogonala om

$$(\vec{u} | \vec{v}) = 0.$$

Eftersom skalärprodukten är posititivt definit är $(\vec{u}|\vec{u}) > 0$ för varje $\vec{u} \neq \vec{0}$. Det garanterar att $|\vec{u}| = \sqrt{\langle \vec{u} | \vec{u} \rangle}$ ger ett väldefinierat positivt tal. För $\vec{u} = \vec{0}$ får vi

$$\langle \vec{0} | \vec{0} \rangle = \langle 0\vec{0} | \vec{0} \rangle = 0 \langle \vec{0} | \vec{0} \rangle = 0.$$

Alltså har nollvektorn längd 0 och alla andra vektorer positiv längd.

Hur vet vi att vinkeln är väldefinierad? Det vill säga, hur vet vi att

$$-1 \leq \frac{\langle \vec{u} | \vec{v} \rangle}{|\vec{u}||\vec{v}|} \leq 1?$$

Det följer av Cauchy-Schwarz olikhet som ges i nedanstående sats.

Sats 4.4: Cauchy-Schwarz olikhet

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\vec{u}, \vec{v} \in \mathcal{V}$. Då gäller

$$|\langle \vec{u}|\vec{v}\rangle| \leq |\vec{u}||\vec{v}|$$

och $|\langle \vec{u}|\vec{v}\rangle| = |\vec{u}||\vec{v}|$ om och endast om $\vec{u}$ och $\vec{v}$ är parallella.

Vi kommer att bevisa Cauchy-Schwarz olikhet i slutet av detta avsnitt.

Exempel 4.1.5

Utrusta vektorrummet av alla polynom $\mathcal{V} = \mathcal{P}$ med skalärprodukten

$$\langle f|g\rangle = \int_{-1}^{1} f(x)g(x)dx.$$

Sätt $f(x) = 1$, $g(x) = x$ och $h(x) = 1 + x$, det vill säga $h = f + g$. Vi beräknar längderna av polynom f, g, h.

$$\begin{aligned}
\langle f|f\rangle &= \int_{-1}^{1} 1dx = 2 && \text{ger } |f| = \sqrt{2}. \\
\langle g|g\rangle &= \int_{-1}^{1} x^2dx = \frac{2}{3} && \text{ger } |g| = \sqrt{\frac{2}{3}}. \\
\langle h|h\rangle &= \int_{-1}^{1} (1+x)^2dx = \frac{8}{3} && \text{ger } |h| = 2\sqrt{\frac{2}{3}}.
\end{aligned}$$

Genom att beräkna ytterligare skalärprodukter kan vi bestämma vinklarna mellan f, g och h.

$\langle f|g\rangle = \int_{-1}^{1} xdx = 0$ ger att vinkeln mellan f och g är

$$\cos^{-1}\left(\frac{0}{\sqrt{2}\cdot\sqrt{\frac{2}{3}}}\right) = \cos^{-1}(0) = \frac{\pi}{2}.$$

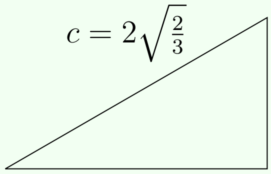
$\langle f|h\rangle = \int_{-1}^{1} (1+x)dx = 2$ ger att vinkeln mellan f och h är

$$\cos^{-1}\left(\frac{2}{\sqrt{2}\cdot 2\sqrt{\frac{2}{3}}}\right) = \cos^{-1}\left(\frac{\sqrt{3}}{2}\right) = \frac{\pi}{6}.$$

$\langle g|h\rangle = \int_{-1}^{1} x(1+x)dx = \frac{2}{3}$ ger att vinkeln mellan g och h är

$$\cos^{-1}\left(\frac{\frac{2}{3}}{\sqrt{\frac{2}{3}}\cdot 2\sqrt{\frac{2}{3}}}\right) = \cos^{-1}\left(\frac{1}{2}\right) = \frac{\pi}{3}.$$

Notera nu förhållandet $f+g=h$. Dessutom är det rät vinkel mellan f och g. Om f, g och h hade varit vektorer i planet skulle de därför bilda en rätvinklig triangel där kateterna ges av f och g medan hypotenusan av h. Faktiskt finns en sådan triangel med just de länger och vinklar som vi räknade ut.



Vi kan bekräfta det genom Pythagoras sats:

$$a^2 + b^2 = 2 + \frac{2}{3} = \frac{8}{3}, \quad c^2 = 4\frac{2}{3} = \frac{8}{3}.$$

Faktiskt är triangeln hälften av en liksidig triangel med sidlängd $2\sqrt{\frac{2}{3}}$. Alltså är de ickeräta vinklarna $\frac{\pi}{6}$ respektive $\frac{\pi}{3}$.

Exemplet ovan antyder att vi kan tänka på $\vec{u}$, $\vec{v}$ och $\vec{u}+\vec{v}$ som sidorna i en triangel trots att vektorerna är i ett helt annat vektorrum än planet $\mathbb{R}^2$ eller rummet $\mathbb{R}^3$. Vi kan jämföra längderna av $\vec{u}$, $\vec{v}$ och $\vec{u}+\vec{v}$ med hjälp av följande formel som påminner om cosinussatsen:

$$|\vec{u}+\vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2 + 2\langle\vec{u}|\vec{v}\rangle$$

Den följer från att skalärprodukten är bilinjär:

$$\begin{aligned}|\vec{u}+\vec{v}|^2 &= \langle\vec{u}+\vec{v}|\vec{u}+\vec{v}\rangle = \langle\vec{u}|\vec{u}+\vec{v}\rangle + \langle\vec{v}|\vec{u}+\vec{v}\rangle\\ &= \langle\vec{u}|\vec{u}\rangle + \langle\vec{u}|\vec{v}\rangle + \langle\vec{v}|\vec{u}\rangle + \langle\vec{v}|\vec{v}\rangle\\ &= |\vec{u}|^2 + |\vec{v}|^2 + 2(\vec{u}|\vec{v}).\end{aligned}$$

I fallet då $\vec{u}$ och $\vec{v}$ är ortogonala ger formeln ovan Pythagoras sats.

Sats 4.5: Pythagoras

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\vec{u}, \vec{v} \in \mathcal{V}$ ortogonala. Då gäller

$$|\vec{u} + \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2.$$

I en triangel kan inte någon sida vara längre än summan av de två andra sidorna. Samma sak gäller för våra trianglar av vektorer enligt följande sats.

Sats 4.6: triangelolikheten

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\vec{u}, \vec{v} \in \mathcal{V}$. Då gäller

$$|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|.$$

Bevis: Formeln för $|\vec{u} + \vec{v}|^2$ ger tillsammans med Cauchy-Schwarz olikhet att

$$\begin{aligned} |\vec{u} + \vec{v}|^2 &= |\vec{u}|^2 + |\vec{v}|^2 + 2\langle \vec{u} | \vec{v} \rangle \\ &\leq |\vec{u}|^2 + |\vec{v}|^2 + 2|\vec{u}||\vec{v}| \\ &= (|\vec{u}| + |\vec{v}|)^2 \end{aligned}$$

Då följer att $|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$.

4.1.4 Ortogonal projektion

Vi har tidigare studerat ortogonal projektion i planet och rummet. Vi generaliserar nu dessa begrepp till vektorrum utrustade med skalärprodukt.

Sats 4.7

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\vec{u}, \vec{v} \in \mathcal{V}$ så att $\vec{u} \neq \vec{0}$. Då finns entydigt bestämda vektorer $\vec{x}, \vec{y} \in \mathcal{V}$ så att

a) $\vec{x}$ är parallel med $\vec{u}$.

b) $\vec{y}$ är ortogonal med $\vec{u}$.

c) $\vec{v} = \vec{x} + \vec{y}$.

Bevis: Villkor a) betyder att $\vec{x} = t\vec{u}$ för något t. Då gäller c) om och endast om $\vec{y} = \vec{v} - t\vec{u}$. Det återstående villkoret b) säger då

$$0 = \langle \vec{y} | \vec{u} \rangle = \langle \vec{v} - t\vec{u} | \vec{u} \rangle = \langle \vec{v} | \vec{u} \rangle - t \langle \vec{u} | \vec{u} \rangle$$

vilket är uppfyllt om och endast om

$$t = \frac{\langle \vec{v} | \vec{u} \rangle}{\langle \vec{u} | \vec{u} \rangle}.$$

Beviset ger en formel för $\vec{x}$, nämligen

$$\vec{x} = \frac{\langle \vec{v} | \vec{u} \rangle}{\langle \vec{u} | \vec{u} \rangle} \vec{u}.$$

Denna vektor kallas den *ortogonala projektionen* av $\vec{v}$ på $\vec{u}$. Vi skriver därför

$$\operatorname{proj}_{\vec{u}} \vec{v} = \frac{\langle \vec{v} | \vec{u} \rangle}{\langle \vec{u} | \vec{u} \rangle} \vec{u}.$$

Notera att $\vec{y} = \vec{v} - \operatorname{proj}_{\vec{u}} \vec{v}$.

Vi använder nu ortogonal projektion för att bevisa Cauchy-Schwarz olikhet, det vill säga

$$|(\vec{u} | \vec{v})| \leq |\vec{u}||\vec{v}|.$$

Om $\vec{u} = 0$ så är $|\langle \vec{u} | \vec{v} \rangle| = 0$ och $|\vec{u}||\vec{v}| = 0$. Annars ger Pythagoras sats att

$$|\vec{v}| \geq |\operatorname{proj}_{\vec{u}} \vec{v}|$$

Men

$$|\operatorname{proj}_{\vec{u}} \vec{v}| = \frac{|\langle \vec{v} | \vec{u} \rangle|}{|\vec{u}|^2} |\vec{u}| = \frac{|\langle \vec{v} | \vec{u} \rangle|}{|\vec{u}|}.$$

Så

$$|\vec{v}||\vec{u}| \geq |\langle \vec{v} | \vec{u} \rangle|.$$

Likhet gäller om och endast om $|\vec{v}| = |\operatorname{proj}_{\vec{u}} \vec{v}|$, det vill säga $|\vec{v} - \operatorname{proj}_{\vec{u}} \vec{v}| = 0$ vilket är det samma som $\vec{v} - \operatorname{proj}_{\vec{u}} \vec{v} = \vec{0}$. Men det är precis fallet då $\vec{v} = \operatorname{proj}_{\vec{u}} \vec{v}$, alltså då $\vec{v}$ är parallell med $\vec{u}$.

4.2 Symmetriska och positivt definita matriser

Kom ihåg standardskalärprodukten av två vektorer

$$\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

i $\mathbb{R}^n$ ges av

$$\langle \vec{x} | \vec{y} \rangle = x_1 y_1 + \cdots + x_n y_n = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \vec{x}^\top \vec{y}.$$

I det här avsnittet ska vi gå igenom vilka andra möjliga skalärprodukter det går att definiera på $\mathbb{R}^n$. Innan vi gör det påminner vi om transponatet av en matris som vi redan använt för enskilda kolonnvektorer.

Definition 4.8

$$\text{Om } A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \text{ är en } (m \times n)\text{-matris,}$$

$$\text{så kallas } A^\top = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \cdots & a_{mn} \end{pmatrix} \text{ transponatet av } A.$$

Notera att kolonnerna i A står som rader i $A^\top$ och vice versa.

Exempel 4.2.1

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 4 & 5 \end{pmatrix}^\top = \begin{pmatrix} 1 & 3 \\ 2 & 4 \\ 3 & 5 \end{pmatrix}$$

Vi påminner även om följande räkneregler för transponatet av en matris.

Sats 4.9

Låt A och A' vara $(m \times k)$-matriser, och B vara en $(k \times n)$-matris. Då gäller

a) $(A^\top)^\top = A$

b) $(A + A')^\top = A^\top + A'^\top$

c) $(\lambda A)^\top = \lambda A^\top$

d) $(AB)^\top = B^\top A^\top$

4.2.1 Skalärprodukter på $\mathbb{R}^n$

Vi ska nu undersöka vilka skalärprodukter som kan definieras på $\mathbb{R}^n$. Vi inleder med ett exempel.

