LINJÄR ALGEBRA OCH NUMERISK ANALYS

Ivar Gustafsson och Kjell Holmåker

Innehåll

0 Inledning

När man tillämpar matematik för att lösa konkreta problem ställs man ofta inför ekvationer av olika slag. Det kan röra sig om ekvationssystem (linjära eller icke-linjära) med ofta ett stort antal obekanta eller om differentialekvationer (ordinära eller partiella).

De första fyra kapitlen i detta kompendium handlar om linjär algebra. Detta är ett ämne som tillhandahåller en apparat för att hantera och analysera linjära ekvationer. När det gäller linjära ekvationssystem vill man t.ex. veta om ekvationerna är oberoende av varandra eller om vissa av dem är konsekvenser av de övriga, hur många lösningar det finns, hur de kan erhållas, osv. För linjära system av ordinära differentialekvationer är man likaså intresserad av hur många lösningar det finns och av metoder att finna dessa lösningar. I detta kompendium presenteras en begreppsapparat som gör det möjligt för oss att analysera och besvara sådana frågor. Begrepp och metoder från linjär algebra används idag inom allt fler ämnesområden, inte bara teknik och naturvetenskap utan också t.ex. samhällsvetenskap och ekonomi. Bland tillämpningsämnen där linjär algebra används kan nämnas Fourier-analys, matematisk fysik, partiella differentialekvationer (ortogonalitet, egenvärden), kvantmekanik (linjära operatorer, ortogonalitet, egenvärden), kretsteori, reglerteori (linjära system), osv.

Vi förutsätter att läsaren har läst en inledande kurs i ämnet, där man studerat

geometriska vektorer,
matriser,
lösning av ekvationssystem med Gauss-elimination,
determinanter.

Speciellt vill vi påminna om följande två resultat:

Sats 0.1. *Ett homogent linjärt ekvationssystem med fler obekanta än ekvationer har alltid icke-triviala lösningar.*

Sats 0.2. *En $n \times n$-matris A är inverterbar om och endast om* $\det A \neq 0$.

Den senare delen av kompendiet, kapitel 5-11 behandlar numerisk analys och numeriska metoder. Vi studerar hur vetenskapliga och tekniska problem från vitt skilda discipliner formuleras matematiskt och löses (approximativt) med hjälp av algoritmer avsedda för datorimplementering. Frågor om snabbhet hos metoderna och noggrannhet i den erhållna approximativa lösningen spelar en central roll.

Den matematiska analysen klarar bara av mycket speciella, ofta grovt förenklade problem medan de numeriska teknikerna är mycket mer generella. På så sätt utgör den numeriska analysen en brygga mellan verkligheten och matematiken.

Programsystemet MATLAB används flitigt i samband med exempel och övningar.

1 Linjära rum, vektorrum

1.1 Definition, räkneregler och exempel

I matematiken möter man i olika sammanhang storheter som på ett naturligt sätt kan adderas och multipliceras med tal så att vissa enkla räknelagar gäller. Några exempel är geometriska vektorer, vektorer i R^n, matriser i $R^{m\times n}$ och funktioner. Trots de yttre olikheterna har dessa och andra exempel någonting gemensamt, och för att fånga detta har man infört det generella begreppet linjärt rum.

Definition 1.1. Ett *linjärt rum* är en icke-tom mängd V, vars element kan adderas parvis och multipliceras med tal. Talen kallar vi *skalärer* och de kan vara reella eller komplexa. Här behandlar vi nästan enbart fallet att talen är reella, dvs så kallade *reella linjära rum.* För additionen använder vi symbolen $\oplus$ och för multiplikation med tal (skalär) använder vi symbolen $\odot$. Den exakta innebörden av detta är följande:

(1) För alla $u \in V$ och $v \in V$ finns ett väldefinierat, entydigt element $u \oplus v \in V$.

(2) För alla $u \in V$ och $\alpha \in R$ finns ett väldefinierat, entydigt element $\alpha \odot u \in V$.

Dessa två operationer, addition och multiplikation med tal (eller *skalär*) skall uppfylla följande räknelagar:

(3) $u \oplus v = v \oplus u$ för alla $u, v \in V$. (Kommutativ lag)

(4) $(u \oplus v) \oplus w = u \oplus (v \oplus w)$ för alla $u, v, w \in V$. (Associativ lag)

(5) Det finns ett element $\mathit{0} \in V$ (*nollelementet*) så att $\mathit{0} \oplus u = u \oplus \mathit{0} = u$ för alla $u \in V$.

(6) För varje $u \in V$ finns ett element $-u \in V$ (*additativ invers*) sådant att $u \oplus (-u) = (-u) \oplus u = \mathit{0}$.

(7) $\alpha \odot (\beta \odot u) = (\alpha\beta) \odot u$ för alla $\alpha,\ \beta \in R,\ u \in V$. (Associativ lag)

(8) $\alpha \odot (u \oplus v) = \alpha \odot u \oplus \alpha \odot v$ för alla $\alpha \in R,\ u,\ v \in V$. (Distributiv lag)

(9) $(\alpha + \beta) \odot u = \alpha \odot u \oplus \beta \odot u$ för alla $\alpha,\ \beta \in R,\ u \in V$. (Distributiv lag)

(10) $1 \odot u = u$ för alla $u \in V$.

Notera att vi i (7) och (9) har den vanliga multiplikationen respektive additionen i R.

Linjära rum kallas också *vektorrum.* Elementen i ett linjärt rum kallas *vektorer.*

Då skalärerna är komplexa gäller egenskaperna (axiomen) (1) – (10) med C i stället för R (där C betecknar mängden av alla komplexa tal) och man talar då om ett *komplext linjärt rum* (vektorrum). Alla resultat som vi presenterar i detta kapitel gäller även för komplexa linjära rum. All matris- och determinantkalkyl gäller också för matriser med komplexa element.

Exempel 1.1. Några exempel på linjära rum.

(a) De geometriska vektorerna i (tvådimensionella) planet eller (tredimensionella) rummet.

(b) R^n, dvs. mängden av alla n-tipler

$$x = (x_1, x_2, \ldots, x_n)$$

av reella tal, där vi definierar, med $y = (y_1, y_2, \cdots, y_n)$

$$\begin{aligned} x \oplus y &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \\ \alpha \odot x &= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n). \end{aligned}$$

Elementen i R^n representeras som *kolonnvektorer*, dvs. $n \times 1$-matriser.

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Motsvarande *radvektorer*, dvs. $1 \times n$-matrisen, skriver vi med transponat, dvs x^T.

(c) Mängden $R^{m \times n}$ av alla $m \times n$-matriser med reella element och addition och multiplikation med skalär definerade enligt den vanliga matrisalgebran, dvs $A \oplus B = A + B$ och $\alpha \odot A = \alpha A$.

(d) Mängden $F(I)$ av alla funktioner på ett intervall I. Addition och multiplikation med skalär definieras på det naturliga sättet, dvs.

$$(f \oplus g)(t) = f(t) + g(t), \quad (\alpha \odot f)(t) = \alpha f(t) \quad \text{för alla } t \in I.$$

Detta är inget konstigt, bortsett från kanske ring-operatorerna. Exempelvis är det uppenbart att vi med $\sin \oplus (3 \odot \cos) \in F(I)$ menar den funktion för vilken det gäller att $(\sin \oplus (3 \odot \cos))(t) = \sin(t) + 3\cos(t), \quad \forall t \in I$.
Nollelementet är nollfunktionen, $\mathit{0}(t) = 0$ för alla $t \in I$ och $-f$ är definierad av att $(-f)(t) = -f(t)$ för alla t i I. Axiomen (3) – (10) blir uppfyllda för $F(I)$ på grund av att motsvarande räknelagar gäller för reella (eller komplexa) tal.

(e) I (b) kan alla tal få vara komplexa, så att C^n (dvs. mängden av alla n-tipler av komplexa tal) är ett komplext linjärt rum. Om vi emellertid begränsar oss till reella skalärer, så ser vi att C^n också är ett reellt linjärt rum (dvs. ett linjärt rum med reella skalärer).

(f) Mängden R^+, dvs mängden av positiva reella tal, med operationerna $u \oplus v = uv$ och $\alpha \odot u = u^\alpha$ för $\alpha \in R$. Vi noterar först att $u \oplus v$ och $\alpha \odot u$ tillhör R^+ med de givna definitionerna. Räknereglerna är enkla att bekräfta med nollelementet $\mathit{0} = 1 \in R^+$ och additativa inversen $-u = u^{-1} \in R^+$. ■

De två första exemplen kan sägas ha stått modell för det allmänna begreppet linjärt rum. Därför används ett geometriskt språkbruk, och man kan hela tiden ha en geometrisk bild i tankarna.

I de konkreta exemplen i Exempel 1.1 är egenskaperna (1) – (10) mer eller mindre självklara. Dessutom har man i dessa fall ytterligare egenskaper som kan förefalla lika självklara. En del av dessa kan visas vara konsekvenser av egenskaperna (1) – (10) och gäller alltså för godtyckliga linjära rum, medan andra egenskaper kan vara specifika för det speciella exemplet och sakna motsvarighet i det allmänna fallet. Som en illustration skall vi härleda några konsekvenser av (1) – (10).

Vi börjar med att konstatera att entydigheten i (1) leder till att

$$u = v \Leftrightarrow u \oplus w = v \oplus w, \quad \forall w \in V \tag{1.1}$$

Vi konstaterar också att nollelementet (nollvektorn) $\mathit{0}$ i (5) är entydigt bestämd. (Observera att en formulering som "det finns ett element ... så att ... " betyder "det finns minst ett element ... så att ... ".) Om nämligen (även) $\mathit{0}^*$ uppfyller (5) dvs. $\mathit{0}^* \oplus u = u \oplus \mathit{0}^* = u$ för alla u, får vi $\mathit{0}^* \oplus u = \mathit{0} \oplus u$ som enligt (1.1) leder till $\mathit{0}^* = \mathit{0}$.

Låt oss vidare titta på en ekvation $w \oplus u = v$ (u och v givna). Den har den entydiga lösningen $w = v \oplus (-u)$, som vi skriver $w = v \ominus u$. Att w är en lösning följer av att

$$w \oplus u = (v \oplus (-u)) \oplus u \overset{(4)}{=} v \oplus ((-u) \oplus u) \overset{(6)}{=} v \oplus \mathit{0} \overset{(5)}{=} v.$$

Å andra sidan är w den enda tänkbara lösningen, ty anta att w^* är lösning. Då gäller:

$$w^* \oplus u = v \overset{(1.1)}{\Rightarrow} (w^* \oplus u) + (-u) = v \oplus (-u) \overset{(4)}{\Rightarrow}$$
$$w^* + \oplus(u \oplus (-u)) = v \oplus (-u) \overset{(6)}{\Rightarrow} w^* \oplus \mathit{0} = v \oplus (-u) \overset{(5)}{\Rightarrow}$$
$$w^* = v \oplus (-u) = w.$$

Följande allmänna konsekvenser av räknelagarna lämnas som övning:
- Den additativa inversen $-u$, definierad i (6), är entydigt bestämd.
- $0 \odot u = \mathit{0}$, där nollan i vänsterledet är det reella talet 0 och nollan i högerledet är nollvektorn i V.
- $\alpha \odot \mathit{0} = \mathit{0}$ (nollvektorn) $\forall \alpha \in R$.
- $-u = (-1) \odot u$.

För fullständighetens skull ger vi ett exempel på en mängd som inte är ett linjärt rum. Fler exempel finns bland övningarna.

Exempel 1.2. Mängden $V = R^{n\times n}$ med räknereglerna $A \oplus B = A + B$ (vanlig matrisaddition) och $\alpha \odot A = A^\alpha,\ \alpha \in R$ är inte ett linjärt rum. Det räcker att hitta en räkneregel som inte gäller. För kontroll av (9) finner vi:
$(\alpha+\beta)\odot A = A^{\alpha+\beta} = A^\alpha A^\beta$ medan $\alpha\odot A\oplus\beta\odot A = A^\alpha + A^\beta$. Eftersom för t.ex. $\alpha = \beta = 1$ och $A = I$ det gäller att $A^\alpha A^\beta \neq A^\alpha + A^\beta$ så gäller inte (9) generellt. ■

1.2 Underrum

Definition 1.2. Ett *underrum* M till ett linjärt rum V är en delmängd M av V sådan att M självt är ett linjärt rum m.a.p. samma operationer.

Eftersom räknelagarna (3) – (10) i Definition 1.1 gäller i hela V gäller de också i en delmängd M under förutsättning att $\mathit{0} \in M$. Därför behöver man bara verifiera (1) och (2) för M för att inse att M är ett underrum av V, dvs. man skall verifiera att M är slutet under addition och multiplikation med tal. Vi formulerar detta som en sats.

Sats 1.1. *En icke-tom delmängd M av ett linjärt rum V är ett underrum av V om och endast om M har egenskaperna*

$$\begin{aligned} u, v \in M \quad &\Rightarrow\ u \oplus v \in M,\\ u \in M,\ \alpha \in R \quad &\Rightarrow\ \alpha \odot u \in M, \end{aligned} \tag{1.2}$$

eller, ekvivalent, egenskapen

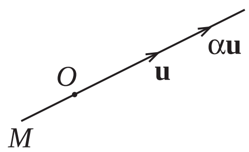
$$u, v \in M,\ \alpha, \beta \in R\ \Rightarrow\ \alpha \odot u \oplus \beta \odot v \in M.$$

Anmärkning. Tag ett $u_1 \in M$ (M är icke-tom) och $\alpha = 0$ i (1.2) så fås att $0\odot u_1 = \mathit{0} \in M$. Därmed gäller (5) för M, om (1.2) gäller. På samma sätt med $u_1 \in M$ och $\alpha = -1 \in R$ gäller att $-u_1 = (-1) \odot u_1 \in M$, så (6) gäller.

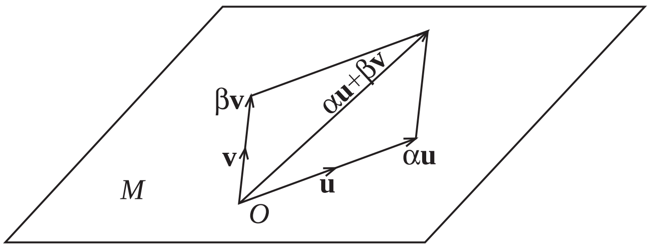
Exempel 1.3. I det åskådliga, tredimensionella rummet, där vi som vanligt identifierar punkter och ortsvektorer, är följande delmängder underrum:

(a) Origo, dvs den mängd vars enda element är $\mathit{0}$, är underrum av dimension 0.

(b) Räta linjer genom origo är underrum av dimension 1.



(c) Plan genom origo är underrum av dimension 2.



Även i allmännare situationer kan man ha denna figur i tankarna och föreställa sig ett underrum som någonting som är analogt med ett plan genom origo.

(d) Hela tredimensionella rummet. ■

Exempel 1.4. Låt V vara det linjära rummet av alla reella $n \times n$-matriser (jfr Exempel 1.1 (c)), och låt M vara den delmängd som består av alla *symmetriska* $n \times n$-matriser dvs matriser $A \in R^{n \times n}$ sådana att $A^T = A$. Då är M ett underrum av V, ty om A och B är symmetriska, så är också $A + B$ symmetrisk ty $(A+B)^T = A^T + B^T = A + B$, och αA är symmetrisk om α är ett reellt tal ty $(\alpha A)^T = \alpha A^T = \alpha A$. ■

Exempel 1.5. I tillämpningarna möter man ofta s.k. *funktionsrum*, dvs. olika underrum av $F(I)$ (se Exempel 1.1 (d)). Några exempel:

$C[a,b]$ = mängden av alla funktioner som är kontinuerliga på $[a,b]$.

$C^k(a,b)$ = mängden av alla funktioner som är k gånger kontinuerligt deriverbara i (a,b).

$C^\infty(a,b)$ = mängden av alla funktioner som har kontinuerliga derivator av alla ordningar i (a,b).

$\mathcal{P}$ = mängden av alla polynom.

$\mathcal{P}_n$ = mängden av alla polynom av grad högst n.

$L^2(a,b)$ = mängden av alla reellvärda funktioner f sådana att f^2 är integrerbar på (a,b). (Bokstaven L kommer från namnet Lebesgue och s.k. Lebesgueintegrerbara funktioner.)

För att visa att en mängd av funktioner är ett linjärt rum räcker det att verifiera att $f \oplus g$ och $\alpha \odot f$ tillhör mängden så snart f och g gör det. Mängden $C[a,b]$ blir ett linjärt rum, därför att $f \oplus g$ och $\alpha \odot f$ är kontinuerliga, om f och g är kontinuerliga. För C^k och C^∞ används också att $f \oplus g$ och $\alpha \odot f$ är deriverbara så snart f och g är det. För att visa att $L^2(a,b)$ är slutet under addition använder vi den enkla men nyttiga olikheten $\alpha\beta \le \frac{1}{2}(\alpha^2 + \beta^2)$ för godtyckliga reella tal α och β. Denna följer av att $\alpha^2 + \beta^2 - 2\alpha\beta = (\alpha - \beta)^2 \ge 0$. Vi får också $(\alpha + \beta)^2 = \alpha^2 + \beta^2 + 2\alpha\beta \le 2(\alpha^2 + \beta^2)$. Om f och g tillhör $L^2(a,b)$, så är $[f(x) + g(x)]^2 \le 2f^2(x) + 2g^2(x)$, vilket medför att $f \oplus g$ tillhör $L^2(a,b)$. Att $\alpha \odot f$ tillhör $L^2(a,b)$, om f gör det, inses omedelbart. ■

Exempel 1.6. Låt V vara det linjära rummet $C[0,1]$ (se Exempel 1.5). Sätt

$$M = \{f \in V : f(0) = f(1) = 0\}.$$

Då är M ett underrum av V, ty om f och g tillhör M och α är ett tal, så är $f \oplus g$ kontinuerlig och satisfierar

$$(f \oplus g)(0) = f(0) + g(0) = 0 + 0 = 0, \quad (f \oplus g)(1) = f(1) + g(1) = 0,$$

och $\alpha \odot f$ är kontinuerlig och satisfierar

$$(\alpha \odot f)(0) = \alpha f(0) = \alpha 0 = 0, \quad (\alpha \odot f)(1) = \alpha f(1) = \alpha 0 = 0. \qquad \blacksquare$$

I exemplen ovan rörde det sig om rum av funktioner av en variabel, men naturligtvis gäller motsvarande för funktioner av flera variabler.

Ofta är det uppenbart hur operationerna $\oplus$ och $\odot$ ska definieras och då kommer vi att utesluta ringarna om missuppfattning inte kan ske. Det gäller t.ex. i kommande exempel där det egentligen skulle stå $\oplus$ mellan funktionerna men vi skriver helt enkelt $+$. Vidare skriver vi αu där det formellt är $\alpha \odot u$.

Exempel 1.7. Ett exempel på ett randvärdesproblem för partiella differentialekvationer är: För en given kontinuerlig funktion f, bestäm en funktion u sådan att

$$\begin{cases} \dfrac{\partial^2 u}{\partial x^2} + \dfrac{\partial^2 u}{\partial y^2} = f(x,y) & \text{för } x^2 + y^2 < 1, \\ u(x,y) = 0 & \text{för } x^2 + y^2 = 1. \end{cases}$$

Det är naturligt att kräva att u är kontinuerlig för $x^2 + y^2 \leq 1$ och har kontinuerliga partiella derivator av andra ordningen för $x^2 + y^2 < 1$. Låt V beteckna mängden av alla funktioner $u(x,y)$, som har dessa egenskaper och satisfierar randvillkoret $u(x,y) = 0$ för $x^2 + y^2 = 1$. Visa att V är ett linjärt rum.

Lösning. Låt u och v tillhöra V och sätt $w = \alpha u + \beta v$. Då gäller att w är kontinuerlig för $x^2 + y^2 \leq 1$ och har kontinuerliga andraderivator för $x^2 + y^2 < 1$ (eftersom detta gäller för u och v). Vidare är

$$w(x,y) = \alpha u(x,y) + \beta v(x,y) = 0 \quad \text{för } x^2 + y^2 = 1.$$

Alltså tillhör också w mängden V, vilket visar att V är ett linjärt rum. $\blacksquare$

Exempel 1.8. Låt V vara mängden av alla reella kontinuerliga funktioner f definierade på $[0,\infty)$ sådana att $\int_0^\infty f^2(x)e^{-x}dx < \infty$. Visa att V är ett underrum av $C[0,\infty)$.

Lösning. Vi måste visa att mängden V är sluten under addition och multiplikation med skalär (jfr diskussionen av $L^2(a,b)$ i Exempel 1.5). Antag att f och g tillhör V. Vi vill visa att $\int_0^\infty [f(x)+g(x)]^2 e^{-x} dx$ är konvergent. Vi använder olikheten $(\alpha+\beta)^2 \leq 2(\alpha^2+\beta^2)$ från Exempel 1.5 och får

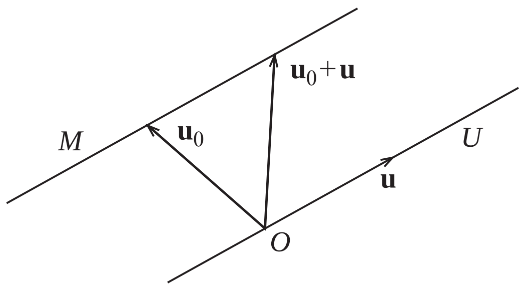
$$\int_0^\infty [f(x)+g(x)]^2 e^{-x} dx \leq 2\int_0^\infty [f^2(x)+g^2(x)]e^{-x} dx =$$
$$= 2\int_0^\infty f^2(x)e^{-x} dx + 2\int_0^\infty g^2(x)e^{-x} dx < \infty.$$

Alltså gäller att $f+g \in V$. Vi har också

$$\int_0^\infty [\alpha f(x)]^2 e^{-x} dx = \alpha^2 \int_0^\infty f^2(x)e^{-x} dx < \infty,$$

varför $\alpha f \in V$. Alltså är V ett underrum. ■

Adjektivet ”linjär” i termen ”linjärt rum” är bildat från ordet ”linje”, och därför kan det kanske tyckas konstigt att (räta) linjer (och likaså plan) i allmänhet inte är linjära rum utan bara sådana som går genom origo. I det allmännare fallet använder man ofta adjektivet ”affin”. Eftersom en godtycklig rät linje är en translation (parallellförflyttning) av en linje genom origo (se fig.), så gör vi följande definition:



Definition 1.3. En delmängd M till V kallas *affin*, om det finns en vektor $u_0 \in V$ och ett underrum U av V så att

$$M = u_0 \oplus U = \{u_0 \oplus u : u \in U\}.$$

1.3 Underrum som spänns av en mängd vektorer

Låt mängden $v_1, \ v_2, \ \ldots, v_p$ vara vektorer i det linjära rummet V och betrakta mängden av alla *linjärkombinationer* av dessa vektorer dvs vektorer på formen (som sagts ovan skippar vi ring-operatorerna)
$\alpha_1 v_1 + \alpha_2 v_2 + \ \ldots \ + \alpha_p v_p$, där $\alpha_j \in R, \ \ j = 1, \ \ldots, \ p$.
Denna mängd av linjärkombinationer, H, är ett underrum till V och betecknas
$H = Span\{v_1, \ v_2, \ \ldots, \ v_p\}$ eller $H = Span\{v_j\}_{j=1}^p$. Vi säger att underrummet spänns av de aktuella vektorerna.
Exempelvis vet vi att två vektorer i R^3, som inte är parallella, spänner upp ett underrum (ett plan) av dimension 2.

1.4 Fundamentala underrum i samband med matriser

Den grundläggande teorin för linjära ekvationssystem $Ax = b$, där $A \in R^{m\times n}$, $x \in R^n$ och $b \in R^m$ får anses bekant men vi påminner ändå om följande sats:

Sats 1.2. *Antag att x_p är en lösning till $Ax = b$ (en partikulärlösning). Då är allmänna lösningen till $Ax = b$*

$$x = x_p + x_h,$$

där x_h är allmänna lösningen till det homogena ekvationssystemet $Ax = 0$.

Vid studiet av linjära ekvationssystem spelar två speciella linjära underrum en viktig roll.

Definition 1.4. Låt $A \in R^{m\times n}$.
Nollrummet för A är mängden

$$N(A) = \{x \in R^n : Ax = 0\},$$

dvs. alla lösningar till det homogena ekvationssystemet $Ax = 0$.
Värderummet för A är mängden

$$V(A) = \{y \in R^m : \text{det finns ett } x \in R^n \text{ så att } Ax = y\} = \{Ax : x \in R^n\},$$

dvs. $V(A)$ är mängden av alla y för vilka ekvationssystemet $Ax = y$ är lösbart.

Att $N(A)$ är ett underrum till R^n visas genom egenskaperna (1.2) i Sats 1.1 och att $V(A)$ är underrum till R^m inses enklast genom att $V(A) = Span\{a_1,\ a_2\ ,\ \ldots\ ,\ a_n\}$ där $a_j,\ j = 1,\ \ldots\ ,\ n$ är kolonnerna i A. Av denna anledning kallas även $V(A)$ för *kolonnrummet* till A.

På samma sätt definieras *radrummet* för A, med beteckninig $Row(A)$, som det underrum av R^n som spänns av raderna i A. Uppenbarligen är radrummet till A lika med kolonnrummet till A^T.

Av sats 1.2 framgår att lösningsmängden till systemet $Ax = b$ är $x_p + N(A)$, en affin mängd alltså, enligt Definition 1.3.

I MATLAB får man en ortogonal bas för kolonnrummet till A genom kommandot *orth(A)* och en ortogonal bas för nollrummet till A genom kommandot *null(A)*.

Det är på sin plats att påpeka att man kan utföra beräkningen ”matris gånger vektor” på olika sätt. Produkten $b = Ax$, där A har elementen a_{ij}, kan utföras som ”rader i A gånger x” dvs.
(1) $\quad b_i = \sum_j a_{ij}x_j,$
eller som
(2) $\quad b = \sum_j x_j a_j,$

där $a_j,\ i = 1,\ \ldots\ ,\ n$ är kolonnerna i A, dvs b beräknas som en linjärkombination av A:s kolonner.

Vi vet att Gausselimination (rad-reduktion) bygger på elementära *radoperationer*. Metoden fungerar eftersom radrummet förblir oförändrat vid dessa elementära radoperationer, beviset lämnas som övning.

Naturligtvis kan man också utföra motsvarande kolonnoperationer. Vi gör en formell definition.

Definition 1.5. Låt A vara en matris. En (elementär) *kolonnoperation* är en transformation av A av någon av följande tre typer:

(i) Att låta två kolonner byta plats.

(ii) Att multiplicera en kolonn med ett tal $c \neq 0$.

(iii) Att addera en multipel av en kolonn till en annan kolonn.

Eftersom kolonnoperationer på A motsvarar radoperationer på A^T så gäller uppenbarligen också att kolonnrummet (värderummet) förblir oförändrat vid elementära kolonnoperationer.

I följande exempel använder vi begrepp som ”linjärt oberoende”, ”att spänna ett rum” och ”bas”. Det förutsätts att dessa begrepp är bekanta vad gäller rummet R^n. För allmänna linjära rum tas de upp i avsnitt 1.6.

Exempel 1.9. Låt

$$A = \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ -1 & 2 & 3 & 0 & 5 \\ 2 & 1 & -1 & -3 & -5 \end{pmatrix}.$$

(a) Bestäm en bas för $N(A)$. (b) Bestäm en bas för $V(A)$.

Lösning. (a) $N(A)$ bestäms genom att man löser det homogena ekvationssystemet $Ax = \mathit{0}$. Detta sker genom att koefficientmatrisen undergår elementära radoperationer ($\overset{\text{RE}}{\sim}$ kan utläsas ”är radekvivalent med”):

$$A = \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ -1 & 2 & 3 & 0 & 5 \\ 2 & 1 & -1 & -3 & -5 \end{pmatrix} \boxed{1}\,\boxed{-2} \overset{\text{RE}}{\sim} \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 5 & 5 & -1 & 5 \\ 0 & -5 & -5 & -1 & -5 \end{pmatrix} \boxed{-5}\,\boxed{5}$$

$$\overset{\text{RE}}{\sim} \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & -11 & 0 \\ 0 & 0 & 0 & 9 & 0 \end{pmatrix} \boxed{-\tfrac{1}{11}}\,\boxed{-9} \quad \overset{\text{RE}}{\sim} \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = A'.$$

En radoperation ändrar inte ekvationssystemets lösningsmängd, så $Ax = 0$ är ekvivalent med $A'x = 0$. Av utseendet på A' framgår att x_3 och x_5 bör väljas som parametrar, varefter x_1, x_2 och x_4 blir bestämda. Man får lösningen

$$\begin{cases} x_1 = s + 3t \\ x_2 = -s - t \\ x_3 = s \\ x_4 = 0 \\ x_5 = t \end{cases}$$

eller i vektorform $x = su_1 + tu_2$, där

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 3 \\ -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Man ser att u_1 och u_2 är linjärt oberoende, ty om $su_1 + tu_2 = 0 = x$, så ger tredje och femte ekvationerna direkt $s = x_3 = 0$, $t = x_5 = 0$. Alltså är u_1 och u_2 en bas för $N(A)$.

(b) $V(A)$ kan bestämmas genom att man utför elementära kolonnoperationer och stöder sig på det faktum att dessa inte förändrar kolonnrummet. Kolonnekvivalens betecknas med $\overset{\text{KE}}{\sim}$.

$$A = \begin{pmatrix} 1 & 3 & 2 & -1 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ -1 & 2 & 3 & 0 & 5 \\ 2 & 1 & -1 & -3 & -5 \end{pmatrix} \overset{\text{KE}}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 2 & 1 \\ -1 & 5 & 5 & -1 & 5 \\ 2 & -5 & -5 & -1 & -5 \end{pmatrix}$$

$$\boxed{-3} \quad \boxed{-2} \quad \boxed{1} \qquad\qquad \boxed{-1} \quad \boxed{-2} \quad \boxed{-1}$$

$$\overset{\text{KE}}{\sim} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -1 & 5 & 0 & -11 & 0 \\ 2 & -5 & 0 & 9 & 0 \end{pmatrix} = \tilde{A}.$$

Det är tydligt att första, andra och fjärde kolonnerna i $\tilde{A}$ är linjärt oberoende och spänner kolonnrummet för $\tilde{A}$. Alltså är

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \\ 2 \end{pmatrix}, \; v_2 = \begin{pmatrix} 0 \\ 1 \\ 5 \\ -5 \end{pmatrix}, \; v_3 = \begin{pmatrix} 0 \\ 0 \\ -11 \\ 9 \end{pmatrix}$$

en bas för $V(A) = V(\tilde{A})$. ■

1.5 Linjära avbildningar, linjära transformationer

Definition 1.6. Låt U och V vara reella linjära rum. En avbildning (funktion, tranformation) från U till V kallas *linjär*, om

$$F(u \oplus v) = F(u) \oplus F(v) \quad \text{för alla } u, v \in U$$

och

$$F(\alpha \odot u) = \alpha \odot F(u) \quad \text{för alla } \alpha \in R,\ u \in U,$$

eller ekvivalent, om vi även skippar ring-operatorerna,

$$F(\alpha u + \beta v) = \alpha F(u) + \beta F(v) \quad \text{för alla } \alpha, \beta \in R,\ u, v \in U. \tag{1.3}$$

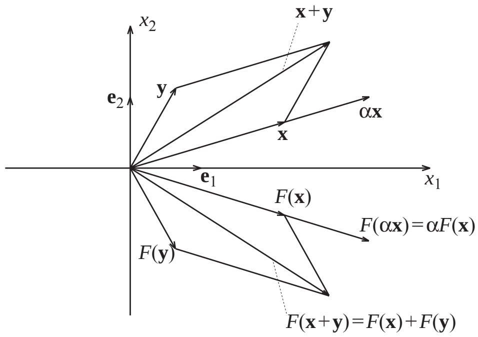
Om $U = V$ säger man att F är en linjär avbildning på V.

Genom upprepad användning av (1.3) finner man att

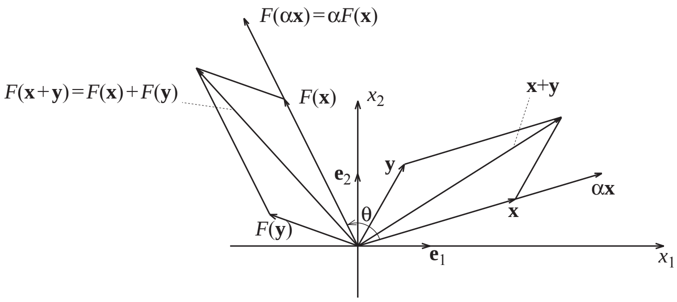
$$F\big(\sum_{i=1}^{k} \alpha_i u_i\big) = \sum_{i=1}^{k} \alpha_i F(u_i) \tag{1.4}$$

för godtyckliga $\alpha_i \in R$, $u_i \in U$, $k \geq 1$. Vidare ger $\alpha = 0$ i definitionen att $F(\mathit{0}) = \mathit{0}$.

Exempel 1.10. Spegling i en linje genom origo i R^2. I figuren illustreras spegling i x_1-axeln.



Exempel 1.11. $U = V =$ ett plan med en ON-bas e_1, e_2, $F(x) =$ rotation av vektorn x vinkeln θ i positiv led. Linjariteten illustreras i figuren.



■

Exempel 1.12. $U = R^n$, $V = R^m$,

$$F(x) = Ax, \quad x \in R^n,$$

där A är en $m \times n$-matris. Linjäriteten följer direkt av regler i matrisalgebran, med $y \in R^n$:

$$F(\alpha x + \beta y) = A(\alpha x + \beta y) = \alpha Ax + \beta Ay = \alpha F(x) + \beta F(y). \qquad \blacksquare$$

Exempel 1.13. (a) $U = C^1[0,1]$ = mängden av alla kontinuerligt deriverbara funktioner på $[0,1]$, $V = C[0,1]$,

$$F(f) = Df = f'.$$

Här följer linjariteten av räknereglerna för derivata:

$$F(\alpha f + \beta g) = D(\alpha f + \beta g) = \alpha f' + \beta g' = \alpha F(f) + \beta F(g).$$

(b) Låt $P(t) = t^n + a_{n-1}t^{n-1} + \cdots + a_1 t + a_0$ vara ett polynom. Då är

$$\begin{aligned} F(f) = P(D)f &= (D^n + a_{n-1}D^{n-1} + \cdots + a_1 D + a_0)f \\ &= f^{(n)} + a_{n-1}f^{(n-1)} + \cdots + a_1 f' + a_0 f \end{aligned}$$

en linjär avbildning från $U = C^n(-\infty, \infty)$ = mängden av alla n gånger kontinuerligt deriverbara funktioner på $(-\infty, \infty)$ till $V = C(-\infty, \infty)$. Detta visas som i (a). $\blacksquare$

Exempel 1.14. (a) $U = V = C[a,b]$. För $f \in C[a,b]$ definieras $F_1(f) = g_1$ och $F_2(f) = g_2$, där funktionerna g_1 och g_2 i $C[a,b]$ ges av

$$g_1(x) = \int_a^x f(t)dt, \quad g_2(x) = \int_a^b (x-t)^2 f(t)dt, \quad x \in [a,b].$$

Att verifiera att F_1 och F_2 är linjära lämnas som övning.

En avbildning som (likt dessa och avbildningarna i Exempel 1.13) avbildar en funktion på en funktion brukar kallas en *operator*.

(b) $U = C[a,b]$, $V = R$, $\quad F(f) = \displaystyle\int_a^b f(x)dx.$

En avbildning som (likt denna) avbildar en funktion på ett tal brukar kallas en *funktional*.

Helt analogt med definitionerna av nollrum och värderum för matriser, Definition 1.4, gör vi motsvarande definitioner för avbildningar.

Definition 1.7. Låt F vara en linjär avbildning från ett linjärt rum U till ett linjärt rum V. *Nollrummet* för F är

$$N(F) = \{u \in U : F(u) = 0\},$$

och *värderummet* för F är

$$V(F) = \{v \in V : v = F(u) \text{ för något } u \in U\} = \{F(u) : u \in U\};$$

Precis som i Sats 1.2 visar man att $N(F)$ och $V(F)$ är underrum till U respektive V. Om $U = R^n$, $V = R^m$ och $F(x) = Ax$, med $A \in R^{m\times n}$, så är $N(F) = N(A)$ och $V(F) = V(A)$. I Exempel 1.10 är $N(F) =$ origo och $V(F) = R^2$, i Exempel 1.13 (a) är $N(F) =$ mängden av konstanta funktioner och $V(F) = V$.

En ekvation

$$F(u) = v, \tag{1.5}$$

där $v \in V$ är givet och $u \in U$ sökt, kallas en *linjär ekvation*. Då $v = \mathit{0}$ kallas ekvationen (1.5) *homogen*, annars *inhomogen*. Beroende på F kan (1.5) vara ett linjärt ekvationssystem, en differentialekvation, ett system av differentialekvationer, en differensekvation, en integralekvation, osv. Analogt med Sats 1.2 gäller följande sats:

Sats 1.3. *Antag att u_p är en känd lösning till* (1.5) *(en partikulärlösning). Då är $u \in U$ en lösning till* (1.5) *om och endast om u är av formen*

$$u = u_p + u_h, \quad \textit{där } u_h \in N(F). \tag{1.6}$$

Bevis. Då $F(u_p) = v$, gäller följande ekvivalenser:

$$\begin{aligned}&F(u) = v = F(u_p) \;\Leftrightarrow\; F(u - u_p) = F(u) - F(u_p) = \mathbf{0} \;\Leftrightarrow\\ &u - u_p \in N(F),\end{aligned}$$

vilket visar (1.6). ■

Anmärkning. Enligt Sats 1.3 är lösningsmängden till (1.5) den affina mängden $u_p + N(F)$; se Definition 1.3. I detta sammanhang kan anmärkas att en avbildning som är av formen $v_0 + F(u)$, där $v_0 \in V$ är konstant och F är linjär, brukar kallas en *affin avbildning*.

Vi kommer att återkomma till linjära avbildningar i kapitel 3, speciellt kommer vi att titta på så kallade ortogonala avbildningar samt se på hur man kan beskriva linjära avbildningar med hjälp av matriser.

1.6 Linjärt oberoende, linjärt beroende, bas och dimension

Begreppen som vi introducerar i detta avsnitt stöter man på i någon form i samband med geometriska vektorer. Så är t.ex. tre (geometriska) vektorer linjärt oberoende om de inte ligger i ett och samma plan. En sådan uppsättning vektorer kan användas som en bas i rummet, och detta leder som bekant till begrepp som komponenter (av en vektor) och koordinater (för en punkt). Motsvarande kan göras i ett godtyckligt vektorrum (av ändlig dimension). Dimensionen är det antal vektorer som ingår i en bas. Ett centralt resultat i detta avsnitt är att varje bas innehåller lika många vektorer, vilket inte trivialt. För att komma fram till det behöver vi en del grundläggande satser. Bokstaven V betecknar genomgående ett reellt linjärt rum (men allting gäller också för komplexa linjära rum).

Definition 1.8. Låt $u_1, \dots, u_n$ vara vektorer i V. En *linjärkombination* av $u_1, \dots, u_n$ är en vektor av formen

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n$$

för vissa skalärer $\lambda_1, \dots, \lambda_n$.

Vi har redan sett i avsnitt 1.3 att mängden av alla linjärkombinationer av $u_1, \dots, u_n$ utgör ett underrum av V och att detta underrum betecknas $\mathrm{Span}\{u_1, \dots, u_n\}$.

Definition 1.9. En uppsättning vektorer $u_1, \dots, u_n$ i V sägs *spänna* (eller *generera*) V om varje element i V är en linjärkombination av $u_1, \dots, u_n$, dvs. $V = \mathrm{Span}\{u_1, \dots, u_n\}$.

Exempel 1.15. R^n spänns av vektorerna

$$e_1 = (1, 0, \dots, 0),\ e_2 = (0, 1, \dots, 0), \dots,\ e_n = (0, 0, \dots, 1),$$

ty ett godtyckligt $x \in R^n$ kan skrivas

$$x = (x_1, x_2, \dots, x_n) = x_1 e_1 + x_2 e_2 + \cdots + x_n e_n. \qquad \blacksquare$$

Definition 1.10. En uppsättning vektorer $u_1, \dots, u_n$ i V sägs vara *linjärt beroende* om nollvektorn kan skrivas som en icke-trivial linjärkombination av $u_1, \dots, u_n$, dvs. om ett samband av formen

$$\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_n u_n = \mathit{0} \tag{1.7}$$

är möjligt med något $\lambda_i \neq 0$. Om å andra sidan relationen (1.7) endast är möjlig då $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, så sägs $u_1, \dots, u_n$ vara *linjärt oberoende.*

Exempel 1.16. Vektorerna $e_1, \dots, e_n$ i R^n (se Exempel 1.15) är linjärt oberoende, ty om

$$\mathit{0} = \lambda_1 e_1 + \cdots + \lambda_n e_n = (\lambda_1, \dots, \lambda_n),$$

så är alla $\lambda_i = 0$. $\blacksquare$

Följande lemma är bekant för den som studerat geometriska vektorer och vektorer i R^n, men gäller alltså generellt för linjära rum:

Lemma 1.1. *Vektorerna $u_1, \dots, u_n$ är linjärt beroende om och endast om någon av dem kan skrivas som en linjärkombination av de övriga.*

Bevis. Antag att $u_1, \dots, u_n$ är linjärt beroende så att (1.7) gäller med någon koefficient, säg λ_i, skild från noll. Då kan vi dividera med λ_i och lösa ut u_i som en linjärkombination av de övriga vektorerna:

$$u_i = -\frac{\lambda_1}{\lambda_i} u_1 - \cdots - \frac{\lambda_{i-1}}{\lambda_i} u_{i-1} - \frac{\lambda_{i+1}}{\lambda_i} u_{i+1} - \cdots - \frac{\lambda_n}{\lambda_i} u_n.$$

Omvänt, om någon av vektorerna, säg u_i, är en linjärkombination av de övriga, så är

$$u_i = \alpha_1 u_1 + \cdots + \alpha_{i-1} u_{i-1} + \alpha_{i+1} u_{i+1} + \cdots + \alpha_n u_n$$

för vissa tal $\alpha_1, \ldots, \alpha_{i-1}, \alpha_{i+1}, \ldots, \alpha_n$. Överflyttning av u_i till högerledet ger ett samband av formen (1.7) där åtminstone någon koefficient är skild från noll (nämligen $\lambda_i = -1$). ■

Lemma 1.2. *Antag att $u_1, \ldots, u_m$ är linjärt oberoende vektorer i V. Om $v \in V$ inte tillhör underrummet $U = \mathrm{Span}\{u_1, \ldots, u_m\}$, så är $u_1, \ldots, u_m, v$ linjärt oberoende.*

Bevis. Antag

$$\alpha_1 u_1 + \cdots + \alpha_m u_m + \beta v = \mathit{0}. \tag{1.8}$$

Om $\beta \neq 0$, kan v skrivas som en linjärkombination av $u_1, \ldots, u_m$, vilket strider mot att $v \notin U$. Alltså är $\beta = 0$, och

$$\alpha_1 u_1 + \cdots + \alpha_m u_m = \mathit{0}.$$

Då $u_1, \ldots, u_m$ är linjärt oberoende, så måste vi ha $\alpha_1 = \cdots = \alpha_m = 0$. Då är alla koefficienter i (1.8) noll, och $u_1, \ldots, u_m, v$ är linjärt oberoende. ■

Definition 1.11. En uppsättning vektorer $u_1, \ldots, u_n$ i V sägs vara en *bas* för V om vektorerna är linjärt oberoende och spänner V.

Definition 1.12. Ett linjärt rum V sägs ha *dimensionen* n, skrivet $\dim V = n$, om n är det maximala antalet linjärt oberoende vektorer i V. Om inget sådant n finns, dvs. om man kan finna delmängder av V bestående av godtyckligt många linjärt oberoende vektorer, sägs V vara *oändligdimensionellt.*

Anmärkning. Det triviala rummet $\{\mathit{0}\}$ innehåller inga linjärt oberoende vektorer, varför dimensionen är 0.

Sats 1.4. *Antag att V har dimensionen $n > 0$. Då finns minst en uppsättning av n linjärt oberoende vektorer ur V. Varje sådan uppsättning är en bas för V.*

Bevis. Låt $u_1, \ldots, u_n$ vara en godtycklig uppsättning av n linjärt oberoende vektorer ur V. Existensen av minst en sådan uppsättning följer från Definition 1.12. Låt $v \in V$ vara godtyckligt. Om v inte tillhör $\mathrm{Span}\{u_1, \ldots, u_n\}$, så ger Lemma 1.2 att $u_1, \ldots, u_n, v$ är linjärt oberoende. Men detta strider mot maximaliteten av n i Definition 1.12. Alltså är v en linjärkombination av $u_1, \ldots, u_n$. Då $v \in V$ är godtyckligt, visar detta att $u_1, \ldots, u_n$ spänner V och alltså är en bas för V. ■

Vi har nu visat att varje linjärt rum av ändlig dimension har någon bas. Vi önskar bevisa att alla baser för ett rum V har lika många element. Då behöver vi följande resultat, som är centralt i samband med begreppen bas och dimension.

Sats 1.5. *Antag att V spänner av vektorerna $u_1, \ldots, u_n$. Då är varje uppsättning av fler än n vektorer i V linjärt beroende.*

Bevis. Låt $v_1, \ldots, v_p$ vara p vektorer i V, där $p > n$. Eftersom $u_1, \ldots, u_n$ spänner V kan man skriva

$$v_i = a_{1i}u_1 + \cdots + a_{ni}u_n = \sum_{j=1}^{n} a_{ji}u_j, \quad i = 1, \ldots, p,$$

för vissa koefficienter a_{ji}. Betrakta ekvationen

$$x_1v_1 + \cdots + x_pv_p = 0, \tag{1.9}$$

som kan skrivas

$$\sum_{i=1}^{p} x_i\Big(\sum_{j=1}^{n} a_{ji}u_j\Big) = 0,$$

eller

$$\sum_{j=1}^{n}\Big(\sum_{i=1}^{p} a_{ji}x_i\Big)u_j = 0. \tag{1.10}$$

Ekvationen (1.10), och därmed (1.9), är satisfierad för varje val av $x_1, \ldots, x_p$ sådant att

$$\sum_{i=1}^{p} a_{ji}x_i = 0 \quad \text{för } j = 1, \ldots, n. \tag{1.11}$$

Men (1.11) är ett homogent ekvationssystem med n ekvationer och p obekanta, där $p > n$. Enligt Sats 0.1 har (1.11) en icke-trivial lösning, dvs. det finns tal $x_1, \ldots, x_p$, ej alla noll, så att (1.9) satisfieras. Detta betyder att $v_1, \ldots, v_p$ är linjärt beroende. ■

Sats 1.6. *Antag att V är ändligdimensionellt. Då har alla baser för V lika många element, och detta antal är lika med dimensionen för V.*

Bevis. Klart om $\dim V = 0$. Om $\dim V = n > 0$, så finns enligt Sats 1.4 någon bas med n element. Betrakta så en annan bas (vilken som helst) och antag att den består av m vektorer. Eftersom n är det största antal linjärt oberoende vektorer, som kan förekomma i V, och eftersom vi har m linjärt oberoende vektorer i den andra basen, så måste $m \leq n$. Dessa m vektorer spänner också V, och skulle vi ha $n > m$, så skulle Sats 1.5 ge att de n vektorerna i den första basen vore linjärt beroende, vilket är orimligt. Alltså är $m = n$, dvs. alla baser för V har n element. ■

Exempel 1.17. (a) R^n har dimensionen n, ty enligt Exemplen 1.15 och 1.16 har R^n en bas bestående av n element, nämligen $\{e_i\}_{i=1}^{n}$.

(b) C^n med komplexa skalärer har på samma sätt dimensionen n, ty $e_1, \ldots, e_n$ är en bas också för C^n. Det blir dock annorlunda om vi betraktar reella skalärer. Det komplexa talet $x + iy$ kan identifieras med punkten (x, y) i R^2 (det komplexa planet), varför $\dim C = 2$

och $\dim C^n = 2n$ med reella skalärer. I det senare fallet är vektorerna $e_1, ie_1, \ldots, e_n, ie_n$ en bas.

(c) Rummet P_n, polynom av grad $\leq n$, har dimension $n+1$ ty vektorerna (polynomen) $p_1 = 1,\ p_2 = t,\ ...,\ p_{n+1} = t^n$ är linjärt oberoende och spänner rummet dvs de är en bas med $n+1$ vektorer.

(d) Rummet $C[a,b]$ är oändligdimensionellt, ty vi kan finna godtyckligt många linjärt oberoende funktioner i rummet. Exempelvis är funktionerna $1, t, \ldots, t^n$ kontinuerliga och linjärt oberoende för varje n, ty ett polynom som inte är nollpolynomet har bara ändligt många nollställen. ■

Vi avslutar detta avsnitt med några nyttiga satser rörande bas och dimension. En uppsättning vektorer, som spänner V, kan tunnas ut till en bas för V, och en uppsättning linjärt oberoende vektorer kan utvidgas till en bas för V. Detta är innebörden i det två följande satserna.

Sats 1.7. *Antag att V spänns av $u_1, \ldots, u_n$, och antag att m är det maximala antalet linjärt oberoende vektorer bland $u_1, \ldots, u_n$. Efter eventuell omnumrering av vektorerna kan vi anta att $u_1, \ldots, u_m$ är linjärt oberoende. Då gäller att $u_1, \ldots, u_m$ är en bas för V, så att $\dim V = m$.*

Bevis. Om $m = n$ är saken klar, så antag $m < n$. Låt $U = \text{Span}\{u_1, \ldots, u_m\}$. Då är alltså $u_1, \ldots, u_m$ en bas för U. Antag $m < k \leq n$. Då följer av Lemma 1.2 att $u_k \in U$ (i annat fall vore $u_1, \ldots, u_m, u_k$ linjärt oberoende, vilket skulle strida mot maximaliteten av m). Alltså gäller att alla vektorerna $u_1, \ldots, u_n$, och därmed alla linjärkombinationer av dem, tillhör U. Med andra ord gäller att varje vektor i V tillhör U, varför $U = V$. ■

Sats 1.8. *Antag att $\dim V = n > 0$ och antag att $u_1, \ldots, u_m$ med $m < n$ är linjärt oberoende vektorer i V. Då kan man finna vektorer $u_{m+1}, \ldots, u_n$ så att $u_1, \ldots, u_n$ är en bas för V.*

Bevis. Låt $U_m = \text{Span}\{u_1, \ldots, u_m\}$. Då är $\dim U_m = m < n$, och vi kan finna en vektor $u_{m+1} \in V$ sådan att $u_{m+1} \notin U_m$. Enligt Lemma 1.2 är $u_1, \ldots, u_m, u_{m+1}$ linjärt oberoende. Om $m+1 < n$ upprepas detta resonemang tills vi har funnit n linjärt oberoende vektorer $u_1, \ldots, u_n$. Enligt Sats 1.4 utgör dessa en bas för V. ■

Sats 1.9. *Antag att V är ändligdimensionellt med $\dim V > 0$ och att U är ett underrum av V. Då är $\dim U \leq \dim V$. Om $\dim U = \dim V$, så är $U = V$.*

Bevis. Låt $n = \dim V$. Fler än n linjärt oberoende vektorer i U betyder fler än n linjärt oberoende vektorer i V, vilket är omöjligt. Alltså är $\dim U \leq n$ enligt Definition 1.12. Om $\dim U = n$ så finns en bas för U bestående av n vektorer. Men dessa vektorer måste då också utgöra en bas för V enligt Sats 1.4. Då fås att $U = V$. ■

1.7 Grundläggande teori för linjära ekvationssystem

I Exempel 1.9 bestämde vi baser för nollrummet $N(A)$ och värderummet $V(A)$ till en matris. Vi fann att $\dim N(A) + \dim V(A) = 2 + 3 = 5 =$ antalet kolonner i A. Detta är ingen tillfällighet utan en konsekvens av följande sats, som är en av hörnstenarna i den linjära algebran.

Sats 1.10 (Dimensionssatsen)**.** *Låt A vara en $m \times n$-matris. Då gäller att*

$$\dim N(A) + \dim V(A) = n.$$

Bevis. Antag $\dim N(A) = p$, där $0 < p < n$, och låt $e_1, \ldots, e_p$ vara en bas för $N(A) \subseteq R^n$. Enligt Sats 1.8 kan vi finna vektorer $e_{p+1}, \ldots, e_n$ så att $e_1, \ldots, e_n$ blir en bas för R^n. Varje $x \in R^n$ kan skrivas på formen

$$x = \lambda_1 e_1 + \cdots + \lambda_n e_n,$$

och vi får

$$\begin{aligned} Ax &= \lambda_1 Ae_1 + \cdots + \lambda_p Ae_p + \lambda_{p+1} Ae_{p+1} + \cdots + \lambda_n Ae_n = \\ &= \lambda_{p+1} Ae_{p+1} + \cdots + \lambda_n Ae_n, \end{aligned}$$

ty $Ae_i = 0$, eftersom $e_i \in N(A)$, för $i = 1, \ldots, p$. Alltså spänns $V(A)$ av vektorerna $Ae_{p+1}, \ldots, Ae_n$. Låt oss visa att de är linjärt oberoende. Antag

$$\alpha_{p+1} Ae_{p+1} + \cdots + \alpha_n Ae_n = \mathit{0},$$

och sätt $y = \alpha_{p+1} e_{p+1} + \cdots + \alpha_n e_n$. Då är $Ay = \mathit{0}$, dvs. $y \in N(A)$, och y kan skrivas på formen

$$y = \alpha_1 e_1 + \cdots + \alpha_p e_p.$$

Men då är

$$\alpha_1 e_1 + \cdots + \alpha_p e_p - \alpha_{p+1} e_{p+1} - \cdots - \alpha_n e_n = \mathit{0},$$

och då $e_1, \ldots, e_n$ är linjärt oberoende, så är $\alpha_i = 0$ för alla i. Speciellt är $\alpha_{p+1} = \cdots = \alpha_n = 0$, vilket visar att $Ae_{p+1}, \ldots, Ae_n$ är linjärt oberoende. Därmed är dessa vektorer en bas för $V(A)$, och det följer att $\dim V(A) = n - p$, dvs. $\dim V(A) + \dim N(A) = n$. Med en lätt modifiering av ovanstående resonemang ser man att detta gäller även i fallen $p = 0$ och $p = n$, varför satsen är fullständigt bevisad. ■

Låt *kolonnrangen* för A vara lika med dimensionen för värderummet till A och låt *radrangen* för A vara lika med dimensionen för radrummet till A. Med hjälp av dimensionssatsen kan vi nu bevisa att kolonnrangen och radrangen för en matris är lika.