Exempel 4.2.2

Är följande en skalärprodukt på $\mathbb{R}^2$?

$$\langle \vec{x} | \vec{y} \rangle = 2x_1y_1 + 2x_2y_2 + x_1y_2 + x_2y_1, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \ \vec{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Vi börjar med att skriva om $\langle \vec{x} | \vec{y} \rangle$ på följande sätt

$$\begin{aligned} \text{Om } \langle \vec{x} | \vec{y} \rangle &= x_1(2y_1 + y_2) + x_2(y_1 + 2y_2) \\ &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2y_1 + y_2 \\ y_1 + 2y_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned}$$

Vi skriver

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

så att $\langle \vec{x} | \vec{y} \rangle = \vec{x}^\top A \vec{y}$. Vi kan utnyttja att varje (1×1)-matris är lika med sitt transponat för att visa att $\langle \vec{x} | \vec{y} \rangle$ är en symmetrisk operation:

$$\begin{aligned} \langle \vec{y} | \vec{x} \rangle = \vec{y}^\top A \vec{x} &= (\vec{y}^\top A \vec{x})^\top = (A\vec{x})^\top (\vec{y}^\top)^\top, \\ &= \vec{x}^\top A^\top \vec{y} = \vec{x}^\top A \vec{y} = \langle \vec{x} | \vec{y} \rangle. \end{aligned}$$

På grund av symmetri räcker det att kolla bilinjäritet i ett argument:

$$\begin{aligned}\langle\vec{x}|\vec{y}+\vec{z}\rangle &= \vec{x}^{\top}A(\vec{y}+\vec{z}) = \vec{x}^{\top}A\vec{y}+\vec{x}^{\top}A\vec{z} = \langle\vec{x}|\vec{y}\rangle+\langle\vec{x}|\vec{z}\rangle\\ \langle\vec{x}|\lambda\vec{y}\rangle &= \vec{x}^{\top}A(\lambda\vec{y}) = \lambda(\vec{x}^{\top}A\vec{y}) = \lambda\langle\vec{x}|\vec{y}\rangle\end{aligned}$$

Då återstår att undersöka om operationen är positivt definit, det vill säga $\vec{x}\neq\vec{0}$ medför $\langle\vec{x}|\vec{x}\rangle>0$. För att göra det kvadratkompletterar vi

$$\begin{aligned}\langle\vec{x}|\vec{x}\rangle &= 2x_1x_1+2x_2x_2+x_1x_2+x_2x_1\\ &= 2(x_1^2+x_1x_2+x_2^2)\\ &= 2((x_1+\tfrac{1}{2}x_2)^2-\frac{1}{4}x_2^2+x_2^2)\\ &= 2((x_1+\tfrac{1}{2}x_2)^2+\frac{3}{4}x_2^2)\geq 0.\end{aligned}$$

I den sista olikheten gäller likhet om och endast om

$$\begin{cases} x_1+\tfrac{1}{2}x_2=0,\\ x_2=0,\end{cases}$$

vilket bara sker då $\vec{x}=\vec{0}$. Alltså är operationen är positivt definit och därmed en skalärprodukt.

Motiverade av exemplet ovan gör vi följande definition.

Definition 4.10

Låt A vara en $(n\times n)$-matris. Då kallas A

a) symmetrisk om $A^{\top}=A$,

b) positivt definit om $\vec{x}^{\top}A\vec{x}>0$ för alla $\vec{x}\in\mathbb{R}^n$, $\vec{x}\neq 0$.

Då är

$$\langle\vec{x}|\vec{y}\rangle=\vec{x}^{\top}A\vec{y}$$

en skalärprodukt på $\mathbb{R}^n$.

Exempel 4.2.3

a) Matrisen $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ är symmetrisk och positivt definit. Den ger upphov till skalärprodukten vi såg i föregående exempel.

b) Enhetsmatrisen $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ är symmetrisk och positivt definit. Denna ger upphov till standardskalärprodukten på $\mathbb{R}^2$. På liknande sätt motsvarar enhetsmatrisen av storlek n standardskalärprodukten på $\mathbb{R}^n$.

c) Matrisen $A = \begin{pmatrix} 2 & 2 \\ 0 & 2 \end{pmatrix}$ är inte symmetrisk.

d) Matrisen $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ är symmetrisk och positivt definit eftersom $\vec{x}^\top A\vec{x} = 2x_1^2 + x_2^2$.

e) Matrisen $A = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ är symmetrisk men inte positivt definit eftersom $\vec{x}^\top A\vec{x} = 2x_1^2 - x_2^2$ som är negativt till exempel då $x_1 = 0$ och $x_2 = 1$.

I exemplen ovan där A är en diagonalmatris ser vi att egenvärdena (alltså elementen på diagonalen) spelar roll för om A är positivt definit eller inte. Detta samband gäller generellt. Om $\vec{x}$ är en egenvektor till A med egenvärde λ så är

$$\vec{x}^\top A\vec{x} = \vec{x}^\top(\lambda\vec{x}) = \lambda\vec{x}^\top\vec{x}.$$

Om A dessutom är positivt definit gäller alltså $\lambda > 0$.

Exempel 4.2.4

Matrisen $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ är inte positivt definit eftersom $\lambda = -1$ är ett negativt egenvärde.

Faktiskt gäller även det omvända påståendet.

Sats 4.11

Låt A vara en symmetrisk $(n \times n)$-matris. Då är A positivt definit om och endast om alla egenvärden för A är positiva.

Enligt resonemanget ovan är det klart att A positivt definit implicerar att alla egenvärden för A är positiva. Den andra implikationen följer av spektralsatsen som vi ska gå igenom senare.

4.2.2 Matrisen för en skalärprodukt

Vi har sett att matriser ger upphov till skalärprodukter på $\mathbb{R}^n$. Nu ska vi se att alla skalärprodukter kan beskrivas med matriser så länge vi väljer en bas i vårt vektorrum.

Sats 4.12

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\underline{b} = (\vec{b}_1, \ldots, \vec{b}_n)$ vara en bas i $\mathcal{V}$. Sätt

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \quad \text{där } a_{ij} = \langle \vec{b}_i | \vec{b}_j \rangle.$$

Då är A en symmetrisk positivt definit matris och

$$\langle \vec{u} | \vec{v} \rangle = (\vec{u}_{\underline{b}})^\top A \vec{v}_{\underline{b}}.$$

Bevis: Vi genomför beviset för $n = 2$. Beviset för övriga n är likande. Skriv $\vec{u}_{\underline{b}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, $\vec{v}_{\underline{b}} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$.

$$\begin{aligned} \langle \vec{u} | \vec{v} \rangle = \langle x_1\vec{b}_1 + x_2\vec{b}_2 | y_1\vec{b}_1 + y_2\vec{b}_2 \rangle &= \begin{matrix} x_1y_1\langle \vec{b}_1|\vec{b}_1\rangle + x_1y_2\langle \vec{b}_1|\vec{b}_2\rangle + \\ x_2y_1\langle \vec{b}_2|\vec{b}_1\rangle + x_2y_2\langle \vec{b}_2|\vec{b}_2\rangle \end{matrix} \\ &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (\vec{u}_{\underline{b}})^\top A \vec{v}_{\underline{b}}. \end{aligned}$$

Att matrisen A är symmetrisk följer av att $a_{ij} = \langle \vec{b}_i | \vec{b}_j \rangle = \langle \vec{b}_j | \vec{b}_i \rangle = a_{ji}$. Att A är positivt definit följer av att $(\vec{u}_{\underline{b}})^\top A \vec{u}_{\underline{b}} = \langle \vec{u} | \vec{u} \rangle > 0$, då $\vec{u} \neq 0$, vilket är det samma som $\vec{u}_{\underline{b}} \neq 0$.

Exempel 4.2.5

Är följande en skalärprodukt på $\mathbb{R}^3$?

$$\langle \vec{x}|\vec{y}\rangle = 3x_1y_1 + 4x_2y_2 + x_3y_3 + 2x_2y_3 + 2x_3y_2$$

Vi börjar med att skriva $\langle \vec{x}|\vec{y}\rangle = \vec{x}^\top A\vec{y}$ där

$$A = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 4 & 2 \\ 0 & 2 & 1 \end{pmatrix}$$

Eftersom A är symmetrisk behöver vi bara undersöka positivt definit. Vi använder återigen kvadratkomplettering.

$$\begin{aligned} \langle \vec{x}|\vec{x}\rangle &= 3x_1^2 + 4x_2^2 + x_3^2 + 4x_2x_3 \\ &= 3x_1^2 + 4(x_2^2 + x_2x_3) + x_3^2 \\ &= 3x_1^2 + 4((x_2 + \tfrac{1}{2}x_3)^2 - \frac{1}{4}x_3^2) + x_3^2 \\ &= 3x_1^2 + 4(x_2 + \tfrac{1}{2}x_3)^2 \geq 0. \end{aligned}$$

I den sista olikheten gäller likhet om och endast om

$$\begin{cases} x_1 & = 0 \\ \quad x_2 + \tfrac{1}{2}x_3 & = 0 \end{cases}$$

vilket är möjligt även om $\vec{x} \neq 0$, till exempel för

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

Alltså är operationen är inte positivt definit och därmed inte en skalärprodukt.

Exempel 4.2.6

Utrusta som tidigare vektorrummet av alla polynom $\mathcal{V} = \mathcal{P}$ med skalärprodukten

$$\langle f|g\rangle = \int_{-1}^{1} f(x)g(x)dx$$

och välj basen $\underline{b} = (1, x, x^2)$.

a) Hitta matrisen A så att $\langle f|g\rangle = f_{\underline{b}}{}^{\top} A g_{\underline{b}}$.

b) Beräkna $\langle 1 + x + x^2 | 1 - x^2\rangle$.

a) Vi utnyttjar att $\int_{-1}^{1} x^m dx = [\frac{x^{m+1}}{m+1}]_{-1}^{1} = \frac{1}{m+1} + \frac{(-1)^m}{m+1}$ och får att

$$\begin{array}{ccc} \langle 1|1\rangle = 2 & \langle 1|x\rangle = 0 & \langle 1|x^2\rangle = \frac{2}{3} \\ \langle x|1\rangle = 0 & \langle x|x\rangle = \frac{2}{3} & \langle x|x^2\rangle = 0 \\ \langle x^2|1\rangle = \frac{2}{3} & \langle x^2|x\rangle = 0 & \langle x^2|x^2\rangle = \frac{2}{5} \end{array}$$

Alltså är

$$A = \begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix}.$$

b) Vi kan nu beräkna $\langle 1+x+x^2|1-x^2\rangle$ med hjälp av koordinaterna i basen $\underline{b}$:

$$\begin{aligned} \langle 1 + x + x^2 | 1 - x^2\rangle &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \frac{4}{3} \\ 0 \\ \frac{4}{15} \end{pmatrix} = \frac{4}{3} + 0 + \frac{4}{15} = \frac{8}{5}. \end{aligned}$$

Alternativt kan vi beräkna direkt att

$$\begin{aligned} \langle 1 + x + x^2 | 1 - x^2\rangle &= \int_{-1}^{1} (1 + x + x^2)(1 - x^2) dx \\ &= \int_{-1}^{1} 1 + x - x^3 - x^4 dx \\ &= 2 - \frac{2}{5} = \frac{8}{5}. \end{aligned}$$

Att räkna ut matrisen A bara för att beräkna skalärprodukten av två vektorer är inte så praktiskt, men om vi vill räkna ut skalärprodukten av många olika vektorer så räcker det att bara räkna ut A en gång.

4.3 ON-baser

Standardbasen i $\mathbb{R}^3$ består av tre vektorer,

$$\vec{e}_1 = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad \vec{e}_3 = \begin{pmatrix}0\\0\\1\end{pmatrix},$$

som har längd 1 och är ortogonala mot varandra. Vi kan uttrycka detta på ett komprimerat sätt genom att skriva

$$\vec{e}_i \cdot \vec{e}_j = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

Att använda koordinater från en sådan bas gör det smidigt att beräkna längder och vinklar. Vi kommer därför generalisera ovanstående situation till godtyckliga vektorrum med skalärprodukt.

4.3.1 Ortonormala vektorer

Definition 4.13

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt. En uppsättning vektorer $\vec{u}_1, \ldots, \vec{u}_n \in \mathcal{V}$ kallas *ortonormal* (eller *ortonormerad*) (ON) om

$$\langle \vec{u}_i | \vec{u}_j \rangle = \begin{cases} 1 & i = j, \\ 0 & i \neq j. \end{cases}$$

I fallet då $i = j$ så säger definitionen att $|\vec{u}_i|^2 = 1$, det vill säga vektorerna $\vec{u}_i$ har längd 1. Sådana vektorer kallas normerade. I fallet $i \neq j$ säger definitionen att $\langle \vec{u}_i | \vec{u}_j \rangle = 0$, det vill säga $\vec{u}_i$ och $\vec{u}_j$ är ortogonala. Eftersom båda villkoren är uppfyllda säger vi att vektorerna är ortonormerade eller ortonormala.

Givet en vektor $\vec{v} \neq \vec{0}$ så är det lätt att ta fram en vektor med samma riktning och längd 1 genom att sätta

$$\vec{u} = \frac{1}{|\vec{v}|}\vec{v}.$$

Då är $|\vec{u}| = \frac{1}{|\vec{v}|}|\vec{v}| = 1$. Detta kallas för att normera vektorn $\vec{v}$.

Mycket blir lättare av att använda ortonormerade vektorer. En av de viktigaste anledningarna är följande.

Anmärkning 4.3.1

Om $\vec{u}_1, \ldots, \vec{u}_n$ är ON, så är

$$\langle c_1\vec{u}_1 + \cdots + c_1\vec{u}_n|\vec{u}_i\rangle = c_1\langle\vec{u}_1|\vec{u}_i\rangle + \cdots + c_n\langle\vec{u}_n|\vec{u}_i\rangle = c_i.$$

Den sista likheten följer av att $c_i\langle\vec{u}_i|\vec{u}_i\rangle = c_i$ och $c_j\langle\vec{u}_j|\vec{u}_i\rangle = 0$ om $i \neq j$.