Sats 1.11 (Rangsatsen)**.** *För en godtycklig $m \times n$-matris A gäller att kolonnrangen och radrangen är lika.*

Bevis. Genom en följd av elementära radoperationer (som vid fullständig Gauss-elimination) kan A transformeras till en matris av formen

$$A' = \begin{array}{c} \begin{array}{ccccccccc} j_1 & j_2 & & & j_3 & & j_{r-1} & j_r & \end{array} \\ \underbrace{\left(\begin{array}{ccccccccc} 1 & 0 & * & * & 0 & \dots & 0 & 0 & * \\ 0 & 1 & * & * & 0 & \dots & 0 & 0 & * \\ 0 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & * \\ -- & -- & -- & -- & -- & -- & -- & -- & -- \\ 0 & 0 & & \dots & & 0 & 1 & 0 & * \\ 0 & 0 & & \dots & & 0 & 0 & 1 & * \\ 0 & 0 & & \dots & & 0 & 0 & 0 & 0 \\ -- & -- & -- & -- & -- & -- & -- & -- & -- \\ 0 & 0 & & \dots & & 0 & 0 & 0 & 0 \end{array}\right)}_{n} \end{array} \begin{array}{l} \\ \left.\begin{array}{c} \\ \\ \\ \\ \\ \\ \end{array}\right\} r \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} m-r \end{array} \tag{1.12}$$

Före kolonn nummer j_1 kan finnas ett antal kolonner bestående av idel nollor (i extremfallet att A är nollmatrisen, så är det allt som finns), och efter kolonn nummer j_r kan finnas ett antal kolonner, som avslutas med $m-r$ nollor. Då ekvationerna $Ax = 0$ och $A'x = 0$ är ekvivalenta, så gäller att $N(A) = N(A')$. Eftersom radrummet bevaras så är radrangen för A och A' lika. Nu ger dimensionssatsen att detsamma gäller för kolonnrang, ty

$$\dim V(A') = n - \dim N(A') = n - \dim N(A) = \dim V(A).$$

Radrangen för A' är r, ty det är uppenbart att de r första raderna i A' är linjärt oberoende (titta speciellt på elementen nummer $j_1, j_2, \ldots, j_r$ i raderna). Likaså är kolonnrangen för A' lika med r, ty kolonnerna nummer $j_1, j_2, \ldots, j_r$ är linjärt oberoende och övriga kolonner är linjärkombinationer av dessa. Av det ovan sagda följer nu att både radrangen och kolonnrangen för A är lika med r. ■

Definition 1.13. Det gemensamma värdet på radrang och kolonnrang kallas matrisens *rang* och betecknas $\operatorname{rang} A$.

Anmärkning. Beviset i Sats 1.11 visar att rangen för A dvs $\dim V(A)$ också kan beskrivas som antalet pivotelement efter fullbordad Gauss-elimination vid lösandet av ekvationssystemet $Ax = 0$. Det står också klart att $\dim N(A)$ är mängden av fria variabler vid lösandet av detta homogena system.

Alternativt bevis för Sats 1.11. Man kan bevisa Sats 1.11 mer direkt utan att behöva åberopa dimensionssatsen. Som förut konstaterar vi att

$$Ax = 0 \Leftrightarrow A'x = 0,$$

där A' är matrisen (1.12). Om a_j är kolonnerna i A och a'_j är kolonnerna i A' så kan detta skrivas

$$x_1 a_1 + \cdots + x_n a_n = 0 \Leftrightarrow x_1 a'_1 + \cdots + x_n a'_n = 0.$$

Alltså bevaras linjära samband mellan kolonnerna vid elementära radoperationer. Detta betyder att om vissa av kolonnerna i A' är linjärt oberoende, så gäller detsamma för motsvarande kolonner i A (och omvänt), och om en viss kolonn a'_{j_0} är en linjärkombination av vissa andra a'_j, så gäller detsamma för motsvarande kolonner i A (och omvänt). Som förut konstaterar vi att kolonnerna nummer $j_1, j_2, \ldots, j_r$ i A' är linjärt oberoende och att övriga kolonner är linjärkombinationer av dessa. Då gäller samma sak för kolonnerna i A. Därmed är kolonnerna nummer $j_1, \ldots, j_r$ i A en bas för $V(A)$, så att kolonnrangen för A är r. Som förut är radrangen för A också lika med r, och satsen följer. ■

Exempel 1.9 (forts.). Med ledning av ovanstående alternativa bevis kan vi ge en annan lösning till del (b). Vi kan nämligen utnyttja radoperationerna i (a)-delen för att bestämma $V(A)$ såväl som $N(A)$. Vi ser att för den transformerade matrisen A' i (a) gäller att första, andra och fjärde kolonnerna utgör en bas för $V(A')$. Alltså är (enligt ovanstående bevis) första, andra och fjärde kolonnerna i A en bas för $V(A)$. ■

Sats 1.12. *Låt A vara en $n \times n$-matris. Då gäller att* $\operatorname{rang} A = n$ *om och endast om* $\det A \neq 0$.

Bevis. Antag att A transformeras till trappstegsformen (1.12) (med $m = n$) genom en följd av elementära radoperationer. Vid en enstaka radoperation multipliceras matrisens determinant med ett tal skilt från noll, så att $\det A = c \det A'$, där $c \neq 0$. Om $r = \operatorname{rang} A = n = m$, så måste pivotelementen 1 i A' ligga utefter huvuddiagonalen (se (1.12)), så att $\det A' = 1$ (eftersom A' är uppåt triangulär). Alltså är $\det A \neq 0$. Om å andra sidan $r < n$, så är minst ett diagonalelement i A' noll. Alltså är $\det A' = 0$ och $\det A = 0$. ■

Med hjälp av dimensionssatsen, Sats 1.10, kan vi nu ge en koncis sammanfattning av teorin för linjära ekvationssystem.

Sats 1.13. *Låt A vara en $m \times n$-matris. Sätt* $r = \operatorname{rang} A$.

(a) *Ekvationssystemet $Ax = b$ är lösbart om och endast om b tillhör det r-dimensionella underrummet $V(A)$ av R^m. Speciellt är $Ax = b$ lösbart för alla $b \in R^m$ om och endast om $r = m$.*

(b) *Antag att systemet $Ax = b$ är lösbart. Lösningsmängden är då $(n - r)$-parametrig. Speciellt är lösningen entydigt bestämd om och endast om $r = n$.*

Bevis. (a) Det första påståendet följer direkt av definitionen av $V(A)$ och av att $\dim V(A) = r$. För det andra påståendet utnyttjar man också att R^m är det enda m-dimensionella underrummet av R^m (Sats 1.9).

(b) Lösningsmängden beskrivs i Sats 1.2. Mängden av alla lösningar x_h till det homogena ekvationssystemet $Ax = 0$ är förstås inget annat än nollrummet $N(A)$. Enligt Sats 1.10 är $\dim N(A) = n - \dim V(A) = n - r$. Därför ingår $n - r$ parametrar i allmänna lösningen x_h. Specialfallet följer av att $\{0\}$ är det enda 0-dimensionella underrummet av R^n. ■

Sats 1.14. *Låt A vara en $n \times n$-matris. Då är följande utsagor ekvivalenta:*

(i) $Ax = 0$ *har endast lösningen* $x = 0$.
(ii) $Ax = b$ *är lösbar för alla* $b \in R^n$.
(iii) $Ax = b$ *är entydigt lösbar för alla* $b \in R^n$.
(iv) $\operatorname{rang} A = n$.
(v) *Kolonnerna i A är linjärt oberoende.*
(vi) *A är inverterbar.*
(vii) $\det A \neq 0$.

Bevis. Sats 1.13 visar att (i), (ii) och (iii) var och en är ekvivalent med att $\operatorname{rang} A = n$ (använd (b) för (i), (a) för (ii), (a) och (b) för (iii)).

Om kolonnerna i A är $a_1, \ldots, a_n$ så gäller att

$$Ax = x_1 a_1 + \cdots + x_n a_n.$$

Härav framgår att både (i) och (v) innebär

$$Ax = 0 \;\Rightarrow\; x = 0.$$

Alltså gäller att (i) $\Leftrightarrow$ (v).

(vi) $\Leftrightarrow$ (vii) följer av Sats 0.2.

Slutligen har vi att (iv) $\Leftrightarrow$ (vii), vilket visades i Sats 1.12.

Därmed är visat att (i) – (vii) är ekvivalenta. ■

Särskilt ofta används implikationen (i) $\Rightarrow$ (vii) i Sats 1.14. Omvänt kan den formuleras på följande sätt:

Följdsats. *Låt A vara en $n \times n$-matris med* $\det A = 0$. *Då har ekvationssystemet* $Ax = 0$ *icke-triviala lösningar.*

Vi avslutar detta avsnitt med ett exempel som behandlar aktuella begrepp för en given matris.

Exempel 1.18. Vi bestämmer baser för $N(A)$, $V(A)$, $Row(A)$ samt rangen för matrisen $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 6 & 0 \\ 0 & -4 & 2 \\ 3 & 7 & 1 \end{pmatrix}$.

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

Vi ser att $Row(A) = Row(A') = Span\{(1 \ \ 1 \ \ 1), \ (0 \ \ 4 \ \ -2)\}$.
Pivotkolonnerna är nummer 1 och 2. Motsvarande kolonner i A spänner kolonnrummet, dvs. $V(A) = Span\{\begin{pmatrix} 1 \\ 2 \\ 0 \\ 3 \end{pmatrix}, \begin{pmatrix} 1 \\ 6 \\ -4 \\ 7 \end{pmatrix}\}$.
För nollrummet löser vi homogena systemet $Ax = 0 \Leftrightarrow A'x = 0$, som har parameterlösningen $x = t\begin{pmatrix} -3/2 \\ 1/2 \\ 1 \end{pmatrix}$ dvs. $N(A) = Span\{\begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}\}$.
Enligt definitionen är $\operatorname{rang} A = \dim V(A) = 2$.
Vi noterar också att $\dim V(A) + \dim N(A) = 3$, som dimensionssatsen säger. ■

1.8 Koordinater och basbyte

Låt $\mathbf{e} = \{e_1, e_2, \dots, e_n\}$ vara en bas för ett linjärt rum V (observera att vektorernas ordningsföljd är väsentlig). Varje vektor $u \in V$ kan då skrivas på formen

$$u = x_1e_1 + x_2e_2 + \cdots + x_ne_n \tag{1.13}$$

för vissa tal $x_1, \dots, x_n$, eftersom $e_1, \dots, e_n$ spänner V. Denna framställning är entydig, ty om

$$u = x_1e_1 + \cdots + x_ne_n = x'_1e_1 + \cdots + x'_ne_n,$$

så får vi

$$(x_1 - x'_1)e_1 + \cdots + (x_n - x'_n)e_n = 0,$$

varför $x_i - x'_i = 0$ för $i = 1, \dots, n$, eftersom $e_1, \dots, e_n$ är linjärt oberoende. Alltså finns en entydig motsvarighet mellan element i V och element i R^n given av (1.13):

$$u \ \leftrightarrow \ x = (x_1, x_2, \dots, x_n).$$

Denna motsvarighet bevarar addition och multiplikation med skalär, dvs. om $u \leftrightarrow x$, $v \leftrightarrow y$, med $y = (y_1, y_2, \dots, y_n) \in R^n$, och $\alpha \in R$, så gäller att

$$\begin{aligned} u + v \ &\leftrightarrow \ x + y, \\ \alpha u \ &\leftrightarrow \ \alpha x. \end{aligned}$$

Detta följer av att

$$\begin{aligned} u+v &= \sum_{i=1}^n x_i e_i + \sum_{i=1}^n y_i e_i = \sum_{i=1}^n (x_i+y_i)e_i \;\leftrightarrow\; (x_1+y_1,\dots,x_n+y_n) = \\ &= (x_1,\dots,x_n)+(y_1,\dots,y_n) = x+y, \\ \alpha u &= \alpha\sum_{i=1}^n x_i e_i = \sum_{i=1}^n \alpha x_i e_i \;\leftrightarrow\; (\alpha x_1,\dots,\alpha x_n) = \alpha(x_1,\dots,x_n) = \alpha x. \end{aligned}$$

Detta utsträcks sedan till godtyckliga linjärkombinationer, så att om $u_j \leftrightarrow x_j$ och $\alpha_j \in R$ för $j = 1,\dots,k$, så gäller att

$$\alpha_1 u_1 + \dots + \alpha_k u_k \;\leftrightarrow\; \alpha_1 x_1 + \dots + \alpha_k x_k.$$

Anmärkning. Av denna relation framgår att ett antal vektorer $u_1,\dots,u_k$ i V linjärt oberoende om och endast om motsvarande vektorer $x_1,\dots,x_k$ i R^n är linjärt oberoende.

Definition 1.14. Talen $x_1,\dots,x_n$ i (1.13) kallas *koordinaterna* för vektorn u i basen $\mathbf{e}$, och vektorn x kan vi kalla för u:s *koordinatvektor*.

Vi låter koordinatvektorn vara en kolonnvektor, som vi också skriver som

$$\begin{pmatrix} | \\ u \\ | \end{pmatrix}_{\mathbf{e}},$$

eller i kompaktare form som $[u]_{\mathbf{e}}$. Då kan ovanstående skrivas

$$[\alpha_1 u_1 + \dots + \alpha_k u_k]_{\mathbf{e}} = \alpha_1 [u_1]_{\mathbf{e}} + \dots + \alpha_k [u_k]_{\mathbf{e}}. \tag{1.14}$$

Om basen ändras, så ändras förstås också koordinaterna. Låt $\mathbf{e} = \{e_1,\dots,e_n\}$ och $\mathbf{e}' = \{e'_1,\dots,e'_n\}$ vara två baser för ett vektorrum V. Låt en vektor $u \in V$ ha koordinaterna $x = (x_1,\dots,x_n)$ och $x' = (x'_1,\dots,x'_n)$ i baserna $\mathbf{e}$ resp. $\mathbf{e}'$. Eftersom $u = \sum_{j=1}^n x'_j e'_j$ ger (1.14)

$$x = [u]_{\mathbf{e}} = \sum_{j=1}^n x'_j [e'_j]_{\mathbf{e}} = \big([e'_1]_{\mathbf{e}},\dots,[e'_n]_{\mathbf{e}}\big)\begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = Tx',$$

där

$$T = \big([e'_1]_{\mathbf{e}},\dots,[e'_n]_{\mathbf{e}}\big) = \begin{pmatrix} | & & | \\ e'_1 & \dots & e'_n \\ | & & | \end{pmatrix}_{\mathbf{e}}. \tag{1.15}$$

Det sista skrivsättet framhäver att kolonnerna i matrisen T är koordinatvektorerna för $e'_1,\dots,e'_n$ i basen $\mathbf{e}$. Vi kallar T för *transformationsmatrisen* eftersom den beskriver övergången från koordinater i en bas $\mathbf{e}'$ till en annan bas $\mathbf{e}$. Om $T = (t_{ij})$ så är $e'_j = \sum_{i=1}^n t_{ij} e_i$.

Eftersom $e'_1, \ldots, e'_n$ är linjärt oberoende i V, så är kolonnerna i T linjärt oberoende i R^n (se anmärkningen ovan). Enligt Sats 1.14 (ekvivalensen mellan (v) och (vi)) är då T inverterbar. Vi har alltså

$$x = Tx', \quad x' = T^{-1}x, \tag{1.16}$$

där

$$T^{-1} = \begin{pmatrix} | & & | \\ e_1 & \ldots & e_n \\ | & & | \end{pmatrix}_{\mathbf{e}'}.$$

Ibland kan det vara aktuellt att förtydliga mellan vilka baser transformationen sker och då använder vi skrivsättet $T_{\mathbf{e}\leftarrow\mathbf{e}'}$ och det följer att $T_{\mathbf{e}'\leftarrow\mathbf{e}} = T^{-1}_{\mathbf{e}\leftarrow\mathbf{e}'}$. Basbytesformlerna blir då:

$$x = T_{\mathbf{e}\leftarrow\mathbf{e}'}x', \quad x' = T_{\mathbf{e}'\leftarrow\mathbf{e}}x = T^{-1}_{\mathbf{e}\leftarrow\mathbf{e}'}x. \tag{1.17}$$

Ibland har vi situationen att V är ett underrum av något linjärt rum U, och att vektorerna e_i och e'_j är givna med koordinater i någon standardbas $\mathbf{f}$ för U. Då ger (1.14)

$$[e'_j]_{\mathbf{f}} = \sum_{i=1}^{n} t_{ij}[e_i]_{\mathbf{f}} = \big([e_1]_{\mathbf{f}}, \ldots, [e_n]_{\mathbf{f}}\big)\begin{pmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{pmatrix} = \big([e_1]_{\mathbf{f}}, \ldots, [e_n]_{\mathbf{f}}\big)[e'_j]_{\mathbf{e}}, \quad j = 1, \ldots, n,$$

dvs.

$$\begin{pmatrix} | & & | \\ e'_1 & \ldots & e'_n \\ | & & | \end{pmatrix}_{\mathbf{f}} = \begin{pmatrix} | & & | \\ e_1 & \ldots & e_n \\ | & & | \end{pmatrix}_{\mathbf{f}} T. \tag{1.18}$$

Speciellt kan vi ha $U = V$ och ta $\mathbf{f} = \mathbf{e}$ eller $\mathbf{f} = \mathbf{e}'$, vilket ger tillbaka (1.17).

Exempel 1.19. Låt $e_1 = (3, 0, -1)$, $e_2 = (1, -1, -1)$ och $e'_1 = (2, 1, 0)$, $e'_2 = (0, 3, 2)$. Då är $\mathbf{e} = \{e_1, e_2\}$ och $\mathbf{e}' = \{e'_1, e'_2\}$ två baser för underrummet (planet) $x - 2y + 3z = 0$ i R^3. (Kontrollera att vektorernas koordinater satisfierar denna ekvation.) Med standardbasen för R^3 ger då (1.18)

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

Vi kan få T genom att multiplicera med transponatet till matrisen som står framför T.

$$\begin{pmatrix} 3 & 0 & -1 \\ 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 3 & 1 \\ 0 & -1 \\ -1 & -1 \end{pmatrix} T = \begin{pmatrix} 3 & 0 & -1 \\ 1 & -1 & -1 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 1 & 3 \\ 0 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 10 & 4 \\ 4 & 3 \end{pmatrix} T = \begin{pmatrix} 6 & -2 \\ 1 & -5 \end{pmatrix}, \quad T = \frac{1}{14}\begin{pmatrix} 3 & -4 \\ -4 & 10 \end{pmatrix}\begin{pmatrix} 6 & -2 \\ 1 & -5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & -3 \end{pmatrix}.$$

Detta betyder att $e'_1 = e_1 - e_2$ och $e'_2 = e_1 - 3e_2$, vilket naturligtvis är lätt att se direkt i detta enkla exempel. ■

1.9 Några exempel på grundläggande begrepp inom linjär algebra

I detta avsnitt ger vi några exempel, som illustrerar begrepp och resultat som vi gått igenom hittills i detta kapitel.

Exempel 1.20. Linjärt oberoende vektorer (funktioner) Visa att funktionerna

$$x,\ e^x,\ xe^x,\ e^{2x}$$

är linjärt oberoende (a) på R, (b) på ett godtyckligt intervall $[a,b]$ $(a<b)$.

Lösning. (a) Antag att

$$c_1x + c_2e^x + c_3xe^x + c_4e^{2x} = 0 \text{ för alla } x \in R. \tag{1.19}$$

Dividera (1.19) med e^{2x}:

$$c_1xe^{-2x} + c_2e^{-x} + c_3xe^{-x} + c_4 = 0 \text{ för alla } x \in R.$$

Låt här $x \to \infty$ så fås $c_4 = 0$. Sätt $x = 0$ i (1.19):

$$c_2 + c_4 = c_2 = 0.$$

Dividera (1.19) med xe^x (och utnyttja $c_2 = c_4 = 0$):

$$c_1e^{-x} + c_3 = 0 \text{ för alla } x \in R.$$

Låt $x \to \infty$ så fås $c_3 = 0$, och alltså är också $c_1 = 0$. Därmed är alla $c_i = 0$, och funktionerna är linjärt oberoende.

(b) Antag att

$$c_1x + c_2e^x + c_3xe^x + c_4e^{2x} = 0 \text{ för alla } x \in [a,b]. \tag{1.20}$$

Den här gången kan vi varken låta $x \to \infty$ eller sätta $x = 0$ (i allmänhet), utan vi måste resonera annorlunda. Ett sätt är att derivera identiteten (1.20) tre gånger, vilket ger

$$\begin{aligned} c_1 + c_2e^x + c_3(x+1)e^x + 2c_4e^{2x} &= 0 \\ c_2e^x + c_3(x+2)e^x + 4c_4e^{2x} &= 0 \\ c_2e^x + c_3(x+3)e^x + 8c_4e^{2x} &= 0 \end{aligned}$$

för alla $x \in [a,b]$. För varje fixt $x \in [a,b]$ fås följande ekvationssystem (efter division med e^x) i de obekanta c_1, c_2, c_3, c_4:

$$\begin{cases} xe^{-x}c_1 + c_2 + xc_3 + e^xc_4 = 0 \\ e^{-x}c_1 + c_2 + (x+1)c_3 + 2e^xc_4 = 0 \\ c_2 + (x+2)c_3 + 4e^xc_4 = 0 \\ c_2 + (x+3)c_3 + 8e^xc_4 = 0 \end{cases}$$

Systemdeterminanten är

$$\begin{vmatrix} xe^{-x} & 1 & x & e^x \\ e^{-x} & 1 & x+1 & 2e^x \\ 0 & 1 & x+2 & 4e^x \\ 0 & 1 & x+3 & 8e^x \end{vmatrix} \ \boxed{-1} = \begin{vmatrix} xe^{-x} & 0 & -2 & -3e^x \\ e^{-x} & 0 & -1 & -2e^x \\ 0 & 1 & x+2 & 4e^x \\ 0 & 0 & 1 & 4e^x \end{vmatrix}$$

$$= -\begin{vmatrix} xe^{-x} & -2 & -3e^x \\ e^{-x} & -1 & -2e^x \\ 0 & 1 & 4e^x \end{vmatrix} \ \boxed{1}\,\boxed{2} = -\begin{vmatrix} xe^{-x} & 0 & 5e^x \\ e^{-x} & 0 & 2e^x \\ 0 & 1 & 4e^x \end{vmatrix} =$$

$$= \begin{vmatrix} xe^{-x} & 5e^x \\ e^{-x} & 2e^x \end{vmatrix} = 2x - 5.$$

Om vi ser till att välja $x \neq \frac{5}{2}$, så blir alltså determinanten $\neq 0$, och systemet har den entydiga lösningen $c_1 = c_2 = c_3 = c_4 = 0$. Detta betyder att funktionerna är linjärt oberoende. ■

Exempel 1.21. (Koordinater och basbyte) Betrakta följande två baser för rummet $\mathcal{P}_3$ (rummet av polynom av högst grad 3): $\mathbf{b} = \{1,\ t-1,\ t^2-t,\ t^3\}$, $\mathbf{c} = \{1,\ t+1,\ t^2-1,\ t^3+t\}$.
Vi ska
a) bestämma koordinaterna för $p = t^3 - t + 1$ i basen $\mathbf{b}$
b) använda koordinattransformation för att bestämma koordinaterna för p i basen $\mathbf{c}$
c) bestämma koordinaterna i $\mathbf{b}$-basen för polynomet q, som i $\mathbf{c}$-basen har koordinaterna $[q]_{\mathbf{c}} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}$

Lösning:

a): Vi får att $p = -1(t-1) + 1(t^3) = -1b_2 + b_4$ dvs. $[q]_{\mathbf{b}} = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$.

b): I detta exempel är det lätt att direkt uttrycka $\mathbf{b}$-basen i $\mathbf{c}$-koordinater, t.ex. $b_2 = t-1 = -2+(t+1) = -2c_1+c_2$. Detta ger transformationsmatrisen $T_{\mathbf{c}\leftarrow\mathbf{b}} = \begin{pmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Koordinatbytet blir alltså $[p]_{\mathbf{c}} = T_{\mathbf{c}\leftarrow\mathbf{b}}[p]_{\mathbf{b}} = \begin{pmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 0 \\ 1 \end{pmatrix}$.
Som kontroll kan vi skriva: $3 - 2(t+1) + t^3 + t = t^3 - t + 1$.
I allmänhet är det inte så lätt att hitta transformationsmatrisen dirikt. Då kan man gå via

en standardbas, som i ekvation (1.18). I detta exempel skulle vi kunna ta standardbasen för P_3 som $\mathbf{e} = \{1,\ t,\ t^2,\ t^3\}$. Det leder till $T_{\mathbf{e}\leftarrow\mathbf{b}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ och $T_{\mathbf{e}\leftarrow\mathbf{c}} = \begin{pmatrix} 1 & 1 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$, och enligt (1.18) blir då $T_{\mathbf{c}\leftarrow\mathbf{b}} = T_{\mathbf{e}\leftarrow\mathbf{c}}^{-1} T_{\mathbf{e}\leftarrow\mathbf{b}}$. I MATLAB med notationen $\backslash$ (backslash) för invertering (ekvationslösning) kan vi alltså skriva $T_{\mathbf{c}\leftarrow\mathbf{b}} = T_{\mathbf{e}\leftarrow\mathbf{c}} \backslash T_{\mathbf{e}\leftarrow\mathbf{b}}$.

c): Enligt basbytesformeln gäller $[q]_{\mathbf{c}} = T_{\mathbf{c}\leftarrow\mathbf{b}}[q]_{\mathbf{b}} \Leftrightarrow \begin{pmatrix} 1 & -2 & 2 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} [q]_{\mathbf{b}} = \begin{pmatrix} 1 \\ 2 \\ -1 \\ 1 \end{pmatrix}$.

Lösning av ekvationssystemet ger $[q]_{\mathbf{b}} = \begin{pmatrix} 6 \\ 2 \\ -1 \\ 1 \end{pmatrix}$.

Exempel 1.22. (Lagranges interpolationsformel) Låt x_i, $i = 0, \ldots, n$, vara $n+1$ olika punkter, och låt c_i, $i = 0, \ldots, n$, vara givna tal. Visa att det finns exakt ett polynom p av grad högst n sådant att

$$p(x_i) = c_i, \quad i = 0, \ldots, n.$$

Lösning. Sätt

$$p_i(x) = \frac{(x - x_0) \ldots (x - x_{i-1})(x - x_{i+1}) \ldots (x - x_n)}{(x_i - x_0) \ldots (x_i - x_{i-1})(x_i - x_{i+1}) \ldots (x_i - x_n)}, \quad i = 0, \ldots, n.$$

Då gäller att $p_i \in \mathcal{P}_n$, och

$$p_i(x_j) = \begin{cases} 0 & \text{för } i \neq j, \\ 1 & \text{för } i = j. \end{cases} \tag{1.21}$$

Polynomen p_i är linjärt oberoende, ty antag att

$$\sum_{j=0}^{n} \lambda_j p_j(x) = 0 \ \text{ för alla } x.$$

Då fås speciellt för $x = x_i$ (med hjälp av (1.21)) att

$$\sum_{j=0}^{n} \lambda_j p_j(x_i) = \lambda_i = 0, \quad i = 0, \ldots, n,$$

vilket bevisar påståendet. Eftersom $\mathcal{P}_n$ är $(n+1)$-dimensionellt, så utgör alltså de $n+1$ elementen p_j en bas för $\mathcal{P}_n$. Av (1.21) följer också att

$$p(x) = \sum_{j=0}^{n} c_j p_j(x)$$

uppfyller $p(x_i) = c_i$ för $i = 0, \ldots, n$. Polynomet p är entydigt bestämt, ty antag att $q \in \mathcal{P}_n$ är sådant att $q(x_i) = c_i$ för $i = 0, \ldots, n$. Då är $q - p$ ett polynom av grad högst n med $n+1$ olika nollställen. Alltså är $q - p$ nollpolynomet, dvs. $q = p$. ■

Exempel 1.23. (Magiska kvadrater) En magisk kvadrat av ordningen n är en $n \times n$-matris, där summorna längs alla rader, alla kolonner och båda diagonalerna alla är lika med ett och samma tal. Detta tal kallas kvadratens linjesumma. Ett enkelt exempel är

$$\begin{pmatrix} 2 & 9 & 4 \\ 7 & 5 & 3 \\ 6 & 1 & 8 \end{pmatrix},$$

en magisk kvadrat av ordning 3 med linjesumman 15. Vi betecknar mängden av alla magiska kvadrater av ordning n med Mag_n. Om $M \in \mathrm{Mag}_n$, så låter vi $s(M)$ beteckna linjesumman hos M. Vidare låter vi $0\mathrm{Mag}_n$ beteckna mängden av alla $M \in \mathrm{Mag}_n$ med $s(M) = 0$. Om M_1 och M_2 tillhör Mag_n, så är det tydligt att $M_1 + M_2$ också är en magisk kvadrat med $s(M_1+M_2) = s(M_1)+s(M_2)$. Vidare gäller att $\alpha M_1 \in \mathrm{Mag}_n$ med $s(\alpha M_1) = \alpha s(M_1)$. Alltså är Mag_n ett underrum av $R^{n\times n}$, och vidare är $0\mathrm{Mag}_n$ ett underrum av Mag_n. Uppgiften är att bestämma dimensionen av Mag_n. Vi antar att $n \geq 3$.

Antag $M \in \mathrm{Mag}_n$ med $s(M) = m$. Låt E vara $n \times n$-matrisen med ettor på alla platser. Då gäller att $E \in \mathrm{Mag}_n$ med $s(E) = n$. Alltså är $M - \frac{m}{n}E$ en magisk kvadrat med linjesumma $s(M - \frac{m}{n}E) = s(M) - \frac{m}{n}s(E) = m - \frac{m}{n}n = 0$, dvs. $M - \frac{m}{n}E \in 0\mathrm{Mag}_n$. Antag att $\dim 0\mathrm{Mag}_n = q$, och låt $M_1, M_2, \ldots, M_q$ vara en bas för $0\mathrm{Mag}_n$. Då kan alltså $M - \frac{m}{n}E$ skrivas som $\alpha_1 M_1 + \cdots + \alpha_q M_q$ för vissa tal $\alpha_1, \ldots, \alpha_q$, och vi får $M = \alpha_1 M_E + \cdots + \alpha_q M_q + \frac{m}{n}E$. Alltså spänns Mag_n av $M_1, \ldots, M_q, E$. Dessa matriser är linjärt oberoende, ty antag att $\beta_1 M_1 + \cdots + \beta_q M_q + \gamma E = 0$. Då är $\beta_1 s(M_1) + \cdots + \beta_q s(M_q) + \gamma s(E) = \gamma n = s(0) = 0$, ty $s(M_i) = 0$ för $i = 1, \ldots, q$. Alltså är $\gamma = 0$ och $\beta_1 M_1 + \cdots + \beta_q M_q = 0$, varför $\beta_1 = \cdots = \beta_q = 0$, eftersom $M_1, \ldots, M_q$ är linjärt oberoende. Följaktligen är $M_1, \ldots, M_q, E$ en bas för Mag_n, och $\dim \mathrm{Mag}_n = q + 1$. Det gäller nu att bestämma q.

Antag att $A = (a_{ij}) \in 0\mathrm{Mag}_n$. Rad-, kolonn- och diagonalsummorna är då

$$\begin{aligned}
\sum_{j=1}^{n} a_{ij} &= 0, \qquad i = 1, \dots, n, \\
\sum_{i=1}^{n} a_{ij} &= 0, \qquad j = 1, \dots, n, \\
\sum_{i=1}^{n} a_{ii} &= 0, \\
\sum_{i=1}^{n} a_{i,n+1-i} &= 0.
\end{aligned}$$

Lägg A:s element i en vektor X:

$$X = (a_{11}, \dots, a_{1n}, a_{21}, \dots, a_{2n}, \dots, a_{n1}, \dots, a_{nn})^T.$$

Då ligger A i $0\mathrm{Mag}_n$ om och endast om

$$BX = 0,$$

där B är $(2n+2) \times n^2$-matrisen

$$B = \begin{pmatrix}
1\,1 \dots 1 & 0\,0 \dots 0 & \dots & 0\,0 \dots 0 \\
0\,0 \dots 0 & 1\,1 \dots 1 & \dots & 0\,0 \dots 0 \\
\vdots\ \vdots\ \ddots\ \vdots & \vdots\ \vdots\ \ddots\ \vdots & \ddots & \vdots\ \vdots\ \ddots\ \vdots \\
0\,0 \dots 0 & 0\,0 \dots 0 & \dots & 1\,1 \dots 1 \\
1\,0 \dots 0 & 1\,0 \dots 0 & \dots & 1\,0 \dots 0 \\
0\,1 \dots 0 & 0\,1 \dots 0 & \dots & 0\,1 \dots 0 \\
\vdots\ \ \ddots\ \ \vdots & \vdots\ \ \ddots\ \ \vdots & \ddots & \vdots\ \ \ddots\ \ \vdots \\
0 \dots 1\,0 & 0 \dots 1\,0 & \dots & 0 \dots 1\,0 \\
0\,0 \dots 1 & 0\,0 \dots 1 & \dots & 0\,0 \dots 1 \\
1\,0 \dots 0 & 0\,1 \dots 0 & \dots & 0\,0 \dots 1 \\
0\,0 \dots 1 & 0 \dots 1\,0 & \dots & 1\,0 \dots 0
\end{pmatrix}.$$

Det gäller alltså att $\dim 0\mathrm{Mag}_n = \dim N(B)$. Låt $r_1, r_2, \dots, r_{2n+2}$ vara raderna i B. Man ser att

$$r_{2n} = r_1 + r_2 + \dots + r_n - r_{n+1} - \dots - r_{2n-1}.$$

Låt oss stryka rad $2n$ och visa att de övriga raderna är linjärt oberoende. Antag att

$$\lambda_1 r_1 + \dots + \lambda_{2n-1} r_{2n-1} + \lambda_{2n+1} r_{2n+1} + \lambda_{2n+2} r_{2n+2} = \mathit{0}.$$

Genom att betrakta positionerna nummer $2n$ och $n+1$ ser man att $\lambda_2 = 0$ och $\lambda_{n+1} = 0$, och på samma sätt fås att $\lambda_3 = 0, \dots, \lambda_{n-1} = 0$. Positionerna nummer $1, 2$ och $n+2$ ger så ekvationerna

$$\begin{cases} \lambda_1 + \lambda_{2n+1} = 0, \\ \lambda_1 + \lambda_{n+2} = 0, \\ \lambda_{n+2} + \lambda_{2n+1} = 0, \end{cases}$$

varav $\lambda_1 = \lambda_{n+2} = \lambda_{2n+1} = 0$. Härefter ses direkt att även övriga koefficienter blir 0. I fallet $n = 3$ måste detta resonemang modifieras något, men resultatet blir det samma. Vi har nu visat att rang $B = 2n + 1$. Enligt dimensionssatsen blir $q = \dim 0\mathrm{Mag}_n = \dim N(B) = n^2 - (2n + 1)$. Slutligen får vi att $\dim \mathrm{Mag}_n = q + 1 = n^2 - 2n$. ■

1.10 Övningsuppgifter

1. a) Ge exempel på två kvadratiska matriser A och B av samma typ sådana att $AB \neq BA$.
b) Antag att de kvadratiska matriserna A och B uppfyller ekvationen $A^2 + AB = I$. Visa att $AB = BA$.

2. Låt V vara mängden av alla positiva reella tal försedd med operationerna

$$\begin{aligned} a \oplus b &= ab \\ \alpha \odot a &= a^\alpha \end{aligned}$$

Visa att V med dessa operationer är ett vektorrum.

3. Låt V vara mängden R^2 försedd med operationerna

$$\begin{aligned} (x_1, x_2) \oplus (y_1, y_2) &= (x_1 + y_1 + 1,\, x_2 + y_2 + 1) \\ \alpha \odot (x_1, x_2) &= (\alpha + \alpha x_1 - 1,\, \alpha + \alpha x_2 - 1). \end{aligned}$$

Visa att V med dessa operationer är ett vektorrum. Vad blir nollelementet?

4. Låt V vara första kvadranten i xy-planet dvs. $V = \{\begin{pmatrix} x \\ y \end{pmatrix} : \ x \geq 0, \ y \geq 0\}$ med operationerna $u \oplus v = u + v$ (vanlig addition i R^2) och $c \odot v = cv$ (vanlig multiplikation med skalär i R^2) för alla $v, \ u \in V$ och $c \in R$.
a) Om u och v är vektorer i V, gäller det då att $u \oplus v$ är en vektor i V?
b) Om u är en vektor i V, gäller det då att $c \odot u$ är en vektor i V för alla c i R?
c) Är V ett linjärt rum (vektorrum)?

5. Låt V vara unionen av första och tredje kvadranten i xy-planet dvs.
$V = \{\begin{pmatrix} x \\ y \end{pmatrix} : \ xy \geq 0\}$ med operationerna $u \oplus v = u + v$ (vanlig addition i R^2) och $c \odot v = cv$ (vanlig multiplikation med skalär i R^2) för alla $v, \ u \in V$ och $c \in R$.
a) Om u och v är vektorer i V, gäller det då att $u \oplus v$ är en vektor i V?
b) Om u är en vektor i V, gäller det då att $c \odot u$ är en vektor i V för alla c i R?
c) Är V ett linjärt rum (vektorrum)?

6. Konstruera en geometrisk figur som illustrerar att en linje i R^2, som inte går genom origo, inte är sluten under addition (vanlig addition i R^2).

7. Låt M vara mängden punkter på enhetsskivan i xy-planet dvs.
$M = \{\begin{pmatrix} x \\ y \end{pmatrix} : \ x^2 + y^2 \leq 1\}$. Visa att M inte är ett underrum till R^2.

8. Vilka av följande delmängder av $\mathcal{P}_n$ (polynom $p(t)$ av grad $\leq n$) är underrum?
a) $p(t) = at^2$ med $a \in R$
b) $p(t) = a + t^2$ med $a \in R$
c) $p(0) = 0$
d) $2p(0) = p(1)$
e) $p(t) \geq 0 \quad$ då $\quad 0 \leq t \leq 1$
f) $p(t) = p(1-t)$ för alla t.
g) koefficienterna är heltal.

9. Vilka av följande delmängder av R^4 är underrum? Vilka är multiplan?
a) $\{x \in R^4 : 3x_1 - 2x_2 + x_3 - x_4 = 0\}$
b) $\{x \in R^4 : 9x_1 - 2x_2 + x_3 + 4x_4 = 1\}$
c) $\{x \in R^4 : x = s(2,3,4,5) + t(6,7,8,9),\ s,t \in R\}$
d) $\{x \in R^4 : 2x_1 - 3x_2 = 0 \quad \text{och} \quad x_1 - x_2 - 2x_3 - x_4 = 1\}$
e) $\{x \in R^4 : x = (1,2,0,1) + t(0,1,2,2)\}$
f) $\{x \in R^4 : x_1x_3 = 0\}$
g) $\{x \in R^4 : x_1^2 + x_2^2 = 0\}$
h) $\{x \in R^4 : x_1^2 + x_2^2 = 1\}$

10. Avgör om M är ett underrum av V, då
a) $V = C(R)$, $M = \{f \in V : f(-x) = f(x) \quad \text{för alla} \quad x \in R\}$.
b) $V = R^{n \times n}$, $M =$ mängden av alla uppåt triangulära $n \times n$-matriser.
c) $V = C[0,1]$, $M = \{f \in V : f(0) = 1\}$.
d) $V = C[0,1]$, $M = \{f \in V : f(1) = 0\}$.
e) $V = C[a,b]$, $M = \{f \in V : f(a) = f(b)\}$
f) $V = R^{n \times n}$, $M =$ mängden av alla matriser $A \in V$ som kommuterar med en given matris $B \in V$, dvs. $AB = BA$.
g) $V = R^{n \times n}$, $M = \{A \in V : \operatorname{sp} A = 0\}$,
($\operatorname{sp} A =$ spåret för $A = \sum_{i=1}^n a_{ii}$).
h) $V = R^{2 \times 4}$, $M = \{A \in V : FA = O\}$, där $F \in R^{3 \times 2}$ är en fix matris.
i) $V = C(R)$, $M = \{f \in V : \int_{-\infty}^{\infty} f^2(x) e^{-x^2} dx < \infty\}$.

11. Bestäm de gemensamma punkterna till följande tre hyperplan i R^4

$$\begin{aligned} 2x_1 + x_2 - x_3 - 3x_4 &= 75 \\ 3x_1 + 2x_2 + x_3 + 2x_4 &= 50 \\ -4x_1 + 3x_2 + 2x_3 - 4x_4 &= 25. \end{aligned}$$

12. (M) Bestäm på formen $Ax_1 + Bx_2 + Cx_3 + Dx_4 + E = 0$ ekvationen för det hyperplan i R^4 som går genom punkterna
$P_0 = (1, 1, 1, 1)$, $P_1 = (2, 3, 2, 2)$, $P_2 = (4, 5, 4, 6)$, $P_3 = (0, 1, 3, 4)$

13. a) Visa att om M är ett multiplan i R^n och x, y är två skilda punkter i M så innehåller M linjen genom x och y, dvs alla punkter av formen $x + t(y - x)$ där $t \in R$.
b) Visa att om M är en delmängd av R^n med egenskapen:
$x, y \in M$, $x \neq y$, medför att linjen genom x och y ligger i M,
så är M ett multiplan.

14. Visa att om M_1 och M_2 är två plan i R^n av dimension 2, så finns ett multiplan av dimension ≤ 5 som innehåller både M_1 och M_2.

15. (M) Visa att vektorerna $(1, 0, 1, 4), (2, 2, 0, 0), (3, 1, 0, 2)$ och $(4, 1, 1, 6)$ i R^4 är linjärt beroende. Skriv $(3, 1, 0, 2)$ som en linjärkombination av de övriga. Kan $(2, 2, 0, 0)$ skrivas som en linjärkombination av de övriga?

16. Undersök vilka av följande mängder av vektorer som är linjärt oberoende.
a) $\{(1, -7, 2), (0, 5, -2), (1, 1, 1)\}$ i R^3
b) $\{(1, 0, 2, -1), (1, 2, -1, 0), (1, 1, 0, -1)\}$ i R^4
c) $\{(1, 2, -1, -2, 4), (6, 1, 5, 3, 2), (4, -3, 7, 7, -6)\}$ i R^5.

17. Visa att vektorerna $(1, 2, 3, 4), (0, 1, 2, 3), (0, 0, 1, 2), (0, 0, 0, 1)$ bildar en bas i R^4. Bestäm koordinaterna för $(1, 1, 1, 1)$ i denna bas.

18. Bestäm dimensionen av de underrum i R^4 som genereras av vektorerna
a) $(1, 1, 1, 1), (4, 2, -2, -3), (3, 1, -3, -4)$
b) $(1, 0, 2, 1), (0, 2, 2, 4), (1, -1, 1, -1)$
c) $(1, 2, 1, 2), (2, 1, 2, 1), (1, 2, 2, 1)$
d) $(1, 2, 4, 3), (2, 1, 5, 0), (1, 1, 3, 1), (4, 2, 10, 0)$
e) $(1, 1, -2, 3), (3, -2, 4, 1), (4, -1, 0, 3), (-3, 2, 2, 2)$.

19. (M) Visa att vektorerna $(1, 0, 1, 0, 1, 0)$ och $(0, 1, 1, 1, 1, -1)$ genererar samma underrum i R^6 som vektorerna $(4, -5, -1, -5, -1, 5)$ och $(-3, 2, -1, 2, -1, -2)$.

20. Låt i exemplen nedan M vara mängden punkter som definieras genom godtyckliga reella tal a, b och c. Avgör om M är ett vektorrum och om så är fallet ange en mängd vektorer som spänner rummet.

$$\text{a)} \begin{bmatrix} 3a+b \\ 4 \\ a-5b \end{bmatrix} \quad \text{b)} \begin{bmatrix} -a+1 \\ a-6b \\ 2b+a \end{bmatrix} \quad \text{c)} \begin{bmatrix} a-b \\ b-c \\ c-a \\ b \end{bmatrix} \quad \text{d)} \begin{bmatrix} 4a+3b \\ 0 \\ a+b+c \\ c-2a \end{bmatrix}.$$

21. Vektorerna $v_1, v_2, \ldots, v_k$ i R^n är linjärt oberoende. Vad är dimensionen av det underrum som genereras av $v_1 - v_2, v_2 - v_3, \ldots, v_{k-1} - v_k, v_k - v_1$?

22. a) Vad är dimensionen av det linjära rummet $R^{n \times n}$ av alla $n \times n$ matriser?

b) Vad är dimensionen av underrummet av symmetriska matriser i $R^{n\times n}$?
c) Visa att de skevsymmetriska matriserna, dvs de som uppfyller $A^T = -A$, bildar ett underrum i $R^{n\times n}$. Vad är dimensionen av detta underrum?

23. Visa att följande funktioner är linjärt beroende
a) $\sin 2t, \cos 2t, \sin^2 t, \cos^2 t$
b) $\ln(t^6+1), \ln(t^4-t^2+1),\ \ln(t^2+1)$
c) $\sin(t+\alpha),\ \sin(t+\beta),\ \sin(t+\gamma)$ för godtyckliga $\alpha, \beta, \gamma \in R$.

24. Visa att följande funktioner är linjärt oberoende
a) e^t, e^{t^2}, e^{t^3}
b) $\sin t, \cos t, \sin 2t, \cos 2t.$
c) $t, \sin t, \cos 2t, \sin t \cos t$

25. Låt $\mathcal{P}_n$ vara det linjära rummet av polynom av grad $\leq n$. Antag att polynomen $p_1, p_2, \ldots, p_{n+1} \in \mathcal{P}_n$ och har nollställe i origo, dvs $p_i(0) = 0$, $i = 1, \ldots, n+1$. Visa att

$$\{p_1, p_2, \ldots, p_{n+1}\}$$

är linjärt beroende.

26. Låt U_1 och U_2 vara underrum i det linjära rummet V. Visa att $U_1 \cap U_2$ och $U_1 + U_2$ är underrum i V. $(U_1 + U_2 = \{u + v : u \in U_1,\ v \in U_2\})$

27. (M) Bestäm dimensionerna av $U_1 \cap U_2$ och $U_1 + U_2$ för de underrum U_1 och U_2 i R^4 som definieras av

a) $U_1 : \begin{cases} x_1 - 3x_2 + 2x_3 = 0 \\ 2x_1 - 5x_2 + 7x_3 + x_4 = 0 \end{cases}$ $\quad U_2 : \begin{cases} x_1 - 4x_2 + x_3 + 2x_4 = 0 \\ x_1 - 5x_2 - 2x_3 + x_4 = 0 \end{cases}$

b) $U_1 : \begin{cases} x_1 - 2x_2 + x_4 = 0 \\ 3x_1 - 4x_2 - x_3 + 2x_4 = 0 \end{cases}$ $\quad U_2 : \begin{cases} 2x_1 - 2x_2 + 3x_3 + 4x_4 = 0 \\ x_2 + x_3 - 2x_4 = 0 \end{cases}$

c) $U_i : \begin{cases} x_1 + x_2 + x_3 + x_4 = 0 \\ x_1 + 2x_2 - x_3 - x_4 = 0 \end{cases}$ $\quad U_2 : \begin{cases} x_1 - 3x_2 + 2x_3 - x_4 = 0 \\ x_1 + 4x_2 + 2x_3 + 5x_4 = 0 \end{cases}$.

28. Låt U_1 vara det underrum i R^4 som genereras av vektorerna $(1, 2, 0, 1), (1, 1, 1, 0)$ och låt U_2 vara det underrum som genereras av $(1, 0, 1, 0), (1, 3, 0, 1)$. Ange dimensionerna av $U_1 + U_2$ och $U_1 \cap U_2$.

29. Låt U_1 och U_2 vara underrum i det linjära rummet V sådana att det finns en vektor u som tillhör U_1 men inte U_2 och en vektor v som tillhör U_2 men inte U_1. Visa att $U_1 \cup U_2$ inte är ett underrum till V.

30. Låt U_1 och U_2 vara underrum i ett ändligdimensionellt vektorrum V. Antag vidare att

$$\dim U_1 + \dim U_2 > \dim V.$$

Visa att det finns en vektor $u \in U_1 \cap U_2$ och $u \neq 0$.

31. Ange en bas för nollrummet $N(A)$ till följande matriser

a) $A=\begin{bmatrix}1&1&3&3\\2&0&4&2\\3&2&8&7\end{bmatrix}$ b) $A=\begin{bmatrix}1&1&2&-1\\1&4&-7&2\\2&3&1&-1\end{bmatrix}$

c) $A=\begin{bmatrix}1&2&1&4\\2&1&1&3\\0&3&1&5\end{bmatrix}$ d) $A=\begin{bmatrix}1&2&3&1&1\\2&5&5&4&4\\1&2&4&4&2\end{bmatrix}$.

32. Ange för samtliga matriser i uppgift 31 en bas för värderummet $V(A)$.

33. Bestäm matris A så att de givna mängderna utgör värderummet $V(A)$:

a) $\{\begin{pmatrix}2s+3t\\r+s-2t\\4r+s\\3r-s-t\end{pmatrix}, r,\ s,\ t\in R\}$ b) $\{\begin{pmatrix}b-c\\2b+c+d\\5c-4d\\d\end{pmatrix}, b,\ c,\ d\in R\}$

34. Bestäm matris A så att de givna mängderna utgör nollrummet $N(A)$:

a) $\{\begin{pmatrix}t\\s\\s-t\end{pmatrix}, s,\ t\in R\}$ b) $\{\begin{pmatrix}t\\t+u\\-s\\s-u\end{pmatrix}, s,\ t,\ u\in R\}$

35. (M) Beräkna $\dim N(A)$ och $\dim V(A)$ för följande matriser

a) $A=\begin{bmatrix}1&1&3&0\\3&0&3&3\\5&2&9&3\\7&1&9&6\end{bmatrix}$ b) $A=\begin{bmatrix}1&1&2&2\\-1&0&2&1\\2&-1&-8&-5\\2&3&2&3\end{bmatrix}$

c) $A=\begin{bmatrix}1&1&1\\2&6&0\\0&-4&2\\3&7&1\end{bmatrix}$ d) $A=\begin{bmatrix}1&3&0&2&1\\2&-1&1&3&2\\3&1&0&1&-1\\0&1&1&4&4\end{bmatrix}$ e) $\begin{bmatrix}1&1&1&1&1&1\\2&4&6&8&10&12\\3&5&7&9&11&13\end{bmatrix}$.

36. Låt

$$A=\begin{bmatrix}1&2&2&3\\3&2&2&1\\2&1&1&0\end{bmatrix}\quad\text{och}\quad B=\begin{bmatrix}1&1&2&0\\1&2&2&3\\2&1&4&-3\end{bmatrix}.$$

a) (M) Undersök om nollrummen $N(A)$ och $N(B)$ i R^4 har någon gemensam vektor $\neq 0$.

b) Har värderummen $V(A)$ och $V(B)$ i R^3 någon gemensam vektor $\neq 0$?

37. Ett underrum U i R^5 genereras av vektorerna $(1,2,1,2,1),(1,3,1,1,2),(2,1,4,3,-3)$, $(3,2,4,8,-2)$. Bestäm dimensionen av U och ange en matris A sådan att $U=N(A)$.

38. (M) Ange rangen av följande matriser

a) $\begin{bmatrix} 1 & 3 & 5 & -2 \\ 2 & 1 & 5 & 1 \\ 1 & 2 & 4 & -1 \end{bmatrix}$ b) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & -2 & 3 & -4 \\ 3 & 2 & 9 & 4 \\ 4 & -4 & 12 & -8 \end{bmatrix}$ c) $\begin{bmatrix} 1 & 2 & 1 & 0 & 3 \\ -1 & 0 & -2 & 1 & -2 \\ 2 & 1 & -4 & 0 & 3 \\ -1 & 1 & 0 & 1 & -1 \end{bmatrix}$

39. Låt A och B vara matriser med samma antal rader och låt C vara matrisen man får av alla kolonner i A och alla kolonner i B. Visa att

$$\operatorname{rang} C \leq \operatorname{rang} A + \operatorname{rang} B.$$

40. Låt A och B vara matriser av samma typ. Visa att

$$\operatorname{rang}(A + B) \leq \operatorname{rang} A + \operatorname{rang} B.$$

41. Avgör vilka av följande avbildningar $T : R^2 \to R^2$ som är linjära:

$$\begin{aligned} T_1(x) &= (x_2^2, x_2); \\ T_2(x) &= (x_1 + x_2, x_1); \\ T_3(x) &= (x_1, 1). \end{aligned}$$

42. En linjär avbildning $F : U \to V$ (där U och V är linjära rum) är ju en funktion, och därför kan linjära avbildningar adderas och multipliceras med skalär på det sätt som är brukligt för funktioner. Vi definierar alltså

$$\begin{aligned} (F_1 \oplus F_2)(u) &= F_1(u) \oplus F_2(u), \quad u \in U, \\ (\alpha \odot F)(u) &= \alpha \odot F(u), \quad u \in U. \end{aligned}$$

Visa att mängden av alla linjära avbildningar $F : U \to V$ också är ett linjärt rum. Det är viktigt att kontrollera att $F_1 \oplus F_2$ och $\alpha \odot F$ verkligen blir linjära.