Sats 4.14

Om $\vec{u}_1, \ldots, \vec{u}_n \in \mathcal{V}$ är ON så är $\vec{u}_1, \ldots, \vec{u}_n$ linjärt oberoende.

Bevis: Vi studerar beroendeekvationen

$$x_1\vec{u}_1 + \cdots + x_n\vec{u}_n = \vec{0}.$$

Tar vi skalärprodukt med u_i av båda led och använder Anmärkning 4.3.1 så får vi

$$x_i = \langle x_1\vec{u}_1 + \cdots + x_n\vec{u}_n|\vec{u}_i\rangle = 0.$$

Alltså är $\vec{u}_1, \ldots, \vec{u}_n$ linjärt oberoende.

Satsen ovan är praktisk då vi vill visa att en uppsättningar ortonormala vektorer utgör en bas. Närmare bestämt har vi följande sats.

Sats 4.15

Om $\vec{u}_1, \ldots, \vec{u}_n \subseteq \mathcal{V}$ är ON så är följande påståenden ekvivalenta

a) $(\vec{u}_1, \ldots, \vec{u}_n)$ är en bas i $\mathcal{V}$,

b) $\operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_n\} = \mathcal{V}$,

c) $n = \dim \mathcal{V}$.

Bevis: Sätt $\mathcal{U} = \operatorname{span}\{\vec{u}_1, \ldots, \vec{u}_n\} \subseteq \mathcal{V}$. Eftersom $\vec{u}_1, \ldots, \vec{u}_n$ är linjärt oberoende så utgör de en bas i $\mathcal{U}$, vilket ger $\dim\mathcal{U} = n$. Nu följer

$$\dim\mathcal{U} = \dim\mathcal{V} \Leftrightarrow \mathcal{U} = \mathcal{V} \Leftrightarrow (\vec{u}_1, \ldots, \vec{u}_n) \text{ är en bas i } \mathcal{V}.$$

I denna situation kallas $\underline{u} = (\vec{u}_1, \dots, \vec{u}_n)$ en ON-bas i $\mathcal{V}$.

Exempel 4.3.2

Låt $\mathcal{V} = \mathbb{R}^n$ med standardskalärprodukten. Då är standardbasen en ON-bas.

4.3.2 Koordinater i en ON-bas

Genom att välja en bas i ett vektorrum $\mathcal{V}$ kan vektorerna i $\mathcal{V}$ beskrivas med koordinater, vilket gör det möjligt att översätta frågeställningar om vektorer i $\mathcal{V}$ till motsvarande frågeställningar om vektorer i $\mathbb{R}^n$. Om vi nu utrustar $\mathcal{V}$ med en skalärprodukt och $\mathbb{R}^n$ med standardskalärprodukten så kan vi även ställa frågor och länger och längder och vinklar i $\mathcal{V}$ respektive $\mathbb{R}^n$. Om vi ser till att välja en ON-bas i $\mathcal{V}$ så kan även dessa frågor föras över från $\mathcal{V}$ till $\mathbb{R}^n$, med hjälp av följande sat.

Sats 4.16

Låt $\underline{u} = (\vec{u}_1, \dots, \vec{u}_n)$ vara en ON-bas i $\mathcal{V}$ och $\vec{v}, \vec{w} \in \mathcal{V}$. Då är

a) $\vec{v}_{\underline{u}} = \begin{pmatrix} \langle \vec{v} | \vec{u}_1 \rangle \\ \vdots \\ \langle \vec{v} | \vec{u}_n \rangle \end{pmatrix}$

b) $\langle \vec{v} | \vec{w} \rangle = \langle \vec{v} | \vec{u}_1 \rangle \langle \vec{w} | \vec{u}_1 \rangle + \cdots + \langle \vec{v} | \vec{u}_n \rangle \langle \vec{w} | \vec{u}_n \rangle = \vec{v}_{\underline{u}} \cdot \vec{w}_{\underline{u}}$

Bevis: Skriv $\vec{v} = x_1 \vec{u}_1 + \cdots + x_n \vec{u}_n$. Enligt Anmärkning 4.3.1 är

$$\langle \vec{v} | \vec{u}_i \rangle = \langle x_1 \vec{u}_1 + \cdots + x_n \vec{u}_n | \vec{u}_i \rangle = x_i$$

vilket ger a). Dessutom får vi

$$\begin{aligned} \langle \vec{v} | \vec{w} \rangle &= \langle x_1 \vec{u}_1 + \cdots + x_n \vec{u}_n | \vec{w} \rangle = x_1 \langle \vec{u}_1 | \vec{w} \rangle + \cdots + x_n \langle \vec{u}_n | \vec{w} \rangle \\ &= \langle \vec{v} | \vec{u}_1 \rangle \langle \vec{u}_1 | \vec{w} \rangle + \cdots + \langle \vec{v} | \vec{u}_n \rangle \langle \vec{u}_n | \vec{w} \rangle. \end{aligned}$$

Del a) gör det enkelt att räkna ut koordinater med avseende på en ON-bas, vi beräknar helt enkelt skalärprodukten med varje basvektor för att få motsvarande koordinat. Enligt b) kan vi med dessa koordinater sedan beräkna skalärprodukter genom att ta standadskalärprodukter av motsvarande koordinatvektorer. Vi genomför detta i ett konkret exempel.

Exempel 4.3.3

Utrusta vektorrummet $\mathcal{V} = \mathcal{P}_1$ av alla polynom av grad högst 1, med skalärprodukten

$$\langle f|g\rangle = \int_{-1}^{1} f(x)g(x)dx.$$

Då är $f_1(x) = \frac{1}{\sqrt{2}}$ och $f_2(x) = \sqrt{\frac{3}{2}}x$ ortonormala. För att se det beräknar vi

$$\begin{aligned}
\langle f_1|f_1\rangle &= \int_{-1}^{1} \frac{1}{2}dx = \frac{2}{2} = 1 && \text{alltså } |f_1| = 1.\\
\langle f_1|f_2\rangle &= \int_{-1}^{1} \frac{\sqrt{3}}{2}xdx = 0 && \text{alltså } f_1, f_2 \text{ ortogonala.}\\
\langle f_2|f_2\rangle &= \int_{-1}^{1} \frac{3}{2}x^2dx = \frac{3}{2}\cdot\frac{2}{3} = 1 && \text{alltså } |f_2| = 1.
\end{aligned}$$

Om vi nu tar ett annat polynom $g \in \mathcal{V}$, till exempel $g(x) = 1 + x$ så gäller $g = \langle g|f_1\rangle f_1 + \langle g|f_1\rangle f_2$. Vi bekräftar det genom att räkna ut

$$\begin{aligned}
\langle g|f_1\rangle &= \int_{-1}^{1} \frac{1}{\sqrt{2}}(1+x)dx = \frac{1}{\sqrt{2}}\cdot 2 = \sqrt{2}\\
\langle g|f_2\rangle &= \int_{-1}^{1} \sqrt{\frac{3}{2}}(x+x^2)dx = \sqrt{\frac{3}{2}}\cdot\frac{2}{3} = \sqrt{\frac{2}{3}}
\end{aligned}$$

som ger

$$\langle g|f_1\rangle f_1 + \langle g|f_1\rangle f_2 = \sqrt{2}\cdot\frac{1}{\sqrt{2}} + \sqrt{\frac{2}{3}}\cdot\sqrt{\frac{3}{2}}x = 1 + x = g(x).$$

4.3.3 Ortogonal projektion

Vi har tidigare sett hur skalärprodukter kan användas för att introducera ortogonal projektion av en vektor på en annan vektor. Vi ska nu generalisera detta till ortogonal projektion på ett delrum. Vi börjar med att introducera det ortogonala komplementet till $\mathcal{U}$.

Definition 4.17

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\mathcal{U} \subseteq \mathcal{V}$ ett delrum. Då är

$$\mathcal{U}^{\perp} = \{\vec{v} \in \mathcal{V} \mid \langle \vec{u}|\vec{v}\rangle = 0 \text{ för alla } \vec{u} \in \mathcal{U}\}$$

ett delrum som kallas det ortogonala komplementet till $\mathcal{U}$.

Vi bekräftar att $\mathcal{U}^{\perp}$ är ett delrum. Om $\vec{v}, \vec{w} \in \mathcal{U}^{\perp}$ så är

$$\langle \vec{u}|\vec{v} + \vec{w}\rangle = \langle \vec{u}|\vec{v}\rangle + \langle \vec{u}|\vec{w}\rangle = 0 + 0$$

så $\vec{v} + \vec{w} \in \mathcal{U}^{\perp}$. Dessutom är

$$\langle \vec{u}|\lambda\vec{v}\rangle = \lambda\langle \vec{u}|\vec{v}\rangle = 0$$

så $\lambda\vec{v} \in \mathcal{U}^{\perp}$.

Exempel 4.3.4

Låt $\mathcal{V} = \mathbb{R}^n$ med standardskalärprodukten och $\mathcal{U} = \operatorname{span}\{\vec{u}\}$ där

$$\vec{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Då är

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \in \mathcal{U}^{\perp} \text{ om och endast om } 0 = \vec{x} \cdot \vec{u} = x_1 + 2x_2 + 3x_3.$$

Därmed är $\mathcal{U}^{\perp}$ planet genom origo med normallinje $\mathcal{U}$. Vi har tidigare observerat att varje vektor $\vec{x}$ kan skrivas som en summa av två vektorer $\vec{x} = \vec{v} + \vec{w}$, där $\vec{v}$ tillhör linjen $\mathcal{U}$ och $\vec{w}$ tillhör planet $\mathcal{U}^{\perp}$. Nedan bevisar vi en mer generell version av detta påstående.

Sats 4.18: Ortogonalt projektion på delrum

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\mathcal{U} \subseteq \mathcal{V}$ ett delrum. Antag att $(\vec{u}_1, \ldots, \vec{u}_n)$ är en ON-bas i $\mathcal{U}$. Då finns för varje

$\vec{v} \in \mathcal{V}$ entydigt bestämda vektorer $\vec{x}, \vec{y} \in \mathcal{V}$ så att

a) $\vec{x} \in \mathcal{U}$,

b) $\vec{y} \in \mathcal{U}^\perp$,

c) $\vec{v} = \vec{x} + \vec{y}$.

Dessutom är $\vec{x} = \langle \vec{v} | \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v} | \vec{u}_n \rangle \vec{u}_n$.

Bevis: Sätt $\vec{x} = \langle \vec{v} | \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v} | \vec{u}_n \rangle \vec{u}_n$ och $\vec{y} = \vec{v} - \vec{x}$. Då gäller a) och c). För att visa b) räcker det att visa $\langle \vec{y} | \vec{u}_i \rangle = 0$ för varje i. Det följer av Anmärkning 4.3.1:

$$\langle \vec{y} | \vec{u}_i \rangle = \langle \vec{v} | \vec{u}_i \rangle - \langle \vec{x} | \vec{u}_i \rangle = \langle \vec{v} | \vec{u}_i \rangle - \langle \vec{v} | \vec{u}_i \rangle = 0$$

För att visa entydighet antar vi nu att $\vec{v} = \vec{x}' + \vec{y}'$ för $\vec{x}' \in \mathcal{U}$ och $\vec{y}' \in \mathcal{U}^\perp$. Då är $\vec{x} + \vec{y} = \vec{x}' + \vec{y}'$ och $\vec{x} - \vec{x}' = \vec{y}' - \vec{y}$. Men det betyder att vektorn $\vec{w} = \vec{x} - \vec{x}' = \vec{y}' - \vec{y}$ uppfyller $\vec{w} \in \mathcal{U}$ och $\vec{w} \in \mathcal{U}^\perp$. I synnerhet är $\langle \vec{w} | \vec{w} \rangle = 0$ och $\vec{w} = 0$ då skalärprodukten är positivt definit. Alltså är $\vec{x}' = \vec{x}$ och $\vec{y}' = \vec{y}$.

Vektorn $\vec{x} = \langle \vec{v} | \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v} | \vec{u}_n \rangle \vec{u}_n$ kallas den ortogonala projektionen av $\vec{v}$ på $\mathcal{U}$. Vi skriver

$$\mathrm{proj}_{\mathcal{U}}(\vec{v}) = \langle \vec{v} | \vec{u}_1 \rangle \vec{u}_1 + \cdots + \langle \vec{v} | \vec{u}_n \rangle \vec{u}_n.$$

Exempel 4.3.5

Utrusta vektorrummet $\mathcal{V} = \mathcal{C}([-1, 1])$ av alla kontinuerliga funktioner på intervallet $[-1, 1]$ med skalärprodukten

$$\langle f | g \rangle = \int_{-1}^{1} f(x) g(x) dx.$$

Vi kan betrakta $\mathcal{U} = \mathcal{P}_1$ som ett delrum av $\mathcal{V}$ genom att identifiera varje polynom med motsvarande polynomfunktion på intervallet $[-1, 1]$.