43. Låt M vara vektorrummet av alla 2×2 matriser. Definiera operatorn $T : M \to M$ genom

$$T(X) = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} X + X \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

a) Visa att T är linjär.
b) Bestäm nollrummet $N(T)$ och dess dimension.

44. Betrakta avbildningen $T : R^3 \to R^3$,

$$T(x_1, x_2, x_3) = (x_1 - x_2 + 2x_3, 2x_1 + x_2, -x_1 - 2x_2 + 2x_3).$$

a) Visa att T är linjär.

b) Visa att $N(T)$ är en rät linje och bestäm dess ekvation.
c) Visa att värdemängden $V(T)$ är ett plan och bestäm dess ekvation.
d) Bestäm alla urbilder till (3,2,1) och (2,1,3).

45. Definiera avbildningen $T : \mathcal{P}_2 \to R^2$ genom $T(p) = \begin{pmatrix} p(0) \\ p(1) \end{pmatrix}$
a) Visa att T är en linjär avbildning.
b) Bestäm ett polynom $p \in \mathcal{P}_2$ som spänner nollrummet $N(T)$.
c) Beskriv värderummet $V(T)$.

46. Definiera en linjär avbildning genom $T : \mathcal{P}_2 \to R^2$ genom $T(p) = \begin{pmatrix} p(0) \\ p(0) \end{pmatrix}$
a) Bestäm två polynom i $\mathcal{P}_2$ som spänner nollrummet $N(T)$.
b) Beskriv värderummet $V(T)$.

47. Betrakta avbildningen $T : R^{n\times n}$ definierad av $T(A) = A + A^T$.
a) Visa att T är linjär.
b) Låt $B \in R^{n\times n}$ vara symmetrisk ($B = B^T$). Bestäm $A \in R^{n\times n}$ sådan att $T(A) = B$.
c) Beskriv nollrummet $N(T)$.
d) Visa att värderummet $V(T)$ är mängden av symmetriska matriser i $R^{n\times n}$.

48. Definiera avbildningen $T : C[0, 1] \to C[0, 1]$ genom att $T(f)$ är primitiv funktion F till f sådan att $F(0) = 0$. Visa att T är linjär och beskriv nollrummet $N(T)$.

49. Låt $T : V \to W$ vara en linjär avbildning. Visa att om $\{v_1,\ v_2,\ \dots\ ,\ v_p\}$ är linjärt beroende i V så är $\{T(v_1),\ T(v_2),\ \dots\ ,\ T(v_p)\}$ linjärt beroende i W.

50. Låt $B = \{u_1, u_2, u_3\}$ och $B' = \{v_1, v_2, v_3\}$ vara baser i R^3 där

$$\begin{aligned} &u_1 = (2, 1, 1), \quad u_2 = (2, -1, 1), \quad u_3 = (1, 2, 1), \\ &v_1 = (3, 1, -5), \quad v_2 = (1, 1, -3), \quad v_3 = (-1, 0, 2). \end{aligned}$$

a) Bestäm övergångsmatrisen från B till B'.
b) Låt $w = (-5, 8, -5)$. Bestäm w:s koordinater i baserna B och B'.

51. Låt V vara vektorrummet, genererat av $f_1 = \sin x$ och $f_2 = \cos x$. Låt $g_1 = 2\sin x + \cos x$ och $g_2 = 3\cos x$.

a) Visa att $B' = \{g_1, g_2\}$ är en bas i V.
b) Bestäm övergångsmatriserna från B' till $B = \{f_1, f_2\}$ och från B till B'.
c) Låt $h = 2\sin x - 5\cos x$. Bestäm h:s koordinater i basen B.
d) Bestäm h:s koordinater i basen B' med hjälp av b) och c).
e) Bestäm h:s koordinater i basen B' direkt.

52. Betrakta följande bas i R^3: $\mathcal{B} = \{\begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -5 \end{pmatrix}, \begin{pmatrix} 8 \\ -2 \\ 7 \end{pmatrix}\}$.

a) Bestäm basbytesmatrisen från $\mathcal{B}$ till standardbasen.
b) Bestäm basbytesmatrisen från standardbasen till basen $\mathcal{B}$.
c) Bestäm koordinaterna för $x = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}$ i $\mathcal{B}$-basen.

53. Undersök om mängden $\{1 + t^3,\ 3 + t - 2t^2,\ -t + 3t^2 - t^3\}$ är linjärt beroende i $\mathcal{P}_3$. Kan elementen utgöra bas för $\mathcal{P}_3$?

54. Visa att $p_1(t), p_2(t), p_3(t)$ är en bas för $\mathcal{P}_2$ då
a) $p_1(t) = (t+1)^2$, $p_2(t) = (t+2)^2$, $p_3(t) = (t+3)^2$
b) $p_1(t) = \frac{1}{2}(t-2)(t-3)$, $p_2(t) = -(t-1)(t-3)$, $p_3(t) = \frac{1}{2}(t-1)(t-2)$.
Ange också koordinaterna för polynomet t^2 i basen p_1, p_2, p_3 (i de båda fallen).

55. Låt (x_1, x_2, x_3) och (x'_1, x'_2, x'_3) vara koordinaterna för samma vektor i två olika baser e_1, e_2, e_3 och e'_1, e'_2, e'_3 för ett tredimensionellt linjärt rum V. Uttryck x'-koordinaterna med hjälp av x-koordinaterna om

a) $\begin{aligned} e'_1 &= e_1 + e_2 + 3e_3 \\ e'_2 &= 2e_1 + 3e_2 + e_3 \\ e'_3 &= 2e_1 + 3e_2 \end{aligned}$ b) $\begin{aligned} e'_1 &= e_1 + 2e_2 + 3e_3 \\ e'_2 &= e_1 + e_2 + e_3 \\ e'_3 &= 2e_1 + e_2 + e_3. \end{aligned}$

56. Bestäm dimensionen hos det underrum till R^3 som bestäms av att första och tredje elementen i vektorerna är lika.

57. (M) De fyra första Hermite-polynomen är $\{1,\ 2t,\ -2+4t,\ -12t+8t^3\}$. Visa att dessa polynom utgör bas i $\mathcal{P}_3$ samt bestäm koordinaterna för $p = 7 - 12t - 8t^2 + 12t^3$ i denna bas.

58. (M) Utöka följande mängd av vektorer till en bas för R^5: $\{ \begin{pmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{pmatrix}, \begin{pmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{pmatrix}, \begin{pmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{pmatrix} \}$.

59. En matris $A \in R^{3\times 8}$ har rang $A = 3$. Bestäm $\dim N(A)$, $\dim Row(A)$ och rang A^T.

60. Om $A \in R^{7\times 5}$, vad är största möjliga rang hos A? Om $A \in R^{5\times 7}$, vad är största möjliga rang hos A? Om $A \in R^{6\times 8}$, vad är minsta möjliga $\dim N(A)$?

61. Visa att om u och v är två vektorer i R^n som inte är noll så är rang $uv^T = 1$.

62. (M) Låt $A = \begin{pmatrix} 7 & -9 & -4 & 5 & 3 & -3 & -7 \\ -4 & 6 & 7 & -2 & -6 & -5 & 5 \\ 5 & -7 & -6 & 5 & -6 & 2 & 8 \\ -3 & 5 & 8 & -1 & -7 & -4 & 8 \\ 6 & -8 & -5 & 4 & 4 & 9 & 3 \end{pmatrix}$. Bestäm matriser C, N och M vars kolonner är respektive bas för $V(A)$, $N(A)$ och $N(A^T)$ samt en matris R vars rader är bas för $Row(A)$. Forma matriserna $S = [R^T \ N]$ och $T = [C \ M]$ och verifiera att S och T är reguljära (inverterbara).

63. (M) Vektorerna $B = \{1, \cos t, \cos^2 t, \ ... \ , \cos^6 t\}$ och $C = \{1, \cos t, \cos 2t, \ ... \ , \cos 6t\}$ är baser i samma linjära rum.
a) Bestäm övergångsmatrisen $T_{B \leftarrow C}$ med hjälp av de trigonometriska sambanden:
$\cos 2t = -1 + 2\cos^2 t$
$\cos 3t = -3\cos t + 4\cos^3 t$
$\cos 4t = 1 - 8\cos^2 t + 8\cos^4 t$
$\cos 5t = 5\cos t - 20\cos^3 t + 16\cos^5 t$
$\cos 6t = -1 + 18\cos^2 t - 48\cos^4 t + 32\cos^6 t$

b) Använd $T_{B \leftarrow C}$ för att (utan jobbiga partialintegreringar) bestämma den primitiva funktionen $\int (5\cos^3 t - 6\cos^4 t + 5\cos^5 t - 12\cos^6 t)dt$.

1.11 Svar

3. $(-1, -1)$
4. a) ja, b) nej, c) nej
5. a) nej, b) ja, c) nej
8. a), c) d) och f) gäller.
9. a), c) och g) är underrum, a),b),c),d),e) och g) är multiplan.
10. c) nej, övriga ja.
11. $x = [4, 35, -32, 0) + t(-1, 2, -3, 1), \quad t \in R.$
12. $-x_1 + x_2 - 2x_3 + x_4 + 1 = 0.$
15. $(3, 1, 0, 2) = (4, 1, 1, 6) - (1, 0, 1, 4)$; nej
16. a) och b)
17. $(1, -1, 0, 0)$
18. a) 2, b) 2, c) 3, d) 2, e) 3
20. a) nej , b) nej, c) $\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix} \right\}$, d) $\left\{ \begin{pmatrix} 4 \\ 0 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}$,
21. $k - 1$
22. a) n^2, b) $\frac{n(n+1)}{2}$, c) $\frac{n(n-1)}{2}$.
27. a) $\dim(U_1 \cap U_2) = 1$, $\dim(U_1 + U_2) = 3$
b) $\dim(U_1 \cap U_2) = 0$, $\dim(U_1 + U_2) = 4$
c) $\dim(U_1 \cap U_2) = 1$, $\dim(U_1 + U_2) = 3$
28. $\dim(U_1 + U_2) = 3$, $\dim(U_1 \cup U_2) = 1$
31. a) t.ex. $(-2, -1, 1, 0)$, $(-1, -2, 0, 1)$
b) t.ex. $(-5, 3, 1, 0)$, $(2, -1, 0, 1)$
c) t.ex. $(1, 1, -3, 1, 0)$, $(1, 0, -5, 1)$

d) t.ex. $(18,-5,-3,1,0)$ $(8,-3,-1,0,1)$
32. a) t.ex. $(1,2,3),\ (1,0,2)$
b) t.ex. $(1,1,2),\ (1,4,3)$
c) t.ex. $(1,2,0,3),\ (2,1,3,0)$
d) t.ex. $(1,2,1),\ (2,5,2),\ (3,5,4)$
33. a) $A=\begin{pmatrix} 0 & 2 & 3 \\ 1 & 1 & -2 \\ 4 & 1 & 0 \\ 3 & -1 & -1 \end{pmatrix}$, b) $A=\begin{pmatrix} 1 & -1 & 0 \\ 2 & 1 & 1 \\ 0 & 5 & -4 \\ 0 & 0 & 1 \end{pmatrix}$
34. a) $A=\begin{pmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix}$, b) $A=\begin{pmatrix} 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & -1 \end{pmatrix}$.
35. a) $\dim N(A)=2$, $\dim V(A)=2$
b) $\dim N(A)=1$, $\dim V(A)=3$
c) $\dim N(A)=1$, $\dim V(A)=2$
d) $\dim N(A)=2$, $\dim V(A)=3$
e) $\dim N(A)=4$, $\dim V(A)=2$
36. a) $N(A)\cap N(B)\neq\{0\}$, b) ja
37. dim = 3, t.ex. $A=\begin{pmatrix} -6 & 1 & 2 & 1 & 0 \\ 0 & -1 & 1 & 0 & 1 \end{pmatrix}$.
38. a) 2, b) 2, c) 3.
41. endast T_2.
43. b) $N(T)=\left\{\begin{bmatrix} a & 0 \\ -a & 0 \end{bmatrix};\ a\in R\right\}$, $\dim N(T)=1$
44. b) $t(-2,4,3)$, $t\in R$, c) $x_1-x_2-x_3=0$
d) till $(3,2,1)$: $(\frac{5}{3},-\frac{4}{3},0)+t(-2,4,3)$
till $(2,1,3)$: inga
45. b) $p=t(t-1)$, c) $V(T)=R^2$
46. b) $p_1=t,\ p_2=t^2$, c) $V(T)=\{(a,a),\ a\in R.$
47. b) $A=\frac{1}{2}B$, c) $N(T)=\{\begin{pmatrix} 0 & b \\ -b & 0 \end{pmatrix}\}$.
48. $N(T)=\{0\}$.
50. a) $\begin{bmatrix} 3 & 2 & 5/2 \\ -2 & -3 & -1/2 \\ 5 & 1 & 6 \end{bmatrix}$

b) $[w]_B=\begin{bmatrix} 9 \\ -9 \\ -5 \end{bmatrix}_B$, $[w]_{B'}=\begin{bmatrix} -7/2 \\ 23/2 \\ 6 \end{bmatrix}_{B'}$
51. b) $\begin{bmatrix} 2 & 0 \\ 1 & 3 \end{bmatrix}$ resp. $\frac{1}{6}\begin{bmatrix} 3 & 0 \\ -1 & 2 \end{bmatrix}$

c) $\begin{bmatrix} 2 \\ -5 \end{bmatrix}_B$

d) och e) $\begin{bmatrix} 1 \\ -2 \end{bmatrix}_{B'}$

52. $T_{E\leftarrow B} = \begin{pmatrix} 3 & 2 & 8 \\ -1 & 0 & -2 \\ 4 & -5 & 7 \end{pmatrix}$, $T_{B\leftarrow E} = \frac{1}{8}\begin{pmatrix} -10 & -54 & -4 \\ -1 & -11 & -2 \\ 5 & 23 & 2 \end{pmatrix}$, $[x]_B = \frac{1}{8}\begin{pmatrix} -122 \\ -25 \\ 53 \end{pmatrix}$.

53. de är linjärt oberoende men inte bas ty de är för få för att spänna rummet.

54. a) koordinaterna för t^2 är $(3, -3, 1)$

b) koordinaterna för t^2 är $(1, 4, 9)$

55. a) $\begin{cases} x_1' = 3x_1 - 2x_2 \\ x_2' = -9x_1 + 6x_2 + x_3 \\ x_3' = 8x_1 - 5x_2 - x_3 \end{cases}$, b) $\begin{cases} x_1' = -2x_2 + x_3 \\ x_2' = -x_1 + 5x_2 - 3x_3 \\ x_3' = x_1 - 2x_2 + x_3 \end{cases}$

56. dim$=2$.

57 koordinaterna för p: (3,3,-2,1.5).

58. Bas för R^5: $\left\{ \begin{pmatrix} -9 \\ -7 \\ 8 \\ -5 \\ 7 \end{pmatrix}, \begin{pmatrix} 9 \\ 4 \\ 1 \\ 6 \\ -7 \end{pmatrix}, \begin{pmatrix} 6 \\ 7 \\ -8 \\ 5 \\ -7 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$.

59. $\dim N(A) = 5$, $\dim Row(A) = 3$, rang$A^T = 3$.

60. 5, 5, 2.

62. Använd lämpliga funktioner i MATLAB.

63. a) $T_{B\leftarrow C} = \begin{pmatrix} 1 & 0 & -1 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & -3 & 0 & 5 & 0 \\ 0 & 0 & 2 & 0 & -8 & 0 & 18 \\ 0 & 0 & 0 & 4 & 0 & -20 & 0 \\ 0 & 0 & 0 & 0 & 8 & 0 & -48 \\ 0 & 0 & 0 & 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 32 \end{pmatrix}$.

b) B-koordinaterna $(0, 0, 0, 5, -6, 5, 12)$ förs över till C-koordinaterna $(-6, 55/8, -69/8, 45/16, -3, 5/16, -3/8)$ och integralen blir
$-6t + \frac{55}{8}\sin t - \frac{69}{16}\sin 2t + \frac{15}{16}\sin 3t - \frac{3}{4}\sin 4t + \frac{1}{16}\sin 5t - \frac{1}{16}\sin 6t + C$.

2 Skalärprodukt

2.1 Definitioner och grundläggande satser

I föregående kapitel studerade vi linjära rum, dvs. mängder försedda med en linjär struktur, som gjorde det möjligt att addera vektorer och multiplicera dem med skalär. Enbart med tillgång till dessa operationer härledde vi ett antal intressanta resultat. Ett typiskt exempel på ett linjärt rum är det åskådliga rummet av geometriska vektorer. I det fallet har vi en rikare struktur med möjlighet att t.ex. mäta avstånd och vinklar. Vi kommer nu med hjälp av begreppet skalärprudukt att generalisera detta till allmännare linjära rum.

Det får antas bekant att för geometriska vektorer definieras en produkt, skalärprodukten, som med hjälp av vektorernas koordinater i en bas bestående av parvis vinkelräta (ortogonala) vektorer av längden 1 (en s.k. ON-bas) kunde skrivas $x \cdot y = x_1y_1 + x_2y_2 + x_3y_3$. Läsaren har säkert också sett att denna formel låter sig generaliseras till R^n. För $x = (x_1, x_2, \ldots, x_n)$ och $y = (y_1, y_2, \ldots, y_n)$ i R^n definierar vi alltså skalärprodukten

$$x \cdot y = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \tag{2.1}$$

Om x och y identifieras med kolonnvektorer, kan $x \cdot y$ skrivas som en matrisprodukt (radvektor gånger kolonnvektor):

$$x \cdot y = x^T y.$$

Med hjälp härav och räknelagar för matrisprodukt får vi direkt följande egenskaper:

(i) $x \cdot y = y \cdot x$ för alla $x, y \in R^n$,
(ii) $(\alpha x) \cdot y = \alpha(x \cdot y)$ för alla $x, y \in R^n, \alpha \in R$,
(iii) $(x + y) \cdot z = x \cdot z + y \cdot z$ för alla $x, y, z \in R^n$,
(iv) $x \cdot x \geq 0$ för alla $x \in R^n$ och $x \cdot x = 0$ endast för $x = \mathit{0}$.

När det gäller geometriska vektorer är skalärprodukten definierad med hjälp av geometriska begrepp som längd och vinkel. I R^n med $n \geq 4$, där vi inte har någon geometrisk intuition att stödja oss på, går vi den motsatta vägen och definierar längd och vinkel med hjälp av skalärprodukten. Hur detta går till visas nedan. Om man utgår från skalärprodukten (2.1) visar det sig – föga överraskande med tanke på ovanstående – att standardbasen $e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_n = (0, 0, \ldots, 1)$ blir en ON-bas. Nu är det ju inte alltid man har en ON-bas, och för geometriska vektorer vet vi att skalärprodukten i koordinatform inte längre blir $x_1y_1 + x_2y_2 + x_3y_3$ om basen inte är ON. Därför är (2.1) inte den enda kandidaten till en skalärprodukt i R^n Hur skalärprodukten än ser ut måste dock egenskaperna (i) – (iv) förbli gällande, och det tar vi fasta på. Man kan också vilja införa en skalärprodukt i linjära rum som inte är R^n Även detta kan göras på många olika sätt, och det enda man egentligen kan säga generellt är att motsvarigheterna till egenskaperna (i) – (iv) måste gälla. Vi gör därför följande formella definition:

Definition 2.1. Låt V vara ett reellt linjärt rum. En *skalärprodukt* i V är en reellvärd funktion $\langle u, v\rangle$ av två variabler u och v i V med följande egenskaper:

(i) $\langle u, v\rangle = \langle v, u\rangle$ för alla $u, v \in V$,
(ii) $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$ för alla $u, v \in V, \alpha \in R$,
(iii) $\langle u_1 + u_2, v\rangle = \langle u_1, v\rangle + \langle u_2, v\rangle$ för alla $u_1, u_2, v \in V$,
(iv) $\langle u, u\rangle \geq 0$ för alla $u \in V$, och $\langle u, u\rangle = 0$ endast för $u = \mathit{0}$.

Exempel 2.1. Några exempel på skalärprodukt.

(a)
$$\langle x, y\rangle = x \cdot y = x_1y_1 + \cdots + x_ny_n$$

är en skalärprodukt i R^n enligt ovan. Den kallas *standardskalärprodukten.*

(b)
$$\langle x, y\rangle = x_1y_1 + 2x_2y_2 + \cdots + nx_ny_n$$

är också en skalärprodukt i R^n.

(c)
$$\langle x, y\rangle = x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2$$

är en skalärprodukt i R^2. Det är endast egenskap (iv) som inte är uppenbar. Vi har

$$\langle x, x\rangle = x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1 + x_2)^2 + x_2^2 \geq 0,$$

och likhet inträffar om och endast om $x_1 + x_2 = 0$ och $x_2 = 0$, dvs. $x_1 = x_2 = 0$.

Dessa exempel är olika specialfall av följande.

(d) Låt A vara en *symmetrisk* $n \times n$-matris. Antag också att A är *positivt definit* (se avsnitt 4.5), dvs. att $x^TAx \geq 0$ för alla $x \in R^n$, och $x^TAx = 0$ endast för $x = \mathit{0}$. Då definierar $\langle x, y\rangle = x^TAy$ en skalärprodukt i R^n. Låt oss verifiera egenskaperna (i) – (iv) i Definition 2.1:

(i) $\langle y, x\rangle = y^TAx = (y^TAx)^T = x^TA^Ty = x^TAy = \langle x, y\rangle$, eftersom A är symmetrisk.
(ii) $\langle \alpha x, y\rangle = (\alpha x)^TAy = \alpha x^TAy = \alpha\langle x, y\rangle$.
(iii) $\langle x + y, z\rangle = (x + y)^TAz = (x^T + y^T)Az = x^TAz + y^TAz = \langle x, z\rangle + \langle y, z\rangle$.
(iv) Att $\langle x, x\rangle = x^TAx \geq 0$ för alla x med likhet endast för $x = \mathit{0}$ är precis definitionen av att A är positivt definit.

(e) Vi inser att skalärprodukten i c) motsvarar x^TAy med $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$.

(f) Låt V vara det linjära rummet $C[a, b]$ av alla reella kontinuerliga funktioner på $[a, b]$. Då är

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt$$

en skalärprodukt i V. ■

(g) Med polynom $p,\ q \in \mathcal{P}_n$ och distinkta punkter $t_1 \neq t_2 \neq\ \ldots\ \neq t_{n+1}$ är $\langle p, q\rangle = p(t_1)q(t_1) + p(t_2)q(t_2) +\ \ldots\ + p(t_{n+1})q(t_{n+1})$ en skalärprodukt i $\mathcal{P}_n$.

Låt oss notera några konsekvenser av Definition 2.1. Genom upprepad användning av (ii) och (iii) finner vi att

$$\begin{aligned}\langle \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k, v\rangle &= \langle \alpha_1 u_1, v\rangle + \langle \alpha_2 u_2, v\rangle + \cdots + \langle \alpha_k u_k, v\rangle = \\ &= \alpha_1\langle u_1, v\rangle + \alpha_2\langle u_2, v\rangle + \cdots + \alpha_k\langle u_k, v\rangle.\end{aligned} \tag{2.2}$$

Med användning av (i) får vi också

$$\langle u, \beta v\rangle = \beta\langle u, v\rangle, \tag{2.3}$$

$$\langle u, \beta_1 v_1 + \beta_2 v_2 + \cdots + \beta_m v_m\rangle = \beta_1\langle u, v_1\rangle + \beta_2\langle u, v_2\rangle + \cdots + \beta_m\langle u, v_m\rangle, \tag{2.4}$$

och genom en kombination av (2.2) och (2.4)

$$\begin{aligned}\Big\langle \sum_{i=1}^{k}\alpha_i u_i, \sum_{j=1}^{m}\beta_j v_j\Big\rangle &= \sum_{i=1}^{k}\alpha_i\Big\langle u_i, \sum_{j=1}^{m}\beta_j v_j\Big\rangle = \sum_{i=1}^{k}\alpha_i\sum_{j=1}^{m}\beta_j\langle u_i, v_j\rangle = \\ &= \sum_{i=1}^{k}\sum_{j=1}^{m}\alpha_i\beta_j\langle u_i, v_j\rangle.\end{aligned} \tag{2.5}$$

Ur (i) och (ii) med $\alpha = 0$ fås

$$\langle \mathit{0}, v\rangle = \langle v, \mathit{0}\rangle = 0 \;\text{ för alla } v \in V. \tag{2.6}$$

I fortsättningen av kapitel 2 betecknar V ett reellt linjärt rum försett med en skalärprodukt. I R^n använder vi alltid standardskalärprodukten om inget annat sägs.

I analogi med vad som gäller för geometriska vektorer kan vi införa begreppen längd och vinkel med hjälp av skalärprodukt.

Definition 2.2. Med *längden* eller *normen* av vektorn $u \in V$ menas

$$\|u\| = \sqrt{\langle u, u\rangle}. \tag{2.7}$$

Observera att detta är väldefinierat p.g.a. (iv). Med *avståndet* mellan u och v menas $\|u-v\|$.

Ur (2.7) och (2.5) får vi

$$\|u+v\|^2 = \langle u+v, u+v\rangle = \langle u, u\rangle + \langle u, v\rangle + \langle v, u\rangle + \langle v, v\rangle,$$

dvs.

$$\|u+v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2. \tag{2.8}$$

Vidare följer av (2.7), (ii) och (2.3) att

$$\|\alpha v\| = |\alpha|\|v\|. \tag{2.9}$$

Exempel 2.2. Vi bestämmer längden av $f = t$ och $g = t^2$ i $C[0,1]$ med skalärprodukt $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. Från definitionen får vi $\|f\| = \sqrt{\int_0^1 t^2\,dt} = \sqrt{1/3}$, $\|g\| = \sqrt{\int_0^1 t^4\,dt} = \sqrt{1/5}$. $\blacksquare$

Definitionen av vinkel hänger på följande viktiga olikhet.

Sats 2.1 (Cauchy-Schwarz' olikhet)**.**

$$|\langle u, v\rangle| \le \|u\|\|v\| \quad \textit{för alla } u, v \in V. \tag{2.10}$$

Likhet gäller om och endast om u och v är linjärt beroende.