Vi har sett tidigare att $f_1(x) = \frac{1}{\sqrt{2}}$ och $f_2(x) = \sqrt{\frac{3}{2}} x$ är ortonormala. Eftersom $f_1, f_2 \in \mathcal{U}$ och $\dim \mathcal{U} = 2$ är (f_1, f_2) en ON-bas i $\mathcal{U}$.

Låt nu $h \in \mathcal{V}$ ges av $h(x) = e^x$. Den ortogonala projektionen av h på $\mathcal{U}$ är polynomet

$$f(x) = \langle h|f_1\rangle f_1 + \langle h|f_1\rangle f_2.$$

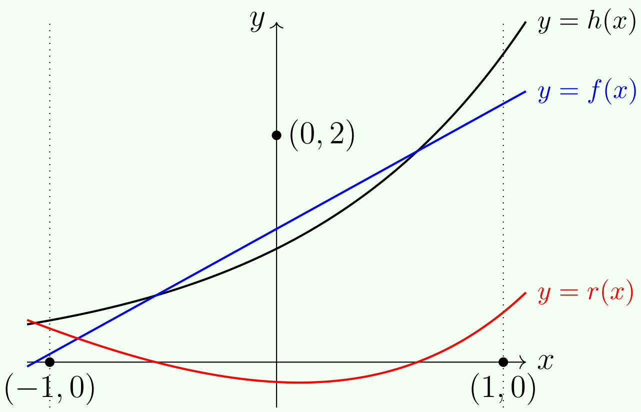
Vi beräknar

$$\begin{aligned}
\langle h|f_1\rangle &= \int_{-1}^{1} \frac{1}{\sqrt{2}} e^x dx = \frac{1}{\sqrt{2}}(e - e^{-1}) = \frac{e - e^{-1}}{\sqrt{2}}, \\
\langle h|f_2\rangle &= \int_{-1}^{1} \sqrt{\frac{3}{2}} x e^x dx = \sqrt{\frac{3}{2}} \cdot 2e^{-1} = \sqrt{6}e^{-1}.
\end{aligned}$$

Alltså är

$$f(x) = \frac{e - e^{-1}}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}} + \sqrt{6}e^{-1} \cdot \sqrt{\frac{3}{2}} x = \frac{e - e^{-1}}{2} + 3e^{-1}x$$

Den ortogonala projektionen f är det polynom av grad 1 som har minst avstånd till h, det vill säga så att $|f - h|$ är så litet som möjligt. För att illustrera plottar vi f, h och $r = f - h$ på intervallet $[-1, 1]$.



Notera att $y = r(x)$ är förhållandevis nära x-axeln, vilket beror på att linjen $y = f(x)$ är en god approximation av kurvan $y = h(x)$. Utanför intervallet $[-1, 1]$ kommer $h(x)$ och $f(x)$ avvika mer och mer från varandra. Det är förväntat eftersom vår skalärprodukt bara tar hänsyn till $x \in [-1, 1]$. Om vi vill få en god approximation på ett annat intervall behöver vi ändra till en annan skalärprodukt.

Ovanstående exempel visar på vikten av att kunna bestämma en ON-bas i ett delrum. Det finns en allmän metod för det som kallas Gram-Schmidt ortonormalisering. Metoden går ut på att från en given bas $(\vec{u}_1, \dots, \vec{u}_n)$ i ett delrum $\mathcal{U}$ steg för steg byta ut vektorerna $\vec{u}_1, \vec{u}_2, \dots$ mot ortonormala vektorer $\vec{f}_1, \vec{f}_2, \dots$ så att deras spann inte ändras. Att metoden fungerar beror på följande sats, som förklarar hur vi tar oss från steg k till $k+1$.

Sats 4.19: Gram-Schmidt ortonormalisering

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och $\vec{u}_1, \dots, \vec{u}_n \in \mathcal{V}$ linjärt oberoende. För varje $k < n$ gäller att om

$$\vec{f}_1, \dots, \vec{f}_k \text{ är en ON-bas i } \operatorname{span}\{\vec{u}_1, \dots, \vec{u}_k\}$$

så är

$$\vec{f}_1, \dots, \vec{f}_k, \vec{f}_{k+1}, \text{ är en ON-bas i } \operatorname{span}\{\vec{u}_1, \dots, \vec{u}_k, \vec{u}_{k+1}\}$$

där

$$\vec{f}_{k+1} = \frac{1}{|\vec{u}|}\vec{u}, \quad \vec{u} = \vec{u}_{k+1} - (\vec{u}_{k+1}|\vec{f}_1)\vec{f}_1 - \dots - (\vec{u}_{k+1}|\vec{f}_k)\vec{f}_k.$$

Bevis: Att $\vec{u}_1, \dots, \vec{u}_n$ är linjärt oberoende garanterar att $\vec{u} \neq 0$ så $\vec{f}_{k+1}$ är väldefinierad och har längd 1. Notera att $\vec{f}_1, \dots, \vec{f}_k, \vec{f}_{k+1} \in \operatorname{span}\{\vec{u}_1, \dots, \vec{u}_k, \vec{u}_{k+1}\}$, så det räcker att visa att $\vec{f}_1, \dots, \vec{f}_k, \vec{f}_{k+1}$ är ortonormala. Vi har enligt antagande att $\vec{f}_1, \dots, \vec{f}_k$ är ortonormala, så det återstår att visa att $\langle \vec{f}_{k+1} | \vec{f}_i \rangle = 0$ för $i \leq k$. Men det följer från $\langle \vec{u} | \vec{f}_i \rangle = 0$ som är en konsekvens av att $\vec{u}$ är $\vec{u}_{k+1}$ minus projektionen av $\vec{u}_{k+1}$ på $\operatorname{span}\left\{\vec{f}_1, \dots, \vec{f}_k, \vec{f}_{k+1}\right\}$.

Metoden är lättast att förstå genom ett exempel.

Exempel 4.3.6

Låt $\mathcal{U} \subseteq \mathbb{R}^4$ vara delrummet som ges av $\mathcal{U} = \operatorname{span}\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ där

$$\vec{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix}, \quad \vec{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}.$$

Vi använder Gram-Schmidt ortonormalisering för att hitta en ON-bas i $\mathcal{U}$.

Den första basvektorn får vi genom att normera $\vec{u}_1$ så vi sätter

$$\vec{f}_1 = \frac{1}{|\vec{u}_1|}\vec{u}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\\0\end{pmatrix}.$$

Nu använder vi satsen för $k = 1$ och beräknar

$$\begin{aligned}\vec{u}_2 - \langle\vec{u}_2|\vec{f}_1\rangle\vec{f}_1 &= \begin{pmatrix}1\\0\\2\\1\end{pmatrix} - \frac{1+0+2+0}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\\0\end{pmatrix}\\ &= \begin{pmatrix}1\\0\\2\\1\end{pmatrix} - \begin{pmatrix}1\\1\\1\\0\end{pmatrix} = \begin{pmatrix}0\\-1\\1\\1\end{pmatrix}.\end{aligned}$$

Denna vektor är ortogonal mot $\vec{f}_1$, vilket är lätt att kontrollera så vi normerar för att få

$$\vec{f}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix}0\\-1\\1\\1\end{pmatrix}.$$

Nu använder vi satsen för $k = 2$ och beräknar

$$\begin{aligned}&\vec{u}_3 - \langle\vec{u}_3|\vec{f}_1\rangle\vec{f}_1 - \langle\vec{u}_3|\vec{f}_2\rangle\vec{f}_2\\ &= \begin{pmatrix}0\\0\\1\\1\end{pmatrix} - \frac{0+0+1+0}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\\0\end{pmatrix} - \frac{0+0+1+1}{\sqrt{3}}\cdot\frac{1}{\sqrt{3}}\begin{pmatrix}0\\-1\\1\\1\end{pmatrix}\\ &= \begin{pmatrix}0\\0\\1\\1\end{pmatrix} - \frac{1}{3}\begin{pmatrix}1\\1\\1\\0\end{pmatrix} - \frac{2}{3}\begin{pmatrix}0\\-1\\1\\1\end{pmatrix} = \frac{1}{3}\begin{pmatrix}-1\\1\\0\\1\end{pmatrix}\end{aligned}$$

Denna vektor är ortogonal mot $\vec{f}_1$ och $\vec{f}_2$, vilket igen är lätt att kontrollera. Det återstår bara att normera vilket ger oss

$$\vec{f}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

Sammanfattningsvis har vi

$$\vec{f}_1 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{f}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 0 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{f}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

vilket är tre ortonormala vektorer som spänner upp $\mathcal{U}$. Alltså är $(\vec{f}_1, \vec{f}_2, \vec{f}_3)$ en ON-bas i $\mathcal{U}$.

En viktig konsekvens av Gram-Schmidt ortonormalisering är att varje vektorrum av ändlig dimension med en skalärprodukt har en ON-bas. Det följer av att vi kan ta vilken bas som helst och ortonormalisera.

Sats 4.20

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt så att $\dim \mathcal{V} < \infty$. Då finns en ON-bas i $\mathcal{V}$.

4.4 Isometrier och spektralsatsen

Vi har sett att två vektorrum $\mathcal{V}$ och $\mathcal{W}$ kan kopplas samman med hjälp av en linjär avbildning $F : \mathcal{V} \to \mathcal{W}$. Vi antar nu att $\mathcal{V}$ och $\mathcal{W}$ har utrustats med varsin skalärprodukt, så att vi kan prata om längder och vinklar i både $\mathcal{V}$ och $\mathcal{W}$. För att kunna jämföra dessa längder och vinklar måste F på något sätt vara kompatibel med skalärprodukterna. I det här avsnittet reder vi ut mer precist hur det fungerar.

4.4.1 Isometrier

En avbildning som bevarar avstånd en isometri (eller isometrisk). I den här kursen studerar vi linjära avbildningar och för dessa är att bevara avstånd det samma som att bevara längder. Mer precist gör vi följande definition.

Definition 4.21

Låt $\mathcal{V}$ och $\mathcal{W}$ vara vektorrum med skalärprodukter. En linjär avbildning

$$F : \mathcal{V} \to \mathcal{W}$$

kallas en isometri om $|F(\vec{v})| = |\vec{v}|$.

En linjär isometri bevarar även skalärprodukter och därmed även vinklar. Mer precist har vi följande sats.

Sats 4.22

F är en isometri om och endast om

$$\langle F(\vec{u})|F(\vec{v})\rangle = \langle \vec{u}|\vec{v}\rangle.$$

Bevis: Om $\langle F(\vec{u})|F(\vec{v})\rangle = \langle \vec{u}|\vec{v}\rangle$ så gäller särskilt att

$$|F(\vec{v})|^2 = \langle F(\vec{v})|F(\vec{v})\rangle = \langle \vec{v}|\vec{v}\rangle = |\vec{v}|^2,$$

vilket ger $|F(\vec{v})| = |\vec{v}|$ så F är en isometri.

Vi har tidigare stött på formeln

$$|\vec{u} + \vec{v}|^2 = |\vec{u}|^2 + |\vec{v}|^2 + 2\langle \vec{u}|\vec{v}\rangle.$$

Vi använder nu samma formel för att beräkna $|F(\vec{u} + \vec{v})|^2$:

$$|F(\vec{u} + \vec{v})|^2 = |F(\vec{u}) + F(\vec{v})|^2 = |F(\vec{u})|^2 + |F(\vec{v})|^2 + 2\langle F(\vec{u})|F(\vec{v})\rangle$$

Om F är en isometri så gäller $|F(\vec{u} + \vec{v})|^2 = |\vec{u} + \vec{v}|^2$, $|F(\vec{u})|^2 = |\vec{u}|^2$ och $|F(\vec{v})|^2 = |\vec{v}|^2$, så genom att jämföra ovanstående ekvationer får vi

$$2(F(\vec{u})|F(\vec{v})) = 2(\vec{u}|\vec{v}) \text{ och } (F(\vec{u})|F(\vec{v})) = (\vec{u}|\vec{v}).$$

Exempel 4.4.1

a) Låt $\mathcal{V} = \mathcal{W} = \mathbb{R}^2$ med standardskalärprodukten. Då är avbildningen $F : \mathbb{R}^2 \to \mathbb{R}^2$ som roterar varje vektor med vinkeln α en isometri.

b) Låt $\mathcal{V} = \mathcal{W} = \mathbb{R}^3$ med standardskalärprodukten. Låt $F : \mathbb{R}^3 \to \mathbb{R}^3$ vara spegling i planet $x_1 + 2x_2 + 3x_3 = 0$. Då är F en isometri.

c) Låt $\mathcal{V} = \mathbb{R}^2$ och $\mathcal{W} = \mathbb{R}^3$ med standardskalärprodukterna. Då är $F : \mathbb{R}^2 \to \mathbb{R}^3$ som ges av

$$F\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

en isometri. Denna avbildning realiserar helt enkelt varje vektor i $\mathbb{R}^2$ som motsvarande vektor i x_1x_2-planet.

Notera att i exemplen ovan är avbildningarna i a) och b) bijektiva, medan den i c) är injektiv. I allmänhet gäller följande.

Sats 4.23

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en isometri. Då är F injektiv. Om dessutom $\dim \mathcal{V} = \dim \mathcal{W} < \infty$, så är F bijektiv.

Bevis: Om $F(\vec{v}) = 0$ så är $|\vec{v}| = |F(\vec{v})| = 0$ och $\vec{v} = 0$. Alltså är kärnan till F noll, vilket ger att F är injektiv.

Vi har även följande satser som är lätta att bevisa.

Sats 4.24

Låt $F : \mathcal{V} \to \mathcal{W}$ vara en inverterbar isometri. Då är $F^{-1} : \mathcal{W} \to \mathcal{V}$ en isometri.

Bevis: $|F^{-1}(\vec{v})| = |F(F^{-1}(\vec{v}))| = |\vec{v}|$.

Sats 4.25

Låt $F : \mathcal{U} \to \mathcal{V}$ och $G : \mathcal{V} \to \mathcal{W}$ vara isometrier. Då är $G \circ F : \mathcal{U} \to \mathcal{W}$ en isometri.

Bevis: $|G(F(\vec{u}))| = |F(\vec{u})| = |\vec{u}|$.

En koordinatavbildning som kommer fån ON-bas ger ett viktigt exempel på en isometri.

Sats 4.26

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och en ON-bas $\underline{u} = (\vec{u}_1, \ldots, \vec{u}_n)$. Utrusta $\mathbb{R}^n$ med standardskalärprodukten. Då är koordinatavbildningen

$$F : \mathcal{V} \to \mathbb{R}^n, \quad F(\vec{x}) = \vec{x}_{\vec{u}}$$

en isometri.

Bevis: Enligt Sats 4.16 är

$$\langle \vec{v} | \vec{w} \rangle = \vec{v}_{\underline{u}} \cdot \vec{w}_{\underline{u}},$$

vilket ger att F är en isometri.

Vi avslutar detta avsnitt med att studera isometrier på $\mathbb{R}^n$ med standardskalärprodukten.

Sats 4.27

Låt $F : \mathbb{R}^n \to \mathbb{R}^n$ vara en linjär avbildning med matris A i standardbasen. Då är följande påståenden ekvivalenta

a) F är en isometri,

b) $A^\top A = I$.

Bevis: Kom ihåg att standardskalärprodukten $\mathbb{R}^n$ kan beräknas genom

$$\langle \vec{x} | \vec{y} \rangle = \vec{x}^\top \vec{y}.$$

Då gäller även

$$(F\langle \vec{x})|F(\vec{y})\rangle = (A\vec{x})^\top (A\vec{y}) = \vec{x}^\top A^\top A\vec{y}.$$

Alltså är F är en isometri om och endast om $\vec{x}^\top A^\top A\vec{y} = \vec{x}^\top \vec{y}$, vilket gäller för varje $\vec{x}$ och $\vec{y}$ om och endast om $A^\top A = I$. För att se det kan vi ta $\vec{x} = \vec{e}_i$ och $\vec{y} = \vec{e}_j$. Då blir $\vec{x}^\top A^\top A\vec{y}$ elementet på plats (i, j) i $A^\top A$.

Motiverade av ovanstående sats inför vi följande definition.

Definition 4.28

En $(n \times n)$-matris A kallas ortonormal om $A^\top A = I$.

Namnet ortonormal kommer från följande sats som ger ett antal villkor ekvivalenta med $A^\top A = I$.

Sats 4.29

För en $(n \times n)$-matris A är följande påståenden ekvivalenta.

a) $A^\top A = I$.

b) $AA^\top = I$.

c) Kolonnerna i A bildar en ON-bas.

d) Raderna i A bildar en ON-bas.

e) A är inverterbar med invers $A^\top = A^{-1}$.

Bevis: Kolonnerna i A är raderna i $A^\top$, så när vi beräknar $A^\top A$ får vi precis alla skalärprodukter av kolonnerna i A. Villkoret $A^\top A = I$ är då precis att kolonnerna är ortonormala. Därmed är a) ekvivalent med c). På liknande sätt är b) ekvivalent med d). Eftersom A är kvadratisk säger $A^\top A = I$ (eller $AA^\top = I$) att $A^\top = A^{-1}$. Alltså är a) ekvivalent med d) och b) ekvivalent med d). Nu följer att alla villkoren är ekvivalenta.

I vissa sammanhang kallas ortonormala matriser för ortogonala, vilket kan vara förvirrande då det även i dessa sammanhang ingår att kolonnerna (respektive raderna) ska ha längd 1.

Exempel 4.4.2

Låt återigen $F : \mathbb{R}^2 \to \mathbb{R}^2$ vara rotation med vinkeln α. Då är F en isometri. Vi har sett att matrisen för F i standardbasen är

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

och

$$\begin{aligned}
A^{\top}A &= \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \\
&= \begin{pmatrix} \cos^2\alpha + \sin^2\alpha & -\cos\alpha\sin\alpha + \sin\alpha\cos\alpha \\ -\sin\alpha\cos\alpha + \cos\alpha\sin\alpha & \sin^2\alpha + \cos^2\alpha \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.
\end{aligned}$$

Alltså gäller $A^{\top}A = I$. Övriga villkor i Sats 4.29 är också lätta att kolla.

Vi kan generalisera situationen ovan lite genom att undersöka på matrisen för en godtycklig avbildning i en ON-bas.

Sats 4.30

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt och en ON-bas $\underline{u} = (\vec{u}_1, \ldots, \vec{u}_n)$. Låt $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning med $(F)^{\underline{u}}_{\underline{u}} = A$. Då är F en isometri om och endast om A är ortonormal.

Bevis: Vi har sett att koordinatavbildningen från $\mathcal{V}$ till $\mathbb{R}^n$ är en isometri. Därmed är dess invers också en isometri. Genom att sammansätta med dessa avbildningar kan vi översätta mellan $F : \mathcal{V} \to \mathcal{V}$ och $G : \mathbb{R}^n \to \mathbb{R}^n$, $G(\vec{x}) = A\vec{x}$. Eftersom sammansättning av isometrier ger en isometri följer det att F är en isometri om och endast om G är det, vilket är det samma som att A är ortonormal.

4.4.2 Symmetriska avbildningar

I förgående avsnitt såg vi hur isometrier förhåller sig till ortonormala matriser. Samma förhållande gäller mellan symmetriska avbildningar och symmetriska matriser. Kom ihåg att en matris A kallas symmetrisk om $A^{\top} = A$. Det följer att A måste vara kvadratisk. Vi börjar med att översätta villkoret $A^{\top} = A$ till ett villkor som involverar skalärprodukter.

Sats 4.31

En matris A är symmetrisk om och endast om

$$(A\vec{x}) \cdot \vec{y} = \vec{x} \cdot (A\vec{y})$$

för alla $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Bevis: Vi beräknar vänsterled

$$(A\vec{x}) \cdot \vec{y} = (A\vec{x})^\top \vec{y} = \vec{x}^\top A^\top \vec{y}$$

och högerled

$$\vec{x} \cdot (A\vec{y}) = \vec{x}^\top A\vec{y}$$

Denna likhet gäller för alla $\vec{x}$ och $\vec{y}$ om och endast om $A^\top = A$.

Motiverade av denna sats gör vi följande definition.

Definition 4.32

Låt $\mathcal{V}$ vara ett vektorrum med en skalärprodukt. En linjär avbildning

$$F : \mathcal{V} \to \mathcal{V}$$

kallas symmetrisk om $\langle F(\vec{u})|\vec{v}\rangle = \langle \vec{u}|F(\vec{v})\rangle$ för alla $\vec{u}, \vec{v} \in \mathcal{V}$.

Sats 4.33

Låt $\underline{u} = (\vec{u}_1, \ldots, \vec{u}_n)$ vara en ON-bas i $\mathcal{V}$ och $F : \mathcal{V} \to \mathcal{V}$ vara en linjär avbildning med $(F)_{\underline{u}}^{\underline{u}} = A$. Då är F symmetrisk om och endast om A är symmetrisk.

Bevis: Som tidigare har vi att koordinatavbildningen från $\mathcal{V}$ till $\mathbb{R}^n$ är en isometri. Via den översätts villkoret $\langle F(\vec{u})|\vec{v}\rangle = \langle \vec{u}|F(\vec{v})\rangle$ till $(A\vec{x}) \cdot \vec{y} = \vec{x} \cdot (A\vec{y})$, vilket är ekvivalent med att $A^\top = A$.

Vi avslutar med spektralsatsen som ger en karaktärisering av symmetriska avbildningar respektive matriser. Vi formulerar den på två sätt, men de är i stort sett ekvivalenta då vi kan översätta mellan symmetriska avbildningar och symmetriska matriser genom att välja en ON-bas.

Sats 4.34

Låt $\mathcal{V}$ vara ett ändligt dimensionellt vektorrum med en skalärprodukt och $F : \mathcal{V} \to \mathcal{V}$ en linjär avbildning. Då är följande villkor ekvivalenta

a) F är symmetrisk,

b) $\mathcal{V}$ har en ON-bas av bestående av egenvektorer till F.

Sats 4.35

Låt A vara en $(n \times n)$-matris. Då är följande villkor ekvivalenta

a) A är symmetrisk,

b) Det finns en ortonormal matris T så att $D = T^{-1}AT$ är en diagonalmatris.

Bevis: Vi visar att b) implicerar a). Beviset för att b) implicerar a) är betydligt längre och gås inte igenom här. Observera att $T^{-1} = T^\top$ eftersom T är ortonormal. Villkoret $D = T^{-1}AT$ ger $A = TDT^{-1} = TDT^\top$. Nu får vi att

$$A^\top = \left(TDT^\top\right)^\top = \left(T^\top\right)^\top D^\top T^\top = TDT^\top = A.$$

Här följer $D^\top = D$ eftersom D är en diagonalmatris.

Låt A vara symmetrisk. Hur hittar vi T ortonormal och D diagonal så att $T^{-1}AT = D$ (vilket är det samma som $T^\top AT = D$)? Process att ta fram T och D kallas ortonormal diagonalisering och kan genomföras i följande steg.

a) Bestäm egenvärdena $\lambda_1, \ldots, \lambda_k$ till A. Då fås D genom att sätta egenvärdena med multipliciteter på diagonalen.

b) Bestäm för varje i en bas i varje egenrummet $\mathcal{E}_{\lambda_i}(A)$.

c) Använd Gram-Schmidt ortonormalisering för att hitta en ON-bas i $\mathcal{E}_{\lambda_i}(A)$.

d) Sätt de resulterande vektorerna som kolonner i T.

Spektralsatsen garanterar att matrisen är diagonaliserbar så den algebraiska och geometriska multipliciteten för λ_i är samma. Om denna multiplicitet är 1 har vi $\dim \mathcal{E}_{\lambda_i}(A) = 1$. I så fall är steg b) lätt: ta en godtycklig nollskild egenvektor och normera. Hur kan vi vara säkra på att vi till slut får en ON-bas, det vill säga att T är ortonormal? Det är beror på att $A^\top = A$ garanterar att egenvektorer från olika egenrum är ortogonala. Att det är så är en konsekvens av följande sats.

Sats 4.36

Låt $F : \mathcal{V} \to \mathcal{V}$ vara en symmetrisk avbildning, $F(\vec{u}) = \lambda\vec{u}$ och $F(\vec{v}) = \mu\vec{v}$. Om $\lambda \neq \mu$ så är $(\vec{u}|\vec{v}) = 0$.

Bevis: Vi har att

$$(F(\vec{u})|\vec{v}) = (\lambda\vec{u}|\vec{v}) = \lambda(\vec{u}|\vec{v}) \quad (\vec{u}|F(\vec{v})) = (\vec{u}|\mu\vec{v}) = \mu(\vec{u}|\vec{v}).$$

Eftersom F är symmetrisk är alltså $\lambda(\vec{u}|\vec{v}) = \mu(\vec{u}|\vec{v})$, vilket ger $0 = \lambda(\vec{u}|\vec{v}) - \mu(\vec{u}|\vec{v}) = (\lambda-\mu)(\vec{u}|\vec{v})$. Från $\lambda \neq \mu$ följer nu $(\vec{u}|\vec{v})$.