Bevis. Om $u = 0$, så gäller (2.10), ty båda leden är 0 (jfr (2.6)). Antag i fortsättningen att $u \neq 0$. För godtyckligt $t \in \mathbb{R}$ är

$$q(t) = \|tu + v\|^2 \ge 0.$$

Med hjälp av (2.8) och (2.9) kan vi skriva

$$q(t) = \|tu\|^2 + 2\langle tu, v\rangle + \|v\|^2 = t^2\|u\|^2 + 2t\langle u, v\rangle + \|v\|^2 = at^2 + 2bt + c,$$

där

$$a = \|u\|^2 > 0,\ b = \langle u, v\rangle,\ c = \|v\|^2.$$

Funktionen $q(t)$ har minimum för $t = -\frac{b}{a}$ (studera $q'(t)$ eller kvadratkomplettera), och minimivärdet är

$$q\big(-\frac{b}{a}\big) = a\big(\frac{b}{a}\big)^2 - 2b\frac{b}{a} + c = c - \frac{b^2}{a} \ge 0. \tag{2.11}$$

Då $a > 0$, fås härav $ac \ge b^2$ eller $|b| \le \sqrt{a}\sqrt{c}$, vilket är samma sak som (2.10). Om likhet gäller här, så råder likhet i (2.11), varför $\| -\frac{b}{a}u + v\|^2 = 0$, $v = \frac{b}{a}u$, dvs. u och v är linjärt beroende. Omvänt, om u och v är linjärt beroende och $u \neq 0$, så är $v = \lambda u$ för något tal λ, och då fås likhet i (2.10). $\blacksquare$

Exempel 2.3. Två exempel på Cauchy-Schwarz' olikhet (2.10).

(a) I R^n fås

$$\Big|\sum_{i=1}^n x_i y_i\Big| \le \Big[\sum_{i=1}^n x_i^2\Big]^{1/2}\Big[\sum_{i=1}^n y_i^2\Big]^{1/2}.$$

(b) I $C[a,b]$ med skalärprodukt som i Exempel 2.1 (f) fås

$$\Big|\int_a^b f(x)g(x)\,dx\Big| \le \Big[\int_a^b f^2(x)\,dx\Big]^{1/2}\Big[\int_a^b g^2(x)\,dx\Big]^{1/2}. \qquad \blacksquare$$

Definition 2.3. Om $u, v \neq \mathit{0}$ är $-1 \leq \frac{\langle u,v\rangle}{\|u\|\|v\|} \leq 1$ enligt Sats 2.1, och vi definierar *vinkeln* mellan u och v som den vinkel θ som uppfyller

$$\cos\theta = \frac{\langle u, v\rangle}{\|u\|\|v\|}, \quad 0 \leq \theta \leq \pi. \tag{2.12}$$

Sats 2.2 (Triangelolikheten)**.**

$$\|u + v\| \leq \|u\| + \|v\| \quad \textit{för alla } u, v \in V. \tag{2.13}$$

Likhet gäller om och endast om $u = 0$ eller $v = \lambda u$ med $\lambda \geq 0$.

Bevis. Enligt (2.8) och Sats 2.1 är

$$\begin{aligned}\|u + v\|^2 = \|u\|^2 + 2\langle u, v\rangle + \|v\|^2 \leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \leq \\ \leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 = (\|u\| + \|v\|)^2.\end{aligned}$$

Rotutdragning ger (2.13). Påståendet om när likhet inträffar följer också direkt ur Sats 2.1. ■

Exempel 2.4. Vi bestämmer vinkeln θ mellan $f = t$ och $f = t^2$ i $C[0, 1]$ med skalärprodukt $\langle f, g\rangle = \int_0^1 f(t)g(t)\, dt$, jämför Exempel 2.2. Vi bestämmer först skalärprodukten $\langle f, g\rangle = \int_0^1 t^3\, dt = 1/4$ och med $\|f\|$ och $\|g\|$ från Exempel 2.2 får vi enligt definitionen av vinkel $\cos\theta = \frac{1/4}{\sqrt{1/15}} = \frac{\sqrt{15}}{4}$ och $\theta = \arccos\frac{\sqrt{15}}{4} \approx 0.2527$. ■

Exempel 2.5. Låt $u \neq \mathit{0}$ och $v \neq \mathit{0}$ vara två vektorer i ett linjärt rum med skalärprodukt. Anta att följande relationer gäller mellan längder av vektorer: $\|u\| = \frac{2}{\sqrt{3}}\|v\| = 2\|u - v\|$. Vi vill bestämma vinkeln α mellan u och v.
Lösning: Från definition av längd och linjäritet hos skalärprodukt får vi
$\|u - v\|^2 = \langle u - v, u - v\rangle = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle$, dvs. $\langle u, v\rangle = \frac{1}{2}\|u\|^2 + \frac{1}{2}\|v\|^2 - \frac{1}{2}\|u - v\|^2$.
Definitionen av vinkel och de givna relationerna mellan längder ger nu
$\cos\alpha = \frac{\langle u,v\rangle}{\|u\|\|v\|} = \frac{1}{2}\frac{\|u\|}{\|v\|} + \frac{1}{2}\frac{\|v\|}{\|u\|} - \frac{1}{2}\frac{\|u-v\|^2}{\|u\|\|v\|} = \frac{1}{2}\frac{2}{\sqrt{3}} + \frac{1}{2}\frac{\sqrt{3}}{2} - \frac{1}{2}\frac{1}{2}\frac{1}{\sqrt{3}} = \frac{\sqrt{3}}{2}$
och därmed blir vinkeln $\alpha = \arccos\frac{\sqrt{3}}{2} = \frac{\pi}{6}$.
Vi noterar att vinkeln stämmer för en liksidig triangel i planet men ett sådant geometriskt resonemang räcker inte för bevis, eftersom uppgiften gäller ett allmänt linjärt rum, men kan användas för kontroll. ■

2.2 Ortogonalitet. Ortogonalprojektion

Fallet $\theta = \frac{\pi}{2}$ i formel (2.12) är av särskilt intresse.

Definition 2.4.

(a) Två vektorer u och v i V är *ortogonala* om $\langle u, v\rangle = 0$. En mängd av vektorer kallas en *ortogonalmängd*, om vektorerna i mängden är parvis ortogonala.
(b) En vektor $u \in V$ kallas *normerad* om $\|u\| = 1$.
(c) En mängd av vektorer kallas en *ortonormerad* mängd, eller en *ON-mängd*, om den är en ortogonalmängd, vars samtliga vektorer är normerade.
(d) En ortogonalmängd, som är en bas för ett linjärt rum, kallas en *ortogonalbas*. Om basen är en ON-mängd, kallas den en *ON-bas*.

Exempel 2.6. Polynomen $p_1 = 1$ och $p_2 = 2t-1$ är ortogonala i $C[0,1]$ med skalärprodukt $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$ ty $\langle p_1, p_2\rangle = \int_0^1 2t-1\,dt = 0$. ■

Exempel 2.7. Vi vill göra $u_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $u_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $u_3 = \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}$ till en ON-bas för R^3. Vektorerna är en ortogonalmängd så det återstår endast att normaliserade den:
$e_1 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $e_2 = \frac{1}{\sqrt{5}}\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $e_3 = \frac{1}{\sqrt{30}}\begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}$. ■

Lemma 2.1. *En ON-mängd $e_1, \dots, e_n$ är linjärt oberoende. Om $e_1, \dots, e_n$ är en ON-bas för V, så gäller att varje vektor $u \in V$ kan skrivas*

$$u = \langle u, e_1\rangle e_1 + \langle u, e_2\rangle e_2 + \cdots + \langle u, e_n\rangle e_n. \tag{2.14}$$

Bevis. Antag att en vektor u är en linjärkombination av $e_1, \dots, e_n$, så att

$$u = \lambda_1 e_1 + \lambda_2 e_2 + \cdots + \lambda_n e_n \tag{2.15}$$

för vissa tal $\lambda_1, \dots, \lambda_n$. Multiplicera (2.15) skalärt med $e_j, 1 \le j \le n$, så fås

$$\langle u, e_j\rangle = \sum_{i=1}^{n} \lambda_i \langle e_i, e_j\rangle = \lambda_j.$$

Om vi väljer $u = \mathit{0}$ i (2.15), så får vi $\lambda_j = \langle \mathit{0}, e_j\rangle = 0$ för alla j, dvs. $e_1, \dots, e_n$ är linjärt oberoende. Om $e_1, \dots, e_n$ är en ON-bas för V, så kan varje $u \in V$ skrivas på formen (2.15), och vi får $\lambda_j = \langle u, e_j\rangle$ för alla j, dvs. vi har visat (2.14). ■

Anmärkning Det är lätt att inse att om e_1, ... , e_n är en ortogonalbas, ej nödvändigtvis normerad, så blir formeln motsvarande (2.14)

$$u = \frac{\langle u, e_1\rangle}{\langle e_1, e_1\rangle} e_1 + \frac{\langle u, e_2\rangle}{\langle e_2, e_2\rangle} e_2 + \cdots + \frac{\langle u, e_n\rangle}{\langle e_n, e_n\rangle} e_n.$$

Exempel 2.8. Mängden $e_1 = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$, $e_2 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}$, $e_3 = \begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}$ är en ortogonalmängd i R^3, se exempel 2.7, och kan alltså väljas som bas i R^3. Vi vill skriva $y = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

som en linjärkombination av $e_1,\ e_2,\ e_3$. Enligt formeln i anmärkningen ovan får vi
$y = \frac{\langle y,e_1\rangle}{\langle e_1,e_1\rangle}e_1 + \frac{\langle y,e_2\rangle}{\langle e_2,e_2\rangle}e_2 + \frac{\langle y,e_3\rangle}{\langle e_3,e_3\rangle}e_3 = \frac{0}{6}e_1 + \frac{3}{5}e_2 + \frac{-6}{30}e_3 = \frac{3}{5}\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} - \frac{1}{5}\begin{pmatrix} -5 \\ -2 \\ 1 \end{pmatrix}$ dvs. y har koordinaterna $(0,\ 3/5,\ -1/5)$ i basen $e_1,\ e_2,\ e_3$. Alternativt kan vi använda den normerade ON-basen från Exempel 2.7 och använda formel (2.14). ■

Sats 2.3. *Låt V vara ändligdimensionellt. Då har V en ON-bas.*

Bevis. Beviset består av en beskrivning av en metod, *Gram-Schmidts ortogonaliseringsmetod*, som också är användbar för konkreta räkningar.

Antag $\dim V = n$ och antag att $v_1, \dots, v_n$ är en bas för V. Sätt $e_1' = v_1$ och ansätt

$$e_2' = \alpha_{21}e_1' + v_2. \tag{2.16}$$

Bestäm α_{21} så att e_2' och e_1' blir ortogonala. Multiplicera alltså (2.16) skalärt med e_1':

$$0 = \langle e_2', e_1'\rangle = \alpha_{21}\langle e_1', e_1'\rangle + \langle v_2, e_1'\rangle = \alpha_{21}\|e_1'\|^2 + \langle v_2, e_1'\rangle.$$

Då $\|e_1'\| = \|v_1\| \neq 0$, måste

$$\alpha_{21} = -\frac{\langle v_2, e_1'\rangle}{\|e_1'\|^2}.$$

Av (2.16) framgår att $e_2' \neq 0$ (annars vore v_1 och v_2 linjärt beroende) och även att e_1', e_2' genererar samma underrum som v_1, v_2. Vi fortsätter på samma sätt. Antag att $e_1', \dots, e_k'$ $(1 \leq k \leq n-1)$ har konstruerats så att $e_1', \dots, e_k'$ är parvis ortogonala och så att $e_1', \dots, e_k'$ genererar samma underrum som $v_1, \dots, v_k$. Ansätt

$$e_{k+1}' = \alpha_{k+1,1}e_1' + \cdots + \alpha_{k+1,k}e_k' + v_{k+1}. \tag{2.17}$$

Talen $\alpha_{k+1,1}, \dots, \alpha_{k+1,k}$ skall väljas så att e_{k+1}' blir ortogonal mot $e_1', \dots, e_k'$. Multiplicera (2.17) skalärt med e_j' för $j = 1, \dots, k$, så fås, eftersom $\langle e_i', e_j'\rangle = 0$ för $i \neq j$,

$$0 = \langle e_{k+1}', e_j'\rangle = \alpha_{k+1,j}\|e_j'\|^2 + \langle v_{k+1}, e_j'\rangle,$$
$$\alpha_{k+1,j} = -\frac{\langle v_{k+1}, e_j'\rangle}{\|e_j'\|^2}, \quad j = 1, \dots, k.$$

Formel (2.17) tillsammans med förutsättningarna om $e_1', \dots, e_k'$ visar att $e_1', \dots, e_{k+1}'$ genererar samma underrum som $v_1, \dots, v_{k+1}$ (speciellt är $e_{k+1}' \neq 0$). Processen fortsättes tills $e_1', \dots, e_n'$ har konstruerats. Sammanfattningsvis får vi alltså

$$\begin{aligned}
e_1' &= v_1, \\
e_2' &= v_2 - \frac{\langle v_2, e_1'\rangle}{\|e_1'\|^2}e_1', \\
&-\,-\,-\,-\,- \\
e_n' &= v_n - \frac{\langle v_n, e_1'\rangle}{\|e_1'\|^2}e_1' - \cdots - \frac{\langle v_n, e_{n-1}'\rangle}{\|e_{n-1}'\|^2}e_{n-1}',
\end{aligned}$$

och dessa vektorer utgör en ortogonalbas för V. Sätt slutligen

$$e_i = \frac{e_i'}{\|e_i'\|}, \quad i = 1, \ldots, n.$$

Då blir $e_1, \ldots, e_n$ en ON-bas för V. ■

Exempel 2.9. Vi ska besämma en ON-bas för $V(A)$ där $A = \begin{pmatrix} -1 & 6 & 6 \\ 3 & -8 & 3 \\ 1 & -2 & 6 \\ 1 & -4 & -3 \end{pmatrix}$ med rang $A = 3$. Kolonnerna, $a_1,\ a_2$ och a_3, är uppenbarligen linjärt oberoende, eftersom rangen är full, och det gäller alltså att göra dem till ON-mängd. Vi använder Gram-Schmidts process enligt beviset ovan:

$$e_1' = a_1 = \begin{pmatrix} -1 \\ 3 \\ 1 \\ 1 \end{pmatrix}$$

$$e_2' = a_2 - \frac{\langle a_2, e_1 \rangle}{\langle e_1, e_1 \rangle} e_1 = \begin{pmatrix} 6 \\ -8 \\ -2 \\ -4 \end{pmatrix} - \frac{-36}{12} \begin{pmatrix} -1 \\ 3 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \\ 1 \\ -1 \end{pmatrix}$$

$$e_3' = a_3 - \frac{\langle a_3, e_1 \rangle}{\langle e_1, e_1 \rangle} e_1 - \frac{\langle a_3, e_2 \rangle}{\langle e_2, e_2 \rangle} e_2 = \begin{pmatrix} 6 \\ 3 \\ 6 \\ -3 \end{pmatrix} - \frac{6}{12} \begin{pmatrix} -1 \\ 3 \\ 1 \\ 1 \end{pmatrix} - \frac{30}{12} \begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 3 \\ -1 \end{pmatrix}$$

Normering ger nu ON-basen $e_1 = \frac{1}{\sqrt{12}} \begin{pmatrix} -1 \\ 3 \\ 1 \\ 1 \end{pmatrix}$, $e_2 = \frac{1}{\sqrt{12}} \begin{pmatrix} 3 \\ 1 \\ 1 \\ -1 \end{pmatrix}$, $e_3 = \frac{1}{\sqrt{12}} \begin{pmatrix} -1 \\ -1 \\ 3 \\ -1 \end{pmatrix}$

■

Sats 2.4 (”Pythagoras’ sats”)**.** (a) *Om u och v är ortogonala, så är*

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2.$$

(b) *Om $\{u_1, u_2, \ldots, u_n\}$ är en ortogonalmängd i V, så är*

$$\|u_1 + u_2 + \cdots + u_n\|^2 = \|u_1\|^2 + \|u_2\|^2 + \cdots + \|u_n\|^2.$$

Bevis. (a) följer ur (2.8), ty $\langle u, v \rangle = 0$.

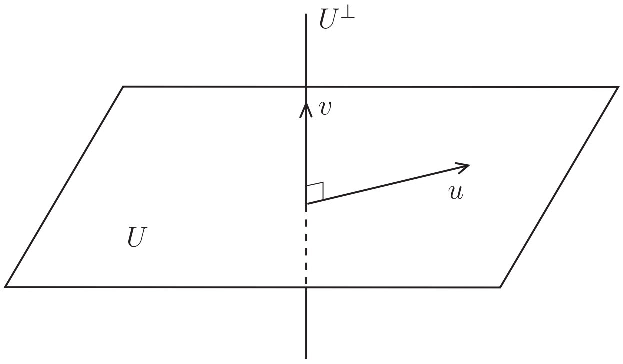
(b) Enligt formel (2.5) är

$$\|u_1 + u_2 + \cdots + u_n\|^2 = \left\langle \sum_{i=1}^{n} u_i \, , \sum_{j=1}^{n} u_j \right\rangle = \sum_{i=1}^{n} \sum_{j=1}^{n} \langle u_i, u_j \rangle = \sum_{i=1}^{n} \|u_i\|^2,$$

ty $\langle u_i, u_j\rangle = 0$ då $i \neq j$. ■

Definition 2.5. Låt U vara ett underrum av V. *Det ortogonala komplementet* $U^\perp$ till U är mängden av alla vektorer i V som är ortogonala mot alla vektorer i U. Med andra ord är

$$U^\perp = \{v \in V : \langle v, u\rangle = 0 \text{ för alla } u \in U\}.$$



Sats 2.5. *Det ortogonala komplementet $U^\perp$ är ett underrum.*

Bevis. Låt $v_1, v_2 \in U^\perp$, $\alpha, \beta \in \mathbb{R}$. Enligt definition är $\langle v_1, u\rangle = 0$ och $\langle v_2, u\rangle = 0$ för alla $u \in U$. Alltså är

$$\langle \alpha v_1 + \beta v_2, u\rangle = \alpha\langle v_1, u\rangle + \beta\langle v_2, u\rangle = 0 \text{ för alla } u \in U,$$

vilket visar att $\alpha v_1 + \beta v_2 \in U^\perp$, och därmed är $U^\perp$ ett underrum enligt Sats 1.1. ■

I många sammanhang har man ett underrum U av V, och man vill kunna dela upp en given vektor $u \in V$ som $u = u' + u''$, där $u' \in U$, $u'' \in U^\perp$. Det är inte alltid en sådan uppdelning är möjlig, men är den det, så går den bara att göra på ett sätt.

Lemma 2.2. *En eventuell uppdelning $u = u' + u''$, där $u' \in U$, $u'' \in U^\perp$, är entydigt bestämd.*

Bevis. Antag $u = u_1' + u_1'' = u_2' + u_2''$, $u_i' \in U$, $u_i'' \in U^\perp$, $i = 1, 2$. Då är $u_1' - u_2' = u_2'' - u_1''$; kalla det gemensamma värdet w. Då gäller dels att $w \in U$ (U underrum), dels att $w \in U^\perp$ ($U^\perp$ underrum). Alltså är w ortogonal mot sig själv, dvs. $0 = \langle w, w\rangle = \|w\|^2$, varför $w = \mathit{0}$, och $u_1' = u_2'$, $u_1'' = u_2''$. ■

Definition 2.6. Komponenten u' i en uppdelning $u = u' + u''$, $u' \in U$, $u'' \in U^\perp$, kallas för *ortogonalprojektionen* av u på U.

Eftersom varje vektor i U är ortogonal mot alla vektorer i $U^\perp$, så är $U \subseteq (U^\perp)^\perp$, och därmed gäller också att komponenten u'' är ortogonalprojektionen på $U^\perp$.

Lemma 2.3. *Om ortogonalprojektionen på U existerar för varje $u \in V$, så gäller att $(U^\perp)^\perp = U$.*

Bevis. Enligt ovan är $U \subseteq (U^\perp)^\perp$. Tag ett godtyckligt $u \in (U^\perp)^\perp$ och skriv $u = u' + u''$ med $u' \in U$, $u'' \in U^\perp$. Då är $0 = \langle u, u''\rangle = \langle u', u''\rangle + \langle u'', u''\rangle = \|u''\|^2$, och alltså är $u'' = 0$, och $u = u' \in U$. Därmed är $(U^\perp)^\perp = U$. ■

Följande sats visar att ortogonalprojektionen på U kan karakteriseras som den vektor i U som ligger närmast u.

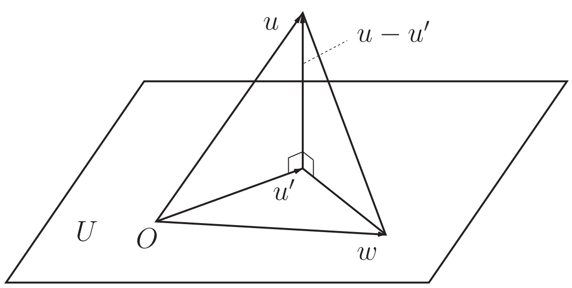
Sats 2.6. *Låt $u \in V$. Följande två egenskaper är ekvivalenta för en vektor $u' \in U$:*

(i) $u - u' \in U^\perp$.
(ii) $\|u - u'\| \leq \|u - w\|$ *för alla* $w \in U$.

Bevis. (i) $\Rightarrow$ (ii). Antag att $u - u' \in U^\perp$, och låt w vara en godtycklig vektor i U. Vektorn $u' - w$ ligger i U och är därmed ortogonal mot $u - u'$. Enligt Pythagoras' sats (Sats 2.4 (a)) är

$$\|u - w\|^2 = \|(u - u') + (u' - w)\|^2 = \|u - u'\|^2 + \|u' - w\|^2 \geq \|u - u'\|^2.$$

Alltså är $\|u - w\| \geq \|u - u'\|$. Likhet inträffar f.ö. endast för $w = u'$.



(ii) $\Rightarrow$ (i). Antag att (ii) gäller, och sätt $u'' = u - u'$. Vi vill visa att $u'' \in U^\perp$. Låt $w \in U$ vara godtyckligt. För alla reella tal t gäller att $u' - tw \in U$, och $\|u - (u' - tw)\| = \|u'' + tw\| \geq \|u''\|$ enligt (ii). Alltså har funktionen $\varphi(t) = \|u'' + tw\|^2$ minimum för $t = 0$. Vi har (jfr beviset för Sats 2.1)

$$\varphi(t) = \|u''\|^2 + 2t\langle u'', w\rangle + t^2\|w\|^2,$$

och

$$\varphi'(0) = 2\langle u'', w\rangle.$$

Men det måste gälla att $\varphi'(0) = 0$. Alltså är $\langle u'', w\rangle = 0$, och det följer att $u'' \in U^\perp$. ■

Vi visar nu att ortogonalprojektionen alltid existerar om U är ändligdimensionellt.

Sats 2.7. *Antag* $\dim U = k$. *För* $u \in V$ *existerar ett entydigt* $u' \in U$ *så att* $u - u' \in U^\perp$. *Om* $k = 0$ *är* $u' = 0$. *Om* $k > 0$ *och om* $e_1, \ldots, e_k$ *är en ON-bas för* U, *så ges* u' *av formeln*

$$u' = \langle u, e_1 \rangle e_1 + \langle u, e_2 \rangle e_2 + \cdots + \langle u, e_k \rangle e_k. \tag{2.18}$$

Bevis. Entydigheten följer av Lemma 2.2. För att visa existensen av ortogonalprojektionen u' räcker det att verifiera att u' givet av formel (2.18) duger då $k > 0$ ((2.18) motiveras av (2.14)). Observera att U har en ON-bas enligt Sats 2.3. Det är tydligt att $u' = \sum_{i=1}^k \langle u, e_i \rangle e_i \in U$. För $j = 1, \ldots, k$ är

$$\begin{aligned} \langle u - u', e_j \rangle = \langle u, e_j \rangle - \langle u', e_j \rangle = \langle u, e_j \rangle - \sum_{i=1}^k \langle u, e_i \rangle \langle e_i, e_j \rangle = \\ = \langle u, e_j \rangle - \langle u, e_j \rangle = 0. \end{aligned}$$

Alltså är $u - u'$ ortogonal mot alla basvektorer e_j och därmed mot alla linjärkombinationer av dessa. Följaktligen gäller att $u - u' \in U^\perp$. ■

Följdsats. *Då* U *är ändligdimensionellt finns ett entydigt bestämt* $u' \in U$ *med egenskaperna* (*i*) *och* (ii) *i Sats 2.6. Vidare är* $(U^\perp)^\perp = U$ *enligt Lemma 2.3.*

Enligt Pythagoras' sats är $\|u\|^2 = \|u'\|^2 + \|u - u'\|^2 \geq \|u'\|^2$. En ny tillämpning av Pythagoras' sats på (2.18) visar att $\|u'\|^2 = \sum_{i=1}^k \langle u, e_i \rangle^2 \|e_i\|^2 = \sum_{i=1}^k \langle u, e_i \rangle^2$. Alltså är

$$\sum_{i=1}^k \langle u, e_i \rangle^2 \leq \|u\|^2,$$

en olikhet som kallas *Bessels olikhet*.

Då även V är ändligdimensionellt gäller följande sats.

Sats 2.8. *Om* $\dim V = n$, *så är*

$$\dim U + \dim U^\perp = n, \tag{2.19}$$

Bevis. Låt $k = \dim U$ och $p = \dim U^\perp$. Satsen klar om $k = 0$ eller $p = 0$, ty $\{0\}^\perp = V$, så antag $k > 0$ och $p > 0$. Låt $e_1, \ldots, e_k$ vara en bas för U och $f_1, \ldots, f_p$ en bas för $U^\perp$. Enligt Sats 2.7 kan varje $u \in V$ skrivas på formen

$$u = u' + u'' = \alpha_1 e_1 + \cdots + \alpha_k e_k + \beta_1 f_1 + \cdots + \beta_p f_p$$

för vissa entydigt bestämda koefficienter $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_p$. Detta innebär att $e_1, \ldots, e_k$, $f_1, \ldots, f_p$ är en bas för V. (Att de är linjärt oberoende beror på att om $u = 0$, så måste $\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_p = 0$ p.g.a. entydigheten.) Alltså är $k + p = n$, vilket är (2.19). ■

Exempel 2.10. Detta exempel illustrerar vad som kan inträffa om U är oändligdimensionellt. Tag $V = C[0,1]$ och låt U vara underrummet av alla $f \in V$ sådana att $f(0) = 0$. Använd den vanliga skalärprodukten $\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt$. För $h \in U^\perp$ gäller att

$$\langle h, f\rangle = \int_0^1 h(t)f(t)\,dt = 0 \qquad \text{för alla } f \in U.$$

Funktionen $th(t)$ tillhör U, varför vi måste ha

$$\int_0^1 t[h(t)]^2\,dt = 0.$$

Detta är möjligt endast om $h(t) = 0$ för alla $t \in [0,1]$ (eftersom h är kontinuerlig). Alltså är $U^\perp = \{0\}$, varför ortogonalprojektionen av en funktion f på U inte existerar (såvida inte f redan ligger i U). ■

Vi kommer att möta en viktig tillämpning av ortogonalprojektion i nästa avsnitt (minsta kvadratmetoden). En annan tillämpning kommer vi att finna senare i samband med studiet av Fourier-serier; se avsnitt 2.5.

Exempel 2.11. Låt $U = Span\{\begin{pmatrix} -4 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}\}$. Vi vill skriva $u = \begin{pmatrix} 6 \\ 4 \\ -5 \end{pmatrix}$ som $u = u' + u''$ med $u' \in U$ och $u'' \in U^\perp$.