Exempel 4.4.3

Låt

$$A = \begin{pmatrix} 5 & 1 & 2 \\ 1 & 5 & 2 \\ -2 & 2 & 2 \end{pmatrix}$$

Vi ser att $A^\top = A$. Spektralsatsen garanterar därmed att det finns en ON-bas av egenvektorer till A, det vill säga vi kan hitta T ortonormal och D diagonal så att $T^{-1}AT = D$. Vi börjar med att bestämma egenvärdena till A.

$$\begin{aligned}
\chi_A(\lambda) &= \begin{vmatrix} 5-\lambda & 1 & 2 \\ 1 & 5-\lambda & 2 \\ -2 & 2 & 2-\lambda \end{vmatrix} \begin{matrix} \\ 1 \\ \\ \end{matrix} = \begin{vmatrix} 6-\lambda & 6-\lambda & 0 \\ 1 & 5-\lambda & 2 \\ -2 & 2 & 2-\lambda \end{vmatrix} \\
&= (6-\lambda)\begin{vmatrix} 1 & 1 & 0 \\ 1 & 5-\lambda & 2 \\ -2 & 2 & 2-\lambda \end{vmatrix} = (6-\lambda)\begin{vmatrix} 1 & 0 & 0 \\ 1 & 4-\lambda & 2 \\ -2 & 4 & 2-\lambda \end{vmatrix} \\
& \begin{matrix} -1 \end{matrix} \\
&= (6-\lambda)\begin{vmatrix} 4-\lambda & 2 \\ 4 & 2-\lambda \end{vmatrix} = (6-\lambda)((4-\lambda)(2-\lambda)-8) \\
&= (6-\lambda)(\lambda^2 - 6\lambda + 8 - 8) = (6-\lambda)(\lambda^2-6\lambda) \\
&= -\lambda(\lambda-6)^2.
\end{aligned}$$

Alltså har vi två egenvärden $\lambda_1 = 6$ med multiplicitet 2 och $\lambda_2 = 0$ med multiplicitet 1.

Nu bestämmer vi en ON-bas i varje egenrum. Vi börjar med $\lambda_1 = 6$ och beräknar

$$\left(\begin{array}{ccc|c} -1 & 1 & -2 & 0 \\ 1 & -1 & 2 & 0 \\ -2 & 2 & -4 & 0 \end{array}\right) \begin{array}{c} \\ 1\;2 \\ \\ \end{array} \sim \left(\begin{array}{ccc|c} 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} s-2t \\ s \\ t \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t\begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \qquad s, t \in R.$$

Vi sätter

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}.$$

Dessa vektorer är inte ON, så vi använder Gram-Schmidt ortonormalisering. Först normerar vi $\vec{v}_1$ och sätter

$$\vec{f}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

Sen beräknar vi

$$\vec{v}_2 - \langle \vec{v}_2 | \vec{f}_1 \rangle \vec{f}_1 = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} - \frac{-2}{\sqrt{2}} \cdot \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

och normerar

$$\vec{f}_2 = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

Vi fortsätter nu med $\lambda_2 = 0$ och beräknar

$$\left(\begin{array}{ccc|c} 5 & 1 & -2 & 0 \\ 1 & 5 & 2 & 0 \\ -2 & 2 & 2 & 0 \end{array}\right) -\tfrac{1}{2} \sim \left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 5 & 1 & -2 & 0 \\ 1 & 5 & 2 & 0 \end{array}\right) -5 \; -1$$

$$\sim \left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 6 & 3 & 0 \\ 0 & 6 & 3 & 0 \end{array}\right) -1 \; \tfrac{1}{6} \sim \left(\begin{array}{ccc|c} 1 & -1 & -1 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right) 1$$

$$\sim \left(\begin{array}{ccc|c} 1 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 \end{array}\right),$$

$$\text{Lösningar: } \vec{x} = \begin{pmatrix} t \\ -t \\ 2t \end{pmatrix} = t \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \qquad t \in R.$$

Vi sätter

$$\vec{v}_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \quad \text{och normerar} \quad \vec{f}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Vår ON-bas av egenvektorer är

$$\vec{f}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{f}_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{f}_3 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Notera att $\vec{f}_1 \cdot \vec{f}_3 = 0$ och $\vec{f}_2 \cdot \vec{f}_3 = 0$. Vi sätter

$$T = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} \end{pmatrix} \qquad D = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Då gäller $T^{\top}T = I$ och $T^{-1}AT = D$.

5 Tillämpningar

5.1 Kvadratiska former

Vi ska nu studera en viss typ av funktioner som kallas kvadratiska former.

Definition 5.1

En funktion $Q : \mathbb{R}^n \to \mathbb{R}$ kallas en *kvadratisk form* om

$$Q(\vec{x}) = \sum_{i \leq j} q_{ij} x_i x_j, \quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{där} \quad q_{ij} \in \mathbb{R}.$$

Observera att $Q(\lambda\vec{x}) = \lambda^2 Q(\vec{x})$, vilket är en förklaring till varför vi använder ordet kvadratisk.

Exempel 5.1.1

- Låt $n = 2$ och $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, då är

  $$Q(\vec{x}) = q_{11}x_1^2 + q_{22}x_2^2 + q_{12}x_1x_2$$

  en kvadratisk form.

- Låt $n = 3$ och $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$, då är

  $$Q(\vec{x}) = q_{11}x_1^2 + q_{22}x_2^2 + q_{33}x_3^2 + q_{12}x_1x_2 + q_{13}x_1x_3 + q_{23}x_2x_3$$

  en kvadratisk form.

Trots att kvadratiska former inte är linjära kan vi använda linjär algebra för att analysera dem. Vi börjar med att se hur vi kan beskriva Q med hjälp av en symmetrisk matris

5.1.1 Matrisen till en kvadratisk form

Låt $Q(\vec{x}) = \sum_{i \leq j} q_{ij} x_i x_j$ vara given. Skriv

$$a_{ij} = \begin{cases} q_{ij} & i = j \\ \frac{q_{ij}}{2} & i < j \\ \frac{q_{ji}}{2} & i > j \end{cases} \quad \text{och} \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Då är $A^\top = A$ och $Q(\vec{x}) = \vec{x}^\top A \vec{x}$.

Exempel 5.1.2

Låt $n = 2$ och $Q(\vec{x}) = 3x_1^2 + 5x_2^2 + 2x_1x_2$. Då är $A = \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix}$. Vi bekräftar att $Q(\vec{x}) = \vec{x}^\top A \vec{x}$ stämmer.

$$\begin{aligned} \vec{x}^\top A \vec{x} &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} 3x_1 + x_2 \\ x_1 + 5x_2 \end{pmatrix} = 3x_1x_1 + x_1x_2 + x_2x_1 + 5x_2x_2 = Q(\vec{x}). \end{aligned}$$

Eftersom $A^\top = A$ kan vi tillämpa spektralsatsen. Låt därför $\underline{u} = (\vec{u}_1 \cdots \vec{u}_n)$ vara en ON-bas av egenvektorer till A. Alltså gäller

$$A\vec{u}_i = \lambda_i \vec{u}_i.$$

Sätt $T = T_{\underline{e}}^{\underline{u}}$, det vill säga basbytesmatrisen från ON-basen $\underline{u}$ till standardbasen $\underline{e}$. Då gäller

$$T^{-1} = T^\top \qquad T^\top A T = D \qquad D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}.$$

Sätt $\vec{y} = \vec{x}_{\underline{u}}$. Då är $T\vec{y} = T_{\underline{e}}^{\underline{u}} (\vec{x})_{\underline{u}} = (\vec{x})_{\underline{e}}$. Men eftersom $\underline{e}$ är standardbasen har vi $(\vec{x})_{\underline{e}} = \vec{x}$. Alltså är

$$T\vec{y} = \vec{x} \qquad \vec{y} = T^\top \vec{x}.$$

Fördelen med att använda koordinaterna i basen $\underline{u}$ är att vi då får en ekvation utan blandade termer. Mer precist får vi att

$$\begin{aligned} Q(\vec{x}) &= \vec{x}^\top A \vec{x} = (T\vec{y})^\top A T \vec{y} = \vec{y}^\top T^\top A T \vec{y} = \vec{y}^\top D \vec{y} \\ &= \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \end{aligned} \qquad \text{där } \vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

Mycket information om den kvadratiska formen Q finns antalet positiva respektive negativa termer i $\lambda_1 y_1^2 + \cdots + \lambda_n y_n^2$, det vill säga antalet positiva respektive negativa egenvärden till A. Detta kallas ibland för formens signatur.

Exempel 5.1.3

Låt $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Då är

$$\chi_A(\lambda) = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda) - 1 = (3-\lambda)(1-\lambda).$$

Alltså har vi två egenvräden $\lambda_1 = 3$ och $\lambda_2 = 1$. Vi kan därmed redan nu sätta $D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$.

För att bestämma T behöver vi hitta en ON-bas av egenvektorer till A. Vi börjar med egenvärdet $\lambda_1 = 3$.

$$\left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right) \; 1 \; \boxed{-1} \sim \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Efter normering får vi första basvektorn $\vec{f}_1 = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Vi fortsätter med egenvärdet $\lambda_1 = 1$.

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right) \; \boxed{-1} \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right),$$

Efter normering får vi andra basvektorn $\vec{f}_2 = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Vi sätter

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

och kontrollerar vårt svar

$$\begin{aligned} T^{\top}AT &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 3 & -1 \\ 3 & 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} = D. \end{aligned}$$

Den kvadratiska formen $\vec{x}^{\top}A\vec{x}$ kan nu skrivas om som $3y_1^2 + y_2^2$ där $T\vec{y} = \vec{x}$.

5.1.2 Andragradskurvor

Låt $Q : \mathbb{R}^2 \to \mathbb{R}$ vara en kvadratisk form och $c \in \mathbb{R}$. Då är lösningen till ekvationen

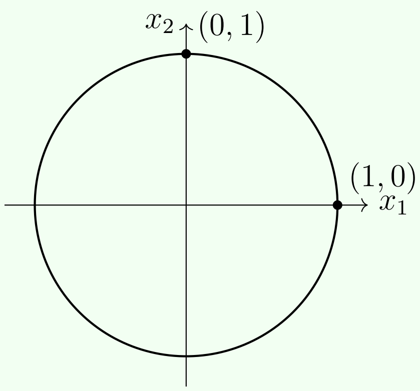
$$Q(\vec{x}) = c$$

en andragradskurva. Fallet $c = 0$ är lite speciellt och kommer inte behandlas. Om $c \neq 0$ kan vi dela båda led med c och få en likande ekvation med 1 i högerledet. Det räcker därmed att behandla fallet $c = 1$. Detta är inte alltid det smidigaste för beräkningar, men vi kommer i alla fall alltid skriva om så att $c > 0$.

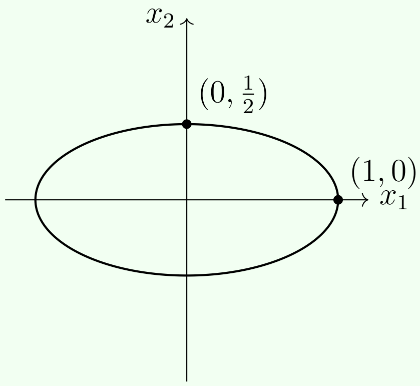
Låt oss nu studera några vägledande exempel.

Exempel 5.1.4

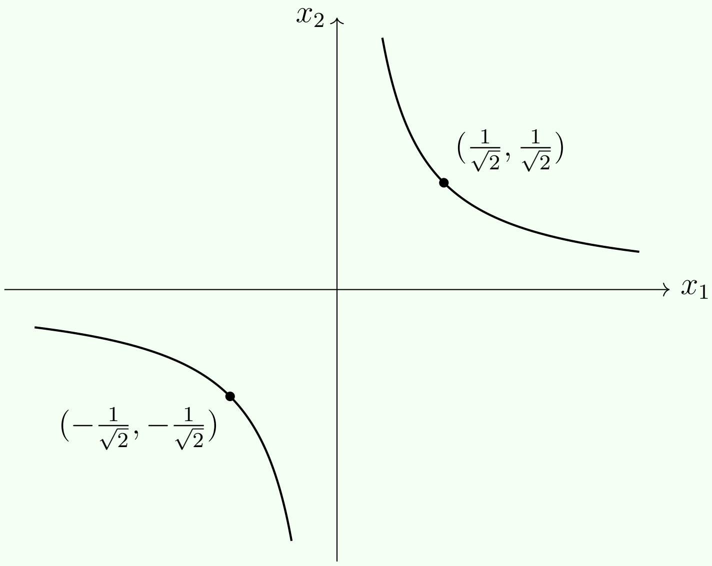
a) Ekvationen $x_1^2+x_2^2 = 1$ beskriver en andragradskurva. Närmare bestämt är detta enhetscirkeln.



b) Ekvationen $x_1^2 + 4x_2^2 = 1$ beskriver en andragradskurva. Vi kan skriva om den som $x_1^2 + (2x_2)^2 = 1$, och ser att den motsvarar föregående ekvation omskalad i x_2-led. Till exempel motsvarar lösningen $(0, 1)$ till den föregående ekvationen lösningen $(0, \frac{1}{2})$ till $x_1^2 + 4x_2^2 = 1$. Alltså beskriver ekvationen en ellips.



c) Ekvationen $2x_1x_2 = 1$ beskriver en andragradskurva. I detta fall kan vi skriva om ekvationen som $x_2 = \frac{1}{2x_1}$. Kurvan är alltså grafen till funktionen $f(x) = \frac{1}{2x}$, som är en hyperbel.



I vart och en av exemplen har vi beräknat några lösningar och markerat motsvarande punkter. Notera att kurvan i de första två exemplen ligger symmetriskt kring koordinataxlarna. Det har att göra med att vi inte har några blandade termer. För varje lösning (x_1, x_2) får vi därför tre ytterligare lösningar $(-x_1, x_2)$, $(x_1, -x_2)$, $(-x_1, -x_2)$. I nästa exempel gör vi ett koordinatbyte för att skriva hyperbeln $2x_1x_2 = 1$ på liknande sätt.

Exempel 5.1.5

Vi studerar nu åter $2x_1x_2 = 1$. Sätt $Q(\vec{x}) = 2x_1x_2$ så att $Q(\vec{x}) = \vec{x}^\top A\vec{x}$, där

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

Vi bestämmer egenvärdena till A:

$$\chi_A(\lambda) = \begin{vmatrix} 0-\lambda & 1 \\ 1 & 0-\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda-1)(\lambda+1).$$

Alltså har vi två egenvärden $\lambda_1 = 1$ och $\lambda_2 = -1$.

Vi bestämmer nu en ON-bas av egenvektorer till A. Vi börjar med egenvärdet $\lambda_1 = 1$.

$$\left(\begin{array}{cc|c} -1 & 1 & 0 \\ 1 & -1 & 0 \end{array}\right) \ 1 \ \boxed{-1} \sim \left(\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Efter normering får vi första basvektorn $\vec{f}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Vi fortsätter med egenvärdet $\lambda_1 = -1$.

$$\left(\begin{array}{cc|c} 1 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right) \ \boxed{-1} \sim \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right)$$

Efter normering får vi andra basvektorn $\vec{f}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Vi sätter

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

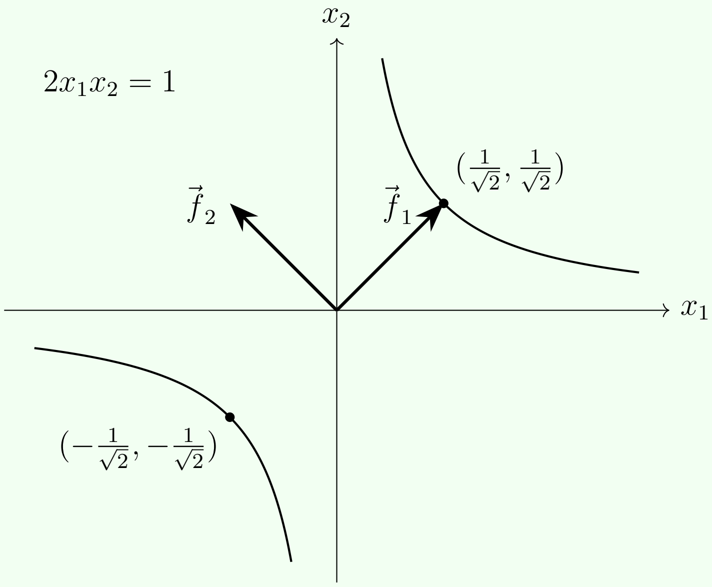
och kollar att det stämmer

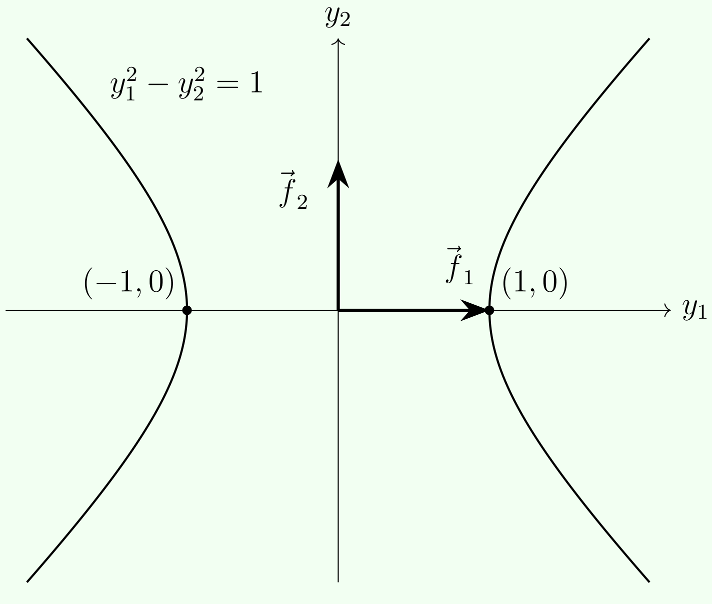
$$\begin{aligned} T^\top AT &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = D. \end{aligned}$$

Nu genomför vi vårt koordinatbyte:

$$T\vec{y} = \vec{x} \qquad \vec{y} = T^\top\vec{x}.$$

Då är $\vec{x}^\top A\vec{x} = 1$ ekvivalent med $\vec{y}^\top D\vec{y} = 1$. Med andra ord beskriver $2x_1x_2 = 1$ och $y_1^2 - y_2^2 = 1$ samma kurva i olika koordinatsystem. Vi plottar dom efter varandra och ser om det stämmer. I varje koordinatsystem ritar vi också in $\vec{f}_1$ och $\vec{f}_2$.





Som vi ser får vi x_1x_2-koordinaterna genom att rotera y_1y_2-

koordinaterna med vinkel $\frac{\pi}{4}$. Mycket riktigt är i detta exempel

$$T = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{pmatrix}.$$

I allmänhet kommer T ges av någon isometri i planet. Notera att i y_1y_2-koordinaterna ligger kurvan symmetriskt kring koordinataxlarna. Det gör det lättare att resonera om kurvan. Till exempel är det klart att punkterna $(y_1, y_2) = (1, 0)$ och $(y_1, y_2) = (-1, 0)$ är de som ligger närmast origo. Dessa punkter motsvarar $(x_1, x_2) = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ och $(x_1, x_2) = (-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$. Eftersom koordinatbytet är en isometri är avstånden samma. I båda fall ligger punkterna på avstånd 1 från origo.

Vi sammanfattar nu vad som gäller i allmänhet. Om $Q : \mathbb{R}^2 \to \mathbb{R}$ är en kvadratisk form, så finns en symmetrisk matris A så att

$$Q(\vec{x}) = 1 \Leftrightarrow \vec{x}^\top A\vec{x} = 1 \Leftrightarrow \vec{y}^\top D\vec{y} = 1 \Leftrightarrow \lambda_1 y_1^2 + \lambda_2 y_2^2 = 1$$

där

$$T^\top = T^{-1} \quad T^\top AT = D = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad T\vec{y} = \vec{x}.$$

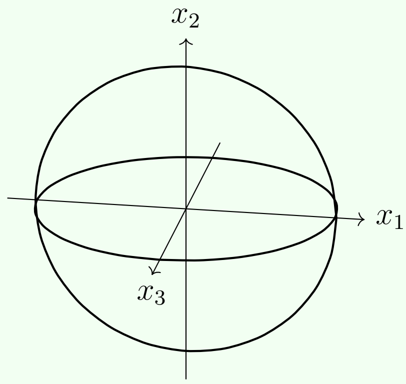
- Om $\lambda_1 > 0$ och $\lambda_2 > 0$ så beskriver $Q(\vec{x}) = 1$ en ellips.
- Om $\lambda_1 > 0$ och $\lambda_2 < 0$ så beskriver $Q(\vec{x}) = 1$ en hyperbel.

5.1.3 Andragradsytor

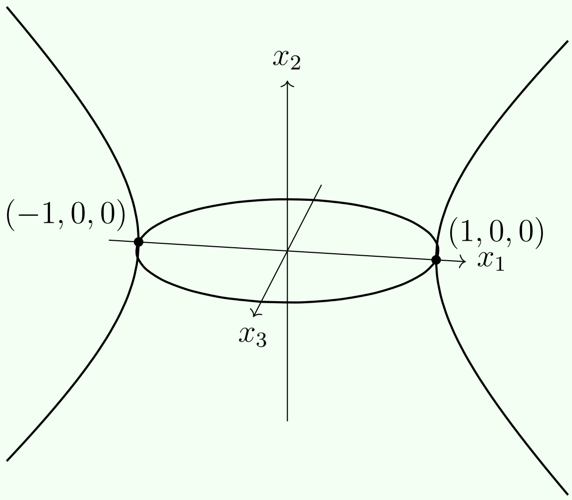
Låt $Q : \mathbb{R}^3 \to \mathbb{R}$ vara en kvadratisk form och $c \in \mathbb{R}$. Då är lösningen till ekvaionen $Q(\vec{x}) = c$ en andragradsyta. Som i fallet av andragradskurvor kommer vi anta att $c > 0$.

Exempel 5.1.6

a) Ekvationen $x_1^2 + x_2^2 + x_3^2 = 1$ beskriver en andragradsyta. Närmare bestämt är detta enhetssfären. Nedan plottar vi skärningen med x_1x_2-planet och med x_1x_3-planet som båda är cirklar.

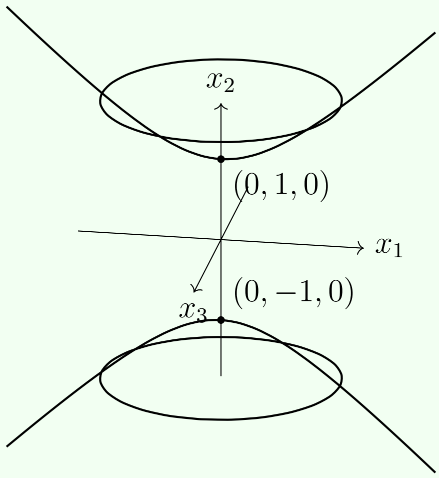


b) Ekvationen $x_1^2 - x_2^2 + x_3^2 = 1$ beskriver en andragradsyta. Vi kan skriva om den som $x_1^2 + x_3^2 = 1 + x_2^2$. Sätter vi $x_3 = 0$ får vi en hyperbel i x_1x_2-planet som skär x_1-axeln i punkterna $(1, 0, 0)$ och $(-1, 0, 0)$. Sätter vi $x_2 = 0$ så får vi en cirkel i x_1x_3-planet. Låter vi sedan x_2 variera så får vi för varje fixerat x_2-värde en cirkel i ett pan som är parallellt med x_1x_3-planet. Alltså beskriver ekvationen en hyperbelboloid. Vi plottar skärningen med x_1x_2-planet och med x_1x_3-planet.



c) Ekvationen $x_1^2 - x_2^2 + x_3^2 = -1$ beskriver en andragradsyta. Högerledet är negativt så vi börjar med att skriva om som

$-x_1^2+x_2^2-x_3^2=1$. Vi kan även skriva om som $x_1^2+x_3^2=x_2^2-1$. Sätter vi $x_3=0$ får vi en hyperbel som skär x_2-axeln i punkterna $(0,1,0)$ och $(0,-1,0)$. För varje fixerat x_2-värde större än 1 (eller mindre än -1) får vi en cirkel i ett pan som är parallellt med x_1x_3-planet. Denna yta är också en hyperbelboloid.



Notera att hyperboloiderna skiljer sig på så sätt att den första är sammanhängande och den andra har två delar. Det första fallet kallas för en enmantlad hyperboloid. Den andra fallet för en tvåmantlad hyperboloid.

Som för andragradskurvor kan vi alltid skriva om en andragradsyta $Q(\vec{x})=1$ som $\lambda_1 y_1^2+\lambda_2 y_2^2+\lambda_3 y_3^2=1$ genom en ortonormalt koordinatbyte: $T\vec{y}=\vec{x}$.

- Om $\lambda_1>0$, $\lambda_2>0$ och $\lambda_3>0$ så beskriver $Q(\vec{x})=1$ en ellipsoid.
- Om $\lambda_1>0$, $\lambda_2>0$ och $\lambda_3<0$ så beskriver $Q(\vec{x})=1$ en enmantlad hyperboloid.
- Om $\lambda_1>0$, $\lambda_2<0$ och $\lambda_3<0$ så beskriver $Q(\vec{x})=1$ en tvåmantlad hyperboloid.

I y-koordinaterna är punkterna närmast origo de som ligger på den axel som motsvarar störst egenvärde. Om till exempel λ_1 är störst är dessa punkter $(\frac{1}{\sqrt{\lambda_1}}, 0, 0)$ och $(-\frac{1}{\sqrt{\lambda_1}}, 0, 0)$. För att hitta motsvarande punkter i x-koordinaterna kan vi använda $\vec{x} = T\vec{y}$.

5.2 System av linjära differentialekvationer

I det här avsnittet ska vi använda diagonalisering för att lösa system av linjära differentialekvationer. Sådana system är vanliga i matematisk modellering och dyker upp i många tillämpningar. Vi börjar med ett exempel.