Lösning: $u_1 = \begin{pmatrix} -4 \\ -1 \\ 1 \end{pmatrix}$ och $u_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$ är en ortogonalbas för U. Vi använder en variant av (2.18) som gäller i fallet ortogonalbas men inte nödvändigtvis ON-bas:

$$u' = \frac{\langle u, u_1\rangle}{\langle u_1, u_1\rangle}u_1 + \frac{\langle u, u_2\rangle}{\langle u_2, u_2\rangle}u_2 = \frac{-33}{18}\begin{pmatrix} -4 \\ -1 \\ 1 \end{pmatrix} + \frac{-1}{2}\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 22 \\ 4 \\ -7 \end{pmatrix}$$

och $u'' = u - u' = \frac{1}{3}\begin{pmatrix} 18 \\ 12 \\ -15 \end{pmatrix}$ ■

Exempel 2.12. Betrakta samma linjära rum V som i Exempel 1.8, dvs. rummet av reella funktioner f som är kontinuerliga på $[0,\infty)$ och som är sådana att $\int_0^\infty f^2(t)e^{-t}\,dt < \infty$.

a) Visa att

$$\langle f, g\rangle = \int_0^\infty f(t)g(t)e^{-t}\,dt$$

är en skalärprodukt i V.

b) Visa att $\mathcal{P} \subseteq V$. Bestäm en ON-bas för $\mathcal{P}_2$ med användande av skalärprodukten i a).

c) Bestäm ortogonalprojektionen av e^{-t} på $\mathcal{P}_2$.

Lösning. a) Låt $f, g \in V$. Med användande av olikheten $ab \leq \frac{1}{2}(a^2+b^2)$ från Exempel 1.8 har vi

$$\begin{aligned}\int_0^\infty |f(t)g(t)|e^{-t}\,dt &\leq \int_0^\infty \frac{1}{2}[f^2(t)+g^2(t)]e^{-t}\,dt = \\ &= \frac{1}{2}\int_0^\infty f^2(t)e^{-t}\,dt + \frac{1}{2}\int_0^\infty g^2(t)e^{-t}\,dt < \infty,\end{aligned}$$

varför $\int_0^\infty f(t)g(t)e^{-t}\,dt$ är (absolut) konvergent. Egenskaperna hos en skalärprodukt är nu lätta att verifiera. Låt $f, g, h \in V$ och $\alpha \in R$. Då gäller:

(i) $\langle g, f\rangle = \langle f, g\rangle$,

(ii) $\langle f+g, h\rangle = \int_0^\infty [f(t)+g(t)]h(t)e^{-t}\,dt =$
$= \int_0^\infty f(t)h(t)e^{-t}\,dt + \int_0^\infty g(t)h(t)e^{-t}\,dt = \langle f, h\rangle + \langle g, h\rangle$,

(iii) $\langle \alpha f, g\rangle = \int_0^\infty \alpha f(t)g(t)e^{-t}\,dt = \alpha\int_0^\infty f(t)g(t)e^{-t}\,dt = \alpha\langle f, g\rangle$,

(iv) $\langle f, f\rangle = \int_0^\infty f^2(t)e^{-t}\,dt \geq 0$ och $\langle f, f\rangle = 0 \Leftrightarrow f^2(t)e^{-t} = 0$ för alla $t \Leftrightarrow f = 0$.

b) Då $\int_0^\infty t^n e^{-t}\,dt < \infty$ för alla $n \geq 0$, gäller att alla polynom ligger i V. Använd Gram-Schmidts metod och utgå från basen $1, t, t^2$ för $\mathcal{P}_2$. Tag $p_0(t) = 1$. Ansätt $p_1(t) = \alpha p_0(t) + t = t + \alpha$ och bestäm α så att $\langle p_1, p_0\rangle = 0$:

$$0 = \langle p_1, p_0\rangle = \int_0^\infty (t+\alpha)\cdot 1\cdot e^{-t}\,dt = \alpha + 1.$$

Alltså är $\alpha = -1$ och $p_1(t) = t - 1$. Ansätt $p_2(t) = \beta p_0(t) + \gamma p_1(t) + t^2$ och bestäm β och γ så att $\langle p_2, p_0\rangle = \langle p_2, p_1\rangle = 0$:

$$\begin{aligned}0 &= \langle p_2, p_0\rangle = \beta\langle p_0, p_0\rangle + \gamma\langle p_1, p_0\rangle + \langle t^2, p_0\rangle = \\ &= \beta\int_0^\infty 1^2 e^{-t}\,dt + \int_0^\infty t^2 e^{-t}\,dt = \beta + 2, \quad \beta = -2, \\ 0 &= \langle p_2, p_1\rangle = \beta\langle p_0, p_1\rangle + \gamma\langle p_1, p_1\rangle + \langle t^2, p_1\rangle = \\ &= \gamma\int_0^\infty (t-1)^2 e^{-t}\,dt + \int_0^\infty t^2(t-1)e^{-t}\,dt = \gamma + 4, \quad \gamma = -4.\end{aligned}$$

Alltså är $p_2(t) = t^2 - 4t + 2$. Slutligen normerar vi funktionerna genom att sätta $q_i = p_i/\|p_i\|$, $i = 0, 1, 2$. Då blir q_0, q_1, q_2 en ON-bas för $\mathcal{P}_2$. Vi har

$$\begin{aligned}\|p_0\|^2 &= \int_0^\infty 1^2 e^{-x}\,dt = 1, \quad \|p_0\| = 1, \\ \|p_1\|^2 &= \int_0^\infty (t-1)^2 e^{-t}\,dt = 1, \quad \|p_1\| = 1, \\ \|p_2\|^2 &= \int_0^\infty (t^2-4t+2)^2 e^{-t}\,dt = 4, \quad \|p_2\| = 2.\end{aligned}$$

En ON-bas är alltså $q_0(t) = 1$, $q_1(t) = t - 1$, $q_2(t) = \frac{1}{2}t^2 - 2t + 1$. I räkningarna har vi upprepade gånger använt integralerna $\int_0^\infty t^n e^{-t} dt = n!$.

c) Då q_0, q_1, q_2 är en ON-bas för $\mathcal{P}_2$, ges ortogonalprojektionen av $f(t) = e^{-t}$ på $\mathcal{P}_2$ av projektionsformeln

$$\text{proj. av } f = \langle f, q_0\rangle q_0(t) + \langle f, q_1\rangle q_1(t) + \langle f, q_2\rangle q_2(t).$$

Här är

$$\begin{aligned}
\langle f, q_0\rangle &= \int_0^\infty e^{-t}e^{-t}\, dt = \int_0^\infty e^{-2t}\, dt = \frac{1}{2}, \\
\langle f, q_1\rangle &= \int_0^\infty e^{-t}(t-1)e^{-t}\, dt = \int_0^\infty (t-1)e^{-2t}\, dt = -\frac{1}{4}, \\
\langle f, q_2\rangle &= \int_0^\infty e^{-t}\big(\frac{1}{2}t^2 - 2t + 1\big)e^{-t}\, dt = \int_0^\infty \big(\frac{1}{2}t^2 - 2t + 1\big)e^{-2t}\, dt = \frac{1}{8}.
\end{aligned}$$

Alltså är

$$\text{proj. av } f = \frac{1}{2}\cdot 1 - \frac{1}{4}(t-1) + \frac{1}{8}\big(\frac{1}{2}t^2 - 2t + 1\big) = \frac{1}{16}t^2 - \frac{1}{2}t + \frac{7}{8}. \qquad \blacksquare$$

2.3 Ekvationssystem. Minsta kvadratmetoden

Låt A vara en $m \times n$ -matris. Vi skall nu se hur man med hjälp av ortogonalitet kan ge en ny tolkning av ekvationssystemet $Ax = 0$. Som vanligt använder vi oss av standardskalärprodukten (2.1).

Ur matrisproduktens definition följer att $Ax = 0$ är liktydigt med att x är ortogonal mot samtliga rader i A (eller kanske hellre radernas transponat, om vi vill att alla vektorer skall vara kolonnvektorer). Därför har vi följande ekvivalenser:

$Ax = 0 \;\Leftrightarrow$

x är ortogonal mot samtliga (transponerade) rader i $A \;\Leftrightarrow$

x är ortogonal mot alla kolonner i $A^T \;\Leftrightarrow$

x är ortogonal mot det underrum i R^n som genereras av kolonnerna i $A^T \;\Leftrightarrow$

x är ortogonal mot $V(A^T) \;\Leftrightarrow$

$x \in V(A^T)^\perp$.

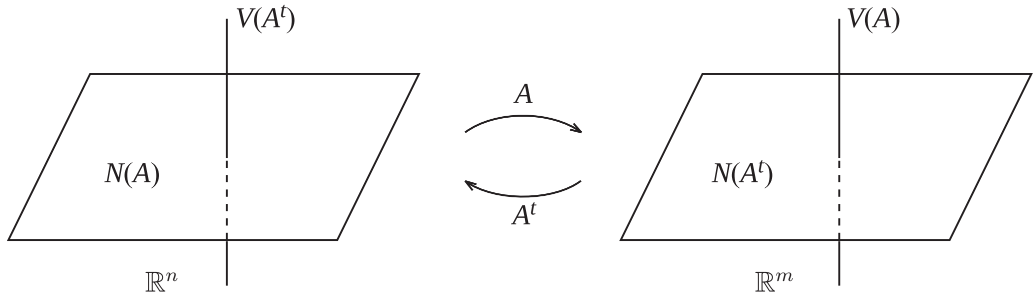
Sats 2.9. *För en godtycklig $m \times n$ -matris A gäller att nollrummet $N(A)$ är lika med det ortogonala komplementet till radrummet $V(A^T)$, dvs.*

$$N(A) = V(A^T)^\perp, \quad N(A)^\perp = V(A^T). \tag{2.20}$$

Det gäller också att

$$N(A^T) = V(A)^\perp, \quad N(A^T)^\perp = V(A). \tag{2.21}$$

Bevis. Den första likheten i (2.20) bevisades ovan. Den andra är bara en omskrivning härav enligt formeln $(U^\perp)^\perp = U$, som gäller för rum av ändlig dimension. Formel (2.21) följer genom att (2.20) tillämpas på matrisen A^T. ■



Med hjälp av Sats 2.8, Sats 2.9 och dimensionssatsen (Sats 1.11) kan vi ge ett nytt bevis för att radrang = kolonnrang:

$$\begin{aligned}&\text{Radrangen för } A = \dim V(A^T) = n - \dim V(A^T)^\perp = \\ &= n - \dim N(A) = \dim V(A) = \text{ kolonnrangen för } A.\end{aligned}$$

Exempel 2.13. Låt U vara det underrum i R^5 som definieras av

$$\begin{cases} x_1 + 2x_2 - x_3 - 3x_4 + 2x_5 = 0, \\ -2x_1 + 3x_2 + x_3 + x_4 - x_5 = 0. \end{cases}$$

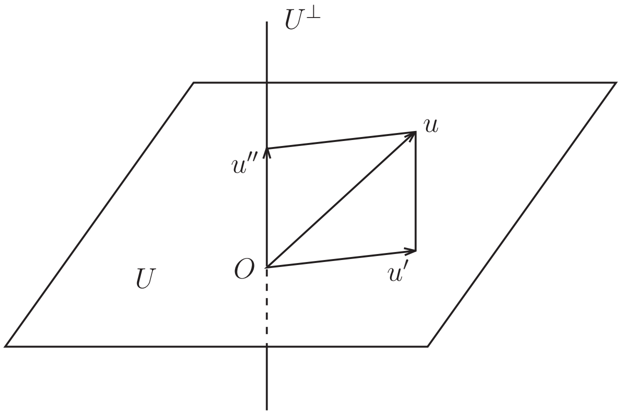
Bestäm det minsta avståndet från $(2, -1, 3, 4, 1)$ till U.

Lösning. Låt A vara matrisen

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

Då är U nollrummet till A, $U = N(A)$. Dela upp den givna vektorn:

$$u = (2, -1, 3, 4, 1) = u' + u'', \quad u' \in U, \quad u'' \in U^\perp. \tag{2.22}$$



En lösningsmetod är att först bestämma en bas för $N(A)$ på vanligt sätt, därefter skaffa sig en ON-bas med hjälp av Gram-Schmidts metod, varefter man får den ortogonala projektionen u' på U med formel (2.18). Slutligen får man det sökta avståndet som $\|u''\| = \|u - u'\|$. Något enklare är det att i stället projicera på $U^\perp$, ty u'' i uppdelningen (2.22) är ju ortogonalprojektionen på $U^\perp$, och $U^\perp$ har lägre dimension än U. Enligt (2.20) är $U^\perp = N(A)^\perp = V(A^T)$, dvs. $U^\perp$ genereras av vektorerna $v_1 = (1,2,-1,-3,2)$, $v_2 = (-2,3,1,1,-1)$. Vi skaffar en ON-bas för $U^\perp$ genom Gram-Schmidts metod. Vi får $e_1' = v_1$ och $e_2' = v_2 - \frac{v_2 \cdot e_1'}{e_1' \cdot e_1'} e_1' = v_2 - \frac{-2}{19} = \frac{1}{19}(-36, 61, 17, 13, -15)$.
Med $e_i = e_i'/\|e_i'\|$, $i = 1,2$ får vi en ON-bas för $U^\perp$.

$$\begin{aligned}
e_1 &= \frac{1}{\sqrt{19}}(1,2,-1,-3,2), \\
e_2 &= \frac{1}{10\sqrt{57}}(-36,61,17,13,-15).
\end{aligned}$$

Ortogonalprojektionen på $U^\perp$ är

$$\begin{aligned}
u'' &= (u \cdot e_1)e_1 + (u \cdot e_2)e_2 = \\
&= \frac{1}{19}(-13)(1,2,-1,-3,2) + \frac{1}{5700}(-45)(-36,61,17,13,-15) = \\
&= \frac{1}{380}(-152,-703,209,741,-475) = \frac{1}{20}(-8,-37,11,39,-25).
\end{aligned}$$

Det sökta avståndet är

$$\|u''\| = \frac{1}{20}\sqrt{8^2 + 37^2 + 11^2 + 39^2 + 25^2} = \frac{1}{2}\sqrt{37}. \qquad \blacksquare$$

Betrakta nu ett ekvationssystem $Ax = b$ som eventuellt saknar lösning (detta är ju i regel fallet för s.k. *överbestämda system* med fler ekvationer än obekanta). Ett sådant system uppkommer t.ex. då man försöker anpassa en rät linje till ett antal mätpunkter. Om det nu inte går att lösa systemet exakt, är man intresserad av att hitta en så bra approximativ lösning som möjligt. Man kan tänka sig många olika kriterier på vad som menas med en bra approximativ lösning. Ett val, som leder till enkla räkningar och därför används mycket i praktiken, är följande: Man söker ett x sådant att kvadratsumman på avvikelserna $[Ax]_i - b_i$ blir så liten som möjligt, dvs. så att

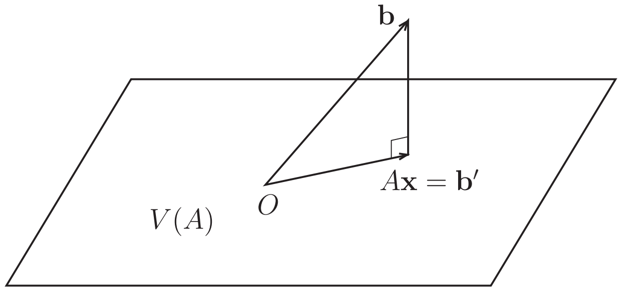
$$([Ax]_1 - b_1)^2 + ([Ax]_2 - b_2)^2 + \cdots + ([Ax]_m - b_m)^2 \tag{2.23}$$

blir så litet som möjligt. Denna approximationsmetod kallas *minsta kvadratmetoden*, och man talar om approximation i minsta kvadratmetodens mening. Nu är ju (2.23) ingenting annat än $\|Ax - b\|^2$, så uttryckt i geometriska termer innebär minsta kvadratmetoden att man minimerar avståndet $\|Ax - b\|$. Alla vektorer av formen Ax utgör ju värderummet $V(A)$, så därför gäller det att hitta en vektor $Ax \in V(A)$, som ligger så nära b som möjligt. Enligt Sats 2.6 ges lösningen av $Ax = b'$, där b' är den ortogonala projektionen av b på $V(A)$ (b' existerar enligt Sats 2.7). Eftersom $b' \in V(A)$ är denna ekvation lösbar, och det visar sig möjligt att lösa den utan att beräkna b'.

Sats 2.10. *Det finns (minst) en lösning x i minsta kvadratmetodens mening till $Ax = b$. Det gäller att x är en sådan lösning om och endast om x satisfierar*

$$A^T Ax = A^T b,$$

ett system av ekvationer, som ibland kallas normalekvationerna.



Bevis. Vi har redan visat (med hjälp av Satserna 2.6 och 2.7) att x är en lösning i minsta kvadratmetodens mening om och endast om x satisfierar $Ax = b'$, där b' är ortogonalprojektionen av b på $V(A)$. Eftersom $Ax \in V(A)$ för godtyckligt x, så följer av ortogonalprojektionens entydighet att $Ax = b'$ om och endast om $b - Ax \in V(A)^\perp$. Men $V(A)^\perp = N(A^T)$ (se (2.21)), och därför har vi nu visat att $Ax = b'$ om och endast om $b - Ax \in N(A^T)$, dvs. om och endast om $A^T(b - Ax) = 0$, eller $A^T Ax = A^T b$. ■

Följdsats. *Ortogonalprojektionen av b på $V(A)$ fås som $b' = Ax$, där x satisfierar $A^T Ax = A^T b$.*

Exempel 2.13 (forts.). En alternativ lösningsmetod är att använda ovanstående följdsats om ortogonalprojektion på ett värderum. Vi hade att u'' var ortogonalprojektionen av u på $V(A^T)$. Alltså är $u'' = A^T x$, där x satisfierar $AA^T x = Au$. Detta ekvationssystem blir

$$\begin{cases} 19x_1 - 2x_2 = -13 \\ -2x_1 + 16x_2 = -1 \end{cases}$$

med lösning $x_1 = -7/10$, $x_2 = -3/20$. Vi får $u'' = \frac{1}{20}(-8, -37, 11, 39, -25)$ som förut. ■

Det ekvationssystem med matris AA^T som vi löste detta exempel brukar man kalla *normalekvationerna av andra slaget.*

Det är uppenbart att $A^T A$ är en symmetrisk matris. I detta sammanhang passar det bra att nämna följande egenskaper för värderummen och nollrummen till $A^T A$.

Lemma 2.4. *För en godtycklig $m \times n$-matris A gäller*

(a) $V(A^T A) = V(A^T)$, *och speciellt* $\operatorname{rang} A^T A = \operatorname{rang} A$,
(b) $N(A^T A) = N(A)$.

Bevis. Egenskap (a) följer direkt av Sats 2.10. Trivialt är $V(A^T A) \subseteq V(A^T)$, och Sats 2.10 visar att ett godtyckligt element i $V(A^T)$ (som kan skrivas på formen $A^T b$) också ligger i $V(A^T A)$. Alltså är $V(A^T A) = V(A^T)$. Speciellt är då $\operatorname{rang} A^T A = \dim V(A^T A) = \dim V(A^T) = \operatorname{rang} A^T = \operatorname{rang} A$. Sedan följer (b) av (a) och Sats 2.9, ty (eftersom $A^T A$ är symmetrisk)

$$N(A^T A) = N((A^T A)^T) = V(A^T A)^\perp = V(A^T)^\perp = N(A). \qquad ■$$

Normalekvationerna blir ofta illa konditionerade. Vi kommer i kapitel 9 att presentera några metoder att lösa det aktuella *minsta-kvadratproblemet*

$$min_x \|Ax - b\|,$$

som har bättre numeriska egenskaper t.ex. QR-faktorisering. Då används *ortogonala matriser*, ett begrepp som vi introducerar i nästa avsnitt. I MATLAB får vi lösningen till ekvationssystem med kommandot \ och programmet väljer lösningsmetod beroende på karaktären av ekvationssystemet, Gausselimination vid ett (inverterbart) kvadratiskt system och QR-faktorisering vid ett överbestämt system dvs. med $A \in R^{m \times n}$ med $m > n$.
Vi avslutar detta avsnitt med ett par exempel på användning av minsta-kvadratproblem, fler tillämpningar finns i kapitel 9.

Exempel 2.14. Statistik, linjära modeller.
En linjär modell beskrivs av en designmatris X, som beskriver hur modellen byggs upp, en vektor y som utgör observationer av det fenomen som modellen avses beskriva samt en vektor β som innehåller de parametrar i modellen som man vill optimera med avseende på. Anta att vi har n parametrar i modellen och gör m observationer. Då gäller $X \in$

$R^{m\times n}$, $y \in R^m$, $\beta \in R^n$ och vi får ett linjärt ekvationssystem $X\beta = y$ med m ekvationer (observationer) och n obekanta (parametrar). Om vi antar att vi gör fler observationer än vi har parametrar så blir $m > n$ och ett sådant ekvationssystem kallar vi *överbestämt*. Normalekvationerna för att lösa systemet i minsta kvadratmening blir $X^T X\beta = X^T y$.
Ett enkelt, konkret exempel på en modell är den linjära ekvationen $\beta_1 + \beta_2 x = y$ med mätningar (x_i, y_i), $i = 1, \ ... \ , m$, dvs anpassning av en rät linje till m givna punkter i planet. Här är alltså $X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ .. & .. \\ .. & .. \\ 1 & x_m \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ y_2 \\ .. \\ .. \\ y_m \end{pmatrix}$.
Med fem mätningar med (x, y)-koordinaterna $(1, 2)$, $(2, 3)$, $(3, 5)$, $(4, 6)$ och $(5, 8)$ får vi

$$X = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}, \ y = \begin{pmatrix} 2 \\ 3 \\ 5 \\ 6 \\ 8 \end{pmatrix}, \ X^T X = \begin{pmatrix} 5 & 15 \\ 15 & 55 \end{pmatrix}, \ X^T y = \begin{pmatrix} 24 \\ 87 \end{pmatrix}.$$

Normalekvationssystemet $\begin{pmatrix} 5 & 15 \\ 15 & 55 \end{pmatrix}\beta = \begin{pmatrix} 24 \\ 87 \end{pmatrix}$ har lösningen $\beta = \begin{pmatrix} 0.3 \\ 1.5 \end{pmatrix}$ med minsta norm $\|X\beta - y\| = 0.5417$.
Om vi i stället vill anpassa med en kvadratisk ekvation $\beta_1 + \beta_2 x + \beta_3 x^2 = y$ så blir modellen fortfarande linjär eftersom modellparametrarna β förekommer linjärt i modellen. En lämplig övning för läsaren kan vara att sätta upp motsvarande överbestämda linjära ekvationssystem. ■

Exempel 2.15. Keplers ekvation
Astronomen Kepler tyckte sig finna följande samband för vinkeln ϕ till solen mellan två positoner för en planet motsvarande tidpunkterna t_0 och t: $\phi - e\sin\phi = a(t - t_0)$
Här är e och a två parametrar, som beror på vilken planet det handlar om (och vilket solsystem man betraktar). Om vi har mätningar av vinkel vid m olika tidpunkter (t_i, ϕ_i), $i = 1, \ ... \ , \ m, \ m > 2$ så får vi ekvationer $a(t_i - t_0) + e\sin\phi = \phi$, $i = 1, \ ... \ , \ m$ dvs. ett överbestämt linjärt ekvationssystem för att bestämma parametrarna e och a.
På matrisform blir systemet $\begin{pmatrix} t_1 - t_0 & \sin\phi_1 \\ t_2 - t_0 & \sin\phi_2 \\ .. & .. \\ .. & .. \\ t_m - t_0 & \sin\phi_m \end{pmatrix} \begin{pmatrix} a \\ e \end{pmatrix} = \begin{pmatrix} \phi_1 \\ \phi_2 \\ .. \\ .. \\ \phi_m \end{pmatrix} \Leftrightarrow Ax = b.$
I MATLAB löser vi alltså ett sådant system (med inlagda konkreta data) med kommandot $x = A\backslash b$ och får minsta kvadratlösningen dvs den med minsta norm $\|Ax - b\|$. ■

2.4 Basbyte mellan ON-baser. Ortogonala matriser

Låt $e_1, \ldots, e_n$ vara en ON-bas för V. Om två vektorer u och v i V har koordinaterna $x_1, \ldots, x_n$ respektive $y_1, \ldots, y_n$ i denna bas, så blir deras skalärprodukt (se formel (2.5))

$$\langle u, v\rangle = \Big\langle \sum_{i=1}^{n} x_i e_i\,, \sum_{j=1}^{n} y_j e_j \Big\rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} x_i y_j \langle e_i, e_j\rangle = \sum_{i=1}^{n} x_i y_i, \tag{2.24}$$

eftersom $\langle e_i, e_j\rangle = \delta_{ij}$, där som vanligt

$$\delta_{ij} = \begin{cases} 0, & \text{då } i \neq j, \\ 1, & \text{då } i = j. \end{cases}$$

Alltså är skalärprodukten $\langle u, v\rangle$ lika med standardskalärprodukten $x \cdot y$ i R^n mellan motsvarande koordinatvektorer $x = (x_1, \ldots, x_n)$ och $y = (y_1, \ldots, y_n)$ (jämför den inledande diskussionen i avsnitt 2.1).

Låt oss återvända till basbytesformeln (1.16). Antag nu att båda baserna $e_1, \ldots, e_n$ och $e'_1, \ldots, e'_n$ är ON-baser. Kolonnerna i transformationsmatrisen T är koordinatvektorerna för $e'_1, \ldots, e'_n$ med koordinater i basen $e_1, \ldots, e_n$ (se (1.18)), och enligt det ovan sagda är de alltså parvis ortogonala (med standardskalärprodukten i R^n) och har längden 1. Denna egenskap hos en matris förtjänar ett särskilt namn.

Definition 2.7. En kvadratisk matris A kallas *ortogonal*, om $A^T A = I$.

Observation. A är ortogonal om och endast om dess kolonner är parvis ortogonala och har längden 1. Låt nämligen kolonnerna vara $a_1, a_2, \ldots, a_n$. Då är

$$A^T A = \begin{pmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{pmatrix} (a_1, a_2, \ldots, a_n) = \begin{pmatrix} a_1^T a_1 & a_1^T a_2 & \ldots & a_1^T a_n \\ a_2^T a_1 & a_2^T a_2 & \ldots & a_2^T a_n \\ -- & -- & -- & -- \\ a_n^T a_1 & a_n^T a_2 & \ldots & a_n^T a_n \end{pmatrix},$$

dvs. $[A^T A]_{ij} = a_i^T a_j = a_i \cdot a_j$. Eftersom A är ortogonal om och endast om $[A^T A]_{ij} = \delta_{ij}$, så följer påståendet. ■

Exempel 2.16. Från Exempel 2.7 följer att matrisen $A = \frac{1}{\sqrt{30}} \begin{pmatrix} \sqrt{5} & 0 & -5 \\ -2\sqrt{5} & \sqrt{6} & -2 \\ \sqrt{5} & 2\sqrt{6} & 1 \end{pmatrix}$ är en ortogonal matris. ■

Om A är ortogonal, så är $A^T A = I$, varför $A^{-1} = A^T$. Men då är också $AA^T = I$, vilket leder till resultatet att även A^T är ortogonal. Alltså är även A:s rader parvis ortogonala med längden 1.