5.2.1 En enkel modell för ett vattenverk

I det här avsnittet ger vi en modell för ett vattenverk baserat på system av linjära differentialekvationer. För att det inte ska bli för komplicerat är modellen är förenklad. Vattenverket har tre delar och vatten passerar genom genom dessa i ordning med en hastighet av r kubikmeter per timme. Delarna i vattenverket är följande:

1. Luftning. Här sänks halten kolsyra i vattnet. Total volym vatten i denna del: a kubikmeter.

2. Filtrering. Här filtreras partiklar bort för att få renare vatten. Total volym vatten i denna del: b kubikmeter.

3. Desinficering. Här dödas bakterier i vattnet, till exempel genom UV-strålning. Total volym vatten i denna del: c kubikmeter.

Det har skett en olycka och vatten som förorenats med ett miljögift har kommit in i vattenverket. Allt eftersom vattnet passerar vattenverket sprider giftet sig mellan de olika delarna och till slut ut i dricksvattnet. Vi vill nu modellera detta förlopp så att vi kan säga något om var giftet befinner sig efter en viss tid t. Vi låter $y_i(t)$ beteckna mängden gift i del i efter t timmar.

Vi gör en förenkling och antar att giftet hela tiden är jämt fördelat i vardera del. Koncentrationen av gift i del 1 är $y_1(t)/a$ och eftersom vattnet passerar från del 1 till del 2 med r kubikmeter per timme kommer mängden gift förflytta sig med en hastighet av $\frac{r}{a}y_1(t)$. Eftersom mängden gift minskar är förändringen i del 1 negativ, det vill säga $y_1'(t) = -\frac{r}{a}y_1(t)$. I del 2 kommer gift in från station 1 och försvinner ut till station 3. Därför får vi $y_2'(t) = \frac{r}{a}y_1(t) - \frac{r}{b}y_2(t)$. På liknande sätt är $y_3'(t) = \frac{r}{b}y_2(t) - \frac{r}{c}y_3(t)$. Sammanfattningsvis uppfyller funktionerna $y_i(t)$ följande ekvationssystem.

$$\begin{cases} y_1'(t) &= -\frac{r}{a}y_1(t) \\ y_2'(t) &= \frac{r}{a}y_1(t) \quad -\frac{r}{b}y_2(t) \\ y_3'(t) &= \qquad\qquad \frac{r}{b}y_2(t) \quad -\frac{r}{c}y_3(t) \end{cases}$$

Eftersom varje ekvation är en differentialekvation och dessutom är linjär kallas ovanstående för ett system av linjära differentialekvationer. I nästa avsnitt går vi igenom hur dessa kan lösas mer allmänt. Efter det ska vi återkomma till vårt vattenverk och se hur det går med giftet.

5.2.2 System av linjära differentialekvationer

Ett system av linjära differentialekvationer är ett ekvationssystem på följande form:

$$\begin{cases} y_1'(t) &= a_{11}y_1(t) &+ a_{12}y_2(t) & \cdots &+ a_{1n}y_n(t) \\ y_2'(t) &= a_{21}y_1(t) &+ a_{22}y_2(t) & \cdots &+ a_{2n}y_n(t) \\ \vdots & \vdots \quad \vdots & \vdots & \ddots & \vdots \\ y_n'(t) &= a_{n1}y_n(t) &+ a_{n2}y_2(t) & \cdots &+ a_{nn}y_n(t) \end{cases}$$

Vi kan skriva om det som en matrisprodukt

$$\begin{pmatrix} y_1'(t) \\ y_2'(t) \\ \vdots \\ y_n'(t) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

som mer kortfattat skrivs

$$\vec{y}' = A\vec{y}.$$

Där $\vec{y}$ är en vektorvärd funktion och $\vec{y}'$ är dess derivata. Vi kan tänka på $\vec{y}$ som en vektor som förändras över tid. Det kan till exempel vara positionen för en partikel som rör sig i rummet (om $n = 3$) eller så beskriver $\vec{y}$ mängden gift i ett vattenverk som förändras över tid.

Det finns flera sätt att hitta lösningarna till ovanstående system. I det här avsnittet går vi igenom en metod som bygger på att diagonalisera A. Innan vi kommer dit tar vi några enkla fall.

Vi börjar med fallet $n = 1$. Då ges lösningarna av följande sats.

Sats 5.2

Lösningarna till differentialekvationen $y' = ay$ är

$$y(t) = ce^{at} \quad \text{där} \quad c \in \mathbb{R}.$$

Bevis: Om $y(t) = ce^{at}$ så är $y'(t) = ace^{at} = ay(t)$.
Låt y vara en godtycklig funktion så att $y' = ay$. Sätt

$$z(t) = y(t)e^{-at}.$$

Då är

$$\begin{aligned} z'(t) &= y'(t)e^{-at} - ay(t)e^{-at} \\ &= ay(t)e^{-at} - ay(t)e^{-at} = 0. \end{aligned}$$

Alltså är $z(t) = c$ för något $c \in \mathbb{R}$ och $y(t) = ce^{at}$.

Vi undersöker nu vad som sker ifall matrisen A redan är diagonal.

Sats 5.3

Lösningarna till differentialekvationssystemet

$$\vec{y}' = D\vec{y} \quad \text{där} \quad D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix} \quad \text{är}$$

$$\vec{y}(t) = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} \quad \text{där} \quad c_i \in \mathbb{R}.$$

Bevis: Evkationen $\vec{y}' = D\vec{y}$ motsvarar

$$\begin{cases} y_1'(t) &= & \lambda_1 y_1(t) \\ y_2'(t) &= & \lambda_2 y_2(t) \\ \vdots & \vdots & \vdots \\ y_n'(t) &= & \lambda_n y_n(t). \end{cases}$$

Varje ekvation kan lösas för sig med hjälp av föregående sats så att

$$\begin{cases} y_1(t) &= c_1e^{\lambda_1 t} \\ y_2(t) &= c_2e^{\lambda_2 t} \\ \vdots & \vdots \quad \vdots \\ y_n(t) &= c_ne^{\lambda_n t}. \end{cases}$$

5.2.3 Exempel med diagonalisering

Vi återgår nu till vårt vattenverk. För att få konkreta siffror antar vi att $\frac{r}{a} = 4$, $\frac{r}{b} = 1$, $\frac{r}{c} = 5$. Det betyder att del 2 är den som har störst volym, vilket är rimligt då filtrering tar längre tid än de andra delarna. Vi kan förvänta oss att giftet stannar kvar längre i denna del. Enligt vårt tidigare resonemang är vi intresserade av lösningarna till följande system

$$\begin{cases} y_1'(t) &= -4y_1(t), \\ y_2'(t) &= 4y_1(t) - y_2(t), \\ y_3'(t) &= y_2(t) - 5y_3(t). \end{cases}$$

Skriv om ekvationen som $\vec{y}' = A\vec{y}$ där

$$A = \begin{pmatrix} -4 & 0 & 0 \\ 4 & -1 & 0 \\ 0 & 1 & -5 \end{pmatrix}.$$

Vi visar nu hur lösningarna kan fås genom att diagonalisera A.

Genom att beräkna $\chi_A(\lambda)$ får vi fram egenvärdena $\lambda_1 = -4$, $\lambda_2 = -1$, $\lambda_3 = -5$. För varje egenvärde får vi fram egenvektorerna genom att lösa $(A - \lambda I)\vec{x} = \vec{0}$. Vi väljer en nollskild egenvektor till varje egenvärde och får en egenbas $\underline{v} = (\vec{v}_1, \vec{v}_2, \vec{v}_3)$ där

$$\vec{v}_1 = \begin{pmatrix} -3/4 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 0 \\ 4 \\ 1 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Vi sätter

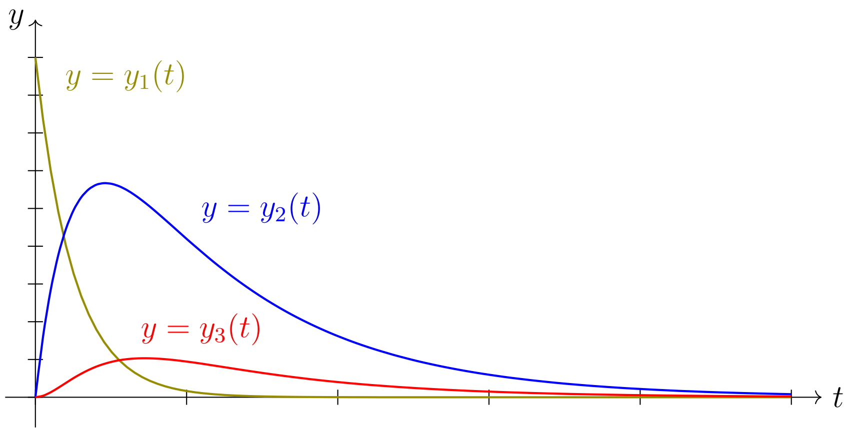
$$T = \begin{pmatrix} -3/4 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 1 & 1 \end{pmatrix} \text{ och } D = \begin{pmatrix} -4 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -5 \end{pmatrix}$$

så att
$$T^{-1}AT = D.$$
Låt $\vec{z} = T^{-1}\vec{y}$ så att $\vec{y} = T\vec{z}$ och $\vec{y}' = T\vec{z}'$. Då är
$$\vec{y}' = A\vec{y} \Leftrightarrow T\vec{z}' = AT\vec{z} \Leftrightarrow \vec{z}' = T^{-1}AT\vec{z} \Leftrightarrow \vec{z}' = D\vec{z}.$$
Alltså är
$$\vec{z}(t) = \begin{pmatrix} c_1e^{-4t} \\ c_2e^{-t} \\ c_3e^{-5t} \end{pmatrix} \text{ och } \vec{y}(t) = T\vec{z}(t) = \begin{pmatrix} -3/4 & 0 & 0 \\ 1 & 4 & 0 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} c_1e^{-4t} \\ c_2e^{-t} \\ c_3e^{-5t} \end{pmatrix}$$
Det vill säga
$$\begin{cases} y_1(t) &= -3/4c_1e^{-4t} \\ y_2(t) &= c_1e^{-4t} + 4c_2e^{-t} \\ y_3(t) &= c_1e^{-4t} + c_2e^{-t} + c_3e^{-5t} \end{cases}$$
Vilka värden c_1, c_2, c_3 antar beror på vilken lösning det handlar om. Om vi till exempel vill hitt lösningen till
$$\begin{cases} y_1'(t) &= -4y_1(t) \\ y_2'(t) &= 4y_1(t) - y_2(t) \\ y_3'(t) &= y_2(t) - 5y_3(t) \end{cases} \quad \text{där} \quad \begin{cases} y_1(0) &= 9 \\ y_2(0) &= 0 \\ y_3(0) &= 0 \end{cases}$$
så kan vi bestämma c_1, c_2, c_3 genom att sätta in $t = 0$ och matcha mot initialvärdena. Det ger oss
$$\begin{cases} -3/4c_1 &= 9 \\ c_1 + 4c_2 &= 0 \\ c_1 + c_2 + c_3 &= 0 \end{cases}$$
Observera att detta ekvationssystem på matrisfrom är precis $T\vec{c} = \vec{y}(0)$ där
$$\vec{c} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$
Genom att lösa ekvationssystemet får vi ut
$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} -12 \\ 3 \\ 9 \end{pmatrix}$$

Så att lösningen till slut är

$$\begin{cases} y_1(t) &= & 9e^{-4t} & & & & \\ y_2(t) &= & -12e^{-4t} & + & 12e^{-t} & & \\ y_3(t) &= & -12e^{-4t} & + & 3e^{-t} & + & 9e^{-5t} \end{cases}$$

Nedan plottar vi lösningarna



Som vi kan se förs giftet snabbt över från del 1 till del 2 där det stannar en längre tid. I del 3 blir nivån aldrig särskilt hög eftersom giftet försvinner ut i dricksvattnet. Vid $t = 4$ har nästan allt giftet passerat genom vattenverket.

5.2.4 Allmän lösningsmetod

I allmänhet kan vi göra på samma sätt för att lösa $\vec{y}' = A\vec{y}$ så länge A är diagonaliserbar. Uppdelat i steg är metoden följande.

- Hitta T inverterbar och D diagonal så att $T^{-1}AT = D$.
- Sätt $\vec{y} = T\vec{z}$. Då är $\vec{y}' = T\vec{z}'$ och

$$\vec{y}' = A\vec{y} \Leftrightarrow \vec{z}' = D\vec{z}$$

  Om D har diagonalelementen $\lambda_1, \ldots, \lambda_n$ så är lösningarna

$$\vec{z} = \begin{pmatrix} c_1 e^{\lambda_1 t} \\ c_2 e^{\lambda_2 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{pmatrix} \quad \text{där } c_i \in \mathbb{R}^n \quad \text{och} \quad \vec{y} = T\vec{z}$$

- Vi har att
$$\vec{z}(0) = \vec{c} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$
så lösningen som uppfyller $\vec{y}(0) = \vec{y}_0$ fås genom att välja $\vec{c} = T^{-1}\vec{y}_0$.