Vid basbyte mellan två ON-baser blir alltså transformationsmatrisen T ortogonal. Då är $T^{-1} = T^T$, så att den andra ekvationen (1.16) förenklas till

$$x' = T^T x.$$

Exempel 2.17. Betrakta följande två ON-baser i R^3: $B = \{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\}$
$C = \{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\}$, där B är standardbasen. Vi vill bestämma
a) koordinaterna för vektorn $x = (6, 4, 2) \in R^3$ i basen C.
b) det element y i R^3 som har koordinaterna $(1, 2, 1)$ i C-basen.
Lösning:

a) Vi bestämmer först överföringsmatrisen $T_{B \leftarrow C} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}$. Nu får vi

$[x]_B = \begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix}$, $[x]_C = T_{C \leftarrow B}[x]_B = T^T_{B \leftarrow C}[x]_B = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}\begin{pmatrix} 6 \\ 4 \\ 2 \end{pmatrix} = \begin{pmatrix} \sqrt{2} \\ 5\sqrt{2} \\ 2 \end{pmatrix}$.

b) $y = [y]_B = T_{B \leftarrow C}[y]_C = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3/\sqrt{2} \\ 1/\sqrt{2} \\ 1 \end{pmatrix}$.

2.5 Några tillämpningar på skalärproduktrum

I avsnitt 2.3 såg vi några tillämpningar på minsta kvadratmetoden, som utgick ifrån att vi ville bestämma parametrar i en linjär modell så att modellen anpassades till gjorda mätningar. Vi lade då lika stor vikt vid de olika mätningar. Det finns situationer då man vill lägga större vikt vid vissa mätningar och vi ska studera hur det kan gå till.

Viktad minstakvadrat

Vi erinrar oss att en linjär modell med parametrar, som ska bestämmas, kan formuleras som ett överbestämt (om antalet observationer överstiger antalet parametrar) linjärt ekvationssystem $Ax = b$. Här är alltså $x \in R^n$ parametrarna och $b \in R^m$ är observerade värden på modellen och $A \in R^{m \times n}$ beskriver hur modellen ser ut. Minsta kvadrat-metoden för ekvationssystemet innebär att lösningen $\hat{x}$ väljs så att $\hat{b} = A\hat{x}$ blir så nära b som möjligt dvs så att $\|\hat{b} - b\|$ minimeras vilket är samma sak som att $\|\hat{b} - b\|^2 = < \hat{b} - b, \hat{b} - b >$ minimeras där $< \cdot, \cdot >$ är standardskalärprodukten i R^m. Hur man får lösningen beskrevs i ansnitt 2.3. Detta innebär, för de olika observationerna b_i, $i = 1, \ \ldots \ , m$, att $(\hat{b}_1 - b_1)^2 + (\hat{b}_2 - b_2)^2 + \ \ldots \ + (\hat{b}_m - b_m)^2$ minimeras.
Om man nu vill lägga olika vikt vid de olika observationerna så bör man i stället minimera $w_1^2(\hat{b}_1 - b_1)^2 + w_2^2(\hat{b}_2 - b_2)^2 + \ \ldots \ + w_m^2(\hat{b}_m - b_m)^2$ för olika vikter $w_i, i = 1, \ \ldots \ , m$. Detta

motsvarar att vi använder skalärprodukten $< x, y >= w_1^2 x_1 y_1 + w_2^2 x_2 y_2 + \ \dots \ + w_m^2 x_m y_m$ i stället för standardskalärprodukten. Med matrisen $W = diag(w_1, w_2, \ \dots \ , w_m)$ (matris med de angivna elementen i diagonalen) blir denna skalärprodukt $< x, y >= x^T W^2 y$ och viktad minsta kvadrat är alltså att minimera $\|W\hat{b} - Wb\|^2$ eller, vilket är samma sak, minimera $\|W\hat{b} - Wb\|$. Det överbestämda linjära ekvationssystem som vi ska lösa i minsta kvadrat-mening är $WAx = Wb$ och motsvarande normalekvationer $(WA)^T WAx = (WA)^T Wb$ ger oss lösningen $x = \hat{x}$. Normalekvationerna används mest vid handräkning, har man tillgång till matematisk programvara som t.ex. MATLAB så kan man få lösningen direkt med *backslash*.

Exempel 2.18. Vi ska anpassa en rät linje $x_1 + x_2 t$ till följande mätningar, punkter (t_i, y_i), $i = 1, \ \dots \ , 5$ i planet: (1,1.9), (2,2.7), (3,5.1), (4,5.9), (5,8.1), där de två sista mätningarna är osäkrare än de övriga och betraktas som hälften så viktiga.

Vi får modellmatrisen $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \\ 1 & 5 \end{pmatrix}$, högerledet $b = y = \begin{pmatrix} 1.9 \\ 2.7 \\ 5.1 \\ 5.9 \\ 8.1 \end{pmatrix}$ och väljer $W = diag(2, 2, 2, 1, 1)$.

Systemet att lösa är alltså $WAx = Wb$ som lätt löses med MATLAB: $x = (W * A) \backslash (W * b)$.

Vi får då minsta kvadrat-lösningen $x = \begin{pmatrix} 0.0859 \\ 1.5636 \end{pmatrix}$. ■

Vi har så långt studerat givna modeller men inte diskuterat val av själva modellen. Vi gör en lätt introduktion till uppgiften att välja rätt modell i ett enkelt fall.

Trendanalys

Vi har värden av en funktion $f(t)$, som beskriver något fenomen i tiden, i n diskreta tidpunkter $t_1, \ t_2, \ \dots \ , \ t_n$. Vi vill hitta en trend hos funktionen och anpassar t.ex. polynom av olika gradtal k till punkterna för att se om f är bra anpassad med linjärt, kvadratiskt, kubiskt etc polynom. Med den diskreta skalärprodukten $< f, g >= \sum_{i=1}^n f(t_i) g(t_i)$ (standardskalärprodukten i R^n för funktionsvärdena) blir uppgiften att minimera $\|\hat{f} - f\|, \hat{f} \in \mathcal{P}_k$, dvs $\hat{f}$ ska vara bästa approximation av f i underrummet $\mathcal{P}_k$. Om $\{p_1, \ p_2, \ \dots \ , \ p_{k+1}\}$ är ortogonal bas för $\mathcal{P}_k$ så blir lösningen $\hat{f} = \frac{<f,p_1>}{<p_1,p_1>} p_1 + \ \dots \ + \frac{<f,p_{k+1}>}{<p_{k+1},p_{k+1}>} p_{k+1}$.

Exempel 2.19. Vi vill anpassa funktionen $f(t)$ med de givna fem funktionsvärdena $f(-2) = 3, \ f(-1) = 5, \ f(0) = 5, \ f(1) = 4$ och $f(2) = 3$ med polynom $\hat{f}$ i $\mathcal{P}_2$ så att $\hat{f}$ är bästa approximation av f med avseende på skalärprodukten $< f, g >= \sum_{i=1}^5 f(t_i) g(t_i)$.

Polynomen $p_1 = 1, \ p_2 = t$ och $p_2 = t^2 - 2$ är ortogonala med avseende på denna skalärprodukt så lösningen blir

$\hat{f} = \frac{<f,p_1>}{<p_1,p_1>} p_1 + \frac{<f,p_2>}{<p_2,p_2>} p_2 + \frac{<f,p_3>}{<p_3,p_3>} p_3 = \frac{20}{5} + \frac{-1}{10} t + \frac{-7}{14}(t^2 - 2) = 4 - 0.1t - 0.5(t^2 - 2)$. ■

En viktig tillämpning på bästa approximation i ett underrum får vi den s.k. Fourieranalysen. Vi gör en första introduktion till detta rätt stora ämne.

Fourier-serie-approximation
Vi betraktar det linjära rummet $V = C[0, 2\pi]$ med skalärprodukt $< f, g >= \int_0^{2\pi} f(t)g(t)\ dt$ och underrummet $\mathcal{F}_n = Span\{1,\ \cos t,\ \cos 2t,\ \dots,\ \cos nt,\ \sin t,\ \sin 2t,\ \dots,\ \sin nt\}$, där de ingående funktionerna är ortogonala med avseende på den givna skalärprodukten, se övning 42.
Fourier-approximationen av ordning n av en given funktion $f \in V$ är bästa approximation $\hat{f}$ i underrummet $\mathcal{F}_n$. Denna kan vi skriva
$\hat{f} = proj_{\mathcal{F}_n} f = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + \quad + a_n \cos nt + b_1 \sin t + b_2 \sin 2t + \quad + b_n \sin nt$,
där koefficienterna $a_k,\ k = 0,\ \dots,\ n$ och $b_k,\ k = 1,\ \dots,\ n$ kallas *Fourier-koefficienter*.
Dessa beräknas enligt:
$\frac{a_0}{2} = \frac{<f,1>}{<1,1>} = \frac{1}{2\pi} \int_0^{2\pi} f(t)\ dt$
$a_1 = \frac{<f,\cos t>}{<\cos t,\cos t>} = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos t\ dt$
...
$a_n = \frac{<f,\cos nt>}{<\cos nt,\cos nt>} = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt\ dt$
$b_1 = \frac{<f,\sin t>}{<\sin t,\sin t>} = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin t\ dt$
...
$b_n = \frac{<f,\sin nt>}{<\sin nt,\sin nt>} = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt\ dt.$

Exempel 2.20. Vi ska bestämma Fourier-approximationen av andra ordningen till funktionen:
$f = \begin{cases} 1, & 0 \le t < \pi \\ -1 & \pi \le t \le 2\pi \end{cases}$
Vi får $\hat{f} = \frac{a_0}{2} + a_1 \cos t + a_2 \cos 2t + b_1 \sin t + b_2 \sin 2t$ med kofficienterna uträknade enligt:
$a_0 = \frac{1}{\pi} \int_0^{2\pi} f\ dt = 0$
$a_1 = \frac{1}{\pi} \int_0^{2\pi} f \cos t\ dt = \frac{1}{\pi}[\int_0^{\pi} \cos t\ dt - \int_{\pi}^{2\pi} \cos t\ dt] = 0$
$a_2 = \frac{1}{\pi} \int_0^{2\pi} f \cos 2t\ dt = \frac{1}{\pi}[\int_0^{\pi} \cos 2t\ dt - \int_{\pi}^{2\pi} \cos 2t\ dt] = 0$
$b_1 = \frac{1}{\pi} \int_0^{2\pi} f \sin t\ dt = \frac{1}{\pi}[\int_0^{\pi} \sin t\ dt - \int_{\pi}^{2\pi} \sin t\ dt] = \frac{4}{\pi}$
$b_2 = \frac{1}{\pi} \int_0^{2\pi} f \sin 2t\ dt = \frac{1}{\pi}[\int_0^{\pi} \sin 2t\ dt - \int_{\pi}^{2\pi} \sin 2t\ dt] = 0$
Svaret blir alltså $\hat{f} = \frac{4}{\pi} \sin t$. ■

2.6 Övningsuppgifter

1. Bestäm $\cos v$, där v är vinkeln mellan vektorerna $3t + 1$ och $5t^2 + 3$ i det linjära rummet $\mathcal{P}_2$ försett med skalärprodukten $< f, g >= \int_{-1}^{1} f(t)g(t)dt$.

2. Visa att $\int_0^{\frac{\pi}{2}} \sqrt{\sin x \cos x}\, dx \le 1$.

3. För vilka värden på a är vektorerna $(a, 1, 1)$ och $(a, 1, a)$ ortogonala med avseende på skalärprodukten $x_1y_1 + 2x_2y_2 + 3x_3y_3$ i R^3?

4. Visa att $< x, y >= 2x_1y_1 - 2x_1y_2 - 2x_2y_1 + 5x_2y_2$ är en skalärprodukt i R^2. Bestäm en ortonormerad bas för R^2 med avseende på denna skalärprodukt.

5. Är någon av följande två kandidater en skalärprodukt på $C^1[a, b]$:

$$\begin{aligned} < f, g > &= \int_a^b f'(t)g'(t)dt \\ < f, g > &= \int_a^b f'(t)g'(t)dt + f(a)g(a)? \end{aligned}$$

(f' betecknar derivatan av f.)

6. Låt $A \in R^{n\times n}$ vara inverterbar. Visa att för $u \in R^n$ och $v \in R^n$ så är $< u, v >= (Au) \cdot (Av) = (Au)^T(Av)$ en skalärprodukt i R^n.

7. Visa att för en skalärprodukt och motsvarande norm i ett linjärt rum gäller:
$< u, v >= \frac{1}{4}\|u + v\|^2 - \frac{1}{4}\|u - v\|^2$.

8. Betrakta rummet $\mathcal{P}_2$ med skalärprodukten i Exempel 2.1(g) med avseende på punkterna $t_1 = -1,\ t_2 = 0,\ t_3 = 1$. Beräkna $< p, q >,\ \|p\|$ och $\|u\|$ då $p = 4 + t$ och $q = 5 - 4t^2$. Bestäm även ortogonala projektionen av q på det rum som spänns av p.

9. Låt $u \neq 0$ och $v \neq 0$ vara två vektorer i ett linjärt rum med skalärprodukt sådana att $\|u\| = \|v\| = \|u - v\|$. Bestäm vinkeln mellan u och v.

10. Låt $\|u\| > \|v\|$. Visa att u och $u - v$ inte är ortogonala.

11. Låt a och b vara positiva tal och låt $u = \begin{pmatrix} \sqrt{a} \\ \sqrt{b} \end{pmatrix}$ och $v = \begin{pmatrix} \sqrt{b} \\ \sqrt{a} \end{pmatrix}$. Använd u, v och Cauchys olikhet för att få fram en relation mellan det geometriska medelvärdet $\sqrt{ab}$ och det aritmetiska medelvärdet $(a + b)/2$.

12. Betrakta rummet $V = C[0, 1]$ med skalärprodukt enligt Exempel 2.1(f).
Beräkna $< f, g >,\ \|f\|$ och $\|g\|$ då $f = 1 - 3t^2$ och $g = t - t^3$.

13. Låt $V = C[-1, 1]$ med skalärprodukt enligt Exempel 2.1(f). Bestäm en ortogonal bas för det underrum som spänns av vektorerna 1, t och t^2. (Polynomen i denna bas kallas Legendre-polynom.)

14. Låt K_n vara den n-dimensionella kuben

$$K_n = \{x \in R^n; 0 \leq x_j \leq 1, \quad j = 1, \ldots, n\}.$$

a) Beräkna längden av diagonalen i K_n från $(0, \ldots, 0)$ till $(1, \ldots, 1)$.
b) Visa att diagonalen i a) bildar samma vinkel θ_n med alla kantlinjer i K_n. Vilket värde närmar sig θ_n då $n \to \infty$?

15. Ange samtliga vektorer i R^4 som är ortogonala mot de båda vektorerna $u_1 = (1, 2, 1, 3)$ och $u_2 = (2, 5, 1, 4)$.

16. (M) Ange (genom handräkning) en ortonormerad bas i det underrum i R^4 som genereras av vektorerna
$w_1 = (1,1,1,1)$, $w_2 = (1,2,2,1)$, $w_3 = (2,3,1,6)$.
Jämför med vad MATLAB ger.

17. (M) Bestäm (genom handräkning) en ortonormerad bas i det underrum i R^4 som ges av

$$x_1 + 2x_2 - x_3 + 4x_4 = 0.$$

Jämför med vad MATLAB ger.

18. (M) Låt U vara det tvådimensionella underrum i R^5 som genereras av vektorerna $(1,2,0,0,0)$ och $(1,0,3,0,0)$. Bestäm (genom handräkning) en ortonormerad bas för det ortogonala komplementet $U^\perp$ till U. Jämför med vad MATLAB ger.

19. Ange den ortogonala projektionerna av $u = (0,4,4,0)$ på det underrum i R^4 som genereras av vektorerna $(1,1,1,1)$ och $(1,-1,1,-1)$.
Bestäm avståndet från u till underrummet.

20. Beräkna avståndet från (0,2,0,2,1) till det underrum i R^5 som genereras av (1,1,1,1,1) och (1,2,1,0,1).

21. (M) Bestäm (för hand) en ON-bas för värderummet $V(A)$ till matrisen $\begin{pmatrix} 1 & 2 & 5 \\ -1 & 1 & -4 \\ -1 & 4 & -3 \\ 1 & -4 & 7 \\ 1 & 2 & 1 \end{pmatrix}$.
Jämför med vad MATLAB ger.

22. Låt U vara ett hyperplan i R^n med ekvationen

$$a_1x_1 + \ldots + a_nx_n + b = 0.$$

Visa att avståndet d från en godtycklig punkt $x = (x_1, \ldots, x_n)$ till U ges av formeln

$$d = \frac{|a_1x_1 + \ldots + a_nx_n + b|}{\sqrt{a_1^2 + \ldots + a_n^2}}.$$

23. (M) Beräkna det minsta avståndet från (0,2,4,1,6) till skärningen mellan de två hyperplanen $x_1 + x_2 - x_3 + x_5 = 0$ och $x_1 - x_2 + 3x_3 + x_4 - x_5 = 0$ i R^5.

24. Låt $e_1, \ldots, e_n$ vara en ortonormerad mängd i ett linjärt rum V med skalärprodukt.
a) Verifiera att om $u \in V$ och $u' = \sum_{j=1}^n < u, e_j > e_j$, så är vektorerna $u - u'$ och u' ortogonala.

b) Visa Bessels olikhet dvs. att

$$\sum_{j=1}^{n} < u, e_j >^2 \leq \|u\|^2 \,.$$

25. U_1 och U_2 är underrum i ett linjärt rum V med skalärprodukt. Visa att

$$(U_1 + U_2)^{\perp} = U_1^{\perp} \cap U_2^{\perp}.$$

$(U_1 + U_2 = \{u + v : u \in U_1,\ v \in U_2\})$

26. a) Visa att $< f, g >= \int_0^{\infty} f(t)g(t)e^{-t}dt$ definierar en skalärprodukt på $\mathcal{P}_n$.
b) Konstruera en ortonormerad bas för $\mathcal{P}_2$ (med avseende på skalärprodukten i a)) genom att tillämpa Gram-Schmidts metod på basen $1, t, t^2$.

27. (M) Tillämpa minsta kvadratmetoden på följande överbestämda ekvationssystem

a) $\begin{cases} x_1 + 2x_2 = 3 \\ -2x_1 + x_2 = -4 \\ x_1 - 3x_2 = -2 \\ -x_1 + x_2 = -1 \\ 2x_1 + x_2 = 5 \end{cases}$ b) $\begin{cases} x_1 + x_2 + x_3 = 4 \\ -x_1 + x_2 + x_3 = 0 \\ -x_2 + x_3 = 1 \\ x_1 + x_3 = 2 \end{cases}$

28. (M) Ange den andragradskurva $y = f(x) = ax^2 + bx + c$ som, i minsta kvadratmening, bäst ansluter till följande värden på x och y

a)

x	-2	-1	0	1	2
y	2	1	1	2	3

b)

x	-1	0	1	2
y	2	2	1	0

(Det gäller alltså att välja f så att $\sum(f(x_j) - y_j)^2$ blir så liten som möjligt.)

29. Anpassa modellen $y = A\cos x + B\sin x$ i minsta kvadratmening till mätdata:
$(1, 7.9),\ (2, 5.4),\ (3, -0.9)$.

30. För att undersöka ett flygplans egenskaper vid start har man uppmätt dess höjd över marken varje sekund från $t = 0$ till $t = 12$. Höjderna (i fot) mättes till:
0, 8.8, 29.9, 62.0, 104.7, 159.1, 222.0, 294.5, 380.4, 471.1, 571.7, 686.8, 809.2.
Bestäm en kubisk anpassning $y = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$ till dessa data i minsta kvadratmening. Använd anpassningen för att approximera planets hastighet vid $t = 4.5$.

31. Låt e_1, e_2, e_3, e_4 vara en ortonormerad bas för det linjära rummet V med skalärprodukt och sätt

$$\begin{aligned} 3e_1' &= e_1 + 2e_2 + 2e_4 \\ 3e_2' &= 2e_1 - e_2 + 2e_3 \\ 3e_3' &= 2e_2 + e_3 - 2e_4 \\ 3e_4' &= -2e_1 + 2e_3 + e_4. \end{aligned}$$

Verifiera att också e_1', e_2', e_3', e_4' är en ortonormerad bas och uttryck koordinaterna i denna bas med hjälp av koordinaterna i den ursprungliga basen. (Observera att för en ortogonal matris är $T^T = T^{-1}$.)

32. Bestäm talen a,b,c så att matrisen

$$\frac{1}{7}\begin{bmatrix} 2 & 6 & b \\ 3 & 2 & c \\ 6 & a & 2 \end{bmatrix}$$

blir ortogonal.

33. Visa att om A och B är två ortogonala $n \times n$ matriser så är också AB ortogonal.

34. Visa att determinanten av en ortogonal matris är 1 eller -1.

35. Låt A vara en ortogonal matris av typ $n \times n$ där n är udda. Låt $\det A = 1$. Visa att $\det(A - I) = 0$

36. Låt A vara en skevsymmetrisk $n \times n$ matris dvs $A = -A^T$. Visa att
a) Ax är ortogonal mot x, dvs $(Ax)^T x = 0$, för varje $x \in R^n$.
b) $I - A$ är inverterbar
c) $(I + A)(I - A) = (I - A)(I + A)$
d) $T = (I - A)^{-1}(I + A)$ är en ortogonal matris.

37. Verifiera att om T är ortogonal och $T + I$ är inverterbar, så är $A = (T - I)(T + I)^{-1}$ skevsymmetrisk, för definition se föregående övning.

38. Bestäm andragradspolynomet $p(t) = at^2 + bt + c$ så att integralen

$$\int_{-1}^{1} (t^4 - p(t))^2 dt$$

blir så liten som möjligt.

39. Betrakta rummet $C[-2, 2]$ med skalärprodukt $< f, g >= \int_{-2}^{2} f(t)g(t)\ dt$. Visa att $q_1,\ q_2,\ q_3$ är en ortogonal mängd, där $q_1 = 1,\ q_2 = t$ och $q_3 = 3t^2 - 4$.

40. (M) Bestäm bästa minsta kvadratanpassning med modellen $y = \beta_1 + \beta_2 t + \beta_3 t^2$ till datapunkterna $(-2, 0),\ (-1, 0),\ (0, 2),\ (1, 4),\ (2,\ 4)$, där den första och sista punkten är osäkrare än de övriga och ges vikten hälften så mycket.

41. Betrakta skalärprodukten i Exempel 2.1(g) och punkterna
$t_1 = -2,\ t_2 = -1,\ t_3 = 0,\ t_4 = 1,\ t_5 = 2.$
Visa att polynomen $\{1,\ t,\ t^2-2,\ \frac{5}{6}t^3 - \frac{17}{6}t\}$ är ortogonala. Anpassa med hjälp av dessa polynom en kubisk trend till datapunkterna
$(-2,3),\ (-1,5),\ (0,5),\ (1,4),\ (2,3).$

42. Betrakta rummet $C[0,2\pi]$ med skalärprodukt $< f, g >= \int_0^{2\pi} f(t)g(t)\ dt.$
a) Visa att $\sin mt$ och $\sin nt$ är ortogonala om $n \neq m$.
b) Visa att $\sin mt$ och $\cos nt$ är ortogonala för alla (positiva heltal) m och n.
c) Visa att $\|\sin nt\| = \pi$ och $\|\cos nt\| = \pi$ för alla $n > 0$.

43. Bestäm tredje ordningens Fourier-approximation till $f(t) = 2\pi - t$.

44. Bestäm första, andra och tredje ordningens Fourier-approximation till
$f(t) = 3 - 2\sin t + 5\sin 2t - 6\cos 2t$.

45. Bestäm tredje ordningens Fourier-approximation till $f(t) = \sin^2 t$.

2.7 Svar

1. $\frac{7}{6\sqrt{6}}$.
3. $a = -1$ eller $a = -2$.
5. Ja, den andra.
8. $< p, q >= 28,\ \|p\| = \sqrt{50},\ \|q\| = \sqrt{27},\ proj_p\ q = \frac{28}{50}(4+t).$
9. $\frac{\pi}{3}$
11. $\sqrt{ab} \leq \frac{a+b}{2}$.
12. $< f, g >= 0,\ \|f\| = \frac{2}{\sqrt{5}},\ \|g\| = \frac{2\sqrt{2}}{\sqrt{105}}$.
13. $\{1,\ t,\ t^2 - 1/3\}$.
14. a) $\sqrt{n}$, b) $\frac{\pi}{2}$
15. $t(-3,1,1,0) + s(-7,2,0,1), \quad t, s \in R$
16. t.ex. $\frac{1}{2}(1,1,1,1),\ \frac{1}{2}(-1,1,1,-1),\ \frac{1}{\sqrt{10}}(-2,1,-1,2)$
17. t.ex. $\frac{1}{\sqrt{2}}(1,0,1,0),\ \frac{1}{\sqrt{3}}(-1,1,1,0),\ \frac{1}{\sqrt{33}}(2,4,-2,-3)$
18. t.ex. $\frac{1}{7}(6,-3,-2,0,0), (0,0,0,1,0), (0,0,0,0,1)$
19. $(2,2,2,2)$, avstånd 4.
20. 2.
21. $\{\frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \\ -1 \end{pmatrix}\}$.
23. $\sqrt{13}$.
26. b) $\{1,\ t-1,\ \frac{1}{2}t^2 - 2t + 1\}$

27. a) $x_1 = 2$; $x_2 = 1$, b) $x_1 = 1.6$; $x_2 = 0.6$; $x_3 = 1.2$
28. a) $\frac{1}{70}(25x^2 + 21x + 76)$ b) $\frac{1}{20}(-5x^2 - 9x + 37)$
29. $A = 2.3421,\ B = 7.4475.$
30. $\beta = (-0.8558,\ 4.7025,\ 5.5554,\ -0.0274)$, $v(4.5) = 53.0387.$
31. $3x_1' = x_1 + 2x_2 + 2x_4$, $3x_2' = 2x_1 - x_2 + 2x_3$, $3x_3' = 2x_2 + x_3 - 2x_4$, $3x_4' = -2x_1 + 2x_3 + x_4.$
32. $a = -3,\ b = 3,\ c = -6.$
38. $\frac{6}{7}t^2 - \frac{3}{35}.$
40. $y = 2 + \frac{3}{2}t.$
41. $\hat{y} = 4 - 0.1t - 0.5(t^2 - 2) + 0.5(\frac{5}{6}t^3 - \frac{17}{6}t).$
43. $\hat{f} = \pi + 2\sin t + \sin 2t + \frac{2}{3}\sin 3t.$
44. Första ordning: $\hat{f} = 3 - 2\sin t$. Andra och tredje ordning: $\hat{f} = f$.
45. $f = \sin^2 t = \frac{1}{2} - \frac{1}{2}\cos 2t$, $\hat{f} = f$.